A High Average-Current Electron Source for the Jefferson Laboratory Free Electron Laser

Fay E. Hannon M.Eng.

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

to the
Department of Engineering
University of Lancaster

October 29, 2008

Declaration

I hereby declare that this thesis is of my own composition, and that it contains no material previously submitted for the award of any other degree. The work reported in this thesis has been executed by myself, except where due acknowledgement is made in the text.

Fay E. Hannon M.Eng.

Abstract

A High Average-Current Electron Source for the Jefferson Laboratory Free Electron Laser.

The spectral output power from the Jefferson Laboratory infra-red free electron laser is primarily limited by the performance of the electron injector. Free electron laser power is directly proportional to the electron beam current and at present the electron injector is limited to 10mA average current. To date the highest laser power achieved has been 14.2kW and the next goal is to reach 100kW. For this to occur a new electron injector has been designed that is capable of producing over 100mA average current. This thesis describes an investigation into the behaviour of this injector through simulation.

Given that the layout of the injector is fixed, this thesis aims to find suitable operating regimes for various electron bunch charge scenarios. By determining the important features the electron beam must have at the exit of the injector, and the limitations of each component, this information was used to form an optimisation problem that could be solved to find the best operation point.

To improve the simulation of electron bunches being launched from a photocathode, measurements were performed on a similar injector to evaluate the thermal energy and response time of the cathode. These values are a function of the laser wavelength used with the photocathode and so were repeated over a range of wavelengths from infra-red to green. The injector at Cornell University was used to take measurements of the electron beam that could then be compared against simulation to benchmark the code.

The brightness and quality of electron beams in linac-based light sources, such as at Jefferson Laboratory, are limited by the properties of the beam in the injector. It is therefore important to have knowledge of the phase space distribution of the electron beam in addition to the rms emittance, to provide an insight into high brightness formation mechanisms. A tomography technique has been successfully used to reconstruct the transverse phase space of the electron beam delivered from the Cornell University ERL DC gun. The gun is similar to that in the 100mA JLab injector, therefore a tomography diagnostic could in future be applied to that case.

Acknowledgements

I would like to thank my advisor, Richard Carter, for his support and guidance in completing this thesis. It has been a challenging but rewarding undertaking, and Richard has steered me adeptly through to its completion.

I embarked upon this post graduate study part-time whilst working as a graduate at Daresbury Laboratory. Without the encouragement and assistance from Mike Dykes, Jim Clarke and Susan Smith, and their commitment to further education, this thesis would not have been feasible. I express my gratitude to all members of the Accelerator Science and Technology Centre for their help along my journey. During my time at Daresbury I made a lasting friendship with Bruno Muratori, who merits special thanks. He has contributed immeasurably to the completion of this project. The giving of his time, advice, encouragement and unfailing interest, warrant my respect and deepest gratitude.

In 2005 I relocated to Jefferson Laboratory to continue my studies full time under the guidance of Bob Rimmer. Bob has been an excellent mentor to me throughout my time at JLab. I have benefitted greatly from our enlightening discussions, his sharp mind and unwavering moral support. There are other colleagues from Jefferson Lab that I would like to thank for their expertise and technical advice. The FEL group, in particular George Neil, Gwyn Williams, Carlos Hernandez, Steve Benson, Dave Douglas, Pavel Evtushenko and Michelle Shinn, for their patience in answering my many questions. I am indebted to Alexander Stolin from the Radiation Detector and Medical Imaging Group for furthering my understanding of reconstruction algorithms. Finally I wish to extend my appreciation to Geoffrey Krafft for helping me through some mathematically challenging times.

I owe a debt of gratitude to Maury Tigner for agreeing to my secondment to Cornell University to take measurements on the ERL prototype. I enjoyed working with the ERL group immensely, from whose extensive collective knowledge I benefitted greatly. I would specifically like to thank Bruce Dunham and Ivan Bazarov for their guidance and confidence in me. I am particularly indebted to Ivan for his inspiration and the extension of my knowledge and horizons.

It is unavoidable that some omissions have been made in these acknowledgements. I would therefore like to thank all who have contributed to this project but are not specifically mentioned here.

On a personal note, I wish to convey my appreciation to Andrew Kimber for his forbearance and encouragement throughout this project. Finally, I want to express an especial debt of gratitude to my parents, to whom this thesis is dedicated. Without the love and support they have given me throughout my academic career, this project would not have been possible.

Contents

List of Figures

CHAPTER 1

Introduction

This thesis will describe the progression of design and development of a high brightness, high average-current electron injector. The electron source is designed for operation with a next generation light source, that produces tuneable, short pulse length, synchrotron radiation from Free Electron Lasers (FELs).

1.1 Energy Recovery Linac Light Sources

The exploitation of synchrotron radiation has evolved over several decades since the first observation in 1947 from the General Electric 70MeV synchrotron [1]. Synchrotron radiation is generated spontaneously when high energy electrons are deflected by magnetic fields, and in early machines the radiation was extracted from ports on the bending magnets. State-of-the-art 3rd generation light sources now emit radiation from electrons circulating in specifically designed undulators and multipole wigglers which increase the spectral flux by many orders of magnitude.

Synchrotron radiation is used to investigate the structural properties of matter [2]. Irradiating materials can provide information on the arrangement of atoms,

how the structure changes during chemical or biological processes, and electronic properties. Other uses include X-ray imaging, tomography and lithography.

The total synchrotron radiation power from storage ring machines is proportional to the beam current, the beam energy to the fourth power, and inversely proportional to the bending radius of the magnets. The spectral brightness (number of photons per unit volume) of the photon beam defines the optical quality, and is inversely proportional to the transverse electron beam emittance. This describes the dimensions and divergence of the beam and will be discussed further in chapter 2. Electron beams with smaller emittances therefore result in higher brightness photon beams and this is a primary goal of modern light sources [3].

Storage rings are by nature equilibrium devices and the emittance is proportional to the beam energy squared and the bend angle cubed. The beam emittance grows with multiple passes around the synchrotron which imposes practical limits on the minimum emittance, and trying to further reduce it becomes complex. Energy Recovery Linacs (ERL) were considered a method of meeting the demand for increased photon output power and spectral brightness. In contrast to storage rings, the electrons in an ERL make only a few circuits of the machine before being discarded and therefore the emittance does not reach the equilibrium state for a given beam energy.

The beam properties in an ERL can be similar to those achieved in a Linear Accelerator (LINAC), which are single pass machines, where the electrons are disposed of at their full energy in either an experimental target or some dedicated beam dump. The brightness of the electron beams (number of electrons per unit volume) in these machines is $\sim$2 orders of magnitude higher than that in storage rings and is primarily determined by the performance of the electron injector.

The concept of energy recovery is not a new one, it was suggested in 1965 by M Tigner [4]. The principle of an ERL is that the electron bunch is decelerated before it is dumped and the energy recovered from the electron beam. In this way the wall-plug power of the machine is much reduced by comparison. An additional benefit is that the design of the beam dump becomes easier with reduced beam energy. Twenty two years after conception the first demonstration employing a superconducting (SC) linac was performed at Stanford using the HEPL facility in 1987 [5]. The electron beam in this experiment was accelerated and decelerated in the same SC cavity. Following this, energy recovery was achieved in a second LINAC, separate from the accelerating cavity. In addition, the electron beam had been disrupted by a FEL (discussed in section 1.2) [6]. It wasn't until 2000 that same-cell energy recovery was successfully demonstrated with a free electron laser at Thomas Jefferson National Accelerator Facility (JLab) on the IR-DEMO [7]. Following the success of these ERLs, the popularity of such apparatus has considerably increased.

1.2 Free Electron Lasers

In parallel with advancements to synchrotron machines there has been a programme to develop the potential of free electron laser sources, for example [8]. A free electron laser is created using a specifically designed insertion device in which the production of synchrotron radiation is increased by multi-particle coherence.

Insertion devices have spatially periodic magnetic fields acting perpendicular to the electron axis of motion. This causes the electrons to follow a transverse sinusoidal path, thus producing spontaneous radiation. There are two main types of insertion device: undulators and wigglers. The deflection parameter, which calculates the amplitude of the electron oscillation, determines which

type of insertion device it is:

$$K = 0.0934 B_0 \lambda \tag{1.1}$$

where B_0 is the peak magnetic field in Tesla and λ is the period in mm. A wiggler will have a deflection parameter of above 1 and an undulator below 1. However, there is no strict boundary between the two. Wigglers tend to have higher peak magnetic fields and longer periods than undulators. The radiation from wigglers is very similar to that from bending magnets only with N times more flux, where N is the number of periods.

Traditionally insertion devices were placed in storage rings to produce tunable radiation many times brighter than that from the bending magnets and have the additional benefit of reducing the transverse emittance. A FEL source is an insertion device operated in a specific regime, whereby the production of synchrotron radiation is extended further, into a region where output power increases by many orders of magnitude from that obtained with conventional undulators and wigglers through coherent emission. The coherent synchrotron radiation from FEL sources has been used extensively since the first device was demonstrated in 1976 at Stanford University [9]. Since then, FELs have been used to create radiation spanning the spectrum from radio waves to the ultra-violet (UV), both in storage rings and LINACs [10].

In the last 10 years the combination of FEL and ERL has evolved, and three facilities are currently operating. These are at JLab, the Japan Atomic Energy Research Institute (JAERI) [11] and the Budker Institute for Nuclear Physics (BINP) in Russia [12]. Presently, all these machines have an IR FEL installed, but this is likely to change in the next few years as multi-wavelength ERL/FEL proposals start to be commissioned.

1.3 The FEL Project at Jefferson Laboratory

From 1995 to 2001 Thomas Jefferson National Accelerator Facility designed and built an ERL, the IR-DEMO, designed specifically to produce high average-power coherent infra-red (IR) radiation from a FEL. The design of the facility is shown schematically in figure 1.1.

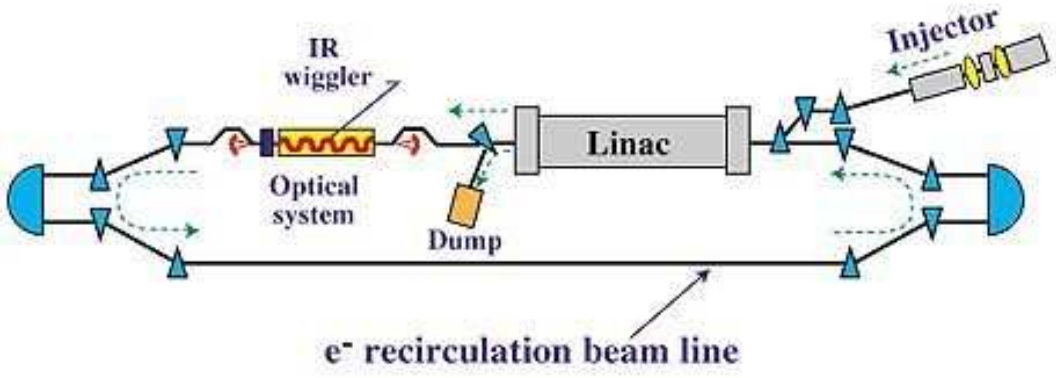

Figure 1.1: Schematic Layout of the IR-DEMO

Initially the FEL itself was a short (∼1m) permanent magnet wiggler that was designed to produce 1kW of infra-red (3μm) power continuously from electron bunches with a repetition rate of 75MHz (the fundamental frequency of the LINAC is 1.497GHz). Within a year of operation the IR-DEMO broke the world record by demonstrating average lasing powers of 1.72kW, and eventually produced over 2kW maximum. Between 2001 and 2004 the machine was redesigned with a new wiggler. A longer optical klystron electromagnetic wiggler was installed before the machine was launched as a user facility. This wiggler produced slightly longer wavelength IR radiation (5μm), but at a much higher power. Routinely, 8.4kW of power could be produced. The 10kW milestone was achieved with that configuration, but only at 25% duty factor [13]. For a short period of time in 2004 a second electromagnetic wiggler was installed,

producing 4kW power at 75MHz at shorter infra-red wavelengths. More recently, from 2005 the wiggler has been replaced with a permanent magnet wiggler which reaches sub micron wavelengths and has delivered 14kW power and programmes are in place to further this still.

The JLab machine layout can be broken down into four sections: The electron injector, the main accelerating LINAC, the beam transport system and the FEL. Electrons are emitted from a photo cathode placed inside a static electric field. The electric field, created by a negative potential applied to the cathode structure, then accelerates the electrons away from the cathode surface, see figure 1.2.

Electron bunches are transported through the injector into the main LINAC for further acceleration to the machine energy of 160MeV. The beam transport system, consisting of a series of magnets, guides the electron beam through the FEL for lasing and back to the LINAC. On re-entering the LINAC the electrons are decelerated before being dumped at the injection energy of $\sim$10MeV. In this way approximately 99% of the energy introduced from the LINAC can be recovered.

The FEL operated as a user facility for two years producing continuous wave (at 75MHz) infra-red radiation, and also the world record for Terahertz radiation as a byproduct from the bending magnets in the beam transport system [14]. In 2005 a programme began to upgrade the machine with an ultraviolet FEL in a second branch, as shown in figure 1.3. The UV wiggler beamline is currently under construction and aims to produce 100W radiation at 0.35μm.

The limiting factor in achieving even higher power FEL output is the current that can be generated by the injector and supported by the machine. The present JLab power supply is limited to 10mA operation. The average FEL output power scales linearly with average current. However the saturation power (maximum from one electron bunch), is heavily dependent on the peak current

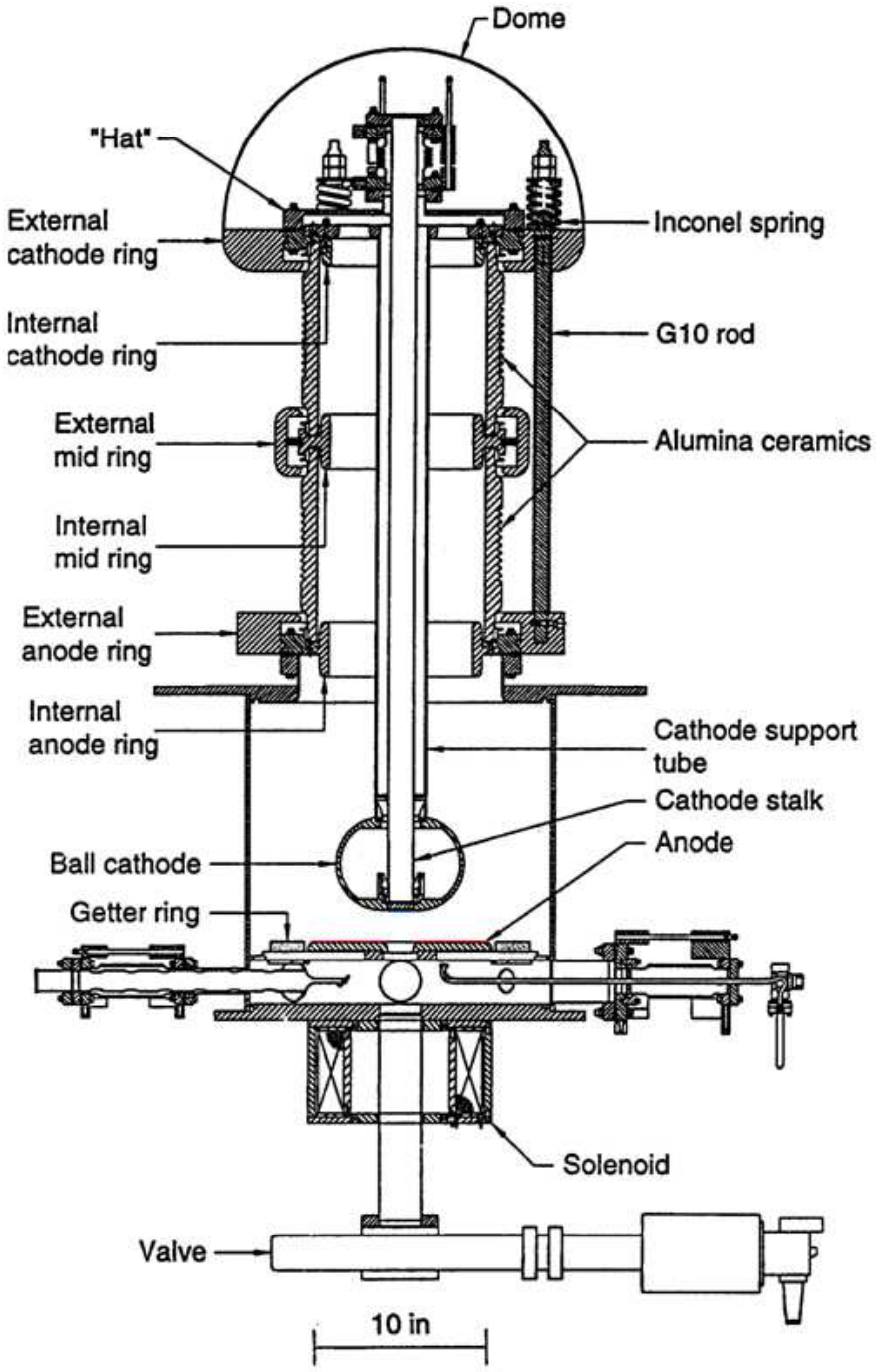

Figure 1.2: Schematic Top View of the IR-DEMO electron gun

and emittance of the electron bunch in the FEL. Both the average beam current and peak bunch current are largely defined by the injector performance, discussed in detail in chapter 2. In addition to this, the main LINAC must be capable of accelerating average currents over 10mA. The beam transport system should also be designed to minimise degrading effects on the bunch properties before the FEL. To date, the maximum average current that has been achieved in

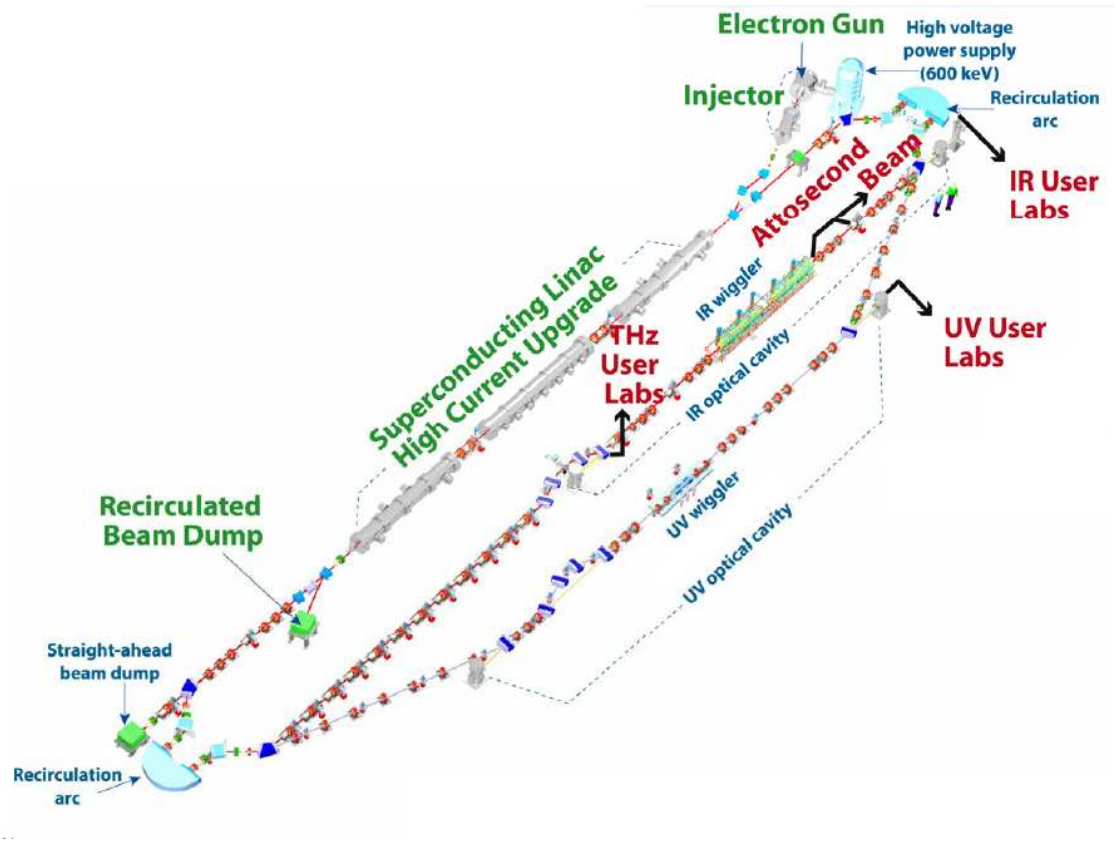

Figure 1.3: Schematic Layout of the FEL Upgrade

a continuous pulse mode is from the JLab FEL injector and is 9.1mA [15]. Each electron bunch had 120pC charge and came at a 75MHz repetition rate filling every 20th RF bucket in the LINAC.

1.4 Scope of this Thesis

It has been proposed to design a replacement injector capable of producing 100mA average current with improved beam quality to increase the FEL output power from 10kW to 100kW. The primary issue in achieving this goal is to build an electron injector that can not only provide the order of magnitude increase in current, but also maintain good electron beam quality and produce lasing at the required wavelength.

The mechanical design of a demonstration 100mA injector was created by Advanced Energy Systems (AES) in 2004. The basis of this design was to use a

DC photo-injector and place superconducting accelerating cavities as close as possible to the exit of the electron gun. The premise was to reduce the distance over which space charge forces in the electron bunch could degrade the quality. The initial design comprised of a DC gun, followed by a solenoid and seven accelerating cavities. This design subsequently changed due to funding issues and was merged with a second project to investigate the operation of a 3rd harmonic cavity [16]. How this new layout would operate in terms of electron beam properties was not investigated completely before manufacture began.

The aim of this thesis is to suggest the optimum set up of the injector through simulation, given that the mechanical design and layout was fixed. The parameters of the injector specification are explained and assessed with other injector projects. With the figures of merit defined, optimisation and particle simulation tools were utilised to determine the best component settings for differing operation scenarios.

To improve and validate the simulation tools used to model the JLab/AES 100mA injector, measurements were made at Cornell University with a similar injector operating in the same regime. The thesis reports the results from measurements of photocathode performance and electron beam dynamics. The results from the photocathode experiments were incorporated into the simulation of the JLab 100mA injector to produce a more realistic model.

Various methods have been employed to measure the phase space of the electron beam and compare it to that from simulation. Finally, a phase space tomography experiment was implemented and compared to simulation to support the findings. The general principle of tomography measurements previously performed was extended to include space charge dominated beams.

1.4.1 Organisation and Contribution

This thesis can roughly be divided into three sections. The first section contains information and definitions that are used throughout the thesis. The second section relates to results from measurements of electron beams that validate the simulation tools used. Simulation is then used to model the performance of the JLab FEL upgrade injector. The final section comments on the predicted performance of the injector, concluding with suggestions for future developments.

Background Information

Chapter 2 gives background to the notation used throughout the thesis to describe the properties of electron beams. The figures of merit that are used to compare the quality of electron bunches are examined in the context of what is required for the upgrade injector for the JLab FEL. Following this, **chapter 3** discusses methods of generating and accelerating electron beams, particularly those used at JLab and Cornell University. An overview and the general properties of photocathodes are described to give a framework to the measurements of photocathodes described in **chapter 5**.

Chapter 4 gives an overview of the simulation tools commonly used to model the behaviour of electron bunches, specifically in injectors. The simulation tool used throughout this thesis to model both the JLab upgrade FEL and the Cornell University injectors is described in detail, with particular attention paid as to how the figures of merit are calculated.

Simulation and Experiment

Chapter 5 describes the set up of the Cornell University ERL injector and diagnostic beamline. The diagnostic beamline provided means to measure the properties of the electron beam as it evolved. The properties of two photocathodes were investigated to derive the thermal energy and response time as a function

of illumination laser wavelength. The measurement programme was devised by the Cornell ERL group, and the author of this thesis contributed to the measurement taking and data analysis of the thermal energy experiment as part of the ”beam running” team. The result of this research was published in [17]. Although the author did not participate in the response time measurements, the results are briefly repeated in **chapter 5**, as they are later used in simulation.

The various methods used to measure thermal energy are given, as these are again required for the phase space tomography described in **chapter 6**. The general principles of tomography and reconstruction algorithms are discussed. These are then applied to tomography on electron bunches with no space charge forces in the Cornell injector, published in [18]. The author created the experimental schedule and designed the diagnostic beamline layout. The experiment was first performed virtually using simulation, and then compared to the results obtained from electron beam measurements. The tomography method was then extended to include electron bunches that had space charge forces. The validity and limitations of this were investigated. The results reported in **chapter 6** are entirely the work of the author.

Chapter 7 details a benchmarking experiment where the properties of electron bunches with dominant space charge forces are compared with the results of the modelling code used in this thesis. The measurements were again taken using the Cornell University injector. The results of a comparison with other simulation codes is reported in [19], to which the author contributed data.

Using the results of **chapters 5, 6** and **7** to improve the simulation of electron bunches, the JLab FEL upgrade injector (described in **chapter 8**) was modelled in **chapter 9**, solely by the author. A multivariate optimisation technique was used to predict the best performance achievable from the injector, given the constraints of the design. This was published by the author in [20]. The results of research are compared with the specification of the device.

To conclude, **chapter 10** discusses the results of the measurements taken using the Cornell injector, and the simulations performed of the JLab/AES upgrade injector. The performance of the latter, is finally evaluated against the predictions and requirements of other injector systems.

CHAPTER 2

Electron Beam Dynamics and FEL Requirements

This chapter introduces some of the figures of merit often used to characterise electron beams and how they relate to the performance of the machine. Defining the parameter space, or injector specification, is application dependent. In the context of the JLab FEL upgrade, achieving the wavelength, brightness and power of the FEL is the prime driver for the specification of the injector.

2.1 Electron Beam Description

The motion of charged particles in a beam can be described completely by six degrees of freedom in *phase space*, namely the position ($x,\ y,\ z$) and momentum ($p_x,\ p_y,\ p_z$) in Cartesian coordinates. Phase space is the space in which all possible states of a system are represented. The coordinate system used in defining electron beams is shown in figure 2.1, where the longitudinal, horizontal and vertical axes are defined throughout this thesis by z, x and y respectively. Liouville's theorem states that the six dimensional phase space density of non-interacting particles is conserved in time.

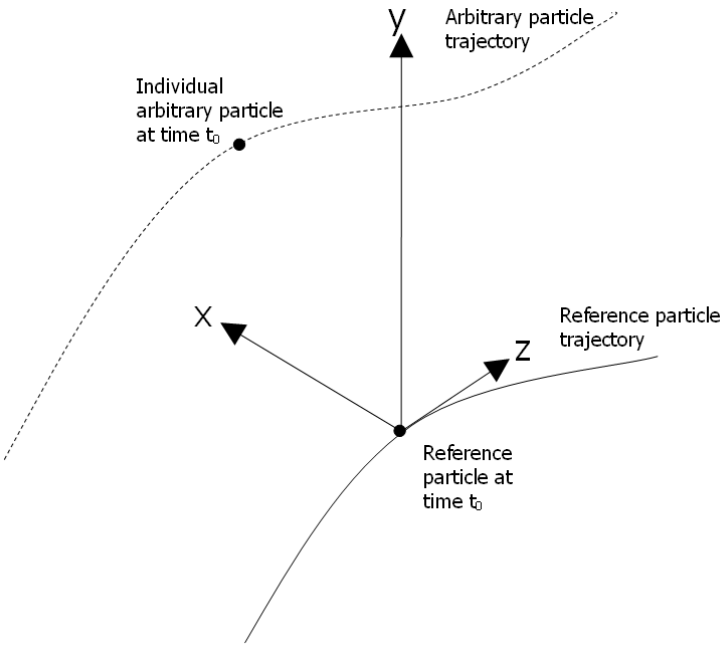

Figure 2.1: The electron beam coordinate system

Often, to extract useful information about the bunch, it is more convenient to consider projections of the 6-D hyper-volume onto three orthogonal 2-D planes (about the bunch centre). Providing there is no coupling between the axial and transverse motion or acceleration, the 4-D transverse sub-space is also governed by Liouville invariance. Therefore, the total phase space density can be written as the product of the three projected densities.

$$f(x, p_x, y, p_y, z, p_z) \quad = \quad f(x, p_x).f(y, p_y).f(z, p_z)$$

This implies that the area of the projected density function is a conserved quantity also. These areas are therefore a good measure of the electron beam, and are usually expressed in terms of the normalised beam *emittance*. When comparing electron sources the emittance is commonly used as the figure of merit. The emittance is a measure of the electron bunch size and divergence and is therefore often likened to the entropy of the bunch. Unfortunately there are a number of different definitions of emittance which can lead to some confusion.

In practice it is not possible to measure the full distributions in order to perform the integrals required for the emittance calculation. This is shown below for the transverse, horizontal direction:

$$\varepsilon_{n,x} = \frac{1}{\pi mc} \int \int f(x, p_x) dx dp_x$$

Rather than circumscribe the entire distribution, a *selection* of the particles is enclosed by a rms phase space ellipse, thus describing the central core of the distribution. Therefore it is an inexact figure of merit. In a relativistic transport system with no acceleration and only linear focusing and steering forces the area of the ellipse is preserved and remains constant, even though the form of the ellipse may change whilst moving through the accelerator. Calculated emittance from beam simulations often provides a snapshot in time of the electron bunch. In reality the emittance is measured as it passes through a fixed longitudinal point, z.

2.1.1 Horizontal Particle Motion

A simple description of the horizontal particle motion (assuming no momentum deviation) is given by Hill's equation [2]:

$$\frac{d^2x}{dz^2} + K(z)x = 0 \tag{2.1}$$

The general solution to this equation is:

$$x = \sqrt{\beta_x(z)\varepsilon_x} \cos\left(\phi_x(z) - \phi_0\right) \tag{2.2}$$

where ε_x and ϕ_0 are arbitrary constants that are dependent on the initial conditions. $(\phi_x(z) - \phi_0)$ is the phase advance, ε_x the emittance and $\beta_x(z)$ is the am-

plitude variation. Equation 2.2 describes a pseudo-harmonic oscillation with varying amplitude.

The emittance can also be written in terms of the horizontal particle motion and its first derivative. The derivation is given in appendix A.

$$\gamma_x x^2 + 2\alpha_x x x' + \beta_x x'^2 = \varepsilon_x \tag{2.3}$$

given

$$\alpha_x = -\frac{1}{2}\beta_x' \quad , \quad \gamma_x = \frac{1+\alpha_x^2}{\beta_x} \tag{2.4}$$

The ellipse parameters, α_x, β_x and γ_x are called the Twiss parameters, and these determine the orientation and shape of the ellipse. ε_x is a measure of the ellipse area divided by π. A particle whose motion is defined by equation 2.1 will move along an ellipsoidal contour given by equation 2.3. A second particle with a smaller amplitude function ($\beta_x(z)$) will also follow an ellipse, but it will be inside that of the first particle. Therefore if an ellipse is drawn around a sample of the beam, all those particles inside it will remain therein.

2.1.2 Beam Matrix

The equation of the phase space ellipse can be written as such by introducing the beam matrix, $\boldsymbol{\sigma}$:

$$\begin{pmatrix} x & x' \end{pmatrix} \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}^{-1} \begin{pmatrix} x \\ x' \end{pmatrix} = 1 = \mathbf{X}^T \sigma^{-1} \mathbf{X} \tag{2.5}$$

Since $\sigma_{12} = \sigma_{21}$ this can be expressed:

$$\sigma_{22} x^2 - 2\sigma_{12} x x' + \sigma_{11} x'^2 = \det \boldsymbol{\sigma} \tag{2.6}$$

Comparing this to equation 2.3 yields the following relationships between the Twiss parameters and the beam matrix:

$$\boldsymbol{\sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} = \varepsilon_x \begin{pmatrix} \beta_x & -\alpha_x \\ -\alpha_x & \gamma_x \end{pmatrix} \tag{2.7}$$

$$\varepsilon_x = \sqrt{\det \boldsymbol{\sigma}} = \sqrt{\sigma_{11}\sigma_{22} - \sigma_{12}^2} \tag{2.8}$$

The transfer matrix, $\mathbf{R}$, for the transverse x plane that describes the particle motion through a beamline consisting of non dispersive elements [1] is given by:

$$\begin{pmatrix} x \\ x' \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \begin{pmatrix} x_i \\ x_i' \end{pmatrix} \tag{2.9}$$

Using the identities $\mathbf{I} = \mathbf{R}^{-1}\mathbf{R} = \mathbf{R}^T\mathbf{R}^{T^{-1}}$ and inserting into equation 2.5 at the initial starting position, i yields:

$$\begin{aligned} \mathbf{X}_i^T(\mathbf{R}^T\mathbf{R}^{T^{-1}})\sigma_i^{-1}(\mathbf{R}^{-1}\mathbf{R})\mathbf{X}_i &= 1 \\ (\mathbf{X}_i\mathbf{R})^T(\mathbf{R}^T\sigma_i\mathbf{R})^{-1}\mathbf{R}\mathbf{X}_i &= 1 \\ \mathbf{X}^T(\mathbf{R}^T\sigma_i\mathbf{R})^{-1}\mathbf{X} &= 1 \end{aligned}$$

Therefore the beam matrix at the final position, f is given by:

$$\sigma_f = \mathbf{R}\sigma_i\mathbf{R}^T \tag{2.10}$$

Expanding this for the first beam matrix element gives:

$$\sigma_{11,f} = R_{11}^2\sigma_{11,i} + 2R_{11}R_{12}\sigma_{12,i} + R_{12}^2\sigma_{22,i} \tag{2.11}$$

[1] components that cause transverse motion depending on particle energy

where $\sigma_{11} = \langle x^2 \rangle$, $\sigma_{12} = \langle xx' \rangle$, and $\sigma_{22} = \langle x'^2 \rangle$. Therefore given the transfer matrix $\mathbf{R}$ and the starting conditions, σ_i, the sigma matrix at the final position can be calculated.

2.1.3 Phase Space Emittance

A statistical definition of the phase space normalised (n) rms emittance is given in equation 2.12. A normalised emittance accounts for the dependence on beam energy.

$$\varepsilon_{n,x,rms} = \frac{1}{\pi m_0 c}\sqrt{< x^2 >< p_x^2 > - < xp_x >^2} \tag{2.12}$$

where m_0 is the electron rest mass and c is the speed of light. Here $<>$ defines the second central moment of the particle distribution. i.e.

$$\begin{aligned} < x^2 > &= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2 \\ < p_x^2 > &= \frac{1}{n}\sum_{i=1}^{n} p_{x,i}^2 - \left(\frac{1}{n}\sum_{i=1}^{n} p_{x,i}\right)^2 \\ < xp_x > &= \frac{1}{n}\sum_{i=1}^{n} x_i p_{x,i} - \frac{1}{n^2}\left(\sum_{i=1}^{n} x_i \sum_{i=1}^{n} p_{x,i}\right) \end{aligned}$$

The factor of π shown in equation 2.12 is generally omitted from emittance equations but is sometimes acknowledged in the units (π m rad).

2.1.4 Geometric Emittance

From the rms normalised emittance the geometric (or beam) emittance can be written as:

$$\varepsilon_{rms} = \frac{m_0 c}{< p_z >}\varepsilon_{n,rms} \tag{2.13}$$

where $< p_z >$ denotes the mean longitudinal momentum.

During acceleration of the electron bunch the longitudinal momentum increases while the transverse momentum remains constant. Thus, the geometric emittance decreases as $< p_z >$ increases, whilst the normalised phase space emittance does not show this feature. This is *adiabatic damping*. The geometric emittance can be directly related to the Twiss parameters given in equation 2.3.

2.1.5 Trace Space Emittance

Experimentally it is not normally the transverse momenta that are measured but the divergence x', y'. By using the identity $x' = p_x/p_z$, $y' = p_y/p_z$, a good approximation for paraxial beams[2], the normalised trace space (tr) emittance is given by substituting into equation 2.12 (ignoring a factor of π):

$$\varepsilon_{n,x,tr,rms} = \frac{< p_z >}{m_0 c}\sqrt{< x^2 >< x'^2 > - < x\, x' >^2} \tag{2.14}$$

Finally, the trace space emittance is given as:

$$\varepsilon_{x,tr,rms} = \sqrt{< x^2 >< x'^2 > - < x\, x' >^2} \tag{2.15}$$

The statistical emittance can be related to the Twiss parameters as shown in figure 2.2.

$$\begin{aligned} x_{rms} &= \sqrt{\varepsilon_{rms}\beta}, \quad x_{int} = \sqrt{\tfrac{\varepsilon_{rms}}{\gamma}} \\ x'_{rms} &= \sqrt{\varepsilon_{rms}\gamma}, \quad x'_{int} = \sqrt{\tfrac{\varepsilon_{rms}}{\beta}} \\ Area &= \pi\varepsilon_{rms} \end{aligned}$$

An example of a trace space ellipse superimposed onto a simulated distribution of particles is shown in figure 2.3. The curved tails of the distribution are from the space charge forces in the electron bunch and are typical of what is seen

[2]Beams with small divergence and particles close to the axis

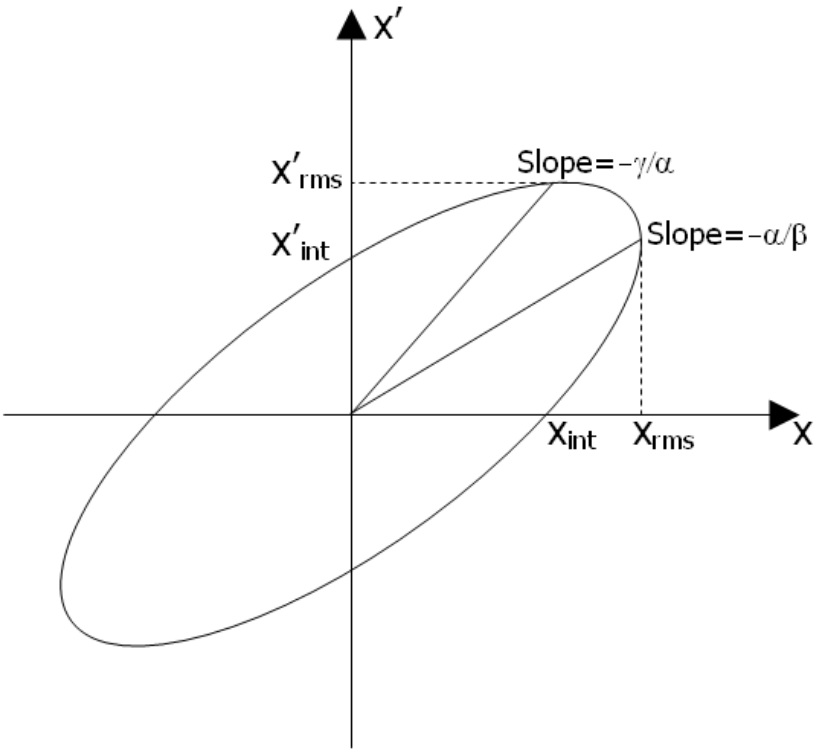

Figure 2.2: The relation between statistical measures and Twiss parameters

experimentally. Figure 2.3 demonstrates that in some cases an rms ellipse does not describe the core of the distribution very well, particularly when the distribution has outlying points. For this reason it is desirable to know the shape of the distribution in trace space, in addition to the emittance, as the tails can be responsible for beam loss or low gain FEL lasing for example. Therefore when the distribution of particles in trace space becomes very disorganised, measuring rms emittance can be less meaningful.

To add additional confusion, the terms trace and phase space are often used interchangeably, so it is important to note the method of calculation and the axes used on graphs. However, in most situations the numerical difference between trace and phase space is negligible. Throughout this thesis the term phase space is used, though the numerical calculation is that of trace space.

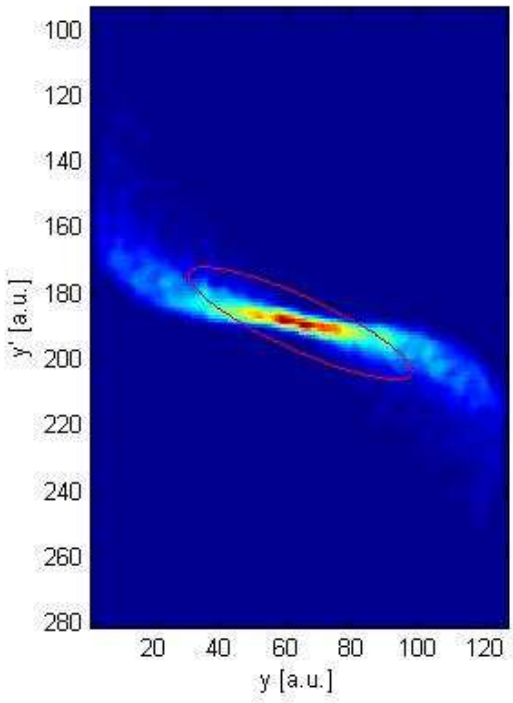

Figure 2.3: Trace space ellipse superimposed onto a particle distribution

2.1.6 Thermal Emittance

When the emittance of a real system is measured, the thermal emittance is always included. This is the source emittance that is present at the cathode and imposes a lower limit for the normalised emittance that can be achieved by an injector. The normalised rms thermal emittance depends on the emitting area, the momentum distribution, and the angular distribution of the emitted electrons. The energy and divergence distributions are functions of the cathode material and photon energy (for photocathodes). To create an accurate model of a specific photocathode, it is important to understand the photoemission process in that particular material. Typically, the thermal emittance is measured from experiment, usually of the order of $0.5\mu m$, and this value can be used in modelling. The thermal emittance becomes important when striving for sub-micron total emittances in the region of FELs. The total emittance is a combination of that induced by space charge or RF and the intrinsic thermal emittance.

The laser spot size can be set to reduce any emittance growth in the low energy region due to space charge forces or RF components. The thermal emittance can be calculated by assuming the electrons from the cathode are emitted uniformly and isotropically, within a radius r in the presence of an accelerating field. The angular distribution, x', then has a Maxwell-Boltzmann distribution and $\langle x'^2 \rangle$ can be calculated as kT/mc^2, where the cathode is at a temperature T, and $\langle x^2 \rangle = \sigma_0^2 = r^2/4$ [21]. Inserting into equation 2.14 with $\langle xx' \rangle = 0$ yields the expression for thermal emittance, equation 2.16.

$$\varepsilon_{th,n,rms} = \sigma_0 \sqrt{\frac{kT}{m_0 c^2}} = \sigma_0 \sqrt{\frac{E_{kin}}{m_0 c^2}} \tag{2.16}$$

where σ_0 is the rms of the emitting area (m) and E_{kin} is the average thermal energy (eV).

2.1.7 Emittance Compensation

Although it is not possible to reduce the thermal emittance once the electrons have left the cathode, it is feasible to minimise the emittance growth from linear space charge forces. Emittance growth from the cathode is a combination of space charge forces in the bunch and transverse components of the accelerating field. Carlsten [22] notes that by careful positioning of a lens, it is possible to eliminate the emittance growth due to linear space charge in the bunch after a drift. This assumes that, to first order, the transverse space charge forces act as a defocusing lens. For a typical Gaussian electron bunch the strength of the defocusing depends on the longitudinal position in the electron bunch, as the space charge is strongest in the middle of the bunch and decreases towards the ends. The compensation can be seen if the slice emittance of the distribution is used. The slice emittance is calculated by dividing the electron bunch up into slices longitudinally from tail to head. The emittance of each slice is then calculated. An example of an electron beam from a DC gun with a compensation

solenoid at two different locations is shown in figure 2.4. The phase space is plotted at a distance of 70cm from the cathode and 108cm where the emittance is smallest in this case. The beam has been divided into 6 longitudinal slices and the ellipse drawn for each of these. Figure 2.4 (a) shows that the ellipses are all at different orientations, whilst the ellipses in the compensated case (b) have very similar angles except those from the very head and tail of the distribution. The minimum total emittance occurs as these slice ellipses all align. Variations in the energy distribution along the electron bunch, and non-linear space charge forces that act on the extremes of the distribution, contribute to incomplete compensation.

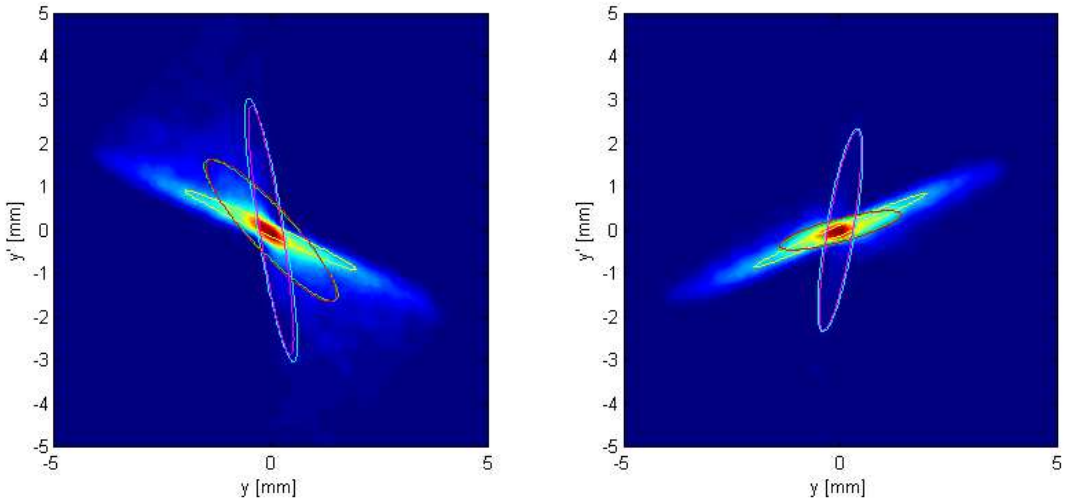

Figure 2.4: Phase space and slice emittance ellipses of a non compensated (a) and a compensated (b) beam

2.1.8 Longitudinal Emittance

In a similar fashion to the transverse emittance, the normalised longitudinal emittance can be calculated using:

$$\varepsilon_{n,z,rms} = \frac{1}{\pi m_0 c}\sqrt{< z^2 >< p_z^2 > - < zp_z >^2} \tag{2.17}$$

2.1.9 Brightness

The concept of emittance can be extended to include longitudinal bunch information of the current density. This is called the brightness of the beam or bunch [23], and is defined as:

$$B_n = \frac{2I}{\pi^2 \varepsilon_{n,x} \varepsilon_{n,y}} \tag{2.18}$$

B_n is the normalised brightness (A/m^2). Again, the factor of π is generally omitted. I is the current in the bunch and either I_{peak} (peak current) or I_{avg} (average current) can be used to give peak or average normalised brightness respectively. The following section will show the importance of having a high peak current for high gain FELs and that a high average brightness is required for oscillator FELs. It is clear from equation 2.18 that a combination of small emittance and large current will result in a high brightness beam.

2.1.10 Usage

For injector simulation usually the normalised rms geometric, or phase space emittance, is calculated to give meaningful results from the modelling. It is important to note whether the emittance is calculated at a particular time or position. Experimentally, it is normally the trace space emittance that is measured at a particular longitudinal position. Therefore the electron bunch is projected onto a transverse plane as it passes through that location, for example when it hits a view screen. The emittance measured can consequently be slightly different from that simulated at the *time* the centre of the bunch reaches that same location (as there is some longitudinal distribution).

Care should also be taken when using the trace space emittance as it can display non-physical behaviour in regions with large energy spread or divergence [24]. Since in real systems it is the trace space emittance that is commonly measured, simulation should be checked to see that the difference between phase and trace space emittances is minimal for comparison.

Optimising the emittance depends on the application. Sometimes the slice or some percentage of the beam emittance is a more useful measure of beam quality. This can, for example, be important for short wavelength amplifier FELs where it is thought that only part of the electron bunch contributes to the lasing process.

2.2 FEL Theory

In a free electron laser, relativistic electrons propagate through a spatially periodic, alternating (transverse) magnetic field, superimposed with an optical field. The optical field can either be seeded (from a laboratory laser or another synchrotron radiation source), or generated spontaneously from the electron beam itself. The periodic magnetic field is provided by either a wiggler or undulator. The undulator field causes electron bunches to oscillate transversely and spontaneously emit synchrotron radiation, as shown in figure 2.5.

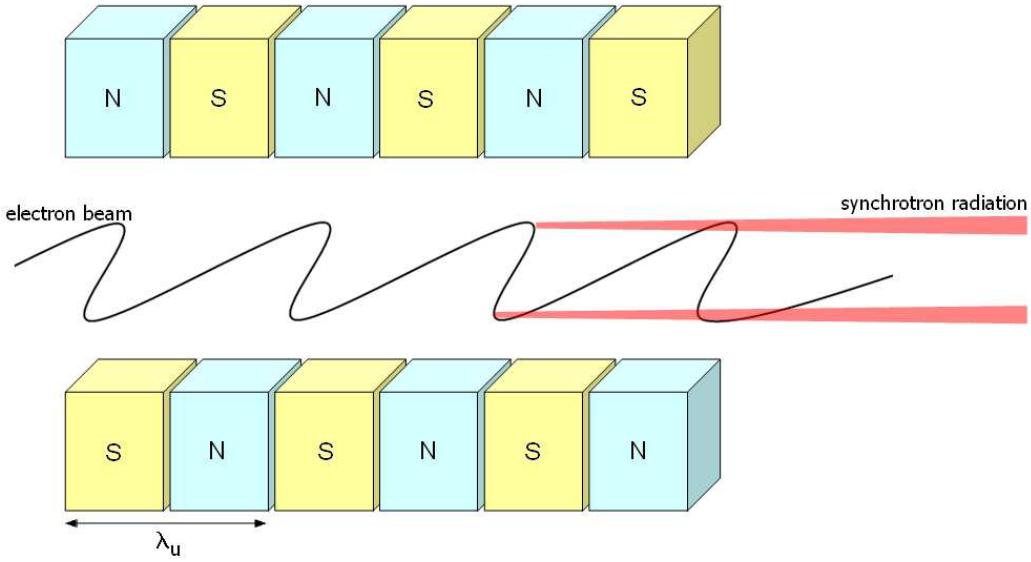

Figure 2.5: Synchrotron radiation from an insertion device

The forward radiation from an undulator is emitted in a cone of angle $\theta \sim \frac{1}{\gamma}$, and has a peak, or resonant, wavelength of [25]:

$$\begin{aligned} \lambda_r &= \frac{\lambda_u}{2\gamma^2}(1+\frac{K^2}{2}) \\ K &= \frac{eB_u\lambda_u}{2\pi m_e c}, \gamma = \frac{E}{m_e c^2} \end{aligned} \tag{2.19}$$

where λ_u is the undulator period (m), K the undulator deflection parameter, γ the resonant energy, E the electron energy (eV), B_u the undulator peak magnetic field (T), c the speed of light (ms^{-1}), and e the elementary charge (C).

As the electron bunch propagates through the undulator, the transverse motion of the electrons couples to the transverse electric component of the optical field, giving rise to energy transfer. If the relationship between the electron and electric field is correct, the optical field gains energy from the electron motion. Whether the electrons absorb from, or relinquish energy to the optical field depends on the phase. If the electron bunch is input into the undulator at the resonant energy, the electrons have a random phase distribution along the bunch, so initially both processes occur simultaneously resulting in no net FEL gain. The energy transfer gives rise to an energy density modulation of the electron bunch, which in turn causes longitudinal bunching as the path length through the undulator is energy dependent. The increased bunching results in more coherent emission, which induces more bunching, and so on, until the modulation reaches a maximum, and the emission is fully coherent. Since most of the electrons within a very short bunch have very nearly the same phase they emit with a high degree of coherence. The average phase between the optical field and the electrons actually remains constant as the electrons pass through the undulator because they slip back one radiation wavelength every undulator period. This occurs because of the longer path length encountered by the electrons and their less than the speed of light velocity.

FELs can be configured in two ways, either as an amplifier or an oscillator. In an amplifier, as its name suggests, the spontaneous radiation is amplified as it passes through the undulator on a single pass. For the radiation to reach saturation, the undulators need to be long enough for this to occur. Saturation occurs when electrons that have given energy to the optical field begin to reabsorb it. With an oscillator, or low-gain FEL, the radiation is partially confined between two mirrors either side of the undulator, so that it traverses the FEL cavity many times, interacting with the electron beam. Because of the requirement for mirrors, the oscillator can only be used for radiation which can be reflected efficiently, making it difficult to realise areas of the short wavelength spectrum beyond the UV using this method [10]. In a low-gain FEL the undulator is relatively short and to achieve maximum gain, the electrons should be injected at an energy slightly higher than the resonant energy. Roughly 25% of the radiation emitted in a single pass is permitted to escape the optical confinement, whilst the remaining portion sustains the FEL process. Saturation in this situation occurs when the power extracted from the electron beam equals that coupled out from the optical cavity.

In the first experiments, the FELs in the amplifier (high-gain) configuration were seeded with input lasers to begin the FEL process. It was later realised that an electron beam instability (i.e. noise within the electron bunch) could result in the spontaneous exponential growth of radiation to a high gain regime without a seed laser. In this scenario, the portion of the spontaneous emission that fulfils the resonant lasing criteria (i.e. equation 2.20) is amplified along the undulator. This mode of operation, termed Self Amplified Spontaneous Emission (SASE), gave the potential to extend FEL sources to the higher energy end of the spectrum. The undulators for high-gain FELs tend to be very long so that saturation can be achieved. The interaction between the radiation and the electrons again causes a charge density modulation as some electrons gain, and others lose energy. For short wavelength FELs, the wavelength is short compared to

the electron bunch length, and so results in micro-bunching on the scale of a radiation wavelength within the electron bunch.

2.2.1 Gain Degradation

The gain of a FEL can simply be determined by the ratio of electron beam power to optical laser power. The strength of coupling between the electron beam and the optical field is given by the small gain coefficient (single-pass, continuous electron beam with negligible energy spread and transverse emittance). Following reference [26], for a planar undulator this is:

$$g_0 = \frac{16\pi}{\gamma}\lambda_r L_u \frac{J}{I_A} N_u^2 \xi F(\xi) \tag{2.20}$$
$$\xi = \frac{1}{4}\frac{K^2}{1+K^2/2},\ F(\xi) = (J_0(\xi) - J_1(\xi))^2$$

where L_u is the undulator length, N_u the number of periods, I_A is the Alfvén current (17kA for electrons), J the current density, and J_k are Bessel functions. It can be seen from equation 2.20 that the gain is proportional to the peak electron beam current and inversely proportional to beam energy. In real systems, beam emittance and energy spread degrade the quality of the FEL interaction by reducing the gain per unit undulator length. The maximum ideal gain from a system is $G_{max} = 0.27\pi g_0$ [26]. By correlating the effect of gain degradation with electron beam parameters it is possible to determine the criteria for electron beam quality [27]. Typically the normalised transverse emittance must be:

$$\varepsilon_{n,rms} \leq \frac{\gamma\lambda_r}{2\pi} \tag{2.21}$$

where $\gamma = \frac{1}{\sqrt{1-\beta^2}}$ and $\beta = \frac{v}{c}$. This implies for lasing at nm wavelengths, a transverse emittance of $\sim 0.5\mu$m is required, and this is close to the thermal emittance of most photocathodes. Equation 2.21 suggests that by increasing the

energy of the electron beam the requirements on emittance become less stringent. However, not only is the wavelength dependent on electron energy, but also the small signal gain is inversely proportional to it, which would lead to a weaker coupling of the FEL process.

2.3 Conclusions

Low emittances are required for lasing, and the shorter the FEL wavelength, the smaller the emittance needed. The small emittance from the cathode must therefore be preserved through the accelerator until it reaches the FEL. The output power from a FEL is proportional to the current of the electron bunch, so this must be simultaneously maximised. A high peak current implies a high charge density which increases the FEL gain.

The emittance and brightness are limited by a number of contributing factors. One factor is the thermal emittance of the electron source. The thermal distribution of the emitted electrons determines the lower limit to the achievable emittance. There are also the space charge forces within an electron bunch, that act to expand the bunch in the first stages of acceleration, before the relativistic effects at higher energy make space charge negligible. This process works in both longitudinal and transverse planes and becomes worse with increasing bunch charge. The transverse forces are not entirely linear, and so cannot be compensated for fully with the use of solenoids. A final consideration is that all electrons within a bunch do not experience the same accelerating electric fields from RF cavities. This gives rise to an energy spread along the electron bunch.

The trade-off between charge density within a bunch and the emittance must be carefully balanced within the injector. Higher charge densities are beneficial to the lasing process, as are low emittances. However at the low energies in the injector, a high charge density serves to degrade the emittance due to space

charge forces. A poor emittance in turn will degrade the lasing process in the FEL.

CHAPTER 3

High Brightness, High Average-Current Electron Sources

The production of electrons for acceleration can be achieved in many ways, from traditional thermionic guns to novel combinations of photo and field emission sources. A brief overview of these electron sources is given in the following sections.

Once electrons have been produced they must then be accelerated by an electric field, and so all cathodes have an applied field at the surface to move the electrons in the required direction. Again, methods vary for different applications. The most common for use within an ERL or other high brightness machine are discussed in this chapter.

3.1 Electron Sources

Cathodes can be generalised as surfaces that emit electrons under the influence of some stimulating energy. This energy could be in the form of heat, light, kinetic energy from incident particles, or alternatively, the electrons may be emitted in the presence of high electric fields at the surface. There are many different classifications of cathode depending on how electrons are stimulated or on the

geometry. For example there are hot cathodes, cold cathodes, photocathodes, field emitters, secondary emitters, and hollow cathodes to name a few. Ideally the cathode should emit electrons freely and plentifully with zero momentum spread, and should have an infinite lifetime.

3.1.1 Thermionic Sources

Thermionic emission is the oldest method of liberating electrons from a material. Electrons are effectively evaporated from a heated surface. To escape the material, electrons must have a component of velocity perpendicular to the surface through their kinetic energy. The kinetic energy of an electron must be at least equal to the work done in passing through the surface. This minimum energy is known as the work function, ϕ (eV), and is material specific. The work function is the minimum energy needed to remove an electron from a solid to a point immediately outside the solid surface. The energy to overcome the work function arises purely from the thermal energy of the system. The saturated emission from a heated cathode is given by the Richardson/Dushman equation:

$$J_0 = A_0 T^2 e^{\frac{-e\phi}{kT}} \tag{3.1}$$

Where J_0 is the saturated thermal emission current density (A/mm^2), A_0 is Dushman's constant (the theoretical value is 120.4 A/cm^2/K^2 which is not attained for real materials [28]). k is Boltzmann's constant (8.6$\times$10^{-5} eV/K), e the charge of an electron (eV) and T is the cathode temperature (K). Current is dependent on both the temperature and the emitting area of the cathode, as is thermal emittance (discussed in chapter 2). A good cathode however, will produce a high current and a low thermal emittance, so there is a trade-off to be made. Equation 3.1 does not take into account any electric field at the cathode surface that is applied to accelerate the electrons away. As a result, a voltage (from an anode adjacent to the cathode) or electric field must appear in the equations. It is found that when an electric field is applied to the surface of an emitter, the

emission increases. The applied field reduces the energy barrier that an electron must overcome in order to escape the surface. In effect the work function is reduced by $\sqrt{\frac{eE}{4\pi\varepsilon_0}}$, where E is the applied field and ε_0 is the permittivity of free space. The emission is therefore given by:

$$J = J_0 e^{\sqrt{\frac{eE}{4\pi\varepsilon_0}}} \tag{3.2}$$

Thermionic sources are simple devices and avoid the need for the complex laser drive systems required for photocathodes, and as a result remain popular in some laboratories. However producing flexible, high peak current pulse trains is difficult, as electrons are constantly emitted from the cathode surface. Pulsed electric fields can be used to suppress emission from the surface, but there are limitations on the repetition rate and duty cycle resulting in relatively long pulses ($\sim$ns). Therefore a significantly more complex injector stage is required to reduce the bunch length. Nevertheless, a thermionic gun has been used for the JAERI FEL [29, 30] with 230kV DC acceleration. Subsequent compression stages are used to shorten the bunch length. Such an electron source has also been proposed for the Spring-8 Compact SASE Source (SCSS) with DC acceleration to 500keV [31]. The cathode is pulsed to produce 1.6μs (FWHM) 500keV bunches, which are shortened by a number of bunching stages to 12ps (FWHM) at 50MeV.

Secondary Emission

Secondary emission occurs when electrons with sufficient energy bombard others in the lattice structure of the cathode and cause other electrons to be emitted from the surface. For many cathode materials high secondary emission is accompanied by short lifetime. When a cold cathode is used, that is one which is not actively heated, all the emission is a combination of field (discussed in the next section) and secondary. Cold emitters must have a low work function at

operating temperatures to emit at all, and the presence of an electric field can lower it further. The amount of secondary emission depends on the energy and angle of incidence of the primary electron and also the material of the cathode itself. Due to the emission process, the thermal energy of the electrons emitted is lower than that from room temperature thermionic cathodes.

3.1.2 Field Emission Sources

As the electric field at the surface of the cathode is increased to the 10^3 - 10^4 MV/m level, it is found that electron emission increases rapidly [32]. Furthermore, the increase is almost independent of temperature. With a high applied field the potential barrier at the surface is very narrow, and even though the kinetic energy is not sufficient, electrons can escape the surface via tunnelling. This is the Schottky effect, and after onset, field emission increases exponentially with applied field.

High fields at cathode surfaces can be achieved through the geometry of the cathode. For example, needle cathodes which have very narrow tips ($\sim 1\mu$m) can have a field gradient on the surface of $10^3 - 10^4$ MV/m for an applied cathode-anode voltage of just 50kV. These single-needle cathodes are being developed for use with table-top FELs [33], and use a laser to gate the electron emission through the photo-electric effect. In this way very short pulses, that roughly follow the laser temporal profile can be obtained [34, 35]. As an alternative, arrays of needle cathodes can be used as field emitting cathodes, see figure 3.1.

The emission from these cathodes is gated through a surface layer that locally suppresses the field with a bias voltage [36]. To further enhance the field at the tip, carbon nanotubes can be grown on the vertex [37]. As the emitting area is so small, and the geometry of the cathode provides almost parallel electric field

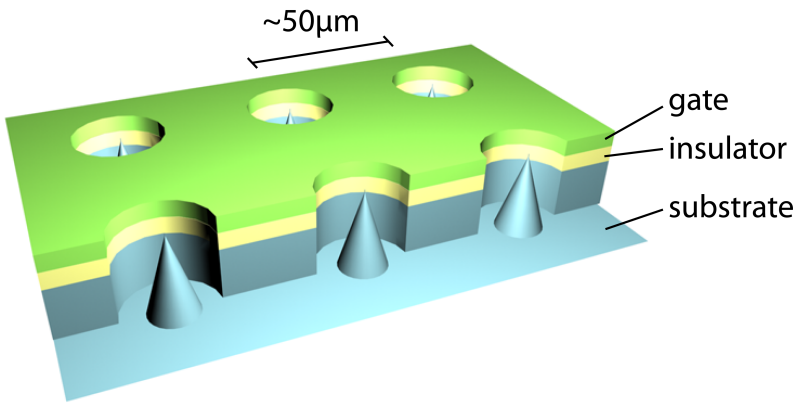

Figure 3.1: Field Emission Array

gradients, the emittance from these cathodes promises to be excellent. As yet, they are untested in an accelerator style gun.

3.1.3 Ferroelectric Sources

Ferroelectric materials exhibit a permanent electric dipole moment in their bulk and can spontaneously produce electrons without an applied external extraction field. The electron emission results from spontaneous bulk polarisation switching, which generates a high surface electric field that expels the layer of compensating electrons. In changing the dipole moment, electrons can be separated from the crystal lattice and emitted from the material. A change in the direction of the dipole moment can be induced by the application of a submicrosecond external electric field, irradiation with a laser, acoustic waves or heating. The use of a laser to stimulate electron emission is the most promising method of producing short electron pulses, however the shape of emission is not directly controlled by the laser. The electrons are not produced via the photoelectric effect so the wavelength need not be matched to the cathode. An example of such a cathode being used in an RF gun is given in [38]. In this scenario the current produced is high, the duty cycle is low, and the pulse length long, making this inappropriate for FEL machines.

3.1.4 *Photoelectric Sources*

Photoemission is described as the escape of electrons from a material after the absorbtion of photons above a threshold frequency. The energy of the photons must be greater than that of the work function for electrons to be emitted (i.e. $hv \geq \phi$, h - Plank's constant, v - photon frequency, ϕ - work function). Any additional energy from the photon, over the work function, is imparted as kinetic energy to the emitted electron (contributing to thermal emittance). The photoemission process from a solid, particularly a semiconductor, is commonly described by the three-step model [39]:

1. A photon is absorbed into the material and an electron-hole pair is created. This is followed by electron thermalisation to the bottom of the conduction band through electron-phonon[1] collisions.
2. The electron transports to the surface by diffusion.
3. The electron passes through the surface layer and escapes into the vacuum if it has enough energy to overcome the potential barrier, otherwise it will recombine.

It is obvious that the more electrons emitted for a given intensity of photons the better the photoemitter. This figure of merit is called the quantum efficiency (QE) and is used when describing the photoemission material, or photocathode, in an accelerator. The photo-current from a photocathode is given by 3.3.

$$\begin{aligned} q &= \eta E_{laser} \frac{e\lambda}{hc} \\ i &= \eta P_{laser} \frac{e\lambda}{hc} \end{aligned} \tag{3.3}$$

Where, q is the charge (C), i current (A), η quantum efficiency, E_{laser} laser energy (J), P_{laser} laser power (W), λ laser wavelength (m), e elementary charge

[1] Thermal energy transferred by lattice vibrations in waves. The waves behave like particles (phonos) that possess energy and momentum.

(C), h Plank's constant (J s). The equation shows that the quantum efficiency of a photocathode is dependent not only on the power of the laser, but also on the wavelength (assuming the work function condition is met). Some photocathodes, such as metals and alkali tellurides, have their highest QE when illuminated with ultra-violet light. Even then, the QE can be much less than one percent. Lasers at this wavelength with power high enough to get adequate current are costly, and the frequency multiplication stages required to get ultraviolet light, degrade the spacial and temporal distributions. Therefore the laser-cathode combination must be considered. At the moment UV lasers that can operate at the power and repetition rates for a 100mA machine do not exist. It is also important to note that some portion of the laser light can be reflected from the surface of the photocathode, so not all photons are absorbed in the bulk. For a high quantum efficiency photocathode, the electrons must be excited to an energy greater than the work function. In addition, the ratio of absorption length for the photon, to the escape length of the electron must be small. If the absorption takes place over a long length compared to the escape length, only a few electrons will make it to the surface, and the QE will be low. This escape length is determined by the dominant scattering mechanism in the material. In metals, the scattering is mostly through inelastic collisions with conduction band electrons, and the escape length is approximately equal to the electron-electron mean free path. In each collision a photoelectron will lose a significant amount of its energy and therefore is less likely to have enough energy to overcome the surface barrier. This electron is then lost for the photoemission process. By contrast, semiconductors have electron-phonon scattering as the dominant mechanism for photoelectrons. The electrons lose only a small portion of their energy on colliding with a phonon, and therefore have a much greater chance of reaching the surface with sufficient energy to overcome the potential barrier. As a result, the escape length can be much greater than that in a metal, which yields a higher quantum efficiency.

The response time of a photocathode is another important parameter to consider. This again can be pictured using the three-step model. It is determined by the time distribution between the absorbtion of photons and the emission of electrons. The escape depth in metals is very short, so the response time is very fast, on the order of femtoseconds. In this situation the temporal distribution of an electron bunch emitted will be almost identical to the distribution of the laser used. Semiconductors, with longer escape depths, can have response times as long as several nanoseconds. This gives rise to the trend that high quantum efficiency materials tend to have longer response times, as the escape lengths are generally longer. Because the photon absorption is a function of the laser wavelength, the response time for a given material can vary as a function of wavelength. A long response time leads to the temporal length of emitted electron bunches being greater than that of the incident laser used, and also the shape of the distribution no longer directly follows that of the laser.

Of the high QE semiconductor photocathodes, there are three commonly used broad classes [40]:

- Alkali antimonides: Cs_3Sb, K_2CsSb [41, 42].
- Alkali tellurides: Cs_2Te [43], KTe, CsKTe [44].
- Negative electron affinity (NEA) semiconductors: GaAs, GaAsP.

The alkali based semiconductors have a positive electron affinity (PEA). The electron affinity of a material is the difference in energy between the bottom of the conduction band and the vacuum level outside. In a PEA photocathode the electrons must escape before they are thermalised to the bottom of the conduction band, from which they cannot escape. In NEA cathodes the vacuum level is below that of the conduction band minimum (CBM). The electrons can escape from the bottom of the conduction band and may become thermalised before emission as the conduction band energy is below that of the vacuum. This is shown schematically in figure 3.2. For this reason NEA cathodes have the ad-

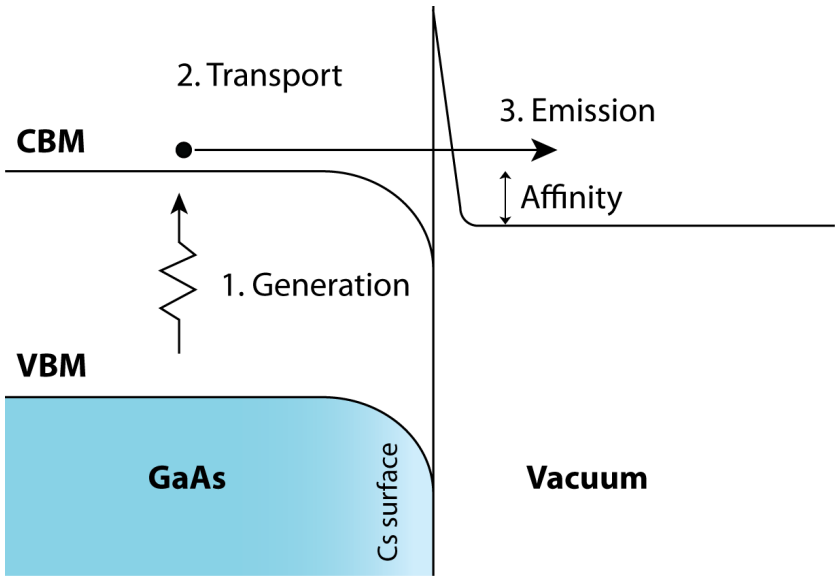

Figure 3.2: Band Structure and 3 Step Photoemission in NEA GaAs

ditional advantage of emitting electron bunches with very low thermal energy or emittance. Bulk gallium arsenide has a slightly positive electron affinity of 4eV. However, the work function of GaAs can be lowered to a point where the band gap electrons can escape, by building a dipole moment surface layer with caesium and an oxidant (e.g. O_2, NF_3, N_2O). The two NEA semiconductors important to the studies in this thesis are GaAs and GaAsP. They have both been used in various DC guns around the world, in particular at SLAC, JLab and Cornell [45, 46]. The thermal energy of these cathodes is low, below 250meV, and depends on several factors; the illuminating wavelength, the degree of negative affinity, and also the band structure within the photocathode material. There is a trade-off between improving thermal energy and the quantum efficiency; as one becomes higher the other decreases. A disadvantage of GaAs photocathodes is that they have been measured to have a relatively long photoemission tail ($\sim$300ps) when excited with photons near the band gap of 1.4eV [47]. Since the absorption length in GaAs is a function of photon energy, the use of shorter radiation wavelengths for excitation may be preferable due to a faster photoemission response time. The response time for these wavelengths will be discussed in chapter 5.

Given the excellent properties of these GaAs photocathode materials, it is unfortunate that they cannot be used effectively in normal conducting RF guns due to very short lifetime (a few seconds) [48]. The primary disadvantage of these semiconductors is that they are easily contaminated and damaged. They must be operated under very high vacuum (10^{-11} Torr) if they are to survive. Cathode poisoning can occur from residual gas in the vacuum and reduces the QE, as does damage from ions that are accelerated towards the cathode and bombard it. The vacuum in RF guns is generally a few orders of magnitude worse than in DC guns because of the fabrication material and the electric fields inside, so GaAs cathodes are not really appropriate. Robust cathodes, such as metals, which effectively never degrade, and the alkali cathodes which have longer lifetimes are used in RF injectors. However, the vacuum conditions in a SRF gun are comparable to that in a DC gun, therefore theoretically GaAs cathodes could be used [49]. Although the lifetime of all the semiconductor cathodes are short compared to that of metal they can usually be regenerated through a heat cycle, or a new Cs layer. Ion back bombardment is not such a problem in RF guns as it is in DC guns, as the accelerating electric field is oscillating. In a DC gun ions tend to be accelerated back towards the centre of the cathode, creating a damage spot. Damaged alkali cathodes in RF guns can still operate with a high QE, but may start to exhibit field emission under the high gradients within the cavity.

Niobium and lead have also been considered for use as photocathodes. Niobium is an obvious choice for a SRF gun where the cavity itself could be used [50], but like other metals it has a low QE and the laser required has a UV wavelength. A high repetition rate of high power laser pulses on the superconducting niobium could cause enough heating to quench the cavity. To try and avoid these problems, lead deposited onto the back wall of a niobium cavity has been tested as it has a slightly higher QE (0.5%) [51]. This may be promising for pro-

ducing modest current but so far it does not meet the requirements of a high average current injector.

Recently there has been a novel development in the area of high QE cathodes. A group at Brookhaven National Laboratory have proposed, and started to prototype, a diamond amplifier photocathode. The principle of this cathode, as described in depth in [52], is briefly summarised here. Primary electrons are produced by a normal photocathode and are accelerated to a few keV over a short distance. These then strike a diamond window, where the large secondary electron yield of diamond multiplies the number of electrons by two orders of magnitude. The advantage of such a scheme is that the laser power on the photocathode can be greatly reduced. The diamond window also separates the cathode from gun, so the cathode is not poisoned. This additionally means that it can be used in a SRF gun which is sensitive to contaminants from the cathode. Testing of this cathode is already underway which shows favourable results [53], but the secondary electron yield has been lower than anticipated.

3.1.5 Electron Source Conclusions

Thermionic cathodes have a long, well tested history of use as a robust source of electrons. Although a thermionic gun has been used as the source for the JAERI FEL, the repetition rate was low (20MHz) and the duty factor 1%. The emittance from the gun was reported as 13μm which is adequate for the long 24μm wavelength FEL. This type of electron source is not suitable for the JLab FEL which lases at a much shorter wavelength, 1μm, and at a higher repetition rate (75MHz, 100% duty factor). Additionally, field emission and ferroelectric sources are also not appropriate as they are far too immature a technology to be considered as upgrade options. This leaves photocathodes as the best choice for the Jlab FEL electron source. The decision of cathode material largely depends on the laser wavelength and desired type of acceleration (details in the next section). High quantum efficiency is also beneficial in a cathode. Table 3.1

summarises the properties of the different types of photocathode. The lower risk choice would be a GaAs cathode because of the high quantum efficiency and relative ease in obtaining a high power, high repetition rate, visible length laser.

	Metal	Alkali Semi.	NEA Semi.	Diamond Amplifier
QE	$\sim$0.1%	$\sim$1%	$>$2%	$>$100%?
Laser λ	UV	UV	Visible$\rightarrow$IR	*
Lifetime	Very long	Long	Short	Very long
Response time	Prompt	Prompt	$f(\lambda)$	*

*source dependent

Table 3.1: Comparison of Photocathodes

3.2 Electron Injectors

The method of accelerating electrons produced from a cathode generally falls into one of two categories; DC or RF acceleration, and both methods are used with thermionic and photocathodes. The benefits of each technique are discussed in the following sections.

3.2.1 DC Acceleration

In a DC gun an electrostatic field is applied between the cathode and anode, and is often operated continuously. Emitted electrons flow from the cathode towards an anode, which is maintained at a positive potential with respect to the cathode (i.e. the anode is normally grounded and the cathode has a negative potential for safety). A hole in the anode allows the electrons to escape from the accelerating structure. The energy the electrons gain is the product of their charge and the potential difference between cathode and anode. The higher the energy, the more resistant the bunch appears to be to internal space charge forces as viewed in the laboratory frame. Therefore the higher the potential that can be sustained, the better the performance of the gun.

The maximum static voltage that can be applied is limited by the electrical breakdown of the gas the electrodes are in. Insulating gases such as SF_6 are commonly used around high voltage parts, such as the power supply to increase the breakdown gradient to approximately 10MV/m. The cathode and anode of an electron gun are situated in vacuum, so there is no gas to permit electrical breakdown. However, in areas where the local gradient is high, field emission may begin. Sharp edges can lead to very high local gradients, so the electrodes must be carefully designed to minimise this and also be very smooth. The electrodes on the JLab DC gun are hand polished to a surface roughness of $< 1\mu$m for this reason. Not only is this extremely time consuming, but it is also not fool-proof. The slightest scratch can induce field emission, a process which can change the surface itself through electrons being ripped from the material. Whiskers can form on the surface, which induce more field emission, and so on until it is no longer possible to operate at that gradient [54]. To try and minimise this effect, the electrodes are conditioned by slowly increasing the voltage, whilst watching the current drawn from the power supply and the vacuum level. Emission can be seen by a rise in either of these values. Field emission sites can mostly be removed by ramping of the voltage until they are pulled from the surface or melted by emission current. To ensure that operation is not disrupted by breakdown of this sort, the electrodes are conditioned to a higher voltage than for nominal operation. This limit on maximum gradient is one of the disadvantages of DC acceleration, as electrons typically exit the structure with only a few hundred keV energy. At this low energy, the effect of space charge forces within the electron bunch is very pronounced, and so further acceleration is required before the beam is degraded further. An advantage of DC acceleration is that the fields are purely electrostatic, and these can easily be shaped to give focusing via the cathode-anode geometry. Furthermore, the shape and construction of DC guns allows for many vacuum pumps to be attached, and so the vacuum

achieved is superior to that in RF guns. Photocathodes can often be sensitive to poor vacuum, so improving the vacuum increases the number of viable options.

DC guns are most commonly used with both thermionic and photocathodes. Thermionic guns evolved first, and have been used as the electron sources for synchrotrons since their conception, but cannot be used to produce the high brightness beams required for short wavelength, high power FELs. It is possible to pulse a high voltage DC supply for use with thermionic or field emission sources to generate pulsed, rather than continuous electron beams. There is however, a limit to the frequency with which the DC field can be switched on and off (MHz), and therefore a minimum pulse length achievable using this method. Subsequent bunch compression is then required to generate short (ps) bunches. Through pulsing the DC field it becomes possible to sustain higher gradients in the gun without breakdown, as they are only there transiently for a few ns. Recently a pulsed DC gun, that can produce in excess of 1GV/m gradient in a small gap (mm), has been developed and is in the early stages of being tested [55]. A very short electron bunch is produced in this instance, from a photo cathode and accelerated to 2MeV by a DC field. It is then immediately accelerated further by RF cavities. This method has the potential to produce high peak current electron bunches, but not at a very high repetition rate, so average current is low.

This thesis will largely concentrate on the continuous field DC photocathode guns that are in operation at JLab and Cornell. JLab currently operates a 350kV DC gun at the FEL facility. Additional bunching and accelerating stages are required to achieve the required parameters from the injector. The JLab/AES upgrade injector will use a gun that is geometrically identical to that in the FEL, but has the potential to be operated at higher gradients (to produce 500keV electron bunches). The Cornell university ERL demonstration injector has a similar design to that used at JLab, but aims for a higher operating voltage of 750kV.

Solenoids are used to compensate for space charge forces within the electron bunch, and buncher and accelerating cavities are used to shorten the bunch length and increase the energy of the electrons respectively. These injectors will be discussed further in chapters 5 and 8.

3.2.2 RF Acceleration

The electrons exiting DC guns that are currently in use in accelerators all have an energy below 1MeV, and the injector usually includes a further acceleration stage of a series of radio frequency cavities. RF cavities are shaped metallic structures that can support either travelling or standing wave electromagnetic fields within. The physical dimensions of the resonant structure determine the fundamental frequency of the RF field. More commonly for continuous wave electron acceleration, standing wave cavities are used. One of the most simple resonant structures is the pillbox cavity, which is a closed cylinder. Maxwell's equations are used to describe the permissable standing wave patterns, or modes inside the cavity. There are two types of mode, transverse electric (TE) or transverse magnetic (TM). For the TE modes, the on axis longitudinal electric field is zero, and so is not suitable for acceleration. The mode pattern is described by three subscripts $\mathrm{TM}_{m,n,v}$, where m is the number of periods around the circumference, n is the number of radial zeros in the field, and v is the number of half period variations along the electron axis, z. The lowest TM mode is $\mathrm{TM}_{0,1,0}$, and has all electric field lines parallel to the electron axis. This is shown schematically in figure 3.3. The $\mathrm{TM}_{0,n,v}$ modes all have zero magnetic field on axis. If the length of the cavity is chosen such that the electron is inside the cavity for the duration of the accelerating half of a RF period, maximum acceleration is achieved. Several cavities can be placed in succession, and if the on axis electric field has 180° difference in each one (π mode), then the electron bunch will see only accelerating fields as it passes through each cavity.

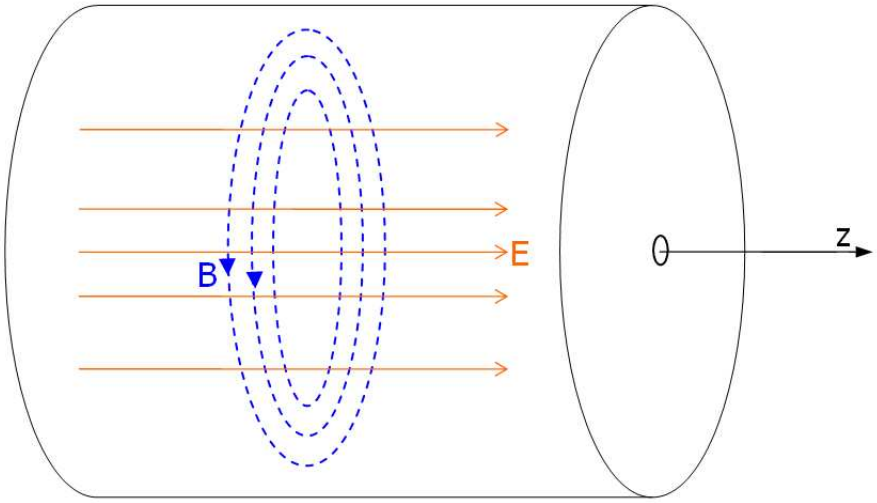

Figure 3.3: Pillbox cavity with an electromagnetic field in the $TM_{0,1,0}$ mode

The principle of an RF gun, is simply to place a cathode inside an RF cavity. When the electric field is accelerating, electrons can be extracted and accelerated away from the cathode. Early RF guns used a thermionic cathode placed inside a cavity, and the electrons would be suppressed when the field was decelerating, and extracted when accelerating. The first thermionic RF gun was built at Stanford High Energy Physics Laboratory in 1983 and was used for an infra-red FEL [56]. The Duke University FEL still operates using a thermionic RF gun, despite the progress made with photocathodes, due to its robust nature [57]. The pulse length of a thermionic RF gun is limited by the frequency of the RF. For low GHz frequencies, the pulse length is of the order of 100ps long, and the temporal shape is determined by a combination of the RF field and charge density saturation. The advantage of photocathodes is that the longitudinal distribution and length can be shaped by the incident laser pulse.

The first demonstration of a photocathode RF gun was made shortly after the thermionic RF gun in 1988, again at Stanford [58]. Traditionally RF guns were made from copper and operated in a macro-pulse mode, where the RF would be applied for a few μs at a time, and the electron bunches in the macro-pulse

were dependent on the fundamental frequency. The ratio of time with RF applied to no RF would give the duty factor, and this was usually one or two percent. When the RF is applied the vacuum inside the cavity degrades. The electric fields inside the cavity can be high enough for field emission from some surfaces. Field emitted electrons can then impact with the walls of the cavity, heating it and dislodging impurity ions or secondary electrons. The resistance of the cavity itself results in heating, which in turn causes desorption of impurities. Prolonged conditioning can improve the vacuum, but the macro-pulse mode permits high electric fields to be attained momentarily, whilst a modest vacuum is maintained. Since the first RF photocathode gun, this technology has subsequently provided the brightest electron bunches and has been used for numerous low repetition rate accelerators (100's of Hz), such as linear colliders and SASE FELs. RF guns that can operate with a continuous wave are the ultimate goal, and higher rep-rate (MHz) and duty factor RF guns are being developed to remove the heat loading through many cooling channels [59, 60, 61]. To overcome the problems of resistive heating in cavities degrading the vacuum, superconducting RF (SRF) cavities were developed. This was achieved first by Stanford University [62]. The RF can be supported in these structures continuously, and the peak on-axis gradients can be higher compared to CW normal conducting cavities. Of course the added complexity of the cryogenic system required to keep the cavities superconducting, adds to the cost significantly. Despite this, SRF cavites are now routinely used for acceleration of particles to high energy. Only recently has this technology been applied to an electron gun. Projects at FZ Rossendorf [63, 64], Brookhaven National Laboratory [65] and JLab [50, 51] are investigating the feasibility of SRF photocathode guns. Difficulties arise with trying to place a photocathode inside the cavity. Using the superconducting niobium or a deposition of lead on the cavity as a photocathode has the associated problem of the laser heating up the cavity and possibly causing the cavity to quench. Inserting a cathode on a plug into the back wall

of the cavity, has the disadvantage of the surface no longer being smooth, and there are spaces for the RF field to leak out of the cavity. A final drawback is that all electron guns require a focusing solenoid field to counter-act the space charge forces in the electron bunches. With DC and normal conducting RF guns, it is simple to place a solenoid close to the cathode to act on the bunches almost immediately. In the case of SRF cavities, all magnetic fields are expelled in order to achieve superconductivity, so the solenoid must be placed some distance away from the cavity. The distance from the cathode to the entrance of the accelerator is considerably longer than that of a normal conducting RF gun, as the surrounding cryogenic container is bulky. Therefore, the emittance compensation is not ideal. It has recently been suggested that a $TE_{0,2,1}$ mode can be excited in a cavity that would provide an on-axis magnetic field to apply some focusing to the electron bunches whilst still in the gun cavities [66], but this technology is still unproven.

3.3 Conclusions

Both DC and RF guns each have their own distinct advantages. Where a continuous stream of electron bunches is required, only DC guns can provide this at this time, as operational RF guns are limited to repetition rates below 1kHz due to heating and poor vacuum conditions. Conversely, the emittance and brightness from DC injectors does not match the capabilities of RF guns, which routinely operate with higher bunch charge. The goal of 4th generation light sources is to achieve up to 100mA average current and $< 1\mu$m emittance. With careful management of the power loading inside RF guns, it looks feasible to create an operational CW RF gun in the future. Superconducting RF guns are the least developed technology, but may ultimately offer the highest quality electron beams.

The electron gun is not the only issue with a 100mA injector. The subsequent acceleration in booster cavities must also be considered. New designs will be required for these also if they are to accelerate 100mA beams. It is clear that whichever technology is chosen, further technical developments will be required to meet the requirements of next generation machines.

JLab have chosen to modify the design of the FEL DC gun for a 100mA injector as the technology is more developed locally. Improvements to the vacuum and the photocathode systems have been made, and a power supply procured that can provide 100mA at 500kV. The electron beam quality is better from photocathodes than from thermionic cathodes, but further testing is required to see what the lifetime of high quantum efficiency cathodes will be.

CHAPTER 4

Modelling a High Brightness Electron Injector

There are many codes available for modelling the beam dynamics of electron injectors where the space charge forces of electron bunches are dominant. Choosing the best code for a given application depends on several factors. The computation time, accuracy, personal preference, whether 2D or 3D modelling is required, and what accelerator components are to be included, all have to be considered. Code comparisons as well as benchmarking exercises have been conducted for those most commonly used [67, 68]. A brief summary of some of these codes is given in table 4.1.

Code	**Type**
PARMELA	2D/3D
ASTRA	2D/3D
HOMDYN	2D
GPT	2D/3D
TREDI	3D
IMPACT-T	3D

Table 4.1: Summary of Injector Simulation Tools

The codes listed above are the most popular codes in use at present for injector modelling. The 2D codes are only really applicable to modelling situations when the electron beam and the fields acting upon it are cylindrically symmetrical. HOMDYN [69] uses multi-envelope equations to predict the motion of

sections of the electron bunch. It assumes that the electron bunch can be represented by a uniformly charged cylinder of varying length and radius. HOMDYN is ideal for conducting quick parameter studies with limited accuracy, as computation time is very short. PARMELA 2D [70] and ASTRA 2D [71] are similar, in that they both Lorentz transform the electron bunches into the average rest frame of the bunch, and evaluate the effects of static forces on ring shaped cells containing macro particles. The assumption with these codes is that the cells have a uniform charge density. Simulation time for the same injector is longer using PARMELA than ASTRA. IMPACT-T [72] and GPT [73] are relatively new 3D codes that can incorporate more features and so are useful for simulations beyond injectors alone. These 3D space charge algorithms can take a long time to run because they require a large number of macro-particles to adequately simulate the beam dynamics. Fully consistent codes, which can include and calculate the electro-magnetic components directly, such as MAFIA [74] and VORPOL [75] are extremely resource intensive.

Traditionally at JLab the injectors and accelerators have been modelled using PARMELA. The open source code policy of PARMELA has occasionally led to errors being introduced through modification to some versions, resulting in non-realistic simulations. Advantages of using PARMELA are that it interfaces well to electro-magnetic field solvers POISSON and SUPERFISH, that were developed alongside PARMELA, plus bending magnets can be included in the simulation.

At the time of writing, the original PARMELA source code and licence was no longer available. This was the primary reason for choosing ASTRA as the simulation tool for the 100mA FEL injector. ASTRA is a very user friendly code, as component field maps are simple to implement, and input parameters bear direct relationships to parameters that are optimised experimentally. The drawbacks are that bending magnets cannot be included, so modelling a circular

accelerator is extremely difficult, and the closed-source code is frequently updated without an increment in the version number.

4.1 ASTRA Basics

The program ASTRA (A Space-charge TRacking Algorithm) is designed to track macro-particles through user defined external fields whilst including the effects of the space charge forces on the particle cloud [71].

ASTRA is both a 2D and 3D code, so it is possible for the user to select the space charge algorithm. The 2D algorithm is used to model cylindrically symmetric particle bunches, and can be used to simulate the emission of particles from a cathode, including the effects of mirror charges and Schottky enhancement if desired. The 3D algorithm can simulate beam dynamics for non-symmetric bunches or those with a large transverse aspect ratio. Unfortunately, it is not possible to use the 3D algorithm for emission from a cathode; a disadvantage, as bunch asymmetries will be most dominant in low energy regions near the cathode.

One-dimensional field maps can be introduced, or for complex (asymmetric) components, it is possible to define 3D field maps. For simple, cylindrical elements such as RF cavities and solenoids, basic on-axis field maps can be used for a good approximation and reduced computation time. The current version of ASTRA (dated 14th March 2008) includes cavities (both travelling and standing wave), solenoids and quadrupoles. Bending elements are not included as the average bunch orbit would deviate from the longitudinal axis.

4.1.1 Particle Distributions

The initial particle distribution can be user-defined or can be created using the program *generator*, which comes as part of the ASTRA suite. Here the proper-

ties of the electron bunch emitted, from a cathode in this case, can be defined in terms of distributions (Gaussian, uniform etc) and rms values and truncation parameters. It is possible to combine multiple distributions to try and accurately represent the real-world particle bunches emitted from cathodes. The number of macro-particles that will be used to represent the electron bunch in the injector is also defined here. The trade-off between computation time and accuracy must be made. For preliminary simulations using a smaller number of macro-particles, 1000 for example, will demonstrate the general behaviour of each modelled component. Once the general, useful parameter space has been defined, it is then possible to increase the number of macro-particles representing the electron bunch, to increase the accuracy. Figure 4.1(a) shows how the emittance estimate of a simulation converges with an increasing number of macro particles. The emittance is calculated 1.2m from the cathode of the Cornell DC gun with a 370G solenoid placed at the gun exit. The time for the simulation also increases, as shown in 4.1(b). The particle distribution file contains

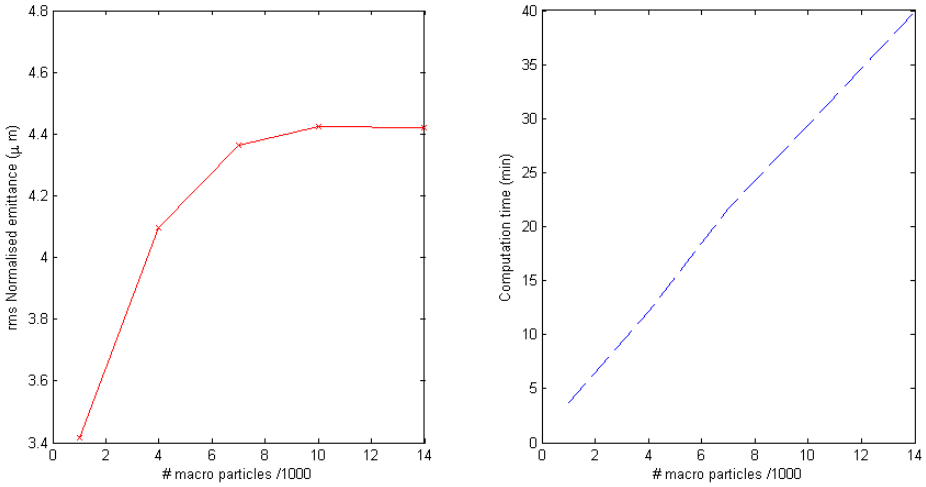

Figure 4.1: Convergence of emittance calculation and computation time using 2D space charge calculation

a value of the 6 canonical variables, (x, y, z, p_x, p_y, p_z), for each macro-particle. Each variable can be defined by a distribution described by a probability function, e.g. Gaussian, uniform, or radial. The function is defined using rms values

and the cut-off. See figure 4.2 for example transverse distributions. These distributions can also be applied to the longitudinal and transverse momenta. In this way the thermal emittance of a cathode can be included into the model.

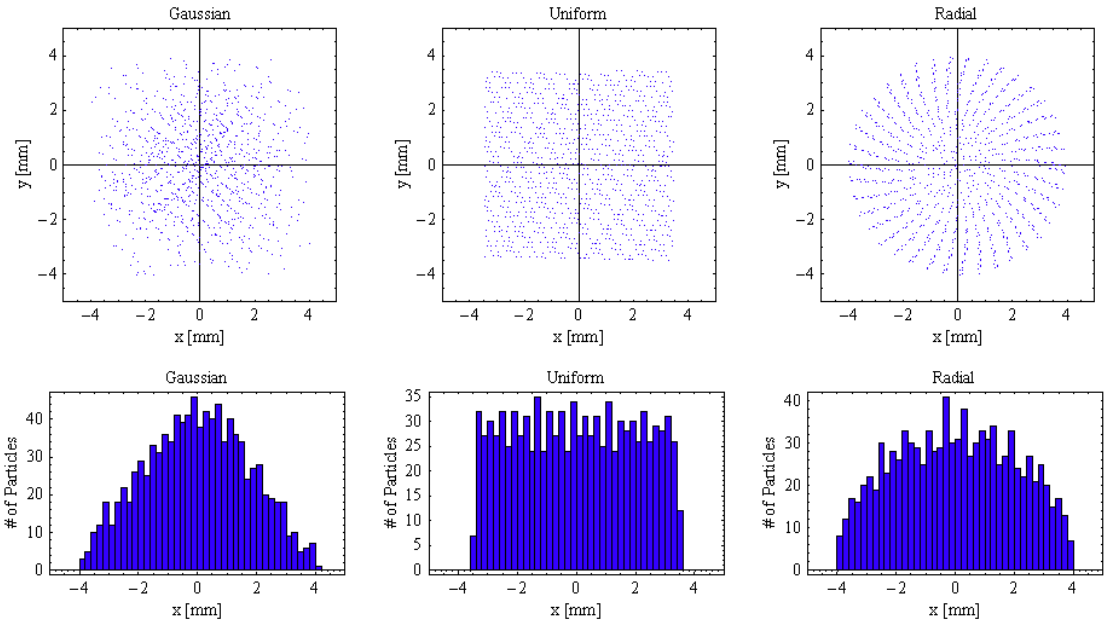

Figure 4.2: Examples of particle distributions $\sigma_x = \sigma_y = 2$mm (3σ cut off for Gaussian)

4.1.2 Field Maps

For rotationally symmetric fields a simple on-axis field map is used. For standing wave RF and static fields, a table of longitudinal position and longitudinal field (z, E_z) is read into ASTRA. The radial electric and magnetic field components are deduced from the first to third derivatives of the on axis field. The maximum field amplitude and phase of the wave are defined by the user. Prior to particle tracking, the energy gain of a single reference particle is calculated for each cavity as a function of the phase. The user defined phase is given relative to the maximum energy gain phase, with a cosine function. e.g. a phase $\phi = 0$, would result in maximum energy gain, and $\phi = -90°$ would give no net energy gain, but longitudinal bunching (as the tail of the bunch sees an accelerating field and the head sees a decelerating field). This definition is used throughout

this thesis. It should be noted that for low energy particles the phase of maximum energy gain does not correspond to the phase of the RF when the field is at a maximum in the cavity. This is because the electron bunch does not move at the speed of light, so it effectively slips backwards with respect to the RF phase.

Solenoid fields are implemented in a similar way, with a table of z, B_z. For this element only the maximum field amplitude is required and the field is scaled. Again, off axis radial and longitudinal components of the solenoid field are derived from the derivatives of the on axis field.

For the simulations of the injectors in this thesis only 1D field maps have been used. All the cavities simulated in this thesis have standing wave fields. The on-

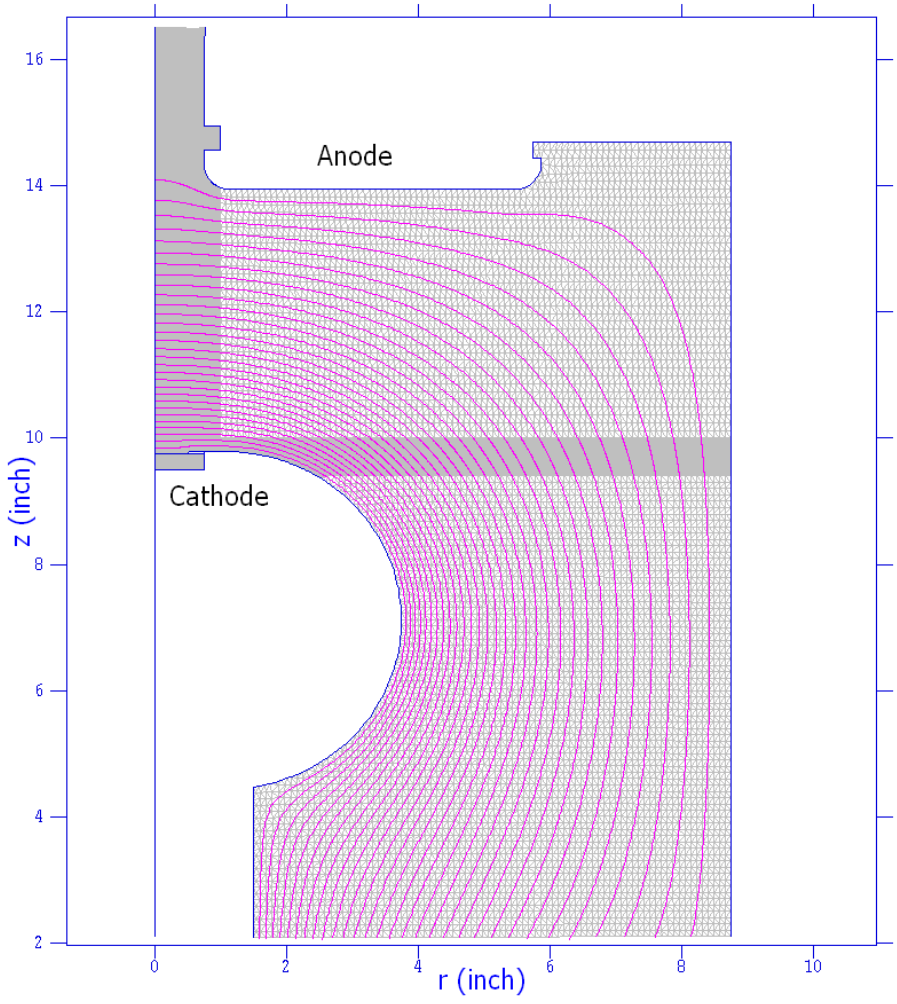

Figure 4.3: Jlab Gun geometry modelled in SUPERFISH. The design is axially symmetric around r= 0". Electric field contours are shown in pink

axis fields for each component has been modelled using SUPERFISH or POISSON [76]. As an example, the electrostatic field profile in the JLab FEL DC gun was generated from SUPERFISH. The geometry (white area), mesh (grey) and contours of equal gradient (pink) are shown in figure 4.3.

4.1.3 Space Charge Calculation

The calculation of the space charge fields is dependent on the aspect ratio of the electron bunch. Earlier versions of ASTRA only considered the 2D case where the particle bunch is cylindrically symmetric. In this scenario the space charge forces are computed by creating a cylindrical grid around the particle bunch. Radial rings and longitudinal slices are used to grid the extent of the particle bunch, see figure 4.4. An automatic procedure scales the size of each ring so

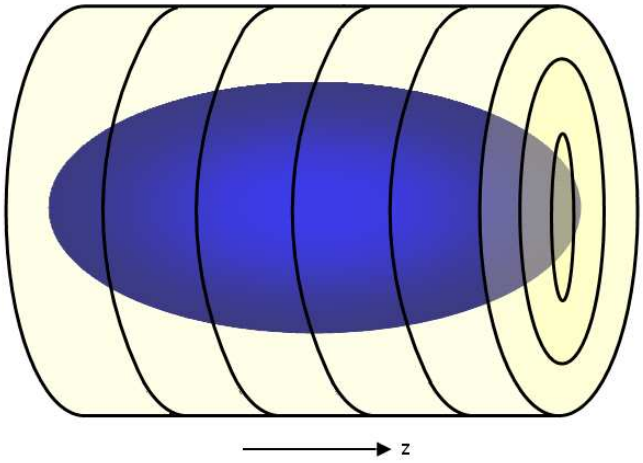

Figure 4.4: Cylindrical Grid Used for 2D Space Charge Calculation

that they always enclose the entire bunch as the dimensions change. The inner ring radius can be made larger so that the proportion of particles contained within remain comparable to the number in the outer rings. If the longitudinal profile has a long tail for example, the longitudinal cells around the tail may have a small number of particles, which decreases the accuracy (as statistical measurements are less meaningful). To improve the statistical properties it is possible to merge longitudinal cells together. The accuracy of this method is

thus lower for bunches where the transverse dimensions are disproportionate (eg. flat beams).

The grid is then Lorentz transformed into the average rest frame of the bunch, where, to a good approximation, the particle motion is non-relativistic. A static field calculation is made by integrating over the rings and finding the field contribution at the centre point of each cell. The field is then summed, and then transformed back into the laboratory frame. The field at any point from the ring centre points, is then calculated by interpolation. The space charge field is treated like an external field for tracking. The tracking of particles is based on a Runge-Kutta integration of fourth order with a fixed time step, which sums the internal and external forces. ASTRA does not include any effects from beampipe walls. It is assumed that the beam is in an infinite vacuum.

To simulate the emission of particles from a cathode, the particles are introduced to the tracking according to a timing spread of the initial distribution. The space charge field is scaled for the increase in charge as more particles are emitted. During emission of particles from a cathode, a mirror charge is included in the space charge calculation. The mirror fields are calculated in the rest frame of the mirror bunch at the Lorentz transformed distance between it and the real bunch. Once the contribution to space charge effects is less than 1%, the mirror charge is ignored.

The later versions of ASTRA include a 3D space charge algorithm that can handle non-symmetric beam distributions [77]. This algorithm calculates the space charge field on the intersections of a 3D cartesian mesh, as shown in figure 4.5. The grid density is user-defined and must be fine to accurately describe local variations. In addition, to avoid statistical problems, a large number of macro particles are required to reduce the number of empty cells. The space charge calculation takes place as an electrostatic computation in the rest frame of the bunch. The charge distribution inside a grid cell is represented by 3 constant

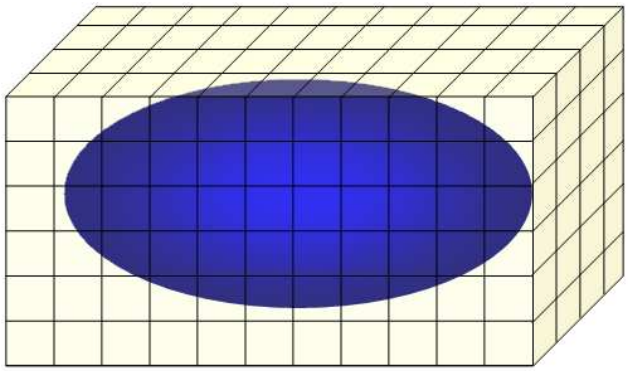

Figure 4.5: Cartesian Grid Used for 3D Space Charge Calculation

line charges along the axes of the cell. The potential of the space charge field is derived from a convolution of the charge density of the grid and analytic Greens functions. The rest frame electric field components are calculated from the derivative of the potential, and transformed back into the laboratory frame. As mentioned previously, emission from a cathode can not be calculated using the 3D space charge algorithm, and image charge forces cannot be included. During the 2D calculation of the JLab DC gun, the mirror charges are still included at a distance of 13cm from the cathode, which is almost at the gun anode positioned at 15cm.

4.1.4 Emittance Calculation

As discussed in chapter 2 the emittance of an electron injector is a critical figure of merit. It is therefore important to understand which emittance is being calculated by particle tracking codes. As output, ASTRA can be configured to calculate either the normalised phase space or trace space emittance as described in sections 2.1.3 and 2.1.5. In the case where the bunch has a small energy spread and low beam divergence in a drift space, these two emittances are almost the same [24].

Due to ASTRA tracking being based on time steps, the default emittance is calculated at a particular time step with individual particles having differing lon-

gitudinal coordinates. i.e. the x, p_x, x' etc coordinates at a snap-shot in time are used.

Most emittance diagnostics, see chapter 6 for examples, measure the beam properties at a given position, recording the transverse particle positions as they pass through at different times. ASTRA contains two further options to calculate emittance. Firstly, the emittance calculation can be such that the particle information is recorded as it passes through a specific longitudinal position, so that it is not calculated until the entire bunch has passed. Alternatively it can be calculated at the time when the average bunch particle passes the measurement point. All other particles are projected to that longitudinal position taking into account their position and momentum.

4.2 ASTRA in This Thesis

For the majority of simulation scenarios throughout this thesis, the electron bunch to be modelled is assumed to be cylindrically symmetric. The electron bunches are almost always launched from the centre of a cathode inside a DC gun with an axially symmetric distribution, and for these reasons the 2D space algorithm is utilised. For additional accuracy, all the field maps that are inserted use off-axis expansion expressions to 3rd order, rather than the nominal 1st order. This should improve the simulation for those particles that significantly deviate from the central axis (more than a few mm).

The results for transverse emittance reported in this thesis, are all projected trace space emittances calculated at a longitudinal position z. For speed of simulation the particles are projected onto the transverse plane when the reference particle reaches position z. The reference particle begins at the centre of the electron bunch at the cathode. If there is any deviation from this emittance estimate, it will be explicitly stated.

CHAPTER 5

100mA Cornell University Injector

Cornell University have an operational demonstration of a DC electron injector as described in the following sections. The components and electron beam dynamics are similar to those being developed at JLab. The measurements and simulation that were performed using the Cornell injector are therefore from a regime that is expected at JLab.

5.1 The Cornell University Accelerator Programme

The Cornell Electron-positron Storage Ring (CESR) is an electron-positron collider that has been operating since 1979 when it collided its first beams [78]. It consists of a 768 meter storage ring in which the particles circulate and collide with each other at an interaction point. The synchrotron used to accelerate the particles to a typical energy of 5GeV is situated inside the storage ring, in the same tunnel, as shown in figure 5.1.

The synchrotron was actually constructed first, being completed in 1968. It was not until the 1970's that the proposal was made to use the synchrotron as an injector to a colliding beam machine. The synchrotron radiation that is generated as a by-product of particle transport in the storage ring has always been used for X-ray research in addition to the collider experiments. The future of CESR

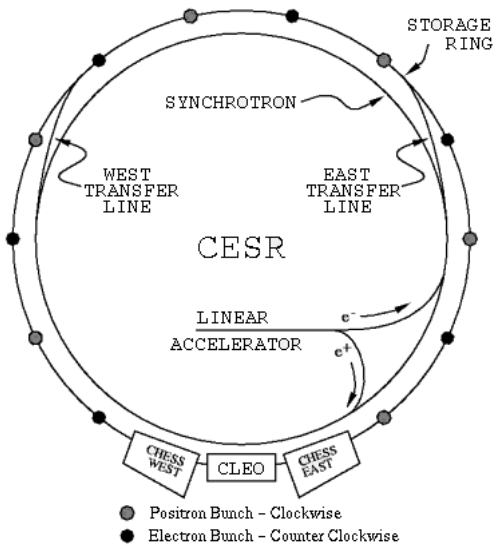

Figure 5.1: CESR schematic

lies in this research area, as Cornell University plan to phase out the high energy physics programme over the coming years. A proposal was made jointly between Cornell University and Jefferson Laboratory to develop an innovative ERL based X-ray source on the CESR site. The planned ultra-bright X-ray light source will exceed the spectral brightness of third generation synchrotron facilities by two to three orders of magnitude and also generate ultra-short X-ray pulses with flexible distributions that can promote new X-ray science [79]. To achieve this, a multi-GeV machine is required that can operate CW with extremely low emittance (2μm), pushing the limits of current state-of-the-art technology.

In February 2005, Cornell University was awarded $18 million to begin development of the ERL project [80]. The proposed layout of the ERL is shown schematically in figure 5.2, depicting the CESR tunnel and layout of a possible linear ERL extension.

Electrons are to be emitted from an injector that has been optimised for very low emittance and short electron pulses (location 1). These would subsequently be accelerated in the first linac (2) to half their maximum energy. The electrons

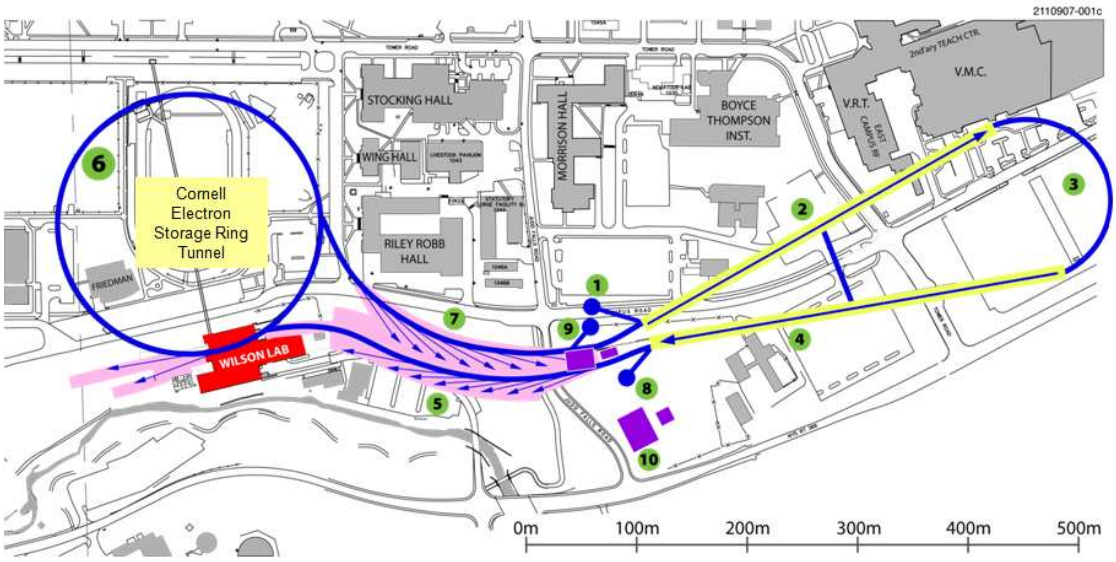

Figure 5.2: Layout of the upgrade of the existing ring accelerator to an ERL

would then follow around a return loop (3) and into a second linac, located in the same straight tunnel as the first, for further acceleration (4). An arc (5) is to be used to merge the electrons into the CESR ring (6), where they will travel clockwise until another arc (7) injects them back into the first linac, 180° out of phase, where they will be decelerated to half their energy. The return loop leads the electrons again to the second linac section where they are once more decelerated back to their low injection energy, at which they are finally dumped (8).

5.2 The ERL Prototype

In order to facilitate the development of the technologies required to build the ERL and address the challenges involved, a prototype ERL will be used [81]. A 100MeV, 100mA CW ERL 'Phase 1' machine is currently under construction, beginning with the electron injector. The prototype, figure 5.3, is very similar in appearance to the JLab IR-DEMO and will use similar technology. The parameter list for the prototype is given in table 5.1.

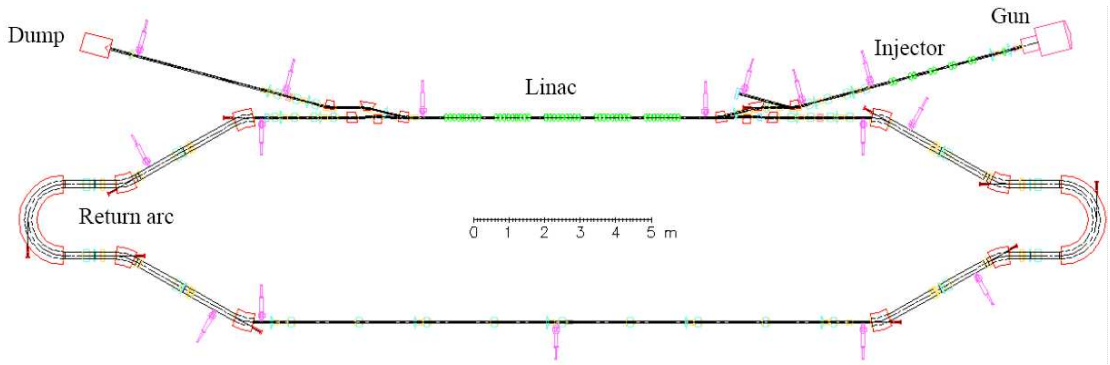

Figure 5.3: Phase I ERL prototype at Cornell University

Parameter	Value
Energy (MeV)	100
Transverse norm. rms emittance (μm)	2 at 77pC
Energy spread rms (%)	0.02 - 0.6
Average current (mA)	100
Bunch charge (pC)	1 - 400
Bunch length rms (ps)	0.1 - 2
Repetition rate (GHz)	1.3

Table 5.1: Cornell ERL Phase 1 parameter list

Construction of the prototype began with the electron gun and a diagnostic beamline, used to measure all properties of the electron bunches emitted. The beamline has been changed many times to accommodate different measurements by altering the position of, or adding new components. A schematic of the beamline used for transverse beam tomography measurements (described in chapter 6) is shown in figure 5.4.

The photocathode, which is located inside a DC gun (not shown), is situated on the right of the schematic. Directly following this are some small corrector magnets (not shown), that are used to centre the electron beam into the middle of the first solenoid. This solenoid is used for space charge compensation [22] and focuses the beam through the smallest aperture in the beamline; that of the light box. The light box contains two mirrors that are used to direct the laser beam onto the cathode and to remove any reflected light. There are two further

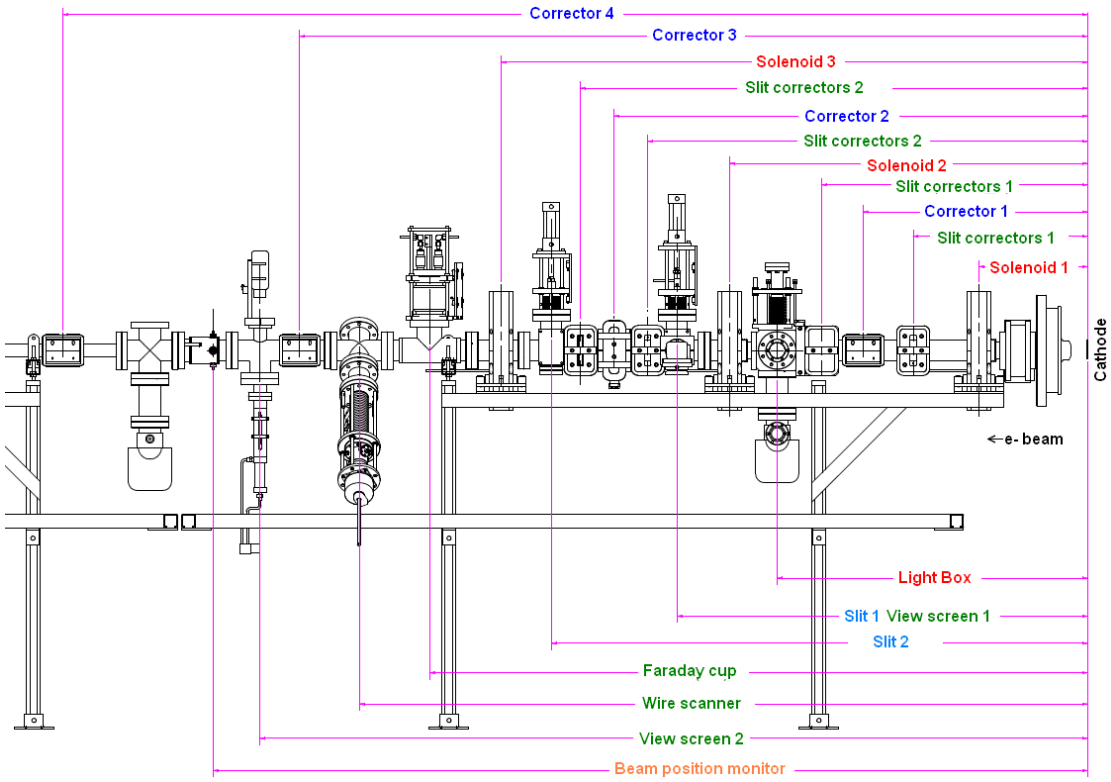

Figure 5.4: Diagnostic beamline layout

solenoids used to thread the electron bunches down the beamline onto diagnostics, or into the beam dump. A selection of corrector magnets are used for horizontal and vertical steering of the beam. The complete diagnostic suite consists of 4 view screens, a wire scanner, a Faraday cup, a beam position monitor and 2 horizontally collimating slits. The view screens are used to directly image the electron beam at various distances from the cathode. The wire scanner has three wires on a fork that pass through the electron beam giving information of the transverse profile in both the horizontal, x, vertical, y and x/y directions. The Faraday cup gives a direct reading of the current in the electron beam, and combined with a reading of the laser power, can be used to determine the quantum efficiency of the photocathode. Finally, the 2 slits are used to measure the vertical phase-space of the electron beam (at the location of the first slit), and thus give a direct measure of vertical emittance.

The beamline pictured is primarily for transverse beam measurements. Not shown is the deflecting cavity which is used for longitudinal measurements [82]. This is a 1.3GHz copper cavity, figure 5.5, excited in the TM_{110} dipole mode. When an electron bunch is injected into such a cavity at the zero-crossing of the

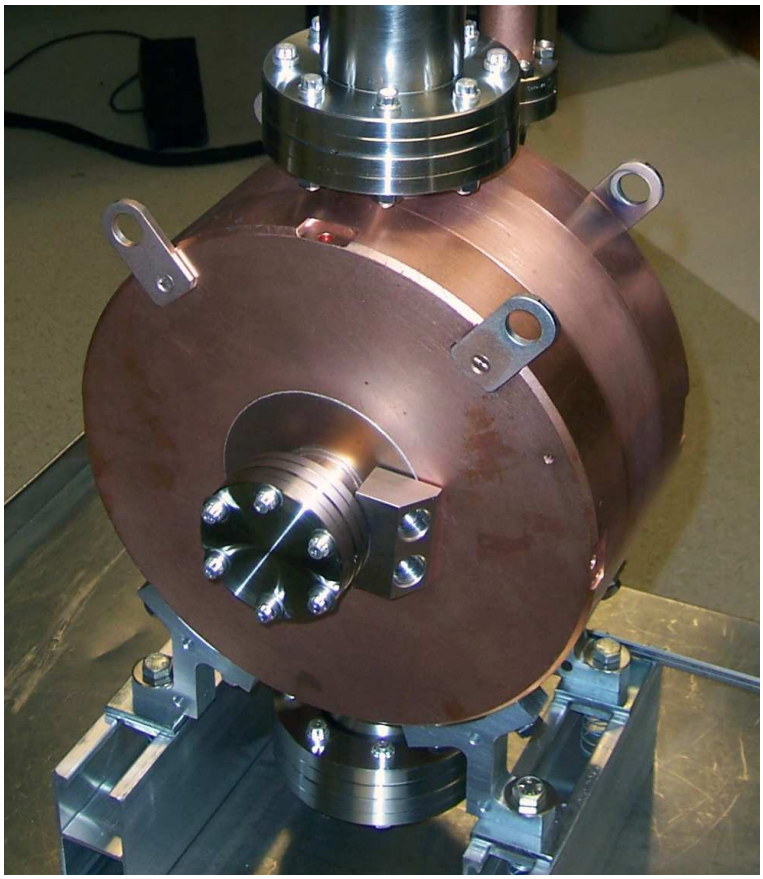

Figure 5.5: 1.3GHz deflecting cavity

RF, it experiences a transverse momentum kick that corresponds to the longitudinal position along the bunch. The strength of the kick is approximately linear with the arrival time of the particles. As a result, the longitudinal distribution of the beam is deflected into the transverse plane and can be imaged on a profile monitor located downstream. This diagnostic is useful for imaging the charge density along the bunch, and the longitudinal profile, a method typically used to measure the response time of the photocathodes at various wavelengths. An example of a longitudinal beam profile is shown in figure 5.6 with arbitrary units. The head of the electron bunch is shown at the bottom of 5.6 (a) and the tail at the top. The profile from this GaAs cathode is shown in 5.6 (b). The

ripples are produced from the laser pulse shaping technique that expands the laser pulse to 30ps (FWHM). The distribution at the head of the electron bunch is different from that at the tail. There is no sharp cut-off at the tail and this is the result of the response time of GaAs, discussed in section 5.6.

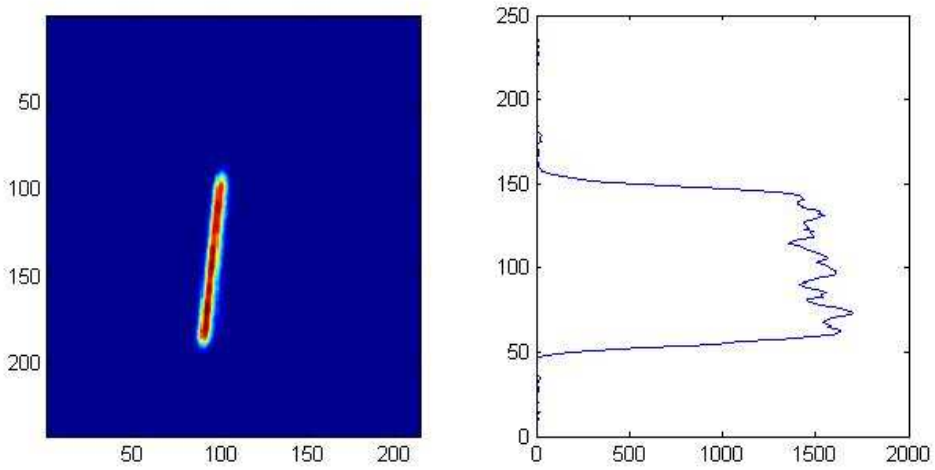

Figure 5.6: Example of a temporal profile measurement of a GaAs cathode with 520nm laser, in arbitrary units

5.3 The Electron Gun

In order to achieve the transverse emittance of 2μm required in the undulators of the full scale ERL, it will be necessary to generate a 1.5μm emittance from the injector, and even less than that from the electron gun. Similar to JLab, Cornell have opted to use a DC electron gun as their source, with a caesiated GaAs cathode. However, that is where the similarity ends. Table 5.2 lists notable differences between the two JLab and Cornell schemes, all of which have been constructed.

The Cornell gun is designed to reach higher voltages than both of the JLab guns. It also employs a Pierce electrode to introduce some transverse focusing to the electron beam as it is emitted from the cathode [83], shown in figure 5.7. The cathode needs to deliver an order of magnitude increase in average current over

Parameter	JLab FEL Gun	JLab Upgrade Gun	Cornell Gun
Voltage (kV)	350	500	750
Average Current (mA)	10	100	100
Ceramic orientation	Horizontal	Horizontal	Vertical
Max gradient on electrode *(MV/m)	7.8	7.8	6.8
Load-lock	No	Semi	Yes

*for 500kV

Table 5.2: Comparison of DC electron guns

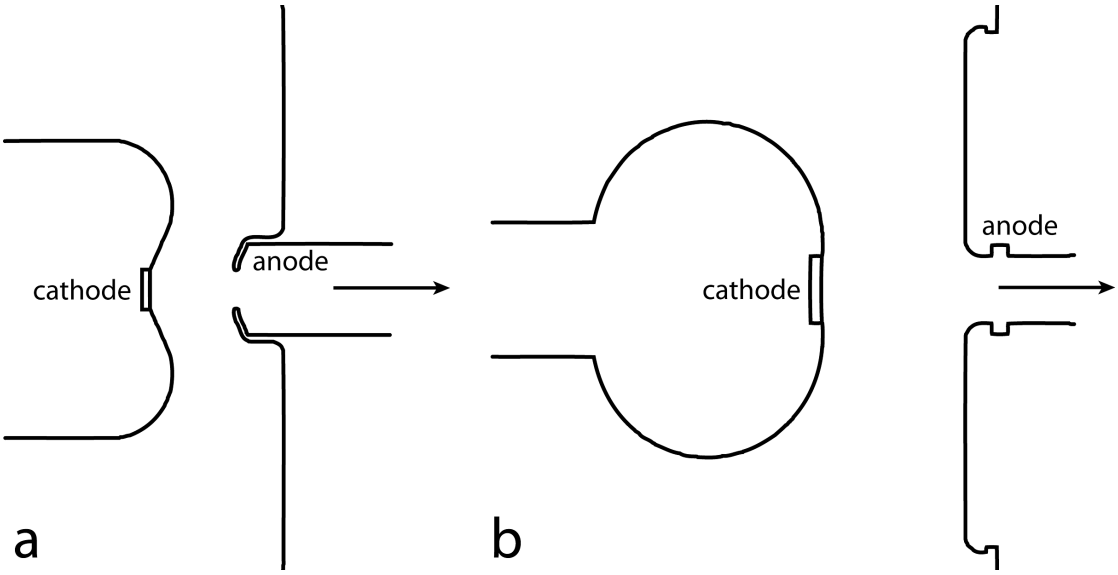

Figure 5.7: Comparison of the Cornell (a) and JLab (b) gun geometries

the JLab FEL gun. Consequently, the cathode lifetime will be reduced and therefore regular replacements or re-caesiations must occur. For this reason, the gun has been designed with a load-lock cathode system. This ensures the vacuum is not broken when introducing a new cathode. It follows that the fewer times that the gun chamber is vented, the less time is spent conditioning the electrodes to higher voltages. Another convenience of the load-lock is the cathode can be completely withdrawn from the gun chamber into a cathode preparation system located behind the gun. Here the cathode can be heat or hydrogen cleaned and re-caesiated without the possibility of contaminating the gun itself. The preparation chamber can hold spare cathodes that can be interchanged whilst old cathodes are processed. To accommodate the load-lock system, the ceramic

and electrodes are positioned vertically in contrast to the JLab gun, see figures 5.8 and 5.9.

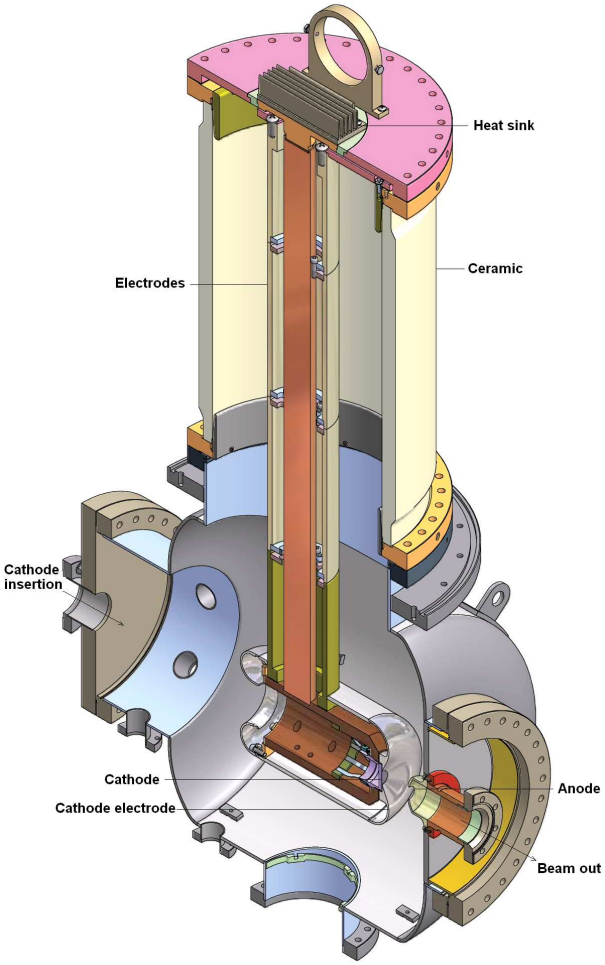

Figure 5.8: Cross-section of the Gun

Figure 5.9: Image of the gun as constructed

To illustrate the effect of the Pierce geometry in the Cornell gun, figures 5.10 (a) and (b) show the longitudinal and radial fields respectively. The fields are each calculated from POISSON models of the cathode-anode geometry with a voltage of -500kV on the cathode. It can be seen from the longitudinal field maps that the cathode-anode distance is much shorter for the Cornell gun. This has the benefit of accelerating the electron bunch in a shorter distance, so the effects of space charge beam blow up are reduced. The radial field component is shown as a function of distance from the cathode, and is calculated at a radius of 2mm from the cathode centre to show how it would affect the edges of

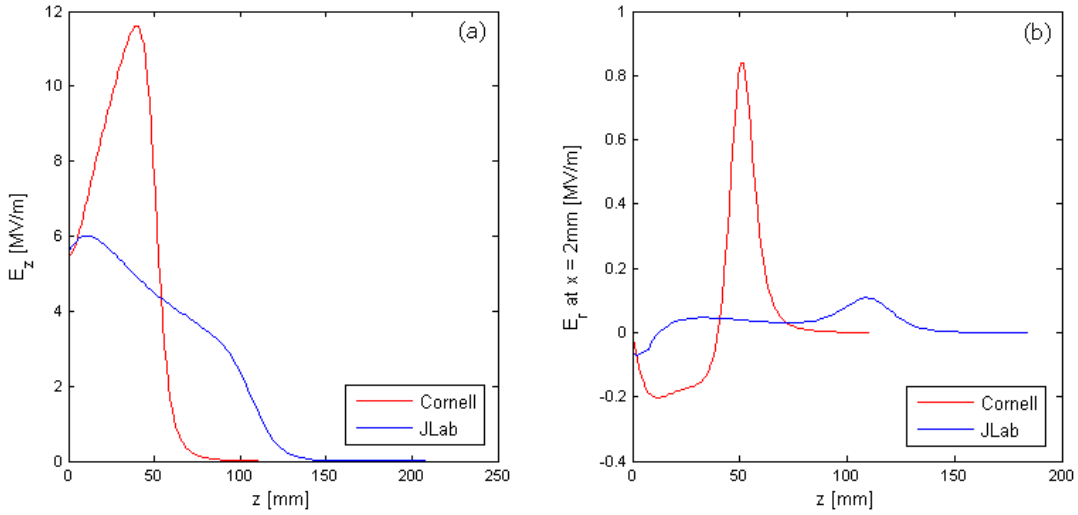

Figure 5.10: Comparison of the longitudinal (a) and radial (b) electric field on axis in the Cornell and JLab guns

a 4mm diameter electron beam emitted from the cathode. The Pierce electrode will demonstrate greater transverse focusing of the electron beam at the cathode, than from the flat JLab geometry. This will decrease the emittance induced in the gun region, which in turn will benefit the emittance downstream in the machine. As mentioned in chapter 2, the ultimate emittance is limited by the thermal emittance of the cathode. This, added to any beam induced emittance growth, is what culminates at the undulators, and as such is an important factor that must be understood.

5.4 The Photocathode

Cornell, like many other labs, is using a gallium arsenide (GaAs) semiconductor photocathode. This choice of cathode has propagated from one laboratory to another, and is therefore mostly used with DC guns for historical reasons. Originally GaAs was chosen for its properties of high quantum efficiency, low thermal emittance, and also because, when illuminated with polarised light, it emits polarised electrons. The design of the Cornell electron gun is a next generation design of that at JLab and for that reason uses the same cathode material,

even though polarised electrons are not required. As an alternative, a gallium arsenide phosphide (GaAsP) cathode will be prepared and tested to investigate any advantages in thermal emittance and response time that it may possess. The thermal emittance or energy of electrons is dependent on the chemical composition of the cathode, and so different materials or surface preparations produce various energies. As discussed in chapter 3, negative electron affinity (NEA) photocathodes are known to have a naturally low thermal emittance because of the emission process. A prime motivator for the diagnostic beamline on the ERL prototype was to perform a detailed study of these emission properties to characterise the cathode performance before it is operated in a user facility.

5.5 Thermal Energy Measurement

The thermal energy of GaAs has been previously measured at JLab on the 100kV CEBAF injector using laser wavelengths between 514 and 840nm [46]. The thermal energy was measured using a wire-scanner technique and found to be 34$\pm$2meV and 103$\pm$6meV at 840nm and 514nm respectively. Separate measurements have also more recently been conducted at Nagoya University, Japan for wavelengths between 700 and 950nm [84]. Here the emittance was measured using a pepper-pot method, at similar electron energies of 100-200keV. For bulk GaAs though, the thermal emittance measured was double that of the JLab experiment. A more complex definition of the equation for thermal emittance was used which includes band gap, affinity and photon energy information. Some difference can be associated with the surface treatments that each cathode was subjected to; preparing NEA cathodes is subjective. These measurements were repeated for the Cornell cathodes to determine the definitive behaviour for this case [85, 17].

The GaAs wafers used at Cornell had a Zn doping between 2×10^{18} and 2×10^{19}cm^{-3} and were heat cleaned and activated in the load-lock chamber behind

the gun. A "yo-yo" deposition [86], so called because of alternating Cs and NF_3 on the wafer surface, was used to give a NEA surface and gave typical quantum efficiencies of 10% measured at 532nm. The GaAsP photocathode surface was prepared in a similar way. The GaAsP was epitaxially grown onto a GaAs substrate to a thickness of 2μm, with a 2μm transition layer of graded phosphorus concentration. The phosphorous concentration in the surface layer was 45% with a p-doping level of $\sim 3 \times 10^{18}$cm^{-3}.

5.5.1 Measurement Theory

Recall from chapter 2 that if only the transverse, uncoupled motion of a particle is considered, the transfer matrix through non-dispersive elements is given by:

$$\begin{pmatrix} x \\ x' \end{pmatrix} = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \begin{pmatrix} x_0 \\ x'_0 \end{pmatrix} \tag{5.1}$$

and in terms of the beam matrix and transfer matrix elements:

$$\sigma_{11} = R_{11}^2 \sigma_{11,0} + 2R_{11}R_{12}\sigma_{12,0} + R_{12}^2 \sigma_{22,0} \tag{5.2}$$

where $\sigma_{11} = \langle x^2 \rangle$, $\sigma_{11,0} = \langle x_0^2 \rangle$, $\sigma_{12,0} = \sigma_{21,0} = \langle x_0 x'_0 \rangle$, and $\sigma_{22,0} = \langle x'^2_0 \rangle$. According to equation 5.2 three measurements of the beam size at the end of the beamline for different magnet settings would be sufficient to determine the initial sigma matrix, $\boldsymbol{\sigma_0}$, if the elements of the transfer matrix $\mathbf{R}$ are known:

$$\underbrace{\begin{pmatrix} \langle x_1^2 \rangle \\ \langle x_2^2 \rangle \\ \langle x_3^2 \rangle \end{pmatrix}}_{A} = \underbrace{\begin{pmatrix} R_{11,1}^2 & 2R_{11,1}R_{12,1} & R_{12,1}^2 \\ R_{11,2}^2 & 2R_{11,2}R_{12,2} & R_{12,2}^2 \\ R_{11,3}^2 & 2R_{11,3}R_{12,3} & R_{12,3}^2 \end{pmatrix}}_{M} \underbrace{\begin{pmatrix} \langle x_0^2 \rangle \\ \langle x_0 x'_0 \rangle \\ \langle x'^2_0 \rangle \end{pmatrix}}_{B} \tag{5.3}$$

$$\mathbf{B} = \mathbf{M}^{-1}\mathbf{A} \tag{5.4}$$

The effective rms (trace space) emittance is given by equation 2.15, repeated below:

$$\varepsilon_{x,rms} = \sqrt{< x_0^2 >< x_0'^2 > - < x_0\, x_0' >^2} \tag{5.5}$$

Although 3 measurements are mathematically sufficient to derive the emittance, it is preferable to take more measurements of the beam size at different transfer matrix settings in order to be less sensitive to errors. By measuring the beam size for n sets of transfer matrix settings and performing a least squares fit, the estimates for $\langle x_0^2 \rangle, \langle x_0 x_0' \rangle, \langle x_0'^2 \rangle$ are improved. The best estimates are obtained by minimising the Euclidian norm squared of the residual, $\mathbf{MB} - \mathbf{A}$:

$$||\mathbf{MB} - \mathbf{A}||^2 = ([\mathbf{MB}]_1 - \mathbf{A}_1)^2 + ([\mathbf{MB}]_2 - \mathbf{A}_2)^2 + ... + ([\mathbf{MB}]_n - \mathbf{A}_n)^2 \tag{5.6}$$

Using the fact that the squared norm of $\mathbf{V}$ is $\mathbf{V^T V}$ this can be rewritten as:

$$(\mathbf{MB} - \mathbf{A})^T (\mathbf{MB} - \mathbf{A}) = (\mathbf{MB})^T \mathbf{MB} - \mathbf{A}^T \mathbf{MB} - (\mathbf{MB})^T \mathbf{A} + \mathbf{A}^T \mathbf{A} \tag{5.7}$$

The two central terms on the right hand side are equal to one another and the minimum occurs when the derivative with respect to $\mathbf{B}$ equals zero.

$$\begin{aligned} \frac{d}{d\mathbf{B}} \left((\mathbf{MB})^T \mathbf{MB} - 2\,(\mathbf{MB})^T \mathbf{A} + \mathbf{A}^T \mathbf{A} \right) &= 0 \\ 2\mathbf{M}^T \mathbf{MB} - 2\mathbf{M}^T \mathbf{A} &= 0 \end{aligned} \tag{5.8}$$

$$\begin{aligned} \mathbf{M}^T \mathbf{MB} &= \mathbf{M}^T \mathbf{A} \\ \mathbf{B} &= \left(\mathbf{M}^T \mathbf{M} \right)^{-1} \mathbf{M}^T \mathbf{A} \end{aligned} \tag{5.9}$$

The vector $\mathbf{B}$ then contains the estimate of $\langle x_0^2 \rangle, \langle x_0 x_0' \rangle, \langle x_0'^2 \rangle$. These can be inserted into equation 5.5 to yield the transverse phase space emittance.

If the errors $\sigma_{\mathbf{A},n}$ in the measurement of the beam size vector $\mathbf{A}$ are to be included, both $\mathbf{A}$ and $\mathbf{M}$ are normalised by the error, as described in [87] to give

$\hat{\mathbf{A}} = \mathbf{A}/\sigma_{\mathbf{A},n}$ and $\hat{\mathbf{M}} = \mathbf{M}/\sigma_{\mathbf{A},n}$. The error in the emittance result is given by:

$$\sigma^2_{\varepsilon_x} = (\nabla_{\mathbf{B}}\varepsilon_x)^T(\hat{\mathbf{M}}^T\hat{\mathbf{M}})^{-1}\nabla_{\mathbf{B}}\varepsilon_x \tag{5.10}$$

where:

$$\nabla_{\mathbf{B}}\varepsilon_x = \begin{pmatrix} \langle x_0'^2\rangle/2\varepsilon_x \\ -\langle x_0 x_0'\rangle/\varepsilon_x \\ \langle x_0^2\rangle/2\varepsilon_x \end{pmatrix} \tag{5.11}$$

If there is no contribution to the beam size from space charge forces (or other effects such as wake fields), the measured emittance is entirely thermal. Using equation 5.12 and plotting thermal emittance, $\varepsilon_{th,n,rms}$, against the rms size of the emitting area, σ_0, the thermal energy can be obtained from the best fit slope to the data.

$$\varepsilon_{th,n,rms} = \sigma_0\sqrt{\frac{E_{kin}}{m_0c^2}} \tag{5.12}$$

5.5.2 *Experimental Set Up*

To perform studies of the photocathode thermal energy, the diagnostic beamline was specifically laid out for the purpose. The layout is shown in figure 5.11.

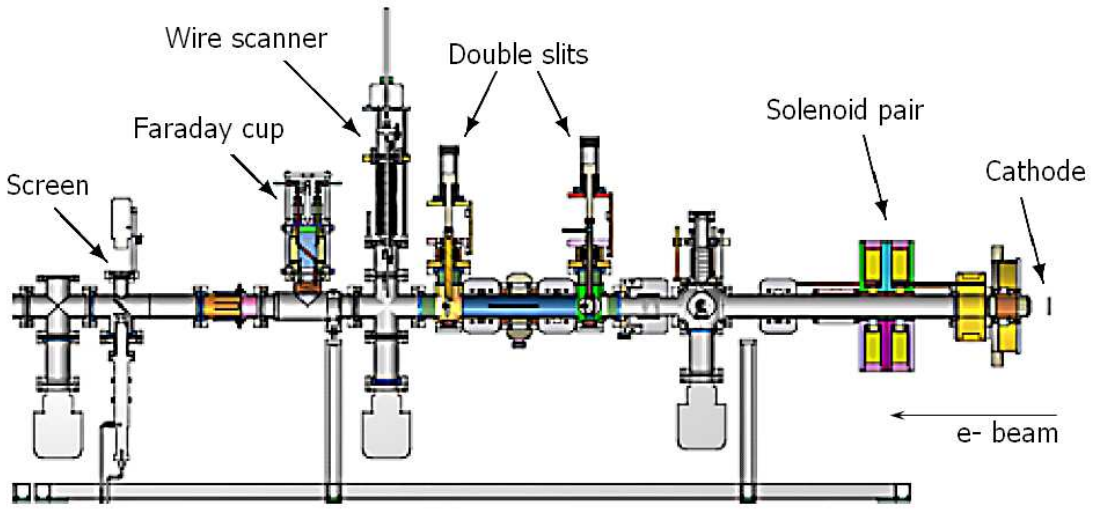

Figure 5.11: Diagnostic beamline layout for the thermal energy measurement

The solenoid pair directly after the gun consist of two identical solenoids back-to-back. These are counter-wound so that the total field integral is zero and the electron beam motion is uncoupled, thus the transfer matrix becomes simplified. Rather than analytically generating the solenoid matrix, it was calculated from the measured magnetic field profile shown in figure 5.12. The error between the current set in the solenoids and the actual current was measured to be less than 2%.

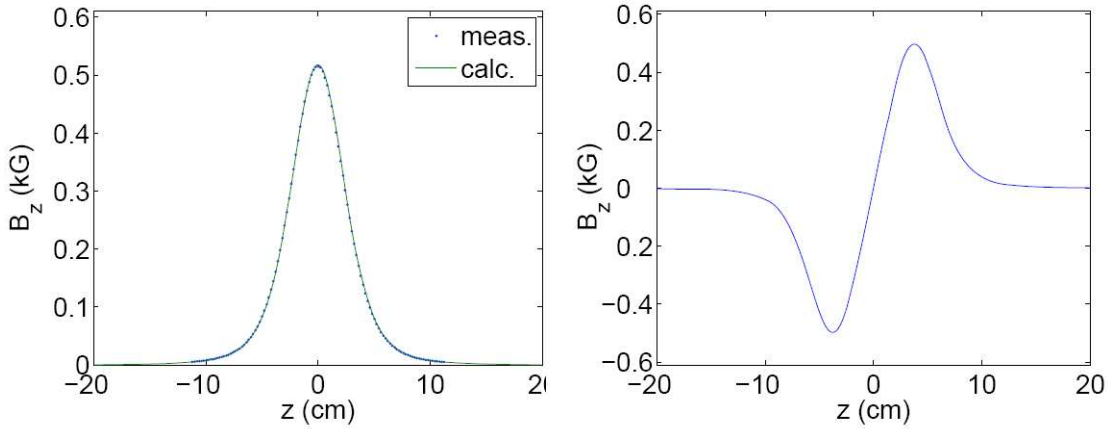

Figure 5.12: Calculated and measured magnetic field of a single (left) and two back-to-back (right) solenoids with 5A excitation current

The 4 x 4 matrix for a solenoid (without dispersion) that would transform x, x', y, y' coordinates from one point to another is given by:

$$\mathbf{R}_{sol} = \begin{pmatrix} C^2 & \frac{SC}{k} & SC & \frac{S^2}{k} \\ -kSC & C^2 & -kS^2 & SC \\ -SC & -\frac{S^2}{k} & C^2 & \frac{SC}{k} \\ kS^2 & -SC & -kSC & C^2 \end{pmatrix} \tag{5.13}$$

where $S = \sin(kl_{eff})$, $C = \cos(kl_{eff})$ and $k = \frac{B_0}{2B\rho}$. $B\rho$ is the magnetic rigidity, B_0 is the peak magnetic field, and l_{eff} is the effective length of the solenoid. The

matrix for a drift of length L is:

$$\mathbf{R}_{dri} = \begin{pmatrix} 1 & L & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & L \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.14}$$

The complete matrix for the beamline is the product of all the component matrices: $\mathbf{R} = \mathbf{R}_{dri,2}\mathbf{R}_{sol,2}\mathbf{R}_{dri,1}\mathbf{R}_{sol,1}$. When the solenoids are counter-wound this yields:

$$\begin{aligned} R_{11} &= R_{33} &&= C^2 - S^2 + L_1 L_2 k^2 S^2 - (L_1 + 2L_2)kSC \\ R_{12} &= R_{34} &&= (L_1 + L_2)C^2 - L_2 S^2 + 2SC/k - L_1 L_2 SCk \\ R_{21} &= R_{43} &&= L_1 k^2 S^2 - 2kSC \\ R_{22} &= R_{44} &&= C^2 - S^2 - L_1 kSC \end{aligned}$$

The two off-diagonal 2 x 2 matrices of the full 4 x 4 matrix are all zero, which shows that the motion between the two transverse planes is decoupled in this instance.

The measurements of the electron beam size were performed using the wire scanner, which is strung with 20μm wire and would give the beam profile in the horizontal, x, and the diagonal x/y direction as it moved through the electron beam. A third wire that would infer the beam profile in the y direction was damaged and therefore not used. The current measured from the wire was connected to an amplifier and sampled together with an encoder at 500Hz. Additionally, a BeO view screen was used with a 12-bit CCD camera to give an image of the transverse electron beam properties, from which the profile could also be determined.

All the beam profile data, taken from either the wire scanner or screen, was fitted with a super-Gaussian curve with free parameters A, B, n, s and x_0:

$$y = Ae^{-\frac{(x-x_0)^n}{s^n}} + B \tag{5.15}$$

The rms beam size is given by $s\sqrt{\Gamma\left(\frac{3}{n}\right)/\Gamma\left(\frac{1}{n}\right)}$, where Γ is the Gamma function. An example of the fit to the profile data is shown in figure 5.13. The error in the rms beam size estimate is calculated from the 68.5% confidence level of the fit, and at largest was 2%.

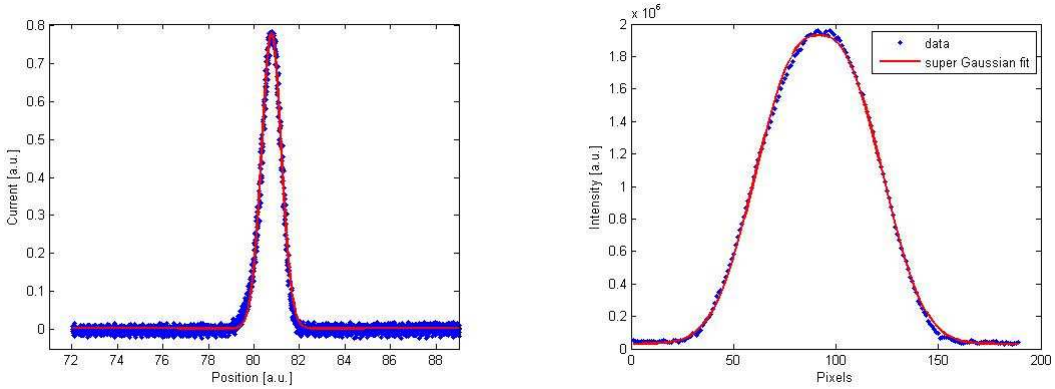

Figure 5.13: Fit to (left) wire scan and (right) screen profiles

To cover the wavelength range of 458 to 860nm four different laser systems were used; an argon-ion laser (458, 488, 514.5nm), a solid-state green (532nm), a helium-neon laser (633nm) and a titanium-sapphire laser (710-860nm). None of the lasers were operated in a pulsed mode. In a similar fashion to [46] a laser illuminated aperture was imaged 1:1 onto the centre of the photocathode with an achromatic 1m focal length lens. To be sure of the beam size on the cathode, for each measurement the laser was diverted after the lens onto a 10-bit SPIRICON camera, placed at the same distance as the photocathode. A typical set of laser spot images for the different apertures used is shown in figure 5.14. The measured laser profiles were found to be largely uniform, and the SPIRICON software that accompanied the camera was used to generate the rms spot sizes.

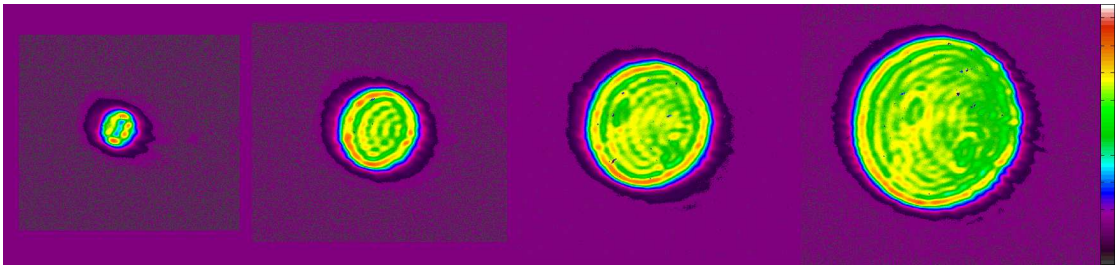

Figure 5.14: Laser beam profiles for apertures or 0.5mm, 1mm 1.5mm and 2mm (532nm)

To measure the thermal emittance of the cathode there must be no emittance contribution from space charge forces in the electron bunch. The power of the laser was therefore set to ensure the current was sufficiently low (typically less than 0.1μA).

5.5.3 Measurement Procedure

The following method was applied to measuring the thermal energy of both GaAs and GaAsP cathodes at several different laser wavelengths:

1. Measure the laser profile using the SPIRICON camera and software.
2. Measure the electron profile for a range of solenoid current settings using either the wire scanner or view screen.
3. Analyse the solenoid scan data and fit a curve, then estimate the thermal emittance.
4. Plot the thermal emittance versus laser spot size for different apertures.
5. Generate the line of best fit to the plot and calculate the thermal energy for that laser wavelength.

The gun voltage used in the majority of these experiments was 250kV, however a check at 200kV gave similar results within 1%. The typical quantum efficiency was 5%. An example of a solenoid scan and data fit is shown in figure 5.15. The errors in measuring the emittance sum to 4% and 5% for the wire scanner and screen techniques respectively.

From a set of solenoid scans using different apertures, a thermal emittance-laser spot size plot could be created. An example of this is shown in figure 5.16. The thermal energy is then estimated from the line of best fit to this plot.

Both methods of measuring the profile and inferring the emittance from the solenoid scan yielded similar results, as shown in figure 5.17. The thermal emittance from the wire scanner and screen measurements were calculated as 113 ± 8meV and 114 ± 8meV respectively for GaAs with 532nm wavelength. Also in the figure are two points from a further method of determining the emittance. This method, described in detail in section 6.1, utilises the slit and screen diagnostics to give a direct image of the phase space which can be used to estimate the emittance. The error in this method was estimated at 7%, and as the figure shows, lends some credence to the other measurement techniques as the agreement is very good.

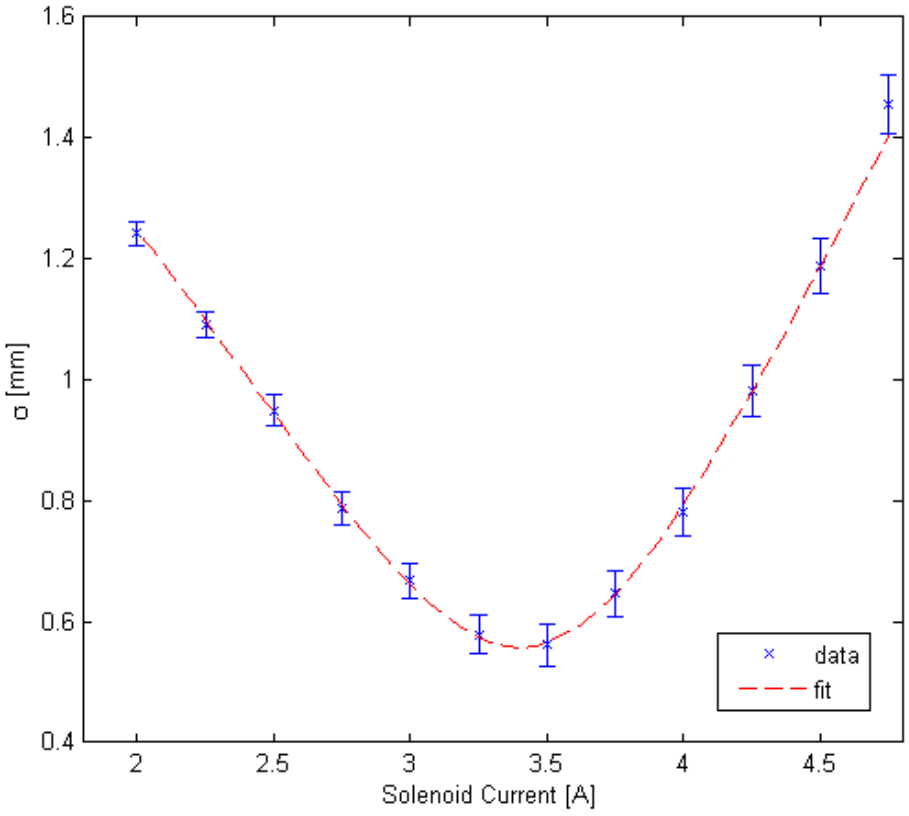

Figure 5.15: Solenoid scan and fit, (860nm, GaAs)

5.5.4 *Results*

The results of the thermal emittance measurements at various wavelengths for GaAs and GaAsP are shown in figures 5.18 and 5.19 respectively. For the GaAs cathodes the thermal energy measured was similar to that reported in [46]. Due to time limitations most of the measurements taken with the Ti:sapphire laser at wavelengths of 710 to 865nm were performed using one aperture (1.5mm) rather than the multiple aperture method, but nevertheless follow the trend. Near the band gap wavelength of ~870nm the thermal energy is close to room temperature as expected. The thermal energy at the wavelength of the ERL laser, 520nm is estimated to be 120±8meV for GaAs.

GaAsP is a semiconductor with two band gaps; direct and indirect. As the fraction of P to As is increased both the band gap energies increase, but the direct

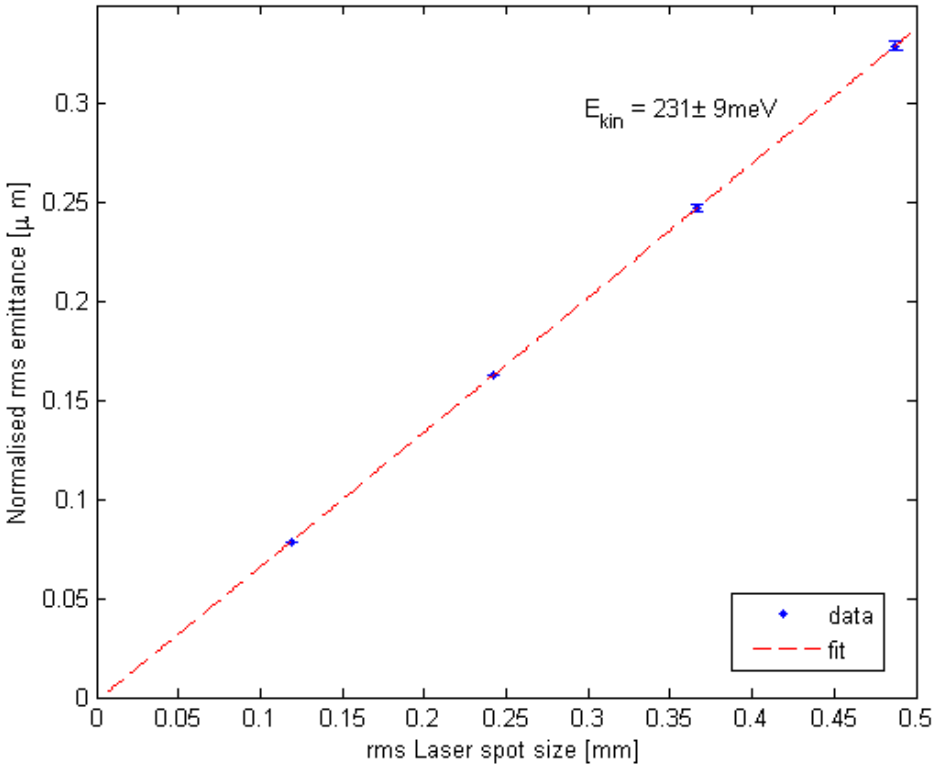

Figure 5.16: Thermal emittance of GaAsP as a function of laser spot size at 458nm

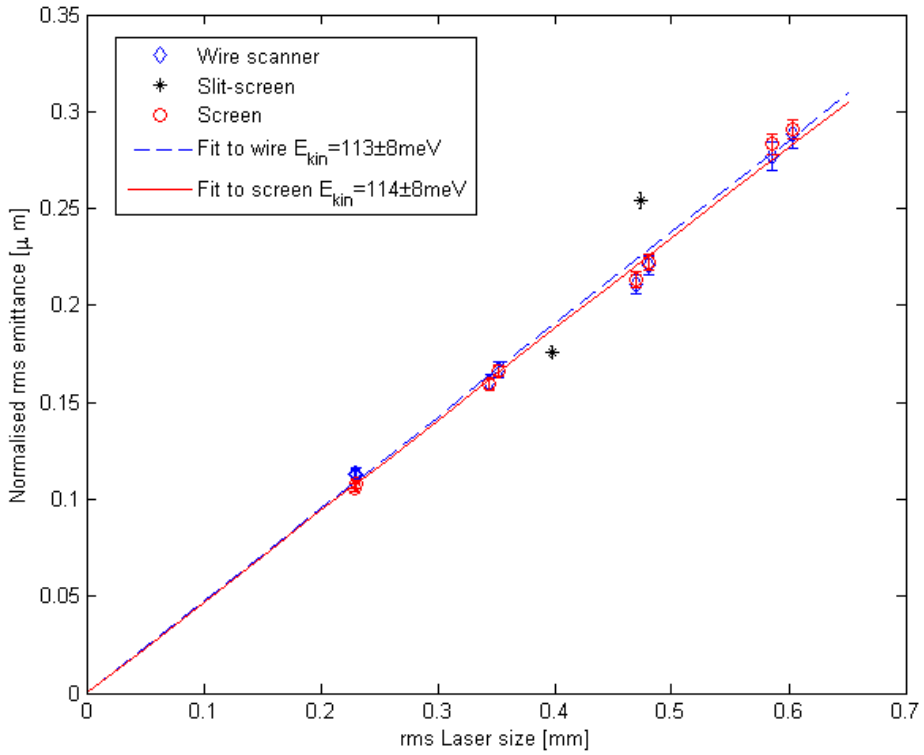

Figure 5.17: Comparison of various thermal emittance measurement techniques for GaAs at 532nm

band gap increases at a greater rate. At a mole concentration of 45% P, the band gaps cross and the semiconductor undergoes a transition from direct to indirect band gap. The wavelength of the band gap at the cross over is approximately 630nm, shorter than that for GaAs. Therefore, the thermal energy of the GaAsP cathode was expected to be lower than that of GaAs. The results of the GaAsP measurements do not support this assumption. The measurements were taken on two different occasions with the same cathode. The experiment was repeated because the first results were taken with a very low quantum efficiency ($< 1\%$). The second set was taken after the cathode had been re-caesiated. Both data sets reveal a higher thermal energy than GaAs at the same wavelength. At the present time the large thermal energy of GaAsP is not fully understood. It may indicate that a large fraction of the emitted electrons are coming from the direct, rather than indirect band gap minimum. Further investigation of GaAsP

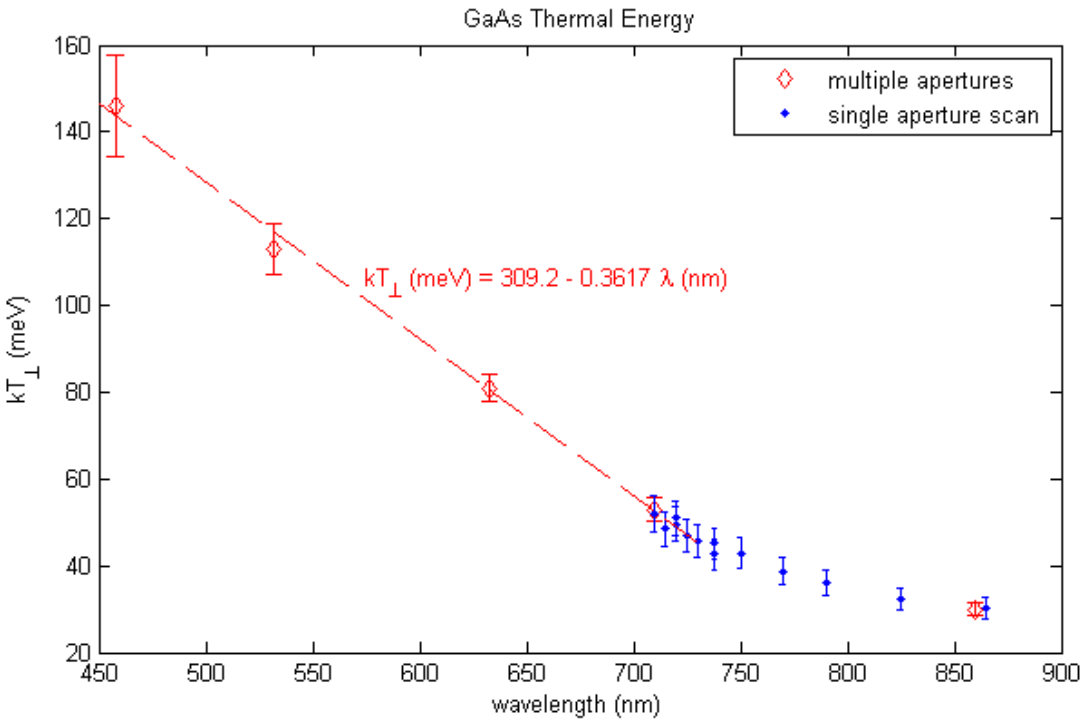

Figure 5.18: Measured thermal energy for GaAs as a function of wavelength

cathodes with differing phosphorus surface concentrations would be required to determine the emission process.

5.6 Response Time Measurements

When modelling injectors with simulation codes, the longitudinal profile of the electron bunches is often assumed to be the same as that of the laser. For GaAs based cathodes this is not the case in reality. GaAs is not a prompt emitter, as discussed in chapter 3, and the response time is dependent on the laser wavelength. Any long emission tail from the cathode can cause problems in an accelerator, producing halo electrons, longitudinal asymmetry and increasing the longitudinal emittance. Knowledge of the response time and distribution is important in mitigating these effects. For cathodes that exhibit tails, it may be possible to tailor the laser pulse shape to offset the response distribution and create the desired longitudinal shaping.

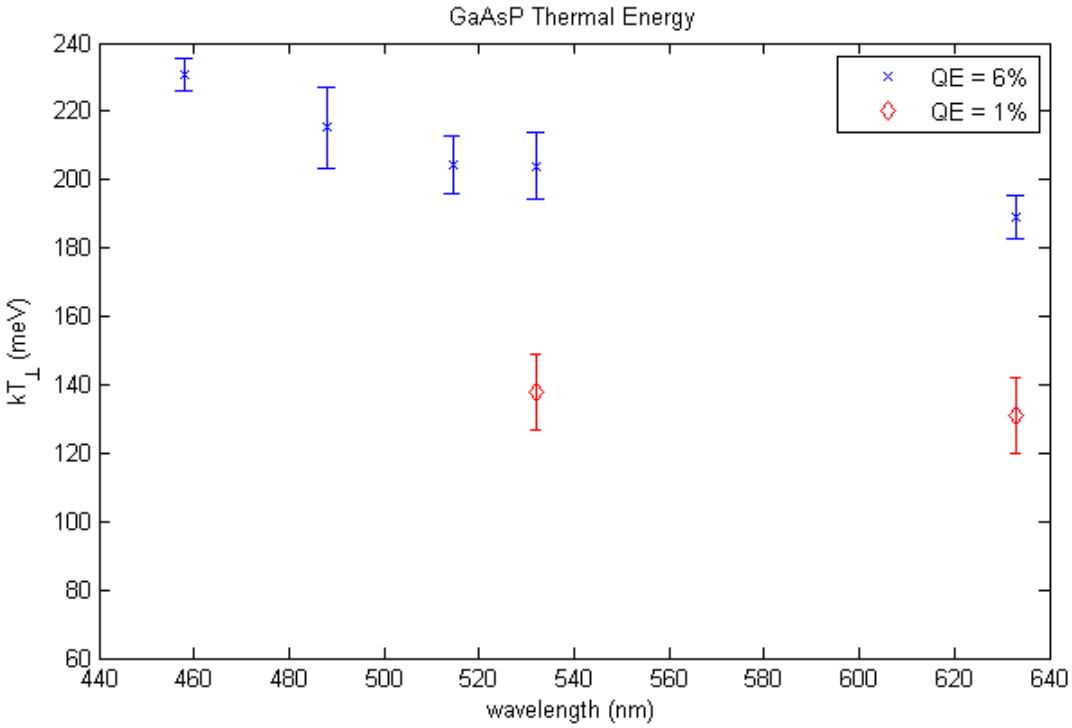

Figure 5.19: Measured thermal energy for GaAsP as a function of wavelength

The response of bulk GaAs as a function of time has a sharp leading edge followed by a long tail, as described in [47]. The reference describes how the three-step model was used to create a diffusion model that fits the measured photoemission response curves from GaAs. The theoretical model reproduced the measured temporal response profile well, and was used to form an analytical expression to fit the measurements made at Cornell [17]. The emission in this reference was measured with a wavelength of 840nm, which is near the band gap, and resulted in a long temporal tail of over 100ps. The absorption length is shorter at 520nm, and so the response time should also be shorter for the Cornell case.

From reference [47], the analytic solution for the photoemission current as a function of time is given by:

$$I(\kappa) \propto \frac{1}{\sqrt{\pi\kappa}} - \exp(\kappa)\mathrm{erfc}(\sqrt{\kappa}) \tag{5.16}$$

with normalised time $\kappa = t/\tau$, where $\tau \equiv \alpha^{-2}D^{-1}$ is the characteristic time; a function of optical absorption and electron diffusion. D is the electron diffusion constant, α is the optical absorption coefficient and the complementary error function $\text{erfc}(\eta) \equiv \frac{2}{\sqrt{\pi}} \int_{\eta}^{\infty} e^{-\zeta^2} d\zeta$. This is shown graphically in figure 5.20 (a). This function is singular at $t = 0$, and the actual temporal response is given

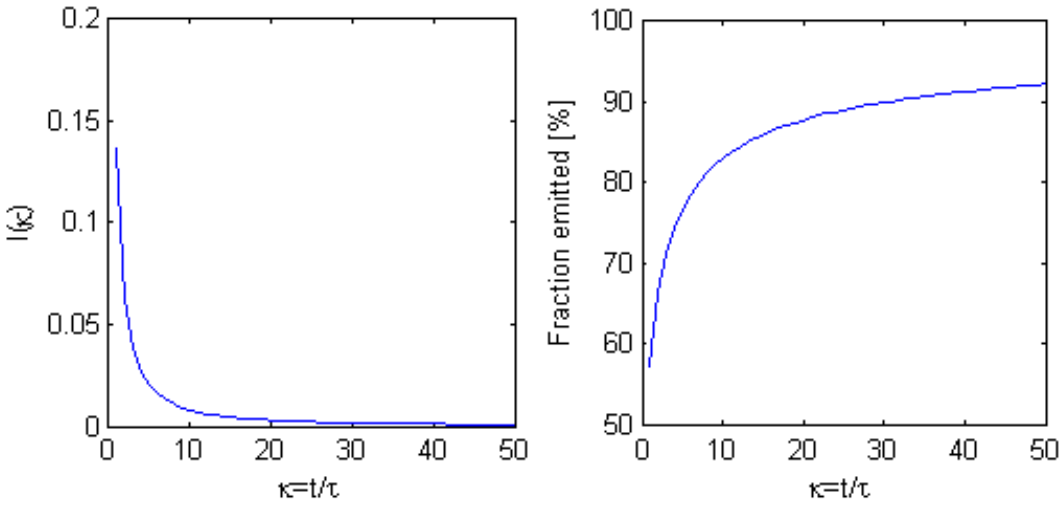

Figure 5.20: Response of GaAs to a delta light pulse (a), percentage of emitted electrons as a function of time (b)

by the convolution of the laser longitudinal pulse shape with the response to a delta light pulse given by equation 5.16. The response is characterised by an initial fast component followed by a relatively long tail. The fraction of the photoemission current $P(\kappa)$ emitted after time $t = \kappa\tau$ is calculated from the integral of equation 5.16, given by:

$$P(\kappa) = 1 - \exp(\kappa)\text{erfc}(\sqrt{\kappa}) \tag{5.17}$$

The characteristic time τ corresponds to the photoemission of 57% of the pulse, while 10τ gives 83%. This is shown in figure 5.20 (b). The shorter the characteristic time τ, the more prompt the emission.

5.6.1 Measurement Procedure

To give some measure of the rms response time of the photocathodes without the deflecting cavity, which was not initially available, an original method that utilised the transverse measurement beamline was implemented. The methodology was developed by I.V. Bazarov and reported in [17]. The author of this thesis contributed a minor role in the measurement of the response time. The analysis and results are important for the simulations described in later chapters and are therefore summarised and reproduced below.

A very short electron bunch with strong space charge forces was created using a sub ps laser pulse. This was used to infer the photoemission response time by fitting solenoid scan data with results of particle tracking simulation using the response time as the fit variable. The method for the experiment was to perform a solenoid scan as described in section 5.5 for a short electron bunch of known charge and initial transverse profile. A fit was then made to the solenoid scan through simulation. The transverse space charge force in the electron bunch is proportional to the peak current, which is a function of the bunch length (for a given charge); so by measuring the transverse beam size, the bunch length at the cathode could be inferred. The shape of the beam size versus solenoid field curve changes as a function of the response time parameter τ, as shown in figure 5.21. The simulation was made for 100fC bunches with a gun voltage of 250kV. The thermal energy was 150meV and the laser spot size was 120μm rms on the cathode. Without solenoid fields applied, the beam size increases with decreasing response times. This is on account of the peak current increasing due to the reduced emission time, and the transverse space charge forces becoming larger.

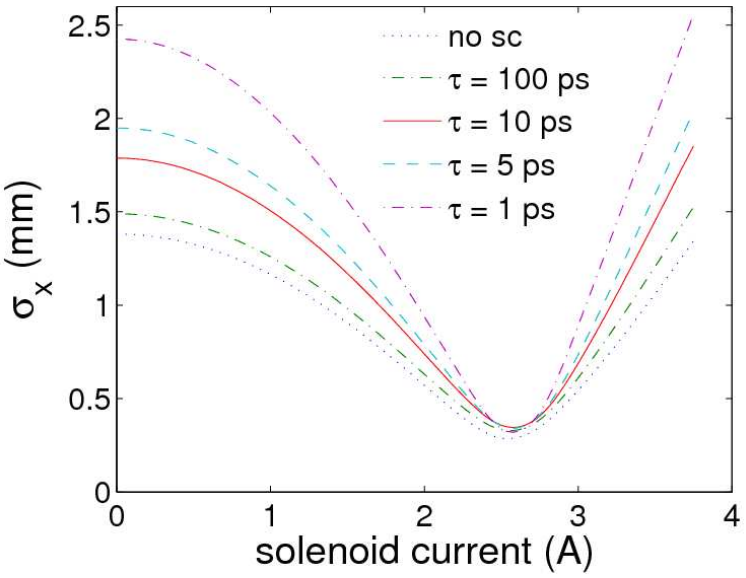

Figure 5.21: Simulated beam size as a function of solenoid current for different values of τ

5.6.2 Results

The results of the measured response time are reproduced in tables 5.3 and 5.4 for GaAs and GaAsP cathodes respectively. The response was measured at two different gun voltages of 200 and 250kV; the agreement to within experimental error between the two values support the validity of this method.

Wavelength (nm)	τ (ps)	Comment
860	76±26	V_{gun} = 200 kV
860	69±22	V_{gun} = 250 kV
785	11.5±1.2	V_{gun} = 200 kV
785	9.3±1.1	V_{gun} = 250 kV
710	5.8±0.5	V_{gun} = 200 kV
710	5.2±0.5	V_{gun} = 250 kV
520	<1	upper estimate placed
460	<0.14	upper estimate placed

Table 5.3: Results of data fitting for GaAs response time

Wavelength (nm)	τ (ps)	Comment
520	<1	upper estimate placed
460	<0.14	upper estimate placed

Table 5.4: Results of data fitting for GaAsP response time

For both GaAs and GaAsP at the shorter wavelengths of 520 and 460nm, the fitted response time could only be assigned an upper limit. This was because good fits to the data could be made by using the initial temporal distribution of the laser alone.

5.7 Conclusions

The thermal energy of electrons emitted from GaAs and GaAsP was measured using two complementary techniques to good agreement. As expected, the thermal energy increased with decreasing laser wavelength. At the Cornell ERL laser wavelength of 520nm, the thermal energy of GaAs was estimated to be 120$\pm$8meV. This is in good agreement with those published in [46], yet, there remains some discrepancy with those in [84]. This could be due to the equation chosen to describe thermal emittance. The result from the GaAsP cathode was unexpected as the thermal energy was higher than that from GaAs at all wavelengths and showed a strong dependence on quantum efficiency. It is possible that the structure of the GaAsP cathode is not as well understood and that there may be more electron scattering before emission; regardless, this requires further investigation.

The measured photoemission response from the two cathodes show a strong dependence on the laser wavelength. By using ASTRA it was possible to estimate the rms response time of each cathode. At the wavelength of the ERL laser, the response of the cathode follows that of the laser pulse. The upper limit on the characteristic response time is so short that for a laser pulse duration of longer than 7ps, the effect would be negligible (see fig 5.6). The Cornell ERL will use pulse lengths of $\sim$12ps (30ps FWHM), so for simulation purposes it is reasonable to use the longitudinal profile of the laser to model the response of electrons from the cathode.

The results of these experiments show that NEA GaAs based photocathodes have the potential to be used as high brightness electron sources. Through measuring the thermal emittance, a lower limit on the emittance performance of a photoinjector using a NEA cathode can be set as a function of laser wavelength. Both GaAs and GaAsP are found to be prompt emitters (sub ps) at 520nm. This information can be used to choose the optimum laser wavelength with which to operate these cathodes. This information will be used in later chapters to make simulation more realistic.

CHAPTER 6

Phase Space Tomography

The previous chapter examined a method of measuring the emittance, or area occupied by the particles in transverse phase space, from rms beam properties inferred from a solenoid scan. Notably, this approach does not provide any information on the structure of the distribution within phase space. Knowledge of this is of interest when tuning the accelerator and for high power lasing in FELs.

Tomographic reconstruction algorithms offer the possibility to reconstruct multidimensional density distributions from measurements of various projections of these distributions. Applying this technique to electron beams means that by measuring projections of the 2D transverse plane at different angles of phase space, the divergence of the beam can be inferred. Plots of transverse size versus divergence give the phase space of the beam.

The phase space can be measured directly by using a small slit and a screen diagnostic. The electron beam is moved relative to the slit, which permits a small amount of beam to pass through the slit and propagate to a screen, without any space charge effects contributing to the divergence. The measured divergence on the screen can be used to give the phase space at the position of the slit. This apparatus, however, can take up a lot of valuable space in a beamline. An

alternate method is to use the magnets already available in the beamline and tomography techniques to reconstruct the phase space. Using the Cornell diagnostic beamline it was possible to use phase space tomography and cross-check the results with alternate methods.

6.1 Measuring Phase Space

There are numerous methods of measuring the phase space directly. Those covered in this section are mainly used when the particle energy is low as they all assume some diagnostic which intercepts the beam and allows a small portion to propagate. With increasing energy, the diagnostics must become thicker to prevent unwanted particles penetrating through. As will become apparent in the following descriptions, the error introduced in the measurement increases with the thickness of the diagnostic. All of the following methods of measuring phase space are commonly used in accelerators and described in [21, 23, 87].

6.1.1 The Slit-Screen Method

The principle of this method is to sample the electron beam transversely with a very narrow slit that intercepts it, allowing only a small percentage of the total charge to pass through, forming a beamlet. The charge in a beamlet should be so small that the effect of space charge on it is negligible. This implies that the evolution of the beamlet is dependent entirely on the transverse momenta of the electrons passing through the holes. A screen downstream of the slit is used to image the beamlet. The divergence is then inferred from any offset of the beamlet from the centre of the slit, divided by the distance between the slit and screen. This is shown schematically in figure 6.1. This method has been used on numerous accelerators, at PITZ [88], FLASH [89], and Daresbury [90] for example.

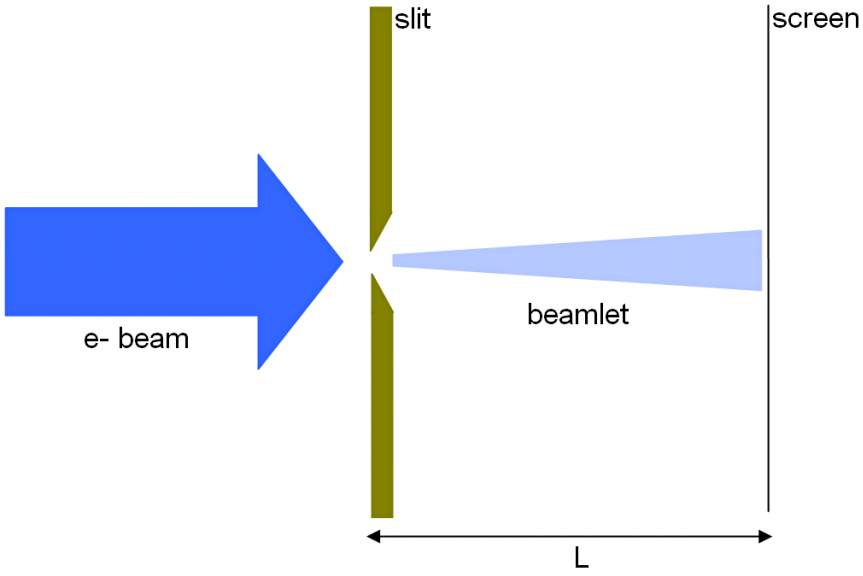

Figure 6.1: Slit-screen diagnostic

The image formed on the screen from the beamlet has a cigar-like shape, as shown in an example of a screen image from the Cornell diagnostic beamline, figure 6.2. The slit will only give information on the divergence of the portion of the beam passing through it. In order to obtain the entire phase space either the slit must move over the beam or vice versa. On the Cornell diagnostic beamline the latter approach was used. Two identical air-core corrector magnets were

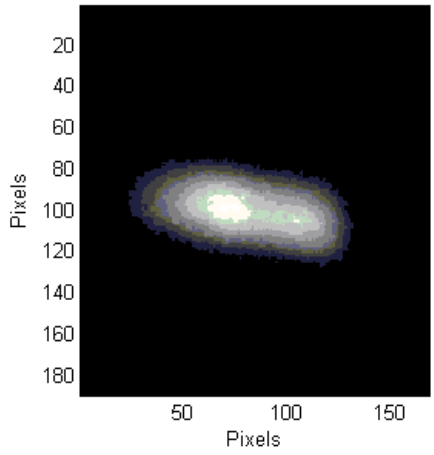

Figure 6.2: Example of a measured beamlet image

placed in front of the slit. The first corrector would impart a vertical kick to the electron beam. Some distance downstream the second corrector would cancel

this kick, and thus the electron beam could be moved up and down vertically. The slit was vertically collimating and was 25μm wide and 200μm thick. Some error is introduced into the measurement of the beamlet as it is assumed that the slit is infinitely small. If the distance between the slit and screen is large by comparison to the slit width, the error will be a small fraction of a percent. In the Cornell beamline, the distance between slit and screen is 2.89m and so the error from this is negligible. The ratio of slit width to depth determines the angular acceptance of the beam. The thinner the material, the greater the beam divergence that passes through the slit can be. Unfortunately the angular acceptance must be offset against the ability to stop those electrons not passing through the slit from travelling to the screen. Minimising the stopping thickness can be achieved by manufacturing the slit from a high density metal such as copper. The slit at Cornell was created by bringing two razor edged pieces of copper together, as shown in the schematic above and in figure 6.3. This keeps the angular acceptance large. There is also a large water cooled guard slit which is used to stop most of the electrons and prevent the diagnostic slot from heating enough to close the razor-slit gap.

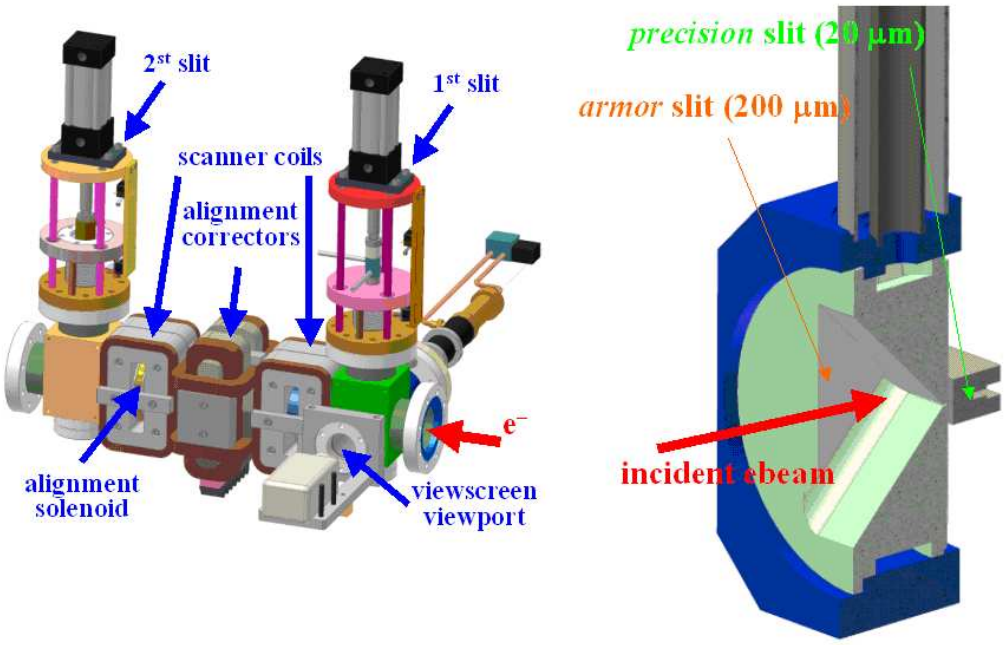

Figure 6.3: The Cornell slit diagnostic

The slit diagnostic can only measure the phase space in one transverse plane so an orthogonal slit apparatus would be required to measure the other plane. The Cornell beamline was only capable of measuring the vertical phase space using this method.

6.1.2 Phase Space Measurement

The phase space measurement was automated via the control interface EPICS (Experimental Physics and Industrial Control System) and MATLAB. One measurement would take less than a minute to complete. The corrector strengths were varied in small incremental steps to move the electron beam over the slit vertically. At each corrector setting an image was taken of the BeO view screen for analysis with MATLAB later. The correctors were carefully calibrated so that the vertical offset as a function of current in the coils was known to within 2%, and this value was recorded along with the associated beamlet.

An intensity profile in the vertical plane was then calculated for each beamlet image by summing the intensity of pixels in each row. This profile, when converted from pixels to mm and then divided by the slit-screen length, corresponds to the divergence information (in mrad) for that part of the beam sampled by the slit.

To correct for any alignment errors in the slit, that would lead to a rotated beamlet image on the screen, the average angle from the horizontal axis was calculated from all beamlet images. Figure 6.2 shows such a rotation that resulted from the slit not being parallel to the horizontal axis. The angle calculated from each image was weighted by the intensity in the beamlet, divided by the total intensity from all beamlets. The average of these angles was then used to rotate all the images before the reconstruction.

To eliminate any contribution to the emittance from background noise, a threshold could be applied to the image that would set pixel values below a user defined percentage of the maximum intensity to zero. If there are enough beamlet profiles of different parts of the electron bunch, it is possible to interpolate between points and generate a good representation of phase space with a contour or surface plot.

The resolution in the vertical plane, which now corresponds to the vertical divergence, y', is given by the calibration of the view screen from pixels to mm divided by the slit-screen distance. The resolution for the screen used is $\pm 15\mu$rad. In the horizontal plane, the error is calculated from the calibration of the corrector magnets and is estimated at $\pm 170\mu$m.

This experiment was performed twice to estimate the thermal emittance of the electron beam at two different laser spot sizes. These measurements were taken after the slit had been realigned, resulting in average beamlet angles of less than 2°. As described in section 5.5.2, the laser spot size was measured for each aperture before starting.

6.1.3 Phase Space Results

As with previous measurements of thermal emittance, the current in the electron beam was kept low enough during the experiment for space charge forces to be ignored. The following results from 1.5mm and 2mm laser spot sizes, shown below in figure 6.4, were used as an additional method to measure the thermal emittance of the electron beam with the 532nm laser. The contours represent lines of equal charge density increasing from blue to red at the centre. To remove the noise from the phase space, the beamlet profiles were subject to a threshold of 1.3% and 2.8% for the 1.5mm and 2mm case respectively, chosen to *just* eliminate background noise from stray laser light or X-rays. The emittance was calculated from the profile data rather than from the contour plot.

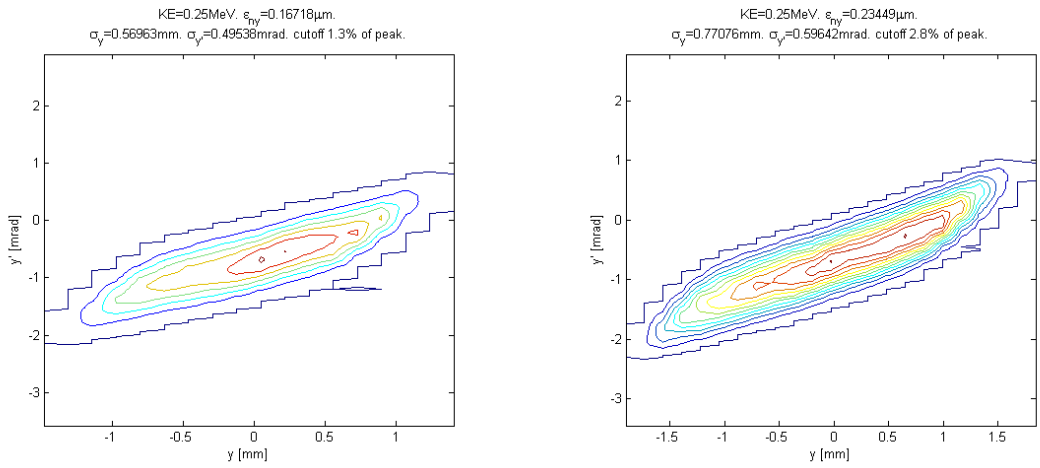

Figure 6.4: Phase space of the electron beam from a 1.5mm (left) and 2.0mm (right) emitting area on the cathode (532nm)

For the 1.5mm laser spot diameter the emittance was found to be 0.167μm and 0.234μm at 2mm. This was within 10% of the estimated thermal emittance from the previous chapter for the corresponding laser wavelength and spot size.

6.1.4 The Double Slit Method

A second system is used at Cornell to create phase space plots. It is similar to that described above, but uses a second slit and pair of correctors after the first, figure 6.5. As before, the first set of correctors move the beam over the first slit.

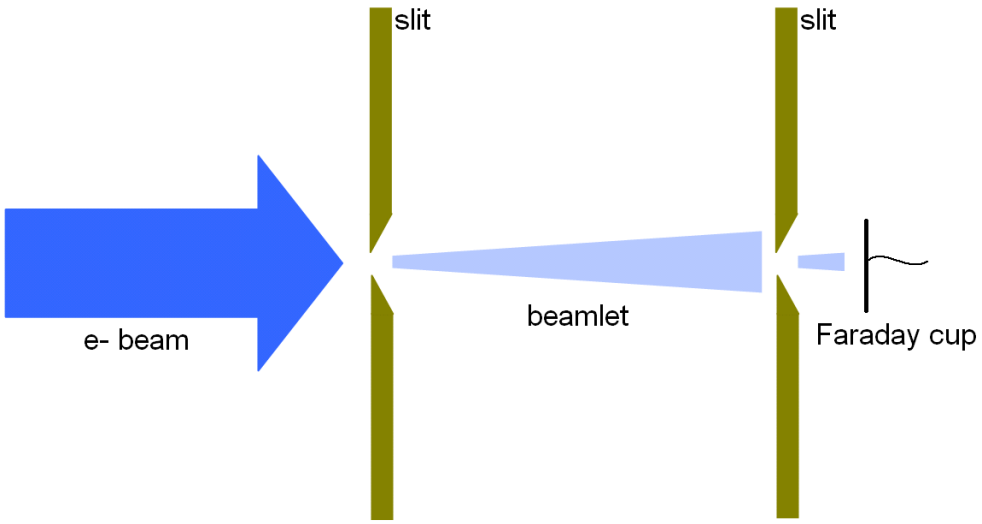

Figure 6.5: Double Slit diagnostic

The second set of correctors move the beamlet over the second slit. Rather than

generating the profile from a screen, it is created by measuring the current on a Faraday cup as the beamlet passes over the second slit. The benefit of this method being the Faraday cup is not sensitive to stray light from the laser as the screens are. The Faraday cup is used with an amplifier and can give a better resolution than that available from the screen. Additionally, when the electron bunch charge is high, even the small charge in the beamlets is enough to saturate the screens so they become unusable for profile measurement. The Faraday cup diagnostic has no such problems and therefore is useful for measuring the phase space of space charge dominated beams. A typical result from such a phase space measurement is shown in figure 6.6. Note that the emittance in figure 6.6a is within 4% of that measured by the solenoid scan method in chapter 5.

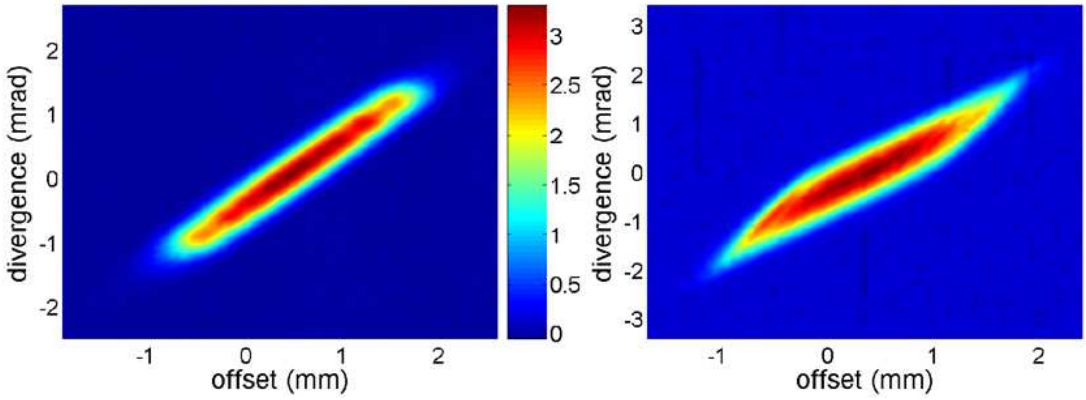

Figure 6.6: Phase space for GaAs (left) and GaAsP (right), (current in nA, 532nm, 1.5mm aperture at 250keV). GaAs ε_y = 0.180μm, GaAsP ε_y = 0.237μm

To calculate the emittance, an improved method is used on the phase space image rather than from beamlet information. Noise and the background subtraction are important when estimating the rms emittance using 2D intensity maps. To avoid introducing error, a self-consistent, unbiased rms emittance analysis tool (SCUBEEx) was used [91] . In this method, an ellipse that contains the data is varied in size. The region outside the ellipse is treated as noise, and its average value is subtracted from the whole data. The rms emittance is calculated as a function of the encompassing ellipse area. When the ellipse is large enough to include the full beam, the calculated emittance should not depend

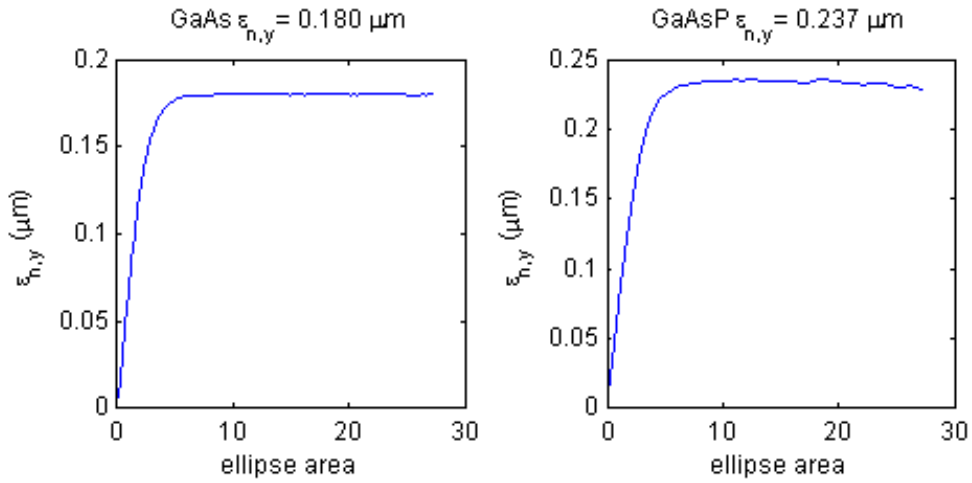

Figure 6.7: Calculated emittance from SCUBEEx, GaAs (left) and GaAsP (right)

on its size assuming the noise is uncorrelated in nature. The value measured in this way represents 100% of the beam rms emittance. The output from the SCUBEEx analysis on the above phase space images is shown in figure 6.7. This method gives a better representation of the emittance than that measured using the single slit technique described previously.

Further results from measurements using this diagnostic and emittance estimation technique will be shown later as a comparison to phase space images from tomography.

6.1.5 Phase Space Conclusions

Generating a phase space plot in this way gives useful information of the beam distribution and emittance at the location of the slit. The disadvantages of this method are that the phase space is created over an average of many electron bunches, and so will include any jitter that may be on the electron beam into the reconstruction. Additionally with a single slit, it is only possible to reconstruct the phase space of one transverse dimension; the vertical at Cornell; and at one longitudinal position, that of the slit. An alternative method is to use a multi-slit, which, as its name implies, has a series of slits in parallel, so that the

beamlets from each slit can be measured simultaneously so neither the beam nor slit need be moved. The downside of this approach is that the total number of beamlets is reduced because of physical restrictions on the distance between each slit. Sometimes also, the beamlets overlap once they have reached the screen, so more complex analysis is required to retrieve the phase space information. Despite these issues, the multi-slit method is widely used [92][93][94]. A further extension of this idea is to use a pepper-pot mask [95]. This diagnostic has an array of small holes that creates round beamlets. The advantage of this method is that the phase space of both transverse planes can be estimated for a single electron bunch [96]. A final benefit of all these techniques is that they can easily be applied to space charge dominated electron bunches as the methodology and reconstruction technique do not change.

6.2 Tomography

Tomography used as a technique to reconstruct images from sets of profiles is most commonly known from the medical physics arena. The first experiments utilised X-rays to form images of tissue based on their X-ray attenuation coefficient. The process of inferring information from density distributions that cannot be measured directly is ideal for use with electron beams where distributions of phase space are unknown.

Tomography is closely related to a theorem by Radon, who has shown that an object can be completely reconstructed from an infinite set of all its projections. Of course, in practice, it is not possible to collect an infinite number of projections, and so some error is introduced when the reconstruction is performed. The aim is to reduce this error through the correct choice of reconstruction algorithm for the problem. A projection can be calculated by integrating some distribution, $f(x, y)$, along a line. This is shown schematically in figure 6.8. The

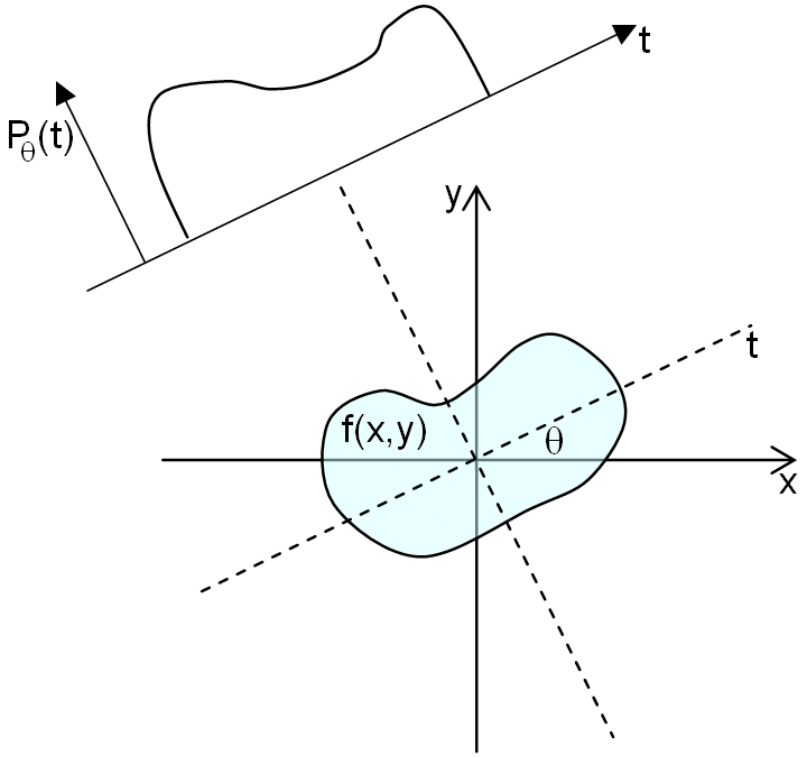

Figure 6.8: A distribution and its projection

equation of a line and its integral is:

$$x\cos(\theta) + y\sin(\theta) = t \tag{6.1}$$

$$P_\theta(t) = \int_{(\theta,t)} f(x,y)ds \tag{6.2}$$

The projection, or Radon transform, of $f(x,y)$ expanded using the delta function is given below, and forms the basis of the tomography reconstruction procedures used later in this chapter.

$$P_\theta(t) = \int\int f(x,y)\delta(x\cos(\theta) + y\sin(\theta) - t)dxdy \tag{6.3}$$

A set of Radon transforms of the distribution for angles between 0 and 180° is required for the different reconstruction algorithms used to recreate phase space. Note that projections from 180 to 360° are just the reverse of those from 0 to 180°.

There have been many methods employed previously in phase space reconstruction experiments. A comprehensive list of tomography experiments performed at laboratories world wide can be found in [97, 98]. Excluding the multi-turn tomography measurements in synchrotrons, where Gaussian approximations are made, all have used quadrupoles to rotate the phase space. One method is to use a quadrupole scan combined with tomographic image reconstruction techniques [99]. The quadrupole scan as an emittance measurement technique is a variant of the solenoid scan featured in chapter 5. If the quadrupoles are set to give the correct rotation in phase space and projections are taken, a reconstruction can be made [100, 101]. A second approach is to have a well known repeating lattice of quadrupoles, where the phase advance is defined for each cell. Profile measurement devices are placed in each cell location for use with reconstruction [102, 98]. A shortcoming of this latter method is that only a few projections can be taken, and this will limit the choice of reconstruction algorithm used.

Using the Cornell diagnostic beamline, the tomography phase space reconstruction technique was extended to consider solenoids as the elements used for rotating the phase space [18]. As the energy of the electron bunches from the gun is low (< 250keV), it was possible to reconstruct phase space from beams both with and without space charge. Simultaneously a project at the University of Maryland also investigated the use of solenoids for a space charge dominated tomography experiment [103].

6.2.1 Tomography of Electron Beams

In this study only tomography of the transverse electron plane will be considered, though similar techniques can be applied to longitudinal phase space [104]. Although the equations will refer to the horizontal transverse plane, x, they are equally valid for the vertical, y. For a charged particle beam, the aim is to determine the two dimensional phase space distribution, $\mu\left(x_0, x_0'\right)$ at some

location z_0 along the beamline. Usually when an electron beam is measured, the four dimensional transverse phase space (x, x', y, y'), is projected onto a two dimensional surface (x, y). If $\mu\left(x_1, x_1'\right)$ is the phase space distribution at a second position z_1 and the system is linear, the phase space at z_1 can be calculated by mapping from the starting location with the transfer matrix $\mathbf{R}$:

$$\begin{pmatrix} x_1 \\ x_1' \end{pmatrix} = \mathbf{R} \begin{pmatrix} x_0 \\ x_0' \end{pmatrix} \tag{6.4}$$

where:

$$\mathbf{R} = \begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \tag{6.5}$$

Integrating the phase space distribution with respect to x_1' yields the beam profile along x_1:

$$P(x_1) = \int \mu(x_1, x_1')dx_1' \tag{6.6}$$

This can be expanded using the Dirac delta function:

$$P(x) = \int\int \mu(x_1, x_1')\delta(x_1 - x)dx_1dx_1' \tag{6.7}$$

Using Liouville's theorem which states that the phase space density remains constant along the lines of trajectory of the system, assuming that the transverse and longitudinal motion is uncoupled:

$$\mu(x_0, x_0') = \mu(x_1, x_1') \tag{6.8}$$

This is a valid assumption for long electron bunches. After some substitution it is possible to get the projection in terms of the starting density distribution and the transfer matrix components:

$$P(x) = \int\int \mu(x_0, x_0')\delta(R_{11}x_0 + R_{12}x_0' - x)dx_0dx_0' \tag{6.9}$$

R_{11} and R_{12} are scaling the x and x' directions respectively. To write the projection in the Radon transform format, the following relations are introduced:

$$\cos(\phi) = \frac{R_{11}}{\sqrt{R_{11}^2 + R_{12}^2}}, \sin(\phi) = \frac{R_{12}}{\sqrt{R_{11}^2 + R_{12}^2}} \tag{6.10}$$

The rotation of phase space (phase advance) is given by:

$$\tan(\phi) = \frac{R_{12}}{R_{11}} \tag{6.11}$$

The projection in the form of the Radon transform with rotation angles is given by:

$$P(x)s = \int\int \mu(x_0, x_0')\delta(\cos(\phi)x_0 + \sin(\phi)x_0' - u)dx_0dx_0' \tag{6.12}$$

where $u = x/s$ and $s = \sqrt{R_{11}^2 + R_{12}^2}$. This shows a simple relationship between the projection and the Radon transform. Comparing this to the Radon transform, equation 6.3, shows that the x coordinate of the measured profile must be scaled with $1/s$, and the projection with s. These equations form the basis of the quadrupole scan method, where the matrix is varied by changing the strength of quadrupoles between z_0 and z_1. For each matrix the rotation of phase space, and the scaling can be calculated from the matrix elements. Generally more than one quadrupole is needed to achieve rotation over a full 180°. As quadrupoles are focusing in one plane whilst simultaneously defocusing in the other, different settings will be required to recreate the horizontal and vertical phase space. Using solenoids ensures that both transverse planes can be reconstructed from the same set of measurements.

For electron beams without space charge it is easy to calculate the transfer matrix between the construction point, z_0, and the measurement point, z_1. For a beamline consisting entirely of solenoids, it is simply the product of the solenoid and drift matrix elements. As with the thermal emittance measurements the transfer matrix was calculated using the field maps of solenoids rather than

thin lens approximations. When space charge must be included, generating the matrix is more involved as a defocusing term must be introduced. This is discussed further in section 6.4.

6.2.2 Beamline Layout

For tomography to work well, the beamline needs to be capable of producing enough profiles for the reconstruction. In practice this means that the magnets and layout must have the flexibility to produce a set of matrices that will give a good range of rotations for the projections, given by equation 6.11. These ideally should be spaced equally between 0 and 180°. In addition, the scaling must be within a reasonable range to give measurable beam sizes, and the beam size at the measurement position must be able to fit on the diagnostic so a full projection can be calculated. The beamline was therefore designed to perform with these constraints.

Initially a beamline with two solenoids was investigated. The magnet positions were somewhat bound by the location of other elements, such as the light box and slit diagnostic, which could not be moved. An unconstrained nonlinear optimisation was performed using a MATLAB script, which would find settings for the two solenoid fields that would give a particular rotation, and meet the other beam size constraints. Using this method it was possible to find matrices that would give 180 rotations equally spaced between 1 and 180° for the case where there was no space charge to consider. Unfortunately, when the same process was applied to the scenario *with* space charge, it was not possible to find solutions for as many rotations; and those found were not evenly distributed. As expected, the more the current in the bunch was increased, the fewer rotations could be found. A further solenoid was added to the beamline which gave greater flexibility and allowed for more rotations to be possible. The final beamline layout had three relatively equally spaced solenoids between the electron gun and the view screen used for measuring the profile. The view screen

was 2.21m from the cathode. A schematic of this is shown in figure 5.4. There is a limit to the space charge permissible before a full set of 180, 1° rotations can no longer be made, as there is limited flexibility with the magnets and layout to create the desired matrix. Nevertheless, it is not essential to have this many rotations, nor is it necessary to have them equally spaced. The number and spacing of rotations is dependent on the reconstruction algorithm used and also the shape of the distribution to be reconstructed. If only a few rotations can be found it may still be possible to generate an accurate reconstruction.

6.2.3 Reconstruction Algorithms

The most common reconstruction algorithm used for phase space tomography is the filtered back projection algorithm (FBP) [97, 101]. It is widely used because the mathematics involved are easily understood and programable. The algorithm utilises the Fourier slice theorem. The 1D Fourier transformed projections are summed together and an inverse 2D Fourier transform reveals the original distribution. Appendix B gives the mathematical algorithm for filtered back projection. All reconstruction algorithms contain some error in the reconstruction either because there are not enough projections, or assumptions are made to estimate the missing data. For the FBP algorithm to give a good reconstruction, many projections are required. An example of this is shown for a complex image in figure 6.9. Although the image is more complex than that of phase space, which is usually 2D-Gaussian or blob shaped, it shows the limitations of the algorithm. In addition to many projections being required for reconstruction, streaking artifacts are introduced. These can be seen more clearly in figure 6.10 for a simple ellipse. The images are normalised for comparison such that the total intensity is the same in each. If the emittance of the image were taken, the noise introduced from the streaks, which take both positive and negative values, would be included. Using the SCUBEEx technique here does not help, as the noise is correlated and not random. A useful feature of this algorithm is

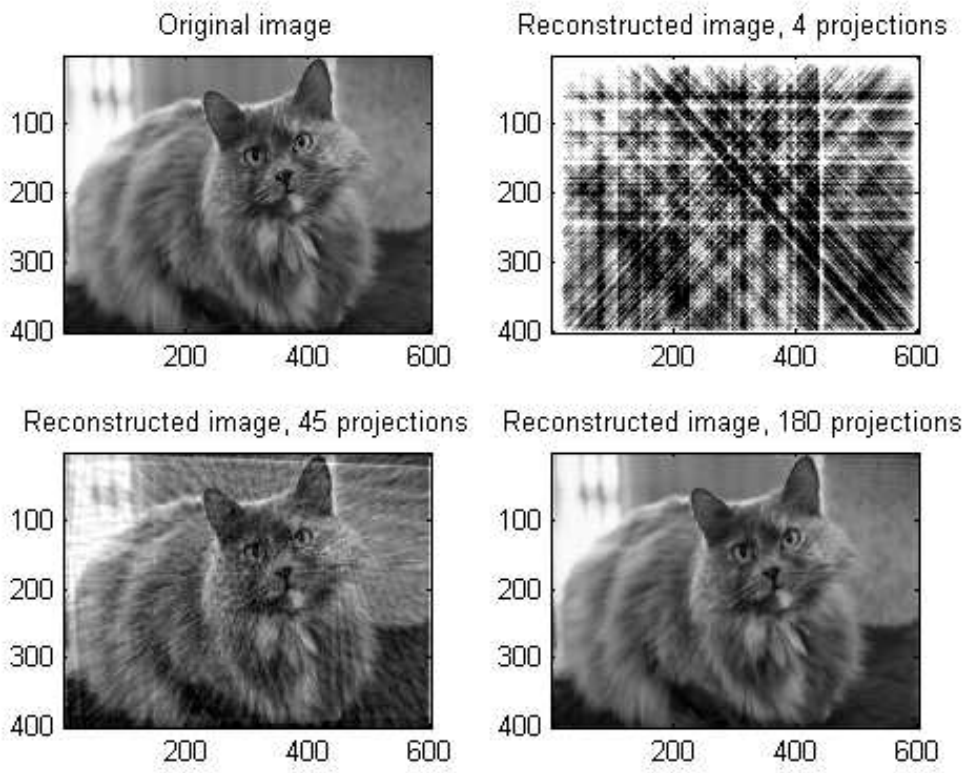

Figure 6.9: Reconstruction of a complex image using the FBP algorithm for differing numbers of projections

that with many projections, it can be implemented quickly to give a qualitative measure of the phase space distribution. Filtering the reconstructed image can also improve the quality, as briefly discussed in appendix B. All the images displayed in this thesis that have been reconstructed using the FBP algorithm, have been filtered using a Hann filter. Quantifying how good the reconstruction is can be difficult because of the artifacts. A simple error estimation can be given as the mean square deviation of the solution:

$$\delta_r = \sqrt{\frac{\sum_i \sum_j (f(x_i, y_j) - f_r(x_i, y_j))^2}{\sum_i \sum_j f(x_i, y_j)^2}} \tag{6.13}$$

where $f(x, y)$ is the original distribution and $f_r(x, y)$ the reconstructed one, i and j are the suffix for the x, y grid. The error in the reconstruction of the ellipse image using this measure is given as δ_r = 97, 12, and 5% for 4, 45, and 180

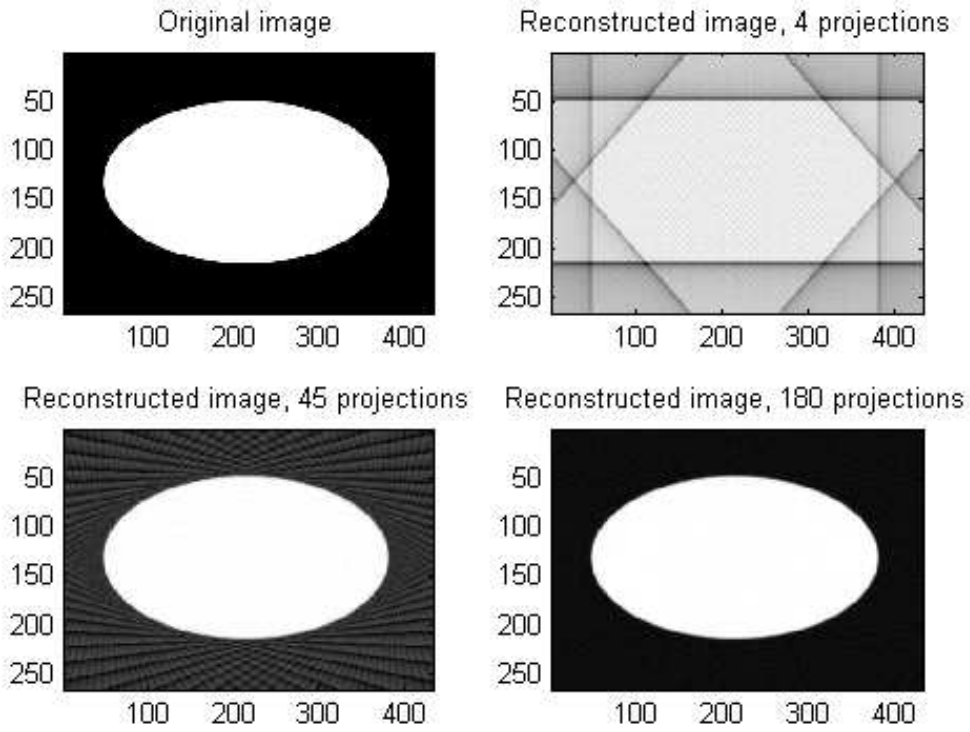

Figure 6.10: Reconstruction of an ellipse using the FBP algorithm for differing numbers of projections

projections respectively. Figure 6.11 shows the reconstruction error as a function of the number of projections used in the reconstruction.

Choosing the correct algorithm to use is largely dependent on the problem being solved. Some algorithms are better at reconstructing Gaussian distributions, whilst others are suited to detailed distributions. Popular reconstruction algorithms used for phase space tomography, in addition to the FBP method, are the maximum entropy (MENT) algorithm, used at Los Alamos, PITZ and Tokyo University [105, 106, 107], and the maximum likelihood - expectation maximisation (MLEM), used at Kyoto University [108].

Iterative reconstruction methods start with an estimate of an object function, and establish a relation between that and the measured projections. Then a minimisation problem is formed to measure the distance between the model-generated projections and those measured. The algorithms differ in the way

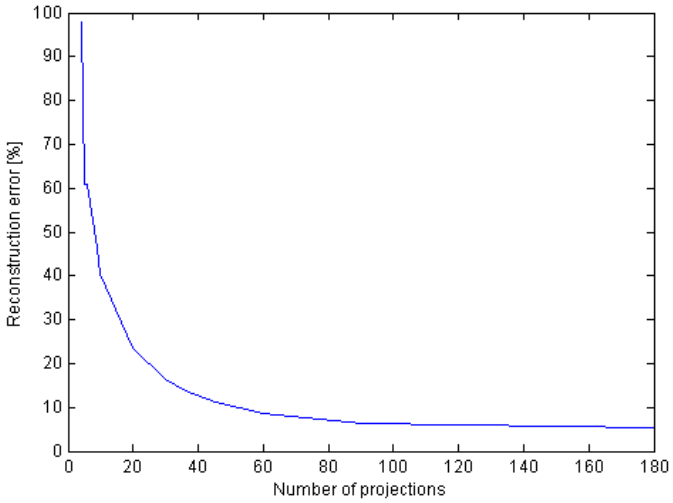

Figure 6.11: Reconstruction error as a function of number of projections for the ellipse image

the measured and estimated projections are compared and the kind of correction applied to the current estimate to create a new estimate. The MLEM is one such iterative method [109]. The algorithm is designed to compute the most likely distribution, given the measured projections. The mathematics of this algorithm are shown in appendix C. The algorithm is terminated after a user defined number of iterations and works well with complex images, see figure 6.12. An optional penalty function can be used to reduce the impact of noise in the projection being visible in the reconstruction. This is a median root prior (MRP) algorithm [110] described in appendix C. The projections, and their corresponding angles are used as input into the algorithm, as well as the weighting for the penalty. The advantage of this method is that fewer projections are needed to reconstruct simple shapes, but, the time taken to make the reconstruction increases as more iterations are required. The trade-off between number of projections and number of iterations is shown in figure 6.13. With increasing iterations the streaking artifacts are reduced as the reconstructed image approaches the shape of the original.

With 18 evenly spaced projections, and 20 iterations, the original image is very well described by the reconstruction. The reconstruction error, summarised in

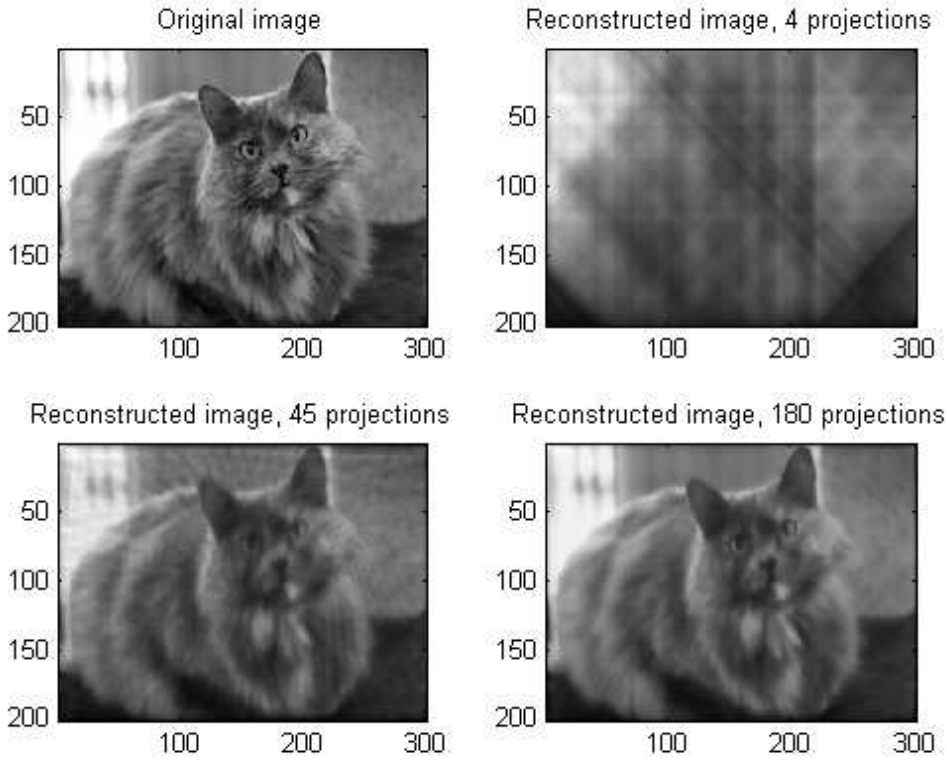

Figure 6.12: Reconstruction using the MLEM algorithm for differing numbers of projections, 70 iterations

6.1, is less than that from the 45 projection FBP reconstruction. For the case of 45 projections, the error introduced in the reconstruction as a function of number of iterations is shown in figure 6.14, reaching a minimum of 2.4% at 100 iterations. It is not always possible to create projections at every angle, making the MLEM algorithm very beneficial.

Iterations / Projections	5	10	20
4	27.61	20.57	18.68
18	25.12	15.14	10.80
30	25.04	14.94	10.58
45	25.03	14.92	10.56

Table 6.1: Error [%] in reconstructions using MLEM algorithm

The reconstructed image appears to be cleaner using the MLEM algorithm, and the minimum error reached with increasing projections and iterations is less

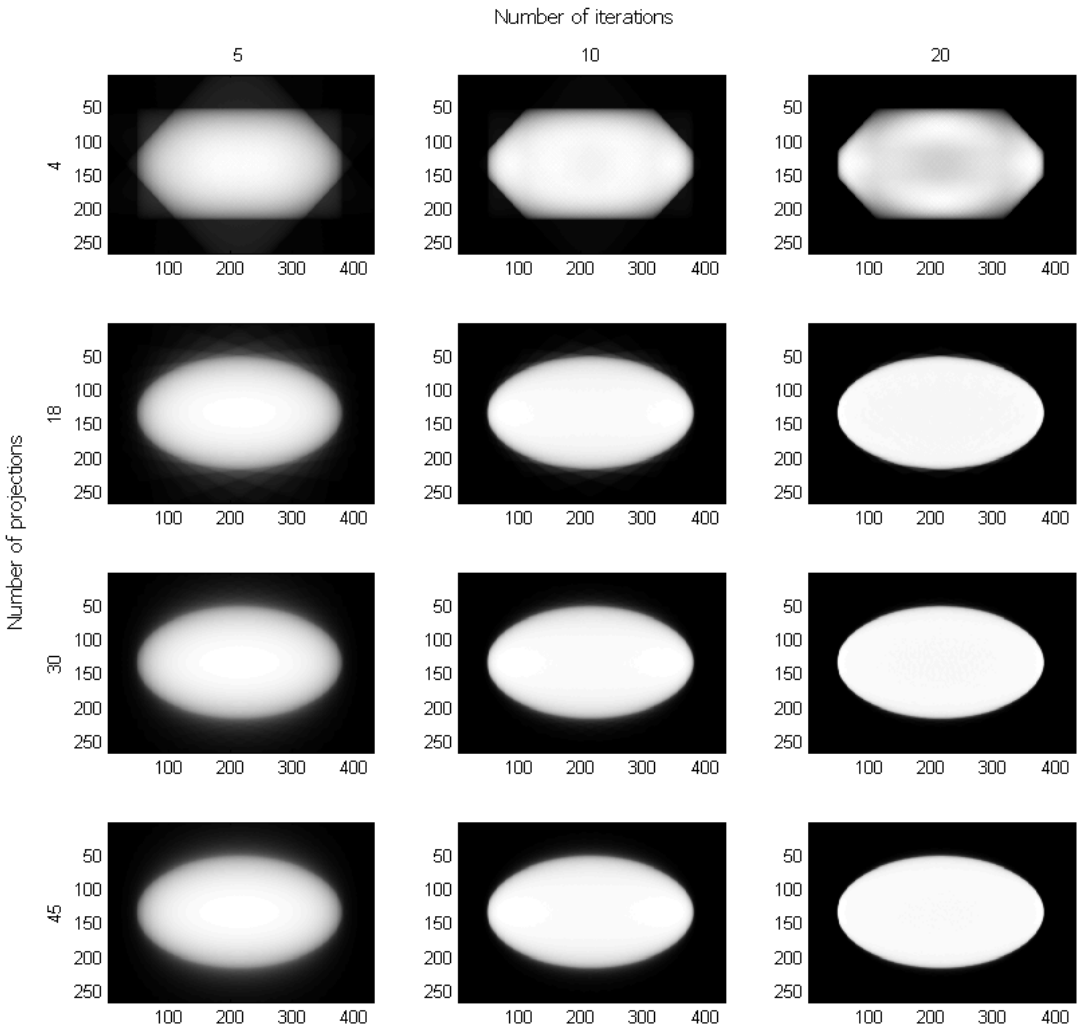

Figure 6.13: Reconstruction using the MLEM algorithm for differing numbers of projections and iterations

than that for the 180 projection FBP reconstruction. The algorithm is well suited to reconstructing the ellipse shapes of phase space. The measure that will be taken from the phase space images of electron beams, will be the emittance. For the original picture this is 3.441 in arbitrary units. The 180 projection FBP reconstruction gives 3.351 (2% error) and the 45 projection (20 iterations) MLEM gives 3.440 (0.04% error).

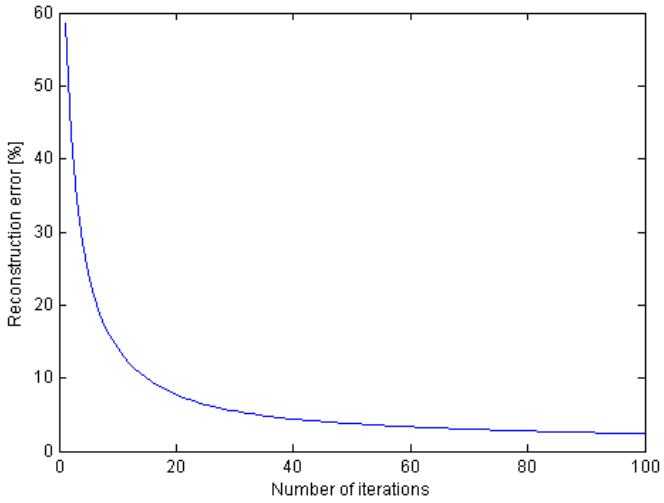

Figure 6.14: Reconstruction error as a function of number of iterations using the MLEM algorithm for 45 projections of the ellipse image

6.3 Phase Space Tomography Without Space Charge

6.3.1 Virtual Experiment

Before measuring beam data, a virtual experiment was performed to verify the performance of the algorithms with beam-like distributions. The solenoid settings were found for 18 rotations in 10° steps from 0 to 170°. Care was taken to ensure that the beam size at the position of the screen was not too large, and that the full beam fit onto the available area. ASTRA was used to calculate the evolution of the electron beam, given the field settings of the solenoid for each rotation. The initial conditions for the particle distribution were for a 2mm top-hat transverse profile, and a 40ps rms uniform distribution longitudinally. No space charge was included in the simulation to model a very low charge bunch. The number of macro-particles used was considerable - 10,000. This was to ensure that when the distribution was used to create what would look like an image from a screen, it would be smooth. The thermal emittance for the wavelength of 520nm was included, as this was the wavelength used in experiments. The particle distribution generated from ASTRA at the position of the screen (2.21m from the cathode) for each rotation was used to generate a

profile. A particle image is created in MATLAB by binning the particles into a pixel array (using actual pixel sizes from that screen). The number of particles in each bin correspond to the intensity, and a gentle blur function (convolution of the image with a 3×3 square) was applied to smooth the image. The result of this is shown in figure 6.15. The profiles were then taken from the set of

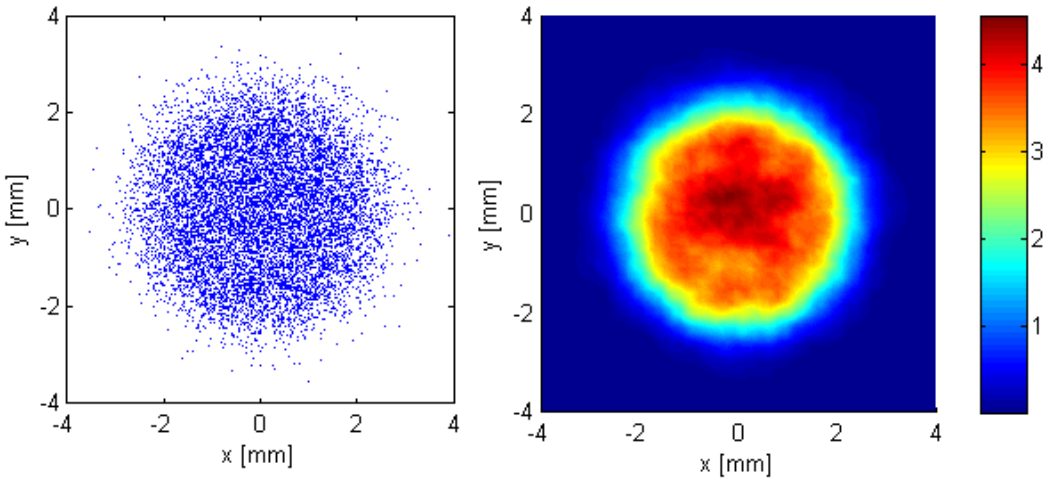

Figure 6.15: Sample particle distribution at the position of the screen (left) and the corresponding image created with intensity scale in arbitrary units (right)

images, scaled according to equation 6.12 and then used for the reconstruction. The phase space reconstructed using the FBP and MLEM algorithms is shown in figure 6.16. This can then be compared to the phase space directly taken from the ASTRA simulation at the reconstruction position, shown in figure 6.17.

As this virtual experiment only used a small number of projections, the MLEM algorithm performed better than the FBP. The emittance of each image should be equal to the value of thermal emittance included in the simulation, as the bunch charge was so low that space charge did not contribute to emittance growth. The emittance given directly from the ASTRA simulation was 0.243μm, and that from the image generated from the particle distribution was 0.292μm. The 6.6% increase in emittance is due to the process of blurring the particle

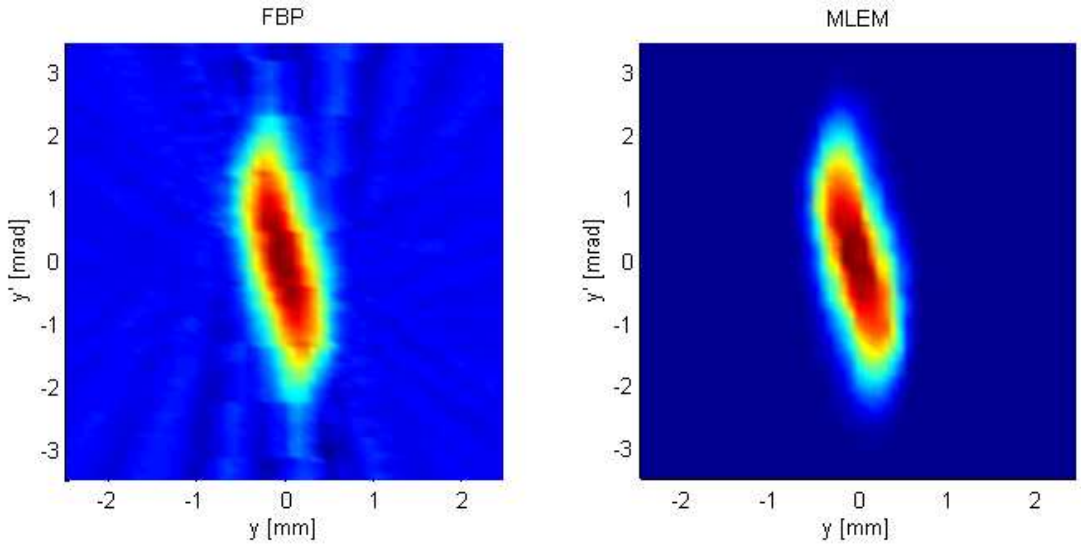

Figure 6.16: Reconstructed phase space using 18 projections, FBP (left) and the MLEM with 30 iterations (right)

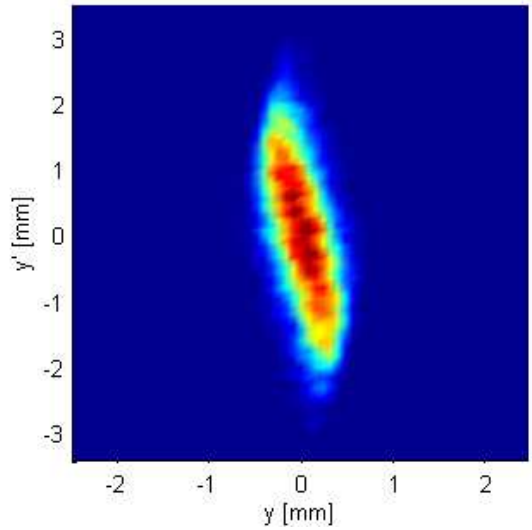

Figure 6.17: ASTRA phase space at the reconstruction location

distribution to create an image, and also the projection of particles in the simulation onto a plane. The particle distribution that is created from ASTRA is centred around the reference particle, so the head of the bunch longitudinally is already inside the solenoid field and the tail remains inside the gun field. It is not feasible to record the position and angle of each particle as it passes through the longitudinal position.

It was not possible to extract a meaningful estimate of emittance from the FBP reconstructed image due to negative pixel values created by the algorithm. The MLEM reconstructed image had a projected emittance of 0.276μm. The MLEM algorithm produces a reconstruction to within 13.6% of that generated directly by simulation, and so shows the reconstructed method works well, as all the phase space features are represented.

6.3.2 Measurement Procedure

Following the success of the simulated experiment, a measurement was made under the same conditions using the diagnostic beamline. The solenoid field settings for this experiment were constrained to be as low as possible, whilst still producing the required rotations in phase space. This was due to a misalignment error that precluded the electron beam passing directly through the centre of the first solenoid. Passing through a solenoid off axis causes non-symmetric distortions in the phase space and also a transverse offset, so reducing the field minimised this error.

Before a tomography experiment was attempted, a test to find the error in using the matrix approximation for the beamline was performed. This measurement used a horizontal corrector magnet, the first solenoid and the first screen in the beamline. For different solenoid settings the angle of the incoming electron beam was changed using the corrector magnet. The beam passed through the solenoid and reached the screen with some transverse offset. The change in offset as a function of incoming angle can be calculated from equation 6.4:

$$\Delta x_1 = \Delta x_0 \mathbf{R}_{11} + \Delta x_0' \mathbf{R}_{12} \tag{6.14}$$

$\Delta x_0 = 0$ as this is not varied. This gives:

$$\frac{\Delta x_1}{\Delta x_0'} = \mathbf{R}_{12} \tag{6.15}$$

$\Delta x_0'$ was calculated by increasing the current in the corrector magnet, with no solenoid fields, and measuring the offset at the screen. A line of best fit was made, which gave the offset in terms of corrector magnet current. This was found to be 34.85mm/A, and when divided by the distance between magnet and screen yielded the divergence, $\Delta x_0' = 31.1$ mrad/A.

For each solenoid setting, between 0 and 3.5A, the offset at the screen was measured as a function of corrector magnet current, to give Δx_1. The result of finding the matrix coefficient was compared with that calculated numerically from the field map of the solenoid. This is shown in figure 6.18. Some error is intro-

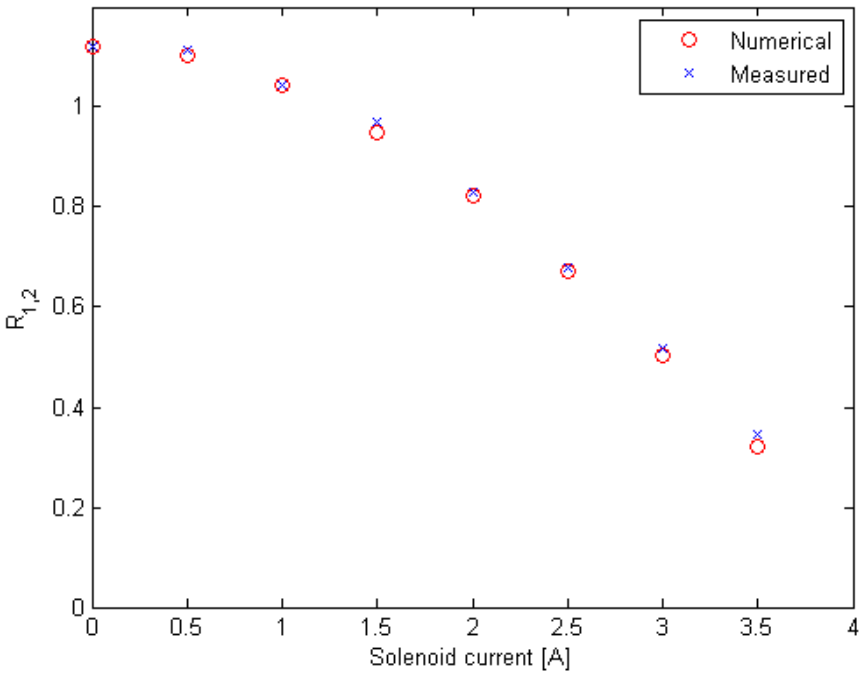

Figure 6.18: Comparison of calculated and measured values of $\mathbf{R_{1,2}}$

duced into this method by the electron beam passing through the solenoid off centre. Higher solenoid fields cause more distortion off axis, so the current was limited to a maximum of 3.5A in the solenoid windings. The maximum error was at 3.5A and was 7.5%, for lower currents the error was less than 3.2%. This demonstrated that the beamline was modelled correctly for the tomography experiment, and the maximum solenoid setting permitted was 3.5A.

The tomography experiment was partially automated using EPICS and MATLAB. For each rotation, the solenoids were cycled between their minimum and maximum currents to avoid errors from hysteresis before they were set. The beam spot was then manually positioned on the screen, away from any area with background noise created from stray laser light for example. A region of interest was selected around the beam using imaging software and the image within this area recorded using a 12-bit camera. The positioning of the spot at each rotation was necessary, as slight solenoid misalignments caused the beam to drift. Once the data set was complete, the images were then analysed and used for reconstruction. Each image was subject to a threshold to eliminate some of the background noise, and the beam spot was centred around the mean position. An example of this is shown in figure 6.19.

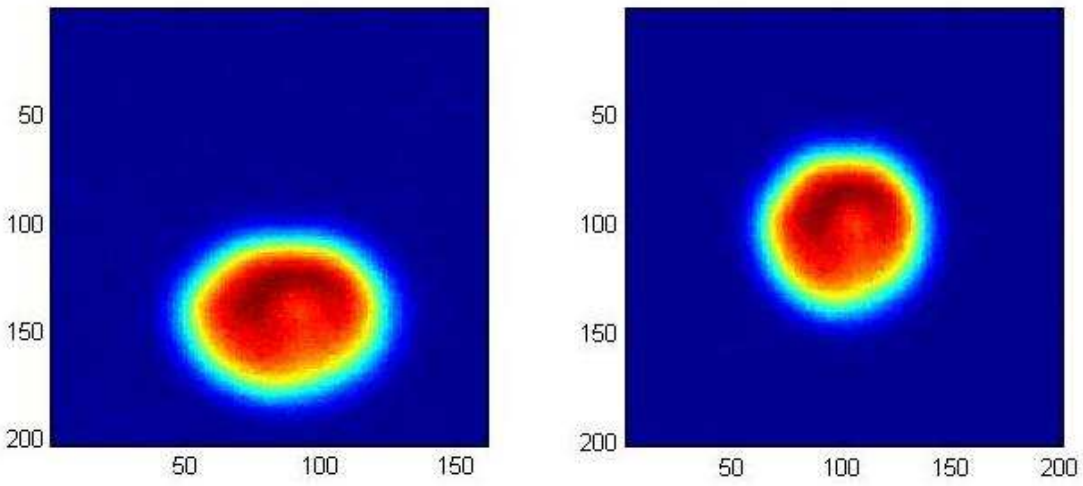

Figure 6.19: Example of image manipulation, image before (left) and after (right)

Unlike quadrupoles, solenoids produce a coupling in motion between the x and y planes due to the rotation of the electron beam, so 4x4 transfer matrices are required. Any 4x4 matrix can be expressed as the product of two affine matrix operations: scaling and rotation. The solenoid 4x4 transfer matrix can therefore be written as a product of decoupled thick lens $\mathbf{R}_{dec}$ and rotation matrices

$\mathbf{R}_{rot}(kL)$:

$$\mathbf{R}_{sol} = \mathbf{R}_{dec}\mathbf{R}_{rot} \tag{6.16}$$

$$\begin{pmatrix} C^2 & \frac{SC}{k} & SC & \frac{S^2}{k} \\ -kSC & C^2 & -kS^2 & SC \\ -SC & -\frac{S^2}{k} & C^2 & \frac{SC}{k} \\ kS^2 & -SC & -kSC & C^2 \end{pmatrix} = \begin{pmatrix} C & \frac{S}{k} & 0 & 0 \\ -kS & C & 0 & 0 \\ 0 & 0 & C & \frac{S}{k} \\ 0 & 0 & -kS & C \end{pmatrix} \begin{pmatrix} C & 0 & S & 0 \\ 0 & C & 0 & S \\ -S & 0 & C & 0 \\ 0 & -S & 0 & C \end{pmatrix} \tag{6.17}$$

where $C = \cos kL$, $S = \sin kL$, kL is the Larmor angle, and $k = B_z/(2B\rho)$ for a region of uniform axial magnetic field of magnitude B_z and length L. For a beamline consisting entirely of solenoids the 2x2 $\mathbf{R}_{dec}$ matrix is simply the product of the corresponding decoupled solenoid and drift matrix elements. Additionally, the total Larmor angle is calculated as $\theta_L = \int B_z/(2B\rho)$. For non symmetric beams, the image was rotated by the Larmor angle of the solenoids to re-orientate the image on the x and y axes. Both the $\mathbf{R}_{dec}$ matrix elements and θ_L were calculated using a field map of the solenoids rather than a thin lens approximation for increased accuracy. Once the obtained x, y images are rotated by -θ_L, the problem of tomography is reduced to the usual 2D phase space reconstruction with both x, x' and y, y' distributions available simultaneously.

The projection along each axis was then taken by summing the intensity of the pixels in each row or column. The projections are then scaled by the factor given by equation 6.12. These modified projections are then used for the reconstruction with both the FBP and MLEM algorithms.

6.3.3 Results

The results of the vertical phase space reconstruction at 15cm from the cathode are given in figure 6.20. The horizontal reconstruction was almost identical as a round beam was used. For this experiment the 520nm laser was used with an aperture of 2mm. The laser operated CW and was used with filters so that the photocurrent measured on the Faraday cup was less than 1μA. A total of 18

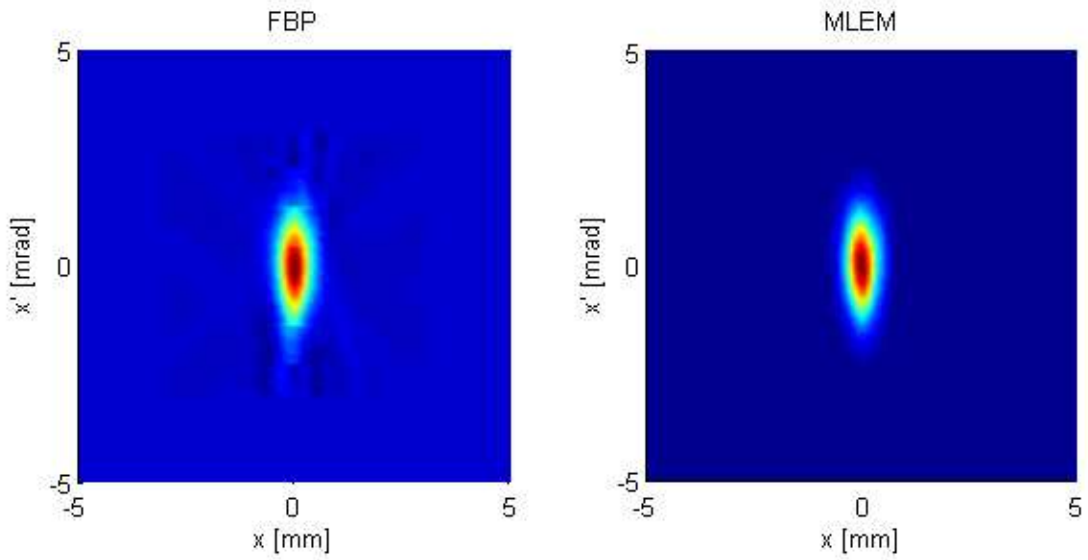

Figure 6.20: Typical phase space reconstruction

projections were collected and used for the reconstruction. This can be directly compared to the results of the virtual experiment as the simulation modelled the starting conditions of the beam used in this experiment. The emittance calculated from the MLEM reconstruction is 0.258μm in the horizontal plane and 0.287μm vertically. The projections were subject to a 1% threshold to reduce noise from the image without removing the tails of the distribution. The MLEM algorithm used 70 iterations and excluded the penalty function for the MRP filter. The horizontal and vertical normalised emittances are within 6% and 18% of the thermal emittance expected ($0.243 \pm 0.06\mu$m) for the laser spot diameter on the cathode and wavelength. The average emittance from 5 tomography experiments was 0.304μm horizontally and 0.332μm vertically.

To demonstrate that this method applies to non-Gaussian beams, an unusual electron distribution was created. Instead of using a normal aperture in the laser beam which would be projected onto the cathode, a mask was placed at that location. The mask was a 2.6mm diameter aperture with a 0.6mm wire bisecting it. The shape of the transverse laser profile, as measured using the 10-bit SPIRICON camera is shown in figure 6.21. This mask was imaged 1:1, to within 10%, onto the cathode where the electrons would be emitted in two

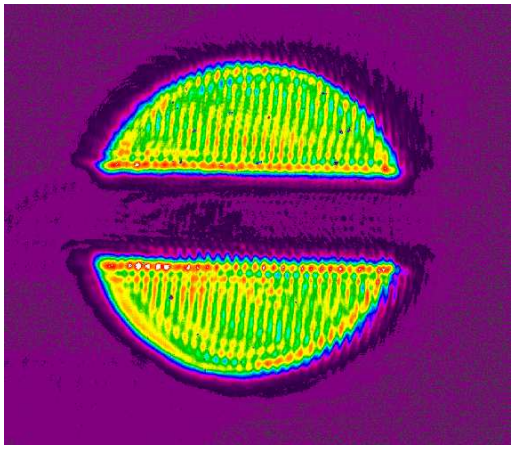

Figure 6.21: Laser profile with feature aperture

distinct lobes. For this investigation a 532nm laser was used, again with low power so the electron beam current was below 1μA. In this experiment it was important to rotate each image by the Larmor angle. The Larmor angle was calculated by integrating the measured field map of the solenoid and was found to be 10.2° per Amp. An example of the data manipulation is shown in figure 6.22. 33 projections were used to create the reconstruction shown in figure 6.23.

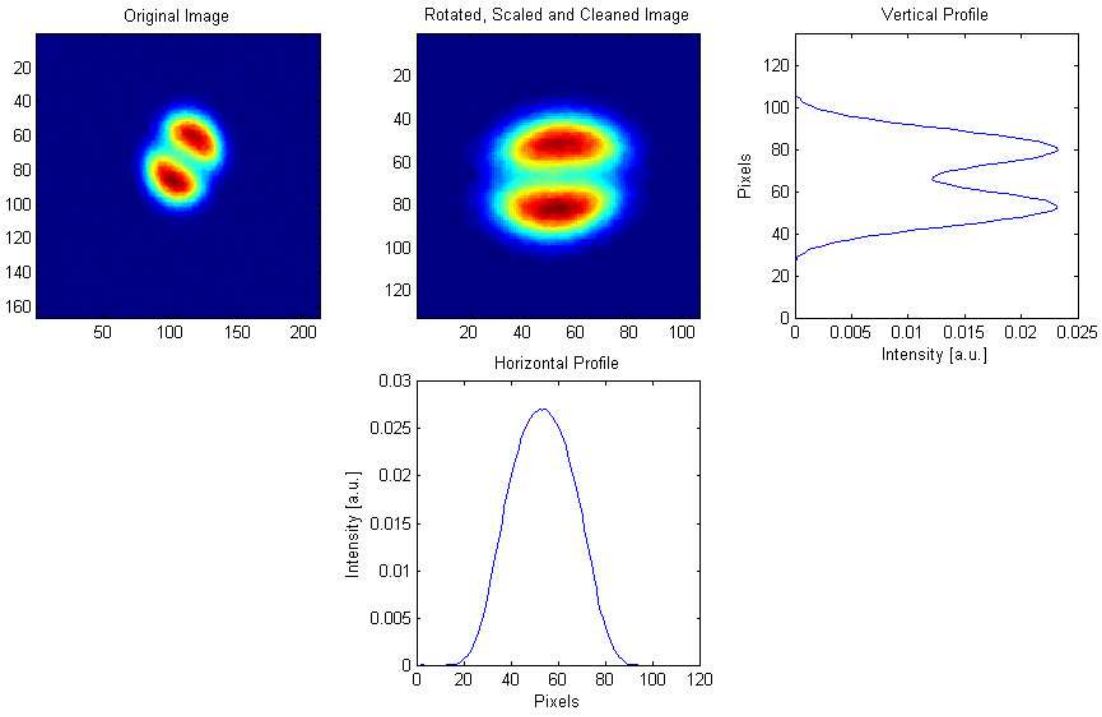

Figure 6.22: Original image (left), processed image (centre), vertical (right) and horizontal (bottom) profiles

The vertical emittance calculated from the MLEM reconstructed image in figure 6.23 is 0.539μm. The emittance calculated from ASTRA simulation is 0.480μm. To compare the reconstruction with experiment, the emittance was measured using the double slit method described previously. The phase space was measured at a position after the first solenoid which results in a different orientation of the phase space. The solenoid was set to zero, so there would be no rotation and the measured phase space is shown in figure 6.24. The emittance calculated from the image (with 1.5% cut-off threshold) and using SCUBEEx was 0.450μm and 0.445μm respectively. The reconstruction overestimates the emittance by 18%.

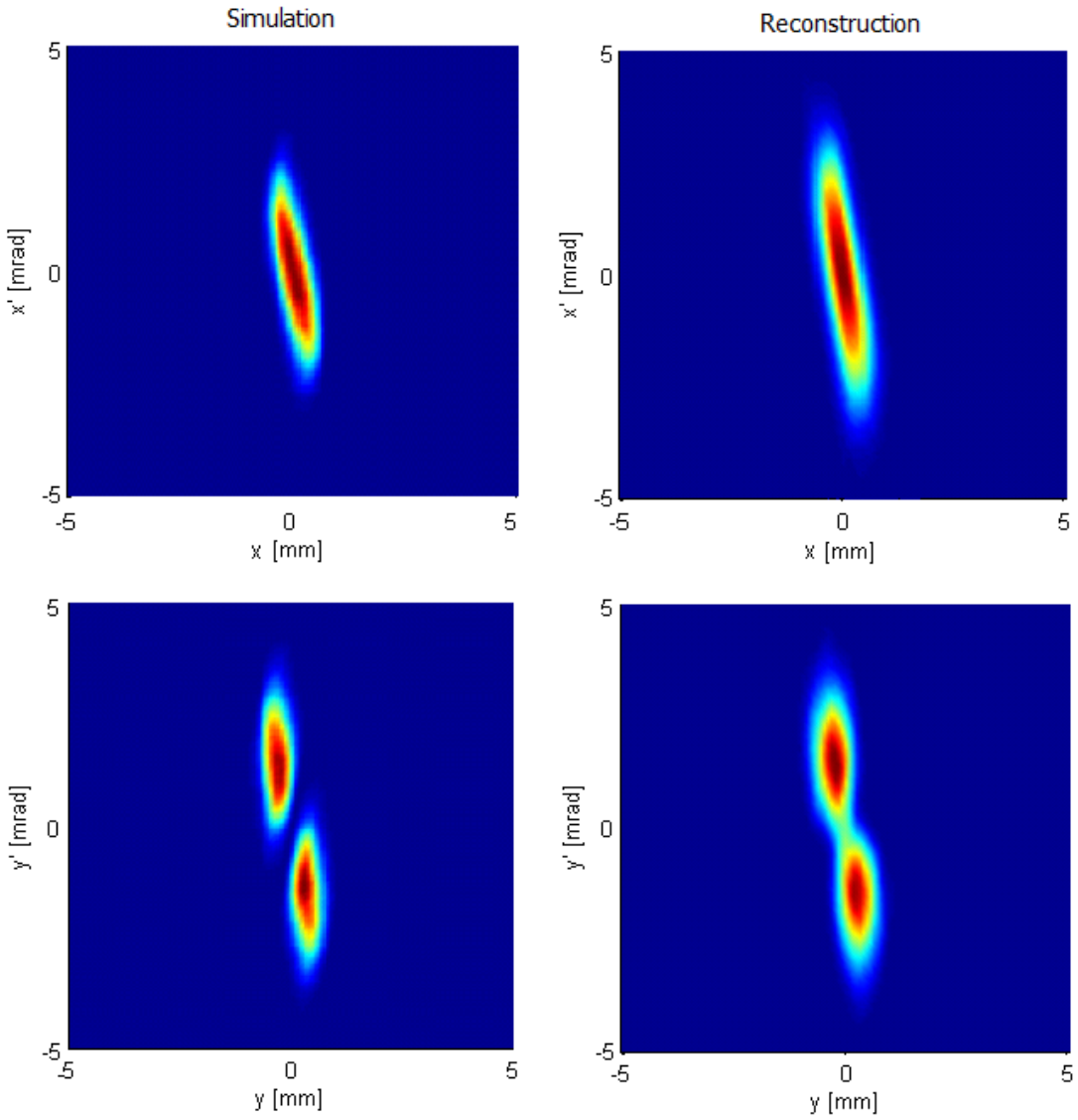

Figure 6.23: ASTRA (left) reconstructed (right) horizontal (top) and vertical (bottom) phase space

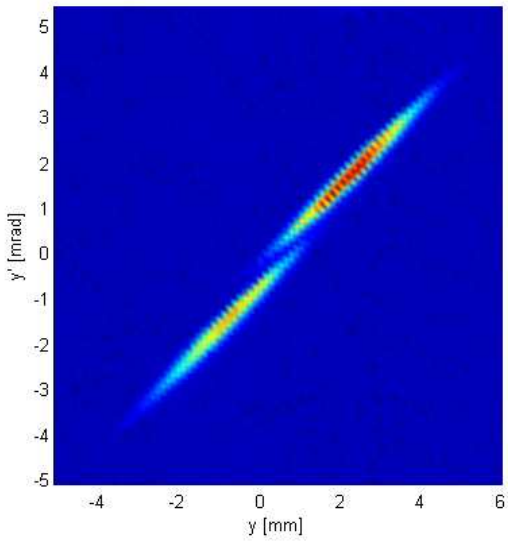

Figure 6.24: Phase space result from double slit scan

6.4 Phase Space Tomography With Space Charge

The evolution of an electron beam along a beamline including the influences of linear space charge, can be estimated using the beam envelope equation [111, 112].

$$a_x'' + k_x(z)^2 a_x - \frac{I_p}{I_a(\gamma\beta)^3(a_x + a_y)} - \frac{\varepsilon_x^2}{a_x^3(\gamma\beta)^2} = 0 \tag{6.18}$$

Where a_x is the beam half width, $k_x(z) = \frac{B(z)}{2B\rho}$ is the external focusing in a solenoid, and $B\rho$ is the magnetic rigidity. The third term in the equation is the space charge contribution, and the fourth term is the emittance term. The equation assumes that the electron beam is infinitely long and uniform with a current I_p. Finally, I_a is the Alfvén current (17kA for electrons) and β, γ are the relativistic beta and gamma respectively. It is also assumed in the equation that the particle distribution ($x, x'y, y'$) will produce a uniform elliptical distribution in phase space (xx' and yy') [113]. Kapchinskij and Vladimirskij (K-V) described such a distribution [114, 112] in which the charge density remains uniform.

To use the beam envelope equation, the starting conditions need to be known. That is: $\sigma_{x,y}, \sigma'_{x,y}, \varepsilon_{x,y}$ and I_p at position z_0 are required. For experimental purposes, if the electron bunch is uniform and of a known length, the current in the bunch, I_p, can be measured with a Faraday cup. To determine the other three starting conditions a solenoid scan can be performed. This will give a measure of the beam size as a function of solenoid field at the measurement location, z_1. Given that the solenoid field and position are well known, a minimisation problem can be solved using the beam envelope equation and the starting conditions as variables. An estimate of the starting conditions are made and the beam size is calculated at position z_1 for each solenoid field using equation 6.18. The difference between measured and calculated beam size is used to define how good the starting conditions are. By using an iterative algorithm to minimise this difference, the estimates for the starting conditions can be found. Once these are known, a matrix can be generated by breaking the beamline into small lengths

and including a defocusing lens of strength $\frac{I_p}{I_a(\gamma\beta)^3(a_x+a_y)a_x}$ at each increment. The complete transfer matrix will then include the effects of linear space charge.

6.4.1 Virtual Experiment

The beginning of the tomography section in the diagnostic beamline starts at the exit of the electron gun at 15cm from the cathode. Ideally the laser produces a 2.6mm diameter, radially uniform distribution transversely, and a 30ps FWHM uniform longitudinal distribution. Simulation shows that by the time the electron bunch reaches the end of the gun, the longitudinal shape is no longer uniform due to the effect of mirror charges at the cathode. With increasing charge per bunch the effect is more noticeable. Once the bunch exits the gun it expands longitudinally due to the space charge forces and energy spread, thus changing the average bunch current. The starting conditions of the electron bunch include an emittance originating from the thermal energy of electrons at the cathode. The transverse momenta distributions therefore do not satisfy the properties of a K-V distribution. Increasingly equation 6.18 becomes a poor approximation for the evolution of the electron beam.

The following virtual experiment explores the effect of bunch length and current on the ability to reconstruct phase space using a tomography method. The starting conditions at the gun exit were estimated from simulation for four different bunch charges, as shown in table 6.2. A thermal emittance of 0.3150μm was included at the cathode. The emittance and transverse size at the gun exit increase with increasing bunch charge. At the lowest charge of 0.5pC the emittance is entirely thermal and space charge forces negligible. The peak current is calculated by assuming a uniform longitudinal distribution of length $2\sqrt{3} \times \sigma_z$.

These parameters were used to generate the initial particle distribution, making the assumption that the distribution was radially uniform transversely, and uniform longitudinally so that the charge density would initially be uniform

Bunch charge (pC)	σ_z (mm)	σ_x (mm)	$\epsilon_{x,n}$ (μm)	I_{peak} (A)
0.5	1.95	0.317	0.315	0.0164
1	1.97	0.352	0.316	0.0325
5	2.13	0.610	0.337	0.1482
10	2.33	0.887	0.377	0.2686

Table 6.2: Beam properties at the exit of the gun for different bunch charges

throughout the electron bunch. For each bunch charge the solenoid field settings for 16 rotations were found as before, except for including the defocusing effect of space charge into the transfer matrix. An ASTRA simulation was performed for each rotation using 5000 macro particles. The resulting distributions were used to create an image that could be used for the reconstruction as described in section 6.3.1.

To demonstrate how the bunch length affects the reconstruction it was doubled and quadrupled for each case. To keep the starting current the same, the bunch charge was scaled accordingly.

6.4.2 Results

Figure 6.25 shows a comparison of the initial phase space distribution and that which was reconstructed. It can be seen that as the bunch length decreases and the bunch charge increases, the reconstructed phase space becomes less accurate. This is because the beam envelope equation begins to make a poor approximation to the simulated beam envelope, and therefore the linear space charge term included into the transfer matrix does not represent the beam dynamics of the simulation. As the electron bunch gets shorter, particularly for the higher current scenarios, curly tails and misaligned features start to appear in the phase space. This is a direct consequence of the initial bunch not being a K-V distribution. Different longitudinal slices will experience different space charge forces. The slice emittance, described in section 2.1.7, varies along the length of the bunch and they fail to align in phase space.

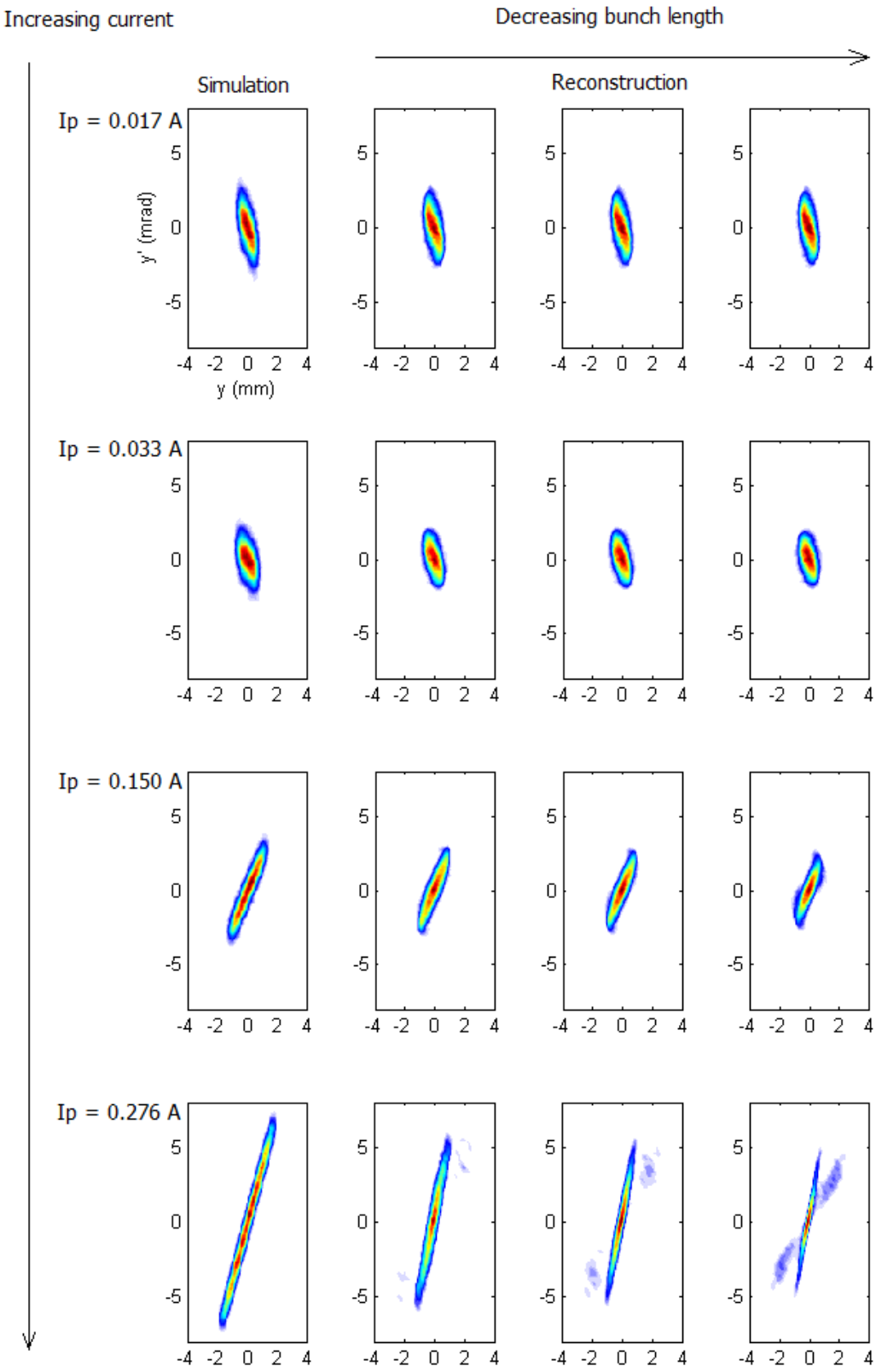

Figure 6.25: Tomographic reconstruction of phase space compared with that directly generated by simulation (left)

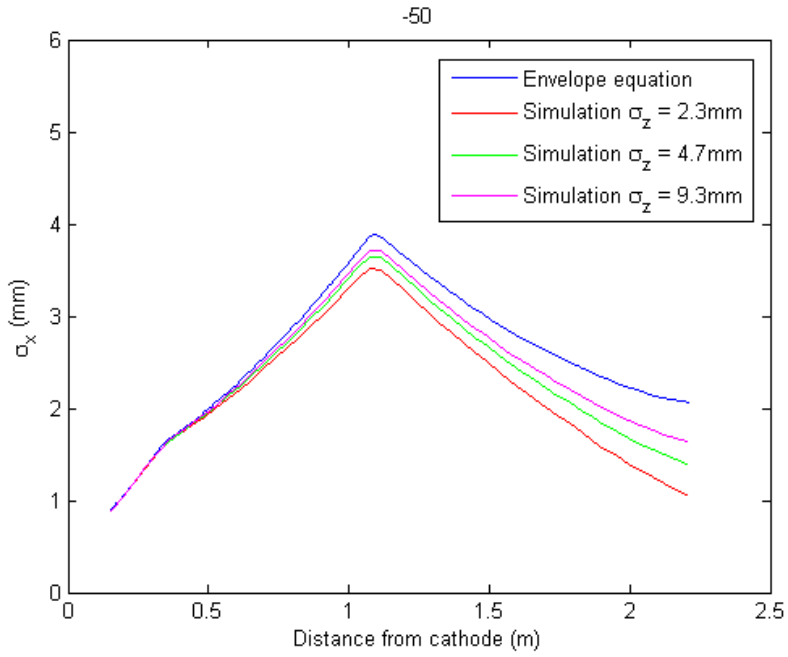

Figure 6.26: Comparison of simulation and the envelope equation for a space charge dominated bunch

Figure 6.26 shows a comparison of the beam envelope approximation with that from simulation for the 10pC bunch charge. In this case the electron bunch evolution is almost completely dominated by the space charge forces. The solenoid field settings correspond to that for a rotation of -50°, which show a typical mismatch between the simulation and analytic equation. It can be seen that even at the longest bunch length, the beam envelope equation overestimates the transverse size. This is accounted for by the space charge forces within the electron bunch becoming increasingly non-linear as it progresses along the beamline. The consequence of this is that the profile of the distribution changes shape as it evolves along the beamline, rather than just size. Figure 6.27 shows a comparison of the desired beam profile required for an accurate reconstruction and that achieved from simulation. Not only is the shape very different in this case, but the absolute size is smaller. The reconstructed phase space therefore underestimates that derived directly from the initial particle distribution.

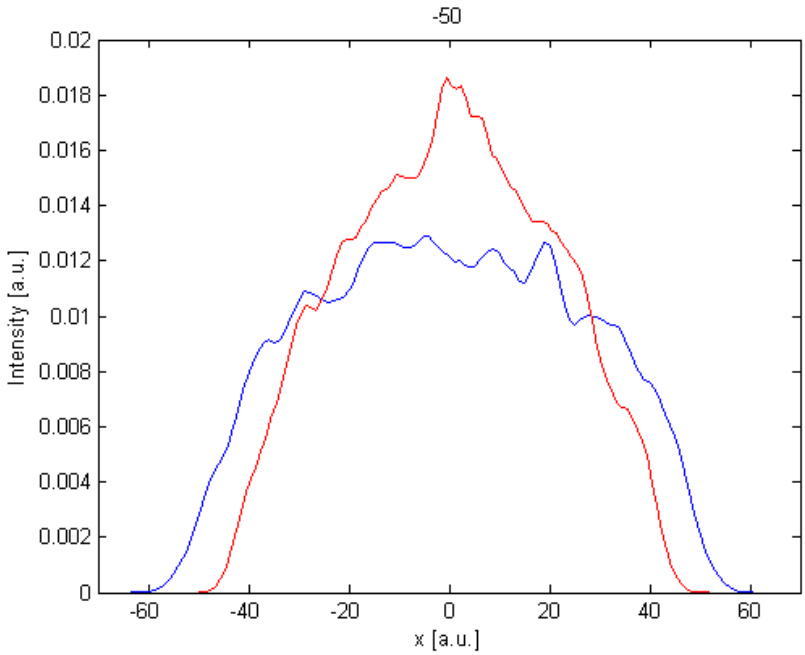

Figure 6.27: The desired beam profile (blue) and that created by simulation (red)

6.4.3 Tomography with Space Charge

A single experiment to reconstruct phase space with a space charge dominated beam was performed. Due to time limitations the properties of the electron bunch at the exit of the gun were inferred from simulation rather than measurement. Solenoid settings were found to give 18 rotations, however the maximum permissable solenoid field had to be increased to achieve this. The electron beam at the cathode had a 2.6mm diameter and 8.66ps rms bunch length. The bunch charge was measured with a Faraday cup and found to be 20pC. The reconstructed and simulated phase space is shown in figure 6.28. The reconstruction of this electron beam is rather poor, and does not represent the phase space expected. This is largely due to errors in the experimental set up. The first error was to come from the laser incident on the cathode. It was not possible to uniformly illuminate the 2.6mm aperture, which resulted in a non uniform electron distribution. Additionally the centroid of the laser did not align with the centre of the aperture, or the central axis of the beamline. The second error arose from a misalignment of the electron beam in the solenoids. Although a

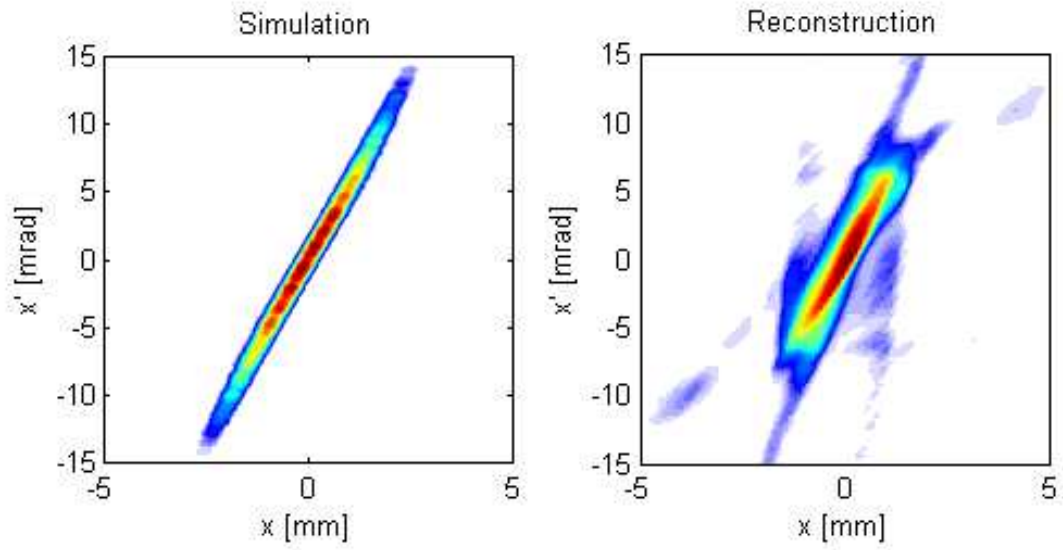

Figure 6.28: Comparison of simulated (left) and reconstructed (right) phase space

beam based alignment technique was used prior to the experiment, it was not possible to keep the electron beam on axis through all three of the solenoids. The effect of the electron beam passing through the solenoids off axis was particularly noticeable when the solenoid strength was high, as was the case for a large number of the images. Figure 6.29 shows a typical image captured on the view screen. The asymmetric nature of the electron beam implies that the beam envelope model does not apply to this experiment. This in turn results in error being introduced into the scaling and rotations used to reconstruct the phase space from the captured images.

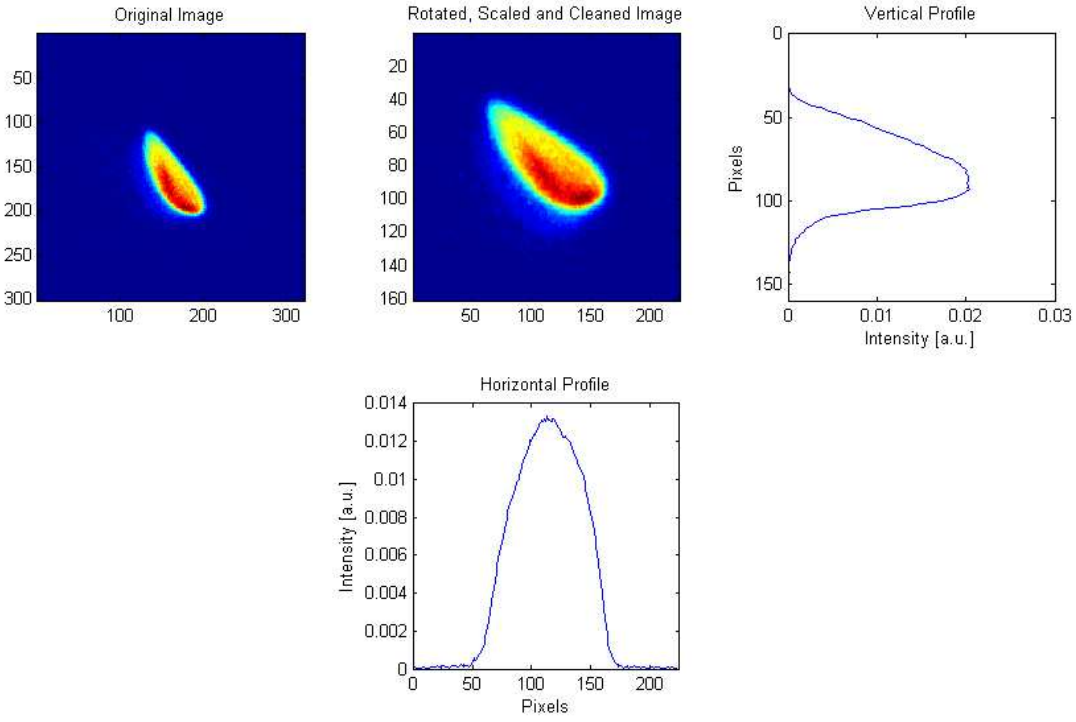

Figure 6.29: Typical image measured on the view screen

6.5 Conclusions

Direct measurements of phase space give useful information about the shape of the distribution in the electron bunch. Particle distributions in accelerators are generally not symmetric or perfectly Gaussian in shape once they have been transported through non-linear components, and detrimental effects such as wake fields and beam break up have acted on the beam. Phase space distributions can be informative about these effects, as well as yielding the beam emittance.

For low energy beams interceptive diagnostics can be used to measure the phase space directly. With increasing energy, these devices are not appropriate. For this reason tomography techniques are very useful in reconstructing the average phase space. It is possible to design a machine such that the transport magnets and some method of measuring the transverse projections, either screens or wire scanners, can be used for a tomography diagnostic. When only a limited number of rotations can be made it is then important to choose an appropriate reconstruction algorithm that can give a good estimate of the phase space.

In this tomography experiment at low energy and with an electron beam with negligible space charge, three solenoids were used to rotate the phase space. The electron beam was reconstructed at a longitudinal position just after the gun, where the emittance contribution is entirely thermal. The reconstructed phase space represented the features expected through simulation well. This was demonstrated with the reconstruction of the two lobe electron beam which clearly had two separate halves of phase space. The measured emittance from the reconstruction was within 22% of that expected for the round beam with only 18 projections, and 18% for the two lobe beam.

Phase space tomography of emittance dominated beams, will produce the features of phase space that cannot be inferred from emittance measurements using

solenoid or quadrupole scan, particularly for non Gaussian beams. Solenoids can be successfully used as an alternative to quadrupoles for tomography experiments that produce the transverse phase space in both planes simultaneously. The reconstructions show the features of phase space well. It has been demonstrated that the method can be used to obtain quantitative information about the phase space (e.g. rms emittance), which was found in agreement to that measured by a direct method: however attention to the details of the reconstruction algorithm and image processing is required.

Finally, an attempt was made to apply tomography to space charge dominated bunched beam ($\sim$20pC/bunch) employing a 50MHz 520nm laser [115]. The transfer matrix was augmented using linear space charge forces [101]. Results of the tomography reconstruction in this case were inconsistent, thought to be due to several factors such as: the difficulty in obtaining sufficient rotation angles, and the fact that a simple linear space charge is insufficient to describe bunched beams with changing aspect ratio as found in the Cornell system. If the phase space or transverse shape of the electron beam deviates too far from a K-V distribution, this technique is not valid. Typically space charge dominated beams are found in the injectors of electron accelerators, and to measure this phase space it is more appropriate to use one of the interceptive diagnostics described at the beginning of this chapter. The following chapter describes the measurement of phase space for space charge dominated beams by using the double slit technique.

CHAPTER 7

Space Charge Measurements

Using a simple configuration of the DC gun and solenoid at Cornell University, measurements of the transverse phase space were taken at different electron bunch charges where space charge forces are significant. The results of the measurement were compared against simulation with a number of different 3D space charge codes, published in [19]. In the publication, 3D codes were used because of the asymmetry that was observed in the transverse phase space measurements. This was due to a non symmetric electron distribution at the cathode that arose from an asymmetric laser profile. For this reason 2D codes were not considered for simulations. However the distribution is not greatly asymmetric. Therefore, in this chapter, ASTRA was used to model the transverse phase space despite only having a 2D space charge algorithm at the cathode. The validity of using this modelling tool for space charge simulations is also discussed.

7.1 Experimental Set Up

The experiment was designed to make a comparison between space charge calculations and direct measurements of the electron beam in a simple system of electron gun and solenoid. Measurements were taken at 3 different bunch charges, 0.5pC, 20pC and 80pC, to give an impression of how increasing space

charge forces change the phase space. The electron gun was operated at 250kV and the current in the solenoid, at 33.5cm from the cathode, was varied as beam size and phase space measurements were taken. At the lowest bunch charge, the phase space was measured at one solenoid setting of 3.7A to provide the value of the cathode thermal emittance. For each of the higher bunch charges, two sets of measurements were made. To begin with the current in the solenoid was scanned over a range of values, whilst the transverse beam size was measured from images taken on a view screen. The second set of measurements involved a double slit phase space measurement at a number of solenoid settings, from which the emittance and beam size could be inferred.

The transverse laser distribution was captured using a camera, detailed in the next section, before each measurement. The laser intensity stability was measured to be 2% rms, and the pointing stability was 60 micron rms in each transverse direction. The laser spot was incident on an aperture, which was imaged without magnification onto the photocathode. The quantum efficiency of the cathode throughout the data taking was approximately 6%, with about 10% peak to peak variation over the illuminated area. These variations in QE could not be accounted for in the simulation as there was no reliable method of mapping them.

7.2 Simulation

The initial conditions of the simulation must represent those of the experiment as closely as possible in order to try and reproduce the measured properties of the electron beam.

The electron bunch emitted from the cathode is defined through the laser transverse and longitudinal shape, and the thermal energy of the cathode material. The transverse shape was measured by diverting the laser with a mirror onto a

camera positioned at the same distance as the cathode. The longitudinal structure of the laser pulse was inferred from a longitudinal measurement of the electron bunch, using the RF deflecting cavity and view screen. As described in chapter 5, the thermal emittance of the cathode was measured from the double slit phase space and is shown in the following results section.

For all measurements the same 2.6mm aperture was imaged 1:1 onto the cathode, however, the aperture was not uniformly illuminated. The laser power was changed to set the desired electron bunch charge for each case. A typical image of the transverse laser distribution is shown in figure 7.1. The horizontal and vertical profiles, also shown, are not uniform or Gaussian in shape and are not symmetric within the aperture. Additionally, the pointing stability of the

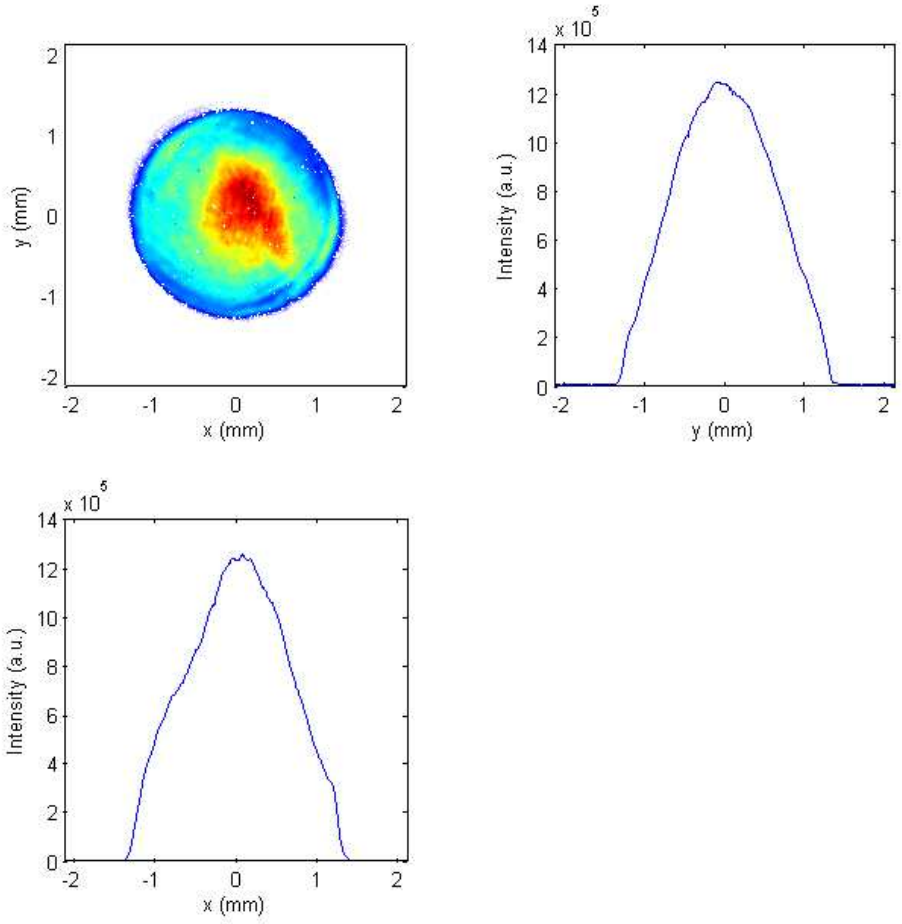

Figure 7.1: Laser intensity distribution with horizontal and vertical profiles

laser would cause a fast jitter, and the average peak position of the distribution would wander slowly with time. For these reasons, the transverse laser images for each measurement are different. Multiple images were recorded and one that best represented the mean centroid position was used to generate the particle distribution. The transverse laser intensity from the image was used as a 2 dimensional probability density function. A Monte Carlo sampling technique was used to find the transverse coordinates of particles to be used in simulation. The result for 10,000 particles is shown in figure 7.2. The mean of the particle distribution was set to equal zero for the simulation rather than symmetrically positioning the edges. This was a realistic scenario, as the laser beam

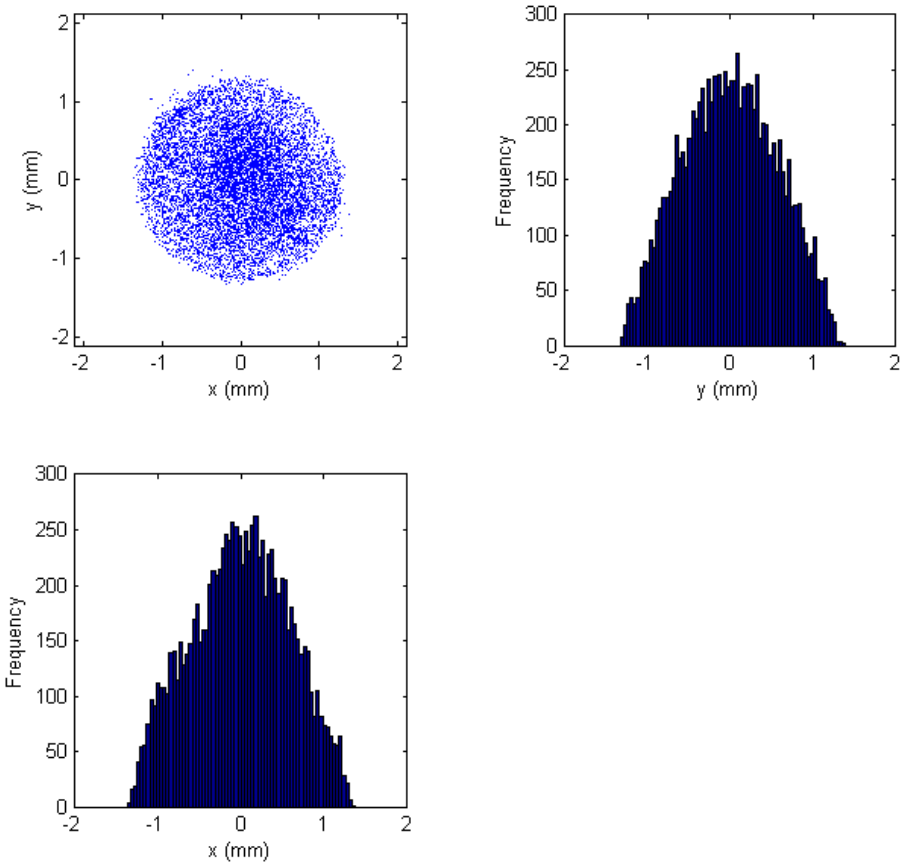

Figure 7.2: Particle distribution for simulation with 10,000 particles. The horizontal and vertical histograms are also shown

was positioned on the cathode such that the resulting electron beam centroid would not move with different solenoid strengths.

The laser pulse length is created by temporally stacking laser pulses using birefringent crystals. The drive laser provides a Gaussian pulse of 1ps rms duration [116]. The birefringent crystal splits this pulse into two smaller pulses with differing polarisations, and the temporal difference between the two is determined precisely by the length of the crystal. By having a series of crystals of different lengths, it is possible to stack the pulses to create a longer profile that is more uniform. The longitudinal laser profile was inferred from a measurement of the electron beam longitudinal properties. An RF deflecting cavity was used to flip a negligible charge electron bunch so that the longitudinal dimension could be projected onto a view screen, as shown in figure 7.3(a). A calibration image was then taken with only one birefringent crystal inserted, creating two pulses of known separation time, shown in figure 7.4. As described in [116], the resolution of the measurement was 1.5ps. Rather than use the measured profile to determine the shape of the particle distribution, the actual profile was

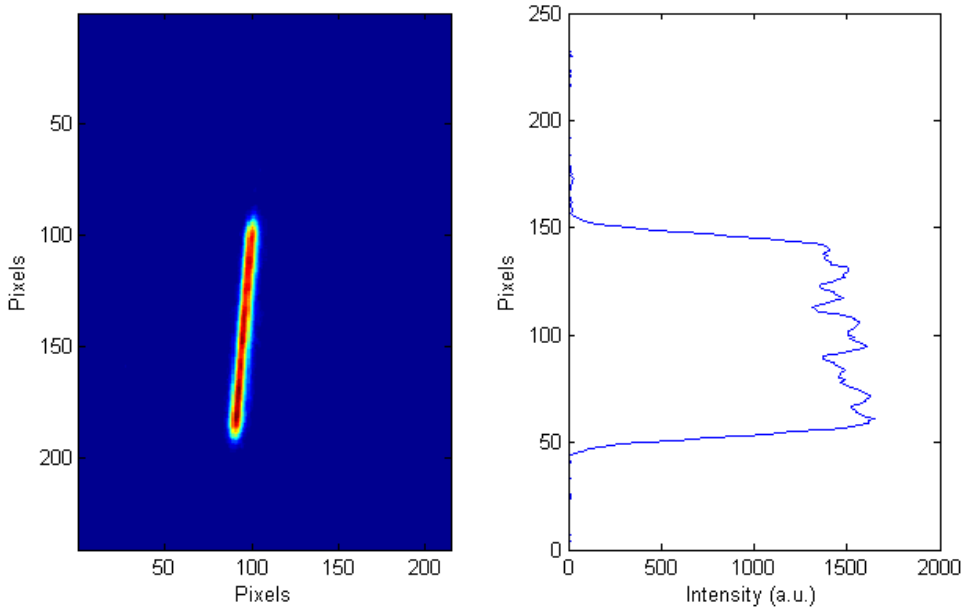

Figure 7.3: Longitudinal distribution of the electron beam a) view screen image, b) profile

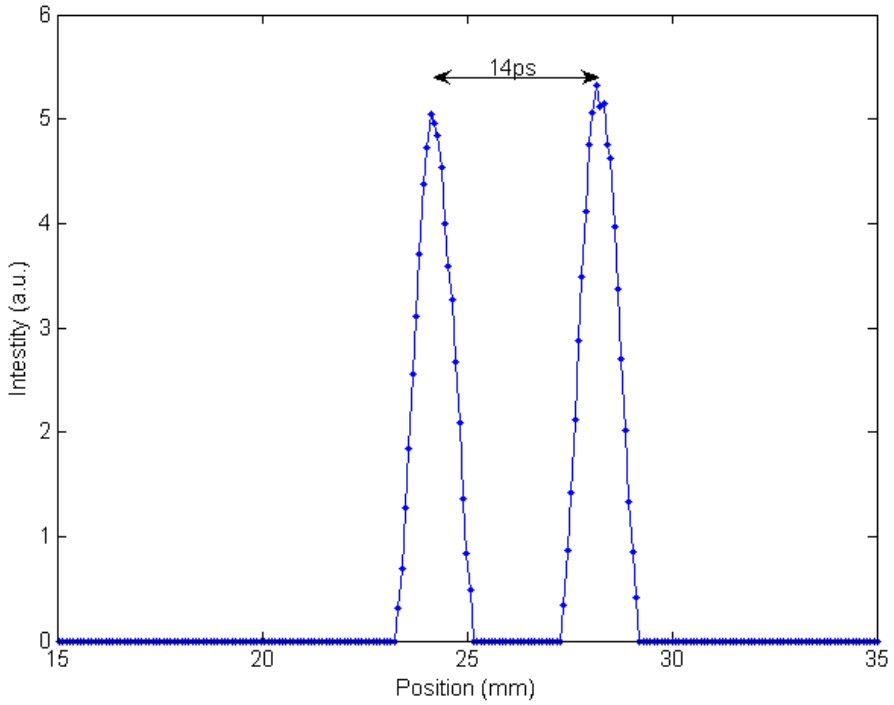

Figure 7.4: Calibration longitudinal profile with one crystal inserted

calculated. Three birefringent crystals were utilized to create a profile of 8 superimposed Gaussian distributions. The profile was fitted with the sum of 8 Gaussian functions to determine the mean and amplitude of each, as described in [19]. The real temporal profile was generated from assigning a sigma of 1ps to each Gaussian, as shown in figure 7.5. This profile was then used to determine the longitudinal properties of the particle distribution, figure 7.6.

Finally, the charge measured on the Faraday cup was used to determine the charge of the macro-particles, and the transverse momenta set to give the required thermal emittance for the simulation. This procedure was applied to the laser distributions for each measurement taken.

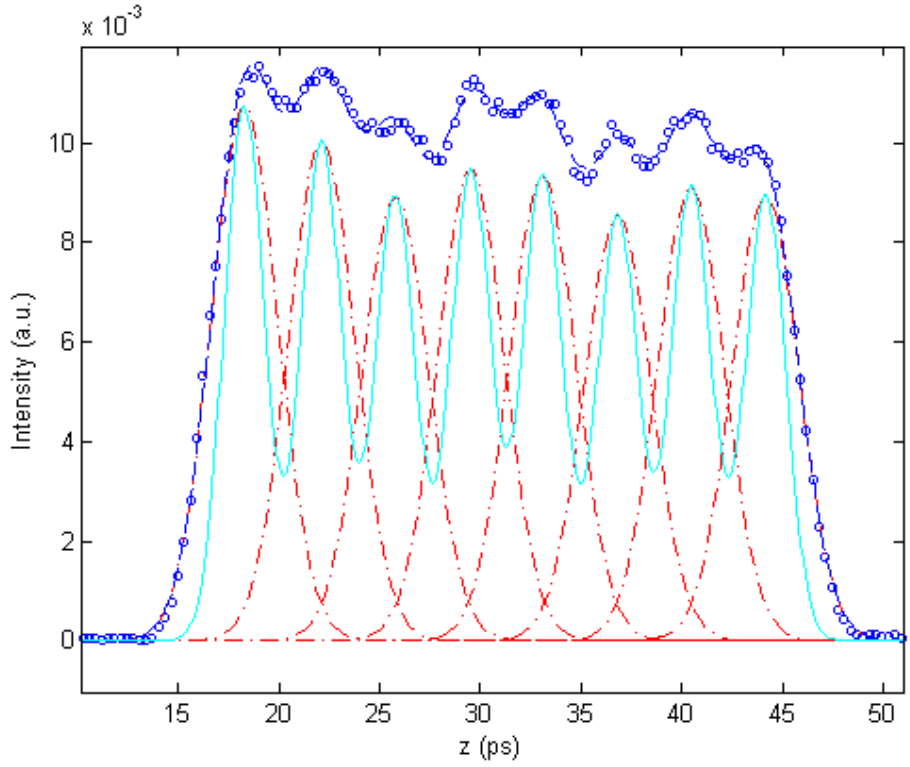

Figure 7.5: Measured temporal profile (blue marker), 8 Gaussian fit (blue line), individual Gaussian distributions (red line), actual profile (cyan)

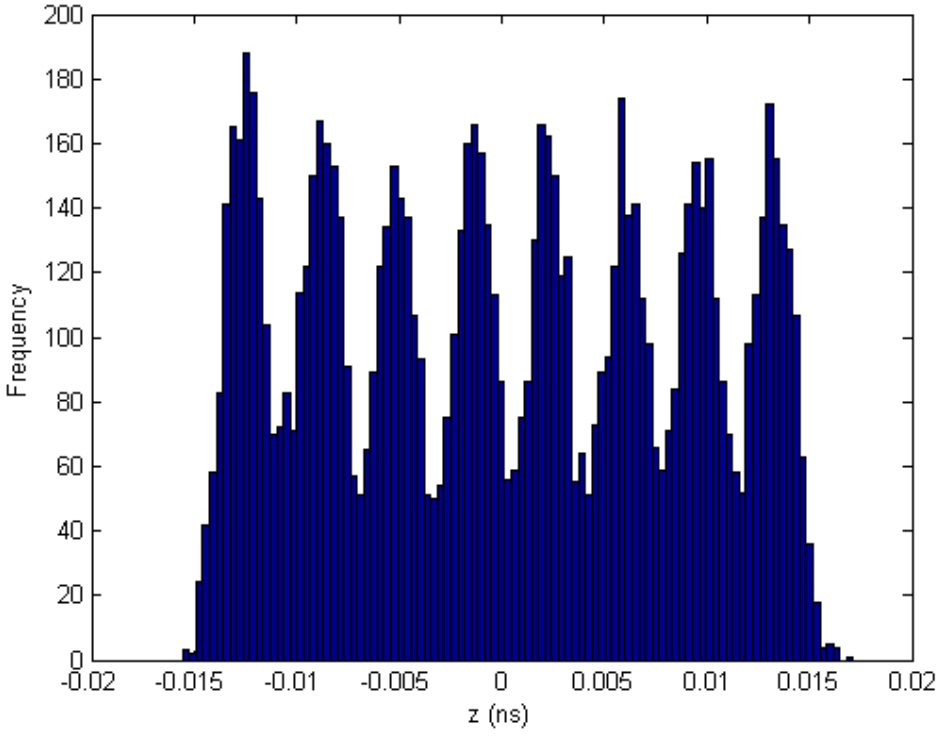

Figure 7.6: Longitudinal profile histogram of 10,000 particles

7.3 Measurement Results

The phase space measurement of the 0.5pC electron bunch is shown in figure 7.7. Using the SCUBEEX method described in chapter 6 to estimate the emittance from the phase space, the thermal emittance is calculated as $0.343 \pm 0.002\mu$m, which is within 10% of that measured previously. This value was used in the simulation of the higher charge cases.

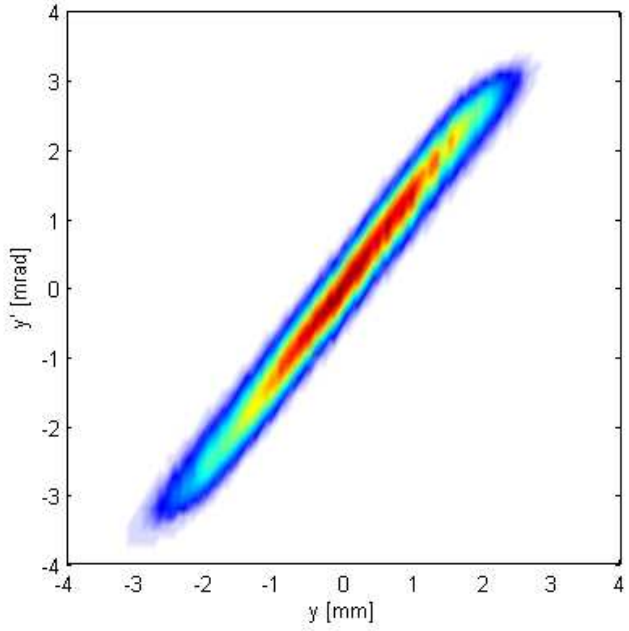

Figure 7.7: Measured vertical phase space for a 0.5pC electron bunch, solenoid current = 3.7A

7.3.1 20pC Results

A comparison between the vertical beam size and normalised emittance at several solenoid settings are shown in figures 7.8 and 7.9 respectively. The simulated electron beam properties match the trend of that measured experimentally.

The phase space at 20pC cannot be well represented by an ellipse. Consequently the SCUBEEX method used previously would not provide a reliable result from the data. Instead a threshold of 0.5% was applied to the phase space image,

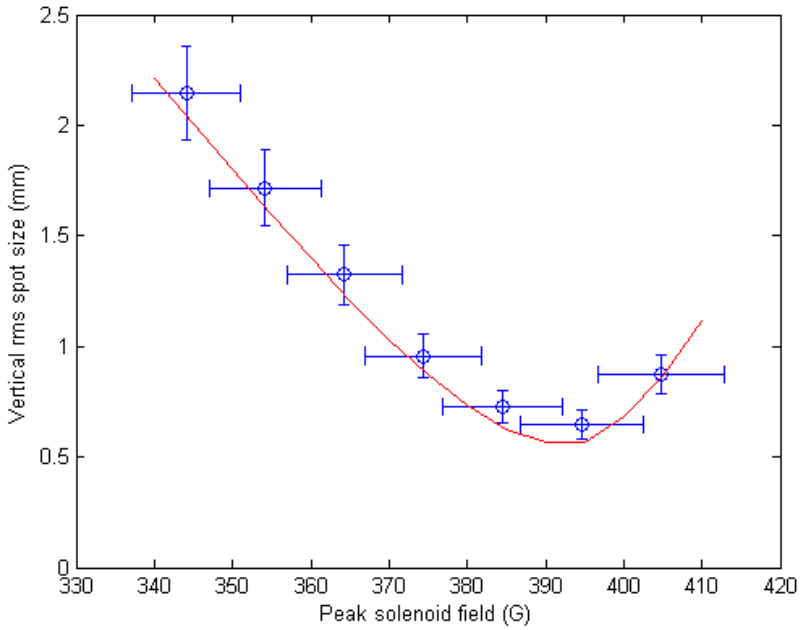

Figure 7.8: Vertical rms post size as a function of solenoid field (blue - data, red - simulation) for 20pC bunches

just enough to eliminate the majority of the background noise. The emittance and rms beam size were then numerically calculated from the modified image. The estimated error in the calculation is 10%. An additional error exists in the calibration from the current set in the windings of the solenoid and the peak field used in the ASTRA simulation. The static field solver POISSON gives this calibration factor as 103.1G/A. The measured peak field was within 2% of this value. The position of the minimum in the simulated beam size depends quite critically on the calibration as the measurement range is so small. By minimising the difference between the measured data and the simulated data using the calibration as a variable, gave the best value at 101G/A which is within 2.1% of that predicted by the model.

Figure 7.10 shows a comparison of measured versus simulated phase space at six solenoid settings. The double slit and Faraday cup method was used to measure the phase space. The streaking features that are visible in the phase space

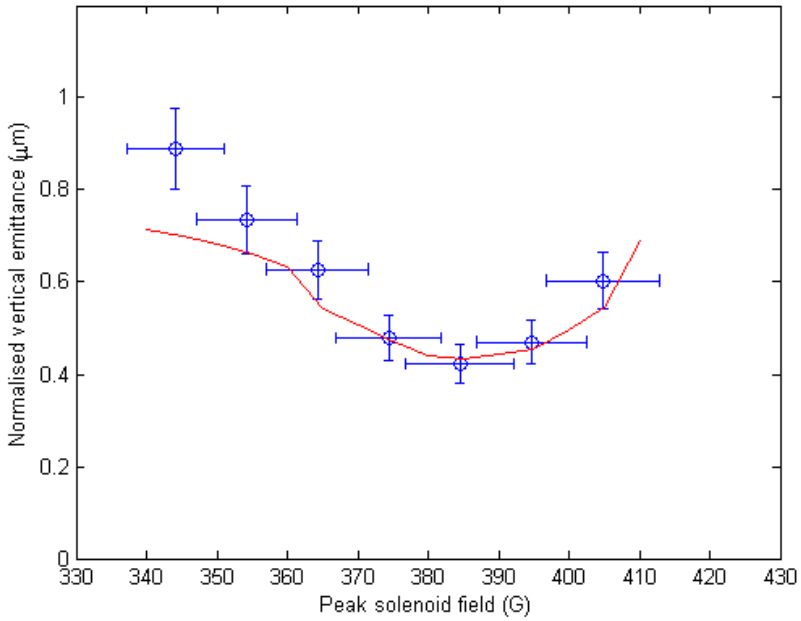

Figure 7.9: Normalised vertical emittance as a function of solenoid field (blue - data, red - simulation) for 20pC bunches

originate from the laser spot jitter. For illustration purposes the simulated particle distribution is blurred, as described in section 6.3.1, to produce an image in the same format as the measurement. Qualitatively the simulation represents the measured vertical phase space well.

Figure 7.10: Simulated (left) and measured (right) vertical phase space for 20pC electron bunches at 1.244m from the cathode with different solenoid settings

7.3.2 80pC Results

Figures 7.11 and 7.12 show the comparison of rms beam size and emittance measurements as a function of solenoid strength. The simulated beam size is

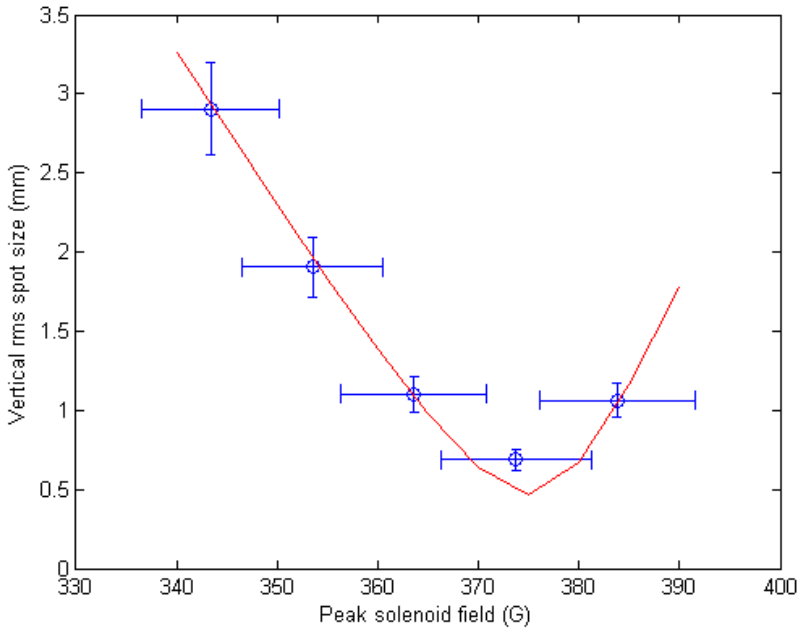

Figure 7.11: Vertical rms post size as a function of solenoid field (blue - data, red - simulation) for 80pC bunches

within the experimental error of that measured, however the simulated emittance does not represent that calculated very well. This could be due to the large error introduced in estimating the emittance from such noisy data. The vertical phase space, shown in figure 7.13, is dominated by the streaking effect from the laser and at higher solenoid fields it becomes more asymmetric. The simulation more accurately represents the data at lower solenoid field strength.

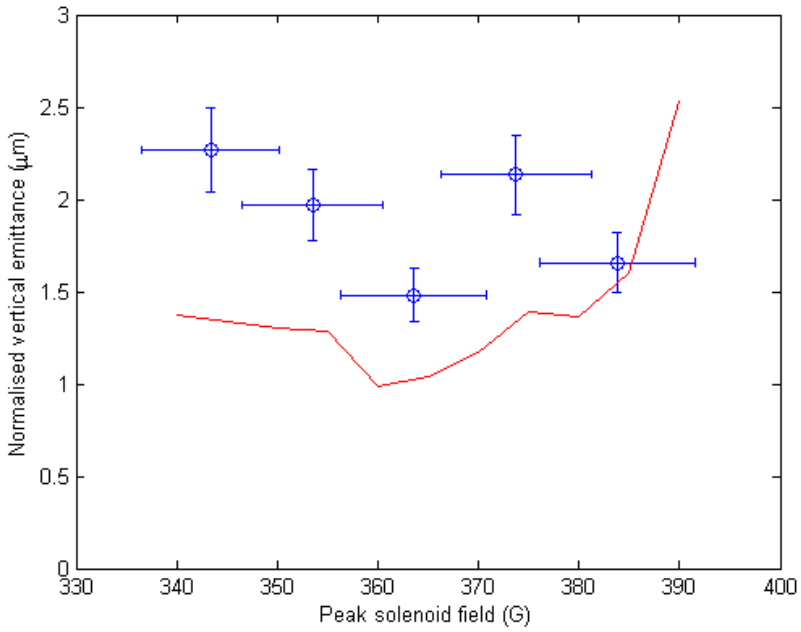

Figure 7.12: Normalised vertical emittance as a function of solenoid field (blue - data, red - simulation) for 80pC bunches

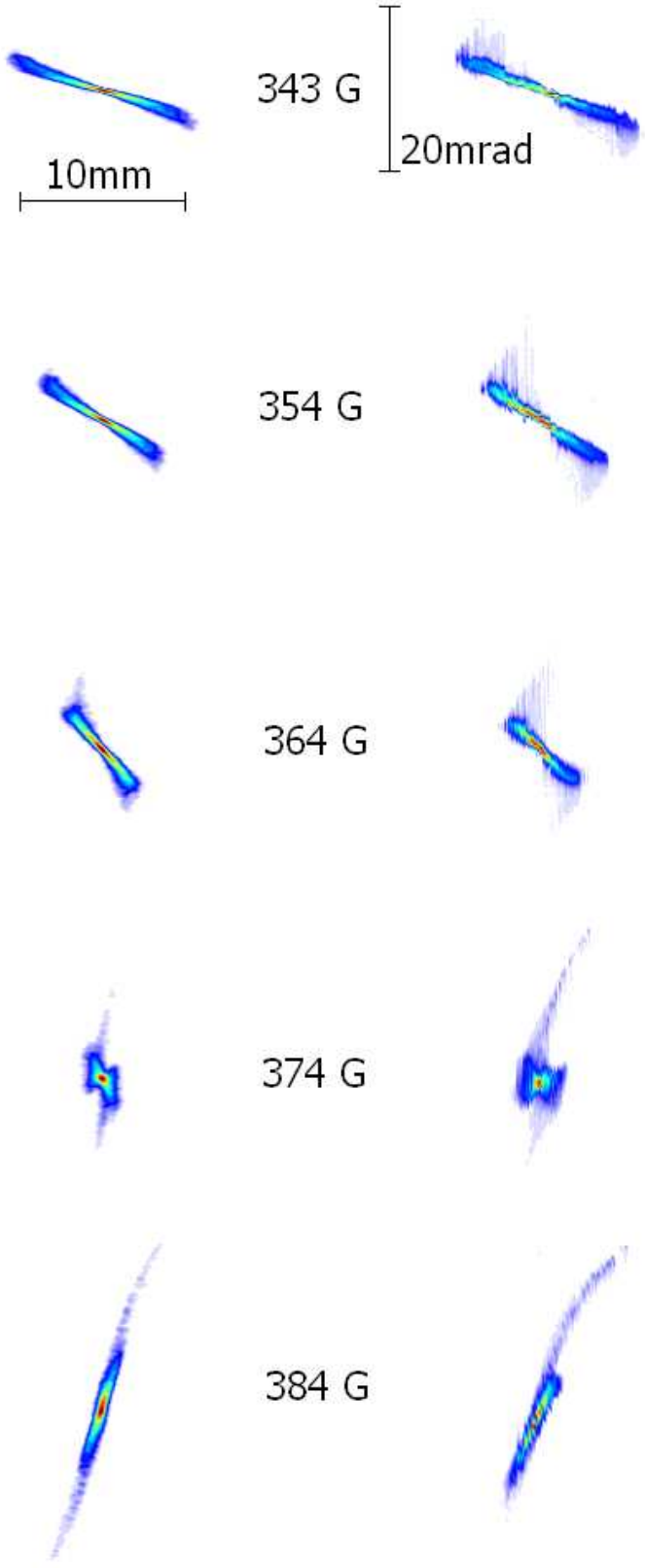

Figure 7.13: Simulated (left) and measured (right) vertical phase space for 80pC electron bunches at 1.244m from the cathode with different solenoid settings

7.4 Conclusions

The agreement between simulation with ASTRA and measurement in this experiment was good. Despite the fact that the 2D space charge algorithm was required because of particles being launched from a cathode, the phase space qualitatively represented the data well. The asymmetry in the laser distribution manifests in a non-symmetric phase space, which is also seen in the simulation. The discrepancy in the emittance comparison with 80pC electron bunches could be attributable to the difficulty in calculating the emittance of such noisy data. This result is consistent with [19], where the 3D simulation codes only make a slight improvement in describing the 80pC case.

In the following chapters ASTRA is used to model the behaviour of the JLab/AES injector, where space charge forces are noticeable due to the low energy, as with the injector used in this experiment. Furthermore, the injector simulation is of an idealised situation. The electrons emitted from the cathode will be from a perfectly symmetric distribution, and the longitudinal profile will have no ripples, which should improve the accuracy. The results of the previous sections demonstrate that ASTRA can be appropriately used in this situation and be relied upon to give a realistic insight into the behaviour of the injector and electron beam properties.

CHAPTER 8

The JLab/AES 100mA Injector

8.1 Design and Layout

As the DC photo-injector appears a promising candidate for producing high CW average current, a joint venture between JLab and Advanced Energy Systems (AES) was undertaken in 2003 to design a high current injector. The initial design philosophy was to use a DC photoinjector closely coupled to a SRF booster. The principle being, that by accelerating the electron bunches to relativistic energies as soon after the gun as possible, the degrading effects of space-charge forces within the bunches should be reduced. Of course problems can arise when trying to accelerate high average current. For example, the SRF cavities of the booster must be capable of accelerating 750mA without beam break up. For this reason, low frequency (748.5MHz) single cell cavities were used. Lower frequency cavities increase the current threshold for the onset of beam break up [117]. With single cells the degrading higher order modes excited in the cavities can be extracted more easily, leaving only the fundamental accelerating mode. Any additional modes can introduce asymmetry into the electron bunch and degrade the transverse emittance.

Originally the JLab/AES injector was designed with two distinct modes of operation envisaged; firstly 100mA CW running with 133pC electron bunches at the fundamental RF frequency, and secondly, up to 750mA using 1nC bunches.

The final injector design consists of a 500kV DC gun, electrostatically identical to that used in the JLab FEL facility. The gun is followed by an emittance compensation solenoid which focuses the electron beam into a 748.5 MHz SRF cryomodule. The module consists of three single cell cavities with the specification to accelerate the beam to 7MeV and a 3rd harmonic cavity used for correcting some longitudinal non-linearity and, to some extent, bunching. This injector design differs from those at Cornell, Daresbury Laboratory and the JLab FEL which all include a dedicated buncher cavity to reduce the bunch length. The traditional layout focuses the electron beam out of the gun into a buncher cavity using the solenoid. A second solenoid then matches the beam into a booster cavity, where the energy is raised.

The evolution of the early design has been described previously [118, 16, 119], and final changes to the component positioning were not modelled. This resulted in the requirement for the re-optimisation and simulation recorded in the following chapter 9. The injector has undergone many transformations from its initial conception. The first design used a new DC gun design, with a short cathode-anode gap, allowing for a higher electric field on the cathode and thus more initial acceleration. This was followed by a solenoid and then seven single cell SRF cavities. Through many iterations, the gun design was changed to use a copy of the JLab FEL gun, only operated at 500kV rather than 350kV. Further alterations occurred as a result of funding issues, whereby the seven SRF cavities were reduced to three, and the project was merged with another, which was to demonstrate the use of a 3rd harmonic cavity. The layout of the final design is shown in figure 8.1.

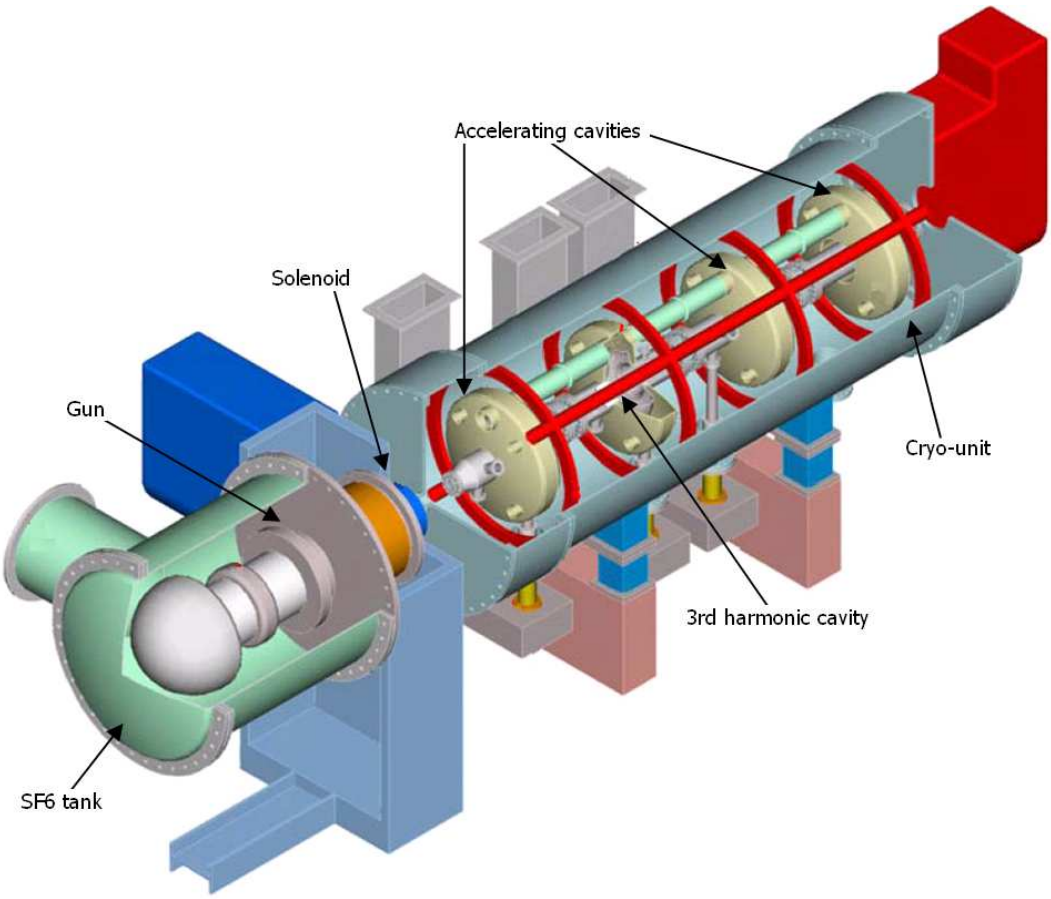

Figure 8.1: Layout of the JLab/AES Injector

The design of the original seven cell injector was constrained to meet a specification defining the limits on electron beam properties. This specification is shown in table 8.1 alongside those currently achievable in the JLab FEL injector.

	Low Charge	High Charge	Existing FEL
Bunch charge (nC)	0.133	1	0.122
Bunch rep. rate (MHz)	748.5	748.5	75
Average current max. (mA)	100	750	9.1
Gun energy (keV)	500	500	350
Injector energy (MeV)	7	7	9
Transverse norm. emittance (μm)	5	5	10
Longitudinal norm. emittance (keV mm)	15	45	9.5
Bunch length (mm)	1.8	1.8	1.8
Energy spread (%)	0.15	0.7	0.5

Table 8.1: Specification for the JLab/AES injector

The original (short cathode-anode gap, seven cell) injector was modelled by AES using PARMELA and was shown to meet the specification in both high and low charge cases [120]. Manufacture of the injector began in 2005 before the final layout (long gun, three cells and 3rd harmonic cavity) had been modelled, and no results were available that would meet the specification in all aspects. In 2008 the gun was assembled and conditioned to 500kV. The maximum bunch charge achieved to date is 1nC (10nA), however the current is presently limited to 10μA because of the permissable radiation limit. The cryo-module has been constructed but RF testing has not been completed. Chapter 9 describes the complete modelling of the final JLab/AES injector design and its performance compared to that of the specification.

The JLab FEL injector routinely demonstrates that it can deliver 1nC electron bunches, but both the maximum repetition rate of the laser, and the high voltage (HV) power supply limit the maximum current that can be drawn from the gun. Increasing the current available from the gun is not limited by the technology of the laser or HV supply, as the requirements are within the bounds of that available. Adverse effects may occur by increasing the repetition rate of the laser, as this increases the power loading on the cathode and hence the temperature rises as more laser power is absorbed. This in turn will increase the thermal emittance of the emitted electrons. Heating inside the gun chamber will also serve to degrade the vacuum quality which subsequently will destroy the quantum efficiency of the cathode. With a poor QE, more laser power is required to deliver the same bunch charge. In this way it is possible to envisage a situation that would progressively worsen. The JLab gun design has no active cooling of the cathode in the design at present, but a modification may be necessary if the temperature rise due to the incident laser is too much.

Heating is not the only challenge with the electron gun. A key design change is the operation at 500kV. At JLab the existing gun has been conditioned to 420kV,

at which point field emitters on the HV surfaces draw current from the power supply and degrade the vacuum. Careful preparation of the HV components is required to avoid being limited by field emission.

Individually considered, the beam parameter specifications are not overly demanding. Smaller emittances, both longitudinally and transversely, shorter bunch lengths and lower energy spreads have been achieved in other machines. These have not been operated in a continuous electron bunch mode at high average current. They are pulsed machines, and as such, are not subject to the limitations of CW running. The difficulty in meeting the specification with this design, as will be seen in the following chapter, is the limited number of components. Trying to simultaneously achieve all parameters will prove problematic with so few degrees of freedom.

CHAPTER 9

Simulations of the JLab/AES Injector

This chapter discusses the modelling and optimisation of the JLab/AES injector. Given that the physical layout of the injector was fixed before being modelled, only the settings of the individual component fields and initial electron distribution could be changed.

The electron beam was first manually modelled through the injector using the particle tracking program ASTRA. Then an optimisation program was applied to the problem to improve performance further. Both operating modes were investigated, and the final working points compared to the specification.

9.1 Methodology

The JLab/AES injector consists of seven physical components; the laser, gun, solenoid, three accelerating cavities and a 3rd harmonic cavity. With so few components, it is realistic to try and achieve a design-point solution by optimising the set-points of each component and simulating the layout by hand. This was the method originally applied to solve this problem. The injector was modelled in a piecewise fashion. A good setting was found for each component

in the order in which they are experienced by the electron bunch, starting first with the gun and solenoid.

To begin with the best case scenario was modelled by making the following assumptions: the injector was assumed to be axially symmetric, implying that the electron bunches would be launched on-axis, and all the electric and magnetic fields that act upon the bunches would be cylindrically symmetric. The higher order effects (beyond 3rd order) from these fields off axis were also neglected in preliminary modelling. In modelling the emission of electrons from the cathode it was assumed that the electron bunch properties would directly mimic the laser shape, i.e. the emission is prompt and quantum efficiency uniform. In addition, no thermal emittance was included, so that the initial transverse momenta of the electrons were zero. By excluding these features computation time was reduced. The effect of including these real world attributes into the simulation is discussed later.

The first iteration of modelling was performed using the space charge tracking code ASTRA, with the on-axis field maps of each component being derived from either POISSON or SUPERFISH [76]. To show the extent of the fields and how they overlap, the superimposed normalised field maps as a function of distance from the cathode are shown in figure 9.1.

Note that the solenoid field extends into the gun chamber, overlapping with the electric field from the gun. The solenoid is placed physically as close as possible to the exit of the gun to compensate for the emittance growth due to the space charge forces within the electron bunch. The space following the solenoid, between it and the first accelerating cavity, is occupied with essential vacuum components and also the end can of the cryo-module. Within the cryo-module the fields of cavities 3 and 4 overlap, but as the field level at that point is less than 1% of the maximum, any cross-talk between the cavities is negligible.

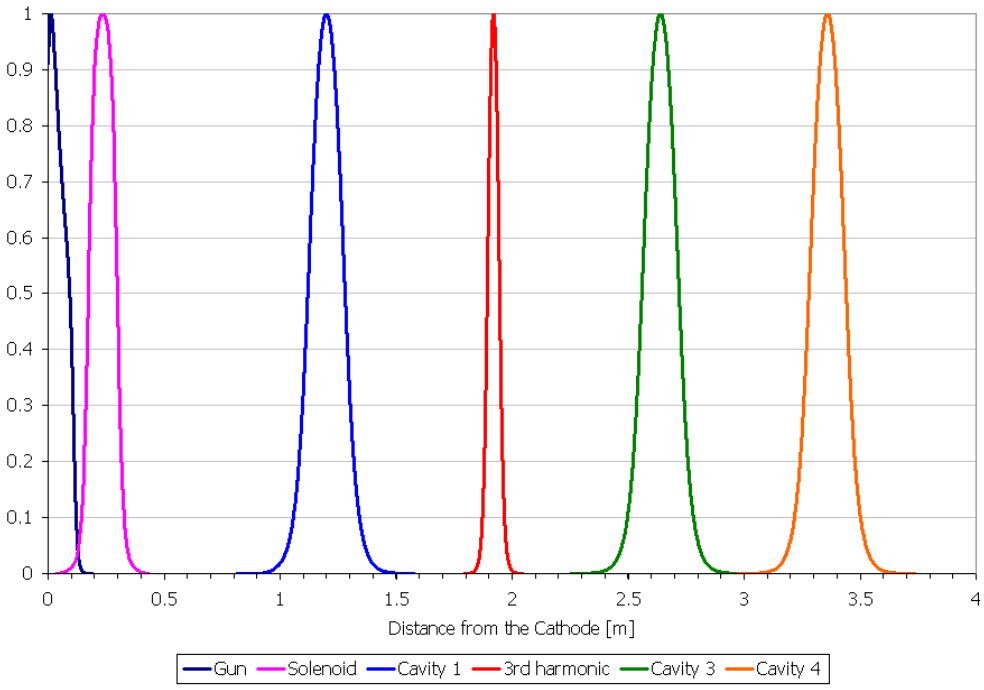

Figure 9.1: Normalised On-axis Field Maps as a Function of Distance from the Cathode

The optimisation philosophy was to achieve small transverse and longitudinal emittances at the end of the injector, and to try and preserve a linear energy chirp along the bunch length, to allow for further bunch compression in a magnetic injection chicane. The injector was set up one component at a time: beginning with the gun and solenoid, then the first accelerating cavity and so on. Between each pair of components the beam properties for different settings were monitored to find the best point for operation.

The starting conditions for the electron beam emitted from the cathode affects how the bunches evolve downstream. In reality the emitted bunch depends on the character of the laser and the emitting properties of the cathode material. The DC gun will operate with a caesiated GaAs photocathode. The results from the Cornell measurements on the response time of GaAs show that it can be considered a prompt emitter if used with a laser with a 520nm wavelength. Initially the injector will be tested using a 527nm laser, so the longitudinal electron distribution can be assumed to follow the laser temporal shape. Transversely

the profile was modelled as being radially uniform, and has been assumed to have a 5mm diameter (rms 1.25mm). The longitudinal profile is Gaussian with a rms time of 20ps. This is similar to the properties of the laser used in the JLab FEL injector [121, 122].

9.2 Manual Set-Up Results

The 135pC, low charge scenario was modelled first, as the space charge effects are smaller than with a 1nC bunch, which means that there is greater flexibility when optimising the component parameters.

The electrons are launched from a cathode inside the gun cavity, and the bunch immediately expands both transversely and longitudinally under space charge forces, see figure 9.2. The solenoid positioned directly after the gun focuses the

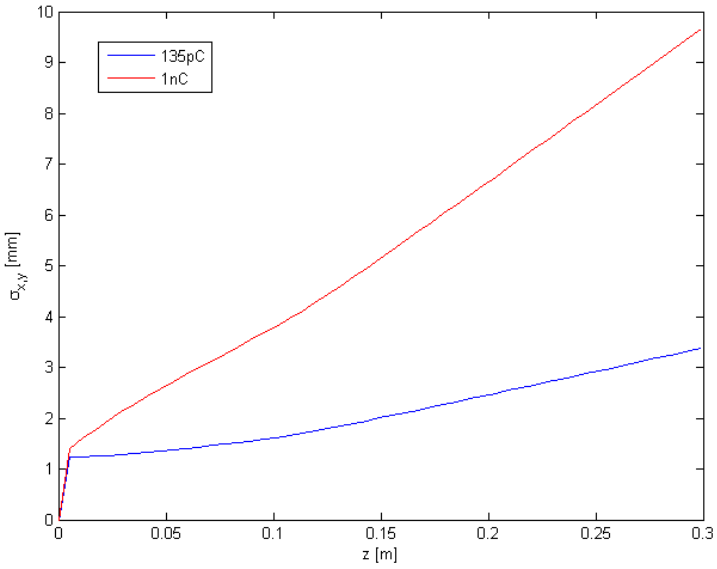

Figure 9.2: Transverse beam size as a function of distance from the cathode

bunch transversely into the first accelerating cavity, whilst the bunch length is still increasing. Despite the focusing in the transverse plane, the bunch length is not heavily dependent on the strength of the field in the solenoid (in this case

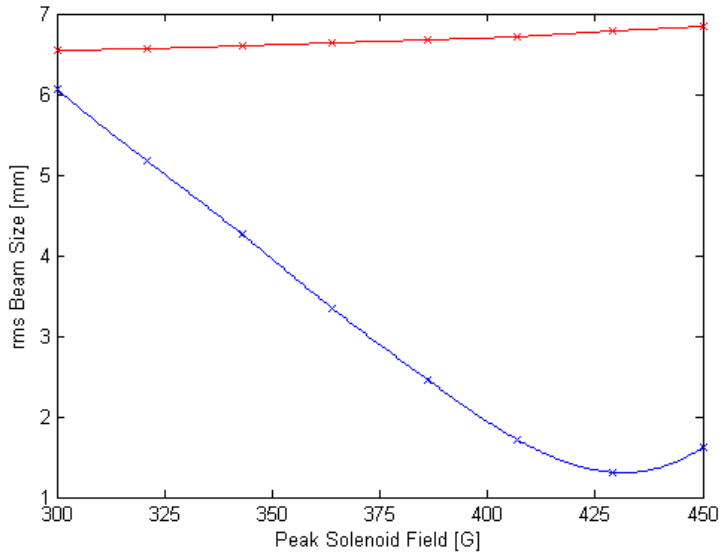

Figure 9.3: Transverse (blue) and longitudinal (red) rms beam sizes as a function of solenoid field strength at 1.2m from the cathode for 135pC

because the bunch is long), so it enters the first cavity with an rms length of $\sim$6.5mm, see figure 9.3.

In the absence of a dedicated buncher cavity the first accelerating cell must be used to impose further bunching while the beam is still malleable at low energy. Velocity bunching in RF cavities is more pronounced the lower the energy of the electrons. The phase of the RF at which the electron bunch enters, determines whether it is accelerated, decelerated, compressed or expanded. This is shown in figure 9.4. An electron that enters at any phase where the accelerating field (E_z) is positive will be accelerated. The maximum acceleration is at phase 'b' for relativistic particles. Electron bunches that experience the field at position 'a' have a zero net energy gain, but are bunched. The compression occurs as the head of the electron bunch is actually decelerated, whilst the tail is accelerated. The variation in energy along the length of the bunch causes it to become compressed as it progresses along the beamline. Care must be taken not to over-compress the bunch, as once the electrons in the tail have overtaken the head, it is impossible to reverse the process. To reduce the bunch length from 6.5mm at the entrance of the first accelerating cell to the specification of 1.8mm at the end

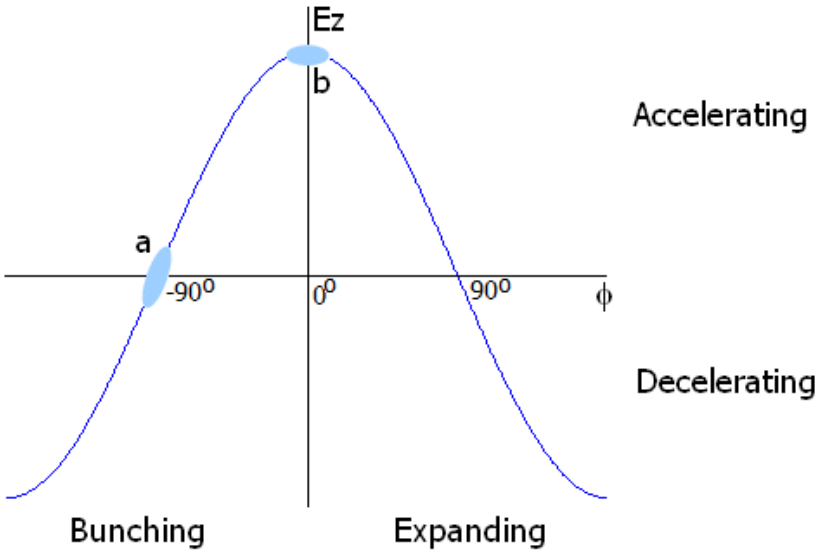

Figure 9.4: Schematic of how energy is imparted to the electron bunch from the RF field

of the injector, strong bunching must take place in the first accelerating cavity. To achieve this, the electron bunch is simply injected at a negative phase with respect to the maximum energy gain phase. A phase of 0 and -90 degrees in the simulation corresponds to maximum energy gain and maximum bunching respectively. Additionally, it is also necessary to accelerate the electron bunch in the first cell as there are only 3 cavities available to increase the energy to the design value of 7MeV. This means that the phase chosen should be between 0 and -90 such that both acceleration and bunching occur. A larger gradient gives more acceleration (and bunching for a given phase), however, the energy spread and non-linearity in the bunch is also increased (due to the sinusoidal nature of the RF field), and so longitudinal emittance grows. Recovering the longitudinal properties in the following components is complex. Compression of the electron bunch longitudinally at low electron energies results in space charge blow-up transversely. At non-relativistic energies, the coupling between the longitudinal and transverse phase space is pronounced. The bunch properties are very sensitive to changes in the cavity gradient, phase and solenoid field strength.

Varying the solenoid strength changes the transverse dimensions of the bunch entering the first accelerating cell. Rather than focus the bunch to a minimum before entering the first cell, a modest setting is chosen so as not to increase the bunch length once in the accelerating cavities. Further transverse compensation can be achieved through RF focusing in the cavities. That is, given the correct phase, the off axis electrons are bent towards the axis due to the transverse RF electric field components off axis, shown in figure 9.5.

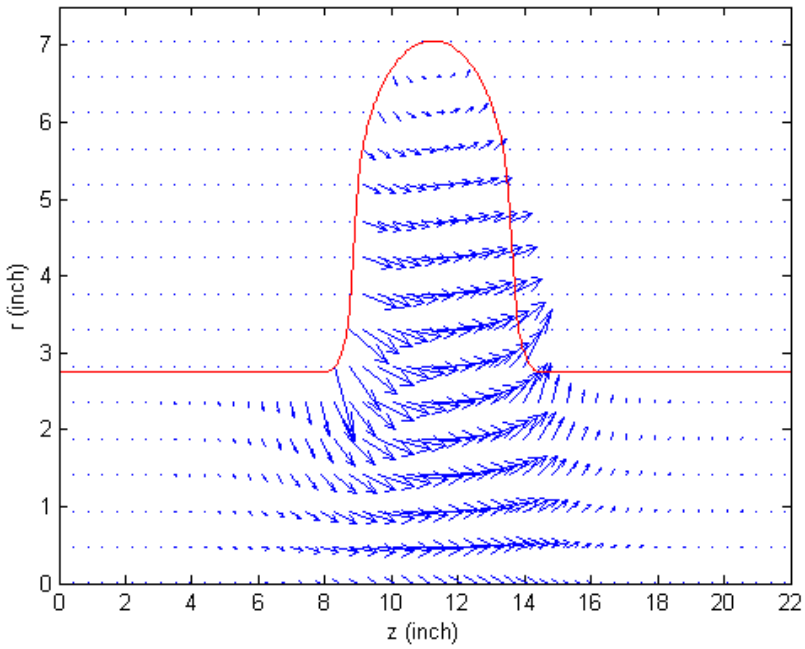

Figure 9.5: Vector plot of the off axis electric field in the accelerating cavity. Geometry (red)

The 3rd harmonic cavity is normally used to restore some linearity to the longitudinal phase space. This means that the energy an electron has is directly proportional to its longitudinal position within the bunch. The electron bunch is injected into the 3rd harmonic cavity with a negative phase without decelerating the bunch (-90 $< \phi <$ 0). The more negative the phase used, the shorter the bunch length and longitudinal emittance, but also the less energy gain. There is of course a trade-off with transverse emittance growth which increases with decreasing bunch length. It is not possible to operate the 3rd harmonic cavity with

a positive phase as this creates a transversely convergent bunch, again from off axis electric fields. This combined with the RF focusing in the following two accelerating cells results in transverse cross-over and non-laminar flow.

The final two cavities are used to accelerate the bunch to relativistic energies ($\beta \sim$0.9) before it leaves the cryo-module. The RF focusing in these two cavities, when the bunch is injected near the maximum accelerating phase, is strong. The phase space linearity of the resultant electron bunch is consequently highly dependent on the relation between the 3rd harmonic and final accelerating cavity settings, both longitudinally and transversely.

The evolution of the electron bunch through the injector for the first-pass modelling can be seen in figure 9.6. The progression of the transverse horizontal phase space after each component is shown in figure 9.7, and remains linear.

The results shown have a final mean energy of just less than 5.5MeV, which is below the design energy, specified at 7MeV. The transverse properties of the bunch meet the specification, but the longitudinal emittance is over twice that specified. Manually trying to increase the exit energy, whilst maintaining modest emittances and bunch lengths, proved problematic due to the complexity of the interactions between the components. The process of finding one operating point for the injector by hand was very time consuming, despite this being the easier of the two operational modes (lower space charge). The next section investigates an improved method of finding the best configuration of the injector given all the conflicting specifications for electron bunches of any charge.

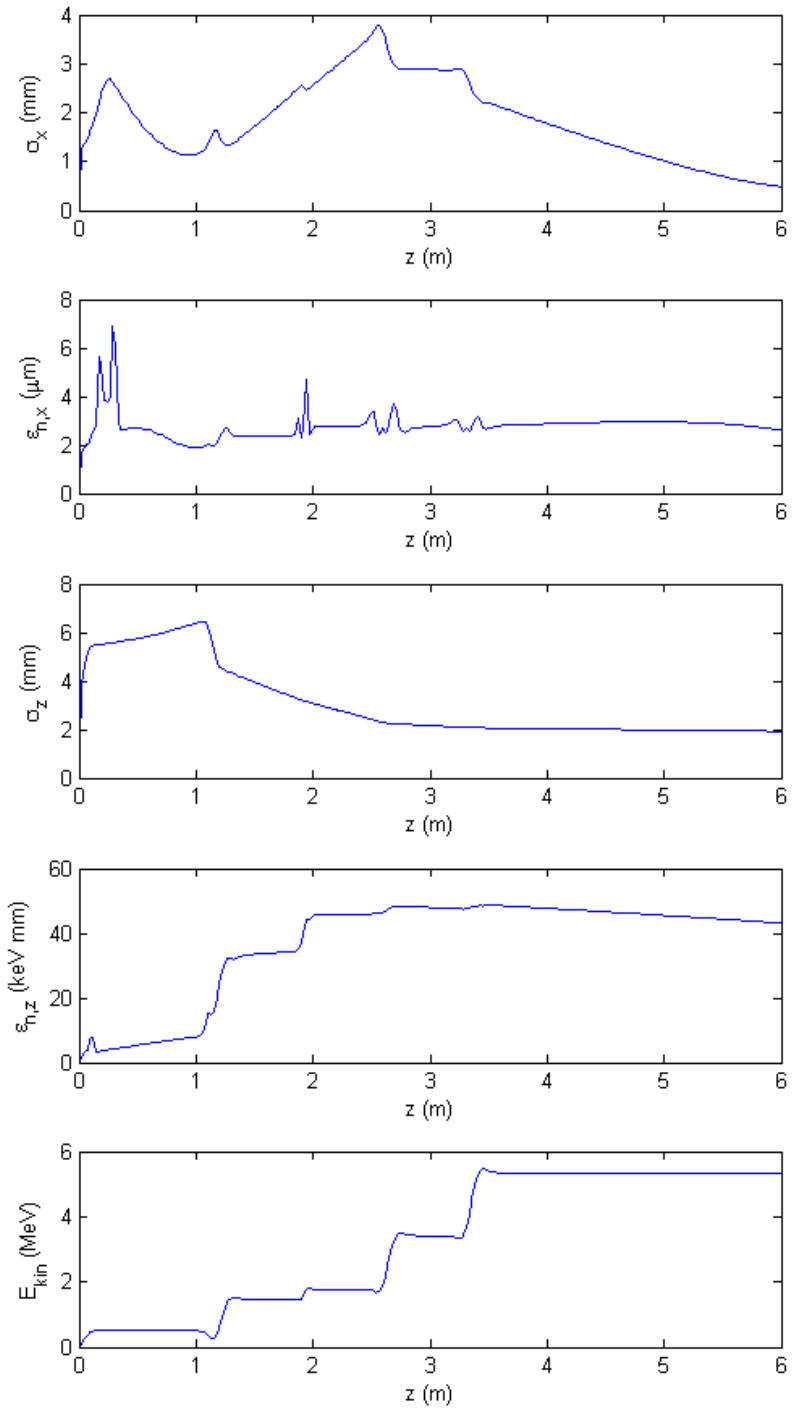

Figure 9.6: Bunch evolution as a function of distance from the cathode. From top to bottom σ_x, $\varepsilon_{n,x}$, σ_z, $\varepsilon_{n,z}$,E $_{kin}$

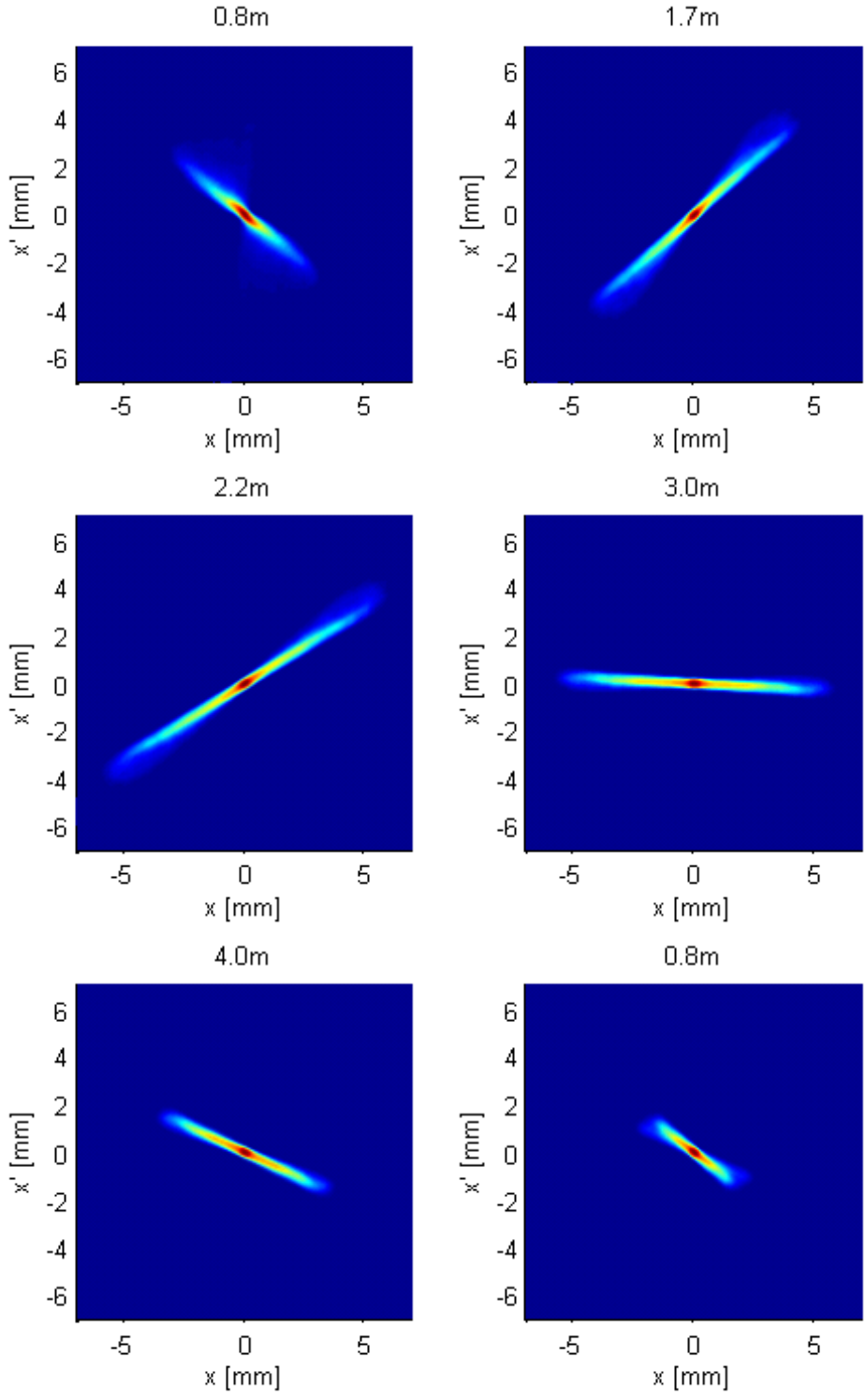

Figure 9.7: Transverse phase space as a function of distance from the cathode

9.3 Optimisation Algorithm

In an effort to further improve the performance of the AES/JLab injector, the multivariate optimisation program, that was first used at Cornell University, was applied to this geometry. The details regarding the use of an evolutionary algorithm to explore the parameter space and find optimal solutions are covered in depth in [123]. At Cornell, the optimisation was used to design an injector from scratch. Applying this technique to the JLab/AES injector differs in that the layout is already fixed, and the optimisation is used to find the capabilities of the predefined design. The problem of finding the optimal set up for the JLab/AES injector which meets the specification is well suited to be solved through an optimisation algorithm. As it becomes difficult to visualise how the electron bunch changes under the influence of each component in the injector, whilst trying to meet a specification and remain within constraints, having this process quantified and automated makes it possible to find a true optimum.

9.3.1 Evolutionary Algorithms

Evolutionary or genetic algorithms are so called because of their close parallels with the theory of biological evolution, using techniques inspired by crossover, mutation, selection, and inheritance. In this way, from a population, those members that are better in some way are more likely to be selected and preserved to the next generation (inheritance). Ultimately an optimum set is reached after a number of generations; where all members of the population are equally good (Pareto-optimum set). Pareto solutions are those for which improvement in one objective can only occur to the detriment of at least one other objective [124]. In a similar way to human genetics, crossing (combining) and mutation are introduced at each generation, that is, the properties of more successful members are combined and varied to create the next population generation.

Optimisation algorithms, in favourable situations, save the user from conducting an exhaustive study of the entire parameter space. When the number of parameters or variables is large this quickly becomes impractical. Many algorithms will only consider one solution, and by comparing it with the previous solution, decide which direction to take with the following trial (a gradient technique). The disadvantage of this method is that it is common to find a local, rather than the global optimum. The starting conditions for the optimisation are usually chosen randomly, so the variable space is not necessarily fully sampled. By contrast, genetic algorithms maintain a pool of solutions rather than just one. The next trial solutions are not computed through a gradient method but by introducing crossing and mutation at each generation. In this way it is more likely that the full variable space is sampled.

All genetic algorithms follow a similar sequence in reaching a set of optimum solutions. Initially a population of trial solutions are generated with each population member having a set of randomly assigned variable values within some bounds. The objective function and constraints are evaluated for each solution. A fitness value is assigned to each member, dependent on how well it compares to the objectives, constraints or other solutions. A fixed number of the solutions are selected for a mating pool. This pool then contains the parent solutions for the next generation. The selection is biased towards those solutions with a better fitness value, and so they are represented in the pool more than poorer solutions. A crossing operator then creates a new generation (offspring) of trial solutions from the mating pool by taking variable settings from one solution and combining with another. Then a mutation operator can be used to diversify the new generation by introducing some spread into the variable settings. In this way the same values are not identical in every generation. The offspring solutions are then calculated and they then become the parents for the next generation, and so on.

Injector design is rather time intensive as the computation time for each simulation can be long (hours, depending on complexity), even with the faster codes. These simulations then need to be repeated for many different conditions. Parallel processing can ease the burden somewhat, by computing a number of solutions simultaneously. Evolutionary algorithms naturally lend themselves to the use of parallel processing, as a population of trial solutions can be calculated each on an individual processor. Only after each trial solution has been computed is there a need to compile the information for selection for the mating pool, crossing and mutation. The optimisation program developed at Cornell used this procedure. The algorithm was based on a modified version of PISA [125] (Platform and programming language independent Interface for Search Algorithms), developed at the Institute of Technology, Zurich (ETH) [126]. PISA separates the optimisation problem into two modules. One part, called the *variator*, contains those things specific to the optimisation problem. For example evaluation and variation of the solutions. The second module, the *selector*, contains those parts of the problem that are independent of the problem; mainly how members are selected for the mating pool. The variator module creates an initial population of members and calculates the objective values of each. The selector module then chooses a collection of parent individuals which, by comparison, are promising. The variator then introduces some diversity into these individuals in order to get a new population of offspring. The selection module again chooses the parents and so on. A benefit of the program being partitioned in this way is that the selector module can be interchanged, so that different evolutionary algorithms can be used. PISA has a suite of optimisation algorithms that can be used for a variety of problems.

The SPEA2 (Strength Pareto Evolutionary Algorithm 2) was successfully used at Cornell, and so was utilised for the JLab/AES injector design also. SPEA2 [127, 124] is an elitist multi-objective evolutionary algorithm. Elitist algorithms allow some of the better individuals to carry to the next generation unaltered.

An external population is set aside to contain the best non-dominated solutions from each generation. A non-dominated solution is one in which it is no worse in *all* objectives than another solution and it is strictly better in at least one objective. In this way the selection operator preserves the best solution from any generation in an archive so that there is always the possibility that it will be represented in the next generation. This implies that good solutions do not get lost, and they always have a chance in the mating pool. The size of the external population is fixed so that it does not become too large with many generations. Solutions from each generation are compared to those in the external population and any dominated solutions are replaced with the non-dominated ones. Additionally a clustering algorithm is used so that solutions in less crowded areas are kept, and those in clustered regions are discriminated against. This helps convergence of the solution towards an evenly distributed Pareto-optimum front.

9.3.2 Problem Definition

The Jlab/AES injector design had a total of 10 free parameters that could be varied; all cavity gradients and phases, except the gun, the solenoid peak field and the transverse rms laser spot size. Those that were kept constant included the component positioning and the bunch emission length from the cathode. Preliminary tests showed that optimised solutions converged at the practical limit imposed on the gun voltage from the available power supply (500kV), so this was kept constant.

Several constraints were applied to the optimisation to ensure sensible output and realistic solutions. In particular, the exit energy was constrained to be above 7MeV to meet the specification. Initial optimisation runs used a small number of macro particles (1000) to simulate the injector in ASTRA to reduce computation time. Once an optimised region had been found it was possible to then run a more detailed simulation to improve accuracy. The clusters on which the optimisation problem was run had 128 nodes available. The optimisation was

computed on one of 5, 128 node 3GHz clusters. For this reason, population sizes were a multiple of 128. Again, after the solution had converged, it was possible to increase the population size to show the extent of the Pareto-front. Run time on the cluster was limited to 24 hours, which limited the number of generations that could be calculated in one run, however the output of one run could be used as input into another. A flow diagram for the optimisation is shown in figure 9.8. Optimisations were limited to having 2 or 3 objectives, as more than this become difficult to visualise.

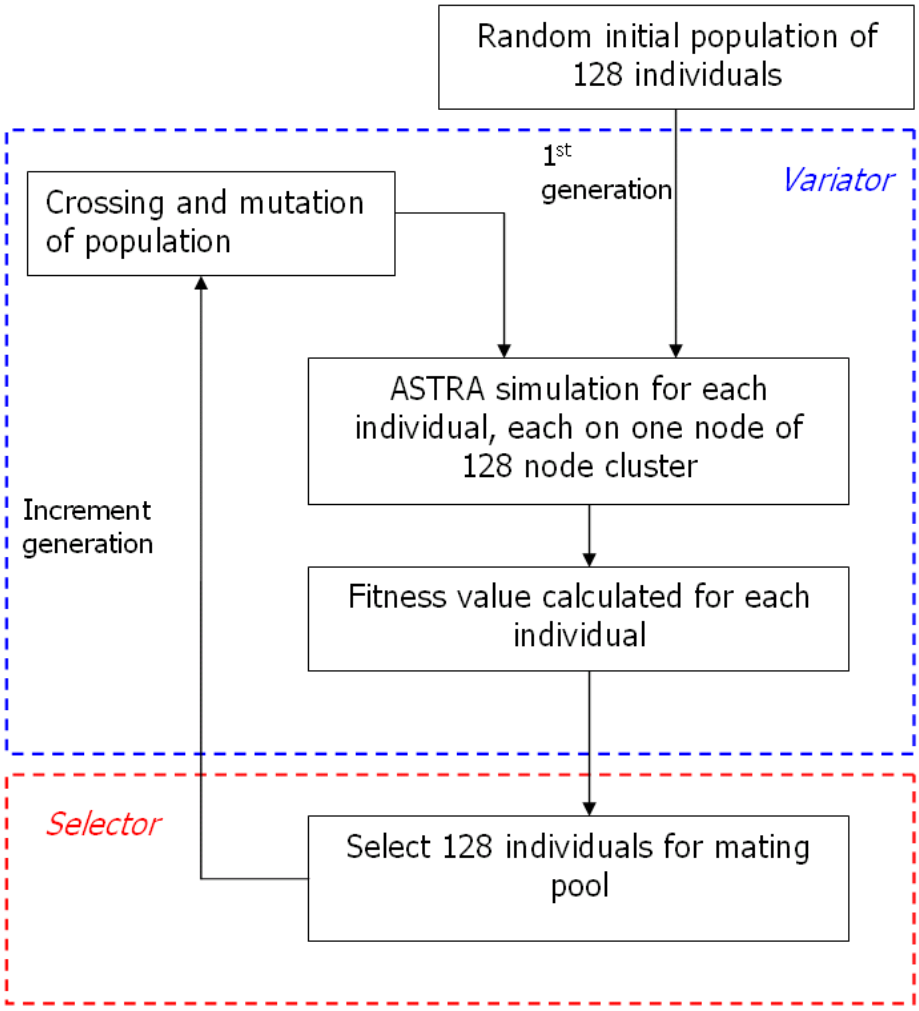

Figure 9.8: Flow diagram for the optimisation

9.4 Results

The aim of the first optimisation problem was to try and meet all criteria of the specification at 135pC, as this was not achieved through manual investigations. The objective was to minimise transverse and longitudinal emittance simultaneously. The variables and constraints for the problem are listed in table 9.1. The variables, particularly those for each cavity, were chosen to cover a realistic range of achievable values. The peak, on-axis gradient in the accelerating cavities was estimated to be 30MV/m, which relates to an accelerating gradient of about 16MV/m. The cavities have a specification to exceed this value. The phases were set to determine which regime each should be operating in. Test runs determined that the maximum gun voltage would produce better solutions, so this was fixed at the highest value of 500kV. The constraints were used to guide the optimisation in the correct direction, and to ensure that sensible beam properties were realised at the exit of the injector, 5m from the cathode. For this optimisation, the constraints were set to be larger than the specification to speed up the process of finding valid solutions, and because test runs indicated that the specification may not be met.

The compromise between longitudinal and transverse performance at the end of the injector is shown in figure 9.9, where each solution on the front is equally optimal. The figure shows that these solutions do not meet the desired values of longitudinal and transverse emittance, however they do all meet the energy requirement. A general trend found in the solutions from the optimal front was the electron emission area from the cathode was at the upper limit of $\sigma_r = 2$mm. When considering space charge forces alone, as in this simulation, a larger electron bunch has lower space charge forces, and so it seems sensible to maximise this variable. As the thermal emittance was not initially included in the simulation, and this is proportional to beam size at the cathode, the real world optimum for this variable may be to have a smaller spot size.

Variables	Min	Max
Transverse spot size rms (mm)	1	2
Peak solenoid field (T)	0.01	0.05
Max. gradient in cavity 1 (MV/m)	5	30
Phase of cavity 1 (deg)	-100	50
Max. gradient in 3rd harmonic (MV/m)	5	50
Phase of 3rd harmonic (deg)	-180	0
Max. gradient in cavity 3 (MV/m)	5	30
Phase of cavity 3 (deg)	-90	90
Max. gradient in cavity 4 (MV/m)	5	30
Phase of cavity 4 (deg)	-50	50
Constraints at 5m from cathode		
Bunch length rms (mm)	< 3	
Energy spread rms (keV)	< 70	
Mean energy (MeV)	> 7	
Longitudinal emittance (keV mm)	< 50	
Transverse emittance (μm)	< 10	
Transverse size rms (mm)	> 0.4	
Longitudinal correlation	Negative	
Transverse correlation	Negative	

Table 9.1: Variables and constraints for optimisation

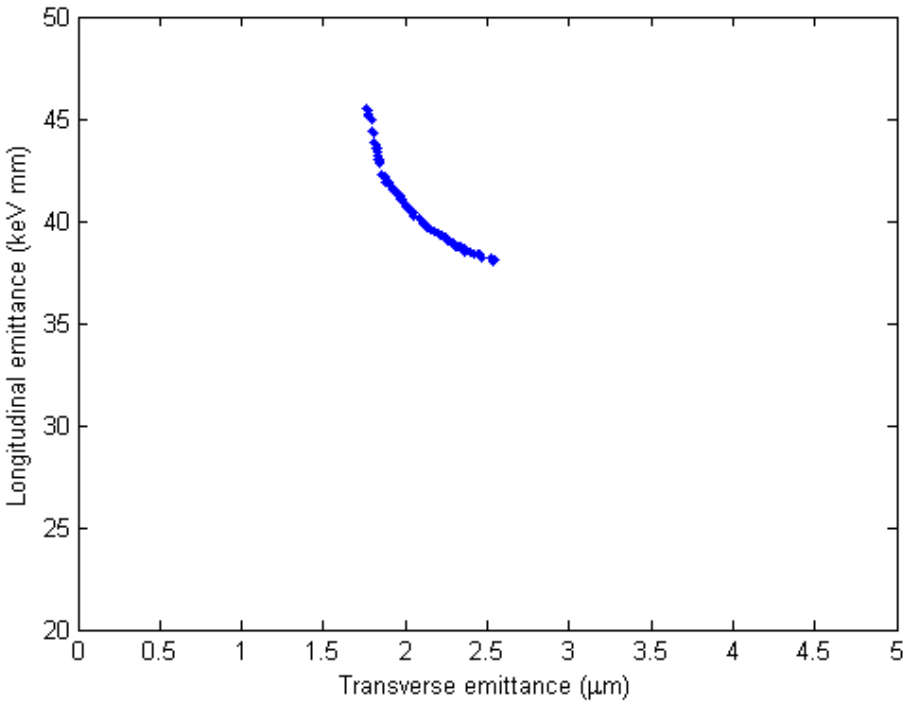

Figure 9.9: Two objective optimisation results for 135pC bunch

The range of values of each variable from all the solutions also gives an insight into the critical or sensitive parameters. The range and standard deviation of the variables for the optimum solutions are shown in table 9.2. For all solutions the

Variable	Max	Min	Avg.	std
Transverse spot size rms (mm)	2	2	2	0
Peak solenoid field (T)	0.039	0.038	0.039	0
Max. gradient in cavity 1 (MV/m)	23.80	23.14	23.41	0.15
Phase of cavity 1 (deg)	-26.72	-29.87	-28.80	0.87
Max. gradient in 3rd harmonic (MV/m)	16.72	7.05	11.96	1.94
Phase of 3rd harmonic (deg)	-70.90	-79.87	-75.13	2.54
Max. gradient in cavity 3 (MV/m)	30.00	29.99	30.00	0
Phase of cavity 3 (deg)	40.55	37.58	38.71	0.79
Max. gradient in cavity 4 (MV/m)	23.18	21.57	22.29	0.32
Phase of cavity 4 (deg)	6.56	-15.96	-9.07	5.13

Table 9.2: Range of variables for 135pC solutions

solenoid strength, first and third cavity gradient and phase are almost the same, the standard deviation being less than 2% of the average value. The average peak gradient of the second accelerating cavity is the maximum permitted. The phase of the first cavity is such that it is both bunching and accelerating for all solutions. The gradient of the third harmonic cavity is comparatively low, and the solutions with a higher gradient tend to have a lower transverse and larger longitudinal emittance. Increasing the phase of the third harmonic cavity has the same effect. This component affects the electron beam parameters the most and is largely responsible for the variation in transverse versus longitudinal emittance in the optimal front. Therefore it should be finely controlled during operation of the injector.

9.4.1 Two Objective Optimisation

The optimal fronts from optimisations of 135pC, 500pC and 1nC electron bunches are shown in figure 9.10. The constraint that the exit energy should be at least 7MeV was applied to each optimisation and met in all cases. The additional constraints applied to the longitudinal and transverse emittances are

given in table 9.3. The figure shows that as the bunch charge is increased so do the achievable lower limits of longitudinal and transverse emittance. Not one of the solutions meets the original specification in all aspects.

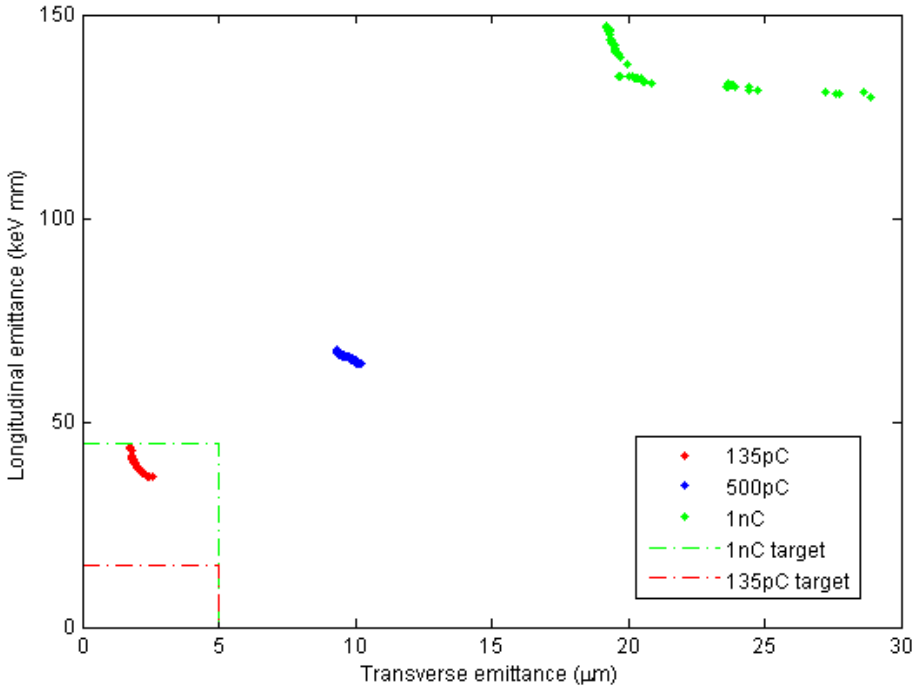

Figure 9.10: Two objective optimisation results

Constraints at 5m from cathode	500pC	1nC
Bunch length rms (mm)	< 3	< 3
Energy spread rms (keV)	< 100	< 100
Mean energy (MeV)	> 7	> 7
Longitudinal emittance (keV mm)	< 70	< 150
Transverse emittance (μm)	< 20	< 30
Transverse size rms (mm)	> 0.4	> 0.4
Longitudinal correlation	Negative	Negative
Transverse correlation	Negative	Negative

Table 9.3: Constraints for the 500pC and 1nC optimisation

Reaching the target mean energy of 7MeV is probably the dominant constraint as all solutions are at the 7MeV lower limit. The original design had 7 accelerating cells with which to achieve 7MeV, with only 3 cells, the gradient required per cell is much larger. The range of values for each variable of all the solutions gives some insight into what is optimum for the problem defined. For example,

in the 1nC case, for all solutions the final two accelerating cells tend to have a peak on axis field at the maximum limit of 30MV/m in order to achieve 7MeV final energy.

9.4.2 *Three Objective Optimisation*

The 7MeV specification was devised when the layout consisted of 7 accelerating cavities, where it was assumed that each cell would give roughly 1MeV acceleration. Since the design was truncated to 3 accelerating cavities, the 7MeV goal is ambitious. Increasing the gradient in the accelerating cavities to achieve 7MeV can have the effect of increasing the off axis focusing of the beam. This section discusses the results of a 3 objective optimisation where the 7MeV constraint is relaxed.

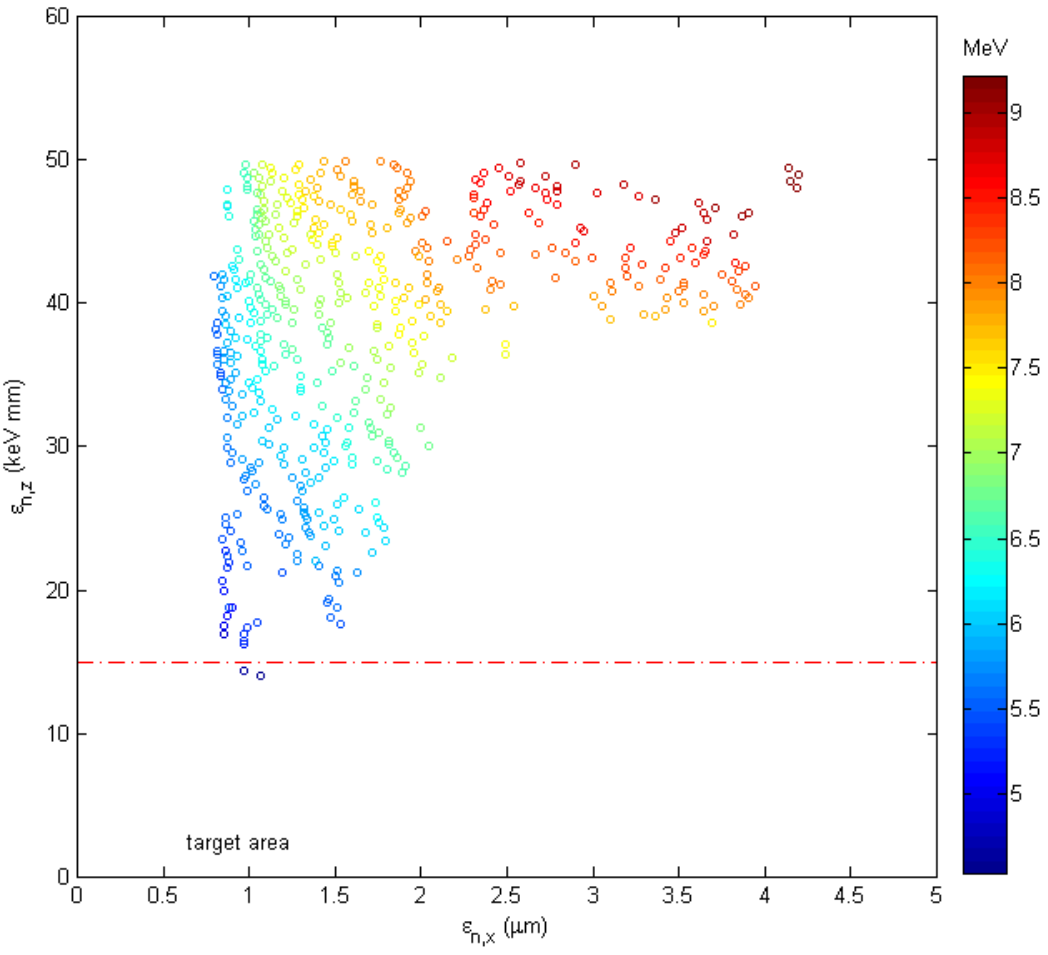

Figure 9.11: Three objective optimisation results for 135pC

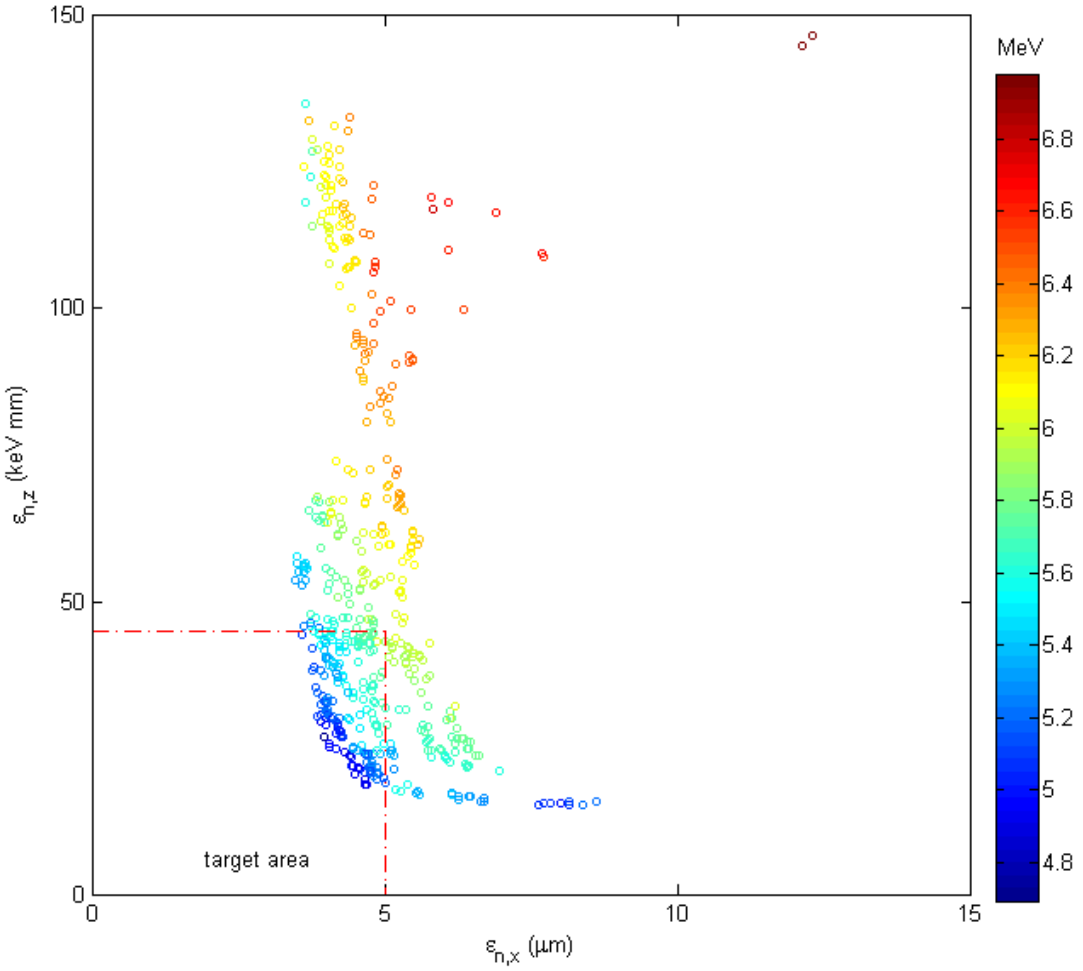

Figure 9.12: Three objective optimisation results for 1nC

Figures 9.11 and 9.12 show the plots of a 3 objective optimisation for the 135pC and 1nC scenarios respectively where the colour depicts the energy of the solution. The goal was again to minimise the longitudinal and transverse emittance, as well as maximising the output energy. There was still the constraint that the final energy should be above 4.5MeV, a value chosen so the energy would be comparable to that from an RF gun.

For the low charge case, the scatter plot shows that as the energy is reduced so is the minimum possible longitudinal and transverse emittance. When the electron bunch charge is 1nC, the solutions have an average energy of 5.7MeV, shown by the crowding of points towards the bottom of the plot. For both cases, at the lowest energy there are several solutions that meet the specification:

therefore if the 7MeV exit energy requirement is relaxed it is possible to improve the longitudinal and transverse emittance to within specification. Not all of these solutions meet the other specifications such as bunch length and energy spread, but there are a few that meet all specifications at lower beam energy.

9.4.3 Laser Pulse Duration

All of the optimisations so far have assumed that the longitudinal distribution from the cathode had a Gaussian distribution with a sigma of 20ps. This was chosen to model the properties of the laser that would be operated with the injector when it is tested. Changing the longitudinal profile of the laser is not too difficult, and so it would be reasonable to modify should there be some performance benefit to be had.

To investigate the effects of varying the longitudinal profile, a series of 2 objective optimisations were performed for both the low and high charge case. Firstly an additional variable was introduced which was the rms pulse length from the cathode, and could take any value between 10 and 20ps. Secondly the longitudinal profile need not be Gaussian, and as Cornell reports that beam quality is improved through using a uniform longitudinal distribution, this was also simulated. The constraints used in the following optimisations were more stringent than previously used and were set to the parameters in the specification.

Figures 9.13 and 9.16 show the results of these optimisations for the 135pC and 1nC scenario respectively.

For the 135pC case, there was an improvement to both longitudinal and transverse emittance with the addition of emission time as a variable. Furthermore, all the solutions have a beam energy above the 7MeV specification. In order to achieve this, the trend for all solutions was to minimise the longitudinal pulse length from the cathode. All solutions therefore have an emission time of 10ps

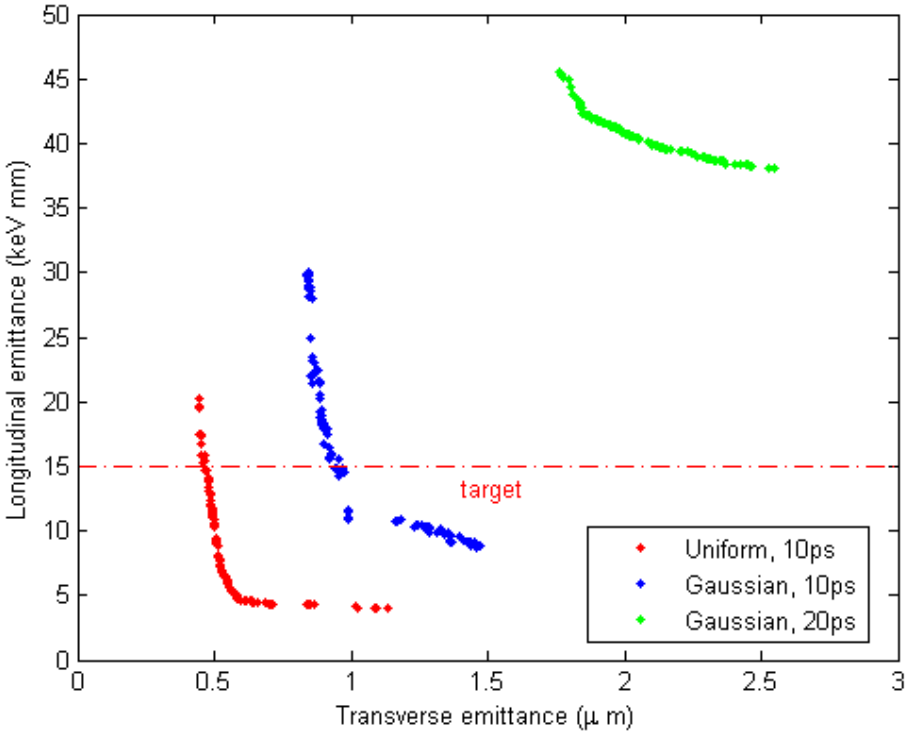

Figure 9.13: 135pC optimisation with the longitudinal distribution at the cathode as a variable

rms for both the Gaussian and uniform profiles. Figure 9.13 additionally shows that there is merit in using the uniform distribution over the Gaussian profile. This is because the space charge forces are more evenly distributed in the uniform beam. Figure 9.14 shows a solution chosen from the centre of each optimal front. The transverse evolution is similar for both uniform and Gaussian initial solutions, but the longitudinal profile shows shorter bunch length and emittance for the uniform case. It can be seen from the longitudinal phase space in figure 9.15 that the longitudinal emittance is smaller for the uniform distribution. For both the Gaussian and uniform scenarios, all aspects of the specification are met for all solutions of the optimisation.

In the 1nC scenario shown in figure 9.16, improvements are again made by permitting the emission time to be lower, but the longitudinal emittance is still larger than desired. For that reason, reducing the emission time alone will not result in the target beam parameters being met. Some combination of reduced

energy requirement and shorter emission time will be necessary to reach the specification in terms of longitudinal and transverse emittance.

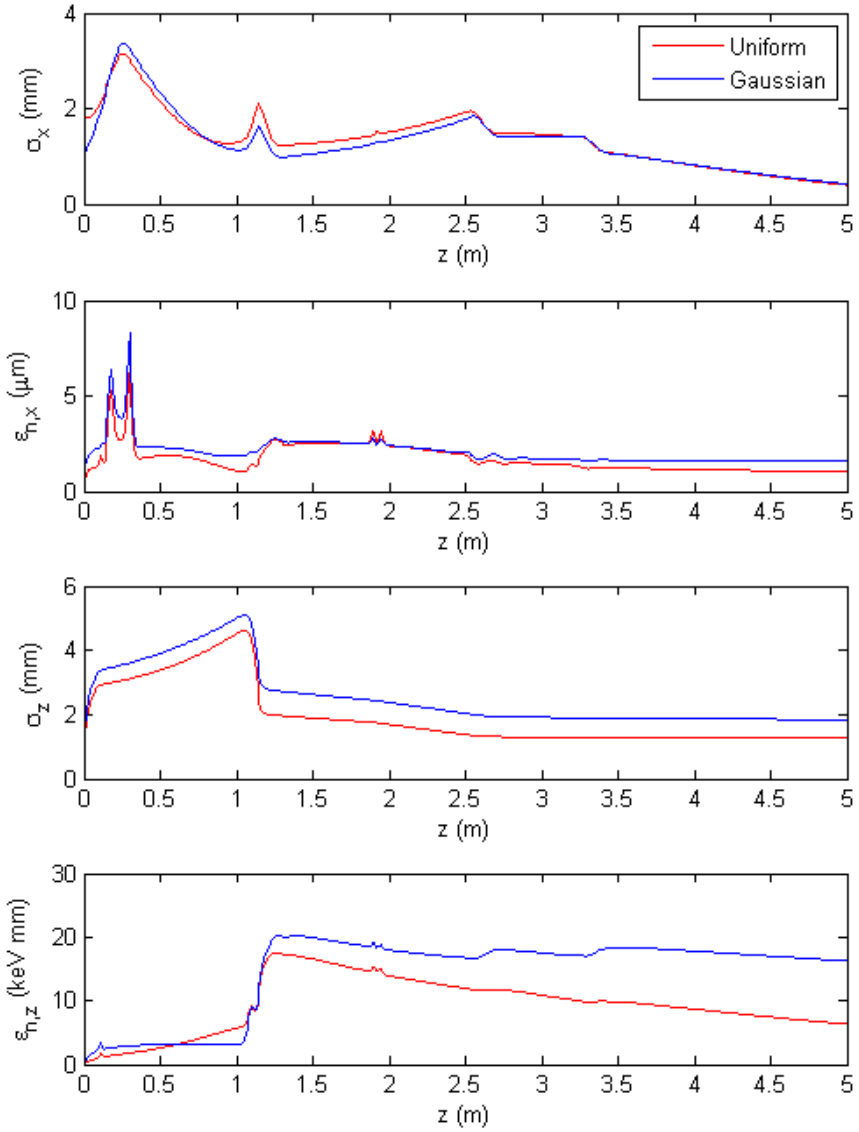

Figure 9.14: Bunch evolution as a function of distance from the cathode for uniform and Gaussian temporal distributions (135pC)

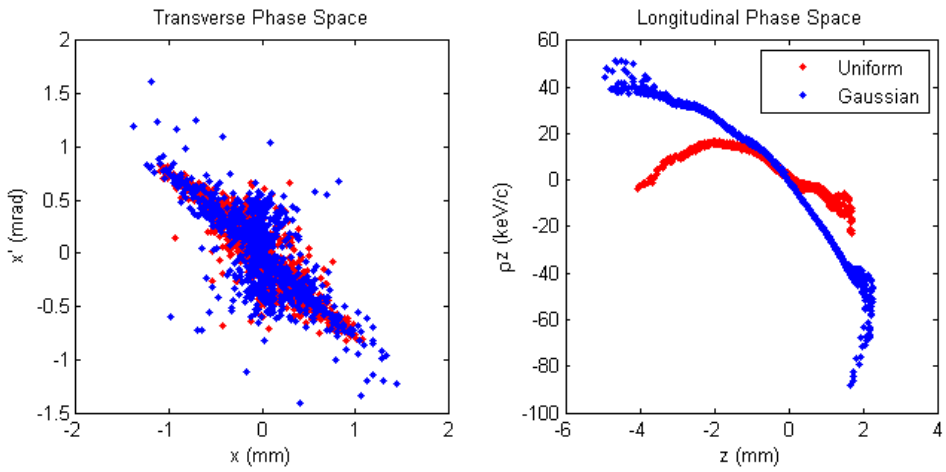

Figure 9.15: Phase space at the injector exit (5m)

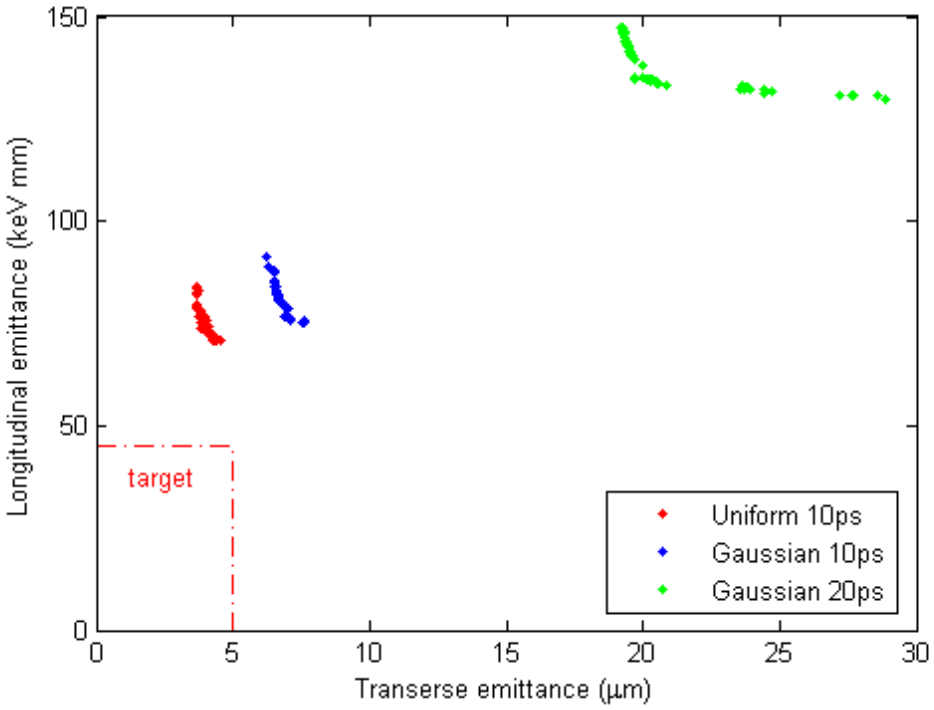

Figure 9.16: 1nC optimisation with the longitudinal distribution at the cathode as a variable

9.4.4 *Photocathode Properties*

The last optimisation runs include the thermal emittance values obtained from the experiments performed at Cornell University described in chapter 6. For the wavelength of 527nm, the estimated thermal energy is 118meV, which was included into the simulation. The final few generations were performed using 50,000 macro particles in the simulation to improve the accuracy of the results. For both the high and low charge cases, a uniform longitudinal distribution was used and emission time was a variable.

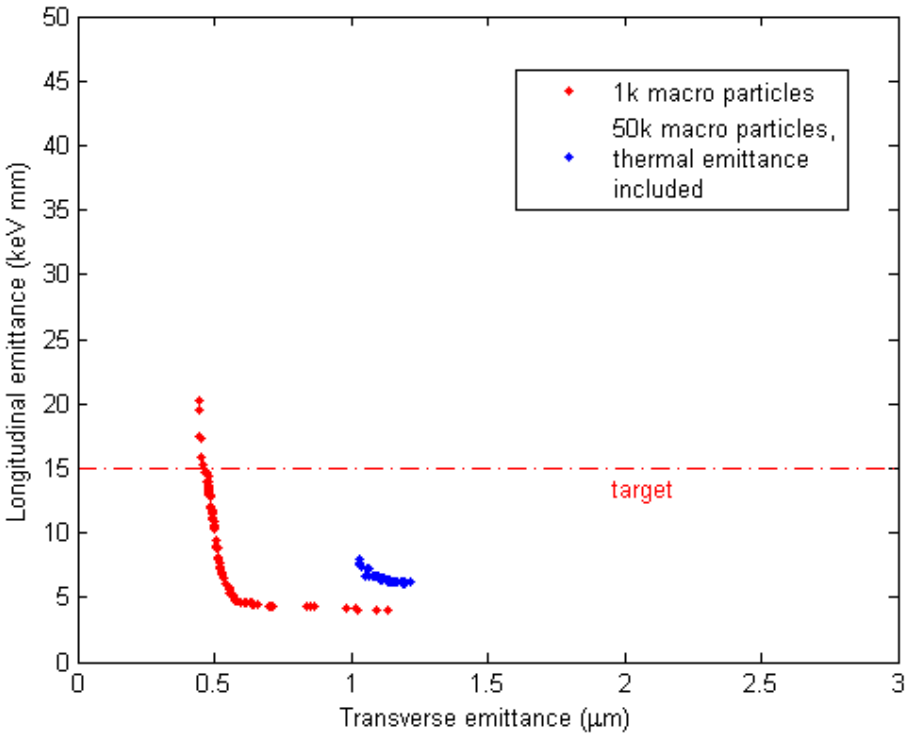

Figure 9.17: 135pC optimisation with thermal emittance included

As expected, the inclusion of thermal emittance shifts the optimal front to higher values of transverse emittance, shown in figure 9.17 for the 135pC case. The longitudinal emittance remains almost the same, the slight increase arising from both improved accuracy in the simulation and reduced emission area.

As with the case which excludes the thermal emittance, the components in the injector behave in a similar manner. The main difference is the emission area on the cathode. Rather than being maximised at the $\sigma_r = 2$mm limit, all solutions have a value within the range of $\sigma_r = 1.76 \pm 0.12$mm. This occurs because of the dependance of thermal emittance on emitting area.

A few solutions from the 135pC optimal front have been randomly chosen, and the beam evolution from the cathode to the exit of the injector is shown in figure 9.18. Since the extent of the optimal front is so small, the variation in beam parameters between the solutions is also small.

The final optimisation of the 1nC case included the thermal emittance of the cathode, but also had a reduced energy requirement in order to achieve the specification. A constraint of 5MeV was applied, as figure 9.12 suggested that solutions that would meet the other aspects of the specification could be found in this region. Figure 9.19 shows the optimal front for this scenario.

The emission area for the solutions of this optimisation is not minimised to reduce any increase in transverse emittance introduced by the inclusion of thermal energy. The reason being the transverse emittance was already well within the desired range and in this case it was more important to reduce the space charge forces within the bunch. By changing the energy constraint another degree of freedom was introduced into the optimisation which resulted in a lower longitudinal emittance.

To summarise, table 9.4, shows the component settings and expected beam properties for 135pC and 1nC bunch charges. The solution in each case is chosen from the knee of the optimum front to give typical parameters.

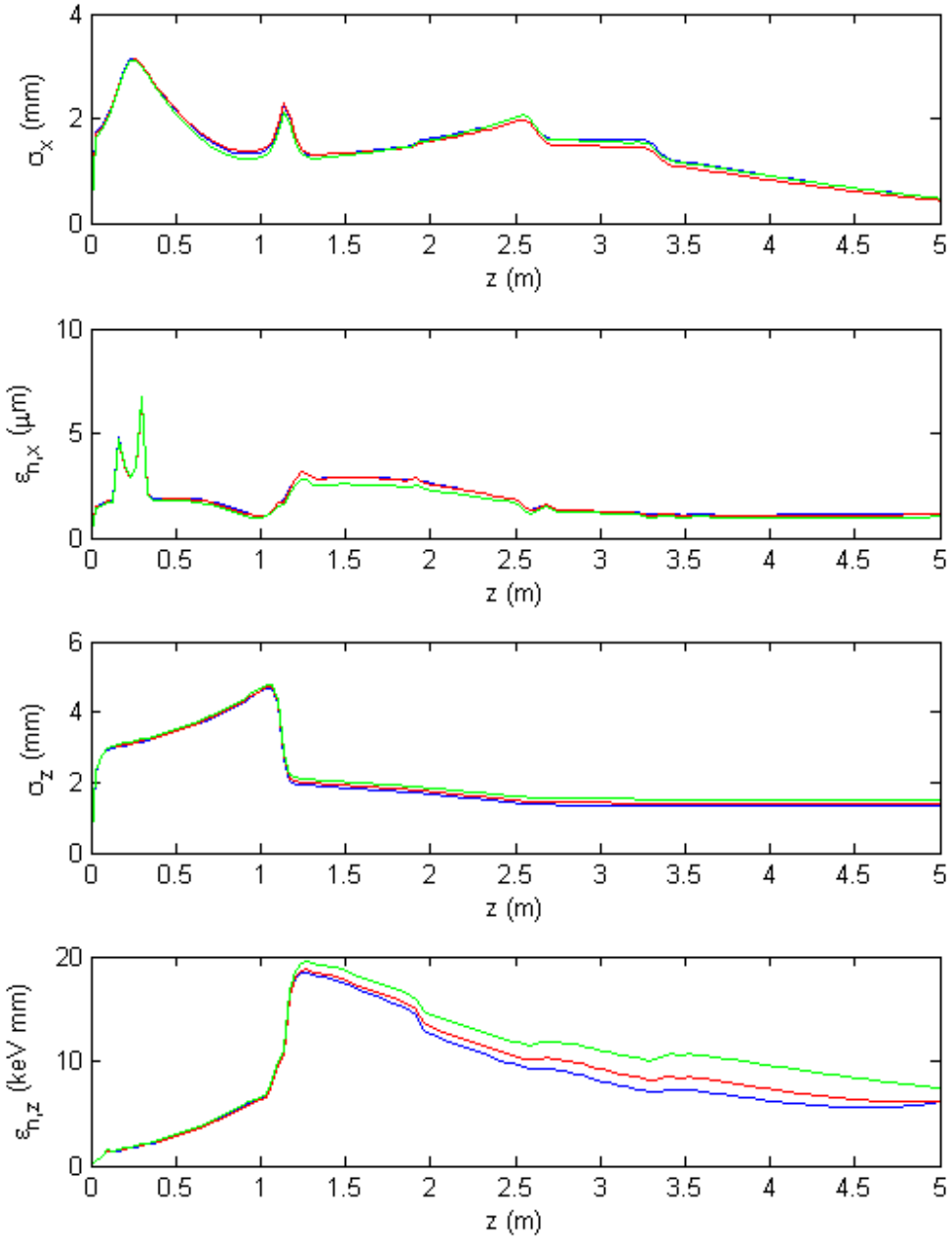

Figure 9.18: 135pC bunch evolution as a function of distance from the cathode for 3 solutions chosen from the optimal front

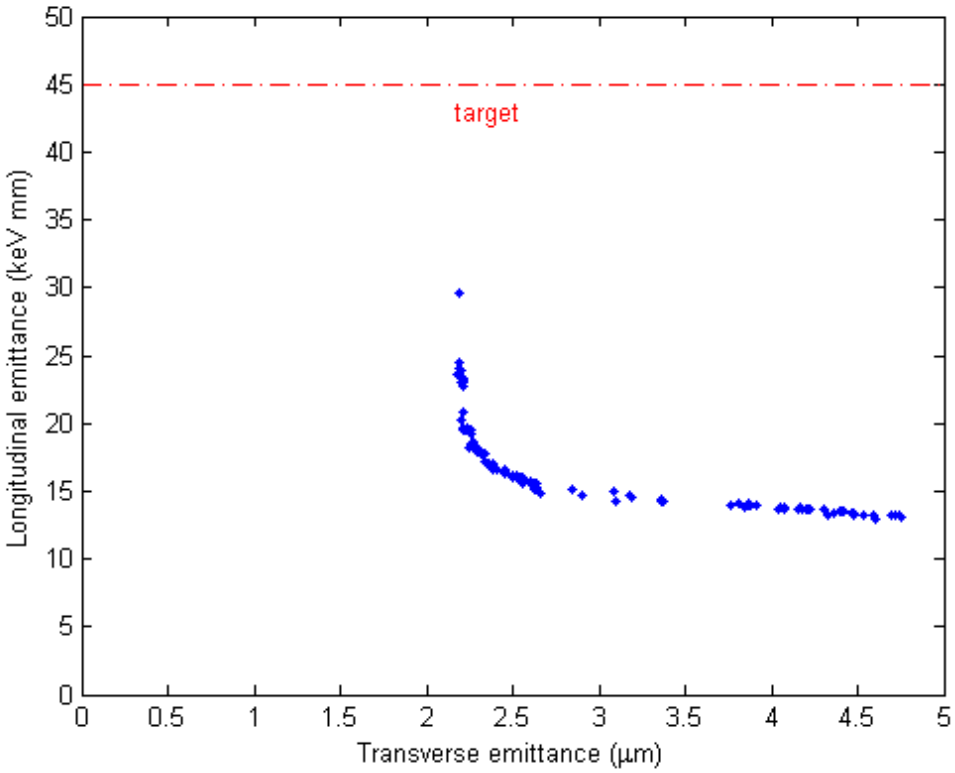

Figure 9.19: 1nC optimisation with thermal emittance included (all solutions have an energy above 5MeV)

Variables	135pC	1nC
Transverse spot size rms (mm)	1.8	2
Emission time rms (ps)	10	19
Peak solenoid field (G)	411	434
Max. gradient in cavity 1 (MV/m)	24.8	19.0
Phase of cavity 1 (deg)	-32.6	-33.3
Max. gradient in 3rd harmonic (MV/m)	6.8	17.5
Phase of 3rd harmonic (deg)	-148.9	-176.4
Max. gradient in cavity 3 (MV/m)	22.0	18.3
Phase of cavity 3 (deg)	15.1	-4.4
Max. gradient in cavity 4 (MV/m)	27.0	18.3
Phase of cavity 4 (deg)	15.0	-3.5
Simulated properties at 5m from cathode		
Bunch length rms (mm)	1.4	2.9
Energy spread rms (keV)	8.8	97.5
Mean energy (MeV)	7.0	5.1
Longitudinal emittance (keV mm)	6.3	17.1
Transverse emittance (μm)	1.1	2.3

Table 9.4: Final operating points for each component and expected electron beam properties

9.4.5 *Sensitivity*

The results of the optimisation study show that some parameters are almost the same for all solutions. The emission area and time, as well as the solenoid field for example. In practice it may not be possible to set the injector to exactly these values, so it becomes important to understand the precision to which each parameter can be set and how any variation may effect the properties of the electron beam at the exit of the injector.

Experience from the JLab FEL injector shows that the voltage of the DC gun can be controlled precisely to within 1kV and any slight variation has a negligible effect on the electron beam. The gradient in RF components can be set to within $\pm$ 50kV/m and the phase to $\pm$ 0.5°. The solenoid field can be set within $\pm$ 10G of the desired value. Finally the precision of the laser diameter and pulse length is estimated to be $\pm 50\mu$m and $\pm$ 0.5ps respectively.

Given the tolerance of each component, it was possible to simulate the effect of variation in the parameter value from the required setting. The desired set-point for each parameter was chosen to be the average value from the solutions of the optimisation. Fifty simulations were made where the parameter values would take a random value within the tolerance around the desired setting. For the 135pC case the variation in longitudinal and transverse emittance around the desired point is shown in figure 9.20. The range of both emittance values is small, and all remain within the specified region. The most sensitive output parameters are the electron bunch energy spread and bunch length which change by up to 15% ($\Delta\sigma_{\Delta E_{kin}} = 2$keV) and 8% ($\Delta\sigma_z = 0.11$mm).

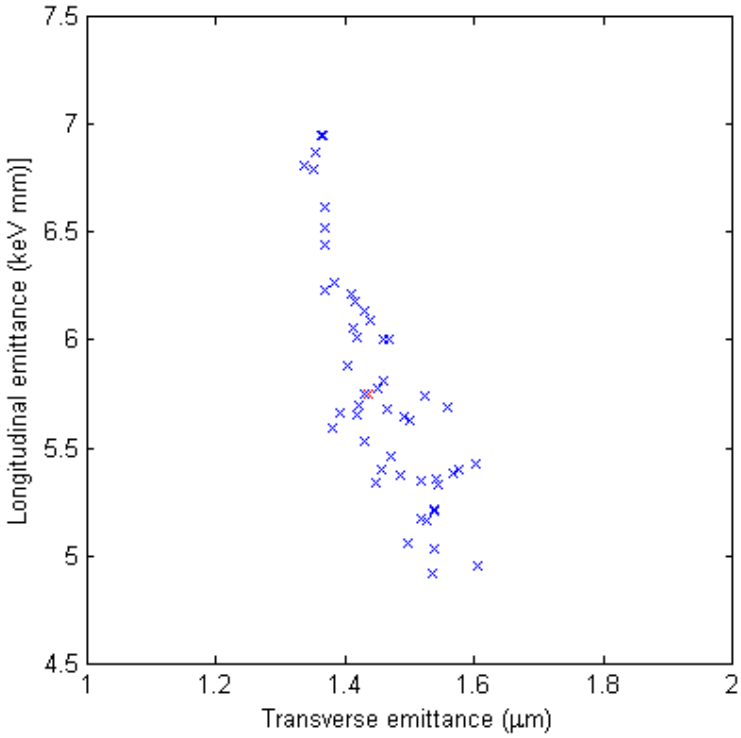

Figure 9.20: 135pC sensitivity due to variation in parameter settings (set-point shown in red)

9.5 Conclusions

Optimisation techniques for complex systems where there are many variables and objectives, provide an effective way of sampling the entire parameter space for the best solutions. With a large number of variables it quickly becomes impractical to explore all combinations and is even more difficult when there is more than one objective to optimise on. The results of optimising the JLab/AES injector also show that the expected operating regime is not always optimal, particularly with reference to the 3rd harmonic cavity in this case.

By comparison to manually finding an operating point for the injector, the optimisation program demonstrated improved results for the same objectives. It was also shown that to keep within the constraints of the original specification it would not be possible to achieve each one with the injector components and layout as it is being constructed. To achieve the required emittance, bunch length etc., flexibility of other aspects of the specification and starting conditions was essential. Reducing the required energy of the electron bunches at the end of the injector and introducing another objective to maximise the energy whilst meeting all other aspects of the specification improved the output properties. It emerged though, that it was still not possible to be within all items of the specification. The transverse properties of electron bunches at lower energy were well within the defined target area, but the longitudinal parameters were more problematic. The bunch length and energy spread were too large, and without a buncher cavity which could reduce these, the only alternative was to reduce the emission time from the cathode. The inclusion of emission time as a variable had a large effect on the longitudinal beam properties, and the optimisation showed that shorter was better. There is a physical limit as to how short the emission time can be because of the charge limit and the response time of the cathode. In addition, with very short bunches the transverse beam size will grow due to space charge forces. By reducing the bunch length from

the cathode, the accelerating cavities do not need to do so much bunching and so there is more flexibility to use them in other regimes. With the high charge case, minimising the emission time from the cathode improved the longitudinal emittance but not enough to be within specification. Therefore as the bunch charge is increased, the output energy must also be reduced to achieve other beam properties.

To conclude, if the expected tolerances on keeping the set values of each component in the injector can be maintained, the variation in the electron beam properties at the end will be minimal and stay within the specification.

CHAPTER 10

Conclusions

High average current electron beams are necessary to produce higher power lasing in the JLab IR FEL. The order of magnitude increase in required current prompted the design of the low frequency electron injector that would accelerate 100mA, without the beam break up problems associated with higher frequency RF cavities.

Construction had begun on the JLab/AES injector before the electron beam dynamics were fully understood. This unusual situation gave rise to the investigation of finding the best operation point for an injector with a fixed layout.

As the injector design consisted of six physical components, it was considered a reasonable undertaking to find an operating scenario by investigating the effects of each component in turn. The result of this study culminated in the realisation that it would not be possible to bring about all aspects of the beam specification simultaneously. It was particularly difficult to meet the longitudinal criteria of low emittance, short bunch length and high energy at the end of the injector. This inability to meet the specification was confirmed through a multivariate optimisation which was used to investigate the entire parameter space for the best solution. The results from the optimisation of the JLab/AES injector in chapter 9, show that with carefully chosen laser properties, the specification

determining the beam quality requirements can be met for the low charge case. Through shortening the emission time of the electron bunch from the cathode, the accelerating cavities did not have to provide so much bunching. For the 1nC, high charge scenario, the energy the electrons have exiting the injector must be reduced in addition to the emission time, in order to preserve other aspects of the beam quality. As the maximum energy of the electron bunches leaving a DC gun is low, and the bunch length long, compared to that from an RF gun, it requires longitudinal bunch compression. Intuitively the first cell of the JLab/AES injector must be used to shorten the electron bunches, which was confirmed by the optimisation results. The optimisation also showed that the 3rd harmonic cavity could be more effectively used in an unorthodox way. By using it for acceleration purposes as well as to reduce the longitudinal emittance induced in the first cavity, the beam properties could be brought to within specification whilst maximising the output energy.

Simulation is only as good as the assumptions used and the starting conditions that are defined. ASTRA is an effective tool for modelling electron injectors that operate in the low energy, high space charge regime described here. The results of the tomography experiment and space charge phase space measurement show that the finer features of transverse phase space can be well represented through simulation, providing the initial conditions are accurate. The benchmarking of ASTRA with the measurement of phase space showed excellent quantitative agreement for the 20pC bunch case. The data became more noisy with increased bunch charge, and because of this the 80pC bunches showed good qualitative agreement, but there was some discrepancy numerically. For these simulations it was found to be important to include the thermal energy measurement of GaAs, as this contributes to the overall emittance measured from phase space.

The measured thermal emittance of GaAs and GaAsP cathodes shows a dependence on the wavelength of the laser. At shorter wavelengths the thermal energy increases. In order to reduce the emittance from an electron injector it would be necessary to minimise this. On account of the response time of GaAs being more prompt and the quantum efficiency higher at shorter wavelengths, there has to be a compromise made. The measurement of GaAsP yielded higher thermal energy values than expected, so further investigation into the effect of P doping levels is required.

The thermal emittance is the theoretical lower boundary for transverse emittance, therefore it should be accounted for in simulation. When this was included into the optimisation of the JLab/AES injector, it was found that there was an optimum value for the emitting area of the cathode. Excluding the thermal emittance, the optimum was found at the maximum area permitted, as this would reduce the space charge forces within the electron bunch. The thermal emittance is proportional to the emitting area, so a balance must be made between the two conflicting factors, reducing the thermal emittance and reducing space charge forces.

Phase space tomography techniques are advantageous for situations where either space for diagnostics is limited or interceptive devices are not appropriate. The tomography experiments described in this thesis show that a tomography diagnostic can be made by utilising the solenoid magnets of the beamline. Solenoid magnets have the advantage of acting on both transverse planes in the same way. Quadrupoles are more usually used as tomography diagnostics in working machines, and tend to be placed after the LINAC in a high energy region. As quadrupoles are focusing in one plane, whilst defocusing in the other, it is only possible to reconstruct one transverse phase space at a time. Solenoids are commonly found in injectors with DC electron guns, which make them appropriate for reconstructing the phase space in these regions. For the

case where the electron bunches have negligible space charge, the reconstructed phase space shows excellent agreement with both measured and simulated results. If space charge is a dominant factor in the evolution of the electron bunch, the situation becomes more complex. The transport matrix from which the rotation and scaling of the projections come must include the effects of space charge. This can be achieved by introducing a defocusing term to approximate the effect of increasing the transverse beam size from space charge forces within the electron bunch. The virtual experiment showed that the reconstructed phase space became increasingly inaccurate with higher bunch current and shorter bunch length. When the thermal emittance is included into the simulation in a realistic way, the space charge forces within the bunch become increasingly non linear as the bunch travels along the beamline. The defocusing space charge term introduced into the transfer matrix does not accurately represent the forces calculated by simulation. This is because it only represents the linear space charge forces. When short electron bunches are considered the bunch expands, and noticeably changes the current within the bunch. The beam envelope equation used in this thesis does not account for this effect and therefore does not approximate the results of simulation well. Finally, due to the complexity in making a tomography measurement for beams with high space charge it may be more beneficial to use an alternative diagnostic. The tomography experiment requires knowledge of the initial conditions that are not always available. If the electron energy is not too high, a slit based phase space measurement would be much less time consuming and more convenient.

10.0.1 Future Work and Outlook

The optimisation procedure reported here was used to find the best performance of the JLab/AES injector with the initial constraint that the layout was already determined. The electron beam properties could be improved considerably if there was greater freedom introduced into the problem. For example:

if the original design with seven accelerating cavities were used it would be possible to manipulate the electron beam more, whilst gradually increasing the energy in each cell. Additional benefits could be gained from using a buncher cavity, designed specifically for bunch compression, or even including superconducting solenoids into the cryo-module. These components could be incorporated into the optimisation, with the addition of their longitudinal position from the cathode as variables.

One paramount conclusion from this thesis reveals that the initial properties of the electron bunches have a large impact on the achievable properties elsewhere in the machine. It follows that it may be possible to improve the electron beam quality further by investing time into a specifically designed electron gun. For example, with a DC gun, the shape of the electric field could be optimised to counteract the effects of space charge forces. To some extent this was achieved with the Cornell DC gun, where a Pierce electrode was used for this purpose. It maybe however, that more complex geometries could have significant benefits. Through an extension of the multivariate optimisation program it would be possible to introduce a static field solver such as POISSON into the loop and optimise the electrode shape for improved beam properties. Expanding this idea further, a comparison could be made by introducing an RF gun into the optimisation.

Whilst the electron beam properties from this injector are good enough to be used with a machine with a long wavelength FEL such as that at JLab, it would not be suitable for X-ray FELs without some redevelopment. The emittance and bunch length are too large for these applications. An injector that could be used with any FEL facility, would have to combine the possibility of high average current and high brightness. DC guns are limited in their energy and achievable bunch length, whilst normal conducting RF guns have not yet been operated CW in a FEL facility. In principle, a superconducting RF gun would

overcome both of these issues, but inherently has others. For example there is a difficulty in inserting a cathode inside a superconducting gun, and magnetic fields cannot be placed close in case of quenching. If solutions to these problems can be found, this would provide a means to creating a flexible next generation electron injector that could be used with any of the proposed light sources.

APPENDIX A

Derivation of Emittance Equation

The general solution to the equation for particle motion is:

$$x = \sqrt{\beta_x(z)\varepsilon_x}\cos\left(\phi_x(z) - \phi_0\right) \tag{A.1}$$

Let $\phi = \phi_x(z) + \phi_0$ and $\beta_x(z) = w^2$. The differential of the general solution becomes (dropping the x sub-scripts):

$$x' = \sqrt{\varepsilon}w'\cos\phi - \sqrt{\varepsilon}w\phi'\sin\phi \tag{A.2}$$

The second differential is:

$$x'' = \sqrt{\varepsilon}w''\cos\phi - 2\sqrt{\varepsilon}w'\phi'\sin\phi - \sqrt{\varepsilon}w\phi''\sin\phi - \sqrt{\varepsilon}w\phi'^2\cos\phi \tag{A.3}$$

Inserting equations A.3 and A.1 into equation 2.1 yields the general equation:

$$\begin{aligned}\sqrt{\varepsilon}w''\cos\phi - 2\sqrt{\varepsilon}w'\phi'\sin\phi - \sqrt{\varepsilon}w\phi''\sin\phi - \sqrt{\varepsilon}w\phi'^2\cos\phi \\ +K(z)\sqrt{\varepsilon}w\cos\phi = 0\end{aligned} \tag{A.4}$$

$\sin\phi$ and $\cos\phi$ are linearly independent, and so for the general equation to hold,

$$\begin{aligned} 2w'\phi' + w\phi'' &= 0 \\ \frac{d}{dz}(\phi' w^2) &= 0 \\ \Rightarrow \phi' w^2 &= k \\ \phi &= \int \frac{k}{w^2} \end{aligned} \tag{A.5}$$

yielding an equation for the phase advance.

Rearranging equation A.1 gives,

$$\cos\phi = \frac{x}{\sqrt{\varepsilon\beta}} \tag{A.6}$$

Inserting into equation A.2 gives an equation for $\sin\phi$,

$$\begin{aligned} x' &= \frac{1}{2}\sqrt{\varepsilon}\frac{\beta'}{\sqrt{\beta}}\frac{x}{\sqrt{\varepsilon\beta}} - \frac{\sqrt{\varepsilon}}{\sqrt{\beta}}\sin\phi \\ x' - \frac{1}{2}\frac{\beta'}{\beta}x &= -\frac{\sqrt{\varepsilon}}{\sqrt{\beta}}\sin\phi \\ \frac{\sqrt{\beta}}{\sqrt{\varepsilon}}\left[x' - \frac{1}{2}\frac{\beta'}{\beta}x\right] &= \sin\phi \end{aligned} \tag{A.7}$$

Combining

$$\begin{aligned} \cos^2\phi + \sin^2\phi &= 1 \\ \frac{x^2}{\varepsilon\beta} + \frac{\beta}{\varepsilon}\left[x' - \frac{1}{2}\frac{\beta'}{\beta}x\right]^2 &= 1 \\ \frac{x^2}{\beta}\left[1 + \frac{\beta'^2}{4}\right] - \beta' x x' + \beta x'^2 &= \varepsilon \end{aligned} \tag{A.8}$$

This is the equation of an ellipse, and can be generalised by

$$\gamma_x x^2 + 2\alpha_x x x' + \beta_x x'^2 = \varepsilon_x \tag{A.9}$$

APPENDIX B

Filtered Back Projection

B.1 Fourier Slice Theorem

The Slice Theorem tells us that the 1D Fourier transform of the projection function $P(\theta, t)$ is equal to the 2D Fourier transform of the image evaluated on the line that the projection was taken on.

The Fourier transform of a 2D function $f(x, y)$ is given by:

$$\begin{aligned} F(u,v) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-2\pi j(ux+vy)}\,dxdy \\ u &= w\cos\theta,\ v = w\sin\theta \\ w^2 &= u^2+v^2 \end{aligned} \tag{B.1}$$

Similarly the Fourier transform of a 1-D projection, $P_\theta(t)$, of function $f(x, y)$ at angle θ is:

$$S_\theta(w) = \int_{-\infty}^{\infty} P_\theta(t)e^{-2\pi jt}\,dt \tag{B.2}$$

When $\theta = 0$:

$$F(u,0) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-2\pi j(ux)}\,dxdy \tag{B.3}$$

$$= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(x,y)\, dy \right] e^{-2\pi j(ux)}\, dx \tag{B.4}$$

$$= \int_{-\infty}^{\infty} P_\theta(x) e^{-2\pi j(ux)}\, dx \tag{B.5}$$

$$\tag{B.6}$$

Note that:

$$S_{\theta=0}(u) = F(u,0) = F(w,0) \tag{B.7}$$

This can be extended to give the proof that:

$$S_\theta(w) = F(w,\theta) \tag{B.8}$$

B.2 Filtered Back Projection Algorithm

The inverse Fourier transform of $F(u,v)$ is given by:

$$f(x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u,v) e^{2\pi j(ux+vy)}\, du dv \tag{B.9}$$

$$= \int_{0}^{2\pi} \int_{0}^{\infty} F(w,\theta) e^{2\pi jw(x\cos\theta + y\sin\theta)} w\, dw d\theta \tag{B.10}$$

$$\tag{B.11}$$

Recall that the equation of a line is given by $t = x\cos\theta + y\sin\theta$. The above integral can be split into two by considering θ from 0 to 180° and 180 to 360° and using the property that $F(w,\theta+\pi) = F(-w,\theta)$, the equation can be expressed:

$$f(x,y) = \int_{0}^{\pi} \left[\int_{0}^{\infty} F(w,\theta)|w| e^{2\pi jw(t)}\, dw \right] d\theta \tag{B.12}$$

$$= \int_{0}^{\pi} \left[\int_{0}^{\infty} S_\theta(w)|w| e^{2\pi jwt}\, dw \right] d\theta \tag{B.13}$$

Using the Fourier Slice Theorem gives equation B.13 which can be used to give an estimate of $f(x,y)$, given the transformed projection data $S_\theta(w)$. Equation B.13 represents a Ram-Lak filtering operation, where the frequency response of

the filter is given by $|w|$. Given an infinite set of projections $P_\theta(t)$ the distribution $f(x, y)$ can be completely reconstructed.

This algorithm can be simply implemented in MATLAB using an in-built function called 'iradon()'. This takes a sinogram and the angles for each projection as input. A sinogram is simply an array of all the projections. It is also possible to introduce a filter into the iradon function that will change the frequency response of the filtering operation by multiplying the Ram-Lak filter. The filters available are given in table B.1, which also shows the error (given by equation 6.13) for the reconstruction of the ellipse in figure 6.10.

Projections / Filter	4	45	180
Ram-Lak	113	13.2	4.1
Shepp-Logan	108	12.7	4.2
Cosine	112	11.9	4.7
Hamming	99	11.2	5.1
Hann	97	11.1	5.3

Table B.1: Error [%] in reconstructions using FBP algorithm with various filters

APPENDIX C

Maximum Likelihood Expectation Maximisation Algorithm

The principle behind the MLEM algorithm is to create an initial guess of the distribution that is being described by the projections. Projections of the guess are taken and then compared to those measured. A relation is established between these, and correction projections are generated. The correction projections are then used to modify the initial guess, and a new guess is created. This process is repeated until the guess converges. This distribution then represents that being described by the projections.

C.1 Maths

The MLEM algorithm as described by Lange and Carson [128] is given by:

$$f_j^{(k+1)} = \frac{f_j^{(k)}}{\sum_{i=1}^{n} a_{ij}} \sum_{i=1}^{n} \frac{g_i}{\sum_{j'=1}^{m} a_{ij} f_{j'}^{(k)}} a_{ij} \tag{C.1}$$

$f_j^{(k)}$ is the current estimate, and $f_j^{(k+1)}$ the next estimate.

There are i projections of the measured distribution, and the reconstructed image will have m pixels. $\sum_{j'=1}^{m} a_{ij} f_{j'}^{(k)}$ gives the projection of the estimate at the angle corresponding to i. $\frac{g_i}{\sum_{j'=1}^{m} a_{ij} f_{j'}^{(k)}}$ gives the ratio of measured projection to estimated projection. $\sum_{i=1}^{n} \frac{g_i}{\sum_{j'=1}^{m} a_{ij} f_{j'}^{(k)}} a_{ij}$ is the back projection of this ratio for pixel j. Finally, a_{ij} is a weighting that corresponds to the probability of each bin contributing to the image.

C.2 MATLAB Implementation

```
function recon = MLEM2(Isino, theta, ite, beta)

if nargin < 4
    beta = 0;
end

[bins, views] = size(Isino);

if views ~= length(theta)
    error('Number of rotations does not match that of sinogram')
end

init = ones(bins); % create initial guess image
init = init / sum(init(:)) * sum(Isino(:));
err = [];
err(1) = 0;

for i=1:ite
    expect = makesino(init, theta); % create sino of guess image
    exp_sum = zeros(bins, views); % correction sino
    exp_sum(expect>0) = Isino(expect>0)./expect(expect>0);
    exp_sum(Isino==0) = 0; %no NaN
    exp_sum(expect==0) = 0;
    err(i+1) = sum((exp_sum - expect).^2) /sum(exp_sum);
    pimg = imgproject2(exp_sum, theta); % recon on correction sino

    if beta > 0
        init = mrp(init, beta, [3 3]); % filter the guess image
    end

    init = pimg.*init; % update the guess image
    init(init<0) = 0;
```

```
end

recon = init/ views;

end
%~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
function pimg = imgproject2(sino, theta)

[bins,views] = size(sino); pimg = repmat(sino(:,1)',bins,1);

pimg = imrotate(pimg,-theta(1),'bilinear','crop');

for i = 2:views
    sri = repmat(sino(:,i)',bins,1);
        % rotating the sinogram
    sri = imrotate(sri,-theta(i),'bilinear','crop');
        % new projected image is obtained from the old PIMG by adding
        % the sinogram row image to it.
    pimg = pimg + sri;
end

pimg = flipdim(pimg,1);

end
%~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
function sino = makesino(img, phi)

y = size(img,1);

% Creating the projections of the sinogram by rotating the image and
% summing the columns of the image
sino = zeros(y,length(phi));

for i = 1:length(phi)
    sino(:,i) = sum(imrotate(img,-phi(i),'bilinear','crop'));
end

sino(sino(:)<(max(sino(:))/10000)) = 0;

end
%~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
function pen = mrp(img, beta, fsize)

[ySz xSz] = size(img);
```

```
filt_img = medfilt2(img,[fsize(1)fsize(2)]);

B = beta*img + filt_img*(1-beta); nonzero = find(B);

cp = zeros(ySz,xSz);
cp(nonzero) = filt_img(nonzero)./B(nonzero);
cp(img == 0) = 0; pen = cp .* img;

end
```

References

[1] F.R. Elder, A.M. Gurewitsch, R.V. Langmuir and H.C. Pollock. Radiation from Electrons in a Synchrotron. *Physical Review*, 71:829–830, 1947.

[2] E. Wilson. *An Introduction to Particle Accelerators*. Oxford, 2001.

[3] I.V. Bazarov. Overview of energy recovery linacs. *Particle Accelerator Conference*, pages 382–384, 2005.

[4] M. Tigner. A Possible Apparatue for Electron Clashing-Beam Experiments. *Nuovo Cimento*, 37:1228, 1965.

[5] T.I. Smith, H.A. Schwettman, R. Rohatgi, Y. Lapierre and J. Edighoffer. Development of the SCA/FEL for use in Biomedical and Materials Science Experiments. *Nuclear Instruments and Methods A*, 259, 1:1–7, 1987.

[6] R.W. Warren et al. Recent Results from the Los Alamos Free Electron Laser. *Nuclear Instruments and Methods A*, 259(1-2):8–12, 1987.

[7] G.R. Neil et al. Sustained Kilowatt Lasing in a Free-Electron Laser with Same-Cell Energy Recovery. *Physical Review Letters*, 84(4):662, 2000.

[8] EuroFEL Collaboration. http://www.eurofel.org.

[9] D.A.G. Deacon, L.R. Elias, J.M.J. Madey, G.J. Ramian, H.A. Schwettman and T.I. Smith. First Operation of a Free-Electron Laser. *Physical Review Letters*, 38:892894, April 1977.

[10] W.B. Colson, J. Blau, and A. Kampouridis. Free electron lasers in 2006. *Free Electron Laser Conference*, pages 756–758, 2006.

[11] R. Hajima et al. First Demonstration of Energy-Recovery Operation in the JAERI Superconducting Linac for a High-Power Free-Electron Laser. *Nuclear Instruments and Methods A*, 507:115–119, 2003.

[12] N.A. Vinokurov et al. First Lasing at the High-Power Free Electron Laser at Siberian Center for Photochemistry Research. *Nuclear Instruments and Methods A*, 528:15–18, 2004.

[13] S. Benson et al. High Power Lasing In The IR Upgrade FEL At Jefferson Lab . *Free Electron Laser Conference*, page 229, 2004.

[14] G.L. Carr, M.C. Martin, W.R. McKinney, K. Jordan, G.R. Neil and G.P. Williams. High-power terahertz radiation from relativistic electrons. *Nature*, 420:153–156, 2002.

[15] C. Hernandez et al. Performance and Modeling of the JLAB IR FEL Upgrade Injector. *Free Electron Laser Conference*, page 558, 2004.

[16] A. Todd et al. State of the Art Electron Guns and Injector designs. *Particle Accelerator Conference*, 2005.

[17] I.V. Bazarov, B.M. Dunham, F.E. Hannon , Y. Li, X. Liu, T. Miyajima, D. Ouzounov, C.K. Sinclair. Thermal emittance and response time measurements of negative electron affinity photocathodes. *Journal of Applied Physics*, 2007.

[18] F.E. Hannon et al. Phase Space Tomography Using the Cornell ERL DC Gun. *European Particle Accelerator Conference*, 2008.

[19] I.V. Bazarov, B.M. Dunham, C. Gulliford, Y. Li, X. Liu, C.K. Sinclair, K. Soong, F.E. Hannon. Benchmarking of 3D space charge codes using direct phase space measurements from photoemission high voltage DC gun. *Journal of Applied Physics*, 2008.

[20] F.E. Hannon and C. Hernandez-Garcia. Simulation and Optimisation of a 100mA DC Photo-Injector. *European Particle Accelerator Conference*, pages 3550–3552, 2006.

[21] S. Humphries. *Charged Particle Beams*. John Wiley and Sons, 1990.

[22] B.E. Carlsten. New Photoelectric Injector Design for the Los Alamos National Laboratory XUV FEL Accelerator. *Nuclear Instruments and Methods A*, 285:313–319, 1989.

[23] C. Lejeune and J. Aubert. Emittance and Brightness: Definitions and Measurements. *Advances in Electronics and Electron Physics*, 13 A:159–258, 1980.

[24] K. Flöttmann. Some Basic Features of the Beam Emittance. *Physical Review Special Topics - Accelerators and Beams*, 6, 2003.

[25] Z. Huang and K-J Kim. Review of x-ray free-electron laser theory. *Physical Review Special Topics - Accelerators amd Beams*, 10, 2007.

[26] *CERN Accelerator School: Synchrotron radiation and free-electron lasers*, Brunnen, Switzerland, July 2003.

[27] W.B. Colson, C. Pellegrini, A. Renieri, editor. *Laser Handbook Vol. 6*. Elsevier Science Publishers B.V., 1990.

[28] S. Seely. Work Function and Temperature. *Physical Review*, 59:75, 1941.

[29] N. Nishimori et al. A Thermionic Electron Gun System for the JAERI Superconducting FEL. *European Particle Accelerator Conference*, page 1672, 2000.

[30] M. Sawamura et al. Status and Development for the JAERI ERL-FEL for High-Power and Long-Pulse Operation. *European Particle Accelerator Conference*, page 1723, 2004.

[31] SCSS X-FEL Conceptual Design Report. Technical report, SPRING-8, 2005.

[32] A.P. French and E.F. Taylor. *An Introduction to Quantum Physics*. Chapman and Hall, 1997.

[33] C. Hernandez-Garcia and C. A. Brau. Photoelectric field emission from needle cathodes. *Nuclear Instruments and Methods A*, 429:257–263, 1999.

[34] C. Hernandez Garcia and C.A. Brau. Pulsed photoelectric field emission from needle cathodes. *Nuclear Instruments and Methods A*, 483:273276, 2002.

[35] J.W. Lewellen and C.A. Brau. RF Photoelectric injectors using needle cathodes. *Nuclear Instruments and Methods A*, 507:323326, 2003.

[36] R. Ganter et. al. Low Emittance Gun Project Based on Field Emission. *Free Electron Laser Conference*, pages 602–605, 2004.

[37] W. Shao et.al. Micro-gated-field emission arrays with single carbon nanotubes grown on Mo tips. *Applied Surface Science*, 253:75597562, 2007.

[38] I.V. Khodak and V.A. Kushnir. RF gun with ferroelectric cathode. *Nuclear Instruments and Methods A*, 562:662665, 2006.

[39] R.L. Bell. *Negative Electron Affinity Devices*. Oxford University Press, 1973.

[40] R.L. Sheffield. High Brightness Electron Sources. *Eurpoean Particle Accelerator Conference*, page 882, 1996.

[41] D.H. Dowell et al. Results from the average power laser experiment photocathode injector test. *Nuclear Instruments and Methods A*, page 167, 1994.

[42] P. Michelato et al. R & D Activity on High QE Alkali Photocathodes for RF Guns. *European Particle Accelerator Conference*, page 1049, 1996.

[43] L. Monaco et al. High QE Photocathode Performance During Operation at FLASH/PITZ Photoinjectors. *Particle Accelerator Conference*, page 2763, 2007.

[44] S. Schreiber et al. On the Photocathodes used at the TTF Photoinjector. *Particle Accelerator Conference*, page 2071, 2003.

[45] P. Sez et al. Polarization Studies of Strained GaAs Photocathodes at the SLAC Gun Test Laboratory. *European Particle Accelerator Conference*, page 887, 1996.

[46] B.M. Dunham et al. Emittance Measurements for the Illinois/CEBAF Polarized Electron Source. *Particle Accelerator Conference*, pages 1030–1032, 1995.

[47] P. Hartmann et al. A diffusion model for picosecond electron bunches from negative electron affinity GaAs photocathodes. *Journal of Applied Physics*, 86(4):2245–2249, August 1999.

[48] P.V.Logatchov et al. High Power Test Of GaAs Photocathode In RF Gun. *European Particle Accelerator Conference*, page 1450, 1998.

[49] D. Holmes et al. Superconducting RF Photocathode Gun for Low Emittance Polarized Electron Beams. *Proceedings 13th International Workshop on RF Superconductivity*, page 1450, 2007.

[50] T. Rao et al. Photoemission Studies on BNL/AES/JLAB all Niobium Superconducting RF Injector. *Particle Accelerator Conference*, 2005.

[51] J. Smedley et al. Progress on Lead Photocathodes for Superconducting Injectors. *Particle Accelerator Conference*, 2005.

[52] I. Ben-Zvi et al. Secondary Emission Enhanced Photoinjector. Technical report, 2004. Internal Report C-A/AP/149.

[53] X. Chang et al. Recent Progress on the Diamond Amplified Photo-Cathode Experiment. *Particle Accelerator Conference*, pages 2044–2046, 2007.

[54] M.A. Allen, et al. RF Superconducting Materials Research at SLAC. *Particle Accelerator Conference*, page 1009, 1969.

[55] F.B. Kiewiet et al. A DC/RF Gun for Generating Ultra-Short High-Brightness Electron Bunches. *European Particle Accelerator Conference*, 2002.

[56] G.A. Westenskow and J.M.J. Madey. *Laser and Particle Beams*, 2:223, 1984.

[57] Y. Kim et al. Commissioning of S-band RF GUN and Linac for the Mark-III FEL Facility at Duke University. *Free Electron Laser Conference*, pages 411–413, 2006.

[58] M. Curtin et al. First demonstration of a free-electron laser driven by electrons from a laser-irradiated photocathode. *Nuclear Instruments Methods A*, 296:127, 1990.

[59] D. H. Dowell, et al. First Operation of a Photocathode Radio Frequency Gun Injector at High Duty Factor. *Applied Physics Letters*, 63 (15):2035–2037, 1993.

[60] F. Marhauser. High Power Tests of a High Duty Cycle, High Repetition Rate RF Photoinjector Gun for the BESSY FEL. *European Particle Accelerator Conference*, pages 68–70, 2006.

[61] J.W. Staples et. al. The LBNL Femtosource (LUX) 10 kHz Photoinjector. *Particle Accelerator Conference*, pages 2092–2094, 2003.

[62] H.A. Schwettman et al. Measurements at High Electric Field Strengths on Superconducting Accelerator Cavities. *V International Conference on High Energy Accelerators*, page 690, 1966.

[63] D. Janssen et al. On the Way to a Superconducting RF-Gun: First Measurements with the Gun Cavity. *Nuclear Instruments Methods A*, 445:408, 2000.

[64] D. Janssen et al. First Operation of a Superconducting RF Gun. *Nuclear Instruments Methods A*, 507:314, 2003.

[65] M. Cole et al. Progress Toward an all Niobium Superconducting RF Photocathode Electron Gun. *Particle Accelerator Conference*, page 2272, 2001.

[66] D. Janssen et al. Status of the Cell Rossendorf Superconducting RF Gun. *Free Electron Laser Conference*, page 359, 2004.

[67] C. Limborg, Y. Batygin, L. Giannessi, M. Quattromini, J.P. Carneiro, K. Flöttmann. Code Comparison for Simulations of Photo-Injectors. *Particle Accelerator Conference*, page 3548, 2003.

[68] L. Giannessi. Simulation codes for high brightness electron beam free-electron laser experiments. *Physical Review Special Topics - Accelerators and Beams*, 6(114802), 2003.

[69] M. Ferrario et al. Multi-bunch Energy Spread Induced by Beam Loading in a Standing Wave Structure. Technical report, 1995. TESLA Report 1995-06.

[70] L. Young and J. Billen. The Particle Tracking Code PARMELA. *Particle Accelerator Conference*, 2003.

[71] K. Flöttmann. ASTRA - A Space charge TRacking Algorithm, http://www.desy.de/˜mpyflo/Astra_dokumentation/Manual_part1.pdf.

[72] J. Qiang, S. Lidia, and R.D. Ryne. Three-dimensional quasistatic model for high brightness beam dynamics simulation. *Physics Review Special Topics - Accelerators and Beams*, 9(044204), 2006.

[73] S.B. van der Geer and M.J. de Loos. GPT- General Particle Tracer.

[74] Computer Simulation Technology, (CST). http://www.cst.com/Content/Products/MAFIA/Overview.aspx.

[75] Tech-X. http://www.txcorp.com/products/VORPAL.

[76] K. Halbach and R. F. Holsinger. LANL. SUPERFISH a computer program for evaluation of RF cavities with cylindrical symmetry, Particle Accelerators,1976, Vol. 7 (213).

[77] K. Flöttmann, S. Lidia, P. Piot. Recent improvements to the ASTRA particle tracking code. *Particle Accelerator Conference*, 2003.

[78] Cornell's laboratory is at the crossroads . *CERN Courier*, January 2002.

[79] I. Bazarov et al. The Energy Recovery Linac (ERL) as a Driver for X-Ray Producing Insertion Devices. *Particle Accelerator Conference*, pages 230 – 232, 2001.

[80] S. Moss. NSF awards Cornell $18 million to develop a new source of X-rays. *Cornell Chronical*, 36(23), February 2005.

[81] I. Bazarov et al. Phase I Energy Recovery Linac at Cornell University. *European Particle Accelerator Conference*, pages 664 – 666, 2002.

[82] S. A. Belomestnykh, V. D. Shemelin, K. W. Smolenski, V. Veshcherevich. Deflecting Cavity for Beam Diagnostics in ERL Injector . *Particle Accelerator Conference*, pages 2331–2333, 2007.

[83] B. M. Dunham, C. K. Sinclair, I. V. Bazarov, Y. Li, X. Liu, and K. W. Smolenski. Performance of a Very High Voltage Photoemission Electron Gun for a High Brightness, High Average Current ERL Injector. *Particle Accelerator Conference*, pages 1224–1226, 2007.

[84] N. Yamamoto et al. Thermal Emittance Measurements for Electron beams Produced from Bulk and Superlattice Negative Electron Affinity Photocathodes. *Journal of Applied Physics*, 102, 2007.

[85] I.V. Bazarov, B.M. Dunham, F.E. Hannon , Y. Li, X. Liu, T. Miyajima. Thermal Emittance Measurements from Negative Electron Affinity Photocathodes. *Particle Accelerator Conference*, pages 1221–1223, 2007.

[86] F.E. Hannon. Cathode Caesiation Methods at J-Lab. *ASTeC Internal note: erlp-inj-gun-rpt-003*, 2006.

[87] M.G. Minty and F. Zimmermann. *Measurement and Control of Charged Particle Beams*. Springer, 2003.

[88] V. Miltchev. Modelling the Transverse Phase Space and the Core Emittance Studies at PITZ. *Free Electron Laser Conference*, pages 556–558, 2005.

[89] The TTF FEL Team. SASE FEL at the TESLA Facility, Phase 2. *Internal Note: TESLA-FEL 2002-01*, 2002.

[90] D.J. Holder, C. Gerth, F.E. Hannon and R.J. Smith. ERLP Gun Commissioning Beamline Design. *LINAC Conference*, pages 93–95, 2004.

[91] M.P. Stocklia, R.F. Welton, R. Keller. Self-consistent, unbiased root-mean-square emittance analysis. *Review of Scientific Instruments*, 75(5):1646–1649, 2004.

[92] C.H. Wang et al. Slits Measurement of Emittance at TTF. *International Conference on Accelerator and Large Experimental Physics Control Systems*, pages 284–286, 1999.

[93] K. Abrahamyan et al. Transverse Emittance Measurements at the Photoinjector Test Facility at DESY Zeuthen (PITZ). *Diagnsotic Instruments for Particle Accelerators Conference*, pages 225–227, 2003.

[94] C.M. Bhat et al. Envelope and Multi-Slit Emittance Measurements at Fermilab A0-Photoinjector and Comparison with Simulations. *Particle Accelerator Conference*, page 2936, 2007.

[95] F.E. Hannon. Pepper-Pot Design for the ERLP Gun Diagnostic Beamline. *ASTeC Internal note: erlp-inj-gun-rpt-006*, 2004.

[96] J.G. Power et al. Pepper-Pot based Emittance Measurements for the AWA Photoinjector. *Particle Accelerator Conference*, pages 4393–4395, 2007.

[97] D. Stratakis et al. Phase Space Tomography of Beams With Extreme Space Charge. *Particle Accelerator Conference*, pages 2025–2029, 2007.

[98] F. Löhl. *Measurements of the Transverse Emittance at the VUV-FEL*. Diploma thesis, University of Hamburg, DESY, Hamburg, May 2005.

[99] O. Sander et al. Beam Tomography in two and four dimensions. *Linear Accelerator Conference*, 1979.

[100] M. Geitz et al. Phase Space Tomography at the Tesla Test Facility. *Particle Accelerator Conference*, pages 2175–2177, 1999.

[101] D. Stratakis et al. Tomography as a diagnostic tool for phase space mapping of intense particle beams. *Physical Review Special Topics - Accelerators amd Beams*, 9, 2006.

[102] D.J. Holder, B.D. Muratori, S. Khodyachykh, A. Oppelt, F.E. Hannon. A Phase Space Tomography Diagnostic for PITZ. *European Particle Accelerator Conference*, pages 1091–1093, 2006.

[103] D. Stratakis et al. Tomography phase-space mapping of intense particle beams. *Physics of Plasmas*, 14, 2007.

[104] S. Hancock et al. Longitudinal Phase Space Tomography with Space Charge. *European Particle Accelerator Conference*, 2000.

[105] C.T. Mottershead et al. Maximum Entropy Beam Diagnostic Tomography. *Particle Accelerator Conference*, page 1970, 1985.

[106] G. Asova et al. Phase Space Tomography Diangostics at the PITZ Facility. *International Computational Accelerator Physics Conference*, page 194, 2006.

[107] R. Hajima et al. Reconstruction of electron beam distribution in phase space by using parallel maximum entropy method. *Nuclear Instruments and Methods, A*, 389:65–68, 1997.

[108] H. Zen et al. Experimental Study on Effect of Energy Distribution on Transverse Phase Space Tomography. *Free Electron Laser Conference*, pages 668–670, 2006.

[109] L.A. Shepp and Y. Vardi. Maximum Likelihood Reconstruction for Emission Tomography. *IEEE Transactions on Medical Imaging*, MI-1(2), October 1982.

[110] S. Alenius. Generalization of Median Root Prior Reconstruction. *IEEE Transactions on Medical Imaging*, 21(11), November 2002.

[111] A.W. Chao, M Tigner. *Handbook of Accelerator Physics and Engineering*. World Scientific, 1999.

[112] M. Reiser. *The Theory and Design of Charged Particle Beams*. Wiley-VCH, 2004.

[113] J.D. Lawson. *The Physics of Charged-Particle Beams*. Oxford Science Publications, 1988.

[114] I.M. Kapchinskij and V.V. Vladimirskij. Limitations of Proton Beam Current in a Strong Focusing Linear Accelerator Associated with the Beam Space Charge. *International Conference on High Energy Accelerators*, 1959.

[115] D. Ouzounov et al. The Laser System for the ERL Electron Source at Cornell University. *Particle Accelerator Conference*, pages 530–532, 2007.

[116] S. Zhou et al. Efficient temporal shaping of ultrashort pulses with birefringent crystals. *Applied Optics*, 2007.

[117] E. Pozdeyev et al. Multipass Beam Breakup in Energy Recovery Linacs. *Nuclear Instruments and Methods A*, 2005.

[118] A. Todd et al. High Power Electron Beam Injectors For 100kW Free-Electron Lasers. *Particle Accelerator Conference*, 2003.

[119] J. Rathke et al. Design and Fabrication of an FEL Injector Cryomodule. *Particle Accelerator Conference*, 2005.

[120] Advanced Energy Systems. Superconducting RF Injector for High-Power Free-Electron Lasers (FEL). Technical report, 2004.

[121] H. Lui et al. Design of a High Charge CW Photocathode Injector Test Stand at CEBAF. *European Particle Accelerator Conference*, page 942, 1996.

[122] S. Benson and M. Shinn. Development of an Accelerator Ready Photocathode Drive Laser at CEBAF. *European Particle Accelerator Conference*, pages 1052 – 1054, 1996.

[123] I.V. Bazarov and C.K. Sinclair. Multivariate optimization of a high brightness dc gun photoinjector. *Physical Review Special Topics - Accelerators and Beams*, 8(034202), 2005.

[124] Kalyanmoy Deb. *Multi-Objective Optimization using Evolutionary Algorithms*. John Wiley & Sons, 2001.

[125] S. Bleuler, M. Laumanns, L. Thiele, and E. Zitzler. PISA A Platform and Programming Language Independent Interface for Search Algorithms. *Evolutionary Multi-Criterion Optimization International Conference*, pages 494–508, 2003.

[126] http://www.tik.ee.ethz.ch/sop/pisa/. Website.

[127] E. Zitzler and M. Laumanns and L. Thiele. SPEA2: Improving the Strength Pareto Evolutionary Algorithm. Technical Report 103, Computer Engineering and Networks Laboratory (TIK), Swiss Federal Institute of Technology (ETH) Zurich, Gloriastrasse 35, CH-8092 Zurich, Switzerland, May 2001.

[128] P. Bruyant. Analytic and Iterative Reconstruction Algorithms in SPECT. *Journal of Nuclear Medicine*, 43(10):1343, 2002.