

BROADBAND ANALOG OPTO-ELECTRONIC  
BLIND SOURCE SEPARATION

by

PAUL CARSON SMITH

B.E., University of Colorado, 2000

M.S., University of Colorado, 2000

A thesis submitted to the  
Faculty of the Graduate School of the  
University of Colorado in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
Department of Electrical and Computer Engineering  
2005

This thesis for the Doctor of Philosophy degree by  
Paul Carson Smith  
has been approved for the  
Department of  
Electrical and Computer Engineering  
by

Zoya Popović

Dana Anderson

May 26. 2005.

The final copy of this thesis has been examined by the signatories, and we  
find that both the content and the form meet acceptable presentation  
standards of scholarly work in the above mentioned discipline.

Smith, Paul Carson (Ph.D., Electrical Engineering)

Broadband analog opto-electronic blind source separation

Thesis directed by Professor Zoya Popović

This thesis addresses the problem of high bandwidth blind source separation and demonstrates an analog opto-electronic implementation of a processor capable of solving this problem for two linearly mixed signals. Using only the assumption of signal independence, real-time separation exceeding 20 dB is demonstrated. The opto-electronic circuit is capable of processing signals with 100 kHz bandwidth. The theory presented here show that the dynamics of the core portion of the system depend on statistical characteristics of the signals. If at least one of the signals is sub-Gaussian, the dynamical properties of the system are such that one of the original signals obtains all the system gain and is extracted from the mixture.

A dynamic feedback loop containing an electro-optic phase modulator operating in the nonlinear regime provides processing functions analogous to higher-order statistical correlation of the signals. The correlation portion of the processor is based on photorefractive two-beam coupling in BaTiO<sub>3</sub> at an optical wavelength of 532 nm. The behavior of the feedback loop is described by an open-loop gain analysis and a steady-state stability analysis. The system demonstrates “winner-takes-all” competition between the two input signals.

The bandwidth of the processor presented here is limited by electronic amplifiers and can exceed 100 MHz with the higher bandwidth off-the-shelf components. The high bandwidth makes this processor applicable to communication blind source separation scenarios that are difficult to solve using traditional digital signal processing. In addition, this processor can be used directly at the microwave carrier frequency of tens of gigahertz.

# Dedication

*To my parents, brother, and friends, who always believe in me.*

# Acknowledgments

I would like to acknowledge, offer thanks, and otherwise recognize the positive influence of a few dozen or more friends and colleagues.

*At the University of Colorado:*

Thanks very much to Zoya Popović and Dana Anderson for teaching me, the financial support and the opportunity

*Past friends in the Antenna Lab:*

Joe Hagerty, Paja, Jim Vian, Michael Forman, and Manoja Weiss

*Past Friends in the Optics Lab:*

Shane, Hongke, and Ying-Ju

*My contemporaries and friends in suffering:*

Ben, Matt, Steve, Patrick, Alan, and Jason

*Most importantly, my past lab partners:*

Edeline Fotheringham and Stefania Romisch

*And my current lab partners:*

Marty Baylor and Karl Erickson

*And to my friends and counselors:*

Leslie Czaia and Rachael Tearle

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Overview . . . . .	1
1.2	Problem Definition . . . . .	6
1.3	Opto-electronic ICA . . . . .	8
1.4	Thesis Outline . . . . .	13
<b>2</b>	<b>Opto-electronic ICA Theory</b>	<b>14</b>
2.1	Introduction . . . . .	14
2.2	The opto-electronic feedback loop . . . . .	16
2.3	Sinusoidal signal analysis . . . . .	18
2.3.1	Behavior of two input signals . . . . .	18
2.3.2	Sinusoidal dynamic evolution . . . . .	21
2.3.3	System steady state behavior . . . . .	22
2.3.4	Steady state amplitude . . . . .	25
2.4	Arbitrary probability density function analysis . . . . .	28
2.4.1	Behavior of two input signals . . . . .	28

2.4.2	Gain of various probability density functions . . . . .	31
2.4.3	Dynamic evolution of arbitrary PDF signals . . . . .	34
2.4.4	<i>Lotka-Volterra</i> simplifications and Kurtosis . . . . .	35
2.4.5	<i>Lotka-Volterra</i> steady state solutions . . . . .	36
2.4.6	Complete system dynamics . . . . .	38
2.4.7	High gain system dynamics . . . . .	42
2.5	Summary . . . . .	45
<b>3</b>	<b>100 kHz bandwidth opto- electronic ICA implementation</b>	<b>49</b>
3.1	Overview . . . . .	49
3.2	System description . . . . .	50
3.3	Input channels . . . . .	51
3.3.1	Signal generation . . . . .	53
3.3.2	Input electronics . . . . .	55
3.3.3	Input optics . . . . .	59
3.4	ICA feedback loop . . . . .	61
3.4.1	Electro-optic modulator . . . . .	63
3.4.2	Photorefractive medium and loop beam power . . . . .	69
3.4.3	Photodetector . . . . .	72
3.4.4	Electronic Amplifiers . . . . .	73
3.5	Homodyne detection circuitry . . . . .	73
3.5.1	Homodyne theory . . . . .	73
3.5.2	Homodyne circuit . . . . .	75

3.6	Measurement circuitry . . . . .	79
3.7	Summary . . . . .	81
<b>4</b>	<b>100 kHz ICA experimental results</b>	<b>82</b>
4.1	Introduction . . . . .	82
4.2	Open-loop gain measurements . . . . .	83
4.2.1	Sinusoidal signal measurement . . . . .	86
4.2.2	Gaussian signal measurement . . . . .	90
4.3	Sinusoidal signal separation . . . . .	91
4.3.1	Experimental setup . . . . .	92
4.3.2	Sinusoidal signal separation measurement . . . . .	92
4.3.3	Loop loss measurement . . . . .	95
4.3.4	Phase error results . . . . .	96
4.3.5	Signal winner measurements . . . . .	98
4.4	Arbitrary PDF signal separation . . . . .	99
4.4.1	At least one sub-Gaussian signal separation measurements . . . . .	100
4.4.2	Gaussian signal separation measurements . . . . .	102
4.4.3	Laplace signal separation measurements . . . . .	103
4.5	Time-varying measurements . . . . .	103
4.5.1	Time-varying system setup . . . . .	103
4.5.2	Time-varying signal separation . . . . .	104
4.5.3	Time-varying mixtures . . . . .	105

4.6	FM signal and Gaussian noise separation . . . . .	109
4.6.1	Frequency modulation probability density function . . .	109
4.6.2	Digital modulation and demodulation of an FM signal	113
4.6.3	Analog modulation and demodulation of an FM signal	115
4.6.4	System noise measurements . . . . .	116
4.6.5	FM signal separation results . . . . .	118
4.7	Ones density binary signal separation . . . . .	118
4.7.1	Binary data probability density function . . . . .	120
4.7.2	Ones density demodulation . . . . .	122
4.7.3	System noise measurements for binary signal separation	122
4.7.4	Binary signal separation results . . . . .	123
4.8	Summary of experimental results . . . . .	124
<b>5</b>	<b>Related and Future Work</b>	<b>125</b>
5.1	Thesis contributions . . . . .	125
5.2	Future work . . . . .	126
5.3	20-MHz bandwidth electro-optic modulator . . . . .	127
5.3.1	EOM crystal design . . . . .	127
5.3.2	EOM matching circuit design . . . . .	130
5.3.3	EOM mount design . . . . .	138
5.4	Summary . . . . .	141
	<b>Bibliography</b>	<b>144</b>



# Tables

4.1	Comparison of kurtoses of music signals . . . . .	110
4.2	Measured Signal-to-Noise Ratio of System Components . . .	117

# Figures

1.1	The generalized communications environment, with $N$ transmitted signals . . . . .	2
1.2	A high-level description of the ICA procedure. . . . .	8
1.3	The optical Principal Component Analysis (PCA) system. . .	10
1.4	A two-channel optoelectronic blind-source-separation system.	12
1.5	Generalized block diagram of the optoelectronic ICA system.	12
2.1	Representation of opto-electronic ICA . . . . .	15
2.2	Schematic of the dynamic ICA feedback loop . . . . .	16
2.3	Open-loop gain as a function of input signal amplitude . . . .	26
2.4	Open-loop gain as a function of input signal amplitude with feedback loop loss . . . . .	28
2.5	Four probability density functions that are analyzed in the feedback loop . . . . .	32
2.6	Open-loop gain of harmonic, uniform, Gaussian, and Laplace density functions . . . . .	34
2.7	Illustration of kurtoses . . . . .	37

2.8	Harmonic vs. harmonic system dynamics . . . . .	40
2.9	Uniform vs. uniform system dynamics . . . . .	41
2.10	Gaussian vs. Gaussian system dynamics . . . . .	42
2.11	Laplace vs. Laplace system dynamics . . . . .	43
2.12	Harmonic vs. Laplace system dynamics . . . . .	44
2.13	Harmonic vs. Gaussian system dynamics . . . . .	45
2.14	Gaussian vs. Laplace system dynamics . . . . .	46
2.15	Harmonic signal open-loop gain curve . . . . .	47
2.16	High gain harmonic vs. harmonic system dynamics . . . . .	48
3.1	Generalized block diagram of the opto-electronic ICA system	51
3.2	Single input channel schematic . . . . .	52
3.3	Signal space diagrams of whitened signals . . . . .	54
3.4	Low pass filter magnitude and phase characteristics . . . . .	58
3.5	Low pass filter used for input channel phase correction . . . . .	59
3.6	Carrier suppression schematic . . . . .	61
3.7	Input channel optical setup . . . . .	62
3.8	Opto-electronic feedback loop . . . . .	62
3.9	Feedback loop electro-optic modulator structure . . . . .	64
3.10	Feedback loop electro-optic modulator and wave plate setup . . . . .	65
3.11	Feedback loop electro-optic modulator half-wave voltage measurement . . . . .	66
3.12	Electro-optic modulator output polarizations . . . . .	67

3.13	Loop modulator output intensity . . . . .	68
3.14	Photorefractive crystal gain saturation . . . . .	71
3.15	Measured photorefractive gain . . . . .	72
3.16	Homodyne detection circuitry . . . . .	76
3.17	Homodyne detection circuitry beam phase shifts . . . . .	78
3.18	Output performance and evaluation circuitry . . . . .	80
4.1	Open-loop system setup . . . . .	84
4.2	Open-loop gain measurement using digital signal analyzer . . . . .	85
4.3	Open-loop gain measurement using analog correlator . . . . .	86
4.4	Sinusoidal open-loop gain . . . . .	87
4.5	Two-channel sinusoidal open-loop gain results . . . . .	88
4.6	Open-loop gain with saturated photorefractive gain . . . . .	89
4.7	Gaussian signal open-loop gain . . . . .	91
4.8	Closed-loop sinusoidal signal measurement setup . . . . .	93
4.9	Sinusoidal signal separation . . . . .	94
4.10	Sinusoidal signal separation with phase error . . . . .	97
4.11	Random signal separation with arbitrary PDFs . . . . .	101
4.12	Random signal separation with Gaussian signals . . . . .	102
4.13	Correlation voltage as a function of time . . . . .	104
4.14	1-degree per second rotating mixture . . . . .	106
4.15	20-degrees per second rotating mixture . . . . .	107
4.16	30 degree per second rotating mixture . . . . .	108

4.17	Generic super-Gaussian music signal . . . . .	111
4.18	Spectrum of FM signal and Gaussian noise mixture . . . . .	114
4.19	Ones density modulation . . . . .	119
4.20	Ones density modulation setup. . . . .	121
4.21	Circuit of fifth-order elliptic filter for ones density demodulation	123
5.1	Electro-optic crystal dimensions . . . . .	129
5.2	Fourth order low pass Chebychev filter . . . . .	132
5.3	Second order bandpass Chebychev filter . . . . .	133
5.4	Two-stage impedance transform circuit . . . . .	133
5.5	Lossless EOM impedance matching circuit . . . . .	135
5.6	Lossy EOM impedance matching circuit . . . . .	136
5.7	Lossy EOM impedance matching circuit frequency response .	136
5.8	Lossy EOM impedance matching circuit voltage and current responses . . . . .	137
5.9	Electro-optic modulator mount design . . . . .	139
5.10	EOM matching circuit layout . . . . .	140
A.1	ICA procedure . . . . .	157

# Chapter 1

## Introduction

### 1.1 Overview

The signal processing power of independent component analysis (ICA) is well established through a variety of demonstrations in audio processing, medical signal processing, image processing, and many other fields [1]. The ICA algorithm has been implemented in software or, in some cases, by a powerful digital signal processor [2], [3]. The signals separated are generally either stored data or have bandwidths of less than 50 kHz. This thesis describes an opto-electronic system designed to perform ICA on radio frequency (RF) bandwidth signals at the carrier frequency.

Our goal is to apply the system demonstrated in this thesis to multiple signals modulated onto multiple radio-frequency carriers that are incident on an antenna array from unknown directions. Figure 1.1 shows a sketch of

the envisioned scenario, where  $N$  transmitted signals,  $s_i(t)$ ,  $i = 1 \dots M$ , are incident on  $M$  antennas, each of which receives a different linear combination (mixture) of the  $N$  signals,  $\tilde{s}_j(t)$ ,  $i = 1 \dots N$ , due to the fact that the antennas are spatially separated. The classical formulation of the blind signal separation problem assumes that the number of receiving antennas is equal to or greater than the number of sources,  $M \geq N$  [4]. In this case, there are  $N$  unknown signals, and  $M \times N$  unknown mixture coefficients, resulting in a total of  $N(M + 1)$  time-varying unknowns that need to be determined from  $M$  measured mixtures. Therefore, some assumptions need to be made in order to obtain the unknown original signals.

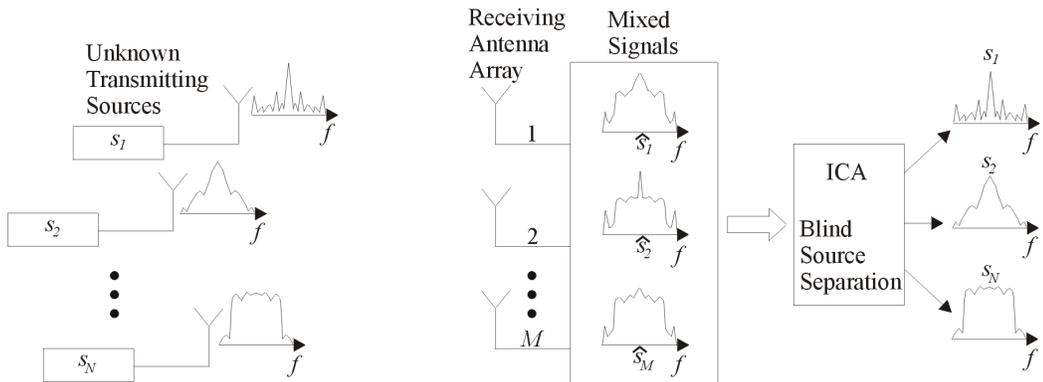


Figure 1.1: The generalized communications environment, with  $N$  transmitted signals. Each of the transmitted signals is incident upon a receiving array of  $M$  antennas. The signals are linearly mixed during propagation and detection. The observed signal mixtures are the only information available to the blind-source-separation system. The separation system uses statistics to recover the original signals from the mixtures.

The primary assumption that is made in every ICA algorithm or system

is that the original signals are statistically independent. Although independence, in the strictest sense, is defined as the joint probability density function of two random variables being equal to the product of the probability density functions of each variable, in this thesis, independence will be defined as signals that are uncorrelated to every order:

$$\langle s_1^a s_2^b \rangle = \langle s_1^a \rangle \langle s_2^b \rangle \quad (1.1)$$

where  $a$  and  $b$  are integers. The assumption of independence is usually satisfied in a communication environment where each transmitting source has a different oscillator or clock. The independence assumption also leads to a second stipulation: only one of the original signals may have a Gaussian distribution. This criterion comes from the fact that all Gaussian signals are correlated at higher orders ( $a + b \geq 4$ ). There is enough information to determine both the original signals (up to a scaling constant) if the original signals are uncorrelated up to fourth order. This means that  $a + b \leq 4$  for all  $a$  and  $b$ .

ICA is used in a diverse range of applications [5]: separation of multiple audio signals in hearing aids [6], [7], [8], [9]; isolating the fetal heartbeat from the mother's heartbeat [10]; functional magnetic resonance imaging (fMRI) analysis [11]; reduction of noise in image enhancement techniques [12]; analyzing astronomical data [13], [14]; facial recognition [15], [16], [17]; machine malfunction tracking in an environment with high vibration levels [18]; and observing stock market trends [19].

In all the listed applications, either the signal bandwidth is narrow, limited to tens of kHz, or real-time processing is not necessary (e.g. in astronomy). The attempts to use ICA in communications scenarios all require more information about the transmitted signals, such as the CDMA code [20], [21], [22], [23]. There are relatively few systems that attempt real-time ICA processing [24], [25], [26]. For wireless communications and radar scenarios that could benefit from real-time ICA, state-of-the-art digital signal processing does not have sufficient computational power to process the necessary signal bandwidths, which are typically above a few MHz. This problem arises from the fact that higher-order statistical component estimation, described in appendix 2, requires large processing power. For example, in [27], 60 MHz DSP electronics was required for ICA of 2 signals with 22 kHz bandwidth. Therefore, in order to separate signals with bandwidths of hundreds of MHz, scaling indicates that 4 orders of magnitude improvement in DSP electronic bandwidth is required. In the communication scenario outlined in figure 1, the carrier signals may have any common microwave communication frequency, from hundreds of megahertz to tens of gigahertz. The signals may be downconverted into the kilohertz range or remain in the gigahertz regime depending on the application. The type of modulation used is transparent to the operation of the system, provided that it creates signal with a sub-Gaussian distribution function, as will be discussed in chapter 2. This difference between signal bandwidths is a primary motivation for an opto-electronic ICA system.

Although the ICA system presented in this thesis relies on the same assumption as the ICA algorithms, it is fundamentally different in operation. Instead of following an algorithmic procedure, it is based on the nonlinear, dynamic operation of a partly optical, partly electronic feedback loop. The primary nonlinearity in the feedback loop is an over-driven electro-optic phase modulator (EOM), used to generate harmonics of the modulated signal.

The second essential component of the feedback loop is a photorefractive crystal [28]. This crystal acts like a coherent amplifier for the phase-modulated laser beam. The amplification process can be used to provide a correlation between the harmonics generated by the EOM and the signals themselves. While it is not intuitively obvious, this correlation makes it possible for signals with different statistical distributions to experience different amplification within the feedback loop. The dynamics of the system are hidden within the combination of the electro-optic modulator and the photorefractive crystal. We have theoretically and experimentally demonstrated that the feedback loop actually prefers to allow the oscillation of only one independent signal present within a mixture of input signals.

The primary reason for opto-electronic processing is signal bandwidth, which is limited by the bandwidth of the operating components and the physical size of the feedback loop. With high-speed components and a small loop, the bandwidth of the system can be pushed to hundreds of megahertz, approaching a gigahertz. The second motivating factor is the inherent parallelism of optics, which allows linear scaling with the number of input signals.

To separate  $N$  signals,  $N-1$  feedback loops are required. In contrast, computerized ICA algorithms become highly inefficient with only a modest number of inputs such as eight [1].

## 1.2 Problem Definition

ICA is applicable to a large variety of problems because it makes very few assumptions about the input signals or mixture. A mathematical description of the problem is:

$$\tilde{s}(t) = \mathbf{A}(t) s(t) \quad (1.2)$$

where  $s(t) = [s_1(t) \dots s_N(t)]$  is a vector of the unknown original signals,  $\tilde{s}(t) = [\tilde{s}_1(t) \dots \tilde{s}_M(t)]$  is a vector of the signal mixtures, and  $\mathbf{A}(t)$  is the unknown and in general time-varying linear  $M \times N$  mixing matrix. For the case of a two-signal system, Equation 1.2 becomes:

$$\begin{bmatrix} \tilde{s}_1(t) \\ \tilde{s}_2(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix}. \quad (1.3)$$

Only  $\tilde{s}_1(t)$  and  $\tilde{s}_2(t)$  are measured as the signals received by two antennas in the communications scenario highlighted in Figure 1.1. The two original signals,  $s_1(t)$  and  $s_2(t)$ , and all four mixing coefficients,  $a_{mn}(t)$ , are unknown. The problem of obtaining the original signals,  $s_1$  and  $s_2$ , reduces to inverting an unknown matrix. As previously stated, ICA algorithms approach this matrix inversion by using signal statistics.

All ICA algorithms must somehow measure the independence, at least up to fourth order, of the signal mixtures [29]. Like most problems in science, this one has many possible strategies, which have been expressed in different ICA algorithms, for arriving at the solution [30]. Among these algorithms, independence measures such as signal kurtosis [31], negentropy [32], minimization of mutual information [33], and maximum likelihood [34] have been tried, as described in appendix A (see [35] for comparison).

To increase the speed of the ICA algorithms, additional assumptions about the input signals can be made. For example, the input signals can be assumed to have zero mean. This assumption does not affect the generality of ICA [1]. In addition, the signals are usually assumed to have symmetrical probability density functions. This means that there are no odd order correlations in Equation 1.1. However, the ICA framework can be modified to include asymmetric signals.

These added assumptions imply that there are only two orders of independence to investigate: second order,  $a+b = 2$ , and fourth order,  $a+b = 4$ . The procedure of ICA is usually to split the problem into two sections, corresponding to a second-order decorrelation and a fourth-order decorrelation [33], [36], as shown in Figure 1.2. The second-order decorrelation is known as Principal Component Analysis (PCA) [37]. Although PCA technically is part of finding the independence of the signals, it is usually considered a preprocessing step to the actual ICA algorithm. After the second-order decorrelation, the signal mixtures are orthogonal, i.e.,  $\tilde{s}_1(t) \cdot \tilde{s}_2(t) = 0$ . Immediately following

PCA, the signal mixtures are normalized to have the same power. Applying PCA and normalization together is known as “whitening” the signals.

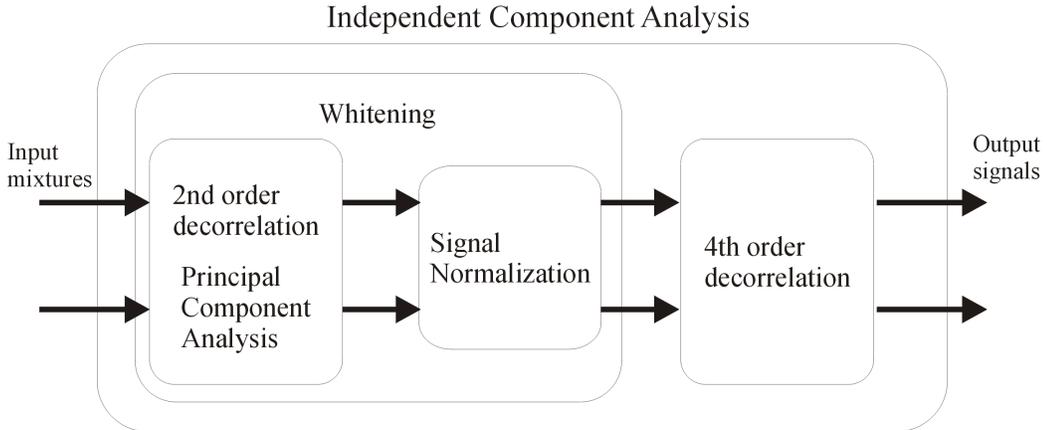


Figure 1.2: A high-level description of the ICA procedure. The ICA algorithms usually divide the problem into two parts: a second-order decorrelation of the signals followed by normalization (together called whitening) and a final fourth-order decorrelation.

### 1.3 Opto-electronic ICA

Like many ICA algorithms, the opto-electronic ICA processor performs the signal separation by first whitening the mixtures and then decorrelating them to higher orders. The PCA processor (shown in Figure 1.3) is an important initial step in the ICA and was primarily developed by Dr. Edeline Fotheringham, as described in detail in [38] and her thesis [39]. This thesis will just address the conceptual points of her work necessary for the understanding of the ICA processor design.

PCA requires that a second-order correlation is performed on the input signals. This is a computationally intensive process when implemented in the digital domain. There are, however, physical domains in which a correlation between signals is computed both quickly and in parallel with other correlations. This rapid correlation is one of the strengths of dynamic holography.

Dynamic holography, also known as two-beam coupling, is exploited for its correlation feature in the optoelectronic PCA system. Two-beam coupling in a photorefractive crystal is a method of amplifying one signal by using another stronger signal. The amplification only occurs when the two signals are correlated. Through two-beam coupling, an all-optical feedback loop gives gain to only one of the principal components present in the component mixture. The other components that do not receive gain from the two-beam coupling process are unable to oscillate within the feedback loop. Although the PCA processor described in [38] is designed to output only one of the principal components, it is possible to modify it to output all of the principal components that will later be used as the inputs to an optoelectronic ICA system.

This PCA system highlights a principle that is fundamentally different between an opto-electronic system and a computer algorithm. The system relies on the analog principle of two-beam coupling to provide the necessary decorrelation rather than on digitizing the signals so that they can be stored and decorrelated by a digital signal processor.

Since these systems are based on such a different concept as compared to

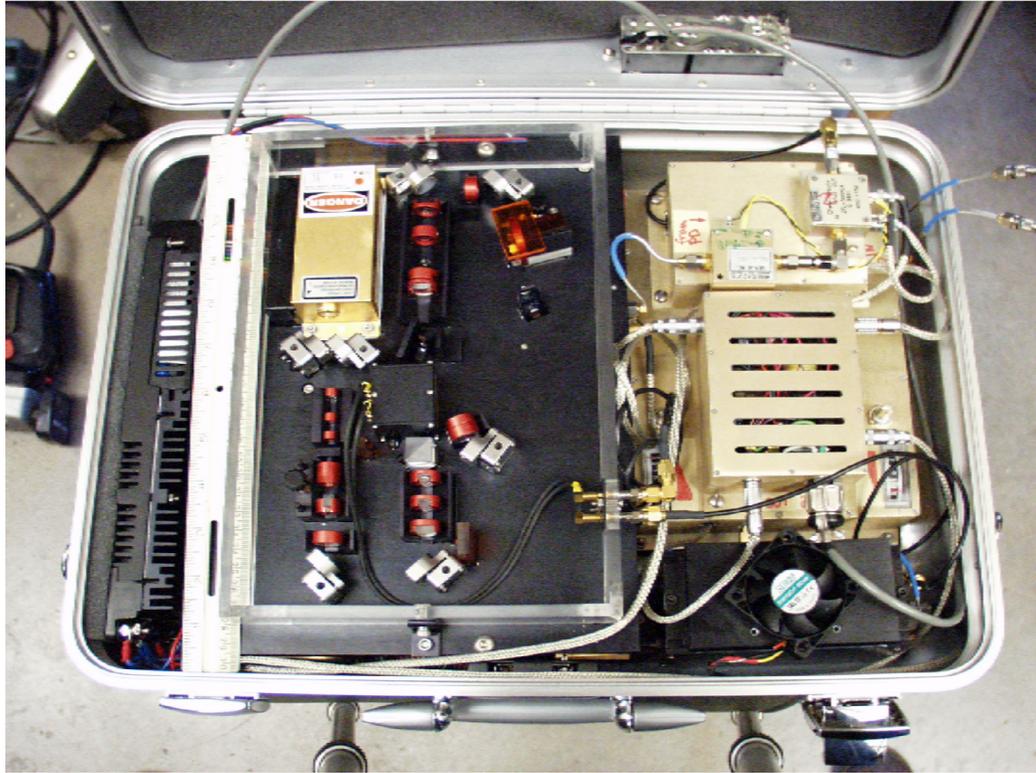


Figure 1.3: The optical Principal Component Analysis (PCA) system. This optical system has been reduced in size to fit into a common briefcase and uses two electrical input signal mixtures at a 10 GHz carrier frequency. The two cables located on the right hand side are inputs from the two-element antenna array. The output, also electrical, is one of the principal components of the input signals. This system can be expanded to output both of the principal mixtures as well, making it compatible with the ICA system analyzed in this thesis.

the step-by-step method of a computer algorithm, they must have a different infrastructure built around them. The essential components of the infrastructure used both in the PCA system (shown in Figure 1.4) and in the ICA system described in this thesis, are shown in Figure 1.5. The input signals

can be generated via nearly any means of signal generation, including arbitrary waveform generators, computer outputs, synthesizers, CD players, etc. The signals are then mixed using an electronic mixing circuit. As described in Equation 1.2, this combines the signals into two linear mixtures that will be input into the system. The signals are modulated onto an optical carrier (laser beam) using either electro-optic modulation or acousto-optic modulation. The usual, unsurpressed carrier, electro-optic modulation presents a problem, however, because it introduces a false correlation between the input signals. This false correlation can be removed through carrier suppression, described in [40], and the carrier suppressed signals can now be processed using opto-electronic ICA (Figure 1.5). The final step is to convert the signals back into electrical signals which can then be demodulated as required by their communications scheme. This is generally done using simple photodetectors. These generalized steps are shown in Figure 1.4.

The ICA system block diagram in Figure 1.5 generates the input signals and mixes them with a computer. The computer output comprises both the original signals and the mixtures. The original signals are used only to evaluate the performance of the ICA system. The mixtures are modulated on the laser beams in two identical input channels. The brain of the system is an optoelectronic feedback loop that performs the signal separation and will be described thoroughly in subsequent chapters. Like its PCA counterpart, it is designed to output only one of the two signals, although it can be modified to output both of the signals. Because it is difficult to directly detect phase

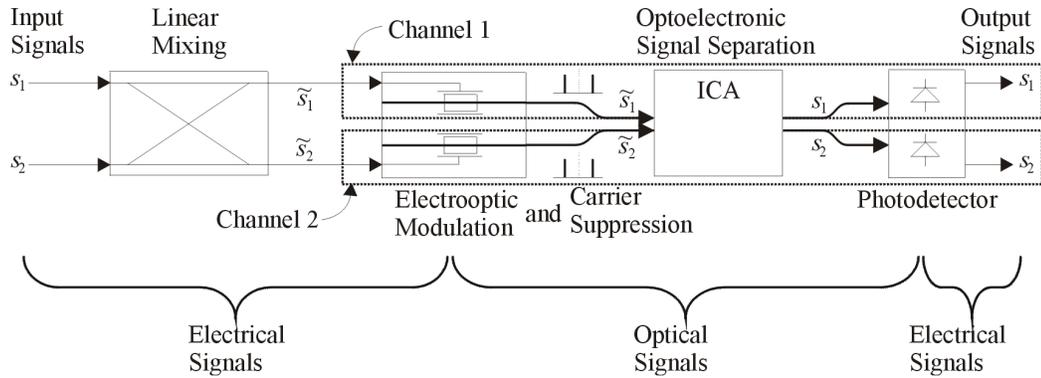


Figure 1.4: A two-channel optoelectronic blind-source-separation system. The electrical inputs are converted to optical signals through optical phase modulation, separated by the ICA system, and outputted as electrical signals.

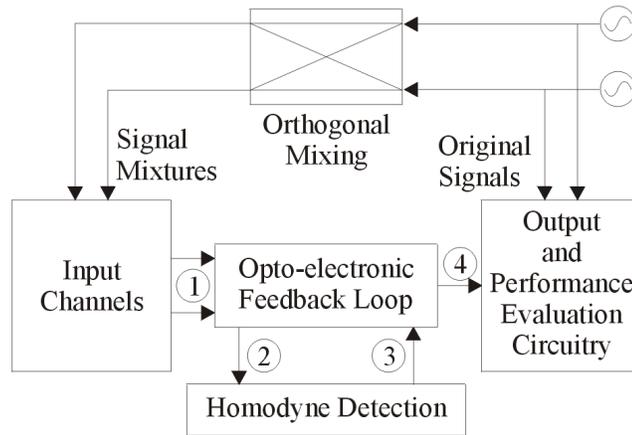


Figure 1.5: Generalized block diagram of the optoelectronic ICA system. The input mixtures and the original signals are created by a computer. The four principal subsystems are the input channels, the optoelectronic feedback loop, the homodyne detection, and the output.

modulated signals, they must be first converted to amplitude modulated signals by interfering them with a homodyne beam.

## 1.4 Thesis Outline

This thesis presents the first demonstration of an opto-electronic ICA system to the best of our knowledge. The thesis is organized as follows: chapter 2 gives a detailed description of the feedback loop and the theoretical behavior of the loop, chapter 3 gives a description of the 100 kHz bandwidth ICA system, chapter 4 gives the experimental results of the system. The prospect of future work, increasing the sytem bandwidth, and a complete blind source separation system are described in chapter 5. Also included is appendix A, which describes ICA theory.

# Chapter 2

## Opto-electronic ICA Theory

### 2.1 Introduction

Computer algorithms that implement ICA have been thoroughly analyzed [30], [41], [31]. This chapter is meant to lay out a theoretical foundation for the new opto-electronic ICA processor.

The problem that is being addressed here is not the complete blind separation of an arbitrary number of signals, but rather the separation of one signal from an orthogonal mixture of signals (Figure 2.1). It is assumed that the usual PCA and normalization preprocessing of the signals has already been performed. The system is designed to allow one (and only one) of the signals to oscillate within the opto-electronic feedback loop described in the next section. In applications that require both the signals to be recovered, the second signal can be retrieved by suppressing the first signal from one of

the input mixtures. This procedure is discussed in chapter 5.

This chapter is organized in the following way. It first describes in detail the opto-electronic feedback loop and each of the components that are instrumental in the actual separation process. A brief description of the method used to analyze the feedback loop is also given. The feedback loop is analyzed in two scenarios. The first is the specific case of exactly two sinusoidal input signals. This analysis shows that in steady state operation, only one of the two signals is able to oscillate in the loop. The second is the more general case of an arbitrary number of input signals with random probability distributions. This analysis shows that in steady state operation, only one of the signals oscillates in the loop provided that at least one of the signals has a sub-Gaussian distribution.

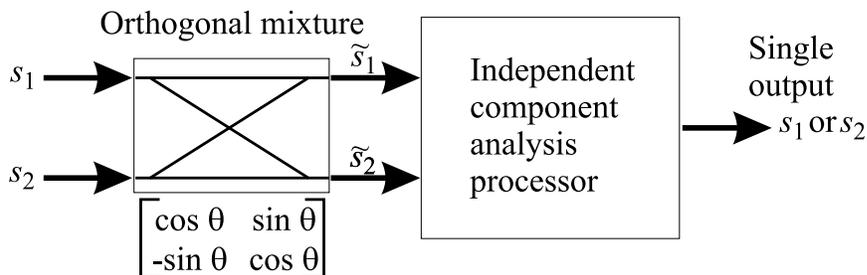


Figure 2.1: Schematic representation of ICA. The system described in this paper is designed to take orthogonal, normalized (whitened) input mixtures (two shown here) and separate out one of the original signals.  $\theta$  is any angle and  $\tilde{s}_1$  and  $\tilde{s}_2$  represent the orthogonal signals.

## 2.2 The opto-electronic feedback loop

A schematic of the heart of the opto-electronic ICA processor, a feedback loop first described in [42], is shown in Figure 2.2.

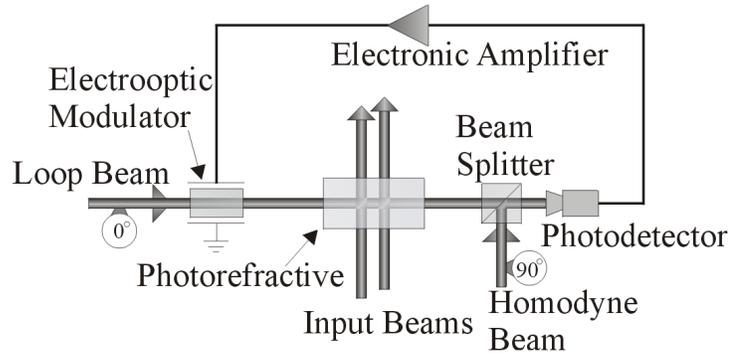


Figure 2.2: Schematic of the dynamic ICA feedback loop. The loop and input beams come from the same laser source. The input beams are modulated with the whitened input mixtures. The electrooptic feedback loop serves to saturate the phase modulator, producing higher order harmonics that will be correlated with the input signals in the photorefractive crystal.

The feedback loop is half optical, as it utilizes a loop laser beam, input laser beams, and a homodyne beam (all at a 532 nm wavelength and from the same source), and half electronic, as demonstrated by the photodetector and the electronic amplifier. The signal mixtures, which have been uncorrelated and have normalized power levels, are modulated on the input beams as shown at the bottom of Figure 2.2. The procedure for modulating these beams is discussed in chapter 3. There is the same number of input beams as there are signal mixtures (four in Figure 2.2). The feedback loop signal is the output and can be extracted from the loop anywhere, but is most

conveniently tapped just after the electronic amplification. The tapped signal contains -50 dB of the oscillating signal power.

The input beams provide gain to the loop beam through two-beam coupling in the photorefractive crystal. In two-beam coupling, holographic amplification can only occur between beams that are temporally correlated. This process is described more thoroughly in the next section. The signal that is amplified on the loop beam is a phase modulated signal, which cannot be detected by an intensity photodetector. The phase modulated signal is converted to an amplitude modulated signal by adding a homodyne beam just before detection. The homodyne beam is merely a laser beam with the same frequency as the original laser carrier. The homodyne beam is equivalent to the local oscillator in a coherent detection radio. In order to produce an amplitude modulation, the homodyne beam must be phase shifted by 90 degrees from the original loop beam laser carrier. After detection, electronic amplification increases the signal voltage on the electro-optic phase modulator. The phase modulator, which is inherently nonlinear, modulates the signal on the loop laser beam, completing one round trip of the loop.

Recall from chapter 1 that ICA can only be performed if the signals are independent. The definition of independence, Equation 1.1, implies that the signals are de-correlated to all orders. The dynamical behavior of the feedback loop is a result of the combination of the nonlinear electro-optic modulator with photorefractive two-beam amplification. Nonlinear phase modulation generates signal harmonics on the loop laser beam. Holographic

amplification then ensures that the correlated portions of these harmonics are amplified in the feedback loop.

A mathematical analysis of the feedback loop begins by describing the gain of each of the components of the loop, which determine the open-loop system gain, used to generate equations for the time-evolution of the input signal slowly-varying amplitudes. The stability of the steady state solutions is found using perturbation theory. The analysis theoretically predicts the amplitudes of each of the signals oscillating in the feedback loop when the system is stable.

## 2.3 Sinusoidal signal analysis

### 2.3.1 Behavior of two input signals

The analysis procedure described above is carried out in this section on two sinusoidal signals,  $s_1$  and  $s_2$ , of frequencies  $\omega_1$  and  $\omega_2$ , and slowly-varying amplitudes  $v_1(t)$  and  $v_2(t)$ :

$$\begin{aligned} s_1 &= v_1(t) \sin(\omega_1 t) \\ s_2 &= v_2(t) \sin(\omega_2 t) \end{aligned} \tag{2.1}$$

The amplitudes of the signals,  $v_1(t)$  and  $v_2(t)$  evolve on a time scale determined by the photorefractive time constant, which is on the order of seconds. The signals vary much faster, on the order of ten microseconds, for signals with a 100 kHz bandwidth.

The goal of this analysis is to find the evolution of the slowly-varying amplitudes. It is a simple calculation to show that these signals are independent, as required by ICA, if  $p\omega_1 \neq q\omega_2$  where  $p$  and  $q$  are integers:

$$\langle \sin^m(p\omega_1 t) \sin^n(q\omega_2 t) \rangle = \langle \sin^m(p\omega_1 t) \rangle \langle \sin^n(q\omega_2 t) \rangle \quad (2.2)$$

and where  $\langle \rangle$  denotes the statistical expectation value.

Let us now assume that some combination of these signals is present in the feedback loop just before the phase modulator. The signals modulate the laser beam electric field,  $E$ , as:

$$E = E_0 \exp(-i(v_1 \sin(\omega_1 t) + v_2 \sin(\omega_2 t))) \quad (2.3)$$

where  $E_0$  is the loop beam amplitude and is normalized to 1. Here, the explicit time dependence on the slowly varying amplitudes,  $v_1$  and  $v_2$ , has been omitted. The loop amplitude signal can be rewritten as a Bessel expansion:

$$E = \sum_{m=-\infty}^{\infty} J_m(v_1) \exp(-im\omega_1 t) \sum_{n=-\infty}^{\infty} J_n(v_2) \exp(-in\omega_2 t) \quad (2.4)$$

where  $J_m$  is the  $m$ th order Bessel function of the first kind.

The loop signal and the input signals will interfere in the photorefractive to produce a holographic grating. The photorefractive time constant is on the order of milliseconds to seconds. Such a slow time constant implies that a grating can only be generated by the crystal between two beams that are the exact same frequency (within 1 to 10 Hz) since the interference pattern between two beams with different frequencies will be moving.

In addition to the correlation, the photorefractive crystal provides a second advantage when placed in the feedback loop. As seen in Figure 2.2, there are multiple input beams that are coupled with the loop beam within the holographic medium. If now a second input beam is added, the same power transfer will occur provided the total power in the two input beams is the same as in the single-input beam case. This implies that the feedback loop gives the same gain to signals of any orthogonal mixture (a mixture that has the same power in both inputs). Since the system input mixtures are already whitened, the system will behave the same way for any input mixture. This implies that the actual signal mixture is irrelevant, provided that it is orthogonal, and the analysis can proceed assuming that the signals always receive the same gain as they would if they were present on separate beams.

The correlation requirement implies that the only two components of  $E(t)$  that will be amplified in the feedback loop are those that are present at  $\omega_1$  and  $\omega_2$ :

$$\begin{aligned} E_{\omega_1} &= J_0(v_2) J_1(v_1) \exp(-i\omega_1 t) \\ E_{\omega_2} &= J_0(v_1) J_1(v_2) \exp(-i\omega_2 t) \end{aligned} \tag{2.5}$$

The signal output from the photorefractive crystal will have only a field proportional to these two terms, when  $m = 1$  and  $n = 0$  or  $m = 0$  and  $n = 1$ . Since the input signals are already assumed to be normalized to the same power, the power in each of the two components of Equation 2.5 are also of

the same power, and the amplitudes are also normalized to 1.

### 2.3.2 Sinusoidal dynamic evolution

At this point in the analysis, it becomes instructive to expand the Bessel functions to the first three terms. This simplifying assumption is not made in the following section with a more general analysis. The Bessel functions expand as:

$$\begin{aligned} J_0(v) &= 1 - \frac{v^2}{4} + \dots \\ J_1(v) &= \frac{v}{2} - \frac{v^3}{16} + \dots \end{aligned} \tag{2.6}$$

The homodyne detection is set up so that the received signal is directly proportional to the phase modulation on the loop beam. The detected signal is also linearly amplified by the electronic amplifier. The resulting round trip signal amplitudes,  $v_{1return}$  and  $v_{2return}$  are:

$$\begin{aligned} v_{1return} &= G \left( 1 - \frac{v_2^2}{4} \right) \left( \frac{v_1}{2} - \frac{v_1^3}{16} \right) \\ v_{2return} &= G \left( 1 - \frac{v_1^2}{4} \right) \left( \frac{v_2}{2} - \frac{v_2^3}{16} \right) \end{aligned} \tag{2.7}$$

where the linear loop gain is  $G$ . The two equations shown here represent the amount that the amplitude of each signal,  $s_1$  and  $s_2$ , changes in one round trip of the feedback loop. A set of gain equations, based on the initial and final signal amplitudes can be found with the assumption that one round trip of the feedback loop takes one unit of time, where  $\Delta v_1$  and  $\Delta v_2$  are the

amplitude differences in  $v_1$  and  $v_2$ :

$$\begin{aligned}\Delta v_1 &= G \left(1 - \frac{v_2^2}{4}\right) \left(\frac{v_1}{2} - \frac{v_1^3}{16}\right) - v_1 \\ \Delta v_2 &= G \left(1 - \frac{v_1^2}{4}\right) \left(\frac{v_2}{2} - \frac{v_2^3}{16}\right) - v_2\end{aligned}\tag{2.8}$$

The difference gain equations can be turned into differential equations:

$$\begin{aligned}\frac{\partial v_1}{\partial t} &= \left(\frac{G}{2} - 1\right) - G\frac{1}{8}v_1v_2^2 - G\frac{1}{16}v_1^3 \\ \frac{\partial v_2}{\partial t} &= \left(\frac{G}{2} - 1\right) - G\frac{1}{8}v_2v_1^2 - G\frac{1}{16}v_2^3\end{aligned}\tag{2.9}$$

We now have a set of two coupled differential equations showing the time evolution of the signal amplitudes.

### 2.3.3 System steady state behavior

Next in the analysis, Equation 2.9 is used to predict how the feedback loop will behave given the two inputs. The evolution of Equation 2.9 will eventually reach a steady state solution, when amplitudes  $v_1$  and  $v_2$  are stationary. This steady state solution can be found by setting both differentials to zero. We find that there are three possible solutions:

$$v_1 = v_2 = 0,\tag{2.10}$$

$$v_1 = v_2 = \left(\frac{\left(\frac{G}{2} - 1\right)}{\frac{3}{16}G}\right)^{\frac{1}{2}},\tag{2.11}$$

$$v_1 = 0, \quad v_2 = \left( \frac{\left(\frac{G}{2} - 1\right)}{\frac{1}{16}G} \right)^{\frac{1}{2}} \quad \text{or} \quad v_2 = 0, \quad v_1 = \left( \frac{\left(\frac{G}{2} - 1\right)}{\frac{1}{16}G} \right)^{\frac{1}{2}}. \quad (2.12)$$

The first solution is the trivial case where neither of the signals is oscillating in the feedback loop. The second solution is the case where both signals are oscillating with equal amplitudes within the loop. The third solution is the case where only one of the two signals is oscillating in the loop, and the other is not present.

It is of interest to find out which of these solutions is stable, and what the conditions of stability are on each. To accomplish this task, we turn back to the initial dynamic Equation 2.9. The amplitudes will be perturbed by small amounts,  $\epsilon_1$  and  $\epsilon_2$ , (also dependent on time,  $t$ ) for the two signals, respectively. In the case of a stable solution, the signals decay back to the steady state solution after this perturbation. If the solution is unstable, the signals will deviate from the steady state solution.

The initial equations are:

$$\begin{aligned} \frac{\partial(v_1+\epsilon_1)}{\partial t} &= \left(\frac{G}{2} - 1\right) - G\frac{1}{8}(v_1 + \epsilon_1)(v_2 + \epsilon_2)^2 - G\frac{1}{16}(v_1 + \epsilon_1)^3 \\ \frac{\partial(v_2+\epsilon_2)}{\partial t} &= \left(\frac{G}{2} - 1\right) - G\frac{1}{8}(v_2 + \epsilon_2)(v_1 + \epsilon_1)^2 - G\frac{1}{16}(v_2 + \epsilon_2)^3 \end{aligned} \quad (2.13)$$

These can be rearranged as:

$$\begin{aligned} \frac{\partial v_1}{\partial t} + \frac{\partial \epsilon_1}{\partial t} &= \left[\left(\frac{G}{2} - 1\right) - G\frac{1}{8}v_1v_2^2 - G\frac{1}{16}v_1^3\right] + \left(\frac{G}{2} - 1\right)\epsilon_1 - \frac{G}{8}(2v_1v_2\epsilon_2 + v_2\epsilon_1) - \frac{G}{16}(3v_1^2\epsilon_1) \\ \frac{\partial v_2}{\partial t} + \frac{\partial \epsilon_2}{\partial t} &= \left[\left(\frac{G}{2} - 1\right) - G\frac{1}{8}v_2v_1^2 - G\frac{1}{16}v_2^3\right] + \left(\frac{G}{2} - 1\right)\epsilon_2 - \frac{G}{8}(2v_2v_1\epsilon_1 + v_1\epsilon_2) - \frac{G}{16}(3v_2^2\epsilon_2) \end{aligned} \quad (2.14)$$

where all higher orders of the perturbation terms,  $\epsilon^n, n \geq 2$ , have been omitted.

Since the analysis is occurring at the steady state value, the  $\frac{\partial v}{\partial t}$  terms are equal to 0, and the components of Equation 2.14 that are in brackets will drop out (see Equation 2.9), leaving only a set of coupled differential equations for the perturbations. If these perturbations decay to zero amplitude over time, then the steady state is stable. The perturbation differential equations can be represented as:

$$\frac{\partial \epsilon}{\partial t} = \begin{pmatrix} \left(\frac{G}{2} - 1\right) - G\frac{1}{8}v_2^2 - G\frac{3}{16}v_1^2 & -G\frac{1}{4}v_1v_2 \\ -G\frac{1}{4}v_1v_2 & \left(\frac{G}{2} - 1\right) - G\frac{1}{8}v_1^2 - G\frac{3}{16}v_2^2 \end{pmatrix} \epsilon. \quad (2.15)$$

The solution to this system of equations takes the form:

$$\epsilon(t) = C_1 \begin{pmatrix} \xi_1^{(1)} \\ \xi_2^{(1)} \end{pmatrix} \exp(\lambda_1 t) + C_2 \begin{pmatrix} \xi_1^{(2)} \\ \xi_2^{(2)} \end{pmatrix} \exp(\lambda_2 t) \quad (2.16)$$

where  $C_1$  and  $C_2$  are constants,  $\lambda_1$  and  $\lambda_2$  represent the eigenvalues of the matrix in Equation 2.16, and  $\xi^{(1)}$  and  $\xi^{(2)}$  are the eigenvectors. From this solution, it is clear that the perturbations will decay if both eigenvalues are negative.

Now that we have a definition of the stability, we can solve for the original three cases. The first trivial case of neither signal oscillating in the loop is stable for  $G < 2$ . This implies that there must be enough gain within the feedback loop to make the signals oscillate. For the second case of both signals

oscillating in the feedback loop with equal amplitudes, one eigenvalue is negative while the other is positive. This implies that this case is never stable for two sinusoidal signals. The third case, with only one signal oscillating in the system, has two negative eigenvalues. This case is therefore stable for two sinusoidal signals.

### 2.3.4 Steady state amplitude

To determine the amplitude of the oscillating signal will be in terms of physical parameters in the feedback loop, the open-loop gain is found from Equation 2.5. From this equation, the gain is proportional to the first order Bessel function,  $J_1$ . This implies that the gain is zero when  $J_1(v_1) = 0$ , at  $v_1 = 3.83$  volts/V .

The feedback loop will oscillate at steady state when the gain is unity, assuming that there is no loss in the loop. The curve determining the open-loop gain is shown in Figure 2.3. The y-axis shows the initial gain of the feedback loop to be  $G$ , or the linear loop gain. This parameter can be adjusted by the gain of the electronic amplifiers,  $G_e$ , the linear gain of the photorefractive crystal,  $G_p$ , and the loss of the loop,  $L$ , where  $G = G_e + G_p - L$ .

If the photorefractive crystal gain is removed from the system, and the total gain is still greater than unity, the loop will break into spontaneous oscillation. This is undesirable behavior of the system, so we will assume that the electronic gain is set so that the total gain is just below unity. In

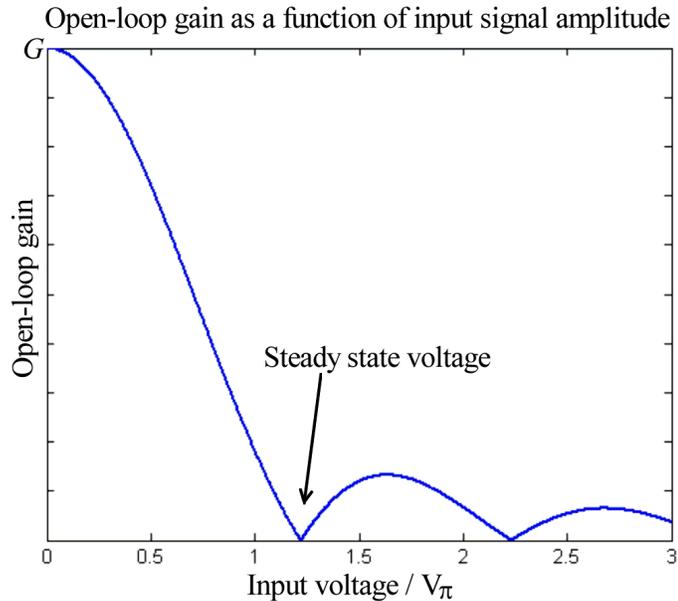


Figure 2.3: Open-loop gain of the opto-electronic feedback loop (Figure 2.2) as a function of input signal amplitude. The gain that a sinusoidal signal receives in the feedback loop is determined by its amplitude according to the Bessel relationship found in Equation 2.5.

essence, the electronic gain almost completely compensates for the loss in the loop. The amount by which  $G$  is larger than unity is therefore dependent solely on the photorefractive gain. If the photorefractive gain is 300, then the initial gain on this curve will be approximately 300.

Figure 2.3 shows that the gain of the feedback loop decreases as the input signal amplitude increases. The gain reaches a value of zero when the input amplitude is  $\pm 1.22$  V or  $3.83/\pi$ . The gain is unity just below this voltage, but  $\pm 1.22$  V is a close approximation. We now know that the signal in the feedback loop will be stable when it is oscillating with a drive voltage

of  $\pm 1.22$  V . The signal amplitude is therefore dependent on the electro-optic modulator and is much higher than the common drive voltages that are used with electro-optic modulators. Recall that higher order correlations perform the independent component analysis, and harmonics of the signals are produced by overdriving the modulator.

If the total system gain is decreased by lowering  $G$  in Figure 2.3, the steady state input voltage will also be decreased. The amount of decrease is shown by Figure 2.4. The oscillation voltage is shown by the intersection of the loss curve with the gain curve. In a real system, we will most likely not be able to operate at the voltage of  $\pm 1.22$  V because this will require the gain to be right on the verge of making the system spontaneously oscillate. Instead, it is more likely that the gain will have to be lowered purposefully to stabilize the system. This means that the actual oscillating voltage will be slightly less than the ideal  $\pm 1.22$  V.

The analysis of two sinusoidal input signals shows that the steady state behavior of the feedback loop is such that it allows only one of the signals to exist in the loop. The amplitude of the signal is determined by the total loop gain and the half-wave voltage of the modulator.

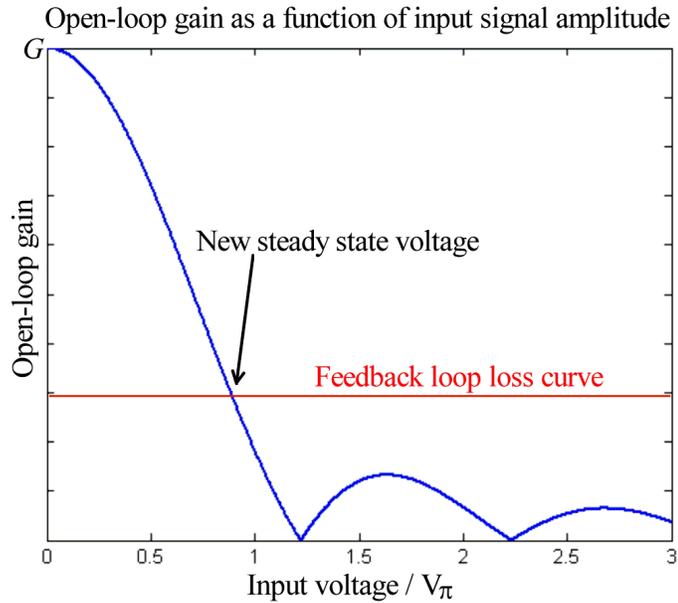


Figure 2.4: Open-loop gain as a function of input signal amplitude with feedback loop loss. As the feedback loop loss increases or the photorefractive gain decreases, the steady state voltage of the oscillating signal decreases.

## 2.4 Arbitrary probability density function analysis

### 2.4.1 Behavior of two input signals

While the two sinusoidal input case provides intuition into the behavior of the nonlinear feedback loop, it does not address the important issue of multiple unknown signals with arbitrary probability density functions. Similarly to the sinusoidal case, we will first assume a drive voltage,  $V(t)$ , on the electro-

optic modulator that consists of a superposition of  $N$  signals:

$$V(t') = \sum_{n=1}^N v_n(t') s_n(t). \quad (2.17)$$

Here, the slowly varying envelopes of the signals are given by  $v_n(t')$ , and the signals themselves are  $s_n(t)$ .  $t'$  is used to represent the slowly varying time frame, while  $t$  represents the quickly varying signals, which vary more than three orders of magnitude more quickly. Without loss of generality, the signals are again assumed to be normalized to the same power.

The loop laser beam is phase modulated by the signal  $V(t)$ :

$$E(t, t') = E_0 e^{iV(t)} e^{-i\omega_c t} + [E_0 e^{iV(t)} e^{-i\omega_c t}]^* \quad (2.18)$$

where  $E_0$  is the amplitude of the loop beam and  $\omega_c$  is the laser frequency.

As in the sinusoidal case, the signal writes a grating with the correlated terms of the input beams in the photorefractive crystal and is amplified by that correlation. Mathematically, this type of gain is represented by:

$$G_{photorefractive} \propto \langle e^{iV(t)} s_n(t) \rangle_{\tau} \quad (2.19)$$

where  $\tau$  is the photorefractive time constant (milliseconds to seconds) and represents the integration time over which the signals are correlated [43]. As in the two-signal sinusoidal case, the actual mixture of the signals on the two input beams is irrelevant provided that it is an orthogonal mixture.

Continuing along the feedback loop, the signals is detected through the homodyne detection scheme and amplified by the electronic amplifiers. This

new signal is represented as:

$$\tilde{v}_n = -\Re \left\{ i \left\langle e^{iV(t)} s_n(t) \right\rangle_\tau \right\} \kappa v_n. \quad (2.20)$$

where  $\kappa$  is the linear gain in the feedback loop, and  $\tilde{v}_n$  is the return signal amplitudes. The real part of the complex function and the negative sign are a consequence of the homodyne detection method. The gain can be split into two terms: the linear  $\kappa$  term, and the nonlinear term,  $A_n$ :

$$A_n(\vec{v}) = -\text{Re} \left\{ \frac{i}{v_n} \left\langle e^{iV(t)} s_n(t) \right\rangle_\tau \right\} \quad (2.21)$$

where  $A_n$  is dependent on all the signal amplitudes,  $v_1, v_2, \dots, v_N$ , since it contains  $V(t)$ . The amplitudes  $v_1, v_2, \dots, v_N$  are expressed as a multidimensional vector,  $\vec{v}$ .

Up to this point, the gain of the system has been represented using the time-averaged correlation due to the photorefractive gain. Instead of a time-varying signal, the input signals can be viewed as stochastic variables with probability density function  $p_n(s_n)$ . This implies that the time-averaged correlation can be expressed as a function of the signal probability density function:

$$\left\langle e^{iV(t)} s_n(t) \right\rangle_\tau \rightarrow \int_{-\infty}^{\infty} e^{iV(t)} s_n p_n(s_n) ds_n \quad (2.22)$$

and the nonlinear gain term becomes:

$$A_n(\vec{v}) = -\text{Re} \left\{ \frac{i}{v_n} \int_{-\infty}^{\infty} e^{iV(t)} s_n p_n(s_n) ds_n \right\}. \quad (2.23)$$

While this function seems more complex, it can actually be expressed very compactly using the characteristic function,  $\phi(v)$ , which is the Fourier transform of the probability density function:

$$\phi(v) = \int_{-\infty}^{\infty} e^{ivs} p(s) ds \quad (2.24)$$

where  $s$  is a random variable with probability density  $p$  [44]. The characteristic function of a multidimensional set of random variables,  $\Phi(\vec{v})$  is simply the multidimensional Fourier transform, which reduces to the product of the individual characteristic functions when the variables are independent:

$$\Phi(\vec{v}) = \phi(v_1) \phi(v_2) \dots \phi(v_N). \quad (2.25)$$

Now the nonlinear gain term reduces to:

$$A_n(\vec{v}) = -\frac{1}{v_n} \text{Re} \{ \partial_n \Phi(\vec{v}) \}. \quad (2.26)$$

This is the open-loop gain of the feedback loop for the signal amplitudes  $v_n(t)$ . Equation 2.26 shows that the gain of each signal is dependent on both the amplitude of that signal itself, and the derivative of the multidimensional characteristic function. In more general terms, the gain in the loop is determined by the probability density function of the input signals.

## 2.4.2 Gain of various probability density functions

As an example of how the gain in the feedback loop depends on the characteristic function, let us take four different probability distributions, and

plot their gains. The four distributions are: harmonic, uniform, Gaussian, and Laplace, shown in Figure 2.5. These functions are represented by the probability distributions:

$$p_L(v) = \frac{1}{2} \exp(-|v|) \quad \text{Laplace} \quad (2.27)$$

$$p_G(v) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{|v|^2}{2}\right) \quad \text{Gaussian} \quad (2.28)$$

$$p_U(v) = \begin{cases} \frac{1}{2A} & \text{if } -A \leq v \leq A \\ 0 & \text{otherwise} \end{cases} \quad \text{Uniform} \quad (2.29)$$

$$p_H(v) = \frac{1}{\pi\sqrt{1-v^2}} \quad \text{Harmonic.} \quad (2.30)$$

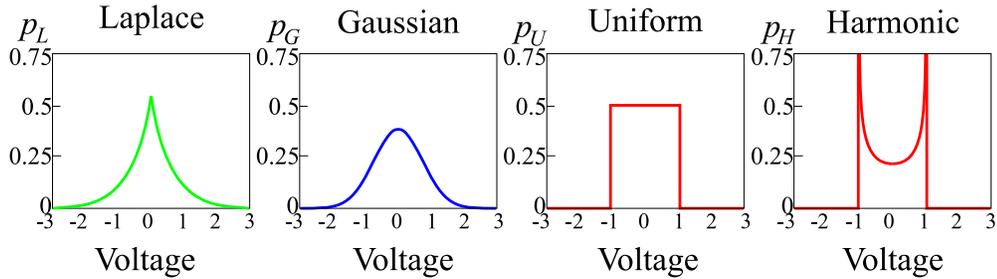


Figure 2.5: Four probability density functions that are analyzed in the feedback loop.

While Laplace, Gaussian, and uniform distributions are common in literature on statistics, a harmonic distribution is not often seen. This is the distribution of a sinusoidal signal, common for communications signals, and

is therefore of practical importance. The distribution type does not change is the sinusoidal carrier is modulated by frequency modulation (FM), phase modulation (PM), and any phase-shift keying scheme (BPSK, QPSK).

The characteristic function of each of the distributions in Equation 2.27 - Equation 2.30 is:

$$\phi_L(v) = \left(1 + \frac{v^2}{2}\right)^{-1} \quad \text{Laplace} \quad (2.31)$$

$$\phi_G(v) = \exp\left(\frac{-v^2}{2}\right) \quad \text{Gaussian} \quad (2.32)$$

$$\phi_U(v) = \frac{\sin(v\sqrt{3})}{v\sqrt{3}} \quad \text{Uniform} \quad (2.33)$$

$$\phi_H(v) = J_0(v\sqrt{2}) \quad \text{harmonic} \quad (2.34)$$

where  $J_0$  is the zeroth order Bessel function.

The gains for each of these probability distributions in the feedback loop is shown in Figure 2.6. This figure shows that the distribution that has the highest gain within the feedback loop is the harmonic distribution, followed by the uniform, Gaussian, and Laplace. The figure is normalized to the maximum gain of each probability density function, which is equal for all the functions.

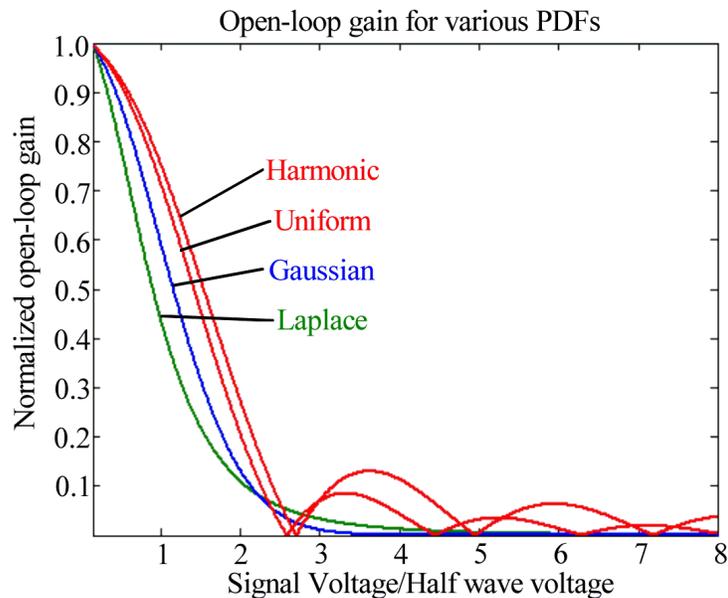


Figure 2.6: The open-loop gains of harmonic, uniform, Gaussian, and Laplace density functions. Each of these functions has the same maximum gain and is normalized to unity.

### 2.4.3 Dynamic evolution of arbitrary PDF signals

In a similar method as was used in the two-signal sinusoidal case, the open-loop gain can be turned into a time-evolution equation for the input signal amplitudes:

$$\tau \frac{\partial \vec{v}}{\partial t} = -(\vec{v} + \kappa \text{Re} \{ \nabla \Phi \}) \quad (2.35)$$

where  $\nabla$  is the multidimensional gradient operator.

#### 2.4.4 *Lotka-Volterra* simplifications and Kurtosis

By itself, Equation 2.35 is not intuitive. It represents a particle moving in a plane and subject to a restoring force given by the derivative of the characteristic function. We can make some assumptions to simplify this equation and provide insight into the underlying dynamics. First, let us assume that there are only two input signals. This assumption again gives us two coupled differential equations. Second, the characteristic function can be expanded in terms of the moments of the signal probability density function:

$$\phi(v) = 1 + iv \langle s \rangle - \frac{1}{2}v^2 \langle s^2 \rangle - \frac{1}{3!}iv^3 \langle s^3 \rangle + \frac{1}{4!}v^4 \langle s^4 \rangle + \dots \quad (2.36)$$

Third, we will transform the equation from an amplitude evolution equation to a power evolution equation. Lastly, we will assume that the two input signals have the same probability distribution functions. This implies that they will have the same fourth moments,  $\langle s_{1,2}^4 \rangle$ . With these added assumptions, we find that the coupled equations become the *Lotka-Volterra* equations [45]:

$$\begin{aligned} \frac{\tau}{2\kappa} \frac{\partial P_1}{\partial t} &= \left( \alpha - \frac{1}{3} \langle s_{1,2}^4 \rangle P_1 - P_2 \right) P_1 \\ \frac{\tau}{2\kappa} \frac{\partial P_2}{\partial t} &= \left( \alpha - \frac{1}{3} \langle s_{1,2}^4 \rangle P_2 - P_1 \right) P_2 \end{aligned} \quad (2.37)$$

where  $P_1$  and  $P_2$  are the signal powers, and  $\alpha = 1 - \frac{1}{\kappa}$  is a measure of the linear gain. The dynamic equations have one pivotal coefficient: the fourth order expectation value.

Another statistical measure, known as the signal kurtosis,  $\gamma$ , also uses the fourth-order expectation value. Kurtosis is generally used as a measure of the Gaussianity of a signal, and is defined as:

$$\gamma = \frac{\langle s^4 \rangle}{\langle s^2 \rangle^2}. \quad (2.38)$$

Figure 2.7 shows a comparison of several different probability densities, each with normalized variance, and their kurtosis. The Gaussian distribution always has a kurtosis of 3. Signals that have a kurtosis between 0 and 3 are known as sub-Gaussian because they have a kurtosis lower than that of a Gaussian. These signals have probability densities that are flatter in the middle and have higher amplitudes at edge voltages. These include the binary distribution (the distribution of a binary,  $\pm 1$ , signal), the harmonic distribution, the uniform distribution, and the  $\exp(-|s|^3)$ . Likewise, signals that have a kurtosis that is greater than 3 are known as super-Gaussian signals. These signals have distributions that are more peaked than a Gaussian and have less energy in the wings of the distribution. Examples of these signals include a Laplace and a voice signal, as shown in Figure 2.7.

#### 2.4.5 *Lotka-Volterra* steady state solutions

Equation 2.37 have three possible steady-state solutions, depending on the  $\frac{\langle s_{1,2}^4 \rangle}{3}$  term: a monostable solution if the term is greater than unity, a metastable solution if the term is equal to unity, and a bistable solution if the term is

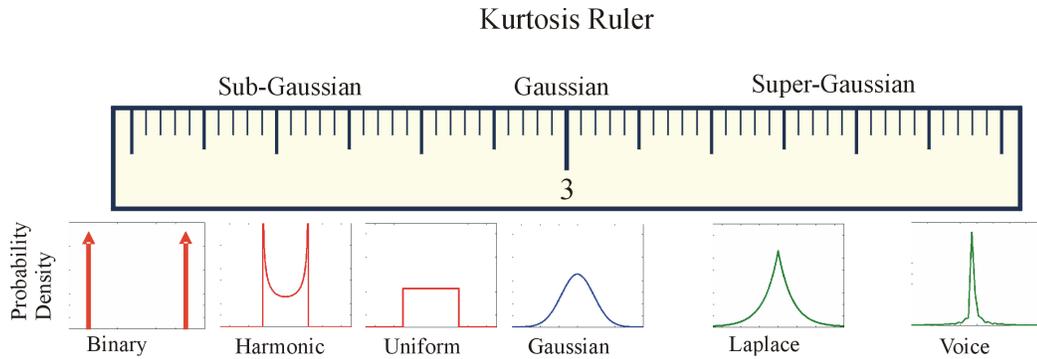


Figure 2.7: Illustration of kurtoses. The kurtosis, a measure of a signal's Gaussianity, increases for each PDF shown from left to right. (The PDFs are shown for descriptive purposes only- they are not to scale on the kurtosis ruler.)

less than unity. Since this term is based on a fourth-order expectation, it is also directly related to the kurtosis of the input signals. This implies that the dynamics of the system can be cast in terms of the Gaussianity of the input signals:

- If the system inputs are both super-Gaussian, the stable signal that exists in the loop will be an equal combination of both inputs.
- Gaussian inputs will result in a random orthogonal combination of the signal.
- If both signals are sub-Gaussian, only one of the two signals, chosen with equal probability, can occur.

The sub-Gaussian winner-takes-all solution is the desired steady-state operation of the loop because only one signal is present in the loop, and therefore the signals are separated out of the mixture.

## 2.4.6 Complete system dynamics

The dynamics of the system when two signals with different probability density functions are competing can be found directly from the original evolution, Equation 2.35. Here, we will make state-space vector diagrams representing the force on an input vector of signal amplitudes. To make a state-space diagram, two input signal probability density functions are first chosen. Next, input amplitudes,  $\vec{v} = v_1 s_1 + v_2 s_2$ , are inserted into Equation 2.35, resulting in the differential change in amplitudes. A vector corresponding to the direction and magnitude of the differential change is placed on the diagram at position  $(v_1, v_2)$ .

The diagrams have two axes, representing the voltage amplitude of each of the input signals. All points along the horizontal axis represent finite voltage amplitude for  $s_1$ , and zero amplitude for  $s_2$ . Similarly, all points along the vertical axis represent zero amplitude for  $s_1$ , and finite amplitude for  $s_2$ . A point off either axis means that both signals have finite amplitudes. The arrows spaced throughout the diagram represent the direction in which the amplitudes will change given some starting position. All arrows on the plot are reduced to a unity length, but color coated to represent their actual length. The light gray arrows represent shorter vectors, and the dark gray

and black arrows represent long vectors. Steady states are present where all vectors point either into or away from a point on the diagram. If the vectors are pointing away from a location, it is an unstable steady state solution. If the vectors are pointing toward a location, it is a stable steady state solution.

Figure 2.8 shows the state-space diagram for two harmonic distribution (sub-Gaussian) signals. There are stable steady state solutions along both the horizontal and vertical axes. Like in the sinusoidal two-input analysis, this means that either signal oscillating in the feedback loop is stable, but both cannot coexist at the same time. Again, this is the desired operation of the system and it corresponds to solution found by looking at the *Lotka-Volterra* simplification.

The competition between two uniform distribution signals is shown in Figure 2.9. This competition also shows stable steady states when either one of the signals is oscillating.

A competition between two Gaussian distribution signals results in a metastable solution, as shown in Figure 2.10. In this case, there is a ring of states that the system can “fall” into. There is no truly stable solution, however, and the system will randomly shift through the signal amplitudes represented by the ring.

A competition between two super-Gaussian distributions, both Laplace, is shown in Figure 2.11. A monostable solution with both signals oscillating with equal amplitudes is the only stable solution in this case. This again corresponds to the result found from the *Lotka-Volterra* simplification.

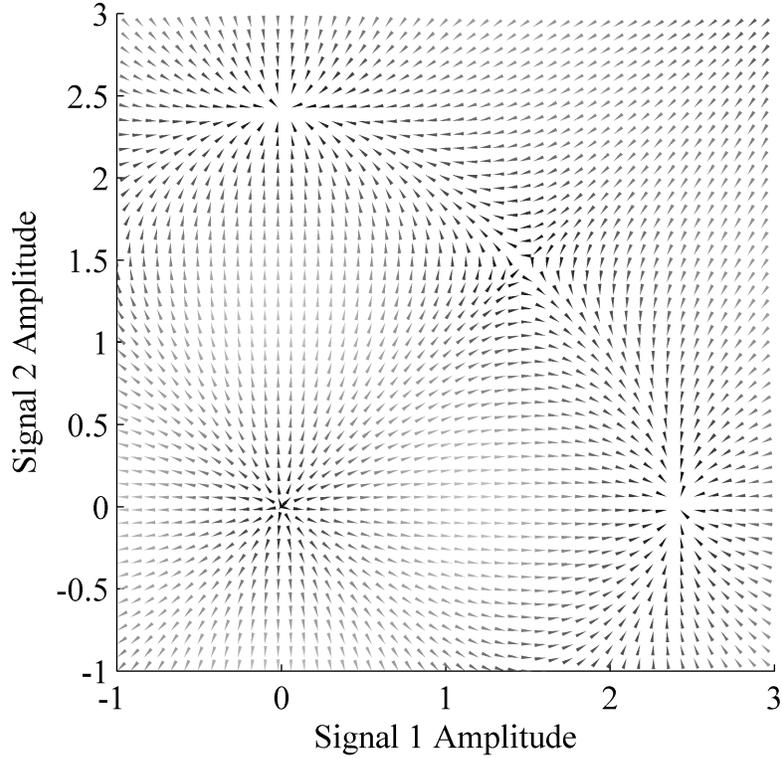


Figure 2.8: Harmonic vs. harmonic system dynamics. The evolution of two harmonic signals moves from the unstable equilibrium at point  $v_1 = 0, v_2 = 0$  to a stable equilibrium where either signal 1 or signal 2 exists.

It is interesting to investigate what happens to the system when one of the inputs is sub-Gaussian (harmonic) while the other is super-Gaussian (Laplace). In this case, represented by the state space diagram in Figure 2.12, the system has only one stable solution: the sub-Gaussian signal oscillates in the feedback loop.

In the case of a Gaussian signal competing with a sub-Gaussian (har-

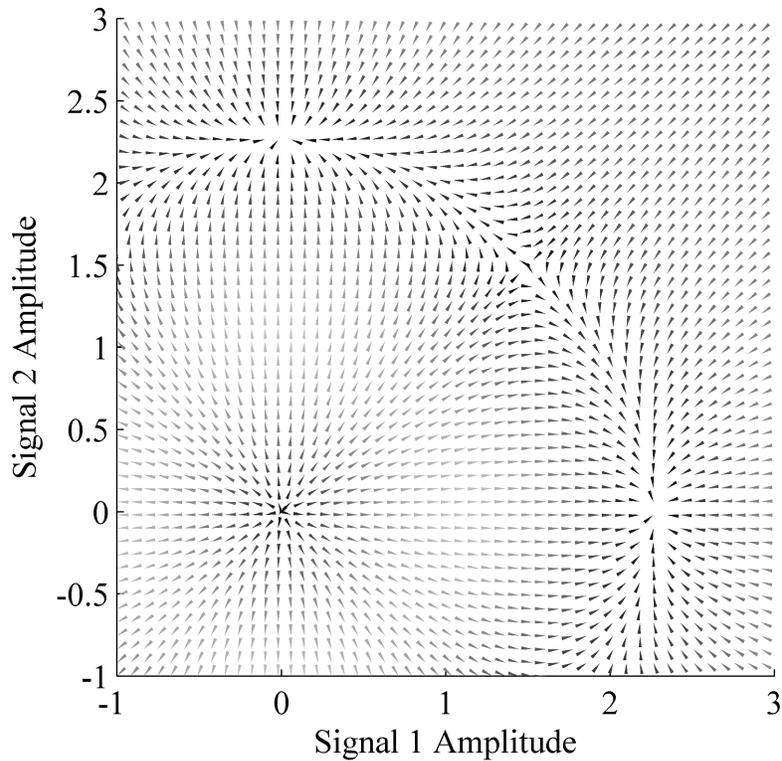


Figure 2.9: Uniform vs. uniform system dynamics. The evolution of two uniform signals moves from the unstable equilibrium at point  $v_1 = 0, v_2 = 0$  to a stable equilibrium where either signal 1 or signal 2 exists.

monic) signal, the sub-Gaussian signal wins, as shown in Figure 2.13.

In the case of a Gaussian competing with a super-Gaussian (Laplace) signal, the Gaussian signal wins, as seen in Figure 2.14. This is an interesting case that shows how the system behaves at the boundary between one of the signals moving from sub to super-Gaussian.

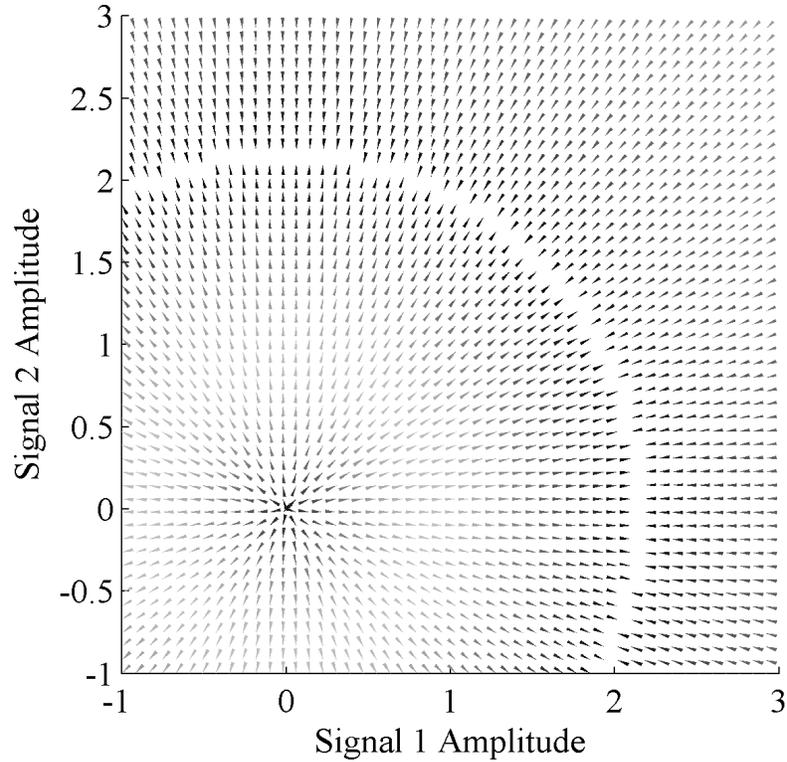


Figure 2.10: Gaussian vs. Gaussian system dynamics. There is no stable solution for two competing Gaussian signals. Rather, there is a metastable ring where both signals coexist.

### 2.4.7 High gain system dynamics

There is the possibility of the system oscillating at different steady state values as well, as can be seen in Figure 2.15. This figure is a reproduction of the open-loop gain curve found in Figure 2.6 for a harmonic signal. If the gain of the feedback loop is sufficiently high, it can be seen that the gain curve can be made to cross the loss curve multiple times. This is also demonstrated

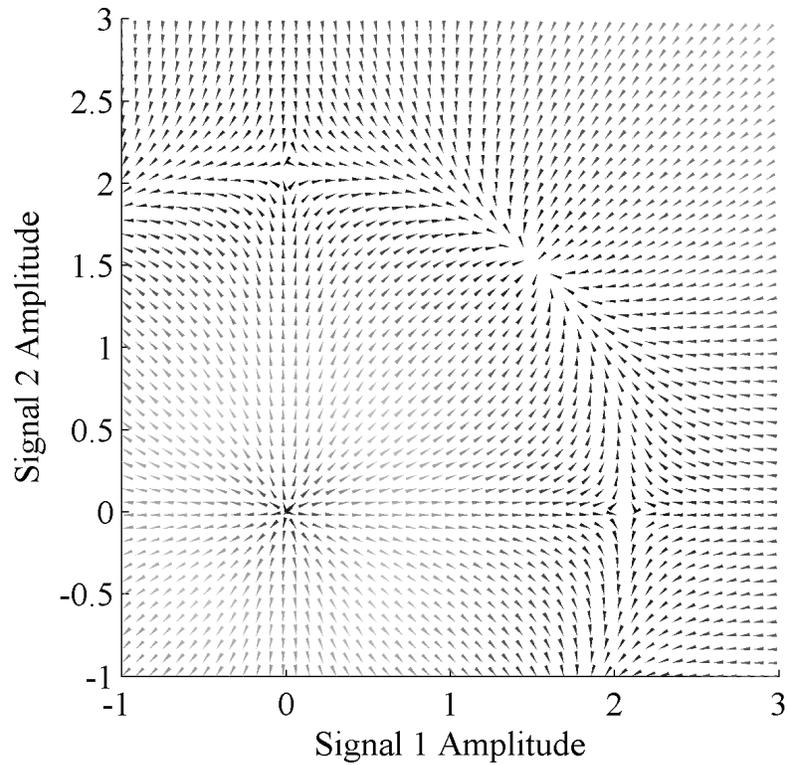


Figure 2.11: Laplace vs. Laplace system dynamics. The stable solution in the two Laplace signals corresponds to an equal amplitude of both signals coexisting in the feedback loop.

by the state space diagram in Figure 2.16, which allows the signals to have higher voltages in the feedback loop. This curve shows multiple stable steady states with both signals oscillating in the loop at higher voltages.

These steady state solutions are not desired since the signals are not separated, which can be seen from the stable solutions that do not lie on either signal axis. To prevent the system from oscillating at these points,

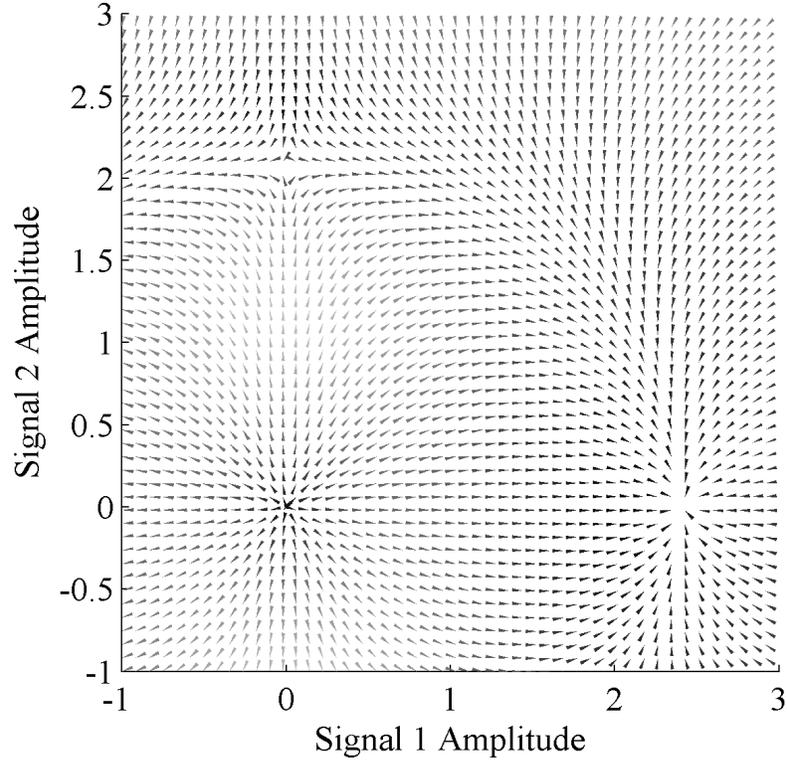


Figure 2.12: Harmonic vs. Laplace system dynamics. The only stable solution in this case is the harmonic signal existing alone.

it is necessary to stay out of the high-gain regime. We can estimate what the boundary between these two regimes is by picking the initial gain ( $G$  in Figure 2.15) so that the open-loop gain will only cross the loss curve in one place. As can be seen from Figure 2.16, the first unwanted steady state occurs at  $s_1 = 4.8V_\pi$ , and  $s_2 = 2.4V_\pi$ .

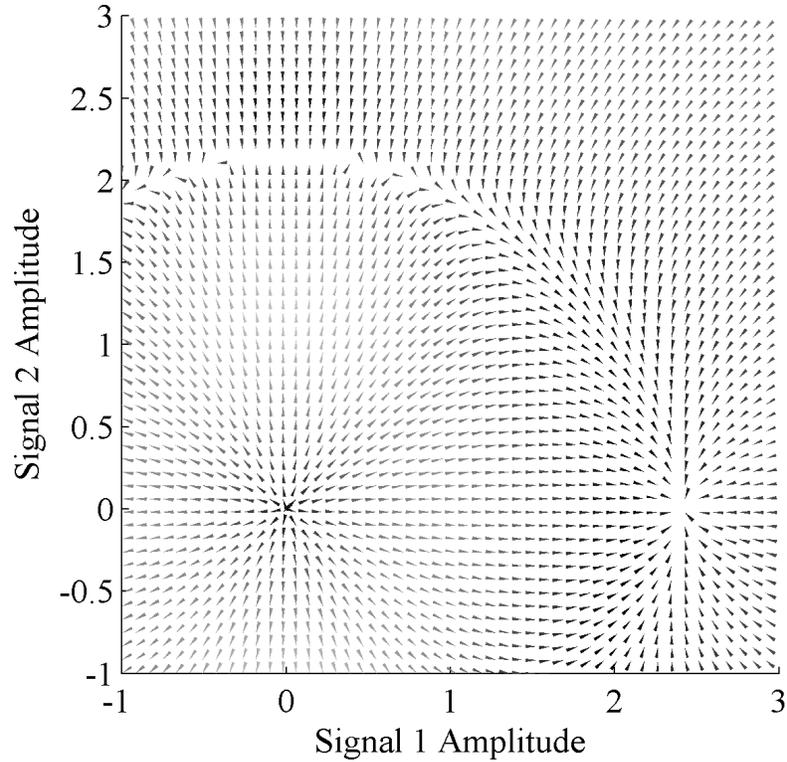


Figure 2.13: Harmonic vs. Gaussian system dynamics. The only stable solution in this case is the harmonic signal existing alone.

## 2.5 Summary

In summary, the analysis presented in this chapter shows that the proposed opto-electronic system is capable of selecting one signal from a mixture of input signals. This signal will oscillate in the feedback loop without crosstalk from the other signals. At least one of the input signals must be sub-Gaussian for this desired behavior.

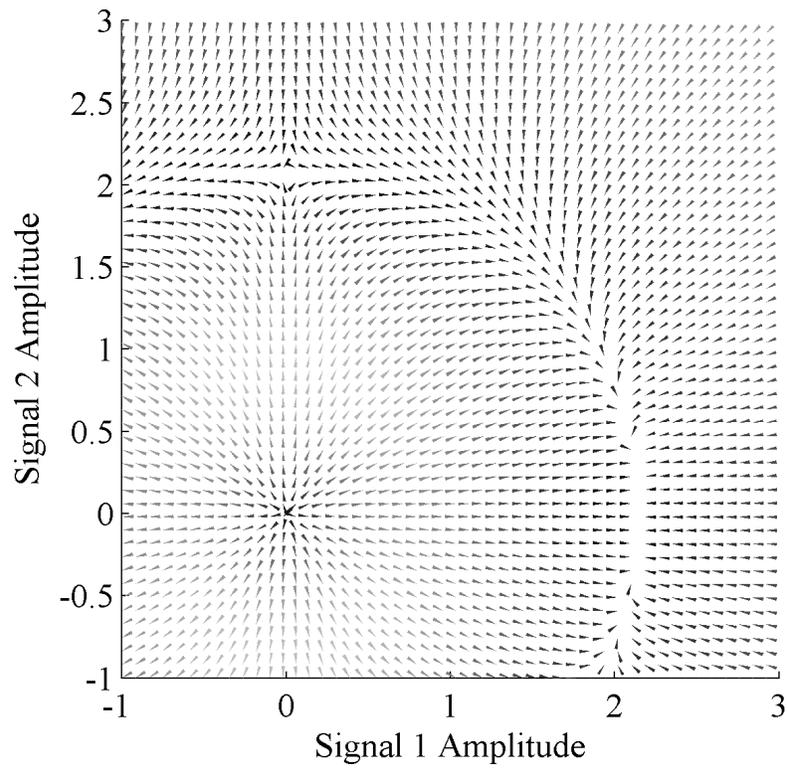


Figure 2.14: Gaussian vs. Laplace system dynamics. The Gaussian signal dominates in the case of a Gaussian vs. super-Gaussian competition.

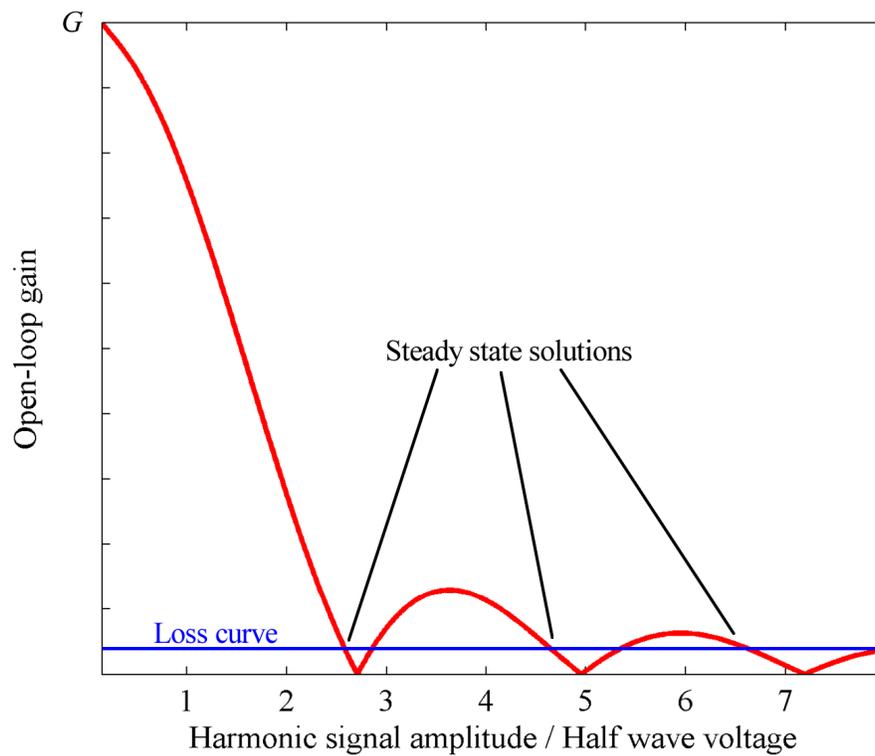


Figure 2.15: Harmonic signal open-loop gain curve. If the gain,  $G$ , is high enough, there are multiple steady state solutions for the amplitude of the oscillating signal.

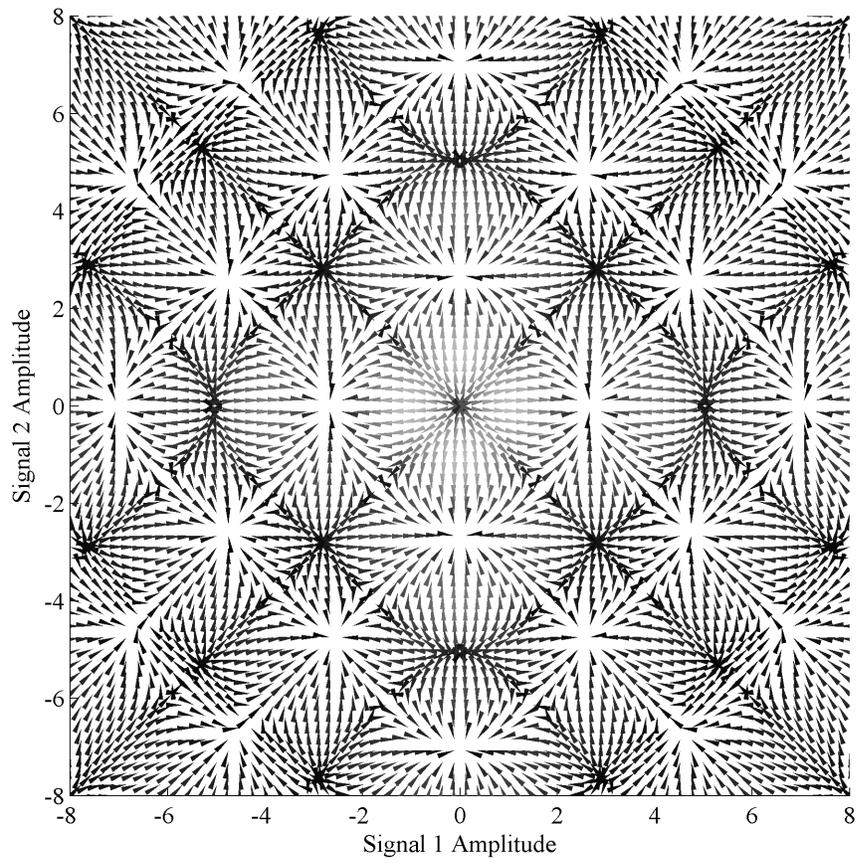


Figure 2.16: High gain harmonic vs. harmonic system dynamics.

# Chapter 3

## 100 kHz bandwidth opto-electronic ICA implementation

### 3.1 Overview

The focus of this chapter is on the design and implementation of a 100 kHz opto-electronic ICA system. While the ultimate goal of our opto-electronic processor is to perform ICA on RF signals, it was more appropriate in terms of cost to design the first system for kilohertz bandwidth signals. These signals can be generated and manipulated by computer, which provides fine control over the signal probability density functions. All components of the feedback loop in this design (electronic amplifiers, electro-optic modulators, and photodetector) must have a signal bandwidth of 100 kHz or greater.

The proof of concept system described is designed to separate one signal

from two mixtures of two signals and justify the mathematical framework in chapter 2. It has two input channels that each carry one of the input mixtures that are orthogonalized and normalized to have the same power. The system presented here is therefore only able to carry out the fourth order decorrelation portion of ICA (see Figure 1.2) The remainder of this chapter is dedicated to describing the design of a real electro-optical implementation of a narrow band analog ICA processor. Chapter 4 describes experimental characterization of the system.

## 3.2 System description

Figure 3.1 shows a block diagram of the ICA system. There are four primary subsystems to be considered: the signal generation and input circuit, the opto-electronic feedback loop, the homodyne detection circuit, and the output and performance evaluation circuitry. In order to properly evaluate the performance of the system, there are four signals output by the computer: the signal mixtures (inputs to the system), and the original signals themselves, for comparison at the output. There are two identical input channels that modulate the electrical signal onto an optical carrier and preprocess it for input into the feedback loop. The feedback loop is the heart of the analog ICA processing system. It is this portion of the system that was analyzed in chapter 2. The feedback loop signal is phase modulated, and is converted to a detectable amplitude modulation by combination with

a homodyne beam. The homodyne beam phase must be kept at a 90 degree offset from the original carrier phase. To compensate for environmental drift, the phase of the homodyne beam is controlled using a piezoelectric movable mirror. The single output of the system is compared with both original signals in order to evaluate the amount of crosstalk in the feedback loop. Each of these subsystems will be described in the next sections.

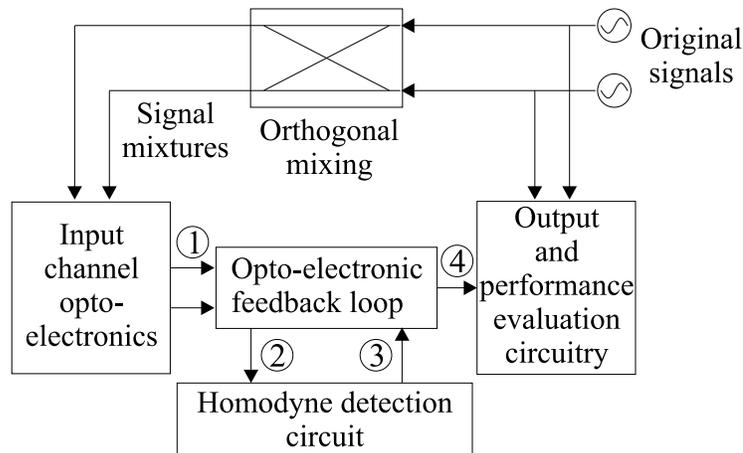


Figure 3.1: Generalized block diagram of the opto-electronic ICA system. The input mixtures and the original signals are created by a computer (not shown). The four principle subsystems are the input channel opto-electronics, the opto-electronic feedback loop, the homodyne detection circuit, and the output circuit.

### 3.3 Input channels

A schematic diagram of one of the two identical input channels is shown in Figure 3.2. The computer-generated signal mixtures are electronically

amplified in order to drive an electro-optic phase modulator of a 532 nm laser beam common to both channels. As a result, there are five optical signals in the system: the optical carrier, and the upper and lower sidebands for both of the input signals. The carrier is unwanted in the processor as it introduces correlation between the signals. To remove this correlation, an optical carrier suppression circuit has been added to each of the input channels. Two-beam coupling in a photorefractive crystal has been shown to be a very effective way to remove the carrier from the two input channels, as detailed in [40].

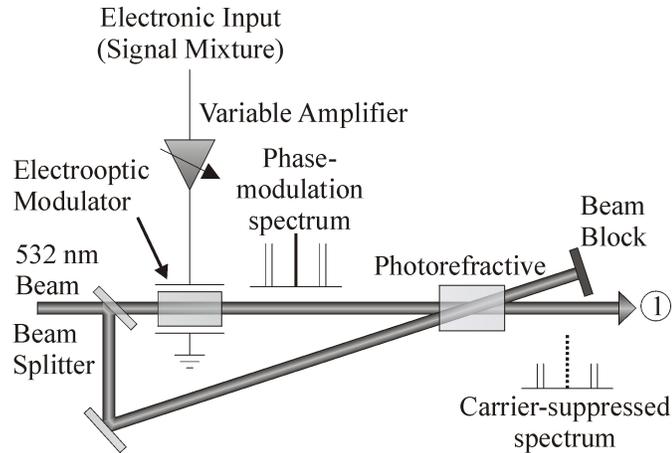


Figure 3.2: One input channel. The 100 kHz amplifier has a nominal gain of 200 and a  $\pm 200$  V output, which provides the high voltage required to drive the electro-optic phase modulator. The double-sideband modulated laser beam (Coherent Verdi, 532 nm) is carrier suppressed by a photorefractive carrier suppression circuit to eliminate unwanted false correlation between the two input channels. While the schematic shown here uses two sinusoidal signals for ease of representation, any signal type can be used in the actual implementation.

### 3.3.1 Signal generation

The signals are digitally created in Matlab, and output as a .wav file to the sound card, a four channel Soundblaster Audigy. It has 44.1 kHz digital-to-analog conversion, with 24 bits of resolution. The signal mixtures are output to the left and right speaker outputs. When played as an audio sound file, the signal amplitude is approximately 8 volts zero-to-peak at maximum volume.

The signals,  $s_1$  and  $s_2$ , are generated in Matlab as vectors with individual element amplitudes between -1 and +1 and length  $44100 \cdot T$ , where  $T$  is the signal length in seconds. These signals are normalized to have the same variance, or:

$$\langle s_1^2 \rangle = \langle s_2^2 \rangle \quad (3.1)$$

This condition ensures that both signals have the same power.

Once two signals in question have been created, they are mixed together using an orthogonal mixing matrix:

$$\mathbf{M} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \quad (3.2)$$

The two signal mixtures,  $\tilde{s}_1$  and  $\tilde{s}_2$ , are given by:

$$\begin{bmatrix} \tilde{s}_1(t) \\ \tilde{s}_2(t) \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix} \quad (3.3)$$

The matrix  $\mathbf{M}$  is a simple rotation matrix that has only one coefficient: the amount of rotation,  $\theta$ . As seen in the signal space diagrams in Figure 3.3, if  $\theta$  is 0 degrees, or any integer multiple of 90 degrees, then each input mixture

carries one signal without contamination from the other signal. If  $\theta$  is 45 degrees, or any odd integer multiple, then there is an equal amount of each signal present in both mixtures. In other words, one mixture is given by  $\tilde{s}_1 = \frac{1}{\sqrt{2}}(s_1 + s_2)$ , and the second mixture is given by  $\tilde{s}_2 = \frac{1}{\sqrt{2}}(-s_1 + s_2)$ . There is the same power in each signal in both the 0 degree mixture and in the 45 degree mixture.

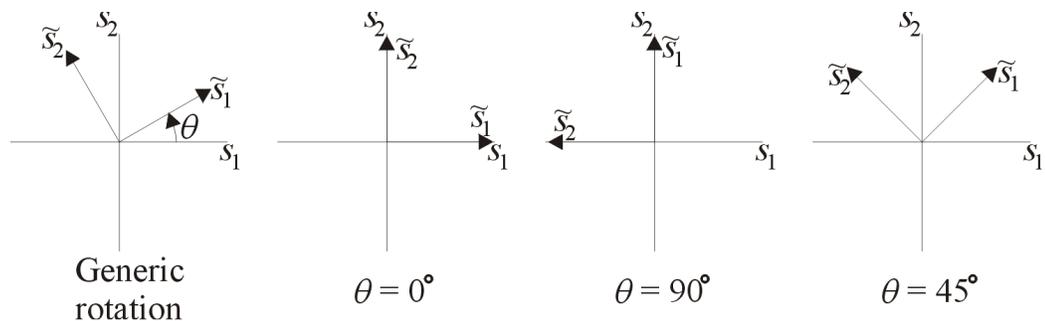


Figure 3.3: Signal space diagrams of whitened signals. In signal space, orthogonal signals are defined by the parameter  $\theta$ . If  $\theta$  is an integer multiple of 90 degrees, then there is only one signal present in each mixture. If  $\theta$  is an odd integer multiple of 45 degrees, then there is an even amount of each signal present in the mixtures.

The original signals,  $s_1$  and  $s_2$ , generated in Matlab can be either sinusoidal signals with any frequency up to 22.05 kHz (Nyquist sampling criterion) or completely random signals with arbitrary probability density functions. The random number generator in Matlab generates a number,  $R$ , between 0 and 1 which must be mapped to an arbitrary probability density,  $p_{arbitrary}(x)$ . The new random number,  $r$ , with the desired probability density

is found by solving the following equation for  $r$ :

$$R = \int_{-\infty}^r p_{arbitrary}(x) dx \quad (3.4)$$

The solution to Equation 3.4 is found numerically in Matlab. The accuracy of the probability density resulting from this mapping is verified by generating a histogram from several thousand random numbers.

### 3.3.2 Input electronics

Once the signals have been generated, they are input into the system through a series of electronic amplifiers represented by the variable amplifier in Figure 3.2. The first amplifier is an Ithaco 1201 preamp. These devices have a variable gain from 1 to 1000. They have an output voltage of 20 volts peak-to-peak. In addition, they have a high pass and low pass filter on the output signal. These filters are set to DC and max ( 500 kHz) respectively in order to maintain the original signal structure.

The 20 volt output of the preamplifiers is not high enough to drive the electro-optic modulators (made by Gsanger) that have a half wave voltage of 900 volts. A second stage of amplification is therefore used. There have been two different high-voltage amplifiers used to drive the electro-optic modulator. Both of these amplifiers are described here since they have both been used during data collection. The first is the Trek Model 601b-3 amplifier with an output voltage of 1000 volts peak-to-peak and a gain of 100. This high-voltage amplifier has a bandwidth of 8 kHz. While the amplified signal

easily has a high enough voltage to drive the phase modulators, the bandwidth is still smaller than the desired 100 kHz, prompting the acquisition of the second FLC 400D amplifier. This amplifier has a gain of 20 and an output of 400 volts peak-to-peak with a signal bandwidth of 100 kHz.

It is crucial that the two input channels be very well matched in both gain and phase. If there is either a gain or phase mismatch, the input mixtures will not be orthogonal, and will violate the basic assumption made in the analysis in chapter 2. Mismatches are manifested as reduced signal separation in the output of the system. The Ithaco variable gain preamplifiers adjust the input channels to compensate for gain mismatch in the high-voltage amplifiers.

The phase mismatch was measured to be 14 degrees between the two Trek amplifiers at 1.5 kHz. The first method to compensate for this phase mismatch is to digitally delay one input mixture with respect to the other in Matlab. The smallest increment that one signal can be shifted by, however, is determined by the sampling frequency of the computer sound card. This smallest shift is  $\frac{1}{44.1 \text{ kHz}}$ , or 22.6  $\mu\text{s}$  corresponding to a 12 degree phase shift at 1.5 kHz. This means that the smallest possible step size is the same order of magnitude as the required phase compensation.

The second method for phase mismatch compensation is to output one of the signal mixtures twice from the computer. One of the outputs will have a 0 degree phase shift, and the other will have a 90 degree phase shift. By changing the amplitude of these two signals and summing them together, it

is possible to create an arbitrary phase shift, as seen from:

$$\cos(\omega t + \phi) = \cos \phi \cos \omega t + \sin \phi \cos\left(\omega t - \frac{\pi}{2}\right) \quad (3.5)$$

Using this method, the required phase shift is obtained using amplitude variations from the computer, given by  $\cos \phi$  and  $\sin \phi$ , which is very accurate due to the 24-bit resolution of the sound card.

The drawbacks to this method of phase shifting are twofold. First, each time a different phase shift is required, we must generate a completely new set of signals. While the phase mismatch between the two amplifiers tends to remain relatively stable, it varies while the amplifiers are warming up. Second, this method requires the use of three of the four output channels of the sound card: one channel for  $\tilde{s}_1$ , and two channels for the in-phase and quadrature components of  $\tilde{s}_2$ . This method cannot be used when all four channels are required for evaluation of the system performance, described in section 3.6.

The third method of compensating the phase mismatch between the two high voltage amplifiers is to include a filter in one of the channels [46]. A first-order, low-pass filter was chosen for this task. The filter has the affect of decreasing the signal amplitude at higher frequencies and also shifting the phase by some amount at higher frequencies. This is illustrated in Figure 3.4. The low-pass filter, shown in Figure 3.5, uses a potentiometer to accurately control the exact phase shift at 1.5 kHz. To compensate for the decrease in signal amplitude at 1.5 kHz, the input preamp gain is increased. This gives

a precise, analog control over the phase mismatch between the two channels.

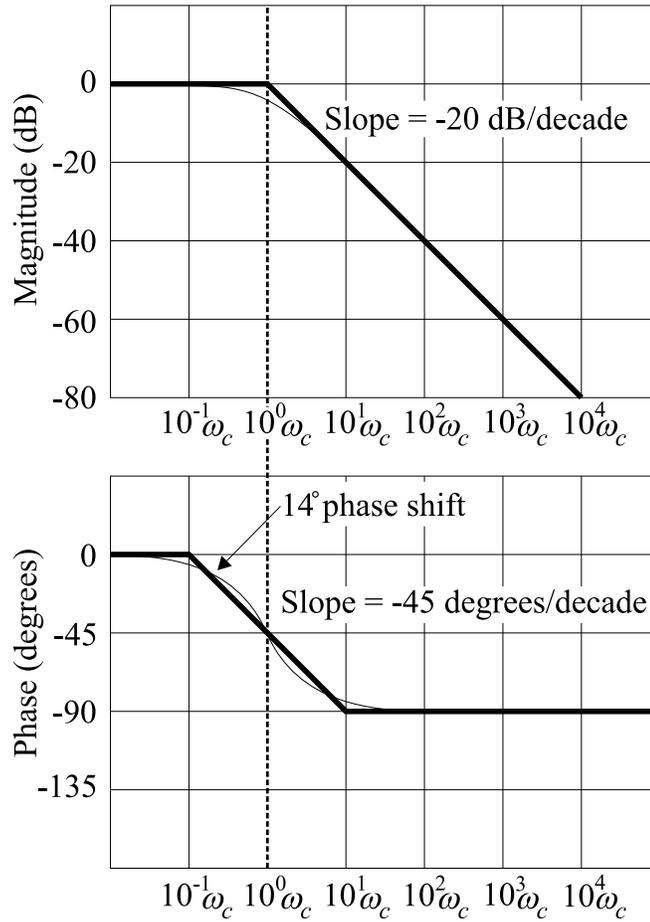


Figure 3.4: Low pass filter magnitude and phase characteristics. A low pass filter both attenuates signals at high frequencies and provides a phase shift. At a phase shift of approximately -14 degrees, there is very little amplitude attenuation.

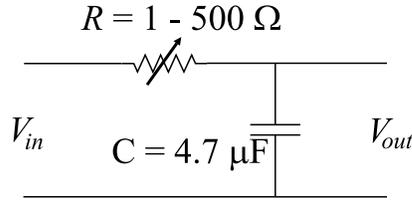


Figure 3.5: Low pass filter used for input channel phase correction.

### 3.3.3 Input optics

As seen in Figure 3.2, the electronic signal mixtures phase-modulate an optical carrier. The electro-optic modulator is a Gsanger LM24P lumped element phase modulator with Brewster windows. Using the travel-time limitation [47], the electro-optic modulators have a bandwidth of 550 MHz (assuming a 6 cm  $LiNbO_3$  crystal length). This bandwidth is substantially higher than the bandwidth of the driving amplifiers, and therefore not a limitation for the 100 kHz system. The measured capacitance of the modulators is approximately 70 pF, which is a very large impedance (approximately 20 k $\Omega$ ) at 100 kHz. This capacitance is low enough to have negligible effect on 100 kHz signals and does not limit the bandwidth of the system.

It is important that the input phase modulation is linear as this is the assumption used in Equation 2.19. The FLC amplifiers drive the electro-optic modulators at 400 volts peak-to-peak. Therefore, at maximum drive voltage,  $J_1\left(\frac{400\pi}{900}\right)$ , or 54% of the power is in the first signal harmonic, while only  $J_3\left(\frac{400\pi}{900}\right)$ , or 5% of the power is in the third harmonic.

The carrier suppression circuit is also included in the optical portion of

the input channels. This circuit is discussed in detail in [40] and [48], and only a brief overview will be presented here. The carrier suppression scheme is shown in Figure 3.6. In this setup, two beams are mixed in a photorefractive crystal. One of the two beams is phase modulated, while the other is the optical carrier. The carrier signal in each of the two beams will write a grating within the crystal. If the intensities of the two beams is set properly (so that the total intensity in both paths is the same) all of the carrier in the phase modulated beam will be transferred to the other beam, resulting in -70 dB of carrier suppression [40] using this method. 3 dB of the sideband power will also be transferred to the other beam and is lost. This beam, which contains the carrier signal from both inputs to the photorefractive, is not needed in the system, and blocked. The carrier suppression in the ICA system has been measured to be over -30 dB suppression.

The design of the optical portion of the input channels is shown in Figure 3.7. The initial half-wave plate in Figure 3.7 is used to control the relative power between the two arms of the carrier suppression circuit. Since balanced power in the two arms is vital to the carrier suppression, precise analog control over this power ratio is required. The photorefractive crystal is specifically cut to be used with horizontally polarized light. Therefore, in order to minimize the number of half-wave plates, the electro-optic modulator in each input channel is rotated by 90 degrees so that it modulates horizontally polarized light. The Brewster windows are designed to reflect all unwanted vertical polarization in this configuration. The lenses are used

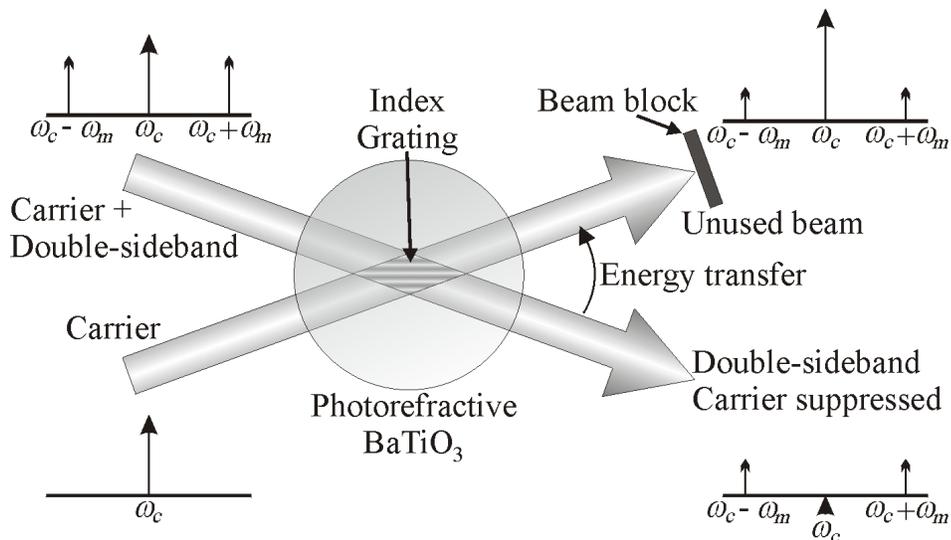


Figure 3.6: Carrier suppression schematic. Carrier suppression is accomplished through two-beam coupling in a photorefractive crystal of  $BaTiO_3$ . One of the input beams contains both the carrier and the modulation sidebands, while the other contains only the carrier. A photorefractive grating is written between the two carrier signals, and the carrier energy is transferred from the carrier suppressed beam to the other beam.

to focus the both the signal beam and the carrier beam to approximately a 0.5 mm waist located in the photorefractive crystal.

After all of the input channel preprocessing, the signals input to the processor feedback loop are double-sideband, carrier suppressed signals.

### 3.4 ICA feedback loop

The central processing portion of the opto-electronic ICA system is the feedback loop shown in Figure 3.8 (a reproduction of Figure 2.2). The primary

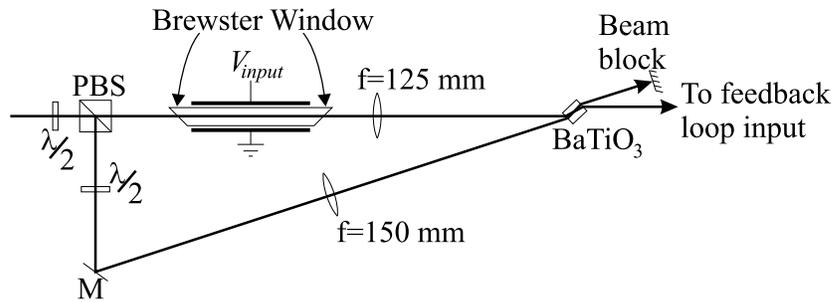


Figure 3.7: Input channel optical setup.

components of the feedback loop have been described in section 2.2. The input signals shown at the bottom of the figure are modulated with the input mixtures, as described in section 3.3. The next several sections describe the components of the feedback loop.

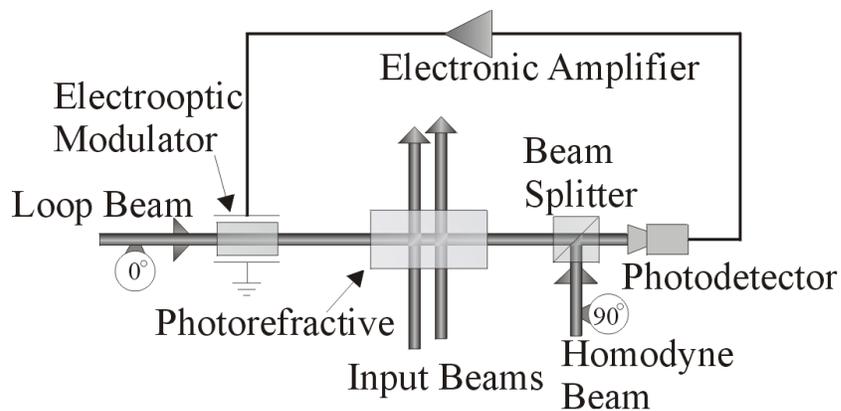


Figure 3.8: Opto-electronic feedback loop.

### 3.4.1 Electro-optic modulator

The electro-optic modulator in the feedback loop is a Quantum Technology, Inc. bulk electro-optic modulator with no Brewster windows. Its length of approximately 10 cm gives a travel-time bandwidth limitation of 330 MHz and a measured capacitance of 110 pF.

The modulator (Figure 3.9) does not have the bottom electrode connected directly to ground. It uses four crystals oriented at 90 degrees with respect to each other to compensate for birefringence due to temperature variation. The double electrode structure means that the modulator is used differently for the two amplifiers that have been used in the system, the Trek 601C and the FLC A400DI. When used with the Trek, the second electrode of the modulator is shorted directly to ground. The FLC amplifier has two outputs that are 180 degrees out of phase with each other, but each with 20 times amplification. The voltage across the electro-optic crystal can therefore be doubled to 800 volts peak-to-peak by plugging each output into one of the electrodes.

The modulation setup used in the opto-electronic feedback loop is shown in Figure 3.10. The half wave voltage of this modulator was measured in this setup by placing a DC voltage on the electrode with the half-wave and quarter-wave plate removed. As the voltage is increased, the modulator changes the output polarization of the laser beam. The polarization modulation is changed into an amplitude modulation with a horizontally-oriented

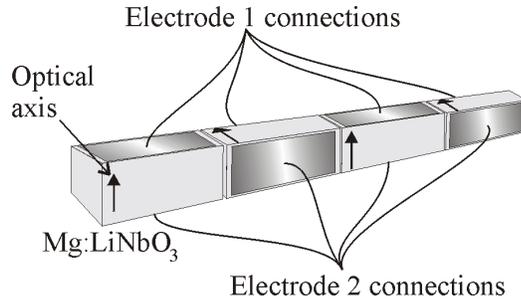


Figure 3.9: Feedback loop electro-optic modulator structure. The four electro-optic modulator crystals and electrodes are rotated by 90 degrees with respect to each other to compensate for temperature induced birefringence.

polarizer. The output intensity versus input voltage curve is shown in Figure 3.11. The amount of voltage that is required to get a  $\pi$  radian phase shift is 290 V peak-to-peak. As described in chapter 2, the modulator must be driven at  $\pm 1.22V_\pi$  in order to achieve the modulation depth necessary for the steady state, winner-takes-all ICA solution. This means that the output amplifiers must have an output voltage of at least 700 volts peak-to-peak.

As is the case with the input channels, the photorefractive crystal in the feedback loop is cut to be used with horizontally polarized light. In addition, the modulated output signal from the electro-optic modulator should have the same structure as the input beams in order to be able to write gratings with the input beams in the photorefractive crystal. In other words, the output of the modulator should be horizontally polarized, double-sideband, carrier-suppressed signals. This requirement dictates the modulator setup shown in Figure 3.10.

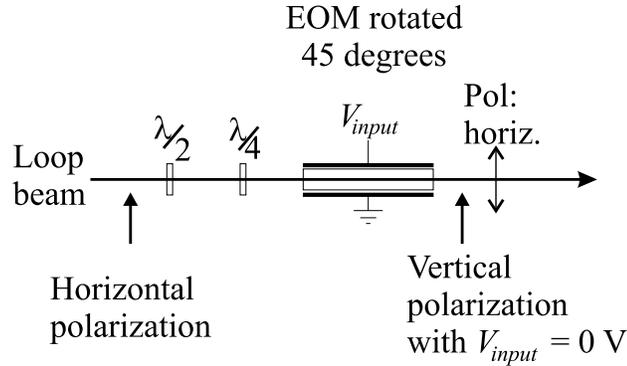


Figure 3.10: Feedback loop electro-optic modulator and wave plate setup. A half-wave plate and a quarter-wave plate pair are used to compensate for the natural birefringence of the modulator. The feedback loop electro-optic modulator is rotated by 45 degrees to phase modulate only one polarization component of the input beam. The output intensity of this circuit is minimized when there is a 0 volt DC signal present on the modulator.

The electro-optical phase modulator is designed to modulate only one of the two polarizations traveling through the modulator. For example, if +45 degree linearly polarized light travels through the modulator, only the vertical component will be phase modulated. This will change the output polarization of the modulator as shown in Figure 3.12, from +45 degree linear to right circular to -45 degree linear to left circular as the phase shift goes from 0 to  $2\pi$  radians. If a polarizer is placed in the polarization modulated beam, it will be converted to an intensity modulation. If the polarizer is placed at -45 degrees, then there will be zero intensity output when there is zero volts across the crystal, and the +45 degree linearly polarized beam is unmodulated. There will be 50% intensity transmitted at both right and

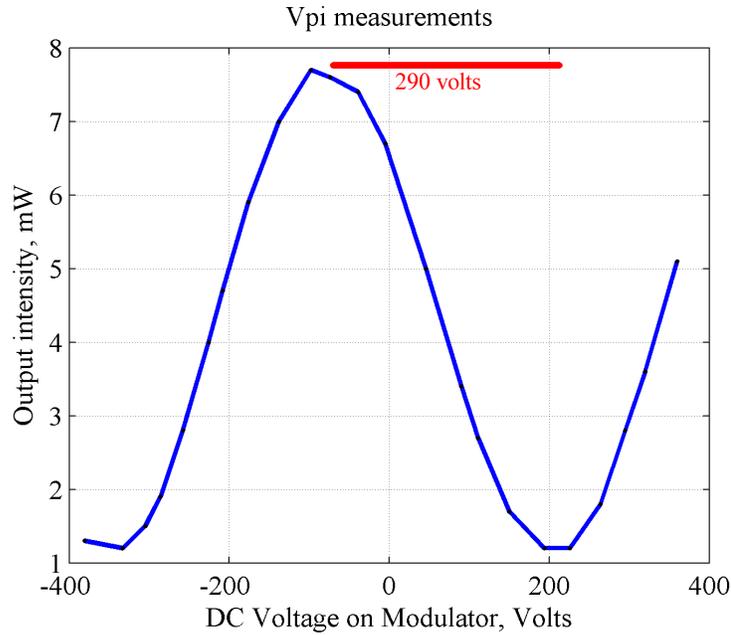


Figure 3.11: Feedback loop electro-optic modulator half-wave voltage measurement. The half-wave voltage of the loop modulator is made without the half-wave and quarter-wave plates shown in Figure 3.10. A DC voltage is applied to the EOM electrode, and the resulting intensity output from the polarizer is measured. This plot shows a half-wave voltage of 290 V.

left circular polarizations, and 100% transmission at +45 degree linear polarization. This intensity modulation can be mapped as shown in the plot in Figure 3.13a. As the input voltage (on the horizontal axis) increases, the polarizer intensity output follows a sinusoidal profile.

The position on the sinusoidal intensity output modulation is imposed by the position of the output polarizer. If the polarizer is placed at a -45 degree rotation, then zero volts (no phase shift) will result in 0% intensity output. A sinusoidal change in applied voltage will produce an intensity output as

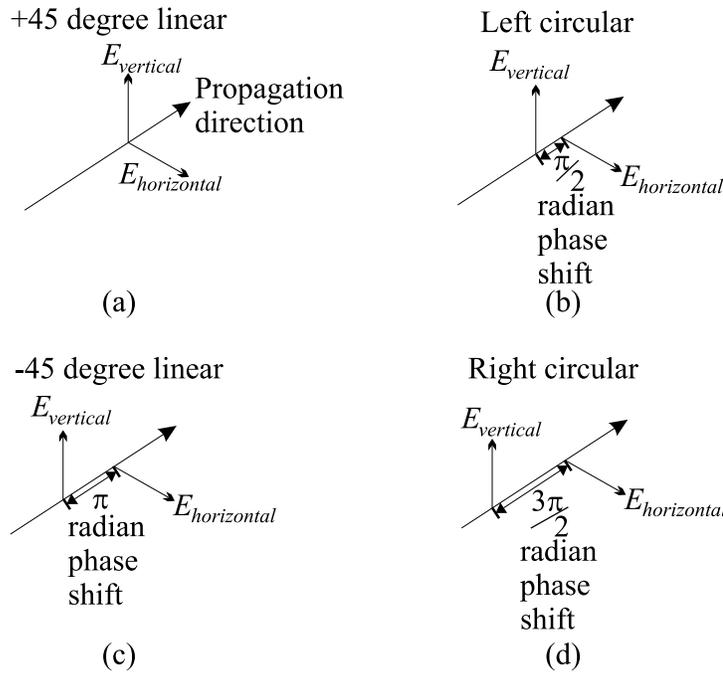


Figure 3.12: Electro-optic modulator output polarizations. By phase shifting only the vertical polarization component of an input electric field, the electro-optic modulator polarization can be changed to (a) +45 degree linear, (b) left hand circular, (c) -45 degree linear, and (d) right hand circular.

shown in Figure 3.13b. When biased at the point of 0% intensity with zero volts, the output intensity is a double-sideband, carrier-suppressed signal as desired.

The second stipulation, that the output must be horizontally polarized, is easily achieved simply by rotating the whole setup by 45 degrees so that the output polarizer is horizontally polarized.

In a real electro-optic modulator, the crystal usually has some birefringence, which actually rotates the output polarization to some unknown ellip-

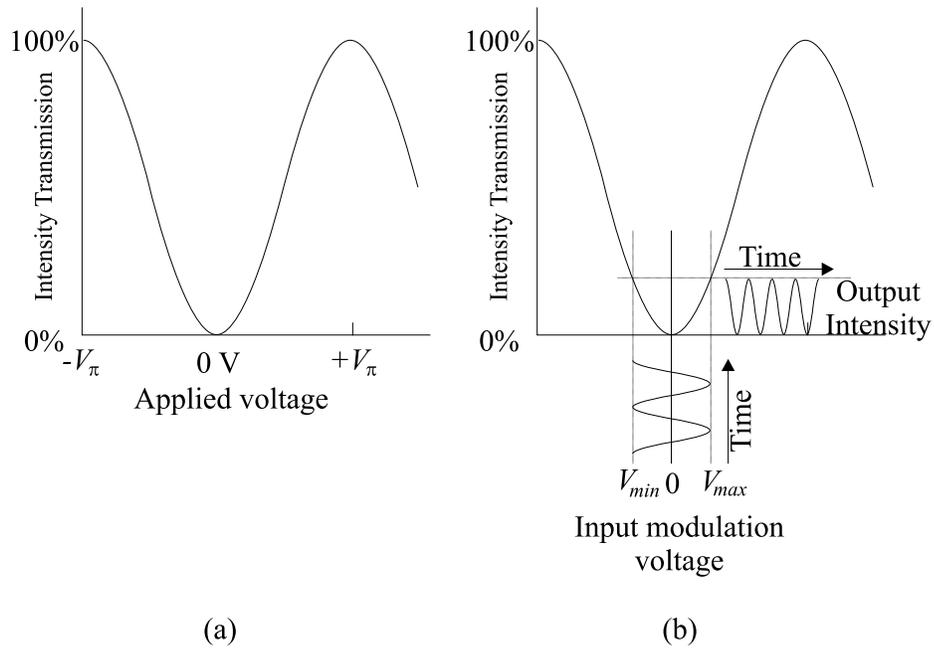


Figure 3.13: Loop modulator output intensity. (a) The output intensity of the loop modulator and polarizer follows a sinusoidal profile. (b) If the modulator is biased so that 0 applied volts results in 0% intensity output, a sinusoidal input will result in a double-sideband, carrier-suppressed output intensity signal.

tical polarization. The output intensity transmitted by the polarizer is therefore not exactly 0%, as required by the double-sideband, carrier-suppressed condition. To achieve this, the beam polarization must be vertical after the birefringent effects of the modulator. To compensate for the unwanted modulator birefringence, a half-wave plate and a quarter-wave plate are placed in front of the electro-optic modulator, as shown in Figure 3.10. Using this combination, it is possible to change any unknown elliptical polarization to a known linear polarization (vertical in this case). This wave plate pair can

be placed either in front of or behind the modulator.

The alignment procedure for the electro-optic setup in Figure 3.10 is straight forward. First, the output polarizer is set horizontal and the modulator is rotated by 45 degrees. The wave plate pair is rotated until there is zero intensity transmission from the polarizer when there is no signal on the modulator. This procedure insures that the modulator is biased at the bottom of the sinusoidal curve, as depicted in Figure 3.13b so that it produces the proper carrier-suppressed, double-sideband signal, and still has the required horizontal polarization.

### 3.4.2 Photorefractive medium and loop beam power

It is the combination of the electro-optical modulator and the photorefractive crystal that makes the dynamics discussed in chapter 2 possible. As discussed in chapter 2, the nonlinearity of the feedback loop is provided by the electro-optical modulator while the correlation is produced by the photorefractive crystal. The system described here uses barium titanate ( $\text{BaTiO}_3$ ) as the photorefractive material. It is 0 degree cut with respect to the optical axis. The maximum gain occurs when the input beams are at an angle of 18 degrees in free space.

It has already been stated in section 2.3.1 that the photorefractive crystal provides coupling between two beams and allows energy transfer between the coherent portions of the two beams. The increased signal due to the power transfer from one beam (the pump beam) to the other (the signal

beam) is defined to be *photorefractive* gain. For BaTiO<sub>3</sub>, it is not uncommon to see gain numbers on the order of 10,000, or 40 dB (power gain). The only possible way to see this much gain, however, is if the pump beam has enough power to give to the signal beam (see Figure 3.14). In order to maximize the photorefractive gain, therefore, the ratio of the powers between the pump beam and the signal beam must be greater than the gain of the photorefractive. If the ratio of powers in the two beams is less than the possible photorefractive gain, then the pump beam will be depleted, causing a nonlinear saturation in the amplification of the signal beam.

The analysis of opto-electronic ICA in chapter 2 assumes that the loop beam will be amplified by the correlated portions of the input beam and the amplification is linear. In other words, the photorefractive gain must not be saturated because this introduces a second nonlinearity to the feedback loop. The unsaturated regime can be reached in two different ways. First, a photorefractive crystal with low gain is chosen. A low gain crystal is relatively thin. The crystal that is used in the low-bandwidth system is approximately 3 mm wide. Second, the power ratio between the loop beam and the input beams should be much higher than the gain of the crystal. This is more easily achieved with a low gain crystal. These two approaches to having a linear gain due to the photorefractive crystal have been followed. The loop beam power is set to approximately 130 nW, and the input beam powers are each set to approximately 3 mW, a ratio of 23,000. This means that if the gain is significantly less than 23,000, the crystal should be in the linear

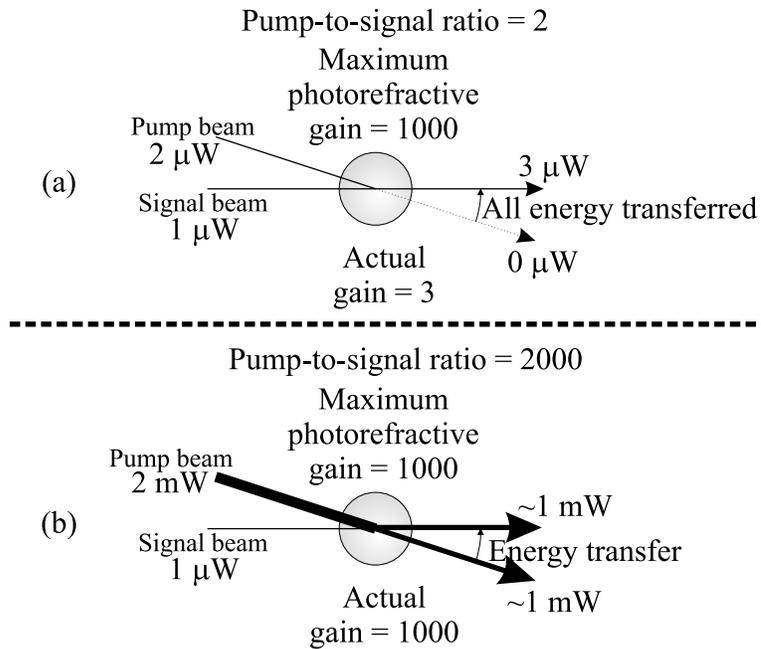


Figure 3.14: Photorefractive crystal gain saturation. (a) The pump-to-signal ratio is smaller than the photorefractive gain, limiting the actual observed gain and saturating the crystal. (b) The pump-to-signal ratio is larger than the photorefractive gain, allowing the maximum increase in the signal beam.

regime.

To measure the gain of the photorefractive crystal, the ratio of the input beam power versus the output beam power can be measured. The gain ratio will stay constant as the input power is varied if the gain is in the linear regime. The gain measurement of the photorefractive crystal used in the system is shown in Figure 3.15. The constant slope of the curve shows that the gain is indeed in the linear regime, and that the gain is approximately 330. This gain value is, as required, less than the beam power ratio of 23,000.

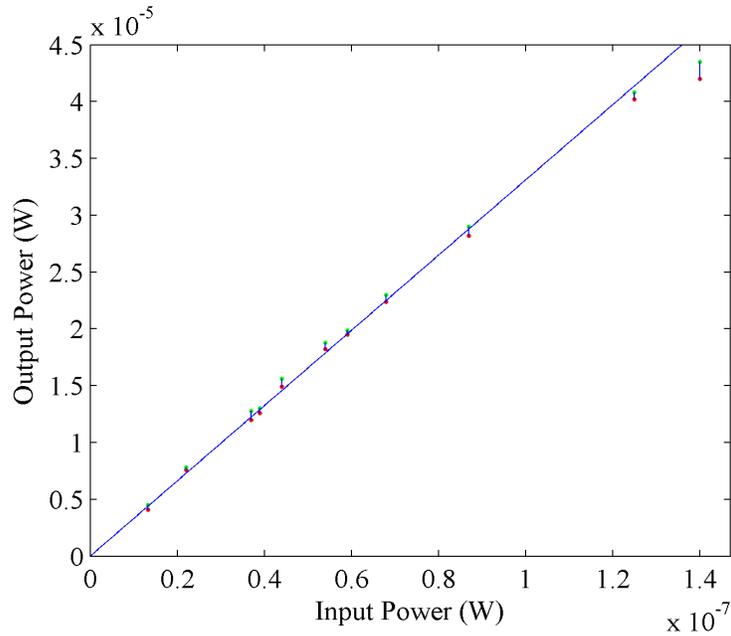


Figure 3.15: Measured photorefractive gain. This plot shows a gain of 330. The gain is linear if the input beam power is kept below 130 nW.

### 3.4.3 Photodetector

The signal in the feedback loop is converted from an optical signal to an electronic one via the photodetector. There have been two primary photodetectors that have been used in the low-bandwidth ICA setup. The first of these is a single photodiode FFD-100 EG&G detector with a bandwidth of approximately 3 MHz. The second photodetector is a differencing photodetector that uses two PIN-5DI photodiodes and has a bandwidth of approximately 100 kHz. The signal from these photodiodes is differenced. This scheme is used to remove the laser amplitude noise from the signal. One of

the diodes sees only the laser beam, along with its noise, while the other sees both the feedback loop signal and the laser noise. When the two are differenced, the laser amplitude noise cancels between the two diodes.

### **3.4.4 Electronic Amplifiers**

Once the loop signal has been detected, it must be electronically amplified before it is reimposed on the beam by the electro-optic modulator. The amplification is accomplished by a set of two amplifiers. The first amplifier is an Ithaco 1201 preamp, which is discussed in section 3.3.2. Two high-voltage amplifiers have been used as well, the Trek 603b-3 amplifier, also discussed in section 3.3.2, and the FLC 400DI. The FLC amplifier is very similar to the ones used on the input channels, except that there is an inverter placed between two of the original FLC 400D amplifiers. These amplifiers can therefore be used to double the amount of voltage (to 800 V peak-to-peak) if they are used as a floating voltage. They have the same bandwidth as the input FLC amplifiers as well (100 kHz).

## **3.5 Homodyne detection circuitry**

### **3.5.1 Homodyne theory**

In order to change the loop phase modulated signal into a detectable amplitude modulation, a homodyne detection scheme is used. Homodyne detection

is a reintroduction of the laser carrier (shifted by 90 degrees from the original carrier) to the loop beam. Mathematically, a phase modulated signal is represented as:

$$f_c(t) = A \cos[\omega_c t + \Delta\theta \cos(\omega_m t)] \quad (3.6)$$

where  $\omega_c$  is the carrier frequency,  $\omega_m$  is the modulation frequency,  $A$  is the carrier amplitude, and  $\Delta\theta$  is the modulation depth. The modulated signal in this case is  $\cos(\omega_m t)$ . Note that there is no amplitude modulation present in Equation 3.6.

The modulation depth,  $\Delta\theta$ , is assumed to be very small, or  $\Delta\theta \ll 1$ . With this simplification and some algebraic manipulation, Equation 3.6 can be written as:

$$f_c(t) = A \cos(\omega_c t) - A\Delta\theta \sin(\omega_c t) \cos(\omega_m t) \quad (3.7)$$

The carrier suppressed signal is represented as:

$$f_{c-C.S.}(t) = -A\Delta\theta \sin(\omega_c t) \cos(\omega_m t) \quad (3.8)$$

where the term containing  $\omega_c$  is removed.

The reintroduction of the carrier (the homodyne beam) at the photodiode is represented as:

$$f_{detected}(t) = [-A\Delta\theta \sin(\omega_c t) \cos(\omega_m t)] \cos(\omega_c t + \varphi) \quad (3.9)$$

where the reintroduced carrier has a phase of  $\varphi$  from the original carrier.

With a bit of manipulation, this can be rewritten as:

$$f_{detected}(t) = \frac{A\Delta\theta}{2} \cos(\omega_m t) \sin \varphi + \frac{A\Delta\theta}{2} \cos(\omega_m t) \sin(2\omega_c t - \varphi) \quad (3.10)$$

The second term in this detected signal has a frequency of  $2\omega_c$ , which is filtered out. The first term of Equation 3.10 is a reproduction of the original modulation signal,  $\cos(\omega_m t)$ , except that it is multiplied by some amplitude coefficients. In particular, the amplitude of the detected signal is maximized if the phase of the reintroduced carrier is 90 degrees from the original carrier. Notice that as the phase shift of the carrier goes from 90 degrees back to 0 degrees, the amplitude modulation disappears. The system will therefore be required to maintain a 90 degree phase shift on the homodyne beam throughout the system operation.

### 3.5.2 Homodyne circuit

The homodyne detection circuit that maintains the 90 degree phase shift on the carrier against environmental drift (including temperature changes, air currents, and even dust) is shown in Figure 3.16. The loop beam is split (at point 2 in Figure 3.16) just before modulation into the loop beam and the homodyne beam. The homodyne beam is recombined with the loop beam at point 3. The other components are necessary to actively maintain the 90 degree carrier shift.

The 90 degree phase shift is generated by the second mirror in the interferometer arm, which sits on a piezo-electric crystal. There is a position translation in the mirror if a voltage is placed across the piezo-electric crystal. The mirror provides a  $2\pi$  radian phase shift for 532 nm laser light for every 2.5 volt increase in the piezo voltage.

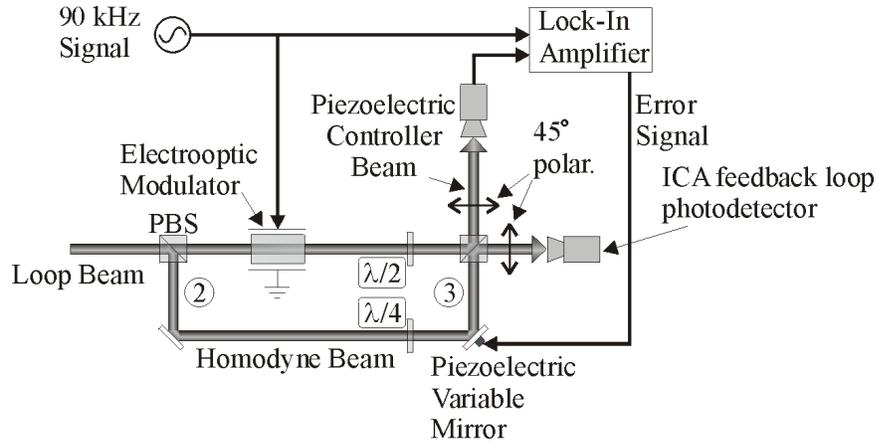


Figure 3.16: Homodyne detection circuitry. This circuit utilizes a feedback loop to provide the proper carrier phase to one of the interferometer arms. The phase is set with a piezoelectric mirror controlled by the lock-in amplifier’s error signal.

Figure 3.16 shows a control feedback loop to maintain the proper voltage on the piezo-controlled mirror to maintain the 90 degree phase shift. The objective is to maximize the signal reaching the loop photodetector. Likewise, this implies a minimization, or ideally the complete disappearance, of a signal that has a carrier 90 degrees off from the carrier in the loop beam, as described in Equation 3.10. It is actually the minimized signal that will be observed and tracked. The signal used for stabilizing the piezo-controlled mirror feedback loop is a 90 kHz signal modulated onto the loop beam. This ensures that there is always a known signal present on the loop beam to track.

Several polarization techniques are used to track the minimum of the 90 kHz signal that are highlighted in Figure 3.17. A quarter wave plate is placed

in the homodyne arm to provide a 90 degree shift between the two polarizations of the homodyne beam. One of these polarizations is be used in the ICA feedback loop detection, and one is be used in the piezo-controller feedback loop. The horizontally polarized homodyne beam mixes with a portion of the loop beam and is detected by photodetector 2 in Figure 3.17. The vertically polarized portion of the homodyne beam mixes with the loop beam and is detected at photodetector 1, the ICA feedback loop photodetector. Similarly, a half wave plate is placed in the loop beam arm of the interferometer to provide a 180 degree phase shift between the two polarizations of the signal beam. The half wave plate is used so that the vertical polarization of the loop signal will be mixed with the horizontal polarization of the homodyne beam.

The phase of the homodyne beam is adjusted by changing the voltage on the piezo-controlled mirror. If the voltage is set so that the 90 kHz signal detected on photodetector 2 disappears, then the signal on photodetector 1 is maximized.

The signal from photodetector 2 is fed into a lock-in amplifier that is phase-locked to the 90 kHz signal. This allows very precise detection of just the 90 kHz signal, and filters out all noise sources that may also be present on the laser beam. The output of the lock-in amplifier is governed by the equation:

$$V_o = V_i A \cos(\phi) \tag{3.11}$$

where  $V_i$  is the amplitude of the detected 90 kHz signal,  $A$  is a linear gain

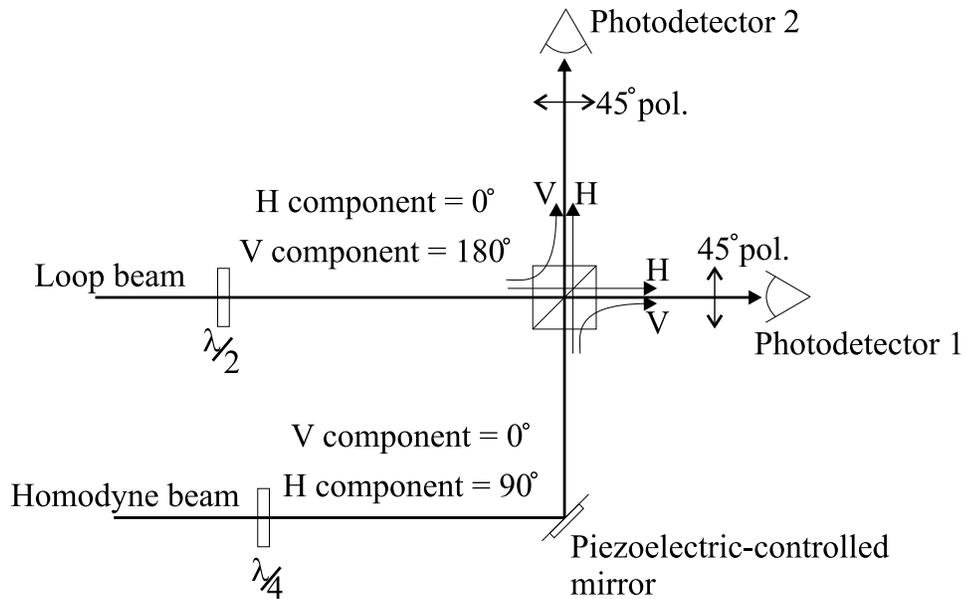


Figure 3.17: Homodyne detection circuitry beam phase shifts. The quarter-wave plate creates a 90 degree phase shift between the horizontal and vertical components of the homodyne arm. Similarly, a 180 degree phase shift is created by the half-wave plate in the loop arm beam. When the signal detected on photodetector 2 is minimized, the signal on photodetector 1 is maximized.

setting, and  $\phi$  is the phase difference between the reference 90 kHz signal and the detected 90 kHz signal. The phase difference between the two signals is 0 degrees, which maximizes the cosine term. The output of the lock-in amplifier is therefore directly proportional to the amplitude of the 90 kHz detected signal,  $V_i$ . If the amplitude drops to 0 volts, then the output of the lock-in is also 0 volts, at the exact voltage when the carrier phase has the correct shift. The lock-in output can therefore be used as an error signal to maintain the 90 degree phase shift on the carrier. This signal is used directly

as the voltage on the piezo-controlled mirror.

If the piezo-controller feedback loop is operating properly, a stable 90 kHz signal can be observed on the ICA feedback loop photodetector, and no signal is present on the piezo-controller loop photodetector.

### 3.6 Measurement circuitry

The opto-electronic ICA system output is the same as the signal in the feedback loop after it reaches a steady state. In order to evaluate the system performance, the output signal is compared to the original unmixed input signals. The unmixed signals are output on the third and fourth channels of the computer sound card. These original, unmixed signals are used as shown in Figure 3.18. The evaluation circuitry is connected to the feedback loop at point 4 in Figure 3.1.

The signal that exists in the feedback loop,  $L(t)$ , consists of some combination of the initial signals,  $s_1(t)$  and  $s_2(t)$ . The proportion of each signal present in the loop is measured using an analog correlation of the loop output with both the initial signals, *s<sub>1</sub> correlation voltage* and *s<sub>2</sub> correlation voltage*. Each correlator is an analog 50 kHz multiplier followed by a low pass filter with a cutoff frequency of 1 Hz allowing a 1 second integration time. This filter time constant is on the same order as that of the photorefractive crystal, ensuring that the observed correlation changes with the slowly varying signal amplitudes. If the DC level at the filter output is at zero volts, then the initial

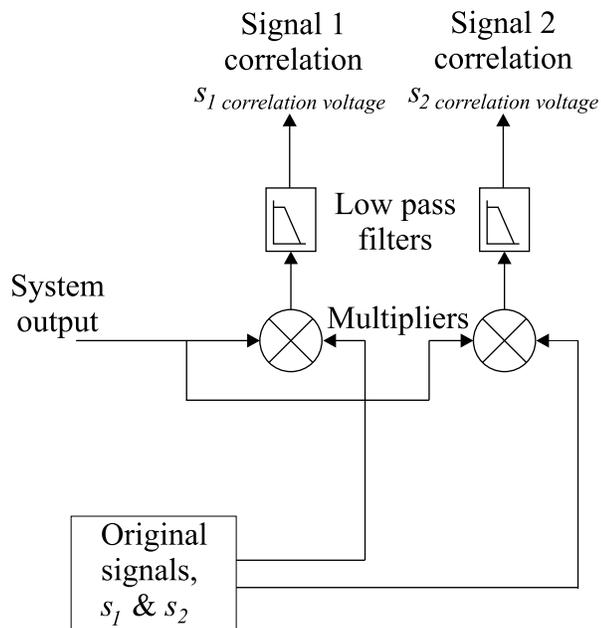


Figure 3.18: Output performance and evaluation circuitry. This subsystem performs correlation of the system output with the original signals. The separation performance is calculated as a ratio of these two correlations.

signal and the loop signal are uncorrelated. The achieved signal separation is determined by the ratio of the output DC voltage for both initial signals as given by the formula:

$$Signal\ separation = \left| 20 \log \left( \frac{s_1\ correlation\ voltage}{s_2\ correlation\ voltage} \right) \right| \quad (3.12)$$

The signal separation parameter is used in chapter 4 as an evaluation of the system performance.

## 3.7 Summary

In summary, the implementation of the 100 kHz bandwidth opto-electronic ICA system uses several sub-systems in order to achieve a physical realization of the theory described in chapter 2. The input channels are designed to linearly modulate the computer-generated orthogonal signal mixtures onto laser beams before they are input into the feedback loop. The optical carrier is suppressed from the phase modulated spectrum to eliminate false correlation between the signals using photorefractive carrier suppression. The feedback loop photorefractive crystal is set so that its gain is linear by decreasing the loop signal intensity. The ICA feedback loop signal is detected using a homodyne detection scheme. A second feedback control loop maintains the proper carrier phase against environmental variations. In order to evaluate the system performance, the output is compared to the original signals using two analog correlators.

# Chapter 4

## 100 kHz ICA experimental results

### 4.1 Introduction

The theory described in chapter 2 can be used to verify the system described in chapter 3 through experiment. This chapter details the experimental data generated by the opto-electronic ICA system, in particular the open-loop gain for sinusoidal and Gaussian signals. The sinusoidal steady-state signal separation performance as described in section 2.3.4 is measured. In addition, signals of arbitrary probability densities, as described in section 2.4.6 are separated. The verification of the system dynamics includes experiments with super-Gaussian and Gaussian signals (section 2.4.6). This chapter also describes experiments with time-varying mixtures, and with different types

of communication signals, including frequency modulated signals competing against Gaussian noise and two ones-density bitstreams competing against each other.

## 4.2 Open-loop gain measurements

The experimental setup for measuring the open-loop gain is shown in Figure 4.1. The feedback loop has been opened between the loop photodetector and loop preamplifier. The input signal,  $v_{in}$  is a small amplitude signal that can be input into the loop preamplifier. The signal is detected by the loop photodetector after traversing the feedback loop. The feedback loop gain,  $A$ , is defined as

$$A = \frac{v_{out}}{v_{in}}. \quad (4.1)$$

As stated in chapter 2, there are two types of gain present in the feedback loop: linear electronic gain, generated by the electronic amplifiers, and non-linear gain generated by the combination of the electro-optic modulator and the two-beam coupling in the photorefractive crystal. Two-beam coupling gain requires a second input beam into the photorefractive crystal. Ideally, both input channels are identical and either can be used. The input channel driving voltage is set to the level that it will be at when the system is operating with a closed feedback loop:  $\pm 200$  volts driving the input modulator.

The output signal can be measured in one of two ways, depending on the type of input signal that is selected. If a sinusoidal signal is used, then

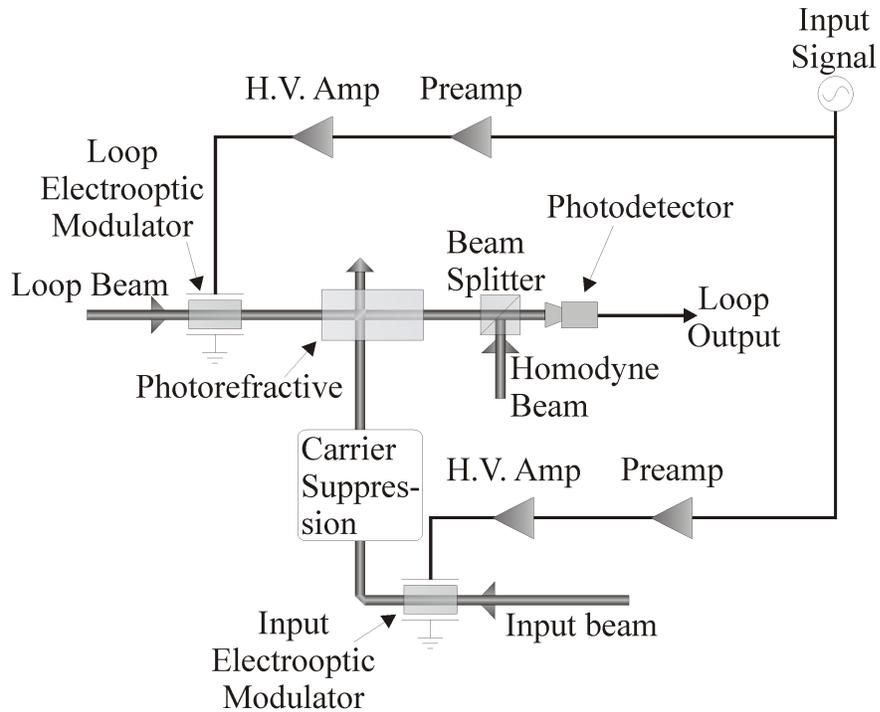


Figure 4.1: Open-loop system setup. The feedback loop in the open-loop gain measurement is broken between the loop photodetector and the loop preamplifier. The input signal is input directly into the feedback loop and on one of the two input channels. Both the input signal amplitude and probability density function can be varied.

the output signal amplitude can be measured on a digital signal analyzer, Figure 4.2. In this case, the digital signal analyzer is used to both generate the input signal and to compute the gain from Equation 4.1.

If a random signal with an arbitrary probability distribution function is used, the output signal amplitude can no longer be read off the digital signal analyzer. In this case, an analog correlator must be used to compute the gain. This correlator is identical to the ones that are described in section

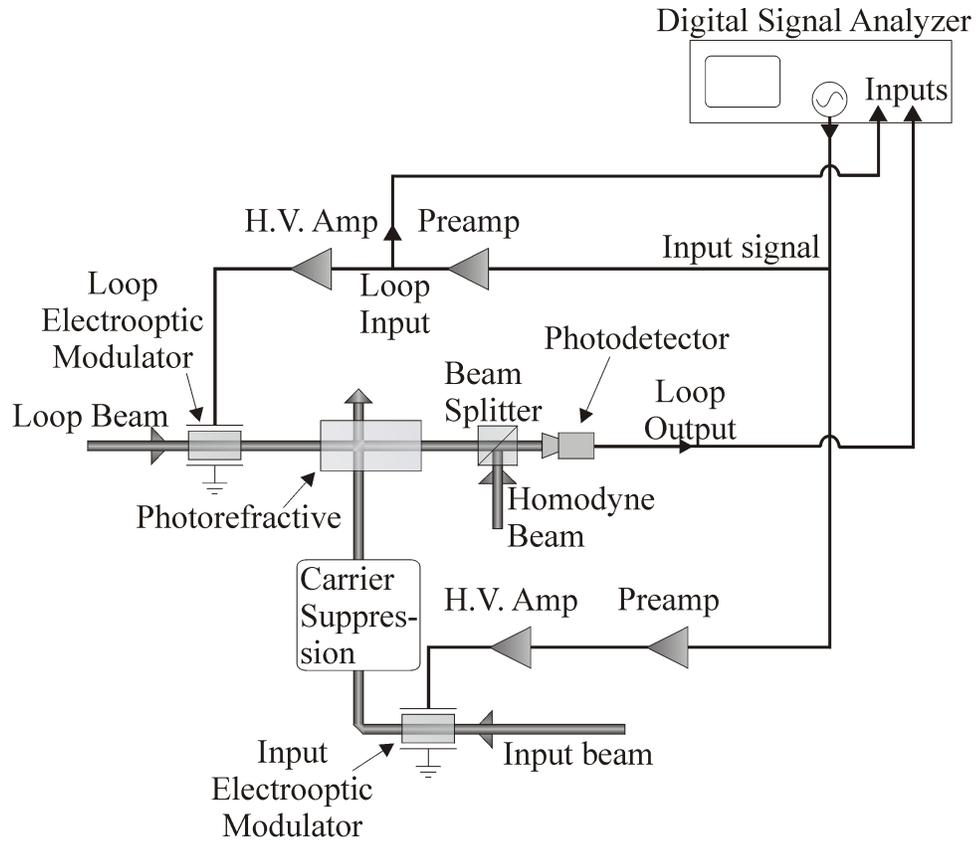


Figure 4.2: Open-loop gain measurement using digital signal analyzer. The digital signal analyzer is used to generate both the input sinusoidal signal and perform the input to output signal ratio.

3.6, and consists of an analog multiplier followed by a low pass filter. This measurement setup is shown in Figure 4.3. In this case, two correlations are preformed: on the input signal and on the output signal. The input signal correlation with itself gives  $V_{in}$ , and acts as a calibration voltage for the correlator. The output correlation measurement gives  $V_{out}$ . The gain is found by taking the ratio of the two DC correlation voltages.

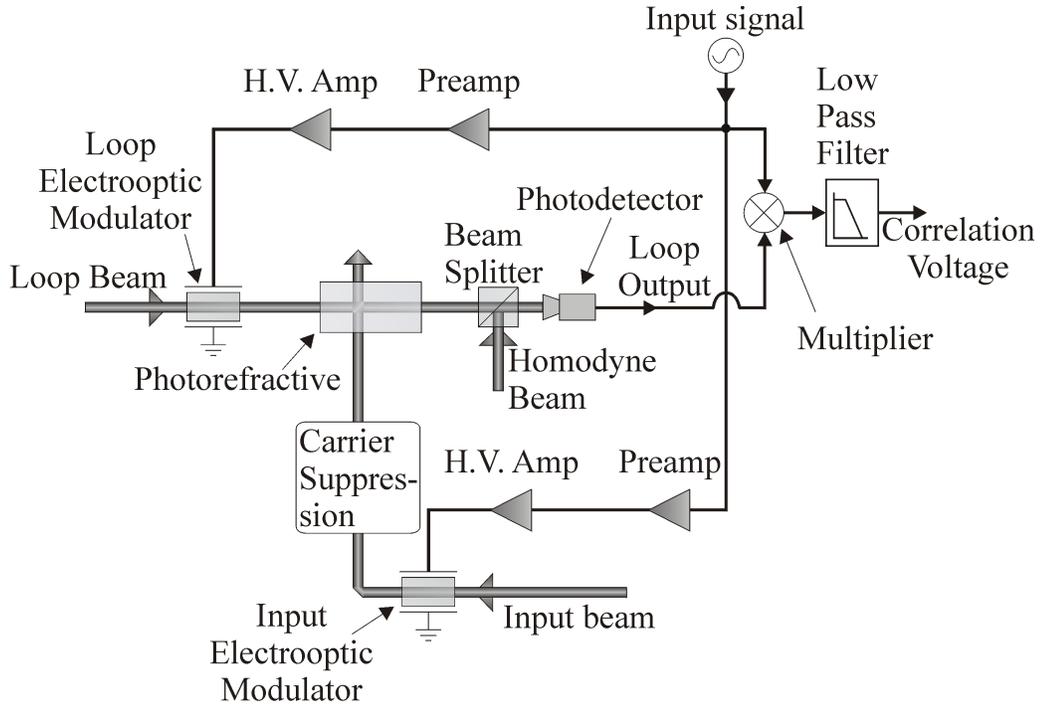


Figure 4.3: Open-loop gain measurement using analog correlator. An analog multiplier and 1 Hz cutoff low pass filter are used to make an output correlator, which enables open-loop gain testing on random signals with arbitrary probability distribution functions.

### 4.2.1 Sinusoidal signal measurement

Using the digital signal analyzer setup described in Figure 4.2, the gain curve shown in Figure 4.4 has been measured by sweeping the input signal amplitude from 0 to 42 mV. The horizontal axis has three sets of labels: the measured voltage after the photodetector, the computed voltage after electronic amplification, and normalized to  $V_\pi$ . The solid line shows the gain computed from Equation 2.26 assuming an initial gain,  $G_0$ , of 160 and fitted

to the proper  $V_\pi$  axis. The parameter  $G_0$  is linearly dependent on the preamp gain, which cannot be set to the proper value until the loop is closed (see section 4.3.3). As a result,  $G_0$  varies somewhat in the plots shown in this section, but always stays within a reasonable range, from approximately 100 to 300.

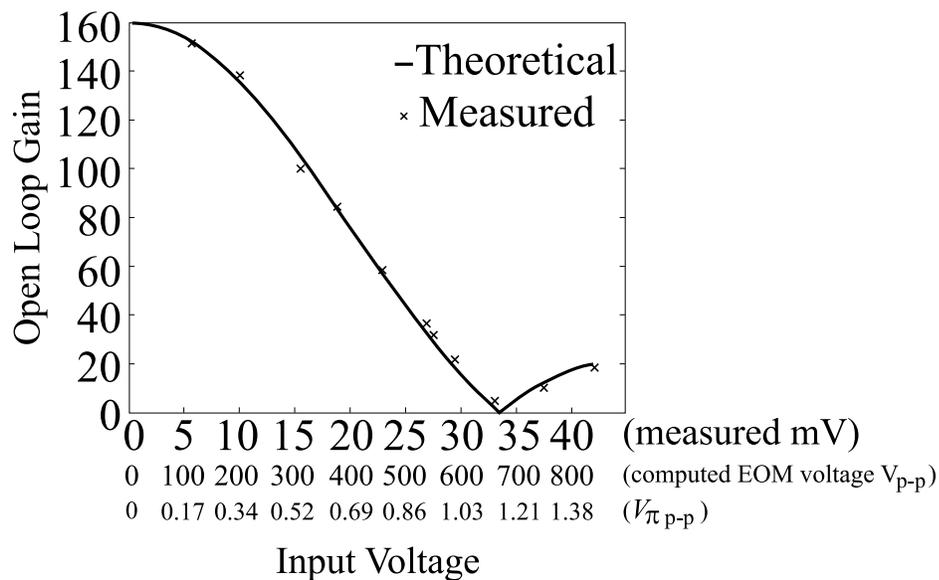


Figure 4.4: Sinusoidal open-loop gain: measured (x) and theoretical (solid line).

The gain of the two input channels must be closely matched. Figure 4.5 shows a gain measurement using both channel 1 and channel 2. By varying the position of the beams in the loop photorefractive crystal, the gain of each channel can be changed until they are precisely matched. Once this gain match is accomplished, the position of the input beams and loop beams

are fixed, and gain variations in the two channels are compensated by the variable preamplifiers.

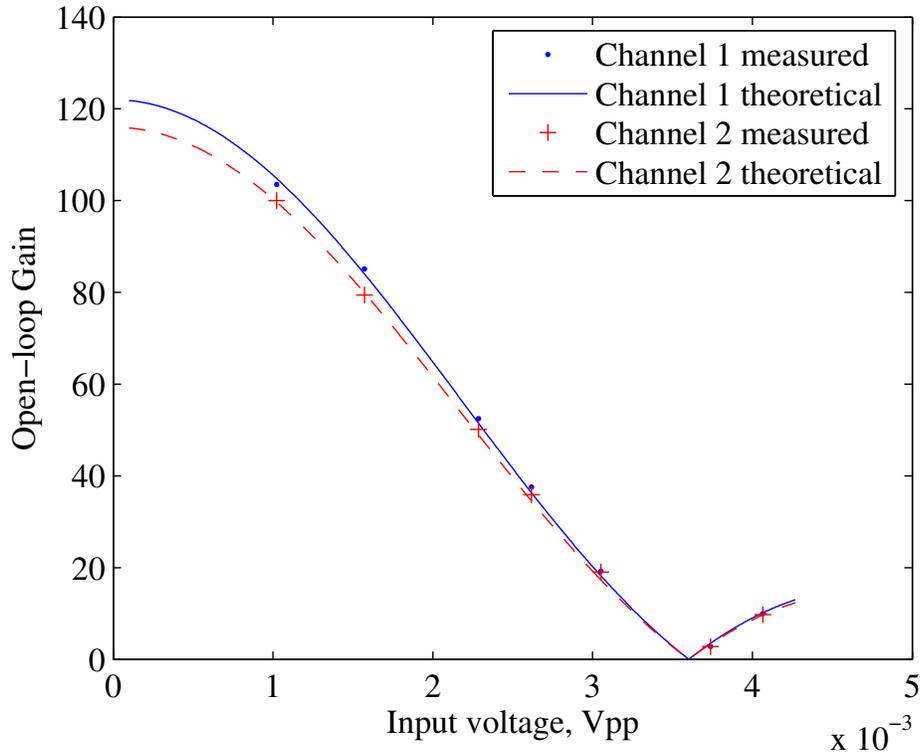


Figure 4.5: Two-channel sinusoidal open-loop gain results. This plot shows the measured and theoretical open-loop gain for both input channels. The gains in each channel must be adjusted until they are closely matched.

If the gain provided by the photorefractive crystal is saturated, as discussed in section 3.4.2, the total gain of the feedback loop drops for lower input voltages, as shown in the curve in Figure 4.6. Also shown in Figure 4.6 is a comparison of the data taken by the digital signal analyzer setup (Fig-

ure 4.2) and the correlator setup (Figure 4.3). The dots are DSA data, and the crosses are correlator data. These two methods show good agreement with each other. This comparison indicates that the correlation measurement method is as accurate as the digital signal analyzer measurement and that all scaling factor differences between the two measurements have been accounted for properly.

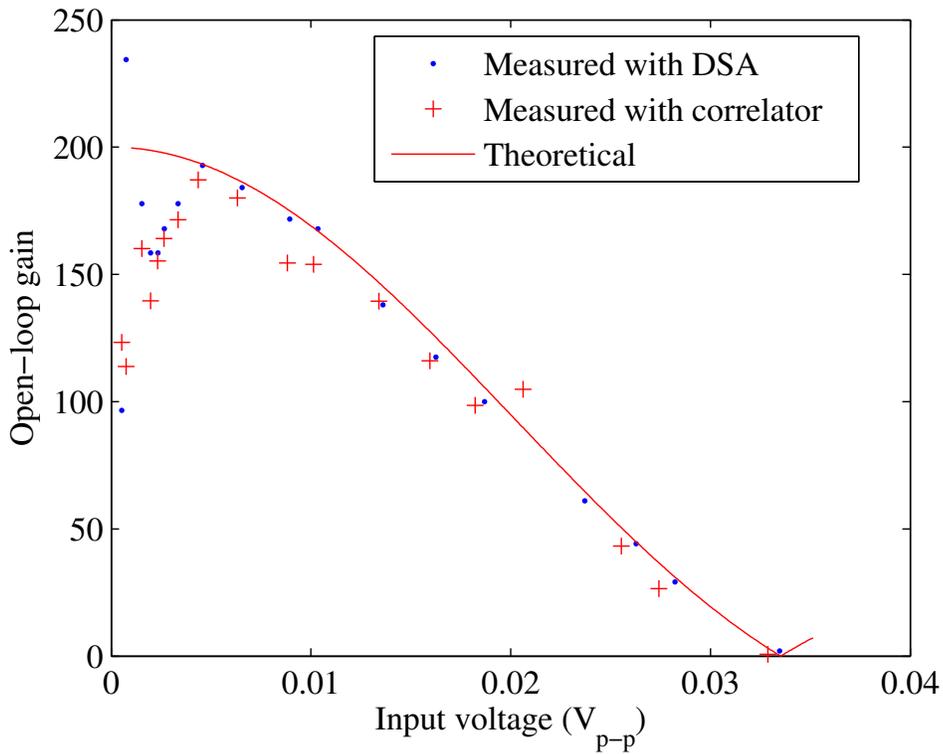


Figure 4.6: Open-loop gain with saturated photorefractive gain. This plot shows decreased gain at lower input voltages when the photorefractive gain is beginning to saturate. The crosses and dots demonstrate the match between the digital signal analyzer and analog correlator measurement techniques.

## 4.2.2 Gaussian signal measurement

One of the most important conclusions from the chapter 2 theory is that the gain of the feedback loop is dependent on the probability density function of the input signal. The gain of a completely random signal with a Gaussian PDF, measured using the setup of Figure 4.3, is shown in Figure 4.7. The data points are the measured data, and the solid line is a theoretical fit using the gain curve from Equation 2.26 given a Gaussian input. The maximum gain on this plot is shown to be approximately 8, which is lower than all values found in the sinusoidal case. The reason for this is that the input signal, generated by the Agilent 33250A arbitrary function generator, needs to have an amplitude of 6.1 volts zero-to-peak in order to have the same variance as the sinusoidal signals. However, the highest voltage that the function generator is capable of achieving is 1.56 volts, almost 4 times too small. This severely limits the gain of the photorefractive because the input beams are less intense, causing saturation more easily than in the sinusoidal case. However, despite this limitation, the measured data follows the theoretical data very well.

The data in this section confirm that the open-loop gain of the system is dependent on the probability density function of the input signal. In addition, the measured gain closely represents the theoretical gain of Equation 2.26.

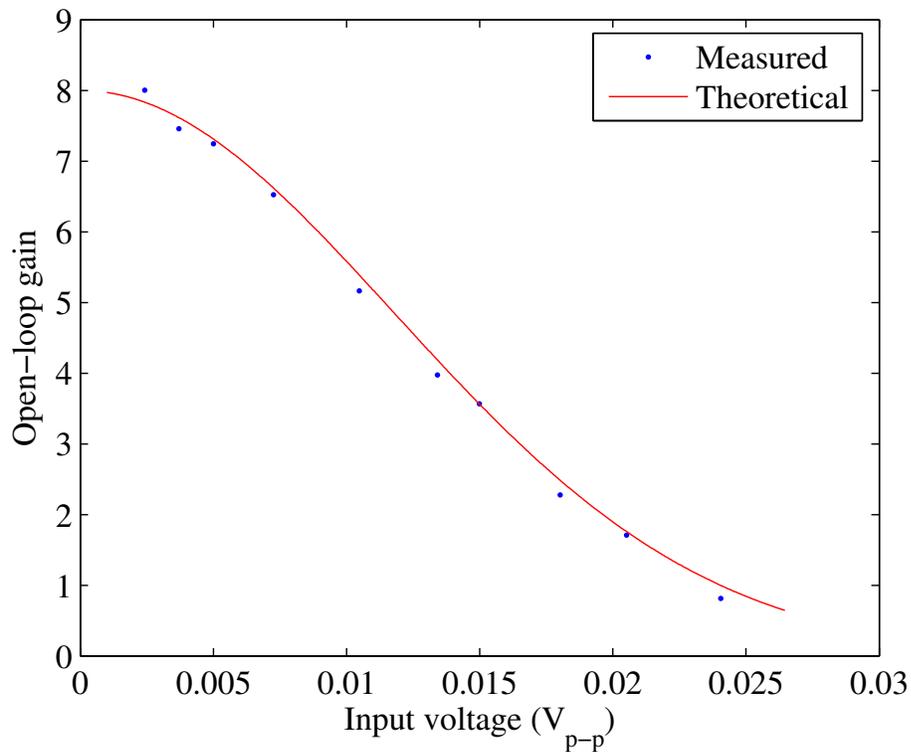


Figure 4.7: Gaussian signal open-loop gain. This plot shows the measured and theoretical open-loop gain of a random signal with Gaussian distribution.

### 4.3 Sinusoidal signal separation

The ultimate goal of the opto-electronic ICA feedback loop is to perform signal separation on two input signals. The sinusoidal input signals used in this section are always generated and mixed by computer. The output is observed in two possible places in the feedback loop: either taken directly from the photodetector, or after amplification by the loop preamp.

### 4.3.1 Experimental setup

The setup used to perform the sinusoidal signal separation is shown in Figure 4.8. This is almost exactly the same setup as is described in the previous section except that the loop is now closed. The photodetector output is fed to the digital signal analyzer to determine the signal separation value, and the preamp output is observed on an oscilloscope to determine the qualitative separation. The oscilloscope measurement proves to be necessary during the initial tuning stages despite the fact that it not used to record data because the signal on the oscilloscope changes in real time as the system is tuned. In contrast, the signal viewed on the digital signal analyzer has approximately one second of delay. The delay makes swift tuning of the input channel gain and phase difficult when using only the DSA.

The resulting data is represented on plots of signal separation versus mixture rotation angle. Recall from section 3.3.1 that an orthogonal signal mixture can be given by one parameter,  $\theta$ . To show that the system behaves the same for all input mixtures, the separation is measured for input mixtures that range from -45 degrees to + 45 degrees.

### 4.3.2 Sinusoidal signal separation measurement

Using two sinusoidal signals, with frequencies of 1.5 and 1.6 kHz, and equal powers, the signal separation shown in Figure 4.9 is measured. To gather this data, all signal mixtures are prepared and stored on computer. The signal

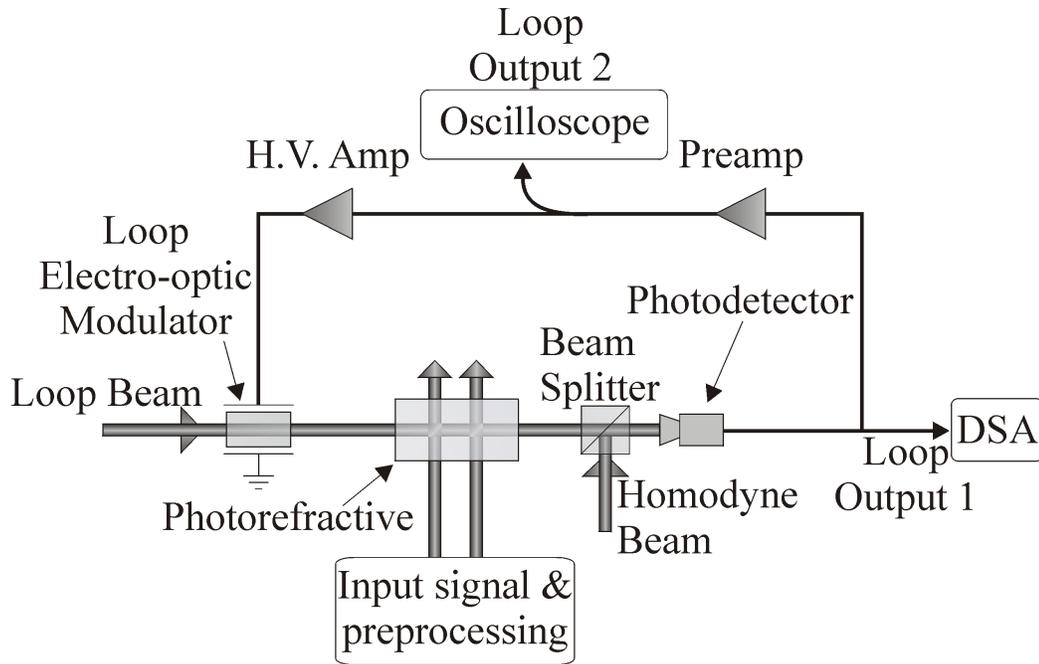


Figure 4.8: Closed-loop sinusoidal signal measurement setup. The closed-loop experimental setup for sinusoidal signals requires only the digital signal analyzer to determine the signal separation value. The pre-amplified signal is also observed on the oscilloscope for ease of adjustment.

mixture is rotated from +45 to -45 degrees in 5 degree steps. Figure 4.9 shows 20 dB signal separation for all mixtures.

Upon further investigation of this data, it appears that the system achieves higher signal separation for mixture rotations with a positive  $\theta$ . The reason for this is that the system is tuned to achieve the best separation possible at the beginning of the rotation measurements and not changed throughout the remainder of the measurement. The decreased system performance at negative mixture angles is due to environmental drift changing the relative

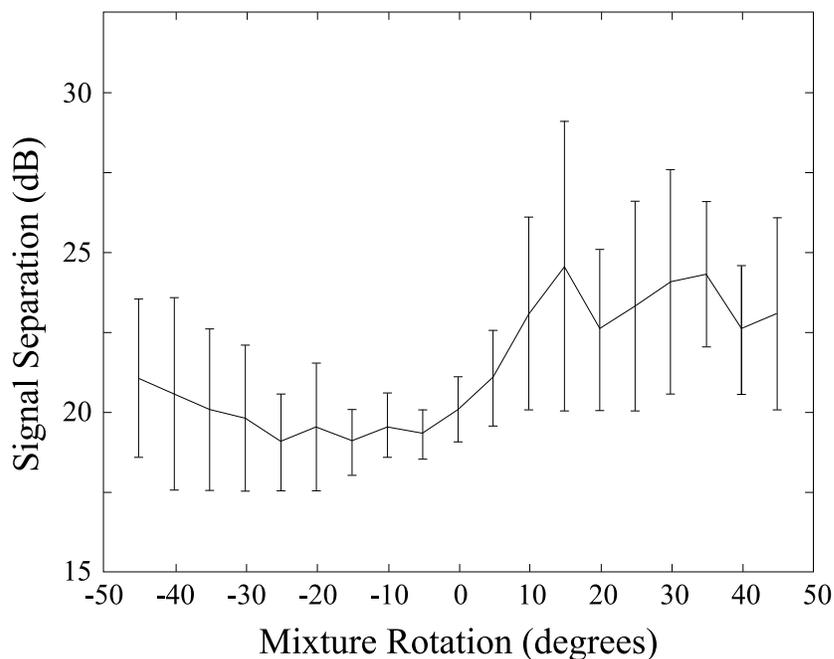


Figure 4.9: Sinusoidal signal separation. This plot shows the measured signal separation for mixtures between  $-45$  and  $45$  degrees on two sinusoidal signals, measured with the digital signal analyzer.

photorefractive gain in the two input channels.

The best signal separation that has been observed for two sinusoidal signals is on the order of 30 dB, although this high separation value is difficult to achieve for more than a minute or two before it drifts to a more consistent value of 20 dB. At a separation value of 30 dB, the measurement is limited by the system noise, which is described in more detail in section 4.6.4.

To quantify the high sensitivity environmental drift has on signal separation, particularly at high separation values, the following experiment was performed on a 45 degree mixture of signals. First, the input channel gains

are set to provide the maximum signal separation. This is measured at 20 dB. Next, in less than 2 seconds, (shorter than the change due to environmental drift) the gain on preamp 2 is varied by 6.5%, from 18.2 to 17.0. This change, the smallest that can be accurately reproduced by hand on the preamps, causes 12 dB decrease in signal separation. The new signal separation value is approximately 8 dB. This experiment shows that extremely small changes in the gain can cause high signal separation variations.

### 4.3.3 Loop loss measurement

The proper way to set the loop preamplifier gain is by using the loop electronic gain to completely compensate for the loop loss. In order to accomplish this, the loop loss must be measured. To measure the loss, the photorefractive two beam coupling is turned off by blocking the input beams. At this point, the gain and the loss in the feedback loop are both linear. If the electronic gain in the loop is increased, eventually the total loop gain will rise above unity, and the loop will break into spontaneous oscillation. It is at the threshold between no oscillation and spontaneous oscillation that the gain and the loss in the loop are equal. Ideally, the gain in the loop is set to just below this point.

We have measured the gain at the point when the loop breaks into spontaneous oscillation to be approximately 50 dB. The gain is therefore be reduced to around 40 dB during system operation, less than the point that it breaks into spontaneous oscillation. The gain provided by the photorefractive is

high enough to overcome the loss. In other words, the addition of the photorefractive gain makes the total gain higher than the oscillation threshold.

#### 4.3.4 Phase error results

In section 3.3.2, the different methods for varying the phase of the input signals is described. However, it is convenient to know how the system behaves when there is a phase mismatch between the two channels. If the input mixture angle is at 0 degrees, then a phase error in one of the channels merely phase delays one of the two signals, making it a new signal. In this case, a phase error should have no affect on the signal separation. The phase error between the two channels should affect the system the most when the signal mixture is at  $\pm 45$  degrees. In this case, each signal is present on each channel. Signal 1 on channel 1 will therefore be phase delayed with respect to signal 1 on channel 2. As these two portions of signal 1 become more out of phase with each other, they also become more uncorrelated, resulting in a reduced signal separation.

Based on the heuristic model of phase error between the two channels, undiminished signal separation is expected at 0 degrees mixture rotation, while reduced signal separation is expected at  $\pm 45$  degrees. A signal separation graph, presented in Figure 4.10, was made with phase error of approximately 14 degrees due to the Trek 601b phase error discussed in 3.3.2. This plot shows exactly the pattern expected for a phase error.

It is possible to use the difference in signal separation at 0 degrees and

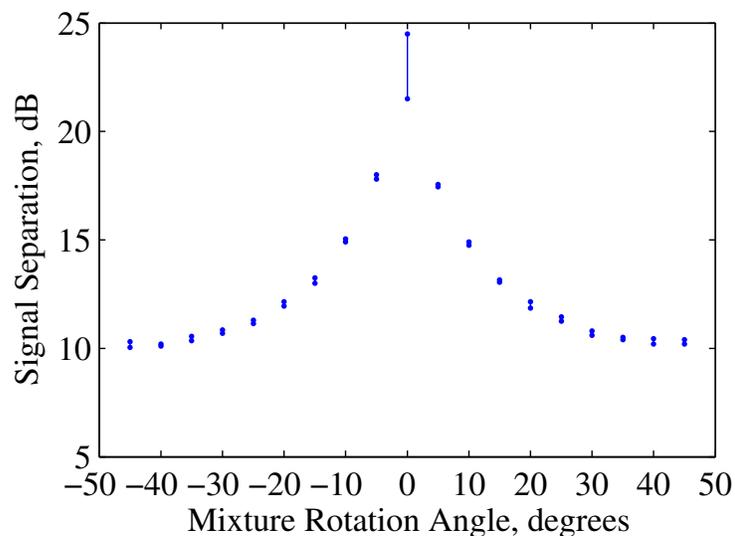


Figure 4.10: Sinusoidal signal separation with phase error. This plot shows the measured signal separation for mixtures between  $-45$  and  $45$  degrees on two sinusoidal signals with a  $14$  degree phase error between the two input channels. The signal separation at  $0$  degrees is still near  $20$  dB, but it decreases at  $\pm 45$  degrees.

$45$  degrees to calibrate out phase error in the channels. Two signal mixtures, at  $0$  degrees and at  $45$  degrees, are generated. The phase on one channel is tuned while the  $45$ -degree mixture is input into the system. Once the signal separation at  $45$  degrees reaches the signal separation at  $0$  degrees, the phase is properly set.

This method for setting the phase of the input channels is also useful for setting the gain. Again, using the same logic as for phase, the gain error between the channels is the worst for an input mixture of  $45$  degrees. Therefore, when the  $45$ -degree mixture has the same separation as the  $0$ -

degree mixture, both gain and phase are properly set.

### 4.3.5 Signal winner measurements

A gain error between the two channels also appears in a different way for a 0 degree mixture. In this mixture, a gain error will increase the power present in one signal with respect to the other. Recall from section 2.4.1 that the gain in the photorefractive crystal is independent of the signal mixture rotation angle provided that the signals are orthogonally mixed. The power mismatch between the signals invalidates this assumption. The signal with more power will receive more gain within the photorefractive medium. As a result, with a 0-degree mixture, the signal that has more power will be more likely to win the competition.

Section 2.3.3 demonstrated that the system should randomly choose one of the two input signals with equal probability. If there is a gain mismatch between the two channels, the system will preferentially choose one signal over the other.

In an attempt to measure the randomness of the signal selection, the following experiment is performed. First two signals with a zero degree mixture are input into the system and the system selects one of the signals. Then, the inputs are turned off, the photorefractive gratings are given time to dissipate with the help of some bright white light that is shined on the photorefractive crystal. Next, the inputs are turned back on and the system re-selects one of the signals. Ideally, the probability of selecting one signal after several mea-

surements is 50%. After 11 tests, we have shown that the system selected signal 1 six times, and signal 2 five times.

## 4.4 Arbitrary PDF signal separation

After the completion of the sinusoidal signal separation experiments, it is necessary to test the system with completely random signals that have arbitrary probability density functions, as described in section 2.4. These random signals are generated by computer, as described in section 3.3.1, and detected using the correlation scheme in section 3.6. The accuracy of the correlation scheme is verified using sinusoidal signals, as shown in section 4.2.2. Ideally, the two signals output from the computer must be completely uncorrelated. Using the analog correlators, the correlation of the two generated signals is below the measurement noise floor.

The probability density functions used in the arbitrary PDF measurements are:

$$p_L(x) = \frac{1}{2} \exp(-|x|) \quad \text{Laplace} \quad (4.2)$$

$$p_G(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-|x|^2}{2}\right) \quad \text{Gaussian} \quad (4.3)$$

$$p_E(x) = A \exp\left(\frac{-|x|^3}{2}\right) \quad \text{Exponential 3} \quad (4.4)$$

$$p_H(x) = \frac{1}{\pi\sqrt{1-x^2}} \quad \text{Harmonic} \quad (4.5)$$

These four density functions represent super-Gaussian, Gaussian, and two sub-Gaussian densities respectively, as described in section 2.4.6. The difference between the functions used in section 2.4.6 and in these measurements is that the uniform density has been replaced by the  $\exp(-|x|^3)$  density. The reason for this difference is that the densities from Equation 4.2 to Equation 4.4 can all be represented by the function:

$$p(x) = A \exp\left(\frac{-|x|^\lambda}{B}\right) \quad (4.6)$$

by changing  $\lambda$  from 1 to 3, and choosing  $A$  and  $B$  to maintain a unity area integral of a PDF.

#### 4.4.1 At least one sub-Gaussian signal separation measurements

Figure 4.11 shows signal separation plots for eight combinations of the signals represented by Equation 4.2 through Equation 4.5. In Figure 4.11, the winning signal is always the signal with the lowest kurtosis, as expected from the theory in section 2.4.6. This implies that the sub-Gaussian signals, the harmonic and the  $\exp(-|x|^3)$ , will always win over the Gaussian and super-Gaussian signals. If two sub-Gaussian signals compete either one can win, although the system shows a preference for the signal with harmonic PDF, which has slightly lower kurtosis.

The large variations in the signal separation, from 20 dB to over 25 dB arise from the high sensitivity of the system to small differences in channel gains, as described in section 4.3.2.

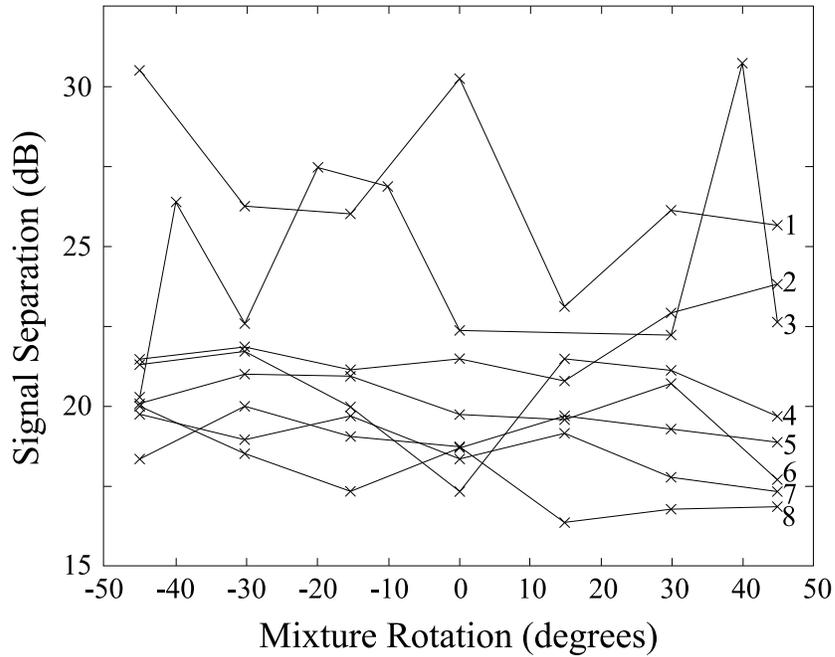


Figure 4.11: Random signal separation with arbitrary PDFs. Computer generated random noise signals with super-Gaussian, Gaussian, and sub-Gaussian densities have been tested. In every case, the winning signal has lower kurtosis. Line number: 1. Laplace vs. Harmonic 2. Harmonic vs. Harmonic 3. Gaussian vs. Harmonic 4. Harmonic vs.  $\exp(-|x|^3)$  5.  $\exp(-|x|^3)$  vs.  $\exp(-|x|^3)$  6. Laplace vs. Gaussian 7. Laplace vs.  $\exp(-|x|^3)$  8. Gaussian vs.  $\exp(-|x|^3)$ .

#### 4.4.2 Gaussian signal separation measurements

The theory in section 2.4.6 predicts that the system will not be able to adequately separate two purely Gaussian signals. The Gaussian signal separation experiment results in the separation plot shown in Figure 4.12.

This plot shows that the signal separation is not constant for any input mixture rotation. The separation values range from nearly 0 dB to 10 dB. In addition, this data changes as it is recorded, making data collection difficult for these signals. This data is consistent with the prediction that the system is unable to separate two Gaussian signals.

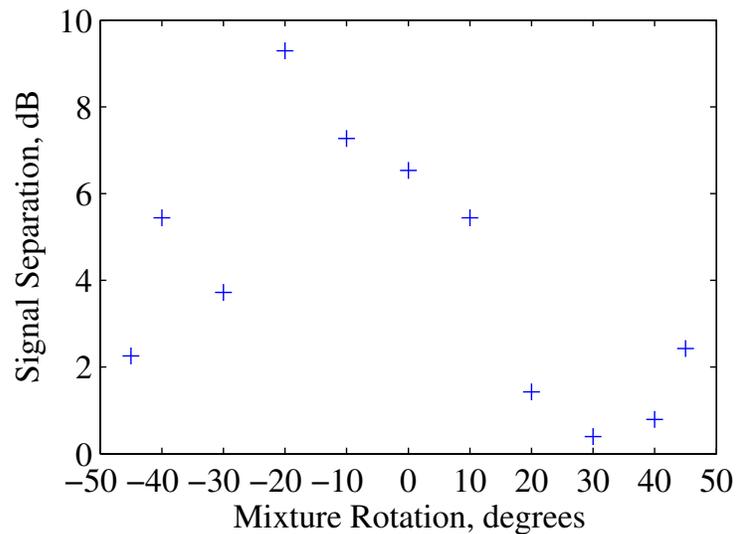


Figure 4.12: Random signal separation with Gaussian signals. The signal separation for random signals with Gaussian probability densities varies between 0 and 10 dB. The separation for each of these points also changes with time.

### **4.4.3 Laplace signal separation measurements**

Section 2.4.6 theory predicts that the system is unable to separate two super-Gaussian signals as well. However, unlike the purely Gaussian case, the super-Gaussian signals have equal amplitudes at the system output. A 0 dB signal separation has been measured with both a 0-degree mixture and a 45-degree mixture.

## **4.5 Time-varying measurements**

All of the separation data thus far has been presented as signal separation versus input mixture angle. The signal separation value is only one data point for every input mixture. While the system is designed to operate at steady state, when the output should only contain one signal, it is important to observe the system behavior as it varies in time.

### **4.5.1 Time-varying system setup**

In order to track the signal separation as a function of time, it is necessary to read the correlation measurement voltages more rapidly than the system adapts to the signals. This rapid reading of the multimeters is accomplished by a GPIB link to a computer. The two multimeters can be probed at approximately 2.8 times per second. This rate is slightly faster than the 1-Hz correlation filters and is fast enough to measure the signal separation.

## 4.5.2 Time-varying signal separation

Figure 4.13a shows the correlation voltage of the two signals (1.5 and 1.6 kHz sine waves) over a period of 33 seconds. The signals are turned on at approximately 4 seconds. In this experiment, the correlation voltage for both signals is negative. This is due to the random permutation ambiguity described in appendix 1. If this same experiment is run at a later time, one or both of the correlation voltages could be positive.

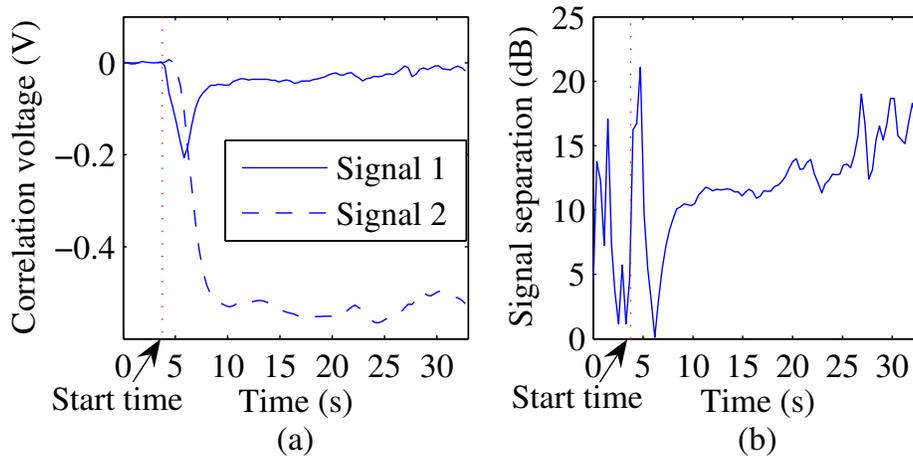


Figure 4.13: Correlation voltage as a function of time. The correlation voltage and signal separation in this plot are measured in time. The signals are turned on at 4 seconds and have achieved 10 dB separation by 8 seconds.

Both signals grow at approximately the same rate for the first 2 seconds, from a correlation voltage of 0 volts to -200 mV. After two seconds, one of the signals begins to be suppressed. Figure 4.13b shows the signal separation computed from this data. The data is inaccurate when both signals are

turned off, resulting in noise for the first 4 seconds of the plot. The separation drops to 0 dB after the signals have been turned on but before they are separated. During the next two seconds, the separation increases to 11 dB. The separation then slowly increases from 11 to approximately 15 dB during the rest of the measurement. In this experiment, the system does not perform as well as it did during the signal separation measurements described in section 4.3.2, likely due to a misalignment of the interferometer.

### 4.5.3 Time-varying mixtures

The mixing matrix will be time varying if the signals are moving or if the propagation environment changes in time [49], [50]. The time required to separate two signals in the system is dictated by the photorefractive time constant. Essentially, it takes around one second to set up a grating within the photorefractive crystal. This time period is convenient for long integration times of the signals, but limits how quickly the signal mixture can change.

To measure how easily the system can track one signal with a changing mixture, the following experiment is performed. The input signal mixture rotation angle is increased at different rates from 1 degree per second to 60 degrees per second. The correlation voltages are recorded, and the signal separation is computed from the data.

Figure 4.14 shows signal 1 or signal 2 winning for a 1-degree per second rotation angle increase. As can be seen from the output correlation voltages,

the system tracks one signal through a mixture that varies at this rate.

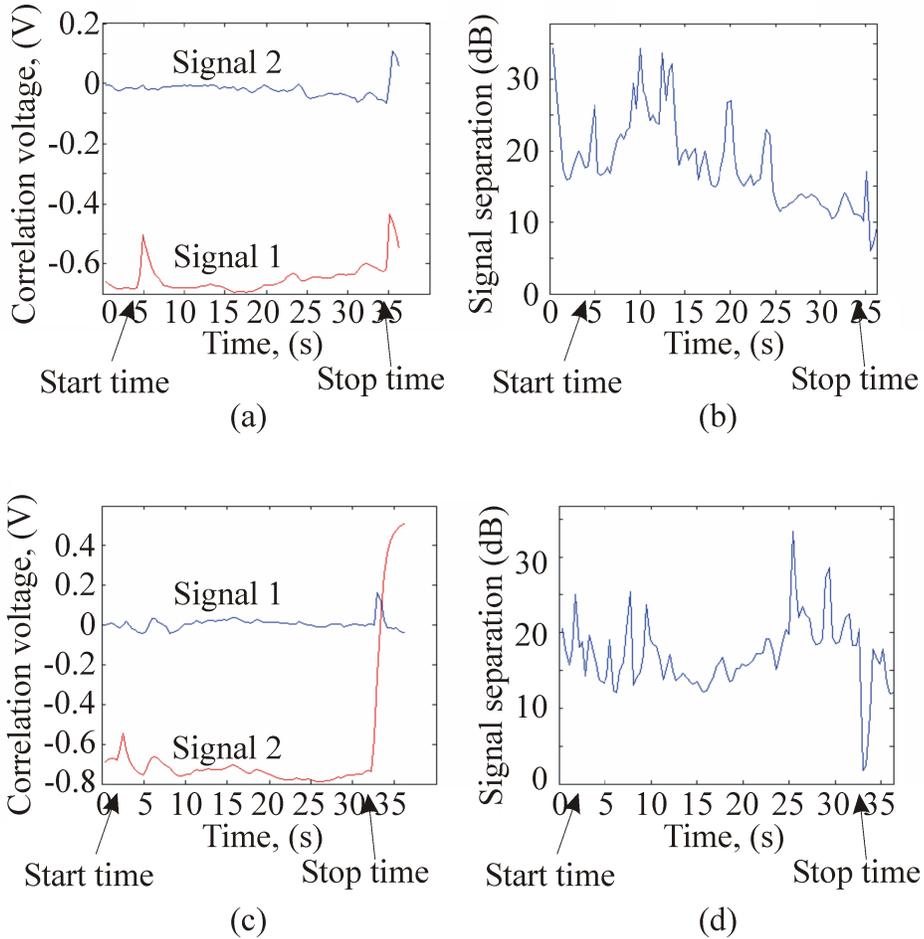


Figure 4.14: 1-degree per second rotating mixture. Plots (a) and (b) show signal 1 winning with a rotating mixture. (c) and (d) likewise show signal 2 winning.

Figure 4.15 shows a 20-degrees per second rotation. Here, the system still tracks a single signal, but the signal separation periodically varies between 5 dB and 20 dB. Figure 4.16 shows an increase in the rotation rate to 30-

degrees per second. In this case, the system is unable to keep up with the mixture variation, and the winning signal alters periodically between signal 1 and signal 2.

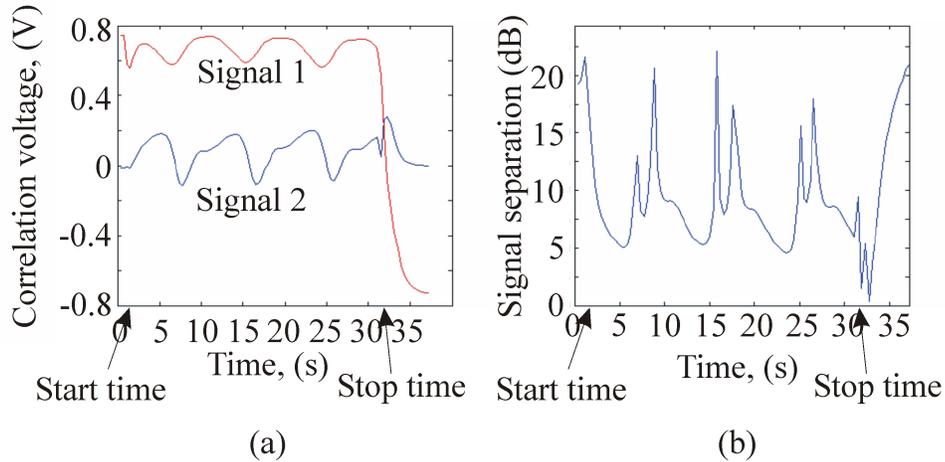


Figure 4.15: 20-degrees per second rotating mixture. (a) The correlation voltage and (b) signal separation as a function of time. A mixture that changes at the rate of 20-degrees per second shows an average signal separation of approximately 10 dB. However, the system is still able to track one of the signals. The peaks in the signal separation value correspond to mixture angles that are multiples of 90 degrees, when the signals are not mixed on the channels.

These plots show experimentally that the system is able to track a signal mixture that is changing at 20-degrees per second, but not at 30-degrees per second. While this experiment highlights the system’s ability to track through time varying mixtures, it is difficult to give physical insight into a moving orthogonal mixture. If two signals are transmitted from two moving sources, and the received mixture is whitened, then it is this whitened

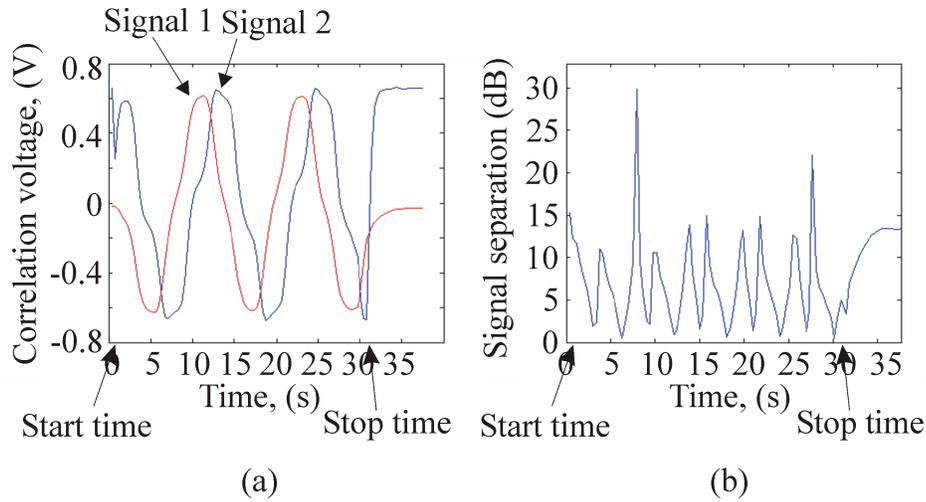


Figure 4.16: 30-degrees per second rotating mixture. (a) The correlation voltage and (b) signal separation as a function of time. A mixture that changes at the rate of 30 degrees per second shows an average signal separation of approximately 10 dB. In this case, the system is unable to track one of the signals. The winning signal changes from signal 1 to signal 2 as the mixture rotates.

mixture that can change at 20-degrees per second and still be separated. However, there are many variables that could cause the original mixture to vary extremely rapidly, including the source distance from the receiver, time-varying transmitted power, and the propagation environment. It is outside the scope of this thesis to quantify the rate of change in signal mixtures for all these variables.

## 4.6 FM signal and Gaussian noise separation

The opto-electronic ICA system can also be used with data carrying signals to make an audio demonstration of its separation ability. The first example is a frequency-modulated (FM) carrier that is mixed with Gaussian noise. The modulating signal is an arbitrarily chosen music signal. Ideally, when the output of the system is demodulated, the music will be heard and the noise suppressed by 20 dB.

### 4.6.1 Frequency modulation probability density function

As described in section 2.4.6, if at least one of the input signals is sub-Gaussian, the opto-electronic ICA system will separate the two signals. In a mixture of a Gaussian signal with a sub-Gaussian signal, the Gaussian signal will be suppressed.

It is interesting to find what the kurtosis of a music signal is. To answer this question, the kurtosis of nine different music signals (with normalized variance) is computed based on Equation 2.38, and the results are shown in Table 4.1.

Table 4.1 illustrates different aspects of the probability densities of music signals. First of all, they are all super-Gaussian (kurtosis  $>3$ ). This generally occurs because there is a substantial amount of low-amplitude time in a song. Low amplitude time, as seen in Figure 4.17, is the time when the

Table 4.1: Comparison of kurtoses of music signals

Title	Kurtosis
<i>Ave Maria</i> , (Bach)	4.2554
Pachelbel's Canon, (Pachelbel)	5.8162
<i>Can't Buy Me Love</i> , (The Beatles)	3.3499
<i>Fur Elise</i> , (Beethoven)	8.9962
<i>Hallelujah Chorus</i> , (Handel)	5.8620
<i>Moonlight Sonata</i> , (Beethoven)	11.3876
<i>Midnight Train to Georgia</i> , (Gladys Knight)	4.6341
<i>Help!</i> , (The Beatles)	4.7747
<i>Johnny B. Goode</i> , (Chuck Berry)	3.2629

signal amplitude is close to zero. By concentrating the energy near zero, the PDF becomes more peaked, and therefore is more super-Gaussian, as seen in Figure 2.7. This table also illustrates that music signals that have more dynamic range (more amplitude levels), such as the *Moonlight Sonata*, generally are more super-Gaussian than songs with less dynamic range, such as *Johnny B. Goode*.

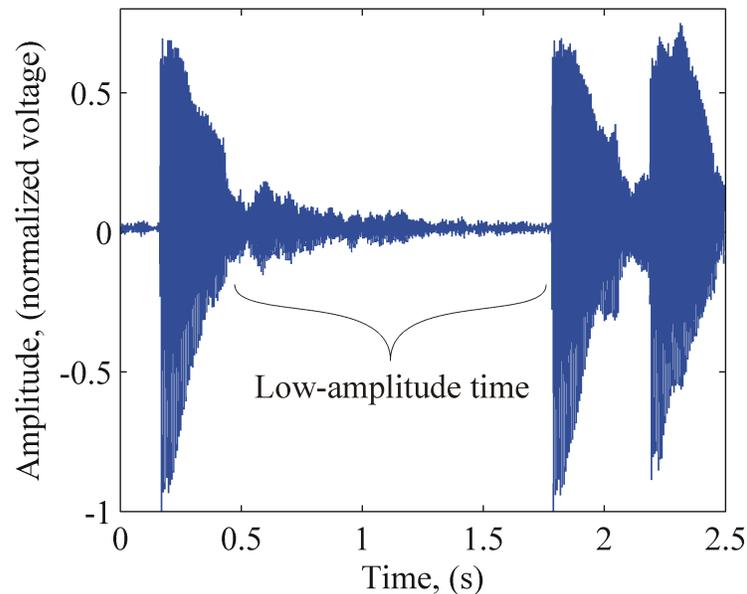


Figure 4.17: Generic super-Gaussian music signal. In this common music signal (from Ray Lynch, *Deep Breakfast*), the amount of time that the signal spends with low amplitude outweighs the high-amplitude time. This tends to make music signals super-Gaussian.

It is clear, however, that if these raw, super-Gaussian music signals are input in the system, they will have less gain than a competing sub-Gaussian

signal and will be suppressed. Since we would like to have the music signal win the competition, suppressing them is not desirable. It is therefore necessary to alter the probability density function of the music signals from super to sub-Gaussian.

A frequency modulated signal is always sub-Gaussian. An FM signal,  $f_c(t)$ , is described by the function [51]:

$$f_c(t) = A \cos \left[ \omega_c t + B \int_0^t f_m(\tau) d\tau \right] \quad (4.7)$$

where  $A$  is the carrier amplitude,  $\omega_c$  is the carrier frequency,  $f_m(t)$  is the modulating signal, and  $B$  is proportional to the modulation depth. This equation shows that the amplitude of the carrier never is changed, but the frequency varies according to the modulation signal. As a result, the probability density function is exactly the same as the probability density function of the carrier itself:

$$p_{Harmonic}(x) = \frac{1}{\pi \sqrt{1-x^2}}. \quad (4.8)$$

This probability density function is the probability density function of a sinusoidal, and has a kurtosis of 1.5 (or an excess kurtosis of -1.5). Thus, all signals, whether super-Gaussian, Gaussian, or sub-Gaussian, will be converted to a sub-Gaussian signal if they are frequency modulated.

## 4.6.2 Digital modulation and demodulation of an FM signal

FM signals can be generated in two different ways: digitally (in Matlab and output as a sound signal as described in section 3.3.1); and using a voltage controlled oscillator (VCO), which will be described in the following section.

The digital generation of an FM signal exactly follows Equation 4.7. The signal to be modulated is chosen. It is then digitally integrated by a cumulative summation. The resulting signal is then simply added to the phase of the carrier signal. Experimentally, the modulation depth,  $B$  is less than 0.3 in order to get non-distorted demodulation. The carrier frequency,  $\omega_c$  must be chosen to be smaller than the Nyquist frequency of the output sampler. For the Audigy sound card, with an output of 44.1 kHz, this means that the carrier frequency must be smaller than 22.05 kHz. However, there is a second concern when generating FM signals. The frequency of the sidebands must also be lower than the Nyquist frequency. This implies that the bandwidth of the modulated signal plus the carrier frequency should ideally be lower than 22.05 kHz. If this is not the case, then there will be aliasing in the demodulated signal. The carrier frequency is chosen to be 15 kHz, and the sound signal bandwidth is limited to 7 kHz.

The second signal, a Gaussian noise source, is also generated in Matlab to have the same variance as the FM voice signal. This determines the amount of power in the Gaussian signal. We choose the noise spectrum and the FM

spectrum to overlap, and furthermore the noise to be bandlimited to about 6 kHz centered at the FM carrier. A measurement of the spectrum of the signal mixture is shown in Figure 4.18.

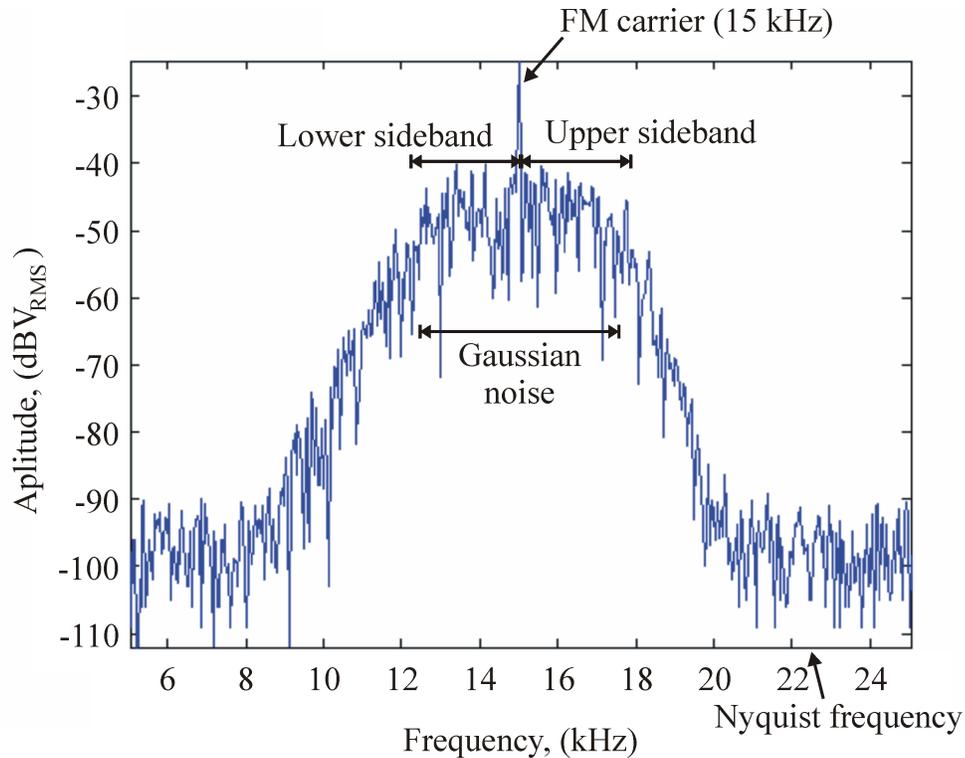


Figure 4.18: Spectrum of FM signal and Gaussian noise mixture. The bandlimited Gaussian noise, centered on the FM carrier, completely overlaps the signal spectrum.

The output of the system is recorded and digitally demodulated. The recording is made on the Audigy sound card at the same time as the signal mixtures are input into the system. The record sampling frequency can be set at either 44.1 kHz or 48 kHz. While either frequency is high enough to

prevent signal aliasing, the 48 kHz recording frequency generates a 3.9 kHz beat tone with the 44.1 kHz output frequency, and is therefore not used.

The recorded and saved signal is demodulated after the recording is made. The drawback to this method is that the music is not played in real time as the system separates the signals. The signal demodulation follows a common FM demodulation scheme. First, the signal is multiplied by a copy of the carrier. This carrier must have the same phase as the recorded carrier. The proper phase is found by searching through all possible phases (from 0 to 180 degrees) in 1-degree increments until the output signal is maximized. The resulting signal is low-pass filtered to remove the frequency doubled components after multiplication with the carrier. A tenth order Butterworth filter with a cutoff frequency of 8 kHz (programmed in Matlab) is used for this task. Both the music and the noise can be heard simultaneously when the signal is demodulated. The ICA system then suppresses the noise at the time constant of the photorefractive.

### **4.6.3 Analog modulation and demodulation of an FM signal**

The frequency modulation can also be done in the analog domain. In this case, the music signal is output from the computer and fed into a voltage controlled oscillator. To demodulate the signal, the carrier is multiplied with the FM signal, the output is low pass filtered, and the correct carrier phase

is found with a phase locked loop.

The advantage of this method is that it can be done in real time. In other words, the signal separation can directly be heard without having to record the output signal first. The disadvantage to an analog modulation and demodulation is that the signals cannot be mixed by computer, and a special mixing circuit is required.

#### 4.6.4 System noise measurements

In making signal recordings of demodulated music signals, it was found that the system added noise. In an effort to remove the noise from the recordings the added noise from each component in the system is characterized.

The measurements of each component (made with a digital signal analyzer), using a 15 kHz input sine wave, are shown in Table 4.2. There are two ways to improve the system SNR: either by lowering the noise or increasing the signal strength. By careful alignment of the interferometer, the signal strength is increased. The signal apparent on the loop photodetector is at  $-32 \text{ dBV}_{rms}$  and the noise floor is at  $-85 \text{ dBV}_{rms}$ . This implies that the best signal separation that can be measured above the noise floor is about 25 dB, approximately equal to the the measured SNR of the worst components from Table 4.2. This is above the average 20 dB signal separation, which implies that the limiting noise is the suppressed Gaussian noise.

Table 4.2: Measured Signal-to-Noise Ratio of System Components

System component	SNR (dB)
Photodetector	25
Microphone	25
Function Generator	40
Channel 1 HV Amp	39
Channel 2 HV Amp	39
Loop HV Amp 1	39
Loop HV Amp 2 and Inverter	39
Loop Adder	39
Ithaco Preamps	40

### 4.6.5 FM signal separation results

The experimental music recordings of signal separation convey independent component analysis to the person listening. The signal separation is measured at 17 to 20 dB. Qualitatively, this means that if the music and noise sound equally loud when mixed, after ICA, the noise is approximately 100 times quieter and cannot be heard. The recording time is generally under 1 minute before drift of the input channel gains causes reduced signal separation. The drift problem can be compensated for with a dynamic feedback gain control and is not a limiting factor.

## 4.7 Ones density binary signal separation

A second example of independent component analysis is separation of two digitally modulated voice signals. The modulation scheme that is chosen is a ones density modulation, the output of a sigma-delta analog-to-digital converter (ADC). Figure 4.19 shows a sinusoidal signal before and after ones density modulation. In this type of modulation, the signal is highly oversampled at many times the Nyquist frequency. The ones density bitstream contains more ones to represent a high voltage input, and more zeros to represent a low voltage input. The time average across many periods of a ones density bitstream is a representation of the original signal.

There are two types of sigma-delta ADCs that are commonly used: a time sampling version [52], and a free-running version [53], and both use a

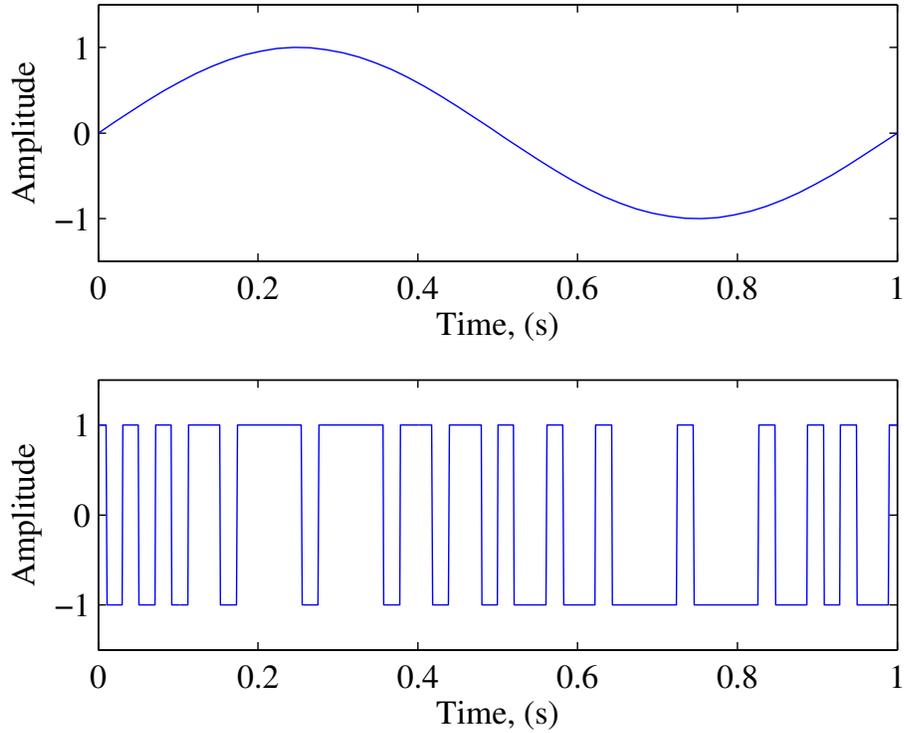


Figure 4.19: Ones density modulation. These plots represent a sinusoidal signal (upper plot) that is ones density modulated (lower plot). The average number of ones and zeros in the digital signal corresponds to the voltage of the original signal. The high-amplitude portions of the original signal are represented by more ones.

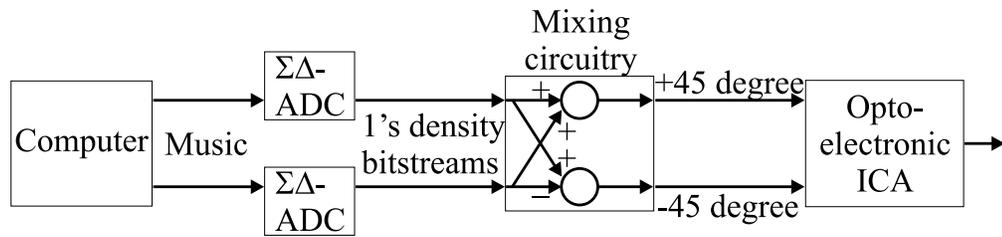
feedback loop to produce the output modulation. The time sampling version uses sample-and-hold circuits to generate a periodic bitstream. The effective number of bits in this version of ADC is determined by its order (feedback loop filter) and its oversampling rate [52]. The free-running sigma-delta ADC does not use a sample-and-hold circuit, and generates the output bitstream

at as high a frequency as allowed by the feedback loop. The actual output of this free-running version is a periodic square wave with duty cycle that is modulated by the analog input signal. Again, in this version, the free-running average of the bitstream represents the modulated signal.

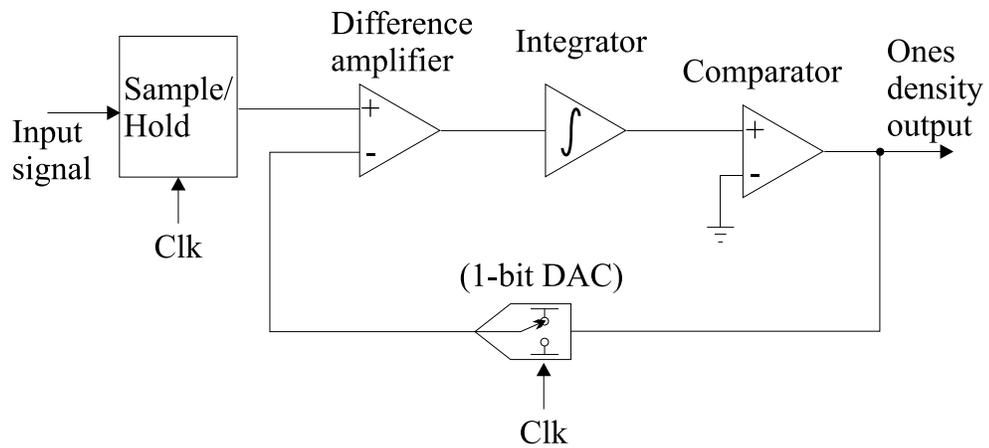
The experimental setup for binary signal separation is shown in Figure 4.20. The music signals are generated and output by the computer. These signals are input into two identical free-running sigma-delta modulators. The resulting pair of ones density bitstreams are mixed via an analog mixing circuit which consists of a single adder and subtracter. The only signal mixture that is achievable using this circuit is a 45 degree mixture of the two inputs. It is these two mixtures that are input into the opto-electronic ICA system. It should also be possible to generate and mix these signals entirely in the computer, although this experiment was not attempted.

### 4.7.1 Binary data probability density function

Like most modulation techniques, a ones density modulation changes the probability density function of the modulated signal. In order to be used in the system, the resulting density must be sub-Gaussian, as described in section 2.4.5. If the input signal has a zero mean, then its ones density modulation is composed of 50% ones and 50% zeros. In this case, the probability density is represented by two delta functions located at +1 and -1, each with an area of  $\frac{1}{2}$  (see binary pdf in Figure 2.7). Using Equation 2.38, the resulting kurtosis is 1 (or an excess kurtosis of -2), implying that it is a sub-Gaussian



(a)



(b)

Figure 4.20: Ones density modulation setup. (a) The music signals are generated via computer, but they are encoded and mixed with analog sigma-delta converters and mixing circuitry. The resulting 45-degree signal mixture is input into the ICA system. (b) The generalized sigma-delta analog-to-digital converter to generate ones density bitstream. The free-running modulator does not use the initial sample and hold circuit or clock input.

density.

### 4.7.2 Ones density demodulation

The demodulation of a ones density bitstream is just the time average of the bitstream itself. While this computation can be done in many ways, the simplest is to filter the bitstream with a low pass filter. The filter cutoff frequency must be set above the bandwidth of the modulated signal, but below the sampling frequency. For the sigma-delta ADCs used in the system, the sampling frequency was chosen to be anywhere between 20 kHz and approximately 70 kHz. The filter cutoff is chosen to be at approximately 8 kHz.

The low pass filter, design in ADS, is a 5th order elliptic filter, shown in Figure 4.21. The response of this type of filter is -40 dB/decade at the cutoff. A filter with such a sharp cutoff is desirable in the 100 kHz opto-electronic system simply because the sampling frequencies and signal bandwidths are very similar.

### 4.7.3 System noise measurements for binary signal separation

As in the FM signal experiments, the binary signal separation measurements contain background noise that is not added by the input signals. In a similar manner to the FM signal separation experiments, the input and output signal-to-noise ratio must be checked for each component of the system. In these experiments, it is found that there is a substantial amount of laser noise

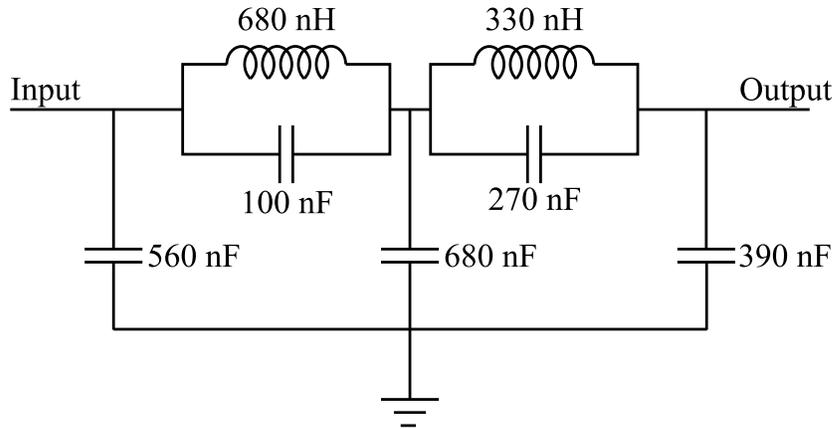


Figure 4.21: Circuit of fifth-order elliptic filter for ones density demodulation.

present. This noise is generally reduced by adding a differencing photodetector to the system. The differencing photodetector is described in section 3.4.3. The final signal-to-noise ratio of the system is on the order of 20-25 dB, a reasonable number assuming 20 dB signal separation.

#### 4.7.4 Binary signal separation results

The final recorded music using the binary ones density modulation scheme has between 15 and 20 dB signal separation. Again, there is a minor fluctuation in the separation value due to environmental drift, but recordings of approximately 1 minute can be made. There is a low 400 Hz noise that can be heard in the background of the output that has not been tracked down yet.

## 4.8 Summary of experimental results

This chapter has shown that the 100 kHz bandwidth opto-electronic ICA system follows the theory set out in chapter 2. The system achieves 20 dB signal separation for two signals provided one of them is sub-Gaussian. The feedback loop gain is dependent on the probability density function of the input signals, and the signal with the highest gain is the winner. If both signals have the same densities, and therefore the same gain, then the system randomly selects one of the two signals. Nearly constant signal separation for all input mixtures is observed, with variations due mainly to environmental drift. The system adaptation time is on the order of 1 to 2 seconds, and it is able to track one signal in a time-varying mixture of signals, provided the mixture does not change too rapidly. Communications signals, including frequency modulated signals and binary encoded signals have also been used for audio demonstrations of independent component analysis.

# Chapter 5

## Related and Future Work

### 5.1 Thesis contributions

The primary contributions of this thesis are as follows.

- Design and integration of the opto-electronic independent component analysis system (chapter 3). This design includes the layout of all optical components and supporting electronics.
- Design of input channels, including carrier suppression circuitry.
- Design of homodyne feedback loop to maintain the 90 degree phase shifted carrier and produce detectable amplitude modulation. The ICA feedback loop design is also my work.
- Design of ICA feedback loop.

- While the theory was pioneered by Dr. Dana Anderson, the complete system testing in order to verify the opto-electronic theory is a contribution of this thesis. This includes the open-loop gain measurements, and both sinusoidal and arbitrary probability density function signal competition. The testing work includes design of all of the supporting software, and the design of all the experiments.

## 5.2 Future work

Many communications signals, including television signals, cell phone CDMA signals, and radar signals, require a bandwidth of several megahertz. It is in this high bandwidth domain that the opto-electronic ICA processor has a true advantage over digital signal processors. This chapter investigates the next version of the system, which will be able to separate signals with a bandwidth of 20 MHz and a carrier of 200 MHz. Although the high-frequency version of the ICA system is not yet complete, the initial stages of the design have begun, but will not be examined in detail in this thesis. The limitation from the previous 100 kHz design were the electro-optic modulators and high voltage amplifiers. This chapter will discuss the design and implementation of electro-optic modulators for the high-frequency system. Not discussed is the initial designs for modularized sub-systems, including the carrier suppression sub-system and the homodyne detection sub-system.

Among the items that are also going to be implemented in the future of

the opto-electronic ICA project is the addition of an opto-electronic principal component analysis (PCA) system that performs the whitening preprocessing step. It is also important in the communications environment that both signals are recovered from the ICA system, instead of just the one that oscillates in the feedback loop. A multi-channel scheme is necessary to increase the number of signals from two to four. The multi-channel system requires that each signal be separated out of the initial input mixtures.

Future students must also examine the theory behind opto-electronic blind signal separation in greater depth. With the combination of PCA and ICA, it is possible to trace added noise through the system. This will provide a base noise floor estimate for signal separation. In addition, it may be possible to evaluate how rapidly the signals and mixtures can change in time, and the degradation of the signal separation. In addition, the theory behind an under-constrained system, with more signals than mixtures, may be evaluated.

## **5.3 20-MHz bandwidth electro-optic modulator**

### **5.3.1 EOM crystal design**

The bandwidth limitation in the 100-kHz opto-electronic ICA processor is due to the electro-optic modulators and the high-voltage amplifiers. The

modulator used in the main loop has a half-wave voltage of 280 volts, and a required drive voltage of approximately 700 volts (see section 3.4.1). It is this high drive voltage that is difficult to achieve with a wide bandwidth amplifier and it is necessary to lower the required EOM drive voltage.

The half-wave voltage of an electro-optic modulator is given by the equation [47]:

$$V_{\pi} = \frac{\lambda d}{\ell n_0^3 r_{33}} \quad (5.1)$$

where  $d$  is the thickness of the electro-optic crystal,  $\ell$  is its length (Figure 5.1),  $n_0$  is its index of refraction,  $r_{33}$  is the electro-optic coefficient, and  $\lambda$  is the free-space wavelength of the light. For magnesium doped lithium niobate,  $n_0 = 2.28$ , and  $r_{33} = 30.9 \times 10^{-12}$ . Equation 5.1 shows that the required voltage is directly proportional to the thickness of the crystal and inversely proportional to the length of the crystal. Ideally, for a lumped element modulator, the thinnest, longest crystal possible should be used to reduce the half-wave voltage.

The chosen crystal dimensions are  $d = 300 \mu\text{m}$  and  $\ell = 30 \text{ mm}$ . These dimensions lead to a half-wave voltage of 14.5 volts. From section 2.3.4, the drive voltage required for a harmonic signal on the loop modulator is therefore  $\pm 17.7$  volts. In addition, to save material, the crystal width is only 1 mm.

In optical considerations, a 532 nm Gaussian laser beam must be focused into the crystal without clipping the edges of the beam on either the input or output face. This can be done with a beam waist of 0.25 mm, and a Rayleigh

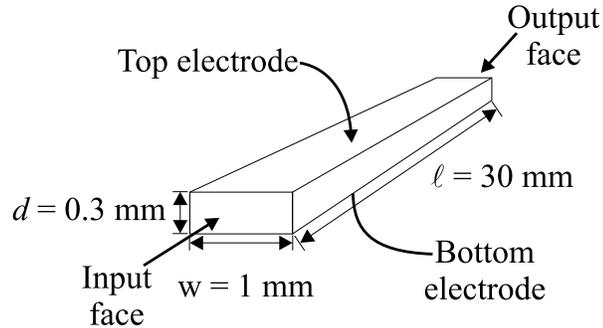


Figure 5.1: Electro-optic crystal dimensions. The electro-optic crystal of  $\text{LiNbO}_3$  is  $30 \times 1 \times 0.3 \text{ mm}$ . The input and output faces are anti-reflection coated at  $532 \text{ nm}$ . The electrodes are located on the top and bottom faces as depicted. The input light must be vertically polarized if the crystal is in illustrated orientation.

range of  $84 \text{ cm}$  in  $\text{LiNbO}_3$  (see [54] for Gaussian beam parameters).

It is interesting to note that the half-wave voltage is linearly proportional to the thickness of the crystal. Therefore, if the crystal is polished to  $0.33 \text{ mm}$ , or  $10\%$  thicker than expected, then the half-wave voltage also increases by  $10\%$ . Because power is proportional to the voltage squared, this implies that the driving amplifier needs to have nearly  $21\%$  higher power. If we specify that the crystal thickness tolerance will be within  $10\%$ , then an amplifier that can drive least  $21\%$  higher power than required for the ideal design should be chosen.

### 5.3.2 EOM matching circuit design

Most amplifiers that are designed to work with a center frequency of 200 MHz and a bandwidth of 20 MHz have an output impedance of  $50 \Omega$ . The EOM electrode is mainly capacitive with a small amount of series resistance due to the dielectric loss and input wire resistance.

The capacitance and resistance are found using Agilent's Advanced Design System (ADS) software. Using the geometry of the electrode from Figure 5.1, a loss tangent,  $\tan \delta$ , for  $\text{LiNbO}_3$  of 0.0013 [55], and the conductivity,  $\sigma$ , for copper of  $5.81 \times 10^7 \text{ S/m}$ , the series resistance is simulated to be  $0.036 \Omega$ . The simulation also accounts for field fringing at the edges of the electrode, giving a capacitance value of 56 pF. The series combination of capacitance and resistance can also be turned into a parallel combination of 56 pF and  $5.5 \text{ k}\Omega$ , which is used in the following analysis.

When designing a matching network for this load, the objective is to get as much voltage as possible across the electro-optic crystal with the smallest amount of input power. One method to increase the voltage across a component over a narrow bandwidth is to place it in a resonant circuit. Theoretically, with a completely lossless network, it is possible to increase the voltage by any factor desired. However, the trade off is that if the voltage amplification is higher, then the insertion loss is also higher and a lower voltage is applied to the resonator. Because the electrode impedance is also capacitive, the impedance match is narrow band.

Ideally, the reflection coefficient should be as small as possible to allow most of the signal to be transmitted over the desired bandwidth. The Bode-Fano criterion [56], [57], can be used to find the reflection coefficient for an arbitrary lossless matching network:

$$\int_0^\infty \ln \frac{1}{|\Gamma(\omega)|} d\omega < \frac{\pi}{RC} \quad (5.2)$$

where  $R$  and  $C$  are the resistance and capacitance respectively, and  $\Gamma(\omega)$  is the frequency dependent reflection coefficient. The minimum reflection coefficient,  $\Gamma_m$  can be found using the formula [58]:

$$\Delta\omega \ln \frac{1}{\Gamma_m} \leq \frac{\pi}{RC}. \quad (5.3)$$

where  $\Delta\omega$  is the bandwidth. Equation 5.3 implies that there is a trade off between a wider bandwidth and the minimum reflection coefficient. For the electrode impedance described above, with a bandwidth of 20 MHz, the minimum reflection coefficient that is achievable is  $\Gamma_m = 0.92$ . In other words, with a perfect matching network, 92% of the voltage will be reflected in order to achieve a constant match across 20 MHz.

The method for designing the a matching network is found in [59]. A 20 MHz bandpass filter is designed that contains the impedance of the electrode as part of its structure. One approximation that must be made is that the filter cannot have an infinite number of segments. Instead, a Chebychev bandpass filter with an order  $n = 2$  is used. Off-the-shelf components with values as close as possible to the actual design values are also used. In

the Chebychev filter design, if the load resistance is fixed ( $5.5 \text{ k}\Omega$  in this case) and the first capacitor of the filter is also fixed ( $56 \text{ pF}$  for the EOM), then all other components of the network are also already defined, including the source resistance (see Figure 5.2). The final filter design is shown in Figure 5.3a. The filter response is shown in Figure 5.3b. This filter has a 3-dB bandwidth of 49 MHz. The source resistance,  $R_s$ , is actually  $50 \text{ }\Omega$ , not the required  $103 \text{ }\Omega$  required by the filter design, and therefore must be transformed. The impedance transform is also non-ideal, and consists of a two-stage ladder network of capacitors and inductors, shown in Figure 5.4. The bandwidth of this segment is over 400 MHz, which is larger than the filter bandwidth.

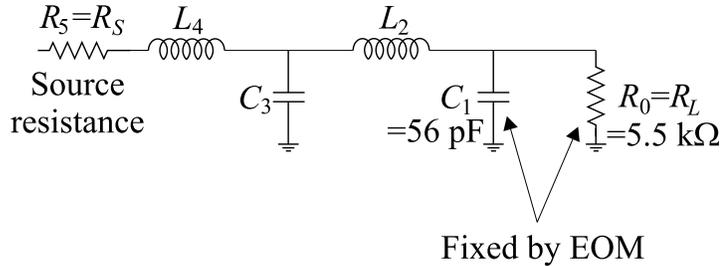


Figure 5.2: Fourth order low pass Chebychev filter. In a Chebychev filter, two parameters can be chosen. In this case,  $R_0$  and  $C_1$  are given by the electro-optic modulator impedance. All other parameters are then set by the Chebychev design.

The final lossless matching network is shown in Figure 5.5. As can be seen in Figure 5.5b and c, this filter response is not perfectly flat, or perfectly sharp. The voltage gain of this circuit is approximately 1.45, and the

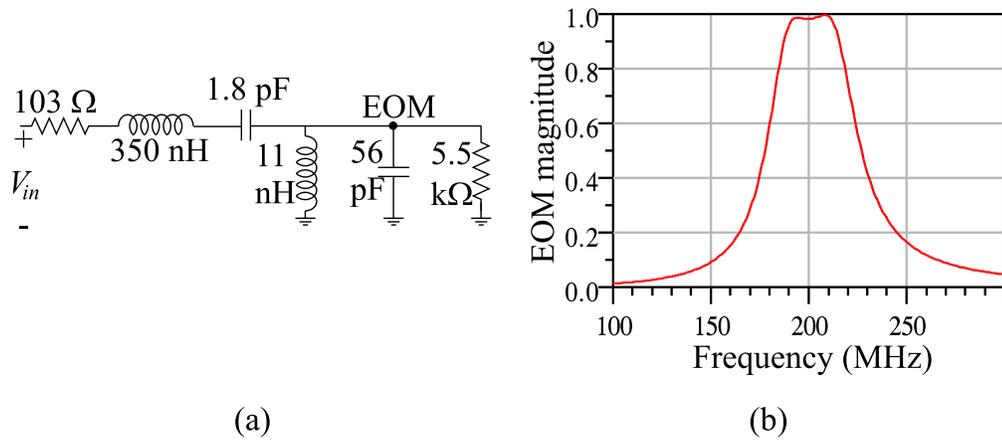


Figure 5.3: Second order bandpass Chebyshev filter. The actual Chebyshev bandpass filter design has a 3 dB bandwidth of 49 MHz.

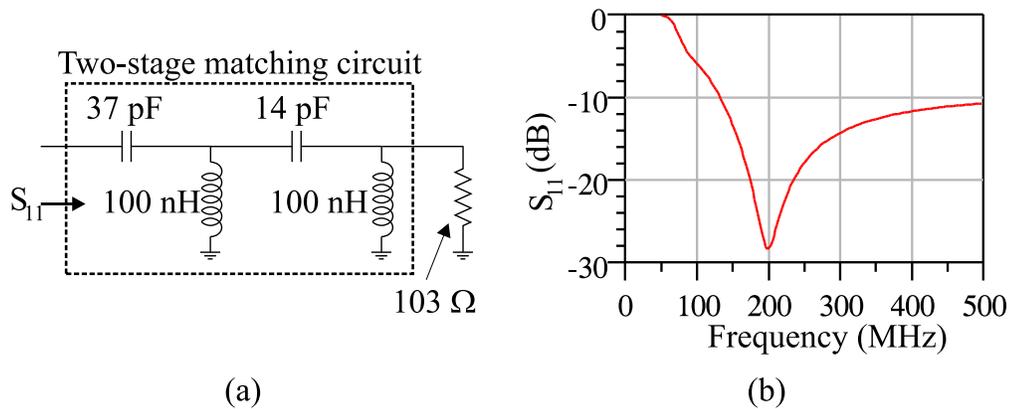


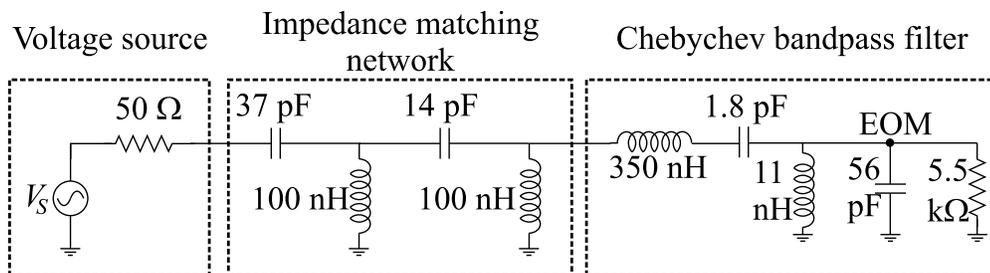
Figure 5.4: Two-stage impedance transform circuit. (a) The ladder network shown here has a 3-dB bandwidth of over 400 MHz. (b) Simulated network frequency response.

reflection coefficient is approximately 0.96, higher than the ideal 0.92. The discrepancy is caused because the filter is only a second order design. The design widens the bandwidth of the filter from the design parameter of 20

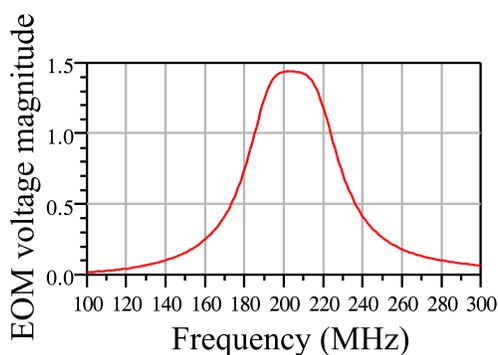
MHz to nearly 50 MHz and decreases the passband transmission. The lossless matching circuit shown in Figure 5.5 is also relatively large as it contains 7 reactive elements, including 4 inductors, which occupy approximately 3 mm by 3 mm of circuit space. In addition, this circuit is not tolerant to minor changes in element values. A 10% change in the electro-optic capacitance causes a 25 MHz shift of the impedance match center frequency. The center frequency can be brought back to 200 MHz by tuning the two filter inductors and capacitor. However, tunable inductors are generally low-frequency components that take up nearly 7 mm<sup>2</sup> of circuit space. It is this reason that the lossless matching circuit design is unused in the final electro-optic modulator.

A second way to make the matching circuit is to add a small series resistance or large parallel resistance to the electrode structure. This method adds loss to the matching network. A wide bandwidth impedance match is, in principle, easier to design with more resistance in the electrode structure. The trade off is that the circuit draws more power because of the added resistance. Ideally, the circuit should reach an impedance match at least as good as was achieved by the Chebyshev network, but still consume a small enough amount of power for a common amplifier to supply. (This number will be defined later.)

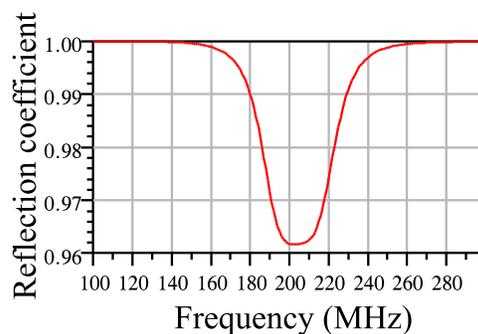
The circuit chosen is shown in Figure 5.6. This circuit is a two stage ladder network impedance match to a load that contains a 150  $\Omega$  resistor in parallel with the modulator electrode. It is designed to operate at a center



(a)



(b)



(c)

Figure 5.5: Lossless EOM impedance matching circuit. (a) The lossless impedance matching circuit contains seven reactive components to achieve a 48 MHz 3 dB bandwidth with a voltage gain of 1.4. (b) Modulator voltage response curve for a normalized source of 1 volt. (c) Magnitude of reflection coefficient. In the ideal case with an infinite number of stages and the precise component values, the reflection coefficient would be 0.92.

frequency of 200 MHz, and with a bandwidth of 20 MHz. The elements are chosen with values that are easily purchased. The reflection coefficient (simulated in ADS) is shown in Figure 5.7. This matching circuit is a two-stage match, which serves to make the bandwidth approximately 25% wider than its one-stage counterpart.

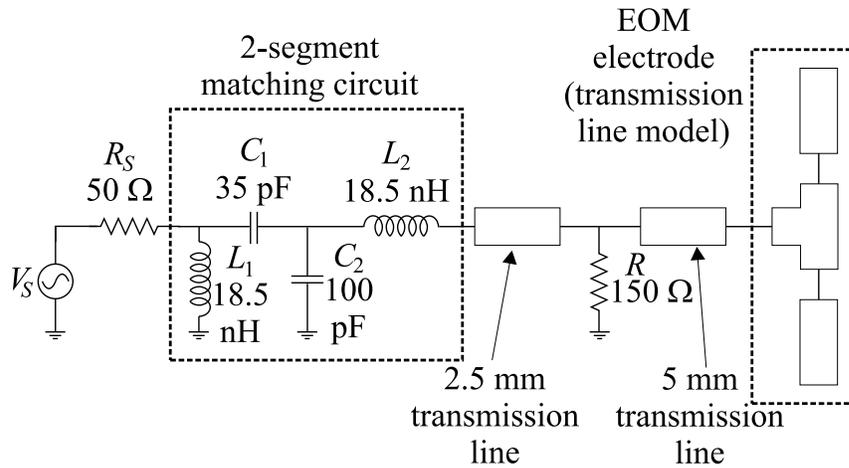


Figure 5.6: Lossy EOM impedance matching circuit. The impedance matching circuit shown here adds a 150  $\Omega$  resistor in parallel with the EOM electrode. The added loss increases the required power, but widens the matching bandwidth.

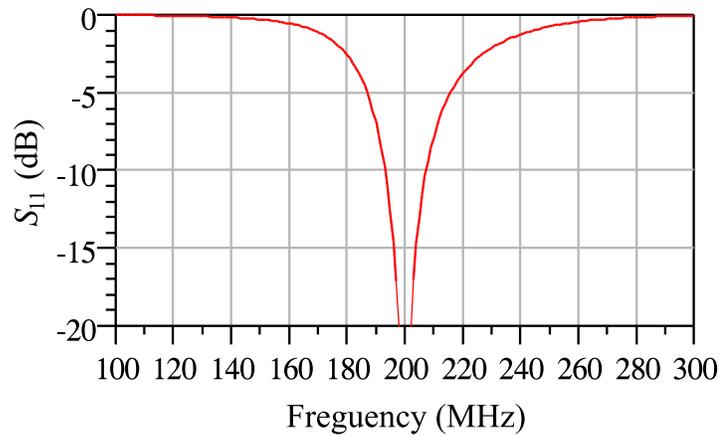


Figure 5.7: Lossy EOM impedance matching circuit frequency response. The reflection coefficient ( $S_{11}$ ) of the impedance matching circuit shown in Figure 5.6 has a center frequency of 200 MHz and a 3-dB bandwidth of 41 MHz.

The amount of power drawn by this circuit can also be found. If the voltage across the electrode is set to the required 17.7 volts, at 200 MHz, the amount of current drawn is 190 mA. Both the voltage and current plots are shown in Figure 5.8. This implies that the circuit amplifier must output 1.8 watts. In addition, all components must be able to handle the specified current. The resistor has a 25 W capability, and the inductors have 2.4 A capacity.

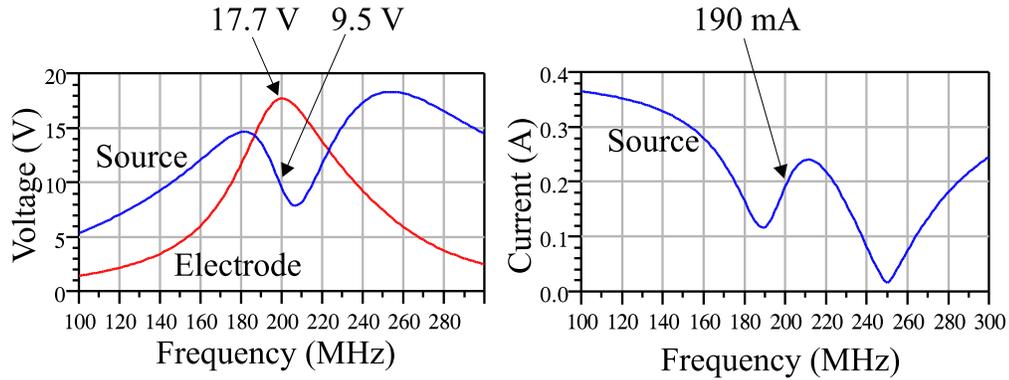


Figure 5.8: Lossy EOM impedance matching circuit voltage and current responses. To achieve 17.7 volts at the EOM electrode, the source must have an output of 9.5 V (zero-to-peak) and 190 mA current. This implies a drive power of 1.8 W.

While the match is adequate at 200 MHz, it is necessary to realize that the element values are not going to be exactly accurate. To compensate for these inaccuracies, small, variable capacitors are placed in parallel with the original capacitors. The average capacitance of each pair adds up to the capacitance shown in Figure 5.6. Simulation results show that by varying

these two capacitances, it is possible to completely compensate for up to 10% error in all the other component values. Therefore, this circuit is tolerant to error in component values.

### 5.3.3 EOM mount design

The mount of the electro-optic modulator should be as small as possible given several constraints on the design. First, the mount must hold the electro-optic crystal firmly without any added stress because the crystals are extremely fragile. Second, there must be a good electrical ground on the lower crystal electrode. Third, the crystal must be located close to the edge of the mount in order to reduce the size of the carrier suppression module (see section 5.3). Fourth, the crystal should be thermally isolated from the heat produced by the added resistor as the modulator is being used. Lastly, connecting and reconnecting the source to the EOM should not add any stress to the mount.

The final mount design is shown in figure Figure 5.9. The mount consists of three brass pieces: a base, a back wall, and a slider. The crystal rests on the base of the mount that acts as electrical ground. This piece is the only piece of brass that the crystal will be touching, and is lapped to a  $9\ \mu\text{m}$  surface roughness. In addition, the brass base is gold coated to minimize the amount of oxidation. The upper electrode is made from a piece of gold coated silicon that is cut from a  $500\ \mu\text{m}$  thick wafer. The wafer is also optically flat and does not put any stress on the crystal. The silicon electrode is gold

coated on all four sides, and the electrode contact is made by a ribbon piece of copper that had been silver epoxied to the top electrode surface. On top of the stack is a piece of Sorbothane rubber that is also designed to minimize the stress on the crystal. The entire stack is pressed down by the slider piece of brass. The slider consists of two pegs that hold it horizontal while being moved up and down, and relieve the amount of rotational stress caused by tightening the screws that lock it in place.

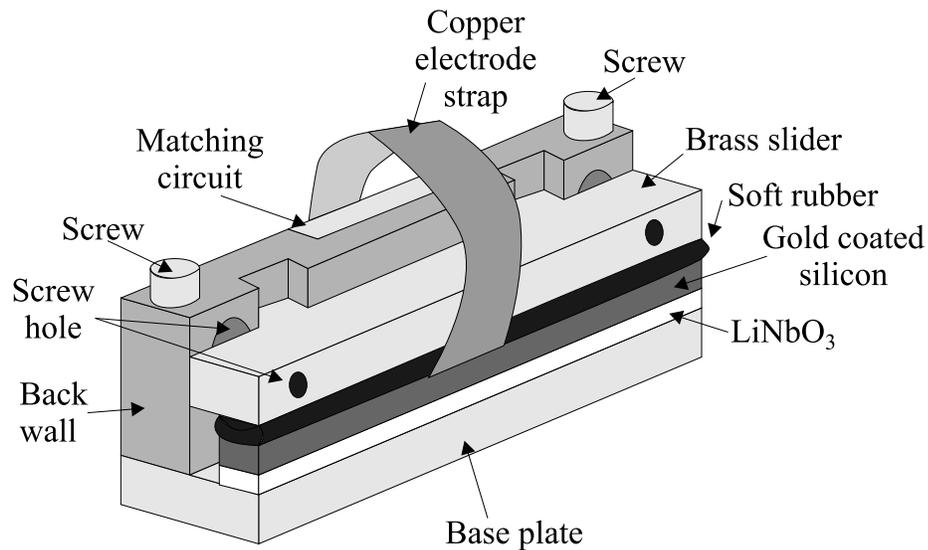


Figure 5.9: Electro-optic modulator mount design. The electro-optic modulator mount is 30 mm long, 4 mm wide, and 7 mm tall. It is made from three brass pieces, and the matching circuit sits against the back wall.

The matching circuit is placed on the back of the mount sidewall. The layout of the matching circuit is shown in figure Figure 5.10. The circuit is 7.3 mm long by 7.5 mm wide and made on an FR4 substrate for strength.

The resistor is a 25 Watt resistor, which is completely removed from the mount except for two connecting wires. These wires add inductance to the design, but this is calibrated out with the variable capacitors. The resistor sits on top of a ceramic post with a miniature heat sink.

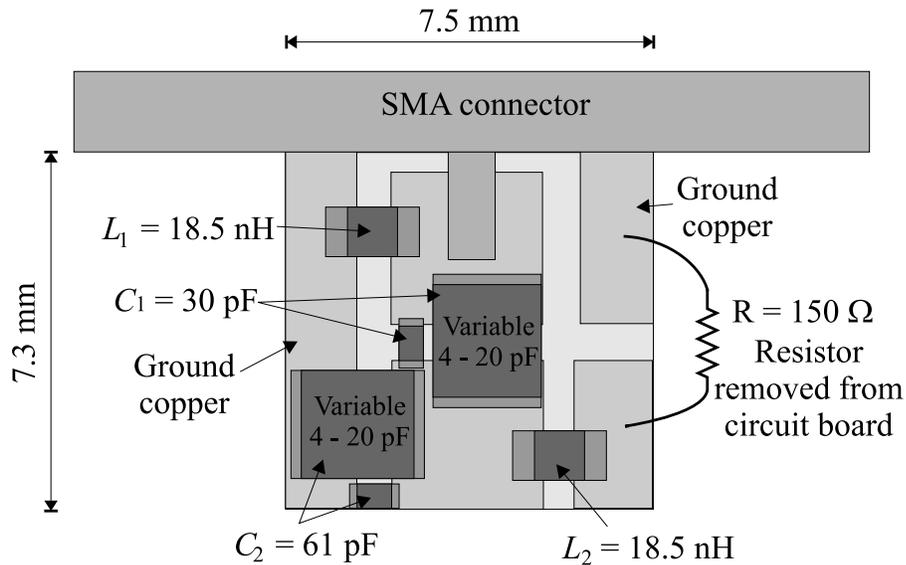


Figure 5.10: EOM matching circuit layout. The matching circuit is 7.3 mm  $\times$  7.5 mm. The resistor is removed from the circuit to keep excess heat away from the  $\text{LiNbO}_3$  crystal.

The SMA connector is not directly touching the mount, but rather connected by three short wires, one to the center connector, and two to the outside ground. The wires remove the mechanical stress caused by connecting and disconnecting the circuit.

## 5.4 Summary

This thesis describes an opto-electronic blind source separation system that is capable of separating an unknown mixture of two completely unknown signals. The separation technique that the system uses is similar to independent component analysis in that it uses higher order statistics to separate the signals. The system achieves a signal separation of over 20 dB provided that at least one of the signals is sub-Gaussian.

Chapter 2 describes the theory underlying opto-electronic ICA. It begins with the case of two sinusoidal signals and shows that the dynamics of the system are such that only one signal is able to oscillate in the opto-electronic feedback loop. The analysis is extended to two arbitrary probability density function signals, and shows that the system gain is dependent on their joint characteristic function.

Chapter 3 is a description of the opto-electronic ICA system. The input beams are phase modulated with the two mixtures and carrier suppressed. The ICA feedback loop contains another electro-optic phase modulator that is over-driven to produce signal harmonics, and a photorefractive crystal that amplifies the correlated portion of the loop signal with the input mixtures. The phase modulated signal is mixed with a 90 degree shifted carrier to convert it into a detectable amplitude signal. The 90 degree phase shift is maintained by a feedback control loop that sets the voltage on a piezoelectric movable mirror. The system performance is measured by correlating

the output signal with the desired unmixed original signals.

Chapter 4 provides a description of the experimental data used to verify the theory presented in chapter 2. The open-loop gain is measured, and signal competition between both sinusoidal and arbitrary PDF signals is quantified. In addition, the system performance from initial conditions to steady state is measured in real time, and the signal mixtures are allowed to vary with time while the system tracks one component. The system is demonstrated using both frequency modulated and ones density encoded signals.

A 20 MHz bandwidth opto-electronic ICA system requires different components than the previously demonstrated 100 kHz design. Among these components are high-bandwidth electro-optic modulators and driving amplifiers. The modulators must have a low driving voltage to achieve the desired bandwidth, leading to a design of thin, long crystals. The driving amplifiers require an input impedance of  $50 \Omega$ , which necessitates broadband matching circuits on the modulators.

In addition, a high-bandwidth system should use modularized circuits, including the carrier suppression circuit. Modularized subsystems will facilitate the building of the system for the next version, the multi-channel system. These sub-systems must be as small as possible to reduce the environmental drift in the interferometry.

Different homodyne detection schemes can be used in the high-bandwidth system as well. This reduces the need for a high-speed phase locked loop in the piezo controller feedback loop. In addition, it removes the piezo controller

error signal from the input signal bandwidth.

The next steps in the life of the opto-electronic ICA project are well defined. In order to be viable in a real world communications scenario, the system must be a complete blind signal separation processor. This system must include a PCA processor for whitening. Secondly, both of the input signals must be recoverable, implying the design of a component removal system. It is desirable to increase the number of communication channels from two to at least four to demonstrate the scalability of the system. The next step in increasing the signal bandwidth is to move to several hundred megahertz. A jump of this magnitude requires a new design for the electro-optic modulators. However, modulators with this bandwidth exist for longer wavelength light and there is no reason that they could not be made at 532 nm as well. The feedback loop must be made smaller in order to utilize the bandwidth of this type of modulator.

# Bibliography

- [1] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent component analysis*, Wiley, New York, NY, 2001. [1](#), [6](#), [7](#), [155](#)
  
- [2] G. Cauwenberghs, Milutin Stanacevic, and G. Zweig, “Blind broadband source localization and separation in miniture sensor arrays,” *ICA 2001*, 2001. [1](#)
  
- [3] M. H. Cohen and A. G. Andreou, “Analog CMOS integration and experimentation with an autoadaptive independent component analyzer,” *IEEE Transactions on Circuits and Systems - II: Analog and Digital Signal Processing*, vol. 42, no. 2, pp. 65–77, 1995. [1](#)
  
- [4] S. Haykin, *Unsupervised Adaptive Filtering, Volume 1: Blind Source Separation*, Wiley, New York, NY, 2000. [2](#)
  
- [5] A. Mansour and M. Kawamoto, “Ica papers classified according to their applications and performances,” *IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences*, vol. E86-A, no. 3, pp. 620–633, 2003. [3](#)

- [6] A. K. Barros F. Itakura, T. Rutkowski, A. Mansour, and N. Ohonishi, “Estimation of speech embedded in a reverberant environment with multiple sources of noise,” *ICA 2001*, 2001. 3
  
- [7] Davies ME Mitianoudis N, “Audio source separation of convolutive mixtures,” *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 5, pp. 489–497, 2003. 3
  
- [8] T. Nishikawa, H. Saruwatari, and K. Shikano, “Blind source separation of acoustic signals based on multistage ICA combining frequency-domain ICA and time-domain ICA,” *IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences*, vol. E86-A, no. 4, pp. 846–858, 2003. 3
  
- [9] J. Anemuller and B. Kollmeier, “Adaptive separation of acoustic sources for anechoic conditions: A constrained frequency domain approach,” *Speech Communications*, vol. 39, pp. 79–95, 2003. 3
  
- [10] V. Zarzoso and A.K. Nandi AK, “Noninvasive fetal electrocardiogram extraction: Blind separation versus adaptive noise cancellation,” *IEEE Trans. On Biomedical Imaging*, vol. 48, no. 1, pp. 12–18, 2001. 3
  
- [11] P. G. Calhoun V and T. Adali, “Independent component analysis applied to fMRI data: A generative model for validating results,” *Journal of VLSI Signal Processing Systems for Signal Image and Video Technology*, vol. 37, no. 2–3, pp. 281–291, 2004. 3
  
- [12] M. Haritopoulos, H. Yin, and N. M. Allinson, “Image denoising using self-organized map-based nonlinear independent component analysis,” *Neural Networks*, vol. 15, no. Special Issue, pp. 1085–1098, 2002. 3

- [13] J. Delabrouille, J.-F. Cardoso, and G. Patanchon, “Multidetector multicomponent spectral matching and applications for cosmic microwave background data analysis,” *Monthly Notices of the Royal Astronomical Society*, vol. 346, no. 4, pp. 1089–1102, 2003. 3
- [14] D. Manio, A. Farusi, C. Baccigalupi, F. Perrotta, A.J. Banday, L. Bedini, C. Burigana, G. De Zotti, K.M. Gorshi, and E. Salerno, “All-sky astrophysical component separation with fast independent component analysis (FASTICA),” *Monthly Notice of the Royal Astronomical Society*, vol. 334, pp. 53–68, 2002. 3
- [15] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, “Face recognition by independent component analysis,” *IEEE Transactions on Neural Networks*, vol. 13, no. 6, 2002. 3
- [16] X.B. Yu and G. Baciù, “Face recognition from color images in presence of dynamic orientations and illumination conditions,” *Biometric Authentication, Proceedings Lecture Notes in Computer Science*, vol. 3072, pp. 227–233, 2004. 3
- [17] J. Kim, J. Choi, and J. Yi, “Face recognition based on locally salient ICA information,” *Biometric Authentication, Proceedings Lecture Notes in Computer Science*, vol. 3087, pp. 1–9, 2004. 3
- [18] D.S. Lee, J.Y. Choi, and D.H. Choi, “Ica based thermal source extraction and thermal distortion compensation method for a machine tool,” *International Journal of Machine Tools and Manufacture*, vol. 43, pp. 589–597, 2003. 3
- [19] A. E. Capobianco, “Hammerstein system representation of finan-

- cial volatility processes hammerstein system representation of financial volatility processes,” *The European Physical Journal B*, vol. 27, pp. 201–211, 2002. 3
- [20] T. Ristaniemi T and J. Joutsensalo, “Advanced ica-based receivers for block fading DS-CDMA channels,” *Signal Processing*, vol. 82, pp. 417–431, 2002. 4
- [21] T.Y. Yang and W.B. Mikhael, “A general approach for image and co-channel interference suppression in diversity wireless receivers employing ICA,” *Circuits Systems and Signal Processing*, vol. 23, no. 4, pp. 317–327, 2004. 4
- [22] I. Kostanic and W. Mikhael, “Independent component analysis based QAM receiver,” *Digital Signal Processing*, vol. 14, no. 3, pp. 241–252, 2004. 4
- [23] T. Ristaniemi T and J. Joutsensalo, “Independent component analysis with code information utilization in DS-CDMA signal separation,” *Global Telecommunications Conference - Globecom’99*, 1999. 4
- [24] A. Celik, M. Stanacevic, and G. Cauwenberghs, “Mixed-signal real-time adaptive blind source separation,” *Proc. IEEE Int. Symp. Circuits and Systems*, May 2004. 4
- [25] A.B. Lim, J.C. Rajapakse, and A.R. Omondi, “Comparative study of implementting ICNNs on FPGAs,” *ICA 2001*, pp. 177–182, 2001. 4
- [26] K. Cho and S. Lee, “Offset-tolerant design of analog chips for indepen-

- dent component analysis,” *IEICE Trans. on Electronics*, vol. E87C, no. 8, pp. 1382–1387, 2004. [4](#)
- [27] G. Erten and F.M. Salam, “Voice extraction by on-line signal separation and recovery,” *IEEE Trans. On Circuits and Systems II: Analog and Digital Signal Processing*, vol. 46, no. 7, pp. 915–922, 1999. [4](#)
- [28] P. Yeh, *Introduction to Photorefractive Nonlinear Optics*, Wiley-Interscience, New York, NY, 1993. [5](#)
- [29] J.-F. Cardoso, “Blind signal separation: Statistical principles,” *Proceedings of the IEEE*, vol. 86, no. 10, pp. 2009–2025, 1998. [7](#)
- [30] J.-F. Cardoso, “On the stability of source separation algorithms,” *Journal of VLSI Signal Processing Systems for Signal Image and Video Technology*, vol. 26, no. 1–2, pp. 7–14, 2000. [7](#), [14](#)
- [31] A. Hyvarinen and E. Oja, “Independent component analysis: algorithms and applications,” *Neural Networks*, vol. 13, pp. 411–430, 2000. [7](#), [14](#)
- [32] J.-F. Cardoso, “High-order contrasts for independent component analysis,” *Neural Computation*, vol. 11, no. 1, pp. 157–192, 1999. [7](#)
- [33] J.V. Stone, *Independent component analysis, a tutorial introduction*, MIT Press, Cambridge, Massachusetts, 2004. [7](#), [155](#)
- [34] J.-F. Cardoso, “Infomax and maximum likelihood for blind source sep-

- aration,” *IEEE Signal Processing Letters*, vol. 4, no. 4, pp. 112–114, 1997. [7](#)
- [35] X. Giannakopoulos, J. Karhunen, and E. Oja, “A comparison of neural ICA algorithms using real world data,” *ICA 1999*, pp. 888–893, 2001. [7](#)
- [36] A. Cichocki and R. Thawonmas, “On-line algorithm for blind signal extraction of arbitrarily distributed, but temporally correlated sources using second order statistics,” *Neural Processing Letters*, vol. 12, no. 1, pp. 91–98, 2000. [7](#)
- [37] I.T. Jolliffe, *Principal component analysis*, Springer, New York, NY, 2002. [7](#)
- [38] D. Z. Anderson and E. Fotheringham, “A miniature photorefractive circuit for principal component extraction,” *Applied Optics-IP*, vol. 42, no. 23, pp. 4736–4746, 2003. [8](#), [9](#)
- [39] Edeline Fotheringham, *Microwave signal processing with photorefractive dynamic holography*, Ph.D. thesis, University of Colorado at Boulder, Boulder, CO, 2003. [8](#)
- [40] D. Z. Anderson, V. Damiao, D. Popovic, Z. Popovic, S. Romisch, and A. Sullivan, “-70 db optical carrier suppression by two-beam coupling in photorefractive media,” *Applied Physics B*, vol. 72, pp. 743–748, 2001. [11](#), [52](#), [60](#)

- [41] J.-F. Cardoso and P. Comon, “Independent component analysis, a survey of some algebraic methods,” *ICA 2001*, 2001. [14](#), [155](#)
- [42] D.Z. Anderson, “Holographic scheme for independent component analysis,” *Journal of Applied Physics*, vol. 95, no. 7, pp. 3272–3278, 2004. [16](#)
- [43] D. Z. Anderson, R. W. Brockett, and N. Nuttall, “Information dynamics of photorefractive two-beam coupling,” *Physical Review Letters*, vol. 82, no. 7, pp. 1418–1421, 1999. [29](#)
- [44] A. Leon-Garcia, *Probability and random processes for electrical engineering*, Addison-Wesley, Reading, Massachusetts, 1994. [31](#)
- [45] C. Benkert and D.Z. Anderson, “Controlled competitive dynamics in a photorefractive ring oscillator: “winner-takes-all” and the “voting-paradox” dynamics,” *Physical Review A*, vol. 44, no. 7, pp. 4633–4638, 1991. [35](#)
- [46] R.E. Thomas and A.J. Rosa, *The analysis and design of linear circuits*, Prentice Hall, Englewood Cliffs, NJ, 1994. [57](#)
- [47] A. Yariv, *Optical electronics in modern communications*, Oxford University Press, New York, NY, 1997. [59](#), [128](#)
- [48] E. Fotheringham, S. Romisch, P.C. Smith, D. Popovic D, D.Z. Anderson, and Z. Popovic, “A lens array with adaptive optical processing,” *IEEE Transactions on Antennas and Propagation*, vol. 50, no. 5, pp. 607–617, 2002. [60](#)

- [49] D.T. Pham and J.-F. Cardoso, “Blind separation of instantaneous mixtures of nonstationary sources,” *IEEE Transactions on Signal Processing*, vol. 49, no. 9, pp. 1837–1848, 2001. [105](#)
- [50] R. Everson and S.J. Roberts, “Non-stationary independent component analysis,” *Artificial Neural Networks*, vol. 1, pp. 503–508, Sept. 1999. [105](#)
- [51] J. Dunlop and D.G. Smith, *Telecommunications engineering*, Van Nostrand Reinhold (UK) Co, Ltd., Berkshire, England, 1984. [112](#)
- [52] J.C. Candy and G.C. Temes, *Oversampled delta-sigma data converters, theory, design and simulation*, IEEE Press, New York, NY, 1992. [118](#), [119](#)
- [53] James A. Cherry and W. Martin Snelgrove, *Continuous-time delta-sigma modulators for high-speed A/D conversion : theory, practice and fundamental performance limits*, Kluwer Academic Pub., Boston, MA, 2000. [118](#)
- [54] B.E.A. Saleh and M.C. Teich, *Fundamentals of photonics*, Wiley, New York, NY, 1991. [129](#)
- [55] K.K. Wong, *Properties of Lithium Niobate*, IEE/INSPEC, London, 2002. [130](#)
- [56] H.W. Bode, *Network analysis and amplifier feedback design*, Van Nostrand, New York, NY, 1945. [131](#)

- [57] W. Chen, *Broadband Matching Theory and Implementations*, World Scientific, Singapore, 1988. [131](#)
- [58] D.M. Pozar, *Microwave Engineering*, Wiley, New York, NY, 2005. [131](#)
- [59] G.L. Matthaei, L. Young, and E.M.T. Jones, *Microwave filters, impedance-matching networks, and coupling structures*, McGraw-Hill, New York, NY, 1964. [131](#)
- [60] P. Comon, “Independent component analysis: A new concept?,” *Signal Processing*, vol. 36, no. 3, pp. 287–314, 1994. [154](#)
- [61] C. Jutten and J. Herault, “Blind separation of sources, part i: An adaptive algorithm based on neuromimetic architecture,” *Signal Processing*, vol. 24, pp. 1–10, 1991. [154](#)
- [62] Y. Liu and W. Mikhael, “Practical frequency domain ICA algorithm capable of solving permutation and gain ambiguities for digital communication system,” *Electronics Letters*, vol. 40, no. 13, pp. 839–840, 2004. [154](#)
- [63] A. Mansour, A.K. Barros, and N. Ohnishi, “Blind separation of sources: Methods, assumptions and applications,” *IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences*, vol. E83-A, no. 8, pp. 1498–1512, 2000. [154](#)
- [64] H. Saruwatari, S. Kurita, and K. Takeda, “Blind source separation combining frequency-domain ICA and beamforming,” *ICA 2001*, pp. 2733–2736, 2001. [154](#)

- [65] H. Sawada, R. Mukai, S. Araki, and S. Makino, “A robust and precise method for solving the permutation problem of frequency-domain blind source separation,” *IEEE Transactions on Speech and Audio Processing*, vol. 12, no. 5, pp. 530–538, 2004. [154](#)
- [66] T. Nishikawa, H. Abe, H. Saruwatari, K. Shikano, and A. Kaminuma, “Overdetermined blind separation for real convolutive mixtures of speech based on multistage ICA using subarray processing,” *IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences*, vol. E87-A, no. 8, pp. 1924–1932, 2004. [155](#)
- [67] T. Yamaguchi and K. Itoh, “An algebraic solution to independent component analysis,” *Optics Communications*, vol. 178, pp. 59–64, 2000. [155](#)
- [68] S.A. Cruces-Alvarez, A. Cichocki, and S.I. Amari, “From blind signal extraction to blind instantaneous signal separation: Criteria, algorithms, and stability,” *IEEE Transactions on Neural Networks*, vol. 15, no. 4, pp. 859–873, July 2004. [156](#)

# Appendix A

## Fundamentals of ICA

Independent component analysis (ICA) is a powerful processing technique that uses the assumption of signal independence to discover the unknown signals within a mixture. Although the rigorous method for was created by Comon, in [60], it was originally developed as a back-propagating neural network by Herault and Jutten [61]. As described in chapter 1, independence is defined as:

$$\langle s_1^a s_2^b \rangle = \langle s_1^a \rangle \langle s_2^b \rangle \quad (\text{A.1})$$

where  $a$  and  $b$  are integers, and  $s_1$  and  $s_2$  are the signals, and the angular brackets represent the time average of the signal. While  $s_1$  and  $s_2$  are described as functions of time in this thesis, they can also be functions of space or of frequency [62], [63], [64], [65]. Using the two-signal problem statement

defined in chapter 1:

$$\begin{bmatrix} \tilde{s}_1(t) \\ \tilde{s}_2(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix}. \quad (\text{A.2})$$

where  $\tilde{s}_1$  and  $\tilde{s}_2$  are the mixtures and  $a_{mn}$  are the mixing coefficients. More mixtures than coefficients is known as the over-determined ICA problem [66], and is not addressed in this thesis. In general, this system represents two equations and six unknowns. On the most fundamental level, ICA is a technique to generate four more equations based on the assumption of independence [41], [67]:

$$\langle s_1 s_2 \rangle = \langle s_1 \rangle \langle s_2 \rangle \quad (\text{A.3})$$

$$\langle s_1^1 s_2^3 \rangle = \langle s_1^1 \rangle \langle s_2^3 \rangle \quad (\text{A.4})$$

$$\langle s_1^2 s_2^2 \rangle = \langle s_1^2 \rangle \langle s_2^2 \rangle \quad (\text{A.5})$$

$$\langle s_1^3 s_2^1 \rangle = \langle s_1^3 \rangle \langle s_2^1 \rangle \quad (\text{A.6})$$

While these correlation equations provide an intuitive understanding to the solution of the blind source separation problem, they are difficult to simultaneously solve for arbitrary signals. Instead, different measures of signal independence that are based on one quantity are used. These measures include *kurtosis* (section 2.4.4), *maximum likelihood estimation*, *entropy*, and *mutual information* [1], [33] to name a few.

The traditional ICA algorithm assumes that there exists a new mixing matrix,  $\mathbf{B}$ , that can be applied to the signal mixtures in order to find the original signals:

$$\begin{bmatrix} s_1(t) \\ s_2(t) \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \tilde{s}_1(t) \\ \tilde{s}_2(t) \end{bmatrix}. \quad (\text{A.7})$$

This assumption is valid if the mixing matrix,  $\mathbf{A}$ , is invertible.

Algorithmically, the coefficients of matrix  $\mathbf{B}$  are randomly chosen, and the independence of the resulting signals is calculated using one of the methods suggested above. The algorithm is then required to scan through the possible values of  $\mathbf{B}$  until the independence is maximized. For a more efficient algorithm, the gradient of the independence measure chosen is calculated for each of the weights,  $\mathbf{B}$ . The gradient provides a method for estimating new values of the unmixing matrix in each iteration of the algorithm. This method is known as the *gradient ascent* approach [68].

A useful illustration of the ICA algorithm applied to two signals is shown in Figure A.1. This figure shows two depictions of each step of the algorithm. A time domain representation of the two signals is shown in the first two rows of this figure. The third row of the figure shows signal space plots for each step of the process. In a signal space plot, the axes represent the two channels of the system. A channel is one of the input pathways to the system. The signals are represented as the vectors in the diagram. The projection of a signal vector onto a channel axis represents the amount of that signal present in the mixture on the channel. In this depiction, each vector lying along one

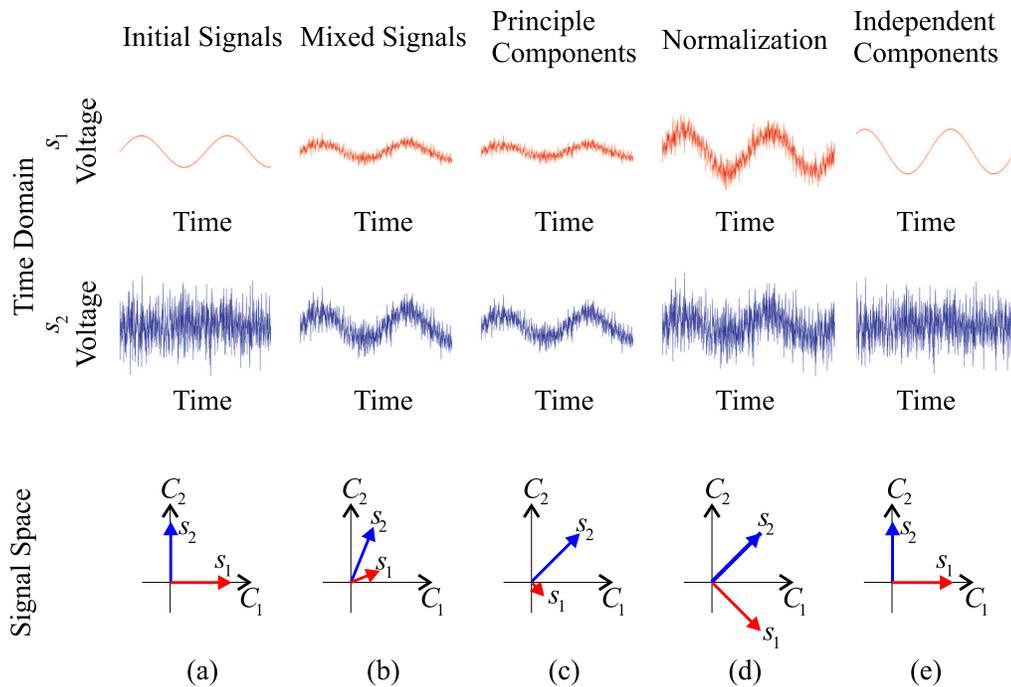


Figure A.1: ICA procedure. This graphical representation demonstrates the complete procedure of ICA in both the time domain and the signal space domain. Two initial time-varying signals (a) are arbitrarily mixed together (b). The eigenvectors of the input signals are found in (c), and normalized to have the same variance (d). The final ICA procedure, a rotation, returns the signal to its initial state (e).

of the channel axes represents separated signals.

Consider two independent signals: a sine wave,  $s_1$ , and a random signal,  $s_2$ , with a Gaussian probability density function, as depicted in Figure A.1(a). These signals, initially separated, each exist on only one channel, shown in the signal space diagram. Without loss of generality, the signals are assumed to have zero mean.

The signals are mixed by a random matrix in Figure A.1(b), resulting in the two system inputs. In the example shown here, the arbitrary real matrix used is:

$$\mathbf{A} = \begin{bmatrix} 0.5 & 0.1 \\ 0.8 & 0.4 \end{bmatrix}. \quad (\text{A.8})$$

The first step to descramble the two signals, shown in Figure A.1(c), is known as principle component analysis (PCA). From the initial mixtures, two new uncorrelated mixtures are chosen. Mathematically, the time average of the product of the two mixtures is zero:

$$\langle \tilde{s}_1 \tilde{s}_2 \rangle = 0. \quad (\text{A.9})$$

The mixtures are now orthogonal, although still not representations of the original signals.

Figure A.1(d) shows the normalization of the power of the two mixtures. This step, together with the previous step, is known as whitening or equalization of the signals.

The final step of the procedure, shown in Figure A.1(e), is to find two original independent signals, by measuring the independence as given by Equation A.1:

There is enough information to find the signals if the correlations up to fourth order are used, or  $a + b \geq 4$ . This inequality is satisfied only when the signals are completely separated on each channel. The final ICA step is the most difficult part of the procedure. It requires an accurate measurement of

the statistical quantities, which implies that a long integration time for the averaging is necessary compared to the signal variation time.