BOSTON UNIVERSITY

COLLEGE OF ENGINEERING

Thesis

LOW ION DOSE IMAGING IN THE HELIUM ION MICROSCOPE UNDER NEYMAN TYPE A STATISTICS

by

JEFF CRALEY

B.S., Virginia Tech, 2011
B.A., Virginia Tech, 2011

Submitted in partial fulfillment of the

requirements for the degree of

Master of Science

2015

Approved by

First Reader

Vivek Goyal, PhD
Associate Professor of Electrical and Computer Engineering

Second Reader

W. Clem Karl
Professor of ECE, BME, and SE

Third Reader

Anna Swan
Assistant Professor of ECE and MSE

Acknowledgments

For his role as my academic and research adviser I would like to thank Vivek Goyal for his help and support. Karl Berggren and his group, especially Richie Hobbs and Vitor Manfrinato, have provided expertise and help with the HIM. I'd like to thank Michelle and Picabo the cat. Also my family. This work was supported in part by the U.S. National Science Foundation under Grant 1422034.

Jeff Craley

Student

ECE Department

LOW ION DOSE IMAGING IN THE HELIUM ION MICROSCOPE UNDER NEYMAN TYPE A STATISTICS

JEFF CRALEY

ABSTRACT

The Helium Ion Microscope (HIM) is a new tool capable of imaging at resolutions not possible in the ubiquitous Scanning Electron Microscope (SEM). However, along with this improvement in imaging, the helium ions used in HIM cause greater damage to samples. Reducing ion doses decreases the amount of damage incurred at the expense of increased noise in the final image.

Motivated by the damage to samples inherent in HIM imaging, the possibility of imaging at low ion doses is investigated. A two-parameter Neyman Type A model of pixelwise Ion induced Secondary Electron (iSE) emission is introduced. This model takes into account randomness in both the Poisson emission of helium ions and the Poisson emission of iSE per each incident helium ion. HIM images are simulated for low ion doses. Using the Neyman Type A model, an image formation algorithm is developed using maximum likelihood (ML) estimation. This estimation procedure is regularized by sparsity in the wavelet domain to account for spatial correlations inherent in real world samples.

Contents

List of Tables

List of Figures

List of Abbreviations

ET	.....	Everhart-Thornley
FIB	.....	Focused Ion Beam
FIM	.....	Field Ion Microscope
FPI	.....	First Photon Imaging
HIM	.....	Helium Ion Microscope
IONiSE	.....	(Monte Carlo Simulated) Ion Induced Secondary Electrons
iSE	.....	Ion Induced Secondary Electron
GFIS	.....	Gas Field Ion Source
KE	.....	Kinetic Emission
ML	.....	Maximum Likelihood
MSE	.....	Mean Squared Error
PE	.....	Potential Emission
PGF	.....	Probability Generating Function
PMF	.....	Probability Mass Function
PMT	.....	Photomultiplier Tube
RBS	.....	Rutherford Backscattering Spectroscopy
SE	.....	Secondary Electron
SEM	.....	Scanning Electron Microscope
SNR	.....	Signal-to-Noise Ratio
SPAD	.....	Single Photon Avalanche Diode
SpaRSA	.....	Sparse Reconstruction by Separable Approximation
SPIRAL	.....	Sparse Poisson Image Reconstruction Algorithm
SRIM	.....	Stopping and Range of Ions in Matter
TV	.....	Total Variation

Chapter 1

Introduction

1.1 Motivation

The helium ion microscope (HIM) is one of the newest tools in focused ion beam (FIB) microscopy. When compared to the ubiquitous scanning electron microscope (SEM), the HIM is capable of producing images both at higher resolutions and with better contrast characteristics. These improvements are direct results of the use of helium ions in the imaging process, which are much more massive than the electrons used in SEM.

However, the benefits in imaging with an HIM can be offset by the damage produced in the imaging process. The larger helium ions impinging on the sample can cause damage at a far greater rate than the electrons used in an SEM. The damage can be especially bad for organic samples but occurs for many different types of materials. Damage in HIM imaging can be controlled by using lower ion doses. However this method introduces a trade-off between image quality and sample damage, as the noise in the HIM image increases as fewer impinging ions are used.

1.2 Problem Overview

In this thesis a model of image formation in the HIM following Neyman Type A statistics is introduced. This model accounts for randomness in the emission of helium ions as well as randomness in the emission of electrons from the sample. Using

this model, an image is formed by performing regularized maximum likelihood (ML) estimation to recover secondary electron yield values at each pixel. The estimation is regularized by sparsity in the wavelet domain to capture spatial correlations in samples. This process is performed by solving the minimization problem

$$\begin{aligned} \hat{\eta} = \quad & \underset{\eta \in \mathbb{R}^{n \times n}}{\arg\min}\ F(\eta, Y, \lambda) + \tau \|W^T \eta\|_1 \\ & \text{subject to } \eta \geq 0 \end{aligned} \tag{1.1}$$

where η represents an image of local secondary electron yields, $F(\eta, Y, \lambda)$ is the negative log-likelihood of the Neyman Type A distribution, W represents a wavelet transform, λ is the mean number of helium ions emitted per pixel, and τ is a regularization parameter. The function $F(\eta, Y, \lambda)$ will be specifically defined in Chapter 3 and the minimization problem will be discussed in Chapter 4.

In comparison with traditional HIM imaging, the estimation based technique developed in this thesis redefines image formation in the HIM in two main regards. First, in traditional HIM the noise due to randomness in the imaging process is typically suppressed by increasing the ion dose. Instead of relying on higher ion doses, the image formation technique developed here takes into account both randomness in the helium ion beam and randomness in the emission of electrons. By using a statistical model of image formation regularized by spatial correlations, imaging to take place at lower ion doses can occur.

Secondly, traditional HIM imaging uses the signal of collected electrons to create an image at every pixel location. Changes in the electron signal are due to changing secondary electron yields across the sample. This yield value η is a physical parameter depending on properties of the material and the shape of the sample. Instead of using this signal as the image itself, the image formation procedure detailed here estimates the secondary electron yield at every pixel. Thus the image formed is an estimate

a physical property of the sample where traditional HIM imaging uses the electron signal as a stand-in for this parameter. In expected value, the iSE represents a scaling of the η parameter, however at low ion doses the randomness in the imaging process degrades the iSE signal. In low ion dose imaging, it becomes important to form an image directly using the η parameter to avoid this degradation.

Estimation under the Neyman Type A likelihood distribution poses several challenges. The Neyman Type A distribution is multimodal, which results in a non-convex negative log-likelihood. However, in the regions of valid secondary electron yields for the HIM microscope we find this function to be pseudoconvex, justifying the use of subgradient methods. The solution set is restricted to be non-negative, adding further complication to the minimization.

1.3 Prior Work

In First-Photon Imaging (FPI) (Kirmani et al., 2014), depth and reflectivity images are formed for a scene using only one detected photon at each pixel using a Single-Photon Avalanche Detector (SPAD). Through the combination of accurate probabilistic models and the exploitation of spatial correlations within the scene, FPI is able to perform robust 3D imaging. The FPI image acquisition system used a variable dwell time and was only able to perform one measurement at a time. Recently the FPI method was extended to use a SPAD array and a fixed dwell for an overall reduction in imaging time (Shin et al., 2015). The approach taken in FPI inspired the approach taken here. Through the use of accurate statistical models and spatial regulation, HIM imaging at low ion doses was performed.

The Sparse Reconstruction by Separable Approximation (SpaRSA) (Wright et al., 2009) framework was developed to perform sparse reconstructions using subgradient methods. The SpaRSA framework uses a series of subproblems that are solvable

using iterative shrinkage/thresholding algorithms when the original problem may be difficult to solve directly. The Sparse Poisson Intensity Reconstruction Algorithms (SPIRAL) (Harmany et al., 2012) framework applies this approach to signal reconstruction under Poisson noise. The approach taken to image regularization in HIM under the Neyman Type A distribution uses an approach very similar to that taken by SPIRAL.

1.4 Outline

In the Chapter 2, a brief overview of the HIM is presented. This includes a discussion of its development and working principle. Discussion of the secondary electron signal generated in the microscope and the types of damage produced in the imaging process is also included. Finally the Neyman Type A statistical model for the generation of the secondary electrons is introduced in the context of the HIM.

In Chapter 3 the Neyman Type A distribution is discussed in more detail. A description of its properties and behavior is presented. The estimation algorithm requires the use of the gradient and hessian of the negative log-likelihood. These formulas are derived. Finally, an argument for the validity of gradient based regularization techniques under the assumptions regarding physically possible secondary electron yields is made.

Chapter 4 includes a description of the estimation procedure used in this thesis. The procedure is adapted from SPIRAL. The overall approach is presented, followed by the solution to the series of subproblems created.

In Chapter 5 and Chapter 6 experimental results are presented. Chapter 5 focuses on a simulated imaging experiment. In this experiment, noisy images are simulated from a ground truth HIM image. The performance of the image formation algorithm is compared to a naive image formation procedure. Chapter 6 focuses on the application

of the image formation procedure to real world data. This chapter details an attempt to fit real HIM images to the Neyman Type A model and the use of the estimation procedure on the resulting images.

In Chapter 7 conclusions are presented. The major contributions of the project as well as the experimental results are revisited. Finally, avenues for future work on the project are discussed.

Chapter 2

Helium Ion Microscope Background

2.1 Introduction

The basic working principle in HIM microscopy, in other types of FIB microscopy, and in the SEM is practically the same regardless of the impinging particle used. To create a beam, an ion or electron source creates ionized particles and accelerates them downward towards the sample. Next, the beam is focused either by an electrostatic or magnetic lens. The beam is swept across the sample and at each location some signal is measured. While this general working principle is the same across all particle microscopy, the properties of the particles composing the beam as well as the kind of signal collected have a major effect on the imaging process.

Developed by the ALIS Corporation, which was subsequently purchased by Zeiss, the HIM, shown in Figure 2·1, is a recently developed tool in FIB microscopy (Ward et al., 2006). Using the same working principle as the SEM, the HIM improves upon imaging performance. Most notably the HIM is able to achieve much smaller resolutions than possible in SEM. Because of the short wavelengths of the large helium ions used, the diffraction limit in the HIM is smaller than that of the SEM. However, this increase in performance comes at the price of increased damage to the sample being imaged. The much heavier helium ions cause more damage than the smaller electrons.

This chapter discusses the HIM as an imaging tool. In many cases it will be desirable to compare the HIM to the SEM. This will include discussion of the physical

Figure 2·1: Zeiss Orion HIM, image from (Zeiss, 2015)

principles at work in the microscope as well as the signals generated. This chapter will also contain a discussion of the damage caused by the HIM. Finally the Neyman Type A model of the electron statistics will be introduced. This distribution will be discussed in further detail in Chapter 3.

2.2 The Ion Beam

The HIM was predated by the field ion microscope (FIM). Developed by Erwin Müller in the 1950's, the FIM was the first instrument capable of imaging an individual atom. The FIM consists of a positively charged needle surrounded by neutral helium gas particles. The tip of the needle, which is cryogenically cooled, is on the order of 100 nm (Ward et al., 2006). Particles from the imaging gas are drawn towards the tip of the needle. The particles become ionized in regions called ionization disks that form at each atom on the surface of the needle. Due to the strong electric fields in these regions, quantum tunneling of electrons from the helium particles into the needle is possible. The now positively charged helium ions are accelerated from the positively charged needle in a direction roughly perpendicular to the surface of the

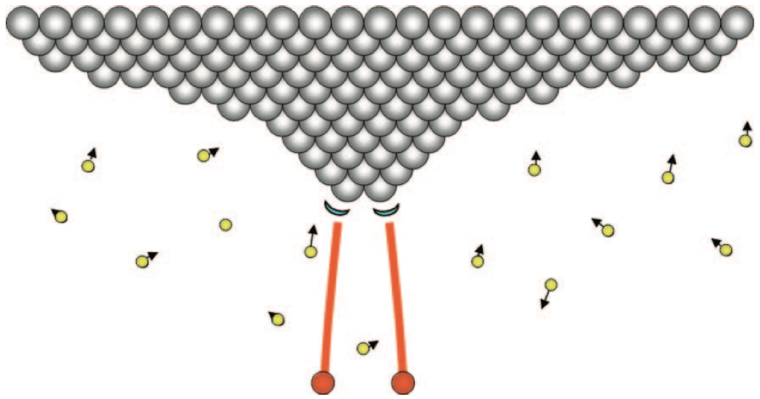

Figure 2·2: Ion source surrounded by helium, image from (Ward et al., 2006)

needle. A screen is placed in the path of the accelerated helium ions and an image of the needle is formed.

Because of the relatively large and spherical shape of the needles used in FIM, the ion sources are incapable of forming a tightly focused beam. A new type of ion source was necessary to make the HIM possible. In 2006, the ALIS corporation developed their gas field ion source (GFIS) for HIM. Through a proprietary process, the tip of the needle is shaped into a three sided pyramid. At the apex of the pyramid, three atoms called the trimer form a stable tip where ionization takes place. By using a pyramidal structure shown in Figure 2·2, the electric field is concentrated at these three atoms and ionization takes place mostly at the trimer. Thus three distinct beams are created at each atom at the tip of the trimer.

In HIM, the helium gas is shared only by the atoms at the tip of the needle in the GFIS. Because ionization takes place at a smaller number of atoms in GFIS, more helium atoms ionize per atom at the trimer. This results in a brighter, more focused beam. Additionally, the beam can be apertured such that only ions from one atom of the trimer make up the beam. In this fashion, the GFIS can produce a tightly focused, bright ion beam with little energy spread (Ward et al., 2006).

A cross section showing the beam column is shown in Figure 2·3. The ions are

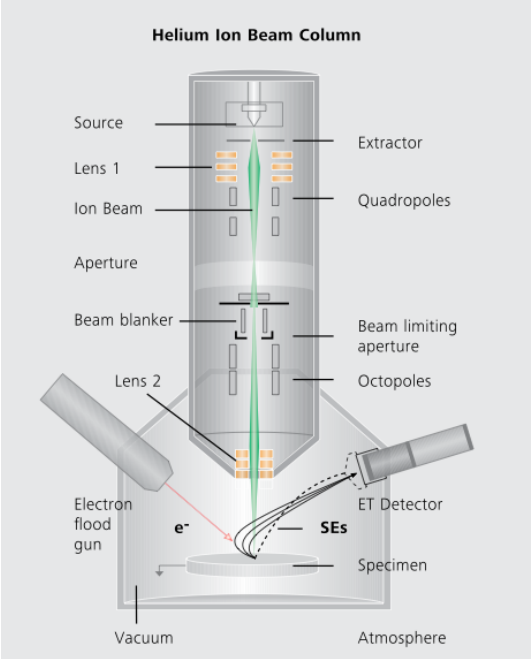

Figure 2·3: Cross Section of the Zeiss Orion HIM, image from (Zeiss, 2015)

accelerated downward towards the sample by an accelerator voltage. This is set by the user at the beginning of the imaging process. By adjusting the accelerator voltage, the speed, and thus the energy of the ions in the beam can be controlled. Next the ion beam passes through a series of electrostatic lenses that focus the beam on the sample. By applying a varying voltage to the lens, the beam is raster scanned across the sample.

2.3 Secondary Electron Signal

When an impinging helium ion beam strikes the sample many different physical processes occur, allowing for several possible modes of imaging. Some of the impinging helium ions are reflected by the sample and bounce backwards through a process called Rutherford backscattering. These ions can be collected and are sometimes used for identifying materials present in the sample through a technique called Rutherford

backscattering spectroscopy (RBS). In addition, photons and x-rays are produced by impinging helium ions. These signals are available for imaging as well.

The most widely used signal in HIM, as well as in SEM and other types of FIB microscopy, is the secondary electron (SE) signal. When charged particles impinge upon a sample, atoms within the sample are ionized and eject electrons. These electrons are collected, most commonly using an Everhart-Thornley (ET) detector. Due to SE contrast changes caused by material and topography, the SE based images are easy to interpret by the human visual system. This ease of interpretation along with the relatively high SE yield makes SE imaging the most commonly used imaging mode in SEM, HIM, and other kinds of FIB microscopy.

The ET detector works by applying a positive voltage to a grid to attract electrons. The electrons then strike a scintillator. The scintillator emits photons which are captured and amplified by a photomultiplier tube (PMT). The intensity of the output of the PMT is what determines the pixel brightness at any given location. As the beam is raster scanned across the sample, the PMT intensities form the brightness at each pixel, thus creating the HIM image.

When any ion impinges upon a sample, the material begins to slow the particle. The rate at which the particle is slowed is called the stopping power. It was proposed by Bethe in the 1930's that the number of electrons ejected is proportional to the stopping power:

$$\eta_{SE} = -\frac{1}{\epsilon}\frac{dE}{dS}. \tag{2.1}$$

In this expression, ϵ is a parameter that varies depending on the material. E represents change in energy while S represents distance.

To demonstrate the advantages to HIM, a discussion of the SE signal in SEM is useful. In SEM, SE generated by the initial electron strike are called SE1. The SE1 are generally created closer to the surface and nearer to the location that the impinging

electron struck the sample. Therefore the SE1 represent the signal of interest, as they are generated close to the location currently being imaged. As the original electron travels into the surface, it is often backscattered. After backscattering, the electron can create more SE, denoted SE2. The SE2 signal is generally two to three times larger than the SE1 signal. Because the original electron has often traveled some distance before producing SE2, the SE2 are less localized. Because both SE1 and SE2 are collected by the ET detector, the signals are indistinguishable from each other. This results in a weak signal in the presence of a noisy background. In addition, the impinging electron can travel through a large interaction volume, which also reduces the possible resolution. The combination of these effects limits the resolution possible in SEM, increases the noise in the image, and degrades its contrast.

In the HIM, ion induced secondary electrons (iSE) have few of the adverse properties present in the SEM. While backscattered helium ions can create secondary electrons, this portion of the signal is miniscule compared to iSE generated by unbackscattered ions. Thus the SE2 signal that degrades imaging in the SEM is not present in the HIM. In addition, the stopping power for helium ions is greater than the stopping power for electrons. Thus the helium ion travels a smaller distance inside the sample. This results in a smaller interaction volume and an iSE signal that is local to the point of impingement. As seen in equation (2.1), the stopping power is also proportional to the number of iSE generated. Because the stopping power of helium is higher, the iSE yield is roughly 3-5 times higher than SE yield in the SEM. Where SE yield maximums are roughly 1.5, iSE yields can be on the order of 1-10 for impinging ions in the 1-10 MeV energy range (Joy, 2013). All of these attributes contribute to form a generally better quality image.

2.4 Damage

In all kinds of ion beam microscopy, some damage to the sample is unavoidable. Whenever charged particles are incident upon a sample, they have some kind of effect on the sample. However, when compared to SEM, the damage caused in HIM is much greater. This is due to the much larger size of the helium ions, which are roughly 7,300 times as massive as electrons.

In SEM, many different kinds of damage can occur. After impinging upon the sample, some electrons from the beam remain lodged inside the sample. If the number of SE ejected from the sample is different from the number of electrons that remain in the sample, charging occurs. Charging can both degrade the quality of the SEM image and cause damage to the sample. In order to avoid charging, samples can be coated with a conducting material which effectively destroys the sample for uses other than imaging. While metals generally do not need coating and do not suffer from charging damage, insulators can suffer from these kinds of damage. In addition, organic materials and polymers can be chemically altered during the imaging process (Joy, 2013).

HIM samples suffer both more and a wider variety of damage. While the much heavier helium ions are typically moving much slower than electrons in SEM, their increased mass causes much greater damage. Similar to SEM, the HIM imaging process causes charging. As positively charged helium ions are lodged in the sample, negatively charged electrons are ejected. This leads to an accumulation of positive charge in the sample that can again degrade the image and cause damage. Through a process called sputtering, atoms can be knocked off of the surface sample by impinging helium ions. While the sputtering process can useful for nanofabrication and milling, in normal imaging situations it is generally an undesired side effect. In addition, the heavier helium atoms can be lodged in the crystaline samples causing interstitials.

Conversely, if the binding energy of a crystal lattice is low enough, impinging helium ions can completely knock atoms out of the crystal lattice causing vacancies. These point defects, sputtering, and damage due to charging are all prevalent in HIM. Thus the damage in HIM imaging is much greater than the damage in SEM. This increase in damage motivates our image reconstruction algorithm.

2.5 Electron Statistics

Some discussion of electron number statistics is required to justify our modeling of the imaging process. When charged ions impinge on a material, secondary electron emission can occur due to two main processes. Kinetic emission (KE) is the result of quickly moving helium ions causing ionization in the sample. Potential emission (PE) of iSE is the result of potential energy interactions between the ion and the sample. Typically the electron number statistics are heavily dependent on the different kinds of electron emission that occur in any given ion/sample interaction. Furthermore the ratio of PE and KE emission will be affected by the type of ion and sample material. A discussion of these processes is available from Vana et al. (Vana et al., 1995).

Electron number statics have been derived for different ions impinging on different samples in (Eder et al., 1997) and (Vana et al., 1995). Specifically, these studies evaluate the electron number statistics for H^+, He^+, He^{2+}, N^{3+}, N^{4+}, O^{5+}, O^{6+} impinging on gold and a similar array of ions impinging on gold and lithium flouride. Different probability distributions were fit to the measured numbers of iSE per impinging ion. Binomial, Poisson, and Polya distributions were fit to the measured electron number distributions with differing levels of accuracy depending on the ion and sample. It was shown however that for cases where iSE generation was purely the result of KE, the Poisson distribution model was the most appropriate.

In this thesis, the model of the imaging process assumes Poisson distributed iSE

generation statistics. This assumption is justified by noting that PE emission for impinging He^+ ions is only significant for ions travelling less than about 10^7 cm/sec. This corresponds to He^+ ions with energies of only a few hundred eV. At a few keV and higher, iSE generation for He^+ ions is almost entirely due to kinetic energy emission. In HIM, ions are travelling in the regime where only KE emission is significant (Ramachandra et al., 2009). As such it is assumed that iSE generation is Poisson distributed.

The assumption that iSE statistics are determined by the Poisson distribution is echoed in (Castaldo et al., 2011) and (Orloff et al., 2011). To complete the model of electron emission, it is assumed He^+ emission in the ion beam is Poisson distributed as well. Thus at each pixel, a Poisson number of He^+ ions impinge upon the sample. For each impinging He^+ ion, a Poisson number of SE are emitted. This model of image formation takes into account both randomness in the source of helium and randomness in the emission of iSE per impinging ion. The distribution that arises from this combination of Poisson distributions is the Neyman Type A distribution. The Neyman Type A distribution and its properties will be discussed in Chapter 3.

Other methods of modeling iSE yields have been developed as well. Ramachandra et al. developed a method for simulating the imaging process in HIM called IONiSE. Their simulation uses the widely used SRIM system to simulate the trajectory of impinging ions. When the ion enters the sample, some scattering angle is computed. The ion then travels some distance until the next scattering event and its trajectory is recomputed. The energy loss is computed along each step, depending on the stopping power of the ion in the sample material. If, at the end of a given step, the ion is outside the sample or its energy dips below a predetermined level, the trajectory for a new ion is begun. Along each step of the trajectory, the stopping power is used to determine some amount of iSE generated according to Bethe's formula shown in equation (2.1).

Depending on the depth within the sample, the electron diffuses out of the sample with some probability. Combining primary ion trajectory and iSE generation modeling, the IONiSE system is capable of simulating iSE yields for ions impinging at different energies on different sample materials (Ramachandra et al., 2009).

Chapter 3

Neyman Type A Distribution

3.1 Background

As developed in Chapter 2, iSE generation in the HIM can be modeled using the Neyman Type A distribution. The Neyman Type A distribution was first conceived by Neyman in 1939. Neyman was interested in a class of distributions he termed "contagious" distributions, useful in representing certain biological processes. In particular, Neyman used the Type A distribution to model the distribution of larvae in a given area of a field. He assumed that both the distribution of egg clusters and the number of larvae produced by each cluster were Poisson distributed (Johnson et al., 1992). The Neyman Type A distribution has also been proposed for pulse, particle, and photon counting as well (Teich, 1981). As such, the total number of counts Y is given by

$$Y = \begin{cases} 0 & \text{if } M = 0 \\ \sum_{m=1}^{M} X_m & \text{if } M \geq 1 \end{cases} \tag{3.1}$$

where $M \sim \text{Poisson}(\lambda)$ and $X_m \sim \text{Poisson}(\eta)$ and each emission event X_m is independent.

The resulting distribution is a Poisson-stopped summed-Poisson distribution, which can be represented as

$$Y \sim \text{Poisson}(\lambda) \bigvee \text{Poisson}(\eta). \tag{3.2}$$

The mean of the distribution is $\mu = \lambda\eta$ and the variance is $\sigma^2 = \lambda\eta + \lambda\eta^2$. The proba-

bility generating function of the Neyman Type A distribution can be derived from the probability generating functions (PGF) of its two component Poisson distributions.

$$G_X(z) = \exp[\eta(z-1)] \tag{3.3}$$

$$G_M(z) = \exp[\lambda(z-1)] \tag{3.4}$$

Thus the PGF for the Neyman Type A distribution is

$$G_Y(z) = G_M(G_X(z)) = \exp[\lambda(e^{\eta(z-1)} - 1)]. \tag{3.5}$$

Alernatively, the Neyman Type A distribution can be regarded as a mixture of Poisson distributions. The total number of counts is Poisson with parameter $m\eta$ where m is Poisson with parameter λ. Thus the probability of some number of counts y can be expressed as

$$p_Y(y) = \sum_{m=0}^{\infty} \frac{e^{-m\eta}(m\eta)^y}{y!} \cdot \frac{e^{-\lambda}\lambda^m}{m!} = \frac{e^{-\lambda}\eta^y}{y!} \sum_{m=0}^{\infty} \frac{(\lambda e^{-\eta})^m m^y}{m!}. \tag{3.6}$$

Useful properties can be derived from both the PGF and PMF however mainly (3.6) will be used.

Examples of different Neyman Type A distributions with mean 10 are shown in Figure 3·1. As the λ parameter is decreased, the probability of zero counts increases. It can be observed that the Neyman Type A distribution is multimodal, where in many cases a mode exists at $Y = 0$ and near $Y = \eta\lambda$. However, the number of modes is not bounded. Figure 3·2 shows a multimodal Neyman Type A distribution with $\lambda = 3$ and $\eta = 10$. Modes are visible in the figure at roughly integer values multiplied by the η parameter. This holds in general, as modes usually appear around 0 and at integer multiples of η for most values of λ and η.

The multimodality of the Neyman Type A distribution presents problems for

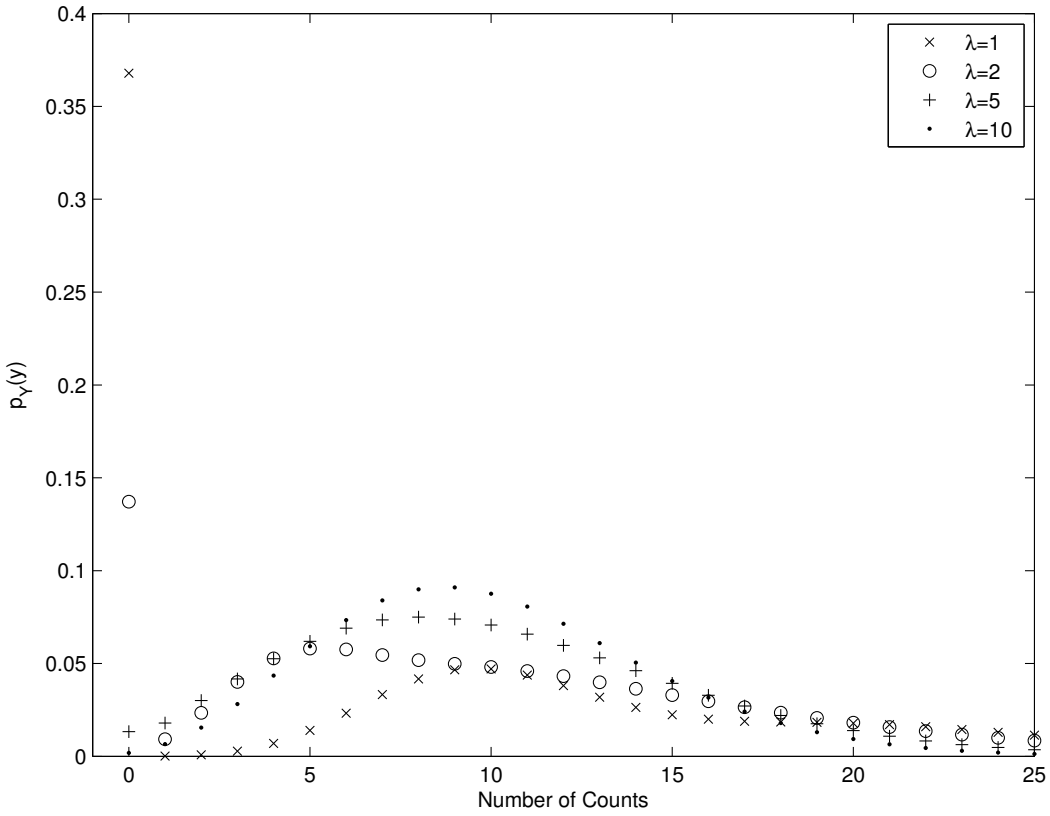

Figure 3·1: Comparison of different Neyman Type A distributions with mean $\eta\lambda = 10$

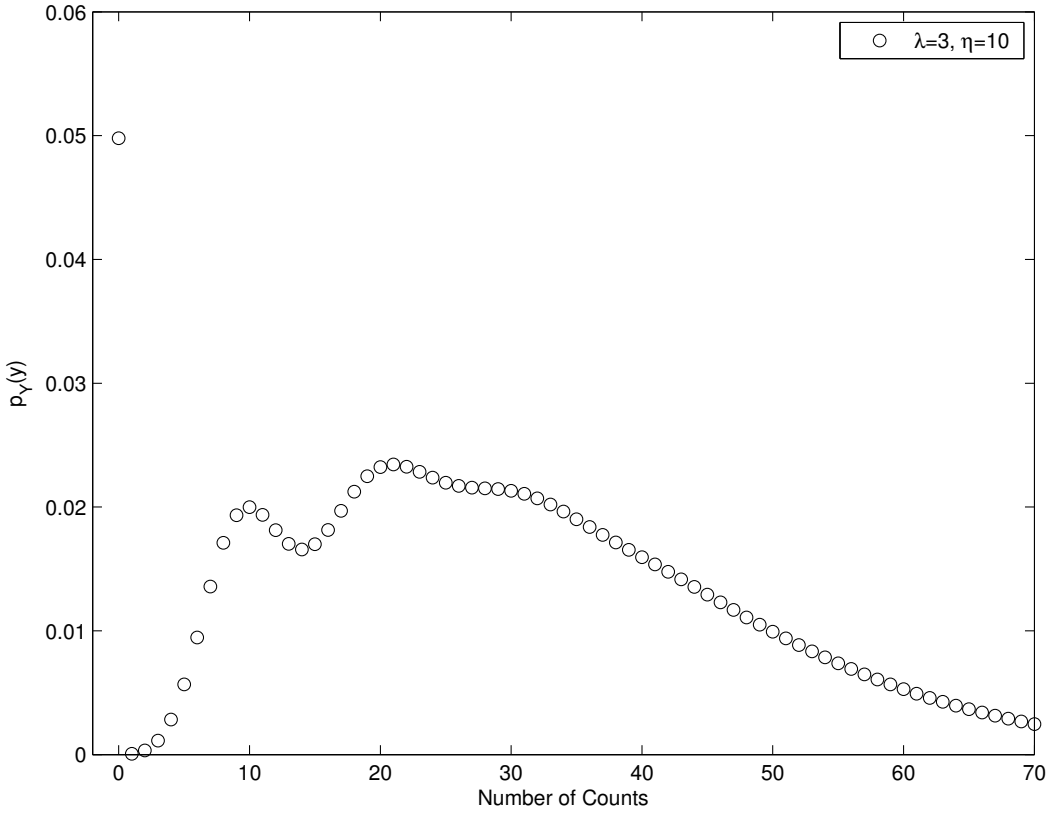

Figure 3·2: Example of a multimodal Neyman Type A distribution

assuring the convergence of the algorithm. However, the range of η determined to be physically valid is limited by the maximum iSE yield in the HIM. As such, the estimation procedure must be well behaved over this range of η and λ. Section 3.3 will argue for the existence of a region in which pseudoconvexity holds for the negative log-likelihood function. This region is shown to include the physically reasonable values of λ and η.

3.2 Negative Log-Likelihood and its Derivatives

To find computable formulas, the infinite sum in (3.6) must be expressed in closed form. Following (Boyadzhiev, 2009),(Roman, 1984) a useful family of polynomials called the exponential polynomials can be derived by considering the expression

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Differentiating both sides and multiplying by x we have

$$xe^x = \sum_{k=0}^{\infty} \frac{kx^k}{k!}.$$

Continuing this pattern of differentiation and multiplication, the exponential polynomials $\phi_n(x)$ of degree n can be defined as

$$x\left(xe^x\right)' = (x + x^2)e^x = \sum_{k=0}^{\infty} \frac{k^2 x^k}{k!}$$

$$\left(x\frac{d}{dx}\right)^n e^x = \phi_n(x)e^x = \sum_{k=0}^{\infty} \frac{k^n x^k}{k!}. \tag{3.7}$$

It can be shown that the exponential polynomials can be represented as

$$\phi_n(x) = \sum_{k=0}^{n} S(n,k)x^k$$

where $S(n,k)$ are the Stirling numbers of the second kind.

These relations make it possible to express $p_Y(y)$ as a polynomial. Using equation (3.7), the infinite sum in equation (3.6) can be written in the form of an exponential polynomial

$$\sum_{m=0}^{\infty} \frac{(\lambda e^{-\eta})^m m^y}{m!} = \phi_y(\lambda e^{-\eta}) \exp\left(\lambda e^{-\eta}\right).$$

Thus the PMF from equation (3.6) can be expressed using the exponential polynomials as

$$p_Y(y) = \frac{e^{-\lambda}\eta^y}{y!}\phi_y(\lambda e^{-\eta}) \exp(\lambda e^{-\lambda}). \tag{3.8}$$

This closed form version of the PMF allows for the computation of Neyman Type A probabilities.

To define the negative log-likelihood function $F(\eta)$ the infinite sum version of the PMF in equation (3.6) can be used. This leads to easier derivation of the gradient of $F(\eta)$. Defining the the likelihood function $\ell(\eta|y,\lambda) = p_Y(y|\eta,\lambda)$, $F(\eta)$ can be defined as

$$F(\eta) = -\log\left[\ell(\eta|y,\lambda)\right] = \lambda - y\log(\eta) + \log(y!) - \log\left(\sum_{k=0}^{\infty} \frac{(\lambda e^{-\eta})^k k^y}{k!}\right). \tag{3.9}$$

Taking the first derivative of $F(\eta)$, we find

$$\begin{aligned}
\frac{d}{d\eta}F(\eta) &= -\frac{y}{\eta} + \frac{\sum_{k=0}^{\infty} \frac{k(\lambda e^{-\eta})^k k^y}{k!}}{\sum_{k=0}^{\infty} \frac{(\lambda e^{-\eta})^k k^y}{k!}} \\
&= -\frac{y}{\eta} + \frac{\sum_{k=0}^{\infty} \frac{(\lambda e^{-\eta})^k k^{y+1}}{k!}}{\sum_{k=0}^{\infty} \frac{(\lambda e^{-\eta})^k k^y}{k!}} \\
&= -\frac{y}{\eta} + \frac{\exp(\lambda e^{-\eta})\phi_{y+1}(\lambda e^{-\eta})}{\exp(\lambda e^{-\eta})\phi_y(\lambda e^{-\eta})} \\
&= -\frac{y}{\eta} + \frac{\phi_{y+1}(\lambda e^{-\eta})}{\phi_y(\lambda e^{-\eta})}.
\end{aligned} \tag{3.10}$$

This function will be used to compute the gradient in the estimation algorithm.

Taking the second derivative,

$$\begin{aligned}\frac{d^2}{d\eta^2}F(\eta) &= \frac{y}{\eta^2} + \frac{\left(-\sum_{k=0}^{\infty}\frac{k(\lambda e^{-\eta})^k k^{y+1}}{k!}\right)\cdot\left(\sum_{k=0}^{\infty}\frac{(\lambda e^{-\eta})^k k^y}{k!}\right) + \left(\sum_{k=0}^{\infty}\frac{k(\lambda e^{-\eta})^k k^y}{k!}\right)^2}{\left(\sum_{k=0}^{\infty}\frac{(\lambda e^{-\eta})^k k^{y+1}}{k!}\right)^2} \\ &= \frac{y}{\eta^2} + \frac{-\phi_{y+2}(\lambda e^{-\eta})\cdot\phi_y(\lambda e^{-\eta})\cdot[\exp(\lambda e^{-\eta})]^2 + [\phi_{y+1}(\lambda e^{-\eta})\cdot\exp(\lambda e^{-\eta})]^2}{[\phi_y(\lambda e^{-\eta})\cdot\exp(\lambda e^{-\eta})]^2} \\ &= \frac{y}{\eta^2} + \frac{-\phi_{y+2}(\lambda e^{-\eta})\cdot\phi_y(\lambda e^{-\eta}) + (\phi_{y+1}(\lambda e^{-\eta}))^2}{(\phi_y(\lambda e^{-\eta}))^2}.\end{aligned} \tag{3.11}$$

These equations will be necessary for the estimation algorithm. Namely, negative log-likelihood will be used for the computation of the objective function, the first derivative will be used to compute the gradient, and the second derivative will be used in the determination of the step size taken in each iteration.

3.3 Region of Pseudoconvexity

The estimation scheme is complicated by the fact that the Neyman Type A distribution is multimodal, resulting in a non-convex negative log-likelihood function. In order to justify the use of a subgradient method in the optimization algorithm, an argument will be made for pseudoconvexity at low values of η. Furthermore, the iSE yield values that we have considered valid in this experiment lie within this region. This justifies the use of a subgradient method in the estimation algorithm. Thus while convergence can not be strictly guaranteed, in most cases the algorithm will converge. A general overview of convergence will be given in Appendix A while an argument for the pseudoconvexity of the negative log-likelihood in a range of η and λ values is made here.

A differentiable function f is defined to be pseudoconvex on some convex set X if for all $x, y \in X$ such that

$$\nabla f(x)\cdot(y-x) \geq 0$$

then

$$f(y) \geq f(x). \tag{3.12}$$

Under pseudoconvexity, $f(x^*)$ is a local minimum if and only if $\nabla f(x^*) = 0$. For subgradient methods, this means that the algorithm will cease to move at a minimum. In addition, the negative gradient at any point $x \neq x^*$ will point in a descent direction. Furthermore if the function is pseudoconvex over some convex region, a slightly relaxed Slater's condition holds and we can use duality to solve the minimization problem. Slater's condition states that when solving a convex optimization problem over a convex set, strong duality holds and there is no duality gap. This holds true if the objective function is pseudoconvex as well.

To determine the region of pseudoconvexity, the first derivative of the negative log-likelihood will be examined. Both the roots of the function and the behavior of the function as η goes to zero need to be considered. A loose argument for the existence of a pseudoconvex region between 0 and some boundary $\tilde{\eta}$ can be made by first noting that as

$$\eta \to 0_+, \quad \frac{y}{\eta} > \frac{\phi_{y+1}(\lambda e^{-\eta})}{\phi_y(\lambda e^{-\eta})}$$

thus

$$\lim_{\eta \to 0_+} \frac{d}{d\eta} F(\eta) = -\infty. \tag{3.13}$$

Near $\eta = 0$, $\nabla F(\eta) < 0$ and thus the objective function is decreasing. Assuming the existence of some lowest positive root $\bar{\eta}$, it can be noted that for $\eta < \bar{\eta}$, $\frac{d}{d\eta} F(\eta)$ is negative. At $\bar{\eta}$ there exists a local minimum of the objective function and for $\eta > \bar{\eta}$ the objective function is increasing. Thus there is some pseudoconvex region for $\eta \in [0, \tilde{\eta}]$ or possibly on $\eta \in [0, \infty]$. For an example, see Figure 3·3.

To find an upperbound on the pseudoconvex region, the roots of $\nabla F(\eta)$ must again be investigated. The presence of a root where $\nabla F(\eta)$ transitions from positive

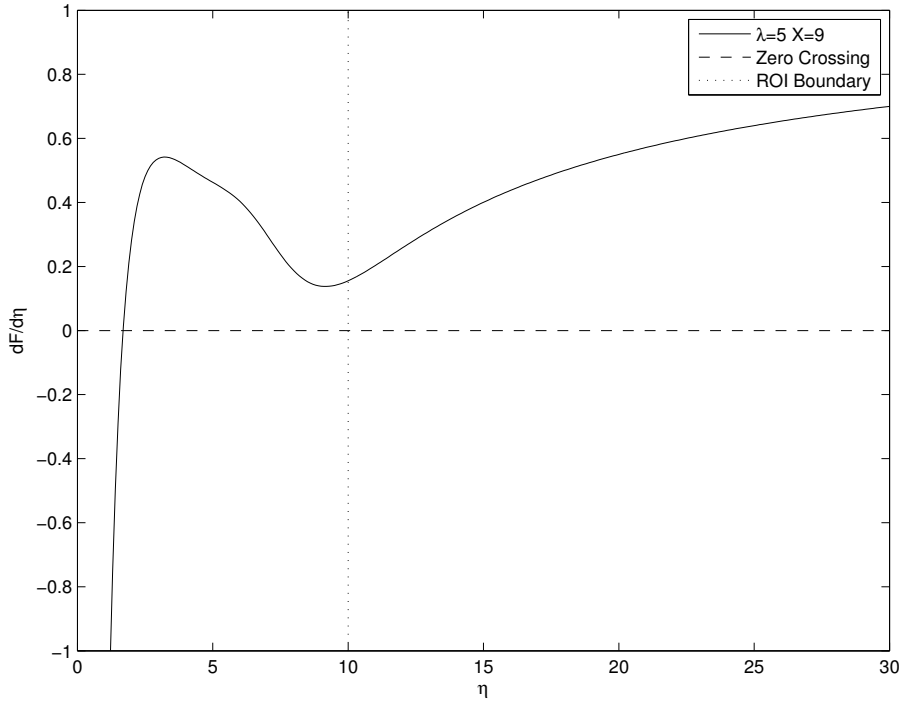

Figure 3·3: First derivative of the objective function with $\lambda = 5$, $Y = 9$

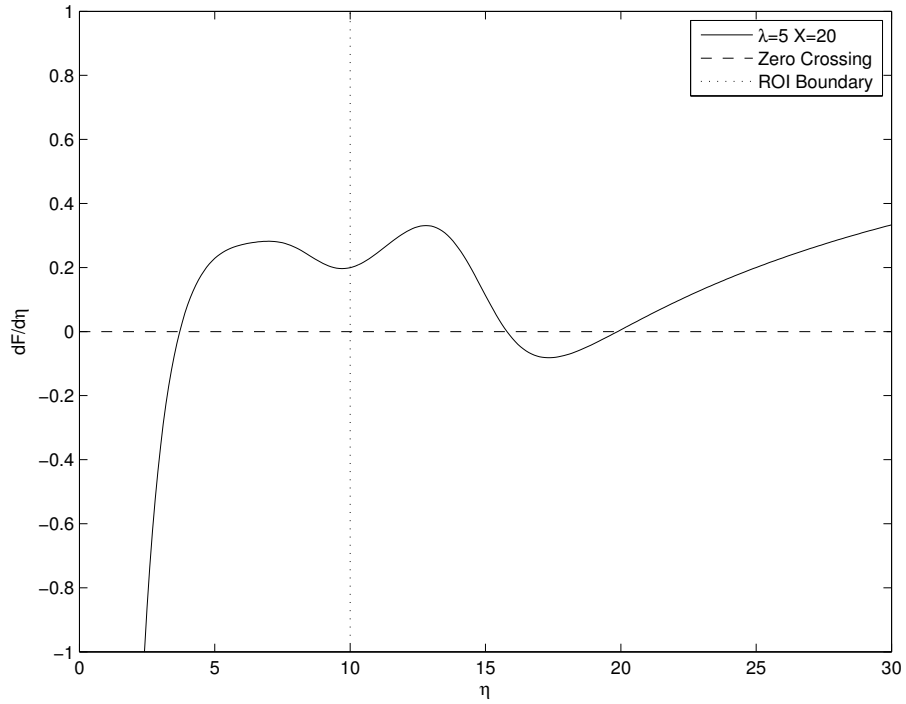

Figure 3·4: First derivative of the objective function with $\lambda = 5$, $Y = 20$

to negative indicates a local maximum. Thus the upper boundary on the pseudoconvex region $\tilde{\eta}$ must occur at such a root. Because the first root of $\nabla F(\eta)$ occurs at a transition from negative to positive, the second lowest root serves as the upper boundary of the pseudoconvex region. See Figure 3·4 for an example of $\nabla F(\eta)$ with multiple roots.

In addition to searching for the second largest root, another search was performed for roots occurring below $\eta\lambda$. It was assumed that for some number of counts measured, the root sought would more likely be at roughly $\eta = Y/\lambda$. Any roots forming a reasonable amount below this point were also deemed to effectively bound the pseudoconvex region. An example is shown in Figure 3·5. In this case, the $\lambda = 3$ and the number of counts $Y = 60$. It is expected that a root will form near $\eta = 20$ and indeed one does appear. However, two roots exist below this value as well. As such

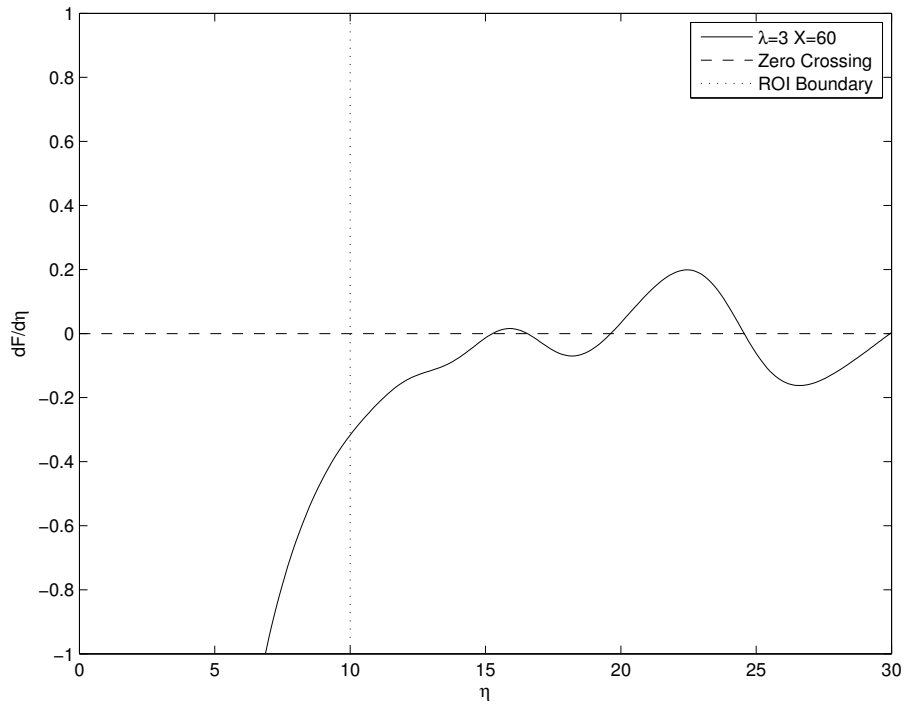

Figure 3·5: First derivative of the objective function with $\lambda = 3$, $Y = 60$

the upper boundary on the pseudonconvex region is considered to be the lowest root.

The region of pseudoconvexity was computed empirically using Matlab. For a range of possible count values and over the range of λ values used in the simulated imaging experiment, the values of the function $\nabla F(\eta)$ in equation (3.10) was computed. For each value of λ, the number of counts was varied and the upper boundary for each case was recorded. The minimum value of this upper boundary was recorded.

Figure 3·6 shows the upper boundaries computed for η over which the function is pseudoconvex. Values of (3.10) were computed for η on $[0, 15]$. The number of counts was varied from $Y = 0, 1, \ldots, 100$ over a range of λ on $[0, 10]$. For the simulated imaging experiment we considered values of iSE yields to be valid up to $\eta = 10$, which corresponds with measured iSE yields in HIM. Thus we consider the region of pseudoconvexity to include values of mean ions $\lambda \geq 3$ and $\eta \in [0, 10]$. This region lies

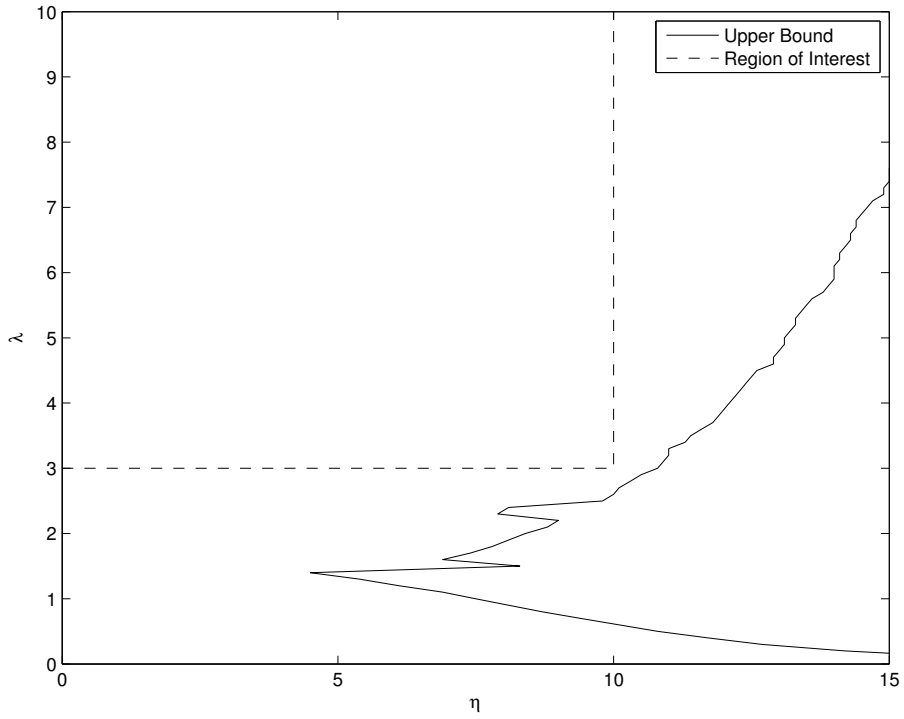

Figure 3·6: Boundaries for η

within the boundary of pseudoconvexity and thus the use of subgradient methods in justified for the HIM image formation estimation problem.

Chapter 4

Estimation Algorithm

4.1 Introduction

In traditional HIM imaging, the randomness of the helium ion beam establishes a lower limit to the ion dose used in the imaging process. Defining the signal-to-noise ratio (SNR) as the ratio of the mean to the standard deviation, the SNR becomes

$$SNR = \frac{\mu}{\sigma} = \frac{\sqrt{\lambda\eta}}{\sqrt{1+\eta}}. \tag{4.1}$$

To suppress noise, an HIM user might simply increase the ion dose, at the expense to damage to the sample. The image is then formulated directly from the iSE signal at each pixel.

By using estimation, image formation can be performed at lower ion doses while maintaining image quality. The Neyman Type A model described in Chapter 3 will be used as part of a regularized ML estimation procedure that estimates the iSE yield at every pixel. Regularization will be applied by assuming that sample sparsity can be accurately captured by the ℓ_1 norm of the wavelet transform. By penalizing complexity in the wavelet domain, an acceptable solution must thus be sparse in the wavelet transform.

By taking into account randomness in the beam, the need for higher ion doses is suppressed. In addition, the mean iSE yield $\eta_{i,j}$ is taken to be a parameter being estimated every pixel. This change in image formation both drastically improves the

performance at low ion doses and reformulates the imaging procedure to represent the reconstruction of a physical parameter. Traditional imaging relies on the variation of iSE yields to create contrast changes in the iSE signal. The estimation procedure presented here instead estimates the iSE yield values at each pixel. As η is the parameter directly responsible for changes in image intensity, this approach directly attempts to generate an image of a physical property of the sample as opposed to a raw iSE signal.

The regularized ML estimation problem can be performed by solving the minimization problem

$$\begin{aligned} \hat{\eta} = \quad & \underset{\eta \in \mathbb{R}^{n \times n}}{\arg\min}\ F(\eta, Y, \lambda) + \tau \|W^T \eta\|_1 \\ & \text{subject to } \eta \geq 0. \end{aligned} \tag{4.2}$$

This formulation performs a regularized ML estimation subject to sparsity in the wavelet transform. It is assumed that $\|W^T \eta\|_1$ accurately captures spatial correlation within the scene. Thus a prior of wavelet transform sparsity is imposed on our estimation. The degree of regularization can be controlled using the τ parameter. While this parameter has no direct physical meaning, higher τ values result in smoother estimations. A τ value of 0 results in a maximum likelihood estimation with no regularization.

The minimization problem shown in equation (4.2) is too difficult to be solved directly. Instead a numerical method must be adopted to find the solution $\hat{\eta}$. To solve this problem, an algorithm adapted from the Sparse Poisson Intensity Reconstruction ALgorithm (SPIRAL)(Harmany et al., 2012) was used. The SPIRAL framework was originally designed for reconstruction of signals under Poisson noise. Based originally on the SpaRSa framework (Wright et al., 2009), SPIRAL is capable of performing reconstructions under a variety of sparsity penalties. These include the total variation (TV) penalty, sparsity in the canonical basis, sparsity in some orthonormal basis of

choice, and recursive dyadic partition based penalties.

While the adaptation of the SPIRAL algorithm to the Neyman Type A distribution is capable of reconstructions under these penalties as well, concern will only be placed on sparsity penalty in an orthonormal basis, specifically the wavelet transform. The development of the algorithm and its convergence properties follows directly from the SPIRAL framework. Please see (Harmany et al., 2012) and a (Wright et al., 2009) for a full discussion of convergence. An abridged version is presented in Appendix A.

4.2 Overall Algorithmic Framework

To solve the minimization problem, a sequence of Taylor series approximations is used at each iteration. By minimizing these quadratic approximations of $F(\eta^k)$, the optimal image is approached iteratively. From the original minimization problem (4.2) a sequence of subproblems is solved. F^k denotes the quadratic approximation of $F(\eta^k)$ at a given iteration k.

$$\begin{aligned} \eta^{k+1} = \quad & \underset{\eta \in \mathbb{R}^n}{\arg\min}\, F^k(\eta) + \tau \|W^T \eta\|_1 \\ & \text{subject to } \eta \geq 0 \end{aligned} \tag{4.3}$$

$$F^k(\eta) = F(\eta^k) + (\eta - \eta^k)^T \nabla F(\eta^k) + \frac{\alpha_k}{2} \|\eta - \eta^k\|_2^2 \tag{4.4}$$

F^k is a second order approximation of the objective function at location η^k. In place of the Hessian, $\nabla^2 F(\eta^k)$ is replaced with a scaled version of the identity matrix $\alpha_k I$, where α_k contains information regarding the curvature of $F(\eta)$ at η^k.

By neglecting the constant term in (4.4), the subproblem can be reformulated as the equivalent subproblem

$$\begin{aligned} \eta^{k+1} = & \underset{\eta \in \mathbb{R}^n}{\arg\min}\, \phi^k(\eta) = \frac{1}{2} \|\eta - s^k\|_2^2 + \frac{\tau}{\alpha_k} \|W^T \eta\|_1 \\ & \text{subject to } \eta \geq 0 \end{aligned} \tag{4.5}$$

where

$$s^k = \eta^k - \frac{1}{\alpha_k}\nabla F(\eta^k). \tag{4.6}$$

This formulation can be recognized as a scaled gradient descent with regularization penalty. In the original SPIRAL algorithm, the step size α_k is chosen through a modified Barzilai-Borwein method. Setting $\delta^k = \eta^k - \eta^{k-1}$, the modified Barzilai-Borwein method chooses

$$\alpha_k = \frac{(\delta^k)'\nabla^2 F(\eta^k)\delta^k}{\|\delta^k\|_2^2}. \tag{4.7}$$

This approach has been modified for the pseudoconvexity of the Neyman Type A reconstruction problem. Because the objective function is non-convex, the Hessian takes negative values. Choosing α_k according to equation (4.7) can lead to negative values of α_k. Instead the method has been modified. Defining $\{\nabla^2 F(\eta^k)\}_+ = \max\{\nabla^2 F(\eta^k), 0\}$, where the maximum is taken elementwise, the stepsize is chosen as

$$\alpha_k = \frac{(\delta^k)'\{\nabla^2 F(\eta^k)\}_+\delta^k}{\|\delta^k\|_2^2}. \tag{4.8}$$

This uses curvature information to determine the stepsize parameter.

At each iteration k, the objective function ϕ^k is guaranteed to be monotonic up to some number M of previous iterations. Thus,

$$\phi(\eta^{k+1}) \leq \max_{i=[k-M]_+,\ldots,k} \phi(f^i) - \frac{\sigma\alpha_k}{2}\|\eta^{k+1} - \eta^k\|_2^2. \tag{4.9}$$

Should this criterion fail, α_k is multiplied by some predetermined parameter ν until the monotonicity criteria is satisfied. By setting $M = 0$, monoticity in the objective function is strictly enforced. The algorithm is summarized below in Algorithm 1.

Algorithm 1 Compound-SPIRAL

1: **Initialize** Choose η (α multiplier), $\sigma \in (0,1)$, $M \in \mathbb{Z}_+$, $0 < \alpha_{min} \leq \alpha_{max}$, choose initial solution η^0. Start iteration counter $k \leftarrow 0$.
2: **repeat**
3: choose $\alpha_k \in [\alpha_{min}, \alpha_{max}]$
4: $\eta^{k+1} \leftarrow$ solution of subproblem
5: **while** η^{k+1} does not meet termination criteria **do**
6: $\alpha_k \leftarrow \eta\alpha_k$
7: $\eta^{k+1} \leftarrow$ solution of subproblem
8: $k \leftarrow k+1$
9: **until** Acceptance criterion is reached

4.3 The Subproblems

A solution of the quadratic subproblems defined above in (4.3) will be derived in this section. The vector of wavelet coefficients is defined as $\theta \triangleq W^T\eta$. An equivalent subproblem can be expressed in terms of the wavelet transform coefficients of η.

$$\begin{aligned} \theta^{k+1} \triangleq \quad & \arg\min_{\theta\in\mathbb{R}^n} \phi^k(\theta) \triangleq \frac{1}{2}\|\theta - s^k\|_2^2 + \frac{\tau}{\alpha_k}\|\theta\|_1 \\ & \text{subject to } W\theta \geq 0 \end{aligned} \tag{4.10}$$

Note that the non-negativity of η must be guaranteed by enforcing the non-negativity of $W\theta$. This maintains the non-negativity of η.

Defining $u - v = \theta$, we eliminate the free variable θ and impose sign constraints on u and v.

$$\begin{aligned} (u^{k+1}, v^{k+1}) \triangleq \quad & \arg\min_{\theta\in\mathbb{R}^n} \phi^k(\theta) \triangleq \frac{1}{2}\|u - v - s^k\|_2^2 + \frac{\tau}{\alpha_k}\|u + v\|_1 \\ & \text{subject to } u, v \geq 0, \quad W(u - v) \geq 0 \end{aligned} \tag{4.11}$$

Though the number of variables is doubled, rewriting the subproblem in this form allows us to take the derivative of the objective function. This allows its Lagrangian

dual to be found.

$$L(u,v,\lambda_1,\lambda_2,\lambda_3) = \frac{1}{2}\|u-v-s^k\|_2^2 + \frac{\tau}{\alpha_k}\mathbb{1}^T(u+v) - \lambda_1^T u - \lambda_2^T v - \lambda_3^T W(u-v) \quad (4.12)$$

Begin by taking the gradient of the Lagrangian with respect to u and v.

$$\nabla_u L = u - v - s^k + \frac{\tau}{\alpha_k}\mathbb{1}^T - \lambda_1^T - W^T\lambda_3 \quad (4.13)$$

$$\nabla_v L = -u + v + s^k + \frac{\tau}{\alpha_k}\mathbb{1}^T - \lambda_2^T + W^T\lambda_3 \quad (4.14)$$

By setting each gradient equal to zero, the following optimality conditions arise:

$$u - v = s^k + \lambda_1 - \frac{\tau}{\alpha_k}\mathbb{1} + W^T\lambda_3 \quad (4.15)$$

$$\lambda_2 = \frac{2\tau}{\alpha_k}\mathbb{1} - \lambda_1. \quad (4.16)$$

These expressions will be used to derive the final form of the Lagrangian dual function. First substitute for λ_2.

$$L(u,v,\lambda_1,\lambda_2,\lambda_3) = \frac{1}{2}\|u-v-s^k\|_2^2 + \frac{\tau}{\alpha_k}\mathbb{1}^T(u+v) - \lambda_1^T u - (\frac{2\tau}{\alpha_k}\mathbb{1}^T - \lambda_1^T)v - \lambda_3^T W(u-v)$$

$$L(u,v,\lambda_1,\lambda_3) = \frac{1}{2}\|u-v-s^k\|_2^2 - (\lambda_1 - \frac{\tau}{\alpha_k}\mathbb{1} + W^T\lambda_3)^T(u-v) \quad (4.17)$$

Next substitute in for u and v.

$$L(\lambda_1,\lambda_3) = \frac{1}{2}(\lambda_1 - \frac{\tau}{\alpha_k}\mathbb{1} + W^T\lambda_3)^T(\lambda_1 - \frac{\tau}{\alpha_k}\mathbb{1} + W^T\lambda_3) - (\lambda_1 - \frac{\tau}{\alpha_k}\mathbb{1} + W^T\lambda_3)^T(s^k + \lambda_1 - \frac{\tau}{\alpha_k}\mathbb{1} + W^T\lambda_3)$$

$$L(\lambda_1, \lambda_3) = -\frac{1}{2}(s^k + \lambda_1 - \frac{\tau}{\alpha_k}\mathbb{1} + W^T\lambda_3)^T(s^k + \lambda_1 - \frac{\tau}{\alpha_k}\mathbb{1} + W^T\lambda_3) + \frac{1}{2}(s^k)^T(s^k) \quad (4.18)$$

And finally arriving at the final form of the Lagrangian dual function

$$g(\lambda_1, \lambda_3) = -\frac{1}{2}\|s^k + \lambda_1 - \frac{1}{2}\mathbb{1} + W^T\lambda_3\|_2^2 + \frac{1}{2}\|s^k\|. \quad (4.19)$$

In order to simplify this expression, define $\gamma = \lambda_1 - \frac{1}{2}\mathbb{1}$ and $\lambda \triangleq \lambda_3$. Also recall $\lambda_1, \lambda_2, \lambda_3 \geq 0$. From (4.16) we have

$$\frac{2\tau}{\alpha_k}\mathbb{1} - \lambda_1 \geq 0 \rightarrow 0 \leq \lambda_1 \leq \frac{2\tau}{\alpha_k}\mathbb{1} \quad (4.20)$$

$$-\frac{\tau}{\alpha_k}\mathbb{1} \leq \gamma \leq \frac{\tau}{\alpha_k}\mathbb{1} \quad (4.21)$$

Thus the Lagrange dual problem of the subproblems can be defined as

$$\begin{aligned} \underset{\gamma,\lambda\in\mathbb{R}^n}{\text{minimize}} \quad & h(\lambda, \gamma) \triangleq \frac{1}{2}\|s^k + \gamma + W^T\lambda\|_2^2 - \frac{1}{2}\|s^k\| \\ \text{subject to} \quad & \lambda \geq 0,\ -\frac{\tau}{\alpha_k}\mathbb{1} \leq \gamma \leq -\frac{\tau}{\alpha_k}\mathbb{1}. \end{aligned} \quad (4.22)$$

The reconstruction of θ^{k+1} from the optimal values of the dual is

$$\theta^{k+1} \triangleq u^{k+1} - v^{k+1} = s^k + \gamma^\star - W^T\lambda^\star. \quad (4.23)$$

To minimize the dual problem (4.22), use alternating minimization over γ and λ. Take the partial derivatives with respect to each variable and set them equal to zero.

$$\nabla_\lambda h(\lambda, \gamma) = W(s^k + \gamma + W^T\lambda) \qquad \rightarrow \qquad \lambda = -W(s^k + \gamma) \quad (4.24)$$

$$\nabla_\gamma h(\lambda, \gamma) = s^k + \gamma + W^T \qquad \rightarrow \qquad \gamma = -s^k - W^T\lambda \quad (4.25)$$

Extra insurance is needed to maintain λ and γ to the bounds set in (4.22). Definining the operator mid$\{a, b, c\}$ to chose the component-wise middle element and j as the

current iteration.

$$\begin{aligned} \gamma^{(j)} &= \text{mid}\{-\frac{\tau}{\alpha_k}\mathbb{1}, -s^k - W^T\lambda^{(j-1)}, \frac{\tau}{\alpha_k}\mathbb{1}\} \\ \lambda^{(j)} &= \big[-W(s^k+\lambda)\big]_+ \end{aligned} \tag{4.26}$$

Equation (4.26) gives the alternating calculations performed at each subiteration.

Chapter 5

Simulated Results

5.1 Experiment

To demonstrate the effectiveness of the image formation algorithm, simulated HIM images at different numbers of mean impinging ions were generated. Using these low dose images, the image formation technique according to the two-parameter Neyman Type A distribution will be compared to a naive regularization assuming Poisson statistics.

First ground truth images were generated from an actual HIM image. It was assumed that the pixel intensities in the HIM image were linearly related to the true secondary electron yield values at each pixel. To create a ground truth image of secondary electron yields, the pixel values in this original HIM image were linearly scaled to a range of yield values deemed suitable for the experiment. The image selected is shown in Figure 5·1. The image is of gold nanoparticles on silicon.

From the original image a 256 by 256 pixel section of the lower left hand corner was selected to form the ground truth image. Two different ground truth images were generated for the experiment using the same original HIM image. For one image, pixel intensities $I_{i,j} \in [0, 255]$ were scaled to $\eta_{i,j} \in [0, 10]$, while in the second image $\eta_{i,j} \in [0, 5]$ was used. The choice of the range of η values was motivated by real world values of secondary electron yields in the HIM. The ground truth image is shown in Figure 5·2.

After ground truth images were produced, images at low helium ion doses were

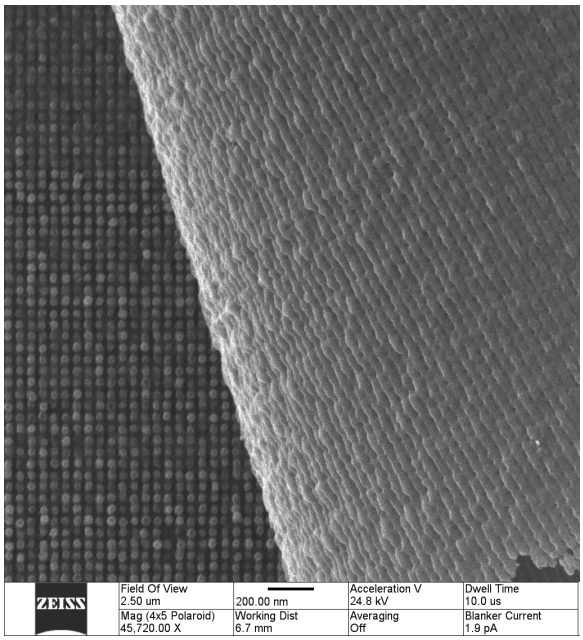

Figure 5·1: Original HIM image of gold nanoparticles

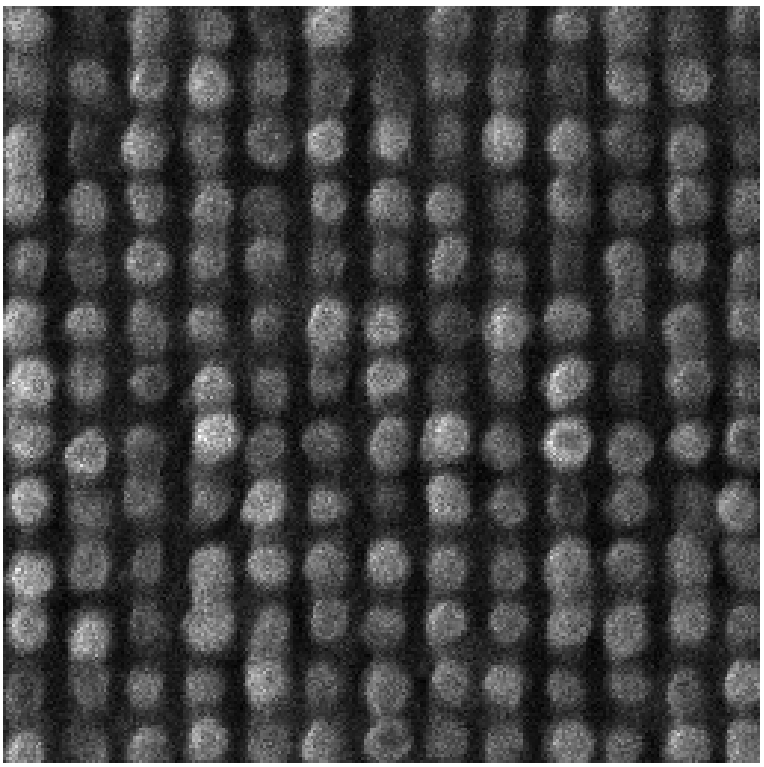

Figure 5·2: Ground truth image

generated. A range of different λ values were used. It was assumed that by varying the dwell time and beam current, the value of the parameter λ can be accurately controlled by the microscope user. The values discussed in the following section are $\lambda = 3, 5, 7$. These values for λ fall into the pseudoconvex region where the estimation algorithm is well behaved. At each value of λ, a series of emission matrices are generated with $M_{i,j} \sim \text{Poisson}(\lambda)$. The matrix of iSE counts is then generated according to $Y_{i,j} \sim \text{Poisson}(m_{i,j}\eta_{i,j})$. This generates the Neyman Type A distributed image in parameters λ and η.

The generated images of iSE counts were then used in the regularized ML estimation algorithm according to our Neyman Type A method and the naive assumption of Poisson statistics. The SPIRAL algorithm was used to perform the regularization under the Poisson assumption. Under the naive Poisson assumption, it is assumed that each point in the electron count matrix is governed by $Y_{i,j} \sim \text{Poisson}(\lambda\eta_{i,j})$. Both regularization schemes were initialized with $\eta_0 = Y/\lambda$. For both methods, a Daubchies length 4 wavelet transform penalty was used. For both methods, the penalization parameter τ was varied by 0.1 and a series of images were created.

5.2 Results

The two regularization techniques were compared both on mean squared error (MSE) performance and in terms of visual quality. Figure 5·3 shows the effect of the regularization parameter τ on the MSE for estimation using the two distributions and over the two different ranges of η. As τ was increased the MSE dropped sharply, leveled out to some degree, and then began to increase. For each regularization technique, the best performance in terms of MSE was about the same.

Figure 5·4 shows MSE for unregularized ML estimation for different values of λ. It should be noted that for an unregularized estimation, i.e. one for which $\tau = 0$, the

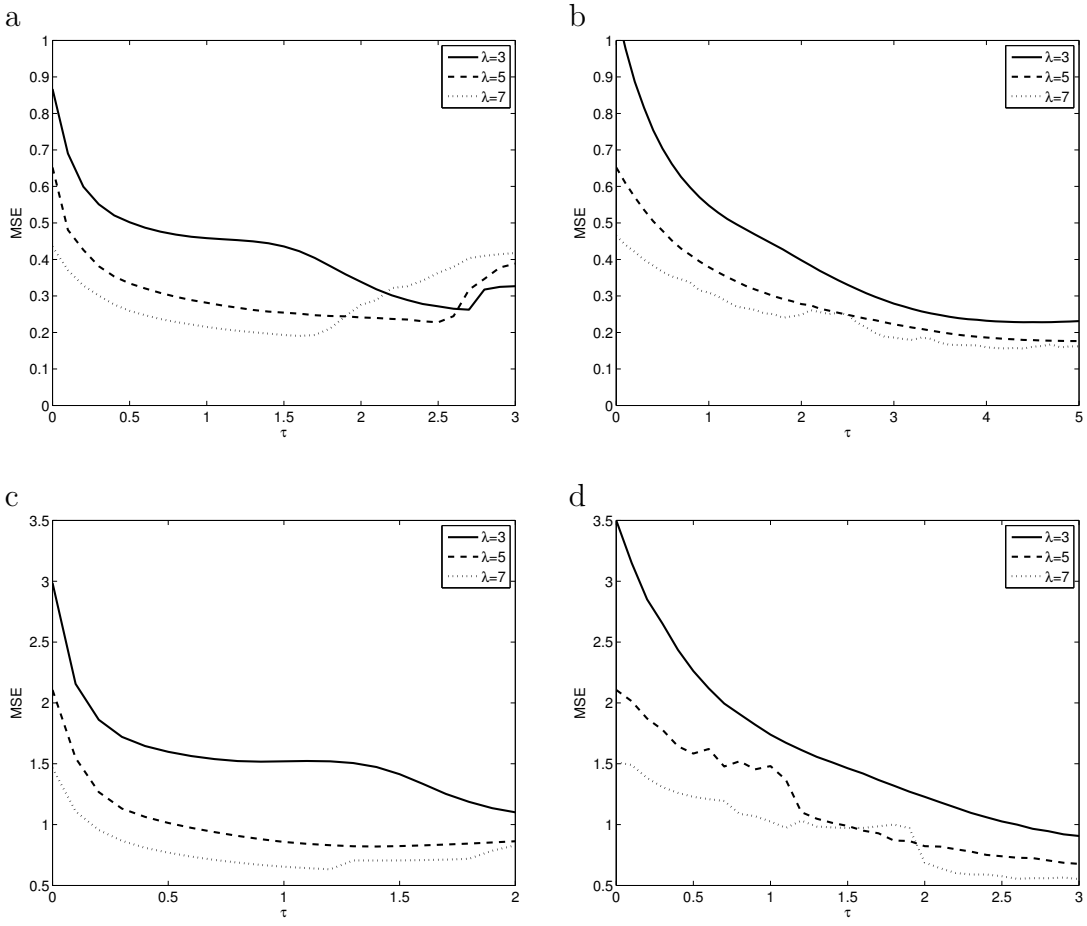

Figure 5·3: MSE vs. τ for (a) our method for η on $[0, 5]$ (b) naive SPIRAL method η on $[0, 5]$ (c) our method for η on $[0, 10]$ (d) naive SPIRAL method η on $[0, 10]$

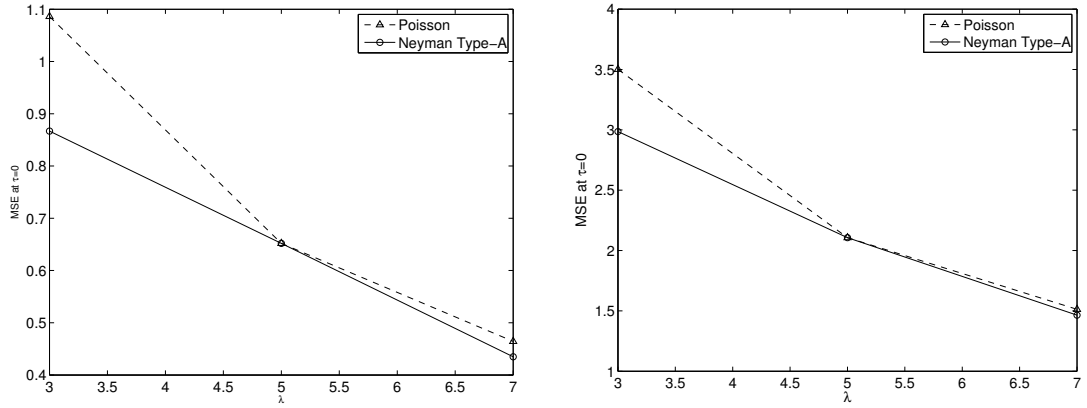

Figure 5·4: Unregularized ML MSE vs. λ for η on $[0, 5]$ (left) and η on $[0, 10]$ (right)

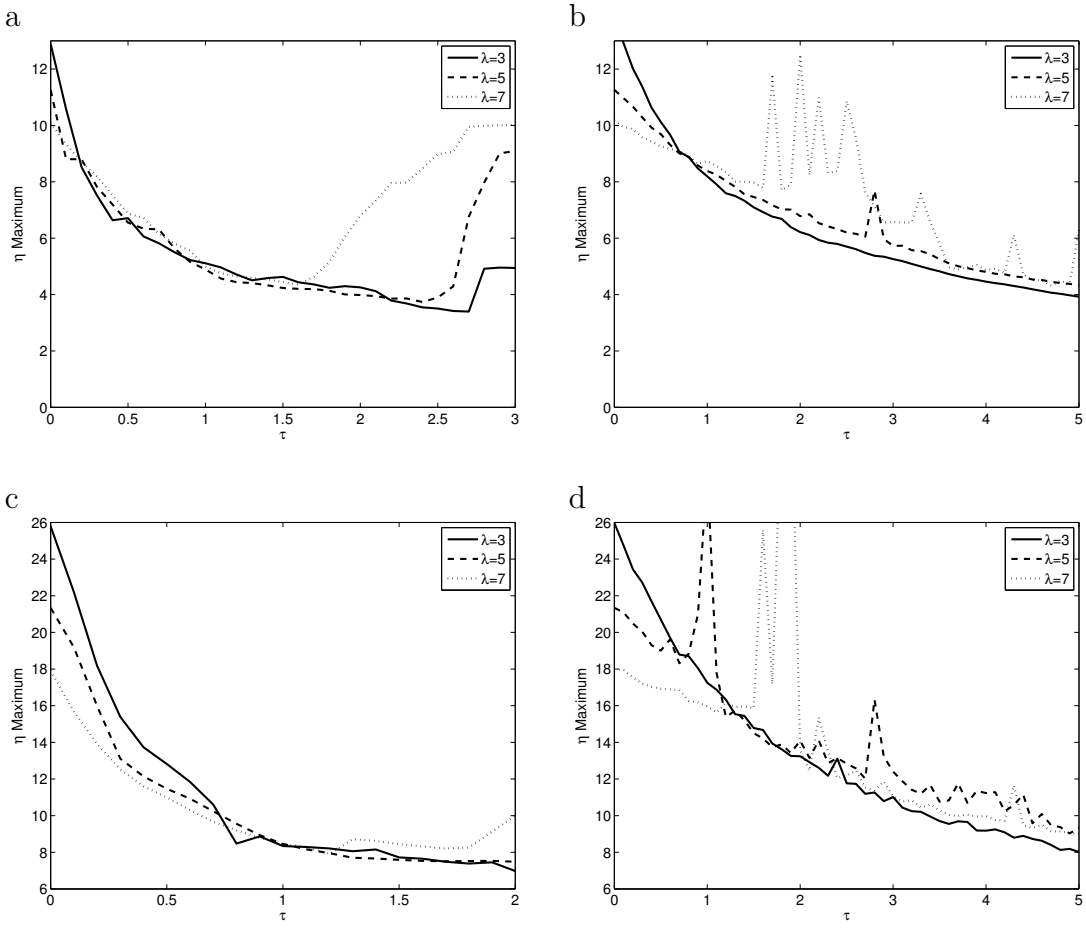

Figure 5·5: Maximum values of $\hat{\eta}$ vs. τ for (a) our method for η on $[0, 5]$ (b) naive SPIRAL method η on $[0, 5]$ (c) our method for η on $[0, 10]$ (d) naive SPIRAL method η on $[0, 10]$

Neyman Type A ML estimation outperforms the naive Poisson ML estimation. This improvement reinforces the Neyman Type A model choice for image formation. The gain in performance for estimation that does not use spatial information shows that the Neyman Type A distribution more accurately represents image formation in the HIM.

At the best performing MSE value of τ, the resulting image is oversmoothed. This was true for both the Poisson and Neyman Type A estimations. In order to compare estimation under the two distributions, a Neyman Type A image with good visual qualities and reasonable MSE performance was selected. As a point of comparison, a Poisson regularized image with similar MSE was chosen. This strategy allowed for the comparison the estimation techniques both as an estimator of $\hat{\eta}$ and in terms of image quality.

The resulting images are displayed in Figure 5·6 and Figure 5·7 for η ranges on $[0, 5]$ and $[0, 10]$, respectively. The ground truth image is shown in Figure 5·2 for visual comparison. The reconstructions are presented scaled on $[\eta_{min}, \eta_{max}] \rightarrow [0, 1]$ for each reconstruction.

The Neyman Type A estimation technique better preserves the contrast of the original image. In the naive Poisson regularizations, the images are duller and have poorer contrast characteristics than the Neyman Type A method. This is generally caused by the Poisson estimation's inability to account for differing numbers of ion strikes. A pixel with a large number of ion strikes will generally result in a large number of iSE counts. The Poisson regularization will assign this pixel a high $\hat{\eta}$ value because of the assumption that all pixels are impinged upon by the same number of ions.

The Neyman Type A regularization is better suited to handling differing numbers of ion strikes. The longer tail of the Neyman Type A distribution makes the distri-

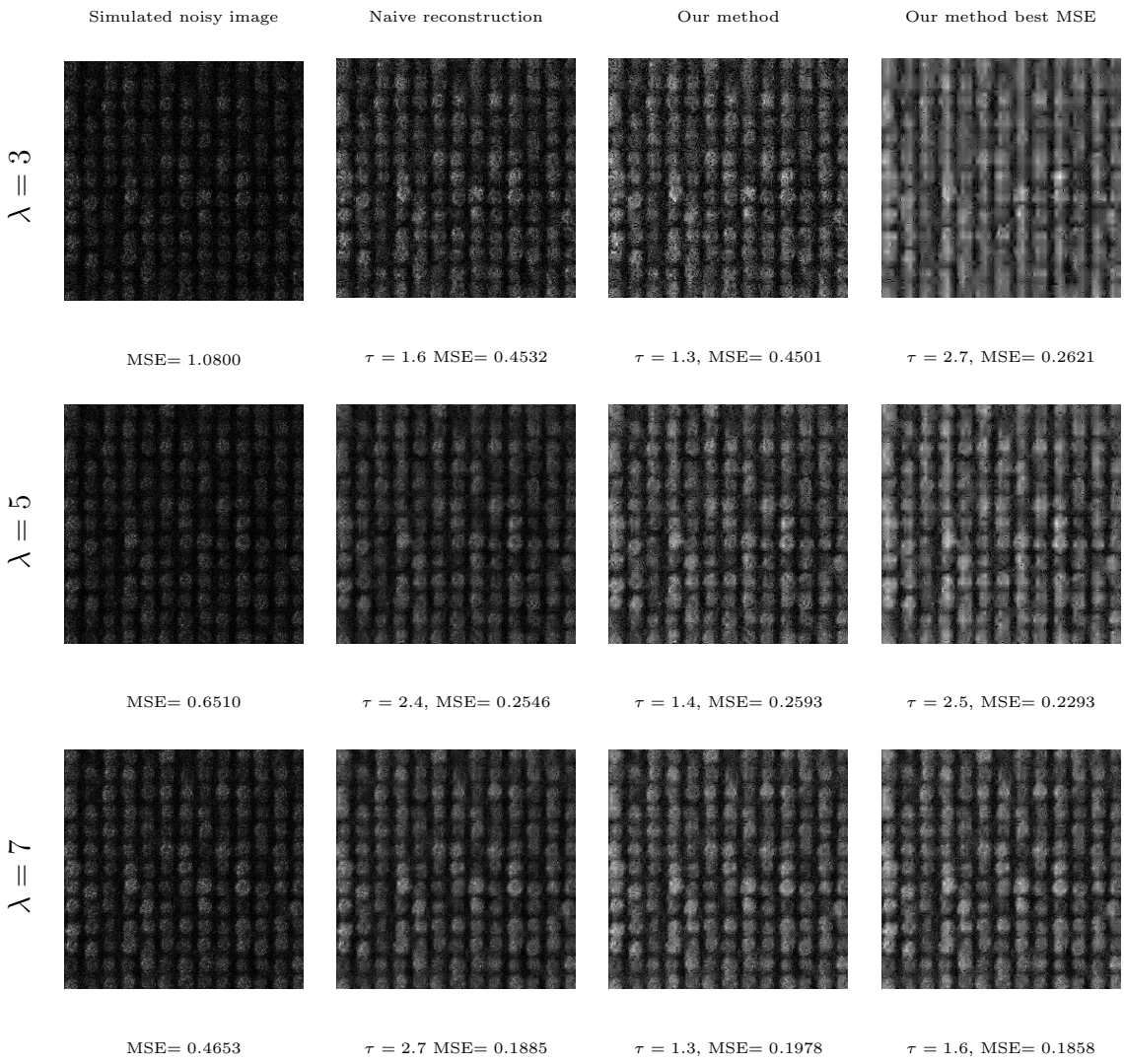

Figure 5·6: Noisy and regularized images for η scaled on $[0, 5]$

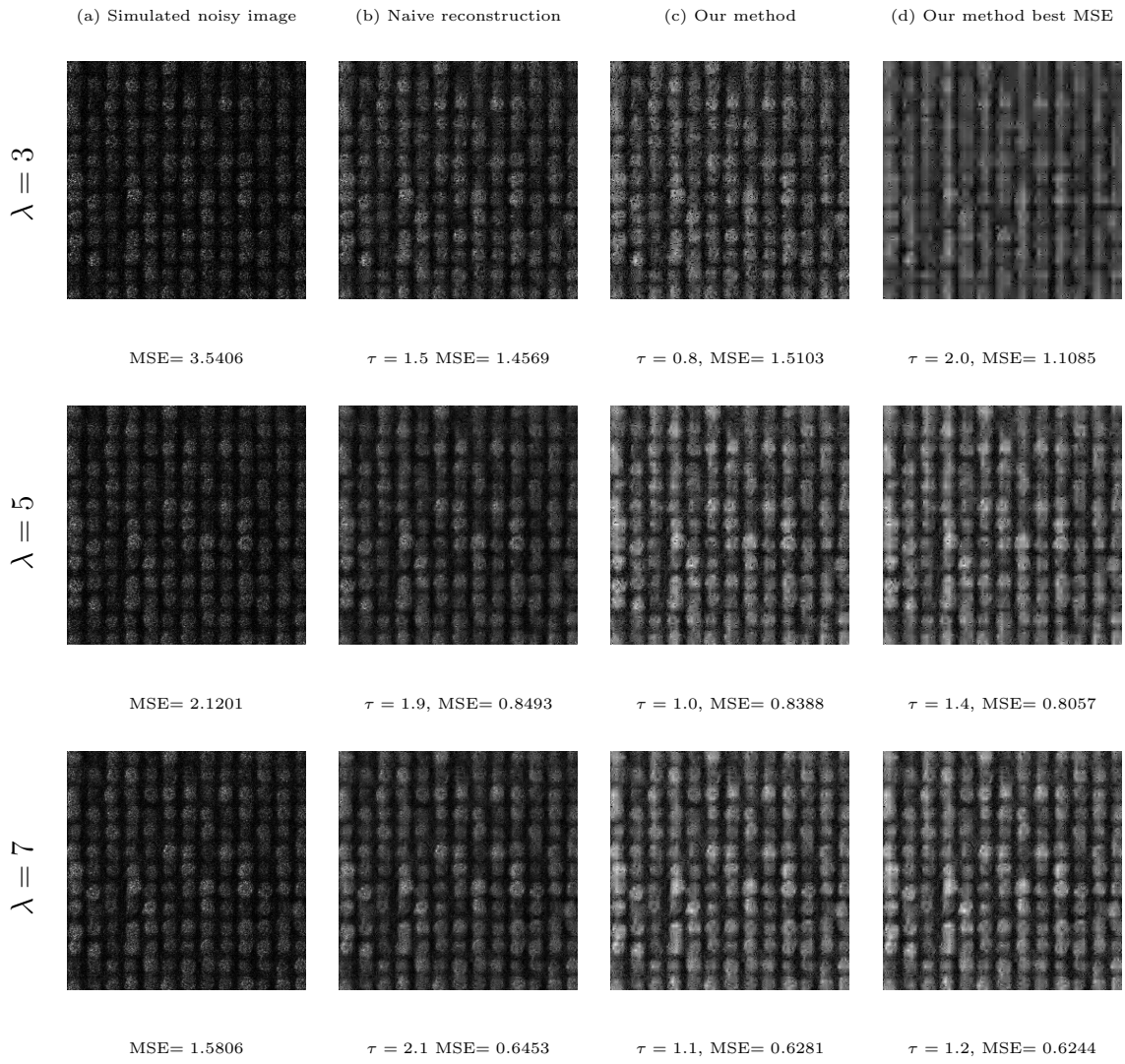

Figure 5·7: Noisy and regularized images for η scaled on $[0, 10]$

bution more able to account for these differences in impinging ions between pixels. The effect of regularizing for spatial correlation is less influenced by very large or very small differences between neighboring pixels. This effect can be seen in Figure 5·5. Estimation under Neyman Type A statistics brings $\hat{\eta}$ to within the predetermined range of ground truth η values more quickly than Poisson regularization. The higher maximum values under the Poisson regularization degrade image contrast and result in a poorer quality image.

5.3 Summary

Through this experiment, the gains from the novel image formation method are apparent. In a scenario using no regularization the Neyman Type A estimator outperforms a naive Poisson estimator. This reinforces the choice of model and demonstrates the gains possible even without including spatial information in the estimation process.

When regularization is included, the performance of the two estimators in terms of best MSE is about the same. However, this result is misleading as the images produced are so oversmoothed that they are useless as visual representations of the scene. Thus some amount of MSE performance must be traded for visual quality.

When comparing regularized Neyman Type A estimation and regularized naive Poisson estimation, the Neyman Type A estimation has greater visual quality when compared with a Poisson estimation at the same MSE. This is because of the Neyman Type A distribution's ability to account for the randomness in the amount of helium ions emitted. The Neyman Type A estimator can more easily use spatial information to bring the value of $\hat{\eta}_{i,j}$ to a correct range. The Poisson distribution is unable to account for variation in the beam and therefore makes much higher estimates of $\hat{\eta}_{i,j}$ at some pixels. These high values degrade contrast over the entire image. The resulting naive estimation is lacking in visual quality when compared to the Neyman Type A

estimation.

Chapter 6

Experimental Results

6.1 Introduction

In real world HIM imaging, the image presented to the user is not a direct representation of iSE counts. The image represents the some manipulated version of the iSE signal that may be subject to scaling, clipping, gamma correction, noise in the detector, and other effects. Thus in order to apply the Neyman Type A estimation procedure developed in Chapter 4 to real world data, iSE counts must be derived from HIM images.

This chapter details an experiment to apply the Neyman Type A estimation method to real world HIM images. A series of images at different beam currents and dwell times were taken, resulting in images with a range of λ values. Using a section of each image that was assumed to be flat and of a constant iSE yield, image intensities were compared to the PMF of the Neyman Type A distribution. Through this comparison, iSE counts were derived for each image. Using these new images representing iSE counts, the Neyman Type A estimation procedure could be applied.

Because of the lack of data regarding the actual iSE counts measured and the iSE yields for the sample used, quantitative analysis of the results was not possible. Instead, the results presented were judged based on visual quality alone. While the procedure developed here offers insight into the HIM, further experimentation will require more direct access to the data.

6.2 Image Acquisition

The Zeiss Orion HIM was used to acquire a series of images at a variety of dwell times and beam currents. The first objective was to infer iSE count information from a series of images taken at very low mean ions. In total, 30 images were taken at each setting. The number of mean ions per pixel was much lower than those values typically used in HIM imaging. A sample of silicon marked with a diamond-tipped scribe was used. A sample image taken at a dwell time of 1000ns with a beam current of 823fA is shown in Figure 6·1. Note the noise present in the image, as this image reflects a value of λ of roughly 5.137. All images used are 512 by 512 pixels.

While the Zeiss Orion microscope is capable of beam currents below 1pA, the standard beam current measuring system in the microscope is unable to measure the low currents needed for our experiment. To accurately measure the beam current, a transimpedance amplifier was connected to the beam blanker measurement port. The transimpedance amplifier was then connected to an oscilloscope where beam currents were measured for each series of images. The beam currents and dwell times used are detailed in Tables 6.1, 6.2, and 6.3. For each set of beam currents I_B and dwell times t_D, the λ mean ion parameter was calculated.

$$\lambda = \frac{I_B \cdot t_D}{q_{He}} \tag{6.1}$$

A sample calculation for a beam current of 115fA and a dwell time of 500ns is shown below in equation (6.2).

$$\lambda = \frac{115fA \cdot 500ns}{q_{He}} = \frac{115 \times 10^{-15} A \cdot 500 \times 10^{-9} s}{1.602 \times 10^{-16}} = 0.3277 \tag{6.2}$$

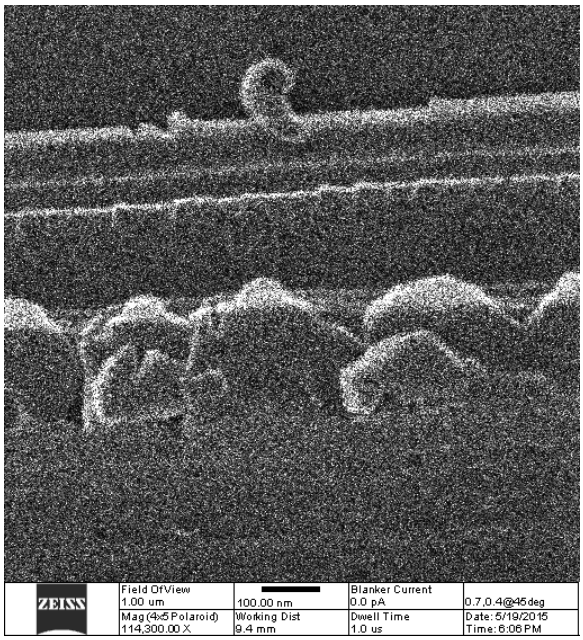

Figure 6·1: Noisy image of the silicon sample taken at 1000ns dwell time with beam current 823fA

Dwell Time (nS)	100							
Beam Current (fA)	115	197	300	400	500	590	700	823
λ (ions per pixel)	0.0717	0.1229	0.1872	0.2496	0.3121	0.3682	0.4369	0.5137

Table 6.1: Beam currents and λ parameter for 100ns dwell time

Dwell Time (nS)	500							
Beam Current (fA)	105	182	280	390	490	592	695	823
λ (ions per pixel)	0.3277	0.5680	0.8739	1.217	1.529	1.847	2.169	2.568

Table 6.2: Beam currents and λ parameter for 500ns dwell time

Dwell Time (nS)	1000							
Beam Current (fA)	105	182	280	390	490	592	695	823
λ (ions per pixel)	0.6554	1.136	1.747	2.434	3.058	3.685	4.338	5.137

Table 6.3: Beam currents and λ parameter for 1000ns dwell time

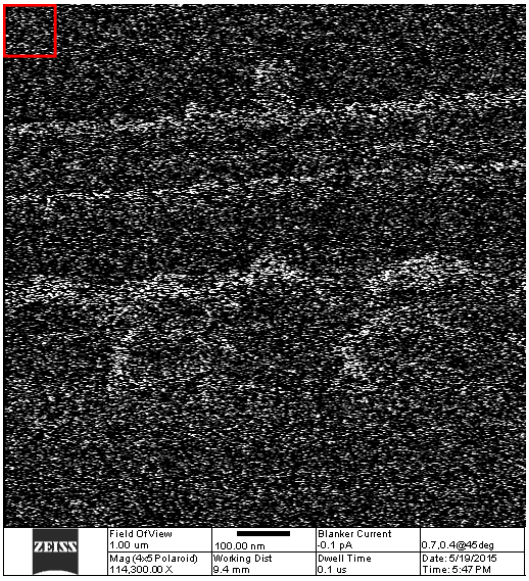

Figure 6·2: 100ns 400fA image with flat silicon region shown in red square

6.3 Procedure

Because of non-linearities in the ET detector and in the image generation software of the HIM, it was impossible to directly infer iSE counts from the pixel intensities. Instead, a small region of each image was compared to a Neyman Type A distribution. A fit of the image intensities to specific numbers of iSE counts was performed based on the distribution. A 50 by 50 pixel region of the upper left corner of each image was cropped. This region consisted of flat silicon that had not been marked by the diamond-scribe. An example is shown in Figure 6·2. It was assumed that this region was perfectly flat with a constant iSE yield $\eta = 0.63$. This value was selected from (Joy, 2013) and represents an experimental of the iSE yield of flat silicon at a beam energy of 30 keV.

This region was examined for all images in the series of 30 images for each beam current and dwell time setting. A histogram of image intensity values was generated for the combined 30 images in the series. An example for 100ns dwell time and 400fA

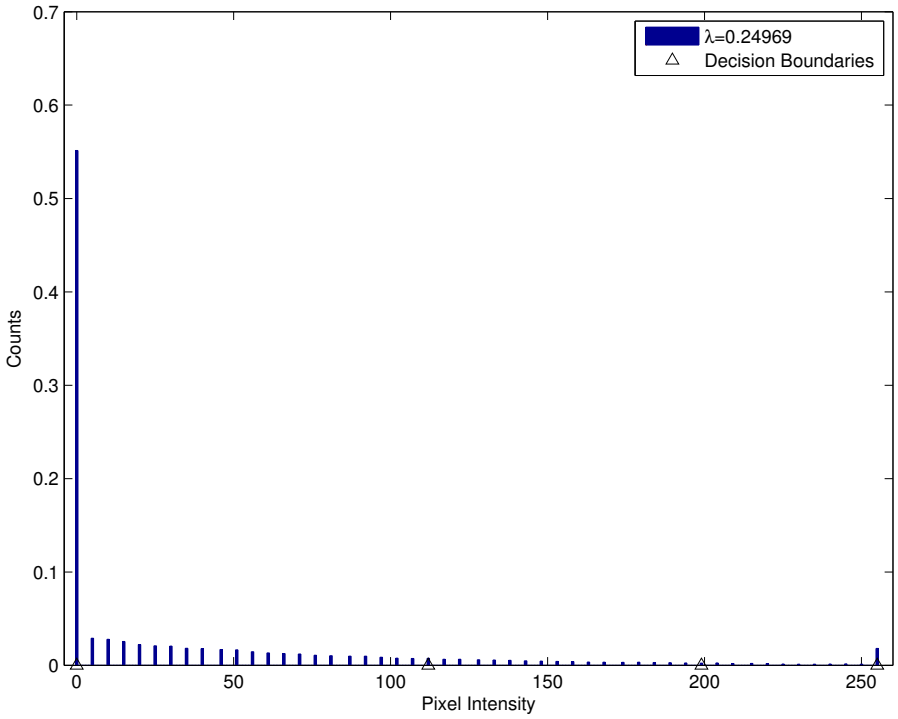

Figure 6·3: 100ns 400fA small region image intensity histogram

beam current is shown in Figure 6·3. The number of counts at each intensity was divided by the total number of counts to generate an empirical distribution of image intensity values over the flat region.

Using the λ value calculated from equation (6.1) and $\eta = 0.63$ given by (Joy, 2013), a PMF and CMF of the Neyman Type A distribution was calculated. It was assumed that the actual numbers of iSE generated in the flat region followed this distribution. It was also assumed that the image intensities at each pixel represented some scaling of numbers of iSE generated. Thus based on the Neyman Type A PMF, some percentage of each iSE count would be present in the image of the small region. Ranges of image intensities were then assigned to specific iSE count values based on equal mass between the empirical and calculated CMFs. An example of this assignment procedure is shown in Figure 6·4. The decision boundaries for this assignment are shown in Figure 6·3 as well. Thus the image intensities on $[0, 255]$

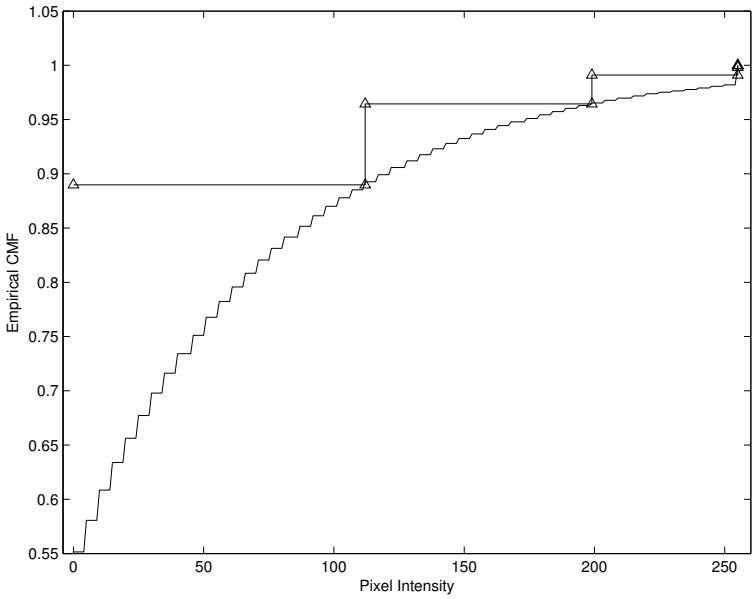

Figure 6·4: Empirical image intensity CMF shown with steps of Neyman Type A CMF values. Each step represents a change in iSE count assigned.

were mapped to positive integer values representing iSE counts.

The decision regions calculated from the flat region of the sample were then applied to the rest of the image. Through this process, all 30 images of the series were mapped to integer counts of iSE. An example of one such image is shown in Figure 6·3. This image is the mapped version of Figure 6·1.

These individual images represent in only a very small number of mean ions per pixel. In order to generate images with higher numbers of mean ions suitable for estimation, some number of images were added together to create composite images with higher values of λ. An example of such an image is shown in Figure 6·6. Composite images with roughly integer values of λ were sought in order to accommodate the integer values of λ used in the estimation procedure. For this example image, 24 individual images at an original $\lambda = 0.2496$ were used to create a composite image with $\lambda = 5.9925$.

The Neyman Type A estimation procedure developed earlier in this paper was

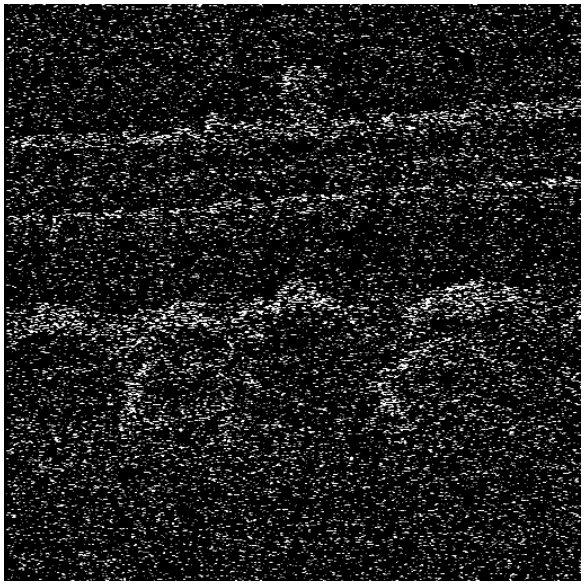

Figure 6·5: Thresholded image representing iSE counts

then used to estimation iSE yield values for the composite images. An integer range of λ values for a set of beam currents and dwell times were used. The value of the regularization parameter τ was varied over a range to produce images with differing degrees of regularization. An example regularized image is shown in Figure 6·7. The results are detailed in the following sections.

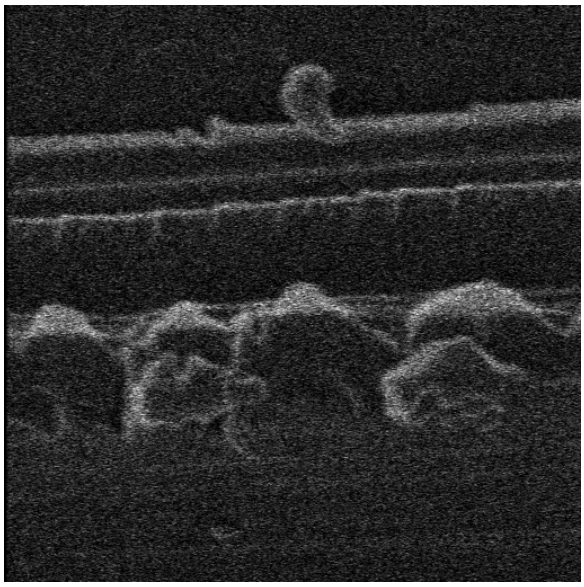

Figure 6·6: Composite image for 100ns 400fA with λ roughly equal to 6

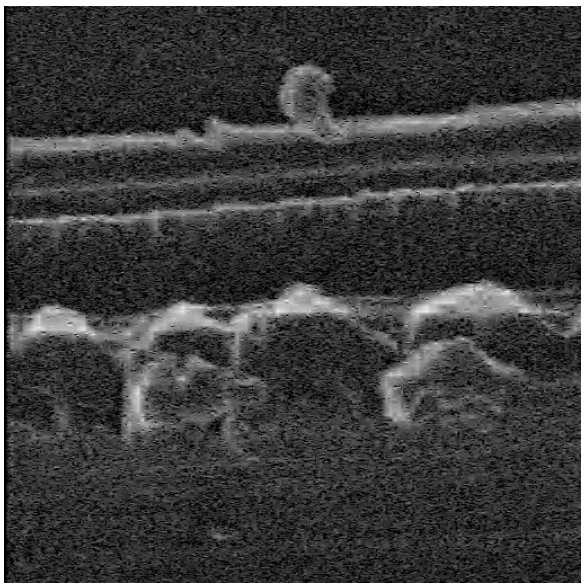

Figure 6·7: Regularized image for 100ns 400fA with λ roughly equal to 6 for $\tau = 1.4$

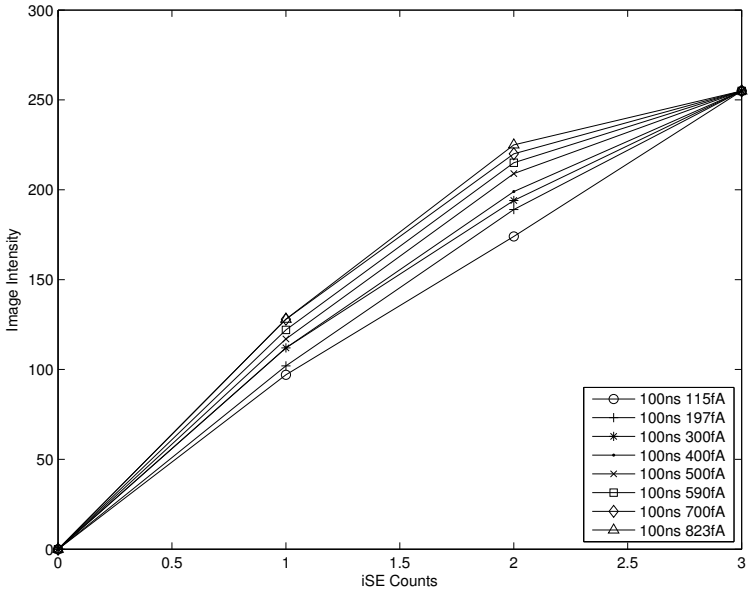

Figure 6·8: Decision boundaries for 100ns dwell time

6.4 Data Fitting Results

Computed decision boundaries are shown in Figures 6·8, 6·9, and 6·10 for dwell times 100ns, 500ns, and 1000ns respectively. For each beam current and dwell time setting, the computed decision regions appear to be roughly linear in image intensity. This is especially true for 500ns dwell time, where decision boundaries were clustered close together in a linear fashion for varying beam currents. For 100ns, the decision boundaries are monotonically increasing with beam current. However, for 1000ns boundaries seem to decrease as beam current is increased at higher numbers of iSE counts. This is perhaps due to clipping in the image formation process. The highly linear nature of the computed quantization regions reinforces the choice of the Neyman Type A model for image formation. Furthermore, the clustering of the decision boundaries at specific locations also reinforces the choice of model.

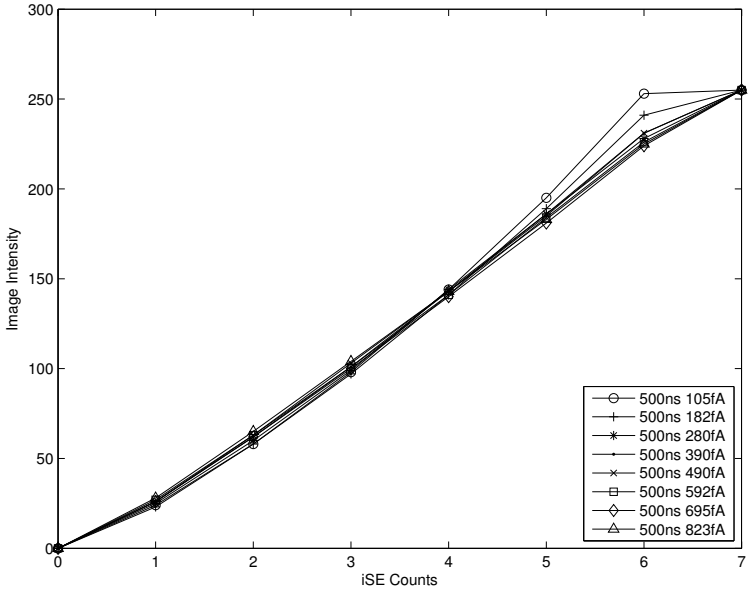

Figure 6·9: Decision boundaries for 500ns dwell time

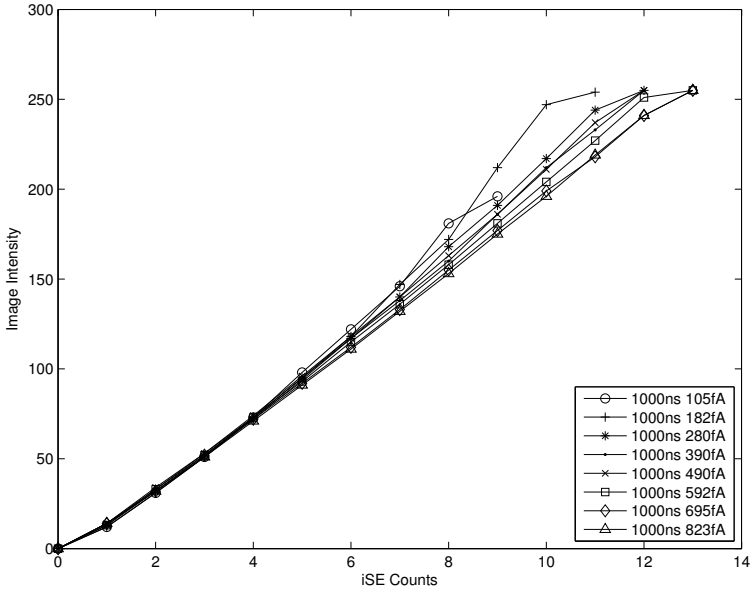

Figure 6·10: Decision boundaries for 1000ns dwell time

Average and composite images for values of $\lambda \in \{1, 3, 5, 7\}$ are shown in Figures 6·11-6·14, 6·15-6·18, 6·19-6·22, and 6·23-6·26 for dwell times and beam currents of 100ns at 400fA, 100ns at 500fA, 500ns at 105fA, and 500ns at 182fA respectively. These particular beam current and dwell time settings were selected based on the number of mean ions for a single image. These settings had λ values for individual images that were capable of approximating integer numbers of λ when combined. The average images were created simply by averaging intensity values at each pixel while the composite images have first been mapped to iSE counts and then added together. The images show an improvement in contrast due to the mapping process for dwell times of 100ns.

This improvement in image quality also reinforces the choice of imaging model. The average images loosely correspond to a similar philosophy used in traditional HIM imaging. In the average images, the iSE signal is simply added over the course of the images. Noise is suppressed by the greater number of helium ions at each pixel. The composite images represent direct measures of iSE counts. The increased contrast represents the removal of noise in the ET detector and restoration of the iSE count signal.

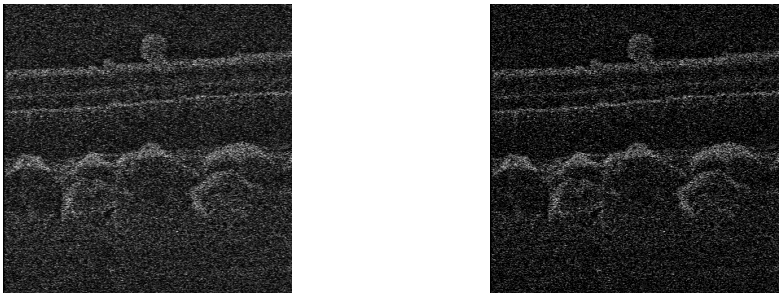

Figure 6·11: Average and composite images for 100ns dwell time at 400fA beam current, $\lambda = 1$

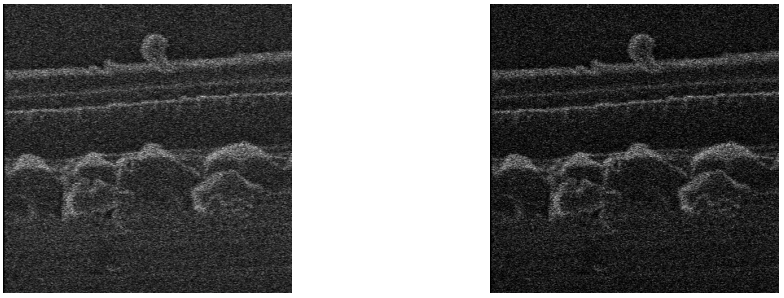

Figure 6·12: Average and composite images for 100ns dwell time at 400fA beam current, $\lambda = 3$

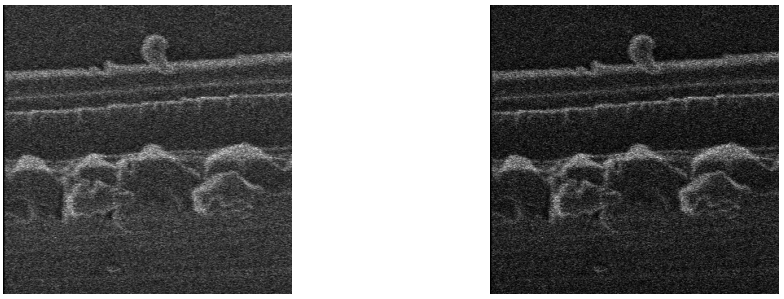

Figure 6·13: Average and composite images for 100ns dwell time at 400fA beam current, $\lambda = 5$

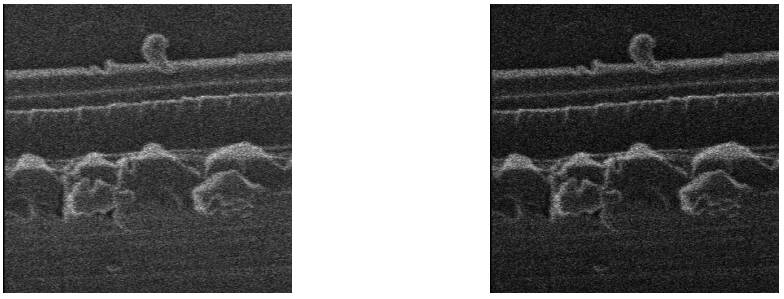

Figure 6·14: Average and composite images for 100ns dwell time at 400fA beam current, $\lambda = 7$

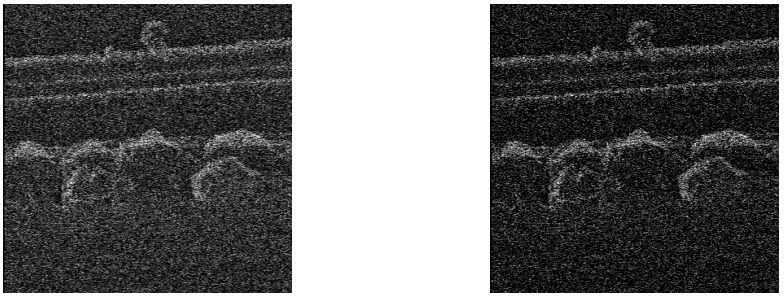

Figure 6·15: Average and composite images for 100ns dwell time at 500fA beam current, $\lambda = 1$

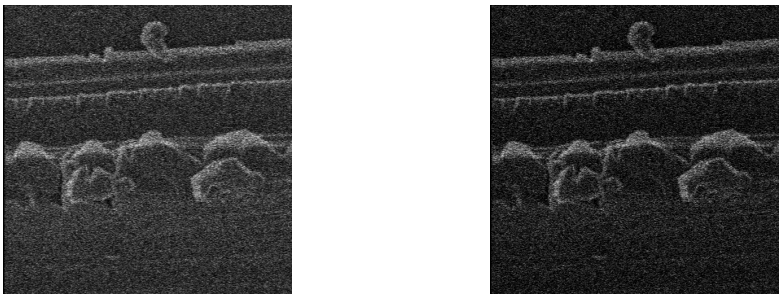

Figure 6·16: Average and composite images for 100ns dwell time at 500fA beam current, $\lambda = 3$

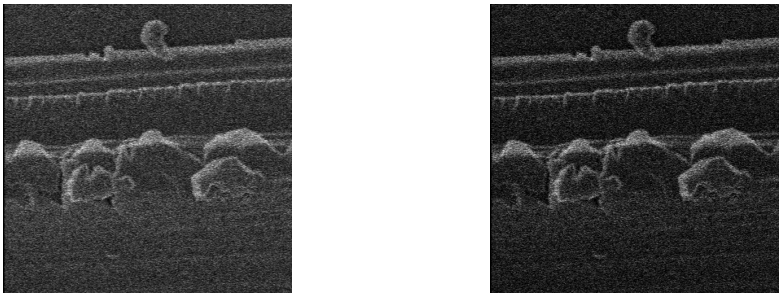

Figure 6·17: Average and composite images for 100ns dwell time at 500fA beam current, $\lambda = 5$

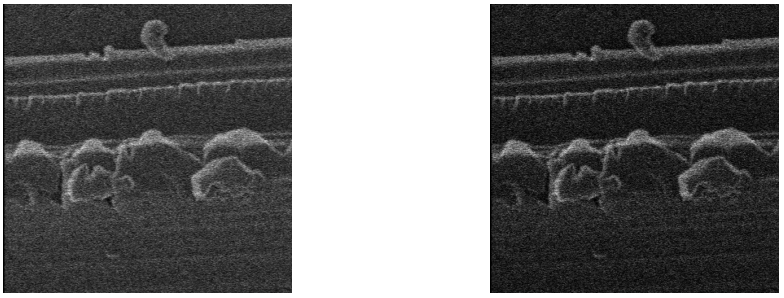

Figure 6·18: Average and composite images for 100ns dwell time at 500fA beam current, $\lambda = 7$

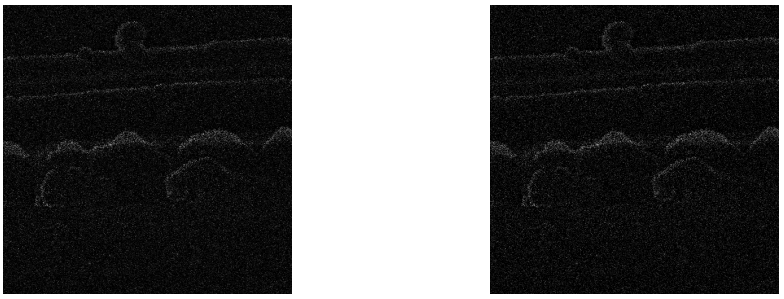

Figure 6·19: Average and composite images for 500ns dwell time at 105fA beam current, $\lambda = 1$

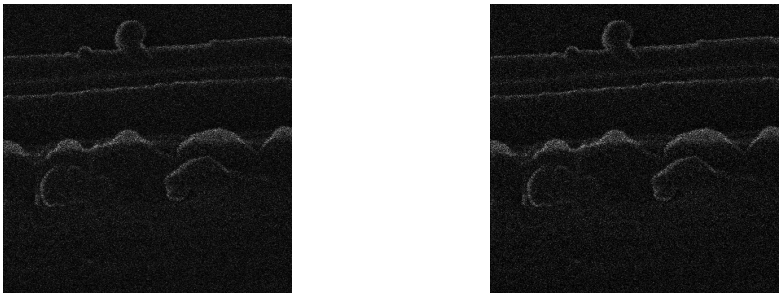

Figure 6·20: Average and composite images for 500ns dwell time at 105fA beam current, $\lambda = 3$

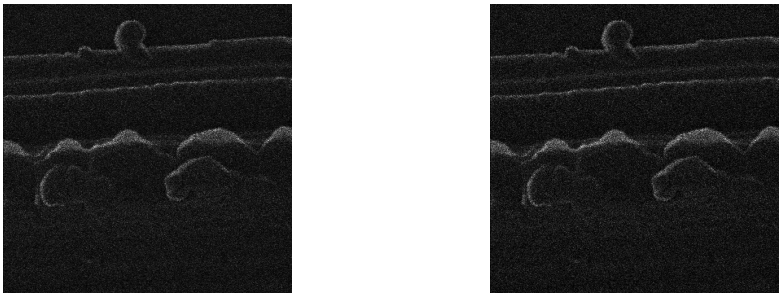

Figure 6·21: Average and composite images for 500ns dwell time at 105fA beam current, $\lambda = 5$

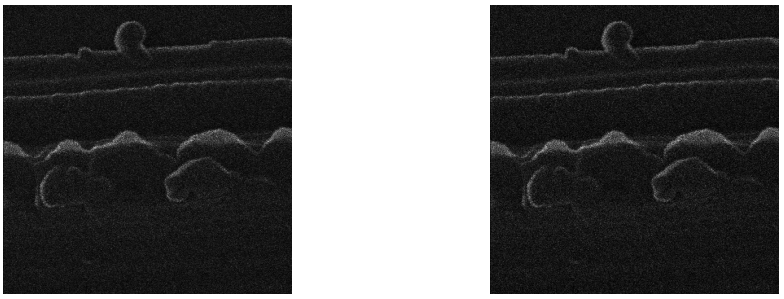

Figure 6·22: Average and composite images for 500ns dwell time at 105fA beam current, $\lambda = 7$

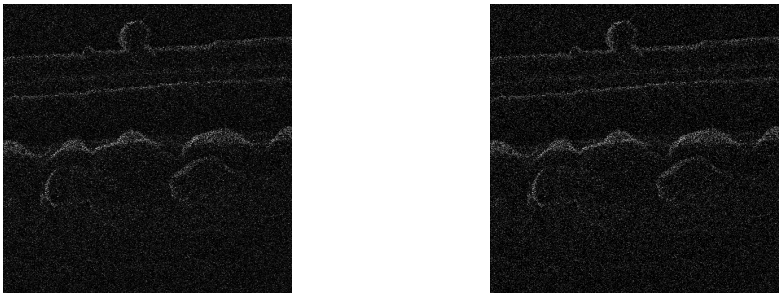

Figure 6·23: Average and composite images for 500ns dwell time at 182fA beam current, $\lambda = 1$

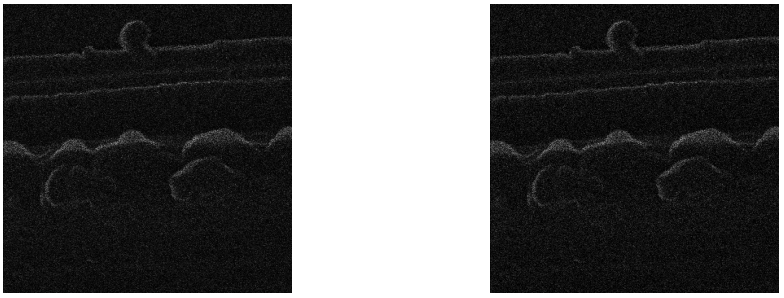

Figure 6·24: Average and composite images for 500ns dwell time at 182fA beam current, $\lambda = 3$

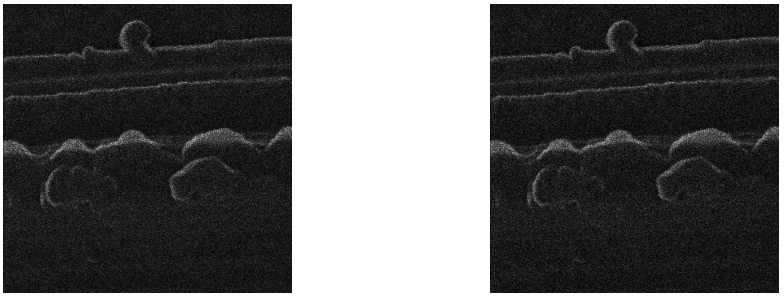

Figure 6·25: Average and composite images for 500ns dwell time at 182fA beam current, $\lambda = 5$

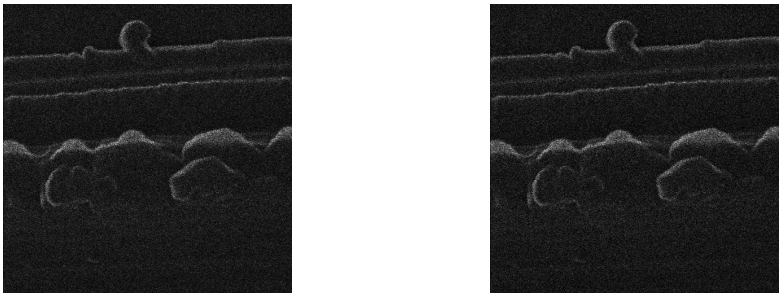

Figure 6·26: Average and composite images for 500ns dwell time at 182fA beam current, $\lambda = 7$

6.5 Estimation Results

Because ground truth iSE yield data from the given sample is unavailable, it is impossible to judge the performance of the estimation procedure in terms of MSE. Instead, the images can only be compared in terms of visual quality. For each beam current

and dwell time setting used in the experiment, an image representing "best" visual performance was selected for each regularization. The image selected corresponded to a value of τ which produced an image with good contrast, brightness, and a minimum of oversmoothing while demonstrating the effect of the regularization process.

For dwell times of 100ns, the performance between the Neyman Type A and Poisson estimations are visually very similar. In general, the smoothing process has the strongest effect in bright regions. In these regions impulsive noise caused by pixels with zero impinging ions is suppressed. As the τ parameter is increased, similar smoothing can occur in the flat, less bright regions as well. However, when τ is increased to these levels image features become degraded by oversmoothing. The brightness of the regularized images is generally better than the original composite images. This is the result of a decrease in the maximum value of $\hat{\eta}$ by the regularization process, improving brightness and contrast.

For dwell times of 500ns, the performance of the Neyman Type A estimation is slightly better than the Poisson performance. The same general increase in brightness occurs, as the maximum value of $\hat{\eta}$ is decreased by the estimation process. For values of $\lambda = 1$ and $\lambda = 3$ the Neyman estimation procedure is able to create images that are in general brighter with generally better visual qualities than the original composite images. However as λ is increased to 5 and 7, the composite images are interpretable in their original state and estimation only results in an increase in brightness.

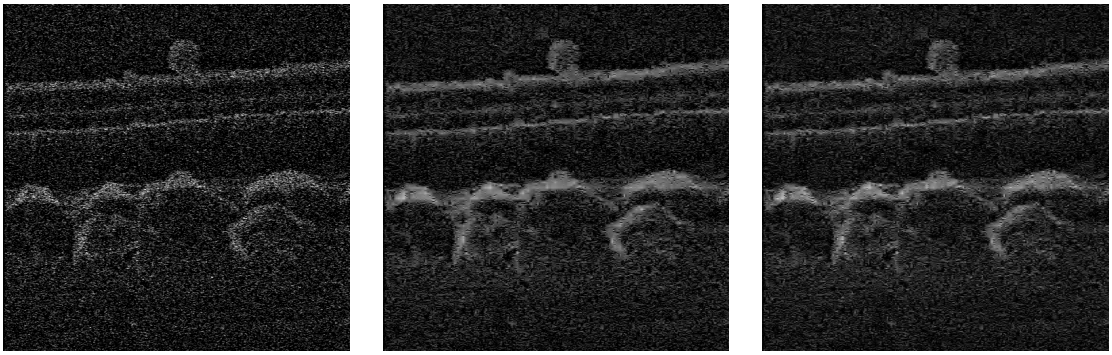

Figure 6·27: Composite (left) and Neyman regularized (middle) and Poisson regularized (right) images for 100ns 400fA, $\lambda = 1$

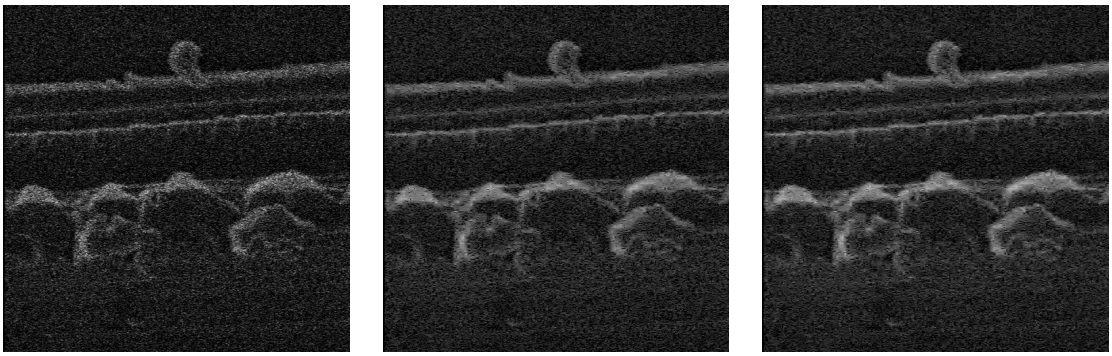

Figure 6·28: Composite (left) and Neyman regularized (middle) and Poisson regularized (right) images for 100ns 400fA, $\lambda = 3$

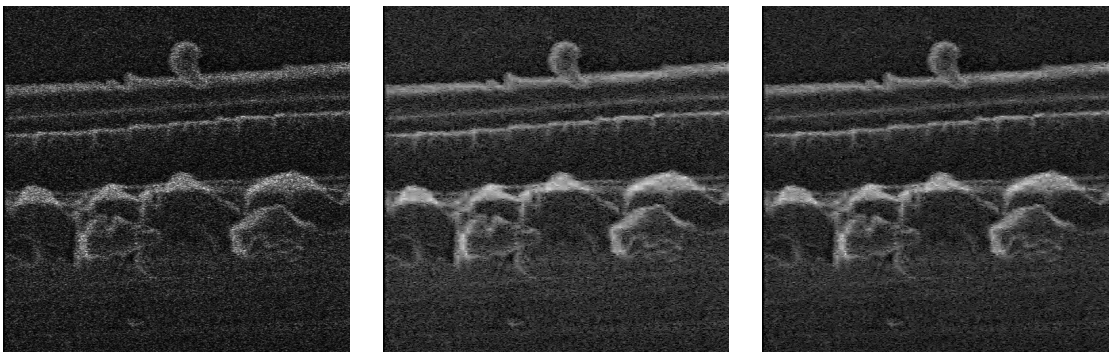

Figure 6·29: Composite (left) and Neyman regularized (middle) and Poisson regularized (right) images for 100ns 400fA, $\lambda = 5$

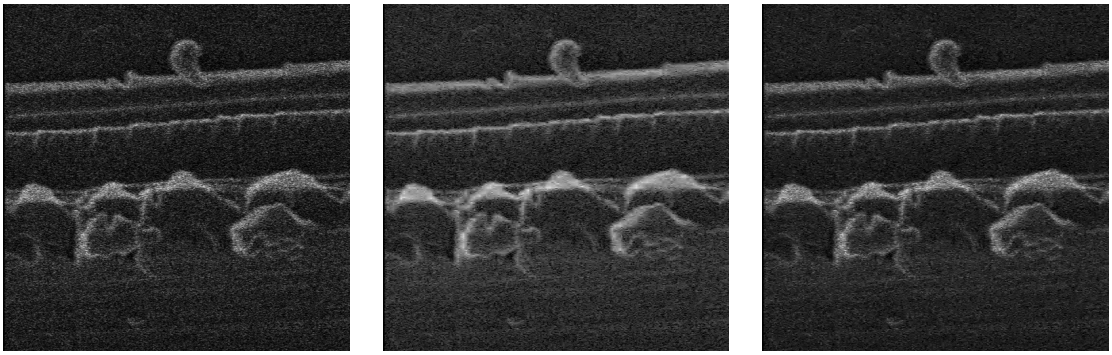

Figure 6·30: Composite (left) and Neyman regularized (middle) and Poisson regularized (right) images for 100ns 400fA, $\lambda = 7$

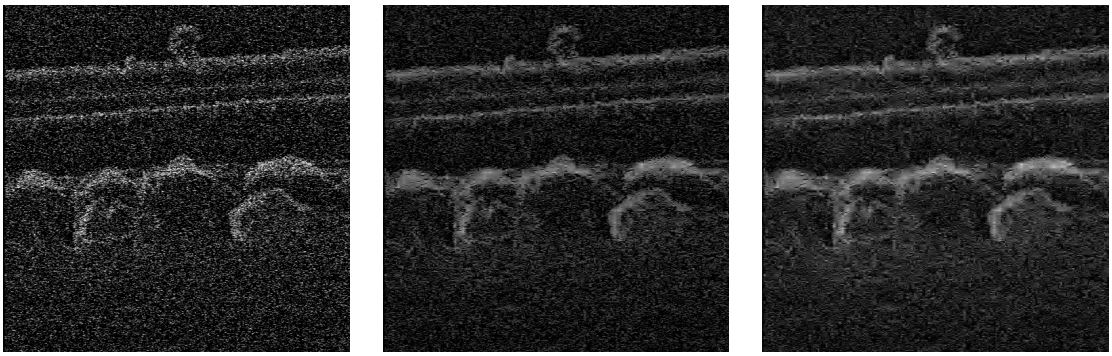

Figure 6·31: Composite (left) and Neyman regularized (middle) and Poisson regularized (right) images for 100ns 500fA, $\lambda = 1$

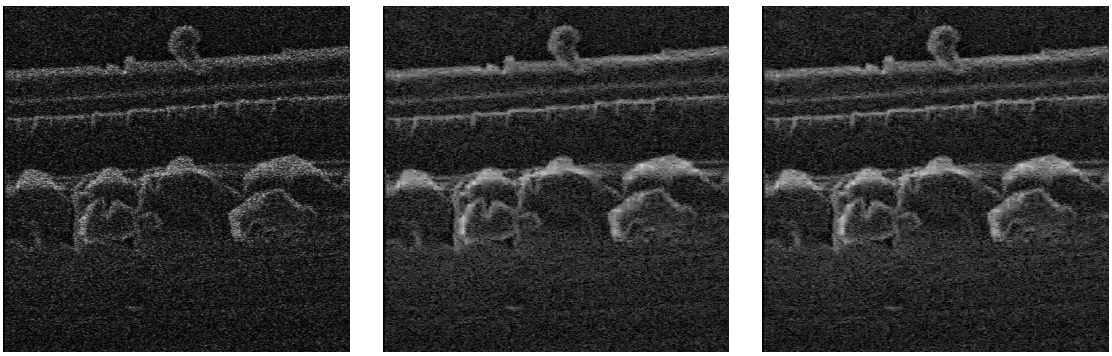

Figure 6·32: Composite (left) and Neyman regularized (middle) and Poisson regularized (right) images for 100ns 500fA, $\lambda = 3$

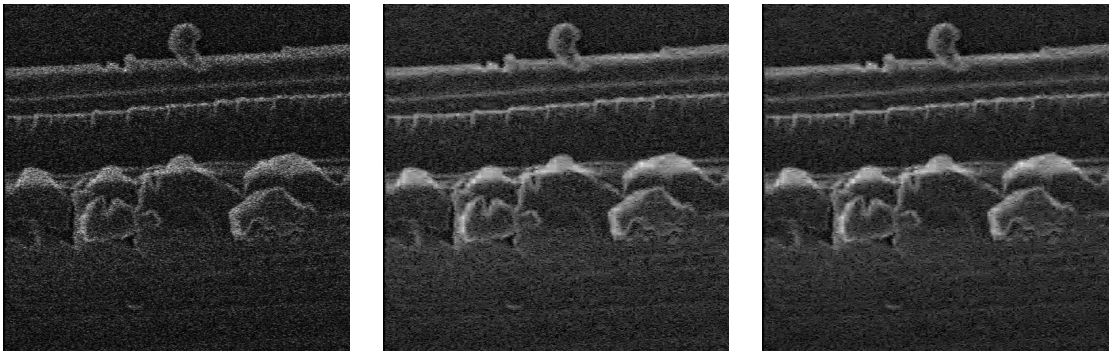

Figure 6·33: Composite (left) and Neyman regularized (middle) and Poisson regularized (right) images for 100ns 500fA, $\lambda = 5$

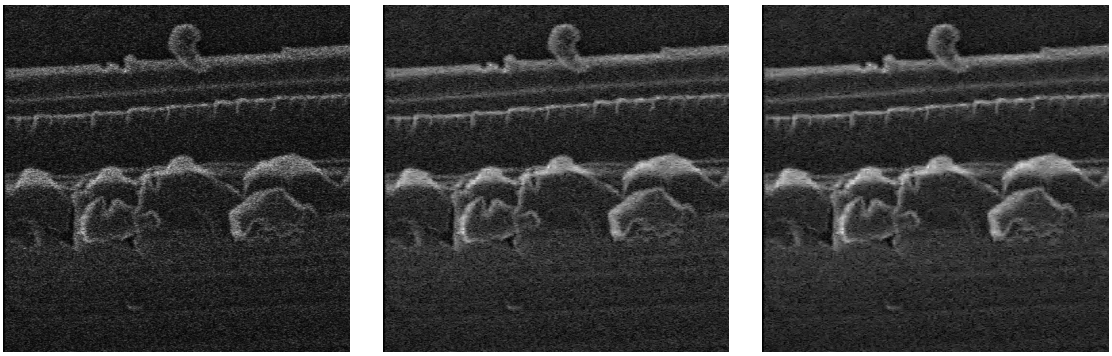

Figure 6·34: Composite (left) and Neyman regularized (middle) and Poisson regularized (right) images for 100ns 500fA, $\lambda = 7$

Figure 6·35: Composite (left) and Neyman regularized (middle) and Poisson regularized (right) images for 500ns 105fA, $\lambda = 1$

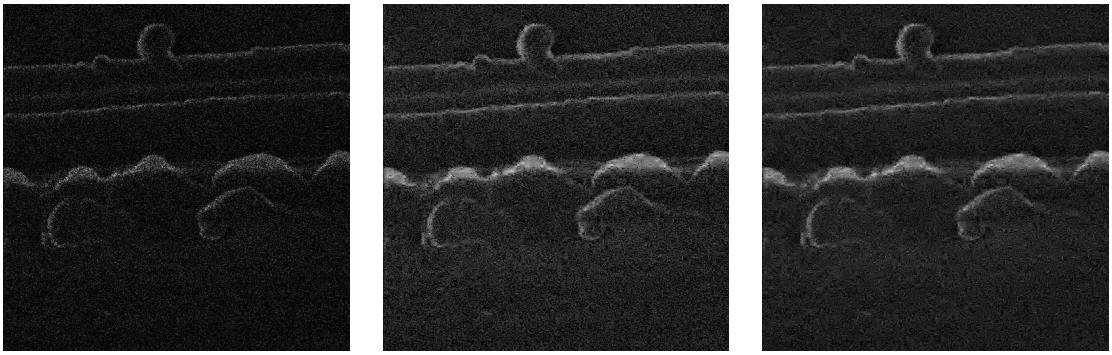

Figure 6·36: Composite (left) and Neyman regularized (middle) and Poisson regularized (right) images for 500ns 105fA, $\lambda = 3$

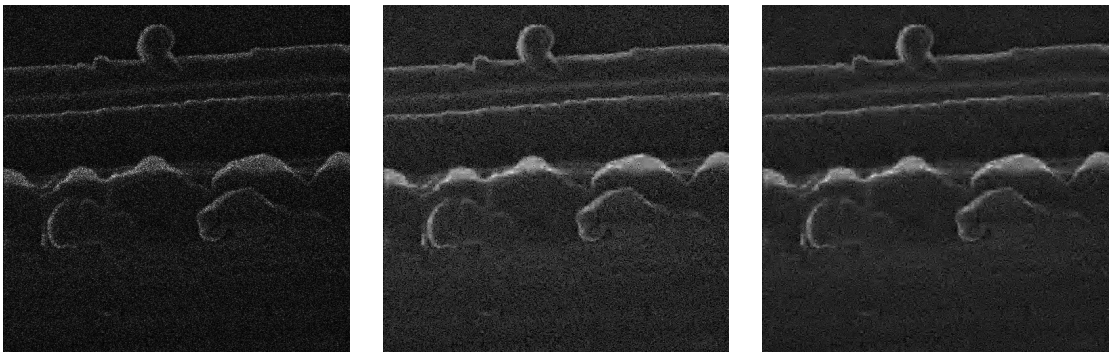

Figure 6·37: Composite (left) and Neyman regularized (middle) and Poisson regularized (right) images for 500ns 105fA, $\lambda = 5$

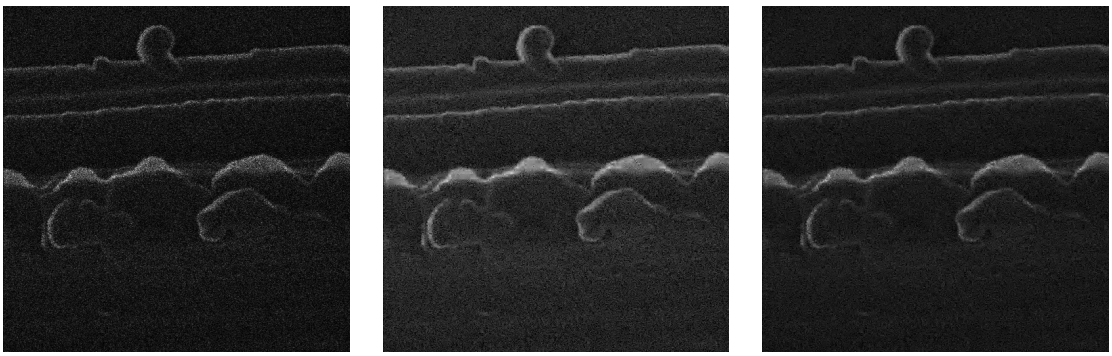

Figure 6·38: Composite (left) and Neyman regularized (middle) and Poisson regularized (right) images for 500ns 105fA, $\lambda = 7$

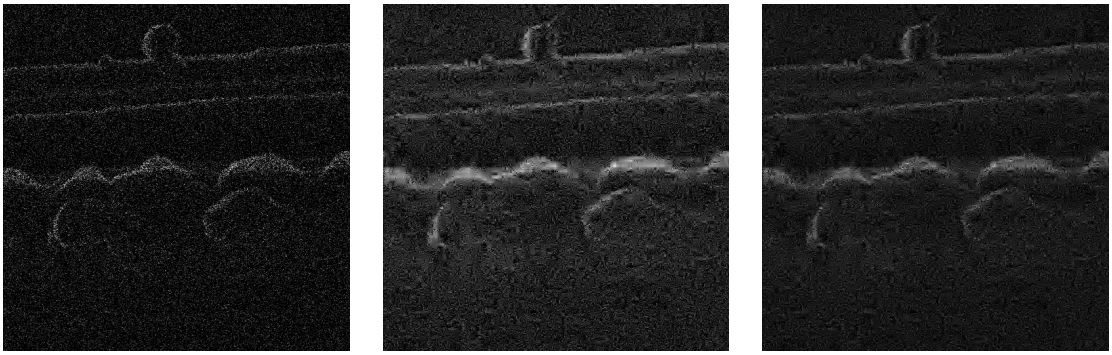

Figure 6·39: Composite (left) and Neyman regularized (middle) and Poisson regularized (right) images for 500ns 182fA, $\lambda = 1$

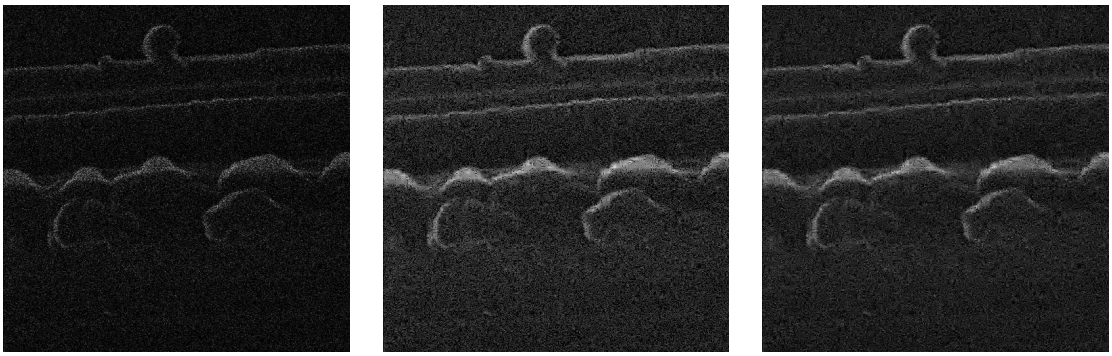

Figure 6·40: Composite (left) and Neyman regularized (middle) and Poisson regularized (right) images for 500ns 182fA, $\lambda = 3$

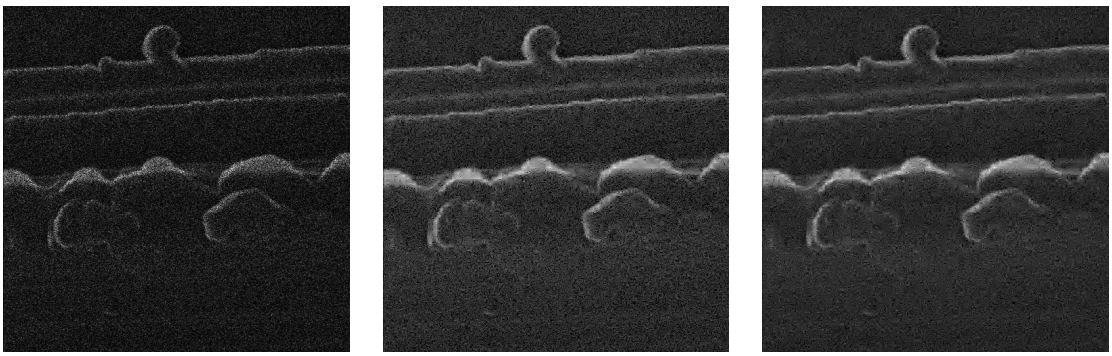

Figure 6·41: Composite (left) and Neyman regularized (middle) and Poisson regularized (right) images for 500ns 182fA, $\lambda = 5$

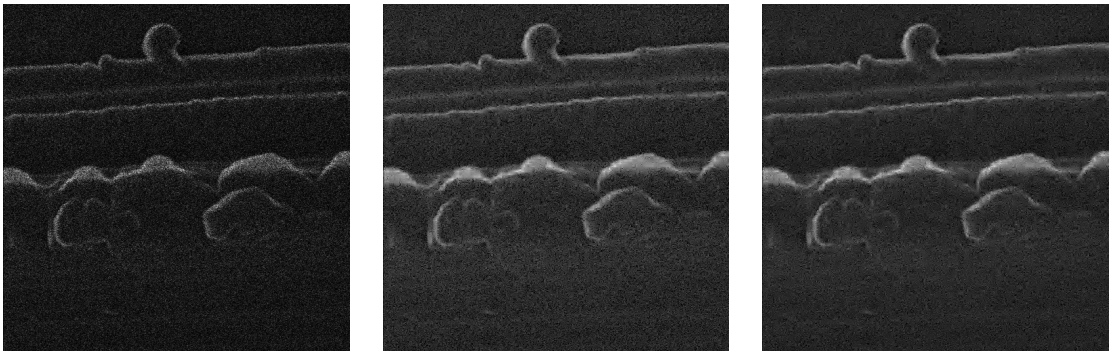

Figure 6·42: Composite (left) and Neyman regularized (middle) and Poisson regularized (right) images for 500ns 182fA, $\lambda = 7$

6.6 Summary

This chapter has detailed the application of the Neyman Type A estimation procedure to real world HIM data. This process is complicated by the unavailability of iSE count measurements and a ground truth image of iSE yields. In order to develop iSE count images, quantization was performed based on a Neyman Type A fit of the image intensities. The linearity as well as the clustering of quantization boundaries observed reinforced the choice of the Neyman Type A distribution as an accurate model of image formation. Furthermore, the improvements in imaging through quantization also reinforce the choice of model. Thus by incorporating the Neyman Type A model of iSE statistics, the image formation process can be improved even without the use of the regularized ML estimation procedure.

The use of the estimation algorithm provides some improvement of visual quality. However, the improvement is minor when compared with the original images. The increase in performance is about the same for both the Neyman Type A estimation and the Poisson estimation. Without having a ground truth of iSE yields or an actual measurement of iSE counts, quantitatively determining the performance is not possible.

Chapter 7

Conclusion

7.1 Conclusion

The major development in this thesis is a method of estimating iSE yields in HIM samples given an image of iSE counts. This approach to HIM imaging differs from traditional HIM imaging in two main respects. First, the randomness in the beam is directly accounted for in the estimation process. In traditional HIM imaging, the effect of this randomness is suppressed by simply increasing helium ion dose. The estimation method presented here includes beam randomness and is thus able to perform imaging at lower ion doses.

Secondly, in traditional HIM image formation, the image is a representation of the iSE signal at a given pixel. Changes in this signal are due to changes in $\eta_{i,j}$ over the sample. However, the iSE signal includes randomness due to the emission of helium ions and the emission of iSE. The estimation procedure developed here directly estimates $\eta_{i,j}$. Thus the image formation procedure attempts to form an image of the physical parameter responsible for changes in contrast. By comparison traditional image formation accepts the iSE signal as a sort of proxy for the change in iSE yield over the sample.

The simulated imaging experiments show that improvement in low ion dose imaging is possible using Neyman Type A regularized ML estimation. When compared with naive estimations assuming Poisson statistics, the Neyman Type A method performs better in terms of contrast and brightness. The ability of the Neyman Type

A estimation algorithm to take into account both very high and very low iSE counts allows it to preserve contrast while estimating physically realistic iSE yield values.

When evaluating real data from the microscope, the choice of the Neyman Type A distribution was reinforced. The decision boundaries created by the quantization process were linear and clustered at similar points. This indicates that the Neyman Type A model is an accurate representation of the imaging process. Furthermore, the process of quantizing the image data into iSE yields from image intensities improves the quality of combinations of images. This further reinforces the choice of image formation model. Thus there is an apparent gain in imaging just by using the Neyman Type A model in the imaging process without even applying an estimator.

Real data was used in the estimation procedure with limited increase in performance. Without an available ground truth or an actual measure of iSE counts, the results were impossible to evaluate in terms of MSE. Instead only visual quality was considered when evaluating the performance of the estimation. While some gain in brightness and contrast was noted as an effect of the estimation procedure, the visual performance was about the same as the original images.

7.2 Future Work

This project is open for further development in many directions. Many of these directions stem from the difficulty in detecting explicit numbers of iSE counts in the HIM. One possible experiment could verify the current method of mapping image intensities to specific numbers of iSE counts. Such an experiment might use a sample containing two flat regions of constant iSE yield. Using each flat region to create a mapping of image intensity to iSE counts, it would be possible to validate our method of mapping intensities to counts.

However, the ability to directly count iSEs would improve the ability to work with

real world data. Due to opacity in the imaging process, this data is not currently available. Direct access to the ET detector may allow for measurement of iSE counts at a given pixel. However it is difficult to determine what may be possible without more direct access to the instrument.

A method of estimation that uses the number of helium ion emissions could also improve the performance. In the case of real world data, a series of images were quantized to iSE counts. Considering a single pixel across the series of images, if the pixel had a non-zero count of iSE in an image then it can be inferred that at least one helium ion struck the sample at that pixel. If there were no iSE counts, than a strike could not be directly inferred. Over the series of images, some minimum number of ion strikes could be determined for every pixel. Using this information could improve the performance of the estimation procedure.

Another potential avenue of investigation is in the application of our method, or a similar type of estimation to a different type of microscopy. Direct electron counting is becoming available in transmission electron microscopy and has been investigated in (Moldovan et al., 2009) and (McMullan et al., 2014). Use of direct electron counting detectors would make regularization using electron statistics less difficult. While these detectors are not available for HIM, similar regularization as proposed here may be possible in transmission electron microscopy.

Appendix A

Convergence

This appendix outlines the general convergence results for the minimization algorithm. This discussion relies heavily on the development of convergence for the SPIRAL algorithm. For a more detailed overview of the convergence please see (Harmany et al., 2012). Only the basic details of convergence will be presented here as the main results follow directly the proof in (Harmany et al., 2012).

The original minimization problem is reformulated in order to place the non-negativity constraint in the objective function.

$$\begin{aligned} \underset{\eta\in\mathbb{R}^n}{\text{minimize}} \quad & \Phi \triangleq F(\eta) + \tau\text{pen}(\eta) \\ \text{subject to} \quad & \eta \geq 0 \end{aligned} \tag{A.1}$$

Defining $\rho : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty, \infty\}$ to be

$$\rho = \tau\text{pen}(\eta) + \delta_+(\eta) \tag{A.2}$$

where

$$\delta_+(\eta) = \begin{cases} 0 & \text{if } \eta \geq 0 \\ \infty & \text{otherwise} \end{cases} \tag{A.3}$$

the original problem can be considered as

$$\underset{\eta\in\mathbb{R}^n}{\text{minimize}}\ F(\eta) + \rho(\eta) \tag{A.4}$$

Some assumptions must be made to for the convergence of the algorithm

- (A1) The negative log-likelihood function is pseudoconvex in the region the algorithm is being applied to
- (A2) ρ is proper convex and continuous on $\mathbb{R}_+^n$
- Φ is coercive

The assumption (A1) has been argued in Section 3.3.

Next, it is shown that for any $\bar{\eta} \in \mathbb{R}_+^n$ that is not critical and $\bar{\alpha} \geq \alpha_{min}$, there exists some $\epsilon \geq 0$ such that for all j sufficiently large

$$\|\delta^{k_j+1}\|_2 = \|\eta^{k_j+1} - \eta^{k_j}\|_2 \geq \epsilon$$

where $\eta_{j\in\mathbb{Z}_+}^{k_j}$ is a subsequence with $\lim_{j\to\infty} \eta^{k_j+1} = \bar{\eta}$ and $\alpha_{min} \leq \alpha_{k_j} \leq \bar{\alpha}$. This condition shows that at a non-critical point, the subproblem will take some step.

It is also necessary to show that the sequence $\{\eta^k\}_{k\in\mathbb{Z}_+}$ generated by the algorithm is such that $\lim_{k\to\infty} \delta^{k+1} = 0$. As this occurs, there is a value of the objective function $\bar{\Phi} \in \mathbb{R}$ such that $\lim_{k\to\infty} = \bar{\Phi}$.

Taken together, these conditions ensure that the must converge to some critical point. The algorithm will continue to iterate until a critical point has been reached. By the pseudoconvexity of the objective function, this point must be a minimum.

References

Boyadzhiev, K. N. (2009). Exponential polynomials, Stirling numbers, and evaluation of some gamma integrals. *Abstract and Applied Analysis*, 2009.

Castaldo, V., Hagen, C., and Kruit, P. (2011). Simulation of ion imaging: Sputtering, contrast, noise. *Ultramicroscopy*, 111(8):982–994.

Eder, H., Vana, M., Aumayr, F., and Winter, H. P. (1997). Precise total electron yield measurements for impact of singly or multiply charged ions on clean solid surfaces. *Review of Scientific Instruments*, 68.

Harmany, Z. T., Marcia, R. F., and Willett, R. M. (2012). This is SPIRAL-TAP: Sparse Poisson Intensity Reconstruction ALgorithms–theory and practice. *IEEE Transactions on Image Processing*, 21(3):1084–1096.

Johnson, N. L., Kotz, S., and Kemp, A. W. (1992). *Univariate Discrete Distributions*. John Wiley and Sons, New York, 2nd edition.

Joy, D. C. (2013). *Helium Ion Microscopy: Principles and Applications*. Springer.

Kirmani, A., Venkatraman, D., Shin, D., Colao, A., Wong, F. N. C., Shapiro, J. H., and Goyal, V. K. (2014). First-photon imaging. *Science*, 343(6166):58–61.

McMullan, G., Faruqi, A., Clare, D., and Henderson, R. (2014). Comparison of optimal performance at 300 keV of three direct electron detectors for use in low dose electron microscopy. *Ultramicroscopy*, 147:156–163.

Moldovan, G., Jeffery, B., Nomerotski, A., and Kirkland, A. (2009). Imaging modes for direct electron detection in TEM with column parallel CCD. *Nuclear Inst and Methods in Physics Research A*, 607(1):13–16.

Orloff, J., Swanson, L., and Utlaut, M. (2011). Fundamental limits to imaging resolution for focused ion beams. *Journal of Vacuum Science and Technology. B, Microelectronics Processing and Phenomena*, 14(6):3759–3763.

Ramachandra, R., Griffin, B., and Joy, D. (2009). A model of secondary electron imaging in the helium ion scanning microscope. *Ultramicroscopy*, 109(6):748–757.

Roman, S. (1984). *The Umbral Calculus*. Academic Press, New York.

Shin, D., Kirmani, A., Goyal, V. K., and Shapiro, J. H. (2015). Photon-efficient computational 3d and reflectivity imaging with single-photon detectors. *IEEE Transactions on Computational Imaging*, 1(2).

Teich, M. C. (1981). Role of the doubly stochastic Neyman type-A and Thomas counting distributions in photon detection. *Applied Optics*, 20(14):2457–2467.

Vana, M., Aumayr, F., Lemell, C., and Winter, H. (1995). Ion-induced electron emission from solid surfaces: information content of the electron number statistics. *International Journal of Mass Spectrometry and Ion Processes*, 149:45–57.

Ward, B., Notte, J. A., and Economou, N. (2006). Helium ion microscope: A new tool for nanoscale microscopy and metrology. *Journal of Vacuum Science & Technology*, 24(6):2871–2874.

Wright, S. J., Nowak, R. D., and Figueiredo, M. A. T. (2009). Sparse reconstruction by separable approximation. *IEEE Transactions on Signal Processing*, 57(7):2479–2493.

Zeiss (2015). Orion nanofab three ion beams for total flexibility in sub-10 nm fabrication. Technical report, Carl Zeiss Microscopy GmbH.

CURRICULUM VITAE

Jeff Craley

Education

2014-2015 MS Electrical Engineering - Boston University

2005-2010 BS Aerospace Engineering, BA English - Virginia Tech

Work

2014-2015 Research Assistant - Boston Univeristy

2010-2014 Math and Physics Tutor - Private Tutoring

2006-2007 Civil Engineering Intern - GKY & Assoc.