Inverse Analysis of Light Scattered by Soot Aggregates

by

David William Burr

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Applied Science
in
Mechanical Engineering

Waterloo, Ontario, Canada, 2010

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The impact of soot on human health and the environment is a function of its size and morphology. Thus, it is important to have a method to quickly and accurately determine the aggregate size distribution of soot.

Elastic light scattering is considered as a method to determine the aggregate size distribution in an aerosol. The relationship between the scattered light and the aggregate size distribution is derived and a robust inversion method is presented. The method is validated against artificial data. It is then applied to experimental data from a flame condition at which a distribution obtained from TEM analysis exists, and found to work quite well.

Finally, optimization is applied to the experimental angles at which light is measured. The results showed that there is indeed an optimal angle setup, and that the error at that optimal setup is reduced over other angular setups.

Acknowledgements

I would first like to thank my Lord and Saviour, Jesus Christ, without whom I wouldn't be where I am today. In and above providing the means for my eternal salvation, He has been my rock through everything that life has thrown my way, providing guidance, love, and support. I can say with certainty that I would not have been able to complete this thesis, or the Masters Degree it is associated with, without Him.

I would also like to thank my supervisor, Kyle Daun, for all the help he has provided along the way; whenever I got stuck he would point me in an appropriate direction or have a new angle that I could try to get around the difficulty, and he was almost always helpful when it came to writing papers and this thesis. I am also extremely grateful that he offered me a position on his research staff when I arrived at the University of Waterloo.

I thank my fellow graduate students Adam Horsman and Sina Haji Taheri, who made life interesting (sword fights, Journey breaks, and general shenanigans stand out in my mind) and always listened as I tried to explain what I was doing, providing help if they were able.

Thanks also go to Oliver and Kevin, as well as the other guys at the National Research Council, for their help and guidance during my time there. I learned a great deal and thoroughly enjoyed getting some time in the lab.

I must also thank Mark and Keri and the boys for allowing me to invade their home for two years during my time here. Their support is greatly appreciated

Finally, I would like to thank my parents, who always believed in me and challenged me to do the best that I could in everything I set my hand to. Without their love and support I would likely not have made it this far.

Table of Contents

List of Figures

List of Tables

Nomenclature

Variables

${}_1F_1$	Hypergeometric function
$\mathbf{A}$	Discrete kernel matrix
a_{λ_l}	Absorption coefficient
C	Scaling coefficient
C_{exp}	Experimental scaling parameter
$\bar{C}_s^{agg}$	Average scattering cross section of a primary particle
$\bar{C}_{vv}^{agg}$	Average vertical polarization cross section of a primary particle
C_{vv}^{p}	Polarization cross section of a primary particle
D_f	Fractal dimension
d_p	Aggregate primary particle diameter
$F(\boldsymbol{m})$	Complex scattering function
$f(\cdot)$	Structure factor
f_v	Soot volume fraction
$g(\theta)$	Angular light scattering data
$\mathbf{g}$	Discrete measured light scattering data
$I(q)$	Intensity of light
I_{λ_l}	Source function
$I_{[0,\infty)}$	Indicator function
i_{λ_l}	Radiation intensity
k_g	Fractal prefactor
$K(\theta, N_p)$	Kernel function

K	Extinction coefficient
$\mathbf{L}$	Smoothing matrix
$\boldsymbol{m}$	Complex index of refraction
N_{agg}	Aggregate number density
N_p	Number of primary particles per aggregate
$N_{p,g}$	Geometric mean
$\mathrm{P}(\mathbf{g}\|\mathbf{x})$	Likelihood function
$P(N_p)$	Probability density function of the aggregate size distribution
q	Scattering wave vector
R_g	Aggregate radius of gyration
$\mathbf{x}$	Discrete probability density function of the aggregate size distribution
x_p	Size parameter

Greek Symbols

Γ	Gamma function
γ	Regularization (weighting) parameter for prescribed distribution
$\Delta\mathrm{s}$	Path of intersection between the detector view and the laser beam
$\Delta\omega_{\mathrm{laser}}$	Solid angle subtended by the laser beam
$\delta\mathbf{g}$	Perturbation of the measured data
$\delta\mathbf{x}$	Perturbation of the recovered distribution
η	Wave number
θ	Angle formed between the laser beam and the detector view
κ	Optical thickness
λ_l	Wavelength of incident light

λ	Regularization parameter
σ	Standard deviation
σ_g	Width of a lognormal distribution
σ_i	Singular values of the Kernel matrix
σ_s	Scattering coefficient
ϕ	Scattering phase function
Ω	Scattering albedo
ω	Finite solid angle

Chapter 1
Introduction

1.1 Fundamentals of Soot

Soot aggregates form through incomplete combustion; the presence of soot implies inefficiencies in a combustion process, and a potential risk to human health and the environment. This chapter will present the background of the study of soot aerosols, explain why soot is of interest, discuss the properties of soot, and finally put forward the basic experimental methodology on which this thesis is based.

The study of soot aerosols has shown two important reasons for understanding soot aggregate morphology: to achieve greater efficiency for combustion devices and to minimize particulate matter in the air that can have an adverse affect on the environment and human health.

One must understand how soot is formed to further understanding in these two areas. Soot particles are unburned fragments of fuel, made up mainly of carbon and hydrogen, which agglomerate into nanospheres. These spheres then coalesce to form soot aggregates through aerosol dynamics (Puri, et al. 1993). These aggregates are long chains of particles, called mass-fractals, that cannot be represented by simple geometric shapes. Two typical soot aggregates imaged via transmission electron microscopy (TEM) are shown in Figure 1.1.

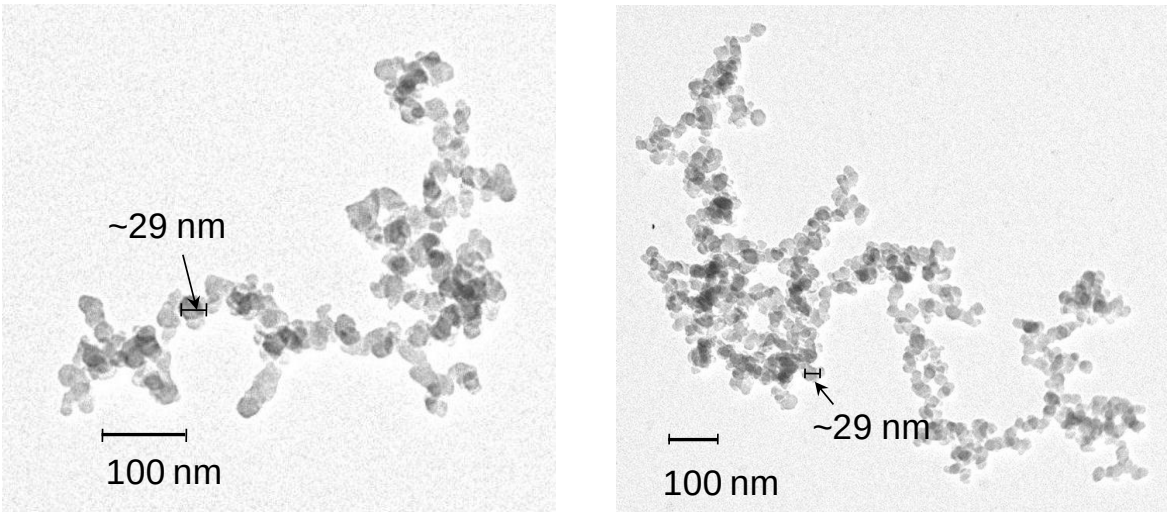

Figure 1.1 –Typical soot aggregates, courtesy of Dashan Wang, National Research Council, Canada

1.2 Motivation of Research

Soot is potentially harmful and represents inefficiencies in a combustion process. To understand the effect soot has in these two areas, an overview of the current research and usage of soot in four fields is presented: its role in combustion applications, its adverse affect on the environment, its potential harm to humans if inhaled, and its purposeful creation in some industrial applications.

1.2.1 Role in Combustion

Soot can be useful in certain combustion processes. For instance, soot is very efficient at radiating energy compared to gas-phase radiation, which is useful in a boiler. The soot within the flame radiates the energy of combustion to the walls where it is transferred to the working fluid. However, this principle is undesirable in an internal combustion engine, where losing heat to the walls decreases the efficiency of the engine and requires that materials capable of resisting higher temperatures be used (Shaddix and Williams 2007).

The radiative efficiency of soot can have devastating effects in certain situations. For example, this property can greatly increase the damage caused by a pool fire, where the fuel source is a pool of liquid hydrocarbons. Because of the way the air moves around a pool fire the flame is fuel rich, which promotes soot formation. The soot then rises above the pool and radiates most of its energy back to the liquid. This vaporizes more fuel, which generates more soot, and the cycle continues. This radiative feedback mechanism increases the rate at which the fuel combusts and also increases the radiant loading of surrounding surfaces. This effect makes such fires difficult to extinguish, and in many cases they must simply be allowed to burn until the fuel source is depleted. This is often the case when vehicles that have a great deal of fuel on board crash (Shaddix and Williams 2007).

Furthermore, the presence of soot in any combustion reaction inherently means that the reaction is not continuing as efficiently as possible, as soot is formed from unburned fuel. There is a drive to use fuel as efficiently as possible, and a method to quantify the soot aggregates present within the exhaust gas in situ would be invaluable.

1.2.2 Environment

Soot has been tentatively linked to climate change. Once cooled, soot is very dark in colour and is highly absorptive. If soot settles on a surface it can greatly increase the heat absorbed. This can raise the temperature of the surrounding area (this is similar to the “heat island” effect in cities due to dark-coloured roads and buildings) and in the case of snow or ice, may promote melting. It is speculated that soot deposition is one of the reasons why the polar ice caps are melting faster than the current models predict (Jacobson 2002). Some larger soot aggregates can form condensation nuclei for water droplets in the atmosphere, promoting cloud

formation. Increasing cloud formation can lead to global dimming, where the increased cloud cover reflects more of the sun's energy, reducing the energy that reaches the surface of the planet. This is especially important because light is necessary for photosynthesis in plants and hence for all life. A decrease in the amount of light may potentially reduce the productivity of a plant. As well, increasing clouds might cause an increase in rainfall, changing the weather patterns of a given area. These two effects combined might have some small affect on crops, but most research so far points to the change in crop productivity being minor with small to moderate changes in global dimming. In fact, global dimming might, in some specific cases, cause an increase in crop productivity (Stanhill and Cohen 2001).

Other studies have suggested that global dimming may have a much stronger effect. One of the largest factors in climate change is thought to be the Surface Solar Radiation (SSR), which is the amount of solar energy reaching the surface of the earth. The SSR is directly affected by global dimming (Wild 2009). In turn, global dimming is thought to be affected by the presence of particulate matter in the atmosphere (above and beyond the somewhat minor effect of cloud nucleation by particulate matter (Wild 2009)). A reduction in the SSR can potentially lead to reduced winds and rainfall, such as was seen in the sub-Sahara in northern Africa in the 1980's, which caused drought and famine that affected millions (Folland, Palmer and Parker 1986). Research is ongoing as global dimming directly counteracts global warming, making it appear to be a good thing. There is a tentative link between global dimming and disastrous climate change, however, and a great deal of research is being carried out on the topic to quantitatively determine its impact. One of the biggest problems is a lack of data on the particulate matter in the atmosphere going back far enough to make conclusive arguments (Wild 2009).

1.2.3 Health Effects

Soot plays an important role in human health. The effects of soot aggregates on human health and the environment have been debated for years. Pope and Dockery (2006) performed a comprehensive review of the studies concerning the effects of particulates on human cardiovascular health that had been carried out up to that time. They noted that extreme cases of air pollution were linked with an increase in both the frequency and severity of cardiovascular disease. The existence of this link is no longer in question, although there is still uncertainty as to the level of particulate matter in the air that causes morbidity. For a number of decades after the link was determined it was thought that mid- to high-level concentrations of particulate matter were needed to cause disease. Recent work, however, has shown that even low levels of particulate matter can also potentially cause disease; this has lead to the conclusion that aggregate morphology is just as important as the particulate mass and volume fractions in human cardiovascular health.

This is particularly worrisome given the increase in smog in urban centres. Smog is made up of particulate matter as well as water droplets condensed around aggregates. Such particulate matter, when inhaled, has different effects on the body, depending on aggregate size. Very small particulates are generally less soluble than large aggregates, which may allow them to translocate from the lungs directly into the bloodstream across the pleural membrane rather than being trapped in the lungs as larger particles tend to be. Long-term build-up of aggregates in the lungs can lead to breathing trouble as well as infection, and even death (Pope and Dockery 2006).

1.2.4 Industrial Applications

While soot can have adverse effects on human health and the environment, some industrial processes utilize the beneficial properties of specifically generated particles, known as carbon black. These particles differ from flame-generated soot in their chemical makeup; carbon black is near to 97% carbon, whereas soot is generally less than 60% carbon. For years these aggregates have been added to certain rubbers, generally as hardening agents in tires, as well as for pigment in ink used in printer toner. The industrial processes that produce carbon blacks must generate aggregates of a specific size, as the desired functionality depends on the size of the aggregate. As well, soot aggregates were the original source for carbon nanotubes and bucky balls, which are of significant interest to those studying nanotechnology, due to their unique engineering properties (Shaddix and Williams 2007, Carbon Black User's Guide 2004).

1.3 Experimental Methodology

This section presents a more in-depth look at the fractal nature of soot aggregates and lays out light scattering theory, which is used to recover the probability density function of the aggregate size distribution. The ill-posedness of the system is then considered, along with methods for managing it.

1.3.1 Light Scattering From Fractal Soot Aggregates

When flames were first studied scientifically, it was assumed that soot aggregates could be modelled as spheres, or some other such simple shape. This can be seen in the early work in the field of light scattering from soot aerosols (Gans 1928, P. Debye 1944, LaMer and Sinclair 1949, Engelhard and Friess 1949). It wasn’t until much later that the theory of fractals, developed by Benoit Mandelbrot (1977), was applied to describe soot aggregates (Forrest and

Witten 1979). Fractals are recursive in nature, such that viewing an aggregate at different magnifications reveals a similar structure. Modeling soot aggregates as fractals has provided a mathematically rigorous framework in which to work.

If a fractal nature is assumed, the number of primary particles per soot aggregate, N_p, is related to the radius of gyration, R_g, of an aggregate by

$$N_p = k_g \left(\frac{2R_g}{d_p} \right)^{D_f} \tag{1.1}$$

where k_g is the fractal prefactor, D_f is the fractal dimension, denoting the "openness" of an aggregate ($D_f = 3$ is a solid sphere and $D_f = 1$ is a line), and d_p is the diameter of the primary particles in an aggregate. It is generally assumed that d_p is constant (Brasil, Farias and Carvalho 1999), so the aggregate size distribution for an aerosol varies by N_p alone.

To determine the impact of a soot aerosol on human health and the environment, the number of aggregates in the aerosol as a function of the number of primary particles per aggregate, called the aggregate size distribution, must be determined. One way to do this is to collect aggregates from the flame and view them using electron microscopy as done, for example, by Dobbins and Megaridis (1987). However, this method is time consuming, and certain aggregate sizes may be preferentially attracted to the sampling slide. A sampling of TEM images are shown in Figure 1.1. Another method is to shine light through the aerosol and measure the angular distribution of the scattered light, using the experimental apparatus shown in Figure 1.2. The laser shines a specified wavelength of collimated and vertically polarized light at through the flame atop the burner. The light is scattered by the soot, and the detector measures

this scattered light. The measurement is then converted into the intensity of the light scattered in the measurement direction, which is the experimental data used throughout this thesis.

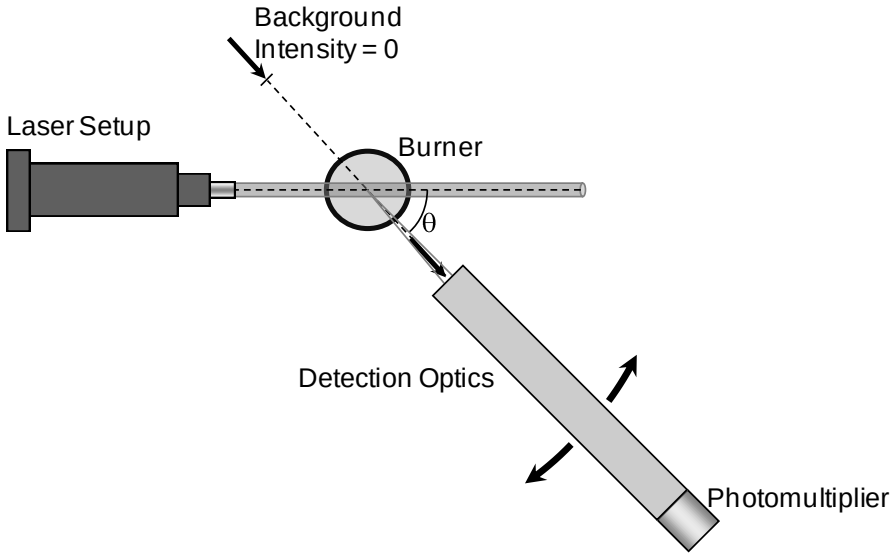

Figure 1.2 – Experimental Apparatus

The angular scattering measurements are strongly related to the aggregate size distribution. Each aggregate scatters light in a different angular distribution, depending on the morphology. The angular scattering from three different sized aggregates is shown in Figure 1.3. An aerosol is made up of a wide range of aggregates of varying sizes, and the measured signal is the integration of the scattering from all aggregate size classes combined. This makes recovering the aggregate size distribution difficult, and necessitates measuring the scattering at multiple angles, to help narrow down the possibilities.

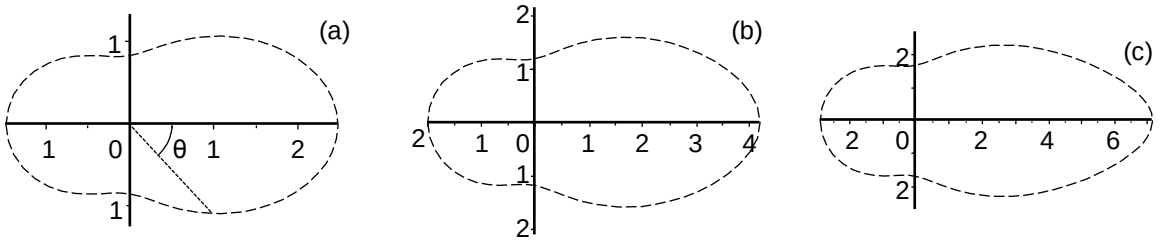

Figure 1.3 – Angular scattering for an aggregate of (a) 50 primary particles, (b) 125 primary particles, and (c) 250 primary particles

1.3.2 Ill-posedness and Regularization

The dependence of the scattered light on the probability density function of the aggregate size distribution is described by a Fredholm integral equation of the first kind,

$$g(\theta) = C \int_1^\infty K(\theta, N_p)\, P(N_p) dN_p \tag{1.2}$$

where $g(\theta)$ is the angular light scattering data, C is a scaling constant, $K(\theta, N_p)$ is the kernel function that contains system information, and $P(N_p)$ is the probability density function of the aggregate size distribution. The term $P(N_p)$ physically describes the probability of a given aggregate in an aerosol having a certain number of primary particles. The determination of this unknown distribution is the end goal of this thesis.

What makes this quantity so difficult to recover, however, is the ill-posedness of the problem. Hadamard (1923) was the first to define a mathematically well-posed problem as one that has a solution that is both unique and stable. Inversion of Eq. (1.2) is ill-posed because its solution is not unique or stable. Instead, there are an infinite number of solutions that explain the data within experimental accuracy. This is caused by a 'blending' property of the kernel, whereby specific details in the aggregate size distribution are smoothed or smeared into much more subtle details in the measured data. These smeared, subtle effects are generally much

smaller in magnitude, and can easily be lost among the inherent noise in the measurements. While the inversion process can amplify such subtle details, in doing so the noise would also be amplified, leading to large variations in the aggregate size distribution recovered.

Essentially, ill-posedness arises from the fact that the governing equations provide inadequate information to uniquely specify a solution due to the smearing of the kernel function. Accordingly, one must add additional assumed information to recover a solution, a process called regularization. There are two main classes of regularization methods: linear and iterative. A popular linear method, Tikhonov regularization, work by converting Eq. (1.2) into a matrix equation,

$$\mathbf{Ax} = \mathbf{g} \tag{1.3}$$

and adding another matrix to the system. In this equation, **A** is the discrete form of the kernel matrix, **x** is the discrete aggregate size distribution, and **g** is the measured light scattering data at various discrete angles. The second matrix adds another set of equations to the system that push the solution in the direction of what is assumed to be the true solution. For instance, a matrix **L** that promotes smoothness can be added to give

$$\begin{bmatrix} \mathbf{A} \\ \lambda \mathbf{L} \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{g} \\ \mathbf{0} \end{bmatrix} \tag{1.4}$$

where λ is the regularization parameter that controls the importance of the additional equation on the solution. The solution to this augmented matrix equation will best satisfy both the original equation and the additional equation.

Another popular linear method is Truncated Singular Value Decomposition (TSVD). This method works by decomposing the matrix equation into orthonormal basis vectors and considering the singular values, which are weights that determine how strongly the solution

depends on a given basis vector. If the singular value is small (relative to the other singular values) then it is truncated, and the solution does not depend on the associated basis vector. This prevents large and highly oscillatory basis vectors from dominating the inverse problem.

One final linear method is Conjugate Gradient Iteration, which is an iterative linear method that is guaranteed to solve an $n \times n$ matrix equation in n steps or less. The method works by forming the residual between the current solution and the true solution, which is used to form both the step length and step direction for the next iteration. The step is taken, and the process repeated, up to n times. This method is only applicable to matrices that are both symmetric and positive definite.

Iterative methods are the other major class of regularization. These methods start from an initial guess of the solution and progressively move towards the true solution. Many such methods use the residual between the current guess and the measured data to determine the step length at each iteration, and some form of the Jacobian matrix to determine the step direction (Hansen 1998).

A problem arises when adding one piece of information is not adequate to recover the solution. Bayesian inference was employed to add multiple assumed characteristics to the system, known as priors, in a mathematically rigorous manner. The probability of a given solution being correct is

$$P(\mathbf{x}|\mathbf{g}) = \frac{P(\mathbf{g}|\mathbf{x})}{P(\mathbf{g})} P_{\text{prior}}(\mathbf{x}) \tag{1.5}$$

where $P(\mathbf{x}|\mathbf{g})$ is the probability of $\mathbf{x}$ being correct given $\mathbf{g}$, $P(\mathbf{g}|\mathbf{x})$ is the probability of $\mathbf{g}$ occurring given a distribution $\mathbf{x}$ (which is also called the likelihood function), $P(\mathbf{g})$ is the

marginal probability of the data and $P_{\text{prior}}(\mathbf{x})$ is the probability that the solution is correct given a prior. Other regularization methods can be derived from Bayesian inference; for instance, Eq. (1.4) is derived from this equation when $P_{\text{prior}}(\mathbf{x})$ is a smoothness prior.

Multiple priors (from multiple known features of the system) are simply multiplied onto Eq. (1.5) to give an overall probability of the solution being correct. This method can always be written in an iterative form, but in many cases a linear form can also be derived. Due to the complexity of the problem, the linear form is generally solved iteratively as well, such that a solution is found and checked for validity, something is changed and the problem is solved again. This is advantageous as solving a linear equation many times takes far less computational power than solving a non-linear iterative equation.

1.4 Scope of Problem

The present study will recover the aggregate size distribution from an ethylene co-flow diffusion flame using light scattering data measured at different angles. Bayesian inference will be employed to add information to the system in a mathematically rigorous manner. The angles at which the scattering data should be collected will be optimized to produce the least ill-posed system, although theses optimal angles are not tested experimentally.

1.5 Outline of Thesis

The remainder of this thesis is divided into six main sections.

Chapter 2 presents a review of the relevant literature in two distinct areas: first, the use of light scattering to determine properties of soot aggregates; and second, a review of the inversion techniques that have been employed to solve for the soot aggregate size distribution.

Chapter 3 presents the derivation of the governing equations, beginning at the radiative transfer equation. As well, a comparison of the different structure factors is laid out, sources of uncertainty are mentioned, and the nonlinearity of the governing equation is discussed

In Chapter 4 the experiment is explained in detail, covering each optical component the light travels through between the laser and the detector. The spatial resolution of the measurements is discussed, as are sources of uncertainty.

In Chapter 5 the method for applying regularization to the system is laid out, both for the linear and the nonlinear case, an artificial data set is generated and inverted as a proof of concept, and the experimental data is then inverted.

Chapter 6 presents the optimization of the experimental procedure and apparatus, considering the size and conditioning of the matrix equation. The sensitivity matrix is derived and applied to the system to optimize the angle set, and the resulting angles are shown, theoretically, to reduce the regularization and perturbation errors inherent in the system.

Finally, Chapter 7 summarizes the results obtained in Chapters 5 and 6, and considers the benefits of employing regularization to invert this problem. Recommendations for future work are presented.

Chapter 2
Literature Review

2.1 Introduction

The goal of this chapter is to present the literature relevant to the current research. There are two key sections to this review: the development of light scattering theory from soot aggregates, and application of inverse analysis to infer aerosol characteristics from light scattering experiments. In the section covering light scattering from soot aerosols, the development of light scattering as a diagnostic technique for recovering the size of soot aggregate size distributions will be presented. The section covering inverse analysis of light scattering from aerosols will deal with both linear and nonlinear regularization schemes as applied to the recovery of the soot aggregate size distribution from light scattering data, and lay out other methods that can add additional information to the system, reducing the amount of regularization required.

2.2 Light Scattering From Soot Aerosols

Using scattered light to infer attributes of aerosolized nanoparticles was pioneered by Tyndall (1869). His work showed that microscopic particles scatter concentrated beams of light. He employed this technique to determine if there were such particles in a given sample of air. Soon after that, Lord Rayleigh expanded the theory, as reviewed by Gans (1928). In the early 1940's a number of scientists began to explore the use of light scattering to determine particle size within an aerosol (P. Debye 1944, LaMer and Sinclair 1949, Engelhard and Friess 1949).

Debye (1946) was one of the early proponents of light scattering as an experimental technique, although he did little work with soot aggregates. He developed one of the first angular light scattering measurement instruments (Debye 1946) along with techniques to determine homogeneities in solids via light and x-ray scattering (Debye and Bueche 1949). This led to the use of x-rays as a diagnostic technique for inconsistencies and structural deformities in human bones (Patel and Khanduala 1984).

Tabibian et al. (1956) looked at light scattering from single spheres, using Mie theory to correlate the measured scattering to the physical spheres that caused it. Mie theory was developed by Gustov Mie (1908) to work with single spheres that are too large to satisfy Rayleigh scattering, and that are approximately the same size as the wavelength of light. The theory is an analytical solution of Maxwell's equations dealing with scattering of electromagnetic radiation, and is covered in many sources, for example (van de Hulst 1981, Bohren and Huffman 1983). They found that Mie theory almost exactly explained the scattering they observed experimentally, and therefore concluded that Mie theory was accurate for spherical particles. They did notice that the refractive index of the scatterers should be known very accurately to apply the theory.

Early work showed that the aggregation of colloids formed aggregates that were not spherical (see, for instance, (Wada 1967), (Granqvist and Buhrman 1976)). Thus, while an approximation could be made that soot aggregates were spherical with an effective radius, satisfying Mie scattering, the resulting predicted scattering did not match experimental observations (Dobbins and Megaridis 1991).

Because aggregates are long chains of primary particles classified as mass-fractal in nature, a more accurate theory was needed to describe how they scattered light. First, a model for the scattering and absorption cross-sections was needed. These cross-sections essentially describe the 'apparent' area of an object that is presented to incoming radiation in relation to the object's capacity to affect the path of the light (Siegel and Howell 2002). There was a great deal of work in this field, such as that undertaken by Charalampopoulos and Chang (1991), Ku and Shim (1991), and Dobbins and Megaridis (1991), resulting in equations that relate light scattering and absorption by soot aggregates to the size of the aggregates. Eymet et al. (2002) presented a straightforward representation of these cross sections, which were used in this thesis.

The physics of light scattering by an aggregate is not trivial. The intensity of scattered light is dependent on the size of the aggregates, the electromagnetic properties of the soot, the wavelength of the scattered light, and the angle of scattering. The variations in the scattering can be split into different regimes, pictured in Figure 2.1, which is a plot of the modulus of the scattering wave vector squared, defined as $q \equiv 4\pi \sin(\theta/2)/\lambda_l$ where λ_l is the wavelength of the incident light, versus the relative intensity of the light, $I(q)/I(0)$.

The Rayleigh regime consists of very small particles, much smaller than the wavelength of light, and small-angle scattering. The scattering of light in this regime is invariant with respect to the scattering wave vector, and hence the relative intensity of the light is also invariant, giving a horizontal line graphically.

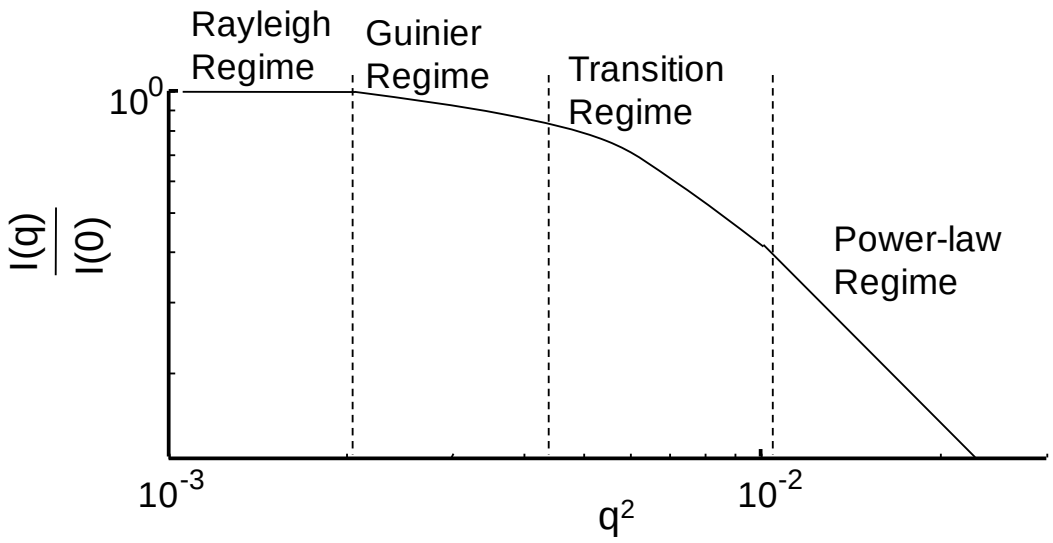

Figure 2.1 – Light scattering regimes

The Guinier regime is defined as occurring when $qR_g \ll 1$, where R_g is the radius of gyration of the aggregate. This describes small-angle scattering where

$$\frac{I(q)}{I(0)} \approx 1 - \frac{q^2 R_g^2}{3} \tag{2.1}$$

and was originally considered for small-angle x-ray scattering by André Guinier (1955), but then expanded to aggregates in general. The relative intensity in the Guinier regime is almost linear when plotted against q^2 on the horizontal axis, giving rise to Eq. (2.1). This allows the effective radius of gyration to be easily determined.

For large aggregates and angles, the slope again becomes linear with a value of $-D_f$, as the relative light scattering becomes invariant to the size of the aggregates. This section is known as the power-law regime. Only two measurements in this regime are required to fully describe it.

The physics underlying the transition regime is quite ambiguous. It is bounded by the Guinier regime and the power-law regime. There is no simple equation to describe it, and there

are no particularly useful known graphical features in this regime. The regime is exaggerated in Figure 2.1 to show its positioning.

A function known as a structure factor is defined that carries information about the underlying structure of the particulates that scattered the incoming light. This structure factor must span all of the regimes because it must describe the scattering of light in every regime. A great deal of work has been done on developing a structure factor that is consistent with fractal aggregate theory (Sorensen 2001, Weiz and Oliveria 1984, Schaefer, et al. 1984, Freltoft, Kjems and Sinha 1986, Lin, et al. 1989). A more complete review of the structure factor is carried out in Section 3.3.

The scattering of light from soot aggregates falls under Rayleigh-Gans scattering, which works for composites consisting of smaller elements. Each smaller element is assumed to be a Rayleigh scatterer that scatters independently from the rest of the particles in the aggregate, and that the refractive index of each smaller element is similar to that of the surroundings.

When Farias et al. (1995) assembled these light scattering concepts into a consistent theory, it was named the Rayleigh-Debye-Gans Fractal Aggregate (RDG-FA) theory, which encompasses the light scattering work of Debye and the Rayleigh-Gans scattering nature of soot aggregates together. This theory is accurate for single scattering aggregates consisting of spherical primary particles that do not overlap. A number of studies have considered the validity of the RDG-FA theory, finding it to provide scattering properties within 10% to the more accurate but more computationally intense integral equation formulation for scattering (IEFS) theory (Farias, Köylü and Carvalho 1996, Eymet, et al. 2002). These studies also found that

assuming an equivalent volume sphere is only valid for aerosols with little aggregation, or for small aggregates.

2.2.1 Soot Characterization from Light Scattering Experiments

In contrast to other metrology techniques, light scattering can be used to characterize a flame in a non-intrusive way. Jagoda et al. (1980) used light scattering to characterize polystyrene and polypropylene flames, and found that the concentration of soot aggregates was largest near to the fuel-rich portion of the flame front, where the actual combustion was occurring. In the central region of the flame the concentration was much lower.

Schaefer et al. (1984) and Martin et al. (1986) applied the theory of fractal aggregates to soot particulates, using both light and x-ray scattering measurements, to recover the fractal dimension, which was in agreement with the theoretical power-law decay for the structure factor.

Sorensen et al. (1995) used light scattering to characterize the crossover or transition regime between Guinier and power-law regimes, which allowed them to calculate the soot aggregate distribution width in the aerosol. However, the fractal dimension and the structure factor cut-off function must be known accurately to employ this method.

In characterizing light scattering in a flame, Sorensen's review of the subject (Sorensen 2001) should be considered. While it is by no means an exhaustive review, it is extensive and is a very good starting place. The review begins with the physics of light scattering and builds a heuristic model of light scattering. The review includes a great deal of discussion on formulating both a monodisperse and a polydisperse structure factor, the complex index of refraction, and how light scatters from a fractal aggregate. Note, however, that Sorensen (2001) does not provide an explicit formulation for the scattering cross sections, nor a derivation of the kernel

function, neither of which are necessary if a specific distribution type (such as a scaling, self-preserving distribution) is chosen to represent the soot aggregate size distribution in the aerosol.

2.2.2 Expanding the Range of Measurements

One of the drawbacks of a light scattering experiment is that the range that the scattering measurements can be taken at is physically restricted. It is useful to measure data in each of the regimes to properly characterize the scattering, and a number of researchers have put effort into expanding the angles at, or the regimes in, which scattering can be measured.

Sztucki et al. (2007) made use of x-ray scattering at small angles to greatly increase the range of the scattering wave vector. They also employed a custom-designed analyzer crystal that allowed them to make narrow-angle measurements, well below what is generally possible using commercially available equipment. They used this apparatus to study the growth of soot primary particles, which they found to reach a maximum radius of gyration of approximately 27 nm; they also studied the growth of aggregates in the aerosol, which they found to have a maximum radius of gyration of approximately 250 nm. The small angle x-ray measurements allowed for a complete characterization of the power-law regime so that the fractal dimension could be determined experimentally.

Oltmann et al. (2009) made use of mirrors to expand the angular range at which scattering measurements could be made. The method allows for almost simultaneous measurements over a wide range of angles, which might be employed using a pulsed laser to study turbulent flames.

2.3 Inversion of Soot Aerosol Data

The end goal of taking light scattering measurements from soot aerosols is to characterize the flame that scattered the light. This can take the form of determining the soot aggregate concentration, the fractal dimension, or, as is the goal of this thesis, the aggregate size distribution.

2.3.1 Applications

Using simulated aggregates Mountain and Mullholland (1988) showed how light scattering measurements could be employed to find the soot aggregate concentration, fractal dimension, size and radius of gyration, allowing for a more complete characterization of the aerosol.

Soot aggregates, of course, are not single spheres, but chains of spheres that are varying lengths. Dobbins and Jizmagian (1966) worked to develop a mean scattering cross section for a polydisperse aerosol, which they found to depend on the second and third moments of the distribution. They employed this technique to find the mean particle size, although they were unable to determine the distribution shape of the aerosol.

2.3.2 Linear Inversion Methods

Regularization theory has a long and rich background, both in mathematics and in the light scattering community. One of the first researchers to attempt to find a general inverse for an integral equation of the first kind, such as that given in Eq. (1.2), was Phillips (1962). Phillips created a linear inverse method that worked reasonably well if the unknown distribution were smooth. The method employed a simple quadrature rule method (Baker 1977) to convert the kernel function into a matrix, and added a smoothing matrix to augment the system. The purpose

of the smoothing matrix was to augment the information contained in the kernel with additional, prior information that the recovered distribution should be smooth.

Twomey (1963) took the work of Phillip (1962) and improved it by rewriting the equation such that only one matrix inversion need by computed, increasing the speed at which such an inversion could be completed. Furthermore, Twomey's method allowed for the computation of a solution in the case where the matrix produced by quadrature of the kernel was not square, greatly increasing its practicality.

In a recent paper, di Stasio et al. (2006) used Tikhonov regularization on small-angle x-ray scattering to recover the aggregate radius of gyration, related to the number of primary particles in an aggregate by Eq. **(**1.1), at different heights above burner. They made use of a software package called GNOM (Svergun, Semenyuk and Feigin 1988) to recover the distribution of aggregate radii of gyration. Because they were looking at low heights in the flame, little aggregate had occurred, and they were thus recovering the radius of gyration of non-aggregated primary particles or very small aggregates.

2.3.3 Iterative Inversion Methods

Grassl (1971) used an iterative method on spectral attenuation data to recover the aggregate size distribution found in aerosols. His work led to the iterative work of Twomey (1975). This work was originally purposed to find aggregate size distributions in the atmosphere, which are not smooth, so that the Phillip-Twomey linear inversion method was not applicable.

The distributions generated using Twomey's iterative method tend to be rough (Crump and Seinfeld 1982), which lead Markowski (1987) to propose an improved algorithm that added a second layer of iteration. After the Twomey iterative method had reached convergence, a

smoothing algorithm was applied and the smoothed function was used as the initial guess for Twomey's iterative algorithm. This was repeated until both a smoothness criterion and a convergence criterion for Twomey's method were met.

Newton's method was applied to the problem to form an iterative method, resulting in what is commonly known as Landweber iteration, (Hansen 1998, Baker 1977). This iterative method uses something akin to the Jacobian matrix from the Newton method as the kernel, but also employs a weighting scheme that can be set heuristically to promote convergence on a given problem. However, there are a number of schemes that have been presented to generate the weighting factors in a more rigorous manner. The stopping criterion is not well defined, however, which can lead to some problems in implementing this method.

Inversion of the soot aerosol system requires the addition of some form of information into the system to augment what is already there. One of the most common ways of doing this is to assume that the distribution is governed by a certain distribution type, such as a lognormal distribution or a scaling distribution (Sorensen 2001). This procedure introduces an inherent bias into the analysis, as the underlying distribution might not be the assumed shape. Furthermore, as shown by Link et al. (2010), there is a large set of parameters that give the same residual value, so that finding the best fit values is difficult, if not impossible.

2.3.4 Compound Methods to Stabilize the Inversion

Rather than using regularization, the ill-posedness of the inversion can be reduced by adding additional measurement data to the problem that is different from, and complementary to, the light scattering data. A number of such compound methods proposed. Combining data from relative light scattering and absolute light scattering, extinction measurements, or laser-induced

incandescence (LII) measurements can help to quantify the aggregate size distribution or the variables that are used to construct it.

D'Alessio et al. (1975, 1977) characterized the nucleation zone of a methane-oxygen flame using light scattering at different fuel to air ratios. They found that soot aggregates grew larger as they rose in the flame, both in terms of increasing numbers of primary particles and surface growth on the particles themselves. One prominent result presented was that scattering and extinction ratios allow for a more accurate measure of particle size than does the non-symmetry between the forward and backward scattering.

Other techniques have been developed that use LII to determine certain properties of the aerosol particles. Laser-induced incandescence makes use of a pulsed laser beam to heat refractory nanoparticles (most often aerosolized soot) to incandescence, and the decay of the radiance is measured. The decay curve can be compared to that of an aerosol of known volume fraction to obtain the particle volume fraction of the aerosol. While such techniques are not used directly in this thesis, valuable information can be obtained from them (Snelling, Smallwood, et al. 2005, Liu, Thomson and Smallwood 2008),. By analysing the time-resolved incandescence decay, the primary particle diameter of the soot aggregates can also be determined, which is useful in light scattering analysis. Techniques have been developed to incorporate both LII and light scattering diagnostics into a single instrument, such as that of Reimann et al. (2009).

Link et al. (2010) considered both relative and absolute intensity measurements, to overcome the inherent coupling between the distribution width and the fractal dimension. They point out that certain parameters, such as refractive index scattering function, $F(\boldsymbol{m})$ and the primary particle diameter must be known accurately in order to make use of this method.

Iyer et al. (2007) looked at combining elastic light scattering with extinction measurements to determine the aggregate size distribution without knowing the primary particle diameter or the aggregate morphology. The paper presents a good overview of RDG-FA theory for soot aggregates as well as the common method of assuming a distribution type to fit the light scattering data in order to determine the aggregate size distribution.

2.4 Summary

Soot aggregates have been of interest for many years, both because their size and presence are an indication of the efficiency of a flame and because they can adversely affect human health and the environment. The study of soot has been progressing for more than half a century. However, the advent of the fractal theory in 1977 for soot aggregates has spurred a great deal of attention to the problem of characterizing the soot aggregate size distribution. Similarly, light scattering from aerosols has been investigated since the late nineteenth century, but it was only once the fractal nature of soot aggregates was known that its rapid advance began as a diagnostic for combustion applications. Finally, inversion as applied to light scattering from fractal aggregates also has a long history, yet it also advanced rapidly once soot was recognized to be a fractal aggregate. All of these pieces had to be in place before the inversion of light scattering from fractal aggregates could be attempted in a mathematically rigorous way.

Chapter 3
Derivation of Light Scattering Equations

3.1 Introduction

In order to use the light scattering data in recovering the aggregate size distribution, there must be a model that relates the scattered data to the size distribution of soot aggregates that generated it. This chapter will begin by deriving such a relationship between the angular distribution of scattered light and the aggregate size distribution, noting the assumptions made, and will then consider a number of structure factors that have been reported in literature. Following that, the uncertainty in some of the system variables will be recognized. Finally the linearity of the problem will be discussed.

3.2 Derivation of the Light Scattering Equations

The experimental optics setup that was used to measure elastic light scattering are shown in Figure 3.1. The derivation of the governing equation for light scattering from a soot aerosol begins with the Radiative Transfer Equation (RTE) (Siegel and Howell 2002),

$$i_{\lambda_l}\left(\kappa_{\lambda_l},\omega\right) = i_{\lambda_l}(0,\omega)e^{-\kappa_{\lambda_l}} + \int_0^{\kappa_{\lambda_l}} I_{\lambda_l}\left(\kappa^*_{\lambda_l},\omega\right)e^{-\left(\kappa_{\lambda_l}-\kappa^*_{\lambda_l}\right)}d\kappa^*_{\lambda_l} \tag{3.1}$$

where the optical depth, or the opacity of the medium, is given by

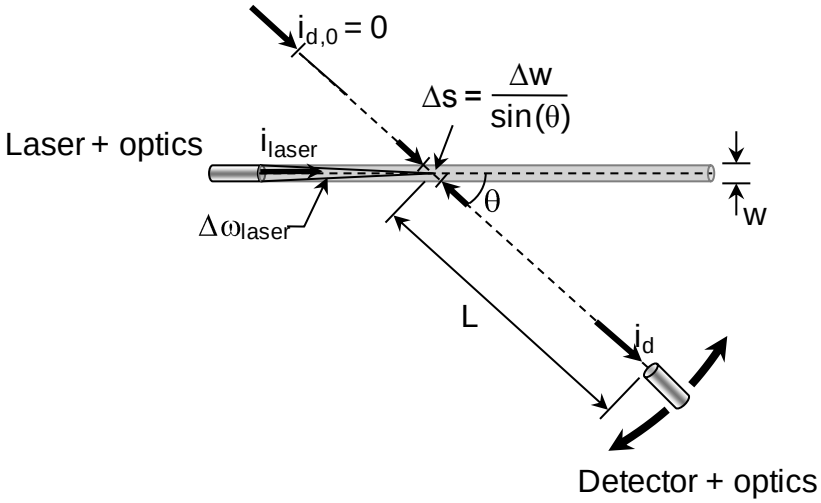

Figure 3.1 – Experimental optics for multiangle elastic light scattering

$$\kappa(s) = \int_0^S K(S^*)dS^* = \int_0^S \left[a_{\lambda_l}(S^*) + \sigma_s(S^*)\right] dS^* \tag{3.2}$$

and the source function is defined by

$$I(\kappa, \omega_d) \equiv (1 - \Omega) i_{\mathrm{b}}(\kappa) + \frac{\Omega}{4\pi} \int_0^{4\pi} i(\kappa, \omega_i)\, \phi(\omega_{\mathrm{d}}, \omega_{\mathrm{i}}) \mathrm{d}\omega_{\mathrm{i}} \tag{3.3}$$

where K is the extinction coefficient, a_{λ_l} is the absorption coefficient, σ_s is the scattering coefficient, κ is the optical thickness, i_{λ_l} is the radiation intensity, i_{b} is the blackbody intensity, $i_{\lambda_l}(0, \omega)$ is the background intensity, I_{λ_l} is the source function, ω is a finite solid angle, λ_l is the wavelength of the radiation, ϕ is the scattering phase function and Ω is the scattering albedo.

Equation (3.1) gives the intensity of light measured in a specified direction from all sources, including scattering, absorption, emission, and background blackbody radiation. The source function in Eq. (3.3) is the source of the intensity scattered or emitted along the optical path under consideration.

To simplify the governing equations, the absorption and emission of light between the probe volume and detector is assumed to be negligible relative to the intensity of the scattered light. Emission can be neglected since the wavelength of laser light used in the measurement is shorter than the characteristic emission wavelengths of soot in the flame. Neglecting the absorption is reasonable because the flame is optically thin, as demonstrated by Snelling et al. (2005). That the flame is optically thin can be shown by properly calibrating the laser to a known scatterer, such as nitrogen gas, and then comparing the known measurement to that of the scattering from a soot aerosol.

The third assumption is that the laser light undergoes only single scattering between the laser and the detector. The laser light is vertically polarized on the laser side, and single scattering dictates that it should still be vertically polarized on the detector side. This was verified by looking at the vertical-horizontal scattering measured by the detector, which was negligible in comparison to the vertical-vertical scattering. This indicated that the scattering from the soot aggregates was only single scattering, since multiple scattering would produce a strong horizontally polarized component.

The fourth assumption made is that the value of the background intensity that reaches the detector is zero. In the experiment, this was ensured by subtracting the value of the background intensity from each measurement. The frequency of the detector measurements was 10kH and the frequency of the laser was 5kH, such that every other measurement was strictly the background radiation. The measurements of the laser intensity could then be reduced by the background measurements. The background intensity was far less than one percent of the laser laser intensity, however, so even if the background intensity is not removed from the measurements the term $i_{\lambda_l}(0, \omega)$ in the RTE can be taken to be zero.

The fifth assumption is that the laser beam obeys a top-hat profile (constant intensity across the beam width). In reality the beam obeys a Gaussian profile. However, experimental work by researchers at the National Research Council Canada suggested that the difference between the two profiles was negligible. This hasn't been proven conclusively, but is assumed true throughout this work.

Since emission is negligible, the $i_{\mathrm{b}}(\kappa)$ term is zero, and the source function simplifies to

$$I(\kappa, \omega_d) \equiv \frac{\Omega}{4\pi} \int_0^{4\pi} i(\kappa, \omega_i)\, \phi(\omega_{\mathrm{d}}, \omega_{\mathrm{i}}) \mathrm{d}\omega_{\mathrm{i}} \tag{3.4}$$

where Ω is a function of the wavelength of light and the properties of the soot itself. The derivation and value of the scattering albedo is considered in more detail in Appendix B.3.

Adding this to the RTE, along with the other assumptions mentioned gives

$$i_{\lambda_l}(\kappa_{\lambda_l}, \omega) = \int_0^{\kappa_{\lambda_l}} \left[\frac{\Omega}{4\pi} \int_0^{4\pi} i(\kappa, \omega_i)\, \phi(\omega_{\mathrm{d}}, \omega_{\mathrm{i}}) \mathrm{d}\omega_{\mathrm{i}} \right] e^{-\left(\kappa_{\lambda_l} - \kappa^*_{\lambda_l}\right)} d\kappa^*_{\lambda} \tag{3.5}$$

If the substitution from Eq. (3.2) is applied to the system, and letting $d\kappa(s) = \sigma_s(s)ds$, Eq. (3.5) becomes

$$i_{\lambda_l}(\kappa_{\lambda_l}, \omega) = \int_0^{L} \left[\frac{\Omega}{4\pi} \int_0^{4\pi} i(s, \omega_i)\, \phi(\omega_{\mathrm{d}}, \omega_{\mathrm{i}}) \mathrm{d}\omega_{\mathrm{i}} \right] \sigma_{\mathrm{s}}(\mathrm{s}) \exp\left[- \int_{s^*}^{L} \sigma_s(s^*) ds^* \right] ds \tag{3.6}$$

where L is the length from the laser beam to the detector. Now, because of the assumption of single scattering, the laser light scattered in the direction of the detector can only come from within the path of the beam itself, so the scattering coefficient is equal to zero everywhere else. Thus, the exponential term in Eq. (3.6) becomes unity, indicating that there is no loss of signal between the scattering event and the light reaching the detector. As well, the inscattering integral in Eq. (3.6) is equal to zero in every direction except the direction of the laser beam. This gives

$$i_{\lambda_l}(\kappa_{\lambda_l}, \omega) = \left[\frac{\Omega}{4\pi}\int_0^{4\pi} i(s, \omega_i)\, \phi(\omega_\mathrm{d}, \omega_\mathrm{i})\mathrm{d}\omega_\mathrm{i}\right]\sigma_\mathrm{s}\Delta\mathrm{s} \approx \frac{\Omega}{4\pi} i_\mathrm{laser}\phi(\omega_\mathrm{d}, \omega_\mathrm{laser})\Delta\omega_\mathrm{laser}\sigma_\mathrm{s}\Delta s \tag{3.7}$$

In this equation, $\Delta\omega_{\mathrm{laser}}$ is the small, but finite, solid angle that is subtended by the laser beam. If the laser beam is assumed to have a uniform intensity over the entire width, known as a top-hat profile, then the length of the path of intersection between the detector view and the laser beam is given by $\Delta s = \Delta w/\sin(\theta)$, where θ is the angle formed between the laser beam and the detector. Furthermore, the scattering phase function does not depend on the laser angle and detection angle, but rather the angle formed by the intersection between the laser path and the detection path, θ. Thus, Equation (3.7) becomes

$$i_{\lambda_l}(\kappa_{\lambda_l}, \omega) \approx \frac{\Omega\, i_\mathrm{laser}\phi(\omega_\mathrm{d}, \omega_\mathrm{laser})\Delta\omega_\mathrm{laser}\sigma_\mathrm{s}}{4\pi}\frac{\Delta w}{\sin(\theta)} = C_1\frac{\phi(\theta)\sigma_\mathrm{s}}{\sin(\theta)} \tag{3.8}$$

where C_1 is a coefficient containing all terms that are independent of both θ and N_p,

$$C_1 = C_{exp}\mathrm{i}_\mathrm{laser}\Delta\omega_\mathrm{laser}\Delta w\,\frac{\Omega}{4\pi} \tag{3.9}$$

where C_{exp} is a scaling parameter containing all scaling parameters known from the experimental apparatus.

Equation (3.8) gives the basic relationship between the observed scattered light and the system that scattered it. However, more simplifications can be made by implementing an expression for each of the functions. For vertically polarized light, the scattering phase function is given by

$$\phi(\theta) = 4\pi \frac{\bar{C}_{vv}^{agg}}{\bar{C}_{s}^{agg}} \tag{3.10}$$

where $\overline{C}_{s}^{agg}$ and $\overline{C}_{vv}^{agg}$ are the average scattering and vertical polarization cross sections of the soot aggregates, respectively (Eymet, et al. 2002). As well, the scattering coefficient is given by

$$\sigma_s = N_{agg} \bar{C}_{s}^{agg} \tag{3.11}$$

where N_{agg} is the aggregate number density. If these two terms are multiplied together, it gives

$$\phi(\theta)\sigma_s = 4\pi \frac{\bar{C}_{vv}^{agg}}{\bar{C}_{s}^{agg}} N_{agg} \bar{C}_{s}^{agg} = 4\pi \bar{C}_{vv}^{agg} N_{agg} = 4\pi N_{agg} \int_1^\infty C_{vv}^{agg} P(N_p) dN_p \tag{3.12}$$

where $P(N_p)$ is the aggregate size distribution in the soot aerosol, and N_p is the number of primary particles in an aggregate. The vertical polarization cross section is given by

$$C_{vv}^{agg} = N_p^2 C_{vv}^{p} f\left[q(\theta) R_g(N_p)\right] \tag{3.13}$$

where $C_{vv}^{p} \equiv x_p^6 F(\boldsymbol{m})/\eta^2$ is the polarization cross section of a primary particle, $q(\theta) \equiv 2\eta \sin(\theta/2)$ is the modulus of the scattering wave vector, $F(\boldsymbol{m}) = |(\boldsymbol{m}^2 - 1)/(\boldsymbol{m}^2 + 1)|^2$ is the complex scattering function, $\boldsymbol{m}$ is the complex index of refraction, $\eta \equiv 2\pi/\lambda_l$ is the wavenumber, $x_p \equiv \pi d_p/\lambda_l$ is the size parameter, d_p is the primary particle diameter, R_g is the aggregate radius of gyration, and $f(\cdot)$ is the structure factor.

Substituting all of these terms into Eq. (3.8) gives

$$i_{\lambda_l}(\kappa_{\lambda_l}, \omega) = C_1 \frac{1}{\sin(\theta)} 4\pi N_{agg} \int_1^\infty N_p^2 C_{vv}^{p} f\left[q(\theta) R_g(N_p)\right] P(N_p) dN_p \tag{3.14}$$

All of the terms containing either θ or N_p, except for the aggregate size distribution (the determination of which is the overarching goal of this thesis), are gathered together into one term, called the kernel function,

$$K\left(\theta, N_p\right) = \frac{N_p^2 f\left[q(\theta) R_g\left(N_p\right)\right]}{\sin(\theta)} \tag{3.15}$$

and all the other terms are collected into a coefficient,

$$\begin{aligned} C &= 4\pi C_1 N_{agg} \left(\pi d_p/\lambda_l\right)^6 F(\boldsymbol{m})/(2\pi/\lambda_l)^2 \\ &= C_{exp} \mathrm{i}_{\mathrm{laser}} \Delta\omega_{\mathrm{laser}} \Delta \mathrm{w}\, N_{agg} \pi^4 d_p^6 F(\boldsymbol{m}) \Omega/(4\lambda_l^4) \end{aligned} \tag{3.16}$$

resulting in the following Fredholm integral equation of the first kind, Eq. (1.2),

$$g(\theta) = C \int_1^{\infty} K\left(\theta, N_p\right) P\left(N_p\right) dN_p$$

where $i_{\lambda_l}\left(\kappa_{\lambda_l}, \omega\right)$ is written as $g(\theta)$.

3.3 Structure Factors

The structure factor contains information about the form of the aggregate that scattered the light. In literature, a number of structure factors have been proposed. For instance, Sorensen (2001) lists seven distinct structure factors that could be considered, but suggests that the hypergeometric structure factor is the best choice. The hypergeometric function is defined as

$${}_1F_1[a; b; c] = 1 + \sum_{i=1}^{\infty} \frac{\Gamma(a+i)}{\Gamma(a)} \frac{\Gamma(b)}{\Gamma(b+i)} \frac{c^i}{i!} \tag{3.17}$$

and a Matlab implementation of it can be found in Appendix A.6. Γ is the Gamma function, the implementation of which can be found in a sub-function in the implementation of the hypergeometric function. In the case of the structure factor for the soot aerosol system, $a = D_f/2$, $b = 3/2$, and $c = -\left(qR_g\right)^2/D_f$.

One consideration in using the hypergeometric function is the use of the cutoff function, as noted by Sorensen (2001). At a critical value of qR_g, which is the product of the scattering

wave vector and the radius of gyration of an aggregate, the hypergeometric function is replaced by the cutoff function

$$f\left[q(\theta)R_g(N_p)\right] = C\left(qR_g\right)^{-D_f} \tag{3.18}$$

where the value of C is dependent on the cutoff function used. In this work, the Gaussian cutoff function suggested by Sorensen was used, such that $C = {D_f}^{D_f/2}\left[\Gamma(3/2)/\Gamma\left(\left(3-D_f\right)/2\right)\right]$. The value of qR_g at which this cutoff is implemented is very much up to debate. Sorensen suggests that the cutoff should be applied for $qR_g > 1$, whereas Eymet et al. (2002) suggest that the cutoff should be applied at $qR_g > 1.5D_f$.

Another problem with the hypergeometric function is that there appears to be no simple implementation, and to determine accurate results the sum must be taken to as high a value as possible. This is complicated due to the transcendental Gamma function, which is written as an infinite product that must be approximated to a large number of terms to achieve a reasonable degree of accuracy. These two facts combine to produce a function that requires a great deal of computational expense to calculate the structure factor. A simpler form of the structure factor that is computationally inexpensive yet sufficiently accurate is needed.

Such a structure factor was proposed by Yang and Köylü (2004),

$$f\left[q(\theta)R_g(N_p)\right] = \left[1 + \frac{8\left(qR_g\right)^2}{3D_f} + \left(qR_g\right)^8\right]^{-D_f/8} \tag{3.19}$$

which is a heuristic fit to the power-law and Guinier scattering regimes, spanning the transition regime between the other two regimes.

A comparison of the hypergeometric and Köylü structure factors was carried out. Each structure factor is a variable of qR_g, so a range of values was chosen to represent the entirety of the problem. The modulus of the scattering wave vector is defined as $q \equiv 4\pi\lambda^{-1}\sin(\theta/2)$ and, because soot aggregates are fractals, the fractal scaling law applies, given by Eq. (1.1). From TEM experiments that have considered this specific flame, aggregates that contain more than 500 primary particles are rare, so that number is taken as the upper bound. Experimentally, the maximum angle that can be considered is $\theta = 160^{\circ}$. As well, because $k_g = 2.4$, $D_f = 1.72$, $d_p = 29$ nm and $\lambda_l = 527$ nm are the values applicable to this flame, an upper limit for qR_g of roughly 7.5 was calculated.

For a range of qR_g, the two structure factors were calculated with the cutoff taken at $qR_g > 7.5$. The results are plotted in Figure 3.2 (a). The two agree quite well, save for a small spike around $qR_g = 7$. This spike is caused by numerical error creeping into the hypergeometric function, as the simulations were not capable of computing the sum to enough accuracy. This is the case because the factorial term and the term c^i both reach the upper limit at which a computer is capable of computing them. If an additional term is used to approximate the Gamma function, one of these terms goes to infinity, whereas in reality there is a finite value for each of them that cannot be calculated. The greatest discrepancy between the hypergeometric and Köylü structure factor occurs within the transition regime, which is notoriously difficult to properly assess (Sorensen, Lu and Cai 1995). However, if the Eymet cutoff value is used the result is closer, although there is still a jump when the hypergeometric function transitions to the cutoff function. This is shown in Figure 3.2 (b).

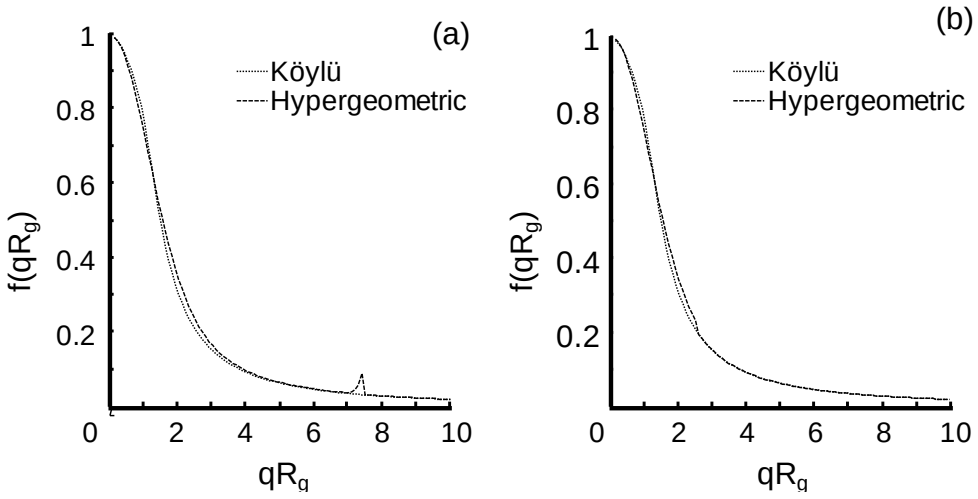

Figure 3.2 – Comparison of Structure Factors: (a) Cutoff at $qR_g = 7.5$ and (b) Eymet Cutoff

We conclude that the Köylü structure factor is sufficiently close to the more accurate, yet computationally expensive hypergeometric structure factor for the purpose of this analysis. The main variation comes in the transition range, where both functions simply span the distance between the Guinier and power-law regimes, because characterization of the transition regime is especially challenging.

3.4 Sources of Uncertainty

There are a number of sources of uncertainty in the above equations, some of which can be mitigated and some of which cannot. The first source of uncertainty is the primary particle diameter. When RDG-FA theory is applied, it is almost always assumed that all soot aggregates in an aerosol contain uniformly-sized particles. This is not actually the case. Brasil et al. (1999) note that the primary particle diameter in an aerosol can vary by up to 25%, although in a single aggregate the size tends to be constant. Reimann et al. (2009) considered the primary particle size distribution by looking at soot aggregates using transmission electron microscopy (TEM).

They fit a lognormal distribution to the resulting histogram. This was ignored in this treatment of the problem, as it greatly increases the complexity of the calculations.

One of the greatest areas of uncertainty in dealing with soot aggregates is the complex index of refraction, $\boldsymbol{m}$. As a complex number, the complex index of refraction is made up of two parts: the real part is the speed of light in the material relative to the speed of light in a vacuum, and the complex part is related to the light that is lost due to absorption. This variable contains a great deal of uncertainty, both in the value it should take and what factors affect it. Sorensen (2001) noted almost a dozen different values reported in the literature, and there are many more. He pointed out that the uncertainty in determining the value of the refractive index may be greater than the actual value of the refractive index itself. The refractive index is of particular concern if the experiment is dealing with absolute scattering, whereas it is less of a concern in the case of relative scattering, which was used throughout this thesis. The complex index of refraction is still found in the value of the constant C in the case of this derivation, so that it is a necessary value.

Another source of uncertainty is the fractal dimension, D_f. Recent work by Link et al. (2010) has shown that the fractal dimension is a strong function of the distribution width by means of inversion of light scattering measurements through an assumed distribution. A plot of the distribution width versus the fractal dimension reveals a long, wide valley of near-identical residuals of the fit solutions for lognormal distributions. Even if the distribution width is set, there are a wide range of fractal dimensions that are capable of reconstructing the problem.

3.5 Nonlinearity of the Governing Equation

Up to this point the coefficient in Eq. (1.2) has been taken to be a constant. However, contained in this coefficient, C, is the aggregate number density, N_{agg}, which is defined as

$$N_{agg} = \frac{f_v}{\pi d_p^2 \overline{N}_p} = \frac{f_v}{\pi d_p^2 \int_1^\infty N_p P(N_p) dN_p} \tag{3.20}$$

where f_v is the soot volume fraction. In this relation the aggregate number density is a function of the soot volume fraction, which can be obtained using line-of-sight-attenuation (LOSA) measurements, and the aggregate size distribution. Thus, when the soot volume fraction is used to determine the aggregate number density, the scaling coefficient, C, is actually a function of N_p, and thus Eq. (1.2) is nonlinear.

N_{agg} can be determined in another fashion. For instance, a condensation particle counter (CPC) can be used to accurately determine the number aggregate density of an aerosol (Agarwal and Sem 1980), but it cannot be done at the same time as a light scattering measurement. Hence, two separate experiments must be carried out to prevent this experiment from becoming nonlinear. Another method, which has been used in this thesis, is to treat the coefficient C as an additional unknown in the multiangle angular scattering experiment. This is done iteratively, as part of the solver working to find the aggregate size distribution. In this manner C is treated as just another unknown value and solved along with everything else, which greatly decreases the complexity of the problem by making Eq. (1.2) linear in N_p.

Chapter 4
Experiment

4.1 Introduction

This chapter will describe the experimental setup used to obtain the light scattering data used throughout this thesis. As with any experiment, there were many things to take into account and optimize to properly carry out the experiment, and some of these will be discussed. First, the setup of the experiment will be presented. Second, the details of burner will be considered, including position with respect to the optics. Third, the spatial resolution of the detector will be considered. Fourth, other sources of error will be noted. This experimental setup was built by Oliver Link at the Institute for Combustion Processes and Environmental Technology, National Research Council Canada, and used by the author between May and June, 2010, to collect data for use in this thesis. The specific angle and measurement set used throughout this work can be seen in Appendix B.6.

4.2 Experimental Setup

The basic experimental setup used to measure scattered intensity from soot aggregates is described by Link et al. (2010). The laser source was a diode-pumped, Q-switched Nd-YLF laser that operated at a wavelength of 527 nm. The average power of the laser was 400mW at 5 kHz. The laser beam was expanded and collimated through two lenses to a diameter of 3mm before traveling through a half-wave plate that rotated the linear polarization, reducing the energy transmitted. The beam then traveled through a polarizer in order to control its energy and orientation, before a final wave plate ensured that the polarization was vertical. The beam then

traveled through a final lens that focused it down to between 150µm and 200µm. Before reaching the aerosol, the beam passed through shielding to block any light scattered by the optical components. This setup is shown in Figure 4.1.

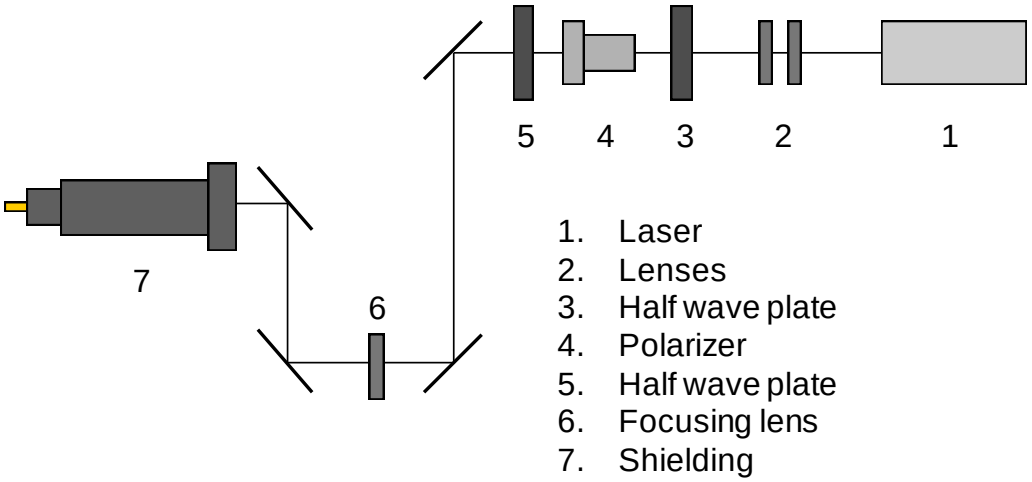

Figure 4.1 – Experimental setup: laser

The detector optical components were mounted on a planar rotation stage capable of moving the detector to various angles around the flame. The forward direction, the direction of propagation of the laser beam, was denoted as 0°. The rotation stage was limited to movement approximately between 10° and 160° due to the physical constraints of the setup. The laser was aligned such that its beam passed through the axis of rotation of the rotation stage. The burner was positioned so that the feature of interest was located at the centre of rotation of the rotation stage. The overall setup of the detector optics is shown below in Figure 4.2.

The scattered laser intensity seen by the detector was reduced by placing one or more filters in series in the path of the light prior to the detector optics. After passing through the filters the light travelled through a collection lens that collimated the light. The collection lens had a focal length of 20 cm, and was situated 20 cm from the center of rotation of the rotation

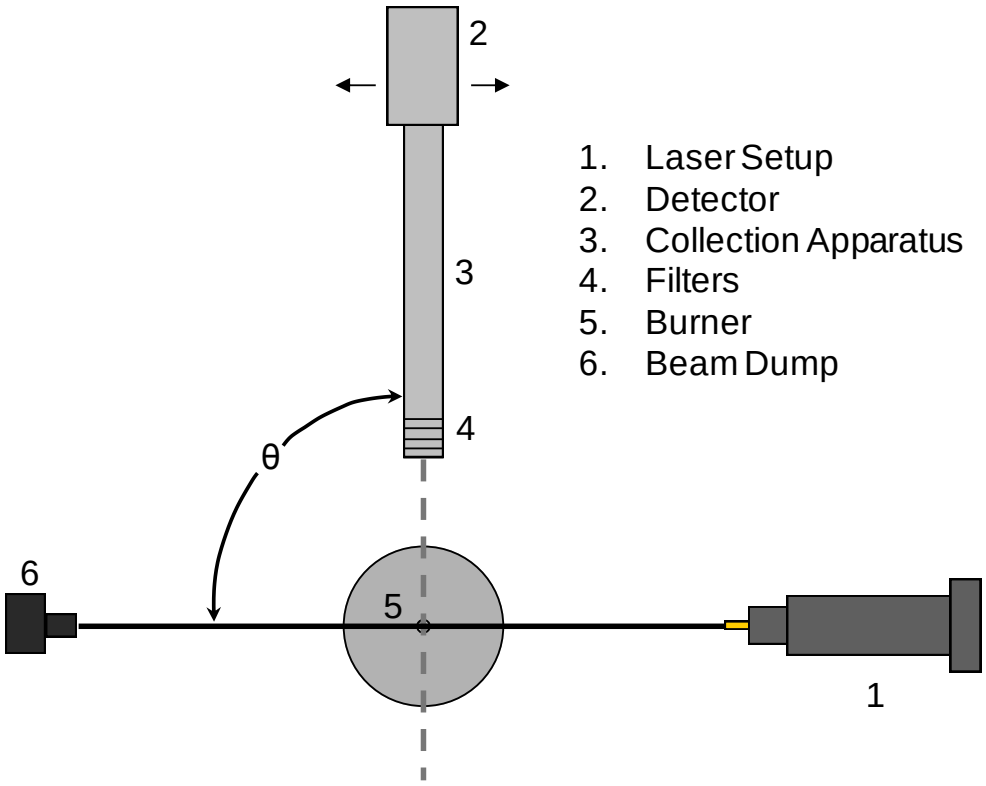

Figure 4.2 – Experimental setup: detector optics

stage on a horizontal plane. An aperture decreased the diameter of the lens to 17mm. The collimated light passed through a polarizer that ensured the light reaching the detector was strictly vertically polarized. Following the polarizer the light travelled through a bandpass filter set to 527nm. The light then passed through a second lens, also with a focal length of 20 cm, which imaged the light onto a final aperture, selected as a 50 μm slit, and then on the photomultiplier. This setup is shown in Figure 4.3.

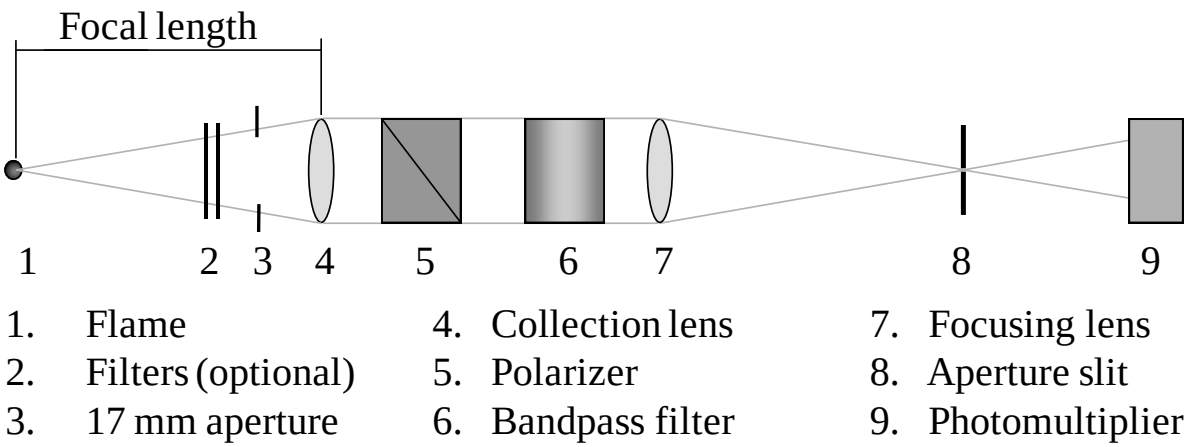

Figure 4.3 – Experimental setup: collection apparatus

4.3 The Burner

The soot source that was used in this experiment was the Gülder laminar co-annular diffusion flame (Snelling, Thomson, et al. 1999). The burner consisted of coaxial tubes, one carrying fuel and the other carrying air. The inner tube, carrying ethylene at a flow rate of 326 sccm (standard cubic centimetre per minute), was made of steel with an inner diameter of 10.8 mm and an outer diameter of 12.7 mm. The outer tube, carrying air at a flow rate of 284 slpm (standard litre per minute), was made of aluminum with an inner diameter of 88.4 mm.

The burner could be moved in three degrees of freedom through the use of multiple translation stages. Each stage corresponded to one of the principle axes of the setup: the first stage was aligned parallel to the direction of propagation of the laser, the second stage was aligned to be parallel to the collection optics when set at 90˚ to the laser propagation direction, and the third stage allowed vertical movement of the burner, allowing different measurement heights within the flame. The movement of each stage was controlled by a micrometer screw, allowing for fine adjustments of the burner position. The stages were attached to a basic rack that

was levelled to ensure that the flame didn't vary in the horizontal plane due to any misalignment of the burner itself.

4.4 Positioning the Burner

It is important, when considering spatially resolved scattering measurements, to properly position the portion of the flame that is of interest. Having a refined spatial resolution is of no use if it is pointing at the wrong position in the flame. First, the center of rotation of the rotation stage must be found. This was done by centering a pin above the burner and viewing the movement of the tip of the pin with respect to the laser beam as the stage rotated through a full sweep. The pin was allowed to move in two axes on the plane defined by the laser beam propagation direction and the collection optics, and was adjusted until the tip did not move as the rotation stage turned. A schematic of the pin setup is seen in Figure 4.4 (a).

Once the pin was set in place, a second laser was added to shine through the collection optics in the reverse direction (in the direction from the photomultiplier to the burner) and the optics adjusted until the laser intersected the pin head. The vertical height of the collection optics was adjusted to be of the same height as the laser beam.Because the slit aperture was used rather than a circular aperture, however, the exact vertical alignment of the collection optics and the laser beam was not as essential as in the case when a circular aperture was used. The alignment setup is seen in Figure 4.4 (b).

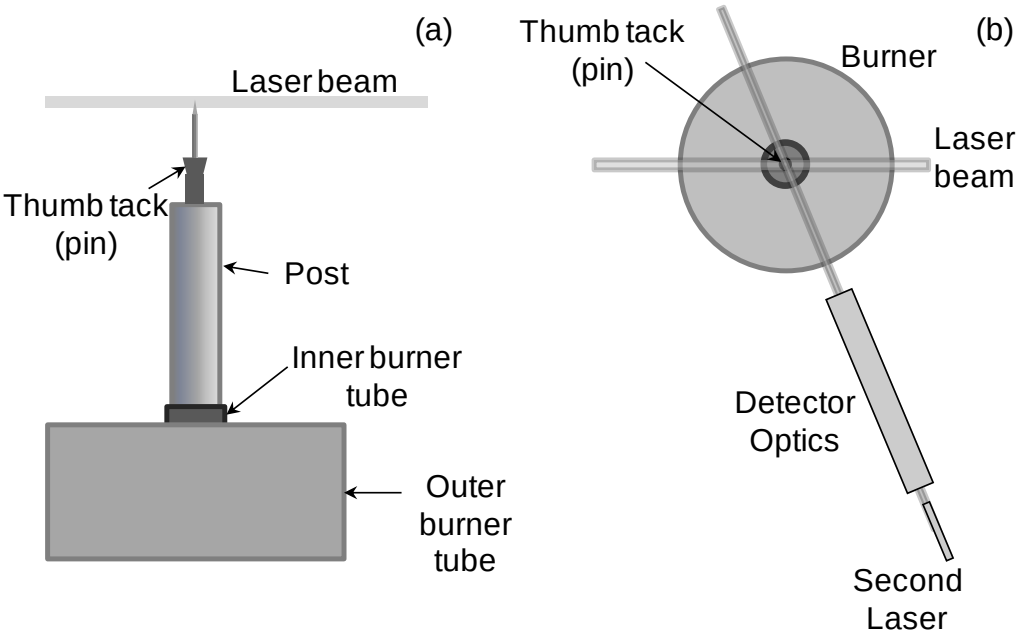

Figure 4.4 – Positioning the laser: (a) the pin setup, and (b) aligning detector optics

4.5 Spatial Resolution

One of the key requirements of laser diagnostics on soot aerosol systems is having a refined spatial resolution. The flame produced by the Gülder burner has strong gradients in soot concentration and, as a consequence, in the laser intensity measured by the detector. A horizontal scan through the flame at a height above the burner of 42 mm is seen in Figure 4.6. Equation (1.2) was derived assuming that only single scattering takes place. That is, light scattered from an aggregate travels directly to the detector without encountering any other aggregates. Therefore, any light that reaches the detector must have been scattered from the path of the laser beam. The intersection of the detector viewing angle and the laser beam is called the sample volume, as shown in Figure 4.5.

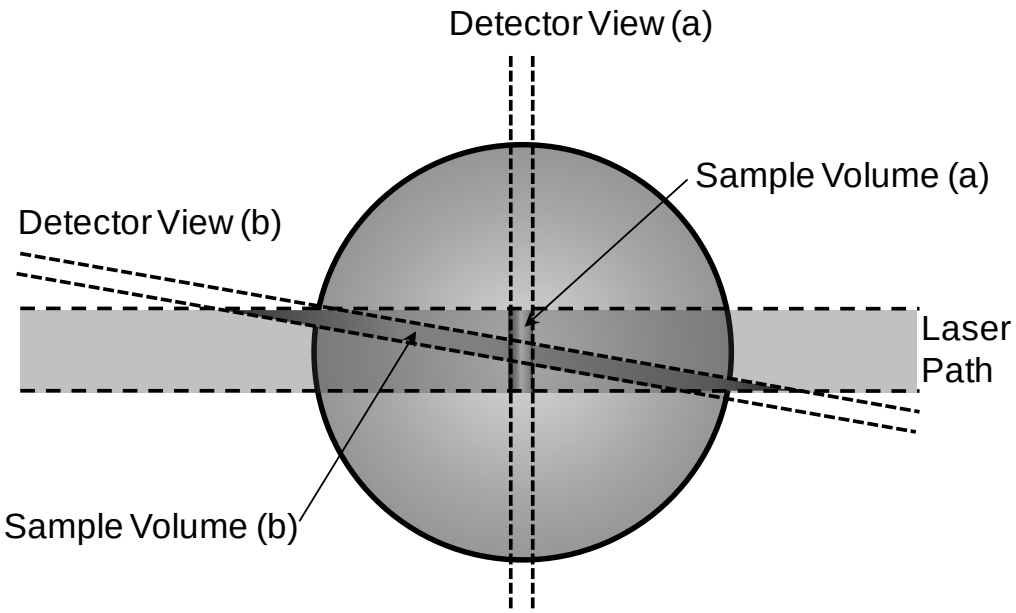

Figure 4.5 – Sample Volume at (a) 90˚ and (b) 10˚

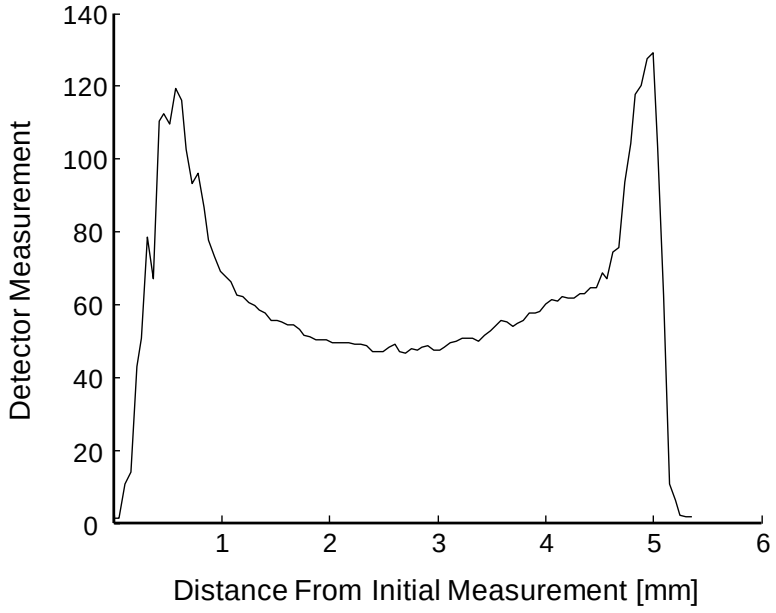

Figure 4.6 – Horizontal scan through the Gulder burner at 42mm.

One of the assumptions made in using light scattering data is that the flame within the sample volume was invariant with respect to intensity. If the angle the collection apparatus makes with the laser propagation direction is denoted as θ, the sample volume length in the

direction of the laser beam propagation is defined as the beam width multiplied by the cosine of θ and added together with the width of the aperture in the collection optics before the photomultiplier, which is all divided by sine of θ. As is seen in Figure 4.5, the length of the sample volume in the laser propagation direction grows larger the closer the angle becomes gets to 0˚. It is important, then, that the length of the sample volume not exceed the length of the roughly flat section in the centre of the flame, which is defined as not varying by more than 10%.

It was found that using a 50μm slit on the detector in conjunction with a 200μm laser beam width produces a sample volume length that does not exceed 1.5 mm at 10°. The flat portion of the flame at 42mm HAB is found to be roughly 1.8mm, such that the sample volume will actually fit within it. It is therefore acceptable to use angles as small at 10° if need be. It should be noted that the intensity of the flame is not exactly constant across this central region, which is a potential source of error in the system.

Light scattering measurements are taken at different angles around the burner, but no one has really provided an in-depth analysis of how to choose the angles. One suggestion was made by Sorensen (2001) to choose the angles uniformly in the scattering wave vector (q) domain. Following that advice, twenty-one angles uniformly spaced in the q-domain were chosen and angular measurements taken. For completeness, $\theta = 10°$ and $\theta = 160°$ were included to make use of the entire angular range, increasing the total number of angles at which light scattering data was measured to twenty-three. This choice of angle set is reconsidered in Chapter 6.

4.6 Error Sources

For the purpose of repeatability it was assumed that throughout the experiment the laser source was always operating at its specified wavelength and output power. It is possible that the

power of the laser oscillated or varied over the course of a number of experiments, but no data was collected to determine if this was the case. One could determine this by measuring the laser intensity over a long period of time (some hours at least) to determine if there is a change. The laser was always allowed a few minutes to 'warm up' before actual measurements were taken, to ensure that if it took some time to reach full power the results would not be skewed.

Another concern was flame stability. As noted by Kevin Thomson in his PhD thesis - (Thomson 2000), some form of shroud should be placed around the flame to ensure that air movement in the room does not disturb it. This is crucial, as even the opening of a door across the room can cause the flame to waver by up to a centimetre. Because of the significant increase in intensity in the annular region of the flame, seen in Figure 4.6, that amount of movement could easily cause the annulus to move into the sample volume, strongly skewing the measurements. Due to this instability measurements were generally taken at times when no one else would be in the lab, although portable wall divisions were eventually moved around the experiment to ensure that extraneous air currents didn't affect the experiment.

Chapter 5
Inversion of Light Scattering Data

5.1 Introduction

The angular distribution of light scattered from a soot aerosol contains information about the system that can be employed to determine the aggregate size distribution within the aerosol. However, these two quantities are related by a Fredholm integral equation of the first kind, which is inherently ill-posed. Thus, it is necessary to add additional information to the problem, through regularization, to recover the aggregate size distribution. This chapter will first demonstrate the ill-posedness of this problem. Then, it will introduce Tikhonov regularization, one of the most popular regularization techniques, which will then be demonstrated by recovering the aggregate size distribution from an artificially generated data set. Finally, the chapter concludes with a more sophisticated and generalized technique called Maximum a Posteriori (MAP) inference, which will be applied to recover the soot aggregate size distribution from the experimental data set.

5.2 Ill-posed Problem

Light scattering from a soot aerosol is related to the aggregate size distribution by a Fredholm integral equation of the first kind, Eq. (1.2),

$$g(\theta) = C \int_1^{\infty} K(\theta, N_p)\, P(N_p) dN_p \tag{1.2}$$

the deconvolution of which is ill-posed. Hadamard (1923) gave the criterion for a well-posed problem: the solution must exist, the solution must be unique, and the solution must be stable, i.e. adding small amounts of noise to the measured data should produce equally small variations in the recovered solution. At the time, Hadamard considered problems that violated these conditions to be nothing more than mathematical curiosities, something that would not be found in nature. It has since been discovered that such problems, called ill-posed problems, are found in most fields of science and are ubiquitous. The discussion in Section 1.3.2 on the smoothing effect of the kernel posits that the inversion of light scattering data from a soot aerosol is an ill-posed problem. To deal with the inherent ill-conditioning of the matrix equation, more information must be added to the problem through a process called regularization.

5.2.1 Presumed Distribution Type

The most prevalent way of regularizing this particular problem in the literature is to assume that the aggregate size distribution is a specific distribution type, such as a lognormal distribution, defined as

$$\mathrm{P_{lognormal}}(\mathrm{N_p}) = \frac{1}{\mathrm{N_p}\,\ln(\sigma_g)\sqrt{2\pi}} \exp\left[-\frac{\left(\ln(\mathrm{N_p}) - \ln(\mathrm{N_{p,g}})\right)^2}{2\ln(\sigma_g)^2}\right] \tag{5.1}$$

where σ_g is the width of the distribution and $N_{p,g}$ is the geometric mean. Assuming such a distribution greatly reduces the ill-posedness of the problem, by limiting the degrees-of-freedom to two. The objective, then, is to identify the set $(\sigma_g, N_{p,g})$ that minimize the residual of Eq. (1.3). Figure 5.1 (a) plots the residual function, $\left\|\mathbf{A}\mathbf{x}_{\text{lognormal}} - \mathbf{g}\right\|_2^2$, over the domain of these two variables. Three points have been chosen in the long narrow valley, which is made up of solutions that solve the problem within experimental accuracy. Figure 5.1 (b) shows the lognormal distributions corresponding to these three points, which vary greatly in shape, and yet all accurately reconstruct the light scattering data. This shows that the inverse analysis of elastic light scattering by soot aggregates violates Hadamard's second criterion, i.e. the solution is not unique, even with an assumed aggregate size distribution type.

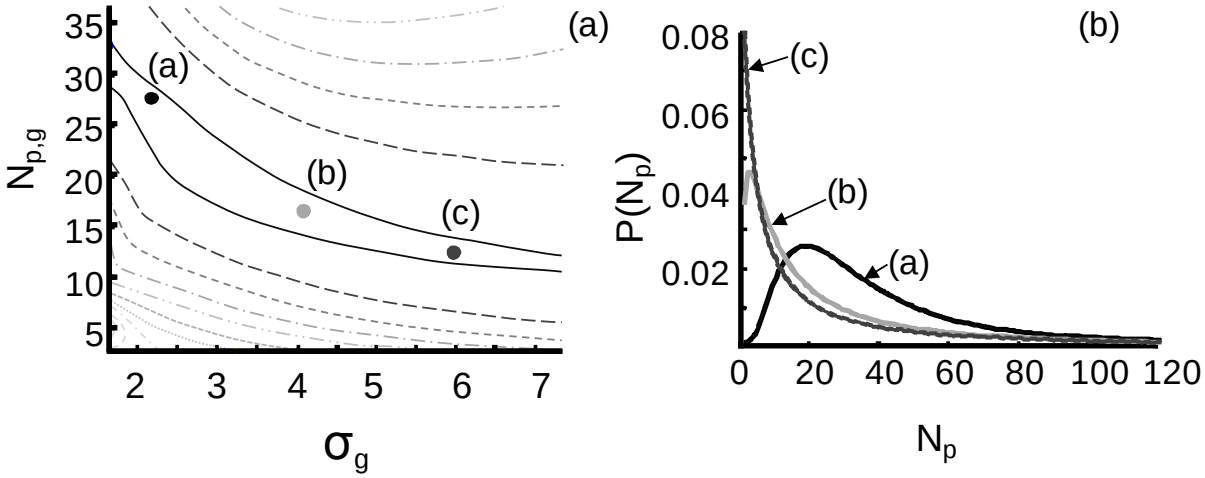

Figure 5.1 – Multiple lognormal distributions that all recover the Elastic Light Scattering (ELS) data accurately.

5.2.2 Discrete Ill-Conditioned Matrix Equation

To quantify the ill-posedness of this problem, Eq. (1.2) is first converted into a matrix equation, Eq. (1.3),

$$\mathbf{Ax} = \mathbf{g} \tag{1.3}$$

where **A** is generated by

$$A_{i,j} = \frac{C}{\Delta N_p} \int_{N_{p,j}-\Delta N_p}^{N_{p,j}+\Delta N_p} K\left(\theta_i, N_p^*\right) \mathrm{d}N_p^* \tag{5.2}$$

A perturbation analysis is now carried out on the matrix equation. First, define the perturbation in the measured data and the recovered solution as $\delta\mathbf{g}$ and $\delta\mathbf{x}$, respectively. The matrix equation is perturbed by these errors,

$$\mathbf{A}(\mathbf{x} + \delta\mathbf{x}) = \mathbf{g} + \delta\mathbf{g} \tag{5.3}$$

The left hand side is expanded and the equation rearranged to give

$$\mathbf{A}\delta\mathbf{x} = (\mathbf{g} - \mathbf{Ax}) + \delta\mathbf{g} \tag{5.4}$$

The original matrix equation, Eq.(1.3), is taken to hold true, so that the bracketed term in Eq. (5.4) is zero. The remaining terms can be rearranged to give

$$\delta\mathbf{x} = \mathbf{A^{-1}}\delta\mathbf{g} \tag{5.5}$$

The two-norm is applied to both sides of the equation, giving

$$\|\delta\mathbf{x}\|_2 = \|\mathbf{A}^{-1}\delta\mathbf{g}\|_2 \tag{5.6}$$

and the right-hand side is separated into two separate norms, producing the inequality

$$\|\delta\mathbf{x}\|_2 \leq \|\mathbf{A}^{-1}\|_2 \|\delta\mathbf{g}\|_2 \tag{5.7}$$

If the same analysis is done to Eq. (1.3) the following is found

$$\begin{gathered} \|\mathbf{A}\mathbf{x}\|_2 = \|\mathbf{g}\|_2 \\ \Longrightarrow \ \|\mathbf{A}\|_2 \|\mathbf{x}\|_2 \geq \|\mathbf{g}\|_2 \\ \Longrightarrow \ \frac{1}{\|\mathbf{x}\|_2} \leq \frac{\|\mathbf{A}\|_2}{\|\mathbf{g}\|_2} \end{gathered} \tag{5.8}$$

Equations (5.7) and (5.8) are multiplied together, so that

$$\frac{\|\delta\mathbf{x}\|_2}{\|\mathbf{x}\|_2} \leq \|\mathbf{A}\|_2 \|\mathbf{A}^{-1}\|_2 \frac{\|\delta\mathbf{g}\|_2}{\|\mathbf{g}\|_2} \tag{5.9}$$

where $\|\mathbf{A}\|_2 \|\mathbf{A}^{-1}\|_2$ is defined as the condition number of the matrix $\mathbf{A}$. This term quantifies the ill-conditioning of this matrix equation, and by extension the ill-posedness of original integral equation. Equation (5.9) is an upper bound on the perturbation in the system. If the matrix is invertible (and hence non-singular) then the condition number will be close to or equal to one. Thus there is a one-to-one correlation between the error in the data and the error in the solution. A small error in the data begets a small error in the solution, and a large error in the data produces a large error in the solution.

However, in the case of an ill-posed problem, the condition number is generally very large. Thus a small error in the data can lead to a large error in the solution, violating Hadamard's third criterion, i.e. stability. It should be noted that this is simply an upper bound of the error; it is possible, though highly unlikely, that in some cases a small error in the data will produce a small error in the solution.

The upper limit of the condition number is infinity, which occurs when the matrix is singular. In this case there are an infinite number of solutions that satisfy the governing equations and hence there is no unique solution. In linear algebra terms, there is a non-trivial nullspace that gets mapped by the matrix to zero. Hence, any linear combination of the vector or vectors in the nullspace added to the solution is also a solution; and thus there are an infinite number of solutions that can exactly satisfy the problem. This violates Hadamard's second criterion that the solution must be unique.

5.3 Tikhonov Regularization

When the governing Fredholm integral equation of the first kind is recast as an ill-conditioned matrix equation, it is possible to add extra information to the problem using linear regularization. One of the earliest regularization techniques is known as Tikhonov regularization, which was developed by Andrey Tikhonov (1943, 1963, Tikhonov and Arsenin 1977).

5.3.1 Smoothing Matrix

Tikhonov regularization works by augmenting the matrix equation with a smoothing matrix, which can be defined in several different ways. The most common option is appending the identity matrix to the kernel matrix, **A**, which is known as standard Tikhonov regularization. The equation is given as

$$\begin{bmatrix}\mathbf{A}\\ \lambda\mathbf{I}\end{bmatrix}\mathbf{x} = \begin{bmatrix}\mathbf{g}\\ \mathbf{0}\end{bmatrix} \tag{5.10}$$

where λ is the regularization parameter that determines the extent of regularization that is used. Using the identity matrix to augment the kernel matrix promotes a small solution as the augmented section of the matrix equation is satisfied when the recovered solution is zero, i.e. the prior information used here is that the solution should be small in magnitude. This makes sense

because the unregularized solution (Figure 5.4) is highly oscillatory with high frequency, and so if the magnitude of **x** is kept small, this should remove many of the unwanted components that have large noise inherent in them.

Another option for the smoothing matrix is the discrete approximation of a differential operator. Using a first order finite difference approximation, the first derivative is

$$f_i' = \frac{f_i - f_{i-1}}{\Delta x} \tag{5.11}$$

which, in matrix form is given as

$$\begin{bmatrix} 1 & -1 & 0 & & & \\ 0 & 1 & -1 & \cdots & 0 & \\ 0 & 0 & 1 & & & \\ & \vdots & & \ddots & \vdots & \\ & 0 & & \cdots & 1 & -1 \end{bmatrix} \tag{5.12}$$

having the same number of columns as the kernel matrix, and one less row. This matrix promotes a smooth solution, and is satisfied when **x** is uniform. The solution promoted by this matrix equation is smooth, such that the variations between two adjacent points in the vector **x** do not vary greatly from each other in value. This variance is of course relative and the variation can sometimes be very large, for example when the kernel matrix dominates the problem.

A third option is the first order finite difference approximation of the second derivative,

$$f_i'' = \frac{f_i - 2f_{i-1} + f_{i-2}}{(\Delta x)^2} \tag{5.13}$$

giving the smoothing matrix

$$\begin{bmatrix} 1 & -2 & 1 & & & & \\ 0 & 1 & -2 & \cdots & & 0 & \\ 0 & 0 & 1 & & & & \\ & \vdots & & \ddots & & \vdots & \\ & 0 & & \cdots & 1 & -2 & 1 \end{bmatrix} \tag{5.14}$$

which promotes a distribution that is C^2 continuous (Hansen 1998).

5.3.2 Regularization Parameter

It is crucial that the presumed information added to the system (prior) have an appropriate emphasis relative to the original ill-posed equations, which is done using the regularization parameter.

For the case of Tikhonov regularization, the regularization parameter scales the smoothing matrix. The larger the parameter, the more weight is given to the smoothing over the original ill-posed problem. Because there is obviously a point where too little regularization is used (when the regularization parameter is equal to zero) and a point where the regularization is too great (when the regularization parameter is so great that the solution is a constant value), and because the parameter varies continuously between these two limiting cases, the mean value theorem is applied to the problem to show that there should indeed be an optimal value somewhere between the two extremes.

A number of methods have been proposed to find this optimal value for the regularization parameter, two of which will be discussed here. The first method is called the Discrepancy Principle (Morozov 1966), which works under the assumption that the residual, $\|\mathbf{A}\mathbf{x}_\lambda - \mathbf{g}\|_2$, cannot be known to a greater degree of accuracy than the error in the system. Hence, λ should be chosen such that $\|\mathbf{A}\mathbf{x}_\lambda - \mathbf{g}\|_2 = \delta_e$, where δ_e is the upper bound for the two-norm of the error vector, $\|\mathbf{e}\|_2 \leq \delta_e$. This method is useful if there is an estimate of the error. Unfortunately, in

this problem the noise contaminating the light scattering data is not sufficiently characterized to make such a prediction.

Another method of determining the value of the regularization parameter is the L-Curve Criterion (Hansen and O'Leary 1993). This method makes use of an interesting graphical feature of the plot of $\log_{10}\|\mathbf{A}\mathbf{x}_\lambda - \mathbf{g}\|_2$ vs. $\log_{10}\|\mathbf{L}\mathbf{x}_\lambda\|_2$. The plot of these two values generates a curve that is the shape of an uppercase L, where the vertical segment of the curve denotes where the problem is dominated by the original matrix equation and the horizontal segment of the curve is dominated by the regularization. These two sections of the curve are referred to as undersmoothing and oversmoothing, respectively.

For the undersmoothing section of the curve, the regularization parameter is slowly increased, which begins to smooth the curve and move it towards the uniform solution that minimizes $\|\mathbf{L}\mathbf{x}_\lambda\|_2$, but has little affect on the residual, $\|\mathbf{A}\mathbf{x}_\lambda - \mathbf{g}\|_2$. After a certain point, however, the curve becomes oversmoothed, so that it is nearing a uniform value at all points, such that increasing λ does little to change $\|\mathbf{L}\mathbf{x}_\lambda\|_2$, but greatly increases $\|\mathbf{A}\mathbf{x}_\lambda - \mathbf{g}\|_2$. This is because in the oversmoothed region the curve no longer works to solve the original problem but instead is trying to solve the augmented problem, causing the solution to depart from the true solution. Thus, the optimal value of λ is found at the intersection of these two segments, at the corner of the L-Curve. An example of an L-Curve is seen in Figure 5.2

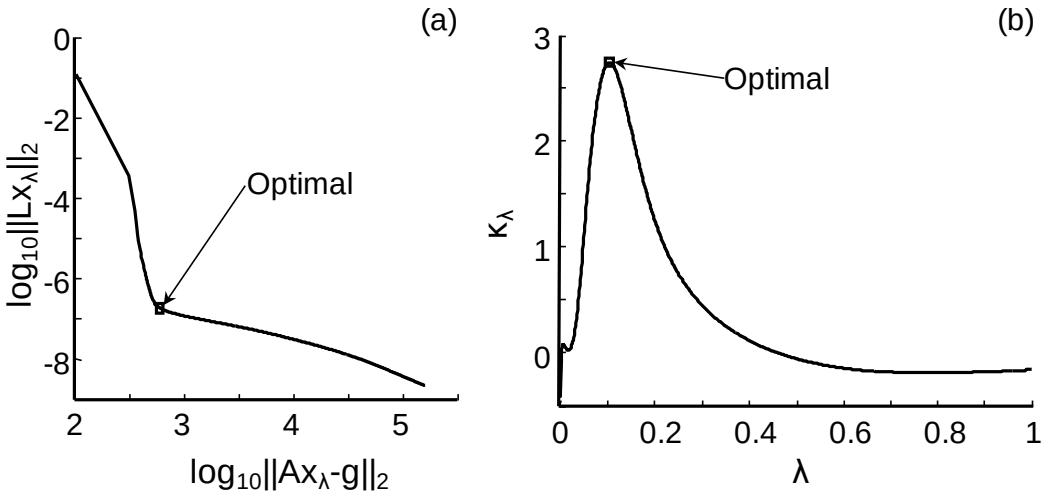

Figure 5.2 – (a) Typical L-Curve and (b) curvature function for this problem

To find the corner of an L-Curve, the curvature function is employed, defined as (Kline 1967)

$$\kappa = \frac{\rho'\eta'' - \rho''\eta'}{((\rho')^2 + (\eta')^2)^{3/2}} \tag{5.15}$$

This equation is applied at every value of λ to determine the curvature of the L-Curve, with $\rho_\lambda \equiv \log_{10}\|\mathbf{Ax}_\lambda - \mathbf{g}\|_2$ and $\eta_\lambda \equiv \log_{10}\|\mathbf{Lx}_\lambda\|_2$. The curvature function is then plotted as a function of λ and the maximum value determined. Obviously, the maximum value of the curvature function is the point of maximum curvature of the L-Curve, which should correspond to the corner. This point is denoted in the plot of the curvature function shown in Figure 5.2.

This method is not perfect. The value of the regularization parameter is taken at discrete points so the derivatives required for the curvature function must be determined using an approximate derivative method, such as finite differences. This introduces some numerical error into the equation and can occasionally lead to the corner being chosen slightly off from where it

appears that it should be. One way to get around this is to calculate higher order approximations for the derivatives, which can reduce some of the numerical noise.

5.4 Application to Artificial Data

Before a method is employed to solve experimentally obtained data, it should be validated by recovering a known distribution using synthetic data. By taking a distribution reported in the literature and numerically integrating Eq. (1.2), an artificial data set was generated with a known distribution that could then be used as a test case for regularization methods. This was done using the numerical integration scheme presented in Appendix B.1.

From previous experiments (Tian, et al. 2004) there was a large set of TEM images of soot aggregates captured from this flame at the same height and operating conditions as the light scattering experiment that could be analyzed. A histogram was formed by splitting the aggregates into bins based on the number of primary particles, which were summed to form points of the cumulative distribution function (CDF). The CDF of a lognormal distribution was then fitted to the histogram CDF by means of a least-squares fitting algorithm, and the result was taken as the distribution for the artificial data set. This procedure is described in more detail in Appendix B.4.

Once a distribution had been chosen and numerically integrated to produce an artificial data set, artificial noise was added. In Tikhonov regularization, the vertical portion of the L-Curve is generated by the noise inherent in the problem, hence noise is required to accurately determine the optimal amount of regularization needed to reproduce the system. The noise was taken to be normally distributed with a 3% standard deviation relative to the intensity at each

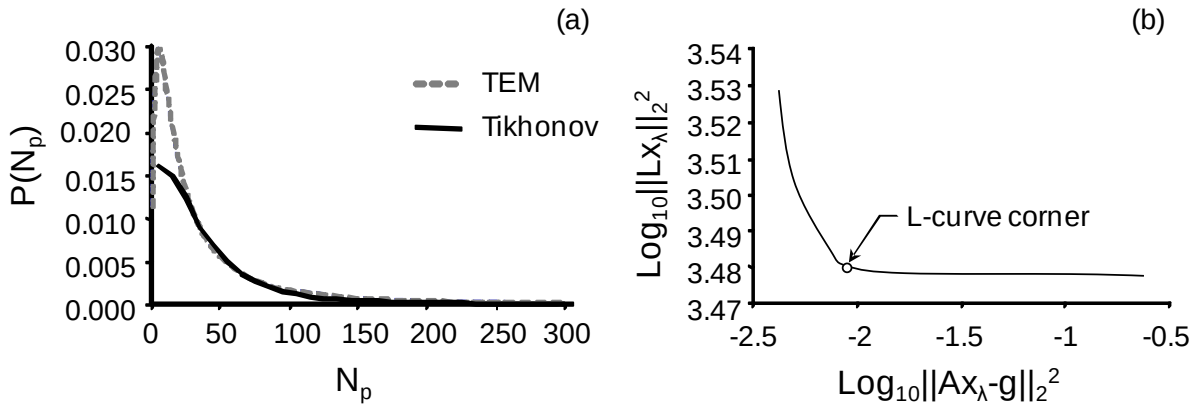

Figure 5.3 – a) Recovered distribution using Tikhonov regularization on artificial data
b) L-Curve for the artificial case

angle. The data were contaminated with this noise and Tikhonov regularization applied. The resulting distribution is in Figure 5.3 (a), with the L-Curve shown in Figure 5.3 (b).

This method is able to recover the features of the original curve, for the most part, but one feature that cannot be accurately recovered is the peak found at the small N_p section of the curve. The reason for this discrepancy can be seen by examining the kernel function, Eq. (3.15),

$$K(\theta, N_p) = \frac{N_p^2 f\left[q(\theta) R_g(N_p)\right]}{\sin(\theta)} \tag{3.15}$$

which has a strong dependency on the number of primary particles squared. Thus, the value of the kernel for large aggregates will dwarf the value for small aggregates. This essentially overwhelms any information the smaller aggregates provide and makes recovering the distribution in the small aggregate range very difficult. Outside of the small aggregate range the recovered distribution matches the original distribution quite well.

Recovering a known distribution from an artificial data set has proven to be possible, although the solution is by no means a perfect match to the distribution used to generate the data. Thus, while Tikhonov regularization is a powerful inversion method, it is not necessarily capable

of recovering the aggregate size distribution in this case, and a more powerful method should be considered that is able to provide additional more information about the system.

5.5 Maximum a Posteriori Inference

Tikhonov regularization is unable to accurately reconstruct the aggregate size distribution over the entire range of aggregate sizes under consideration, and hence an alternative method of regularization is sought that is capable of adding additional assumed information to the system.

A powerful regularization method is Maximum a Posteriori (MAP) inference, which is derived from Bayes's theorem, named for Thomas Bayes who first laid the groundwork for the theory (Bayes 1763). It is sometimes referred to as the principle of inverse probability (Zellner 1971). This technique is useful in cases when a set of data is available from an experiment and the objective of the analysis is to infer the process that generated the data, such as determining the aggregate size distribution of a soot aerosol from elastic light scattering data. The basic form of Bayes's theorem is given as

$$P(\mathbf{x}|\mathbf{g}) = \frac{P(\mathbf{g}|\mathbf{x})}{P(\mathbf{g})} P(\mathbf{x}) \tag{1.5}$$

where $P(\mathbf{x}|\mathbf{g})$ is the probability that the distribution $\mathbf{x}$ is correct given an observed data set $\mathbf{g}$, $P(\mathbf{g}|\mathbf{x})$ is the probability that the data set $\mathbf{g}$ is correct given a distribution $\mathbf{x}$ (which is also known as the likelihood function), $P(\mathbf{g})$ is the marginal probability of the data, and $P(\mathbf{x})$ is the prior probability of the distribution being correct, which will be denoted as $P_{\text{prior}}(\mathbf{x})$ to indicate that this is information that is known prior to applying regularization. Usually $P(\mathbf{g})$ is neglected because it doesn't affect the distribution that maximizes Eq. (1.5), but only scales it.

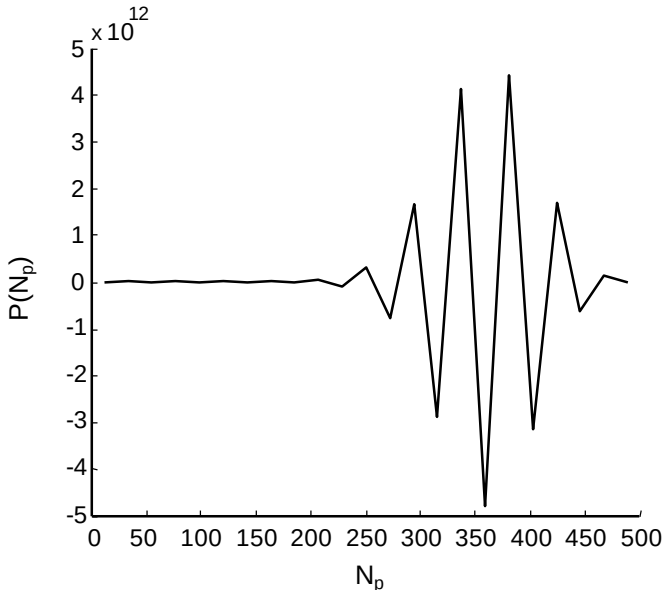

Figure 5.4 – Recovered distribution with no regularization

If no assumptions are made about the nature of **x**, then no prior information can be added and $P_{\text{prior}}(\mathbf{x}) = 1$, which is called an uniformed prior. This is equivalent to directly solving the matrix equation given in Eq. (1.3) with no regularization, which produces the nonphysical results shown in Figure 5.4.

If, however, information is known a about the system, then informed priors can be generated, and a best-estimate of the original distribution can be determined. If a distribution exists that satisfies the original matrix equation, Eq. (1.3), then that distribution is a MAP estimate (Andrews and Hunt 1977). Because a single solution exists, there must therefore be a most likely distribution. The method of applying an algorithm to determine the most likely original distribution from observed data is known as MAP estimation or MAP inference (Rosenfeld and Kak 1982).

If the matrix equation is assumed to be linear, such that $\mathbf{g} = \mathbf{Ax}$, and assuming that the data has been contaminated with Gaussian-distributed error that has the same variance for each data point $(\sigma = \sigma_j)$, then the likelihood function is

$$P(\mathbf{g}|\mathbf{x}) \propto \exp\left[-\frac{1}{\sigma^2}\|\mathbf{Ax}-\mathbf{g}\|_2^2\right] \tag{5.16}$$

The optimal solution distribution is found when this equation is maximized, which occurs when $\|\mathbf{Ax}-\mathbf{g}\|_2^2$ is minimized.

In this problem the noise inherent in the data is relative to the intensity of the scattering, hence the variance is also a function of the intensity of the scattering. To express this problem in the form of Eq.(5.16), the matrix equation must be scaled so that the variance is the same at each measurement angle. This is accomplished by scaling the measured data to equal one and scaling each row of the matrix equation by the same amount,

$$A_{i,j}^{*} = \frac{A_{i,j}}{g_i} \tag{5.17}$$

so that the equation then becomes

$$\mathbf{A}^{*}\mathbf{x} = \mathbf{1} \tag{5.18}$$

This is assumed to have been done from this point on; however, **g** will not be written as a vector of ones but will continue to be written as **g**, with the assumption that it has been scaled.

Prior information is needed to solve for the aggregate size distribution, and one piece of prior information that is known about this system is that the distribution should be smooth, as is

seen in the Tikhonov regularization smoothing matrix. This prior is known as the Gibbs prior, and can be written as

$$P_{\text{Gibbs}}(\mathbf{x}) \propto \exp\left[-\beta\|\mathbf{Lx}\|_2^2\right] \tag{5.19}$$

where $\mathbf{L}$ is the Tikhonov smoothing matrix. Similar to the regularization parameter from Tikhonov regularization, β determines the appropriateness of this prior, allowing it to be scaled to affect the solution more or less, which can be determined heuristically. If the Gibbs prior and the likelihood function are added to Eq. (1.5), the objective function becomes

$$P(\mathbf{x}|\mathbf{g}) \propto \exp\left[-\frac{1}{\sigma^2}\|\mathbf{Ax}-\mathbf{g}\|_2^2 - \beta\|\mathbf{Lx}\|_2^2\right] \tag{5.20}$$

By minimizing Eq. (5.20) Bayesian inference with the Gibbs prior is equivalent to Tikhonov regularization (Hansen 1998).

Bayesian analysis allows for the incorporation of other prior information. For example, by definition the probability density of the aggregate size distribution must be nonnegative. This can be enforced by

$$P_{\text{nonneg}}(\mathbf{x}) = \prod_{j=1}^{n} I_{[0,\infty)}(x_j) \tag{5.21}$$

where I is the indicator function, which is equal to zero when x_j is outside the interval indicated and unity otherwise. This is multiplied onto Eq. (5.20) to give

$$P(\mathbf{x}|\mathbf{g}) \propto \exp\left[-\frac{1}{\sigma^2}\|\mathbf{Ax}-\mathbf{g}\|_2^2 - \beta\|\mathbf{Lx}\|_2^2\right] \times \left[\prod_{j=1}^{n} I_{[0,\infty)}(x_j)\right] \tag{5.22}$$

which gives a zero probability when $\mathbf{x}$ contains any negative terms and a positive probability otherwise.

Finally, if need be, a prior may be added that promotes a certain type of distribution. This is to be avoided if at all possible; the recovered distribution should depend on the original matrix equation to as great an extent as possible, so that the solution is not biased by the priors. If, however, adding the nonnegative prior to Tikhonov regularization is insufficient by itself to accurately reconstruct the size distribution, then promoting a certain distribution type may be required. This is done by using the prior (Johnson 2004)

$$P_{\text{pres}}(\mathbf{x}) \propto \exp\left[-\alpha\|\mathbf{x}-\mathbf{x}_{\text{dist}}(\boldsymbol{\theta}^*)\|_2^2\right] \tag{5.23}$$

where α is a regularization parameter that allows the relative importance of this prior to be set, and $\mathbf{x}_{\text{dist}}$ is the best fit of a prescribed distribution, such as a lognormal distribution, to the current solution $\mathbf{x}$ in an iterative manner. The vector $\boldsymbol{\theta}^*$ contains the parameters of the prescribed distribution, fitted using a Kolmogorov-Smirnov goodness-of-fit statistic (Press, et al. 2007), (Kolmogorov 1933), the algorithm of which is explained in Appendix B.2.

If all of these priors are added together, and the substitutions $\lambda^2 \equiv 2\sigma^2\beta$ and $\gamma^2 \equiv 2\sigma^2\alpha$ are made, the objective function becomes

$$P(\mathbf{x}|\mathbf{g}) \propto \exp[-\|\mathbf{A}\mathbf{x}-\mathbf{g}\|_2^2 - \lambda^2\|\mathbf{L}\mathbf{x}\|_2^2 - \gamma^2\|\mathbf{x}-\mathbf{x}_{\text{dist}}(\boldsymbol{\theta}^*)\|_2^2] \times \left[\prod_{j=1}^{n} I_{[0,\infty)}(x_j)\right] \tag{5.24}$$

Due to the monotonic nature of the natural logarithm, this can be rewritten as

$$\log\left[P(\mathbf{x}|\mathbf{g})\right] \propto -\|\mathbf{A}\mathbf{x}-\mathbf{g}\|_2^2 - \lambda^2\|\mathbf{L}\mathbf{x}\|_2^2 - \gamma^2\|\mathbf{x}-\mathbf{x}_{\text{dist}}(\boldsymbol{\theta}^*)\|_2^2 + \log\left[\prod_{j=1}^{n} I_{[0,\infty)}(x_j)\right] \tag{5.25}$$

where the optimal $\mathbf{x}$ maximizes this equation. Many optimization schemes work by minimization rather than maximization, so this is rewritten as

$$\mathbf{x}^* = argmin\left\{\|\mathbf{Ax}-\mathbf{g}\|_2^2 + \lambda^2\|\mathbf{Lx}\|_2^2 + \gamma^2\|\mathbf{x}-\mathbf{x}_{\text{dist}}(\boldsymbol{\theta}^*)\|_2^2 - \log\left[\prod_{\text{j}=1}^{\text{n}} I_{[0,\infty)}(\text{x}_\text{j})\right]\right\} \tag{5.26}$$

This equation can be rewritten in an iterative linear form, where each step is linear but the steps are taken iteratively, as

$$\mathbf{x}^* = argmin\left\|\begin{bmatrix}\mathbf{A}\\ \lambda\mathbf{L}\\ \gamma\mathbf{I}\end{bmatrix}\mathbf{x} - \begin{bmatrix}\mathbf{g}\\ \mathbf{0}\\ \gamma\mathbf{x}_{\text{dist}}(\boldsymbol{\theta}^*)\end{bmatrix}\right\|_2^2 \quad \text{s.t. } \mathbf{x} \geq 0 \tag{5.27}$$

The first step of this objective function must be taken with $\gamma = 0$, as $\mathbf{x}_{\text{dist}}(\boldsymbol{\theta}^*)$ is fitted to the distribution from the previous step. Hence, the first step is the nonnegative Tikhonov solution. Generally, the value of λ is chosen heuristically for this method, as there is no graphical method synonymous with the L-curve that can solve two variables simultaneously.

5.6 Application to Experimental Data

Within this regularization framework, measurement data from the actual experimental apparatus was inverted using strictly Tikhonov regularization. The goal was to use the least amount of regularization possible to solve the problem so that the solution would be dependent mainly upon the data available from the problem and not the priors that were added.

There is one main problem with solving for the aggregate size distribution of the soot aerosol system; the scaling coefficient C. As noted in Chapter 3, this coefficient can be determined from experimental measurements provided the scattering signal is calibrated and the soot aggregate number density is known; in the case of this experiment these values were not known. Hence, C is treated as another variable to be solved for during the regularization process.

First, Tikhonov regularization was applied to the problem. This was done by choosing a range of values of the scaling coefficient and setting a constant value of the regularization

parameter, λ. The problem was inverted using Tikhonov regularization with the first order smoothing matrix at each value of λ and C. The residual, $\mathrm{F}\left(\mathbf{x}_\lambda(C)\right) = \|\mathbf{A}\mathbf{x}_\lambda - \mathbf{g}\|_2^2 + \lambda^2\|\mathbf{L}\mathbf{x}_\lambda\|_2^2$, was calculated at each set of values as well, and the results are shown below in Figure 5.5 (a). This shows that there is no optimal solution when using strictly Tikhonov regularization, with no other priors. In other words, any value of C is almost equally able to solve the problem. Thus, more information is needed to determine the C coefficient.

Next, the nonnegative prior was added to the problem, and the same procedure was undertaken, with the residual $\mathrm{F}\left(\mathbf{x}_\lambda(C)\right) = \|\mathbf{A}\mathbf{x}_\lambda - \mathbf{g}\|_2^2 + \lambda^2\|\mathbf{L}\mathbf{x}_\lambda\|_2^2 \ \text{s.t.}\, \mathbf{x}_\lambda \geq 0$ employed. The results were much the same, as shown in Figure 5.5 (b). Even with both smoothing and nonnegativity added as priors, an optimal value of the scaling coefficient cannot be recovered. Therefore, more information was needed, added in the form of a prescribed distribution.

To add a prescribed distribution, however, a value for the regularization parameter λ had to be chosen. Considering Figure 5.5, λ should be chosen to be somewhere between 0.001 and 0.1; hence, heuristically, λ has been chosen to be equal to 0.01.

By adding the prescribed distribution, a minimum value for the scaling parameter appears, allowing both the scaling parameter, C, and the weighting parameter for the prescribed distribution, γ, to be determined simultaneously. The residual in this case is $\mathrm{F}\left(\mathbf{x}_\lambda(C)\right) = \|\mathbf{A}\mathbf{x}_\lambda - \mathbf{g}\|_2^2 + \lambda^2\|\mathbf{L}\mathbf{x}_\lambda\|_2^2 + \gamma^2\|\mathbf{x}_\lambda - \mathbf{x}_{\mathrm{dist}}\|_2^2 \ \text{s.t.}\, \mathbf{x}_\lambda \geq 0$. The plot of the scaling parameter versus the residual is shown below in Figure 5.6, with the minimum denoted by a diamond. The optimal value of the scaling coefficient, C, is determined to be 0.508×10^{-5}.

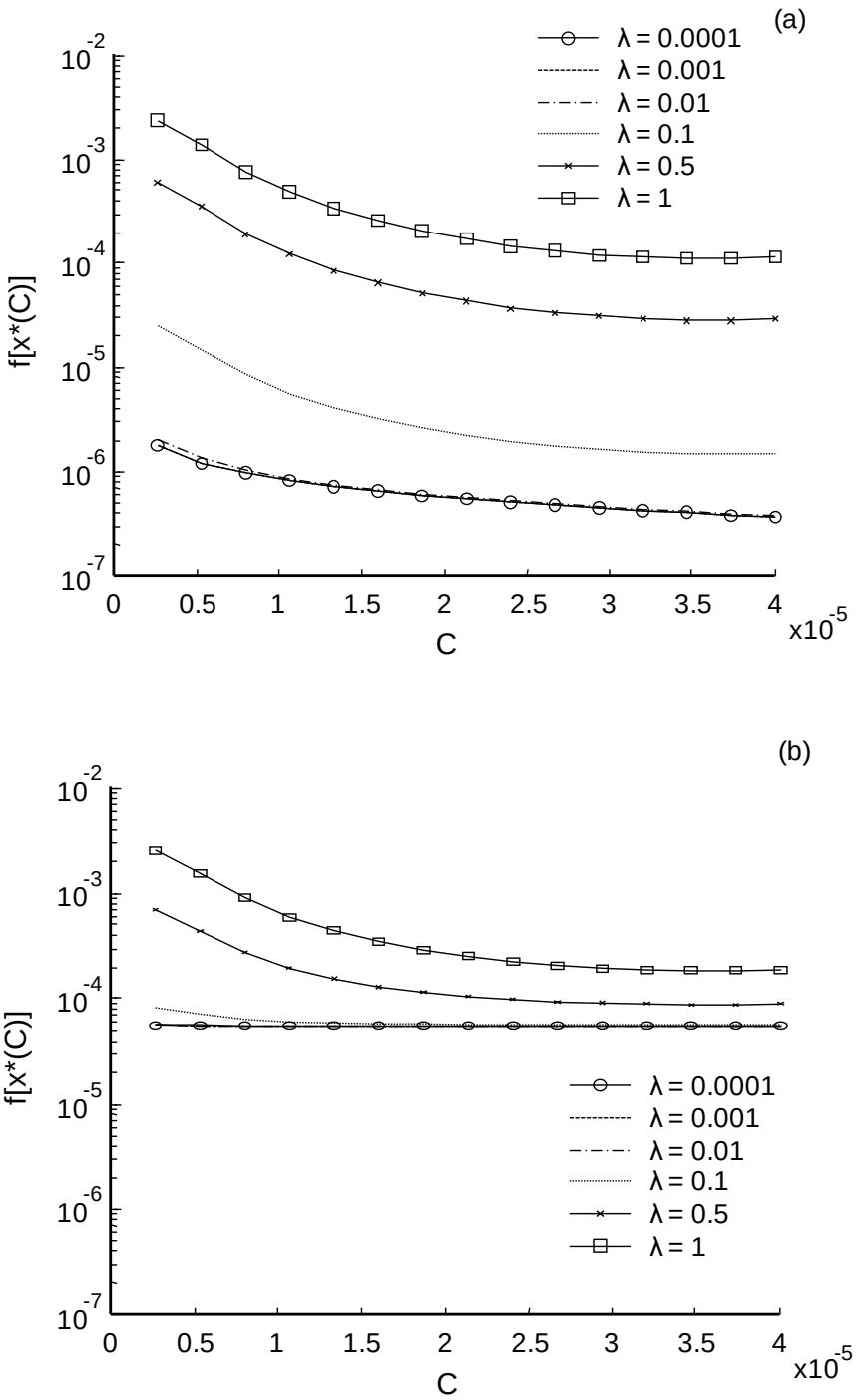

Figure 5.5 – Residual plotted over a range of λ and C values for (a) Tikhonov regularization, and (b) Tikhonov regularization with nonnegativity prior added

These residuals were used, rather than simply $\mathrm{F}\left(\mathbf{x}_\lambda(C)\right) = \|\mathbf{A}\mathbf{x}_\lambda - \mathbf{g}\|_2^2$, because the system being solved is the augmented matrix given in Eq. (5.27). Because MAP inference is being employed to solve this problem, the total probability should be used, rather than only the likelihood function. Hence, Figure 5.6 is generated with the augmented matrix equation.

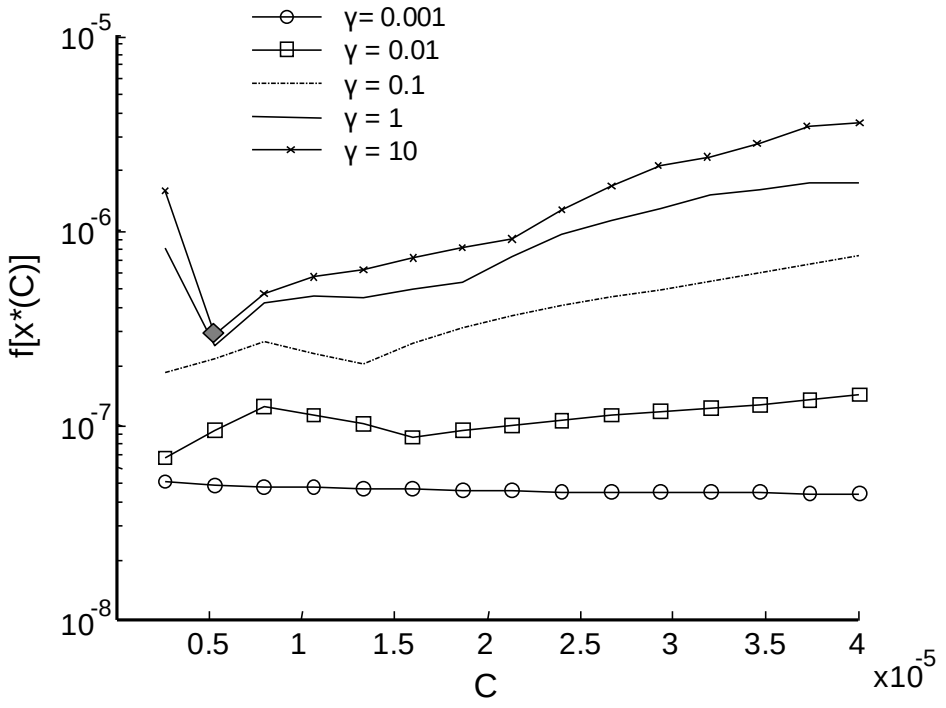

Figure 5.6 – Residual plotted over a range of values for the scaling parameter with nonnegativity and a prescribed distribution

By setting $\lambda = 0.01$, $\gamma = 10$, and $C = 0.508$, the probability density function of the aggregate size distribution was recovered using Eq. (5.27). The resulting distribution is given in Figure 5.8, plotted alongside the lognormal distribution fitted to the TEM histogram for this flame. Also plotted is the fitted prescribed lognomal distribution, $\mathbf{x}_{\mathrm{dist}}(\boldsymbol{\theta}^*)$. The recovered distribution was then used in Eq. (1.3) to reconstruct the scattering data, which is plotted in Figure 5.7 against the actual measured data.

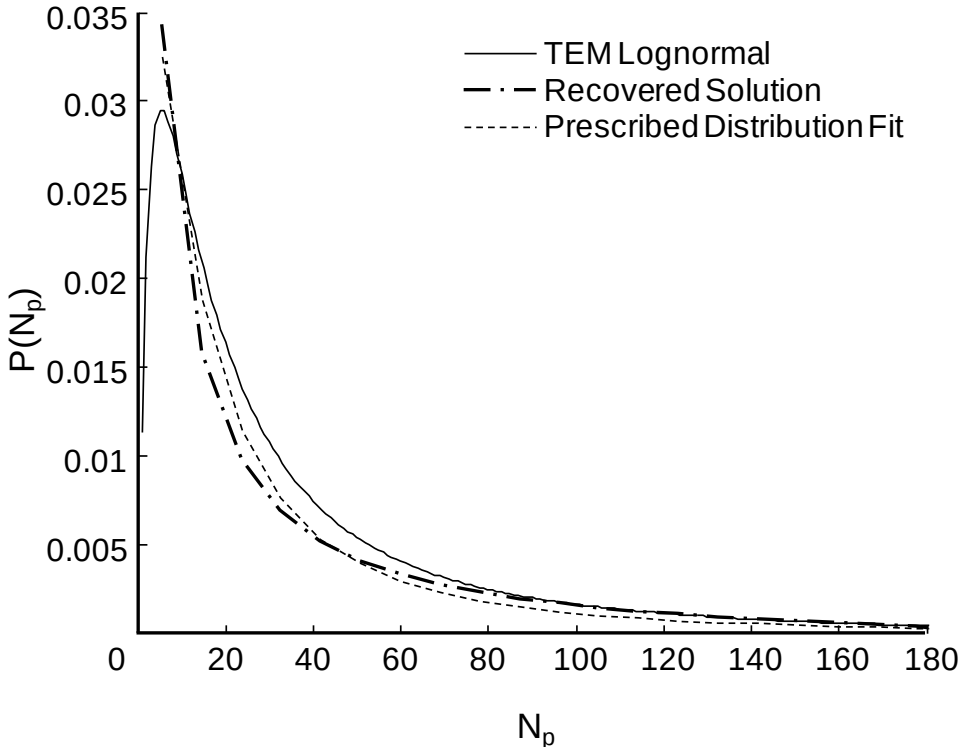

Figure 5.8 – Recovered distribution

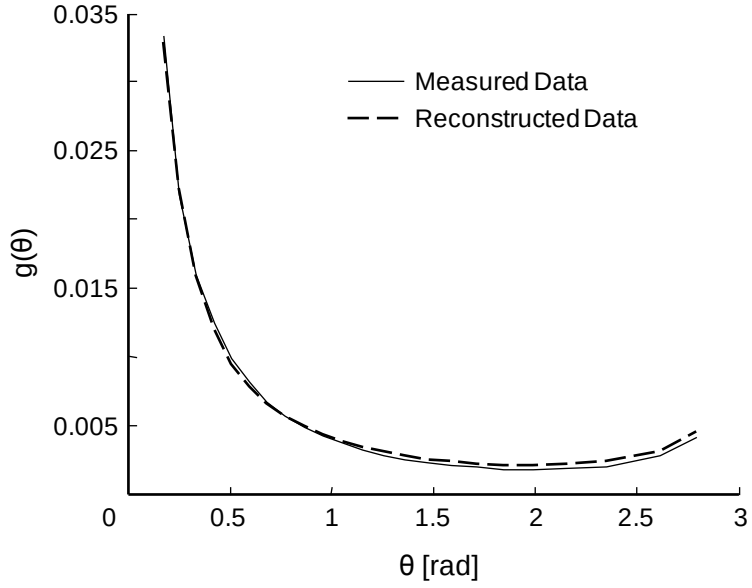

Figure 5.7 – Reconstructed scattering data

5.7 Discussion of the Recovered Solution

This chapter showed how the probability density function of the aggregate size distribution of soot aggregates is recovered from multiangle scattering data. Due to the severe ill-posedness of the problem there are an infinite number of solutions that accurately reconstruct the measured light scattering data, so it is necessary to impose extra information in the form of assumed solution attributes, or *priors*, through regularization.

The recovered solution displays similar trends to the expected solution from TEM analysis of this flame. There is an initial spike of smaller aggregates, and then the distribution slowly decreases towards zero. There is a section in the middle of the curve that doesn't match at all well, but there is the initial spike, and the tail of the curve also matches the TEM distribution quite well. Overall, this technique has been quite successful.

A great deal of regularization was needed to actually solve this problem. Three distinct priors were added before the problem became specific enough to determine the scaling parameter and the two regularization parameters required. Furthermore, these values were determined by inspection in a heuristic manner. A more rigorous, and automated, method should be developed before this technique is put into use outside of a laboratory.

The recovered distribution does reproduce the light scattering data to within experimental accuracy. Thus, this result is an entirely plausible solution, recovered without forcing a specific distribution type. That is one of the key points of using regularization over assuming a specific distribution type. This method promotes a lognormal distribution, but does not force it. This is made clear in Figure 5.8, where the final solution and the prescribed distribution are both shown. The final recovered distribution is not exactly the lognormal distribution that was used in Eq.

(5.27), but has deviated from that distribution because the solution of the objective function led in that direction. This is also supported by Figure 5.9, which is a plot of the objective function $\mathrm{F}(\mathbf{x}) = \|\mathbf{Ax} - \mathbf{g}\|_2^2$, where $\mathbf{x}$ is constrained to be a lognormal distribution. One interesting note is that the prescribed distribution, $\mathbf{x}_{dist}$, does not lie in the valley of solutions that all recover the measured data to within experimental accuracy, although it is very close. This provides evidence that the solution can depart from the optimal lognormal solution if the other priors suggest that such a move would be beneficial.

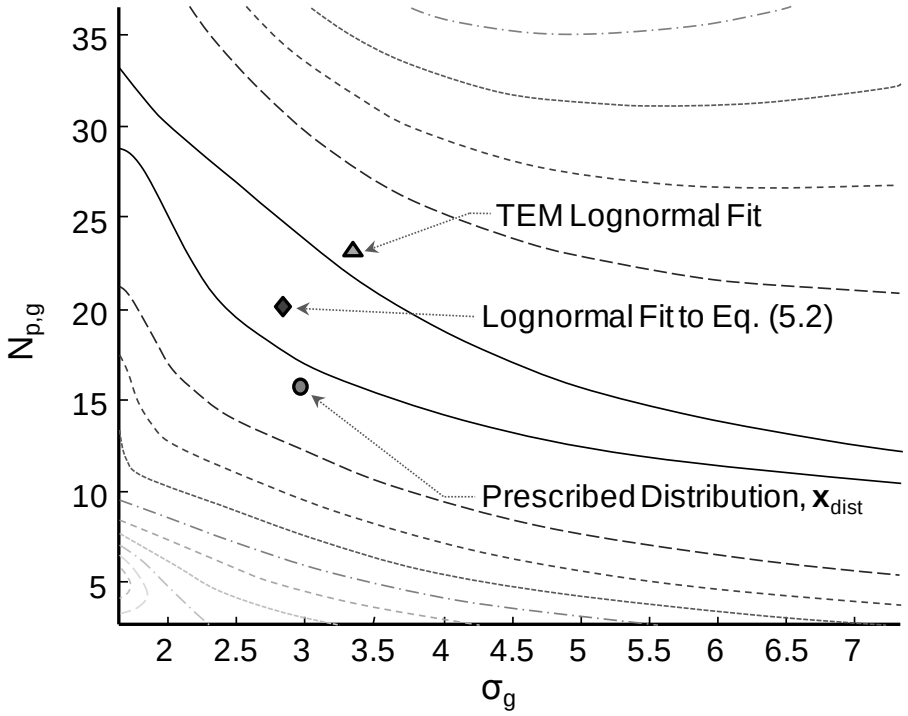

Figure 5.9 – Objective function of lognormal distributions with three distributions of interest marked.

Chapter 6
Optimal Design of Experiment

6.1 Introduction

Most researchers who work on light scattering from soot aggregates heuristically choose the angles at which light scattering measurements are made, most often at uniform angular increments. But one might wonder if there is indeed a better solution, a more precise method for determining the optimal angles in this experiment, or even if such a set of optimal angles exist. As described in Chapter 3, Sorensen (2001) suggested that angles be chosen to be uniform in the q-domain, rather than in the angular domain, because the inverse of the scattering wave vector is the length scale of the scattering experiment, giving physical insight into the scattering mechanics of the aggregates. This was the method used to determine the experimental angles at which light scattering measurement data were taken for this experiment.

This chapter will determine if there is an optimal configuration of angles that produce the best defined system, consider the size of the matrix equation that should be chosen in Eq. (1.3), introduce a rigorous method for determining an optimal angle set, and justify the choice of such an angle set by demonstrating that it reduces the error inherent in the problem.

6.2 Optimal Size of the Kernel Matrix

The kernel matrix is simply the discrete form of the kernel function, and while the number of rows of the matrix is set by the number of angular measurements, the number of columns can be chosen however one pleases. Is there an optimal way to set the number of columns of this matrix? What is the advantage of having more or less columns, or is there even an advantage?

The first consideration is the refinement of the aggregate size distribution. The final solution should be as refined as possible, which suggests adding more columns to the kernel matrix. Recalling the basic matrix equation

$$\mathbf{Ax} = \mathbf{g} \tag{1.3}$$

it is apparent that the more columns there are in **A**, the more points there are in the discrete distribution **x**. If there is a set upper bound to the number of primary particles in an aggregate, these additional points will improve the solution resolution and produce smoother curves, as well as allow for smaller features of the curves to appear. An example of this is shown in Figure 6.1, where two curves of the same function, a bimodal distribution, were generated with vastly different numbers of points.

To understand how this affects the system, however, one should look at the singular value decomposition of the kernel matrix. Any matrix, real or complex, square or rectangular, can be decomposed using singular value decomposition into

$$\mathbf{A} = \mathbf{U\Sigma V}^{\mathrm{T}} \tag{6.1}$$

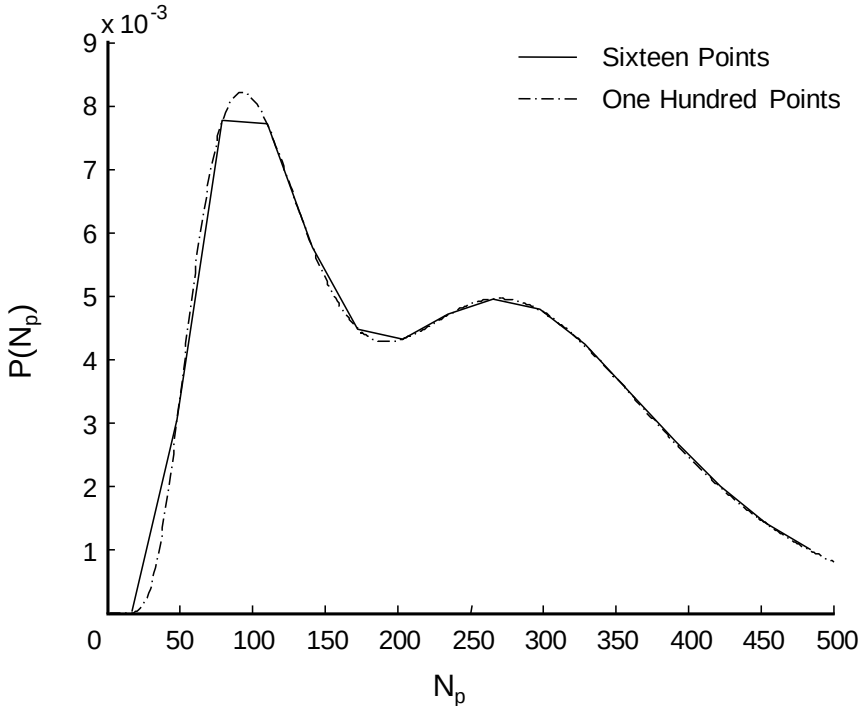

Figure 6.1 – Bimodal distribution at two different levels of refinement.

where the matrix $\mathbf{V}$ is an orthonormal basis for the input vector, $\mathbf{x}$, and $\mathbf{V}^T$ is the conjugate transpose of $\mathbf{V}$. The matrix $\mathbf{U}$ is an orthonormal basis for the space that contains the measured data, $\mathbf{g}$. $\mathbf{\Sigma}$ is a diagonal matrix and contains the singular values of the system. The singular values are essentially weights that determine how strongly the solution depends on each of the basis vectors of $\mathbf{V}$. If the singular values are small the measured data does not depend greatly on the corresponding basis vector. Furthermore, if any of the singular values are equal to zero, the matrix is singular and cannot be inverted using normal algebra. If a square matrix is singular, there exist one or more basis vectors $\mathbf{v}_k$ that map to zero. Hence, $\mathbf{A}\mathbf{v}_k = \mathbf{0}$ and so there exist an infinite set of linear combinations of these vectors that exactly satisfy the problem. Suppose the vector $\mathbf{x}^*$ is an exact solution, such that $\mathbf{A}\mathbf{x}^* = \mathbf{g}$; then the solution $\mathbf{A}(\mathbf{x}^* + \mathbf{v}_k) = \mathbf{g}$ is also an exact solution. Hence, the solution is not unique, and Hadamard's second criterion is violated, so this is an ill-posed problem.

If the **A** matrix is defined to be rectangular, of size m × n such that m $< n$, there are only m singular values. Hence, **Σ** will be of the form

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{\mathrm{m}} & 0 & \cdots & 0 \end{bmatrix} \tag{6.2}$$

Thus, every vector V_i such that $i > m$ will have no effect on the solution, and will be a portion of the solution space that is unreachable. This is demonstrated by the rank-nullity theorem, which states that for an m × n matrix **A**, m ≤ n

$$\mathrm{nullity}(\mathbf{A}) = \mathrm{n} - \mathrm{m} \tag{6.3}$$

The nullity of **A** is defined as the number of nonzero vectors that, when multiplied by **A**, produce the zero vector. This is one of the reasons that this problem is very ill-posed when formed as a rectangular matrix equation. There is an infinite set of vectors that are scalar multiples of the null basis vectors which, when added to the solution vector, adequately solve the problem. In mathematical form this is given as

$$\mathbf{A}(\mathbf{x} + \mathbf{x}_{\mathrm{null}}) = \mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{x}_{\mathrm{null}} = \mathbf{A}\mathbf{x} + \mathbf{0} = \mathbf{g} \tag{6.4}$$

The problem is that $\mathbf{x}_{\mathrm{null}}$ can change the shape of the aggregate size distribution greatly yet have no effect on the reconstructed measuring data, and hence no effect on the objective function, Eq. (5.16).

To reduce the nullity of **A** the matrix should be defined as close to square as possible. This reduces the nullity to zero, or close to zero, and reduces the section of the basis that cannot be mapped to or, in the case of the forward problem, mapped from. However, reducing the nullity to zero doesn't necessarily solve the problem. The other thing to look at is the actual singular values themselves.

6.3 Singular Values

The singular values for the experimental data set with 23 angles chosen uniformly in the q-domain are shown in Figure 6.2. This range of singular values is exactly what is expected from an ill-posed problem (Hansen 1998). Attempting to invert this problem using the singular value decomposition method gives a great deal of insight into the problem these singular values pose.

Using singular value decomposition as defined in Eq. (6.1), Eq. (1.3) can be rewritten as

$$\mathbf{U\Sigma V^T x = g} \tag{6.5}$$

Thus, $\mathbf{x}$ can be written as

$$\mathbf{x = V\Sigma^{-1}U^T g} \tag{6.6}$$

recalling that the inverse of an orthogonal matrix is simply the transpose of the matrix. This can be written as a sum of vector multiplication,

$$\mathbf{x} = \sum_{j=1}^{n} \frac{\mathbf{u}_j^T \mathbf{g}}{\sigma_j} \mathbf{v}_j \tag{6.7}$$

where $\mathbf{v}_j$ and $\mathbf{u}_j$ are column vectors of $\mathbf{V}$ and $\mathbf{U}$, respectively. It should be noted that because $\mathbf{\Sigma}$ is a diagonal matrix the inverse is simply the diagonal matrix whose entries are the reciprocals of the diagonal entries of $\mathbf{\Sigma}$. Thus, if there are very small singular values, as in the case shown in Figure 6.2, the corresponding basis vectors of $\mathbf{V}$ and $\mathbf{U}$ will dominate the problem, and will generally cause the solution of $\mathbf{x}$ to be very large and oscillatory. For this reason regularization is added to the problem, to damp out these highly unrealistic solutions.

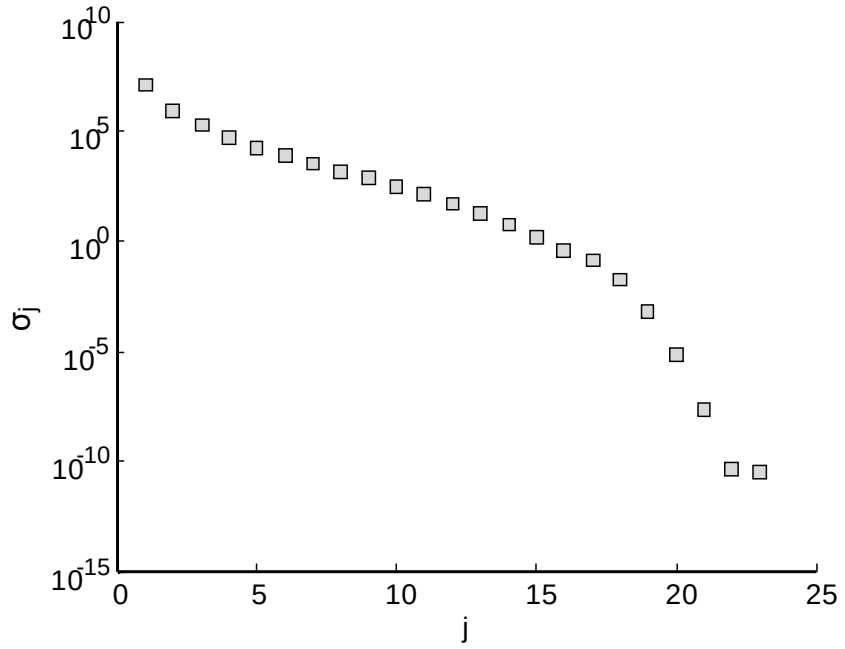

Figure 6.2 – Singular values of the kernel matrix defined with experimental angles.

The question remains as to whether the error in the solution can be reduced, however, and hence decreasing the amount of regularization required, by means of optimizing the angles at which the light scattering is measured.

6.4 Sensitivity Matrix

In order to optimize the angles of the problem, an objective function must first be formed that varies with the set of angles used and is optimal when the set of angles is optimal. The starting point of determining this objective function is defining the ellipse of possible solutions. This is given as (Press, et al. 2007)

$$\mathbf{x} = \sum_{j=1}^{n} \frac{\mathbf{u}_j^T \mathbf{g}}{\sigma_j} \mathbf{v}_\mathrm{j} \pm \frac{\mathbf{v}_1}{\sigma_1} \pm \frac{\mathbf{v}_2}{\sigma_2} \pm \cdots \pm \frac{\mathbf{v}_\mathrm{n}}{\sigma_\mathrm{n}} \tag{6.8}$$

This decomposition shows that the errors inherent in the recovered solution are linearly independent; the vectors $\mathbf{v}_\mathrm{i}$ are orthogonal basis vectors that form the principal axes for the confidence ellipse.

The ellipse formed by the error terms in Eq. (6.8) is not a confidence interval, but instead generates the shape of the ellipse that follows the contours of the objective function. The column vectors of $\mathbf{V}$ are the eigenvectors of the problem, and the reciprocal of the singular values give the relative length of the ellipse along a principle axes. Thus, if the ratio between the largest and smallest eigenvalues is two, which implies that the length of the error ellipse along the direction of the eigenvector corresponding to the largest eigenvalue should be two times larger than that in the direction of the eigenvector corresponding to the smallest eigenvalue. This is shown in Figure 6.3.

To determine a confidence interval, we begin with the Chi-Squared function, defined as

$$\Delta\chi^2 = \sum_{i=1}^{n} Z_i^2 \tag{6.9}$$

where Z_i is an independent random variable that is normally distributed. This random variable must have zero mean and unitary standard deviation. Consider the logarithm of the likelihood function (Eq. (5.16)); this function is normally distributed with a mean of zero. However, the standard deviation of $\mathbf{Ax - g}$ is not unitary, and so σ must be included to scale the equation to have unitary standard deviation. Thus, Eq. (6.9) can be written as

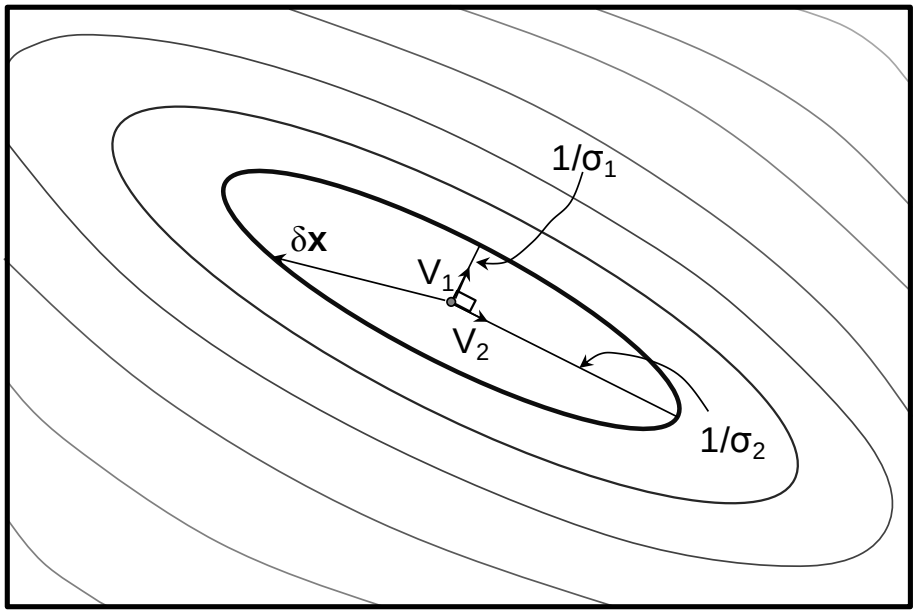

Figure 6.3 – Ellipse of solutions in two dimensions

$$\Delta\chi^2 = \sum_{i=1}^{n} \left(\frac{\mathbf{A}_i \mathbf{x} - \mathrm{g_i}}{\sigma_i} \right)^2 \tag{6.10}$$

where $\mathbf{A}_i$ is a row vector of $\mathbf{A}$. Equation (5.16) is the summation of the squares of normal distributions, producing the probability of the data being correct given a certain distribution. Similarly, integrating the chi-square probability distribution (Özisik and Orlande 2000) from zero to $\Delta\chi^2$ with Equation (6.10) as the argument gives the probability of a vector $\mathbf{x}$ producing a value that is smaller than $\Delta\chi^2$.

Because the normal distributions have all been scaled, in Eq. (5.17), the standard deviations are all the same, i.e. $\sigma_i = \sigma$ for all i. Hence, this becomes

$$\Delta\chi^2 = \frac{1}{\sigma^2} \sum_{i=1}^{n} (\mathbf{A}_i \mathbf{x} - \mathrm{g_i})^2 \tag{6.11}$$

If this is rewritten in matrix form, it becomes

$$\Delta\chi^2 = \frac{1}{\sigma^2}[\mathbf{Ax} - \mathbf{g}]^{\mathrm{T}}[\mathbf{Ax} - \mathbf{g}] \tag{6.12}$$

Decomposing $\mathbf{x}$ is into $\mathbf{x}^{\mathrm{ex}} + \delta\mathbf{x}$, then gives

$$\Delta\chi^2 = \frac{1}{\sigma^2}[\mathbf{A}(\mathbf{x}^{\mathrm{ex}} + \delta\mathbf{x}) - \mathbf{g}]^{\mathrm{T}}[\mathbf{A}(\mathbf{x}^{\mathrm{ex}} + \delta\mathbf{x}) - \mathbf{g}] \tag{6.13}$$

and by the assumption that $\mathbf{Ax}^{\mathrm{ex}} = \mathbf{g}$ this simplifies to

$$\Delta\chi^2 = \frac{1}{\sigma^2}[\mathbf{A}\delta\mathbf{x}]^{\mathrm{T}}[\mathbf{A}\delta\mathbf{x}] = \frac{1}{\sigma^2}\delta\mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}\delta\mathbf{x} \tag{6.14}$$

where $\delta\mathbf{x}$ defines the boundaries of an ellipse that forms a confidence interval. The percentage of the solutions that fall within the confidence interval is determined by the value of $\Delta\chi^2$; a list of specific values can be found in both Press et al. (2007) and Özisik and Orlande (2000).

This equation is equivalent to the equation given by Press et al. (2007) and Emery and Fadale (1996),

$$\Delta\chi^2 = \delta\mathbf{x}^{\mathrm{T}}\boldsymbol{\Gamma}^{-1}\delta\mathbf{x} \tag{6.15}$$

where $\boldsymbol{\Gamma}$ is the covariance matrix. Özisik and Orlande (2000) define the covariance matrix as

$$\boldsymbol{\Gamma} \equiv \sigma^2(\mathbf{J}^{\mathrm{T}}\mathbf{J})^{-1} \tag{6.16}$$

where $\mathbf{J}$ is the sensitivity matrix and σ is the standard deviation. Because of the linearity of the soot aerosol system, the sensitivity matrix is identical to the kernel matrix. This is derived in Appendix B.5.

Considering Eq. (6.14), both $\Delta\chi^2$ and σ^2 are constants; hence, to reduce the length of the vector $\delta\mathbf{x}$ from the centre of the ellipse to the edge the term $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ should be maximized. There are a number of potential ways to maximize this matrix. Maximizing the determinant as suggested

by Özisik and Orlande (2000) is the method used in this work. Thus, the objective function employed to shrink the size of the ellipse is given by

$$\mathbf{f} = \max\left(\det(\boldsymbol{\Gamma}^{-1})\right) = \max\left(\det(\mathbf{J}^{\mathrm{T}}\mathbf{J})\right) = \max\left(\det(\mathbf{A}^{\mathrm{T}}\mathbf{A})\right) \tag{6.17}$$

This objective function works by maximizing the singular values of the kernel matrix **A**. By definition,

$$\det(\mathbf{A}^{\mathrm{T}}\mathbf{A}) = \prod_{j=1}^{n} \sigma_j^2 \tag{6.18}$$

where the σ_j's are the singular values of the kernel matrix, **A**. Thus, by maximizing the determinant of the kernel matrix, the singular values are made as large as possible, and hence the inversion is made less ill-posed. As well, the condition number of a matrix is the ratio of the largest to smallest singular values, so by making the small singular values as large as possible, the conditioning of the kernel matrix is made as large as possible.

Next, a simple test case will be undertaken to demonstrate some of these concepts.

6.4.1 A Two-Dimensional Example

To begin, define a two dimensional linear system,

$$\mathbf{Bx} = \mathbf{a} \tag{6.19}$$

with the 'kernel' matrix

$$\mathbf{B} = \begin{bmatrix} 4 & 2.5 \\ 4 & 2 \end{bmatrix} \tag{6.20}$$

and the 'data' vector

$$\mathbf{a} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} \tag{6.21}$$

The solution to this system is readily obvious as

$$\mathbf{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{6.22}$$

and Eq. (6.7) is able to reconstruct this answer quite accurately. This system was chosen because it is ill-posed and the values in the solution vector are the same. Hence relative error in the solution vectors will also be identical. This is important because it allows this system to be written in the form of Eq. (6.14). Computing the singular value decomposition provides two important pieces of information: first, the singular values for this matrix; and second, the basis vectors for the solution, the columns of the **V** matrix. These are given as

$$\mathbf{\Sigma} = \begin{bmatrix} 6.4927 & 0 \\ 0 & 0.3080 \end{bmatrix}, \mathbf{V} = \begin{bmatrix} -0.8710 & 0.4914 \\ -0.4914 & -0.8710 \end{bmatrix} \tag{6.23}$$

The data is contaminated with 3% normally distributed error relative to the data, and Eq. (6.19) is inverted using Eq. (6.7). This process is repeated one hundred times, with the resulting vectors **x** plotted in Figure 6.4, along with the ellipse defined by Eq. (6.8). As well, the true solution is plotted as a square.

Next, a confidence interval must be determined. Equation (6.8) is not a confidence interval because it does not scale with error in the measured data; the ellipse is always the same size and shape. It does, however, provide the ratio of the lengths of the axes of the ellipse, which follow the contours of the objective function and remain consistent.

If the axes of the ellipse in Eq. (6.8) are defined as $x_1' = 1/\sigma_1$ and $x_2' = 1/\sigma_2$, then in general, for any error ellipse in this test case, the ratio of the axes should be $x_2/x_1 = \sigma_1/\sigma_2$, which is the condition number of the matrix **B**.

Written in matrix form, Eq. (6.16) becomes

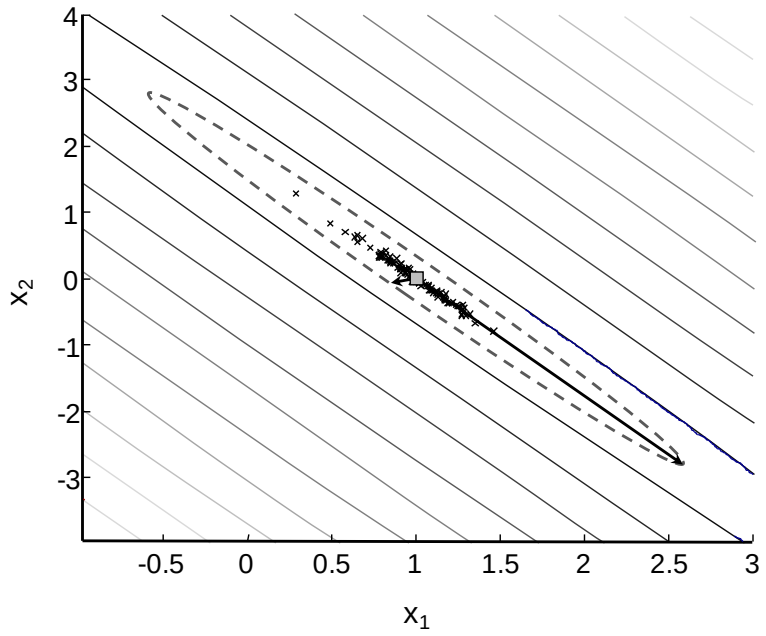

Figure 6.4 – Solutions to two-dimensional test case, 3% normal error, Eq. (6.8) ellipse shown

$$\Delta\chi^2 = \frac{1}{\sigma^2}\begin{bmatrix}\delta\mathbf{x}_1\\ \delta\mathbf{x}_2\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\mathrm{B}_{11} & \mathrm{B}_{21}\\ \mathrm{B}_{12} & \mathrm{B}_{22}\end{bmatrix}\begin{bmatrix}\mathrm{B}_{11} & \mathrm{B}_{12}\\ \mathrm{B}_{21} & \mathrm{B}_{22}\end{bmatrix}\begin{bmatrix}\delta\mathbf{x}_1\\ \delta\mathbf{x}_2\end{bmatrix} \tag{6.24}$$

The goal, then, is to determine the length of the principle axes of the ellipse.

If a confidence interval is chosen, there is a corresponding $\Delta\chi^2$ value that is specified. A list of some common values are given in Press et al. (2007) and Özisik and Orlande (2000). The standard deviation must be determined, which is simply the standard deviation inherent in the measured data. This is determined by solving the problem a great number of times and

determining the standard deviation of each point in the solution vector. This should be the standard deviation of the artificial noise added to the data, scaled by the solution vector. Because the solution vector was chosen so that each value is identical, the standard deviations should also be identical, and can simply be multiplied into the entire matrix in Eq. (6.24).

To determine the values of $\delta\mathbf{x}$, consider Figure 6.5. By employing similar triangles,

$$\begin{aligned} \frac{\delta x_{1,v}}{\delta x_{1,h}} &= \frac{v_{2,1}}{v_{1,1}} \;\rightarrow\; \delta x_{1,v} = \delta x_{1,h}\frac{v_{2,1}}{v_{1,1}} \\ \frac{\delta x_{2,v}}{\delta x_{2,h}} &= \frac{v_{2,2}}{v_{1,2}} \;\rightarrow\; \delta x_{2,v} = \delta x_{2,h}\frac{v_{2,2}}{v_{1,2}} \end{aligned} \tag{6.25}$$

If $\mathbf{B}^{\mathrm{T}}\mathbf{B}$ is written as $\mathbf{D}$, and substituting in Eq. (6.25), then along the vector labelled $\delta\mathbf{x}_1$ in Figure 6.5, Eq. (6.24) becomes

$$\Delta\chi^2 = \frac{1}{\sigma^2}\left(\mathbf{D}_{1,1}{\delta x_{1,h}}^2 + \mathbf{D}_{1,2}{\delta x_{1,h}}^2\left(\frac{v_{2,1}}{v_{1,1}}\right) + \mathbf{D}_{2,1}{\delta x_{1,h}}^2\left(\frac{v_{2,1}}{v_{1,1}}\right) + \mathbf{D}_{2,2}{\delta x_{1,h}}^2\left(\frac{v_{2,1}}{v_{1,1}}\right)^2\right) \tag{6.26}$$

where ${\delta x_{1,h}}^2$ can be isolated to give

$${\delta x_{1,h}}^2 = \sigma^2\Delta\chi^2\left(\mathbf{D}_{1,1} + \mathbf{D}_{1,2}\left(\frac{v_{2,1}}{v_{1,1}}\right) + \mathbf{D}_{2,1}\left(\frac{v_{2,1}}{v_{1,1}}\right) + \mathbf{D}_{2,2}\left(\frac{v_{2,1}}{v_{1,1}}\right)^2\right)^{-1} \tag{6.27}$$

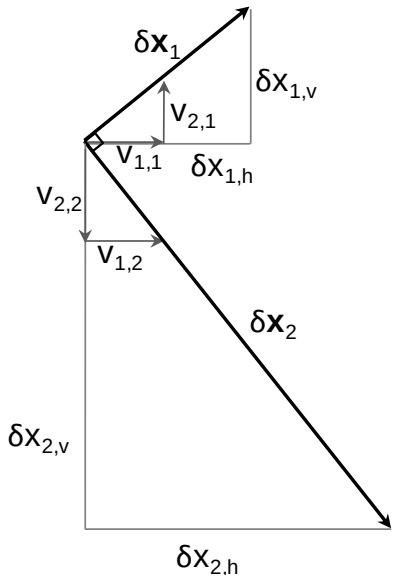

Figure 6.5 – Principle axes of a confidence interval

By solving Eq. (6.27) for 3% normally distributed relative error and specified confidence interval, the horizontal component of one principle axis of the ellipse is determined. By employing Eq. (6.25) the vertical component of the axis can also be calculated, and hence by Pythagorean Theorem the length of that principle axis of the ellipse is found. This process can then be repeated along the other axis, which completely defines the error ellipse. A plot of both the 90% and 99% confidence intervals is shown in Figure 6.6, derived with a 3% standard deviation in the measured data.

This simple example demonstrates that for a given confidence interval, $\Delta\chi^2$, the error ellipse can be reduced in size by maximizing the sensitivity matrix. That is, by making the matrix **B** larger, the matrix **D** is also made larger; hence the principle axes of confidence interval determined through Eq. (6.27) are shrunk.

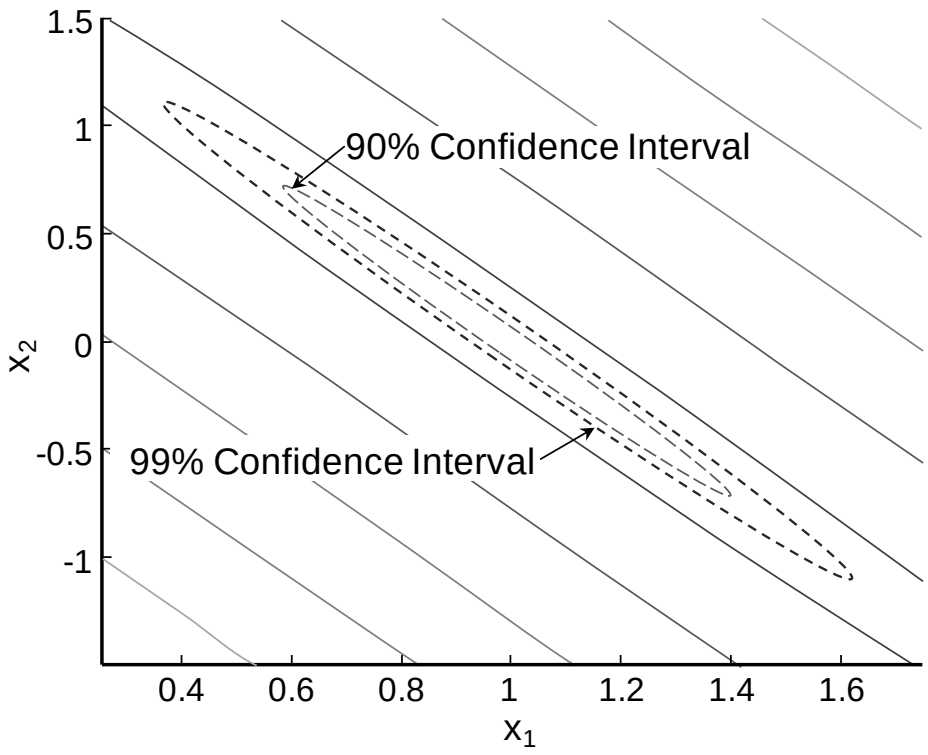

Figure 6.6 – 90% and 99% confidence intervals for the two-dimensional test case

6.4.2 Application to the Soot Aerosol System

The objective function given in Eq. (14) was applied to the soot aerosol system with twenty-three angles (based on the author's experimental work) using the nonlinear programming algorithm fmincon in Matlab to determine the best angle set at which light scattering measurements should be taken. This algorithm uses a sequential quadratic programming method, which estimates the Hessian of the Lagrangian to determine the step direction, and performs a line search to determine the step length. The objective function, however, is by no means smooth. By choosing values for twenty two of the angles, and varying the remaining angle, the objective function shown in Figure 6.7 is found (the angle chosen to vary was the twentieth). It is quite easy to see that around the angle in question the objective function is very noisy, and thus finding the true minimum using non-linear programming should prove difficult.

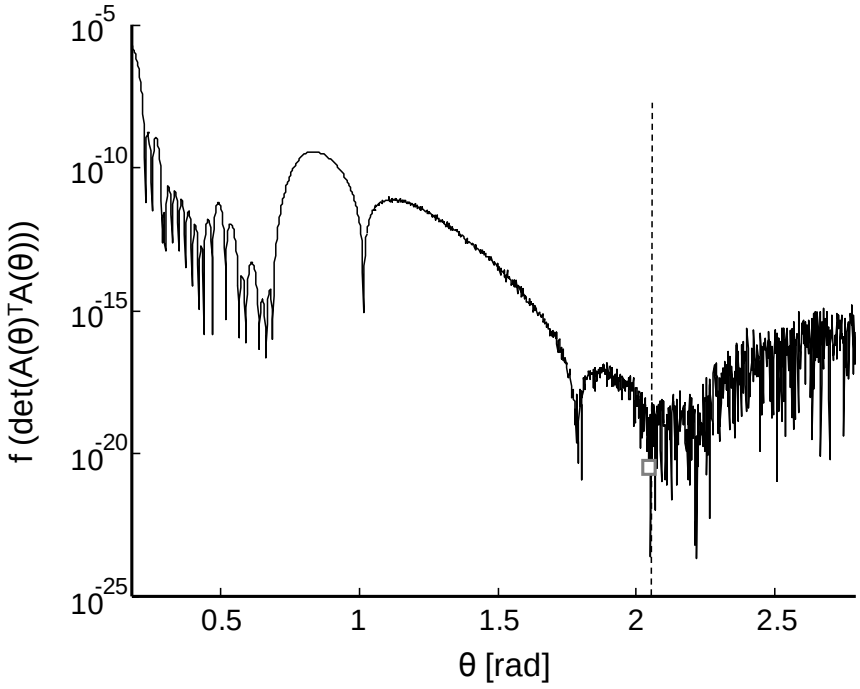

Figure 6.7 – Objective function, varying the twentieth angle over the acceptable range, with the original value of said angle noted

Due to the inherent noise in the objective function, another approach was adopted to determine the optimal angle set. Because nonlinear programming generally works by finding the best downhill direction to step in, it can potentially find a local minimum rather than the global minimum. Furthermore, it can find different local minima depending on the initial condition chosen. To overcome this, a simulated annealing procedure was applied to the problem.

When a metal is heated, its molecules have high energy, and are able to move quite freely. If the metal is then cooled quickly (quenched) the molecules drop down into the closest low-energy state, whatever that happens to be. The lowest energy state is one where every molecule is lined up in a crystal lattice, but in quenching the molecules form smaller sections of crystalline lattice that aren't all aligned properly, and are at a higher energy state than a perfect lattice. If the metal is instead allowed to cool slowly, the molecules have time to align

themselves, and they reach the lowest energy state. This occurs because the molecules are able to temporarily jump up to a higher energy state and then drop into an even lower energy state than the one they were in.

A simulated annealing program attempts to replicate this behaviour (Press, et al. 2007). Because the objective function is noisy, there are many local minima that are not the global minimum. Simulated annealing algorithms allow the objective to move in a direction other than a nearby minimum occasionally, if the move will bring it to a lower energy state. Thus, in places where nonlinear programming algorithms would get 'stuck' in a local minimum, the simulated annealing algorithm can continue optimizing. Using this algorithm, a new angle set was determined.

Three sets of angles are listed in Table 6.1: a set of angles that are uniform in θ; a set of angles that are uniform in q; and the set of angles found by applying simulated annealing. The singular values of the kernel matrix formed by each of these angle sets are shown in Figure 6.8. From Figure 6.3 we can see that the length of the principle axes of the error hyperellipse is a function of the reciprocal of the singular values. Thus, by considering the singular values shown in Figure 6.8 it can be seen that for the optimal angle set determined using simulated the principle axes of the error hyperellipse have been made smaller. As well, the angle set that is uniform in the q-domain produces a smaller error ellipse than the angle set that is uniform in the θ-domain.

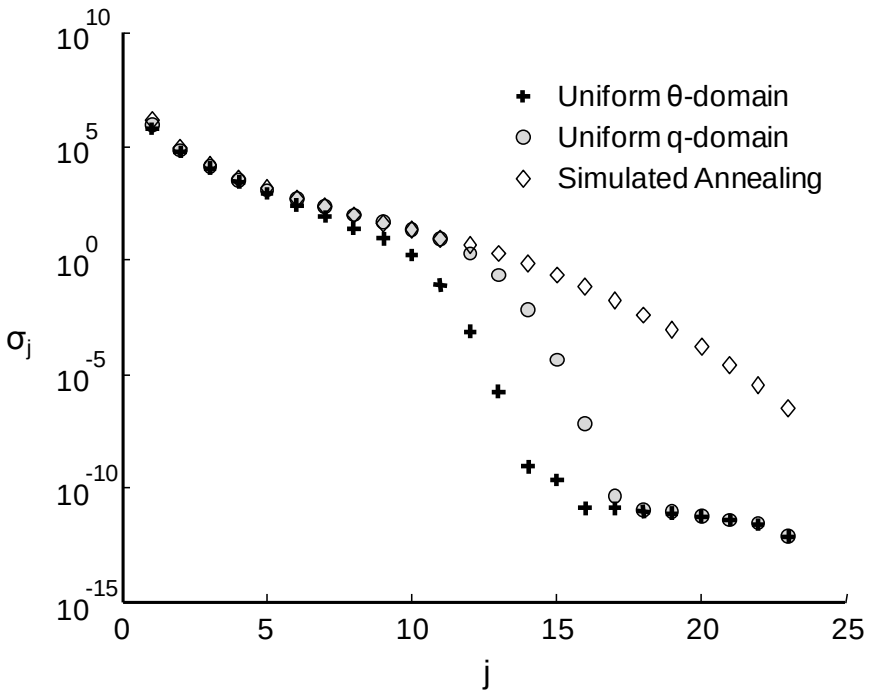

Figure 6.8 – Singular values of the kernel matrix using three angle sets

Figure 6.9 shows the plot of the relative intensity versus modulus of the scattering wave vector squared for the three angle sets, and shows interesting graphical features that may explain why certain angle sets are more optimal. The angles that are uniformly spaced in the θ-domain, shown in Figure 6.9 (a), and the q-domain, shown in Figure 6.9 (b), are clustered towards the large end of the curve. However, as mentioned in Section 2.2, the power-law regime should be a straight line when plotted in this way, and hence only requires two points to fully describe it. Therefore, a great number of points from these angle sets are not optimally placed, as they provide duplicate information to the problem. Likewise, there are few points in the transition regime, which suggests that this regime is not fully described and might therefore contain more information that could reduce the ill-posedness of the problem. It should be noted that the regime positioning in Figure 6.9 is not exact, as it is very difficult to properly determine where the transition regime is located.

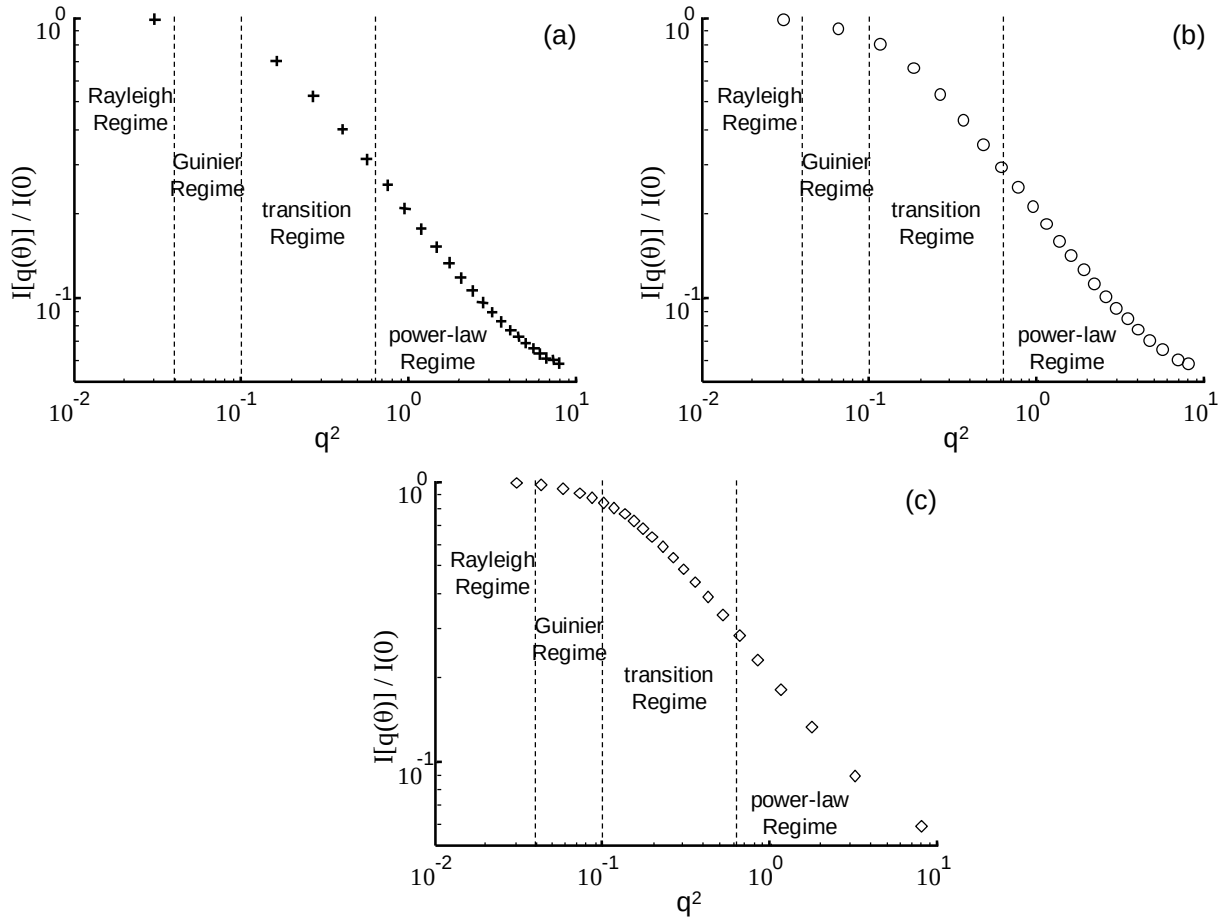

Figure 6.9 – Plot of angular scattering distributions: (a) uniform in θ-domain; (b) uniform in q-domain; (c) optimal

.

The optimal angle set and resulting relative scattering distribution is shown in Figure 6.9 (c). There are more points throughout the transition regime than in the uniform angle data sets, which will help gather as much information from that regime as possible. The Guinier regime is also a straight line when plotted in this manner, but its boundaries are difficult to determine, and any extra points in that regime are likely in the transition regime. The number of points in the power-law regime is far fewer than the other angle sets, but more than expected. This suggests that by having fewer angles the ill-conditioning of the matrix equation could be further reduced.

6.5 Reducing the Error

The point of finding an optimal angle set is to reduce the error inherent in the problem, hence reducing the need for regularization. There are two types of error that optimizing the problem should reduce: the perturbation error and the regularization error. If a perturbation analysis is done on Eq. (1.3), both of these error terms appear.

$$\begin{aligned} \mathbf{A}(\mathbf{x}^{\mathbf{ex}} - \delta\mathbf{x}) &= \mathbf{g}^{\mathbf{ex}} + \delta\mathbf{g} \\ \mathbf{x}^{\mathbf{ex}} - \delta\mathbf{x} &= \boldsymbol{A}^{\#}\mathbf{g}^{\mathbf{ex}} + \boldsymbol{A}^{\#}\delta\mathbf{g} \\ \delta\mathbf{x} &= (\mathbf{x}^{\mathbf{ex}} - \boldsymbol{A}^{\#}\mathbf{g}^{\mathbf{ex}}) - \boldsymbol{A}^{\#}\delta\mathbf{g} \end{aligned} \tag{6.28}$$

In this derivation, $\boldsymbol{A}^{\#}$ is the regularized pseudo-inverse of $\mathbf{A}$. The term $\boldsymbol{A}^{\#}\delta\mathbf{g}$ is defined to be the perturbation error, and the term $\mathbf{x}^{\mathbf{ex}} - \boldsymbol{A}^{\#}\mathbf{g}^{\mathbf{ex}}$ is defined to be the regularization error.

6.5.1 Perturbation Error

Determining whether the perturbation error is reduced by working with an optimized data set, a perturbation analysis is performed. In an artificial problem, the error in the data, $\delta\mathbf{g}$, is slowly increased, and the corresponding error in the solution, $\delta\mathbf{x}$, is plotted.

This was accomplished by forming three separate data sets, $\mathbf{g}$, using the three sets of angles given in Table 6.1. These angles were used, along with a known distribution, to numerically integrate Eq. (1.2). The kernel matrix was also formed for each of these angle sets.

The problem was first solved with no error at all, finding the least amount of regularization required. Then a small amount of error was added, and the problem solved again. This was repeated with increasing error until an arbitrary stopping point was reached. The results are shown in Figure 6.10.

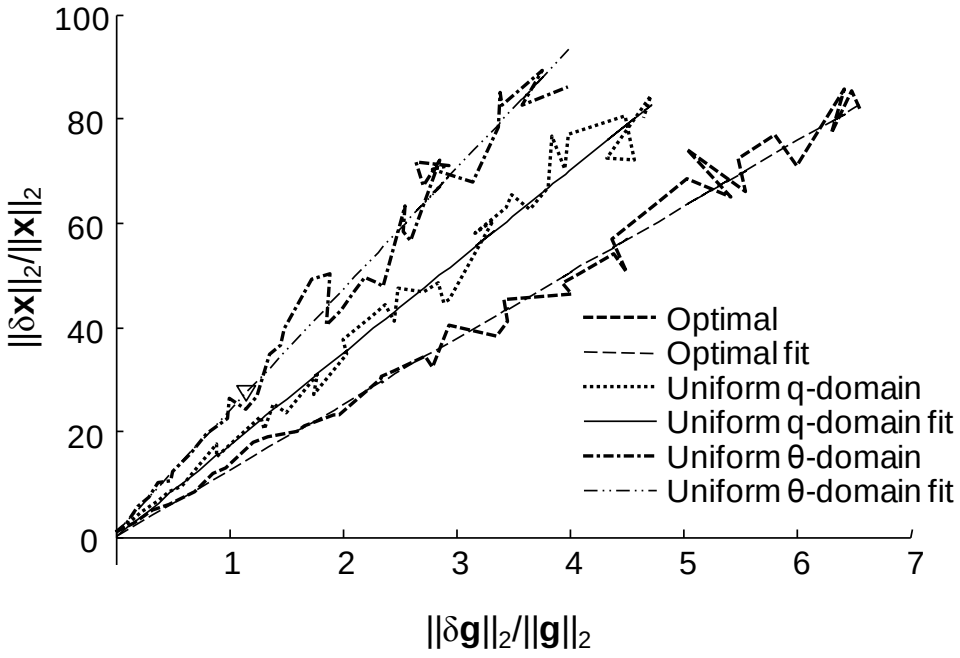

Figure 6.10– Perturbation analysis with linear fits

As expected, the magnitude of the error in the distribution, **x**, increases with error in the measured data, **g**. Employing the optimal angle set with a set amount of error in the data produces less error in the solution distribution than does the same amount of error added to the angle set that is uniform in the q-domain, which in turn has produces less error in the solution than does the data set with uniform angles in the θ-domain. Thus there is wisdom in picking angles in the q-domain over the θ-domain, as suggested by Sorensen (2001), but there is potentially an even better angle set.

Table 6.1– Angle sets used throughout this chapter.

Uniform in θ	Uniform in q-domain	Optimal
10.000	10.000	10.000
23.044	14.455	11.778
29.565	19.314	13.670
36.087	24.208	15.288
42.609	29.147	16.740
49.130	34.143	18.114
55.652	39.206	19.461
62.174	44.350	20.829
68.696	49.590	22.244
75.217	54.943	23.749
81.739	60.430	25.369
88.261	66.075	27.144
94.783	71.906	29.127
101.304	77.962	31.386
107.826	84.288	34.009
114.349	90.948	37.125
120.870	98.028	40.929
127.391	105.653	45.757
133.913	114.016	52.175
140.435	123.450	61.283
146.957	134.625	75.635
153.478	149.400	102.013
160.000	160.000	160.000

6.5.2 Regularization Error – No Noise in the Data

The next step is to prove that the regularization error is reduced for the optimal solution as well. This is done in two parts: first, with no error in the problem; and then with error added. Recalling the recovered solution shown in Figure 5.4, the soot aerosol problem cannot be inverted using normal algebraic methods, and as such some regularization must be added. To begin, regularization was added by employing truncated SVD (TSVD). This works by eliminating the small singular values and the corresponding basis vectors that cause highly oscillatory solutions. Singular values are systematically removed from the problem until the solution matches the exact solution, which is known because this is an artificial case.

This was done for each of the angle sets for a bimodal distribution. The purpose of using a bimodal distribution is that the distribution contains higher frequencies. Regularization tends to remove higher frequency solutions, and hence if the correct solution could be recovered, the procedure is very likely robust. For the case of the angle set that is uniform in the θ-domain the solution was recovered by truncating thirteen singular values out of the original twenty-three. For the angle set that is uniform in the q-domain only ten singular values were truncated to produce the solution. Finally, for the optimal angle set, only six singular values needed to be truncated to reproduce the original bimodal distribution. This is shown in Figure 6.11.

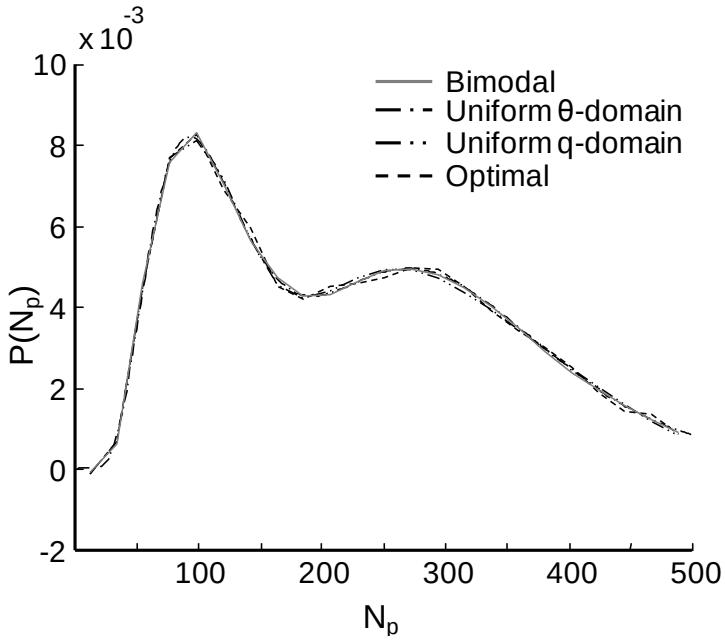

Figure 6.11 – Regularization analysis – recovering a bimodal distribution with TSVD

In determining the number of singular values that needed to be truncated, two things have been determined: first, that the optimal angle set does appear to reduce the amount of regularization required, and hence the regularization error, and second, that the level of truncation needed to invert the problem is decreased. This is important because standard Tikhonov regularization, using the identity matrix, can be written in the same manner as TSVD by using filter factors. For TSVD the filter factors are defined as

$$f_i = \begin{cases} 1, & i \leq k \\ 0, & i > k \end{cases} \tag{6.29}$$

where k is defined as the truncation parameter (Hansen 1998). For example, using the optimal data set $k = 23 - 6 = 17$. However, there is a definition for the filter factors for standard Tikhonov regularization, given as

$$f_i = \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \tag{6.30}$$

If the TSVD solution is known, then the regularization parameter for standard Tikhonov can be determined as $\lambda = \sigma_k$, and using this value the problem can be inverted. This may be somewhat of a convoluted method to solve the problem, but doing so gives a clear indication of the different levels of regularization required to recover a solution using the different angle sets. This is presented in Table 6.2. Thus, using the optimal angle set when no error is present in the problem reduces the regularization required to solve the problem.

Table 6.2 – Regularization required to recover a bimodal distribution with no error.

Angle Set	**Number of Singular Values Truncated**	**Regularization Parameter, λ**
Uniform in θ-domain	13	9.875
Uniform in q-domain	11	2.232
Optimal	7	0.0670

6.5.3 Regularization Error – With Noise in the Data

There is, of course, noise in the actual problem. Therefore noise must be added to the problem. This, unfortunately, complicates the problem greatly, as error in the problem increases the difficulty of recovering a solution, as seen in Chapter 5.

To determine if the optimal angle set still produced better results with noise present in the data, the perturbation error and the regularization error, originally defined in Eq. (6.28), were calculated for a range of λ. However, the actual expression used for the regularization error is given by Hansen (2010) as

$$x_{reg} = \sum_{i=1}^{n} \left(1 - f_i^{\lambda}\right)\left(\mathbf{v}_i^T \mathbf{x}^{exact}\right)\mathbf{v}_i \tag{6.31}$$

where f_i^{λ} are the filter factors for standard Tikhonov regularization for a given λ. Similarly, the perturbation error is defined as

$$x_{pert} = \sum_{i=1}^{n} f_i^{\lambda} \frac{\mathbf{u}_i^T \delta \mathbf{g}}{\sigma_i} \mathbf{v}_i \tag{6.32}$$

These values were determined over a range of λ's and plotted to produce a graph, shown in Figure 6.12. The total error in the problem is the sum of the regularization and perturbation errors, and the optimal amount of regularization is given at, or around, the value of the regularization parameter where both errors have the same value. This point has the smallest combined error, and hence is the value of the regularization parameter that should be employed in recovering the aggregate size distribution through Tikhonov regularization. This is shown graphically in Figure 6.12.

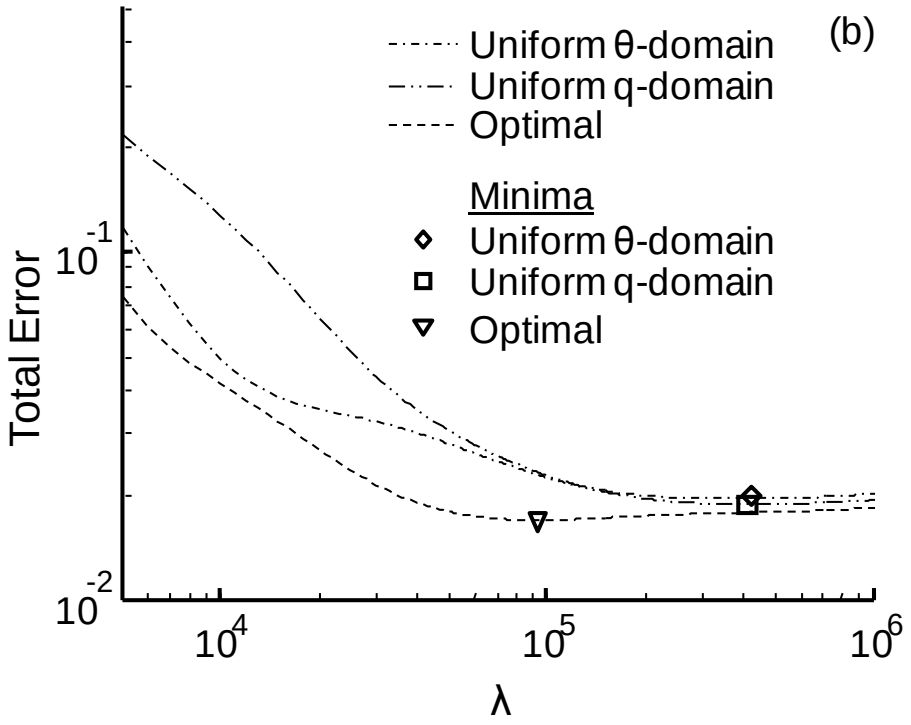

Figure 6.12 – Total Error – noisy data

This shows that the optimal angle set requires less regularization than the angle set that is uniform in the q-domain, which requires less regularization than the angle set that is uniform in the θ-domain. As well, although the difference is less than 10%, the total error in the solution also follows this trend, and hence the optimal angle set reduces the total error, although the reduction is minimal. Thus, optimizing the problem has had the desired effect, reducing the error inherent in the system and reducing the need for regularization. However, it should be noted that optimization of only the angles cannot reduce the need for regularization of the noise-free system to zero.

6.6 Summary

This procedure has shown that the conditioning of the matrix equation can be made better by the proper choice of measurement angles. Optimizing the matrix equation reduces the regularization error and the perturbation error, which in turn imply that less regularization is needed to properly invert the problem, and furthermore, error inherent in the data will produce marginally smaller error in the solution than with other data sets. This is important because the solution to the soot aerosol should depend on the original kernel matrix to as great an extent as possible, to reduce potential bias from the added priors; optimizing the measurement angles helps to achieve this. Hence, the optimization of this experimental setup has, theoretically, been shown to be a success. Obviously, the next step would be to test this experimentally.

Chapter 7
Conclusions

7.1 Summary of Results

The overall goal of this thesis was to use regularization to recover the soot aggregate size distribution from elastic light scattering data. This was shown to be possible, in theory, using an artificially generated data set. The resulting distribution matched the original distribution quite well for most of the range chosen, but in the small aggregate range the two distributions departed. This is due to the nature of the underlying physics, as the light scattered is proportional to the number of primary particles squared. Thus, small aggregates affect the light scattering measured at each angle very little, making it very difficult to recover the distribution properly in that range.

However, when put into practice with experimentally obtained data, the results were less than desirable. While the scaling constant was known for the artificial case, in the experimental case it was unknown and had to be solved for as a variable. This was done for standard Tikhonov regularization and it was discovered that there was no distinct optimal value for the scaling constant. Adding the nonnegativity constraint did nothing to improve the situation, and so a prescribed distribution was added to the regularization to properly recover the aggregate size distribution. However, the solution was quite good, compared to a distribution obtained by collecting and sizing thousands of aggregates via TEM analysis.

Optimization has also been applied to the problem, something that has not, to the best of the author's knowledge, been done before. It was shown that by optimizing the angles that light scattering is measured at, the ill-posedness of the problem can be reduced. However, the reduction in the ill-posedness is small, which is to say that the underlying system is so badly conditioned that the optimization has only a small effect.

7.2 Benefits of Proposed Method

There are substantial benefits to solving this problem using regularization. Unlike techniques that force a distribution, regularization only promotes a distribution, allowing the solution to vary from the prescribed distribution type if such a solution is better. This method also allows for the regularization to be dialled accordingly if specific information is known about the problem.

Furthermore, solving the problem by regularization provides a rigid mathematical framework in which to work. In the case when a distribution is forced to be lognormal, there is a long, narrow valley of solutions that all satisfy the problem. Finding the true minimal solution is very difficult if not near-impossible due to the error in the system. Even a small amount of error in the system can lead to a large amount of error in the recovered solution parameters, which can shift the solution far away from its true value. However, least-squares fitting, as is commonly used, does not have a rigorous method for pushing the solution towards the true solution in the long narrow valley of almost equally valid solutions.

Converting the system to a matrix equation, and looking at it in view of singular value decomposition provides a great deal of insight into the difficulties in actually recovering a solution. The singular values vary in size by up to sixteen orders of magnitude, showing that in

the forward problem certain basis vectors are important while in the inverse problem other vectors are important. This method allows for some optimization to be done to reduce the range of singular values and hence reduce the ill-posedness of the problem.

7.3 Recommendations for Future Work

The optimization performed on this system was successful, in that the error inherent in the problem was reduced. However, a great deal of work can still be done to improve on these results.

7.3.1 Experimental Work

First and foremost, the optimization should be experimentally validated. Though numerical analysis has shown that the optimal angles produced a better solution, it was outside of the power of the author to experimentally validate this. Hence, setting up the experiment and generating data for the optimal angle set and using that data to recover the soot aggregate size distribution should be carried out to determine if the optimization truly had any effect.

7.3.2 Other Optimization

While optimizing the angles proved effective numerically at reducing the ill-posedness of the problem, other variables could potentially be optimized, such as the laser wavelength. The wavelength of the laser is crucial in determining the positioning of the measurements in the q-domain; while the angular range is set by physical constraints, the wavelength could be varied to shift the range of values in the q-domain towards the smaller regimes and away from the power-law regime. Thus, an optimization algorithm could be designed to optimize the wavelength of the laser and the angle set simultaneously, allowing for the best possible experimental setup to be determined.

Another potential avenue of optimization could be determining the range and placement of the aggregate sizes (N_p) that are used to generate the kernel matrix. While initial optimization on this set of sizes suggested that uniform was best, the work was by no means in-depth and future could potentially improve the conditioning of the kernel matrix by means of optimizing its columns.

7.3.3 Other Flames

Obviously this research is in its early stages, and consideration for other flames, and other aerosols laden with soot, is still some way off. However, the end goal of this research is to produce a device that can accurately measure the aggregate size distribution from any soot laden aerosol by means of light scattering. Hence, expanding the work to include other techniques that are able to determine the myriad of parameters that are needed to recover the aggregate size distribution at the same time would be beneficial. For instance, the fractal dimension and primary particle diameter must both be known prior to attempting to recover the aggregate size distribution, and having a technique to determine those values using some form of measurement at the same time that light scattering is measured would potentially make the process more robust.

References

Agarwal, J K, and G J Sem. "Continuous Flow, Single-Particle-Countin Condensation Nucleus Counter." *Journal of Aerosol Science* 11 (1980): 343-357.

Andrews, H C, and B R Hunt. *Digital Image Restoration.* Eaglewood Cliffs, New Jersey: Prentice-Hall, Inc., 1977.

Baker, Christopher T. *The Numerical Treatment of Integral Equations.* Oxford: Clarendon Press, 1977.

Bayes, Thomas. "Essay Towards Solving a Problem in the Doctrine of Chances." *The Philosophical Transaction* 53 (1763): 370-418.

Bohren, Craig F, and Donald H Huffman. *Absorption and Scattering of Light by Small Particles.* New York: John Wiley & Sons, 1983.

Brasil, A M, T L Farias, and M G Carvalho. "A Recipe for Image Characterization of Fractal-Like Aggregates." *J. Aerosol Sci.,* 1999: 1379-1389.

"Carbon Black User's Guide." *International Carbon Black Association.* June 2004. http://www.carbon-black.org (accessed October 14, 2010).

Chapra, Steven C, and Raymond P Canale. *Numerical Methods For Engineers.* New York: McGraw-Hill, Inc., 1988.

Charalampopoulos, T T, and H Chang. "Effect of Soot Agglomeration on Radiative Transfer." *Journal of Quantitative Spectroscopy and Radiative Transfer*, 1991: 125-134.

Crump, James G, and John H Seinfeld. "Further Results on Inversion of Aerosol Size Distribution Data: Higher-Order Sobolev Spaces and Constraints." *Aerosol Science and Technology* 1, no. 4 (1982): 363-369.

D'Alessio, A, A Di Lorenzo, A Borghese, F Beretta, and S Masi. "Study of Soot Nucleation Zones of Rich Methane-Oxygen Flames." *Sixteenth Symposium (International) on Combustion*, 1977: 695-708.

D'Alessio, A, A Di Lorenzo, A F Sarofim, F Beretta, S Masi, and C Venitozzi. "Soot Formation in Methane-Oxygen Flames." *Fifteenth Symposium (International) on Combustion* 15, no. 1 (1975): 1427-1438.

Debye. "A Photoelectric Instrument for Light Scattering Measurements and a Differential Refractometer." *Journal of Applied Physics* 17 (1946): 392-398.

Debye, P. "Light Scattering in Solutions." *Journal of Applied Physics*, 1944: 338-342.

Debye, Peter P, and A M Bueche. "Scattering by an Inhomogeneous Solid." *Journal of Applied Physics* 20 (1949): 518-525.

di Stasio, S, J B' Mitchell, J L LeGarrec, L Biennier, and M Wulff. "Synchrotron SAXS <in situ> identification of three different size modes for soot nanoparticles in a diffusion flame." *Carbon* 44 (2006): 1267-1279.

Dobbins, R A, and C M Megaridis. "Morphology of Flame-Generated Soot As Determined by Thermophoretic Sampling." *Langmuir*, 1987: 254-259.

Dobbins, Richard A, and Constantine M Megaridis. "Absorption and scattering of light by polydisperse aggregates." *Applied Optics,* 1991: 4747-4754.

Dobbins, Richard A, and G Stephen Jizmagian. "Optical Scattering Cross Sections for Polydispersions of Dielectrix Spheres." *Journal of the Optical Society of America* 56 (1966): 1345-1350.

Emery, A F, and T D Fadale. "Design of Experiments Using Uncertainty Information." *Transactions of the ASME,* 1996: 532-538.

Engelhard, H, and H Friess. "Über die Brauchbarkeit des Mieeffektes zur Teilchengrößenbestimmung weißer Aerosole, insbesondere grobdisperser weißer Aerosole." *Kolloid-Z* 44 (1949): 245.

Eymet, V, A M Brasil, M El Hafi, T L Farias, and P J Coelho. "Numerical investigation of the effect of soot aggregation on the radiative properties in the infrared region and radiative transfer." *Journal of Quantitative Spectroscopy & Radiative Transfer,* 2002: 697-718.

Farias, T L, M G Carvalho, Ü Ö Köylü, and G M Faeth. "Computational Evaluation of Approximate Rayleigh-Debye-Gans/Fractal-Aggregate Theory for the Absorption and Scattering Properties of Soot." *Journal of Heat Transfer,* 1995: 152-159.

Farias, T L, Ü Ö Köylü, and M G Carvalho. "Range of validity of the Rayleigh-Debye-Gans theory for optics of fractal aggregates." *Applied Optics,* 1996: 6560-6567.

Folland, C K, T N Palmer, and D E Parker. "Sahel rainfall and worldwide sea temperatures, 1901-85." *Nature* 320, no. 17 (1986): 602-607.

Forrest, S. R., and T. A. Witten. "Long-range correlations in smoke-particle aggregates." *Journal of Physics A,* 1979: L109-L117.

Freltoft, T, J K Kjems, and S K Sinha. "Power-law correlations and finite-size effects in silica particle aggregates studied by small-angle neutron scattering." *The American Physical Society*, 1986: 269-275.

Gans, R. *Handbuch der Experimental Physik.* Leipzig: Akademische Verlagsgesellschaft, 1928.

Granqvist, C G, and R A Buhrman. "Ultrafine metal particles." *Journal of Applied Physics* 47, no. 5 (1976): 2200-2219.

Grassl, H. "Determination of Aerosol Size Distributions from Spectral Attenuation Measurements." *Applied Optics* 10, no. 11 (1971): 2534-2538.

Guinier, André, Gérard Fournet, Christopher B Walker, and Kenneth L Yudowitch. *Small-Angle Scattering of X-Rays.* New York: John Wiley & Sons, Inc., 1955.

Hadamard, Jacques. *Lectures on Cauchy's Problem in Linear Partial Differential Equations.* New York: Yale University Press, 1923.

Hansen, Per Christian. *Discrete Inverse Problems: Insight and Algorithms.* Philadelphia, PA: SIAM, 2010.

—. *Rank-Deficient and Discrete Ill-Posed Problems.* Philadelphia, PA: SIAM, 1998.

Hansen, Per Christian, and D P O'Leary. "The use of the L-curve in the regularization of discrete ill-posed problems." *SIAM Journal of Science and Computing* 14 (1993): 1487–1503.

Iyer, Suresh S, Thomas A Litzinger, Seong-Young Lee, and Robert J Santoro. "Determination of soot scattering coefficient from extinction and three-angle scattering in a laminar diffusion flame." *Combustion and Flame* 149 (2007): 206-216.

Jacobson, Mark Z. "Control of fossil-fuel particulate black carbon and organic matter,." *Journal of Geophysical Research* 107, no. D19 (2002): ACH 16-1:22.

Jagoda, I J, G Prado, and J Lahaye. "An Experimental Investigation into Soot Formation and Distribution in Polymer Diffusion Flames." *Combustion and Flame* 37 (1980): 261-274.

Johnson, Valen E. "A Bayesian χ^2 Test for Goodness-of-Fit." *The Annals of Statistics* 32, no. 6 (2004): 2361-2384.

Kline, Morris. *Calculus: An Intuitive and Physical Approach.* Mineola, New York: John Wiley & Sons, 1967.

Kolmogorov, A N. "Sulla determinazione empirica di une legge di distribuzion." *Giorn. dell'Istit. degli att* 4 (1933): 83-9.

Köylü, Ü Ö, and G M Faeth. "Optical Properties of Overfire Soot in Buoyant Turbulent Diffusion Flames at Long Residence Times." *Journal of Heat Transfer,* 1994: 152-159.

Köylü, Ü Ö, C S McEnally, D E Rosner, and L D Pfefferle. "Simultaneous Measurement of Soot Volume Fraction and Paricle Size/Microstructure in Flames Using a Thermophoretic Sampling Technique." *Combustion and Flame* 110 (1997): 494-507.

Köylü, Ü Ö, G M Faeth, T L Farias, and M G Carvalho. "Fractal and Projected Structure Properties of Soot Aggregates." *Combustion and Flame* 100 (1995): 621-633.

Ku, J C, and K -H Shim. "Optical Diagnostics and Radiative Properties of Simulated Soot Agglomerates." *Journal of Heat Transfer*, 1991: 953-958.

LaMer, V K, and D Sinclair. "Office of Scientific Research and Development, Reports No. 1668 and 1857." *Chemical Reviews* 44 (1949): 245.

Lin, M Y, H M Lindsay, D A Weitz, R C Ball, R Klein, and P Meakin. "Universality of Fractal Aggregates as Probed by Light Scattering." *Proceedings of the Royal Society of London.*, 1989: 71-87.

Link, O, D R Snelling, K A Thomson, and G J Smallwood. "Development of absolute intensity multi-angle light scattering for the determination of polydisperse soot aggregate properties." Beijing, China, 2010.

Liu, Fengshan, Kevin A Thomson, and Gregory J Smallwood. "Effects of soot absorption and scattering on LII intensities in laminar coflow diffusion flames." *Journal of Quantitative Spectroscopy and Radiative Transfer* 109 (2008): 337-348.

Mandelbrot, Benoit B. *Fractals: Form, Chance and Dimension.* San Francisco: W.H. Freeman, 1977.

Markowski, G R. "Improving Twomey's Algorithm for Inversion of Aerosol Measurement Data." *Aerosol Science and Technology* 7 (1987): 127-141.

Martin, James E, Dale W Schaefer, and Alan J Hurd. "Fractal geometry of vapor-phase aggregates." *Physical Review A* 33 (May 1986): 3540-3543.

Mie, Gustov. "Contributions to the Optics of Turbid Media, Particularly of Colloid Metal Solutions." *Annalen der Physik* 25, no. 3 (1908): 377-445.

Modest, Michael F. *Radiative Heat Transfer.* New York: McGraw-Hill, Inc., 1993.

Morozov, V A. "On the solution of functional equations by the." *Soviet Mathematics* 7 (1966): 414–417.

Mountain, Raymond D, and G W Mulholland. "Light Scattering from Simulated Smoke Aggregates." *Langmuir* 4 (1988): 1321-1326.

Oltmann, Hergen, Jörg Reimann, and Stefan Will. "Wide-angle light scattering (WALS) for soot aggregate characterization." *Combustion and Flame*, 2009.

Özisik, M N, and H RB Orlande. *Inverse Heat Transfer.* New York: Taylor and Francis, 2000.

Patel, A, and B Ch Khanduala. "A study of an infected bone tissue (osteomyelitis) by small-angle X-ray scattering." *Journal of Biochemical and Biophysical Methods* 9 (1984): 95-102.

Phillips, David L. "A Technique for the Numerical Solution of Certain Integral Equations of the First Kind." *Journal of the ACM* 9, no. 1 (1962): 84-97.

Pope, C. Arden, and Douglas W. Dockery. "Health Effects of Fine Particulate Air Pollution: Lines That Connect." *Journal of Air and Waste Management Association*, 2006: 709-742.

Press, W H, S A Teukolsky, W T Vetterling, and B P Flannery. *Numerical Recipes: The Art of Scientific Computing.* 3rd. New York: Cambridge University Press, 2007.

Puri, R, T F Richardson, R J Santoro, and R A Dobbins. "Aerosol Dynamic Processes of Soot Aggregates in a Laminar Ethene Diffusion Flame." *Combustion and Flame* 92 (1993): 320-333.

Reimann, J, S -A Kuhlmann, and S Will. "2D aggregate sizing by combining laser-induced incandescence (LII) and elastic light scattering (ELS)." *Applied Physics B* 96 (2009): 583-592.

Rosenfeld, Azriel, and Avinash C Kak. *Digital Picture Processing.* New York: Academic Press, 1982.

Schaefer, Dale W., James E. Martin, Pierre Wiltzius, and David S. Cannell. "Fractal Geometry of Colloid Aggregates." *Physical Review Letters*, 1984: 2371-2375.

Shaddix, Christopher R., and Timothy C. Williams. "Soot: Giver and Taker of Light." *American Scientist*, 2007: 232-239.

Siegel, Robert, and John R Howell. *Thermal Radiation Heat Transfer.* 4th Edition. New York, New York: Taylor & Francis, 2002.

Snelling, David R, Gregory J Smallwood, Fengshan Liu, Ömer L Gülder, and William D Bachalo. "A calibration-independent laser-induced incandescence technique for soot measurement by detecting absolute light intensity." *Applied Optics* 44 (2005): 6773-6785.

Snelling, David R, Kevin A Thomson, Greg J Smallwood, and Ömer L Gülder. "Two-Dimensional Imaging of Soot Volume Fraction in Laminar Diffusion Flames." *Applied Optics* 38, no. 12 (1999): 2478-2485.

Sorensen, C M. "Light Scattering by Fractal Aggregates: A Review." *Aerosol Science and Technology,* 2001: 648-687.

Sorensen, C M, N Lu, and J Cai. "Fractal Cluster Size Distribution Measurement Using Static Light Scattering." *Journal of Colloid and Interface Science* 174 (1995): 456-460.

Stanhill, Gerald, and Shabtai Cohen. "Global dimming: a review of the evidence for a widespread and significant reduction in global radiation with discussion of its probable causes and possible agricultural consequences." *Agricultural and Forest Meterorology* 107 (2001): 255-278.

Svergun, D I, A V Semenyuk, and L A Feigin. "Small-angle-scattering-data treatment by the regularization method." *Acta Crystallographica A* 44 (1988): 244-50.

Sztucki, M, T Narayanan, and G Beaucage. "In Situ study of aggregation of soot particles in an acetylene flame by small angle x-ray scattering." *Journal of Applied Physics*, 2007: 114304:1-7.

Tabibian, Richard M, Wilfired Heller, and Joseph N Epel. "Experimental Investigations on the Light Scattering of Colloidal Spheres." *Journal of Colloid Science* 11 (1956): 195-213.

Thomson, Kevin Austin. *Soot temperature measurements by spectral emission methods in a laminar diffusion flame.* Waterloo, Ontario: University of Waterloo, 2000.

Tian, K, Kevin A Thomson, Fengshan Liu, David R Snelling, Greg J Smallwood, and D Wang. "Distribution of the Number of Primary Particles of Soot Aggregates in a Nonpremixed Laminar Flame." *Combustion and Flame* 138 (2004): 195-198.

Tikhonov, Andrey N, and Vasiliy Y Arsenin. *Solutions of Ill-Posed Problems.* Washington, D.C.: V.H. Winston & Sons, 1977.

Tikhonov, Andrey Nikolayevich. "О решении некорректно поставленных задач и методе регуляризации [Solution of incorrectly formulated problems and the regularization method]." *Proceedings of the USSR Academy of Sciences* 151 (1963): 501–504.

Tikhonov, Andrey Nikolayevich. "Об устойчивости обратных задач [On the stability of inverse problems]." *Proceedings of the USSR Academy of Sciences* 39, no. 5 (1943): 195–198.

Twomey, S. "Comparison of Constrained Linear Inversion and an Iterative Nonlinear Algorithm Applied to the Indirect Estimation of Particle Size Distributions." *Journal of Computational Physics*, 1975: 188-200.

Twomey, S. "On the Numerical Solution of Fredholm Integral Equations of the First Kind by the Inversion of the Linear System Produced by Quadrature." *Journal of the ACM*, 1963: 97-101.

Tyndall, John. "On the Blue Colour of the Sky, the Polarization of Skylight, and on the Polarization of Light by Cloudy Matter Generally." *Proceedings of the Royal Society of London* 17 (1869): 223-233.

van de Hulst, H C. *Light Scattering by Small Particles.* New York: John Wiley & Sons, 1981.

Wada, Nobuhiko. "Preparation of Fine Metal Particles by Means of Evaporation in Helium Gas." *Japanese Journal of Applied Physics* 6, no. 5 (1967): 553-556.

Weiz, D. A., and M. Oliveria. "Fractal Structures Formed by Kinetic Aggregation of Aqueous Gold Colloids." *Physical Review Letters,* 1984: 1433-1437.

Wild, Martin. "Global dimming and brightening: A review." *Journal of Geophysical Research* 114 (2009): D00D16.

Yang, Bo, and Ümit Ö Köylü. "Soot processes in a strongly radiating turbulent flame from laser scattering/extinction experiments." *Journal of Quantitative Spectroscopy & Radiative Transfer* 93 (2004): 289-299.

Zellner, Arnold. *An Introduction to Bayesian Inference in Econometrics.* New York: John Wiley & Sons, 1971.

Appendix A: MATLAB Code

This section contains the MATLAB code that was used in preparing for and writing this thesis.

A.1 Lgnrm.m

```
function Logi = Lgnrm(sigma,mu,NP)
% This calculates the value of the lognormal function at discrete points,
% given by NP

% Inputs
% ---------
% sigma - standard deviation (or log of the width) of the distribution
% mu - log of the geometric mean
% NP - aggregate size under consideration (can be either a vector or a
%   single point)

% if we've defined this badly, the initial value of NP is zero, and
% anything/0 is a really bad thing, so I added a line to just make the
% value zero.
if NP(1) == 0
    NP(1) = 100;
    Logi = exp(-1*(log(NP)- mu).^2./(2*sigma^2))./(NP.*sigma.*sqrt(2.*pi()));
    Logi(1) = 0;
else
   % this is what the expression looks like if we don't consider NP = 0
    Logi = exp(-1*(log(NP)- mu).^2./(2*sigma^2))./(NP.*sigma.*sqrt(2.*pi()));
end
end
```

A.2 Kernel3.m

```
function Phase = Kernel3(np,theta,scal)
% This function determines the value of the kernel function

% Inputs
% -----------
% np - the aggregate size in question
% theta - the angle under consideration
% scal - do we consider the sample volume in the kernel? Scal = 1 -> no

% Df - fractal dimension
% kg - fractal prefactor
% a - primary particle diameter (or radius?)
% lambda - laser wavelength
Df = 1.72; kg = 2.4; a = 29; lambda = 527;
% eta -  wavenumber
eta = 2*pi()/lambda;
```

```
%rg - radius of gyration
rg = a*((np/kg)^(1/Df))/2;
% q - scattering wave vector
q = 2*eta*sin(theta/2);
% this one is self-explanatory
qrg = q*rg;
%fqrg - form factor or structure factor (taken from Koylu and Faeth
fqrg = (1 + 8*(qrg).^2./(3*Df) + (qrg).^8).^(-Df/8);
%Phase - Kernel Function
if scal == 1
    Phase = np^2*fqrg;
else
    Phase = np^2*fqrg/sin(theta);
end
end
```

A.3 miniquad.m

```
function Ker = miniquad(npmax,npmin,theta,scal,delnp)
% This is a simple 6-point Gauss quadrature that integrates a strip around
% a given NP value
 % Inputs
% ----------
% npmax is the upper bound of the strip.
% npmin is the lower bound of the strip.
% theta is the angle at which we're looking.
% scal determines if we consider sample volume in the kernel: scal = 1 ->
%   no, we don't.
% delnp is the width of the strip.

% Gaussian weighting factors
C = [0.171324492, 0.360761573, 0.467913935, 0.467913935, 0.360761573, ...
    0.171324492]';
% NP values, chosen by Gaussian quadrature
NP(1) = ((npmax + npmin) + (npmax-npmin)*(-0.932469514))/2;
NP(2) = ((npmax + npmin) + (npmax-npmin)*(-0.661209386))/2;
NP(3) = ((npmax + npmin) + (npmax-npmin)*(-0.238619186))/2;
NP(4) = ((npmax + npmin) + (npmax-npmin)*(0.238619186))/2;
NP(5) = ((npmax + npmin) + (npmax-npmin)*(0.661209386))/2;
NP(6) = ((npmax + npmin) + (npmax-npmin)*(0.932469514))/2;
% Pre-allocating
Phi = zeros(6,1);
% Determining the Kernel function at each of the above six points.
for i = 1:6
    Phi(i) = Kernel3(NP(i),theta,scal);
end
% Gauss quadrature gives that dNP = (max_NP - min_NP)*dNP_n/2
Lgn = Phi*(npmax-npmin)/2;
% Multiplying by the weighting factors.
Lgn2 = Lgn.*C;
% Summing to get the integral over the strip centred at NP.
Ker = sum(Lgn2)/delnp;
end
```

A.4 gausste5.m

```
function Integ2 = gausste5(logs,mnp,q1,sigma,mu,scal)
% This function takes in a number of parameters, and returns a numerically
% integrated integral equation, defined as the integration of a kernel
% function (dealt with in the function miniquad.m) multiplied by an initial
% distribution (given by logs).

% Inputs
% ------------
% logs - the shape you've chosen as your initial distribution. For
%   instance, if we choose lognormal, logs = @Lgnrm
% mnp - maximum NP, or the upper limit to the dependent variable you're
%   numerically integrating
% q1 - the set of angles you're working with
% sigma - standard deviation (or log of the width) of the distribution
% mu - log of the geometric mean
% scal - determining if we work the sample volume in to the data or the
%   kernel (or either). Scal = 1 -> sample volume dealt with in data, and not
%   included in the kernel.

format long
% Num is the number of angles currently being used
num = length(q1);
% Letting the number of intervals the integral is split up into start at 1
n=2;
% Initial difference set so it will go past the first iteration:
dif = 1000;
% Error - this can be changed depending on how accurate an integration is
% required
epsilon = 1e-5;

%**********************************************************************
% Gaussian Weighting Factors
C = [0.171324492, 0.360761573, 0.467913935, 0.467913935, 0.360761573,...
    0.171324492]';
%**********************************************************************
% Pre-allocating size of some vectors for speed.
Integ3 = ones(1,num);
NP = zeros(6,1);
Integ = zeros(n,1);
Phi = zeros(6,num);
%**********************************************************************
% The main while loop
while dif > epsilon && n<1e5
     % This for loop goes through all the different strips we have split this
   % function in to.
     for j=2:n
          % Here we find the points, for each interval chosen, that are used
      % by the six-point Gaussian Quadrature. j defines the interval in
      % question. On a plot, these are the x-axis points
          NP(1) = (mnp/n*j + mnp/n*(j-1))/2 + (mnp/n*j - mnp/n*(j-1))/2*...
               (-0.932469514);
          NP(2) = (mnp/n*j + mnp/n*(j-1))/2 + (mnp/n*j - mnp/n*(j-1))/2*...
```

```
            (-0.661209386);
        NP(3) = (mnp/n*j + mnp/n*(j-1))/2 + (mnp/n*j - mnp/n*(j-1))/2*...
            (-0.238619186);
        NP(4) = (mnp/n*j + mnp/n*(j-1))/2 + (mnp/n*j - mnp/n*(j-1))/2*...
            (0.238619186);
        NP(5) = (mnp/n*j + mnp/n*(j-1))/2 + (mnp/n*j - mnp/n*(j-1))/2*...
            (0.661209386);
        NP(6) = (mnp/n*j + mnp/n*(j-1))/2 + (mnp/n*j - mnp/n*(j-1))/2*...
            (0.932469514);
    % The distance between successive values of NP
        delnp = NP(2)-NP(1);
    % Reallocation
        L2 = zeros(6,num);
        L3 = zeros(6,num);
    % This for loop goes through the num different angles
        for i=1:num
      % This loop does a mini Gaussian integration, also a six-point
      % Gauss quadrature, at each value of NP for more accuracy
            for k = 1:6
                Phi(k,i) = miniquad(NP(k)+delnp/2,NP(k)-
delnp/2,q1(i),scal,delnp);
            end
        % The value of the integral at each of the six points over the
    % current strip
        L2(:,i) = Phi(:,i).*logs(sigma,mu,NP).*(mnp*j/n - mnp*(j-1)/n)./2;
    % Multiplying the Gaussian weights by the values of the integral
    % function
        L3(:,i) = L2(:,i) .* C;
    % Summing up the six points we have calculated.
        Integ(j,i) = sum(L3(:,i));
        end
    end
    % More reallocating.
    Integ2 = zeros(1,num);
    dif1 = zeros(1,num);
    % This for loop goes through each angle, and sums up all the different
  % strips we've integrated.
    for i=1:num
        Integ2(1,i) = sum(Integ(:,i));
    end
    % Here we define our difference, or the residual between our current
  % integral and the previous one. This is divided by the value of the
  % previous integral to give the relative residual, rather than the
  % absolute. As well, we take the absolute value in the off chance that
  % the previous integral is bigger than the current value, giving a
  % negative residual and ending the loop early.
    dif1(1,i) = abs((Integ2(1,i) - Integ3(1,i))/Integ3(1,i));
    % Setting the previous integral equal to the current integral, for the
  % next step.
    Integ3 = Integ2;
    % We double n for the next step. This is faster than a linear increase,
  % and if twice as many strips doesn't give a significantly better
  % result, then we've likely integrated well enough.
    n = 2*n;
  % Displaying n so that we can keep track of where we are.
    disp(n);
```

```
   % We take the average of the differences calculated above, and use that
   % to determine if we've reached convergence. Some parts of the
   % integration may have, but we need the entire integral to do so.
     dif = sum(dif1(1,i))/num;
   % We print this so that we can see something happening, and determine
   % how close it is to convergence
     disp(dif);
end
% Finally, set it equal to it's transpose so that it is a column vector,
% and return it as the answer. This could of course be done much earlier.
Integ2 = Integ2';
end
```

A.5 Ker.m

```
function Phase = Ker(deltanp,qb,num2,samp,NP)
% calculating a full kernel matrix

% Inputs
% ----------
% deltanp - the distance between the discrete NP values
% qb - the angles under consideration
% num2 - the number of discrete aggregate sizes (NP's) under consideration
%      - num2 = length(NP)
% samp - is the sample volume accounted for in the kernel? samp = 1 -> no
% NP - the discrete aggregate sizes under consideration

% The Kernel matrix
%---------------------------
% The number of angles under consideration
num = length(qb);
% Df - fractal dimension
% kg - fractal prefactor
% a - primary particle diameter (or radius?)
% lambda - laser wavelength
Df = 1.72; kg = 2.4; a = 29; lambda = 527;
% eta -  wavenumber
eta = 2*pi()/lambda;
% Predefining variables of given sizes.
Phase = zeros(num,num2);
npab = zeros(6,1);
rg = zeros(6,1);
qrg = zeros(6,1);
fqrg = zeros(6,1);
Pha = zeros(6,1);
% Gaussian integration, six-point quadrature - points
xd =[-0.932469514,-0.661209386,-0.238619186,0.238619186,...
    0.661209386,0.932469514];
% Gaussian integration, six-point quadrature - weights
C = [0.171324492,0.360761573,0.467913935,0.467913935,0.360761573...
    ,0.171324492]';
%=========================================
```

```
% moving over the rows and columns of the Kernel matrix
for i = 1:num
    for j = 1:num2
    % the maximum and minimum number of primary particles in this
    % particular 'bin'
        Npminj = NP(j)-deltanp/2;
        Npmaxj = NP(j)+deltanp/2;
    % the scattering wave vector at the angle qb(i)
        q = 2*eta*sin(qb(i)/2);
    % Another integration, this time over the strip of the function
    % that we're dealing with.
        for k = 1:6
      % This is the number of primary particles that we are currently
      % dealing with, as determined by Gaussian quadrature
            npab(k) = ((Npmaxj + Npminj) + (Npmaxj-Npminj)*xd(k))/2;
      % rg - radius of gyration
            rg(k) = a*((npab(k)/kg)^(1/Df))/2;
      % qRg - self explanatory
            qrg(k) = q*rg(k);
      % fqrg - form factor or structure factor: from Koylu
            fqrg(k) = (1 + 8*(qrg(k)).^2./(3*Df) + (qrg(k)).^8).^(-Df/8);
      % This is the value of the kernel matrix at a specific point in
      % the strip
            if samp == 1
                Pha(k) = npab(k)^2 * fqrg(k)*(Npmaxj-Npminj)/2;
            else
                Pha(k) = npab(k)^2 * fqrg(k)/sin(qb(i))*(Npmaxj-Npminj)/2;
            end
        end
    % This is integrating over the strip, giving the value of the
    % kernel matrix at the current i and j.
        Phase(i,j) = Pha'*C/deltanp;
    end
end
end
```

A.6 hypergeo.m

```
function g = hypergeo(D,qRg)
    % calculating the hypergeometric structure factor, as suggested by
  % Sorensen (2001).

  % number of iterations to go through. Matlab can't handle much beyond
  % this number, for larger qRg values.
    nu = 170;
  % the three parameters that define the hypergeometric function
    a = D/2;
    b = 3/2;
    z = -((qRg)^2)/D;
  % reallocating
    g2 = zeros(nu,1);
  % two dummy variables that are the gamma functions of a and b
```

```
    braca2 = Gams(a);
    bracb2 = Gams(b);
  % This is the actual hypergeometric function
    for i = 1:nu
    % two more dummy variables that are the ratio of two gamma
    % functions
        braca = Gams(a+i)/braca2 ;
        bracb = Gams(b+i)/bracb2;
    % the ratio of the ratios multiplied by the other variable
        g2(i) = braca/bracb * z^i/factorial(i);
    end
  % the value of the hypergeometric function is the sum of all these
  % values plus one. Adding more values theoretically increases the
  % accuracy of the function.
    g = sum(g2) + 1;
end

function mul = Gams(tau)
  % This integrates a gamma function by using the fact that n! =
  % Gamma(n+1).
  % we want to calculate (n-1)! = Gamma(n)
    tau2 = tau - 1;
  % how many iterations do we go through before we call it quits
    p = 30000;
  % defining this as unity, as this is a multiplication term
    mul = 1;
  % and actually calculating the gamma function at n. If p is larger, it
  % becomes more accurate, but takes longer to compute.
    for k = 1:p
        mul = mul*((k+1)/k)^tau2*((k)/(tau2 + k));
    end
end
```

A.7 Kolmogorovfits.m

```
function dif = Kolmogorovfits(theta)
  % This is a simple implementation of the Kolmogorov-Smirnov goodness
  % -of-fit statistic, which finds the maximum difference between a given
  % distribution and a lognormal distribution defined by the values given
  % in theta.

  % loading the distribution, xd; the length of the distribution, num2d;
  % and the discrete vector of number of primary particles, NPd; I saved
  % these values in the worksheet I am calling from, using
  % save Oct6a xd num2d NPd
    load Oct6a xd num2d NPd
    x = xd;
    num2 = num2d;
    NP = NPd;

    %-----------------
```

```
    % Generating the initial lognormal distribution
    phi0 = Lgnrm(theta(1),theta(2),NP);
  % this is just in case the initial distribution is poorly defined.
    if phi0(1) == inf
        phi0(1) = 10^3;
    end
  % Preallocating
    cdfx = zeros(num2,1);
    cdfphi = zeros(num2,1);
  % The summation of the heights of each distribution, as a rough 'value
  % under the curve'
    sumx = sum(x);
    sumphi = sum(phi0);

  % converting to a Cumulative Distribution Function
    for i = 1:num2
        if i == 1
            cdfx(i) = x(i)/sumx;
            cdfphi(i) = phi0(i)/sumphi;
        else
            cdfx(i) = cdfx(i-1) + x(i)/sumx;
            cdfphi(i) = cdfphi(i-1) + phi0(i)/sumphi;
        end
    end
  % Finding the distance between the two graphs
    D1 = zeros(num2);
    for i = 1:num2
        D1(i) = abs(cdfx(i) - cdfphi(i));
    end
  % the Goodness of fit is the difference between the two.
    dif = max(D1(:));
end
```

Appendix B:

Numerical Algorithms

B.1 Numerical Integration

Equation (1.2) has no analytical solution; there is no simple way to determine the intensity of the light scattered at a specified angle for a given distribution of aggregates. To generate an artificial data set with a known distribution, a numerical integration scheme must be implemented. For the purposes of this work, the six-point Gauss quadrature scheme was chosen (Chapra and Canale 1988).

A function is integrated using this scheme by choosing six optimally chosen points on the interval [-1,1] and evaluating the function at each point. A change of variables must be carried out to transform the optimal points from the unit interval to the interval of interest, as follows

$$x = \frac{(a+b)+(a-b)x_d}{2} \tag{B.1.1}$$

where the value of x_d is the optimal integration point on the interval [-1,1], and the variables a and b are the limits of integration. As well, this change of variables requires that the change

$$dx = \frac{b-a}{2}dx_d \tag{B.1.2}$$

be carried out as well. These changes allow Eq. (1.2) to be rewritten as

$$\begin{aligned} g(\theta) &= C\int_1^{500} K(\theta, N_p)\,P(N_p)dN_p \\ &= C\int_{-1}^{1} K\left(\theta, \frac{(501)+(499)N_{p,d}}{2}\right) P\left(\frac{(501)+(499)N_{p,d}}{2}\right)\left(\frac{599}{2}\right)dN_{p,d} \end{aligned} \tag{B.1.3}$$

Once these substitutions are made, the integral is not evaluated, but rather the integrand is evaluated at each value of x_d and the resulting values are multiplied by a weighting factor and summed,

$$I \approx c_0 f(x_0) + c_1 f(x_1) + \cdots + c_5 f(x_5) \tag{B.1.4}$$

where $f(\cdot)$ is the integrand evaluated at the given point. The values of the optimal arguments, x_d, and the weighting factors, c_d, are given in Table B.1.

Table B.1 - Weighting factors and function arguments for Gauss six-point integration

Index (d)	Weighting Factor (c_d)	Optimal Argument (x_d)
0	0.171324492	-0.932469514
1	0.360761573	-0.661209386
2	0.467913935	-0.238619186
3	0.467913935	0.238619186
4	0.360761573	0.661209386
5	0.171324492	0.932469514

This process can be repeated on increasingly fine intervals until convergence between the integrated values at two consecutive steps is found. That is to say, the function is integrated over the entire interval using this procedure; the interval is then split in half and both halves are integrated and the values summed. If the variation between the two integration values is within a set tolerance the procedure is stopped; otherwise, the halves are split in half and each quarter is integrated. The intervals are continually split in half and the integrand integrated over all the intervals until convergence is met.

B.2 Kolmogorov-Smirnov Goodness-of-fit Statistic

The Kolmogorov-Smirnov goodness-of-fit statistic works by determining the maximum distance between a distribution of interest and a reference distribution. In the case of work carried out for this thesis, the reference distribution was a lognormal distribution. This works by computing the cumulative density function (CDF) of each distribution, defined as

$$\mathrm{CDF}\left(\mathrm{x}\left(\mathrm{N_p}\right)\right) = \int_{-\infty}^{\mathrm{N_p}} \mathrm{x}\left(\mathrm{N_p}\right)\mathrm{dN_p} \tag{B.2.1}$$

The CDF determines how much of the probability density function (PDF) has occurred up to a given $\mathrm{N_p}$. The CDF is used rather than the PDF because small blips away from the area of interest (for instance, at large $\mathrm{N_p}$) have less effect on the overall distribution than for a PDF.

In the discrete case, such as was used in this thesis, the CDF of the reference distribution and the distribution of interest were calculated at each of a set of discrete points, by

$$\mathrm{CDF}\left(\mathrm{x_p}\right) = \sum_{\mathrm{i=1}}^{\mathrm{p}} \mathrm{x_i} \Big/ \sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{x_i} \tag{B.2.2}$$

where p is the size of interest, with $p \leq n$.

Once the two CDF's are generated it is a simple thing to find the difference between them and then find the maximum difference, which is the goodness-of-fit parameter, $\mathrm{F_{K-S}}(\mathbf{x})$.

$$\mathrm{F_{K-S}}(\mathbf{x}) = \max\left|\mathrm{CDF}(\mathbf{x}) - \mathrm{CDF}(\mathbf{x}_{\mathrm{reference}})\right| \tag{B.2.3}$$

This process is depicted below in Figure B.0.1.

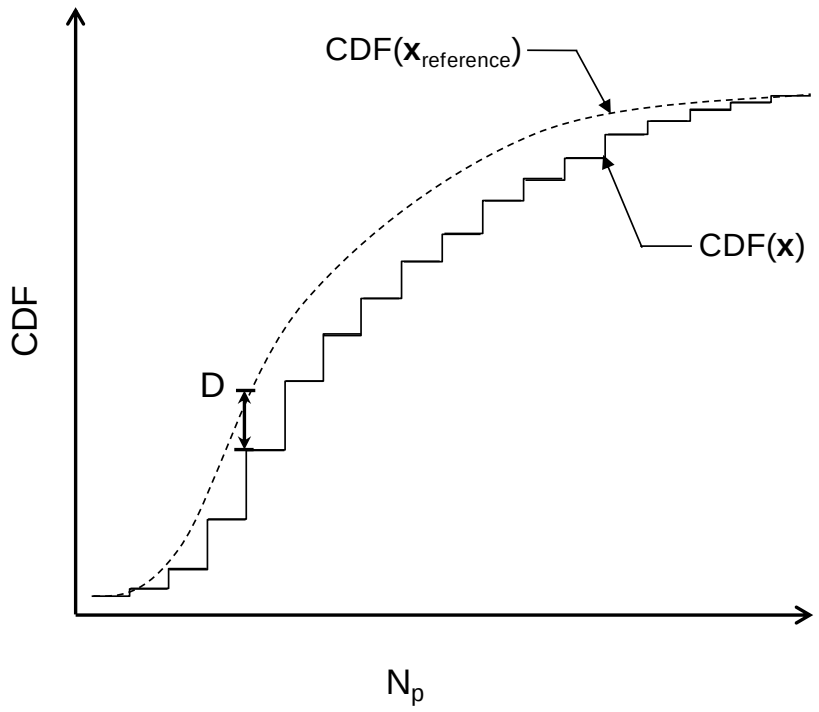

Figure B.0.1 - Fitting a reference lognormal CDF to the distribution of interest

B.3 The Scattering Albedo of Soot

Some consideration should be given to the scattering albedo, which appears in Eq. (3.3) in Section 273.2. The scattering albedo is a constant property of soot that depends on the underlying soot properties, including fractal dimension, D_f, number of primary particles, N_p, primary particle diameter, d_p, the wavelength of the incident radiation, λ, the complex index of refraction, $\boldsymbol{m}$, and others.

The scattering albedo is defined as

$$\Omega = \frac{\sigma_s}{a_\lambda + \sigma_s} \tag{B.3.1}$$

where σ_s is the scattering coefficient and a_λ is the absorption coefficient. From Modest (1993) the scattering and absorption coefficients are defined as

$$\begin{aligned} \sigma_s &= N_{agg}\bar{C}_s^{agg} \\ a_\lambda &= N_{agg}\bar{C}_a^{agg} \end{aligned} \tag{B.3.2}$$

where N_{agg} is the aggregate number density and $\bar{C}_s^{agg}$ and $\bar{C}_a^{agg}$ are the aggregate scattering and absorption cross sections, respectively.

From Eymet, et al. (2002) the cross sections are given as

$$\begin{aligned} C_s^{agg} &= N_p^2 C_s^p g\left(k, R_g, D_f\right) \\ C_a^{agg} &= N_p C_a^p \end{aligned} \tag{B.3.3}$$

for the aggregate cross sections and

$$C_s^p = \frac{8\pi x_p^6}{3\eta^2}\left|\frac{\boldsymbol{m}^2-1}{\boldsymbol{m}^2+2}\right|^2$$
$$C_a^p = \frac{4\pi x_p^3}{\eta^2}\,\mathrm{Im}\left(\frac{\boldsymbol{m}^2-1}{\boldsymbol{m}^2+2}\right) \tag{B.3.4}$$

for the primary particle cross sections. In these equations, $x_p \equiv \pi d_p/\lambda$ is the particle size parameter and $\eta \equiv 2\pi/\lambda$ is the wavenumber.

Finally, from Köylü and Faeth (1994), the aggregate total scattering factor, $g\left(k, R_g, D_f\right)$, is defined as

$$\begin{aligned} g\left(k, R_g, D_f\right) &= 1 - 2\left(\eta R_g\right)^2/3 \qquad \left(\eta R_g\right)^2 \le 3D_f/8 \\ g\left(\eta, R_g, D_f\right) &= \frac{\beta}{2}(3 - 3\beta + 2\beta^2) - \frac{\left(\eta R_g \beta\right)^2}{3}(3 - 4\beta + \beta^2) + \left(2\eta R_g\right)^{-D_f} \\ &* \left[\frac{3}{2 - D_f} - \frac{12}{\left(6 - D_f\right)\left(4 - D_f\right)} - 3\beta^{1-D_f/2}\right. \\ &* \left.\left(\frac{1}{2 - D_f} - \frac{2\beta}{4 - D_f} + \frac{2\beta^2}{6 - D_f}\right)\right] \qquad \left(\eta R_g\right)^2 > 3D_f/8 \end{aligned} \tag{B.3.5}$$

where $\beta = 3D_f/\left(8\eta^2 R_g^2\right)$ is the aggregate scattering parameter and R_g is the aggregate radius of gyration.

Thus, over all aggregate sizes, the scattering and absorption coefficients become

$$\sigma_s = N_{agg}\int_1^\infty N_p^2 \frac{8\pi x_p^6}{3\eta^2}\left|\frac{\boldsymbol{m}^2-1}{\boldsymbol{m}^2+2}\right|^2 g\left(\eta, R_g, D_f\right) dN_p$$
$$a_\lambda = N_{agg}\int_1^\infty N_p \frac{4\pi x_p^3}{\eta^2}\,\mathrm{Im}\left(\frac{\boldsymbol{m}^2-1}{\boldsymbol{m}^2+2}\right) dN_p \tag{B.3.6}$$

Integrating equations (B.3.6) and substituting into equation (B.3.1) gives a value for the scattering albedo of $\Omega \approx 0.0787$, which is in the range expected (Thomson 2000).

B.4 TEM Histogram Analysis

Transmission electron microscopy (TEM) was used to image soot aggregates captured from a flame through thermophoresis. The sampling was carried out at a height of 42mm above the burner at the same operating conditions detailed in Chapter 4. There were 3238 aggregates collected at this height and imaged, allowing the formation of an accurate histogram.

The number of primary particles in the aggregates captured in these images was then computed (making use of computer software to help this along) by means of the projected area method (PAM) (Köylü, McEnally, et al. 1997). This method makes use of the equation

$$N_p = f_a \left(\frac{A_a}{A_p} \right)^{\varepsilon_a} \tag{B.4.1}$$

where A_a is the projected area of the aggregate measured from the TEM image, $A_p = \pi d_p^2/4$ is the projected area of a primary particle, f_a is a projected area prefactor, and ε_a is an empirical exponent of the projected area (Köylü, Faeth and Farias, et al. 1995), (Brasil, Farias and Carvalho 1999).

Using this technique the size of each aggregate imaged by TEM is calculated. The size distribution is then created by splitting the aggregates up into uniformly spaced bins, and scaling each bin by the total number of aggregates. This produces a probability density function, as shown in Figure B.0.2.

To fit a lognormal distribution (Eq. (5.1)) to this distribution, the cumulative density function (CDF) is generated. This function gives the probability of an aggregate in the distribution being of a specific size or smaller. Mathematically, this is given as

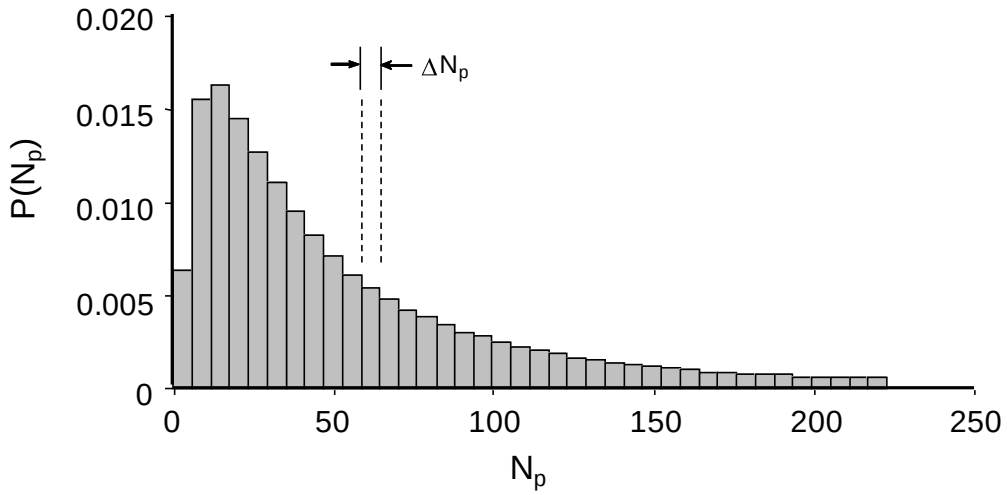

Figure B.0.2 – TEM histogram

$$\mathrm{CDF}\left(x\left(N_p\right)\right) = \int_{-\infty}^{N_p} x\left(N_p\right) dN_p \tag{B.4.2}$$

or, in the case of a discrete TEM distribution,

$$\mathrm{CDF}\left(x_p\right) = \sum_{i=1}^{p} x_i \Big/ \sum_{i=1}^{n} x_i \tag{B.4.3}$$

where x_i is the number of aggregates in each discrete strip of the distribution shown in Figure B.0.2 and n is the total number of strips. The CDF of the TEM histogram is shown in **Error! Reference source not found.**.

A similar method is carried out for a lognormal distribution to produce another CDF. A least-squares fitting procedure is applied to fit the lognormal distribution to the TEM data, perhaps using as an objective the Kolmogorov-Smirnov goodness-of-fit statistic (Appendix B.2). The resulting fitted distribution is shown in Figure B.0.4.

The benefit of fitting the CDF rather than the probability density function (PDF) is that it is an unbiased estimator (Press, et al. 2007). As well, while different distributions give different CDF's, all CDF's agree on the minimum and maximum values of 0 and 1, respectively. As well,

in the case of the soot aggregate size distribution, the minimum value occurs at the point $N_p = 0$ for all distributions.

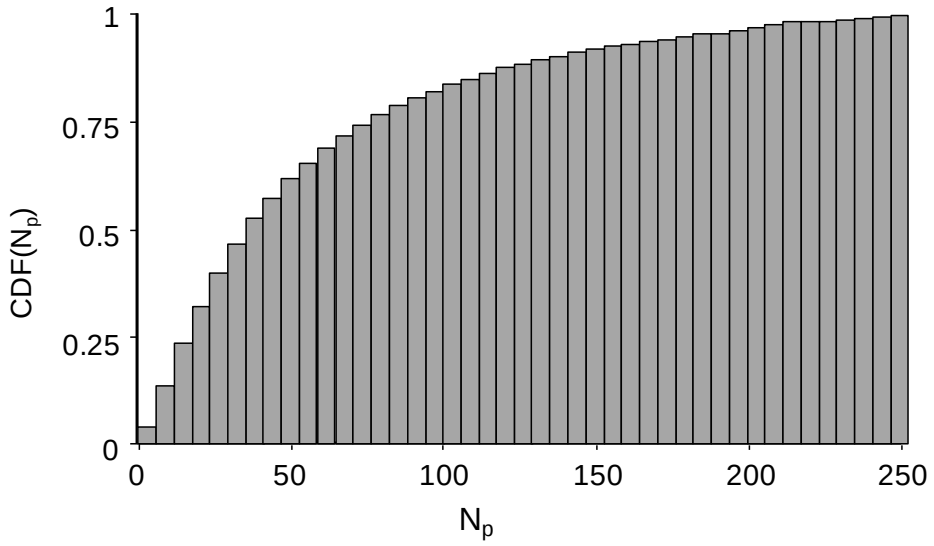

Figure B.0.3 – CDF of the TEM histogram

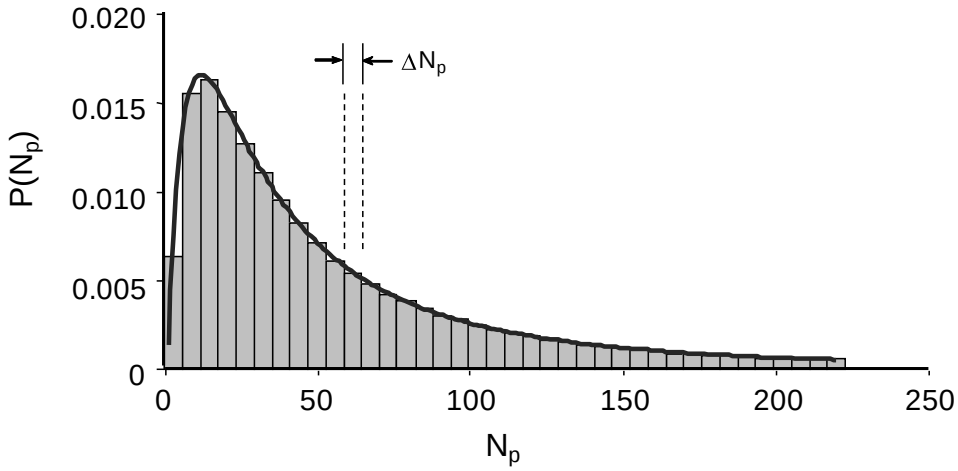

Figure B.0.4 – TEM histogram with fitted lognormal distribution

B.5 Derivation of the Matrix J

To determine a set of optimal angles, Özisik and Orlande (2000) suggest maximizing the determinant of the matrix $\mathbf{J}^{\mathrm{T}}\mathbf{J}$, where **J** is the sensitivity matrix. The form of the sensitivity matrix is derived starting with the ordinary least-squares objective function, defined as (Özisik and Orlande 2000)

$$S(\mathbf{x}) = \sum_{i=1}^{n} [g_i - \mathbf{A}_i \mathbf{x}]^2 \tag{B.5.1}$$

This equation can be rewritten in matrix form as

$$S(\mathbf{x}) = [\mathbf{g} - \mathbf{A}\mathbf{x}]^{\mathrm{T}} [\mathbf{g} - \mathbf{A}\mathbf{x}] \tag{B.5.2}$$

Next, this objective function must be minimized, which is accomplished by finding the value of the unknown parameter, **x**, at which the derivative of $S(\mathbf{x})$ is equal to zero.

$$\frac{\partial S(\mathbf{x})}{\partial x_1} = \frac{\partial S(\mathbf{x})}{\partial x_2} = \cdots = \frac{\partial S(\mathbf{x})}{\partial x_n} = 0 \tag{B.5.3}$$

This can be rewritten in gradient matrix form by taking the derivative of the matrix equation, Eq. (1.3), with respect to **x**.

$$\nabla S(\mathbf{x}) = 2\left[-\frac{\partial (\mathbf{A}\mathbf{x})^{\mathrm{T}}}{\partial \mathbf{x}}\right][\mathbf{g} - \mathbf{A}\mathbf{x}] = 0 \tag{B.5.4}$$

where $\partial(\mathbf{A}\mathbf{x})^{\mathrm{T}}/\partial\mathbf{x}$ is given by

$$\frac{\partial (\mathbf{A}\mathbf{x})^{\mathrm{T}}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{bmatrix} [\mathbf{A}_1\mathbf{x} \quad \cdots \quad \mathbf{A}_n\mathbf{x}] \tag{B.5.5}$$

The sensitivity matrix, **J**, is defined as

$$\mathrm{J}(\mathbf{x}) = \left[\frac{\partial(\mathbf{A}\mathbf{x})^{\mathrm{T}}}{\partial \mathbf{x}}\right]^{\mathrm{T}} = \begin{bmatrix} \frac{\partial \mathbf{A}_1\mathbf{x}}{\partial \mathrm{x}_1} & \cdots & \frac{\partial \mathbf{A}_1\mathbf{x}}{\partial \mathrm{x}_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \mathbf{A}_n\mathbf{x}}{\partial \mathrm{x}_1} & \cdots & \frac{\partial \mathbf{A}_n\mathbf{x}}{\partial \mathrm{x}_n} \end{bmatrix} \tag{B.5.6}$$

where

$$\mathbf{A}_i\mathbf{x} = \mathrm{A}_{i,1}\mathrm{x}_1 + \mathrm{A}_{i,2}\mathrm{x}_2 + \cdots + \mathrm{A}_{i,n}\mathrm{x}_n \tag{B.5.7}$$

If each value x_i is assumed to be independent of every other value x_j, then Eq. (B.5.6) can be written as

$$\mathrm{J}(\mathbf{x}) = \begin{bmatrix} \mathrm{A}_{1,1} & \cdots & \mathrm{A}_{1,n} \\ \vdots & \ddots & \vdots \\ \mathrm{A}_{n,1} & \cdots & \mathrm{A}_{n,n} \end{bmatrix} \tag{B.5.8}$$

which is just the kernel matrix, **A**. Thus, the sensitivity matrix is, in this case, simply the kernel matrix, which is also called the Jacobian Matrix. Thus, to find the optimal angles, the determinant of $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ is maximized.

B.6 Experimental Data

The following data was collected by the author between May 13 and June 30, 2010, at the National Research Council, Canada.

Angle	Measurement
10.000	642.430
14.455	627.188
19.314	606.754
24.208	589.385
29.147	556.212
34.143	534.844
39.206	496.267
44.350	478.798
49.590	431.245
54.943	401.522
60.430	368.817
66.075	342.999
71.906	318.337
77.962	290.651
84.288	268.234
90.948	245.670
98.028	226.149
105.653	208.422
114.016	191.990
123.450	177.023
134.625	164.570
149.400	153.725
160.000	148.592