A Laser Based Straightness Monitor for a Prototype Automated Linear Collider Tunnel Surveying System

Gregory Richard Moss

Jesus College, Oxford

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy
at the University of Oxford

Michaelmas Term, 2013

Abstract

A Laser Based Straightness Monitor for a Prototype Automated Linear Collider Tunnel Surveying System

Gregory Richard Moss

Jesus College, Oxford

Thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Michaelmas Term, 2013

For precise measurement of new TeV-scale physics and precision studies of the Higgs Boson, a new lepton collider is required. To enable meaningful analysis, a centre of mass energy of 500GeV and luminosity of $10^{34}\text{cm}^{-2}\text{s}^{-1}$ is needed. The planned 31km long International Linear Collider is capable of meeting these targets, requiring a final emittance of 10μm rad horizontally and 35nm rad vertically. To achieve these demanding emittance values, the accelerator components in the main linacs must be aligned against an accurately mapped network of reference markers along the entire tunnel. An automated system could map this tunnel network quickly, accurately, safely and repeatedly; the Linear Collider Alignment and Survey (LiCAS) Rapid Tunnel Reference Surveyor (RTRS) is a working prototype of such a system.

The LiCAS RTRS is a train of measurement units that accurately locate regularly spaced retro-reflector markers using Frequency Scanning Interferometry (FSI). The unit locations with respect to each other are precisely reconstructed using a Laser Straightness Monitor (LSM) and tilt sensor system, along with a system of internal FSI lines. The design, commissioning, practical usage, calibration, and reconstruction performance of the LSM is addressed in this work.

The commissioned RTRS is described and the properties of the LSM components are investigated in detail. A method of finding the position of laser beam spots on the LSM cameras is developed, along with a process of combining individual spot positions into a more robust measurement compatible with the data from other sub-systems. Laser beam propagation along the LSM is modelled and a robust method of reconstructing CCD beam spot position measurements into positions and orientations of the LSM units is described. A method of calibrating LSM units using an external witness system is presented, along with a way of using the overdetermined nature of the LSM to improve calibration constant errors by including data taken from unwitnessed runs.

The reconstruction uncertainty, inclusive of both statistical and systematic effects, of the LSM system is found to be of $5.8\mu\text{m} \times 5.3\mu\text{m}$ in lateral translations and $27.6\mu\text{rad} \times 34.1\mu\text{rad}$ in rotations perpendicular to the beam, with an uncertainty of 51.1μrad in rotations around the beam coming from a tilt-sensor arrangement.

Acknowledgements

I would like to express my gratitude to the many people who have been invaluable along the winding journey to the completion of this work.

I am grateful for Armin Reichold's unwavering support, instruction and guidance through the years as the project came to fruition. I am heavily indebted Patrick Brockill to for his unending enthusiasm while deploying his exceptional analytical talents to help me in any theoretical issue I encountered, while I thank John Dale for his commitment to grinding out difficult results while remaining good company throughout. I am grateful to Paul Coe for his sharp eyes and mind when helping me polish my work, and I'm appreciative of Matt Warden, Andrew Lancaster and all my other friends in physics who I've shared my time with.

I would like to thank Tony Handford and Mike Tacon and everyone involved in the mechanical construction of the RTRS. I also appreciate the careful design and construction of the electronic components by Mark Jones, Roy Wastie and Mike Dawson from central electronics. I recognise the guidance, instruction and technical assistance all of us received from Markus Schlösser, and for the efforts and achievements of Johannes Prenting in co-ordinating the effort at DESY.

I appreciate the diligence of Sue Geddes and Kim Proudfoot in smoothing out all administrative bumps in my way, and I am glad to have had Russell Allcock, Richard Smith and everyone else working behind the scenes to keep the computing systems running smoothly.

I treasure the support from all my friends and colleagues, especially Lisa and Riona who kept me strong in tough times.

Finally, I would like to express my love for my family, who always trust and believe in me.

Contents

1 Physical Motivation

This chapter provides an introduction to the International Linear Collider and the importance of its accurate alignment, finishing with the target uncertainty of the reference network.

1.1 The Standard Model

The Standard Model [1] is the great achievement of 20th century particle physics; it explains the myriad of particles observed in nature and three of the four fundamental forces of interaction. All the components of the Standard Model have been experimentally verified, including the Higgs mechanism with the existence of its scalar boson finally confirmed with a mass around $125\,{}^{\mathrm{GeV}}\!/_{c^2}$ in March 2013 [2].

1.2 Physics at the TeV Scale

The Standard Model has proved to be incredibly robust and has remained effectively unchanged for more than thirty years. However, while the Standard Model is self-consistent and has been extremely successful at matching experimental results to date, it is not complete. The Standard Model does not include gravity and is incompatible with General Relativity in extreme conditions. Furthermore, there is no good dark matter candidate and no credible treatment of the hypothesised dark energy, which together account for 95% of the mass-energy of the universe [3]. Another issue is the prevalence of matter over antimatter in the universe, which the Standard Model does not explain. More subtle issues are related to the 'insertion by hand' of its nineteen parameter values, and the 'fine-tuning' of some parameters required to match observations from the universe we live in.

It is only now that we are beginning to achieve the experimental conditions and precisions where the Standard Model may start to break down. There are a wide number of possibilities for what lies beyond the Standard Model. The Large Hadron Collider (LHC) has broken new ground in finding the Higgs boson and will observe its properties over the coming years, as well as probing a full range of new possibilities, such as extra dimensions, Supersymmetry, CP violation, and the behaviour of quark-gluon plasma.

1.3 The International Linear Collider

The hadronic nature of the collisions in the LHC mean it is an excellent machine for discovery, but many precision measurements will need to be made by a complementary leptonic machine.

This machine is the International Linear Collider.

The collisions of the composite protons in the LHC produce large numbers of particles from hadronic background interactions, complicating analysis. As only individual components of each proton interact in a collision, the collisions occur over a wide range of centre-of-mass energies and typically have a large momentum component along the beam line; this complicates analysis and limits detector design. The hadronic interaction couplings also lead to large cross-sections for uninteresting interactions: careful triggering is required to exclude all but one in a million events before any analysis is possible. Furthermore, there is no control over the polarisation, limiting studies in this area.

None of the above problems are of significant concern for the ILC: the e^+e^- collisions that would occur would have precisely controlled, tunable properties, producing cleaner and relatively easy to analyse interactions; this would enable a vast range of physics to be examined [4].

1.3.1 Need for a Linear Machine

A circular machine cannot practically be used for electron-positron collisions at centre of mass energies much above 200GeV due to the overwhelming energy loss by synchrotron radiation. Synchrotron radiation is the emittance of quanta of light due to transverse acceleration of a charged particle[1]. The power radiated by an accelerating particle is given by

$$P_\gamma = \frac{1}{6\pi\epsilon_0}\frac{e^2a^2}{c^3}\gamma^4 , \tag{1.1}$$

where P_γ is the radiated power, a is the (radial) acceleration of the particle and γ is the relativistic dilation [5, p111-114]. The factor $\gamma^4 \equiv (E/m_0c^2)^4$ is the critical component of the equation: for protons, synchrotron radiation is of little concern, but for the much lighter electrons it rapidly becomes prohibitive: for a hypothetical 250km circumference super-LEP, a 500GeV particle would radiate 220GeV per turn, giving a total radiated power of 100MW for a typical number of particles!

Therefore a circular electron-positron collider of the energies required to make new observations is impractical; a linear collider is the only currently feasible solution. Such a collider design comes with its own difficulties, though these are not insurmountable.

---

[1] Analogous radiation resulting from a change in the particle's speed is called bremsstrahlung.

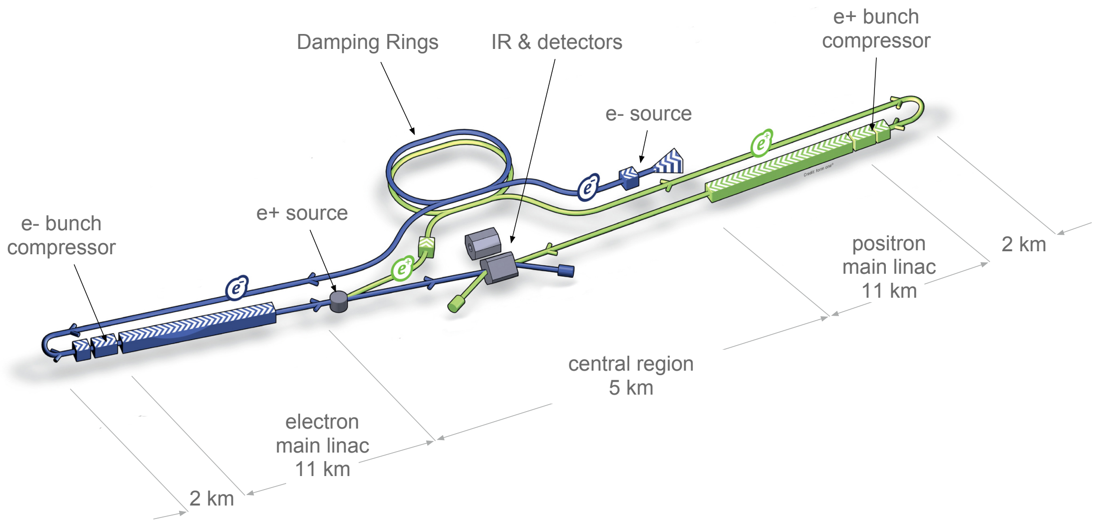


Figure 1.1: Schematic layout of the ILC from the Technical Design Report [6, p9].

1.3.2 ILC Design

With a linear collider the particles only get one chance to collide and must get to their collision energy in one pass through the machine. Accelerating gradients must be very high ($>30\,\mathrm{MV}/\mathrm{m}$) to produce a machine of a reasonable size and in order to maintain a reasonable luminosity the beam cross-section must be very small at the interaction point. This second point, combined with the inability to cool the beam once it is in the main linac, puts very tight requirements on the alignment of the components of the machine.

The required specifications for an effective linear collider have been extensively studied by the Global Design Effort and a construction-ready technical design with a centre of mass energy of 500GeV has been produced [6], including an analysis of the spectrum of physics that would be possible to probe with such a device. Figure 1.1 gives a schematic layout of the current design.

The process to produce electron-positron collisions at the interaction point (IP) is as follows:

1. A beam of electrons with 80% polarisation is produced at the electron source near the centre of the collider. The beam is bunched and accelerated to 76MeV.

2. The electrons are accelerated to 5GeV in a superconducting linac and fed into the central damping ring.

3. The damping rings cool the beams, reducing the emittances by up to five orders of magnitude. They also allow active compensation of pulse to pulse variations in the beam properties.

4. The electrons enter the Ring to Main Linac section which transports the beam to the start of the main linac. It also longitudinally compresses the bunches by a factor of 30 to 45,

collimates the beam halo and rotates the beam polarisation to any desired angle.

5. The electrons enter the main linac with an energy of 15GeV where they are accelerated to 250 GeV using 1.3GHz superconducting RF cavities powered by conventional klystrons.

6. The electrons exit the main linac after 11km and enter the 2.2km long Beam Delivery System which focuses the beam, measures its properties and collides it with the oncoming positron beam.

The positrons are processed in an almost identical way in the opposite direction, though are generated from pair production from photons obtained by passing the main electron beam over an undulator, rather than the photocathode gun used for the electrons. This results in a machine that is about 31km long with a centre-of-mass energy range of 200-500GeV, but with the potential to double this with a future upgrade. The comprehensive baseline design for the ILC can be found in the Volume 3.II of the Technical Design Report [7].

1.3.3 Accelerator Parameters

Table 1.1 gives some parameters of interest for different configurations of the ILC. There are two parameters that are critical: the first is the centre of mass energy which dictates what interactions are possible and their respective rates, while the other is the luminosity; this is an indication of *how many* collisions will be produced per unit time per unit cross-section. For comparison, LEP achieved 209GeV centre-of-mass energy with a peak luminosity of $0.01 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$ using beams with dimensions $200,000$nm by $2,000$nm [8].

The luminosity of a collider is given by the following expression [9]:

$$\mathcal{L} = \frac{N_1 N_2 f n_b}{4\pi\sigma_x\sigma_y} \cdot W \cdot e^{\frac{B^2}{A}} \cdot S \cdot \mathcal{H}_D, \tag{1.2}$$

where N_1 is the number of particles in a bunch of the first beam, N_2 is the number of particles in a bunch of the second beam, f is the number of bunch trains per second, n_b is the number of bunches in a bunch train, and σ_x & σ_y are the r.m.s. beam dimensions at the IP. W is a reduction factor due to beam offsets relative to each other, while S is a reduction due to the crossing angle of the beams. $e^{\frac{B^2}{A}}$ is a compound effect of the crossing angle *and* a relative offset, where A and B are functions of the crossing angle, beam offset, beam width in the crossing plane and bunch length. $\mathcal{H}_\mathcal{D}$ is the disruption/enhancement parameter (which describes the pinching effect of beam-beam interactions).

Table 1.1: Some ILC parameters [6, p11] for the first phase configuration, a full energy baseline configuration, and two potential upgrades. For the upgrades, improvements from the baseline parameters are highlighted.

Parameter	Symbol	Units	1st Stage	Baseline	$\mathcal{L}$ Upgrade	E_{CM} Upgrade A
Centre-of-mass energy	E_{CM}	GeV	250	500	500	**1000**
Bunch train collision rate	f_{rep}	Hz	5	5	5	4
Number of bunches	n_b		1312	1312	**2625**	**2450**
Bunch population	N	$\times 10^{10}$	2	2	2	1.74
Main linac average gradient	G_a	MVm^{-1}	31.5	31.5	31.5	**38.2**
Estimated AC power	P_{AC}	MW	129	163	**204**	**300**
Electron RMS energy spread	$\Delta p/p$	%	0.190	0.124	0.124	**0.083**
Positron RMS energy spread	$\Delta p/p$	%	0.152	0.070	0.070	**0.043**
Electron polarisation	P_-	%	80	80	80	80
Positron polarisation	P_+	%	30	30	30	20
Horizontal emittance	$\gamma\epsilon_x$	nm	10000	10000	10000	10000
Vertical emittance	$\gamma\epsilon_y$	nm	35	35	35	**30**
Horizontal beta function at IP	β_x^*	mm	13.0	11.0	11.0	22.6
Vertical beta function at IP	β_y^*	mm	0.41	0.48	0.48	**0.23**
RMS horizontal beam size at IP	σ_x^*	nm	729	474	474	481
RMS vertical beam size at IP	σ_y^*	nm	7.7	5.9	5.9	**2.8**
Luminosity	$\mathcal{L}$	10^{34}cm^{-2}s^{-1}	0.75	1.8	**3.6**	**3.6**

Due to the very small cross-sections of interesting events, the luminosity must be of order 10^{34}cm^{-2}s^{-1} to give a significant number of interactions over the lifetime of the machine, which is considered around five to ten years. The beam size in a single dimension is simply given by:

$$\sigma = \sqrt{\epsilon\beta(z)}, \tag{1.3}$$

where σ is the beam r.ms. width, ϵ is the beam emittance and $\beta(z)$ is the machine beta function [5, p31] at distance z in the accelerator. $\beta(z)$ is in theory only a function of the position in the accelerator[2], while ϵ is in theory independent and is a constant property of a bunch unless there are velocity dependent dissipative forces present.

Accelerators are designed to have very small beta functions at the interaction points but it is just as important to reduce the emittance to as small a value as practically possible.

1.3.4 Beam Shape

Parameter $\mathcal{H_D}$ in equation 1.2 is also highly important. At low bunch charge densities its value is unity, but as the charge density of the beams increases, the electromagnetic field of the opposing bunch attracts the particles inwards. While this pinching effect may seem beneficial, the acceleration causes radiation of *beamstrahlung* photons which causes an energy spread in the

---

[2] $\beta(z)$ is only a property of the geometry of the machine for a closed loop; for a linear collider, the properties of the beam at the input are relevant.

particles of the bunches and pair creation background [9, p63]. Care must be taken to minimise these beam-beam interactions as if this spread reaches about 20-30%, the desirable property that the incoming particle energies are known is lost, limiting the types of analysis available [5, p167].

The ILC enables large numbers of particles in a bunch without excessive beam-beam effects by having flat beams, with their width being a hundred times their height; as essentially all the steering happens in the horizontal plane, this is the logical orientation. Spreading the charge cloud out minimises beam-beam effects, but puts very tight emittance limits in the vertical direction.

1.4 Emittance

1.4.1 Emittance Definition

Consider a single transverse axis at a point in the ring of a circular particle accelerator with a single particle orbiting. Each orbit, the particle will have a different transverse position at this point, as well as a different transverse momentum. Over time, plotting these values against each other will trace out an ellipse. The ellipse changes shape as a function of position in the accelerator but, barring external effects, the ellipses for the same particle will all have the same area. This constant area ellipse, the equation of which is the Courant-Snyder invariant, is a consequence of Liouville's Theorem [5, p45-48].

When considering a bunch, different particles will trace concentric ellipses of the same shape but different sizes, down to a particle on the ideal orbit which is a point at centre. For an electron beam, which typically has a Gaussian distribution, the 'beam emittance' is defined as the area of the ellipse that has the width of one standard deviation. (i.e. one that contains 68.2% of particles.) The dimensions of the ellipse are simply related to the beta function at that point and the emittance of the particle according to equation 1.3, meaning the area of the phase-space ellipse is π times the beam emittance.

1.4.2 Non-conservation of Emittance

One caveat is that relativistic dilation affects Liouville's Theorem, meaning $\epsilon^* = \beta\left(v\right)\gamma\left(v\right)\epsilon$ is the value conserved as the energy of the beam is changed: a process called adiabatic damping. Therefore, for highly relativistic beams, the transverse beam size shrinks by a factor $\sim \sqrt{E}$ as the energy is increased.

As noted in section 1.3.1, electrons radiate quanta of radiation when accelerated. This means Liouville's theorem no longer holds and, for a circular machine, the beam emittance is determined by an equilibrium between the higher emittance particles radiating more (cooling) and the random jumps in momentum caused by the quantised nature of the radiation (heating). This effect is used to reduce the beam emittance by up to five orders of magnitude in the ILC damping rings before they are transferred to the main linac [6, p15].

1.4.3 Causes of Emittance Growth in the Main Linac

Other causes of emittance growth are related to accelerations on particles caused by various effects, such as kicks caused by misaligned components and the effect of the fields generated by the particles in the bunches. The primary sources of emittance growth identified and considered in the ILC Reference Design Report [10, Section 2.8.2] are:

- Dispersion: If quadrupoles are laterally misaligned, then the particles experience a transverse ‘kick’ which increases the emittance due to dispersion.
- X-Y Coupling: If a quadrupole is skewed then the horizontal (X) and vertical (Y) emittances will be mixed slightly, which is an issue due to the large asymmetry of the beam emittances.
- Cavity tilts: If an accelerating cavity is tilted then a transverse kick would be applied to the beam.
- Single-bunch wakefields: Mirror charges in the beam-pipe cause a wakefield which can slightly affect the tail of a bunch.
- Multi-bunch wakefields (Higher-order modes): Similarly, wakefields from leading bunches can affect trailing bunches, potentially causing variation in the path taken by different bunches in the bunch-train.

The beam-beam effects can be minimised by careful design of the machine, while the machine-beam effects can be minimised by precise alignment of the machine components and beam based tuning.

1.5 Component Alignment

There are four steps to produce an aligned machine.

1. The active components of the machine are fiducialised. (The internals such as sensor positions or magnetic fields are mapped out relative to ‘fiducial markers’, which are externally visible.)
2. The tunnel is surveyed and a reference network is produced.

3. The components of the accelerator are installed into the tunnel and aligned, by combining measurements of their fiducial markers with respect to the tunnel reference network.
4. When the accelerator is functional, beam based alignment methods are used to make final adjustments and to find optimal beam trajectories.

It is the reference network production of step 2 that is under consideration in this work, although there is the potential for similar automation to be applied to the component installation of step 3.

1.5.1 Reference Network Requirements

The uncertainty of the reference network is critically important, though the exact precision and accuracy required of the reconstruction is difficult to calculate. Initially, a working requirement of a 1σ deviation of the markers of the reference network to be within 200 μm vertically and 500 μm horizontally from their nominal positions over any 600 metre length of tunnel was used, taken from the TESLA Technical Design Report [11]. Later work [12, p167-169] suggests that error curves over the whole tunnel network, rather than just for an arbitrary section, need to be considered: differing combinations of residuals producing almost identical error curves over 600 metres produced acceptable and unacceptable dispersion matched steering performance. The requirement in the TDR uses a 200μm per 600m requirement for the misalignment of component fiducials [7, p238]. This is possibly conservative, but given the difficulty in assessing the precise requirements, it is a sensible target.

2 LiCAS Design and Measurement Principles

This chapter describes a proposed automated survey system to produce a tunnel reference network. The general principles of network measurements are introduced. Current alignment methods and the potential of automation are discussed. The automated solution proposed by the LiCAS group is a train that travels the length of the tunnel making measurements of regularly spaced markers on the walls. The mechanisms and physical principles of the systems used inside the train are explained.

2.1 Tunnel Measurement Principles

2.1.1 Conventional Surveying

Surveying is the production of a network of marker points that can then be used as a reference. It is a mature field and there are many different techniques that have been in use for decades [13], including GPS [14]. Readings made using hydrostatic levelling system, laser trackers, photogrammetry, stretched wires, tachometers and GPS can all be combined in a large model to give a final reference marker network. This process involves many measurements of each marker from multiple positions to give a well constrained and low error network, taking many man-hours of work.

With a well defined network of reference points established, accelerator components can be installed correctly by measuring their fiducial markers with reference to the network and adjusting each component until the desired positions and orientations are achieved [15].

2.1.2 Comparison With an Automated System

When taking measurements over large distances significant systematic errors can arise due to the air; any optical measurement will have to deal with motion of the air and thermal gradients. Even a small thermal gradient of 0.1K/m would displace a beam by 4.5 *millimetres* over a distance of 600m [11, partII, p250-256]. The proposed RTRS negates this problem by carrying its own portable vacuum.

Another of the difficulties with conventional methods is that they take a significant amount of time to perform and require many skilled operators. An automated system would drastically cut the amount of time taken to complete the survey and hence decrease the amount of accelerator down-time significantly [12, p20]. Similarly, an automated system could run while the machine is operating, without risk to its operators from radiation. This has the potential to transform the network measurement from an occasional, intrusive process to a continuous, background process - provided the maintenance time for the instrument itself is small enough.

Finally with a suitable upgrade and some standardisation, the RTRS could measure accelerator components in addition to the tunnel network, aiding installation, as well as monitoring the change over time.

2.2 Proposed Automated Alignment System

2.2.1 Method

The measurements made by the Rapid Tunnel Reference Surveyor (RTRS) can be used to reconstruct the positions of the markers all the way along the tunnel. Figure 2.1 gives the principles of operation.

1. Retro-reflectors are placed at regular intervals along the tunnel wall.
2. A train consisting of N units is positioned at the start of the tunnel.
3. Each unit of the RTRS is positioned so it can measure one retro reflector, while maintaining the internal alignment of all units with each other.
4. The units measure the position of their retro-reflector with respect to themselves.
5. The units measure their position and orientation with respect to their neighbouring units.
6. This accurately gives the position of N markers with respect to each other, in the local co-ordinates of the train.
7. The train moves on one step, measuring one new marker and the N-1 overlapping markers an additional time.
8. Steps 3-7 are repeated until the entire length of the tunnel is measured.

Each segment can be superimposed on the previous N-2 segments to produce a final chain which links them all together, bringing all measurements to one common frame. To completely constrain all degrees of freedom of the overlap, four cars are needed. Further additional cars provide redundancy to reduce errors and provide self consistency measurements.

In practice, each segment is not analysed individually: the entire set of wall marker measurements can be processed at the same time to produce a final reference network that is more

Table 2.1: The sensitivity of the different subsystems of the RTRS to the three position axes (X, Y & Z; see figure 3.6) and three orientation components (rotations around the three defined axes). A bullet point indicates good sensitivity while ± indicates limited sensitivity.

Subsystem	X	Y	Z	Rx	Ry	Rz
LSM	•	•		•	•	
Internal FSI	±	±	•	±	±	
Tilt Sensors				•		•

self-consistent and has smaller errors. Indeed, there is the potential to directly include the subsystem measurements themselves into a comprehensive model, again performing the minimisation of all elements in one step.

2.2.2 Limitations of Prototype

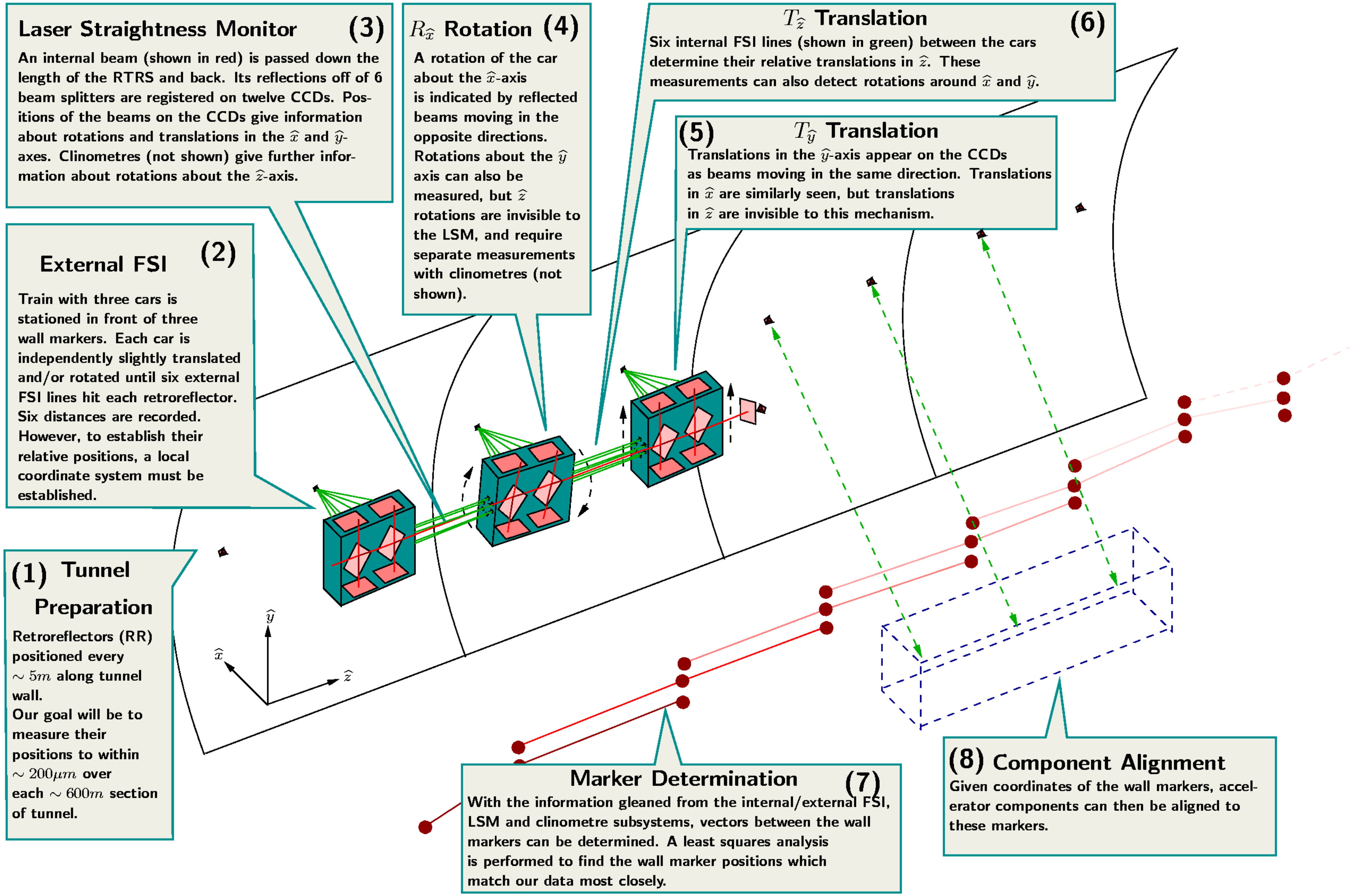
Due to budget and time limitations, the prototype was designed with only three cars. This gives a rotational degree of freedom around the axis along the tunnel, which is only constrained by the long distance comparison of tilt sensor values.

The prototype is only designed to measure a linear arrangement of reference markers, which is not sufficient to create a useful network on its own, though can be combined with other systems like a hydrostatic levelling system [17]. A comprehensive system would need to have further markers positioned to break the linear symmetry; for example on the floor and roof, or with a zig-zag pattern. Adding more external FSI lines to make these measurements in a future iteration would not be a significant extra challenge but is not necessary to demonstrate the procedure.

There was no attempt to include fiducialisation apparatus in the prototype, which is left for a future investigation.

2.3 Techniques Used

The LiCAS group's proposal uses a variety of techniques to enable the measurement of the marker and subsequent reconstruction to produce a usable network. The external frequency scanning interferometry subsystem is used to pinpoint the position of each retro-reflector relative to its corresponding measurement unit. The other three subsystems enable the relative positions and orientations of the three cars to be known. An overview of the system can be found in figure 2.1 while a table summarising the different sensitivities can be found in table 2.1.



Laser Straightness Monitor (3)
An internal beam (shown in red) is passed down the length of the RTRS and back. Its reflections off of 6 beam splitters are registered on twelve CCDs. Positions of the beams on the CCDs give information about rotations and translations in the $\widehat{x}$ and $\widehat{y}$-axes. Clinometres (not shown) give further information about rotations about the $\widehat{z}$-axis.

$R_{\widehat{x}}$ Rotation (4)
A rotation of the car about the $\widehat{x}$-axis is indicated by reflected beams moving in the opposite directions. Rotations about the $\widehat{y}$ axis can also be measured, but $\widehat{z}$ rotations are invisible to the LSM, and require separate measurements with clinometres (not shown).

$T_{\widehat{z}}$ Translation (6)
Six internal FSI lines (shown in green) between the cars determine their relative translations in $\widehat{z}$. These measurements can also detect rotations around $\widehat{x}$ and $\widehat{y}$.

(5) $T_{\widehat{y}}$ Translation
Translations in the $\widehat{y}$-axis appear on the CCDs as beams moving in the same direction. Translations in $\widehat{x}$ are similarly seen, but translations in $\widehat{z}$ are invisible to this mechanism.

External FSI (2)
Train with three cars is stationed in front of three wall markers. Each car is independently slightly translated and/or rotated until six external FSI lines hit each retroreflector. Six distances are recorded. However, to establish their relative positions, a local coordinate system must be established.

(1) Tunnel Preparation
Retroreflectors (RR) positioned every $\sim 5m$ along tunnel wall.
Our goal will be to measure their positions to within $\sim 200\mu m$ over each $\sim 600m$ section of tunnel.

Marker Determination (7)
With the information gleaned from the internal/external FSI, LSM and clinometre subsystems, vectors between the wall markers can be determined. A least squares analysis is performed to find the wall marker positions which match our data most closely.

(8) Component Alignment
Given coordinates of the wall markers, accelerator components can then be aligned to these markers.

Figure 2.1: LiCAS systems overview [16].

2.3.1 Frequency Scanning Interferometry

Frequency Scanning Interferometry (FSI) is a technique that uses the interference between two coherent beams from the same source to make highly accurate distance measurements. When used with a calibrated reference interferometer, it can be used to produce an absolute reading. A detailed analysis of the theory, capabilities and usage of FSI in the RTRS can be found in [18] and [12].

In the RTRS, FSI is the basis of two subsystems. The External FSI system is used to find the position of the wall marker being observed, relative to the observing unit. It consists of a ring of six uncollimated beam launches pointing in the direction of a retro-reflector wall marker, reconstructing the position of the marker in the unit's frame of reference from those measurements, with a precision of between one and three microns after combined minimisation [12, p172]. The Internal FSI system is primarily used to measure the distances between the cars, but has some sensitivity to their relative rotations. Each internal FSI connection consists of a ring of six collimated parallel beams, travelling between launches in one one unit and retro-reflectors mounted in the adjacent unit. The reconstruction precision is impressive at $0.27\mu m$ in Z-translation and even rotations reconstructed to $O(10)\mu rad$, though these values are subject to significant systematic errors [12, p171].

2.3.2 Laser Straightness Monitor

The Laser Straightness Monitor (LSM) enables relative positions of the cars to be found along the X and Y axes. It also enables the rotations of the units to be found around the x and y axes. Historically, stretched wires have been used for this type of purpose, including a similar prototype to the LiCAS RTRS [19]. However, wires suffer from the effects of sagging under their own weight, vibration, and buffeting by air currents.

Using a laser avoids these mechanical issues and uncertainties due to refractive index gradients can be avoided by propagating the beam in a vacuum. The basic principle is to precisely measure the position of a small fraction of the beam reflected via a pellicle beam splitter multiple times in each unit using CCDs. The lateral position and two rotation angles can be calculated from the beam spot positions. The LSM is the focus of this work and is covered in extensive detail in chapter 4.

2.3.3 Tilt Sensors

Rotation around the Z-axis is constrained by using tilt sensors. In addition, there is an extra tilt sensor on each car which measures the rotation around its X-axis which is very similar to the 'beam' X-axis measured by the LSM, allowing for overdetermination and cross-checking. Care must be taken to account for the fact that the tilt sensors are gravity referenced, rather than the local axis system, relative measurements returned by the other sub-systems (§8.6).

3 Implementation

This chapter gives an overview of the RTRS built and developed by the Oxford LiCAS group and introduces the systems required for effective operation.

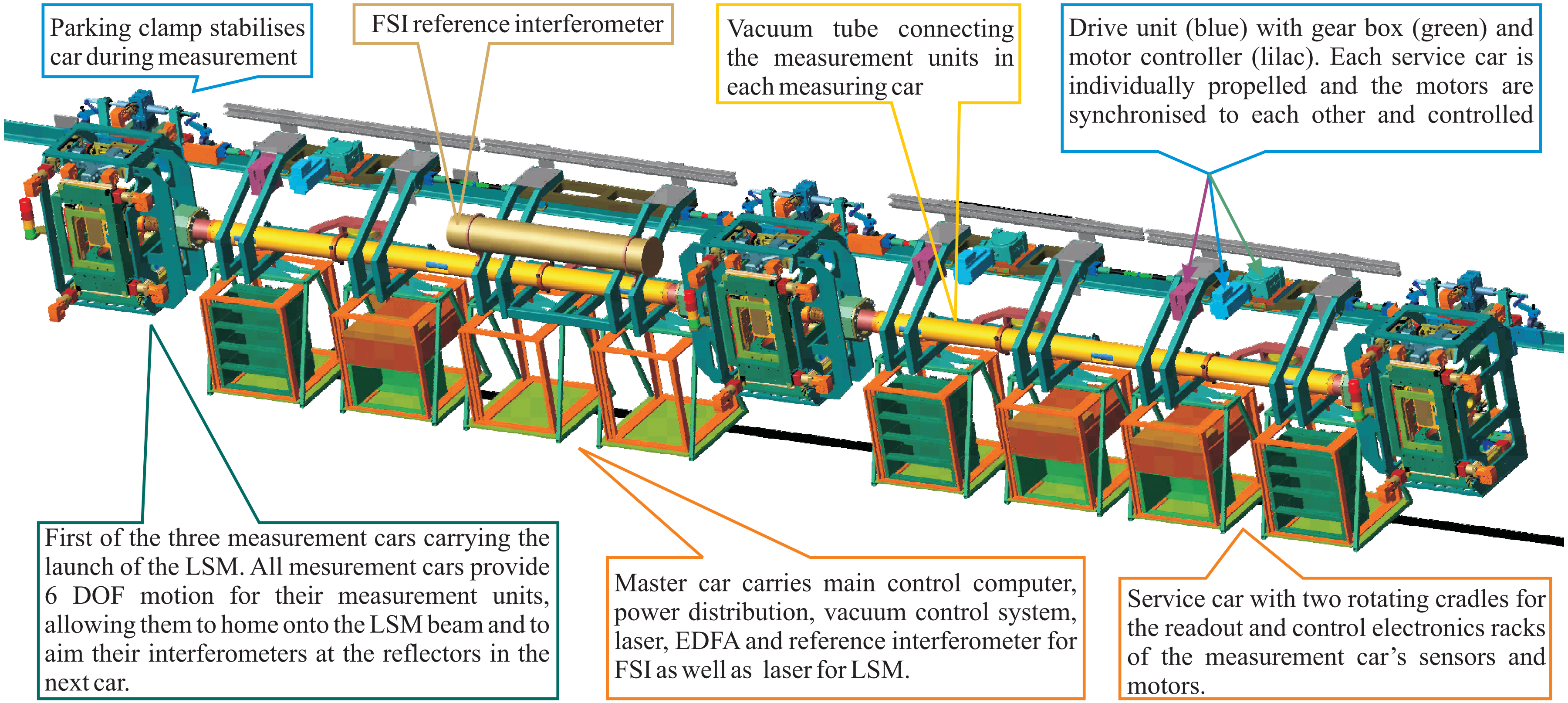
3.1 System Overview

The main design overview can be found in figure 3.1, while a view of the physical RTRS can be found in figure 3.2. The RTRS consists of a train of three powered measurement cars. Each car houses a single measurement unit in a motorised 'cage' which provides 6-D motion. Each measurement unit is capable of measuring both the location of a wall-mounted retro-reflector, and its own position & orientation relative to its nearest neighbours and a common internal laser line. The measurement units are linearly connected via vacuum pipes, given flexibility via a system of bellows and ball-and-socket joints. A vacuum pump is mounted at the end of the train at the base of car 3, which is used to pump the system down to its working vacuum before use. Nearby service cars hold power distribution and electronic equipment, with drive motor assemblies housed above each allowing motion along a monorail. A master car containing interlock electronics, power distribution, interlock relay, the FSI laser and Erbium Doped Fibre Amplifier (EDFA), the pressure sensor and the master control computer is also part of the train. The FSI reference interferometer is mounted on the train above the master car. Exchangeable end-caps contain the launch and retro-reflection optics for the LSM, allowing the LSM beam to travel along the core of the vacuum system. The master control computer is accessed via remote desktop over WiFi, enabling totally remote operation.

3.2 Measurement Units

The three measurement units hold all the instrumentation used to produce data for tunnel reconstruction. Figure 3.3 shows a unit in a clean room ready to have the covering plates put on, with the uncovered internals of the launch end-cap connected.

Each measurement unit houses four cameras and two pellicles for the LSM system, six external FSI lines, six internal FSI launches on one side and six retro-reflectors on the other, two fine tilt sensors, a dual-axis coarse tilt sensor and temperature sensors throughout. Each is machined out of a single block of Invar-36 steel. Invar-36 was chosen as it has an extremely small coefficient of thermal expansion of less than $1.3 \times 10^{-6}\mathrm{K}^{-1}$ [21]. This, in combination with the reasonable thermal conductivity and large thermal mass, causes the geometry of the measurement units to



Parking clamp stabilises car during measurement

FSI reference interferometer

Vacuum tube connecting the measurement units in each measuring car

Drive unit (blue) with gear box (green) and motor controller (lilac). Each service car is individually propelled and the motors are synchronised to each other and controlled

First of the three measurement cars carrying the launch of the LSM. All mesurement cars provide 6 DOF motion for their measurement units, allowing them to home onto the LSM beam and to aim their interferometers at the reflectors in the next car.

Master car carries main control computer, power distribution, vacuum control system, laser, EDFA and reference interferometer for FSI as well as laser for LSM.

Service car with two rotating cradles for the readout and control electronics racks of the measurement car's sensors and motors.

Figure 3.1: Annotated schematic of the RTRS (from [20]). Car 1 is on the far left, car 2 in the centre, and car 3 is on the far right.

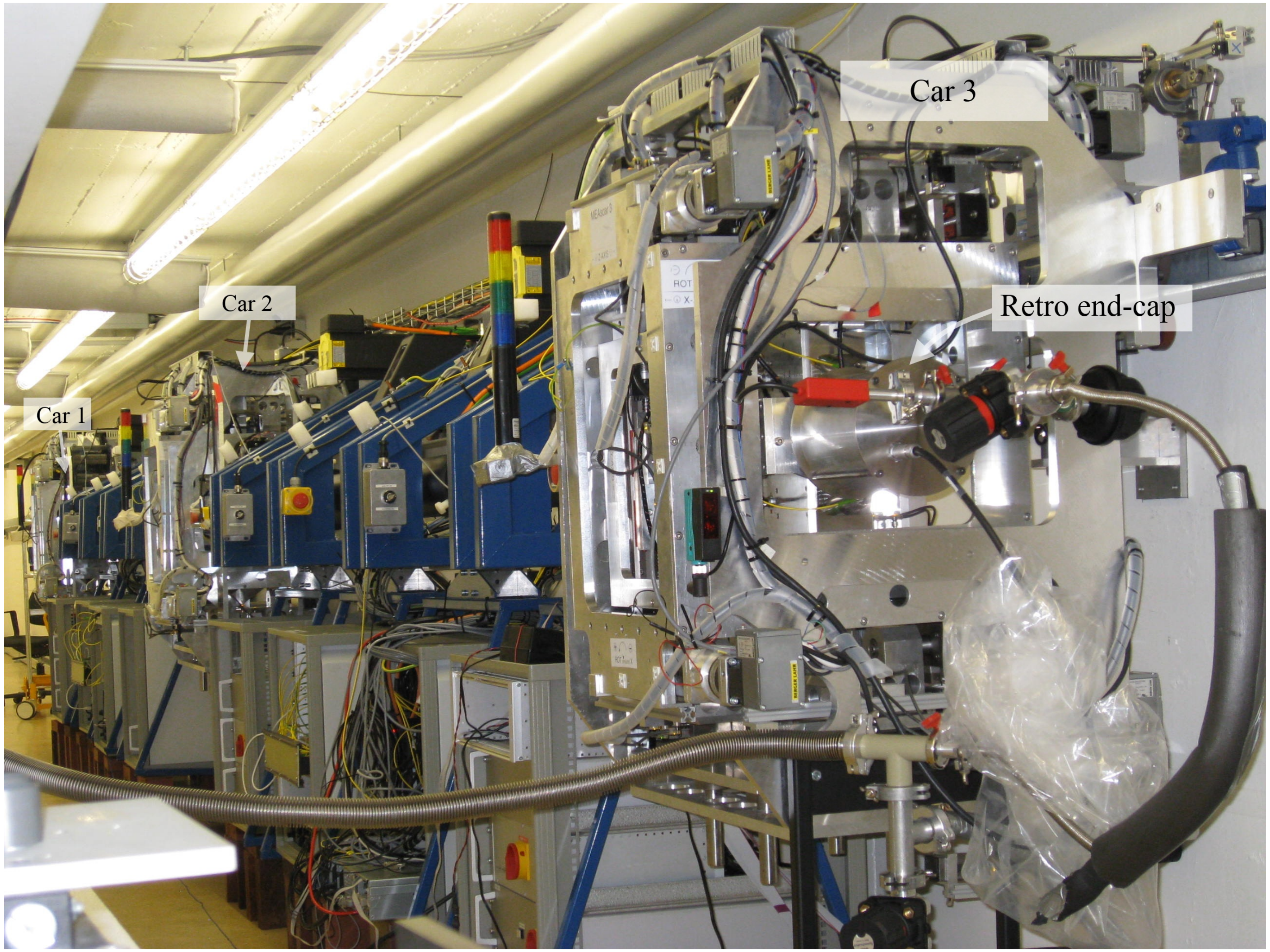


Figure 3.2: View along the train from right hand side of car 3. The aluminium framework of the measurement car 3 motor cage dominates the right-hand side of the image, with the other two measurement cars visible in the background. Connected to the retro-reflector end-cap is the vacuum pump that evacuates the system, with another connection to an external FSI interferometer under calibration. The service cars and master car are seen supported by their blue supports.

change very little as the temperature varies. To maintain the vacuum of less than 1mbar, the numerous access points on each unit are sealed using o-rings, electrical cables are routed through using commercial feed-throughs while the internal FSI fibre feed-throughs were custom made. Figure 3.4 shows the collimators and retro-reflectors installed into each unit that make up the optics of the internal FSI system.

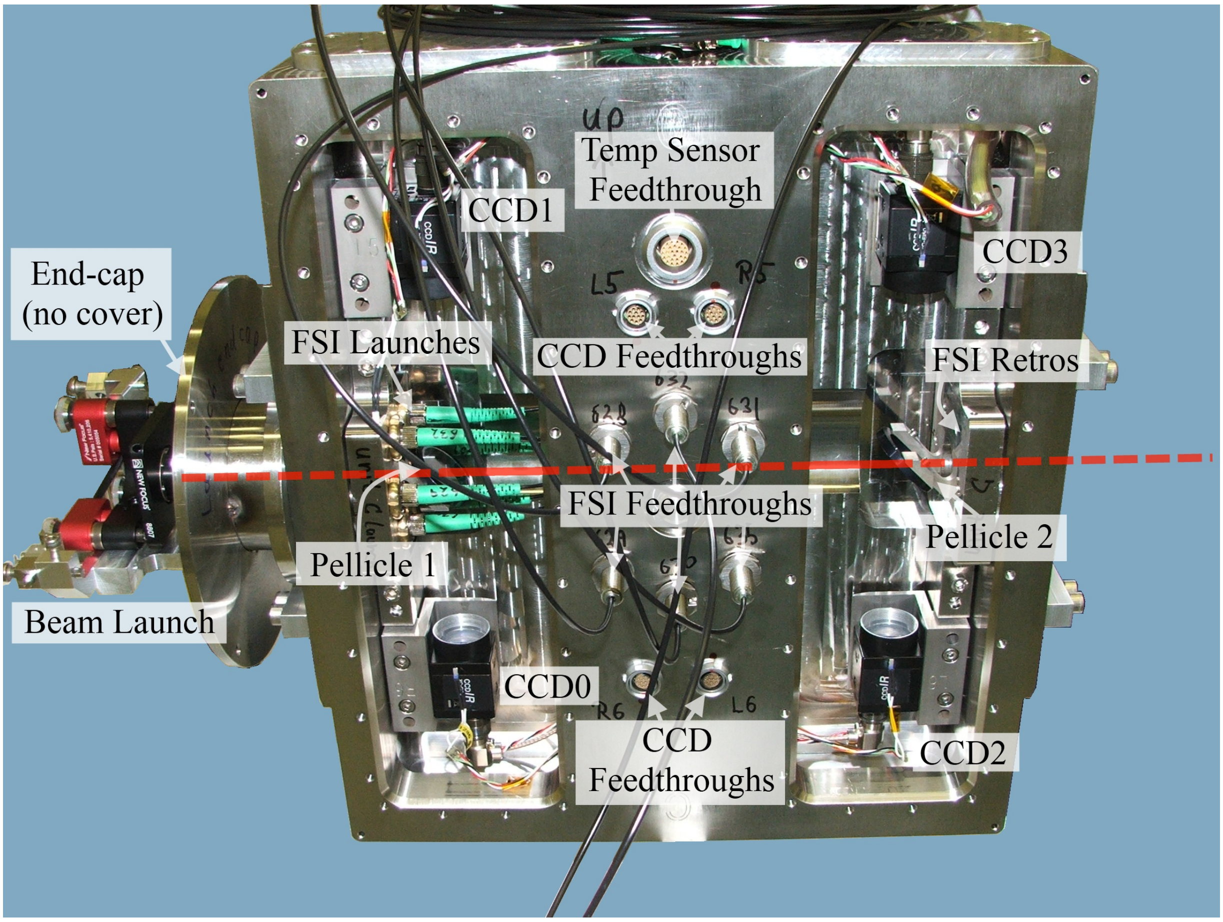


Figure 3.3: Unit 1 during assembly with uncovered launch end-cap fitted. The red ray shows the path taken by the transmitted laser beam. Each pellicle reflects a small amount of the light vertically down for the outgoing beam and up for the anti-collinear return beam.

3.2.1 Tilt Sensors

The coarse, two-axis tilt sensor is only used for crude alignment purposes when the fine sensors are out of range. More important are the two, single axis fine tilt sensors which can be seen installed onto a measurement unit in figure 3.5. These sensors are Schaevitz LSOP/LSOC-1 DC-Operated, Gravity Referenced Inclinometers, shown to be suitable in a similar apparatus [19, p9]. When supplied with a stable DC voltage, provided by batteries, the sensors give output signals of up to 20mA, which are read by a connected Analogue to Digital Converter (ADC, § 3.4).

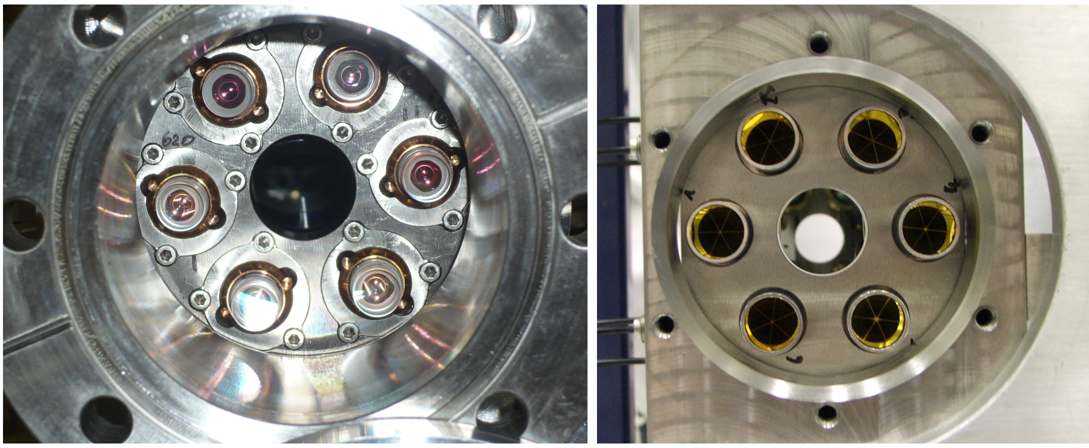


Figure 3.4: Internal FSI optics during assembly. The left-hand image shows the six sets of launch optics for the internal FSI lines with a central hole for the LSM beam. A pellicle can just be made out inside the central hole for the LSM beam. The right-hand image shows the internal retro-reflectors from the other side of the car, again with a central hole.

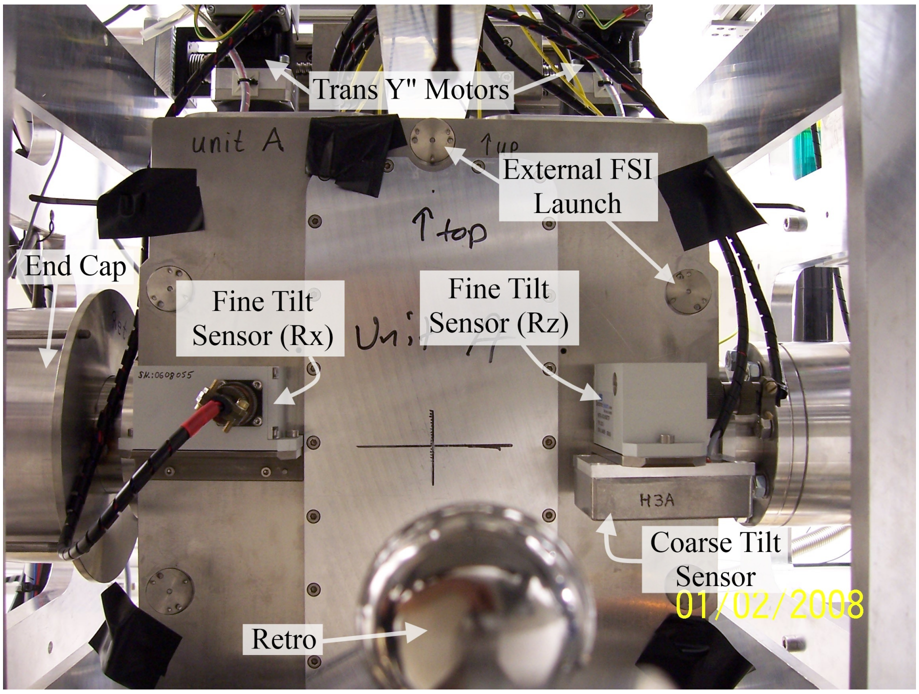


Figure 3.5: The wall facing side of unit A, mounted in car 3. Five of the six external FSI lines for the unit can be seen with the back of the retro-reflector in the foreground. The two fine tilt sensors are half-way up, with the coarse sensor (H3A) under the right-hand fine tilt sensor. Part of the retro-reflector end-cap can be seen on the left of the image, with the rotary vacuum seal connecting the unit to the bellows just off screen to the right. Two stepper motors used for Y'' translations (§3.3.1) can be seen at the top of the image while part of the unit cradle frames the shot.

The specifications [22] give the resolution as 0.1 arc seconds (0.48μrad) and a maximum non-repeatability of 14μrad and a maximum zero offset of 35μrad. Another factor in the measurement uncertainty was the performance of the readout: there were two readouts made using the ELMBs; one 'coarse' measurement with a precision of 45μrad and a range of ±17mrad, and a 'fine' measurement with a precision of 12μrad and a range of 1.7mrad. Comparison of relative tilts with results from the FSI subsystem [12, Section 6.8], which also has a precision of 12μrad, gives a measured combined error of 18.1μrad which is consistent with two component errors.

3.2.2 Temperature Sensors

The temperature was measured throughout the train, to enable the detection and possible correction of temperature dependent effects. There were 105 separate PT100 temperature sensors for this purpose, each with an accuracy of 0.05°C [23] when measured using the ADCs via the four-wire method [24] .

3.3 Measurement Cars

Each measurement unit is housed in its own measurement car, which consists of an aluminium superstructure containing the motors and framework needed to enable 6D motion. Figure 3.6 shows a side view of the train, looking at measurement car 1, while a three-quarters view from the opposite end can be found in figure 3.2.

3.3.1 Adjusting the Units

Accurate, stable, six degrees of freedom motion of the units is achieved using ten linear stepper motors and one rotational stage, powered by the same type of stepper motor. The motors chosen were ICLA IFA Intelligent Compact Drive Fieldbus stepper motors [25]. These motors are equipped with brakes that engage when no current is supplied to them; this enables the unit to be held fast with minimal vibration. The position and orientation of the unit can be understood as being controlled by adjusting three 'shells', with motion of an inner shell occurring in the frame of reference of the outer. These dynamic car-local axes are modelled with intrinsic rotations to understand motion induced by the motors, though are only roughly in the same directions as the lab system eventually defined as part of the calibration (§9.6.7). Axis directions that are dependent on one outer shell rotation are denoted by a prime while a double-prime denotes two dependencies. The axis hierarchy is Trans-X / Rot-Z, Trans-Z$'$ / Rot-X$'$, then Trans-Y$''$ / Rot-Y$''$.

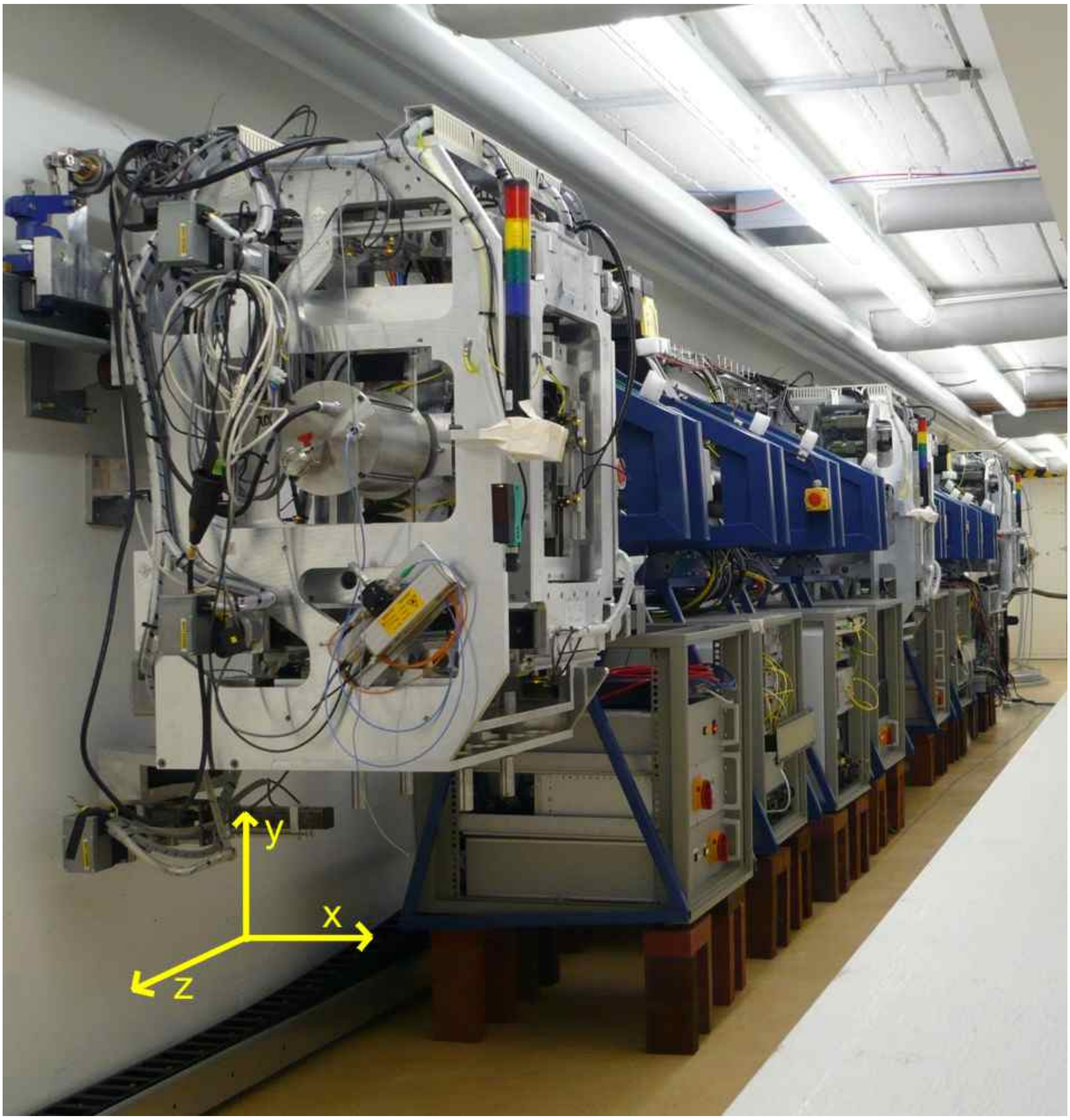


Figure 3.6: View down the train from the left hand side of car 1, dominated by the motor cage framework, The circular launch end-cap can clearly be seen, connected to the the 57FCM infra-red LSM laser by the blue fibre-optic cable. The pillar of lights is used to alert users to the current status of the train, warning them of car alignment or train motion. The bricks supporting the support crates are a temporary installation as part of a stability experiment. The general directions of the lab axis system are shown.

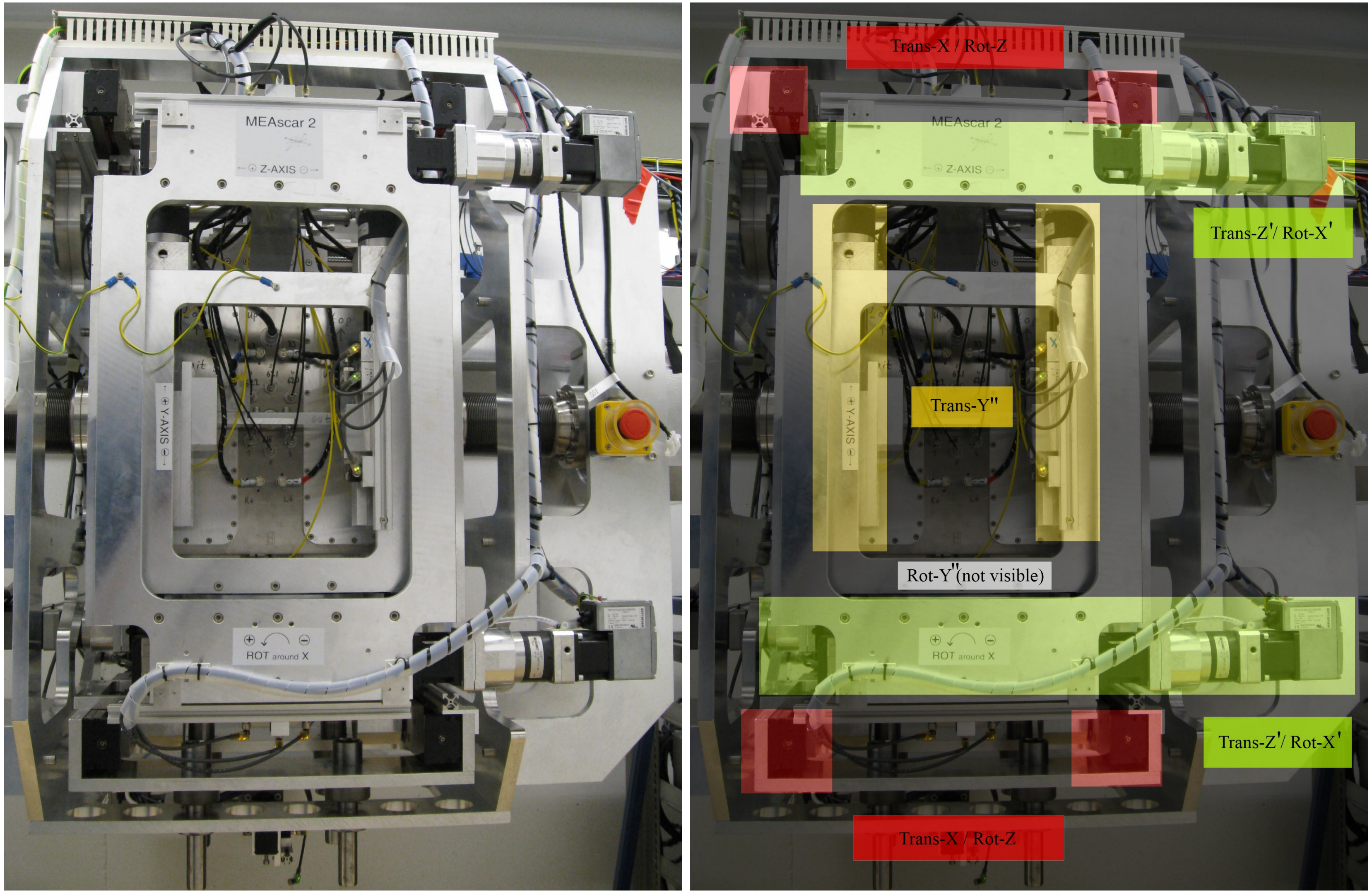


Figure 3.7: Unit 2 mounted inside the motorised cage with the various stepper motors labelled. There are two more Trans-Z′ / Rot-X motors near the wall and the Rot-Y″ turntable is not visible.

The outermost shell is adjusted by four stepper motors perpendicular to the tunnel wall, two at the top of the cage forming a pair and two at the bottom, also forming a pair. Motors in a pair are controlled as one. Thus, by moving both pairs by the same amount in the same direction an X-translation would be achieved, and by moving pairs in the opposite direction to each other a rotation around the Z-axis is produced. Contained inside the outer cage are four motors pointing roughly down the tunnel. When they move together a Z'-translation is produced and when they move in opposing directions a rotation about the X'-axis is seen. Inside the second shell is a rotation stage driven by a single motor; this rotates the unit around the Y''-axis. Finally, mounted on the stage are two motors that give almost vertical motion along the Y''-axis. Figure 3.7 shows the positions of the stepper motors and labels the associated axes.

The stepper motors have a high torque and could do significant damage if they attempt to move the unit too far. To prevent this, there are three layers of protection: first, the control software has soft limits, then there are limit switches to inform the control software of excess motion, finally there are hard physical stops which prevent motion. Due to the compound nature of the movements, particularly the link between a rotation around one axis with a translation along another, the safe hyper-volume is actually quite complicated. The limit switches and soft limits were set so that no combination of motor positions could produce an invalid position. This significantly reduced the range of each unit, making alignment more difficult but the range is still enough to cope with all stations in the tunnel.

3.4 Service Cars

Adjacent to each measurement car is a Service Car which contains a cradle of electronics and a rack mounted data acquisition computer, as well as a second cradle containing an interlocked electrical power supply system for the electronics and the measurement unit.

Each service car contains an Embedded Local Monitor Board (ELMB); a general purpose ADC board designed for use in the ATLAS experiment [26]. For the RTRS they were used to digitise the signals from the vacuum pressure gauge, temperature sensors, fine tilt sensors and picomotor limit switches. They are polled at up to 15Hz over a CANbus.

There is a fibre-optic splitter tree used for the FSI lines, and a crate of custom electronics for measuring the internal and external FSI signals [12, p32].

3.5 Master Car

A Master Car, consisting of two crates, houses much of the remaining infrastructure. The first holds the FSI laser's mains power supply as well as the Interlock Relay Unit. The second houses the master computer, the vacuum pressure monitor and the common FSI components. These are the FSI laser itself, the Erbium Doped Fibre Amplifier (EDFA) and the master splitter tree which then supplies light to the secondary splitter trees in the service cars [12, p40]. Finally, mounted on the train near the master car is the FSI reference interferometer, needed to enable precise absolute distance measurement.

3.6 Launch and Retro End-Caps

End-caps are attached to the outer vacuum flange of the measurement units at either end of the train. The launch end-cap has a fibre feed-through for the optical fibre containing the LSM laser beam. The LSM laser source and attenuator to provide the beam for the launch end-cap is mounted on a plate attached to the measurement car outer framework (figure 3.6). Internally, the fibre is connected to a collimator housed in a New Focus 8807 [27] motorised optical mount, which is used to direct the beam along the core of the train (figure 3.8; the launch is discussed further in § 4.3). There is an an electrical feed-through to send in the end cap for the signals to and from the motorised mount. Finally, there is a vacuum flange which can be used to attach the vacuum pump when the end-cap is mounted in car 3.

The retro end-cap is mounted at the other end of the train. It has an identical vacuum flange and electrical feed-through, but no need for a fibre-optic feed-through. Inside is the retro-reflection assembly designed to reflect the beam back in an anti-collinear fashion (§ 4.5). To enable the required adjustment of the assembly, it is mounted on a New Focus 8071 4-axis motorised stage [27].

Picomotors feature a variable step size that changes depending on conditions like load and direction. Furthermore, picomotors do not have encoders and lock up when they reach the limit of their travel. To avoid this possibility, limit switches were mounted on both the mount and the stage with small rockers attached surrounding the protruding end of the picomotor axle (figure 3.8). These limit switches were monitored using an ELMB, enabling the controlling program to halt motion at a safe limit.

The picomotors are controlled by sending telnet commands from the master car to an 8752 Intelligent Picomotor Ethernet Controller [27] over an ethernet connection. This controller is

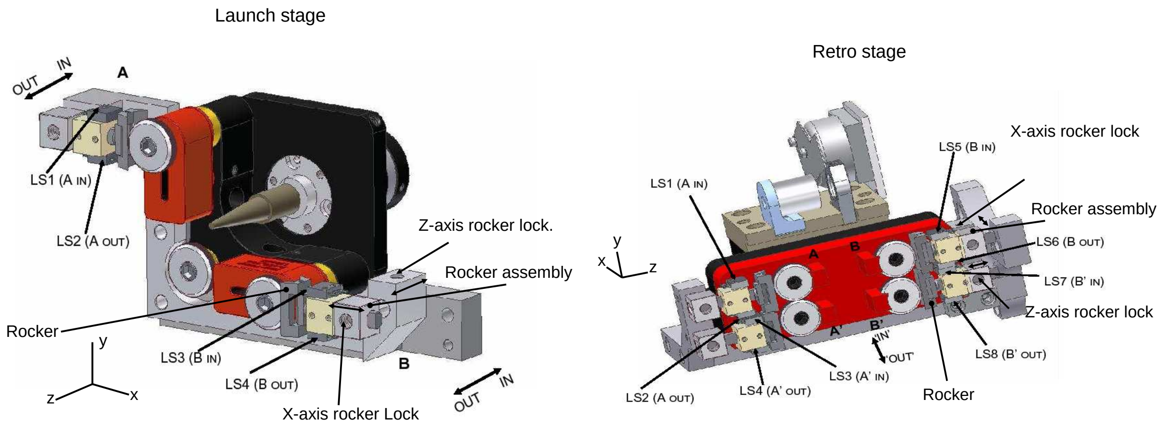


Figure 3.8: Left: Diagram of the LSM launch assembly, with the limit switches labelled. The back of the beam launch collimator is visible with the tan cone representing the fibre-optic cable connection. Right: Side on view of the retro-reflection stage assembly. The picomotors, with the circular heads visible, move in the X-axis, tipping the adjustable rockers when they reach their limit. The 'primed' motors cause the breadboard to move along the Y-axis. The laser beam enters from the right, interacting with the $\frac{1}{2}$ inch semi-mirror before being reflected off the cylindrical retro-reflector. The shutter is shown in the raised position.

connected to three 8753 Intelligent Picomotor Driver units via a daisy chain which in-turn are connected to the individual picomotors through the feed-throughs in the end-caps.

The end-caps are interchangeable and short different pins in the electrical feed-through to enable the system to be aware which configuration is in use. This enables checking for systematic errors by comparing the behaviour with both set-ups, and aids calibration of the normal launch car by allowing it to be calibrated with the end caps swapped to give better CCD coverage.

3.7 Propulsion

To drive the train along the rail, there are three Baldor BSM90B-275AA 1.7kW motors [28], each mounted above the service cars adjacent to the associated measurement car (figure 3.9). The motors run off the 230V mains system, managed by a MINTDriveII Baldor local controller unit. Each motor is connected through a 16:1 gearbox to provide the torque needed to propel the train. The drive wheels do not support the weight of the train; this task is performed by smaller wheels directly underneath. Although each car has a drive wheel encoder, they can be subject to slippage. Therefore, a separate precision encoder wheel is present on car 1 which enables the distance travelled to be measured accurately, allowing the train to stop at the correct positions automatically. This car is configured as the master car and has a RS485 link to the master car computer. Communication between the motors is carried over a CAN bus enabling the three drive assemblies to move synchronously, using torque sharing or trajectory copying. Under normal usage the maximum acceleration is 0.1 m/s^2, maximum deceleration is 0.2 m/s^2 and the top speed is 1 m/s. When stationary, the cars are kept stable by a clamp and a foot that extends to push the lower wheels away from the wall (figure 3.9).

3.8 Vacuum System

The RTRS has a large evacuated volume which is needed for the LSM and internal FSI systems to function accurately and with high degree of precision. The large number of seals, about 90, and the need to accommodate 6D motion of the units, add further complexity. Inside the car, the measurement unit is connected to a bellows via a rotary seal, which enables the unit to move with little resistance (figure 3.9). The other side of the bellows is connected to a ball-and-socket joint on the edge of the car which joins onto the 20cm diameter vacuum tube and its rigid outer shielding. This enables the measurement cars, including the frames, to move and rotate relative to each other while maintaining a fixed separation.

At the extremities of the train, end caps are directly attached to the measurement units. Although cheaper and more flexible than having separate apparatuses, attaching the end-caps directly does have the disadvantage of the vacuum exerting a strong pull of about 30kN inwards. The motorised cages hold this force, but the units suffer a 5mm shift inwards after pumping down.

The vacuum is created using a dry scroll pump affixed to car 3. The pump produces a vacuum of 5.9×10^{-2} millibar but its large vibrations when activated make effective measurements impossible. With the pump off, the leak rate was measured at roughly 0.3×10^{-2} millibar/hour. This is comfortably low enough to not significantly affect the readings on the time-scales of the experiments performed; motion of hundreds of microns caused by air fluctuations would still be reduced by 99.9% after 300 hours.

3.9 Safety

There are a number of areas where steps were taken to ensure safe usage of the train. On the train all systems are linked to the Interlock Relay Unit (IRU) which monitors the safety systems and takes the appropriate actions if a problem is found.

There are a number of safety features that are present to keep the drive system safe. At two points on the train there is an emergency kill switch which will cut all power when pressed. The stepper motor brakes will rapidly engage, while the high gearing of the drive motors leads to

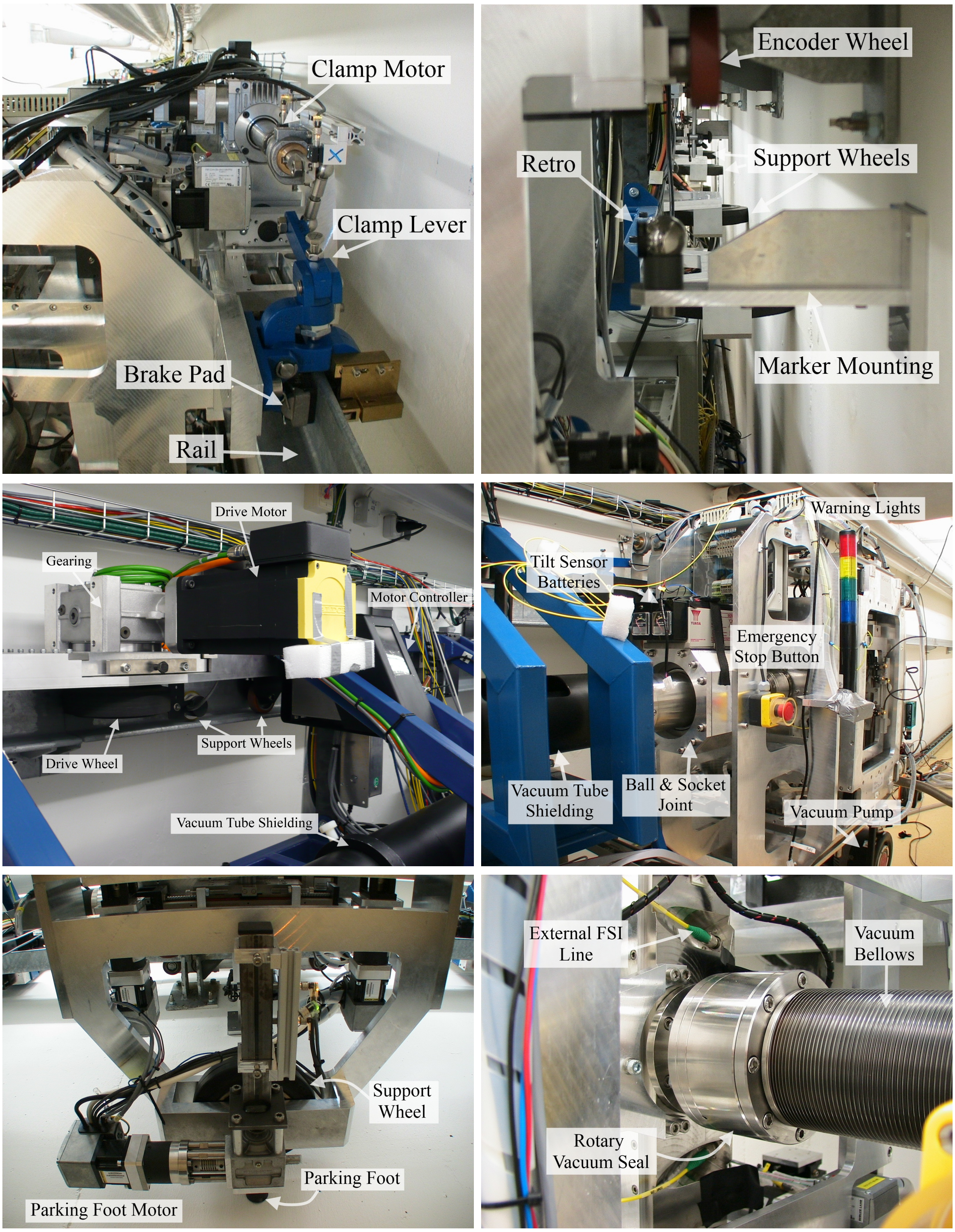


Figure 3.9: Views of various external components. Top-left: A view along the top of the RTRS. When in position the cars are held fast by a clamp which can be seen as the blue apparatus here. Top-right: A view along the wall immediately underneath the rail. A wall-mounted retro-reflector is visible in the foreground, with unpowered support wheels also apparent. Centre-left: One of the three train drive assemblies, with the vacuum tube shielding visible at the bottom right. Centre-right: A view of car 3, showing the vacuum tube connecting to the a ball-and-socket joint in the centre of the image. The vacuum pump can just be seen in the bottom right corner. Bottom-left: Extendible parking 'foot' extended during data taking to provide stability. Bottom-right: A vacuum bellows and rotary seal connected to a measurement unit. Two yellow external FSI fibre optic cables are visible.

rapid stopping when the power is cut. On either end of the train is an optical proximity sensor. If an obstruction is less than one metre away the drive motors will be disabled. There are limit switches at either end of the rail which again, are linked to the IRU to prevent travel off the end of the rail. A further hard stop is positioned after the limit switches as a fail-safe.

The entire tunnel has an interlock system which is connected to the train interlock system. The drive motors, stepper motors and lasers can all be enabled individually using the tunnel interlock system. Any unauthorised entry will halt all train motion and disable the lasers, starting with the EDFA to avoid damage to it. While the external FSI falls to a safe intensity within $50cm$, it is still in open air. The internal FSI and the LSM are collimated beams but in normal usage are completely enclosed. However, during some experiments, the enclosure was not complete, necessitating care over laser safety and making the interlock critical. Laser goggles are needed when the collimated beams had the potential to escape. As the two wavelengths are very different and some of the goggles did not adequately block both wavelengths, only one laser was enabled when there was a risk of free beams.

The units also have a system of warning lights that inform any people present of the current action being performed; either the drive system being engaged or the brakes being disengaged for the stepper motors (figure 3.9).

4 The Laser Straightness Monitor in Detail

This chapter covers details of the hardware set-up of the Laser Straightness Monitor used in the RTRS. It follows the path of the laser beam and examines the components encountered and their performance characteristics.

4.1 Overview

The LSM consisted of a laser beam travelling in vacuum down the length of the train, before returning in an anti-parallel direction. Each car had two pellicle beam splitters at 45° to the beam which reflected about 8% of the incident light onto a CCD. As a measurement unit changed orientation, the position of the beams on the CCDs changed accordingly. Using accurate knowledge of the geometry of the units, the relative positions and orientations can be reconstructed.

Figure 4.1 shows a simple schematic of some of the possible motions of a LSM unit, showing the effect on the beam spot position on the CCDs. It can be seen that a translation moves the beam spots on the CCDs the same amount, while a rotation around the centre of symmetry of the unit only causes motion on CCD 1 and CCD 2.

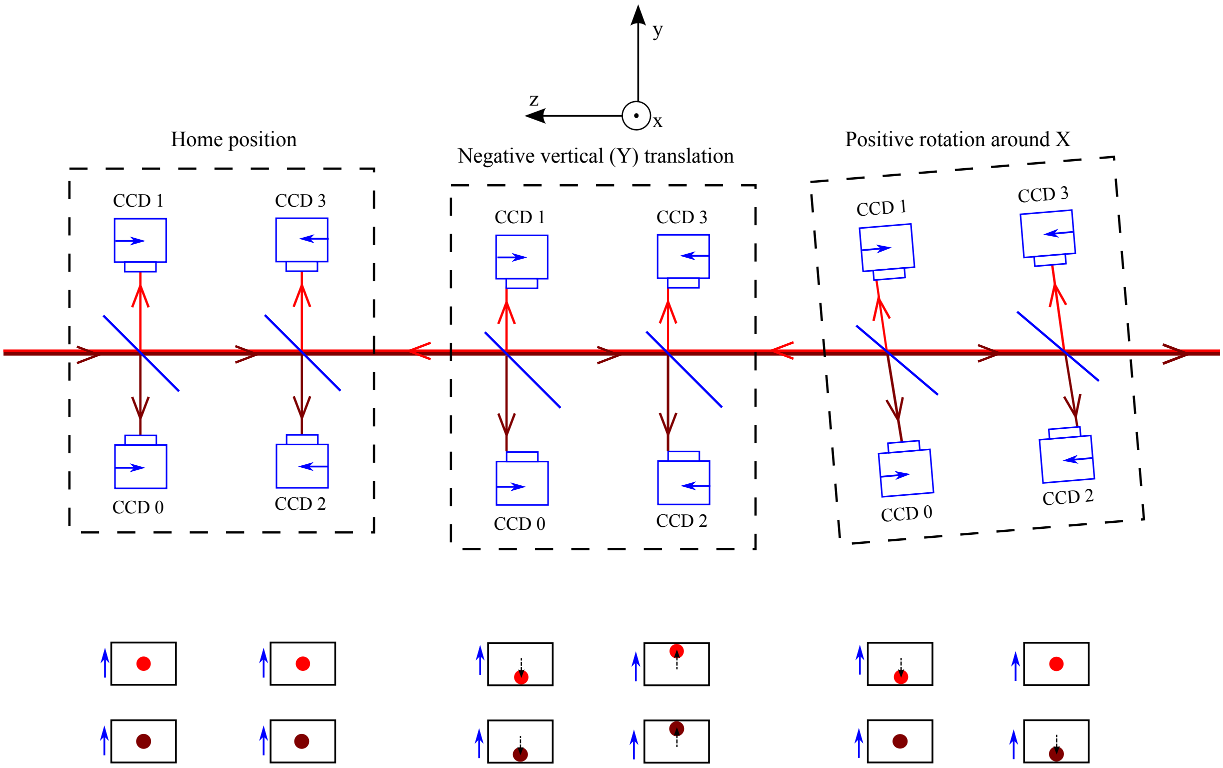


Figure 4.1: LSM Schematic showing a unit at its home position, another translated vertically down, and a final unit rotated around the X-axis. The motion of the beam spots on the CCDs from their home positions are also demonstrated.

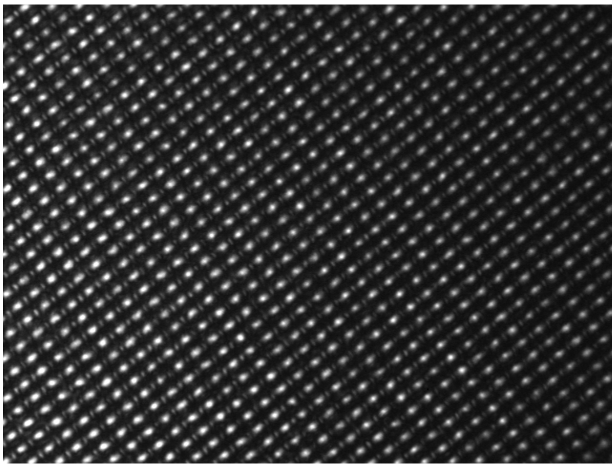


Figure 4.2: Image produced by passing uncollimated light through a crossed pair of diffraction gratings.

4.2 Laser

The laser chosen was the Schäfter and Kirchhoff 57FCM-A-178-20-A-2-28-150 [29] with the Y03 diode, due to its low coherence length, wavelength detectable by CCDs, fibre coupled output and reasonable price.

4.2.1 Spectrum

The Y03 diode had a central wavelength of 783nm which, although outside the visible range, is easily detectable with the cameras selected when their IR filters are removed. The large longitudinal coherence length of laser light is potentially a problem in that it gives rise to various interference effects. However, the Y03 diode had a spectral half width of 12nm at low current ($P_{out} = 2$mW), which resulted in a very short coherence length of 50 microns. At high current ($P_{out} = 15$mW) the half width fell to 2.2nm with an associated coherence length of 200 microns. The laser was used in the low current regime to minimise interference, though some issues were later encountered (chapter 6). Figure 4.2 shows an image produced by passing an uncollimated beam through a pair of crossed gratings, centred on the left-hand side of the image. While there is single-slit like interference, there is no clear long-range interference pattern associated with diffraction gratings.

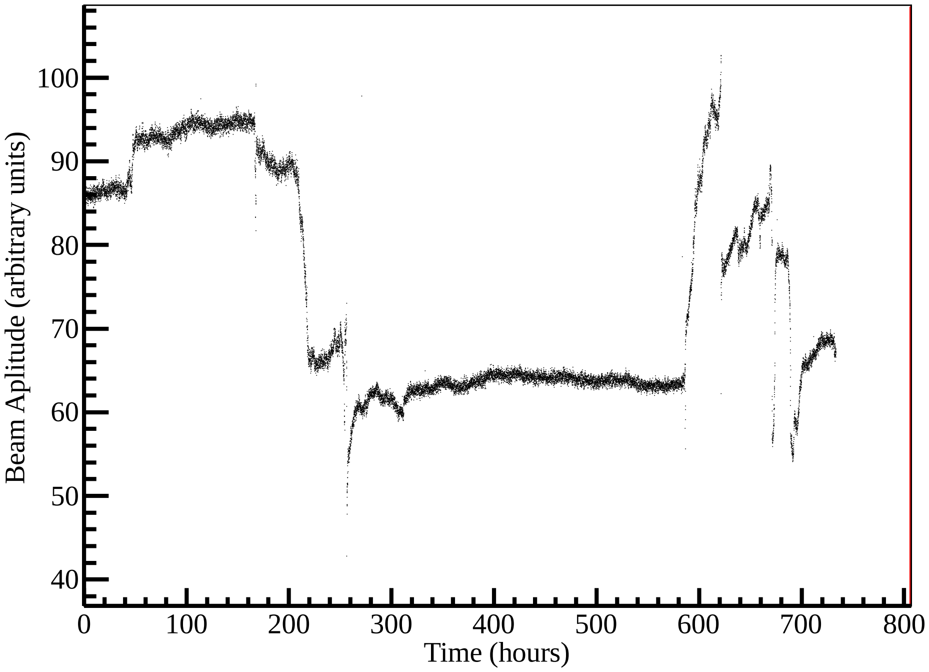
4.2.2 Stability

Wavelength stability was not considered a strict requirement as the only the intensity of the beam was being measured. However, as discussed in section 4.6.2, wavelength changes do affect the relative reflectance somewhat and mode hops were observed as the power was changed. This is still only a very minor concern as the beam fitting methods cope well with fitted beam spot amplitude variations and the power of the laser in normal use was kept constant at a low value, where the beam wavelength spectrum did not hop noticeably.

Intensity stability is desirable, as fixing the beam spot amplitude when fitting makes the process more robust, but there is another effect: if the amplitude varies too much then the images formed on some cameras may have some saturated pixels or others may have too low a brightness beam to fit well. Saturated beam spots are not a particular problem unless the beam is asymmetrical because saturated pixels are ignored. However if an asymmetric part becomes saturated then a shift will be produced due to its biasing effect being altered (§6.2). Figure 4.3 shows the significant variation in laser output over very long time-scales, though with a substantial period of stability. To produce the figure, the laser was left on in the train for the month of December 2008 while being observed by all 12 cameras, sampling every 131.5 seconds, producing 20000 sets of images. The images were fitted at 1:4 scale (chapter 6) and the intensity for each set plotted. It demonstrates both rapid 'hops' and more slow variations in intensity, as well as 5% rapid intensity fluctuation. The lower plot demonstrates wide variation in brightness measured by the cameras, though this is mainly due to the different gain settings and pellicle thicknesses (§4.6.2) rather than the reduction in amplitude caused by beam divergence and loss due to sampling. The fact that the all camera signals correlates strongly demonstrates that the majority of the variation is caused by power output rather than frequency shift. As such, adjusting the attenuation before a period of use is sufficient to keep the cameras within their intensity ranges.

4.3 Beam Launch

The fibre from the laser was directed into the launch end-cap using a custom made feed-through. Once inside it led to a Schäfter and Kirchhoff 60FC-T-4-G25-10 collimator [29] held in a New Focus 8807 [27] motorised mount (figure 4.4). This enabled the beam to be aligned after the end cap is installed. In principle, this operation only needs to be done once per installation, though in practice it was used much more than this due to its use in alignment: when the RTRS was moved to a new position in the tunnel for the first time, the beam often did not even hit the CCDs on car 2. To realign, the beam launch was adjusted to sweep the beam until it appeared on the outgoing CCDs in car 2. Then the launch unit itself was rotated to the orientation that would enable the collimator to be returned to its original orientation with the beam still hitting the car 2 cameras. The picomotors in the mount gave an angular resolution of 0.7 μrad [27] in



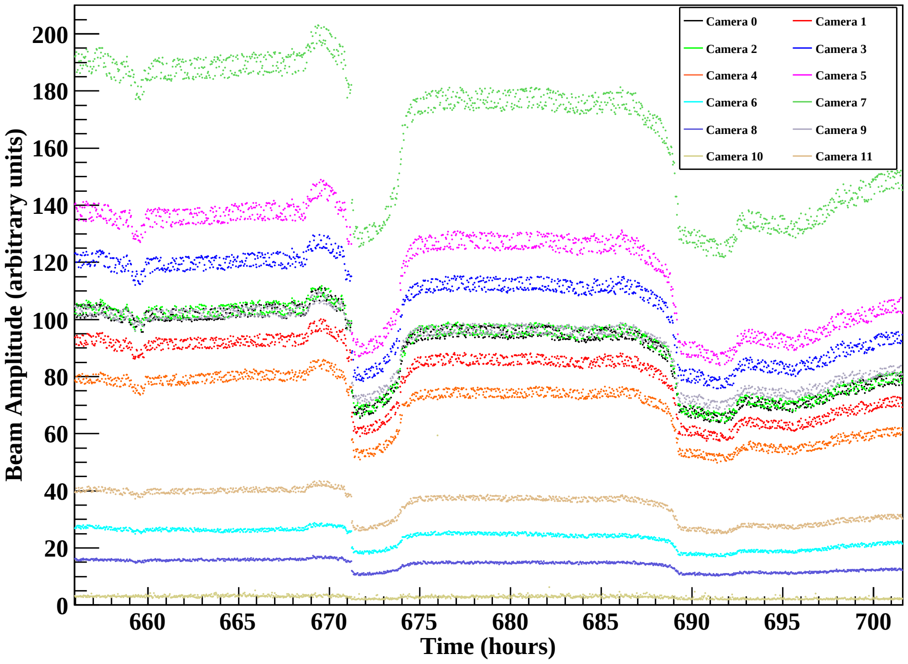


Figure 4.3: Long term beam intensity stability. The upper plot is the average of all cameras over the whole range while the lower plot shows the readings for each camera and is zoomed on a more variable region.

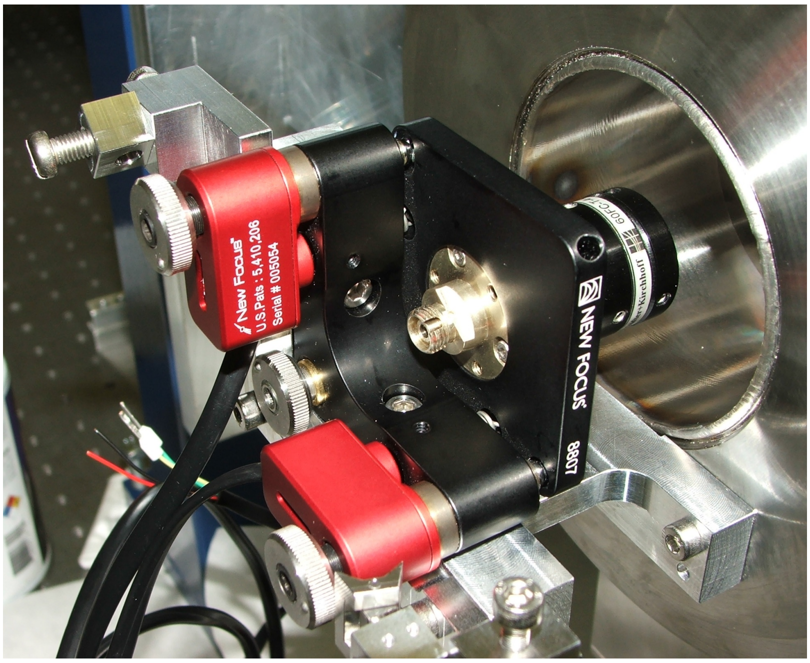


Figure 4.4: Beam Launch assembly during the fitting of the limit switch assemblies.

both axes, which is sufficient for normal use.

4.4 Beam Propagation

The beam propagated through the vacuum with a roughly Gaussian profile. The irradiance distribution of a Gaussian beam is:

$$I(r,z) = I_0 \left(\frac{w_0}{w(z)}\right)^2 \exp\left(-\frac{2r^2}{w(z)^2}\right) , \tag{4.1}$$

where r is the radial distance from the central axis, z is the axial distance from the beam waist, I_0 is the central intensity at the beam waist ($z = 0$) and $w(z)$ is the radius as which the intensity is $\frac{1}{e^2}I_0$, given by:

$$w(z) = w_0\sqrt{1 + \left(\frac{z}{z_R}\right)^2} , \tag{4.2}$$

where

$$z_R = \frac{\pi w_0^2}{\lambda} . \tag{4.3}$$

The size and position of the beam waist are important as the long range characteristics of the beam depend on them. The LSM launch collimator was adjusted to produce the beam waist near to the half-way point in the optical path: at the retro end-cap. However, the control over the collimator focus was not sufficient to precisely set the waist location.

Any obstructions in the way of the beam cause diffraction effects further along the optical path. Possible sources of distortions of the beam include:

- Clipping inside the collimator on the launch.
- Clipping of the beam by the edges of the pellicles or the return mirror.
- Effects from the retro-reflector, if used.
- Interference from dust on the return mirror.
- Interference from dust on the pellicles.
- Interference from dust on the CCDs.

All the effects get ‘washed out’ as the beam propagates. Issues caused by the above effects were avoided as much as possible, and were reduced as part of the fitting process in chapter 6. Interference effects from dust particles are briefly investigated in appendix A.

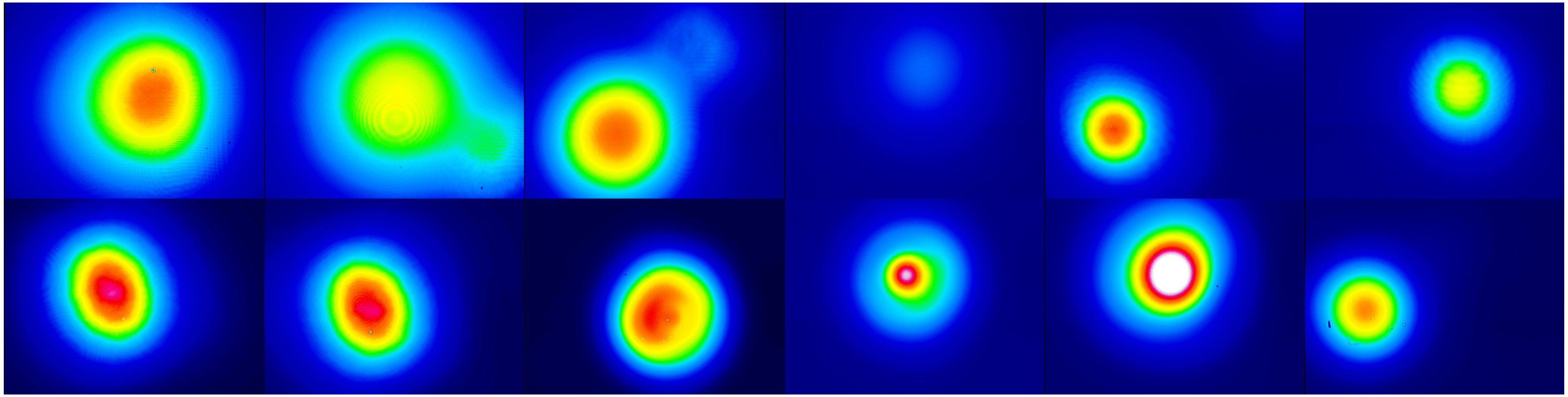
4.5 Beam Retro-Reflection

At the far end of the train, the beam was reflected back in a second end-cap. The end-cap contained a 50% reflectance mirror with a corner cube retro-reflector placed behind, mounted on a New Focus 8071 motorised 4-axis stage [27]. A remote controllable shutter in between the two other components allowed blocking of the light to the retro-reflector. The stage had four internal picomotors, each with a linear increment of less than 30nm, resulting in an angular resolution of 0.7μrad, which is sufficient for use in the LSM (§3.6).

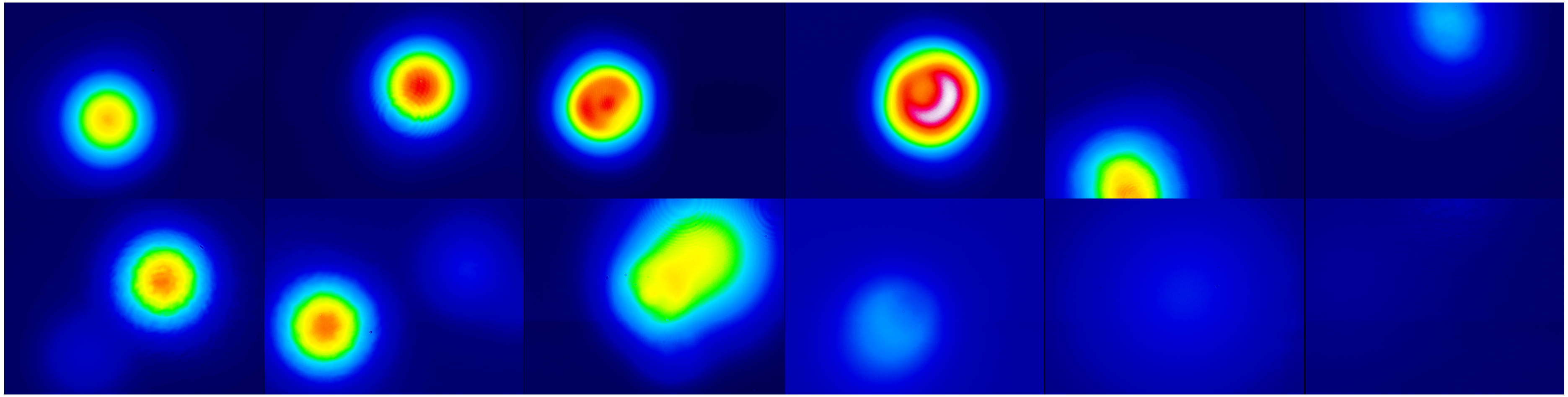
The original intention was to compare the beam produced by the retro-reflector - which is anti-parallel but has a lateral offset if off-centre, with the beam produced by the mirror - which is not anti-parallel but gives no offset, and adjust the stage until the two beam spots are on top of each other on multiple CCDs. However, for operational reasons and complications related to beam distortion by the retro-reflector, only the half-mirror was ever used. An alternative method of producing an anti-collinear beam was developed using reconstruction (§9.2.3).

4.6 Beam Splitters

The laser beam was sampled at twelve points along the RTRS' length by reflecting a fraction of the beam onto CCDs located above and below. To minimise disruption of the beam by refraction,



(a) False colour images of the beam as it propagates from its launch in car 1 (the two bottom left images), through car 2 (the two central bottom images) and car 3 (bottom right) before its return journey (along the top, ending with camera 1 at the top left). Taken on 27th June 2008.



(b) False colour images of the beam as it propagates from its launch in car 3 from top right to top left, before returning, producing the images from bottom left to bottom right. Taken on 8th October 2008. The distortion of the third image on the bottom row is caused by a ghost beam (§6.4) in the centre.

Figure 4.5: Unnormalised beam profiles for the RTRS in the two configurations, demonstrating significant intensity variation and beam distortion.

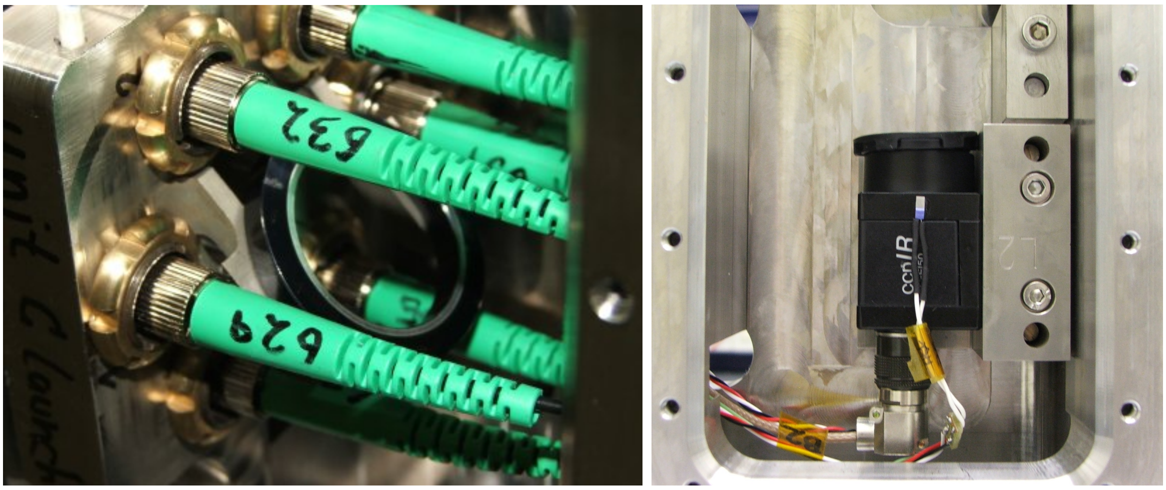


Figure 4.6: Pellicle beam splitter with Internal FSI Launch fibres (left) and a mounted CCD camera with dust cap (right). The camera cabling and PT100 temperature sensor can also clearly be seen.

1 inch diameter, 2μm thickness, uncoated nitrocellulose pellicle beam splitters [30] were used, mounted at a 45 degree incidence (figure 4.6).

4.6.1 Beam Walk from Beam Splitter Transmission

As the laser beam passes through a beam splitter, it refracts twice (figure 4.7), producing a small walk given by the formula:

$$W = \frac{d \sin(\theta_i - \theta_t)}{cos\theta_t}\,, \tag{4.4}$$

$$\theta_t = \arcsin\left(\frac{1}{n_2}\sin\theta_i\right)\,. \tag{4.5}$$

Nitrocellulose has a refractive index of 1.50 [31], so using Snell's law the transmitted angle in a vacuum can be found to be 28.1°, leading to the walk produced by each pellicle to be 0.66μm. The absolute value of the walk is not important as it is absorbed into the calibration constants determining the positions of the CCDs hit after transmission (§9.) Potentially more important is the change in walk as the angle of the unit changes with respect to the laser beam. However, as the angular range of the LSM was ±20mrad, equation 4.4 leads to the total range of the offset caused by each pellicle to be 5nm, which is negligible.

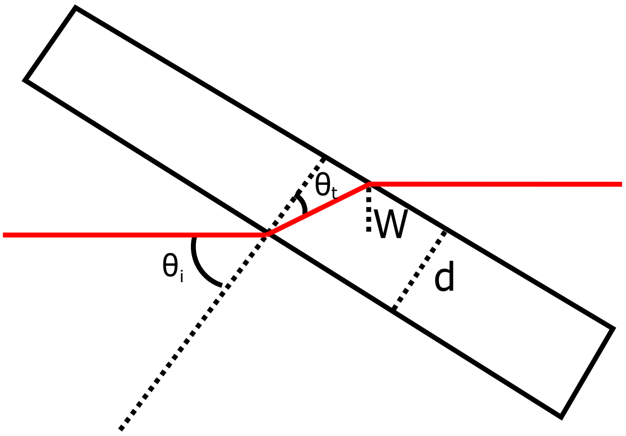


Figure 4.7: Walk of a light ray produced by a plate at an angle.

4.6.2 Reflection from a Beam Splitter

The thin nature of the beam splitters compared to the coherence length of the laser light means their reflectance is not trivial. Reflection occurs off both surfaces of the pellicle as described by the Fresnel Equations [32, Section 4.6.2]. For non-magnetic dielectric media the reflectances are given by:

$$r_\perp = -\frac{\sin(\theta_i - \theta_t)}{\sin(\theta_i + \theta_t)}, \tag{4.6}$$

$$r_\parallel = +\frac{\tan(\theta_i - \theta_t)}{\tan(\theta_i + \theta_t)}, \tag{4.7}$$

where θ_i is the angle of the incoming beam from normal, θ_t is the angle of the transmitted beam, $\perp$ refers to the part of the wave with the **E** field perpendicular to the plane of incidence and $||$ refers to the part which is parallel. Thin film interference between both reflections leads to an intensity of the reflected beam of [32, section 9.6]:

$$I_r = I_i \frac{2r^2(1-\cos\Delta)}{(1+r^4) - 2r^2\cos\Delta}, \tag{4.8}$$

where Δ is the phase shift caused by the extra optical path in the beam splitter of refractive index n and is given by $\Delta = 4\pi nd\cos(\theta_t)/\lambda$ for a wavelength λ and the geometry in figure 4.7. Combining equation 4.8 with the expressions for r (equations 4.6 & 4.7) gives the reflectance of the pellicles for each polarisation as a function of incident angle. Figures 4.8 & 4.9 show the reflectance as a function of angle and thickness respectively for 783nm light. Figure 4.9 shows

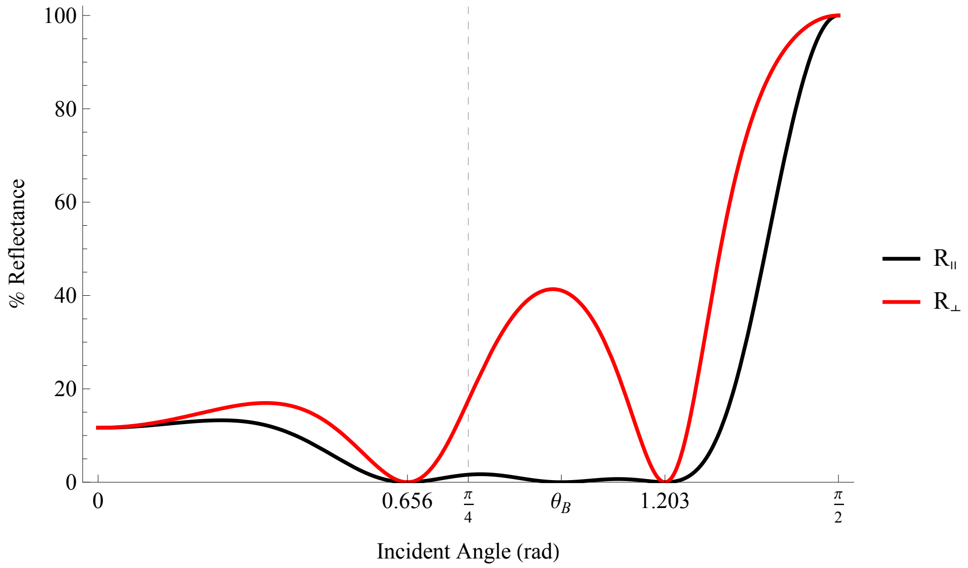


Figure 4.8: Pellicle reflectance in both perpendicular (red) and parallel (black) polarisations for a 2μm thick pellicle as a function of incident angle.

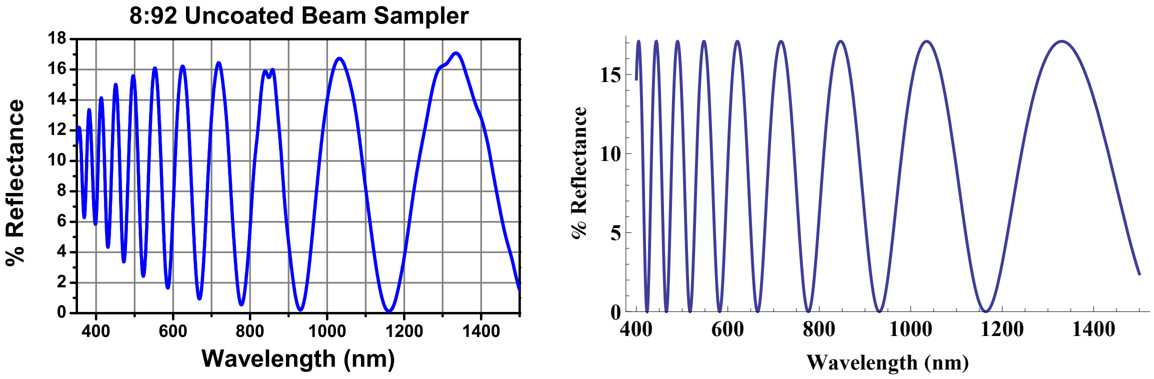


Figure 4.9: A plot of the measured reflectance of a 8:92 uncoated pellicle beam splitter with unpolarised light at 45° incidence as measured by Thorlabs [30] (left). Theoretical plot with a pellicle thickness of 1.76 microns (right).

the transmission of the pellicle as determined by experiment by Thorlabs [30]. The plot is an excellent match for the theoretical calculation for a pellicle thickness of 1.76μm.

There is very little variation in thickness over an individual pellicle, giving a focal length of more than 1km when measured, meaning lens effects of the pellicles can be ignored [33]. However, the variation in thickness between different pellicles is significant. While the data-sheet for the pellicles does not give a tolerance for the thickness, a Thorlabs engineer gave an estimate of $\pm 0.2\mu$m, reinforced by the theoretical match of 1.76μm to their experimental results presented in figure 4.9. Additionally, the 10nm range of central wavelengths that the laser may hop to has an effect on the reflectance. Figure 4.10 demonstrates the variation in reflectance over the expected thickness range for both polarisations of light, giving bands for the wavelength range; the reflected

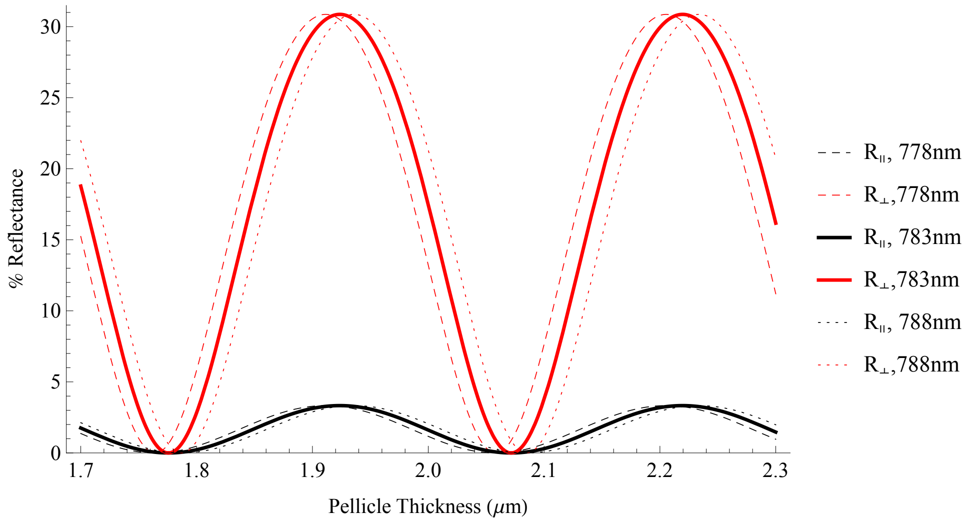


Figure 4.10: Theoretical reflectance for a beam incident at 45° for a pellicle of varying thickness for both polarisation directions, for the central wavelength and wavelengths differing by ±5nm.

light is mostly perpendicularly polarised with significant variation in polarisation-averaged total reflectance from zero to 16%, averaging to the 8% claimed in the specifications [34].

The angle of incidence has a range of ±40mrad, which is enough to cause significant variation in reflectance for pellicles for thicknesses where the rate of change in reflectance is high. Figure 4.11 shows the calculated reflectance of pellicles of a range of different thicknesses under unpolarised light over the angular range of the LSM, with bands showing the effect of shifting the wavelength ±5nm from the nominal 783nm[3]. The effects are similar in magnitude, but are both minor for pellicles that have thicknesses corresponding to high reflectance. Figure 4.10 suggests that pellicles can be screened for this with roughly 20% within the high-reflectance, low-variation range. However, due to time and resource constraints not all pellicles used were optimal, leading to low reflectance of unit 2, pellicle 2 and unit 3, pellicle 2, corresponding to low beam amplitudes for CCD 7 with the launch in car 1, and low amplitudes for CCD 6 and CCD 10 with the launch in car 3 (figure 4.5).

4.7 Camera Performance

The camera used in the RTRS was the Sony XC EI50-CE [35] with an ICX249AL CCD sensor [36]. This camera was chosen as it:

- functioned in a vacuum without overheating or exhibiting other issues.
- was sufficiently sensitive at the laser wavelength of 783nm to image the beam clearly.

---

[3] Note the opposite effect of a wavelength increase for the 2.00μm pellicle and the 2.15μm pellicle.

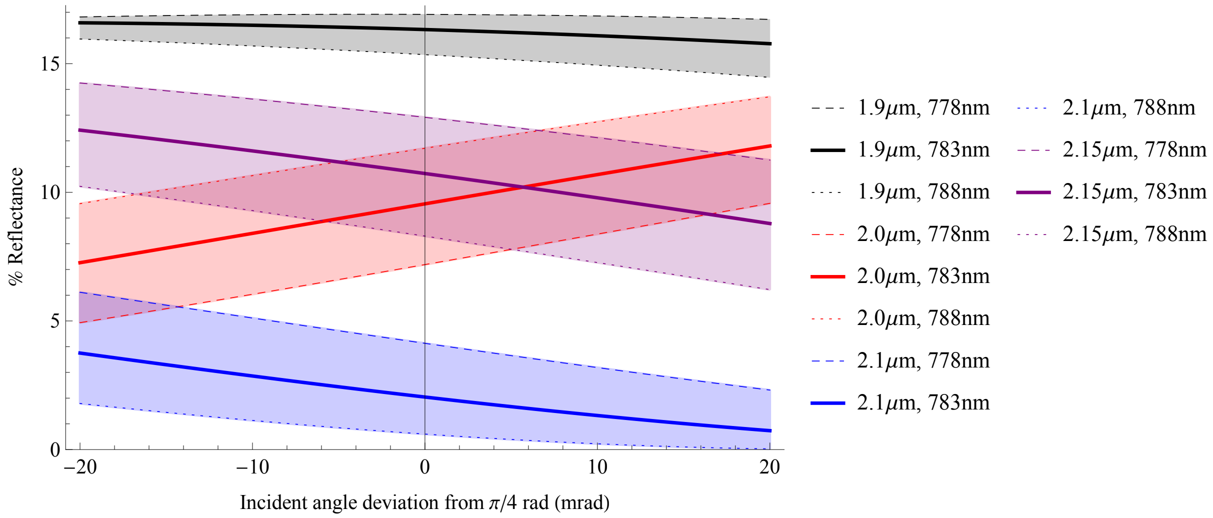


Figure 4.11: Reflectance bands for different thicknesses of pellicle and laser wavelengths over the angular range of the LSM.

- had a CCD size large enough to give a good working range for the unit.
- was vibration tolerant enough to withstand extended use in the RTRS.
- was able to take images in rapid succession, at up to 25Hz.
- was within budget.

The XC EI50-CE is an analogue camera so requires a frame grabber to acquire images. For initial experiments a National Instruments IMAQ PCI-1409 [37] frame grabber was used. Images were grabbed at a 10 bit resolution using LabVIEW [38].

4.7.1 CCD Dimensions

The cameras were measured with an optical smart scope with a ±1.5μm measurement uncertainty, giving measurements of the dimensions of the CCD comparable to the reference values. The pixel sizes were calculated by using the number of rows and columns given in the ICX249AL data-sheet (752×582). As can be seen in table 4.1, the measured values match very well with the manufacturer's specifications. However, there are additional columns of pixels on the CCD which, although marked out separately with a different background, are still functional. Additionally, some trouble was experienced with some rows at the left-hand extreme of the CCD. Therefore, the final region of interest for image acquisition was 760×572 pixels.

Table 4.1: Measured and specification CCD dimensions in microns, with the active area used for acquisition also given.

	Measured	Specification	Active Length
CCD Width	6468.13 (43)	6467.2	6536
CCD Height	4829.73 (43)	4830.6	4747.6
Pixel Width	8.6012 (74)	8.6	
Pixel Height	8.2985 (74)	8.3	

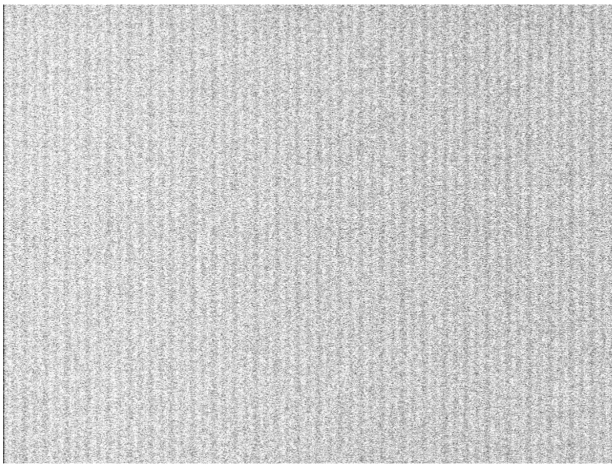


Figure 4.12: Normalised background image taken at maximum gain demonstrating static noise.

4.7.2 Noise in Darkness

To find the behaviour of the CCDs in complete darkness, sets of 200 images were taken at 3Hz with the cameras set to varying gain levels. There were two components to the noise: a static background pattern with dynamic random noise overlaid. Figure 4.12 presents a typical image, normalised to fill the full image grey-scale range to make the pattern visible. A measurement of the dynamic noise was made by first creating a reference image by averaging 200 individual images. Then the difference of each individual image from this reference was found and the residuals histogrammed, giving a noise spectrum. Noise profiles for both minimum and maximum gain can be found in figure 4.13. They are generally Gaussian with some distortion for the high gain spectrum.

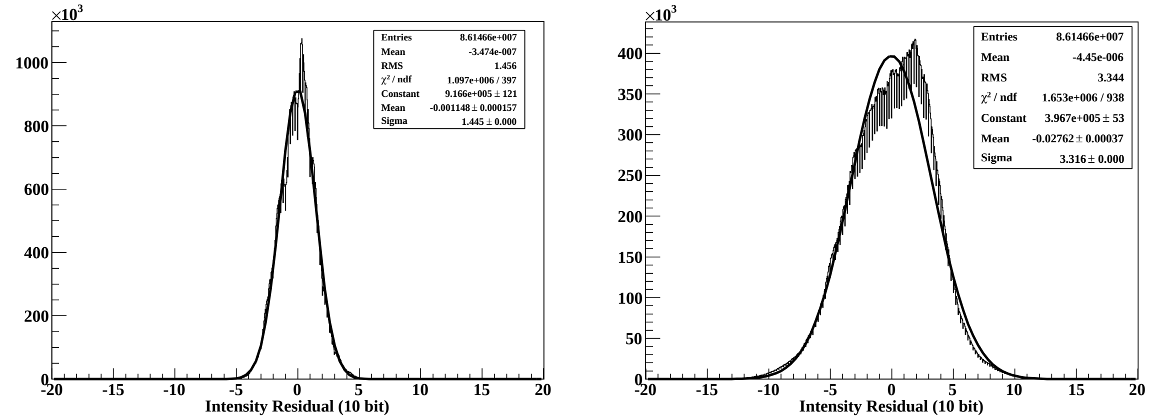


Figure 4.13: Pixel counts for intensity residuals at zero incident light for minimum gain (left) and maximum gain (right). The near-Gaussian distributions are highlighted by applying a simple Gaussian fit.

4.7.3 Noise Analysis with an Incident Beam

To produce a CCD noise spectrum as a function of the amount of incident light, a camera was exposed to an uncollimated beam in a vacuum vessel, via a 50:50 beam splitter[4] (figure 4.14). The camera was left on until the temperature had stabilised at 34°C. Then three series of 200 images were taken at 3Hz with the gain set to minimum, half-way and maximum. The beam brightness adjusted to fill the range of the camera without saturation for each gain setting. A similar process to section 4.7.2 was used to create noise profiles, but this time 64 separate noise spectra were generated based on the value of each pixel in the average image: pixels with a value in the average image in the range of 0-15 create the darkest spectrum, with pixels in the range 1008-1023 creating the noise spectrum for the brightest pixel values.

Figure 4.15 gives some typical noise profiles for pixels with different light levels and gains, all taken with the same beam position. They show the increase in noise as the gain and intensity are increased. The profiles are fitted well by Gaussians, though the intensity range when the camera is at maximum gain shows some substructure. This substructure causes the chi-squared per degree of freedom to be higher than expected when using the square root of the bin value as the error.

4.7.4 Modelling CCD Noise

With a defined process of generating noise spectra for different intensities, and having seen the spectra are well represented by Gaussians, it is appropriate to model the noise over the intensity

---

[4] The transmitted beam was used to produce the image seen in figure 4.2, though this second camera was disabled for the noise analysis.

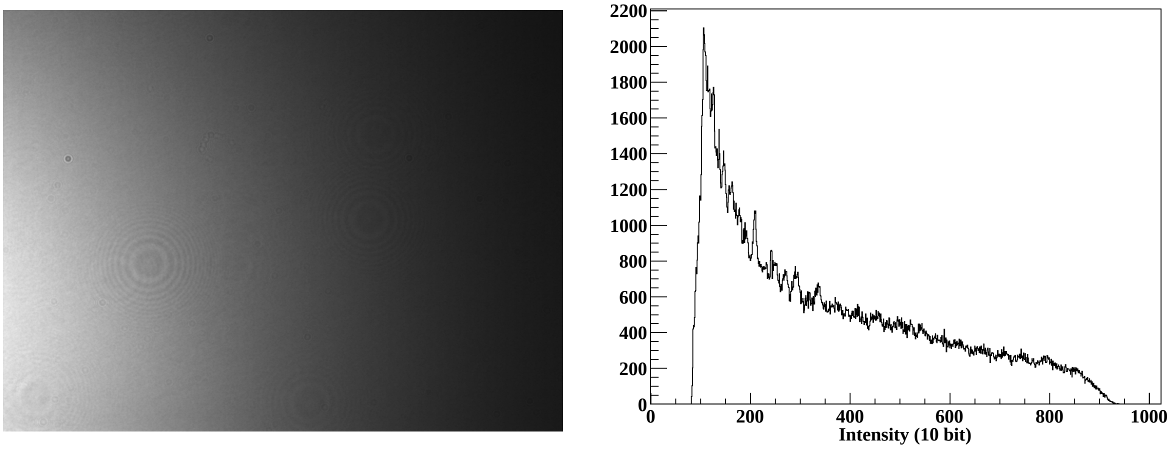


Figure 4.14: Typical image used for noise analysis taken at medium gain (left), with the resulting pixel count distribution over the intensity range shown (right).

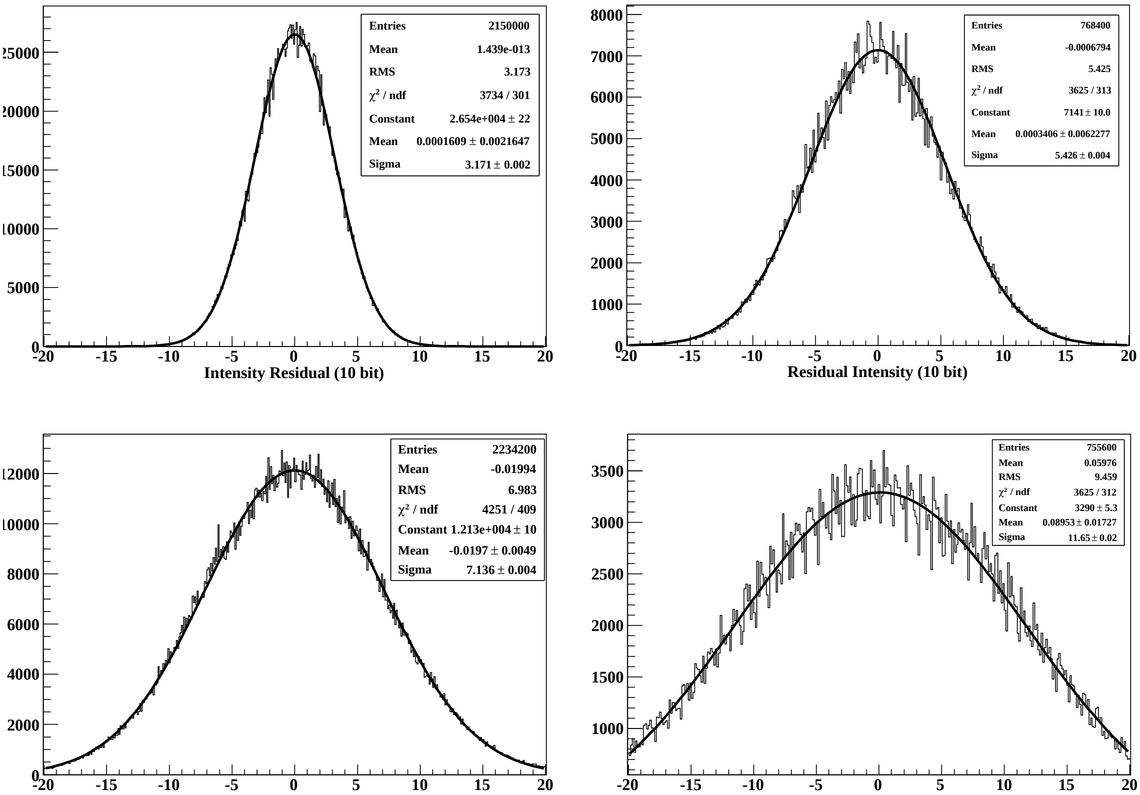


Figure 4.15: Pixel count distributions for pixels with a mean amplitude in the range 240-255 (left), 752-767 (right) for minimum gain (top row) and maximum gain (bottom row). Gaussians fits have been superimposed to highlight the nature of the distributions.

range.

The sample variance (S^2_{n-1}) of the residuals was calculated for each intensity range bin using the raw data, and the error on the sample variance was found by the mean squared error estimator:

$$MSE(S^2_{n-1}) = \frac{2}{n-1}\sigma^4 \,, \tag{4.9}$$

where n is the number of number of data points (pixels in the range × number of images) and σ is the sample standard deviation of intensity residuals. Brighter ranges have far fewer pixels due to the intensity profile of the beam, and there are no pixels in the very lowest ranges due to the black offset of the frame grabber.

The sample standard deviations were then plotted against the bin central intensity for the different gain settings; figure 4.16 contains the resultant plots for low-gain and high-gain. The error on the points increases with intensity as there are fewer pixels with the required average intensity value. This was particularly true at the highest intensities leading to very poor precision near saturation. The plots were fitted with the function:

$$N(I) = \alpha\sqrt{I-\beta}\,, \tag{4.10}$$

where α corresponds to a noise scale factor and β corresponds to the intensity offset needed to give a zero-intensity pixel the noise value observed. Equation 4.10 describes the general trend very well except that:

1. There is an oscillatory structure to the noise on top of the square root trend.
2. At the highest intensities the noise reduces slightly, suggesting saturation of a part of the amplification chain. (However, the errors on these values are large so there is low confidence in this observation.)

The oscillatory structure is due to the frame-grabber: changing the grabber to a Matrox Solios XA [39] later gave similar spectra with a different oscillation pattern (figure 4.18, right).

Table 4.2 summarises the camera behaviour: the first two columns give the black pixel intensities and their r.m.s from the dark set-up, followed by two columns with the fit parameters for equation 4.10 from the data with an incident beam. It also gives values for the dark RMS extrapolated from the fitted function which are fairly consistent with the observed values. While the x-offsets do not seem to be particularly constant across the gain levels, their effect quickly becomes small as the intensity increases.

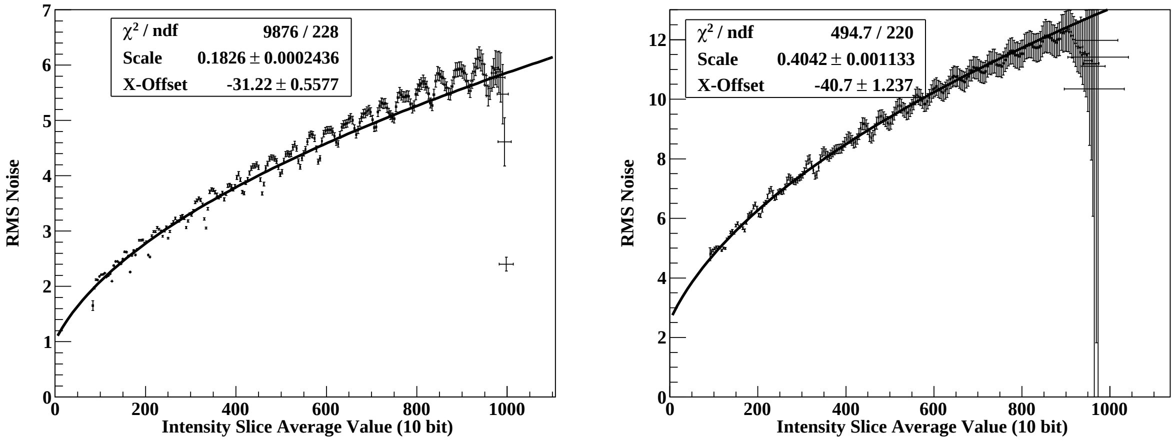


Figure 4.16: Sample Standard Deviation of the intensity readings against intensity, with the camera set to minimum gain (left) and maximum gain (right). There were four intensity levels per bin.

Table 4.2: Noise analysis summary, showing the observed black pixel values and the parameters of best fits of function 4.10. The expected Standard Deviation of a pixel with intensity value of zero calculated from the fitted function is also given for comparison.

Gain Setting	Average black pixel value	Black pixel SD	Noise Scale (α)	Noise X-Offset (β)	Extrapolated Black pixel SD
low	26.655	1.456	0.1826	-31.2	1.389
medium	28.944	1.838	0.2354	-30.9	1.821
high	37.769	3.344	0.4042	-40.7	3.581

4.7.5 Cameras In the Train

The process of producing CCD specific, intensity dependent pixel noise estimates was applied to each CCD in the RTRS using sets of 250 images taken at 25Hz every time the RTRS was prepared for a series of experiments. Images were taken rapidly to minimise effects of beam motion and electronic interference, which are investigated in detail in chapter 7. With the beam motion a fraction of a percent of the beam width, and the largest interference pixel contribution being 0.2 units, both effects are negligible compared to the noise.

The frame grabber used in each car of the train was the Matrox Solios XA, 4-way PCI-X analogue frame grabber [39], using Matrox MIL-Lite drivers for code-level access. This frame grabber has a bit-depth of eight if all four channels are used, as is the case in the RTRS. At very low intensity levels, the granularity of the grabber significantly affects the noise spectra generated due to the low resolution compared to signal noise (figure 4.17). This effect disappears at an

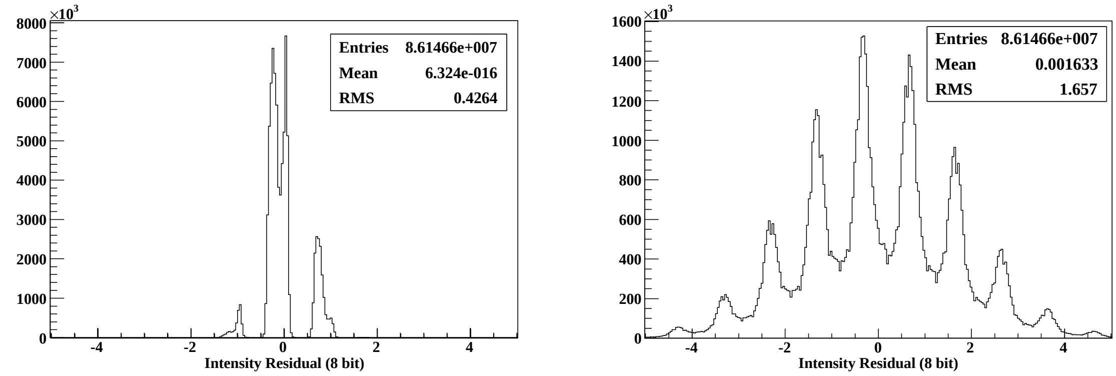


Figure 4.17: Pixel counts of the intensity residuals, produced by cameras at different gain settings: car 2, CCD 1, low gain (left) and car 3, CCD 2, high gain (right). Digitisation effects are immediately apparent.

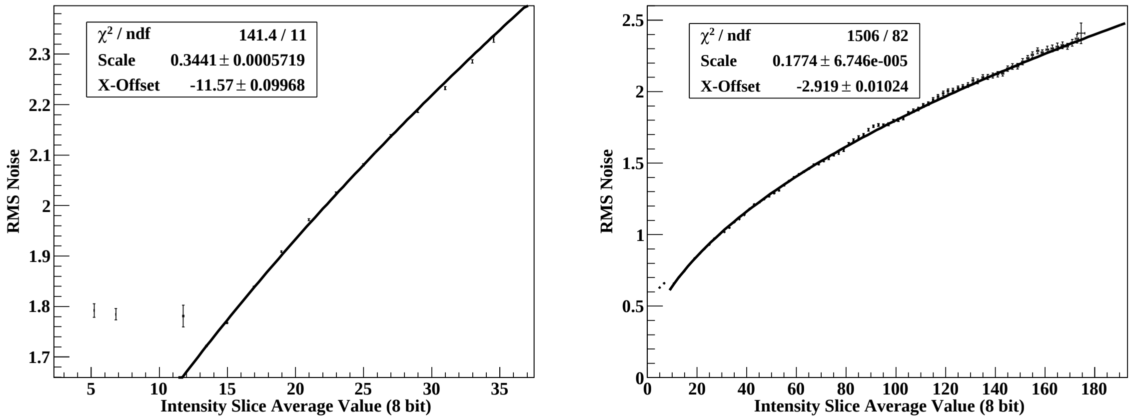


Figure 4.18: RMS noise levels for car 3, CCD 0 (left) and car 3, CCD 1 (right). Car 3, CCD 0 has high gain with a dim beam, while car 3, CCD 1 has low gain with a brighter beam. Fits using equation 4.10 for data with an intensity over 10/255 are shown.

intensity of around 10/255 and so the noise modelling function was set to only apply to noise values above an intensity of 12/255; below this intensity a simple constant was used.

Figure 4.18 shows the noise versus intensity plots for two of the cameras in unit 3 with the beam launch end-cap attached from that end. The (right hand) plot for CCD 1 shows the RMS distribution with very low gain; the camera was close to the source, the beam was small and only had a single sample taken off, so is bright. The (left hand) plot for CCD 0 is quite different: by the time the beam has travelled the length of the train and back, the beam width is twice as large, has had nine samples taken off and a reflection off a half-mirror, making it much dimmer (figure 4.5b). Even with the gain increased, the maximum intensity recorded was 35/255. Both plots show the square-root dependence with the minimum cut-off, with different RMS noise scales due to the different gain settings.

5 RTRS Set-up and Usage

This chapter contains an overview of the control of the RTRS motion and alignment.

5.1 Train Control and Data Acquisition

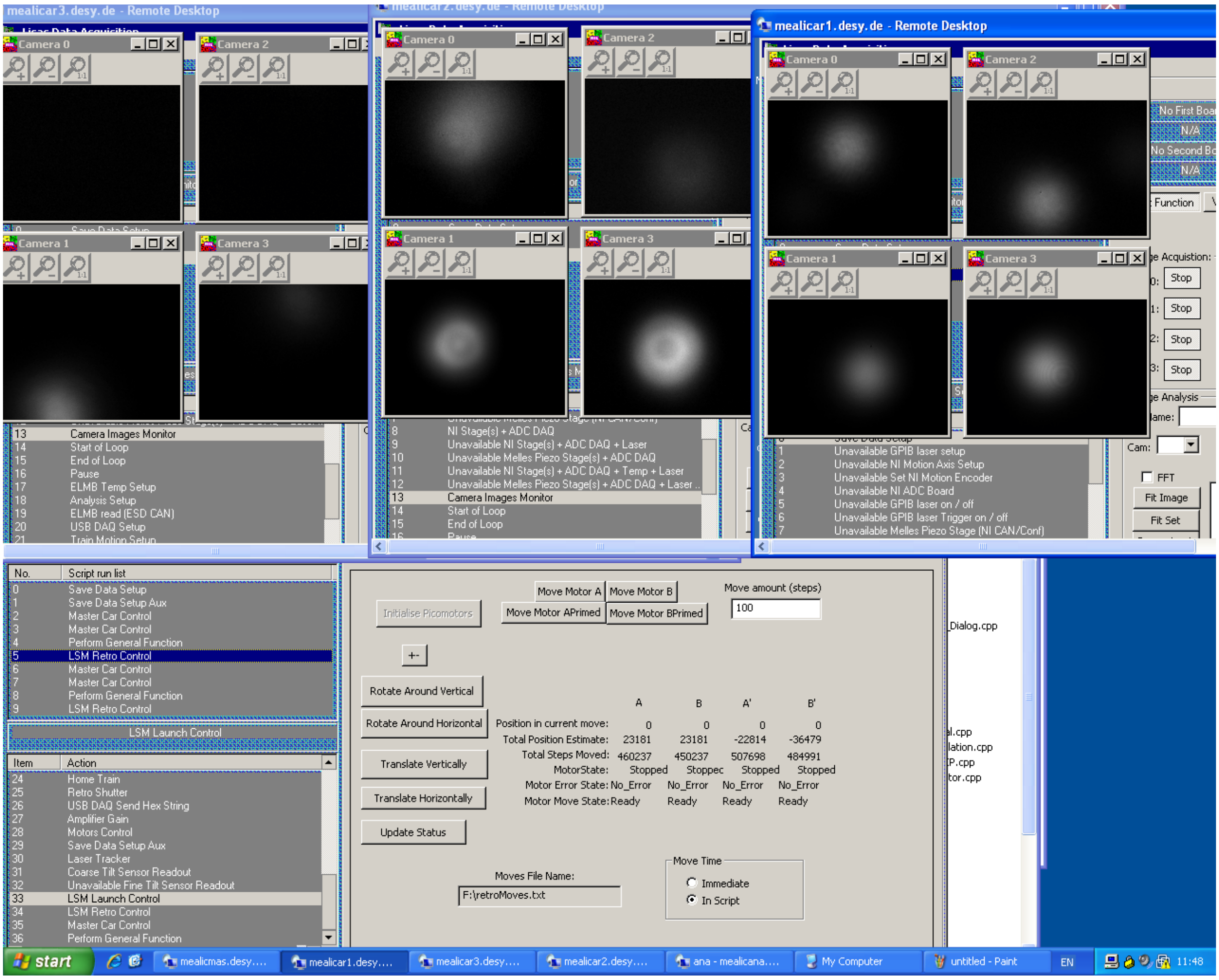
5.1.1 DAQ and Control Program

All the functions of the RTRS were controlled using software custom written in C++ using the Microsoft Foundation Class library [40]. The software was run on each of the three data-acquisition computers, in a slave mode, synchronised over a network connection behind a switch via another instance running on the master car. Data was [41] made available for all units via a shared hard drive on the master computer.

The control software was based on modular 'script items' for each action required, such as 'LSM Retro Control' for adjusting the position and orientation of the end-cap reflection assembly, 'Camera Images Monitor' for taking images from the CCDs and 'Read USB DAQ and Save' for taking interferometer data. There were also program flow script items, synchronisation items and a 'Perform General Function' item for invoking functions in DLLs for which no script item was written, such as LSM reconstruction. Some script items are only relevant for particular computers and were disabled when not useful; for example, the train drive motor control item is only enabled on the master computer.

Script items were selectable from an 'action list' from where they could be dragged and dropped into a 'script run list'. Most script items had two modes: an 'immediate' mode, visible when the script item was selected in the action list, and a 'script mode' which is enacted instead when the script is run. The 'immediate mode' was used to perform live actions manually, while the the 'script mode' used the same controls panel to set up desired actions in advance. When suitable scripts are ready on all computers, the script on the master computer with its synchronisation components is run.

Figure 5.1 shows the instances of the software running on the four different cars, displayed on a laptop screen using remote desktop. The script item visible at the bottom is for adjusting the retro assembly, while the top three windows contain measurement car remote desktop sessions, with the live images from the cameras being displayed.



No.	Script run list
0	Save Data Setup
1	Save Data Setup Aux
2	Master Car Control
3	Master Car Control
4	Perform General Function
5	LSM Retro Control
6	Master Car Control
7	Master Car Control
8	Perform General Function
9	LSM Retro Control

LSM Launch Control

Item	Action
24	Home Train
25	Retro Shutter
26	USB DAQ Send Hex String
27	Amplifier Gain
28	Motors Control
29	Save Data Setup Aux
30	Laser Tracker
31	Coarse Tilt Sensor Readout
32	Unavailable Fine Tilt Sensor Readout
33	LSM Launch Control
34	LSM Retro Control
35	Master Car Control
36	Perform General Function

Move amount (steps): 100

	A	B	A'	B'
Position in current move:	0	0	0	0
Total Position Estimate:	23181	23181	-22814	-36479
Total Steps Moved:	460237	450237	507698	484991
MotorState:	Stopped	Stoppec	Stopped	Stopped
Motor Error State:	No_Error	No_Error	No_Error	No_Error
Motor Move State:	Ready	Ready	Ready	Ready

Moves File Name: F:\retroMoves.txt

Move Time
○ Immediate
● In Script

Figure 5.1: Train control and data acquisition software in use.

5.2 Train Alignment Procedure

When the train was moved to a new stop in the tunnel, the units needed to be aligned so that the internal FSI beams hit the correct retro-reflectors, the LSM beam travelled the length of the train and back again, and the tilt sensors were in-range.

The rail on the wall was sufficiently uneven to create situations where the LSM beam did not hit the CCDs in unit 2 immediately after a move. In the absence of an automated system to pre-align the RTRS the following procedure was devised. It assumes the launch end-cap was connected to the 'launch' unit and the retro end-cap was connected to the 'back' unit:

1. The units were rotated until the tilt sensors gave a near-zero reading. (Less than 100 microradians).

2. The units were manually translated until there was line-of-sight through externally mounted sights.

3. The laser brightness was increased substantially, making it easy to see the beam a long way from its centre and also to see reflections off the vacuum tube.
4. The picomotor mount containing the launch optics was manually adjusted until the beam was seen on the outgoing cameras in unit 2.
5. While observing the cameras on the back car, the beam was moved over the 1 inch aperture of the beam splitter in car 2 by adjusting the picomotor mount containing the launch optics.
6. If step 5 failed to produce any response on the back car cameras, the back car was moved in a spiral pattern in the X-Y plane until the beam was found.
7. The laser brightness was reduced to its standard level.
8. The laser launch was then adjusted back to its nominal position and the rotation of the launch car adjusted to compensate, keeping the laser hitting the back car CCDs.
9. The launch unit was then restored to zero tilt sensor rotation if possible, with all units translated to maintain the beam on the back unit CCDs. The exact car translations varied depending on the limits of their travel.
10. The units were adjusted until the LSM beam was central in the first outgoing CCD hit on each car; this was done manually by translating unit 2 and the back unit. If the translations required would put the cars at a limit, the launch unit was adjusted to avoid this.
11. The retro stage was then adjusted to give an anti-collinear beam (§ 5.2.1).

Once the train was aligned, the motor positions for that stop were saved. Later, when the train returned to the stop, the saved positions were used to provide a good starting point for the alignment. When a position had been saved the procedure was much simpler:

1. The cars were all adjusted to their saved positions; the outgoing beam is now at least partially on-CCD for all cars.
2. The units were adjusted until the beam is central in the first outgoing CCD hit on each car; this was done manually by translating unit 2 and the back unit. If the translations required would put the cars at a limit, the launch unit was adjusted to avoid this.
3. The return beam was then adjusted to give an anti-collinear beam.

A further complication was that the internal FSI subsystem requires unit 2 to be orientated to point the FSI beams correctly at the FSI retro-reflectors on unit 1, while unit 3 needs to be correctly aligned to point at unit 2. The optimal relative orientations were found and the functional range of $0.75mrad$ confirmed [12, p115-118]. If the LSM launch was in car 3, then the LSM launch mount was adjusted so that its nominal home position corresponded to the aligned state of the unit 3 FSI subsystem for a central beam on the first CCD hit on unit 2. (unit 2, CCD 3.) The correct orientation of unit 2 to point the FSI beams at the retro-reflectors on unit 1 with the beam in the centre of unit 1, CCD 3 was saved. When moved to a new stop, car 2 was adjusted to this orientation by performing a reconstruction (chapter 8) on it and making the necessary correction.

5.2.1 Producing an Anti-collinear Beam

After the units are aligned so that the outgoing beam can travel the length of the train, the mirror in the retro end-cap needs to be adjusted to produce an anti-collinear beam. In principle, if all units are at zero x-rotation as indicated by the tilt sensors and the launch optics are unadjusted, only the y-rotation should need adjustment. In practice, it was not always possible to level each unit or there was enough variation in orientations to need both rotation axis adjusted.

To produce an anti-collinear beam, the retro end-cap mirror alignment was performed using the method developed as part of the calibration experiment in section 9.2.3. The method was to calculate the beam spot positions expected on the return CCDs of the unit holding the retro end-cap if the mirror was normal to the beam, then adjust the mirror until those positions were observed. Final adjustments were then made based on the beam positions on the return CCDs in the unit with the launch optics.

6 Beam Spot Fitting

To achieve the reference network performance discussed in section 1.5.1, the LSM must reconstruct the position and orientation to roughly $1-5\mu m/\mu rad$ [42]. Translating these requirements into the precision of the laser beam spot positions on the cameras is not trivial; the error propagation of the system is analysed in chapter 10. However, the translational error in reconstruction is roughly half the beam spot fit precision, leading to a beam position uncertainty requirement of around O($1\mu m$). Fitting the beams at a level of around one tenth of a pixel is difficult as many different effects need to be accounted for; both static and dynamic. This chapter covers the production of a position reading from a single image on each camera.

6.1 Conventions

When describing the properties of the beam spot on a camera, there are a few conventions that need defining to avoid confusion.

6.1.1 Beam Position

The co-ordinates of a fitted position on a CCD can be confusing. As cameras are normally used with lenses, unlike in the RTRS, the illumination pattern on the CCD is flipped with respect to the image saved to file. This means:

- The (0,0) pixel on the CCD chip is in the bottom-left hand corner when facing the chip.
- The (0,0) pixel is at the top-left hand corner of the image displayed.

A simple schematic of the CCD can be found in figure 6.1 with an example image. The image is then fitted to give a beam position in pixel-space, before being scaled to a physical length co-ordinate. Fitting is performed using the pixel co-ordinates of the image: increasing the X-co-ordinate moves the beam to the right and increasing the Y-co-ordinate moves the beam down on the image (and hence up on the CCD).

When the fitted position is converted into a physical co-ordinate, the origin is set to be the centre of the CCD. This was chosen as it is symmetrical and makes the set-up and visualisation of a ray-tracing model simpler.

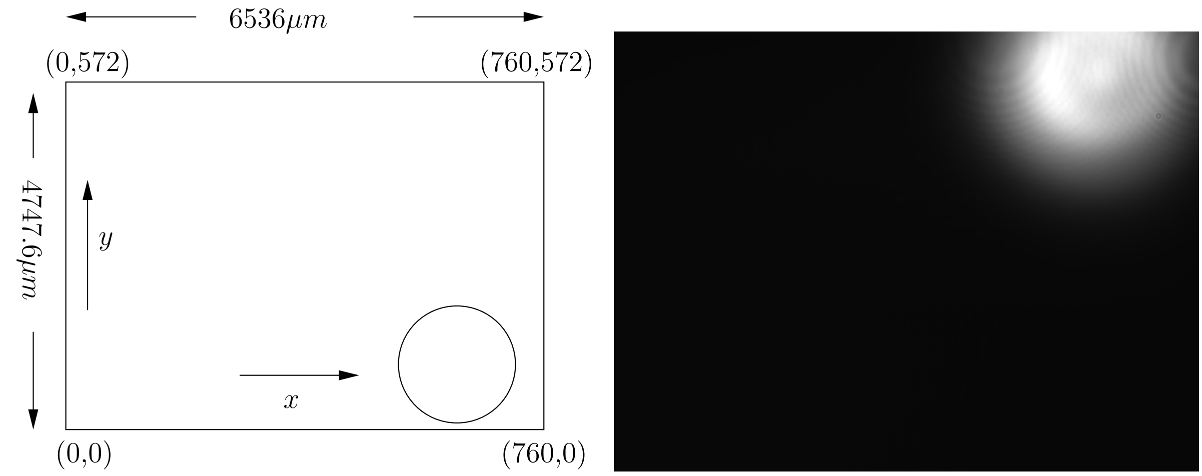


Figure 6.1: Simple diagram of the CCD chip with beam hitting the bottom right of the chip represented by a circle. The image produced by the camera is also shown.

6.1.2 Beam Width

There are numerous possible definitions of beam width. The one chosen to describe the results of the beam fits was the $\frac{1}{e^2}$ half-width; this corresponds to the parameter w in the Gaussian propagation equation, equation 4.1.

6.1.3 False Colour

For clarity, some images are displayed using false colours; the intensity mapping can be seen in figure 6.2. The look-up table was produced by selecting a 16 colour LUT and interpolating up to 256 colours. The image processing software ImageJ [43] was used to produce the false colour images. The display of residuals is either made with the same LUT when produced as an image or with a similar ROOT palette; in either case, the scale is shown.

6.2 Factors Influencing the Beam Spot Position Measurement

There are a number of challenges to overcome to ensure the process of turning the images into a beam spot position is accurate; some caused during beam propagation, some during the process of creating an image on the CCD.

- Dust

Despite the units being assembled in clean room conditions, dust managed to penetrate the units. This was mostly due to the need to open the units up in the poor air of the tunnel. There are two places where dust affects the LSM; dust on the CCDs themselves causes small (characteristic radius of order(30)μm) interference rings and dark patches in the images and dust on the pellicles

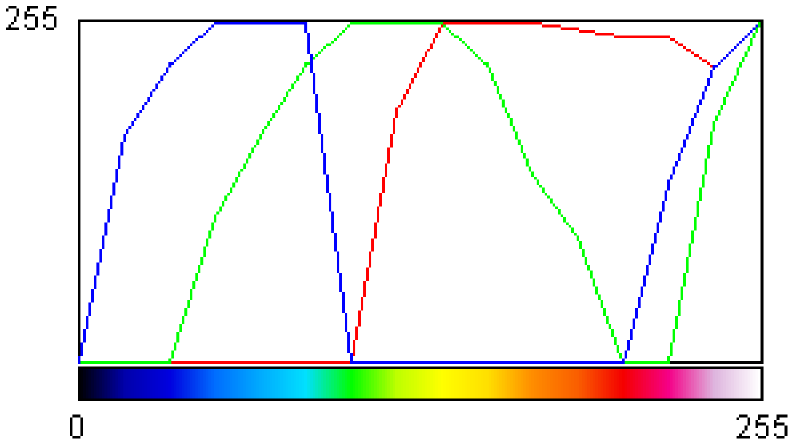


Figure 6.2: False colour intensity mapping, manually created to enable easy interpretation of images.

causes much larger (radius of order(700)μm) rings on the image. A detailed analysis of the effects of dust is given in appendix A.

- Beam Clipping

In some configurations the laser beam was clipped by the edge of a pellicle, which caused severe disruption to the beam. The effect was seen in either the ghost beam, the real beam or both. Figure 6.3 demonstrates the effect of clipping on the beam imaged. If clipping was observed the train was readjusted to a better alignment.

- Thin film interference at pellicles

Thin film interference combined with the different thicknesses of different pellicles caused significant variation of reflectivity. This effect is covered in detail in section 4.6.2.

- Beam Shape

For the cameras close to the collimator the beam has non-Gaussian elements, seen in figures 6.22 and 6.23. This adds an offset to the fitted position which is absorbed into the calibration constants of those CCDs. However, the offset only remains constant if the beam is entirely on the CCD and if the number of saturated pixels is constant, typically at zero. The effects of beam shape are investigated further in section 6.10.

- Beam Brightness

If the beam is too dim then the fit quality can be poor, due to the beam brightness being of a similar order to the CCD noise. If the beam gets too bright, some pixels saturate and are not

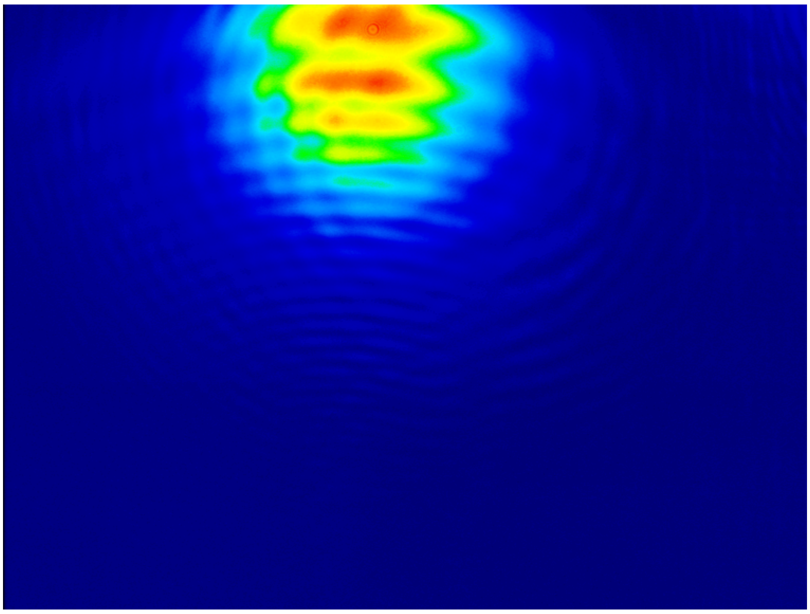


Figure 6.3: A 20-shot averaged image from Car 1 camera 2 demonstrating clipping (large arcs in a horizontal pattern) and part of a pellicle dust ring.

useful for fitting. Furthermore, if the level of saturation of an asymmetric beam changes, the offset caused by the asymmetry will change, leading to systematic errors (§6.10.1). As noted in section 4.2.2, the output power and spectrum of the laser change over time; meaning both the absolute brightness and the relative brightness of the beam on each CCD is variable. This can be kept within acceptable bounds by manually adjusting the attenuation before a run to give a non-saturating beam, and allowing the beam brightness to be a free parameter when fitting.

- Static Background

As noted in section 4.7, the cameras did not produce a constant dark reading for all pixels. Unless corrected, this would give a varying distortion on the reported beam position, so their dark background patterns were found and removed (§6.3.1).

- Poorly Performing Pixels

In some cameras the first three columns of pixels did not measure the incoming intensity well and so were excluded from use for all cameras.

- Electronic Cross-Talk

Electronic interference was observed when grabbing multiple cameras at the same time. Dynamic interference lines caused significant disruption of the fitting. While the effect on a single pixel is less than one intensity level, the overall coherent pattern is enough to shift the fitted position of the beam by up to five microns (§7.1).

6.3 Pre-Processing

There were numerous artefacts on the images that could be reduced or eliminated by pre-processing the images before attempting to fit. This sub-section describes these artefacts and the pre-processing used to minimise their effects.

6.3.1 Static Background Subtraction

As noted in section 4.7, the cameras produce a spatially varying but temporally static background image as well as time-dependent Gaussian noise. The static background component is found by recording 240 images for each camera with the laser turned off. An average image is produced and subtracted from all future images before analysis. A typical background image, normalised to demonstrate the pattern, is shown in figure 4.12. Background images were retaken if the camera gain was altered or before the start of any large data-taking session.

6.3.2 Fourier Filtering

Once the static background was removed, the image was filtered in the Fourier domain. The image is transformed into the Fourier domain using a 2D Discrete Fourier Transform in the FFTW libraries [44] wrapped with the FFTW++ [45] C++ classes. The image arrays are padded to make the number of rows and columns both a power of two for FFT efficiency and to reduce artefacts appearing when filtering. In the case of a full scale image the padded dimensions were 1024×1024 pixels. The image is padded with extra rows and columns which are assigned an intensity according to the roll-off function:

$$I = I_0 \left[1 - \left(\frac{r}{r_{\max}} \right)^2 \right]^2 , \tag{6.1}$$

where I_0 is the intensity of the pixel at the edge of the data, r is the distance from the edge and $r_{\max}$ is the number of extra pixels needed to add on that side. The corner rectangles are filled using the same function, using the end column generated from the padding operation as the starting values. This function ensures a smooth transition to zero at the edge and is used elsewhere [46]. A superior algorithm would take into account the gradient of the data near the edge, but this was not implemented due to time constraints and the acceptable performance of the roll-off of equation 6.1.

To remove the high frequency components of the data (noise and interference fringes) while preserving the low frequency components (the beams) a low-pass filter is applied. A simple step function is not used as this produces severe ringing. (This effect is covered in depth in [47]). Instead, the data is low-pass filtered using a Butterworth filter of the form:

$$H = \left[1 + \left(\frac{D}{D_0}\right)^{2n}\right]^{-1} , \tag{6.2}$$

where D is the spatial frequency number ($\sqrt{k_x^2 + k_y^2}$), D_0 is the cut-off frequency and n is the order of the filter. The data is then Fourier transformed back to give an image with a significant reduction in artefacts and much reduced noise.

6.3.3 Filtering Parameters

To determine the optimal parameters needed to reduce the effect of unwanted components while minimising distortion of the beam, some analysis was required.

The Fourier transform of a Gaussian function is also a Gaussian [32, p521], the width in frequency space being inversely proportional to the real space width. Consider the Gaussian beam $f(x)$ and its Fourier transform $F(k)$ where k is angular frequency.

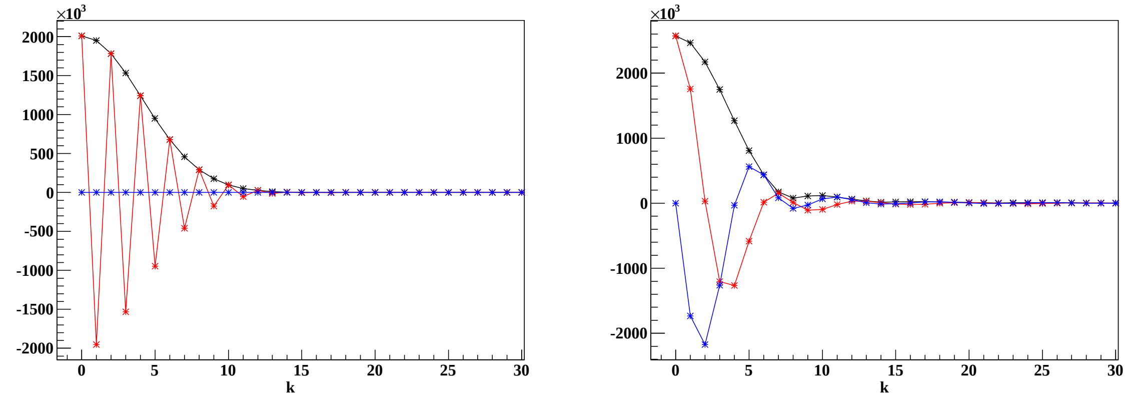
$$f(x) = \frac{1}{a\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x}{a}\right)^2\right) , \tag{6.3}$$

$$F(k) = \int_{-\infty}^{+\infty} \left(\frac{1}{a\sqrt{2\pi}} \exp\left(-\frac{x^2}{2a^2}\right)\right) e^{-ikx} dx = e^{-\frac{1}{2}(ak)^2} . \tag{6.4}$$

Note that the half-width at $\frac{1}{\sqrt{e}}$ intensity, the standard deviation, of the beam is a while in k-space it is $\frac{1}{a}$. In a discrete transform with N samples in real space, in k-space there are N evenly spaced spatial frequencies from 0 to $N-1$ cycles per sample.

Figure 6.4 shows the first 31 k_x-value coefficients for the first row ($k_y = 0$) of the spectrum of a 40 pixel wide beam with the beam centred in the image, and with some of the beam off-CCD. Figure 6.5 shows part of the $k_y = 0$ region of spectrum for a simulated dust interference pattern with no beam present before filtering, and a smaller region after the application of a 2nd order, $k = 16$ cut-off filter. Note the large amount of the signal apparent in the higher frequencies before filtering and the significant reduction in signal above $k = 20$ in the filtered spectrum. Table 6.1 shows that the filter removes less than one percent of the signal while removing the majority of the interference pattern.

To determine the optimal values of cut-off frequency and order, a number of simulations were performed. First a typical beam of standard deviation 100 pixels and amplitude of 200/255,



(a) Centred beam

(b) Beam on edge of camera

Figure 6.4: The first 31 spatial frequency (k_x values, with $k_y = 0$) coefficients of the Fourier spectrum for a simulated 40 pixel wide Gaussian beam, demonstrating the effect of truncation. The left-hand plot is for a central beam with negligible truncation, while the right-hand plot is for a beam centred 10 pixels from the left-hand edge of the image. The black points are the magnitude, the red points are the real part and the blue points are the imaginary part; the spectrum is un-normalised. Lines have been added to the discrete values to aid the eye.

Table 6.1: Effect of a Butterworth filter (cut-off: 10, order: 1) on the r.m.s pixel values of image components. (In arbitrary units.)

	Beam in centre	Beam on edge	Dust spot
RMS pixel value before filter	0.0318546	0.025847	0.00301
RMS pixel value after filter	0.0315911	0.025596	0.00058
reduction	0.99%	0.97%	80.73%

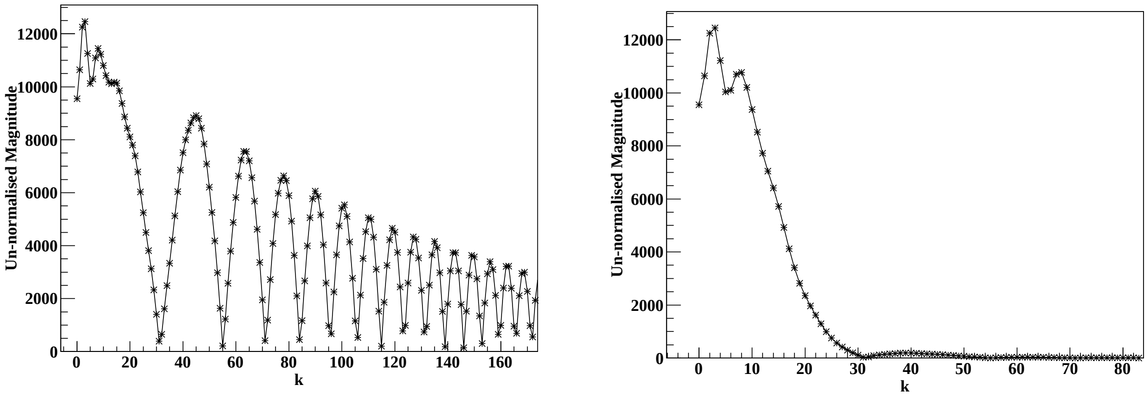


Figure 6.5: The first spatial frequency (k_x values, with $k_y = 0$) coefficients of the Fourier spectrum of a dust speck, demonstrating the reduction of high k-value components due to spatial filtering. The left-hand plot shows the spectrum created for a dust interference pattern of characteristic radius 30 pixels. The right-hand plot shows the effect of applying a 2nd order Butterworth filter with a cut-off frequency (k value) of 16.

centred at (50, 250)[5] was generated, then noise from the the CCD 0 spectrum was added to the image. The image was filtered using a variety of orders and cut-off frequencies, and the difference between the resultant image and the original, noiseless image plotted. A comparison of orders can be found in figure 6.6 which shows the large and far-reaching distortion by the first order filter, with the significantly smaller and more localised distortions caused by the second and third order filters. The final settings used were a 3rd order Butterworth filter with a cut-off k-value of 32 which gives artefact reduction with minimal beam distortion. The behaviour of fitting of filtered images is further examined in section 6.3.4 and section 6.8.

6.3.4 Filtering Effectiveness

Figure 6.7 shows the filtering of a real beam. It demonstrates a clear reduction in the prominence of artefacts, as well as a general smoothing of the noise. The difference image also gives an indication of the increased level of noise experienced by bright pixels (§4.7), which nevertheless is effectively smoothed away. There is no visible distortion of the beam near the edge of the image, showing the effectiveness of the filter choice and edge roll-off.

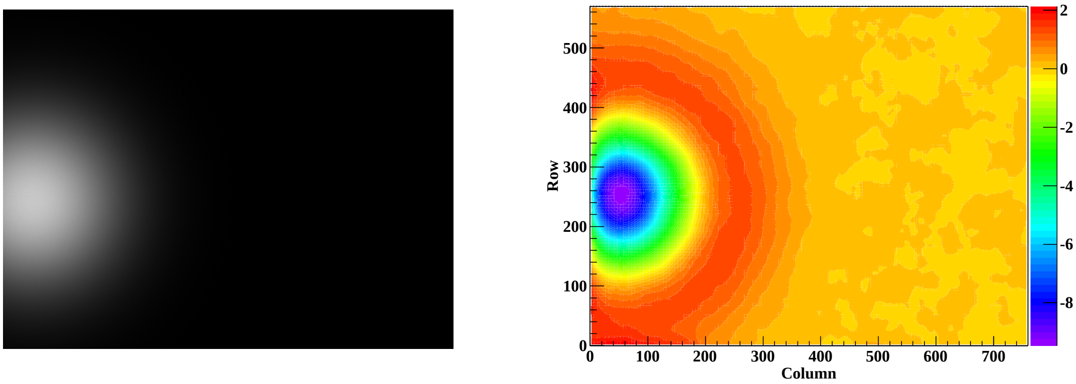
Filtering clearly has an effect on the noise spectrum of the cameras as found in section 4.7.3. Figure 6.8 shows a more extreme case; a camera at high gain with an amplitude range of 11-13 of of 255, giving a dark, noisy image with significant digitisation effects. Filtering reduces the RMS five-fold and eliminates the digitisation effect on the spectrum. For the effect on the noise over different pixel brightness ranges, compare figure 6.9, which shows the RMS noise produced for pixels of different amplitudes, with the unfiltered versions in figure 4.18. There is a significant reduction in noise, particularly for high gain and low amplitudes, and the noise function is more linear. The noise spectrum generation procedure of section 4.7 still works with filtering applied, after background subtraction.

6.4 Ghost Beams

Due to the arrangement of the cameras, light can reflect off the CCD cover-plate and hit the camera on the opposing side of the unit. The amplitude of the reflection was reduced by coating the cover-plate of the CCDs with an anti-reflection coating. However, this was only partially successful as this only stops reflections from the front face of the plate; light still reflected off the back face and the silicon surface itself. The size and brightness of the ghosts can be seen in table 6.3 and table 6.4 towards the end of the chapter, which gives typical parameters found

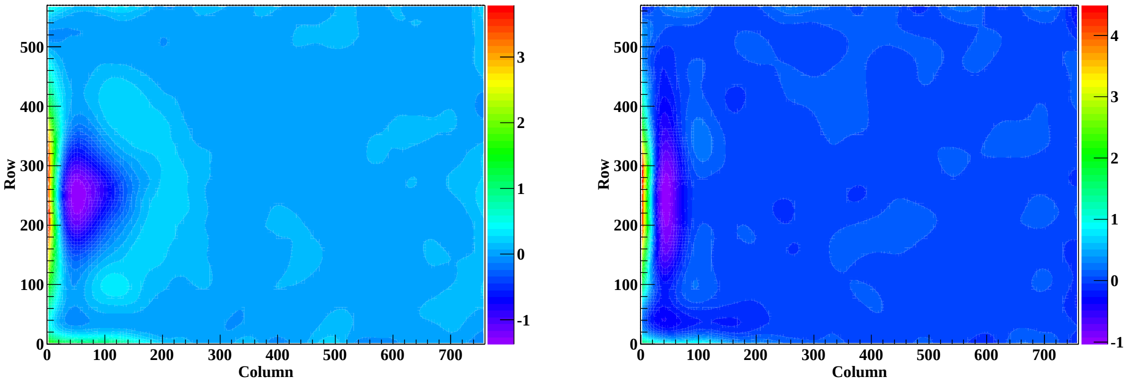
---

[5] This produces a beam with a significant fraction off-CCD.



(a) Original Image

(b) 1st order



(c) 2nd order

(d) 3rd order

Figure 6.6: The distorting effect of Butterworth filtering on an ideal, dust-free beam: the colours represent the pixel intensity (8 bit) residuals against the original image (top left, flipped for ease of comparison) of a typical beam after the application of a Butterworth filter with a cut-off of 10 for the first three filter orders.

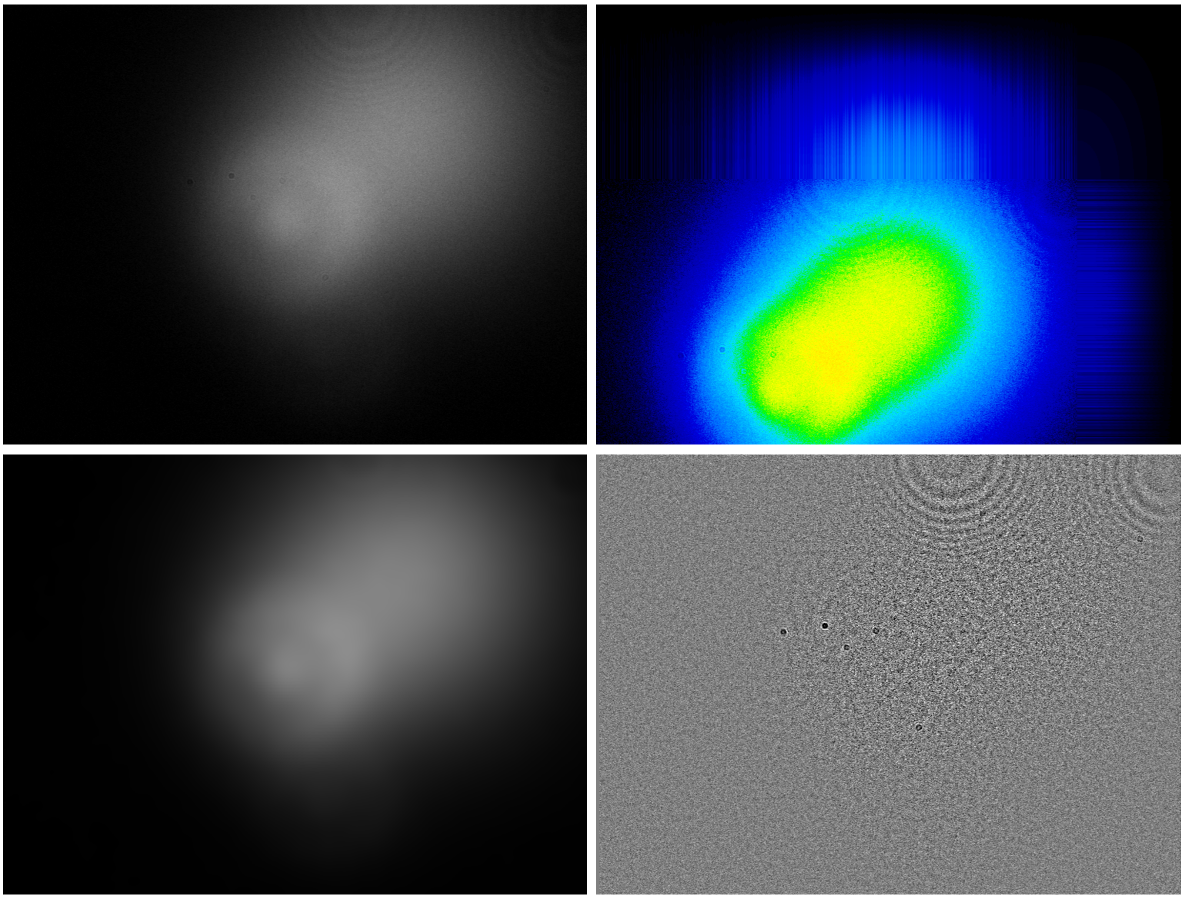


Figure 6.7: The effect of filtering on a representative beam image: Top-left: A typical image taken from Car 2, CCD 0 showing noise, dust interference rings, on-CCD dust and a prominent ghost beam, with the static background removed. Top-right: false colour image showing part of the beam and the roll-off padding. Bottom-left: image after filtering. Bottom right: Difference between the top-left and bottom-left images, with the amplitude scaled up by a factor of ten.

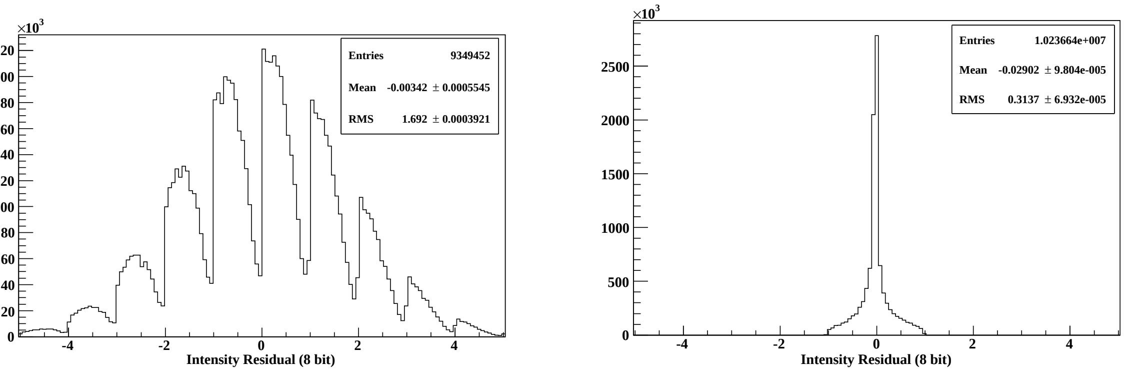


Figure 6.8: Pixel counts of intensity residuals against an average image from car 2, CCD 0 data. Left: the resultant noise spectrum shown at high gain before filtering. Right: the resultant spectrum after filtering.

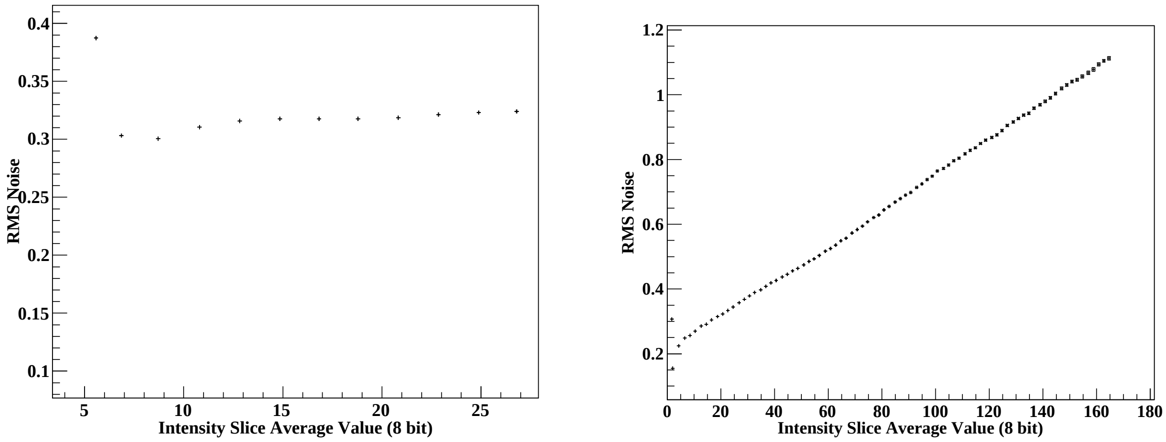


Figure 6.9: RMS noise levels for a high gain, low beam intensity camera: car 3, CCD 0 (left); and a low gain high beam intensity camera: car 3 CCD 1 (right) after filtering. The anomalous point in the car 3, CCD 0 image is due to low statistics.

during a run. The total spatially integrated irradiance on the CCD from the primary beam, $\mathcal{I}_p$, is as follows:

$$\mathcal{I}_p = (1 - R_p)^n \, R_p R_r \mathcal{I}_0 \, , \tag{6.5}$$

which uses the rough approximation that the pellicles each reflect the same fraction of light. R_p is the intensity reflectivity co-efficient of the pellicle, R_r is the reflectivity of the mirror in the retro-reflection assembly (if encountered, unity if not), I_0 is the initial spatially integrated irradiance and n is the number of pellicle interactions the beam undergoes before hitting the CCD. Ghost beams were only observed on the return cameras and their integrated irradiance, $\mathcal{I}_g$, is given by:

$$\mathcal{I}_g = R_p R_c \, (1 - R_p)^n \, \mathcal{I}_0 \, , \tag{6.6}$$

where R_c is the reflectivity of the CCD. With the launch in car 1, the camera with the brightest ghost beam relative to the primary is CCD 1, where the light for the primary beam travels the furthest and has the most pellicle interactions (11 transmissions and a reflection) and has hit the end-cap retro mirror before hitting the CCD;

$$\mathcal{I}_{p_1} = 0.92^{11} \times 0.50 \times 0.08 \times \mathcal{I}_0 = 0.016 \mathcal{I}_0 \, . \tag{6.7}$$

Conversely, the ghost has travelled the least distance and had the fewest interactions (1 pellicle transmission, 1 pellicle reflection at 8% , and 1 CCD reflection at 4%);

$$\mathcal{I}_{g_1} = 0.08 \times 0.92 \times 0.04 \times \mathcal{I}_0 = 0.0029\mathcal{I}_0\,. \tag{6.8}$$

In addition, the change of the beam width as it propagates means the intensity of the beam centre is accordingly different for the primary beam. In the case of camera 1, the primary half-widths were 3022μm and 2875μm in the X-axis and Y-axis respectively, while the ghost half widths were 1688μm and 2010μm (table 6.3). Therefore the central amplitude ratio of the beams would be expected to be:

$$\frac{I_{g_1}}{I_{p_1}} = \frac{\mathcal{I}_{g_1}}{\mathcal{I}_{p_1}} \times \left(\frac{A_{p_1}}{A_{g_1}}\right)\,, \tag{6.9}$$

$$\frac{I_{g_1}}{I_{p_1}} = \frac{0.0029}{0.016} \times \left(\frac{3022 \times 2875}{1688 \times 2010}\right) = 47\%\,. \tag{6.10}$$

Table 6.2 shows the expected primary and ghost beam properties for the various cameras where ghost beams were present; for other cameras the ghost beams were not visible, due to the beam missing the camera. The agreement with the observed amplitude ratios is poor, although the general trend is predicted correctly. The assumption of similar reflectivity of pellicles, and potentially the CCDs themselves, is a poor one. However, a full understanding of the ghost beam intensities is not required to be able to account for them in fitting.

6.5 Seeding Methods

A variety of methods were developed to determine the properties of the beam hitting each CCD without using fitting. The methods described below are useful for providing good starting values for full fitting (§6.6.5.)

6.5.1 Simple Beam Parameter Estimates

The simplest method is to find the brightest row and column in an image: summing the intensities for each column and row and returning the ones with the largest total. The brightest column or row might not correspond to the centre of the beam due to multiple beams overlapping. It also cannot account for asymmetrical beams, leading to errors of the order of hundreds of microns. However it always rapidly produces a rough initial guess.

Table 6.2: Observed and expected ghost/primary amplitude ratios for a selection of cameras. The top section is from real fit results with data from stability runs, while the lower section is calculated using equation 6.9.

	Launch in car 1				Launch in car 3		
CCD	1	3	5	9	0	2	4
Primary Amplitude	106	126	180	152	161.6	165.4	123.4
Ghost Amplitude	73	60	37.8	14.5	14.2	22.69	84.5
Measured Amplitude ratio	**69%**	**48%**	**21%**	**10%**	**9%**	**14%**	**68%**
Area Ratio	39%	31%	70%	48%	102%	106%	53%
Primary pellicle transmissions	11	10	9	7	6	7	8
Primary intensity fraction	1.6%	1.7%	1.9%	2.2%	2.4%	2.2%	2.1%
Ghost pellicle transmissions	1	2	3	5	5	4	3
Ghost intensity fraction	0.29%	0.27%	0.25%	0.21%	0.21%	0.23%	0.25%
Expected Amplitude ratio	**47.2%**	**50.7%**	**18.9%**	**19.8%**	**8.6%**	**9.7%**	**22.9%**

Estimates of the X and Y widths of the beam were made by reading the pixel intensities in a line out from the centre found above, and stopping when the value read became smaller than the initial value by a factor of $\sqrt{e}$.[6] The process is performed in both the positive and negative directions and uses the smaller width, unless one direction encounters the edge before falling to the required value. This gives a reasonable estimate of the width (typically correct to the order of 50 microns) very rapidly and is used for seeding more accurate fitting methods. Moment analysis of images was also performed, but did not give better results than simple estimates and was particularly vulnerable to errors when a ghost beam was present.

6.5.2 Box Analysis

Box analysis works by placing a ‘box’ defining a region of interest on the image, and summing the intensity of all the pixels it covers. The box is then moved one row or column at a time in the direction that increases the summed intensity. The process is stopped when no direction can be found that increases the sum. This process was used to produce an effective seed for two dimensional fitting. Box fitting has the advantage that it can distinguish two overlapping beams but the process must be started with at least some of the beam in the box. The process is fast as only the intensity around the edge must be summed for each step. Box fitting with a square box

---

[6] This is the intensity fraction at half the beam width for a Gaussian beam, or one ‘Standard Deviation’ if the Normal Distribution analogy is considered.

with a size of 60 pixels was found to be effective at producing a good beam position for use for seeding later fits, although no method was developed for finding the beam widths using boxes - the simple width algorithm described above was adequate.

6.6 Fitting Beams

The best results were found by fitting the beam to a model, as this method is far less affected by noise and interference. This fitting was based on the concept of least squares presented by C.F. Gauss [48, Third Section]. Least Squares Fitting is covered in more detail in appendix B and the particular minimisation method is covered in section 6.6.3.

6.6.1 Projection Fitting

Projection fitting simply projects the image onto both axes and fits two separate 1-D Gaussians to them. This has the advantage that it is extremely fast for a fit as it is conducting two 3 (4 if an offset is allowed) parameter fits, one with 760 data points and one with 572 data points. The full two-dimensional fit detailed later conducts one 6-parameter fit with 434,720 data points. However, projection fitting is more sensitive to artefacts and struggles with multiple beams. This can be understood by considering beams whose centres are close in the X-axis, but are fairly separated in the Y-axis. It is trivial for a two dimensional fit to find the beam parameters but significantly more difficult for the X-centres to be found on the projection, particularly if the relative intensities are not known; there are also complications in linking the pairs of results together. Similarly, an interference pattern may be such that the chi-squared cannot be lowered significantly by changing the parameters in the two dimensional fit and so is effectively ignored. However, the same pattern may affect a projection fit: for example, if the pattern is a long way from the beam centre in one axis but not another.

6.6.2 Two Dimensional Fitting

The main method of fitting used in this work tries to reproduce the image using parametrised two-dimensional functions. To remove any effects from saturated pixels they are excluded from the fit: all pixels that had an intensity of 254 or 255 before the background was subtracted, are marked as bad and excluded.

To give the error estimate on a pixel by pixel basis, a map of the pixel intensity errors is created by using the parametrised noise function given in section 4.7.5, and adjusted in section 6.3.4.

This map is then used during the minimisation as a cache of the errors associated with each pixel value. The fitting procedure is summarised below:

1. Start with a blank image of the correct dimensions.
2. Add a constant intensity to all pixels, set by the fitter.
3. Add the parametrised intensity distributions of one or two beams to the image, with beam parameters set by the fitter.
4. Loop over all the valid pixels calculating the difference from the data, dividing by the error estimate for each individual pixel, then squaring.
5. The sum is the χ^2.

The beam parameters are adjusted by the fitter until a minimum is found or an error state reached.

6.6.3 Fitting Package Used

The package used to perform the minimisation was Minuit2 [49], which is integrated into the ROOT framework [50]. Minuit2 is a robust and accurate function minimiser with multiple minimisation strategies, described more in appendix B.7. The Migrad variable metric strategy was found to perform well, given reasonable starting values, and was therefore used for all beam fitting. There are a few minor pitfalls that need to be avoided:

- It will only find a local minimum. This means that initial estimates and careful use of limits are very important for our minimisation. For example, if the initial estimate of beam position is well away from the correct one, the chi-squared can be reduced by reducing the beam amplitude or simply moving it off screen! This issue was avoided by having good initial estimates.
- Parameter limits can bias the minimum position and its error estimates, particularly if the minimum is near a limit. To avoid this issue, if limits were used then upon initial convergence, a final fit was performed with all parameters free.
- The Migrad method relies on the use of derivatives, so can struggle with discontinuities and may give unexpected answers when a relationship with a one-to-many mapping is used. This did not affect beam fitting, though was relevant for unit co-ordinate reconstruction later.

6.6.4 Beam Functions Used

The fitting package was written to be able to use different functions to fit the beams. A few different functions were implemented:

1. Two dimensional Gaussian with different X and Y widths.

$$I = I_0 \exp\left[-\frac{1}{2}\left(\frac{x-x_0}{\sigma_x}\right)^2 - \frac{1}{2}\left(\frac{y-y_0}{\sigma_y}\right)^2\right] \tag{6.11}$$

2. Two dimensional Gaussian with the Y width a multiple of the X width.

$$I = I_0 \exp\left[-\frac{1}{2}\left(\frac{x-x_0}{\sigma}\right)^2 - \frac{1}{2}\left(\frac{y-y_0}{R\sigma}\right)^2\right] \tag{6.12}$$

3. Two dimensional Gaussian with the y' width a multiple of the x' width and the axes rotated by an angle, θ, relative to the axes of the image.

$$\begin{aligned} I &= I_o \exp\left[-\frac{1}{2}\left(\frac{(x-x_0)\cos\theta + (y-y_0)\sin\theta}{\sigma}\right)^2\right. \\ &\left. -\frac{1}{2}\left(\frac{-(x-x_0)\sin\theta + (y-y_0)\cos\theta}{R\sigma}\right)^2\right] \end{aligned} \tag{6.13}$$

4. A stored template image of the beam with the brightness, X and Y positions as free parameters.

The second function is more effective than the first, as it is faster to fit due to the natural correlation of the widths, and enables the ratio to be fixed at strategic points in a minimisation. It was used for almost all fitting cases. The third function was implemented for use when the axes of the beam spot ellipse were not aligned with the CCD axes. This was only evident on the first two cameras hit (figure 4.5a). The final function removes the need for an analytical form by using a reference image of the beam, the centre of which had already been determined by fitting with function 2. Sub-pixel precision can be achieved by interpolating. While promising, it was not used extensively due to the difficulty of obtaining a good reference image due to stationary artefacts like dust and ghost beams.

6.6.5 Seeding Procedure

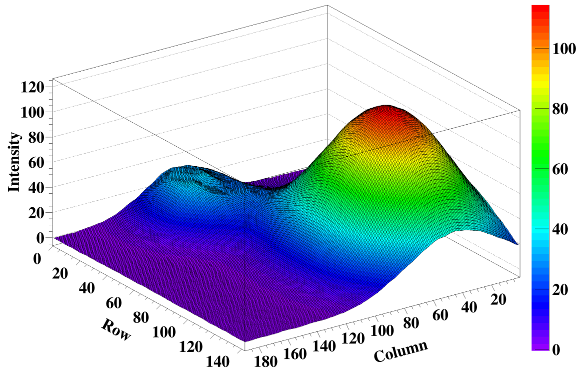
It is important to seed the fits well otherwise the fit may take a long time to find the correct minimum, find an incorrect minimum or fail to minimise at all. The chain of procedures to create a good seed for the main fit is:

1. A starting central point is found by using brightest row and column in the image.
2. A box fit is performed around the seed point; this reduces the effect of two intersecting beams.
3. The central intensity in the optimal box position is read and a beam width estimate is performed using the simple method described in section 6.5.1.

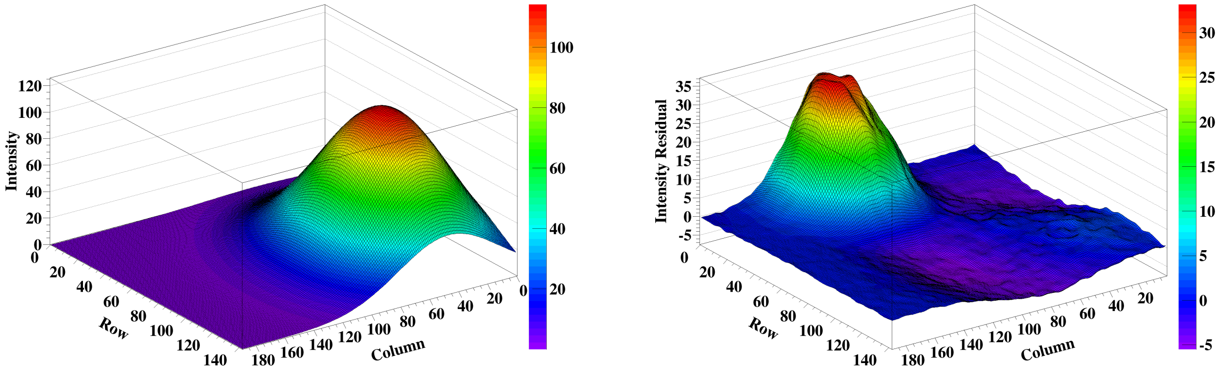
The resulting parameter values can then be used for the starting point for the full fit described in section 6.6.2. If the image is expected to have ghost beams (table 6.2) then they also need to be seeded, which is done as follows:

1. A primary beam with the characteristics found using the procedure above is computed.
2. The beam is then subtracted from the image. This leaves the secondary beam as the brightest part of the image.
3. The secondary beam is found by finding the brightest row and column and making a width estimate. A Box fit was not found to be needed.

In principle this procedure could be repeated for multiple beams, though in practice the greatest number of beams on a camera was two. Figure 6.10 shows the process in action, with intermediate images showing the properties of the seed beams and the residuals with the real image. The result is an excellent set of initial parameters, with the seeds correct to a small fraction of the beam width and height.

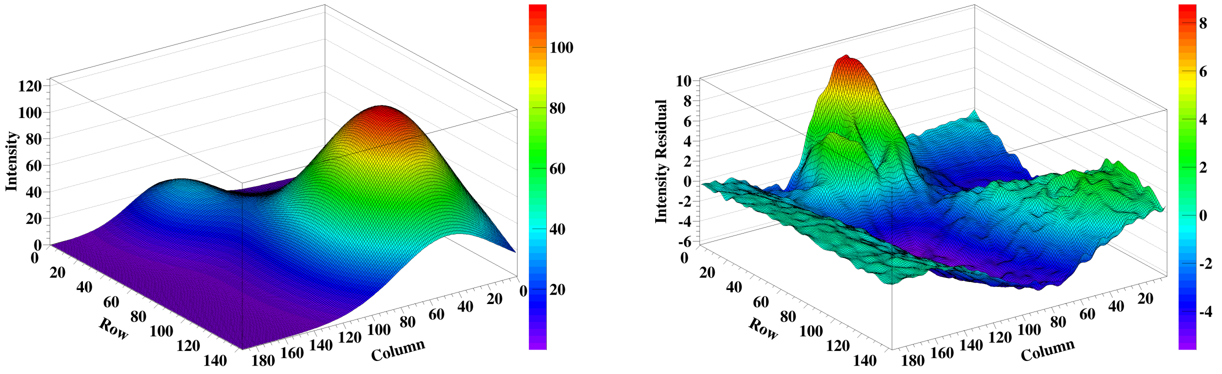


(a) Pixel intensity distribution of the image to be fitted.



(b) Primary seed beam intensity distribution.

(c) Residual after subtracting primary seed beam.



(d) Intensity distribution of both seed beams combined.

(e) Residual after subtracting both seed beams.

Figure 6.10: The fit seeding process for a typical image at one quarter scale, taken from car 2, camera 1 with the launch in car 1. (a) A view of the image to be analysed, showing a primary beam with an amplitude of around 100 units, and a narrower 'ghost' beam with an amplitude of around 30 units. (b) The primary seed beam. (c) The residual after removing the seed beam from the initial image. (Note the reduction in the intensity scale.) (d) The intensity distribution produced by the combination of both seed beams. (e) The residual after both seed beams have been subtracted.

6.7 Full Fitting Procedure

6.7.1 Re-sampling Beams

The images can be re-sampled to a fraction of the original size, reducing the number of pixels and hence dramatically improving the speed of the fit. The re-sampling was performed by grouping pixels into ‘superpixels’ containing 16 subpixels. Using a 1:4 ratio means there are no residual pixels unused in a 760 × 572 image. There are a couple of adjustments that need to be made if the image is re-sampled. The pixel error estimate is reduced by a factor of the square root of the number of sub-pixels in the super-pixel. There is also an offset introduced by the rescaling; a function filling pixel according to the function $f(i, j)$ at full size would need to fill according to the function $f(i', j')$ where:

$$i' = i + \left(\frac{S-1}{2S}\right); \quad j' = j + \left(\frac{S-1}{2S}\right), \tag{6.14}$$

with S being the scale factor. The re-sampled images give fits close to the final full-size results: figure 6.11 shows a negligible bias and an RMS residual of less than 0.03 microns. This suggests that the number of pixels is in excess of what is needed for precise fitting and the precision is limited by other factors. However, all the images were fitted at full scale, using 1:4 scale fits as a seed.

6.7.2 Method

The full procedure for fitting an image is as follows:

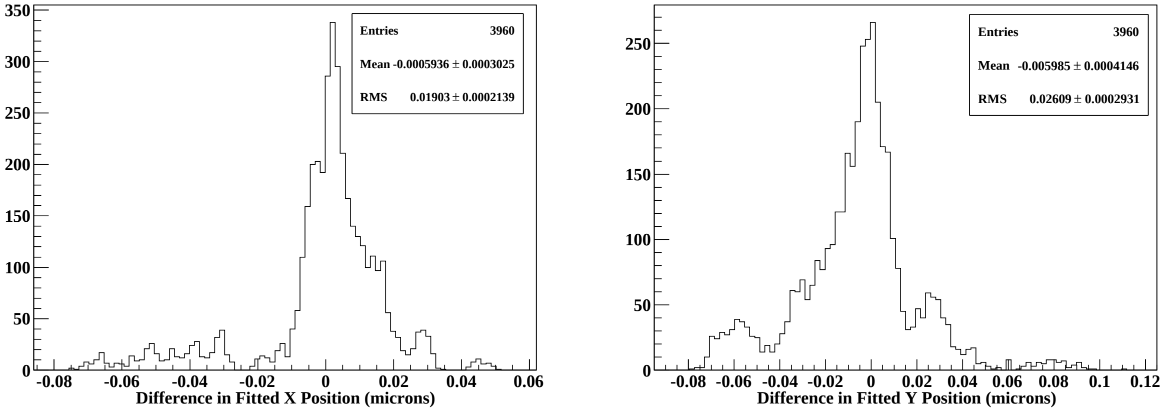


Figure 6.11: The pixel count against X-position (left) and Y-position (right) residuals between the 1:4 scale image fits and the full scale image fits for car 2, CCD 0 over a set of 99 different orientations with 40 images per run.

1. The image to be fitted is loaded from file
2. The background image corresponding to the camera used and the train configuration is subtracted from the image.
3. The region of interest of the image is set.
4. The image is filtered in the Fourier domain if set to do so (§6.3).
5. The seed parameters are created using the procedure in 6.6.5 with the images at one quarter scale. Whether or not to fit ghosts depends on the CCD and is set by a set-up file.
6. The parameters found in the step above are then used as the seeds for a full 2D fit at one quarter scale, with the beam function(s) dictated by set-up file.
7. The results from the step above are used to seed the full sized 2D fit which is now performed.

The whole process takes around 20 - 30 seconds for a one beam fit and around 80 - 120 seconds for a two beam fit, though multiple images are fitted at once. When the minimisation has finished, the results are checked for issues reported by the minimiser; such as an invalid minimum or convergence failures; or parameters away from expected values such as too wide, width mismatch, too dim, coincident primary/ghost, too bright, partially-off-CCD or far off-CCD. If an error is encountered, the seeding and minimisation is repeated using an auxiliary parameter set from the set-up file. Minor issues like partially-off-CCD are only warnings and do not trigger a re-fit. For an image which is expected to have a ghost beam present, the auxiliary set-up is a single beam; this means that if the ghost beam is off-CCD then the primary beam is fitted correctly second time around. After a repeat fit, the error states are compared and the most valid minimisation is chosen, comparing chi-squared values if the error states are the same. This was found to give robust fitting performance, with a few failures per thousand with typical beams.

6.8 Gaussian Fitting Simulation

To produce an estimate of the fitting performance, the process can be performed on images which have been generated computationally. A simulation generated beams with centres at various positions on or near the CCD, adding dust pattern interference and CCD noise, then fitting the resulting image. A grid of 51 by 51 positions was created starting 100 pixels (860 microns) to the left of the edge of the CCD and 100 pixels (830 microns) below the CCD, and ending at the corresponding positions at the opposite corner. Then for each point on the grid:

1. The intensity distribution of a Gaussian beam with typical parameters (amplitude 200/255, width 100 pixels) was added to a blank image, with the centre at the corresponding point in the grid.
2. Interference patterns were generated and added to the image for two pieces of dust placed at (150, 400) and (400, 200).
3. Noise was added to the image using the spectrum of CCD 0.
4. The image was fitted and the results written to file.
5. The image was then filtered with a 3rd order Butterworth filter set with a cut-off k-value of 32.
6. This filtered image was then fitted and the results written to a second file.

Figure 6.12 shows a typical image generated by the simulation. Figure 6.13 shows the fitted position residuals for all the points in the grid while figure 6.14 shows a histogram of the residuals only for beams with a centre on-CCD. Figure 6.13 shows that when the beam centre is not on-CCD, the fit error increases with distance quickly, especially if Fourier filtering is enabled. As beams are not expected to be off-CCD the following observations are only for parts of the grid where the beam centre is within the light sensitive area, marked with a green rectangle. Figure 6.13 shows that the dust has a significant pushing effect on the fit positions. For ring 'A' the residual is up to six microns but is reduced to 4 microns after filtering. The interference ring 'B' affects the fit by up to 4 microns without filtering and 1.5 microns with filtering. The improvement is also evident in figure 6.14 where the tails of the distributions are reduced, leading to a lower overall RMS residual. Similar reductions are apparent for other fit parameters.

However, figure 6.15 considers only fits where beam is well away from dust specks; the range (25,35) to (43,43) in the grid (marked with a red rectangle in figure 6.13). Here, filtering has a slightly detrimental effect; for the X-position the RMS residual increases from 0.07 microns to 0.12 microns, while the Y-position RMS residual remained at 0.11 microns. A small constant offset does appear with filtering: the X-position mean residual increases from 0.009 microns to 0.08 microns, while the Y-position mean residual increases from -0.001 to 0.16 microns. This suggests that filtering is not helpful if beam spot artefacts can be eliminated completely, and any reported fit precision better than 0.1 microns should be viewed with caution.

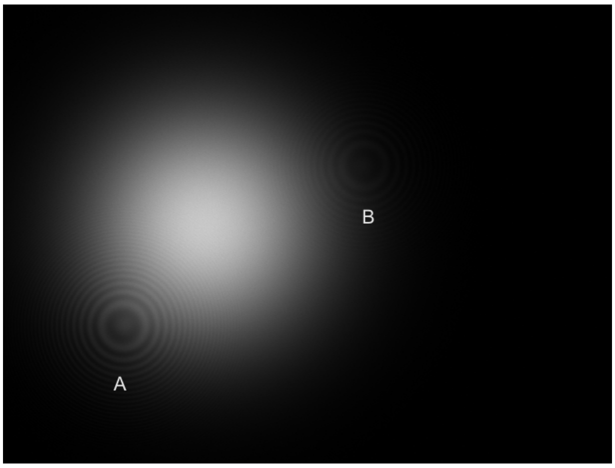


Figure 6.12: A typical simulation image. The two dust rings in the simulation are labelled 'A' and 'B'; 'B' has low visibility.

6.9 Fitting of Real Beams

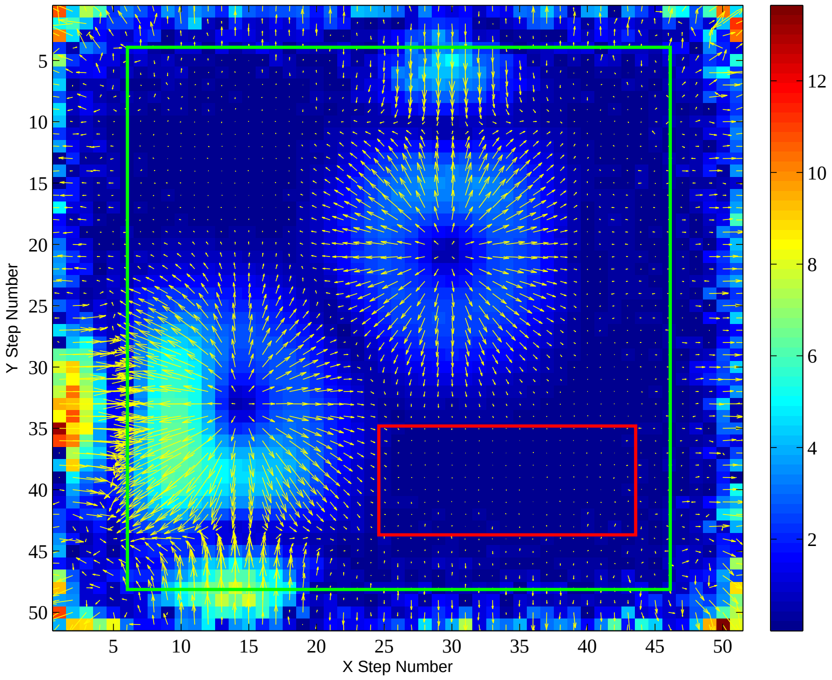
Figure 6.16 shows shows a typical beam spot image (car 2, CCD 1) with its fit and the resultant residuals with filtering disabled, while figure 6.17 shows the same plots with filtering enabled. They demonstrate the significant reduction in noise and interference artefacts by the filter, leaving only the low spatial frequency beam shape anomalies. The fits give a good representation of the beams hitting the CCD with no positional bias apparent.

Figure 6.18 shows the distribution of intensity residuals with and without filtering for two different areas of the image, one containing the primary beam and one containing only dark pixels, as well as for the whole image. The residual distribution for the whole image appears to be a superposition of the two sub-areas analysed, which is reasonable as they contain most of the CCD area between them. With filtering, there is a almost complete elimination of Gaussian noise (and the digitisation effect seen in the dark region) down to at least a 0.15 unit RMS, leaving the residual distribution due to the beam shape anomalies to dominate.

6.10 Reference Beam Simulations

6.10.1 Beam Saturation

The effects of saturation were simulated using template beams, produced by averaging a set of 40 images with the static background removed. The first set was taken from car 1, CCD 1, and the second from car 2, CCD 1, both with the launch in car 3 (figure 6.19). The car 1, CCD 1 beam is generally symmetrical and Gaussian in appearance, with deviations less than 1% of the calculated beam brightness, apart from a single dust speck. The car 2, CCD 1 beam has clear



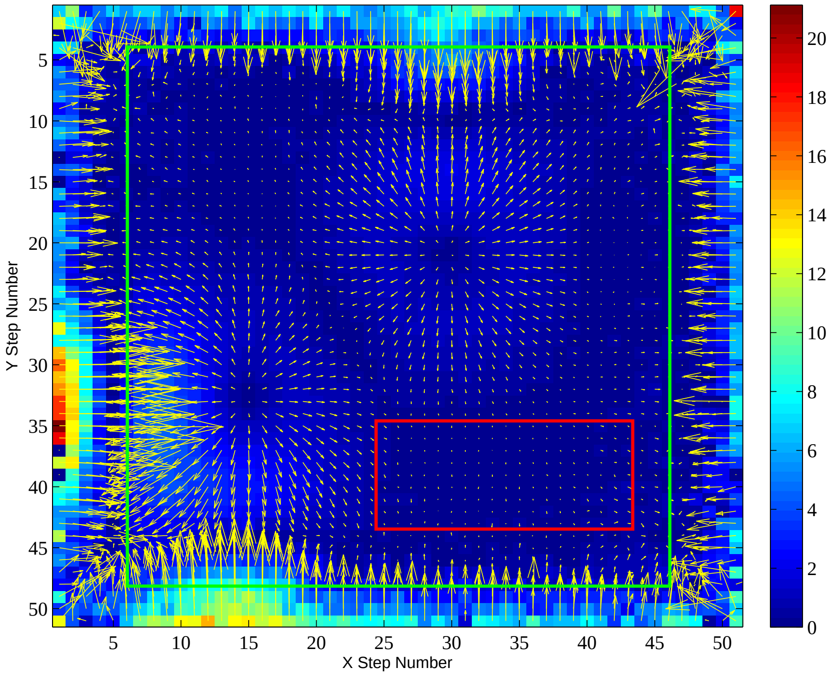


Figure 6.13: Unfiltered (top) and filtered (bottom) beam spot fit simulations. Points within the green rectangle are where the beam centre is on the CCD. The colour scale shows the difference in microns between the fitted position and the truth in microns. The arrows show the direction of the error; a right pointing arrow corresponds to a fitted X-position being more positive than the truth. An up-pointing arrow corresponds to a fitted Y-position being more negative than the truth. While the colour schemes cover a different range, the arrows are to the same scale.

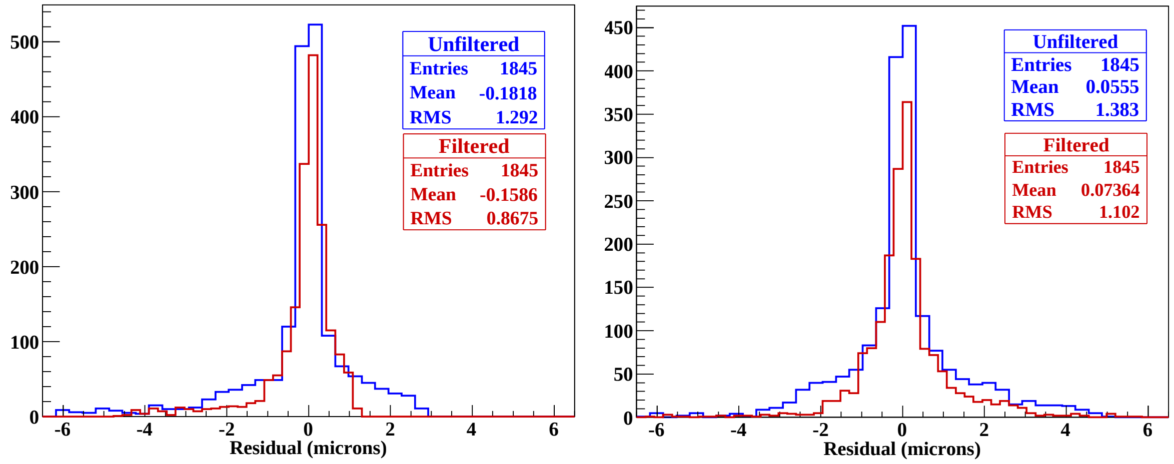


Figure 6.14: X-position residual distribution (left) and Y-position residual distribution (right) for the beam fitting simulation using data where the beam centre is on-CCD. The results for the filtered images (in red) show a 33% reduction in the RMS residual for X-positions and a 20% reduction in the RMS residual for Y-position over the unfiltered images (in blue). Filtering does not introduce any noticeable bias.

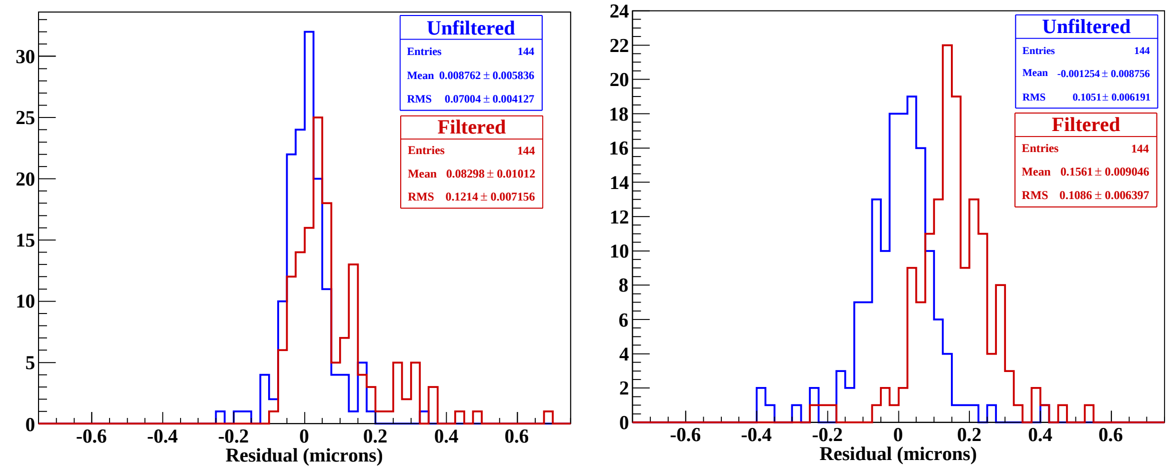


Figure 6.15: Histograms of the distribution of residuals of the X-position (left) and Y-position (right) for filtered and unfiltered simulations using only areas far from dust spots.

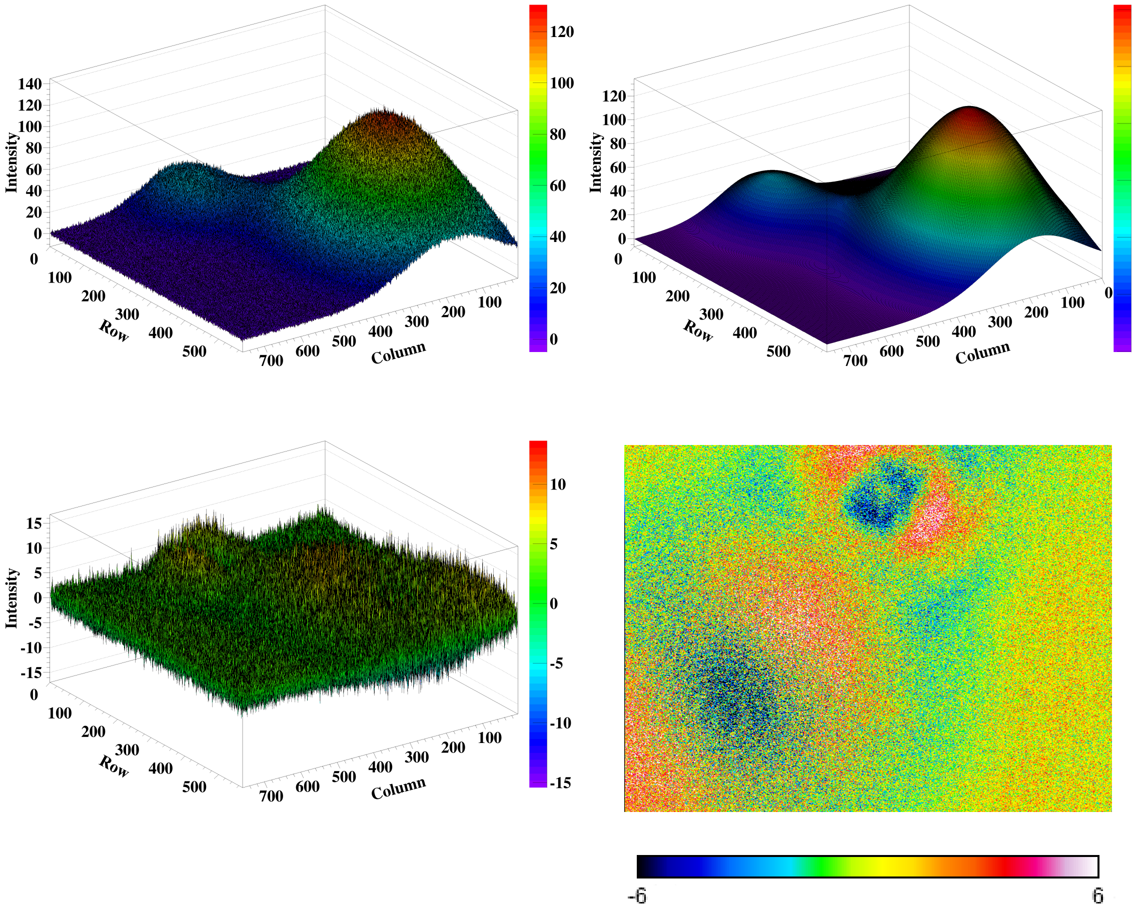


Figure 6.16: Fitting a single unfiltered image taken from Car 2, CCD 1. Vertical scales are in pixel grey-scale units. Top left: Original 8-bit image. Top right: Fitted beam profiles. Bottom left: Fit residuals. Bottom right: Another view of the same residuals, clipped to a narrower range.

non-Gaussian properties, with 15% deviations from the calculated brightness; these appear to be the lobes of a TEM_{01} mode. The difference plot in figure 6.19 shows the real car 2, CCD 1 beam is much dimmer than a Gaussian in the centre, with lobes which are brighter. Both show slight effects of dust, with an on-CCD speck particularly visible in the car 1 residual image, though there is no reason to expect issues due to this. The outer ring visible in the car 1, CCD 1 image is appears to be like a Laguerre-Guassian-10 mode [51].

The pixel intensity values in the reference images were multiplied by a scale factor and fitted repeatedly as scale factor was increased from 0.025 to 4 in steps of 0.025. Figure 6.20 gives the number of saturated pixels, the fitted intensity:scale-factor ratio, the beam width, and the difference in fitted position from the fit of the unscaled beam, each plotted against the scale factor applied to the reference beam. All show distinct differences between the two beams, with filtering having a negligible effect. The latter is expected as the beams are almost free from dust and have

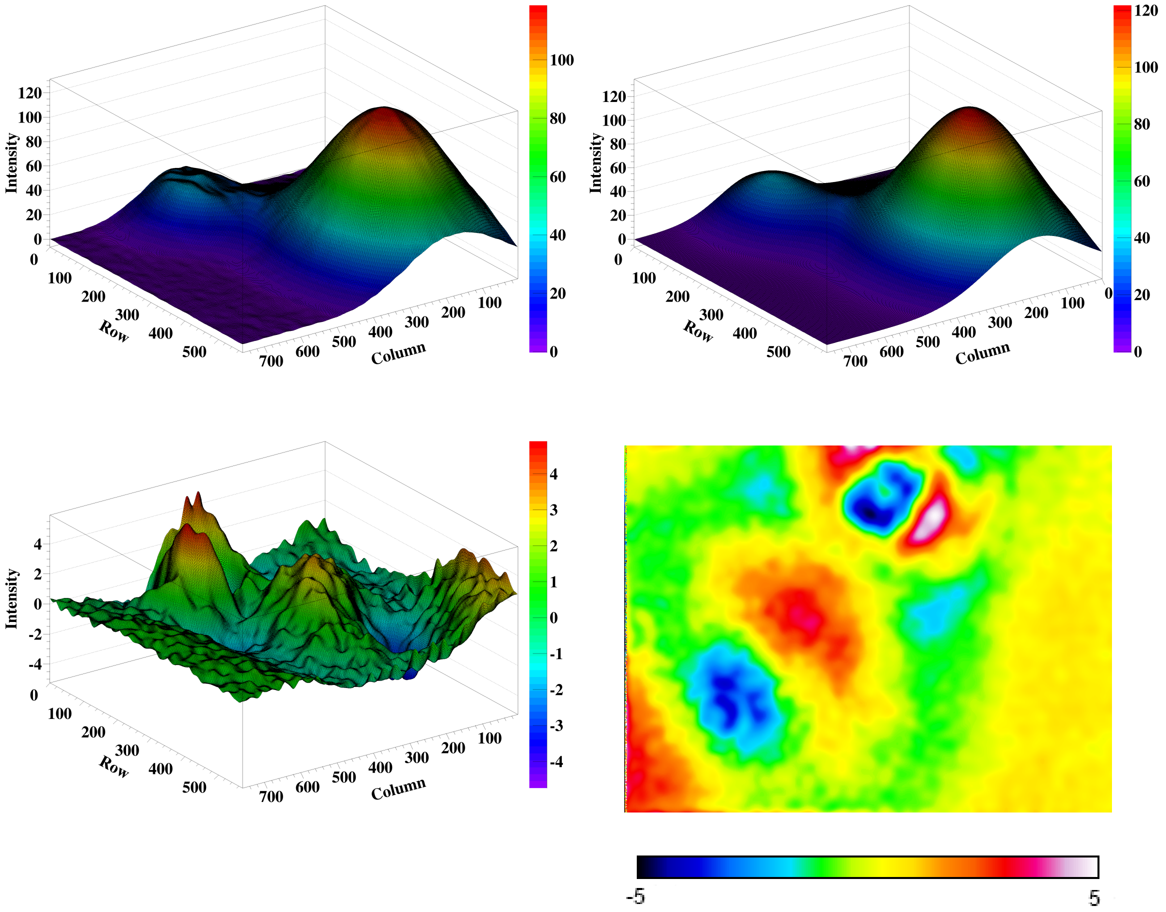
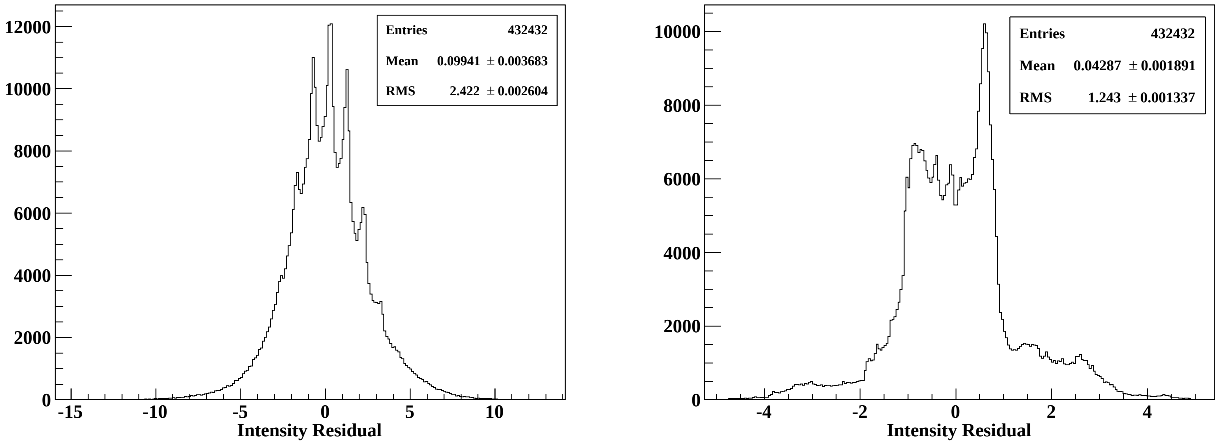


Figure 6.17: Fitting a single filtered image taken from Car 2, CCD 1. Vertical scales are in pixel grey-scale units. Top left: Original 8-bit image. Top right: Fitted beam profiles. Bottom left: Fit residuals. Bottom right: Another view of the same residuals, clipped to a narrower range. The first three columns of the image are outside the region of interest and are not part of any processing or fitting.

the noise level reduced from the effect of averaging the images. There is a minor issue with this approach; scaling the image brightness also scales the noise pattern linearly, which is not the behaviour found in section 4.7, meaning the weightings applied to the difference pixels will be slightly incorrect.

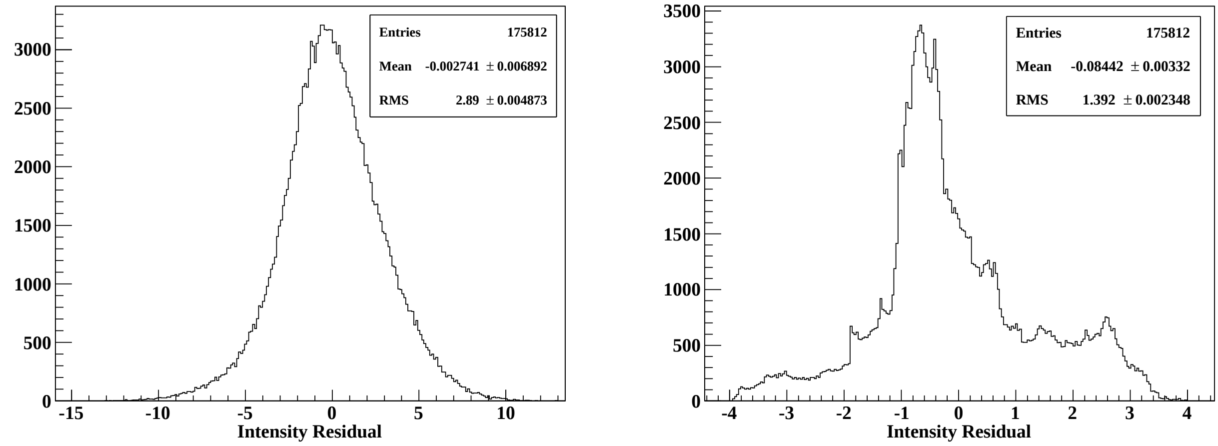
There is a small variation in fitted intensity ratio (up to 1%) over the possible working range of relative beam brightness variations (corresponding to scale factors of roughly 0.75 to 1.25) that might be encountered under normal usage (§4.6.2). Both images start to saturate at a scale factor of 1.5; this affects the fitted brightness more, particularly for the asymmetric beam due to the saturation of the beam centre. The centre is anomalously dark compared to the more Gaussian periphery; with the saturated centre ignored, only the periphery is used to fit.

The fitted beam width shows similar behaviour: less than 1% variation over the expected



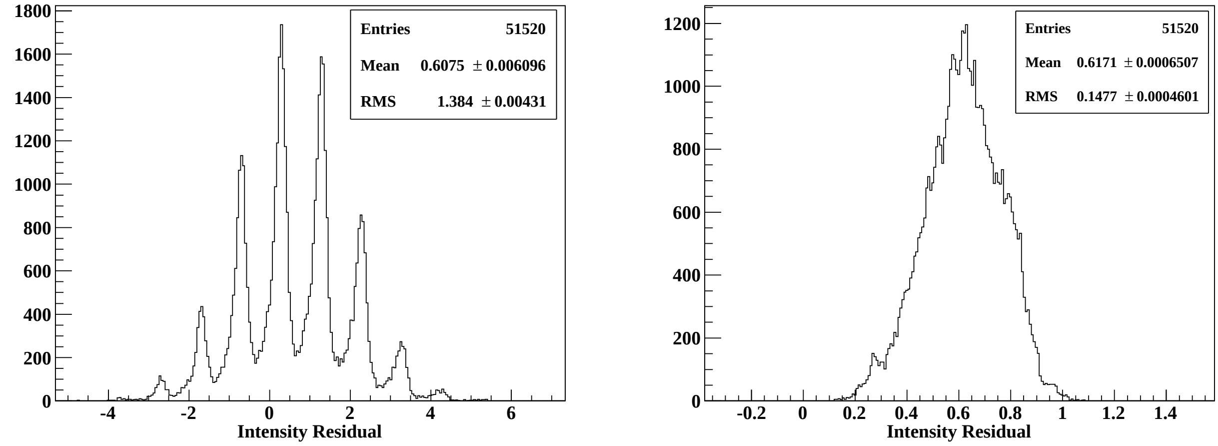
(a) Whole image, unfiltered.

(b) Whole image, filtered.



(c) Primary beam region, unfiltered.

(d) Primary beam region, filtered.



(e) Unilluminated region, unfiltered.

(f) Unilluminated region, filtered.

Figure 6.18: Intensity residual distributions for the fit of a single image from car 2, CCD 1: (a), (c) & (e) without filtering, and (b), (d) & (f) with filtering. The top row is for the whole image (except the first 3 columns which are always excluded), the second row is for a window enclosing the primary beam (4-550 in X, 250-572 in Y), and the bottom row is for a window away from either beam (600-720 in X, 250-572 in Y).

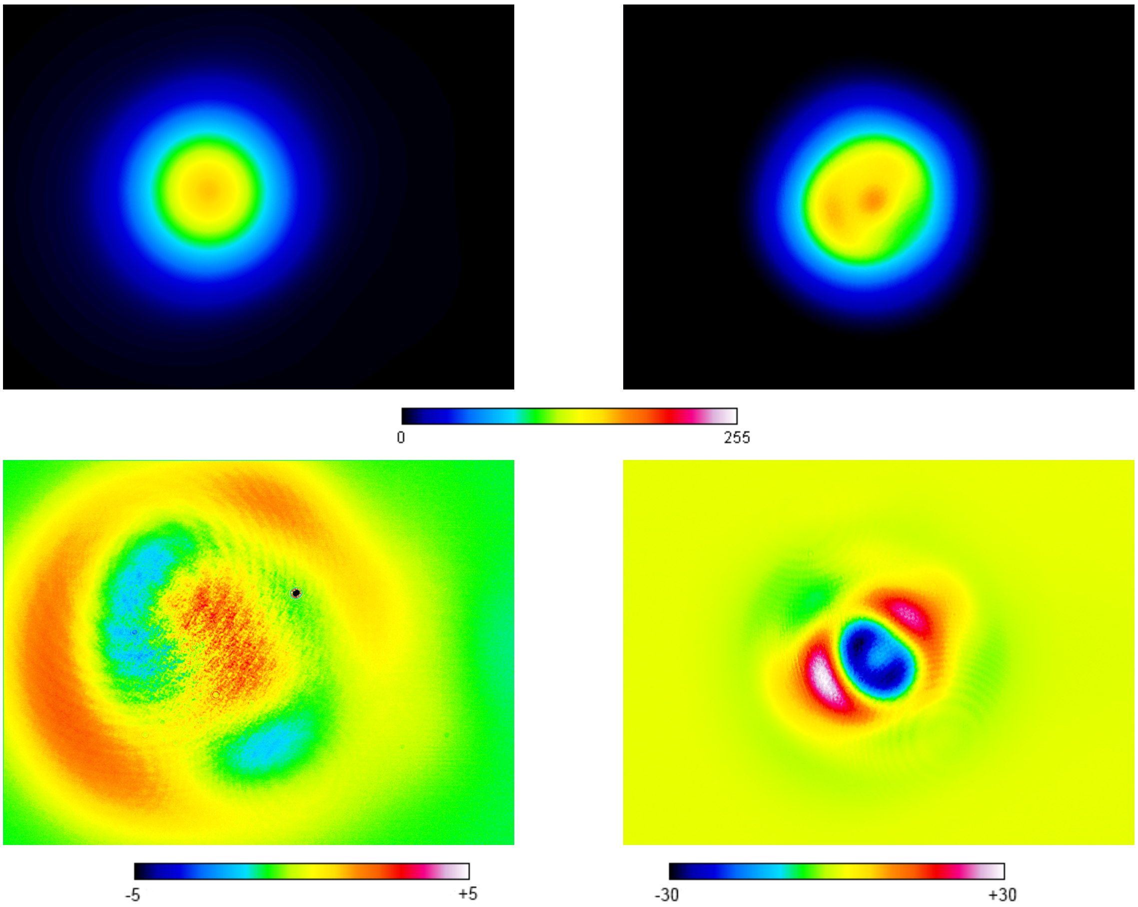


Figure 6.19: Car 1 CCD 1 (left) and car 2, CCD 1 (right) beams used for simulations of saturation and edge effects. Top row: Beam intensity profiles over a full CCD 8-bit image. Bottom row: Intensity residual values from unfiltered fits, highlighting beam anomalies in the car 2, CCD 1 in particular. (Note the different colour scales.)

range, with a change in behaviour during saturation, particularly so for the asymmetrical beam. The reduction in beam width with saturation for the asymmetric beam agrees with the central area being anomalously dim, pulling down and spreading out the fitted beam.

The variation in fitted position is the most important parameter, and exhibits a range of 0.5μm (X-axis) and 2.5μm (Y-axis) for the car 1 beam and less than 0.5μm in both axes for the car 2 beam over the 0.75 to 1.25 scale factor range. The larger variation of the Gaussian beam is surprising and not well understood, possibly caused by a combination of the template beam being off-centre and the possible effect of noise scaling. The motion of the difference for the car 2 beam agrees with the left-hand lobe (visible in the images of figure 6.19) saturating first, pulling the fitted position right, then the saturation of the right-hand lobe reduces its pull, enabling the beam to start to return left. The 2.5μm motion is not insignificant, suggesting that care should be taken to keep the laser brightness constant and also minimise the variation during rotation by selecting pellicles with a thickness corresponding to high-reflectivity (§4.6.2).

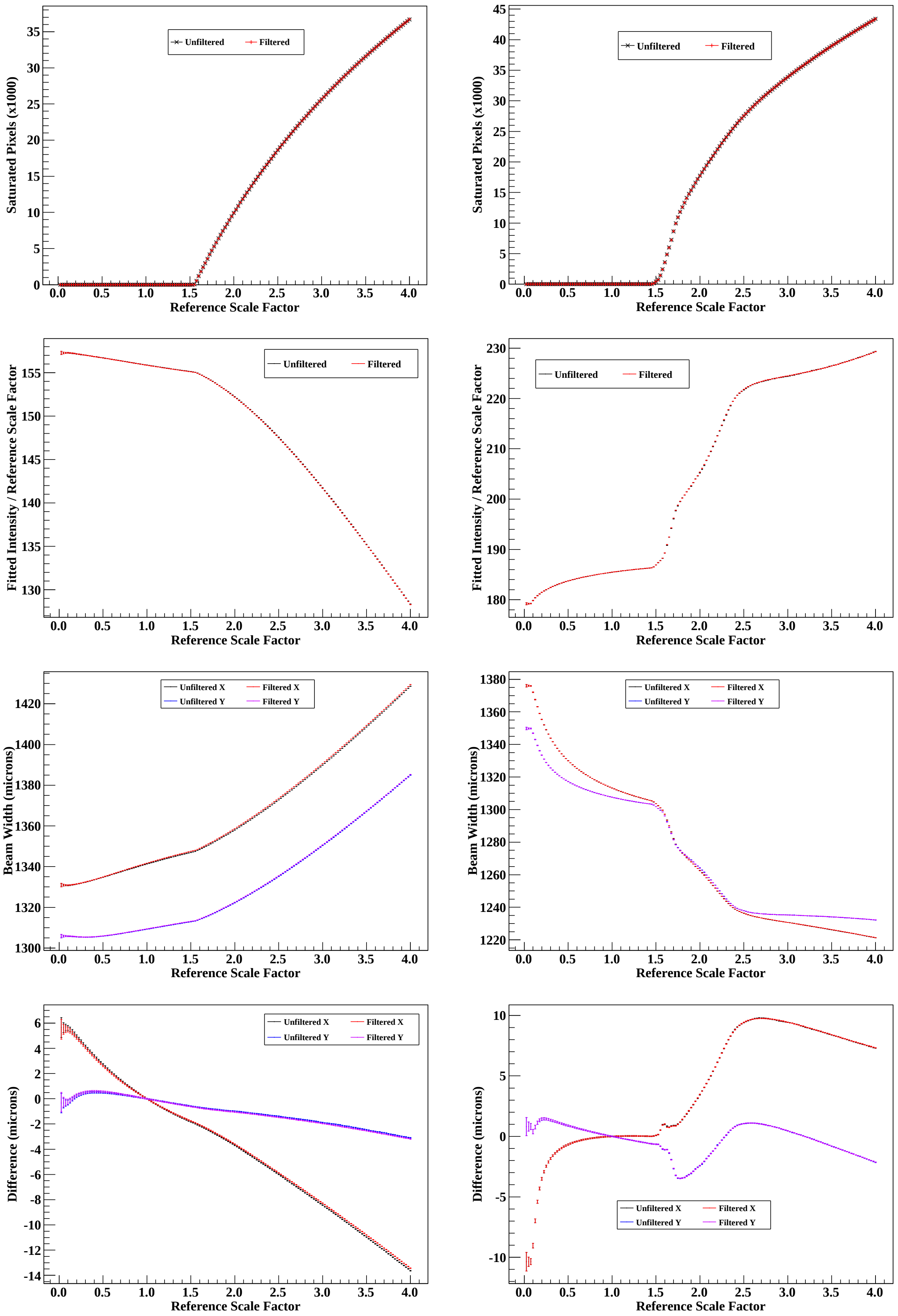


Figure 6.20: The effect of reference image brightness scaling on beam fitting for images from car 1 CCD 1 (left) and car 2, CCD 1 (right). Top row: The number of saturated pixels for unfiltered (black) and filtered (red) fitting. Second row: Fitted beam intensity / scale factor ratio for unfiltered (black) and filtered (red) fitting. Third row: Fitted beam widths along the X-axis (black: unfiltered, red: filtered) and Y-axis (blue: unfiltered, purple: filtered). Bottom row: Difference of fitted position from fit (for both X and Y) of unscaled image along the X-axis (black: unfiltered, red: filtered) and Y-axis (blue: unfiltered, purple: filtered).

6.10.2 Edge Effects

To investigate the effects of moving a real beam over the area of a CCD, the same template beams from section 6.10.1 were moved horizontally across a blank image. There is a minor problem with this approach as this means some columns are completely black. However, the effect of this should be minimal as they are naturally far from the beam.

Figure 6.21 gives the behaviour of the fitted intensity, beam width, and the difference (for both X and y) between the fitted position and the reference beam, against the reference beam position in the X-axis. (Note that the reference image has a fitted X-position of -656.8μm for car 1 and -156.5μm for car 2, which are corrected for, meaning the reference beam with an X-position of zero is in the centre of the CCD.)

While the beam is effectively entirely on-CCD, the intensity, width and difference from reference fit are all constant. The intensity and beam width only change significantly when the centre of the beam is very near the edge of the CCD but the fitted beam position is more affected: there is a shift on the order of microns when the centre of the beam is closer than 2mm to the edge. This is concerning, as the $1/e^2$ width of the beam is about 1.4mm, suggesting significant sensitivity to effects at the periphery of the beam.

6.11 Beam Profiles from all CCDs

Figure 6.22 shows the typical beam spot images with the launch in car 1 from figure 4.5a, normalised to have a common beam spot amplitude. Underneath the images are the residuals of their fits, normalised so the colour range covers ±10% of the beam amplitude. Forty of these sets of images taken over two minutes are used for each set of extrapolated points in a run. Table 6.3 shows the properties of the beam fits from a single run, consisting of forty images, including the standard deviations of the fitted parameters. Note that the even numbered cameras measure the outgoing beam, while the odd numbered cameras measure the return beam. With this configuration and particular orientation of the cars, there were four ghost beams present.

Similarly, figure 6.23 shows the typical beam spot images (seen in figure 4.5b) and fit residuals with the launch in car 3, while table 6.4 shows the beam properties. The fit for CCD 10 for this run is very poor, caused by the very low intensity, wide primary beam and the edge of a ghost failing to be detected; these fits were all marked as bad, but are shown for completeness. Note that the fits for car 1 with the launch in car 3 are well described by a Gaussian - with very minor effects from dust and interference effects from the end-cap mirror.

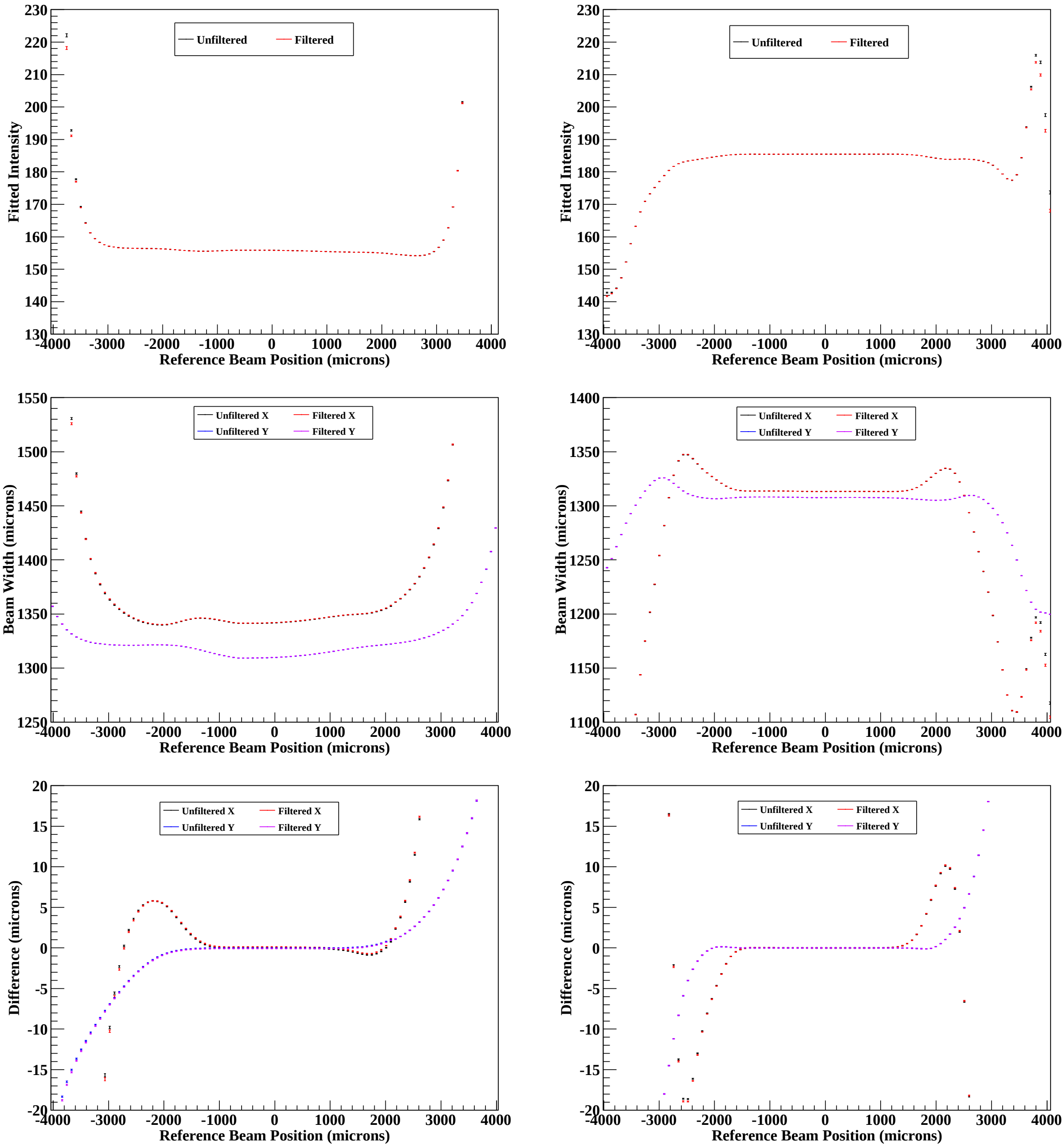


Figure 6.21: The effect of moving the reference beam centre across a CCD, on beam fits (both unfiltered and filtered) for images from car 1 CCD 1 (left) and car 2, CCD 1 (right). Top row: Fitted beam intensity. Middle row: Fitted beam width. Bottom: Difference in fit position from the fit of the reference image. The active CCD width is $\pm 3268\mu$m.

Both sets of residual images show an elliptical beam for the first two CCDs hit within 1m of the launch optics, significant beam artefacts for the images about 5m away (and also on the next CCD for the first set), before becoming much more well represented by a Gaussian on subsequent cameras; the remaining significant artefacts are in the ghost beams. The 5m artefacts appear to be similar to a TEM_{01} cylindrical mode beam brought into focus, causing a central peak and an outer ring, with the latter giving an anomalously wide fitted beam. This artefact dissipates as the beam propagates further. Figure 6.24 shows the fitted beam widths as a function of optical distance travelled from the launch, and fits the results with a Gaussian beam propagation. The agreement is poor due to the aforementioned beam anomalies on the cameras at the shorter distances.

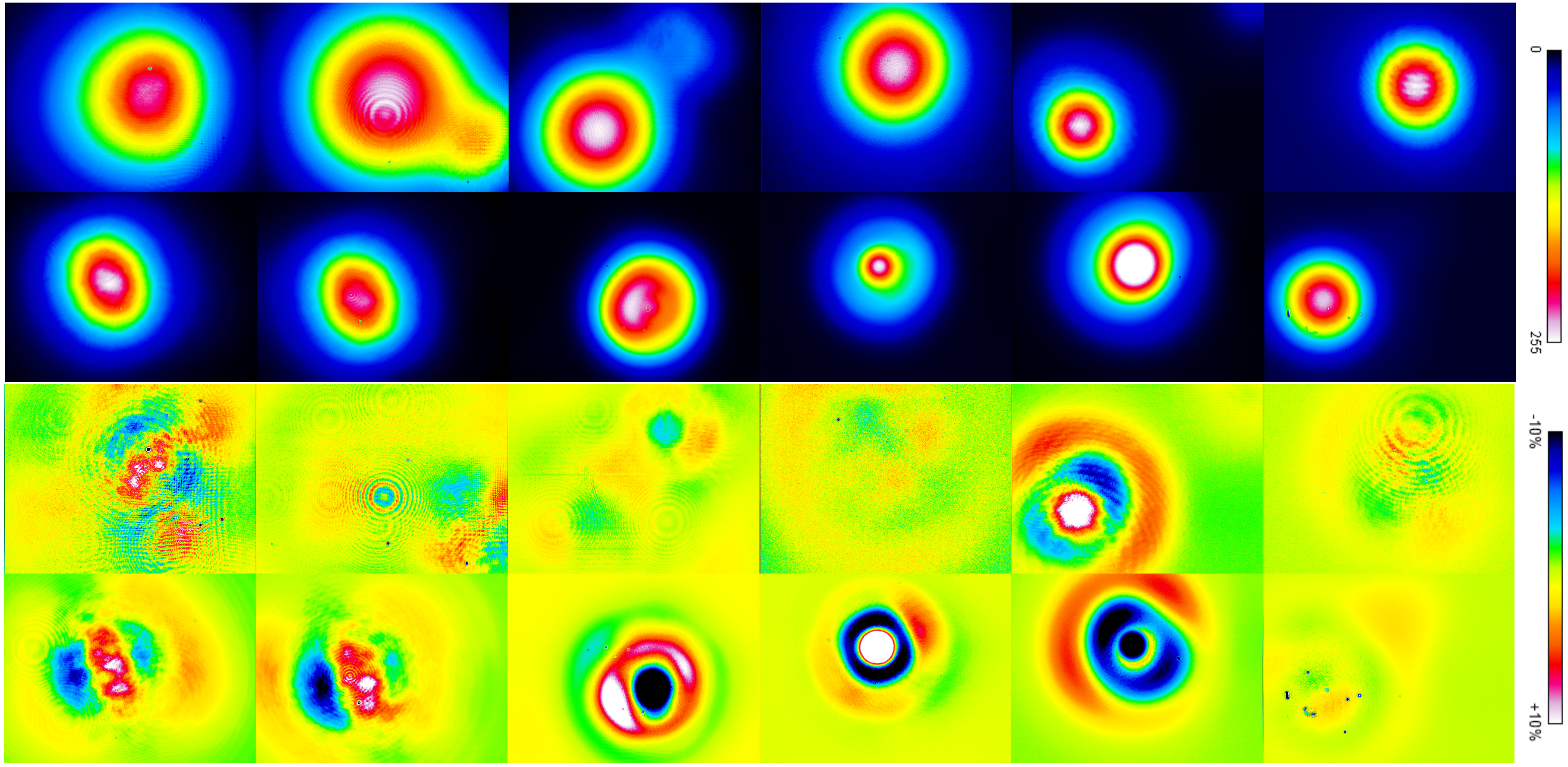


Figure 6.22: Unfiltered, normalised beam intensity profiles (top) and beam fit residuals (bottom) for all twelve RTRS CCDs from the stability run performed on 27th June 2008. For the beam profiles, the amplitudes of the images from figure 4.5a have been scaled to fill the 0-255 range (figure 6.2). The bottom images are of the residuals after the fit has been subtracted, normalised so the colour range spans ±10% of the beam peak intensity. Significant deviations from a simple Gaussian can be seen on half the images. Interference rings from dust on the pellicles are also clearly visible, as well as some specks from dust on the CCDs themselves.

Table 6.3: Typical fitted beam parameters with the launch in car 1. Taken from a set of forty images from a stability run performed on 27th June 2008. The Y-X asymmetry percentage is $100 \times (\sigma_y - \sigma_x)/\sigma_x$. Note the amplitude of CCD 8 implies saturation.

Camera		0	1	2	3	4	5	6	7	8	9	10	11
Optical Distance	m	0.234	18.952	0.435	18.750	4.734	14.52	4.935	14.250	9.234	9.952	9.435	9.750
Primary amplitude	8-bit	198	106	194	126	225	180	136	36.3	286	152	163	122
Primary amplitude SD	μm	6.8	4.2	6.6	4.3	8.3	6.0	4.4	1.7	6.3	5.8	8.0	6.0
Primary X width	μm	1674.0	3022	1697.7	2949	1593.2	2076	1592.8	1959	1467	1455.8	1424.7	1468.9
Primary X width SD	μm	2.3	25	1.9	19	2.5	4.0	1.6	9.6	13.1	2.1	1.3	3.2
Primary Y width	μm	1788.7	2875	1795.9	2935	1628.7	2052	1581.6	2003	1481	1457.7	1363.7	1482.2
Primary Y width SD	μm	0.59	21	0.72	14	3.8	11	1.7	14	9.8	1.7	2.4	3.1
Y-X asymmetry	%	6.9	-4.8	5.8	-0.48	2.2	-1.2	-0.71	2.2	0.9	0.13	-4.3	0.9
Ghost amplitude	8-bit		73		60		37.8				14.5		
Ghost amplitude SD	μm		2.6		2.7		1.4				0.44		
Ghost X width	μm		1688		1573		1748				1033		
Ghost X width SD	μm		12		46		26				26		
Ghost Y width	μm		2010		1692		1697				980		
Ghost Y width SD	μm		68		23		14				41		
Ghost Y-X asymmetry	%		19		7.6		-2.9				-4.8		

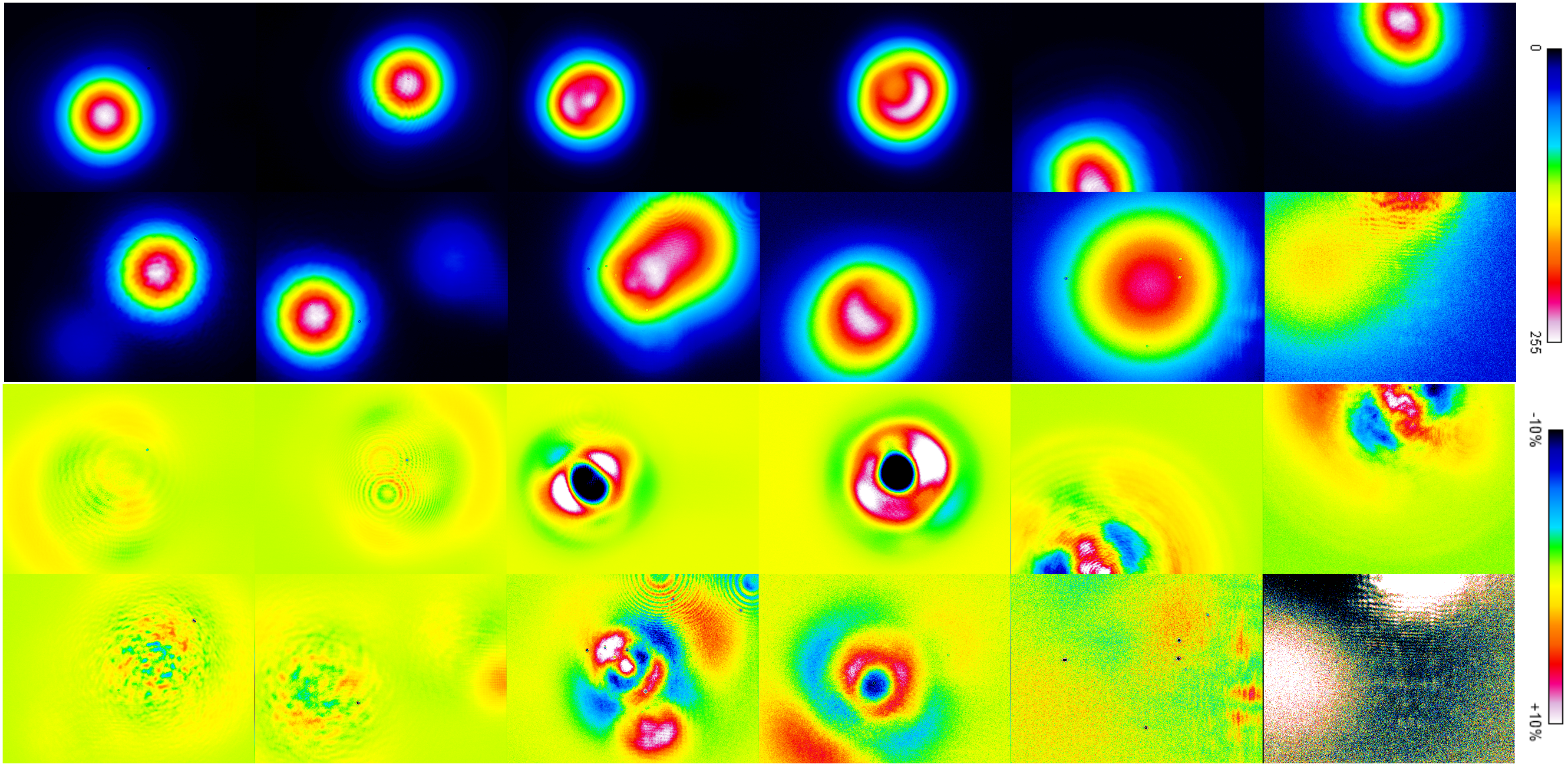
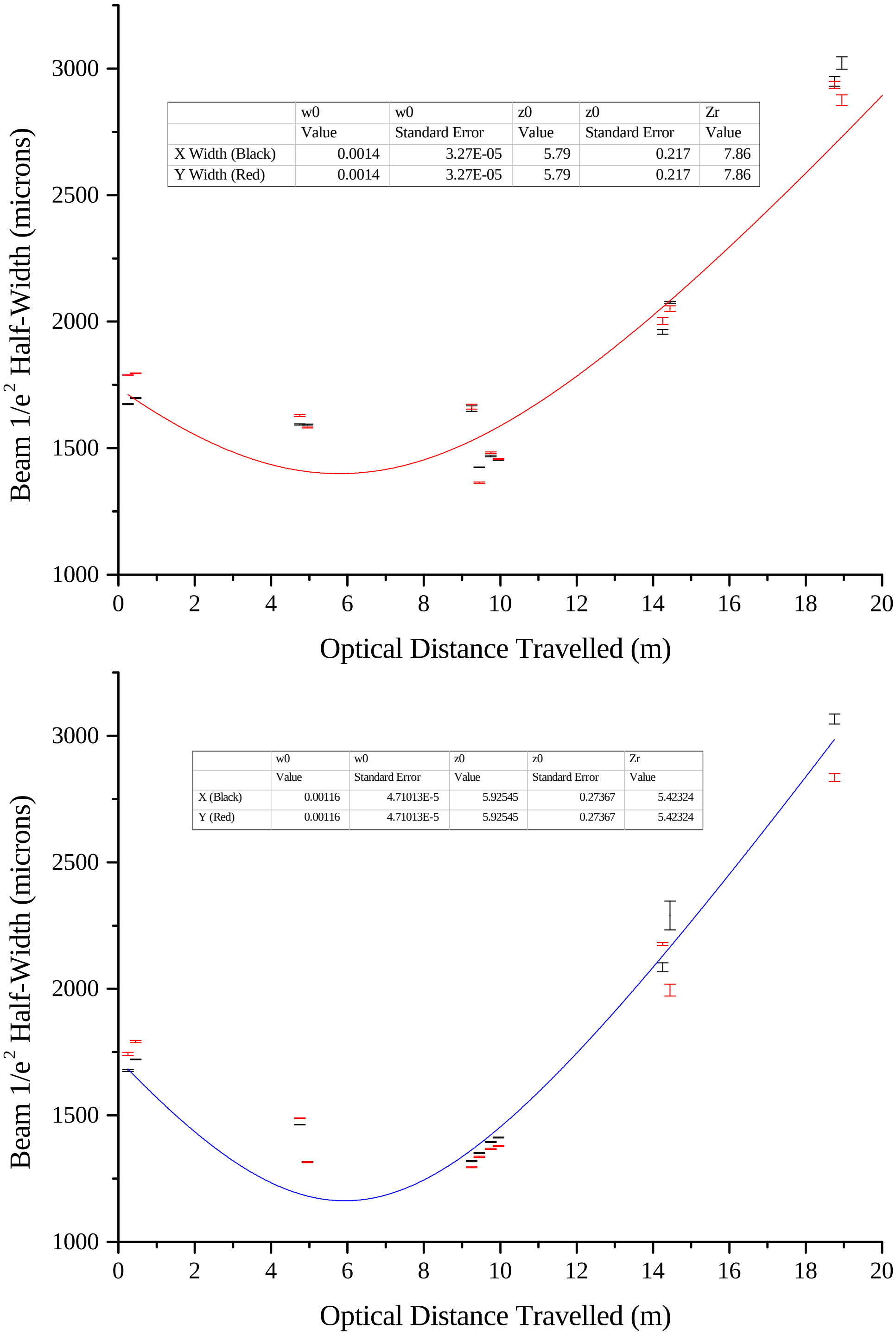


Figure 6.23: Unfiltered, normalised beam intensity profiles (top) and beam fit residuals (bottom) for all twelve RTRS CCDs from the stability run performed on 25th October 2008. For the beam profiles, the amplitudes of the images from figure 4.5a have been scaled to fill the 0-255 range (figure 6.2). The bottom images are of the residuals after the fit has been subtracted, normalised so the colour range spans ±10% of the beam peak intensity.

Table 6.4: Typical fitted beam parameters with the launch in car 3. Taken from a set of forty images from a stability run performed on 25th October 2008. The Y-X asymmetry percentage is $100 \times (\sigma_y - \sigma_x)/\sigma_x$. *Fit marked as bad.

Camera		0	1	2	3	4	5	6	7	8	9	10*	11
Optical Distance	m	9.750	9.435	9.952	9.234	14.250	4.935	14.520	4.734	18.750	0.435	18.952	0.234
Primary amplitude	8-bit	161.6	145.2	165.4	188.3	123.4	224.1	37.5	268.3	25.59	141.8	3.38*	55.14
Primary amplitude SD	μm	1.1	0.95	1.2	1.2	2.6	2.9	1.6	3.5	0.35	0.87	0.21	0.35
Primary X width	μm	1395.0	1351.5	1412.9	1319.3	2085	1315.9	2290	1463.0	3066	1721.4	4300*	1677.4
Primary X width SD	μm	1.8	2.1	2.0	1.9	18	1.4	57	1.1	20	1.7	6.7	3.8
Primary Y width	μm	1367.6	1335.9	1379.3	1294.9	2177	1315.5	1995	1489.0	2835	1792	4470*	1743
Primary Y width SD	μm	2.8	2.5	1.8	2.2	5.9	1.8	23	1.4	16	4.5	160	6.3
Y-X asymmetry	%	-2.0	-1.3	-2.4	-1.8	4.4	-0.03	-13	1.8	-7.5	4.1	3.9	3.9
Ghost amplitude	8-bit	14.20		22.69		84.5							
Ghost amplitude SD	μm	0.15		0.20		0.91							
Ghost X width	μm	1368		1505		1505							
Ghost X width SD	μm	11.8		8.3		5.4							
Ghost Y width	μm	1418		1377		1597							
Ghost Y width SD	μm	18.5		7.3		5.8							
Ghost Y-X asymmetry	%	3.6		-8.5		6.2							



	w0	w0	z0	z0	Zr
	Value	Standard Error	Value	Standard Error	Value
X Width (Black)	0.0014	3.27E-05	5.79	0.217	7.86
Y Width (Red)	0.0014	3.27E-05	5.79	0.217	7.86

	w0	w0	z0	z0	Zr
	Value	Standard Error	Value	Standard Error	Value
X (Black)	0.00116	4.71013E-5	5.92545	0.27367	5.42324
Y (Red)	0.00116	4.71013E-5	5.92545	0.27367	5.42324

Figure 6.24: Fitted beam widths as a function of optical distance travelled inside the RTRS for a run taken with the launch in car 1 (top) and a separate run with the launch in car 3 (bottom). The points with error bars are fitted beam widths from the 12 CCDs against the CCD optical propagation distance. The lines are the best fit of the Gaussian propagation equation (equation 4.2).

7 Producing a Beam Position Measurement

Taking a useful measurement for the LSM requires producing a set of data that can be used in combination with measurements made by the other subsystems. This means that the set of beam spot positions produced must be for a specific point in time and have good error estimates. To produce a beam spot position measurement for a CCD, sets of 40 images were taken and fitted. The individual fit results can then be interpolated or extrapolated to match the measurement time of the other sub-systems, with a calculated statistical error estimate. Secondly, taking multiple images allows dynamic electronic cross-talk effects to be mitigated and any motion of the system during the measurement window to be observed and corrected over short time-scales.

7.1 Electronic Cross-Talk and Timings

As noted in chapter 6, there was electronic cross-talk on the images produced by the CCDs. The effects of this cross-talk need to be understood and mitigated before measurements can be considered reliable. To enable this, the manifestation of the cross-talk was observed, as well as its effects on the results of the fits. An appropriate way of minimising the effects was found and implemented.

7.1.1 Observations of Cross-Talk

To study the cross-talk, plots were created by taking a four second set of images at 25Hz taken with no incident beam using CCD 0 on car 1. An average image of the set was created and subtracted from each individual image to be analysed. Therefore, without cross-talk, one would expect the deviation from zero of any pixel to be entirely due to the noise of the CCD, with no overall structure. However, when the brightness of the image was normalised, there was a clear time dependent structure to the noise, consisting of vertical stripes that move across the image. Histograms for the row and column totals were created for all the images and an animation produced which made the motion clear.

The column totals histograms showed up to three roughly 100 pixel thick vertical stripes, with one of the stripes having two associated lines a couple of pixels thick. The stripes moved at different rates and only occurred when other cameras in the unit are connected, demonstrating that it was electronic cross-talk in the system. The actual amplitude of the cross-talk lines is less than one intensity unit; meaning the effect on an individual pixel per line is about 0.2 units out of 255. If every pixel were one unit brighter the peaks would have an amplitude of 572; by

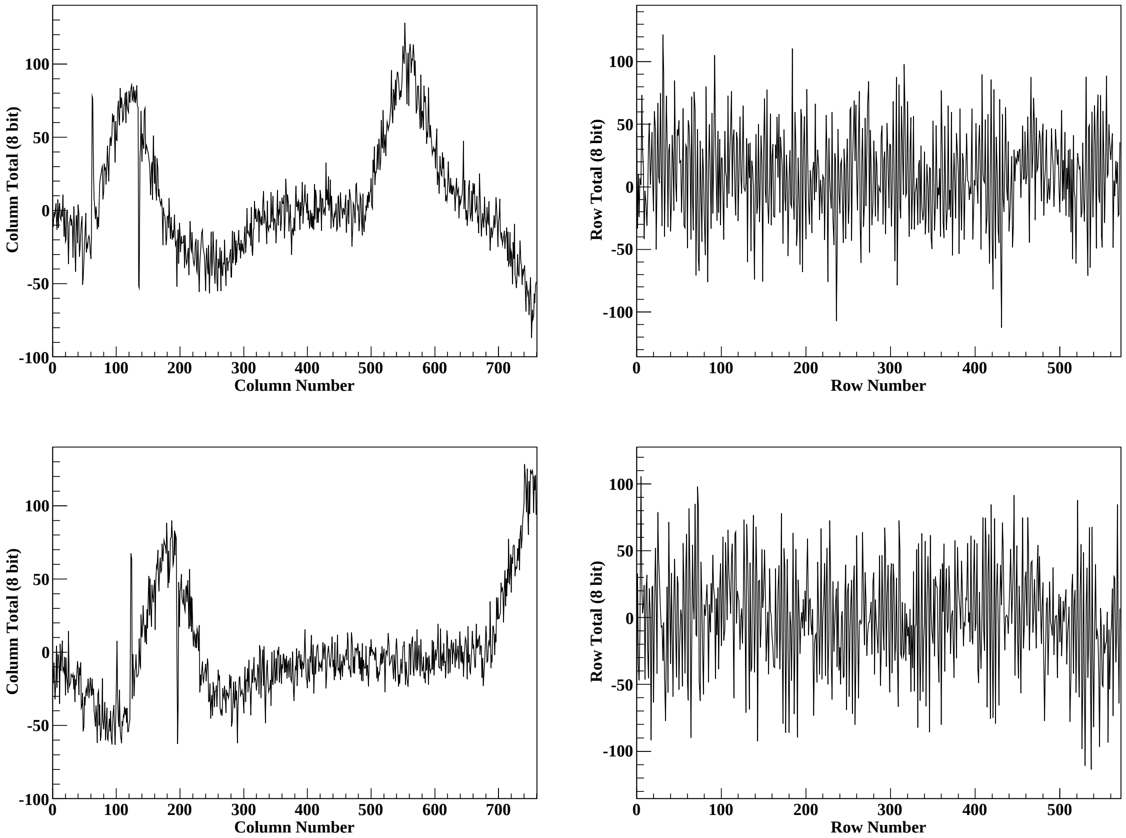


Figure 7.1: Histograms of the column intensity totals (left) and the row intensity totals (right) with no incoming beam. The upper histograms use an image taken exactly one second earlier than the lower histograms.

contrast a typical beam has a column total amplitude of around sixty thousand units. The row totals histograms show no structure with a distribution matching the noise characteristics of the camera ($\sqrt{760} \times 0.8 = 22$ units expected intensity SD, assuming a random walk for the 760 pixel width). The average value of the row totals varied from -50 to 50 depending on how many cross-talk lines are present in the image.

Some example histograms can be seen in figure 7.1. They show two bands for the column totals: the motion of the left-hand cross-talk band to the right is clear, as is the faster motion of the right-hand cross-talk band. The first cross-talk band has a complex structure, with two troughs, a central peak and two sharp spikes clearly visible. The second cross-talk band is more plain. For the row totals, there is no clear pattern discernible from the background noise. However, the average value is significantly lower in the later image as more of the first band's trough is on-CCD.

The process was repeated with an incident beam (see figure 7.2), again with CCD 0 on car 1. This camera was chosen as it is the closest to the beam launch and is mounted on the same rigid

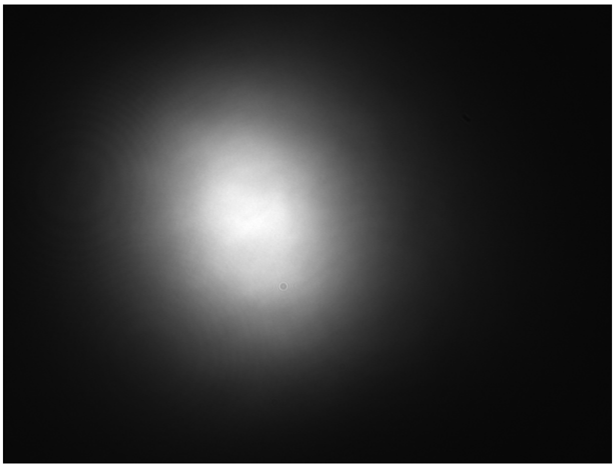


Figure 7.2: Typical image observed in the rapid acquisition cross-talk investigation.

support structure. This means that any motion of the car is irrelevant and any motion due to instability of the beam launch would be negligible. Images were taken for a period of 20 seconds, with very similar results, with the additional observation that the magnitude of the cross-talk was not dependent on pixel brightness. The cause of the cross-talk is thought to be the beat of the other connected CCDs' readout clocks interfering on the frame grabber.

7.1.2 Effect on Fitted Position

The effect of these cross-talk lines is significant. To produce a clear demonstration of the amount of uncertainty created, a baseline fit position distribution was made using as stable beam as possible with minimal cross-talk. The feed from CCD 0 on car 1 was captured with the other three cameras in the unit disconnected, with a 40 second duration sequence taken at 25 Hz, producing 1000 images. The images were fitted using the projection method described in section 6.6.1, with no filtering. The projection method is faster and more sensitive to cross-talk than the full fit method, but the period of any dynamic offset will be the same. Histograms of the fitted X and Y-positions can be found in figure 7.3. The fitting precision was excellent: analysis shows a RMS variation of 46 *nanometres* in X and 115 *nanometres* in Y and there seems to be no clear time dependent structure (figure 7.4).

A 400 second sequence of 10000 images was taken with all four cameras in the unit connected and being read out. Figure 7.5 demonstrates the degradation in performance the cross-talk causes: the X-position RMS variation increases to 1480nm and the Y-position RMS increases to 450nm. The reason becomes clear in figures 7.6 and 7.7; there was movement of the fitted position with an X-motion range of ten microns and a Y-motion range of two microns, both with a frequency of about 0.2Hz, shown in figure 7.7. Further structure is seen at slower time-scales in figure 7.6.

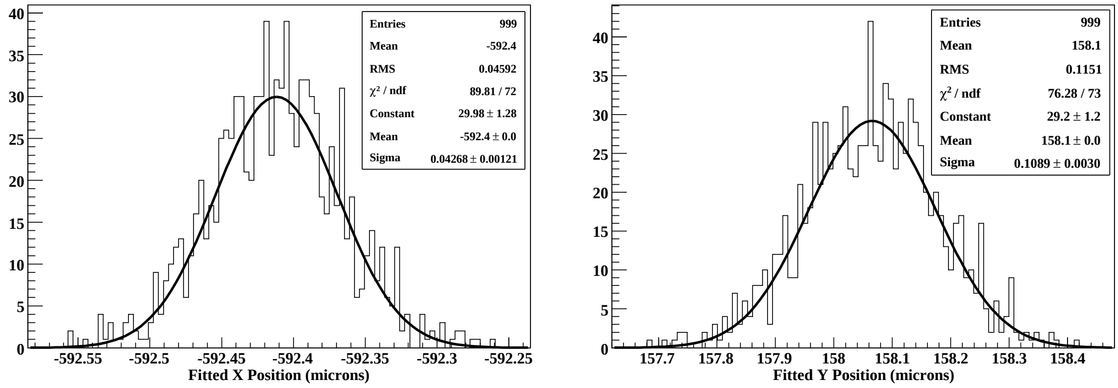


Figure 7.3: The distribution of fitted positions in X and Y from 1000 images taken at 25Hz by CCD 0, car 1 as the only connected camera. The 1σ widths of the Gaussian distributions of the fits are 43nm and 109nm respectively.

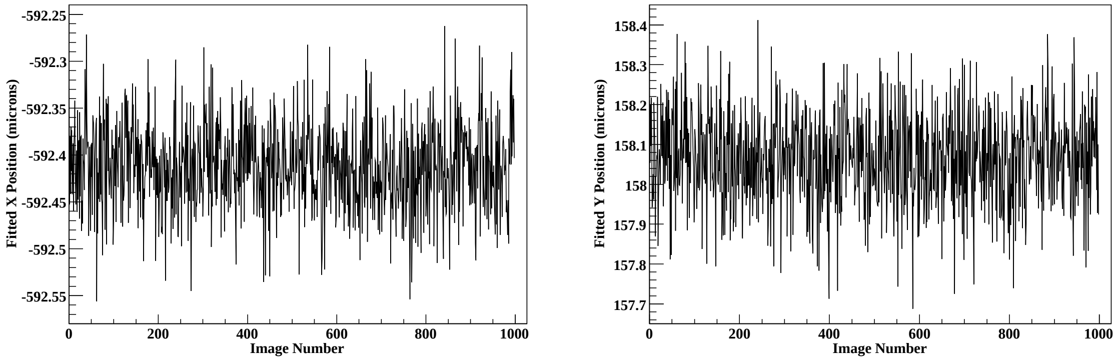


Figure 7.4: Fitted positions in X and Y from 1000 images taken at 25Hz by CCD 0, car 1 as the only connected camera as a function of image number.

The results were further analysed in frequency space. Both sets of data were re-based to be in seconds, rather than image number, filtered using a symmetric Hamming [52, section III and section V:D] window to reduce spectral leakage, before being Fourier transformed and plotted in figure 7.8 using Matlab [53]. The resultant spectra clearly demonstrate the frequencies of the cross-talk. Both the X and Y fits exhibit spikes at 0.21, 0.42 and 0.63Hz, consistent with the beating hypothesis (§7.1.1), as well as numerous lower frequency peaks. There were two more higher frequency peaks for the X-spectrum, which can be observed as the dimples in the peaks in the left hand plot in figure 7.7.

Attempts were made to fit and cancel the stripes but this could not be achieved reliably within the time available. Instead it was decided to take multiple images and use them in combination, either by producing an average image to fit, or fitting each individually and combining the results. The images have to be taken over a time frame that reduces the effect of the cross-talk as much as practically possible. The process of choosing this period is described in the following section.

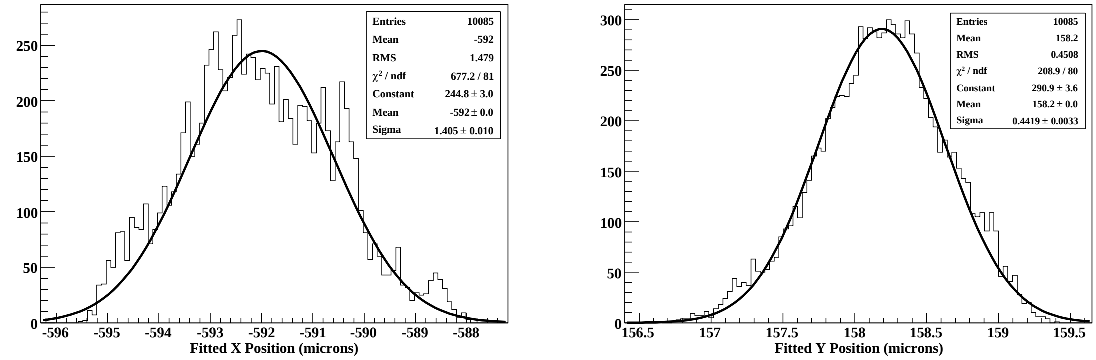


Figure 7.5: Fitted positions in X and Y for 25Hz data with interfering cameras connected. The 1σ widths of the Gaussian distributions of the fits are 1480nm and 450nm respectively.

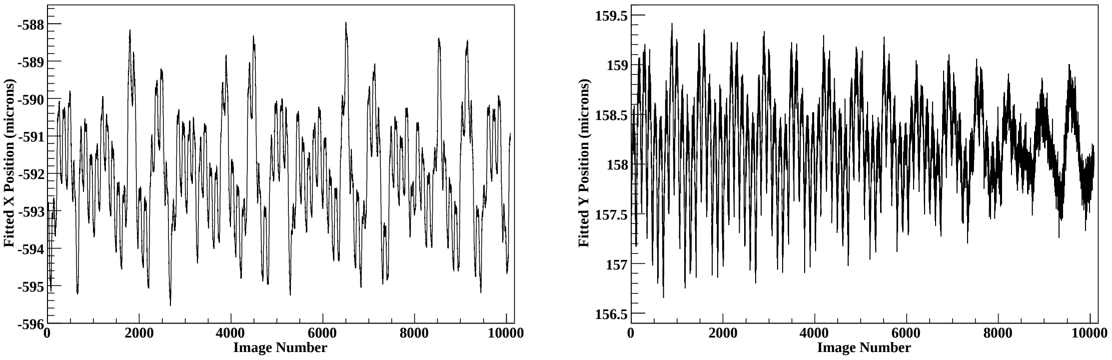


Figure 7.6: Fitted positions in X and Y for 25Hz data with interfering cameras connected, as a function of image number.

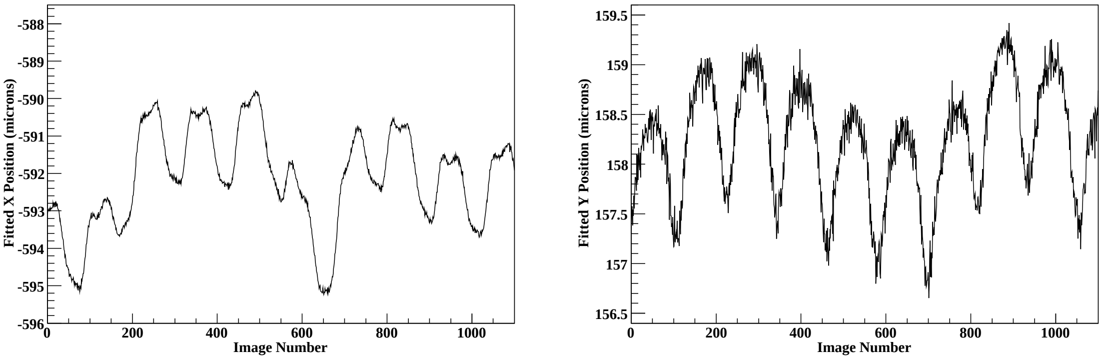
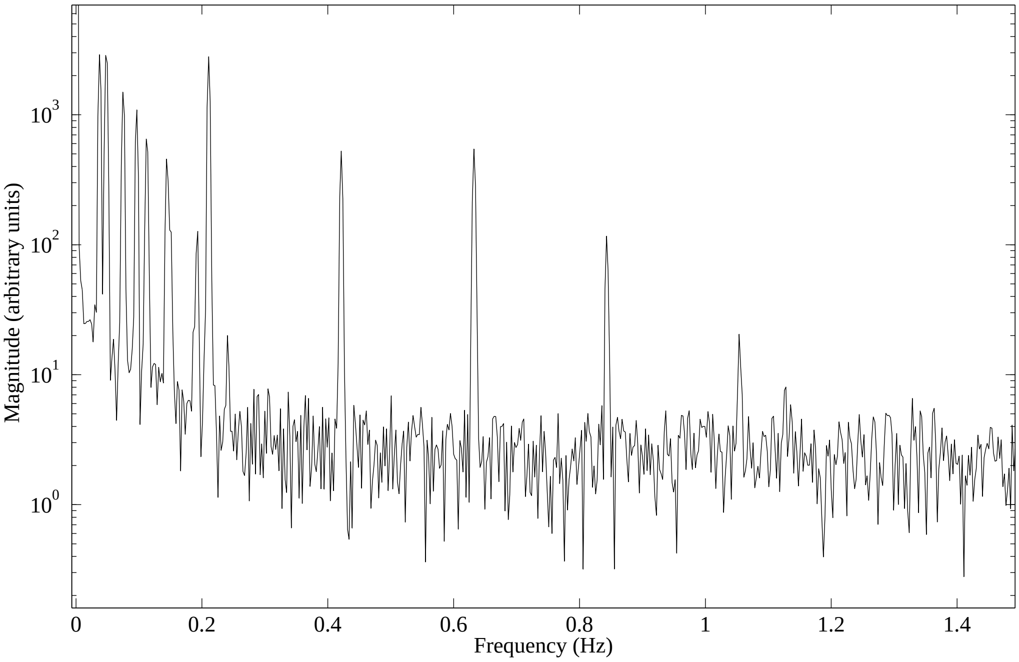
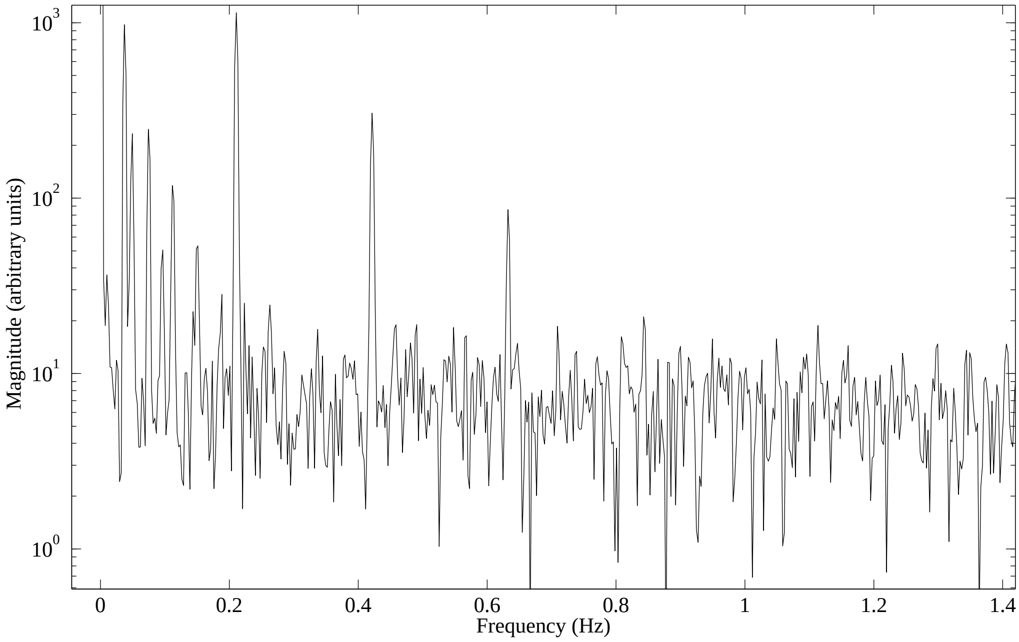


Figure 7.7: Fitted positions in X and Y for 25Hz data with interfering cameras connected, as a function of image number for the first 40 seconds.



(a) Frequency spectrum for fitted X-position



(b) Frequency spectrum for fitted Y-position

Figure 7.8: Fourier spectrum of beam 'motion' due to cross-talk for fitted X-position and fitted Y-position. There are three peaks clearly in common with the only significant differences occurring at high and low frequency. Only frequencies below 1.4Hz are shown as above this value there are no discernible peaks.

7.1.3 Finding the Optimal Period for Image Averaging

The data for the ten thousand images analysed in section 7.1.2 was analysed to determine a suitable spacing in time for the images used to create each average. A running average of both the X-positions and Y-positions was calculated by averaging over a window with a defined duration. For this investigation, the number of images in the window is kept the same (at 40); only the duration of the window is changed in steps of 40 from 40 to 9000 time spacings. The method is described further below:

1. A set of evenly spaced images filling the entire window was picked, surrounding the image of interest.
2. Their fitted X-position and Y-position values were averaged and the value assigned to the image of interest.
3. This process was repeated for all points where a full window exists.
4. The standard deviations of the X and Y-positions was calculated from the set of running averaged values .
5. This gave a single pair of results: an X standard deviation and a Y standard deviation. The process was then repeated with the whole range of different window durations. (e.g. using every 2nd image gives a duration of 80 steps, every 3rd image gives a a duration of 120 steps and so on.)
6. The two sets of standard deviations against window duration were then plotted, allowing the identification of window durations away from peaks.

The entire investigation was then repeated using 20 images in the window, though over the same time to see if aliasing effects arose. The resultant plot of the fit position standard deviation against the number of time steps in the window can be found in figure 7.9. The RMS value rapidly drops until a window size of about 600 time steps (24 seconds), before rising slightly and then falling to a minimum at 2000 steps (80 seconds). Between 2000 and 2500 steps the different number of samples have huge differences in their standard deviation due to an aliasing effect - both spike later at around a window size of 4700 steps (190 seconds).

There are a number of regions where there is a low position standard deviation for both 20 and 40 images per window; the region around 3000 steps duration is adequate and was used for all further work. This corresponds to an image every three seconds and is easy to synchronise

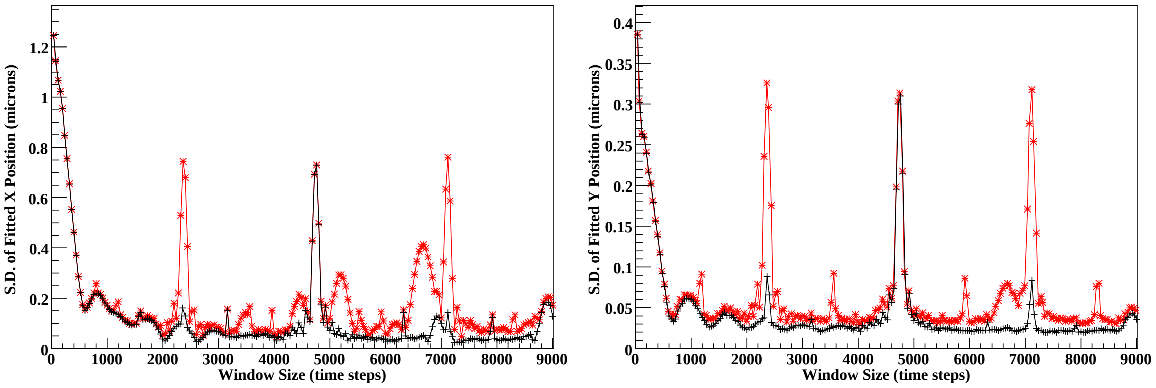


Figure 7.9: The standard deviation of the running average fitted position against window size for 20 points per average (red with stars) and 40 points per average (black with crosses) along the X-axis (left) and Y-axis (right).

across multiple cars with the coarse timing available in the DAQ. As expected, the frequency of 0.33Hz is also well away from any of the peaks in the Fourier spectra of the cross-talk 'motion' (figure 7.8).

7.2 Time Point Estimation

For the full reconstruction of the RTRS, the measurements must refer to a common moment in time. In most runs the FSI sweep, with a one second duration, was performed three seconds before the first LSM images are taken. Therefore, it is appropriate to extrapolate the LSM data to the time of the FSI sweep. However, both the units and the laser beam can move slightly during data taking with a time-scale on the order of seconds (figure 7.10). This means it is not optimal to assume a static unit, so an extrapolation to the correct time needs to be made, rather than a simple average. As the cross-talk covered in section 7.1 is on a much shorter time-scale, its effects are essentially random over the forty points taken during a run. Therefore, it should not give any particular bias meaning the extrapolation should give an accurate value, as long as the function used to model the motion is appropriate for the time-scale of the extrapolated point. To extrapolate a position to a time T = -3.0 s that can be used in reconstruction the following method is used:

- For every camera, each image in a set of 40 is fitted.
- Beam spot fits marked as invalid by the fitting algorithm are removed. (This is typically about 0.5% of images, though almost entirely confined to cameras with poor beams.)
- Outliers are found and removed by taking a simple standard deviation and marking beam

spots where the centre is more than 2.5 standard deviations away from the mean. This removes the small fraction of a percent of images where the minimisation has converged to a poor value, often many standard deviations away.

- The motion of the parameter being studied (for example the fitted X-position) is modelled by a polynomial and fitted to the positions produced in the step above using the method described in appendix B.2. The order of the polynomial is customisable, but an order of three was found to be sufficient to accurately describe the motion over the course of the data taking and the target point in time. It is assumed that the errors of the individual fits do not change in the set as the beam motion is small so a weighting matrix of unity is used.
- The chi-squared value is taken and divided by the number of degrees of freedom. The square root of this chi-squared per degree of freedom is the expected RMS error for each point.
- The point projection procedure described in appendix B.3.5 is used with the weighting matrix filled diagonally with the RMS error found in the step above (raised to the power -2). The projected time is three seconds before the first image time.
- The process is then repeated for the other parameters of interest, such as the beam intensity or width.

For all bar the first testing runs and later calibration runs, the images were taken in two sets of 20 images with the other measurements, such as FSI and tilt sensors, taken in the middle; the method described above is also suitable for estimation in this case. This has the advantage that the precision is higher when estimating a point within the batches, and the growth of the error does not depend as strongly on the order of the polynomial. Figure 7.11 shows the projected error as a function of the time for a typical measurement with one set of 40 images and another with two sets of 20 images.

The projected X-position error values were histogrammed using 99 collections of images produced from car 3, CCD 0 from a car 2 calibration run with the launch in car 1; the results of which can be seen in figure 7.12. The plot shows a typical error of 0.32μm - well within desired limits. This was repeated for all cameras, with the resultant mean error estimates plotted against optical distance plotted in figure 7.13. The abrupt increase in error after 10 metres is due to the reduction in amplitude on reflection off the mirror, any instability of the mirror itself, edge effects from passing through more apertures, the increased size of the beam, and the prominence of ghost beams on the image; all of which lower the effectiveness of the fitting.

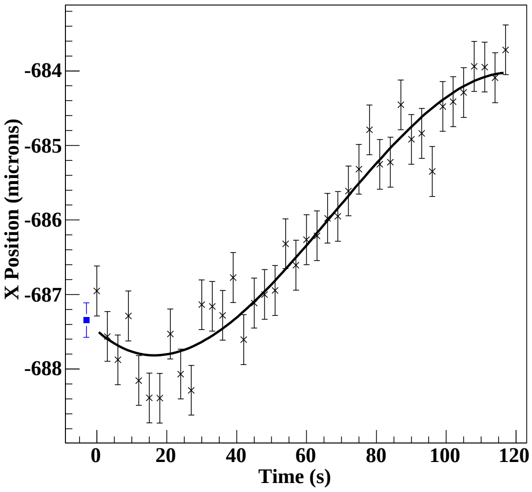


Figure 7.10: Typical beam motion during LSM data taking. The forty points (black crosses), with errors, are the fitted X-positions from images taken three seconds apart from a single camera. The line denotes the fitted third order polynomial and the projected point (blue square) at $t = -3$ is shown with its error estimate.

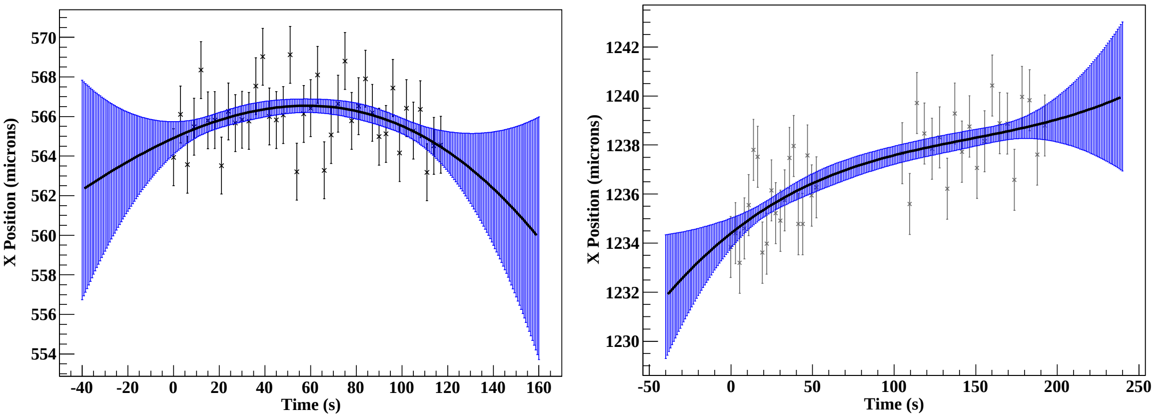


Figure 7.11: The error estimates (vertical blue lines) for projections of the X-position for car 1, CCD 1 (launch in car 3), to different points in time using a 3rd order polynomial (black line), with data taken in one batch of 40 images (left) and two batches of 20 images (right).

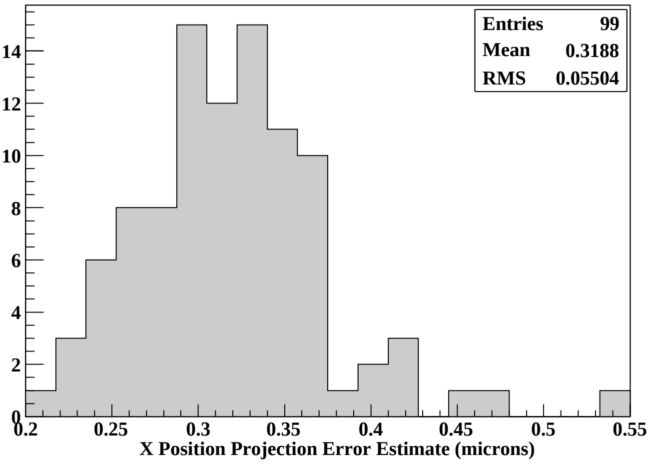


Figure 7.12: Calculated errors on the projected X-position of the beam hitting car 3, camera 8, (launch in car 1) at t = -3s for a collection of 99 sets of 40 images taken during a car 2 calibration run.

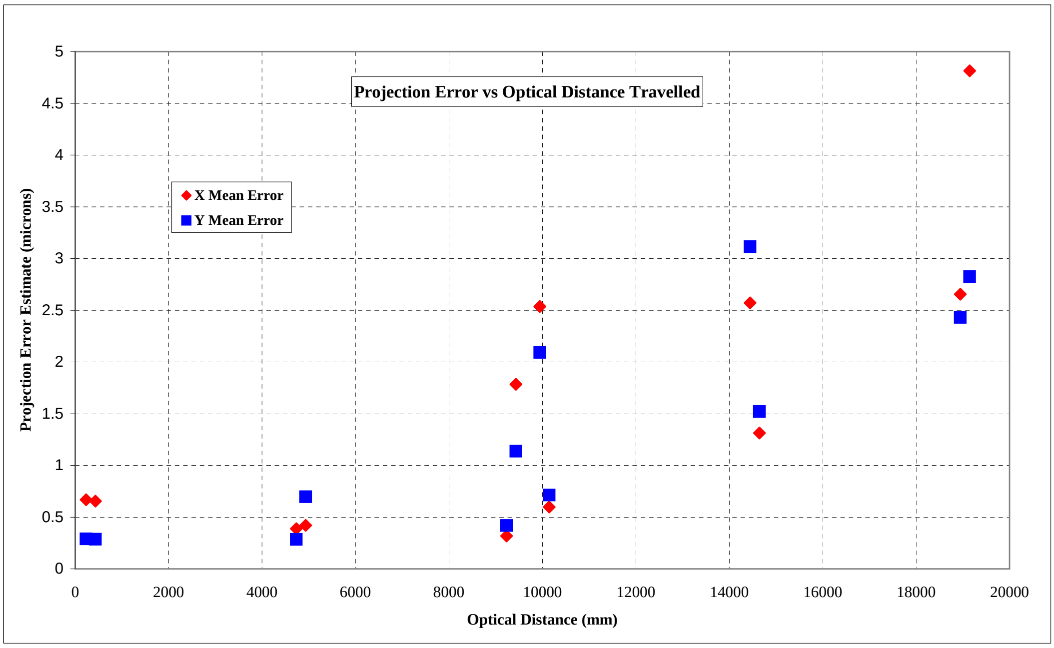


Figure 7.13: Beam spot position error estimates for the X and Y-positions of the projected points against the optical distance, with the launch in car 1.

8 Principles of Reconstruction

This chapter describes the process of turning beam spot position measurements on the CCDs into reconstructed positions and orientation co-ordinates of the units containing them. An accurate model of the propagation of light in the LSM system is described, which is then used as part of a minimisation process to give the desired co-ordinate reconstruction of a single unit.

8.1 Modelling LSM Components

A ray-tracing framework was built in C++ to trace the path of a light ray until it, and all ‘child’ rays spawned from it, have been terminated: either by an LSM component or at infinity. This framework was used to gain a qualitative understanding of the behaviour of the system via the viewer described in section 8.3.3, live reconstruction as part of the alignment process, reconstructions of large numbers of runs, and limited exploration of the full system.

A simpler version of the framework was also implemented in Mathematica [54], which is covered in section 8.4. This framework was used to quantitatively investigate the behaviour of the reconstruction minimisation, investigate symmetries, calibration, and detailed investigation of the overall performance of the system under a set of simplifying assumptions.

8.1.1 Components of the LSM

The ray-tracer reduces almost all important elements of the LSM to planes with a defined origin, axes directions and size. Objects also had other properties modelled: transmissivity and reflectivity, though only with single, constant values. It implements the following entities:

Laser	The launch of the beam from the collimator is represented by a Laser object, which spawns a light ray in a defined direction from its centre.
Ray	A Ray object is used to model the path from a component to the next one hit. It has a brightness, start point, direction and length, and keeps a list of all child rays. Beam width was not modelled.
BeamSplitter	The pellicle beam splitters are modelled by planes which spawn a transmitted and a reflected ray when hit by a beam. Default transmissivity is 0.92 and default reflectivity is 0.08.
Mirror	Similar to a BeamSplitter but has zero transmissivity.

CCD	The light sensitive face of the camera is represented by a zero transmissivity plane that records the positions of all hits on it.
Virtual CCD	A special case of CCD which allows light to pass through and inverts its x-axis readout (§8.1.4).

A corner-cube retro reflector was modelled by using 3 plane mirrors with a special case for intersections, though was not used in practice. Refraction of the beam in the pellicle beam splitters was not modelled as the offset caused was found to be negligible (§4.6.1).

8.1.2 Ray tracing

A 'ray' vector representing the beam is extended in the direction of propagation to the first interaction with an optical component. At an intersection the hit component processes the ray, noting the local position of the hit; if appropriate, it spawns another ray or rays as children of the first. These are then processed recursively, until there are no unprocessed rays above a threshold brightness. This produces a system with the lowest generation child rays typically either hitting a low-reflectivity component (CCD) or having no further intersections.

8.1.3 Standard Set-ups

The components that make up an LSM can be placed in their true physical positions and orientations, and the ray-trace performed. This gives the most complete description of the system, but is the most complex and has symmetries that complicate reconstruction later. For example, any translation of a BeamSplitter has an analogous transformation of the CCDs which would cause the same laser beam spot motion of the primary beam on the CCDs. Figure 8.1 shows a schematic of a ray-trace of the unit motions demonstrated in figure 4.1, with the calculated beam spot positions on the CCDs shown. Standard set-ups can also be used to model the beam brightness, though a more thorough treatment would be required to correctly model the variable pellicle reflectivity and transmissivity discussed in section 4.6.2. The standard set-ups are the only ones that can correctly model the ghost rays reflected of the CCD faces, though no full investigation of them was made.

8.1.4 Virtual Set-ups

It is possible to take advantage of component symmetries to produce a simplified model of the set-up [55], which we refer to as a Virtual Set-up. If the relevant CCDs are reflected in the plane of their associated beam splitter, and modified to allow light rays to continue through them after an intersection, they give almost identical beam spot positions as the standard CCDs. The only difference is the mirror operation causes a change in the handedness of the local co-ordinate system: this can be correct by applying the transformation $x \to -x$ to the reported beam spot hit position. The beam amplitude is not correctly found, but this is of little concern for reconstruction. Light rays are set to ignore the virtual CCDs when they hit them on the non-light sensitive side.

This virtual model method also has the advantage that it is easier to visualise the motion of the spots as long as consideration is given to the orientation and inversion of the reflected cameras - for example it becomes immediately clear that CCD 0 and CCD 3 should see almost the same beam spot motion, as their virtual images are almost on top of each other (figure 8.1).

Removal of the beam splitters speeds up the tracing calculations and simplifies reconstruction and calibration later, due to the removal of degeneracies and therefore symmetries in the minimisation.

8.1.5 Simplified Virtual Set-up

A further simplification is possible by allowing the back face of the cameras to be sensitive and removing the mirror entirely, though this is only valid if the return beam is anti-collinear. (The transformation $x \to -x$ CCD hit transformation is still required.) The advantage in this simplification is that it reduces number of components in the problem and hence degrees of freedom to a minimum. This was the ray-tracing method used for reconstruction and calibration in Mathematica.

8.2 Co-ordinate System

It is extremely important to have a consistent choice of co-ordinate system throughout. The 'Lab' co-ordinate axes in the tunnel are defined as:

- positive x is away from the wall being measured
- positive y is vertically up
- positive z is left looking at the wall

This produces a right handed co-ordinate system. The origin is arbitrary as everything is in the Lab frame, or a frame related to the Lab frame. This co-ordinate system definition is only approximately the same as the co-ordinate system of the LSM and is only useful as a rough guide:

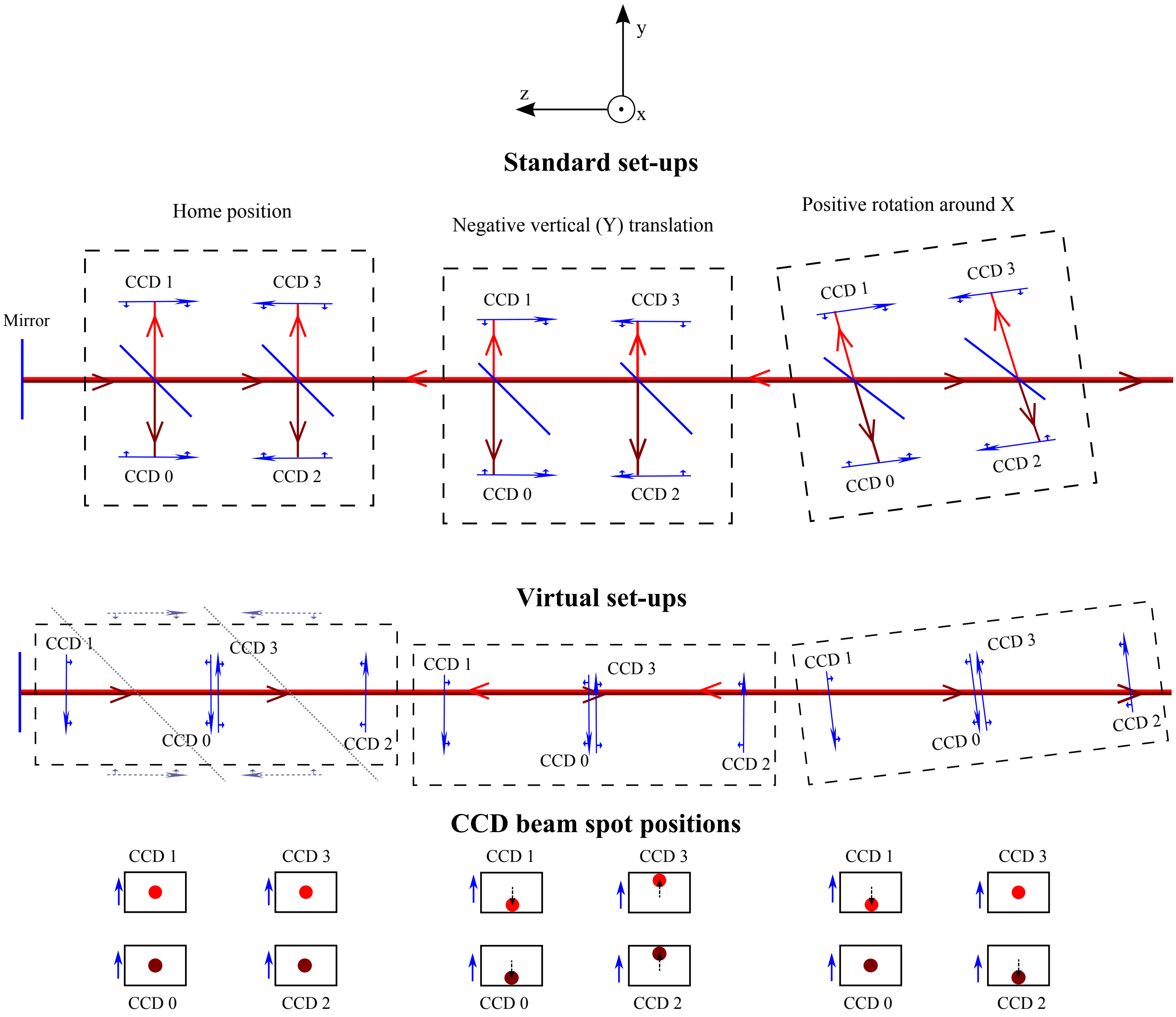


Figure 8.1: Ray trace of standard set-ups (top row) and virtual set-ups (bottom row) with a unit in its home position, translated vertically, and rotated around the x-axis. The beam launch is off screen to the right, producing an outgoing beam travelling to the left, and an anti-collinear return beam travelling to the right. The 'up' direction of the CCDs is marked with an arrowhead, while the light sensitive faces is marked with two tiny arrowheads. The location of the physical CCDs and the mirror planes are shown for the virtual set-up in the home position. The virtual set-up can be seen to be the exact analogue of the standard set-up. Beam spot displacements from the home positions are marked for both the translation and the rotation.

before calibration, there is no link between the Lab frame and the frame of the LSM components, initially defined with the launch in Car 3 according to the following rules:

- The origin is the centre of symmetry of the nominal Car 3.
- Positive z is along the direction of propagation of the outgoing beam when Car 3 is in its nominal home position.
- Positive y is the vertical direction of the nominally positioned Car 3, CCD 3.
- The x axis is formed by the cross product of y and z.

The above definition is sufficient when dealing with the LSM in isolation, though reconstructions in this frame are not externally useful as there is no link to anything physical. This link is established later in section 9.6, via the used of an attached platform and a laser tracker.

8.2.1 Rotation Matrix Convention

When parametrising the rotation of components, it is important to choose an appropriate convention for the three characteristic angles that define the rotation (appendix D.2). Two of the angles are periodic modulo 2π, while the third has a range of π. In the case of using the z-x-z(α, β, γ) Euler formalism (appendix D.2.1), typically α and γ have a range of $-\pi$ to π while β has a range of 0 to π.

A periodic function is not a problem as a 2π subtraction can always be made, but with a poor choice of rotation parametrisation convention, a very small change in orientation produces very different characteristic angles. Consider the following case using z-x-z Euler angles: the unit is almost in its reference position of no rotation, but slightly rotated around the x-axis by an angle δ_{Rx}; the Euler angles would then be $(0, \delta_{Rx}, 0)$. A very similar orientation is one where the unit is rotated by a small angle δ_{Ry} around the y-axis; the Euler angles would then be $(-\pi/2, \delta_{Ry}, \pi/2)$, very different from the earlier set. These discontinuities and numerical instabilities make fitting difficult, as fit methods typically use gradients to find the direction to a minimum. There is a final potential difficulty in that if the angle β is zero, then only the combination $\alpha - \beta$ is unique. Similarly, if β has a value of π, then only the combination $\alpha + \beta$ is unique. This symmetry, called gimbal lock, has the potential to cause numerical instabilities when fitting, though in practice this was not encountered.

The above issues were avoided by using the z-y'-x''(x-y-z) Tait-Bryan formalism [56, p154] used by G. Grzelak [57] in his studies of the LiCAS RTRS using the simple SIMULGEO [58]

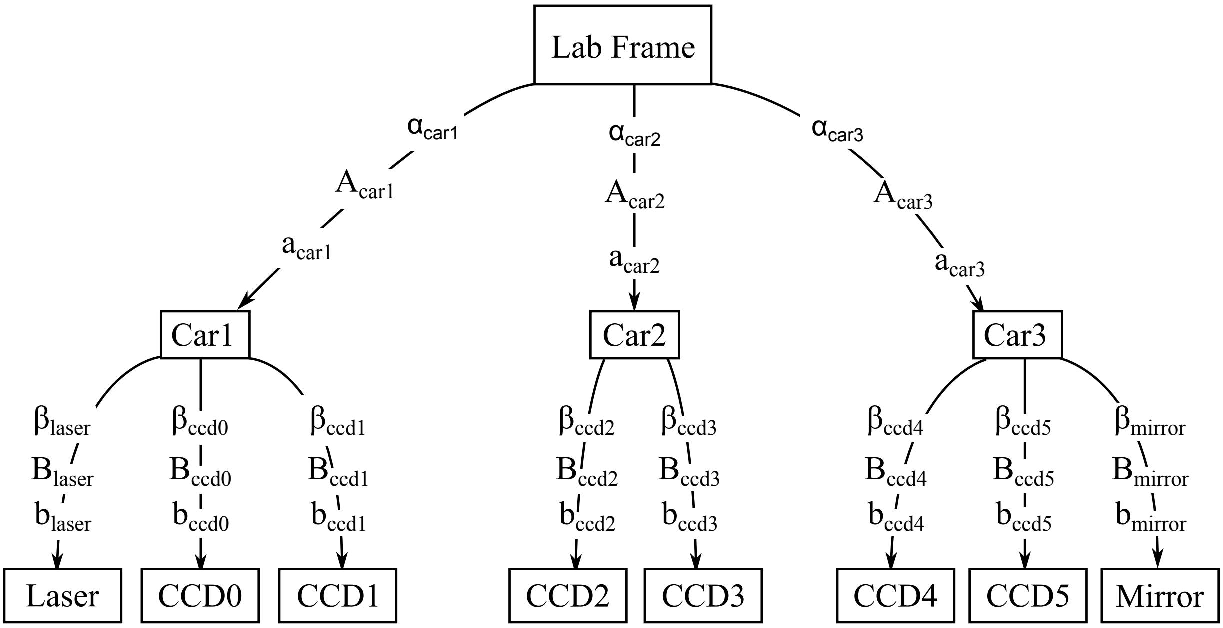


Figure 8.2: C++ Ray-tracer frame hierarchy for a 2 CCD per car train. α and β are translations, **A** and **B** are rotations, and **a** and **b** are further translations.

framework, which does not suffer from numerical instability problems for any orientations of the components modelled. Both the C++ ray-tracer and Mathematica ray-tracer are capable of taking input in multiple rotation conventions when building up a rotation matrix, but only the Tait-Bryan x-y-z formalism was used (appendix D.2.2).

8.3 The C++ Ray-Tracer

8.3.1 Frames of Reference

The C++ ray-tracer allows the construction of a hierarchy of frames of reference, with each frame having a position and orientation in its parent's frame. This enables easy movement of collections of components by moving a common parent frame. Components also have a local position in their own frame, which is also passed on to any child frames of that component. While technically redundant, this extra translation allows simpler descriptions of local positions. For example, if a unit may vibrate along its own length in its own 'Car' frame, rather than along a direction in the lab frame. Figure 8.2 demonstrates a typical hierarchy used when modelling the whole train.

8.3.2 Ray Tracing

The C++ ray tracer works by spawning a ray from the Laser object in the Lab frame, then for each component:

1. Calculates the intersection point with the ray in the Lab frame.

2. Calculates the distance to intersection from the ray start to the plane.
3. Calculates the local intersection point on the component.

The tracer then:

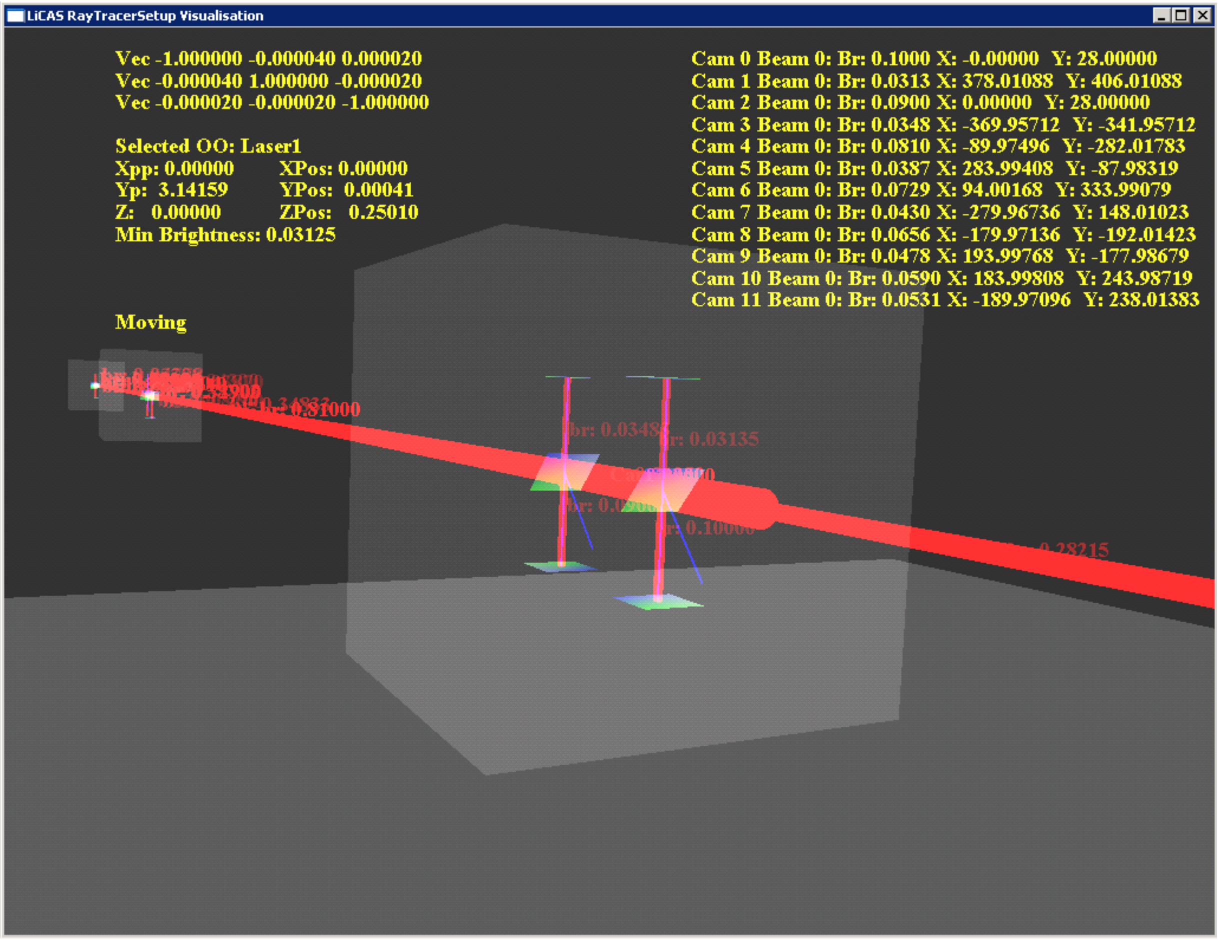
1. Discards components with a negative or zero intersection distance.
2. Discards components where this point is outside the defined dimensions of the component.
3. Discards components where the dot product of the normal with the ray is not negative. (i.e. hitting the component from behind.)
4. Picks the component with the shortest distance, sets the ray length to that value, and nominates the component to process the interaction.

The object nominated then registers the local position of the hit and spawns child rays if appropriate, their brightness set by the reflectivity of the object. (If no object is hit then the ray is terminated.) The ray-tracer then repeats the whole process on the child rays that are above the threshold brightness. The tracer keeps a list of CCD objects and outputs their local hit positions when polled. The different set-ups are loaded from, and saved to, a human-readable text file enabling significant flexibility in the use of the tracer.

This gives the ability to define a set-up, including a position and orientation for a grouping (e.g. an LSM car), and generate beam spot positions on the CCDs modelled. The changes in beam spot positions when altering the position and orientation of a unit are of particular interest. Normal usage involves keeping the internal geometry the same, while changing the position and orientation of a single car: i.e. only changing the α, $\mathbf{A}$ and $\mathbf{a}$ matrices in figure 8.2.

8.3.3 Ray-tracer Set-up Viewer

A simple OpenGL [59] viewer was written to enable the easy visualisation of LSM set-ups. This was useful to observe the effects of transformations and to check the equivalence of different set-ups. The viewer also supports modification of the components and groups of components in a frame. It is helpful to give a qualitative understanding of the behaviour of the laser beam under various operations, such as motion of a single component or a group of components. Quantitative output is also given, enabling more precise understanding. Figure 8.3 shows a typical image from the viewer.



```
Vec -1.000000 -0.000040 0.000020
Vec -0.000040 1.000000 -0.000020
Vec -0.000020 -0.000020 -1.000000

Selected OO: Laser1
Xpp: 0.00000        XPos: 0.00000
Yp:  3.14159        YPos:  0.00041
Z:   0.00000        ZPos:   0.25010
Min Brightness: 0.03125


Moving
```

```
Cam 0 Beam 0: Br: 0.1000 X: -0.00000  Y: 28.00000
Cam 1 Beam 0: Br: 0.0313 X: 378.01088  Y: 406.01088
Cam 2 Beam 0: Br: 0.0900 X: 0.00000  Y: 28.00000
Cam 3 Beam 0: Br: 0.0348 X: -369.95712  Y: -341.95712
Cam 4 Beam 0: Br: 0.0810 X: -89.97496  Y: -282.01783
Cam 5 Beam 0: Br: 0.0387 X: 283.99408  Y: -87.98319
Cam 6 Beam 0: Br: 0.0729 X: 94.00168  Y: 333.99079
Cam 7 Beam 0: Br: 0.0430 X: -279.96736  Y: 148.01023
Cam 8 Beam 0: Br: 0.0656 X: -179.97136  Y: -192.01423
Cam 9 Beam 0: Br: 0.0478 X: 193.99768  Y: -177.98679
Cam 10 Beam 0: Br: 0.0590 X: 183.99808  Y: 243.98719
Cam 11 Beam 0: Br: 0.0531 X: -189.97096  Y: 238.01383
```

Figure 8.3: The Ray-Tracer Set-up Viewer, shown inspecting a three car standard set-up with the launch in car 1. Car 1 is represented by the translucent box in the foreground, while the other cars are in the background on the left-hand side of the image. Light ray brightness is indicated by thickness, as well as being explicitly labelled. Plane origin normals are displayed in blue. The yellow text on the right-hand side describes the beam spot brightness, X-position and Y-position for all CCDs hit. The yellow text on the left-hand side shows information about the currently selected Optical Object or Frame: its normal vector in the Lab frame, and its position and rotation in its parent frame. The minimum ray brightness to trace is also shown.

Table 8.1: The motion of the beam spot on the image on-screen produced by motions of a unit.

Unit motion	CCD 0	CCD 1	CCD 2	CCD 3
+ve in X	Left	Right	Right	Left
+ve in Y	Down	Down	Up	Up
+ve in R_x''	Unchanged	Up	Up	Unchanged
+ve in R_y'	Unchanged	Right	Left	Unchanged

8.3.4 Beam Spot Motion on CCDs

The motion of the beam spots that result from the four discrete movements the LSM is sensitive to are summarised in table 8.1. The location of the beams on CCD 0 and CCD 3 barely change as the unit is rotated; this is because of the very small lever-arm distance of the virtual CCD face to the centre of rotation. As can be seen by inspecting the design drawing in appendix C, the distance from the centre of the pellicle to the centre of rotation is nominally 100.7mm while the perpendicular distance from the pellicle to the camera face is 99.5mm. This gives a lever arm of only 1.2mm, so while a rotation of 10 milliradians would give a motion of 2mm on CCD 1 and CCD 2, the motion on CCD 0 and CCD 3 would be 12μm. This made manual alignment easier, as beam spot motion on the 'inner' cameras is essentially entirely due to translations. However, it also means that CCD 0 and CCD 3 do not form a pair that could be used alone to derive the orientation of the unit.

8.3.5 Reconstruction Method

Having defined a method of producing sets of beam spots after applying different positions and orientations of a unit to an initial geometry, a fitting process was developed to do the reverse: find the unit co-ordinates that best describe a set of beam spots from a starting geometry.

The ray-tracer was extended to allow for the set-up file to designate any of the parameters of any component or frame as fit parameters, with a starting value and a rough guess of the distance to minimum, as well as the procedure for any minimisation; the framework allowed for the temporary fixing of fit parameters, as well as application and removal of parameter limits during the minimisation, controlled by the set-up file. A framework was created for loading the input beam spot positions, associated errors, the LSM model and parameters to be fitted into a form suitable for chi-squared minimisation using the Minut2 minimiser (appendix B.7). This method reconstructs one set of beam spots at a time, treating the geometry as precisely known and only propagating errors from the beam spot measurements.

8.4 Reconstruction in Mathematica

Further insight can be gained by performing the reconstruction in Mathematica using a custom-written minimiser. Fine control over the process, and Mathematica's sophisticated analysis and visualisation functions enable a more thorough understanding of the system and are necessary for the difficult calibration process described in chapter 9.

8.4.1 Minimisation

The reconstruction in Mathematica was performed for a single unit using the simplified virtual set-up: only modelling the CCDs, fixing the outgoing beam, assuming the return beam is anti-collinear, and only tracking the primary beam. This allows the removal of the mirror and CCD beam direction sensitivity from the ray-trace, reducing the calculation to the intersection of a single beam with four planes. The four planes are coupled together in a single frame, which has a single translation and rotation from the Lab frame. The equations producing the expected observations (beam spot positions) given the unit X-position, Y-position, rX-rotation and ry rotation, are capable of processing multiple runs' worth of data at once, giving the following expression for $\mathbf{L} = \mathbf{F}(\mathbf{X})$ for N[7] runs:

$$\begin{pmatrix} beam\,hit\,pos\,recon[run\,1, ccd\,1, axis\,1]\;(hr[1,1,1]) \\ hr[1,1,2] \\ hr[1,2,1] \\ hr[1,2,2] \\ hr[1,3,1] \\ hr[1,3,2] \\ hr[1,4,1] \\ hr[1,4,2] \\ hr[2,1,1] \\ \vdots \\ hr[N,4,2] \end{pmatrix} = \mathbf{F} \begin{pmatrix} unit\,location\,recon[run\,1, x]\;(r[1,1]) \\ unit\,location\,recon[run\,1, y]\;(r[1,2]) \\ unit\,rotation\,recon[run\,1, rx]\;(r[1,4]) \\ unit\,rotation\,recon[run\,1, ry]\;(r[1,5]) \\ r[2,1] \\ r[2,2] \\ r[2,4] \\ r[2,5] \\ \vdots \\ r[N,5] \end{pmatrix}, \tag{8.1}$$

where $\mathbf{L}$ is the vector of observables, $\mathbf{X}$ is the vector of unknown parameters, and $\mathbf{F}$ is the functional form of the model. Components of $\mathbf{L}$ are labelled with their run number, CCD number, and CCD local axis number (1 denoting the X-axis and 2 denoting the Y-axis). Components of $\mathbf{X}$ are labelled with their run number, and 6D axis number (1: X-position, 2: Y-position, 3: Z-position, 4: rotation around X / Rx, 5: rotation around Y / Ry, 6: rotation around Z / Rz).

---

[7] Note that all labelling in Mathematica starts from one, rather than zero.

This equation is then used in the least-squares fitting process described in appendix B, typically taking 4-5 iterations to find a minimum of the weighted objective function. In this formulation, there are no parameter degeneracies so constraints are not needed; potential degeneracies such as the z-translation reconstruction are avoided by simply not including them in the model while degeneracies such as the axis system origin have been implicitly set to fixed values rather than being parametrised. **F** is non-linear so is not easy to represent analytically, but the linearised approximation found via the process described in appendix B.5.1 can be interpreted as a matrix. This matrix describes the change in observables for small changes in the parameters, given a particular set of parameter values. It is therefore called the *design matrix*, and has the symbol **A**. Figure 8.4 demonstrates a graphical representation of the values of the design matrix at the function minimum for a three run reconstruction. Inspection of the design matrix with reference to equation 8.1 confirms:

- There are more observations than model parameters, required for successful minimisation. (The matrix is tall.)
- The matrix has a block diagonal form, meaning there is no connection between the different runs.
- The motions of the beam spots on the CCDs agree with the observations in table 8.1.
    - There is a 1:1 mapping between translations of the unit and CCD beam spot motion
    - There is roughly a 5:1 mapping between rotations of the unit in radians and the CCD beam spot motion in metres on the 'outer' virtual CCDs, due to their roughly $0.2m$ distance from the centre of rotation.
    - Rotations lead to very small beam spot motion on the 'inner' CCDs.

The absence of non-zero values outside the central blocks associated with each run mean that the problem is factorisable, meaning each run can be reconstructed individually and the result will be the same as if they had all be reconstructed together. This is expected as the geometry is fixed and there is nothing else connecting each run.

8.4.2 Covariance Matrix Analysis

The covariance matrix, $\mathbf{\Sigma}$, of the parameter vector at minimum also provides information about the behaviour of the system. As demonstrated in appendix B.2.2, $\mathbf{\Sigma} = \left(\mathbf{A}^{\mathbf{T}}\mathbf{P}\mathbf{A}\right)^{-1}$, where **P** is

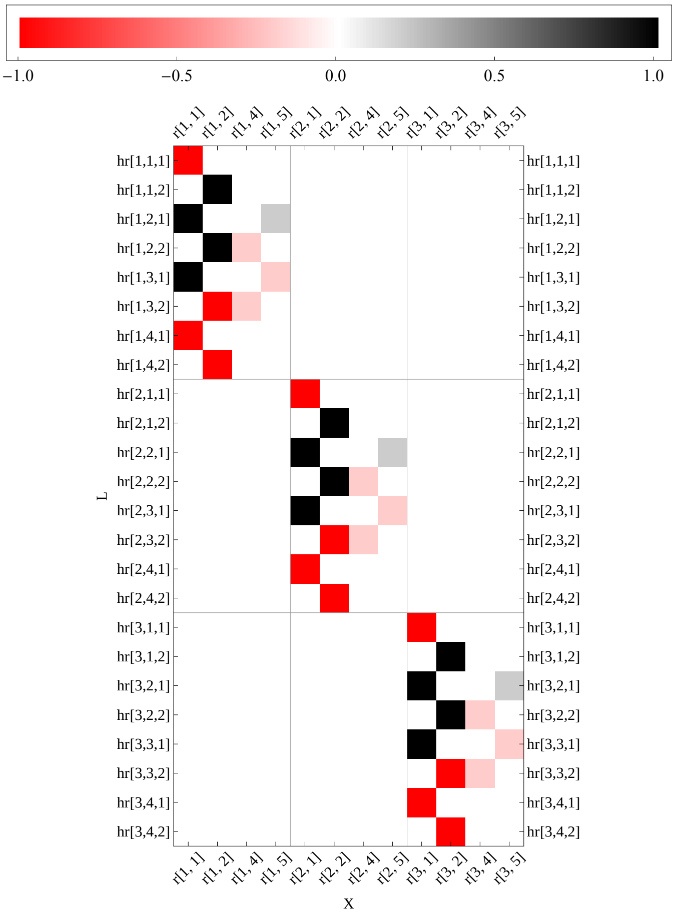


Figure 8.4: Graphical representation of the design matrix at the minimum for a three run reconstruction using only CCD data. Parameters associated with different runs are partitioned by the faint lines.

the weighting matrix of the measurements. The covariance matrix is a square symmetric matrix, with the diagonal elements, σ_{ii}, giving the variance of the reconstructed model parameters and the off-diagonal elements, σ_{ij}, giving the covariances between them. The off-diagonal elements are typically much smaller than the diagonal elements and there can be significant variation in the magnitude of the diagonals. To enable easier comparison of the matrix elements, the Pearson correlation matrix, ρ, is produced by normalising the elements of the covariance matrix by the standard deviations of the pairs of parameters contributing to the element. ($\rho_{ij} = \sigma_{ij}/\sqrt{\sigma_{ii}\sigma_{jj}}$.) This normalises all the diagonals to unity and gives the off-diagonal elements values between -1 and 1. Figure 8.5 shows the Pearson correlation matrix for the three run reconstruction. Note the differing correlations between the X-position and RY-rotation, and Y-position and RX-rotation. The correlation is due to rotations mixing in translations from the other axis according to the sine of the rotation. The correlations are variable as the effect depends on the specific position and orientation of the unit. Apart from this, there are no other clear correlations.

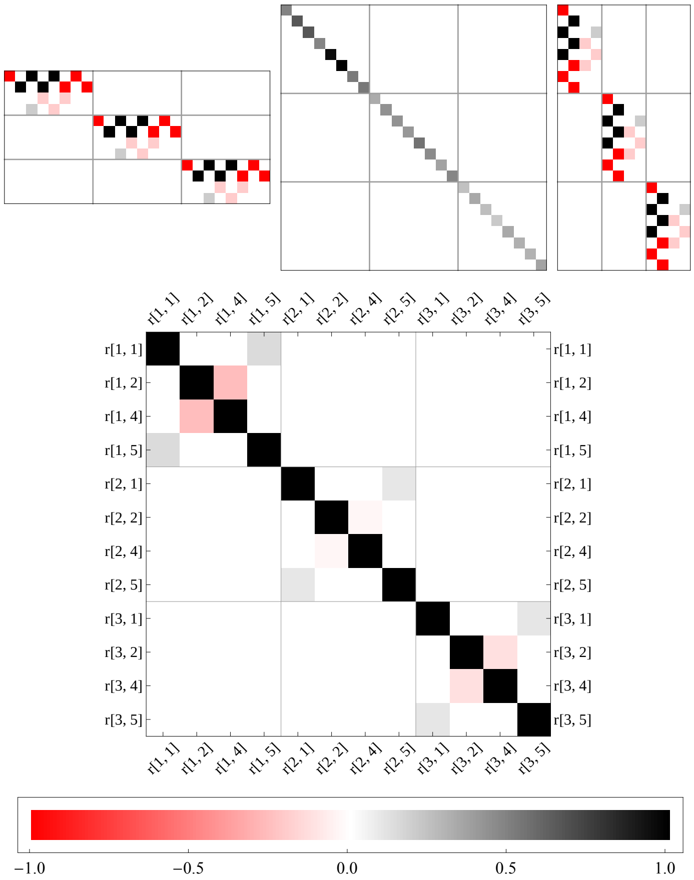


Figure 8.5: Graphical representation of $\mathbf{A^TPA}$ (top) at the minimum for a three run reconstruction using only CCD data, which after inverting and normalisation becomes the the correlation matrix (bottom). Parameters associated with different runs are partitioned by the faint lines.

8.4.3 Eigensystem analysis

Further information about the behaviour of the system at the minimum can be derived by inspecting the eigensystem of the $\mathbf{A^TPA}$ matrix. The eigenvectors of this matrix can be considered as unique basis vectors in the chi-squared space. Eigenvectors with large eigenvalues correspond to steep valleys in chi-squared space, while eigenvectors with small eigenvalues correspond to shallow valleys, and hence are directions along which motion causes little change in the chi-squared value. To aid visualisation, the eigenvectors can be imagined as collective motions of the components of the model in real space; sometimes these directions are easily understood, such as the motion of a few components, though they can be significantly more complex.

Figure 8.6 shows the eigenvalues and a representation of the eigenvectors of the system at minimum. All the eigenvalues are large and of a similar order so the system can be seen to be well defined and numerically stable. The linear-scaled eigenvector component plot shows that each eigenvector is essentially entirely along a single $\mathbf{X}$ component; all the primary components have magnitudes of more than 0.999 while the largest secondary component of any eigenvector

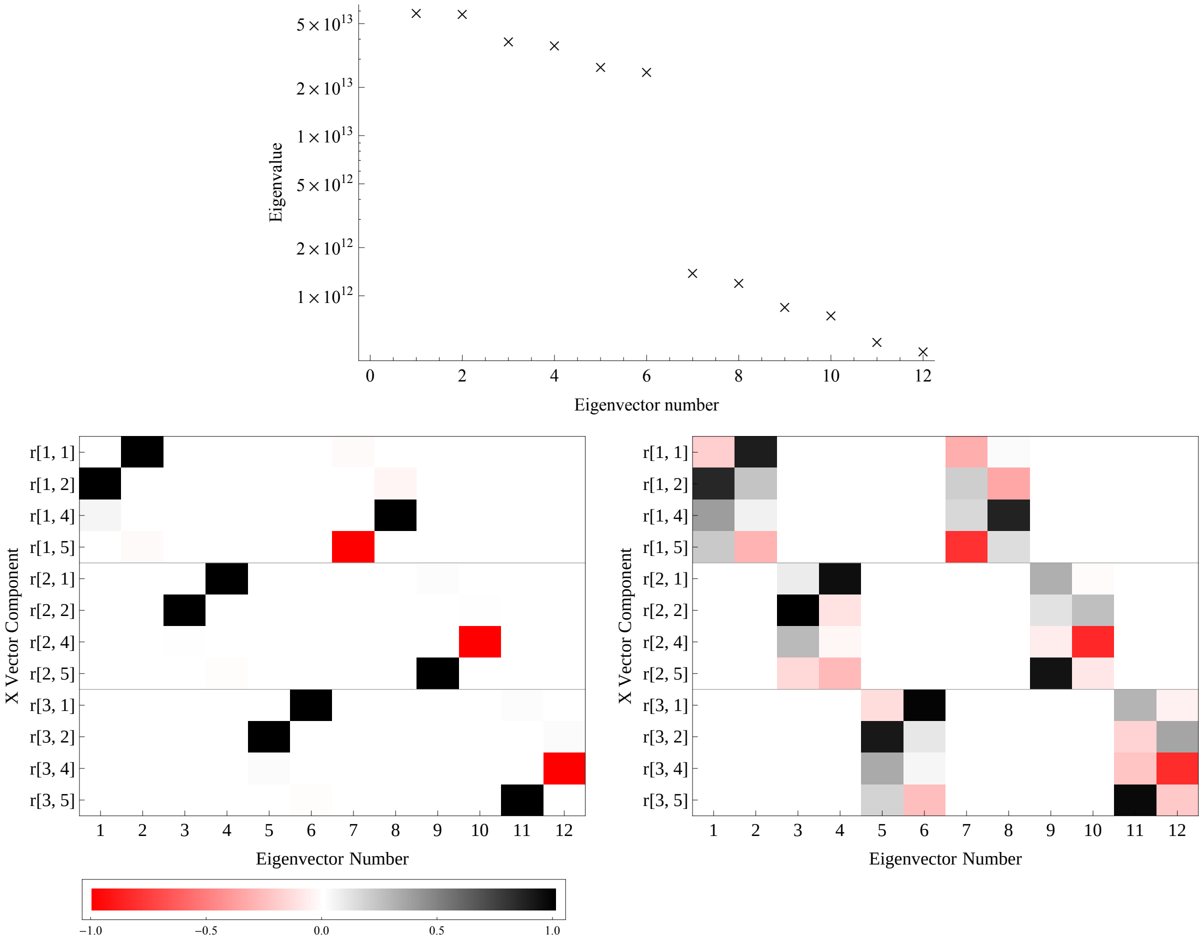


Figure 8.6: Eigenvalues (top) and associated eigenvector components with a linear colour scale (left) and a non-linear colour scale (right). Vector components are read vertically.

was 0.04. When matched with the eigenvalues plot, the eigenvectors corresponding to translations have larger eigenvalues than the eigenvectors corresponding to rotations, meaning that the system is more sensitive to rotations. As the eigenvectors correspond almost exactly to the model parameters, this can be confirmed by looking at the reconstruction parameter error estimates in section 8.5.2. The non-linear representation shows that the eigenvectors are specific to individual runs, with the largest mixing between the X-position and Ry rotation or Y-position and Rx rotation[8].

8.5 Comparison of Methods

8.5.1 Set-up

To confirm the equivalence of the three methods, reconstruction of a single unit was performed with the unit modelled by a standard set-up in C++, a virtual set-up in C++, and a simplified

---

[8] The order of eigenvectors matching the run numbers is coincidental.

Table 8.2: Ray-tracing component parameters used for reconstructing car 1 calibration data for a standard set-up with normal CCD and BeamSplitter objects (top), and a virtual set-up using TransparentCCD objects (bottom). The simplified virtual set-up used in the Mathematica reconstruction is similar to the virtual set-up shown, but with no mirror, the laser set via a source vector, and CCD rotations set to zero. Starred values are the parameters to be fitted. (Values calculated from the design drawings of appendix C.)

Component	Parent Frame	Position (m)			Orientation (rad)		
		X	Y	Z	Rx	Ry	Rz
Car1	Lab	0*	0*	0	0*	0*	0
CCD0	Car1	0	-0.0995	0.1007	$-\frac{\pi}{2}$	0	0
CCD1	Car1	0	0.0995	0.1007	$\frac{\pi}{2}$	π	0
CCD2	Car1	0	-0.0995	-0.1007	$-\frac{\pi}{2}$	π	0
CCD3	Car1	0	0.0995	-0.1007	$\frac{\pi}{2}$	0	0
BS0	Car1	0	0	0.100672	$\frac{\pi}{4}$	0	0
BS1	Car1	0	0	-0.100672	$\frac{\pi}{4}$	0	0
Mirror	Lab	0	0	-1	0	0	0
Laser	Lab	0	0	10	0	π	0

Component	Parent Frame	Position (m)			Orientation (rad)		
		X	Y	Z	Rx	Ry	Rz
Car1	Lab	0*	0*	0	0*	0*	0
TCCD0	Car1	0	2.8×10^{-5}	0.001172	π	π	0
TCCD1	Car1	0	2.8×10^{-5}	0.200172	0	π	π
TCCD2	Car1	0	-2.8×10^{-5}	-0.200172	0	0	0
TCCD3	Car1	0	-2.8×10^{-5}	-0.001172	0	π	0
Mirror	Lab	0	0	-1	0	0	0
Laser	Lab	0	0	10	0	π	0

virtual set-up in Mathematica. For continuity, the results presented are for the corrected car 1 calibration data developed in chapter 9. A summary of the component set-ups calculated from the design drawings (appendix C) can be found in table 8.2. Note that the mirror is not placed in the car 1 frame of reference and is assumed to remain static throughout.

8.5.2 Results

The three reconstructions gave identical results, up to the level of numerical precision: the mean difference between the Mathematica reconstruction and the C++ reconstructions were all of order nm/nrad or less. Similarly, the standard deviation of the differences between the values created by the different reconstruction processes were of order nm/nrad or less. The error estimates produced by the minimisations had a similar level of agreement.

Table 8.3 shows the resultant parameter ranges and means of the error estimates calculated by the reconstructions. The values are within reasonable ranges but at this point there is no easy

Table 8.3: Parameter ranges and mean error estimates for simple reconstruction using a nominal model and per-run, statistical-only error estimates from beam spot projection for a set of 53 runs taken during the calibration procedure of chapter 9.

Co-ordinate	Parameter	Min. Value	Max. Value	Mean Error Estimate
X-translation	r[i,1] (μm)	-1051.1	2656.0	0.15
Y-translation	r[i,2] (μm)	-1674.8	1044.2	0.14
Rx-rotation	r[i,4] (μrad)	-3782.0	2439.1	0.98
Rx-rotation	r[i,5] (μrad)	-3285.0	4868.0	1.08

way of checking the values; the co-ordinate system that has been used for reconstruction has no link to the real world. This is because the co-ordinate system is based on a nominal geometry and has no accurately tractable relation to tangible elements of the LSM or any other part of the RTRS or tunnel. This is addressed in chapter 9.

Additionally, the fits produce a very large chi-squared per degree-of-freedom value of 1.7 million, indicating the results are not yet valid. The high chi-squared is due to a combination of the unit not having the assumed nominal geometry, and the statistical error estimates of the beam spot fit precision found by the method of section 7.2 not being representative of the true uncertainty, which is dominated by a quasi-systematic contribution. Both these issues are resolved in the calibration process of chapter 9.

8.5.3 Parameter Error Estimates

Figure 8.7 shows the individual error estimates of the reconstructed parameters for each run, showing fairly minor variations, caused by the differing beam spot position errors for each run, and the effect of the different orientations.

The error estimates are very small but could only be achieved if the sole input errors were the statistical errors of the CCD beam spot positions, meaning they act as a lower limit of what could be achievable. In reality, they are not a true reflection of the parameter uncertainties: firstly, the error estimates for the CCD beam spot positions are later found out to be too small (see sections 9.4 and 9.5); secondly, the error generation only propagates the statistical errors and assumes the geometry is known exactly. This is not the case: the geometry is not known exactly and the uncertainties in the component placement must be included in a full error propagation. This is investigated further in chapter 10.

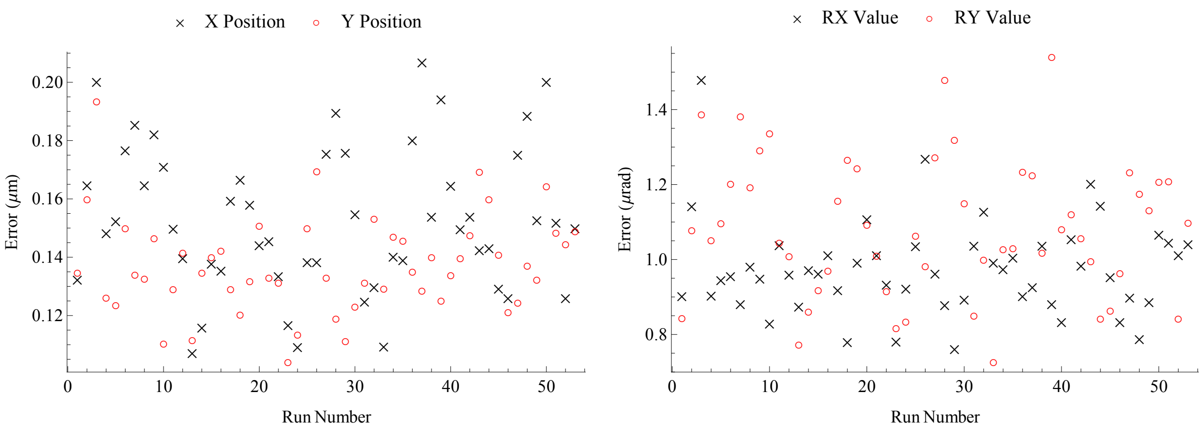


Figure 8.7: Error estimates for the translational (left) and rotational (right) fit parameters for each of the runs reconstructed, using only statistical error estimates on the beam spot positions.

8.6 Tilt Sensors

While not strictly part of the LSM, it is necessary to include the tilt sensors to measure the rotation around the z-axis of an LSM unit. The tilt sensors were included in the Mathematica model as follows:

If the tilt sensors can be assumed to lie directly along the principal axes of the system, and that the direction of gravity is anti-parallel to the y-axis, $\mathbf{e_2}$, then they can be modelled using the following process. A tilt sensor can be defined with a characteristic vector, $\mathbf{e}'$, normal to the axis of rotation sensitivity. If the zero rotation value, is defined to give a horizontal vector, $\mathbf{e}$, then the angle measured, shown in figure 8.8, can be found by simple trigonometry:

$$\sin\theta = \mathbf{e_2}.\mathbf{e}' \,, \tag{8.2}$$

which is simply the second component of $\mathbf{e}'$. For the tilt sensor measuring the Z-rotation (roll), $\mathbf{e} = \mathbf{e_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. Therefore if a general rotation $\mathbf{R}$ is applied to produce $\mathbf{e}'$, the reported angle from the tilt sensor will be

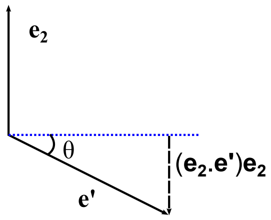


Figure 8.8: Tilt sensor measurement.

$$\sin\theta_r = \mathbf{e_2}.\mathbf{R}\mathbf{e_1} = (\mathbf{R})_{21}\,. \tag{8.3}$$

Similarly, for the X-rotation (pitch), $\mathbf{e} = \mathbf{e_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, giving

$$\sin\theta_p = \mathbf{e_2}.\mathbf{R}\mathbf{e_3} = (\mathbf{R})_{23}\ . \tag{8.4}$$

This is a convenient form as the sensor readout provides a voltage proportional to $\sin\theta$, and if the the Tait-Bryan formalism for defining the rotation matrix is used, $\mathbf{R}_{21}$ and $\mathbf{R}_{23}$ are simple (equation D.26 in appendix D.2). However, the tilt sensors are not necessarily aligned with the $\mathbf{e_1}$ and $\mathbf{e_3}$ axes, be it due to positioning tolerances or an internal offset. Each can have its own characteristic misalignment rotation, $\mathbf{R_{0p}}$ or $\mathbf{R_{0r}}$, applied before the above procedure. Note that the direction of a vector only has two degrees of freedom, so the misalignment rotation can be parametrised with two variables: for the X-rotation (pitch) sensor, this means setting the misalignment rotation around the z-axis to be zero, while for the Z-rotation (roll) sensor the misalignment around the x-axis is set to zero. This gives the values reported by the X-rotation ($\sin\theta_p$) and Z-rotation ($\sin\theta_r$) tilt sensors follows:

$$\sin\theta_p = \mathbf{e_2}.\mathbf{R}\left\{\mathbf{R_{0p}}\right\}\mathbf{e_3} = (\mathbf{R}\mathbf{R_{0p}})_{23}\,, \tag{8.5}$$

$$\sin\theta_r = \mathbf{e_2}.\mathbf{R}\left\{\mathbf{R_{0r}}\right\}\mathbf{e_1} = (\mathbf{R}\mathbf{R_{0r}})_{21}\,. \tag{8.6}$$

Similarly, if the direction of gravity doesn't correspond to the $\mathbf{e_2}$ vector, this would manifest itself as a common misalignment of the two sensors. However, there is no way to distinguish this from the misorientation of the sensor, so it is not dealt with separately. Another possible problem is an electronic zero-offset, which for the sensor is expected to be of the order $35\mu rad$ [22]. While this is not quite the same as a misalignment around the rotation axis as there is no 'leaking' of orthogonal rotations into each other, the difference is completely negligible so it will be absorbed into the misalignment rotation.

Finally, while the output voltage is proportional to the sine of the tilt angle, the constant of proportionality is not known precisely and will need to be found as part of the calibration process.

9 Calibration of an LSM Unit

This chapter introduces a system to calibrate units of the LSM using an external witness. The calibration is performed on unit 1 after overcoming various technical challenges, and the behaviour of the calibration is analysed.

The witnessing system consisted of a set of four retro-reflector marker points on a calibration platform fixed to an LSM unit, each measured by a laser tracker. The platform was modelled, enabling the position and orientation co-ordinates of the unit/platform system to be found via least-squares minimisation. The unit was moved over a set of co-ordinates that covered a good fraction of the acceptance of the unit's LSM system, with CCD and tilt sensor data from all three units recorded in addition to the marker measurements made with the laser tracker. The unit 1 CCD data was initially poor, even after cutting outliers and applying simple corrections using the data from unit 2. The results from the witnessing system were used to improve the CCD data enough to be usable in calibration. An initial full calibration minimisation was performed using data from all systems, giving preliminary calibration constants as part of its output. These preliminary constants were used to improve the model used for correction, leading to further improved CCD data and enabling the best laser tracker error paradigm to be chosen. This data was then used for a second iteration of the calibration minimisation, giving the final results. The properties of this final calibration were explored to confirm validity and robustness.

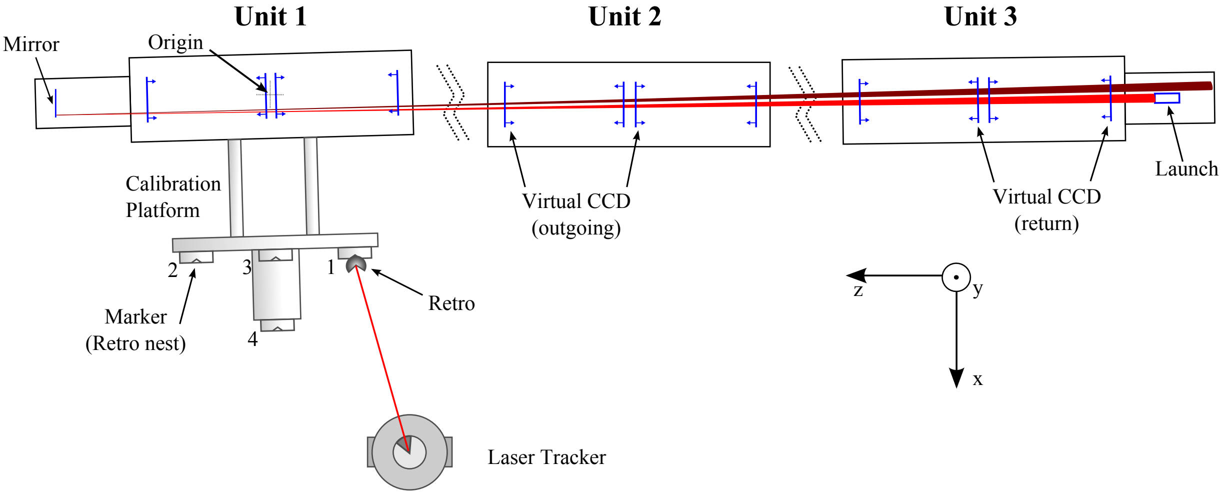


Figure 9.1: Plan view schematic for car 1 calibration, showing the important LSM components in blue, a misaligned beam in red with a darker return beam, and the calibration platform with measurement of the retro in the marker 1 position in progress. (Not to scale).

9.1 Overview

The procedure to calibrate a unit effectively is complex and many steps are needed. In principle, the best method to calibrate the LiCAS RTRS is to model all of its subsystems exercising all the units, devise a calibration routine consisting of one or many procedures to cover all of them, and combine everything into a universal, simultaneous minimisation. However, this is beyond the scope of this investigation; here, only the calibration of a single unit's LSM and tilt-sensor subsystems are demonstrated.

Figure 9.1 gives a schematic of the unit 1 calibration, showing the main elements of the set-up. Unit 1 was chosen for this investigation as the beam spots are well described by a Gaussian so fit well, allowing effort to be concentrated on finding a sound calibration process. The calibration process is most difficult for the end unit because it is more affected by beam jitter and settling motion from the pull of the vacuum, and because the return beam needs to be realigned for each orientation due the end-cap being attached to the unit.

1. To ensure effective calibration, CCD and tilt sensor data needs to be taken over a reasonable fraction of the acceptance of the LSM. There are a number of issues to be overcome:
    (a) The beam was not stable, but had a jitter enough to give 100 micron motion on the unit 1 CCDs, caused by motion of the beam launch.
    (b) The measurement of the platform markers (see step 2) takes a time that is comparable to the time-scale of the beam jitter; long enough to preclude simple averaging and too long to accurately model the motion.
    (c) A complete model of the LSM system would account for the beam jitter as all LSM components from the other cars would be included. However, to keep the model relatively simple, only car 1 was modelled and the data put into this model was corrected in advance.
    (d) The data can be corrected by observing the beam spot motion on the cameras of unit 2, and applying the relative difference with a geometric scale factor. This relies on the assumption that unit 2 is a stationary witness, which is sound.
    (e) Outliers are then removed before projecting data from multiple images to give a single set of beam spot positions per run. This removal means it is slightly superior to apply the above correction after projection than before.

2. An external witnessing system is needed to monitor the motion of the LSM unit as it moves through the set of positions and orientations.
    (a) The witnessing system consists of a laser tracker observing a retro-reflector as it is moved between a set of four marker positions. These marker positions are four 'nests' mounted on a calibration platform baseplate which is rigidly attached to the side of the unit.
    (b) The expected observations of the laser tracker were modelled and a co-ordinate system chosen, with the arbitrary origin set to a position that allowed later results to be easily interpreted.
    (c) The model was used in a minimisation to produce a set of calibration platform baseplate reconstructed co-ordinates, as well as a set of laser tracker constants and marker constants.
    (d) The input errors from the laser tracker were not well known. Three possible error paradigms were tried, each producing slightly different outputs.
    (e) Using the minimisation output, marker locations can be calculated and compared with data transformed into the same frame to check for consistency.
    (f) For each error paradigm the observables calculated by the models were compared with the original data, with the RMS residuals checked against the input errors. It was not clear at this stage which paradigm to use so steps 3-5 were repeated for each.
3. The CCDs and the tilt sensors of the unit were modelled. Expected beam spot positions were generated from the results of the platform reconstruction by augmenting the platform model with the CCD and tilt sensor geometry. This links the cameras with the platform in the frame defined by the laser tracker, though the laser launch position and propagation direction are not not initially known well.
    (a) At first, the laser beam was modelled as travelling along the Z-axis, and the CCDs were located in their nominal positions. The co-ordinate system definition and origin chosen for the reconstruction in step 1 were such that each axis is through the centre of the LSM unit and in a physically meaningful direction.
    (b) The calculated beam spot positions on the CCDs were then compared with the corrected data by looking at residuals.

   (c) A strong correlation of the X-residual and the X-position was removed by applying a global scale factor to beam spot X-hit-positions.
   (d) The necessity of the beam jitter correction was confirmed by calculating the residuals with uncorrected data. To look for improvements to the geometrically calculated scale factor, a variable post-projection version of the jitter correction was trialled but no clear improvement was found.
   (e) Four relatively inconsistent runs were identified and removed from the analysis.
   (f) Using the correlation of residuals with Z-axis motion, a 1.25 milliradian deviation of the beam direction from the Z-axis was included into the geometry of the LSM model.

4. With the model adjusted to allow for the particulars of the calibration set-up an initial ‘pre-calibration’ step was performed.
   (a) This was a full calibration using the data with the corrections previously described and the model with the laser beam angle adjustment.
   (b) The input CCD errors were set to blanket average values of 15 microns, taken from the RMS residuals of step 3.
   (c) Step 4.(b) gave a set of calibration constants describing the laser tracker, baseplate, markers, CCDs and tilt sensors as well as the run-by-run co-ordinates of the unit.

5. Applying the calibration constants found by the pre-calibration to the model, and repeating the platform-only reconstruction of step 3, better corrections to the data were found.
   (a) The pre-calibrated model was now accurate enough to allow the post-projection jitter correction scan to give a meaningful result;
   (b) This result was used to adjust the centres of rotation of the pre-projection beam jitter correction to axis-decoupled empirical values.
   (c) This adjustment enabled the regeneration of better corrected CCD data and the selection of the best laser tracker error paradigm.
   (d) Another bad run was identified and removed.
   (e) A final, better CCD RMS residual set was found using the best corrected, expurgated CCD data, and the individual values used for future error estimates.

6. A full calibration was performed with the best corrected, expurgated data, model and error estimates.
    (a) This gave the final calibration results.
    (b) An additional full calibration was performed using the statistically derived per-run CCD errors, as the error estimates serve as a guide to the expected achievable performance. However, the parameter values themselves are unreliable.
7. The properties of the calibration system are examined in detail.
    (a) The results of the per-run calibration are briefly examined: model parameters and their errors appear reasonable given the input. Observables generated by the minimisation, when compared with the input data, give residuals that do not match the input errors well: the reason is the over-optimistic error estimates for the CCD beam spot fit positions.
    (b) The calibration using empirical, quasi-systematic (§9.5) CCD beam spot fit position error estimates is examined. The calibration constants and reconstructed values, as well as the error estimates associated with those values were inspected. The results are consistent and there are no clear biases.
    (c) The covariance matrix for the model parameters was analysed and the relative errors between typical parameters shown. No concerns arose from inspecting the matrix.
    (d) A method of analysing the eigensystem of the covariance matrix of the minimised model is presented. A sample of structural weaknesses were identified and their physical meaning explored. A simple method of examining the structural contribution to the error on a model parameter is presented with an example.
8. Calibration robustness was investigated by altering the number of runs included in the calibration.
    (a) The error estimates of the constants changed in line with expectations, while there was relatively little change in the error estimates of the reconstructed co-ordinates.
    (b) The calibration constants and reconstructed co-ordinates changed by a non-negligible amount, though still small enough to maintain confidence in the approach.

A series of flowcharts summarising the steps required to calibrate the system can be found in figures 9.2, 9.3 and 9.4.

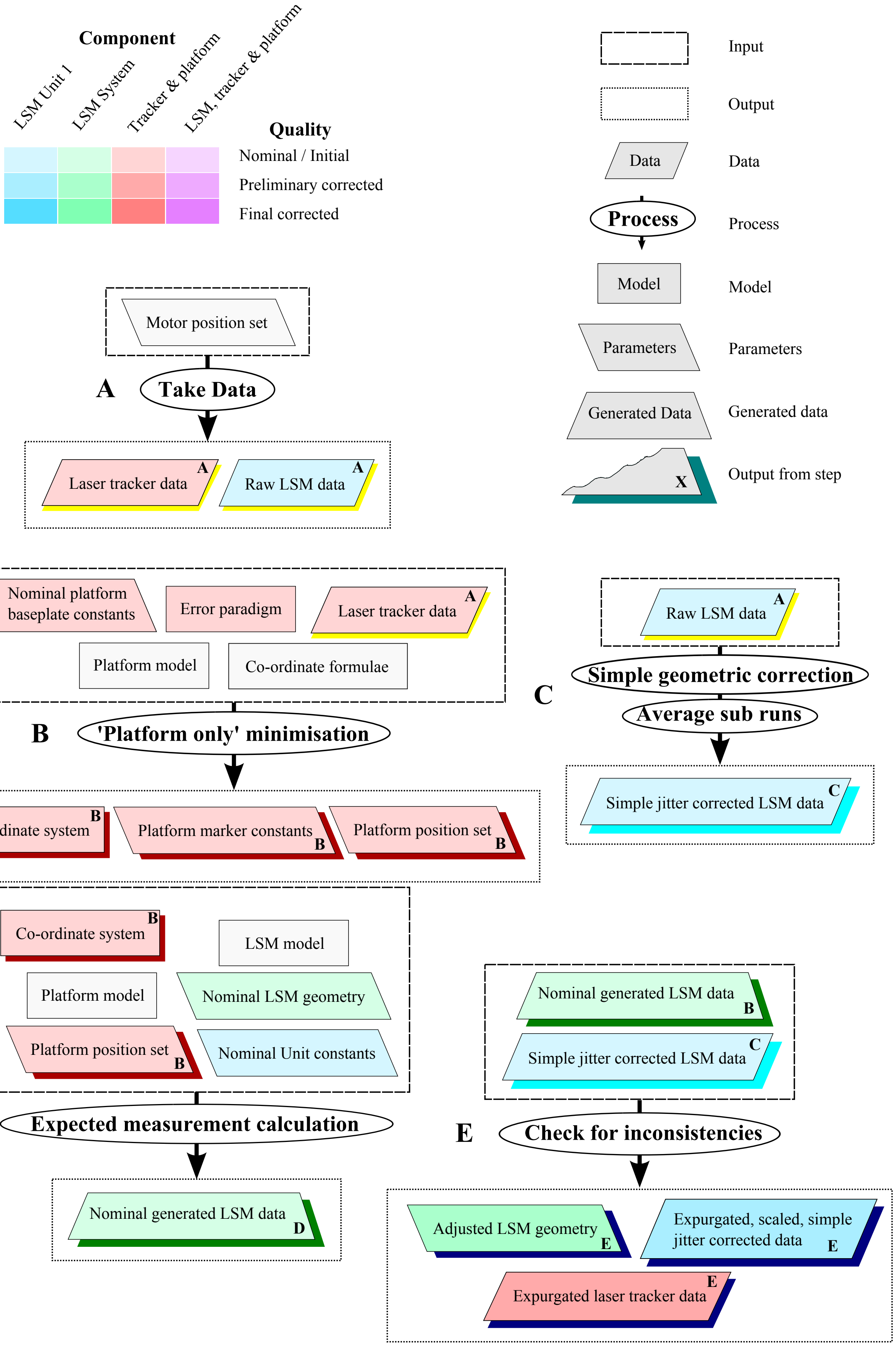


Figure 9.2: Calibration flowcharts covering the initial steps needed to produce suitable data. The component the information pertains to is denoted by the colour, with the saturation deepening as the results are improved. Steps are marked by letters, and results from steps are marked with a coloured drop-shadow and letter.

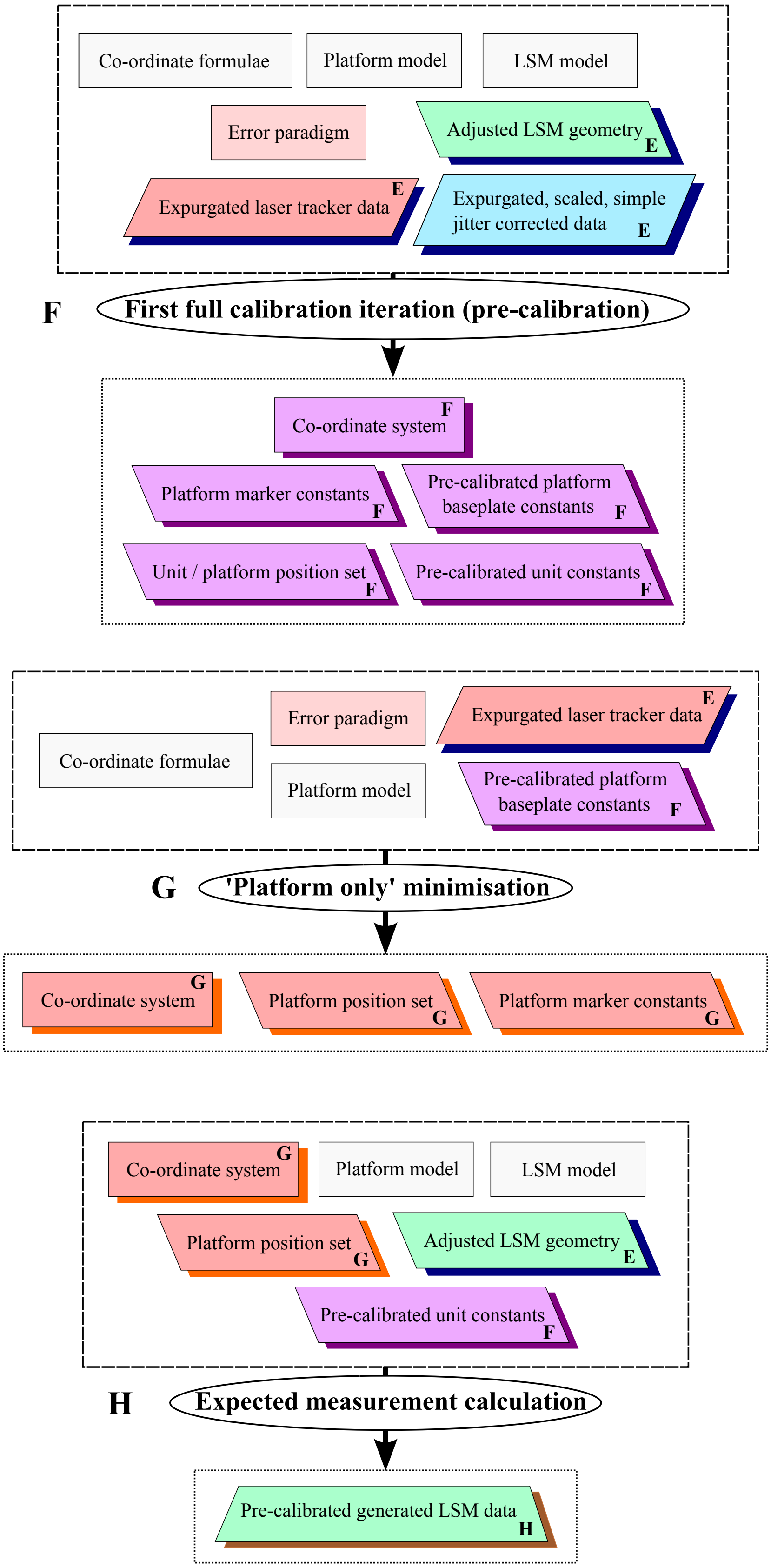


Figure 9.3: Calibration flowcharts covering first iteration of calibration and further steps needed to improve the data.

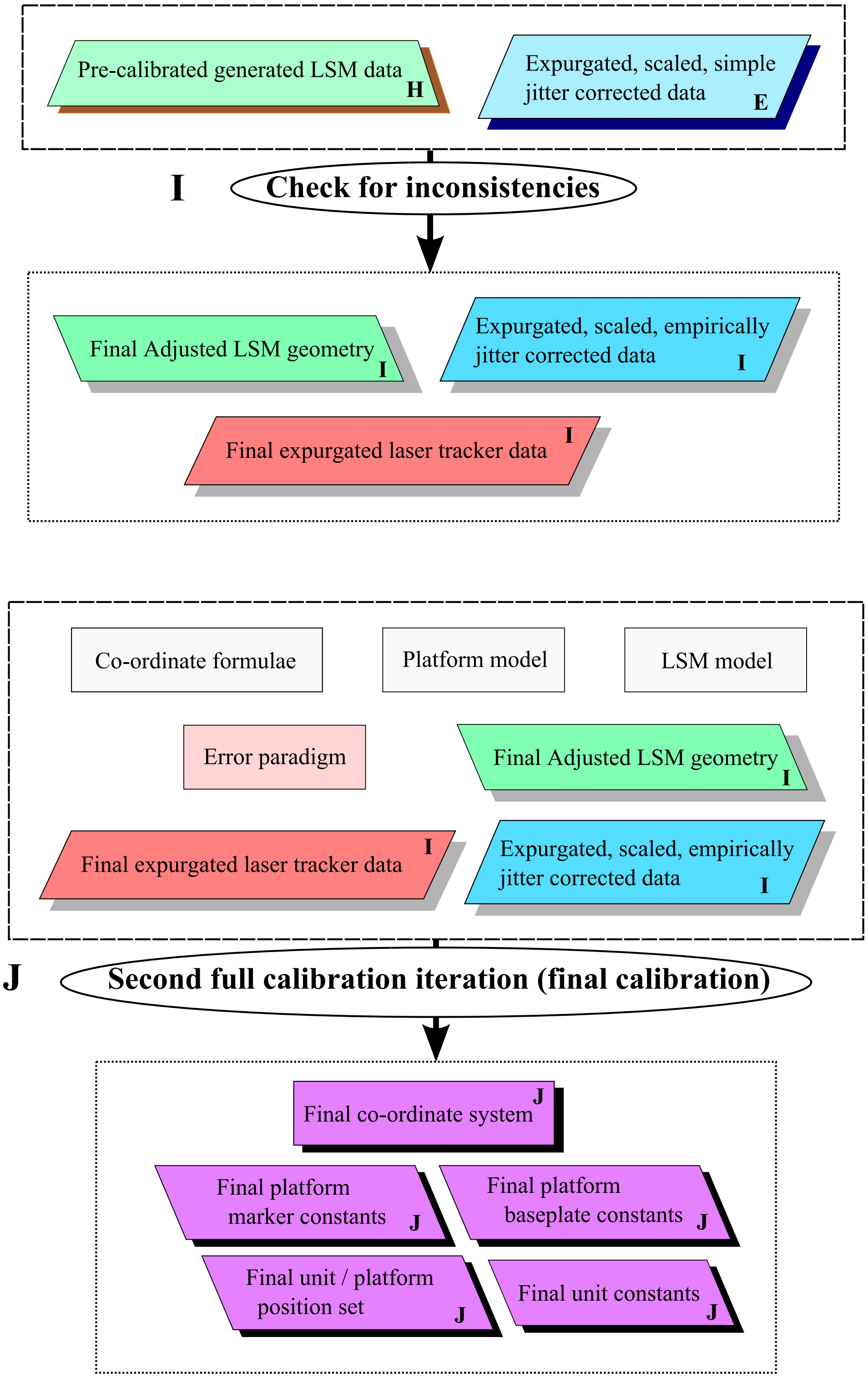


Figure 9.4: Calibration flowcharts covering data and model improvements and the final calibration iteration.

9.2 Taking Calibration Data

9.2.1 Unit Position Coverage

In order to produce an effective calibration, the unit being calibrated needs to be positioned to give measurements over a large range of position/orientation space. The different sub-systems can measure over different ranges of positions and orientations; for this calibration there is no strict need to have the restrictive (±0.75mrad) internal FSI system in range and the external FSI range is much greater than for the other subsystems.

Therefore the restrictions used to define the range of unit motion were: keeping the beams on all four CCDs of dimensions $6468\mu\text{m} \times 4830\mu\text{m}$, and keeping the tilt sensors within their ±1 degree range. As the components are not perfectly aligned, the range of valid positions and orientations is reduced from a theoretical optimum. Care needs to be taken to account for the compound effect of translations with rotations, as the latter only affect CCD 1 and 2. Finally, as unit 1 is attached to an end-cap, the return beam re-alignment process after a move puts limits on large orientation changes between points; if a return CCD beam spot position is near the edge of a CCD, then any large rotation change is likely to take the beam spot off-CCD before the retro mirror is realigned. (The mirror alignment process is detailed in section 9.2.3.)

A program was written that randomly generated positions and orientations that fitted all the restrictions, with the limits conservatively set at ±2mm for lateral translations and with a total CCD beam spot range of ±2mm. Motion along the Z-axis was limited to ±10mm and the Rz rotation to ±35mrad. It was not clear what the best distribution to fill the parameter space was, so a constant probability distribution was used. To avoid the problems with return beam realignment detailed above, the point cloud was checked for potential issues and regenerated until a valid set was found. A model of the action of the stepper motors on the motion of the unit was created and then used to create the motor moves necessary to move through the point cloud. Although the stepper motors did not have encoders and some variation in step size was expected, the resultant positions and orientations of the unit were in line with expectations. Table 9.1 gives the parameter ranges found from the platform reconstruction of section 9.6, while the coverage of the cameras is shown in figure 9.5. They show that there is a reasonable range of values, though there is some potential for increased coverage of the CCDs and a greater parameter volume.

9.2.2 Data Taking Procedure

The general calibration procedure was as follows:

Table 9.1: Total parameter ranges found via platform-only reconstruction

Axis	Units	Range
X	μm	3246
Y	μm	2315
Z	μm	21078
Rx	μrad	6200
Ry	μrad	7210
Rz	μrad	35880

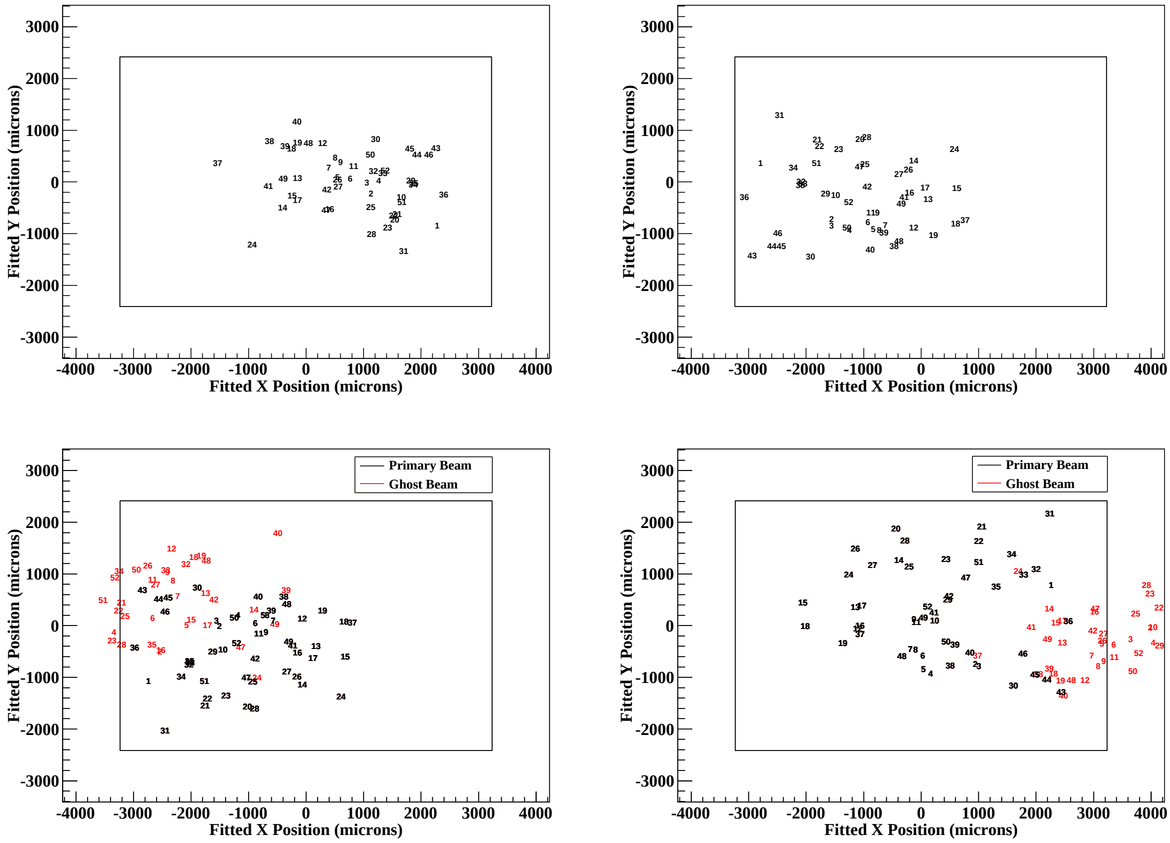


Figure 9.5: Uncorrected CCD beam positions for each set, of the car 1 calibration run for (anti-clockwise from the top right) cameras 3, 1, 0 & 2. For the return cameras the positions of any ghost beams are shown in red. The dimensions of the CCD are shown in each plot. (See table 8.1 for beam spot motions caused by unit motion.)

1. The unit was moved to the position and orientation dictated by the stepper motor moves file.
2. The four nests on the platform were measured by the laser tracker following a single retro while the units took temperature and tilt sensor data. The LSM return mirror is automatically adjusted to give an anti-parallel beam during this step (89 to 93 seconds).
3. The External FSI data was taken (1 second).
4. 3 seconds later, 40 sets of 12 CCD's LSM images were taken, with a 3 second gap between each (123 seconds).

The process took 300 seconds for each of the fifty runs in the calibration, plus three discarded runs at the end, though the total process took just under seven hours to complete.

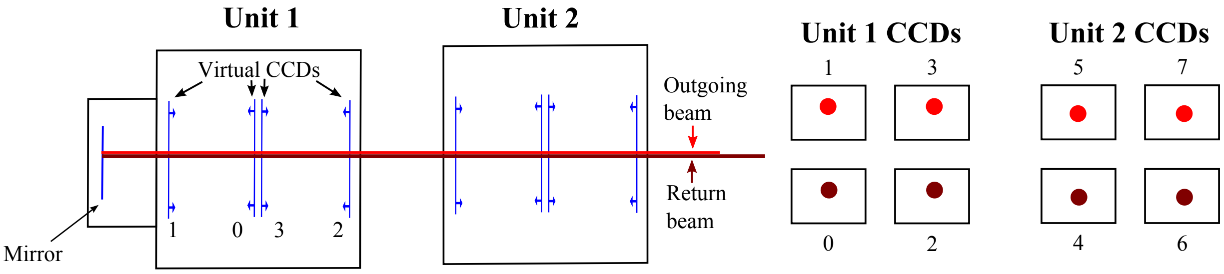
9.2.3 Return Beam Alignment

When calibrating either unit one or unit three, there is an additional complication in that moving the car moves the return mirror and hence the returning beam. Therefore the mirror has to be adjusted to bring the returning beam back into an anti-collinear state. The procedure worked as follows:

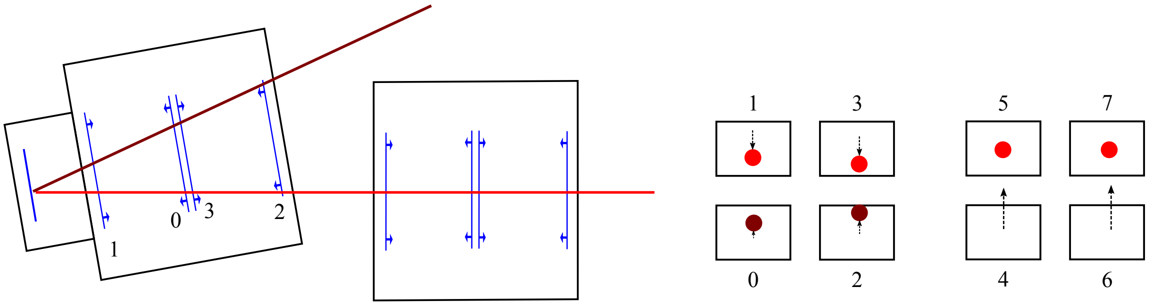
1. Take CCD images for the two outgoing cameras in unit 1: CCDs 1 & 3.
2. Reconstruct the orientation of the unit using a model containing only those two cameras.
3. Trace the beam through a model with a 4-camera set-up, assigned the orientation found by step 2, but with the mirror positioned to give an anti-parallel beam. This gives two calculated return beam positions.
4. Adjust the mirror so that the observed return beams match the calculated beams produced in step 3.

A schematic of the process can be found in figure 9.6. This is enough to give beam spots that are on-CCD on the other cars.

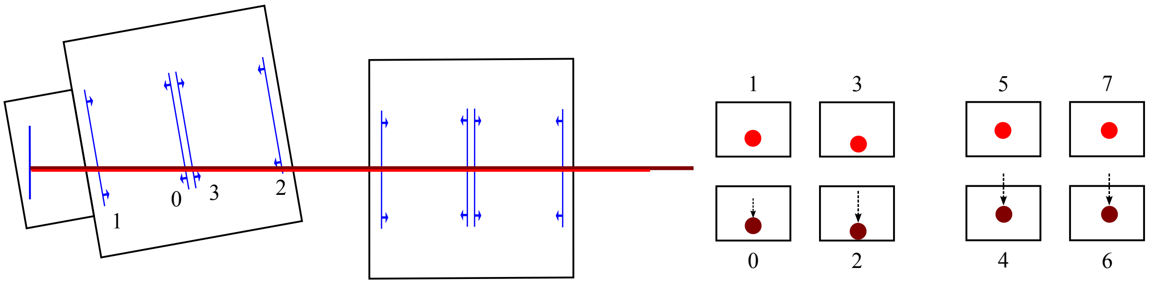
Fine adjustment is then performed by fitting and comparing the beam spot on the other cars with a reference fit position, and adjusting the mirror. Due to difficulties fitting the poor beam on the returning CCDs of the launch unit (car 3), the central unit (car 2) was used. This was still good enough to give beams with an RMS spread of angles of 10 microradians. This remaining



The unit starts in the position from a previous run, with the beam normally incident on the mirror



Immediately after the move the mirror is not aligned, meaning the return beam misses the return cameras on car 2 and 3. The beam still needs to be present on cameras 0 & 2 for the next step to be effective. Note that the beam-spot motion on the cameras is in the opposite direction to the motion if the mirror had not moved.



The expected beam spot positions on cameras 0 & 2 are calculated from the measurements from cameras 1 & 3. The mirror is then adjusted to make the real 0 & 2 beam spots coincide with these calculated positions. The return beam now hits cameras 4 & 6 on car 2, enabling a further minor adjustment to the mirror using the beam spot position on these cameras.

Figure 9.6: Mirror alignment process.

variation in beam direction can be corrected before the data is used (§9.3.3). As alignment was performed automatically during step 2 of the calibration procedure, it did not interfere with data taking.

A further problem is that while both the initial and final orientations of the unit may be in range of the LSM cameras in that unit, the intermediate state with the misaligned mirror may produce beams which are not on-CCD. As can be seen in figure 9.6, the beam spot on a CCD during a rotation moves in the opposite direction to where it will end up after realignment. To prevent the potential problem of an intermediate beam between two moves being off-CCD in unit 1 and precluding the alignment process, the expected CCD positions produced in the intermediate step were calculated and checked for validity when generating the point cloud. With this precaution in producing practical point clouds, the alignment procedure was 100% effective. This procedure could potentially give a slight bias away from the CCD edges and affects the position and orientation valid ‘volume’, but could be replaced with a multi-step move if greater parameter ranges were used. For the ranges used in the calibration experiment, the point cloud only required a few regenerations before finding an acceptable set of 100, suggesting minimal bias.

9.2.4 Environmental Changes During Calibration

The calibration experiment took seven hours to perform, giving the potential for slow changes during the process to manifest themselves. The external temperature of the tunnel, measured by the laser tracker was a constant 24.5°C, with the humidity ranging from 47.0% to 53.25%. Air pressure slowly increased from 998.6mbar to 999.8mbar. The average camera temperature rose linearly from 33.85°C to 34.24°C at run 41, before dropping to 34.21°C over the 80 minute break until run 42, and finally rising to 34.22°C for run 50. The vacuum pressure rose linearly from 7.6×10^{-2}mbar at a rate of 8×10^{-3}mbar/hr, passing 10.0×10^{-2}mbar at run 29. These small variations will have had a negligible effect on the measurements taken. The beam intensity as measured by the first CCD hit, CCD 11, showed a slight fall from a starting value of 65.5 to a value of 61.5 at the end of the calibration, the fractional chance replicated by the fits on the other CCDs. Again, this was not enough to cause any concern.

9.3 Transforming Raw LSM Data for Calibration

For the calibration to be useful for calibrating all the subsystems, the measurements taken must all be comparable with each other. The projection method detailed in section 7.2 is appropriate for reconstruction, where all the non-LSM readings are co-incident to within a second. It is less suitable for calibration, as the measurement of the markers in the platform took 89-93 seconds per run before the LSM data was taken. The beam motion cannot be modelled easily over such time-scales and is significant, so it is not appropriate to try to project LSM measurements to a point in time or range of times using the polynomial fit.

A more comprehensive calibration would model the motion of all three cars at the same time, as well as allow for motion of the mobile optical components and distortions of the platform. However, this is a significantly greater task, both in theory and in computing resources needed, and so is beyond the scope of this investigation. Instead, it is more practical to attempt to correct the data on the car being calibrated by observing the beam motion on the other cars. Then the corrected data can be subjected to the projection process to transform the collection of into a single set of beam spot values per run. Even with restricting the minimisation to a single car with adjusted data, a minimisation using 50 runs typically took five to ten hours on a modern machine and used around 30GB of RAM, both scaling more than quadratically with the number of sub-runs analysed.

9.3.1 Data Pre-Processing

In addition to beam motion requiring correction (see section 9.3.2), in about 0.5% of cases the beams were fitted poorly; either for no clear reason or due to poor ghost beam fitting. A poor fit is characterised by beam spot properties noticeably different from similar fits and a higher chi-squared value. Although the range of the calibration was chosen so that the primary beam was on all the CCDs, for the return cameras the ghost beam was frequently barely on the camera at all, or even absent entirely.

In principle, the fitting algorithm is capable of detecting completely off-CCD ghost images and automatically re-fit without a ghost beam spot in the model, but it struggled to tell which was a better fit as the chi-squared value is often not lower for the single beam version. To counter this, the fit residual images were manually checked and a handful of images were forced to be fitted with a single beam. Revisiting figure 9.5, it can be seen that there are visible ghosts on about two thirds of the images for the two cameras which measured the return beam. Fortunately there is only one case (run 47, CCD 0) where the ghost is close to the primary beam and manual inspection showed a good fit. There were no instances of the ghost being mistaken for the real beam.

To counter the rare remaining poor fits, before projection the beam spot fits have a simple check for outliers: the standard deviation of the positions of the good beams in the beam set is taken, and any beams outside of six standard deviations are marked as bad. The process is repeated five times with the remaining good beams. This method was sufficient to mark all outliers that were manually checked.

The final method to produce the corrected, projected data for use in the calibration is as follows:

1. Take each collection of 12 cameras' worth of 40 fitted beam positions, i.e. 40 sub-runs taken during a single run, ignoring positions marked as unreliable by the fitter.
2. Mark any outlier beam positions as bad.
3. Correct the car 1 beam positions using the process described in section 9.3.2 and section 9.3.3.

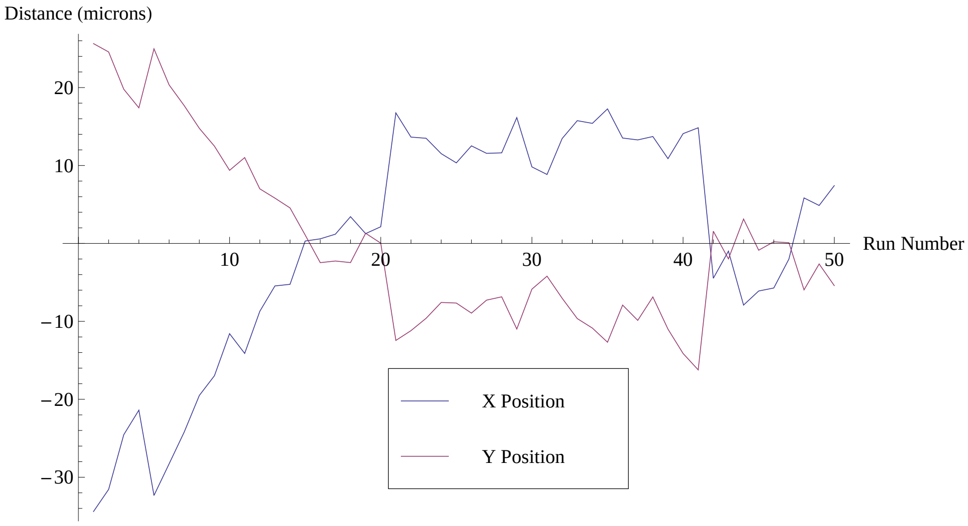


Figure 9.7: The deviation from the mean position of the beam on car 2, camera 3 over the course of the car 1 calibration run.

4. Mark any outliers as bad again to catch any problems with the correction.
5. Project each collection into a single set of 12 cameras' worth of beam spot fit positions.

The laser tracker data is converted to radians and metres. The yaw measurement is converted to be right-handed and transformed to the range 0 to 2π.

9.3.2 Correction of outgoing beam positions

The most significant issue was a dynamic motion of the angle of the incident laser beam from its mean direction of the order of 10μrad. While a problem for the rotations which are of the order 2-20 milliradians with a target uncertainty of 5μrad, it manifests itself as a more serious 100 micron translational motion on car 1. Figure 9.7 shows the motions of the beams on car 2, camera 3. This is the outgoing camera with the virtual image almost on the centre of symmetry of the unit. It shows the beam move to point further away from the wall and more towards the floor. Note the extremely high correlation between the axes, suggesting a single rotation about an axis 35-40 degrees from vertical, potentially caused by the beam launch.

To correct the outgoing CCD beam spot data, the following calculation produces the necessary correction on the car 1 cameras. It assumes the motion is caused by a car 3 rotation and car 2, camera 3 is treated as the reference:

$$C_o = -d_{2o}\frac{l_{13}+\delta_i}{l_{23}}\,, \tag{9.1}$$

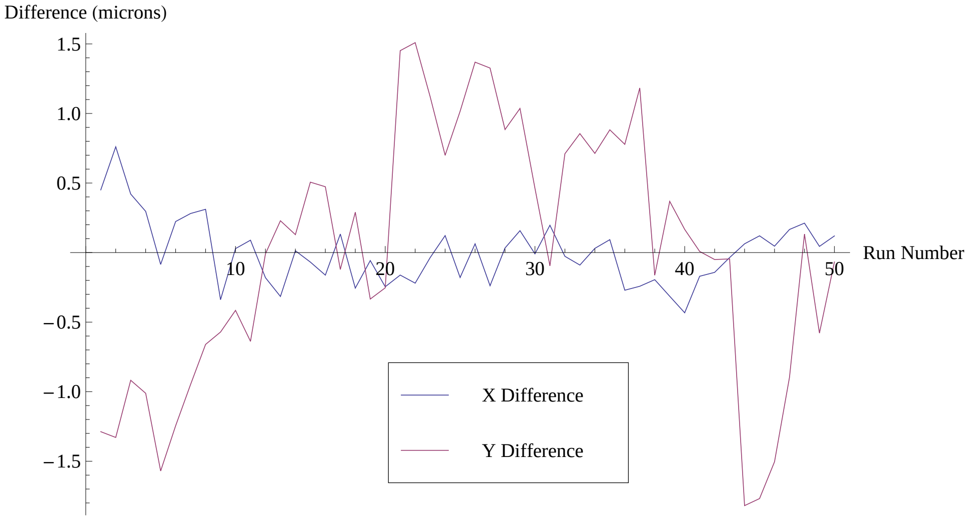


Figure 9.8: The difference between the deviation from mean of the beams on car 2, camera 1 and car 2 camera 3 over the course of the car 1 calibration run.

where d_{2o} is the outgoing beam position deviation on car 2, camera 3, l_{13} is the distance from the centre of car 1 to car 3, δ_i is the adjustment into l_{13} based on the virtual camera position with respect to the car centre, and l_{23} is the distance from car 2 to car 3. The r.m.s of the corrections produced is 30 microns in X and 22 microns in y.

The stability of car 2's orientation can be checked by observing the output of car 2's tilt sensors. As car 2 was levelled for the experiment, the fine readout can be used. The Rx measurement gave no overall motion with a standard deviation of 4.8μrad. The Rz measurement gave a slow drift of 1.2μrad over the calibration, with a RMS deviation from a linear fit of 0.2μrad.

Figure 9.8 shows this difference between the fitted beam positions on the two car 2 outgoing cameras: car 2, CCD 1 and car 2, CCD 3. Inspecting the difference removes effects from the translational motion of the car, be it absolute or due to the beam moving, leaving only rotational effects. The differences had a standard deviation of 0.23 microns in X and 0.90 microns in y. A clear correlation with beam position of figure 9.7 is evident, though the magnitudes are different, suggesting the centres of rotation may be different for the different axes.

A simple estimate of the expected difference due to angular motion of the beam was made by considering the 40-50 micron motion in figure 9.7, corresponding to a 10μrad rotation if the centre of rotation is considered to be around the launch 5 metres away. Assuming a 20cm spacing between the virtual CCDs gives a 2 micron difference; roughly the same as the observed values.

If the correction described by equation 9.1 is applied to the data, the standard deviation rises to 0.58 microns in X, but falls to 0.62 microns in Y. This implies this is the limit of car 2's physical and fit stability when used for correction.

9.3.3 Correction of return beam positions

The correction discussed in section 9.3.2 is valid for the return cameras, under the assumption that the return beam is anti-collinear. This is not necessarily the case as the realignment process of section 9.2.3 is not perfect, so an additional correction is appropriate. This is given by

$$C_r = -\left(d_{2r} - d_{2o}\right) \frac{l_m - \delta_i}{l_m + l_{12}}\,, \tag{9.2}$$

where d_{2r} is the return beam position deviation on car 2, camera 0, l_m is the distance from the centre of the unit to the mirror, and l_{12} is the distance from car 1 to car 2. This gives an additional correction to the data with an RMS of approximately 3 microns (camera 0) or 5 microns (camera 2), though the overall RMS correction including the outgoing part remains at 30 microns (X) and 22 (Y) microns. The correction is much smaller as the centre of the second rotation is known to be the mirror, which is much closer to the car 1 cameras.

9.3.4 Limitations of Corrections

Correcting the data relies on the following assumptions:

- The motion of the units not being calibrated during data taking is minimal
    - The corrections assume that car 2 does not move at all and there is no translational motion of car 3. The checks in section 9.3.2 show car 2 is stable enough for this to hold. Translational motion of car 3 can be absorbed into the centre of rotation.
- The centre of rotation of the beam launch is the same for both X and Y axes and does not change
    - In reality, the beam jitter is caused by both the launch optics moving and car 3 itself rotating and there is no overall reason for a common centre of rotation or it remaining constant.
    - If the centre of rotation is off by 10cm, this will give a 2% error in the correction, which would correspond to an RMS residual of around 0.5 microns on car 1; a similar value to the precision of the beam spot measurement itself.

- The directions of the axes for car 2 are the same as those of car 1.
  - While this is clearly not true, the small level of rotation will mean the mismatch will be of the order milliradians, meaning any mixing will be negligible.
- Any significant remaining beam spot motion after correction can be modelled accurately.
  - As the model used for correction is imperfect, residual beam spot motion may remain which may leave structure in the sub-run by sub-run data. Manual fine tuning of the correction allows for a significant reduction of remaining structure.

The correction also couples the measurements taken with car 1 to the measurements taken with car 2, increasing the error on the corrected car 1 data slightly (as shown in figure 9.9) and reducing the robustness. There is potential to improve the correction model by analysing the data produced, which is covered in section 9.7.4.

9.3.5 Projection Details and Beam Spot Fit Precision

Figure 9.9 demonstrates the advantage of correcting the data before making the projection. The uncorrected data clearly shows beam motion over the course of a measurement run. When extrapolating to nearby periods of time it is reasonable to fit with a polynomial, though this is clearly not valid for more distant times as the motion does not tend to infinity. Table 9.2 shows some properties of the projections of figure 9.9. For the uncorrected data, the projections using different orders produce values that are not consistent with the error estimates and there is a significant drop in the RMS residual as the order of the fit increases. This shows that the simple polynomials are not good models for the data. With the corrected data, the different projections are consistent with each other and the RMS residual barely improves with increasing order. Note the RMS residual is actually higher due to the random error on the car 2 camera 3 measurement used in the correction; the RMS error of the corrected data around the order 2 polynomial for car 2 camera 3 is 0.47 microns in the X-axis and 0.32 microns in the Y-axis. This suggests that the majority of the beam motion has been removed and that it is appropriate to use a zero-order projection, effectively a simple average, to produce the final beam position measurement. Further beam spot hit position mesurements presented were all calculated using this method.

Figure 9.10 shows typical error estimate distributions for the point projection in both the X-axis and Y-axis, taken from 53 runs of CCD 1 data, for a zero order polynomial projection. Table 9.3 shows the results of the analysis for all of the car 1 beam fits, averaged across all 53

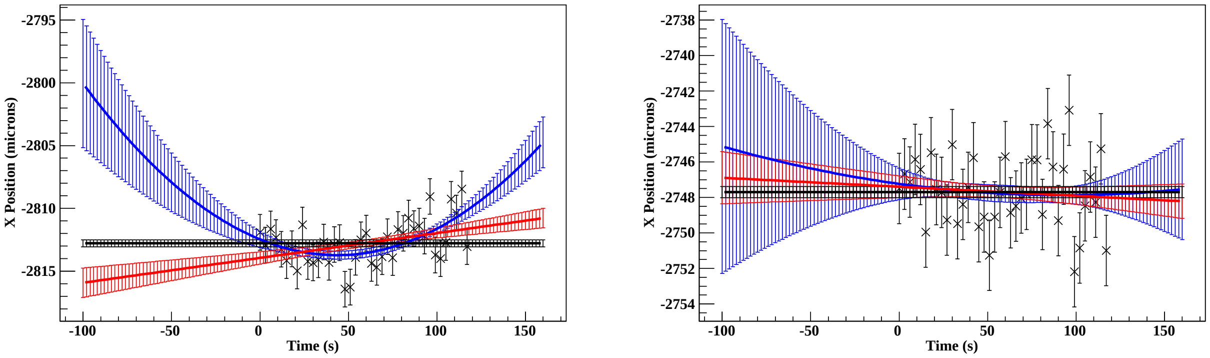


Figure 9.9: Typical projections with uncorrected (left) and corrected (right) data. The example shown is for the CCD 3 X-position. Best-fit polynomials of three orders (black: 0, red: 1, blue: 2) are presented with the associated error on point estimates.

Table 9.2: Properties of projections of a the beam X-position on car 1 CCD 3 for a the first set of 40 images in the car 1 calibration data.

	Polynomial order	0	1	2
Uncorrected	RMS residual (μm)	1.72	1.59	1.42
	t=$-100s$ Projected value (μm)	-2812.8	-2815.9	-2800
	Error estimate (μm)	0.27	1.2	5.1
Corrected	RMS residual (μm)	2.00	1.99	1.99
	t=$-100s$ Projected value (μm)	-2747.7	-2746.9	-2745
	Error estimate (μm)	0.32	1.47	7.2

runs. The mean RMS residual for the data for the projection functions of different orders is given, as well as the calculated error for a point at $t = -45$ seconds. In the case of the zero-order polynomial, the calculated error is the sample standard deviation. There is not a strict $\sqrt{40}$ ratio between the RMS residual and the sample standard deviation as some of the data-points are discarded as outliers before projection according to section 9.3.1. Comparing the ratio of the mean RMS residuals for the zero-order and second-order fits gives a 13% reduction in mean RMS residual for uncorrected data, while only a 4.6% reduction for corrected data. This demonstrates a clear reduction of beam motion hinted at by the individual run above and although not perfect, the corrected zero-order projection beam position values are used for all remaining calibration work.

The error estimates produced by the projection method give a reliable estimate of the statistical precision of each individual beam position measurement after correction, specific to the CCD, axis and run number. Table 9.4 summarises these mean values found over the fifty runs.

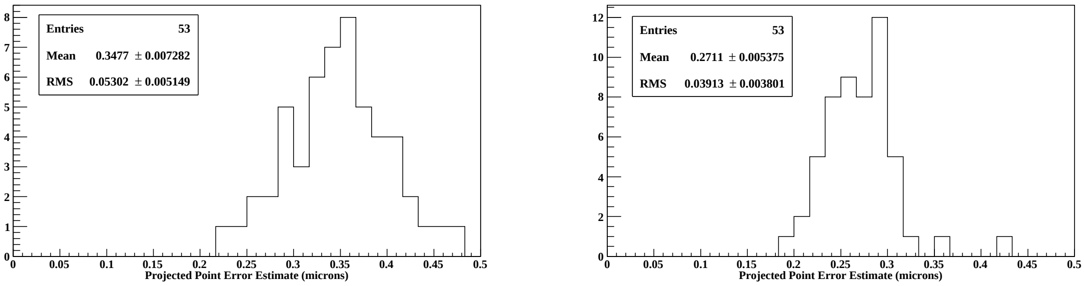


Figure 9.10: Error estimates for fitted CCD 1 beam hit positions in the X-axis (left) and Y-axis (right).

Table 9.3: The mean CCD fitted beam spot position RMS residuals, and the mean error on a $t = -45$s projection for all car 1 CCDs, using 53 runs' worth of uncorrected and corrected calibration data for both the X and Y axes. Ghost beams have not been corrected. The representative values for the errors used in the calibration are in bold.

		Uncorrected			Corrected		
	Order of extrapolation	0	1	2	**0**	1	2
CCD 0, X	mean RMS residual	1.46	1.37	1.35	2.33	2.28	2.22
	mean projected error	0.23	0.69	2.02	**0.37**	1.15	3.34
CCD 0, Y	mean RMS residual	2.34	2.10	2.01	2.49	2.41	2.35
	mean projected error	0.37	1.05	3.00	**0.40**	1.21	3.53
CCD 0 (ghost), X	mean RMS residual	26.98	26.40	24.48	-	-	-
	mean projected error	6.94	20.94	58.54	-	-	-
CCD 0 (ghost), Y	mean RMS residual	64.89	63.05	57.28	-	-	-
	mean projected error	18.58	55.55	154.06	-	-	-
CCD 1, X	mean RMS residual	1.31	1.23	1.22	2.20	2.16	2.10
	mean projected error	0.21	0.62	1.81	**0.35**	1.08	3.12
CCD 1, Y	mean RMS residual	1.43	1.23	1.18	1.71	1.68	1.65
	mean projected error	0.23	0.62	1.76	**0.27**	0.84	2.46
CCD 2, X	mean RMS residual	1.18	1.06	1.05	2.09	2.05	1.99
	mean projected error	0.19	0.53	1.57	**0.33**	1.03	2.99
CCD 2, Y	mean RMS residual	1.43	1.22	1.15	1.63	1.57	1.55
	mean projected error	0.23	0.61	1.73	**0.26**	0.79	2.33
CCD 2 (ghost), X	mean RMS residual	18.44	18.23	17.90	-	-	-
	mean projected error	3.40	11.08	34.56	-	-	-
CCD 2 (ghost), Y	mean RMS residual	4.28	4.16	4.07	-	-	-
	mean projected error	0.73	2.25	6.69	-	-	-
CCD 3, X	mean RMS residual	1.19	1.11	1.09	2.11	2.06	2.00
	mean projected error	0.19	0.55	1.62	**0.33**	1.03	2.99
CCD 3, Y	mean RMS residual	1.35	1.15	1.09	1.63	1.59	1.57
	mean projected error	0.21	0.57	1.63	**0.26**	0.79	2.34

Table 9.4: The mean values of beam spot fit precision estimates for a zero order polynomial projection from a set of 53 runs' worth of calibration data.

CCD Number	Mean error estimate (μm)	
	X-axis	Y-axis
0	0.37	0.40
1	0.35	0.27
2	0.33	0.26
3	0.33	0.26

9.4 CCD Data Consistency Checks

The calibration minimisation is very sensitive to incorrect or outlying data. Therefore, it is important to make a number of checks on the data used to see that it conforms to expectations.

9.4.1 Coincident Virtual Cameras

As the virtual images of CCD 0 and CCD 3 are only 2mm apart, if the return beam is anti-collinear then the movement of the beam spots on these cameras, after transforming to account for inversion, should be very similar. Plotting the difference can reveal fitting issues or problems with the return beam. Figure 9.11 shows the CCD 0 − CCD 3 values, for the calibration runs. Runs later excluded after the analysis in sections 9.7.5 and 9.8.3 are highlighted and show some deviation from the typical values. There is no clear correlation of the difference with any known model parameter or data value, but the RMS difference is 12.9 microns in the X-axis and 13.7 microns in the Y-axis; these are far higher than the values expected from the the small angular range of table 9.1 and the sub-micron fit precision found in table 9.4. This issue is investigated further in section 9.5.

9.4.2 Equivalent Pairs of Cameras

In a similar manner to the coincident virtual CCDs, the difference between the beam positions on CCDs 0 and 2 should be the same as the corresponding difference between CCDs 1 and 3. This can be understood by realising that an LSM system can be made with a single pair of CCDs and that of the six possible combinations, four are easy to follow: the 0-2 (return) pair should give the same values as the 1-3 (outgoing) pair, which is equivalent to the 0-1 pair giving the same results as the 2-3 pair. Figure 9.12 shows the (CCD 0 − CCD 2) − (CCD 1 − CCD 3) values for

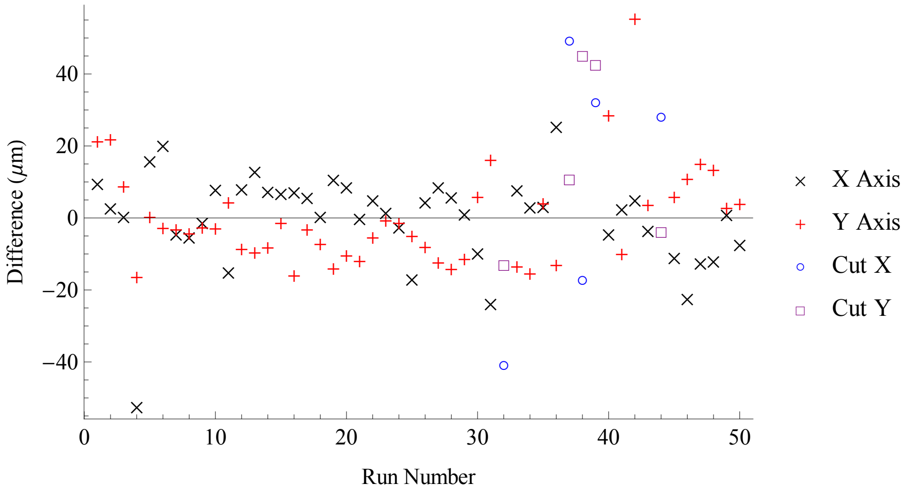


Figure 9.11: Difference between fitted beam positions on CCD 0 and CCD 3 for the runs used in calibration, normalised to a mean of zero. Runs marked bad via later checks are highlighted with open circles and squares.

the calibration runs, with the cut runs also shown. The RMS difference is 23.9 microns in the X-axis and 20.8 microns in the Y-axis. These values are broadly consistent with CCDs with a beam fit precision of 12 microns; this is similar to the findings of the 0-3 pairing. There is some correlation of the position with run number up to about run 30, with slightly higher variance data thereafter, but overall the data is self-consistent to the level expected from the earlier coincident virtual CCD check.

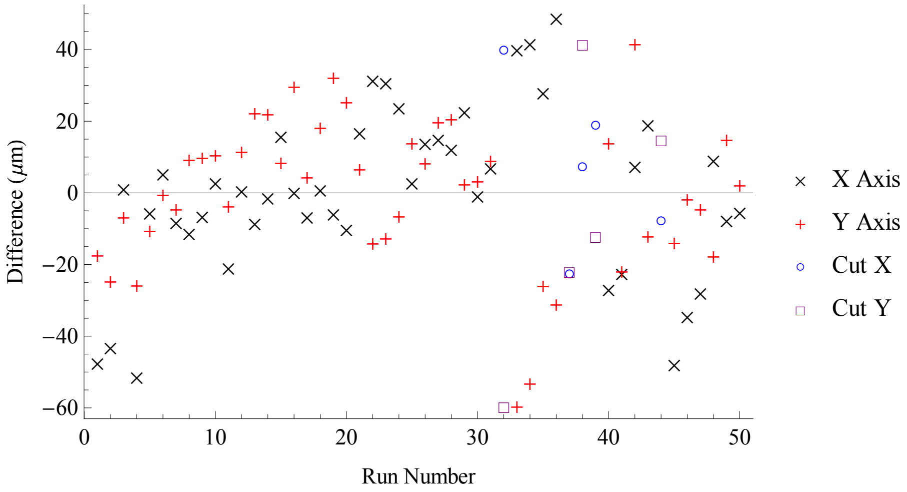


Figure 9.12: The residual from the 1-3 (outgoing) pair subtracted from the 0-2 (return) pair, normalised to a mean of zero. Runs marked bad via later checks are highlighted with open circles and squares.

9.4.3 Additional Checks

Other simple checks can be made to see if the CCD data will give approximately the right reconstruction.

- The motion of the beams on CCD 0 and CCD 3, which are close to the centre of rotation of the unit, should nearly correspond with the translation of the unit: a 20mrad rotation of the unit, enough to be near the limit of the 'outer' cameras would only give a beam motion of order 20μm on the 'inner' cameras.
- The CCD 0 $-$ CCD 1 difference and the CCD 2 $-$ CCD 3 difference should nearly correspond to the rotation of the unit,after division of the lever-arm distance between the virtual CCDs. $(0.2m)$

Therefore plotting these combinations, with suitable transformations to account for orientation of the CCDs, will give a rough prediction of the translations and rotations. Figure 9.13 contains the results of these crude reconstructions, again with runs later found to be outliers highlighted. The different methods agree with each other, give values consistent with the expected range, and when checked against figure 9.36 produced as part of the final calibration also show general agreement. There is generally good agreement between the two different cameras for the translations and there is also generally good agreement with the two different pairings. However, some of the bad runs show a noticeable difference in the Ry-rotation in particular, suggesting that for those particular bad runs, the CCD data might be bad; this observation is later confirmed in section 9.7.5.

9.5 Quasi-Systematic Errors

The data checks highlight a serious issue with the CCD beam spot projected fit position data: although the beam fit precision is consistently about 0.3 microns (see table 9.4), the self-consistency arguments of sections 9.4.1 and 9.4.2 suggest a value an order of magnitude greater. Later corrections do improve the consistency of the measurements slightly (§9.7.7), though they remain an order of magnitude above expectations. There is a significant difference between the repeatability (having the same position and orientation) of a measurement and the comparability (having different position and orientations) of separate measurements.

9.5.1 Background

Consider an observable, O, which can be measured with a precision of σ_o producing a set of individual readings. If O is constant during the measurement, then the readings can be combined

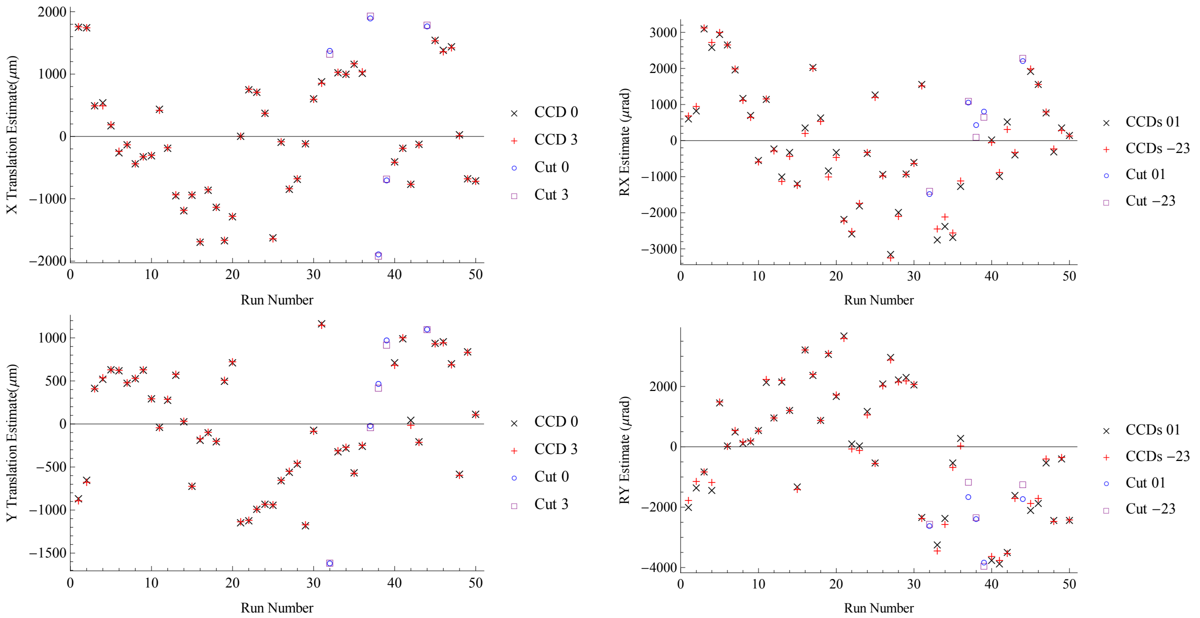


Figure 9.13: Crude translation reconstruction values obtained using data from CCD 0 and CCD 3 (left), and crude rotation values obtained by the CCD 0-CCD 1 pair and CCD 2-CCD3 pair. Runs marked bad via later checks are highlighted with open circles and squares.

to give a best estimate $\hat{O}$, with a precision of $\sigma_o/\sqrt{N}$, where N is the number of readings. $\hat{O}$ may be subject to an offset caused by the experimental set-up, which is the systematic error.

If O is not a constant, but a function of an unknown parameter p, giving $O(p)$, then as long as p has a constant value of p_0 during each reading of a single measurement, then the measurement will give an estimate $\hat{O}(p_0)$, again with precision $\sigma_o/\sqrt{N}$. However, a subsequent measurement, taken at a different value of p, would give a different value: $\hat{O}(p_1)$ with the same precision. If the existence of x and the dependence of O on it is unknown, the values of $\hat{O}$ are unlikely to be consistent with each other: as the underlying true value of O is different; it is the model that is incomplete. Figure 9.14 shows example measurement values for a function that does not depend on p but has a large measurement error, and example values for a function that does depend on p but has a small measurement error. Both would have similar variances on the mean value of O, but the first is due to the measurement error, while the second is due to deficiencies of the model. As long the systematic offset function is oscillatory, has no long-range structure with p, and is sampled at many different values of p, it has the same effect as a statistical error. Therefore, statistical error propagation is appropriate, and the errors can be considered as quasi-systematic. There is still room in the paradigm for non-dynamic systematic errors, meaning calibration is still possible for this subset.

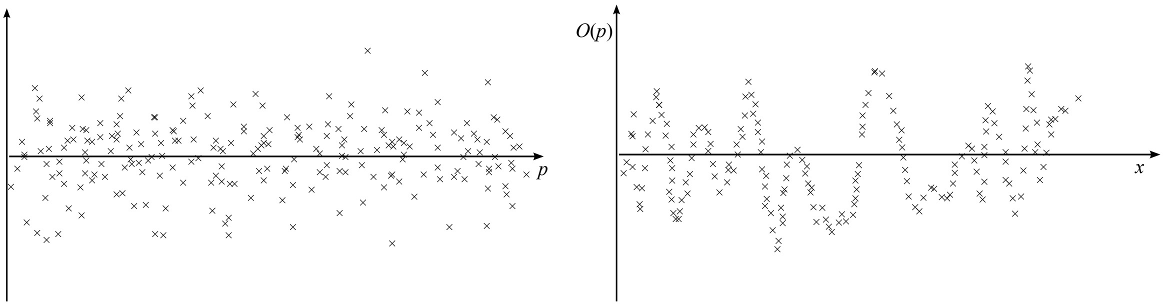


Figure 9.14: Example distributions exhibiting large statistical-only uncertainties (left) and a combination of a small statistical and large quasi-systematic (right) uncertainties.

9.5.2 Investigation using LSM Stability Data

On a CCD, the deviation from the truth of the beam spot fit position is a function of numerous variables. We expect the most significant un-modelled variable to be the position of the beam on the CCD, with the majority of the deviation caused by the effects of dust biasing the fit or possibly edge effects. (See chapters 6 and 7.) It was deemed too difficult to model the distortions to fitted beam positions by dust, particularly as the patterns were not always stationary and it would require significant extra data to build up a good model of each CCD's particular dust distribution. These effects therefore become a quasi-systematic error, with an apparently random dynamic offset, which is actually dependent on the pre-correction beam fit position, and hence indirectly the position and orientation of the unit and beam jitter.

Although significant effort has been made to reduce these effects, there appears to be a residual offset function with magnitude of around 10 microns, swamping the statistical precision of 0.3 microns.

Data from a stability experiment can be analysed to check this hypothesis. With the RTRS set up identically to the car 1 calibration (figure 9.1), 200 runs' worth of LSM data was collected over a 35 hour period with all cars left untouched. As can be seen in the left-hand plot in figure 9.15, which shows the beam motion on car 1, there was significant slow beam motion over the course of the experiment similar to the motion observed in the calibration experiment. (See figure 9.7, noting the CCD 3 Y-axis inversion in figure 9.15.) There was a general settling process over the first 40-50 runs (8 hours) but continued motion thereafter. The total range is 350 microns in X-axis and 1200 microns in the Y-axis: far more than the equivalent values of 100 microns and 80 microns in the calibration experiment, found by doubling the car 2 motion of figure 9.7. This is about one beam half-width or 25% of the CCD height. There is a strong correlation of motion between the axes and the same pattern is seen in the difference between the two camera values consistent with the launch optics rotating. (See the right-hand plot in figure 9.15.)

The geometric correction method of section 9.3.2 could be applied, but for the 0-3 comparison a simpler method is available: to remove the effect of the different magnitude of beam spot motion on the two cameras, a simple minimisation on the standard deviation of the difference between the two CCDs against a scale factor applied to the fitted CCD 0 beam spot position was performed. This simply reduces the beam spot motion on CCD 0 until it matches that of CCD 3. It does not rely on guessing the beam jitter centre of rotation but only works if the beam jitter is large compared to the 0-3 residual, and hence requires the unit to be stationary to avoid the quasi-systematic effects. Similarly, the correlation of the 0-3 difference with the 0-3 sum as a function of the scale factor was solved to find the zero-correlation value, giving identical results. Note that although this scaling method artificially reduces the noise on CCD 0, this effect is small as the scale factors found are near unity with values of 0.932 and 0.948. The optical path length to CCD 0 is roughly 9.77m while the path length to CCD 0 is 9.25m, giving an expected scale factor of 0.95, assuming the launch optics are the centre of rotation.

Table 9.5 shows the effect of this correction: with the beam motion removed the standard deviation of the difference is reduced down to sub-micron levels. Assuming that both CCDs contribute equally to the standard deviation and that the individual errors are uncorrelated, representative CCD beam fit position error estimates of 0.39 microns in the X-axis and 0.29 microns in the Y-axis are produced. These values are an excellent match to the calculated precision in table 9.4, confirming the hypothesis. Interestingly, as can be seen in the bottom plot in figure 9.15, there is still some structure in the difference over time, suggesting a residual unmodelled contribution - a small quasi-systematic error. The slowly varying structure could be caused by the dust as the beam moves 25% of the CCD height, but over the time-scale of 35 hours, there is potential for various other effects to come in to play. However, the complexity of this particular quasi-systematic is apparent, even if its magnitude is very small.

The difference in the X and Y correction scale factors is no cause for concern as the overall centre of rotation does not have to be common between the axes. While the strong correlation suggests a single source, possibly the beam launch optics, an additional physical translation such as the launch unit settling would break the degeneracy, even if it is correlated with the rotation. Using the reference distance between the CCD images of 0.52m, a simple geometric calculation gives the distance from the unit 1 centre to the centre of rotation of the beam jitter corresponding to the scale factors found. Table 9.5 also shows the optical distance from the centre of beam rotation to CCD 3 that is implied for the geometrical model to produce the minimum standard

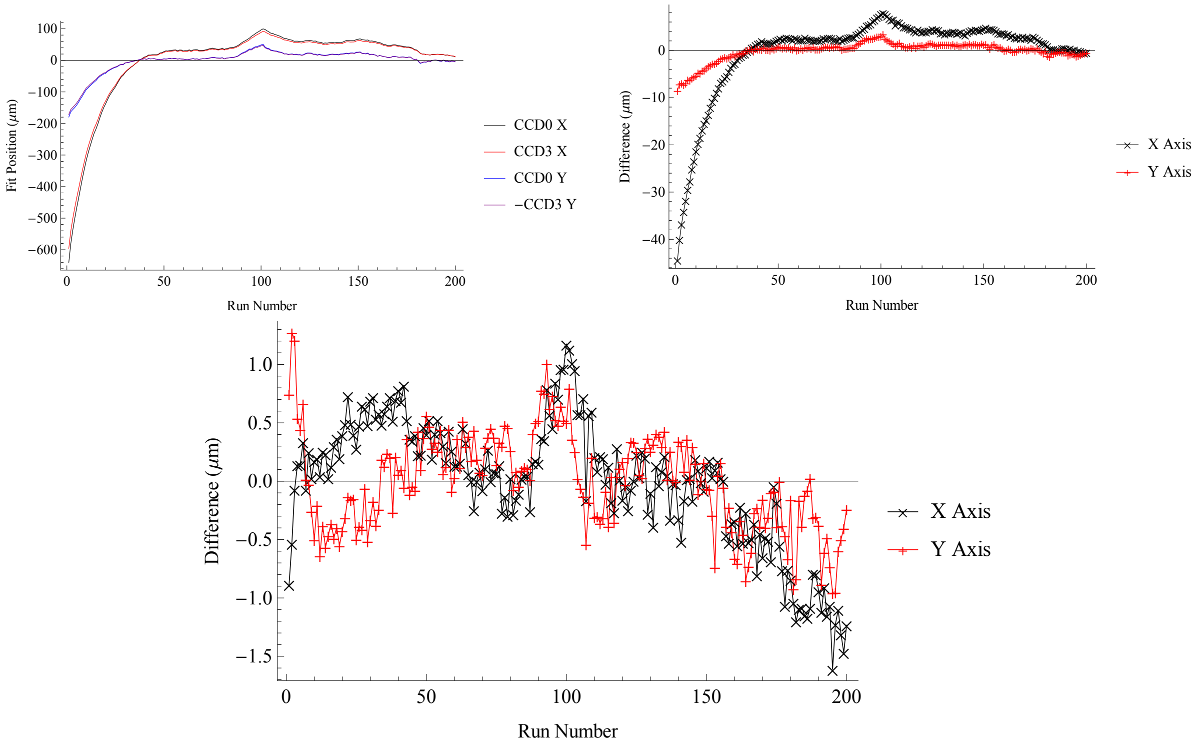


Figure 9.15: Fitted beam spot positions normalised to a mean of zero (left) for CCD 0 and CCD 3 over a 35 hours stability run and the difference between the positions before (right) and after correction (bottom). The RMS differences are 8.6μm in the X-axis and 2.0μm in the Y-axis for the uncorrected data, and 0.55μm in the X-axis and 0.41μm in the Y-axis for the corrected data.

deviation CCD 0 scale factor. The values are roughly consistent with the location of car 3, about 9 metres away, with the larger effect of the motion in the X-axis leading to a closer centre of rotation, possibly caused by an X-translation of the launch unit.

9.5.3 Effect on Calibration

Having observed that the 0-3 difference has no clear correlation with any known parameter and with no reliable way of reducing the effect found, it is reasonable to treat the CCD beam spot position quasi-systematic errors as dominated by the quasi-systematic effects. (Further checks are made later in section 9.9.2.) Any parameter values produced using the per-run statistical errors produced by the projection process are likely to have biases caused by incorrect weighting of the data and should not be considered reliable. However, the per-run precision estimates are still useful in determining the parameter errors possible in the absence of quasi-systematic errors, and offer a guide to the repeatability of measurements.

Having elected to use the quasi-systematic errors, there is still the problem of finding out what their values are! The values estimated by the process of section 9.4.1 are only valid for the combination of CCD 0 and CCD 3. Estimates of these quasi-systematic errors are found later in

Table 9.5: The effect of scaling CCD 0 data on the standard deviation of the 0-3 beam spot fit position difference on stability run data taken over 35 hours. The scale factor associated with the minimum values for both CCD axes is given as well as the scale factor found by minimising the correlation of the difference with the sum. The geometrical scale factor distance is the distance to the beam centre of rotation assuming a 0.52m optical distance between the CCDs. The bold values are comparable with the CCD 0 and CCD 3 values of table 9.4.

	X-axis	Y-axis
Uncorrected Standard Deviation (μm)	8.63	2.01
Minimum Standard Deviation (μm)	0.55	0.41
$1/\sqrt{2}\times$Minimum Standard Deviation (μm)	**0.39**	**0.29**
Min SD CCD 0 scale factor	0.932	0.948
Zero Correlation CCD 0 scale factor	0.932	0.948
Geometrical scale factor distance (m)	7.1	9.6

section 9.8.3 by comparing expected beam spot position values calculated using the positions and orientations found using an external witnessing system. This means that while the error estimates can be specific to the CCD and the axis, the value is constant for every run. These final errors used in the final calibration, given in table 9.16, are consistent with the 0-3 values of section 9.4.1.

9.6 Witnessing Unit Motion

Witnessed calibration works by performing measurements with a unit while observing it using an external system. If the external system's results are treated as a yardstick, then the problem may be factorisable and the reconstruction made by the two systems may be done separately. That is, the calibration would be performed by measuring a set of translations and orientations of the unit with the external system and then adjusting the calibration constants of the unit until its reconstructions differ by as little as possible from the externally measured values. However, this approach does not take into account the uncertainties of the external system which is potentially a serious shortcoming if the systems are of comparable precision. A better approach is to combine the two systems; both the external system and the measurement units are modelled together and the resultant model used in one combined minimisation. Therefore the parameters of the models, including the sought after calibration constants, are adjusted until all the readings are as consistent as possible with each other when weighted appropriately.

The motion of the unit was witnessed by attaching a custom built measurement platform to it and measuring its motion by means of its four retro-reflector nests using a laser tracker

Table 9.6: Laser tracker errors from the manufacturer's specifications, assuming 1m distance, as used by DESY Geodesy Group, and as measured with a simple experiment.

	Specifications (1σ)	DESY Group	Experiment
Distance d	2.5μm	10μm	3μm
Elevation (pitch) θ_P	6μrad	0.3mgon (4.7μrad)	7.4×10^{-4}deg (12.9μrad)
Azimuth (yaw) θ_Y	6μrad	0.3mgon (4.7μrad)	8.6×10^{-4}deg (15.0μrad)

(figure 9.16) which remained stationary throughout. To make a measurement of the position of the platform, and hence the position of the unit, a laser tracker positioned one metre away was set to follow and measure the position of a retro reflector. This reflector was then moved from the central mount to the bottom-right, bottom-left and top mounts before being returned to the central mount again. From these five readings the position and orientation of the platform can be calculated with some redundancy.

9.6.1 Laser Tracker Performance

The laser tracker used was a Leica LTD 800 [60] in interferometric mode (IFM) and using double layer measurements[9]. The laser tracker gives three values per measurement; the azimuthal/yaw angle, θ_Y, (around vertical), the angle of elevation/zenith/pitch, θ_P, (around horizontal), and the distance (table 9.6). This information is sufficient to produce the location of a retro-reflector relative to the tracker in the tracker's frame.

The exact errors on the measurements made by the laser tracker are not trivial to find. Table 9.6 gives three appropriate choices. The datasheet values used are the 1σ repeatability numbers as any systematic offset will be absorbed into the laser tracker position in the minimisation. A distance of one metre has been assumed for the angular precision. The DESY group values are based on the experience of the group [61]. Finally, a simple experiment was conducted where the platform was rigidly mounted on a wall and repeatedly measured with the tracker. The standard deviations of the measurements were calculated and entered into the table. These errors will be important later in determining the weightings assigned to their measurements during minimisation.

---

[9] IFM mode starts the retro off at a known location and counts interference peaks as it is moved without breaking the beam to maintain sub-micron precision. A double layer measurement involves automatically measuring a second time after a rotation of the laser tracker head by 180 degrees in the azimuthal direction and correcting the zenith. This removes systematic errors from the encoders and leads to greater accuracy.

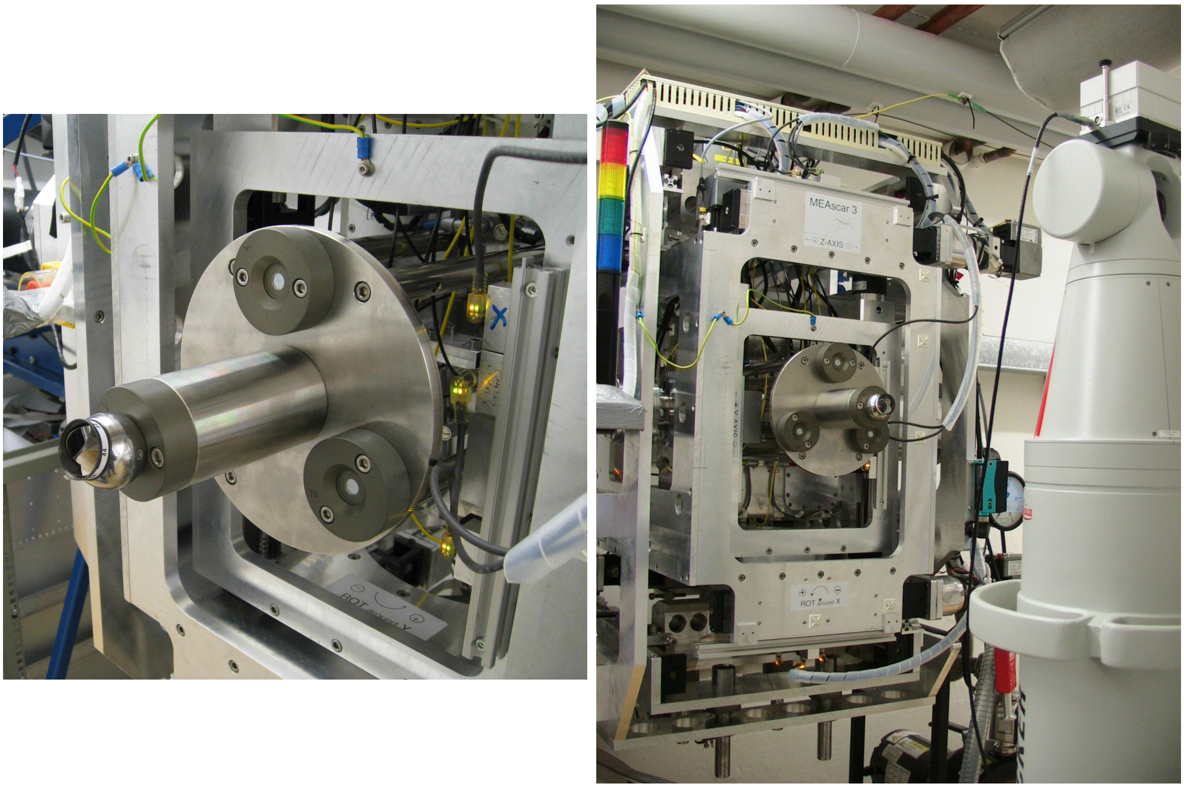


Figure 9.16: Calibration Platform with three of the four retro-reflector nests visible with the retro-reflector sited in marker 4 (left). Laser tracker observing unit. (right)

9.6.2 Calibration Platform

The calibration platform is designed to be stable during the calibration and to have a large enough distance between the retro nests to enable laser tracker measurements to produce a precise enough reconstruction. As can be seen in figure 9.16, it consists of an invar disc, mounted to the side of the measurement unit by four 200mm invar rods. On the plate are positioned three retro-reflector nests in an equilateral triangle, each 70mm from the centre. A fourth retro nest is 99mm further out, mounted on an invar pillar. With the laser tracker about one metre away, the uncertainty ellipsoid of each marker reading has major axes of $3 \times 13 \times 15$ microns ($x'y'z'$), with its axes roughly in line with the lab system.

9.6.3 Platform Data Checks

The laser tracker data taken during the calibration experiment needed to be checked for anomalies and that it does not deviate much from the expected performance. The laser tracker readings for each marker can be transformed into Cartesian co-ordinates and checked for inconsistencies. The '+' markers in figure 9.18 of section 9.6.8 shows the data transformed into a single, common frame allowing the range of motion to be seen. The frame was defined by figure 9.17, with the

Table 9.7: Standard deviations in microns of the distances between platform markers for the car 1 calibration dataset. Marker zero is the first measurement of marker four.

1	**2**	**3**	**4**	**Marker**
2.68	2.42	1.48	1.33	**0(4)**
	2.58	1.89	1.78	**1**
		1.34	2.30	**2**
			1.89	**3**

arbitrary x-offset chosen to be a convenient value representative of the distance to the centre of the unit (0.32m). The points show the fairly even coverage of points over the legal volume. The greater freedom of motion along and around the Z-axis is also apparent.

To check the combination of rigidity of the platform and stability of the laser tracker, the reconstructed marker positions 1, 2 and 3 for each run were used to define the x-y plane of a co-ordinate system for that run, with the origin being the mean of those positions. If there was no tracker error and the platform was perfectly rigid, the points for each run would all lie on top of each other after the transformation. The positions calculated for the markers followed a random walk in the plane with no overall motion, with ranges of between 4μm and 10μm.

Table 9.7 shows the standard deviations of the calculated distances between the markers for the car 1 calibration dataset. The value of around 2μm is significantly better than the expected value of more than 20μm that combining the experimentally derived error ellipse discussed in section 9.6.2 would suggest. This suggests that there may be correlation of errors or that conditions were favourable for measurement.

9.6.4 Modelling Individual Laser Tracker Marker Measurements

It is important to be able to calculate the expected tracker readings for a single marker, given the position of the marker and the laser tracker as well as the orientation of the laser tracker in a global frame. This will enable the decoupling of the positions from the tracker's frame and allow a more convenient axis system to be chosen later; paving the way for a more complete model. The individual marker measurements are modelled as follows:

Consider the laser tracker having a position $\vec{\sigma}$ in the global frame and an orientation $\mathbf{S}$, where $\mathbf{S}$ is a rotation matrix, and a marker point having a position $\vec{\rho}$. The vector $\tilde{\rho} - \tilde{\sigma}$ is then the position of the marker with respect to the tracker. The distance measured is trivially $|\vec{\rho} - \vec{\sigma}|$, though the angles measured by the laser tracker are slightly more involved as it measures angles in its own, primed, frame. The unit vector $\hat{\vec{z}}'$, is defined as the home direction of the tracker which is simply the global $\hat{\vec{z}}$ axis rotated into the the tracker frame. $\hat{\vec{z}}'$ can be rotated into the $\widehat{\vec{\rho}-\vec{\sigma}} \equiv {}^{(\vec{\rho}-\vec{\sigma})}/_{|\vec{\rho}-\vec{\sigma}|}$ direction by a rotation matrix, $\mathbf{R}$. Following the Tait-Bryan YXZ formalism of appendix D.2, the rotation matrix can be described by the pitch, θ_P, and yaw, θ_Y, angles, noting that the roll is immaterial and set to zero.

$$\begin{aligned} \widehat{\vec{\rho}-\vec{\sigma}} &= \mathbf{R_{y'}}(\theta_Y)\mathbf{R_{x'}}(\theta_P)\hat{\vec{z}}' \\ &= \mathbf{R_{y'}}(\theta_Y)\mathbf{R_{x'}}(\theta_P)\mathbf{S}\hat{\vec{z}}. \end{aligned} \tag{9.3}$$

However, the forms of the pitch and yaw matrices are much more convenient in the global frame. The rotation can be removed from the right-hand side of equation 9.3 by noting that a rotation about a primed axis can be achieved by rotating back to the unprimed frame, applying the desired rotation in the unprimed frame, then rotating back to the primed frame:

$$\mathbf{R_{y'}}(\theta) \equiv \mathbf{S}\mathbf{R_{y}}(\theta_Y)\mathbf{S}^{-1} \tag{9.4}$$

$$\mathbf{R_{x'}}(\theta) \equiv \mathbf{S}\mathbf{R_{x}}(\theta_P)\mathbf{S}^{-1}\,. \tag{9.5}$$

This can be used in the opposite manner to transform the rotations to be in the lab frame:

$$\begin{aligned} \mathbf{S}^{-1}\widehat{\vec{\rho}-\vec{\sigma}} &= \mathbf{S^{-1}}\mathbf{R_{y'}}(\theta_Y)\mathbf{S}\mathbf{S^{-1}}\mathbf{R_{x'}}(\theta_P)\mathbf{S}\hat{\vec{z}} \\ &= \mathbf{R_{S^{-1}y'}}(\theta_Y)\mathbf{R_{S^{-1}x'}}(\theta_P)\hat{\vec{z}} \\ &= \mathbf{R_{y}}(\theta_Y)\mathbf{R_{x}}(\theta_P)\hat{\vec{z}}. \end{aligned} \tag{9.6}$$

The right-hand side can be written explicitly by considering the expanded Tait-Bryan matrix of equation D.26 and setting $\mathbf{R_x}(\theta_R)$ to zero:

$$\mathbf{S^{-1}}\widehat{\vec{\rho}-\vec{\sigma}} = \begin{pmatrix} \cos\theta_P \sin\theta_Y \\ -\sin\theta_P \\ \cos\theta_P \cos\theta_Y \end{pmatrix} \tag{9.7}$$

This can be used to give the two angles as measured by the laser tracker, by picking out the elements by applying the dot product with the axis definition vectors, achieved by left-multiplying by them in their row vector form. Including the distance measurement for completeness gives the general equations for the laser tracker measurements[10]:

$$\begin{aligned} d &= |\vec{\rho} - \vec{\sigma}| && (9.8) \\ \sin\theta_P &= -\hat{\vec{y}}\mathbf{S}^{-1}\widehat{\vec{\rho} - \vec{\sigma}} && (9.9) \\ \tan\theta_Y &= \frac{\hat{\vec{x}}\mathbf{S}^{-1}\widehat{\vec{\rho} - \vec{\sigma}}}{\hat{\vec{z}}\mathbf{S}^{-1}\widehat{\vec{\rho} - \vec{\sigma}}} . && (9.10) \end{aligned}$$

9.6.5 Modelling the Calibration Platform

With a method of generating laser tracker readings found, a model of the whole system can be built up. The components of $\vec{\sigma}$ were parametrised as simply:

$$\vec{\sigma} = \begin{pmatrix} l[1] \\ l[2] \\ l[3] \end{pmatrix} . \qquad (9.11)$$

The laser tracker orientation matrix, $\mathbf{S}$, can be parametrised using any reasonable convention; the Tait-Bryan ZYX was chosen for consistency with unit motion:

$$\mathbf{S} = \mathbf{R_z}(l[6])\mathbf{R_y}(l[5])\mathbf{R_x}(l[4])\,. \qquad (9.12)$$

This is equivalent to the intrinsic rotation around the Z-axis, then the new Y′-axis, then the X″-axis. The locations of each of the four markers were denoted by $m[marker\#, axis]$. To aid later analysis and constraint, the concept of a baseplate point was introduced; this is defined to be average of the locations of marker 1, marker 2 and marker 3. While written as input variables, $b[axis]$, they are not independent and will need constraining later. With a set of marker home-positions defined with the help of a baseplate centre point, the whole group can be rotated then translated to a new set of positions in space:

$$\vec{\rho}[i] = \begin{pmatrix} P[1] \\ P[2] \\ P[3] \end{pmatrix} + \mathbf{R_z}(P[6])\mathbf{R_y}(P[5])\mathbf{R_x}(P[4]) \begin{pmatrix} m[i,1] \\ m[i,2] \\ m[i,3] \end{pmatrix} , \qquad (9.13)$$

---

[10] Note that to give the correct, full range, value of the the angle θ_Y, the two argument inverse tangent function should be used.

where i runs from 1 to 4. Equations 9.8, 9.9 and 9.10 can then be used to calculate the tracker readings for a given set of input parameters, giving a set of observables $d[1..4]$, $p[1..4]$ and $y[1..4]$ which are the distance, pitch and yaw values for each marker.

During a single calibration experiment the values of l, b and m will not change; this restricts the laser tracker to be stationary and the markers to remain rigidly fixed in a non-deforming baseplate, which itself moves freely. Therefore these are the calibration constants of the tracker-platform system. However, there are many different orientations the LSM unit will be set to, so $P[1..6]$ is extended to $P[1..n, 1..6]$. These are the dynamic parameters, or reconstructed co-ordinates. Similarly, the observables are extended to $d[1..n, 1..4]$, $p[1..n, 1..4]$ and $y[1..n, 1..4]$. This expression can then generate expected laser tracker readings for many runs at the same time for a given laser tracker position and orientation and marker home positions, with the baseplate home position arising from the others. The final expression has the form $\mathbf{L} = \mathbf{F}(\mathbf{X})$:

$$\begin{pmatrix} distance[run\,1, marker\,1] \\ d[1,2] \\ d[1,3] \\ d[1,4] \\ pitch[run\,1, marker\,1] \\ p[1,2] \\ p[1,3] \\ p[1,4] \\ yaw[run\,1, marker\,1] \\ y[1,2] \\ y[1,3] \\ y[1,4] \\ d[2,1] \\ \vdots \\ y[50,4] \end{pmatrix} = \mathbf{F} \begin{pmatrix} laser\,tracker\,location[x]\ (l[1]) \\ l[2] \\ l[3] \\ laser\,tracker\,rotation[rx]\ (l[4]) \\ l[5] \\ l[6] \\ calibration\,platform\,base\,home\,location[x]\ (b[1]) \\ b[2] \\ b[3] \\ marker\,home\,location[marker\,1, x]\ (m[1,1]) \\ m[1,2] \\ m[1,3] \\ \vdots \\ m[4,3] \\ Platform\,location[run\,1, x]\ (P[1,1]) \\ P[1,2] \\ P[1,3] \\ Platform\,rotation[run\,1, rx]\ (P[1,4]) \\ P[1,5] \\ P[1,6] \\ P[2,1] \\ \vdots \\ P[50,6] \end{pmatrix}, \tag{9.14}$$

where $\mathbf{F}$ is the (very large) functional form of the model and is not expressible as a matrix. The symbol '$\vdots$' has been used to indicate elided entries.

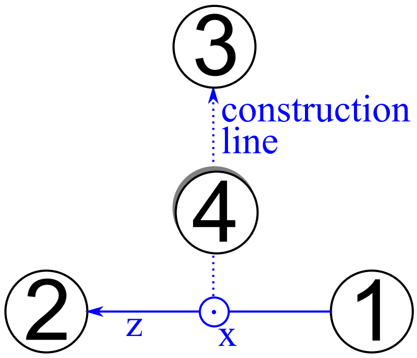


Figure 9.17: Axis system definition

9.6.6 Platform Reconstruction

Platform reconstruction is now simply finding the parameters of the model described in equation 9.14 which produce observables that best fit the laser tracker data. This can be achieved with a non-linear least-squares minimisation. Appendix B covers the theory behind least-squares minimisation, with section B.5 covering a steepest descent minimisation method for solving a non-linear problem. An implementation of this method using analytical derivatives, constraints, SVD over-constraint removal, Gaussian elimination solution, and a backtracking line-search, was manually implemented in Mathematica [54]. The input errors on the laser tracker data were taken from the three paradigms of table 9.6, meaning that there were three possible options to choose from.

9.6.7 Symmetries and Constraints

There are a number of symmetries of the system which need to be constrained to enable successful minimisation:

- There are 3 degrees of freedom used to define the axis directions
- There are 3 degrees of freedom used to define an origin.
- The LSM unit is rotated around an unknown origin and translated along an unknown axis system. This gives another 6 degrees of freedom.
- The location of the calibration platform baseplate centre relative to the centre of the LSM unit is not known well. With only laser tracker data available, allowing the position to vary introduces a symmetry analogous to the LSM origin symmetry. This means arbitrary fixed values have to be used for platform-only reconstruction.

- There is the potential for a length scale-factor freedom, which is often used in networks, but the model does not require it.

As noted in appendix B.3.2, the system is not solvable while under-constrained. A co-ordinate system must be defined from our measurements to produce six constraints. When a unit has been calibrated, the constants produced are expressed in this co-ordinate system and hence so will the reconstructions made in use. Therefore, it makes sense to choose a coordinate system that will give reconstruction values that are easy to interpret. The coordinate system chosen was one with the the origin at the centre of the unit with orthogonal, right-handed axes.

To define the axis directions, consider the platform as seen in figure 9.16 with the marker nests labelled according to the schematic in figure 9.17. Over the course of the calibration, the marker points move over a fairly small volume of 0.5 × 1.5 × 2.5cm (figure 9.18). The relative differences between average positions of the markers over the whole calibration were used to define the directions. The axis labels are not critical, though for consistency with other work, the Y-axis was chosen to be roughly vertical, with the Z-axis roughly along the tunnel and the X-axis roughly away from the tunnel wall. The Z-axis was defined first using markers 1 and 2. A Y-like construction line was then found connecting the midpoint of markers 1 and 2 with marker 3. The vector product of this construction line is then found to give an X-axis which is orthogonal to the Z-axis. The Y-axis is then simply the vector product of the Z-axis and the X-axis:

$$\begin{aligned}
\hat{\mathbf{x}} &= \left(\bar{\mathbf{M}}_3 - \frac{(\bar{\mathbf{M}}_1 + \bar{\mathbf{M}}_2)}{2}\right)^{\wedge} \wedge \hat{\mathbf{z}}, && (9.15)\\
\hat{\mathbf{y}} &= \hat{\mathbf{z}} \wedge \hat{\mathbf{x}}\,, && (9.16)\\
\hat{\mathbf{z}} &= \widehat{\bar{\mathbf{M}}_2 - \bar{\mathbf{M}}_1}\,, && (9.17)
\end{aligned}$$

where $\bar{\mathbf{M}}_i$ is the vector describing the mean location in the global frame of the retro-reflector over all runs when placed in nest i. In principle, equations 9.15 - 9.17 yield nine constraints. However, as the axis system can be considered as a reference system with a 3D rotation applied, there must only be three non-redundant equations. Some equations are trivial to exclude - as the right-hand sides of the equations are normalised and the left-hand sides are the respective unit vectors, three equations are simply $1 = 1$. The remaining redundancy can be removed inside the minimisation with the singular value decomposition method described in appendix B.3.4.

In addition to setting the axis directions, the position of the origin needs to be defined. A simple constraint is to take the average position of markers 1, 2 & 3 and fix it to be a defined point in space. This is defined to be the baseplate location, $b[1..3]$. If the platform alone is being reconstructed, this point is arbitrary, though it is convenient to use the design drawings to give an origin at the centre of symmetry of the attached LSM unit. Later, when camera data is available, this common centre can be allowed to float in the minimisation.

Having defined the axis system and the origin, there is still a translational symmetry and rotational symmetry where the unit and the laser tracker move in step. This can be constrained by fixing one of the objects - the platform average position and associated rotation angles were set to be zero. Figure 9.26 contains a graphical representation of these symmetry-breaking constraints, as well further constraints added later to break near-symmetries.

9.6.8 Platform Reconstruction Results

The platform reconstruction method worked successfully, and was used to produce three sets of results: one for each error paradigm. The minimisation typically took 5-7 iterations to complete when given rough estimates of initial parameter values, though took roughly 20 minutes to complete and used several GB of memory, caused by the calculation of analytical derivatives.

Although the output of the model is laser tracker readings and the model parameters are as defined in equation 9.14, it can easily be interrogated to produce marker locations. Figure 9.18 shows the resultant locations when using the 'Specification' error paradigm, overlaid on the values found by individually transforming the raw data to the same plat. There is a close agreement; the points are indistinguishable in the plot and the locations themselves fit with expectations.

Table 9.8 shows the values of the tracker-platform calibration constants, the error estimates on those values, and the change in the constants when changing to the DESY and experimentally derived measurement error values. The table shows that changing the measurement error estimates significantly changes the model parameter error estimate, but not the value of the parameter itself. Note that the m[1,1], m[2,1] and m[3,1] errors are very small due to the constraints that fix them to define the origin via the baseplate: the baseplate parameters, b[1], b[2] and b[3] were constrained to their values which were found in a full calibration run later. Their values are also identical for this reason.

Of course, the static parameters (tracker-platform calibration constants), of the platform and tracker are of secondary importance. Figure 9.9 shows the effect on the 'reconstruction parameters' that vary from run to run: the platform position (P[i,1], P[i,2] & P[i,3]) and orientation (P[i,4],

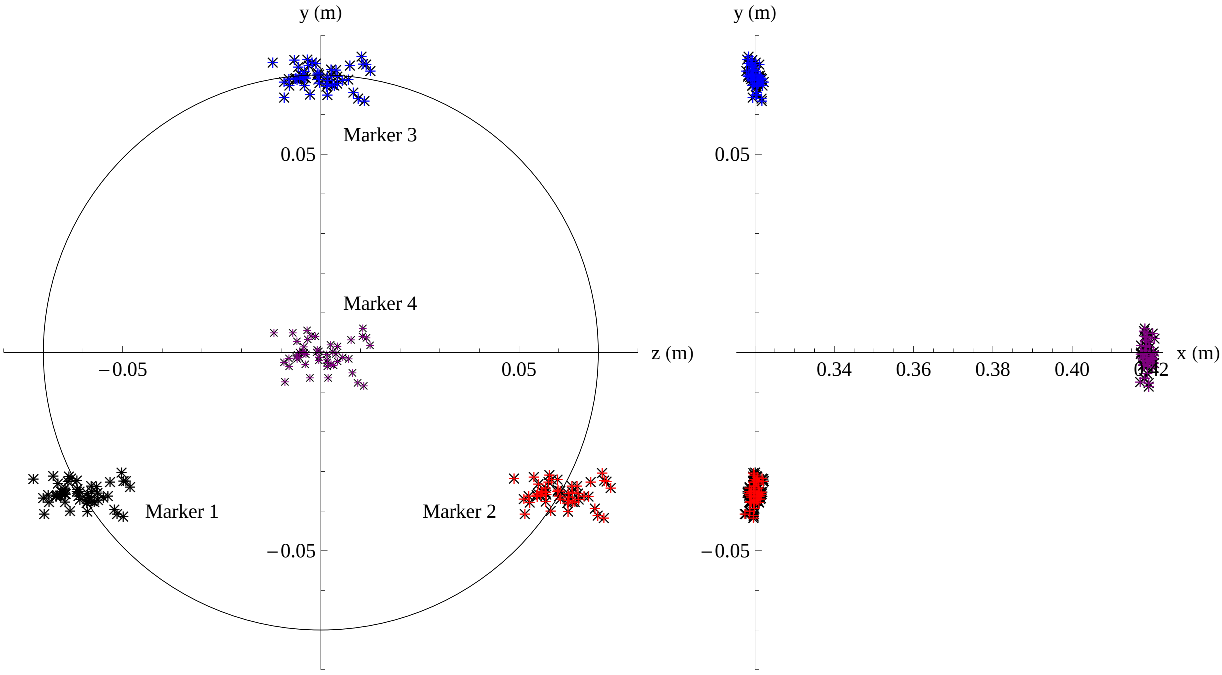


Figure 9.18: Platform reconstruction showing the values produced by the reconstruction ('×'s) as well as the individual marker locations transformed into the 'average' car frame ('+'s). The left-hand plot shows the Y-Z plane, facing *away* from the wall (along the X-axis) with a reference 70mm circle shown. The right-hand plot shows the y-x plane against the Z-axis.

P[i,5] & P[i,6]). Again, there is significant variation in error estimate, but there is also large change in the parameter values themselves. The RMS of the change in the 50 sets of parameters when moving from one error paradigm to another is up to six microns and seventeen microradians.

With the platform reconstructed, the residual of the data and the calculated observables from the model can be found and checked for consistency. Table 9.10 shows the RMS of the resultant residuals for the three error paradigms discussed in section 9.6.1, with the ratio of the error estimate and the RMS residual given. The χ^2/DoF from the minimisation is also given. The table shows that the input errors appear pessimistic, particularly so in the case of the experimentally derived values. Note that for the reconstruction parameters, the RMS residual should be less than the underlying measurement error: part of the measurement residual from the truth can be absorbed into a contribution to the parameter value. e.g. if for a single run the tracker, positioned roughly along the X-axis, happened to report distances that were smaller than the truth, the reconstruction for that run would produce the unit further out than reality but the residuals would be reduced. A simple simulation showed about a 20% reduction in expected RMS residuals from this effect for the system. The ratios of the RMS residuals to the input measurement errors are broadly the same for the different observables in each individual error paradigm. A large variation would be evidence for unrepresentative input errors, but all paradigms appear acceptable.

Table 9.8: Effect of changing input measurement errors on run-independent elements of **X**. The constants found are given in metres apart for the orientation components of the laser tracker which are in radians. All other values are in microns/microradians. The Error columns contain the calculated error estimate from the minimisation. The other columns show the small difference in calibration constants that appears when changing error paradigm; all are much less than the errors. (*=Parameter fixed.)

Parameter	Spec. errors (S)		DESY errors (D)		Experimental errors (E)		
	Value	Error	Error	D-S	Error	E-S	E-D
	(m)	(μm)	(μm)	(μm)	(μm)	(μm)	(μm)
l[1]	0.871136	0.85	3.2	0.1	1.1	-0.001	-0.1
l[2]	0.193274	2.4	8.8	-0.3	3.3	-0.002	0.3
l[3]	0.017872	2.4	9.6	-0.2	3.0	0.006	0.2
l[4] (rad)	0.002825	6.6	8.9	0.1	13.8	-0.009	-0.2
l[5] (rad)	2.61240	4.5	15.2	0.3	6.9	-0.01	-0.3
l[6] (rad)	-0.007247	5.4	16.4	-0.4	9.2	-0.01	0.4
b[1]*	0.319992	-	-	-	-	-	-
b[2]*	-0.000248568	-	-	-	-	-	-
b[3]*	0	-	-	-	-	-	-
m[1,1]	0.3200040	0.009	0.011	0.007	0.014	-0.002	-0.009
m[1,2]	-0.0352644	0.20	0.22	0.002	0.41	-0.0003	-0.002
m[1,3]	-0.0605912	0.43	0.45	0.003	1.03	0.003	-0.0001
m[2,1]	0.3200040	0.009	0.011	0.007	0.014	-0.002	-0.009
m[2,2]	-0.0352647	0.20	0.22	0.002	0.41	-0.00009	-0.002
m[2,3]	0.0606604	0.43	0.41	0.01	1.04	-0.005	-0.02
m[3,1]	0.3200040	0.009	0.011	0.006	0.013	-0.002	-0.009
m[3,2]	0.0697834	0.39	0.44	-0.004	0.83	0.0004	0.004
m[3,3]	-0.0000692	0.55	0.66	-0.01	1.26	0.002	0.02
m[4,1]	0.419014	0.42	1.51	0.05	0.62	-0.02	-0.07
m[4,2]	-0.0002541	0.64	1.77	-0.04	1.14	0.007	0.05
m[4,3]	-0.0000861	0.64	1.61	-0.03	1.30	-0.002	0.02

Table 9.9: Means of the reconstruction parameter error estimate and the RMS of the difference in reconstructed values, with different input measurement error schemes.

Parameter	Spec. (S)	DESY (D)		Experimental (E)		
	$\overline{\sigma_{P[-,k]}}$	$\overline{\sigma_{P[-,k]}}$	$rms(\mathrm{D}-\mathrm{S})$	$\overline{\sigma_{P[-,k]}}$	$rms(\mathrm{E}-\mathrm{S})$	$rms(\mathrm{E}-\mathrm{D})$
P[i,1] (μm)	1.3	4.4	1.3	1.9	0.05	1.3
P[i,2] (μm)	8.2	9.8	4.3	12	1.7	6.0
P[i,3] (μm)	8.2	8.4	3.6	15	2.4	6.0
P[i,4] (μrad)	28	23	4.6	62	3.0	7.2
P[i,5] (μrad)	23	24	10	39	6.9	17
P[i,6] (μrad)	23	28	12	32	5.0	17

A minimisation was attempted using the RMS residuals from the experimental reconstruction as the input errors to see if it gave the same values as the output, as if a steady state was produced then it would suggest the relative weightings were correct. However, this was not the case as the minimisation gave very uneven RMS residual distributions. There is potential in using this method to find balanced error estimate ratios for the different readings but this was not investigated further.

Overall, the platform reconstruction process was found to perform well and produce reasonable, self-consistent results. The minimisation process itself was robust and had plenty of potential for extension to cover more complex problems.

9.7 Adjustment of LSM Data Using Platform-Only Results

With the positions and orientations of the unit as reconstructed using laser tracker data known, we can insert CCDs into the model and generate the CCD hit position values that would be found if the tracker reconstruction was correct and the locations the CCDs are inserted are accurate. This can then be compared with the real data recorded by the cameras and used to further correct the LSM CCD data and model, including improving the jitter correction using the car 2 data.

9.7.1 Modelling of CCDs

The modelling of the beam spot positions on the CCDs has been discussed in section 8. The calibration process for a single unit isolates the other cars from the system by making corrections to the data. This makes the model simple as it only needs to contain a single unit with an immobile laser beam. Furthermore, fixing the return beam to be anti-collinear means the model

Table 9.10: Input measurement error schemes and RMS residuals of the modelled laser tracker readings. The chi-squared per degree of freedom is also given. (There were 294 degrees of freedom in the minimisation.)

	Spec.			DESY			Experimental		
Measurement	Input	RMS res.	Ratio	Input	RMS res.	Ratio	Input	RMS res.	Ratio
d[1] (μm)	2.5	0.66	0.26	10	1.51	0.15	3	0.36	0.12
d[2] (μm)	2.5	0.79	0.32	10	2.35	0.23	3	0.38	0.13
d[3] (μm)	2.5	0.65	0.26	10	2.13	0.21	3	0.29	0.10
d[4] (μm)	2.5	0.64	0.26	10	1.69	0.17	3	0.57	0.19
p[1] (μrad)	6	1.93	0.32	4.7	1.70	0.36	12.9	1.96	0.15
p[2] (μrad)	6	1.96	0.33	4.7	1.91	0.40	12.9	1.89	0.14
p[3] (μrad)	6	2.87	0.48	4.7	2.59	0.55	12.9	2.90	0.22
p[4] (μrad)	6	2.12	0.35	4.7	0.68	0.14	12.9	2.79	0.22
y[1] (μrad)	6	2.50	0.42	4.7	2.44	0.51	15	2.61	0.17
y[2] (μrad)	6	2.80	0.47	4.7	2.17	0.46	15	3.15	0.21
y[3] (μrad)	6	1.73	0.29	4.7	1.75	0.37	15	2.00	0.13
y[4] (μrad)	6	2.07	0.35	4.7	0.67	0.14	15	3.19	0.21
χ^2/DoF		0.244			0.235			0.0591	

can be further simplified by removing the mirror entirely and registering hits on the back of the 'return' CCDs.

Due to the design of the unit and calibration platform, as well as simple pre-alignment in preparation of calibration, the direction of propagation is very close (within 2 milliradians) to the Z-axis defined in figure 9.17. Therefore, the direction of propagation is initially set to be directly along the Z-axis and is later adjusted manually in section 9.7.6.

9.7.2 Potential Issues of Correction Adjustment

There are a number of caveats in this adjustment approach:

1. The true positions of the cameras are not yet known, so the nominal positions from the design specifications must be used, at least for the first iteration of this process.
2. Later, the process can be repeated using a partially calibrated model, but if that model is found using a flawed process, then biases may be locked in.
3. The reconstruction using the platform is not perfect, so the correction amount that gives the best fitting real data may not be the physically correct one.
4. There may be legitimate correlations of the residuals from the real data and the generated data with the beam motion on car 2: for example, both car 1 and car 3 could have a

collective motion. If this were the case, then the model and corrections that give the lowest RMS residuals and no correlation would not be the correct choice.

5. As discussed in section 9.3.5, if there is beam motion the projection is less likely to give an accurate result. Therefore a calculated adjustment to the correction may not have quite the right value when actually applied before the data is projected.

The impact of point 1 is that the improvements to the CCD data may not be the best possible, but if they are sufficient to allow a successful initial calibration, then the process can be repeated, mitigating it. Point 2 is only a minor effect as the differences to the corrections when changing the model are small, though not negligible compared to the magnitude of the corrections themselves. The risk from point 3 will lessen as the number of runs increases, though if there is a systematic bias in the platform reconstruction using laser tracker data this will persist. However, this will only ever be a minor effect as the largest source of uncertainty is the beam jitter, which is unrelated to the unit 1 motion. The impact of point 4 is likely to be minimal as there is no particular reason for correlated inter-car motion, particularly as car 2 and car 3 are fixed for the duration of the experiment.

9.7.3 CCD Scale Factors

Upon the insertion of a nominal unit into the model and the generation of expected beam spot positions on the CCDs, an issue immediately became apparent: the standard deviation of residuals in the CCD's X-axes were far higher (approximately 50 microns) than those along the Y-axes (20 microns). In addition, there was a strong correlation (with a modulus of approximately 0.85) between the CCD X-residual and the reconstructed platform x-location for each CCD.

This was caused by a scale error in the reported CCD beam X-position; there is no other reasonable combination of possible factors that could give the same effect. The source of this error is not understood, but introducing a scalar multiplier to the calculated CCD X-hit position and minimising the standard deviation of residuals with regard to this parameter gave minima with similar RMS residuals to the y-axes. Similarly, the value of the scale factor with zero correlation between the residual and platform X-position was also found. Individual minimisations were performed for each camera; table 9.11 shows the effects of introducing this scale factor and demonstrates that both the minimum residual standard deviation, and zero correlation approaches lead to the same scale factors. The process was also performed for the Y-axis though there was negligible improvement.

Table 9.11: RMS residuals, and correlations of residuals with the platform location and the effect of scaling the camera model estimates.

	CCD0	CCD1	CCD2	CCD3
RMS X-residual (μm)	51.9	38.5	58.3	49.6
X-residual correlation with platform-x	-0.90	0.80	0.86	-0.86
RMS Y-residual (μm)	15.7	19.5	19.6	18.9
Y-residual correlation with platform-y	0.22	0.06	0.10	-0.05
minimum scaled X-residual (μm)	22.4	19.7	21.4	25.2
minimum residual model X-scale factor	1.048	1.036	1.046	1.044
zero-correlation model X-scale factor	1.048	1.036	1.046	1.044
minimum scaled Y-residual (μm)	15.3	19.4	19.6	18.9
minimum residual model Y-scale factor	0.995	0.999	1.001	0.998
zero-correlation model Y-scale factor	0.995	0.999	1.001	0.999

With the need apparent, a common X-scale factor was added to the model as individual camera scale factors are unphysical. For the remainder of the analysis using platform data only, the X-scale value was set at 1.043. A Y-scale factor was included into the model but fixed at 1. It is possible for these common factors to be included in the full calibration later, which was done. This is not the case for the beam motion correction and the beam propagation direction correction in sections 9.7.4 and 9.7.6 as they are extra factors in the experiment rather than properties of the simplified 1-car system. However, the CCD X-scale factor is serious enough to need finding to a good approximation before proceeding with other corrections. The scale factor found at this point was sufficient and in fact was within 0.01 of the final value.

9.7.4 Adjustment of Beam Motion CCD Correction

With the x-scaling issue minimised, the residuals of the real data with the values generated inserting CCDs into the platform reconstructions can be analysed further. Figure 9.19 shows these residuals, offset for clarity. The X-residuals still showed significant motion correlated with the run number: the beam X-positions for CCD 1 and CCD 2 both rose by about 30 microns until around run 25. This is similar to the motion of the beam on car 2, suggesting the correction had not worked entirely (figure 9.7). This was partly due to the car 2 data not having the X-scale factor of section 9.7.3 applied, which was done from this point onwards, but also potentially due the centre of rotation of the laser beam not being in the assumed position of the beam launch.

A simple way to check the effectiveness of the beam motion correction is to start with uncorrected car 1 CCD data and add the car 2, CCD 3 beam spot position deviation multiplied by a simple scale factor, α, rather than calculating the factor using geometric considerations.

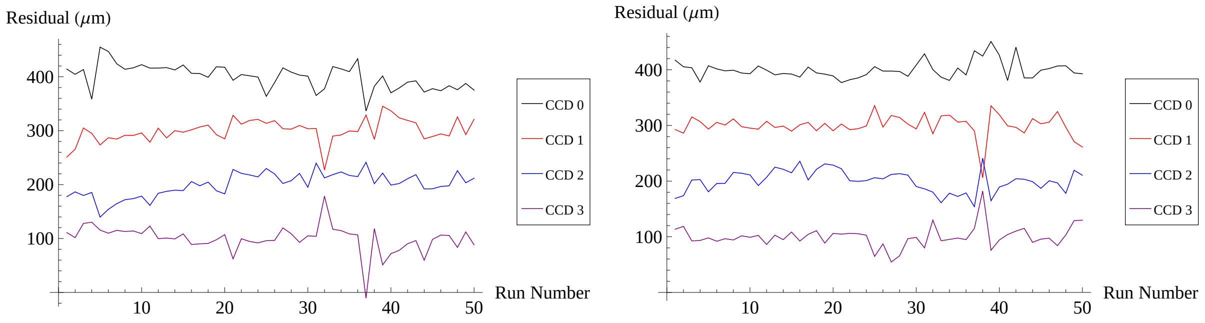


Figure 9.19: X (left) and Y (right) residuals for inserted model cameras against run number. The values have been offset by multiples of 100 to enable clearer comparison.

Under the sound assumption that car 2 is stationary, this will cancel out angular beam motion. (Additionally, there is a parity factor required due to the relative orientations of the cameras; see table 8.1.) The return correction is applied as normal as the source of that motion is known to be the mirror. The RMS residuals of the real data with this correction against the generated data using inserted CCDs into the platform reconstructed values can then be found for different values of α. Similarly, the correlation of the residuals with the car 2, CCD 3 beam spot position deviation can be calculated for different values of α. If the beam jitter is caused by rotation centred in car 3 then the values of α that give the minimum RMS residuals and zero correlation with the car 2, CCD 3 data should be around -2 due to the relative distances of the units, with minor differences of about ± 0.05 due to the different positions and hence optical distances of the virtual CCDs.

The top four plots of figure 9.20 show the results of this procedure, with α ranging from minus four to zero. There is a clear improvement in the RMS residuals of up to 15 microns and a reduction in correlation as α approaches 2, though there is a significant spread in the minimum RMS α values. The variation of the minimum RMS values of α vary from 1.14 to 2.37, far beyond the expected physical range (table 9.12).

The procedure was repeated using data that has had the geometric corrections of section 9.3.2 applied. This time, if the correction is optimal, the minima of the RMS residuals should occur with an α of zero. This should also be the value giving zero correlation with the car 2, CCD 3 deviation. The pattern seen above persists, with a large spread of values of α, particularly in the X-axis.

A different correction factor for the X-axis and Y-axis is reasonable, but a large variation between cameras is unphysical. This variation is primarily due to using a nominal model for the

Table 9.12: The effect of beam spot position correction on RMS residuals of CCD data with inserted cameras. The minimisation used specification laser tracker errors and 1.25 mm/m beam propagation X component, ignoring runs 32, 37, 38 & 39. The cameras were inserted at the nominal model positions.

CCD number and axis	0 X	1 X	2 X	3 X	0 Y	1 Y	2 Y	3 Y
Uncorrected RMS residual (μm)	33.9	24.8	20.9	31.7	24.6	29.8	27.5	29.9
after applying best α-correction (μm)	19.4	9.0	11.8	14.7	12.9	12.9	17.2	13.8
associated α value	-1.84	1.52	1.14	-1.86	1.74	2.23	-1.78	-2.20
zero correlation α value	-1.85	1.54	1.14	-1.87	1.86	2.37	-1.91	-2.33
with geometrical correction (μm)	19.5	10.6	16.7	14.7	12.7	13.5	16.9	14.6
after applying best α-correction (μm)	19.5	9.0	12.0	14.7	12.7	12.8	16.9	14.0
associatedα value	0.07	-0.37	-0.77	0.00	-0.04	0.36	0.08	-0.19
zero correlation α value	0.07	-0.37	-0.78	-0.00	-0.04	0.38	-0.10	-0.38

CCD positions and means that using this method to adjust the centre of the laser jitter rotation is not valid. However, the need for such a correction is now clear and with no robust results to suggest deviation, the rotation was set to the beam launch optics position.

9.7.5 Identification of Bad Runs

Figure 9.19 shows that the real CCD data from some runs have significant deviations from the calculated value for multiple cameras. To find runs where the data is in poor agreement the following procedure was undertaken.

- Starting with the platform-only reconstruction results using the experimental error paradigm, CCDs and Tilt sensors were inserted into the model in their nominal positions.
- The residuals of the generated observables against the real motion-corrected measurements were found for each run.
- The residuals were then normalised by subtracting the means and divided by the measurement error estimate. The subtraction is necessary as the CCDs and tilt sensors were not part of the minimisation, leading to biases in the residuals. The tracker error estimates from the experimental paradigm were used, while representative values of 20 microns and 45 microradians were used for the cameras and tilt sensors.
- The sum-square of the normalised residuals of physical and generated observables for each run were taken and plotted. This is similar to the chi-squared contribution of each run to a full minimisation, apart from the mean subtraction.

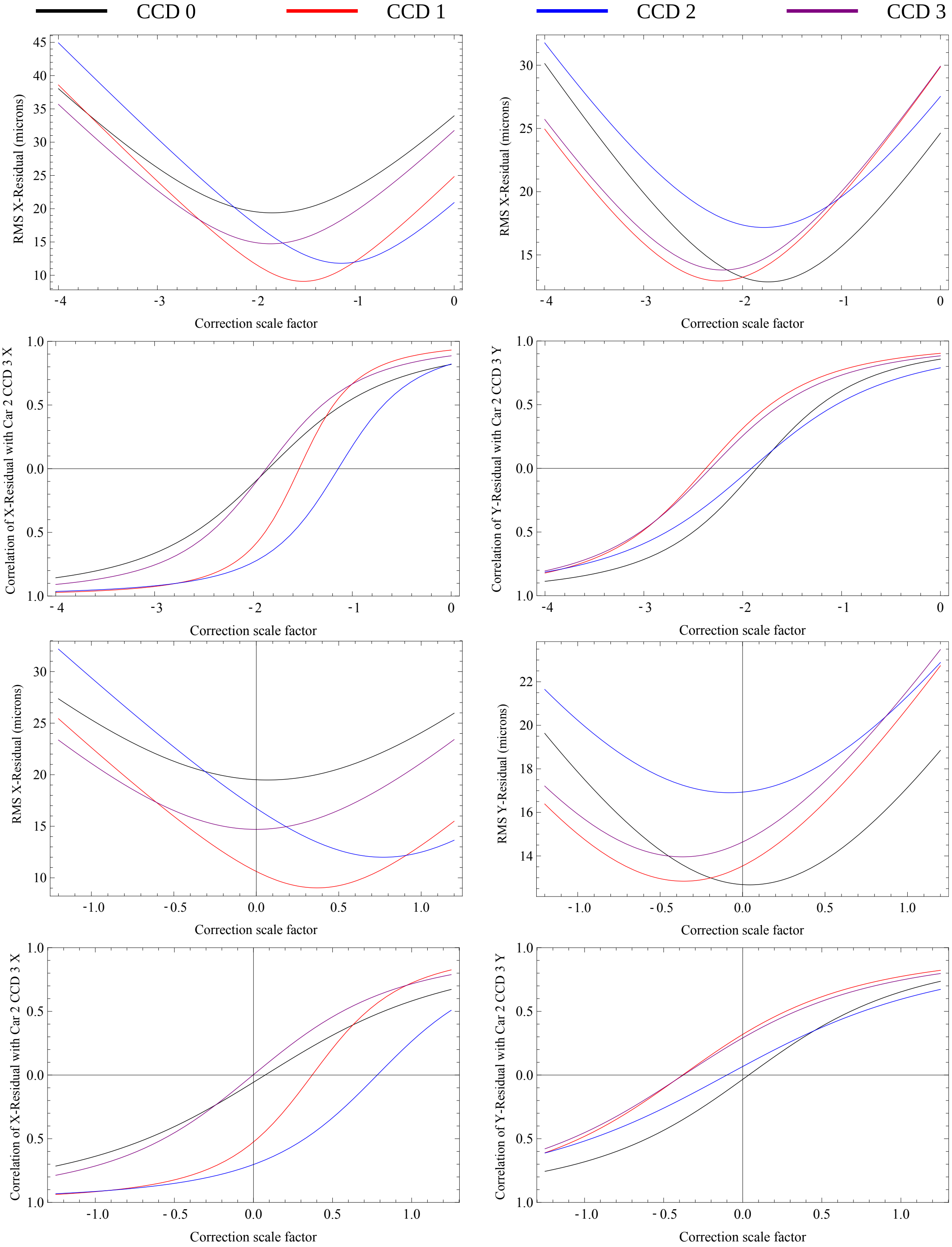


Figure 9.20: Effect on RMS residuals (odd rows) and correlation with car 2 camera 3 (even rows) in the X-axis (left) and Y-axis (right), with uncorrected data (top two rows) and corrected data (bottom two rows), upon the application of scaled correction factor α with a nominal model. The data has been flipped where necessary to remove the effects of camera inversion, making the lines directly comparable.

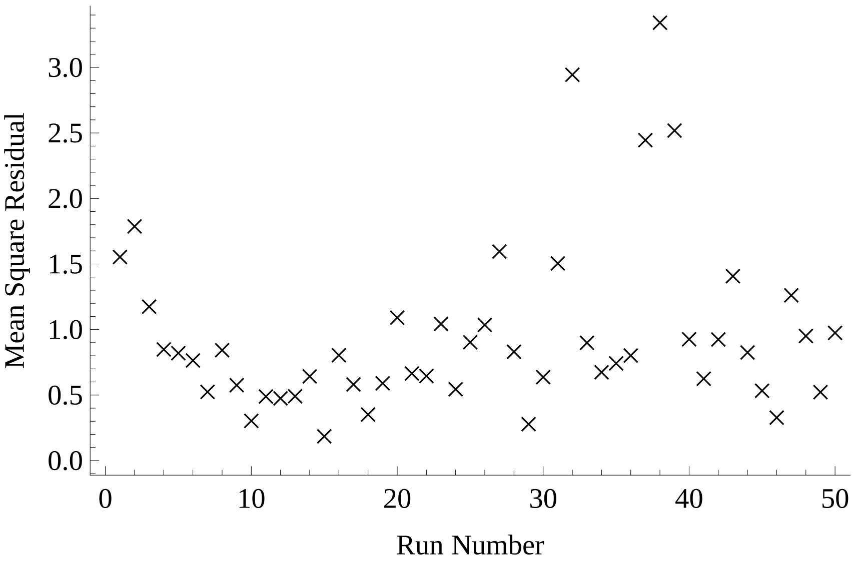


Figure 9.21: Mean square normalised residual of runs for platform-reconstructed data with calculated CCD and tilt sensor readings inserted.

The resultant plot, shown in figure 9.21, clearly shows runs 32, 37, 38 & 39 as anomalous. No common or distinctive feature of the system was found for these runs, so they were marked as bad. Repeating the process for the other two laser tracker error paradigms also gave these runs as stand-outs. Similarly, if the residuals are normalised by the total RMS for each parameter across all runs instead of the prescribed measurement error, which has the effect of reducing the weighting of the CCD residuals, the same four runs still stand out. These anomalous runs were excluded from the input datasets and the minimisations were repeated using the expurgated data. Further work was carried out exclusively on these new, expurgated results.

9.7.6 Adjustment of Modelled Laser Propagation Direction

With the modelled beam set to be paraxial, the residuals of the inserted CCDs show a clear correlation against the motion of the platform in the Z-axis. (See figure 9.22, noting the expected parity from table 8.1 has been applied to make the points comparable.)

Figure 9.23 shows the effect of changing the laser beam propagation direction from $\begin{pmatrix} 0 & 0 & -1 \end{pmatrix}$ to $\begin{pmatrix} x & y & -1 \end{pmatrix}$, where x and y are altered over a range of a few thousandths. The beam modelled is adjusted to ensure there is no resultant translation at the axis system origin. Table 9.13 shows the resultant minimum residuals and associated beam direction component values. These were found by fitting a parabola to the residuals and taking the minimum, and taking a linear fit of the residual-z gradient and finding the X-axis intersection. As expected, the gradient of this linear fit is simply the camera scale value used. (1.043 in X and 1 exactly for y.) Each plot is made

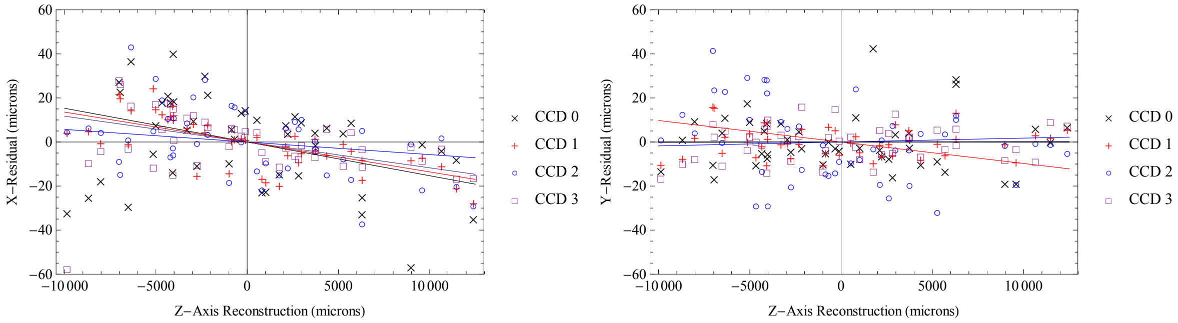


Figure 9.22: Residual correlation with Z-axis motion with a paraxial modelled beam for X-axis residuals (left) and Y-axis residuals (right). Residuals and fits have been inverted where appropriate for comparison and linear fits have been added.

Table 9.13: The minimum residuals and associated laser propagation direction component, as well as the propagation components producing zero correlation.

CCD number and axis	0 X	1 X	2 X	3 X	0 Y	1 Y	2 Y	3 Y
minimum RMS residual (μm)	19.5	10.5	16.7	13.7	12.7	13.4	16.2	14.4
min RMS laser component ($^{\mathrm{mm}}/_{\mathrm{m}}$)	1.27	1.84	1.35	0.53	0.07	0.50	-0.93	0.52
zero-correlation laser component ($^{\mathrm{mm}}/_{\mathrm{m}}$)	1.27	1.83	1.35	0.55	0.08	0.47	-0.91	0.48

at the optimal angle found for the other axis, which was decided by simply taking the mean of the zero correlation values. As can be seen from the table, there is significant variation between the cameras, caused by the low statistics and weak correlation. Averaging the values gives 1247 microradians in X for minimum RMS residual and 1251 microradians for zero-correlation. The corresponding Y-axis values are 42 and 31 microradians. For all future minimisations the values of 1250 microradians and 30 microradians were used.

9.7.7 CCD Residuals With Corrected Data

With a method found for improving the CCD data, discarding bad runs and adjusting the laser propagation direction in the model, it can be repeated for each possible laser tracker error paradigm. Table 9.14 compares the RMS residuals of the CCDs inserted at the nominal model positions with the real, corrected data. For the CCD X-axes, the results are all very similar; while for the Y-axes the DESY error paradigm performs noticeably worse. The results of a simulation, mimicking the procedure with generated data and zero error CCD beam spot measurements, are also given. They show fair agreement, apart from the inner CCDs along the X-axes, suggesting that the CCD beam spot position error is the dominant factor for those measurements. The simulation shows that with ideal CCDs, there is a significant change in the RMS residual as

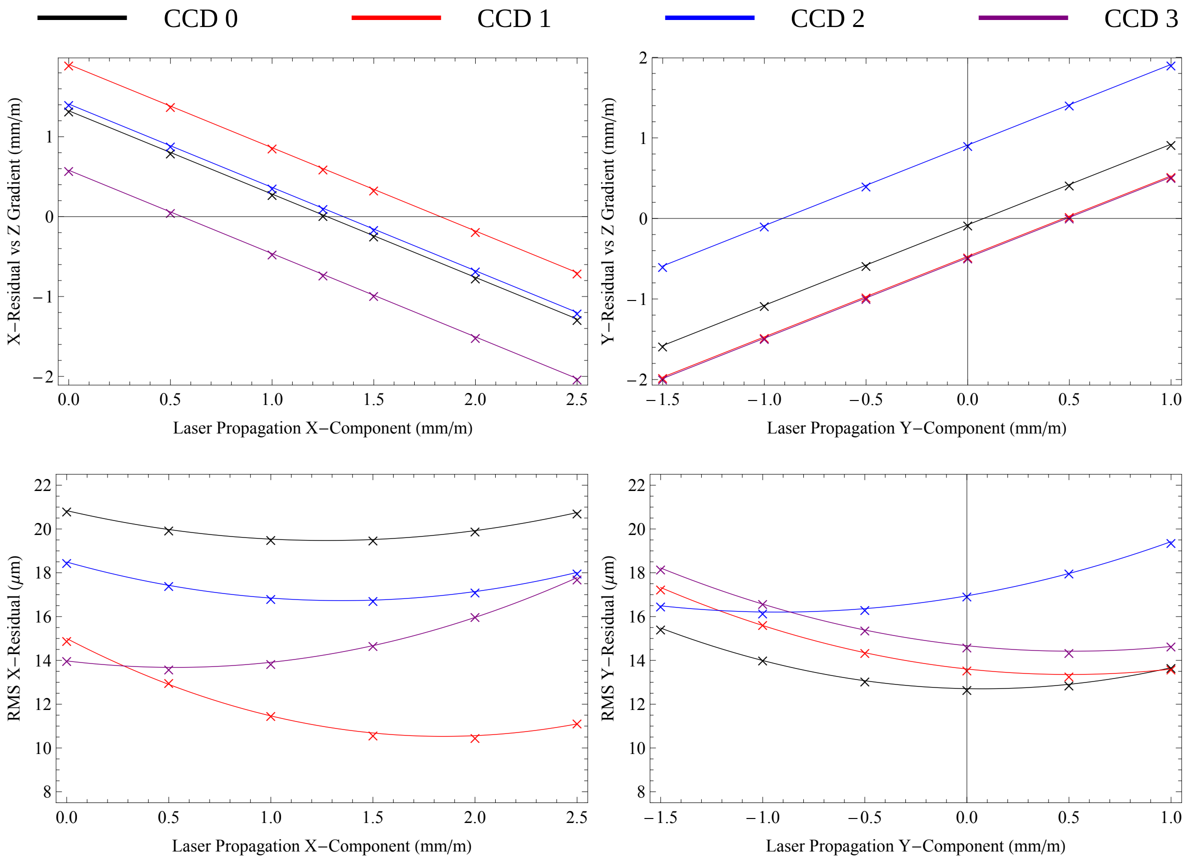


Figure 9.23: The gradient of a linear fit to the calculated X (left) and Y (right) residuals against reconstructed Z position plotted verses the corresponding component of laser propagation (top). The corresponding r.m.s residuals are also included at the bottom. The gradients have been flipped where appropriate to enable comparison and representative functions have been fitted.

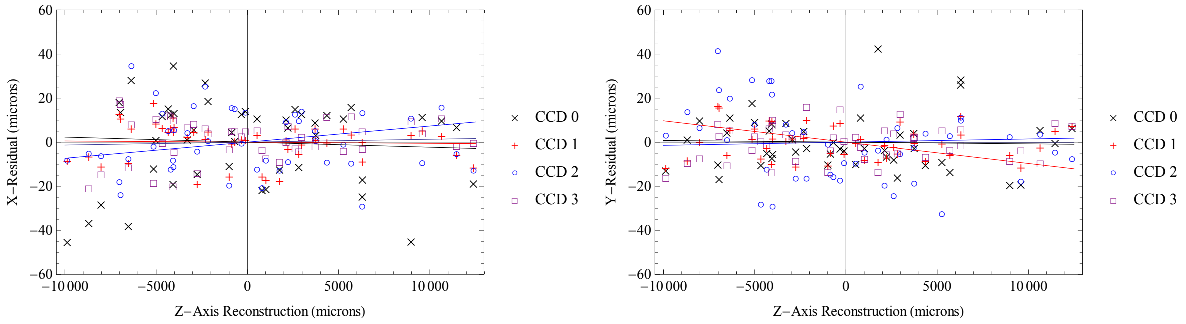


Figure 9.24: Residual correlation with Z-axis motion with an adjusted beam for X-axis residuals (left) and Y-axis residuals (right). Residuals and fits have been inverted where appropriate and linear fits have been added. Correctly modelling the laser propagation direction minimises the residual correlation with Z-axis motion.

Table 9.14: The RMS residuals in microns for model inserted cameras based on reconstructed positions using the three different laser tracker error paradigms after all adjustments and corrections. Simulated values are also given for comparison.

CCD number and axis		0 X	1 X	2 X	3 X	0 Y	1 Y	2 Y	3 Y
Real	Spec	19.5	10.2	16.7	13.9	12.7	9.0	17.0	9.2
	DESY	19.3	10.0	16.9	13.7	14.8	11.0	17.7	11.0
	Experimental	19.5	10.1	16.5	13.9	12.1	8.4	17.0	8.8
Simulation	Spec	1.4	4.4	4.4	1.4	7.6	9.3	8.8	7.6
	DESY	4.6	6.8	6.4	4.6	9.9	11.2	10.0	9.9
	Experimental	2.1	7.1	7.2	2.1	11.6	16.5	15.6	11.6

the error paradigm is altered underlining the importance of having the best weightings in the minimisation.

9.8 Full Calibration

With the ability to reconstruct the platform precisely using the laser tracker measurements demonstrated and the CCD data improved to a sufficient degree, the CCD and tilt sensor data and model can be incorporated into the minimisation step to enable full calibration.

9.8.1 Full Model

The CCD model required is covered in section 9.7.1, with details in section 8, while the tilt sensor model is covered in section 8.6. The final expression uses in the minimisation has the form:

$$
\begin{pmatrix}
\mathit{distance}[\mathit{run}\,1, \mathit{marker}\,1]\ (d[1,1]) \\
d[1,2] \\
d[1,3] \\
d[1,4] \\
\mathit{pitch}[\mathit{run}\,1, \mathit{marker}\,1]\ (p[1,1]) \\
p[1,2] \\
p[1,3] \\
p[1,4] \\
\mathit{yaw}[\mathit{run}\,1, \mathit{marker}\,1]\ (y[1,1]) \\
y[1,2] \\
y[1,3] \\
y[1,4] \\
\mathit{beam\,hit\,pos}[\mathit{run}\,1, \mathit{ccd}\,1, \mathit{axis}\,1]\ (h[1,1,1]) \\
h[1,1,2] \\
h[1,2,1] \\
h[1,2,2] \\
h[1,3,1] \\
h[1,3,2] \\
h[1,4,1] \\
h[1,4,2] \\
d[2,1] \\
\vdots \\
h[50,4,2]
\end{pmatrix}
= \mathbf{F}
\begin{pmatrix}
\mathit{laser\,tracker\,location}_x\ (l[1]) \\
l[2] \\
l[3] \\
\mathit{laser\,tracker\,rotation}_{rx}\ (l[4]) \\
l[5] \\
l[6] \\
\mathit{calibration\,platform\,base\,home\,location}[x]\ (b[1]) \\
b[2] \\
b[3] \\
\mathit{marker\,home\,location}[\mathit{marker}\,1, x]\ (m[1,1]) \\
m[1,2] \\
m[1,3] \\
\vdots \\
m[4,3] \\
\mathit{camera\,location}[\mathit{ccd}\,1, x]\ (c[1,1]) \\
c[1,2] \\
c[1,3] \\
\mathit{camera\,location}[\mathit{ccd}\,1, rx]\ (c[1,4]) \\
\vdots \\
c[4,6] \\
\mathit{camera\,x\,scale\,factor}\ (cs[1]) \\
cs[2] \\
\mathit{tilt\,sensor\,scale\,x}\ (tsx) \\
tsz \\
\mathit{x\,tilt\,sensor\,y\,rotation}\ (tax[1]) \\
\mathit{x\,tilt\,sensor\,z\,rotation}\ (tax[2]) \\
\mathit{z\,tilt\,sensor\,x\,rotation}\ (taz[1]) \\
\mathit{z\,tilt\,sensor\,y\,rotation}\ (taz[2]) \\
\mathit{Platform\,location}[\mathit{run}\,1, x]\ (P[1,1]) \\
P[1,2] \\
P[1,3] \\
\mathit{Platform\,rotation}\ [\mathit{run}\,1, rx]\ (P[1,4]) \\
P[1,5] \\
P[1,6] \\
P[2,1] \\
\vdots \\
P[50,6]
\end{pmatrix},
\tag{9.18}
$$

where **F** is the (very large) functional form of the model. Figure 9.25 gives a graphical representation of part of the design matrix, which is the linearised approximation of **F**, for a particular set of parameter values. Inspection of the rows of the matrix confirms expected behaviour: for example, for the distance measurement to marker 4 for the run 2 ($d[2,4]$), moving the laser tracker in the X-axis ($l[1]$) has a near 1:1 mapping, while the marker home x-location ($m[4,1]$) and platform/unit x-location ($P[2,1]$) have a -1:1 mapping. The effect on $d[2,4]$ for motion of components in the Y-axis is similar but with a smaller magnitude, and the effect of motion in the Z-axis is small. This is consistent with the position of the laser tracker relative to the markers, which was roughly 90cm away in X-axis, 20 cm above in Y-axis, and in-line in the Z-axis (figure 9.16 and table 9.17).

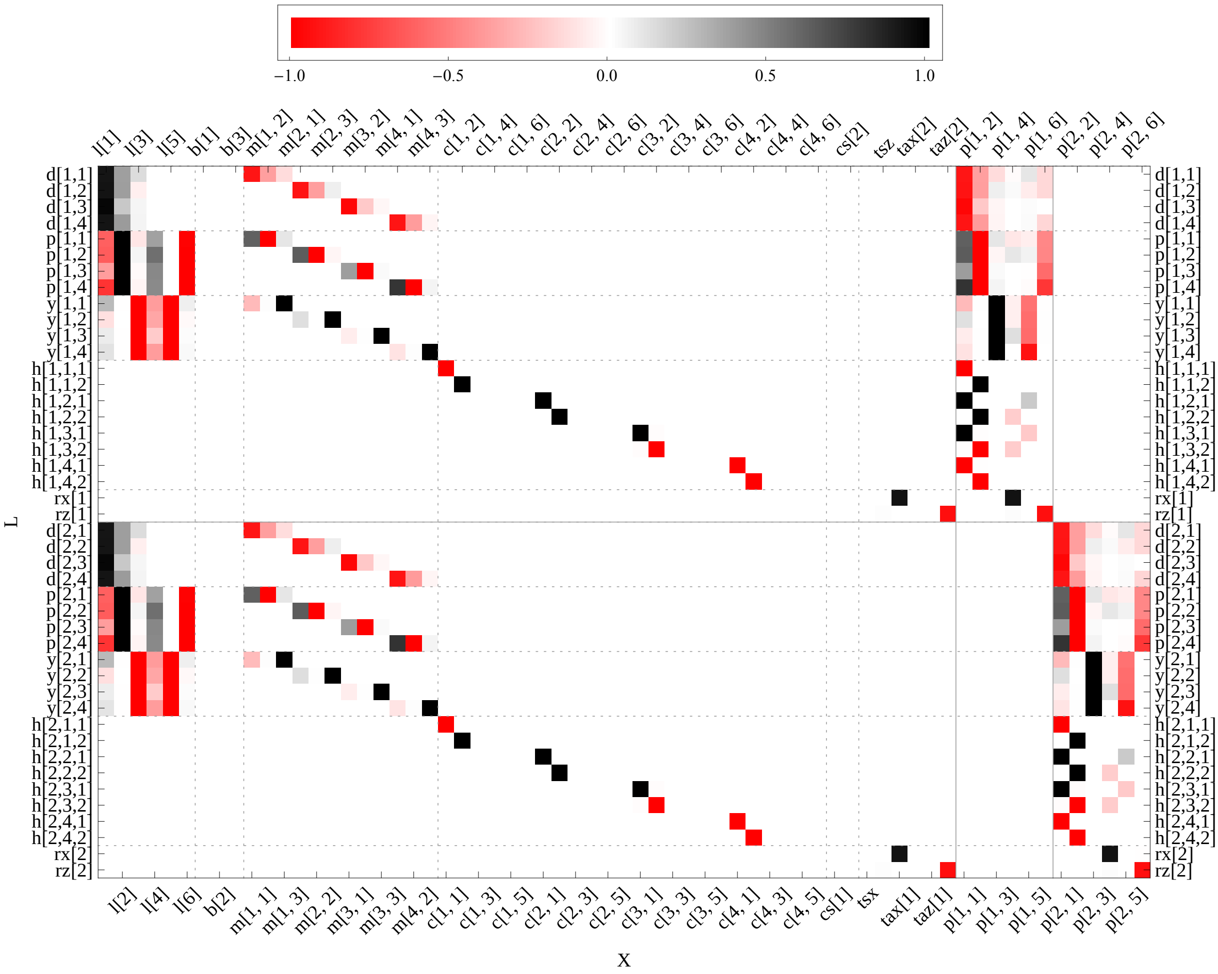


Figure 9.25: Graphical representation of an extract of the design matrix at the minimum for a calibration experiment, describing the behaviour for the first two runs. Element blocks associated with the different runs are partitioned by solid faint lines, while dotted lines divide different classes of measurement. Note that more than two runs are needed to make the system overdetermined, as demonstrated by the matrix width being larger than its height. (Note that some elements are too small to see with the linear scale used, such as the CCD and tilt sensor scale factors, but are still significant.)

9.8.2 Near Symmetries

As well as perfect symmetries like the translational invariance, there are also near symmetries which cause difficulties when trying to minimise. Mathematically, near symmetries manifest themselves as eigenvectors of the C-matrix (see equation B.28) with very small eigenvalues, typically many orders of magnitude smaller than the bulk of the eigenvalues. With near-symmetries, the C-matrix is still invertible, unlike the singular C-matrix found if there are any remaining perfect symmetries, but a large condition number causes inaccuracies when using limited precision numbers. In addition, the error estimate on any **X** element value which is a significant component of a small eigenvector will be so large that the result for that variable will be effectively meaningless. Therefore, the near-symmetry must be broken, either with a constraint or a measurement included into the minimisation.

One example is a rotation of a virtual CCD about a vector in the plane of its face: if the rotation was around the CCD's X-axis by an angle θ, then during a Y-translation, the laser beam would move over the face of the CCD with a scale factor of $1/\cos\theta$, meaning it would give translations that were too large. However, for small values of θ, the scale factor tends to $1+\theta^2$ where $\theta \ll 1$. For example, if the camera were mounted with an angle of 1mrad, this would give a 1ppm residual, which, over the 4mm range of the camera, would be undetectable. While this means that the calibration constant is not needed to a high precision, (a requirement of a 1 micron residual over a the range of the camera would require an angular tolerance of 16mrad,) it will still cause numerical instability in the minimisation process. The rotations of all virtual CCDs were limited to be only around its Z-axis, perpendicular to its face. Figure 9.26 gives a graphical representation of part of the constraint matrix for a calibration.

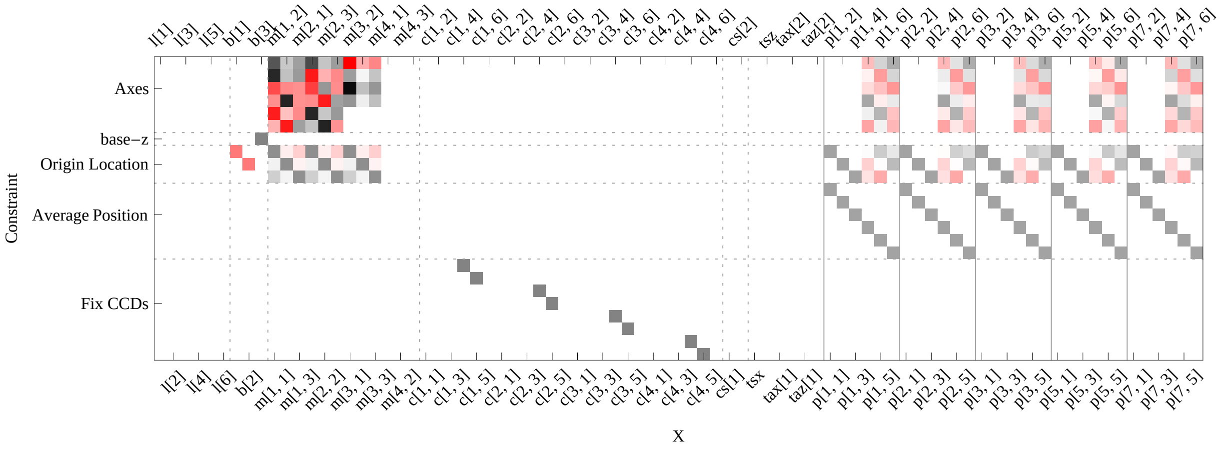


Figure 9.26: Graphical representation of an extract of the constraint matrix at the minimum for a calibration experiment, describing the constraints for the first five runs. Element blocks associated with the different runs are partitioned vertically by solid faint lines, while the dotted vertical lines divide different classes of measurement. The dotted horizontal lines separate different classes of constraint. (Non-linear scale.)

9.8.3 Initial Calibration and Further LSM Data Adjustment

A full calibration was performed using a common quasi-systematic error estimate for the CCD beam positions of 15 microns in both axes, as well as tilt sensor errors of 45 microradians, using the forty six runs that were considered good. These intermediate results are not discussed here but the techniques used on the final results in section 9.9 are also applicable and showed no serious problems. Figure 9.27 shows the mean-square normalised residuals for each run from the calibration. Unlike figure 9.21, the sum of all the runs is the chi-squared value of the minimum.

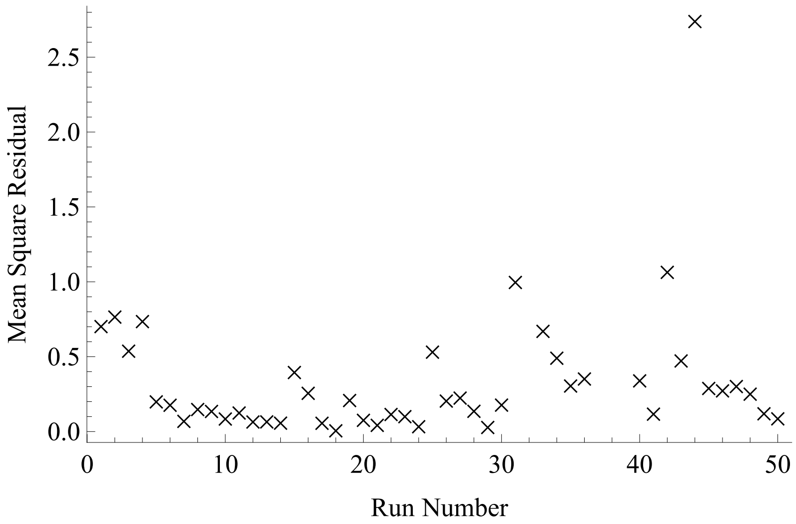


Figure 9.27: Mean square normalised residuals (chi-squared contributions) from each run for full calibration data.

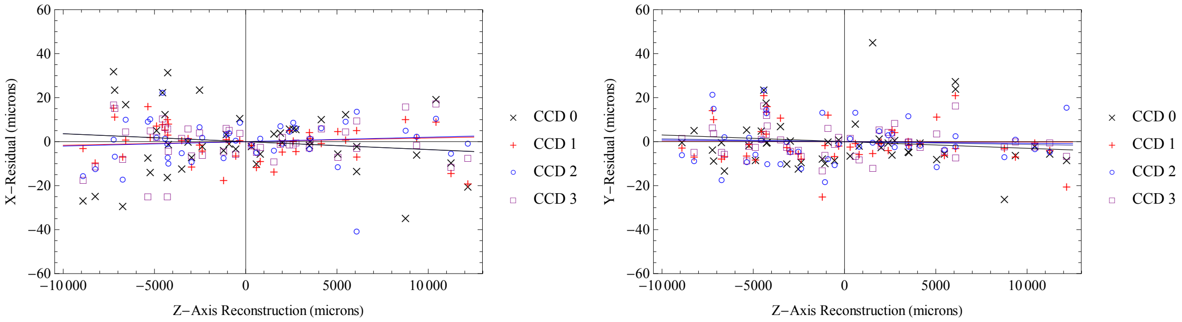


Figure 9.28: Residual correlation with Z-axis motion with a paraxial beam for X-axis residuals (left) and Y-axis residuals (right) using data generated using a pre-calibrated model. Residuals and fits have been inverted where appropriate and linear fits have been added.

It shows that run 44 disproportionately adds to the chi-squared and hence it was marked as bad and ignored in future analysis.

The intermediate calibration constants were applied to the model to generate another set of CCD readings from platform-only data as in section 9.7. The same analyses were re-run and the corrections updated, producing a noticeable improvement on the RMS residuals. Figure 9.28 shows the correlation of the residuals with the Z-position using a beam direction vector of $\begin{pmatrix} 1250 \times 10^{-6} & 30 \times 10^{-6} & 1 \end{pmatrix}$; the correlation is negligible and the fits for the different CCDs in good agreement compared to figure 9.24.

Table 9.15 shows results of applying the beam motion correction process of section 9.7.4 with the intermediate model for the three different laser tracker error paradigms. It is the analogue of table 9.12 but showing all three error paradigms and omitting the uncorrected data.

The average value of the scale factor (α), after accounting for the positive direction differences

Table 9.15: The effect of beam spot position correction on RMS residuals of CCD data with inserted cameras. The minimisation used specification laser tracker errors and a 1.25 mm/m beam propagation X component, ignoring runs 32, 37, 38 & 39. The cameras were inserted at positions found by preliminary calibrations with the respective tracker error paradigms and a 15 micron blanket CCD error estimate.

	CCD number and axis	0 X	1 X	2 X	3 X	0 Y	1 Y	2 Y	3 Y
Spec.	with geometrical correction (μm)	16.4	10.2	12.9	13.3	12.9	9.38	9.7	8.3
	after applying best α-correction (μm)	16.2	8.3	11.4	13.2	12.5	8.6	9.1	7.7
	associatedα value	0.14	-0.39	-0.40	0.13	0.26	0.32	-0.28	-0.26
	zero correlation α value	0.14	-0.39	-0.40	0.13	0.28	0.33	-0.29	-0.27
DESY	with geometrical correction (μm)	16.2	10.2	13.0	13.1	14.6	11.3	11.0	10.5
	after applying best α-correction (μm)	16.1	8.6	11.8	13.0	14.4	10.8	10.6	10.2
	associatedα value	0.09	-0.36	-0.35	0.08	0.21	0.26	-0.24	-0.21
	zero correlation α value	0.10	-0.36	-0.36	0.08	0.23	0.26	-0.25	-0.21
Exp.	with geometrical correction (μm)	16.3	10.0	12.6	13.3	12.5	9.0	9.7	7.9
	after applying best α-correction (μm)	16.2	8.1	11.1	13.2	12.0	7.9	9.1	7.1
	associatedα value	0.14	-0.39	-0.39	0.13	0.29	0.36	-0.29	-0.29
	zero correlation α value	0.14	-0.39	-0.40	0.13	0.32	0.38	-0.30	-0.30

($+--+$ for the X-axis and $++--$ for the Y-axis), was -0.29 for the X-axis and $+0.31$ for the Y-axis, with much better agreement between the cameras. This means the centre of rotation for the X-axis and Y-axis of the beam launch can reasonably be decoupled, with the centre for the X-axis correction being 6.5m from car 2 and the Y-axis correction centre being 3.75m. This gives the smaller correction needed for the X-axis and the larger one needed for the Y-axis.

The tracker error paradigm with the smallest errors is the experimentally derived set. The experimental error values will be used exclusively from this point forward. Although the Y-axis values are noticeably lower than the values found by the simulation (table 9.14) they are still broadly consistent.

Figure 9.29 shows the correlations with car 2, camera 3 reference and the RMS residuals using a preliminarily calibrated unit. (See figure 9.20 for comparison.) It shows the results for uncorrected data, and for data that has had the centre of rotation set at 3.75m from the reference camera for the x-correction and 6.5m from the reference camera for the y-correction.

Table 9.16 shows the final RMS residuals achieved after all corrections. There is a fair amount of variation, particularly in the X-axis, but this can be explained by the differences in beam uniformity, dust, noise level, beam brightness, interference, edge effects and mirror stability for the different CCDs. A slightly better result for the Y-axis values is expected and is observed.

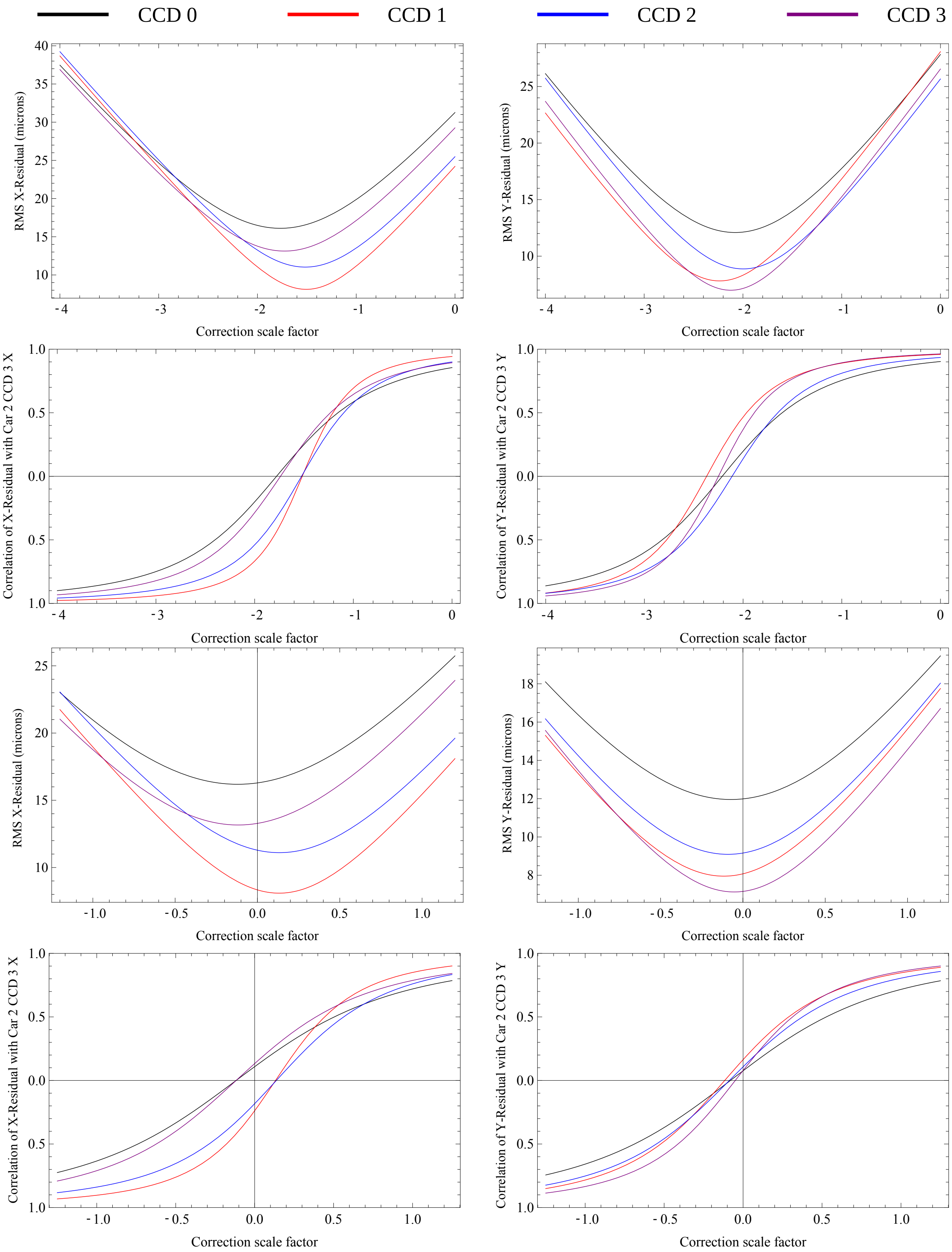


Figure 9.29: Effect on RMS residuals (odd rows) and correlation with car 2 camera 3 (even rows) in the X-axis (left) and Y-axis (right), with uncorrected data (top two rows) and corrected data (bottom two rows), upon the application of scaled correction with a partially calibrated model. The data has been flipped where necessary to remove the effects of camera inversion, making the lines directly comparable.

Table 9.16: Final RMS residuals in microns of corrected CCD data with generated data from inserted CCDs, using the experimentally derived laser tracker error set, with a calibrated model and all corrections applied.

CCD number	X-axis RMS residual	Y-axis RMS residual
0	16.2	12.0
1	8.1	8.0
2	11.1	9.1
3	13.2	7.1
Mean	12.3	9.0

9.9 Final Calibration Results and Analysis

A second iteration of the calibration process can now be performed using the final, corrected input data and the most suitable, experimental, laser tracker error paradigm. For this calibration the camera $c[i,4]$ and $c[i,5]$ values (rotations around their local X and Y axes) were set to zero and the platform base Z-position was fixed at zero.

9.9.1 Using Per-Run Statistical Errors

If only the per-run statistical errors are considered, then the results give the best error estimates achievable if a way to eliminate the quasi-systematic errors was found. Table 9.17 includes the results of a full calibration using the forty five good runs where the error estimates for each value of the CCD data are the statistical error estimates produced by the projection process described in section 9.3.5, giving generally sub-micron values. A measurement of value 1 with precision 10^{-5} was applied in the minimisation to the camera Y-scale to prevent the value minimising to an invalid minimum, far away from the values obtained with lower camera weightings; this is a symptom of having unrepresentative errors.

The calibration constants found have values that are consistent with the quasi-systematic results (§9.9.2) and the error estimates are very precise: the camera X and Y positions are found to less than a micron, the camera Z-positions are found to a couple of tenths of a millimetre and the $c[i,6]$ camera angles are found to a fraction of a milliradian. The results also give error estimates for the reconstruction of the platform locations which are very precise compared to the quasi-systematic estimates (table 9.18).

However, the results are misleading; table 9.19 shows the standard deviation of the residuals of the data and the model in the same way as table 9.10, with the χ^2/DoF also shown. It clearly shows that, even though the minimisation has adjusted to match the cameras by worsening the agreement with the laser tracker, the RMS residual for the CCD beam spot fit positions are an order of magnitude worse than the input error. The final χ^2/DoF at 224 is well away from unity, meaning it is not a sound representation of the system. It is not entirely clear why the RMS residual of the 'inner' cameras (CCD 0 & CCD 3, corresponding to h[1,j] and h[4,j],) is around 8 microns while the 'outer' cameras have values around 4 microns. It is likely that the underlying errors of table 9.16 used for the quasi-systematic calibration are sound, but the minimisation can adjust the reconstructed orientation of the unit to accommodate the outer cameras better - masking the error by a slightly incorrect reconstruction. As the inner virtual cameras are essentially on top of each other, they cannot benefit from this and give values consistent with the results of the coincident CCD checks in section 9.4.1.

9.9.2 Using Quasi-Systematic Errors

Having demonstrated that the errors are dominated by unknown dynamic systematics (§9.5), it is reasonable to attempt to characterise their effects using statistical methods to give an overall measurement uncertainty. The dynamic effects of the quasi-systematics could manifest themselves via:

- The transformation of all measurements to a common frame.
- The lack of knowledge of the 'true' measurement errors of the laser tracker.
- Imperfections of the LSM beam corrections.
- Fit distortions such as dust and edge effects.

As found in section 9.4, the dominant contribution comes from the beam fit distortions, though the results of 9.16 should account for all contributions. However, it is important to check that there are no clear issues with the approach. Figure 9.30 shows the correlations of the residuals of the fits with the reconstructed unit co-ordinates ($P[i,j]$), as well as the output observables. The correlations do not exhibit any clear pattern and the overall distribution of correlations is Gaussian-like. There is a slight positive correlation of the yaw residuals of marker 2 with the platform Z-position and a similar but negative correlation for marker 4. This is connected with similar correlations of the residuals with the yaw value for the individual markers, but isn't extreme. As no significant correlation with any parameter is present, it is apparent that the dynamic offsets introduced by the listed factors introduce an uncertainty which *behaves* like a statistical error.

Even though there is the potential for its elimination if a greater understanding is found, normal statistical error propagation can be expected to give a reasonable estimate of the resultant uncertainty. As such, it is appropriate to use the CCD errors in table 9.16 along with the experimentally derived laser tracker errors of table 9.6 and representative errors of 45 microradians for the tilt sensors.

Table 9.17 gives the calibration constants and their error estimates produced by the minimisation. It also has an additional column showing the results of repeating the calibration having fixed the baseplate's parameters ($b[i]$) at the values found by the first minimisation, which eliminates the uncertainty added by the axis system flexibility. Although the significant uncertainty of the location of the baseplate directly adds to the laser tracker and platform marker position uncertainties, highlighted in the table, the results show that it does not bleed into the calibration constants of the unit itself. Thus, there is a decoupling of the uncertainty of the axis system itself and the components in the axis system. The results give the positions of the CCDs to within two microns in the important directions (X and Y) and 1 mm in the far less sensitive Z-axis. The magnitude of the difference from the nominal positions is somewhat surprising, at up to one millimetre, but this is borne out by the images of figure 4.5b where the beam spots are far from the centre of the CCDs when aligned and levelled. The only angle that has a noticeable effect ($c[i, 6]$) is found to 2 milliradians, with the values ranging from 2 to 11 milliradians. The poorly understood camera scaling factors are found to 0.1% in the X-axis, while the value found for the Y-axis is consistent with a value of unity. The tilt sensor scale factors are found to 0.6% around the X-axis (tsx) and 0.1% around the Z-axis (tsz), while their orientations are found to O(10μrad) in the axes which produce the zero-offsets in the measurements and far worse in the axes which only have second order effects. These values are in line with expectations considering the geometry and the tilt-sensor error estimates. Overall, the set of error estimates give no cause for concern over the method's applicability.

Table 9.18 shows the means of the error estimates for the unit/platform positions produced as part of the calibration. The values are consistent with what would be roughly expected from the errors on the measurements. Comparing with the platform-only reconstructions of table 9.9, there is an improvement in the error estimate in the parameters which the LSM measures more precisely than the tracker, with a small improvement in other errors: we see little change in the X-position error, which is well defined by the laser tracker distance measurement. However, there is a 60% reduction in the Y-position error, and a 30% reduction in Z-position error. There is also a more than 50% reduction in the angular errors. The reduction in the Z-axis error is not well understood, but thought to be caused by the reduction of leakage of the other errors, as there is no direct Z-axis translational measurement apart from with the laser tracker platform.

Table 9.19 shows the RMS residuals of the data with the values produced by the model. The data again suggests that the error estimates for the laser tracker measurements are slightly too pessimistic with the ratio of the RMS residuals to their input errors less than unity and less than the 0.8 that would be expected after the 20% absorption of the residual into a reconstruction parameter residual discussed in section 9.6.8. However, the overall χ^2/DoF is reasonable with a value of 0.35. The laser tracker ratios are much more reasonable than with the per-run errors because the minimisation does not force the unit to be consistent with the CCD beam spot positions at all costs. The CCD residuals are not significantly worse than the per-run values, showing that it is the CCD measurement inconsistencies that dominate the per-run minimisation. The correlation of the residuals were checked against all the reconstructed parameters, producing a Gaussian-like distribution centred on zero with a standard deviation of 0.2 and no outliers. Table 9.20 summarises the RMS residuals for the CCD beam spot fit positions from the calibration. It suggests that CCD 0 is significantly worse than the others, consistent with table 9.16, though direct comparison is difficult as the input errors affect the residuals.

There are no obvious issues with the results, the values produced are within expectations and the constants are smaller than the dynamic uncertainties so the calibration can be considered successful.

Table 9.17: Calibration constants and associated errors for full calibrations. The first column contains the results for setting the input CCD hit position errors as the values in table 9.16; The next column contains the resultant errors, with the third column giving the errors produced when the baseplate parameters are fixed. The final columns contains the results when the camera errors are the per-run values whose means are summarised in table 9.4. Starred values are fixed and highlighted values are of particular interest.

			Final Settings (+ fixing baseplate)			Per-Run CCD Errors	
	Constant	Units	**Value**	**Error** $\times 10^{-6}$	Error$\times 10^{-6}$	Value	Error $\times 10^{-6}$
Laser Tracker	l[1]	m	0.87057	**258.**	**1.2**	0.87093	243.
	l[2]	m	0.19324	**114.**	**3.4**	0.19324	39.
	l[3]	m	0.017823	**3.6**	3.1	0.017822	3.11
	l[4]	rad	0.002759	**15.**	15.	0.002754	14.5
	l[5]	rad	2.612160	**7.9**	7.2	2.612164	7.19
	l[6]	rad	-0.006932	**10.1**	9.7	-0.006935	9.68
Base	b[1]	m	0.31944	**259.**	**-**	0.31981	243.
	b[2]	m	-0.00060	**114.**	**-**	-0.000608	38.3
	b[3]	m	0*	-	**-**	0*	-
Platform Markers	m[1,1]	m	0.31945	**259.**	**0.01**	0.31982	243.
	m[1,2]	m	-0.03561	**114.**	**0.44**	-0.035623	38.3
	m[1,3]	m	-0.0605911	**1.1**	1.1	-0.0605914	1.09
	m[2,1]	m	0.31945	**259.**	**0.01**	0.31982	243.
	m[2,2]	m	-0.03561	**114.**	**0.44**	-0.035624	38.3
	m[2,3]	m	0.0606602	**1.1**	1.1	0.0606601	1.1
	m[3,1]	m	0.31945	**259.**	**0.01**	0.31982	243.
	m[3,2]	m	0.06943	**114.**	**0.87**	0.069424	38.2
	m[3,3]	m	-0.0000691	**1.3**	1.3	-0.0000687	1.32
	m[4,1]	m	0.41847	**259.**	**0.65**	0.41883	243.
	m[4,2]	m	-0.00060	**114.**	**1.2**	-0.000613	38.3
	m[4,3]	m	-0.0000858	**1.4**	1.4	-0.0000862	1.37
CCD 0	c[1,1]	m	0.0009548	**3.0**	2.9	0.0009543	**0.44**
	c[1,2]	m	-0.0004248	**2.8**	2.8	-0.0004305	**0.87**
	c[1,3]	m	-0.0037	**923.**	918.	-0.00333	**165.**
	c[1,4]	rad	π*	**-**	**-**	π*	**-**
	c[1,5]	rad	π*	**-**	**-**	π*	**-**
	c[1,6]	rad	0.0007	**1970.**	1960.	0.00298	**421.**
CCD 1	c[2,1]	m	0.0005101	**1.9**	1.7	0.0005082	**0.71**
	c[2,2]	m	-0.0000209	**1.9**	1.8	-0.0000155	**0.33**
	c[2,3]	m	0.1992	**758.**	750.	0.2003478	**523.**
	c[2,4]	rad	0*	**-**	**-**	0*	**-**
	c[2,5]	rad	π*	**-**	**-**	π*	**-**
	c[2,6]	rad	3.1429	**1860.**	1860.	3.15037	**436.**
CCD 2	c[3,1]	m	0.0005125	**2.3**	2.1	0.0005130	**0.72**
	c[3,2]	m	-0.0003150	**1.8**	1.6	-0.0003208	**0.62**
	c[3,3]	m	-0.2033	**966.**	960.	-0.2043	**552.**
	c[3,4]	rad	0*	**-**	**-**	0*	-
	c[3,5]	rad	0*	**-**	**-**	0*	**-**
	c[3,6]	rad	-0.0120	**1460.**	1450.	-0.01312	**425.**
CCD 3	c[4,1]	m	0.0009978	**2.6**	2.5	0.0009957	**0.51**
	c[4,2]	m	0.0002911	**2.1**	2.0	0.0002980	**0.78**
	c[4,3]	m	-0.0016	**720.**	710.	-0.00126	**164.**
	c[4,4]	rad	0*	**-**	-	0*	-
	c[4,5]	rad	π*	**-**	-	π*	-
	c[4,6]	rad	0.0036	**1490.**	1450.	0.0055	**421.**
Scale	cs[1]	-	1.0422	**1090.**	1090.	1.0431	**358.**
	cs[2]	-	0.9976	**2580.**	2540.	0.9809	1930.
Tilt Sensors	tsx	-	1.077	**6690.**	6680.	1.090	6090.
	tsz	-	1.0600	**1200.**	1030.	1.0593	1210.
	tax[1]	rad	0.000736	**9.0**	8.5	0.0007445	8.4
	tax[2]	rad	0.0049	**1110.**	1110.	0.0050	985.
	taz[1]	rad	-0.0159	**5130.**	5120.	-0.0034	4860.
	taz[2]	rad	0.007242	**11.6**	10.1	0.007237	10.9

Table 9.18: Means of the parameter error estimates for the reconstructed platform positions and orientations produced as part of the calibration.

Parameter	**Final**	Base-Fixed	Per-Run
P[i,1] (μm)	**1.9**	1.7	0.5
P[i,2] (μm)	**4.3**	4.3	1.3
P[i,3] (μm)	**9.6**	9.6	4.2
P[i,4] (μrad)	**24.0**	24.0	4.1
P[i,5] (μrad)	**24.0**	24.0	4.3
P[i,6] (μrad)	**14.3**	13.6	9.5

Table 9.19: Full calibration input measurement error schemes and standard deviations of residuals. Camera hit locations are labelled h[<camera number +1>, <axis number>]. rx is the x-tilt sensor reading and rz is the z-tilt sensor reading. The CCD input errors given above for the per-run sets are the mean values.

	Final			Per-run		
Measurement	Input	RMS resid.	ratio	Input	RMS resid.	ratio
d[1] (μm)	3.	0.8	0.26	3.	4.3	1.4
d[2] (μm)	3.	1.0	0.33	3.	4.4	1.5
d[3] (μm)	3.	0.9	0.31	3.	7.1	2.4
d[4] (μm)	3.	0.9	0.29	3.	4.2	1.4
p[1] (μrad)	12.9	3.5	0.27	12.9	14.1	1.1
p[2] (μrad)	12.9	3.1	0.24	12.9	14.5	1.1
p[3] (μrad)	12.9	3.8	0.30	12.9	14.8	1.1
p[4] (μrad)	12.9	3.5	0.27	12.9	19.6	1.5
y[1] (μrad)	15.	3.1	0.21	15.	5.3	0.35
y[2] (μrad)	15.	3.2	0.22	15.	5.2	0.34
y[3] (μrad)	15.	3.7	0.25	15.	6.8	0.45
y[4] (μrad)	15.	3.9	0.26	15.	5.8	0.39
h[1,1] (μm)	16.2	14.3	0.89	0.34	9.2	27.0
h[2,1] (μm)	12.0	5.2	0.44	0.4	4.6	11.4
h[3,1] (μm)	8.1	8.4	1.04	0.32	3.9	12.2
h[4,1] (μm)	8.0	8.9	1.11	0.28	6.8	23.9
h[1,2] (μm)	11.1	10.5	0.94	0.3	9.5	31.9
h[2,2] (μm)	9.1	2.4	0.27	0.26	3.6	13.6
h[3,2] (μm)	13.2	4.6	0.35	0.3	3.1	10.5
h[4,2] (μm)	7.1	3.4	0.48	0.27	4.2	15.9
rx (μrad)	45.	30.0	0.67	45.	45.4	1.01
rz (μrad)	45.	33.4	0.74	45.	43.2	0.96
χ^2/DoF		0.35			224	

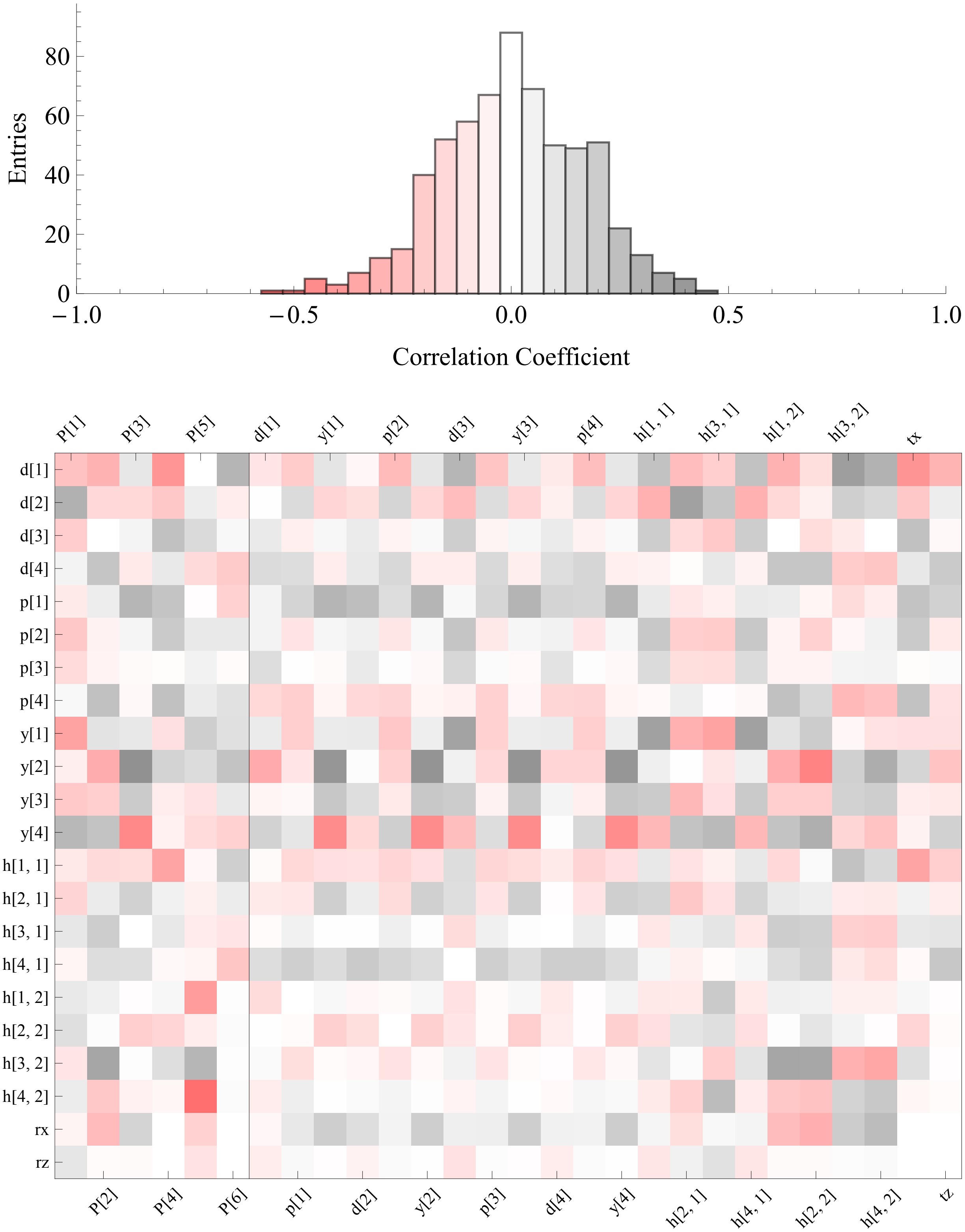


Figure 9.30: Correlations of residuals with model parameter values and with calculated model measurements. The lower image is a graphical representation of the correlation of the measurement residuals, down the vertical axis, against the sets of reconstructed co-ordinates (first six columns) and the calculated measurements themselves. The colour mapping (red to white to black) is given in the top histogram, which also shows the overall distribution.

Table 9.20: Final output RMS residuals in microns of CCD data from the full calibration.

CCD number	X-axis RMS residual	Y-axis RMS residual
0	14.3	10.5
1	5.2	2.4
2	8.4	4.6
3	8.9	3.4
Mean	9.2	5.2

9.9.3 Covariance Matrix Analysis

The covariance matrix of the model parameters is also a useful source of information on the behaviour of the system and whether it has minimised correctly. This matrix is very large and has a significant variation in magnitude of values, so again is normalised to give the Pearson correlation matrix. Figure 9.31 contains a graphical representation of the sub-matrix of the constants and first five runs of the calibration Pearson correlation matrix.

Looking at the centre of the sub-matrix, we see some anti-correlation between the camera X-position and Z-position. Looking at the reconstructed unit locations and rotations at the bottom-right of the sub-matrix, there is a correlation between the Z-axis reconstruction and the rotation around the Y′-axis, as well a correlation between the Y-axis translation and rotation around the Z-axis.

As can also be seen in the eigensystem analysis of section 9.9.4, there is also a clear collective motion in the X-axis and Y-axis. This is the cause of the large error estimates of the marker and laser tracker X and Y values seen in table 9.17 and their significant reduction by artificially fixing the baseplate position. Perhaps a more appropriate measurement of error when the axis system itself is not well defined is the relative error on a difference measurement between two positions. The error estimate in this case is simply

$$\sigma_{a-b}^2 = \sigma_a^2 + \sigma_b^2 - 2\sigma_{ab}\,, \tag{9.19}$$

which accounts for any correlation between the values. The errors for more complex measurements, such as more complex combinations of parameters can be found in exactly the same way as the point estimation performed in section 7.2 with the mathematics discussed in appendix B.3.5. Table 9.21 demonstrates the significant reduction in error estimates for a few correlated difference measurements, while confirming the quadrature result for measurements where little correlation

Table 9.21: Example error estimates of parameters and of relative measurements between them demonstrating the effects of correlation.

Parameter	Error (μm)	Parameter	Error (μm)	Measurement	Error (μm)
l[1]	258.5	m[3,1]	258.5	l[1]−m[3,1]	1.20
l[2]	114.1	m[3,2]	114.0	l[2]−m[3,2]	3.93
l[3]	3.1	m[3,3]	1.3	l[3]−m[3,3]	3.50
l[1]	258.5	p[3,3]	2.7	l[1]−p[3,3]	258.5
c[3,1]	2.1	p[3,3]	2.6	c[3,1]−p[3,3]	3.5

is seen.

Figure 9.32 gives the sub-matrix of the resultant calibration constants only. The mean value of the off-diagonal elements is 0.02, with a mean absolute value of 0.11, indicating little correlation between the constants. This should be expected with a properly set up problem; if there were very strong correlations, there would be small eigenvectors of the minimisation as the parameters would be almost degenerate. These would be observed with the eigensystem analysis and could then be fixed with a constraint. Examining the correlations, we see a strong correlation with the tilt sensor scale factors and the zero-offset rotation around the sensitive axis, which is entirely expected. Similarly, the negligible correlation of the Z-rotation with the anything else is reassuring as the cameras are almost blind to this.

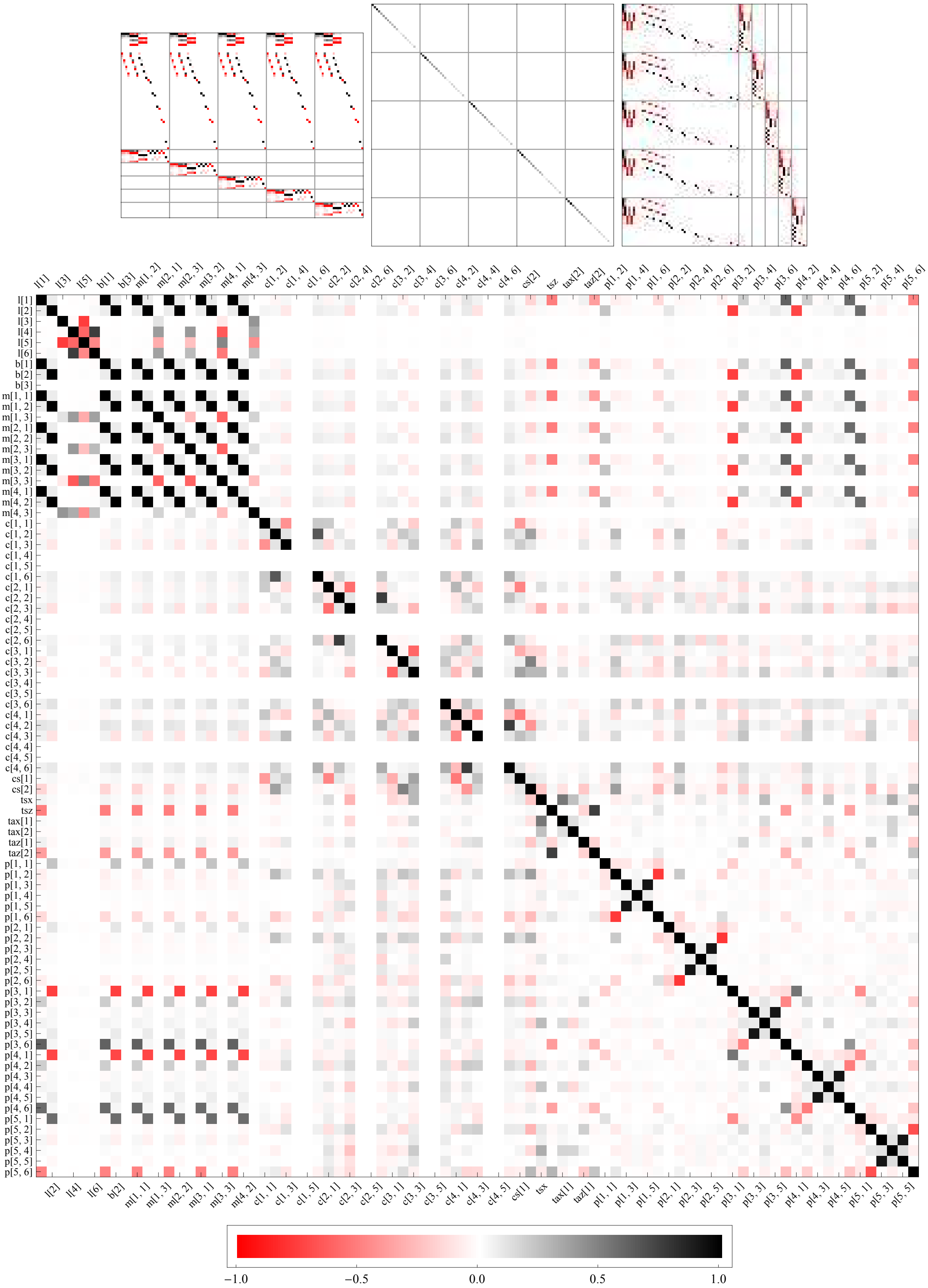


Figure 9.31: Top: Graphical representation of $\mathbf{A}^\mathbf{T}\mathbf{PA}$ at the minimum for the first five runs of a calibration session. Bottom: Correlation matrix extract, showing only the first 5 runs.

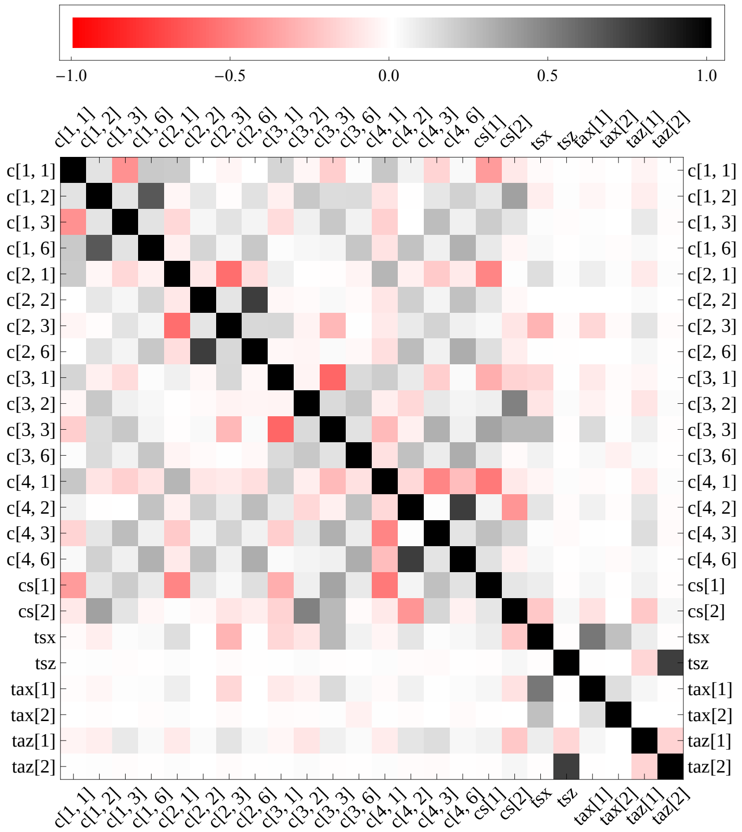


Figure 9.32: Graphical representation of the correlation matrix of the camera and tilt sensor calibration constants found by a full 45 run calibration.

9.9.4 Eigensystem Analysis

Inspecting the eigensystem of the constrained, weighted **C**-matrix (equation B.28) reveals various properties of the system. Figure 9.33 shows the values of the eigenvalues at the minimum. Negative eigenvalues are caused by constraints - eigenvalues for model parameters are always positive, while eigenvalues for constraint equations may be positive or negative. The right-hand side of the constraint equations have been multiplied by 1×10^{11}: while in theory this has no effect on the minimisation (§B.3.2), it reduces the condition number of the matrix, enabling easier computation. Small eigenvalues, along with large error estimates and numerical accuracy warnings from the fit, can be used to identify components of **X** that are poorly defined and produce near-symmetries (§9.8.2).

In the case of the calibration, the platform base Z-location is particularly poorly defined - the majority of the effect of any residual would be to turn already small X and Y rotations, reconstructed from the platform, into phantom z-translations which the cameras are completely blind to. Adding a ‘measurement’ generated from design diagrams breaks the symmetry, but the resultant minimised **X** simply snaps to the measurement with an error two orders of magnitude

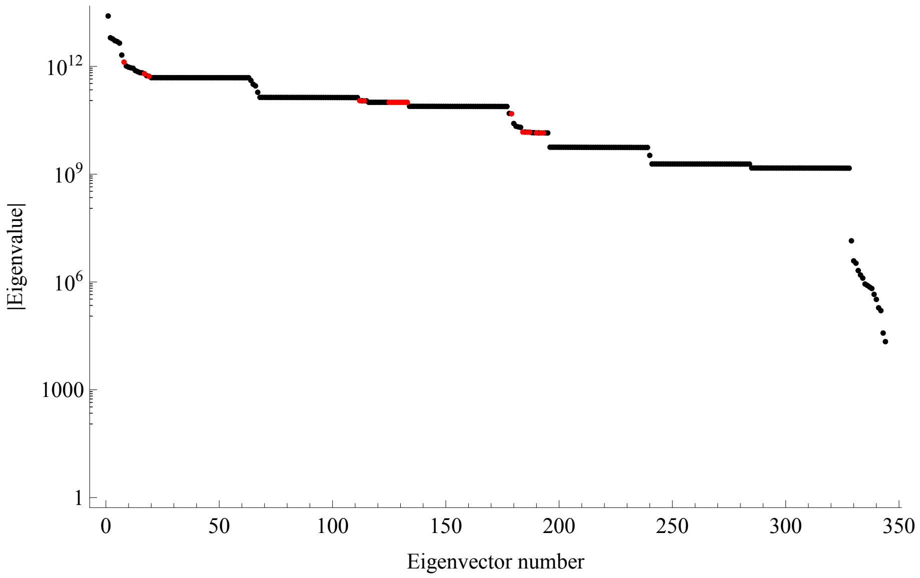


Figure 9.33: Eigenvalues for car 1 calibration; positive eigenvalues are in black, while negative eigenvalues corresponding to constraints are in red. The largest values are associated with laser tracker positions and constraints, the main plateaus mostly correspond to unit co-ordinate reconstructions, while the smallest eigenvalues are generally related to the CCDs and tilt sensor co-ordinates and the focus of figure 9.34.

larger than similar parameters. The eigenvector associated with the smallest eigenvalue is a unit vector in the direction of the baseplate Z-position parameter, meaning it is most efficient to fix the parameter at the design value and remove it from the equation.

One distinct group eigenvalues are significantly smaller than the others; figure 9.34 is a close-up of this tail of small eigenvalues and shows a visualisation of their associated eigenvectors for the vector elements that have non-negligible values. It demonstrates two classes of eigenvectors: virtually single component eigenvectors, and eigenvectors corresponding to co-ordinated motion.

For the near single-valued eigenvectors, the tilt sensor scale values both feature, along with the rotation angles which only have a second order effect on the value read. The scale value for the x-rotation tilt sensor is less well known (table 9.17) even though it has a contribution from the cameras because the range of motion is much smaller (figure 9.36). This directly leads to a smaller eigenvalue.

For the group motion eigenvectors, the clearest are the global motions in the X-axis and Y-axis seen in eigenvectors 330 and 329. Examining the motion in eigenvector 329, we see a collective motion of the laser tracker, the baseplate and the markers and a very slight motion of the cameras in the Z-axis. This can be visualised as demonstrating that an experiment with the tracker adjusted slightly higher, the platform mounted on the unit the same amount higher and with slight adjustments to the CCDs, but going through the exact same motions, would give very similar results.

9.9.5 Error Matrix Contribution Eigensystem Analysis

The eigensystem can also be used to find what has the largest contribution to the error of any component of **X**. This can be useful in highlighting the effects of system weaknesses caused by near symmetries. Consider the eigensystem identity:

$$\mathbf{C} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{-1}\,, \tag{9.20}$$

where **Q** is a matrix composed of columns of the eigenvectors and Λ is a diagonal matrix of the associated eigenvalues. Recalling that $\mathbf{C}^{-1}$ is the covariance matrix, we can find the error estimate for parameter $\mathbf{X}_i$ as follows:

$$\sigma_{i,k} = \left(\mathbf{C}^{-1}\right)_{ik} = \left(\mathbf{Q}^{-1}\boldsymbol{\Lambda}^{-1}\mathbf{Q}\right)_{ik}\,. \tag{9.21}$$

If we rewrite this in terms of the components we have:

$$\begin{aligned}\left(\mathbf{C}^{-1}\right)_{ik} &= \sum_{j,j'}\left(\mathbf{Q}^{-1}\right)_{i,j}\Lambda^{-1}_{j,j'}\mathbf{Q}_{j'k}\\ &= \sum_{j,j'}\left(\mathbf{Q}^{-1}\right)_{i,j}\frac{1}{\lambda_{j,j'}}\delta_{j,j'}\mathbf{Q}_{j'k}\\ &= \sum_{j}\left(\mathbf{Q}^{-1}\right)_{i,j}\frac{1}{\lambda_{j}}\mathbf{Q}_{jk}\,.\end{aligned} \tag{9.22}$$

Of particular interest are the diagonal elements (i.e. σ_{ii}) which correspond to the simple error estimate on that parameter. In effect, equation 9.22 is simply summing all the ith components of the basis eigenvectors, weighted by the inverse of their eigenvalue. By comparing the contributions to the sum of the different components, the most significant ones can be found and the eigenvectors associated with the eigenvalues can be analysed to give insight into the system weaknesses that increase the error.

For example, figure 9.35 shows the inverse-weighted eigenvalues and the most significant elements of their corresponding eigenvectors for a representative parameter; the CCD 1 X-position. Unsurprisingly, the largest contributor, eigenvector 13, is almost entirely the parameter itself. However, the second largest contributor is eigenvector 194, which corresponds to a collective X-

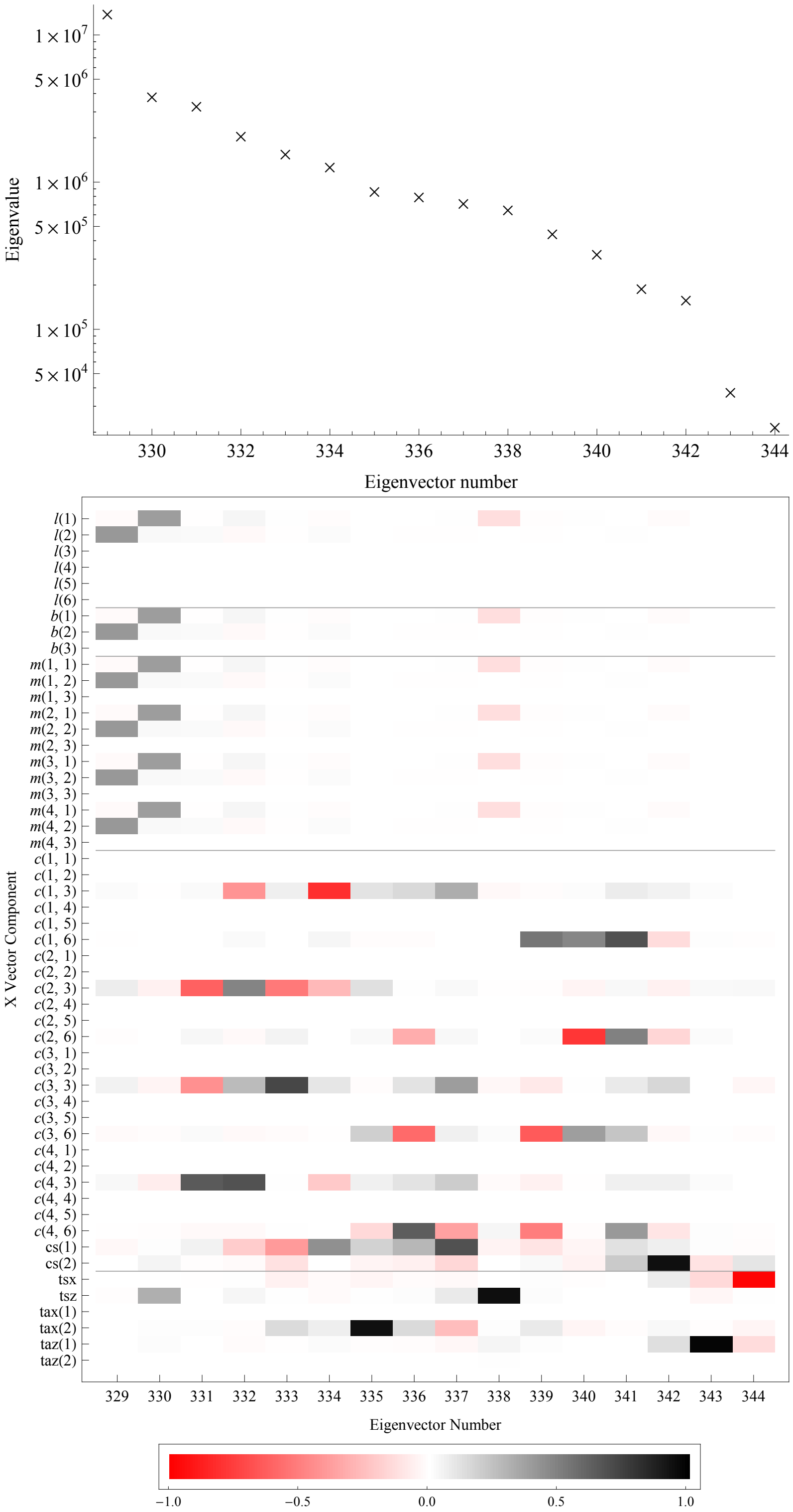


Figure 9.34: Top: Small Eigenvalues from the car 1 calibration. All eigenvalues are positive. Bottom: A graphical representation of the eigenvectors associated with the eigenvalues, showing only the components of **X** with significant values.

translation of the cameras, with a corresponding translation in the reconstructed run X-positions, as well as a lesser analogous motion in the Y-axis. The third eigenvector, 333, is an opposing z-motion for the outer cameras (CCD 1 and CCD 2), along with a change in the CCD X-axis scale factor. The fourth eigenvector, 337, is a complex weakness with the CCD X-axis scale factor being the largest contributor. The analysis produces similar results for the other parameters, with the self-eigenvector contributing between 20% and 63% for the CCD parameters, suggesting significant scope for an improvement in the error estimate by reducing the symmetries of the system. Collective motion eigenvector contributions analogous to eigenvector 194 were seen for all CCD parameters, typically with roughly a 10-20% weighting. Finally there were varying numbers and weightings of eigenvectors with few, hard to interpret components, like 333 and 337.

9.10 Altering the Number of Runs

It is important to check the performance of the calibration by observing the change in the calibration values as the number of runs used to calibrate is altered. The change in these values was found for a range of twenty five runs to forty five runs in steps of five runs, each a superset of the last.

One difficulty is that the origin and the axis directions that define the frame of reference are defined as part of the calibration, by using all the available runs (§9.6.7). When comparing calibrations comprising of different numbers of runs, the mean positions of the markers will be shifted by many tens to hundreds of microns, significantly affecting the axis origins and directions.

A reasonable approach might be to fix the baseplate positions to the values derived from the calibration with the most runs. However, this would not fix the axis directions themselves; the constraints used to define the axis directions are too complex to be fixed in the same way. Instead, the constraints that define the axes can be applied to only a common subset of the data, present in all calibrations. To ensure a common subset, a fixed random order set of the forty-five valid runs was used, with each calibration using the same n randomly chosen runs numbers in its axis constraints. n was set to be 25 for this comparison.

Table 9.22 shows the calibration constant values found by the 45-run calibration with their error estimates, as well as the difference between the parameters and those found with the other calibration runs. The calibration constants of the 45-run calibration are the same as the ones found in table 9.17, but expressed in a different frame of reference due to the different axis defining constraints. The differences between the values of the constants from the 45-run calibration and the other calibration runs have similar values to the respective errors, suggesting the errors

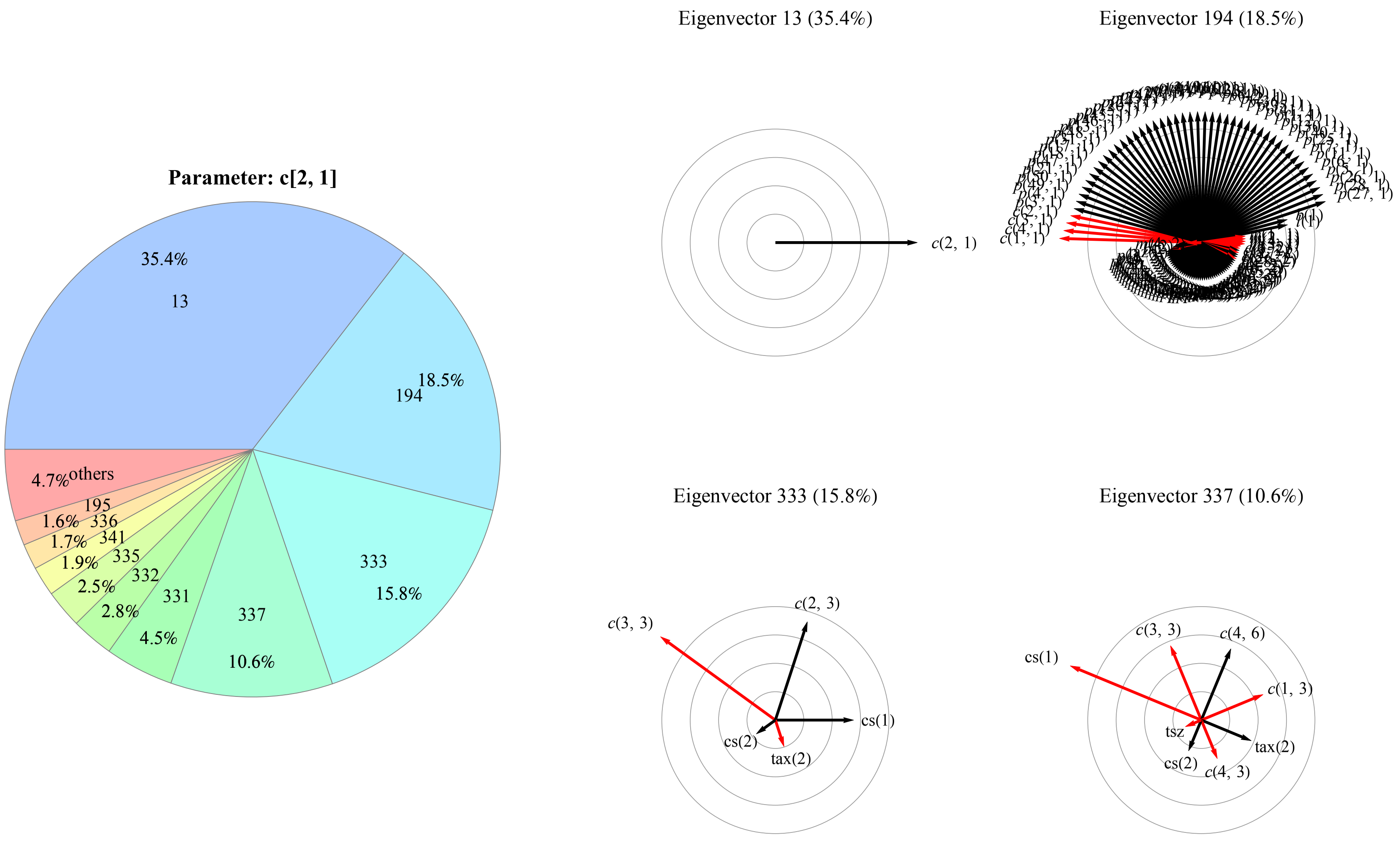


Figure 9.35: Inverse-weighted eigenvectors contributing to the error on the CCD 1 X-axis (c[2,1]) value. Representations of the relative values of the most significant components of the first four most important eigenvectors are presented as radial arrows proportional to the component size, with red arrows signifying a negative value. A component with a value of more than 10% of the largest is considered relevant.

Table 9.22: Calibration constants and errors. Starred values are fixed. The values are given at higher precision than is warranted under the error estimates of table 9.23 due to the small differences between the calibrations.

		45 runs		40-45	35-45	30-45	25-45
Constant	Units	Value	Error×10^{-6}	Diff.×10^{-6}	Diff.×10^{-6}	Diff.×10^{-6}	Diff.×10^{-6}
l[1]	m	0.8708293	258	17.5	29.9	-1.52	-19.9
l[2]	m	0.1920496	114	-22.4	-21.0	144	90.3
l[3]	m	0.0180116	3.6	0.04	0.14	-0.11	-0.68
l[4]	rad	0.0028512	15.0	-0.56	-1.67	0.20	-0.03
l[5]	rad	2.6123219	7.9	0.09	0.22	0.06	1.14
l[6]	rad	-0.0082522	10.1	-0.94	-2.2	-0.61	-1.08
b[1]	m	0.3194702	259	17.4	29.7	-1.63	-20.2
b[2]	m	-0.0005359	114	-22.1	-20.4	145	90.8
b[3]	m	0*					
m[1,1]	m	0.3194814	259	17.4	29.7	-1.63	-20.2
m[1,2]	m	-0.0355514	114	-22.1	-20.4	145	90.9
m[1,3]	m	-0.0605912	1.1	-0.14	-0.13	-0.08	0.04
m[2,1]	m	0.3194820	259	17.4	29.7	-1.63	-20.2
m[2,2]	m	-0.0355522	114	-22.1	-20.4	145	90.9
m[2,3]	m	0.0606601	1.1	0.06	0.11	0.11	0.07
m[3,1]	m	0.3194818	259	17.4	29.7	-1.63	-20.2
m[3,2]	m	0.0694959	114	-22.2	-20.5	145	90.7
m[3,3]	m	-0.0000689	1.3	0.08	0.03	-0.03	-0.11
m[4,1]	m	0.4184918	259	17.4	29.6	-1.70	-20.2
m[4,2]	m	-0.0005413	114	-22.0	-20.3	145	90.8
m[4,3]	m	-0.0000860	1.4	-0.03	-0.07	-0.13	-0.34
c[1,1]	m	0.0009881	3.0	-0.34	-0.84	0.74	1.14
c[1,2]	m	-0.0003544	2.8	0.20	0.10	0.91	-0.79
c[1,3]	m	-0.0035831	923	103	-183	-877	-965
c[1,4]	rad	π*					
c[1,5]	rad	π*					
c[1,6]	rad	0.0021899	1970	-125	-756	-898	-2050
c[2,1]	m	0.0005100	1.9	-0.71	-0.14	0.47	1.15
c[2,2]	m	0.0000325	1.9	0.05	-0.34	-0.52	-0.80
c[2,3]	m	0.199193	758	331	270	-311	-256
c[2,4]	rad	0*					
c[2,5]	rad	π*					
c[2,6]	rad	3.1440997	1860	-149.9	-748	-780	-1270
c[3,1]	m	0.0005780	2.3	-0.57	0.11	0.62	0.09
c[3,2]	m	-0.0002295	1.8	-0.37	-0.16	-0.03	-0.58
c[3,3]	m	-0.2029504	966	280	50.6	-388	-237
c[3,4]	rad	0*					
c[3,5]	rad	0*					
c[3,6]	rad	-0.0109285	1460	-415	-444	24.1	-2100
c[4,1]	m	0.0010307	2.6	-0.34	1.00	1.94	4.2
c[4,2]	m	0.0003615	2.0	0.01	-0.38	-0.55	-1.19
c[4,3]	m	-0.0017594	720	129	97.8	-278	-291
c[4,4]	rad	0*					
c[4,5]	rad	π*					
c[4,6]	rad	0.0048331	1490	-121	-338	-387	-1280
cs[1]	-	1.0438823	1090	361	-346	-180	-832
cs[2]	-	0.9974933	2580	-129	-203	252	-814
tsx	-	1.0782875	6690	1650	2030	3030	4650
tsz	-	1.0599094	1200	-22.1	549	975	836
tax[1]	rad	0.0008095	9.0	3.74	3.4	2.6	2.2
tax[2]	rad	0.0047726	1110	330	450	844	784
taz[1]	rad	-0.0160560	5130	-199	164	-1830	-2100
taz[2]	rad	0.0086068	11.6	1.23	6.0	11.5	8.6

Table 9.23: Calibration constant errors. Starred values are fixed.

Constant	Units	Error estimate with number of runs				
		45	40	35	30	25
l[1]	μm	258	267	291	322	339
l[2]	μm	114	120	131.909	148	153
l[3]	μm	3.6	3.7	3.8	4.0	4.2
l[4]	μrad	15.0	15.8	16.8	18.0	19.5
l[5]	μrad	7.9	8.2	8.6	9.1	9.7
l[6]	μrad	10.1	10.6	11.2	12.0	13.0
b[1]	μm	259	267	291	322	339
b[2]	μm	114	120	132	147	153
b[3]	μm					
m[1,1]	μm	259	267	291	322	339
m[1,2]	μm	114	120	132	147	153
m[1,3]	μm	1.1	1.2	1.2	1.3	1.5
m[2,1]	μm	259	267	291	322	339
m[2,2]	μm	114	120	132	147	153
m[2,3]	μm	1.1	1.2	1.2	1.3	1.5
m[3,1]	μm	259	267	291	322	339
m[3,2]	μm	114	120	132	147	153
m[3,3]	μm	1.3	1.4	1.5	1.6	1.8
m[4,1]	μm	259	267	291	322	339
m[4,2]	μm	114	120	132	147	153
m[4,3]	μm	1.4	1.5	1.6	1.7	1.8
c[1,1]	μm	3.0	3.2	3.3	3.6	3.8
c[1,2]	μm	2.8	2.9	3.1	3.2	3.6
c[1,3]	μm	923	992	1020	1120	1190
c[1,4]	μrad					
c[1,5]	μrad					
c[1,6]	μrad	1970	2050	2150	2270	2500
c[2,1]	μm	1.9	1.9	2.0	2.1	2.2
c[2,2]	μm	1.9	2.0	2.0	2.1	2.2
c[2,3]	μm	758	815	832	898	954
c[2,4]	μrad					
c[2,5]	μrad					
c[2,6]	μrad	1860	1940	2040	2160	2380
c[3,1]	μm	2.3	2.4	2.5	2.6	2.8
c[3,2]	μm	1.8	1.9	1.9	2.0	2.2
c[3,3]	μm	966	1030	1070	1150	1220
c[3,4]	μrad					
c[3,5]	μrad					
c[3,6]	μrad	1460	1550	1610	1740	1930
c[4,1]	μm	2.6	2.7	2.8	3.0	3.2
c[4,2]	μm	2.0	2.2	2.2	2.4	2.6
c[4,3]	μm	720	776	802	878	932
c[4,4]	μrad					
c[4,5]	μrad					
c[4,6]	μrad	1490	1550	1620	1720	1900
cs[1]	$\times10^{-6}$	1090	1110	1170	1210	1310
cs[2]	$\times10^{-6}$	2580	2810	2980	3270	3770
tsx	$\times10^{-6}$	6690	7300	7550	8460	9210
tsz	$\times10^{-6}$	1200	1220	1330	1490	1570
tax[1]	μrad	9.0	9.6	9.9	10.6	11.5
tax[2]	μrad	1110	1150	1230	1400	1500
taz[1]	μrad	5130	5570	5750	6460	6970
taz[2]	μrad	11.6	12.2	13.5	15.1	16.6

are an accurate representation of the system. Table 9.23 shows the constant parameter error estimates for the different calibration experiments, with typically a thirty percent increase when comparing the 45-run calibration errors with the 25-run calibration errors. A $\sqrt{n}$-like scaling seems reasonable for these parameters which are common to all runs.

Figure 9.36 presents the reconstructed unit co-ordinates produced by the 25-run calibration and the 45-run calibration with the axes constrained by the common runs. It demonstrates the range of motion over the different axes and highlights which runs were randomly assigned to be part of the smaller calibration set. Visual comparison with figure 9.13 shows the results are in line with rough estimates. They do not show any clear difference between the residuals of the axis defining runs and the others, suggesting that the constraint does not introduce any particular bias.

Table 9.24 gives the mean of the reconstruction parameter error estimates for the different calibration experiments. It shows only a minor variation with the number of runs used when calibrating. There is a general increase as the number of runs used in calibration declines but the amount is only small and not all parameters follow the pattern, contrasting with the fairly uniform and significant degradation of the error estimate of the constant parameters seen in table 9.23. This is due to the majority of the error for an individual reconstruction coming from the uncertainty of the measurements directly applicable, rather than the indirect uncertainty of the other runs via the common constants.

It is important to see how the reconstructed co-ordinates themselves change to assess the effect of variations in the calibration process which give rise to systematic errors. Figure 9.37 shows the how the reconstructed co-ordinates produced by the calibrations using 25, 30, 35 and 40 runs differ from the co-ordinates produced by the 45-run calibration. There is no clear pattern with run number, though it is clear that the largest average squared deviation is for the 45-25 difference. Table 9.25 shows the RMS difference of reconstructed parameters between the various calibration sets for all common runs. When the runs used for defining the axis system were treated separately, there was no significant change in RMS difference values between the axis-defining and other runs, confirming the suggestion of no bias introduced by having only some runs involved in the axis constraints. Unsurprisingly, altering the number of runs by a larger amount has a greater effect, both in absolute terms (e.g. 45-40 vs. 45-25) and relative terms (e.g. 45-35 vs. 40-30 or 35-25). The RMS differences in table 9.25 are significantly smaller than the error estimates of the parameters seen in table 9.24, giving no particular cause for concern. Note that there is zero overall bias in any of the parameters due to the axis definition. However, an r.m.s difference of

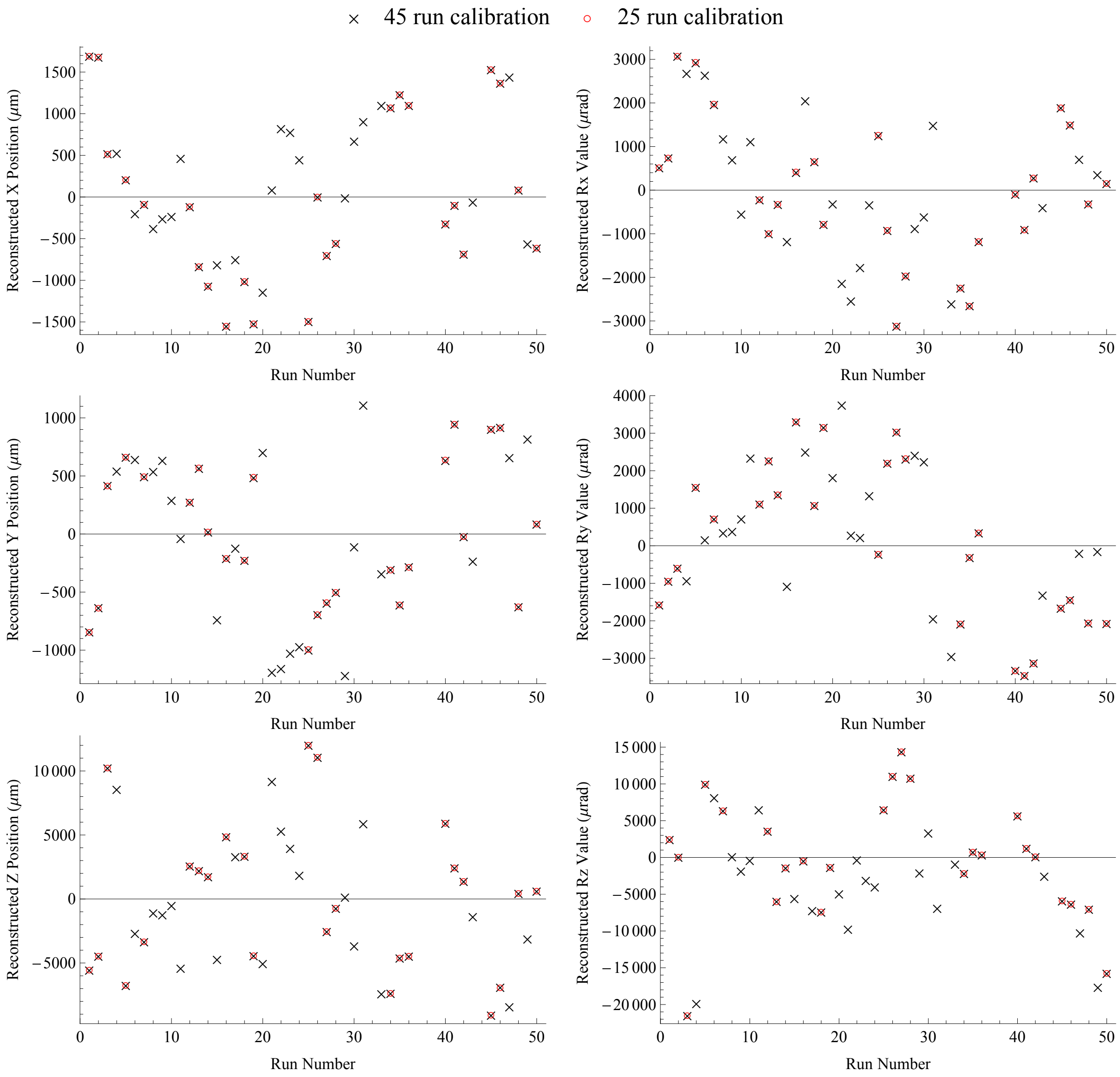


Figure 9.36: Reconstructed unit co-ordinates for both 25 runs and 45 runs with the axes defined by the common set of 25 runs.

order 1 micron or 3 microradians is not completely insignificant, so for this system the full set of calibration runs was used for future work.

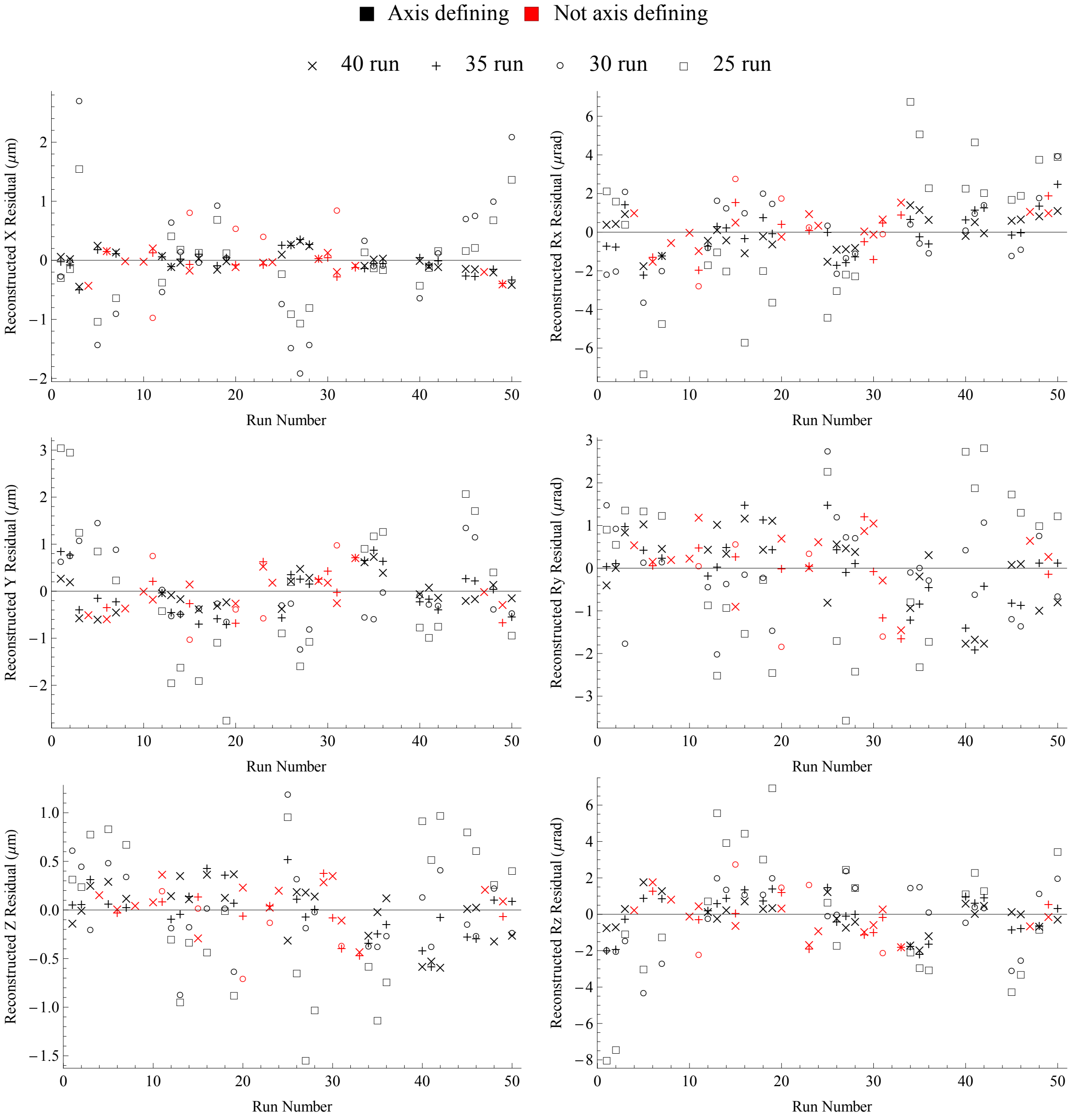


Figure 9.37: Residuals when compared with the reconstructed unit co-ordinates of the 45-run calibration set for calibrations of different run lengths. Runs which were used to define the axes system have been marked in black while runs that were not are in red.

Table 9.24: Error estimates for reconstructed unit co-ordinates for calibrations using different numbers of runs.

Calibration Runs	45	40	35	30	25
P[i,1] (μm)	1.9	1.9	1.9	2.0	2.0
P[i,2] (μm)	4.3	4.3	4.4	4.5	4.6
P[i,3] (μm)	9.7	9.7	9.7	9.6	9.6
P[i,4] (μrad)	24.3	24.3	24.3	24.4	24.6
P[i,5] (μrad)	24.2	24.1	24.1	24.1	24.1
P[i,6] (μrad)	14.4	14.5	14.7	14.9	15.2

Table 9.25: RMS differences of reconstructed parameters between pairs of calibrations.

Comparison sets	45-40	45-35	45-30	45-25	40-35	40-30	40-25	35-30	35-25	30-25
Common runs	40	35	30	25	35	30	25	30	25	25
P[i,1] (μm)	0.19	0.19	1.02	0.65	0.07	1.20	0.84	1.20	0.83	0.44
P[i,2] (μm)	0.36	0.50	0.74	1.55	0.31	0.96	1.52	0.82	1.22	1.17
P[i,3] (μm)	0.27	0.25	0.43	0.77	0.29	0.58	0.95	0.38	0.81	0.65
P[i,4] (μrad)	0.87	1.15	1.81	3.67	0.86	1.66	2.87	0.93	3.14	3.59
P[i,5] (μrad)	0.83	0.82	1.10	1.87	0.77	1.64	2.52	1.19	2.18	1.99
P[i,6] (μrad)	0.89	1.13	1.90	3.76	0.63	2.43	3.61	2.08	2.97	2.81

10 Calibrated Reconstruction

With the unit calibrated, it can accurately reconstruct its position and orientation co-ordinates, apart from in the Z-direction, using the cameras and tilt sensors alone. Errors from both calibration constants and measurements need to be propagated carefully, and there is scope for using the overdetermined nature of reconstruction to improve performance.

10.1 Application of Calibration Constants

The constants found in the calibration process can be applied directly to the model with no account taken for the error estimates[11]. Applying the constants in this way has the important advantage that the co-ordinates reconstructed for different runs are all completely independent from each other. However, the calibration constant error estimates are not propagated into the final parameter uncertainties and must be accounted for as a separate systematic contribution.

Table 10.1 gives the mean values of the error estimates for reconstructed co-ordinates of the same forty five runs used to calibrate the unit; using either the representative quasi-systematic input CCD beam spot position errors (column 1), or the per-run statistical errors (column 2)[12]. The error estimates of column 1 and column 2 look reasonable when the considering the input beam position quasi-systematic errors (table 9.20) and per-run statistical errors (table 9.3) in combination with the calibration constant covariances (table 9.17).

For comparison, the platform only error estimates produced when using only the laser tracker measurements (column 3, see §9.6.8), and when combining the laser tracker measurements and quasi-systematic LSM measurements as part of the calibration (column 4, see table 9.18.) are also shown. When considering the quasi-systematic results, the laser tracker platform-only errors are better in the X-translation and Z-rotation, similar for the Y-rotation, and noticeably worse for the Y-translation and X-rotation. This may mean a better witnessing system would be required to improve calibration constant values strongly dependent on these measurements. Similarly, calibration would become more difficult if the precision of the CCD beam measurements were to improve as the laser tracker platform reconstruction errors would overwhelm the contributions from the CCDs.

The per-run errors are potentially a guide to what is achievable with the system if systematics

---

[11] In practice, it was more straightforward to add measurements of the constants to the **L** vector with a very high weighting in the **P** matrix. This still has the effect of fixing them at the input values.

[12] The data used for reconstruction could be the uncorrected 3rd order projected data for time $t = -3$s but it is appropriate to use the corrected zero-order order projected data used for calibration to enable comparison. The quasi-systematic errors will be identical while the per-run statistical errors have already been found to have very similar values around 0.3μm.

Table 10.1: Mean error estimates of reconstruction parameters with fixed calibration constants, for both the error estimates which include quasi-systematic effects and the over-optimistic per-run errors. The values found for platform-only reconstruction using only laser tracker measurements, and output as part of the calibration minimisation (using laser tracker, tilt sensor and CCDs with quasi-systematic errors) are shown for comparison. Note that reconstruction is blind to Z-axis motion (r[i,3]).

Co-ordinate	Parameter	Q-S Errors	Per-Run	Platform	Calibration
X-translation	r[i,1] (μm)	**5.5**	0.15	1.9	1.9
Y-translation	r[i,2] (μm)	**4.3**	0.14	10.4	4.3
RX-rotation	r[i,4] (μrad)	**25.5**	0.98	61.9	24.0
RY-rotation	r[i,5] (μrad)	**32.3**	1.03	38.8	24.0
RZ-rotation	r[i,6] (μrad)	**47.7**	47.3	32.4	14.3
	χ^2/DoF	0.61	807		

related to the fitted beam spot positions on the CCDs could be eliminated, but are inappropriate for use as the high chi-squared per degree of freedom value of 807 confirms. This is significantly down from the order of millions originally seen in section 8.5.2, however.

It is appropriate to check the reconstructed unit co-ordinates against those found from the platform-only reconstruction, and the full calibration. Table 10.2 contains the RMS differences between the LSM reconstructions and the platform-only reconstructions (column 2), as well as the RMS differences between the LSM reconstructions and the unit locations output as part of the calibration process (column 4). The differences are consistent with the error estimates (columns 3 and 5), apart from being too high in the X-axis. The X-axis discrepancy is due to the unit Z-position being fixed at zero. The laser beam propagation direction is not exactly along the Z-axis, but has a 1.25×10^{-3} X-component and a 0.03×10^{-3} Y-component (§9.8.3). Without the internal FSI included in the minimisation to determine the Z-position, the reconstructed position will be incorrect by the leakage from the Z-translation into the X and Y translations. The 3.6mm range of Z-positions used translates to a 4.5μm range of systematic contributions in the X-axis and a negligible 0.1μm contribution in the Y-axis.

10.1.1 Altering the number of calibration runs

If the reconstruction is performed with calibration constants obtained from calibrations using different numbers of runs (§9.10) then the reconstructed co-ordinates change slightly, but there was no noticeable change in the error estimates. Table 10.3 shows the RMS difference in reconstructed

Table 10.2: RMS of the difference between the reconstruction values produced by the reconstruction with the a platform-only reconstruction for the same runs (left) and as a side effect of the calibration itself (right), as well as the mean error estimates from those runs. (Note that the LSM errors include the contributions from calibration constant uncertainties; see section 10.2.1. Note also that the errors for the differences themselves have not been estimated.)

Parameter		vs. Platform-only		vs. Calibration Output	
	Error	RMS diff	Error Estimate	RMS diff	Error Estimate
r[i,1] (μm)	5.8	9.4	1.9	9.0	1.9
r[i,2] (μm)	5.3	8.8	10.4	7.2	4.3
r[i,4] (μrad)	27.6	20.8	61.9	11.9	24.0
r[i,5] (μrad)	34.1	31.0	38.8	20.0	24.0
r[i,6] (μrad)	51.1	27.2	32.4	35.3	14.3

Table 10.3: Root mean square differences of reconstructed co-ordinates between pairs of reconstructions using calibration constants from calibrations using varying numbers of runs; larger batches are always a superset of smaller batches. The errors on the reconstruction parameters are given for context.

Parameter	Error	45-40	45-35	45-30	45-25	40-35	40-30	40-25	35-30	35-25	30-25
r[i,1] (μm)	5.8	0.53	0.58	0.70	1.22	0.56	1.13	1.33	1.08	0.96	1.11
r[i,2] (μm)	5.3	0.44	0.54	0.80	1.55	0.32	1.12	1.48	0.96	1.16	1.31
r[i,4] (μrad)	27.6	1.08	1.31	2.10	4.28	0.94	1.85	3.26	1.13	3.59	4.08
r[i,5] (μrad)	34.1	1.18	1.00	1.13	4.01	0.84	1.60	4.02	0.94	3.99	4.64
r[i,6] (μrad)	51.1	0.37	4.12	7.28	6.50	4.36	7.37	6.55	4.02	3.82	1.18

co-ordinates using these different sets of calibration constants, and is the reconstruction-only analogue of table 9.25. The behaviour is in line with expectations: the RMS differences are much smaller than the error estimates but are roughly in the same proportions, and increase as the fractional change in number of calibration runs used increases. RMS differences are also a reasonable guide to the magnitude of error contribution from the uncertainties on the calibration constants.

10.2 Advanced Calibration Methods

10.2.1 Calibration Constant Error Propagation

The minimisation process can be modified to allow for the fact that the calibration constants are not known exactly. This is achieved by letting the calibration ‘constants’ appear as in the parameter vector and treating the calibration output values as measurements of these constants:

$$\begin{pmatrix} camera\,location\,measurement_{camera\,1,x} \\ \vdots \\ cm_{4,rz} \\ beam\,hit\,position_{run\,1,camera\,1,x} \\ \vdots \\ hr_{n,4,y} \end{pmatrix} = \mathbf{F}. \begin{pmatrix} camera\,location_{camera\,1,x} \\ \vdots \\ c_{4,rz} \\ Unit\,reconstruction_{1,x} \\ \vdots \\ r_{n,rz} \end{pmatrix} . \qquad (10.1)$$

The uncertainty on the calibration constants is included in the weighting matrix, and will correctly propagate into the reconstruction errors themselves. Note that the matrix block corresponding to the calibration constants does not have to simply be the diagonal of $1/{\sigma_i}^2$, but is in general the inverse of the relevant sub-matrix of the full output covariance matrix from the calibration.[13] Thus, correlations between the calibration 'constants' discovered during the calibration will also be propagated. The reconstruction parameter error estimates then correctly account for the calibration constant uncertainties. A simple approach of repeated minimisations while individually adjusting the constants to their extremes of range is equivalent to this method in the case that the calibration constant covariance matrix is diagonal. When applied to the reconstruction, assuming uncorrelated input calibration constant errors leads to slightly pessimistic error estimates compared to the full propagation as the correlations cancel some of the errors. Table 10.4 shows the mean error estimates produced for the same set of 20 runs using the fixed constant, uncorrelated and full error propagation methods; it demonstrates the underestimate of errors when assuming perfect constants and the overestimate of errors when ignoring correlations.

Figure 10.1 shows a graphical representation of $\mathbf{A^T P A}$ at minimum (top) and a visualisation of the correlation matrix that results after the normalisation process of section 8.4.2 (bottom). Visible in the correlation matrix and hence present in the covariance matrix, are correlations with the calibration constants. Note that the calibration constant error matrix is almost identical to the sub-matrix produced by the calibration (figure 9.32). It is not actually identical due to an 'autocalibration' effect caused by the overdetermined nature of the problem.

10.2.2 Autocalibration

If a system making a measurement is overdetermined, i.e. it takes more readings than the minimum needed to produce each measurement, the extra information taken when making measurements can pull the calibration 'constant' values away from their calibration 'measurements'. This has

---

[13] The weighting matrix reduces to a diagonal matrix of the $1/{\sigma_i}^2$ values in the case of uncorrelated constants.

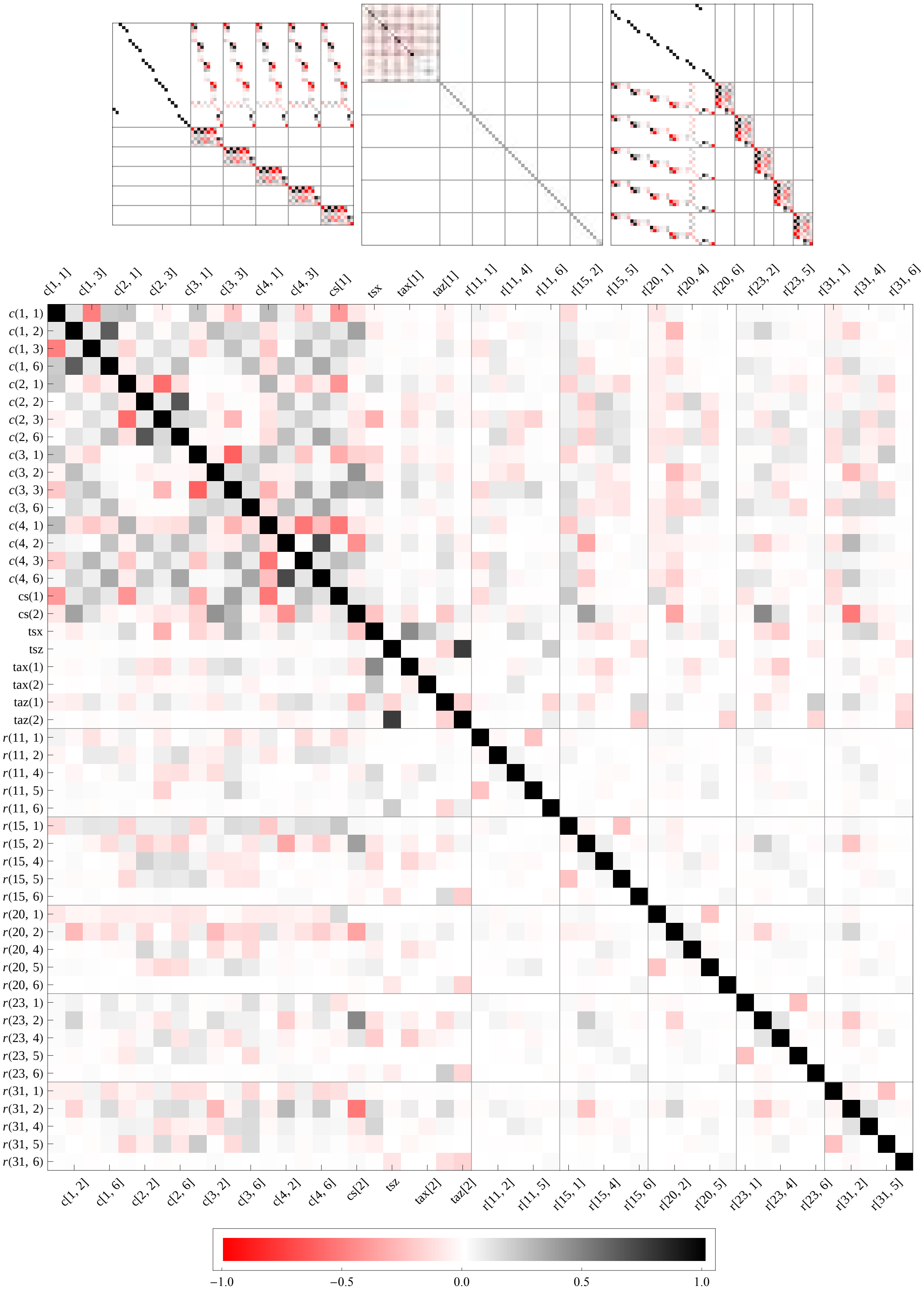


Figure 10.1: Graphical representation of $\mathbf{A}^{\mathrm{T}}\mathbf{PA}$ (top) at the minimum for a five run reconstruction using CCD and tilt sensor data. Calibration constants are included as parameters with an associated error sub-matrix. The scale is non-linear to highlight small contributions. After including constraints, inverting, extracting the error sub-matrix and normalisation, this becomes the the correlation matrix (bottom).

Table 10.4: Mean parameter error estimates for a 20 run reconstruction, for differing methods of calibration constant error propagation: assuming perfect constants, assuming uncorrelated errors, and a full input error matrix. Calibration was performed using a separate 25 runs of data.

Parameter	Perfect	Uncorrelated	Full
r[i,1] (μm)	5.5	6.0	5.8
r[i,2] (μm)	4.3	5.3	5.3
r[i,4] (μrad)	25.6	27.8	27.6
r[i,5] (μrad)	32.3	35.0	34.1
r[i,6] (μrad)	47.7	53.1	51.1
χ^2/DoF	0.587	0.519	0.541

the effect that if a batch of reconstructions are made together, then each run can contribute to the calibration constants for the reconstruction as a whole; updating both their values and covariance matrix elements. However, this does mean that the reconstructed co-ordinates for a given set of input beam spot positions are no longer unique; the same set would give slightly different reconstruction values depending on other runs reconstructed at the same time. This can be seen in figure 10.1 where in contrast to the simple reconstruction (see figure 8.5), there are slight correlations between different runs. This loss of constancy can be considered undesirable, but can be avoided by fixing the values of the constants during minimisation, producing $\mathbf{X}_{min,fixed}$. The error matrix can then be calculated separately for $\mathbf{X}_{min,fixed}$ with the constraints holding the constants at their fixed values removed.

If the user can accept the adjustment of individual reconstructions when seemingly unrelated runs are added, this autocalibrating effect can potentially be utilised to improve the values and errors for the calibration constants and the reconstructed co-ordinates. If this adjustment is not desired the process can still be used: nominated batches of runs, possibly taken especially for this purpose, could be used to give improved calibration constant values and covariances; the calibration constants could then be fixed at these improved values and used for further reconstructions, while the covariance matrix could be used to generate the input weighting matrix for the final error propagation discussed above.

This effect has yet further potential: for example, if the values of the constants drift over time, or dramatically after a disruptive event, the impact will be reflected in the new values of the constants produced as part of the reconstruction. This would partially mitigate any offsets in the reconstructions that might have been caused by using out of date constants. The values of the constants themselves could be monitored to see the change over time; if they deviate beyond a chosen amount from the original values, the system could be recalibrated.

10.2.3 Combined Calibration and Reconstruction

The most comprehensive way to include all information is to combine the calibration and reconstruction into one large minimisation. If the calibration and reconstruction problem is not easily factorisable, this would be the only fully valid approach. However, this comes at the cost of requiring significantly more computing resources, particularly memory. The results of this combined calibration and reconstruction are compared to the other methods of calibrated reconstruction in section 10.2.4.

10.2.4 Behaviour of New Methods

To analyse the effect of autocalibration, and combining calibration with reconstruction, part of the data from the calibration run can be used to calibrate, while the rest of the data can be used to reconstruct. It is important not to use the same data for reconstruction and calibration; otherwise the measurements taken by the CCDs and tilt sensors will be double counted in the overlapping runs.

Table 10.5 shows the effect on error estimates of including extra runs; either as part of the calibration data (the 'Extra Calibration' columns, previously discussed in section 9.10), as reconstruction data included as part of the calibration (the 'Combined' columns), or as a separate reconstruction minimisation with floating constants (the 'Autocalibration' columns). The effect of increasing the runs in the extra calibration is fairly uniform for all parameters. However, when the extra runs are added for reconstruction only, many constant errors remain unaffected. The unaffected constants are all the parameters related to the laser tracker and calibration platform (l[_], b[_] & m[_,_], with _ standing for all indices), the Z-rotation tilt sensor parameters (tsz & taz[_]), and the camera scale parameters (cs[1] & cs[2]). This lack of effect is expected, as without laser tracker data there is no way to find these parameter values. More interestingly, the reduction in error estimate for the camera parameters is less uniform; the c[1,_] constants fall at about 75% of the rate in the full calibration sets, but the c[2,_] and c[3,_] constant errors, particularly in the X and Z axes, barely reduce. The c[4,_] constants behave more similarly to the c[1,_] constants, but only drop at 50% of the rate of the full calibration sets. The reason for the asymmetry is not fully understood, though it is likely that CCD 0 and CCD 3 [14] can

---

[14] c[1,_] and c[4,_] correspond to CCD 0 & 3 when referenced outside of Mathematica

autocalibrate with each other as their virtual images are in almost the same space, whereas the outer cameras are sensitive to the added effect of rotations with only the X-rotation tilt sensor to help constrain them. Conversely, the X-tilt sensor error decreases as it is starting from a much larger absolute value; in effect, the camera measurements continue to calibrate the X-tilt sensor.

The errors found using the autocalibration method behave in a similar manner. In fact, they are almost identical to the results of the combined method to two significant figures; the only difference is a single value different by 1% for the tilt sensor scale in a single run. This is encouraging as it shows that the results of the calibration effectively capture all the information given, and that it is therefore appropriate to factorise the problem into calibration followed by reconstruction with negligible loss of information. This is useful as the full calibration analysis takes a significant amount of time (hours), while the reconstruction is far faster (minutes). The factorisation is also interesting as it allows the combination of the results of full calibration with possible multiple auto-calibrations. The covariance matrix used to generate the input calibration constant weighting matrix for the autocalibration was compared with the output covariance matrix by normalising both and observing the change in correlations. About two thirds of the constants had a negligible change, while one third had a change between ±0.1.

Table 10.6 shows that the improvement in the errors on the constants only translates to a negligible improvement in the error estimates on the reconstructed co-ordinates. This is consistent with the negligible improvement in co-ordinate reconstruction errors observed when altering the number of runs used to calibrate (§10.1.1).

Similarly, table 10.7 shows the limited effect of autocalibration on the reconstructed co-ordinate values themselves. The data presented in table 10.7 is produced by selecting a set of ten runs to reconstruct. Reconstruction is performed using a set of fixed constants from a 45-run calibration. These ten reconstructions are then treated as the reference values[15]. Then, the ten reconstructions are made using a fixed set of calibration constants from a 25-run calibration. Next, the ten reconstructions are performed allowing the constants to float (autocalibration). Finally, a set of 20 runs containing the original 10, are reconstructed; the superfluous runs are discarded and the results for the 10 common runs are compared. In principle, the difference between the reconstructed co-ordinates and the 'reference' co-ordinates should reduce as more data is used for autocalibration, leading to reduction in both the mean difference and the RMS difference in the parameter values. However, the only noticeable reduction in both the RMS and mean residual is

---

[15] If the assumption is made that the 45-run calibration gives the best estimate of the underlying calibration constants, then the reconstructed co-ordinates found when using these values are the best estimates.

for the rotation around the X″-axis.

The lack of improvement in errors or differences demonstrates that the calibration of the unit has been effective, with limited scope for improvement, as the quasi-systematic errors on the beam spot position readings overwhelm any refinement of calibration constants.

10.2.5 Autocalibration with One Poor Constant

If one constant is known poorly but others are known well, then the autocalibration effect may help to define the poor constant. Reconstructions were performed using the 25-run calibration constant values and errors, but setting $c[1,1]$ to have a value of zero and an error estimate of 10mm, with all other elements in the $c[1,1]$ row or column of the covariance matrix set to zero. Table 10.8 shows the effect autocalibration has on the value of the constant, its error, and the most affected reconstructed co-ordinate for a single run; the reconstruction was performed with set of 1, 2, 5, 10 and 20 runs that were not used in the calibration as well as with the constants fixed. The 'calibration' column shows the 'true' value and error of $c[1,1]$ found from calibration. The '1 Full' column and '20 Full' column show the results of the reconstruction for $c[1,1]$ and $r[23,1]$ for the case where the full input covariance matrix has been used with no change to the $c[1,1]$ measurement value or covariance. The autocalibration effect works very well: even when reconstructing a single run in isolation, the $c[1,1]$ value is within five microns of the value produced by the calibration, with an appropriate error. As more runs are added the $c[1,1]$ error improves and the value remains close to the expected value. Similarly, the reconstructed co-ordinate most affected by changing $c[1,1]$, is more than 100μm away from the 'true' value with fixed constants, but moves to within 0.4μm with two or more runs reconstructed at the same time. The situation is artificial in that any constant that is found well by this process should also be found to good precision in a well designed calibration experiment as the calibration experiment would be even more overdetermined. However, if calibration constants are found by direct measurement of the components, the autocalibration effect demonstrated is useful for checking consistency and potentially finding the values of constants that might not be easy to measure well.

10.2.6 Uncalibrated Reconstruction

The behaviour of the uncalibrated unit can also be investigated if appropriate estimates of the nominal calibration constant uncertainties can be found. As there is no ensemble of units from which such errors could be derived statistically, the only option is to compare the calibrated values with the nominal values and assume the difference is representative.

Table 10.5: Percentage calibration constant error compared with 25-run values, for extra calibration, combined calibration and reconstruction, and discrete reconstruction with floating constants.

	Constant	Extra calibration				Combined (25 +)				Autocalibration (25+)			
	Runs	30	35	40	45	5	10	15	20	5	10	15	20
Laser Tracker	l[1]	95	86	79	76	100	100	100	100				
	l[2]	96	86	78	75	100	100	100	100				
	l[3]	95	91	88	85	100	100	100	100				
	l[4]	92	86	81	77	100	100	100	100				
	l[5]	93	89	85	82	100	100	100	100				
	l[6]	92	86	81	77	100	100	100	100				
Base	b[1]	95	86	79	76	100	100	100	100				
	b[2]	96	86	78	75	100	100	100	100				
	b[3]												
Platform Markers	m[1,1]	95	86	79	76	100	100	100	100				
	m[1,2]	96	86	78	75	100	100	100	100				
	m[1,3]	91	84	79	75	100	100	100	100				
	m[2,1]	95	86	79	76	100	100	100	100				
	m[2,2]	96	86	78	75	100	100	100	100				
	m[2,3]	91	85	79	75	100	100	100	100				
	m[3,1]	95	86	79	76	100	100	100	100				
	m[3,2]	96	86	78	75	100	100	100	100				
	m[3,3]	91	84	79	75	100	100	100	100				
	m[4,1]	95	86	79	76	100	100	100	100				
	m[4,2]	96	86	78	75	100	100	100	100				
	m[4,3]	91	85	79	75	100	100	100	100				
CCD 0	c[1,1]	93	87	82	79	95	90	86	84	95	90	86	84
	c[1,2]	91	86	83	79	95	92	90	87	95	92	90	87
	c[1,3]	94	86	84	78	96	90	88	85	96	90	88	85
	c[1,4]												
	c[1,5]												
	c[1,6]	91	86	82	79	95	92	89	87	95	92	89	87
CCD 1	c[2,1]	95	92	90	87	99	99	99	99	99	99	99	99
	c[2,2]	94	91	89	87	98	96	95	94	98	96	95	94
	c[2,3]	94	87	85	79	100	99	99	99	100	99	99	99
	c[2,4]												
	c[2,5]												
	c[2,6]	91	86	81	78	97	96	94	93	97	96	94	93
CCD 2	c[3,1]	95	90	87	84	99	98	97	97	99	98	97	97
	c[3,2]	93	89	86	83	97	95	94	93	97	95	94	93
	c[3,3]	94	87	84	79	100	99	99	98	100	99	99	98
	c[3,4]												
	c[3,5]												
	c[3,6]	90	84	81	76	98	96	95	93	98	96	95	93
CCD 3	c[4,1]	93	88	84	81	95	92	89	87	95	92	89	87
	c[4,2]	91	86	83	78	97	95	94	92	97	95	94	92
	c[4,3]	94	86	83	77	97	92	91	88	97	92	91	88
	c[4,4]												
	c[4,5]												
	c[4,6]	91	86	82	79	97	95	93	92	97	95	93	92
Scale	cs[1]	92	89	85	83	100	100	100	100	100	100	100	100
	cs[2]	87	79	75	68	100	100	100	100	100	100	100	100
Tilt Sensors	tsx	92	82	79	73	95	88	86	83	95	88	**87**	83
	tsz	95	85	78	76	100	100	100	100	100	100	100	100
	tax[1]	92	85	83	78	93	88	86	82	93	88	86	82
	tax[2]	93	82	77	74	94	83	78	76	94	83	78	76
	taz[1]	93	82	80	74	100	100	100	100	100	100	100	100
	taz[2]	91	82	74	70	100	100	100	100	100	100	100	100

Table 10.6: Error estimates for reconstructed parameters for reconstructions using floating constants with increasing numbers of runs, each a superset of the last. Only the first five runs are considered for the average to maintain comparability and the results are given to excessive precision to show the very slight improvement.

Parameter	5	10	15	20
r[i,1] (μm)	5.855	5.853	5.853	5.851
r[i,2] (μm)	5.495	5.492	5.491	5.488
r[i,4] (μrad)	27.422	27.354	27.323	27.287
r[i,5] (μrad)	34.161	34.131	34.113	34.096
r[i,6] (μrad)	50.325	50.324	50.323	50.322

Table 10.7: Root mean square and mean residuals between a common set of ten reconstructions using fixed 45 run calibration constants and constants from a 25-run calibration: reconstruction with fixed constants (first column), 10-run reconstruction with floating constants (second column), and 20-run reconstruction with floating constants (third column).

Parameter	RMS Residual			Mean Residual		
	25 fixed	25+10	25+20	25 fixed	25+10	25+20
r[i,1] (μm)	1.40	1.37	1.42	0.52	0.51	0.52
r[i,2] (μm)	1.36	1.35	1.46	0.24	0.26	0.26
r[i,4] (μrad)	5.00	4.66	4.11	0.57	0.42	0.33
r[i,5] (μrad)	4.11	3.31	4.15	-0.97	-0.98	-1.01
r[i,6] (μrad)	7.63	7.53	7.52	3.40	3.37	3.36

Table 10.10 shows the difference between the calibrated values and the nominal values ('Difference' column). The differences were then adjusted to remove collective camera motion ('Adjusted' column): the co-ordinate system for the calibrated unit is unrelated to that of the nominal system. These difference values are averaged across CCDs and rounded to give a best guess at the build tolerances ('Estimate' column).

The values from the 'Estimate' column were then used as input errors for the calibration constants in the reconstruction, this gave a reasonable estimate of the performance of an uncalibrated system. Addressing the autocalibration effect first, the final two columns of table 10.10 show the output errors of the floating constants for a reconstruction of a single run, and for a batch of 20 runs. Some errors show an immediate improvement with only a single run, caused by the constraints of the reconstruction minimisation; most notably a fifty percent reduction in the CCD X and Y position errors. The effect of the autocalibration on the constants is evident between the

Table 10.8: The effect of autocalibration on the output value and error estimate of a calibration constant $c[1,1]$ and an affected reconstructed co-ordinate from an arbitrarily chosen run, as autocalibration is enabled and as more runs are reconstructed simultaneously. For the reconstruction, the $c[1,1]$ calibration constant input value was set to zero with a 1cm error estimate and zero correlations. The other input calibration constant values and covariances were from the 25-run calibration. The 'Fixed' column is for a reconstruction where the calibration constants were fixed, whereas the columns '1' to '20' indicate reconstructions of varying numbers of runs with autocalibration enabled. The 'Calibration' column shows the constant properties the 25-run calibration produced, while the 'Full' columns shows the values produced for reconstructions using the full input calibration constant error matrix and values without alterations. All values are in μm. *Fixing the calibration constants gives an incorrect error estimate.

	Fixed	1	2	5	10	20	Calibration	1 Full	20 Full
c[1,1] Value	0	885.6	889.5	893.0	889.0	893.0	890.5	890.3	891.1
c[1,1] Error	10,000	17.11	12.26	7.89	5.85	4.52	3.83	3.80	3.22
r[23,1] Value	896.53	785.23	784.8	784.4	784.8	784.2	-	784.6	784.6
r[23,1] Error	5.45*	6.20	6.02	5.91	5.87	5.85	-	5.82	5.82

1-run reconstruction set and the 20-run reconstruction set, particularly for the translations along and rotations around the Z-axis, as well as the X-rotation tilt sensor. The reduction in calibration constant error values is not particularly significant for most parameters. Comparing the error reductions with the changes seen in table 10.5 reveals that fewer constants are affected; this has not been investigated in detail but is possibly caused by the large errors on the other constants such as the CCD scale factors which have no potential for improvement by autocalibration.

Table 10.9 shows reconstructed co-ordinate error estimates for a single run, and the average errors over 20 runs with different numbers of runs reconstructed in each batch. Unsurprisingly, the uncertainty on the calibration constants dominates the contribution from beam spot measurement errors, giving rise to much larger uncertainties than the values of the calibrated unit in table 10.4. Table 10.9 also demonstrates the limited effect on reconstruction error estimates of autocalibration in this case. This is consistent with the limited change in the calibration constants errors.

Given the geometry of the LSM, autocalibration appears to only act to 'even out' calibration constant errors - if all are large, then there is little scope for improvement using autocalibration. In addition, in normal day-to-day usage, where the phase space of co-ordinates is constrained by other sub-systems like the internal FSI, there is even less scope for improvement.

Table 10.9: Reconstructed co-ordinate error estimates for an uncalibrated unit when reconstructing in different sized batches. The left-hand set is the error estimate on run 23 only for reconstructions with between 1 and 20 different runs in a batch. The right hand set is the mean error estimate for all 20 runs, with reconstructions performed in 4 batches of 5, 2 batches of 10, and a single batch of 20.

Parameter	Run 23 only					All 20 runs		
	1	5	10	15	20	$4 \times \mathbf{5}$	$2 \times \mathbf{10}$	$1 \times \mathbf{20}$
r[i,1] (μm)	142.9	143.0	143.1	142.8	142.6	134.6	135.1	134.4
r[i,2] (μm)	125.2	125.2	125.2	125.2	125.2	125.1	125.1	125.1
r[i,4] (μrad)	574.2	573.2	573.2	572.8	572.9	573.3	573.6	573.3
r[i,5] (μrad)	886.2	879.8	875.2	874.2	874.	883.5	897.7	878.2
r[i,6] (μrad)	10010.	9996.	9994.	9997.	9996	10009	10008	10005

Table 10.10: Difference between calibrated values from the 45/25 calibration and nominal values. The 'Adjusted' column has had collective motion removed while the 'Estimate' column contains the final best estimate. The two right-hand columns show the values of the constants that result when the 'Estimate' errors are used as part of an autocalibrating reconstruction using 1 run or 20 runs.

Parameter	Units	Difference	Adjusted	Estimate	1 run	20 runs
c[1,1]	μm	889	211	250	126.	125.
c[1,2]	μm	-428	335	250	126.	125.
c[1,3]	μm	-4800	-2500	2000	2000.	1360.
c[1,4]	μrad					
c[1,5]	μrad					
c[1,6]	μrad	1100	1100	5000	5000.	3240.
c[2,1]	μm	398	280	250	217.	215.
c[2,2]	μm	-25	68	250	170.	170.
c[2,3]	μm	-1000	1300	2000	2000.	1790.
c[2,4]	μrad					
c[2,5]	μrad					
c[2,6]	μrad	1900	1900	5000	5000.	3690.
c[3,1]	μm	493	-185	250	217.	215.
c[3,2]	μm	-262	-168	250	170.	170.
c[3,3]	μm	-30	700	2000	2000.	1790.
c[3,4]	μrad					
c[3,5]	μrad					
c[3,6]	μrad	-11600	-11600	5000	5000.	3270.
c[4,1]	μm	932	254	250	126.	126.
c[4,2]	μm	344	437	250	125.	125.
c[4,3]	μm	-400	1900	2000	2000.	1250.
c[4,4]	μrad					
c[4,5]	μrad					
c[4,6]	μrad	4000	4000	5000	5000.	2950.
cs[1]	-	0.0429	0.0429	0.05	49800.	48800.
cs[2]	-	-0.0023	-0.0023	0.0025	2500.	2500.
tsx	-	0.077	0.077	0.075	74900.	11900.
tsz	-	0.060	0.060	0.075	75000.	74100.
tax[1]	μrad	730	730	750	573.	572.
tax[2]	μrad	4780	4780	5000	5000.	2080.
taz[1]	μrad	-16000	-16000	20000	20000.	20000.
taz[2]	μrad	8966	8966	10000	10000	9990

11 Summary and Future Work

11.1 Summary

The International Linear Collider must be aligned precisely against a reference network to reach its luminosity target. Mapping this network rapidly and accurately is a difficult problem for which an automated system may be a good approach.

A prototype automated tunnel mapping system, the LiCAS RTRS, has been commissioned in a test tunnel in DESY. A core component of the RTRS is a Laser Straightness Monitor (LSM) which is used to reconstruct the positions and orientations of the units relative to each other in 4 degrees of freedom. Tilt sensors and a system of FSI lines measure the remaining translation along, and rotation around the longitudinal axis.

This work focusses on the design, component performance, laser beam spot position location, reconstruction method, calibration method, calibration behaviour, and final reconstruction performance of the LSM. I was exclusively responsible for all work after the initial design, including writing the 100,000 lines of code required for fitting, reconstruction, calibration and analysis.

I also greatly assisted in the assembly, installation, commissioning, post-installation fixes and improvements, maintaining and improving DAQ software, improving motor-control software, general calibration, and day-to-day usage of the RTRS; these contributions are not covered in this thesis.

11.1.1 Hardware Implementation and Usage

The RTRS is a 3-car train of measurement units connected by a vacuum tube, with a laser beam travelling along its length and back (§3.1). Each measurement unit is housed in a motorised cage which provides 6D motion (§3.3.1), enabling the units to be adjusted to counteract irregularities in the monorail along which the RTRS travels (§3.7). Each measurement unit reconstructs the position of a wall marker relative to itself using Frequency Scanning Interferometry (§2.3.1). These local reconstructions need to be combined into a single frame, meaning the relative positions and orientations of the measurement units must also be found.

The LSM enables the reconstruction of four of the six degrees of freedom of each unit by sampling the laser beam and measuring the beam spot positions on multiple cameras (§4.1). The low-coherence laser source (§4.2) is partially reflected off pellicle beam splitters (§4.6) onto CCDs which have had their performance carefully analysed (§4.7). A sophisticated and flexible DAQ and train control program enables easy production and use of operational scripts remotely (§5).

11.1.2 Beam Spot Fitting

A reliable method of finding the positions of beam spots hitting CCDs has been developed using least-squares fitting (§6.7), giving a best-case precision of $40 - 100$nm (§7.1.2). This is overwhelmed by electronic interference from the multi-channel analogue frame grabber (§7.1.1) and laser beam jitter, which is countered by careful temporal interpolation or extrapolation of fits from multiple images; this gives a repeatability precision of around 0.5μm for outgoing CCDs, rising to 3μm for return CCDs with poor quality beam spots (§7.2). Significant further issues are caused by position-dependent systematic errors in the fits of approximately $2 - 10\mu$m, primarily caused by dust and edge effects (§6.10).

11.1.3 Reconstruction

A method of reconstructing measurement unit positions and orientations has been developed, using least-squares minimisation to match the output of a ray-tracing model with the CCD data; either tracing a replication of the physical LSM geometry, a 'virtual' model, or a simplified virtual model. The first two methods were implemented in C++(§8.3) and the third method in Mathematica (§8.4). These methods have been shown to give identical results when used under the same conditions (§8.5), giving statistical-only errors of 0.15μm and 1.0μrad when reconstructing with typical data (§8.5.2). A method of visually and numerically inspecting the behaviour of the linear equation system in Mathematica has been developed based on the $\mathbf{A^TPA}$ matrix; its eigensystem (§8.4.3), its inverse and its normalised inverse (§8.4.2). Tilt sensors have been modelled in Mathematica and included in the minimisation (§8.6).

11.1.4 Calibration

A witnessed calibration procedure has been developed (§9.1), combining LSM and tilt sensor data and measurements from a witnessing system (§9.6) to give the calibration constants for a single unit at a time. For this purpose, a calibration platform was attached to the unit with four retro-reflector nests that were measured using a laser tracker.

The LSM data was found to require correction to reduce the effects of beam drift and jitter (§9.3), producing averaged measurements compatible with the 90 second time-scale of the measurement of the calibration platform by the laser tracker (§9.3.5). This gave a statistical error on the beam spot positions for the calibration experiment of about 0.3μm (table 9.4).

Data consistency was checked, showing that the statistical errors generated by the projections underestimate the overall beam spot position uncertainty (§9.4).

The correction was confirmed to cancel beam walk over longer periods of time (§9.5.2); when applied to a dataset taken over 35 hours with a stationary unit and beam walk totalling 700μm, correction left error estimates of 0.55μmand 0.41μm though with some irreducible structure (fig. 9.15). With a mobile unit giving beam spot positions over a significant range of the CCDs (§9.2.1), the systematic errors from dust and edge effects were found to act as random offsets, so were propagated statistically as 'quasi-systematic errors' (§9.5), leading to a best estimates for the overall beam uncertainties of between 7μm and $16\mu m$ (table 9.16).

After assessing the performance of the laser tracker (§9.6.1) and the identification of three possible error paradigms, a model of laser tracker measurements of the platform was produced (§9.6.5). Least-squares minimisation was used to reconstruct the co-ordinates of the unit/platform system, giving error estimates of $1 - 10\mu$m and $20 - 60\mu$rad (table 9.9). For minimisation, axes were defined and symmetries caused by the system geometry were identified and constrained (§9.6.7).

The platform reconstructions were used to correct the data further, remove bad runs and refine the beam propagation direction (§9.7), reducing r.m.s. beam spot position residuals from $20 - 50\mu$m to $8 - 20\mu$m.

The LSM, tilt sensor, and platform models were combined to produce a comprehensive model (§9.8.1). Near symmetries were identified and constrained (§9.8.2) and an initial calibration iteration was performed using a least-squares minimisation. The results from this first step were used to correct the data further and enabled the selection of a laser tracker error paradigm (§9.8.3), giving final r.m.s. residuals of $7 - 16\mu$m (table 9.16).

A second iteration of the minimisation using the data with the best corrections produces the final calibration results. The most important component positions were found with uncertainties of $2 - 3\mu$m (§9.9.2). The calibration procedure has been found to be robust (§9.9.3,§9.9.4), with the results changing in line with expectations as the amount of input data is varied: variations of the constants of order 1μm and a $\sqrt{n}$-like behaviour of the calibration constant error estimates. (§9.10). A method of finding contributions to parameter errors from the geometry of the system by analysing the covariance matrix has been demonstrated (§9.9.5).

11.1.5 Calibrated Reconstruction

The importance of correctly accounting for calibration constant errors in reconstruction has been demonstrated (table 10.4) and a method of correctly propagating such errors has been shown (§10.2.1).

Altering the number of runs used for calibration was shown to have a knock-on effect on the reconstruction. Reconstructed co-ordinates changed in a way consistent with expectations: a typical r.m.s. difference between sets was of order 1μm (§10.1.1). There was a negligible change in the reconstructed co-ordinate error estimates.

A 'combined' method of calibration and reconstruction in one large minimisation is introduced (§10.2.3). This has been shown to behave the same as a 'factorised' autocalibration method which calibrates and reconstructs separately, but preserves information about the calibration constants using their full covariance matrix (§10.2.4). This method allows the calibration constants to 'float' and be adjusted as part of the unit co-ordinate reconstruction process, allowing the overdetermined nature of the LSM to contribute by increasing self-consistency.

When using the autocalibration method or combined calibration and reconstruction, a reduction in calibration constant error estimates of typically $0 - 10\%$ was observed as more runs were reconstructed together (table 10.5). However, there was a negligible impact on the reconstruction error estimates (table 10.6). The co-ordinate values generated by reconstructions when using autocalibration change as more reconstructions are performed at once, but have not been demonstrated to shift towards values expected from a more comprehensively calibrated unit (table 10.7). However, there may be potential for autocalibration be used to monitor the stability of calibration constants over time (§10.2.2). Autocalibration was shown to work well when a single calibration constant was set to be very poorly defined, which may potentially be of use if calibration is performed by incomplete component measurement rather than a calibration experiment (§10.2.5).

The final performance of a single unit reconstruction is given in table 11.1 and is representative of the expected behaviour of the other units. The translation into overall tunnel marker reconstructions at a single stop and eventually for a complete tunnel is beyond the scope of this work.

11.2 Observations, Suggestions and Potential Future Work

While the principle of using an LSM as part of an automated alignment system has been shown to be sound, a production version would end up looking very different from the prototype system.

Table 11.1: Mean LSM reconstruction error estimates

Co-ordinate	Parameter	Total Error
X-translation	r[i,1] (μm)	5.8
Y-translation	r[i,2] (μm)	5.3
Rx-rotation	r[i,4] (μrad)	27.6
Ry-rotation	r[i,5] (μrad)	34.1
Rz-rotation	r[i,6] (μrad)	51.1

11.2.1 RTRS Construction

Observations and suggestions for a future RTRS:

- Systematic errors in beam position measurements are the limiting factor in LSM performance.
  - Use an LED or superluminescent diode as the light source. This would significantly reduce the various problems caused by interference, such as interference patterns and significant beam splitter reflectivity variation.
  - Effort should be made to produce a ‘clean’, stable Gaussian beam with no ghosts on CCD or electronic interference. This would include ensuring dust cannot enter the system during use, having large apertures, component positioning to avoid unwanted reflections appearing on CCDs, and stable launch optics.
  - Large CCD surface areas are recommended; pixel noise performance and resolution are secondary. This will increase the range without biases caused by edge effects.
  - Returning each unit to the same position and orientation for each wall-marker measurement would significantly reduce the importance of some systematic errors. This would require a larger active spatial range of the external FSI subsystem when measuring the wall markers.
- Maintaining a vacuum added enormous mechanical and operational complexity to the RTRS.
  - If the inner volume was filled with a low refractive index gas such as helium, which has a refractivity 12% that of air [62], this may reduce thermal systematics to acceptable levels. With a thermally conducting pipe, the temperature inside would be more

  uniform than exposed air too. A positive pressure would significantly reduce the problem of dust entering the system.
  - If a vacuum is required, using end-caps attached to units is not recommended due to the large forces exerted by air pressure acting unevenly against the motor cage. Separate, rigidly held launch and retro assemblies would be superior, though increase the number of degrees of freedom in the system and may complicate alignment further.

- Heavy invar measurement units also added to mechanical complexity.
  - Lighter units, possibly made out of carbon fibre, could be mounted onto hexapods, greatly reducing the complexity of the unit motion.

- Initial alignment of the RTRS at stop can be quite difficult as the LSM beam typically did not reach the next car. Furthermore, an arrangement that produces a central beam on the car 2 outgoing cameras may produce a beam out of the dynamic range of the final car.
  - Increasing the aperture and dynamic range of the units would reduce this problem.
  - The addition of a secondary, diverging beam that could be swept over all possible angles would aid initial location, while more refined and co-ordinated unit motor control would aid adjustment to find a valid multi-car arrangement.
  - The coupling of the return beam back down the launch optics can be measured and used to aid the production of an anti-collinear beam.

11.2.2 Geometrical Observations

- Before producing a LSM, significant effort should be spent finding a suitable geometry using full simulation and analysis methods, with consideration of both normal use and calibration.

- A significant amount of over-determination aids cross-checking, allows improved analysis and reduces errors. It also gives the potential for auto-calibration.

- Symmetries resulting from the geometry of the systems introduce weaknesses into both calibration and reconstruction minimisations, increasing parameter uncertainties and difficulty in finding an appropriate minimum.
  - The system should be made geometrically irregular; for example skewing some CCD by 45 degrees and 90 degrees would have broken some of the camera X-scale factor

symmetries, while having a CCD with the incident beam significantly away from normal would break a tilt-and-scale symmetry.

- Angular sensitivity can be improved by increasing the effective separation of CCDs, possibly using internal mirrors to keep units small.
  - There is also potential for using the CCD faces themselves as semi-mirrors and measuring a single beam reflected off a pellicle at multiple distances, aiding both alignment and adding further redundancy.
- If a secondary reflecting surface is used, further sensitivity can be gained by making it convex rather than a plane to give greater beam motion, at the cost of careful analysis of non-uniformity.
- A four car train is necessary for tunnel reconstruction without relying on long-range Z-rotation sensor values.
- A single line of reference markers is not sufficient to give a robust network.
- Any production version is likely to have a much larger inter-car distance to reduce the scaling of systematic errors, compounding the construction issues previously covered.

11.2.3 Beam Fitting Suggestions

- A bias map could be produced experimentally for each CCD by moving a reference beam across its face and comparing the fit position to the known beam motion.
  - This would then correct for the position dependent offsets caused by edge effects.
- The beam spot model could be made more representative
  - The dust interference patterns could be included as part of the fit
  - The non-Gaussian elements of the beam could be modelled and accounted for
- A region of interest window that tracks the beam spot position would ensure the same number of pixels contribute to each fit, reducing edge effects.

11.2.4 Modelling Suggestions

- The ray-tracing models could be improved to model additional beam properties during propagation; including beam width, shape, interference and correctly finding modelling reflectivity would enable the identification of practical problems before a system was built.
- The calibration and reconstruction models could be improved to include other subsystems and calibration set-ups, such as exercising a unit while taking both LSM and FSI data.
- The Mathematica minimisation procedure can be improved to use more efficient minimisation techniques; for example the DFP algorithm used by Minuit (§B.7) or by using sparse matrices for larger problems. Avoiding analytical calculation of derivatives would give orders of magnitude performance improvements.

11.2.5 Calibration Suggestions

- There could be a factory calibration of each unit individually under controlled lab conditions, giving significantly more control over the experiment. This was not performed due to time constraints.
- The eigensystem analysis (§8.4.3) and covariance matrix analysis (§8.4.2) can be used to find weaknesses of the design, enabling the design of more effective LSM and calibration geometries. For example, the best place to position the laser tracker could be calculated.
    - The contributions to calibration constant errors and reconstructed co-ordinate errors from individual input errors can also be found by analysing the **C**-matrix further, enabling the specification of input parameter errors for desired output errors.
- Auto-calibration may have some potential for the following, particularly if the geometry of the LSM is designed with it in mind:
    - improvement in calibration constants using self-consistency if calibration is achieved by component measurement, particularly if some constants are measured to low accuracy.
    - the detection of components straying from their calibrated positions over time.

References

[1] J. Beringer et al. Review of Particle Physics. *Phys. Rev. D*, 86:010001, Jul 2012. doi: 10.1103/PhysRevD.86.010001. URL `http://link.aps.org/doi/10.1103/PhysRevD.86.010001`. 1.1

[2] New results indicate that particle discovered at CERN is a Higgs boson, March 2013. URL `http://press.web.cern.ch/press-releases/2013/03/new-results-indicate-particle-discovered-cern-higgs-boson`. 1.1

[3] N. Aghanim, C. Armitage-Caplan, et al. Planck 2013 results. I. Overview of products and scientific results. *arXiv preprint arXiv:1303.5062*, 2013. 1.2

[4] H. Baer, T. Barklow, et al. The International Linear Collider Technical Design Report-Volume 2: Physics. *arXiv preprint arXiv:1306.6352*, 2013. URL `http://www.linearcollider.org/ILC/Publications/Technical-Design-Report`. 1.3

[5] E. J. N. Wilson. *An introduction to particle accelerators.* Oxford University Press, 2001. 1.3.1, 1.3.3, 1.3.4, 1.4.1

[6] T. Behnke, J. E. Brau, et al. The International Linear Collider Technical Design Report-Volume 1: Executive Summary. *arXiv preprint arXiv:1306.6327*, 2013. 1.1, 1.3.2, 1.1, 1.4.2

[7] C. Adolphsen et al. The International Linear Collider Technical Design Report-Volume 3. II: Accelerator Baseline Design. *arXiv preprint arXiv:1306.6328*, 2013. 1.3.2, 1.5.1

[8] G. Arduini and et. al, editors. *Electron-positron collisions at 209 GeV in LEP*, 2001. 1.3.3

[9] W. Herr. Concept of luminosity in particle colliders. In *Intermediate Accelerator Physics Course*, 2010. 1.3.3, 1.3.4

[10] N. Phinney, N. Toge, N. Walker, and et al. ILC Reference Design Report, Volume 3: Accelerator. Technical report, ILC GDE, 2007. 1.4.3

[11] R. Brinkmann. TESLA Technical Design Report. Technical report, 2001. 1.5.1, 2.1.2

[12] J. Dale. *A Study of Interferometric Distance Measurement Systems on a Prototype Rapid Tunnel Reference Surveyor and the Effects of Reference Network Errors at the International Linear Collider.* PhD thesis, University Of Oxford, 2009. 1.5.1, 2.1.2, 2.3.1, 3.2.1, 3.4, 3.5, 5.2

[13] J. Gervaise. High precision survey and alignment techniques in accelerator construction. Technical report, SIS-74-2675, 1974. 2.1.1

[14] W. M. Welsch, editor. *Heft*, volume 20, October 1985. Internationale Federation of Surveyors. 2.1.1

[15] H. M. Durand et al. Validation of the CLIC alignment strategy on short range. 2012. 2.1.1

[16] P. Brockill. LiCAS-RTRS schematic and measurement procedure. Private communication, 2007. 2.1

[17] W. Schwarz, editor. *Concept for the Alignment of the planned Linear Collider at DESY*, Proceedings of the Fifth International Workshop on Accelerator Alignment, Advanced Photon Source, Argonne National Laboratory, October 1997. IWAA. 2.2.2

[18] J. Green. *Development of a Prototype Frequency Scanning Interferometric Absolute Distance Measurement System for the Survey & Alignment of the International Linear Collider.* PhD thesis, University Of Oxford, 2007. 2.3.1

[19] A. Herty and J. Albert. High precision survey and alignment of large linear colliders - horizontal alignment. Technical report, DESY, 2002. 2.3.2, 3.2.1

[20] G. Grzelak, A. Reichold, et al. The LiCAS-RTRS a rapid and cost efficient survey system for the ILC. *Proceedings of 9th International Workshop on Accelerator Alignment (IWAA 06)*, 2006. 3.1

[21] *Invar36 Technical Datasheet.* Carpenter Technology Corporation, 2011. URL `http://cartech.ides.com/datasheet.aspx?i=103&e=183&c=TechArt`. 3.2

[22] Schaevitz LSOP/LSOC, DC-Operated, Gravity-Referenced Inclinometer, 2006. URL `http://www.sherbornesensors.com/uploads/files/Files/LSOP-LSOC%27L%27-Iss3.pdf`. 3.2.1, 8.6

[23] C. Uribe-Estrada. ELMB setup for LiCAS. Private communication, June 2007. 3.2.2

[24] T. Kuphaldt. Kelvin (4-wire) resistance measurement, 2012. URL `http://www.allaboutcircuits.com/vol_1/chpt_8/9.html`. 3.2.2

[25] *ICLA IFA Intelligent Compact Drive Fieldbus servo motor technical documentation.* Berger Lahr, 2003. 3.3.1

[26] J. R. Cook and G. Thomas. *ELMB128 Documentation*, February 2005. 3.4

[27] *New Focus Motion control manual.* Newfocus, 2007. 3.6, 3.6, 4.3, 4.5

[28] *AC Servo Motors and Servo Rated Gearheads.* Baldor, 2003. URL `http://www.baldor.com/support/Literature/Load.ashx/BR1202-E?LitNumber=BR1202-E`. 3.7

[29] *Fiber-Coupled Laser Sources.* Schäfter and Kirchhoff, 2004. URL `http://www.sukhamburg.com/products/Fiberoptics.html`. 4.2, 4.3

[30] D. Jenkins. Pellicle Properties, 2009. URL `http://www.thorlabs.com/NewGroupPage9.cfm?ObjectGroup=898`. 4.6, 4.9, 4.6.2

[31] Y. Cui and R. M. A. Azzam. Determination of the refractive index and thickness of transparent pellicles by use of the polarization independent absentee-layer condition. *Appl. Opt.*, 35:5040–5043, 1996. 4.6.1

[32] E. Hecht. *Optics, Fourth Edition.* Addison Welsey, 2002. 4.6.2, 4.6.2, 6.3.3

[33] R. Bingham. Lens Effects of RTRS Pellicles. Private communication, 2005. 4.6.2

[34] *Pellicle BeamSplitters.* Thorlabs, 2009. URL `http://www.thorlabs.com/catalogPages/826.pdf`. 4.6.2

[35] Sony, 2001. URL `http://pro.sony.com/bbsc/ssr/cat-camerasindustrial/cat-ciindustrial/product-XCEI50CE/`. 4.7

[36] *Diagonal 6mm (Type 1/3) CCD Image Sensor for CCIR B/W Video Cameras.* Sony, 2001. URL `http://www.sony.net/Products/SC-HP/datasheet/01/data/E99526C3Z.pdf`. 4.7

[37] *IMAQ PXI/PCI-1409 User Manual.* National Instruments, June 2003. URL `http://www.ni.com/pdf/manuals/322811b.pdf`. 4.7

[38] *LabVIEW.* National Instruments, 2003. URL `http://www.ni.com/labview/`. 4.7

[39] Matrox, August 2004. URL `http://www.matrox.com/imaging/en/products/frame_grabbers/solios/solios_ea_xa/`. 4.7.4, 4.7.5

[40] Microsoft, April 2003. URL `http://msdn.microsoft.com/en-us/library/d06h2x6e%28v=vs.71%29.aspx`. 5.1.1

[41] N. Gray. Data is a singular noun, March 2012. URL `http://purl.org/nxg/note/singular-data`. 5.1.1

[42] G. Grzelak. Global Simulation Of The LiCAS/RTRS Survey System For The ILC. In *IWAA 2008*, 2008. URL `http://www-pnp.physics.ox.ac.uk/~licas/page_talks/IWAA2008/iwaa2008_grzelak_licas_sim_proc_ver2.pdf`. 6

[43] W. Rasband. ImageJ, 1997-2011. URL `http://imagej.nih.gov/ij/`. 6.1.3

[44] M. Frigo and S. G. Johnson. The Design and Implementation of FFTW3. In *Proceedings of the IEEE*, volume 93, page 216 to 231, 2005. 6.3.2

[45] J. Bowman. FFTW++: Fast Fourier Transform C++ Header Class for FFWT3 Library, 2005. URL `http://www.math.ualberta.ca/imaging/fftw++/`. 6.3.2

[46] Y. Ge, Q. Cheng, and S. Zhang. Reduction of edge effects in spatial information extraction from regional geochemical data: a case study based on multifractal filtering techniques. *Computers and Geosciences*, 2004. 6.3.2

[47] E. Hewitt and R. E. Hewitt. The Gibbs-Wilbraham phenomenon: An episode in fourier analysis. *Archive for History of Exact Sciences*, 21(2):129–160, 1979. ISSN 0003-9519. doi: 10.1007/BF00330404. URL `http://dx.doi.org/10.1007/BF00330404`. 6.3.2

[48] C. F. Gauss. *Theory of Motion of the Heavenly Bodies Moving about the Sun in Conic Sections.* Little, Brown and Co, 1857. 6.6

[49] F. James. *Minuit Tutorial, Function Minimisation*, 2004. URL `http://seal.web.cern.ch/seal/documents/minuit/mntutorial.pdf`. 6.6.3

[50] R. Brun and F. Rademakers. ROOT - An Object Oriented Data Analysis Framework. In *Proceedings AIHENP'96 Workshop, Lausanne.*, pages 81–86. Nucl. Inst. & Meth. in Phys., 1996. URL `http://root.cern.ch/`. 6.6.3

[51] A. Siegman. *Lasers.* University Science Books, 1986. 6.10.1

[52] F. J. Harris. On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform. *Proceedings of the IEEE*, 66(1):51–83, January 1978. 7.1.2

[53] *MATLAB Version 7.2.0.294 (R2006a).* Mathworks, Natick, Massachusetts, 2006. 7.1.2

[54] I. Wolfram Research. Mathematica 9.0, 2013. 8.1, 9.6.6

[55] G. Grzelak and B. Szczygieł. Virtual setups for the LiCAS LSM. Private Communication, 2005. 8.1.4

[56] H. Goldstein, C. Poole, and J. Safko. *Classical Mechanics.* Addison-Wesley, 3rd edition, June 2001. 8.2.1, D.2.1

[57] G. Grzelak. Simulation of the Performance of the LiCAS Train. In *IWAA 2004*, 2004. 8.2.1

[58] L. Brunel. SIMULGEO, Simulation and Reconstruction Software for Opto-Geometrical Systems. Technical Report CMS NOTE 1998/079, CERN, 1998. URL `http://cmsdoc.cern.ch/cgi-doc/Grep_Search_F/documents/allxx.html?STRING=NOTE1998079`. 8.2.1

[59] M. Woo, J. Neider, T. Davis, and D. Shreiner. OpenGL programming guide: the official guide to learning OpenGL, version 1.2, 1999. 8.3.3

[60] Leica Laser Tracker for Hand Tools, 2003. URL `http://www.leica-geosystems.com/media/new/product_solution/en_L3_LT%28D%29800.pdf`. 9.6.1

[61] M. Schlösser. Typical laser tracker measurement precision. Private Communication, 2008. 9.6.1

[62] G. Kaye and T. Laby. *Tables of Physical and Chemical Constants.* Longman Scientific and Technical, 1989. 11.2.1

[63] D. Wells and E. J. Krakiwsky. *The Method of Least Squares.* 1971. URL `http://gge.unb.ca/Pubs/LectureNotes.html`. B, B.3.2, B.5.1

[64] E. J. Krakiwsky. A Synthesis of Recent Advances in the Method of Least Squares. 1975. URL `http://gge.unb.ca/Pubs/LectureNotes.html`. Ref. [6] in SIMULGEO manual. B.5.1, B.5.1

[65] J. N. S. J. Write. *Numerical Optimization.* Springer, 1999. B.5.2

[66] R. Fletcher and M. Powell. A Rapidly Convergent Descent Method for Minimization. *Computer Journal*, 6:163–168, 1963. B.7, B.7

[67] M. Winkler, F. James, L. Moneta, and A. Zsenei. MNLineSearch, April 2005. URL `http://seal.web.cern.ch/seal/MathLibs/Minuit2/html/classROOT_1_1Minuit2_1_1MnLineSearch.html`. B.7

[68] R. Fletcher. A New Approach to Variable Metric Algorithms. *Computer Journal*, 13:317–322, 1970. B.7

A Dust Interference Analysis

A.1 Observed Interference Patterns

A speck of dust on a pellicle beam splitter causes an series of concentric rings to develop in the CCD image. To analyse the pattern produced, a typical image with a clear dust ring was fitted. (Shown in figure A.1, left.) An image with the fit and static background removed was then created and a region of interest (from (200,338) to (600,358)) decided (figure A.1, right). A plot of the sum of each of the columns in the region of interest can be found in figure A.2 clearly showing the oscillatory pattern.. The peaks were found using the ROOT TSpectrum peak finding algorithm, with the two central false positives and the poor peaks after column 350 removed. The process was repeated using the negative of the values to find the minima. Figure A.3 shows the radius of the for each ring away from the centre. Minima are counted has having half integral order. Figure A.4 shows the intensity of the peaks against the column number.

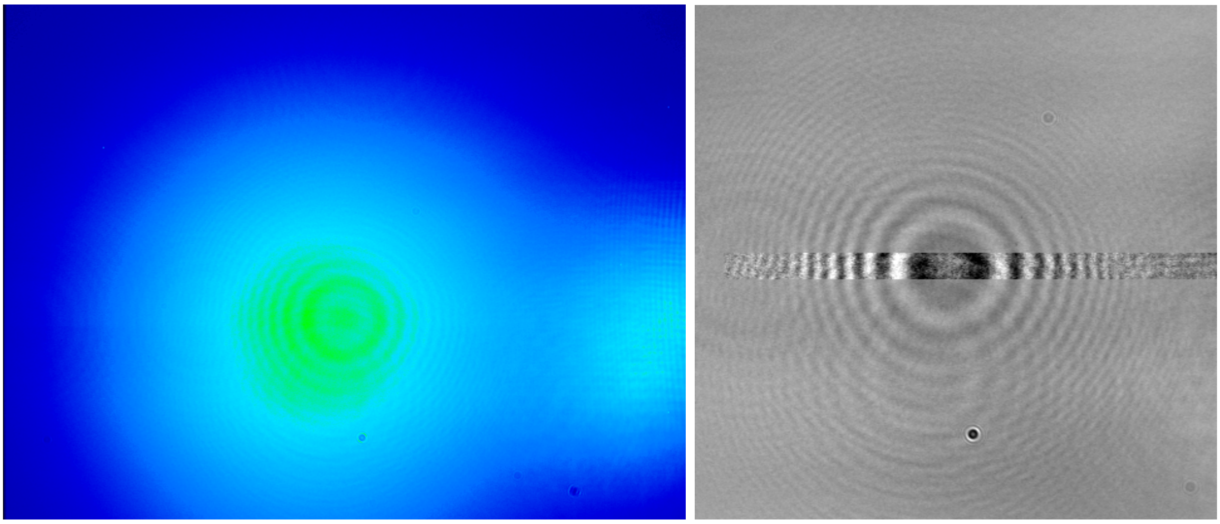


Figure A.1: A false colour image from car 1 CCD 3 with a clearly visible interference pattern. The right hand image shows part of the residual image after fitting making the interference pattern much clearer as well as bringing others to visibility. The region used for the intensity plot is highlighted. The small specks are due to dust on the CCD itself.

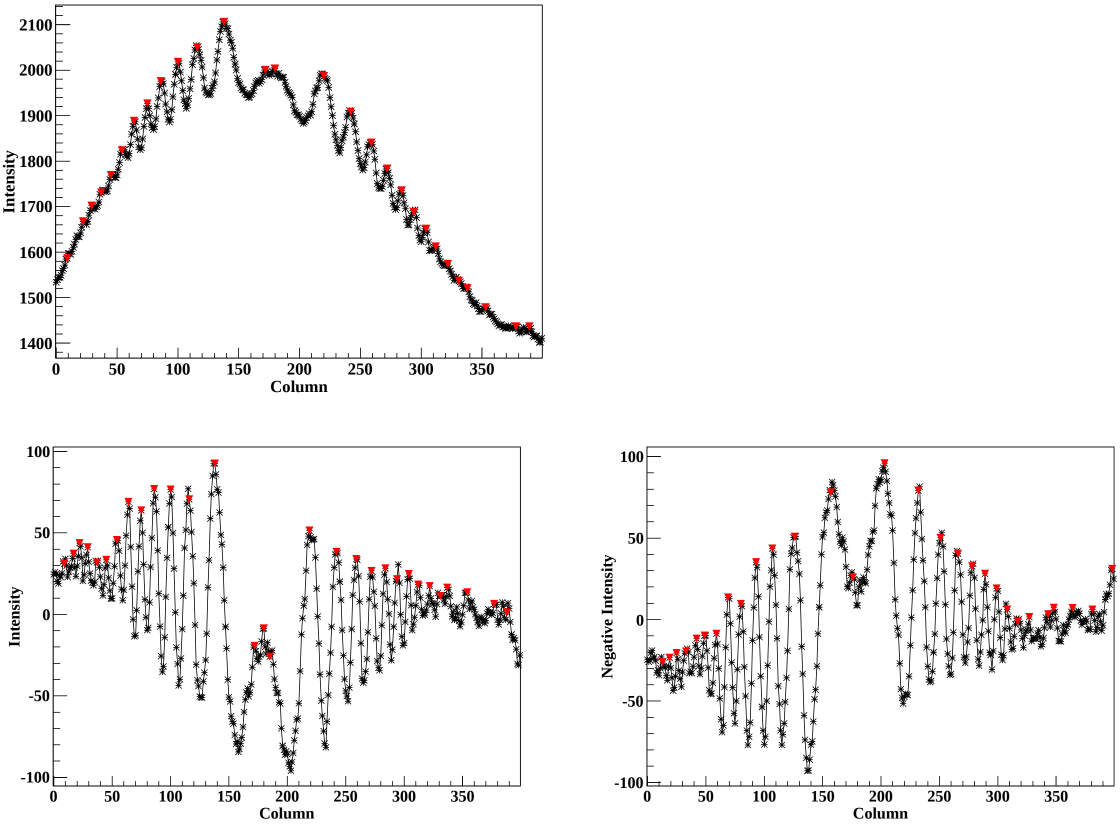


Figure A.2: The summed intensity of the columns of the strip through the interference pattern (top left), with the fitted beam subtracted (bottom left) and the inverted, subtracted plot (bottom right) for an image from car 1 CCD 3. The peaks found have been marked.

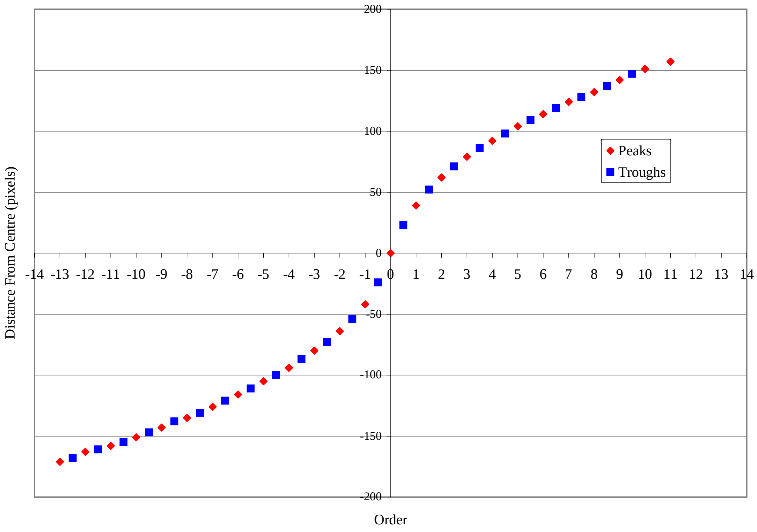


Figure A.3: The distance of the interference peaks to the centre of the beam against the 'order' of the fringe for an image from car 1, CCD 3.

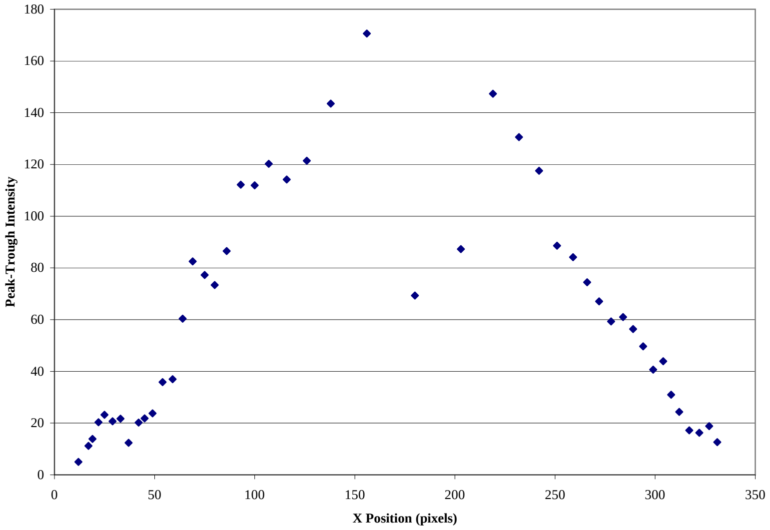


Figure A.4: Peak intensity plotted against distance from centre in pixels for an interference pattern from car 1, CCD 3. The phase shift causes the low central values.

A.2 Analysis of interference

The interference can be modelled as small particle scattering the incident light to give a spherical wave. This wave then interferes with the unscattered plane wave to produce rings.

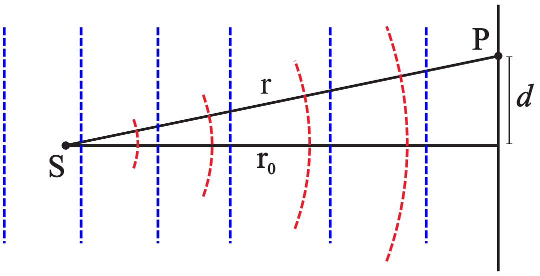


Figure A.5: A plane wave (blue) scattering off an imperfection (S) producing a spherical scattered wave (red).

Consider a plane wave hitting a piece of dust S, at the origin with a screen at distance r_0 in the direction of propagation of the plane wave. At point P, at perpendicular distance d the distance from the dust is r (figure A.5). The plane wave can be written as:

$$P = A_p e^{ikx-\omega t}\,, \tag{A.1}$$

where A_p is the real amplitude, k the linear wave number, x the X-co-ordinate, ω the angular frequency and t the time. The spherical wave can be represented with:

$$S = \frac{A_s}{r} e^{ikr+i\kappa-\omega t}\,, \tag{A.2}$$

where A_s is the real amplitude, r the distance from the origin, and κ is the phase shift is caused by the scatter. The amplitude at point P is therefore:

$$\Sigma = A_p \exp\left(ikr_0 - \omega t\right) + \frac{A_s}{\sqrt{r_0^2 + d^2}} \exp\left(ik\sqrt{r_0^2 + d^2} + i\kappa - \omega t\right) . \tag{A.3}$$

The intensity can then be found by taking the modulus-squared:

$$I = \frac{A_s^2}{r_0^2 + d^2} + A_p^2 + \frac{2A_sA_p}{\sqrt{r_0^2 + d^2}} \cos\left(k\sqrt{r_0^2 + d^2} + \kappa - kr_0\right) . \tag{A.4}$$

If $d \ll r_0$ then a binomial expansion produces the expression:

$$I = \frac{A_s^2}{r_0^2} + A_p^2 + \frac{2A_sA_p}{r_0} \cos\left(\frac{kd^2}{2r_0} + \kappa\right) . \tag{A.5}$$

Note that any phase shift from the scatter is still present. The maxima can therefore be found when:

$$\frac{kd^2}{2r_0} + \kappa = 2n\pi \, , \tag{A.6}$$

with n being the order number. Rearranging and introducing the wavelength gives the following expression of the ring radius:

$$d = \sqrt{2\lambda r_0 \left(n - \frac{\kappa}{2\pi}\right)} \, , \tag{A.7}$$

which clearly shows the square-root dependence on ring order seen in figure A.3. If the square of the ring radius is plotted against the order (see figure A.6 for the car 1, CCD 3 example), then the gradient of the fit gives the distance from the speck of dust to the CCD, according to the formula:

$$\frac{\mathrm{d}d^2}{\mathrm{d}n} = 2\lambda r_0 \, . \tag{A.8}$$

Table A.1 shows the results of this calculation for a dust ring from three different CCDs. There is some variation, but due to the angle of the beam splitter this is expected to be Y-dependent.

In principle, this effect could be used to help calibrate the unit by adding measurements of the pellicle-CCD distance. The process could be extended further by making three or more deliberate markings at the edge of the pellicles, outside the normal usage. If positioned accurately, they could then be used to determine the relative positions and orientations of the pellicle and its

Table A.1: Dust ring positions and calculations of the optical distance travelled, assuming a 783nm beam, for three different dust specks.

Car	Camera	X-centre	Y-centre	d^2 gradient	optical path
		mm	mm	mm^2	mm
1	3	0.01	0.51	0.1708	109.1
2	0	0.66	-2.38	0.1804	115.2
2	1	1.08	-1.70	0.1724	110.1

associated CCDs. This was not pursued in this work due to time constraints and the abstraction away of the pellicles in the virtual models of the unit described in section 8.1.4.

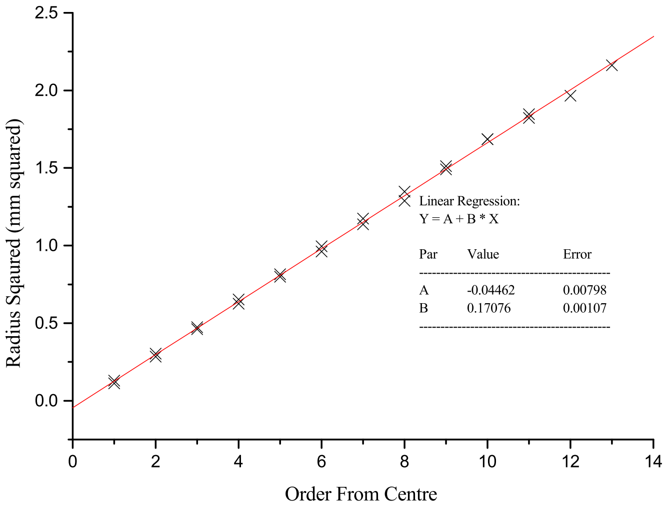


Figure A.6: The square of the radius of the maxima plotted against the order of the fringe for an image from car 1, CCD 3. The maxima for both directions in x are plotted together.

A.3 Dust on the CCD Surface

Specks of dust on the CCDs themselves also cause interference patterns in the same way. They all have a constant size of 3 pixels from the centre to the first peak but the low intensity and small size of the patterns preclude a detailed analysis; a second ring is not visible. These patterns are much less of an issue as there are a number of mitigating factors; the small size of the patterns mean they are very effectively removed by the Fourier filtering methods described in the section 6.3. Secondly, it is possible to clean the CCD surfaces in situ; while necessitating opening up of the unit and requiring a fair amount of care; touching the surface does not destroy it, unlike with the pellicles.

B Least-Squares Fitting

The method of least squares is used extensively throughout this investigation: point extrapolation and interpolation was performed using custom written C++ code; unit reconstruction and calibration using custom written Mathematica code; beam fitting and a second unit reconstruction framework using Minuit2 in C++ code. The greatest descent least squares method described here follows the procedure demonstrated by [63], while the method used by Minuit2 described in section B.7 is a quasi-Newtonian method.

B.1 Definitions

Consider the case of having a vector of measurements, **L**, and a mathematical model taking a vector of parameters, **F**(**X**). If the measurements have no errors and the model perfectly describes the measurements then the following is true for a particular **X**.

$$\mathbf{L} = \mathbf{F}(\mathbf{X})\,. \tag{B.1}$$

Note that **X** does not need have the same number of elements as **L**. In order for the problem to be solvable **X** needs to have fewer elements than **L** (though this is not sufficient). However, in the real world the measurements are not perfect and, assuming that there are more than the minimum for it to be solvable, the relation will not hold exactly. We can represent these residuals by the misclosure vector; **V**.

$$\mathbf{F}(\mathbf{X}) - \mathbf{L} = \mathbf{V}\,. \tag{B.2}$$

We can then minimise the amount of inconsistency to produce our best estimate of the true **X**: if the elements of **L** all have the same error then a measure of how well **F**(**X**) matches **L** is simply:

$$\Phi_{unweighted} = \left[\mathbf{F}(\mathbf{X}) - \mathbf{L}\right]^{\mathbf{T}} \left[\mathbf{F}(\mathbf{X}) - \mathbf{L}\right]\,. \tag{B.3}$$

This is called the objective function, and simply gives the sum of squares of the differences between the model's predictions for the data points and the points themselves. If the errors on the measurements are not identical then the objective function above needs to be altered to give different weighting to the different data points. This is achieved by inserting the weighting matrix **P**. This then produces the objective function:

$$\Phi = [\mathbf{F}(\mathbf{X}) - \mathbf{L}]^{\mathbf{T}} \, \mathbf{P} \, [\mathbf{F}(\mathbf{X}) - \mathbf{L}] \, . \tag{B.4}$$

In our case, as the measurement errors are assumed to be uncorrelated, the weighting matrix is diagonal and is described thus:

$$\mathbf{P} = \begin{pmatrix} \frac{1}{{\sigma_1}^2} & 0 & 0 \\ 0 & \frac{1}{{\sigma_2}^2} & 0 \\ 0 & 0 & \ddots \end{pmatrix} . \tag{B.5}$$

where σ_i is the statistical error of the ith measurement in the measurement vector. We then simply minimise Φ to give the $\hat{\mathbf{X}}$ that fits the data best. Any minimum will have the condition:

$$\left. \frac{\partial \Phi}{\partial X_i} \right|_{X_i = X_{i\,min}} = 0 \, , \tag{B.6}$$

for all i.

B.2 Linear Case

Consider the case where $\mathbf{F}(\mathbf{X})$ is linear. Then $\mathbf{F}(\mathbf{X})$ can be represented by $\mathbf{AX}$ where $\mathbf{A}$ is a matrix of the appropriate dimensions. This matrix is called the *design matrix* as it describes the system:

$$\mathbf{L} = \mathbf{AX} \, . \tag{B.7}$$

Now the objective function given by equation B.4 becomes:

$$\Phi = [\mathbf{AX} - \mathbf{L}]^{\mathbf{T}} \, \mathbf{P} \, [\mathbf{AX} - \mathbf{L}] \, . \tag{B.8}$$

B.2.1 Minimum

To find the minimum the condition given in equation B.6 is solved. Combining equations B.6 and B.8 with chain rule differentiation gives:

$$\frac{\partial \boldsymbol{\Phi}}{\partial \hat{\mathbf{X}}} = \frac{\partial \boldsymbol{\Phi}}{\partial (\mathbf{A}\hat{\mathbf{X}} - \mathbf{L})} \frac{\partial (\mathbf{A}\hat{\mathbf{X}} - \mathbf{L})}{\partial \hat{\mathbf{X}}} = 0 \tag{B.9}$$

with the substitution

$$\mathbf{R} = \mathbf{A}\hat{\mathbf{X}} - \mathbf{L} \tag{B.10}$$

producing

$$\frac{\partial \mathbf{\Phi}}{\partial \hat{\mathbf{X}}} = \frac{\partial \mathbf{\Phi}}{\partial \mathbf{R}} \frac{\partial \mathbf{R}}{\partial \hat{\mathbf{X}}} \tag{B.11}$$

The matrix product differentiation is performed using the identity (row-vector formalism):

$$\frac{\partial \left(\mathbf{R}^{\mathbf{T}}\mathbf{P}\mathbf{R}\right)}{\partial \mathbf{R}} = \mathbf{R}(\mathbf{P} + \mathbf{P}^{\mathbf{T}}) = 2\mathbf{R}^{\mathbf{T}}\mathbf{P}\,. \tag{B.12}$$

where the simplification is due to $\mathbf{P}^{\mathbf{T}} = \mathbf{P}$. With row-vector $\frac{\partial \mathbf{R}}{\partial \hat{\mathbf{X}}}$ is $\mathbf{A}$, giving:

$$2(\mathbf{A}\hat{\mathbf{X}} - \mathbf{L})^{\mathbf{T}}\mathbf{P}\mathbf{A} = 0\,. \tag{B.13}$$

The equation is divided by 2 and transposed:

$$\mathbf{A}^{\mathbf{T}}\mathbf{P}^{\mathbf{T}}(\mathbf{A}\hat{\mathbf{X}} - \mathbf{L}) = 0\,. \tag{B.14}$$

The results is separated and the identity $\mathbf{P}^{\mathbf{T}} = \mathbf{P}$ is applied:

$$\mathbf{A}^{\mathbf{T}}\mathbf{P}\mathbf{A}\hat{\mathbf{X}} = \mathbf{A}^{\mathbf{T}}\mathbf{P}\mathbf{L}\,. \tag{B.15}$$

Solving for $\mathbf{X}$ by left-applying $\left(\mathbf{A}^{\mathbf{T}}\mathbf{P}\mathbf{A}\right)^{-1}$ giving:

$$\hat{\mathbf{X}} = \left(\mathbf{A}^{\mathbf{T}}\mathbf{P}\mathbf{A}\right)^{\mathbf{-1}}\mathbf{A}^{\mathbf{T}}\mathbf{P}\mathbf{L}\,. \tag{B.16}$$

which is the unbiased estimator for $\mathbf{X}$ at the stationary point, which due to the configuration of the the problem, is the minimum.

B.2.2 Covariance

The covariance matrix at the minimum can also be found easily. Consider:

$$\mathbf{L} = \mathbf{A}\mathbf{X}\,, \tag{B.17}$$

where $\mathbf{X}$ is the vector of parameters, $\mathbf{A}$ is a matrix which contains the model that is linear in $\mathbf{X}$ (or is a linear approximation). Now, consider a simple expression:

$$\mathbf{AB} = \mathbf{I}\,. \tag{B.18}$$

Therefore

$$\mathbf{BL} = \hat{\mathbf{X}}\,, \tag{B.19}$$

where $\hat{\mathbf{X}}$ is the unbiased estimator of $\mathbf{X}$. Applying the covariance law gives:

$$\mathbf{\Sigma}_{\hat{\mathbf{X}}} = \mathbf{B}\mathbf{\Sigma}_{\mathbf{L}}\mathbf{B}^{\mathbf{T}}\,. \tag{B.20}$$

where $\mathbf{\Sigma}_{\mathbf{L}}$ is the covariance matrix of the measurements. Applying equation B.18 gives:

$$\mathbf{\Sigma}_{\hat{\mathbf{X}}} = \mathbf{A}^{-\mathbf{1}}\mathbf{\Sigma}_{\mathbf{L}}\mathbf{A}^{-\mathbf{T}} \tag{B.21}$$

which is better expressed as:

$$\mathbf{\Sigma}_{\hat{\mathbf{X}}} = \left(\mathbf{A}^{\mathbf{T}}\mathbf{\Sigma}_{\mathbf{L}}^{-\mathbf{1}}\mathbf{A}\right)^{-\mathbf{1}}\,. \tag{B.22}$$

If $\mathbf{\Sigma}_{\mathbf{L}}^{-\mathbf{1}}$ is the weighting matrix, $\mathbf{P}$, so:

$$\mathbf{\Sigma}_{\hat{\mathbf{X}}} = \left(\mathbf{A}^{\mathbf{T}}\mathbf{P}\mathbf{A}\right)^{-\mathbf{1}}\,. \tag{B.23}$$

If the weighting matrix is only known down to a scale factor then it can be parametrised as follows:

$$\mathbf{\Sigma}_{\hat{\mathbf{X}}} = \hat{\sigma}_0^2\left(\mathbf{A}^{\mathbf{T}}\hat{\mathbf{P}}\mathbf{A}\right)^{-\mathbf{1}}\,. \tag{B.24}$$

where $\hat{\sigma}_0^2$ is the chi-squared per degree of freedom normalisation factor given by:

$$\hat{\sigma}_0^2 = \frac{\hat{\mathbf{V}}^{\mathbf{T}}\hat{\mathbf{P}}\hat{\mathbf{V}}}{n-u}\,, \tag{B.25}$$

where $\mathbf{V}$ is the misclosure vector, n is the number of observations and u is the number of model parameters, meaning $n - u$ is the number of degrees of freedom. If the input error matrix has the correct absolute values of the error estimates then the value of $\hat{\sigma_0^2}$ is near unity.

The design matrix (**A**) may be complicated so the covariance matrix might not be a simple diagonal matrix, but have significant off-diagonal elements, corresponding to parameter error correlations. The diagonal elements are the variances of the model parameters.

B.3 Symmetries, Constraints and Point Projection

B.3.1 Symmetries

If a system has degeneracies then $\mathbf{A^TPA}$ will be singular. In a physical model, these degeneracies represent transformations that leave the value of the function unchanged; a global transformation, for example. This has the exact analogue in field theories as gauge transformations. The symmetry can be removed by adding constraints to the system, making the combined matrix invertible, and allowing the method to proceed.

B.3.2 Outer Constraints

Outer constraints are a fixed set of restrictions to the system identified in advance. It is implemented by introducing a constraint matrix, $\mathbf{A_2}$, describing the restriction into the model [63, p142-146]. This matrix needs to be introduced in such a way to to reflect the equation:

$$\mathbf{G}(\mathbf{X}) \equiv \mathbf{F}(\mathbf{X}) - \mathbf{L} = 0\,, \tag{B.26}$$

which is simply a re-expression of equation B.1. The constraint matrix is defined by:

$$\mathbf{A_2} = \alpha \frac{\partial \mathbf{G}}{\partial \mathbf{X}}\,, \tag{B.27}$$

where α is a scalar constant, used to make it clear that there is no dependence of the result upon scaling the equation. We define the matrix $\mathbf{C}$ by:

$$\mathbf{C} = \begin{pmatrix} A^TPA & A_2^T \\ A_2 & 0 \end{pmatrix}. \tag{B.28}$$

with the expected value of any constraint appended onto the $\mathbf{X}$ vector. This new combined matrix is still invertible. Now the parameter vector can be found using the relation analogous to equation B.16:

$$\mathbf{X} = \mathbf{C^{-1}} \begin{pmatrix} A^TPL \\ 0 \end{pmatrix}. \tag{B.29}$$

The covariance-squared matrix is simply $\mathbf{C^{-1}}$, though the sub-matrix corresponding to the model parameters can be examined in isolation.

B.3.3 Inner Constraints

Inner constraints are where the symmetries are identified in each step of the minimisation and dynamic constraints are added to remove them. In effect, this means that the zero-valued eigenvectors are calculated repeatedly, then used as constraints in the same way as above. The main differences with outer constraints are that the eigenvectors, and hence constraints, can change during the minimisation and are not necessarily easy to interpret into a physical symmetry. Due to the nature of explicitly using calculated eigenvectors, inner constraints will always give the lowest trace of the error matrix at the end of the minimisation, though this does not mean the minimum is better, just expressed differently. Inner constraints can be used in combination with outer constraints and can be very useful when the properties of the target of the minimisation are not fully understood.

B.3.4 Solving an over-constrained problem

Sometimes a set of outer constraints will contain fewer actual restrictions than the number of equations. This will introduce a degeneracy in $\mathbf{A_2}$ matrix and lead to a non-invertible $\mathbf{C}$ matrix. To reduce the $\mathbf{A_2}$ matrix to the minimal number we can perform a singular value decomposition as follows: $\mathbf{A_2}$ is an $m \times n$ matrix with $m > n$ which can be written as the product of three matrices as follows:

$$\mathbf{A_2} = \mathbf{UWV^T} \tag{B.30}$$

where $\mathbf{U}$ is a dense $m \times m$ matrix, $\mathbf{W}$ is a $m \times n$ matrix with the singular values along the diagonal and zero elsewhere, and $\mathbf{V}$ is a dense $n \times n$ matrix. Note that for both $\mathbf{U}$ and $\mathbf{V}$, the transpose is also the inverse.

If the singular values along the diagonal along the $\mathbf{W}$ matrix are arranged in order of magnitude, it is easy to discard the rows of the matrix with zero singular values, and similarly from the matrix $\mathbf{U^T A_2}$. This reduced matrix, along with the iteration correction with the same process applied, can then be used instead of $\mathbf{A_2}$ to produce the change in $\mathbf{X}$.

B.3.5 Point Projection

The method used to introduce outer constraints can be amended to produce estimates of the function output for given input as part of the minimisation.

A variable for the desired output value for the function being projected at the point of interest is added to the **X** vector. This has the effect of adding a column of zeros to the **A** matrix, as clearly the output of the modelling function itself cannot depend on this parameter. Then, a constraint is added which contains a representation of the linear function being minimised, with a value of -1 for the element corresponding to the result parameter. As the whole line is required to be zero, then this forces the result into that parameter. If the model produces more than one output value then additional parameters will be required with additional constraints, though the method is the same.

The error matrix calculation is unaffected so the error estimate on the function for a given input value can be simply read out from the appropriate matrix element.

B.4 Polynomial Fit Example

A simple example is a polynomial fit for a dataset of n pairs of t values with dependent y values. We want to know the coefficients of an order u polynomial (where $u < n$) with coefficients x_0 to x_u that best describes the relationship. The **A** matrix contains columns of the t values raised to the power of the polynomial term associated with the coefficient:

$$\begin{array}{cccccc} \mathbf{V} & + & \mathbf{L} & = & \mathbf{A} & \mathbf{X}, \\ \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \end{pmatrix} & + & \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix} & = & \begin{pmatrix} 1 & t_1 & t_1^2 & \dots & t_1^u \\ 1 & t_2 & t_2^2 & \dots & t_2^u \\ 1 & t_3 & t_3^2 & \dots & t_3^u \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_n & t_n^2 & \dots & t_n^u \end{pmatrix} & \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_u \end{pmatrix}. \end{array}$$

Note that although the polynomial can be of order greater than one, the minimisation itself is linear in the abscissa. Equation B.16 can then be used to find the coefficients, given a suitable weighting matrix, with equation B.24 producing the covariance matrix. Extending to make a projection introduces the constraint matrix:

$$\mathbf{A_2} = \begin{pmatrix} 1 & \tau & \tau^2 & \dots & \tau^u & -1 \end{pmatrix}, \tag{B.31}$$

where τ is an explicit time where a y estimate is desired. The **A** matrix gains a null column and the $\mathbf{X}$ vector gains the parameter that will give the estimate $y(\tau)$. The problem being solved is easier to visualise if the elements are appended, though this expression itself is not solved as the values appended to the $\mathbf{L}$ vector are exact and are not measurements, breaking any possible weighting matrix:

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} 1 & t_1 & t_1^2 & \dots & t_1^u & 0 \\ 1 & t_2 & t_2^2 & \dots & t_2^u & 0 \\ 1 & t_3 & t_3^2 & \dots & t_3^u & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & 0 \\ 1 & t_n & t_n^2 & \dots & t_n^u & 0 \\ \mathbf{1} & \tau_{\mathbf{n}} & \tau_{\mathbf{n}}^{\mathbf{2}} & \dots & \tau_{\mathbf{n}}^{\mathbf{u}} & \mathbf{-1} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_u \\ \hat{\mathbf{y}}_\tau \end{pmatrix} . \tag{B.32}$$

Instead, equations B.28 and B.29 are then used to find the polynomial parameters and the estimate of $y(\tau)$, with the error-squared values read off the corresponding elements of $\mathbf{C^{-1}}$.

B.5 Non-Linear Case

B.5.1 Linearisation

In the case that the problem is non-linear, then the above method will not work as $\mathbf{F}(\mathbf{X})$ cannot be represented by $\mathbf{AX}$. Instead, an initial estimate of $\hat{\mathbf{X}}$, $\mathbf{X_0}$ is made, which is then iteratively updated until it cannot be improved any more to a given tolerance. There are many methods to achieve this. One such method, described in [63, p101], is to use the first order Taylor series of the function to linearise the function. The linearised function is then used as in section B.2 to give the starting point for the next iteration. This is a steepest descent minimisation. (Note that the second order term in the Taylor series can be used to give a more rapid convergence, at the cost of requiring second derivatives. This, Newton's Method, is covered in section 6.6.3.) Consider the initial estimate and the adjustment, giving the total solution vector:

$$\bar{\mathbf{X}} = \mathbf{X_0} + \boldsymbol{\Delta}\mathbf{X} \tag{B.33}$$

The first-order method considers the misclosure equation, B.2, with the expansion applied, giving:

$$\mathbf{F}(\mathbf{X_0}) + \left.\frac{\partial \mathbf{F}}{\partial \boldsymbol{\Delta}\mathbf{X}}\right|_{\mathbf{X_0}} \boldsymbol{\Delta}\mathbf{X} - (\mathbf{L} + \mathbf{V}) = \mathbf{0}\,, \tag{B.34}$$

which can be expressed as:

$$\mathbf{W} + \mathbf{A}\boldsymbol{\Delta}\mathbf{X} - \mathbf{V} = 0\,, \tag{B.35}$$

with $\mathbf{W} = \mathbf{F}(\mathbf{X_0}) - \mathbf{L}$ and $\mathbf{A} = \frac{\partial \mathbf{F}}{\partial \mathbf{\Delta X}}\big|_{\mathbf{X_0}}$. This can then be used to derive the normal equations in a similar way to the linear case. (See [64, p108-116] for a general case using Lagrange's method and a more familiar back-substitution method.) This leads to the expression of the value of the adjustment vector:

$$\mathbf{\Delta X} = -\left(\mathbf{A^T P A}\right)^{-1} \mathbf{A^T P}\left(\mathbf{F}(\mathbf{X_0}) - \mathbf{L}\right) \tag{B.36}$$

with the covariance matrix given by:

$$\mathbf{\Sigma_X} = \sigma_{\mathbf{0}}^{\mathbf{2}} \left(\mathbf{A^T P A}\right)^{-1} \tag{B.37}$$

as before [64, p116-120]. However, as $\mathbf{F}(\mathbf{X})$ is not linear, this is not precisely the solution, only an improved estimate. Hence the process is repeated until the desired tolerance, measured by the change in Φ or $\mathbf{X}$, is met.

B.5.2 Line Search

One difficulty with non-linear problems is that the update to the iteration of $\mathbf{X}$ produced by equation B.36 may overshoot, leading to slow convergence. To mitigate this, a line search can be performed along the direction of $\Delta\mathbf{X}$ to give a scalar multiplier, α, to adjust the initial update to a better one. In principle, we want to minimise the objective function, B.4, constrained along the direction of $\mathbf{\Delta X}$:

$$\mathbf{\Phi}\left(\mathbf{X_0} + \alpha\mathbf{\Delta X}\right) = \left(\mathbf{F}(\mathbf{X_0} + \alpha\mathbf{\Delta X}) - \mathbf{L}\right)^{\mathbf{T}} \mathbf{P}\left(\mathbf{F}(\mathbf{X_0} + \alpha\mathbf{\Delta X}) - \mathbf{L}\right) \tag{B.38}$$

In practice, minimising this for α may be very time consuming, even if we limit our search to looking for a local minimum. Instead, we find a solution that is good enough using an inexact line search; the requirements are not very strict seeing as the solution will only be used as a temporary iteration or the distance will be so small as to not make a noticeable difference.

'Good enough' is a flexible term. A simple requirement that a new point must be better than the current one may not be sufficient to give rapid minimisation. When there is an estimate of the upper limit to work with, which is the case in many minimisations, then a commonly used pair of requirements are the Wolfe Conditions [65, Section 3.1]:

$$\begin{aligned} f\left(\mathbf{X_0}+\alpha\mathbf{\Delta X}\right) &\leq f\left(\mathbf{X_0}\right)+c_1\alpha \left.\frac{\partial f\left(\mathbf{X}\right)}{\partial \mathbf{X}}\right|_{\mathbf{X}=\mathbf{X_0}}^{\mathbf{T}} \mathbf{\Delta X}\,, &\text{(B.39)}\\ \frac{\partial f\left(\mathbf{X_0}+\alpha\mathbf{\Delta X}\right)}{\partial \mathbf{X}}^{\mathbf{T}} \mathbf{\Delta X} &\geq c_2 \left.\frac{\partial f\left(\mathbf{X}\right)}{\partial \mathbf{X}}\right|_{\mathbf{X}=\mathbf{X_0}}^{\mathbf{T}} \mathbf{\Delta X}\,, &\text{(B.40)}\end{aligned}$$

where $0 < c_1 < c_2 < 1$. These two additional scalar constants are generally related to the type of search direction algorithm, with c_1 usually having a value of 1×10^{-4} and c_2 having a value of 0.9 for Newton's or quasi-Newtonian methods. The first equation ensures that the decrease is sufficient while the second ensures that the gradient is more positive than the one at the current position. If the gradient is more negative, then increasing α should give quick gains.

A simple method is a backtracking line search, which is to start at $\alpha = 1$ and geometrically reduce it until the sufficient decrease condition is true. This has the advantage that the problem of very small steps is not an issue, meaning the second Wolfe condition is not needed.

B.6 Potential Issues

B.6.1 Local Minima

With a non-linear function there may be multiple local minima. Which minimum is found by the procedure depends on the starting conditions, so it is often important to have good estimates of the initial values. Techniques to find global minima, such as simulated annealing, are available though were not found to be necessary for the minimisations performed in this investigation.

B.6.2 Calculating Derivatives

Analytically calculating derivatives can be extremely time-consuming and memory intensive, particularly if second derivatives are required as is the case for many more advanced minimisations. The steepest descent method described above only requires first derivatives, and was suitable for the calibration problem it was used for, but even these may not be practical to calculate for larger problems. To avoid this issue, many more sophisticated minimisation techniques use an estimate for the Hessian matrix, which is typically updated each iteration (§B.7).

B.6.3 Inverting Matrices

If the function being minimised has many input variables and output measurements, then the **C** matrix may become very large. However, the matrix does not need to be inverted each iteration;

instead equation B.36 can be solved just as a series of linear equations, for which there are numerous other approaches. For example, Mathematica has a dedicated function to solve the problem $m.x = b$ based on Gaussian Elimination which is far faster for large systems. The error estimates do require the inversion of the **C** matrix but only need to be calculated at the end of the minimisation.

B.7 The Minuit Minimiser

The primary algorithm used by the Minuit fitting package is the Migrad algorithm. It is based on the method described by R. Fletcher and M.J.D. Powell [66] which itself used the work of W.C. Davidon, producing the Davidon-Fletcher-Powell (DFP) algorithm. These approaches and derivative works are all quasi-Newtonian methods. Newton's method is to assume that the function can be approximated by a quadratic in each dimension. Then, given a suitable starting point, the change in input vector to get to the stationary point can be found by considering the Taylor expansion. As long as the function is reasonably approximated by the second order Taylor expansion, the new point will be closer to the stationary point and so repeated iterations will rapidly converge to the correct answer. The main advantage over the steepest descent methods is that typically fewer steps are needed, particularly if there is significant variation in the chi-squared dependence on model parameters.

The process is easily understood as follows: consider the one-dimensional case with a function $f(x)$ at point x_n. Its Taylor expansion is:

$$f_T(x_n + \Delta x) = f(x_n) + f^{'}(x_n)\Delta x + \frac{1}{2} f^{''}(x_n)\Delta x^2 \,. \tag{B.41}$$

Now, at a stationary point $\frac{df_T}{d\Delta x} = 0$, so clearly the following will hold:

$$f^{'}(x_n) + f^{''}(x_n)\Delta x = 0 \,. \tag{B.42}$$

So the step needed to get to the stationary point of the quadratic form is:

$$\Delta x = -\frac{f^{'}(x)}{f^{''}(x)} \,. \tag{B.43}$$

If the function being analysed is not a quadratic, the process needs to be repeated until the desired precision is achieved.

Extending to multiple dimensions is easy - the differential simply becomes the gradient operator and equation B.43 becomes:

$$\Delta\mathbf{X} = -\left(\mathbf{G}(\mathbf{X})\right)^{-1} \nabla F(\mathbf{X}) \,, \tag{B.44}$$

where $\mathbf{G}$ is the Hessian matrix such that $G_{ij} = \partial^2 F/\partial x_i \partial x_j$. A line search is then performed along $\mathbf{\Delta X}$ to get to the next point in the iteration.

In quasi-Newtonian methods, the Hessian matrix is not used, as it may be difficult to calculate (and expensive to invert). Instead, the very similar procedure below is performed:

1. Start with a point $\mathbf{X_n}$ and function $F(\mathbf{X})$.
2. Find the value, or estimate of, gradient $\nabla_{\mathbf{x}} F(\mathbf{X_n}) \equiv \mathbf{g}(\mathbf{X_n})$ and an estimate of the Hessian matrix.
3. Aim to find a point $\mathbf{X_{n+1}}$, where $F(\mathbf{X_{n+1}}) < F(\mathbf{X_n})$ using equation B.44, possibly with a scalar multiplier.
4. Start again from this new point until $F(\mathbf{X_n}) - F(\mathbf{X_{n+1}}) < t$, where t is the fit tolerance.

As explained above, and in [66], $\mathbf{G^{-1}}$ can be used to calculate the displacement from the current point to the minimum: $\mathbf{X_0} = -\mathbf{G^{-1}}(\mathbf{X})\mathbf{g}(\mathbf{X}) + \mathbf{X}$. The approach the DFP algorithm uses is not to try and calculate $\mathbf{G}^{-1}$, but to have a running estimate $\mathbf{H}$, initially the identity matrix, which is improved each iteration along with the estimate of $\mathbf{X_0}$. i.e. the direction of the next $\mathbf{X}$ from the current one is $-\mathbf{Hg}(\mathbf{X})$. Then the new estimate for $\mathbf{X}$ is produced by altering the scalar multiplier to minimise $F(\mathbf{X} - \alpha\mathbf{Hg}(\mathbf{X}))$. Minuit uses quadratic interpolation or cubic interpolation to perform this line search [67]. The method used by Minuit to update its estimate of the inverse Hessian is the final update to the DFP algorithm and is described in [68].

C Design Drawings

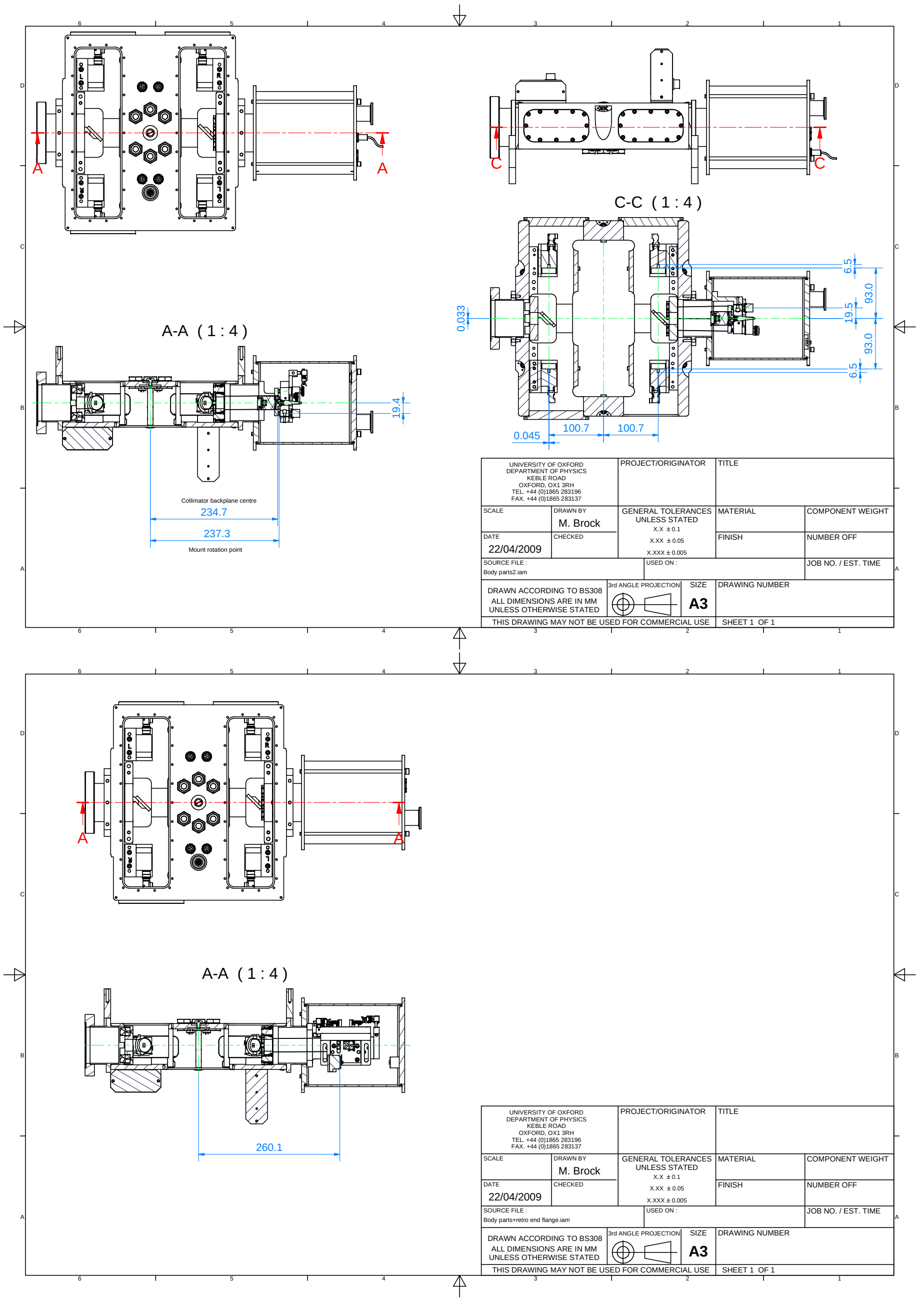


Figure C.1: Measurement units with the Launch End-Cap attached (top) and with the Retro End-Cap attached (bottom). Some relevant specification measurements are highlighted.

D Trigonometry

D.1 Background

The transformations from frame to frame need to be carefully considered. Defining the frame column vectors as follows:

$$\mathbf{e_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} ; \quad \mathbf{e_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} ; \quad \mathbf{e_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{D.1}$$

These will then transform to the frame vectors in the rotated (primed) frame by pre-applying the rotation matrix **R** to them. Due to the definitions of the initial frame vectors **R** can then be expressed simply as a matrix whose 3 columns are the primed frame vectors.

$$\mathbf{R} = \left(\mathbf{e}_1^{'}, \mathbf{e}_2^{'}, \mathbf{e}_3^{'}\right) \tag{D.2}$$

e.g:

$$\mathbf{e}_1^{'} = R^i{}_1 \mathbf{e}_i \tag{D.3}$$

Therefore the same point in space can be expressed in the two different frames:

$$X^i \mathbf{e}_i = X^{'j} \mathbf{e}_j^{'} \tag{D.4}$$

by using equation D.3 we can simply produce:

$$X^{'k} R^i{}_k \mathbf{e}_i = X^{'j} \mathbf{e}^{'}{}_j \tag{D.5}$$

which trivially leads to:

$$X^{'k} R^i{}_k \mathbf{e}_i = X^{'j} \mathbf{e}_j^{'} = X^i \mathbf{e}_i \tag{D.6}$$

reducing to:

$$X^{'j} R^i{}_j = X^i \tag{D.7}$$

This simply means that if an object in the primed co-ordinate system has co-ordinates $\mathbf{X}^{'}$ then it has co-ordinates $\mathbf{R}\mathbf{X}^{'}$ in the unprimed system. This can also be expressed as the matrix equation:

$$\mathbf{R}\mathbf{X}^{'} = \mathbf{X} \tag{D.8}$$

$$\mathbf{X}^{'} = \mathbf{R}^{\mathbf{-1}}\mathbf{X} = \mathbf{R}^{\mathbf{T}}\mathbf{X} \tag{D.9}$$

In the 3-frame system

$$X^{i}\mathbf{e}_{i} \equiv X^{'j}\mathbf{e}_{j}^{'} \equiv X^{''k}\mathbf{e}_{k}^{''} \tag{D.10}$$

with the relations:

$$\mathbf{e}_{j}^{'} = R^{i}{}_{j}\mathbf{e}_{\mathbf{i}},\ \mathbf{e}_{k}^{''} = R^{'j}{}_{k}\mathbf{e}_{\mathbf{j}}^{'} \tag{D.11}$$

and equation D.4 expands to:

$$X^{''l}\mathbf{e}_{l}^{''} = X^{''k}R^{'j}{}_{k}\mathbf{e}_{j}^{'} = X^{''k}R^{'j}{}_{k}R^{i}{}_{j}\mathbf{e}_{\mathbf{i}} \tag{D.12}$$

which again reduces to the equivalent to D.7:

$$X^{''k}R^{'j}{}_{k}R^{i}{}_{j} = X^{i} \tag{D.13}$$

giving the equations:

$$\mathbf{R}\mathbf{R}^{'}\mathbf{X}^{''} = \mathbf{X} \tag{D.14}$$

$$\mathbf{X}^{''} = \left(\mathbf{R}^{'}\right)^{\mathbf{-1}}\mathbf{R}^{\mathbf{-1}}\mathbf{X} \tag{D.15}$$

D.1.1 Active and Passive Rotations

Another convention that needs to be decided is whether rotations are taken as ‘active’ or ‘passive’. Active rotations are where the rotation is applied to the object, while passive rotations are when the rotation is applied to the frame the object is in: one is the inverse of the other. The ray-tracers written both used active rotations with the following matrices:

$$\mathbf{R}_x(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \tag{D.16}$$

$$\mathbf{R}_y(\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \tag{D.17}$$

$$\mathbf{R}_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{D.18}$$

D.1.2 Translations

Entities in the ray tracer need to have an associated position in space as well. To this end, each frame (apart from the top-level or 'master' frame) has a position in its parent frame, expressed in the parent's co-ordinate system. In addition to this 'position in parent frame' there is a 'local position' which applies to the object (e.g. a CCD or a child frame) contained in this frame. While the ray-tracer internally makes the translations and applies the rotations in separate steps, the translations being addition of vectors and the rotations being multiplication of matrices the different operations can be combined more easily by using homogeneous co-ordinates and affine transformations.

If a unit element is added to the end of each position vector then translations can be realised via matrix multiplication. The transformation given in equation D.19 produces a rotation by **R** followed by a translation along **t** in the primed co-ordinate system.

$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} & & & t_x \\ & \mathbf{R} & & t_y \\ & & & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \tag{D.19}$$

D.2 Explicit formalisms

The rotation matrix has three degrees of freedom. There are many methods used to parametrise the rotation matrix with difference conventions appropriate for different situations. It is logical to decompose the matrix into three separate rotations, each with a single free parameter.

D.2.1 ZXZ Euler Formalism

A common method of decomposition is into rotate around the Z-axis, then X-axis (or Y-axis), then Z-axis again [56, p150-154]. Consider an axis system XYZ: To transform it into a new orientation using a set of Euler angles the following operations are performed in order:

- Rotate XYZ by α around the Z-axis.
- Rotate XYZ by β around the new X'-axis.
- Rotate XYZ by γ around the new Z”-axis.

which is the intrinsic form. Note that the effect is identical to the following:

- Rotate XYZ by γ around the Z-axis.
- Rotate XYZ by β around the original X-axis.
- Rotate XYZ by α around the original Z-axis.

which is the extrinsic form.

$$\begin{aligned} \mathbf{R} &= \mathbf{R_{z''}}(\gamma)\mathbf{R_{x'}}(\beta)\mathbf{R_z}(\alpha) && \text{(D.20)} \\ &\equiv \mathbf{R_z}(\alpha)\mathbf{R_x}(\beta)\mathbf{R_z}(\gamma)\,, && \text{(D.21)} \end{aligned}$$

D.2.2 Tait-Bryan Formalism

The rotation matrix for Tait-Bryan XYZ (pitch-yaw-roll) formalism is constructed as follows:

$$\begin{aligned} \mathbf{R} &= \mathbf{R_{z''}}(\theta_R)\mathbf{R_{y'}}(\theta_Y)\mathbf{R_x}(\theta_P) && \text{(D.22)} \\ &\equiv \mathbf{R_x}(\theta_P)\mathbf{R_y}(\theta_Y)\mathbf{R_z}(\theta_R)\,, && \text{(D.23)} \end{aligned}$$

where the extrinsic matrix notation of D.25 is far more convenient as the matrices themselves have a constant form. This formalism was used in almost all cases of this work. The rotation matrix for Tait-Bryan YXZ (yaw-pitch-roll) formalism is constructed as follows:

$$\begin{aligned} \mathbf{R} &= \mathbf{R_{z''}}(\theta_R)\mathbf{R_{x'}}(\theta_P)\mathbf{R_y}(\theta_Y) && \text{(D.24)} \\ &\equiv \mathbf{R_y}(\theta_Y)\mathbf{R_x}(\theta_P)\mathbf{R_z}(\theta_R)\,, && \text{(D.25)} \end{aligned}$$

Explicitly expanding out the rotation matrices gives the following expression:

$$\mathbf{R} = \begin{pmatrix} \cos\theta_R\cos\theta_Y + \sin\theta_R\sin\theta_Y\sin\theta_P & \cos\theta_R\sin\theta_Y\sin\theta_P - \sin\theta_R\cos\theta_Y & \cos\theta_P\sin\theta_Y \\ \sin\theta_R\cos\theta_P & \cos\theta_R\cos\theta_P & -\sin\theta_P \\ -\cos\theta_R\sin\theta_Y + \sin\theta_R\cos\theta_Y\sin\theta_P & \cos\theta_R\cos\theta_Y\sin\theta_P + \sin\theta_R\sin\theta_Y & \cos\theta_P\cos\theta_Y \end{pmatrix}, \tag{D.26}$$

which can be useful when the total rotation matrix is known and the values of the characteristic parameters are required.

D.3 Spherical-Polar coordinates

In some circumstances it is natural to express positions in terms of spherical polar coordinates; for example measurements by a laser tracker are literally the two angles and radial distance with the tracker at the origin. There are numerous different conventions in use to describe points.

D.3.1 Leica Laser Tracker conventions

The laser tracker uses a right-handed axis system with the z axis being vertical. The radial distance, reported as “Z” is measured in millimetres. The inclination, confusingly both reported as “elevation” and “Phi” is measured in degrees from the z-axis, with a range of 0 to 180. The azimuthal angle, reported as “Theta” is measured in degrees *clockwise* from the X-axis and was not ranged, though in practice the value reported was taken modulo 360. Table D.1 summarises the conventions encountered.

Table D.1: Spherical Polar Coordinate Conventions

	‘Standard’ physics	Leica LT	LiCAS
Units	radians, m	degrees, mm	radians, mm
Vertical axis	Z	Z	Y
Inclination measured down from	Z	Z	Y
Azimuthal angle measured from	X	X	Y
Azimuthal angle measured	anti-clockwise	clockwise	anti-clockwise