Cluster Counting in Drift Chambers for Particle Identification and Tracking

by

Jean-François Caron

B.Sc. Hon. Physics, University of Calgary, 2006

M.Sc., The University of British Columbia, Vancouver, 2009

A THESIS 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)

December 2015

Abstract

Drift chambers are a type of gaseous ionization detector used in high-energy physics experiments. They can identify charged particles and measure their momentum. When a high-energy charged particle crosses the drift chamber, it ionizes the gas. The liberated electrons drift towards positive-high-voltage wires where an ionization avalanche amplifies the signal. Traditional drift chambers use only the arrival time of the cluster of charge from the closest ionization for tracking, and use only the integral of the whole signal for particle identification.

We constructed prototype drift chambers with the ability to resolve the charge cluster signals from individual ionization events. Different algorithms were studied and optimized to best detect the clusters. The improvements to particle identification were studied using a single-cell prototype detector, while the improvements to particle tracking were studied using a multiple-layer prototype. The prototypes were built in the context of initial work for the now-cancelled Super$B$ project, but the results apply to any drift chambers used in flavour-factory experiments.

The results show that the choice of algorithm is not as critical as properly optimizing the algorithm parameters for the dataset. We find that a smoothing time of a few nanoseconds is optimal. This corresponds to bandwidth of a few hundred megahertz, indicating that gigahertz-bandwidth electronics are not required to make use of this technique.

Particle identification performance is quantified by the fraction of real pions correctly identified as pions with at most 10% of real pions mis-identified as muons. In our single-cell prototype, the performance increases from 50%

to 60% of pions correctly identified when cluster counting is combined with a traditional truncated-mean charge measurement, compared to the charge measurement alone.

Tracking performance is quantified by the single-cell resolution: the uncertainty in measuring the distance of charged particle tracks from a given sense wire. In our multiple-layer prototype, the single-cell tracking resolution using traditional methods is measured to be $\sim 150\,\mu$m. With cluster counting implemented, the resolution is unchanged, indicating that the additional cluster information is not useful.

Preface

The first half of my research project was a study of the single-cell prototype particle tracking detectors which we call the TRIUMF prototypes. They are described in Section 4.2 and were built at TRIUMF by Philip Lu, Rocky So, and Wayne Faszer; the amplifiers used were designed and constructed by Jean-Pierre Martin at the Université de Montréal. A beam test was performed with the help of fellow Super*B* collaboration members and TRIUMF staff: Christopher Hearty, Philip Lu, Rocky So, Racha Cheaib, Jean-Pierre Martin, Wayne Faszer, Alexandre Beaulieu, Samuel de Jong, Michael Roney, Riccardo de Sangro, Giulietto Felici, Giuseppe Finocchiaro, Marcello Piccolo, Wyatt Gronnemose, and Steven Robertson.

With the data from the beam test, I performed a particle-identification study and the results were published in the paper titled "Improved particle identification using cluster counting in a full-length drift chamber prototype", published by Elsevier in the journal *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Volume 735, on January 21 2014, Pages 169-183 [1]. The article was written by me, except for the section describing the amplifiers which was written by Jean-Pierre Martin. The coding and analysis in the paper was entirely done by me, with consultation from the above-mentioned people in the Super*B* collaboration. The content of that article is included in Chapter 4, with permission from the publisher and co-authors. Jerry Va'vra lent us the MCPs for our TOF system, and Hirohisa Tanaka lent us the oscilloscope for our data acquisition. In addition to the authors of the resulting article, Wyatt Gronnemose and Steven Robertson

assisted during the beam test.

The second half of my research project involved the multi-cell prototype called "proto 2" or the Italian prototype, and is described in Chapter 7. The Italian prototype was constructed at the Laboratorio Nazionale di Frascati by myself and Riccardo de Sangro, Giulietto Felici, Giuseppe Finocchiaro, and Marcello Piccolo. The beam test was performed at TRIUMF by the same people from the TRIUMF prototype beam test. The coding and analysis for tracking was done entirely by me in consultation with Christopher Hearty.

Several contributions were made to open-source programs. The most notable is the addition of a new interpolation method to the GNU Scientific Library [2]. Many bug reports and some patches were submitted to the developers of ROOT [3] and Garfield [4]. The largest patch involved corrections to gender-specific pronouns used in documentation; it is described in Appendix A.4. The analyses in this work also made extensive use of these open-source programs: Python [5], IPython [6], PyROOT [7] and NumPy [8].

This work was supported by the Natural Sciences and Engineering Research Council of Canada and TRIUMF.

My supervisory committee members were Christopher Hearty, Janis McKenna, Joanna Karczmarek, and Kris Sigurdson. Sherry Leung assisted with editing and proofreading. Figures 1.3, 1.4, 7.4, 7.9, and 7.18 were made by Alon Hershenhorn.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank all my physics teachers and instructors over the years. Each of you have contributed piece-by-piece to my overall understanding of the physical world, and hopefully you continue to inspire others.

I want to thank my partner Sherry Leung for tremendous emotional support, and my family for setting me up to succeed in my endeavours.

While personal achievement is partially based on individual achievement, I must acknowledge that part of it is based on privilege. My work on this project was greatly assisted by my being born in Canada as a white cis-gendered male, heterosexual, into a middle-class family, and other categories. I did not choose to have this privilege, but I benefit from it regardless, and I am glad to be using it for something I believe is good.

Chapter 1

Introduction

1.1 Particle Physics

Particle physics is the study of the smallest and most-indivisible constituents of the physical world. What are currently called atomic physics and nuclear physics used to be seen the same way that particle physics is seen now. As we learned more about the constituents of the objects of study (atoms, nuclei), these fields revealed even more elementary objects. As far as we know, the elementary particles of particle physics cannot be divided or reduced any further.

In a sense, current high-energy particle physics is the limit of reductionist science. Reductionism is an approach where a complex system is investigated by breaking it down into smaller independent components. One can see a steady progression of this approach from early chemistry to atomic physics, nuclear physics, and now particle physics.

The terms "elementary" and "fundamental" are often added to particle physics to mean explicitly those particles that are indivisible even at the highest energies imaginable. There are still-divisible entities that are perfectly well-described at certain energies by particle models, such as protons. Though the terms are interchangeable, I prefer the term "elementary" because it has more reductionist implications - they are the simplest systems available for study.

In order to access the energy scales and sizes relevant to particle physics, enormous energies are required. Studying these interactions is a fruitful part of astrophysics, as these energies are routinely present in cosmic ray interactions in the upper atmosphere. Unfortunately cosmic ray collisions are not controlled enough and do not have the high rates required to study the rare and exotic interactions and species in particle physics. The ongoing success of particle physics experiments has resulted in the latest projects being some of the largest, most complex, and most powerful machines in the world.

1.2 Flavour Physics

Flavour physics is a branch of particle physics concerned with the transformation of quarks and neutrinos (elementary particles) into other kinds of quarks and neutrinos. It could be described romantically as a modern alchemy. Probably the most interesting result of flavour physics research is the discovery of quark mixing and CP violation. A good popular explanation of these phenomena can be found in the Fall 2006 issue of the physics@mit online newsletter [9].

Quark mixing is a phenomenon in which a given type of quark (up, down, strange, charm, top, bottom) is observed to spontaneously transform into another type. The explanation relies on the idea of two different "views" of quarks. In the "interaction" view, the relevant quarks are the ones we usually see in tables of the elementary particles. Quarks produced in interactions (including when we observe them) use this view. The other view is the "propagation" view, and in this view the relevant quarks are mixtures (quantum superpositions) of the quarks in the interaction view. The propagation view is used when quarks are moving through space and time. The propagation quark states are not given clever names, and since they have a majority component that is one of the interaction quarks, we typically just use the majority component's name with a prime ($'$). The mixture is symmetric, so one could think of an up quark as a mixture of up$'$, charm$'$ and top$'$ states, or one could think of an up$'$ as a mixture of up, charm, and

top quarks.

When a particle with a given set of quarks is created in an interaction, it is in a "pure" interaction state: its quarks are unambiguously up, down, or whatever. When a given particle is measured, e.g., detected in a particle detector, it is also unambiguously in a pure interaction state. However since there is always a time lag between these two events, the particle that we detect might not be the same kind as the one that was originally created. Along the way, some of the quarks might have transformed. The types of particles that can be observed are not ridiculous: energy, charge, and other physical quantities have to be conserved, but nevertheless this is an unusual process. No other known physics can change the identity of a particle simply by having it travel through space.

The most interesting consequence of quark mixing is CP symmetry violation. CP stands for "charge and parity" and represents a total inversion of charge and the spatial coordinate system. If you have a spin-up electron moving to the left and apply the CP operation, you now have a spin-down positron moving to the right. For a long time it was believed that the laws of particle physics were invariant under the CP transformation, i.e., that CP was a symmetry of nature. The strong and electromagnetic interactions do exhibit CP symmetry (as far as we know), but the weak interaction does not. The weak interaction is the one responsible for the quark mixing explained above.

A particular result of CP symmetry violation is found in the decay of neutral kaons. There are two states in which you can find a neutral kaon, called K^0 and $\bar{K}^0$, with different quark compositions ($d\bar{s}$ and $s\bar{d}$ respectively). Because of quark mixing, a K^0 can transform into a $\bar{K}^0$ and vice-versa, and kaons produced in experiments are quantum superpositions of these states. Neutral kaons predominantly decay via the weak interaction, but there are two kinds of decays: one with a short lifetime ($\sim 10^{-9}\,\mathrm{s}$) resulting in two pions and one with a long lifetime ($\sim 10^{-8}\,\mathrm{s}$) resulting in three pions. In an experiment where you let the short-lived components of neutral kaons decay away, there are still anomalous "long-lived" decays of the remaining neutral kaons to two pions. This would not be possible if the weak interaction did

not violate the CP symmetry. This is exactly the mechanism by which CP violation was discovered in 1964 [10], leading to a Nobel Prize being awarded to Cronin and Fitch in 1980.

An intuitive explanation (i.e., oversimplification) of CP violation is that there is a difference between matter and anti-matter, or equivalently, that physical processes that involve the weak interaction are fundamentally different going forwards or backwards in time. This description relies on the CPT theorem which is quite important for quantum field theory, but in some exotic theories (e.g., non-local theories, or those that break Lorentz invariance) there is no equivalence between CP violation and time-reversal.

1.3 Drift Chambers

In order to perform experiments in particle physics, we need devices that can detect the particles and measure their properties. There are a wide variety of such devices that detect different kinds of particles or measure different things, so a typical large particle physics experiment will employ many different types of detectors in order to fully characterize the physical processes that are happening. Drift chambers are one such kind of device [11].

Drift chambers consist of a volume of gas with thin metal wires strung throughout. When a high-energy charged particle crosses the gas (in $\sim 0.5\,\text{ns}$), it has a chance of interacting with the gas atoms or molecules. If the interaction frees electrons from the gas atoms, this is called ionization. The charged atoms that are left behind after freeing the electrons are called ions. Metal wires are held at high and low voltages by external electronics, creating an electric field to which the electrons and ions will react.

Drift chambers only detect charged particles; neutral particles are invisible to a drift chamber. They can determine the trajectory that a charged particle took when it passed through the detector. We call the trajectory the "track". If the drift chamber is in a magnetic field, its measurement of the track can be used to determine the particle's momentum. Drift chambers can also reasonably distinguish between different kinds of charged particles. For example they can tell the difference between electrons, pions, and kaons,

even if these all have the same momenta. "Reasonably" here means that the identification is sometimes ambiguous, and generally other types of detectors will be used to more confidently identify particles. An advantage of drift chambers is that their signals are very fast to obtain, thus the drift chamber in a big experiment is often used as part of a "trigger". The trigger is a signal to the whole experimental system that an interesting particle collision or event has happened, and thus to record the data from that interaction. Most well-understood but uninteresting events (called background events) are not recorded, saving considerable computer resources.

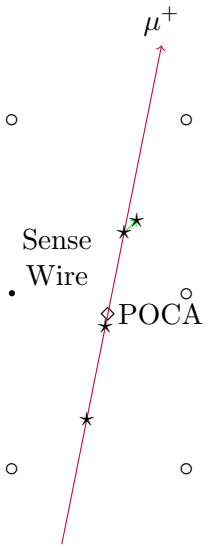
The negatively charged electrons will drift towards the high voltage wires (hence the name drift chambers), and the positively charged ions will drift towards the low voltage wires. The ions are much heavier, so they take much longer to drift all the way to the wire ($\sim 100\,\text{ns}$ for electrons, $\sim 1\,\mu\text{s}$ for ions). For most of the volume of the gas, the electrons drift at a constant speed, with the electrostatic attraction to the wires balanced by collisions with the gas. These collisions slightly change the direction in a random way. The average trajectory of the drifting electrons still goes towards the wire, but the side-to-side deviations from this path can affect the eventual timing measurement that will be made. This effect is diffusion. While the electrons have a constant drift speed, the electrons are slow enough that the collisions with the gas do not ionize it. This changes when the electrons get very close to the wires.

The high-voltage wires are very thin ($\sim 25\,\mu\text{m}$ is typical), while the low-voltage wires are typically thicker ($\sim 100\,\mu\text{m}$). Very thin wires can have huge electric fields near their surfaces because the electric field varies as V/r, with the minimum r being the radius of the wires. The electric field is so large near the surface (e.g., within a few μm) that the drifting electrons accelerate to high enough speeds to ionize more gas particles and liberate more electrons. These new electrons are also accelerated by the electric field, ionize more gas, and so on. Eventually the cluster of electrons is collected on the wire and the process finishes. This self-amplifying process is called an avalanche and is one of the critical processes in the operation of drift chambers. The ions do not form avalanches because the low-voltage wires

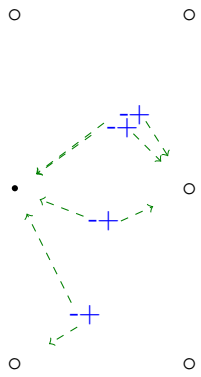
are much thicker and the ions are much heavier, so the electric fields do not accelerate the ions enough to further ionize the gas. This is intentional, because the avalanches create huge numbers of electron-ion pairs. An ion avalanche at a low-voltage wire would create lots of electrons that would eventually drift back to the high-voltage wire and create yet-more electrons, and vice versa.

The large number of electrons and ions created by the avalanche process can be thought of as two clusters of moving charge - one negatively charged and one positively charged (see Figures 1.1 and 1.2 for an illustration). These moving charges near the sense wire induce a current, which is picked up by the electronics at the end(s) of the wires. There, amplifiers will further increase the amplitude of the signal before sending it to a digitizer and finally into some processing systems so that the signal can be used. The electrons from the avalanche only move $\sim 1\,\mu$m before hitting the sense wire, while the ions must drift across the whole cell. This means that the majority of the induced signal is actually produced by the ion's movement [12]. The largest current is induced immediately when the charges are moving fast in the strong electric field near the sense wire ($\sim$ ns), but the ions drift for $\sim \mu$s before finally reaching the field wires on the edge of the cells. Thus the recorded signal has a sharp spike at the beginning followed by a long tail.

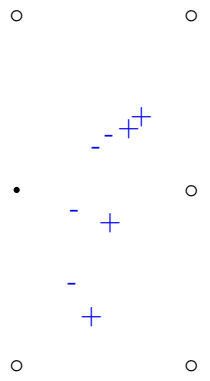
Each ionization event from the original high-energy charged particle produces a separate cluster, but almost always multiple avalanches will hit the same wire. The signal is recorded as a series of voltage measurements at regular time intervals, and the individual clusters would look like spikes with a sharp leading edge and a slower decay (e.g., Figure 7.18). The time at which the first spike is seen in the signal is called the arrival time. The number of spikes gives the number of original ionizations that resulted in a cluster hitting the wire. The integral of the entire signal yields a quantity proportional to how many electrons hit the wire in total.



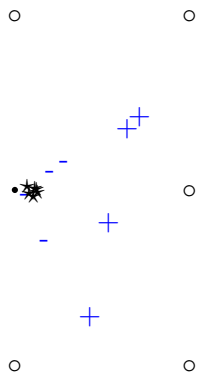
(a) Charged particle track, primary ionizations and delta-ray ($t \sim$ ns).



(b) Electrons and ion pairs created by ionization ($t \sim$ ns).

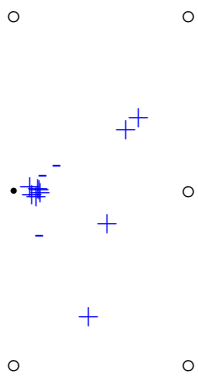


(c) Electrons drift towards the sense wire, ions are very slow by comparison ($t \sim 10\,$ns).

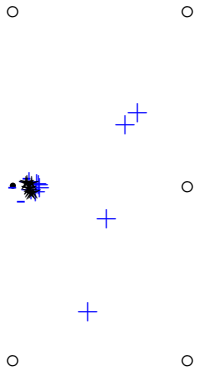


(d) Electron from the closest ionization creates an avalanche at the sense wire ($t \sim 10\,$ns).

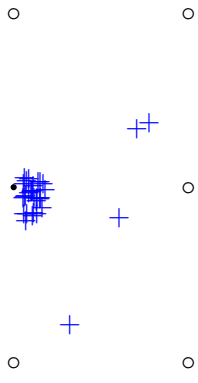
Figure 1.1: Illustration of the ionization, drift, and avalanche process. The figures show half of a drift chamber cell, with the sense wire on the left and field wires shown as open circles. The stars represent ionization sites, and the + and - symbols represent charges (not necessarily single electrons or ions).



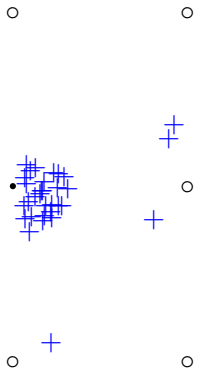
(a) Remaining electrons continue drifting, cloud of ions remains at sense wire from avalanche ($t \sim 10\,\text{ns}$).



(b) The remaining electrons also avalanche. Lots of ions are present at the sense wire ($t \sim 100\,\text{ns}$).



(c) Electrons are gone, ions continue slow drift towards field wires ($t \sim \mu\text{s}$).



(d) The ions continue to drift until they hit the field wires ($t \sim 10\,\mu\text{s}$).

Figure 1.2: Continued illustration of the ionization, drift, and avalanche process from Figure 1.1.

1.3.1 Tracking

If the drift properties of the gas are well-known (from calculation, simulation, or previous experiments), then the recorded signal can be used to infer information about the original high-energy charged particle. The most important of these is the time-to-distance relation: it is the distribution of arrival times for signals from high-energy particles passing at a given distance from the wire. If this is known, then a recorded signal can be used to infer how far the particle passed from the wire. The time-to-distance relation depends mostly on the gas composition and wire voltages, but also on the operating pressure and temperature of the drift chamber. Other examples of useful gas properties are diffusion coefficients and ion mobilities.

The arrangement of the wires in a drift chamber is tailored to the requirements of the experiment. Most often this is a regular grid of "cells" where a single high-voltage wire is surrounded by low-voltage wires (see Figure 2.3 for a diagram). The high-voltage wire in the centre of each cell through which currents are induced by the motion of electrons and ions is called a sense wire. The low-voltage wires are typically not instrumented as the ions do not create avalanches.

With a sense wire grid and a time-to-distance relation, the arrival times of the signals in all the cells crossed by a high-energy charged particle can be used to track a particle. Schematically, each individual sense wire's arrival time allows us to draw a locus of points or contour around the wire with the corresponding distance. The path taken by the particle is the line or curve that comes the closest to being tangent to all the contours simultaneously. Normally such a path would be required to be a straight line (e.g., Figure 7.8). Often drift chambers are operated in a strong magnetic field, which causes the charged particles to take helix-shaped paths. The radius of curvature of a particle's trajectory gives a measure of the particle's momentum, so this is highly desirable. Unfortunately the magnetic field also makes the drifting electrons inside the gas follow curved paths, so the design and calibration of the chamber is more complex.

The time-to-distance relation only gives information about the track in

the two dimensions perpendicular to the wire. From it and an arrival time, we determine the absolute distance from the sense wire of the charged particle's path through the cell. In general there is a full cylinder (or cylinder-like contour) around the wire that the track could touch. Two single-wire techniques are used to also determine where the charged particle passed along the axis of the wire itself (the so-called z-coordinate), and a third technique uses the information from multiple wires at the same time.

The first involves instrumenting both ends of the sense wire. The signals from a charge cluster hitting the wire propagate in both directions, but because the long sense wire has electrical resistance, the signal on each side of the wire will have a different amplitude (see Figure 1.3). The ratio of the amplitudes of the signals at the two ends of the wire gives a measurement of distance of the ionization event along the sense wire. A downside of this technique is the requirement of electronics on both ends of each drift chamber wire.

The second technique involves instrumenting only one end of the wire and deliberately letting signals reflect from the far end of the drift chamber back to the instrumented side (see Figure 1.4). The delay between the first signal and its reflection can be used to deduce the distance along the wire. This technique is less effective than the former because the reflected signal may be ambiguous when multiple clusters are involved, but it has the advantage of needing no additional readout electronics.

The third technique involves combining information from multiple wires. If the wires are not laid out all parallel to each other (so-called axial layout), then which specific cells are traversed by a charged particle will depend on the distance along the general direction of the sense wire axes. This technique (called stereo layout) is often used, because it has no drawbacks during operation. It does complicate the design and construction of the chamber however.

Another method ignores the wires and instead uses instrumented cathode strips on the inner and outer shells of the drift chamber. These strips are aligned to measure the z-coordinate independently of the drift chamber signals. Such a technique was employed for example by the CLEO [13] and

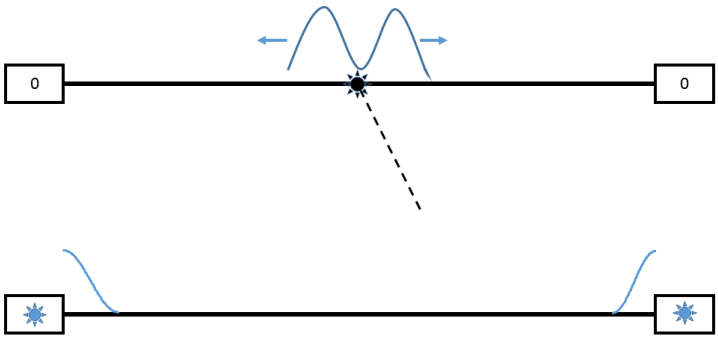


Figure 1.3: Illustration of the charge division technique. In the upper picture, back-to-back signals are created at a point on the wire. In the second picture, the signals have arrived at the instrumented ends, and the signal strength is attenuated based on the distance travelled.

Belle [14] detectors.

A final method for determining the z-coordinate is simply to use an additional detector for that purpose. If this extra detector is also a drift chamber, it is called a "Z-chamber".

1.3.2 Particle Identification

While tracking relies on a good understanding of the drift of electrons and ions in the gas, particle identification depends on understanding the ionization behaviour of the gas. The probability of ionizing a gas particle depends almost entirely on the speed of the charged high-energy particle in the chamber. This is represented by the Bethe formula[15, 16] which describes how much energy is deposited in the gas by the passage of the particle:

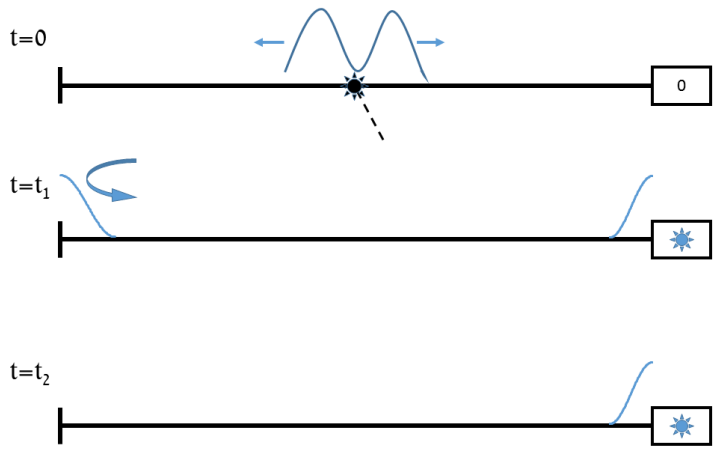


Figure 1.4: Illustration of the time delay technique. In the upper picture, back-to-back signals are created at a point on the wire. The signal propagating towards the instrumented end of the chamber arrives at time t_1. The reflection off the far end of the chamber reaches the instrumented end at time t_2.

$$\left\langle -\frac{dE}{dx} \right\rangle = Kz^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 W_{\max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right]. \tag{1.1}$$

In the above formula, Z and A are the atomic number and atomic mass of the atoms in the gas, K is a combination of physical constants, and z is the charge of the high-energy particle (in units of the elementary charge). m_e is the mass of the electron, c is the speed of light, β is the speed of the high-energy particles in units of c, and

$$\gamma = \frac{1}{\sqrt{1-\beta^2}} \tag{1.2}$$

is called the Lorentz factor. I is the mean excitation energy of the gas atoms ($I \approx 10\,\mathrm{eV}Z$), and $\delta(\beta\gamma)$ is called the density correction. The density correction becomes important only at very high energies and is more important

for higher-density absorbing materials like liquids and solids than for gases. $W_{\max}$ is the maximum kinetic energy that can be imparted on a free electron in a single collision by the high-energy particle. $W_{\max}$ depends very slightly on the mass of the high-energy particle, but this is only relevant at very high energies.

For the typical particle energies expected in the drift chamber of a flavour-physics experiment, the Bethe formula depends on β^2 alone, the other terms being essentially constant. If the momentum p of the particle is known, for example from measuring the curvature of its tracks in a magnetic field (as described in Section 1.3.1) or from another detector, then the mass m can be determined:

$$pc = \frac{mc^2\beta}{\sqrt{1-\beta^2}}. \tag{1.3}$$

The mass uniquely determines the kind of particle.

The probability of ionizing a particle determines how many ionization events will occur within the boundaries of a cell. Each ionization event will liberate one or a few electrons which drifts and creates an avalanche. The recorded signal is the combination of the charge clusters from all the ionization events. Thus the number of clusters or the integrated signal can be used to determine the particle's type. With a fixed input impedance Ω for the data acquisition, the integral of the signal over time is the charge deposited on the wire.

$$\frac{1}{\Omega}\int V(t)dt = C \tag{1.4}$$

The traditional method of particle identification is to use the integrated current, as this is simple to perform (and can even be done before digitization using analog electronics). A particle crossing a drift chamber will ionize gas in multiple cells along its track, so to improve the resolution, the charges from all the cells in a track are added up and averaged over the track length. Even with many cells in a track, the difference in the sum of signal integrals between different types of particles can be relatively small. Unfortunately the charge-per-unit-track-length statistic cannot be used directly, because

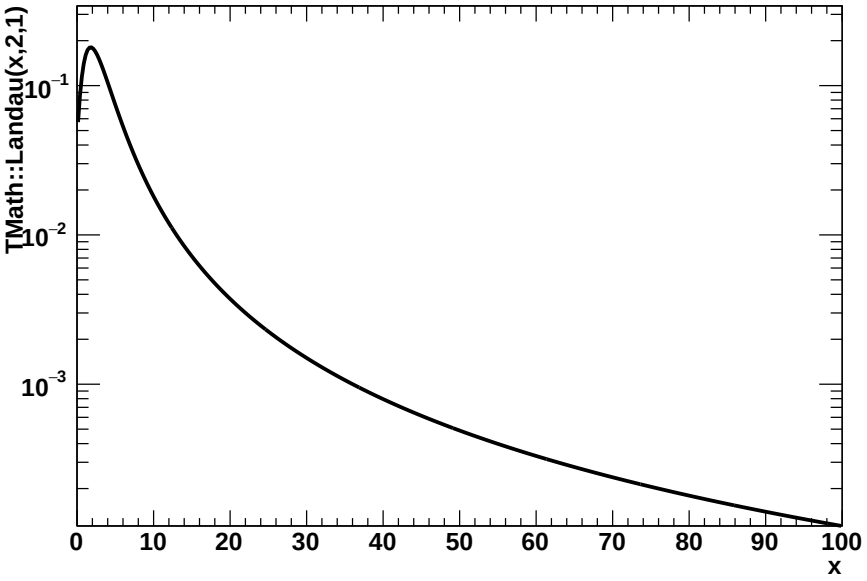


Figure 1.5: Plot of a Landau distribution plotted with parameters $\mu = 2$, $\sigma = 1$ using ROOT's `TMath::Landau` numerical calculation. Note that μ and σ are *not* the mean and standard deviation. The μ parameter determines the location of the peak (most likely value), while σ controls the width.

it follows a difficult-to-use distribution. The probability density function is called a Landau distribution and one of its features is that the mean and standard deviation are not well-defined (see Figure 1.5 for an illustration). The calculations for the mean, standard deviation and indeed all of its moments are divergent (they are infinity). With Normal-distributed quantities, one would ask "how close is this particular measurement to the mean value of this quantity, in units of the standard deviation?" With the Landau distribution, there is no obvious way to do it, so alternate methods of quantifying the data must be used. The divergent mean and standard deviation of the Landau distribution are due to rare large fluctuations that contribute to the sums and integrals in the calculations.

The reason that the total charge statistic follows a Landau distribution

is because it results from the combination of several physical processes. The initial ionization is a Poisson process, so the number of individual ionization events (each contributing a cluster of charge) follows a well-behaved Poisson distribution. Each primary ionization can also release multiple electrons. Though the most likely number is one, there can be a reasonable probability that up to a few electrons are released at a time [12]. The precise distribution depends on the drift chamber gas and incident particle energy. In addition to all this, the primary electrons that are produced in the ionization by the passage of the charged particle can sometimes have rather large momenta. These are called "delta rays", and can travel a short distance in the drift chamber cell and make their own "secondary" ionizations. The electrons from the secondary ionization also go through the drift and avalanche processes, and appear indistinguishable from primary signals. The derivation of the theoretical deposited charge involves the convolution of all these processes, and generally includes integrals over the number of electrons released and over electromagnetic frequencies up to infinity. In a real physical situation obviously there are limits on these quantities, but this does not change the qualitative statement that large rare fluctuations prevent the use of simple statistical treatments. The empirical measurement of the deposited charge is well-modelled by a Landau distribution, if you have a cut-off of some kind (e.g., a histogram with finite range).

A robust method to deal with the Landau-distributed quantities is to perform a so-called "truncated mean". Instead of using the total charge of all the signals per unit track length, first the individual charges from the different cells are sorted and only the values below a certain percentile rank (typically $\sim$ 80%) are used. The rest of the values are discarded, and these will contain the large-but-rare fluctuations from the delta rays. The resulting truncated mean statistic is well-behaved and can be easily analyzed. Unfortunately this method necessarily throws away information, effectively reducing the number of cells that can contribute to the charge measurement.

An alternative to the deposited-charge statistic is the number of clusters. This skips the variations due to the number of primary electrons and

the amplification process, so the statistic should be well-behaved. Unfortunately resolving the clusters can be difficult, as the electronic signals from multiple clusters can pile up and look like large-amplitude single clusters. Further, spurious noise in the signal can fool the cluster-counting algorithms into counting fake clusters, and some algorithms may be too complex to run online (i.e., as fast as the events are read out in the detector). The choice and tuning of the cluster-counting algorithm is critical to obtain good performance. Certain design choices can make this process easier, for example fewer clusters will pile up if a gas with a slower drift velocity is chosen.

Once a suitable statistic is found (either truncated mean of the deposited charge per track length or number of clusters per track length), it is a simple matter to build up distributions of these quantities for different particle species. This can be done by using known radiation sources or processes in which the particle identities are known, or an existing particle identification device such as a time-of-flight counter. A time-of-flight counter consists of a pair of charged particle detectors placed a large distance apart so that the time delay between the coincident signals is measurable. The distance between the detectors is measured precisely, so the time delay is converted into a speed. With a known particle momentum (e.g., from a bending magnet, tracker, or from the source), the speed gives the mass and thus identity of the particle.

Once one has validated distributions of the truncated mean of the deposited charge or number of clusters for the various relevant particle types, one can measure the same statistic for an unknown particle and assign a likelihood of being a given particle. For drift chambers, the only relevant types of particles are electrons (and positrons), protons (and antiprotons), deuterons, charged pions, muons, and kaons. Other charged particles are not long-lived enough to leave measurable tracks in the detector. Depending on the energy range, possibly only one or two particle species may be relevant, simplifying the analysis.

Chapter 2

Design Considerations of Drift Chambers

2.1 Overall Design

Drift chambers are general-purpose detectors of high-energy charged particles, but each is designed with a specific experiment in mind. They are typically one-of-a-kind, designed and custom-built by researchers and technicians in particle physics and are not commercially produced by companies (like e.g., photomultiplier tubes). The various design choices are made considering the requirements of the experiment at hand. Because of the dependence on these details, the range of designs of drift chambers that can be found in the literature is quite broad.

For Super*B* and the prototypes used in this work, the primary design consideration was the amount of material. Because of the types of analyses done with B-factory data, very precise identification and determinations of tracks of charged particles with $\sim 1\,\mathrm{GeV/c}$ or less momentum is very important. The main source of uncertainty in these measurements is the elastic scattering of the charged particle on material in the drift chamber. The term used is "multiple scattering" because often there are multiple such scatterings. Multiple scattering also affects the measurements made by other detector components, because it changes the momentum of the

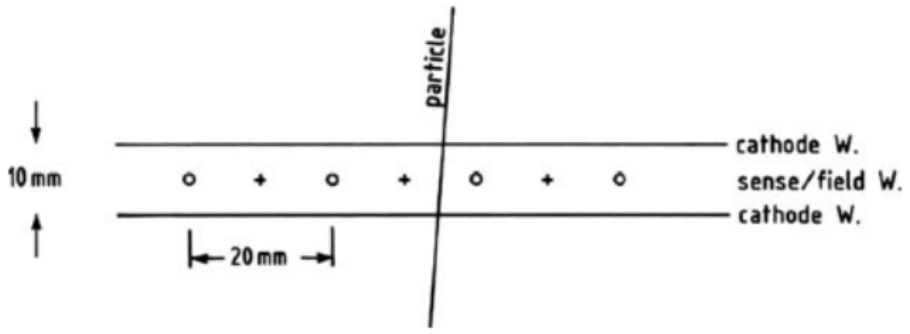


Figure 2.1: Schematic of a planar chamber layout. Taken from [12] where it is Figure 11.2, on page 364.

particle while it is traversing the drift chamber. The material involved is everything in the drift chamber: the gas, the wires, and the external casing, and even the front-end electronics.

There are a few canonical categories used to describe drift chamber designs, and three of which will be described here.

Planar chambers feature just one (or just a few) layer of sense wires (see Figure 2.1). Since a single high-energy particle will likely only excite one cell, a single planar chamber cannot be used for tracking by itself. They are often stacked (with the wire orientations orthogonal) or combined with other detectors (see [17] for an example).

Jet chambers have sense wires arranged in "rays" or "spokes" to better capture the signals left by jets of particles coming from the hadronization of high-energy quarks and gluons. In a jet chamber the cells are rather wide, so a low-diffusion gas and high electric fields are needed. An example jet chamber design can be seen in Figure 2.2.

Cylindrical drift chambers with a "small cell" design have become standard for flavour physics experiments. An example is the ARGUS drift chamber, whose cell layout can be seen in Figure 2.3. The small cells measure around 1 cm on each side and are designed to "tile" nearly the whole cross-section of the drift chamber gas. Counterintuitively, drift chambers with

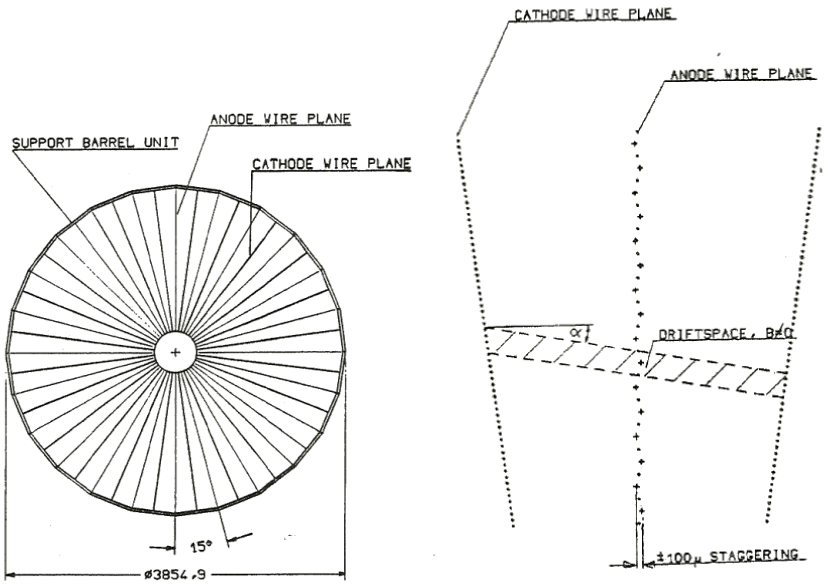


Figure 2.2: Schematic of the OPAL Central Jet Chamber. The figure on the left shows the arrangement of the radial sections. The figure on the right shows the individual wires in part of a section. Taken from [18].

small cells require less material than large-cell chambers, even though the number of wires is greater. This is because other designs like jet chambers require high electric fields far away from the wires in the large cells, so the wires are made thicker to avoid having spontaneous ionization of the gas by the electric field near the wire surface. The thicker wires need higher tension to reduce gravitational sag, and thus the endplates must be thicker and stronger as well. Small cells have other advantages: electron diffusion is less important because it is proportional to the drift distance; the charge accumulation on each wire is lower, delaying ageing; the tiling effect of the small cells reduces the amount dead space in the detector; and the lower tension in the wires means result in less creep.

Creep behaviour of the wire material should be well-understood. Creep is

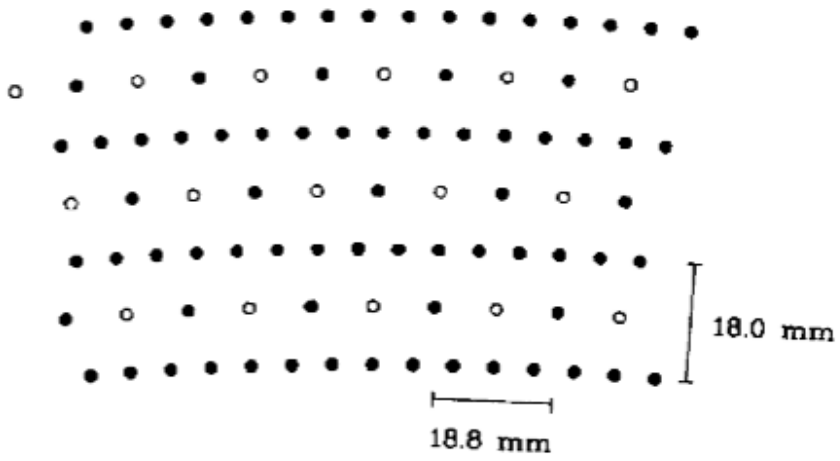


Figure 2.3: Schematic of a cylindrical chamber layout for the ARGUS drift chamber. The empty circles are the sense wires, and the black circles are field wires. This configuration is identical to the one planned for Super*B*. Taken from [12] where it is Figure 11.14, on page 380.

the slow permanent deformation of a material experiencing sustained forces. In the case of a wire, the creep effect would be to slightly lengthen, lowering the tension and increasing the gravitational sag. To compensate for material creep the wires are given extra tension upon stringing so that they have the proper tension after a certain expected amount of creep. More details about creep are given in Section 2.3.

2.2 Gas Composition

Generally drift chambers use a mix of gases, the major component typically being a noble gas. The minority component(s) is typically a simple hydrocarbon (e.g., ethane) or carbon dioxide, and is called a "quencher". The process of ionization produces ultraviolet photons, which can cause undesired further ionization via the photoelectric effect. The quencher's role is to

absorb these ultraviolet photons, so quencher gasses are chosen to strongly absorb ultraviolet radiation. The ratio of the gas components varies quite a lot, from $50:50$ [17] to $97:3$ [19]. Sometimes additional minor gas components are added, usually to address ageing. For example it is believed that a small component of water vapour helps reduce or temporarily reverse the effect of drift chamber ageing. See Section 2.5 for more details.

A traditional primary gas of choice for drift chambers is argon, but drift chambers that need to minimize multiple-scattering now use helium as the primary gas, as it is the lightest monatomic gas. Helium unfortunately has a high electron diffusion constant (i.e., the amount that the electrons diffuse per unit drift time), so it must be used with a small-cell design to minimize the total diffusion [20]. Helium has a much higher ionization potential than argon, the downside of which is that higher wire voltages are required to create measurable avalanches on the sense wires (the electrons must be accelerated to higher momenta before ionizing more atoms). It was revealed with some surprise that good results could be obtained from a helium-based drift chamber [21]. A benefit of helium's higher ionization potential is that the low-energy photon background from the beam has a smaller chance of triggering an avalanche and being registered as a signal. For high-rate environments such as flavour factories this is essential [22].

To illustrate how helium can reduce multiple-scattering, consider that the momentum resolution of a charged particle with momentum p can be well-parametrized by a function of the following form [23]:

$$\frac{\sigma_p}{p^2} = \sqrt{A^2 + \left(\frac{B}{p}\right)^2}). \tag{2.1}$$

Momentum resolution is reported as the uncertainty on a momentum measurement σ_p divided by the square of the momentum p. This is because the uncertainty of the momentum measurement generically is proportional to p^2, so the ratio better shows the dependence on other factors [12]. In Equation 2.1, A is the fixed contribution from the measurement errors like time resolution and wire position. For an argon-based gas, A is typically around

0.5 to 1.0 %/GeV, while for helium-based gases it can be significantly worse. The second term with B is the contribution from multiple scattering, and is inversely proportional to the momentum of the charged particle. B for an argon-based gas is typically around 0.7%, but for a helium gas it is around 0.3%.

For high momentum p, the dominant contribution to σ_p/p^2 is from the first term, so the difference in B values is negligible. Thus drift chambers built to detect very high energy particles will use a gas mixture based on its A value. In flavour-factory experiments like Super$B$, most charged particles to be detected have lower energies (less than 1 GeV/c), so the B contribution can be dominant. Most flavour-factory drift chambers use helium as the primary gas component because it has a smaller B value even though its A value is worse than argon.

2.3 Wire Materials

Several materials have been used for drift chamber wires. The materials must be non-magnetic in order to not react to the external magnetic field that allows us to measure particle momenta. They also need to be strong enough to avoid breaking, because some tension is applied to minimize the gravitational sag. The wires should be conductive enough to carry the raw electronic signals to the external electronics, and should not react with the gas.

As mentioned in Section 2.2, multiple scattering is only a concern for flavour-factory drift chambers, as another process dominates the momentum resolution at higher momenta. In argon-based drift chambers for which multiple scattering is not a concern, the gas alone contributes so much material that the wire's material is negligible. For helium-based drift chambers, the material in the wires is quite significant.

The "creep" behaviour of the material should be well-understood. Creep is the slow permanent deformation of a material experiencing sustained forces. In the case of a wire, the creep effect would be to slightly lengthen the wire, lowering the tension and increasing the gravitational sag. To com-

pensate for creep the wires are given extra tension upon stringing so that they have the proper tension after a certain amount of creep. Creep is a bigger factor in aluminium wires than other materials, and experiments using aluminium wires generally do a study of the creep behaviour during development [24].

The material used for sense wires and other wires is generally different. Sense wires must be very thin in order to generate strong electric fields near their surface for electron multiplication. The field wires are typically much thicker so that the surface fields are smaller to prevent the onset of the Malter effect (see Section 2.5). The thin sense wires must be made of stronger and more conductive material, since mechanical strength and conductivity both decrease with wire diameter. The mechanical and electrical limits to the sense wire diameters are about 20 μm [25]. Thinner than this, the wires are difficult to string and tend to break when attached to the feedthroughs on the endplates. Thinner wires having higher electrical resistance may be advantageous for the charge-division method of determining the z-coordinate [26].

There are a few canonical wire materials used in drift chambers. The materials used in a few historical drift chambers are shown in Table 2.1. Stainless steel and copper-beryllium alloys are used because they have similar thermal expansion coefficients to aluminium [25]. Aluminium is the typical material used for endplates and supporting structures, and having the whole chamber expand and contract at the same rate during temperature fluctuations is obviously desirable. Unfortunately stainless steel is not normally antimagnetic, so special non-magnetic wires must be obtained [27].

For small-cell low-mass drift chambers, the canonical material of choice for the sense wires is gold-plated tungsten-rhenium. The tungsten-rhenium material has high strength, can be made into very thin wires, and has acceptable creep characteristics. Gold coating is needed to improve the conductivity for the electronic signal and to provide a non-reactive surface.

The Italian prototype proto 2 used in this study had six layers of cells with molybdenum sense wires instead of tungsten-rhenium (Figure 5.1). Molybdenum has roughly the same mechanical properties as tungsten-rhenium,

Table 2.1: Table of wire materials used in the drift chambers of various historical experiments.

Experiment	Sense Wire	Field Wire	Notes
ISR Final Design	Ni-Cr		High resistivity for charge division [26]
ISR Initial Design	25 μm stainless steel	100 μm Cu-Be	High-resistivity sense wire for charge division, special non-magnetic stainless steel [27]
UA1	35 μm Ni-Cr	100 μm Au-coated Cu-Be	[28]
ARGUS	30 μm W	76 μm Cu-Be 76 μm field	Majority of scattering is from the gas [19]
CLEO-II	20 μm Au-coated W	100 μm Au-coated Al and Cu-Be	Al for inner layers, Cu-Be for outer to save money, as Alu wire is expensive [29]
CLAS at CEBAF	20 μm Au-coated W	150 μm Au-coated Al	[30]
KEDR	28 μm Au-coated W	150 μm Au-coated Ti	[31]
CDC at SLD		150 μm Au-coated Al	Sense wires unspecified [32]

but has much lower density and electrical resistivity. The density of molybdenum is nearly half that of tungsten, and tungsten's resistivity is roughly 9.5 $\mu\Omega$cm [33] while molybdenum's is 5.3 $\mu\Omega$cm [34]. Though the gold coating carries much of the signal, a lower-resistivity wire material may be beneficial. The other two layers in the eight-layer prototype were the canonical tungsten-rhenium for comparison. No detailed study of the performance of the molybdenum cells in proto 2 was performed. Only one other drift chamber seems to have ever used molybdenum wires: the inner drift chamber for the TOPAZ experiment at TRISTAN [35]. Only the TOPAZ field wires are molybdenum, the sense wires are the more standard tungsten-rhenium. They do not state their reasons for the choice of molybdenum, but it does not appear to be for the minimization of multiple scattering based on the kinds of experiments done at TRISTAN.

The canonical choice for field wires is larger-diameter aluminium. Large diameters are used to avoid the Malter effect (Section 2.5). Aluminium is chosen because it has a low density for a metal, so the large-diameter wires contribute less material. Most experiments use gold-coated aluminium, for conductivity and to provide an inert surface. BaBar used gold-plated aluminium, and Belle used bare field wires. Bare aluminium wires tend to have more surface defects and are subject to oxidation, but no conclusive analysis has been done to determine whether the gold coating is significant. Belle's experience shows that the gold coating is not essential, and the gold coating contributes a non-negligible portion of the total material of the wires. Belle-II is continuing with bare aluminium field wires, and Super$B$ was also designed to use bare wires.

2.4 Outer Structure

The outer structure of a cylindrical drift chamber consists of the inner support tube, the exterior barrel, and the endplates (see Figure 2.4). The outer structure provides the physical structure to hold the wires, provides a gas seal, and provides a grounded electromagnetic shield for the inside of the chamber.

Because of the large number of wires under tension (typically thousands), the structure has to resist forces on the order of tonnes, even if each wire has a reasonable tension on the order of grams.

The endplates of large drift chambers are commonly made of aluminium, but some are made of composite materials like carbon fibre. A finite-element analysis is done of the endplate deformation in order guarantee that the structure will not collapse, and in order to adjust wire tensions to account for a changing endplate position and shape as more wires are attached. The shape of the endplates also varies: some are flat, some conical or hemispherical, and some with wedding-cake-like steps [36].

The outer cylinder is always load-bearing, and takes most of the force from the endplates. Typical materials are again aluminium or carbon fibre. A construction strategy must be devised to hold the endplates during

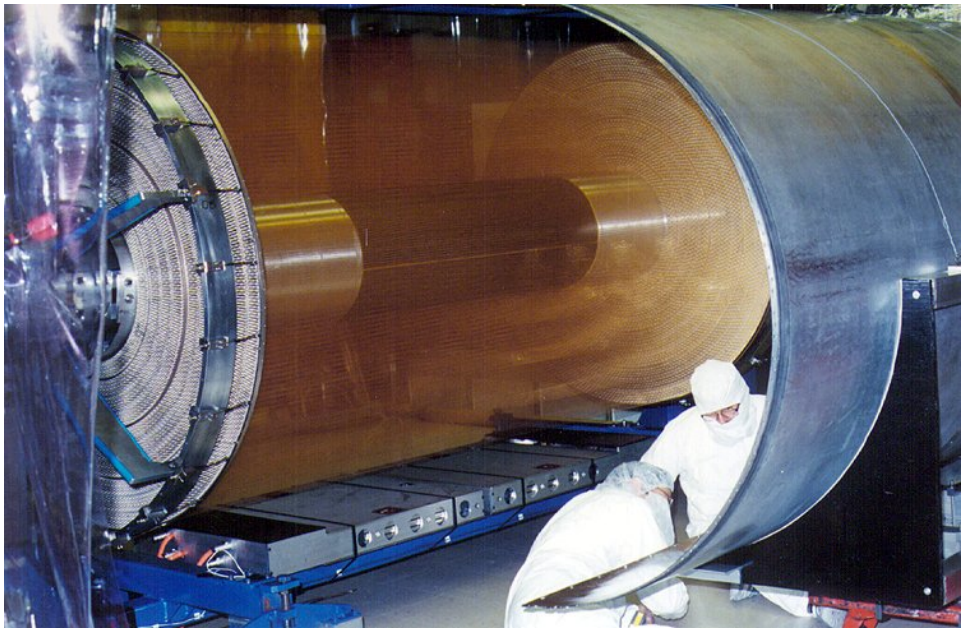


Figure 2.4: Photo of the BaBar drift chamber during construction. The inner cylinder and endplates are visible, as is the external mechanical support for the endplates (the blue claw-like pieces). The outer cylinder plates are being installed, after which the external supports will be removed. Photo from Chris Hearty.

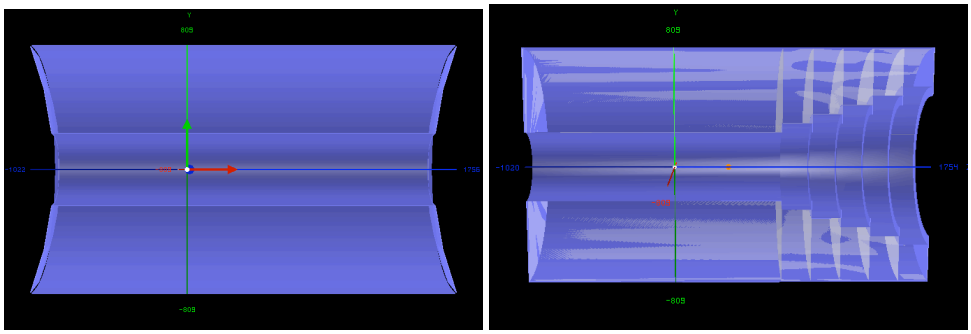


Figure 2.5: Example of two endplate designs considered for Super*B*. The design on the left has concave spherical endplates, while the one on the right is a stepped "wedding-cake" design. Taken from the Super*B* Technical Design Report [37]

stringing before the outer cylinder can be attached. For example external struts may be attached to support the endplates. After stringing, the struts are removed as the outer cylinder plates are attached.

The inner cylinder of the chamber is sometimes load-bearing, but sometimes it is only a ground plane and gas seal. In some designs, the inner cylinder is the beam-pipe itself. Even for load-bearing designs, the material tends to be thinner and lighter, such as carbon fibre or porous aluminium cells. BaBar used a beryllium inner cylinder [38]. Beryllium is a strong material that has a remarkably long radiation length (the average distance that particles will travel before interacting), but is difficult to machine and its dust is extremely toxic. Non-load-bearing inner cylinders may consist simply of a conductive foil (e.g., aluminized mylar) for grounding and gas sealing.

2.5 Ageing

Ageing is a term used to describe the gradual degradation of the performance of a drift chamber over the lifetime of an experiment. The observable phenomena are: a reduction in the gas gain when operating parameters that affect the gas gain are held constant (e.g., the wire voltages and gas composition) and persistent currents on the wires in the absence of traversing particles. A consequence of lower gain is lower-amplitude signals relative to the intrinsic noise of the system, and thus reduced accuracy of measurements of the signal arrival times and charge integrals. When threshold algorithms are used (e.g., for arrival times and cluster counting), the thresholds selected at the beginning of the experiment may be too high after some ageing has occurred, spoiling all the measurements. The information in this section is mostly taken from an extensive review of drift chamber aging by Jerry Va'Vra at SLAC [39]. Research is ongoing on understanding ageing, both from a microscopic perspective [40] and from a pragmatic perspective [41] that wishes only to mitigate it. During the beam tests described in this thesis, ageing tests were performed by Rocky So. The work forms part of his PhD thesis [42].

The mechanism of ageing is complex, but can generally be understood as a buildup of material on the sense and field wires. During normal operation of the drift chamber, there is a lot of ionization of the gas from primary and secondary ionizations, and the avalanche process. For most drift chamber gas mixtures, the ions of the main component (argon or helium) are non-reactive, but the ions, fragments, and radicals from the quencher gases can participate in rather complex chemistry. Contaminants in the gas will also play a role, if present (e.g., outgassing molecules from a glue or cleaning product). For example, polymers can form from the fragments of broken-up hydrocarbon molecules. The final products will accumulate on the surfaces inside the drift chamber: the wires and the outer structure.

The build-up of material on the drift chamber inner surfaces can have several effects. On the very thin wires, the material can effectively increase the wire diameter. The material build-up is typically non-conductive, so charges arriving from ionization processes can build up and slowly diffuse through to the metallic surfaces below. In all these cases, the result is a reduced signal amplitude from smaller avalanches and insulated conductors.

The extra material may also increase the gravitational sag of the wire. This is likely negligible, as the wire masses are $\sim 10 - 100\,\mathrm{mg}$ and the deposits are negligible. No study has yet examined this possibility.

Extreme cases of ageing can result in so-called "dark currents". These are persistent currents in the wires of a drift chamber even when no high-energy charged particles are crossing the chamber. The mechanism causing these dark currents is called the Malter effect [43]. The effect is that the non-conductive coating on the cathodes (the electrodes towards which the ions drift, i.e., field wires) prevents the ions from being collected by the cathode. The positive charge thus accumulates and attracts electrons from inside the cathode. If the electric field around the cathode and the positive charge buildup is strong enough to overcome the work function of the surface, these electrons are ejected into the gas rather than recombine with the ions. The electrons then drift towards the sense wires and cause avalanches, releasing more ions. This self-sustaining and self-amplifying process explains the currents which are present even without high-energy particles crossing

the chamber.

Since ageing is produced by the ionization process in the chamber, the duration is quantified using "deposited charge" accumulation: the total amount of charge deposited on the sense wires per wire length. This means that drift chambers operating in high-rate environments will age faster, all other things being equal. Ageing is a primary design consideration for drift chambers intended for high-rate experiments or that will operate for a long time [44].

The chemical deposits responsible from ageing come from the ionization of the quencher, so a smaller quencher fraction is likely to reduce the rate of ageing. Some quencher is always needed as described in Section 2.2, so it cannot be eliminated entirely. The choice of quencher may also be important, but among the lower hydrocarbons (methane, ethane, propane, butane, the most common quenchers) no differences are observed.

Larger wire diameters provide a larger surface area for the material to build up, which would reduce the ageing effect. Unfortunately thicker wires would mean more tension is needed (and thus stronger endplates and outer structure), and higher operating voltages would be needed to provide the same gas gain. Unless the material is changed, thicker wires also increase the amount of material in the chamber, leading to more multiple scattering.

The overall gas flow rate can influence ageing. Some materials like plastics and glues can slowly release contaminants into the drift chamber volume. A drift chamber or its gas system may also have hard-to-detect leaks where contaminants can enter. Thus most chambers continuously flow gas through the chamber, filtering and re-circulating some of the gas. As an example, the BaBar drift chamber flowed 15 L/min of gas, of which 12.5 L/min was recycled and filtered [38]. The BaBar drift chamber was also kept at a small 4 mbar pressure above the ambient atmosphere so that any leaks would be to the outside, thus preventing contamination. Sealed gas ionization detectors do exist (e.g., [45]) but they require careful sealing procedures and high-purity gases.

Small quantities of impurities in the gas are observed to have significant effects on ageing [39]. Some of these impurities can have positive effects, by

delaying or even apparently reversing the ageing effect. There is a bit of folk wisdom associated with these additives, as some were discovered by accident and not systematically studied. For example it is thought that molecular oxygen in the gas can combine with hydrocarbon radicals to produce non-reactive molecules. Similarly the addition of water vapour is thought to "reverse" the effect of ageing by embedding itself in the insulating deposits on the wires and increasing their conductivity [41].

An example of ageing can be found in the report of the performance of the BaBar detector [46]. Over the lifetime of the BaBar experiment the drift chamber wires accumulated $\sim 34\,\mathrm{mC/cm}$ of charge, and the gain (after adjusting for changes in voltage) was observed to drop by $0.337 \pm 0.006\,\%/(\mathrm{mC/cm})$. To compensate, the voltages on the sense wires were increased periodically to maintain roughly the same gain. The voltages cannot be increased arbitrarily, as the wires interact electrostatically with the fields. The effect is that the wires' gravitational sag is enhanced by large voltages, and the sag can be large enough to break the wire or make it touch other wires. The effect depends on the wire tensions, but the voltages required for instability are on the same order as the typical operating voltages [12].

Chapter 3

Super$B$

3.1 Introduction

Cutting-edge accelerator-based high-energy particle physics experiments can be classified into two types. So-called "energy-frontier" experiments collide particles at heretofore-unreached centre-of-mass energies or energy densities [47]. A popular example of this approach is seen in the ATLAS and CMS experiments at the LHC, where protons or ions are collided at the highest energies ever obtained in a laboratory. Modern energy-frontier experiments are huge enterprises, involving entire consortia of countries and requiring significant industrial support. The LHC and its experiments are among the most complex machines ever constructed by humanity. It is admirable that these machines are not used for war, as with many other devices in this category. There are real technical, economic, and political reasons why energy-frontier experiments are difficult to improve. Other examples of energy-frontier experiments include the Tevatron[48] and the cancelled Superconducting Super Collider[49]. The only currently planned future energy-frontier experiments are China's Circular Electron Positron Collider (CEPC [50]) and the International Linear Collider (ILC [51]). The ILC has been in planning stages for decades, and some consider it an eternally-hypothetical experiment that may never be built.

The other type of cutting-edge particle experiment is the so called "intensity-

frontier" type [52]. Rather than trying to directly reach the energies required to produce new particles or unlock new physics processes, high-precision measurements are made of reactions at more moderate energies but at extremely high rates. Through higher-order interactions and virtual particles, currently unobserved particles and unknown new physics can contribute to the results. A famous example of this is the ARGUS experiment being able to constrain the mass of the at-the-time undiscovered top quark despite not being able to produce real top quarks[53, 54]. If there are indeed new physics processes or particles, their presence would slightly shift the values of reaction rates from those expected under the regular standard model. In order to measure these slight shifts, lots of data are required, hence the focus on high-precision measurements at high rates. Intensity-frontier experiments tend to be of a much smaller scale than energy-frontier ones, capable of being undertaken by single countries (albeit only rich first-world countries) and by consortia of universities. Examples of intensity-frontier physics are the BaBar and Belle experiments, which focused on the production of B mesons to study CP violation and other flavour physics topics.

Super$B$ was a planned particle physics experiment of the intensity-frontier type led by the national Italian particle and nuclear physics institute, the Instituto Nazionale di Fisica Nucleare [55]. It was to be a so-called super-flavour-factory, and a showcase of new technologies designed to tease out the details of physics beyond the standard model. One of the repeated promotional phrases was that when high-energy experiments (e.g., the LHC) find evidence of new physics, Super$B$ will ascertain exactly what kind of new physics has been found.

Super$B$ was to be the successor to BaBar, and indeed would have reused some of the components of the detector, and many senior scientists from BaBar were involved in the planning and design stages of Super$B$. Unfortunately during these initial stages of the project, the budget ballooned to 1 billion euros from an original estimate of 350 million, and the newly-appointed Italian government cancelled the project. Much work was done before the cancellation in predicting the physics results obtainable with Super$B$ (e.g., [56]), and many prototype components were designed and con-

structed [37, 57–59].

A successor to Belle is currently being developed in Japan, called Belle II. Compared to Super$B$ and BaBar, Belle II is a less ambitious upgrade of Belle, but it has the advantage of being in the same location as its predecessor and undertaken by a country in a better financial situation than Italy. It is less ambitious in the sense of mostly following the design of the original Belle experiment with incremental upgrades [60]. In comparison, Super$B$ was to be in a different country than its ancestor, was to be built at a brand-new facility with a new accelerator using new technology [61], and had significant changes in its design.

3.2 Super$B$ Drift Chamber

From the original conception of Super$B$ until its cancellation, much initial work was done to determine the overall design and features of the accelerator and detector. The most detailed description can be found in the Technical Design Report [37]. Much of the design was inherited from the experience of the BaBar [38] and Belle [14] experiments.

For example BaBar found that the hexagonal cells combined with the super-layer design resulted in unnecessary dead space between layers. Dead space is a region of gas with no associated sense wire, and does not contribute to the measurements. The Super$B$ drift chamber was designed to have square or rectangular cells to mitigate this, but the trade-off is that the time-to-distance relation has more severe dependence on the track angle. Another example is the material chosen for the field wires. In the BaBar drift chamber, gold-coated aluminium wires were used because it was assumed that the surface of bare aluminium wires would be too rough and would cause premature ageing of the chamber. Belle used bare aluminium wires and observed no such effect, and thus had a reduced amount of material in the drift chamber volume, which is highly desirable. Thus the Super$B$ drift chamber was to use bare aluminium wires too.

This section will present those aspects of the Super$B$ drift chamber which influenced the design of the prototypes used in our studies. Other details

about the construction, electronics, cooling, high-voltage, and structural integration into the detector can be found in the Technical Design Report and are not repeated here. The Super$B$ project was cancelled before a final design was chosen for the drift chamber, so some choices were still preliminary. Most design choices were optimized using simulation programs (either GEANT or a Super$B$-specific program called FastSim).

The Super$B$ drift chamber has three purposes: to precisely measure the momentum of charged particles (see Section 1.3.1), to identify said charged particles (see Section 1.3.2), and to provide a trigger for the whole experiment. The charged particles involved generally have momenta around or below 1 GeV/c, even though the centre-of-mass energy of the colliding beams is $\sim$ 10 GeV. This is because the very massive resonances and particles created at the interaction point are very short-lived, and the only particles that make it to the drift chamber itself are daughters possibly several-generations-removed from the original. All these daughter particles share the energy budget (along with neutral particles that are not detected by the drift chamber), bringing us into the sub-GeV range. This is also why charged particle identification is very important, because the original particles are only observed through reconstruction by adding up the momenta of the daughter particles which are actually observed in the detector. If a single particle in the drift chamber is misidentified, it can completely spoil the inferred measurements of the parent particle.

Because the majority of particles to be detected have momenta around 1 GeV/c, the dominant contribution to the uncertainty of measurements made in the drift chamber is from multiple-scattering. Multiple-scattering is when the charged particle crossing the drift chamber interacts elastically with the material in the chamber. The material may be from the outer shell, end-plates, wires, or gas (see Figure 2.4 showing the different components). The elastic interaction means that no new particles are created, but the charged particle exchanges a bit of momentum with the material that it hits. This causes a slight kink in the track, or a slight change in energy. Multiple elastic scattering thus makes the tracks jagged and not conforming to our helical expectations, and spoiling the momentum measurement. Most

of the design choices for the Super$B$ drift chamber were made with the aim of reducing the amount of material in the chamber in order to minimize multiple scattering. Reducing multiple scattering also helps those detector components outside the drift chamber, because the momentum of particles is less affected by crossing the drift chamber. Another benefit is that the whole structure actually has a reduced weight, easing structural constraints.

The total length is 3092 mm, but with space allotted for the electronics and endplates, the axial wires would be 2557 mm long. The inner and outer radii of the chamber are 270 mm and 809 mm, respectively. These parameters are largely defined by the positions and sizes of the other detector components such as the vertex tracker and the calorimeter. The inner wall is made of carbon fibre, and the design called for it to be as thin as possible. Convex endplates were determined to allow slightly longer wires compared to concave, which allows the drift chamber to detect particles with more extreme angles.

As mentioned earlier, BaBar used hexagonal cells. This has the advantage of requiring fewer field wires per cell compared to square cells, because of the way hexagons pack together. However there are several disadvantages to the hexagonal cells. The voltages required in a hexagonal configuration mean that thicker field wires need to be used to avoid the onset of the Malter effect (described in Section 2.5). In a square-cell design, the voltages can be made lower, because there are more field wires, so the wires can be thinner. A consequence of the thicker wires in the hexagonal cell case is that a greater tension must be applied to achieve acceptable gravitational sag ($\sim 200\,\mu$m). When the tensions of all the wires are added up, the higher-tension but less-numerous wires in a hexagonal configuration actually apply more force to the endplates than the lower-tension but more-numerous wires. Thus the hexagonal cells would force a design with stronger and thicker endplates and supporting cylinders.

There are other considerations here too: hexagonal cells are closer to symmetric circular cells, so their time-to-distance relation is more uniform as a function of angle than square cells which have deeper and narrower corners. A track clipping the corner of a square cell would leave a signal

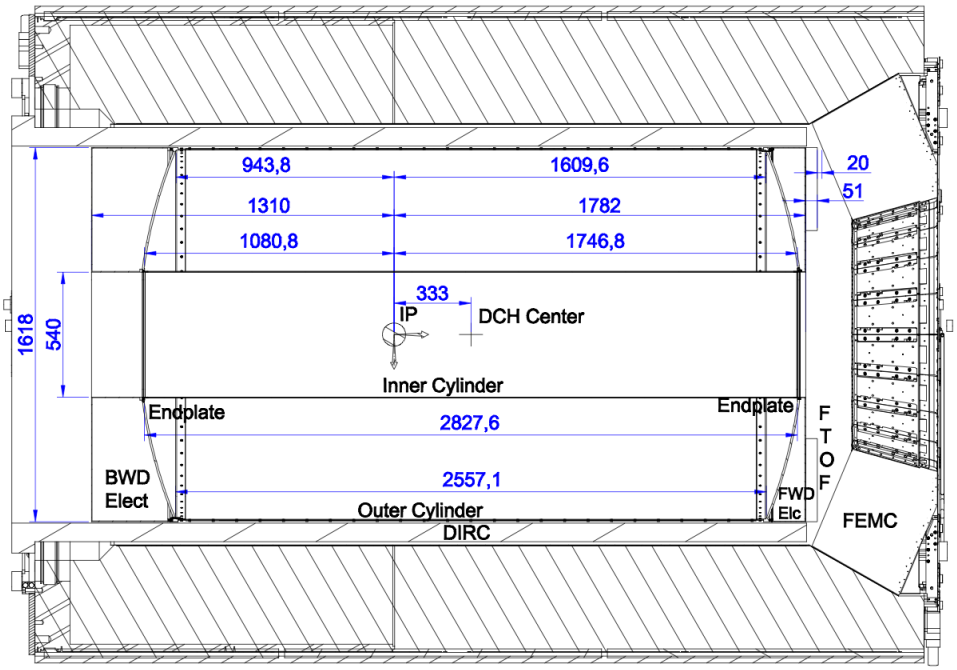


Figure 3.1: SuperB drift chamber schematic, taken from the Technical Design Report [37]. Dimensions are in millimetres. IP indicates the interaction point, while FWD Elect and BWD Elect refer to electronic component modules. DCH is drift chamber while DIRC, FTOF, and FEMC are other detector components.

with an unusually long arrival time. BaBar used stereo wire layers (layers of wires not parallel to the drift chamber axis, but rotated by an angle) to provide a measurement of the z track coordinate, but the hexagonal cell design does not allow adjacent cell layers to have different stereo angles (the wires would touch). Thus between each layer of cells that changed angle, they had to add a space with no cells and extra "guard" wires to correct the electric fields in the adjacent layers. This dead space is clearly detrimental, and the extra wires ruin one of the main advantages of a hexagonal layout: the smaller number of wires per cell. A square cell design allows adjacent layers to have different stereo angles, so although the Super$B$ drift chamber is roughly the same size as BaBar's, it is able to support 44 layers of cells

instead of just 40.

Since the cell layout is of concentric rings of approximately square cells (or square-ish, given the curvature), the cell sizes must change slightly as a function of radius in order to fit. The cell sizes range from 10.4 mm to 19.2 mm, with the smaller cells in the innermost layers. The inner cells are smaller because they are exposed to higher levels of background particles, being closer to the beam. Smaller cells in the same space means more cells, resulting in a reduced background intensity per cell.

The gas chosen for Super$B$ is a 90 : 10 ratio of helium and isobutane. Helium is chosen as the primary component because it is the lightest noble gas, and thus minimizes the amount of material in the chamber. Helium has a relatively low electron drift velocity compared to other drift chamber gasses like argon. With the wire voltages used at Super$B$, the electron velocities would be "unsaturated". Saturated electron velocities result when the acceleration from the electric field is balanced by the collisions with gas molecules, so the electron drift velocity does not vary much with changing electric fields and does not fluctuate much with gas temperature or pressure. A mostly-constant drift velocity would make the time-to-distance relations and calibrations much simpler. Nevertheless, the reduction of multiple-scattering from choosing helium supersedes, and the unsaturated gas problem can be dealt with with careful calibration and by controlling the gas temperature and pressure. A positive consequence of the slower electron drift velocity in helium is that the charge clusters hitting the sense wires will be more spread out in time. This makes the job of identifying individual clusters easier, as they are less likely to overlap.

Isobutane is the quencher gas that is responsible for re-absorbing ultraviolet photons released when the drifting electrons ionize the gas, or from backgrounds related to the beam. Without a suitable fraction of quencher, the photons could propagate and ionize more gas and the whole chamber would become a Geiger tube. The quencher is nearly always a hydrocarbon gas, and the specific choice is mostly made based on availability and convenience.

The outer structure of the Super$B$ drift chamber is made entirely of car-

bon fibre. The outer structure is responsible for supporting the huge forces (2 tonnes, or 19.6 kN) from the wires with minimal deformations, and also provides a gas-tight envelope. The amount of material should be minimized to reduce the effect on particles that cross into and out of the chamber. Using finite-element analysis programs, it was determined that a convex hemispherical shape would minimize the thickness of the endplates, and would fortuitously also maximize the lengths of the inner wires, which slightly increases the range of particle angles that can be detected. The final design, which includes considerations for the weakening of the endplates by drilling the holes for the wires, is only 8 mm thick. For comparison, a flat endplate would need to be 52 mm thick to have the same deformations.

The inner cylinder of the Super*B* drift chamber is designed to be non-load-bearing in order to minimize the material that particles must cross before entering the drift chamber. Its only role is then to provide gas tightness. The drift chamber operates at atmospheric pressure, but the pressure seal and inner cylinder are designed to accommodate up to 20 mbar of differential pressure to survive rapid weather changes. The 270 mm radius cylinder is composed of a 3 mm thick honeycomb structure sandwiched between two 90 μm sheets, all of carbon fibre. The carbon fibre sheets are also each covered by a 25 μm aluminium foil for shielding against stray electromagnetic fields.

In addition to providing gas tightness, the outer cylinder must support the entire force of the wires against the endplates. In order to allow for stringing of the wires, the cylinder is made of two half-shells, and these will only be installed after stringing. During stringing, a special external support frame is used to keep the endplates fixed. See Figure 2.4 for a photo of the BaBar drift chamber during construction, showing the same technique that would be used for Super*B*. The outer cylinder is structured the same way as the inner one, but with a 6 mm thick honeycomb structure and 1 mm thick sheets.

The final Super*B* design called for gold-plated molybdenum sense wires. Molybdenum has roughly the same mechanical strength as the more typical tungsten-rhenium, but has less resistivity and is slightly less dense (see

Section 2.3 for more details). The result is a better-quality signal recorded by the electronics, and 1.6 kN less force on the endplates. The properties of molybdenum wires however were not fully explored, so the original concept for the Italian prototype (see Chapter 5) was to test the use of molybdenum wires, and more work was needed to test the creep and breaking strengths. The diameter of the sense wires was chosen to be as small as possible while still allowing them to be handled for construction. 20 μm is the thinnest practical wire for this purpose. Thinner sense wires allow lower voltages to be used, which also allow thinner field wires that do not avalanche, and thus less tension and less weight on the endplates. Using molybdenum sense wires increases the radiation length of the whole gas and wire system to 545 m from 480 m when using tungsten wires.

The field wires are chosen to be 90 μm thick bare aluminium wires. This diameter is the smallest that keeps the surface field below 20 kV/cm. This is considered a safe value to avoid the Malter effect from manifesting itself (Section 2.5).

Chapter 4

Particle Identification Study

This chapter is a literal embedding of a paper published in 2012 [1]. The references cited can be found in the bibliography at the end of this thesis. The last few sections appeared as appendices in the original paper, so some of the references have been modified to suit, along with a few other formatting changes and typographical corrections. Only a brief overview of the work will be described here, as sufficient detail is provided in the paper.

Two prototype drift chambers were built at TRIUMF. These are single-cell chambers, 2.7 m in length. The cell layout can be seen in Figure 4.2. In a beam of π^+, e^+, and μ^+ particles at ~ 210 MeV the prototypes were tested with different configurations of: wire diameters, amplifiers, connector cables, and wire voltages. An external time-of-flight (TOF) system was used to provide unambiguous particle identification. To study particle identification, we constructed "tracks" of 40 cells being crossed by a particle by composing 40 single-cell events that were identified to be the same kind of particle by the external TOF system.

The composite tracks were analyzed using a traditional truncated-mean measurement of the integrated charge. We also do particle identification using cluster counting, and various algorithms are tested. The algorithm parameters are optimized based on the performance on real data.

The main result is that cluster counting is found to improve the pion identification efficiency from 50% to 60% when requiring 90% muon rejection

efficiency, compared to the truncated-mean technique alone. It was found that optimal smoothing times are $\sim 5\,\mathrm{ns}$, so that amplifiers and digitizers of only hundreds of MHz bandwidth would be sufficient to implement cluster counting (previously it was thought that at least 1 GHz would be required). Secondary results are that all the algorithms tested are equally effective when their parameters (thresholds, smoothing times) are properly optimized, and that attempting to use the cluster times themselves does not improve PID performance. Unfortunately the analysis comparing amplifiers, wire choice, and cables was not able to provide good conclusions.

4.1 Introduction

This paper describes the development and testing of a prototype drift chamber whose purpose is to evaluate the feasibility of a "cluster-counting" technique [62] for implementation in a high luminosity e^+e^- experiment. Cluster counting is expected to improve particle identification (PID) by reducing the effect of fluctuations in drift chamber signals. These are due to gas amplification and the fluctuation in the number of primary electrons per ionization site. There may also be improvements in tracking resolution, but this is left for a later study. The requirement of fast electronics and larger data sizes may make the technique impractical in terms of capital costs, available space near the detector, and computing power. To date the technique has not been deployed in an operating experiment. This work demonstrates that a cluster-counting drift chamber is a feasible option for an experiment such as Super$B$ [55, 57]. Super$B$ was cancelled after the experiments described in this paper, but the results are applicable to any drift chamber that is used for particle identification. The design of our prototype chambers was strongly influenced by the demands of Super$B$, which are described in the Technical Design Report [37].

4.1.1 Drift Chambers

Drift chambers are general-purpose detectors that can track and identify charged particles [12, 63]. They consist of a large volume of gas with in-

strumented wires held at different voltages. When charged particles move through the chamber they ionize the gas particles. The electrons from these primary ionizations drift towards the wires held at high positive voltage, while the ions drift towards the grounded wires. The sense wires are very thin ($\sim 20\,\mu$m), such that the strong electric field accelerates the electrons enough to cause further ionization near the sense wire. The new electrons ionize further into an avalanche, which is registered as an electronic signal on the sense wire. The amplification of the low-integer number of primary ionization electrons into a detectable signal on the wire is called the gas gain.

The energy loss of a heavy ($m \gtrsim 1\,\mathrm{MeV/c^2}$) charged particle from primary ionizations depends on its speed, as given by the Bethe formula [15] and various corrections [16]. The speed measurement is combined with the independent momentum measurement from tracking, giving the particle's mass, which is a unique identifier. To measure speed, we measure or estimate a quantity proportional to the number of primary ionizations. A traditional drift chamber accomplishes this by measuring the total ionization per unit length of the track, which is proportional to the integral of the electronic signal on the sense wires belonging to a track. The theoretical probability distribution function for the total ionization is a Landau distribution, which has an infinite mean and standard deviation [12]. The consequence is that if one takes the average of a number of samples (e.g. 40 measurements of deposited charge in a track), the resulting distribution is non-Gaussian and is dependent on the number of samples taken. Instead of the mean of the distribution, one can use the most probable value for the total ionization. This is accessed by a truncated mean technique. Our truncated mean procedure is described in Section 4.5.2.

4.1.2 Cluster Counting

The conventional technique described above is sensitive to gas gain fluctuations as well as the statistical fluctuations in the number of primary electrons produced in each ionization event. Moreover, the truncated mean procedure

that is typically used discards a substantial fraction of the available information. None of these disadvantages exist if the number of primary ionizations can be measured more directly.

Technique

The cluster-counting technique involves resolving the cluster of avalanching electrons from each primary ionization event. This is done by digitizing the signal from the sense wire in each cell and applying a suitable algorithm. The rise time of the signal from a cluster is approximately 2 ns, so electronics with sufficiently high bandwidth are required.

In principle, clusters can be detected as long as they do not overlap completely in time. This can happen irrespective of the electronics involved due to the probabilistic nature of the ionization process. Overlapping clusters are more likely for highly oblique tracks. Complex algorithms which consider signal pulse heights might disentangle even overlapping clusters, but the algorithms tested in this work do not.

An optimal algorithm would have a high efficiency for identifying true clusters and a low rate of reporting false clusters (due to noise for example).

PID

In traditional drift chambers using the integrated signal, the signal amplitude is determined by the convolution of the probability of primary ionization, the number of primary electrons produced, and the variations in gas gain. This results in a long-tailed distribution that is typically dealt with by the truncated mean procedure. Conversely, if clusters are perfectly identified, then the only variation is from the primary ionization, which is a Poisson process. No cluster counts need to be discarded to allow for a proper statistical treatment. In reality some counted clusters will be missing or fake, the rate of these being caused by gas gain fluctuations, noise level, and the time separation capabilities of the electronics. The idea is that the sensitivity to these effects is small. The difficulty arises from the need to optimize a cluster-counting algorithm that may have many parameters.

A difficulty with both charge integration and cluster counting is the presence of δ-rays [12]. These are electrons produced in primary ionizations that travel far in the gas before further ionizing, such that they create their own separate ionization cluster. The production of δ-rays at a given momentum depends only on the particle speed ($\propto 1/\beta^2$) [16]. This inflates the charge integral and the cluster count with only a weak dependence on the species of the original particle, the result is a decreased PID resolution in general. The presence of δ-rays is one of the reasons why a truncated mean is used in the charge integration method. While cluster counting is also affected by δ-rays, the effect is less pronounced, allowing all of the data to be used.

Cluster Timing

Any cluster-counting algorithm that uses a digitized signal is able to report not only the number of clusters in a cell, but also the arrival time of each of those clusters. In the oversimplified case of a linear and homogenous drift velocity and infinite cells, the average spacing in time between consecutive clusters would simply be proportional to the inverse of the number of clusters in the cell. In a more realistic scenario, the average spacing between clusters is useful information that is not one-to-one with the number of clusters. We can exploit the lack of perfect correlation and use the cluster timing information to further improve our ability to identify particles.

Tracking

For tracking, cluster counting may also improve performance, but in a much lesser degree and more subtle manner than as for PID. A traditional drift chamber uses only the arrival time of the overall signal in determining the distance of closest approach from a sense wire. Unfortunately this arrival time measurement is vulnerable to noise, gas gain fluctuations (small initial clusters may be missed), etc. If the first few clusters are resolved, then while the first cluster arrival time is still the primary datum, the second cluster arrival time can be used as a consistency check. If the second cluster arrives much too late, then the chance that the first cluster was a fake is greater, so

a smaller statistical weight can be assigned to that cell when reconstructing the whole track. This paper deals only with the PID improvements and does not address tracking.

4.2 Apparatus

In this section we describe the prototype drift chambers that were built, the custom signal amplifiers and the various types of cables that were tested. We also describe the experimental setup in the test beam, the data acquisition system, and the devices used for external PID and triggering.

4.2.1 Prototype Drift Chambers

We built two nearly identical full-length (2.7 m) single-cell drift chambers, called chamber A and chamber B (Figure 4.1). The only difference between the two chambers is the diameter of the sense wires: 20 μm for Chamber A and 25 or 30 μm for Chamber B. More details about the wires are given below.

The wire layout creates a square cell 15 mm wide in a 10×10 cm cross-section casing (for a gas volume of $2.7 \times 10^4\,\mathrm{cm}^3$). Figure 4.2 shows a cell diagram including the dimensions and wire locations. The aluminium casing of the chambers has five large windows on two sides of the cell to allow particles to enter and exit unimpeded. The windows are made of thin ($\sim$ 20 μm) aluminium, protected by aluminized Mylar.

Different amplifiers are mounted on the endplates of the drift chambers, connected directly to the sense wires. The amplifiers vary in their gain, input impedance, and bandwidth. They are described in more detail in Section 4.2.2.

We had the option of including a termination resistor to ground on the non-instrumented side of the chamber. The required termination resistance to prevent reflection of signals is 390 Ω. Runs were taken with and without termination, to see the effect of reflected signals on PID performance. A circuit diagram showing our termination is in Figure 4.3.

Runs were taken with chambers A and B strung with 20 μm and 25 μm

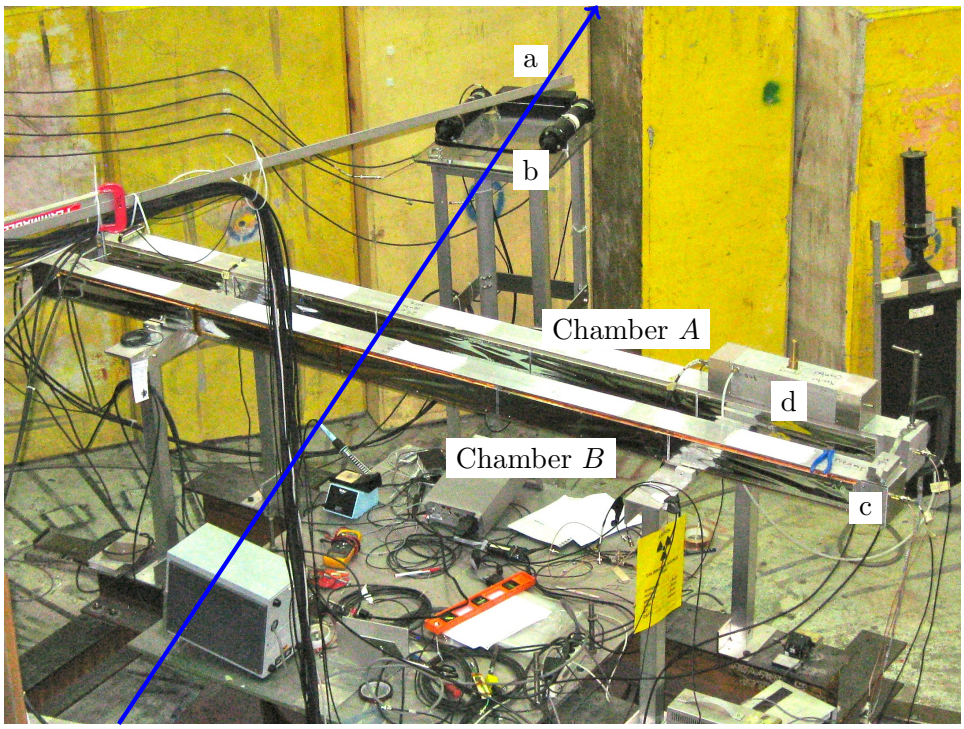


Figure 4.1: Photo of the prototype chambers mounted during our beam test. The far scintillator (Section 4.2.6) and additional PMTs (Section 4.2.7) are visible in the background (labelled a and b respectively). The amplifier shielding boxes (c) are on the right side of the picture. The smaller monitoring chamber (labelled d) (Section 4.2.8) is on top of Chamber B. The blue arrow shows the path of the particle beam through our prototypes and two of the scintillators.

gold-plated tungsten sense wires, respectively, and gold-plated aluminium field wires. For some later runs, chamber B was re-strung with a $30\,\mu$m sense wire. The wires are connected to the endplates by the same crimp-pins and feedthroughs that were used in the BaBar drift chamber [38].

The gas chosen for the test was a mixture of helium and isobutane in a $90:10$ volume ratio. Helium was chosen because it reduces the effect of multiple scattering compared to the more typical argon [23]. Multiple scattering of the charged particles is the dominant contribution to the tracking

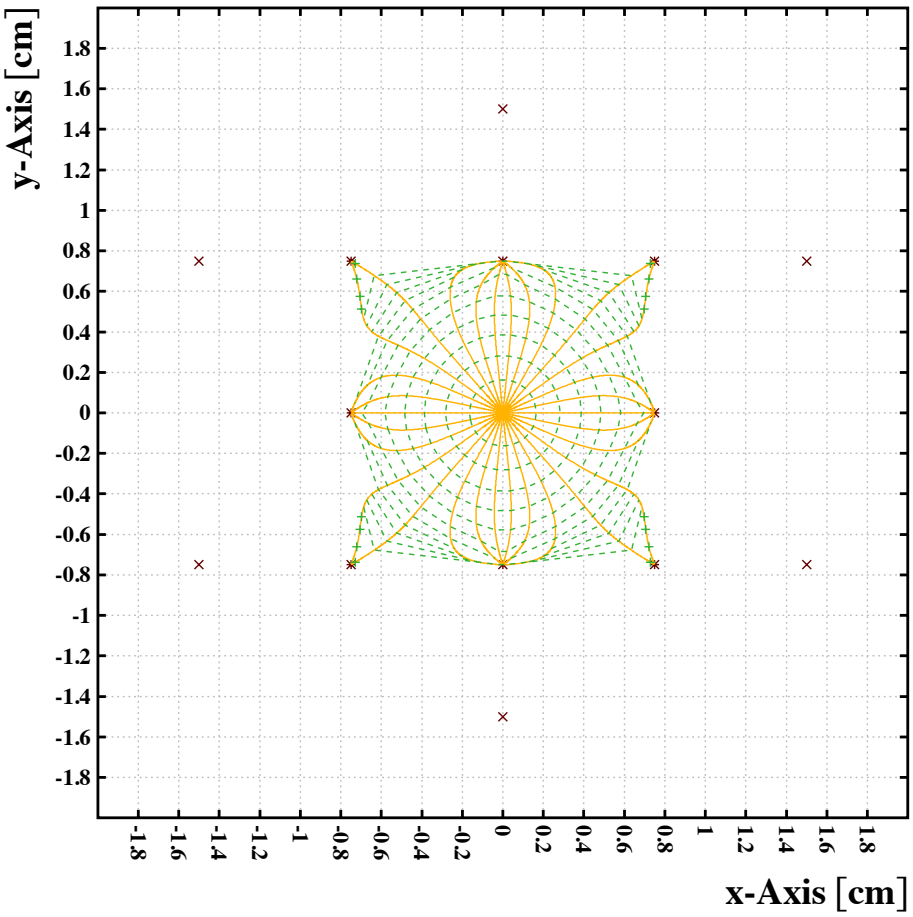


Figure 4.2: Garfield [4] simulation of isochrones for electron drift times in our prototypes, with 90 : 10 helium and isobutane. The isochrone intervals (dashed lines) are 50 ns. The full orange lines are the drift paths. The central point is the sense wire at high voltage, while the 8 points in a square around it are the field wires at ground. The extra 6 points outside the cell are bias wires to simulate the presence of an infinite network of cells. The wire voltages are 1820 V and 1054 V for the sense wire and bias wires respectively.

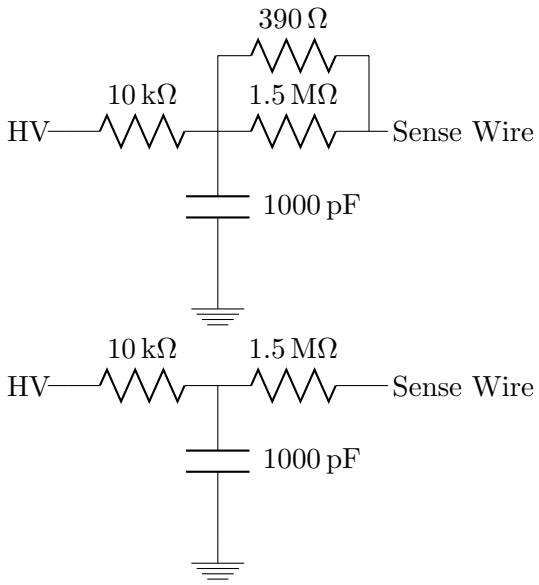


Figure 4.3: Circuit diagram representing the high voltage connection to the sense wire with (top) and without (bottom) termination resistor. Without the 390 Ω resistor, the signal can bounce.

resolution at a B-factory like Super*B*. In consideration of the rest of the Super*B* detector, using helium reduces the number of radiation lengths represented by the drift chamber. With isobutane as the quench gas, we are able to operate the chamber with a large helium fraction, further reducing the amount of material. Helium also exhibits a lower drift velocity and ionization density, which also makes it an ideal choice for cluster counting as the incoming clusters will be less likely to overlap in the digitized signal.

The chambers are operated at room temperature and atmospheric pressure. We measured the temperature and pressure during the data taking periods, we did not use these at any level of the analysis.

4.2.2 Amplifiers

We used custom made amplifiers in order to achieve the bandwidth required for cluster counting. The amplifiers are based on the AD8354 RF gain block from Analog Devices. These have a reasonably low power consumption

($\sim 140\,\mathrm{mW}$ for the whole unit) and a bandwidth of 2.7 GHz. These devices have $50\,\Omega$ input and output impedance, and a fixed gain of 20 dB. The simplest configuration that we investigated was with two AD8354s in cascade. This provides very good bandwidth performance, but the input impedance of $50\,\Omega$ creates a large mismatch with the characteristic impedance of the drift chamber cells (around $370\,\Omega$) and the signal to noise ratio is not optimal. So, an emitter follower stage was added at the input, using a low noise RF transistor (BFG425). This was configured either with $370\,\Omega$ input impedance, or with $180\,\Omega$, as a compromise between impedance matching and tolerance to stray capacitance. We also tried a configuration with an additional low gain ($2\times$) inverting stage (with a BFG425 transistor), having $370\,\Omega$ input impedance. In this case, a single AD8354 gain block was used. The $370\,\Omega$ configuration gave the best overall results. A schematic of the amplifier setup is shown in Figure 4.4.

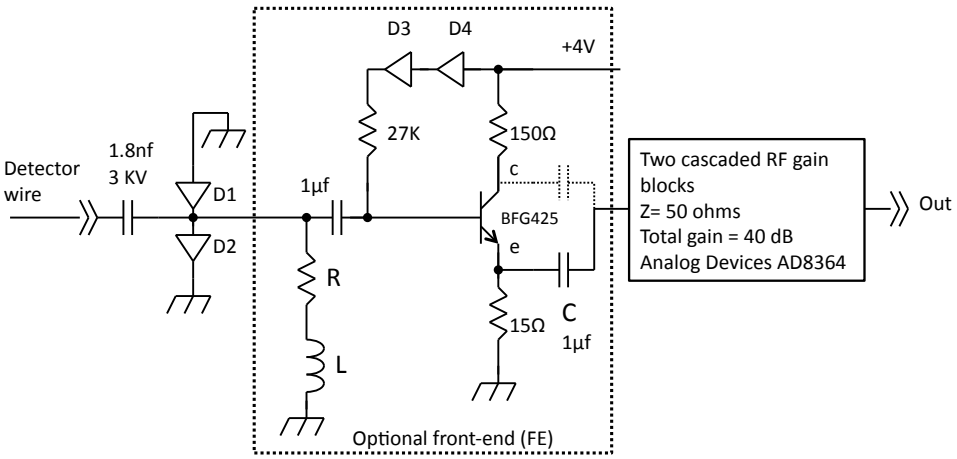
In our final analysis, only the $50\,\Omega$ and $370\,\Omega$ amplifiers are considered. The data runs using the $180\,\Omega$ amplifiers gave signals which were of low enough quality that a full analysis was not possible.

4.2.3 Wire Voltages

The correct voltage settings for the guard wires in the cell were determined using the computer program Garfield [4]. The guard wire voltages are chosen to make the sensitive region of our cell behave as if it were part of an infinite array of identical cells. These voltages scale linearly with the chosen sense wire voltage.

The sense wire voltages were tuned to obtain roughly equal-amplitude pulses for all combinations of chamber and amplifier. This was done empirically by looking at the fraction of events on the oscilloscope (Section 4.2.8) that saturated the full voltage range. The voltage was tuned until this fraction was $\sim 15\,\%$.

The resulting voltage for chamber A ($20\,\mu$m sense wire) using one of the $50\,\Omega$ amplifiers is 1700 V. The corresponding electric field at the wire surface is calculated by Garfield to be $217\,\mathrm{kV/cm}$.



Preamplifier type	FE	R	L	C position
50 Ω direct	not used	----	----	----
180Ω, non inverting	used	180 Ω	15nHy	e
370Ω, non inverting	used	370 Ω	68nHy	e
370Ω, inverting	used	370 Ω	68nHy	c

Figure 4.4: Simplified schematic of the amplifiers used in the experiment.

4.2.4 Cabling

For some of the runs we varied the type of signal cable used to connect the output of the amplifiers to the data acquisition system. We used two different types of Sub-Miniature RG-59/U cables (models 1855A and 179DT from Belden) and Miniature Coax (model 1282 from Belden), all with 75 Ω impedance. The lengths were all 10 m, which is the distance between the amplifiers and digitizers for Super*B*. From the signal-propagation perspective, the 1855A is a better cable than the 179DT, having less signal attenuation (34 db/100m versus 70 dB/100m at 1 GHz). From the perspective of mechanical integration with the rest of the detector however, the 179DT cable would be preferable to the 1855A, being lighter, thinner, and allowing a smaller minimum bend radius (25.4 mm versus 38.1 mm).

We also took data with a header connector between the amplifier and the signal cable to simulate a connector through the real drift chamber bulkhead. The header connector has 20 pins with a 2.54 mm pin spacing. Only two

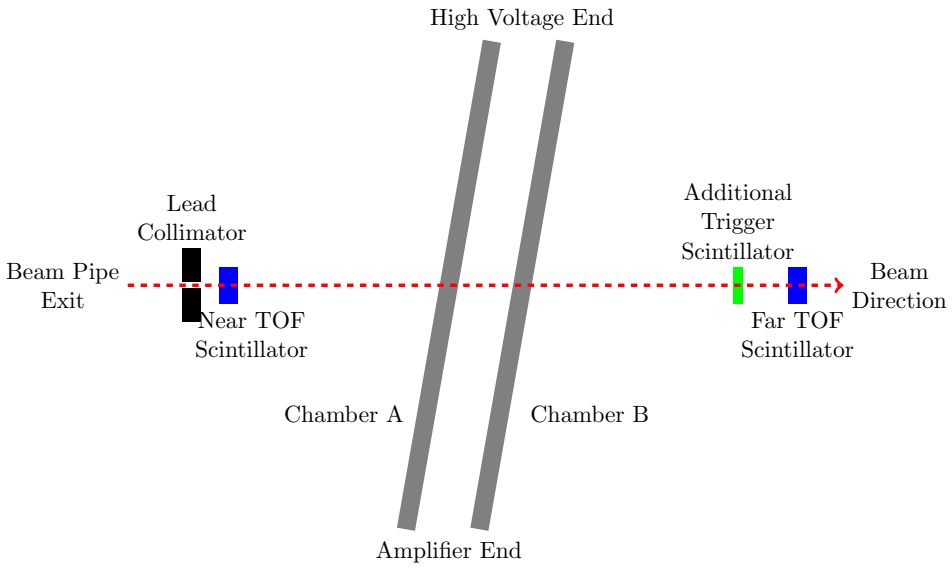


Figure 4.5: Schematic of beam test setup at the TRIUMF M11 facility. The distances in this schematic are not to scale, though the drift chamber proportions are correct.

pins are used in the connector to connect the ground and signal parts of an additional 30 cm 1855A cable which is inserted in our signal cable length using regular BNC connectors.

4.2.5 Test Beam

Data were collected at the TRIUMF M11 beam [64], which simultaneously delivers positrons, positive muons, positive pions at a tuneable momentum range of 100 to 350 MeV/c. We block residual protons from upstream using a slab of polypropylene at the mouth of the beam pipe (6.35 mm thick at 210 MeV/c). We can determine the beam populations using the time-of-flight system described in Section 4.2.6.

The prototypes were mounted on a rotating and moveable table, which allowed us to take runs at different dip angles and positions along the length of the sense wires. A schematic of the beam test setup is in Figure 4.5 and a photo of the test hall is in Figure 4.1.

Most of the data were collected at 210 MeV/c, a relatively low momentum for a high-energy particle physics experiment. At this momentum however the Bethe formula separation of pions and muons is similar to the separation of pions and kaons at 2 GeV/c. This is confirmed by our simulations at both momenta, described in Section 4.3. High-efficiency separation of pions and kaons at 2 GeV/c is critical for high-precision measurements and reconstructions at a high-energy particle experiment like Super*B*.

4.2.6 Time of Flight

An external time of flight (TOF) system was used to identify the particles independently of the prototypes. The beam's momentum spread is small enough that a histogram of the TOF shows distinct peaks corresponding to the species of the particles in the beam. The TOF system consists of two counters ~ 4 m apart, one upstream of the prototypes and one downstream (Figure 4.5). The counters are $12.7 \times 12.7 \times 220$ mm BC-404 scintillators each read out by two Burle 8501-1 64-channel micro-channel plates (MCPs), one on each end of the scintillator block. The scintillators are roughly the same size as the beam spot. The MCPs have 25 μm pores. Each of the 64 channels in the MCPs have an active region of 6×6 mm. We gang together four of the channels to form one combined signal. This signal from each MCP goes to an Ortec 935 constant-fraction discriminator (CFD) with no pulse height correction applied. Each is then delayed by a given time in order to separate the pulses, then they are combined and recorded in a single channel of our oscilloscope.

The signals from the MCPs are used as part of the trigger. Additional signals from photomultiplier tubes are used and are described in Section 4.2.7.

The unscaled TOF is obtained by determining the arrival time of each pulse from the MCPs as recorded by the oscilloscope. The first two pulses are from the two ends of the upstream counter, while the following two are from the downstream counter. These pairs are averaged, then the difference is taken. There are arbitrary delays associated with the MCP signals, so

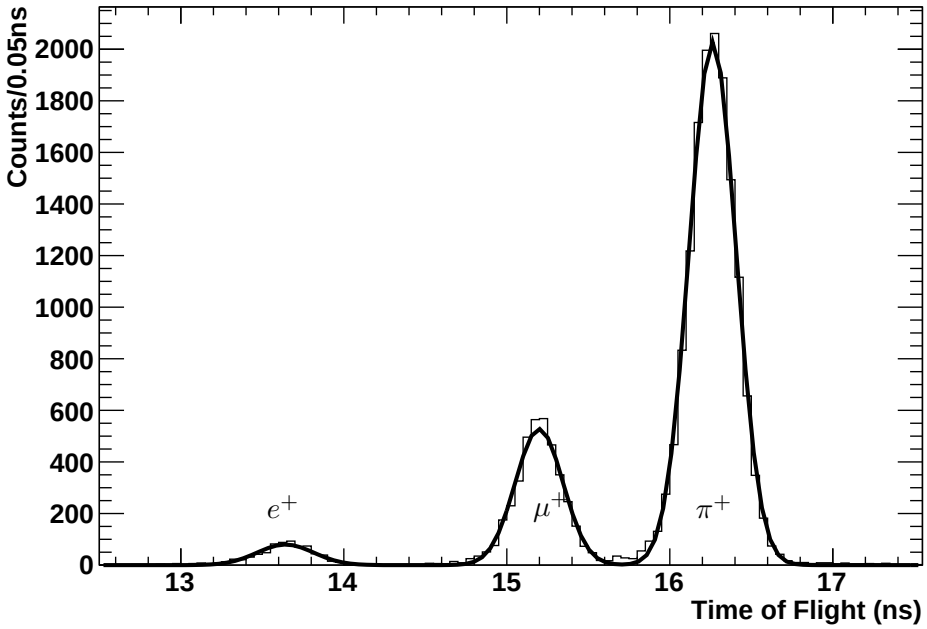


Figure 4.6: Time-of-flight histogram for a run at 210 MeV/c beam momentum. The three peaks correspond to positrons, muons, and pions, in increasing TOF order. The fit is to the sum of three Gaussian distributions.

the TOF quantity is scaled to be physically meaningful. A run is chosen and a histogram of the TOF quantity is made, where the positrons, muons, and pions are clearly resolved as Gaussian peaks. We fit the positron peak with a Gaussian distribution. The beam momentum is high enough that the positrons may be treated as moving at the speed of light, and the actual distance between the two counters is well-measured.

We were able to achieve TOF resolutions of ~ 60 ps per MCP. For a 210 MeV/c beam (Figure 4.6) the separations of the Gaussian peaks are greater than the $3\,\sigma$ ranges used to identify particles in our track composition process described in Section 4.5.2. A sample trace of the actual TOF signal is shown in Figure 4.7, where the first four pulses are from the MCPs.

We fit the TOF distribution with the sum of three Gaussians and count how many particles are within $3\,\sigma$ of each peak. For the run shown in

Figure 4.6, we find that of all the physical triggers 3.8% are positrons, 20.5% are muons, and 75.7% are pions.

4.2.7 Trigger

The TOF signals are also used as part of the trigger system for the oscilloscope. It was noted that with only the upstream and downstream counters, many events contained no signals in the drift chambers (i.e. oscilloscope traces with just normal noise, no clusters). In addition, the TOF histogram showed six peaks, though only three were expected. The six peaks appeared to be in two similar groups, shifted in TOF value. The conjectured origin of the higher-TOF valued population was beam particles passing through the upstream counter but angled downwards, scattering off of the metal table, then passing through the downstream counter, bypassing the chambers entirely and taking a longer path.

A third scintillator strip 3 mm thick was placed between the prototypes and the downstream counter (Figure 4.5), instrumented with photomultiplier tubes. The coincidence of the three (upstream, downstream, strip) was required for a physical trigger. This additional requirement removed the extraneous TOF population and many of the events with no drift chamber signals. Part of the trigger signal can be seen in Figure 4.7 in the upper trace. The third scintillator was not digitized and thus is not visible in the figure.

The coincidence rate is ~ 30 Hz, while the signal rate on the upstream counter without requiring coincidences ranges from a few kHz to tens of kHz, depending on beam line settings. We also introduced an asynchronous trigger based on a pulse generator whose frequency was tuned to $\sim 15\%$ of the total trigger rate. These asynchronous triggers are uncorrelated with real beam events. They provide a sample of empty events for monitoring and measuring baseline voltages and noise levels during the run.

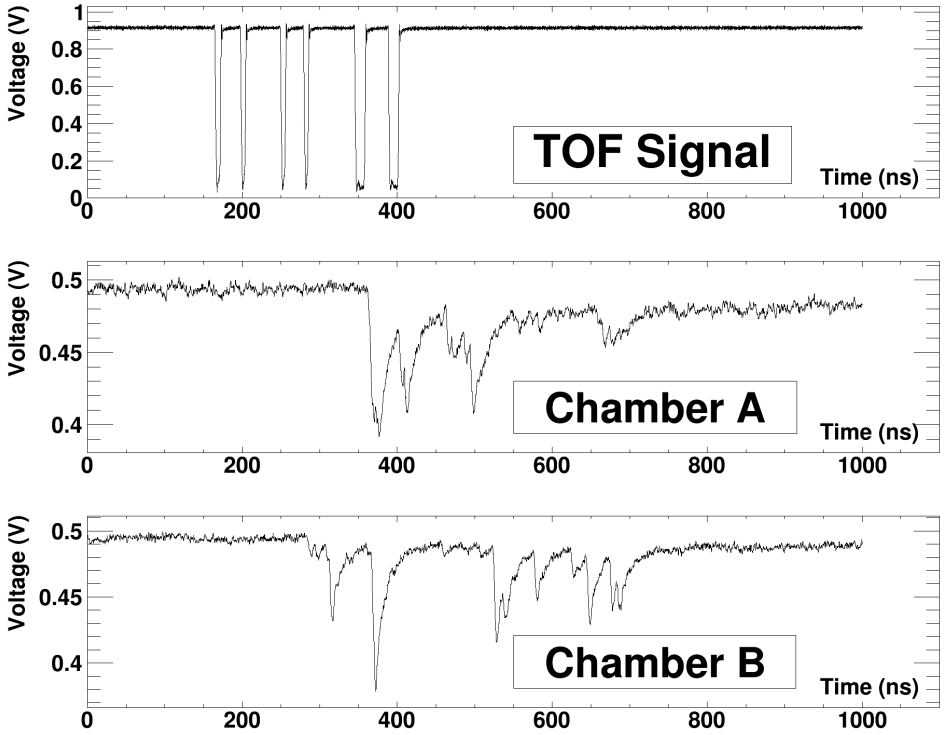


Figure 4.7: Oscilloscope traces for a run at 210 MeV/c beam momentum. The first is the TOF signal, with four initial pulses from the TOF MCPs, and two additional pulses from the extra trigger PMT. The TOF value identifies this particle as a pion. The second and third traces are from prototypes A and B, respectively. The cluster structure is clearly evident in these signals.

4.2.8 Data Acquisition System

Our data acquisition system consisted of a LeCroy WavePro 740Zi, an oscilloscope with 4GHz bandwidth. Data were written to an external USB hard disk in a proprietary binary format and then converted into ROOT [3] files for analysis. The oscilloscope writes one file per active channel per trigger. We used one channel for the time-of-flight system and one channel for each prototype sense wire, meaning we had three small files written per trigger. Each channel read 20002 samples with 50 ps spacing, for a trace duration of

$\sim 1\,\mu$s. The biggest bottleneck was the filesystem (Microsoft NTFS), which does not perform well with directories having tens of thousands of files. The overall rate of events written to disk was $\sim 12\,$Hz.

We used the MIDAS [65] data acquisition system to automatically record temperature and atmospheric pressure as well as the current in a small monitoring chamber. The monitoring chamber was connected in series with the primary chambers on the gas line, and was exposed to an ^{55}Fe source. The monitoring chamber wire voltages were held fixed, allowing us to monitor the gas and environmental conditions by tracking changes in the gas gain.

4.3 Simulations

We used a gaseous ionization detector simulation package called Garfield [4] to simulate tracks through our prototypes. We did not simulate the electronics chain and the data acquisition system, but we are able to get predicted charge depositions and cluster counts for our specific gas mixture and wire configuration.

The charge deposition is not reported directly, but is proportional to the energy lost by charged particles passing through the gas. It is plotted in Figure 4.8 for muons, pions, and kaons. The momentum scale is chosen to illustrate the fact that the difference in energy loss between pions and muons at $\sim 200\,$MeV/c is similar to that between pions and kaons at $\sim 2\,$GeV/c (Section 4.2.5).

The number of primary ionizations is reported directly by the simulation software and can be treated as a "true" number of clusters. It does not depend on the choice of electronics, algorithms, and it does not count δ-rays (Section 4.1.2). The distribution of primary ionizations for muons, pions and kaons is shown in Figure 4.9 and also shows the similarity between muon-pion separation at our beam momentum and pion-kaon separation at higher momenta. It is also important to point out that the absolute number of clusters for muons and pions at 210 MeV/c approximately mirrors that of pions and kaons at 2 GeV/c, not just the difference. The absolute value is important because it is related to our ability to actually resolve the clusters.

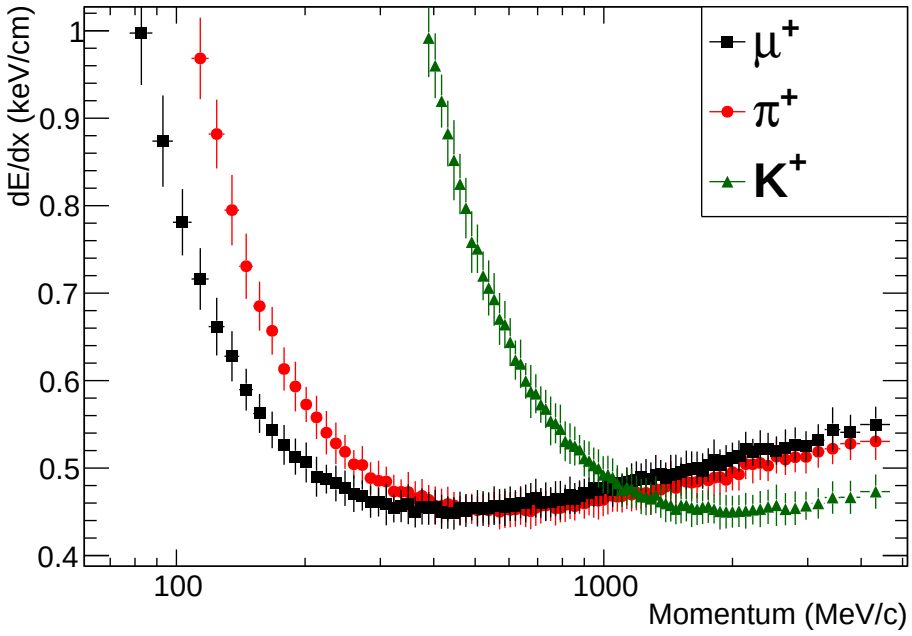


Figure 4.8: Garfield simulation of the energy loss by a charged particle crossing 40 cells of a 90:10 mixture of helium and isobutane. The black squares, red circles, and green triangles represent muons, pions, and kaons, respectively. The marker position is the 70 % truncated mean energy loss, while the vertical error bar on each marker is the RMS of the truncated mean.

4.4 Beam Test Data

The data were taken during August and September 2012. Approximately 200 runs of 30000 events were acquired. A run is a contiguous data-collection period during which no setup parameters are changed. On average, 15 % of the events were from asynchronous triggers and 10 % of the physical triggers did not leave signals in the prototypes.

Various parameters were changed from run to run. These were: the sense wire voltages, amplifiers, signal cable types, beam momentum, angle of incidence of the beam with the chamber, beam position along the sense wire length and presence of a proper termination resistor on the sense wire.

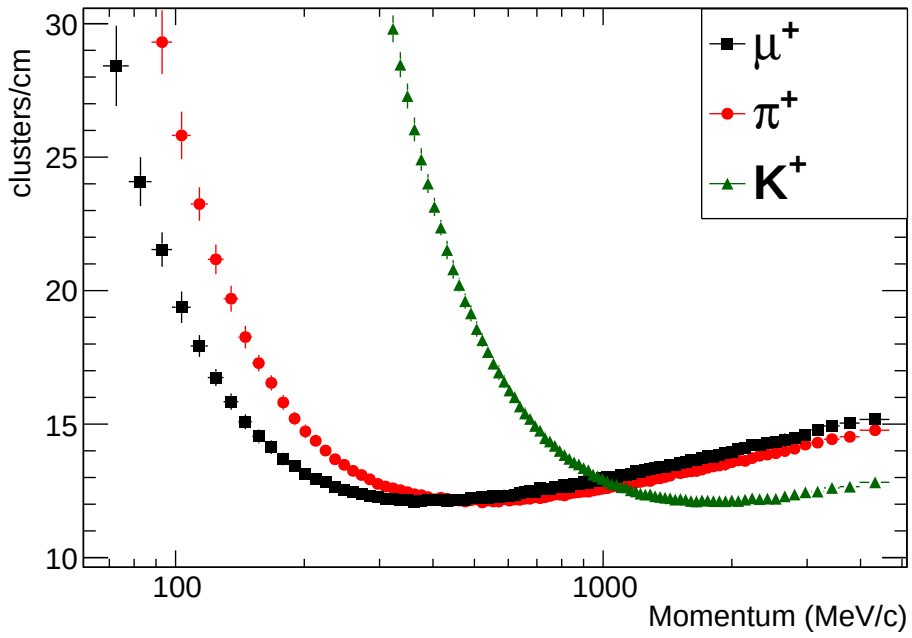


Figure 4.9: Garfield simulation of charge clusters produced by a charged particle crossing a 90:10 mixture of helium and isobutane. The black squares, red circles, and green triangles represent muons, pions, and kaons, respectively. The marker position is the average number of clusters, while the vertical error bar on each marker is the RMS.

In the end, many runs turned out to be recorded using unsuccessful amplifier prototypes and could not be used for a detailed analysis. This analysis uses 20 runs, for a total of 633050 recorded events.

4.5 Analysis

The analysis of the test-beam data is performed in two steps, both of which are done offline (after the data for that run has been fully collected). The first step involves analyzing the signals (voltage as a function of time) from the three oscilloscope channels. The first channel is connected to the time-of-

flight (TOF) system, with voltage pulses corresponding to a particle crossing the scintillators before and after the drift chambers. The second and third oscilloscope channels are connected to the amplifiers on the sense wires of the two drift chambers.

The second step of analysis involves constructing multi-cell "tracks" from the single-cell events using a composition process (described in Section 4.5.2). Single-cell events are taken from the same run, same chamber, and having a TOF consistent with the same particle type. Forty of these are used to build up a track as if it were traversing a full Super$B$-size drift chamber.

4.5.1 Single-Cell Information

This section describes in detail the first stage of analysis in which we deal with single-cell events. The time-of-flight is measured, the signal is adjusted for baseline drift and basic quality controls are imposed. In this stage we also perform the charge integration and use cluster-counting algorithms to count clusters on the drift chamber signals.

Time of Flight

The time-of-flight is determined by applying a simple threshold-over-baseline algorithm to the oscilloscope trace from the channel connected to our scintillator MCPs and PMTs. A valid TOF signal consists of four identified pulses, while an asynchronous trigger has zero pulses. Events with one, two, or three TOF pulses are rejected, and represent the small fraction of events from asynchronous triggers with a pulse in one of the TOF counters.

Baselining and Signal Confirmation

The baseline voltage for each drift chamber is simply the average voltage of the entire signal from the previous asynchronous trigger. The RMS deviation from this baseline is also measured. The mean of these RMS deviations is $\sim 2\,\mathrm{mV}$. Signals from physical triggers have amplitudes on the order of hundreds of mV above the baseline.

The real particle events are tested for the presence of an actual signal by a threshold algorithm, where the baseline and threshold levels are determined by the previous asynchronous trigger measurements. Real particle events that have no signal in the chambers are rejected. These are from events where a real particle crossed the scintillators, but either missed one or both drift chambers, or did not interact within them.

Charge Integration

A charge integration is performed for the remaining asynchronous and physical events, starting at the time of the threshold crossing mentioned in Section 4.5.1 (or at an arbitrarily chosen time for asynchronous events), integrating for a fixed duration. The distribution of start times for a sample run is shown in Figure 4.10. If the duration is too short, then some pulses may be missing or the tail of the last pulse may be clipped. If the duration is too long, then unnecessary noise is also integrated, reducing the resolving power of the charge measurement. Different equipment combinations give different pulse tail decay times, so the duration must be optimized empirically. A typical optimal value is $\sim 600\,\text{ns}$, as shown in Figure 4.11. The optimization of the integration time is described in Section 4.6.1.

From the integrated charge we subtract a pedestal calculated from the previous asynchronous trigger. This pedestal is a charge integration with the same integration time, but a fixed starting time. The result is a baseline-subtracted charge, which should have a smaller systematic error than the raw charge integral. The distribution of integrated charges for physical triggers and asynchronous triggers is shown in Figure 4.12. The physical triggers are shown separately for each species in Figure 4.13.

Cluster Counting

Cluster-counting algorithms can vary in complexity, efficiency, and in their rate of reporting fake clusters. Here we briefly describe the various algorithms, but precise definitions can be found in Section 4.8.

The algorithms involve two forms of smoothing of the oscilloscope traces

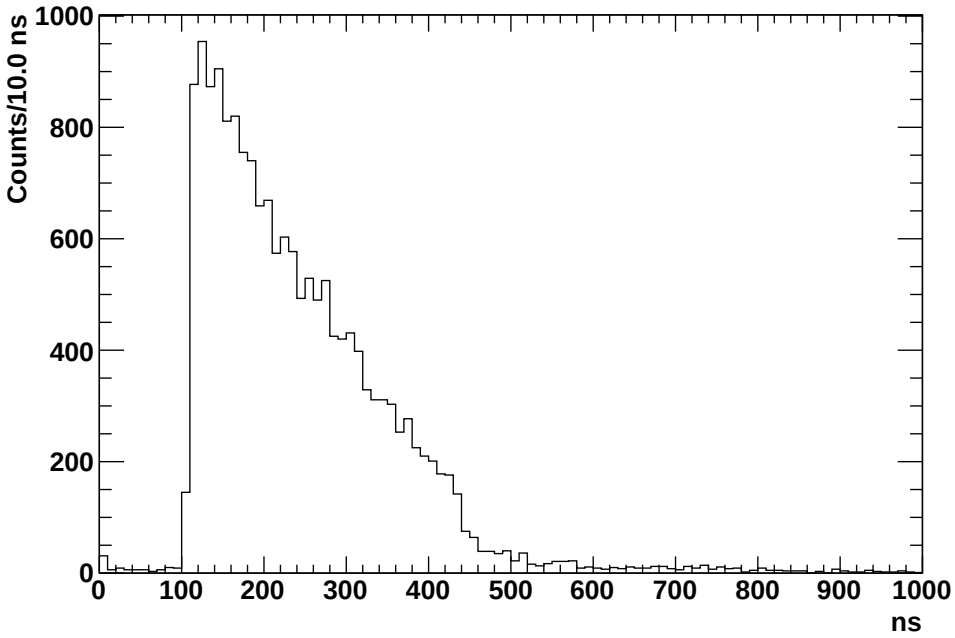


Figure 4.10: Time at which the charge integration begins in Chamber A for a run at 210 MeV/c.

(Figure 4.14). The first is a "boxcar smoothing" where each sample is replaced with the average of itself and the $n-1$ previous samples. The second is a true averaging procedure, where the number of points in a trace is reduced and each point is the average of n points.

All of the algorithms involve some kind of transformation of the smoothed signal, and a threshold-crossing criterion. The transformed signals for the various algorithms are shown in Figure 4.15. One of the most basic cluster-counting algorithms is the "Threshold above Average". It subtracts the non-smoothed signal at time t from the boxcar-smoothed signal at time $t-1$, then applies a threshold.

A more general algorithm (of which the previous is a special case) is the "Smooth and Delay" algorithm. It involves smoothing two copies of the signal by different amounts, delaying one of the copies by a certain number of frames, then taking the difference and applying a threshold. This algorithm has four parameters, and is thus more difficult to optimize.

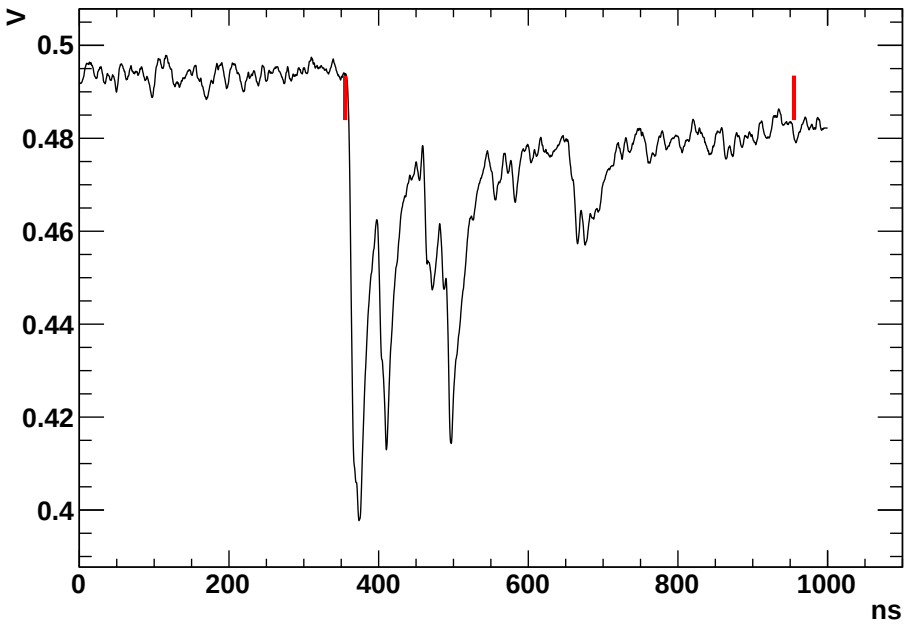


Figure 4.11: Sample event in chamber A showing 600 ns integration time.

The two algorithms above essentially implement a first-derivative method. We also implemented a second-derivative method. This one uses the true averaging procedure rather than the "boxcar smoothing". The first derivative is first calculated by taking the difference between consecutive smoothed samples. The second derivative is then calculated by taking the difference between consecutive first derivative values. Each time, we divide by the time interval represented by a sample, to keep the units consistent. The number of clusters counted using the second derivative is shown for each particle species in Figure 4.16.

All of the threshold algorithms in principle trigger on the leading edge of cluster signals. However it is noticeable that real cluster pulses have a very sharp leading edge (approximately 3 ns) and a slower decaying trailing edge (approximately 100 ns). Fake clusters are more symmetric, returning to the baseline voltage faster than the signal from a real cluster. Thus an algorithm was devised that takes cluster candidates from the above algorithms, but

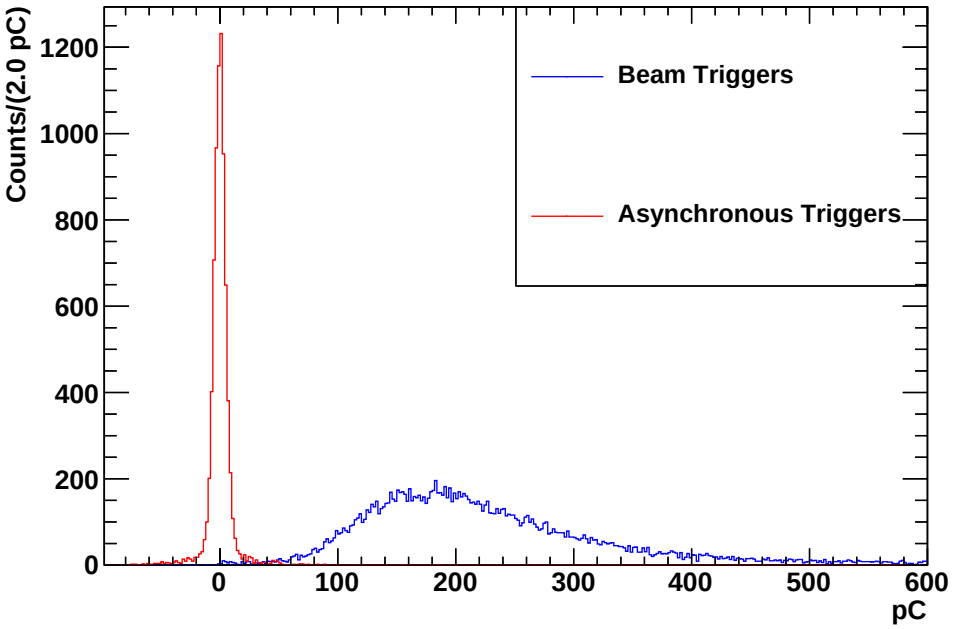


Figure 4.12: Baseline-subtracted charge distributions as identified by the time-of-flight system. The sharp peak on the left is from asynchronous triggers (with no particles in the prototypes), while the broader peak in the middle is from physical triggers with all particle species combined.

requires the pulse to last a minimum duration in order to be confirmed. Pulses that return to baseline too quickly are discarded as fake clusters. This "timeout booster" allows the use of smaller thresholds, which while increasing the efficiency of finding real clusters also admits more fakes. The timeout criterion removes most of the fakes but keeps the real clusters.

As mentioned before, each of the cluster-counting algorithms can return not only the number of clusters, but the actual time at which each cluster was found. We investigated the use of this information, in the form of an average time separation between clusters in each cell.

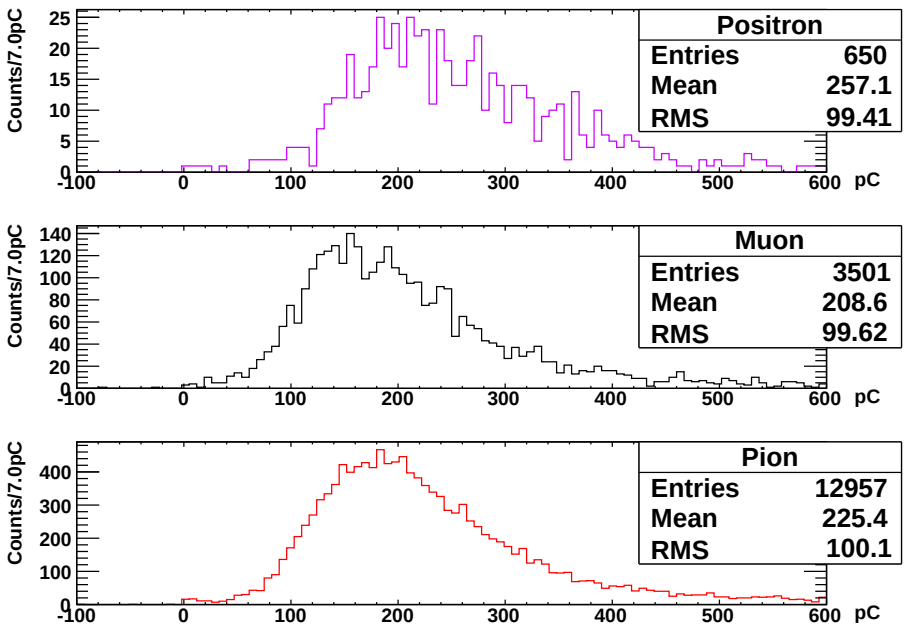


Figure 4.13: Baseline-subtracted charge distributions for each particle species at 210 MeV/c. Note that the sample mean and RMS values indicated in the figure are not representative of the underlying distribution since it does not have well-defined moments.

4.5.2 Track Composition

The prototypes have only a single cell. The traditional method of identifying particles using the truncated mean requires many cells forming a track. Thus we construct tracks from the single-cell events.

To compose a track for a given species of particle, we select (with replacement) random single-cell events that have been identified with the time-of-flight information. We positively identify particles with TOF values within 3 standard deviations of the central values of the three Gaussian peaks corresponding to the particle species. For a typical run with e.g. 3500 single muon events, the number of possible muon tracks is astronomical ($\sim 10^{94}$), and the likelihood of a given track being composed of multiple copies of the same single-cell event is low ($\sim 1\%$). We also form empty tracks by

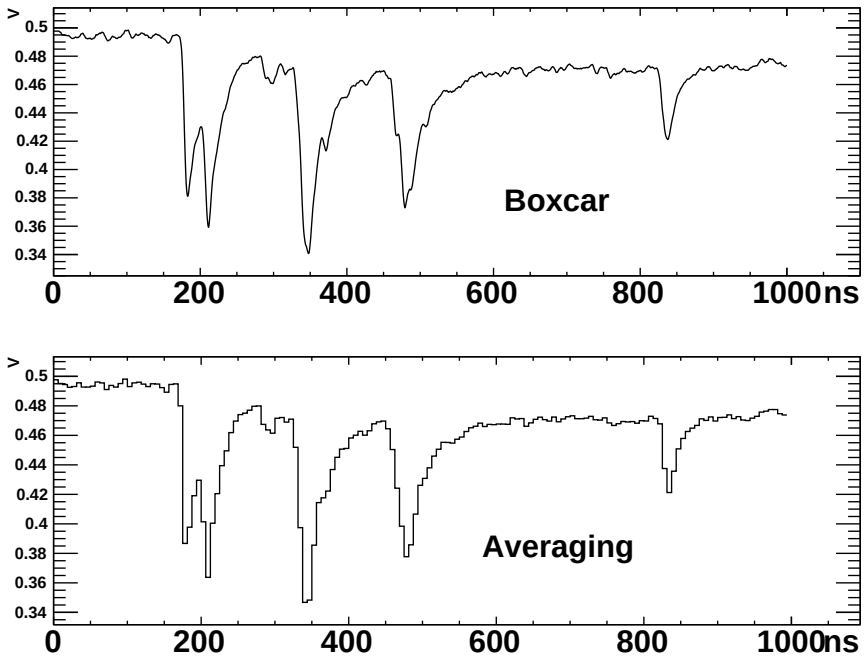


Figure 4.14: The two smoothing algorithms used, each smoothing over 125 frames of 50 ps width, for a smoothing width of 6.25 ns. This event is the same as shown in Figure 4.11

combining the signals from asynchronous events.

The information from each event is combined to form the track information. The track information is the particle species, total number of clusters found per cm of track, and the truncated mean of the charge integrals from each cell. The truncated mean is performed by sorting the list of charge integrals and taking 70 % of the values starting from the beginning of the list. The value of 70 % was roughly optimized to give better separation, for comparison 80 % was used in BaBar [38]. The Super*B* drift chamber design has 40 layers. Thus we use 40 events from our single-cell prototypes events to create a composed track. The 70 % truncated mean was thus done by rejecting the largest 12 integrated charge values from the cells.

In the case of tracks formed from asynchronous events, the list is not sorted, since these values are already Gaussian, but still the same fraction of

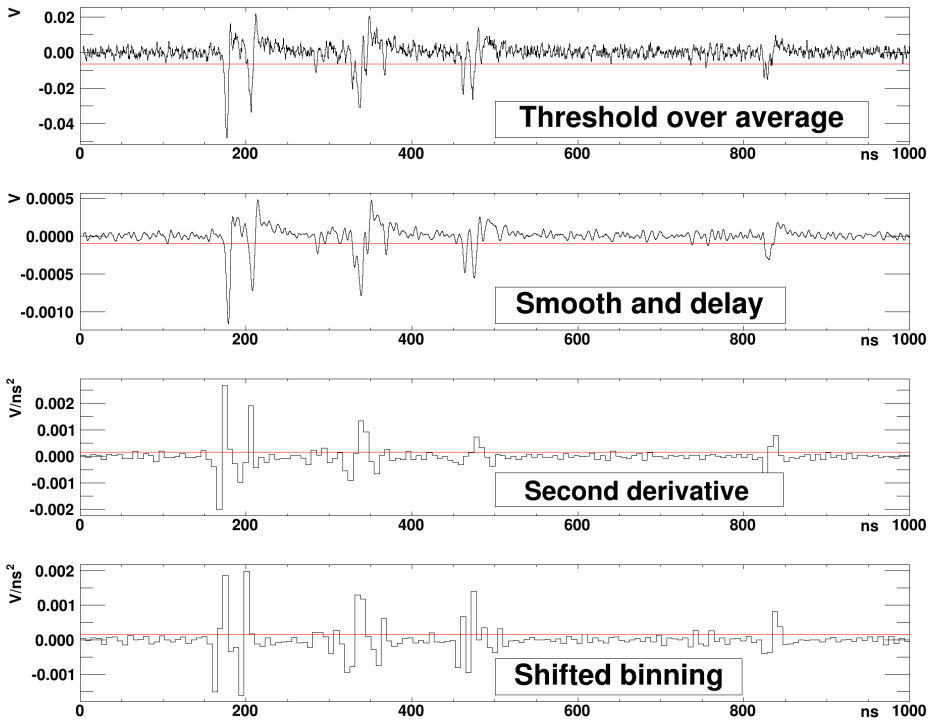


Figure 4.15: Illustration of the quantity on which a threshold is applied in the various cluster-counting algorithms. Each uses a set of parameters (smoothing width, threshold level) that were optimized for this run. The threshold level is indicated by the red horizontal line. The last image is the same as the second derivative, but with the binning shifted, to show that some clusters can be hidden by the binning (e.g. around 480 ns). This event is the same as shown in Figures 4.11 and 4.14

values is discarded. The distribution of truncated mean charge and clusters for the composed tracks is shown in Figures 4.17 4.18, respectively.

We also form the track-wise average time separation between clusters by doing a weighed average of the cell-wise average cluster separation for the events in the track. The weights are the number of clusters in the cells.

It is worth noting that the relative separations of the muon and pion peaks shown in Figures 4.17 and 4.18 are very different. For the truncated

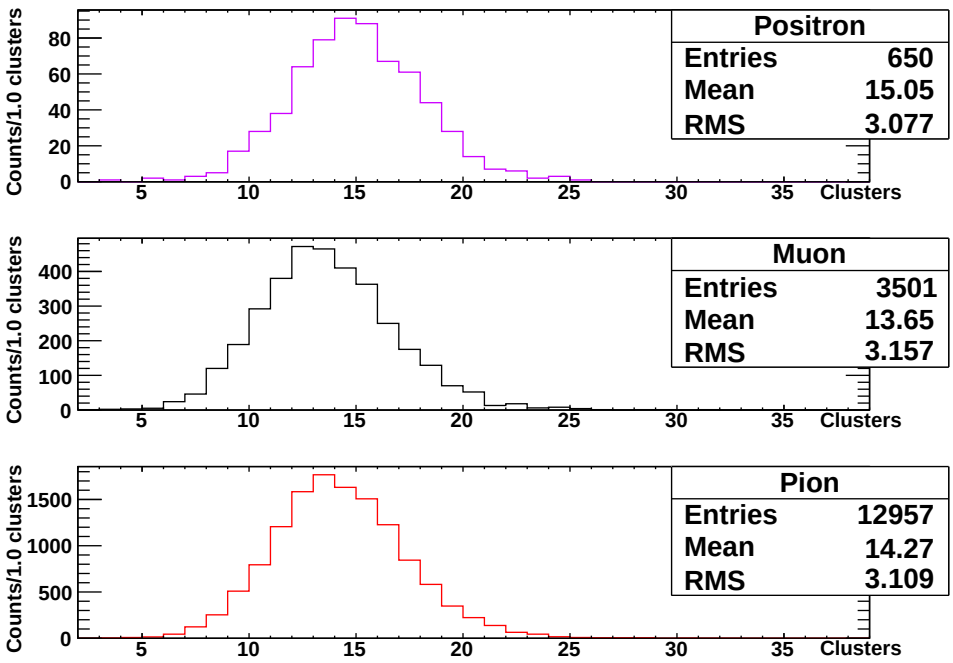


Figure 4.16: Number of clusters found for each species as identified by the TOF system. This is for a 210 MeV/c run using the second-derivative algorithm.

mean of the integrated charges, the relative separation between the peaks (difference in the location of the peaks, divided by the average of the two) is $\sim 10\,\%$, while for the cluster counting it is $\sim 5\,\%$. Naïvely this should mean that the cluster counting technique is less effective. However because the widths of these peaks are also very different, the two techniques turn out to be of comparable power (Figure 4.19).

4.5.3 Combined Likelihood Ratio

In order to combine the information from the truncated mean and the cluster count, we form likelihoods based on fits to the two quantities. These quantities are reasonably Gaussian (for non-empty tracks), so we fit them with Gaussian distributions $G_{s,k}$, for particle species s and measured quantity k. For a given track, the likelihood of the track coming from a particle s is

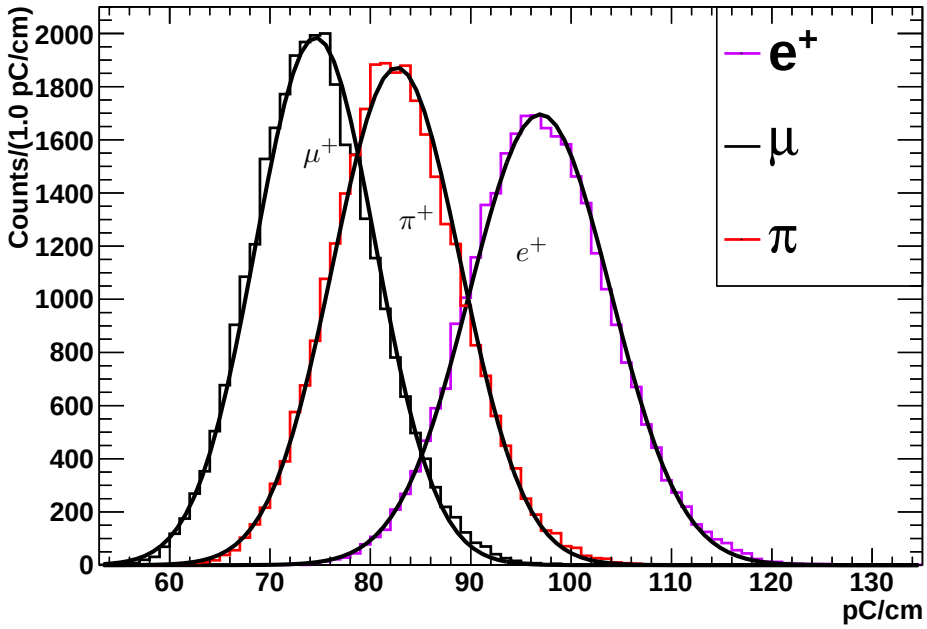


Figure 4.17: Truncated mean of charges (dE/dx) in composed tracks. This is using events from the same run as Figure 4.13. The three peaks from left to right are from muons, pions, and positrons, respectively. Here the particle populations are equal, as we compose an equal number of tracks for each species.

found by evaluating the product of the fitted distribution functions for both ks at the measured values. Thus if the measured truncated mean charge for a track is q and the clusters per cm of track are n, the combined likelihood is

$$L_s(q, n) = G_{s,\text{charge}}(q) \times G_{s,\text{clusters}}(n). \tag{4.1}$$

This combined likelihood ignores any correlation between the two quantities. The correlation is indeed non-zero but is somewhat weak (~ 0.3). Possibly combined likelihood models which make use of the correlation would be more effective, but we did not investigate this.

As mentioned in Section 4.2.5, the ability to identify muons and pions at

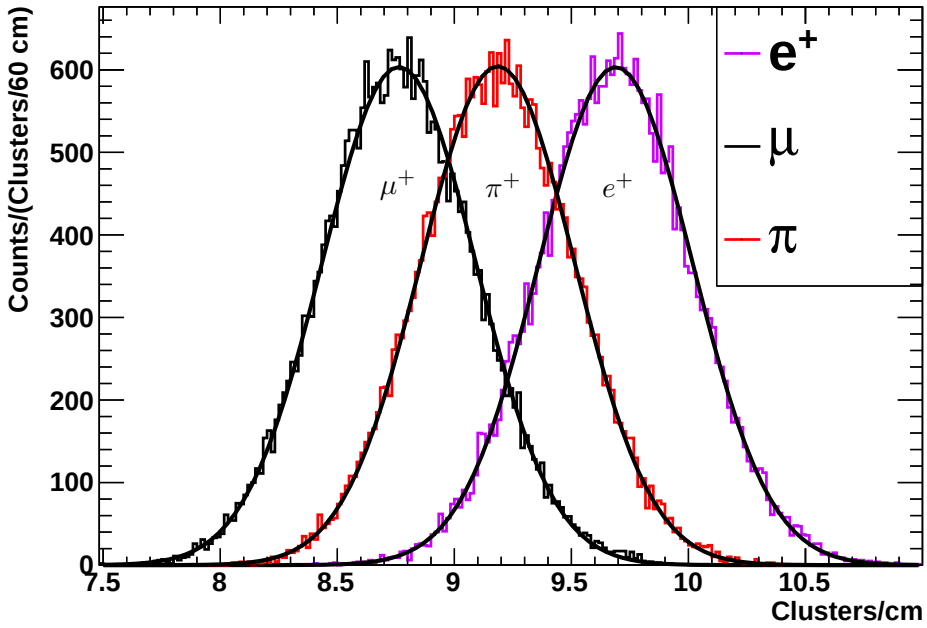


Figure 4.18: Number of clusters per cm in composed tracks. This is using events from the same run as Figure 4.16 and using the second-derivative algorithm. The three peaks from left to right are from muons, pions, and positrons, respectively. Here the particle populations are equal, as we compose an equal number of tracks for each species.

$\sim 210\,\mathrm{MeV/c}$ is our proxy variable for the performance of the prototypes. Thus we form a ratio of the combined likelihoods of being a muon and pion:

$$R(q,n) = \frac{L_\mu(q,n)}{L_\mu(q,n) + L_\pi(q,n)}. \tag{4.2}$$

This quantity's distribution is peaked at 0 for real pions and at 1 for real muons. A cut can be made that maximizes the separation according to some figure of merit. A typical way to demonstrate the performance is by making a rejection-selection efficiency plot. Consider the fraction of real pions that would also be identified as pions by the cut on R, and the fraction of real muons that would not be identified (that is, rejected) as pions by the cut on R. We can thus make a parametric plot of muon rejection efficiency

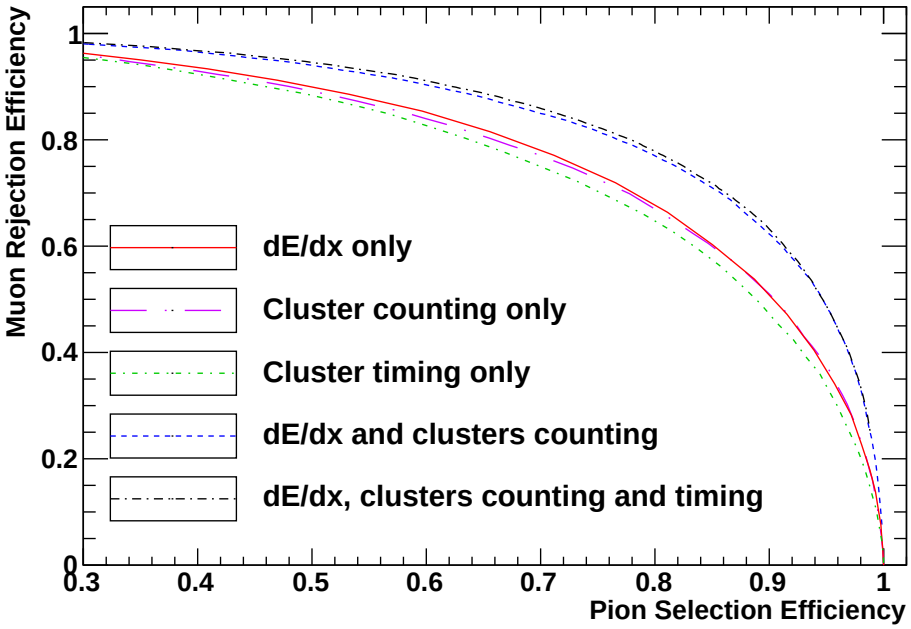


Figure 4.19: Efficiency graph for a run at 210 MeV/c, the same run as earlier figures. The cluster counting is done using the second-derivative algorithm. The upper two curves nearly coincide and are the efficiencies when cutting on the combined likelihood ratios. One combines the truncated mean, cluster count, and cluster separation, the other only truncated mean and cluster count. The lower three curves are the efficiencies when one cuts directly on the truncated mean, cluster count, or cluster separation quantities.

on the vertical axis and pion selection efficiency on the horizontal axis, with the parameter being the chosen R cut value (Figure 4.19). Similar efficiency graphs can be made for cuts directly on the physical quantities of charge and cluster counts.

4.5.4 Figures of Merit

The efficiency graphs are a good way to represent the performance of a particular setup, but they are two-dimensional and difficult to include in summaries. Thus we construct figures of merit in order to quantify the

performance of an equipment choice or algorithm. A convenient method is to set a given background rejection level and state the corresponding signal efficiency. In the muon rejection and pion selection plot, one may thus find the muon rejection efficiency corresponding to 90 % pion selection efficiency, or vice-versa. These figures of merit are easy to interpret physically and correspond to how detector performance is typically quantified in past experiments.

An alternative figure of merit turns out to better differentiate between algorithm parameter choices, but has a much less intuitive physical meaning. It is the maximum excursion on the muon rejection and pion efficiency plot from the origin of the graph. The curves on the graph approach $(0, 1)$ and $(1, 0)$ in the limits of R cut values of 0 and 1 respectively, but the curves can lie above that inscribed by a circle of unit radius. The length of the longest straight line joining $(0, 0)$ and the efficiency curve is taken as the figure of merit. In certain cases the performance is bad enough that the lines lie below that inscribed by a circle, in this case the alternative figure of merit is not meaningful, as it is identically 1.

All three figures of merit can be shown to be equivalent, in the sense that local maxima and minima lie in the same regions of parameter space. The maximum-excursion-from-origin figure gives better separation for those runs where it is meaningful (the majority). It is used for the optimization of algorithms, but the results are presented using the more intuitive figure of merit of pion selection efficiency at 90 % muon rejection.

4.6 Results

In this section we present the results of varying the cluster-counting algorithms, gas gain, various chamber positions, and other equipment choices.

4.6.1 Charge Integration

The time over which to integrate a signal in order to capture the charge deposition on the wire was determined empirically. In principle the optimal value varies from run to run depending on gas gain, dip angle of the beam,

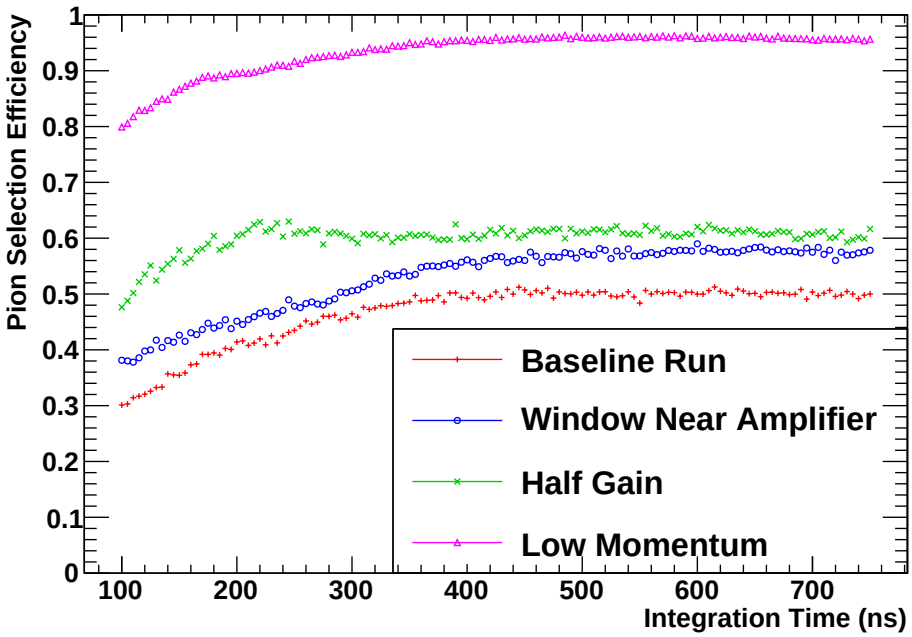


Figure 4.20: Pion selection efficiency using dE/dx only for several runs as a function of charge integration time. The baseline run is at a window 1883 mm from the amplifier, 210 MeV/c and using nominal gain as calculated by simulations. The low-momentum run is at 140 MeV/c. All the runs use the extra termination resistor.

and window position, but we wish to compare runs at different settings. Thus we look at the figure of merit for many different runs and choose a suitable compromise (Figure 4.20). As it turns out, the performance does not vary strongly as a function of integration time once the time is suitably long. We choose an integration time of 600 ns for the rest of the study.

4.6.2 Cluster Counting

The various cluster-counting algorithms have parameters that must be tuned empirically. By iterating this procedure many times using the same run, a "map" of the figure of merit can be created in the algorithm parameter space, the maxima of which are optimal values for the algorithms (Figure 4.21).

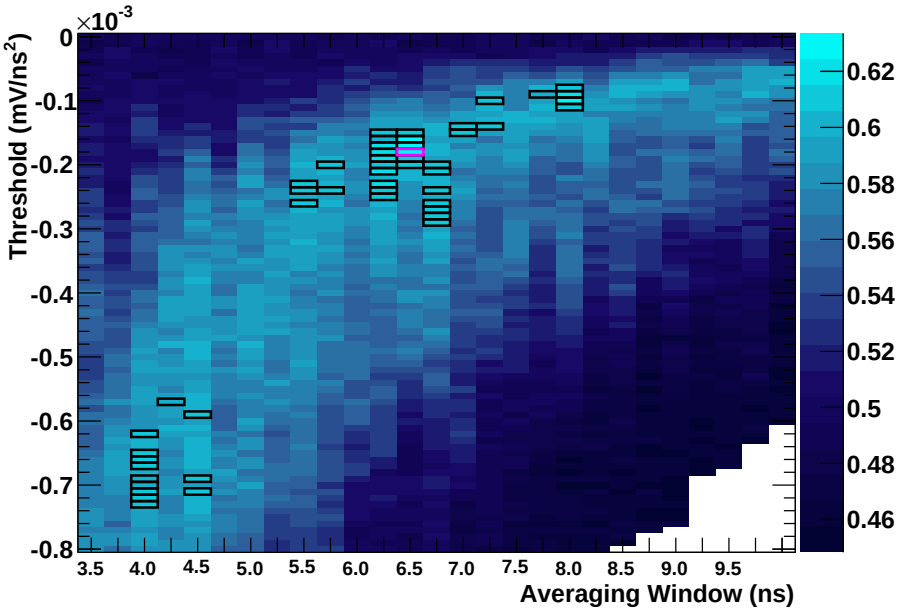


Figure 4.21: Example performance "heat map" for the second-derivative algorithm using a run at 210 MeV/c. In this case, the optimal parameters are an averaging window of ~ 6.5 ns and a threshold of ~ -0.16 mV. The black rectangles are the highest-performance regions, the magenta is the single best. The colour scale refers to the optimized figure-of-merit from Section 4.5.4.

While the figure of merit includes the PID performance from dE/dx and cluster counting, the dE/dx contribution is essentially constant even with the randomness introduced by the track composition process.

The optimal parameters vary from algorithm to algorithm and depend on the run used to optimize the parameters. In an operational experiment, only one set of parameters can be chosen, so some compromise will be necessary. Nevertheless, to compare the algorithms themselves, we may compare the performance of each algorithm when optimized on the same data run.

The chosen run has the following parameters: 10 degree dip angle, window 1883 mm from the amplifier, and 20 μm sense wire. A 370 Ω inverting

amplifier and 1855A Sub-miniature RG59/U signal cable with no extra connector are used. The beam momentum was 210 MeV/c. A total of 30784 triggers were recorded of which 7720 are asynchronous, and 680, 3649, and 13579 are positively identified as positrons, muons, and pions respectively. The remainder have TOF values more than $3\,\sigma$ away from the peaks or have no signal in the chamber.

The dip angle of 10 degrees was chosen rather than 0 in order to avoid space charge effects. The avalanches produced on the wire from the passage of a particle at zero degrees occur all in the same gas volume near the wire. This can affect the overall results and essentially makes 0 degrees a "special" angle. In an operating e^+e^- collider experiment the fraction of tracks suffering from space charge effects is negligible.

We explored a variety of algorithms, which are described in detail in Section 4.8. Each has some kind of threshold as one of the parameters, and some smoothing or averaging duration. A common feature is that the optimal smoothing or averaging duration is $\sim$ 5 ns, which indicates that extremely high sampling rate and bandwidth are not necessary to improve PID with cluster counting. The smoothing times correspond to Nyquist frequencies of $\sim$ 100 MHz. Our amplifiers have much higher bandwidth than this (Section 4.2.2), so using amplifiers with smaller bandwidth but better signal-to-noise ratios should improve overall performance.

In Table 4.1, the figure of merit is the pion selection efficiency for 90 % muon rejection. Here and in later plots, it is difficult to give a good estimate of the systematic uncertainty as many factors were not taken into account. For example the temperature of the gas in the chamber plays no role in our calculations, though the temperature did change during the data taking period. The track composition process involves drawing random numbers, so a contribution to the uncertainty from this can be estimated by composing multiple sets of tracks and seeing the distribution of results. Running the code 100 times yields an RMS deviation from the mean of ~ 0.05. The mean is what is reported in Table 4.1.

In the table, only algorithm *C* uses the "Timeout Booster" technique. We also tried applying the technique to the other algorithms, but it was no-

Table 4.1: Summary of optimal parameters for the various cluster-counting algorithms. The algorithms labelled A, B, C, D, and E are "Signal above Average", "Smooth and Delay", "Signal above Average with Timeout", "Second Derivative", and "Second Derivative (Two Passes)", respectively. The threshold is given with the appropriate units for that algorithm, and τ is the smoothing or averaging time in nanoseconds. Algorithm B has in principle two smoothing times, but the optimal value has them equal. The additional parameter Δ_t for the algorithms B and C are the delay and the timeout, respectively. The figure of merit ϵ_π is the pion selection efficiency for 90 % muon rejection.

Algorithm	Threshold	τ(ns)	Δ_t(ns)	ϵ_π
A	-6.5 mV	3.5		0.62
B	-0.1 mV	2.75	3.75	0.64
C	-3.0 mV	3.5	4.25	0.62
D	0.16 mV/ns^2	6.5		0.64
E	0.15 mV/ns^2	6.25		0.64

ticed that if the algorithm already has reasonable performance, the improvement from the timeout is negligible. Indeed the optimal timeout duration for the "Smooth and Delay" algorithm is zero, yielding the same performance as the bare algorithm.

Overall the best algorithm is the two-pass second derivative algorithm, but it is only marginally better than the other algorithms. The difference is less than the typical variation due to the track composition process.

It is fortuitous that even the simple algorithms have good performance, as they are reasonable to implement using a field-programmable gate array (FPGA) or even analog hardware.

In some sections that follow, the PID performance with optimized cluster counting refers to the use of a cluster-counting algorithm where the parameters were chosen to give the best figure of merit for that run. The optimal parameters vary from run to run, so in each case, we also run the algorithm on a given run using parameters that were optimal for a set of other runs. The other runs each vary in only a single parameter: the window, the HV settings, and the momentum. The average performance using these

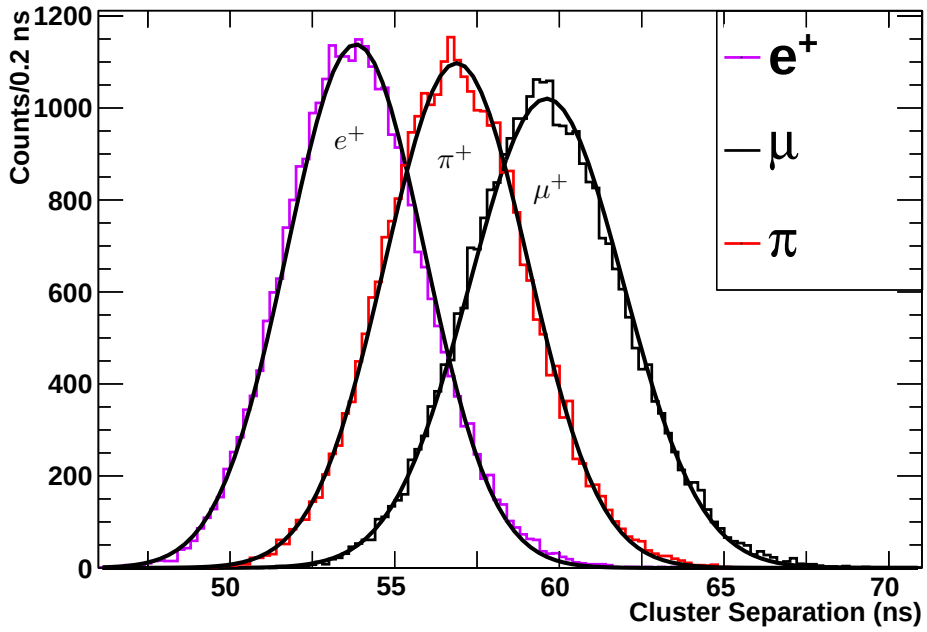


Figure 4.22: Track-wise weighed average of time intervals between clusters, in 50 ps units for each particle species. This is a run at 210 MeV/c. The three peaks from left to right are from positrons, pions, and muons, respectively

non-optimal parameters is labelled "sub-optimal cluster counting" in later figures.

4.6.3 Cluster Timing for PID

In each cell, we take the average of the time intervals between consecutive clusters. In the track composition process, we form a weighted average of the cell-wise averages, with the weights given by the number of clusters in each track. The resulting quantity gives a reasonable separation for each particle type (Figure 4.22).

Unfortunately the performance is not as good as either the traditional charge integration or cluster counting (Figure 4.19). In addition, if we form a tripartite combined likelihood, the improvement relative to the bipartite charge integration and cluster counting combination is negligible. Given the

increased computational complexity of calculating the average separations, it is unlikely that the timing information will be useful for PID purposes in a real particle physics experiment.

4.6.4 Dependence of PID on Gas Gain

The gas gain of the prototypes depends on the choice of sense wire voltage and on the gas. We tested only one gas, a mixture of helium and isobutane in a ratio of 90 : 10. A nominal voltage was selected as described in Section 4.2.3. The actual gas gain for our gas mix and voltages is on the order of 10^5, measured offline using an ^{55}Fe source. The procedure aims to obtain oscilloscope signals with roughly the same amplitude with all the amplifiers. The dependence of gas gain on sense wire voltage is approximately exponential [66]. In our case a $\pm 60\,\mathrm{V}$ change corresponds to a doubling or halving of the gas gain. The resulting performance after doubling and halving the gain is shown in Figure 4.23.

Previous to the experiment, the intuitive notion was that higher gas gains would be better, since the signals would stand out more from the random noise on the chamber wires. It appears however that this is not the case and that indeed better PID performance can be obtained at lower gas gains. Lower performance at higher gas gains is either due to gas effects (e.g. space charge) or to the amplifiers. We did not explore the even lower gains where the performance is expected to decrease again. Data runs using other amplifiers with different gain do show the eventual decrease (Section 4.6.7), so the optimal voltage is not too far from that shown in Figure 4.23 (within $\sim 100\,\mathrm{V}$).

When choosing a gas gain for an experiment the most important features are more often the tracking performance, ageing issues, and operational issues. This is more likely to influence the choice of specific gain, regardless of the PID performance. However, if PID performance is also highly valued, lower gains should be explored.

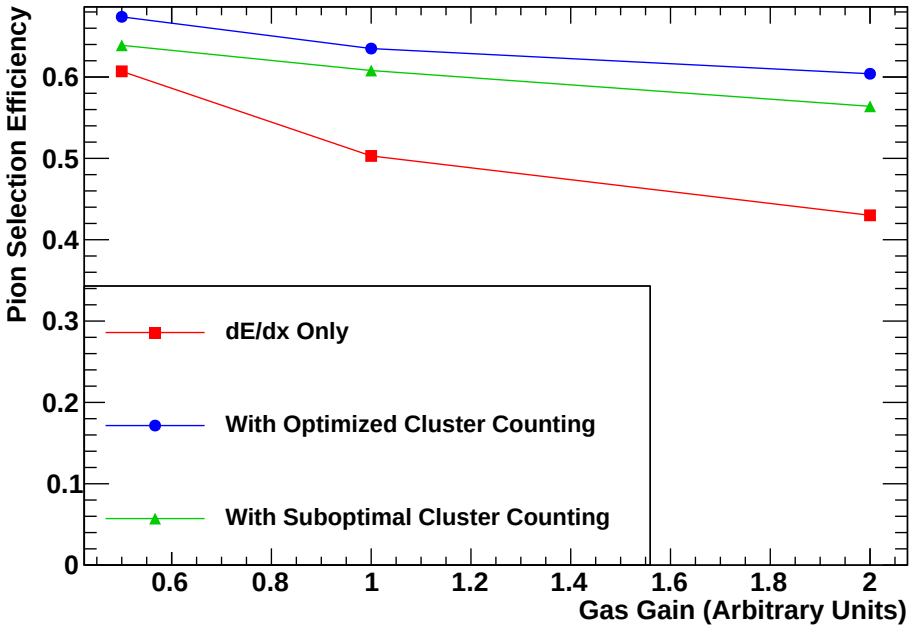


Figure 4.23: Variation in PID performance at the three gas gains that were explored. This is a run at 210 MeV/c using an inverting 370 Ω amplifier.

4.6.5 Momentum

As shown in Figure 4.24, the difference of ionization between pions and muons is greater at lower momenta. This is in agreement with theoretical expectations and simulations. As expected, the improvement from adding cluster counting is most noticeable at the momentum where the overall performance is worst, making the detector response more uniform.

4.6.6 Dependence of PID on Window (Z-position)

The prototypes have five windows at five thin aluminium positions along their 2.7 m length. The reference point is chosen to be the amplifiers, so the high-voltage connectors at the other end of the chamber are at 2700 mm. The centres of the five windows are 283, 816, 1349, 1883 and 2415 mm from the amplifiers.

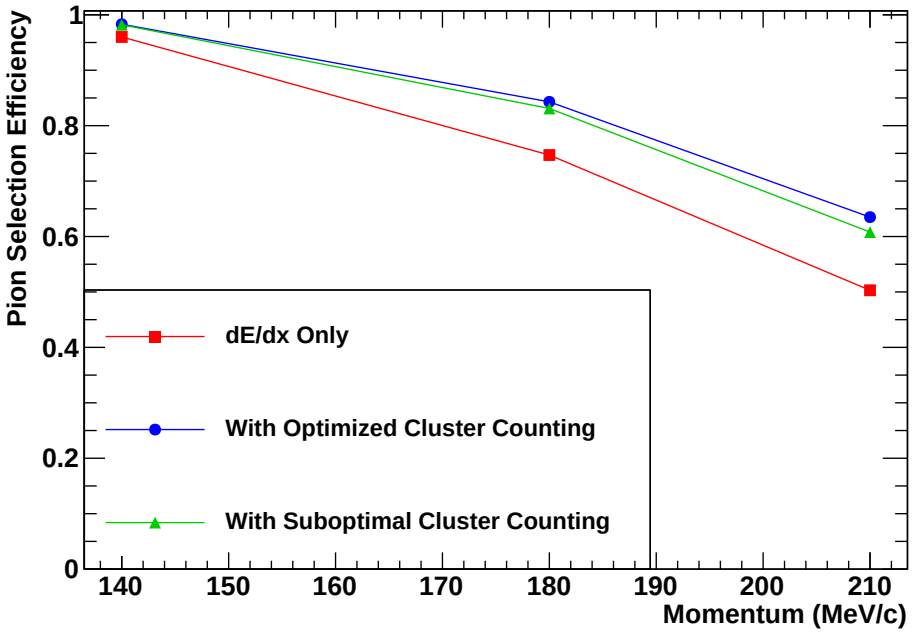


Figure 4.24: Variation in PID performance with momentum. These three runs all use the same amplifier with 370 Ω input impedance.

Most tests were performed at the windows 1349 and 1883 mm from the amplifiers, but a sequence of runs was taken to determine the effect of the signal propagating along the sense wire. The sense wire voltages were chosen as described in Section 4.2.3 at the middle position, but left unaltered for the other windows in the sequence. Thus the oscilloscope and amplifier saturations may change as a function of beam position.

The tungsten wire is very thin and has a non-negligible DC resistance (421 Ω for the 20 μm diameter wire), so it was expected that the performance would be better at the windows closer to the amplifiers. Indeed the runs taken at the two windows closest to the amplifiers have slightly higher efficiencies (Figure 4.25) than at the two furthest windows, but the difference is not large. The variation for this small data set is also not monotonic, the second-closest window to the amplifiers shows inexplicably better performance than the closest.

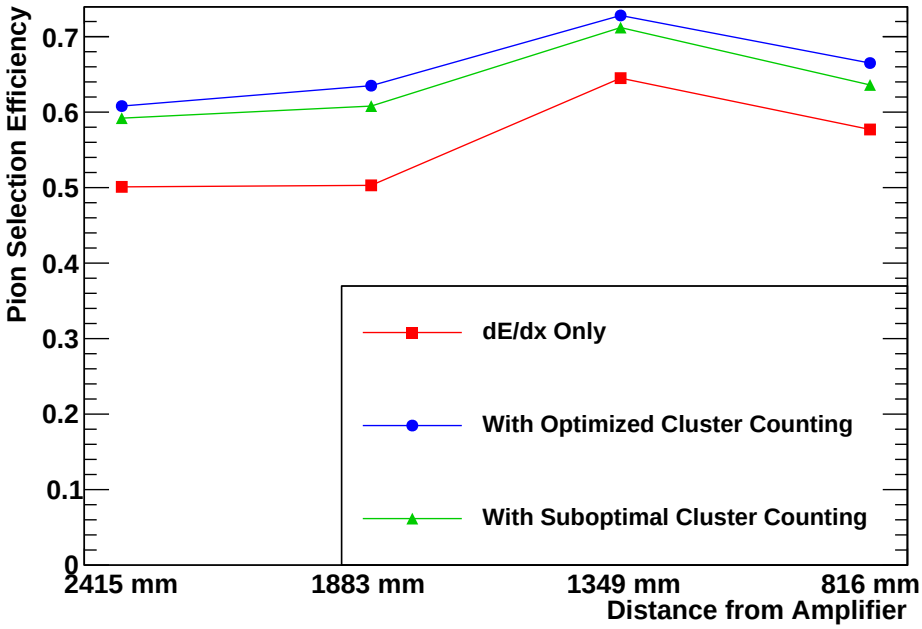


Figure 4.25: Variation in PID performance at the different windows of the prototype. These runs all use the same amplifier with 370 Ω input impedance.

4.6.7 Cables

As mentioned in Section 4.2.4, we tested two different cable types, and the effect of adding an additional header connector to simulate needing to feed through a bulkhead. Unlike the previous sections, we did not compare the performance of the cluster-counting algorithms using parameters optimized on the single run with non-optimal parameters. Thus the individual performance numbers may be optimistic, but the comparison between cable types can still be done. In Figure 4.26 we show the result from several runs using an amplifier with 50 Ω input impedance. The low gain columns have the Chamber *A* sense wire voltage at 1820 V, while the high gain columns are at 1835 V. The high gain voltage was chosen according to Section 4.2.3. Since our gain-doubling voltage is approximately 60 V, the low gain columns have about 84 % the gain of the high gain columns. The voltages are higher

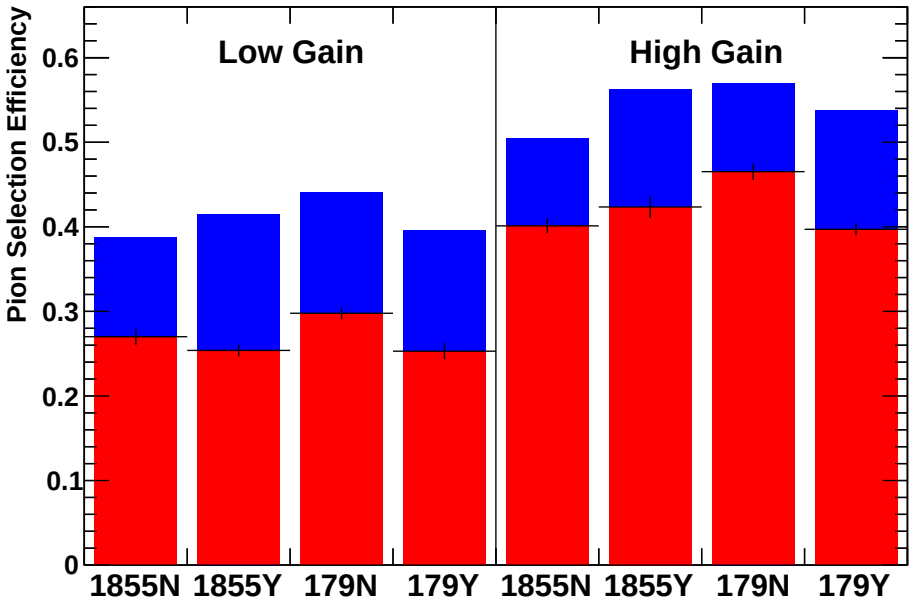


Figure 4.26: Variation in PID performance using different cable types. The red is the performance using charge integration only, the blue on top is the additional performance gain from combining charge integration and cluster counting. All the runs are at 210 MeV/c and use the same amplifier with 50 Ω input impedance. "Y" and "N" designate the presence or absence of an extra header connector. The first four runs are at low gas gain, and the last four are at higher gas gain.

than for the runs described in the earlier sections because the amplifiers have different electronic gain. The cable types 1855 and 179 are described in Section 4.2.4, while the Y and N designations indicate the presence or absence of the extra header connector, respectively.

A general trend to be noticed is that the high gain columns have noticeably better performance than the low gain columns, which is contrary to what was shown in Section 4.6.4. This is likely because these amplifiers have different electronic gains and the selected wire voltages do not lie in the same performance region as the results shown in Section 4.6.4.

The cable type and the inclusion of the header connector only marginally

affect the figure of merit, by an amount less than the typical variation between identical runs and from the track composition process $\sim 5\,\%$. It is tempting to see that the 179N columns are the highest between the two sets, but the difference is not nearly as dramatic as the variation due to gas gain or the additional contribution of cluster counting itself.

4.6.8 Amplifiers

As described in Section 4.2.2, we tested several types of amplifiers, mostly distinguished by their input impedance and gain. We remind the reader that the sense wire voltages used are different for the various amplifiers, and were chosen to get approximately constant signal amplitude as described in Section 4.6.4.

In Figure 4.27, the results from three different amplifiers at two different positions along the sense wire are shown. The input impedance of each amplifier is indicated, and the amplifiers with the same labels are the same for the two different positions. The $370\,\Omega$ "inv" amplifier returns an inverted signal, while the others do not.

There is considerable variation between the amplifiers, but the general result is that the $370\,\Omega$ amplifiers give the best results. This indicates the importance of matching the amplifier input impedance with the impedance and termination of the drift chamber itself. Unfortunately the indication of the best amplifier is not very strong, as a proper study of the optimal gas gain for each amplifier was not done in this experiment. The variation between the amplifiers in Figure 4.27 is of the same order as the variation with gas gain for a single amplifier shown in Figure 4.23. It is possible that the variations seen here are mostly due to gain effects rather than the impedance and implementation details of the amplifiers.

4.6.9 Summary of Results

The studies undertaken attempt to explore a multidimensional parameter space, so the results are difficult to summarize concisely. Here we restate the lessons learned from each study described above.

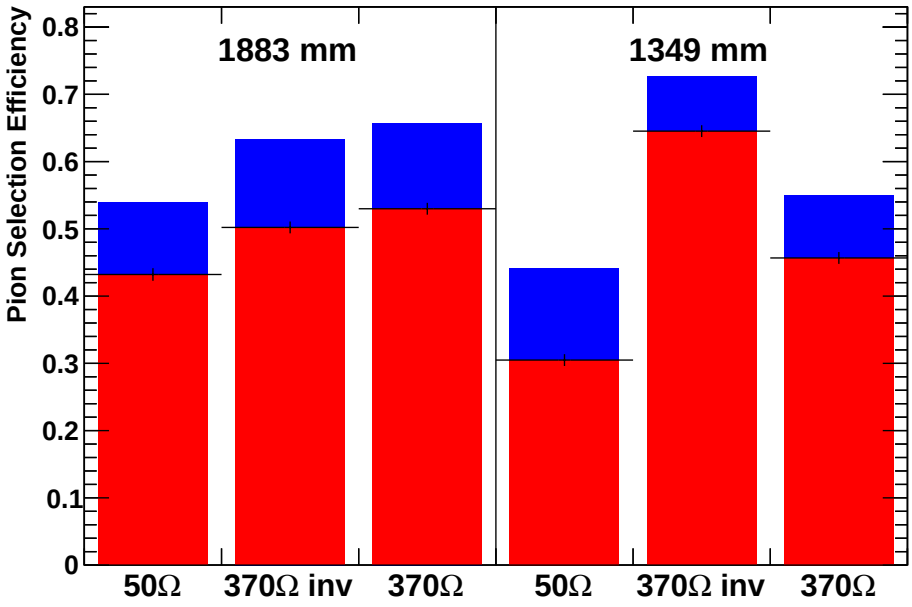


Figure 4.27: Variation in PID performance using different amplifiers. The red is the performance using charge integration only, the blue is the additional performance gain from combining charge integration and cluster counting. The upper labels indicate the beam position along the sense wire, measured from the amplifier. The "inv" label indicates an inverting amplifier.

The various cluster counting algorithms all perform roughly equivalently (Section 4.6.2). Their parameters must be optimized for good performance, but the regions of good performance in parameter-space are quite large. Even sub-optimal parameters only give slightly worse performance. More advanced techniques (such as the timeout booster) can compensate for a less-optimized algorithm, but are unnecessary when the algorithm is optimized properly.

Optimal smoothing for the cluster-counting algorithms is on the order of a few nanoseconds, indicating that a higher sampling rate is unnecessary. The corresponding Nyquist frequency is on the order of hundreds of MHz. This means that the successful implementation of cluster counting does not

depend on getting overly expensive or customized hardware. Indeed the best algorithm studied simply applies a threshold to the second-derivative of the signal, a process that can be done with analog electronics or in an FPGA.

Cluster timing gives results that are slightly poorer than cluster counting used alone (Section 4.6.3). When combined with charge integration and cluster counting however, the improvement is minor compared to charge integration and cluster counting without the cluster timing. Given the additional complexity of storing and calculating average cluster timings, this technique is unlikely to be worth exploring further.

PID performance depends strongly on having the proper wire voltages and thus gas gains (Section 4.6.4). In some configurations, higher gain is not necessarily better, but this is dependent on the choice of amplifier. Thus for a given amplifier and equipment configuration, the optimal gas gain must be carefully determined.

There is not much variation in PID performance as a function of the beam position along the sense wire length (Section 4.6.6). Since the signal is attenuated while travelling along the sense wire, this effect is coupled with the gain of the amplifier and the choice of wire voltages.

The choice of cable types and additional connectors seems to have a negligible effect on the PID performance (Section 4.6.7). Performance is very sensitive to the choice of amplifier (Section 4.6.8), but this is coupled with the sense wire voltage. There is a weak indication that matching the amplifier input impedance with the impedance and termination of the chamber itself gives better performance.

4.7 Conclusions

The general result is clear: implementing cluster counting increases the particle identification capability of a drift chamber. We make no claim of having found the optimal equipment and analysis techniques in the multidimensional parameter space that we explored. Thus we can state that cluster counting improves PID performance even in sub-optimal conditions.

The absolute improvement in the pion selection efficiency at 90 % muon

rejection is generally around 10 % (e.g., from 50 % to 60 %, and see Figure 4.19). The improvement is greatest when the PID performance from charge integration only is poorest, thus making the detector PID response more uniform.

The optimal smoothing times for cluster-counting algorithms are on the order of a few nanoseconds, corresponding to a Nyquist frequency of hundreds of MHz. Thus successful cluster counting can be accomplished even with modest hardware.

All future particle physics experiments that use a drift chamber for PID should strongly consider a cluster-counting option. This study shows that performance gains can be obtained that justify the additional complexity and cost of a cluster-counting drift chamber.

4.8 Cluster-Counting Algorithms

Here are contained precise definitions of the cluster-counting and smoothing algorithms used in this work. We define a signal or trace as a series of voltage samples indexed by a discrete time variable $V(t)$. Though the time variable has units (in our raw format the units are 50 ps), here we treat it as an integer index. In general, a signal will have N samples indexed with integer t running from 0 to $N-1$.

4.8.1 Smoothing Procedures

Two types of smoothing are used in the algorithms. One involves replacing each element of the signal by the average of itself and its neighbours, without reducing the total number of elements. The other reduces the total number of elements, and each element's value is the average of a set of elements in the original signal.

Boxcar Smoothing

The so-called "boxcar smoothing" with n frames substitutes each sample with the average of itself and the previous $n-1$ samples. The first n samples

(0 to $n-1$) are a boundary case, replaced simply by $\tilde{V}_n(n)$.

$$\tilde{V}_n(t) = \begin{cases} \frac{1}{n} \sum_{i=0}^{n-1} V(t-i) & t \geq n, \\ \tilde{V}_n(n) & t < n. \end{cases} \tag{4.3}$$

Averaging

The so-called "true averaging" procedure produces a signal with a reduced number of samples. For an n-frame averaging, the result is a series of $k = N \div n$ voltages (floored division), indexed with the integer $\bar{t}$ running from 0 to k.

$$\bar{V}_n(\bar{t}) = \frac{1}{n} \sum_{i=0}^{n-1} V(s + n\bar{t} + i) \tag{4.4}$$

Here, n is the number of samples that are averaged, $s = N \bmod n$, and N is the total number of samples in the original trace.

This averaging has the potential to "divide" cluster signals if the averaging bin edges lie on top of a cluster (Figure 4.15). Thus it is useful to also shift the smoothing bins by adding $n \div 2$ to the argument of V inside the sum. If the smoothing is done with and without the shift, it is less likely that the same cluster will be divided in both cases, compared to doing the smoothing only one way.

4.8.2 Signal above Average

This algorithm has two parameters: a number of frames for smoothing and a threshold. From the non-smoothed signal at time t is subtracted the n-frame smoothed signal at time $t-1$. If the resulting quantity crosses the threshold Δ downwards, a cluster is identified at that time.

The cluster times found by this algorithm are those t in $\{\max(n, 2)..N\}$

that satisfy

$$\begin{aligned} (V(t) - \tilde{V}_n(t-1) < \Delta) \text{ and} \\ (V(t-1) - \tilde{V}_n(t-2) \geq \Delta). \end{aligned} \tag{4.5}$$

4.8.3 Smooth and Delay

This algorithm has four parameters: two smoothing times, a delay, and a threshold.

Two copies of the original signal are smoothed by different amounts (p and q frames) using the "boxcar smoothing". The q-frame smoothed copy is then delayed by d frames, and the two copies are then subtracted. If the resulting quantity crosses the threshold Δ downwards, a cluster is counted at that time.

The cluster times found by this algorithm are those t in $\{d..N\}$ that satisfy

$$\begin{aligned} \frac{\tilde{V}_p(t) - \tilde{V}_q(t-d)}{d} < \Delta \text{ and} \\ \frac{\tilde{V}_p(t-1) - \tilde{V}_q(t-1-d)}{d} \geq \Delta. \end{aligned} \tag{4.6}$$

The "Signal above Average" algorithm is a special case with $p = 1$, $q = n$, and $d = 1$. Another special case can be constructed with $d = 0$ with the denominator set to 1.

It can be shown that if the two smoothing times are equal ($p = q$), the quantity computed with smoothing q and delay d is identical to that computed with smoothing d and delay q. Thus the parameter range can be restricted to $d > q$ without loss of generality.

4.8.4 Second Derivative

This algorithm has two parameters: a smoothing time and a threshold. It uses the true averaging procedure rather than the "boxcar smoothing", so the time is labelled $\bar{t}$ as in Section 4.8.1. Simply put, the second derivative

is calculated and compared with a threshold.

The second derivative is calculated as follows:

$$\bar{V}''(\bar{t}) = \frac{1}{\delta^2}\Big([\bar{V}(\bar{t}+2) - \bar{V}(\bar{t}+1)] - [\bar{V}(\bar{t}+1) - \bar{V}(\bar{t})]\Big) \tag{4.7}$$

where δ is the time interval corresponding to the n samples that were averaged to do the smoothing.

The times of the clusters found with the second-derivative algorithm are those $\bar{t}$ in $\{0..N \div n\}$ that satisfy

$$\bar{V}''(\bar{t}) < \Delta \text{ and } \bar{V}''(\bar{t}-1) \geq \Delta. \tag{4.8}$$

Because this algorithm uses the true averaging, it suffers from the problem of potentially "dividing" cluster signals between smoothing bins (Figure 4.15). Thus we also implemented a two-pass second-derivative algorithm that looks for clusters a second time on the averaged signal with a delay applied as described in Section 4.8.1. The numbers of clusters found in each pass are added together. It is understood that the resulting cluster count is inflated because many clusters will be double-counted, but nevertheless it is an appropriate variable for identifying particles.

4.8.5 Timeout Booster

The so-called "timeout booster" takes as an input the list of clusters found by one of the above algorithms. It considers these as cluster candidates, and validates or rejects each one in turn.

For a given cluster candidate, the voltage and time in the original waveform at which the cluster-finding algorithm was triggered is recorded. Then following the waveform forward, the voltage is checked to see when it has recovered above the recorded value (the pulses are negative). If the voltage recovered within the timeout window, it is a short-lived pulse and thus rejected as a fake. If the timeout is reached without the voltage recovering, it is long-lived and kept as a real cluster.

For a list of potential clusters t'_i, real clusters satisfy

$$V(t) < V(t'_i) \text{ for all } t \text{ in } \{t'_i..(t'_i + T)\} \tag{4.9}$$

where T is the chosen timeout. The rejection of fake clusters by the timeout procedure permits the use of lower thresholds in the original algorithm. The lower threshold increases the efficiency of finding real clusters (smaller miss rate) but increases the rate of detecting fake clusters. The timeout procedure then eliminates most of the fake clusters, keeping the real ones.

Chapter 5

Multi-Cell Prototype

5.1 Introduction

In 2011 the Italian prototype called "proto 2" was constructed at LNF-INFN in Frascati, Italy. Machining and construction of the mechanical components was done by technical staff at LNF, except for the outer shell which was constructed by a commercial firm. The stringing of the 205 wires was done by hand over the course of several weeks by Giuseppe Finocchiaro and me.

All of the assembly was done in a clean room at LNF, and all inner components were cleaned with acetone. The smaller pieces were also cleaned with an ultrasonic cleaning tank. The clean room was class 10000-equivalent; the designation refers to the maximum allowed number of dust particles greater than 0.5 μm diameter per cubic foot. Class 10000 is not particularly clean as far as clean rooms go, it is only 2 classes cleaner than regular room air (there are 8 classes cleaner than regular air defined by ISO standards). The cleanliness is mostly important for preventing sparking and ageing behaviour in drift chambers built to be used for a long time or with very high rates (see Section 2.2 for a discussion of ageing). For prototype work, this cleanliness is expected to be adequate. Only the wires were omitted from the cleaning because of their fragility.

The prototype was tested at the M11 beamline at TRIUMF [64]. This is

the same facility that was used in the single-cell prototype study described in Chapter 4. The beam is composed of positrons, positive muons and pions, and protons. The beam particles have roughly the same momentum for a given test, and momenta of 140 to 350 MeV were used. In our test we added a polyethylene absorber to block the protons. The analysis of the data taken during the beam test is described in Chapter 7.

In the following sections we describe the wire materials, layout, and stringing; the outer structure of the prototype; and the electronics used to read the sense wire signals.

5.2 Wires

There are 205 wires in total: 28 sense wires, 152 field wires, and 25 guard wires (see Figure 5.1). The guard wires and outer field wires are outside of the active cell areas and serve to mimic the influence of an infinite plane of cells. This makes the fields inside the cells more homogeneous. The guard and field wires are made of bare aluminium, while the sense wires are either tungsten-rhenium or molybdenum, both gold coated.

The wires were inspected with a microscope to determine their surface quality. The the gold-coated tungsten-rhenium and molybdenum wires were very smooth, but the bare aluminium wires were observed to have large $\sim 50\,\mu$m protrusions from the surface (Figure 5.2). The protrusions may have simply been dust on the wire, but we were unable to identify them with the microscope. The $80\,\mu$m diameter aluminium wires had less severe defects than the thicker $120\,\mu$m wires, so these were mostly used throughout the chamber. Some initial outer wires of $120\,\mu$m diameter were installed before the defects were noticed. These were not replaced as they are for guard wires outside the active cells, so their effect should be negligible.

As mentioned above, 9 of the field wires used the thicker aluminium, but the rest are all $80\,\mu$m, and the difference is not expected to change the electric fields in the cells. The sense wires for cells 0 to 6 used the traditional tungsten wire, while all others used the molybdenum wire (see Figure 5.1 for the cell numbering). The intent was twofold: to evaluate the behaviour

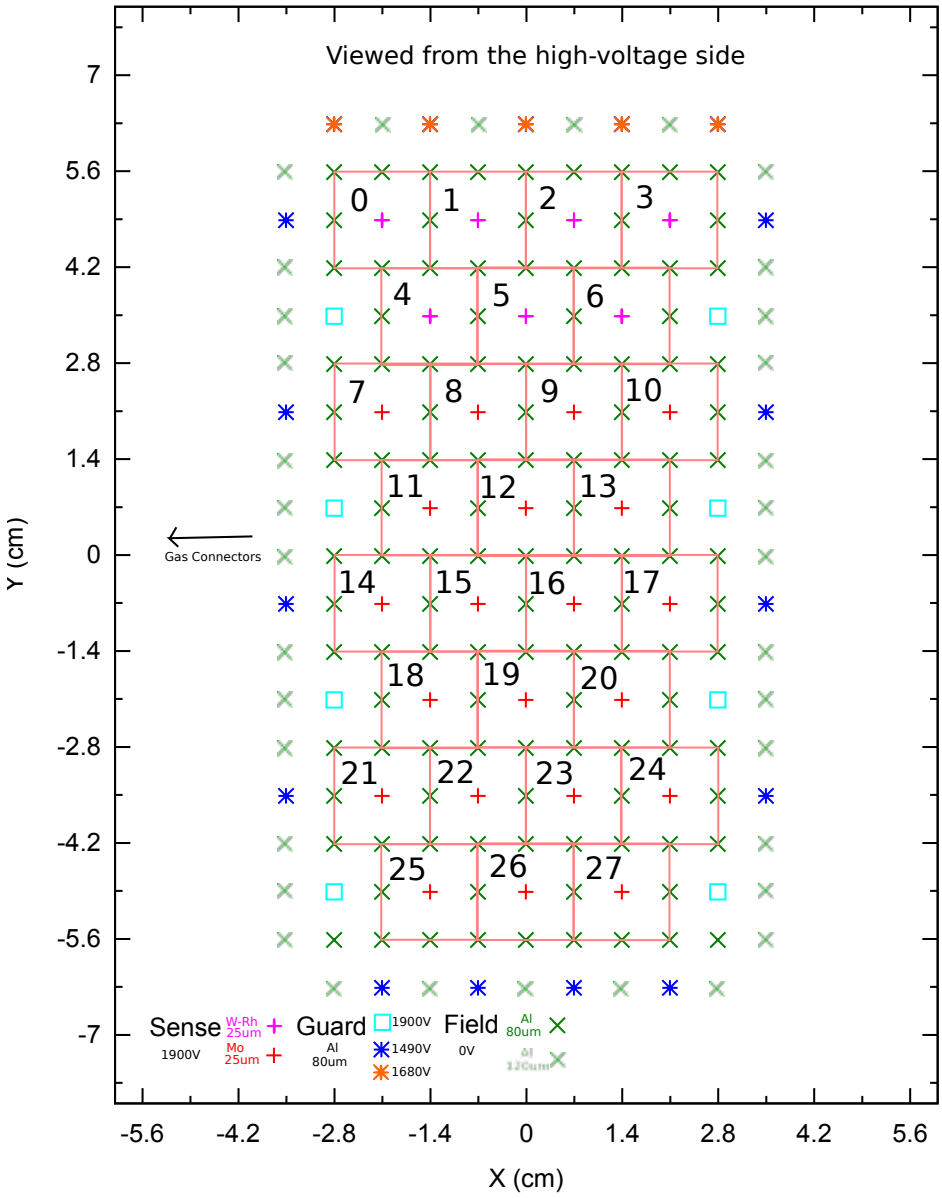


Figure 5.1: Layout of all the wires in proto 2. The type, materials, diameters, and optimal voltages on the wires are indicated in the legend. The numbers in each cell indicate the numbering system used in this work. In the beam tests described in this thesis, the charged particle tracks come from the top of the figure and exit through the bottom.

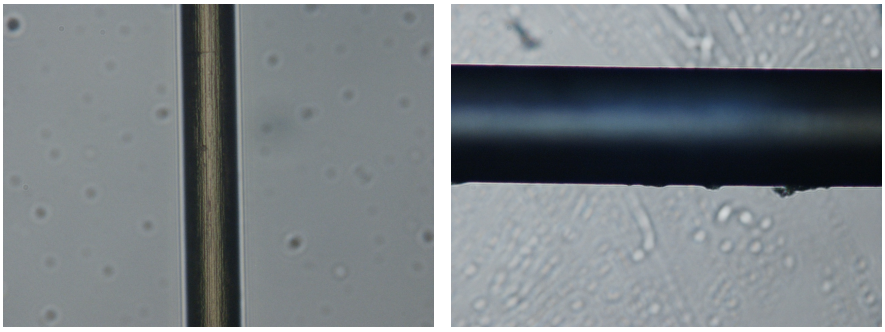


Figure 5.2: Microscope image of the 25 μm diameter gold-coated molybdenum wire (left) and the 80 μm diameter bare aluminium wire (right). The scales are different.

of the molybdenum wire during stringing (e.g., is it easier or harder to use than tungsten), and to compare the performance of the tungsten and molybdenum cells. The performance analysis was never carried out.

To obtain a nominal gravitational sag of 200 μm in the middle of the chamber, we calculated the required tension:

$$T = \frac{g\pi r^2 \rho L^2}{8S}. \tag{5.1}$$

where g is the gravitational acceleration, r is the radius of the wire, ρ is the volumetric mass density of the wire, L is the length of the wire, and S is the desired sag [12]. We calculated T/g to find the mass of a weight to be hung from one end of the wire to provide the tension. For our prototype, masses of 120, 53.0, 39.1, and 21.5 g (grams) were indicated for the thick aluminium, thinner aluminium, tungsten, and molybdenum wires, respectively. The total force on the endplates is thus about 10 kgf (kilogram-force).

The wires are all strung parallel to each other, and each is 2.5 m long. We only read out the signal on one end of the wire, and the far end is terminated to prevent reflections, so the parallel wires give no z-coordinate information, i.e., we have no information about the distance of charged particle tracks along the wire axis. The TRIUMF M11 beam spot size is $\sim 1\,\text{cm} \times 20\,\text{cm}$,

and in all our tests we only expose the chamber to the beam through one window at a time, so all our analyses are effectively two-dimensional.

5.3 Structure

The aluminium structure and surface defects in the internal aluminium surfaces were attached and filled in with a two-part epoxy. This epoxy was the same as used for the LHCb drift tubes (Araldite AY103-1 from Huntsman[67]). Since the epoxy was black, it was later covered with shiny aluminium tape to maintain a conductive surface and so that the chamber would not look ugly.

Electronic feedthroughs are pin-like devices with an inner hollow metallic cylinder and an outer plastic sheath. The inner cylinder can be crimped to hold the wires, and the plastic sheath's diameter is such that it fits snugly into the holes of the endplates. For proto 2, we re-purposed left-over feedthroughs from the KLOE drift chamber [68]. A schematic of a feedthrough with tensioning weight is shown in Figure 5.3.

The feedthroughs inserted into the endplate and the wire ends on the feedthroughs were sealed with acrylic glue. The glue had low viscosity and tended to fill cracks, but may have entered into the drift chamber inner surface before drying. We did not do extensive tests of glues as was done for BaBar [69]. A photograph of the fully-strung chamber is in Figure 5.4.

The outer shell is composed of a 3 mm sheet of aluminium folded into a rectangular box and welded down a side. Six thin windows are machined out of the walls to reduce the material exposed to the beam. Photos of the shell can be seen in Figures 5.5 and 5.6. The endplates into which the wire feedthroughs are inserted are made of a material called permaglass, which is a kind of fibreglass. On top of the permaglass endplates are several layers of material for supporting the electronics. The design ensures that when the electronic connectors are connected and disconnected from the wires, the force is exerted on the feedthroughs and not the wires or crimp pins, which are fragile.

In the lab we tested for gas-tightness using a "sniffer" device, which

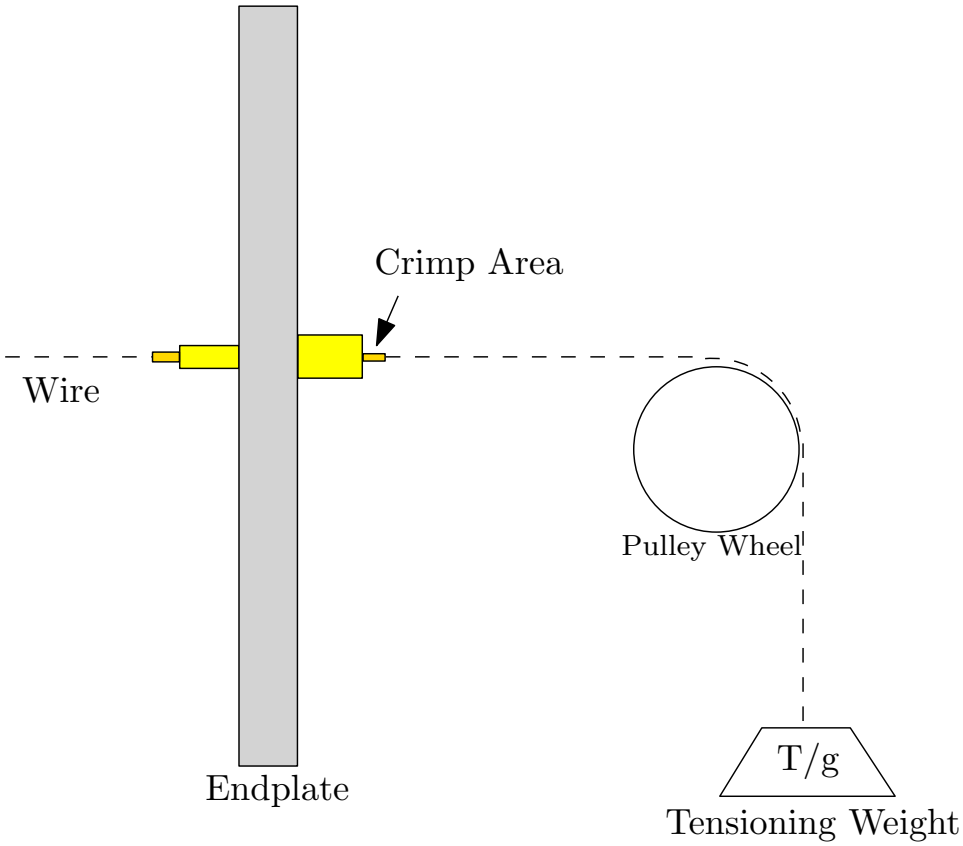


Figure 5.3: Schematic of the chamber endplate with a feedthrough inserted, and a tensioning weight applied before crimping.

samples the air and beeps when detecting certain gases. We flushed the chamber with pure helium, as this is the smallest gas particle in our mixture and would most readily flow through leaks. The sniffer was not very reliable (e.g., the readings would fluctuate rapidly, and sometimes would refuse to detect a known gas coming directly from a hose), but we were able to use it to find a few significant leaks and seal them with epoxy.

During the gas testing, the helium exiting the chamber was vented with a long thin tube leading to an external window in the lab. Unfortunately it was overlooked that this long thin tube would offer resistance to the flow of

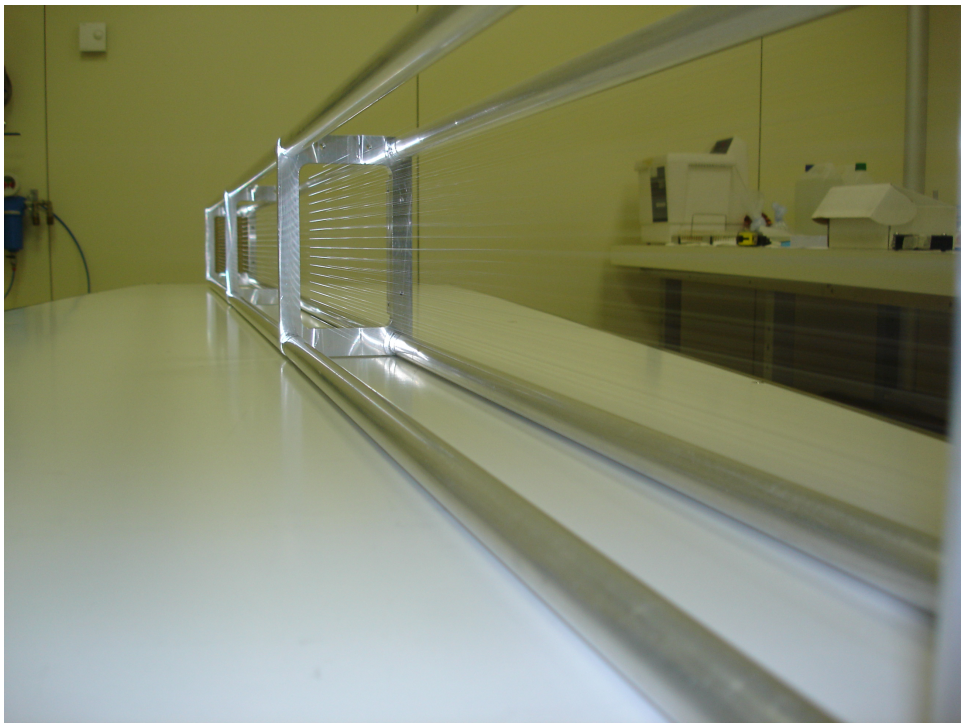


Figure 5.4: Structural frame and wires of the prototype. Still missing are some endplate components and the outer aluminium shell.

helium, much as long thin wire could have significant electrical resistance. At some point the helium gas flow was increased, but rather than flowing out of the tube, it increased the pressure in the drift chamber. The drift chamber is not designed to operate at significant over- or under-pressures relative to the atmosphere, and the thin chamber wall at the large central window exploded. A few wires were broken and had to be re-strung, and a new window had to be welded to close the large hole (see Figure 5.6).

5.4 Electronics

Connectors for the high-voltage power supplies are mounted on one end of the drift chamber (Figure 5.7). Front-end electronics are mounted on the other end. The front-end electronics are composed of four boards with pre-

Figure 5.5: Photo of the outer shell showing the window positions on proto 2. The window sizes are chosen to admit all angles of tracks that cross the 8 layers of cells, i.e., ±20–30°.

amplifiers, each connected to 7 sense wires in adjacent layers. The amplifiers collect charge (integrated current) and output a voltage proportional to it. Their gain is 8 mV/fC and the rise-time of their pulses is about 2.4 ns [37]. The amplifier outputs are connected to digitizers via 10 m long cables (the design length of the cabling for Super$B$). The digitizer takes the continuous stream of voltage and samples it at 1 ns intervals, producing digital voltage measurements that can be easily analyzed using computer programs. The digitizer is a commercial CAEN V1742 with a bandwidth of 500 MHz, and 12-bit analog-to-digital converters. A photo of the instrumented end of the drift chamber is shown in Figure 5.8.

The completed drift chamber was wrapped in a copper foil that was sol-

Figure 5.6: Close-up photo of the large middle window that exploded, after repair. On the upper left one can also see a repaired corner in the 3 mm aluminium sheet.

dered closed and electrically grounded to the aluminium box. This provides additional shielding from electromagnetic interference. The whole detector is mounted on a movable support that can also tilt the chamber along its long axis. This allows us to take data with tracks at multiple angles in 3 dimensions (see Figure 5.9).

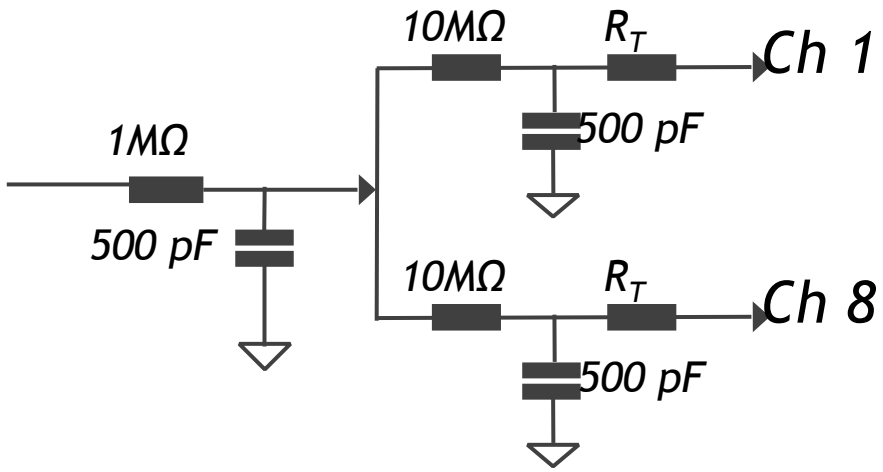


Figure 5.7: Circuit diagram for the Proto 2 high-voltage connections on the sense wires. The filtered connection to the high-voltage power supplies is from the left, and the sense wires are connected on the right. The termination resistor $R_T \sim 300\,\Omega$ is chosen to match the impedance of the sense wire. There are 8 channels per high-voltage connector (only 2 shown here, labelled Ch 1 and Ch 8), but only 7 are used in this prototype due to the layout of the wires. There are 4 connectors for the whole prototype.

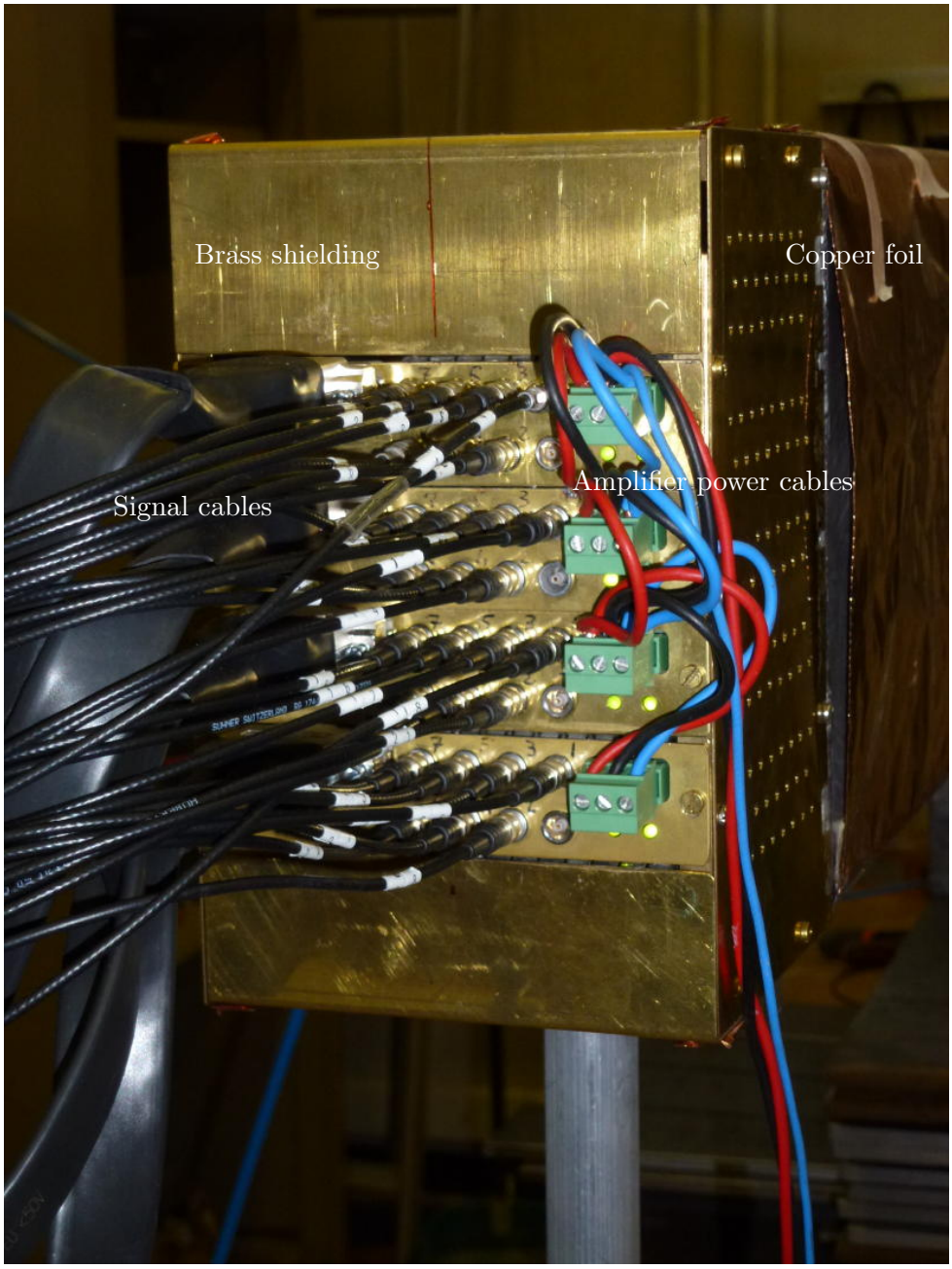


Figure 5.8: The instrumented end of proto 2, showing the 28 connectors from the pre-amplifiers and other electrical connections. The brass box is for shielding.

Figure 5.9: Proto 2 mounted on a movable table and tilting support, for an early beam test at LNF.

Chapter 6

Theoretical Tracking Improvements

In a traditional drift chamber, only the leading edge of the signal from each cell is used for tracking. The arrival time of the signal relative to a global trigger is determined by a threshold algorithm.

From calculation, simulations, and calibrations, a time-to-distance relation can be obtained, mapping arrival times into distances of the track from a sense wire. Different levels of refinement can be used to improve the track distance estimate: information from multiple cells can tell us on which side of a cell the track passed, corrections can be applied for temperature and pressure variations in the gas, etc. All such refinements come from information outside of the single-cell signals.

A tracking refinement exists that uses additional information from the single-cell signal itself. If the arrival times of the individual clusters in the signal are known, a more precise weight can be assigned to the cell when performing the global track fit. This section explains how this single-cell improvement is possible.

6.1 Model

We use a simple model to illustrate, with the understanding that real drift chambers have many additional complexities that may ruin the effect. Part of this research is to evaluate whether indeed the potential improvements to the tracking resolution are measurable in a real drift chamber. Our model consists of a single infinite cell with a single straight track and a sense wire at the origin. The track is produced instantaneously, as if the particle was moving at infinite speed. The model is entirely two-dimensional.

We define the impact parameter of the track, b, as the shortest distance between the wire and the track (see Figure 6.1). This is also known as the distance of closest approach. The unique straight line having length b joining the wire and the track is in general perpendicular to the track itself. The point where this line intersects the track is known at the point of closest approach. We define D_n to be the distance of the n^{th} ionization event as measured along the track from the point of closest approach. The quantity x_n is the actual distance of the ionization event from the wire, which is the sum in quadrature of D_n and b: $x_n = \sqrt{D_n^2 + b^2}$. For the mathematical models constructed, it is also convenient to define D_0, which would be the distance of a fictitious 0^{th} cluster which always lies at the point of closest approach and for which $x_0 = b$.

Drift chamber electronics do not measure distances, but times of arrival for the clusters or signals from the ionization events, however since the time-to-distance relation is monotonic, there is a one-to-one relation between arrival times and distances, so the simple model is treated entirely in terms of distances.

The final simplification is done without loss of generality. Drift chamber electronics have no ability to distinguish ionization events produced before or after of the point of closest approach. In other words, ionization events with $D_n < 0$ look identical to those with $D_n > 0$. Assuming that the ionization events are independent and follow Poisson statistics, we may simply have all ionization events occur at $D_n > 0$ and double the overall ionization density.

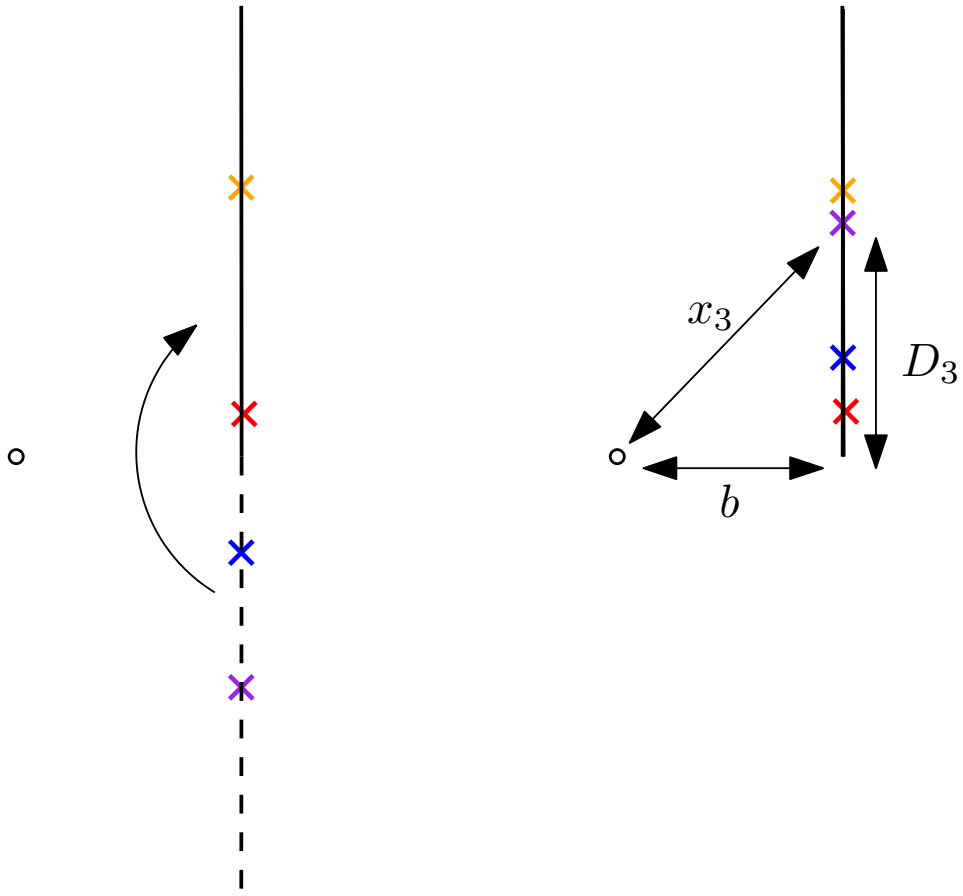


Figure 6.1: Illustration of the final simplification of the model described in Section 6.1, and the distances used in the model. b is the impact parameter of the track, and x_3 is the distance of the 3^{rd} ionization event from the wire. The figure on the left shows a track with ionization events before and after the point of closest approach. The figure on the right is the same track with all ionization events "rectified" to occur after the point of closest approach. The two models are mathematically equivalent, and we use the simpler right-hand model without loss of generality. Another way of stating this is that the model is symmetric under the operation $D_n \to -D_n$ for any n.

6.2 Distances Along the Track

First we define the probability density function for the distance of the fictitious 0^{th} cluster along the track, which always occurs at $D_0 = 0$:

$$f_0(D) = \delta(D). \tag{6.1}$$

$\delta(x)$ is the Dirac delta function. Since the ionization events follow Poisson statistics, the probability density function of the separation of two consecutive ionizations (Δ) follows an exponential distribution.

$$f_\Delta(\Delta) = \begin{cases} \rho e^{-\rho\Delta} & \Delta \geq 0 \\ 0 & \Delta < 0 \end{cases} \tag{6.2}$$

where ρ is double the density of ionization events along the track (explained at the end of Section 6.1).

With Equation 6.1 and Equation 6.2 we can construct the probability density function for the distance of the first cluster along the track, $D_1 = D_0 + \Delta$. The probability density function for the sum of two independent variables is simply the convolution of the two variables' probability density functions:

$$f_1(D) = \int_{-\infty}^{\infty} f_0(x) f_\Delta(D - x) dx \tag{6.3}$$

$$= f_\Delta(D) = \begin{cases} \rho e^{-\rho D} & D \geq 0 \\ 0 & D < 0 \end{cases}. \tag{6.4}$$

Similarly we construct the probability density function for the distance of the second cluster along the track, $D_2 = D_1 + \Delta$.

$$f_2(D) = \int_{-\infty}^{\infty} f_1(x) f_\Delta(D - x) dx \tag{6.5}$$

$$f_2(D) = \begin{cases} \int_0^D \rho^2 e^{-\rho x} e^{-\rho(D-x)} dx & D \geq 0 \\ 0 & D < 0 \end{cases} \tag{6.6}$$

$$f_2(D) = \begin{cases} \rho^2 D e^{-\rho D} & D \geq 0 \\ 0 & D < 0 \end{cases} \tag{6.7}$$

Since $\Delta \geq 0$, we can state that $D_n \geq D_{n-1} > 0$ for all $n > 0$, and $D_0 = 0$, to avoid overly cluttering page with the piecewise notation.

Similarly again we construct $D_3 = D_2 + \Delta$.

$$f_3(D) = \int_{-\infty}^{\infty} f_2(x) f_\Delta(D - x) dx \tag{6.8}$$

$$= \int_0^{D_3} \rho^3 x e^{-\rho x} e^{-\rho(D-x)} dx \tag{6.9}$$

$$= \rho^3 \frac{D^2}{2} e^{-\rho D} \tag{6.10}$$

By obvious pattern-matching and familiarity with iterated integrations, we can generalize to

$$f_n(D) = \rho \frac{(\rho D)^{n-1}}{(n-1)!} e^{-\rho D}. \tag{6.11}$$

This turns out to be an Erlang distribution, which is a special case of a Gamma distribution with integer shape parameter. These are frequently encountered when modelling waiting times and intervals of stochastic processes.

It is important to point out that these distributions for the different clusters are applicable when entire sorted sets of cluster arrival times are available. In other words, if one measures the arrival times of clusters in many tracks, and for each n makes a histogram of the n^{th} cluster in each track, then the distribution will be described by $f_n(D)$ in Equation 6.11 and Figure 6.2.

If instead one is searching for clusters in a track and has so far counted n clusters, the procedure to predict the arrival time of the next cluster the correct distribution is not $f_n(D)$ from Equation 6.11, but instead it would be $f_1(D - D_{n-1})$, the distribution of the next-to-come cluster given the previous one:

$$f_n(D) = f_1(D - D_{n-1}) \tag{6.12}$$

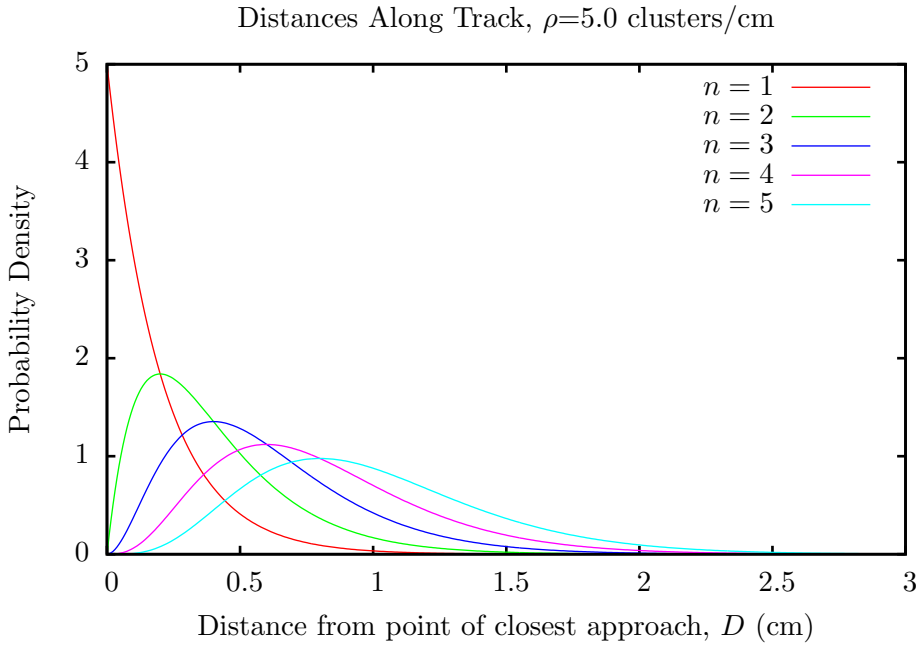


Figure 6.2: Probability distributions of ionization event distances along the track.

$$f_n(D) = \begin{cases} \rho e^{-\rho(D-D_{n-1})} & D \geq D_{n-1} \\ 0 & D < D_{n-1} \end{cases}. \tag{6.13}$$

This is because the ionization events are Poisson-distributed, and thus they are uncorrelated. The shape of the probability density function for the next uncorrelated event clearly cannot depend on the previous events. It only depends on D_{n-1} because Δ cannot be negative, so D_{n-1} serves as the lower limit and shift factor.

6.3 Distances From the Wire

In Section 6.2 we found the probability density function for the distance of the n^{th} ionization event from the point of closest approach along the track (D_n). The point of closest approach of the track is a distance b away from

the actual wire, a distance also called the impact parameter. We wish to know the probability density function for the distance of the n^{th} cluster from the wire itself.

Recalling that we are working in only two dimensions, the distance from the wire given a distance along the track is

$$x_n = \sqrt{D_n^2 + b^2}. \tag{6.14}$$

We exploit the conservation of probability to make the transformation:

$$f_n(D)dD = g_n(x)dx \tag{6.15}$$

$$f_n(D(x))dD(x) = \rho \frac{(\rho\sqrt{x^2-b^2})^{n-1}}{(n-1)!} e^{-\rho\sqrt{x^2-b^2}} \frac{x}{\sqrt{x^2-b^2}} dx \tag{6.16}$$

thus obtaining

$$g_n(x) = \frac{\rho^n}{(n-1)!} x(x^2-b^2)^{\frac{n-2}{2}} e^{-\rho\sqrt{x^2-b^2}}. \tag{6.17}$$

An interesting feature has emerged in the exponent $\frac{n-2}{n}$, so that the functional form drastically changes between $n = 1$, $n = 2$ and $n \geq 3$. For example in the limit of $x \to b^+$ (or $D \to 0^+$), we obtain

$$\lim_{x \to b^+} g_n(x) = \begin{cases} \infty & n = 1 \\ \rho^2 b & n = 2 \\ 0 & n \geq 3 \end{cases} \tag{6.18}$$

It would be interesting to find the critical points and maxima of $g_n(x)$, however this is complicated and the precise answers (involving roots of third-order polynomials) are not very enlightening. We can summarize as follows: g_1 has a maximum only at $x \to D^+$, where it diverges; g_n for $n \geq 3$ has a single finite maximum at finite x, the position of which grows slowly with n. g_2 has interesting maximal behaviour, strongly dependent on the product ρb. Specifically, if $\rho b \geq \frac{1}{2}$, the only maximum is at the boundary $x \to b^+$

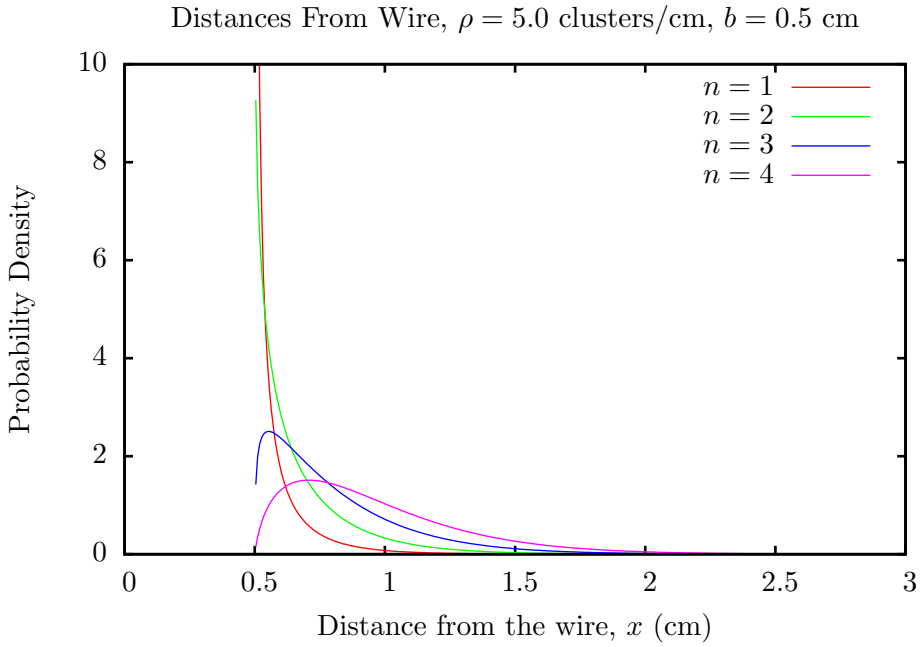


Figure 6.3: Probability distributions of ionization event distances from the wire.

where $g_2(D) = \rho^2 b$. If $\rho b < \frac{1}{2}$ then there are two critical points at

$$x_{\pm}^2 = \frac{1 \pm \sqrt{1 - (2\rho b)^2}}{2\rho^2} \tag{6.19}$$

where x_+ is generally a local maximum and x_- a local minimum. For certain values of ρ and b, this local maximum is also the global maximum, while for others the boundary value at $x \to b^+$ is the global maximum. Unfortunately again the precise expression for these cases is not enlightening.

Again it should be noted that Equation 6.17 was calculated assuming that none of the distances have yet been measured. This is the distribution of distances of clusters one would obtain if one simply made a histogram of the n^{th} measured cluster. In this sense they are "agnostic" probability density functions.

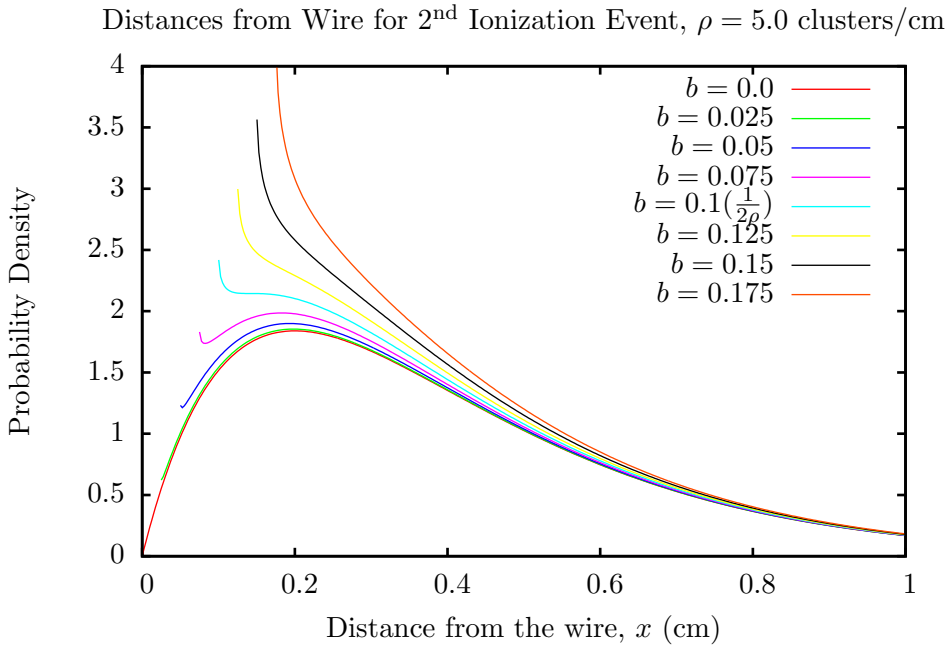


Figure 6.4: Probability distributions of second ionization event distances from the wire.

If the first n ionization events have already had their distances measured, then the probability density function for the distance of the $(n+1)^{\text{th}}$ cluster is not $g_{n+1}(x)$ given by Equation 6.17 but rather a form of $g_1(x)$ shifted by the position of the last measured cluster:

$$h_n(x) = \frac{x_n \rho e^{-\rho(\sqrt{x_n^2 - b^2} - \sqrt{x_{n-1}^2 - b^2})}}{\sqrt{x_n^2 - b^2}}. \tag{6.20}$$

This can be obtained by performing the conservation-of-probability calculation to convert Equation 6.13 to refer to x_n rather than D_n.

6.4 Bayesian Analysis

In Section 6.3 we found the probability distribution functions for the distance of the n^{th} ionization event from the wire, given a known impact parameter b. In a drift chamber experiment, the situation is reversed. There the arrival time of the electron cluster from the n^{th} ionization event is measured and converted into a distance using the time-to-distance relation for that drift chamber cell. From this distance we wish to know the most likely value for the unknown impact parameter. With Bayes' theorem, we can obtain the full probability density function for the impact parameter, given measured distances $\{x_n\}$.

Bayes' theorem states

$$P(\alpha|\beta;\sigma) = \frac{P(\alpha;\sigma)P(\beta|\alpha;\sigma)}{P(\beta;\sigma)} \tag{6.21}$$

where the left hand side is the posterior probability distribution: the probability of an event α given the observed state β and external assumptions σ. The first term in the numerator is the probability of event α given external assumptions σ irrespective of the observation β; this is also called the "prior probability" or just the "prior", because it refers to the state of the experiment before a measurement is done. The second term in the numerator is the probability of observing β given the state α and assumptions σ; this term is called the likelihood. The denominator is the probability of observing β under assumptions σ irrespective of the state α; the denominator is essentially a normalization constant, and is just the integral of the numerator over all possible values of α.

In our case, the left hand side is the probability distribution function for the impact parameter b given a set of observed ionization distances from the wire $\{x_n\}$. The first term in the numerator is the prior probability distribution for the impact parameter, while the second is the previously found probability distribution function for the distances of the ionization events from the wire, given in Equation 6.17. The denominator is the probability of getting a set of measured distances x_n irrespective of the true impact

parameter; as mentioned this normalization term is just the integral of the whole numerator over all possible values of the impact parameter.

Bayes' theorem is very general and in this case we apply it iteratively to build up the probability density function considering one ionization event at a time.

6.4.1 The First Cluster

We calculate the probability density function of the impact parameter b given the measured distance from the sense wire of the first ionization event x_1. Without any additional information, the prior probability density function for the impact parameter is uniform over all positive values. In principle this should be restricted to the actual dimensions of the drift chamber cell volume, though here we maintain the infinite-cell approximation for the calculation. The approximation should have little impact, since we only consider the first few clusters, which necessarily originate closest to the wire. This renders our prior probability distribution function unnormalizable by itself (since its integral diverges), but the denominator in Bayes' theorem fixes this problem automatically.

We get

$$P(b|x_1;\rho) = \frac{Cg_1(x_1)}{P(x_1;\rho)} \tag{6.22}$$

where C is the previously-mentioned unnormalizable term and g_1 is taken from Equation 6.17. Next we calculate the denominator, which is

$$P(x_1;\rho) = \int_0^\infty Cg_1(x_1)db \tag{6.23}$$

$$= \int_0^{x_1} C\rho x_1 (x_1^2 - b^2)^{-\frac{1}{2}} e^{-\rho\sqrt{x_1^2-b^2}} db. \tag{6.24}$$

We make the substitution $z = \sqrt{x_1^2 - b^2}$ to simplify the expression and use

the popular mathematics program Wolfram Alpha [70] to obtain

$$\begin{aligned} P(x_1;\rho) &= C\rho x_1 \int_0^{x_1} \frac{e^{-\rho z}}{x_1^2 - z^2} dz \\ &= C\rho x_1 \frac{2}{\pi}(I_0(\rho x_1) - L_0(\rho x_1)). \end{aligned} \tag{6.25}$$

where I_0 is the modified Bessel function of the first kind at 0^{th} order, and L_0 is the modified Struve function at 0^{th} order. The difference of the two special functions does not reduce to any other special function, but their Taylor series can be combined in a simple way (Appendix A.5).

Thus the final expression for the probability density function of the impact parameter b given a measured distance of the first ionization event from the wire x_1 is

$$P(b|x_1;\rho) = \frac{C\rho x_1 (x_1^2 - b^2)^{-\frac{1}{2}} e^{-\rho\sqrt{x_1^2 - b^2}}}{C\rho x_1 \frac{\pi}{2}(I_0(\rho x_1) - L_0(\rho x_1))} \tag{6.26}$$

$$= \frac{2e^{-\rho\sqrt{x_1^2 - b^2}}}{\pi\sqrt{x_1^2 - b^2}(I_0(\rho x_1) - L_0(\rho x_1))}. \tag{6.27}$$

As stated, the unnormalizable term C cancels out, which is a generic feature of Bayesian analysis. Note that the function is only defined for $b < x_1$, beyond this the probability density is zero.

A plot of Equation 6.27 is shown in Figure 6.5 with three different values of x_1. Unfortunately the numerical calculation becomes unstable with realistic values of ρ in the usual units (cm). Thus this and later plots use "dimensionless" units with $\rho = 1$. The reader may wish to interpret the units of ρ in clusters/millimetre and b in millimetres. In each case, the most likely value for the impact parameter is exactly the distance of the first cluster, and the shape of the function does not change. Indeed the only parameters in Equation 6.27 are the cluster density and the first cluster arrival time. The cluster density is considered fixed, but does vary slightly with the particle species, though that information is not available at the time of measurement.

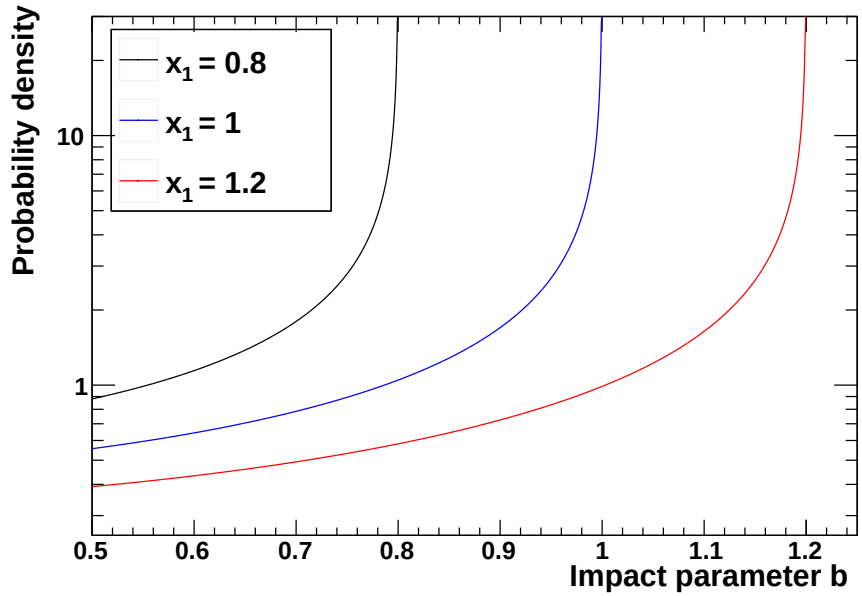


Figure 6.5: Plot of the probability density function of the impact parameter b when considering only a single cluster. The plot shows three different distances x_1 from the sense wire. The mean cluster density of $\rho = 1$.

6.5 The Second Cluster

To consider the information in the second cluster, we simply use Bayes' theorem again, but now the prior probability is the one shown in Equation 6.27 and the likelihood term is Equation 6.20 with $n = 2$. In equations:

$$P_2(b|x_2; x_1, \rho) = \frac{P_1(b|x_1, \rho)P(x_2|b, \rho)}{P(x_2|x_1, \rho)} \tag{6.28}$$

where $P_1(b|x_1, \rho)$ is exactly Equation 6.27 and

$$P(x_2|b, \rho) = \rho x_2 \frac{e^{-\rho(\sqrt{x_2^2-b^2}-\sqrt{x_1^2-b^2})}}{\sqrt{x_2^2 - b^2}}. \tag{6.29}$$

The denominator is again calculated by taking the integral of the whole

numerator. As with the unnormalizable term C in Section 6.4.1, the constant parts of the numerator cancel out with the denominator, since they factor out of the integral. We are left with

$$P_2(b|x_2; x_1, \rho) = \frac{I_2(b, \rho, x_1, x_2)}{\int_0^{x_1} I_2(b, \rho, x_1, x_2) db} \tag{6.30}$$

where

$$I_2(b, \rho, x_1, x_2,) = \frac{e^{-\rho\sqrt{x_2^2 - b^2}}}{\sqrt{x_1^2 - b^2}\sqrt{x_2^2 - b^2}}. \tag{6.31}$$

Unfortunately no closed form expression is available for the integral. Fortunately numerical methods are readily available. We use the default `GSLIntegrator` method through the ROOT wrapper functions, as they are recommended for general use and this function is not particularly nasty [2, 3].

The resulting function has three parameters: ρ which we consider fixed, x_1, and x_2. The plot in Figure 6.6 shows Equation 6.30 computed numerically with different values of x_2, alongside Equation 6.27 computed with the same x_1 value. The main feature is that the most likely value of b is never affected by the position of the second cluster, but that the *shape* of the distribution is affected. Depending on the position of the second cluster, the probability density for b is shifted either away from or towards the most likely value. If the second cluster is in an unlikely place (e.g., very close to x_1 or very far away from the wire) then the width of the distribution is enhanced. If the second cluster is in a more usual place (e.g., within $\sim 1/\rho$ of x_1) then the peak at the most likely value is sharper. This translates directly into increased or decreased confidence in the measured value of the impact parameter.

6.6 More Clusters

The process to include even more clusters is the same as in Section 6.5: the prior is the result of the previous calculation, the likelihood is Equation 6.20 with the appropriate n, and the denominator is the integral of the numerator over b from 0 to x_1.

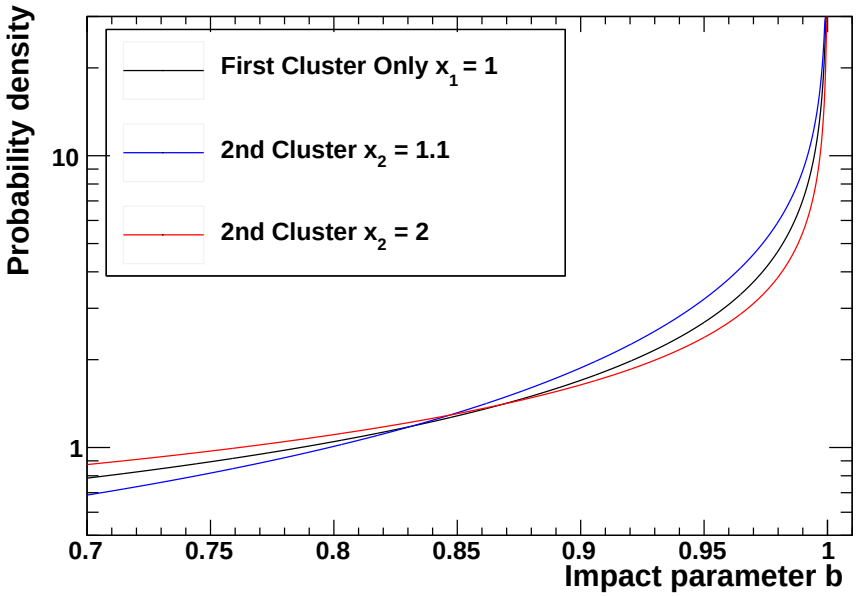


Figure 6.6: Plot of the probability density function of the impact parameter b when considering the first two clusters. The black curve uses the result from Equation 6.27 which only considers the first cluster time. The other two curves use the result from Equation 6.30 considering also the second cluster, at two different 2^{nd} cluster times. The mean cluster density is $\rho = 1$.

The calculation is done numerically again, and the function to be integrated is only slightly more complex:

$$P_3(b|x_3;\rho,x_1,x_2) = \frac{J_3(b,\rho,x_1,x_2,x_3)}{\int_0^{x_1} J_3(b,\rho,x_1,x_2,x_3)db} \tag{6.32}$$

where

$$J_3(b,\rho,x_1,x_2,x_3) = \frac{e^{-\rho\sqrt{x_3^2-b^2}}}{\sqrt{x_3^2-b^2}\sqrt{x_2^2-b^2}\sqrt{x_1^2-b^2}}. \tag{6.33}$$

One can see a pattern emerging, where

$$J_N(b, \rho, x_1, ..., x_N) = \frac{e^{-\rho\sqrt{x_N^2 - b^2}}}{\prod_{k=1}^{N} \sqrt{x_k^2 - b^2}}. \tag{6.34}$$

The result is shown in Figure 6.7 for fixed values of ρ, x_1, and x_2 with two different values of x_3. The corresponding P_1 and P_2 curves are again shown for comparison. It is unfortunately difficult to see the differences between the curves since they are all somewhat similar. The general trend continues from the consideration of the second cluster: the extra cluster does not change the most likely value, but can affect the width of the probability distribution. If the third cluster is in an unlikely place, the width is increased, but if it is consistent with the average cluster density, the width is reduced. The magnitude of the effect from considering the third cluster is on the same order as the change from considering the second cluster.

6.7 Summary

The model presented in this chapter is much too simplified and abstract to make quantitative predictions that are useful for a real drift chamber prototype. The simplicity however strengthens the general qualitative results that directly motivate the study of cluster counting for tracking purposes.

The main result is that indeed there is something to be gained in considering clusters beyond just the first one. While the most likely value of the impact parameter given a set of cluster distances is only dependent on the first cluster's position, the shape of the posterior distribution is affected by the distances of the later clusters. The generic result is that the width of the posterior distribution for the impact parameter b is enhanced when the later clusters are in extremely unlikely places given the average cluster density. If the later clusters are in likely places, the width of the distribution is reduced. The width of the distribution is directly related to our confidence in the measurement of b.

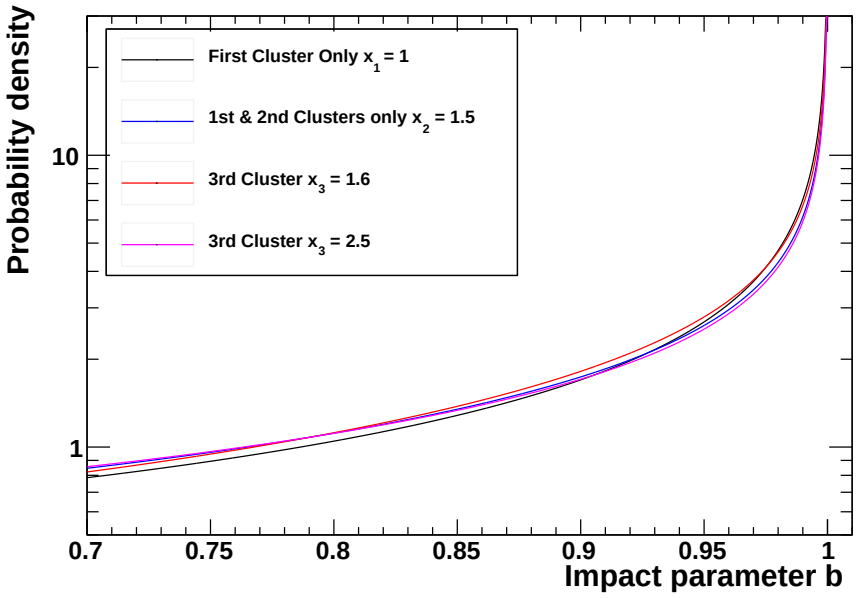


Figure 6.7: Plot of the probability density function of the impact parameter b with a mean cluster density when considering the first three clusters. The black curve uses the result from Equation 6.27 which only considers the first cluster time, the blue curve uses the result from the first two clusters. The red and magenta curves use the information from the first three clusters at two different 3^{rd} cluster times. The average cluster density is $\rho = 1$.

Chapter 7

Tracking

7.1 Overview

This chapter deals with the analysis of data taken at TRIUMF in 2012 using the Italian prototype proto 2. All the coding and analysis presented here was done by me, including the implementation of traditional and cluster-based tracking algorithms. As with the particle identification study presented in Chapter 4, the cluster-based technique is used in combination with the traditional method to try to get better performance.

The traditional tracking procedure is as follows. First, the arrival times of the various signals on the sense wires are determined (Section 7.2). Then, time-to-distance relations are produced using simulations (Section 7.3). These are used with the arrival times to determine how far the track passed from the various sense wires, and subsequently the track parameters are determined (Section 7.4). Using the self-consistency of the tracks, the time-to-distance relations are improved (Section 7.5) and the tracks parameters are re-evaluated. This refinement process iterates two times, after which the final resolution of the prototype drift chamber can be measured (Section 7.8).

The track parameters are determined using the information from all the cells that have signals except one. A special cell (number 12 in Figure 5.1) is excluded from the computation. This cell is chosen because it is central in the detector. A statistical analysis is performed of the difference between

the track distance from cell 12's sense wire according to the fit using the other cells and from the information from cell 12 alone (Section 7.5). This is used to generate empirical corrections to the original time-to-distance relations that were computed theoretically from simulations. With these corrections in hand, we re-interpret the arrival times using the new time-to-distance relations and re-do the track fit procedure. This is iterated one more time to obtain 2^{nd}-order corrected tracks. In the statistical analysis of the difference between the information from cell 12 and the information from the other cells, the standard deviation gives us a measure of the chamber's resolution. This is the final result of the traditional tracking: the resolution of the drift chamber.

Next we implement a cluster-counting algorithm (Section 7.6) and try to use the additional information to improve tracking. Here we don't use just the number of clusters, as in particle identification, but the individual arrival times of each cluster. Similar to the traditional time-to-distance relations, we build up histograms that characterize the distributions of cluster times at different track distances. Then we re-interpret the cluster times by asking "for which track distance is this set of cluster times most likely?" Similar to the particle identification study in Chapter 4 (Section 4.5.3), we form a combined likelihood using the traditional information and the cluster times (Section 7.7). The final measured scatter between the track distance and the distance inferred from cell 12's signals (traditional and cluster information) gives us the new resolution.

What we found is that the resolutions of the drift chamber using traditional tracking and the combined traditional and cluster-counting technique are equivalent. If there is any difference, it is not measurable given the uncertainties in our results (Section 7.8).

7.2 Measuring the Arrival Times

Measuring the arrival time of a signal is non-trivial, because there are many fluctuating factors to account for. The amplitude of the initial cluster can very tremendously, because of the statistical nature of the ionization

avalanche near the wire. Similarly, the baseline voltage (the relative 0 V point) can drift slowly over time, and the signal itself has noise.

To deal with the drifting baseline voltage and the regular signal noise, we first note that our trigger system is set up so that no real pulses are ever recorded in the first ~ 200 ns of each signal (see Figure 7.1). The baseline voltage drift is slow, so we treat it as constant for each event, but it may drift between events. Large changes are actually only observed between entire datasets that were taken after changing voltages or other settings. Thus we take the average voltage over the first 100 ns of the signal as the baseline. After this, all voltages are referred to relative to this baseline, not relative to 0 V. In the same 100 ns region, we calculate the root-mean-square (RMS) deviation from the mean (i.e., the standard deviation). We use this RMS voltage as a measure of the regular noise on the signal.

Since the signal pulses from charge clusters have widely varying amplitudes, we must carefully select a threshold above which we recognize a signal. If the threshold is too high, we will miss pulses that have small amplitudes and ruin the time measurement. If the threshold is too low, we might accidentally identify a random noise spike as the first cluster, again ruining the measurement.

To resolve this, we use thresholds that are proportional to the RMS noise measured in the initial parts of the signal. It was found that thresholds of 4 and $10\,\sigma$ were low enough to catch nearly all signals, and $10\,\sigma$ is clearly high enough to almost never trigger on a noise spike. With this threshold, only one out of every 6.6×10^{22} voltage samples should ever exceed the threshold due to noise, assuming a Gaussian model.

We use two thresholds because a single threshold gives a biased measurement dependent on the amplitude of the signal, since the threshold crossing generally occurs somewhere on the leading edge of the signal. The two threshold crossings are used as points to do a straight-line extrapolation back to the baseline voltage. The point where the extrapolated line crosses the baseline is taken to be the arrival time (see Figure 7.1).

Recall that no real signal pulses arrive for the first ~ 200 ns of each signal. We wish to identify precisely the earliest possible time that signal pulses can

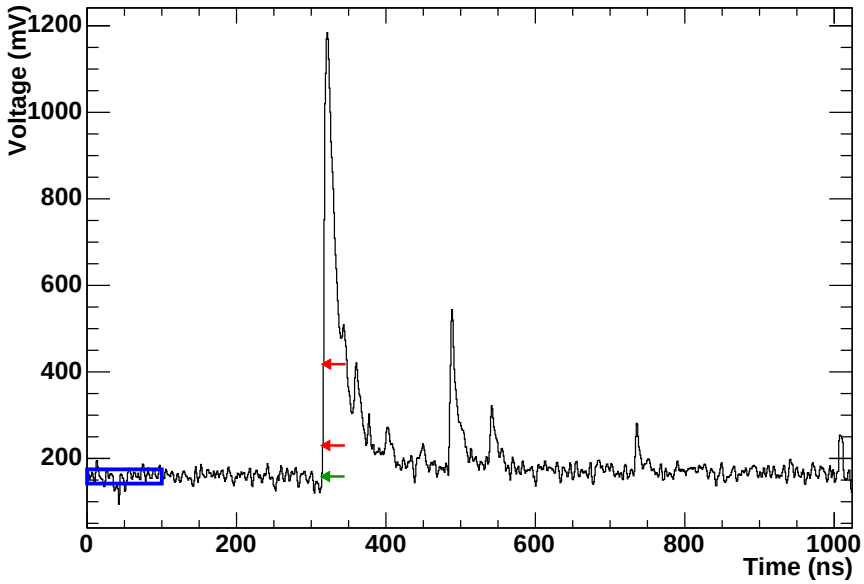


Figure 7.1: Example signal from a cell in proto 2. The small box over the first 100 ns shows the baselining and noise measurement region. The height of the box is $1\,\sigma$ of noise. The three arrows show the upper and lower threshold crossings, and the extrapolated arrival time at the baseline.

arrive. This time would correspond to clusters ionization events immediately adjacent to the sense wire. To do this, we build up a histogram of all the arrival times in a data set. At first, individual histograms were collected for each wire and for different detector configurations, but it was realized that they were all identical. Figure 7.2 shows the distribution of arrival times for all wires in a given data set. The sharp leading edge corresponds to those ionization events that occurred near the sense wire. The other structure relates to the cell geometry. The population above the "knee" at ~ 450 ns corresponds to tracks that do not cross the full width of the cell.

To obtain a precise measurement of the leading edge of the arrival time distribution, we take the bin-by-bin derivative and fit a Gaussian function

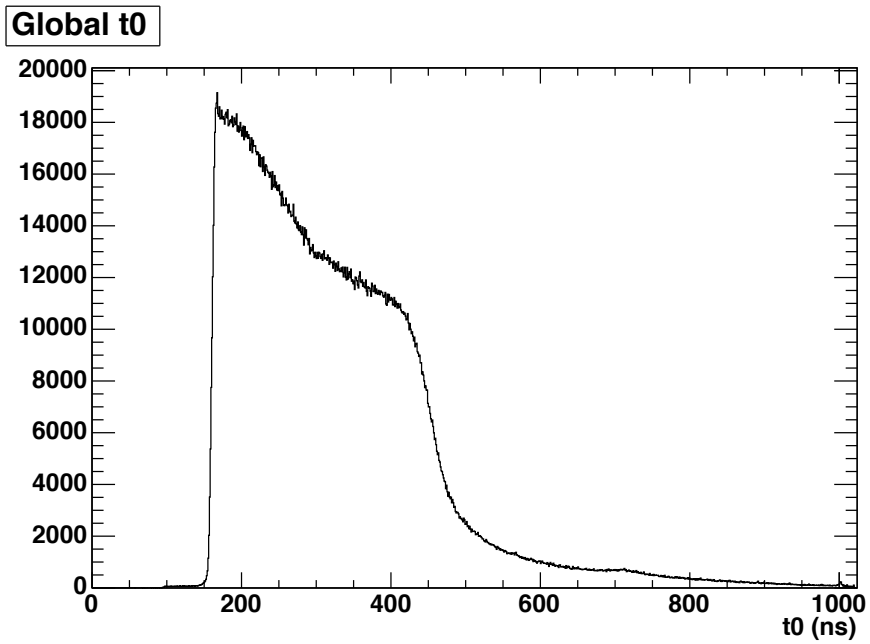


Figure 7.2: Raw arrival time distribution for all signals in all wires.

(see Figure 7.3). The central value of this Gaussian distribution corresponds to the point on the leading edge with the greatest slope, and we take this as the global 0 ns time reference. This is a standard procedure when calibrating drift chamber signal arrival times [71].

In a later study we tried deviating from this time reference by a few nanoseconds in each direction, because it was realized that possibly the real 0 ns time might correspond to some other point on the leading edge, not just the point with the highest slope. Eventually a slightly different value (164.229 ns) was chosen as this was found to slightly improve the strange distribution of track distances observed in the data (see Section 7.4).

With a global time reference, we can now shift all the arrival time measurements so that they are relative to this value. This way signals corresponding to tracks essentially hitting the sense wire will have arrival times of ~ 0 ns, and signals from tracks further away from the wire will have arrival

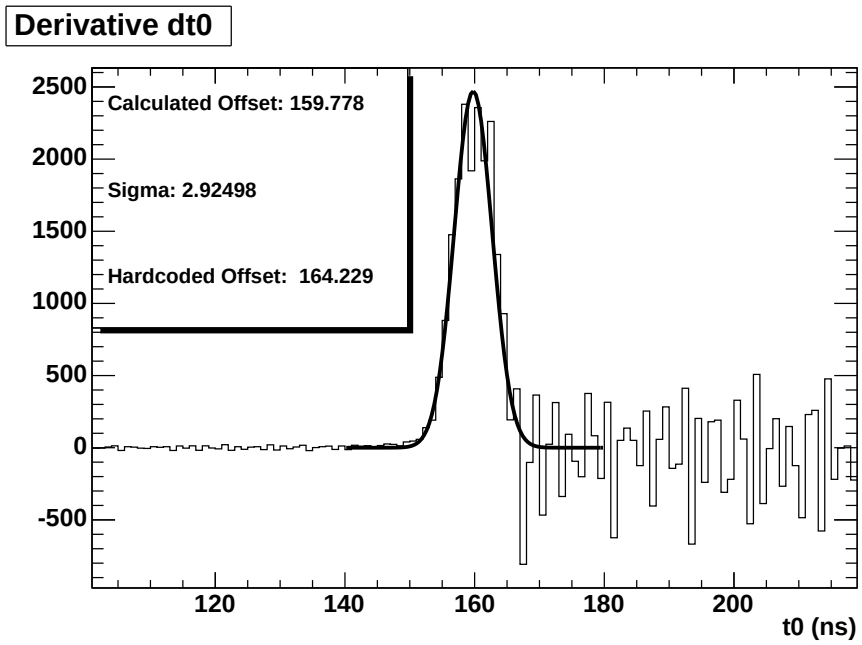


Figure 7.3: Derivative of arrival time distribution for all signals in all wires. This shows only the part on the leading edge, and a Gaussian fit. The mean and standard deviations from the Gaussian fit are labelled "Calculated Offset" and "Sigma". The "Hardcoded Offset" is the actual time offset used in the calculations.

times up to $\sim 500\,\text{ns}$. The next step is to convert these shifted arrival times into distances using the time-to-distance relations.

7.3 Garfield Time-to-Distance Relations

Time-to-distance relations are basically mathematical functions that convert the arrival time of a signal pulse into the distance of closest approach. That distance is measured between the point of closest approach of the high-energy charged particle crossing the drift chamber and the sense wire (see Figure 7.4). Because of the statistical nature of the ionizations along the track, the actual point at which the first recorded ionization event was pro-

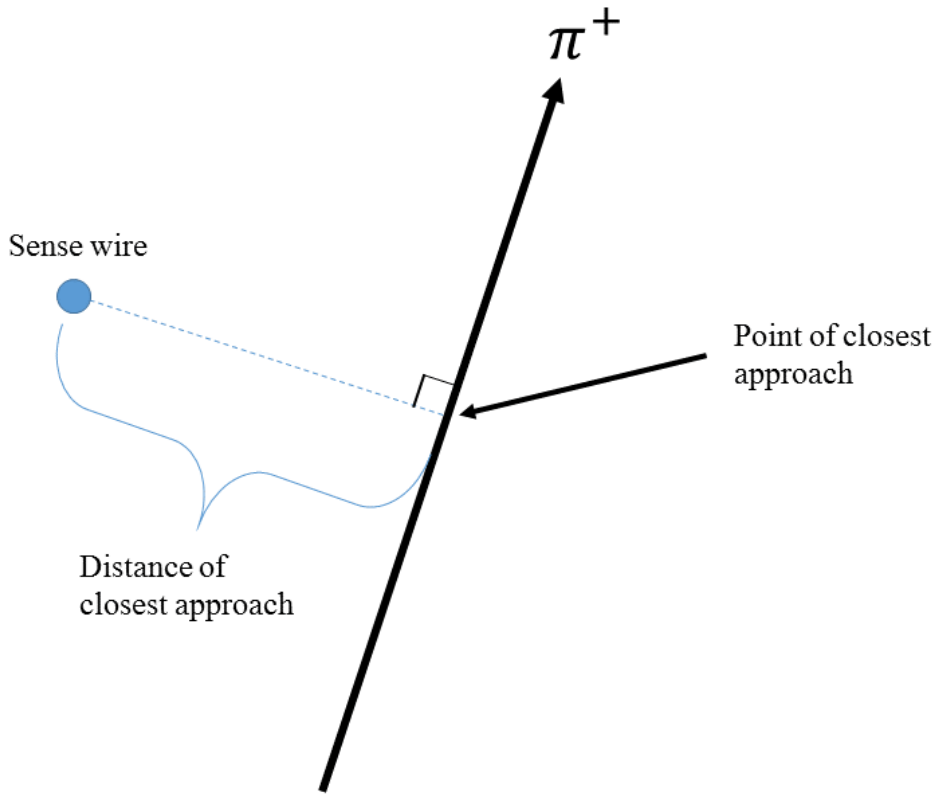


Figure 7.4: Illustration of the terms point and distance of closest approach. The distance of closest approach is also called the impact parameter.

duced may not be the point of closest approach. Thus the time-to-distance relation converts an arrival time into the *most likely* distance for the track. The uncertainty in the measurement is considered when making further calculations.

We use the computer program Garfield [4] to simulate our prototype and produce theoretical time-to-distance relations. Garfield is a commonly used tool for simulating gaseous ionization detectors, including drift chambers. It was used in the design of the prototypes described in this thesis by calculating required wire voltages under various conditions (e.g., with different

gas mixtures).

Garfield simulates individual track events in the chamber. The tracks all have a chosen angle of incidence, and a chosen progression of distances from the wire. For each sense wire, a histogram is built up of the arrival time distribution for small intervals of distances. Garfield prints the mean and RMS of the histogram, along with tables of bin centres and contents. The printed Garfield log file is used to reconstruct the histograms in ROOT [3]. If the histograms were reasonably Gaussian, we could just use the printed mean and RMS values. For tracks near the sense wire the distribution is noticeably non-Gaussian, so the mean and RMS values do not properly describe the distribution. Thus we use the full histograms and fit the arrival time distributions with Novosibirsk functions (Figure 7.5). The peak location is used instead of the histogram mean, and the full-width at half-maximum is used to obtain the resolution instead of the RMS. The Novosibirsk function and the choice of the width parameter is described in detail in Appendix A.1.

The process of using simulated tracks at known distances to determine the distribution of arrival times is very similar to what would be done with an external tracker [72, 73]. If we had an external tracking device (e.g., coincidence counters with narrow windows, or another gaseous detector), we could generate these distributions empirically with real particles from a test beam or from cosmic rays.

The reader may have noticed that although we require a function that converts an arrival time into a track distance, Garfield has provided arrival time distributions for given track distances. The same problem would be encountered if we used empirical distributions from an external tracker. In short, Garfield provides distance-to-time relations, but we require the inverse. Fortunately the relations are monotonic (i.e., the electrons never drift away from the sense wires), so the inversion is mostly a trivial flipping of the x and y axes. An example of a distance-to-time relation and resolution plot are shown in Figure 7.6.

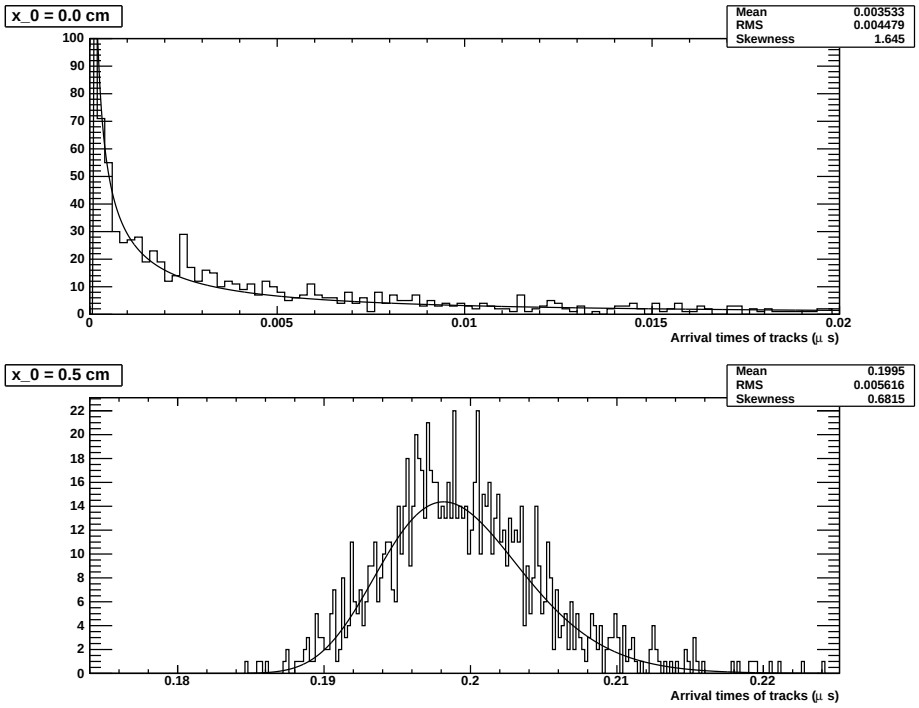


Figure 7.5: These two plots show the Novosibirsk fits to the arrival time distributions for 0° tracks at two different distances from a sense wire.

7.3.1 Interpolation

Garfield only produces histograms of the arrival times for specific requested track positions. We use an interpolation technique to evaluate the time-distance relation and resolution functions in between the track positions considered by Garfield.

It is crucial for the interpolation technique to guarantee monotonicity between the points, because the original data are the expected signal arrival times and uncertainty as a function of track distance. In the analysis we will need the inverse: the track distance and its resolution as a function of signal arrival time. Thus the initial interpolated functions and their derivatives will be inverted and manipulated. If our interpolation functions are non-

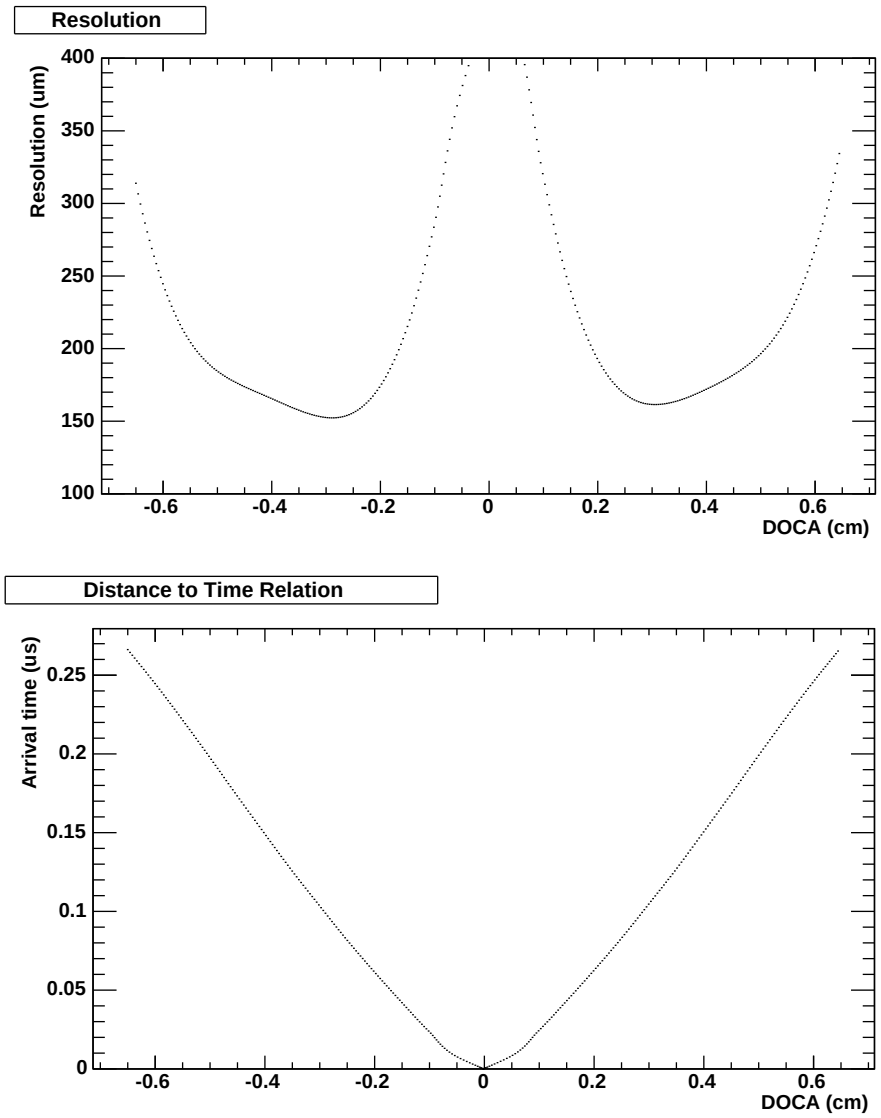


Figure 7.6: Resolution and distance-to-time relation calculated by Garfield. This is for the intended Super$B$ gas of 90 : 10 helium-isobutane mixture. The cell boundaries are at $\pm 0.7\,\mathrm{cm}$. Each point corresponds to a set of Garfield-simulated tracks at a given track distance. The most likely arrival time of signal pulses and the widths of the distributions are the y-values of the lower and upper plots, respectively.

monotonic between the data points, the derivatives will go to zero and this will cause nonphysical spikes to plus or minus infinity in the final function.

Unfortunately the ROOT software only comes with a one interpolation type that guarantees monotonicity between the points: a linear interpolation. We wish to use a more advanced technique, where monotonicity between the points is guaranteed, but which uses curves instead of straight line segments. This is provided by using the so-called Steffen interpolation technique [74]. Monotonicity is guaranteed by sacrificing the smoothness of higher-order derivatives, but Steffen's method maintains continuity of the function itself and its first derivative. The technique also avoids problems present in other interpolation methods such as instability, where a small change in a single point can produce large changes in the interpolation function. A plot showing the interpolations using several techniques is shown in Figure 7.7.

I implemented Steffen's interpolation method using the C programming language as a contribution the GNU Scientific Library (GSL) [2]. It will be available to the public sometime in 2015. Minor modifications had to be made to the ROOT source code to use the new method, which will be contributed to the ROOT project once the new GSL version is released.

7.4 Track Fitting

With the time-to-distance relations in hand, we can convert the signal arrival times into distances of closest approach. Although we know the most likely distance of closest approach (and the uncertainty), we do not yet know the actual angle of the track. Thus we effectively have around each sense wire a circular locus of points to which the track should be tangent.

With the signals from multiple cells, the track should be tangent to all the circles at once, and this is how the ambiguity is resolved (see Figure 7.8). In reality, a track tangent to all the defined circles will be impossible to obtain, because the circles are at the *most likely* distance of closest approach, but the actual track distance can fluctuate.

Given the most likely distances x_i and their uncertainties Δx_i, we use

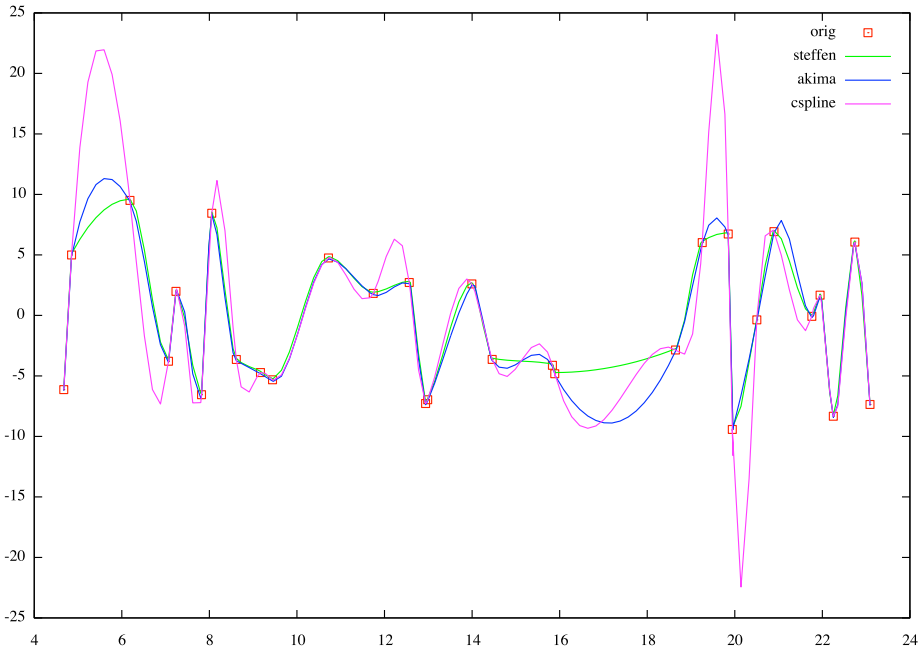


Figure 7.7: A comparison of a few interpolation methods using randomly generated data points. The data points are the red squares. Large oscillations are apparent in the cubic spline, and less severe ones are also visible in the Akima technique, which is already available in GSL and ROOT. The Steffen method preserves monotonicity.

a computer program to minimize the well-known χ^2 function to find the best-fitting track:

$$\chi^2 = \sum_i \frac{1}{N-2} \left(\frac{D_i - x_i}{\Delta x_i} \right)^2 . \tag{7.1}$$

Here the sum over i is over the N active cells from which we extracted arrival times and determined distances of closest approach. D_i is the distance of closest approach for the proposed track from the wire in cell i. In our tests the tracks are straight lines in two dimensions, since there is no magnetic field to curve the tracks and we have no information about the z-coordinate, so x_t depends on two track parameters: θ, the angle of the track relative

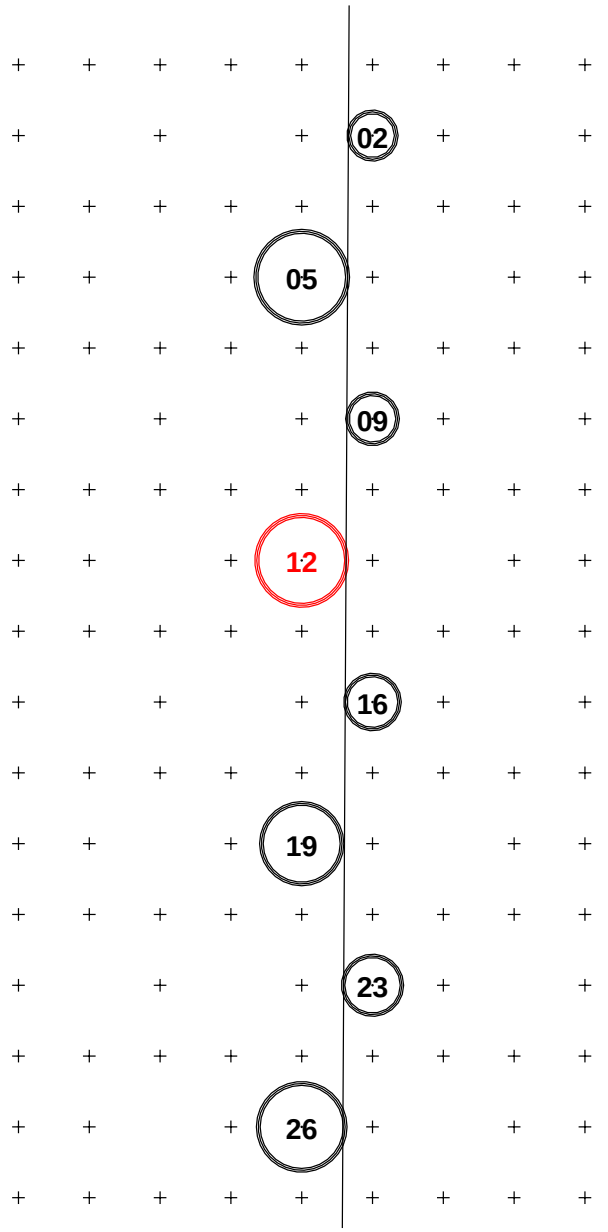


Figure 7.8: Circles to which the track should be tangent, determined from the signal arrival times and the time-to-distance relations. Each cell actually has three concentric circles: the middle corresponds to the most likely distance, while the inner and outer correspond to that distance plus and minus one standard deviation, respectively.

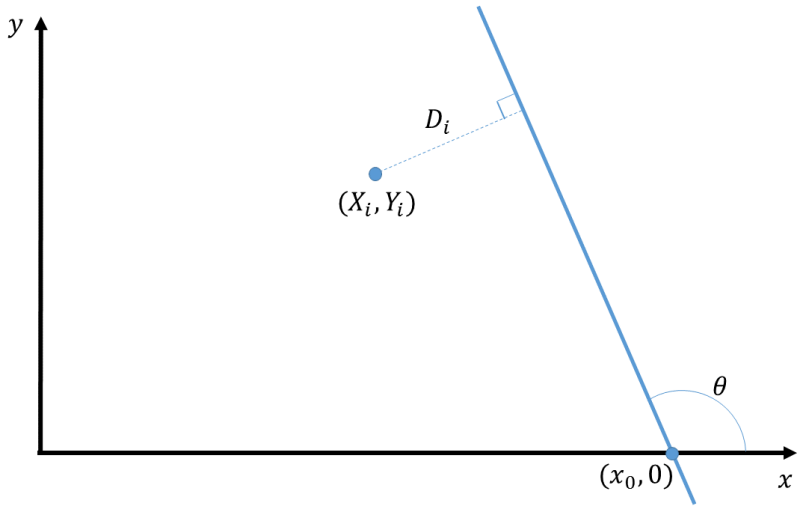


Figure 7.9: Diagram showing track parameters. See text for mathematical definitions.

to the horizontal, and x_0, the x-coordinate of the track at $y = 0$. The dependence is

$$D_i(x_0, \theta) = Y_i \cos\theta + (x_0 - X_i) \sin\theta, \tag{7.2}$$

where X_i and Y_i are the coordinates of the sense wire in cell i. The derivation can be found in Appendix A.2, and an illustration in Figure 7.9.

The best-fitting track corresponds to that choice of x_0 and θ which minimize χ^2 as defined in Equation 7.1. The uncertainties in these parameters are found from the derivative of the χ^2 function at the minimum. The minimization and returning of the uncertainties is done by the `Minuit2` routine in ROOT [3].

7.4.1 Track Initial Parameters

Reasonable initial track parameters must be chosen in order for the fitting program to work properly. For the track angle θ, we note that all of our data runs were taken with the beam at a specific angle to the chamber: 0,

22, or -22 degrees. Since we only analyze data sets with a single alignment angle at a time, we simply use the angle appropriate for the selected data set. In an experiment with tracks at all possible angles, a simple heuristic algorithm would be sufficient to provide an initial guess.

The initial estimate of the parameter x_0 is a function of the sense wire coordinates of each active cell, and of the initial track angle estimate. If the sense wire coordinates are (X_i, Y_i) for cell i, the initial x_0 value is

$$x_0 = \frac{1}{W} \sum_{\text{active}} \frac{1}{\sigma_i} \left(X_i - \frac{Y_i}{\tan\theta} \right) \tag{7.3}$$

where σ_i is the resolution at the location of the track in each cell, used as a weight, and W is sum of these weights. The sum is only done over active cells.

In the case of vertical tracks, the denominator $\tan\theta$ is infinity, so the expression is just the weighted average of the X_i values. For non-vertical tracks, this is the weighted average of the x-coordinates of the sense wires translated down to $y = 0$ along a line parallel to the track. This is illustrated in Figures 7.10 and 7.11 below.

7.4.2 Minimization Troubles

Although a robust algorithm was used for track minimization, the 2-dimensional parameter space that must be searched for the best track has multiple local minima. These are parameter choices that appear to be well-fitting tracks, but that are not the best available. The standard algorithm is not very good at choosing between minima, and sometimes returns less-than-optimal tracks. To try to alleviate this, I produced "heat" maps like the one shown in Figure 7.12. Unfortunately I originally used the default ROOT colour scheme which uses a rainbow spectrum. The rainbow spectrum is well-known to be terrible for human recognition of features [75]. With the rainbow map, the local minima were invisible, and the minimization program gave terrible results. A comparison of the two colour schemes can be seen in Figures 7.12 and 7.13.

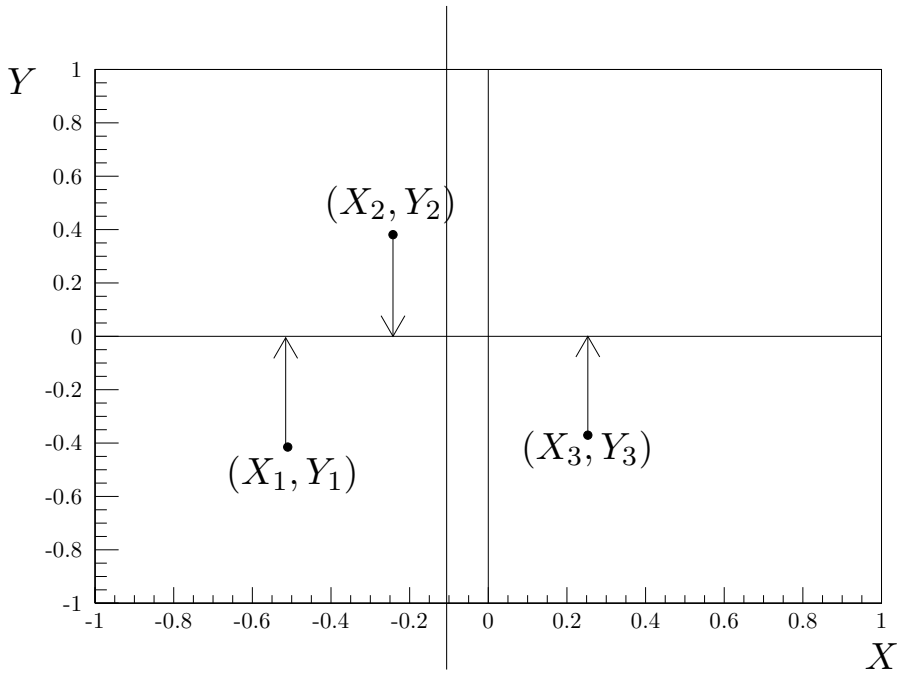


Figure 7.10: Schematic of x_0 calculation for vertical tracks. x_0 is the average of the X coordinates of the active sense wires.

An example of two tracks, one of which is a local minimum in the track-parameter space can be seen in Figure 7.14. The local minima arise because of the geometry of the cells and the ambiguity about which side of the cells the track lies. This results in local minima that tend to be spaced somewhat predictably on the x_0-θ plane. Thus once the regular algorithm has found a minimum, additional minimizations are initiated $\pm 0.7\,\mathrm{cm}$ and ± 0.15 radians away. In cases where these new minimizations find new minima, their χ^2 values are compared and the best one is taken. This generally finds the global minimum, and is a simple way of finding the global minimum in a complicated function.

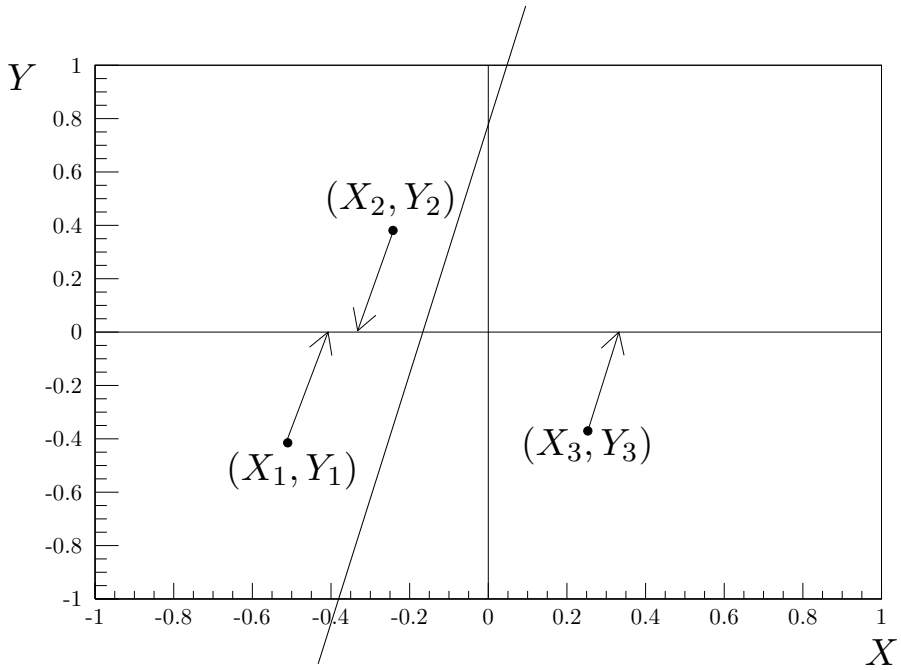


Figure 7.11: Schematic of x_0 calculation for non-vertical tracks. x_0 is found by translating the coordinates of the active sense wires parallel to the track until they reach $Y = 0$, then taking the average of the resulting X coordinates.

7.4.3 Track Parameter Uncertainties

We would like the track fitting to have a negligible contribution to the overall resolution of the drift chamber. D is the distance of closest approach of the best-fit track to a reference cell's sense wire, and ΔD is the uncertainty in that value. Both are functions of the track parameters x_0 and θ and their respective uncertainties. We use cell 12 as the reference cell, because it is in the centre of the prototype and is also used in the next section to improve the time-to-distance relations (Section 7.5).

If ΔD is large compared to the overall tracking resolution, then the track fitting algorithm itself or the geometry of the chamber needs to be modified to improve the resolution. If ΔD is small compared to the overall

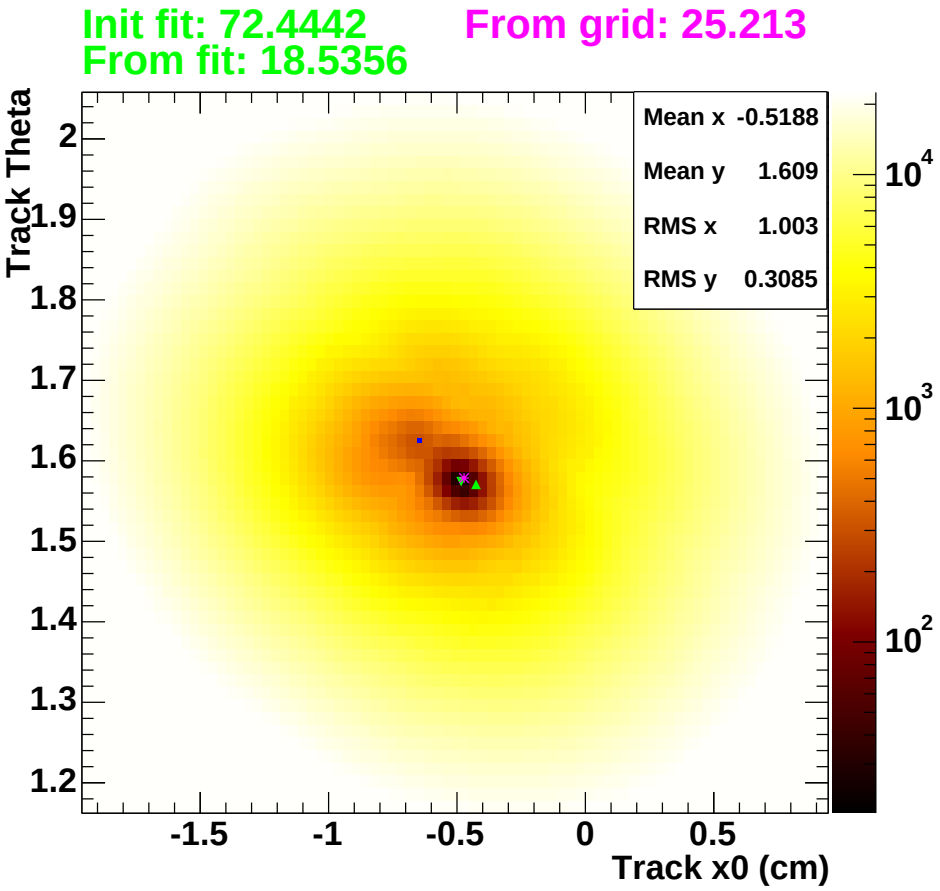


Figure 7.12: Heat-map of the function to be minimized to find the best track. The darker the colour, the better the track. The colour scale is logarithmic. The upwards- and downwards-pointing green triangles are the initial fit parameters and minimum found by the algorithm, respectively. The asterisk symbol is a crude minimum found by a regular grid-search algorithm. The function values at those three points are printed above the plot. A local minimum that is not the global minimum is indicated by a small blue square in a cloud of dark that is slightly separated from the true minimum.

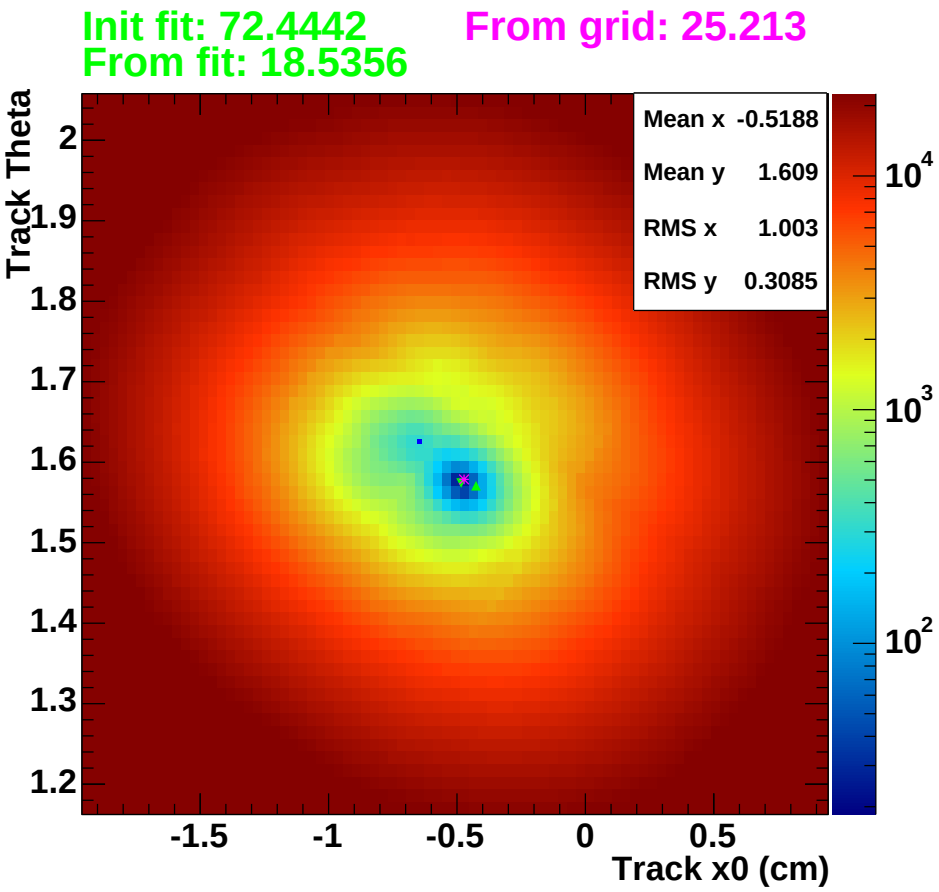


Figure 7.13: The exact same function as plotted in Figure 7.12 but using the default ROOT rainbow colour scheme. The slight "clouds" visible in that figure (local minima) are completely obscured here.

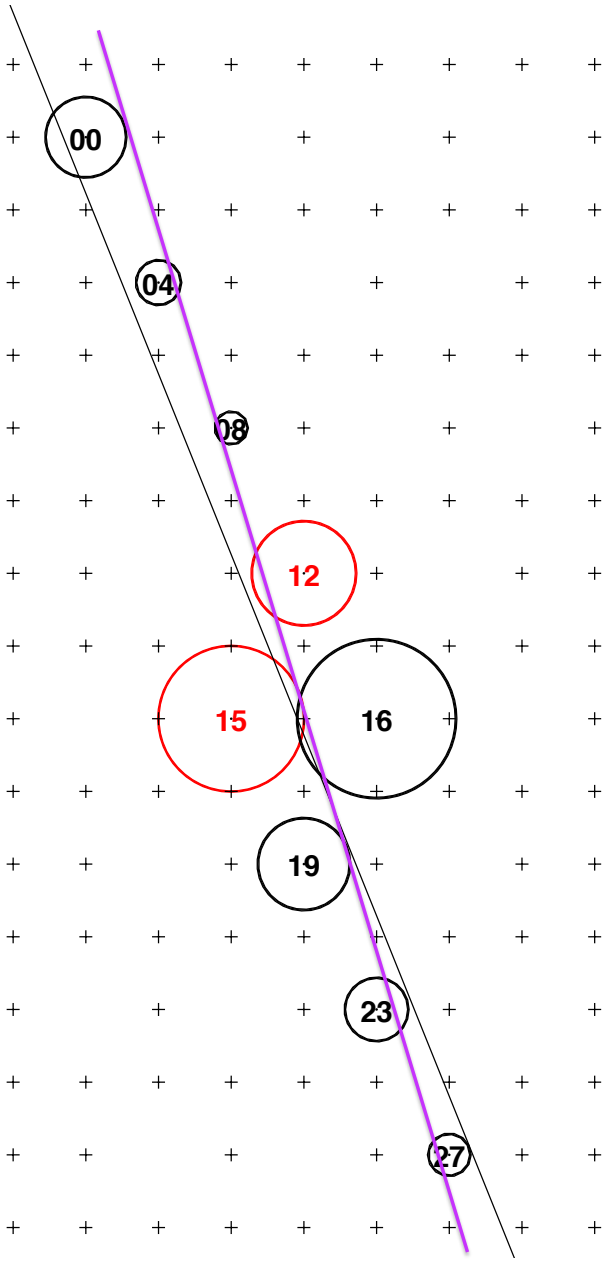


Figure 7.14: Illustration of two potential track candidates. The thin track in black is the local minimum found by the algorithm, while the thicker line in purple is a better minimum found by a simple grid-search.

resolution, then we can conclude that most of the uncertainty comes from the single-cell resolution. In this case, we can attempt to improve the single-cell resolution through more accurate time-to-distance relations and through cluster counting. We expect ΔD to be small from geometric arguments: the track parameters are derived from N measurements, while the single-cell resolution clearly comes from only one measurement. While the N measurements are not all independent, we can roughly expect a ratio of $\sim 1 : \sqrt{N}$ between ΔD and the single-cell resolution. N is the number of cells participating in the track fitting, which in our prototype is ~ 8 for tracks crossing the whole detector.

Strictly speaking, the parameter uncertainties extracted from the χ^2 minimization procedure are only a good choice if the x_i values and their uncertainties follow Gaussian distributions, which they do not when the track is very close to the sense wire. Nevertheless they are used in order to avoid further complicating the process, and since the final resolution measurement that we wish to obtain is not greatly influenced by the individual track parameter uncertainties. The proper method would be to replace the χ^2 minimization with a maximum-likelihood method that uses the Novosibirsk function, then see what variation of parameters would yield a change of $+1/2$ in the likelihood to find the 1σ uncertainty.

From the slightly-improperly calculated uncertainties in x_0 and θ, we can calculate the uncertainty in ΔD. The equation for it can be found in Appendix A.2. ΔD is the uncertainty in D from the track fitting procedure itself, i.e., it is the shallowness of the "bowl" in the 2D χ^2 function that is minimized to find the track. Typical values for ΔD are $\sim 40\,\mu$m, while the overall resolution of the drift chamber is $\sim 200\,\mu$m (shown in Figure 7.16). While this means that ΔD is not negligible, the uncertainty is certainly dominated by the single-cell resolution.

7.5 Iterative Refinement and Resolution Measurement

Once the track parameters and uncertainties have been determined, we can quantify the performance of the detector. Normally this would be done using some kind of external tracking system, to compare the drift chamber's measurement with a known-good measurement. This is similar to what was done in Section 4.5 by comparing the drift chamber's measurement of the energy loss and cluster count to an external time-of-flight system's identification of the particles. In the case of proto 2, no external tracker was available. There was an external tracker built, but it stopped working shortly before proto 2 was shipped to TRIUMF for the beam test.

Fortunately we have another tool at our disposal: self-consistency of the tracks. The measurement of the track parameters using N single-cell signals should be about $1/\sqrt{N}$ better than the single-cell resolution itself. The trick is to perform the track fit using only 7 layers of cells. This provides a "known good" measurement of the track parameters against which we can compare the measurement from the single excluded cell. Though the uncertainty in the track fit using the 7 layers is non-negligible, the dominant uncertainty will still be the single-cell resolution. This is very similar to the process used to calibrate the BaBar drift chamber [38]. The cell in proto 2 chosen for this process is number 12, which is the most-central cell in the chamber. See Figure 5.1 for the wire layout.

To measure the resolution using this method, we make histograms of the difference between the most-likely track distance according to cell 12 and the distance of the track from cell 12 according to the fit using the other 7 layers (Figure 7.15). We exclude tracks that did not cross cell 12. These histograms are made for small intervals of track distances (i.e., distances according to the 7-layer fit). These histograms generally have a Gaussian shape, so we take the mean and standard deviations of Gaussian fits. The standard deviation is our measurement of the resolution of the chamber at that distance from the cell. The mean is a measure of the bias of our time-to-distance relation compared to the "known good" measurements from the

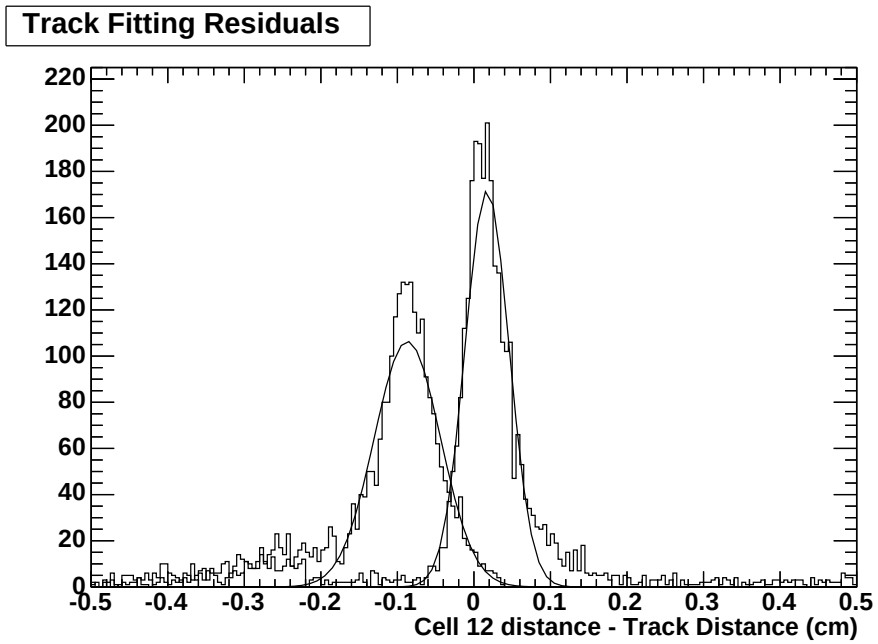


Figure 7.15: Two tracking residual distributions superimposed. The peak on the left is for tracks within 0.003 cm of the wire, while the peak on the left is for tracks between 0.259777 and 0.26777 cm of the wire. The Gaussian fit functions are shown on top.

other cells.

Having measured the bias of the time-to-distance relation, we can subtract the bias from the original time-to-distance relation (from simulation only) as a correction. We can then re-process the data and re-do the track fits, which should improve the tracks. The procedure can be iterated to further improve the tracks. We do two iterations of these improvements, as the average correction is already zero at the second iteration. Figure 7.16 shows the resolution and the corrections for the 3 iterations. 0^{th} order is what we call the process before any improvements are applied, i.e., using just the simulation-produced time-to-distance relation. The sharp corners in the correction functions near the sense wire are to maintain the constraint

that signals that arrive at 0 ns are identified with 0 mm distance from the wire.

Like the Garfield program, our corrections and resolutions are done for small intervals of track distance from the wire. For use in the actual track fitting process, we need a continuous function. For the original Garfield time-to-distance relation we used the monotonic Steffen interpolant. In this case it was desirable to use relatively low-order piecewise polynomials. No library could be found that provided piecewise-continuous polynomials with continuous derivatives, so I coded one myself.

During the iterative improvement process, it was noted that the distribution of track distances was not the expected uniform one. In Figure 7.17 one can clearly see a gap in the middle, an overpopulation at ~ 1.5 mm, and a uniform distribution inside the cell. The tracks with absolute distances greater than ± 7 mm are from bad fits. Many attempts were made to determine the cause of the apparent missing tracks near the sense wire. Clearly this is not a real physical effect, as the beam spot in our tests is $\sim 1\,\text{cm} \times 20\,\text{cm}$ wide and could not possibly miss the wire. Somewhere in the procedure tracks which are close to the wire are being mischaracterized, but we were unable to determine how or why. The Gaussian fits described above to determine the resolutions and corrections to be applied were not reliable very close to the wire, because of the low statistics in the histograms for tracks close to the wire. This is the origin of the tests of shifting the global 0 reference time described in Section 7.2. It was found that shifting the global reference by a few nanoseconds increased the population of tracks near the sense wire by a small amount, but not enough to account for the large gap. The problem is also the origin of the manual enforcement of the constraint that signals with 0 ns arrival times should be identified with 0 mm distance from the wire. The gap was found early on in the development of the analysis, and the constraint was added as an attempt to solve the problem. It did not resolve the issue, but it is a good idea anyways.

In order to continue making progress with the project (as I had been stuck for quite a while at this point), it was decided to ignore the problem and continue the analysis while only considering those tracks that cross the

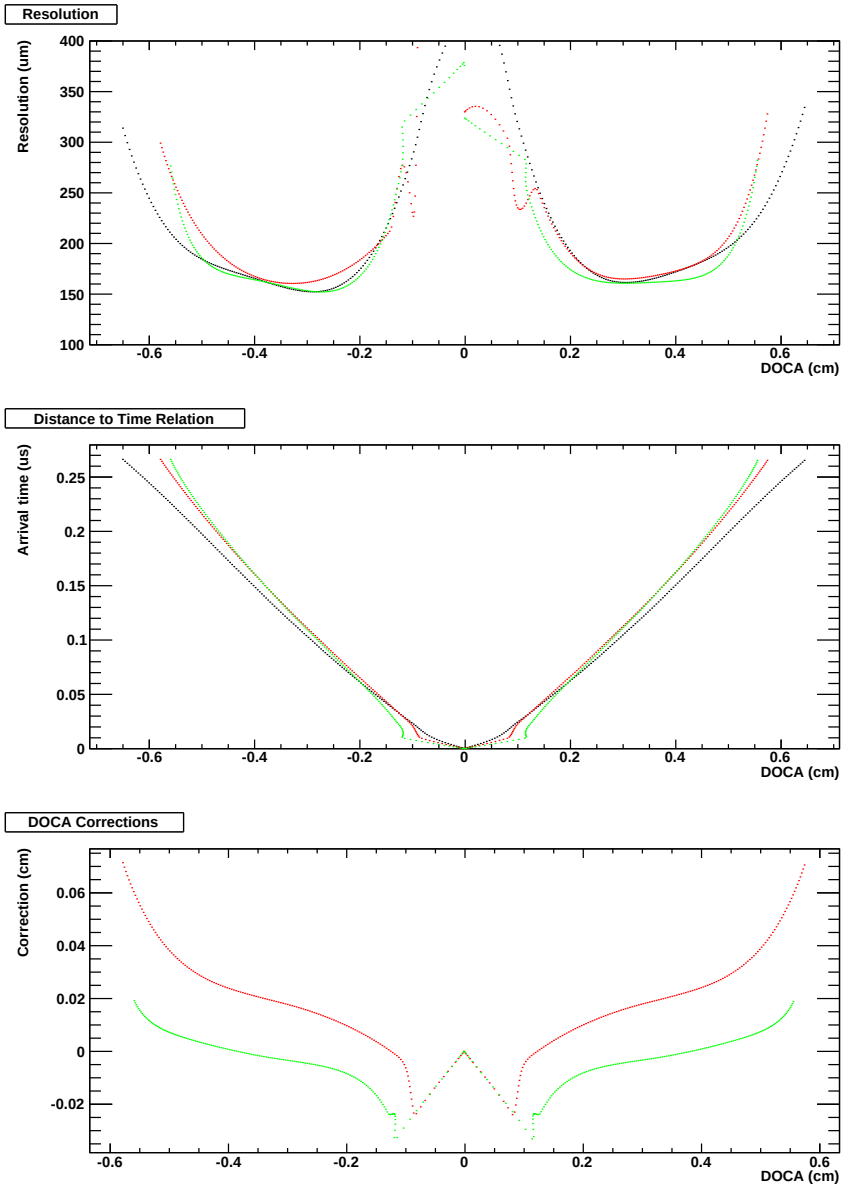


Figure 7.16: Plots showing the resolution of the drift chamber (top), time-to-distance relation (middle), and correction to the time-to-distance relation (bottom) at the various levels of iterative improvement.

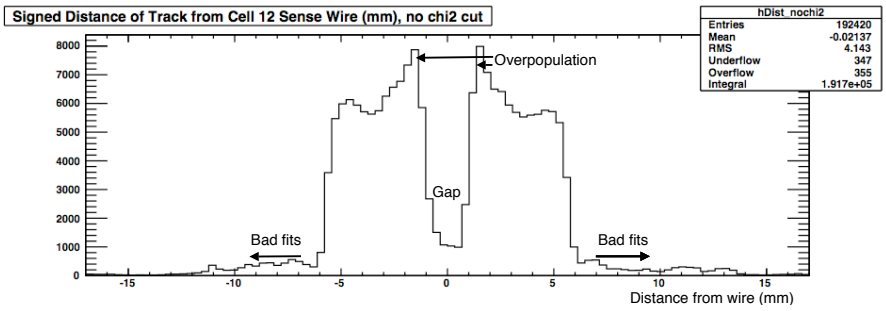


Figure 7.17: Distribution of track distances from cell 12, according to the track fit using the other 7 layers.

middle section of the cells between 0.1 and 0.55 cm. This is not as bad as it sounds, because generically in drift chambers the resolution becomes much worse near the sense wire and near the cell edges. In most cases when reporting a single value for the resolution of a drift chamber, an average over the middle part of the cell or the best resolution is used. In Figure 7.16, we ignore the fluctuating resolution curves near the wire at 0 mm, and we note that the resolution in the middle part of the cell does improve through the iterative improvement process. We can finally claim to have measured a resolution of 160 μm in the best part of the cell in our prototype using traditional tracking means. In comparison, BaBar's drift chamber has a resolution of 100 μm in the best part of its cells.

Possible reasons as to why BaBar's drift chamber had significantly better resolution than our prototype are as follows: the gas mixture is different, and if the Super$B$ gas has a higher drift speed, the resolution will suffer; our prototype used mostly off-the-shelf electronics, while BaBar had custom electronics, so our signals may have been noisier; our prototype only used 7 layers of cells to determine the "known good" tracks, while BaBar had 40 layers, so the contribution to the resolution from the track fitting procedure should be $\sim \sqrt{40/7}$ times smaller than with our prototype.

7.6 Timing Clusters

7.6.1 Motivation

This thesis and the paper presented in Chapter 4 demonstrates that cluster counting improves a drift chamber's ability to identify particles. This was generally believed even before it was shown to work in drift chambers, since similar techniques were used in time expansion chambers (TECs) since the 1970s [62] (a TEC is another type of gaseous ionization detector). Researchers are more skeptical of improvements to tracking from cluster counting, since no comparable device ever used the technique. Only recently have some researchers discussed the concept [76, 77]. The work presented in this chapter is the first time that cluster counting has been used for tracking in a full-scale drift chamber under realistic conditions.

For particle identification, only the number of clusters is used, as this is analogous to the traditional measurement of the integrated charge; both measurements are trying to determine the number of primary ionization events. In the case of tracking, the number of clusters is not relevant. The useful information is the arrival times of the clusters, analogous to the arrival time of the whole signal when doing traditional tracking.

One could imagine a situation where the number of clusters gives useful tracking information: in long thin cells, the angle of the track would affect the track length, so the number of clusters (and indeed the integrated charge) would give a measure of the angle. For our prototype and for typical flavour-factory drift chambers, the cells are mostly circularly symmetric, so the number of clusters does not reveal information about the track, or at least no information that we would expect to be an improvement over traditional tracking.

It is clear that cluster counting can help in particle identification: the traditional method of charge integration requires discarding of data from the truncated mean procedure, while the cluster count measurement has better-behaved statistics and allows us to use the information from all layers. For tracking, the reasoning is much less clear. Why should the times of

the clusters beyond the first contain any useful information at all? After all, the clusters correspond to primary ionizations, which are independent Poisson-distributed events. The first cluster happens to be the one closest to the point of closest approach, so later clusters should simply give a worse measurement of the track distance.

The potential improvement can be qualitatively understood by considering two signals in a drift chamber cell such as shown in Figure 7.18. Both signals have the same arrival time using the traditional method, but the times of the subsequent clusters are very different. The signal at the top of the figure looks suspiciously like a large spurious pulse followed by a real signal from a charged particle. Such spurious signals can come from contamination in the drift chamber (e.g., a charged dust particle accidentally hitting the wire), from a delta-ray coming from an adjacent cell, or other kinds of unplanned and difficult-to-account-for sources. The signal at the bottom looks more like a "typical" signal coming from a charged particle track. Using traditional tracking, these would be effectively identical signals. If we measure the arrival times of the clusters, it will become apparent that the top signal is of a different kind than the bottom. A clever algorithm that does not only consider the arrival time of the first pulse edge may effectively discard the first pulse from the top signal, and properly identify the rest of the signal as coming from a track far away from the wire.

A more mathematical but still qualitative motivation for how cluster counting can help determine a charged particle track is presented in Chapter 6. There, a simple theoretical model of a drift chamber cell is constructed, and a Bayesian statistical analysis is done to show how the posterior probability density function for the track impact parameter (distance from the wire) is affected by the information contained in clusters after the first one.

7.6.2 Overview of Technique

First an algorithm has to be chosen, and its parameters optimized. This is described in Section 7.6.3. Then all the signals that were previously used to determine arrival times for traditional tracking are re-analyzed with the

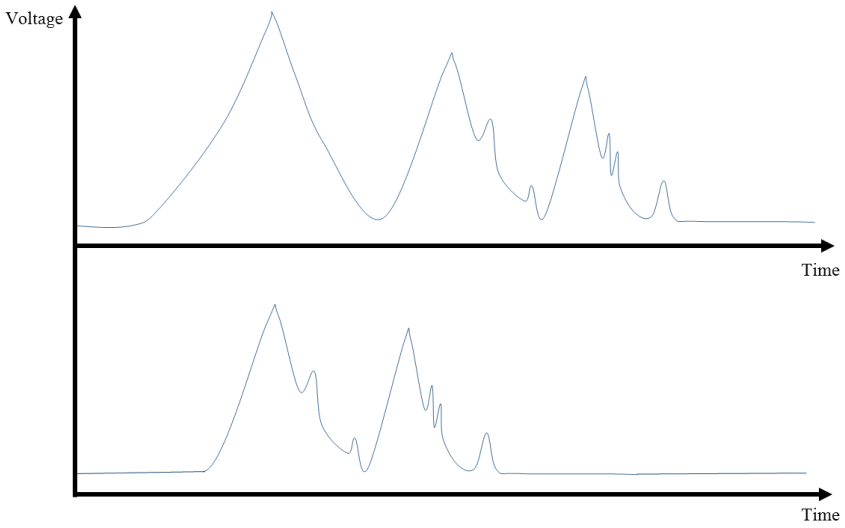


Figure 7.18: Illustration of two signals with the same arrival time using the traditional method, but with different arrival times for subsequent clusters.

cluster-counting algorithm. Using the track distance determined from the traditional track fitting, we make histograms of the cluster arrival times for small intervals of track distance (Figure 7.19). Thus we obtain empirical distribution functions for the cluster arrival times as a function of track distance from the wire.

With these distributions, we now go back to the cluster times in a specific event and ask the question "from which track distance's time distribution is this particular set of times most likely to have been drawn?" To answer this, we calculate the likelihood of drawing that particular set of cluster times from the empirical distribution for each track distance. The likelihood is just a number proportional to the probability, and the exact calculation is shown in Appendix A.3.

The actual quantity calculated is the negative of the log of the likelihood. The logarithm is more convenient numerically because it turns several multiplications of possibly-tiny numbers into a sum of more reasonable numbers.

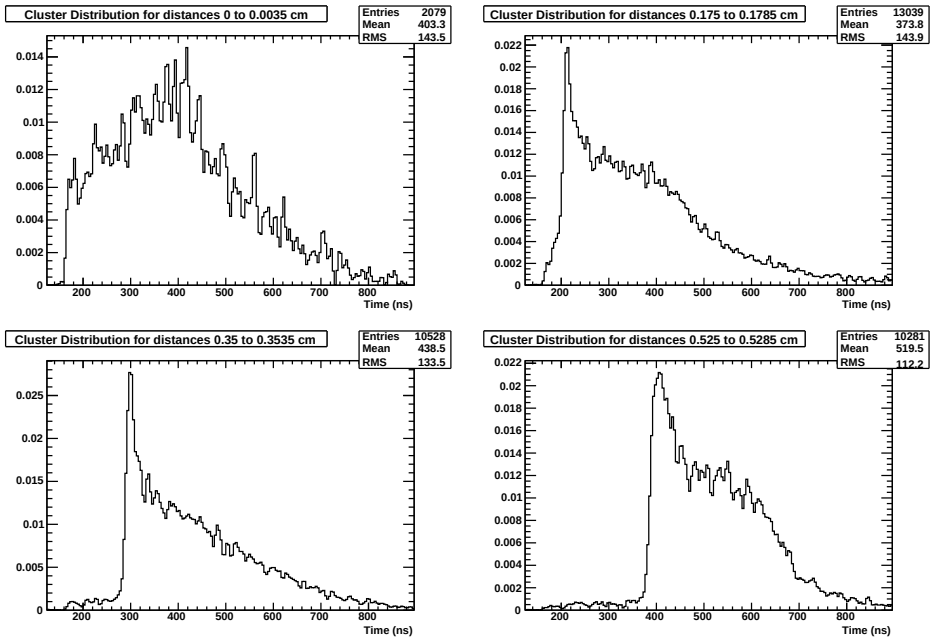


Figure 7.19: Example empirical distribution functions for cluster arrival times for four different track distance intervals.

The negative is a convention because most computer function-optimization programs do minimization, so finding the minimum of the negative log-likelihood is equivalent to finding the maximum of the likelihood itself.

With the negative log-likelihoods (NLLs) calculated, it is a simple matter to find the track distance whose NLL is the minimum. That distance is the cluster-counting equivalent of converting the signal arrival time into a distance using the time-to-distance relation. The equivalent of the resolution - the uncertainty in the time to distance relation - is the shallowness of the minimum in the NLL versus distance graph. The minimum is only searched for in part of the cell. Recall that there was an anomalous absence of tracks with distances very-near the sense wire (described at the end of Section 7.5). This made the residuals and resolution measurements difficult, but it also mars the empirical distribution of cluster times at the extreme edges of the cell. As with the residuals and resolutions, our solution for the sake

of making progress is to effectively cut off the ends of the cell, and only consider minima within the finite boundaries 0.1 to 0.55 cm, as anomalous false minima sometimes are found outside of these boundaries, simply due to statistical fluctuations.

7.6.3 Algorithms and Parameters

This tracking study was done after the particle identification study described in Chapter 4. In that study, we found that the specific choice of algorithm was not so important once the parameters were optimized. Thus for this tracking study, we employ only one algorithm: the "signal above average" described in Section 4.8.2. It was also found that an optimal smoothing time was ~ 5 ns, and this was not re-optimized for the tracking study. The only parameter that was re-optimized was the threshold, because the optimal threshold depends on the combined gas and amplifier gain, which is different in proto 2 than in the single-cell chambers used for the PID study. In the PID study a typical signal amplitude is ~ 100 mV while for the tracking study it is ~ 1000 mV. Rather than find an appropriate scaling between the two experiments, we simply re-do the optimization of the threshold. With the specific algorithm and smoothing time fixed, the optimization is one-dimensional, so this is a simple task. The optimization is done using a grid-search method. Several thresholds are tested and evaluated with a figure of merit, and a graph is made to find the optimal value.

First we determine the empirical distribution functions for the cluster times from the data. This is done for each threshold value under consideration (Figure 7.20). Then we randomly select a track distance, and retrieve the empirical distribution function for that distance. Simulations using random numbers are colloquially called "Monte Carlo" in particle physics, even when the technique is relatively simple. Thus we call the randomly-selected track distance the "MC truth" distance.

We know the average number of clusters found in a signal at the given threshold, so we generate a random cluster count using a Poisson distribution, then randomly draw that number of clusters at times following the

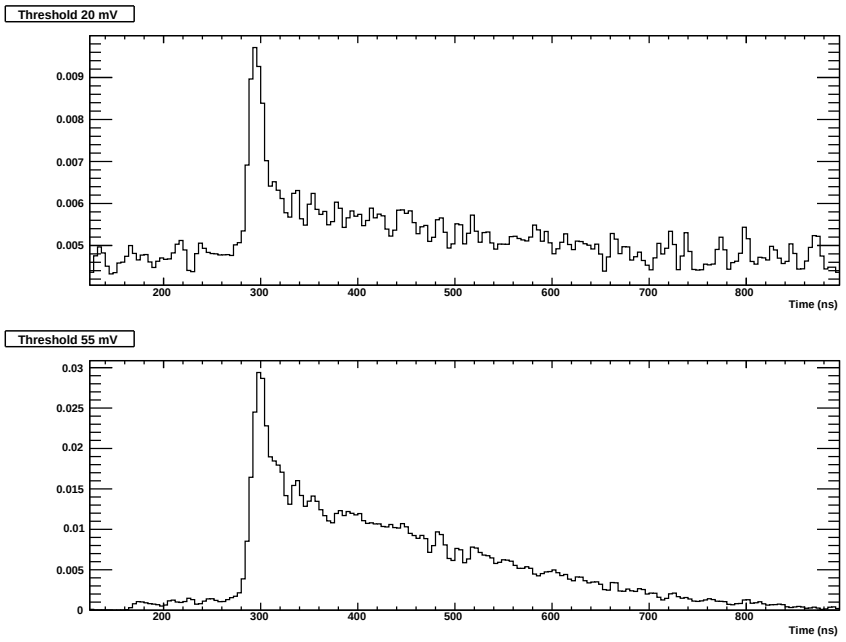


Figure 7.20: A comparison of the empirical distribution functions for the cluster times at two different thresholds. Both are for tracks between 0.35 and 0.3535 cm from the wire.

empirical distribution function at the MC truth distance. These cluster times are then used to calculate the minimum-NLL track distance, which we call the "MC measurement". In other words, we generate a fake signal given the data about a particular track distance, and then ask "how consistent is this with the data at that distance?" We do this many times and make a histogram of the difference between the MC truth and MC measurement distances (Figure 7.21). The resulting histograms (one per threshold tested) contain information about the self-consistency of the cluster information, i.e., whether the cluster time distributions can actually be used to determine a track distance. The central value of the histograms gives a measure of the bias, and the width gives a measure of the uncertainty in a determined distance.

To decide which threshold value is best, we form a figure of merit that

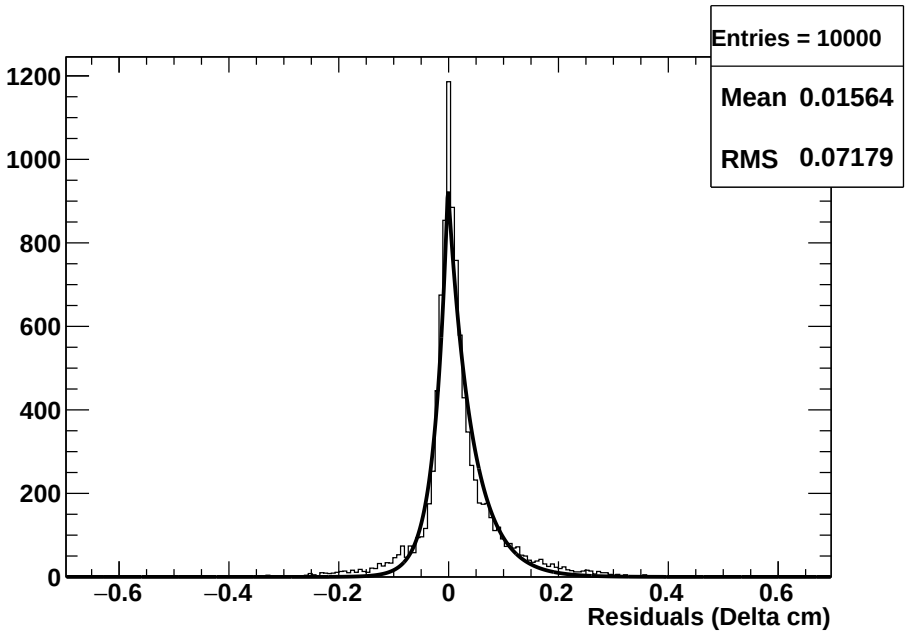


Figure 7.21: An example residual plot between the "MC truth" track distance and the "MC measurement" distance. This is for tracks in the middle part of the cell, using the optimal cluster-counting algorithm parameters (4 ns smoothing, 55 mV threshold). The best-fit line is with an asymmetric Laplace distribution.

measures the width of the histograms of the MC truth and MC measurements. These histograms however are non-Gaussian, so we must be careful with the figure of merit. The histograms appear to follow a kind of asymmetric Laplace distribution. The Laplace distribution is basically two back-to-back exponential distributions centered at μ:

$$f(x) = \frac{1}{2b} \exp -\frac{|x-\mu|}{b} \tag{7.4}$$

where μ is the central value and b is the decay constant. In many of our cases, the decay constants on each side were different, so we fit the following

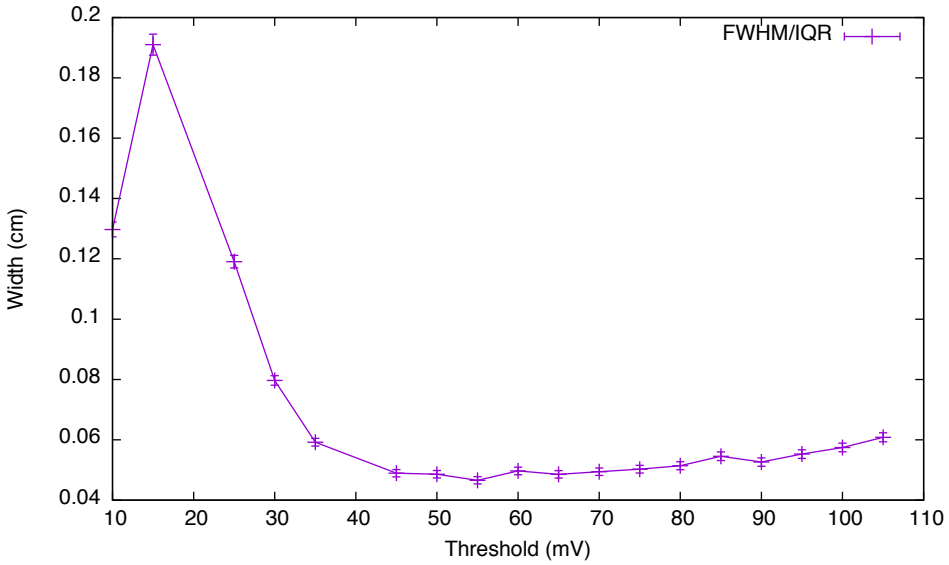


Figure 7.22: Full-width at half-maximum of asymmetric Laplace distribution fits to the MC truth-MC measurement residuals at different threshold values. The error bars are derived from the errors in the fit.

function:

$$f(x) = \frac{A}{b_1 + b_2} \begin{cases} \exp -\frac{x-\mu}{b_1} & x >= \mu \\ \exp -\frac{\mu - x}{b_2} & x < \mu \end{cases} \tag{7.5}$$

where b_1 and b_2 are the decay constants on either side of the central value, and A is an overall normalization. The figure of merit chosen is the full-width at half-maximum of the fitted function, which for the asymmetric Laplace distribution can be shown to be $(b_1 + b_2) \ln 2$. Several other figures of merit were tried: the width of a symmetric Laplace distribution, the inter-quartile range, and the width parameters of a sum-of-Gaussians. All were consistent with the one described above. Given our choice, a smaller figure of merit is better. In all cases, the central value of the residual plots was much smaller than the width, indicating zero or negligible bias in the technique.

In Figure 7.22 one can see the figure of merit plotted over a large range

of threshold values. Thresholds of 10 to 20 mV have a terrible figure of merit, and a plot of the clusters found (Figure 7.23) clearly shows that the measurement is dominated by fake clusters, because the threshold value is well-within the noise of the signal. As the threshold is raised out of the noise the figure of merit improves quite a bit, but then at yet-higher thresholds it gets worse again. The higher thresholds in the plot are high enough to miss pulses from real clusters, so that the connection between track distance and the cluster times becomes less solid. Fortunately the minimum is clearly very shallow, so the choice of a threshold of 55 mV is not too critical; as long as we are within $\sim \pm 10$ mV of the minimum the performance is roughly the same.

7.7 Combined Likelihood

In the particle identification study, the information obtained from cluster counting was combined with the traditional PID from charge integration via a combined likelihood. The same thing is done here. The reason is that clearly any drift chamber built to do cluster counting will also be able to do the measurements with the traditional techniques. The PID study showed that the PID performance was noticeably improved by combining the traditional measurement with cluster counting. Here we attempt to do the same using the cluster times for tracking.

The same method will be used here as in the PID study: combined likelihood. The negative log likelihood was minimized to find the track distance according to the cluster times (Section 7.6.2), but the traditional measurement was only obtained from converting the signal arrival time into a distance (and its uncertainty) using the time-to-distance relation and corrections. To convert the traditional measurement into a likelihood, we assume that the measured distance and uncertainty correspond to the central value and width of a Gaussian distribution. This is a bad assumption for tracks near the wire, as mentioned earlier the distribution of signal arrival times for tracks near the wire are non-Gaussian and required the use Novosibirsk functions (Section 7.3). Conveniently, the tracks near the wire have bad

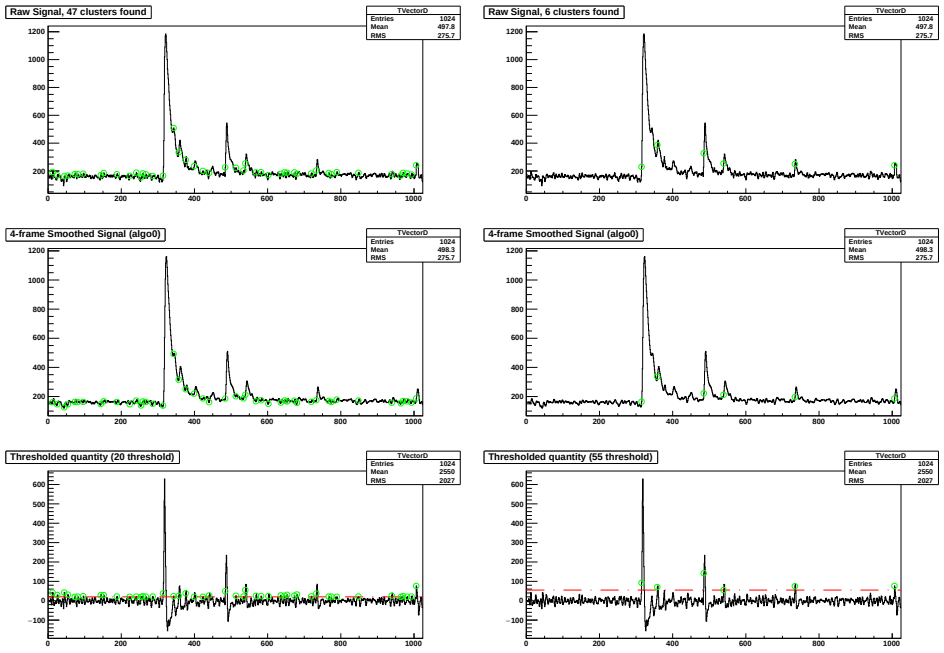


Figure 7.23: The same signal analyzed for clusters using two different thresholds. The three figures on the left use a threshold of 20 mV, those on the right 55 mV (the optimal value). The triplet of plots features the unprocessed signal, the smoothed signal, and the actual derived quantity on which the threshold value is applied, at the top, middle, and bottom, respectively. On each plot the identified clusters are indicated with a green circle. On the derived quantity plot, the threshold is shown with a dashed red line.

statistics and we choose to exclude that part of the cell from our analysis (Section 7.5).

The negative log of the Gaussian distribution is quite simple:

$$-\ln G(x;\mu,\sigma) = \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 \tag{7.6}$$

where μ is the central value and σ is the width. Likelihoods for multiple

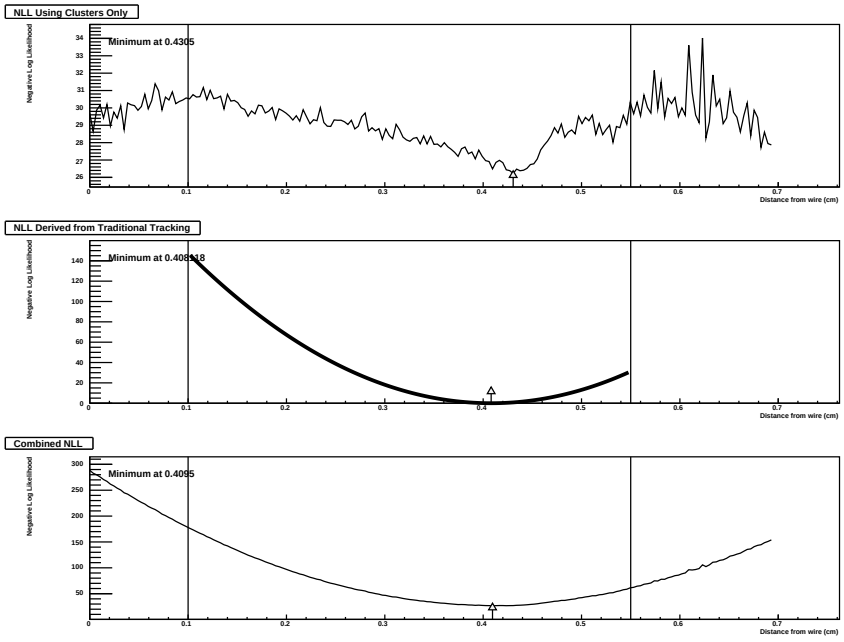


Figure 7.24: Plots showing the negative log likelihoods (NLLs) as a function of track distance computed in three different ways. The top graph shows the NLL using the cluster time information as described in Section 7.6.2. The middle plot shows the quadratic NLL derived from the traditional tracking distance and resolution as described in Section 7.7. The bottom plot shows the sum of these two. On each plot is shown two vertical lines that delimit the range of searching for the minimum. The location of the minimum on each is shown with an arrow.

measurements are multiplied together, so the negative log likelihoods are added together. This is shown in Figure 7.24 where the negative log Gaussian likelihood is shown in the middle plot.

The plots in Figure 7.24 show very typical results: the traditional measurement yields a certain value and uncertainty, and the corresponding measurement using the cluster times yields a mostly-compatible value. This is to be expected, as the arrival time of the first cluster is completely equivalent to the arrival time of the whole signal - though they are measured with

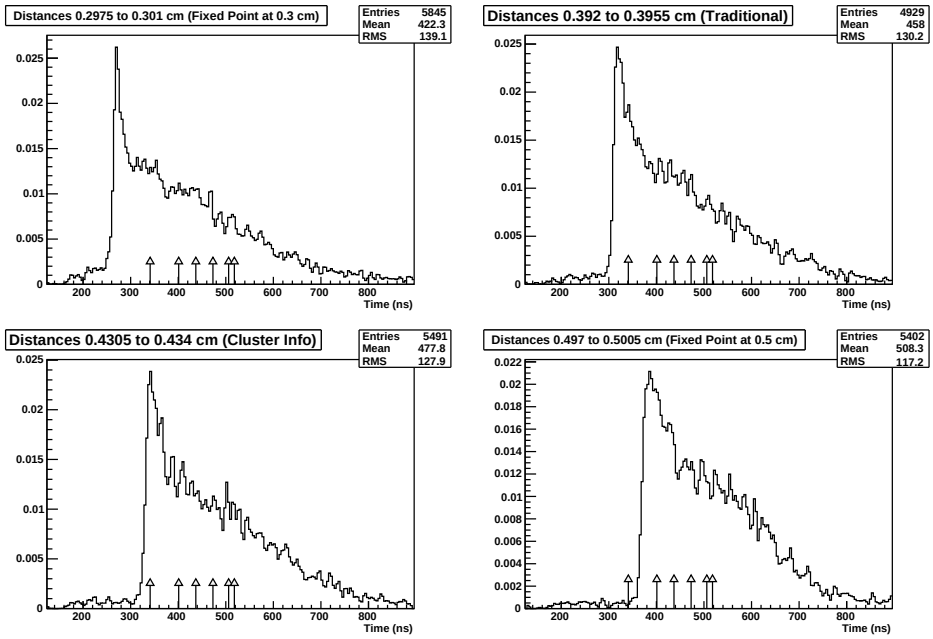


Figure 7.25: Expected cluster arrival time distributions at four different distance intervals. The intervals include an arbitrary point at 0.3 cm, the distance of the track using traditional tracking, the distance of the track using the cluster information in the signal from cell 12, and another arbitrary point at 0.5 cm. The arbitrary points are chosen to be much closer and further from the wire than either found by the algorithms, respectively. The actual cluster times are indicated by the arrows. The track and clusters used in the calculation are the same as in Figure 7.24, where a plot of the calculated likelihood values is shown.

different algorithms. Indeed from Figure 7.25 one can see that the most important factor in determining the likelihood is whether a cluster (or clusters) is found within the peak of the empirical distribution function. The difference between the traditional and cluster-based measurements are due to statistical fluctuations and (hopefully) the additional information contained in the later clusters. In the combined likelihood graph, the location of the minimum is pulled slightly left or right depending on the negative

log likelihood plot of the cluster-based measurement. Typically this pull is small: rarely do we get very incompatible minima from the two techniques. Furthermore, the variations in the negative log likelihood calculated based on the traditional method over the width of the cell are ~ 100, while the variation in the values from the cluster timing calculation are much smaller. Thus the dominant measurement is still the traditional method, unless a particularly deep minimum is found with the cluster technique. Indeed the resulting combined-likelihood plot is still mostly parabolic, even after adding the values from the cluster technique.

The location of the minimum in the combined likelihood graph is the measured value of the track distance with the two methods combined. The distances left and right away from the minimum that the graph crosses the minimum likelihood ± 0.5 are used to measure a 1σ uncertainty on this measurement.

7.8 Results

To quantify the contribution of the cluster-time information to the single-cell tracking resolution, we use the same technique as was used to determine the resolution in the case of traditional tracking. Since the accuracy of the track-fitting using 7 layers is much better than the single-cell resolution even when only using traditional tracking ($40\,\mu$m compared to $160\,\mu$m), and we expect the contribution from cluster-counting to be rather modest, we again use the 7-layer track fit as a "known good" measurement of the track position. We then compare the distance of the track from the wire in cell 12 according to the 7-layer track fit with the distance from the wire according to the combined-likelihood measurement in cell 12 alone.

Histogram of these residuals are made for small intervals of track distance from the wire, and the width of these histograms give the resolution (Figure 7.26). The distributions are very Gaussian-like, so we fit with Gaussian functions and use the σ parameter for the resolution. We do this for the single-cell measurement using just traditional tracking, just the cluster information, and with the combined likelihood.

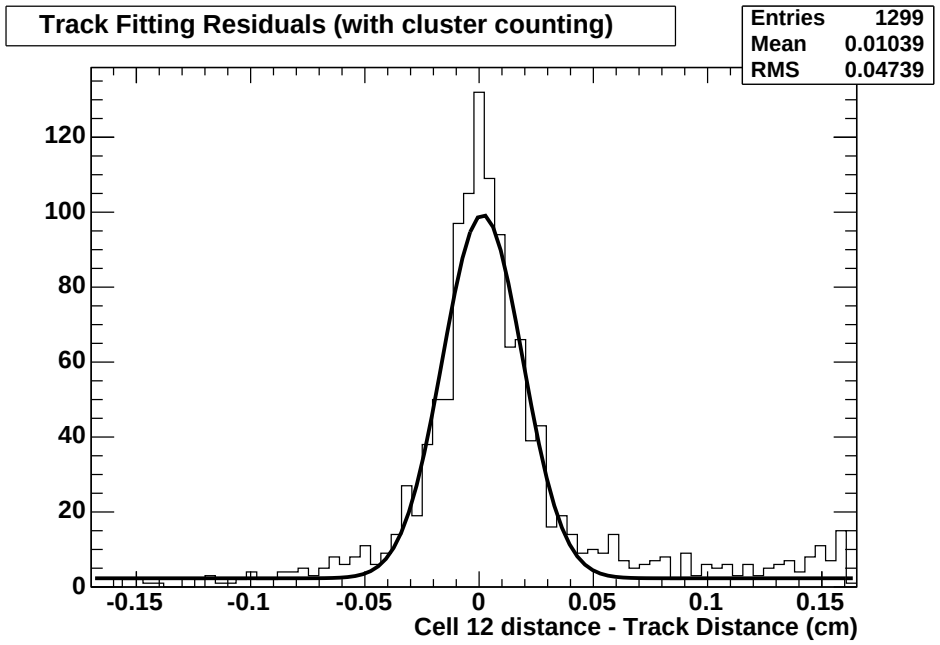


Figure 7.26: Histogram of residuals between the track distance according to cell 12 (using the combined likelihood measurement) and the distance accorindg to the track fit using the other 7 layers. These are for tracks between 0.259 and 0.2625 cm from the wire. Superimposed is the Gaussian fit to the distribution, whose width gives us the resolution in this distance interval.

The resulting resolution as a function of track distance can be seen in Figure 7.27. The vertical values are the resolution - the widths of the Gaussian fits to the residual histograms, while the error bars are the uncertainty on that fit parameter. There is no discernable improvement from the addition of cluster counting using this technique.

The resolutions using the cluster information only (i.e., without traditional tracking) and using the combined likelihood technique are in fact slightly worse than traditional tracking. This may mean that indeed all the valuable information is in the first cluster, and that the later clusters only add noise. The additional information from the later clusters is either

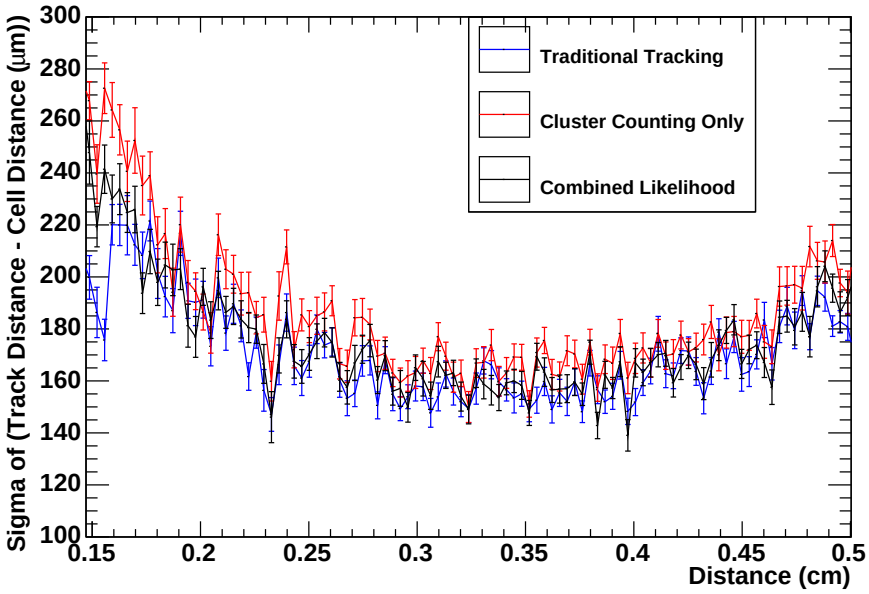


Figure 7.27: Single-cell resolution in cell 12 as determined using three techniques. Note that the bottom of the vertical scale is not at $0\,\mu$m.

not useful, or of such small import that it does not make up for the extra variability from needing to identify the clusters. When averaging over track distances, the degradation in resolution between the traditional technique and the combined likelihood is $\sim 2.4\,\mu$m.

The technique used to make the cluster-based measurement may also not be the only option. One could imagine an algorithm that considers each cluster one at a time, similar to the Bayesian model presented in Chapter 6.

The conclusion is that cluster counting for tracking in proto 2 using the technique presented in this chapter is not beneficial. The limitations of the experiment however do not completely rule out possible benefits in future drift chambers, and it may be worthwhile to keep the option open.

Chapter 8

Conclusion

8.1 Summary of Results

8.1.1 Generalities

The two studies presented in this thesis represent the first analyses of data taken from full-length cluster-counting drift chambers. All previous work in the literature has been with small bench-top test chambers. Previous single-cell chambers have mostly been so-called "drift tubes" with a single sense wire in a metal tube. The particle identification study uses a single-cell chamber, but the chamber is constructed using actual field and guard wires. The tracking study uses the first full-length multicell cluster-counting drift chamber ever constructed. Testing realistic detectors is essential, because even if positive results are obtained in smaller or simpler devices, their design may not necessarily be scalable to large chambers that would actually be used in experiments. For example, a small demonstration chamber can use generic coaxial cables (which are quite bulky) for the high-voltage and sense wire connectors, whereas a large drift chamber has to carefully minimize the wire materials.

The prototypes used in this work were designed to match the design of the eventual Super$B$ drift chamber. All of these designs were preliminary, as the project was cancelled before a final design as chosen. While each

drift chamber is designed for a particular experiment, the choices made for Super$B$ should be similar to those made for other flavour-factory or flavour physics collider experiments. People working on future upgrades to the Belle-II, MEG-II (PSI), or BES-III drift chambers would likely find results from this thesis to be interesting. Indeed MEG-II is already exploring the option of using cluster counting for particle identification in their drift chamber [77].

8.1.2 Particle Identification

The most significant result is the measurement of the improved particle identification (PID) when the cluster count is considered alongside the integrated charge. When requiring 90% of real muons to be rejected (i.e., *not* identified as pions), the efficiency of identifying real pions (i.e., the fraction positively identified) improves from 50% to 60% (Chapter 4). The ionizing behaviour of pions and muons at the momentum of our test beam (210 MeV/c) is almost the same as for pions and kaons at $\sim 1\,\mathrm{GeV/c}$ momentum. This is very important for flavour physics experiments where the typical decay products of the particles under study are pions and kaons at $\sim 1\,\mathrm{GeV/c}$. Higher identification efficiency of the daughter particles means that more real composite particles (e.g., B-mesons) can be reconstructed for the same amount of data collected. Thus implementing cluster counting for PID can boost the effective luminosity of the experiment without increasing background rates, or a specific desired signal significance can be obtained with less integrated luminosity.

Also very significant is the determination that extremely high-bandwidth electronics are not needed to acquire a useful signal for cluster counting. Previously it was thought that at a minimum, one would require 1 ns sampling times and GHz bandwidths to properly identify clusters. In the PID study our data acquisition system was an oscilloscope taking 0.05 ns samples and with 4 GHz bandwidth, but the front-end amplifiers had a bandwidth of 2.7 GHz. In the process of optimizing the parameters of our cluster-counting algorithms, it was found that a smoothing time of $\sim 5\,\mathrm{ns}$ was optimal. This

means that for a real experiment, the data acquisition system could have a sampling rate of only 5 ns, do no smoothing to the waveform, and achieve the same PID improvements as our study. According to the Nyquist sampling theorem [78, 79], the minimum bandwidth required to properly record a signal at a sampling rate of 5 ns is only ~ 100 MHz. This is very fortunate, because high-bandwidth and high-frequency ($\sim$ GHz) electronics are more expensive and consume more power (and thus produce more heat), so accommodating them on the endplate of a large drift chamber with thousands of sense wires would be a difficult engineering challenge. For comparison, the electronics installed in the BaBar drift chamber in 1998 had 200 MHz bandwidth [80].

Secondary results of the particle identification study are that the specific choice of algorithm is not so important as long as the parameters are optimized properly. All algorithms tested gave about the same performance after optimization. This means that an analog device such as an application-specific integrated circuit (ASIC) could be used to count the clusters instead of a programmable digital device like a field-programmable gate array (FPGA). For the quantity required for a drift chamber, ASICs are more costly to design and manufacture, but they are much smaller and consume less power than FPGAs. Considering the time intervals between identified clusters did not improve PID performance when used in a tripartite likelihood with the cluster count and the integrated charge.

8.1.3 Tracking

The main result of the tracking study is that, despite assertions in the literature [76],

there is no measurable improvement in the single-cell tracking resolution when cluster information is considered. Our study used a maximum-likelihood technique and empirical cluster time distributions for the cluster-based tracking. The final measurement was done with a combined-likelihood of both the traditional tracking information and the cluster times. The resolution of the prototype is 150 μm in the best part of the cell, and around

175 μm averaged over the middle of the cell. The resolution near the sense wire and near the edge of the cell are not available due to a still-unexplained gap in the apparent distribution of track distances from the wire.

8.2 Future Improvements

8.2.1 Generalities

Many improvements and corrections to the presented studies are possible. Quite likely if Super$B$ had not been cancelled, more studies would have been done, possibly with new prototypes, before the final design and construction of the drift chamber. These studies would have resolved many of the ambiguities encountered in the work presented here. New work on cluster-counting drift chambers (such as for MEG-II [77]) will have to address these issues.

In the beam tests for both studies, we had data-taking shifts running 24 hours per day in three 8-hour shifts. It was noticeable that the data taken during the graveyard (midnight to morning) shift more frequently had problems than data taken during the other shifts. Problems included missing parameters such as temperature and pressure, or hard-to-interpret notes about anomalies observed. This was particularly apparent for the night shifts done solo at LNF (the data from which was not analyzed in this work). For future beam tests I highly recommend going to a four-shift rotation with 6 hour shifts, and always assign shift-takers in pairs at a minimum.

In both studies, a lot of data was wasted because anomalies in the data were discovered only after the beam tests. For example, in the beam test for the tracking study, it was discovered after data-taking that all the time-of-flight information was somehow uncorrelated with the drift chamber signals. This is likely because of a misconfiguration of the data-acquisition system, and we were unable to correct the problem in the data after the test. During data taking, we had standard sets of histograms and graphs to make after each run (so-called "online" analysis), but the process was tedious and not

automated, so the plan was often not adhered to very consistently. Moreover, those histograms and graphs would not have identified all problems. In the same example as above, our online analysis made histograms of the time-of-flight information and the charge integration measurements individually, and these looked normal. We only noticed that the measurements were uncorrelated on an event-by-event basis when the data was analyzed more carefully after the mean test (so-called "offline" analysis). As a result, no PID study was possible with the proto 2 data from that beam test. Fortunately it was still possible to perform a tracking study using the self-consistency of the tracks, so the data was not all wasted. More effort in designing and implementing the online analysis would reduce the wasted data.

8.2.2 Particle Identification

The beam tests planned with the single-cell prototypes were quite ambitious: we planned to study multiple front-end amplifiers, different cables, wire voltages, and many track positions along the sense wire's length. The combinatoric explosion of permutations meant that not all combinations could be tested. After data taking, we learned that some large fraction of the data was unusable e.g., because of a bad amplifier. So many data runs comparing beam angles and positions were wasted because they were all done using the same untested amplifier. Future beam tests should be less ambitious about the number of parameters to be tested. If the goal is to test amplifiers, all amplifiers should be tested under essentially the same conditions (or the same small set of conditions), and compared to a reference amplifier. When possible, redundant data should be taken, because often one learns of a problem with a given data set only after the beam test is done.

In the beam test with the single-cell prototypes, the temperature and pressure of the gas in the chamber were not controlled. In the hall where the test took place, there is a large door to the outside which was sometimes opened for loading and unloading trucks. There was a large difference in

pressure between the door being opened and closed ($\sim 3\,\mathrm{mbar}$). The temperature fluctuated with the time of the day and over the course of the weeks of the test. Since the drift chamber operates at the same temperature and pressure as the hall, these parameters were monitored and recorded (crudely, with binoculars and a weather station inside the beam test area). Temperature and pressure can affect the gas gains, so a systematic study of the wire voltages would have required more consistent (i.e., automatic) recording of these parameters. The only significant results from the PID study involved analyses of single-run datasets. These data-taking runs were short enough that the environmental conditions did not change appreciably during each run. Thus to avoid unnecessarily complicating the study, we did not use the actual temperature and pressure information in the analysis.

8.2.3 Tracking

I was not involved in the construction of the single-cell prototypes for the PID study, but Giuseppe Finocchiaro and I strung every wire in proto 2. There are several things about the construction of proto 2 that could be improved upon.

While stringing the chamber by hand, we first tried wearing latex gloves to avoid touching the wires and feedthroughs with our bare skin. Unfortunately the loss of dexterity made it extremely difficult. We soon switched to manipulating the wire and feedthroughs with bare (but cleaned) hands in order to finish the construction on schedule. Contaminants from contact with human skin (oils) are mostly a concern for ageing. We only touched the outer parts of the feedthroughs, but again there was no clear indication of what level of cleanliness was required.

In the construction of proto 2, we used an epoxy resin to seal gas leaks and smooth over defects in the aluminium structure. This epoxy is the same that was used for the LHCb straw tubes (Section 5.3), and is presumably non-problematic in a drift chamber environment (e.g., no outgassing of organic molecules), but this was not checked rigorously. Similarly the low-viscosity acrylic used to seal the feedthroughs was not thoroughly checked.

We should have contacted the manufacturers of these adhesives or performed experiments to determine their suitability for use in a drift chamber.

The exact type of acrylic used was actually not written down in an easy-to-find place, which is a serious oversight. We can no longer retroactively check to see if it was an appropriate choice. This was actually a chronic problem with proto 2, and is likely just a different working style of the Italian collaborators and the TRIUMF group. The Italian collaborators have a Principal Investigator, but all team members have equivalent levels of seniority or responsibility. While this has clear advantages in terms of flexibility and collegiality, there is no obvious "boss" to ensure that systematic notes are taken and procedures are laid out and followed. The result is that if detailed notes were taken, they are not consistently available to the whole group. Were Super$B$ still an ongoing project, this would not be a problem as the collaborators would have easy and regular contact with the others. Unfortunately writing this thesis now nearly three years after the cancellation, many of the details are inaccessible. Indeed a collaborative Wiki that was hosted on the INFN website was deleted, and it contained some aggregated notes that I had transferred there. Having better notes about the electronics and data acquisition setup may have allowed us to determine the source of the non-correlation between the time-of-flight and the integrated charge measurement, and thus perform a PID study with proto 2.

The PID study used an external PID device (a time-of-flight detector) which allowed us to use "known good" measurements to compare against the drift chamber measurements. For the tracking study, we did not have an external tracking system, so the "known good" tracking measurement had to come from a track fit using 7 out of 8 layers in the prototype. An external tracker would have allowed us to use all 8 layers and to compare the whole track fit to the external tracker's measurement. We did have an external tracker ready to use (made of stacked drift tubes), but it stopped functioning shortly before the beam test. With more preparation time, we would have been able to arrange an alternate external tracker. Future tracking studies should use an external tracker in order to obtain better results.

The tracking prototype used terminated sense wires. This means that

a termination resistor (R_T in Figure 5.7) was used that minimizes reflected signals at the far end of the drift chamber. While this makes the signal analysis more straightforward (one can ignore the possibility of reflected signals), another effect of the resistor is to reduce the total signal amplitude. Even if the analysis confuses the "true" signal and the reflected one, it is possible that better performance is obtained with unterminated sense wires. The beam test with the single-cell prototypes in the PID study planned to study the difference in PID performance between terminated and unterminated running, but due to data-quality issues, we were unable to obtain results. Any future test of a cluster-counting drift chamber should carefully study the two options to determine which is best.

8.2.4 Other Ideas

Here are a few ideas and suggestions for improving drift chambers.

Nonlinear ADCs and Amplifiers

Due to the large statistical fluctuations of drift chamber signal amplitudes, the amplifiers and analog-to-digital converters (ADCs) inevitably encounter saturating signals. For example, if we have an 8-bit ADC whose full scale is 2 V, any signal above 2 V will register the same as 2 V. The signal is "clipped" at the maximum range of the amplifier or ADC. Thus the large-amplitude signals are badly resolved, and can have bad effects with cluster-counting algorithms that use derivatives. A possible improvement would be to use a non-linear ADC. Most ADCs are linear, and indeed linearity is usually a good property. A linear amplifier is one that maps the digital value to a fraction of the maximum range. For example an 8-bit ADC with a range of voltages from 0 to 2 V would map the digital value 0 to 0 V, the value 255 ($2^8 - 1$) to 2 V, and intermediate values i to $i/255 \times 2\,\mathrm{V}$.

For drift chamber signals, the large-amplitude voltages are more rare than low amplitude ones, but extremely-large amplitudes still occur. A linear ADC effectively "wastes" high-order bits on high amplitudes and still quickly saturates. Another way of saying this is that the ADC has poor

dynamic range. Dynamic range is the ratio of the maximum non-saturating voltage that can be measured to the smallest non-zero voltage that can be measured. Our example 8-bit amplifier above has a dynamic range of 255:1.

ADCs can be built purposely to be non-linear to better match the signal they are meant to digitize. An example is a logarithmic amplifier, where the mapping between the digital value D and the voltage V is

$$V(D) \propto \exp\left(\frac{D}{N}\frac{V_{\max}}{V_{\min}}\right) \tag{8.1}$$

where $V_{\max}$, and $V_{\min}$ are reference voltages in the ADC and N is the number of digital values available (255 for an 8-bit ADC). The actual relation will depend on the design of the amplifier, but dynamic ranges in the thousands are attainable with 8-bit logarithmic ADCs [81, 82]. Another example of a non-linear ADC is a "dual-range" ADC, which has two linear sections. In a dual-range ADC, measurements below or above a given threshold have different resolutions. Like the ADCs, nonlinear amplifiers can also be built [83], with the same benefits as nonlinear ADCs.

Replacing the Truncated Mean

The truncated-mean procedure is very simple, and was originally used in part because it was simple to implement with older electronics. It is likely that a more complex use of the individual charge measurements, without discarding, can give a robust estimate of the primary ionizations. For example if the full long-tailed probability distribution was well-measured with known particle tracks, the later non-truncated measurements could be used to assigned a likelihood of being a given particle. This was investigated during the thesis work of fellow Super*B* student Rocky So [42], who found that the improvement to particle identification performance is modest, compared to the gains found from cluster counting.

Bibliography

[1] J.-F. Caron, et al. Improved particle identification using cluster counting in a full-length drift chamber prototype. *Nucl. Instr. Meth. Phys. Res. A*, 2014. 735(0):169 – 183. doi:10.1016/j.nima.2013.09.028. http://www.sciencedirect.com/science/article/pii/S016890021301262X → pages iv, 40

[2] B. Gough. *GNU Scientific Library Reference Manual - Third Edition*. Network Theory Ltd., 3rd edition, 2009. ISBN 0954612078, 9780954612078. → pages v, 115, 129

[3] R. Brun and F. Rademakers. ROOT — An object oriented data analysis framework. *Nucl. Instrum. Meth.*, 1997. 389(1–2):81 – 86. doi:10.1016/S0168-9002(97)00048-X. http://www.sciencedirect.com/science/article/pii/S016890029700048X → pages v, 55, 115, 126, 132, 185

[4] R. Veenhof. Garfield, recent developments. *Nucl. Instrum. Meth.*, 1998. 419(2-3):726 – 730. doi:10.1016/S0168-9002(98)00851-1. http://www.sciencedirect.com/science/article/pii/S0168900298008511 → pages v, 47, 49, 56, 125

[5] G. Van Rossum. *The Python Language Reference Manual*. Network Theory Ltd., 2003. ISBN 0954161785. http://www.worldcat.org/isbn/0954161785 → pages v

[6] F. Pérez and B. E. Granger. IPython: a System for Interactive Scientific Computing. *Comput. Sci. Eng.*, 2007. 9(3):21–29. http://ipython.org → pages v

[7] J. Generowicz, et al. Reflection-Based Python-C++ Bindings. In *Computing in High Energy and Nuclear Physics (CHEP)*. 2004 . http://www.osti.gov/energycitations/product.biblio.jsp?osti_id=835826 → pages v

[8] S. van der Walt, S. Colbert, and G. Varoquaux. The NumPy Array: A Structure for Efficient Numerical Computation. *Comput. Sci. Eng.*, 2011. 13(2):22–30. doi:10.1109/MCSE.2011.37. → pages v

[9] G. Sciolla. The mystery of CP violation, 2006. Retrieved: July 2015. http://web.mit.edu/physics/news/physicsatmit/physicsatmit_06_sciollafeature.pdf → pages 2

[10] J. H. Christenson, et al. Evidence for the 2π decay of the K_2^0 meson. *Phys. Rev. Lett.*, 1964. 13:138–140. doi:10.1103/PhysRevLett.13.138. → pages 4

[11] F. Sauli. Principles of Operation of Multiwire Proportional and Drift Chambers. *CERN Academic Training Lecture*, 1977. https://cds.cern.ch/record/117989?ln=en → pages 4

[12] W. Blum, W. Riegler, and L. Rolandi. *Particle Detection With Drift Chambers*. Springer, 2008. ISBN 9783540766834. http://books.google.ca/books?id=RgxUlAXv0RAC → pages 6, 15, 18, 20, 21, 30, 41, 42, 44, 93

[13] D. Andrews, et al. The CLEO detector. *Nucl. Instr. Meth. Phys. Res.*, 1983. 211(1):47 – 71. doi:10.1016/0167-5087(83)90556-2. http://www.sciencedirect.com/science/article/pii/0167508783905562 → pages 10

[14] A. Abashian et al. The Belle detector. *Nucl. Instr. Meth. Phys. Res. A*, 2002. 479(1):117 – 232. doi:10.1016/S0168-9002(01)02013-7. http://www.sciencedirect.com/science/article/pii/S0168900201020137 → pages 11, 33

[15] H. Bethe. Zur theorie des durchgangs schneller korpuskularstrahlen durch materie. *Annalen der Physik*, 1930. 397(3):325–400. doi:10.1002/andp.19303970303. → pages 11, 42

[16] J. Beringer, et al. Review of Particle Physics. *Phys. Rev. D*, 2012. 86. doi:10.1103/PhysRevD.86.010001. http://link.aps.org/doi/10.1103/PhysRevD.86.010001 → pages 11, 42, 44

[17] O. Baker, et al. The high momentum spectrometer drift chambers in hall C at CEBAF. *Nucl. Instr. Meth. Phys. Res. A*, 1995. 367(1–3):92 – 95. doi:10.1016/0168-9002(95)00737-7. Proceedings of the 7th International Wire Chamber Conference. http://www.sciencedirect.com/science/article/pii/0168900295007377 → pages 18, 21

[18] E. Gheflin. *An Improved Measurement of the Number of Light Neutrino Families by Single Photon Counting with the OPAL Detector at LEP*. Ph.D. thesis, University of California, Riverside, 1993. http://www.egheflin.com/professional/edwardgheflin_phd_thesis/EdwardGHeflin_PhD_Thesis003.html → pages 19

[19] M. Danilov, et al. The ARGUS drift chamber. *Nucl. Instr. Meth. Phys. Res.*, 1983. 217(1–2):153 – 159. doi:10.1016/0167-5087(83)90124-2. http://www.sciencedirect.com/science/article/pii/0167508783901242 → pages 21, 24

[20] I.-A. Yadigaroglu, et al. Studies of energy loss in helium gas mixtures. *Nucl. Instr. Meth. Phys. Res. A*, 1992. 323(1–2):322 – 326. doi:10.1016/0168-9002(92)90310-Z. http://www.sciencedirect.com/science/article/pii/016890029290310Z → pages 21

[21] W. Zimmermann, et al. Helium-propane as drift chamber gas. *Nucl. Instr. Meth. Phys. Res. A*, 1986. 243(1):86 – 90. doi:10.1016/0168-9002(86)90824-7. http://www.sciencedirect.com/science/article/pii/0168900286908247 → pages 21

[22] M. Agnello, et al. Performance of a large area drift chamber operating with a He:iC4H10 mixture. *Nucl. Instr. Meth. Phys. Res. A*, 1995. 367(1–3):100 – 103. doi:10.1016/0168-9002(95)00740-7. Proceedings of the 7th International Wire Chamber Conference. http://www.sciencedirect.com/science/article/pii/0168900295007407 → pages 21

[23] P. Burchat, et al. Studies of helium gas mixtures in drift chambers. *Nucl. Instrum. Meth.*, 1992. A316:217–222. doi:10.1016/0168-9002(92)90903-H. → pages 21, 46

[24] G. Bencivenni, et al. Creep measurement on aluminum-5056 wires, 1998. KLOE Memo #138. http://web-docs.gsi.de/~webhades/mdc/literatur/KLOE-memo138.pdf → pages 23

[25] D. Schmidt. Technology of drift chambers. *Nucl. Instr. Meth.*, 1980. 176(1–2):39 – 50. doi:10.1016/0029-554X(80)90679-5. http://www.sciencedirect.com/science/article/pii/0029554X80906795 → pages 23

[26] D. Cockerill, et al. Operation of a drift chamber vertex detector at the ISR. *Nucl. Instr. Meth.*, 1980. 176(1–2):159 – 162. doi:10.1016/0029-554X(80)90696-5. http://www.sciencedirect.com/science/article/pii/0029554X80906965 → pages 23, 24

[27] A. Bechini, et al. A modular drift-chamber vertex detector at the CERN ISR. *Nucl. Instr. Meth.*, 1978. 156(1–2):181 – 186. doi:10.1016/0029-554X(78)90711-5. http://www.sciencedirect.com/science/article/pii/0029554X78907115 → pages 23, 24

[28] M. B. Luque, et al. The construction of the central detector for an experiment at the CERN p-p collider. *Nucl. Instr. Meth.*, 1980. 176(1–2):175 – 180. doi:10.1016/0029-554X(80)90699-0. http://www.sciencedirect.com/science/article/pii/0029554X80906990 → pages 24

[29] D. Cassel, et al. Design and construction of the CLEO II drift chamber. *Nucl. Instr. Meth. Phys. Res. A*, 1986. 252(2–3):325 – 330. doi:10.1016/0168-9002(86)91201-5. http://www.sciencedirect.com/science/article/pii/0168900286912015 → pages 24

[30] F. Barbosa, et al. A drift chamber system for a toroidal detector. *Nucl. Instr. Meth. Phys. Res. A*, 1992. 323(1–2):191 – 197. doi:10.1016/0168-9002(92)90287-E. http://www.sciencedirect.com/science/article/pii/016890029290287E → pages 24

[31] S. Baru, et al. The KEDR drift chamber. *Nucl. Instr. Meth. Phys. Res. A*, 1992. 323(1–2):151 – 156. doi:10.1016/0168-9002(92)90281-8. http://www.sciencedirect.com/science/article/pii/0168900292902818 → pages 24

[32] M. Fero, et al. Performance of the SLD central drift chamber. *Nucl. Instr. Meth. Phys. Res. A*, 1995. 367(1–3):111 – 114. doi:10.1016/0168-9002(95)00732-6. Proceedings of the 7th International Wire Chamber Conference. http://www.sciencedirect.com/science/article/pii/0168900295007326 → pages 24

[33] I. Advanced Probing Systems. Tungsten, tungsten-rheium, and beryllium-copper material properties. Retrieved: July 2015. http://www.advancedprobing.com/pdfs/tungsten-rhenium-beryllium-copper-material.pdf → pages 24

[34] Wikipedia. Molybdenum, 2015. Retrieved: July 2015. https://en.wikipedia.org/wiki/Molybdenum → pages 24

[35] A. Imanishi, et al. The TOPAZ inner drift chamber. *Nucl. Instr. Meth. Phys. Res. A*, 1988. 269(3):513 – 521. doi:10.1016/0168-9002(88)90127-1. http://www.sciencedirect.com/science/article/pii/0168900288901271 → pages 24

[36] K. McDonald. Should the drift chamber inner wall be load bearing?, 1996. BaBar Note TNDC-96-20. https://www.slac.stanford.edu/BFROOT/www/Detector/CentralTracker/TNDC_Notes/tndc-96-20.ps.gz → pages 25

[37] M. Baszczyk et al. SuperB Technical Design Report. *ArXiv e-prints*, 2013. http://arxiv.org/abs/1306.5655 → pages 26, 33, 36, 41, 97

[38] B. Aubert et al. The BaBar detector. *Nucl. Instrum. Meth.*, 2002. A479:1–116. doi:10.1016/S0168-9002(01)02012-5. http://www.sciencedirect.com/science/article/pii/S0168900201020125 → pages 27, 29, 33, 46, 65, 140

[39] J. Va'Vra. Review of wire chamber aging. *Nucl. Instr. Meth. Phys. Res. A*, 1986. 252:547–563. doi:10.1016/0168-9002(86)91239-8. → pages 27, 29

[40] A. M. Boyarski. Model of high-current breakdown from cathode field emission in aged wire chambers. *Nucl. Instr. Meth. Phys. Res. A*, 2004. 535(3):632 – 643. doi:10.1016/j.nima.2004.06.164. http://www.sciencedirect.com/science/article/pii/S0168900204016602 → pages 27

[41] A. Boyarski. Additives that prevent or reverse cathode aging in drift chambers with helium-isobutane gas. *Nucl. Instr. Meth. Phys. Res. A*, 2003. 515(1-2):190 – 195. doi:10.1016/j.nima.2003.08.148. Proceedings of the International Workshop on Aging Phenomena in Gaseous Detectors. http://www.sciencedirect.com/science/article/pii/S0168900203024690 → pages 27, 30

[42] R. Y. C. So. *Search for a light Higgs boson with the BABAR detector : with proposals to improve drift chamber aging studies and particle identification*. Ph.D. thesis, University of British Columbia, 2015. http://dx.doi.org/10.14288/1.0166452 → pages 27, 168

[43] L. Malter. Thin film field emission. *Phys. Rev.*, 1936. 50:48–58. doi:10.1103/PhysRev.50.48. http://link.aps.org/doi/10.1103/PhysRev.50.48 → pages 28

[44] M. Y. Dong, et al. Aging effect in the BESIII drift chamber. *ArXiv e-prints*, 2015. http://arxiv.org/abs/1504.04681 → pages 29

[45] R. Rios, et al. Sealed drift tube cosmic ray veto counters. *Nucl. Instr. Meth. Phys. Res. A*, 2011. 637(1):105 – 108. doi:10.1016/j.nima.2010.12.098. http://www.sciencedirect.com/science/article/pii/S0168900210028780 → pages 29

[46] B. Aubert et al. The BaBar detector: Upgrades, operation and performance. *Nucl. Instr. Meth. Phys. Res. A*, 2013. 729:615 – 701. doi:10.1016/j.nima.2013.05.107. http://www.sciencedirect.com/science/article/pii/S0168900213007183 → pages 30

[47] Fermilab. Energy frontier, 2011. Retrieved: July 2015. https://www.fnal.gov/pub/science/experiments/energy/ → pages 31

[48] R. R. Wilson. The Tevatron. *Phys. Today*, 1977. 30N10:23–30. doi:10.1063/1.3037746. → pages 31

[49] M. Riordan. The demise of the superconducting super collider. *Physics in Perspective*, 2000. 2(4):411–425. doi:10.1007/s000160050053. → pages 31

[50] Q. Qin, et al. Overview of the CEPC accelerator. In *55th ICFA Advanced Beam Dynamics Workshop on High Luminosity Circular e+ e-Colliders–Higgs Factory (HF2014)*. 2014 . http://accelconf.web.cern.ch/AccelConf/HF2014/papers/thp3h2.pdf → pages 31

[51] T. Behnke, et al. The International Linear Collider Technical Design Report - Volume 1: Executive Summary. 2013. http://arxiv.org/abs/1306.6327 → pages 31

[52] Fermilab. Intensity frontier, 2010. Retrieved: July 2015. https://www.fnal.gov/pub/science/experiments/intensity/ → pages 32

[53] H. Albrecht, et al. Observation of B^0- $\bar{B}^0$ mixing. *Physics Letters B*, 1987. 192(1–2):245 – 252. doi:10.1016/0370-2693(87)91177-4. http://www.sciencedirect.com/science/article/pii/0370269387911774 → pages 32

[54] D. Du and Z. Zhao. Implications of large B_d^0 - $\bar{B}_d^0$ mixing. *Phys. Rev. Lett.*, 1987. 59:1072–1075. doi:10.1103/PhysRevLett.59.1072. → pages 32

[55] M. Bona et al. *SuperB*: A High-Luminosity Asymmetric e^+ e^- Super Flavor Factory. Conceptual Design Report. *ArXiv e-prints*, 2007. http://arxiv.org/abs/0709.0451 → pages 32, 41

[56] B. Meadows, et al. The impact of SuperB on flavour physics. *ArXiv e-prints*, 2011. http://arxiv.org/abs/1109.5028 → pages 32

[57] SuperB Collaboration. SuperB Progress Reports – Detector. *arXiv:1007.4241*, 2010. page 95. http://arxiv.org/abs/1007.4241 → pages 33, 41

[58] SuperB Collaboration. SuperB Progress Reports – Physics. *ArXiv e-prints*, 2010. http://arxiv.org/abs/1008.1541 → pages

[59] SuperB Collaboration. SuperB Progress Reports – Accelerator. *ArXiv e-prints*, 2010.
http://arxiv.org/abs/1009.6178 → pages 33

[60] T. Abe, et al. Belle II Technical Design Report. *ArXiv e-prints*, 2010.
http://arxiv.org/abs/1011.0352 → pages 33

[61] M. Zobov, et al. Crab Waist Collision Studies for e+e- Factories. *ArXiv e-prints*, 2008.
http://arxiv.org/abs/0802.2667 → pages 33

[62] A. H. Walenta. The Time Expansion Chamber and Single Ionization Cluster Measurement. *IEEE Trans. Nucl. Sci.*, 1979. 26(1):73–80. doi:10.1109/TNS.1979.4329616. → pages 41, 145

[63] G. Charpak, et al. The use of multiwire proportional counters to select and localize charged particles. *Nucl. Instrum. Meth.*, 1968. 62(3):262 – 268. doi:10.1016/0029-554X(68)90371-6.
http://www.sciencedirect.com/science/article/pii/0029554X68903716 → pages 41

[64] E. Blackmore. Operation of the TRIUMF (20-500 MeV) proton irradiation facility. In *Radiation Effects Data Workshop, 2000.* 2000 pages 1–5. doi:10.1109/REDW.2000.896260. → pages 51, 90

[65] S. Ritt and P. A. Amaudruz. The MIDAS Data Acquisition System. 2002. Retrieved: July 2015.
http://midas.triumf.ca → pages 56

[66] D. N. Poenaru and W. Greiner. *Experimental Techniques in Nuclear Physics*. De Gruyter, 1997. → pages 77

[67] R. Arink et al. Performance of the LHCb Outer Tracker. *JINST*, 2014. 9(01):P01002. doi:10.1088/1748-0221/9/01/P01002. → pages 94

[68] M. Adinolfi, et al. The tracking detector of the KLOE experiment. *Nucl. Instr. Meth. Phys. Res. A*, 2002. 488(1–2):51 – 73. doi:10.1016/S0168-9002(02)00514-4.
http://www.sciencedirect.com/science/article/pii/S0168900202005144 → pages 94

[69] C. H. W. Wong. Evaluation of adhesives for use in the babar drift chamber, 1996. BaBar Note TNDC-96-20. https://www.slac.stanford.edu/BFROOT/www/Detector/CentralTracker/TNDC_Notes/tndc-96-55.ps.gz → pages 94

[70] W. Alpha. integrate exp(-rho*z)/sqrt(xˆ2-zˆ2) dz from z = 0 to x - wolfram alpha. Retrieved: July 2015. http://wolfr.am/6n8roHkm → pages 113

[71] A. Andryakov, et al. The full-length prototype of the kloe drift chamber. *Nucl. Instr. Meth. Phys. Res. A*, 1998. 404(2–3):248 – 264. doi:10.1016/S0168-9002(97)01154-6. http://www.sciencedirect.com/science/article/pii/S0168900297011546 → pages 123

[72] G. Cabibbo, et al. The calibration of the space–time relations of the kloe drift chamber. KLOE Note 175, 2001. → pages 126

[73] H. Hirano, et al. A high-resolution cylindrical drift chamber for the KEK B-factory. *Nucl. Instr. Meth. Phys. Res. A*, 2000. 455(2):294 – 304. doi:10.1016/S0168-9002(00)00513-1. http://www.sciencedirect.com/science/article/pii/S0168900200005131 → pages 126

[74] M. Steffen. A Simple Method for Monotonic Interpolation in One Dimension. *Astron. Astrophys.*, 1990. 239:443. → pages 129

[75] L. A. Treinish and B. E. Rogowitz. Why should engineers and scientists be worried about color? *IBM Research Website*, 1996. Retrieved: July 2015. http://www.research.ibm.com/people/l/lloydt/color/color.HTM → pages 133

[76] M. Cascella, F. Grancagnolo, and G. Tassielli. Cluster counting/timing techniques for drift chambers. *Nuclear Physics B - Proceedings Supplements*, 2014. 248–250(0):127 – 130. doi:http://dx.doi.org/10.1016/j.nuclphysbps.2014.02.025. 1st Conference on Charged Lepton Flavor Violation 1st Conference on Charged Lepton Flavor Violation. http://www.sciencedirect.com/science/article/pii/S0920563214000267 → pages 145, 162

[77] G. Cavoto, et al. Study of the single cluster response of a helium-isobutane drift chamber prototype using 8 keV X-rays. *J. Instrum.*, 2015. 10(03):P03012. http://stacks.iop.org/1748-0221/10/i=03/a=P03012 → pages 145, 161, 163

[78] H. Nyquist. Certain topics in telegraph transmission theory. *American Institute of Electrical Engineers, Transactions of the*, 1928. 47(2):617–644. doi:10.1109/T-AIEE.1928.5055024. → pages 162

[79] V. A. Kotelnikov. On the carrying capacity of the ether and wire in telecommunications. In *Material for the First All-Union Conference on Questions of Communication, Izd. Red. Upr. Svyazi RKKA, Moscow*. 1933 . → pages 162

[80] D. Nelson. Babar drift chamber electronics shielding and grounding requirements. BaBar Note 97-74, 1997. https://www.slac.stanford.edu/BFROOT/www/Detector/CentralTracker/TNDC_Notes/tndc-97-74.ps.gz → pages 162

[81] W. Kester. Intentionally nonlinear ADCs. Tutorial, Analog Devices, 2009. Retrieved: July 2015. http://www.analog.com/media/en/training-seminars/tutorials/MT-018.pdf → pages 168

[82] J. Guo and S. Sonkusale. An area-efficient and low-power logarithmic A/D converter for current-mode sensor array. *Sensors Journal, IEEE*, 2009. 9(12):2042–2043. http://dx.doi.org/10.1109/JSEN.2009.2033199 → pages 168

[83] A. J. Peyton and V. Walsh. *Analog Electronics with Op-amps: A Source Book of Practical Circuits*. Cambridge University Press, 2001. Retrieved: July 2015. → pages 168

[84] H. Ikeda, et al. A detailed test of the CsI(Tl) calorimeter for BELLE with photon beams of energy between 20 MeV and 5.4 GeV. *Nucl. Instr. Meth. Phys. Res. A*, 2000. 441(3):401 – 426. doi:10.1016/S0168-9002(99)00992-4. http://www.sciencedirect.com/science/article/pii/S0168900299009924 → pages 180

[85] G. Doyle. Singular "they" and the many reasons why it's correct. Retrieved: July 2015. https://motivatedgrammar.wordpress.com/2009/09/10/singular-they-and-the-many-reasons-why-its-correct/ → pages 185

[86] G. Greer. The silver searcher. Retrieved: July 2015. http://geoff.greer.fm/ag/ → pages 186

[87] E. C. Government of Canada, Department of Justice. Legistics - themself or themselves? Retrieved: July 2015. http://canada.justice.gc.ca/eng/rp-pr/csj-sjc/legis-redact/legistics/p1p30.html → pages 186

[88] Google. Code of conduct - investor relations. Retrieved: July 2015. https://investor.google.com/corporate/code-of-conduct.html → pages 186

[89] J. Spiro. *Defending the Master Race: Conservation, Eugenics, and the Legacy of Madison Grant*. University of Vermont Press, 2009. ISBN 9781584657156. https://books.google.ca/books?id=4NoE2VyfN70C → pages 189

Appendix A

Supporting Materials

A.1 The Novosibirsk Function

The Novosibirsk function was first used to model the energy deposition in a calorimeter by the Belle experiment [84]. Unlike the Belle calorimeter we have no mathematical derivation from physical principles that motivates the use of this particular function. We use it because it is a continuous deformation of a Gaussian distribution with a semi-infinite range, and it works well for our purposes (Section 7.3). Figure A.1 shows the normalized function plotted using a few different parameter values.

Our Novosibirsk function is defined as

$$F(x; N, x_p, \sigma_E, \eta) = N \exp\left(-\frac{1}{2\sigma_0^2}\left(\ln\left(1 + \frac{x - x_p}{\sigma_E}\eta\right)\right)^2 - \frac{\sigma_0^2}{2}\right) \quad \text{(A.1)}$$

where x_p is the location of the peak, N is a normalization factor, η is an asymmetry parameter, and σ_0 is given by

$$\sigma_0 = \frac{2}{\xi}\sinh^{-1}\frac{\eta\xi}{2}. \quad \text{(A.2)}$$

σ_E is the standard deviation of the Gaussian distribution that is a limiting case of the Novosibirsk function as $\eta \to 0$, and ξ is the constant $2\sqrt{\ln 4} \approx 2.35$ which relates the Gaussian distribution's standard deviation to the

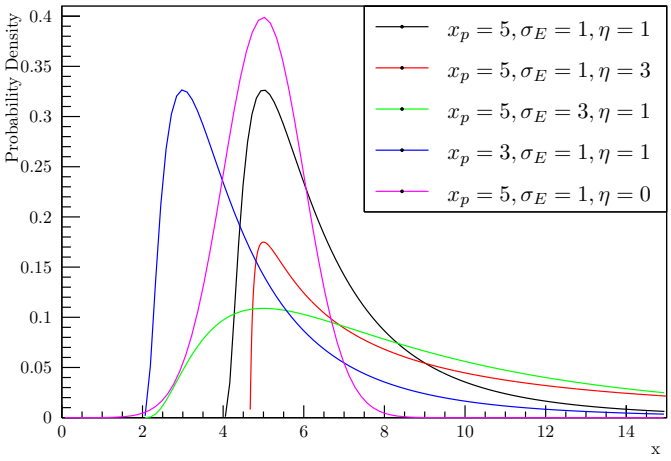


Figure A.1: Novosibirsk function plotted with various parameters. Each distribution has been normalized to unit area. The tallest curve has $\eta = 0$ and corresponds to a Gaussian. It is tallest because it does not have the extended tails from $\eta \neq 0$.

Novosibirsk function's full-width at half-maximum. Our function has the x-axis reversed compared to the original reference.

Here we show the limiting case of the Novosibirsk function. We will need two Taylor expansions for terms in the exponential, one for the logarithm in the numerator, and one for the inverse hyperbolic sine in σ_0.

$$\ln(1-x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots \tag{A.3}$$

$$\sinh^{-1} x = \sum_{n=0}^{\infty} \frac{(-1)^n (2n)!}{2^{2n}(n!)^2} \frac{x^{2n+1}}{2n+1} = x - \frac{1}{2}\frac{x^3}{3} + \frac{1\cdot 3}{2\cdot 4}\frac{x^5}{5} - \cdots \tag{A.4}$$

In the context of the limit $\eta \to 0$, these two series converge absolutely, so we are free to expand the logarithm and inverse hyperbolic sine inside the squared bracket to only the leading orders. The argument of the exponential

part of $F(x)$ becomes

$$\lim_{\eta\to 0} -\frac{1}{2}\frac{\left((\frac{x-x_p}{\sigma_E}\eta) - \frac{1}{2}(\frac{x-x_p}{\sigma_E}\eta)^2 + \text{h.o}\right)^2}{\frac{4}{\xi^2}\left(\frac{\eta\xi}{2} - \frac{1}{6}(\frac{\eta\xi}{2})^3 + \text{h.o.}\right)^2} - \frac{2}{\xi^2}\left(\frac{\eta\xi}{2} + \text{h.o.}\right)^2. \tag{A.5}$$

Again, because the series converge absolutely, we can expand the squared brackets naïvely, keeping only the leading order terms. The extra term coming from the lone $\sigma_0^2/2$ goes to zero in the limit. The argument of the exponential is now

$$\lim_{\eta\to 0} -\frac{1}{2}\frac{(\frac{x-x_p}{\sigma_E}\eta)^2 + \text{h.o.}}{\frac{2^2}{\xi^2}(\frac{\eta^2\xi^2}{4} + \text{h.o.})} = -\frac{1}{2}\left(\frac{x-x_p}{\sigma_E}\right)^2 \tag{A.6}$$

which is immediately recognizable as the argument of the exponential part of a Gaussian distribution with central value x_P and standard deviation σ_E. Thus in the limit $\eta \to 0$, $F(x)$ becomes a regular Gaussian distribution.

The Novosibirsk function also has semi-infinite support as a probability density function. The limit on x is determined by the logarithmic term. If we require that the logarithm remains real, then

$$1 + \frac{x-x_p}{\sigma_E}\eta > 0 \tag{A.7}$$

$$x > x_p - \frac{\sigma_E}{\eta}. \tag{A.8}$$

As we can see, as η goes to zero, this lower limit goes to $-\infty$ as it should for a Gaussian.

A.2 Track Distance From a Wire

Here we derive D, the absolute distance of the track to a given sense wire, and its uncertainty ΔD. These will be functions of the track parameters (x_0 and θ), the uncertainty and correlation of the track parameters (Δx_0, $\Delta\theta$, and $\text{cov}(\theta, x_0)$), and the sense wire coordinates (x_i and y_i).

First we define the slope of the track as

$$m = \frac{1}{\tan\theta} \tag{A.9}$$

and the point of closest approach on the track (x_p,y_p) where:

$$x_p = my_p + x_0 \tag{A.10}$$

and

$$y_p = \sin^2\theta(y_i - m(x_0 - x_i)). \tag{A.11}$$

The actual absolute track distance from the sense wire is then

$$D(x_0, \theta) = \sqrt{(x_p - x_i)^2 + (y_p - y_i)^2} = ||\vec{x}_p - \vec{x}_i||. \tag{A.12}$$

If we focus on the individual squared terms, we will see that D simplifies a lot.

$$\begin{aligned}
(y_p - y_i)^2 &= \left(\sin^2\theta(y_i - \frac{\cos\theta}{\sin\theta}(x_0 - x_i)) - y_i\right)^2 && \text{(A.13)}\\
&= \left((\sin^2\theta - 1)y_i - \sin\theta\cos\theta(x_0 - x_i)\right)^2 && \text{(A.14)}\\
&= \left(\cos^2\theta y_i + \sin\theta\cos\theta(x_0 - x_i)\right)^2 && \text{(A.15)}
\end{aligned}$$

$$\begin{aligned}
(x_p - x_i)^2 &= \left(\frac{\cos\theta}{\sin\theta}y_p + x_0 - x_i\right)^2 && \text{(A.16)}\\
&= \left(\cos\theta\sin\theta y_i + (\cos^2\theta - 1)x_i + \sin\theta x_0\right)^2 && \text{(A.17)}\\
&= \left(\cos\theta\sin\theta y_i + \sin^2\theta(x_0 - x_i)\right)^2 && \text{(A.18)}
\end{aligned}$$

Expanding the squared expressions and combining like terms allows several reductions using $\sin^2\theta + \cos^2\theta = 1$ to finally obtain

$$D = y_i\cos\theta + (x_0 - x_i)\sin\theta. \tag{A.19}$$

Using standard error propagation, we have the uncertainty on D:

$$\Delta D = \left[\left| \frac{\partial D}{\partial \theta} \right|^2 (\Delta\theta)^2 + \left| \frac{\partial D}{\partial x_0} \right|^2 (\Delta x_0)^2 + 2 \frac{\partial D}{\partial \theta} \frac{\partial D}{\partial x_0} \operatorname{cov}(\theta, x_0) \right]^{\frac{1}{2}} \tag{A.20}$$

where $\operatorname{cov}(\theta, x_0)$ is the covariance between θ and x_0.

Writing out the full expression for ΔD is not so illuminating, as none of the terms combine, so it is omitted.

A.3 Cluster Time Likelihood Calculation

The empirical distribution function of cluster times at a given track distance from the wire is stored in a histogram. The population in each bin of the histogram represents the average number of clusters expected there using Poisson statistics. Effectively these are the values μ_i where i indexes the possible cluster arrival times. In our case i corresponds to time in nanoseconds. μ_i is the expected number of clusters arriving at time i.

For a specific signal in a cell, we have a set of real cluster arrival times x_i. i again indexes the time and x_i is the number of clusters determined to have arrived at that time. Practically, x_i is always 1, but the code was made general enough to accommodate greater multiplicities.

The likelihood of having counted x_i clusters at time i is a Poisson distribution with mean μ_i:

$$L_i = \frac{\mu_i^{x_i}}{x_i!} e^{-\mu_i}. \tag{A.21}$$

For the whole set of clusters, we take the product of the likelihoods of all clusters.

In most applications of likelihoods, a more convenient quantity is the negative of the natural log of the likelihood. This turns the product into a sum, and the numerical values are smaller. This simplifies the calculations and makes numerical computer code more stable.

The final formula for the negative log likelihood is

$$L = \sum_{x_i} -x_i \ln \mu_i - \mu_i - \ln \Gamma(x_i + 1) \tag{A.22}$$

where the sum is over x_i, not i: i.e., we only consider the time bins where actual clusters were recorded. The log of the factorial is computed using the LnGamma function in ROOT for numerical precision. We also have a special case where $\mu_i = 0$, where the first logarithm term would give infinity. μ_i is only zero when the reference empirical distribution function has zero expected clusters at that time. This occurs randomly because some times have low statistics, and that bin accidentally has zero clusters in it. This is not such a big problem, and the work-around is to virtually expand the size of the bin. We iteratively sum up the μ_j of adjacent bins and take the average until a nonzero average μ_{i*} is found. Then this average μ_{i*} is used in the formula for the negative log likelihood. Typically only a few bins must be averaged before a suitable μ_{i*} is found.

A.4 Superfluous Gendered Language in Open-Source Project Documentation

Since this project made extensive use of the ROOT framework [3], a large amount of the documentation was read, and it was noticed that there were frequent examples where a default-male user was assumed. As an example, in the source file for the TMinuit class, from which the documentation is automatically generated, one can find this passage (emphasis added):

> The meaning of the parameters `par` is of course defined by the user, who uses the values of those parameters to calculate *his* function value.

There is nothing added by specifying the gender of the user here, and clearly the original author of the sentence did not mean to specify it. They were likely using an implied "generic masculine" that is common in many romance languages. In English, there is a reasonably well-established tradition of using "they" as a gender-neutral singular [85].

I searched through the entire ROOT source code using a tool called The Silver Searcher [86] for occurrences of the word "he", "him", "his", and "himself". In some cases, these were legitimate masculines as they were references to specific scientists with names like Adam. I made no attempt to discover the actual gender preference of these people, and assumed their names were conventional. Other times the "he" was in the phrasing "he or she" or similar. In these cases I made no changes.

In all other cases, I changed the wording to use the gender-neutral "they", "them", "their", and "themself", and sometimes changed the sentence to flow better. "Themself" is more controversial, but it also has an established use [87]. Basically "themselves" sounds extra-plural compared to "they" to me.

I purposely did not search for unnecessary specifications of feminine-gendered generic users, because the reality of our society is that the genders are not treated equally. When reading a manual, seeing a generic masculine is a reflection of our patriarchal state, while seeing a generic feminine is a challenge to it.

The various fixes were relatively simple and took little time. The changes were submitted as a patch by the ROOT developers and current releases now have the fix. It turns out that some of the code distributed with ROOT is from the "cling" project, so I submitted patches for that part of the code to the cling where it was accepted. Some of the code in cling is part of a package for software testing called Google Test. I submitted my patch there as well, but was met with resistance. Some users were supportive, but the most vocal argued with the "ungrammatical" or "neologistic" wording, or claimed that the fixed sentences were unreadable. One supporter reminded the developers of Google Test of this line from the Google Code of Conduct for investors (emphasis added):

> Failure of a Google contractor, consultant or other covered service provider to follow the Code can result in termination of *their* relationship with Google. [88]

In the end I grew quite uncomfortable with the tone of the discussion on the

Google Test forum, and my changes only affected part of the Google Test code anyways, so I gave up.

If the reader is interested in auditing a piece of open-source software as I did, the procedure is quite simple. Install The Silver Searcher or the more-commonly-installed "ack" and download the code for the software to be audited. Navigate to the base directory in your shell, and use the following command:

```
ag -i "\b(he|his|him|himself)\b"
```

If you use ack, use "ack" instead of "ag". The thing in the quotation marks is a regular expression that will match those whole words, and the -i makes it case-insensitive. The program may find a lot of false positives, e.g., if the code deals with stellar astrophysics (the abbreviation for helium is He) or if some developers use "his" as variable names for histograms.

A.5 Combining Bessel and Struve Functions

In Chapter 6 we encountered this form:

$$I_0(z) - L_0(z) \tag{A.23}$$

where I_0 is the modified Bessel function of the first kind at 0^{th} order, and L_0 is the modified Struve function at 0^{th} order. Each of the terms has a simple Taylor expansion:

$$I_0(z) = \sum_{k=0}^{\infty} \frac{\left(\frac{z}{2}\right)^{2k}}{(k!)^2}, \tag{A.24}$$

and

$$L_0(z) = \frac{z}{2} \sum_{k=0}^{\infty} \frac{\left(\frac{z}{2}\right)^{2k}}{(\Gamma(k+3/2))^2}. \tag{A.25}$$

These expansions look remarkably similar, and indeed it turns out they can be combined into a single sum. First we recognize that the factorial in Equation A.24 can be written as a Gamma function, and that the sum can be re-written to be only over even integers.

$$I_0(z) = \sum_{n=0,\text{even}}^{\infty} \frac{\left(\frac{z}{2}\right)^n}{\Gamma(n/2+1)^2} \tag{A.26}$$

Similarly, Equation A.25 can be rewritten to be a sum over only the odd integers.

$$L_0(z) = \sum_{n=1,\text{odd}}^{\infty} \frac{\left(\frac{z}{2}\right)^n}{\Gamma(n/2+1)^2} \tag{A.27}$$

The difference of the two functions can then be expressed as a single alternating sum:

$$I_0(z) - L_0(z) = \sum_{n=0}^{\infty} \frac{(-1)^n \left(\frac{z}{2}\right)^n}{\Gamma(n/2+1)^2}. \tag{A.28}$$

I was unable to find any existing special function which has this Taylor expansion.

A.6 Personal Philosophy of Science

The following passages represent my personal philosophy about the purpose of scientific research and its role in society. I am not a scholar of philosophy of science (nor are most scientists), so the arguments therein are not meant to stand up to rigorous philosophical examination. They are here to express an important part of scientific thought that is often omitted from scientific discource: the motivation behind it all.

A.6.1 Science

In today's world, science is often a bludgeon used to win arguments and silence opponents. Many things are called scientific that I would not consider as such. Perhaps those who misuse the term simply have a different definition of science, so I should make clear what I mean.

Science for me is about seeking truths. We don't have access to absolute truths, but science is one of the methods by which we can convince ourselves that we are not completely deluded. The kind of truth revealed through science relies on logical and temporal consistency. While we obviously can't

claim that the statement "in the presence of massive objects, other massive objects follow certain trajectories" is an absolute truth, it is absolutely consistent with all of our experiences and concepts, and we can reasonably expect that other people following logical consistency would reach the same conclusion.

Science is a process of refinement. What seems perfectly consistent on one day or in one age might later be revealed to be incorrect or imprecise. A scientific process must always be falsifiable or correctable in some way. Scientists must be fully honest about their claims of truth being based on consistency alone.

There are many kinds of truth and ways of accessing it. Science is only one way and it only reveals scientific truths. In many other cases, science is completely the wrong approach, or science can only partially inform arguments. For example in ethics, policy-making and aesthetics, over-reliance on science to the exclusion of other methods can and has led to terrible decisions. For example, in the 19th and early 20^{th} century, then-current scientific knowledge was used to back up decidedly racist and sexist ideologies [89].

In our society, scientific knowledge is seen as superior to other kinds of knowledge, and thus the term "scientific" is often used carelessly as a label to try to strengthen arguments. Arguments that need such bolstering are rarely of the scientific kind to begin with. Unfortunately most people do not understand enough about the philosophy of science to tell the difference, and misuse of scientific terms makes it very difficult to make good decisions.

As a concrete example, consider public reporting about climate change. There is a general consensus in the scientific community that global average temperatures are increasing, that the *rate* of change is increasing, and that this change is caused by human industrial activity. It is also generally believed that increased temperatures will lead to more extreme weather, with a negative impact on human well-being from failed agriculture, erosion, and loss of land to the ocean. Unfortunately in public reporting, those trying to promote this view to the public and those trying to claim that global warming is either false or not a problem each call the other non-scientific and try to back up their claims with scientific terminology. To the general

public this can be confusing and creates the appearance of a scientific debate when in fact there is a clear general consensus on one side. Most of the most-polluting nations in the world are nominally democratic, so if the public does not correctly perceive the threat of continued industrial expansion and unsustainable economies, there is little hope for positive change.

Another example can be found in the marketing of so-called "alternative" medicine and health products. Terms with well-defined meanings in science are abused to sell products. It is possible to find deodorant that claims to contain "no chemicals" (one must suppose it is made of pure energy, or just elementary particles).

Even in otherwise well-meaning groups, science is conflated with other endeavours that it barely resembles. STEM is an acronym meaning Science Technology Engineering and Mathematics. It is often used when referring to recruitment efforts, e.g., trying to get more women and people of colour to get degrees in these fields. At first glance the term seems like an inoffensive combination of disciplines, but no one seems to ask why they are grouped together. While it's true that science and engineering both use mathematical models, engineering as a human endeavour is far removed from science as I define it above. Lumping the four STEM fields together, given our society's obsession with immediate results and productivity, erases the possibility that science and mathematics could be curiosity-driven, exploratory, and completely devoid of marketability. While it's true that many scientific results have led to directly useful applications, it is my belief that this should not be the sole motivation for pursuing science as an individual or for funding science as a government agency or private funder.

A.6.2 Physical Science

Physical science is applying scientific reasoning to physical systems. Physical systems could be intuitively defined as those which can reasonably be assumed to be universal in nature. As an example, a physical scientific result is one which we could imagine an unknown alien race on some other planet coming up with the same result without communicating with us. Physics

such as particles and fields, chemistry of atoms, molecules, and materials, and to some extent, biological systems are all part of physical science.

Complex biological systems and social systems I consider to be outside the realm of physical science, because they are too unique to our specific situation on earth. The way animals behave is influenced by the fact that humans study them, and complex systems can quite reasonably be called unique. Part of the universality of physical systems is that at some level, each of their members is interchangeable or indistinguishable. As far as we can tell (and by the laws of quantum mechanics), every electron in the universe is actually identical to any other. The same cannot be said for markets of goods, human psyches, and even animals and plants.

This universality is what initially drew me into physics and what continues to fascinate me. I like the fact that the results we get, the truths that we claim to access, are independent of political opinions, manipulation, and particular situations. It is important to distinguish the scientific results from the policy decisions around science. A scientific result could be something about the nature of a particle, but a policy decision would be whether or not to devote a lot of resources towards doing an experiment. The first is independent of politics, the second is not.