Detecting Antihydrogen Annihilations in ALPHA-g for Measurement of Gravitational Free Fall

by

Gareth Smith

M.Sc., The University of British Columbia, 2021
B.Sc., Carleton University, 2019

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Physics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

September 2025

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Detecting Antihydrogen Annihilations in ALPHA-g for Measurement of Gravitational Free Fall

submitted by **Gareth Smith** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy** in **Physics**.

Examining Committee:

Dr. Makoto Fujiwara, Senior Research Scientist, Particle Physics Department, TRIUMF
Co-Supervisor

Dr. Takamasa Momose, Professor, Department of Chemistry, University of British Columbia
Co-Supervisor

Dr. Michael Hasinoff, Professor, Department of Physics and Astronomy, University of British Columbia
Supervisory Committee Member

Dr. Christopher Hearty, Professor, Department of Physics and Astronomy, University of British Columbia
University Examiner

Dr. Andrew MacFarlane, Professor, Department of Chemistry, University of British Columbia
University Examiner

Dr. Michel Lefebvre, Professor, Department of Physics and Astronomy, University of Victoria
External Examiner

Additional Supervisory Committee Members:

Dr. Mark Halpern, Professor, Department of Physics and Astronomy, University of British Columbia
Supervisory Committee Member

Abstract

Antihydrogen, the bound state of an antiproton and a positron, is an ideal system for testing fundamental symmetries between matter and antimatter. The Antihydrogen Laser PHysics Apparatus (ALPHA) at CERN has a proven history of producing and trapping antihydrogen atoms, with many precision tests of charge-parity-time (CPT) symmetry. The new ALPHA-g apparatus extends this program to a test of gravity, directly probing the applicability of the weak equivalence principle (WEP) to an antimatter system for the first time.

This dissertation presents the first ALPHA-g results, which confirm that Earth's gravity acts downwards on antimatter. The ratio between the gravitational acceleration of antihydrogen and hydrogen is estimated to be $\overline{g}/g = 0.75 \pm 0.13$ (statistical and systematic) $\pm\ 0.16$ (simulation).

This measurement depends on accurately determining the annihilation positions of antihydrogen atoms released from the ALPHA-g magnetic trap, which is accomplished using a time projection chamber (TPC) detector. To reduce the substantial background due to cosmic rays, a time-of-flight barrel scintillator (BSC) detector was also implemented. The author's primary contributions include the commissioning of the BSC, including devising and implementing time-of-flight measurements, leading to a significantly reduced background rate. Additional contributions include improving the TPC tracking algorithms to reach the required annihilation position resolution.

This dissertation begins with theoretical motivations for testing the WEP using antihydrogen, followed by an overview of the ALPHA apparatus and antihydrogen production. The TPC and BSC detectors are introduced, including the event reconstruction methods, with care to highlight the author's work. Finally, the gravity measurement methodology and results, as well as future perspectives, are discussed.

This first direct test of gravity on antimatter rules out a large difference in the gravitational behaviour of antimatter and matter, furthering our understanding of the fundamental matter-antimatter symmetry, and laying the foundation for future precision WEP tests using antimatter.

Lay summary

The known universe is almost entirely made of matter – mainly protons, neutrons, and electrons. Each of these particles has an antimatter version, a mirror-opposite counterpart. We study these rare particles in the laboratory.

The ALPHA collaboration researches the simplest antimatter atom, antihydrogen. We recently made headlines by confirming that antihydrogen falls downwards towards the Earth due to gravity, like normal matter.

To do this, ALPHA produces and traps a few hundred antihydrogen atoms, and then releases them so they fall and collide with the apparatus. Each antiatom is destroyed in an annihilation, which we can detect with a sophisticated set of particle detectors.

In my PhD, I commissioned one of these particle detectors, and developed algorithms to identify the annihilations. My work contributed directly to the measurement of gravity, and will enable more precise future measurements.

Preface

This dissertation is based on the ALPHA-g experiment and the work of the ALPHA collaboration, an international group of researchers. The apparatus, data collection and analysis is the result of a combined effort from the entire collaboration. My contributions were made within the context of the ALPHA collaboration, and would not have been possible without input from many collaborators. This text is entirely original.

The apparatus and techniques described in Chapter 3 reflect individual contributions from the entire ALPHA collaboration. The ALPHA-g detectors described in Chapter 4 were developed and commissioned by M. Fujiwara, P-A. Amaudruz, K. Olchanski, A. Capra, L. Martin, I. Carli, N. Massacret, D. Duque, P. Woosaree, R. Henderson, P. Lu, D. Gill, A. Olin, S. Menary, and myself. The TPC reconstruction algorithms described in Chapter 5 were primarily developed by A. Capra, K. Olchanski, L. Martin, J. McKenna, and D. Duque, with my own contributions identified throughout the chapter. The BSC reconstruction and calibration algorithms presented in Chapters 6–8 are entirely my own work, and the background rejection algorithms were developed by J. McKenna, L. Golino, and myself. The measurement procedure described in Chapter 9 is the work of the entire ALPHA collaboration. The conceptual development was led by J. Fajans, simulations led by C. So, magnet design and construction led by W. Bertsche and C. So, ECR development led by T. Friesen and A. Powell, magnetron-metry development led by A. Christensen, and escape curve analysis led by G. Bonomi, M. Uroni, S. Stracka, J. Fajans, J. Wurtele, and D. Hodgkinson. The discussion of future measurement schemes is my own. For the duration of my time on the ALPHA collaboration, the spokesperson was J. Hangst, co-spokesperson was N. Madsen, and technical coordinator was C. Ø. Rasmussen.

The gravity measurement presented in Chapter 9 is previously published in [1], of which I am a coauthor. The BSC calibration methods presented in Chapter 7 have been presented at the EXA/LEAP 2024 and PD24 conferences, with peer-reviewed conference proceedings in [2] and [3] respectively.

The use of *Generative AI* in this dissertation was limited to editorial and coding purposes. Selected paragraphs were evaluated by ChatGPT-4o to ensure grammatical correctness and good flow. Furthermore, ChatGPT-4o was used for some formatting tasks, such as inputting provided data into LaTeX tables, or creating BibTeX entries for provided citations. ChatGPT-3 was used to generate code which served as a starting point for the neural network training data pipeline discussed in Section 5.5, quickly introducing me to the syntax for the hdf5 data structure. Finally, ChatGPT-4o was used to generate code which served as a starting point for some figures included in this dissertation, specifically Figures 3.2, 3.4, 3.5, 3.11, 3.12, 4.3, 4.6, 4.7, 5.16, 8.2, 9.1, 9.8, 9.9, and 9.10. In each case, the generative AI output was heavily altered to create the code which produced the desired figure.

Table of contents

Appendices

List of tables

List of figures

Glossary

$\overline{\mathbf{g}}$ The gravitational acceleration of antihydrogen, analogous to $g = 9.81\,\mathrm{m/s^2}$ for matter.

AD The **A**ntiproton **D**ecelerator, CERN's machine which produces a beam of slow antiprotons which is further decelerated by ELENA.

ADC **A**nalog-to-**D**igital **C**onverter, a digitizer which samples an analog signal to store a full waveform.

ALPHA The **A**ntihydrogen **L**aser **PH**ysics **A**pparatus, the CERN-based collaboration of which the author is a member.

ALPHA-2 The second-generation ALPHA trap apparatus designed to perform precision spectroscopy on trapped antihydrogen atoms.

ALPHA-g The new ALPHA trap apparatus designed to measure the gravity of trapped antihydrogen atoms.

alphasoft The reconstruction software developed for the ALPHA detectors.

ASD **A**nalog **S**um **D**iscriminator board, a custom circuit board which sums together the analog signals from multiple SiPMs in the ALPHA-g BSC, and outputs both an analog and a digital discriminated signal.

AWB **A**node **W**ire **B**oard, a custom circuit board used to read out the anode wire current of the ALPHA-g TPC.

bar One of 64 scintillating bars in the BSC.

bar hit A data structure in the BSC reconstruction containing ADC and TDC data from SiPM signals on both ends of a bar.

BDT **B**oosted **D**escion **T**rees, a machine learning multivariate analysis method.

BSC The **B**arrel **SC**intillator, a scintillator-based time-of-flight detector in ALPHA-g.

CERN The European Organization for Nuclear Research, the leading global particle physics laboratory located in Geneva, Switzerland.

cosmic ray A particle originating from outside the Earth, or a secondary particle produced by its interaction in the atmosphere. At ground level, these are primarily muons.

CPT **C**harge-**P**arity-**T**ime, a fundamental symmetry governing the relationship between matter and antimatter.

cyclotron motion The fast circular motion of a charged particle in a constant magnetic field.

DAQ **D**ata **A**c**Q**uisition.

DCCT **D**irect-**C**urrent **C**urrent **T**ransformers, an electrical system used to measure large currents.

ECR **E**lectron **C**yclotron **R**esonance, a technique developed at ALPHA to measure magnetic fields using electron plasmas.

ELENA The **E**xtra **L**ow **EN**ergy **A**ntiproton ring, which further slows antiprotons from the AD and provides them to experiments.

end hit A data structure in the BSC reconstruction containing ADC and TDC data from SiPM signals on one end of a bar.

FPGA **F**ield **P**rogrammable **G**ate **A**rray, a commercially available reprogrammable integrated circuit.

Garfield++ A software toolkit designed for the simulation of gas-based particle detectors [5].

helix The spiral-shaped function fit to tracks in the TPC reconstruction.

Ioffe-Pritchard trap A neutral particle trap which constrains atoms in a magnetic field minima via their magnetic dipole moment. Consists of an octupole magnet and two solenoidal mirror coils.

magnetron motion The slow circular motion of a charged particle due to the $\vec{E} \times \vec{B}$ drift of non-parallel electric and magnetic fields.

magnetron-metry A technique developed at ALPHA to measure magnetic fields based on the magnetron motion of an electron plasma.

MCP **M**icro**C**hannel **P**late, a device used to multiple and accelerate electrons. Combined with phosphor screens and CCD cameras to image charge particle plasmas in ALPHA.

MIDAS **M**aximum **I**ntegrated **D**ata **A**cquisition **S**ystem, a data acquisition system developed and TRIUMF and the Paul Scherrer Institute.

mirror coil The solenoid magnets of the Ioffe-Pritchard trap responsible for the axial confinement of antiatom.

Penning-Malmberg trap A trap with cylindrical electrodes which constrains charged particles using electric fields.

PMT **P**hoto**M**ultiplier **T**ube, a devide used to detect photons.

PWB **P**ad **W**ing **B**oard, a custom circuit board used to read out the cathode pads of the ALPHA-g TPC.

ROOT A data analysis framework developed at CERN.

SDREVC **S**trong **D**rive **R**egime **E**vaporati**V**e **C**ooling, a plasma manipulation technique combining evaporative cooling and the rotating wall technique. It is used at ALPHA to create reproducible plasmas.

SiPM **Si**licon **P**hoto**M**ultiplier, a modern silicon-based photon detector.

spacepoint A three-dimensional point in the TPC reconstruction corresponding to an ionization of the TPC gas.

SVD The **S**ilion **V**ertex **D**etector, the silicon detector used for track and annihilation vertex reconstruction in ALPHA-2.

TDC **T**ime-to-**D**igital **C**onverter, a digitizer which precisely digitizes the start time of a signal.

time walk The effect where for two signals of the same shape but different amplitudes, the larger signal will cross a fixed threshold sooner.

TOF **T**ime-**O**f-**F**light, a method where particle information is extracted from the time between two measurements.

TPC The **T**ime **P**rojection **C**hamber, the drift chamber detector which is used for annihilation track reconstruction in ALPHA-g.

track A group of spacepoints in the TPC reconstruction corresponding to a single charged particle.

TRIUMF A national physics accelerator laboratory in Vancouver, Canada.

vertex The three-dimensional position of an antiproton or antihydrogen annihilation.

WEP The **W**eak **E**quivalence **P**rinciple, the statement that the motion of an object in a gravitational field depends on neither its mass nor composition.

Acknowledgements

Firstly, thank you to my supervisor Makoto. You've given me both the support and feedback I needed to succeed, as well as the freedom and space to explore at my pace. I couldn't have asked for a better supervisor. Thank you as well to the entire ALPHA team at TRIUMF, in particular Andrea, Lars, Konstantin; you helped me hit the ground running at ALPHA.

I thoroughly enjoyed my time at CERN, mostly due to the collection of genuine, hardworking, and fun people at ALPHA. The sense of community and collaboration was awesome; I could (and did!) ask questions about anything to anyone, and the struggles and efforts were shared. In particular, I want to thank the detector team at CERN: Ina, Joseph, Daniel, Pooja, Lukas – we really made it work.

I am incredibly grateful to my parents Rosie and Nigel and my sister Megan for encouraging me and supporting me throughout my journey. And I am thankful for the compassionate and inspiring community of friends I have met during my graduate studies at UBC, without which I surely would have gone insane in the pandemic.

Thank you to Emily for being by my side every step of the way, and for sharing this adventure with me. And of course, to Sally and Charlie.

This project was partially funded by the Canadian Institute of Nuclear Physics 2023 Graduate Fellowship. It was partially performed at the TRIUMF laboratory, which lies on the traditional, ancestral, and unceded territory of the xʷməθkʷəy̓əm (Musqueam) People.

Chapter 1

Introduction

Our current understanding of the motion and interactions of fundamental particles is governed by two theories: Einstein's general relativity describing the force of gravity, and the Standard Model of particle physics describing the weak, strong, and electromagnetic forces. The combination of these theories has proven to accurately model the universe around us, despite the fact that the two frameworks are fundamentally incompatible, with efforts to create a unified theory of quantum gravity so far unsuccessful.

Nevertheless, there are many observations which remain unexplained by the Standard Model and general relativity, indicating that our understanding of the universe is incomplete and motivating searches for new physics. One such unexplained phenomenon is *dark energy*, the unknown force which is required for general relativity to match our observations of the expansion of the universe. Another is the *baryon asymmetry problem*, the dominance of matter over antimatter in our observable universe.

The theoretical description of antimatter in the Standard Model is based on CPT symmetry, an underlying fundamental symmetry between matter and antimatter. However, there is another symmetry between matter and antimatter: their identical gravitational behaviour. This is asserted by the weak equivalence principle (WEP), which is one of the axiomatic assumptions of general relativity.

With unsolved problems indicating that both the Standard Model and general relativity are incomplete, it is important to experimentally verify their underlying assumptions. Testing the applicability of the WEP to antimatter is a natural goal as it lies at the intersection of the two theories, especially since the WEP was formulated before the discovery of antimatter, based solely on observations of matter.

Until recently, a direct test of the WEP on antimatter has never been performed. Such a test requires using electrically neutral antimatter, as electromagnetic forces would otherwise overwhelm gravity. Excluding antineutrons, which are prohibitively difficult to trap and manipulate, the simplest viable system is the antihydrogen atom: the bound state of an antiproton and a positron.

In pursuit of the goal of verifying the fundamental symmetries between matter and antimatter, the ability to perform measurements on cold trapped antiprotons and antihydrogen atoms has been developed at the Antiproton Decelerator (AD) at CERN. A wide range of fundamental properties have been measured and compared to their matter equivalents, such as the antiproton charge-to-mass ratio [6] and magnetic moment [7], the antihydrogen 1S-2S [8] and 1S-2P [9] transition frequencies, and the antihydrogen fine structure [10] and hyperfine structure [11]. These measurements constitute a series of independent, direct tests of CPT symmetry, and have so far found no deviation from the predictions of the standard model.

Recently, the ALPHA (Antihydrogen Laser PHysics Apparatus) collaboration at the AD has extended its focus from CPT symmetry to the WEP. This dissertation presents the first direct measurement of the gravitational free-fall of the antihydrogen atom, conducted using the new ALPHA-g apparatus. Our results [1] are consistent with antihydrogen behaving in accordance with the WEP. This is the first direct test of the WEP which has been carried out on a purely antimatter system in the laboratory, and represents a major milestone towards precision tests.

The ALPHA collaboration is a diverse group from a range of disciplines and with many areas of expertise. During the 15 months I spent at CERN, I took part in the development and data-taking for the ALPHA-g measurement, along with various spectroscopy measurements in the ALPHA apparatus. I contributed to a variety of tasks through shift work, such as optimization and development of particle manipulation techniques, installation of particle detectors and electronics, cryogenic handling, and data analysis. I took initiative to improve the particle tracking algorithms used for the time projection chamber (TPC) detector. I was also responsible for the commissioning of the barrel scintillator (BSC) detector and the development of calibration techniques for this detector, with support from Dr. Andrea Capra and Dr. Ina Carli among others. I was able to successfully implement time-of-flight measurements in this detector for rejecting background events caused by cosmic rays. I then worked closely with Dr. Lukas Golino to include these measurements into the machine-learning-based cosmic ray background rejection algorithm used in ALPHA-g. This led to a factor of 10 improvement in background suppression, which will directly enable new measurement capabilities in ALPHA-g.

Following this introduction, Chapter 2 provides some brief theoretical background on antimatter, CPT symmetry, and the weak equivalence principle. It also provides some motivation for this study, and reviews contemporary experimental efforts towards understanding the gravitational behaviour of antimatter. Chapter 3 describes the apparatus and techniques developed by the ALPHA collaboration to produce and manipulate antihydrogen atoms. Chapter 4 details the particle detector system used to detect and reconstruct the annihilations of antiprotons and antihydrogen atoms. Chapter 5 covers the reconstruction algorithms for the main TPC tracking detector. Chapter 6 presents the commissioning and event reconstruction for BSC detector, Chapter 7 describes the calibration algorithms implemented for this detector, and Chapter 8 details the inclusion of this detector into cosmic ray background rejection algorithms. Chapter 9 reports on the measurement of the effect of gravity on antihydrogen, including the techniques, analysis, and the final results. Finally, Chapter 10 provides concluding remarks.

Chapter 2

Background

This chapter presents the theoretical background and context relevant to the study of the gravitational interactions of antimatter. We begin by reviewing the historical and current understanding of antimatter. We then explore fundamental symmetries in physics, with particular emphasis on CPT symmetry and the weak equivalence principle. The question of how antimatter interacts with gravity is framed within the context of the baryon asymmetry problem, and as an intersection between the standard model and general relativity. Finally, we summarize the theoretical expectations and experimental constraints on the gravitational behaviour of antimatter, and provide an overview of ongoing experimental efforts towards its direct measurement.

2.1 A brief history of antimatter

In 1915, Einstein published his theory of general relativity [12], describing gravity as the curvature of spacetime induced by mass and energy. We will return to this in our discussion of the weak equivalence principle, but it is important to underline that the postulation and discovery of antimatter happened long after our current understanding of gravity was formulated. Einstein's general relativity also serves as a classic success story for theoretical prediction in physics; his field equations give rise to the phenomenon of gravitational waves. Despite Einstein's own skepticism regarding their observability [13], the LIGO collaboration in 2015 confirmed the existence of gravitational waves [14], a century after the theory's publication.

2.1.1 The discovery of antimatter

The story of antimatter begins with a similar prediction and experimental verification. In 1928, Paul Dirac published a relativistic equation describing the electron, reconciling quantum mechanics with Einstein's special relativity [15]. Dirac observed that his equation admitted solutions "equally well to an electron with charge e as to one with charge $-e$" [15]. Although he initially dismissed the negative energy solutions, he later embraced them, identifying them in 1931 as "a new kind of particle, unknown to experimental physics, having the same mass and opposite charge to an electron. We may call such a particle an antielectron" [16].

Dirac was more fortunate than Einstein when it came to experimental validation. Just two years later, in 1933, Carl Anderson observed cosmic ray tracks in a cloud chamber under a constant magnetic field [17]. From the curvature of the tracks, he determined the charge-to-mass ratio of each particle, identifying fifteen instances of Dirac's antielectron. He named this particle the positron.[1]

Dirac predicted the existence of the antiproton during his 1933 Nobel lecture [18] following the discovery of the positron. As further particles were discovered throughout the 1930's and 1940's, it was suspected that each particle might have a corresponding antiparticle. However, antiparticles beyond the positron are not observed naturally, and so it was not until the age of high energy collisions that this theory could be verified.

The antiproton was discovered in 1955 at the Bevatron, a 6.2 GeV proton synchrotron [19]. Protons were collided with a copper target, and negatively charged particles with a momentum of 1.19 GeV were extracted. Their velocity was measured using time-of-flight and Cerenkov counters, confirming their rest mass and hence their identity as antiprotons. The discovery of the antineutron followed a year later in 1956 from a group at the same laboratory [20]. Antiprotons were passed through a liquid scintillator, which would observe the occasional antineutron produced through a charge exchange interaction. The antineutron would then pass undetected through additional scintillator panels, before annihilating in a Cerenkov detector.

As routine production of antiprotons became possible, CERN commissioned the Low Energy Antiproton Ring (LEAR) in 1982 [21]. This machine supplied 100 MeV antiprotons to experiments with a broad physics program. These efforts culminated in the 1995 detection of nine antihydrogen atoms in-flight by the PS210 experiment. Antiprotons passing through a xenon target captured positrons pair-produced in an atom's electric field [22], briefly producing an excited state of antihydrogen in flight. LEAR was decommissioned in 1996, and repurposed into the Low Energy Ion Ring (LEIR), which remains an integral part of the LHC heavy ion program.

The success of LEAR prompted further investment into the new Antiproton Decelerator (AD), which began operation in 2000 [23]. Antiprotons are produced by directing a proton beam into an iridium target, and a

[1] Positron is a contraction of positive electron. He also proposed renaming the electron to the "negatron", a term which regrettably does not see common usage.

magnetic horn focuses antiprotons produced at 3.6 GeV into the decelerator ring. These antiproton bunches are slowed to 5.3 MeV and delivered to experiments such as ATHENA [24], which succeeded in 2002 to produce cold antihydrogen atoms from confined plasmas of antiprotons and positrons. In 2010, the ALPHA collaboration became the first to magnetically trap antihydrogen [25].

2.1.2 Experimental attempts to measure antimatter gravity

Since the mass of a particle and corresponding antiparticle are equal, this suggests that gravity should act equally on both. The work presented in this dissertation joins a long history of attempts to measure the effect of the Earth's gravity on a purely antimatter system (hereafter *antimatter gravity* for brevity), by observing the gravitational free-fall of an antimatter system. While these methods have only placed loose constraints on antimatter gravity, a number of indirect measurements have demonstrated equality between matter and antimatter gravity much more precisely. However, these indirect methods rely on other underlying assumptions, and will be discussed in Section 2.5.3. Here, we discuss only *direct* measurements of antimatter gravity, which involve the observation of the motion of antimatter systems under the influence of gravity.

In the 1960s, Witteborn and Fairbank sought to compare the Earth's gravitational pull on positrons and electrons [26]. Their method was to measure the particle time-of-flight in a vertical drift tube, taking care to shield the particle from all external electromagnetic fields. However, this proved to be prohibitively challenging due to a phenomenon known as patch potentials, local variations in the electric potential along the surface of the metal tube. These can be caused by microscopic impurities in the metal tube's surface causing local charge build up, by grain boundaries, or other effects [27]. The resulting electrostatic forces dwarfed the expected gravitational effect.[2]

Witteborn and collaborators returned to this problem once slow antiprotons became available at LEAR [29]. Their idea was that since antiprotons are much heavier than positrons, they would require less electromagnetic shielding to achieve sensitivity to gravity. However, this project faced issues in trapping and cooling antiprotons while maintaining a sufficiently field-free region for a gravity measurement. The experiment ultimately failed to

[2]Electrons in the conducting walls of the drift tube also experience gravity, forming an electric potential gradient which counteracts gravity. Witteborn and Fairbank were able to confirm that the effect of gravity on these electrons is equal to the effect of gravity on a free electron in the drift tube, to within 9% [28].

produce any results before the shutdown of LEAR [30].

This brings us to the present generation of antimatter gravity experiments. To avoid electromagnetic interference in a free-fall experiment, one must use a neutral form of antimatter. Since antineutrons are difficult to trap and manipulate, the best candidate for a direct measurement of gravity on a bound antimatter system is the antihydrogen atom. Other experiments have been proposed which use an interferometer on a beam of either muonium [31] or positronium [32].

A precursor to the 2023 result [1] was published by the ALPHA collaboration in 2013 [33]. Antihydrogen atoms were released from the ALPHA-2 trap, and a search for an asymmetry in the vertical annihilation positions was performed. An upper limit of 75 was imposed on the ratio of the antihydrogen gravitational and inertial mass. This study was limited by the small vertical distance over which the antiatoms could free fall before annihilating; however, its success paved the way for the design of the current ALPHA-g experiment.

2.1.3 Current antimatter experiments

Today, the Extra Low Energy Antiproton ring (ELENA) further decelerates the AD antiproton beam to 100 keV [34]. This beam remains the world's only source of low energy antiprotons. It fosters an active community of unique experiments, including:

- BASE [35] — tests CPT symmetry by measuring the antiproton magnetic moment [7] and charge-to-mass ratio [6].
- ASACUSA [36] — aims to create a beam of polarized antihydrogen to study the antiproton mass [37] and the zero-field hyperfine splitting.
- PUMA [38] — aims to study the surface nucleon density of radioactive nuclei by transporting antiprotons to CERN's ISOLDE radioactive ion beam facility.
- AEGIS [39] — aims to use a Moire deflectometer to measure Earth's gravity on a beam of antihydrogen.
- GBAR [40] — aims to measure the gravitational free-fall of microkelvin antihydrogen, by photo-ionization of positive antihydrogen ions.

The ALPHA collaboration also remains active at the AD, having performed a wide range of spectroscopic measurements on trapped antihydrogen with the ALPHA-2 experiment. The new ALPHA-g experiment has also yielded recent results in the measurement of gravity acting on antihydrogen, which are presented in this dissertation. A detailed overview of the ALPHA experiment will be given in Chapter 3.

2.2 CPT symmetry

The current generation of AD experiments share a common goal: to study the fundamental symmetries between matter and antimatter. In physics, a symmetry is synonymous with an invariance. If a theory remains unchanged under a certain transformation (such as a translation in space), it is said to possess that symmetry or invariance (in this case, translational symmetry).

There are three fundamental symmetries that were long believed to be obeyed by all physical processes, corresponding to the following transformations:

- Charge conjugation (C): Changes the sign of all charges, effectively exchanging a particle with its antiparticle.
- Parity inversion (P): Reverses the sign of all three spatial coordinates, equivalent to a mirror reflection along all three spatial axes.
- Time reversal (T): Reverses the direction of time.

These symmetries were formalized by Wigner [41, 42] and Dirac [16] from 1927 to 1933. It came as a surprise when in 1957, Wu demonstrated that the weak interaction violates P symmetry [43]. That same year, Landau proposed that the combined CP transformation might still be a true symmetry of nature [44]. This proposal held until 1964, when a group at Brookhaven showed evidence for CP violation in the decay of neutral kaons [45].

Today, only the combined CPT symmetry is thought to hold universally. In 1957, Lüders and Pauli independently proved the CPT theorem, which states that any Lorentz-invariant local quantum field theory must be invariant under a CPT transformation [46]. The assumptions of Lorentz-invariance and locality are foundational to both general relativity and quantum field theory; an observation of CPT symmetry violation would indicate physics beyond these frameworks.

The history of C and CP violation reminds us that even our most deeply held theoretical assumptions must be tested. Since CPT symmetry connects matter and antimatter, its experimental verification necessarily involves antimatter systems, as charge conjugation C corresponds to the symmetry between matter and antimatter. Consequently, many of the experiments at the AD – including the precision spectroscopy program at ALPHA – are designed to search for violations of CPT symmetry.

2.3 The weak equivalence principle

Galileo first formalized the weak equivalence principle (WEP), the statement that the motion of an object in a gravitational field does not depend on its mass or composition, around 1590 [47]. Neglecting air resistance, it guarantees that an apple and a cannonball will fall at the same rate; a fact he is said to have demonstrated by dropping such objects from the Leaning Tower of Pisa. An equivalent statement of the WEP is that an object's gravitational mass (its coupling to gravity) and inertial mass (of Newton's second law $F = ma$) are equal.

In 1907, Einstein extended this idea into what is now called the Einstein Equivalence Principle (EEP) [48], which asserts that the laws of physics are indistinguishable between an accelerated reference frame and one at rest in a gravitational field. These equivalence principles were cornerstones of his 1915 theory of general relativity, which has withstood repeated experimental scrutiny, and remains to this day our most comprehensive theory of gravity.

An observation of a violation of the WEP would necessitate a reformulation of general relativity, and reshape our understanding of gravity. As such, many generations of experiments have searched for WEP violation using a variety of techniques. Some of the early precision tests were performed by Loránd Eötvös from the 1880s to the 1910s [49], using a torsion balance to compare the ratio of gravitational to inertial mass across various materials. Variations on this technique were later used to verify the WEP to a precision of 10^{-13} [50, 51]. The current best measurement of the WEP was performed by the MICROSCOPE experiment, which achieved a precision of 10^{-15} using Earth-orbiting satellites [52].

The WEP asserts that two objects of the same mass will experience the same force of gravity, regardless of their composition; this includes a particle and its antimatter counterpart. Since our understanding of gravity predates the prediction and discovery of antimatter, general relativity does not give any explicit role to antimatter. But we should remember that the WEP was devised based on observations of matter-matter interactions, and is taken as a starting point for general relativity. That antimatter should also obey the WEP makes for a very simple and beautiful theory, but as the last hundred years have demonstrated, it is only by performing experiments that we can verify the laws of nature; beautiful theories do not preclude unexpected results.

2.4 Motivations for the study of the gravity of antimatter

The concept of antigravity, or gravitational repulsion, has long been the subject of science fiction. Beyond pop culture, many theorists have investigated the concept, both historically [53] and recently [54]. It is interesting to consider matter and antimatter which gravitationally repel each other, and whether such a universe could be consistent with our own.

2.4.1 The intersection of gravity and CPT symmetry

There is a solid reason to study the effect of gravity on antimatter from a theoretical perspective. Of the four fundamental forces, the Standard Model describes only three; gravity is left to its own theory, general relativity. Despite decades of effort, a unified theory that consistently combines the two has not been found. Antimatter is deeply connected to CPT symmetry, which arises naturally from the principles of Lorentz-invariance and locality, two axioms of the Standard Model. By testing the WEP on antimatter, we are directly testing the compatibility of the foundations of the two theories. The results of such a test will constrain any unified theory of quantum gravity, and perhaps hint towards how the two theories should be united.

2.4.2 The baryon asymmetry of the universe

There is another motivation for studying this area, based on an unexplained observation. Although the description of antimatter within the standard model entirely matches experiment, there is a gap in our understanding known as the baryon asymmetry problem. As the universe cooled from the very hot and energetic state immediately following the big bang, elementary particles were created via pair production. This mechanism creates particles in pairs – one particle and one antiparticle – and so the total balance between matter and antimatter in the universe is conserved. The problem is that today's universe is overwhelmingly composed of matter. What mechanism gave rise to this dominance of matter? Was less antimatter created, or did some unknown process preferentially remove antimatter from the universe?

Without violating CPT symmetry, such a mechanism is possible if it obeys the three Sakharov criteria [55]:

1. Baryon number violation, i.e. the production of a baryon without the corresponding antibaryon.

2. C and CP symmetry violation.

3. Interactions out of thermal equilibrium.

Baryon number violation is an obvious requirement to explain the imbalance of baryons and antibaryons. However, it has not yet been observed, and is not predicted by the standard model except at very high temperatures [56]. The discovery of a process which does not conserve baryon number, such as proton decay, would be a key piece towards understanding the baryon asymmetry problem.

C and CP symmetry breaking is a requirement for the development of a baryon asymmetry; if C or CP symmetry held, then any process which created a baryon would have a C or CP conjugate process creating antibaryons at the same rate. As we have seen previously, C and CP symmetry breaking has already been observed. However, the degree of symmetry breaking observed so far is too small to be solely responsible for the baryon asymmetry of the universe. Searches for further symmetry breaking processes are underway, for example in the neutrino sector [57].

In a thermal equilibrium, the forward and reverse of a process occur at the same rate; clearly a mechanism to generate a baryon asymmetry must occur outside of thermal equilibrium. Thus understanding the origin of the baryon asymmetry must go hand in hand with understanding the structure and thermal evolution of the universe; understanding how the universe develops will tell us when conditions out of equilibrium would occur. One of the key factors in this development is the gravitational behaviour of the constituents of the universe, and in this way, understanding the gravitational behaviour of antimatter could give us insight into the origins of the baryon asymmetry problem.

2.4.3 Dark energy

There is a final reason to study the gravity of antimatter: we already have indications that our understanding of gravity itself is incomplete. In 1917, Einstein amended his theory of general relativity to include a cosmological constant, which was required in order to keep the size of the universe static [58]. This was soon discarded, when Hubble demonstrated in 1929 that the universe is expanding [59]. However, the idea of a cosmological constant has been revived following observations of the universe's accelerating expansion [60].

The origin of this acceleration is not understood at all, and so has been given the mysterious name of dark energy. A cosmological constant is the leading candidate explanation for dark energy, but not the only one.[3] It is commonly identified with the non-zero energy of the quantum-mechanical vacuum state, which permeates all space. However, the cosmological constant has its own problem: quantum field theory predicts a value which is 120 orders of magnitude greater than the value which would explain the accelerated expansion of the universe [62].

This non-zero quantum vacuum energy corresponds to the creation and annihilation of pairs of virtual particles and antiparticles which occurs continuously in the vacuum. One highly speculative approach is to consider the gravitational interactions between these virtual particles as it relates to the cosmological constant problem [63].

A simpler explanation is that general relativity itself is not correct, and dark energy could arise from a modified theory of gravity. However, in the absence of experimental results that conflict with general relativity, it is difficult to predict how and where such a theory should diverge. Thus testing general relativity in new regimes, such as with antimatter, is essential for identifying possible deviations and guiding modified gravity models.

2.5 Theoretical considerations on repulsive antimatter gravity

It is widely expected among physicists that antimatter should experience gravity identically to matter. This belief is supported by three classic arguments, described by Nieto and Goldman [54], which are frequently cited in discussions of antimatter gravity. We outline these arguments here, taking care to highlight their underlying assumptions. We then consider different antimatter gravity scenarios from a cosmological perspective, and ask if they could lead to an evolution of the universe which matches our observations. Finally, we examine measurements which indirectly constrain the effect of gravity on antimatter.

[3] Indeed, recent DESI results [61] are in tension with a cosmological constant solution.

2.5.1 Arguments for equal matter-antimatter gravitation

The first argument, proposed by Morrison [64], is based on conservation of energy. This argument is stated more simply by Nieto and Goldman [54], whose formulation I describe here. Let us first consider an atom-antiatom pair in a universe where the gravitational behaviour of antimatter and matter are identical. Work can be done to raise the pair to a height h above the Earth's surface, increasing the gravitational potential energy of the pair by $2mgh$. If the pair then annihilates and the resulting photons fall back to the ground, they would be blue-shifted by the Earth's gravitational field, each with energy increased by mgh compared to if they were produced at ground level. This corresponds to the work done to raise the atom-antiatom pair, and so energy is conserved.

Now suppose antimatter and matter repel gravitationally, so that an antiatom on the Earth's surface is repelled upwards with gravitational acceleration $-g$. Then an atom-antiatom pair could be raised to a height h at no net energy cost. If the pair annihilates and the resulting photons fall back to the ground, they would still experience the same blue-shift from the Earth's gravitational field. They would end up with an additional energy of $2mgh$ more than the atom-antiatom pair initially possessed, despite no work being done, in violation of the conservation of energy. Thus, Morrison concluded that matter and antimatter must experience the same gravitational acceleration.

The second classic argument, due to Schiff [65], draws on the results of the Eötvös experiments discussed earlier in the context of WEP tests. The electric field of an atom polarizes the vacuum and generates virtual electron-positron pairs, which contribute significantly to the mass of the atom as quantum-mechanical loop corrections. Since the magnitude of this contribution differs between atoms, any difference in the gravitational mass of electrons and positrons would result in composition-dependent variations in gravitational mass. This would not affect the inertial mass, necessarily leading to a breaking of the WEP. Such deviations are strongly constrained by the Eötvös experiments of Section 2.3, leading Schiff to conclude that antimatter must couple to gravity identically to matter.

The third argument, proposed by Good [66], considers neutral kaon oscillations. A gravitational mass difference between the K^0 and $\bar{K}^0$ would induce regeneration of the long-lived K_L into the short-lived K_S, a CP-violating process. At the time, this was thought to be forbidden. However, CP violation in the neutral kaon system was soon observed experimentally [45], albeit at a smaller magnitude than Good predicted, leaving this argument partially undermined but still of interest.

These arguments have been scrutinized in the decades since, especially with dark energy generating new interest in modified theories of gravity. A recent review by Chardin and Manfredi [67] outlines various theoretical approaches that circumvent, dispute, or reinterpret the constraints posed by Morrison, Schiff, and Good. These debates continue, and there is no consensus as to whether the classical arguments definitively rule out repulsive antimatter gravity.

If these arguments were to be considered indisputable, a direct measurement of the gravitational interaction of antimatter could be seen as redundant or unnecessary. However, there remains significant interest among the community in such a measurement. Each argument relies on underlying assumptions, such as the universality of energy conservation in curved spacetime, or the applicability of general relativity to virtual particles. These arguments can then be considered to raise the stakes: a violation of the weak equivalence principle in antimatter systems would force us to reconsider our underlying assumptions.

2.5.2 Antimatter gravity cosmologies

Even if repulsive gravity between matter and antimatter were not theoretically excluded, it could still be ruled out based on observations of our universe. We must ask whether such a scenario could lead to a universe which is observationally similar to our own, or whether it is ruled out cosmologically. The study of cosmologies with exotic gravitational behaviour of antimatter is not a mature field, and remains largely speculative.

The main cosmology to consider is called the Dirac-Milne universe [68], which contains an equal amount of matter and antimatter, but gives a negative gravitational mass to antimatter. It includes antimatter that is gravitationally repelled by both matter and antimatter, violating CPT symmetry. This mutual repulsion between antimatter particles is necessary to address the observed baryon asymmetry of the universe: if antimatter was mutually repulsive, antistars and other large antimatter systems would not form, allowing it to hide in the cosmic voids. This cosmology can reproduce some cosmological features such as primordial nucleosynthesis rates, but others such as large scale structure have not been reconciled with observations. Such a theory of gravity would be a considerable departure from general relativity, and necessarily violates CPT symmetry, which would have dramatic consequences.

Other more speculative cosmologies with repulsive antimatter gravity have been proposed which would not require CPT breaking [69]. However, none of these models are compatible with the breadth of our observations of the universe, failing to explain features such as baryon acoustic oscillations [70] or the large scale structure of the universe.

Should a difference in gravitational behaviour between matter and antimatter be experimentally confirmed, it would necessitate the development of a new cosmological model capable of reproducing our observed universe. The fact that such a model is not currently known does not rule out the possibility of its existence; rather, it reflects limited theoretical efforts. The existing models can be considered starting points for further development, in the case where experimental results indicate exotic gravitational behaviour of antimatter.

2.5.3 Indirect antimatter gravity constraints

While a direct free-fall-type WEP test has never been performed on antimatter, and cosmological models are not sufficiently developed to rule out exotic gravitational behaviour of antimatter, there are a number of *indirect* limits on antimatter gravity from other experiments. Here, an indirect limit refers to a measurement which must be combined with other assumptions or models in order to constrain antimatter gravity. Such methods are nicely summarized in [71]. The authors conclude that while indirect reasoning can be used to rule out WEP violation in antimatter at the 1% level, this analysis relies on models and assumptions which are also worth testing, and so a direct test is still highly valuable.

One line of reasoning is that only around 1% of the antiproton mass is due to the antiquark mass. Most of the remaining mass originates from binding energy in the gluon field. If only the antiquark mass is treated gravitationally as antimatter, then the maximum possible WEP violation in antiprotons is at the 1% level. A more recent treatment of this logic is presented in [72], which argues that antiquark kinetic energy should also be treated as antimatter. Based on recent lattice QCD calculations, this brings the total gravitational antimatter fraction of the antiproton to around 33%. This argument relies on QCD, and so a precise determination of the gravitational acceleration of antihydrogen could be re-framed in this way as a test of the Standard Model.

One indirect constraint can be applied based on the reasoning of Schiff described in Section 2.5.1, which relies on differing fractions of virtual antiparticles in the nuclei of different elements. This is estimated to constrain antimatter gravity to be equal to matter gravity to within less than 0.5% [71]. However, this method relies on our understanding of the coupling of gravity to virtual particles, which should itself be called into question.

A second category of indirect constraints on antimatter gravity described in [71] rely on time-scaling effects. Should matter and antimatter experience different gravitational accelerations, then a matter and antimatter particle held at the same location will be at different gravitational potentials. As a result, the antimatter particle will exhibit gravitational time dilation, and the internal clocks of the matter and antimatter particles will run at different speeds.

This can result in a number of measurable effects. The first effect is the one proposed by Good (see Section 2.5.1), where the oscillation between kaon states would be shifted from theoretical expectations. A difference between the gravitational acceleration of matter and antimatter of 0.05% has been ruled out based on these observations. Other measurable effects include a difference in the cyclotron frequency between protons and antiprotons, which has been ruled out by the BASE collaboration at the 10^{-7} level [73], or in neutrino oscillations [74] at the 0.02% level.

Whereas gravitational time dilation is usually framed in terms of gravitational potential differences, the analysis of each of these time-scaling measurements relies on the use of an absolute gravitational potential. While this could be valid in general relativity, since we are interpreting a gravity difference between matter and antimatter, we have already departed from the realm of general relativity; it is questionable whether the notion of an absolute gravitational potential is justified in this case. Furthermore, it is not clear which gravitational potential to use; typically the gravitational potential of the galaxy or galactic supercluster is used in these calculations.

In conclusion, there are many indirect constraints which rule out a departure of antimatter from the WEP, at least at the level accessible in the ALPHA-g experiment. However, these indirect constraints each rely on assumptions which are not fully accepted, or at least merit further testing. As such, there is considerable interest among the antimatter community in a direct, free-fall type experiment which can be interpreted in a model-independent fashion.

Chapter 3

Introduction to ALPHA

The ALPHA collaboration was formed by members of the ATHENA collaboration, the first group to produce antihydrogen by combining trapped positrons and antiprotons [24]. The primary physics goal remained: to study the properties of antihydrogen, in order to verify the fundamental symmetries between matter and antimatter. In 2010, ALPHA became the first group to trap antihydrogen atoms [25]. This work continues into the present day with the ALPHA-2 experiment, which has performed a number of spectroscopic measurements on antihydrogen, such as the 1S-2S transition [8], 1S-2P transition [9], fine structure [10], and hyperfine structure [11], among many others.

Building on these successes, ALPHA has recently turned towards a further fundamental symmetry between matter and antimatter: the equivalence of free fall under gravity imposed by the WEP. To this end, the ALPHA-g experiment was commissioned, reusing much of the existing ALPHA infrastructure and technical knowledge. The first results from ALPHA-g are published in [1]; this publication will hereafter be referred to as the *gravity paper*. The details of this measurement are included in this dissertation, as well as in other dissertations [75–78].

This chapter will outline the various systems, components, and techniques which were used to make these measurements. I will not attempt to provide an exhaustive description of the ALPHA apparatus; instead, I will briefly introduce the subsystems and techniques, with a focus on those used for the gravity measurement in ALPHA-g. More detailed descriptions are available in the papers cited above and throughout this chapter. This chapter will focus on the experimental apparatus and individual techniques used; the overall measurement procedure for the gravity paper, as well as its limitations and possible extensions, is described in Chapter 9.

3.1 The AD and ELENA

Antiproton production at CERN begins with a proton beam. Negative hydrogen ions are accelerated to an energy of 160 MeV by the linear accelerator

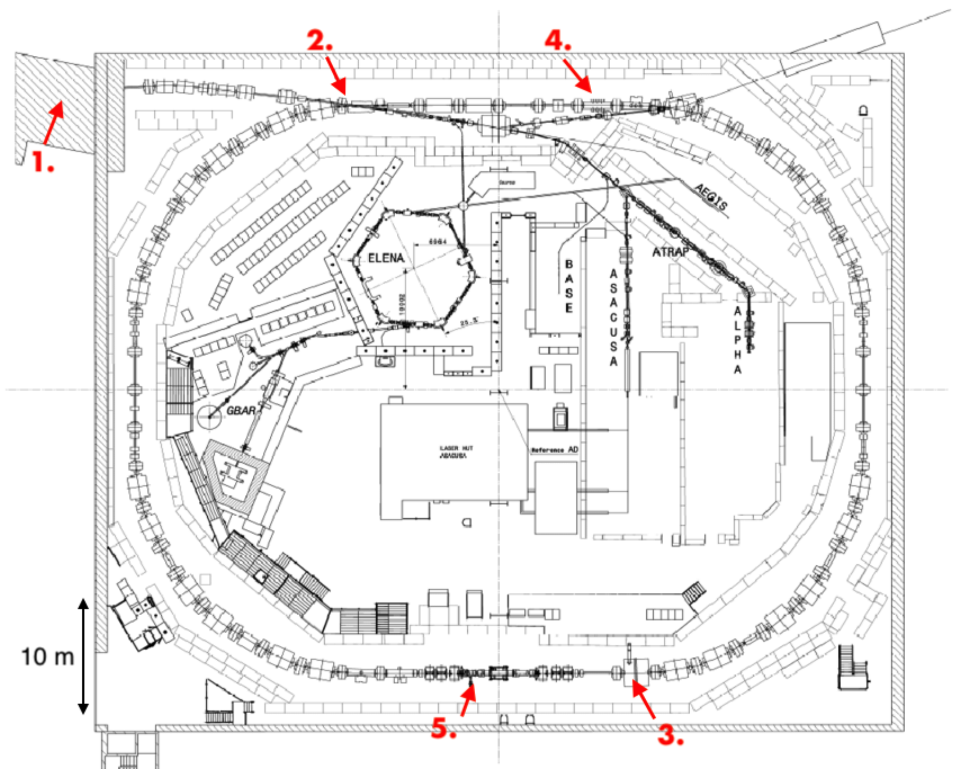

Figure 3.1: A schematic layout of the AD, adapted from [85]. Additional red labels: 1. AD target and magnetic horn location. 2. Beam injection and extraction point. 3. RF cavities. 4. Stochastic cooling. 5. Electron cooling. Also visible are the ELENA decelerator and the ALPHA experimental zone.

LINAC 4 [79]. The electrons are stripped as the beam passes through a stripping foil and into the Proton Synchrotron Booster, an accelerator ring that further accelerates the protons to an energy of 2 GeV [80, 81]. They are then further accelerated to 26 GeV by the proton synchrotron (PS) [82].

This beam is directed into an iridium rod target, where collisions produce a plethora of particle types and energies. A magnetic lens, or horn, collects the antiprotons produced, guiding those with a momentum of 3.5 GeV/c into a beamline. From here, the antiprotons enter the Antiproton Decelerator (AD) [83, 84], a storage ring where they are slowed to a kinetic energy of 5.3 MeV. Dipole magnets bend their trajectory into a ring, and quadrupole magnets focus the beam. The beam deceleration and cooling is facilitated by a number of systems highlighted in Figure 3.1.

Radiofrequency (RF) cavities lower the momentum of the beam in steps from 3.5 GeV/c to 100 MeV/c. The bunch arrival time is synchronized with an oscillating electric field. In this way, the field applied to the antiproton bunches opposes the direction of motion and slows down the bunch.

Between each momentum step, cooling techniques are applied to control the spread of particles in position and momentum space, known as the beam emittance. There are two stochastic cooling stages at 3.5 GeV/c and 2 GeV/c. This process involves using a detector, or "pickup", to detect particles deviating from the desired bunch centre. The pickup signal is processed and fed forward to a "kicker", which applies a corrective electric field.

There are two electron cooling stages at 300 MeV/c and 100 MeV/c. An electron gun is used to create a beam of electrons, and the velocity of this beam is matched to the velocity of the antiproton beam. The two beams are then overlapped. Any antiprotons with higher than average velocity interact with electrons, transferring their energy via Coulomb scattering, and slow down towards the average velocity.

Following the electron cooling at momentum 100 MeV/c, or kinetic energy 5.3 MeV, the antiproton beam is injected into a second storage ring called the Extra Low ENergy Antiproton Ring (ELENA). In this ring, successive decelerations using RF cavities are interleaved with electron cooling stages; a similar technique as the AD. The antiproton beam is slowed to a kinetic energy of 100 keV. It is extracted at one of two extraction points and delivered to experiments such as ALPHA, which require such low kinetic energies to efficiently trap the incoming antiprotons.

An example of an AD and ELENA cycle, including the deceleration and cooling stages, is shown in Figure 3.2. In typical operation, the cycle lasts approximately 120 seconds per bunch delivered. The AD team at CERN have also developed a technique to split each antiproton bunch into four smaller sub-bunches, enabling simultaneous delivery to multiple experiments. As a result, antiprotons were available to ALPHA continuously throughout the run periods.

3.2 Particle trapping

Production of antihydrogen in ALPHA begins with the trapping of antiprotons. A single bunch of antiprotons from ELENA contains approximately 7.5×10^6 antiprotons, of which around 5×10^5 are successfully trapped. Similar numbers of positrons are also generated and captured. These collections of charged particles, known as plasmas, have sufficiently high densities that

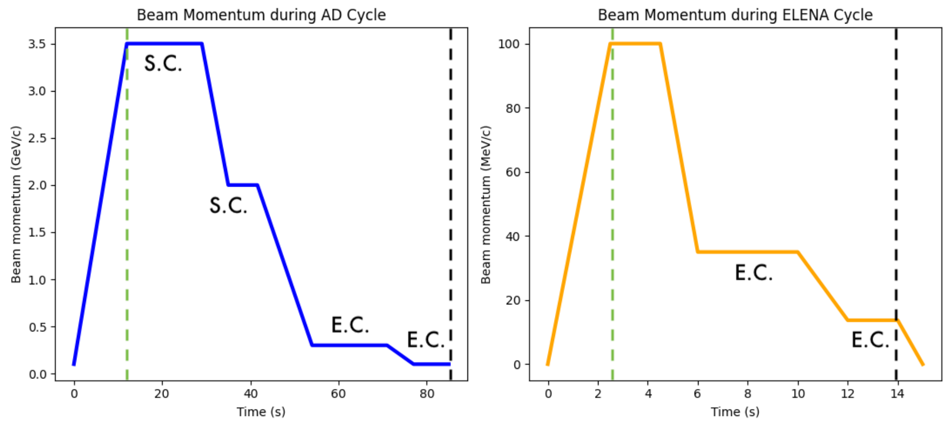

Figure 3.2: Model of beam momentum over an AD (left) and ELENA (right) cycle. Sloped segments indicate beam momentum adjustments using RF cavities. S.C. indicates stochastic cooling periods, and E.C. indicates electron cooling periods. The dashed lines indicate beam injection (green) and extraction (black) times.

the interactions between particles cannot be neglected. The trapping and manipulation of these charged particle plasmas is carried out using a number of Penning-Malmberg traps. The trapping of electrically neutral antihydrogen atoms requires a different type of trap: ALPHA uses a superimposed magnetic minimum trap to confine these atoms, relying on the interaction of the antihydrogen magnetic moment with inhomogeneous magnetic fields. In this section we will discuss the methods used to trap these charged particle plasmas and neutral atoms, both from a theoretical and practical standpoint.

3.2.1 Charged particle trapping

First, let us consider a single charged particle of charge q and mass m, propelled by an electric field $\vec{E}$ towards a minimum in the electrostatic potential V. It experiences a force $\vec{F}$ and subsequent acceleration $\ddot{\vec{x}}$ (using the dot notation for time derivatives) given by:

$$\vec{F} = m\ddot{\vec{x}} = q\vec{E} = -q\vec{\nabla}V \tag{3.2.1}$$

One could trap a charged particle by placing it in a quadratic potential well, so that any motion away from the minimum results in a force returning the particle to the centre. This works well in one dimension; unfortunately,

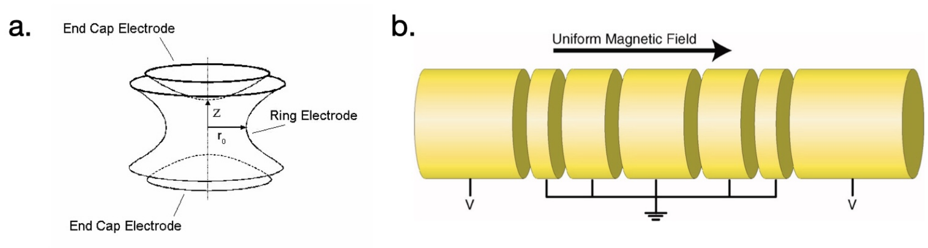

Figure 3.3: (a.) The hyperbolic electrodes of a Penning trap, from [89]. A uniform magnetic field along the Z axis is applied by solenoid magnet. (b.) The cylindrical electrodes of a Penning-Malmberg trap, from [90].

it was proven by Earnshaw in 1842 [86] that no static electric field configuration can produce a three-dimensional potential minimum. Today, this is framed in terms of Poisson's equation $-\nabla^2 V = 0$, a consequence of Maxwell's equations.

Since a charged particle cannot be trapped by static electric fields alone, let us consider now the motion of a particle in a constant magnetic field $\vec{B}$:

$$\vec{F} = m\ddot{\vec{x}} = q\dot{\vec{x}} \times \vec{B} \tag{3.2.2}$$

The particle will experience a force perpendicular to its velocity, resulting in circular motion in the plane perpendicular to the magnetic field, known as cyclotron motion. The frequency of this motion is the cyclotron frequency $\omega_c = qB/m$.

The combination of these two ideas led to the development of the Penning trap [87, 88], where a constant magnetic field confines the particle to circular motion in the radial dimensions, while an electric potential well traps particles axially. The electric potential configuration is achieved by two hyperbolic end caps and a hyperbolic ring electrode, as shown in Figure 3.3.a. The magnetic field is applied by a superimposed solenoid magnet.

Achieving a quadratic electrostatic potential on-axis necessarily also creates a radial electrostatic potential. This leads to a Penning trap potential of the form:

$$V(\vec{x}) = \frac{V_0}{R^2}\left(2z^2 - x^2 - y^2\right) \tag{3.2.3}$$

Here, V_0 sets the magnitude of the potential, R is a characteristic trap dimension, and x, y, z are the components of the position $\vec{x}$. Combining this electrostatic potential with a constant magnetic field $\vec{B} = B_0\hat{z}$ and solving the equations of motion (the sum of equations 3.2.1 and 3.2.2) yields decoupled axial and radial motions. The axial motion is described by a simple harmonic motion where ω_z is called the axial bounce frequency:

$$\ddot{z} = -\frac{4qV_0}{mR^2}z = \omega_z^2 z \tag{3.2.4}$$

The radial motion is more complicated; the radial electric field perturbs the solution from the simple cyclotron motion of a uniform magnetic field. Introducing $u(t) = x(t) + iy(t)$ as a combined spatial coordinate, we can write the equation of radial motion as:

$$\ddot{u} = i\omega_c\dot{u} - \frac{1}{2}\omega_z u \tag{3.2.5}$$

This has two independent solutions $u(t) = u_0 e^{-i\omega_\pm t}$. Here $\omega_c = qB/m$ is the classic cyclotron frequency in the absence of the electrostatic field, and

$$\omega_\pm = \frac{\omega_c}{2} \pm \frac{1}{2}\sqrt{\omega_c^2 - 2\omega_z^2} \tag{3.2.6}$$

The validity of this solution depends on $\omega_c^2 - 2\omega_z^2 > 0$, called the trapping condition, which requires that the magnetic field be strong compared to the radial electrostatic field in order for the particles to be radially confined.

The frequency $\omega_+ \sim \omega_c$ is called the modified cyclotron frequency. The solution $u(t) = u_0 e^{-i\omega_+ t}$ is similar to the motion of a charged particle in a constant magnetic field, in the absence of electrostatic fields. The frequency $\omega_- \sim \frac{\omega_z^2}{2\omega_c}$ is called the magnetron frequency, and the solution $u(t) = u_0 e^{-i\omega_- t}$ is related to the $\vec{E} \times \vec{B}$ drift a charged particle experiences in constant electrostatic and magnetic fields. The general solution to the radial equation of motion is a linear combination of the two solutions. As a result, the motion of a particle trapped in a Penning trap can be broken down into three separate components:

1. The simple harmonic axial motion with frequency ω_z.
2. The slow circular magnetron motion with frequency ω_-.
3. The fast circular cyclotron motion with frequency ω_+.

An example of the path traced by a particle in a Penning trap is shown in Figure 3.4.

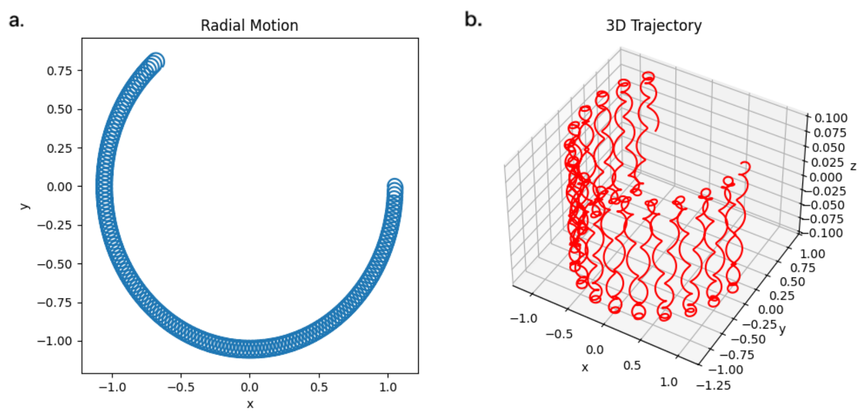

Figure 3.4: Radial (a.) and three-dimensional (b.) motion of a charged particle in a Penning trap. Axes are in arbitrary units; the length scale is set by the strength of the trapping potentials.

3.2.2 Non-neutral plasma trapping

In ALPHA and many other practical applications, one would like to simultaneously trap many charged particles. The hyperbolic geometry of a traditional Penning trap is not well-suited for this, as the electric fields are uniform over only a small spatial volume. Furthermore, the hyperbolic electrodes are mechanically cumbersome and offer limited flexibility for particle manipulation.

A Penning-Malmberg trap offers an alternative solution, formed by a stack of cylindrical electrodes and a superimposed magnetic solenoid shown in Figure 3.3.b. By applying a voltage to two electrodes separated by several grounded electrodes, a confining axial electric potential is produced. Combined with a constant axial magnetic field, this configuration confines charged particles. The electrode cylinder is open at both ends, allowing access for particle loading or ejection in either axial direction.

The main downside to this cylindrical geometry is that the electric field is no longer perfectly quadratic. As a result, the motion of particles in a Penning-Malmberg trap can no longer be described analytically and must be modelled by simulation. However, this geometry allows trapping of significantly larger numbers of particles over an extended axial region. Once multiple particles are introduced, the mutual repulsion between them makes describing the situation mathematically more complicated. This system of interacting particles is referred to as a non-neutral plasma.

A traditional quasi-neutral plasma is a state of matter where the constituent particles have enough energy to overcome binding energies between each other, and contains similar numbers of positive and negative particles. These plasmas exhibit collective effects such as Debye screening, where particles on the plasma skin will rearrange themselves to shield the interior of the plasma from an electric field. By contrast, a non-neutral plasma is typically composed of charged particles of a single species. As a result, it has a net charge and generates significant internal electric fields. Confinement requires a combination of electric and magnetic fields, such as those in a Penning-Malmberg trap, to counteract the internal repulsive forces.

The rotational symmetry of the Penning-Malmberg trap leads to a rotational invariance, and conservation of the canonical angular momentum L of the plasma [91]:

$$L = \sum_{j=1}^{N} \Big(m v_{\theta,j} r_j + \frac{q}{c}\frac{B r_j}{2} r_j \Big) = const. \tag{3.2.7}$$

In a sufficiently strong magnetic field, the second term due to the magnetic field will dominate the first term due to particle angular momentum, and the condition simplifies to a constraint on the mean radius squared of the particles:

$$\sum_{j=1}^{N} r_j^2 = const. \tag{3.2.8}$$

Given an initially trapped plasma, this leads to an upper bound on the fraction of trapped particles which can reach the trap wall [92]. Typically, a confined plasma will reach an equilibrium state with a constant r-dependent rotational velocity and density profile.

3.2.3 Antihydrogen trapping

The Penning-Malmberg trap relies on particles being charged, allowing them to interact with applied electric and magnetic fields. This trap works well for both antiprotons and positrons, but trapping the neutral antihydrogen atom requires a different approach.

A neutral atom can still experience a force in a magnetic field due to its magnetic dipole moment. This is a quantum mechanical property which is a combination of the orbital dipole moment, due to the positron orbit acting as a current loop, and the intrinsic spin dipole moment of the positron and antiproton. The orbital moment is determined by the orbital quantum number of the atomic state; this is zero in the ground state, which is by far the most likely state in a Penning-Malmberg trap. As a result, the magnetic dipole moment is given entirely by the spin dipole moment. The axial component of the dipole moment is then $\mu_z = -g_s \mu_B m_s$, where g_s is the spin g-factor, μ_B is the Bohr magneton, and $m_s = \pm 1/2$ is the spin quantum number of the state.

The Zeeman energy, corresponding to the interaction of the magnetic dipole moment with an external magnetic field, is given by:

$$U = -\vec{\mu} \cdot \vec{B} \tag{3.2.9}$$

where $\vec{\mu}$ is the magnetic dipole moment of the antihydrogen atom and $\vec{B}$ is the applied magnetic field.

The torque resulting from the magnetic field will cause the dipole moment to either align or anti-align with the applied field, and so the potential experienced by the antihydrogen atom will either increase or decrease by μB. Those antiatoms which see a decreased potential by μB are known as high field seekers. Since it is impossible to produce a magnetic field with a local maximum [93], these antiatoms are deemed untrappable. The antiatoms which see an increased potential by μB are known as low field seekers. If moved away from the minimum of a magnetic field, they will experience a force returning them to the minimum.

This is the basis of the Ioffe-Pritchard trap [94]. A quadrupole, sextupole, or octupole magnet produces a field with strength increasing with radius, which creates radial confinement. A solenoid applies a constant axial magnetic field, which is enhanced at the ends by two small mirror coils, confining the atoms axially. The geometry of a simple Ioffe-Pritchard trap is shown in Figure 3.5.

3.3 Annihilation

The process of annihilation occurs when a particle collides with its antiparticle. Both the particle and antiparticle are destroyed, with their mass converted to energy following $E = mc^2$. This results in the creation of other particles (annihilation products) following the conservation of energy, momentum, spin, charge, and baryon and lepton number.

The complexity of annihilation outcomes increases with the energy available in the collision. If either the particle or antiparticle have a large kinetic

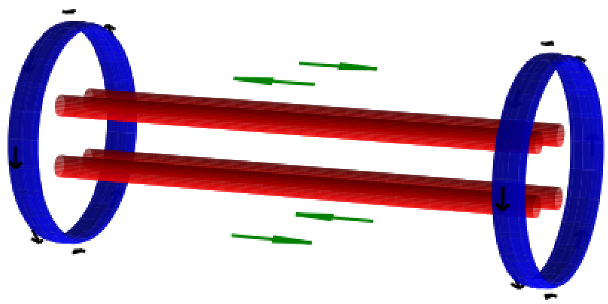

Figure 3.5: A simple Ioffe-Pritchard trap. Ioffe bars (red) with alternating current flow produce a quadrupole magnetic field, while mirror coils (blue) increase the magnetic field at the ends of the trap.

energy, then many particles can be created; this is the principle of electron-positron and proton-antiproton colliders. In ALPHA, the highest kinetic energy involved is the 100 keV with which antiprotons are delivered from ELENA. This is far smaller than the 938.8 MeV/c^2 rest mass of the antiproton, and so we will limit our discussion of annihilation to the case where both particles are (almost) at rest.

3.3.1 Positron annihilation

By far the most probable channels in electron-positron annihilation are the production of either two photons $e^+ + e^- \longrightarrow 2\gamma$ or three photons $e^+ + e^- \longrightarrow 3\gamma$. The electron and positron have a very small rest mass of 511 keV/c^2; the only other particles with low enough rest mass to be created are neutrinos, and these are suppressed by the relative weakness of the weak force.

A low-energy positron is typically slowed down and captured by an electron to form positronium. After a short time (either 0.12 ns or 13.6 ns), the electron and positron annihilate. The lifetime and number of photons produced depends on the alignment of the electron and positron spins. If the spins are anti-aligned (as in para-positronium), conservation of charge parity prohibits the production of an odd number of photons, leading to the dominance of 2γ channel. If the electron and positron spins are aligned (as in ortho-positronium), the production of an even number of photons is prohibited, resulting in primarily 3γ production. It is also possible for a positron to annihilate directly on an electron without forming positronium, in which case the 2γ production is most probable [95].

The relative probability of the 2γ and 3γ annihilations depends on the formation rates of para-positronium and ortho-positronium. With a triplet of three spin configurations, ortho-positronium is approximately three times more likely to be formed than para-positronium, which has only one spin configuration. However, the exact ratio varies slightly depending on the material [96].

The single photon channel $e^+ + e^- \longrightarrow \gamma$ is forbidden by conservation of momentum. In the centre of mass frame of reference, in which the electron-positron pair has zero total momentum, the resulting photon must also have zero momentum. Since the photon energy and momentum are related by $p = Ec$, this photon would have zero energy, which cannot occur.

In the two photon case, the same argument can be applied to show that the photons must be created back-to-back, with 511 keV of energy each. This is incidentally the principle by which positron emission tomography works, and so we can thank conservation of momentum in annihilations for saving countably many lives.[4]

3.3.2 Antiproton annihilation

Antiproton annihilation is a much messier affair than electron-positron annihilation due to two main factors:

1. Protons and antiprotons have a higher rest mass of 938.8 MeV/c^2, increasing the number and types of annihilation products.
2. Protons and antiprotons are ensembles of three quarks, and not elementary particles.

The proton is composed of two u quarks and one d quark, while the antiproton is composed of two $\bar{u}$ antiquarks and a $\bar{d}$ antiquark. Instead of an antiproton annihilation, we should consider the annihilation of an antiquark with a quark. An antiquark in an antiproton can also annihilate with a quark in a neutron, and so we should consider both antiproton-proton and antiproton-neutron annihilations.

The pion final state branching ratios of these two processes are shown in Table 3.1. Almost all annihilations end in a state with some number of charged pions $\pi^\pm$ (the $u\bar{d}$ and $d\bar{u}$ bound states of a quark and an antiquark) and neutral pions π^0 (the $(u\bar{u} + d\bar{d})/\sqrt{2}$ bound state). Antiproton-proton

[4] Positron emission tomography is often cited as a practical application of antimatter.

annihilation results in an equal number of π^+ and π^- mesons due to conservation of charge, while antiproton-neutron annihilation results in one more π^- meson.

(a) Antiproton-proton (from [38, 97])

Pion final state	%
$\pi^0\pi^0$	0.028
$\pi^0\pi^0\pi^0$	0.76
$\pi^0\pi^0\pi^0\pi^0$	3.0
$\pi^+\pi^-$	0.32
$\pi^+\pi^-\pi^0$	6.9
$\pi^+\pi^-\pi^0\pi^0$	9.3
$\pi^+\pi^-\pi^0\pi^0\pi^0$	23.3
$\pi^+\pi^-\pi^0\pi^0\pi^0\pi^0$	2.8
$\pi^+\pi^-\pi^+\pi^-$	6.9
$\pi^+\pi^-\pi^+\pi^-\pi^0$	19.6
$\pi^+\pi^-\pi^+\pi^-\pi^0\pi^0$	16.6
$\pi^+\pi^-\pi^+\pi^-\pi^0\pi^0\pi^0$	4.2
$\pi^+\pi^-\pi^+\pi^-\pi^+\pi^-$	2.1
$\pi^+\pi^-\pi^+\pi^-\pi^+\pi^-\pi^0$	1.9

(b) Antiproton-neutron (from [38, 98])

Pion final state	%
$\pi^-\pi^0$	0.75
$\pi^-n\pi^0$ $(n > 1)$	16.9
$\pi^-\pi^-\pi^+$	0.23
$\pi^-\pi^-\pi^+\pi^0$	17.0
$\pi^-\pi^-\pi^+n\pi^0$ $(n > 1)$	39.7
$\pi^-\pi^-\pi^-\pi^+\pi^+$	4.2
$\pi^-\pi^-\pi^-\pi^+\pi^+\pi^0$	12.0
$\pi^-\pi^-\pi^-\pi^+\pi^+n\pi^0$ $(n > 1)$	6.6
$\pi^-\pi^-\pi^-\pi^-\pi^+\pi^+\pi^+n\pi^0$ $(n \geq 0)$	0.35

Table 3.1: Measured pionic final states and branching ratios for antiproton-proton and antiproton-neutron annihilation. In (b), n refers to the number of π^0 mesons. Note that the total adds to 97.7% for each table, since not all decay channels are shown.

The $\pi^\pm$ rest mass is $139.6\,\mathrm{MeV/c^2}$ and its lifetime at rest is $\tau = 26\,\mathrm{ns}$ [56]. Pions are typically produced with around 100 MeV of kinetic energy; a charged pion at this energy has a mean decay length of $L = \gamma\beta c\tau = 11\,\mathrm{m}$, and so they can be considered stable particles over the length scale of the ALPHA detectors.

The π^0 has rest mass $135\,\mathrm{MeV/c^2}$, but its lifetime at rest is much shorter at $8.5 \times 10^{-17}\,\mathrm{s}$ [56]. The same calculation yields a mean decay length of 36 nm, and so we should consider the decay of the neutral pion. The dominant mode with 99% of the branching ratio is $\pi^0 \longrightarrow 2\gamma$, and most remaining decays are $\pi^0 \longrightarrow e^- + e^+ + \gamma$ [56]. These photons further interact via pair production $\gamma \longrightarrow e^- + e^+$ in the electromagnetic field of a nucleus, with an attenuation length on the order of 1 cm. The resulting state as observed in the ALPHA particle detectors is N π^+ mesons, either N or $N+1$ π^- mesons, and some number of e^-e^+ pairs from displaced photon pair production vertices.

A further complication arises when we consider than our antiprotons are not annihilating on free protons and neutrons, but on those contained within a nucleus. Some number of produced pions may travel further into the nucleus and be absorbed, reducing the observed pion multiplicity of the annihilation. This may have further effects such as causing a nucleon excitation and subsequent decay, or producing additional particles through a further pion-nucleon cascade [99, 100].

3.3.3 Antihydrogen annihilation

The annihilation of antihydrogen atoms can be thought of as the combination of the antiproton annihilation and the positron annihilation processes. When an antihydrogen atom comes into contact with a regular atom, the antiproton will annihilate with a nucleus as discussed in Section 3.3.2 and the positron will annihilate with an electron as discussed in Section 3.3.1.

The positron annihilation contributes only a small fraction of the overall energy and produces only two or three photons. These are often difficult to detect among the higher multiplicity of secondary particles from the antiproton annihilation. In some detection schemes, the back-to-back photon signal of the positron annihilation can be used as a supporting signal to distinguish antihydrogen annihilations. Primarily, however, the signals of the antiproton annihilation and antihydrogen annihilation are practically identical, and the two must be distinguished by a well-designed experimental procedure.

3.4 The ALPHA apparatus

The ALPHA experiment is a mature project which has made many distinct measurements in the same apparatus. Although they are often said to be separate experiments, the modular design means much of the apparatus is shared between ALPHA-2 and ALPHA-g. The ALPHA-2 and ALPHA-g atom traps are variations on the same design: a Penning-Malmberg charged particle trap and a superimposed Ioffe-Pritchard neutral atom trap. The entire ALPHA apparatus is shown in Figure 3.6. This section will describe the components of the ALPHA apparatus relevant for gravity experiments in ALPHA-g: the catching trap, positron accumulator, beamline, beam diagnostics, annihilation diagnostics, and the ALPHA-g trap.

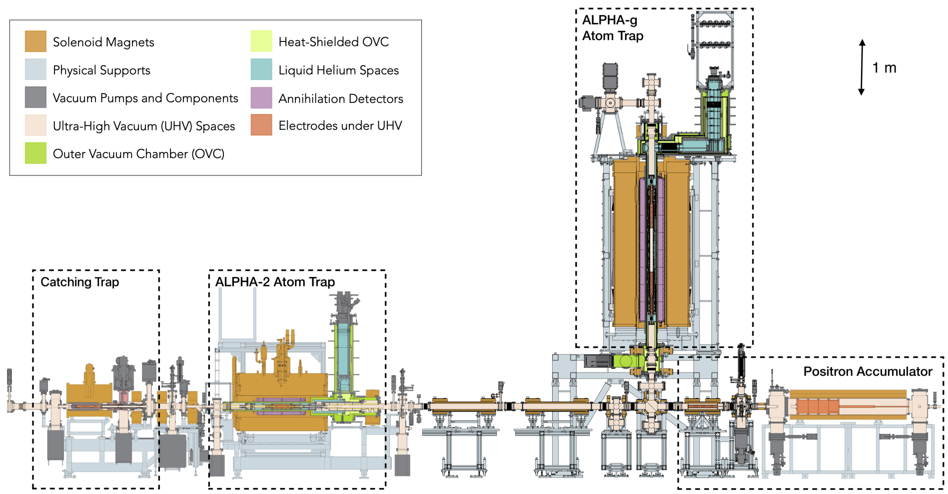

Figure 3.6: Cross-section of the full ALPHA apparatus. Adapted from [101].

3.4.1 Catching trap

The ALPHA catching trap is a Penning-Malmberg trap designed to receive and cool bunches of antiprotons from ELENA. A superconducting solenoid provides a 3 T axial magnetic field for radial confinement, while axial confinement is provided by a set of 20 cylindrical electrode segments. To facilitate the superconducting magnet, the apparatus is enclosed by a closed-system liquid helium cryostat. Scroll pumps, turbo pumps, and ion pumps are used to create a vacuum of 10^{-10} mbar. This is further enhanced by cryopumping which improves the vacuum to better than 10^{-13} mbar [102].

Upon arrival, an antiproton bunch passes through a 1 µm aluminium degrader foil, optimized for the 100 keV ELENA ejection energy [103]. A fraction of the bunch will annihilate on the foil, but the remaining antiprotons will be slowed to even lower energies, facilitating their capture.

The catching trap features two specialized high-voltage electrodes which can be biased to up to ± 5 kV. The downstream electrode is held at -5 kV, and the upstream electrode is ramped to the same voltage immediately after the bunch arrival. This confines antiprotons with energy less than 5 keV between the two high voltage electrodes, as discussed in Section 3.2.1. This process is depicted in Figure 3.7, which shows the resulting axial potential before and after the upstream electrode is raised.

This process also includes loading an electron plasma into the catching

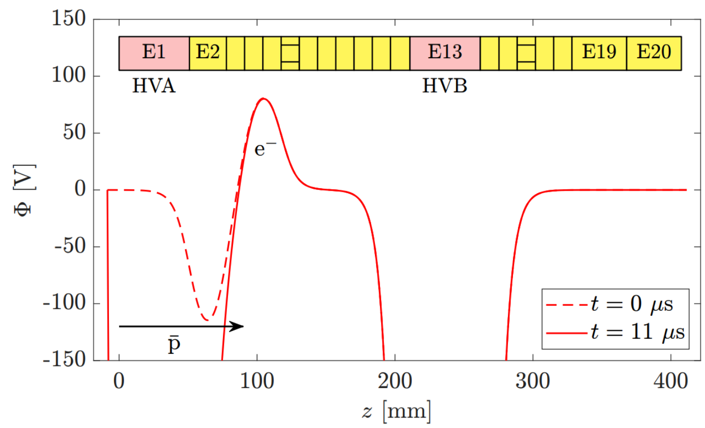

Figure 3.7: The catching trap electrodes (top) and axial potential (bottom). The potential is shown before (dashed) and after (solid) HVA is ramped to $-5\,\text{kV}$. Figure from [104].

trap prior to the antiproton arrival, to facilitate electron cooling. This and other antiproton plasma preparation techniques will be discussed in Section 3.5.1. Following this preparation, a potential gradient is applied using the electrodes, to kick the antiprotons downstream out of the catching trap and into the beamline.

3.4.2 Positron accumulator

The positrons used in ALPHA are created by the β^+ decay of a sodium-22 source with a half-life of 2.6 years [105] and an activity of 2.8 GBq [106]:

$$^{22}_{11}\text{Na} \rightarrow ^{22}_{10}\text{Ne} + e^+ + \nu_e \tag{3.4.1}$$

They typically have energies on the order of an MeV, too high for trapping. To slow down the positrons, a solid Ne moderator is grown on the surface of the source. While the majority of positrons annihilate within the moderator, a small fraction stop near the surface of the moderator and survive. These positrons are released into the vacuum chamber with an energy of around 50 eV.

These moderated positrons are electromagnetically steered into a modified Penning-Malmberg trap known as a Surko trap [107]. A 0.15 T magnetic field confines the positrons radially, and seven cylindrical electrodes provide axial confinement. The diameter of these electrodes increases from 12.7 mm

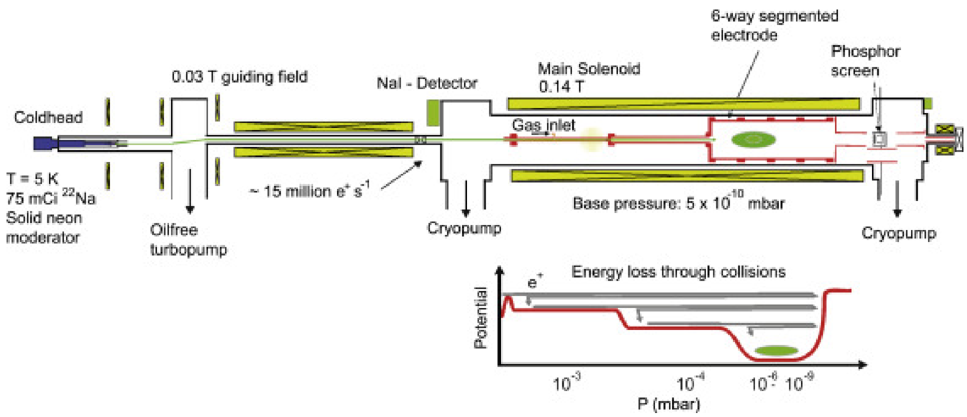

Figure 3.8: A diagram of the positron accumulator, taken from [106]. The inlaid diagram shows the axial potential and pressure.

for the first two electrodes, to 30.5 mm for the third electrode, and finally to 200.7 mm for the last four electrodes. A pressure gradient is established by injecting a nitrogen buffer gas near the narrow electrodes and pumping it away near the wider ones.

Both the gas pressure and electric field are highest near the narrow electrodes, as shown in Figure 3.8. The incoming positrons are initially confined along the full length of the trap. They collide inelastically with N_2 molecules, losing energy via electronic excitation. As their energy decreases, they become increasingly confined to the low-pressure, low-field region within the larger-radius electrodes.

Finally, a technique known as a rotating wall electrode is applied to compress the positron plasma. One of the electrodes is divided into six segments to accommodate this; the procedure will be discussed in Section 3.5.1. The result is a dense, cold plasma which can be transferred to the main ALPHA trap for use in antihydrogen formation. A typical accumulation cycle of 100 seconds yields around 3×10^6 positrons [1].

3.4.3 Beamline

The overall design of the ALPHA apparatus requires the delivery of antiprotons and positrons to both the ALPHA-2 and ALPHA-g atom traps. Since ALPHA-g requires a vertical orientation, while ELENA and all existing ALPHA apparatus is horizontally aligned, transferring both species to ALPHA-g while maintaining the possibility of positron transfer to ALPHA-2 is not a trivial problem.

Moving a plasma between Penning-Malmberg traps is already challenging, as it involves guiding particles through regions with strong magnetic field gradients. To understand the impact of these gradients, consider a particle of charge q and mass m in the trap at a Larmor radius r and with velocity components $v_\perp$ and $v_\parallel$ with respect to the magnetic field B:

$$\frac{v_\perp}{r} = \frac{q}{m}B \tag{3.4.2}$$

As B decreases, r will increase and $v_\perp$ will decrease. Increasing r obviously causes issues for particle transfer; a radial plasma expansion will lead to particles annihilating on the walls of the beamline. Conversely, as B increases, r decreases and $v_\perp$ increases. Since magnetic fields do no work, the total kinetic energy is conserved, and so this increase in $v_\perp$ must correspond to a decrease in $v_\parallel$. If $v_\parallel$ is decreased to zero, then the particle will reverse direction and begin travelling back towards lower B, in a phenomenon known as magnetic mirroring. In these ways, any increase or decrease in B can cause significant losses during a transfer between two Penning-Malmberg traps.

These effects and many others were extensively studied via simulation in the design of the ALPHA beamline, which is detailed in [101]. Seven sections of water-cooled solenoids provide a weaker magnetic field to guide both positrons and antiprotons. The polarity of the current in each solenoid can be reversed, allowing a suitable magnetic field configuration for transfers into either ALPHA-2 or ALPHA-g. The most complicated section of the beamline is called the interconnect, which is responsible for the bending of antiprotons and positrons from opposite sides 90° upwards into ALPHA-g without significant beam degradation, while also having a configuration that allows positrons to be guided through towards ALPHA-2.

3.4.4 Sticks

To monitor and debug the transfer of particles at various stages during preparation and transfer, ALPHA uses a number of diagnostic stations affectionately known as sticks: linear actuators which are able to move a number of different instruments into the beam path. The specific instruments available on each stick vary depending on its location and usage. I will give a brief overview of the most common instruments here.

Electron guns are used to load a plasma of electrons into any of the Penning-Malmberg traps. This can be useful for trap diagnostics, and is also used in electron cooling of antiprotons. A barium oxide filament is heated through ohmic heating by an applied voltage. This excites electrons in the barium oxide, which has a low work function of 1 eV to 2 eV, above which electrons will escape the material. These electrons are then accelerated by a voltage difference between two plates. The emission current, typically a few hundred µA, is monitored and a feedback loop adjusts the applied voltage to maintain a stable emission of electrons.

Faraday cups are a simple diagnostic tool to destructively measure the total charge of a group of charged particles, used for counting the number of charged particles in a plasma. The particles induce a charge on a metal cup as they strike it, which is read out. Around 10^6 elementary charges will create a measurable current.

Microchannel plates (MCPs) are used to image plasmas. An MCP consists of a resistive plate which has many small (5 µm to 20 µm diameter) cylindrical channels through it. A charged particle striking the MCP will produce electrons, which are accelerated through the channels by an applied voltage of around 1 kV. As the electrons strike the walls of the channels, further electrons are dislodged from an emissive coating. The number of electrons is amplified at each successive interaction, increasing the signal by almost 10^6. Behind the MCP is a phosphor screen, a panel coated with a material which creates scintillation light after it is excited by an incident charged particle. These are paired with a mirror at 45° which directs the scintillation light through a vacuum window towards a charge-coupled device (CCD) camera. The highly variable gain of the MCP makes this system suitable for imaging plasmas ranging from one to 10^8 particles.

These tools combine to provide essential feedback for beam alignment, plasma manipulation, and real-time monitoring during antiproton and positron transfer operations.

3.4.5 Annihilation diagnostics

The diagnostic sticks work well for understanding plasmas which are controlled sufficiently to be propelled out of a Penning-Malmberg trap towards them. However, additional diagnostic tools are required for improving the transfer of antiprotons and positrons between traps, as well as understanding losses during particle preparation.

A number of plastic scintillator panels are used for this purpose. These are large (typically 0.4 m × 0.6 m × 2.2 cm) sheets of polyvinyltoluene doped with a florescent scintillator material. When a charged particle passes through the scintillator panel, it will excite a molecule of the scintillator. When it de-excites, the molecule releases a photon of the characteristic wavelength.

A minimum-ionizing charged particle (such as a charged pion from an annihilation) will produce around 2×10^4 photons when passing through one of these panels. Scintillators will be discussed further in Section 4.3.1.

The produced photons will propagate inside the plastic scintillator via total internal reflection before being collected by photosensors mounted on the edge of the panel. The panels surrounding the catching trap and ALPHA-2 feature a light guide and PMT, while the newer panels used for ALPHA-g rely on SiPMs. The details of these technologies will be discussed in Section 4.3.2.

The scintillator panel + photosensor system is used throughout ALPHA to detect charged particles resulting from antiproton annihilation. They are usually positioned in pairs, so that the coincidence between the two panels can be used to suppress false triggers due to electronic noise or stray photons, improving the signal-to-noise ratio. The recovery time of the system is on the order of 50 ns, fast enough to avoid saturation and be able to count many annihilations in quick succession.

In addition to the plastic scintillator panels, ALPHA employs a number of thallium-doped cesium iodide CsI(Tl) detectors. This is a dense inorganic crystal scintillator, which is directly coupled to a PMT to detect ionizing particles in much the same way as the plastic scintillator panels. The CsI(Tl) detectors are particularly efficient at stopping and detecting the 511 keV photons produced in the annihilation of a positron with an electron, and are used primarily for debugging positron losses during transfers between traps.

3.4.6 ALPHA-g

At its core, ALPHA-g is an Ioffe-Pritchard magnetic trap superimposed on a Penning-Malmberg charged particle trap. This is a proven design which has been extensively tested by the ALPHA collaboration with the ALPHA-2 atom trap, discussed elsewhere [106]. Whereas ALPHA-2 was designed as a multi-purpose antihydrogen trap, ALPHA-g was purpose built for the measurement of gravity on antihydrogen.

A schematic overview of ALPHA-g is shown in Figure 3.9.a. The design was intended to be entirely symmetric, to minimize the systematic effects on a measurement of antihydrogen released from the centre of the apparatus by ensuring systematic bias in field gradients or trap structure affects upward and downward-released antihydrogen equally. The initial trap design can be separated into three distinct parts: the bottom trap (a combined magnetic moment and Penning-Malmberg trap), the top trap (a mirror copy of the bottom trap), and the central precision region (a purely magnetic trap with

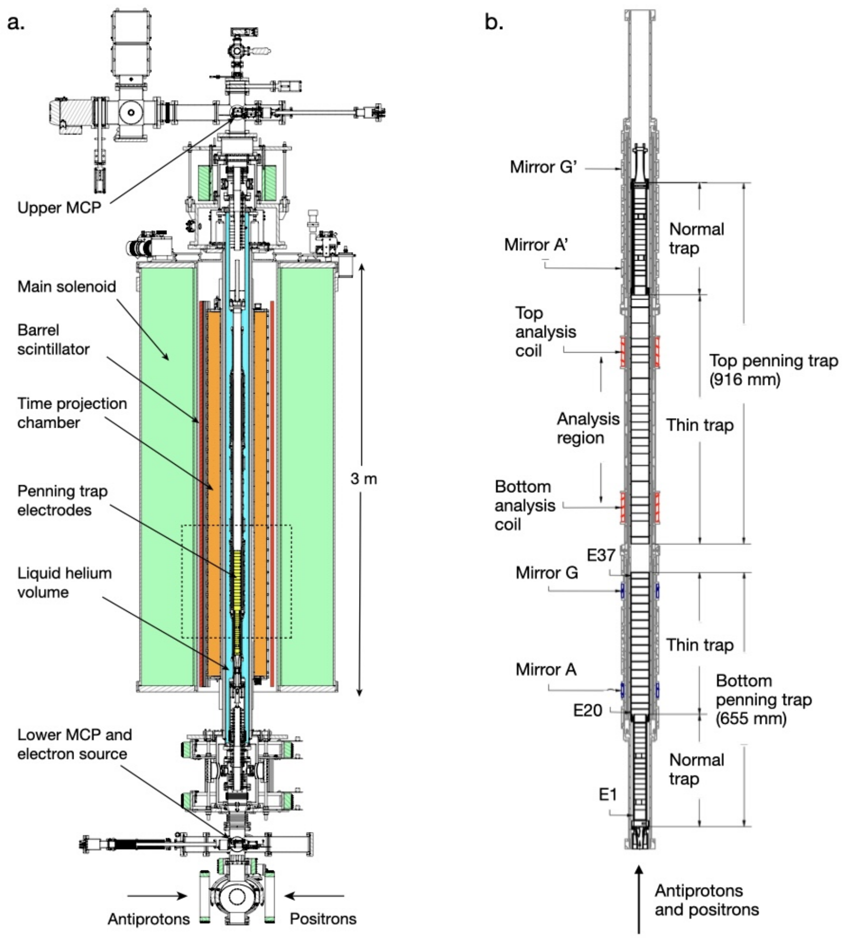

Figure 3.9: (a.) A schematic overview of the full ALPHA-g apparatus, adapted from [1]. The dotted box highlights the bottom Penning trap. (b.) A design schematic for the ALPHA-g Penning-Malmberg and Ioffe-Pritchard traps, adapted from an upcoming publication. This reflects a recent decision to include a Penning-Malmberg trap in the precision region (here, analysis region), due to its utility for magnetic field measurements.

no Penning-Malmberg electrodes).

Recent developments diverge from this initial plan. The complete symmetry between the top and bottom traps has been forgone, since new trap production techniques have made installing a Penning-Malmberg trap in the precision region possible by extending the top trap, as shown in Figure 3.9.b. The magnetic field characterization techniques used in the initial measurement, which will be discussed in Section 9.1.2, rely on electron plasmas trapped within the Penning-Malmberg trap. The additional electrodes allow the extension of these techniques to the precision region. The implications of this modification on the final precision measurement have not been fully explored.

In the first published ALPHA-g measurement of gravity [1], only the bottom trap was commissioned and used. Antiprotons and positrons were combined in this region, and antihydrogen was magnetically trapped there before being released. In future precision measurements, antihydrogen atoms will be formed in either the top or bottom trap, before being transferred to the central precision trap and released.

The electrodes of the bottom and top traps are further subdivided into two sections: the normal trap and the thin trap. The normal trap has an internal diameter 29.6 mm, and is responsible for re-catching the antiprotons and positrons sent into ALPHA-g. The thin trap has electrodes of only 1.5 mm thickness, allowing for an increased internal diameter of 44.35 mm. Whereas charged particle plasmas can easily be confined to a small radius, this is not the case for antihydrogen atoms in the magnetic minimum trap. Increasing the radius of the trapping area allows for higher energy antihydrogen atoms to be trapped.

The ALPHA-g magnet system involves a large number of superconducting magnet coils, which are shown in Figures 3.9 and 3.10. They are enclosed by a liquid helium cryostat, which cools the magnets and also acts as a cryopump. This is supplemented by scroll pumps, turbo pumps, and ion pumps in a scheme similar to the one used in the catching trap.

A 1 T superconducting solenoid magnet provides the axial magnetic field of the Penning-Malmberg trap. This magnet is 3 m long with internal diameter 50 cm, enclosing the full ALPHA-g apparatus. This magnet uses its own closed system liquid helium cryostat. In addition to the 1 T solenoid, two additional *capture solenoids* surround the normal section of the top and bottom trap. These increase the axial magnetic field from 1 T to 3 T, creating a stronger trap for the antiprotons and positrons arriving from the interconnect. The magnetic minimum trap is formed by a set of octupoles and a number of solenoidal mirror coils. The *long octupole* spans the length

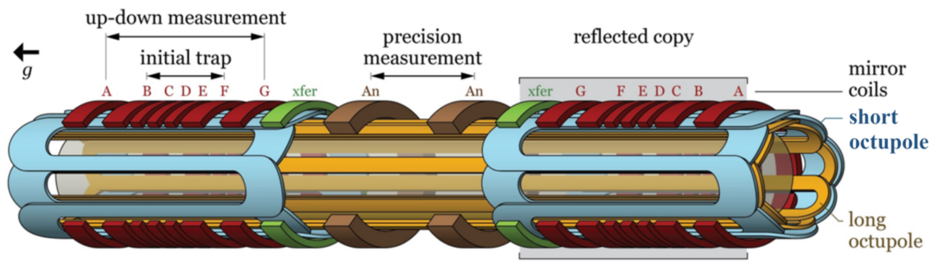

Figure 3.10: The layout of the ALPHA-g magnet system, rotated so that gravity acts in the leftwards direction. The initial measurement was performed between mirror coils A and G, in the region labelled "up-down measurement". Image credit: Dr. Chukman So.

of ALPHA-g, and creates a radial magnetic field minimum. This field is increased by two *short octupole* magnets covering the top and bottom trap. They provide radial confinement to antihydrogen atoms with kinetic energies corresponding to temperatures of up to 0.5 K. Axial confinement in the bottom trap is provided by seven *mirror coils*, labelled A to G. In the initial measurement, only the outermost coils A and G were used to confine the antihydrogen atoms. These were also the coils which were used to test the effect of gravity.

There are two further mirror coils known as the *analysis coils*, which create a trapping region in the precision trap, and will be used in the upcoming precision measurement. The depth of the precision trap is smaller than the two end traps due to the smaller octupole field. This necessitates cooling the antihydrogen from 0.5 K to even lower energies, using techniques which will be discussed in the next section. Two *transfer coils* are used to increase the magnetic field between the top, bottom and precision regions, without which the field would be too small to facilitate the transfer of antihydrogen atoms without significant losses.

The details of how this apparatus is used to measure the effect of gravity are deferred to Chapter 9. This measurement was performed only in the bottom trap, as the top and precision traps are still undergoing installation and commissioning. A simplified version of the procedure was used for the 2023 gravity paper: antihydrogen was formed in the bottom trap, in the region labelled "up-down measurement" in Figure 3.10. Mirror coils A and G were then ramped down so that the antihydrogen could escape axially, and the antiatoms escaping upwards and downwards were counted. The effect of gravity was then extracted from the ratio of the two counts. Future measurements following the installation and commissioning of the full trap will be performed in the precision region; this will be discussed in Chapter 9.

3.5 Particle preparation techniques

The techniques used by ALPHA to manipulate the antiprotons, positrons, and antihydrogen atoms within the apparatus have been developed over many years of experimentation. This section will briefly outline some of the techniques used in the antihydrogen gravity experiment.

3.5.1 Antiproton preparation

Electron cooling

One of the features of a Penning-Malmberg trap is that it has an inherent cooling mechanism called cyclotron cooling. As a trapped particle undergoes circular cyclotron motion as described in Section 3.2.1, it will continuously emit photons via a process called cyclotron radiation. Each photon emitted decreases the kinetic energy E of the trapped particle, at a cooling rate given by:

$$\frac{dE}{dt} = -\frac{q^4 B_0^2}{3\pi\epsilon_0 m^3 c^3} E = -\frac{1}{\tau_c} E \tag{3.5.1}$$

We define τ_c as the cyclotron cooling time; it is a combination of the trapped particle charge q, mass m, and the applied magnetic field B_0. Here, ϵ_0 is the permittivity of free space, and c is the speed of light. This equation is derived from the Larmor formula [108], which gives the power radiated by a non-relativistic accelerating point charge. We make the approximations that the cooling due to the cyclotron motion dominates that due to the magnetron motion, and that all the energy of the charged particle is kinetic. Note that the cooling rate is proportional to the energy, and so the solution is given by $E(t) = E_0 e^{-t/\tau_c}$.

In a 3 T field such as the catching trap or ALPHA-g bottom trap, the antiproton cooling time is $\tau_c = 2 \times 10^9$ s, which is unfortunately far too long to be useful. By comparison, the same calculation for an electron yields $\tau_c = 0.3$ s; electrons held in Penning-Malmberg trap will naturally cool to their minimum energy in seconds. This minimum energy is reached when cyclotron cooling is in a balanced equilibrium with heating from thermal radiation of ohmic electrode heating, or from magnetic field inhomogeneities.

In ALPHA, we take advantage of electron cyclotron cooling to sympathetically cool antiprotons. Before receiving antiprotons from ELENA, a small electron plasma is loaded from the electron gun on the catching trap stick into a potential well in the catching trap, where it undergoes cyclotron cooling. Antiprotons from ELENA are then caught in an overlapping potential well, and exchange energy with the electrons via Coulomb interactions. Over a few hundred milliseconds, the antiproton temperature is lowered to match that of the electron plasma, forming a mixed plasma with a single temperature confined in a potential well of 50 V to 100 V. A similar procedure is used to catch and cool antiprotons sent from the catching trap into ALPHA-g.

Evaporative cooling

One can generally assume that the kinetic energies of charged particles in a plasma follow the thermal Maxwell-Boltzmann distribution. The majority of particles have an energy near the mean, but a small fraction form a higher-energy tail. Partially lowering the electrode potentials providing axial confinement allows the more energetic particles to escape, truncating the high energy tail of the Maxwell-Boltzmann distribution. The remaining particles undergo Coulomb collisions, redistributing energy and restoring thermal equilibrium. The kinetic energy will re-equilibrate to a Maxwell-Boltzmann distribution with a lower average temperature such that the high-energy tail is below the escape energy, as shown in Figure 3.11. This process is known as evaporative cooling (EVC).

In reality, the situation is more complicated than this simple picture, and a full understanding of this process must be achieved by modelling the trajectories of individual charged particles in the trap. This process has been used on neutral atoms to great success, for example leading to the creation of the first Bose-Einstein condensate [109]. However, its previous application to charged particle plasmas has been limited, and was substantially developed by the ALPHA collaboration [110].

Rotating wall electrodes

One commonly-used method to control the size and density of a charged particle plasma is the rotating wall technique [111]. In this method, one electrode in a Penning-Malmberg trap is segmented azimuthally, into six segments in the ALPHA traps. Independent sinusoidal voltages are applied to each segment, with frequencies of order 100 kHz, and a phase offset of

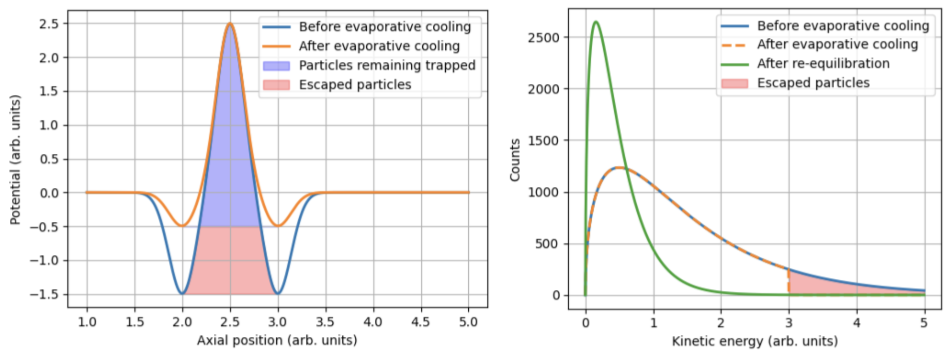

Figure 3.11: Left: Sketch of the axial potential before and after EVC. Right: Sketch of the kinetic energy distribution before EVC, after EVC, and after re-equilibrating to a thermal distribution. The escape threshold energy is determined by the depth of the potential well after EVC.

$2\pi/n$ between adjacent segments as shown in Figure 3.12.

The electric field of the rotating wall applies a torque to the plasma, which changes the total angular momentum of the plasma which would otherwise remain constant as discussed in Section 3.2.2. This also adjusts the radius of the plasma and the rotation frequency. One particular method of applying the rotating technique is known as the strong drive regime (SDR) [112]. Here, the applied rotating wall frequency is close to the initial rotational frequency of the plasma, and the torque applied forces the two frequencies to match. The plasma radius then responds linearly to adjustments in the applied frequency, allowing for precise control, independent of the magnitude and duration of the applied rotating wall voltage. Furthermore, by controlling the plasma radius without changing the plasma length or number of particles, this technique also directly controls the plasma density.

SDREVC

A development made at ALPHA was to combine evaporative cooling and the strong drive regime rotating wall technique. This is known as strong drive regime evaporative cooling (SDREVC) [113]. The strong drive regime rotating wall controls the plasma radius and density, while evaporative cooling selectively reduces the particle number by removing the higher energy particles. The combination of the two techniques results in a plasma with a size, density, and number of particles consistent to within 1% across many trials with different initial parameters.

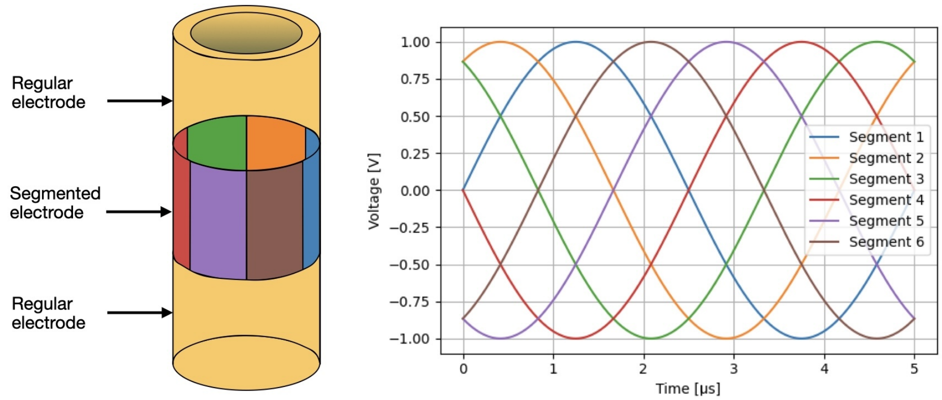

Figure 3.12: Left: A diagram showing segmented electrodes used for rotating wall in the catching trap and ALPHA-g. Right: The voltage applied to each segment during a typical rotating wall procedure.

SDREVC is featured extensively throughout the ALPHA antihydrogen production cycle. It is first used to create reproducible electron plasmas for the electron cooling of antiprotons, both in the catching trap, and in ALPHA-g. After receiving antiprotons, it is used to further cool and shape the mixed electron-antiproton plasma. It is also employed to create a reproducible positron plasma in ALPHA-g, as well as on Be^+ ions, as will be discussed later in this section.

Electron kick-out

After the capture of antiprotons and subsequent SDREVC of the mixed antiproton-electron plasma, both in the catching trap and in ALPHA-g, it is necessary to remove the electrons to obtain a pure antiproton plasma. This is done by exploiting the large mass difference between the two species.

The velocity of a particle with mass m and kinetic energy $E_k = \frac{1}{2}k_B T$, where T is the particle temperature and k_B is the Boltzmann constant, is given by:

$$v = \sqrt{\frac{2E_k}{m}} = \sqrt{\frac{k_b T}{m}} \qquad (3.5.2)$$

Once the antiprotons and electrons are in thermal equilibrium and have the same temperature, the electrons travel $\sqrt{m_p/m_e} = 43$ times faster than the antiprotons.

To make use of this difference, the axial confinement potential is briefly lowered and then rapidly restored. The fast-moving electrons escape during this short time window, while the slower antiprotons remain trapped. This technique is known as an electron kick-out, or e-kick.

While highly effective at removing electrons, e-kicks can cause heating of the remaining antiprotons due to abrupt changes in potential. For this reason, the procedure is staged. After an initial period of SDREVC, a partial e-kick removes most electrons but leaves a small population. These remaining electrons facilitate further SDREVC and re-cool the plasma. Subsequent e-kicks then remove the remaining electrons.

Temperature diagnostics

To determine the temperature of a given antiproton plasma, a destructive measurement is performed. This measurement uses the assumption that the velocity of the antiprotons is thermal, following a Maxwell-Boltzmann distribution with a characteristic plasma temperature. The axial (one-dimensional) kinetic energy E for particles with a Maxwell-Boltzmann velocity distribution follows a χ^2 distribution with one degree of freedom, where T is the temperature and k_B is the Boltzmann constant:

$$f(E) = \frac{1}{\sqrt{\pi E \; k_B T}} \cdot \exp\left(-\frac{E}{k_B T}\right) \tag{3.5.3}$$

To measure the plasma temperature, the confining potential on the electrodes is lowered slowly so that all the antiprotons are able to escape. They are directed towards a diagnostic stick inserted into the beam path, where they annihilate and are detected by the surrounding scintillating panels (see Section 3.4.5). Because the potential barrier is lowered gradually, the time at which each antiproton escapes corresponds to a known potential energy. In this way, the energy distribution of the plasma is reconstructed.

Instead of fitting the full energy distribution with Equation 3.5.3, only the high-energy tail is fit. Here, the $\sqrt{E}$ dependence can be ignored and the distribution can be approximated as an exponential. In practice, the energy distribution is plotted with the counts on a log scale, so that a linear fit can be used on the tail of the distribution, as shown in Figure 3.13.

3.5.2 Positron preparation

Preparation of the positron plasma uses many of the same techniques used for antiprotons. Since the number of positrons transferred from the accumu-

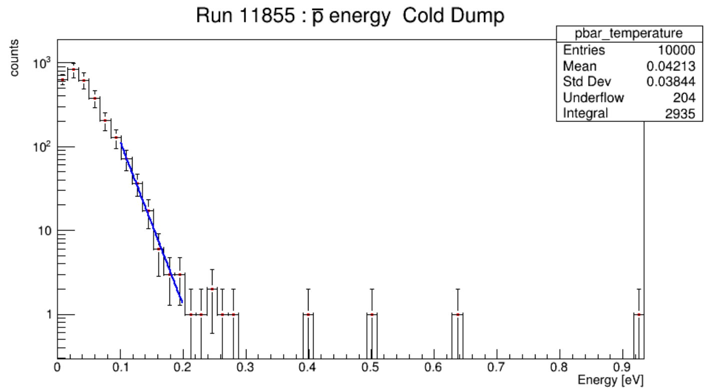

Figure 3.13: Temperature diagnostic of an antiproton plasma. The blue line is a fit to the data from the range of 0.1 eV to 0.2 eV, and corresponds to a temperature of 259 K.

lator can vary significantly between cycles, SDREVC is crucial to forming a reproducible positron plasma. A similar temperature diagnostic is used, with the positron signal measured on a Faraday cup instead of measuring annihilations on a scintillator panel.

Since the depth of the neutral atom trap is only 0.5 T, only the antihydrogen atoms in the low-energy tail of the energy distribution are trapped, making it imperative to obtain the coldest possible sample of antihydrogen. The antihydrogen temperature depends strongly on the temperature of both the antiproton and the positron. As such, it is desirable to cool the positrons beyond cyclotron cooling and evaporative cooling; this led to the recent development of beryllium ion (Be^+) assisted cooling.

This method begins with the creation of a Be^+ plasma using a beryllium ablation source mounted on a diagnostic stick. A pulsed 355 nm laser ablates a beryllium target, generating a plasma of Be^+ ions in the ALPHA-g trap. A 313 nm laser is then used to Doppler-cool this plasma; the basic principle of laser cooling is described in Section 3.5.4. While laser cooling continues, SDREVC is applied to fix the Be^+ plasma size and density.

Simultaneously, a positron plasma is received from the accumulator, and stabilized using SDREVC. The positron and Be^+ plasmas are then merged, forming a mixed-species plasma, The positrons are cooled sympathetically via Coulomb interactions with the Be^+ ions. Laser cooling continues during this phase to maintain the low Be^+ temperature. The choice to use Be^+ ions was made because they are light enough for effective sympathetic cooling with positrons, easy to produce via ablation, and relatively easy to laser cool. The Be^+ assisted positron preparation will be described in detail in an upcoming publication.

3.5.3 Antihydrogen formation

The approach used by ALPHA to create antihydrogen was first demonstrated by ATHENA in 2002 [24], whereby cold plasmas of antiprotons and positrons are slowly mixed. The dominant production mechanism is called three-body recombination:

$$\bar{p} + e^+ + e^+ \longrightarrow \bar{H} + e^+ \tag{3.5.4}$$

Two positrons interact electromagnetically, leaving one positron with low enough energy to be captured by an antiproton and form a bound state. The excess energy and momentum is carried away by the second positron, essential for conservation of energy. Further photons are then produced as the antihydrogen atom radiatively cascades into the ground state.

Application of this mechanism involves spatially overlapping the antiproton and positron plasmas in the Penning-Malmberg trap. This is not trivial since positrons and antiprotons have opposite charges, and so a single electrostatic potential cannot confine both simultaneously; potential wells that confine antiprotons will repel positrons and vice versa. A slow merge process illustrated in Figure 3.14 is used. The antiproton and positron plasmas are initially confined in adjacent potential wells. The electrostatic barrier between them is then gradually lowered, allowing the two species to escape from their trapping potential well and pass through the plasma of the other species. This process takes around 1 s, determined empirically so that the potential manipulations do not excessively heat the antiprotons and positrons.

In a typical mixing cycle, almost all antiprotons are converted into antihydrogen, yielding typically around 150 000 antihydrogen atoms. However, most of these have kinetic energy above the 0.5 K Ioffe-Pritchard trap depth, and immediately escape and annihilate. Only the coldest atoms are retained, typically resulting in around 20 trapped antihydrogen atoms per mixing cycle in ALPHA-g. Because neutral antihydrogen atoms are unaffected by the electrostatic fields of the Penning-Malmberg trap, the mixing process can be repeated multiple times while the magnetic minimum trap remains active, accumulating trapped antihydrogen from successive cycles. Following each mixing cycle, a large electrostatic potential gradient is applied using the Penning-Malmberg electrodes, in order to clear any remaining antiprotons trapped by magnetic mirroring.

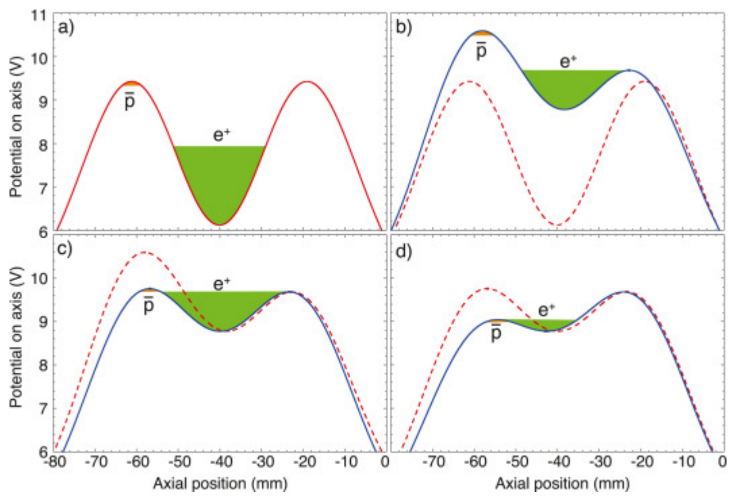

Figure 3.14: Schematic diagram showing antiproton and positron plasmas confined in the ALPHA-2 or ALPHA-g Penning-Malmberg trap in preparation for (a and b), immediately before (c), and during (d) the antihydrogen formation process. From [114].

3.5.4 Antihydrogen cooling

For almost any precision measurement involving antihydrogen, it is beneficial for the sample of atoms to be as cold as possible. This is particularly true for ALPHA's laser spectroscopy experiments, where the sample temperature contributes directly to the spectral linewidth via Doppler broadening, which we will discuss shortly. However, the temperature also plays a role in a measurement of gravity. The lower the antiatom temperature, the smaller the thermal velocity of the atom, and so the more precisely its motion due to gravity can be isolated and measured.

To this end, two antihydrogen cooling mechanisms have been developed. Neither was employed in the initial published measurement of gravity [1], but they will both play a key role in upcoming precision measurements. Both techniques have been successfully demonstrated in the ALPHA-2 apparatus, and their implementation in ALPHA-g is ongoing.

Laser cooling

The first cooling method is known as laser cooling. This is a well-known technique, for which Chu, Cohen-Tannoudji, and Phillips won the 1997 Nobel prize in physics [115]. In an antihydrogen atom, the positron occupies orbits with discrete energy levels. A transition from a higher energy orbital (e.g. 2P) to a lower one (e.g. 1S) releases a photon which carries away the excess energy. Conversely, the absorption of a photon with energy matching the level separation can promote the positron to an excited state. This principle was first demonstrated on the antihydrogen 1S–2P transition by ALPHA in 2018 [9].

In practice, transitions can be excited by photons with a narrow range of energies around the resonance, due to several line-broadening mechanisms. One of the most important is Doppler broadening. If an antihydrogen atom moves with velocity v towards an incident laser photon of frequency f, then in the frame of reference of the atom, the incident photon will have a Doppler shifted frequency f' given by:

$$f' = f\left(1 + \frac{v}{c}\right) \tag{3.5.5}$$

Thus photons with slightly less energy than the atomic resonance can be absorbed by antiatoms moving towards the laser source, and photons with slightly more energy can be absorbed by antiatoms moving away from the laser.

Laser cooling takes advantage of this phenomenon by purposefully detuning the laser to have a lower frequency than the transition frequency, causing antiatoms moving towards the laser source to be preferentially excited. Since photons carry momentum $p = hf$, each photon absorbed gives a momentum kick away from the laser source. In this way, antiatoms travelling towards the laser are preferentially kicked backwards (i.e. opposite to their motion), decreasing the average antiatom momentum.

The energy of the absorbed photon is later re-emitted as a new photon, as the antiatom relaxes to the ground state. However, this emission is isotropic, and so has no effect on the average particle momentum. Over many interactions, this photon absorption and subsequent isotropic emission decreases the overall temperature of a sample of antihydrogen atoms.

This technique was first demonstrated on antihydrogen in ALPHA-2 in 2021 [116]. A pulsed 121.6 nm Lyman-α laser is applied to a sample of trapped antiatoms, detuned by -240 MHz from the 1S-2P resonance, typically over a time period of 2 h to 4 h. The transverse kinetic energy of the antihydrogen atoms can be inferred from the time-of-flight distribution of atoms kicked out of the trap and detected via annihilation; this kickout is performed by driving a transition to an untrappable high-field-seeking state (see Section 3.2.3). Using this technique, the median transverse energy of the antihydrogen atoms was reduced from 15.1 µeV to 1.3 µeV. It is difficult to estimate the change in total kinetic energy as the mixing between transverse and longitudinal energies in the trap is not yet well understood. However, the implementation of laser cooling has already significantly improved recent spectroscopic measurements made in ALPHA-2. Work to implement laser cooling in ALPHA-g is currently underway.

Adiabatic cooling

There is a fundamental limit to the temperature attainable through laser cooling, since the average momentum of an atom cannot fall below the momentum imparted by one photon.[5] To reach even lower temperatures, an alternative method is required.

One such method is adiabatic cooling, where the width of the confining magnetic potential well is slowly expanded, thereby decreasing the temperature of the trapped antihydrogen atom. To understand this effect, consider a harmonic approximation of the axial trap potential $V(z) = \frac{1}{2}m\omega_z^2 z^2$. As the trapping potential is adiabatically (slowly compared to the axial bounce frequency ω_z) expanded, the axial bounce frequency decreases, since the antiatom has a longer distance to travel per oscillation.

This is familiar in a quantum mechanical context, where states $|\psi_n\rangle$ have quantized energies $E_n = (n + \frac{1}{2})\hbar\omega_z$. By decreasing ω_z adiabatically, each particle remains in the n'th state of the perturbed system $|\tilde{\psi}_n\rangle$. Since ω_z has decreased, this state now has a lower energy, corresponding to a decrease in the temperature of the plasma.

In classical mechanics, this concept is encapsulated by adiabatic invariants. These are quantities which remain constant as the Hamiltonian of the system is slowly changed, related to the phase space area of a system. In the case of a harmonic potential, the adiabatic invariant is the ratio of energy to frequency, E/ω [118, 119]. This adiabatic invariant makes the principle behind adiabatic cooling explicit. By slowly expanding the width of the trapping potential, and thereby decreasing ω_z, the energy of each antihydrogen atom is decreased to keep the value of E/ω_z constant.

In reality, the full trap is three-dimensional and non-harmonic, and the motion of the antiatoms is more complex. We rely on simulations to model the cooling effect of adiabatically expanding the magnetic trap, the details of which can be found in [120].

Prospects for adiabatic expansion were a key consideration in the design of ALPHA-g. In the envisioned precision measurement, antihydrogen will be

[5] There are techniques such as Raman cooling [117] which can cool atoms below this recoil limit. However, these require laser systems not currently available at ALPHA.

initially formed between two close mirror coils, such as coils B and D (refer to Figure 3.9). After many mixing cycles, the trap will be adiabatically expanded to between coils A and G. This technique was not employed in the initial gravity measurement, but it has since been demonstrated in the ALPHA-2 apparatus [121]. Adiabatic cooling will be essential for performing upcoming precision measurements of gravity in ALPHA-g.

Chapter 4

The ALPHA-g detectors

Measuring the effect of gravity on antihydrogen requires determining if each antiatom exits the magnetic trap in the upwards or downwards direction; in the absence of other forces, more antiatoms should escape the trap in the direction of gravity. The simplest method for this is to allow the antihydrogen to annihilate on the trap walls, and determine the annihilation positions by detecting the resulting annihilation products. This is the approach taken by the ALPHA-g annihilation detectors.

Detecting the annihilation of a positron is a well-motivated and well-studied problem, as it forms the basis of a common medical imaging technique known as positron emission tomography (PET). A radioactive nuclide is bonded to a drug or biological compound, and injected into the patient. The compound then accumulates in a target area, such as a cancer tumour. There, the radionuclide undergoes beta decay, producing positrons. These positrons quickly annihilate with electrons, producing primarily two photons emitted back-to-back, as discussed in Section 3.3.1. By detecting both photons, the annihilation can be localized along a line between the two detectors. With enough such events, regions of concentrated activity will emerge, indicating the position of a cancer tumour.

A positron annihilation detector using this principle was employed by the ATHENA experiment to verify the first production of cold antihydrogen at CERN [122]. This was used alongside a silicon strip detector which detected the charged pions resulting from the antiproton annihilation. The coincidence of charged pion tracks in the silicon strip detector, together with back-to-back hits in the positron annihilation detector, confirmed the annihilation of an antihydrogen atom.

With the routine production and trapping of antihydrogen atoms fully established, the upgrade to the ALPHA-2 apparatus moved away from the detection of the positron annihilations. Instead, a *silicon vertex detector* (SVD) was installed, capable of precisely reconstructing the annihilation position, or vertex. This detector is shown in Figure 4.1 alongside the other components of the ALPHA-2 trap.

The SVD is composed of individual rectangular silicon microstrip mod-

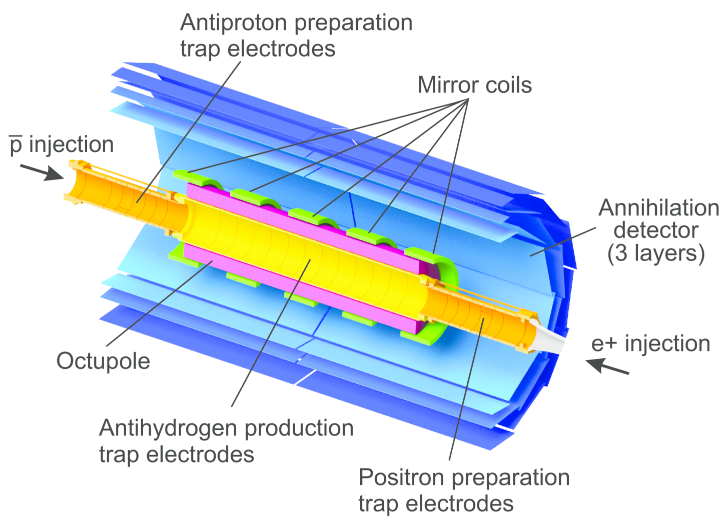

Figure 4.1: The ALPHA-2 apparatus, with the three layers of the silicon vertex detector (SVD) shown in blue. Adapted from [4].

ules, each with a sensitive volume of 5.8 cm × 11.2 cm × 44.8 cm [102]. These modules are arranged in three layers around the trap, with 10, 12, and 14 modules per respective layer [123]. When an antiproton annihilates, it typically produces several charged pions, as discussed in Section 3.3.2. These travel through the silicon layers, leaving hits that can be used to reconstruct the pion tracks in three dimensions. The annihilation vertex is then identified by determining the point of closest approach between all reconstructed tracks.

The ALPHA-2 SVD has been crucial to the experiment's precision spectroscopy program, providing reliable annihilation imaging with high spatial resolution of around 8.5 mm [123]. A full discussion of silicon tracking detectors and the ALPHA-2 SVD is outside the scope of this dissertation; this summary is intended to provide context for the ALPHA-g detectors. Although the detector technologies differ, the principle of reconstructing charged pion tracks to determine antihydrogen annihilation positions remains the same and forms the foundation for the detection strategy in ALPHA-g. This approach has been used for many years in ALPHA-2, and is accepted as the most straightforward and reliable method of antihydrogen detection.

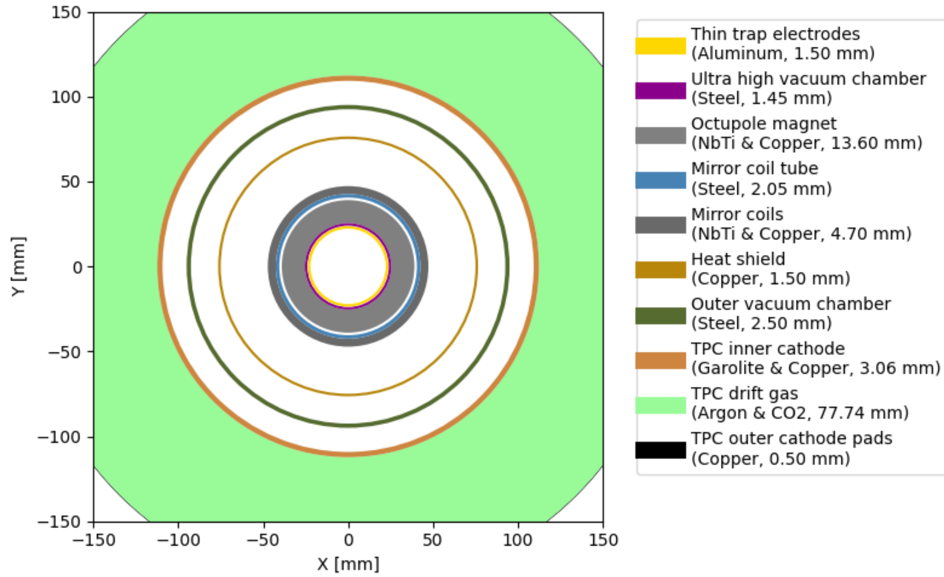

Figure 4.2: Diagram of the main components of the ALPHA-g apparatus inside the time projection chamber detector. Components are listed from innermost to outermost with their respective materials and thicknesses.

4.1 Pion scattering on internal materials

There is a considerable amount of material between the detectors and the trap wall, including vacuum and cryogenics components, as well as the superconducting octupole and mirror coil magnets. A radial cross-section of the apparatus is shown in Figure 4.2, with the materials and thicknesses listed. Before introducing the ALPHA-g detectors, we consider here the effects of this material on an annihilation detection scheme which reconstructs the annihilation position based on charged pion tracks.

While traversing the internal materials, pions lose energy through multiple elastic scatterings on the atomic nuclei. This is a stochastic process, which may lead to an overall large angle deflection of the pion trajectory. This primarily occurs in the octupole magnet and mirror solenoid coils, as they have are the most dense components and have the largest thickness.

In the case of a large angle pion deflection in the material, the ALPHA-g detectors will then observe an altered trajectory which does not point back to the annihilation position. The detector must be able to accurately determine the annihilation position in the presence of highly scattered pions.

The dense material around the trap also causes absorption and scattering of the 511 keV photons from positron annihilations. Using attenuation coefficients from the NIST database [124], we can estimate an attenuation length of around 22 mm in the NbTi/copper magnet coils. This results in around half of the photons from positron annihilation being attenuated in the octupole and mirror coils. This is one of the reasons why detection of positron annihilations was not included in the design of the ALPHA apparatus.

4.2 The time projection chamber

To maximize the gravitational potential difference in the area accessible to the antiatoms, the vertical ALPHA-g apparatus (see Figure 3.9) is over 2 m long, significantly larger than the ALPHA-2 apparatus. This increased size meant that constructing a detector similar to the ALPHA-2 SVD was neither practical or cost-effective. Instead, a different type of charged particle detector called a *time projection chamber* (TPC) was developed, capable of reconstructing the annihilation positions over the larger volume. This section will begin with a broad discussion of TPC technology and ionization drift chambers, before covering specific details of the ALPHA-g TPC. The readout and data acquisition will be covered later in Section 4.4.1.

4.2.1 Ionization chambers

The ionization chamber is one of the oldest concepts for a charged particle detector [125], and many variants have since been developed to suit different applications. The core principle is to use an active medium with a low ionization potential, typically a gas, and then apply an electric field. When a charged particle traverses the chamber, it ionizes the gas atoms along its path, creating many pairs of electrons and positive ions. The applied electric field prevents recombination by pulling the electrons towards the anode and the ions towards the cathode. In some cases (around 34% in argon [126]), the released electron will ionize another nearby atom, creating an ionization cluster. More rarely, it will have enough energy to ionize many atoms, creating an ionization trail known as a delta ray.

One common configuration of ionization chamber is the wire chamber, where the anode is a wire with a large positive voltage applied. The electric field of a cylindrical wire held at a constant voltage is a textbook calculation:

$$E(r) = \frac{\lambda}{2\pi\epsilon_0 \, r} \tag{4.2.1}$$

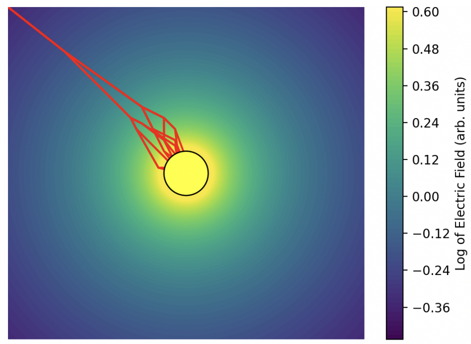

Figure 4.3: A cross-sectional drawing showing the development of an electron avalanche (red) near an anode wire (yellow) in a wire chamber.

Here λ is the charge density of the wire, dependent on the applied voltage. ϵ_0 is the permittivity of free space, and r is the distance from the centre of the wire. As the electrons approach the anode wire, they experience an increasingly strong electric field. The electric field near the wire accelerates the electrons to high enough energies to cause additional ionization. This begins an avalanche, where these newly-freed electrons will ionize subsequent atoms, leading to an exponential multiplication of electrons, shown in Figure 4.3. The electrons are collected on the anode wire, and the ions drifting towards the cathode induce a further current on the anode wire. This current is amplified and read out.

In addition to ionization, incident particles also deposit energy via excitation of the gas atoms, where an electron is raised to a higher atomic energy level. These excitations decay to the ground state by emitting a photon. If the active medium is transparent to the photon wavelength, this can allow an additional readout channel via photon detection. However, these photons can also produce subsequent electrons via the photoelectric effect, and produce after-pulsing or even a continual activation of the wire chamber. To suppress this, a quenching gas is often mixed with the ionization gas, which absorbs excess photons without re-emitting ionizing radiation.

A *multiwire proportional chamber* (MWPC) is a type of ionization chamber which uses multiple anode wires; the wire where a signal is observed gives some spatial knowledge of the ionization position [127]. The signal size is proportional to the energy deposited by the incident charged particle. This design was iterated upon to create the *drift chamber*, where further spatial information can be obtained from the drift time of the electrons and ions.

The ALPHA-g experiment uses a specific variant of a drift chamber known as a *time projection chamber* (TPC). A traditional TPC design has a large drift volume with a constant electric field, which drifts electrons towards multiple anode wires. The specific wire where an avalanche is detected gives one coordinate of the ionization position. The time of the signal then gives another coordinate; since the drift velocity is constant in a constant electric field, the delay is proportional to the distance the electrons travelled. The third coordinate of the ionization position must be obtained through other means: either by using two planes of anode wires at an angle to each other, or by using a segmented cathode plane with readout of the induced charge. Combining these three coordinates allows full reconstruction of the position of each gas ionization, making the TPC an effective tracking detector which has seen widespread use [128, 129].

4.2.2 The ALPHA-g TPC

A conventional TPC uses anode wires at the bottom and/or top of the chamber, resulting in vertical electron drift. However, this layout was impractical for ALPHA-g due to geometrical constraints. To obtain the best annihilation position resolution, the distance between the annihilation and pion detection must be minimized. As such, the TPC must fit in the narrow annular region between the external solenoid and the cryostat, a volume less than 15 cm in radial width. It must also cover the full trap apparatus, over 2 m in length. This aspect ratio made the traditional TPC layout unviable, as electrons would diffuse away from their initial $x - y$ positions over the long vertical drift length, with many being lost on the walls of the TPC.

The ALPHA-g TPC uses a less common geometry shown in Figure 4.4, with a ring of 256 anode wires running along the length of the outer edge of the TPC. A second ring of 256 *field shaping* wires is positioned slightly inside the anode wires. The inner wall of the chamber forms the cathode. Electrons are pulled radially outwards, and ions inwards. In this design, the applied electric field is no longer constant at varying radius. Furthermore, there is a significant deflection from a purely radial drift due to the Lorentz force from the external magnetic field; this deflection has no component in the vertical direction as long as the external magnetic field is purely solenoidal.

In typical operation, the cathode surface at radius 10.9 cm receives -4 kV, the field shaping wires at radius 17.4 cm receive -0.1 kV, and the anode wires at radius 18.2 cm receive 3.2 kV. The field shaping wires ensure that the region over which the electrons and ions induce a charge on the anode wires is roughly constant, regardless of the initial ionization position.

The outer wall of the TPC, at radius 19.0 cm, is finely segmented into

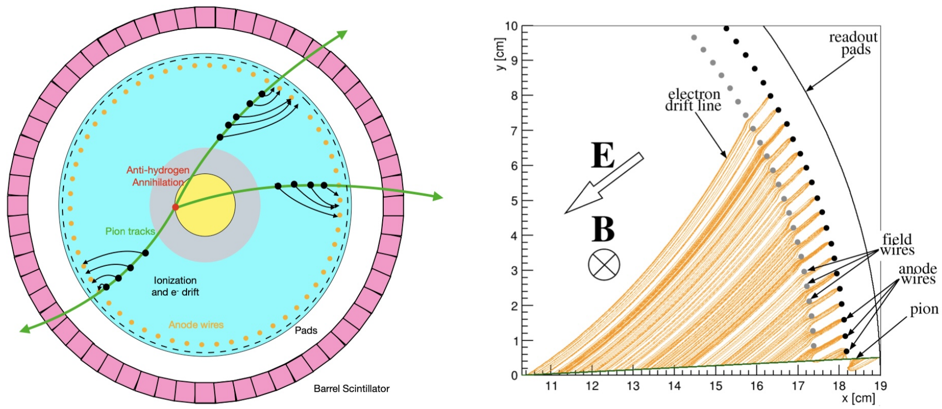

Figure 4.4: Left: A top-down schematic diagram of an antihydrogen annihilation in the ALPHA-g TPC. Shown are the gas region (blue), the anode wires (yellow), and the cathode pads (dashed line), along with electrons (black arrows) drifting from ionization sites (black circles) along pion tracks (green) from an annihilation (red). Right: A Garfield++ simulation (see Section 5.3.1) showing electrons drifting towards the anode wires. From [130].

cathode pads. It is azimuthally divided into 32 sections each with width of 3.7 cm, and these sections are divided into 576 pads with a pitch of 0.4 cm, for a total of 18432 cathode pads. When an electron avalanche occurs near an anode wire, this also induces a charge on the nearby cathode pads. Both the anode wires and the cathode pads are read out, and by matching each anode wire signal to the corresponding cathode pad signals, the location of the avalanche can be determined. Once an avalanche has been located, the three-dimensional position of the corresponding gas ionization is determined using the time projection method.

The gas used in the ALPHA-g TPC is a mixture of 70% argon and 30% carbon dioxide at atmospheric pressure. The argon is an inert ionization gas, and the carbon dioxide is a quenching gas; the ratio was decided empirically for good signal gain and stable operation. The gas mixture is flowed continually through the detector at a rate of 1 L/min to ensure a constant pressure.

The Lorentz deflection due to the external magnetic field causes electron displacement up to 30° from purely radial drift, as shown in Figure 4.4. To measure this effect, a laser calibration mechanism is included in the TPC design. Aluminium strips are located on the inner cathode surface at nine axial positions along the TPC. A pulsed laser is directed towards each of these strips in turn, producing free electrons at known positions and times via the photoelectric effect. These laser calibration measurements were used to validate electron drift simulations in the PhD work of Dr. Pooja Woosaree [77], and are periodically repeated to monitor the performance of the TPC.

4.3 The barrel scintillator

Cosmic rays have been a known source of detector background for the entirety of the ALPHA physics program. These rays, primarily high-energy protons originating from outer space, interact with Earth's atmosphere to produce particle showers [131]. These cascades typically include pions decaying to muons, which are the only charged particle type with a low enough interaction cross-section to reach the Earth's surface. Many of these muons interact with the ALPHA detectors as relativistic minimum ionizing particles. Their detector signatures are typically single, straight tracks, with rare showering interactions or secondary production. These events are identified and discarded via a machine learning analysis which will be discussed in Chapter 8.

The ALPHA-g TPC has a larger volume than the ALPHA-2 SVD, and consequently presents a bigger target for cosmic rays and thus has a higher background rate. To mitigate this, the ALPHA-g detector suite includes a dedicated cosmic ray veto detector, called the *barrel scintillator* (BSC). The detector is positioned in the region outside the TPC and inside the external solenoid. It is designed to measure particle time-of-flight (TOF), to distinguish the direction of charged particles, and identify and reject those originating outside the detector. This section will first review the scintillator and photomultiplier technologies used, before describing the details of the ALPHA-g BSC detector and the measurement principles.

4.3.1 Scintillators

Whereas atomic excitations are an unwanted energy loss mechanism in gaseous ionization detectors, they constitute the working principle of scintillation detectors. A scintillator is a material whose electrons are easily excited by the passage of an ionizing particle; when the electrons return to their ground state they emit photons whose wavelength is fixed by the transition energy. The number of photons produced provides a convenient proxy for the total energy loss in the scintillator. There are a wide variety of scintillating materials used in practical applications, each best-suited to different experimental conditions:

- Liquid noble elements such as argon and xenon combine large atomic masses with high scintillation yields. The size of their atoms and the density of the liquid gives them large cross sections as a target for rare event searches, such as in dark matter experiments.

- Liquid organic scintillators consist of an organic solvent which absorbs energy from the ionizing particle, and a fluorophore which produces scintillation light. The high hydrogen content makes this scintillator type particularly useful for observing neutron-proton scattering, and its scalability makes it a good choice for neutrino detectors.

- Inorganic crystals such as NaI(Tl) or CsI(Tl) have a high proton number and are radiation hard, making them suitable for electromagnetic calorimetry in large collider experiments. In ALPHA, CsI crystal scintillators are used to detect the 511 keV photons from positron annihilation, as discussed in Section 3.4.5.

- Plastic scintillators are inexpensive, have fast decay times, and are easily produced in a variety of shapes. This makes them ideal for timing counters, veto panels, and large area detectors. Plastic scintillator panels are used to detect pions from antiproton annihilations in ALPHA, as discussed in Section 3.4.5.

The type of plastic used in ALPHA is called polyvinyltoluene (PVT), although a variety of different organic polymers can be used. The key structure is the benzene ring, a hexagonal arrangement of six carbon atoms with a unique set of molecular bonds called pi bonds. Here, electrons are delocalized over the full benzene ring, supporting low-lying excited states with energies around 3 eV.

Photons are produced as these excited states decay. However, they cannot travel through the scintillator without being reabsorbed by the same transition in another atom. To address this, the plastic is doped with a small concentration of an atom called a fluorophore. The level structure of these atoms allows them to emit photons with a different wavelength than the one absorbed. The wavelength is chosen such that it is minimally attenuated by the bulk of the scintillator, and can be easily captured by a photon detector.

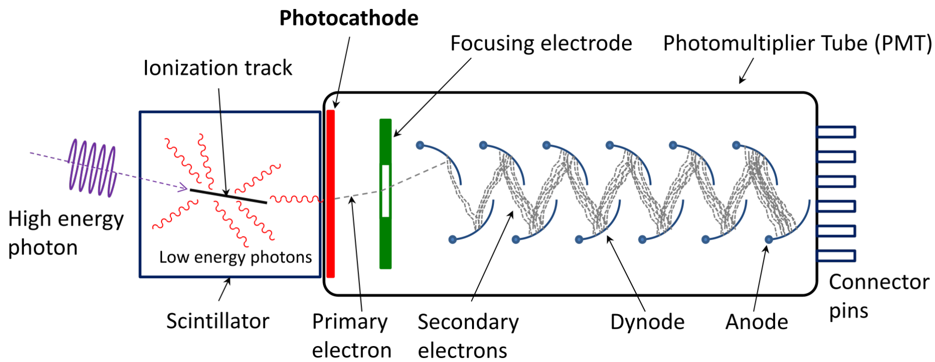

Figure 4.5: Diagram of a photomultiplier tube (PMT), shown in conjunction with a scintillator detecting a high energy photon. Image courtesy: Creative Commons, used under license 3.0. Author: Qwerty123uiop)

4.3.2 Photosensors

Once photons are produced in a scintillator, they must be detected. The classic technology for this is the *photomulitplier tube* (PMT), shown in Figure 4.5. An incident photon strikes the photocathode and is converted to an electron via the photoelectric effect. This electron is accelerated towards a series of dynodes, each held at increasingly high positive voltage. At each dynode, the electron frees multiple secondary electrons; these are in turn accelerated towards the next dynode, creating an exponentially increasing number of electrons analogous to the avalanche discussed in Section 4.2. After several stages, the electron cascade reaches the anode, where the total charge is read out electronically. The combination of a plastic scintillator and a PMT is well-established and has seen widespread general use in ALPHA and elsewhere.

However, PMTs have some notable limitations. They have a large physical footprint, require high voltage power supplies, and are sensitive to magnetic fields as the electrons are deflected away from the dynodes by Lorentz forces. These issues have led to the adoption of a more modern device for photon detection, called the *silicon photomultiplier* (SiPM).

SiPMs are composed of silicon, which is a semiconductor. Electron energy levels form continuous bands rather than discrete levels. In an insulator, the filled valence band is separated from the empty conduction band by a band gap. In a semiconductor, this band gap is small, around 1.1 eV for pure silicon. This allows thermal energy to excite electrons from the valence band

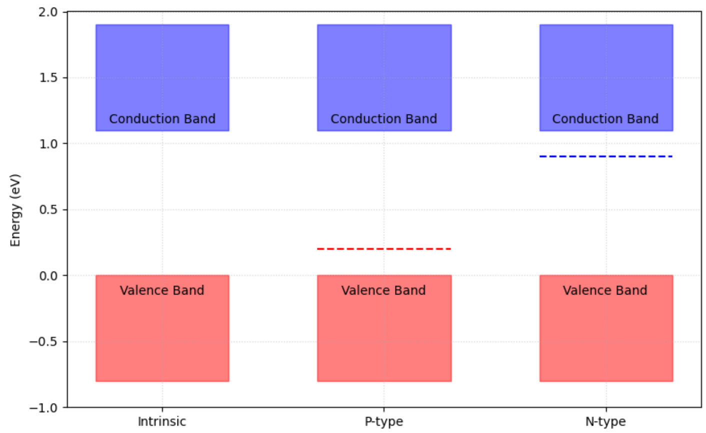

Figure 4.6: Diagram showing the energy levels of intrinsic, p-type, and n-type silicon. Dashed lines indicate additional energy levels of dopant atoms.

to the conduction band, producing an electron-hole pair. Both the electron and the hole have some mobility over the silicon lattice.

Silicon can be doped with ions to add extra electrons or holes, whose energy is slightly above the valence band or below the conduction band. This leads to either n-type silicon, which has extra weakly-bound electrons, or p-type silicon, which has extra vacancies which accept electrons and easily produce holes. These are shown in Figure 4.6.

A silicon diode is formed by combining p-type and n-type silicon. Free electrons produced in the n-type region diffuse across the interface to fill the extra holes in the p-type region. Since both types of silicon begin with a net neutral charge, this creates a buildup of electric charge, causing a potential gradient which opposes and eventually halts the flow of electrons and holes. When an equilibrium is reached, the area around the interface is called a depletion region, as it has very few free electrons and holes. This is shown in Figure 4.7.

In a *silicon photodiode*, a reverse bias voltage is applied to enhance the intrinsic potential gradient, expanding the depletion region. When a charged particle deposits energy in this region, it creates electron-hole pairs and temporarily allows a current to flow; this is the basis of a silicon detector such as the ALPHA-2 SVD.

In a *single photon avalanche detector* (SPAD), a significantly higher reverse bias voltage is applied, which accelerates electrons towards the cathode at the edge of the n-type silicon. Above a certain voltage known as the break-

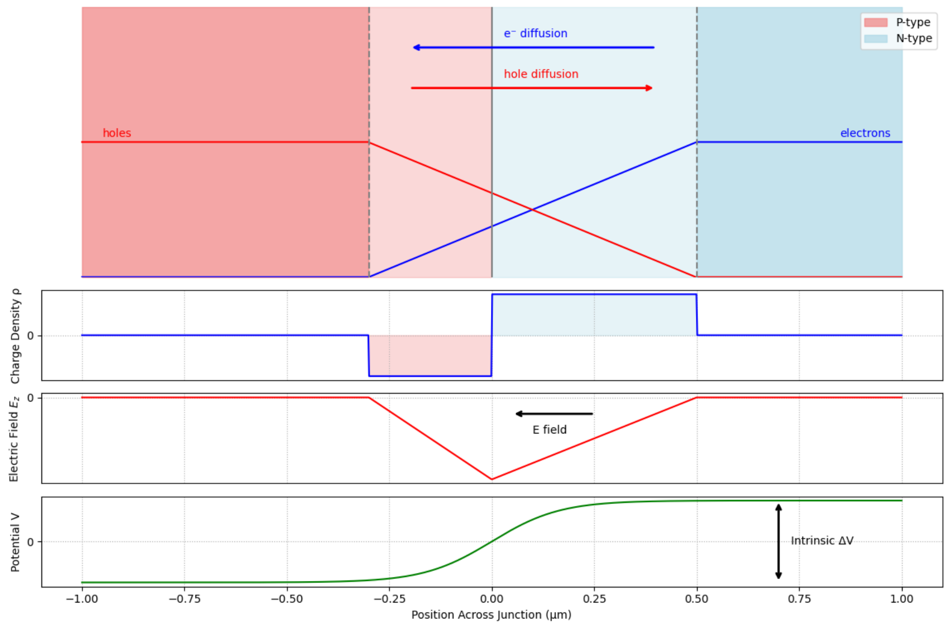

Figure 4.7: Diagram showing the p-n junction in a silicon diode in an unbiased equilibrium, including charge, electric field, and potential. The depletion region is indicated by lighter shading.

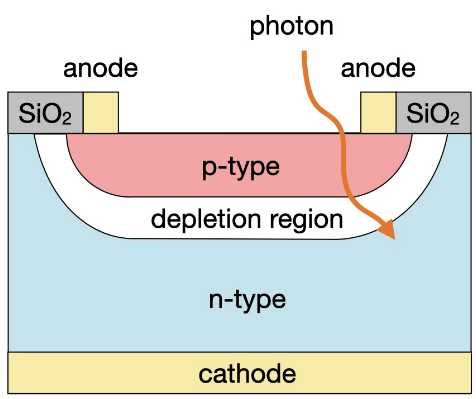

Figure 4.8: Layout of a typical single photon avalanche detector (SPAD). A silicon photomultiplier is a grid of many neighbouring SPAD cells.

down voltage, the electron drift velocity is sufficient to create additional electron hole pairs, causing an avalanche analogous to the gas ionization avalanche in a wire chamber. In this way, a single photon produces a measurable current. The avalanche is eventually halted by a quenching resistor, and the SPAD recovers and returns to its initial state. The sensitivity of the SPAD can also be a downside, as a single thermally-generated electron-hole pair will also cause an avalanche. These avalanches are indistinguishable from those produced by incident photons, causing a significant *dark count rate*. Although there are many varieties of SPAD, a typical geometry is shown in Figure 4.8.

In a SiPM, many SPADs are grouped together, each functioning independently. The total output signal is the sum of the pulses from each triggered cell. This provides some benefits: multiple photons can be detected simultaneously in different cells, the output signal is proportional to the number of incident photons, and dead time is minimized as the other cells are still active while one cell is recovering.

Compared to the PMT, the SiPM is smaller, cheaper, requires lower bias voltage, and works better in a magnetic field. They are used in the newer external plastic scintillator panels in ALPHA, as well as in the BSC.

4.3.3 The ALPHA-g barrel scintillator

To veto cosmic ray background events in the ALPHA-g TPC, it is surrounded by a second detector called the barrel scintillator (BSC). This detector is composed of plastic scintillator bars coupled to SiPMs. This design was chosen for a few reasons:

- Compactness: The BSC must fit in the small (around 2 cm) radial region between the TPC and the external solenoid.
- Speed: Plastic scintillators and SiPMs are fast enough to measure particle time-of-flight across a few centimetres, which is required for background identification.
- Simplicity: They are well-understood technologies which already see use in ALPHA external panels.
- Cost-effectiveness: They are cheaply commercially available, and expertise is available at TRIUMF for electronics design.

The BSC is composed of 64 bars of plastic scintillator. Each bar is trapezoidal in shape, with top width 23.9 mm, bottom width 21.9 mm, and

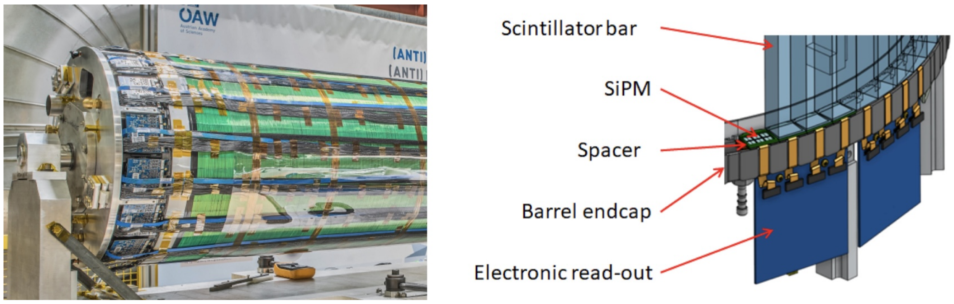

Figure 4.9: Left: A photograph of one end of the BSC, fully assembled and awaiting insertion into ALPHA-g. Image credit: CERN, 2018. Right: A diagram showing the coupling of the scintillator bars to the SiPMs, and accompanying readout electronics. Image credit: Nicolas Massacret.

side length 20 mm. The length of each bar is 2.604 m. They are wrapped in reflective tape to maximize internal reflectivity. The bars are bound together by fishing wire and aluminium clamps into a cylindrical shape with an inner radius of 223 mm and an outer radius of 243 mm, allowing this detector to be placed around the TPC and inside the external solenoid magnet. The layout of the BSC is shown in Figure 4.9.

The plastic scintillator chosen for the bars is EJ-200 from Eljen Technologies [132]. A minimum ionizing particle deposits around 2 $\mathrm{MeVcm^2g^{-1}}$ of energy, or around 4 MeV on a 2 cm path through the 1.023 $\mathrm{gcm^{-3}}$ bar. The light yield of EJ-200 is around 10 000 photons/MeV, for a typical production of around 40 000 photons by an incident charged particle. These photons are created over a few nanoseconds: the scintillation rise time is quoted at 0.9 ns and the decay time at 2.1 ns. The bulk attenuation length of EJ-200 is 3.8 m; neglecting surface attenuation, at least 10 000 of these photons should reach each bar end.

As part of this work, a ray tracing simulation was performed to study the propagation of photons though the scintillator bars. The nominal index of refraction and bulk attenuation lengths were used, while the surface reflectivity was varied to best match the BSC data. This hinted towards a poor surface reflectivity of around 0.9, resulting in $\sim 80\%$ of photons being lost due to surface attenuation. If this model is accurate, around 2000 photons should reach each bar end, with this number varying depending on the scintillation position.

The scintillator bars are coupled to SensL MicroFJ-60035 SiPMs [133], now produced by ON Semiconductor [134]. These have a size of 6 mm by

6 mm, and so an array of six SiPMs is used to cover each bar end, providing 47% surface coverage. They are optically coupled using 1 mm to 2 mm of RTV615 silicon rubber to eliminate air gaps, which was found to increase light transmission by 30%. The maximum efficiency of the SiPM is at 420 nm, a good match with the 425 nm peak emission output of the scintillator.

The breakdown voltage of these SiPMs is 24 V, and they are operated at a reverse bias of 27 V. The variation in breakdown voltage between SiPMs is sufficiently narrow that all SiPMs can be biased using the same voltage. In addition to the standard output, they provide a fast output which has a full width half maximum of 3.0 ns. An air cooling system is used to carry away heat from the readout electronics, which also serves to slightly lower the dark noise rate. A vortex cooler pumps cold air into the volume of the BSC at the top and the bottom, while a fan maintains an air flow in the volume between the TPC and the BSC to minimize heating of the SiPMs due to the TPC electronics.

4.3.4 Detection principles

In this discussion, and in the discussion of the BSC data processing and calibration in Chapters 6 and 7, I will reference some quantities which are measured using the BSC. These symbols are listed in Table 4.1 for reference.

Each BSC bar is read out at both ends to allow determination of the longitudinal position of a scintillation. This can be accomplished by two methods: charge division, and time difference.

In the charge division method, the number of photons arriving at each end of the bar is determined by integrating the SiPM signal. The number of photons decreases with light travel distance: $N = N_0 e^{-x/\lambda}$, where λ is the effective attenuation length, x is the distance travelled, and N_0 is the initial number of photons travelling towards a photosensor. Thus the hit position relative to the bar centre can be determined from the ratio of the top and bottom signal sizes:

$$\begin{aligned} \frac{N^{\text{top}}}{N^{\text{bot}}} &= \frac{N_0 e^{-d^{\text{top}}/\lambda}}{N_0 e^{-d^{\text{bot}}/\lambda}} = e^{(d^{\text{bot}}-d^{\text{top}})/\lambda} \\ \implies z &= \frac{d^{\text{bot}} - d^{\text{top}}}{2} = \frac{\lambda}{2} \ln \frac{N^{\text{top}}}{N^{\text{bot}}} \end{aligned} \tag{4.3.1}$$

Here, we define our coordinate system so that $z = 0$ is at the centre of the bar. Consequently, the distance of the hit from the top and bottom of the bar are given by $d_{\text{top}} = L/2 - z$ and $d_{\text{bot}} = L/2 + z$ respectively,

Symbol	Description
t_i^{top}, t_i^{bot}	The start time of the SiPM signal at the top or bottom of the i'th bar. This is measured by a TDC (see Section 4.4) and includes all time corrections (see Chapter 7).
$t_{\text{TDC},i}^{\text{top}}$, $t_{\text{TDC},i}^{\text{bot}}$	The start time of the SiPM signal at the top or bottom of the i'th bar measured by the TDC, without any time corrections.
N_i^{top}, N_i^{bot}	The number of photons detected by SiPMs at the top or bottom of the i'th bar.
A_i^{top}, A_i^{bot}	The amplitude of the SiPM signal at the top or bottom of the i'th bar. Measured by an ADC (see Section 4.4).
$t_{0,i}$	The bar hit time, the time when a charged particle passes through the i'th bar.
z_i	The bar hit z position, the vertical height where a charged particle passes though the i'th bar. Defined so that $z = 0$ is at the centre of the detector.
d_i^{top}, d_i^{bot}	The distance from the bar hit z position to the top or bottom of the i'th bar.
λ_i	The effective attenuation coefficient of light in the i'th bar.
$v_{\text{eff},i}$	The effective speed of photon prorogation in the i'th bar.
L	The length of each bar ($L = 2.604\,\text{m}$).
$\text{TOF}_{i,j}$	The *time-of-flight* between two bar hits on bars i and j.
$\text{TOF}_{\text{exp},i,j}$	The expected *time-of-flight* between two bar hits on bars i and j, calculated for a cosmic ray travelling at the speed of light with a straight trajectory.

Table 4.1: A list of symbols used when discussing the data handling of the BSC. The subscript i or superscripts top and bot are often dropped if the bar number or bar end is not relevant to the discussion.

where L is the bar length. We also use the assumption that photons are produced isotropically. Instead of using the total number of photons N, one can equivalently use the signal amplitude A or even time over threshold.

For the time difference method, we use the approximation that photons travel with a constant effective speed v_{eff} along the bar. The difference between the scintillation time t_0 and the photon detection time t^{top} or t^{bot} is linear with the distance along the bar between the hit position and the SiPM: $t^{\text{top}} = t_0 + d^{\text{top}}/v_{\text{eff}}$ and $t^{\text{bot}} = t_0 + d^{\text{bot}}/v_{\text{eff}}$. As a result, the hit position along the bar can be determined using the time difference between the top and bottom signals:

$$\begin{aligned} z &= \frac{d^{\text{bot}} - d^{\text{top}}}{2} \\ &= \frac{v_{\text{eff}}}{2}[(t^{\text{bot}} - t_0) - (t^{\text{top}} - t_0)] \\ &= \frac{v_{\text{eff}}}{2}\left(t^{\text{bot}} - t^{\text{top}}\right) \end{aligned} \tag{4.3.2}$$

The main purpose of the BSC is to measure the *time-of-flight* (TOF) of charged particles, defined as the time difference between two hits on different bars. We can find the time t_0 that the charged particle hits the bar from the start time of the top and bottom SiPM signals:

$$\begin{aligned} t_0 &= \frac{1}{2}\left[\left(t^{\text{top}} - d^{\text{top}}/v_{\text{eff}}\right) + \left(t^{\text{bot}} - d^{\text{bot}}/v_{\text{eff}}\right)\right] \\ &= \frac{t^{\text{top}} + t^{\text{bot}}}{2} - \frac{L}{2v_{\text{eff}}} \end{aligned} \tag{4.3.3}$$

Taking the length and effective speed of light to be the same across all bars, the second term is a constant of around 10 ns. Thus the time difference between two bar hits is given by:

$$\begin{aligned} \text{TOF} &= t_{0,i} - t_{0,j} \\ &= \frac{t_i^{\text{top}} + t_i^{\text{bot}}}{2} - \frac{t_j^{\text{top}} + t_j^{\text{bot}}}{2} \end{aligned} \tag{4.3.4}$$

This TOF can be used to distinguish between cosmic ray muons and annihilations, as shown pictorially in Figure 4.10. For relativistic cosmic ray muons, we expect their TOF to be given by $\text{TOF}_{\text{exp}} = D/c$, where D is the distance between where the muon crosses bars A and B, and c is the speed of light. A horizontal cosmic ray passing through the centre of the detector has $\mathrm{TOF}_{\mathrm{exp}} \sim 1.5\,\mathrm{ns}$; typical cosmic rays have $\mathrm{TOF}_{\mathrm{exp}}$ between 1.5 ns and 2 ns. On the other hand, pions produced in an annihilation originate inside the detector, and travel similar distances outwards to the BSC. The TOF, as calculated above, does not represent a true particle path, but rather an *effective TOF*. Since the distance travelled by each pion is similar, this effective TOF will be close to zero. By comparing the measured TOF to the expected cosmic ray $\mathrm{TOF}_{\mathrm{exp}}$, cosmic rays can be distinguished from antihydrogen annihilations.

Using these methods, the longitudinal hit position can be determined using either charge division or time difference, and measuring the effective TOF, the BSC can be used to distinguish between antihydrogen annihilations and cosmic ray muons. To fully implement this scheme, many technical details and higher order effects were addressed. These details, and the results of this background rejection scheme, are described in Chapters 6, 7, and 8.

4.4 Electronics and data acquisition

To handle data collection, ALPHA-g adopted the MIDAS data acquisition (DAQ) framework [135] developed at TRIUMF and at the Paul Scherrer Institute. This software is used in many diverse physics experiments, including current use in ALPHA-2. It handles many aspects of the experiment:

- A web-based user control interface for many experiment subsystems.
- Continuous logging of slow control variables.
- Hardware frontends for TPC and BSC detector services.
- Streaming of digitized data from the TPC and BSC.
- An event builder that assembles TPC and BSC data by timestamp.
- A configuration database for run control settings.
- Data storage on local disks and CERN servers.

The remainder of this section describes how the TPC and BSC signals are processed and digitized by custom ALPHA-g readout electronics before being fed into MIDAS. The full readout chain is shown diagrammatically in Figure 4.11.

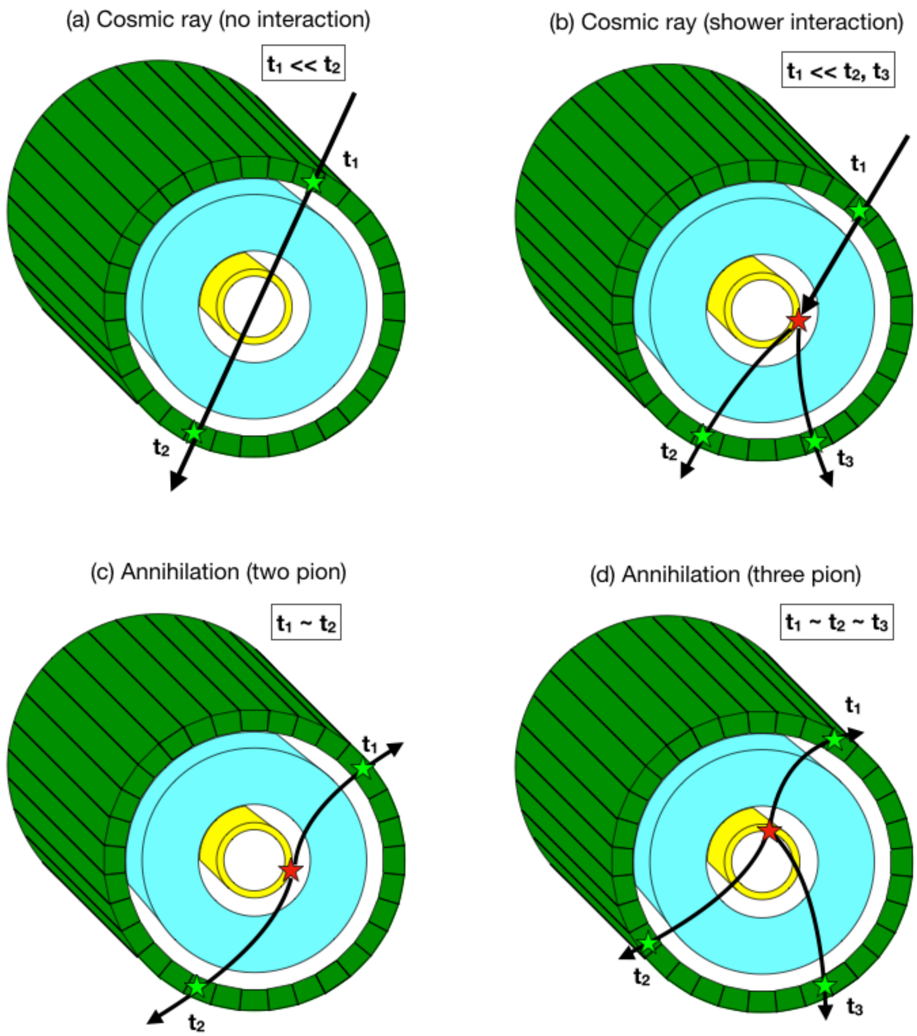

Figure 4.10: Schematic diagrams of particles traversing the BSC (green), TPC (blue), and trap wall (yellow). Inlaid boxes indicate hit times. (a.) A cosmic ray without interacting. (b.) A cosmic ray with a simple shower interaction. (c.) An antihydrogen annihilation with two charged pions. (d.) An antihydrogen annihilation with three charged pions.

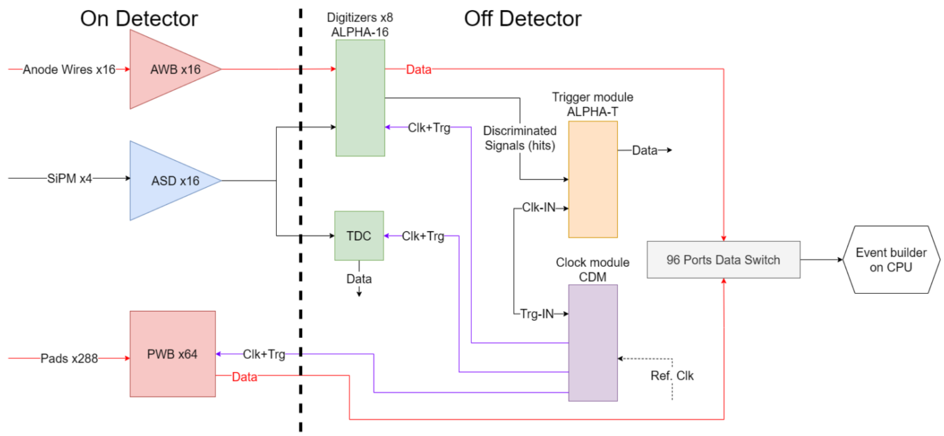

Figure 4.11: Diagram of the readout electronics chain for the TPC and BSC detectors. Taken from [136].

4.4.1 TPC data acquisition

There are two sources of data from the TPC: signals from the 256 anode wires, and those induced on the 18432 cathode pads. Although there were originally plans to read out the anode wire charge from both ends, only the top end of the anode wires are instrumented. This is done by custom pre-amplifiers called *anode wire boards* (AWBs). There are 16 AWBs mounted directly to the top of the TPC, each reading out 16 anode wires. The anode wire signal is decoupled from the high voltage and amplified by the AWB. From the AWB, the anode wire signal is carried off the detector to *analog-to-digital converter* (ADC) digitizers called the Alpha-16, based on the GRIF-16 digitizers previously developed at TRIUMF [137]. When triggered, they digitize all anode wire signals with a sample rate of 62.5 MHz (16 ns sample length) and a dynamic range of 12 bits. The digitized waveforms are transmitted to the DAQ computer via ethernet. There are 8 Alpha-16 digitizers, each receiving signals from 32 wires via 2 AWBs.

The readout of the 18432 cathode pads is significantly more challenging due to the large number of channels. After amplification, these signals are digitized directly on the detector by application-specific integrated circuits (ASICs), using the AFTER switched capacitor array originally designed for the T2K experiment [138]. The capacitors are opened and quickly closed, one at a time, so that they are charged by sampling the cathode pad signal. These capacitors are then discharged in turn, and their charge recorded by an ADC mounted on the printed circuit board. This cathode pad data digitization is the limiting factor for the maximum readout rate of the TPC, nominally 500 Hz.

Each ASIC reads out 72 cathode pads. Four of these ASICs are grouped together on one *pad wing board*, or PWB, each handling 288 cathode pads. There are a total of 64 PWBs needed for all cathode pads. The digitized data from all four ASICs is collected by one central FPGA on each PWB, and then transferred off the detector and to the DAQ computer via a 1 gigabit/second optical fibre and ethernet switch.

The density of readout and digitization electronics on the detector uses significant power and creates excess heat. Copper bus bars supply low voltage to the PWBs, which draw close to 1000 W across the whole detector. The heat created is dissipated by water cooling, flowing through copper pipes mounted beside the bus bars, and the temperature is carefully monitored by temperature probes mounted to the readout electronics.

4.4.2 BSC data acquisition

There are 64 BSC bars, each read out by an array of 6 SiPMs at both ends, for a total of 768 SiPMs. The array of SiPM signals is processed by custom-built *analog sum discriminator* (ASD) boards. Each ASD handles four bar ends, requiring a total of 16 ASDs mounted to each end of the detector.

After initial pre-amplification, the fast output signals from the 6 SiPMs on a bar end are combined by an analog sum circuit. This is followed by some further amplification and pulse shaping. Here, the signal path is split:

- One branch of the signal is passed through an RC low-pass filter pulse shaper, which stretches and lengthens the pulse by an approximate factor of 5. This signal is carried off the detector to the same ADCs which digitize the anode wire signals. Each ADC has 16 channels dedicated to the BSC, which are digitized at 100 MHz (10 ns sample length) and with a dynamic range of 14 bits.

- A second signal branch is passed through a fast comparator on the ASD. A programmable threshold voltage is generated by a digital-to-analog converter (DAC) on the ASD; the comparator outputs a digital pulse when the combined SiPM signal exceeds the threshold. This digital signal is carried off the detector to a *time-to-digital converter* (TDC).

ALPHA-g uses a single 256-channel TRB3 TDC manufactured at the GSI Helmholtz Centre for Heavy Ion Research, which has a precision better than 20 ps [139]. The TDC digitizes the signal arrival time using an FPGA delay line. It uses an externally-supplied 200 MHz clock; when a signal arrives from the BSC, it propagates through the delay line until it is frozen at the start of the next clock cycle. The number of delay elements traversed by the signal indicates the signal time relative to the clock cycle. The leading and falling edge of the signals are digitized separately, although only the leading edge is used in the BSC analysis. When triggered, all TDC events in a roughly 10 µs window are sent from the TDC to the MIDAS event builder.

Further details on the implementation of the BSC readout electronics and data acquisition are provided in Appendix A.

4.4.3 Clock, trigger, and test pulse

Synchronization of the ADCs, TDC, and other readout electronics in ALPHA-g is maintained using a dedicated clock distribution system. A custom clock module based on the clock developed by the GRIFFIN collaboration at TRIUMF [137] provides a stable timing reference. This module receives an input 10 MHz clock from the AD, and distributes derived clocks at 100 MHz and 62.5 MHz to the ADCs, and at 200 MHz to the TDC.

The triggering of the detectors is handled by a custom *ALPHA-T* trigger module developed at TRIUMF, based on the GRIFFIN GRIF-C module [137]. Fast discriminated detector hit information is transmitted from the ADCs to the trigger module. The trigger module processes this data and makes a trigger decision, then distributes a global trigger signal via the clock module to synchronize data acquisition across all subsystems. The triggers are sent directly from the trigger module to MIDAS, allowing for fast counting of annihilations which is not limited by the 500 Hz readout rate of the detectors.

Many trigger configurations are available via the MIDAS run control system. The original design intention was to trigger on a multiplicity of BSC bar hits, and this scheme has been tested successfully. However, the operating configuration for the initial gravity measurement instead uses anode wire signals from the TPC as the trigger source. A trigger is sent when signals are received on two or more AWBs, with a gap of at least one AWB between them with no signal. This trigger condition was chosen to suppress cosmic ray events which create signals on only a single contiguous group of AWBs.

An additional feature implemented is a built-in test pulser system, which can inject known signals into the TPC and BSC electronics to verify system functionality. Each ADC module includes a programmable digital-to-analog converter (DAC) capable of generating a test pulse with adjustable shapes. To test the readout of the TPC anode wires and cathode pads, this signal can be fed into the field wires, which induces a pulse on the anode wires. For the BSC, the test pulse is injected immediately after the first pre-amplification stage of the SiPM signal path, enabling testing of the full analog and digital readout chain.

Chapter 5

TPC event reconstruction

This is the first of four chapters describing my original contributions to ALPHA-g. I have spent a significant fraction of my time improving the reconstruction of the TPC data, resulting in considerable developments which directly lead to the detector performance meeting the resolution and background rate targets. An attempt is made to quantify the effects of these improvements on the gravity measurement in Section 9.4.1.

This chapter is organized as follows. Section 5.1 outlines the data processing strategy for TPC data. Section 5.2 describes my work on the ALPHA-g event display, which was used to inform improvements to the TPC reconstruction. Section 5.3 describes in detail the annihilation position reconstruction algorithm, clearly distinguishing my own improvements and original contributions, primarily in Sections 5.3.4, 5.3.5, 5.3.6, and 5.3.7. Section 5.4 will quantify the results of these improvements. I have also been involved in an effort to develop an alternative vertex finding algorithm based on machine learning and neural networks, which is briefly discussed in Section 5.5.

5.1 TPC data processing

The MIDAS data acquisition system described in the previous chapter reads out the raw data signals from the TPC and BSC into events based on their timestamp. The acquired data is stored in a memory buffer before being written permanently to disk. The data structure follows a MIDAS-specific open-source format. Due to the large number of cathode pads, a substantial quantity of data is produced: around 60 GB per hour for cosmic ray data with the typical trigger configuration, and up to 100 GB per hour during typical antihydrogen production. This data is automatically copied from local ALPHA storage drives to CERN's EOS servers [140] for redundancy and long-term storage.

ALPHA has developed a dedicated alphasoft software suite [141] to analyze this data, primarily written in ROOT/C++ and based on the MIDAS analyzer [135, 142]. This framework has been used for some time to analyze data from the ALPHA-2 SVD. It is modular in structure, allowing different executable programs to be constructed from a list of independently programmed modules. This functionality is used to maintain distinct offline and online analysis programs which share many elements. The offline analysis reads the MIDAS data files stored locally or on the EOS servers and handles the reconstruction steps leading to the determination of the annihilation position. The alphasoft output is a distilled data file in ROOT format which contains higher-level information such as the reconstructed annihilation position and the number of tracks.

The online analysis mode performs the same reconstruction of the annihilation vertex position, but instead reads directly from a MIDAS memory buffer filled in real time by the dedicated MIDAS event builder. It is run on a dedicated computer with 32 cores and 252 GB of RAM so that it can keep pace with the TPC's 500 Hz maximum event rate. The alphasoft suite includes tools for quickly visualizing the time and position of annihilation events, including a real-time online display of annihilation positions, to which I have directly contributed.

A number of summary plots are automatically generated at the end of each run, allowing detector data to be used in real time for the development and debugging of antihydrogen production. Quality-of-data monitoring histograms are also produced both online and offline to assess detector and reconstruction performance.

5.2 Event display

Before discussing the vertex reconstruction algorithm and my contributions to it, it is prudent to cover the ALPHA-g event display. This was a crucial tool for finding issues with the reconstruction and testing solutions, without which many improvements would not have been possible. To this end, I made considerable contributions to the event display in parallel to my work on the reconstruction algorithms. This section will give an overview of the ALPHA-g event display and highlight my contributions. My implementation of the BSC into the event display will be discussed later in Section 6.9.

The event display for ALPHA-g is based on the XSNOED event display used for the SNO experiment [143]. Integrated into the alphasoft reconstruction software as an analysis module, it generates an interactive three-dimensional rendering of each event. Figures 5.1 and 5.2 show the event display across a range of events.

To inform the modifications to the reconstruction algorithms, I spent considerable time debugging and improving the event display, which was not

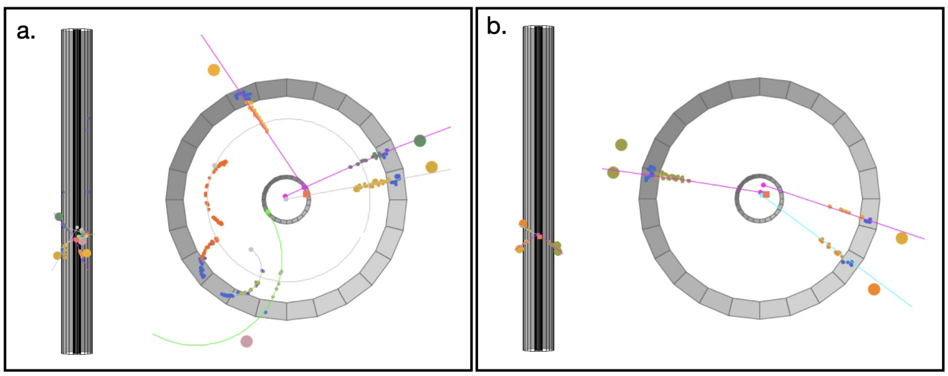

Figure 5.1: The ALPHA-g event display, side-on and top-down views. Spacepoints are shown as small dots, helix fits as curved lines, and the vertex position as a red square. Larger dots represent barrel scintillator hits (Section 6.9). Simple wireframe mesh cylinders mark the outer TPC radius and the trap radius. (a.) An annihilation where a vertex (red square) was found from two helices (purple lines). An attempt was made to add one helix (green line) to the vertex, but this worsened the fit. Many helices failed the selection cuts (grey lines). (b.) A simpler annihilation event. A third helix (cyan line) was successfully added to improve the vertex fit.

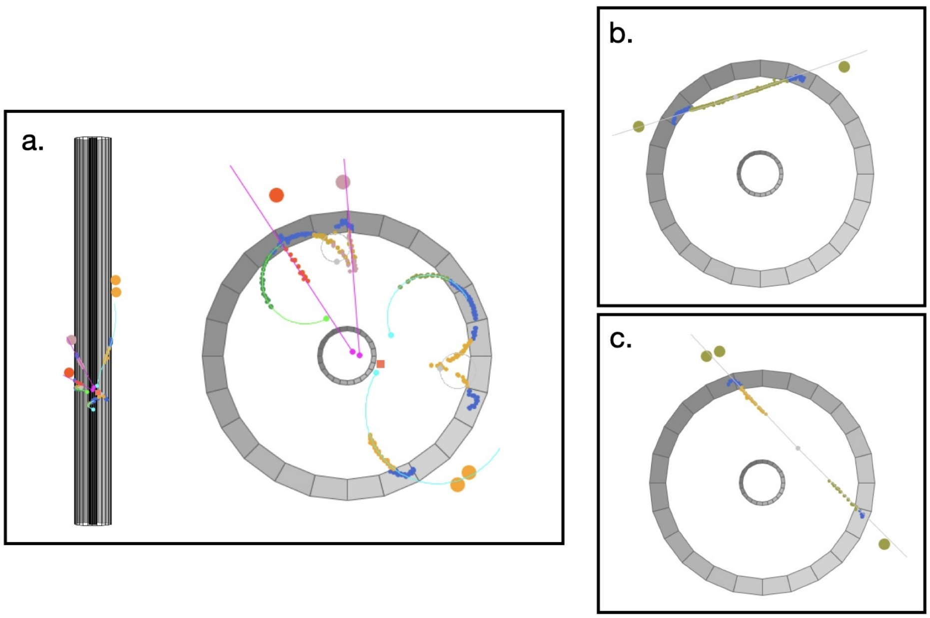

Figure 5.2: The ALPHA-g event display, side-on and top-down views. (a.) An annihilation event where some pair production tracks have been included in the vertex fit. (b.) A cosmic ray which was fit with one double-sided helix. (c.) A cosmic ray which was initially two tracks, but was joined into one by the track joiner.

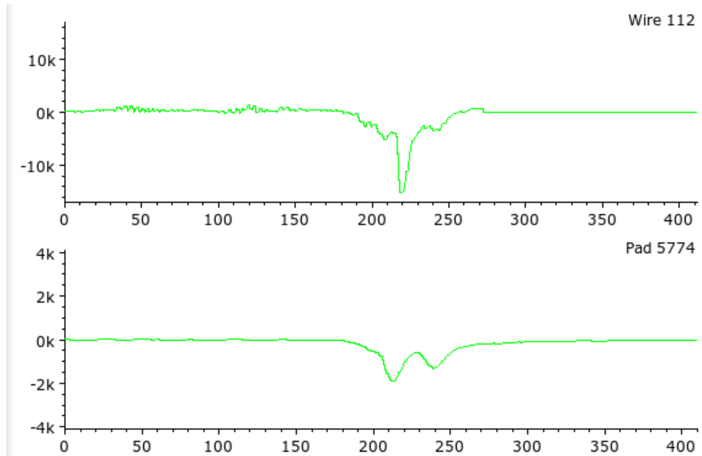

Figure 5.3: The TPC waveform display window in the event display. Both the anode wire waveform (top) and matched cathode pad waveform (bottom) are shown. The x-axis is time in 16 ns bins, and the y-axis is voltage in ADC bits.

previously in a functional state. This work included fixing a number of bugs in the geometry which caused helices and vertices to be drawn in the wrong place. I also added a number of features, such as:

1. An option to colour spacepoints based on their track grouping.
2. Inclusion of the *truth vertex* in simulated events, representing the simulated annihilation position.
3. Options to display only a subset of events matching given criteria, so that the event display could be used to diagnose specific problems.

Alongside the visualization, quantitative information are available in panels for the event, for each spacepoint, and for each helix fit. The visualization is fully interactive: individual spacepoints and helices can be selected to display their corresponding information panel. I added the helix fit information panel and greatly extended the data available in these panels, making them powerful diagnostic tools. A waveform viewer is also built into the event display. Selecting a spacepoint displays the associated anode wire and cathode pad signals, as shown in Figure 5.3.

5.3 Annihilation position reconstruction

The goal of the TPC reconstruction is to determine the three-dimensional position of each antihydrogen atom as it annihilates, called the annihilation vertex. In most cases when antihydrogen is released, the vertex lies along the inside wall of the trap, as this is the first matter the atom will encounter. However, antihydrogen can also annihilate on background gas inside the vacuum chamber, and so the reconstruction algorithm must be able to handle annihilations occurring at any position inside the trap.

The reconstruction proceeds in nine stages, which are illustrated pictorially in Figure 5.4 and as a block diagram in Figure 5.5. To aid the reader, each stage is first introduced in the text with the number corresponding to these diagrams. These stages roughly correspond to individual modules in the alphasoft analysis chain.

This section will describe each stage of the TPC event reconstruction in considerable technical detail. It will also distinguish between three historical versions of the software: the original pre-2022 version, the version used in the gravity paper, and the current improved version.

The reader who is not interested in understanding the minutia of how TPC data is processed to recover annihilation vertices should feel free to skip ahead to Section 5.4 for a summary of the impact of the changes made throughout this work.

5.3.1 Signal deconvolution

A charged particle ionizes the TPC gas along its path, as discussed in Section 4.2.1. A minimum ionizing cosmic ray muon produces around 0.6 ionization clusters per millimetre, as does a charged pion from an annihilation; each cluster consists of one or more electron-ion pairs. Many of these ionization electrons drift towards the same anode wire, or induce signals on the same cathode pads. As a result, the recorded waveforms can be complex, requiring deconvolution to resolve the individual electron arrival times. In typical TPC geometries, this can be mitigated by designing the drift direction to be orthogonal to the expected track directions, so that ionization clusters along a track are more likely to to drift to different wires. Given the geometry of the ALPHA-g TPC, this was not practical, and so the need for accurate signal deconvolution is increased.

The deconvolution procedure is similar for both the anode wire signals and the cathode pad signals. The goal of the **anode wire deconvolution module (1)** and **cathode pad deconvolution module (2)** is to extract

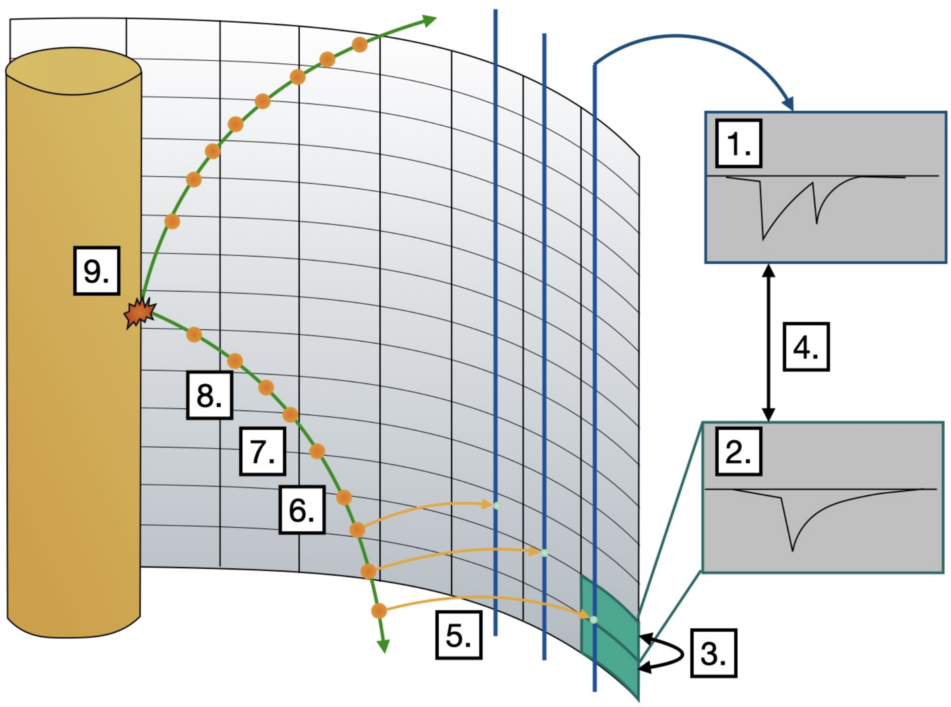

Figure 5.4: A diagram showing the nine stages of the annihilation vertex reconstruction in relation to the TPC geometry. The geometry is not to scale, and the number of wires and pads is significantly reduced for readability. Inlays show simplified anode wire and cathode pad waveforms.
1. Anode wire signals are deconvolved. 2. Cathode pad signals are deconvolved. 3. Nearby pad signals are grouped. 4. Corresponding wire and pad signals are matched. 5. Ionization positions are determined via time projection. 6. Tracks are identified. 7. Tracks are fit. 8. Segmented tracks are joined. 9. The vertex position is determined.

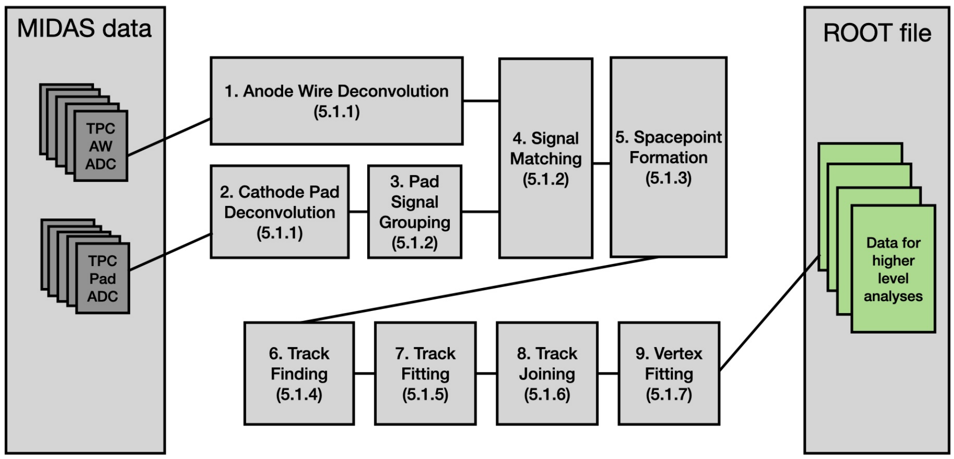

Figure 5.5: A block diagram indicating the reconstruction stages, roughly corresponding to modules in the alphasoft analysis chain. The stages are numbered corresponding to Figure 5.4, and also indicate the section of this dissertation where they are discussed.

signal start times and amplitudes from individual waveforms which may contain many overlapping signals.

The deconvolution is based on template waveforms which correspond to the response of the anode wire and cathode pad to a single electron avalanche. These response functions were produced using Garfield++ simulations of the avalanche process [5] and verified experimentally.

The deconvolution process for the two types of TPC signal is shown in Figure 5.6. It begins by iterating through each bin of the digitized waveforms. For each waveform, if the value in that bin exceeds a given threshold, the response function is subtracted from the waveform, scaled to account for the number of electron-ion pairs in the ionization cluster. This process is repeated for each waveform above the threshold, before proceeding to the next bin.

In the case of anode wires, additional care is taken to account for induction between adjacent wires. An avalanche will induce an opposite-sign signal on neighbouring wires [127]. To account for this, a copy of the response function is also subtracted from the four closest wires on either side. This copy is inverted and scaled by factors of 12%, 4%, 1%, and 0.4% respectively, calculated using Garfield++ simulations.

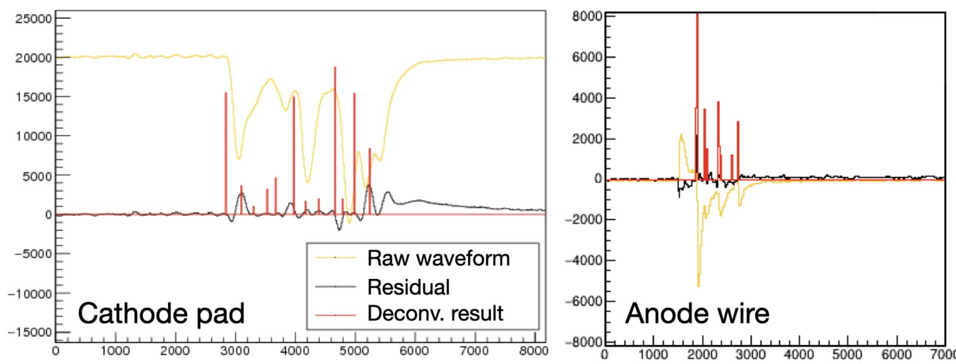

Figure 5.6: Deconvolution of cathode pad (left) and anode wire (right) signals. Yellow indicates the digitized signal, red indicates the number and time of single electron avalanches determined by the deconvolution algorithm, and black shows the residual waveform after subtracting the response functions. The x axis is time in 16 ns time bins, and the y axis is signal amplitude in ADC units. Credit: Lars Martin.

5.3.2 Signal grouping

When an avalanche occurs on an anode wire, a signal is induced on multiple nearby cathode pads. The goal of the **pad signal grouping module (3)** is to group these nearby cathode pad hits together. These combined cathode pad hits are then associated with their corresponding anode wire hit by the **signal matching module (4)**.

To speed up computation given the large number of cathode pad channels, neighbouring pads are not deconvolved simultaneously, as was done for the anode wires. Instead, each pad signal is deconvolved independently, allowing for parallelization of this computationally expensive task. The sign of the induced signal is the same for all pads, enabling the pad hits to be identified individually and grouped later.

As a result of this choice, many pad hits are produced by a single electron avalanche. The pad signal grouping module merges these neighbouring pad signals together. First, deconvolved pad signals are grouped by time: those in the same or adjacent 16 ns time bins are eligible for matching. They are then grouped by pad column, so that only hits in a single vertical column can be grouped together. A blob-finding algorithm then identifies groups of pad signals on contiguous pads. The central z position (i.e. along the TPC axis) of the group of pad signals is then determined, using either a weighted average or by fitting the pad signal heights with a Gaussian function.

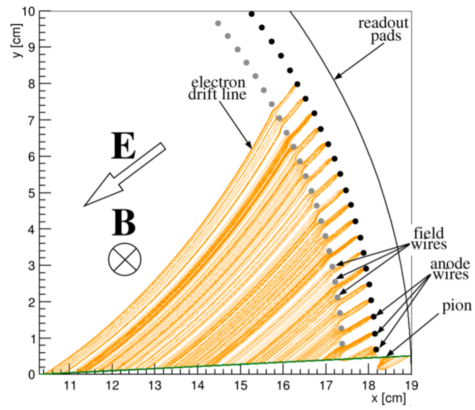

Figure 5.7: Garfield++ simulation of a charged pion (green) in the TPC. The orange lines represent the path of electrons from ionization sites along the pion path to the anode wires. These simulations were used to create space-time relation lookup table. Published in [130].

Following cathode pad grouping, next in the analysis chain is the signal matching module, responsible for combining the pad hit groups with the anode wire hits. The time difference Δt and the difference in the azimuthal coordinate $\Delta\phi$ between each wire hit and each pad hit group are calculated. Each anode wire hit is matched to the pad hit group which minimizes the two-dimensional distance $D = \sqrt{(\Delta t)^2 + (\alpha \cdot \Delta\phi)^2}$, where $\alpha = 3$ is a scaling factor empirically chosen to correctly prioritize Δt (in ns) or $\Delta\phi$ (in radians).

This matching strategy was implemented using a k-d (k-dimensional) tree data structure, which is commonly used for nearest neighbour searches in multidimensional spaces. The technique is similar to a binary tree search, and avoids computing the distance between all pairs of wires and pads, which greatly improves computational performance [144].

5.3.3 Spacepoint reconstruction

After matching the anode wire avalanche signals to the cathode pad signals they induce, the exact position of the avalanche can be determined: the radial r position is fixed at the anode wire radius, the azimuthal ϕ position is given by the wire number, and the axial z position is given by the centre of the group of cathode pad hits. The goal of the **spacepoint formation module (5)** is then to determine the original position of the ionization which created the free electron responsible for the avalanche.

This requires a detailed knowledge of the electron's path through the drift gas as it is guided by both the electric field between the anode wires and inner cathode, as well as the magnetic field created by the external solenoid magnet. To this end, detailed electron drift simulations were performed using Garfield++ [5], illustrated in Figure 5.7. These simulations were used to create a lookup table for the *space-time relation* (STR), the relationship between ionization position and drift time.

To define the drift time, a global zero time for the event is needed. This is taken as the time of the first signal observed on any anode wire, utilizing the assumption that a charged particle traversing the TPC will almost certainly pass close to at least one wire, creating an avalanche with near-zero drift time. The drift time for each avalanche is then defined as the difference between this zero time and the start time of the anode wire signal. Using the STR lookup table, this drift time is converted into a drift length, which gives the r coordinate of the original ionization. The electron drift is not completely radial; there is an azimuthal deflection due to the magnetic field, called the Lorentz angle. This angle is also included in the STR lookup table, and is used to correct the reconstructed ϕ coordinate. The ionization z coordinate remains unchanged from the central z position of the pad hit group.

The final output of this process is a three-dimensional estimate of each ionization position in the TPC gas volume, referred to hereafter as space-points. They are assigned a three-dimensional position uncertainty. The STR lookup table is used to determine this uncertainty from the hit time uncertainty (from the ADC bin width and the deconvolution), avalanche z position uncertainty (from the pad signal weighted average or fit), and wire position pitch.

5.3.4 Track finding

The **track finding module (6)** attempts to group these reconstructed ionization spacepoints into tracks, which correspond to the paths of individual charged particles through the TPC.

The original track finding algorithm, termed the **adaptive finder**, was able to identify pion tracks sufficiently for vertex position reconstruction. However, using the ALPHA-g event display discussed in Section 5.2, I identified some flaws with this method. I addressed these by implementing two further track finding algorithms. The first improved track finder, coined the **iterative finder**, was developed to directly address the issues identified with the adaptive finder, and was used in the data analysis for the 2023

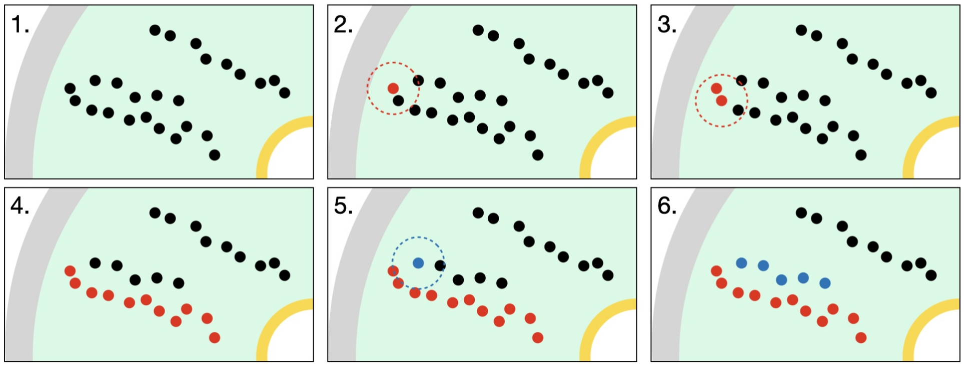

Figure 5.8: Cartoon diagram of the adaptive finder. (1.) Spacepoints in the TPC. (2.) A seed point is chosen close to the anode wires. (3.) The next point within a cutoff distance (dashed circle) is added to the track. (4.) This is repeated until the track is completed, adapting the cutoff distance as required. (5.) A seed point for a new track is chosen. (6.) This is repeated for all possible seed points at large r.

gravity paper. However, some issues still persisted. The second improved track finder, called the **exhaustive finder**, introduced a more fundamental change to the track finding technique.

Original method: adaptive finder

The original method of track finding, the adaptive finder, was developed based primarily on simulation and data from a prototype TPC at TRIUMF [145]. An illustration of this track finding procedure is shown in Figure 5.8.

It begins by sorting spacepoints by radial position r, and then proceeds as follows:

1. The largest r spacepoint is chosen as a seed point for a new track.
2. If the next spacepoint is within a fixed cutoff distance from the current point, it is added to the track and becomes the new current point.
3. This is repeated for every remaining point at smaller r.
4. Once there are no further points within the cutoff distance, the cutoff distance is incrementally increased.
5. The track is completed after the cutoff distance is increased to a maximum value.

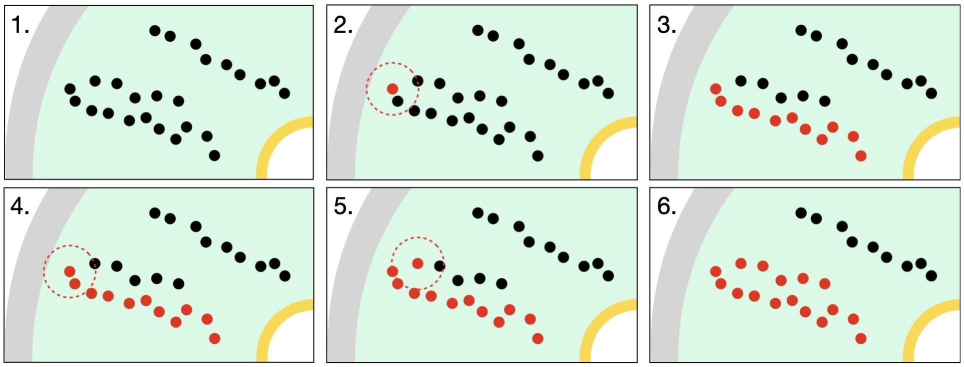

Figure 5.9: Cartoon diagram of the iterative finder. (1.) Spacepoints in the TPC. (2.) A seed point is chosen close to the anode wires. (3.) Spacepoints within a fixed cutoff distance from the previous point are added to the track. (4.) The first point in the track is checked for missed points within the cutoff distance. (5.) The tracking process is restarted from each missed point. (6.) Each point in the track is checked twice before a new track is seeded.

6. The process repeats from the remaining spacepoints with a new seed point, until all unused spacepoints up to a cutoff radius have been tried as a seed point.

In this way, the track is built from large to small r, i.e. from near the anode wires towards the TPC centre. This algorithm was effective at finding tracks which could be used to reconstruct the vertex position. However, some issues became evident once the event display was used to examine the track finding process:

- Spacepoints would often be missed from a track.
- In many cases, these missed spacepoints would form a second, overlapping track (i.e. the blue points in Figure 5.8).
- Tracks would be missed entirely if no spacepoint had large enough r to be chosen as a seed point.

First improvement: iterative finder

To address these issues, I developed a new iterative finder, which is shown schematically in Figure 5.9. It includes three key changes. The first difference is that there is no adaptation of the cutoff distance; the maximum cutoff distance is always used. The second key difference is an additional iterative step once the initial track has been built. Specifically:

1. After the initial track is formed, each of its spacepoints is rechecked for nearby, unused neighbours.
2. If a new point is found within the cutoff distance, it is added, and the track finding process is restarted from that point.
3. The full track is re-scanned twice (i.e. two further iterations).

This iterative augmentation significantly reduced the number of missed spacepoints and prevented the formation of spurious overlapping tracks. The third key change was a retuning of the parameters of the track finding algorithm, which led to fewer missed tracks overall.

Second improvement: exhaustive finder

While the iterative finder was a clear improvement over the adaptive method, it still suffered from two failure modes:

- It would miss tracks that did not extend to large r, since only outer spacepoints were used as seeds.
- It would fail to correctly handle more complicated events, particularly those involving pair production from $\pi^0 \to \gamma \to e^+e^-$, which feature highly curved tracks that meet at a displaced vertex.

Ideally, the tracking algorithm would identify these lepton pair production tracks, and either exclude them or treat them separately. However, the adaptive and iterative finders would fail in this situation. Since the track finders start at high r and move inwards, they tend to find subsets of points which form straight and centrally-pointing false tracks. These false tracks do not represent real particle trajectories, but are still used in the vertex position calculation, often leading to an incorrect vertex.

These flaws became clear during manual inspection using the event display. To address them, I developed a third method: the exhaustive finder, which is illustrated in Figure 5.10. The exhaustive finder proceeds as follows:

1. Precompute the distance between each pair of spacepoints and compare it to a cutoff distance.

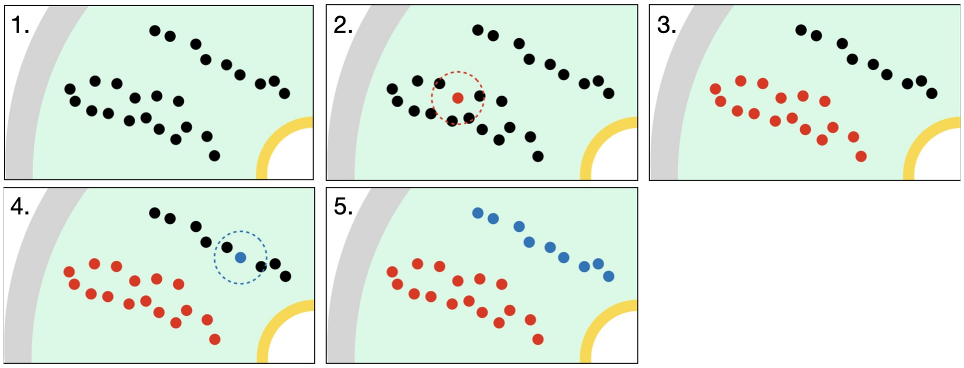

Figure 5.10: Cartoon diagram of the exhaustive finder. (1.) Spacepoints in the TPC. (2.) A seed point is chosen at random. (3.) Points within the cutoff distance are added recursively. (4.) A new seed point is chosen from the remaining points. (5.) This process is repeated until all tracks are found.

2. Randomly choose a seed spacepoint, and recursively add each spacepoint within the cutoff distance to a track.

3. Repeat this with a new seed chosen from the remaining spacepoints, until each spacepoint has been tried as a seed.

This method is less computationally efficient, but the analysis is still far faster than the pad signal deconvolution and matching steps. In exchange for the additional computation, this approach is far more robust. It does not miss tracks which do not approach the anode wires, and does not find false, artificially straight tracks in events with complex topologies.

Summary of track finding improvements

From the initial **adaptive finder**, two improved track finding algorithms were implemented: the **iterative finder** was used for the gravity paper, and the **exhaustive finder** is used in the most recent reconstruction versions. A summary of the three track finding methods is given in Table 5.1.

Both improved track finders were informed by problems observed using the ALPHA-g event display. The iterative finder fixed the issue where many points were missed from a track, frequently resulting in a second (or even third) overlapping track being formed from those points. It also led to fewer missed tracks overall. The exhaustive finder was able to find all possible tracks in the dataset, regardless of their orientation.

Table 5.1: Comparison of track finding algorithms used in ALPHA-g TPC reconstruction. Statistics are approximate and were collected by visual inspection of 50 annihilation events using the event display. Charged pion tracks and other tracks were manually categorized, and the following were counted by hand: missed points in each track, missed tracks, overlapping tracks formed from the spacepoints of a single charged particle, and false radial tracks formed from complicated non-radial trajectories.

Feature	**Adaptive Finder**	**Iterative Finder**	**Exhaustive Finder**
Cutoff distance	$8\,\mathrm{mm} \to 45\,\mathrm{mm}$	$30\,\mathrm{mm}$	$30\,\mathrm{mm}$
Seed point	$r > 168\,\mathrm{mm}$	$r > 168\,\mathrm{mm}$	Any
Method	Single pass	Two refinement iterations	Recursive search
Missed points per track	~ 2.4	~ 0.02	None
Overlapping tracks	$\sim 48\%$	None	None
Missed tracks (pion)	$\sim 38\%$	$\sim 12\%$	None
Missed tracks (other)	$\sim 42\%$	$\sim 28\%$	None
False tracks	$\sim 40\%$	$\sim 21\%$	None
Radial bias	Strong	Strong	None
Computation time per event	$96\,\mu\mathrm{s}$	$139\,\mu\mathrm{s}$	$236\,\mu\mathrm{s}$

The iterative finder was primarily motivated by a desire to improve the annihilation vertex position resolution; we shall see in Section 5.4 that it was successful in this goal. By contrast, the exhaustive finder was motivated by the desire to successfully find all tracks, to extract the most information possible from the detector.

5.3.5 Track fitting

After grouping spacepoints into the tracks, the **track fitting module (7)** fits each track with a helical functional form using independent transverse and axial fits. In the original and gravity paper versions of the reconstruction, these helices were parametrized in terms of the radial variable r. However, this was not compatible with highly curved tracks, which were much more reliably identified by the exhaustive track finder. To address this, the helices were reparameterized, and considerable changes were made to the track fitting logic and initial conditions. This section will first describe the elements common to both the original and current reconstruction versions, then describe the original fitting routine, and finally describe the changes made in the current version.

Helix fit function

The trajectories of charged particles in the ALPHA-g TPC are reconstructed by fitting a helical model to the spacepoints of each track. The mathematical form of the helix is adapted from [146], and represents the expected three-dimensional path of a charged particle in a uniform magnetic field. In the transverse (r–ϕ) plane, the path of the particle traces the arc of a circle due to the Lorentz force from the axial solenoidal magnetic field. In the axial (z) direction, where no forces act on the particle, the motion is linear.

This functional form assumes that the charged particle's energy is constant over its trajectory. The density of the 70:30 $ArCO_2$ gas mixture at atmospheric pressure is $1.7\,\mathrm{kg/m^3}$, and so a minimum-ionizing particle will deposit around $3.4\,\mathrm{keV/cm}$ in the TPC gas. Over a 20 cm path length, which is longer than a typical charged pion path in the TPC, this results in a total energy loss of around 68 keV. Thus the constant energy approximation is valid for the charged pions, which are typically produced in the annihilation with around 100 MeV of kinetic energy.

Mathematically, the helix is described by five free parameters, which are illustrated in Figure 5.11:

- R_c – the radius of curvature in the r–ϕ plane.
- ϕ_0 – the azimuthal angle at the point of closest approach to the origin.
- D – the distance of closest approach to the origin in the r–ϕ plane.
- z_0 – the z coordinate at the point of closest approach.
- λ – the pitch of the helix (slope in z as a function of arc length).

This parametrization takes advantage of the fact that the trajectory is perpendicular to the radius vector at the point of closest approach, meaning the track is uniquely described by (R_c, ϕ_0, D) in the transverse plane. The final two parameters (z_0, λ) describe the z component of the helix.

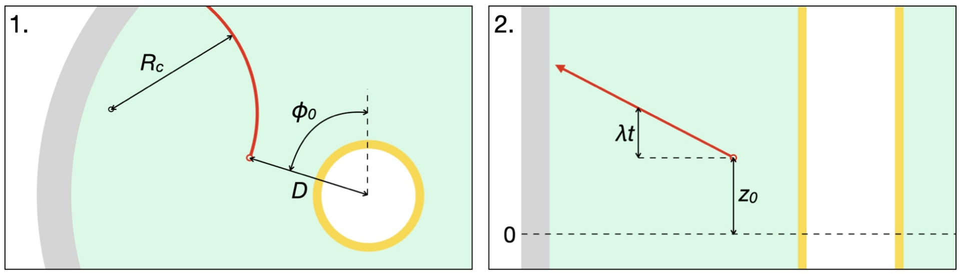

Figure 5.11: The free parameters of the helix functions used in track fitting. (1.) shows the projection of the helix function (red) in the $r-\phi$ plane, illustrating the three radial fit parameters. (2.) shows the projection of the helix function in the $r-z$ plane, illustrating the two axial fit parameters.

The track can be parameterized with respect to a track variable t, which represents arc length from the point of closest approach:

$$\begin{aligned} x(t) &= R_c \sin(t+\phi_0) - (R_c + D)\sin\phi_0 \\ y(t) &= -R_c \cos(t+\phi_0) + (R_c + D)\cos\phi_0 \\ z(t) &= z_0 + \lambda t \end{aligned} \quad (5.3.1)$$

This model was successfully used to fit tracks prior to my contributions. However, the original implementation made a number of further simplifying assumptions, which I will now discuss. I will then outline the changes I made to the track fitting algorithms: first relaxing some of these simplifying assumptions, and then improving the initial fit parameters for more stable track fitting.

Original method: single-valued helix fitting

Given that we are primarily fitting charged pion tracks originating from near the centre of the TPC and propagating outwards, we can impose some constraints on the helix. Firstly, we restrict the track parameter t to be strictly positive or strictly negative, such that the helix extends in only one direction from the point of closest approach. Secondly, we require that the helix does not curve backwards towards the origin. Together, these constraints ensure that the helix is single-valued in the radial coordinate r. This allows us to re-parameterize the helix function in terms of r, as $x(r), y(r), z(r)$.

The helix function is fit to the track in two stages:

- **Transverse ($r-\phi$) fit:** The parameters R_c, ϕ_0, and D are determined by minimizing the transverse residuals:

$$R_T = \sum_i \left[\left(x(r_i) - x_i\right)^2 + \left(y(r_i) - y_i\right)^2 \right] \tag{5.3.2}$$

  where (x_i, y_i) is the i-th spacepoint in the track, and $x(r_i), y(r_i)$ is the helix function evaluated at the radius $r_i = \sqrt{x_i^2 + y_i^2}$ of that spacepoint.

- **Axial (z) fit:** The parameters z_0 and λ are then determined by minimizing the axial residuals:

$$R_Z = \sum_i \left[\left(z(r_i) - z_i\right)^2 \right] \tag{5.3.3}$$

  where z_i is the i-th spacepoint, and $z(r_i)$ is the helix function evaluated at that spacepoint.

This track fit is repeated twice, once with $t > 0$ and once with $t < 0$, and the fit with the lower transverse residual is selected. These correspond to the track extending either clockwise or anticlockwise from the point of closest approach.

Current method: changes to helix fitting routine

The main modification I made to the helix fitting routine was to lift the constraint that the track be single-valued in r. This enables the fitting of helices that curve back on themselves or extend in both directions from the point of closest approach. These are cases that arise frequently in ALPHA-g data for particles other than the charged pions around which the helix fitting procedure was previously designed:

- e^+e^- pairs from neutral pion decay tend to have small momentum, and so are strongly curved in the magnetic field. They frequently produce tracks which curve back on themselves and are not single-valued in r.

- Cosmic ray muons either produce straight line tracks which span both sides of the point of closest approach, or produce two back-to-back tracks on opposite sides of the detector.

To accommodate these cases, I retained the helix parametrization of Equation 5.3.1, but introduced two types of helix:

- A *single-sided* helix which retains the restriction $t > 0$ or $t < 0$, extending in one direction only from the point of closest approach.
- A new *double-sided* helix which has no restriction on t and extends in both directions from the point of closest approach.

This adjustment required a considerable rewriting of the helix fitting framework and routine; since we no longer assume the helix is single-valued in r, we can no longer re-parameterize the helix equations in terms of r.

For the transverse fit, we can no longer evaluate the helix function at the spacepoint radius. Instead, the smallest distance from the spacepoint to the circle defined by (R_c, ϕ_0, D) is found analytically. The transverse residuals are defined as the sum of this distance squared, and the radial parameters are determined by minimizing these residuals:

$$R_T = \sum_i \left[\big(x(t_{\text{closest}}) - x_i\big)^2 + \big(y(t_{\text{closest}}) - y_i\big)^2 \right] \tag{5.3.4}$$

For the axial fit, the residual to be minimized is defined as the difference in z between the spacepoint and the point on the helix that is closest in (x, y):

$$R_Z = \sum_i \left[\big(z(t_{\text{closest}}) - z_i\big)^2 \right] \tag{5.3.5}$$

These changes allow the fitting routine to handle a broader class of tracks more accurately. They also represent a fundamental shift in the tracking philosophy with regards to unwanted tracks from cosmic ray muons and pair production. Previously, these tracks were implicitly excluded by reconstruction algorithms which would fail to find or properly fit them. However, this would often result in mis-reconstructed false tracks which could interfere with the vertex fit. The new philosophy is to successfully find and fit these tracks, and then explicitly exclude them from the vertex fit. This track selection step will be described in Section 5.3.7.

Current method: improved initial fit conditions

Another area where I significantly improved the helix fitting routine is the initialization of the fit parameters. Using the event display to visualize the helices with their initial parameters, I identified and corrected several errors. There were a number of simple mistakes such as a missing factor of π, which caused the initial transverse fit to be placed on the opposite side of the detector, or an omitted minus sign that inverted the initial value of λ. Improving the initial conditions yielded more stable fitting for the gravity paper with fewer failed fits.

In the original reconstruction, the initial conditions for the radial fit were set using the spacepoints with the smallest and largest r. A straight line was drawn between these two points: its angle determined the initial ϕ_0, and its distance of closest approach set the initial D. A very large radius of curvature was used for the initial R_c.

In the current reconstruction, the transverse fit is repeated four times, each with different initial conditions:

1. **r-ordered curved initialization:** The points are ordered by r, and the inner, outer and middle points are found. A circle is drawn through these points and used to define the initial R_c, ϕ_0, and D.

2. **z-ordered curved initialization:** The points are ordered by z, and the top, bottom, and middle points are found. A circle is drawn through these points and used to define the initial R_c, ϕ_0, and D.

3. **Straight positive initialization:** The points are ordered by z, and a straight line is drawn through the top and bottom points. A slight positive curvature is applied.

4. **Straight negative initialization:** The points are ordered by z, and a straight line is drawn through the top and bottom points. A slight negative curvature is applied.

The fit is repeated with each set of initial conditions, and the fit with the lowest χ^2 is used. Each initialization strategy is better suited to a certain type of track; while repeating these calculations is computationally slower, this stage is not a bottleneck, and so the higher number of successfully fitted tracks is worth the marginal increase in analysis time.

For the axial fit, the original reconstruction sorted the points by r, and drew a straight line between the closest point and the middle point to find the initial z_0 and λ. The new method is similar, but takes care to only use spacepoints on one side of the point of closest approach (i.e. either $t < 0$ or $t > 0$). It uses the points with the smallest and largest t value to draw the line $z(t) = \lambda t + z_0$, and uses these λ and t parameters to initialize the axial fit.

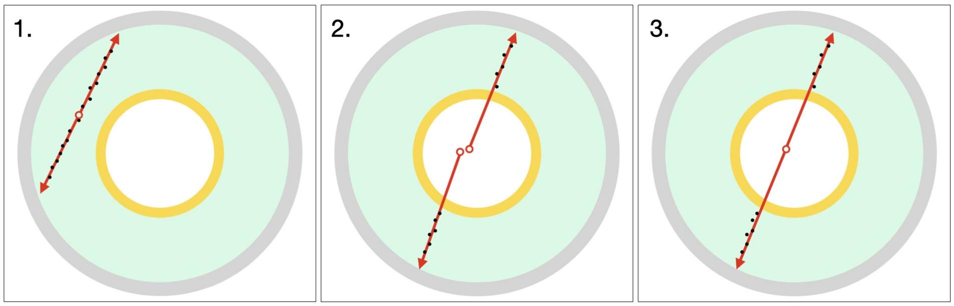

Figure 5.12: Cartoon diagram of cosmic ray tracks in the TPC. (1.) A cosmic ray skims the edge, and is identified as one double-sided track. (2.) A cosmic ray passes through the centre, and is identified as two back-to-back tracks. (3.) The track joining module recombines these into one double-sided track.

5.3.6 Track joining

Although these changes led to improved track finding and fitting behaviour for many types of tracks, there were two further cases which were not addressed properly:

- Cosmic rays passing through the centre of the TPC would produce two back-to-back tracks.
- Delta electrons would create long spiralling tracks which were falsely identified as multiple tracks.

These cases both contribute significantly to the cosmic ray background rate, as they produce enough tracks to form a false annihilation vertex. To address them, I implemented a new reconstruction module called the **track joining module (8)**, with the goal of recombining initially distinct tracks created by the same charged particle. This module proceeds in two steps, each addressing one of the cases above.

Step 1: cosmic ray track joining

Cosmic ray tracks are frequently misidentified as an annihilation with two back-to-back pions. Indeed, these events are the main background in annihilation detection. If these back-to-back tracks could be easily identified as belonging to the same charged particle, then the background rate would immediately be substantially reduced.

When a cosmic ray skims the edge of the TPC, as shown in Figure 5.12 (1.), then the track is fully identified by the exhaustive track finder and can be fit with a single double-sided helix. However, if the cosmic ray passes through the antihydrogen magnetic trap, the spacepoints appear separated into two back-to-back tracks, as shown in Figure 5.12 (2.). Track identification ambiguity can be resolved by combining these into one double-sided track, as shown in Figure 5.12 (3.).

The cosmic ray track joining step attempts to combine each pair of two tracks into a single track. The spacepoint groups of both tracks are merged, and the procedure described in the previous section is used to fit the combined track. The combined helix is used in place of the old ones if the combined goodness of fit is not significantly worse than either of the individual track fit goodness of fits: $R_{T,\text{comb.}} < 2 \cdot R_{T,\text{indiv.}}$ for the transverse fit residuals, and $R_{Z,\text{comb.}} < 1.5 \cdot R_{Z,\text{indiv.}}$ for the axial fit residuals.

Since a set of N tracks has $\frac{1}{2}N(N-1)$ combinatorial pairs, this step is only applied to events with 6 or fewer tracks, which is the case for almost all cosmic ray events. This step was highly successful at combining back-to-back cosmic ray tracks, approximately halving the rate of false annihilation vertices found in cosmic ray events from 13.65 Hz in the original reconstruction to 6.35 Hz in the current reconstruction.

Step 2: cyclotron trapped electron track joining

When a charged particle ionizes the TPC gas, the freed electron will occasionally be released with a large amount of kinetic energy. This *delta electron* is itself an ionizing particle, and travels through the TPC, creating a trail of additional ionizations. Since these electrons have relatively low momentum compared to the initial cosmic ray, they bend more in the external magnetic field, experiencing the type of cyclotron motion discussed in Section 3.2.1. This creates a tight spiral track in the TPC, often with ten or more full rotations over the length of the detector. As they are less massive than pions, these delta electrons transfer less energy per interaction with the gas, and so are more inclined to create atomic excitation rather than ionization. This leads to fewer ionization clusters per unit distance, with larger gaps between ionizations, making these tracks very difficult to identify and reconstruct correctly. They are often fit with a stack of multiple short false tracks on top of each other (see Figure 5.13), which can cause problems for the vertex finding.

These events with delta electrons are rare with an estimated rate of 0.05 Hz, less than 1% of the 6.35 Hz rate of false vertices due to cosmic

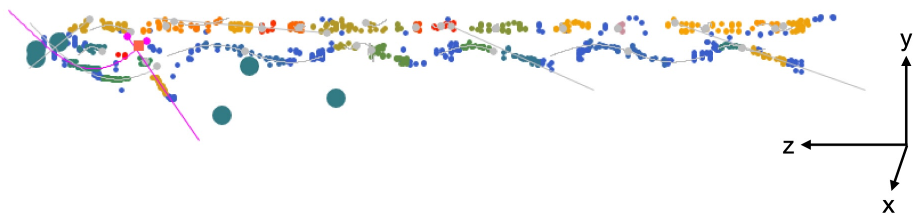

Figure 5.13: The ALPHA-g event display (see Section 5.2) showing an event where a cosmic ray interaction produces two cyclotron-trapped delta electrons which travel the full length of the detector (from left to right here). The colouring of the small dots indicates the grouping of the spacepoints into many distinct tracks. The purple lines are helix fits used in the vertex fit, the grey lines are helix fits excluded from the vertex fit, and the large blue dots are BSC hits (see Section 6.9).

rays. However, they are much more problematic than simple cosmic ray events. A machine learning classification algorithm is used to reject the cosmic ray background, which will be discussed further in Chapter 8. Since these delta electron events are not well represented in the training sample, they are almost always incorrectly classified as annihilation events due to their apparent high number of tracks. These events constituted roughly 40% of the background surviving the machine learning background rejection algorithm.

The second stage of track joining was designed empirically to address these events. It searches for helices fit with a radius of curvature smaller than 25 mm, which was determined using the event display to be larger than a typical delta electron cyclotron radius. If found, it looks for further tracks which have 90% of their spacepoints contained in a vertical column around the centre of the small track. If at least five further tracks are found, three of which have a radius of curvature smaller than 25 mm, then the tracks are joined. The spacepoints are combined into a single track, which is refit with a helix function according to the previous section. This procedure was developed such that it is almost never triggered by annihilation events, while reasonably identifying and fitting cyclotron trapped electrons.

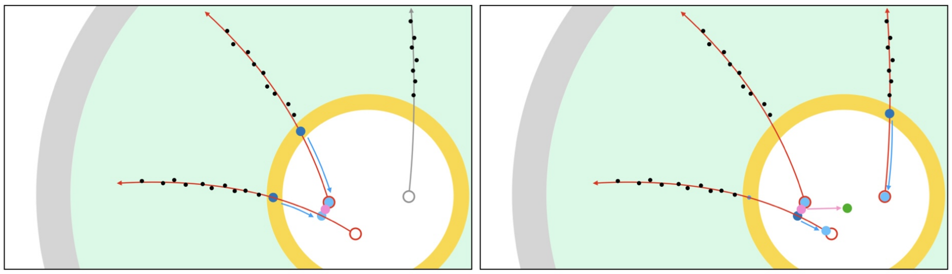

Figure 5.14: Cartoon diagram of the original vertex fit method. The left panel shows the vertex seeding from the closest two helices. A point on each vertex initialized at the trap radius (dark blue), and then allowed to move along the helix to minimize the distance between them (light blue). The seed vertex (pink) is the point halfway between them. The right panel shows the step where additional helices are added in turn. The new vertex (green) is chosen to minimize its distance from all helices.

5.3.7 Vertex fitting

The final reconstruction stage is to determine the annihilation vertex position from the set of fitted helices, performed by the **vertex fitting module (9)**. To accommodate the reparameterization of the fit helices, this stage was also heavily modified between the original and current versions of the reconstruction. This section will compare both reconstruction versions, which are illustrated in Figures 5.14 and 5.15.

For both versions, the vertex calculation proceeds in four steps:

1. A set of selection criteria are applied to the helices, so that only a subset of the helices are used to find the vertex.
2. A seed vertex position is determined analytically from two helices.
3. The seed vertex is refined by a fit using only those two helices.
4. Additional helices are added in turn, and the vertex is refit.

Helix selection

The vertex fitting procedure is designed to recover the common origin of a number of near-straight pion tracks. However, a number of particles such as delta electrons (Section 5.3.6) or pair produced electrons and positrons

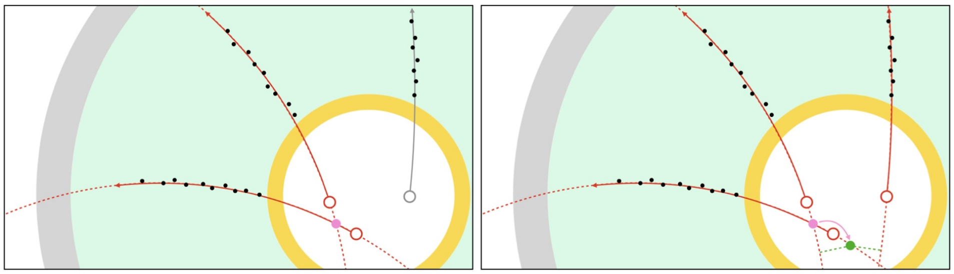

Figure 5.15: Cartoon diagram of the current vertex fit method. The left panel shows the vertex seeding from the closest two helices. The circles traced by the helix projection into the $r-\phi$ plane are used. The seed vertex (pink) is the point closest to both circles, found analytically. The right panel shows the step where further helices are added. The distance squared from each circle to the new vertex (green) is minimized.

(Section 5.3.4) often lead to tracks that originate at a secondary vertex. Furthermore, the track finding and fitting process often yields poorly fit or false helices that do not match a real particle trajectory. To properly determine the vertex position, these tracks must be excluded from the vertex fit.

To do this, only helices passing certain selection criteria are included in the vertex fit:

- Helix fits which did not converge are discarded.
- Cuts are applied on the goodness of fit of both the transverse and axial fits. These are the residuals R_T and R_Z defined in the previous section, divided by the number of degrees of freedom N_{dof}.
- A cut on the minimum number of spacepoints in a track discards poorly constrained helices.
- A cut on the distance of closest approach of the helix discards those which do not pass near the trap, the site of all annihilations.
- A cut on the radius of curvature excludes charged particles with low momentum, such as delta electrons or e^-e^+ pairs, which originate from a displaced vertex.

My changes improved the reconstruction's ability to identify and fit tracks with smaller radius of curvature, and so accurately discarding them became far more important. With these changes, I undertook a campaign to optimize the helix selection criteria for the best vertex position resolution, and for the maximum number of successfully fit vertices. The updated selection criteria are given in Table 5.2.

Table 5.2: Helix selection criteria before and after reconstruction updates.

Selection Criterion	**Original**	**Current**
Transverse goodness of fit R_T/N_{dof}	None	< 60
Axial goodness of fit R_Z/N_{dof}	None	< 110
Number of spacepoints	> 7	> 5
Distance of closest approach	< 110 mm	< 80 mm
Radius of curvature	None	> 50 mm
Fit convergence required	Yes	Yes

Vertex seeding

After discarding the helices not meeting the selection criteria, a *seed vertex* is initialized using the two helices which pass closest to each other. Originally, each pair of helices was used in the following manner:

1. A point along each helix (i.e. a t value) is initialized at the trap radius such that $|\vec{r}_i(t_i)| = r_{\mathrm{trap}}$.
2. The t values are varied to minimize the three-dimensional distance $d = |\vec{r}_1(t_1) - \vec{r}_2(t_2)|$ between the two points.
3. This is repeated for each pair of helices, and the pair with the smallest such distance d is chosen.
4. The seed vertex is set to the point $\vec{s} = \frac{1}{2}[\vec{r}_1(t_1)+\vec{r}_2(t_2)]$ halfway between the two helices.

While this method is possible with my modifications to the helix fits, I took the opportunity to remove this minimization step. Instead, the point of closest approach of each helix to the z axis (the axis of the TPC) is found. The two helices whose points of closest approach are closest to each other are used to seed the vertex. Since each helix traces a circle in the $r - \phi$ plane, the closest point between the two circles can be found analytically and is used to seed the vertex.

Initial vertex fit

The vertex fitting proceeds with an initial fit, using only the two helices which were identified in the vertex seeding step. In the original analysis, a five-parameter minimization is performed: as well as the t value along each helix, the vertex position is also varied. The minimization finds a new vertex position with a minimum distance to the two helices.

In the revised analysis, only the three dimensions of the vertex position are varied. The closest point to the vertex on the circle traced by each helix in the $r-\phi$ plane is found analytically, and the sum of their distances squared to the candidate vertex is minimized. If this distance squared exceeds a fixed threshold, then the vertex fit is aborted and no vertex is found for the event; otherwise the seed vertex position is updated.

Adding further helices

Finally, an attempt is made to incorporate each other helix into the vertex fit. The helices are tried one at a time, and the fit is seeded from the current vertex position. The fitting procedure is identical to the one described in the previous section. In the original analysis, the dimension of the fit is increased by one with each added helix, due to the additional t value for the new helix. In the new analysis, the dimension of the fit is unchanged, since only the vertex position is varied.

After each helix is added, if the total squared distance exceeds the fixed threshold (the same as above), the helix is removed again. Otherwise, the helix is kept and the vertex is updated. In this way, only helices which are compatible with the existing seed vertex are included in the fit.

5.4 Performance improvement results

To benchmark the vertex position resolution of various algorithms, antiprotons are held in a narrow and deep potential well in the ALPHA-g Penning-Malmberg trap. They are held for 2000 s, during which they annihilate on background gas atoms due to the imperfect vacuum.[6] This provides a close approximation of an annihilation point source.

[6] The primary use of this procedure is to measure the quality of the vacuum. Remaining antiprotons are counted by scintillator panels after the 2000 s wait; the number of annihilations is compared with and without the wait to determine the number of antiprotons lost to background gas collisions. This measurement is repeated regularly.

Feature	**Original**	**Gravity paper**	**Improved**
Track finder	Adaptive	Iterative	Exhaustive
Helix initialization	Bugged	Fixed	Improved
Helix selection optimized	·	✓	✓
Helices reparameterized	·	·	✓
Double-sided helices	·	·	✓
Track joiner module	·	·	✓
Vertex fit modified	·	·	✓

Table 5.3: Summary of the improvements made between the three reconstruction version: the original pre-2022 reconstruction, the version used for the gravity paper, and the current improved version. For more information on each item, see Section 5.3.

The potential well is produced by an electrode held at +150 V between two electrodes held at −150 V, where the length of the electrodes is approximately 21 mm. This creates an on-axis potential well $V(z) = -kz^2$, where the well constant is $k \sim 0.35\,\mathrm{V/mm^2}$. An antiproton with a temperature of 1000 K, higher than typical antiproton temperatures used for this measurement, has a kinetic energy of 0.08 eV. This sets a bound on the axial length of the antiproton plasma of around 1 mm in the quadratic well.

This point source is used to quantify the resolution of the many versions of the reconstruction. We will examine three: the original pre-2022 version of the reconstruction, the one used for the gravity paper, and the current improved reconstruction. The improvements made between the versions are summarized in Table 5.3 and described fully in Section 5.3.

The annihilation z position distributions for each of these reconstruction versions are shown in Figure 5.16. A double Gaussian function is fit to each distribution. The integrals of the fit functions are used to determine the amount of data in each population. The detector vertex resolution is defined by the combination of the two σ values of the inner and outer Gaussian fits; it is most insightful to define an *inner resolution* and *outer resolution*.

The use of a double Gaussian reflects two populations of events. Around 60% are simple events with vertices which can be reconstructed easily and precisely from two or more unscattered charged pion tracks which can be traced back to the annihilation vertex. The remaining 40% are those events without two charged pion tracks which can be traced back to the annihilation vertex. This can be either because fewer than two charged pions were

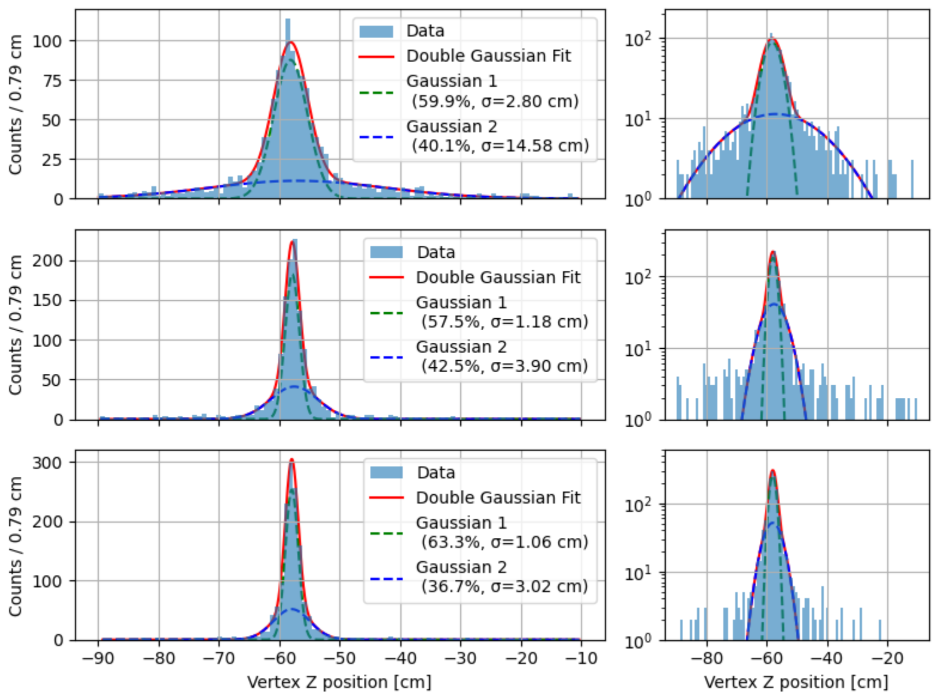

Figure 5.16: Vertex position resolution across three reconstruction versions: the original pre-2022 reconstruction (top), the reconstruction used for the gravity paper (centre), and the current improved reconstruction (bottom). Shown with a linear (left) and logarithmic (right) y-axis. Data shown is from antiprotons annihilating on background gas. The most recent online background rejection cut (see Section 6.7) is used; the same events are included in all three plots.

produced (around 18% of $\overline{p}n$ annihilations and 4% of $\overline{p}p$ annihilations, see Table 3.1), because one or more pions were scattered or absorbed on the internal material of the magnet system, or because reconstruction of one or more charged pion tracks failed. In these events, other tracks from e^-e^+ pairs or scattered pions must be used, which do not originate exactly at the annihilation vertex. As a result, the vertex determination is less precise for these events.

There is a clear improvement in resolution from the original reconstruction to the gravity paper version. This was in large part due to the introduction of the iterative track finder, which improved the overall resolution by roughly a factor of two. Further refinements were made to the helix fit initialization and helix selection, as well as contributions from other collaborators on other aspects of the reconstruction. This ultimately improved the inner resolution from 2.80 cm to 1.18 cm, and the outer resolution from 14.58 cm to 3.90 cm.

The gravity paper requires the localization of annihilations into one of two regions separated by 25.6 cm. The improvements presented here ensured that annihilation mis-classification to the wrong region was not a major experimental systematic; this will be explored further in Section 9.4.1. The improvements to the outer resolution are the most relevant, as the outer resolution describes the events most likely to be misclassified.

The resolution improvement between the gravity paper reconstruction to the current reconstruction is significant but less pronounced: the inner resolution was improved from 1.18 cm to 1.06 cm, and the outer resolution from 3.90 cm to 3.02 cm. This reflects a shift in the main purpose behind the reconstruction changes: while the gravity paper reconstruction was the result of a focused effort to lower the detector vertex resolution, the changes leading to the current version were aimed to create a more physically-motivated reconstruction procedure.

To illustrate this, a number of intermediate metrics are compared across the three reconstruction versions in Table 5.4. Some limitations of the gravity paper reconstruction are visible here: the track selection criteria do not reject any tracks, and each vertex is fit using only two helices. The improved reconstruction fixes these issues, returning to the original intention of incorporating as many useful tracks as possible into the vertex fit, explicitly excluding those which are not useful.

It must be noted that with these improvements, the percentage of annihilation vertices successfully fit decreases from 86% to 67%. Most of those vertices lost were those which were more difficult to constrain; as a result of not reconstructing these vertices, the position resolution of the detector is improved. This was an active choice in the redesign of the reconstruction algorithms. However, should statistics become more important than position resolution in a future experiment, the parameters of the reconstruction can be easily tweaked to recover some of these vertices.

My efforts have led to significant improvements in the detector vertex position resolution, ensuring that the detector resolution did not constitute a major systematic uncertainty to the experiment. In addition, this work has made a tangible improvement to the rate with which cosmic ray background events are mis-identified as annihilation vertices. Furthermore, it has led the reconstructed tracks to match physical particle trajectories more accurately, improving the foundation for studies of ALPHA-g data extending beyond the annihilation z position.

5.5 Future work: machine learning vertex finding

Alongside the algorithmic approach to vertex finding described here, I have also contributed to the development of a second vertex reconstruction method using machine learning. This method follows the analysis chain up until the spacepoint reconstruction stage. However, instead of using helix fitting to determine the annihilation vertex, a neural network is trained to predict the vertex position directly from the spacepoints. My primary contribution to this effort has been the design and implementation of a data pipeline to convert raw detector data and simulation output into a dense data structure suitable for machine learning training.

The initial results of this effort are published in a recent preprint [147]. An architecture called *point net ensemble* is used for the regression. This is based on the point net model which is commonly used for object classification from a three-dimensional point cloud; our novel architecture is trained for regression instead of classification, and combines an ensemble of similar models for enhanced accuracy.

The project, called the *point net ensemble for annihilation reconstruction* (PEAR), has successfully learned the z position of the vertex for simulated annihilations. This method already demonstrates more precise annihilation vertex positions with a lower fail rate compared to the *helix fit method* described in this chapter. The z position resolution on simulation[7] is compared

[7]It should be emphasized that this method was trained and benchmarked using simulated TPC data; a discussion of the TPC simulation is outside the scope of this dissertation. The helix fit method yields an improved vertex resolution when applied to simulation compared to real data, the reason for which is still under investigation.

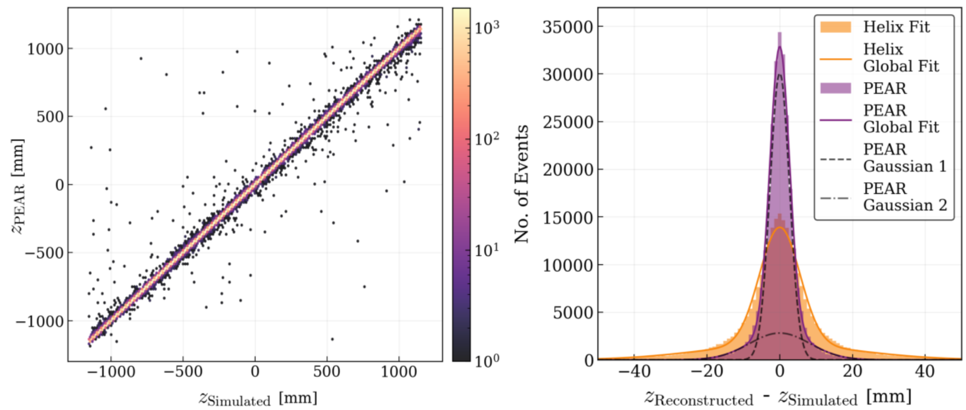

Figure 5.17: Performance of machine learning PEAR vertex finder, compared to the helix fit method of this chapter. Left: Simulated vs PEAR annihilation z position. Right: Residuals for PEAR and helix fit method, fit with a double Gaussian. Taken from [147].

to that of the helix fit method in Figure 5.17, showing a full-with at half-maximum (FWHM) of 6.6 mm compared to the 14 mm FWHM of the helix fit method. If the reconstruction of the vertex (x, y) position is equally successful, and if PEAR performs equally well on real data as on simulation, then the machine-learning based vertex finder should be implemented alongside the algorithmic approach in the upcoming years.

Table 5.4: Comparison of analysis performance across historical snapshots at three times: the original pre-2022 analysis, the analysis used for the gravity paper, and the current improved analysis. Values are averaged over 2 hours of cosmic ray ($\mu^{\pm}$) data and 50 1 s antihydrogen ($\overline{H}$) formation periods (see Section 3.5.3). *Spacepoints* indicates the average number of spacepoints in an event, and *(in tracks)* indicates how many of these were use by the track finding algorithm. *Tracks* indicates the average number of tracks found, *(fit successfully)* indicates how many of these had helix fits which converged, and *(pass selection)* indicates how many of these helices passed the selection criteria for use in vertex fitting. R_T/N_{dof} and R_Z/N_{dof} indicate the average transverse and axial goodness of fit for each track. *% vertices fit* indicates the percent of events for which a vertex position was successfully found. *Candidate helices* indicates the average number of helices available for use in successfully fit vertices, and *used helices* indicates how many of those were actually included in the vertex fit.

Annihilations			
Metric	**Original**	**Gravity paper**	**Improved**
Spacepoints	327.1	297.7	298.4
(in tracks)	99.1	115.5	211.4
Tracks	5.05	3.40	5.55
(fit successfully)	4.30	3.06	5.39
(pass selection)	4.30	3.06	2.35
R_T/N_{dof}	9.41	9.09	4.75
R_Z/N_{dof}	57.18	11.63	7.64
% vertices fit	85.77	85.81	66.96
Candidate helices	4.87	3.43	2.96
Used helices	4.87	2.00	2.62

Cosmic rays			
Metric	**Original**	**Gravity paper**	**Improved**
Spacepoints	196.5	180.8	182.3
(in tracks)	54.7	41.9	137.2
Tracks	2.65	1.48	1.86
(fit successfully)	1.87	0.83	1.49
(pass selection)	1.87	0.83	0.60
R_T/N_{dof}	3.98	3.80	3.34
R_Z/N_{dof}	55.51	8.51	6.43
% vertices fit	55.09	31.98	14.9
Candidate helices	3.23	2.06	2.01
Used helices	3.23	2.00	2.01

Chapter 6

Barrel scintillator commissioning and data processing

The primary project of this dissertation is the commissioning of the barrel scintillator (BSC) detector which is described in Section 4.3. The BSC and associated detector electronics were developed at TRIUMF by other collaborators prior to the start of this work. However, minimal testing had been performed beyond a precursory inspection of the waveforms recorded by the DAQ.

Over the course of this dissertation work, I was responsible for the commissioning of the BSC hardware, as well as the implementation of the low-level data processing routine analogous to the TPC reconstruction algorithm described in the previous chapter. These two tasks were strongly intertwined, and were performed simultaneously; as the data processing was implemented, this often revealed underlying hardware problems, and vice versa. This work ultimately resulted in the successful implementation of time-of-flight measurement using the BSC, which was used to significantly enhance the rejection of the cosmic ray background (discussed in Chapter 8).

The hardware commissioning included a variety of tasks such as:

1. Diagnosing and solving problems leading to dead or poorly-performing channels for the ADC and TDC digitizers.

2. Replacing malfunctioning detector electronics components.

3. Determining optimal values for the SiPM bias voltage, ADC and TDC digitizer threshold voltages, and TDC timing window.

4. Evaluating the in-situ temperature dependence of the SiPMs.

5. Evaluating the z position resolution along each bar using cosmic ray coincidence with two external scintillator panels.

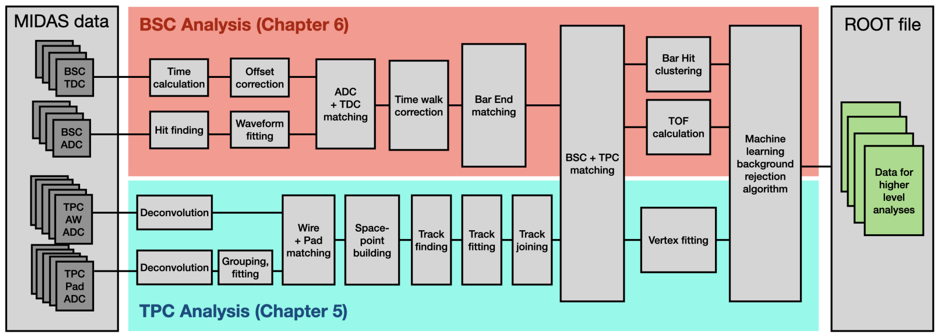

Figure 6.1: The stages of analysis for the BSC and TPC data.

The BSC data processing and event reconstruction is performed in the same *alphasoft* framework as the TPC analysis. BSC reconstruction modules are interleaved with the TPC modules, as the two analyses are interdependent. However, while the TPC analysis was in place and functional before my revisions, the BSC analysis was entirely my own design and implementation.

There are two data sources corresponding to each scintillation event in a bar of the BSC: the waveforms recorded by the ADC, and hit time data from the TDC. The data acquisition system and MIDAS event builder described in Section 4.4 groups these data into events based on their timestamp.

The data processing and event reconstruction procedure is illustrated in Figure 6.1. Each step is handled by an individual module in the *alphasoft* analysis. The data processing begins by extracting *ADC hits* from the ADC waveforms (Section 6.1), and *TDC hits* from the TDC data packets (Section 6.2). These hits are then matched together to form *end hits*, corresponding to a single SiPM-array signal with both waveform and timing information (Section 6.3). Pairs of end hits from opposite ends of the same bar are combined into *bar hits*, which contain all the information from a single scintillation event (Section 6.4), including a calculated scintillation z position along the bar (Section 6.5). These bar hits are then matched to corresponding charged particle tracks in the TPC (Section 6.6). Finally, bar hits are grouped into clusters (Section 6.7), and the time-of-flight (TOF) between pairs of bar hits is calculated (Section 6.8). These data structures are saved in the output ROOT file alongside the vertex information from the TPC, for use in higher level physics analyses.

This section will further discuss the integration of the BSC into the event display (Section 6.9). A description of the online quality of data checks for the BSC will be reserved for Appendix B, and an investigation into pileup events using the BSC is given in Appendix C.

This section will present the data processing for the BSC detector in considerable detail. If the reader is not excited to learn about the signal processing and hit matching algorithms employed to extract hit bar clusters and time-of-flight measurements from the the BSC data, they should feel free to skip this chapter entirely without loss of continuity.

6.1 ADC waveform signal processing

The first task in BSC event reconstruction is to examine each ADC waveform to determine whether a hit occurred and, if so, to extract the amplitude of the SiPM signal. A hit finding algorithm is used which does not rely on any pre-programmed threshold value, instead analyzing the electronic noise and continually updating hit thresholds for each channel. To address ADC saturation which occasionally produces truncated waveforms, a waveform fitting procedure is implemented. Examples of fitted ADC waveforms are shown in Figures 6.2 and 6.3.

6.1.1 Hit finding

Unlike the TPC anode wire or cathode pad signals, it is rare for multiple hits to occur within a single ADC waveform. Since the width of one bin in the ADC waveform is 10 ns, much larger than the charged particle travel time across the BSC, it is not possible to resolve two pions from the same annihilation striking the same bar. As a result, a waveform with two distinct signals only occurs via the coincidence of multiple cosmic rays or antihydrogen annihilations within the 7 µs ADC window. Since this will still occur, albeit rarely, our waveform analysis should be designed to handle this case.

The analysis begins with baseline subtraction. The first 100 samples (1 µs) of the waveform are before the trigger, and so are assumed to contain no signal. The baseline is estimated as the mean of the first 100 samples, and subtracted from all samples.[8]

The simplest technique to find a signal in the ADC waveform, which was initially tested, uses a predefined threshold level. A hit is declared when the

[8]After the end of the SiPM signal, once it returns below a programmable data suppression threshold, the ADC firmware does not send any further samples to MIDAS in an effort to lower the data rate. The MIDAS event builder populates the remainder of the waveform with a constant value based on a similar baseline calculation.

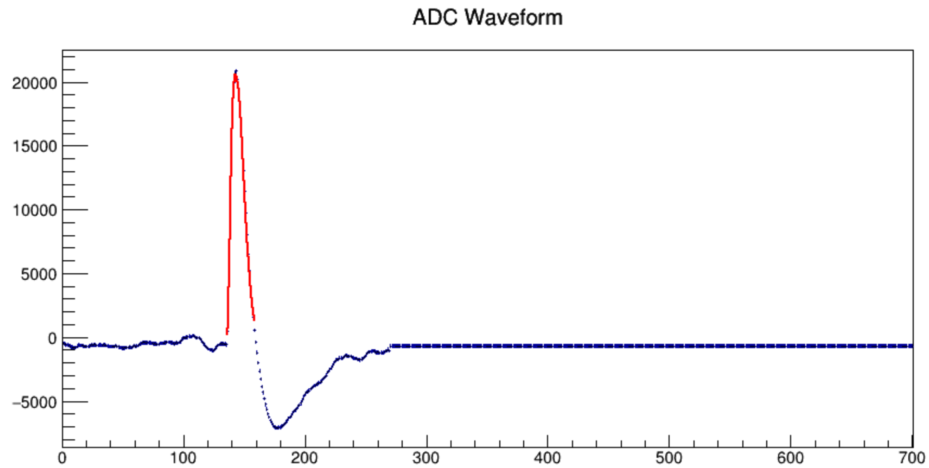

Figure 6.2: An unsaturated ADC waveform. The x-axis is time in 10 ns bins, and the y-axis is voltage in ADC units. The red indicates a fit with the skewed Gaussian of Equation 6.1.1. Data suppression at the DAQ level has zeroed all samples after 280.

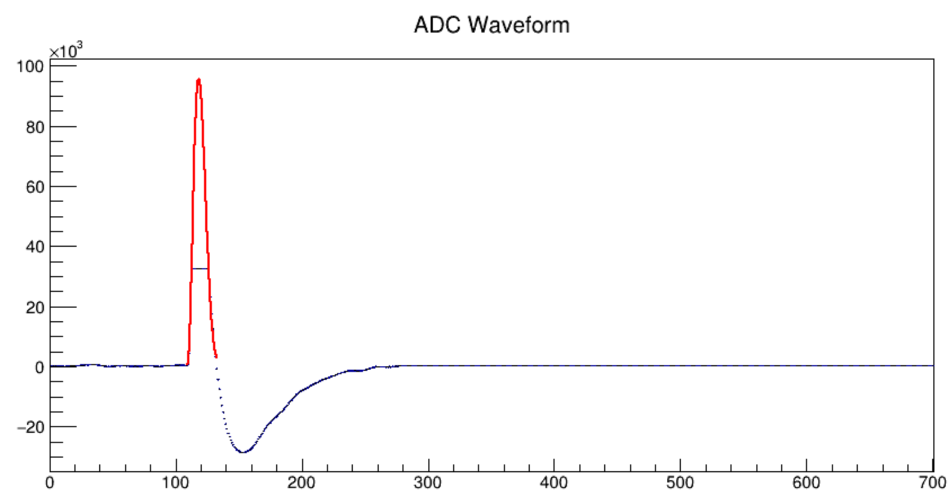

Figure 6.3: A saturated ADC waveform, fit with the skewed Gaussian from Equation 6.1.1. The waveform saturation is visible as the data (blue) flattening at $y = 32768$.

waveform crosses the threshold, and ended when it falls back below. The thresholds were set on a per-channel basis, and regularly recalculated with threshold scans. However, this approach invariably led to missed hits: events where the SiPM signal would trigger the TDC comparator, but would not exceed the predefined ADC threshold.

To find as many ADC hits as possible, a dynamic hit finding algorithm was adapted from the ALPHA-2 SVD calibration approach, where per-channel hit thresholds are recomputed for each run in an ADC calibration routine:

1. The electronic noise level of each waveform is estimated by the standard deviation σ of the first 100 samples (the samples also used to compute the baseline).

2. A *level-one* threshold is set at $+5\sigma$. Hits are saved whenever the waveform crosses this threshold. This threshold is recomputed for every waveform.

3. The amplitude of the hit is defined as the maximum baseline-subtracted value within the hit window, i.e. while the waveform stays above the level-one threshold.

4. A *level-two* threshold for each channel is read from a precalculated ADC calibration file. If the amplitude of the hit is above this threshold, the hit is marked as *good for analysis*.

The computation of the level-two thresholds is handled by a separate ADC calibration routine, which is combined with the TDC calibration routine which will be described in Chapter 7. This calibration routine is performed automatically at the end of each run. In online analysis, the calibration file from the last run is used; otherwise, the calibration file from the run of interest can be used.

The ADC calibration routine takes advantage of electronic noise crossing the level-one threshold:

1. A histogram of hit amplitude is built for each channel, as shown in Figure 6.4.

2. The hits due to electronic noise, those with low amplitude, are fit with a Gaussian.

3. The level-two threshold is set to $\mu + 5\sigma$ using the Gaussian fit parameters.

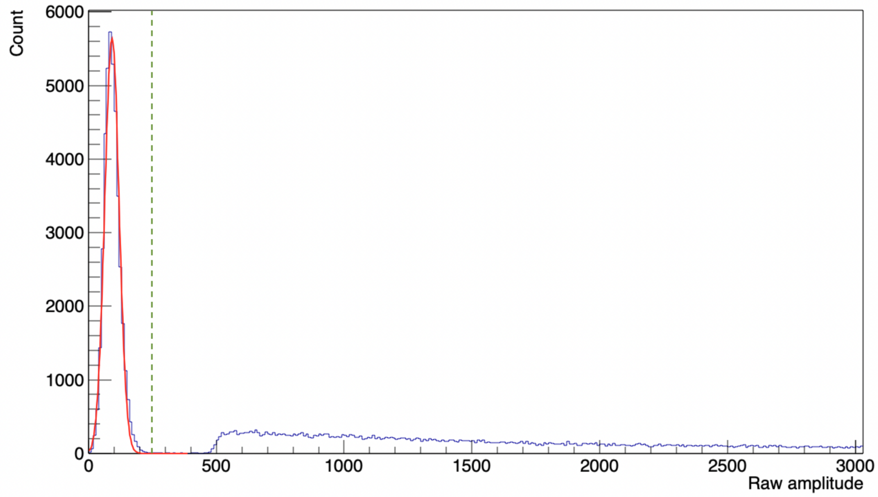

Figure 6.4: The baseline-subtracted amplitude of all ADC hits. The electronic noise portion is fit with a Gaussian, and only hits with amplitude greater than $\mu + 5\sigma$ (the *level-two* threshold indicated by the dashed green line) are used in the analysis.

In this way, the hit thresholds are dynamically adjusted after each run. Accurately setting these thresholds is important to avoid missing the smallest SiPM signals, typically those from scintillations far from the bar end, where most of the scintillation photons have been attenuated.

6.1.2 Waveform fitting

The primary motivation for full digitization of the SiPM waveforms is to extract their amplitude. This is useful for estimating the number of scintillation photons produced, as well as for correcting the time walk effect on TDC time measurements, which will be covered in Chapter 7. For most ADC hits, it is adequate to use the simple amplitude defined as the maximum minus the baseline.

However, a complication is presented by waveform saturation. The ADC has a fixed dynamic range, with a maximum voltage of 1 V corresponding to a maximum ADC value of 32768. This maximum was chosen to exceed most SiPM signals sizes; the 16-bit values represent a trade-off between digitizer cost, data size, and precision. When the true signal exceeds this threshold, such as when a scintillation occurs close to the SiPM, the waveform is truncated at the maximum value, producing a flat top. In this case, the size of the SiPM signal is no longer measurable as the maximum minus the baseline.

There are many approaches to recover an approximate signal amplitude. A common method is to find the length of the flat top; assuming all signals share the same general shape, this can be used to calculate the signal height. Instead, the approach chosen for the BSC was to fit the waveform. A skewed Gaussian function was chosen empirically as it gives a good fit to the waveforms. The fit function is similar to a Gaussian, with an additional parameter k which introduces an asymmetry:

$$f(x) = \begin{cases} A \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} & x > \mu \\ A \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma - k(\mu - x)}\right)^2} & x < \mu \end{cases} \tag{6.1.1}$$

The fit is applied to the waveform within the hit, i.e. where it is above the level-one threshold. Bins are weighted uniformly, except for saturated bins, which are assigned zero weight to exclude them from the fit. The amplitude of the waveform is then defined as the fit parameter A (after baseline subtraction).

This procedure was validated on both saturated and unsaturated waveforms. Figure 6.2 shows an unsaturated waveform fit, while Figure 6.3 illustrates the recovery of an amplitude from a saturated waveform. In typical operation, only the saturated waveforms are fit.

In future work, alternative methods for recovering the amplitude from saturated waveforms could be explored. If the waveforms are shown to have a sufficiently uniform shape, a template function could be found. This could either be used to fit the waveforms directly, or used to calculate the relationship between the amplitude and the length of the flat top. However, this is not a priority; the primary use for the hit amplitude is in correcting for the time walk effect, which is strongest for small signals which do not require fitting.

6.2 TDC event extraction

The TRB3 TDC records time information using three counters with integer values, described in the TRB3 manual [139]:

- The **coarse time counter** tracks the number of 200 MHz clock cycles. It is stored in 11 bits, corresponding to a range of 10.24 µs.

- The 28-bit **epoch counter** increments each time the coarse counter overflows, extending the time range to approximately 45 minutes.
- The **fine time counter** resolves time within a single 5 ns clock cycle. It indicates the number of elements traversed in the FPGA delay line.

There are two methods to calculate precise hit times from these counters. The first uses a fine time linearity approximation, while the second requires precision calibration.

6.2.1 Fine time using linearity approximation

Recall from the discussion in Section 4.4.2 that TDC signals propagate through a series of FGPA delay elements; the number of delay elements traversed by a signal at the start of the next clock cycle is recorded in the fine time counter. The first method to extract the fine time uses the approximation that each FPGA delay element in the TDC has the same signal propagation time. This results in a fine counter which increases linearly with time. The overall hit time can be calculated as:

$$\begin{aligned} t_{\mathrm{TDC}} &= t_{\mathrm{overflow}} + t_{\mathrm{epoch}} + t_{\mathrm{coarse}} + t_{\mathrm{fine}} \\ &= \Big[\big(2^{28}\cdot 2^{11}\cdot N_{\mathrm{overflow}}\big)+\big(2^{11}\cdot c_{\mathrm{epoch}}\big)+\big(c_{\mathrm{coarse}}\big)+\big(\frac{c_{\mathrm{fine}}-m}{M-m}\big)\Big]\cdot 5\mathrm{ns} \end{aligned} \tag{6.2.1}$$

Here, N_{overflow} is the number of epoch counter resets, which must be tracked by the data processing module for runs longer than 45 minutes. c_{epoch}, c_{coarse}, and c_{fine} are the respective counter values, and m and M are the minimum and maximum fine time counter values. These limits must be determined via calibration for each channel.

This method is simple to implement, and correct to within 100 ps, with typical systematic errors of around 20 ps on most channels. This is an acceptable precision for TOF measurements in the BSC; as we shall see in Section 6.8, the final TOF resolution is around 325 ps, and so using the linear approximation does not limit the TOF resolution. This method is available as an alternative in cases where the more precise calibration is not available.

6.2.2 Fine time using delay element calibration

For the highest precision TDC timing, we instead approximate the true propagation time of the individual delay elements. This is done statistically, using

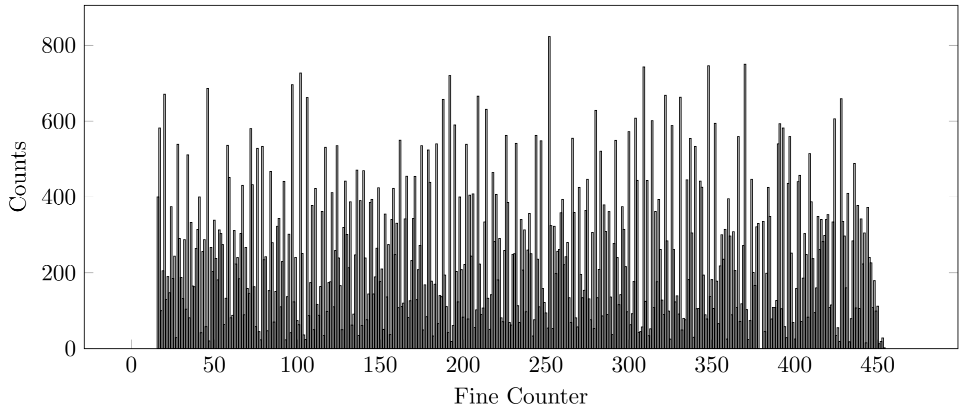

Figure 6.5: Histogram of fine time counter values for one TDC channel, used to determine propagation times for each delay element. Credit: Daniel Duque.

a large sample of randomly distributed data such as SiPM signals from cosmic rays. When the clock signal arrives and the delay line is checked, the probability of the signal currently propagating through a given delay element is proportional to the propagation time of that delay element. A histogram of the fine time counter values can then be used to determine the time of each delay element, shown in Figure 6.5. The propagation time Δt_i of each element is computed from the counts c_i in the corresponding bin as:

$$\Delta t_i = \frac{c_i}{\sum_j c_j} \cdot 5\text{ns} \tag{6.2.2}$$

The propagation times for each delay element for each channel are saved in a lookup table. During event reconstruction, the final t_{fine} term in Equation 6.2.1 is replaced by a fine time calculated from the lookup table Δt_i values of all preceding delay elements:

$$t_{\text{fine}}(c_{\text{fine}}) = \sum_{i<c_{\text{fine}}} \Delta t_i \tag{6.2.3}$$

This calibration is repeated approximately yearly to account for drift in delay element behaviour or degradation of the TDC. A standalone calibration algorithm which populates the fine counter histogram and generates the lookup table was written by PhD student Daniel Duque.

6.2.3 TDC time corrections

After calculating the TDC hit time, two additional corrections are applied to determine the accurate arrival time of the first photon at the SiPM. These corrections rely on careful calibration; Chapter 7 will describe the TDC calibration algorithms used to determine their values.

The first correction is a constant, channel-dependent offset. These time delays could be caused by variations in signal path length, although in our system they are dominated by internal delays within the TRB3 TDC hardware itself.

The second is an amplitude-dependent correction known as the time walk correction. Recall that the digital signal fed to the TDC is generated when the SiPM voltage crosses a fixed reference voltage. For two pulses of identical shape but different amplitudes, the larger amplitude pulse will cross the threshold first, introducing a timing bias known as time walk.

Consider a pulse with a quadratic rising edge, which is a good approximation for our SiPM signals: $f(t) = A \cdot (t - t_0)^2$. We can analytically find the time t_1 where this rising edge crosses a constant threshold value B, in order to recover the true pulse start time t_0:

$$\begin{aligned} f(t_1) &= A \cdot (t_1 - t_0)^2 = B \\ \implies t_0 &= t_1 - \sqrt{\frac{B}{A}} \end{aligned} \tag{6.2.4}$$

This is the primary reason for full ADC waveform digitization: we apply a correction for this effect based on the amplitude of the signal measured by the ADC. This correction has the form $\Delta t = \sqrt{B/A}$, where A is the signal amplitude, and B is the threshold value. The threshold values for each channel are determined using a calibration procedure which will be described in Chapter 7.

Including both the channel dependent time offset δ and the time walk correction $\sqrt{B/A}$, the final hit time used for reconstruction is derived from the hit time reported by the TDC as:

$$t = t_{\text{TDC}} + \delta - \sqrt{\frac{B}{A}} \tag{6.2.5}$$

Both of these corrections are on the order of ns, comparable to or larger than the measured TOF values themselves. As a result, they are crucial to the successful operation of the BSC as a timing detector. These time corrections will be explored further in Chapter 7.

6.2.4 Cable reflections

The transmission boards which receive the TDC signals (see Appendix A) include a design flaw where the differential signal lines are improperly terminated. This causes a reflection of the signal, which generates repeated TDC hits approximately 165 ns after the initial signal. Although there was discussion on modifying the transmission boards with proper termination, ultimately it was decided to address these cable reflections in software.

A final stage of TDC data processing identifies and removes these reflected hits. Any TDC hit occurring between 160 ns and 170 ns after a previous TDC hit on the same channel is marked as a reflection, and excluded from further analysis.

6.3 ADC-TDC matching

The next stage of the event reconstruction aims to match ADC hits with TDC hits. A pair consisting of one ADC hit and one TDC hit is combined into a bar *end hit*, which contains complete information from a single SiPM signal on one end of a bar. This ADC-TDC matching is performed before applying the time walk correction described in the previous section.

Since the data acquisition system compiles events based on timestamp, ADC and TDC hits can usually be matched based on channel number alone, using a known channel mapping. However, to handle cases where multiple hits occur on the same channel within a single readout window, a timing-based matching is also performed.

The ADC hit time is defined as the first sample in the waveform that crosses the level-one threshold (see Section 6.1), and is relative to the event trigger. In contrast, the TDC time is absolute: it is measured relative to the start of the run. Therefore, the key task in this time-based matching scheme is to estimate the time of the event trigger t_{trig} relative to the start of the run, so that ADC and TDC times can be compared.

For each pair of ADC and TDC hits on matching channels, we define a candidate trigger time t_i^{trig} as the difference between the ADC and TDC hit times:

$$t_i^{\text{trig}} = t_i^{\text{TDC}} - t_i^{\text{ADC}} \tag{6.3.1}$$

If both hits are from same SiPM signal, t_i^{trig} should approximate the true event trigger time. A time residual ϵ for each other pair of ADC and TDC hits can be calculated using this candidate trigger time:

$$\epsilon = (t^{\mathrm{TDC}} - t_i^{\mathrm{trig}}) - t^{\mathrm{ADC}} \tag{6.3.2}$$

The candidate trigger time t_i^{trig} which minimizes these residuals over all pairs of hits is selected as the final trigger time t_{trig}.

If there are multiple ADC or TDC hits on the same channel, the time residual ϵ for each pair of ADC and TDC hits is calculated as above, using the final t_{trig}. Each ADC hit is matched to the TDC hit which yields the smallest time residual ϵ. Hits are not reused; a TDC hit matched to one ADC hit cannot later be matched to a different ADC hit. If there are multiple ADC hits in the same channel, they are matched in order of ADC amplitude, so that the larger signals are prioritized. Furthermore, ADC and TDC hits are only paired if $\epsilon < 1\,\mu\mathrm{s}$.

6.4 Bar end matching

Once ADC and TDC hits have been combined into end hits, the next step is to combine two hits on opposite ends of the same bar into *bar hits*. Each bar hit represents the detection of scintillation light in a BSC bar resulting from a single charged particle interaction.

Ideally, each bar hit contains ADC and TDC data from both ends of the bar. However, in practice, some data may be missing. This can be caused by hardware issues such as a dead ADC or TDC channel[9], by photon attenuation if the scintillation light is fully absorbed before reaching one end of the bar, or by differing thresholds between the ADC and TDC.

To account for this, we define different categories of bar hits. A *complete* bar hit has ADC and TDC data from both ends. We also define eight types of *incomplete* bar hits, depending on which data is missing from the bar hit. Hits where one end of the bar has both ADC and TDC data are considered *half-complete*; these are used later in clustering.

The definition of these hit types, as well as their relative frequencies, are summarized in Table 6.1. The frequencies are calculated for cosmic ray data. The probability and type of observed incomplete bar hits is correlated to the hit position along the bar. Since the distribution of cosmic ray hits is roughly uniform along the length of the bar, the listed frequencies approximate an average over the length of the bar.

In most cases, each bar end has no more than one end hit. Matching the top and bottom end hits to create bar hits can then be done trivially

[9]For the 2023 gravity paper, there were four fully dead (ADC and TDC) bar ends due to a lack of bias voltage, and one further dead TDC channel.

Type	ADC 1	TDC 1	ADC 2	TDC 2	Frequency (%)
1	✓	✓	✓	✓	61.8 ± 0.6
2	✓	✓	·	✓	2.63 ± 0.09
3	✓	✓	✓	·	8.5 ± 0.2
4	✓	✓	·	·	10.9 ± 0.2
5	·	✓	·	✓	1.24 ± 0.06
6	·	✓	✓	·	0.21 ± 0.02
7	✓	·	✓	·	0.41 ± 0.03
8	·	✓	·	·	3.3 ± 0.1
9	✓	·	·	·	11.0 ± 0.2

Table 6.1: The definition of the nine bar hit types, depending on the presence of ADC and TDC data, and their relative frequencies in cosmic ray data. The columns ADC 1 and TDC 1 refer to one end hit on either the top or bottom of a bar, while ADC 2 and TDC 2 refer to the end hit on the opposite end. Type 1 (green) represents a complete bar hit, while types 2–4 (yellow) are half-complete. The remaining types (red) are denoted incomplete.

using the channel mapping. However, as in the ADC-TDC matching step, this process should be robust to multiple hits on the same bar during a single readout window. For this reason, a timing-based matching procedure is implemented. The bar hits are created through the following procedure, applied to each bar in turn:

1. Identify the top and bottom end hits which have the smallest TDC time difference. If the difference is less than 1 µs, combine these into a type 1 (complete) bar hit.

2. Repeat this process with remaining unpaired end hits to form additional type 1 bar hits.

3. After all type 1 bar hits have been formed, attempt to create type 2 bar hits by matching remaining end hits with an unpaired TDC hit on the opposite end.

4. Once all type 2 bar hits are formed, repeat this for bar hits of type 3, and so on for the other bar hit types, in order of decreasing completeness.

Only complete or half-complete bar hits are used for clustering, and only complete bar hits are used for TOF calculation. All types of bar hits are saved for use in event visualization and machine learning event classification. Further work is required to recover the timing of incomplete bar hits for use in TOF analysis.

6.5 Bar hit z calculation

At this stage, the z position of each bar hit, i.e. the point where scintillation light was produced, can be calculated. As discussed in Section 4.3.4, there are two available methods: charge division and time difference. Both methods are implemented and recorded, but only z_{TDC} calculated using the TDC time difference is used in the subsequent analysis. The z_{ADC} from ADC charge division is recorded in each event for future studies.

The z_{TDC} position is calculated as:

$$z_{\mathrm{TDC}} = \frac{v_{\mathrm{eff}}}{2}\left(t^{\mathrm{bot}} - t^{\mathrm{top}}\right) \tag{6.5.1}$$

Here, t^{top} and t^{bot} are the TDC times described in Section 6.2, including channel offset and time walk corrections. The parameter $v_{\mathrm{eff}} \sim 0.53c$ is the effective speed of light propagation through the scintillator, which is determined in the TDC calibration routine discussed in Chapter 7.

The z_{ADC} coordinate is calculated using the charge division method as:

$$z_{\mathrm{ADC}} = \frac{\lambda}{2}\ln\frac{A^{\mathrm{top}}}{A^{\mathrm{bot}}} + z_0 \tag{6.5.2}$$

Here, A_{top} and A_{top} are the ADC pulse amplitudes for the top and bottom SiPM signals, described in Section 6.1. The parameter λ is the photon attenuation length in the scintillator, including both bulk and surface absorption, and z_0 is the position where a scintillation creates equal top and bottom signal amplitudes.

The values of these parameters are determined in the ADC calibration routine, which is performed alongside the TDC calibration routine of Chapter 7. As shown in Figure 6.6, the z_{TDC} is plotted against $\frac{1}{2}\ln\frac{A_{\mathrm{top}}}{A_{\mathrm{bot}}}$ for each bar, and the slope and y-intercept of a linear fit yield λ and z_0 respectively. This is performed individually for each bar, with typical values of $z_0 \sim 0.2\,\mathrm{m}$ and $\lambda \sim 0.9\,\mathrm{m}$. These parameters vary between bars by up to 20%.

One notable feature of this plot is the apparent kink at $x \approx -0.2$, suggesting a change in slope of around 30% in some bars. This indicates an inconsistency between the two z reconstruction methods. The origin of this discrepancy remains unresolved. However, a comparison between z_{TDC} and the z_{TPC} of matched TPC tracks showed good agreement, while z_{ADC} was

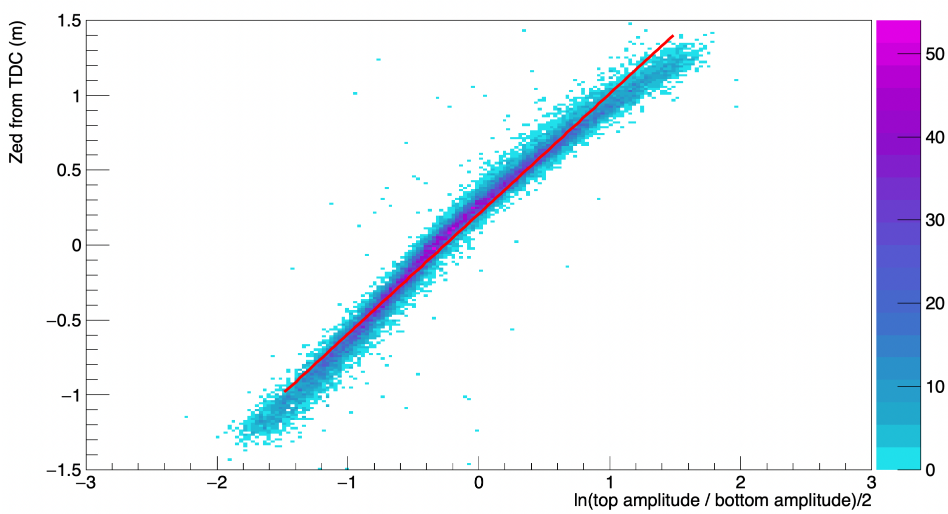

Figure 6.6: z_{TDC} plotted against $\frac{1}{2}\ln\frac{A_{\mathrm{top}}}{A_{\mathrm{bot}}}$ for one bar in cosmic ray data. The red line indicates a linear fit, the parameters of which are used to set λ and z_0 in Equation 6.5.2. The colour scale denotes counts per bin.

systematically inconsistent with both. For this reason, only z_{TDC} is used in subsequent analysis.

6.6 BSC-TPC matching

The next stage of the event reconstruction matches bar hits in the barrel scintillator (BSC) to helices reconstructed in the time projection chamber (TPC). The BSC has full azimuthal coverage of the TPC, with gaps only at the top and the bottom of the detector. As such, nearly every charged particle that produces an ionization track in the TPC is also expected to deposit energy in the BSC and generate scintillation light.

Until this point, the BSC and TPC event reconstruction routines have been entirely independent. The BSC-TPC matching is the first stage where data from the two detectors is combined. This analysis module is placed after the helix fitting and track joining modules, so that the final helices are available. However, it precedes the vertex fitting module. This ordering enables exclusion of unmatched helices from the vertex fit, to reduce contamination from poorly reconstructed tracks. Although this approach has been tested, it did not yield immediate improvements, and so the technique is not adopted in the current analysis.

Each bar hit is matched to the closest helix, with multiple bar hits allowed to be associated with the same helix. This design is compatible with observations made using the event display, where a single charged particle often produces bar hits on multiple neighbouring bars. Furthermore, only complete bar hits are used in matching, although this constraint could be relaxed in the future.

The matching for a given bar hit proceeds as follows:

1. The closest point on each helix to the bar hit is found, using the bar hit z_{TDC} from the TDC time difference method.

2. The transverse $(x-y)$ distance between the bar and this point on each helix is calculated.

3. The bar hit is matched to the helix with the smallest transverse distance, provided the separation in ϕ between the bar centre and the helix point is smaller than 17.2°, approximately the angular width of three bars.

This is a robust procedure which results in 84.7 ± 0.7 % of complete bar hits being successfully matched to a TPC helix. The matching performance is illustrated in Figure 6.7. The residual ϕ between the bar and the matched helix has a full width at half maximum (FWHM) of roughly 5.5°, comparable to the angular width of one bar (5.625°). The residual in z has a FWHM of roughly 12 cm. This can be taken as an indication of the z position resolution of the BSC using the time difference method. While far worse than that of spacepoints in the TPC, this resolution matches design expectations, and is more than sufficient for validating hit matching between the two detectors.

6.7 Bar cluster formation

Since a charged particle can produce bar hits on multiple neighbouring bars, the next event reconstruction stage groups bar hits into *clusters*. These extra bar hits are most commonly produced by charged particles with a sharp angle of incidence to the BSC, which directly pass through multiple bars. However, secondary particles are occasionally produced inside the bars or in surrounding material, leading to additional neighbouring bar hits. Clustering aims to group together all bar hits from a single initial particle, helping to simplify events and create a higher level of abstraction.

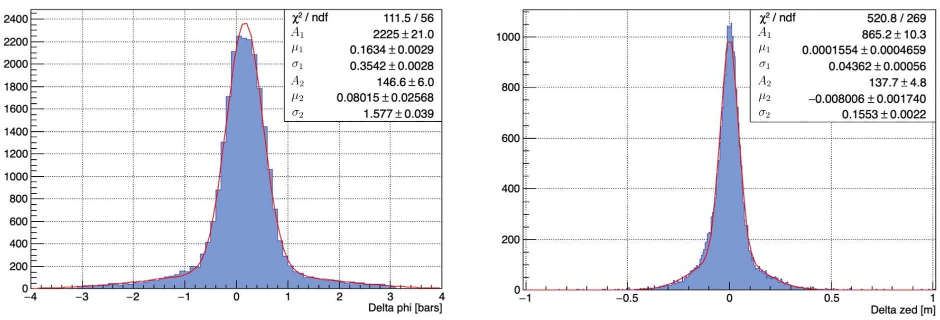

Figure 6.7: Left: ϕ residuals between bar hits and matched TPC helices, fit with a double Gaussian function. Angular units are converted to bars (1 bar = 5.625°). Right: z residuals between bar hits and matched TPC helices, fit with a double Gaussian function.

The main motivation for implementing bar hit clustering was to optimize the identification of cosmic ray background events. There are typically more charged particle tracks in an annihilation, and so the number of bar hits can be used to distinguish these events. However, the number of bar hit clusters is a far more powerful discriminator. Since cosmic rays often strike the BSC at an angle, they are more likely to create bar hits on multiple neighbouring bars. On the other hand, pions produced by annihilations have a mostly radial trajectory, and so mostly strike only a single bar.

Clusters are formed by grouping together all bar hits on contiguous neighbouring bars. A key design decision was the choice of which bar hit types to include in clustering. Including all nine bar hit types increases the number of clusters but introduces significant noise. Using only complete bar hits is cleaner but omits useful information. The final choice was to include only complete or half-complete bar hits, balancing data quality with granularity in the number and size of clusters. A comparison of these clustering strategies is shown in Figure 6.8.

In the real-time online event reconstruction described in Section 5.1, a cut on the number of clusters is used to classify events and reject vertices caused by the cosmic ray background. A comparison of the performance of this discriminator with different cluster definitions is given in Table 6.2.

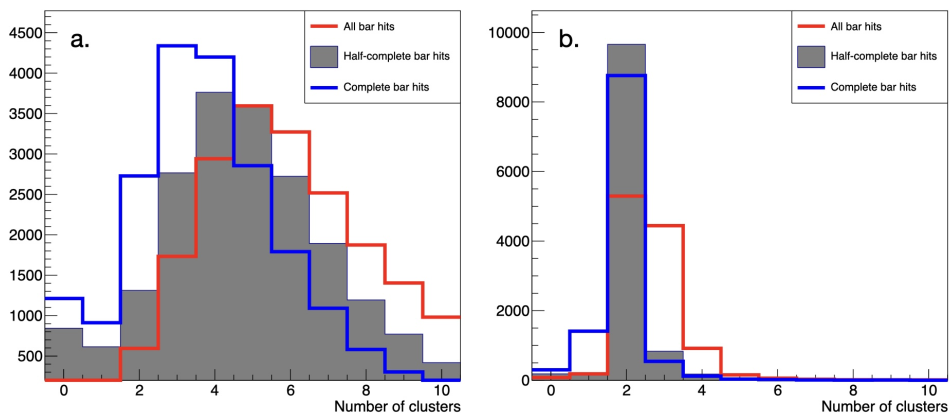

Figure 6.8: Number of clusters per annihilation (a.) or cosmic ray (b.) event. Shaded (grey) is current definition of a cluster, using only complete or half-complete bar hits. Red and blue lines indicate other possible definitions, using all bar hits (red) or only complete bar hits (blue).

	All Bar Hits	Half-Complete	Complete
Number of clusters to pass cut	>= 4	>= 4	>= 3
Annihilations passing cut	87.3%	72.7%	76.1%
Cosmic rays failing cut	89.3%	97.1%	93.5%
Cosmic ray pass cut rate	0.73 Hz	0.20 Hz	0.44 Hz

Table 6.2: Comparison of the performance of the online BSC cluster-based classifier with different cluster definitions. The half-complete bar hit cluster cut is currently used in online analysis.

6.8 Time-of-flight calculation

The final stage of the BSC event reconstruction is the calculation of time-of-flight (TOF) between pairs of bar hits. As described in Section 4.3.4, the time of a single bar hit is calculated from the TDC times at both ends of the bar:

$$t_0 = \frac{t^{\text{top}} + t^{\text{bot}}}{2} - \frac{L}{2v_{\text{eff}}} \tag{6.8.1}$$

Here, t^{top} and t^{bot} are the TDC times including the channel offset and time walk corrections described in Section 6.2. L refers to the bar length of 2.604 m. Although each bar has its own calibrated effective speed of light $v_{\text{eff},i}$ for z reconstruction, a single average value is used for all bars in this timing calculation. This approximation was verified to have a minimal effect on the resulting TOF resolution.

The TOF between two bar hits is calculated by taking the difference between their reconstructed hit times:

$$\text{TOF}_{2,1} = t_2 - t_1 = \frac{t_2^{\text{top}} + t_2^{\text{bot}}}{2} - \frac{t_1^{\text{top}} + t_1^{\text{bot}}}{2} \tag{6.8.2}$$

Note that the second term in Equation 6.8.1 cancels out in this subtraction, due to our approximation of equal v_{eff} across all bars.

For cosmic rays travelling between two bars, the TOF is set by the distance D between the two bar hits and the speed of the cosmic ray. The mean energy of a cosmic ray muon at the surface is 4 GeV [148], corresponding to a relativistic speed of $\beta = 0.9997$. With the approximation that all cosmic rays travel at the speed of light c, we define an expected TOF as:

$$\text{TOF}_{\text{exp}} = \frac{D}{c} \tag{6.8.3}$$

The expected TOF of cosmic rays is used to benchmark the timing performance of the BSC. Figure 6.9 compares the expected TOF to the measured TOF in cosmic ray events. Two populations are visible: a large-TOF group corresponding to muons passing entirely through the BSC, and a small-TOF group due to hits on neighbouring bars. The population separation is due to the trigger condition, which selects only cosmic rays passing through the TPC as well as the BSC.

The residual TOF, defined as the difference between the measured and expected TOF, is plotted in Figure 6.10. The width of the residual TOF distribution is used to characterize the detector resolution. A double-Gaussian

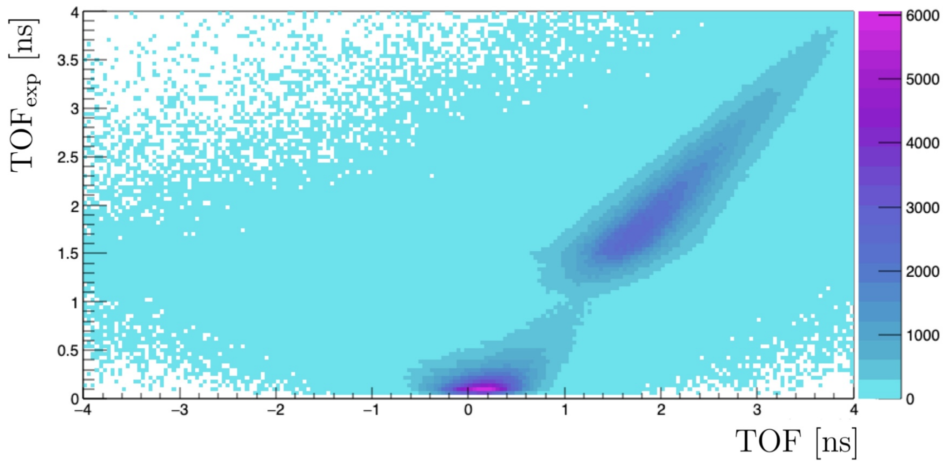

Figure 6.9: Measured vs expected TOF in cosmic ray data. In events with more than two bar hits, the TOF of all pairs of bar hits is included. The sign of the measured TOF is given by the assertion that the higher z position hit is the first one, which will be substantiated in Section 7.2.2.

fit to the residual distribution shows that the inner Gaussian (82% of events) has a width of $\sigma = 325\,\text{ps}$, which is taken as an upper bound on the TOF resolution of the BSC. The outer Gaussian corresponds to more complicated cosmic ray event topologies.

In annihilation events, charged pions originate at the trap wall and travel radially outward. The time difference between two bar hits does not correspond to the time-of-flight of a physical particle. Instead, it represents an *effective TOF*. This is typically small, since all pions travel similar distances from the annihilation point to the bars, at similar velocities of around $\beta = 0.8$ (for a pion with kinetic energy 100 MeV). Thus the effective TOF will be small; when compared to a cosmic ray's expected TOF, these annihilation events show large negative residuals. This difference is the key principle for TOF-based background rejection, which will be covered in Chapter 8.

The two-dimensional distribution of measured (effective) TOF compared to expected TOF is shown in Figure 6.11 for antiproton annihilations. The key feature of this plot is that events are far from the line TOF=TOF$_{exp}$, indicating that TOF can be effectively used to distinguish between the annihilations and cosmic rays.

Similarly to cosmic rays, this distribution contains two populations. The first with TOF$_{exp}$ $\sim 0.1\,\text{ns}$ corresponds to hits on neighbouring or nearby bars

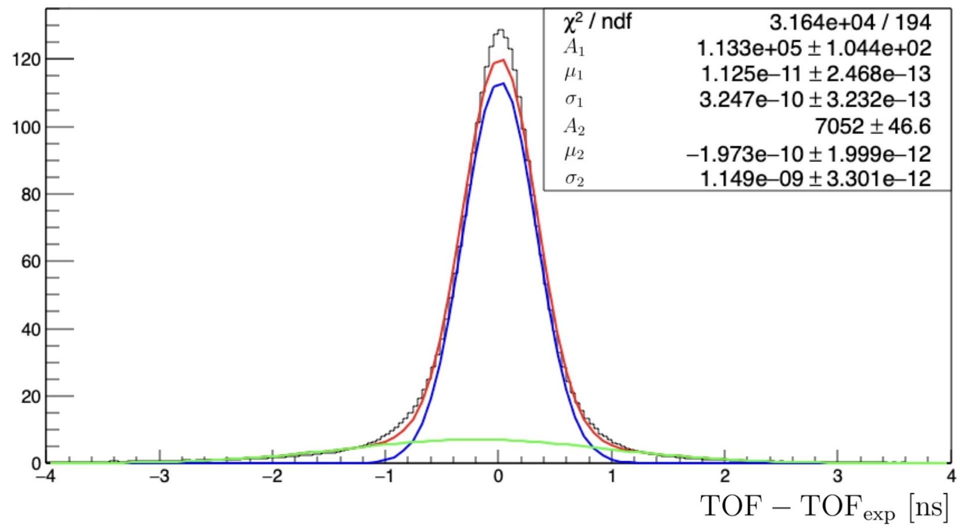

Figure 6.10: The residual TOF, i.e. the diagonal projection of Figure 6.9. The histogram (black) shows the data, which is fit with a double Gaussian (red). The inner Gaussian (blue) contains 82% of the data by total integral and has $\sigma = 325$ ps. The outer Gaussian (green) has $\sigma = 1.149$ ns.

caused by the same charged particle. The second population at $\text{TOF}_{\text{exp}} \sim 1.5$ ns corresponds to hits on opposite sides of the detector. Since annihilation events are much less vertical than cosmic rays, with hits at nearby z positions, the spread in TOF_{exp} is much smaller.

Currently, TOF is calculated only for pairs of *complete* bar hits, i.e., those with both TDC times and ADC waveforms available. In future work, the analysis could be extended to include the significant fraction of incomplete bar hits. For example, if a signal is only recorded on one bar end, the hit time could be inferred using the z position of a matched TPC track to estimate the scintillation light travel time.

6.9 Event display implementation

The development of the ALPHA-g event display was previously discussed in the context of the TPC reconstruction in Section 5.2. The BSC was integrated into the ALPHA-g event display to verify and demonstrate the effectiveness of the BSC analysis. Bar hits are visualized as larger circles positioned in the x-y plane at the centre of the corresponding bar, with their

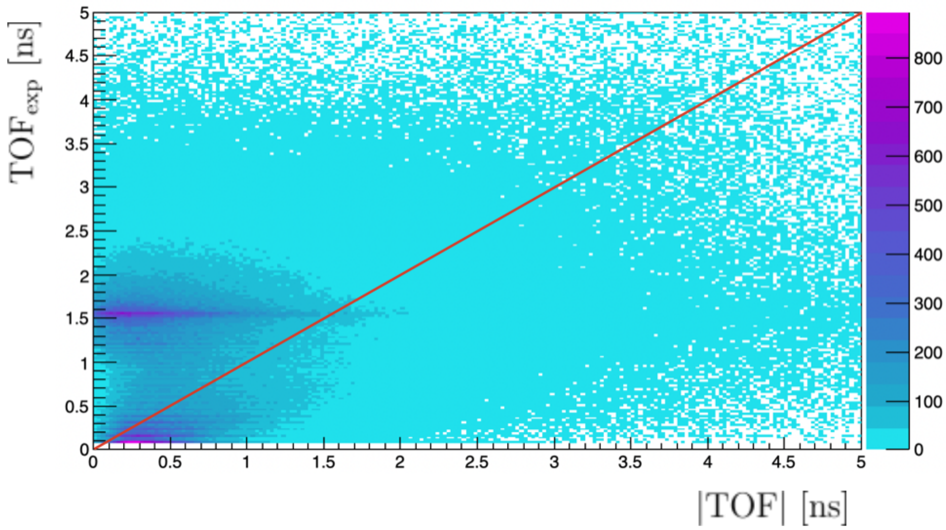

Figure 6.11: Measured vs expected TOF in antiproton annihilation data. Note the change in x axis from Figure 6.9, as the hit ordering based on z position can no longer be used to determine the sign of the TOF. The red line indicates the expectation TOF=TOF_{exp} for cosmic rays.

z position derived using the TDC time difference method. Only complete bar hits are shown by default, with an option to show incomplete hits with an artificial z position. The event display is shown here in Figure 6.12, as well as Figures 5.1 and 5.2.

Several colouring options are available for bar hits. The default assigns each bar hit the same colour as the spacepoints of its matched TPC track, to visually verify track matching. Additional options include colouring by ADC amplitude, TDC time, or by cluster, to demonstrate the cluster finding algorithm. These colouring modes are useful for quickly verifying correct analysis of a given event, and for quickly visualizing the event structure and timing.

Bar hits can be interactively selected by clicking, with an info panel displaying detailed information about the selected hit. Another panel[10] shows the ADC waveforms and TDC hit times for the selected hit, as shown in Figure 6.13. A further panel displays the TDC timing for the selected bar hit, visually indicating the channel offset correction and time walk correction.

[10] The event display uses a "dark mode" by default, which is somewhat easier to interpret, but sadly unsuitable for print.

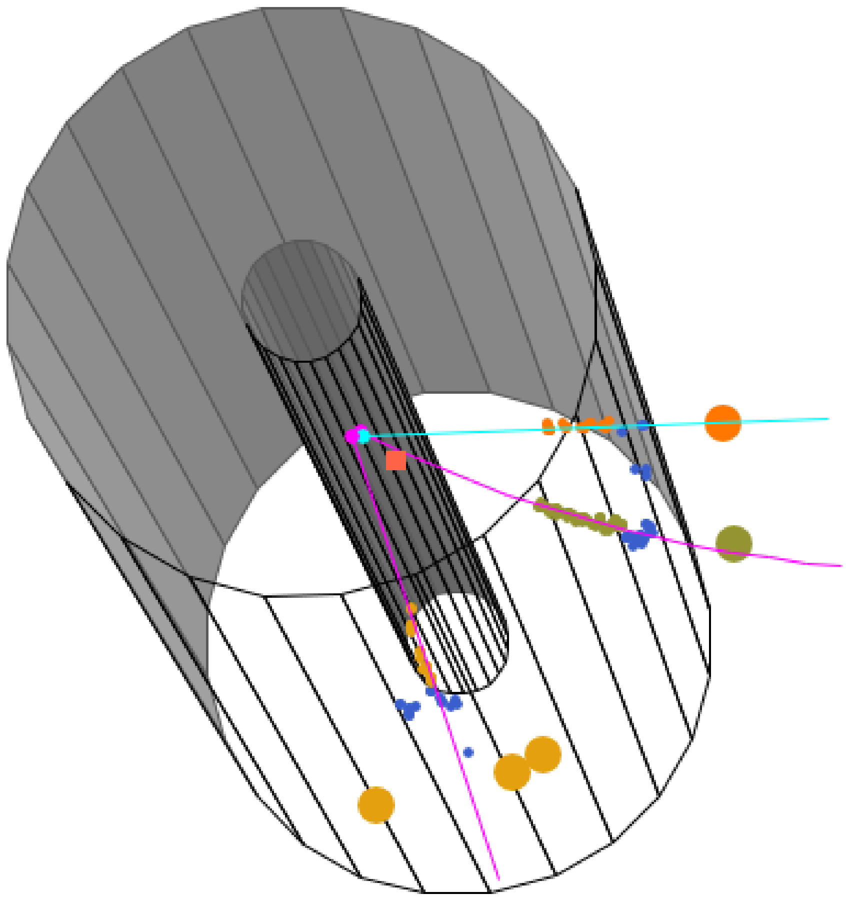

Figure 6.12: Event display of an antiproton annihilation. Large circles correspond to bar hits, while small dots correspond to TPC spacepoints. The colour of the spacepoints indicates the track finding, with blue indicating those not included in a track. The colour of the BSC hits is the same as the colour of the matched TPC track. Two tracks with one matched bar hit each are visible on the right of the image, and a single track is matched to three bar hits in the centre.

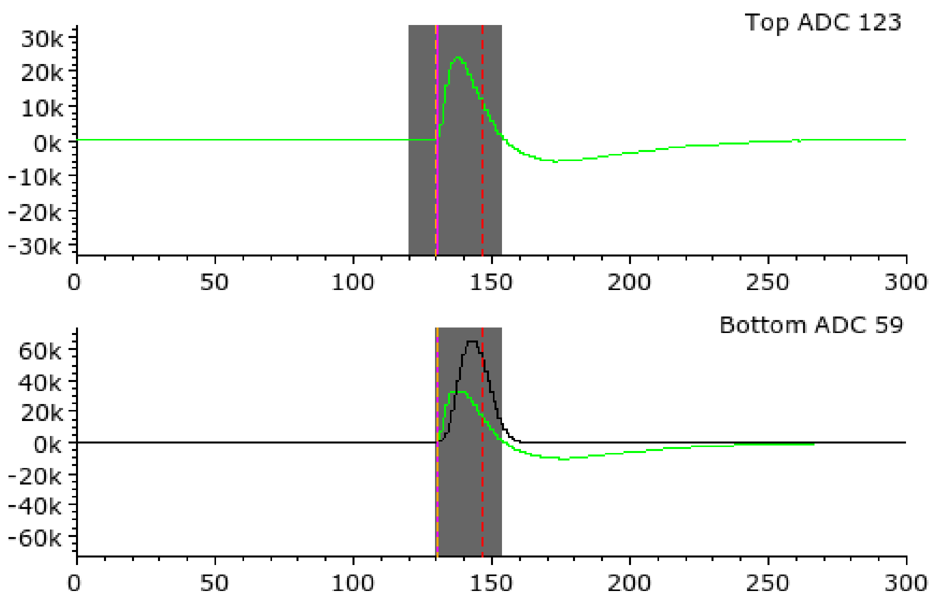

Figure 6.13: BSC waveform panel showing ADC waveforms (green) and TDC hit times (solid pink) from top and bottom of a selected bar hit. Also shown are the region above threshold (grey), the waveform fit in case of saturation (black), other TDC hits on the channel (dashed orange), and 165 ns cable reflections (dashed red) from Section 6.2.4. The x-axis is time in 10 ns bins, and the y-axis is voltage in ADC units.

Chapter 7

Precision time calibrations

As discussed in Section 6.2, there are two corrections which are applied to SiPM signal start time measured by the TDC: individual channel time offsets, and the time walk correction. The form of these corrections is given by Equation 6.2.5, repeated here:

$$t = t_{\mathrm{TDC}} + \delta - \sqrt{\frac{B}{A}} \tag{7.0.1}$$

The meaning of these symbols, and the symbols which will be used throughout this chapter, are given in Table 7.1 for reference.

These time corrections use correction constants δ_i^{top}, δ_i^{bot} and B_i^{top}, B_i^{bot} for each bar end which must be calibrated prior to running the *alphasoft* analysis. Accurate calibration is essential to achieve the time resolution necessary for cosmic ray identification and rejection using the time-of-flight (TOF) method. Initial simulations [149] indicated that a time resolution better than 500 ps is required for effective background suppression. The channel offset correction and time walk correction are each on the order of multiple ns, and so any attempt to measure TOF without first implementing these corrections was futile.

Traditional calibration methods, which are described in Section 7.1, were found to be inapplicable due to the constraints of the ALPHA-g detector system. Instead, a novel calibration procedure was developed to meet the requirements. This procedure updates the values of these calibration constants for each run, based on the cosmic ray data recorded between antiproton injections to ALPHA-g. The calibration algorithm was designed to yield the best possible time resolution, as well as to be robust and require minimum manual intervention.

A schematic diagram of the calibration procedure is shown in Figure 7.1; the calibration steps will be discussed in the remainder of this chapter. Section 7.2 describes the determination of the time walk correction parameters B_i^{top} and B_i^{bot}. Section 7.3 outlines the calibration of the bar hit z position. Section 7.4 details the calculation of the channel offsets δ_i^{top} and δ_i^{bot}. The specifics of the implementation of this algorithm into the ALPHA data

Symbol	**Description**	**Type**
$t^{\text{top}}_{\text{TDC},i}$,$t^{\text{bot}}_{\text{TDC},i}$	The raw hit time measured by the TDC for a hit on the top or bottom for the i'th bar.	Measured
A_i^{top}, A_i^{bot}	The ADC waveform amplitude of a hit the top or bottom of the i'th bar.	Measured
t_i^{top}, t_i^{bot}	The hit time measured by the TDC for a hit on the top or bottom for the i'th bar, with the channel offset and time walk corrections applied.	Calculated
δ_i^{top}, δ_i^{bot}	The channel offset time correction for the top or bottom of the i'th bar.	Calibration constant
B_i^{top}, B_i^{bot}	The time walk correction parameter for the top or bottom of the i'th bar.	Calibration constant
$v_{\text{eff},i}$	The effective speed of photon propagation in the i'th bar.	Calibration constant
$\Delta_i^{\text{bot-top}}$	$\Delta_i^{\text{bot-top}} = \delta_i^{\text{bot}} - \delta_i^{\text{top}}$ — The difference between bottom and top channel offsets of the i'th bar.	Intermediate
$\bar{\delta}_i$	$\bar{\delta}_i = (\delta_i^{\text{bot}}+\delta_i^{\text{bot}})/2$ — The average of the bottom and top channel offsets of the i'th bar.	Intermediate
$\overline{\Delta}_{i,j}$	$\overline{\Delta}_{i,j} = \bar{\delta}_i - \bar{\delta}_j$ — The difference between the average offsets of the i'th and j'th bars.	Intermediate

Table 7.1: A description of the symbols used throughout this chapter. Subscripts and superscripts are dropped when the specific bar end or bar number i is not relevant to the discussion. A distinction is drawn between quantities which are directly measured or calculated in each event, constants which must be calibrated beforehand, and intermediate values used in the calculation of these calibration constants.

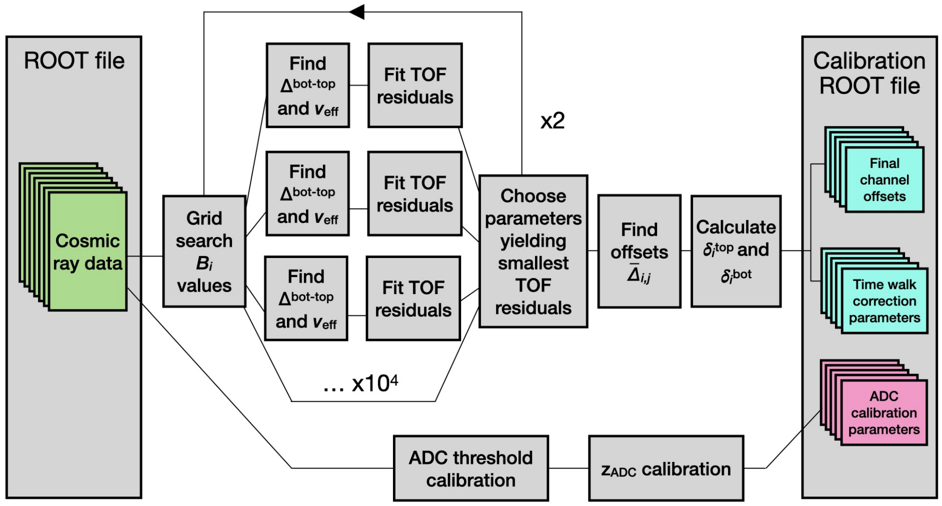

Figure 7.1: Schematic diagram of the BSC calibration algorithm. The steps leading to the calibration of channel time offsets and time walk correction are discussed in this chapter, while the ADC calibration procedure is discussed in Sections 6.1 and 6.5.

processing framework is reserved for Appendix D. This chapter will describe the calibration algorithm and methods in considerable detail, and can be skipped without loss of continuity.

7.1 Issues with traditional methods

In many particle physics experiments, channel offsets are determined using a pulsed calibration source which generates simultaneous signals across many channels. This source can take various forms, such as a laser calibration system [150], low rate bunch crossings in collider experiments [151], or a test pulse injected directly into the readout electronics.

For the BSC, the initial design aimed to use the test pulse generator described in Section 4.4.3 to inject signals after the first stage of SiPM signal amplification. However, this approach encountered several issues.

Firstly, the observed pulses in the ADC were too small for use in timing calibration. In many channels the signal was below the readout threshold. Furthermore, smaller pulses are more susceptible to the time walk effect, and so amplitude variation between channels caused calibration pulse time variation between channels which was difficult to quantify. The test pulse is differentially coupled to the readout electronics, and so increasing the size of the response required a faster rise time in the injected signal. This was achieved to a degree by introducing a fast external pulser.

Secondly, the test pulse is distributed to each ASD board by interface modules called *rear transmission modules* (RTMs), described in Appendix A. Due to a design fault, these modules introduce significant noise into the ADC waveforms when the test pulse distribution is enabled. Furthermore, the behaviour of these modules was inconsistent: with a common input pulse, the output pulse distributed to the ASD boards differs significantly between modules.

A third issue is presented by the TRB3 TDC itself. The magnitude of the channel offsets observed in TDC data cannot be explained by differences in signal path length alone. Instead, the channel offsets are dominated by internal behaviour of the TDC. These channel offsets change significantly after power cycling of the TDC. Compounding this issue, the TDC was found to be unstable during continuous operation, crashing frequently. These crashes are believed to be due to an unstable clock signal provided to the TDC from the CDM (see Section 4.4.3), which would require a significant hardware upgrade to address. With each crash scrambling the internal TDC channel offsets, any one-time or periodic calibration would be quickly rendered obsolete.

Traditional time walk calibration methods also proved unfeasible. These typically rely on events with a known initial time t_0, such as calibration test pulses, or bunch crossing times in collider experiments. The signal time delay $\delta t = t - t_0$ is typically plotted against the pulse amplitude, and the relationship is parameterized, often using the $\Delta t = \sqrt{B/A}$ parametrization from Section 6.2 [152–156].

In ALPHA-g, with the calibration pulser deemed unreliable, there is no mechanism to generate events with a well-defined t_0. Instead, events are spontaneous and self-timed, and so the traditional time walk calibration method is not applicable. A laser-based calibration source could have been considered in the initial BSC designs for in-situ calibration; however, significant problems with the timing calibration were not foreseen, and so such a system was not included.

A further limitation is presented by the nature of the ALPHA-g experiment. Given the fast pace of work at ALPHA, any calibration procedure which takes operator time or requires specific data collection would not be ideal. Such a procedure would be increasingly problematic if it is slow or requires frequent repetition.

With this in mind, cosmic rays were chosen as the source of calibration data. Cosmic ray data is already routinely collected, both in pure background runs used to verify detector functionality, as well as between cycles of antihydrogen formation. Cosmic ray detector signatures are simple and well understood, as they travel at roughly the speed of light and strike typically only two or three bars. The rate of cosmic ray events is sufficiently high for timely collection of a good calibration dataset, and their uniform distribution prevents any bias.

7.2 Time walk correction calibration

The first stage of the calibration algorithm determines the correction for time walk, the effect where smaller SiPM signals will cross the fixed TDC threshold later than larger signals of the same shape. As discussed in Section 7.1, this effect cannot be measured by directly comparing signal time and amplitude, since the true hit time is unknown. However, we can still use the parametrization of the time walk effect

$$\Delta t = \sqrt{B/A} \tag{7.2.1}$$

where Δt is the time correction to be subtracted from a TDC hit time, A is the pulse amplitude, and B is an effective threshold constant which must be calibrated using data.

We use cosmic ray data to determine unique values of B_i^{top} and B_i^{bot} for the two ends of each of the 64 bars. These time walk correction parameters will vary between bar ends based on a number of factors, such as SiPM breakdown voltage and gain, coupling between the scintillator and SiPM, and readout electronics amplification and signal shape.

7.2.1 Determining the time walk correction parameters

The time walk correction parameters are inferred from the cosmic ray data. A brute-force search of different values of B_i^{top} and B_i^{bot} is used. For a given set of B values, the TOF distribution of cosmic rays between two bars is calculated. The time walk correction parameters which lead to the narrowest TOF distribution are chosen; this method ensures that the TOF resolution is minimized.

The computation of B_i^{top} and B_i^{bot} for a given bar i is performed using cosmic rays interacting with bar i and one other specific bar. Usually, the bar directly opposite (bar $i \pm 32$) is chosen. This maximizes the available number of cosmic rays passing through both bars, and provides good time separation between the two hits. In the case where the readout of the opposite bar is malfunctioning, the closest fully functioning bar is chosen instead. The sample is further limited to events with three or fewer bar hit clusters, which minimizes the interference of annihilation events[11] or other events with more complicated topologies, such as scattering inside the detector.

The TOF calculation on this sample of cosmic rays involves four parameters: B_i^{top}, B_i^{bot}, $B_{i\pm32}^{\text{top}}$, and $B_{i\pm32}^{\text{bot}}$. To properly minimize the width of the TOF distribution, all four B values must be varied simultaneously. This is done by a grid search, where 10 different values are tried for each B, for a total of 10^4 combinations. The combination yielding the smallest TOF distribution width is chosen. The grid search is then repeated once, trying a further 10^4 combinations, centred on the current best combination of B values and with a smaller step size.

The approach follows these steps:

1. Apply a time walk correction $\Delta t = \sqrt{B_i/A}$ to each bar end.
2. For each cosmic ray, calculate the expected $\text{TOF}_{\text{exp}} = D/c$ based on the distance between the bar hits, with z from the TDC time difference.
3. Plot the residual TOF, i.e. after subtracting the expected TOF:

$$\text{TOF}_{\text{res}} = \text{TOF} - \text{TOF}_{\text{exp}} = \frac{t_{i\pm32}^{\text{top}} + t_{i\pm32}^{\text{bot}}}{2} - \frac{t_i^{\text{top}} + t_i^{\text{bot}}}{2} - \frac{D}{c} \quad (7.2.2)$$

4. Fit this distribution with a Gaussian and extract the width σ.
5. Repeat this for many different values of B_i, performing a four-dimensional grid search to vary the values of B_i individually for each bar end.
6. Repeat this grid search once with a finer step size.
7. Select the B_i values which minimize the width σ of the residual TOF distribution.

Figure 7.2 shows an example of a residual TOF distribution before and after applying the time walk correction, with the associated Gaussian fits.

[11] The calibration procedure is designed to use pure cosmic ray data. However, the majority of cosmic rays collected during ALPHA-g data taking are not in dedicated cosmic ray background runs, but between antihydrogen formation periods, while waiting to receive antiprotons from ELENA. Low rates of antihydrogen annihilations on background gas are possible during this time period, and so methods to exclude them are still helpful. More information on which time periods are used in long runs is given in Appendix D.

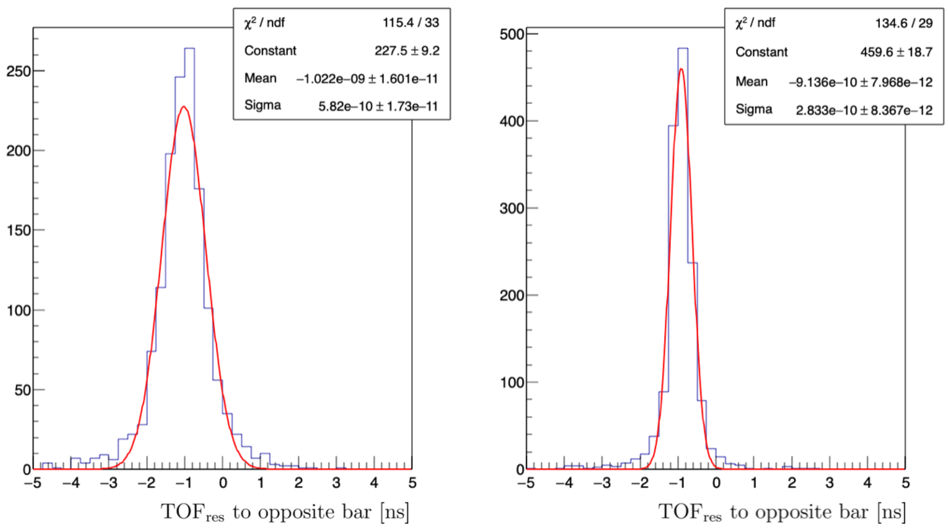

Figure 7.2: Residual TOF distribution from a bar to the opposite bar, with no time walk correction (left) and with the time walk correction parameters which minimize the distribution width (right). Both are fit with a Gaussian (red), with σ improving from 582 ps to 283 ps.

The σ of these fits are plotted for each bar in Figure 7.3. Note that the residual TOF distributions are not centred around zero; there is a large offset, due to the channel time offset between the two bars in question. However, this offset is constant for all events, and does not affect the width of the TOF distribution.

7.2.2 Determining the direction of cosmic rays

The sample of cosmic ray events contains two distinct classes: cosmic rays travelling from bar i to bar $i \pm 32$, and cosmic rays travelling in the opposite direction from bar $i \pm 32$ to bar i. These classes must be separated in order to properly subtract the expected TOF. Due to the potentially large channel time offsets shifting the TOF distribution away from zero, this cannot be done using the sign of the TOF.

Instead, we can identify the direction of the cosmic rays based on the z position of each bar hit. Since cosmic ray muons are unable to penetrate the Earth, the flux of cosmic rays at the surface is highly directional. The angular distribution of cosmic rays at the surface is approximately $I(\theta) \propto \cos^2\theta$ [148], where θ is the zenith angle, with no cosmic rays coming from below. Thus

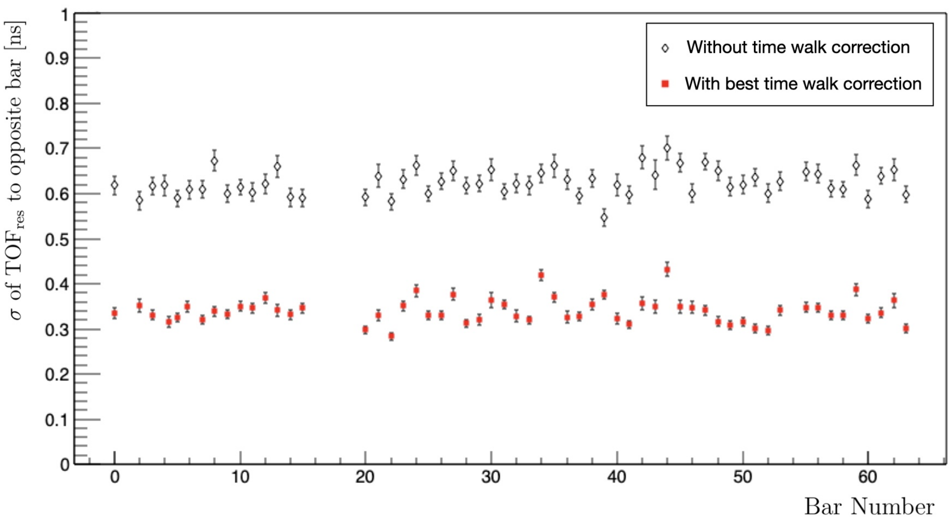

Figure 7.3: σ parameters of TOF_{res} fits, such as in Figure 7.2, for each bar. White diamond indicates with no time walk correction applied, and red square indicates with the final choice of time walk correction parameters.

we can use the vertical position of the hits to determine the direction of the cosmic ray: the hit with the higher z position must have occurred first. This phenomenon is shown in ALPHA-g data in Figure 7.4.

The direction of the cosmic ray, and hence the sign of TOF_{exp}, is determined by the sign of Δz. With the order of hits determined in this way, the residual TOF can be calculated properly and used to determine the time walk correction parameters.

Note that the alignment between the two directions of cosmic rays still has a small unexplained systematic shift, visible as a slight mis-alignment (around 0.5 ns) of the two populations in the right of Figure 7.4. For this reason, the positive Δz component was excluded. When determining the time walk correction parameter for bar i, only cosmic rays hitting first bar i and then bar $i \pm 32$ were used. The cause of this shift requires further investigation.

7.3 Bar hit z position calibration

Since no reference time is available, such as a precise annihilation time, we cannot measure absolute hit times. Instead, only the time difference between two SiPM signals can be measured; this limits our ability to directly

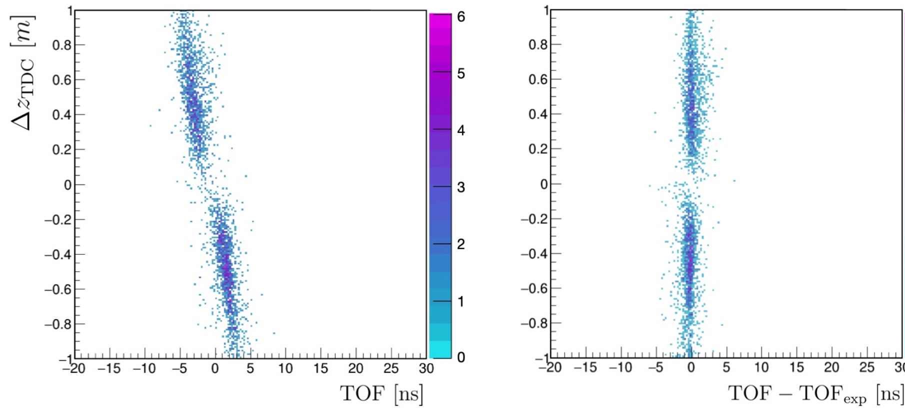

Figure 7.4: Left: The TOF of cosmic rays between a given bar and its opposite, plotted against the difference in hit z position. Note the linear correlation, where cosmic rays appearing to go upwards (positive Δz) have negative TOF, corresponding to an inverted hit ordering. Right: The residual TOF after subtracting the expected TOF, plotted against the difference in hit z position. The sign of the $\mathrm{TOF}_{\mathrm{exp}}$ is determined by the sign of z.

determine the individual channel offsets.

There is one important time difference for which we have a good calibration reference: the difference between the top and bottom of a bar, which is proportional to the z position of the bar hit. The z position of matched TPC tracks is used as an independent reference to calibrate this bar hit z position.

7.3.1 Determining $\Delta^{\text{bot-top}}$ and v_{eff}

Recall the calculation of z position from Equation 6.5.1, written here with the channel offset corrections δ explicitly included:

$$\begin{aligned} z_{\mathrm{TDC}} &= \frac{v_{\mathrm{eff}}}{2}\left[(t^{\mathrm{bot}} - \delta^{\mathrm{bot}}) - (t^{\mathrm{top}} - \delta^{\mathrm{top}})\right] \\ &= \frac{v_{\mathrm{eff}}}{2}\left(t^{\mathrm{bot}} - t^{\mathrm{top}} - \Delta^{\text{bot-top}}\right) \end{aligned} \tag{7.3.1}$$

Here, δ^{top} and δ^{bot} represent the channel offset corrections for a given bar, and $\Delta^{\text{bot-top}} = \delta^{\mathrm{bot}} - \delta^{\mathrm{top}}$ represents the offset difference between the top and bottom of the bar.

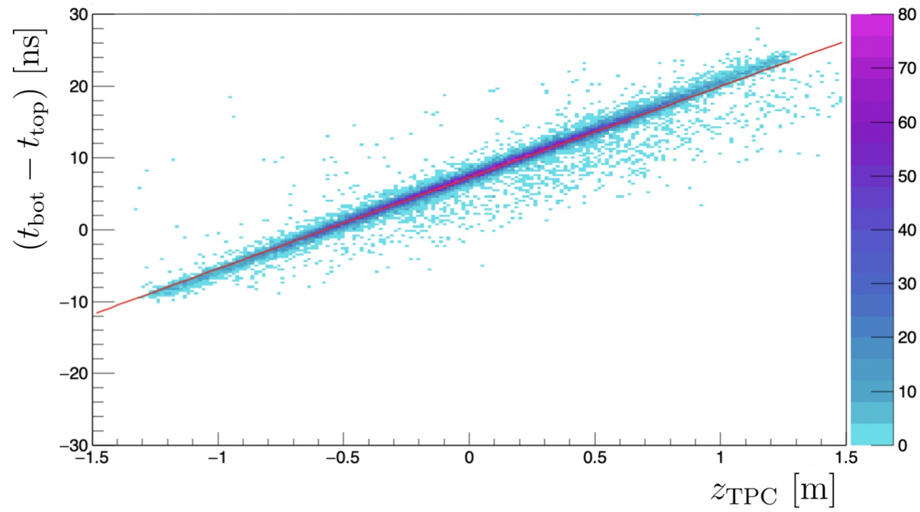

Figure 7.5: The TDC time difference between two ends of an individual bar, plotted against the z position of the matched TPC track. The red line indicates a linear fit, whose parameters are used to define $\Delta^{\text{bot-top}}$ and v_{eff}.

The next stage of the calibration routine is to determine the offset difference $\Delta^{\text{bot-top}}$ for each bar. This is done using cosmic ray bar hits which were successfully matched to TPC tracks. The relationship between $t^{\text{bot}} - t^{\text{top}}$ and z_{TPC} is plotted and fit with a line, shown in Figure 7.5. Setting $z_{\text{TDC}} = z_{\text{TPC}}$ in Equation 7.3.1 allows us to extract the offset and effective speed of light for each bar from the fit parameters:

$$\begin{aligned} v_{\text{eff}} &= \frac{2}{\text{slope}} \\ \Delta^{\text{bot-top}} &= \text{y-intercept} \end{aligned} \tag{7.3.2}$$

Typical values of $\Delta^{\text{bot-top}}$ and v_{eff} are shown in Figures 7.6 and 7.7 respectively. The v_{eff} values are slightly lower than those expected based on the nominal index of refraction of 1.58 [132]. There is a range of path lengths taken by photons to reach the SiPMs, depending on the number of reflections. This discrepancy in v_{eff} corresponds to a non-zero number of reflections for the average photon reaching the SiPMs.

This method takes advantage of the good matching between TPC tracks and BSC hits. It ensures that, to first order, there is no systematic bias in the z position of hits measured in the BSC compared to the TPC. It is

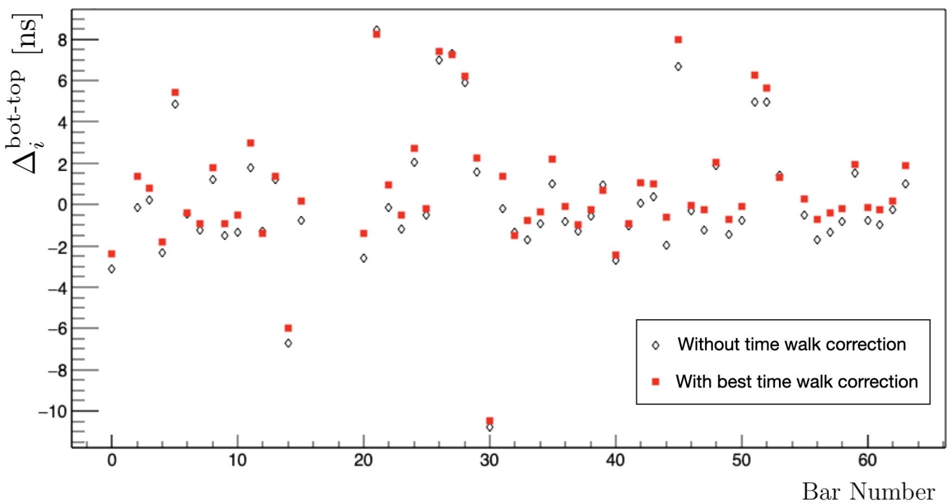

Figure 7.6: Time offsets between the top and bottom of each bar, determined via linear fit (shown in Figure 7.5). Data is from October 26, 2024.

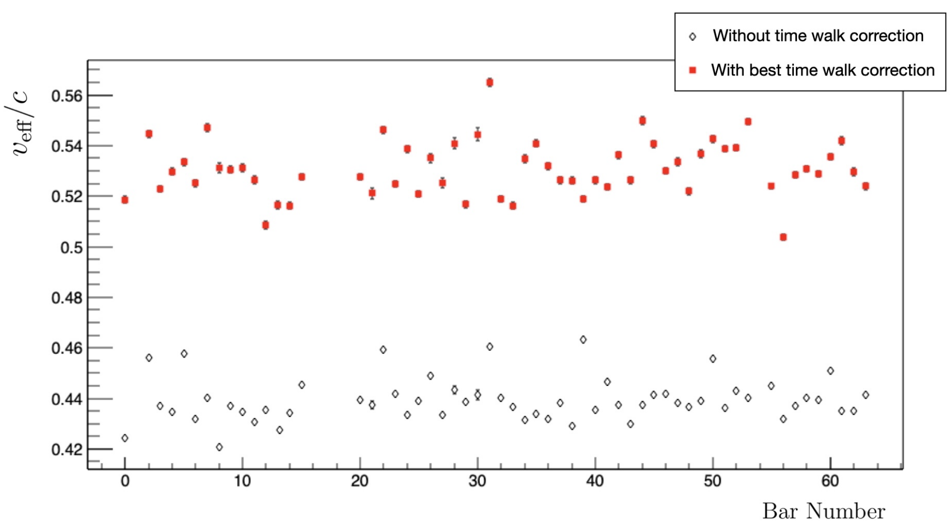

Figure 7.7: Effective speed of photon propagation in each bar, determined via linear fit (shown in Figure 7.5).

important to note that this step defines only the difference in channel offsets $\Delta^{\text{bot-top}} = \delta^{\text{bot}} - \delta^{\text{top}}$, which is used in the calculation of z_{TDC}. It does not fix the sum of channel offsets $\delta^{\text{bot}} + \delta^{\text{top}}$, which is the relevant quantity for calculations of TOF. These channel offsets will be determined individually in a later step.

7.3.2 Interplay between time walk correction and hit z position

As we have seen in Section 7.2, the bar hit z position is used in calibrating the time walk correction. Thus the calibration of the bar hit z position must come before the calibration of the time walk correction. However, the hit z position calibration described in this section is highly impacted by any correction for the time walk effect.

To understand this, consider the case where a large time walk correction Δt is applied to only the bottom end of a bar. The calculation of z_{TDC} following Equation 7.3.1 is then shifted:

$$\begin{aligned} z_{\text{TDC}} &= \frac{v_{\text{eff}}}{2}\left[(t_{\text{bot}} - \Delta t) - t_{\text{top}} - \Delta^{\text{bot-top}}\right] \\ &= \frac{v_{\text{eff}}}{2}\left[t_{\text{bot}} - t_{\text{top}} - \Delta^{\text{bot-top}}\right] - \frac{v_{\text{eff}}}{2}\Delta t \end{aligned} \tag{7.3.3}$$

With this in mind, it should be clear that the calibration of $\Delta^{\text{bot-top}}$ and v_{eff} should be done with the time walk effect properly corrected for. Ideally, the time walk correction would be calibrated first, and only then would $\Delta^{\text{bot,top}}$ and v_{eff} be determined. However, since the z position is used in the time walk calibration procedure, this is not possible. Instead, the calibration of $\Delta^{\text{bot-top}}$ and v_{eff} is repeated for each given set of time walk correction parameters B_i, before building the residual TOF distribution.

Figure 7.6 shows typical values of $\Delta^{\text{bot-top}}$ with and without the time walk correction. There is an interesting systematic increase when the values are calculated after the time walk correction is applied, corresponding to earlier top hit times. The SiPMs on the top of the detector have systematically smaller signals (see discussion in Appendix A), and hence later hit times due to the time walk effect. Correcting for time walk thus makes the top hits disproportionately earlier, increasing the $\Delta^{\text{bot-top}}$ values.

Figure 7.7 shows typical values of v_{eff} with and without the time walk correction. These also exhibit a significant systematic increase when calculated after the time walk correction is applied. To understand this, consider a scintillation event near the top end of a bar. Many photons will quickly reach the top SiPMs, resulting in quick signal with very little time walk effect. Due to attenuation, fewer photons will reach the bottom SiPMs, and those that do will arrive later due to the longer propagation distance. Thus the bottom signal, which is already later, will be made even later by the time walk effect. In this way, time walk has the effect of slowing down the apparent speed of photon propagation in the bar. It is expected, then, that correcting for the time walk effect will increase v_{eff}.

7.3.3 Summary of time walk and hit z position calibrations

For a given bar i, 10 values of each time walk correction parameter B_i^{top}, B_i^{bot}, $B_{i\pm32}^{\text{top}}$, and $B_{i\pm32}^{\text{bot}}$ were tried, in a grid search over 10^4 possible configurations. For each configuration of B values, the calibration of $\Delta_i^{\text{bot-top}}$ and $v_{\text{eff},i}$ was performed first. Then, a residual TOF distribution was built up, using only cosmic rays hitting first bar i and then bar $i \pm 32$, determined using their z position. This distribution was fit, and the width σ of the distribution was saved. The configuration with the smallest width was chosen as the best. A finer grid search of 10^4 more configurations is then performed, centred on those values. After this finer search, the best time walk correction parameters B_i^{top} and B_i^{bot} for bar i are saved, alongside the corresponding values of $\Delta_i^{\text{bot-top}}$ and $v_{\text{eff},i}$.

7.4 Bar to bar offsets

In Section 7.3, we determined the difference between the offsets of the ends of the same bar, $\Delta_i^{\text{bot-top}} = \delta_i^{\text{bot}} - \delta_i^{\text{top}}$. In this section, we will determine the value of each δ_i^{bot} and δ_i^{top} individually. This is done by decomposing each offset as follows:

$$\begin{aligned}
\delta_i^{\text{bot}} &= \frac{\delta_i^{\text{bot}} + \delta_i^{\text{top}}}{2} + \frac{\delta_i^{\text{bot}} - \delta_i^{\text{top}}}{2} = \overline{\delta}_i + \frac{\Delta_i^{\text{bot-top}}}{2} \\
\delta_i^{\text{top}} &= \frac{\delta_i^{\text{bot}} + \delta_i^{\text{top}}}{2} - \frac{\delta_i^{\text{bot}} - \delta_i^{\text{top}}}{2} = \overline{\delta}_i - \frac{\Delta_i^{\text{bot-top}}}{2}
\end{aligned} \tag{7.4.1}$$

With no fixed reference time, we cannot measure the value of each $\overline{\delta}_i$ directly. We can, however, measure the difference between two channels to obtain $\overline{\Delta}_{i,j} = \overline{\delta}_i - \overline{\delta}_j$. The approach taken is to extract these $\overline{\Delta}_{i,j}$ values from cosmic ray data, and then use them to determine the $\overline{\delta}_i$ values, which are finally used to determine the δ_i^{top} and δ_i^{bot} values.

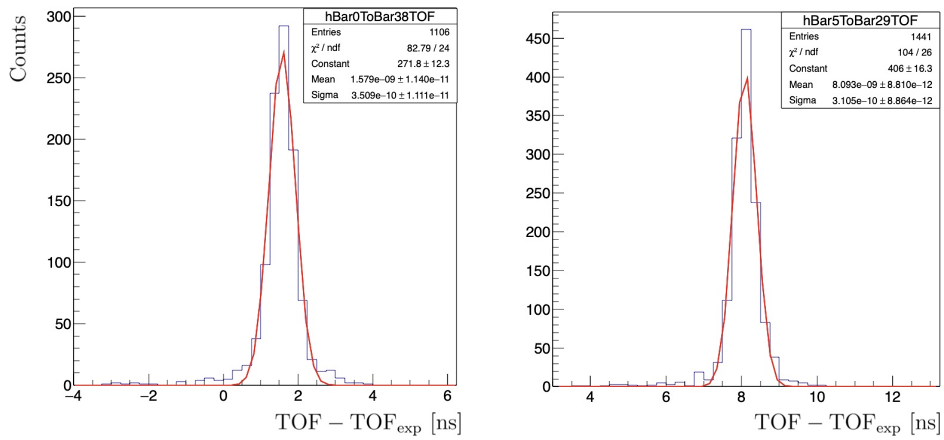

Figure 7.8: Residual TOF distributions from bar 0 to bar 38 (left), and from bar 5 to bar 29 (right). Gaussian fits are used to extract the time offset between the two bars.

7.4.1 Determining time offsets between pairs of bars

The determination of the offset between a given pair of bars i and j is performed using a similar method to the previous section. The residual TOF distribution is built from cosmic rays hitting first bar i and then bar j, with the order of hits determined using the hit z position. Since there are 64 bars, $64 \times 63 = 4032$ residual TOF distributions are populated simultaneously. In events with three or more bar hits, each combination of two hits is used. The same selection criteria as the previous section is applied: events must contain no more than three clusters of bar hits.

The determination of the time offset $\overline{\Delta}_{i,j}$ between bars i and j proceeds as follows:

1. Populate the residual TOF distribution using cosmic rays hitting first bar i and then bar j (determined using hit z positions).
2. Fit the residual TOF distribution with a Gaussian function.
3. Set the offset to the mean of the Gaussian $\overline{\Delta}_{i,j} = \mu_{\text{gaus}}$.

Examples of the residual TOF distributions between pairs of bars are shown in Figure 7.8, populated using 12.75 hours of cosmic ray data.

7.4.2 Determining time offsets of individual bars

At this point, the time offset difference $\overline{\Delta}_{i,j} = \overline{\delta}_i - \overline{\delta}_j$ between most pairs of bars is known. The remaining step is to use these 4032 measured values of $\overline{\Delta}_{i,j}$ to determine the 64 individual channel offsets $\overline{\delta}_i$. These can then be used in Equation 7.4.1 to determine each δ_i^{top} and δ_i^{bot}.

From our measurements of the differences $\overline{\Delta}_{i,j}$, we can determine the individual offsets $\overline{\delta}_i$ as follows:

$$\begin{aligned}\sum_{\text{all } j} \overline{\Delta}_{i,j} &= \sum_{\text{all } j} \left(\overline{\delta}_i - \overline{\delta}_j\right) \\ &= 64 \cdot \overline{\delta}_i - \sum_{\text{all } j} \overline{\delta}_j \\ \implies \overline{\delta}_i &= \frac{\sum_{\text{all } j} \overline{\Delta}_{i,j}}{64} + \frac{\sum_{\text{all } j} \overline{\delta}_j}{64}\end{aligned} \tag{7.4.2}$$

Note that here, the second term is a constant. Since only time differences can be measured in the BSC, adding a constant time to every hit has no effect. Thus the second term can be safely dropped, and the individual channel offset $\overline{\delta}_i$ can be found by averaging all the $\overline{\Delta}_{i,j}$'s.

While mathematically valid, this approach has some practical flaws. It requires successful measurement of all $\overline{\Delta}_{i,j}$ values, and quickly breaks down once we introduce fit failures, malfunctioning bars, or low statistics. A more robust approach is to relax the sum to include only a subset of bars. When calculating the offset $\overline{\delta}_i$ for bar i, we define G_i to be the subset of bars $j \in G_i$ for which the difference $\overline{\Delta}_{i,j}$ was successfully measured. The calculation of $\overline{\delta}_i$ then becomes:

$$\begin{aligned}\sum_{j \in G_i} \overline{\Delta}_{i,j} &= \sum_{j \in G_i} \left(\overline{\delta}_i - \overline{\delta}_j\right) \\ &= |G_i| \cdot \overline{\delta}_i - \sum_{j \in G_i} \overline{\delta}_j \\ \implies \overline{\delta}_i &= \frac{\sum_{j \in G_i} \overline{\Delta}_{i,j}}{|G_i|} + \frac{\sum_{j \in G_i} \overline{\delta}_j}{|G_i|}\end{aligned} \tag{7.4.3}$$

Here, $|G_i|$ represents the size of G_i.

Note that the second term is no longer constant, and so can no longer be dropped from the calculation. As a result, the offset $\overline{\delta}_i$ can no longer

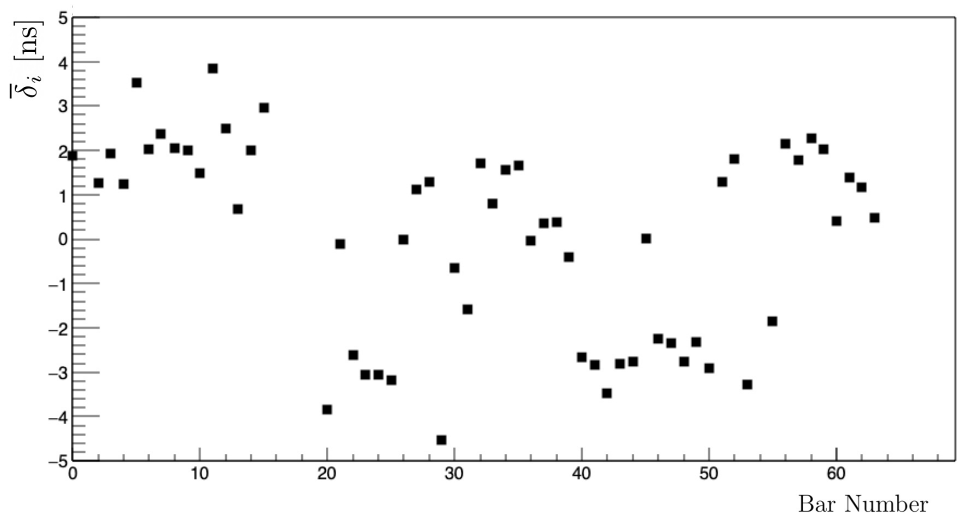

Figure 7.9: Time offsets for each bar, measured on October 26, 2024. Error bars are covered by the symbol.

be calculated analytically. Instead, we have a system of 64 equations in 64 unknowns (the $\overline{\delta}_i$ values), which must be solved simultaneously numerically.

An iterative approach is taken. The $\overline{\delta}_i$ values from Equation 7.4.2 are taken as starting values $\overline{\delta}_i^0$. These are then substituted into Equation 7.4.3 alongside the $\overline{\Delta}_{i,j}$ values to find updated values $\overline{\delta}_i^1$. These updated values are then plugged into Equation 7.4.3 to find $\overline{\delta}_i^2$, and so on, following:

$$\overline{\delta}_i^{n+1} = \frac{\sum_{j \in G_i} \overline{\Delta}_{i,j}}{|G_i|} + \frac{\sum_{j \in G_i} \overline{\delta}_j^n}{|G_i|} \tag{7.4.4}$$

This is repeated for 50 iterations; in practice, fewer than 10 iterations are required for convergence. The set of offsets $\overline{\delta}_i^{50}$ are then taken as the final values of $\overline{\delta}_i$ for each bar. These final offsets are plotted for each bar in Figure 7.9. Note that there is not one unique solution to this system of equations: since we can measure only time differences, a constant can be added to every $\overline{\delta}_i$ without affecting the measurements. This procedure converges on a set of offsets with a mean of $\mu = 0$.

Chapter 8

Cosmic ray background rejection

Cosmic rays are a substantial background for the TPC detector, triggering the detector at a rate of 58 Hz. Of these events, the reconstruction algorithm of Chapter 5 creates false annihilation vertices at a rate of 6.35 Hz; this is improved from 13.65 Hz as a result of my work on the TPC track reconstruction. This is problematic for physics measurements in ALPHA-g, as only a small number (~ 20) of antihydrogen are created per experiment cycle. The general strategy to address this background is to record all events, and distinguish between annihilations and cosmic ray background events afterwards at the analysis stage.

There are two methods used to reject the cosmic ray background in ALPHA-g. The first is with the BSC cluster-based cut described in Section 6.7. This method is used primarily in the live online *alphasoft* reconstruction described in Section 5.1, so that annihilation vertices can be used as a real-time diagnostic for antihydrogen production, with some straightforward background rejection.

The second method is using a machine-learning-based classifier called a *boosted decision tree* (BDT). This method combines many variables, which each provide a small degree of separation between signal and background, into a single strong classifier. A *decision tree* is a classification method which uses a branching structure, with the decision path split into two branches at each node. At each node, one of the input variables is compared to a cut value, and the result of the comparison specifies which of the two branches should be traversed. After reaching the end of the tree, each event is assigned a classifier value. The *boosted decision tree* is a method of training successive decision trees in an ensemble. Events which were misclassified by the previous decision tree are boosted, i.e. given higher weight in training; this focus on the harder-to-classify events results in a more accurate classifier.

This is an established method which has been successfully used to reject cosmic ray background events from the ALPHA-2 silicon vertex detector (SVD) [157]. A version using similar input variables was implemented in the ALPHA-g analysis for the 2023 gravity paper [1]. Following the time calibration work described in Chapter 7, TOF variables from the BSC were added to the ALPHA-g BDT, leading to a new classifier with significantly improved background rejection capability.

For context, Section 8.1 will briefly detail the BDT used in ALPHA-2. Section 8.2 will then describe the adaptation of this BDT to ALPHA-g for the gravity paper, and Section 8.3 outlines how the BSC TOF measurements were later included to further enhance the ALPHA-g BDT. The variables used as inputs to these two versions are listed in Section 8.4, and the final performance of each background rejection method is summarized in Section 8.5. Conlusions are drawn in Section 8.6.

8.1 BDT background rejection in ALPHA-2

Decision trees are regularly used to identify and reject cosmic ray events in the ALPHA-2 SVD [157]. This classifier is periodically retrained using antihydrogen production data as signal and pure cosmic ray background data, in order to keep the model updated as the performance of various detector components changes over time. The variables included are listed in Table 8.1. These variables were selected for their separation between signal and background events, and to be minimally correlated to each other [75].

Variable	**Description**
N_{hit}	Number of SVD hits.
N_{helices}	Number of reconstructed tracks.
$\tilde{N}_{\text{helices}}$	Number of reconstructed tracks used in vertexing.
ϕ_{vertex}	ϕ component of the reconstructed vertex position.
r_{vertex}	r component of the reconstructed vertex position.
δ	Squared residual (measures track straightness).
Σ_{ϕ}	ϕ component of the sphericity tensor eigenvector.
Σ_{z}	z component of the sphericity tensor eigenvector.
$\sqrt{\lambda_1^2 + \lambda_2^2}$	Combination of the two largest sphericity tensor eigenvalues.

Table 8.1: Discriminating variables used to distinguish signal (annihilation) and background (cosmic ray) events in the ALPHA-2 SVD.

The last three variables make reference to the *sphericity tensor* of an event, given by:

$$S_{ab} = \frac{1}{N_{\text{helices}}} \sum_{i=1}^{N_{\text{helices}}} \frac{p_a^i \, p_b^i}{|\vec{p}_i|^2}, \tag{8.1.1}$$

Here, p_i is the momentum of the ith track, and a, b refer to x, y, z. By diagonalizing this matrix, the eigenvalues and eigenvectors are found. The eigenvector Σ corresponds to the largest eigenvalue, and can be thought of as the principle direction of momentum in the event. The sphericity tensor eigenvalues, as well as the first eigenvector, are frequently used for describing event shapes in collider experiments [158].

The classifier yields a score c from -1 (background-like) to 1 (signal-like) for each event. A cut value C is then chosen to enable a binary classification: events with $c < C$ are rejected as background, while events with $c > C$ are accepted as signal. The cut value C is chosen based on the requirements of each individual physics analysis, depending on the required level of signal purity.

8.2 BDT background rejection in ALPHA-g

Based on the success of the BDT classifier in rejecting cosmic rays in the ALPHA-2 SVD, a similar classifier was trained for offline background rejection in ALPHA-g. This effort was lead primarily by Dr. Lukas Golino as part of his PhD project [75]. An initial version of this BDT classifier, using variables inspired by the ones used in the ALPHA-2 classifier, was used in the data analysis for [1]. The variables used in this classifier are primarily based on the TPC; only numbers of hits and hit clusters are used from the BSC. A full description of the variables used in this analysis is given in Section 8.4.

The training of a BDT classifier requires labelled data: events which are known to be either antihydrogen annihilation signal, or cosmic ray background. A sample of 255 000 cosmic ray background events was recorded between experiments when there are no antiprotons in ALPHA-g. A sample of 179 000 annihilation signal events was recorded during antihydrogen production, when a fraction of the newly formed antiatoms annihilate on the walls of the trap because their kinetic energy exceeds the magnetic trap depth. This is an intense 1 s time window with a TPC trigger rate of around 40 000 Hz. To ensure signal purity, this time window is subdivided into 2 ms sub-windows, and only those sub-windows where the detector readout rate is saturated at 500 Hz are used in training.

8.3 Inclusion of TOF in background rejection

After finishing the calibration described in Chapter 7, I worked to implement TOF measurements from the BSC into the BDT classifier. The measured TOF and the residual TOF (after subtracting expected TOF for cosmic rays) were both expected to be powerful discriminatory variables, supplementing the existing variables along a new axis. However, it was not initially clear how to combine the timing information of complicated events with many different hits.

A pragmatic approach was taken. Many quantities with possible discriminating power were calculated for each event utilizing the TOF information from the BSC. A total of 107 new variables were implemented and tested, in addition to the 20 variables used in the gravity paper version. A final set of 20 variables was selected by dropping those with high correlations to each other, as well as those with low separation between signal and background. These final variables are described in Section 8.4. The BDT trained with these 20 final variables had equivalent performance to the BDT trained with the full set of 127 variables. The classifier output is shown in Figure 8.1, which indicates very good separation between annihilation signal and cosmic ray background. The performance of this classifier is discussed in Section 8.5.

The introduction of variables based on TOF adds considerable complexity to the training procedure, as it requires the generation of BSC calibration files for the entire training dataset. A version of the offline calibration pipeline described in Appendix D is used to generate the training data in a reproducible fashion, and perform the BDT training. This allows the classifier to be easily retrained periodically, so that the model can be updated to reflect hardware performance changes.

8.4 Input variables

As described previously, an initial BDT was trained for use in the gravity paper analysis. This initial classifier used 20 input variables, primarily based on TPC data, but including some basic information from the BSC. Many of these variables were inspired by those used successfully in the ALPHA-2 classifier. After completing the TOF calibration for the BSC, I added many additional variables based on BSC data, and retrained a new BDT. The list of variables was then slimmed to the current 20 input variables with no significant decrease in classifier performance.

The variables used as input to the initial and final classifiers are summa-

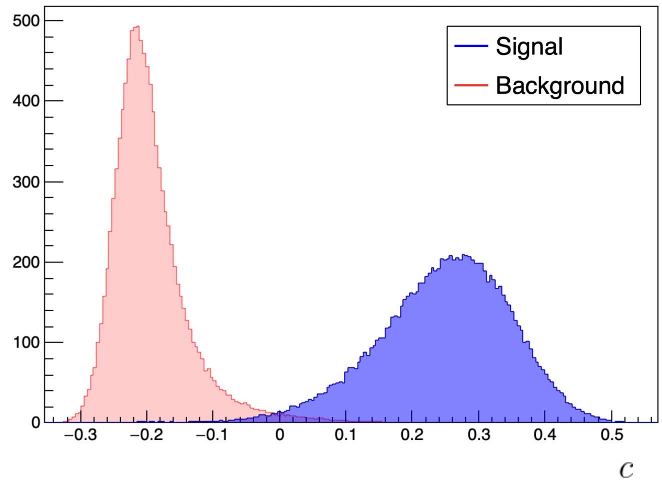

Figure 8.1: Classifier output c of the best-performing BDT trained with TOF and other variables. Classifier values range from -1 (background-like) to 1 (signal-like). The y-axis indicates counts, with data from a portion of the training dataset which was set aside for validation. Signal refers to antihydrogen annihilation data, and background to cosmic ray data.

rized in Table 8.2. Only 8 variables are shared between the new classifier and the initial version used for the gravity paper. While the initial version used only 4 variables based on BSC data, the new classifier uses 15 BSC variables. Of these, 9 rely directly on the calibrated timing information, either for the determination of the bar hit z positions, or for measurements of the TOF between bars. The remaining 6 are simple counting variables, or use only the ADC waveform.

It is notable that three quarters of the 20 variables selected are based on the BSC. This indicates that the BSC is fulfilling its role as an important detector for cosmic ray background rejection. It successfully measures event features which are not captured by the TPC, such as hit time and TOF information, and these features are powerful contributors to the background rejection algorithm.

The distributions of some of these input variables in signal and background are shown in Figure 8.2. These variables were selected to illustrate that the BSC calibration campaign has ultimately lead to a number of TOF-based variables with good separation between annihilation signal and cosmic ray background.

Given that the measurement of the z position of annihilations is central to the gravity measurement strategy, special care was taken to ensure that the classifier output is completely independent of the vertex z position. The annihilations used for training take place in a very small region of ALPHA-g, and so any variable with correlation to z could be used to efficiently recognize annihilation signal based on z position, leading to a highly biased classifier. For this reason, each input variable was checked to ensure it has no correlation with vertex z position in both signal and background.

Table 8.2: Comparison of variables used in the BDT trained for the gravity paper (initial), and the BDT trained incorporating my BSC work (final). Variables using calibrated hit times are highlighted in blue, and other BSC variables are highlighted in green.

Variable	**Description**	**Initial**	**Final**
N_{hit}	Number of TPC spacepoints.	✓	✓
N_{helices}	Number of reconstructed tracks.	✓	✓
N_{pads}	Number of TPC cathode pad hits.	✓	✓
Σ_z^{TPC}	z component of the sphericity tensor eigenvector using TPC tracks.	✓	✓
Σ_ϕ^{TPC}	ϕ component of the sphericity tensor eigenvector using TPC tracks.	✓	✓
N_{clusters}	Number of bar hit clusters.	✓	✓
N_{bars}	Number of bar hits.	✓	✓
N_{ADC}	Number of BSC ADC hits.	✓	✓
$\max(A)$	Largest bar hit ADC amplitude.		✓
$N_{\text{helices}}^{\text{unmatched}}$	Number of reconstructed tracks not matched to a bar hit.		✓
$N_{\text{helices}}^{\text{matched}}$	Number of reconstructed tracks matched to a bar hit.		✓
$\left(\sqrt{\lambda_1^2+\lambda_2^2}\right)^{\text{BSC}}$	Combination of the two largest sphericity tensor eigenvalues, calculated using bar hit positions.		✓
T_θ	Polar angle of the vector from first to last bar hit, determined using TDC time.		✓
max(TOF)	Largest TOF between two bar hits.		✓
min(TOF_4)	Smallest TOF between bar hits separated by at least 4 bars.		✓
$\max(\Delta z)$	Largest Δz between two bar hits.		✓

Variable	**Description**	**Initial**	**Final**
N_{TOF}	Number of pairs of bar hits with valid TOF measurements.		✓
$\max(\Delta\mathrm{TOF})$	Largest difference in TOF between two pairs of bar hits.		✓
$\max(\Delta\mathrm{TOF}_{\text{bar-to-bar}})$	Largest difference in TOF between two pairs of bar hits, calculated using the 4032 directly measured bar-to-bar offsets (see Section 7.4).		✓
$\mathrm{stddev}(\mathrm{TOF})$	Standard deviation of TOF measurements between all pairs of bars.		✓
N_{ends}	Number of bar end hits.	✓	
$\tilde{N}_{\mathrm{helices}}$	Number of tracks used in vertexing.	✓	
N_{AW}	Number of TPC anode wire hits.	✓	
R_{min}	Minimum helix radius of curvature.	✓	
R_{mean}	Mean helix radius of curvature.	✓	
λ_{min}	Minimum helix slope.	✓	
λ_{mean}	Mean helix slope.	✓	
$\sum c_{\pm}$	Integer sum of helix curvature (+1 for right curved tracks, -1 for left curved tracks).	✓	
μ_{points}	Mean spacepoints per track.	✓	
$\mu_{\chi^2_r}$	Mean χ^2 of track radial fit.	✓	
$\left(\sqrt{\lambda_1^2+\lambda_2^2}\right)^{\mathrm{TPC}}$	Combination of the two largest sphericity tensor eigenvalues, calculated using TPC tracks.	✓	
v_{status}	Vertex status, indicates if tracks beyond the first two were added.	✓	

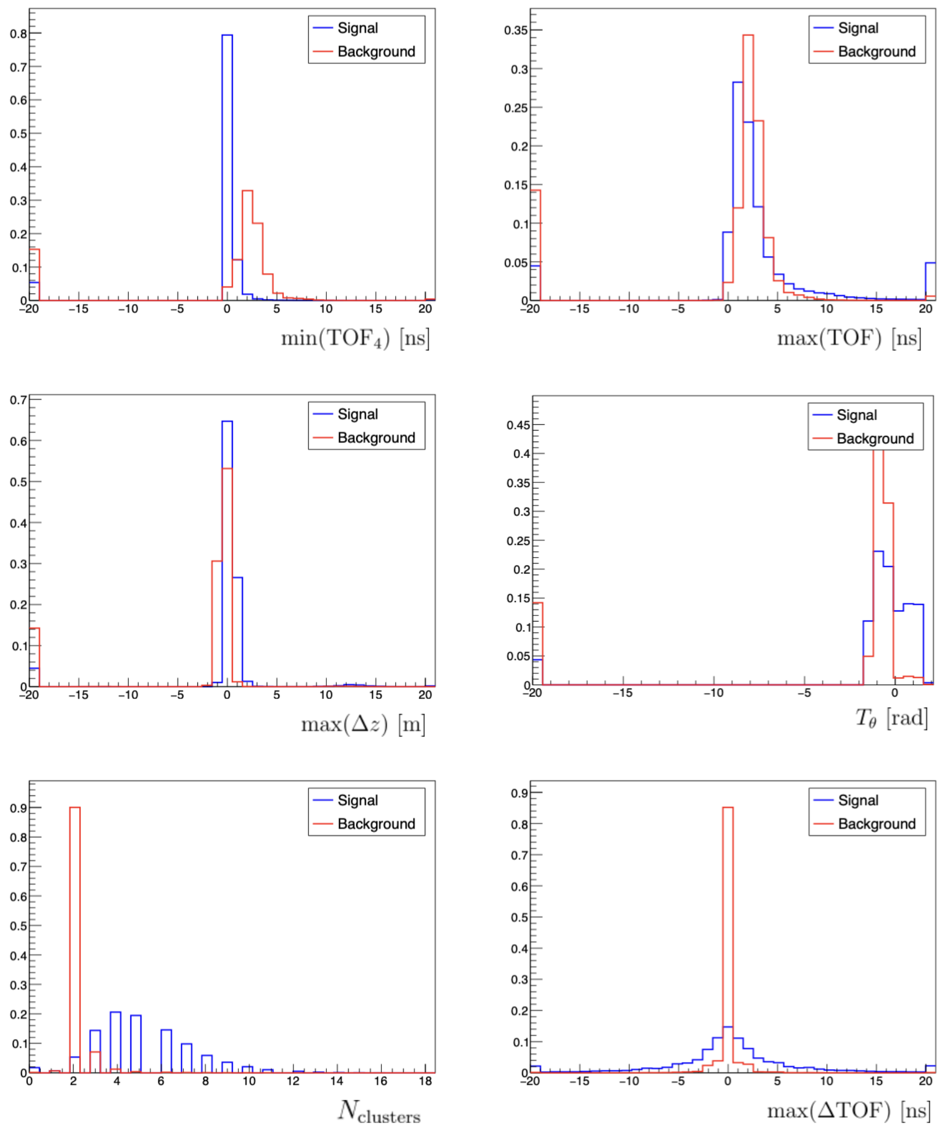

Figure 8.2: Six of the 20 variables used in the final BDT classifier, chosen to illustrate the separation between annihilation signal and cosmic ray background in TOF-based and other BSC variables. Values of -20 are assigned to events with no good TOF pairs.

8.5 Background rejection performance improvement

As this work was carried out concurrently with the development of antihydrogen production in ALPHA-g, the first measurement of gravity was performed before the background rejection work was completed. As such, there have been numerous historical background rejection methods. The annihilation signal acceptance and cosmic ray background rates are shown in Figure 8.3. The following methods are compared:

- **Initial online cut:** During the commissioning of the ALPHA-g traps leading up to the first measurement, a simple cut on the number of bar hit clusters N_{clusters} was used in online analysis (see Section 6.7).
- **Initial BDT:** A BDT classifier was trained for use in the first gravity measurement, using primarily TPC variables modelled after the successful classifier used in ALPHA-2 (see Section 8.2).
- **Improved online cut:** Following many of the TPC track fitting changes described in Section 5.4, the cosmic ray background vertex rate in ALPHA-g was significantly reduced. The same simple cut on N_{clusters} now yields improved performance.
- **Improved BDT:** After the implementation of the BSC calibration routine, a new BDT classifier was retrained using TOF variables (see Section 8.3).

It should be noted that the TPC tracking changes and inclusion of TOF information both had the effect of lowering the background rate of the improved BDT. No intermediate BDT was trained with the improved tracking but without TOF variables, and so the contributions of the two improvements are difficult to disentangle.

8.6 Background rejection conclusions

Overall, the results shown here indicate that an order of magnitude improvement in the background rate of the ALPHA-g detectors has been achieved. This is the result of my work developing the BSC event reconstruction, which includes a substantial TOF calibration algorithm, and my work including TOF measurements into the BDT classifier. My work on the TPC track reconstruction algorithms also benefited the background rate. The impact of

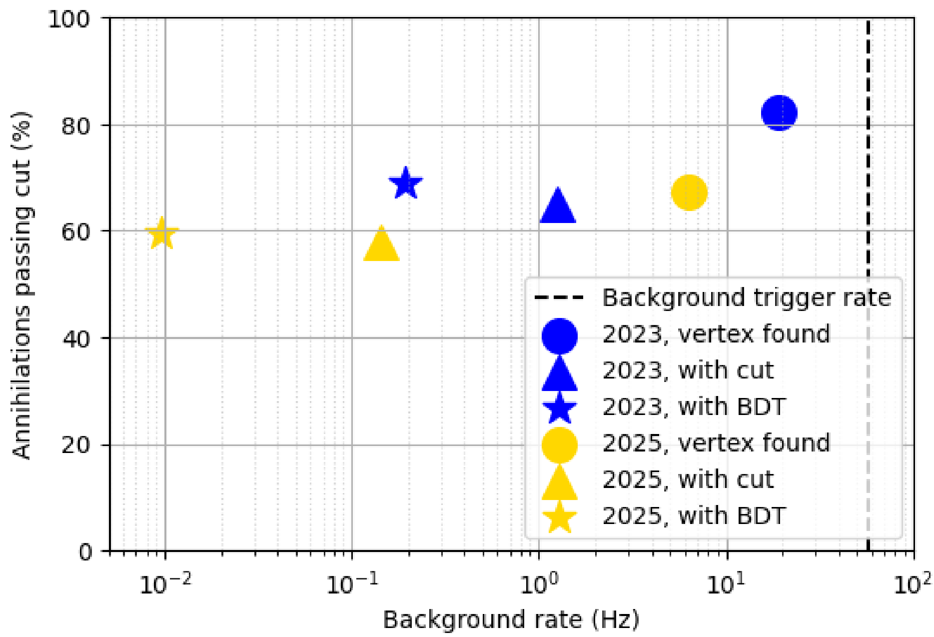

Figure 8.3: Comparison of signal efficiency and background rejection performance between the initial (2023) analysis used for the gravity paper [1], and the improved (2025) analysis at the completion of this work. The dashed line indicates the background detector trigger rate, and the circles indicate events for which a vertex is successfully found. The N_{clusters} cut used for online analysis is indicated by triangles, while the BDT classifier is indicated by the stars. The choice of cut $C = 0.146$ on the classifier output is shown for the 2023 BDT, which was used in the gravity paper, chosen to optimize the significance for an expected 1 count of signal in a period of 8 seconds. For the 2025 BDT, the choice $C = 0.134$ is shown. This choice is arbitrary, and should be re-examined prior to interpreting experimental results.

this improvement on the measurement of antihydrogen gravity in ALPHA-g will be the focus of the next chapter.

This improvement comes at the cost of a lower signal yield: the annihilation vertex is found successfully in around 15% fewer antihydrogen annihilations, translating to around 9% fewer annihilations being accepted by the BDT classifier. This reflects an intentional choice where events with complex topologies, for which the annihilation vertex can not be accurately found, no longer yield a vertex position. If a future analysis requires a high number of vertices, parameters of the vertex reconstruction could easily be re-tuned in order to recover some of these lost vertices.

Chapter 9

Measuring antihydrogen gravity with ALPHA-g

The first results from ALPHA-g, constituting the first direct measurement of gravity acting on an antihydrogen atom, were published in 2023 [1] based on data collected in 2022. The main result of this publication, referred to throughout this dissertation as the *gravity paper*, was that the Earth's gravitational acceleration $\overline{g}$ on antihydrogen has the same sign as the Earth's gravitational acceleration g on regular matter. The ratio of the two constants was determined to be:

$$\overline{g}/g = (0.75 \pm 0.13 \text{ (statistical+systematic) } \pm 0.16 \text{ (simulation)}) \quad (9.0.1)$$

The probability that this measurement is consistent with no gravitational action on antihydrogen is 3×10^{-4}, and the probability of it being consistent with repulsive gravity ($\overline{g} = -g$) is less than 10^{-15}.

The first three sections of this chapter will summarize the key elements of the gravity paper. Section 9.1 describes the experimental procedure, including the magnetic field control and characterization, details of the recorded data set, and the unprocessed results. Section 9.2 presents the statistical analysis used to extract the results quoted above. Section 9.3 addresses statistical and systematic uncertainties on the results, which inform future measurements.

Section 9.4 highlights the author's contributions in the broader context of ongoing and upcoming ALPHA-g measurements. It discusses how the improved track reconstruction and background rejection impacted the gravity paper, and how they will benefit subsequent measurements using the same apparatus. It concludes with a discussion of long-term plans involving the precision trapping region of ALPHA-g, and outlines possible alternative measurement schemes enabled by this work.

9.1 Gravity measurement procedure

The design of the ALPHA-g trap apparatus was presented in Section 3.4.6. The reader should refer back to Figures 3.9 and 3.10 in particular. The measurement procedure of the gravity paper used only a subset of the ALPHA-g trap magnets, as commissioning of the full magnet system is still ongoing. A more precise measurement using the full apparatus will be performed in the future. This section will describe the procedure used in the gravity paper, with upcoming measurements discussed in Section 9.4.

9.1.1 Measurement technique

The measurement strategy begins with the accumulation of antihydrogen atoms in ALPHA-g, as described in Chapter 3. The accumulation cycle is repeated 50 times in each run, trapping around ~ 50 antihydrogen atoms over 3-4 hours. During this time, the antihydrogen atoms are held by the magnetic field of mirror coils A and G (see Figure 3.10), collectively referred to here as the *mirror coils* or simply *mirrors*. They are radially constrained by the combined field of the long octupole magnet and the short octupole magnet.

The key step in the gravity measurement is to slowly release the trapped antihydrogen atoms from their axially confining magnetic potential (shown in Figure 9.1), while keeping them radially constrained. This is done by ramping down the current in the mirror coils simultaneously over 20 seconds, hereafter called the *mirror rampdown*. The antiatoms can then escape the trap axially, either upwards or downwards. They remain radially constrained by the octupole field until they exit the short octuple magnet, at which point they annihilate on the trap walls. The ALPHA-g detector system is used to count the annihilations in the *up* and *down* regions, and the asymmetry between the counts in these regions indicates the net vertical force acting on the antihydrogen atoms.

The success of the experiment depends on precise magnetic field control. Mirror coils A and G are separated by 25.6 cm; the gravitational potential difference (for matter) between the two mirrors is $mg\Delta z = 2.6 \times 10^{-8}\,\mathrm{eV}$. This corresponds to a Zeeman energy $\vec{\mu} \cdot \vec{B}$ for a magnetic field of $4.53 \times 10^{-4}\,\mathrm{T}$. This sets the necessary precision for the control of the magnetic field.

Simulations indicate that under the effect of gravity with $\overline{g} = g$, around 80% of the trapped antiatoms should annihilate in the down region. However, this ratio can be adjusted by holding one of the mirror coils at a slightly

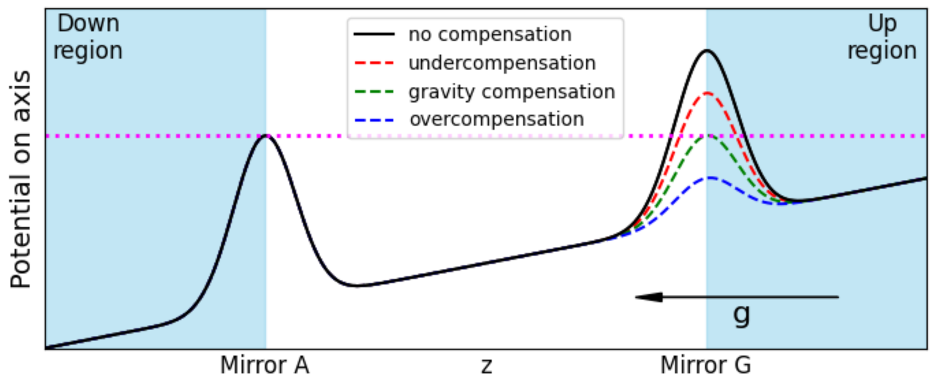

Figure 9.1: A sketch of the on-axis potential (magnetic + gravitational) in the ALPHA-g trap, with different applied magnetic biases compensating for gravity. The dotted purple line indicates the maximum magnetic potential under mirror A, which sets the kinetic energy required for antihydrogen atoms to escape into the down region.

higher current than the other throughout the magnet ramp. This applies a *magnetic bias* which can either counteract or enhance the effect of gravity. For example, by decreasing the field of mirror G by $1g = 4.53 \times 10^{-4}\,\mathrm{T}$ as shown in Figure 9.1, gravity can be compensated for and the balance of annihilations restored to 50% up, 50% down. Here, we introduced the convention of using units of g for magnetic field difference. It should be emphasized that this magnetic bias must be applied before, after, and throughout the mirror rampdown.

To accomplish precise magnetic control, the mirror coils are energized in series by a 70 A power supply, ensuring the two magnets receive equal current regardless of power supply noise or drift. A differential current is then supplied to mirror G alone to generate the magnetic bias. Direct-current current-transformers (DCCTs) are used in combination with a 50 kHz ADC to precisely measure the current flowing through the mirror coils. The measured current is then used to control the output of the magnet power supply using a proportional-integral-derivative (PID)-based loop, as well as to confirm consistent magnet power supply behaviour throughout the experiment.

The reader should note that the one-dimensional picture presented here is highly simplified. In reality, almost all antihydrogen atoms escape the trap off-axis, and so the entire three-dimensional magnetic field profile of the trap should be considered. The gravity compensation is only approximate: for example, mirror coil currents which balance the up and down potentials on-axis (the *gravity compensation* curve in Figure 9.1) will not necessarily balance the potentials at all radii. As such, an accurate interpretation of the experiment relies on the antihydrogen trajectory simulations described in Section 9.2.3, which must be informed and validated by accurate measurements of the magnetic fields throughout the measurement.

9.1.2 Magnetic field measurements

To ensure precise and repeatable control of the magnetic fields, and to characterize the applied bias in each configuration, a magnetic field measurement technique called *electron cyclotron resonance* (ECR) was developed at ALPHA. A small plasma, or reservoir, of electrons is created by an electron gun mounted on the ALPHA-g bottom diagnostic stick (Section 3.4.4), and prepared using the SDREVC technique (Section 3.5.1). For each measurement, a small bunch of electrons is extracted from the reservoir, and moved to the target region of the trap.

As discussed in Section 3.2.1, the electrons undergo cyclotron motion in the Penning-Malmberg trap, with cyclotron frequency $\omega_c = qB/m$ depending on the magnetic field strength. By determining the cyclotron frequency, the local magnetic field strength can be inferred. This is accomplished by injecting microwaves into the ALPHA-g bottom trap. When the microwave frequency matches the cyclotron frequency, the electron plasma undergoes resonant heating.

The temperature of the electron bunch is measured via slow extraction to the MCP mounted on the bottom diagnostic stick, in a procedure similar to the one described for antiprotons in Section 3.5.1. As the confining potential is slowly lowered, the electrons escape, and the kinetic energy of each electron is determined from its arrival time at the MCP. The energy distribution is fit to determine the plasma temperature.

Many small bunches of electrons are extracted from the reservoir and moved to the target location in rapid succession. Each is exposed to microwaves of a different frequency. The temperature at each frequency is measured, and the spectrum is fit to determine the resonance peak frequency, which corresponds to the cyclotron frequency. The magnetic field strength at the target location is determined from this measured frequency. Two example frequency measurements are shown in Figure 9.2. For further details on the ECR technique, see [78] and [1] (Methods).

This technique was used to measure the magnetic field under the mirror coils three times during the measurement procedure: before the mirror

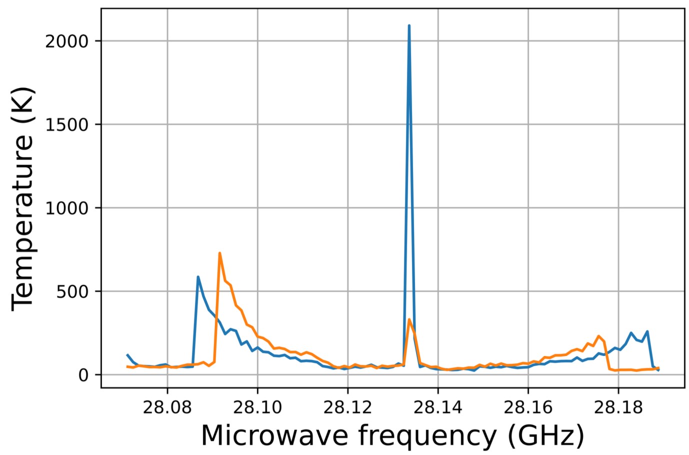

Figure 9.2: Frequency spectra measured in ALPHA-g using ECR. Orange and blue lines represent measurements at the same location in the same magnetic field, but with different Penning-Malmberg confining potentials. For further information, see image source [78].

rampdown, after the mirror rampdown, and after zeroing all mirror coil and octupole magnet currents. This procedure helped ensure that the magnetic field behaved as expected across all trials. In addition, the ECR technique was used to map the magnetic field produced by each of the magnets, and these field maps were used as inputs to the antihydrogen simulations discussed in Section 9.2.3. ECR was also used to study persistent magnetic fields after the end of the rampdown caused by persistent currents in the superconducting mirror coils.

A complementary technique for measuring the magnetic field, called magnetron-metry, was also developed based on the electron magnetron frequency. A small plasma of electrons is displaced from the trap axis, and precesses around the trap centre due to its magnetron motion described in Section 3.2.1. By imaging the plasma using the ALPHA-g MCP after a short time, the number of precessions around the trap centre can be counted to extract the magnetron frequency, and thus determine the local magnetic field. This technique was used to further constrain the magnetic field models used in simulation, and is described further in [27] and [1] (Methods).

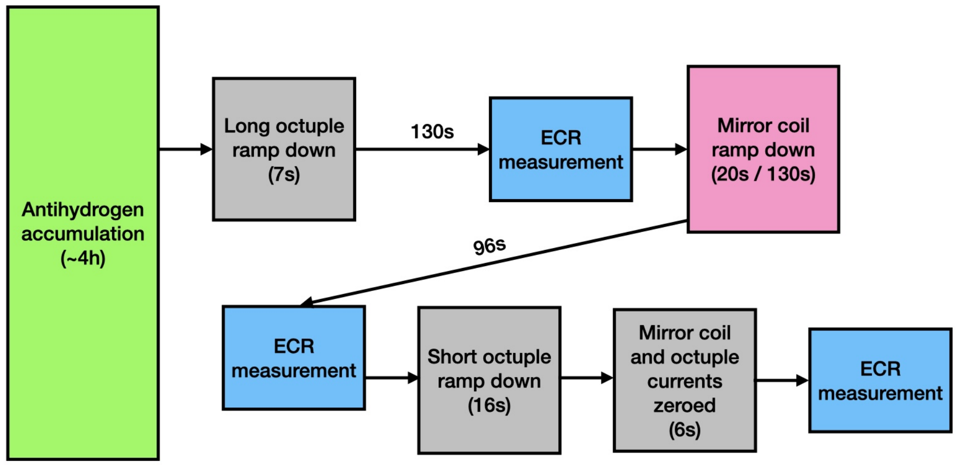

Figure 9.3: A block diagram of the measurement procedure for each run. Antihydrogen accumulation is shown in green, the main mirror rampdown in pink, other magnet ramps in grey, and ECR measurements in blue.

9.1.3 Data collection

The data collection process for the gravity paper was divided into experimental runs. Each run consisted of 3-4 hours of antihydrogen accumulation, followed by a measurement with a specific magnetic bias. A total of 11 distinct bias values were used, each repeated 6 or 7 times depending on the number of collected events. To minimize systematic effects, the bias values were interleaved in a pseudo-random sequence.

Figure 9.3 illustrates the measurement process for a single run. The antihydrogen accumulation period was chosen to achieve a good signal-to-background ratio in the mirror rampdown. Following accumulation, the long octupole magnet was ramped down, leaving the antihydrogen atoms confined only by the mirror coils and short octupole. This step avoids potential systematic bias due to the asymmetry of the long octupole around the bottom trap. However, it lowers the depth of the radial confining potential, which causes slightly more than half of the trapped antihydrogen to escape. The fraction of antihydrogen lost is independent of the magnetic bias, which allows for a calibration of the total number of antihydrogen atoms accumulated in each run.

After an ECR measurement of the initial field, the mirror coil currents are ramped down while maintaining the magnetic bias. Counting the annihilations in the up and down region during this mirror rampdown constitutes the main measurement. A second ECR scan measures the final field where the trap depth is near zero, but the magnetic bias is still maintained. The short octupole magnet is then ramped down, and finally the mirror coil currents and octupole currents are fully zeroed. A final ECR measurement is taken to monitor the background field of the external solenoid magnet.

In addition to the main dataset of 75 runs across 11 magnetic biases, two supplementary datasets were taken. An additional $\pm 10g$ calibration dataset of 12 runs was taken where the applied bias was large enough that almost all antiatoms annihilate in the same (either up or down) region. This dataset was used to determine the relative detector efficiency in the two regions, and to evaluate cross-contamination due to annihilation position mis-identification. A second auxiliary dataset of 18 runs was taken where the mirror rampdown was carried out over 130 s instead of 20 s. This was used to crosscheck the robustness of the result, and the outcome was consistent with the main dataset (see [1]).

The results of these runs are given in Table 9.1, including the number of annihilations in the up and down regions during the mirror rampdown, and the annihilations in the long octupole rampdown. The expected cosmic ray background (using the BDT background rejection of Chapter 8) was determined using dedicated cosmic ray runs with no antimatter in the system, and has been subtracted from the given counts. The z distributions of the annihilation events for each bias are shown in Figure 9.4.

9.2 Analysis procedure

The analysis strategy used in the gravity paper was to construct an *escape curve* — a plot of the probability that an antihydrogen atom escapes the trap downwards (or upwards) as a function of applied magnetic bias. We define P_{dn} as the probability that an antihydrogen atom annihilates in the down region, and the *asymmetry* as $A = 2P_{\mathrm{dn}} - 1$; the escape curve plots these quantities as a function of the applied magnetic bias. The final measured escape curve is shown in Figure 9.5. A naive estimate of $\overline{g}$ can be made from this curve by identifying the bias value for which $P_{\mathrm{dn}} = 0.5$, corresponding to the point where the magnetic bias precisely compensates for gravity, and up/down annihilations are balanced.

A more precise determination of $\overline{g}$ was obtained using simulation. For each assumed value of $\overline{g}$ and for each magnetic bias, the trajectories of many antihydrogen atoms were simulated using a magnetic field model of the ALPHA-g trap. This process generated a family of simulated escape

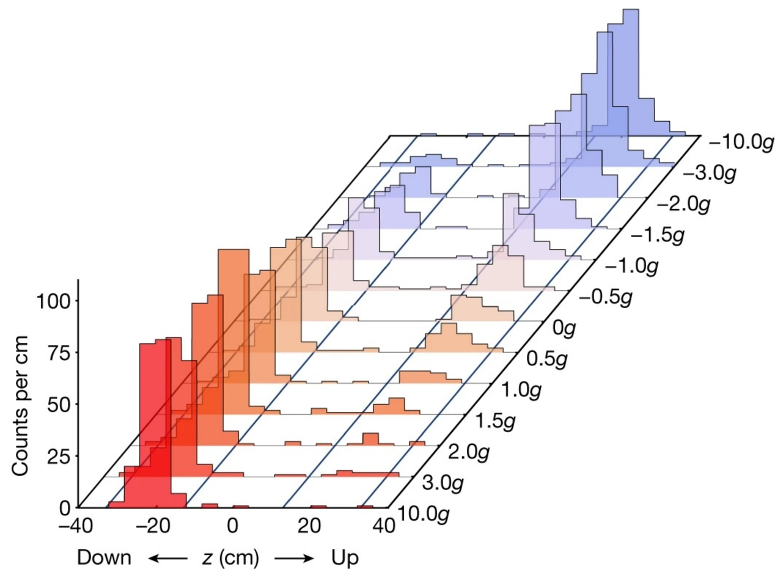

Figure 9.4: Annihilation position z distributions for different applied biases. Lines at $\pm$ 12.8 cm and $\pm$ 32.8 cm indicate the z cuts used to define the up and down regions. Figure from [1].

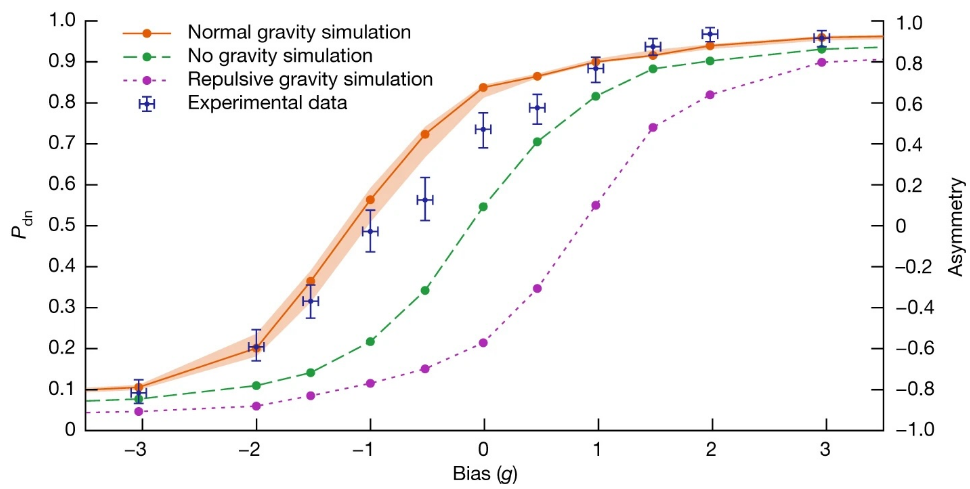

Figure 9.5: The measured escape curve, the probability that an antihydrogen atom annihilates in the down region plotted for different magnetic biases. The data is compared to simulations described in Section 9.2.3. Figure from [1].

Table 9.1: Measured antihydrogen annihilation counts for each magnetic bias. Bias of ±10 indicates calibration dataset, and 130s indicates cross-check with slower mirror rampdown. Event counts have expected cosmic ray background subtracted, and are subject to data selection cuts described in Section 9.2.1. LOc indicates annihilations during the long octupole ramp down. Adapted from [1].

Nominal bias (g)	**Number of trials**	N_{up} (events)	N_{dn} (events)	LOc (events)
−3.0	7	151.7	16.5	199.2
−2.0	7	128.7	33.5	168.2
−1.5	6	128.9	57.7	192.0
−1.0	7	69.7	62.5	183.2
−0.5	7	55.7	67.5	201.2
0	7	36.7	94.5	144.2
0.5	7	36.7	124.5	177.2
1.0	7	17.7	119.5	185.2
1.5	6	13.9	180.7	234.0
2.0	7	6.7	163.5	228.2
3.0	7	7.7	147.5	199.2
−10.0	6	142.9	0.7	169.0
10.0	6	−0.1	185.7	213.0
−2.0 (130 s)	6	22.6	145.2	219.0
−1.0 (130 s)	6	70.4	85.2	249.0
0 (130 s)	6	125.4	26.5	239.0

curves, each corresponding to a different gravitational acceleration. These curves were then compared to the experimental data to quantify compatibility with various $\overline{g}$ hypotheses. Three simulated escape curves are included in Figure 9.5.

To extract P_{dn} for each bias setting from the experimental data, a Bayesian likelihood analysis was employed. This analysis was then extended to estimate the value of $\overline{g}$ by directly incorporating the simulated escape curves.

This section is organized as follows. Section 9.2.1 describes the data selection criteria used to obtain the annihilation counts presented in Table 9.1. Section 9.2.2 outlines the Bayesian methodology used to derive the values of P_{dn} and the final value of $\overline{g}$. Section 9.2.3 describes the antihydrogen trajectory simulations and magnetic field models used to generate the simulated escape curves.

9.2.1 Data selection

To maximize sensitivity to the few antihydrogen annihilation events while minimizing contamination from the cosmic ray background, selection cuts were applied to both the z position of the annihilation vertex and the event time.

The accepted z range for the *down* region is $-32.8\,\mathrm{cm} < z < -12.8\,\mathrm{cm}$, and for the *up* region is $12.8\,\mathrm{cm} < z < 32.8\,\mathrm{cm}$. In this coordinate system, the centre of mirror coils A and G are located at $-12.8\,\mathrm{cm}$ and $12.8\,\mathrm{cm}$ respectively, which define the inner boundaries of the selection regions. The outer boundaries were chosen based on the annihilation distribution of the $\pm 10g$ calibration runs. These spatial cuts are indicated in Figure 9.4.

A time cut was also applied to exclude events outside the main measurement window. Only annihilations occurring during the second half of the 20 s mirror coil rampdown were accepted, since the first half of the rampdown did not contain any annihilations significantly above the background rate. Although there was a significant number of antihydrogen annihilations after the end of the mirror coil rampdown, these were excluded. This time selection is illustrated for the $0g$ bias trials in Figure 9.6.

9.2.2 Bayesian analysis

A Bayesian approach is taken to determine the most likely values of P_{dn} for each bias given the observed annihilation counts. This method is outlined here and described more fully in [159]. Other methods were also explored [76] and yielded similar results.

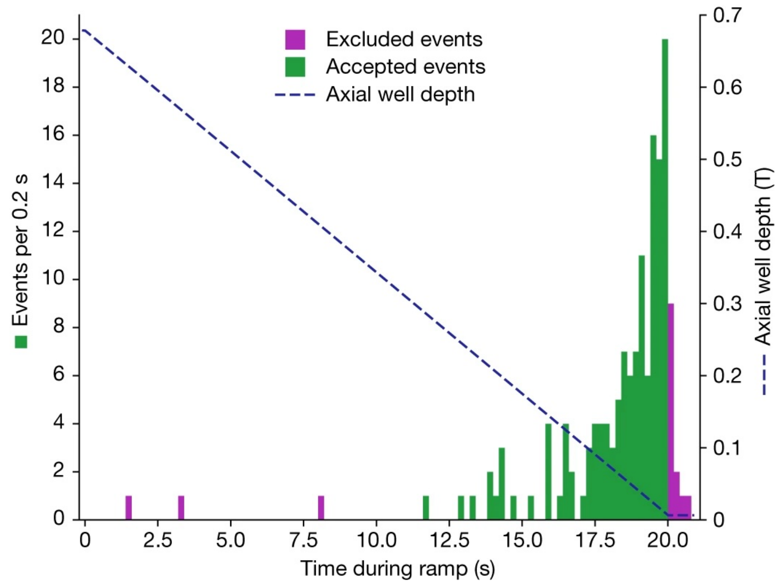

Figure 9.6: Time distribution of annihilations during the mirror coil rampdown. Only events in the window 10 s$< t <$20 s are accepted. Figure from [1].

For a given bias index i, the counts in the up and down regions $n_{\mathrm{up},i}$ and $n_{\mathrm{down},i}$ during the mirror coil rampdown, as well as during the long octupole rampdown $n_{\mathrm{LOc},i}$, are assumed to be independent Poisson-distributed random variables. The means of these distributions are calculated from the model parameters given in Table 9.2. These parameters include the underlying downwards annihilation probabilities p_i for each bias.

Given values for the model parameters, the probability $P(\mathcal{O}|\boldsymbol{\rho}, \boldsymbol{\eta}, \boldsymbol{\lambda}, \beta, \boldsymbol{p})$ of making a set of observations $\mathcal{O}$ (comprising up, down, and long octupole counts from each bias) can be calculated.

The prior distributions for the probabilities p_i are set to uniform probabilities between 0 and 1. The priors for the detection efficiencies and cross-contamination $\boldsymbol{\eta}$ were based on the $\pm 10g$ calibration dataset. Broad uniform priors were assigned to λ_i and β. The priors for the detector background rates ρ_{up}, ρ_{down} and ρ_{LOc} were based on cosmic ray measurements. A global prior probability $P(\boldsymbol{\rho}, \boldsymbol{\eta}, \boldsymbol{\lambda}, \beta, \boldsymbol{p})$ is then assigned to each set of parameters based on the independent prior distributions.

Using Bayes theorem, the posterior parameter probability distribution given a set of observations is:

Table 9.2: Model parameters used in the Bayesian analysis of the gravity paper dataset.

Parameters	**Description**
$\rho_{\text{up}},\ \rho_{\text{down}}$	Cosmic ray background rate in the up and down regions.
ρ_{LOc}	Cosmic ray background rate during long octupole ramp.
$\eta_{u,u},\ \eta_{d,d}$	Detection efficiency in the up and down regions.
$\eta_{u,d},\ \eta_{d,u}$	Cross-contamination between the up and down regions.
β	Proportionality constant between long octupole and mirror rampdown annihilations.
λ_i	Annihilations during long octupole ramp at bias i.
p_i	Down region annihilation probability for bias i.

$$P(\boldsymbol{\rho},\boldsymbol{\eta},\boldsymbol{\lambda},\beta,\boldsymbol{p}|\mathcal{O}) \propto P(\mathcal{O}|\boldsymbol{\rho},\boldsymbol{\eta},\boldsymbol{\lambda},\beta,\boldsymbol{p})\ P(\boldsymbol{\rho},\boldsymbol{\eta},\boldsymbol{\lambda},\beta,\boldsymbol{p}) \tag{9.2.1}$$

The marginalized posterior distribution for each p_i is then obtained by integrating over the other variables (nuisance parameters):

$$P(p_i|\mathcal{O}) \propto \int d\boldsymbol{\rho} \int d\boldsymbol{\eta} \int d\boldsymbol{\lambda} \int d\beta \int_{j\neq i} dp_j\ P(\mathcal{O}|\boldsymbol{\rho},\boldsymbol{\eta},\boldsymbol{\lambda},\beta,\boldsymbol{p})\ P(\boldsymbol{\rho},\boldsymbol{\eta},\boldsymbol{\lambda},\beta,\boldsymbol{p}) \tag{9.2.2}$$

These distributions of $P(p_i|\mathcal{O})$ are determined via the Markov Chain Monte Carlo (MCMC) method, using the STAN package [159]. The maximum of the posterior probability distribution $P(p_i|\mathcal{O})$ is taken as the best estimate of P_{dn} in Figure 9.5.

To determine the most probable value of $\overline{g}$, this analysis was extended by incorporating the simulation results described in Section 9.2.3. The escape curves generated via simulation (which will be shown in Figure 9.7) are given by the same functional form differing by a horizontal shift. These simulated escape curves were aligned and interpolated in order to determine the simulated down annihilation probability $p_{\text{sim}}(\phi)$ as a function of the magnetic bias ϕ, in the absence of gravity.

In this extended analysis, the independent parameters p_i were replaced with the simulated probability $p_{\text{sim}}(\phi_i+a_g)$ evaluated at each bias value, with an additional offset a_g due to the effect of gravity. The posterior distribution of a_g was then determined using the MCMC method by integrating over all other parameters:

$$P(a_g|\mathcal{O}) \propto \int d\boldsymbol{\rho} \int d\boldsymbol{\eta} \int d\boldsymbol{\lambda} \int d\beta \; P(\mathcal{O}|\boldsymbol{\rho}, \boldsymbol{\eta}, \boldsymbol{\lambda}, \beta, a_g) \; P(\boldsymbol{\rho}, \boldsymbol{\eta}, \boldsymbol{\lambda}, \beta, a_g) \tag{9.2.3}$$

The value of a_g with the maximum posterior probability $P(a_g|\mathcal{O})$ is quoted as the final measurement of $\overline{g}$.

9.2.3 Escape curve simulation

To quantify the compatibility of the ALPHA-g measurements with different gravitational scenarios, the observed annihilation distributions were compared to trajectory simulations of antihydrogen atoms escaping from ALPHA-g under different magnetic biases. As outlined in the previous section, the extraction of a value of $\overline{g}$ relies directly on simulation, and so the accuracy of these simulations is central to the experiment.

One crucial input to the simulations is a precise model of the magnetic field in the ALPHA-g trap region. The field model was constructed to match measurements as closely as possible. It is divided into independent models of the magnetic field produced by each magnet: the main external solenoid, and each individual octupole and mirror coil.

The external solenoid field model was initially computed based on the design geometry, and updated to reflect measurements taken with an NMR probe. This field was then perturbed by the insertion of the inner cryostat and internal magnets. The resulting modified field was measured on-axis using ECR (Section 9.1.2), and the field model updated to reflect these in-situ measurements.

The magnetic fields of the octupole and mirror coils were calculated based on the winding geometry measured during construction. These field models were also fine-tuned to match both ECR and magnetron-metry. To reproduce the time evolution of the magnetic fields during magnet ramps, the current histories measured using DCCTs were applied in the simulation. The combined field model matched all ECR and magnetron-metry measurements, both offline and those taken during the measurement procedure, to within 2×10^{-5} T, and to within 1×10^{-6} T in the region under the mirror coils.

The total magnetic field was precomputed and stored in a grid with 0.5 mm spacing, which was then interpolated during the simulation. To model the antihydrogen accumulation cycles, atoms were initialized in a cylinder near the bottom of the trap and allowed to evolve for a random duration between 0 and 4 h. The time evolution followed a stepping algorithm with step sizes of 1 µs or 3 mm, whichever was smaller. The atoms were

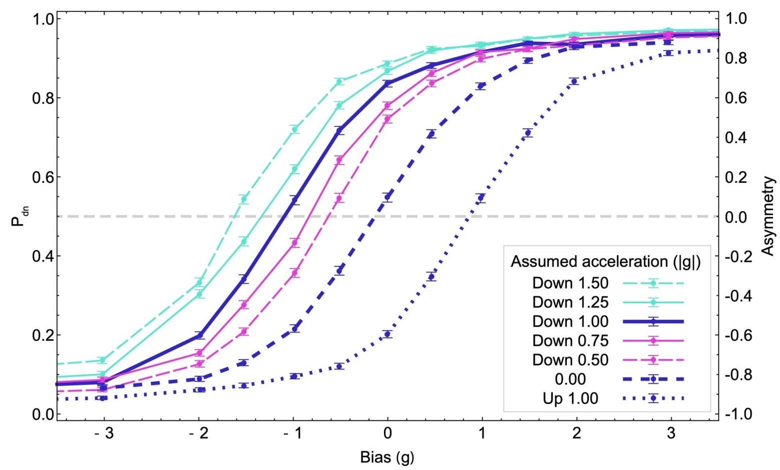

Figure 9.7: Escape curves generated using antiatom trajectory simulations. Each line corresponds to a different assumed value for the antihydrogen gravitational acceleration $\overline{g}$, and each data point represents a simulation with a different magnetic bias applied during the mirror coil ramp down. Figure from [1].

initialized with a 50 K thermal energy distribution and a principal quantum number of 30, then allowed to radiatively cascade to the ground state. After this process, 6726 antihydrogen atoms remained trapped.

These surviving atoms were then used as initial conditions for every mirror coil rampdown simulation. Their trajectories were evolved using the stepping algorithm and time-dependent field models described above, and their annihilation positions and times were recorded. This process was repeated for each of the 11 applied magnetic biases and for a range of assumed values of $\overline{g}$. The resulting escape curves are shown in Figure 9.7.

9.3 Measurement uncertainties and systematics

To contextualize my contributions within the broader ALPHA-g measurement program, it is important to discuss the sources of uncertainty and systematic effects in the gravity paper. There are three types of uncertainty quoted:

- **Magnetic bias uncertainty:** the uncertainty in the magnetic field applied to the antihydrogen atoms during the mirror coil rampdown, which affects the x-axis values of the escape curve.

- **Statistical uncertainty:** the uncertainty in the measured probability of downward annihilation at each magnetic bias, reflected in the y-axis error bars of the escape curve.
- **Simulation uncertainty:** the uncertainty arising from the modelling of antihydrogen trajectories and magnetic fields used to generate the simulated escape curves.

These three components contribute respectively to the horizontal, vertical, and model-based uncertainties shown in Figure 9.5. Each is propagated into the final quoted uncertainty on the gravitational acceleration $\overline{g}$.

9.3.1 Magnetic bias uncertainties

For each trial, the currents in mirror coil A and G were ramped down in a way intended to maintain a constant magnetic bias between the two, as described in Section 9.1.1. However, the actual magnetic bias experienced by the antihydrogen atoms may have deviated from the nominal value. Uncertainties reflecting how well the true magnetic bias is known define the horizontal error bars in Figure 9.5.

The dominant contribution arises from the precision with which the magnetic field difference between the two mirror coils can be measured using the ECR technique. As shown in Figure 9.2, the cyclotron frequency resonance has a finite width. The standard deviation of a Gaussian fit to each spectrum defines the **ECR spectrum width uncertainty**, which was found to be $0.07g$ across all measurements.

A second contribution arises from trial-to-trial variability in the ECR measurements of the magnetic bias. The standard deviation of the measured bias values defines the **repeatability uncertainty**, with a value of $0.014g$.

Due to the influence of background field gradients, the maxima of the magnetic field do not lie directly under the geometric centres of the mirror coils. The peak positions were determined in each configuration via a quadratic fit to ECR measurements at different axial positions. The uncertainty in this fit contributes a **peak field fit uncertainty** of $0.009g$.

Another contribution stems from persistent currents in the superconducting mirror coils, which create residual magnetic fields that decay slowly after the rampdown ends. Because there is a delay of approximately 96 s between the end of the mirror rampdown and the second ECR measurement, the measured magnetic field does not exactly match the field during the rampdown. Offline ECR measurements showed a decay of 6×10^{-5} T over the

96 s following the rampdown, with a decay rate difference between the two mirror coils leading to a **field decay asymmetry uncertainty** of $0.02g$.

The magnetic bias is not constant during the mirror rampdown. Using the field model described in Section 9.2.3, the magnetic bias felt by each antihydrogen atom is calculated based on its annihilation time. The mean of these values defines the x-axis position of each data point in Figure 9.5, and the standard deviation defines a **bias variation uncertainty** of $0.02g$ across all measurements.

This bias determination for each antiatom depends on the accuracy of the magnetic field model, particularly for determining the bias off-axis. The level of agreement between ECR/magnetron-metry measurements and the field model defines a **field modelling uncertainty**, estimated at $0.05g$.

9.3.2 Annihilation counting uncertainties

The second major type of uncertainty arises when determining the probability that antihydrogen atoms annihilate in the up or down detector region at each magnetic bias. This defines the vertical error bars in Figure 9.5.

There are two main contributions to this uncertainty. The first uncertainty is statistical uncertainty in the main measurement dataset, with only around 200 annihilation events for each magnetic bias value. This statistical **finite data size** uncertainty is estimated at $0.06g$ across all measurements.

The second contribution is from uncertainty on the detector efficiency asymmetry, the difference in annihilation detection efficiency between the up and down regions. This is calculated using the $\pm 10g$ calibration dataset, where the applied biases are large enough to force all antihydrogen atoms to one of the two regions. However, less than 200 rampdown annihilations were detected for each calibration bias, leading to a statistical **detector efficiency calibration uncertainty** estimated at $0.12g$.

9.3.3 Simulation uncertainties

While the mirror rampdown measurements agreed well with simulation, there were discrepancies between measurement and simulation in other time periods. In the measurement, fewer atoms survived the mirror coil rampdown, more atoms escaped during the post-rampdown waiting/ECR period, and fewer atoms were released in the final short octupole rampdown.

The trajectory simulations were repeated with a series of systematic perturbations. Some of these perturbations were able to explain the differences between measurement and simulation. The largest changes in the resulting P_{dn} values were used to define the simulation error band in Figure 9.5.

The change to the fraction of antiatoms surviving the mirror coil rampdown was found to be consistent with different initial antihydrogen energy distributions between simulation and experiment. Modifying the initial antihydrogen energy distribution in simulation was found to produce an **antihydrogen initial energy distribution uncertainty** of $0.03g$.

Field model perturbations compatible with ECR and magnetron-metry measurements were explored. These include dipole, quadrupole, sextupole, and octupole perturbations with zero on-axis component applied to the bottom half of the trap field model, as well as misalignment of the axes of the mirror coils, octupole magnets, and external solenoids. These produced the dominant uncertainty of the measurement, a **magnetic field modelling uncertainty** of $0.16g$.

Other simulation perturbations included adding an artificial mechanism for mixing the parallel and transverse energy of the antihydrogen atoms, which was consistent with the increased annihilation count in the post-rampdown waiting period. A final change investigated was the addition of physical vibrations of the magnets. These did not produce a significant effect on the escape curve.

9.3.4 Summary of experimental uncertainties

A summary of the experimental uncertainties considered in the gravity paper is presented in Table 9.3. It should be noted that most of these uncertainties, including the dominant uncertainty, reflect our limited understanding of the magnetic field during the mirror coil rampdown. These will be significantly reduced in future measurements using the precision region of the ALPHA-g trap, which was designed to reduce the magnitude and possible asymmetries of the magnetic fields involved. The other main uncertainties stem from limited annihilation statistics, especially in the calibration dataset. These are already being mitigated by improvements in the antihydrogen trapping rate in ALPHA-g.

9.4 Future measurements and the impact of this work

Although only partially completed at the time of its publication, the work described in Chapters 5-8 has had a substantial impact on the gravity paper.

Table 9.3: Summary of uncertainties affecting the measurement of $\overline{g}$ in the gravity paper. Adapted from [1].

Uncertainty	**Value (units of g)**
Uncertainties in bias determination	
ECR spectrum width uncertainty	0.07
Field modelling uncertainty	0.05
Field decay asymmetry uncertainty	0.02
Bias variation uncertainty	0.02
Repeatability uncertainty	0.014
Peak field fit uncertainty	0.009
Annihilation counting uncertainties	
Detector efficiency calibration uncertainty	0.12
Finite data size uncertainty	0.06
Simulation uncertainties	
Magnetic field modelling uncertainty	0.16
Initial energy distribution uncertainty	0.03

This is evidenced by the list of dominant uncertainties in the previous section, which notably excludes any contributions due to misidentification of the annihilation region or contamination from the cosmic ray background, issues which were mitigated through these improvements. The full implementation of the improved tracking algorithms and TOF-based background rejection was completed for the 2024 data taking run, leading to improvements in multiple metrics.

A significant impact of this work will be on upcoming future precision measurements using the complete ALPHA-g apparatus. This section will discuss future directions for ALPHA-g as the experiment progresses towards a more precise determination of $\overline{g}$, and highlight the role played by the improved tracking and background rejection. Following the medium-term plan to use the precision trap region of ALPHA-g, there are some more speculative alternative approaches which are directly enabled by these improvements to the background rate and position resolution.

Section 9.4.1 will discuss and attempt to quantify the impact of this work on the published gravity paper. Section 9.4.2 will outline ongoing and short-term improvements using only the ALPHA-g bottom trap. Section 9.4.3 will discuss medium-term measurements which will be taken using the full ALPHA-g trap. Finally, Section 9.4.4 will outline potential alternative approaches, and the future of antihydrogen gravity measurements.

9.4.1 Impact on the gravity paper

My work contributed to the gravity paper in two key ways: improvement of the annihilation vertex finding algorithm, and development of background rejection techniques using the BSC.

As discussed in Section 5.4, the annihilation z position resolution for the more poorly-constrained vertices (the larger of two values) was originally 14.6 cm. Given the 25.6 cm distance between the mirror coil centres, this resolution was insufficient to reliably distinguish between up and down annihilations, and would have necessitated a careful treatment of vertex misidentification as a systematic uncertainty. After the introduction of my iterative track finding routine and subsequent refinements, the resolution was improved to around 3 cm for the more poorly-constrained vertices. This greatly reduced the number of misidentified vertices, such that vertex misidentification was no longer a significant source of uncertainty.

In addition, my work on the BSC was crucial for background rejection in the gravity paper. While the BDT background rejection algorithm was being developed, an online cut based on the number of bar hit clusters served as the primary tool for background suppression. Given the trapping rate of less than 2 antihydrogen per 4-minute accumulation cycle, this online cut was instrumental for antihydrogen trapping optimization, which requires accurately measuring the number of trapped antiatoms as the accumulation sequences are incrementally adjusted.

Furthermore, BSC variables were crucial in the machine learning background rejection algorithm used in the gravity paper. Without the BSC, the background rejection performance would have been significantly worse. Although a background rejection BDT was never trained using only the TPC (i.e. without the BSC), we can attempt to estimate the performance of such a classifier. As the online BSC cluster cut reduced the background rate by a factor of 10 (shown in Section 8.5), it is reasonable to estimate that a classifier trained without the BSC would have a similarly worse performance, yielding a background rate of around 2 Hz instead of the 0.19 Hz achieved in the final analysis.

Impact estimate via simulation

To evaluate the quantitative impact of the improvements introduced in this work, a toy model simulation was developed to reproduce the annihilation detection process. Annihilation positions were sampled from Gaussian distributions based on the $\pm 10g$ data. These positions were convoluted with

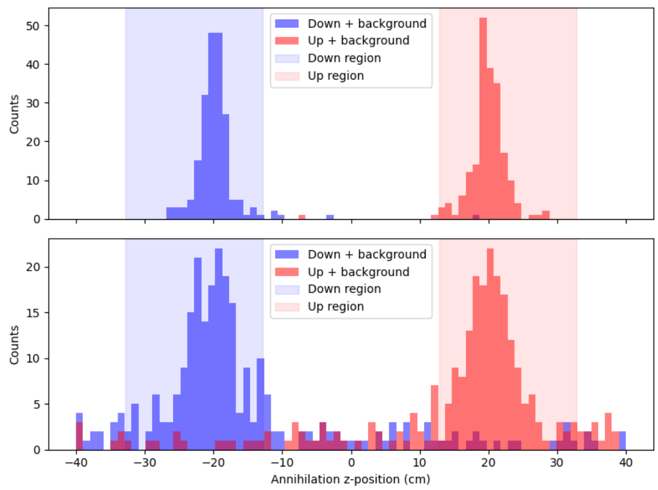

Figure 9.8: 200 simulated up annihilations and 200 simulated down annihilations, under different detector performance assumptions. Top: Gravity paper performance of $\sigma_{\text{inner}} = 1.06\,\text{cm}$, $\sigma_{\text{outer}} = 3.02\,\text{cm}$, $\lambda_{\text{bkg}} = 0.19\,\text{Hz}$. Bottom: Hypothetical worse performance of $\sigma_{\text{inner}} = 2.80\,\text{cm}$, $\sigma_{\text{outer}} = 14.58\,\text{cm}$, $\lambda_{\text{bkg}} = 2\,\text{Hz}$.

a double-Gaussian detector resolution function using the σ_{inner} and σ_{outer} parameters from Figure 5.16. Cosmic ray background events, uniformly distributed in z, were added based on a Poisson distribution with the measured average background rate. The annihilation z distributions of a typical simulation are shown in Figure 9.8, comparing the analysis used in the gravity paper with a hypothetical analysis lacking the vertex resolution and background rate improvements developed in this work.

Two key effects arise from the degraded detector performance. Firstly, the fraction of events reconstructed in the correct annihilation region decreases from $\sim 99\%$ to $\sim 71\%$, reducing measurement statistics. Secondly, the fraction of events reconstructed in the wrong annihilation region increases from $\sim 0\%$ to $\sim 9.5\%$, leading to cross-contamination.

A true P_{down} value was assumed as an input for each simulation. Roughly 200 events were simulated per trial: the number of events simulated in the down and up regions were drawn from Poisson distributions with $\lambda = 200 \cdot P_{\text{down}}$ and $\lambda = 200 \cdot (1 - P_{\text{down}})$ respectively. The z position of each event was sampled from the corresponding Gaussian, and convoluted with the detector performance. After each trial, the events in the up and down regions were

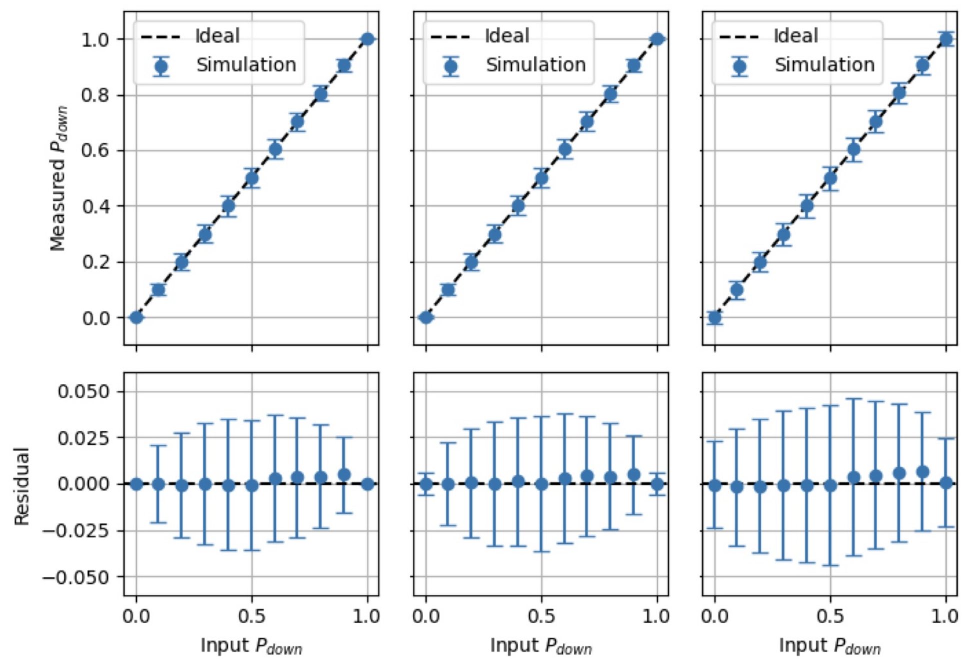

Figure 9.9: The P_{down} measured from simulated data, as a function of the simulation input P_{down} values. (Left:) With perfect detector resolution and zero background. (Centre:) With the detector resolution of 1.06 cm inner Gaussian and 3.02 cm outer Gaussian, and background rate 0.19 Hz, used in the gravity paper. (Right:) With the detector resolution of 2.80 cm inner Gaussian and 14.58 cm outer Gaussian, and background rate 2 Hz, in the absence of this work.

counted, and a measured $P_{\text{down}} = n_{\text{down}}/(n_{\text{up}} + n_{\text{down}})$ was computed.

To quantify the purely statistical uncertainty, for each input $P_{\text{down}} = [0, 0.1, ..., 0.9, 1]$, 5000 trials were performed with no cosmic ray background background and perfect detector resolution. The mean and standard deviation of the measured P_{down} values are shown in the left panel of Figure 9.9. The largest standard deviation is taken as the purely statistical uncertainty $u[P_{\text{down}}]$ (stat.) $= 0.035$, corresponding to the statistical uncertainty $u[\overline{g}] = 0.06g$ quoted in Section 9.3.2.

To include the impact of the finite detector resolution and the cosmic ray background rate into this uncertainty, these simulations were repeated with the resolution and background rate values from the gravity paper (Figure 9.9 centre), and with the hypothetical worse detector performance without my

Input parameters			Simulation results			
σ_{inner} [cm]	σ_{inner} [cm]	Bkg. [Hz]	$u[P_{\text{down}}]$ (stat.)	$u[P_{\text{down}}]$ (stat. & mis.)	$u[P_{\text{down}}]$ (mis.)	$u[\overline{g}]$ (mis.)
0	0	0	0.035	0.035	0	0
1.06	3.02	0.19	0.035	0.036	0.006	0.01g
1.06	3.02	2.0	0.035	0.038	0.019	0.03g
2.80	14.58	0.19	0.035	0.039	0.016	0.03g
2.80	14.58	2.0	0.035	0.042	0.024	0.04g

Table 9.4: Uncertainty values derived from toy model simulation, for the vertex resolution and background rate values from before and after this work. (stat.) indicates Poisson statistical uncertainty, while (mis.) indicates uncertainty from misidentifying the annihilation region. $u[\overline{g}]$ is calculated from $u[P_{\text{down}}]$ by a scaling factor.

improvements (Figure 9.9 right). The largest standard deviations are given as the $u[P_{\text{down}}]$ (stat. & mis.) values in Table 9.4, which summarizes the results of these simulations.

To examine the uncertainty due purely to annihilation region misclassification, i.e. not including the statistical uncertainty, these simulations were repeated where the number of events in the bottom and top regions were fixed exactly to $N_{\text{bot}} = 200 \cdot P_{\text{down}}$ and $N_{\text{top}} = 200 \cdot (1 - P_{\text{down}})$ instead of being drawn from a Poisson distribution. The results from these simulations are reported as the $u[P_{\text{down}}]$ (mis.) values in Table 9.4.

We can conclude that the uncertainty on P_{down} was marginally improved by this work, which prevented an uncertainty on the measured gravitational acceleration of around $u[\overline{g}] = 0.04\text{g}$. It is worth noting that this toy simulation did not include any correction for the detector efficiency asymmetry in the two regions. As we have seen in Section 9.3.2, the uncertainty based on the measurement of the detector efficiency asymmetry using the the $\pm 10g$ calibration samples is dominant over the purely statistical uncertainty. Degraded detector performance would also have a large impact on our ability to measure the detector efficiency asymmetry, leading to a further increase in uncertainty not quantified here.

9.4.2 Repeated measurements in the same apparatus

There are at least four short-term avenues of improvement available using the trap apparatus which was used in the gravity paper, the ALPHA-g bottom trap. Proof of principle demonstrations for three of these were performed

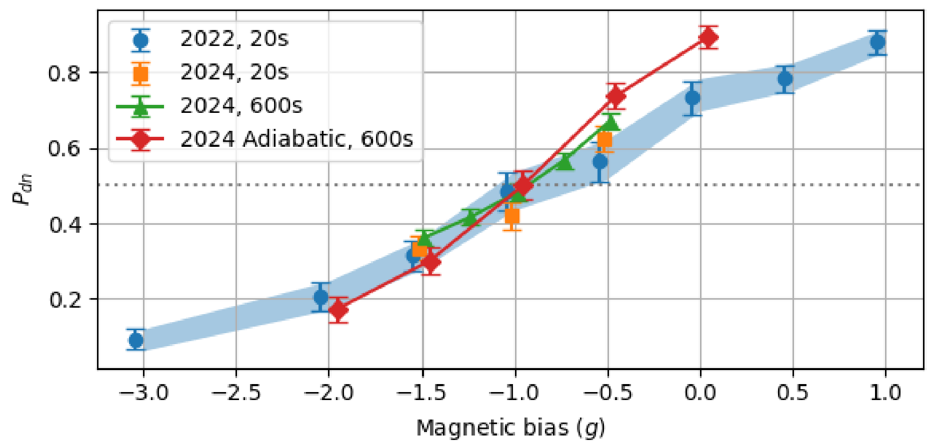

Figure 9.10: Preliminary results from the 2024 data-taking run, showing improved trapping rate only (orange), slower rampdown (green), and adiabatic cooling (red), compared to gravity paper data (blue).

during the 2024 data-taking run: improved antihydrogen production using Be^+, slower mirror coil rampdown speed, and adiabatic expansion. Preliminary results from these methods are shown in Figure 9.10. All new results are compatible with the gravity paper results. The adiabatic expansion yields a slightly steeper escape curve, which corresponds to improved precision in the determination of $\overline{g}$. The fourth avenue of improvement, laser cooling of the antihydrogen atoms, will be attempted in the current or upcoming experimental campaigns.

Improved production rate

As described briefly in Section 3.5.2, a positron plasma cooling method using laser-cooled beryllium ions was developed and implemented after the 2022 measurement. The reduced positron temperatures leads directly to a colder initial temperature distribution for the produced antihydrogen. Consequently, a higher fraction of the antihydrogen atoms have kinetic energies below the trap depth, significantly improving trapping rates from < 2 per cycle to ~ 20 per cycle in the 2024 data taking run.

This has multiple advantages. Firstly, experiments can be conducted more rapidly, reducing susceptibility to long-timescale systematic drifts. Secondly, this enhances the signal-to-background ratio by allowing the experiment to be performed only once per bias. Thirdly, increased annihilation counts per trial reduces statistical uncertainties, which is especially relevant for the detector calibration dataset. Finally, a shorter experiment time allows for a higher number of systematic checks to be performed, such as to validate various aspects of the antihydrogen trajectory simulation.

Mirror rampdown speed

During the mirror coil rampdown, when the confining potential of a mirror coil is lowered below the kinetic energy of an antihydrogen atom, it does not instantaneously escape the trap. It continues on its trajectory, and escapes only if it encounters the magnetic field minimum under the mirror coil. This is highly likely given enough time to explore the phase space. However, if the potential of the second mirror coil drops below the antiatom's kinetic energy before it has a chance to escape through the first, the atom may escape in either direction. This weakens the correlation between escape direction and gravitational bias.

To mitigate this effect and improve directional sensitivity, a slower mirror coil rampdown can be employed. Increasing the ramdpown duration from the 20 s used in the gravity paper improves the likelihood that the antiatom escapes in the preferential direction, leading to a steeper escape curve.

This technique is only feasible as a result of the lower background rates provided by the BSC TOF analysis. Taking advantage of this, a 600 s rampdown was attempted in 2024. However, this technique did not immediately yield a noticeably steeper escape curve. Further trajectory simulations and systematic investigations will be important to understand the level of improvement expected from this change.

Adiabatic expansion cooling

The temperature of the trapped antihydrogen plays a large role in the gravity experiment. Colder antiatoms have reduced thermal motion, allowing their trajectories to be more strongly influenced by gravity and applied magnetic bias, steepening the escape curve.

There is no intrinsic cooling mechanism for trapped antihydrogen. The implementation of adiabatic cooling (described in Section 3.5.4) provides a mechanism to reduce the temperature of the antihydrogen atoms. Practically, however, there are losses in the adiabatic expansion process, and so decreasing the antihydrogen temperature comes at the cost of reduced statistics. Adiabatic cooling has been demonstrated in ALPHA-2 [121], and in 2024, it was first demonstrated in ALPHA-g, producing a marginally steeper escape curve.

Laser cooling

Laser cooling (see Section 3.5.4) is another promising approach to reduce the kinetic energy of trapped antihydrogen, and has been successfully utilized in the ALPHA-2 apparatus. There, laser cooling was typically applied over several hours. The time-of-flight (between a laser pulse and annihilation) during spectroscopy experiments provides a method to estimate the antihydrogen temperature, indicating a temperature of less than 35 mK [160]. When combined with adiabatic expansion, significantly lower temperatures can be achieved. Preliminary attempts to combine these strategies have been made in ALPHA-2.

Efforts are currently underway to implement laser cooling in the ALPHA-g apparatus. The combination of laser cooling and adiabatic cooling promises to provide much cooler antihydrogen for the gravity experiment, improving the steepness of the escape curve and the precision of future determinations of $\overline{g}$.

Laser cooling offers a further benefit to the gravity experiment, beyond a steeper escape curve. The cooling laser in ALPHA-g decreases the average axial velocity of the antihydrogen atoms; however, the axial and transverse kinetic energy mix over time, and so this also leads to a decrease in the transverse temperatures. As a result, the antiatoms will be more radially confined and stay closer to the central trap axis, which is the deepest part of the trapping potential. This is beneficial since the magnetic fields are better understood closer to the trap axis: fringe effects due to the magnet windings are minimized, and the field can be better measured using ECR. With antihydrogen trajectories more constrained to the centre of the trap, the trajectory simulations will be more accurate and present smaller uncertainties. The magnetic bias felt by each antiatom as it exits the trap will also be better known, since only the on-axis magnetic bias is measured throughout the rampdown.

9.4.3 Measurements using the precision trap region

While the gravity paper measurements were performed using only the bottom trap, the most precise measurements will be performed using the full ALPHA-g apparatus. As discussed in Section 3.4.6, the ALPHA-g trapping system is divided into three sections: the top, bottom, and precision magnetic traps. The bottom and top traps are reflected copies of each other, with stronger radial confinement provided by short octupole magnets. The central precision trap has a lower octupole field from only the long octupole magnet, leading to smaller magnetic field uncertainties. Furthermore, the magnet layout is symmetrical around the centre of the precision trap, minimizing systematic uncertainties of the magnetic field.

To enable measurements in the precision trap, a method of transferring antihydrogen atoms between traps must be developed. This will involve slow manipulations of multiple mirror coil currents in order to adiabatically translate the magnetic trapping potential, using a procedure which must be developed.

The improved background rejection and vertex resolution resulting from this work could play a crucial role in the development of this transport technique. Given the antihydrogen trapping rate of around 20 antiatoms per 4-minute cycle, identifying antihydrogen losses requires sensitivity to single annihilations. Furthermore, the improved vertex reconstruction will be very helpful to localize any antihydrogen losses, helping to optimize the transport procedure.

A laser- and adiabatically-cooled antihydrogen atom with 1 mK of kinetic energy travels at approximately 4 m/s, corresponding to an axial bounce time of 100 ms in a 25 cm trap. To avoid unintentional heating or antiatom loss, the mirror coil current adjustments must be slow relative to this timescale. Consequently, this transport process will take several seconds. The improved background rejection developed in this dissertation will be essential for diagnosing losses during antihydrogen transfer, by allowing the transport process to be improved using only the trapped atoms from a single accumulation cycle.

9.4.4 Alternative measurement schemes

All of the improvements discussed thus far retain the same core measurement principle: antihydrogen is confined in a magnetic trap, and the trapping potentials are slowly lowered, allowing the antiatoms to escape under the influence of gravitational and magnetic biases. This rampdown method has proven effective, and will remain the core ALPHA-g measurement principle in the medium-term. However, alternate techniques do exist, each with distinct advantages and systematic uncertainties.

After the rampdown technique has been implemented in the precision trap region, the ALPHA collaboration will explore these alternative measurements schemes. Independent confirmation of the gravitational acceleration

$\overline{g}$ using a different technique will be critical for validating and establishing the robustness of the ALPHA-g results. This will be particularly important to the broader community in the event that the other antimatter gravity experiments at CERN (i.e. GBAR and AEGIS) have not published results.

Even if many underlying components in ALPHA-g are the same, using a new technique will inevitably lead to a different set of systematic measurement uncertainties. A successful cross-check using an alternative method would lend strong credibility to the ALPHA-g results. For this reason, it is important to consider other measurement schemes which could be implemented using the ALPHA-g apparatus.

This section will discuss some alternative and speculative possible measurement schemes. These are preliminary ideas, and their realization will require further considerations and developments, but they are presented here to stimulate the discussion of these approaches. Many of these techniques are directly enabled by my work on the vertex reconstruction and background rejection, and this section will highlight the importance of my work to each potential measurement. It concludes with a discussion of the future of antihydrogen gravity research.

Smaller trapping regions

The improved vertex resolution may allow a class of measurements where the separation between the up and down regions is smaller. Recall from Section 3.4.6 that both the ALPHA-g bottom and top traps are constructed with many independent mirror coils. An escape-curve-type measurement could be performed using antihydrogen trapped between any two mirror coils, for example between mirrors A and B; the distance between these mirror coils is much smaller than the distance between mirrors A and G. As such, the vertex position resolution improvements presented here would be crucial for such a measurement.

These measurements could be used to test some of the systematics of the main measurement. The results from escape curves generated from many small regions within the ALPHA-g trap could be compared to antihydrogen trajectory simulations. This could serve as a cross-check to validate the magnetic field models used in the simulation, independently studying the fields in each small region.

Magnetic heating

An alternative method for constructing an escape curve involves inducing antihydrogen loss not by lowering the confining potentials, but by heating the antihydrogen atoms until they have sufficient energy to escape. This method offers one significant advantage: the axially confining mirror coil fields remain static throughout the experiment. As a result, the magnetic bias between the top and bottom mirror coils remains fixed, and is much easier to characterize using ECR and other techniques. The would significantly reduce the uncertainty arising from knowledge of the magnetic bias.

A possible method to heat the antihydrogen atoms out of the trap is stochastic heating. By applying an oscillatory time-varying magnetic field using the internal mirror coils, the antihydrogen atoms experience random magnetic kicks, each imparting a small amount of momentum. The combined effect of many such kicks is to raise the antihydrogen kinetic energy; after sufficient heating time, the antiatoms will escape the confining well. The kinetic energy imparted by a single interaction must be significantly lower than the trap depth, and so such a measurement would take hundreds of seconds, if not hours. Without the improvements in background rejection described in this work, such a measurement would not be feasible.

Laser heating

Another possible method of inducing antihydrogen escape is laser heating. By tuning the cooling laser frequency slightly above the $1S-2P$ resonance, atoms moving away from the laser source are preferentially excited, transferring momentum and increasing the kinetic energy by around 1.3 mK per photon [160]. If laser excitations occur too rapidly, the resulting momentum kick can introduce a directional bias and potentially interfere with the gravity measurement. To mitigate this, excitation rate must be slow compared to the axial bounce frequency; there are also practical laser power constraints which limit this excitation rate.

Consequently, the laser heating must proceed gradually over multiple hours, similar to the laser cooling process demonstrated in ALPHA-2. This means that the BSC TOF-based background rejection would be crucial to any laser heating measurement. For example, the 0.01 Hz background rate illustrated in Section 8.5 would yield around 100 background counts over a 3 hour measurement. This would be acceptable given the current antihydrogen accumulation rate of ~ 1500 antiatoms over a 4 hour accumulation period.

TOF-based techniques

The TPC has inherently poor time resolution, as it relies on electron drift times on the order of microseconds. By contrast, building upon the calibration work presented here, it is possible to use the BSC to obtain annihilation times accurate to within a nanosecond.

This precise event time could be used directly in a determination of $\overline{g}$. A TOF-based method is proposed by the GBAR experiment [161], where $\bar{H}^+$ ions are created and held in a Penning-Malmberg trap, and then ionized using a pulsed laser. The time between the ionization and annihilation detection is used in combination with the annihilation position to determine the degree to which the antihydrogen atom was accelerated or decelerated by gravity.

A similar TOF-based measurement method is possible in ALPHA-g using existing apparatus and techniques. In ALPHA-2, the 1S-2P resonance is driven to laser cool trapped antihydrogen atoms, as described in Section 3.5.4. Antiatoms can be driven to the 2P_c hyperfine state, the relaxation of which yields antiatoms in both high-field-seeking and low-field-seeking states [116]. Those antiatoms in the high-field-seeking state are immediately ejected as the direction of the confining magnetic force is reversed. The 1S-2P laser is pulsed, with a pulse length of 10 ns.

A possible TOF-based gravity measurement in ALPHA-g could begin with antihydrogen collection in a small trapping region. The pulsed 1S-2P laser is used to drive transitions to untrappable states, ejecting the antiatoms. The time between the laser pulse and the annihilation is then measured using the BSC.

In an axially symmetric trapping potential, the TOF of the antiatoms would be equal in the up and down directions. However, under the acceleration of gravity, the TOF of antiatoms escaping downwards would be shortened, and the TOF of antiatoms escaping upwards would be increased.

For example, a laser-cooled antiatom with 50 mK of axial energy would have an axial speed of roughly 30 m/s, taking 0.33 ms to travel 1 cm. Gravity alters this velocity by around 0.01% over 1 cm, leading to an annihilation time difference of around 100 ns between antihydrogen annihilating 1 cm above and 1 cm below the trapping region. This measurement is feasible using the BSC detector for time measurement, and with the improved vertex position resolution presented in this dissertation.

Atomic fountains

Fundamentally, ALPHA-g is limited by the inhomogeneous magnetic field required to trap antihydrogen. The ultimate precision which is expected to be attainable using ALPHA-g is predicted to be around the 1% level. To surpass this, future antihydrogen gravity experiments will require measurements performed in a region free of large magnetic fields and field gradients.

The HAICU project is taking the first steps towards this next generation of experiments, by constructing an atomic fountain compatible with both hydrogen and antihydrogen. The apparatus uses a hydrogen source and a Zeeman decelerator which supplies a normal-conducting Ioffe-Pritchard quadrupole trap. In this trap, laser cooling will be applied in three dimensions [162].

To measure gravity, antihydrogen atoms will be ejected vertically from the trap, forming an atomic fountain. They will be interacted with while in free fall using lasers to perform Raman interferometry [163]. In this technique, each atom is split into a quantum superposition of two momentum states which follow spatially separated trajectories. These states are then made to reconverge to the same position. Since each path is accelerated differently by gravity, there will be a relative phase shift between the two states, causing interference. The interference pattern can be used to extract the gravitational acceleration. This technique has already been used to measure gravity in matter to a precision of 3×10^{-9}g [164]. If successfully applied to antihydrogen, it would improve upon the gravitational sensitivity of ALPHA-g by several orders of magnitude.

Chapter 10

Conclusions

This dissertation presents the first results from the ALPHA-g experiment, which constrain the acceleration of antihydrogen atoms due to the Earth's gravity. Published in [1], this is the result of work by the entire ALPHA collaboration. My personal contributions outlined here are centred on the commissioning of the antihydrogen detectors, and development of the detector data reconstruction algorithms. I was able to dramatically improve the annihilation position resolution, as well as implement time-of-flight measurements using a newly commissioned scintillator detector, improving the background rate by over an order of magnitude. These improvements mitigated uncertainties in the published result, and will directly enable new techniques in the future.

The measurement presented in this dissertation represents the culmination of a decades-long effort to investigate the effect of gravity on antimatter, a project first undertaken by Witteborn and Fairbank in the 1960s [26, 28]. It is the first direct confirmation of Einstein's weak equivalence principle for antimatter, supporting the long-predicted gravitational equivalence between matter and antimatter. This result constrains many speculative theoretical and cosmological models, and has already impacted the antimatter gravity community [165, 166].

In the ALPHA-g experiment, antiprotons are supplied by CERN's ELENA decelerator to a Penning-Malmberg catching trap, where they are prepared and subsequently transferred to the ALPHA-g apparatus. Simultaneously, positrons from a ^{22}Na source are accumulated, trapped, and also transferred to ALPHA-g. The two particle species are combined to form antihydrogen, some of which is trapped in an Ioffe-Pritchard magnetic trap.

To measure gravity, the vertical confining magnetic potentials are gradually lowered, allowing antiatoms to escape out of either the top or bottom of the trap. The subsequent annihilations are detected, and the ratio of counts in the two regions is plotted as a function of an applied magnetic bias. By comparing the results to antihydrogen trajectory simulations, the gravitational acceleration of antihydrogen was found [1] to be:

$$\overline{g}/g = (0.75 \pm 0.13 \text{ (statistical+systematic) } \pm 0.16 \text{ (simulation)}) \quad (10.0.1)$$

where $g \approx 9.81\,\mathrm{m/s^2}$ is the gravitational acceleration of matter at the Earth's surface.

This measurement depends upon detection of the annihilations, with precise determination of the annihilation positions. This was achieved using a time projection chamber (TPC) detector. The first of my individual contributions was a significant improvement to the track reconstruction and vertex finding algorithms, described in Chapter 5. The track finding algorithm was rewritten to be much more exhaustive, accurately finding different types of tracks which previously avoided reconstruction. The track fitting and vertex fitting functions were then rewritten to properly accommodate the more complete set of tracks. This lead to a vertex reconstruction algorithm which reflects the underlying charged particles much more accurately. This ultimately improved the vertex position outer resolution from approximately 15 cm to 3 cm, effectively eliminating annihilation position misidentification uncertainty as a significant uncertainty in the overall measurement.

Cosmic rays represent a substantial background in the TPC, and TPC data alone was insufficient for accurate background rejection. To address this, a barrel scintillator (BSC) detector was incorporated in ALPHA-g, capable of identifying cosmic rays based on their time-of-flight. The second of my individual contributions was the commissioning of the BSC detector, and the development of the low-level BSC event reconstruction described in Chapter 6. For the published gravity measurement, I used BSC hit information to reduce the cosmic ray background rate to an acceptable level.

This work also included an extensive time-of-flight calibration effort, described in Chapter 7, and the integration of BSC information into a machine-learning-based background rejection algorithm described in Chapter 8. The successful measurement of particle time-of-flight using the BSC has further improved the background rate, reducing the 58 Hz cosmic ray trigger rate to a false vertex rate of 10 mHz. The combination of the improvements to the vertex position resolution and to the cosmic ray background rate was estimated to have prevented an uncertainty on the measured gravitational acceleration of around $u[\overline{g}] = 0.04$g. Beyond these main contributions, I made a number of smaller contributions also described here, such as developing the ALPHA-g event display used to visualize detector data and understand events.

Looking forward, there are several short- and medium-term improvements to the ALPHA-g method which could enable a measurement of $\overline{g}$ with precision at the 1% level. With the improved detector performance presented here, such a measurement would feature a slow release of a large number of laser- and adiabatically-cooled antihydrogen atoms from the ALPHA-g precision trap. These experimental efforts will be complemented by trajectory simulations using a precise field model consistent with an extensive magnetic field measurement campaign.

Beyond the 1% measurement, this work paves the way for alternative measurement strategies in ALPHA-g. The improved detector background rate unlocks a new category of experiments with a duration of hours, such as slowly heating antihydrogen atoms out of a fixed magnetic trap. While these techniques may not significantly outperform the current method in precision, they offer valuable complementary measurements with a different set of systematic uncertainties.

Alongside the detector-focused contributions presented here, I was heavily involved in other aspects of the experiment while on shift at CERN, such as the optimization of antihydrogen production, cryogen management, design of auxiliary systems, and many other tasks essential to the functioning of the ALPHA-g experiment. I also participated in development, data-taking, and analysis for a number of other experiments in the ALPHA-2 apparatus. I have learned from experts across the wide range of disciplines involved in the ALPHA program, and have passed on my knowledge to the next generation of students.

It has been a genuine pleasure to work in the ALPHA collaboration and to directly contribute to our collective understanding of the relationship between antimatter and gravity. While the ALPHA-g results were not surprising from a theoretical perspective, they represent another key piece in our understanding of the universe, direct experimental confirmation of a fundamental symmetry of nature, and another stone overturned in our search for answers to some of the largest unanswered questions in physics.

Bibliography

[1] E.K. Anderson, C.J. Baker, W. Bertsche, et al. Observation of the effect of gravity on the motion of antimatter. *Nature*, 621:716–722, 2023.

[2] G. Smith. Background Rejection for the ALPHA-g Anti-Hydrogen Detectors. *PoS*, EXA-LEAP2024:084, 2025.

[3] G. Smith. Time-of-flight measurement with the ALPHA-g Barrel Scintillator detector for gravitational studies of anti-matter. *J. Instrum.*, 20(05):C05007, May 2025.

[4] A. Capra. Lifetime of magnetically trapped antihydrogen in ALPHA. *Hyperfine Interactions*, 240(1), January 2019.

[5] CERN. Garfield++: simulation of gaseous detectors. `https://garfieldpp.web.cern.ch/garfieldpp/`, 2025. Accessed: 2025-02-04.

[6] M. J. Borchert, J. A. Devlin, S. R. Erlewein, et al. A 16-parts-per-trillion measurement of the antiproton-to-proton charge–mass ratio. *Nature*, 601:53–57, 2022.

[7] C. Smorra, S. Sellner, M. Borchert, et al. A parts-per-billion measurement of the antiproton magnetic moment. *Nature*, 550:371–374, 2017.

[8] M. Ahmadi, B. Alves, C. Baker, et al. Characterization of the 1S–2S transition in antihydrogen. *Nature*, 557:71–75, 2018.

[9] M. Ahmadi, B. Alves, C. Baker, et al. Observation of the 1S–2P Lyman-α transition in antihydrogen. *Nature*, 561:211–215, 2018.

[10] The ALPHA Collaboration. Investigation of the fine structure of antihydrogen. *Nature*, 578:375–380, 2020.

[11] M. Ahmadi, B. Alves, C. Baker, et al. Observation of the hyperfine spectrum of antihydrogen. *Nature*, 548:66–69, 2017.

[12] A. Einstein. Die feldgleichungen der gravitation. *Kgl. Preuss. Akad. Wiss.*, pages 844–847, 1915.

[13] American Scientist. The secret history of gravitational waves, 2024. Accessed: 2024-12-18.

[14] The LIGO Scientific Collaboration and the Virgo Collaboration. Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.*, 116:061102, 2016.

[15] P. A. M. Dirac. The quantum theory of the electron. *Proc. R. Soc. Lond. A*, 117(778):610–624, 1928.

[16] P. A. M. Dirac. Quantised singularities in the electromagnetic field. *Proc. R. Soc. Lond. A*, 133:60–72, 1931.

[17] C. D. Anderson. The positive electron. *Phys. Rev.*, 43:491–494, Mar 1933.

[18] P. A. M. Dirac. Theory of electrons and positrons. In S. Stuart, editor, *Physics 1922–1941: Including Presentation Speeches and Laureates' Biographies*. Elsevier, Amsterdam, December 1965.

[19] O. Chamberlain, E. Segrè, C. Wiegand, and T. Ypsilantis. Observation of antiprotons. *Phys. Rev.*, 100:947–950, Nov 1955.

[20] B. Cork, G. R. Lambertson, O. Piccioni, and W. A. Wenzel. Antineutrons produced from antiprotons in charge-exchange collisions. *Phys. Rev.*, 104(4):1193–1197, 1956.

[21] The LEAR team. Performance of lear. CERN PS technical report CERN-PS-85-31 (LEA), CERN, May 1985.

[22] G. Baur, G. Boero, A. Brauksiepe, et al. Production of antihydrogen. *Phys. Lett. B*, 368(3):251–258, 1996.

[23] S. Maury. The antiproton decelerator: AD. *Hyperfine Interactions*, 109(1-4):43–52, 1997.

[24] M. Amoretti, C. Amsler, G. Bonomi, et al. The ATHENA antihydrogen apparatus. *Nucl. Instrum. Methods Phys. Res. A*, 518(3):679–711, 2004.

[25] G. B. Andresen, M. D. Ashkezari, M. Baquero-Ruiz, et al. Search for trapped antihydrogen. *Phys. Lett. B*, 695(1):95–104, 2011.

[26] Fred C. Witteborn and William M. Fairbank. Experiments to determine the force of gravity on single electrons and positrons. *Nature*, 220(5168):436–440, 1968.

[27] C.J. Baker, W. Bertsche, A. Capra, et al. Measurements of penning-malmberg trap patch potentials and associated performance degradation. *Phys. Rev. Res.*, 6(1):L012008, Jan 2024.

[28] F. C. Witteborn and W. M. Fairbank. Experimental comparison of the gravitational force on freely falling electrons and metallic electrons. *Phys. Rev. Lett.*, 19:1049–1052, Oct 1967.

[29] M. H. Holzscheiter, R. E. Brown, J. Camp, et al. Antimatter gravity and the weak equivalence principle. In J.C. Zorn and R.R. Lewis, editors, *Atomic Physics 12: Twelfth International Conference on Atomic Physics*, New York, 1991. American Institute of Physics.

[30] M. Doser, P. Pérez, and W. Bertsche. Does antimatter fall up? *CERN Courier*, 57(1):39–40, 2017.

[31] T. J. Phillips. The muonium antimatter gravity experiment. *EPJ Web Conf.*, 181:01017, 2018.

[32] G. Vinelli, F. Castelli, R. Ferragut, et al. A large-momentum-transfer matter-wave interferometer to measure the effect of gravity on positronium. *Classical and Quantum Gravity*, 40(20):205024, 2023.

[33] The ALPHA Collaboration and A. E. Charman. Description and first application of a new technique to measure the gravitational mass of antihydrogen. *Nature Communications*, 4:1785, 2013.

[34] C. Carli, D. Gamba, C. Malbrunot, L. Ponce, and S. Ulmer. ELENA: bright perspectives for low energy antiproton physics. *Nuclear Physics News*, 32(3):21–27, 2022.

[35] B. M. Latacz, B. P. Arndt, B. B. Bauer, et al. BASE—high-precision comparisons of the fundamental properties of protons and antiprotons. *Eur. Phys. J. D*, 77:94, 2023.

[36] C. Malbrunot, C. Amsler, S. Arguedas Cuendis, et al. The ASACUSA antihydrogen and hydrogen program: results and prospects. *Phil. Trans. R. Soc. A*, 376:20170273, 2018.

[37] M. Hori, A. Sótér, D. Barna, et al. Two-photon laser spectroscopy of antiprotonic helium and the antiproton-to-electron mass ratio. *Nature*, 475:484–488, 2011.

[38] T. Aumann, W. Bartmann, O. Boine-Frankenheim, et al. PUMA, antiproton unstable matter annihilation. *Eur. Phys. J. A*, 58:88, 2022.

[39] S. Aghion, O. Ahlén, C. Amsler, et al. A Moiré deflectometer for antimatter. *Nat. Commun.*, 5:4538, 2014.

[40] P. Adrich, P. Blumer, G. Caratsch, et al. Production of antihydrogen atoms by 6 keV antiprotons through a positronium cloud. *Eur. Phys. J. C*, 83(11), November 2023.

[41] E. Wigner. Einige folgerungen aus der schrödingerschen theorie für die termstrukturen. *Zeitschrift für Physik*, 43:624–652, 1927.

[42] E. Wigner. Ueber die operation der zeitumkehr in der quantenmechanik. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 1932:546–559, 1932.

[43] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys. Rev.*, 105:1413–1415, Feb 1957.

[44] L. Landau. On the conservation laws for weak interactions. *Nucl. Phys.*, 3(1):127–131, 1957.

[45] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2π decay of the K_2^0 meson. *Phys. Rev. Lett.*, 13:138–140, Jul 1964.

[46] G. Lüders. Proof of the TCP theorem. *Annals of Physics*, 2(1):1–15, 1957.

[47] G. Galilei. *Two New Sciences.* Elsevier, translated by Crew, H. and de Salvio, A., 1638.

[48] A. Einstein. On the relativity principle and the conclusions drawn from it. *Jahrbuch der Radioaktivität und Elektronik*, 4:411–462, 1907.

[49] L. Eötvös. Contributions to the law of proportionality of inertia and gravity. *Annalen der Physik*, 68(11):11–66, 1922.

[50] T. A. Wagner, S. Schlamminger, J. H. Gundlach, and E. G. Adelberger. Torsion-balance tests of the weak equivalence principle. *Classical and Quantum Gravity*, 29(18):184002, August 2012.

[51] S. Schlamminger, K.-Y. Choi, T. A. Wagner, J. H. Gundlach, and E. G. Adelberger. Test of the equivalence principle using a rotating torsion balance. *Phys. Rev. Lett.*, 100:041101, Jan 2008.

[52] P. Touboul, G. Métris, M. Rodrigues, et al. *microscope* mission: final results of the test of the equivalence principle. *Phys. Rev. Lett.*, 129:121102, Sep 2022.

[53] A. Schuster. Potential matter: a holiday dream. *Nature*, 58(1496):367–367, 1898.

[54] M. M. Nieto and T. Goldman. The arguments against "antigravity" and the gravitational acceleration of antimatter. *Phys. Rep.*, 205(5):221–281, 1991.

[55] A. D. Sakharov. Violation of CP invariance, C asymmetry and baryon asymmetry of the universe. *JETP Letters*, 5(1):32–35, 1967.

[56] M. Tanabashi, K. Hagiwara, K. Hikasa, et al. Review of particle physics. *Phys. Rev. D*, 98:030001, Aug 2018.

[57] The T2K Collaboration. Constraint on the matter–antimatter symmetry-violating phase in neutrino oscillations. *Nature*, 580:339–344, 2020.

[58] A. Einstein. Cosmological considerations in the general theory of relativity. *Kgl. Preuss. Akad. Wiss.*, pages 142–152, 1917.

[59] E. Hubble. A relation between distance and radial velocity among extra-galactic nebulae. *Proc. Natl. Acad. Sci. U.S.A.*, 15(3):168–173, 1929.

[60] A.G. Riess, A.V. Filippenko, P. Challis, et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *Astron. J.*, 116(3):1009–1038, 1998.

[61] R. Gsponer, S. Nadathur, N.T.B. Findlay, et al. Desi 2024 VI: Cosmological constraints from the measurements of baryon acoustic oscillations. *J. Cosmol. Astropart. Phys.*, 2025(02):021, 2025.

[62] S. Weinberg. The cosmological constant problem. *Rev. Mod. Phys.*, 61(1):1–23, 1989.

[63] D.S. Hajdukovic. Quantum vacuum and virtual gravitational dipoles: the solution to the dark energy problem? *Astrophys. Space Sci.*, 339:1–5, 2012.

[64] P. Morrison. Approximate nature of physical symmetries. *Am. J. Phys.*, 26(6):358–368, 1958.

[65] L. I. Schiff. Sign of the gravitational mass of a positron. *Phys. Rev. Lett.*, 1:254–255, Oct 1958.

[66] M. L. Good. ${K_2}^0$ and the equivalence principle. *Phys. Rev.*, 121:311–313, Jan 1961.

[67] G. Chardin and G. Manfredi. Gravity, antimatter and the Dirac-Milne universe. *Hyperfine Interactions*, 239(45), 2018.

[68] A. Benoit-Lévy and G. Chardin. Introducing the Dirac-Milne universe. *Astron. Astrophys*, 537:A78, 2012.

[69] D. S. Hajdukovic. Antimatter gravity and the universe. *Mod. Phys. Lett. A*, 35(08):2030001, November 2019.

[70] B. A. Bassett and R. Hlozek. Baryon acoustic oscillations. arXiv, 2009. arXiv:0910.5224.

[71] M. Fischler, J. Lykken, and T. Roberts. Direct observation limits on antimatter gravitation. Technical Report FERMILAB-FN-0822-CD-T, Fermi National Accelerator Lab. FNAL, Batavia, IL (United States), June 2008.

[72] S. Menary. What exactly is antimatter (gravitationally speaking)?, 2024.

[73] S. Ulmer, C. Smorra, A. Mooser, et al. High-precision comparison of the antiproton-to-proton charge-to-mass ratio. *Nature*, 524:196–199, 2015.

[74] H. Minakata and H. Nunokawa. Testing the principle of equivalence by solar neutrinos. *Phys. Rev. D*, 51:6625–6634, Jun 1995.

[75] L.M. Golino. *Machine learning methods for antihydrogen detection.* Ph.D. thesis, Swansea University, 2025.

[76] M. Urioni. *Measurement of Earth's gravitational acceleration on antihydrogen with the ALPHA experiment at CERN*. Ph.D. thesis, Università degli Studi di Brescia, 2024.

[77] P. D. Woosaree. *First measurement of antihydrogen free fall using a radial time projection chamber*. Ph.D. thesis, University of Calgary, 2025.

[78] A. M. W. Powell. *Magnetic field characterisation for gravitational free fall measurements of antihydrogen in the ALPHA-g experiment*. Ph.D. thesis, University of Calgary, 2024.

[79] M. Vretenar, J. Vollaire, R. Scrivens, et al., editors. *Linac4 design report*, volume 6/2020 of *CERN Yellow Reports: Monographs*. CERN, Geneva, 2020.

[80] K. H. Reich. The cern proton synchrotron booster. *IEEE Trans. Nucl. Sci., NS-16: 959- 61(June 1969).*, 01 1969.

[81] J. Speed. Quadrupole design for the 2 GeV upgrade of the CERN PS-booster. *IEEE Trans. Appl. Supercond.*, 28(3):1–4, 2018.

[82] J P Blewett. The proton synchrotron. *Reports on Progress in Physics*, 19(1):37, jan 1956.

[83] S. A. Baird, D. Berlin, J. Boillot, et al. Design study of the antiproton decelerator: AD. Technical Report CERN-PS-96-043-AR; CERN/PS 96-43 (AR), CERN, November 1996.

[84] B. Autin, S. A. Baird, D. Berlin, et al. The antiproton decelerator (ad), a simplified antiproton source (feasibility study). Technical Report CERN-PS-95-36-AR, CERN, November 29 1995.

[85] V. Chohan, C. Alanzeau, M. Angoletta, et al. *Extra Low ENergy Antiproton (ELENA) ring and its transfer lines: design report*. CERN, 04 2014.

[86] S. Earnshaw. On the nature of the molecular forces which regulate the constitution of the luminiferous ether. *Trans. Cambridge Philos. Soc.*, 7:97–112, 1842.

[87] K. Blaum, Y. N. Novikov, and G. Werth. Penning traps as a versatile tool for precise experiments in fundamental physics. *Contemp. Phys.*, 51(2):149–175, March 2010.

[88] H. G. Dehmelt. Hans g. dehmelt – biographical. https://www.nobelprize.org/prizes/physics/1989/dehmelt/biographical/, 1989. Accessed: 2025-04-13.

[89] TRIUMF. TITAN penning trap. https://titan.triumf.ca/equipment/penning_trap/, 2025. Accessed: 04-Feb-2025.

[90] M. Charlton, S. Jonsell, L. V. Jørgensen, N. Madsen, and D. P. van der Werf. Antihydrogen for precision tests in physics. *Contemp. Phys.*, 49(1):29–41, 2008.

[91] D. H. E. Dubin and T. M. O'Neil. Trapped nonneutral plasmas, liquids, and crystals (the thermal equilibrium states). *Rev. Mod. Phys.*, 71:87–172, Jan 1999.

[92] T. M. O'Neil. A confinement theorem for nonneutral plasmas. *Physics of Fluids*, 23(11):2216–2218, 1980.

[93] P. Meystre. *Atom Optics*, volume 33 of *Springer Series on Atomic, Optical, and Plasma Physics*. Springer, New York, 2001. See p. 77.

[94] S. Gov, S. Shtrikman, and H. Thomas. Magnetic trapping of neutral particles: Classical and quantum-mechanical study of a ioffe–pritchard type trap. *J. Appl. Phys.*, 87(8):3989–3998, April 2000.

[95] S. Pevovar, M. Weber, and K. Lynn. Ratio of positron annihilation into three photons versus two. *Phys. Status Solidi C*, 4:3447 – 3450, 09 2007.

[96] B. Jasińska et al. Determination of the 3γ fraction from positron annihilation in mesoporous materials for symmetry violation experiment with J-PET scanner. *Acta Phys. Polon. B*, 47:453, 2016.

[97] C. Ghesquière. An inclusive view on $\bar{p}p \to \pi$ at rest. In *Symposium on antinucleon-nucleon interactions*, pages 436–446. CERN, 1974.

[98] S.J. Orfanidis and V. Rittenberg. Nucleon-antinucleon annihilation into pions. *Nucl. Phys. B*, 59(2):570–582, 1973.

[99] C. Amsler, H. Breuker, M. Bumbar, and et al. Antiproton annihilation at rest in thin solid targets and comparison with monte carlo simulations. *Eur. Phys. J. A*, 60:225, 2024. Published: 17 November 2024.

[100] M. Cahay, J. Cugnon, and J. Vandermeulen. Low-energy antiproton annihilation in nuclei. *Nucl. Phys. A*, 393(3):237–251, 1983.

[101] M. A. Johnson. *Design and commissioning of beamlines for the ALPHA antihydrogen experiment*. PhD thesis, The University of Manchester, Manchester, United Kingdom, 2019. Ph.D. thesis.

[102] R. Akbari, B. Alves, C. Baker, et al. The alpha-2 apparatus - facilitating experimentation with trapped antihydrogen. *Nucl. Instrum. Methods Phys. Res. A.*, 1072:170194, 2025.

[103] S. Fabbri. *Optimization of antiproton capture and delivery for the ALPHA antihydrogen experiment*. Ph.D. thesis, The University of Manchester, December 2021.

[104] P. Granum. *Measuring the properties of antihydrogen*. Ph.D. thesis, Aarhus University, Denmark, June 2022.

[105] R. R. Kinsey et al. The nudat/pcnudat program for nuclear data. In *Proceedings of the 9th International Symposium on Capture Gamma-Ray Spectroscopy and Related Topics*, Budapest, Hungary, October 1996. Data extracted from the NuDat database, version 3.0 (accessed 2025-06-17), National Nuclear Data Center, `https://www.nndc.bnl.gov/nudat/`.

[106] C. Amole, G. B. Andresen, M. D. Ashkezari, et al. The ALPHA antihydrogen trapping apparatus. *Nucl. Instrum. Methods Phys. Res. A.*, 735:319–340, 2014.

[107] T. J. Murphy and C. M. Surko. Positron trapping in an electrostatic well by inelastic collisions with nitrogen molecules. *Phys. Rev. A*, 46:5696–5705, Nov 1992.

[108] J. Larmor. LXIII. On the theory of the magnetic influence on spectra; and on the radiation from moving ions. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 44(271):503–512, 1897.

[109] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell. Observation of Bose-Einstein condensation in a dilute atomic vapor. *Science*, 269(5221):198–201, 1995.

[110] G. B. Andresen, M. D. Ashkezari, M. Baquero-Ruiz, et al. Evaporative cooling of antiprotons to cryogenic temperatures. *Phys. Rev. Lett.*, 105:013003, Jul 2010.

[111] François A. Rotating wall technique and centrifugal separation. In M. Knoop et al., editors, *Physics with Trapped Charged Particles*, pages 221–238. World Scientific, 2016.

[112] J. R. Danielson and C. M. Surko. Radial compression and torque-balanced steady states of single-component plasmas in Penning-Malmberg traps. *Physics of Plasmas*, 13(5):055706, 05 2006.

[113] M. Ahmadi, B. Alves, C. Baker, et al. Enhanced control and reproducibility of non-neutral plasmas. *Phys. Rev. Lett.*, 120:025001, Jan 2018.

[114] R. Akbari, B. Alves, C. Baker, et al. The ALPHA-2 apparatus - facilitating experimentation with trapped antihydrogen. *Nucl. Instrum. Methods Phys. Res. A.*, 1072:170194, 2025.

[115] NobelPrize.org. Additional background material on the nobel prize in physics 1997. `https://www.nobelprize.org/prizes/physics/1997/advanced-information/`, 2025. Nobel Prize Outreach, accessed 4 February 2025.

[116] C. J. Baker, W. Bertsche, A. Capra, and et al. Laser cooling of antihydrogen atoms. *Nature*, 592:35–42, 2021.

[117] M. Kasevich and S. Chu. Laser cooling below a photon recoil with three-level atoms. *Phys. Rev. Lett.*, 69:1741–1744, Sep 1992.

[118] L. D. Landau and E. M. Lifshitz. *Mechanics*, volume 1 of *Course of Theoretical Physics*. Butterworth-Heinemann, Oxford, 3rd edition, 1996.

[119] R. M. Kulsrud. Adiabatic invariant of the harmonic oscillator. *Phys. Rev.*, 106:205–207, Apr 1957.

[120] D. L. Hodgkinson. *On the dynamics of adiabatically cooled antihydrogen in an octupole-based ioffe-pritchard magnetic trap.* Ph.D. thesis, The University of Manchester, 2022.

[121] M. Ahmadi, B. Alves, C. Baker, et al. Adiabatic expansion cooling of antihydrogen. *Phys. Rev. Res.*, 6:L032065, Sep 2024.

[122] M. Amoretti, C. Amsler, G. Bonomi, et al. Production and detection of cold antihydrogen atoms. *Nature*, 419:456–459, 2002.

[123] C. Amole, G. Andresen, M. Ashkezari, et al. Silicon vertex detector upgrade in the alpha experiment. *Nucl. Instrum. Methods Phys. Res. A.*, 732:134–136, 12 2013.

[124] J. H. Hubbell and S. M. Seltzer. Tables of x-ray mass attenuation coefficients and mass energy-absorption coefficients (version 1.4). `http://physics.nist.gov/xaamdi`, 2004. National Institute of Standards and Technology, Gaithersburg, MD. [Online] Available: http://physics.nist.gov/xaamdi [Accessed: 2025-07-19].

[125] G. F. Knoll. *Radiation Detection and Measurement*. Wiley, New York, NY, 4 edition, 2010.

[126] H. J. Hilke. Time projection chambers. *CERN Document Server*, 2010.

[127] G. Charpak. Development of multiwire proportional chambers. *CERN Courier*, 9(6):174–176, 1969.

[128] J. Alme, Y. Andres, H. Appelshäuser, et al. The ALICE TPC, a large 3-dimensional tracking device with fast readout for ultra-high multiplicity events. *Nucl. Instrum. Methods Phys. Res. A.*, 622(1):316–367, October 2010.

[129] K. Majumdar and K. Mavrokoridis. Review of liquid argon detector technologies in the neutrino sector. *Appl. Sci.*, 11(6):2455, March 2021.

[130] A. Capra. ALPHA: precision measurements of antihydrogen. In *Proceedings of the 56th Winter Nuclear and Particle Physics Conference*, Banff, Canada, 2019.

[131] CERN. Cosmic rays: particles from outer space. `https://home.cern/science/physics/cosmic-rays-particles-outer-space`, n.d. Accessed: 2025-02-04.

[132] Eljen Technology. Ej-200, ej-204, ej-208, ej-212 plastic scintillators. `https://eljentechnology.com/products/plastic-scintillators/ej-200-ej-204-ej-208-ej-212`, n.d. Accessed: 2025-02-04.

[133] ON Semiconductor. Microj series datasheet. `https://www.onsemi.com/download/data-sheet/pdf/microj-series-d.pdf`, n.d. Accessed: 2025-02-04.

[134] ON Semiconductor. Customer letter. `https://www.onsemi.jp/site/pdf/Customer%20Letter%20FINAL.pdf`, n.d. Accessed: 2025-02-04.

[135] S. Ritt and P. A. Amaudruz. New components of the MIDAS data acquisition system. In *1999 IEEE Conference on Real-Time Computer Applications in Nuclear Particle and Plasma Physics. 11th IEEE NPSS Real Time Conference. Conference Record (Cat. No.99EX295)*, pages 116–118, 1999.

[136] A. Capra and ALPHA Collaboration. Antihydrogen annihilation detection: the alpha-g radial tpc. In *Proceedings of Science (EXA-LEAP2024)*, volume 480, page 063, 2025. Presented at EXA-LEAP2024, Vienna, Austria, August 26–30, 2024.

[137] A. B. Garnsworthy, C. J. Pearson, D. Bishop, et al. The GRIFFIN data acquisition system. *Nucl. Instrum. Methods Phys. Res. A.*, 853:85–104, 2017.

[138] CERN. Cern report p596. Technical report, CERN, n.d. Accessed: 2025-02-04.

[139] G. Korcyl, L. Maier, J. Michel, et al. A users guide to the trb3 and fpga-tdc based platforms. `http://jspc29.x-matter.uni-frankfurt.de/docu/trb3docu.pdf`, 02 2021. Online; accessed 4-May-2021.

[140] A. J. Peters, E. A. Sindrilaru, and G. Adde. EOS as the present and future solution for data storage at CERN. In *J. Phys. Conf. Ser.*, volume 664, page 042042. IOP Publishing, 2015. Published under licence by IOP Publishing Ltd.

[141] ALPHASOFT Developers. Alphasoft: Software for the alpha experiment. `https://bitbucket.org/expalpha/alphasoft/`, n.d. Accessed: 2025-02-04.

[142] R. Brun and F. Rademakers. ROOT — An object oriented data analysis framework. *Nucl. Instrum. Methods Phys. Res. A.*, 389:81–86, 1997.

[143] P. Harvey. XSNOED: X-Windows SNO event display. `https://www.sno.phy.queensu.ca/~phil/xsnoed/xsnoed.html`. Accessed: 2025-02-04.

[144] J. H. Friedman, J. L. Bentley, and R. A. Finkel. An algorithm for finding best matches in logarithmic expected time. *ACM Trans. Math. Softw.*, 3(3):209–226, September 1977.

[145] A. Capra, P.-A. Amaudruz, D. Bishop, et al. Design of a radial TPC for antihydrogen gravity measurement with ALPHA-g. In *Proceedings of the 12th International Conference on Low Energy Antiproton Physics (LEAP2016)*, volume 18 of *JPS Conf. Proc.*, page 011015, September 2017.

[146] P. Avery. Applied fitting theory IV: formulas for track fitting. Technical Report CBX 92-45, Cornell University, 1992.

[147] A. Ferreira, M. Singh, G. Smith, et al. AI meets antimatter: unveiling antihydrogen annihilations. In *38th conference on Neural Information Processing Systems*, 12 2024.

[148] R. L. Workman and others (Particle Data Group). Review of particle physics. *Progress of Theoretical and Experimental Physics*, 2022(8):083C01, 2022.

[149] A. Capra. Time application of the barrel detector for alpha-g. Technical report, ALPHA Collaboration, 2017.

[150] M. Starič and the Belle II TOP group. Alignment and calibration methods for the Belle II TOP counter. *Nucl. Instrum. Methods Phys. Res. A.*, 876:260–264, 2017.

[151] T. Davidek and the ATLAS Collaboration. Atlas tile calorimeter time calibration, monitoring and performance. *J. Phys. Conf. Ser.*, 928(1):012003, 2017.

[152] Y. Shikaze, S. Orito, T. Mitsui, et al. Large-area scintillator hodoscope with 50 ps timing resolution onboard bess. *Nucl. Instrum. Methods Phys. Res. A.*, 455(3):596–606, December 2000.

[153] E. Pooser, F. Barbosa, W. Boeglin, et al. The gluex start counter detector. *Nucl. Instrum. Methods Phys. Res. A.*, 927:330–342, May 2019.

[154] W. Braunschweig, E. Königs, W. Sturm, and W. Wallraff. A large area time-of-flight system with a resolution of 0.5 ns fwhm. *Nucl. Instrum. Methods*, 134:261–266, 1976.

[155] C. Wu, Y. Heng, Y. Zhao, et al. The timing properties of a plastic time-of-flight scintillator from a beam test. *Nucl. Instrum. Methods Phys. Res. A.*, 555(1-2):142–147, 2005.

[156] F. Carnesecchi. Performance of the alice time-of-flight detector at the lhc. *J. Instrum.*, 14(06):C06023–C06023, June 2019.

[157] A. Capra and ALPHA Collaboration. Machine learning for antihydrogen detection at alpha. *J. Phys. Conf. Ser.*, 1085(4):042007, 2018.

[158] G. Aad, B. Abbott, J. Abdallah, et al. Measurement of event shapes at large momentum transfer with the ATLAS detector in pp collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J. C*, 72:2211, 2012.

[159] D. Hodgkinson, J. Fajans, and J. S. Wurtele. A bayesian data analysis method for an experiment to measure the gravitational acceleration of antihydrogen. *Phys. Sci. Forum*, 9(1), 2023.

[160] A. Evans. *Laser cooling of antihydrogen.* Ph.D. thesis, University of Calgary, Calgary, Canada, 2022.

[161] G. Chardin, P. Grandemange, D. Lunney, et al. Proposal to measure the gravitational behaviour of antihydrogen at rest. Technical report, CERN, Geneva, 2011.

[162] A. Capra. The HAICU experiment at TRIUMF. Presented at TRIUMF Science Week 2024, July 2024. TRIUMF Auditorium, Indico contribution #5631.

[163] B. Barrett, A. Carew, H. C. Beica, et al. Prospects for precise measurements with echo atom interferometry. *Atoms*, 4(3), 2016.

[164] A. Peters, K Y. Chung, and S. Chu. Measurement of gravitational acceleration by dropping atoms. *Nature*, 400(6747):849–852, 1999.

[165] M. Villata. Antimatter gravity and the results of the ALPHA-g experiment. *Annalen der Physik*, 536(12):2300519, 2024.

[166] D.S. Hajdukovic. Antimatter gravity experiments, the astronomical challenges to ΛCDM cosmology and the quantum vacuum as a possible source of gravity in the universe. *Int. J. Theor. Phys.*, 64:93, 2025.

Appendix A

BSC electronics and DAQ implementation

The hardware of the BSC detector is described in Section 4.3, and the central components to the BSC data acquisition (DAQ) are outlined in Section 4.4.2. However, a more detailed discussion of the BSC electronics and DAQ may be useful here. This section will cover components not covered in the main text, and outline some of the practical problems faced during this work. It is intended as a reference for future work on the BSC, or for those who are highly interested in the readout electronics. A diagram of the BSC electronics and DAQ system is shown in Figure A.1.

Power distribution board

The ASD boards require low voltage power, requiring inputs of 5 V, 3 V, and −5 V. These are supplied by a low voltage power supply via a customized power distribution board. This board is also nominally responsible for the distribution of the 27 V SiPM bias voltage from an external high voltage power supply.

Due to a design or fabrication error, the power distribution boards are not able to distribute this high voltage; instead, one wire is removed from each connector and routed into two common connectors which are connected to the high voltage power supply via a custom implementation. The seating of these wires in the common connectors is not secure, and commonly leads to ASD boards without bias voltage. This power distribution board will soon be redesigned.

Rear transmission module

The signals from the 128 ASD boards are carried off the detector by ribbon cables. To distribute these signals to the ADC and TDC, a custom *rear transmission module* (RTM) was designed. Eight of these modules are used, with each module receiving signals from four ASD boards, corresponding to

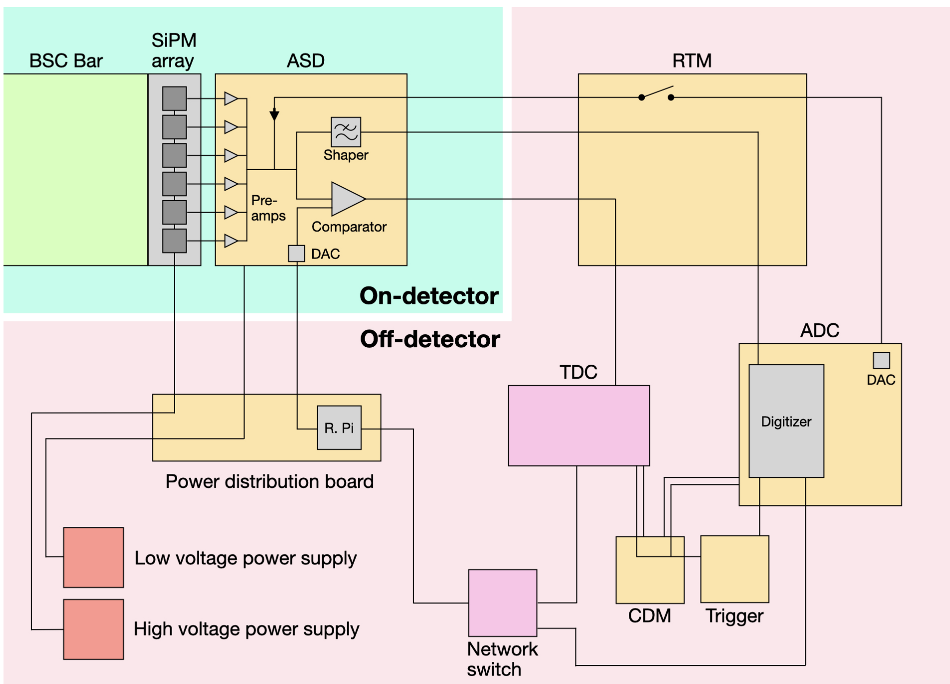

Figure A.1: A diagram of the detector electronics and DAQ system for the BSC. Customized modules and circuit boards developed at TRIUMF are shown in yellow, commercially available units in purple, and power supplies in red.

the top and bottom of eight bars. Each RTM passes the analog SiPM signals to the ADC sitting opposite it in a Versa Module Eurocard (VME) crate. It also passes the digital discriminated signals to the TDC via a twisted pair ribbon cable.

One further feature of the RTM is to pass an electronic test pulse to each ASD board, which is injected directly after the first preamplification stage of the SiPM signal. The test pulse is supplied to each RTM via lemo cable from a digital-to-analog converter (DAC) on an ADC module. The RTM also includes a set of switches to enable or disable this test pulse for each ASD board.

There were numerous problems observed with this test pulse circuit. Despite each RTM being given identical input pulses, there was significant variation in the test pulses observed in the ASD boards. Furthermore, leaving the switches open during data-taking couples a large electronic noise into the system. For these reasons, a test-pulse-based TOF calibration cannot be used with the current RTMs[12].

SiPM gain

One issue is that the gains of the SiPM arrays show variation between each bar end, with some bar ends yielding consistently larger SiPM signals. In particular, the signals are significantly larger on the bottom end of each bar. This SiPM signal amplitude discrepancy is higher than can be explained through factors such as temperature variation.

One working theory was that the physical seating of the BSC bars was not robust. The force of gravity could have pushed the bar into the SiPMs at the bottom, increasing the coupling of photons into the bottom SiPMs. To address this, I performed a compression test using two prototype bars at TRIUMF, with a replica of the DAQ system. The bars were clamped into place and a metal plate pushed into one end of the bars by tightening a pair of screws. Increasing the compression pushing the bars and SiPMs together had no affect on the signal sizes observed by the ADC.

With this theory ruled out, the most likely cause now seems to be a systematic drop in the bias voltage applied to the bottom SiPMs. The redesign of the power distribution board could lead to individual bias voltages being supplied to each SiPM array, which could help to ameliorate this issue.

[12] There are other reasons to want a redesign of the RTMs: poorly soldered ground pins on the ribbon cable connectors often fall out, and the cable retention of the ribbon cables is poor, frequently leading to ADC or TDC channels with no signal.

Bias voltage selection

With the current external bias voltage supply, one common bias voltage must be chosen for all SiPMs in the BSC. The maximum bias voltage recommended by [133] is 30.7 V, corresponding to 6 V over the breakdown voltage. Initially, a bias voltage of 30 V was used: the larger the bias voltage, the larger the SiPM signals, and so the more accurately the signal start time can be found. When a bias voltage of 27 V was tested, the performance of the bar hit cluster background rejection was found to be similar, and so the lower 27 V bias voltage was used instead for the initial gravity measurement.

Once the TOF calibration was completed, the bias voltage was revisited. Cosmic ray data was taken with 27 V, 29 V, and 30 V bias voltage. The TOF calibration algorithm was run using each dataset, and the data analyzed. The results indicated minimal change to the TOF resolution between 27 V and 29 V, but that increasing the bias voltage to 30 V led to larger ADC waveforms, more bar hits, but worse TOF resolution. For this reason, the bias voltage was not changed from 27 V.

This bias voltage selection could be revisited in future work. After adjusting the bias voltage, many further parameters may have to be reoptimized, such as the TDC thresholds. It is also possible that the TOF calibration algorithm should be modified for better performance using data taken with higher bias voltages.

TDC threshold selection

The digital signals sent to the TDC are produced by a comparator which compares the SiPM signal to a reference threshold voltage. This threshold voltage is produced by a digital-to-analog converter (DAC) on the ASD board. The voltage is determined by a Raspberry Pi mounted on the power distribution board, which can specify a unique threshold voltage for each bar end.

To choose the value of the TDC threshold, a threshold scan is regularly performed. The thresholds are decreased from 30 mV to 0 mV in steps, and 5 minutes of cosmic ray data is taken at each step. The TDC hit rate is plotted against the threshold voltage for each channel. The hit rate increases steadily as the threshold is lowered, before a dramatic increase once the threshold reaches the electronic noise level. The threshold for each channel is then set to the voltage immediately before this sharp increase in hit rate.

Malfunctioning TDC channels

The digital signal output from the comparator is carried into one of 8 RTMs, and then into the TDC by twisted pair ribbon cable. The TDC uses four FPGAs to perform the time digitization, with 64 channels per FPGA. Each FPGA has its own circuit board with four connector ports. The ribbon cables from the eight RTMs are fed into eight of these connectors.

After the TOF calibrations were completed, it quickly became obvious that some of the TDC channels had problems. Despite a successful fine time calibration, 11 of the 128 TDC channels used had a curious 6-peaked time structure when probed with the test pulser, and similar timing issues in cosmic ray data. The timing variation on these channels was around 4 ns, which made them unusable for TOF measurements.

To circumvent this issue, the signals from these 11 bar ends were rerouted into different TDC channels. This was done by splitting the corresponding wires from the twisted pair ribbon cable carrying the signal into the TDC, and recombining them into a 9th connector, which was plugged into an empty port on the TDC. This successfully recovered the time measurements on these 11 bar ends, allowing them to be used for TOF measurements.

Appendix B

BSC quality of data

To check the BSC detector quality of data during normal operation, a number of plots are automatically generated during each run. These plots can be monitored by detector experts, to ensure proper functioning of the BSC. These plots have already proven useful for diagnosing both hardware and software issues.

The set of plots is shown here in Figure B.1; the meaning and use of each plot is described in Table B.1. This information is provided primarily as a reference to those who will be working with the BSC in the future.

Table B.1: Descriptions of the quality of data plots shown in Figure B.1.

Row	**Col**	**Description**
1	1	Occupancy plots for the 128 bar ends. Shows ADC hits, TDC hits, and matched ADC+TDC end hits. Used to find bars with no signal, or too many signals, or failed ADC+TDC matching.
1	2	Occupancy plots for the 64 bar ends. Shows all bar hit types, complete bar hits, and bar hits matched to the TPC. Used to validate bar end matching.
1	3	Number of end hits per event, including ADC hits, TDC hits, and ADC+TDC matched.
1	4	Number of bar hits and bar hit clusters per event. Used to validate hit clustering.
2	1	ADC hit amplitude, including raw amplitude and amplitude from waveform fits. Used to validate waveform fitting.
2	2	ADC hit amplitude per channel. Used to monitor variation in gain for each channel.
2	3	Fine time per channel. Used to ensure fine time is calculated correctly for each channel: if some channels do not have a uniform distribution from 0 ns to 5 ns, the fine time calibration of Section 6.2.2 should be repeated.
2	4	Time difference between matched ADC and TDC hits.

Row	**Col**	**Description**
3	1	Bar hit z position from TDC time difference between bar ends.
3	2	Difference between bar hit z position and matched TPC track z position. Used to validate TPC matching and time calibrations.
3	3	TOF between bars plotted against expected cosmic ray TOF.
3	4	Residual TOF (TOF − expected TOF) for each bar. Used to validate TOF calibrations per bar.

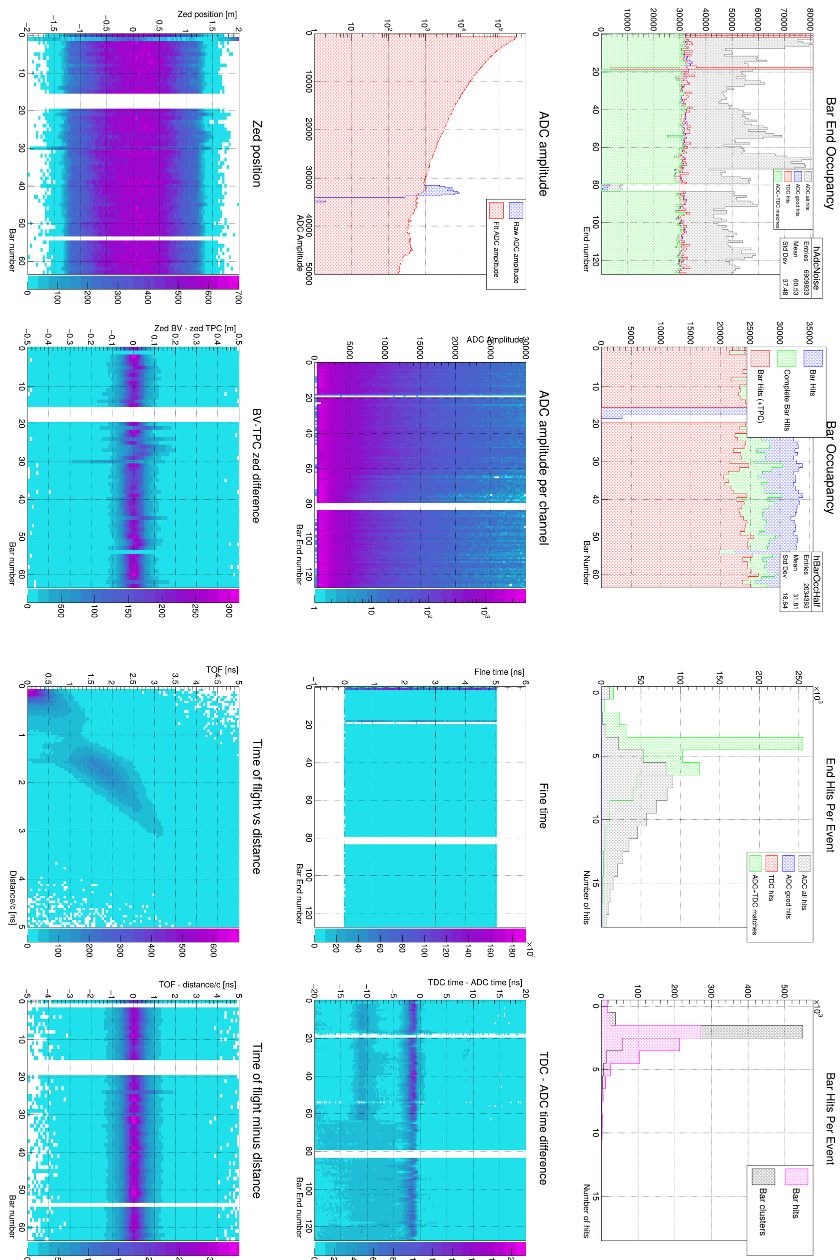

Figure B.1: The quality of data plots automatically generated during regular operation of the BSC. Data is from cosmic rays, with BSC calibration.

Appendix C

Pileup rejection

The maximum drift time of ionization electrons in the TPC is around 4 μs; if a second charged particle crosses the TPC during this time, the two events cannot be separated. This can be problematic for vertex reconstruction; since the time projection step uses the first anode wire signal as the zero time, electron drift times will be miscalculated for the second track.

The 325 ps time resolution of the BSC provides a convenient method to address this. If the effective TOF between a pair of bar hits exceeds 200 ns, the later bar hit is flagged as being caused by pileup. Any TPC track matched to this hit can be excluded from the vertex reconstruction. This modification to the vertex reconstruction was tested, but is current disabled as it did not yield significant improvements to any metric.

Currently, the main use of this pileup tagging is to exclude events containing pileup from the machine learning training set (see Chapter 8). This is crucial as pileup is prevalent in the high rate antihydrogen mixing data which is used for training, but rare in the low rate annihilations typical of gravity experiments.

The theoretical pileup rate can be estimated by modelling detector triggers as a Poisson process. This makes the approximation that there is only one underlying process; if both cosmic rays and annihilations are present in the data, the cosmic rays are neglected. For a Poisson process, the time between successive hits follows an exponential distribution. The theoretical pileup rate can then be obtained integrating this probability density function $P(t) = \lambda e^{-\lambda t}$ over the maximum drift time:

$$P(t < 4\mu s) = \int_0^{4\mu s} \lambda e^{-\lambda t} dt \tag{C.0.1}$$

This relies upon knowledge of the average trigger rate λ. This value is based on the measured detector trigger rate (using the trigger-level data sent from the ADC to the trigger module, as described in Section 4.4.3). This reported trigger rate is not limited by the readout rate of the TPC.

These theoretical pileup rates are compared to the pileup rates measured using the BSC for different ALPHA-g data-taking modes in Table C.1. The

Event Type	**Trigger Rate**	**Theoretical Pileup**	**Measured Pileup**
Cosmic ray μ	78 Hz	0.03%	$0.338 \pm 0.008\%$
$\overline{H}$ release	200 Hz	0.08%	$4.3 \pm 0.1\%$
$\overline{H}$ mixing	44.3 kHz	16.2%	$20.9 \pm 0.5\%$

Table C.1: Comparison of theoretical and measured pileup probabilities for different ALPHA-g data-taking modes.

measured and theoretical rates are roughly comparable, but inconsistent.

To investigate this, the event display was used to examine pileup events. Many of these events included tracks of secondary particles in the TPC, hinting towards other processes which could mimic pileup. For example, the decay of a captured muon creates a secondary electron approximately 2.2 µs after the initial muon track. The time between successive bar hits in the same readout window exhibited a distribution with a similar time constant, consistent with the hypothesis that some of these secondary particles are caused by muon capture.

Events including delayed-secondary-producing processes such as muon capture are also tagged as pileup by the BSC. Since these processes affect only a small fraction of events, and the magnitude of their effect is small, further investigation and proper treatment of them is left for future work.

Appendix D

BSC calibration algorithm implementation

The calibration procedure described in Chapter 7 is implemented into a standalone ROOT macro, which takes as input one or more ROOT files produced by the *alphasoft* analysis of MIDAS data. This macro typically takes 20 minutes to process 3 hours of cosmic ray data. The output of the macro is a new *calibration ROOT file*, which contains:

- Individual time offsets $\overline{\delta}_i^{\text{top}}$ and $\overline{\delta}_i^{\text{bot}}$ per bar end.
- Time walk correction parameters B_i^{top} and B_i^{bot} per bar end.
- Effective speed of photon propagation v_{eff} per bar.
- ADC thresholds per bar end (see Section 6.1).
- z_0 and λ per bar end (see Section 6.5).
- Diagnostic plots to verify the success and quality of each stage of calibration.
- The overall cosmic ray residual TOF distribution with these corrections applied, with a double Gaussian fit (see Figure 6.10). This indicates the TOF resolution achieved via the calibration.

The *alphasoft* event reconstruction will search for the most recent calibration ROOT file, with a run number at most 100 runs before the run number being analyzed, to extract the calibration constants. It is also possible to run the *alphasoft* reconstruction with a manually specified calibration file.

The online mode of *alphasoft* (described in Section 5.1) is used to provide real-time diagnostics, and so it is essential that an up-to-date calibration file is available. After each run, the BSC calibration macro is run automatically using data from that run. If it does not meet the predefined threshold of 3 hours of cosmic ray data, it will include data from additional previous runs to reach this threshold. This automatic calibration procedure was used successfully over the 2024 ALPHA-g data-taking period.

Two countermeasures are taken to prevent contamination of the calibration cosmic ray dataset by annihilations; these allow the calibration to be performed even using runs where antiprotons or antihydrogen were present in ALPHA-g. Firstly, only events with three or fewer clusters of bar hits are used. This removes around 70% of annihilation events, while removing less than 1% of cosmic ray events. The second method is to use our knowledge of the antihydrogen production cycle to automatically exclude time windows where antiprotons are present in ALPHA-g.

As well as being executed automatically after every run for live online calibrations, the calibration macro described in this section is used for offline, publication-quality data. An offline calibration pipeline was developed which is compatible with CERN's *condor* batch computing grid. It was designed to be entirely decoupled from any ROOT files produced by the online *alphasoft* reconstruction, allowing full reproducibility from the MIDAS data files alone. This pipeline performs the BSC calibration and runs the *alphasoft* reconstruction with the resulting calibration file, following these steps for each run:

1. Perform a lightweight analysis of the previous 20 runs to determine the total time without antiprotons in ALPHA-g.
2. Starting with the given run, add runs in reverse chronological order until a pre-defined total time without antiprotons is reached.
3. Analyze these runs with *alphasoft* to produce the ROOT output files.
4. Execute the BSC calibration macro on the produced ROOT output, creating a calibration ROOT file.
5. Check the calibration file to ensure successful calibration, with a TOF resolution better than 500 ps.
6. Run the *alphasoft* analysis again, using the new calibration file.

The most likely cause of calibration failure is if the calibration dataset contains runs from before and after a TDC power cycle (see Section 7.1), i.e. where the channel offsets shift partway through the calibration dataset. In this case, the pipeline looks for a run shortly after which was also analyzed, and for which the calibration was successful. This calibration file is used if one is available; otherwise, the analysis is aborted.

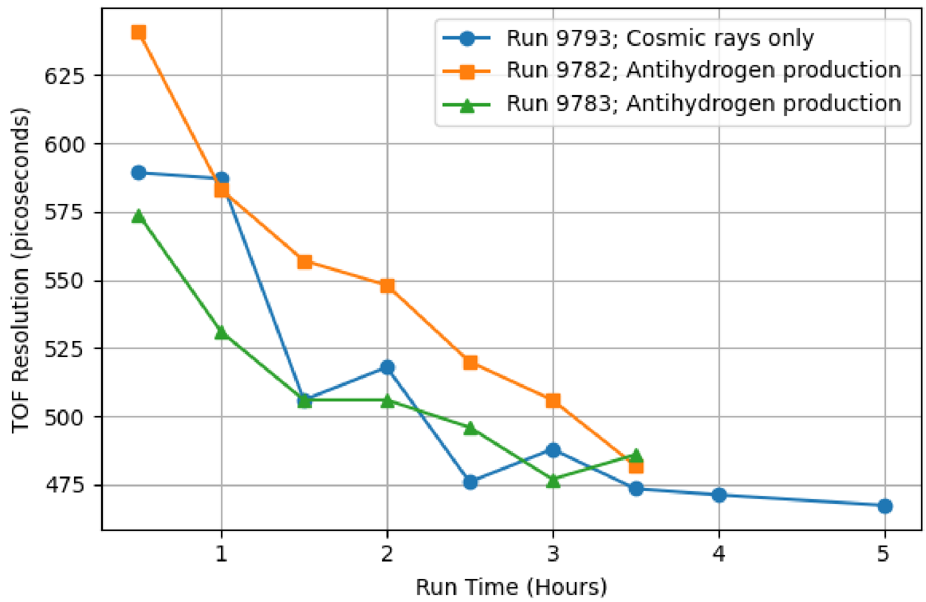

Figure D.1: Time-of-flight resolution dependence on the length of the dataset used for BSC calibration. Based on 2023 data which does not reflect TOF resolution improvements.

One major question is how much cosmic ray data is required for accurate calibration. This is addressed in Figure D.1, using data from 2023 when there were still some hardware issues limiting the TOF resolution. The resolution appears to flatten out with around 3 hours of calibration data, and so the threshold was set to use at least three hours of data for the 2024 run. Figure D.1 should be regenerated using 2024 data to update the data requirements.

A further area of investigation is the stability of the calibration constants over time. While they are known to change after a TDC power cycle, if they are otherwise stable, a single calibration file per inter-power-cycle period could be used for offline analysis.