Visible Spectrometer Utilizing Organic Thin Film Absorption

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
[redacted]
LIBRARIES

by

Laura C. Tiefenbruck

B. S. Electrical Engineering
Northwestern University, 2000

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
APR 0 4 2006
LIBRARIES

Master of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2004 [June 2004]

© Massachusetts Institute of Technology 2004. All rights reserved.

Author .....
Department of Electrical Engineering and Computer Science
May 28, 2004

Certified by ....
Vladimir Bulović
Associate Professor of Electrical Engineering and Computer Science
KDD Career Development Chair
Thesis Supervisor

Certified by .....
Terry P. Orlando
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by .....
Arthur C. Smith
Chairman, Department Committee on Graduate Students

BARKER

Visible Spectrometer Utilizing Organic Thin Film Absorption

by

Laura C. Tiefenbruck

Submitted to the Department of Electrical Engineering and Computer Science
on May 28, 2004, in partial fulfillment of the
requirements for the degree of
Master of Science

Abstract

In this thesis, I modeled and developed a spectrometer for the visible wavelength spectrum, based on absorption characteristics of organic thin films. The device uses fundamental principles of linear algebra to reconstruct spectral components of a signal from the transmission through an organic thin film. Best possible performance of the device is characterized, effects of noise and filtering techniques are observed, and results from several organic films are tested. The implemented device is optimized for cost, spectral reconstruction quality is tested, and guidelines for optimal device performance are proposed.

Thesis Supervisor: Vladimir Bulović
Title: Associate Professor of Electrical Engineering and Computer Science
KDD Career Development Chair

Thesis Supervisor: Terry P. Orlando
Title: Professor of Electrical Engineering and Computer Science

Acknowledgments

In the writing and work performed for this thesis, I have many people to thank:

First and foremost I would like to thank my research supervisor Professor Vladimir Bulović for his academic insights, clever technological ideas and practical suggestions. Additionally I would like to thank the members of the Laboratory for Organic Optics and Electronics who have been so generous is training me in the ways of organic semiconductors, particularly John Kymissis, Alexi Arango and Seth Coe-Sullivan for their experimental suggestions, hands-on lab experience and guidance through my work. I would especially like to thank Conor Madigan and Yaakov Tischler for lending me their expertise on the evaporator and transfer line.

I am thankful for and amazed by the graciousness of Professor Terry Orlando, to whom I owe the completion of this thesis. His scientific abilities have guided me to a satisfying conclusion and his personal warmth has kept my spirits up throughout.

I would like to thank my academic adviser Professor Ron Parker for his encouragement and direction. I would very much like to thank the EECS Department administrative assistant Marilyn Pierce for her good cheer, empathy, humor and indefatigable efforts that she shows every day to graduate students.

I would like to thank the Optics and Quantum Electronics Group, particularly Peter Rakich, Juliet Gopinath and Professor Erich Ippen for the use of their extensive colored glass filter collection.

I would like to recognize the exceptional teaching, counseling and friendship of my instructors and research advisers at Northwestern University, Professor Allen Taflove, Professor Prem Kumar and Professor Alan Sahakian, without tireless hours of personal attention from whom, I would not have pursued a degree, much less a higher degree in science and engineering.

My friends play no small part in my enjoyment of intellectual and personal pursuits in Boston. I would like especially to acknowledge the benefits I reaped from my weekly "Ladies' Lunches" with Heather Stern and Rebecca Reich-Kass, who gave me a sense of belonging at the Institute when there seemed to be no other substantiating

evidence. The members of the MIT Symphony Orchestra, the Brahms Society Orchestra and MacGregor House, particularly the amazing residents of J-Entry, have been a great source of happiness, moral support, fine conversation and lasting memories of my time at the Institute.

My family has been very supportive, both financially and emotionally, of my ambitions. Thanks to Mom, Dad, and even Mark throughout the years. Thanks also to Grandma Ann, Gene, Gary and David, and to Grandma Helen, Rick and Brian.

The mentorship of Dr. Alice White of Lucent Technologies Bell Laboratories has been a solid source of aid for me during my years at MIT. I would like to acknowledge the support of a Lucent Technologies Graduate Research Program for Women fellowship.

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

The detection and selection of optical wavelength signals is a fundamental signature of many scientific systems: biological, chemical, physical. However, at present, neither a simple nor an affordable option exists for many scientists to take these measurements. Systems currently commercially available would be considered expensive for a small laboratory, or for a non-critical instrument, with a minimum estimated cost around 2,000 USD[1].

Consequently, only those who absolutely require precision spectroscopic measurements are able to afford to purchase spectrometers. Those fortunate to have inherited a system or be part of a large, well-established institution may be able to gain access to this equipment, but this is not viable a solution for independent scientists, start-up companies or laboratories on a limited budget that do not particularly require a precision spectrometer but could make use of some spectroscopic measurements.

In addition to expense limitations, currently available systems tend to be unwieldy, requiring dedicated space on an optical table and a computer nearby for control and data collection. The instruments are delicate, should not be transported frequently and the input fiber or internal optics may be easily misaligned by a simple jolt. Nominally, the instruments should be calibrated frequently. None of these characteristics

[1]For example, the Ocean Optics S2000 Miniature Fiber Optic Spectrometer.

are desirable for the casual user of a spectrometer. With all the applications that would serve to benefit from a quick, easy, portable wavelength discriminator, one would guess that there must be a large, untapped market for such a device.

1.2 Prior Art

Solutions to the problems of size have been proposed. Researchers at the Delft University of Technology, Netherlands, have fabricated a CMOS spectrometer utilizing an array of Fabry-Perot etalons, each tuned to an individual resonance wavelength [3]. Researchers at Stanford University are able to micromachine a standing wave MEMS spectrometer, based on a variable-length Fabry-Perot etalon [8]. Another team has demonstrated a MEMS device using a high-dispersion grating in conjunction with a CCD [17]. This approach claims to be low cost, but integrally requires the fabrication of a miniature grating. All these ideas take the bulkiness out of optical signal measurement, but none addresses the delicacy and thus the portability of such instruments, nor the high cost of fabrication and assembly.

1.3 Objectives and Applications

The spectrometer proposed in this thesis would fill the gap where a robust, low-cost alternative is desirable. The goal is to create a system affordable enough to be used by anybody who could benefit from some simple wavelength discrimination. Resolution and signal magnitude would not be as precise as the more expensive systems. However, in the case where a scientist or engineer needs to determine whether or not a signal at a particular wavelength is present or not, our proposed device would be ideal. Similarly, an application where peak wavelength needs to be known but not detailed spectrum shape would be a good match for this device, or the case where shift in peak wavelength is of some significance. The goal of the proposed system is not to compete with high-precision optical grating devices, but to provide an alternative for cases in which a more robust, compact and above all less expensive option

would be preferred.

1.4 Description of Traditional Spectrometers

1.4.1 Grating Spectrometers

Traditional spectrometers operate via a dispersion element, typically a diffraction grating, which separates the light spectrum across a spatial dimension. A scanning photodiode or a CCD array then converts the optical signal to an electrical signal where it can be easily processed.

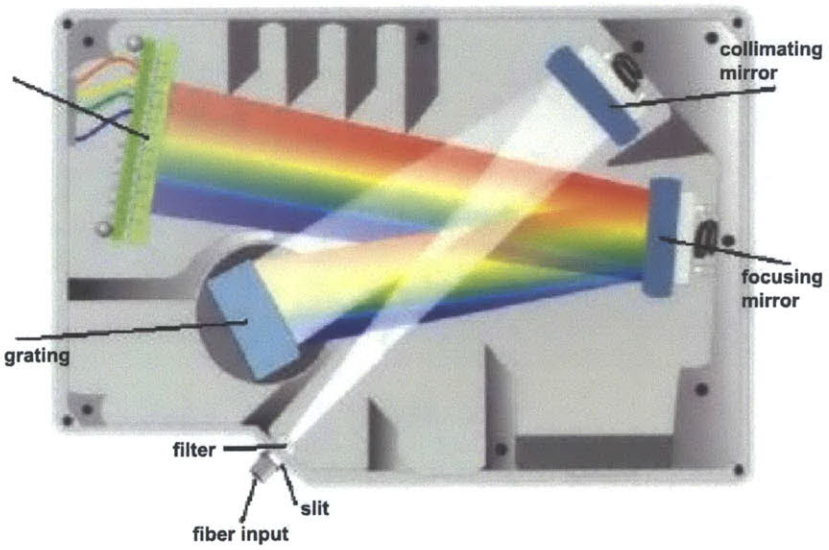

Figure 1-1: Example of a fiber-input grating spectrometer with CCD detector

It is easy to see how this system could be expensive, with the large optics and CCD, and easily misaligned, with several finely-tuned components. The physical limitations of these systems are mainly resolution, bandwidth, and sensitivity.

The resolution of a traditional spectrometer is limited mainly by the grating and the number of pixels on the CCD. Bandwidth limitations are the ultimate range of wavelengths that the grating is able to diffract and that the lenses and mirrors are able to focus on the CCD. Due to the precision machining of optical gratings, only a finite range of wavelengths can be diffracted and dispersed. Clearly, there is a tradeoff

between range and resolution of any given spectrometer.

A perfect grating would have all the grooves strictly parallel and of identical form. Any error in this shape will lead to blurring of the spectrum. However systematic errors in shape lead to periodic spectral lines known as "ghosts."

The angular dispersion of a blazed reflection grating is given as

$$\frac{d\theta}{d\lambda} = \frac{1}{\cos\theta}\frac{m}{d} \tag{1.1}$$

where θ is the angle of incidence, assumed constant, λ is the optical wavelength, m represents the path difference measured in wavelengths between two neighboring grooves. High dispersion is achieved achieved with small spacing d or measurements of large order m [2].

1.4.2 Prism Spectrometers

The operation of a prism spectrometer is very similar to that of a grating spectrometer, with the prism providing chromatic dispersion.

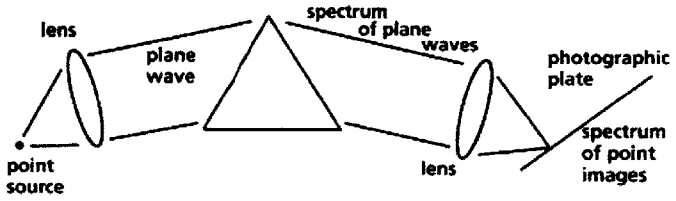

Figure 1-2: Typical prism spectrometer design

A prism with apex angle α and refractive index n diffracts an incident beam by the angle θ_d

$$\theta_d = \theta_i - \alpha + \sin^{-1}[(n^2 - \sin^2\theta_i)^{1/2}\sin\alpha - \sin\theta_i\cos\alpha] \tag{1.2}$$

which may be easily derived by iterating Snell's law twice at the two air-glass interfaces [11].

1.4.3 Fabry-Perot Etalons and Spectroscopy

Spectroscopy using Fabry-Perot etalons is significantly different than the grating and prism methods which use a physical optical device to add dispersion to the incoming signal, spreading it in space where it may be resolved. The Fabry-Perot spectrometers use the feature of finely-tuned resonance of etalons to select and amplify a narrow wavelength range $\Delta\nu$, the entire signal from which is then measured by a single photodetector. Fabry-Perot spectroscopy (or interference spectroscopy) is most often used for resolving detailed structure of spectral lines [7].

A Fabry-Perot interferometer is in its most basic form a pair of flat, parallel, closely spaced mirrors [12]. A Fabry-Perot etalon is a solid, dielectric, extremely transparent slab, also with very flat polished faces. These Fabry-Perot devices can have sharp resonances at discrete frequencies. Multiple reflections inside the resonator cavity create a transmission profile with peaks and dips, with low-loss etalons and cavities having sharp differences in transmission and lossy resonators having a smoother profile. The transmission curves mathematically are defined by Airy functions $\mathcal{A}(\theta)$.

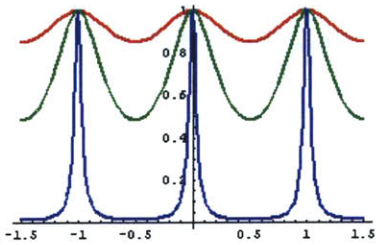

Figure 1-3: Transmission fringes of a Fabry-Perot cavity

The frequency spacing of these peaks is based only on the distance between the faces, by the simple condition

$$\nu_F = \frac{c}{2d} \tag{1.3}$$

Additionally the spectral width of these peaks, or the *finesse* of the resonator can be determined readily by a frequency analysis [11]. Finesse is defined simply as the quotient of the separation of adjacent peaks and the full width half maximum of each

peak.

It should be noted that if the interference fringes of a Fabry-Perot interferometer should approach too closely, the difference between the two cannot be resolved. By Equation 1.3 we see that there is an inverse relation between this frequency spacing and the distance between the cavity interfaces. Thus, in order to achieve high resolving power, the plates must be placed very close to one another. In order to achieve a high-finesse cavity, the plates must be flat and parallel. The eventual limitations of this system are clear. More specifically, the minimum resolvable bandwidth and finesse $\mathcal{F}$ are given by

$$\Delta\nu_{min} = \frac{c}{\mathcal{F}2n_f d} \tag{1.4}$$

$$\mathcal{F} = \frac{\pi\sqrt{R}}{1-R} \tag{1.5}$$

where R is the reflectivity of the mirror or air-cavity interface. This equation merely states that as the bandwidth of the input signal increases, the mth fringe for one wavelength will begin to approach the $(m+1)$th fringe for the opposite wavelength. Even though this system is very sensitive and can resolve a signal with great precision, it's capacity to resolve any large range of wavelengths is limited.

1.5 Slim Format Spectrometer

1.5.1 Theory

The beauty of the proposed design, allowing it to be compact and affordable, is the substitution of expensive gratings and lenses with a variable-absorption element, created from the properties of organic thin films [reference patent app.]. The functional crux of the Slim Format Spectrometer is a variable thickness organic thin film, which we have seen will have a wavelength dependent absorption profile in addition to a thickness dependent transmission profile. From these two properties, a transmission matrix can be created, fully characterizing the organic film in the two variables of

wavelength and thickness.

$$T = \begin{bmatrix} t(\lambda_1, x_1) & . & . & . & t(\lambda_n, x_1) \\ & . & & & . \\ & . & & & . \\ t(\lambda_1, x_m) & . & . & . & t(\lambda_n, x_m) \end{bmatrix} \tag{1.6}$$

Once the film is characterized in these parameters, it is proposed that an unknown signal will have the combined effect of the superposition of each of its wavelength components individually and thus will be able to be reconstructed with a trivial amount of linear algebra.

Specifically, it is known that the transmission of an arbitrary signal through a variable thickness organic thin film will depend both on the wavelength composition of the signal and the thickness of the film. Translated to linear algebra we can see that the detected signal $D(t)$ is a result of multiplication of our known transmission characteristics $T(\lambda, t)$ by the incoming spectrum $S(\lambda)$.

$$D = T \cdot S \tag{1.7}$$

In order to reconstruct the unknown spectrum $S(\lambda)$ from $D(t)$, we simply invert the transmission matrix and multiply through.

$$T^{-1} \cdot T \cdot S = T^{-1} \cdot D$$
$$S = T^{-1} \cdot D$$

Though it may sound foolproof, in fact there is a great deal of uncertainty in the inversion of the transmission matrix $T(\lambda, t)$. First, not all matrices are invertible, so if the characteristic of a particular organic film yields a singular matrix, there is no possibility that an unknown spectrum could be reconstructed using this technique and the particular film. Though not mathematically required to form an orthogonal set

of basis vectors, in order to form a transmission matrix in which a unique absorption coefficient corresponds to a unique wavelength, the wavelength dependence of the absorption coefficient must be monotonic over the range of interest. Practically, many organic absorbers comply with this restraint over some part of the visible range.

1.5.2 Device

Experimentally the instrument may be realized in a number of configurations, falling mainly into two categories: single photodetector or photodetecter array. The configuration requiring only a single photodetector also makes use of a scanning element, either a linear stage on which the variable-thickness film will be scanned between the incoming spectral beam and the photodetector or a disc on which a film with variable thickness as a function of angle will rotate between the photodetector and incoming light. These configurations have the potential to be analog, with resolution limited only by beam spot size, scattering, and system noise.

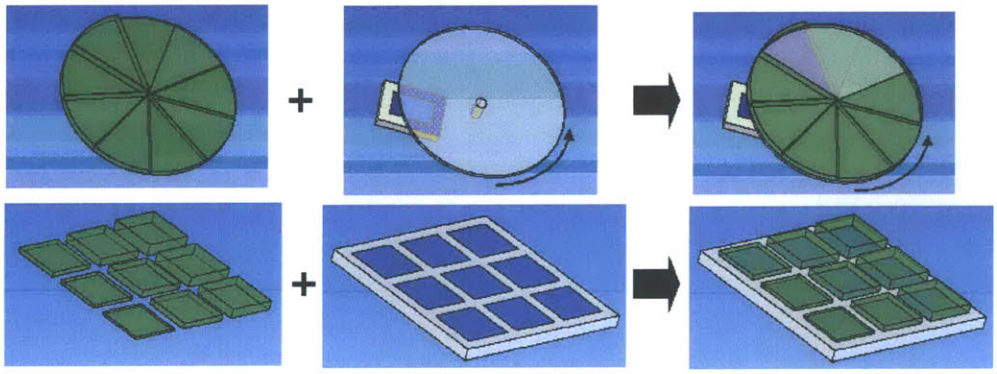

Figure 1-4: (a) single photodetector and analog wheel configuration of Slim Format Spectrometer (b) array photodetector and discrete film thickness configuration

The configuration utilizing an array of photodetectors is very clever in that it allows simultaneous measurement of all points in the detected signal $D(t)$. Variations in the incoming spectrum are eliminated with this simultaneous measurement. This approach is however, necessarily discrete. Care must be taken to produce thin films, each with a unique thickness and that each is well aligned over a single photodetector. In the case that a continuous film is produced and laid over the photodetector array,

a unique thickness may be achieved for each photodetector, but special care must be taken such that the incident light always intersects the same section of the film, to achieve reproducible transmission characteristics. This approach is also limited by the number of individual photodetectors. For example, a large array consisting of 32×32 pixels would have 1024 points of resolution. For our purposes, this is by far a sufficient number, though the cost of the device would scale with the resolution due to the photodetector array.

There are few absolute limitations on a spectrometer system of this sort. In a traditional spectrometer, one would expect there to be limitations on the bandwidth of a grating, or a lens or mirror for that matter. There are limitations to how finely a grating can be manufactured, how many pixels can be fit in a certain space on a CCD and how little chromatic aberration can be expected. In a spectrometer not based so strongly on the optical properties of the components, the limitations arise from altogether separate sources. There is a limit on how accurately film thicknesses can be grown, on the order of Ångstroms. A noise limit exists as to how finely an optical signal can be converted to a voltage. A convenient time limit exists as to how many data points can be collected and how much averaging and filtering may be performed.

The most important trade off is likely the exchange of cost for performance. At what point does the increased cost of a spectrometer outweigh the benefits of increased performance? One may imagine that the most elaborate configuration of the spectrometer would utilize a state of the art, perhaps liquid nitrogen cooled, photodetector - either array or single diode. A specially designed absorptive dye would be used in elaborate packaging to ensure optical stability and protection from environmental elements. Motors with stable rotation frequency or an incredibly precise linear stage would be employed. However the cost of such a system would certainly approach and surpass a traditional spectrometer system. On the other extreme, a noisy system, due to motor jitter, optical scattering, poor photodiode response, and possibly vibrations or drift would potentially still yield a discernable result. The least expensive configuration would make use of everyday household, lab and office components while attempting to minimize the potential for user error. Much of the

objective of this thesis work is to recognize the upper level of functionality for a system with a lower limit of expense. That is, what possibly is the best performance we can expect of an extremely inexpensive system? Is it sufficient? Further, the work attempts to address what minimal improvements can be made in order to greatly enhance the performance of the inexpensive system without upgrading to components that in expense, rival the currently available traditional spectrometer systems.

The least expensive configuration of a Slim Format Spectrometer nevertheless requires a photon to electron conversion and a multiplicative computation, neither of which may be cost-minimized beyond a somewhat expensive threshhold. The fortunate angle of this expense is that most any user who requires a spectrometer likely already owns a computing processor and a photodiode. In any case, a traditional spectrometer also requires these elements. The components then that the Slim Format Spectrometer is able to improve upon are the wavelength selection method – that is to say, the sorting and separating of one wavelength from another. In achieving this task, I believe the Slim Format design is successful.

1.6 Structure of Thesis

First, principles of absorption and the interaction of photons with materials will be explored. Next we will introduce necessary concepts in linear algebra and numerical methods. Simulations and numerical results of the reconstruction of spectra as functions of thin film absorption, center wavelength and variance.

The process for construction the physical experiments will be discussed and results from these trials presented. The effect of filtering and processing on these resulted is discussed. Finally, proposals for next steps are suggested.

Chapter 2

Absorption in Organic Thin Films

2.1 Resonant and Nonresonant Scattering

The proposed spectrometer design uses the absorption properties of organic thin films in place of a grating or dispersive element. The absorption properties, technically *dissipative absorption*, of organic thin films have been well characterized. Physically, absorption is the interaction of atoms with incident electromagnetic waves [7]. Depending on the energy on the incoming light, an atom will react in one of two ways. Genearlly speaking, an atom will scatter the light, redirecting it without otherwise affecting its properties. However, if the energy of the incoming photons matches the amount of energy needed to jump an electron from one state to a higher state, the photon will be absorbed. The photon's energy will be transferred to the atom, and an electron will be allowed to jump from its present state to a higher state. The subsequent conversion of this extra energy to thermal energy gives the entire process the name *dissipative absorption.* The process of transferring energy from a photon to an atom is referred to only as *absorption.*

In the case when an atom has no resonances near the energy of incident photons, say visible light, no electrons change state, but the atom does interact with the light. This exchange is called *nonresonant scattering.*

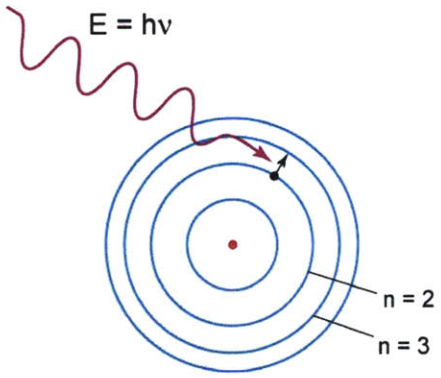

Figure 2-1: Resonant interaction of an atom and a photon

2.2 Absorption in Semiconductors

An analysis of the absorption properties of organic thin films would not be complete without the acknowledging the behavior of these films as semiconductors. The optical absorption properties of semiconductors are well understood [14]. In semiconductor devices, excess carriers are often created for device operation by optical excitation. The main distinction between absorption generally and absorption in a semiconductor is that the semiconductor is often thought of as separated into bands: the conduction band, of higher energy, and the valence band, with lower energy. An energy region exists between the two in which there are few if any states for carriers to occupy. Thus absorption in semiconductors, including organic thin films, is restricted to energy values which would not add into or from this energy region without states.

Generally speaking, the conduction band being of higher energy will have unoccupied states while the states in the valence band are primarily occupied. Thus, a photon of energy greater than that of the forbidden so-called bandgap region is likely to be absorbed, launching an electron from the valence to the conduction band. Light of energy less than the bandgap is unable to be absorbed and will pass through the semiconductor, which is transparent to these frequencies. The bandgap of a semiconductor may be precisely determined by this selective absorption. A monochrometer is swept across a wavelength range and transmission observed. The sharp transition from very little absorption (high transmission) to considerable absorption (lower

transmission) called the *absorption edge* [10], is likely the bandgap energy. For common semiconductors such as silicon and GaAs, bandgap energies are in the range of a few eV, equivalent to the energy of a photon at wavelengths of around 1 micron (infrared). Since the "turn on" energy, or so-called absorption edge, for these materials is below the visible spectrum, these materials also absorb wavelengths throughout the visible spectrum.

Often, predictably, photons of energy a fair bit larger than the bandgap energy will be absorbed. In this case they are often excited to an unoccupied state within the conduction band such that relaxation to a lower energy conduction band state. In this relaxation, the excited electron loses energy to the lattice in scattering until it reaches equilibrium with the other conduction band electrons. In this manner an electron and hole are created. This electron and hole, considered *excess carriers*, are out of balance with their environment. Eventually they must recombine. However, while they exist they contribute to the conductivity of the material.

2.3 Composition of Color

If you've ever seen a color wheel, you'll know that the colors as we typically categorize them do not correspond exclusively to the linearly-progressing single-wavelength colors from red to violet. Instead there is a continuous distribution of perceived colors that wraps around from red through magenta and violet to blue. This distribution is a result of color mixing, as an artifact that several materials are absorbant over a combination of wavelength bands.

The traditional color spectrum can be thought of as a complete set. Orthogonal sets of basis wavelengths may then be chosen, often called *primary colors*, such that any perceivable color may be generated as a combination of these wavelengths. These sets are by no means unique and do not necessarily even need to be monochromatic. The most common set of basis wavelengths is the RGB (red, green, blue) standard used in televisions and computer monitors. This basis set is called an additive set, as the combination of all three colors leads to a white image. Another such set of basis

wavelengths with which you may be familiar is the CYMK (cyan, yellow, magenta, black) set. The CYMK set is a subtractive set in the digital domain such that equal parts of each of the components results in a black image. These orthogonal sets are diagrammed below, as first laid out by James Clerk Maxwell. Notice that cyan can be created as a superposition of green and blue, and magenta a superposition of red and blue. Yellow is obtained by combining red and green, and blue is obtained by mixing cyan and magenta.

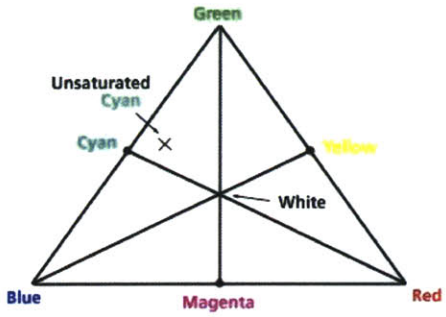

Figure 2-2: James Clerk Maxwell's Color Triangle

The objects we perceive as having color take on their appearance due to selective absorption of energies in the visible spectrum. Objects that appear to us as red, for example, have a preferential absorption of the other colors (green and blue) while nonresonantly scattering red wavelengths.

Many organic molecules have selective absorption over the visible spectrum. These molecules sometimes consist of long chains of alternating single and double bonds, sometimes as chains of alternating single and double bonds turned on themselves into a ring. These molecules all have resonance frequencies in the visible spectrum, and therefore participate in selective absorption. As the energy levels of individual atoms are discrete and very precisely defined, the absorption profile of an atom has very sharp peaks. Proximity of atoms to one another in liquids and solids leads to broadening of these absorption energy bands. Consequently, we can expect that an organic dye will absorb a significant portion of the visible spectrum and appear colored. If the absorption peak were very sharp as is the case in individual atoms, we would expect these materials to appear white as they reflect most wavelengths.

Wavelength dependent absorption will be very important in the realization of an absorption-based spectrometer.

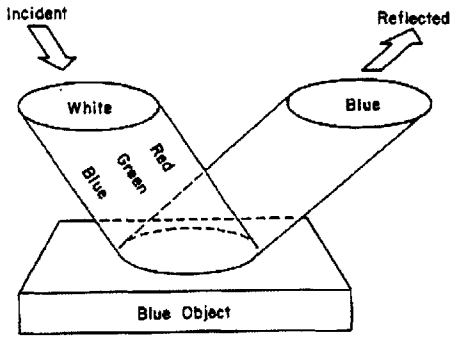

Figure 2-3: Appearance of color as selective absorption

2.4 Beer-Lambert Law

In addition to photon energy contributing to selective absorption of incident light, density of a material has an effect. In a bulk liquid or solid, generally speaking, the more dense the material, the more dissipative absorption one can expect. Intuitively, one can reason that the more dense the material, a photon will have more atomic interactions and thus more opportunities to exchange energy. In the case of thin films, these samples are categorized as such partially because they transmit a large fraction of the light incident upon them. By definition, a thin film is on the order of one micron or less in thickness, which translates to on the order of one thousand or fewer molecular layers. By a very rudimentary argument, it can be said that so long as more than one thousand photons of resonant energy are incident upon a thin film over a time scale less than the dissipative time of a molecule, then certainly the absorption properties of the film have been saturated and some light will pass being only nonresonantly scattered. Perhaps a more common way of thinking is that the incoming flux of photons is very much greater than the number of atoms the population of photons encounters such that several are allowed to pass.

The dependence of absorption, that is fraction of incident photons resonantly, dissipatively absorbed, on the thickness of a thin film, that is the number of molecular

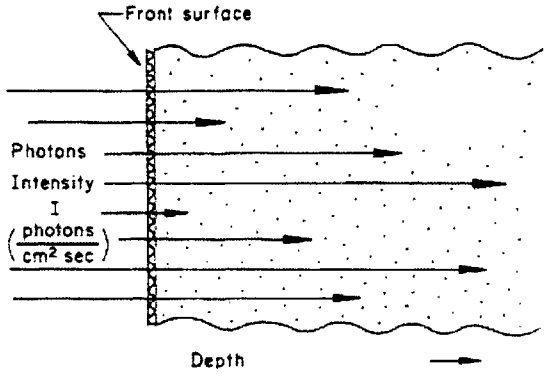

Figure 2-4: Photons as they are absorbed with increasing thickness in a solid

layers, is well characterized. The Beer-Lambert Law (or Beer's Law) defines a linear relationship between absorbance, density of a film and path length of light through the film (in the case of light normally incident on a film, path length is equivalent to film thickness).

$$A = a(\lambda)bc \tag{2.1}$$

A is the measured absorbance through the film, $a(\lambda)$ is the wavelength dependent absorptivity coefficient for a particular material, b is the optical path length and c is the concentration of molecules, analagous to the density of the material.

The full derivation of this expression is included in Appendix A. For present purposes, it suffices to say that absorbance is a function of wavelength squared. Further, transmission is related to intensity by the relation $T = I/I_o$ and $A = -\log T$. Beer's Law in terms of transmission through a thin film is thus

$$T(\nu) = e^{\gamma(\nu)b} \tag{2.2}$$

Where $\gamma(\nu)$ is a well-known gain coefficient in intensity per unit length of the material. This gain coefficient $\gamma(\nu)$ is simply the opposite of the common absorption coefficient, $\gamma(\nu) = -\alpha(\nu)$ Physically, the relation states that intensity of light drops exponentially with interior distance in a medium. In the case of thin films, transmission drops exponentially with increasing film thickness.

Chapter 3

Introduction to Numerical Linear Algebra

Approaching numerical methods takes a certain shift in perspective. Often we think of solving problems as obtaining direct solutions to equations. The aim of numerical methods is not necessarily to obtain expressions for our data but instead to develop an algorithm well-suited to obtaining a solution [6].

3.1 Concepts in Linear Algebra

3.1.1 Familiar Definitions

In thinking about linear algebra and the associated numerical methods, it is helpful to think of standard multiplication operations as sums of products rather than simply algorithmic multiplication of an array by another array, constant or vector.

A Matrix Times a Vector

Take the familiar situation of a matrix multiplied with a vector.

$$b = Ax = \sum_{j=1}^{n} x_j a_j \tag{3.1}$$

Here we represent the jth column of A as a_j. Schematically, this situation is displayed

$$[b] = [a_1|a_2|...|a_n] \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix} \tag{3.2}$$

$$[b] = x_1[a_1] + x_2[a_2] + ... + x_n[a_n] \tag{3.3}$$

Though the difference between these visual representations and the summation formula is subtle, the expressed difference is essential. From mathematical training, most people are used to interpreting the statement $Ax = b$ to mean that A performs a transformation on x to produce an output b. The visual column-based representation attempts to emphasize the difference which is crucial in linear algebra, that in fact x acts on A to produce b.

Range and Nullspace

Purely by definition we know that the range of a matrix A is the set of vectors that can be expressed as Ax for some x [15]. That is, the range is the set of column vectors that can be produced by transforming A with any x. Technically, range of A is the space spanned by the columns of A. Range is also commonly known as column space.

In contrast, the nullspace of A is the set of vectors that satisfy $Ax = 0$. These vectors form a unique vector space basis. The entries of each vector in the nullspace of A give the coefficients of an expansion of zero as a linear combination of columns of A.

$$0 = x_1a_1 + x_2a_2 + ... + x_na_n \tag{3.4}$$

Row space and left nullspace

Less common but still useful are the row space and left nullspace of a matrix. These qualities are analagous to the range and nullspace of a matrix. The row space is simply the column space of the transpose of the matrix. The left nullspace is the set of vectors for which $A^T y = 0$ [13].

Rank

Rank refers to the dimensionality of a matrix, the dimension of the space spanned by its column or rows. A matrix has full rank if the number of linearly independent vectors is equal to its smaller dimension. For an $m \times n$ matrix, the matrix has full rank if for $m \geq n$ it has n linearly independent columns. This matrix can be characterized as a one-to-one mapping function.

Adjoint and Transpose

For complex scalars we use the complex conjugate. For linear algebra, we define the "hermitian conjugate" or "adjoint" as the matrix for which the i, j entry is the complex conjugate of the j, i entry.

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} \Longrightarrow A^* = \begin{bmatrix} a_{11}^* & a_{21}^* & a_{31}^* \\ a_{12}^* & a_{22}^* & a_{32}^* \end{bmatrix} \tag{3.5}$$

In the case where the matrix A is equal to its adjoint A^*, we call A hermitian. Note that hermitian matrices must be square. In the case that all matrix entries are real, the adjoint is simply the transpose, where the rows and columns of A are interchanged. If a real matrix is hermitian such that $A = A^T$, then it is symmetric.

Matrix Inverse

An invertible matrix is a square matrix of full rank. The m columns of a nonsingular (full-rank) $m \times m$ matrix form a basis for the whole space (column space + row

space). We can uniquely express any vector as a linear combination of the columns of an invertible matrix.

If we consider the special case e_j in which the jth entry is 1 and zeroes elsewhere

$$e_j = \sum_{i=1}^{m} z_{ij} a_i = A z_j \tag{3.6}$$

$$[e_1|...|e_m] = I = AZ \tag{3.7}$$

I is then an $m \times m$ matrix with diagonal 1 values known as the identity matrix. Z is defined as the inverse of A. Every square full-rank matrix has a unique inverse A^{-1} such that

$$AA^{-1} = A^{-1}A = I \tag{3.8}$$

A Matrix Inverse Times a Vector

Rather than thinking of the inverse matrix as a complex set of operations on the forward matrix, it can be thought of as the set of coefficients that produces the solution to the equation $Ax = b$. Conversely, x is the vector of coefficients that maps the basis vectors of A onto b. Multiplication of an inverse matrix is simple a change of basis operation, of which there are many types in linear algebra [15]. $A^{-1}b$ is the vector of coefficients of the expansion of b in the basis columns of A.

3.2 Numerical Inversion Algorithms

3.2.1 Elimination Methods

The field of numerical linear algebra has a long history and there exist several well-developed algorithms for solving matrix inversion. The most basic technique is Gaussian elimination. This permutation algorithm solves the problem much as you would by hand, first by transforming the matrix rows into a triangular configuration then

using back substitution to solve each row sequentially [4] [15] This method is simple is all respects and requires that the matrix be square and nonsingular [13].

The next improvement upon Gauss's form is Jordan elimination in which the Jordan form of a matrix is obtained [6] [13]. In this process, the diagonal form of a matrix is obtained so that back substitution is not necessary and the inverse equation coefficients are obtained directly.

Both Gauss and Jordan allow for the solution of the equation $AX = B$ only, where B is a single column b and X is a column x. The method of Aitken allows for a more sophisticated solution to the problem $Y = CA^{-1}B$ [6]. This elimination method combines solutions of the pair of equations

$$\begin{aligned} AX &= B \\ CX &= Y \end{aligned} \tag{3.9}$$

The method can be expanded to quite large arrays of matrices, but ultimately relies on the same elimination methods as the processes of Gauss and Jordan. All three of these require that the matrix to be inverted be square and nonsingular.

3.2.2 Compact Elimination

In addition to the standard elimination and reduction methods, a number of compact elimination methods were developed, to reduce the processing power required to invert large amounts of data.

These methods use the same fundamental procedure as Gauss and Jordan in that a triangular or Jordan form matrix is obtained then the resulting equations solved. However, the compact methods seek to determine whether the triangular and Jordan form matrix may be obtained directly, without every intervening matrix iteration [6].

The methods of Doolittle and Craut examine the steps of Gauss reduction symbolically and simplify to calculate triangular matrices based on matrix elements, based on the observation that the first row of the original matrix is the same as that of the upper triangular matrix. The method of Doolittle makes no attempt to simplify the right-hand, non-matrix element functions, the side of the equality to which matrix

elements are compared. The method of Craut further simplifies the matrix reduction by creating an algebraic expression for the right-hand function equations, thus making back substitution a few computational steps shorter for each matrix row.

3.2.3 Singular Value Decomposition

The *singular value decomposition* combines the approaches of Gaussian elimination to the triangular matrix and the Gram-Schmidt orthogonalization which takes the columns of a matrix and makes them into an orthonormal basis [13].

The singular value decomposition (SVD) is very similar to eigenvalue-eigenvector factorization. In a symmetric matrix, $A = A\Lambda Q^T$. The eigenvalues are in the diagonal matrix Λ and the eigenvector matrix Q is *orthogonal*, $Q^TQ = I$, because eigenvectors of a symmetric matrix can be chosen as orthonormal. However, for most matrices these conditions are not true and are impossible for rectangular matrices. In the singular value decomposition, we allow Q and Q^T to be any orthogonal matrices, not necessarily transposes. Thus, the eigenvalue-eigenvector factorization again becomes possible. The diagonal (but rectangular) matrix in the center is denoted by Σ and its entries by $\sigma_1, ..., \sigma_r$, where r is the rank of A. By definition, any $m \times n$ matrix A can be factored into

$$A = Q_1 \Sigma Q_2^T \tag{3.10}$$

where Q_1 and Q_2^T are orthogonal and Σ is diagonal. The m columns of Q_1 are eigenvectors of AA^T. The n columns of Q_2 are eigenvectors of A^TA. The singular values r on the diagonal of Σ $(m \times n)$ are the square roots of the nonzero eigenvalues of both AA^T and A^TA.

The singular value decomposition is useful in applications where compactness of form is essential. In image processing, for example, only some of the singular values are kept and multiplied into the full image. The rest of the image information is thrown away, significantly compressing the image data size.

3.2.4 Least Squares Minimization

Least squares data fitting has been an indispensable tool for approximately 200 years. The problem is straightforward, seeking to solve the system of equations $Ax = b$ where b is rectangular with more rows than columns. Generally speaking, this problem has no solution. Up until now we have considered only inversion algorithms for matrices of full rank. The assumption behind the least squares minimization problem is that there exists a set of vectors that minimize $\|Ax - b\|_2$, the second norm of the residual $Ax - b$.

In the case of the least squares minimization problem, the system of equations $Ax = b$ is overspecified, with n unknowns and $m > n$ equations. We wish to find a vector x that satisfies $Ax = b$. The heart of the least square minimization is making the residual vector r as small as possible, where $r = b - Ax$.

It does seem challenging to solve a problem that nominally has no solution. The least squares minimization simply minimizes error between the actual solution and the solution found. It is not even necessarily a unique solution. Further, the least squares minimization takes on a number of incantations, the most popular being polynomial fitting but also including the pseudoinverse.

Polynomial Least Squares Fitting

Given a set of distinct points, we wish to fit an $(m-1)$-dimensional polynomial which best fits the data. We call this polynomial the *polynomial interpolant*, which is a polynomial of, at most, degree $m - 1$:

$$p(x) = c_0 + c_1 x + ... + c_{m-1} x^{m-1}. \tag{3.11}$$

We create a system of equations to relate the coefficients c_i to the data x_i, y_i:

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{m-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{m-1} \\ 1 & x_3 & x_3^2 & \dots & x_3^{m-1} \\ & \cdot & & & \cdot \\ & \cdot & & & \cdot \\ & \cdot & & & \cdot \\ 1 & x_m & x_m^2 & \dots & x_m^{m-1} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_{m-1} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \cdot \\ \cdot \\ \cdot \\ y_m \end{bmatrix} \tag{3.12}$$

Fortunately, this system is almost always nonsingular, and can be solved. The polynomial that is the least squares fit is then the one that minimizes the sum of the squares of the deviation from the data

$$\sum_{i=1}^{m} |p(x_i) - y_i|^2 \tag{3.13}$$

This sum of squares is equivalent to the square of the norm of the residual, $\|r\|_2^2$ of the rectangular system of equations similar to equation 3.12 above except with an n-dimensional coefficient vector, where $m > n$.

$$\begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{m-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{m-1} \\ 1 & x_3 & x_3^2 & \dots & x_3^{m-1} \\ & \cdot & & & \cdot \\ & \cdot & & & \cdot \\ & \cdot & & & \cdot \\ 1 & x_m & x_m^2 & \dots & x_m^{m-1} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \cdot \\ \cdot \\ \cdot \\ c_{n-1} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \cdot \\ \cdot \\ \cdot \\ y_m \end{bmatrix} \tag{3.14}$$

3.2.5 Moore-Penrose Pseudoinverse

Though there are a number of well-developed algorithms for inverting a matrix, I used the standard MATLAB function "`pinv`" for my simulations. The function $B = \text{pinv}(A)$ returns the Moore-Penrose pseudoinverse of A. Given an $m \times n$ matrix, this generalized inverse algorithm developed independently by Moore in 1920 and Penrose

in 1955 returns a unique $n \times m$ matrix pseudoinverse [1] [5]. The conditions that are satisfied by the Moore-Penrose pseudoinverse are:

$$ABA = A \tag{3.15}$$

$$BAB = B \tag{3.16}$$

$$(AB)^T = AB \tag{3.17}$$

$$(BA)^T = BA \tag{3.18}$$

Equations 3.17 and 3.18 above state the condition of the matrices being hermitian.

The pseudoinverse is the least squares solution to $Ax = b$. Since it is a generalized inverse, it is computationally more expensive than the simple $\texttt{inv}(A)$ function, though considerably more robust. As an optimization algorithm more than an equation solver, the Moore-Penrose pseudoinverse can provide a satisfactory matrix pseudoinverse for A, even if A is not square, or square and singular.

Chapter 4

Numerical Results

One of the most exciting steps in developing a new device is determining what is possible, what is the best possible device operation that can be expected. In this chapter, I examine the edges of the best case scenarios that a physical device could expect to achieve. The quality of inversion of the transmission matrix is examined, the resolution that the spectrometer can obtain as well as operating parameters in film thickness, random noise and robustness of design.

The simulations were intended to test the feasibility of the instrument, if very high quality materials were obtained and clean performance was achieved. Although realistic conditions were modeled by the addition of random noise (MATLAB function "`randn`"), physical noise in this system turns out likely to be coupled and systematic and thus not completely accounted for in the modeling.

The main categories of simulations demonstrated here are those which determine invertibility of various structures of transfer matrices, those which determine reconstructibility of a smooth input spectrum, those which analyze reconstructibility of various input spectra, those which study the amount of data necessary to reconstruct a spectrum, and those which look at various types of filtering and noise management techniques.

4.1 Absorption Profile

As a first step, we demonstrate the effect of the shape of the wavelength-dependent absorption profile of the organic thin film on both the characteristic transmission matrix and also on the properties of the detected signal of an input spectrum and the reconstructibility of this waveform.

Intuitively, we aim to create an absorption profile that has a smoothly sloping section matched to the wavelength range of interest in our unknown spectrum. Absorption that is too flat in this area would lead to little difference between the absorption of neighboring wavelengths and hence poor spectral resolving power. Absorption that is too steep would lead a very small film thickness difference to yield a large detected signal variation, which in turn may lead to a noisy reconstructed spectrum.

As many organic thin films display an absorption function that is exponentially decreasing with wavelength, we use this profile as a starting point.

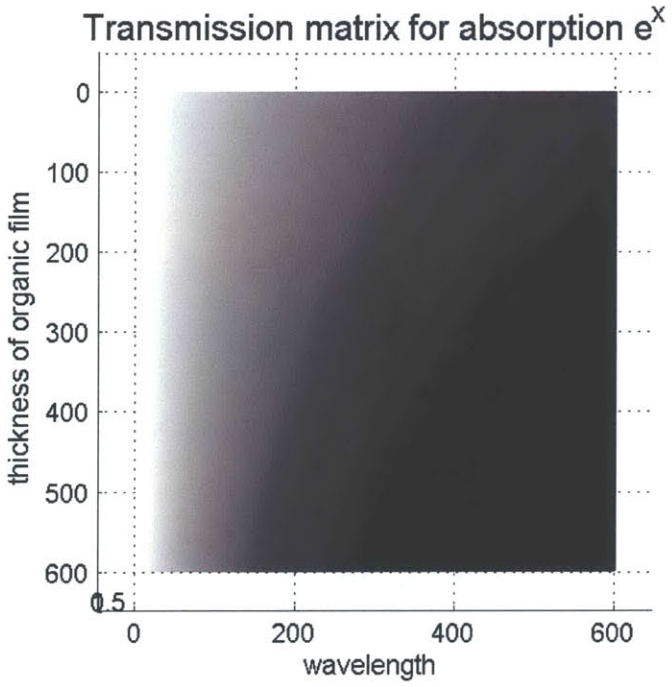

Figure 4-1: Transmission matrix for absorption of e^x, $x = \frac{\lambda - \lambda_{min}}{\lambda_{max} - \lambda_{min}}$

As defined by Equation 2.2, Beer's Law of transmission

$$T(\nu) = e^{\gamma(\nu)b} \tag{4.1}$$

the transmission through a film drops off exponentially with path length b, approximately equivalent to thickness of the film t, and doubly exponentially with wavelength, as the absorption profile is exponential and transmission is exponential with absorption.

To gain an idea for what signals look like through this film, Figure 4-2 shows the detected signal, in arbitrary units, for five Gaussians through an organic thin film. The absoprtion profile is exponentially increasing with wavelength, such that absorption is zero at 400 nm and 100% at 700 nm. The variance of each Gaussian is 30 nm and the thickness of the organic film is linearly increasing from 0 to 400 Å.

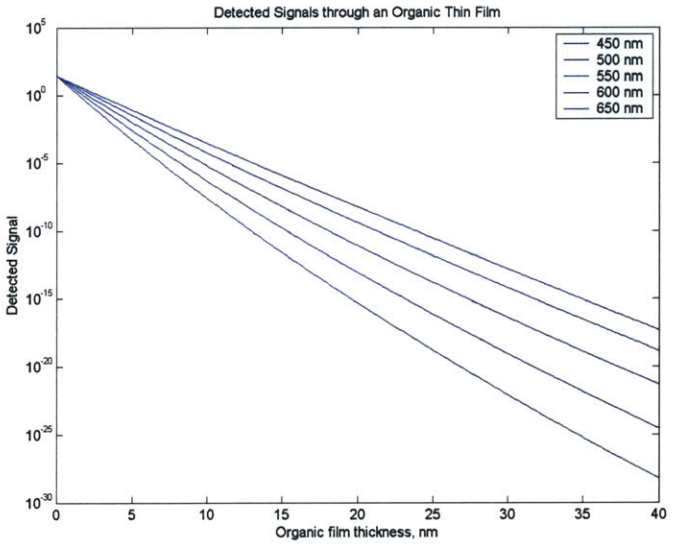

Figure 4-2: Transmission curves at varying peak wavelengths λ_p for absorption profile of e^x, $x = (\frac{\lambda - \lambda_{min}}{\lambda_{max} - \lambda_{min}})$

Each of these curves is roughly exponentially decreasing, as expected with the exponential absorption profile. At the transparent section of the organic thin film, each of the Gaussians exhibits perfect transmission, leaving each signal indistinguishable from the others. As film thickness increases, the signals disperse based on their respective absorption. In this case, the longer wavelengths have higher absorption,

so the bottom curve in Figure 4-2 represents the Gaussian input spectrum with peak wavelength of 650 nm. Progressing toward lesser absorption, we eventually arrive at the top curve with peak wavelength of 450 nm. The goal of the Slim Format spectrometer is to match this detected signal to that of the transmission matrix, an example of which is shown in Figure 4-1, by means of linear algebra and the basis vectors of the transmission matrix to reconstruct the input spectrum numerically.

To demonstrate quality of reconstruction of spectra in a film with exponential absorption dependence, I modeled a Gaussian input spectra, centered toward the red end of the spectrum corresponding to higher absorption, with a variance of 20 pixels. Randomly distributed noise was added to the detected signal at magnitudes varying from 1 to 10^{-20}.

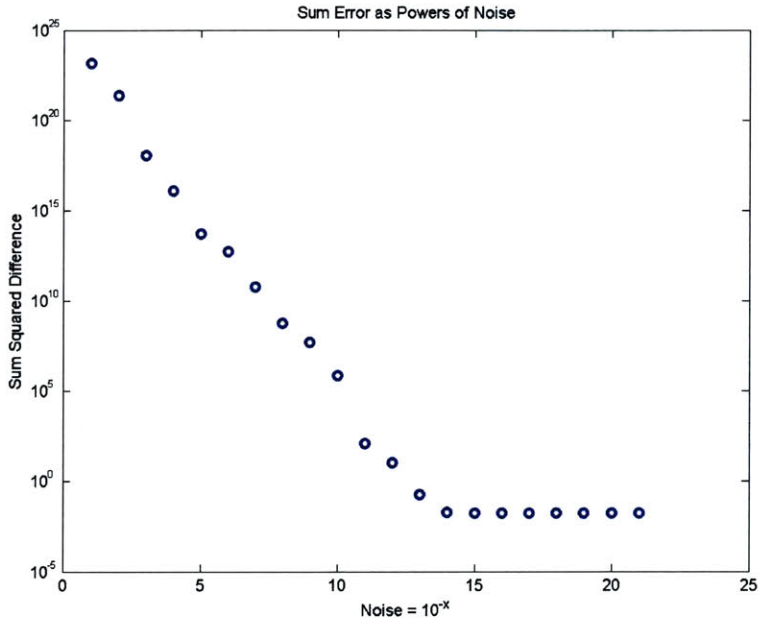

Figure 4-3: Noise versus reconstructibility

The relation of sum squared error between the reconstructed signal and the input spectrum is exponentially decreasing with the order of additive noise. There is a clear threshhold beyond which signal reconstruction will not improve, at noise of approximately 10^{-13}. Given that the detected signal decreased exponentially from an arbitrary magnitude of 40 down to 8×10^{-15}, a thresshold of near that value

is expected. There is an additional interesting feature where additive noise is 10^{-9} and 10^{-10}, which is where the magnitude of the reconstructed signal approaches the magnitude of the input signal. It appears that there is a $-\sin x$ behavior deviating from the exponentially decreasing trend around the pivot of noise magnitude 10^{-10}. A simulation with finer resolution could easily map out this phenomenon. However, it is a second order effect and not necessarily of consequence.

Though the relation of noise to error in spectrum reconstruction is interesting, it is difficult to know what a spectrum with sum squared error of 10^{15} versus a spectrum with sum squared error of 10^{-2} looks like. Figure 4-4 shows four noise scenarios and the visual spectrum reconstruction. The thick line indicates the input spectrum where the thinner line plots the reconstructed spectrum. The first image is where the noise is of equal magnitude to the input signal. Unsurprisingly, the reconstructed signal is completely indecipherable. The second figure shows a detected signal with random noise of magnitude 10^{-11}. Still no signal can be discerned from the reconstructed signal, though noticeably, the reconstructed signal and input spectrum are becoming closer to the same order of magnitude. With a single order of magnitude reduction in the additive noise, from 10^{-11} to 10^{-12}, the reconstructed signal jumps sharply from completely unrecognizable to very prominent. The signal in the third plot of Figure 4-4 is still quite noisy but shows a very clear peak at a wavelength very close to the peak of the input spectrum. At the minimum noise threshhold tested, at a magnitude five orders below the detected signal, the signal construction is shown in the bottom right of Figure 4-4 and is very good. There is still some ringing in the flat part of the spectrum. This ringing is an indication that the detected signal is not a perfect match for the superposition of wavelengths in the Gaussian-distributed signal, and also requires an oscillating amount of contribution from the wavelengths of the transmission matrix in the higher energy part of the spectrum. It is not clear from this example whether the ringing noise is a result of the matrix inversion algorithm, the high-frequency randomly distributed noise or an artifact of matrix multiplication.

To further investigate the effect of absorption profile on spectral reconstruction, I modeled a single Gaussian pulse through series of exponential absorption curves,

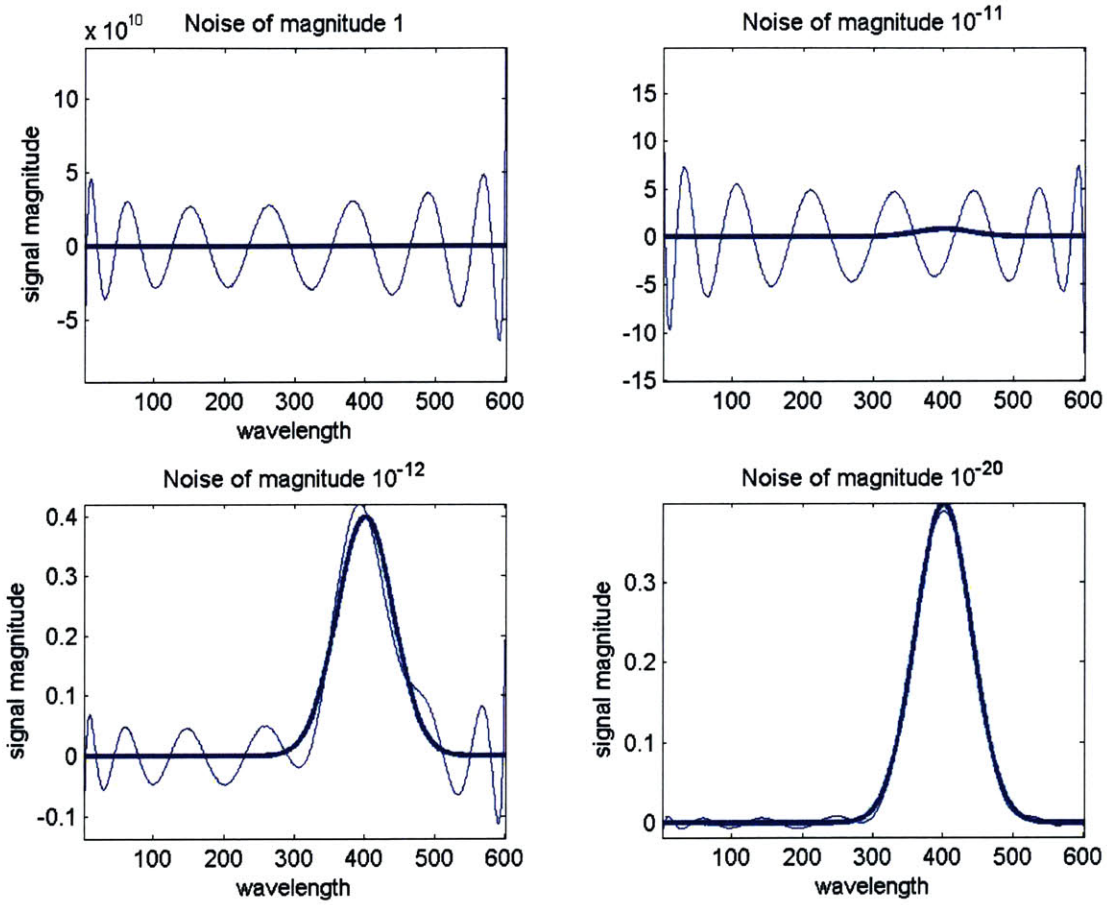

Figure 4-4: Spectrum reconstruction for four noise profiles

$\alpha(\lambda) = x^n$ with $n = 1, 2, ..., 8$. Visually, the spectrum appeared very reproducible through each of these films. Numerically, the error drops steadily as a function of increasing exponent, corresponding to an increasing slope over the wavelength range of the input spectrum. Error is minimized at an exponent of around 6, where it steadily increases again, corresponding to too little absorption by the thin film leading to noisy spectrum reconstruction. Variance of this Gaussian is 10 pixels, centered at a wavelength of $\lambda = 600$ nm. Resolution is 1 nm per pixel and the simulated thickness of the organic film is linear from 0 to 80 nm.

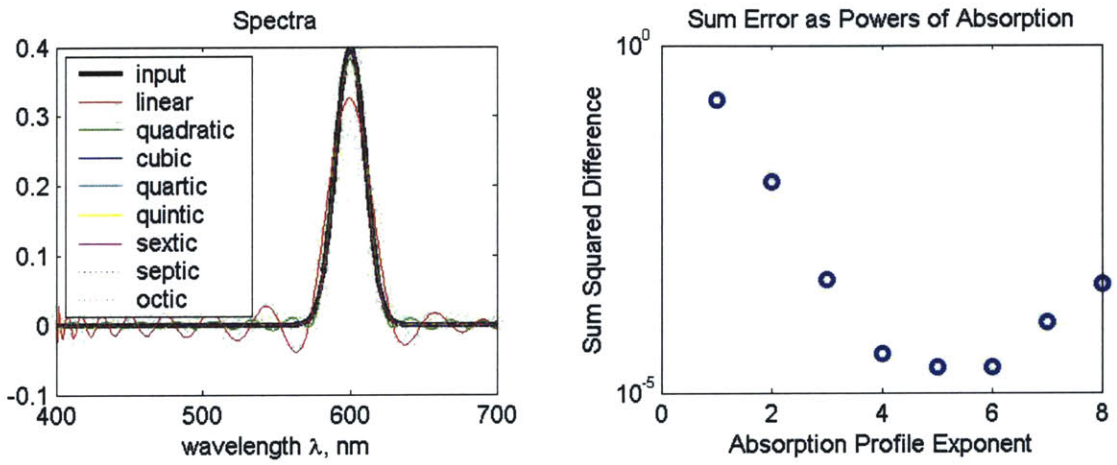

Figure 4-5: Reconstruction of a Gaussian for eight exponentials

Note that for exponents where $n = 3, 4, ..., 8$, very good spectrum construction is demonstrated. With quadratic absorption, a slight decrease in peak magnitude is observed and slight ringing is seen on either side of the peak. With linear absorption, significant peak reduction is observed and significant ringing. Nonetheless, throughout every absorption profile, good spectrum reconstruction is observed. The features of interest have good placement on the absorption curve in each case, with neither too much nor too little absorption or difference from one detected signal data pixel to the next.

4.2 Spectrum Shape

Next we vary the shape of the input spectrum. Waveforms with up to four constructively added Gaussians were tested, over the full range of center wavelengths for each of the eight exponential absorption profiles. The simulated film thickness for each of these trials is linearly increasing from 0 to 150 nm. The absorption profiles are identical to those in Figure 4-5 with $\alpha(\lambda) = e^n$ for $n = 1, 2, ..., 8$. In Figures 4-6 and 4-7, it can be seen that spectrum reconstruction is on the whole very good, for each exponential absorption profile.

The most significant finding of these trials is that spectral features in the short wavelength section of the spectrum, specifically from 400 nm to about 550 nm, suffer much poorer reconstruction overall than features in the long wavelength half of the spectrum. This behavior will be explored further in section 4.3, for now I offer the explanation that absorption is much flatter and weaker for the short wavelengths than for the long wavelengths. Hence, the overall signal attenuation and the difference in attenuation from one wavelength to the next will be much less than in the lower energy section of the spectrum, where the absorption profile slope is sharp and absorption increases.

Additionally, one might expect sharper peaks and dips to be more difficult to resolve than flatter, less severe features. In general, this is not the case for these simulations. Regardless of the variance of a Gaussian, it seems to be very easily reconstructed in the low energy portion of the spectrum and very poorly reconstructed in the high energy half. Certainly if there is to be some error in spectral reconstruction, it tends to be in an inaccurate peak height or wavelength, indicating that these features are in fact more difficult to resolve than the constantly-sloping regions, but it is the presence of the peaks rather than their sharpness that creates the error. Placement of these features in comparison to the corresponding absorption profile wavelength is the critical factor.

The effect of sharp peaks on spectral reconstruction take significance in some instances. For example, if there is an abrupt slope in input signal magnitude over the

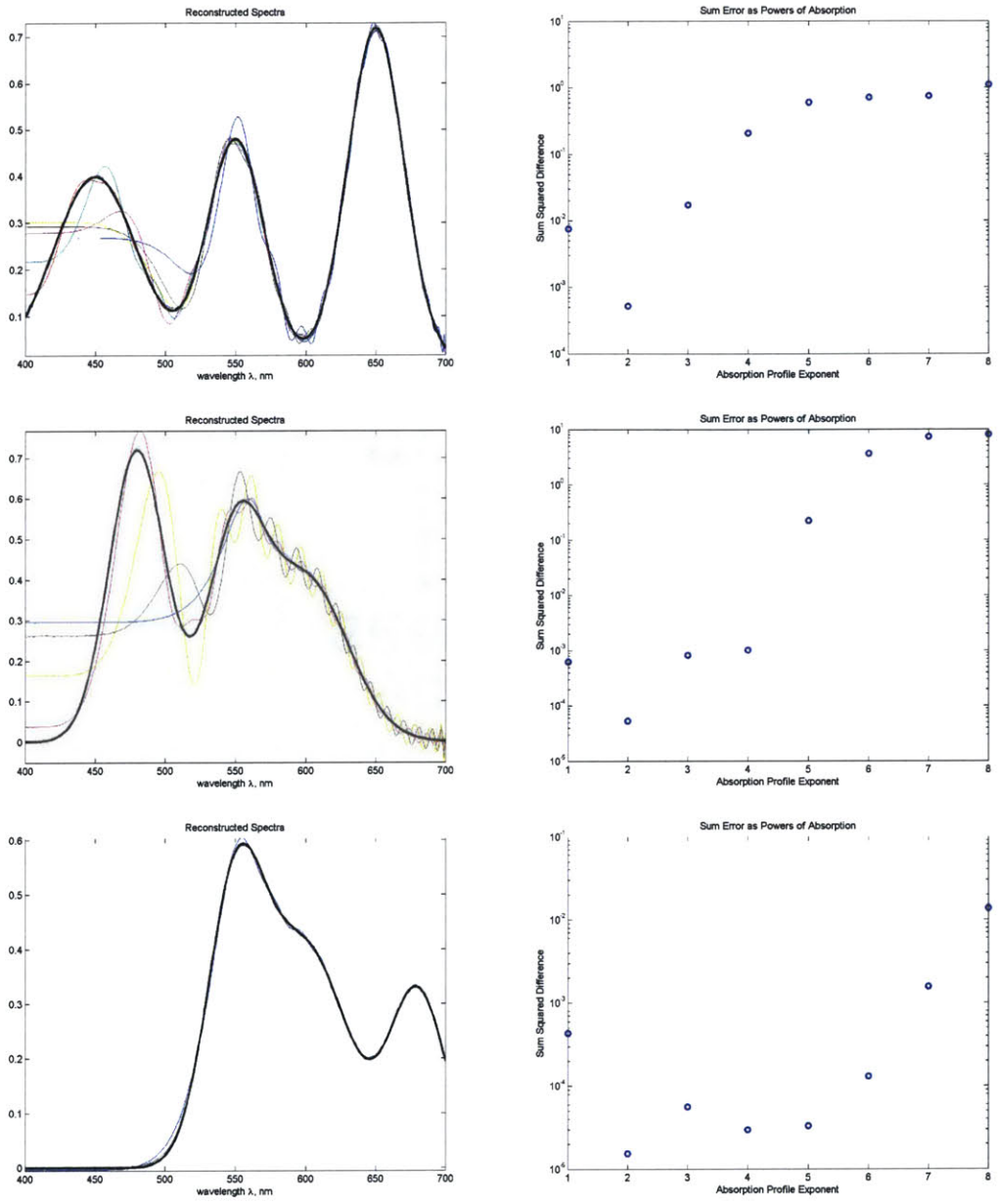

Figure 4-6: Reconstruction of sums of Gaussians for eight exponential absorption profiles

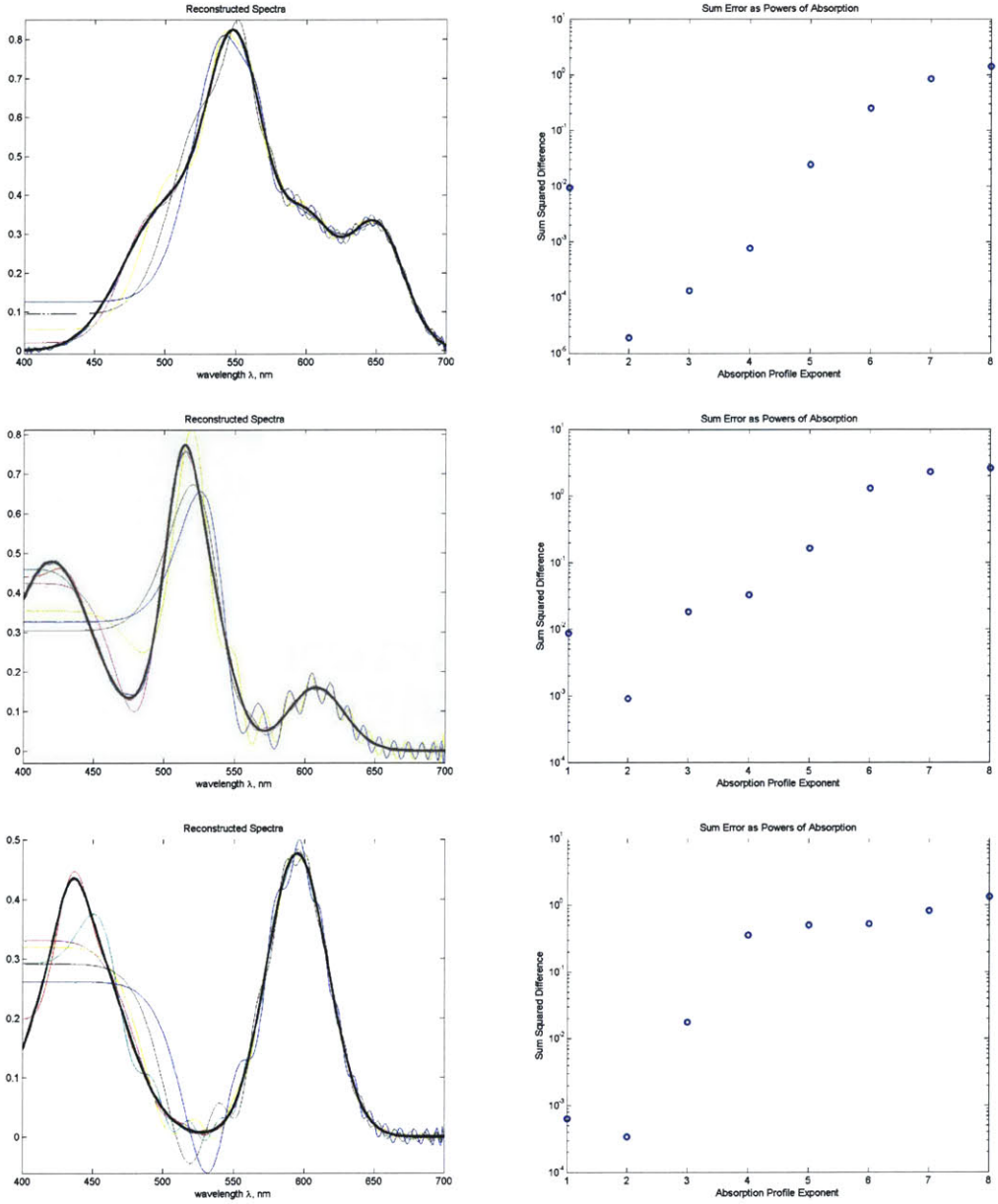

Figure 4-7: Reconstruction of sums of Gaussians for eight exponential absorption profiles

wavelengths $\lambda = 400 - 500$ nm, significant reconstruction error is observed and even residual ringing effects for the remaining flatter portions of the spectrum. This effect is seen in Figure 4-6 (b) and Figure 4-7 (b). Figure 4-6 (c) however, indicates that a sharp peak in the latter half of the spectrum does not induce the same reconstruction error as those for which the peak is over short wavelengths. This effect is especially noticeable in Figures 4-6 (a), 4-7 (b) and (c) where the first Gaussian is extremely poorly reconstructed, not even demonstrating a peak for many of the absorption profiles, while reconstruction of the second peak is recognizable though not clean. Any longer wavelength features are well characterized. Note especially the progression from poor to clean reconstruction in the three peaks of Figure 4-6 (a).

4.3 Film Thickness

So far we have determined that absorption spectrum shape, white noise, and the wavelength match between interesting spectrum features and absorption are all important in achieving the best reconstructibility of the input spectrum.

Further we posit that film thickness and gradient have a significant effect of reconstructibility. As of yet, it is still unclear is what precisely leads to reconstructibility, whether it is slope of the absorption profile or the absolute thickness of the organic film at any one point that creates favorable or unfavorable conditions.

We begin by testing the absolute thickness of the film, holding input spectrum and absorption profile constant. The input spectrum used is a Gaussian of variance 30 nm and center wavelength of $\lambda = 500$ nm. Absorption profile is exponential $\alpha(\lambda) = x^4$, $x = \frac{\lambda - \lambda_{min}}{\lambda_{max} - \lambda_{min}}$, which indicates a function growing quartically from 0 to 1 over the specified wavelength range.

Intuitively we expect absolute thickness to have some effect. We expect an optimal range beyond which and before which signal reconstructibility suffers. By increasing the thickness, the absorption is also increased and the detected signal decreased. Therefore, a very thick film will almost uniformly absorb most of the signal, allowing an almost undetectable portion through. The difference from one wavelength to the

next will be nearly indecipherable as much of the signal is abosorbed in either case. On the opposite extreme, very little of the signal will be absorbed, regardless of wavelength, and little difference is again observed leading to poor and noisy signal reconstruction. Clearly, the absorption profile of the organic film contributes also to the strength of the detected signal and thus will be held constant for this first simulation.

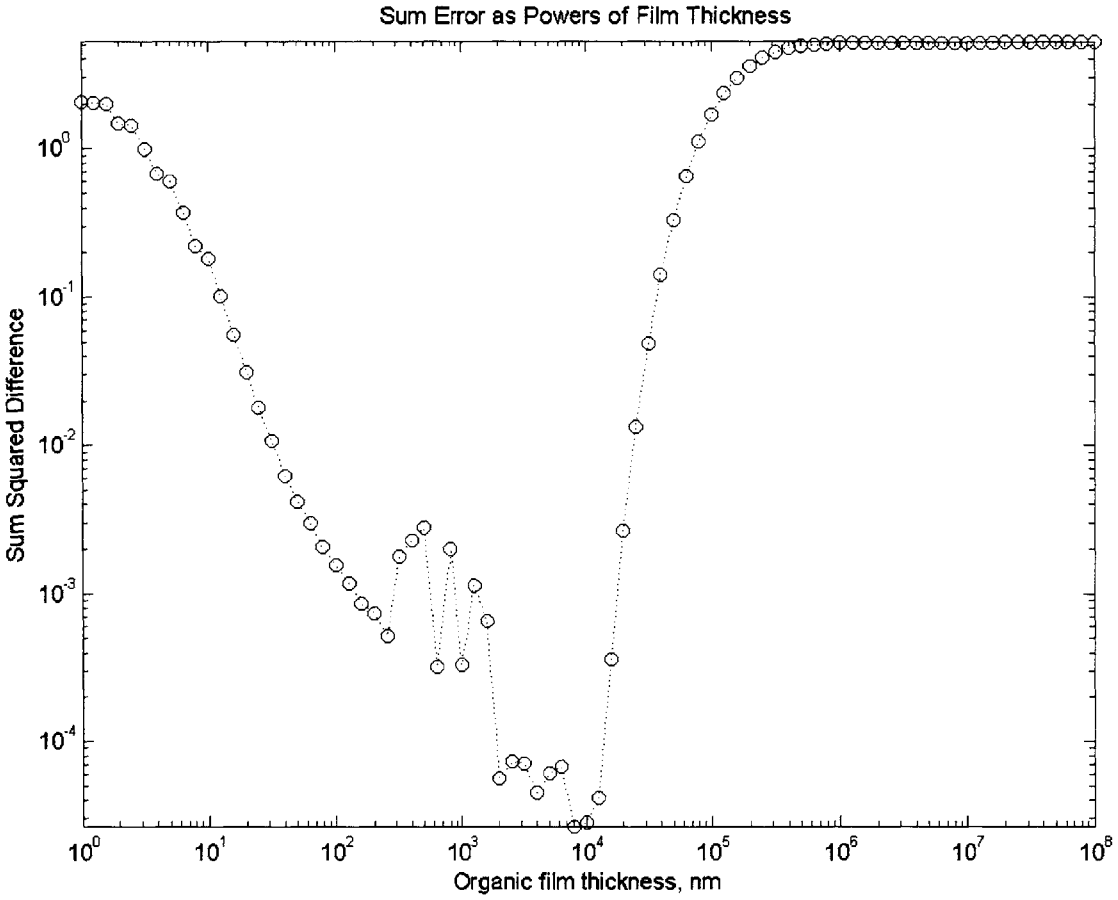

Figure 4-8: Reconstructibility by film thickness for $\alpha(\lambda) = x^4$

Results of the numerical experiment are shown in Figure 4-8. As expected, the trend of reconstruction error plotted against film thickness shows an optimal region bordered by two undesirable regimes. Illustrated on the left is the condition where too little absorption leads to little detected signal difference and poor spectrum reconstruction. On the right is the case with too much absorption and little detected

signal information. In the center is a good balance where good signal reconstruction can be expected.

Since we expect this behavior to be strongly dependent on absorption profile, we run the thickness optimization simulation for each of eight absorption profiles, where $\alpha(\lambda) = (\frac{\lambda-\lambda_{min}}{\lambda_{max}-\lambda_{min}})^n$ and $n = 1, 2, ..., 8$. Indeed we find sum squared difference to be strongly dependent on absorption function for a set film thickness. The set of profiles includes a number of physically interesting behaviors.

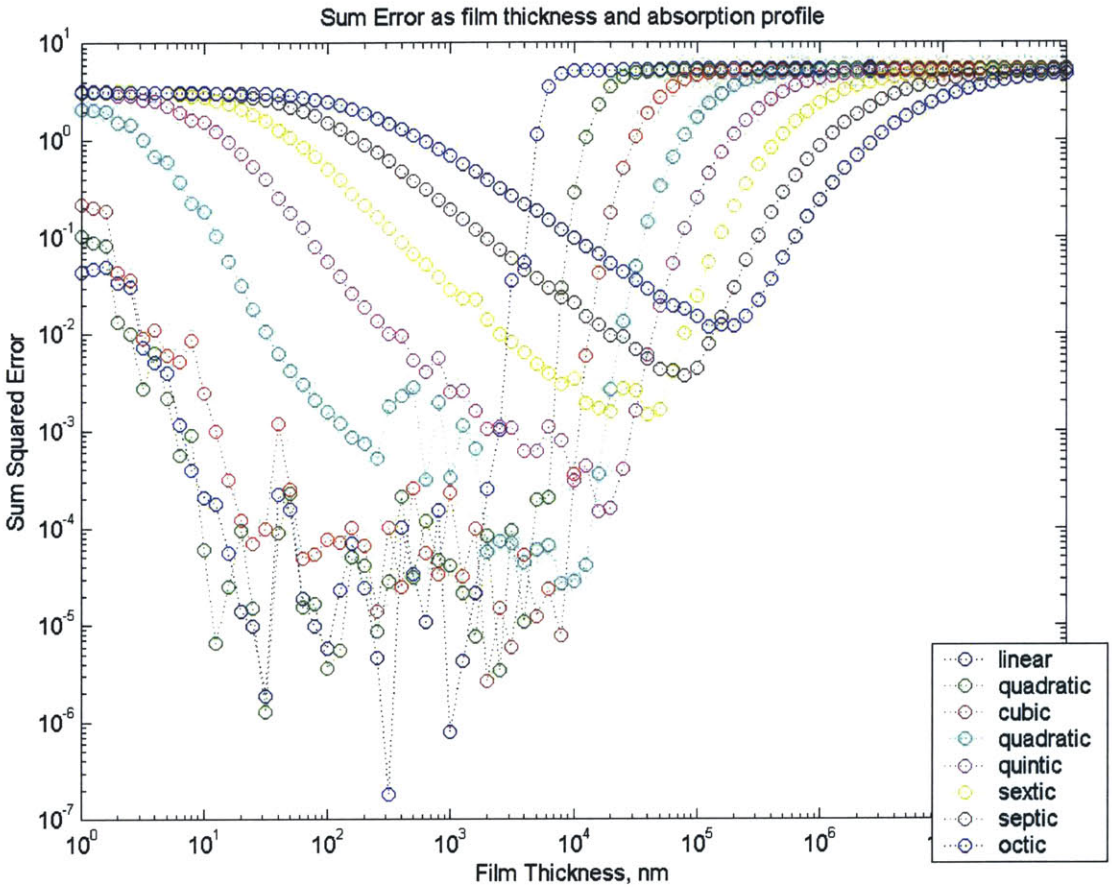

Figure 4-9: Reconstructibility by film thickness for $\alpha(\lambda) = x^n$ where $n = 1, 2, ..., 8$

Note first that the curves progress from linear to octic, left to right. This trend supports our assertion that the lower bound of optimal film thickness is due to insufficient absorption. The linear film with the largest relative absorption will overcome this first and the octic profile will reach it last. In addition the upper bound on the

optimal thickness regime is ordered from linear to octic, confirming the same trend. The linear film will saturate an optimal absorption before the octic absorption profile film.

In addition to proper ordering, it is clear that there is some structure to these curves. Specifically, the lower thicknesses have an exponentially decreasing dependence (linear on the logarithmic scale) on film thickness, up to the optimal point. Then the sum squared error becomes double exponentially increasing (exponential on the logarithmic scale) of varying slopes up to a maximum error, equal for all absorption profiles.

Such a regular distribution indicates a physical relation behind the trends. Noting the tradeoff between exponential and double exponential trends, we are pointed to the transmission/absorption equation, Equation 2.2.

$$T(\nu) = e^{\gamma(\nu)b} \tag{4.2}$$

Here $\gamma(\nu) = -\alpha(\lambda)$ and b represents the optical path length through the media. In our case $b \simeq t$, the thickness of the organic film. Clearly, if there is an optimal transmission value T that will minimize reconstruction error, there would be a tradeoff between α and film thickness t. Looking again at Figure 4-9, we note that the small values of thickness are in the single exponentially diminishing regime. By our proposed equation solution, this would be the regime where the thickness term t dominates the optimization. Once a minimum value is reached, then the double exponential behavior dominates. Mathematically, the absorption $\alpha(\lambda)$ corresponds, indicating that this term is what creates increasing error for larger film thicknesses. Physically, this proposed relation is satisfying and convincing. Transmission in the thinner sections of film are dominated by increasing thickness, while absorption is relatively small. In the thicker sections, the exponential absorption becomes more significant than thickness alone, and the trend for transmission is exponential (on a logarithmic scale). Additionally, we would expect the balance between film thickness and absorption to be reached sooner for the linear case and last for the octic profile,

which is in agreement with the numerical experiment.

To further confirm that reconstructibility is somehow related to optimal transmission value, we examine the points at which a minimum error is achieved for each of the absorption profiles. Looking only at the optimal point between thickness and absorption for each profile, we record thickness of the film for this minimal error. All the numerical experiments use the identical input spectrum, a Gaussian of variance 30 nm centered at 500 nm. Knowing the absorption profile, the optimal thickness and the input spectrum, we are able to reconstruct

For a first order examination, only transmission at the peak wavelength is calculated.

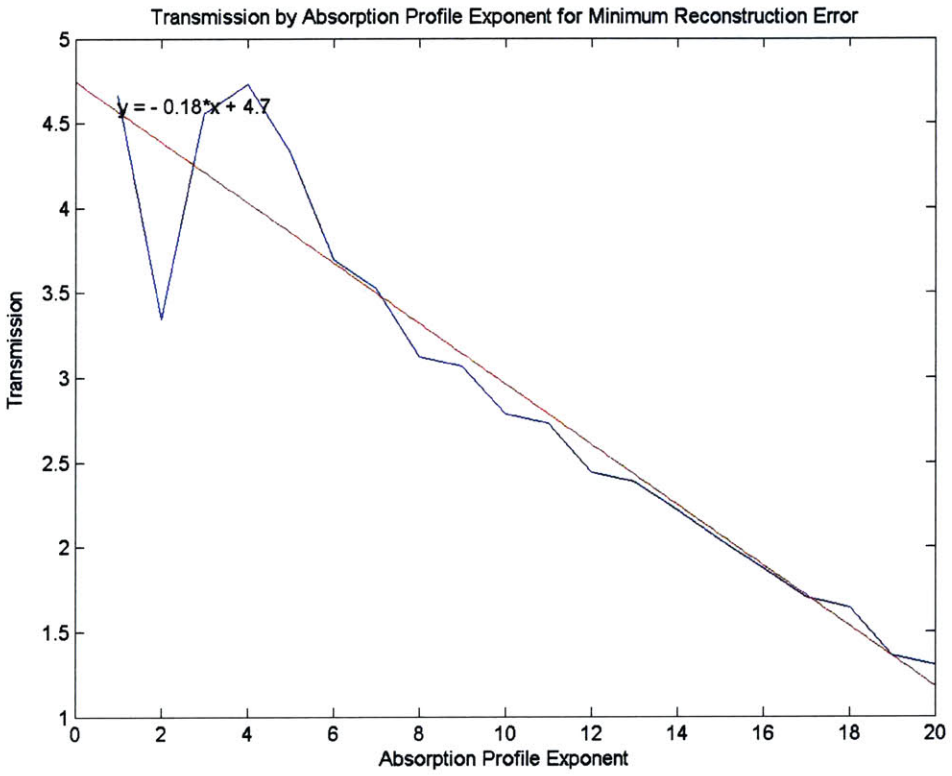

Figure 4-10: Transmission at Reconstruction Error Minima for 20 Exponential Absorption Profiles

For a more complete analysis the integrated transmission for the entire spectrum for each absorption profile should be considered and scaled for each thickness. How-

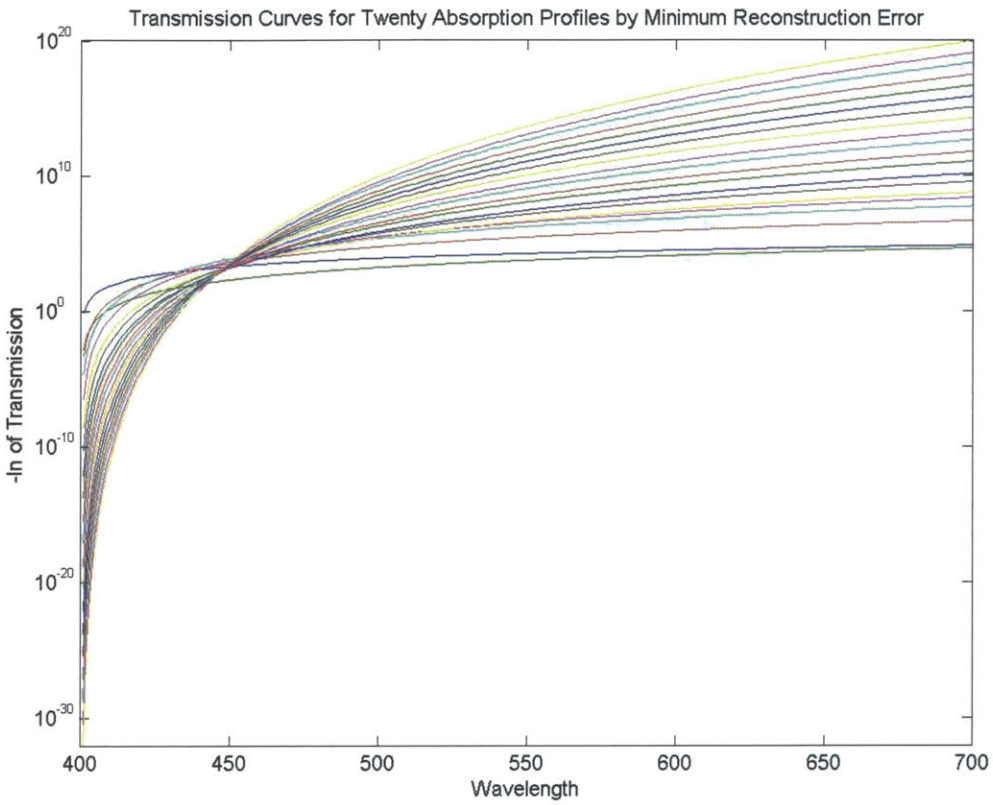

Figure 4-11: Transmission Curves for Twenty Absorption Profiles at the point of minimum spectral reconstruction error

ever, since our input spectra is consistent for each absorption profile and thickness, the values should scale with respect to peak wavelength, and our approximation is valid.

As an additional step in determining whether an "optimal" thickness exists, I plotted the transmission curves for each of twenty absorption profiles at the thickness of minimum reconstruction error. As a simple visual, there appears to be no magical number. The input signal is a Gaussian centered at 500 nm with a variance of 30 nm. The optimal transmission range can be extracted from Figure 4-11 as approximately 10^5 to 10^{10} arbitrary units.

The main lesson learned from this examination of thickness optimization is that for optimal device performance, film thickness should be calculated, and if the device isn't performing as well as otherwise would be expected, too thin or too thick a film are factors to consider.

4.4 Wavelength Optimization

The next test is shifting a uniform Gaussian input across the entire wavelength range and evaluating the sum squared error as a function of peak wavelength and absorption profile. As expected, error is highest for short wavelengths and for higher absorption profile exponents. In Figure 4-12 we see an interesting dependence of peak wavelength and absorption profile on sum squared error. Namely, the error takes a form very similar to the profile found for film thickness for very short wavelengths, possibly indicating a relationship, and then drops sharply as peak wavelength becomes more centered. For higher exponents of absorption profile, the sum squared error is especially high and spectral reconstruction poor. For most values, and for lower exponents of absorption profiles, error is negligible and spectral reconstruction is good. Alternately, it may be viewed that the absorption filter is too thin in the case of high profile exponent, but eventually as the peak wavelength shifts to a thicker section of the spectrum, absorption characteristics become in an operational range and spectral reconstruction again becomes possible. There is little new insight from

these measurements. We merely reiterate that film thickness is a critical factor in reconstruction.

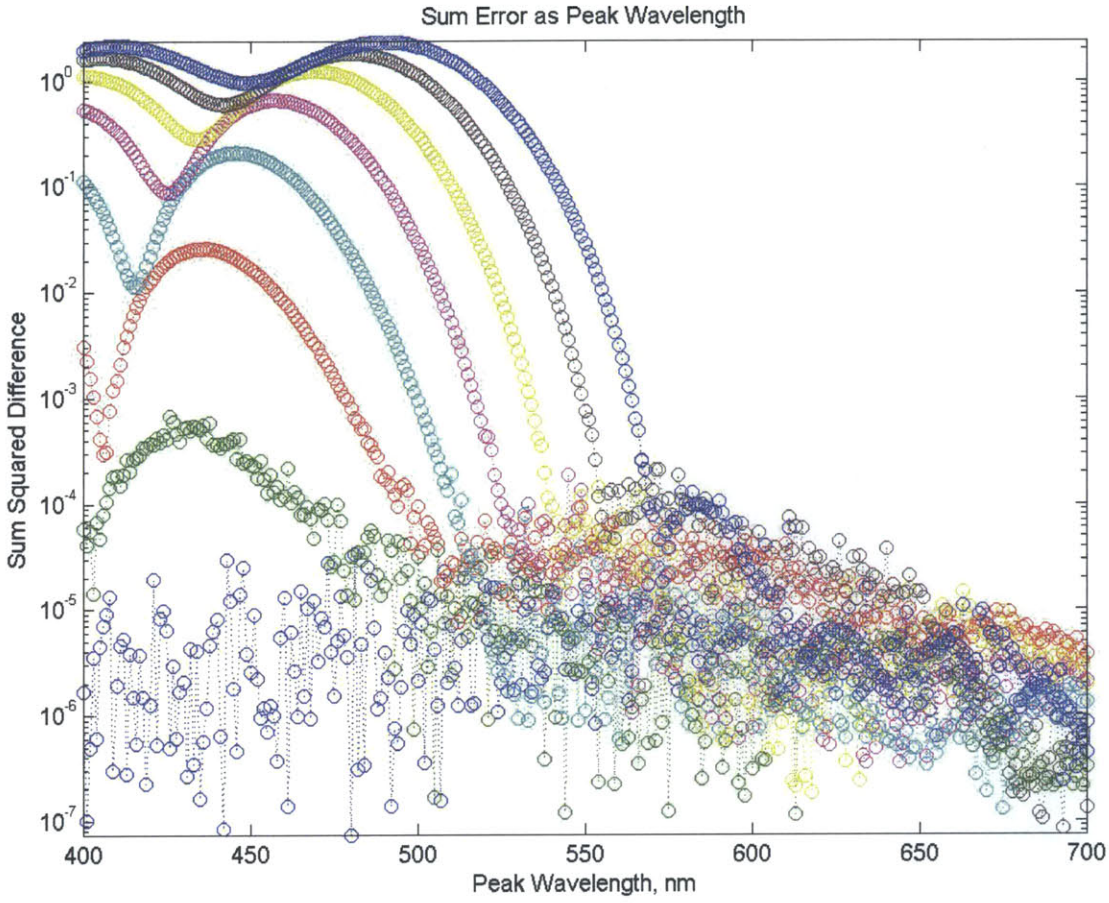

Figure 4-12: Sum Squared Error as a function of center wavelength for 12 exponential absorption profiles

4.5 Spectral Resolution

A final test of basic reconstructibility, we posit that sharp peaks, small variance and high finesse peaks will be more difficult to reconstruct than broad, slowly sloping spectra. If the magnitude of the input spectra differs greatly from one point to the neighboring discretely-resolved point, we imagine that there will be more extreme differences in the transmission matrix, making the measurement more sensitive and more prone to error.

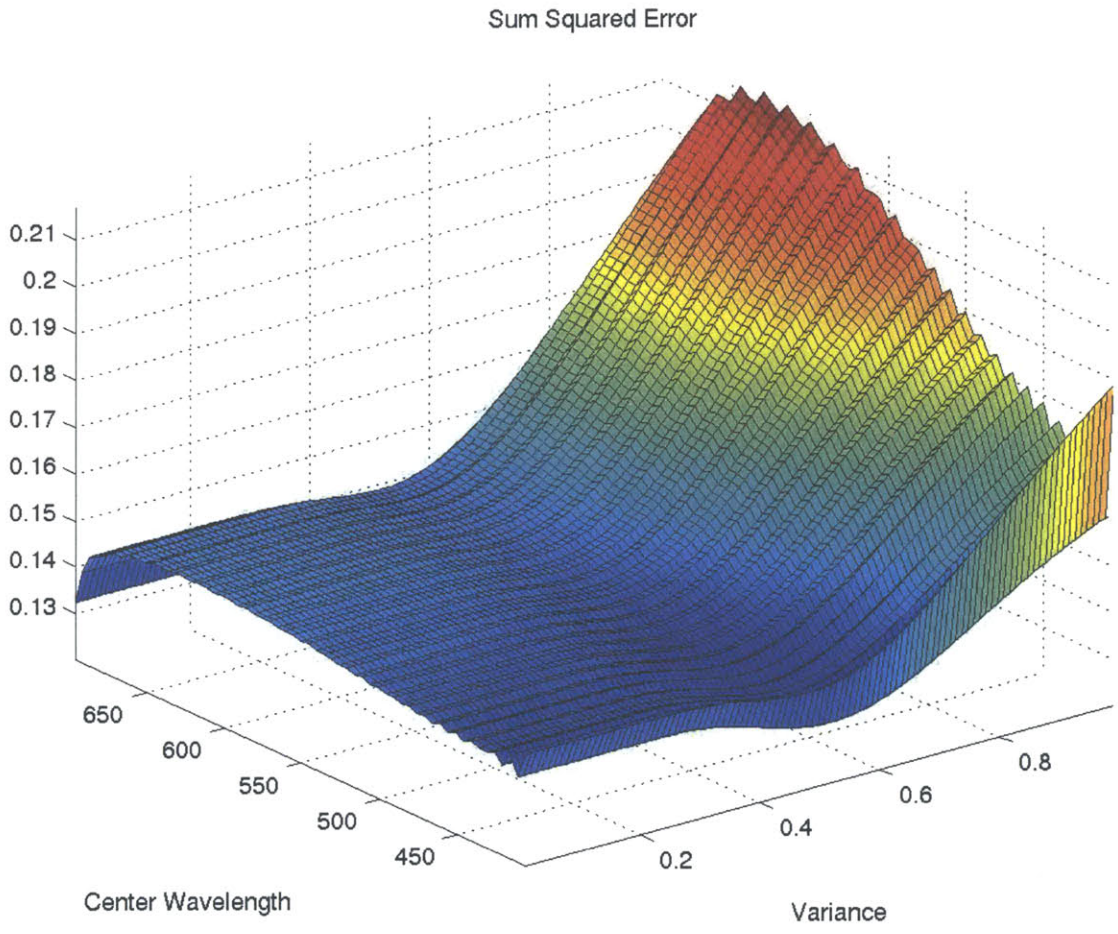

Figure 4-13: Sum Square Error of input Gaussians of varying center wavelength and variance for a quadratic absorption profile

The numerical experiment consists of changing the variance of a single Gaussian input from a single pixel of 5 nm to a broad 60-pixel 300 nm variance, at each center wavelength in 5 nm increments from 400 to 700 nm. The results are shown in Figure 4-13. For extremely narrow spectra, those with variances of up to about 0.4 or 40 pixels, reconstruction is excellent. Though generally counterintuitive, in the case of a quadratic absorption profile, this simulation indicates that broader pulses are more difficult to reconstruct.

The absorption profile used for this simulation was quadratic. The effect of wavelength on reconstruction error was fairly small for narrow input pulses and significant for broader pulses. Across all the variances, there is a trend that shorter wavelengths, which correspond to thinner absorption films, produce better spectral reconstruction. Again, it seems to me that the result here is not on the inherent worth of one shape or location of pulse but merely on the resulting thickness of the transfer matrix signal.

The range of sharp peak reconstruction raises the further issue of how well and for what applications the proposed spectrometer can reasonably be used. Certainly, if you are merely looking to see if a signal exists or not, all possible reconstructions shown here are adequate. If shape or relative magnitude is desired, only a broad peak or broad variations (broad with respect to the number of wavelengths sampled) can be detected. Even in the case of a broad spectrum, an appropriate film material and thickness should be chosen to achieve a monotonic profile over the wavelength region desired and not too little or too much absorption as to be able to resolve the desired features.

4.6 Monotonicity

It has been demonstrated in the above sections that both film thickness and absorption profile are significant factors in stimulating favorable spectral reconstruction. However, each of these models assumes a monotonic spectrum. Physically speaking, many organic thin films have monotonic absorption spectra over a short wavelength range, with the larger trend being non-monotonic. It is thus a very useful bit of knowledge

whether or not the entire wavelength range can be used for spectral reconstruction, or whether we must confine ourselves to the monotonic regions.

We have already seen repeatedly in the simulations above that spectral reconstruction is possible over a wide variety of monotonic transmission matrix operating characteristics. We begin here then with the non-monotonic case.

In Figure 4-14 we begin by testing a non-monotonic absorption profile. With a linear thickness profile, this characteristic alone yields a non-monotonic transmission matrix. The result is very intuitive. In the case of the small variance input Gaussian, the reconstruction is poor. Rather, the reconstruction is ambiguous. The Slim Format algorithm is able to reconstruct a pulse at the location of the input. However, due to the non-monotonic matrix and the similar absorption characteristics on the opposite side of the spectrum, the algorithm cannot distinguish between a pulse centered at 600 nm and one centered around 450 nm. The result is an oscillatory reconstruction with large signal magnitudes on either side of the absorption profile where transmission characteristics are similar. The second "ghost" peak is not exactly opposite the original at 450 nm, as the film thickness also plays a role in transmission characteristics. A slightly thinner film on the short wavelength end of the spectrum, although counter-intuitive, blue shifts the reconstructed peak. An input Gaussian on the side of the absorption peak with thinner film thickness we would expect to yield a red shift.

In the case of a broad peak and a non-monotonic absorption profile, as in the right figure of Figure 4-14, spectral reconstruction also fails. Simply put, the transmission data does not correspond on a one-to-one basis between the transmission matrix and the data signal. Therefore, the algorithm cannot match or reconstruct basis vectors.

Next we attempt a slightly less non-monotonic absorption profile. We create a non-monotonic profile, but place the characteristics of interest around the section of the spectrum that is monotonic. We attempt to determine how much non-monotonicity the reconstruction algorithm can tolerate, if any.

Figure 4-15 depicts very good but not theoretically perfect spectral reconstruction for both the narrow and broad Gaussian inputs. The non-monotonic region of the

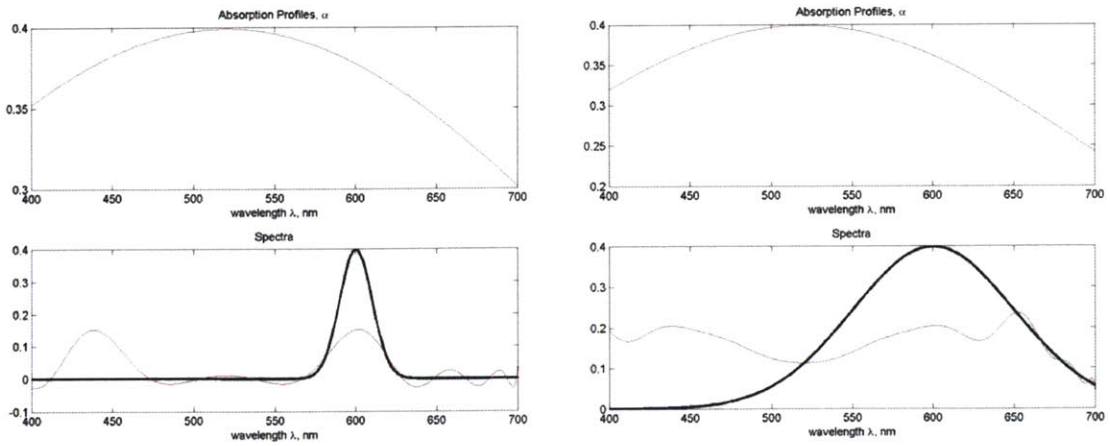

Figure 4-14: Reconstruction of a narrow and broad input Gaussian for a non-monotonic absorption function

transmission matrix seems to have little effect. However, considering what a negligible fraction of the input spectrum overlaps with the non-monotonic transmission matrix region, the result seems to indicate that though basic signal shape and spectral location can be reconstructed, the algorithm is extremely sensitive to non-monotonicity.

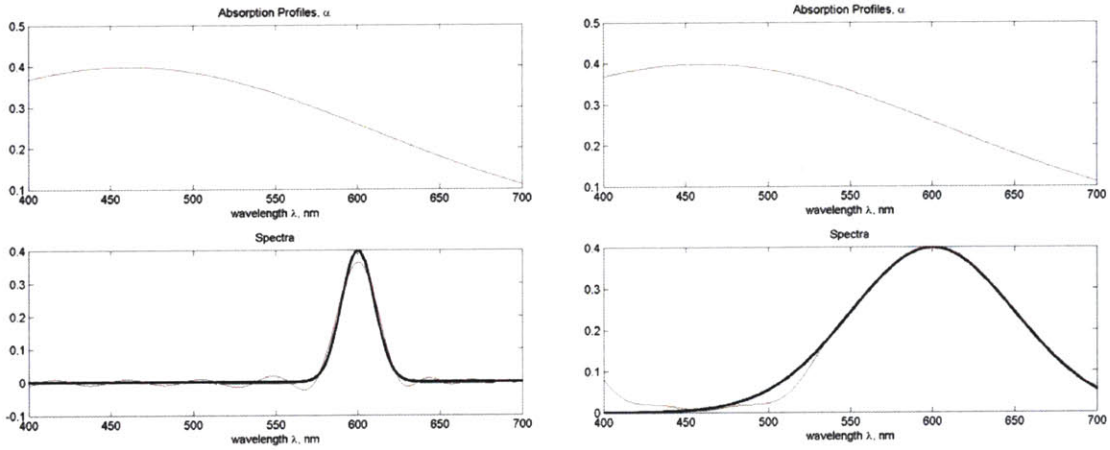

Figure 4-15: Reconstruction of a narrow and broad input Gaussian for a slightly non-monotonic absorption function

In this simulation in particular, I would argue that in the case of both the narrow and broad input Gaussian and also for the narrow Gaussian in Figure 4-14, there is absolutely no signal in the non-monotonic region. Rather than input spectral overlap with non-monotonic regions, what is important is the regions in the transmission

matrix that are similar anywhere in the matrix to regions where there are interesting spectral input characteristics. Although there is no signal in the wavelength region where the transmission matrix is non-monotonic, the transmission matrix looks similar at these wavelengths as it does to wavelengths where there is input signal. It is the similarity of transmission matrix traces that confuses our algorithm, indicating that the Slim Format algorithm is only as powerful as the information contained in the transmission matrix.

4.6.1 Optimization

To determine how robust the algorithm is, I measured the sum squared error as a measure of quality of spectral reconstruction, while scanning both the absorption function and the input Gaussian as functions of center wavelength and of variance. The results are shown in Figures 4-16 and 4-17.

The first figure, Figure 4-16, holds the input Gaussian constant and scans the absorption profile as a function of center wavelength across the visible range and variance from almost nothing to the full 300 nm width of the visible spectrum. The results indicate that variance is much more a determining factor than center wavelength for a particular constant input Gaussion (550 nm center wavelength and variance of 0.3% of the visible range, which translates to 90 nm).

The general trend of the spectral reconstruction is physically intuitive. In the variance dimension, spectral reconstruction is good for small values, dipping even slightly around the half way mark, where signal reconstruction becomes symmetric, then rising sharply as variance increases. My guess is that at small variances, there is plenty of variation in the transmission matrix. Whereas at higher variances, the absorption does not change much – it is uniformly high, and little signal variation leads to poor signal reconstruction.

As for the wavelength dimension, clearly the edges of the considered spectra will have less error, as they hold less information with much of the absorption profile edging off the region of interest, again making the transmission matrix monotonic. The divot in the center of the error trend is again due to symmetry. It is difficult

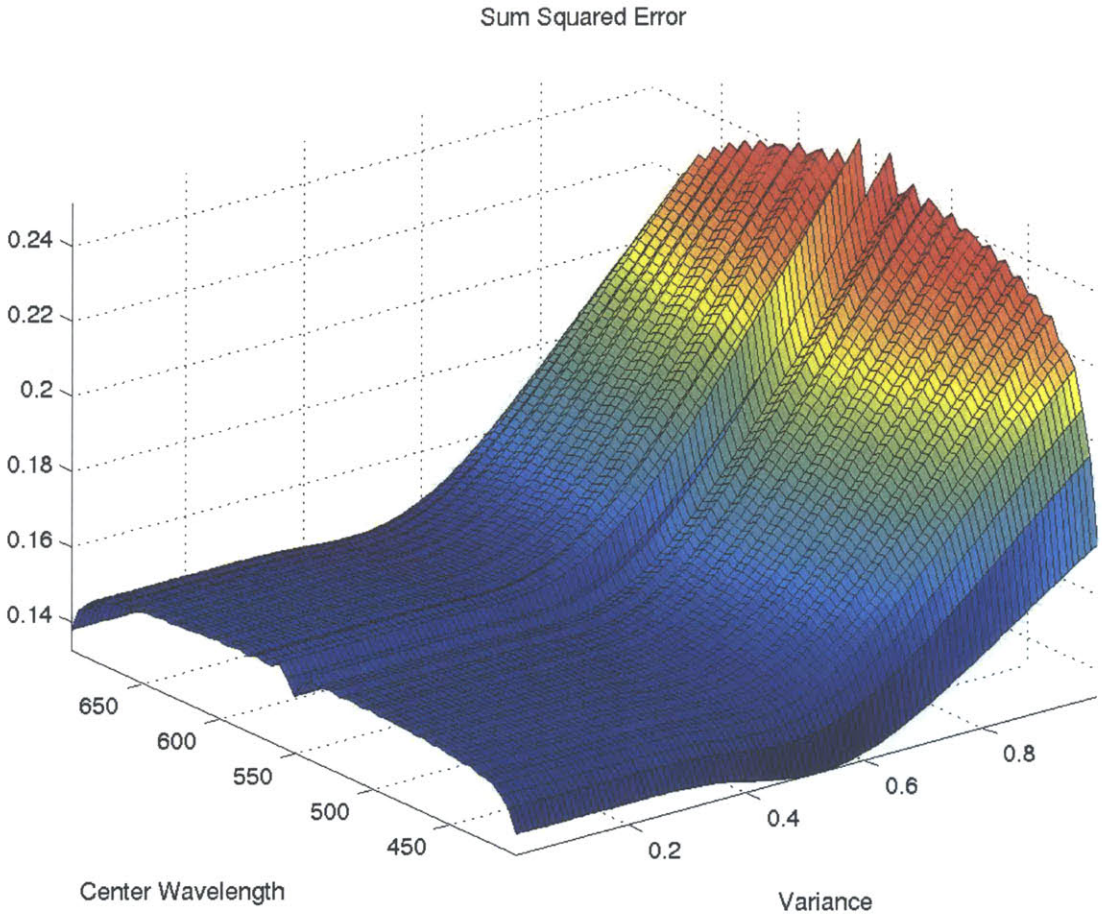

Figure 4-16: Reconstruction Error for varying Gaussian absorption profile center wavelengths and variances, for a fixed Gaussian input spectrum

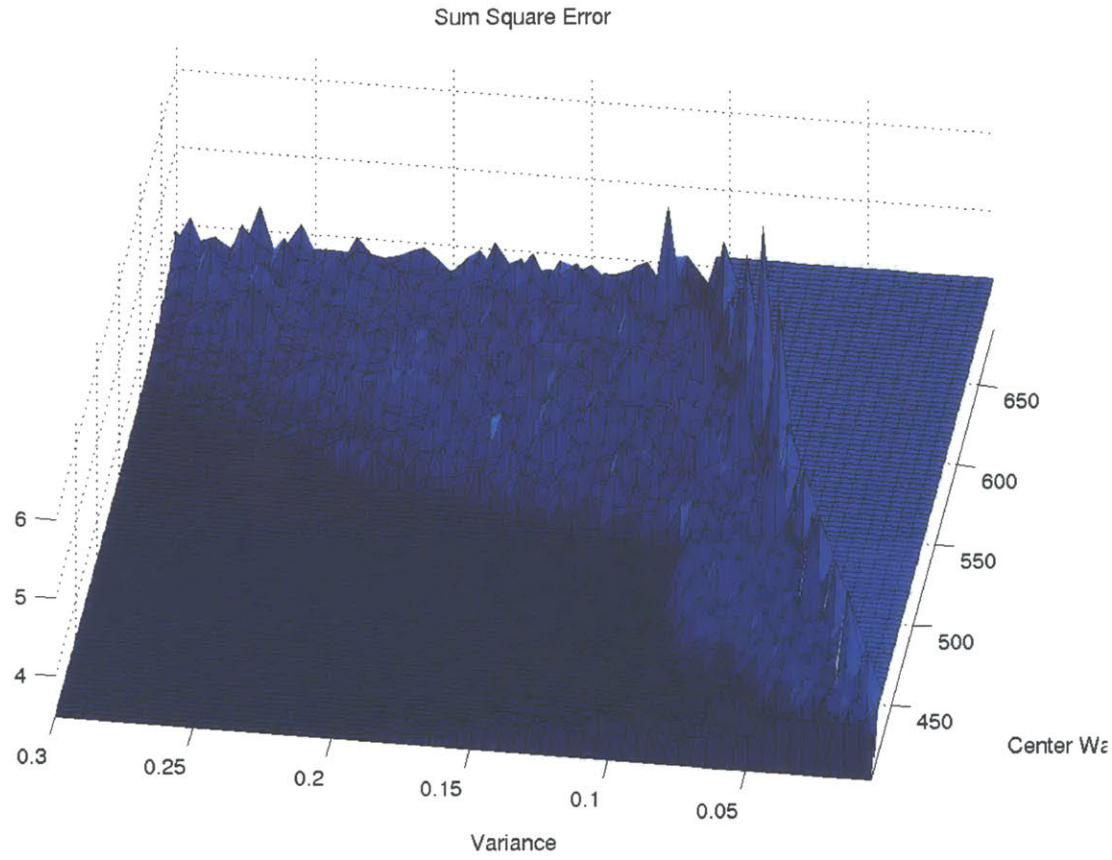

Figure 4-17: Reconstruction Error for varying Gaussian input spectrum center wavelengths and variances, for a fixed Gaussian absorption profile

to make an error around a symmetric absorption profile when both solutions to the reconstruction problem happen to be correct, due to precise symmetric placement of the absorption profile and input Gaussian.

The second figure, Figure 4-17, the absorption profile and thus the transmission matrix, is held constant and the input Gaussian is varied both in center wavelength and variance. It is plain to see that the results of this experiment yield three distinct regions. The first region being the case in which there is very little error and very good spectral reconstruction in the lower left section of the plot. This favorable region yields a combination of low wavelengths, likely corresponding to thin films, and high input variance. In general it is reasonable to assume that broader input spectra may be easier to reconstruct due simply to the small variations from step to step in the data signal. However, this detail cannot explain the general trend seen in Figure 4-17. There is a second distinct region in which considerable reconstruction error is observed, certainly making reconstruction less clear but perhaps not completely disabling it. This region typically has a high center wavelength, corresponding again to thicker films and the full range of input spectrum variances. The third region in which spectral reconstruction is not recommended is a clearly delineated region across all wavelengths from very small to small variances, from 0 to 90 nm.

Both of these optimization experiments serve to confirm further than transmission matrix qualities overall are more important than any contributing factor. Absorption profile is only important in that it contributes to the transmission matrix. The same is true for peak wavelength, variance, film thickness and other physical qualities of the spectrometer system. All of the information used in the reconstruction algorithm is bound tightly into the transmission matrix. It is a powerful tool. However, great care must be taken such that all the information relied upon in this matrix is accurate and unambiguous, else the reconstruction results certainly will be inaccurate or ambiguous.

Chapter 5

Experiments

The experimental procedure consists of two distinct and independent steps: constructing a prototype spectrometer device and testing the device, including considerable data processing.

5.1 Spectrometer Design

The idea for the device, described in Section 1.5.2, is a simple concept of an absorbing thin film, with the absorption a function both of overall intensity and also of wavelength, and a detector. Several configurations are possible, the only restraint being that multiple data points need to be taken for each spectral signal to be reconstructed. Potential device structures considered were array devices, also described in Section 1.5.2, as well as both linear and angular translation configurations. We decided that the linear translational device structure was the most suitable, but the array photodetector structure is also described.

5.1.1 Translational Single-Photodetector Spectrometers

The translational device employs a single photodetector and a organic thin film which varies in thickness in only one dimension. The film is then scanned across the photodetector and a time-dependent signal is recorded, as a function of film thickness.

The advantages of this device configuration are many. Primarily, it's simple in comparison to the array device. Much of the objective of the Slim Format Spectrometer is to create a low-cost device that practically any scientist may construct from materials already in her lab. Very likely, many scientists interested in optics own a visible-light photodiode, whereas relatively few would own a CCD photodetector array. This configuration also allows for a much simpler organic film gradient. Rather than assuring a distinct film thickness for each cell in an array, this device requires only a distinct thickness for each data point across a strip. Such thickness gradients may be created simply by tilting a substrate in an evaporator, rather than creating a substrate mask system with multiple evaporations. Most importantly, removing the constraint of a complicated deposition step, opens the choice of absorption material to a wide range of materials. Any thin film which absorbs in the visible spectrum, or whichever spectral range you intend to measure, may be considered. Again, a scientist is much more likely to have some sort of variable thickness absorption film, such as a plastic gel or piece of glass, than access to an expensive vacuum system. It was our intention by choosing the translational configuration to find an organic film that absorbs across the visible spectrum which would not be as sensitive to light, heat, oxygen and moisture exposure.

5.1.2 Array Photodetector Spectrometers

To clarify, the array device would employ a CCD-like array of photodetectors, with a variable-thickness organic film placed over the front face of the device. Alternately, a more sophisticated version of this array device allows a different thickness film to be grown directly on each photodetector, combining the detection and absorption processes into the same inherent device structure. As a general rule, the accuracy of the thickness of organic films grown in separate thermal evaporator systems may vary on the order of hundreds of Ångstroms (tens of nanometers). Films grown in the same trial in the same evaporator system, may vary on the order of tens of Ångstroms (nanometers). However, the general thickness trend will be preserved, regardless of absolute thickness, in a single growth process. The array photodetector device

structure was ultimately decided against for the prototype due to complexity. As opposed to the translation devices, the array configuration would require development of a high-resolution deposition mechanism. A preliminary run has shown that such a system is possible, but to ensure that each and every photodetector "pixel" is receiving light through a distinct, discrete thickness of organic, significant further experiments would need to be conducted. For the first demonstration of concept, it was thought that a simpler system would suffice.

Additional complications with the array photodetector device formulation include the necessity of the incident beam being spread over the entire detector array, in sufficient but not excessive intensity for data collection, as well as the susceptibility of organics to oxygen, light and moisture. Clearly, in order for each "pixel" to count in the resolution, the exact intensity before organic thin film absorption must be known. Additionally, organic films have a tendency to lose their opto-electronic properties with exposure to oxygen and moisture. They also lose their absorptive properties, saturating with continued exposure to light. A packaging system could reliably delay or prevent these effects. However, again, for a simple demonstration of concept, these issues would be avoided if possible. The additional variable of calibration for both incident intensity and unknown organic film thickness, and the variability of film properties due to deterioration in an oxygen- and moisture-rich atmosphere as well as bleaching with continued exposure to high energy photons, make this configuration even more computationally intensive than the translational systems, which leads to further uncertainty in device performance.

5.2 Prototype

The particular device we settled upon to test the Slim Format concept, uses a rotating disc mounted to a motor. The device follows the translational spectrometer configuration, with organic film variation across the angular θ dimension. The rotating disc allows for data collection as a function of time, which translates to thickness, across a variable thickness film. We chose the rotary as opposed to the linear configuration

as simple motors were thought to be more easily obtainable and much less expensive than linear translation stages. In theory, it would also be possible to collect data much more quickly as the motor spins, rather than resetting and scanning a linear stage.

Our choice of organic film again was selected for cost, convenience, and atmospheric robustness. Coincidentally, or perhaps simply demonstrating the prevalence of organic dyes, color inkjet printer dyes are organic. The dyes are stable in an oxygen- and moisture-rich environment, and withstand direct sunlight without bleaching. Fortunately, printing companies have perfected these dyes chemically. Additionally, they are optimized for use in the visible range. The dyes we chose were the Hewlett Packard #78 cartridge dyes. As for cost and convenience, I would be surprised if any scientist did not have access to a printer with dyes that absorb in the visible spectrum. Furthermore, as a substrate for this organic dye, we chose overhead transparency film. It is already a standard printing material and is transparent in the visible range. In fact, most transparency films have a cutoff wavelength of around 350 nm, where below this wavelength, all light is absorbed, so the film acts as a high pass filter for our experiment and for device operation.

5.2.1 Assembly

The first step in building your very own Slim Format Spectrometer is to obtain a stable substrate, on which to mount the filter you will construct. We ordered 3 inch diameter, 1/8 inch thickness circular discs from Edmund Scientific. Any mechanically stable, transparent, lightweight material that can be mounted to the motor or translation stage will work.

With cost in mind, I chose the Radio Shack $273-255$ High-Speed 12VDC Motor, quoted at $11,500$ RPM and 1.3A. In Boston, this device can be obtained for less than 5.00 USD. Although not a particularly stable motor, nor one that can tolerate heavy loads, it meets our main criteria of low-cost and easily-obtainable by ordinary persons. A photo of the motor is shown in Figure 5-1.

Mounting the substrate and substrate holder to the motor is a creative matter,

Figure 5-1: RadioShack High-Speed 12VDC Motor, 237-255

and often one of personal preference. I drilled a hole, using an ordinary household drill and the smallest bit, through the plastic disc, and then used epoxy, though super glue would hold just as well, to affix the substrate to the motor. Care should be taken to keep the substrate as close to perpendicular as possible, to minimize wobbling in rotation. Also, take care not to let epoxy cover any of the optically active area of the substrate, as epoxy will effect very different absorption characteristics.

5.2.2 Absorption Filters

To create the absorption filter, begin with a graphics editing program. I chose Adobe Photoshop 6.0, due to its ease of use and powerful gradient tools. For a linear gradient, I have found that the latest release of PowerPoint works satisfactorily. However, for a radial gradient, more sophisticated image editors must be chosen.

Photoshop has a number of premade gradients, though it easily allows you to specify a custom gradient, allowing choice of RGB or CYM color and the opacity, as well as a smoothness factor at each location along a linear gradient. This custom

linear gradient may then be applied in a number of geometries, one of which being radial, from 0 to 360 degrees, outward to the edges of the image from a point you specify.

Given that the dyes we have to choose from are of the cyan, yellow, magenta and black basis, I began with a linear thickness gradient in each of these colors. In order to get more interesting and varied transmission matrix characteristics, I fabricated a number of multicolor filters, based on CYMK dyes.

Additionally, it was found to be necessary to create an outer track for a second laser as a tracking mechanism. Since the absorption profile through the absorption filter changes with wavelength and organic film thickness, the trigger which tells the oscilloscope when a data set begins and ends cannot be counted on to remain stable. Consequently, a second trigger track is created which will remain stable with wavelength, such as a greyscale trace, as in Figures 5-2 and 5-3. Operationally, any pattern would work so long as the wavelength of the beam used as a trigger remains constant. Nonetheless, a second gradient is created easily with layers in Photoshop, or by adding multiple images one atop the other in PowerPoint or other more basic image editing software packages.

After the filters images are created, they are printed onto transparency paper, cut with scissors and mounted with tape or glue onto the motor and filter substrate. In one set of experiments I mounted the transparency filter directly to the motor shaft and noticed no degradation of results.

5.3 Experimental System

Once a prototype device has been created, the proper optics and detection electronics must be mounted around the device. As the device will be marketed mainly as simple and user-friendly, so must be the surrounding optics and electronics.

In a block diagram sort of fashion, we simply have an input spectrum, typically a white light source being shone through a sample you wish to test. In our case this sample is a glass filter. However, in laboratory use, this sample would likely be a

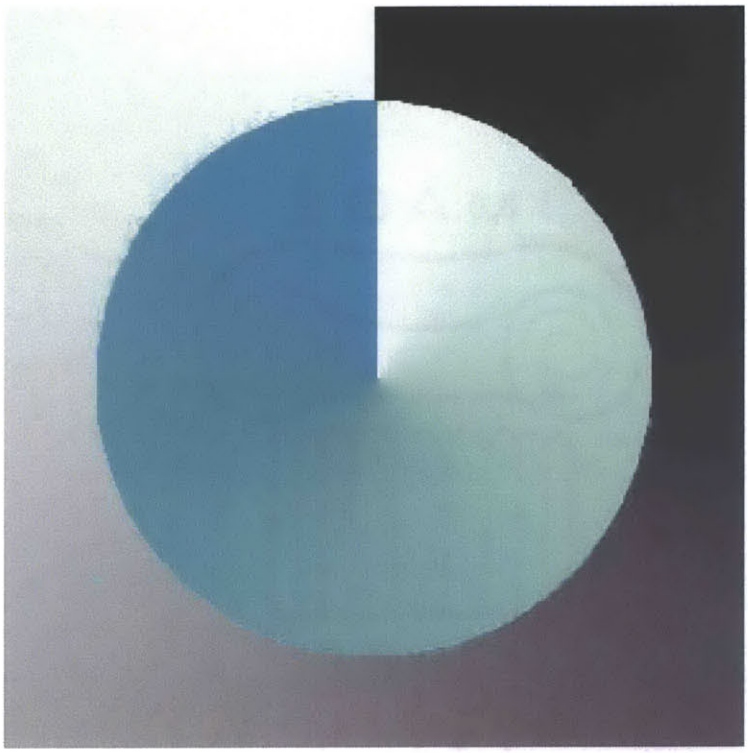

Figure 5-2: Single color absorption filter: Cyan with trigger track

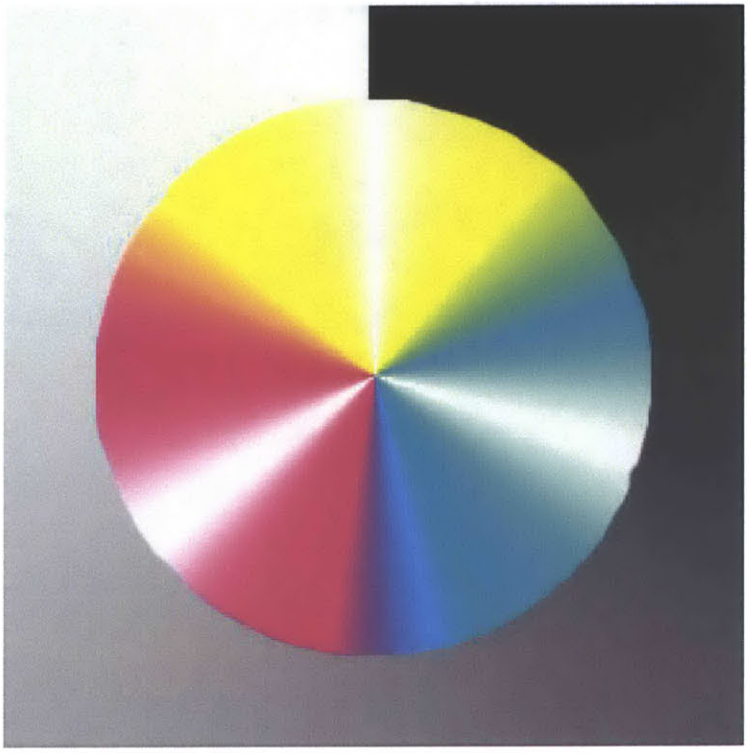

Figure 5-3: Multicolor absorption filter: Cyan, Yellow and Magenta with trigger track

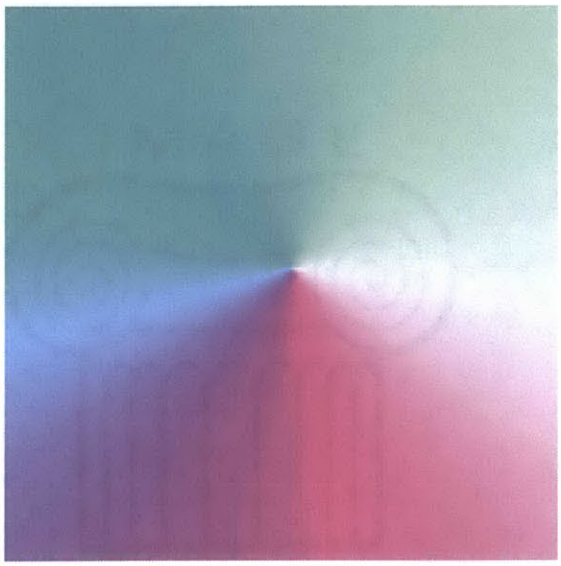

Figure 5-4: Bicolor absorption filter: Cyan and Magenta, blended together and graded from a transparent background

Figure 5-5: Multicolor absorption filter: Cyan, Yellow and Magenta

solution or the output of a fiber or waveguide. This light exiting the test sample is then absorbed again by our absorption spectrometer device, and collected immeditaely by a photodiode. The specifications for the photodiode can be tailored to the application, but it is desirable that the current produced for photons of each energy be somewhat linear and it is essential that the current produced be reproducible.

For practical matters of detection, the current from the photodiode is sent to a current amplifier where the current signal is converted to a voltage. An artifact of the current amplifier we used was that it is also inverting. For example, a large current will be converted to a negative voltage of high magnitude. A smaller current will be converted to a negative voltage of smaller magnitude. This inversion may be removed with processing, but we left it in the data, simply to avoid distortion.

The output of the current amplifier is then fed directly to a digital oscilloscope, where the time-dependent signal is detected. Due to the many sources of variation and noise in our system, such as wobble from the rotating absorption filter, and drift of the input signal, we chose to average the oscilloscope signal before saving the data. Actually, due to the time scale of drift of the various parts of the system, choosing an integration time was challenging. For example, the wobble of the motor varies from rotation to rotation, a very fast time scale over which an integration on the range of seconds could remove. However, on the scale of seconds, the frequency of the motor would drift, and the detected signal, in terms of pixels on the oscilloscope, would represent more or fewer thickness points. Additionally averaging over points which are not of the same thickness on the absorption filter leads to extremely flawed data. As a rule, I chose an integration time which removed the large deviations in signal detection, but was sure to maintain a constant motor frequency, ensuring that the signal was not greatly distorted, especially on the high and low ends where the thick and thin film signals could overlap with varying frequency.

There is a second signal sent through the absorption filter to a separate photodiode used as a trigger for the oscilloscope. I call this signal the trigger signal. The signal is created from a diode laser, so single-frequency light, through the absorption filter, when possible, along a separate single-color filter, adjacent to the main absorption

filter, as in Figures 5-2 and 5-3. The transmitted light from the second light source through the second absorption filter is collected by a second photodiode. The signal is sent directly to the oscilloscope where it is used as the trigger for the data signal. The trigger signal is stable, even as the input spectrum and absorption characteristics change. The signal is used to indicate when a complete rotation of the disk, and thus a complete pass around the absorption filter has been made.

The detected signal as well as the trigger tracking signal, after inversion by the current amplifier and averaging by the digital oscilloscope, was saved manually to a 3.5 inch floppy disk as a comma separated value (.CSV) database import/export file. These files were then imported to MATLAB via a standard function call and a small script, where the data, along with information about the absorption filter, may be used to reconstruct the input data signal.

5.3.1 Operational Detail

On a more detailed level, there are many items of note in the Slim Format test structure. For example, all the optics are mounted on a Newport floating optical table, including the stable white light source lamp. Before the spectrometer is even fit to be used, all optics must be mounted, aligned and the organic thin film absorption filter must be characterized, using not white light, but a monochrometer.

For alignment purposes, each system may have various constraints. Our system, for example had a fixed white light source, and other experiments on the optical table to work around, but the diode lasers and most of the optics, as well as the power sources and the current amplifier were portable. There will likely also be few locations where you can mount the spectrometer device in such a way as to fit two photodiodes directly behind the absorption filter and trigger filter where applicable. Logistics of swapping out a monochrometer for thin film absorption filter characterization and the test sample should be considered before mounting the entire assembly.

The test setup I used was primarily L-shaped, with the white lamp light source on one far end, with the monochrometer and test sample close by. A directing mirror turned the beam approximately 90 degrees, where additional lenses focused the beam

to the smallest spot size possible on the absorption filter. The spectrometer was mounted on an optical lens mount, with the motor clamped into place by a spring loaded mount. The photodiodes were placed in position with a post system of optical mounts. The output from the diodes were then run to the opposite side of the room via cables, to the current amplifier and digital oscilloscope.

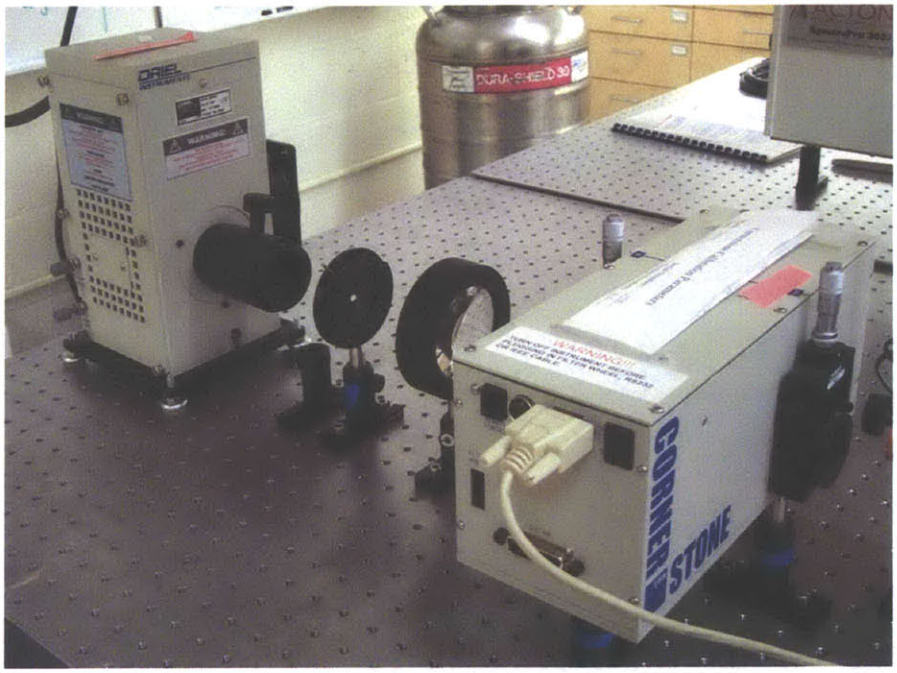

Figure 5-6: White light lamp, focusing and collimating optics and the Cornerstone monochrometer

5.3.2 Test Procedure

Before any experiments could be run, the thin film absorption filter needed to be characterized. For this processes, the white lamp was focused into the monochrometer input, in a standard method which is rarely adjusted. The input slit to the monochrometer was adjusted to allow maximum light intensity throughput, as opposed to sharpness of output wavelength. The output of a monochrometer is a somewhat fuzzy spot, having dispersed somewhat from the slit output. I focused it into the smallest spot possible, which was still somewhat taller than it was wide. I bounced this beam off an inexpensive silvered mirror, coated for visible wavelengths.

I then positioned the beam on the spectrometer such that the smallest thickness difference on the absorption filter would intersect the beam. That is to say, I arranged the photodiode at the top of the filter, such that the tall and thin beam would intersect only a thin section in the thickness and angular dimension, as opposed to arranging the photodiode on the side, which would have the beam intersect a larger range of thicknesses per absorption reading. In principle, so long as the beam shape is kept consistent, it doesn't matter for the measurements or beam reconstruction. However, for the case of resolution, a small incident beam spot size is desirable.

Next we set up the trigger track. A laser, independent of the white light source and monochrometer is mounted, close to the spectrometer, but not interfering with the monochrometer beam. Due to space limitations behind the spectrometer, it was just not possible to line up both the photodiodes along the same radius from the center of the motor. For example, if the filter in Figure 5-2 was used as the absorption filter, the trigger and the absorption cycles would only coincide if the photodiodes are set adjacent to one another. Logistically, the second photodiode was placed opposite the first, and the transfer matrix was adjusted after data collection. Throughout, the spectrometer is held in place by a lens mount attached to the motor base.

Before characterization of the absorption filter can begin, we need to connect the data signal photodiode, the diode behind the absorption filter of interest to the current amplifier. The signal from a photodiode is inherently nonlinear. If this signal was sent directly to the oscilloscope, it is possible that linear algebra would not yield accurate results. In turn the output from the current amplifier is sent to the digital oscilloscope via a BNC cable. The trigger photodiode output is sent directly to the oscilloscope. The magnitude of the trigger signal is irrelevant. The trigger signal needs only to be repeatable for each filter rotation, to track the position of the absorption filter with respect to the photodiode and the incident beam.

Once everything is connected, we turn out the room lights and begin characterization of the absorption filter. As is common with modern monochrometers, the wavelength is chosen via a computer control module. Additionally, LabVIEW drivers can be downloaded for remote control and automation of data collection from the

monochrometer and the digital oscilloscope. Ideally, a program would be written to scan the monochrometer across the entire wavelength range and record the oscilloscope output. However, our monochrometer did not respond reliably, to date, to the written LabVIEW instructions. We were able to set the wavelength accurately and reliably with the included software, but could not sweep. Fortunately, data collection is a relatively quick process. It takes more than the 20 seconds it would take if automated, but characterizing a single absorption filter never took me more than 30 minutes.

The wavelength is set on the monochrometer via computer control, starting at 400 nm, in whatever increments desired - I typically chose 20 nm, through 700 nm. The trigger signal was used to stabilize the data signal on the oscilloscope, digital averaging chosen with good judgment by observing signal variation in magnitude and frequency over time, and the data saved to floppy disk.

The process for taking an absorption trace with an unknown input, such as that through a glass filter or from a solution in the lab, is very similar. The monochrometer is bypassed optically, with silvered mirrors. The optical signal is routed around the monochrometer, through the test sample or doped glass filter, through focusing optics and a pinhole and onto the same optical path as formerly the monochrometer output. The identical trigger path and placement must be used, to ensure that the thickness measurements coincide accurately. The signal photodiode also must not be moved, to ensure accurate relative position to the trigger photodiode.

As previously mentioned, the optical Schott Glass filters, 2 inches by 2 inches, are mounted on a standard Newport 2 inch filter mount and inserted directly into the white light beam. Examples of two of the filters are shown in Figures 5-7 and 5-8. Note that the blue-green filter is much more transparent. The dark red filter, though also 3.00 mm in thickness, appears black. For the experimental trials, I used a range of filters from ultraviolet to near infrared, detailed in Appendix B.

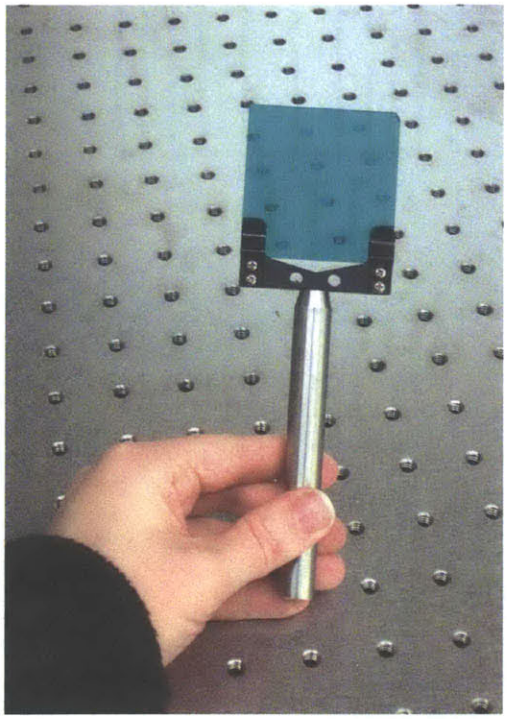

Figure 5-7: A blue-green Schott Glass filter on 2 inch optical filter mount

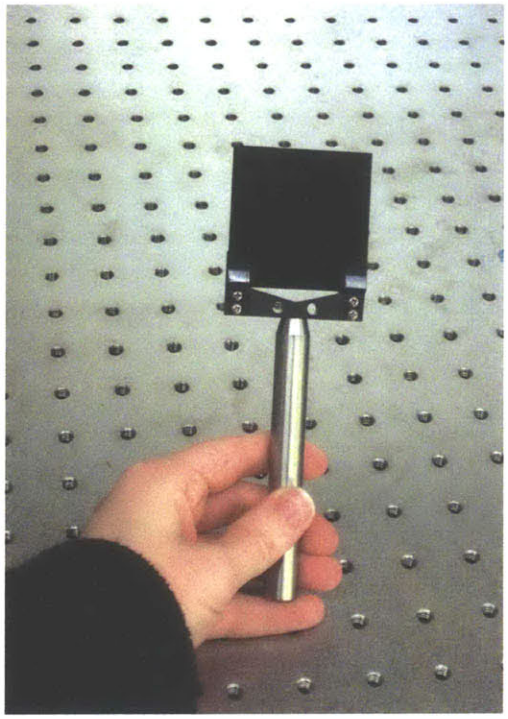

Figure 5-8: A Dark Red, RG-9, Schott optical filter on 2 inch mount

5.3.3 Data Processing

Little processing is done on the raw data extracted from the digital oscilloscope. However, the data saved is in three columns: one of time, one of voltage for the data signal and one of voltage for the trigger signal, as shown in Figure 5-9. The time signal is ignored. The trigger trace is used as a guide to extract only one period of the interesting data signal. Additionally, each data signal is of a slightly different length (number of pixels on the digital oscilloscope). Therefore, each signal must be expanded, interpolated and in a sense scaled to a uniform length.

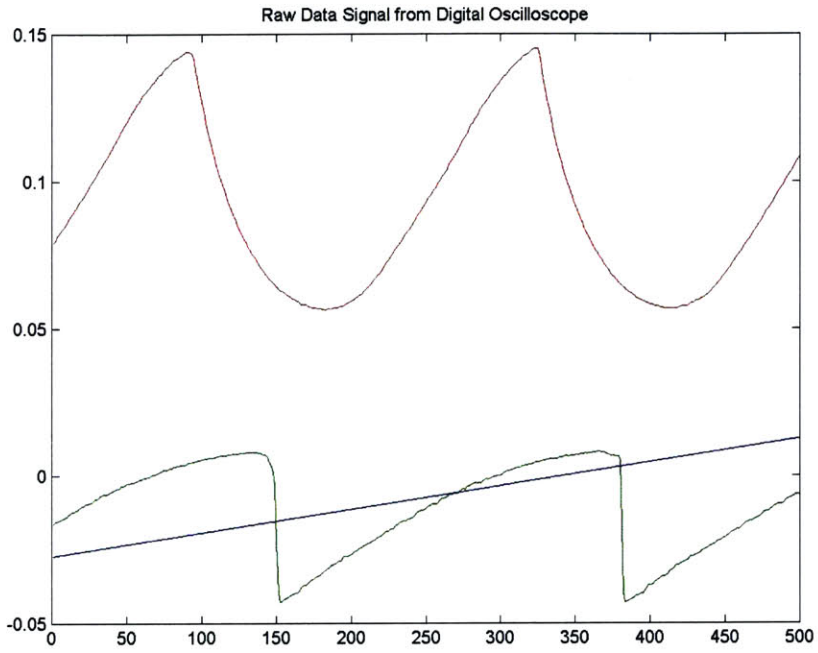

Figure 5-9: Data file 'TEK00012.CSV' from Cyan absorption filter

The first step in processing the oscilloscope data is to import the information into MATLAB. As each important piece of information is stored in a separate data file, each file must be loaded and the interesting data saved into a variable before it can be used. In the case of the monochrometer data, which is used to assemble a transfer matrix of absorption characteristics as a function of both wavelength and thickness, the individual data traces from each monochrometer wavelength need to

be concatonated into a larger grid. The code I used for this process is outlined in Appendix Section C.1. In the case of an unknown data signal, the trace simply needs to be isolated and expanded to the size of the thickness dimension in the corresponding transfer matrix, as outlined by code in Section C.3.

After the data has all been imported, extracted, expanded, scaled and merged into both a transmission matrix and a data signal, all that remains is to take the inverse of the transmission matrix and multiply through by the data signal to create the reconstructed spectrum. Mathematically and numerically the process is very clean. Experimental results are discussed in the following chapter.

Chapter 6

Experimental Results

To walk through the entire process of understanding the operational results of our instance of the Slim Format Spectrometer, we must start at a very simple case and build to the more complex result. The idea itself is elegantly simple. Fortunately, with carefully chosen parameters, the Slim Format Spectrometer is capable of producing useful information.

As described in Sections 1.5.1 and 5.3.2, the basic steps of spectral reconstruction are: to load in the transmission matrix, one monochromatic absorption trace at a time. Then compile them into a two-dimensional matrix, load in the data absorption signal and multiply through to create the reconstructed signal.

For all of the simulations which follow, except when explicitly mentioned otherwise, the transmission matrix shown in Figure 6-1 was used. The first wavelength channel in this matrix begins at 400 nm, in 20 nm increments to 700 nm. It was obtained with the Cyan single-dye filter, also shown in Figure 5-2.

6.1 Monochromatic Light

The first test for the Slim Format Spectrometer is whether or not the algorithm can recognize one of the monochrometer traces taken and already in the transfer matrix. As the input data signal, we choose one of the traces already loaded into the transmission matrix. Fortunately, the reconstruction algorithm is able to recognize

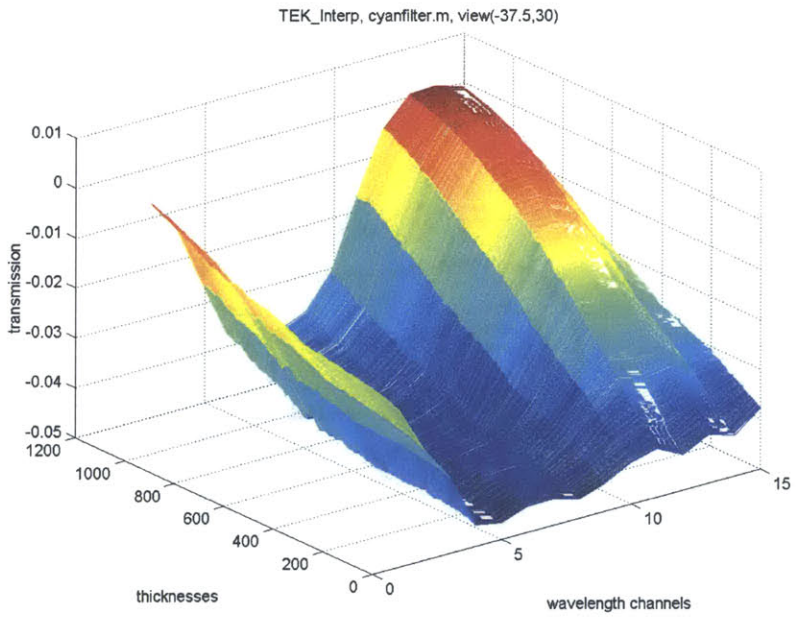

Figure 6-1: Full transfer matrix from 400 to 700 nm, obtained with Cyan absorption filter

that the input spectrum for this absorption trace is of a single wavelength only, the wavelength used to create the input data trace. In this test, note that the input spectrum is simply a delta function at the wavelength corresponding to "12" on the wavelength channel scale of Figure 6-2. The transmission matrix inverted and used in reconstruction is shown in Figure 6-1 and the data signal used is trace number "12" of the same transmission matrix.

With one successful test under our belt, we move to combinations of monochromtic input signals. First we attempt a linear combination of a pair of monochromatic data input signals. As shown in Figure 6-3, the spectrometer is able perfectly to reconstruct this combination of input signals. Even the magnitudes of the signals are equal, consistent with the input.

Perhaps though, the traces are so distinct that the spectrometer had little difficulty differentiating the signals and matching them to wavelength channels in the transmission matrix. As a check, we combine and average two adjacent monochrometer absorption profiles. We combine these two close signals, simulating the case in

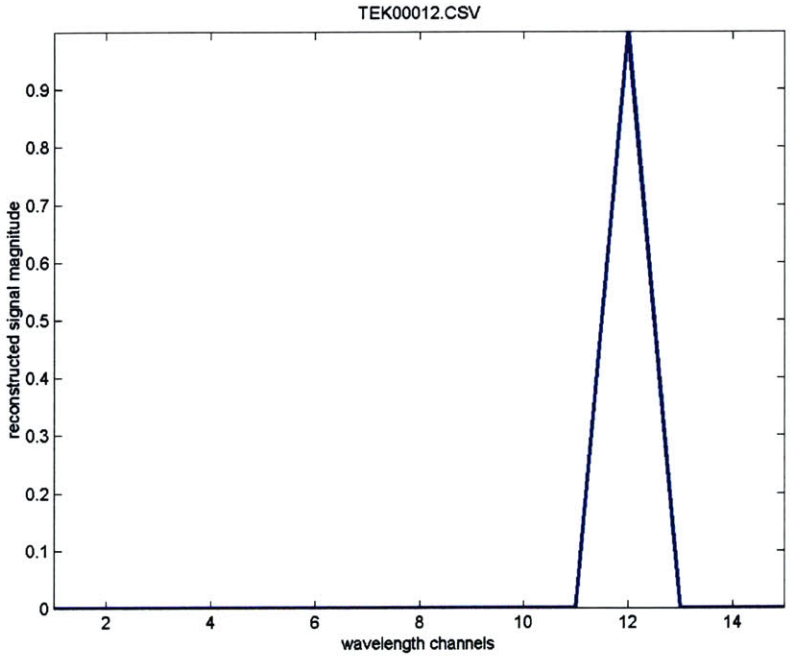

Figure 6-2: Reconstruction of a single monochrometer input

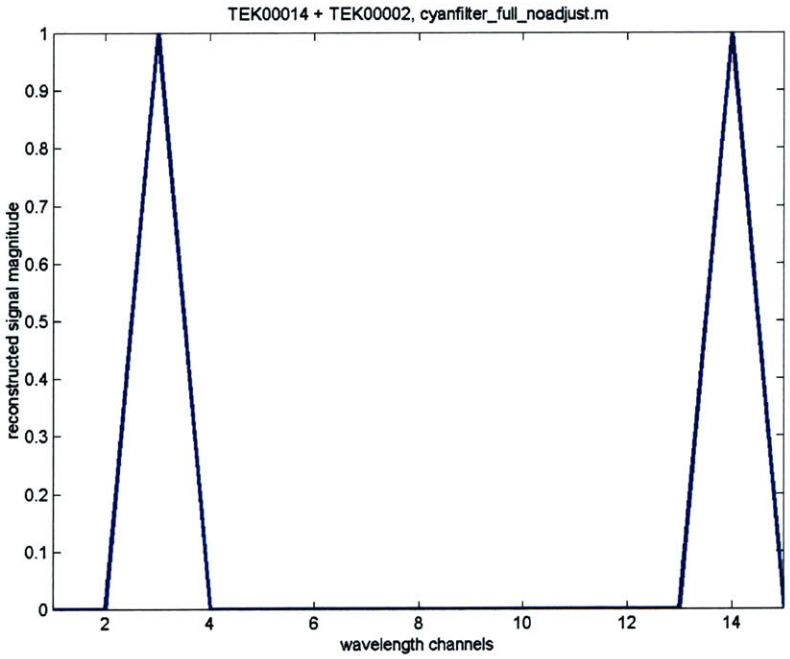

Figure 6-3: Reconstruction of a channel 2 and channel 14 monochrometer inputs

which the data signal is halfway between the first and second wavelength channels. As seen in Figure 6-4, the spectrometer is able to recognize the linear combination of signals. As input, the spectrometer reads back a signal with equal parts of the two adjacent wavelength channels. Additionally, we note that this phenomenon is not particular to any region of the transmission matrix. The spectrometer is able to reconstruct a red signal equally as well as a blue signal.

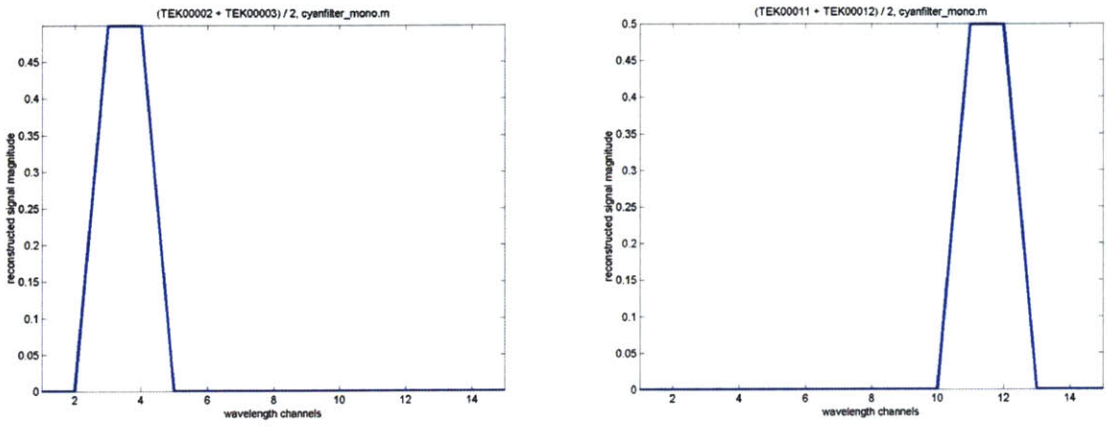

Figure 6-4: Reconstruction of Channels 3 and 4, and Channels 11 and 12

As a final test of monochromatic inputs, we test the Slim Format Spectrometer's ability to reconstruct input signals of varying magnitudes. Figure 6-5 shows the results of these trials, which are equally as clean as the other monochrometer trace reconstruction experiments. In the figure on the left, the input, as indicated by the title, is a combination of a wavelength channel 11 trace of magnitude 2 and a wavelength channel 12 input of unity magnitude. The reconstructed signals depicts exactly this, further confirming that the Slim Format Spectrometer should be capable of not only determining presence or absense of a particular discrete set of wavelengths but also determining their relative magnitude and thus the input spectrum shape. The figure on the left in Figure 6-5 is the reconstruction of an input of combination of wavelength channels 3, 11 and 12 at magnitudes 1.4, 2 and 5. Any simple system then it seems can be managed by the Slim Format Spectrometer. Integer or unity magnitudes are not necessary. Closeness or separation in wavelength is not a factor. The spectrometer is robust in detecting linear combinations of its monochrometer

inputs, in virtually any configuration.

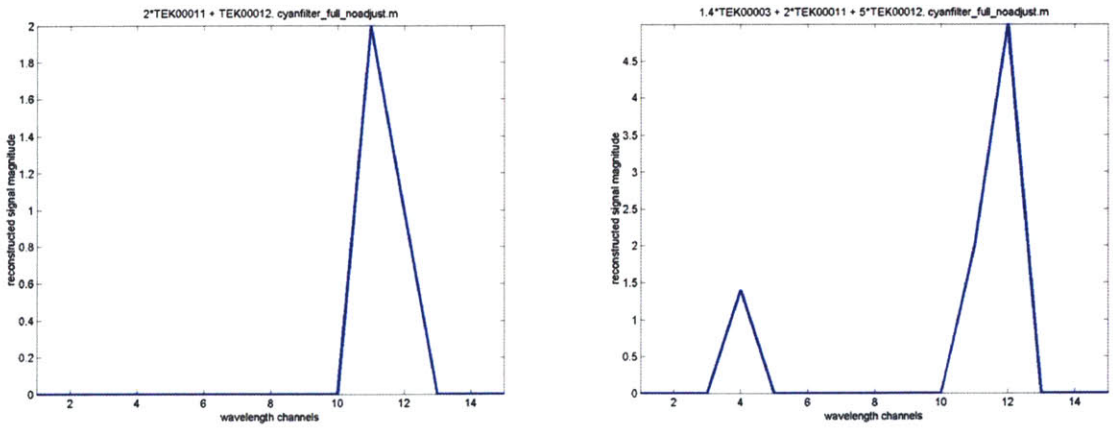

Figure 6-5: Reconstruction of Multiple Monochrometer Inputs and Intensities

It seems reasonable to conclude then that the inversion algorithm used by the Slim Format Spectrometer is capable of recognizing almost any linear combination of inputs of its column vectors. That is to say, if you input any combination of monochrometer traces, the Slim Format Spectrometer will very accurately reconstruct your input, both in wavelength and in intensity.

Intuitively, this result makes sense. The transmission matrix is simply a combination of independent vectors. Inverting the matrix and multiplying through by one of these eigenvectors allows the system to determine the eigenvalues. Similarly the reconstruction process can be thought of one of selective matching. When posed with a data signal, the spectrometer's task is to determine of which basis vectors, of which monochrometer traces the signal is composed. In the case of a perfect one-to-one match, the spectrometer behaves very well. It is a simple linear algebra problem, and an elegant demonstration of the Slim Format concept.

6.2 Data Signals

Although signals composed of exact eigenvectors of the transmission matrix can be reconstructed, we now attempt to reconstruct a more general data signal. We use the green glass band pass filter labeled BP44 (band pass filter, centered at wavelength

440 nm) and use the identical inversion algorithm as above for the monochromatic spectra.

As is evident in Figure 6-6, reconstruction does not work as flawlessly. The input spectrum is clearly centered between 400 and 500 nm with little transmission elsewhere in the spectrum. The reconstructed spectrum, although showing a peak between 400 and 450 nm, shows no preference of spectral shape throughout the remainder of the visible range.

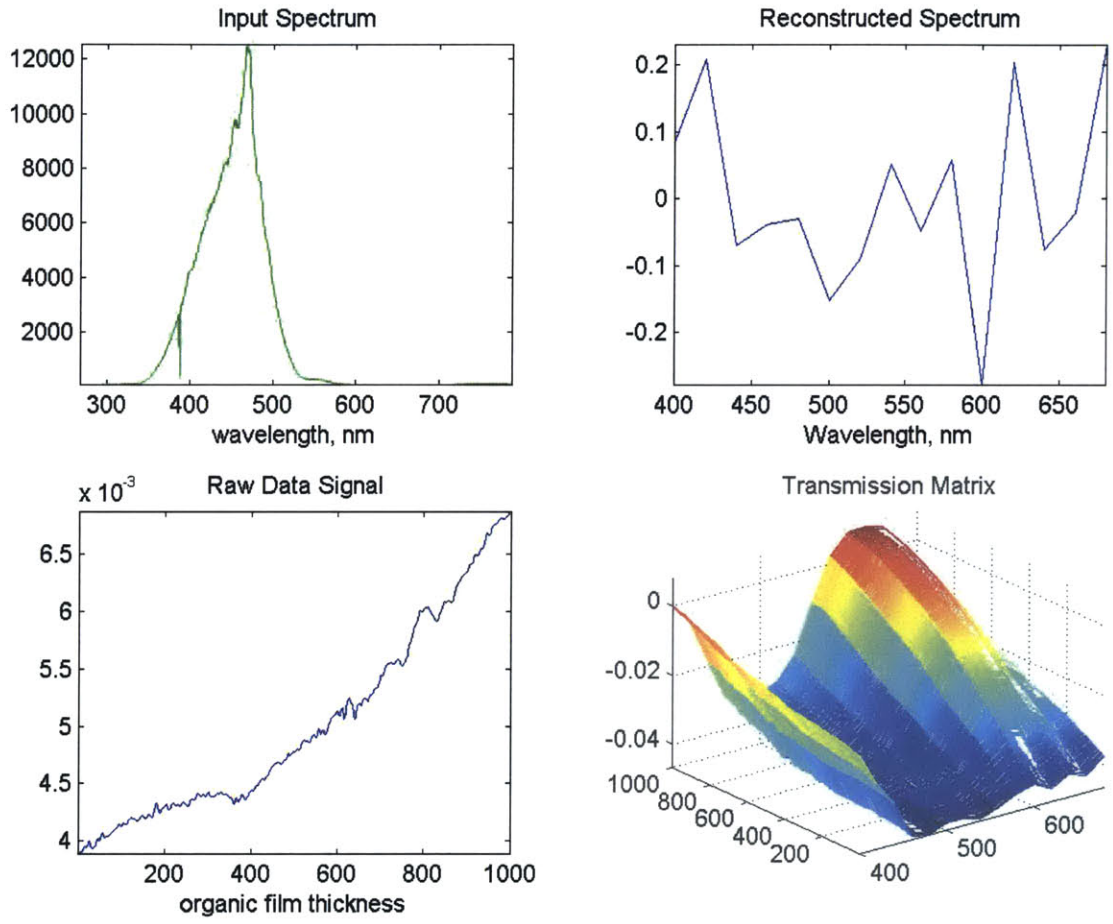

Figure 6-6: Reconstruction of a band pass spectrum

We consult the raw data signal for guidance. The signal, from comparison with the full transmission matrix, looks as though it does indeed fall into the 400 to 500 nm range. If I had to guess, I would place the signal with a strong component of 520 light, or possibly of 460 nm light. There is some ambiguity, as the scales of the plots

do not match (remnant of data-taking procedure). It is unclear whether the concave, steep, upward sloping data trace begins at the minimum, as the 520 trace or fits more in the magnitude range of the 460 trace. Though the transmission matrix has more detailed information that I am able to discern with my naked eye, the ambiguity is intuitively understandable.

6.2.1 Monotonicity

Intuitively thinking about the reconstruction algorithm and how it is more a less a complex matching algorithm, I am not surprised that a non-monotonic transmission matrix yields inconclusive results. In the Cyan filter transmission matrix of Figure 6-1 for example, I cannot differentiate between the downward sloping section of the transmission matrix and the upward sloping section with my naked eye. Although the inversion algorithm has much more information at its disposal, I propose that much of the cause for noise is in ambiguous assignment of absorption shape to a single corresponding wavelength. As was shown theoretically in Section 4.6, an input signal requires a monotonic or nearly monotonic transmission matrix in order to be accurately reconstructed.

The next experimental data processing step is clearly to confine a section of the transmission matrix, over which it is purely monotonic and reattempt input signal reconstruction. We begin by noting visually the range over which the full transmission matrix is monotonic. I arbitrarily chose the middle section, from 480 nm to 600 nm. I manually clipped the transmission matrix in MATLAB and continued with the idential reconstruction algorithm as before.

The results shown in Figure 6-7 indicate that spectral reconstruction succeeds only slightly better than in the case with the full transmission matrix. I argue that the results of Figure 6-7 are better, as clear spectral characteristics beyond white noise are demonstrated. There is a peak at 540 nm with declining intensities on both high and low wavelengths. Certainly it is not a conclusive test, and from examining the original input spectra, it is clear why. Most of the absorption we are using to reconstruct the input signal falls out of the wavelength range of the transmission

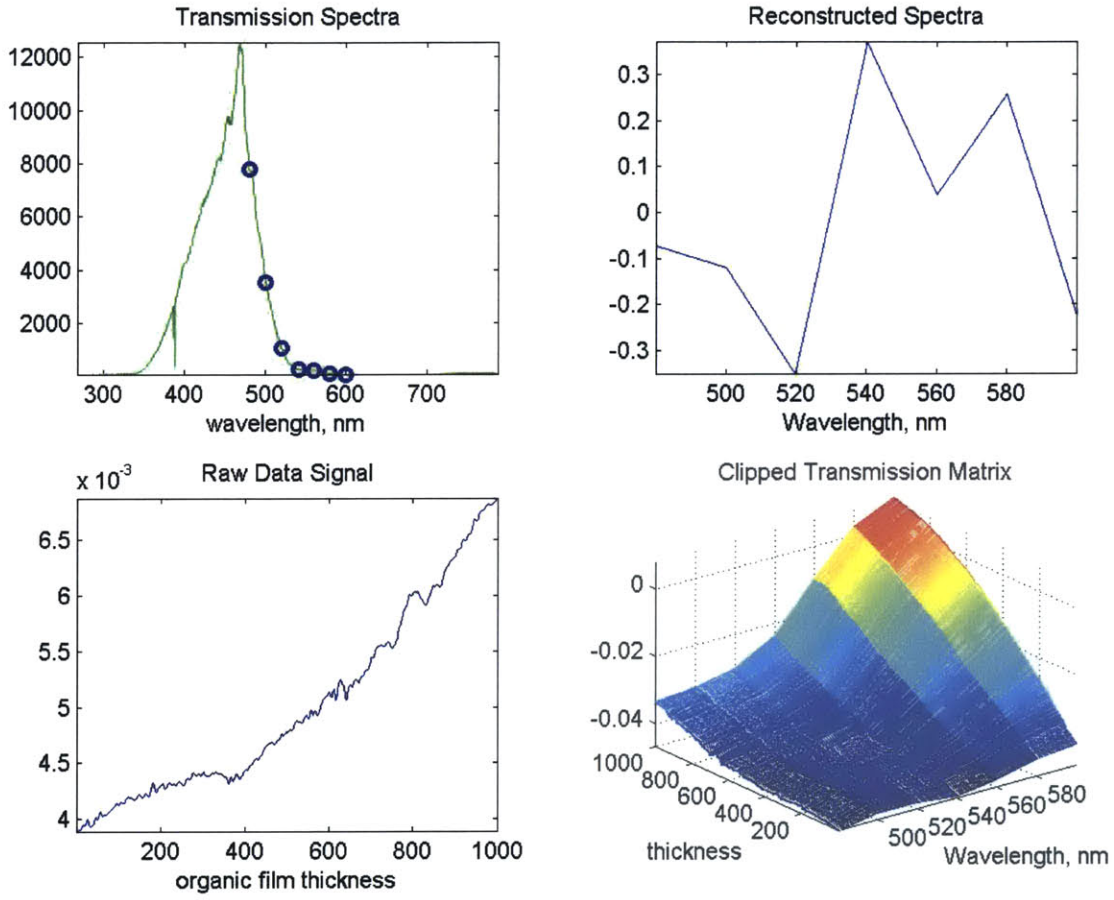

Figure 6-7: Reconstruction of a band pass spectrum with a monotonic transmission matrix

matrix. Monotonicity, though yielding improved results is certainly not a sufficient condition to obtain successful spectral reconstruction.

6.2.2 Bandwidth

In an attempt finally to isolate all the possible variables, I select the section of the transmission matrix that is both monotonic and corresponds as best as possible to our chosen input signal. The transmission matrix spans from 400 to 480 nm and is monotonic in both dimensions. The input spectrum has a slightly larger wavelength range. It spans from about 350 nm to 530 nm. Results of reconstruction are shown in Figure 6-8.

In fact, I think the result is quite good. There is a clear peak, though it is at 440 rather than around 460, but that's not a copious error, especially considering the similarity in these traces in the clipped transmission matrix of Figure 6-8. The reconstructed spectrum shows a clear peak and clear decline in transmission on either side of the peak, with the slope being steeper on the long wavelength side. Each of these characteristics is consistent with the input spectrum from the 440 nm bandpass filter.

6.3 Further Data

6.3.1 Monotonic, Bandwidth-Matched

Given the success of this initial test, I tried the method on other filters which had transmission roughly in the range of monotonic regions of the full transmission matrix. Specifically I chose two green band pass filters, BG18 and the unlabeled Green filter. Results for these trials are shown in Figures 6-9 and 6-10.

In the first case, I think spectral reconstruction is very good, confirming the theory that the algorithm is successful in the case of a monotonic, bandwidth-matched data signal and transmission matrix. The reconstructed spectra shows a clear peak and the rest of the signal on the noise floor.

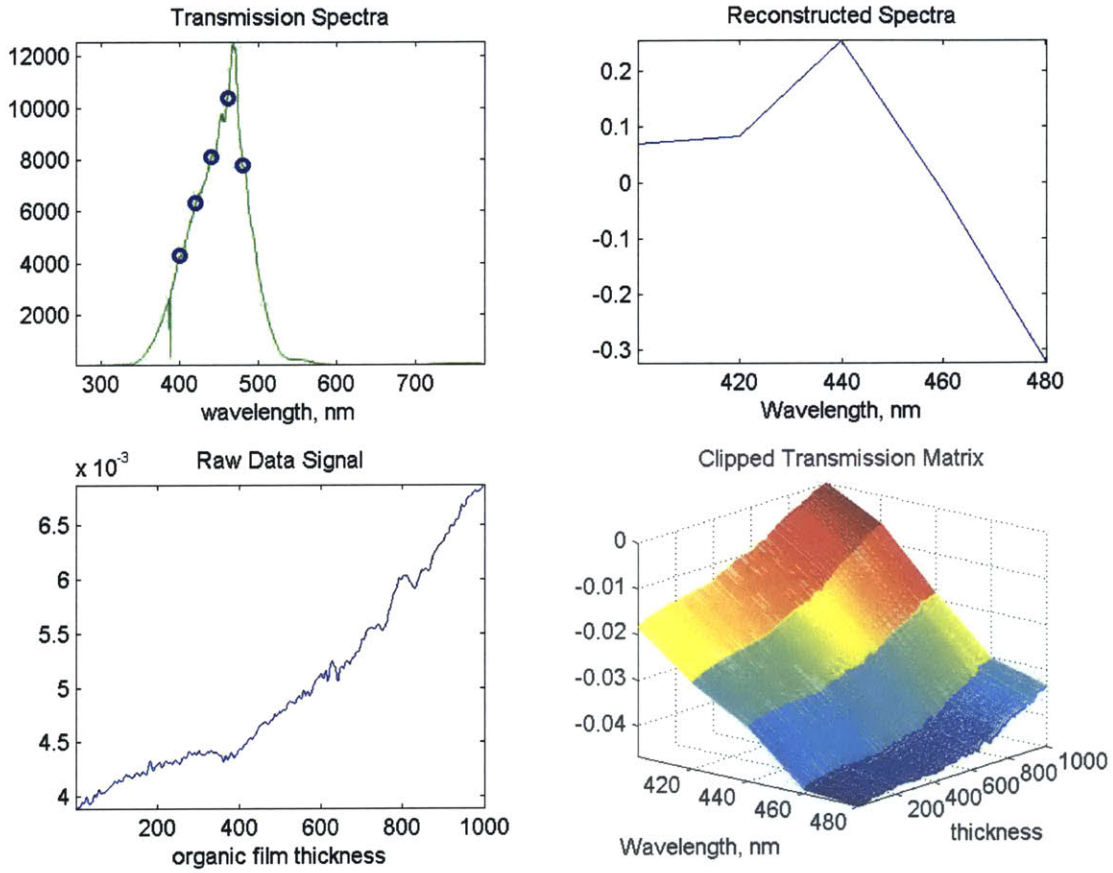

Figure 6-8: Reconstruction of a band pass spectrum with a monotonic transmission matrix and matched spectral range

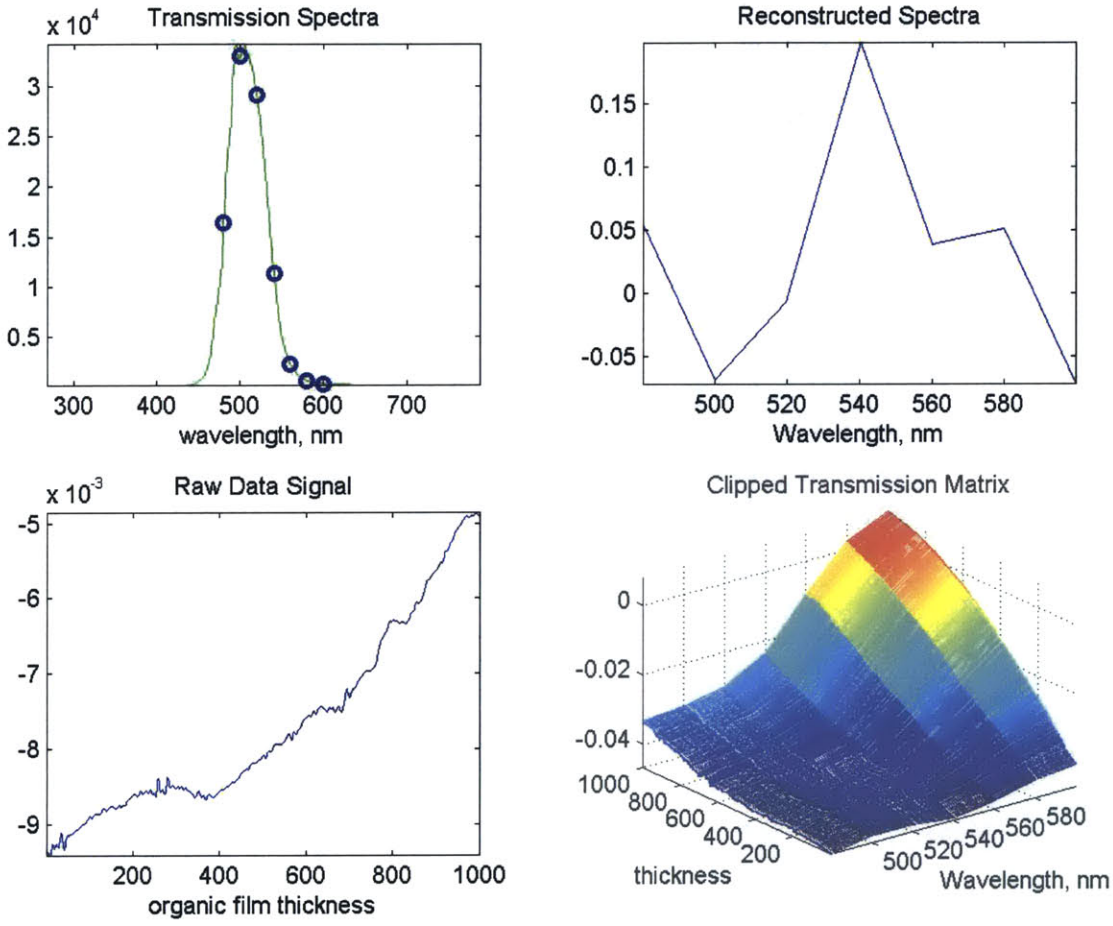

Figure 6-9: Reconstruction of a green band pass spectrum with a monotonic transmission matrix and matched spectral range

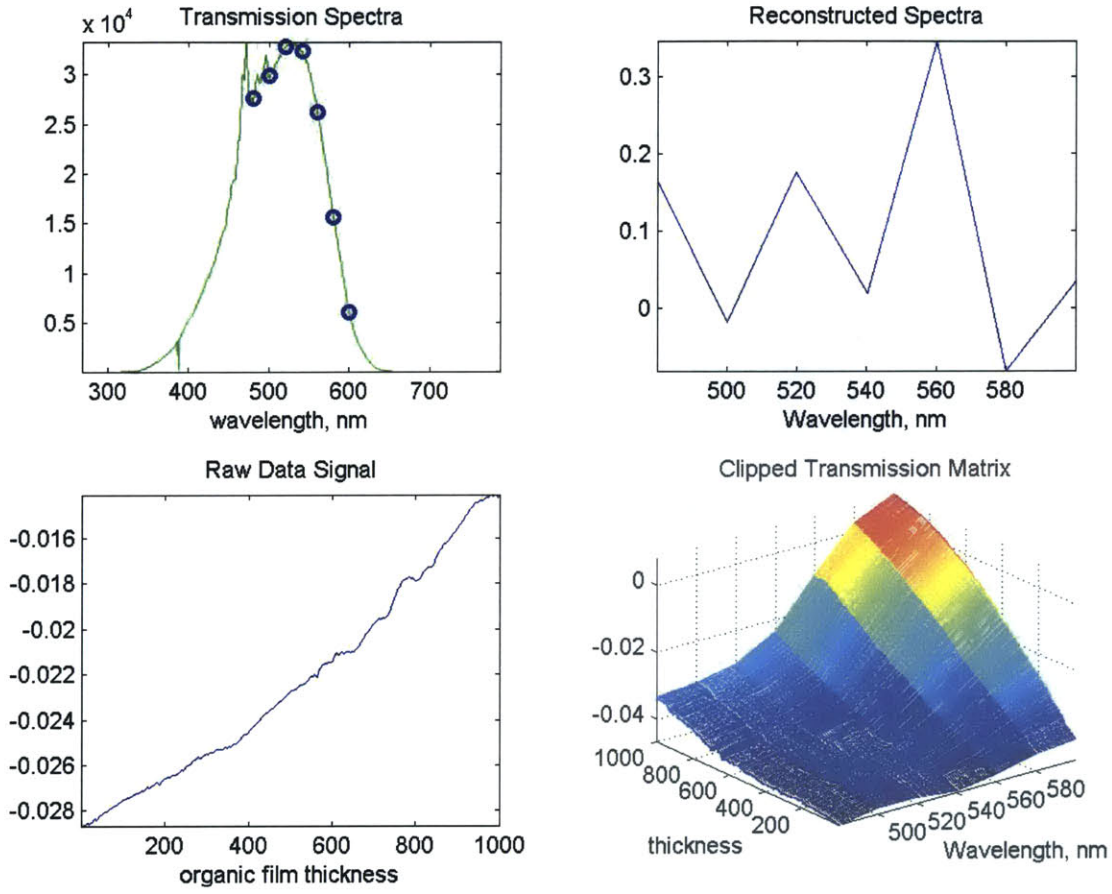

Figure 6-10: Reconstruction of a BG18 band pass spectrum with a monotonic transmission matrix and matched spectral range

The BG18 filter reconstruction is less successful. I attribute the non-indicative reconstructed spectrum on the large amount of data in the absorption signal of the input data that is not in the range of the transmission matrix. Specifically, it is known from the full transmission matrix that the section of the transmission matrix from 400 nm to 480 nm is increasing in transmission, making it easy to confuse with the 560 to 580 nm range. If this absorption data is contributing to the data signal, certainly the Slim Format algorithm will have little ability to discern the appropriate spectrum.

The key then to reconstructing a spectrum accurately is ensuring that the transmission matrix for the chosen filter is monotonic over the wavelength range desired to be measured and that the input signal only has information in the range of the transmission matrix. Choosing a filter could be somewhat complicated, with few dyes having a large range of monotonic absorption. However, limiting the input spectrum should be a simple matter of choosing a band pass filter that only passes the desired wavelengths, or an edge filter or combination of edge filters that allow the desired range only to intersect the absorption filter and be considered in the data signal used for reconstruction.

I would like to mention that the successful theoretical models of Chapter 4 have the advantage of a purely monotonic transmission matrix, in both thickness and wavelength dimensions, and also, that every bit of information in the data signal is covered in the bandwidth of the matrix. The wavelength ranges of both the input data signal and the monochrometer-created transmission matrix are identical, such that no absorption out of range in the data signal contributed to its absorption trace. The sets are equipollent, allowing clearer spectral reconstruction.

6.3.2 Non-Monotonic Bandwidth-Matched

In case there be any confusion, both monotonicity and bandwidth limiting are required for spectral reconstruction. Though the data above considers only the case with monotonicity but not bandwidth limiting, we consider in this section the case where the transmission matrix is limited in wavelength range but is nonetheless not

monotonic.

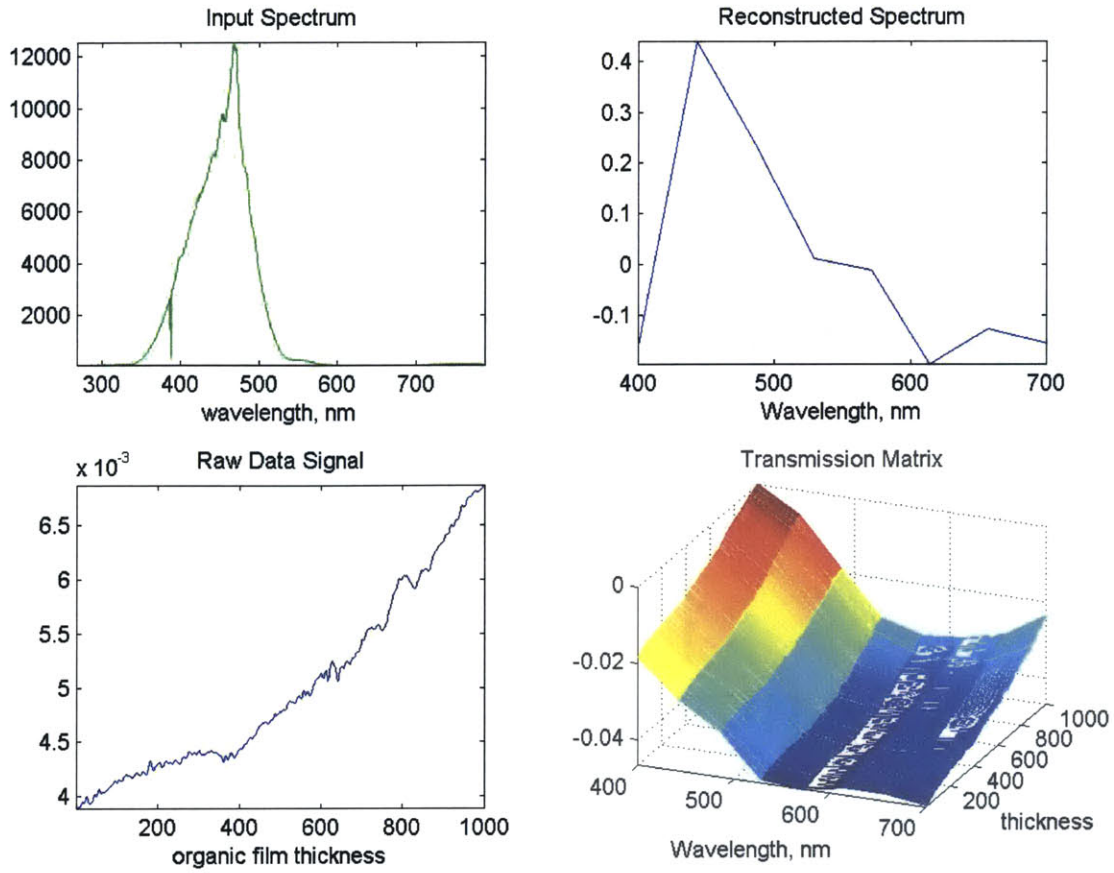

Figure 6-11: Reconstruction of band pass signal, with nonmonotonic but bandwidth-matched transmission matrix

The case of this filter and transmission matrix is interesting. As the transmission matrix is somewhat symmetric, the result of reconstruction turns out to be a very good reproduction of the mirror image of the input signal. Like the simulations in Chapter 4, this particular experiment demonstrates the case in which small error may be obtained but sheer luck, due to the ambiguity in the transmission matrix. In the case particularly, the ambiguity between the actual signal reconstruction and the wavelength that has a similar absorption characteristic but is on the mirror image side of the spectrum in consideration, caused for this example the mirror image of the input signal to be displayed. In general, it is intuitive and has been demonstrated theoretically, that a monotonic transmission matrix is required. If nothing else, this example may be considered to emphasize how important the correct bandwidth is in

signal reconstruction.

6.4 Filtering

One of the methods attempted to increase resolution of the transmission matrix was a combination of interpolation and filtering of the matrix to fill in the spaces in the surface for which we have no experimental data but feel that a guess could be made. It is known from communication theory that no more information could be contained in the reconstructed signal, but it was thought perhaps that spectral shape could be better simulated.

As in the case of the physical experiments, we tested on theory first on a single monochrometer input. Figure 6-12 demonstrates the effect of an interpolation filter on the reconstructed spectrum. Note that the filter acts as a sort of convolution of the original clean signal into a smoother but oscillating signal.

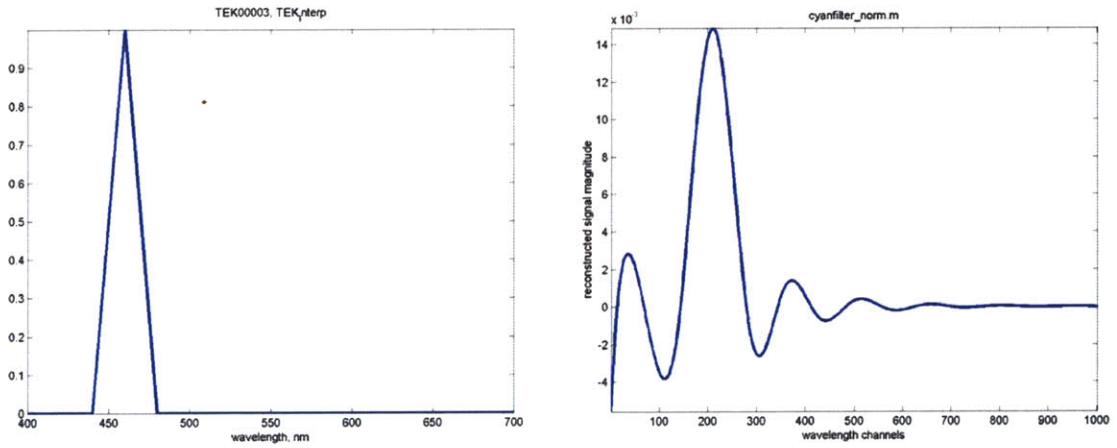

Figure 6-12: Reconstructed impulse function, with standard and interpolated transmission matrices

The process for creating this higher resolution signal is interpolation. A spline method is used on the original transfer matrix until it is expanded to a square matrix. The original and interpolated transfer matrices are shown in figure 6-13. The general shape and quality of the interpolated matrix is kept and the original information seems equivalent. However, the resulting function is much smoother, as if many more

monochrometer traces had been taken.

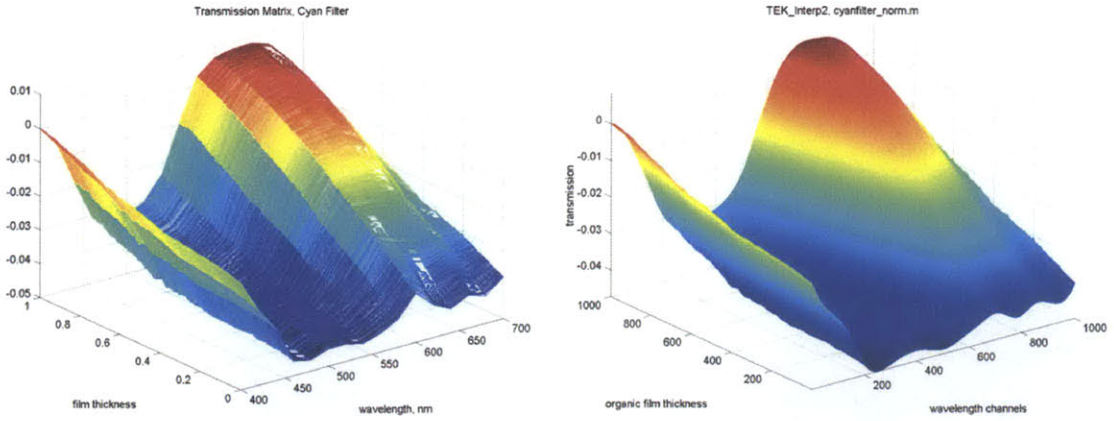

Figure 6-13: Transmission matrix of Cyan filter before and after interpolation

A second test case, as in the experimental procedure, is the combination of impulse functions. A single example is shown in Figure 6-14. As in the case of the single impulse function, the reconstructed spectrum is essentially the same, but the output is smoother and oscillating. The oscillating quality is characteristic of the convolution of an impulse function with the Fourier transform of a rect, or box function, which is a sinc function. These oscillations in signal processing are merely residual from the non-infinite bandwidth of the filter.

Additionally, the width of the impulse functions in the interpolated reconstruction are still highly dependent on the resolution of the impulses in the unfiltered case. Though this concern would be of no consequence in a smooth spectrum, all the same, it is deceptive to think that the width of the detected spectra is as much as 50 nm when in reality it could be infinitesimally thin.

As a final exploratory step, we examine the reconstruction of an impulse function, after the transfer matrix has been completely interpolated, inverted, shifted positive and normalized. The inversion is meant to correct for any distortion created by the current amplifier, which inverts the voltage signal. Normalization is thought to correct for the non-uniform intensity of the white light lamp, powering the monochrometer and for the non-uniform throughput of the monochrometer per wavelength. In the end, however, as can be seen from Figure 6-15, the result is merely to confuse and

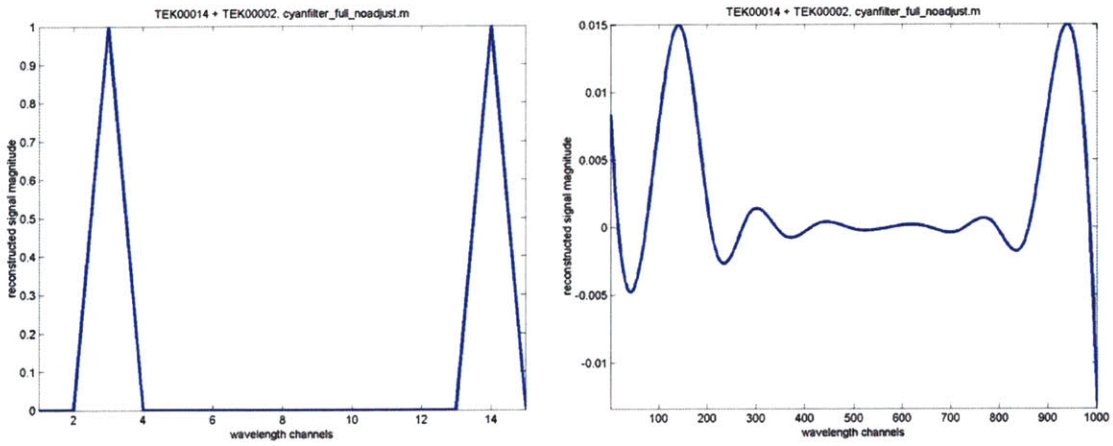

Figure 6-14: Transmission matrix of Cyan filter before and after interpolation

muddle the otherwise clear impulse function.

Though filtering seems a clever, quick and cost-efficient solution to the shortcomings of the Slim Format Spectrometer in terms of resolution, it seems to create more trouble than it solves. The unfiltered results may seem inaccurate or unmeaningful, but the filtered signals are only even more muddled. Even though identical processing is performed on the data signal as was performed on each trace of the transmission matrix, complex signal processing artifacts result in a degradation of the signal information. As a scientist, I would far prefer the blocky but thoroughly accurate results of the low-resolution signal.

6.5 Absorption Filter Variety

I did collect data for multiple absorption filters, all either single color gradients of Cyan, Magenta and Yellow or a combination of these dyes. For the most part, the single-color transfer matrices gave the best data, as the largest part of the spectrum was monotonic.

Consider for example the Cyan-Magenta-Yellow filter shown both in Figure 5-3 and also here in Figure 6-16. There are a few interesting qualities to the transmission matrix. Note first that though each monochrometer wavelength has a very distinct shape, very little if any of the transmission matrix is monotonic in both dimensions.

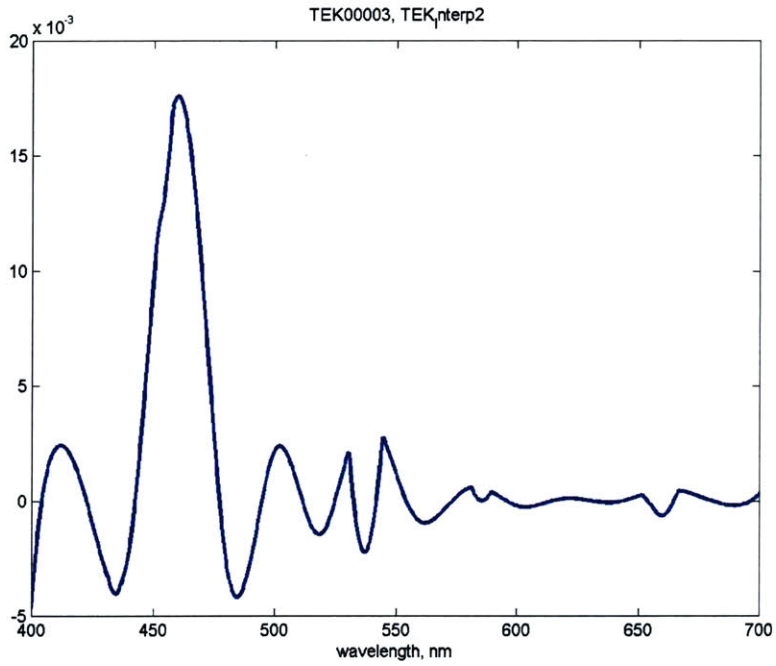

Figure 6-15: Reconstructed impulse function, with interpolated and normalized transmission matrix

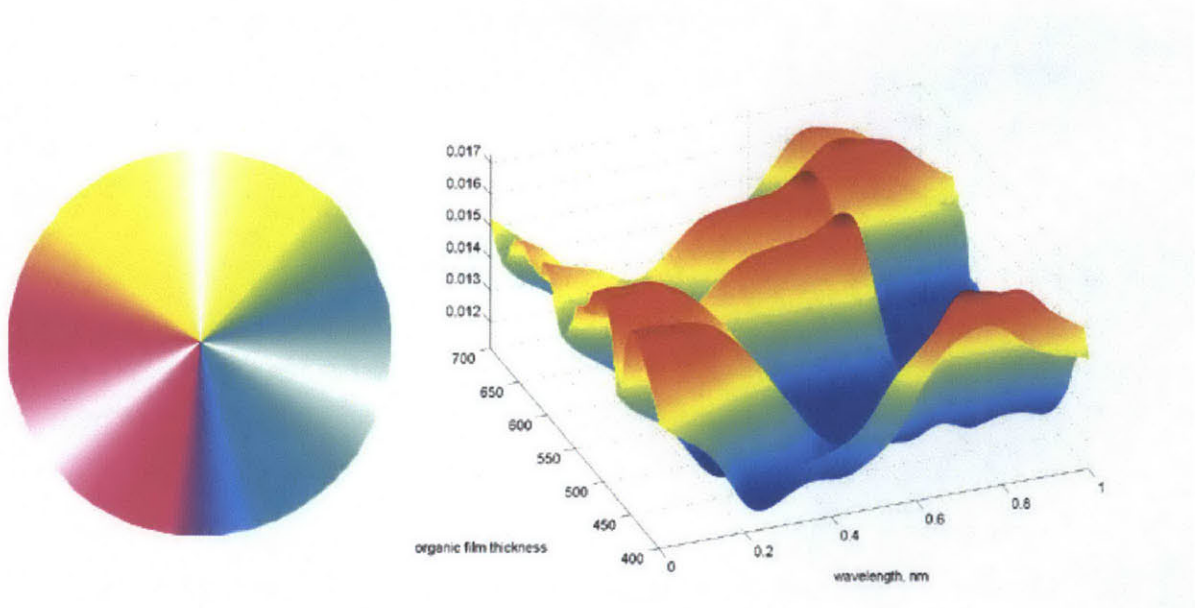

Figure 6-16: Cyan-Magenta-Yellow absorption filter and the corresponding transmission matrix

Although in principle, you would expect such a distinct absorption profile to be easily distinguishable, as likely any comparison tool could match the data signal with the transfer matrix and determine where its main component lies on the wavelength axis. Mathematically, however, in the case of our linear algebra algorithm, it is difficult to distinguish whether a particular wavelength is what combination of which frequencies. Perhaps with only three wavelength choices a better match could be made. As is however, the algorithm seems only to tolerate purely monotonic functions or exact waveform matches.

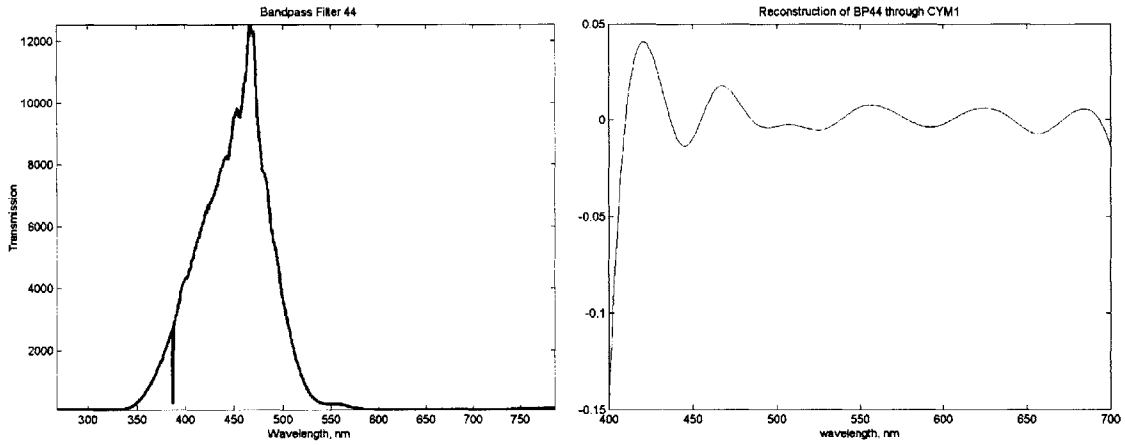

Figure 6-17: Input spectrum from Schott filter BP44 and reconstructed spectrum using CYM absorption filter

Also of note, the transfer matrix does appear quite as you would expect, if you were to compare the spectra of each of the dyes to the absorption filter. Remembering that the transmission matrix is inverted due to the current amplifier, note that the absorption filter is completely transmitting of all wavelengths along three stripes corresponding to the transparent section of the filters. The filter also allows all wavelengths around the color of the dye for three stripes corresponding to cyan, magenta and yellow. The transmission matrix also reveals the absorption to varying extents of the wavelengths not permitted to pass by the particular filter dye. All in all, this filter presents a very satisfying demonstration of the intuitive properties of absorption, transmission and also of spectral reconstruction.

Chapter 7

Future Work

Though it has been demonstrated both theoretically and in preliminary tests that the Slim Format Spectrometer has great potential, a few interesting questions require further investigation.

7.1 Creating a Monotonic Transmission Matrix

Namely, since it has been shown that only monotonic transfer matrices, in both the thickness and wavelength dimensions can yield useful results, we would like to fabricate a transmission matrix that is monotonic over as large a wavelength range as possible. We would like to be able to engineer this matrix while keeping intact the basic operating principles of the spectrometer.

Given the variables that contribute to the matrix are film thickness and wavelength-dependent absorption, these are our two choices in engineering the matrix. Clearly, engineering a material property consists only of obtaining spectral information for a wide variety of materials and choosing one. Again, due to the environmental restrictions that many substances are unstable in oxygen- or moisture-rich environments or when exposed to high-energy photons, the inkjet dyes are still a very good option.

Our remaining option is to compensate low areas of wavelength-dependent absorption with larger film thicknesses. Certainly film thickness can be controlled at a resolution on the filter at least as fine as the wavelength-dependent change in absorp-

tion of the dye. The relation of Photoshop gradient tool to actual film thickness is unknown, but could easily be determined, knowing the governing absorption physics. Certainly this method would have limits, but it seems interesting and potentially useful enough to merit investigation.

Also of interest would be a combination of dyes to yield favorable transmission/absorption qualities. The set of #78 Hewlett Packard inkjet dye spectra are given in Appendix B. The dyes are orthogonal in some regions and coincident in others. Certainly a linear combination of dyes along with an interesting thickness gradient could yield a monotonic transmission characteristic over at least the visible range.

7.2 Device Performance

As the device tested for this thesis was the first of its kind, its performance is not yet optimal. It would well worth the time to try to improve the range and resolution of the device, as well as improve accuracy and simplify the experimental configuration and procedure.

7.2.1 Accuracy

A few simple improvements of physical items would likely improve the accuracy of the prototype Slim Format design. First, replacement of the overhead transparency film, which had visible inclusions, would reduce the scattering and improve the transmission characteristics of the absorption filter.

A more stable motor would reduce a great amount of noise in the data collection, as the frequency of the motor, which determines the thickness dependence of the transmission, would drift on the time scale of a few seconds. It is difficult for a single transmission trace to be taken under such conditions. It is nearly impossible to expect an entire set of data magically to have the same thickness qualities, when the frequency and thus the angular position of the motor are drifting. In the interest of more reproducible position measurements, I recommend switching from a motor,

though much less expensive, to a translation stage, especially if the lab that wishes to use the spectrometer has a translation stage, perhaps for another application. For device performance, reproducibility is more important than speed. Further, the tools required to fabricate a variable-thickness film in one linear dimension rather than radially are much less complicated. In the case of transparencies, the software is much less expensive and more common. In the case of semiconducting materials, thermal evaporation schemes have already been developed which allow for a variable-thickness constant-gradient film. However, such a radial gradient would be very challenging to produce.

As always, a darker test environment, which is very difficult to obtain in the presence of a high intensity white light source, and a more sensitive photodiode would improve the precision of the data collection.

Since the spectral range from which the input signal is composed is of great importance in reconstruction, I would take care to insert a bandpass or a set of high and low pass filters in the optical path of each experiment, to ensure that only visible light or light of the spectrum of interest is contributing to absorption.

7.2.2 Efficiency

Although in improving device performance, more expensive and intricate is often the direction to go, in this case, I would recommend a simpler photodiode for the trigger track, or a simpler method of lining up the rotating filter or translation stage. For example, if a LabVIEW program could be written to automate the characterization of the absorption filter, the time to collect the data signal would become less significant. A researcher could start the automation process and return when it has completed.

Certainly computer automation would allow for more precise characterization of the transmission matrix in the wavelength dimension. With a bit of skill, a program could be written to perform real-time inversion and reconstruction of data signals, once a transmission matrix for each filter has been input. Such a design is of course the final intention for the Slim Format Spectrometer which would be commercially available.

7.2.3 Resolution

On the topic of spectral reconstruction ability, it would be interesting to take transmission matrix measurements at much finer increments. Now that the range at which the spectrometer will likely work has been determined, it would not be difficult nor time-consuming to improve this resolution and examine the abilities of the spectrometer reconstruction algorithm.

7.2.4 Operational Range

The Slim Format Spectrometer concept has been preliminarily shown to be feasible in the visible range. I see no reason that the concept could not be extended to the infrared or ultraviolet wavelength ranges. I would be very interested to find dyes on which to test the range of the device.

Chapter 8

Conclusion

The prototype construction and initial tests of the Slim Format Spectrometer are a modest success. The device shows distinct potential, able to reconstruct spectra that meet a set of specifications. The device however is not yet ready to be built and sold to a general population.

Based on a sound theory of materials physics and linear algebra, simulations show strong indications that given further development, the device design is able to produce accurate, reproducible spectral reconstruction over the visible range.

Specifically, the wavelength- and thickness-dependent absorption in organic thin films has been sufficiently varied to enable wavelength differentiation. The Moore-Penrose pseudoinverse is able to invert any transmission matrix we have obtained to a satisfactory accuracy. Given a transmission matrix and data signal that are both monotonic with matched bandwidths, the Slim Format Spectrometer is able to reconstruct the general spectral location and relative magnitude of any features.

The current drawbacks to the design are its limited range and resolution, both of which I believe can be improved with further development efforts.

Consistent with its purpose, the Slim Format Spectrometer is a viable, low-cost alternative to commercially available, delicate and expensive spectrometer systems.

Appendix A

Derivation of Beer-Lambert Law

There are few distinct approaches to describing the concept of absorption. One approach is derived by quasi-geometrical arguments of the theory of radiative energy transfer [9]. A similar approach cites the complex susceptibility of dielectric materials and further extrapolates to obtain the absorption/gain coefficient for dielectrics [11]. This coefficient is the same, or rather the negated opposite, as the laser gain coefficient, which determines the amplification of light as it passes through a material.

The approach I use here is in fact related to the laser amplifier. Though a change of variables is required, this approach is identical to all the other derivations, but is physically much more straightforward and it explicitly reveals the underlying optics.

A.1 Definitions

We note that in a photon-atom interaction, three interactions are possible. If the atom is unexcited, the photon may be absorbed. If the atom is excited, a photon may be emitted by stimulated emission, leading to amplification. Also possible is the spontaneous emission of a photon from an excited atom, while the atom is in the presence of photons.

Due to the basic probability of photons being absorbed atoms we know that the probability density that an unexcited atom absorbs a single photon. We know the probability density than an atom absorbs a photon in a cavity volume V is

$$p_{ab} = \frac{c}{V}\sigma(\nu) \tag{A.1}$$

where $\sigma(\nu)$ is called the transition cross section. Its shape indicates the relative magnitude with frequency of interactions. We define the area under the transition cross section as S, the transition strength. More commonly used is a quantity called the *lineshape function*, $g(\nu)$, which is the transition cross section normalized by the transition strength.

$$\sigma(\nu) = Sg(\nu) \tag{A.2}$$

One more useful definition is the amount of time that a photon will spend in an excited state before spontaneous relaxation and photon emission into any mode. This transition time varies with the atom and particular state. We define the *spontaneous lifetime* t_{sp} such that the probability density of spontaneous emission of one photon is

$$P_{sp} = \frac{1}{t_{sp}} \tag{A.3}$$

Due to some useful probabilities and integrals, we can extract an expression for S using t_{sp}

$$S = \frac{\lambda^2}{8\pi t_{sp}} \tag{A.4}$$

Equation A.2, the expression for the transition cross section then clearly becomes

$$\sigma(\nu) = \frac{\lambda^2}{8\pi t_{sp}} g(\nu) \tag{A.5}$$

A.2 Gain Coefficient

Consider now the case in which a stream of photons is incident upon a single atom. The photon flux (photons/cm^2-s) is defined as $\phi = \frac{I}{h\nu}$. The number of photons n

in a cylinder of volume $V = hA$, produce a photon flux across the cylinder base of ϕA (photons per second). Photons travel at the speed of light, and therefore at any instant the proposed cylinder contains $n = \phi A = \phi\frac{V}{c}$ photons. The probability density that an atom absorbs one of n photons in the volume is simply $n * p_{ab}$.

We wish to determine the probability density for stimulated emission and absorption. By combining our equation for the probability of absorption of one photon in a volume with the known photon flux in a similar volume, we obtain

$$W_i = \phi\sigma(\nu) \tag{A.6}$$

where W_i is the probability density and $\sigma(\nu)$ is the transition cross section at frequency ν.

Though more detail is readily available, we merely define a population density difference, or a population inversion, N as the difference in the population of the excited state to the unexcited state. That is, the case of lowest energy has a negative N value and the case of a highly excited system yields a positive N value. Furthermore, a positive N causes the photon-flux density $\phi(z)$ to increase with z. Emitted photons stimulate yet more photon emission events. The result of compounded stimulated emission is an exponentially growing $\phi(z)$ photon-flux density.

Written mathematically, this situation is

$$d\phi = NW_i dz \tag{A.7}$$

Defining a gain coefficient $\gamma(\nu)$, we may write this so-called rate equation as

$$\frac{d\phi(z)}{dz} = \gamma(\nu)\phi(z) \tag{A.8}$$

which can be solved and expressed as

$$\gamma(\nu) = N\sigma(\nu) = N\frac{\lambda^2}{8\pi t_{sp}}g(\nu) \tag{A.9}$$

The gain coefficient represents the gain in photon-flux density per unit length.

A.3 Absorption

Harkening back to Equation A.6, we solve this rate equation for ϕ, which is clearly an exponential function. The complete solution is

$$\phi(z) = \phi(0)e^{\gamma(\nu)z} \tag{A.10}$$

.

We now convert to optical intensity $I(z) = h\nu\phi(z)$,

$$I(z) = I(0)e^{\gamma(\nu)z} \tag{A.11}$$

.

A.4 Beer-Lambert Law

Remember now that the amplifier gain coefficient is proportional to N the population inversion $N = N_2 - N_1$. Similarly, we define the attenuation in signal to be the opposite of the signal gain. If the population inversion is negative, photons will be absorbed rather than emitted.

$$\alpha(\nu) = -\gamma(\nu) = -N\sigma(\nu) \tag{A.12}$$

For a total interaction length x, the overall gain of the amplifier $G(\nu)$ is defined as the ratio of the photon-flux density at the output to the photon-flux density at the input,

$$G(\nu) = \frac{\phi(x)}{\phi(0)} = e^{\gamma(\nu)d} \tag{A.13}$$

In the case where γ is negative and α is positive, Equation A.13 represents the Beer-Lambert law of absorption.

Appendix B

Spectra of Glass Filters

For the experiments, I used a set of 20 filter slides. Each slide is approximately 2 inches by 2 inches (5.0 cm by 5.0 cm) and a few millimeters thick. I used a standard 2 inch optical filter mount. Most of the slides were doped glass, manufactured predominantly by the Schott Glass company. Extensive technical details may be obtained online or by request. Two filters were unlabeled but appeared similar to the Schott doped glass filters. The 385 nm filter was manufactured of quartz rather than of glass.

The following spectra were taken with an Ocean Optics spectrometer with a fiber input and the same stable white light lamp that was used for the rest of the experiments. The spectra shown are merely the detected combinatory intensity of the white lamp and the filter, not the pure transfer function of the filter, which would be obtained by dividing out the white light nonuniformities.

These spectra represent the ideals that our thin film spectrometer was intended to recreate. They are the "solutions" to our reconstruction challenge. Certainly our thin film spectrometer never aspired to such fine resolution and accuracy as the Ocean Optics grating spectrometer. The filter spectra are shown here for reference only.

B.1 Schott Glass Filters

The Schott doped glass filters are mainly high pass filters with very accurately engineered edge wavelengths. These filters are useful for removing a specific part of

the spectrum in measurements, such as the laser signal, while allowing all of the interesting data signal to pass. In addition to the high pass filters, I also tested a few band pass filters, in which a certain range of wavelengths is allowed to pass, and band gap filters, in which a certain range of wavelengths is suppressed and filtered out. The transmission spectra of the white lamp source used for all my experiments through these filters is shown below. In order to obtain the transfer function listed in the Schott literature, you would need to divide by the white light spectra, shown in Figure B-21.

The filters are shown below in order of high pass filters, from low to high edge wavelength, then band gap and finally band pass filters. Within these categories, the actual filter code as assigned by Schott Glass Technologies was referenced. In many cases these codes can be guessed intuitively. The OG filters for example stands for "Orange Glass," and they transmit in the orange and red color regions, giving the physical filter an orange-red color. The three digit number following the "OG" code represents the edge wavelength, above which the filter is absorptive. Most of the high and low pass filters use their edge wavelength as a three digit code and are categorized by dopant impurities into the two letter codes, often simply referencing the color that the resulting glass filter appears to humans.

B.2 Inkjet Printer and Transparency Filters

While the first seventeen filters are made of very pure glass, doped with impurities, the trailing four are of a very different class. These are absorption filters I created, identical in spectra to the filters I used in the thin film spectrometer prototype. They are fabricated of inkjet printer dyes, inkjet printed onto transparency sheets. They were created using Adobe Photoshop, scaled to a convenient size, and simply printed with a standard office printer. The thin films of inkjet dye were printed as thick and opaque as possible, with the highest quality setting onto the transparencies. As described in Section 5.2.2, these filters are not optically perfect, but they are very stable in an oxygen and moisture atmosphere and retain a stable transmission spectra

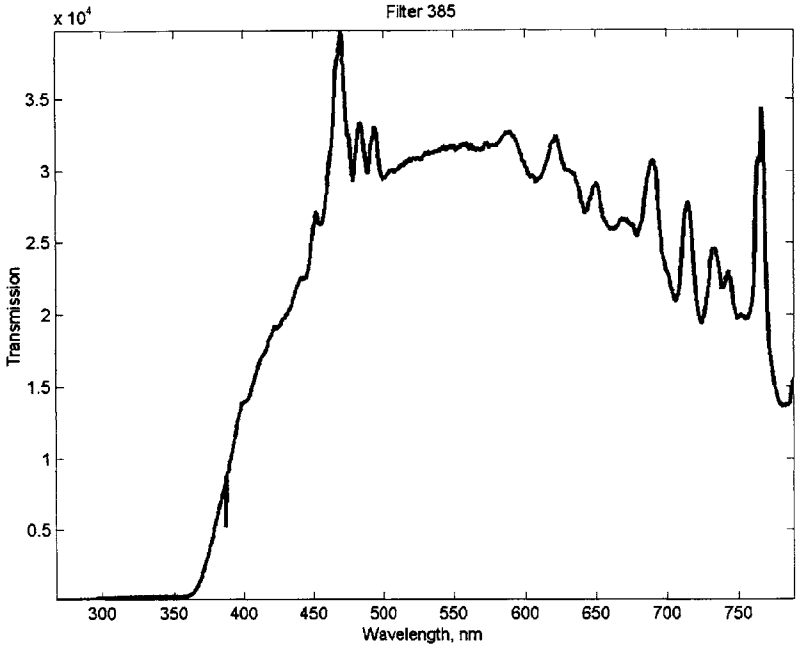

Figure B-1: Colorless quartz filter, high pass with edge at 385 nm

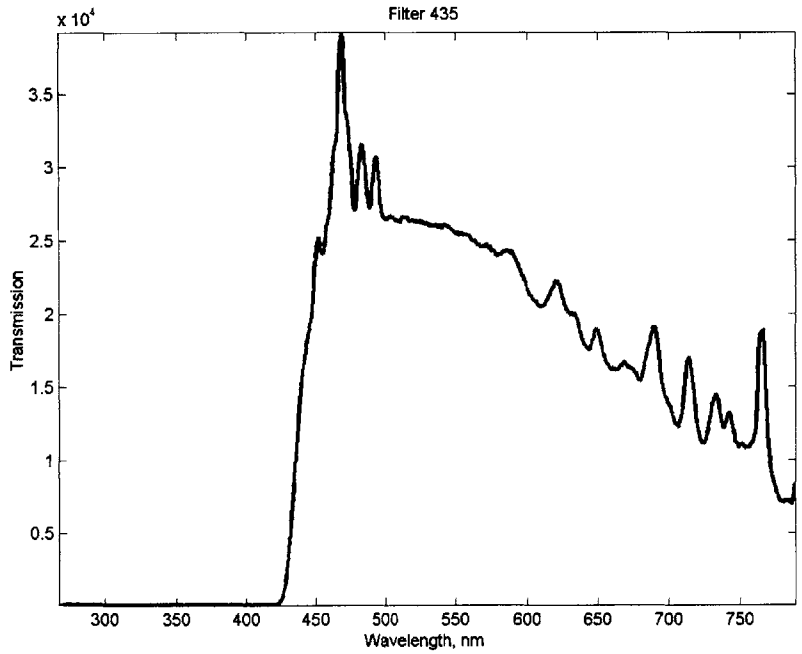

Figure B-2: Yellow-green doped glass filter, high pass with edge at 435

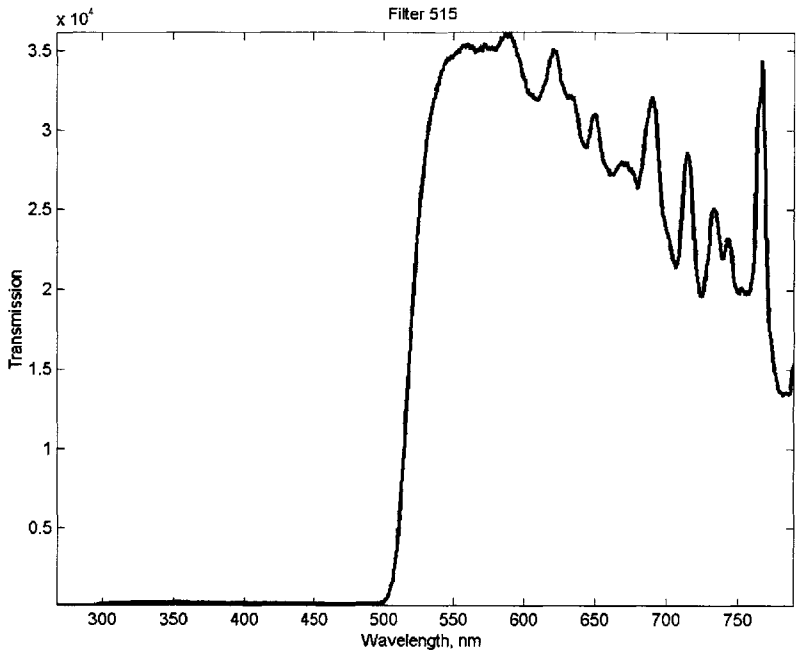

Figure B-3: Yellow-colored doped glass filter, high pass with edge at 515 nm

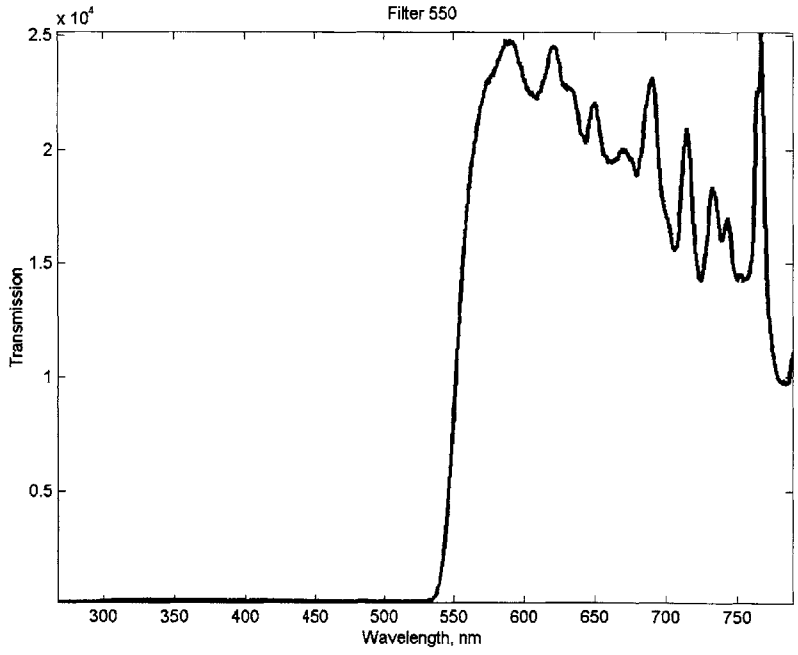

Figure B-4: Orange doped glass filter, high pass with edge at 550 nm

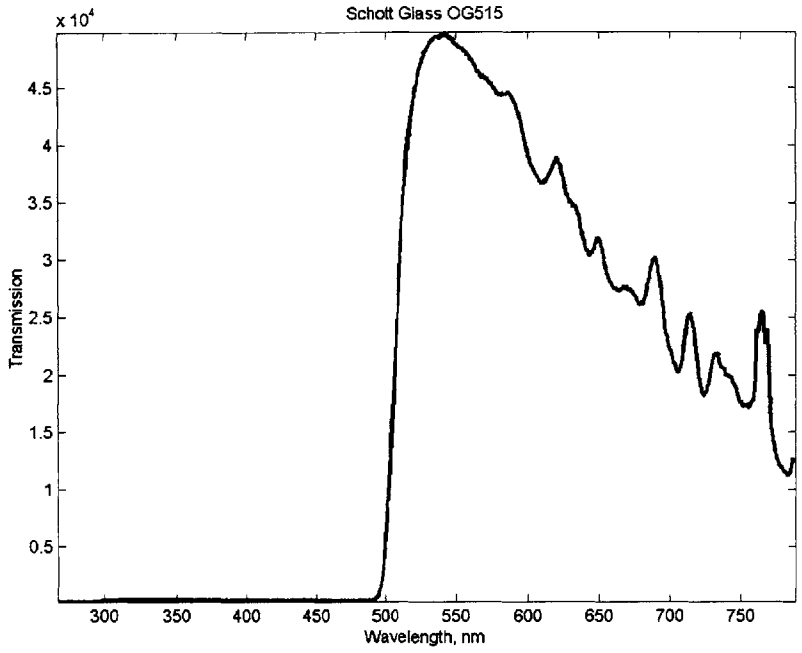

Figure B-5: Schott Glass orange high pass filter, with edge at 515 nm

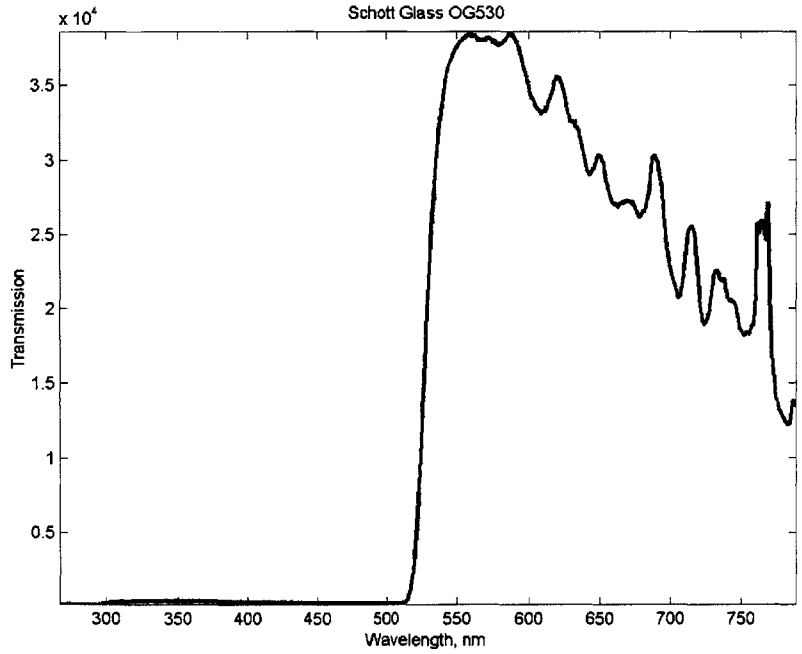

Figure B-6: Schott Glass orange high pass filter, with edge at 530 nm

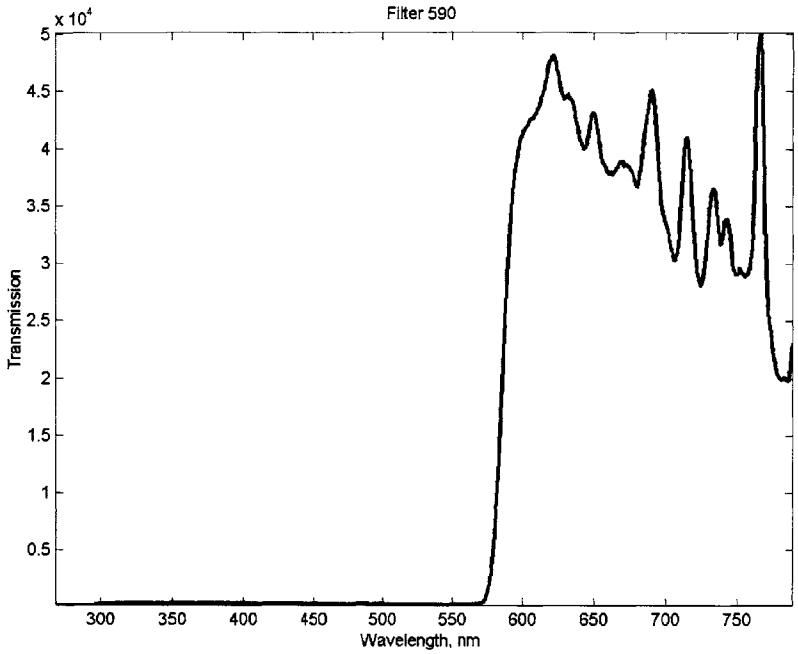

Figure B-7: Red doped glass filter, high pass with edge at 590 nm

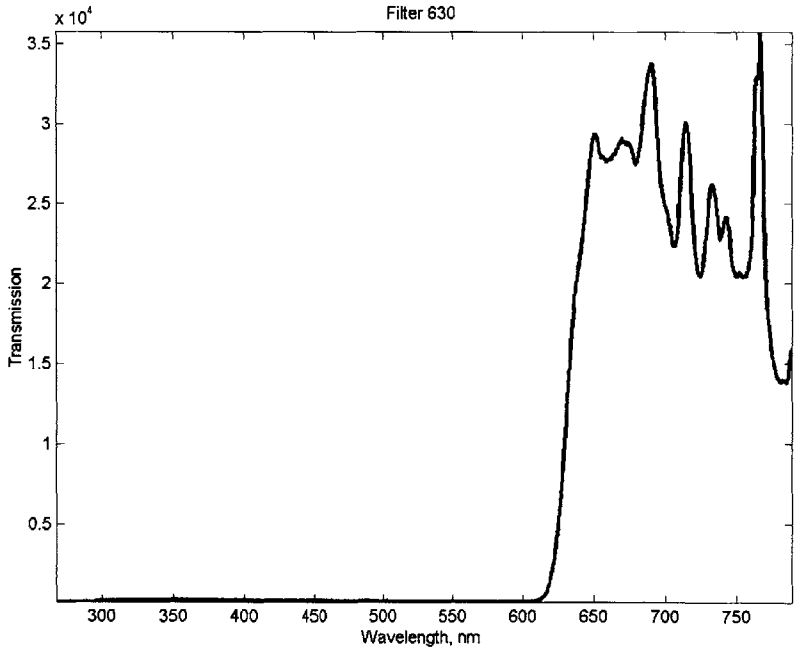

Figure B-8: Deep red doped glass filter, high pass with edge at 630 nm

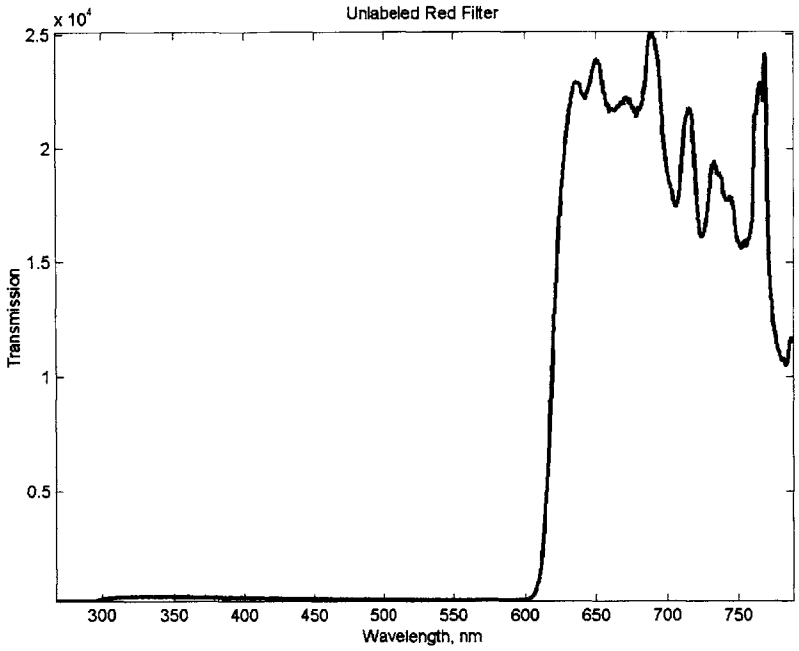

Figure B-9: Red doped glass filter, high pass with edge around 610 nm

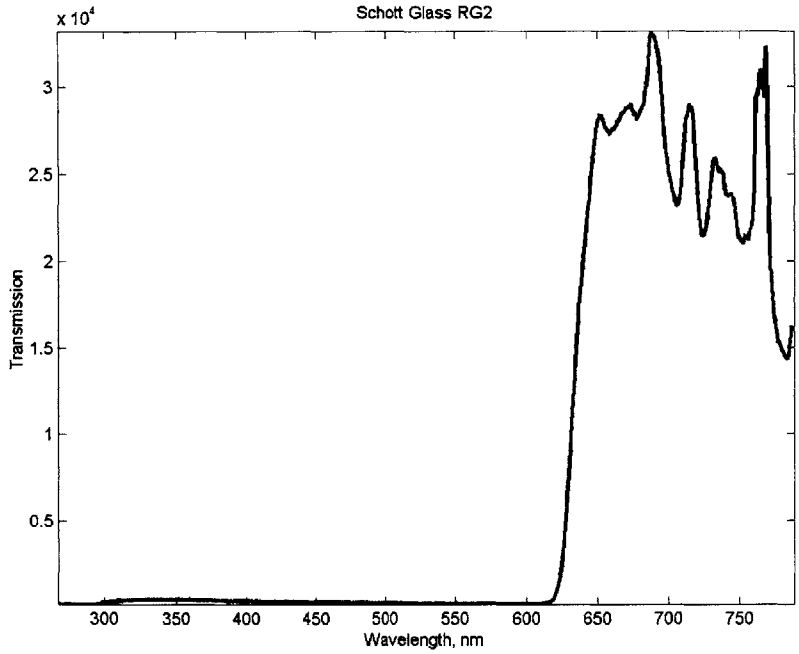

Figure B-10: Schott glass red high pass filter, edge around 625 nm

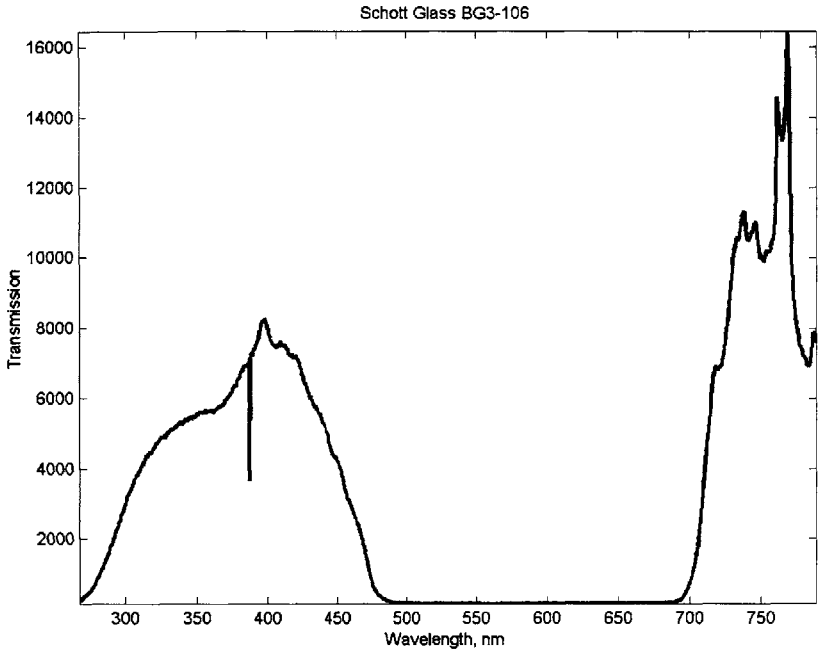

Figure B-11: Band gap filter, omitting transmission throughout most of the visible spectrum

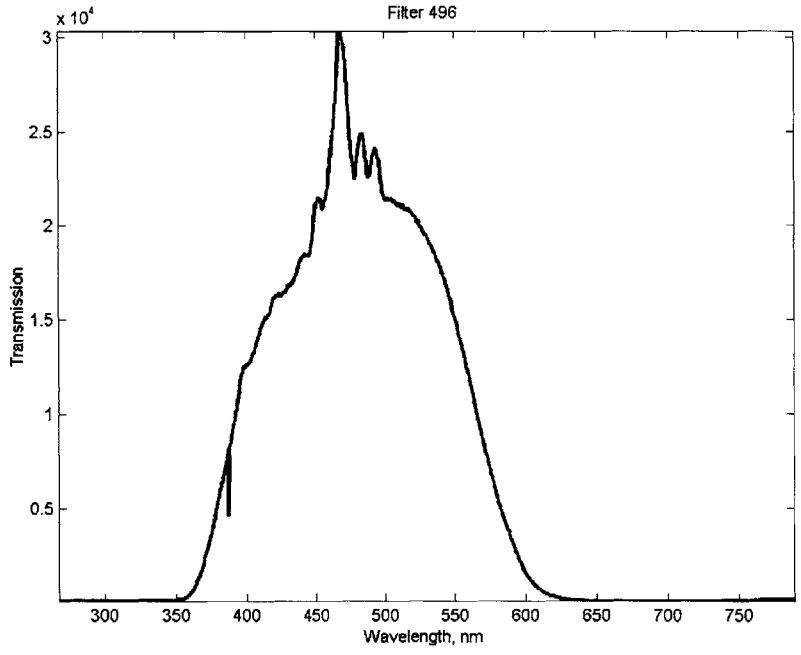

Figure B-12: Blue-colored doped glass filter, band pass filter centered at 496 nm, passing most of the visible spectrum

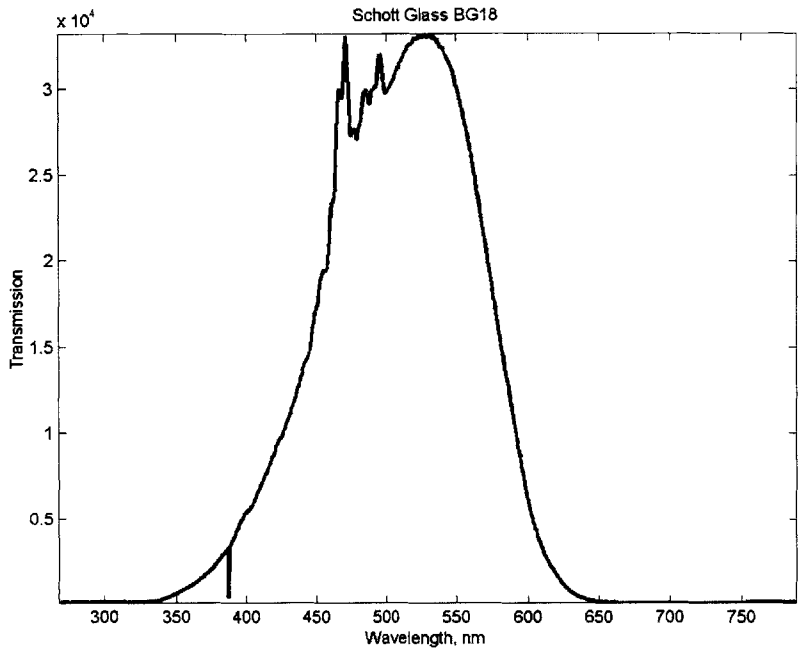

Figure B-13: Band pass filter, allowing transmission from blue to red

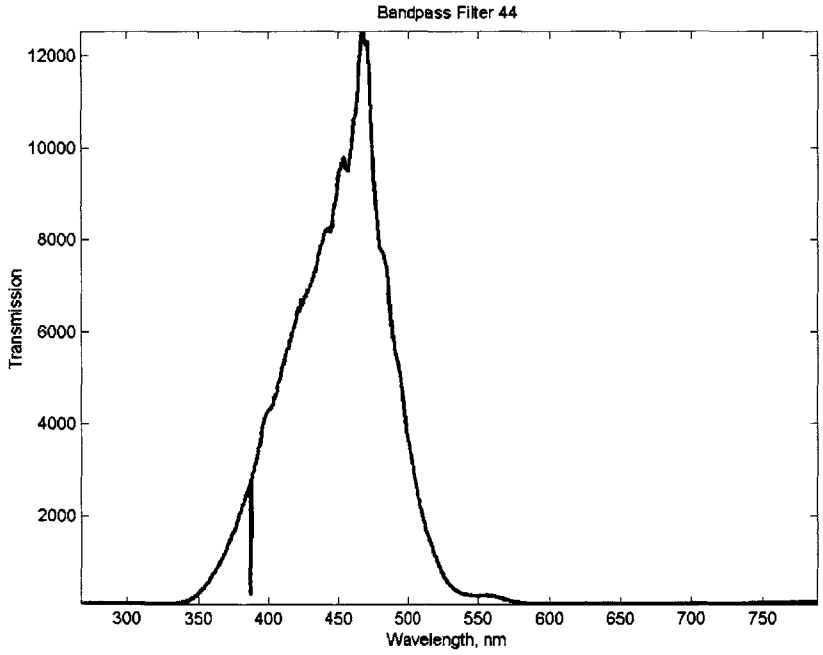

Figure B-14: Sharp band pass filter, transmitting through violet, blue and green

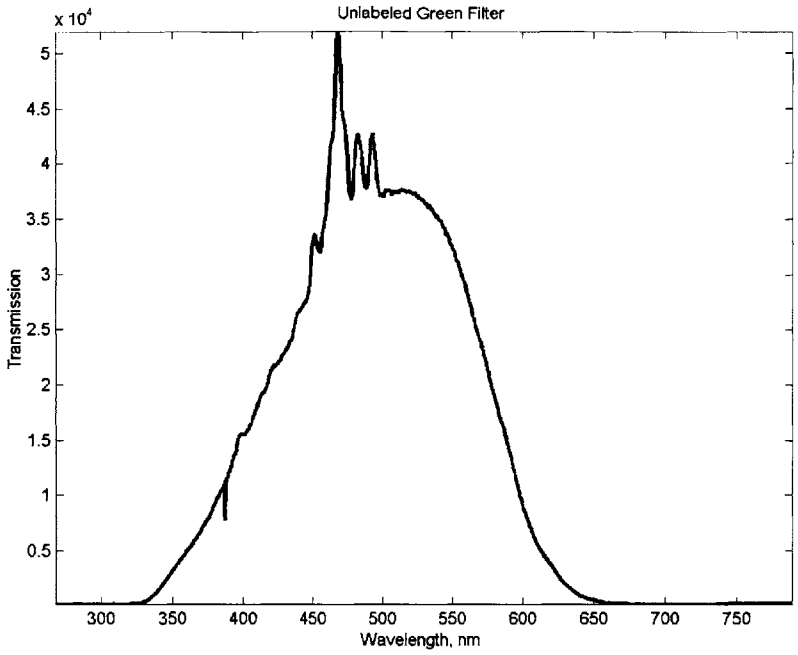

Figure B-15: Green band pass filter, unlabeled, centered at 500 nm, transmitting the entire visible range

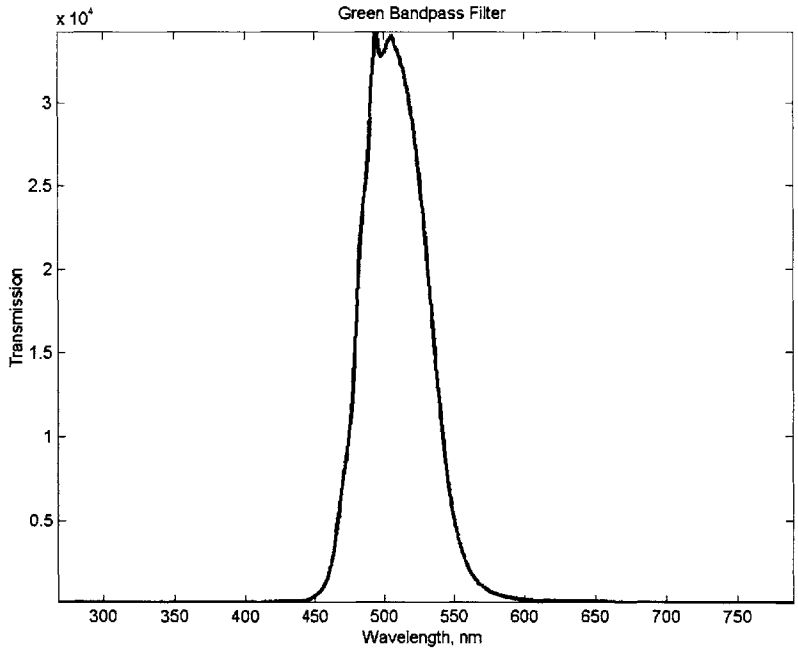

Figure B-16: Narrow band pass, green-colored "notch" filter, centered at 520 nm

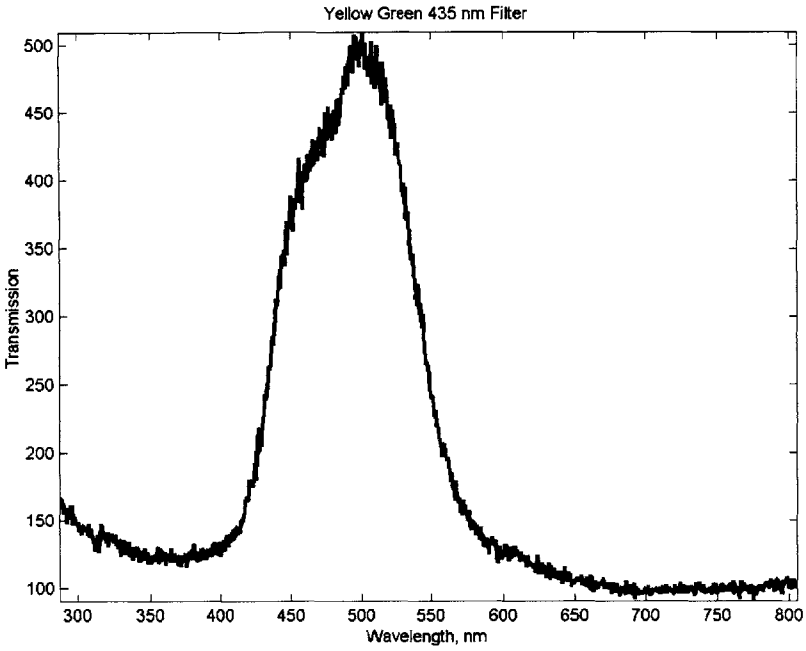

Figure B-17: Yellow-green doped glass filter, band pass centered around 500 nm, transmitting violet, blue and green

with several hours of use.

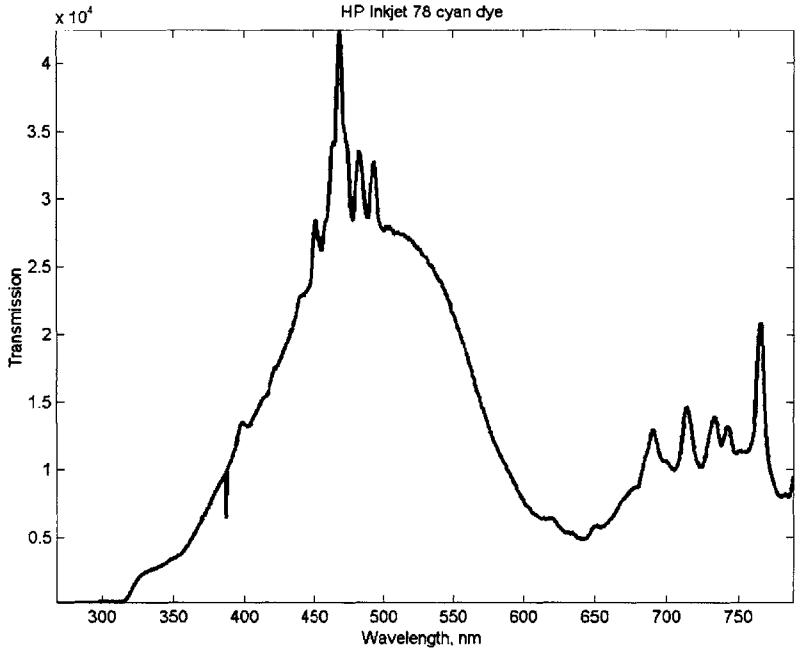

Figure B-18: Inkjet printed filter, HP cartridge No. 78, Cyan dye at maximum opacity

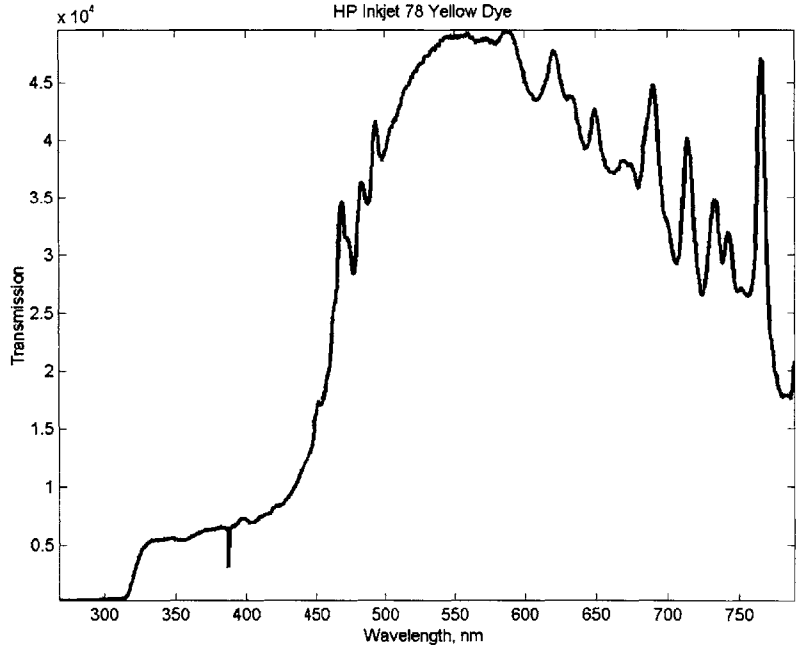

Figure B-19: Inkjet filter printed on transparency film, HP cartridge No. 78, Yellow dye at maximum opacity

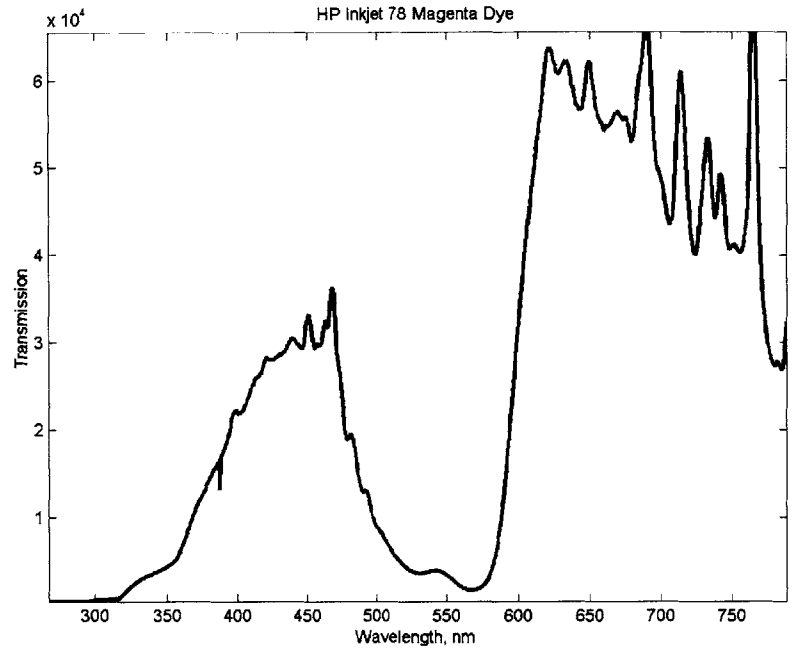

Figure B-20: Inkjet printed filter, HP cartridge No. 78, Magenta dye at maximum opacity

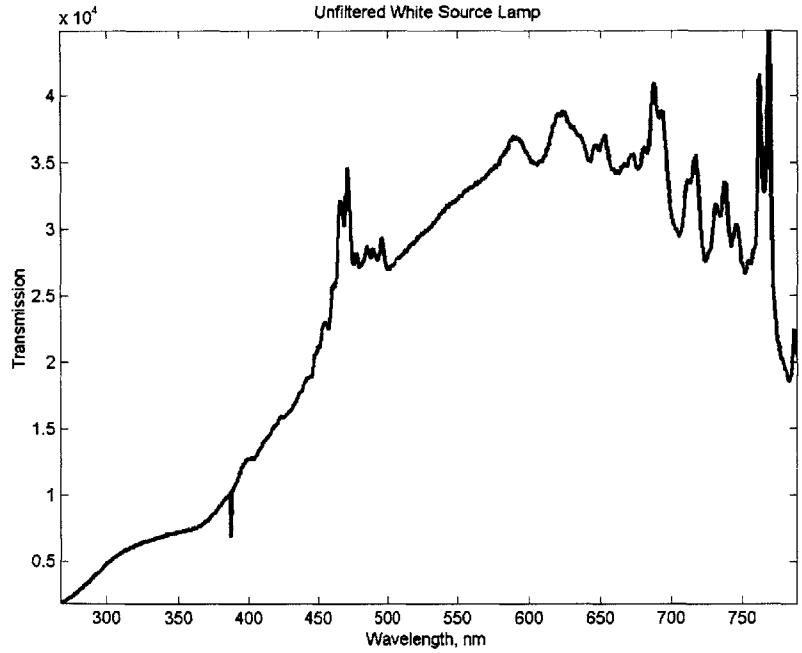

Figure B-21: White Light Stable Lamp Source, unfiltered

Appendix C

Select MATLAB Code

C.1 Read Oscilloscope Data

For this code excerpt, note that the oscilloscope does not allow naming of files. The first file is saved as "TEK00000.CSV' and each additional file is saved with an incremental number. In order for the code below to work properly, all the original oscilloscope data files must be in the same directory as the code file, and the indeces "index_1" and "index_2" must specify the first and second digit of the largest data file number that you wish to be included.

```
% initialize counter
index = 1;

% read oscilloscope data

index_1 = 0;
    for index_2 = 0:1:9;
        filename = sprintf('TEK000%i%i.CSV',index_1, index_2);
        temparray = csvread(filename, 1, 0);
        eval(['TEK000',int2str(index_1), int2str(index_2), ' = temparray']);
        TEK(:,index:index+1) = temparray(:,2:3);
```

```
        index = index + 2;
    end

index_1 = 1; for index_2 = 1:1:5;
    filename = sprintf('TEK000%i%i.CSV',index_1, index_2);
    temparray = csvread(filename, 1, 0);
    eval(['TEK000',int2str(index_1), int2str(index_2), ' = temparray']);
    TEK(:,index:index+1) = temparray(:,2:3);
    index = index + 2;
end
```

C.2 Isolate One Period of Transmission Matrix Signal

```
% Isolate only one period of transfer matrix detected signal

res = 1000;       %expanded resolution of signal

TEKDIFF = diff(TEK);

SZ = size(TEK); SZD = size(TEKDIFF);

TEK_1 = TEK(1:ceil(SZ(1)/2),:);

TEK_2 = TEK(ceil(SZ(1)/2) +1:SZ(1),:);

TEKD_1 = TEKDIFF(1:ceil(SZD(1)/2),:); TEKD_2 =
TEKDIFF(ceil(SZD(1)/2) + 1:SZD(1),:);
```

```
for trace = 1:1:SZ(2)/2
    [First,I] = min(TEKD_1(:,2*trace-1));
    [Final,J] = min(TEKD_2(:,2*trace-1));

    TEK_Temp = cat(2,TEK_1(I:ceil(SZ(1)/2),2*trace-1)',...
    TEK_2(1:J,2*trace-1)');

    [First,I] = min(TEK_Temp);
    [Final,J] = max(TEK_Temp);

    TEK_Temp2 = TEK_Temp(:,I:J);

    SZ_Temp = size(TEK_Temp2);

    TEK_Interp(:,trace) = (interp1(1:SZ_Temp(2),TEK_Temp2,...
    1:(SZ_Temp(2)-1)/res:SZ_Temp(2),'spline'))';
end
```

C.3 Import and Scale Data Signal

```
% Load spectral data
detectedsignal = csvread('TEK00012.CSV', 1, 0);

% Isolate one period
detectedsignal_diff = diff(detectedsignal);

SZ = size(detectedsignal);
SZD = size(detectedsignal_diff);

DS_1 = detectedsignal(1:ceil(SZ(1)/2),:);
```

```
DS_2 = detectedsignal(ceil(SZ(1)/2) + 1:SZ(1),:);

DS_diff1 = detectedsignal_diff(1:ceil(SZD(1)/2),2);

DS_diff2 = detectedsignal_diff(ceil(SZD(1)/2) + 1:SZD(1),2);

[First,I] = min(DS_diff1); [Final,J] = min(DS_diff2);

% Run partial signals together, for single-period signal
DS_Temp = cat(2,DS_1(I:ceil(SZ(1)/2),2)',DS_2(1:J,2)');

% Clip transition points
[First,I] = min(DS_Temp);
[Final,J] = max(DS_Temp);

DS_Temp2 = DS_Temp(:,I:J);

SZ_Temp = size(DS_Temp2);

% Interpolate to Size of transfer matrix
DS_Interp(:,1) = (interp1(1:SZ_Temp(2),...
DS_Temp2,1:(SZ_Temp(2)-1)/res:SZ_Temp(2),'spline'))';
```

C.4 Inversion Function

For the inversion of the transmission matrix, I used the standard Moore-Penrose pseudoinverse. The function is a built-in MATLAB call, such that the conditions for a pseudoinverse, also outlined in Section 3.2.5, are met. Namely,

$$ABA = A \tag{C.1}$$

$$BAB = B \tag{C.2}$$

$$(AB)^T = AB \tag{C.3}$$

$$(BA)^T = BA \tag{C.4}$$

The specific MATLAB function I used allows a tolerance input `tol`. The Moore-Penrose pseudoinverse algorithm in MATLAB is based on a singular value decomposition, `svd(A)`, and any singular values less than `tol` are treated as zero.

The MATLAB function itself is of the form, `B = pinv(A)` returns the Moore-Penrose pseudoinverse of A, and `B = pinv(A,tol)` returns the Moore-Penrose pseudoinverse and overrides the default tolerance, `max(size(A))*norm(A)*eps`.

```
% Use Detected Signals to Reconstruct Spectra
R_DS = pinv(TEK_Interp,9e-13)* DS_Interp;
```

Bibliography

[1] A. Albert. *Regression and the Moore-Penrose Pseudoinverse.* Academic Press, New York and London, 1972.

[2] Max Born. *Principles of Optics.* Cambridge University Press, seventh edition, 1999.

[3] J. H. Correia. Single-chip cmos optical microspectrometer. *Sensors and Actuators*, 82:191–7, 2000.

[4] James W. Demmel. *Applied Numerical Linear Algebra.* Society for Industrial and Applied Mathematics, 1997.

[5] J. A. Fill. The moore-penrose generalized inverse for sums of matrices. *SIAM Journal on Matrix Analysis and Applications*, 21(2):629–35, 1999.

[6] L. Fox. *An Introduction to Numerical Linear Algebra.* Oxford University Press, 1965.

[7] Eugene Hecht. *Optics.* Addison-Wesley, third edition, 1998.

[8] H. L. Kung. Standing-wave transform spectrometer based on integrated mems mirror and thin-film photodetector. *IEEE Journal of Selected Topics in Quantum Electronics*, 8:98–105, 2002.

[9] L. Mandel. *Optical Coherence and Quantum Optics.* Cambridge University Press, Cambridge, 1997.

[10] D. L. Mills. *Nonlinear Optics.* Springer-Verlag, Berlin Heidelberg, 1991.

[11] Bahaa E. A. Saleh. *Fundamentals of Photonics.* John Wiley & Sons, 1991.

[12] Anthony E. Siegman. *Lasers.* University Science Press, Mill Valley, California, 1986.

[13] Gilbert Strang. *Linear Algebra and Its Applications.* Harcourt Brace & Company, 1988.

[14] Ben G. Streetman. *Solid State Electronic Devices.* Prentice Hall, 2000.

[15] Lloyd N. Trefethen. *Numerical Linear Algebra.* Society for Industrial and Applied Mathematics, 1997.

[16] Lloyd N. Trefethen. *Spectral Methods in MATLAB.* Society for Industrial and Applied Mathematics, Philadelphia, PA, 2000.

[17] G. M. Yee. Miniature spectrometers for biochemical analysis. *Sensors and Actuators*, 58:61–6, 1997.

6318-69