Charmed Particle Photoproduction and a Search for Magnetic Monopoles

Michael Koratzinos

Imperial College

London

A thesis submitted to the University of London

for the degree of Doctor of Philosophy.

January 1991

Abstract

A search for Magnetic Monopoles: A large superconductive detector for cosmic magnetic monopoles has been designed, tested and has collected data for a year. It incorporates three coils, one of which has a novel configuration that maximises the sensitive volume but responds with a non-unique signal to a monopole.

The effective area of the detector is $0.18m^2$ and the effective observation time 6600 hours. One event has been seen that is consistent with a monopole passage and cannot be attributed to an extraneous cause. This event was registered in the coil with the poorer signal discrimination. The monopole flux implied from this single event (equal to $7.2\times10^{-12}cm^{-2}s^{-1}sr^{-1}$) is three orders of magnitude higher than the Parker Bound.

Measurement of charmed particle photoproduction: An investigation of charmed particle photoproduction is presented. Data are taken from NA14/2, a high energy photoproduction experiment at CERN, that has collected 17 million triggers during the period 1985-86 at a mean photon energy of 95GeV. High statistics charmed signals have been extracted, with the high resolution silicon vertex detector that was incorporated into the NA14 spectrometer playing an important role.

The channels investigated are $D^0\rightarrow K\pi$ and $D^+\rightarrow K\pi\pi$. The results presented are from 4.3 million triggers directly processed through a 3081/E emulator farm. The values measured for the contributions to the charmed photoproduction cross section over the energy range $40-160$ GeV are:

$$\sigma_{D^0} = 233\pm31\pm43nb \qquad \text{and} \qquad \sigma_{D^+} = 68\pm15\pm13nb$$

The total charm photoproduction cross section estimated from the above two contributions is

$$\sigma_{c_{total}} = 393\pm46\pm84nb$$

The cross section variation with energy is also investigated. The data exclude charmed quark masses of 1.2GeV or below.

Introduction

This thesis covers two independent topics in the field of High Energy Physics, spanning over six years of work in the field. These two topics are presented in chronological order: Part 1 deals with a Monopole Detector experiment situated at the Blackett Laboratory, Imperial College, London, over the period 1983-85. Part 2 refers to a Photoproduction experiment named NA14/2 situated at CERN, the European Centre for Particle Physics, in Geneva, that has collected data in the period 1985-86.

A search for magnetic monopoles

The publication of a candidate monopole event by an experiment at Stanford in 1982 stimulated intense experimental and theoretical activity in the field of magnetic monopoles. Grand Unification theories had already suggested that massive magnetic monopoles (mass of $\approx 10^{16} GeV$) were created in the very early stages of the universe, but astronomical observations limited their present flux to very low, almost undetectable, values.

A number of groups around the world were set to repeat the original Stanford experiment with much larger sensitive areas, and the Imperial College group was one of them. This part of the thesis describes the design, testing, data collection and analysis for the Imperial College monopole detector. A review of monopole theory is also included.

A measurement of charmed particle photoproduction

The charmed quark, the fourth heaviest quark, has been known for some time now (it was discovered in 1974). However, the physics of charmed particles is far from fully exploited to date. NA14/2 is one of the experiments set to exploit charm physics. It is a fixed target experiment situated in the North Area of CERN utilising a photon beam for the charmed particle creation. The photon

beam is derived from the proton beam of the Super Proton Synchrotron (SPS) – an underground accelerator of 8km circumference – at CERN. Charm photoproduction has the advantage of being quite simple in the framework of QCD and therefore it allows QCD predictions to be tested with experimental data.

In this part of the thesis we discuss the analysis leading to a measurement of total charm particle photoproduction cross section. The theory of charm photoproduction is also presented together with a detailed description of the experiment (in both hardware and software levels).

Author's contribution

This thesis covers a wide range of topics, containing the contributions of numerous people. Whenever possible, these contributions are acknowledged by referencing the original publication. The specific contribution of the author in this work includes: The formalization of the superconducting monopole detector theory; some work in the data acquisition system and the signal analysis of the Imperial College monopole detector; the Monte Carlo simulation for the above detector; the 3081/E emulator project of NA14/2; the tagging simulation; and, finally, the analysis leading to the measurement of the charmed particle photoproduction cross section.

Contents

Figures

Tables

Part 1

A Search for Magnetic Monopoles

Chapter 1

MAGNETIC MONOPOLES - INTRODUCTION

1.1 On the monopole nature

1.1.1 Monopoles and Grand Unified Theories

It is now more than 50 years since Dirac first set magnetic monopoles on a sound theoretical foundation [1]. Since then, the notion of magnetic monopoles has evolved considerably from early 'classical' Dirac-type monopoles to the superheavy magnetic monopoles of Grand Unified Theories (GUTs).[1] Current interest in GUT monopoles originated from the work of 't Hooft and Polyakov [2] in 1974, in which they showed that monopoles are a natural and inevitable consequence of any semi-simple, non-abelian gauge group, which when spontaneously broken eventually yields a U(1) factor. Moreover, the prediction that magnetic monopoles exist does not depend on the mechanism of the symmetry breakdown; nor does it matter whether gravitation becomes unified with the other particle interactions at the unification scale [3].

In the SU(5) model of grand unification, monopoles are identified with topological 'knots', or defects in the Higgs field. A monopole corresponds to a configuration of the Higgs field in which the direction of the field in group space at different points in real space is topologically distinct from a configuration in which the Higgs field points in the same direction in group space everywhere in real space, as shown in Figure 1.

---

[1] Grand Unified Theories attempt to unify all fundamental forces of nature. According to these theories, all the fundamental forces are unified above a certain energy threshold. However, below that threshold symmetry breaking occurs and the individual nature of the forces is revealed. The main difficulty of GUTs is the incorporation of the gravitational interaction into the theory due to its, as yet, obscure nature. Detailed account of Grand Unified Theories is beyond the scope of this work. The presentation to follow will, therefore, simply present some results of GUTs that have implications on monopoles without attempting to verify them. The reader should refer to the supplied references for more details and for the original publication bibliography in the case of review articles.

no monopole monopole

→: Direction of Higgs field in group space at a space point.

Figure 1: Higgs field configuration and the GUT monopole

The physical properties of GUT monopoles are manifold. The strength of their magnetic charge, g, is related to the unit of electric charge, e, by the Dirac quantization condition:

$$eg = (n/2)hc \qquad where\ n = \pm 1, \pm 2, \pm 3, ...$$

or

$$g = ne/2\alpha = \left(\frac{137}{2}\right)ne = 68.5ne$$

where α is the fine structure constant. Note that the Dirac condition does not exactly symmetrise electric and magnetic charges. The elementary magnetic charge is predicted to be much stronger than the elementary electric charge. Therefore, two magnetic charges a certain distance apart feel a force which is $(137/2)^2$ greater than that between two electric charges the same distance apart. The dimensionless coupling constant, $g/hc = \alpha(g/e)^2 \approx 34$ would thus be stronger than that of any known elementary force.

The mass of the magnetic monopole, M_m, in minimal SU(5) grand unified theories, is related to the unification mass, M_x by[2]

$$M_m = (1/\alpha_G)M_x \approx 10^{16} GeV$$

where α_G ($\approx$ 1/42 in SU(5)) is the coupling strength of the grand unified interaction at the unification mass scale of the theory. Note that this corresponds to a macroscopic weight of $\approx$ 20 nanograms,

---

[2] Throughout this work, mass, whose units are GeV/c², will simply be given in GeV.

roughly the mass of a bacterium, or 10^6 Joules, roughly the kinetic energy of a charging rhinoceros.

However, monopole mass predictions vary from one theory to another [4]. In supersymmetric theories or supergravity the monopole mass may be somewhat higher, $M_m \approx 10^{16} - 10^{19} GeV$. In Kaluza-Klein theories of higher space-time dimensionality the monopole mass may be close to the Planck mass, $M_p \approx 5 \times 10^{19} GeV$. The monopole mass could also be much lighter than this, e.g. for monopoles in theories of partial unification, masses of order $M_m \approx 10^5 - 10^{10}$ are expected, whereas even in minimal SU(5) or SO(10), monopole masses of $M_m \approx 10^{10}$ may exist.

1.2 Cosmology

If monopoles are extremely heavy as suggested by GUTs, they cannot be produced in accelerator experiments nor in any contemporary sites of the universe. We must therefore look to the early universe (time, t, less than 10^{-35}sec or temperature, T, corresponding to energies greater than 10^{14}GeV) as the birth place of those objects. So the creation of monopoles is very much a cosmological issue. This chapter will only attempt to indicate and briefly discuss the major cosmological issues concerning monopoles, the full treatment being outside the scope of this work. Review articles on the subject can be found in reference [5].

Two basic schemes of monopole creation in the early universe have been proposed:

1. as topological defects in the spontaneous symmetry breaking (SSB) phase transition, or
2. thermal pair production in very energetic particle collisions.

1.2.1 The Kibble process

Process (1) is known as the Kibble process [6]. When the temperature of the universe was above the critical temperature for the SSB transition, the grand unified theory was unbroken, no monopoles were present but there were thermal fluctuations in the direction of the Higgs field. As the universe cooled down below the critical temperature, T_c, of the GUT phase transition, it is thermodynamically favourable for the Higgs field to align itself uniformly over large distances, but causality does not per-

mit the correlation length of the Higgs field, ξ, to be greater than the horizon length, L_h (about 6×10^{-24}cm at that time). And different regions cannot, in general, consistently be brought together without trapping a topological defect, a "hedgehog", where the regions meet. Such a topological defect in the Higgs field, we have seen, is a magnetic monopole. It is expected that of order one monopole will be created per horizon volume. Since the horizon volume contains a net baryon number of $10^{15}GeV/T_c$ which equals about 1000, a monopole abundance of the $O(10^{-3})$ per baryon would be expected. However, such a monopole abundance corresponds to a present mass density of the universe of about $10^{12}\rho_c$, where ρ_c is the critical density of the universe. This is clearly disastrous and is sometimes called the monopole problem. This also implies the universe to be only 30000 years old when it has cooled to its present temperature of 3K. Monopole-antimonopole annihilation for the period after the Kibble process is expected to be small, due to low probability of monopole capture, and therefore ineffective in further reducing the monopole density. (Although there has been a suggestion [7] that annihilations might be enhanced by the presence of e^+e^- plasma, in which case they may be able to reduce the initial abundance to a reasonable level)

There have been many mechanisms proposed for reducing the abundance of SU(5) monopoles, e.g. by delaying the GUT phase transition to a much lower temperature, such as $T_c = 10^{10}GeV$ thus allowing for a much larger horizon length, L_h and, hence, correlation length ξ; or with the introduction of an intermediate phase [8], $10^3 > T > 10^{14}$GeV, in which space behaves like a high temperature superconductor. The monopoles are confined to flux tubes (since flux in a superconductor is expelled and therefore tends to collapse to a flux tube), so there will be in effect a linear potential between monopoles and antimonopoles, which would greatly enhance the annihilation rate. This mechanism goes on till $T < 10^3$GeV, when the universe re-enters its normal phase. Although the flux tubes are very effective at making monopoles and antimonopoles annihilate, some monopoles would possibly be unable to find an antimonopole to pair up with so that when the universe became 'normal' again, a small but perhaps interesting abundance of monopoles might have been left over. Here 'interesting' means large enough to be potentially detectable, but not so as to be ruled out by the astrophysical constraints.

New inflation theories [9] on the other hand, developed for solving problems in cosmology associated with the early universe —and its evolution until today— (e.g. the horizon[3] and flatness[4] problems) offer a solution to the monopole problem as well. In those theories the entire visible universe today started from a single 'bubble', or nucleation site, which 'jumped' due to either thermal fluctuation effects or quantum tunneling from the $<\phi>=0$ state to a $<\phi> \neq 0$ state of the Higgs field, in an epoch when the energy density of the universe was contained predominantly in the form of vacuum energy density. In this scenario the universe supercooled below the critical temperature T_c, driven by a large vacuum energy density,and underwent exponential expansion i.e. 'inflating' in size by a factor of $\approx e^{100}$. In other words, there was a large effective cosmological constant in that period of the early universe which caused it to expand exponentially. Since the period of exponential expansion increases the Higgs correlation length, ξ, to greater than 10^{10}light-yr, monopole production by the Kibble mechanism is highly suppressed, so that ≤ 1 monopole is expected to be produced in this process. Eventually the cosmological constant thermalised; it was turned into radiation and the universe reheated. After reheating, the baryon number of the universe could have been produced by baryon number violating processes, whereas monopole production is negligible.

In more complicated inflationary scenarios, though, monopoles can be produced by the Kibble process during the inflationary period itself (toward the end) [10] or in a subsequent phase transition [11]. Whether an interesting number of monopoles can be produced in this way remains to be seen. It is fair to mention also that, although the inflationary universe has some extremely interesting features, it suffers from the lack of a good candidate for a realistic Higgs potential, which serves as the foundation of the particle physics of the theory [4].

---

[3] The uniformity (to about 1 part in 10^3) of the cosmic background radiation extending over many horizon lengths, something that is not explained by the standard model of cosmology.

[4] $\Omega\ (=\rho/\rho_c)$ is conservatively known to lie in the range $0.01 < \Omega < 10$. But since, within the evolution of the universe according to the standard model of cosmology, the value $\Omega = 1$ is an unstable equilibrium point, this implies that Ω at the time of the GUT transition ($T \approx 10^{14}$ GeV) had to be equal to 1 to within one part in 10^{49}. This fine-tuning of Ω is again hard to explain within the standard model.

1.2.2 Thermal monopole production

Process (2), thermal monopole production in very energetic particle collisions, occurs irrespective of whether the universe underwent inflation. However, the number of monopoles produced in this way is expected to be small due to the enormous suppression by the Boltzmann factor: Monopoles cannot be produced until the SSB has occurred, T $< T_{SSB}$ and $M_{monopole} = T_{SSB}/\alpha \approx 100 T_{SSB}$ hence the Boltzmann factor $e^{-2M_{monopole}/M}$ equals exp(−few 100). However, although thermal production does not look too promising, the uncertainties are such that it is not impossible that an interesting number of monopoles could have been produced this way [12].

To summarize the present status of the theoretical predictions for the monopole abundance − Pick a number, any number! In the absence of a meaningful prediction all we can do is treat the cosmological prediction of the monopole abundance as a free parameter.

1.3 Monopole abundance

1.3.1 Cosmic monopole abundance

Even though interactions of monopoles with each other in the early universe are expected to be relatively rare, interactions with other forms of matter will occur, which will keep them in kinetic equilibrium until the decoupling of that particle species. The interactions of monopoles with e^+,e^- allow them to stay in equilibrium until the epoch of e^+e^- annihilation (T about 0.5MeV or t = 10sec), after which time they effectively cease to interact, except gravitationally. After EM decoupling, T ≈ 1eV or t ≈ 10^{13}sec, matter began to clump due to gravitational instability and the absence of radiation pressure. As structures such as galaxies began to form, monopoles too should have clumped with the (fermionic) matter. However, as they are effectively collisionless they cannot dissipate their kinetic energy and condense into tightly bound objects whose formation involved dissipation, such as disks of spiral galaxies, stars, etc.[5] So we would expect to find monopoles in all structures whose formation did not involve

[5] Later though, after stars have formed, they can capture monopoles which impinge upon them, an issue which will be discussed later.

dissipation, from galactic halos to clusters of galaxies to superclusters. The baryonic densities for those objects relative to the universe as a whole are about 10^5, 100 and a few respectively. Within these objects we would expect a local enhancement of the monopole flux, of the order of the density contrasts. Since we live in a galaxy we would expect a local enhancement of the order of 10^5 over the average monopole flux in the universe. However, the magnetic field of our galaxy would eject monopoles lighter than about 10^{20}GeV in less than the age of the galaxy [13]. Therefore we would not expect to find a concentration of monopoles in our locality. The magnetic fields in clusters are only potent enough to eject monopoles that are lighter than about 10^{15}GeV. Since our galaxy is not a member of a cluster and is only in the outskirts of the Virgo supercluster where the density contrast is about 1, we would expect the monopole flux in our vicinity to be due to monopoles which just happen to be passing through the galaxy (and perhaps an equal number which are bound to the Virgo supercluster). Thus the local flux should be about equal to the average cosmic flux.

It has also been suggested [14] that there might be an enhancement of the local flux of monopoles, particularly near our solar system, of up to six orders of magnitude relative to the average monopole flux in the galaxy, if the sun were to gravitationally capture a cloud of monopoles and keep them in orbit. However, more detailed calculations [15] have shown that the monopole flux in our solar system is unlikely to be enhanced by more than a factor of $\approx$ 50 over the galactic monopole flux.

1.3.2 Monopole capture by astronomical objects

Although monopoles would not be in objects such as stars, planets, etc. *ab initio*, since the formation of these objects clearly involved kinetic energy dissipation, they can be captured by them. Monopoles passing through matter predominantly lose energy by electronic interactions (energy loss due to the eddy currents they induce) [16]. Monopoles less massive than about 10^{20}GeV will lose sufficient energy when passing through neutron stars and white dwarfs to become captured. Monopoles less massive than about 10^{18}GeV will lose sufficient energy in main sequence stars to become captured. Jupiter-sized objects can stop monopoles as massive as 10^{16}GeV, and the earth can stop very slow moving or light monopoles ($\lesssim 10^{15}$GeV). Once captured, monopoles will sink towards the centre of

the object and be supported against gravity by their thermal velocity dispersion or magnetic fields that may be present. The number of monopoles residing in an object also depends upon the importance of monopole−antimonopole annihilation. However, monopole capture does not deplete the cosmic stock of monopoles appreciably; even within the galaxy the mean free path of a monopole is 10^{40}m.

1.4 Monopole velocities

In judging the prospects for detecting a monopole, certain astrophysical considerations are relevant. One of them is their speed relative to the detector. Because of the large monopole mass, their velocities were small at the time of EM decoupling, their velocity dispersion also being small. Today these quantities should be even smaller, because of the redshift from T = 0.5MeV to T $\approx$ 3K. However, they will be accelerated by any gravitational and magnetic field they encounter. The typical peculiar (relative to the Hubble flow) velocities of objects in the universe are of the order 10^{-3}c, implying typical monopole galaxy velocities of this magnitude. Monopoles will be accelerated by gravitational fields of galaxies to O(10^{-3}c), of clusters to O(few 10^{-3}c), and of superclusters to O(10^{-2}c). Monopoles will also be accelerated by galactic magnetic fields. The galactic magnetic field, having a magnitude of about 3μG and a coherence length of roughly 10^{19}m, will accelerate monopoles otherwise at rest to velocities of

$$V_m = 3\times10^{-3}(10^{16}GeV/M_m)^{1/2}c$$

where M_m is the monopole mass. The intergalactic magnetic field strength, B_i, is known to be less than 3×10^{-11}G and will accelerate monopoles to velocities of

$$V_m = 3\times10^{-4}(B_i/10^{-11}G)(10^{16}GeV/M_m)^{1/2}c$$

Of course earth based detectors are not at rest themselves, having velocity components due to the earth's rotation ($\approx 2\times10^{-6}$c), the motion of the earth round the sun ($\approx 10^{-4}$c), and the orbital motion of the solar system through the galaxy ($\approx 7\times10^{-4}$c). Velocity components for monopoles and monopole detectors are summarized in Table 1.

Table 1: Sources of monopole and detector motion

SOURCE	MONOPOLE	DETECTOR
GRAVITATIONAL		
Virgo supercluster	$2 \times 10^{-3}c$	$2 \times 10^{-3}c$
The galaxy	$10^{-3}c$	$7 \times 10^{-4}c$
Solar system	few $\times 10^{-4}c$	$10^{-4}c$
Earth	$3 \times 10^{-5}c$	
MAGNETIC		
galactic B-field	3×10^{-3} $(10^{16}$ GeV$/M_m)^{1/2}c$	
intergalactic B-field	$\leq 3 \times 10^{-4}$ $(10^{16}$ GeV$/M_m)^{1/2}c$	

From the above discussion and the table it is clear that one should expect monopole velocities of at least a few $\times 10^{-3}c$. [In the unlikely case that most of the monopole flux in our neighbourhood is due to an orbiting monopole cloud, typical monopole velocities would be of the order of a few $\times 10^{-4}c$, implying monopole−detector relative velocities of the order a few $\times 10^{-4}c$].

1.5 Flux limits

Because of their three extraordinary properties − macroscopic mass, hefty electromagnetic charge, and ability to catalyse nucleon decay − GUT monopoles, if present even in small numbers, will make themselves astrophysically conspicuous. This leads to stringent astrophysical bounds on their flux.

1.5.1 Mass contribution limits

First consider the mass they contribute. Cosmology favours a flat universe to a high degree. (Ω, the ratio of the cosmic mass density to the critical density is about 1). Big bang nucleosynthesis implies that the baryonic matter alone is not enough to close the universe. For $\Omega_b \geq 0.2$ deuterium is underproduced and helium overproduced compared to what is observed. Monopoles can easily provide the additional mass to achieve closure density. It is known (from estimates of the deceleration parameter and the age of the universe) that $\Omega \leq 2$. Therefore, the mean monopole flux in the universe should be less than

$$F_m < 5\times10^{-15}(\beta_m/10^{-3})(10^{16}GeV/M_m)cm^{-2}sr^{-1}s^{-1}$$

where β_m is the fraction of the speed of light a monopole is travelling with and M_m is the monopole mass.

If monopoles are clustered in galaxies the monopole flux in our vicinity could be bigger than the value quoted above. We obtain a limit by assuming that monopoles cannot contribute more mass density locally than we observe[6]; as indicated by galactic rotation curves, the mass within 30kpc of the centre of the galaxy is less than 10^{12} solar masses. The monopole flux in the galaxy is thus constrained to be less than

$$F_m < 3\times10^{-11}(\beta_m/10^{-3})(10^{16}GeV/M_m)cm^{-2}sr^{-1}s^{-1}$$

In fact we can do slightly better than this. Careful modeling of the galaxy establishes that the contribution of the halo material to the local mass density can be no more than about 1/30 of the total mass density, i.e. the disk component dominates the local mass density. [If monopoles are clustered in the galaxy they must be in the halo since they have no mechanism for dissipating their gravitational energy and condensing into a disk.] This leads to a more stringent limit

$$F_m < 10^{-12}(\beta_m/10^{-3})(10^{16}GeV/M_m)cm^{-2}sr^{-1}s^{-1}$$

---

[6] There is a discrepancy between the amount of matter observed and accounted for. This is known as the dark matter problem, i.e. the apparent mystery that the luminous galactic matter can only account for about 10% of the total galactic mass. Monopoles are therefore a tempting candidate for the solution of the dark matter problem, although there are certainly other possible candidates, such as small dim stars, massive neutrinos, photinos, gravitinos, axions, etc.

At this point we should re-emphasize that the bound based upon the average flux in the universe is probably the relevant one since the galactic magnetic field will eject monopoles as discussed previously.

1.5.2 Monopoles and astrophysical magnetic fields

Because of their magnetic charge monopoles will respond to any magnetic fields they may encounter and usually gain kinetic energy at the expense of magnetic field energy. Parker used this principle to place limits on the flux of monopoles, referred to as the Parker bound [17]. If magnetic monopoles are slow moving, then in traversing the galaxy they will undergo large deflections, gaining kinetic energy at the expense of magnetic field energy. If, on the other hand, monopoles have high velocities then only small deflections in their trajectories are incurred in traversing the galaxy. In this situation, (i.e. for high monopole velocities) it can be shown that for an isotropic distribution of equal numbers of north and south monopoles passing through the galaxy, no net kinetic energy gain or loss occurs for the distribution to first order in B. However, to second order in B there is a net kinetic energy gain by the monopole distribution. [This is so since such monopoles will enter a coherent domain of the galactic magnetic field with an energy larger than the energy they would pick up if accelerated from rest in that domain. Thus to first approximation they are as likely to put energy into the magnetic field as to take it out. But as they cross the domain they tend to be deflected so that there is a second order effect in which they do drain energy from the magnetic field.] This net kinetic energy gain, $<\Delta E_k>$ is given by

$$<\Delta E_k> \approx 1/4(gB_{gal}l)(\beta_m/\beta_{mag})^2$$

where B_{gal} is the galactic magnetic field ($\approx 3\mu$G), $l \approx$ 300pc is the coherence length of the magnetic field and $\beta_{mag} = (2g\,B_{gal}\,l/M_m)^{1/2} = 3\times10^{-3}(10^{16}GeV/M_m)^{1/2}$.

The origin of the galactic magnetic field is believed to be due to dynamo action. The time to generate/degenerate the field is of the order of the galactic rotation period of 10^{8}yr. Requiring that monopoles accelerated by the galactic field, thereby acquiring kinetic energy, do not drain the magnetic field energy in a time shorter than this, results in the following constraints on the monopole flux.

For $\beta_m < \beta_{mag}$:

$$F_m < 6\times10^{-16}(B_{gal}/3\mu G)(3\times10^7 yr/t_{reg})(r/30Kpc)(300pc/l)cm^{-2}sr^{-1}s^{-1}$$

where t_{reg} is the regeneration time for the galactic magnetic field and r the size of the magnetic field region in the galaxy. This portion of the Parker bound is independent of M_m and β_m.

For monopole velocities $\beta_m > \beta_{mag}$:

$$F_m < 6\times10^{-16}(\beta_m/3\times10^{-3})^2(M_m/10^{16}GeV)(B_{gal}/3\mu G)(3\times10^7 yr/t_{reg})(r/30Kpc)(300pc/l)cm^{-2}sr^{-1}s^{-1}$$

This portion of the Parker bound is dependent on both M_m and β_m. The monopole flux limits, F_m, for the Parker bound as a function of monopole velocity expressed in contours of constant M_m are shown in Figure 2.

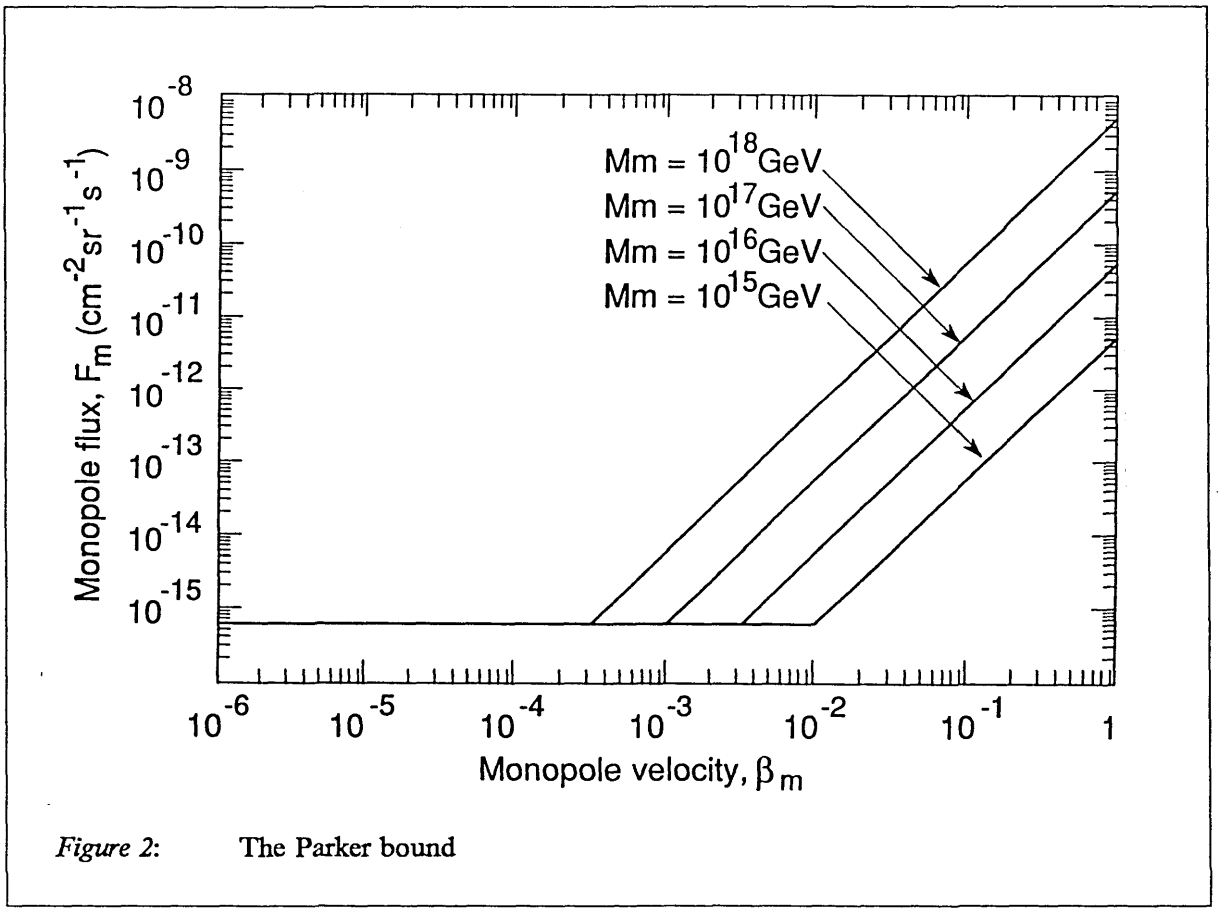


Figure 2: The Parker bound

The Parker bound is subject to a criticism, however. The implicit assumption made in deriving Parker-type bounds is that monopoles respond incoherently to the astrophysical magnetic field. If they can respond coherently (as in the case of magnetic plasma oscillations), then, in an analogy to an LC circuit, energy is transformed from one part of the system to the other in an oscillatory fashion. Thus the kinetic energy they extract from the magnetic field will be returned half a cycle later. In this case the field energy is not drained, and the bound in question can be circumvented [18]. In fact, in some sense, the monopoles are participating in the maintenance of the magnetic field. However, if these oscillations are to persist, they must maintain both spatial and temporal coherence, otherwise the oscillations will undergo Landau damping. This results in a **lower** bound to the monopole flux (if coherent effects are to be important):

$$F_{gal} > 10^{-13}(M_m/10^{16}GeV)(1kpc/l)^2 cm^{-2}sr^{-1}s^{-1}$$

While the Parker bound may be circumvented by coherence effects, there are many issues to be considered before this can be considered to be a serious possibility: how such oscillations got started in the first place; how the required spatial and temporal coherence is maintained in the face of inhomogeneities known to exist in the galaxy; and how the numerous damping mechanisms can be avoided. Although it is not entirely unlikely that some sort of coupled system between the monopoles and the galactic magnetic field could have evolved to a steady state system, when transient terms of the formation of the galaxy have damped out [4], the most difficult problem of such theories at present is the observational one: the present experimental limits on the monopole flux preclude the scenario unless the monopole mass is much smaller than 10^{16}GeV.

The Parker argument has been applied [19] to the survival of the weaker magnetic fields which are observed in rich clusters (10^{-7}G and coherence length $\approx$ 1Mpc). The result is a much more stringent bound

$$F_m < 10^{-18} cm^{-2}sr^{-1}s^{-1}$$

which, however, is somewhat less reliable than the Parker bound as our knowledge of intercluster magnetic fields is not a secure as that of the galactic magnetic fields.

Other authors have applied Parker's logic to magnetic fields in white dwarfs, neutron stars, and peculiar A stars. Also other limits have been obtained by the effect of the galactic electric field induced by the galactic magnetic field, as well as by consideration of the effect of the electric field on extra-galactic cosmic rays [4]. Because of the additional assumptions involved, these bounds, while more stringent than the Parker bound, are much less certain.

1.5.3 Monopole nucleon decay catalysis

The most spectacular effect associated with superheavy magnetic monopoles is their apparent ability to catalyse nucleon decay at a prodigious rate. The rate of energy release of this process is enormous; in a neutron star the energy release is 10^{13}Joule s^{-1} per monopole. Astrophysical objects such as stars, planets, etc., capture some or all of the monopoles that strike them. Once captured they sink to the centre and accumulate there at a rate proportional to the monopole flux. There they catalyse nucleon decay releasing about 1GeV per catalysis event. The energy is thermalized and radiated – in the IR(for planets), visible(for main sequence stars), UV (for white dwarfs), or X-ray (for neutron stars). In the process, some of the energy may also be released in neutrinos.

Using the observed photon flux from a variety of astrophysical objects – the earth, Jupiter, white dwarfs and neutron stars, one can place very stringent limits on the flux of monopoles in the galaxy. The most stringent and probably the most reliable comes from neutron stars [5].

$$F_{gal} < 10^{-21}(\sigma/10^{-28}cm^2)^{-1}cm^{-2}sr^{-1}s^{-1}$$

where σ is the cross section for this process (expected to be of the order of 10^{-28}cm^2, typical value of a strong interaction cross section). This is generally considered reliable, because the various uncertainties due to particle physics and astrophysics are better understood and believed to be smaller for neutron stars. For example, when the catalysis involves nuclei, there are angular momentum barriers, etc.; also at small monopole–nucleon velocities there may be threshold effects. In a neutron star nuclei do not exist, and the relative velocities are of the order of the speed of light.

The neutron star limit quoted above is actually based on three observations: The measured X-ray fluxes from individual objects; the negative results of X-ray searches for bright, nearby stars; and the

use of the measured intensity of the soft X-ray background to limit the integrated luminosity of all the old neutron stars in the galaxy. The limit has also been checked against other possibilities which are found not to be important: e.g. the possibility that neutron stars would eject monopoles they capture, that annihilations might be important, or that essentially all the catalysis energy would be released in neutrinos.

Table 2: Flux limits for typical GUT monopoles

MASS CONTRIBUTION:	
galactic mass contribution	$\leq 10^{-12}$ cm^{-2}sr^{-1}s^{-1}
achieve Universe closure density	$\leq 5 \times 10^{-15}$ cm^{-2} sr^{-1}s^{-1}
MAGNETIC:	
intercluster fields	$\leq 10^{-18}$ cm^{-2}sr^{-1}s^{-1}
Parker bound	$\leq 6 \times 10^{-16}$ cm^{-2} sr^{-1}s^{-1}
magnetic plasma oscillations	$\geq 10^{-13}$ cm^{-2}sr^{-1}s^{-1}
NUCLEON DECAY CATALYSIS:	
white dwarfs	$\leq 10^{-18}$ cm^{-2}sr^{-1}s^{-1}
neutron stars	$\leq 10^{-21}$ cm^{-2}sr^{-1}s^{-1}

The limit based upon catalysis in white dwarfs is also quite stringent (three orders of magnitude above that of the neutron star one, therefore still three orders of magnitude below Parker's limit), with uncertainties that are not too much worse than for neutron stars – and they involve different physics and astrophysics. This is another indication towards the validity of limits below the Parker bound. The limits based on nucleon decay catalysis are not, however, without their criticisms: one might argue that we do not really have enough understanding of the detailed structure of the neutron star or white dwarfs to say confidently how monopoles would affect them; also, the rate for catalysis could be much slower, in which case the catalysis rate rather than the capture rate will be the relevant one in an astrophysical context, and the bounds obtained in this way would be insignificant [20].

A selection of the most important flux limits for typical GUT monopoles (mass 10^{16}GeV and velocity 10^{-3}c) can be seen in Table 2.

1.6 Motivation for monopole search

Monopole search experiments were performed since the early days of modern science without ever obtaining convincing positive results. Moreover, what made the monopole case even more difficult was the lack of good theoretical ground for such objects. Since those days a new monopole identity has emerged, the GUT monopole; GUT theories have given the monopole its missing reason for existence. As we have seen, the theoretical bounds for a monopole flux leave the question of the possibility of monopole detection completely open; it is the astrophysical limits that suggest that such a search will be difficult. Nevertheless, the possibility of confirming a GUT prediction is very tempting, and a lot of experiments have looked for monopoles in recent years. But the interest on monopole experiments increased tremendously when Cabrera, in 1982, reported the observation of a candidate monopole event [21], using a detector that incorporated a novel technique. Various groups around the globe decided to build experiments with increased sensitivity to verify Cabrera's observation. One of those groups was the Imperial College group, whose detector is the subject of this part of this thesis.

Chapter 2

MONOPOLE DETECTORS

The design of a monopole detector must take account of signatures that are unique to a monopole's passage as well as easily discriminate against spurious signals. There exist two major types of detectors, each using a different characteristic feature of the monopole.

One is the ionization detector that relies on monopole interaction with matter at the microscopic (atomic) level; to discriminate monopoles from other particles it relies on the fact that monopoles are expected to be slow moving compared to other elementary particles so that delayed coincidence can be used to veto anything else but monopoles. (Electrically charged particles moving with monopole-like speeds would be absorbed by matter very quickly.)

The other type of detector relies on the monopole's magnetic charge for its detection; more specifically it uses Faraday induction for monopole detection. This is the theoretically simpler method of searching for monopoles: it only relies on a single Maxwell equation and in principle needs no assumptions from more modern physics.

Comparing the two techniques, the use of ionization is clearly simpler from an engineering point of view, using techniques that have been established for years, but is far more complicated from a theoretical point of view. This is due to the fact that monopole interaction with matter is not well understood and that detector efficiency drops dramatically at low monopole velocities. So one has to make extra assumptions about the probable monopole speeds, interaction with matter, etc. But, provided one is confident about these, the ionization type detector has the advantage of a much larger sensitive area to cost ratio over the inductive type.

On the other hand, the inductive detector, as a newly developed technique, has a lot of engineering problems that need to be solved (regarding detector stability, background noise, etc.) and it is much

less cost efficient. But since it is sensitive to the only monopole property that is assured, its magnetic charge, it needs much less theoretical assumptions on the monopole nature. This was the approach that the Imperial College group decided to adopt for its monopole detector.

Using a hybrid detector, incorporating an induction and an ionization detector in coincidence, would have been useful, once monopoles had been detected, for the information it could provide on monopole matter ionization. Had the experiment seen no monopoles on the other hand, such a detector would not have had any clear advantages over the inductive type. Such an arrangement was intended to be used for the Imperial College monopole detector, but the idea was later dropped as too expensive since the introduction of the window frame (WF) geometry that meant that the detector would be sensitive to monopoles intersecting the detector at all angles.

2.1 Inductive detectors

A monopole passing through an isolated closed conducting loop induces a current that generates magnetic flux equal to h/e, equivalent to $2\phi_0$, where ϕ_0 is the flux quantum of superconductivity ($\approx 10^{-7}\mathrm{Gcm}^2$). If the loop has a self inductance L_d the current induced is given by:

$$I_d = \frac{2\phi_0}{L_d}$$

If it misses the loop the induced current is zero. A single turn coil has typical inductance of about $1\mu\mathrm{H}$ so induced currents are of the order of about 1nA.

The easiest way to detect such a signal is by using superconducting techniques. This is because:

1. The signal from a monopole passage through a superconducting coil is a DC one. Therefore, one can work at very low bandwidths to eliminate noise.
2. Johnson noise is low.
3. One has available highly sensitive amplifiers in the form of commercially available SQUIDs (Superconducting Quantum Interference Devices) [22].[7]

---

[7] Theory of SQUID operation will not be discussed here – it will be sufficient to regard them as very sensitive current to voltage converters.

4. The signal is effectively independent of monopole velocity (since we are working at low bandwidths).

5. Superconducting shields can provide a highly stable ambient magnetic field for the detector coil.

On the other hand a superconducting detector suffers both complication and expense from the necessity of the use of cryogenics.

A room temperature induction detector on the other hand is not entirely out of the question. For a bulky coil (i.e. conductor thick compared to the radius of the loop) the Johnson noise can be reduced to the point where a signal to noise (S/N) ratio of 10 is achievable for monopole velocities greater than $1.4\times10^{-4}c$. These calculations have been made by M.Price [23]. He has also done a detailed analysis of the amplifier and integrator for the loop, and again he finds that a S/N ratio of 10 is achievable in principle.

The major difference between a superconducting and a non-superconducting detector is that, whereas in the superconducting case a very low bandwidth is used, a room temperature detector requires fairly high bandwidth. A characteristic frequency of a typical pulse resulting from a monopole passage is of the order of 10MHz. Therefore, the detection problems in the two cases are completely different. The superconducting case has been the subject of intense work over the past few years. Largely motivated by the signal seen by Cabrera, who used a superconducting detector, several groups – of which Imperial College is one – have built detectors based on the same principle. With respect to the other case, however, no one, to my knowledge, has successfully built a non-superconducting detector. If such a detector can be made to work reliably, though, it would have the huge advantage of having a much better sensitive area to cost ratio over the superconducting one.

2.2 Superconducting detectors

In a monopole search we try to measure signals corresponding to fluxes of the order of

$$2\phi_0 = 2.07\times10^{-7} Gcm^2$$

The intrinsic noise of the SQUIDs used for the current to voltage conversion corresponds to about $10^{-2}\phi_0$ through the detector coil (for a typical SQUID-coil arrangement). So a signal from a monopole passage is well within the detecting capabilities of commercially available RF SQUIDS. The change of flux due to a monopole, however, corresponds to a very small change in field compared to typical ambient field values. For a typical loop (one that has an area of the order of $10^2 cm^2$) a change of the order of 10^{-9}G to the ambient magnetic field could mimic a monopole signal. Changes in the Earth's magnetic field (which is about .5G), typically of the order of 10^{-3}G, and other electromagnetic disturbances in a laboratory are orders of magnitude larger than this.

Mechanical movement of the loop in an inhomogenious magnetic field could also fake a monopole signal. For a typical loop sitting in an ambient magnetic field that has spatial variations of the order of 10^{-4}G/cm, a movement of about 1 micron could produce a monopole-like signal.

The major technical problem then is not so much in sensitivity to signal, as in background shielding. the detector has to be sited in a spatially and temporally extremely stable ambient field. A way, therefore, has to be found to reduce the effect of ambient magnetic field spatial and temporal variations.

One way to do this is by attenuating the ambient magnetic field in the detector vicinity. This is achieved by using magnetic shields made from high permeability alloys such as mu-metal. These can attenuate the ambient magnetic field and its temporal fluctuations by about 4 orders of magnitude, an attenuation clearly not adequate by itself. Also, spatial inhomogeneities, although strongly attenuated, continue to be a problem. Two further techniques could be used to decrease the effect of those inhomogeneities: One is giving emphasis to mechanical rigidity of the detector framework structure so as to minimize detector movement in an inhomogeneous magnetic field. The other is to use a higher order

gradiometer[8] design instead of a single loop as the detector coil. A gradiometer is much less sensitive to both spatial and temporal variations of the surrounding magnetic field – the higher its order the less sensitive it is – but there is a price to pay in that in general, the self inductance of the gradiometer would increase with its complexity, hence making the signal from a monopole passage smaller. So there is a limit in how complicated a gradiometer design could be.

Another way of stabilising the field at the detector coil is to surround it with a superconducting shield. When a tube becomes superconducting it traps whatever flux was present inside it at the instant it became superconducting and holds it constant. If there are no holes in the tube this will give rise to infinite dynamic attenuation (although zero static attenuation). In practice holes are needed (for the SQUID RF lines, helium inlet, etc.) but still dynamic attenuation of more than 10^5 is possible.

This scheme is not without its disadvantages, however. Firstly, although the total flux is held constant, redistribution of flux within the superconducting shield can occur and such sudden flux jumps mimic monopole passages. Not much is known about the behaviour of trapped flux at these low fields (the residual field inside good mu-metal shields is about 10^{-5}Gauss) but certainly at higher fields there is evidence that thermal or mechanical shock can cause sudden flux redistribution inside the shield.

Although it now seems that the probability of a flux jump in such low fields is very low, it is potentially the most serious source of spurious signals and one has to make sure that there is enough redundancy in the information collected by the detector to discriminate against such events. Thus accelerometers and temperature monitors are essential in such a monopole detector arrangement. Sampling at a high bandwidth is also desirable. If one believes that shock induced events happen instantaneously, the better time discrimination one gets in this way could be used to time-correlate monopole-like signals in the SQUID outputs with the signals from various detector monitors and thus exclude them as spurious signals caused by environmental disturbances rather than genuine monopole passages. However, shock-induced events happen some time after the shock that caused them (what-

---

[8] A gradiometer is a coil which has segments wound to opposite directions so as to compensate variations in the ambient field to any order, depending on its complexity. An nth order gradiometer is insensitive to the $(n-1)$ spatial derivative of the field. Therefore, a second order gradiometer would compensate fields in up to linear terms.

ever that time may be - milliseconds, seconds or even minutes) so we have to introduce a 'dead time' after each spike in the accelerometer monitor that corresponds to the timescale of such a shock-induced transition; if an event occurs during this 'dead time' period after a spike in the accelerometer monitor this event is not considered as a genuine monopole candidate.

Another disadvantage of using a superconducting shield is that it couples to the detector loop with the result of diluting the - otherwise unique - monopole signal: A monopole passing through the coil will also induce current in the shield. The flux applied to the detector coil will then be reduced to $a2\phi_0$, (where $|a|$ is less than 1) from its value of $2\phi_0$, for the isolated coil case. The inductance of the coil will also be reduced from its free space value L_r to bL_r say. Thus as a result of a monopole passage we must now expect a current change:

$$\Delta I_d = 2a2\phi_0/(bL_r + L_s)$$

where L_s is the inductance of the shield. Moreover, a will depend on the orientation of the monopole trajectory and the point at which it crosses the plane of the detector coil. Even if the monopole misses the loop but passes through the shield the induced current will have a non-zero opposite sign value. (A more quantitative and extensive treatment of monopole signals in the presence of superconducting shields can be found in chapter 3.)

There are a number of ways to deal with this loss of signal uniqueness. One is to have the detector loop diameter much smaller than that of the shield. (This is what Cabrera originally did). Thus the coupling of the shield to the coil is weak and the monopole signal almost unique. However, this is achieved at the expense of sensitive area to available volume ratio. And since the limiting factor for induction detectors is cost, which, amongst other things, determines the size of the cryostat, there is a lot to be gained by optimizing the use of the available volume, and different monopole detector groups have come up with similar solutions to this problem. The idea is to replace the simple coil arrangement with a more complicated gradiometer design. Such an astatic configuration is more weakly coupled to the superconducting shield. Consequently it is possible to increace greatly the fractional area of the shield occupied by the coil without blurring the characteristic signal expected from a monopole too

severely. So whereas Cabrera's original area was 1/16 of that of the shield, the IC astatic coils cover 52% of the shield's cross section. A final advantage of an astatic coil is that its self inductance is smaller than that of a single coil of the same overall area and consequently the induced current is larger..

Using this astatic coil arrangement a ten-fold increase in the detector area over the simple single coil arrangement is possible. This is an improvement, but still most of the available volume inside the superconducting shield is not being used. If it is desired to make use of as much available volume as possible (and we are confident that our detector will be free of spurious signals) then we can use the coupling between the detector and the superconducting shield in a constructive way. The IC detector incorporates such an idea: we call it the window frame (WF) coil [24]. The signal is no longer unique (in fact it is simply related to the length of the monopole track inside the shield) but the sensitive area approaches 2,000cm^2, more than a order of magnitude increase over the astatic coil case.

Another way to get around this superconducting shield problem, is not to use one at all. Then the monopole signal is unique but the residual magnetic field is not as stable as in the superconducting shield case. The use of higher order gradiometers together with emphasis on mechanical stability could probably reduce the noise to acceptable levels. Another virtue of this approach is that the problem of flux jumps, that were feared as a potential cause for spurious events in the superconducting case, cannot occur.

Chapter 3

THEORY OF THE SUPERCONDUCTING DETECTOR

3.1 Single superconducting loop

Consider a single superconducting loop. In the absence of the superconducting shield, the signal of a monopole track would be:

zero, if the track misses the loop, or

$2\phi_0$, if the track passes through the loop.

In the presence of the shield, these signals are modified by the screening currents that are set up in the shield when current flows around the loop. The closer the loop radius to the shield radius, the stronger the mutual inductance between them and, therefore, the more the signal gets attenuated. Consider a single turn loop of area A situated concentrically inside a long cylindrical superconducting shield of unit cross sectional area. The signal of a monopole track inside the superconducting shield and in a direction perpendicular to the detector coil plane will be:

$-2A\phi_0$, if the track misses the loop, or

$2(1-A)\phi_0$, if the track passes through the loop.

This is because the shield now contributes to the signal by a factor A, the ratio of the areas of the shield to the coil. As we let A go to zero, that is let either the coil area to be much smaller than the shield, (which is what Cabrera originally did), or we let the shield radius go to infinity, the problem reduces to the isolated loop case.

To derive a formula for the signal for any track geometry we proceed as follows: We can represent a monopole track by an infinite dipole string. Then, the change of flux produced by a monopole passage is given by:

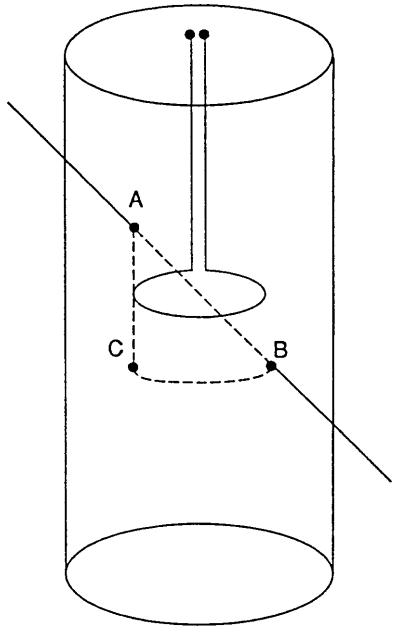


Figure 3: Monopole signal to a coil in a superconducting shield

$$\Phi_m = M \cdot Bdl$$

where B is the field per unit current that would be produced along the track if the coil system carried unit current (the reciprocity theorem of electromagnetism). M, the magnetic dipole moment is given by $M = 2\phi_0/\mu_0$. To calculate the change of flux, consider the line integral of B.dl around the closed path ABC (Figure 3). For a track that does not intersect the loop wiring (and, therefore, any currents)

$$\int_{ABC} B \cdot dl = 0$$

So the signal due to a monopole passage, ϕ_m is just

$$\phi_m = -\left[\int_{BC} Bdl + \int_{CA} Bdl\right]$$

(refer to Figure 3). B does not vary with the radial angle, θ, therefore only the line integral from C to A contributes.

Since the answer only depends on the entry and exit points of a monopole track, a scalar potential can be defined so that the change in flux can be found by simply subtracting two numbers and one does not have to evaluate the line integral every time. This is possible despite the fact that B is not a conservative field provided that one is careful to take into account discontinuities arising from intersecting currents. This potential, V, is arbitrarily taken to be zero at a point far away from the coil plane (at $-\infty$) and is equal to $V=\int Bdz$. The exact solution of $\int Bdz$ is quite involved, requiring the solution of non-trivial integral equations. It can be approximately solved, though, using a technique that gives answers accurate to within 7%, surprisingly accurate for its simplicity [25].

The monopole signal, therefore,is just given by the difference in the value of V over the entry and exit points adding $2\phi_0$ for a track that has intersected the loop.

The shape of the single coil scalar potential for a certain (loop to shield area) ratio can be seen in Figure 4. The asymptotic value of the function at positive distance from the loop plane is equal to A, the ratio of areas of the loop to the shield, as we have discussed previously. In the case of Figure 4, which represents a coil residing in the same position with respect to the shield as the AAP coils (discussed in the next section) incorporated in the actual detector having a diameter of 0.726 of the shield diameter, this is equal to $(0.726)^2=0.527$. Also the gradient of V depends on how close the loop wire is to the shield; the larger the loop radius compared to that of the shield, the sharper the potential rise. This is due to the fact that the screening currents for such a case are more concentrated near the loop plane, therefore their effect, the change in V, is more pronounced in that vicinity.

Therefore, for a single coil arrangement, the bigger the coil used, the less well defined the signal becomes. This is not desirable, since the main aim of a inductive monopole detector is to maximise the area while retaining good monopole signature – which could be a unique signal. There is a way around this problem, however; instead of a single coil we can use a slightly more complicated arrangement which we call the astatic asymmetric pair (AAP).

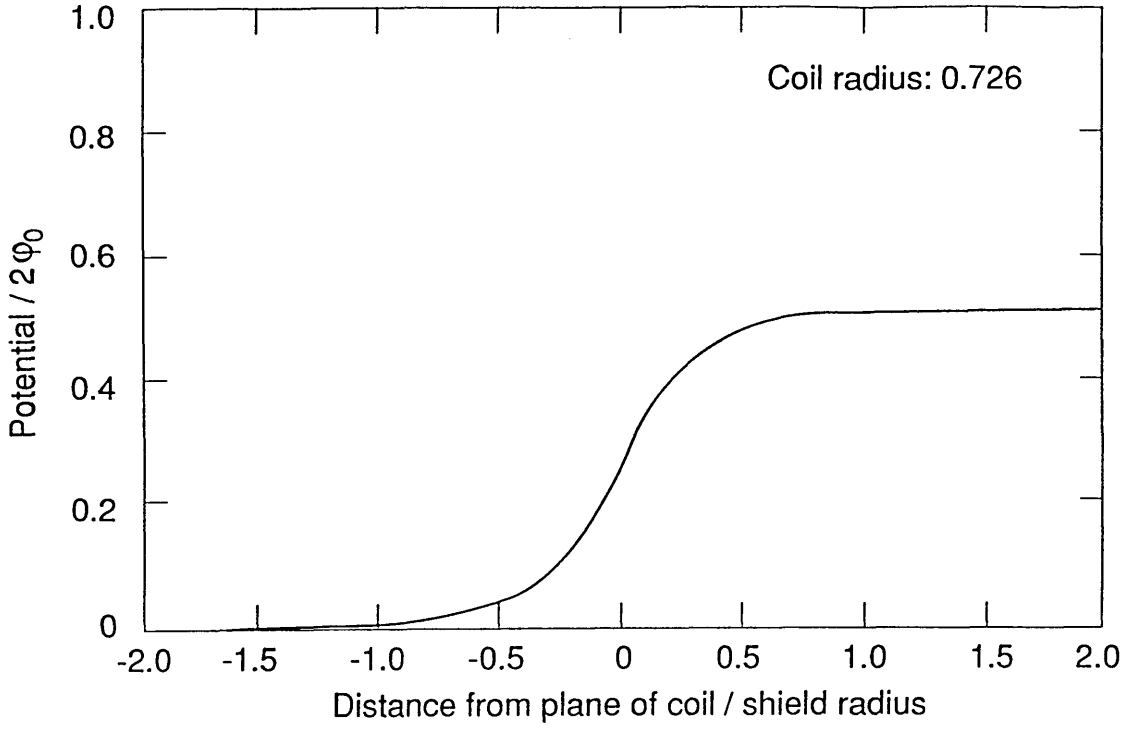


Figure 4: The single coil potential

3.2 The astatic asymmetric pair

The AAP is a coil configuration that tries to minimize the mutual inductance with the shield, so as to permit bigger detector loop radii, and therefore maximize the effective $2\phi_0$ area. This is achieved by having two loops of equal (area × number of turns) connected in series but in opposite sense. In this case there are, therefore, three classes of a monopole signal. One is around zero, corresponding to a near miss, the other is around $2\phi_0$, corresponding to monopole tracks that pass through the main coil but miss the counter turn coil, and the last is around $-2(n-1)\phi_0$, (where n is the number of turns in the counter coil), corresponding to monopoles passing through the counter turn coil. The resulting configuration, the AAP, has a significantly lower mutual inductance with the shield compared to the single coil of the same diameter; the screening currents are now localized to very near the AAP plane. This can be easily seen if one considers the potential function of this arrangement: the potential of the

AAP is a superposition of the potential of the single outside loop, minus the potential of the inside loop times n. As we can clearly see from Figure 4, the potential in this case will vanish to zero far away to either side of the AAP whereas in the vicinity of the loop plane it will have a non-zero value due to the fact that the potential of the counter coil increases less rapidly than the potential of the main loop (since the counter turn coil is further away from the shield). The shape of the resulting potential in the particular case of a 10 turn counter coil (which is what was used in the IC detector configuration), can be seen in Figure 5. It peaks at a distance of .3 of the shield radius and falls rapidly to effectively disappear 2 radii away. Its maximum absolute value is about $0.115\phi_0$, so its maximum contribution to the AAP signal is $0.23\phi_0$.

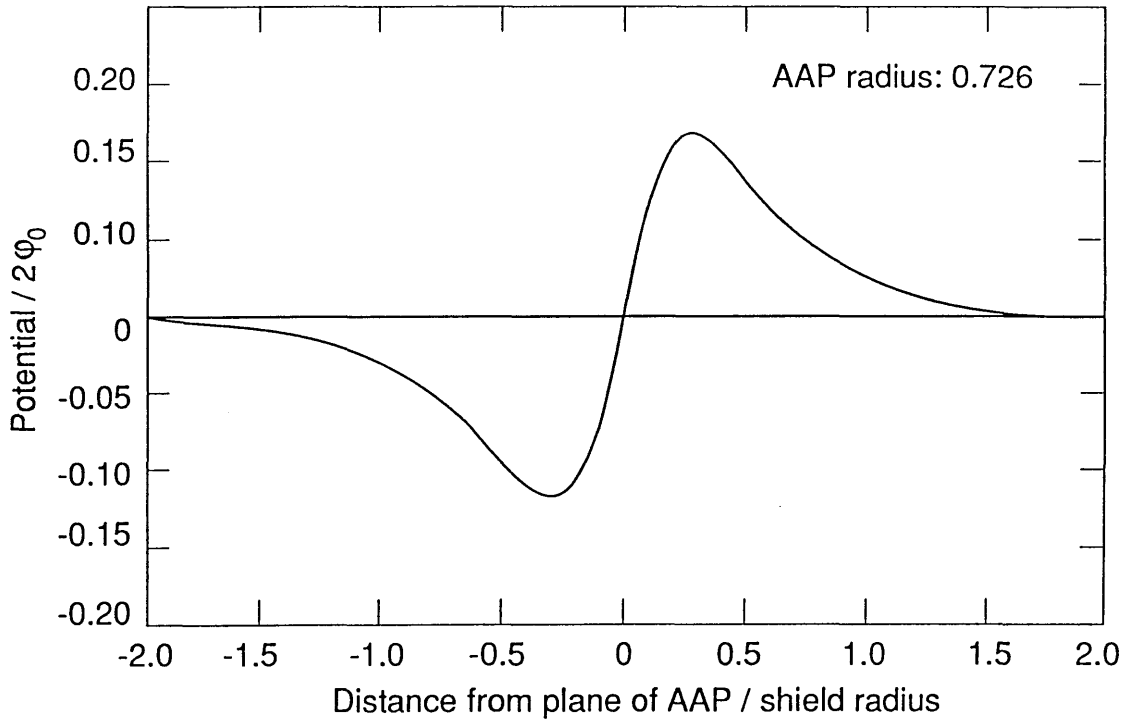


Figure 5: The AAP scalar potential

If one is interested in minimising this maximum contribution to the AAP signal in order to improve the uniqueness of a monopole signal by optimising the number of counter turns, then the

AAP should have two counter turns. This is because n = 2 is the case that the counter loop radius is closest to the main loop radius (and, therefore, the gradients of the two contributions to the overall potential are more closely matched, giving rise to more effective cancellation of the two contributions). Had the above analysis been realized while designing the experiment, an n = 2 AAP coil would have been used instead of the n = 10 one that was actually incorporated.

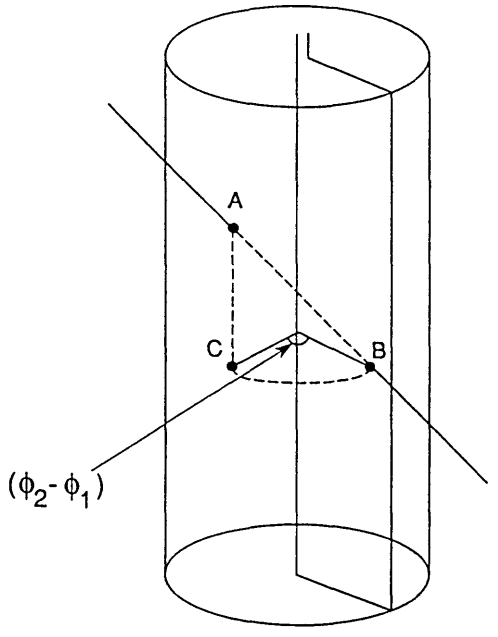


Figure 6: The WF configuration

3.3 The window frame configuration

The window frame configuration tries to maximize the sensitive area of a superconducting monopole detector by tightly coupling to the superconducting shield. It achieves that at the expense of signal uniqueness. The WF configuration can be seen in Figure 6. To calculate the signal of a monopole passage, we proceed as before: The change of flux associated is, as we have seen

$$\Phi_m = M \cdot Bdl$$

To calculate B in the WF arrangement we use the fact that the outside wire is very close to the shield. In this case the induced screening current density due to the return wire will be localized near the wire and inside the screen. This effectively cancels the field produced by the coil wire everywhere except very close to it. This means that B to a very good approximation arises solely from the current in the central wire. This is given by the infinite wire formula:

$$B = \frac{\mu_0 \phi}{2\pi r}$$

Note that the above formula does not get modified by end effects even though the central wire is of finite length. This is because the surrounding screen helps to restore symmetry; one can easily see that by considering a closed loop inside the detector that passes from near the centre and continues towards near one of the end caps. The fact that symmetry does not permit polar angle-varying B fields along the perpendicular and that Ampere's law has to hold ensure that B has to be independent of z and given by the infinite wire formula.

To calculate the change of flux consider the line integral of B.dl around the closed path ABC in Figure 6. For $\delta\phi$ less than π the loop does not intersect any currents so

$$\int_{ABC} B \cdot dl = 0$$

So the signal due to a monopole passage, ϕ_m is just

$$\phi_m = -\left[\int_{BC} Bdl + \int_{CA} Bdl\right]$$

And since the contribution of the track CA is zero this time we finally get

$$\phi_m = \phi_0(\phi_2 - \phi_1)/\pi$$

If $\delta\phi$ is greater than π then the loop encloses unit current from the central wire so

$$\int_{ABC} B \cdot dl = \mu_0$$

Thus

$$\phi_m = \phi_0 \frac{[2\pi - (\phi_2 - \phi_1)]}{\pi}$$

Note that the change in flux is always equal to

$$\delta\phi = \frac{\phi_0(\phi_2 - \phi_1)}{\pi}$$

provided we take the smallest path around BC. In this case the scalar potential, V, is simply a linear function of the asimuthal angle, ϕ, independent of z or any other parameters.

$$V = \phi_0 \frac{\phi}{\pi}$$

and a monopole pass would leave a signal given by

$$V_2 - V_1 = \phi_0 \frac{(\phi_2 - \phi_1)}{\pi}$$

Chapter 4

THE IMPERIAL COLLEGE MONOPOLE DETECTOR

4.1 General outlook

The detector is of the superconducting kind and incorporates three coils mounted in a mechanically rigid non-conducting (tufnol) framework attached to a lead superconducting shield. This shield sits in a liquid helium environment inside a large cryostat surrounded by a set of five mu-metal shields which are used to attenuate the ambient magnetic field. The whole construction sits on air loaded springs that provide some mechanical isolation from the laboratory environment. Monitors of all relevant environmental quantities are attached to the detector and a HP-9816 mini computer handles the on line data analysis as well as the monitoring of the environmental changes and SQUID outputs. The whole detector is situated in a small room adjacent to the main workshop at the basement of the Blackett Laboratory in South Kensington. The location is far from ideal since it is subject to a high level of electromagnetic, magnetic and mechanical disturbances.

A schematic diagram of the detector can be seen in Figure 7. The general philosophy of the experiment was revised after the 1983 monopole conference in Michigan in which no new candidate events were reported. By that time it seemed increasingly unlikely that Cabrera's original candidate event was a real one. Therefore the main aim of the experiment was to reduce the upper bound of the maximum monopole flux rather than expect to see monopoles and measure some of their properties. So the original proposal for the detector layout incorporating three loops in a coincidence-anticoincidence arrangement and good signal discrimination was dropped in favour of a design with much poorer discrimination but 15 times bigger sensitive area. More specifically, the sensitive area for the detector is now about 1,800cm^2 but the probability for a coincidence is only 12% and a

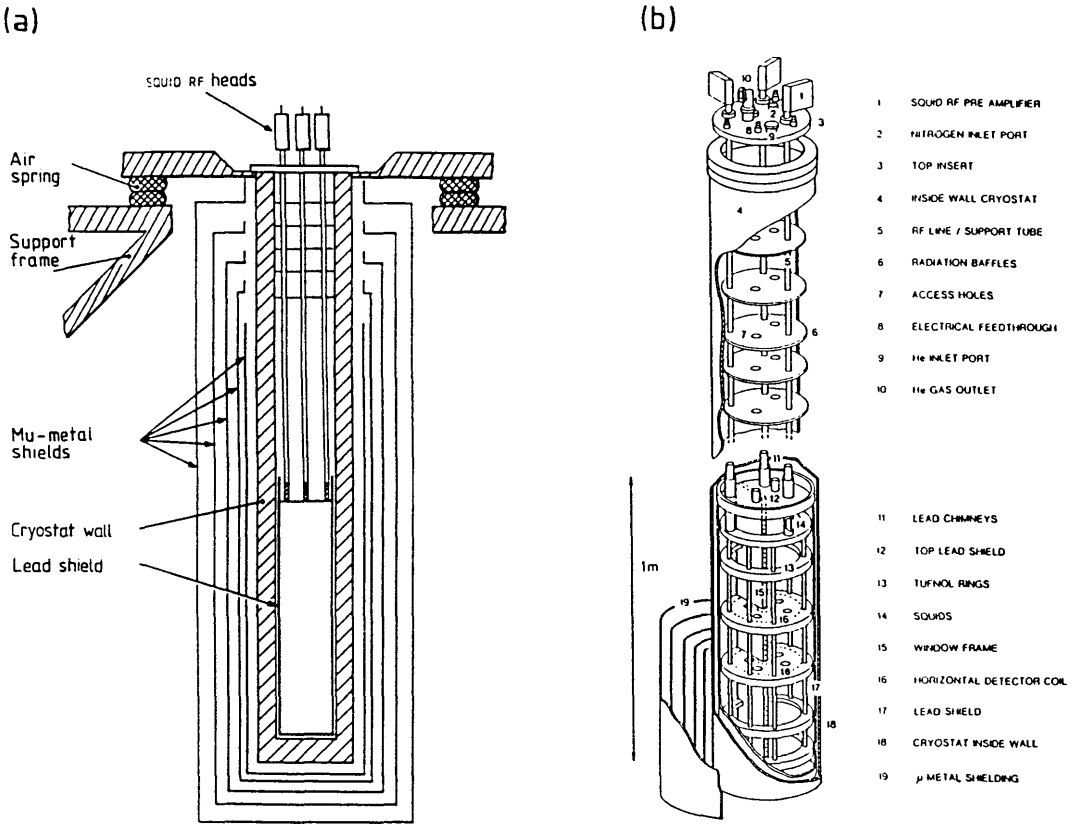


Figure 7: The Imperial College Monopole Detector. This figure shows a schematic cross section of the detector cryostat and magnetic shields. The overall height is about 3m. The detector loop assembly residing inside the lead shield can be seen in Figure 9.

monopole passage would give a range of signals between 0 and ϕ_0. With this arrangement, if monopole-like signals were to be observed, it would be rather hard to support any conclusive arguments concerning their nature, or even to be certain that they did not have spurious causes. If on the other hand no monopole candidates were to be seen, a useful upper bound on the monopole flux could be set, an order of magnitude lower than the one that would have been achieved using the original detector for the same observation time. To use such an approach for a detector, of course, one had to be confident that such an arrangement would be free of spurious events.

Another alternative was to run without a superconducting shield. Such an arrangement would have permitted a sensitive area as big as the present one, with prospects of good signal discrimination

as well as adequate redundancy: The information obtained in the form of two-fold and even three-fold coincidence would give good rejection against spurious events provided that the noise could be kept at a reasonable level. Although preliminary tests looked promising, the idea was dropped since it required a long testing period and time considerations favoured the more conventional approach.

4.2 Cryostat

A large capacity (100 litre) vapour-cooled liquid helium cryostat is used (made by Cryogenic Consultants Ltd.). It has an inner diameter of 250mm and is 2400mm deep; it is of aluminium and fiberglass construction, and has vapour cooled radiation shields. In normal operation of the detector, the helium level is kept above the lead shield, and the upper half of the cryostat serves as a reservoir with usable volume of 45 litres. The boil-off rate (5 litres of liquid helium per day) is low enough to allow intervals between helium refills to be about one week (typical refill intervals were 7 to 8 days). This is an important consideration since the detector is quite unstable after each refill, taking a few hours to settle down. This introduces a dead time typically of the order of $2-3$ hours. This time is usually spent on off-line analysis of the data obtained during the last period (since data acquisition and analysis is performed on the same computer). This dead time could have been avoided if a continuous helium transfer arrangement had been in operation. However, this is a difficult task and was not considered worthwhile.

The helium level should not be allowed to drop below the top of the lead shield, as this results in a considerable increase in SQUID output noise. Helium level monitoring inside the cryostat is made directly, by measuring the resistance of a wire that runs perpendicular to the cryostat; the part of the wire submerged into helium becomes superconducting whereas the rest of the wire remains normal allowing a linear relationship between helium level and resistance to be established. An indirect measurement is obtained from the volume of the helium vapour coming out of the cryostat through a gas meter.

Helium gas boiling off from the cryostat is collected and liquified allowing some 85 to 90% of the original helium to be re-used. This results in a considerable financial saving, but also introduces unde-

sirable pressure fluctuations through the recovery line to the cryostat, especially pressure shocks resulting from the compressor being switched on. This upsets the detector, especially the WF which is most sensitive to pressure changes. To reduce the effect of a pressure shock, a passive low pass filter was introduced in the helium recovery line before the cryostat. This consisted of a large, soft, rubber bag of about $1m^3$ capacity, which introduced a capacitance to the circuit, in series with an impedance in the form of a long narrow rubber tube. After some experimentation using different tube sizes, the final length of the tube was chosen as 10m and its diameter 4mm. The bag fills up completely shortly after refill, but empties itself gradually almost completely (depending on the actual pressure of the recovery system). If the bag is less than full we don't see any pressure variations; this is because the capacitance of the bag is practically infinite when it is not fully expanded (a small change in the amount of helium present in the bag does not change the pressure at all). When the bag is full on the other hand, a small change in the amount of helium present introduces a pressure change in the bag which forces some helium through the narrow tube so that the capacitor 'discharges' slowly through the high impedance line. When the helium recovery compressor runs, we see slow variations in pressure of the order of 2-3mbar. Thus, low frequencies do pass through, and they are visible at the SQUID outputs, but their effect (slow drifts in the SQUID outputs) is easily discriminated against monopole signals.

The cryostat, together with the mu-metal magnetic shields, is mounted on air sprung suspension (springs: Firestone, airmount No. 16). The resonant frequency for vertical oscillations is about 1.5 Hz, and the response is slightly underdamped.

4.3 Magnetic and radio frequency (RF) shielding

4.3.1 mu-metal shields

A set of 5 mu-metal shields is used for ambient magnetic field attenuation (see Figure 7). These initially gave a residual field of about 10^{-3} Gauss which was then reduced to about 10^{-5} Gauss after demagnetising the shields *in situ* . The demagnetisation was done as follows: three turns of wire were passed through the neck at the top and out through a 25mm diameter hole at the bottom of the

shields. A current of 1 Amp at 0.25Hz was reduced slowly to zero. Thus the alternating magnetic field of progressively diminishing amplitude created, drove the shields through a large number of hysteresis loops. The resulting axial field profile inside the cryostat can be seen in Figure 8. The large error bars of the low field points are a result of the sensitivity of the magnetometer used for these measurements being comparable with the magnitude of the field measured in this region. The horizontal component of the field was typically of the same order of magnitude $(1-2nT)$[9], but the sensitivity of our available magnetometer did not allow more detailed measurements.

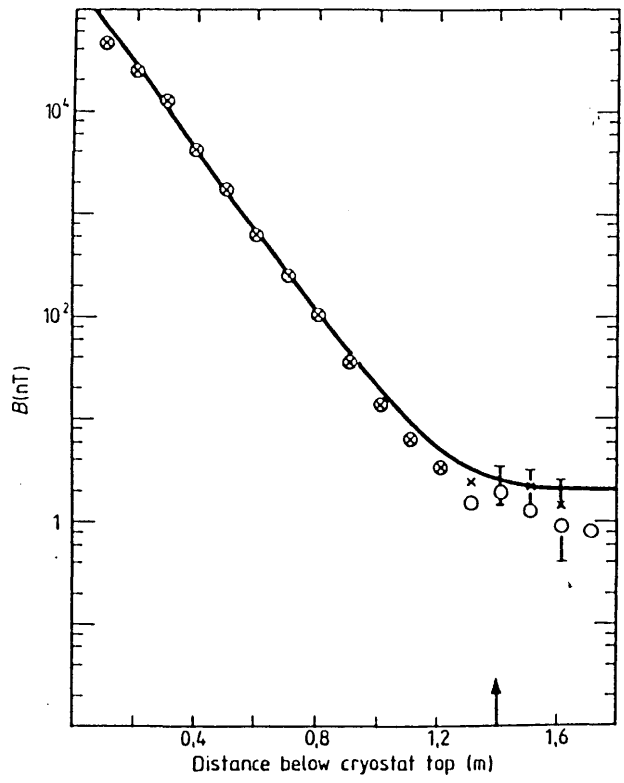


Figure 8: Field profile inside the cryostat

[9] 1 gauss (G) = 10^{-4} tesla (T)

4.3.2 Superconducting shield

A lead superconducting shield was used. It is a cylindrical barrel of height 1m and diameter 25cm and sits in an ambient magnetic field of about 10^{-5} Gauss (the ambient field inside of the cryostat).

It fully encloses the detector coils and has as few holes as possible: there are six small holes at the top of the shield together with five larger ones. The six small holes are for the plastic screws that locate the top of the detector framework to the shield and the five larger ones are access holes (three of them are for the SQUID RF lines/support tubes). Each hole is protected by high lead tubular lids at least four times longer than its diameter. These attenuate variations in flux leakage due to changes in the external magnetic field.

After various tests it was concluded that flux jumps inside the shield would not be as serious a problem as it was originally feared. The lead shield seemed to be extremely quiet.

4.3.3 RF shielding

The RF lines to the SQUIDs run within stainless steel tubes; these, together with the aluminium outer wall of the cryostat, should provide a high degree of RF shielding. It is important, too, to prevent the leakage of RF interference into the cryostat along leads to thermometers, test coils, etc. All of these leads emerge through vacuum seals on the cryostat top plate and in normal operation of the detector, when no connection to them is necessary, they are capped off. On the occasions when the test coils are used, the connections are made through a heavily shielded filter.

4.4 Detector framework

The detector coil support frame is made from 13mm thick tufnol disks and rings of 234mm diameter, joined together by vertical 13mm diameter tufnol rods and also by a rectangular tufnol plate, all held together with tufnol pins. Emphasis was given to mechanical rigidity for reasons already explained. The framework just fits in the lead shield. Vacuum 'apieson' grease is used for easier fitting which solidifies at liquid helium temperatures and holds the framework rigidly attached to the shield. The top of the lead shield is attached to the detector framework by means of plastic screws and is then

soldered to the rest of the shield. As an extra precaution, several litres of a soap/glycerol mixture which becomes glassy at low temperatures, were poured into the shield at final assembly.

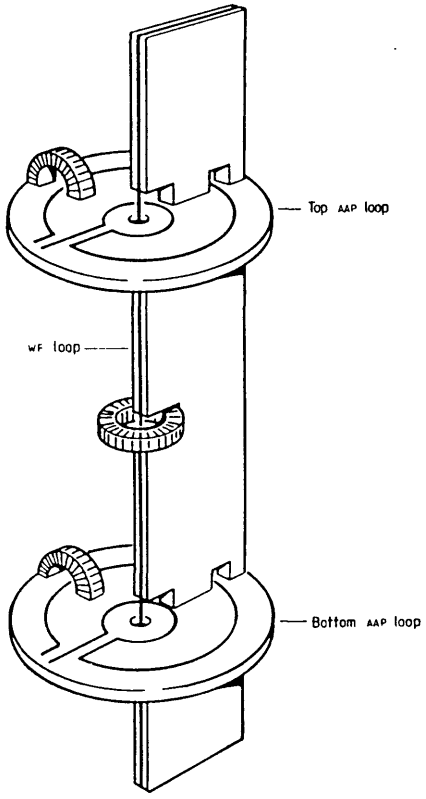


Figure 9: Detector coil and support frame configuration. The additional tufnol rings and rods incorporated in the actual design for mechanical rigidity are not shown for clarity. The three toroids seen are used for calibration.

The framework-shield unit is more or less free to vibrate inside the cryostat – it is only connected to the top insert via the three RF lines/support tubes.

A number of radiation baffles are also used to stop the 300°K radiation of the top insert from reaching the 4.2°K liquid helium so that a low boil-off rate can be achieved.

4.5 Detector coils

The detector coil configuration used can be seen in Figure 9. It consists of two AAP coils of 170mm diameter and the novel WF loop arrangement in partial coincidence. The 4π averaged area of the AAPs is 114cm^2 whereas the WF has an effective 4π averaged area of 1,800cm^2.

One turn of NbTi wire of 75μm diameter is wound around the edge of the vertical plate to form the WF loop, whilst the AAP loops (single turn main loop, 10 turn counter loop) are glued to the upper side of the tufnol discs. In an attempt to reduce the self inductance of each AAP loop, the counter turns are not wound in the coil plane, but in a solenoidal fashion.

4.6 SQUIDs

Three commercial RF SQUIDs are used (System 330, SHE corp.), one per detector loop. Their sensitivity is 2.2, 2.5, and 2.9 Volts per ϕ_0 through the SQUID coil. The quoted value for their maximum RMS noise is $7\times10^{-5}\phi_0 Hz^{-1/2}$ (in current this corresponds to $7\times10^{-12}AHz^{-1/2}$)

4.7 Calibration

One of the virtues of the inductive technique for monopole detection is that it can be calibrated directly – in contrast with the ionization technique. This is done by coupling a known flux through the detector coils. Three toroids, one for each loop, have been used for this purpose. These toroids are of 35mm major diameter and have 8mm^2 cross section. They are wound tightly with about 250 turns of 0.15mm diameter copper wire. A toroidal coil of self inductance L_t and n turns has mutual inductance L_t/n with the single turn detector loop. L_t can be calculated and also measured directly with an accuracy of about 5%, so that, by excitation with a known current, calibration of the detector loops can be made to a similar level of accuracy. (In our case, with $L_t\cong15\mu H$, currents corresponding to a flux quantum ($1\phi_0$) through a detector loop were of the order of 1μA). Another independent way of checking the correctness of the calibration was found to be at our disposal; it was conjectured that a poor connection in the top AAP gave rise to a superconducting weak link in that coil. In a circuit con-

taining a weak link, a circulating current that exceeded the critical current I_c of the weak link would initially decay rapidly and then, as it approached I_c, steps in units corresponding to flux changes of $2\phi_0$ would be visible. Such steps seen in the top AAP on one occasion agreed within 5% with the calibration.

Thus it was found that SQUID I (window frame) has a sensitivity of 8.3mV/ $\phi_0(\pm 5\%)$. This corresponds to a transformer ratio[10] of 265. For SQUIDs II and III (bottom and top AAPs respectively) the sensitivity was $3.1mV/\phi_0(\pm 5\%)$ and $2.7mV/\phi_0(\pm 5\%)$, corresponding to transformer ratios of 935 and 1020 respectively. The fact that the transformer ratio is smaller in the WF is because its inductance was smaller than that of the AAPs. The above, together with the RMS noise (which is discussed later) and the resulting S/N ratios are tabulated in Table 3. For the window frame, since the signal is a whole spectrum of values, the maximum S/N ratio is quoted.

Table 3: Detector coil calibration and indication of s/n ratio

Coil	mV/ϕ_0	RMS noise (@ 1 Hz)	S/N ratio (@ 1 Hz)
WF	8.3 ± 0.4	0.37 mV	22.4 (max)
Bot. AAP	3.1 ± 0.2	0.35 mV	17.7
Top AAP	2.7 ± 0.1	0.50 mV	10.8

4.8 Interference monitoring

In addition to the induction detector outputs, the ambient magnetic field, the amplitude of vibration, the amplitude of RF interference and the pressure above the helium bath are continuously monitored.

[10] Transformer ratio is the ratio of the sensitivity through the SQUID coil to the sensitivity through the detector coil. The sensitivity of this SQUID was $2.2V/\phi_0(\pm 5\%)$ through the SQUID coil. So 265 ϕ_0 in the detector coil couple to 1 ϕ_0 in the SQUID coil itself.

The magnetic field sensor (Domain Micro Systems Ltd., model SAM 3) is a single axis flux gate magnetometer. In a 10Hz bandwidth its peak to peak noise is less than 0.3nT. Because the magnetic shielding of the detector is least effective for the vertical field component, the sensor is mounted with its axis in this direction and outside the mu-metal shields. Field fluctuations in the laboratory are typically 20 to 60 nT, with the nearby lifts being a major source of perturbation. The magnetometer was disconnected towards the end of the main run. Earlier, extensive tests had shown the detector to be immune to external field changes up to two orders of magnitude larger than typical field variations. For this reason, and since no putative event associated with field fluctuations had been seen, the magnetometer was considered redundant.

A single axis piezoceramic accelerometer (D.J. Birchall Ltd., type A/01) is used to monitor vibration. Together with its associated charge amplifier it has a sensitivity of 30V/g, that is reduced by 20 to 30 dB in directions normal to its axis. The output noise in a 10Hz to 5kHz bandwidth is 10mV peak-to-peak, equivalent to 3×10^{-4}g along the axis. The accelerometer is bolted rigidly to the top plate of the cryostat with its axis at 45^0 to the vertical. Because of the large mu-metal shields, the detector is an efficient acoustic antenna, so that the accelerometer records also any sound in the room; normal speech levels give outputs of 10 to 100mV.

The RF SQUID sensors that are used are prone to RF interference, particularly when close to their operating frequency of 19MHz. Although considerable care has been taken with RF shielding, as an additional precaution RF interference is continuously monitored. A wide-band amplifier (Pascall Electronics Ltd., model LNR 404) that has a bandwidth of 1 to 100 MHz and 40dB gain is used with a non-resonant antenna, constructed simply from six wires of lengths between 1 and 20 m. The data acquisition system itself contributes interference of about 4V at output, and other known sources, such as capacitor bank discharges in adjacent laboratories and nearby short-wave transmitters, give outputs of about 8 to 15V.

The pressure sensor above the helium bath is a differential silicon strain pressure gauge (Druck Ltd., Type PDCR 10/L), with a sensitivity of 0.14 mV/mBar. Because short term pressure changes are of main interest, a simple brass can of 2.5 litre volume, which is not temperature controlled, provides

an adequate reference pressure. Typical cryostat pressure variations are 150μBar peak-to-peak on 0.1 to 10 Hz bandwidth; over a longer time scale, the pressure follows that of the atmosphere (but subject to temperature changes of the reference volume).

4.9 Data processing

An HP-9816 mini computer system is used to collect data from the three SQUIDs, the accelerometer, the RF monitor, the magnetometer[11] and the cryostat pressure meter through a multiplexer and a 12-bit A to D converter (Analog Devices Ltd., Model DAS 1128). Data from the 7 channels are sampled at 20 Hz. The recent history – that of the last 400 seconds – is stored at this bandwidth in a cyclic buffer whereas the maximum and minimum values from each recorded channel (the output of the three SQUIDs together with that of the various monitors) is printed in a pen-recorder format once every minute, to form a permanent record of the detector activity. The SQUID sensors, the magnetometer and the pressure sensor are all used with their associated amplifiers set to 10Hz bandwidth. Because much faster pulses could be important to the RF and accelerometer channels, their analogue outputs are taken through peak-hold amplifiers, set to 0.1s time constant. As a back-up in the case of data acquisition software or hardware malfunction, conventional pen recorders are also used for SQUID output monitoring.

Monopoles would traverse the detector loops in times much shorter than the system response time. The apparent speed of an offset, therefore, is important evidence in helping to decide if an event is genuine. Any offset rise-time slower than the system response time can be safely discarded as a spurious event. To decide on the appropriate sampling bandwidth one has to take into account two major factors; one is that the bandwidth has to be high enough for one to be able to discriminate between fast (and therefore possible candidates) and slow (and therefore spurious) events. The other is that the bandwidth has to be low enough to give a good signal to noise ratio for signal detection. Both these factors cannot be satisfied at the same time; data are first sampled at a relatively high bandwidth to get good time resolution and then digitally filtered to a much lower bandwidth for good signal discrimina-

[11] Towards the end of the final run, as we have already discussed, the magnetometer was discarded as redundant

tion.

The time of an offset of magnitude Δu can be located to within about $4\sigma/\Delta u$ samples, where σ is the RMS noise within the full sampling bandwidth; (the factor of 4 comes from converting RMS values to peak-to-peak ones). Therefore it is appropriate to sample the detector outputs to as large a bandwidth as possible to gain more information about the risetime of any events, but only up to the point that such a step could be easily recognisable, that is, it could be located to within a few samples. That implies sampling bandwidths where the noise becomes comparable to the expected signal. In our case the sampling bandwidth was chosen to be 10Hz; at this bandwidth the noise is 0.11 ϕ_0 for the WF loop, and 0.41 and 0.63 ϕ_0 for the AAPs.

On-line analysis is performed in a search for steps in any of the SQUID output channels. This is quite a different task; to allow much smaller offsets to be detected, the noise and therefore the bandwidth has to be reduced. This is done using the standard (digital) matched filter technique. Data from each channel is continuously convoluted with the characteristic signal we are trying to detect, in this case a step function (the characteristic signature of a monopole event), F(i), given by, say:

$$F(i) = -1 \quad \text{for} \quad -N \le i < 0 \quad \text{and} \quad F(i) = 1 \quad \text{for} \quad 0 \le i \le (N-1)$$

were N is related to the correlation time, T_c, of the filter and f_s, the sampling frequency by:

$$f_s = N/T_c$$

In our case N = 2000, giving a correlation time of 100 seconds. Then, if u(j) is the sampled output of one of the SQUID inputs, the result of the convolution, V(j) = u(j) $\otimes$ F(j) could be written as:

$$V(j) = \sum_{k=0}^{N-1} u(j+k) - \sum_{k=1}^{N} u(j-k)$$

From that we can derive a simple recursion relation that gives V(j + 1) from V(j):

$$V(j+1) = V(j) + [u(j+N) - 2u(j) + u(j-N)]$$

This is the relation used by the data acquisition program for quick on-line evaluation of V(j).

An offset Δu in the input produces a triangular output of peak height NΔu. (Because V(j) is calculated N samples later than the sample j, this peak follows the offset after an interval of N/f_s seconds.)

At the same time, the noise is reduced since the bandwidth is effectively narrowed. Since the signal to be detected is a DC one, the matched filter acts as a low pass filter, or an integrator, that effectively reduces the bandwidth to (something of the order) f_s/N thereby reducing the RMS noise by $1/\sqrt{N}$ if we assume that the noise is white. (In fact, this is not exactly correct; there seems to be an excess of low frequency components in the noise spectra. The noise is, however, white to a good approximation.) Therefore, the signal to noise ratio of the output of the filter is a factor $\sqrt{N}$ better than the original S/N ratio.

If the filter product, V(j), increases beyond a set threshold, it triggers the dumping on disk of the recent 300 second history of all channels before the event as well as the history of the 100 seconds after the event for future study. This corresponds to 8000 points per channel. For the 7 channels that are used this corresponds to 112Kbyte of memory or about half the memory space of a 5-inch floppy disk used for data storage. The correlation time for the convolution, as already mentioned, is 100 seconds. In other words, the step analysis is performed at 0.01Hz. The RMS noise for that bandwidth is typically between .15−.20mV. So the trigger levels are set at 1mV, 5 to 6 standard deviations away. For disk space saving reasons, the value of the signal convolution is not stored like the channel values; if needed, part of it can be calculated from the stored data.

4.10 Signal analysis

It is useful for SQUID noise output analysis as well as when doing routine noise checks to be able to do some fast signal analysis. A signal analysis program has been developed for that reason. It analyses data taken from any ADC channel at a user defined maximum bandwidth. It computes and plots (after removing any DC components of the noise and linear drifts) the frequency spectrum, power spectrum, the autocorrelation function and an amplitude histogram. A fast fourier transform (FFT)[12] algorithm is used for fast operation.

---

[12] An FFT is just a quick algorithm to compute the discrete fourier transform (DFT) of a function, exploiting the big arithmetic redundancy of the DFT and optimizing the number of required arithmetic operations.

Chapter 5

MONTE CARLO SIMULATION

The layout of the IC monopole detector is complicated enough not to permit derivation of analytical solutions for all aspects of it. In order to investigate its properties and thoroughly understand the physics of the detector, a Monte Carlo simulation was performed. Some analytical expressions for detector properties had already been derived, and the feedback obtained from comparing them with the Monte Carlo predictions served the dual purpose of testing the Monte Carlo for possible biases and better understanding of the detector physics.

The results we were mainly interested in were basically the signal amplitude probability distributions both for the AAP and the WF arrangements, their 4π averaged area, and the coincidence rates between them.

5.1 Tests

The Monte Carlo simulation was tested for bias by careful comparison with analytically calculated results. Two major tests were made. One was the comparison of the Monte Carlo-predicted 4π averaged area with the theoretical value, and the other was a comparison of the shape of one of the WF signal amplitude probability contributions against the analytically calculated function. The first test searched for any overall biases, whereas the second test, comparing the shape of two functions and not just two numbers, is a more powerful technique to search for any r or z-dependent variations of the generated flux.

5.1.1 4π averaged area

The 4π averaged area of the detector is given by:

$$A = \pi(2LR + R^2)/2$$

where R is the radius and L the half length of the detector (see appendix B). Tracks are generated from a plane that has an area $A_{pl} = 4(R^2 + L^2)$. So the prediction of the Monte Carlo for the 4π averaged area of the detector is simply A_{pl} times the ratio of tracks intersecting the detector to tracks generated.

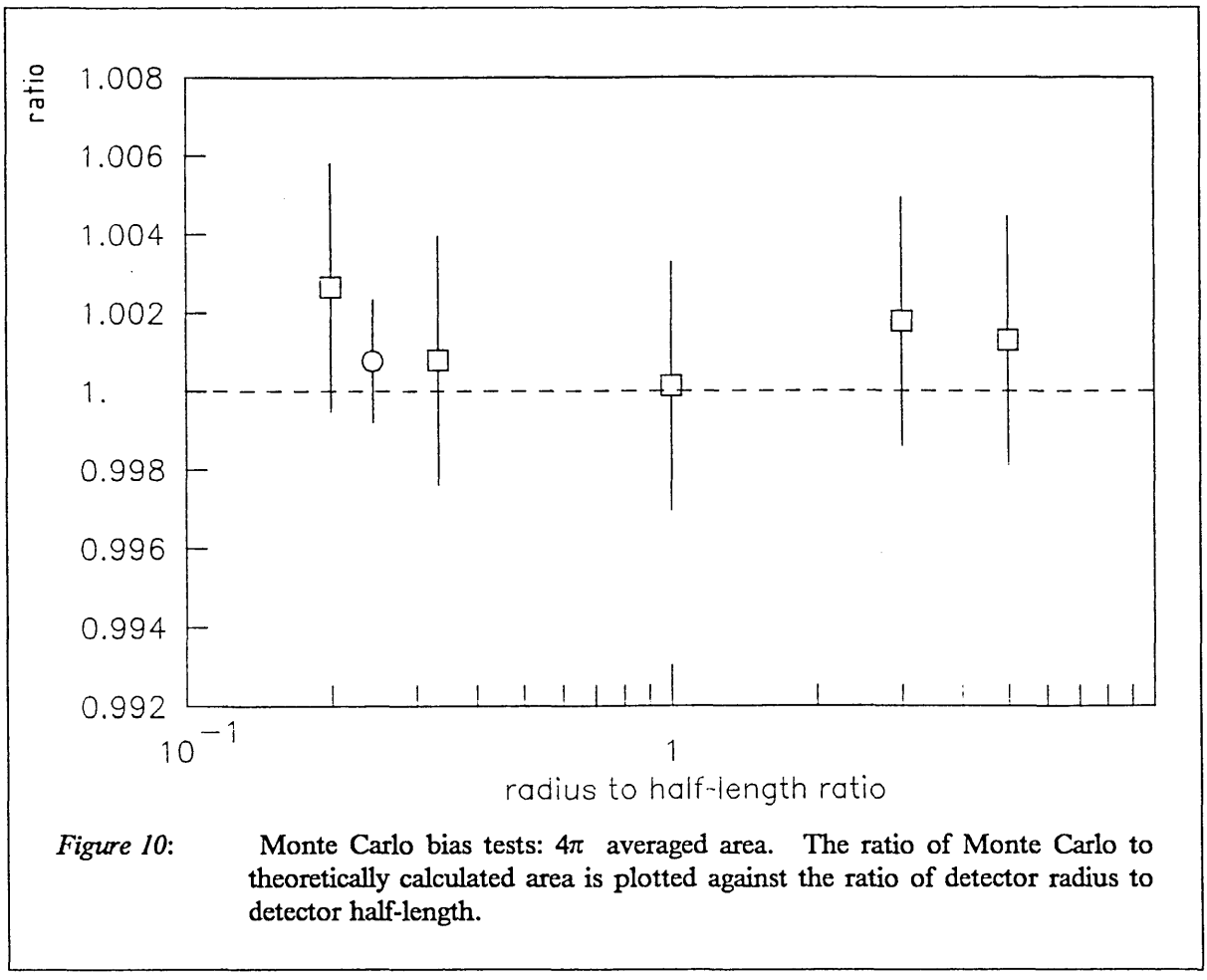


Figure 10: Monte Carlo bias tests: 4π averaged area. The ratio of Monte Carlo to theoretically calculated area is plotted against the ratio of detector radius to detector half-length.

The values of the area of the detector predicted by the Monte Carlo were checked against the theoretical ones for a variety of detector dimensions. The results are plotted in Figure 10. The data point denoted by a circle represents the calculation for the actual detector dimensions; it has higher statistics

than the rest of the points. It can clearly be seen that the predicted values are consistent with theory[13]. Therefore this test ensures that the random track generator routine is behaving sensibly and no direction-averaged biases are present.

5.1.2 Sides contribution

The calculation of the signal height probability distribution function for tracks that intersect the cylindrical sides of the detector, is one of the few detector parameter calculations that can be performed analytically (see appendix A). It is well worth checking it, therefore, against the Monte Carlo prediction

The comparison was done for two different detector diameter to length ratios, one being the actual detector ratio and the other representing a detector squashed in the z direction. The results can be seen in Figure 11 where number of tracks (representing an un-normalised probability distribution) is plotted against produced signal strength in units of ϕ_0.

Again agreement between theory and Monte Carlo is very good.

5.2 Results

The simulation was run with the actual detector dimensions for 10^6 generated tracks, to provide high enough statistics. The main Monte Carlo predictions follow; they are summarised in Table 5.

---

[13] The error bars in the graph denote Monte Carlo statistical error and are calculated as the square root of the number of intersecting tracks.

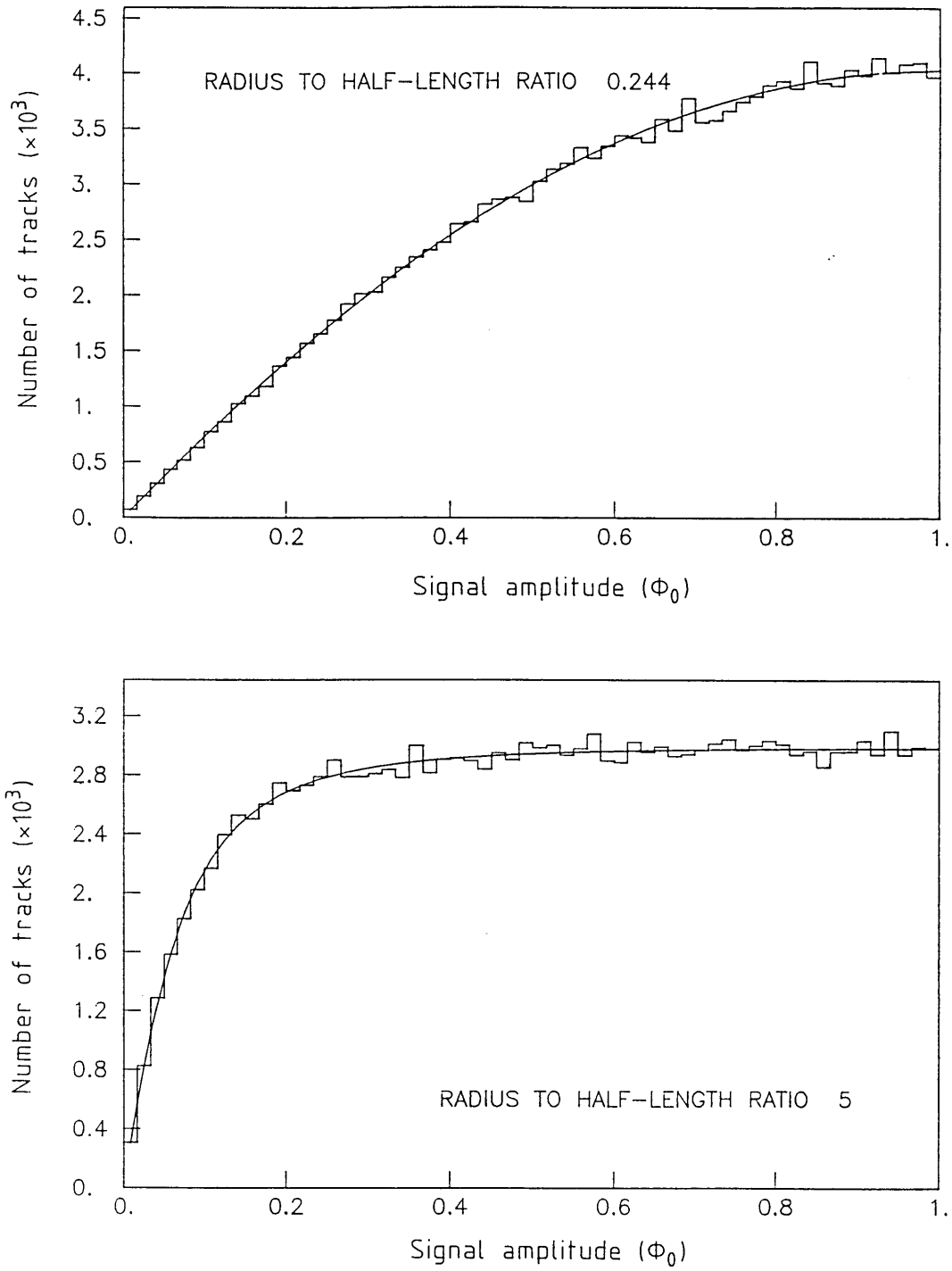


Figure 11: **Monte Carlo bias tests: sides contribution to the WF signal distribution**

5.2.1 Signal amplitude probability distribution

One of the main reasons for this simulation is the calculation of the signal amplitude probability distributions for both the AAPs and the WF loops.

The window frame amplitude probability distribution can be seen in Figure 12. Here the probability, $P(s)/ds$ of a certain signal, s, is plotted against the strength of the signal in units of ϕ_0. The top left histogram is the overall distribution whereas the remaining three graphs show the individual contribution of tracks that intersect the cylindrical sides and one of the end caps (graph(1,2)), of tracks that only intersect the sides (graph(2,1)) and of tracks that pass through both end caps (graph(2,2)). This last contribution is negligible compared with the rest (0.2%), the sides contribution being the most important (see Table 4). Tracks passing through one of the end caps contribute primarily to the low end of the signal probability distribution, whereas the sides contribution gives the sinusoidal-like behaviour. The solid line in the sides contribution graph is the analytically calculated theoretical prediction.

Table 4: Relative contributions to the WF signal

sides	78.4%
sides/end caps	21.4%
end caps	0.2%

The signal probability distribution for the AAPs is shown in Figure 13. Graph (1,1) shows the overall signal probability distribution. The signal has two distinct sections: The low end, around zero, that is produced by the 'near misses', the tracks that pass near the detector loop but not through it, and the high end comprising signals produced by tracks that pass through the detector loop. These low and high ends can be seen more clearly in graphs (1,2) and (2,1). The last graph (graph(2,2)) shows the contribution to the probability distribution of the tracks that pass through the compensating coil of the AAP. These tracks give a signal of about 18 ϕ_0, but this happens in only 10% of the cases when a track intersects the AAP main loop.

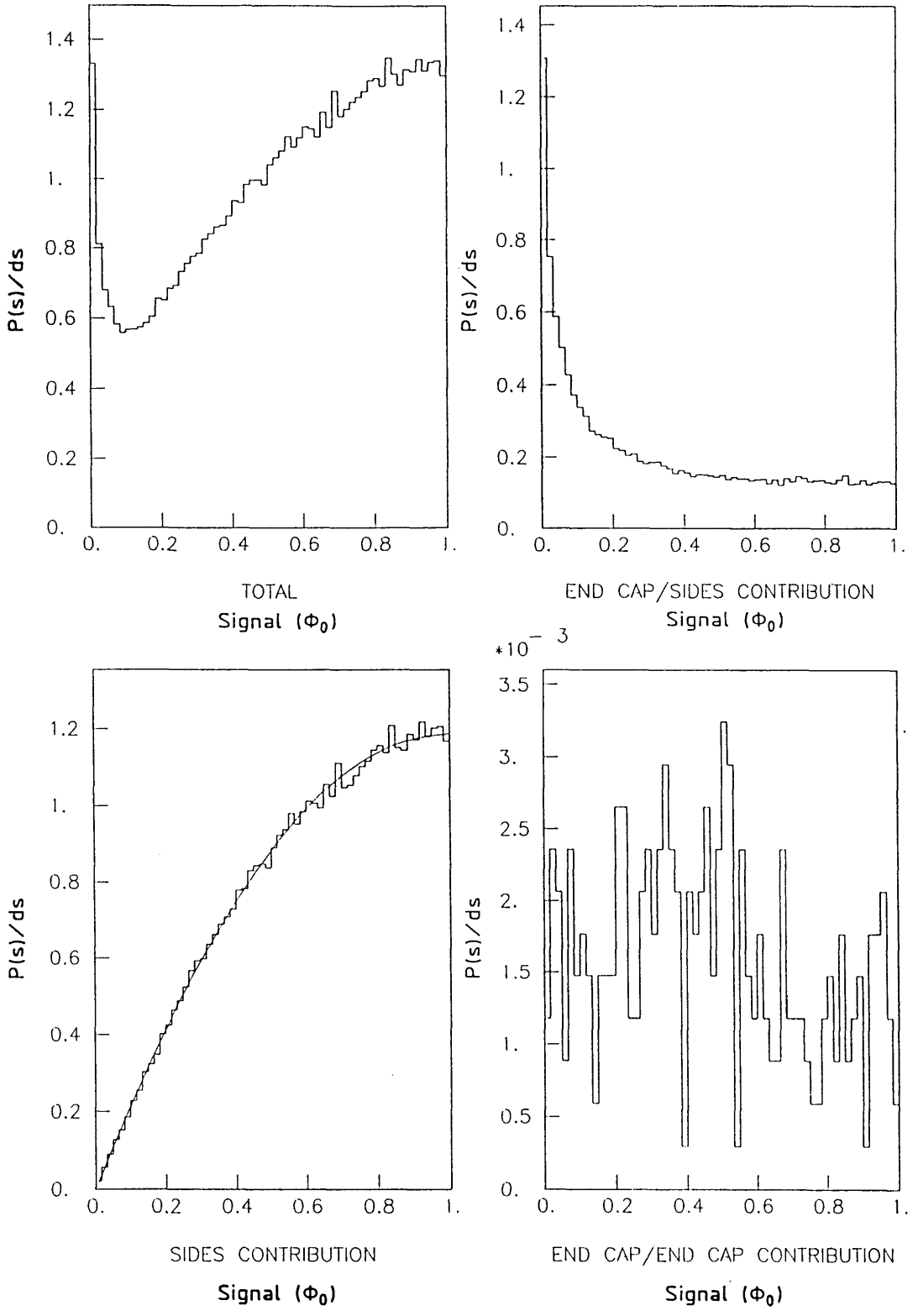


Figure 12: WF amplitude probability distribution

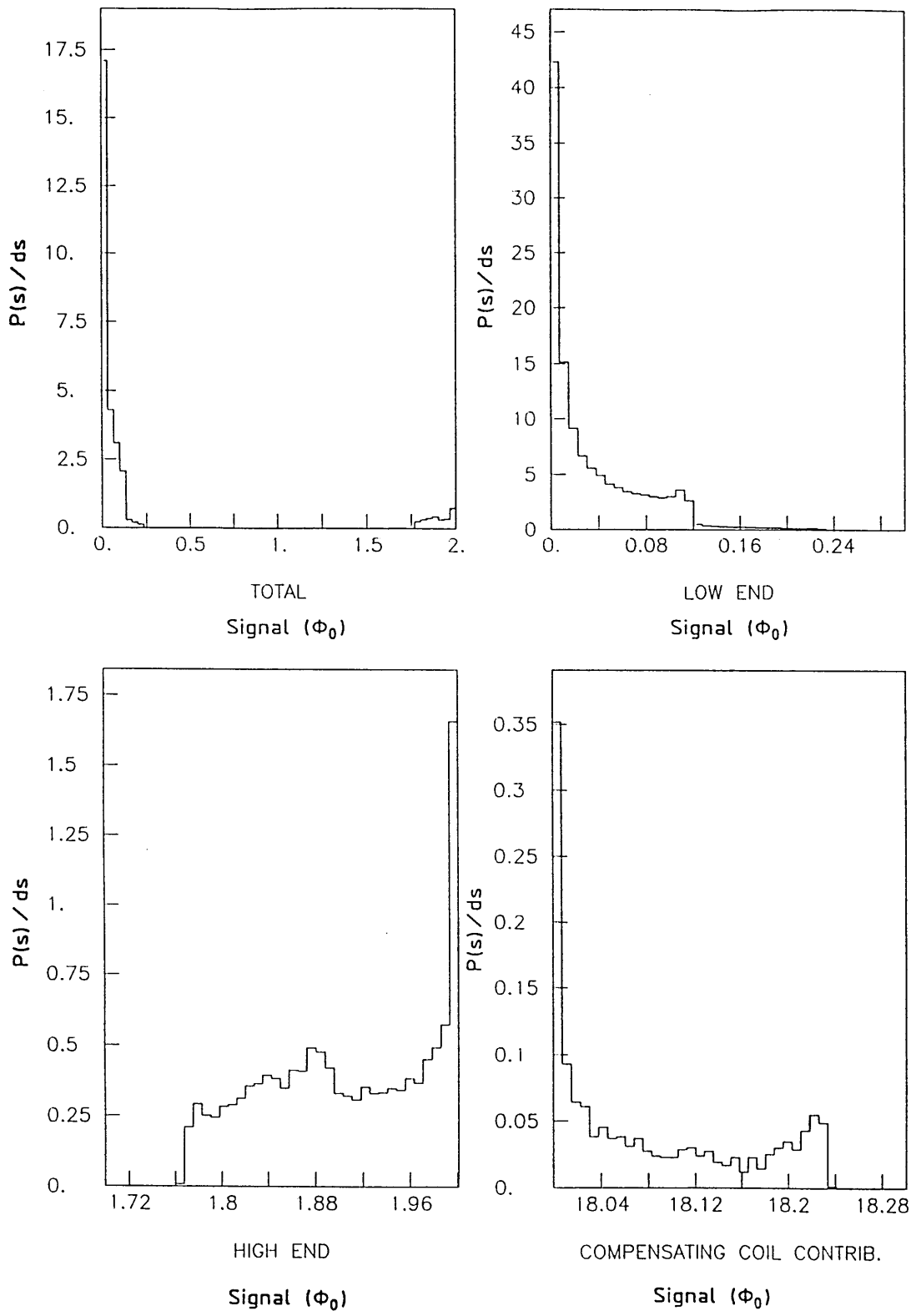


Figure 13: AAP amplitude probability distribution

5.2.2 4π averaged area

The physical 4π averaged area of the detector (WF) is calculated to be 0.1976 m^2 (appendix B). Its sensitive area is less than that, depending on our threshold and the shape of the signal amplitude distribution. For the WF the noise permits a threshold of 1mV, corresponding to 12% of our maximum signal. The 4π averaged area for such a threshold is found to be 0.1806 m^2.

For the AAPs the 4π averaged area is .114 m^2; thresholds are again 1mV. This corresponds to 18.5% and 16% of the 2 ϕ_0 signal respectively (a monopole passage through the main loop gives a signal of 2.7 and 3.1mV respectively). This is too high for the 'near miss' signals to make a contribution to their effective area, but significantly lower than the minimum signal we get for a track that passes through the loop. So the AAP 4π effective area is the same as their physical 4π averaged area.

Table 5: Monte Carlo predictions

4π averaged area:	
WF:	1,806±4 cm^2 (0.12ϕ_0 threshold)
top AAP:	114.6±1 cm^2 (0.37ϕ_0 threshold)
bottom AAP:	114.2±1 cm^2 (0.32ϕ_0 threshold)
Coincidence rates:	
$[P(top)]_{P(WF)}$	= 6.3%
$[P(WF)]_{P(top)}$	= 99.6% (± 0.3%)
$[P(top)]_{P(bot)}$	= 3.3% (± 0.1%)
$[P(top) \times P(bot)]_{P(WF)}$	= 0.2% (± 0.1%)

5.2.3 Coincidence rates

The coincidence rates for AAP and WF signals were calculated. The probability that a track passing through the AAP gives a signal above threshold in the WF is nearly unity, as expected. However, a track seen in the WF has only 12% chance to be detected by any of the AAPs as well. The probability for a coincidence between the top and bottom AAPs is about 3%, whereas the chances of a triple coincidence are minimal (about 0.2%). All the coincidence rates can be seen in Table 5. (The figures in parentheses denote statistical error.)

Chapter 6

TEST AND DATA COLLECTION RUNS

There were four preliminary runs before the main one. These were mainly exploratory, to find out more about the backgrounds, the sources of noise and generally gaining experience in the new techniques involved.

Runs 1 and 2 did not use the lead shield, and a first order gradiometer was used in the place of the WF. Runs 3 and 4 used the lead shield with the window frame. The main results were as follows:

6.1 Run 1

This was the first exploratory run. It demonstrated the need for better mechanical stability and a better gradiometer construction. A few immediately obvious improvements were made and we proceeded to run 2.

6.2 Run 2

Similar layout to run 1; no lead shield and a first order gradiometer in both x and y directions in the place of the WF. Major improvements from run 1 were better mechanical stability and more equally matched areas in the gradiometer. The variation of the amplitude in the three SQUID outputs (i.e. the total noise of the system) was about 10mV RMS (1Hz bandwidth) for the gradiometer and the top AAP and 5mV RMS for the bottom AAP. The fact that the noise in the top AAP and in the gradiometer is approximately the same seems surprising at first, since the gradiometer has some 4 times the area of the AAP. The reason is that, although the AAP is a first order gradiometer as well, for an axially symmetric field – which is probably the case inside the cryostat – the (vertical) gradiometer gives much better compensation than the AAPs. (The AAPs were designed to minimize the coil–

shield coupling and not to compensate for field gradients inside the cryostat. They can only compensate temporal variations and linear spatial gradients.)

Big drifts were present in the output of SQUID 1 (gradiometer). These drifts followed a 24 hour cycle that could be attributed to temperature variations in the room (the heating goes off at night). It is believed that the drift was due to the fact that the permeability of the mu-metal shields, and hence the field inside the cryostat is temperature dependent. The effect was present in the AAPs output as well, but it was of much smaller magnitude. This was (probably) due to the much smaller area of the AAPs together with the different behaviour of the axial (that the AAP is sensitive to) and the radial (that the gradiometer sees) components of the magnetic field inside the cryostat. Slowly varying drifts, in any case, are not a serious problem in monopole detection, since a step in the drifting output is easily recognisable, if the drifts are smooth and there are no sudden jumps in the drift gradient. This seems to be the case if a lead shield (with its characteristic metastable behaviour) is not used. Then flux jumps cannot occur although the output may drift in a smooth manner.

This run demonstrated that running without a superconducting shield is possible. Just using a first order gradiometer and not being very careful about mechanical stability, compensation, thermal currents etc. the RMS noise at 1 Hz was less than 2 ϕ_0. Eliminating the lead shield would give a very clear 2 ϕ_0 signal – as compared to the spectrum of signals (0 to 1 ϕ_0) one gets for the WF arrangement. Of course, the 4π averaged area of the WF is much bigger, but the use of 3 gradiometers in coincidence – which is possible using this technique without sacrificing most of the available volume – would give us about the same 4π averaged area together with a much needed redundancy on the obtained information. More tests were needed, though, before one was confident enough to run without a lead shield for the purpose of collecting data. The project was running behind schedule, however, and it was decided to proceed with the lead shield and the WF as planned.

6.3 Run 3

With the lead shield and WF in place this time, the detector had noisy and quiet periods. In the WF the noise varied from about 0.1 ϕ_0 to about 0.7 ϕ_0 (1 Hz bandwidth). In the AAP (only one of the two was operating in this run) the noise was more constant, about 0.3 ϕ_0.

On further investigation it was found that there was some correlation of the noise in the WF and the state of the helium inside the cryostat. More specifically, when the helium was supercooled, so that the boiling stopped, the detector became quiet. The noise was also reduced when the helium level fell below the level of the SQUIDs. This suggested that the excess noise was due to mechanical vibration caused by the boiling of the helium, and improvement on mechanical rigidity could solve the problem.

Drifts were also present in the WF output as soon as the helium level reached the top of the lead shield. These were thought to be due to mechanical deformation of the lead shield due to the temperature changing as the helium level fell.

6.4 Run 4

This run used the same layout as run 3 but with improved mechanical stability. The noise levels of the SQUID outputs, however, remained at the same levels as for run 3. So mechanical vibration was not the reason for the excess noise. Another possible source of the excess noise is thermal currents in the vicinity of the SQUIDs. Boiling is then a problem, not because of mechanical effects, but because it creates varying temperature gradients.

Another possible source of thermal current is radiation coming down the support tubes from the top insert of the SQUID holders (conduction along the tubes is negligible).

After the run, tests were carried out to investigate the thermal current problem further. The idea was to replace the brass SQUID holders with tufnol ones, in which case the thermal current contribution to the noise should disappear. Also covering all the exposed SQUID cavities with non-superconducting material should attack directly the cause of the production of thermal currents, namely the helium gas bubbles getting in contact with the SQUID cavities inducing thermal gradients.

These tests showed that noise could be reduced by:

1. Using tufnol instead of brass SQUID holders.
2. Placing cotton in exposed SQUID cavities.
3. Not having support tubes.

So the notion that thermal currents were responsible for the excess noise seemed very plausible. A more detailed discussion about thermal noise can be found in chapter 7.

The second important point that was realized during this run was that the detector was extremely sensitive to pressure changes inside the cryostat.

6.5 Run 5

This was our main data collection run. Changes from run 4 were:

1. Tufnol instead of brass SQUID holders.
2. Cotton in SQUID cavities.
3. More and larger (about $1cm^2$) holes in the support tubes to provide a free flow of helium and to eliminate any pressure differences.
4. Radiation shields along the support tubes, around the RF lines consisting of baffles of cotton wool at regular intervals.
5. Non-return valve to isolate the cryostat from the helium liquifier pressure changes together with a low pass filter in the recovery line.
6. Cryostat pressure monitor.
7. Gas operated thermal heat switch in the WF.

The reason for the changes 1 to 6 became obvious from the previous run so we will only amplify change 7: it is a good idea to make sure, sometime after cooldown, that no residual current is flowing through the detector loops. This can be done by driving a fraction of the detector coil wiring normal so that any DC currents flowing around the loop are dissipated. The way to achieve this is by heating a part of the detector coil. The most obvious way to do this is by electrical means, but it is very difficult to perform such an operation without disturbing the system (the switch has to be inside the lead

shield where the detector is extremely sensitive to magnetic fields). So the operation was performed by means of blowing hot air, localised at a small part of the detector loop. The idea was not entirely successful, though, since the thermal power transmitted to the wiring was not localized enough to drive it normal but only after a series of attempts.

The lead shield was bolted to the top of the tufnol framework and filled to a depth of about 2cm with a soap-glycerine solution (which solidifies at helium temperatures and holds the framework rigidly attached to the lead shield).

Table 6: RMS noise

bandwidth	WF	bottom AAP	top AAP
10 Hz	0.91 mV	1.27 mV	1.71 mV
1 Hz	0.37 mV	0.35 mV	0.50 mV
0.1 Hz	0.15 mV	0.16 mV	0.20 mV
bandwidth	WF	bottom AAP	top AAP
10 Hz	$0.11\phi_0$	$0.41\phi_0$	$0.63\phi_0$
1 Hz	$0.04\phi_0$	$0.11\phi_0$	$0.19\phi_0$
0.1 Hz	$0.02\phi_0$	$0.05\phi_0$	$0.07\phi_0$

Units of ϕ_0 through the detector coils

Taking all these measures we managed to stabilize the noise at a value below the 'quiet period' value of the previous run. The noise values can be seen in Table 6.

These noise levels permit to have thresholds of 1mV in the automatic step detecting algorithm of the data acquisition system, which, as already discussed in chapter 4, effectively operates at a bandwidth of 0.01Hz. This threshold corresponds to 0.12 ϕ_0 in the WF which thus covers 91% of the available volume, and about 0.3 ϕ_0 in the case of the AAPs. The detector ran smoothly in this config-

uration. There are on average two to three events per week, triggered by the step detecting algorithm of the data acquisition system that in most occasions can be attributed to trivial causes. Generally the detector was extremely quiet.

Chapter 7

SYSTEM PERFORMANCE AND RESULTS

The detector was cooled down on the 29th of September 1984, and was continuously in operation for one year. It was filled up with helium at regular intervals of about once a week and was not allowed to warm up above 4.2^0K. When the run ended, on the 30th of September 1985, the total effective observation time was 8242 hours. This excludes the time when the detector was not stable: during and a few hours after helium refills, when the helium level dropped below the top of the lead shield, (resulting in big drifts in the detector outputs), or when the data acquisition program was not in operation. The effective observation time is about 94% of the total running time.

A typical monitoring record can be seen in Figure 14. This shows maximum and minimum values in successive 100s periods for the three detector coils, the magnetometer, the accelerometer, the RF monitor and the cryostat helium pressure in a typical data taking day (8 January 1985). Increased activity in the building is evident in the field variation, caused by the lifts, and the level of vibration. The helium recovery compressor caused the pressure variation at about 10:00 h.

In this chapter we shall discuss the performance of the detector and the results obtained after a year's operation [26].

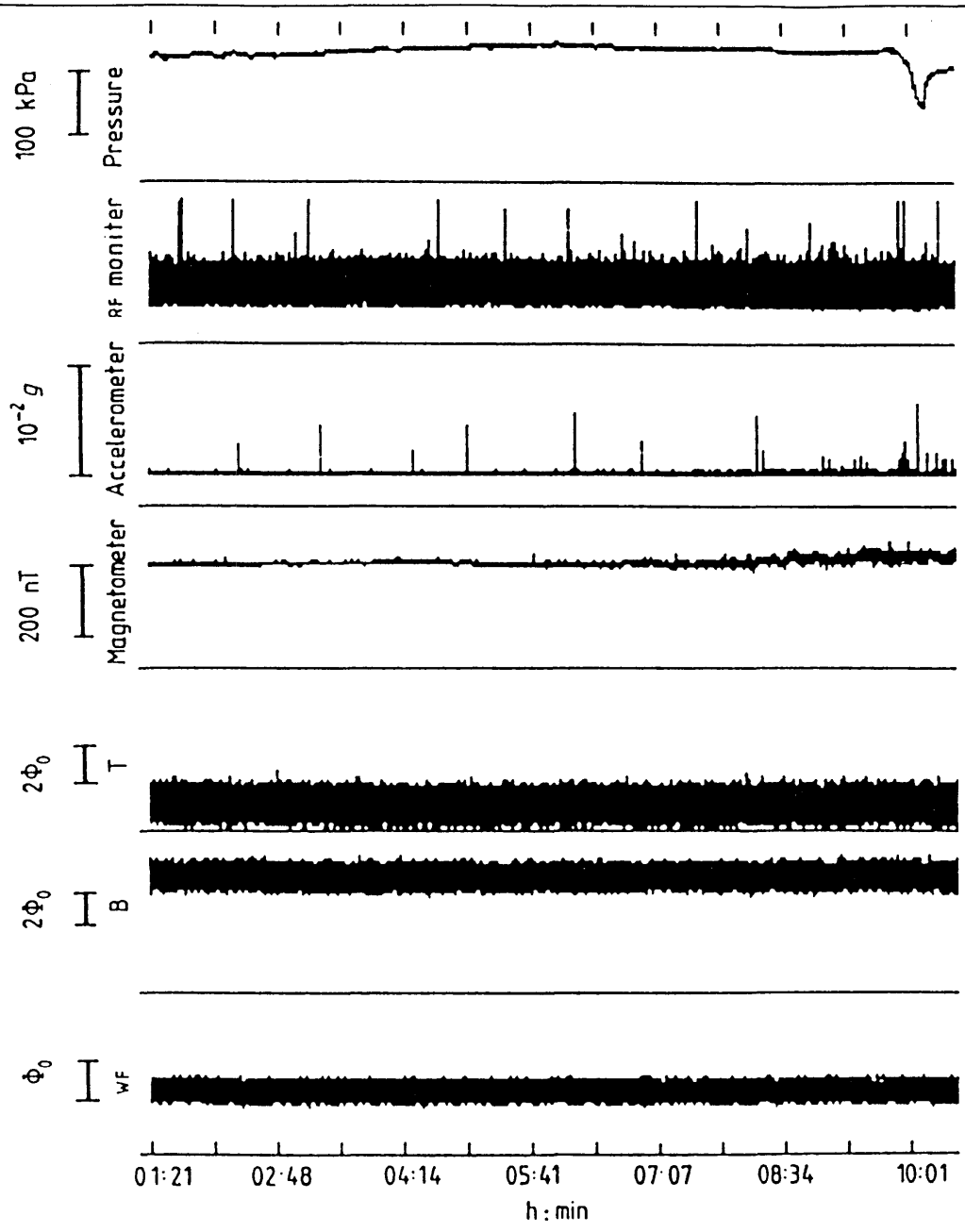


Figure 14: Typical low bandwidth monitoring record

7.1 Detector performance

7.1.1 Magnetic sensitivity

An important test for any inductive monopole detector is how well it attenuates the external magnetic field disturbances. This test was performed by cancelling locally the Earth's magnetic field with a 1m diameter coil placed beneath the shields, thus avoiding the application of any large local fields that might have degraded the performance of the mu-metal shields. Still, the magnitude of the induced disturbances was quite big, some 3 orders of magnitude bigger than typical magnetic field disturbances. With the coil on axis, the test procedure nulled the (local) vertical field component, to which the AAP loops are sensitive.[14] With the external coil off axis, the horizontal field component, to which the WF loop is sensitive, is changed. The tests were made by switching current through the coil at 0.1Hz and then performing a spectral analysis on the detector outputs to look for any excess signal at the frequency of 0.1Hz. In this way very small signals can be detected. No significant response was found in any of the three channels and the upper bound of signals induced was estimated at 0.03 ϕ_0 for the AAP loops and the WF loop. Assuming that the overall compensation is to within 1%, (a very optimistic figure), this result indicates that the dynamic magnetic field attenuation provided by the mu-metal and superconducting shields combined is conservatively estimated to be at least 160dB for longitudinal fields; the absence of response in the WF loop similarly indicates an attenuation of at least 200dB for transverse fields.

7.1.2 Mechanical sensitivity

The mechanical sensitivity of the detector was tested by swinging it with an amplitude of about 10^{-2}rad at its natural frequency of about 0.7Hz. This produced a synchronous response of 0.1 ϕ_0 in the AAP loops, and two or three times larger signal in the WF loop; no permanent offsets were induced. However, sharp knocks to the outside of the cryostat, recorded as 10^{-1}g or more in the

---

[14] The AAPs are sensitive to inhomogeneous magnetic disturbances. They only compensate fields to zeroth order. The disturbance induced in this test when attenuated by the mu-metal shields is very inhomogeneous therefore no compensation is expected.

accelerometer, did produce offsets in the WF of 0.2 to 1 ϕ_0. This confirmed what we already knew, namely that the WF is mechanically more sensitive than the AAPs. This is because the latter are magnetically coupled only weakly to the superconducting shield, whereas three legs of the WF loop are very close to it. Consequently, any small movement of the WF loop relative to the shield would alter its linkage with the trapped flux that is pinned to the shield.

7.1.3 Detector current rise time

It is important to know the time characteristics of putative events and to be sure of the overall system frequency response. This was tested using the toroidal test coils associated with each of the detector loops by exciting them with an appropriate range of frequencies. The response, as expected, was dictated by the bandwidth of the SQUID control unit output filters, which were set to 10Hz. The rise-time (0 to 90%) for a step input was found to be 65ms.

7.1.4 Excess low frequency noise

During the third and fourth run it was found that the SQUID noise levels, particularly at frequencies below 1Hz, were sometimes much higher than expected and depended upon the state and height of the liquid helium within the cryostat. The excess noise disappeared when the helium level dropped below the SQUID sensors and could be suppressed temporarily by closing the helium return line, or by pumping the helium bath to about 4.0K and then returning the pressure above the bath to atmospheric; i.e. the excess noise disappeared when the liquid helium was supercooled inside the cryostat. Moreover, the reversion to the noisy state would be smooth and reproducible. Sometimes just one loop (usually the top AAP) would be noisy, sometimes two (the top AAP and WF loops), and sometimes all three. However, a cross-correlation check between the signals from two noisy channels showed that, while the individual noise spectra were similar, there was no significant correlation between the two. Apparently, the same mechanism was responsible for three uncorrelated local noise sources. Attempts to reproduce this noisy behaviour with a SQUID sensor in a small test rig did not succeed.

The low frequency of this excess noise suggested some kind of thermal mechanism, and SQUID sensors are known to be sensitive to thermal gradients. Such gradients can produce thermoelectric currents, either in the metal case of the SQUID sensor itself or in any other piece of metal in the vicinity. Two further factors are needed: that the metal is inhomogeneous, so that there is some spatial variation of the thermopower, and that the pattern of current flow couples magnetic flux into the SQUID input circuit. In materials such as brass, the material originally used for the SQUID holders, variations of the thermal power at 4^0K of order 10^{-7} V/K are quite usual. When combined with temperature gradients of order 10^{-3} K/cm, these provide large enough currents. Using non-metal SQUID holders, thus avoiding the circulation of thermal currents in the SQUID vicinity, should eliminate this thermally produced excess noise. Also, in the detector, the major heat flow path into the liquid helium inside the lead shield is through the RF return lines; this heat flow would have set up more or less independent convection cells around each of the sensors, thereby producing the noise. We therefore improved the heat insulation of the RF lines, replaced the brass SQUID holders by 'Tufnol' ones, and surrounded the sensors themselves with layers of cotton wool and coilfoil[15]. This combination of modifications proved effective, although we did not test whether they were all necessary. The SQUID noise levels where then within the manufacturer's specification of 1.5×10^{-4} ϕ_0 (RMS) in a DC to 1Hz bandwidth (the Johnson noise of the 2Ω shunt resistors across each SQUID input coil makes a significant additional contribution).

7.1.5 Sensitivity to pressure changes

As already mentioned in chapter 5, the detector loops and the WF in particular were sensitive to pressure changes in the helium return line. A pressure pulse of 50mBar from the helium recovery compressor would induce an offset of 0.5 to 1.5 ϕ_0. Tests indicated that the mechanism responsible was a combination of thermal and mechanical effects. A 50mBar pressure change will stretch the inner bucket of the cryostat, constructed from fiberglass and aluminium, by about 20μm, and some of that displacement would be transmitted to the detector assembly. Thus, mechanical deformation of the

[15] A sheet made of insulated copper wires glued together to provide good thermal conduction but avoiding any closed electrical circuit.

lead shield could change the linked flux to the detector loops causing an offset. After a sudden change in pressure there was also (in addition to the mechanically induced offsets) a slower effect whose timescale suggested some sort of thermal current cause. The WF loop, being more strongly coupled to the lead shield, responds more than the AAP loops.

Although all these are plausible explanations for the pressure induced effects it must be admitted that the subject has not been investigated fully. Instead, the way that the problem of this sensitivity to pressure changes was dealt with, was by inhibiting rapid pressure changes using the low pass filter described in chapter 5. The WF remains sensitive to changes in atmospheric pressure, but these are far too slow to be confused with monopole passages.

7.1.6 Thermal expansion

A more serious problem associated with the detector design is that it moves vertically within the helium bucket as the helium level changes. This is because of differential expansion of the cryostat as the cryostat temperature profile alters and due to the fact that the whole of the detector framework is mechanically overconstrained. The total movement was estimated to be about 2mm. This movement would not be a problem if the detector hung freely within the cryostat, but because they touch,[16] the motion seems to be spasmodic and is a source of intermittent internal mechanical shock. Most of the movement, as indicated by the accelerometer, occurs in the first day or so after the weekly helium refill, but shocks have been seen on other occasions, even late at night, when external sources are rather unlikely.

[16] The lead shield was designed to hang freely inside the cryostat having an external diameter 5mm less than the internal diameter of the cryostat. Manufacturing tolerances, however, together with the fact that the lead shield had heating wires, etc. on its outside and was in the end covered in paper, finally resulted in the lead shield and cryostat being in contact.

7.2 Data analysis and results

About 180 data dumps were triggered by the data acquisition program over the main data taking period. These include some dumps that were triggered when the detector was not stable, (and which were excluded from the effective running period). In most cases the cause was obvious immediately, for example: operator error, gross mechanical shock, or the helium level being allowed to fall below the top of the detector assembly. For a small number of events a more detailed look was needed to establish their cause. A detailed account of the cuts that an event had to pass before it could be regarded as a genuine monopole candidate, and events passing them, will follow in the next sections.

As already discussed, an event dump consists of a 400 second history of the three detector coils and relevant interference monitors, sampled at 10Hz. A typical event dump can be seen in Figure 15. This particular event (event 112 detected at 20:18:44, 19 Apr 1985) is associated with mechanical shock. It was triggered by a step in the WF channel of magnitude of about $0.4\phi_0$ At the time of the event the lab was empty so that the shock was almost certainly caused by differential thermal expansion within the cryostat. Note the time delay between the occurance of the event (18:22) and its detection (18:44) due to the effective bandwidth reduction of the step detecting algorithm.

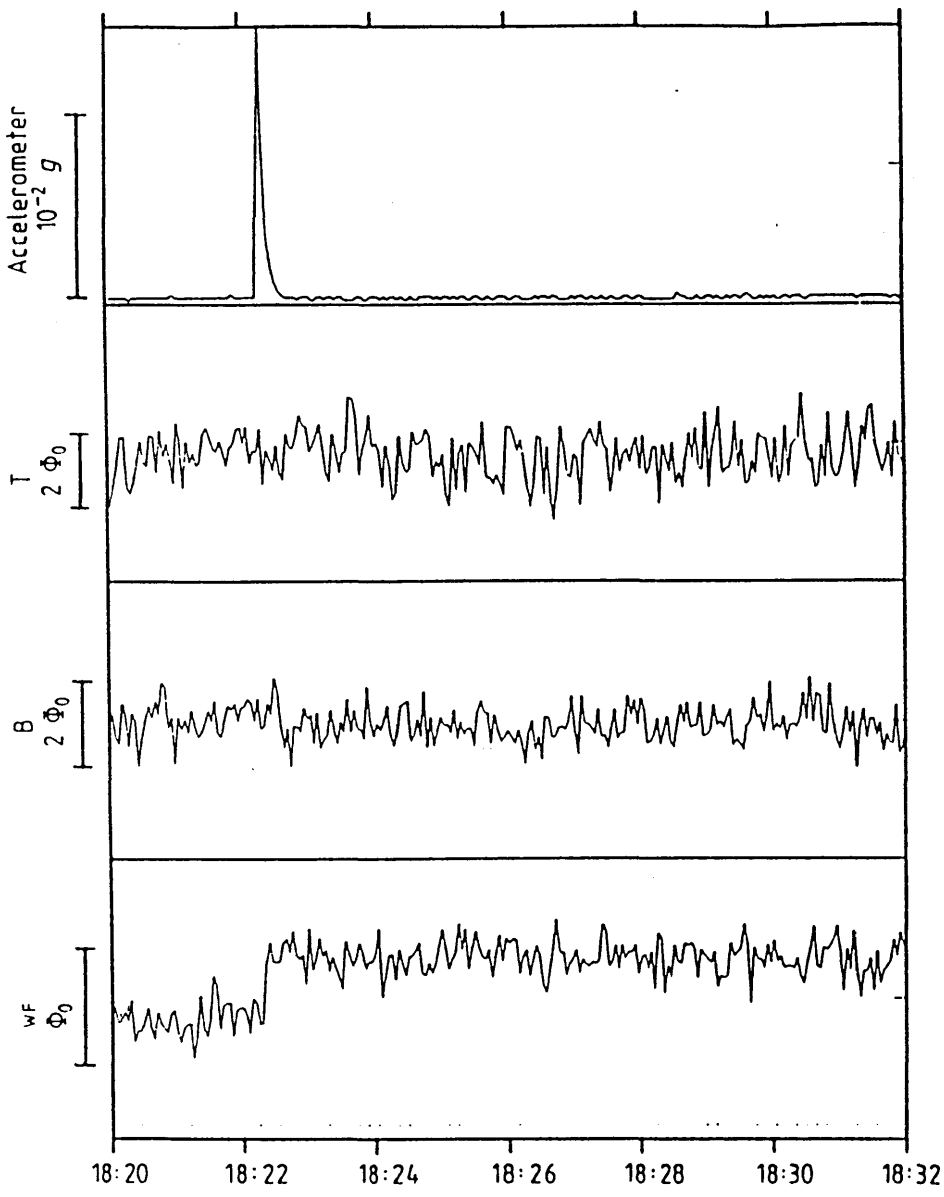


Figure 15: A typical event. A 12 second section of the detailed (10Hz) data record of event 112, detected at 20:18:44, 19 Apr 1985.

7.2.1 Cuts

For an event to be considered as a possible monopole candidate it has to occur during a valid observation time period, i.e. when the detector was stable and operating smoothly, and pass the following cuts:

Magnitude: Steps bigger than the set threshold in the data acquisition program are clearly distinguished from the detector noise. Not all of those signals, though, are compatible with the passage of a monopole. For the AAP loops, the expected monopole signal lies in a narrow range. Therefore, the magnitude of an event in the AAPs has to lie in the range $1.7-2.0\ \phi_0$ or between 18.0 and 18.3 ϕ_0 (within the 5% calibration uncertainty) to be considered a candidate event (see Figure 13 on page 53). For the WF this cut is weaker; here an event can have any magnitude from threshold to 1 ϕ_0.

Coincidence: The detector has only partial coincidence. A signal on the WF has on average a 12% chance of being associated with a signal in any of the AAPs (see Table 5). This depends on the magnitude of the WF signal and can be less than 1% for small WF signals. On the other hand, the chances that an AAP signal is accompanied by a detectable offset to the WF are large (bigger than 99%). So here we can impose an extra cut by demanding that any signal in the AAP must also be accompanied by a signal in the WF.

Monitoring: Offsets can have spurious causes associated with environmental disturbances such as shocks, pressure changes, etc. To guard against those, all important environmental quantities are continuously monitored. A further cut we can impose, therefore, is that a possible monopole candidate should not be associated with a coincident abnormal behaviour in any of the monitored quantities that has been known to be a potential cause for an offset, such as accelerometer spikes, etc. Provided that such abnormal behaviour is not very frequent, the dead time introduced by this cut is negligible.

Rise time: Since the transit time of even a very slow monopole would be fast compared to the rise time of the electronics (65ms for 0 to 90%), we can impose an extra cut by requiring a possible monopole candidate signal to appear as bandwidth limited: This cut becomes less efficient as the signal amplitude gets close to the noise level at the sampling frequency. For an offset much larger than the noise at the sampling frequency, a

slow event is easily recognizable; for a step height magnitude closer to that of the noise, however, only events very much slower than the sampling period could be discarded with this cut. This is a problem only for WF signals; the noise at 10Hz is 0.11 ϕ_0 whereas the threshold is at 0.12 ϕ_0.[17] For the AAPs acceptable signals are at least 4 standard deviations away.

Continuity: The quantities that determine the shape of the detector output, that is the level of the noise, the slope of any drifts present, etc., should not be affected by a monopole's passage. Therefore, an extra cut we can impose is that genuine monopole passages, apart from changing the DC level of the SQUIDs, should not change the shape of the detector output; that is the slope of any drifts present and the RMS noise should be the same before and after the event.

If a step is accompanied by such a change, then it is probably associated with some of those (local) quantities that are responsible for drifts and the noise, such as temperature changes, mechanical motion etc., and therefore should be excluded from consideration as a candidate event.

Malfunction: Events that have been known to have been triggered due to an operator error or a hardware or software malfunction, such as excitation of the calibration coils, fault in one of the SQUID boxes, etc., were immediately excluded as spurious.

7.2.2 Results

Table 7 indicates the causes, as determined after analysis of the high-bandwidth data, of putative events recorded on disk after the step sensing algorithm of the on-line analysis program found an offset higher than threshold. The time when most putative events tend to occur is when the helium level is low. Over such periods, large numbers of events are generated in rapid succession, exceeding the storage capacity of the (5-inch floppy) disks available. (Two disk drives are in operation, so a maximum number of 4 events can be dumped into the disks between operator checks.) Consequently, not all

---

[17] Step detection is done at a much lower frequency, 0.01Hz where noise is much lower than the threshold

such events have been recorded. But in any case, as already mentioned, such periods are excluded from the effective observation time and all the events that occurred over that time are discarded.

Table 7: Causes of putative events

Cause	WF	Top AAP	Bot. AAP
Low helium level	44	7	0
Instability	11	3	1
Pressure fluctuations	17	2	0
Operator	5	4	3
Software errors	8	5	2
RF interference	0	0	0
Mains-borne interference	0	2	0
SQUID unit faults	4	26	0
SQUID unit drifts	0	0	8
Excitation of calibration coils	0	8	0
Coincident mechanical shock	8	0	0
Post-mechanical shock	3	0	0
Unknown	1	0	0
Total	101	57	14

For the few hours following each refill, the detector outputs (mainly the WF and top AAP loops) drift rapidly and are generally unstable causing the data acquisition program to trigger a lot of events. Such periods are again excluded from the effective observation time.

Events due to pressure fluctuations were a significant percent of the early triggers; after the introduction of the low lass filter, however, no more pressure associated events have been seen, although small drifts due to atmospheric pressure changes are still present.

RF interference is a potential cause of offsets in the SQUID outputs especially at the operating frequency of the SQUIDs. Therefore, care was taken to shield the detector against RF. This RF shielding seems to be successful, since no events have been seen that could be associated with RF interference; earlier tests had shown that the detector was pretty insensitive to RF, even at 19MHz, the RF frequency of the SQUIDs. (To see an effect one had to wind the aerial of a 19MHz transmitter around the SQUID RF return lines.)

An occasional intermittent fault in one of the SQUID control units was diagnosed by transferring it from the WF output to the top AAP channel, when it degraded rapidly. It usually produced big offsets (of the order of tens of ϕ_0 or bigger). The unit was fixed and the problem disappeared.

The SQUID control unit of the Bottom AAP is somewhat temperature sensitive, so the difference between day and night temperatures can cause mis-trimming and consequently the noise increases by a large factor, giving rise to occasional triggers of the data acquisition program.

Finally, there were offsets associated with coincident shocks, some of them having a slow rise time.

As it can be seen from Table 7 most of the events dumped to disk from the data acquisition program could be immediately excluded as spurious since they did not pass one or more of the cuts. (usually a spurious event is excluded due to more than one reason – the cuts are not orthogonal to each other)

The above cuts, therefore, leave four unexplained events which were looked at more carefully to see if there was anything that made them suspicious. They all happened in the WF channel, having a magnitude that spanned from 0.14 ϕ_0 to 0.84 ϕ_0. In all cases all the monitors at the time of the event were quiet, and the steps appeared as bandwidth limited. These events can be seen in Table 8.

Inspection of the detailed data for event numbers 34, 35 and 74 revealed the fact that in each of the events there was a preceding mechanical shock [27] (see Table 9). There seems to be a statistical significance between those shocks and the events that followed them; the chance of a random association of the shock with the event was always less than 10%, and for event 74 less than 1%. This

Table 8: Events passing the cuts

event ♯	event date	event time	magnitude
34	6 Nov 84	05:11	$-0.43\phi_0$
35	16 Nov 84	21:43	$-0.41\phi_0$
74	16 Feb 85	15:30	$+0.14\phi_0$
160	11 Aug 85	07:06	$-0.84\phi_0$

strongly suggests that those events were associated with some spurious mechanism that also produces large mechanical shocks sometime before the event (such as: mechanical deformation that results in a quiet movement of the framework of the detector with respect to the shield, mechanical shock that leads to flux redistribution in the shield some time after the shock, etc.). We clearly do not know what the mechanism that caused these events can be, but the strong circumstantial evidence of the earlier shocks suggests that event numbers 34, 35 and 74 are of a spurious nature. In other words, the detector has not been as quiet as we anticipated, and the cuts mentioned earlier are not enough to exclude all spurious events. So we presume that there is a mechanism that sometimes produces clear offsets after a shock. Since we do not know how to guard against it, we have to *a posteriori* impose another cut; namely that an event is disregarded as spurious if it lies in a 150s interval after a major shock recorded in the accelerometer monitor. The length of this interval, 150 seconds, is defined from the maximum delay time between a shock and event that has been seen (see Table 9) and, although arbitrary in the sense that was only defined after the experiment, it is justified from the statistical significance between shocks and offsets in events 34, 35 and 74. This extra cut introduces a considerable dead time; on average we get a few big shocks per hour, therefore the dead time introduced with this cut is of the order 20% of our real observation time. So our **effective** observation time falls to about 6600 hours.

Table 9: Delayed shock events

event number	event size(ϕ_0)	shock (10^{-3} g)	delay between shock and event (s)	Ambient shock frequency
34	−0.43	10	143	1/h
35	−0.41	4	23	10/h
74	+0.14	4	10	1/h

7.2.3 Candidate event

If we therefore exclude the three delayed shock events imposing this extra cut, we are left with one event. This is event number 160, which happened at 7:06 BST on Sunday, 11th of August, 1985 [27]. It was registered in the WF loop and had an amplitude of $0.84 \pm 0.05\phi_0$ (the uncertainty is primarily that in the calibration of the detector loop). No visible offset was recorded in any of the other detector loops. A section of the detailed (10Hz) record of the WF channel when event 160 occurred can be seen in Figure 16.

The noise levels in all three detector channels had their usual magnitudes, and application of a range of statistical tests to the data revealed no anomalies. Nor was any obvious abnormal behaviour to be seen in the channels monitoring vibration, RF interference and cryostat pressure. Inspection of the less detailed monitoring record and the conventional chart records (that serve as a backup) for the days before and after event 160 revealed nothing unusual. The weekly helium refill had been on 6 August with a moderately noisy period on 7−8 August (that always occurs after a refill). From 10 to 12 August conditions were very quiet, but not unusually so for a weekend. As soon as possible after the event occurred, the detector system, including the monitoring devices, was checked thoroughly for normal operation. Moreover, after the end of the main data-taking run on 30 September, the detector, while still at liquid helium temperature, was subjected to tests. Finally, it was warmed slowly to room temperature and an internal inspection was performed. For completeness, the data from this event have been subjected to various statistical tests to make sure that no discrepancies, inconsistencies, or

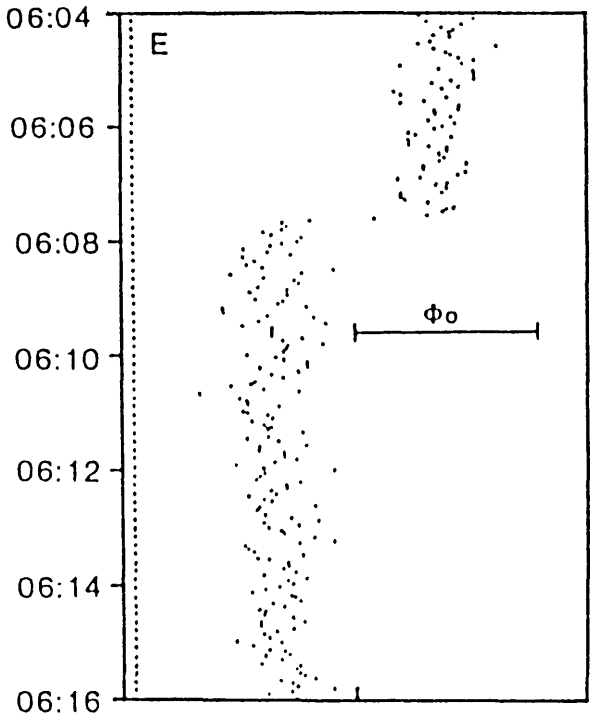


Figure 16: Detailed record of event 160. The points are the sampled levels of the WF coil output. (20 samples per second.)

anything we did not understand would arise that could make the event suspicious. None of the above steps revealed anything abnormal.

Therefore, event 160 has survived all the cuts that classify the event as having a spurious nature, so it should be regarded as a monopole candidate event. Possible causes (all of them highly unlikely) together with a very detailed account of all aspects of this monopole candidate event can be seen in reference [27][18].

We shall here only mention one of the possible explanations, which is unlikely, but not as unlikely as the others. It is the possibility of 'quiet' flux redistribution inside the superconducting shield. When cooled down, the superconducting shield trapped the residual magnetic flux, estimated to correspond to a field of about 10^{-9} T (the field inside the room temperature mu-metal shields). In principle, this flux is thermodynamically metastable and if it were to move, it would simulate the passage of

[18] For some discussion on one of the possible causes of event 160, also see reference [28].

a monopole. Due to WF loop's strong coupling to the shield, such flux motion would be likely to generate a larger signal in the WF loop than in the other two channels. Movement of pinned flux can be activated thermally and so after the initial cooldown we followed the standard procedure of 'annealing' the shield, by maintaining it for some time above its operating temperature. There have also been several occasions when the superconducting shield warmed up slightly when liquid helium refilling was delayed. Thus, any flux that was prone to move would most probably have done so during those periods. Also, even during periods when the helium level is low and the detector is much less stable, we have seen no clear-cut evidence of flux redistribution; there have clearly been events in those periods but the data indicate that causes other than flux motion, such as mechanical movement arising from thermal expansion, were the main culprits. In any case, their magnitude was always less than 0.4 ϕ_0. Thus, although we cannot exclude flux motion in the shield as the cause of event 160, the empirical evidence is that it is unlikely.

Finally, we shall present a detailed account of the statistical analysis of the three detector coil outputs to assess the statistical significance of the step and look for small coincident signals in the other two channels: If some signal of improper size would be seen in either of those two channels, it would indicate the presence of a spurious cause for event 160.

From the detector Monte Carlo we can derive the probability distribution of signal sizes in the AAP loops calculated for those monopole trajectories that give a WF signal close to the magnitude of the signal seen in event 160. Figure 17 plots this probability distribution for monopoles that give a signal in the range $0.78-0.88\ \phi_0$ in the WF loop. About 70% of the tracks give signals that are less than $0.1\ \phi_0$ and it is highly improbable that a signal greater than $0.2\ \phi_0$ would be due to a genuine monopole passage.

To obtain limits for possible signal sizes (together with their statistical significance) for all three detector loops for event 160 we proceed as follows: We reduce the bandwidth to 0.1Hz by averaging over 5 seconds of data (100 points) in an attempt to reduce the noise so that the height of any steps would be more accurately determined. By reducing the bandwidth by 100 times, we could make a fac-

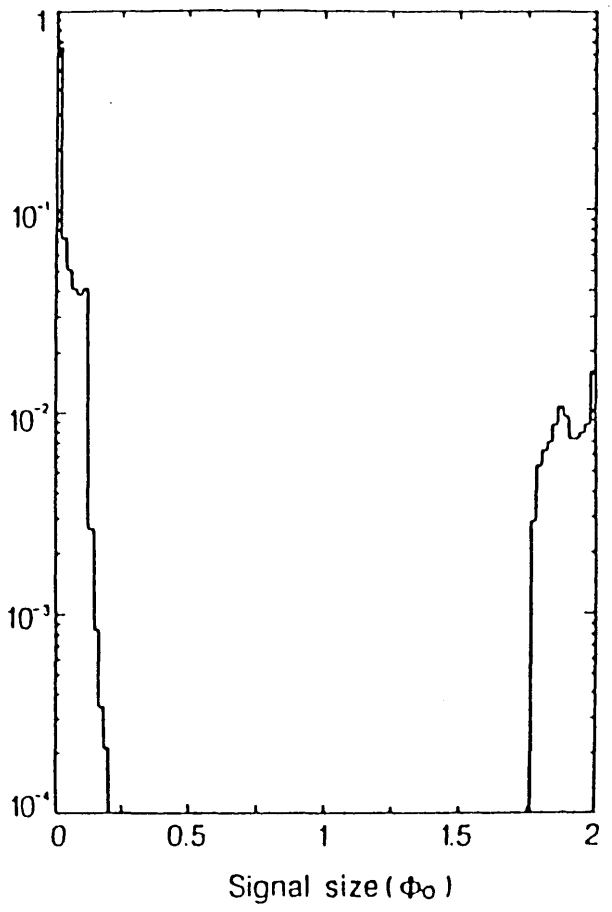


Figure 17: Probability distribution of signal sizes in the AAP loops for event 160

tor of 10 reduction in the RMS noise if we assume that the noise could be approximated by white noise.

After the bandwidth was reduced, different fits were made to assess the significance of steps imposed at the time the event occurred in the WF channel. This was done as follows: Polynomials of various degrees were fitted to the data and the difference in the quality of the fit, when imposing a step and when not, were noted together with the step height. An implicit assumption was made for the fit, namely, that if the step was to be removed, the function that fitted the data should be continuous (i.e. it was assumed that no abrupt change in slope occurred during the event). The justification, as discussed in chapter 6, is that the drift of the detector outputs is caused by environmental parameters that are unrelated to monopole passages; therefore a genuine monopole passage should not induce a change

in slope in any of the SQUID outputs.

Polynomials of 0th, 1st and 2nd order were fitted to the data (10 points before and 10 points after the event, each point being the average of 100 10Hz-bandwidth points and each weighted according to its standard deviation) with and without a step (Therefore, for the 2nd order polynomial fit, the number of degrees of freedom were 16 for a fit with a step, and 17 without one). A tabulation of the results of this analysis can be seen in Table 10. As can be seen, the analysis yields a very significant step in the WF loop, (as expected, of height 0.84 ϕ_0 and nothing significant on the other channels (the χ^2 per degree of freedom when assuming there is a step is not significantly different from the no-step case). Note that the χ^2 is large, implying poor fitting. This is to be expected, however, because of the low frequency noise of the detector outputs that require polynomials of high order for a good fit. This low frequency noise, as previously discussed, is more noticeable in the WF, the bottom AAP being the most quiet channel. Therefore we expect the χ^2 to be highest for the WF data and smallest in the bottom AAP ones. In any case, what we are interested in, is the relative change in χ^2 as the step is changed, and not in the absolute value of the χ^2 itself.

The values for the possible steps for the AAPs are less than our threshold values and bring us in the 'near misses' region of our AAP amplitude probability distribution (see Figure 13. Therefore the probability for a coincidence between the WF and any of the AAPs is bigger than the value quoted in Table 5 for the 0.3 ϕ_0 threshold case and can be seen in Figure 17 as already mentioned. The probability for a coincidence of a WF signal of height between 0.79 and 0.89 ϕ_0 and a signal in one of the AAPs bigger than 0.16 ϕ_0 was calculated to be about 10%, the probability for a triple coincidence being quite small (0.3%). Therefore, the chances of getting a 0.84 ϕ_0 signal in the WF accompanied with a signal smaller than 0.16 ϕ_0 in any of the AAPs (as is the case for event 160) is about 80%. Therefore as far as this analysis is concerned, event number 160 is entirely consistent with a monopole passage. In any case, this method would have been able to rule out a genuine monopole passage only if we could assign a step bigger than, say, 0.20 ϕ_0 in any of the AAP loops.

Table 10: Event #160: step analysis

polynomial order	step height (ϕ_0)	χ^2 per degrees of freedom	
		with step	without step
Window Frame			
0th	-0.838 ± 0.003	3.33	1502
1st	-0.847 ± 0.007	3.47	313
2nd	-0.844 ± 0.000	2.92	328
Top AAP			
0th	-0.056 ± 0.19	1.96	2.09
1st	$+0.002\pm0.37$	1.98	1.87
2nd	$+0.004\pm0.04$	1.57	1.48
Bottom AAP			
0th	-0.011 ± 0.012	0.82	0.80
1st	-0.053 ± 0.024	0.75	0.84
2nd	$+0.051\pm0.024$	1.43	1.00

In conclusion, we can say that we have no evidence that event 160 had a spurious nature. With one event seen in 6600 effective hours of observation, we derive a monopole flux of $1.9\times10^{-12}cm^{-2}s^{-1}sr^{-1}$. This, at the 90% confidence level corresponds to an *upper* bound on the monopole flux of $7.2\times10^{-12}cm^{-2}s^{-1}sr^{-1}$ and a *lower* bound (again at the 90% confidence level) of $9.8\times10^{-13}cm^{-2}s^{-1}sr^{-1}$. This is some 3 orders of magnitude above the Parker bound, suggesting that it is highly unlikely that event 160 could be due to a genuine monopole passage. Had the detector been designed to provide more redundancy in the information supplied in the event of a monopole passage, the puzzle of event 160 might have been resolved.

Part 2

A Measurement of Charmed Particle Photoproduction

Chapter 8

THEORETICAL BACKGROUND

8.1 Photoproduction

The interaction of high energy photons with matter occurs by two mechanisms, reflecting the dual nature of the photon:

1. Point-like: the photon interacts directly with one of the charged partons from the target nucleon. An example is QED-Compton scattering on quarks ($\gamma q \rightarrow \gamma q$).
2. Hadron-like: the photon dissociates into quarks and gluons which then interact with the target. This type of interaction may be further classified:
   a. the photon couples to one of the vector mesons (ρ, ω, ϕ, ψ, etc.) which have the same quantum numbers as the photon. The meson then interacts with the target.
   b. the photon dissociates into a quark antiquark pair, one of which then interacts with the target. This process has similar characteristics to high p_T hadroproduction and it is sometimes termed the anomalous component of the photon's interaction.

8.1.1 Charm photoproduction

In charm photoproduction the situation is simpler, since the momentum transfer leading to the production of a charmed quark of transverse momentum p_T is given by:

$$Q^2 \approx m_c^2 + p_T^2$$

where m_c is the charmed quark mass. Therefore, $Q^2 > m_c^2 \approx 2GeV^2$, which is substantially greater than the hadronisation scale of QCD. Hence, due to the large charmed quark mass, perturbative QCD is

valid even for low p_T, unlike the situation for the general hadronic photoproduction. Therefore QCD can provide predictions without the need to rely on phenomenological models such as the Vector Meson Dominance, for instance.

The situation in charm photoproduction is further simplified by the fact that, to first order, only one diagram contributes, that of the photon gluon fusion process (Figure 18).

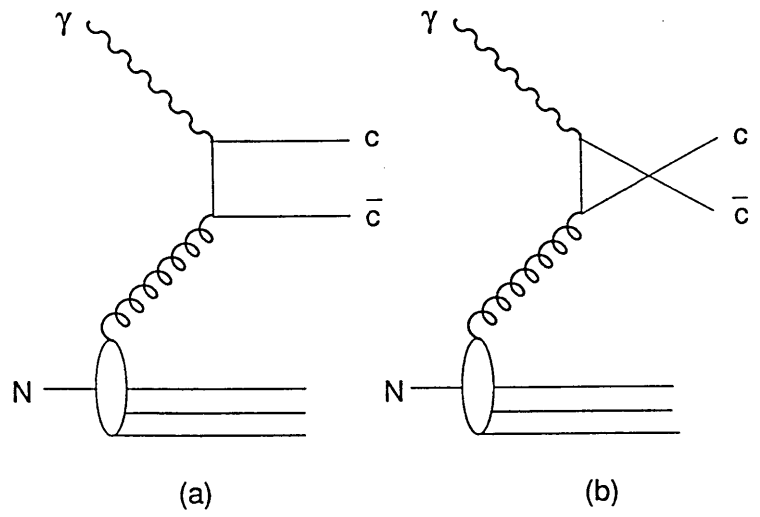


Figure 18: Feynman diagrams for photon gluon fusion

8.1.1.1 Cross section calculation

The standard perturbative QCD formula for the inclusive production of a heavy quark Q of mass m, momentum p and energy E,

$$\gamma(P_1) + H(P_2) \rightarrow Q(p) + X$$

determines the cross section as follows (up to corrections which are suppressed by powers of the heavy quark mass) [29]:

$$\sigma(S) = \sum_j \int dx \hat{\sigma}_{\gamma j}(xS, m^2, \mu^2) F_j^H(x,\mu) + \sum_{i,j} \int dx_1 dx_2 \hat{\sigma}_{ij}(x_1 x_2 S, m^2, \mu^2) F_i^\gamma(x_1,\mu) F_j^H(x_2,\mu)$$

The functions F_i^γ and F_j^H are the number densities of light partons (gluons, light quarks and anti-quarks) in a photon and hadron respectively, evaluated at a scale μ, expressed as a function of the momentum of the target nucleon carried by the parton, x. i and j denote the different light parton flavours. S is the square of the centre of mass energy of the colliding photon-hadron system. The symbol $\hat{\sigma}$ denotes the short distance cross section and is calculable as a perturbation series in $\alpha_S(\mu^2)$. It can be written as

$$\hat{\sigma}_{\gamma j}(s,m^2,\mu^2) = \frac{\alpha_{em}\alpha_S(\mu^2)}{m_2} f_{\gamma j}(\rho, \frac{\mu^2}{m^2})$$

and

$$\hat{\sigma}_{ij}(s,m^2,\mu^2) = \frac{\alpha_S^2(\mu^2)}{m_2} f_{ij}(\rho, \frac{\mu^2}{m^2})$$

where $\rho = 4m^2/s$, s is the square of the partonic centre-of-mass energy, and α_{em} is the electromagnetic fine structure constant. The dimensionless functions f_{ij} and $f_{\gamma j}$ are calculated using their perturbative expansion. The scale μ is *a priori* only determined to be of the order of the mass m of the produced heavy quark.

The first term of the above equation represents the point-like contribution of the photon, while the second term is the hadronic contribution. The separation of the two terms is controlled by the scale μ. Singularities due to the splitting of the photon into massless partons with a transverse momentum less than $O(\mu)$ are reabsorbed into the hadronic component. The splitting of the photon into partons with a transverse momentum larger than $O(\mu)$ is included in the pointlike contribution. It turns out that for the energy range we are considering the hadronic component contribution is small (less than 5% for $50 < E_\gamma < 400\ GeV$)

Contributions to the total charm cross section which, although included in the QCD formula (in the sense that they need not be calculated explicitly had the QCD formula been solved to all orders), are usually mentioned separately are:

- ***Diffractive dissociation***: states described by the first terms in the perturbation expansion of the QCD formula correspond to inelastic final states; On the other hand, diffractive dissociation leads to a different type of final state structure with quasi-elastic excitation of the beam or target particle which then decays into charm-anticharm pairs (Figure 19). The contribution, however, to the cross section is small [30].
- ***Intrinsic charm production***: Charm is expected to be present (at a low level) in the sea of the nucleon, and may therefore lead to a contribution to charm photoproduction through the QCD Compton process $\gamma c \rightarrow gc$. Similarly, the intrinsic charm constituent of the photon may undergo a hard collision with a gluon or a quark in the target nucleon (Figure 20). The resulting cross section, which decreases with energy, is negligible (of the order of 5nb at $E_\gamma \approx 50GeV$ compared to a few hundred nanobarns of the photon gluon fusion mechanism [30]).

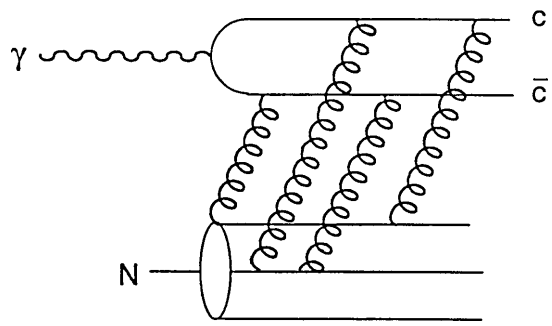


Figure 19: Diffractive dissociation

The first term in the perturbation series which contributes to the pointlike component is $O(\alpha_s \alpha_{em})$. At this order the only process which contributes to the heavy quark production is photon gluon fusion:

$$\gamma + g \rightarrow Q + \bar{Q}$$

The diagrams contributing to the lowest order cross section are shown in Figure 18. The invariant matrix elements squared and the cross sections for these processes have been available in the literature

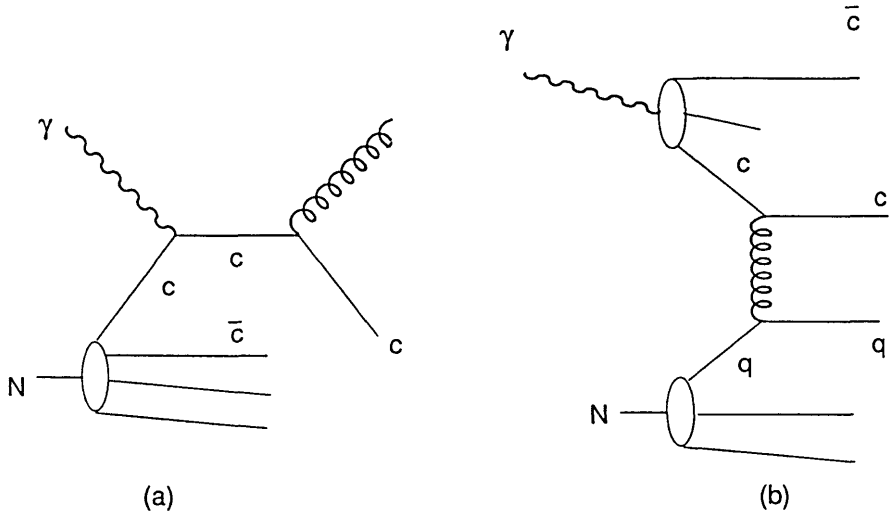


Figure 20: Intrinsic charm: (a) of the nucleon (QCD compton); (b) of the photon

for some time [31]. Recently, a calculation of the inclusive photoproduction cross section for heavy quark production to order $\alpha_S^2\alpha_{em}$ has been published [29]. To this order, the parton sub-processes which contribute to the inclusive cross sections are:

$$\gamma + g \rightarrow Q + \overline{Q} \qquad (order\ \alpha_S\alpha_{em}, \alpha_S^2\alpha_{em})$$

$$\gamma + g \rightarrow Q + \overline{Q} + g \quad (order\ \alpha_S^2\alpha_{em})$$

$$\gamma + q \rightarrow Q + \overline{Q} + q \quad (order\ \alpha_S^2\alpha_{em})$$

$$\gamma + \overline{q} \rightarrow Q + \overline{Q} + \overline{q} \quad (order\ \alpha_S^2\alpha_{em})$$

The formulae from the complete second order calculation are given in reference [29]. In the next section, we concentrate on the phenomenological consequences of those calculations.

8.1.1.2 Results and uncertainties of the QCD calculation

The calculation of the charm cross section depends mainly on the choices for the charmed quark mass, the QCD scale Λ, the mass scale μ, and on the form of the gluon momentum distribution (since the first order and largest contribution is that of gamma gluon fusion). The dependence of the cross section on the charmed quark mass is particularly important at energies accessible to fixed target experiments at present accelerators. The uncertainties in the other parameters lead to smaller variations and we discuss them first.

Choice of μ : If we choose the scale $\mu < 2m_c$ we enter into the region where the gluon distributions are not measured. In order to avoid the problem the value of $\mu \approx 2m_c$ is adopted, as is usually the case in such calculations. The uncertainties associated with the choice of μ will simply be taken to be of the order of a factor of 2 [29]. Note that the cross section does not strongly depend on the choice of μ, but only through the logarithmic variation of α_s with Q^2.

Choice of Λ : The other source of uncertainty is the lack of knowledge of the value of Λ, and the form of the gluon distribution. There is a strong correlation in these parameters [29]. The calculation of both quantities follows the same procedure.

From deep inelastic scattering experiments [32] it has been determined that gluons carry about half of the momentum of the nucleon, so we must have:

$$\int_0^1 xG(x)dx \approx \frac{1}{2}$$

where G(x) is the gluon density and xG(x) is known as the gluon structure function. Naive phase space considerations result in the 'dimensional counting rule' result [33] that for high x values

$$xG(x) \sim (1-x)^{2n_s-1}$$

where n_s is the minimum number of other partons involved (apart from the one participating in the interaction). For low Q^2 values, $n_s = 3$ (the three valence quarks of the nucleon) and hence

$$xG(x) \sim (1-x)^5$$

The above parameterisation gives reasonable agreement with the data at low Q^2 values, and is commonly adopted. The evolution of the structure function with Q^2 leads to an increase of the value of the exponent. High statistics deep inelastic scattering experiments have parameterised the gluon structure function applying the more general $x^{\alpha}(1-x)^{\beta}$ form. The parametrisation of one of these experiments [34] extrapolated to $Q^2 = 10\text{GeV}^2$ is supplied below:

$$G(x,Q_0^2) = 0.24x^{-0.64}(1-x)^{\beta(x)}$$

where

$$\beta(x) = 7.5 + 5.5\ln\{\ln[1/(1-x)]\}$$

The Q^2 evolution of the gluon structure function can be derived using the Altarelli-Parisi evolution equations and is discussed in reference [35]; this method is used for obtaining the gluon structure function used in our cross section calculation.

The value of Λ depends on the number of active flavours; in our case, the active flavours are 4 (u,d,s,c), and Λ is taken to be [29]:

$$\Lambda_4 = 260 \pm 100 MeV$$

The corresponding ranges for Λ_5 and Λ_3 are

$$\Lambda_5 = 170 \pm 80 MeV$$

$$\Lambda_3 = 310 \pm 110 MeV$$

For the cross section calculation the estimates of the errors due to the variation of μ and Λ (and the gluon distribution) are added in quadrature.

The full second order calculation results for the charm cross section can be seen in Figure 21, Figure 22, and Figure 23, corresponding to charmed quark masses of 1.2, 1.5 and 1.8GeV. The outer curves of those figures correspond to the theoretical uncertainties on the choice of μ and Λ. The middle curve, supplied for indicative purposes, corresponds to the values of Λ = 260MeV and μ^2 = 10GeV^2 and should not be taken to correspond to any preferred values; indeed, all curves are obtained using the same value for μ (μ = 10GeV^2). Experimental values for the total cross section (excluding the values obtained from this work) are also shown. From these figures it is clear that even with the

large theoretical uncertainties involved, one can distinguish between different values of the charmed quark mass. The data seem to exclude the low mass value, and the more accurate E691 points [36] seem to favour a heavier charmed quark mass than that which is commonly adopted (usually, the charmed quark mass is taken as half the ψ mass ie 1.5GeV; the value favoured by QCD sum rules is m = 1.46 $\pm$ 0.05 GeV).

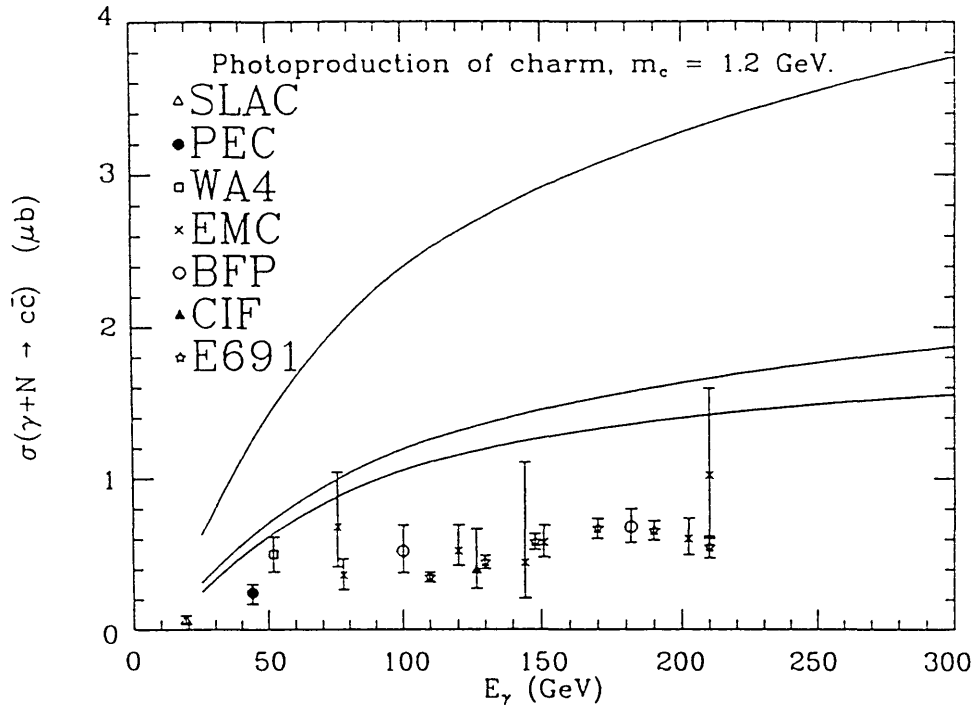


Figure 21: Total charm cross section when the charmed quark mass is fixed at 1.2GeV

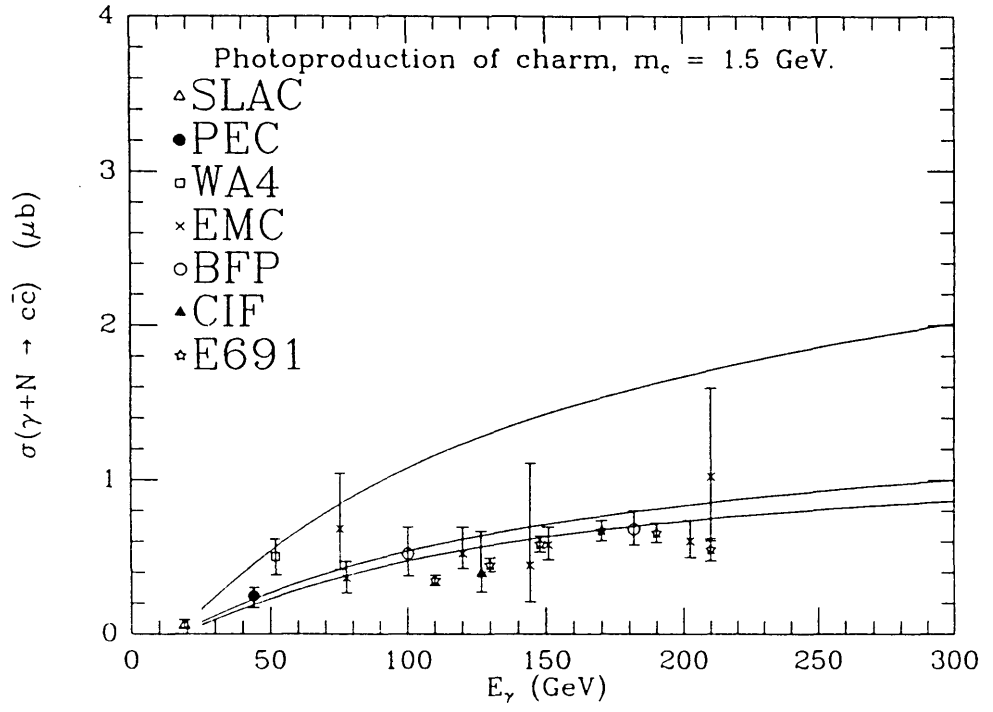


Figure 22: Total charm cross section when the charmed quark mass is fixed at 1.5GeV

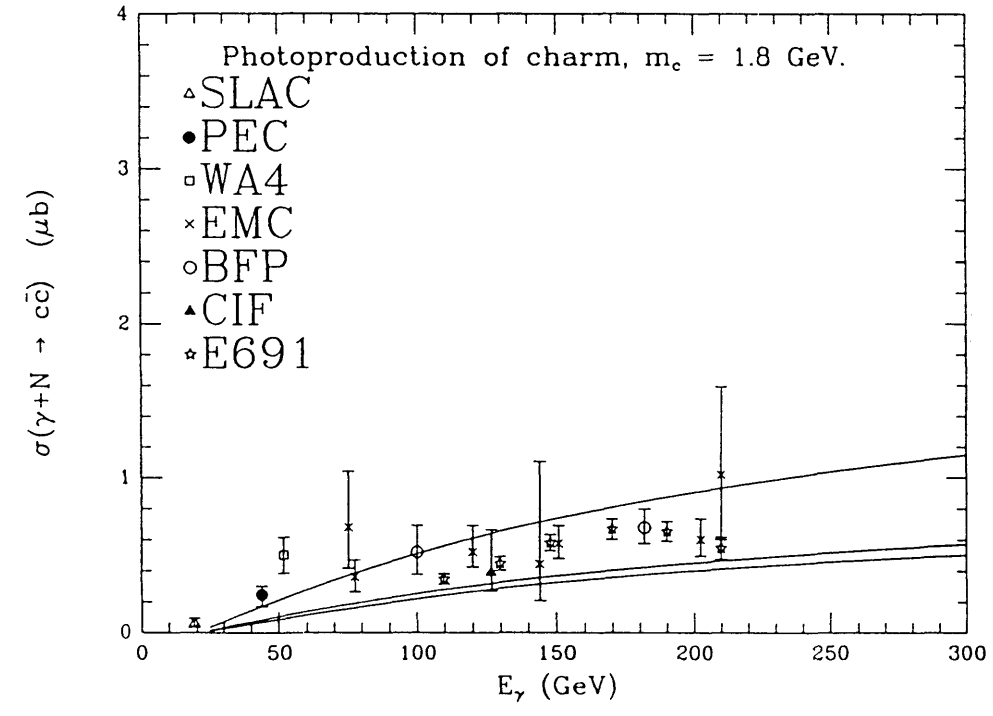


Figure 23: Total charm cross section when the charmed quark mass is fixed at 1.8GeV

8.1.2 Nuclear effects

The above cross section has been calculated for free target nucleons. In this experiment, however, target nucleons are not free but bound in nuclei (silicon), resulting in a 'shadowing' effect (the obscuring effect of the nucleus to individual nucleons). This effect in the measured cross sections is usually parametrised as:

$$\sigma(\gamma A \rightarrow c\bar{c} + X) = A^{\alpha}\sigma(\gamma N \rightarrow c\bar{c} + X)$$

where A is the mass number of the nucleus. α indicates the amount of shadowing present, and it is equal to 1 for no shadowing and 2/3 for geometric shadowing. The high mass of the charmed quark makes charm photoproduction a hard process, and furthermore the total hadronic cross section of the photon is relatively low (compared with that of hadrons), so little shadowing is expected. Experimentally, a value of $\alpha = 0.94 \pm 0.02 \pm 0.03$ has been measured for incoherent Ψ meson photoproduction [37]. For the photoproduction of all hadronic final states $\alpha = 0.920 \pm 0.002$ [38]. For Silicon (A = 28.1), our target material, this difference in the A dependence makes the ratio of hadronic to charm photoproduction differ by 6% ($\pm$11%) from the free target nucleon case.

8.2 Hadronisation mechanisms

Perturbative QCD does not specify the final state hadronisation, that is the way that c and $\bar{c}$ quarks turn into charmed mesons and baryons. Describing completely this process is out of reach given our present understanding of the confinement problem. Therefore, to solve the problem of hadronisation, further assumptions are needed and we shall adopt the framework of the dual parton model (DPM) [39].

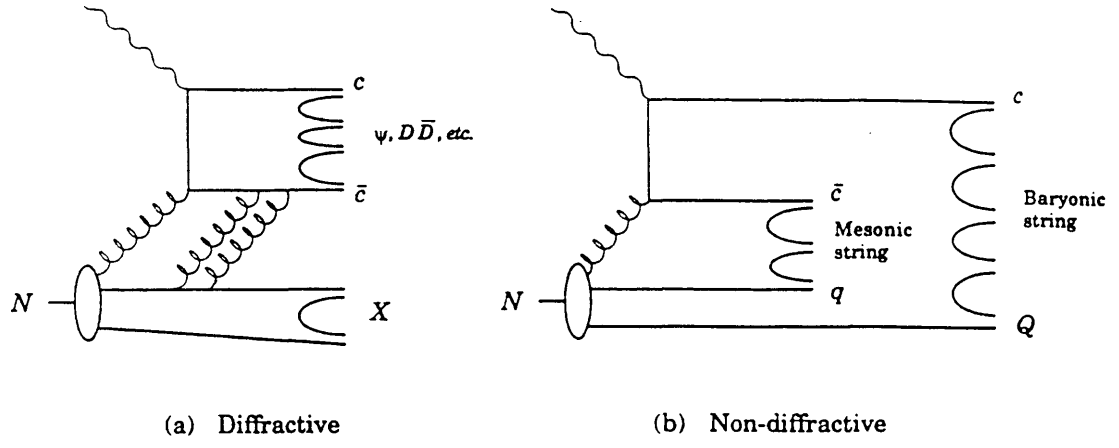
The hadronisation of a $c\bar{c}$ pair can be classified in two broad categories:

- Diffractive events where the projectile and hadronic systems hadronise independently ((a) in Figure 24). This class of events contains interactions which are coherent over nucleons of the nucleus.
- Non-diffractive events in which the charm and anticharm quarks hadronise independently ((b) in Figure 24).

From theoretical arguments it can be shown that the fraction of diffractive events expected is small (a few percent) [30]. A search for coherent events using the information of the active target of NA14 (events where there was no evidence of nuclear break-up) confirms this evaluation, placing an upper limit on the diffractive component at 3% (at 90% confidence level) [40], therefore we shall here only consider non-diffractive events.

8.2.1 The hadronisation scheme

According to the DPM the final state hadronisation is dominated by a single topological graph (b in Figure 24) The hadronisation occurs along two strings, a mesonic string stretched between the anticharm quark and a target spectator quark, and a baryonic string stretched between the charmed quark and the remaining target spectator diquark (denoted by Q in all relevant figures). This process can be simulated using Monte Carlo techniques.



(a) Diffractive (b) Non-diffractive

Figure 24: Hadronisation mechanisms

The NA14 collaboration has developed a Monte Carlo program [41] using the first order QCD calculation for the photoproduction of charm, via the γg fusion mechanism. The gluon structure function $xG(x)\sim(1-x)^5$ is assumed, no Q^2 evolution is introduced, and the mass of the charmed quark is

taken as 1.5GeV. The structure functions of the nucleons are taken from low p_T physics; the quark structure function is taken to be

$$f(x) \sim \frac{1}{\sqrt{x}}(1-x)^{3/2}$$

giving an $x^{-1/2}$ behaviour for the quark and an $x^{3/2}$ behaviour for the diquark at low x ($x_Q = (1-x_q)$ and $x_q << 1$). Hence most of the energy is taken by the diquark as shown in Figure 25, where the masses of the two strings generated by the Monte Carlo simulation are shown. This inequality in the string masses can lead to a particle-antiparticle asymmetry for the charmed particles produced.

The hadronisation is non-diffractive and hadrons are created along the two independent strings using the Lund scheme [42] in which quark-antiquark and diquark-antidiquark pairs are created along the length of the string until the available energy has been exhausted.

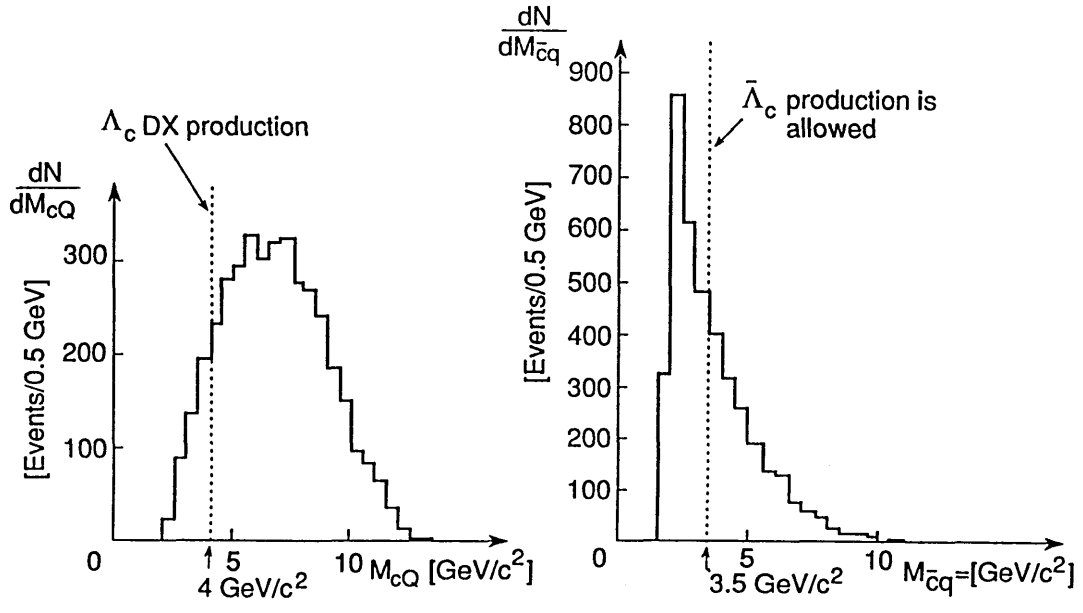


Figure 25: Dual parton model string masses

8.2.1.1 Particle production ratios and particle-antiparticle asymmetries

In the perturbative QCD calculation of charm photoproduction, the charmed quark and anti-quark are produced in an equivalent fashion; the source of any particle-antiparticle asymmetry must therefore be attributed to hadronisation.

Figure 26 shows the results of the Monte Carlo simulation. Here the expected production rates for the various charmed particles are given as a function of the incident photon energy. These rates are needed when deriving the total inclusive charm cross section from a measurement of individual particle production. Below we shall briefly discuss the major Monte Carlo predictions for particle production asymmetries and ratios.

A key issue in particle-antiparticle symmetry is Λ_c production, since Λ_c is the lightest and hence the most abundantly produced charmed baryon.

$\Lambda_c/\overline{\Lambda}_c$ *asymmetry* : The Lund Monte Carlo used for final state hadronisation generates diquarks and antidiquarks with equal probability. However, the Λ_c baryon contains a c-quark and is therefore formed in the baryonic string, whilst the $\overline{\Lambda}_c$ contains a $\bar{c}$-quark, and is formed in the mesonic string. In the mesonic string, $\overline{\Lambda}_c$ production can occur only if the string mass is larger than 3.2GeV (equal to the mass of the Λ_c plus the mass of a nucleon), due to the requirement of baryon number conservation. The baryonic string, however, when above the 2.3GeV Λ_c threshold, may form a low mass system with the charmed quark and the spectator diquark without stretching further to produce additional diquark-antidiquark pairs. This low mass system may then decay into a Λ_c plus additional mesons. This process has a sharp energy dependence, and it is entered manually into the simulation: If the baryonic mass string is less than 4GeV, then only Λ_c baryons are produced; otherwise a *DNX* system (a D meson, a nucleon, plus additional hadrons) is generated. The value of 4GeV was chosen to reproduce the measured Λ_c/D ratio at $E_\gamma = 20GeV$ ($\Lambda_c/D = 0.7$) [43]. A more elaborate method that avoids this *ad-hoc* 4GeV cut [30] yields similar results.

Therefore, at low energies, an excess of Λ_c over $\overline{\Lambda}_c$ production is expected. This excess is also enhanced from the fact that the mesonic string mass distribution is peaked towards lower values compared to the baryonic string. For high Λ_c energies, however, (above 40GeV in our case) Λ_c and $\overline{\Lambda}_c$ production is similar.

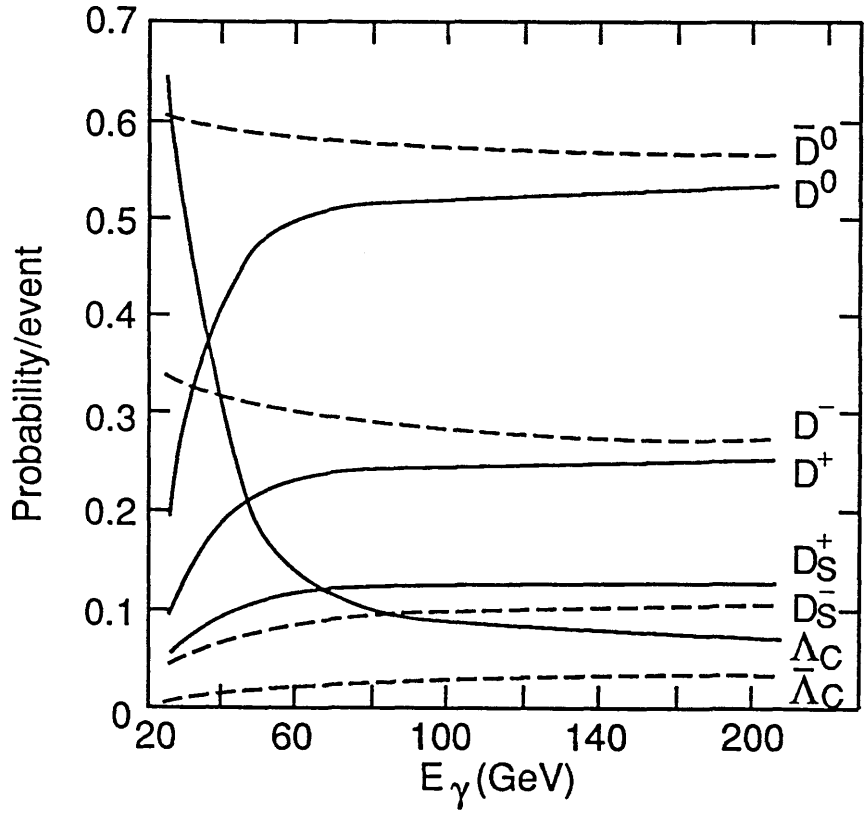


Figure 26: Production rate predictions

$D/\overline{D}$ asymmetry : The $D/\overline{D}$ asymmetry reflects the $\Lambda_c/\overline{\Lambda}_c$ asymmetry. As the rate of baryon production is of the order of 10% of the rate of meson production, the expected excess of $\overline{D}$ is only a few percent.

D_s^+/D_s^- asymmetry : The D_s^+/D_s^- asymmetry is sensitive to the way charmed hadrons are produced: in spite of the fact that, as for the D meson, we would expect an antiparticle excess, there is a

further effect due to phase space limitation in the mesonic string which prevents $D_s^- KX$ production, which leads to an excess of D_s^+ over D_s^-.

D^+/D^0 ratio : The ratio of D^+ to D^0 production depends on the production rate of prompt D^* mesons. This is because the D^{*0} cannot decay to the D^+ (due to kinematics), whereas the D^{*+} decays to D^0 or D^+ with equal probabilities. If we take the branching ratio $Br(D^{*+} \rightarrow D^0\pi) = 0.49 \pm 0.08$ [44] and a ratio of $D^0/D^* = 0.32 \pm 0.01 \pm 0.03$ [36] we predict $D^+/D^0 = 0.41 \pm 0.03$.

D_s^+/D^+ ratio : The production ratio of D_s^+ over D^+ provides information on strangeness production during hadronisation since a charmed quark has to combine with a strange quark of the sea to produce a D_s. Thus D_s^+ production is expected to be suppressed with respect to D^+ production. If we assume the same probability for strange quark production as the one used to account for measurements in e^+e^- annihilation events, we predict a value of 0.38 ± 0.15 for this ratio for incident photon energies greater than 60GeV [45].

Λ_c/D ratio : This production ratio of charmed baryons and antibaryons over D mesons depends on the rate of diquark generation during the process of hadronisation. As threshold effects introduce an asymmetry between Λ_c and $\overline{\Lambda_c}$ giving an excess of Λ_c production from direct fusion of a charmed quark with a target diquark, we should only compare the rates for $\overline{\Lambda_c}$ production to the corresponding production in e^+e^- annihilation. This comparison will be meaningful for high incident photon energies to be away from phase space effects. We predict $\overline{\Lambda_c}/(D^0 + D^+) = 0.03$ at $E_\gamma = 90GeV$.

8.3 Charm decay

Charm decay has been the subject of extensive theoretical study. In this section we shall give a brief overview of the current understanding of weak charm decays.

The weak decay of charmed particles may proceed via three channels:

1. **Purely leptonic** . For a pseudoscalar meson D, the decay is of the form $D \rightarrow l\bar{\nu}_l$. This may occur through the annihilation of the quarks of the meson, as shown in Figure 27.

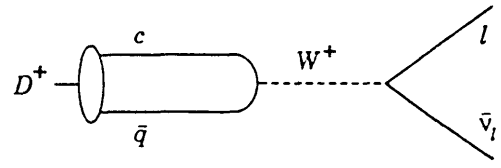


Figure 27: Leptonic decay

These decays are strongly helicity suppressed: due to the (V−A) structure of the weak charged current and the conservation of angular momentum, a spin zero state cannot decay into a massless fermion-antifermion pair via the weak charged current; when the fermion masses are small compared to their energy the process is allowed, but strongly suppressed.

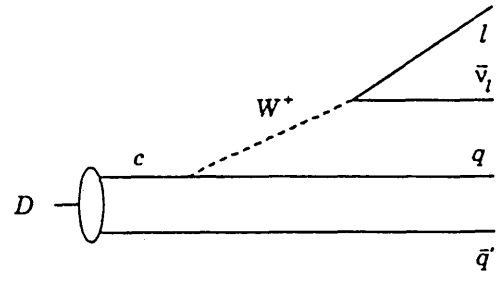


Figure 28: Semileptonic decay

2. **Semileptonic** , i.e. $D \to X l \bar{\nu}_l$. This occurs primarily through the β-decay of the charmed quark (one possible diagram is shown in Figure 28). Decays of this form are known as spectator decays, as the light quark in the charmed meson is assumed not to play any significant part in the decay.

3. **Hadronic** . The weak decay of charmed mesons to exclusively hadronic final states may also occur via a spectator decay; in this case, however, there are other diagrams which may contribute, involving quark annihilation and W exchange. These can be seen in Figure 29. Although the last two processes are helicity suppressed, QCD corrections involving gluon radiation result in allowed decays.

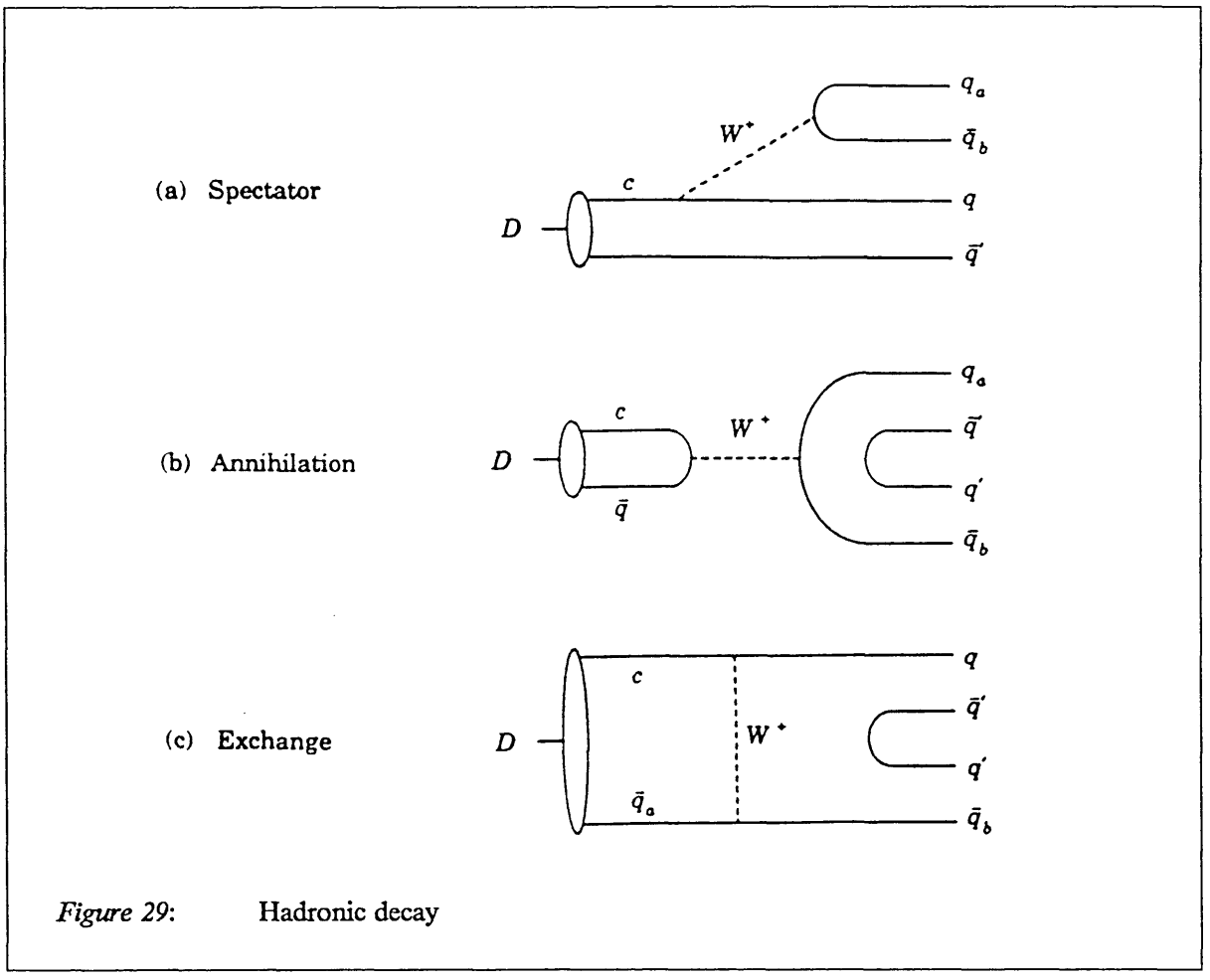


Figure 29: Hadronic decay

8.3.1 The spectator model

The spectator model is motivated by the large mass of the charmed quark, which permits it to be treated as though it were essentially free, with the light quark not contributing to the dynamics of the decay. There is then a direct parallel between the charmed spectator decays and muon decay for which the width is readily calculable in weak interactions and is well measured experimentally. The charm

lifetime, which depends on the fifth power of the charmed quark mass, is predicted to be about 7×10^{-13}s (if $m_c = 1.5GeV$). The measured lifetimes of the D^0, D^+ and D_s are in the range $10^{-13} - 10^{-12}$s, within an order of magnitude from the theoretical prediction. This prediction is, however, not perfect since the lifetimes of the charmed mesons are not exactly the same, as predicted by the model. The inclusion of QCD corrections to the naive spectator model reduces all lifetimes but does not explain the differences. For these to be explained, a more complex picture of charm decay needs to be constructed, where non-spectator diagrams contribute as well.

Charmed particle decay is not the subject of this thesis and the discussion stops here. (see [46] for further details). The essential result is the range of lifetimes of the particles under study: with typical energies of order 30GeV they travel a distance of order $(\frac{E}{m}c\tau \approx 1mm)$ before decay, and a high precision vertex detector is required for their identification.

Chapter 9

EXPERIMENTAL SETUP

9.1 Overview

The NA14 spectrometer is situated at ECN3, one of the underground halls of the North Area of CERN. The experiment utilizes a high energy photon beam created indirectly using the 450GeV protons of the SPS. The photon beam is tagged, i.e. the energy of individual photons can be measured.

The spectrometer can be seen in Figure 30.

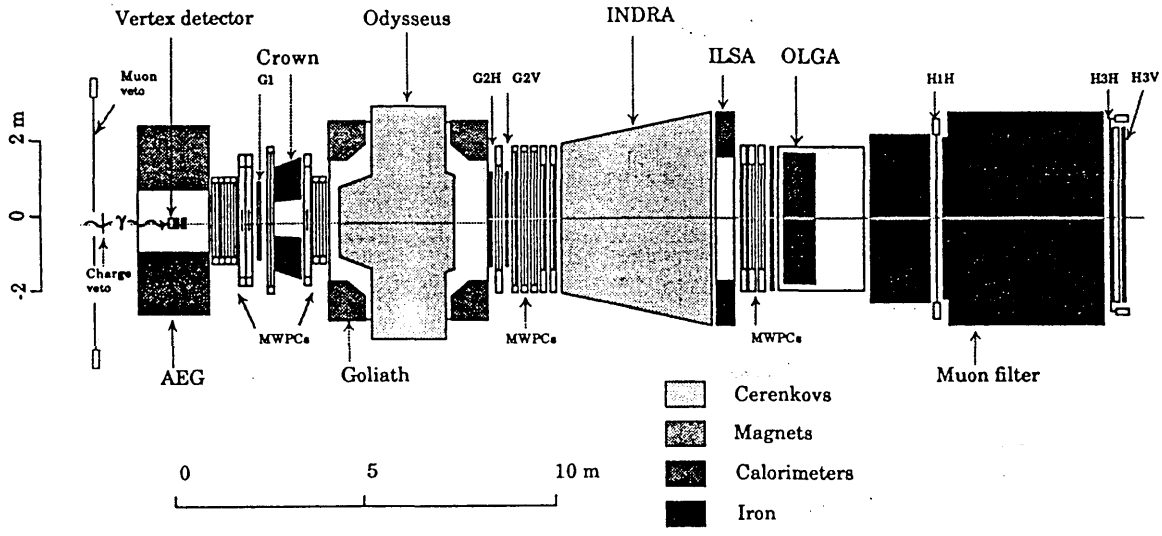


Figure 30: The NA14 Spectrometer

It incorporates a high resolution vertex detector consisting of an active target and a microstrip telescope. Charged particle tracking is provided by three stacks of multi-wire proportional chambers (MWPCs). Momentum and charge measurement are assisted by the use of two room temperature magnets, the AEG and Goliath. The energy of photons and π^0's is measured using a three calorimeter setup, the Crown, ILSA and OLGA, between them covering a wide angular acceptance, matching that of the rest of the apparatus. Particle identification is performed using two Cerencov counters, Odysseus and INDRA, filled with different radiating media. Finally, at the downstream end of the detector, an iron filter helps to identify muons.

The trigger used for data taking is a minimum bias hadronic trigger. Data acquisition is performed by a PDP-11/45 computer and data are written onto 1600bpi tapes. During the main data taking periods of July 1985 and July−September 1986, representing 70 days of total running time, 17 million events were recorded.

The coordinate system chosen to describe the apparatus is Cartesian, with the X-axis along the beam direction and the Y-axis horizontal. Magnetic fields are in the Z-direction. The origin is taken as the centre of one of the spectrometer magnets (the AEG) and corresponds approximately to the position of the active target.

9.2 Beam

9.2.1 Beam production

The experimental beam is produced indirectly using 450GeV protons from the SPS. The SPS cycle is about 14 seconds; there is a 12 second acceleration stage, followed by a 2 second *spill*, the extraction of the protons from the SPS ring by kicker magnets that deflect them into various experiment targets.

The beam is known as the 'Broadband Electron and Gamma' beam. It is produced using a three stage process:

$$p \rightarrow \gamma \rightarrow e^- \rightarrow \gamma$$

This scheme has two main advantages: firstly, it can produce a very pure photon beam with low hadron contamination and secondly, the beam may be *tagged*, i.e. the energy of individual photons may be measured. This is particularly important if production mechanisms are to be studied. The production of the beam is illustrated in Figure 31.

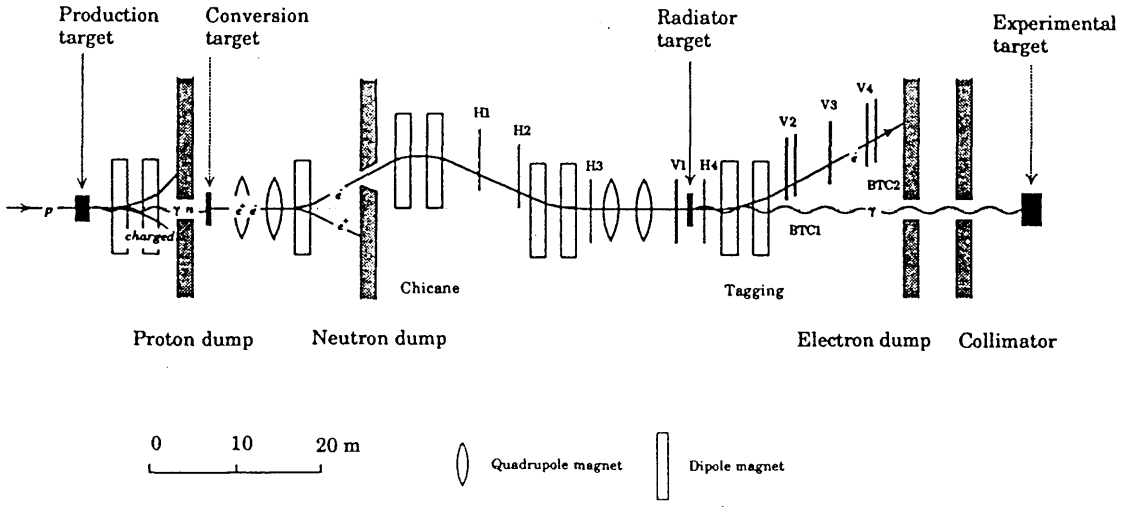


Figure 31: The NA14 beam line

The protons are dumped into a production target (SPS target T6 or T10, depending on the period) consisting of a block of beryllium 100mm in length. About 20% of the protons interact in the target. Amongst the interaction products are neutral pions which decay electromagnetically:

$$\pi^0 \rightarrow \gamma\gamma$$

to give photons that are used in the next stage. The target material and length represent a compromise between the requirements that the protons should be used efficiently, implying a high probability of nuclear interaction, i.e. a short nuclear interaction length; and that the loss through conversion:

$$\gamma \rightarrow e^+e^-$$

of the produced photons should be minimised, implying a long radiation length. The ratio of radiation length to nuclear interaction length is largest for elements of low atomic number, such as beryllium (see Table 11).

Table 11: Properties of materials

Material	Z	Density [g cm^{-3}]	Radiation length [g cm^{-2}]	Nuclear interaction length [g cm^{-2}]
Beryllium	4	1·85	65·2	75·2
Silicon	14	2·33	21·8	106
Lead	82	11·4	6·4	194

A further consideration in the choice of length of the production target is that the creation of muons forming the muon halo, a source of background to the experiment, is greater for longer lengths.

The products of the proton interactions pass through a pair of sweeping magnets that deflect all charged particles, including those protons that have not interacted, into a 4m thick iron wall, the *proton dump*. The photons, and any other neutral particles (such as K^0 and neutrons) pass through an aperture and strike a conversion target consisting of a lead sheet 4mm thick. About 40% of the photons convert to give e^+e^- pairs. Here the conflicting requirements are an efficient conversion of the photons but a minimisation of the contamination from charged particles produced by the interaction of neutral hadrons. A low ratio of radiation length to nuclear interaction is therefore needed, hence the choice of lead. The electrons are selected using a series of five bending magnets in the form of a chicane, the neutral and positive particles being dumped into another iron wall, the *neutron dump*. The aperture through which the electrons pass has adjustable jaws, permitting variation of the momentum range selected; for a typical selection of 120 − 250GeV/c there are about 10^8 electrons per burst with a mean energy of about 150GeV. Electrons rather than positrons are selected to reduce the muon contamination, since about twice as many μ^+ as μ^- are produced in the initial proton interactions. The electron beam is then focussed on the radiator target, a 0.5mm thick lead sheet. Here the electrons undergo bremsstrahlung, the emission of a photon as a result of deceleration in the electric field of a nucleus of the medium, in this case lead. A small percentage (about 14%) of the electrons radiate a photon of energy greater than 40GeV, and it is these photons that are used to provide the experimental beam (the detector trigger efficiency is low for photon energies lower than 40GeV). Increasing the

radiator thickness would increase the photon yield, but also increase the probability of multiple bremsstrahlung (the emission of more than one high energy photon from the same electron) making the measurement of the photon energy by the tagging system inaccurate. The tagging magnets sweep all charged particles out of the beam, and onto the *electron dump*, an iron wall. The photons are not deflected, and pass through an aperture in the electron dump and a collimator, to reach the experiment target.

9.2.2 Tagging

The tagging system determines the energy of an incident photon from the difference of the electron energy before and after radiating. It is implicitly assumed that no multiple bremsstrahlung has occurred. The electron energy is found from a measurement of its momentum through deflection in a magnetic field. Upstream of the radiator target, the field is provided by the chicane magnets, and downstream by the tagging magnets. The electron trajectory is determined using eight scintillator finger hodoscopes, H1−4 and V1−4. Due to the high electron flux, the hodoscopes require fast time resolution and high segmentation to enable the trajectory of the radiating electron to be unambiguously reconstructed. The *back tagging counters* (BTCs) provide timing information to be used for distinguishing between ambiguous tagging solutions. They also provide a time reference for the beam trigger.

For the data taking period of September 1986, the period that provided most of the data analysed for this work, unambiguous tagging solutions are found in about 20% of the events, whilst a further 10% have ambiguous solutions due to multiple trajectories reconstructed in the hodoscopes. The efficiencies are low due to hardware problems experienced with the tagging system during that period of data taking. The granularity of the hodoscopes gives rise to an r.m.s error in the photon energy of about 3GeV. However, multiple bremsstrahlung introduces a much more significant uncertainty in the photon energy estimate, as we shall discuss at some length in chapter 11.

9.2.3 Beam properties

The (bremsstrahlung) beam energy profile is expected to have an approximate (1/energy) shape. It is not, however, directly measurable. The energy profile of *tagged* and *trigger accepted* photons which represents a small percentage of the total number of incident photons, can be seen in Figure 32.

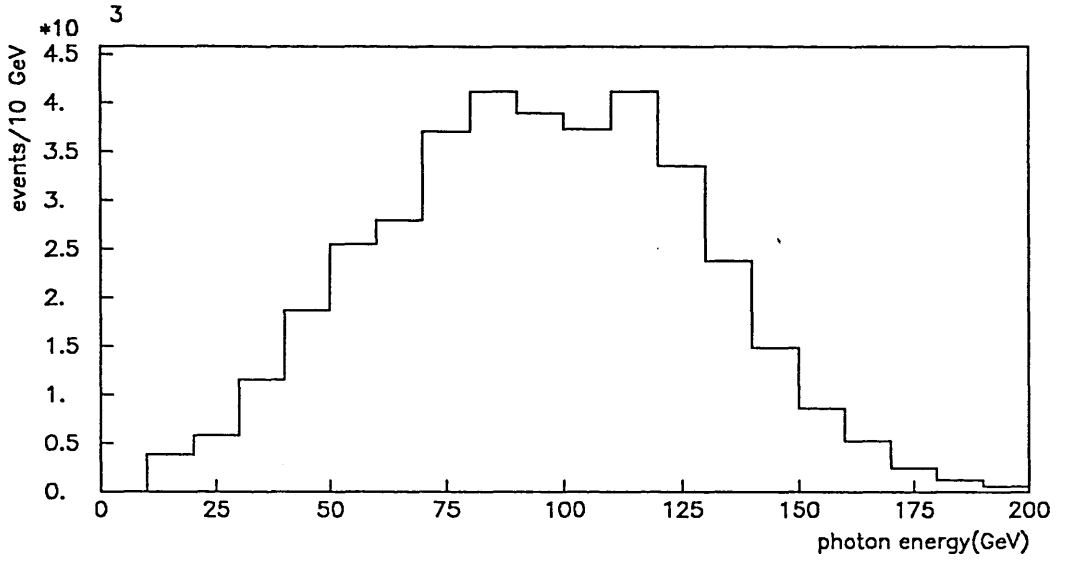


Figure 32: Energy spectrum of tagged and trigger accepted photons

The spectrum has been derived using a Monte Carlo simulation tuned to reproduce the observed tagging answer spectrum. The mean energy is about 100GeV. The beam spot size is approximately 12 mm in the Z projection (FWHM); for the Y projection the beam size is bigger than the active target area included in the trigger; the Y beam size, defined by this trigger requirement, is 28mm (FWHM).

9.2.4 Sources of background

9.2.4.1 Hadronic

The photon beam is very clean of hadronic contamination, to the level of 1 part in 10^5. The contamination is the result of a small number of pions that accompany the electron beam with a ratio π/e of about 10^{-3}, produced by the decay of neutral hadrons, principally Λ's and K^0's. They are swept out of the beam by the tagging magnet, but any neutral particles that they produce through interaction in the radiator will contaminate the beam.

9.2.4.2 Electromagnetic

A serious source of background for the experiment is of electromagnetic origin. 98% of the photons reaching the experimental target have energy less than 50GeV; these low energy photons are a result of either bremsstrahlung in the radiator, or synchrotron radiation caused by the deflection of the electron beam in the downstream tagging magnets. These photons can convert in the experiment target to form e^+e^- pairs, at a rate that would blind the detector if precautions were not taken. Their opening angle, however, is small and therefore they are swept into a horizontal band by the spectrometer magnets. Detectors that would be adversely affected by the pairs are desensitised over a corresponding central region. This electromagnetic background is excluded at the trigger level, as discussed in section 4 of this chapter.

9.2.4.3 Muon halo

Muons are created both in the production target and in the proton dump. They are highly penetrating, since, on the one hand, being leptons they do not interact via the strong force, and on the other being about 200 times as massive as electrons, they lose less energy radiatively. As a result there is a flux of about $10^6 m^{-2}s^{-1}$ muons that penetrate the intervening material and reach the experimental apparatus. The muon halo is excluded at the trigger level by the muon veto, a wall of scintillator hodoscopes.

9.3 Detectors

9.3.1 Vertex detector

The vertex detector of the NA14 spectrometer is a solid state silicon detector. A silicon detector consists of a number of reversely-biased p-n junctions, in other words a number of diode strips. It relies on the property of charged particles to create electron-hole pairs when passing through such a medium. A detailed account on silicon detectors as well as details on construction, testing and operation of the microstrip tracking chamber can be found in references [47] and [46].

The NA14 vertex detector comprises two parts: a multilayer active target [48] and a microstrip tracking chamber [49]. The latter is mainly used for vertex reconstruction in the analysis. The active target complements the microstrip tracking chamber in this task and it is also used in the experimental trigger. A schematic drawing of the detector together with a scale drawing can be seen in Figure 33. The two detectors together with their electronics (the pre-amplifiers of the microstrips) are mounted on a wheeled chariot, which permits movement of and access to the detector from its (inaccessible) position inside the AEG magnet. During data taking, the chariot was locked in its correct position with the active target at the middle of the AEG magnet.

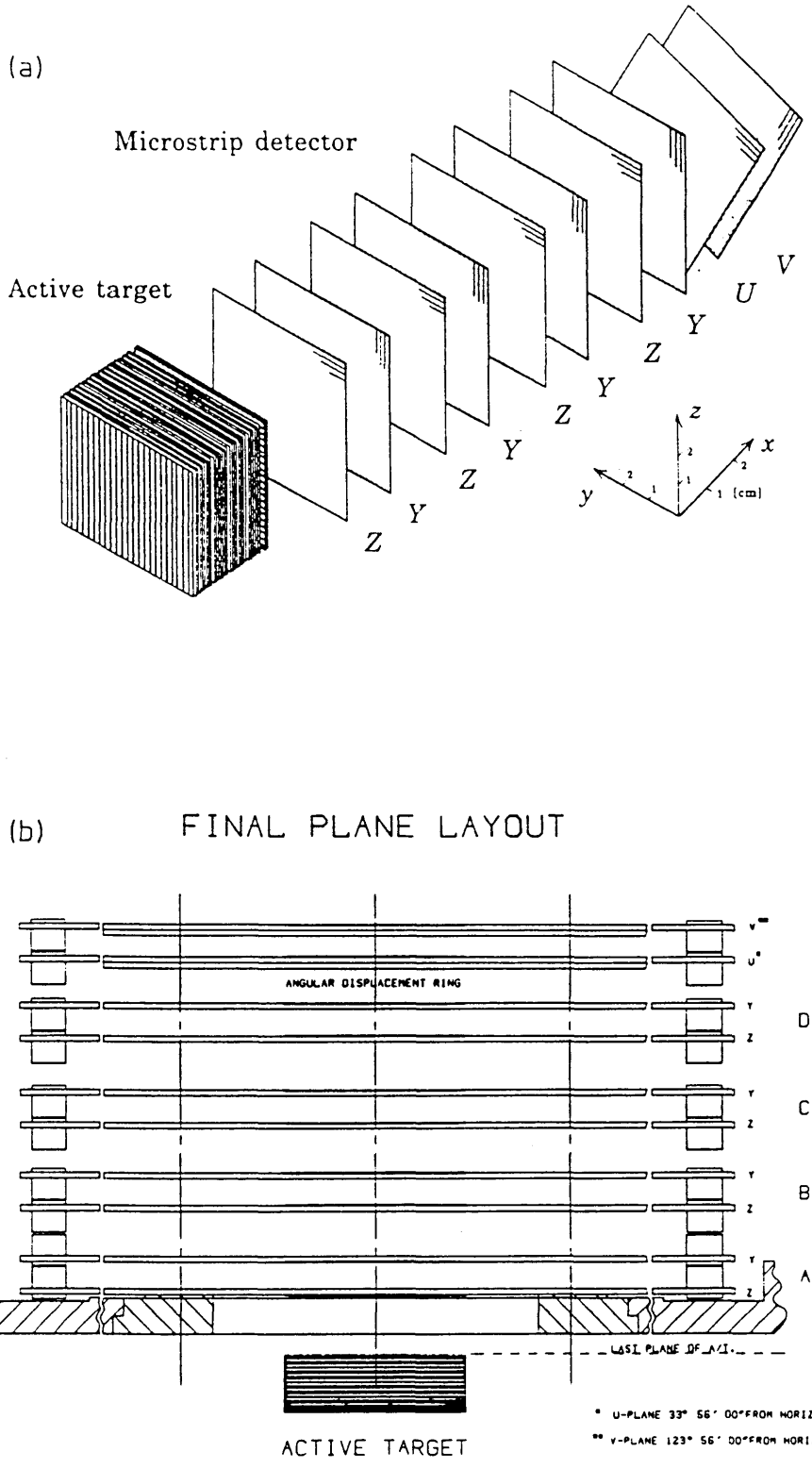


Figure 33: Schematic and scale drawings of the NA14 vertex detector

9.3.1.1 Active target

The active target consists of 32 planes of silicon, each with an active area of 5.0 x 4.4 cm and 300μm thick. The planes are separated by 512μm along the X (beam) axis. They are divided into parallel strips 2mm wide at a pitch of 2.1mm. The first 30 planes are oriented vertically providing information about the y coordinate of the interaction and are known as the Y planes. The last two planes have horizontally oriented strips and are known as Z-planes; these are used in the experimental trigger. The material of the target constitutes about 10% radiation length and 3.2% nuclear interaction length.

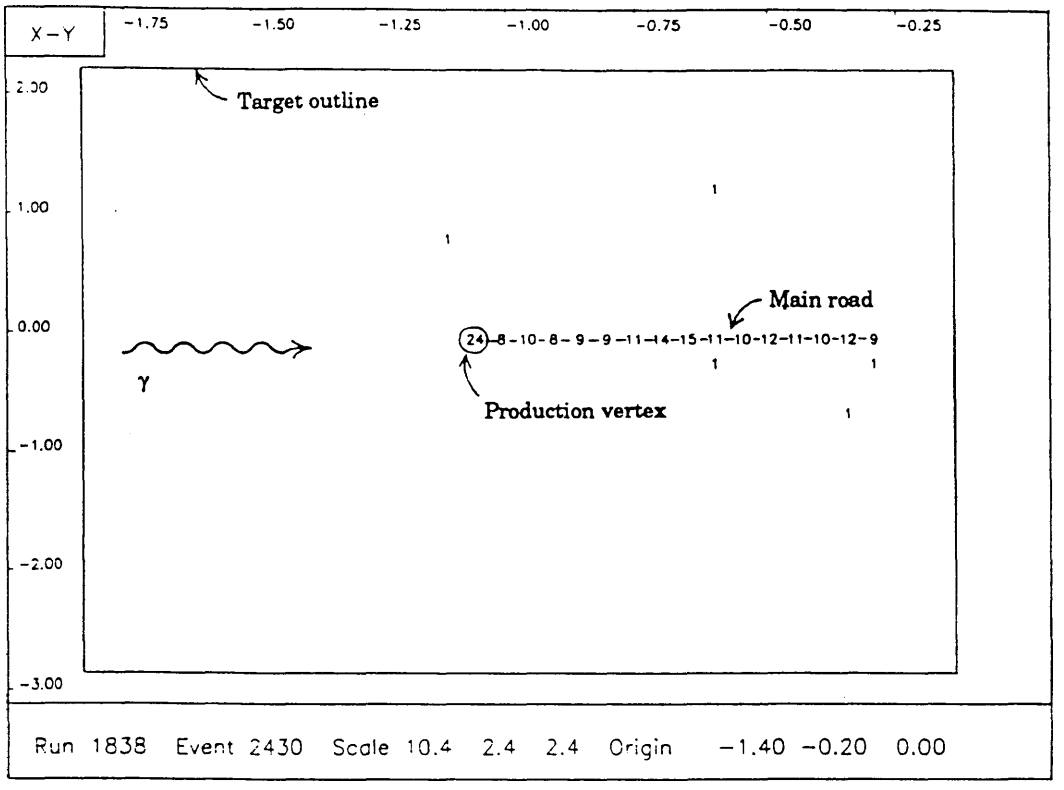


Figure 34: A clean active target event

Each of the strips has analogue readout, providing a direct measure of the ionisation deposited. A typical 'clean' active target event is shown in Figure 34. The rectangle shown is the outline of the target. Pulse heights in the Y-strips are displayed in units of minI (the signal expected for a relativistic

singly charged traversing particle). The row of strips that show activity lie downstream of the production vertex, and are known as the *main road.* For clean events of this type the production vertex is easily localized; it lies in the first strip of the main road. Note the high pulse height in this strip, caused by highly ionising nuclear recoil, which could further aid in the determination of the production vertex position. Since the plane thickness is 300μm, the x-coordinate of the vertex may be determined to within 150μm. However, the proportion of 'clean' active target events, mostly corresponding to coherent production, only corresponds to a few percent of all events; a more typical event is shown in Figure 35. It is an example of an incoherent event, where there has been nuclear break-up, giving rise to 'grey' tracks at large angles to the beam direction. In this case, these tracks make the pattern recognition problem easier; however, the granularity of the active target together with the problem that a backward-going grey track cannot in most cases be easily distinguished from the main road, make primary vertex finding using only the active target information a difficult task.

It was originally intended that a study of the evolution of the charged particle multiplicity along the length of the main road would allow the detection of secondary, decay vertices, since the decay of a particle to charged products would give an increase in multiplicity. Secondary vertices may therefore be detected by searching for a multiplicity step. However this is not an easy task given the Landau fluctuations in the ionisation of individual strips; due to these fluctuations it is necessary to average over a number of strips, hence losing resolution. Therefore, the active target cannot, in most cases, be used to reconstruct vertices by itself; it can complement, however, microstrip information for the determination of the X-coordinate of a vertex *after* the vertex position has been measured using the microstrip tracking chamber information. Operated in this ('post-microstrip') mode, its main contribution is the rejection of secondary interactions and double events. Secondary interaction events are events where a particle from the production vertex has re-interacted in the target, whereas double events are events where two photons interacted in the target. Such events have the same topology as events with a charm decay and therefore cannot be discarded using only the microstrip information. However, secondary interactions are usually accompanied, as we have already seen, by intense activity at the interaction strip due to nuclear recoil. This can be used to reject such events from the charm

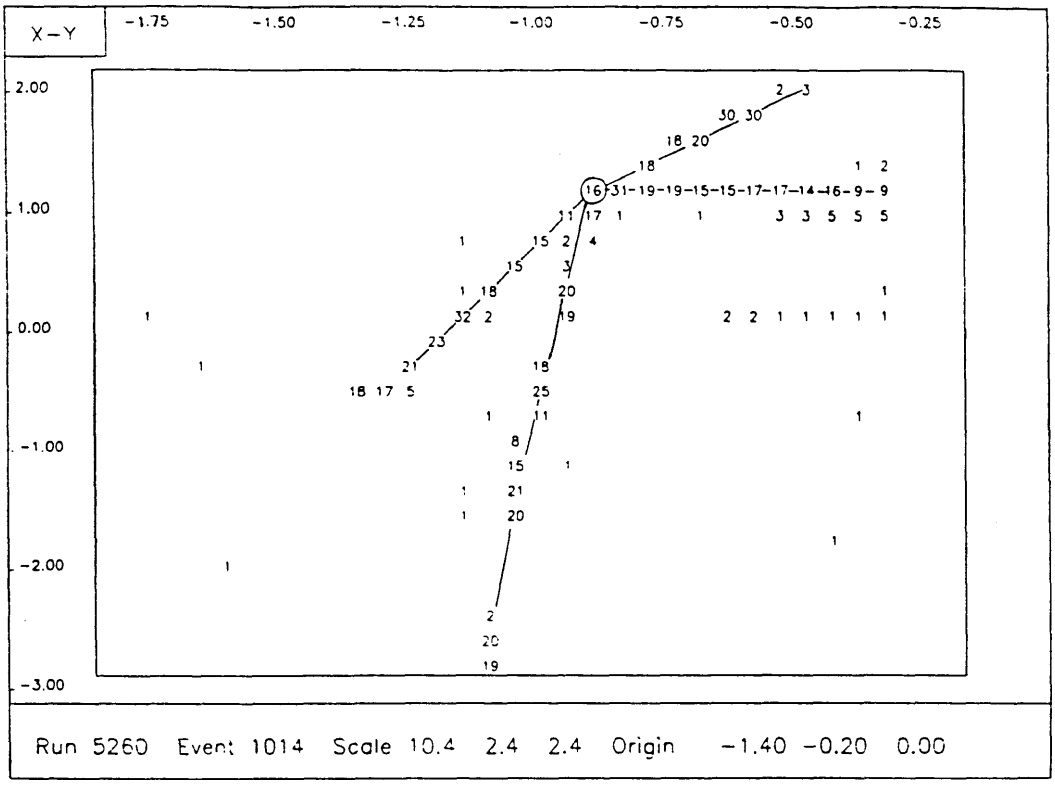


Figure 35: A typical active target event

candidate sample [50]. Another 'post-microstrip' active target usage is the rejection of a primary vertex when reconstructed in a region of no active target activity [46].

9.3.1.2 Microstrip tracking chamber

The microstrip detector consists of 10 planes of silicon of dimensions 5cm × 5cm and 450μm thick. Each plane is divided into 1000 diode strips with a 50μm pitch, giving a total of 10,000 channels to be read out. The individual planes only provide two coordinates of a track passing through them: the x-coordinate of the plane, and a transverse coordinate from the position of the strip that is hit. The first eight planes are therefore arranged such that each plane has its strips orthogonal to its neighbour(s). Planes of similar orientation will then give a projection in one view. Thus two orthogonal views (Y and Z views respectively) are obtained. However, when there is more than one track, the

association of the track projections to form tracks in space is ambiguous. For this reason, a third projection is required. The last two planes, known as the U and V planes, are rotated through approximately 33^0 about the beam axis and are used to remove this ambiguity in track reconstruction.

The area of the planes was limited by the maximum that could be achieved with currently available 3 inch silicon wafers and provides an acceptance of $\pm$250mrad, which closely matches the acceptance of the AEG magnet in which the detector is situated. The 50μm pitch also represented the limit at which strips could be wire-bonded to external printed circuit boards. The positioning of the planes was decided upon after a Monte Carlo study that suggested even $Y-Z$ pair spacing, with the upstream planes as close to the active target as possible, and last planes as far away as could be achieved without loss of acceptance.

The signals from the silicon are fanned out along the tracks of the printed circuit boards (PCBs) and are then transmitted via 50-way ribbon cables to the preamplifiers. After the preamplification, the signals are further amplified, shaped and discriminated. This produces digital signals which are then fed to the data acquisition system. Since for each event the number of hits (typically of the order of 200) is much less than the total number of channels, it is more efficient to transmit addresses for the channels that *have* fired, instead of transmitting the hit/no hit information for *every* channel. Following this 'zero suppression' the information is fed to the data-taking computer so that it may be recorded.

The two stage amplification is necessary due to the requirement of low noise for the system; input capacitance has to be minimised, therefore the first amplification stage has to be as close to the detector as possible, keeping wire length to a minimum. These preamplifiers must therefore have a high packing density due to the limited available space, low power consumption to avoid problems of heat dissipation, and fast shaping time, less than 50ns, due to the high rate of electromagnetic background processes associated with the photon beam. These criteria were met by the MSD2 preamplifiers, designed at CERN. They are of hybrid construction mounted on a 1 inch $\times$ 1 inch ceramic tile that carries four channels. They have a current gain of about 200, rise time of 10ns and shaping time of about 15ns. They are mounted on the detector chariot and reside inside the AEG magnetic field. The

signal from the preamplifiers is fed to the amplifier/discriminator electronics via 3m long 50-way cables individually wrapped in aluminium foil to provide screening from electrical pick up. The amplifier/ discriminator was also designed at CERN to be used in conjunction with the MDS2 and is known as the MSD3. It is also of hybrid construction, with a voltage gain of about 40.

Due to financial constraints, it was decided not to instrument fully all 10,000 channels. Instead, an OR-gate system was devised that permitted each channel of the data acquisition system to read two of the silicon strips. The total number of data acquisition channels required was thereby reduced by a factor of two. The effect of the OR-gate was that for every real hit an 'image' hit is produced somewhere in the detector. These hits complicate the pattern recognition and care was needed to avoid the reconstruction of nonexistent ('ghost') tracks. The decision of which channels to pair was made using a Monte Carlo simulation [51]; the scheme chosen was one that minimised the reconstruction of ghost tracks. This involved the ORing of channels from different regions of orthogonal planes, i.e. the central region of a Y-plane was ORed with the outer regions of a Z-plane. The OR-gate boards were purpose built at the Saclay Institute.

The microstrip data acquisition system is the commercially available PCOS III system from LeCroy. It latches the signals and performs zero suppression as well as having a lot of extra features (like programmable thresholds and fast changing delay) that were generally unnecessary for our application. Due to the intrinsic complexity of the system, some problems were encountered during setting-up prior to data taking [46]. Once resolved, however, the system performed without problems.

9.3.2 Kinematics

Tracking at the NA14 spectrometer is performed by a series of MWPCs. Scintillator hodoscopes are used whenever fast tracking information is needed, i.e. they are used in the trigger logic, but cannot provide very accurate tracking information due to their granularity. To measure the momentum and charge of particles, two magnets are incorporated in the NA14 spectrometer.

9.3.2.1 Hodoscopes

In the NA14 spectrometer three main hodoscopes, named G1, G2H and G2V are used in the experimental trigger as described in section 4 of this chapter. Three more hodoscopes are used in the muon veto discussed in section 3.3.2 of this chapter. Some parameters of these hodoscopes are given in Table 12; both G1 and G2V have their fingers oriented vertically, whilst for G2H the elements are horizontal. In all three there is a central horizontal gap in the area covered to prevent the saturation of the detectors by electromagnetic pairs.

Table 12: Spectrometer hodoscope characteristics

Hodoscope	Distance from target [m]	No. of elements	Element orientation	Element dimensions [mm^2]	Desensitised gap [mm]	Timing Resolution [ns]
G1	1·85	80	V	54×550	± 50	10
G2H	7·5	32	H	1200×70	± 50	15
G2V	8·3	24	V	200×750	± 60	10
H1H	18	36	H	1800×180	no	160
H3H	22	32	H	2190×180	no	160
H3V	23	72	V	180×1450	no	160

9.3.2.2 Multi-wire proportional chambers

MWPCs form the backbone of the NA14 apparatus. There are 21 chambers with a total of 73 planes of wires distributed in three groups or stacks as illustrated in Figure 36.

They are generally situated in regions free of magnetic field to avoid problems of spiralling low energy electrons and track recognition in a magnetic field. An exception is the single chamber positioned within the AEG magnet to provide a track coordinate as close to the target as possible. Stack 1 contains two Orsay and six CPK chambers desensitized in the median horizontal plane; this region is covered by four small 'Fast' chambers, with fast read-out to combat the high rate of pairs. Stack 2 has five chambers which along with stack 1 provide pattern recognition before and after the large Goliath magnet, from which track momenta are measured. Finally, stack 3, with three chambers, enhances the

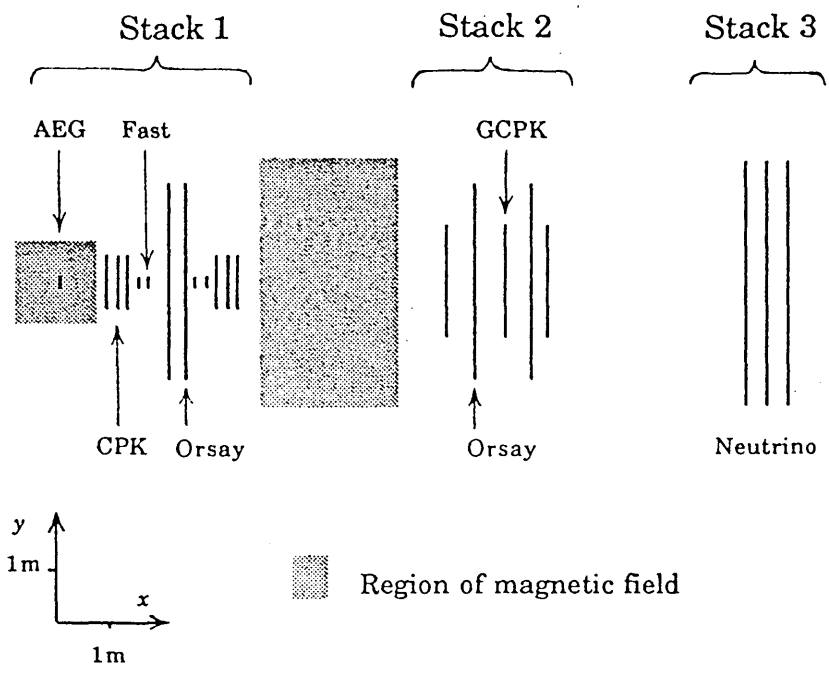


Figure 36: Wire chamber layout

precision of the momentum measurement for high momentum tracks. The main characteristics of the chambers are shown in Table 13.

9.3.2.3 Magnets

The momentum of charged particles may be measured through their deflection in a magnetic field. In a constant field B, a charged particle with charge q and velocity v feels a force

$$\vec{F} = q\vec{v}\times\vec{B}$$

The particle therefore follows a helical trajectory, which in a projection orthogonal to the field is a circle of radius R, given by:

$$R = \frac{p\cos\lambda}{0.003Bq}$$

where λ is the angle of the track relative to the plane of the projection and R is in units of cm, B in Tesla and p is GeV/c. The momentum of the particle may thus be determined from the magnitude of its deflection in the region of magnetic field and its charge from the sense of this deflection. In practice the field is not exactly uniform, but it can be precisely mapped. The trajectories are then calculated

using the Runge-Kutta method.

Table 13: Wire chamber characteristics

Chamber	No.	Distance from target [m]	Area ($y \times z$) [m^2]	De-sensitised region [mm]	Planes per chamber	Orientation* [°]	Wires per plane	Wire spacing [mm]	Gate [ns]
AEG	1	0·1	0·20 × 0·20	no	4	0 90 45 135	186 186 186 186	1	50
CPK	6	1·0 1·1 1·3 3·2 3·4 3·5	1·79 × 0·73	50 to 120	4	104 90 0 14	768 896 368 576	2	120
Fast	4	1·4 1·5 3·0 3·1	0·26 × 0·21	no	2	60 90 or 90 120	256 256 256 256	1	50
Orsay	4	1·7 2·1 8·1 8·8	3·60 × 1·50	5	4	90 60 120 90	1792 1536 1536 1792	2	120
GCPK	3	7·8 8·5 9·0	3·11 × 2·15	no	4	104 14 90 0	1024 896 1024 704	3	120
Neutrino	3	14·0 14·3 14·6	4·33 × 3·75	no	3	60 120 0	1232 1232 1232	3	120

* The angle is measured relative to the horizontal (y) axis, and the order follows the beam direction.

NA14 uses two magnets both providing a dipole field with the axis vertical. The first, named the AEG, surrounds the target. It is required to sweep the abundant low energy electromagnetic pairs out of the acceptance of the detectors. It also provides momentum measurement for tracks with energy less than 5GeV that are outside the acceptance of the second magnet. The second, larger, magnet, named Goliath, provides a precise momentum measurement for energetic particles with $\Delta p/p^2 \approx 5 \times 10^{-4} GeV^{-1}c$ for tracks within its acceptance. The magnet parameters are given in Table 14.

Table 14: Magnet characteristics

	AEG	Goliath
B [T]	1·29	1·3
Distance from target [m]	0·0	5·5
Aperture ($y \times z$) [cm]	$\pm 65 \times \pm 25$	$\pm 120 \times \pm 60$
Acceptance ($y \times z$) [mrad]	$\pm 650 \times \pm 250$	$\pm 160 \times \pm 80$
$\int B \, d\, l$ [Tm]	1·24	3·1

9.3.3 Particle identification

From the multitude of particle types that are produced in a typical interaction, it is only those that are relatively long-lived that reach the detectors, and may be identified. The short-lived particles decay, and can only be identified from their decay products.

Of the long lived particles produced in a hadronic interaction, the most abundant are pions. Also present in significant numbers are kaons, protons, electrons, photons and other neutral hadrons. A number of different detectors are used to distinguish these particle types: the charged hadrons may be identified using Cerenkov detectors, the electrons and photons are identified using electromagnetic calorimeters and for the muons there is a muon filter.

9.3.3.1 Cerenkov counters

A Cerenkov detector is sensitive to the velocity of a particle. If the momentum of the particle is known, then its mass may be determined, thus providing identification. It operates on the principle that when a charged particle travels through a dielectric medium at a speed greater than the speed of light in that medium, some of the light emitted by the excited atoms of the material appears as a coherent wavefront at a fixed angle to the particle trajectory. The minimum velocity for the production of Cerenkov light for a given particle mass translates into a momentum threshold. This light in a Cerenkov counter is collected using mirrors and detected using photomultipliers. Mirror size is determined by light cone dimensions on the one hand, and on the other, by the requirement that tracks do not share mirrors. This leads to smaller size mirrors towards the central region of the detector where

track density is greatest.

Table 15: Cerenkov detector characteristics

Cerenkov	Gas	Refractive index	Momentum threshold [GeV/c]				
			e	μ	π	K	p
INDRA	Dry air	1·000297	0·021	4·34	5·73	20·27	38·52
Odysseus	Freon-12	1·001080	0·011	2·28	3·01	10·63	20·21

NA14 incorporates two Cerenkov detectors, named Indra and Odysseus. They operate in threshold mode, with different radiative media to maximize the momentum range of the particle identification. Their most important features are listed in Table 15. Note that Odysseus lies in the 1.3T field of the Goliath magnet. Alternative sites would have compromised the overall acceptance but it is a source of some complications since:

- The photomultipliers will not function in the magnetic field of Goliath. Therefore, extensive screening is required and, in some cases, some of the photomultipliers had to be moved further out to regions of lower magnetic field. This resulted in a decrease in the amount of light incident on each photomultiplier.
- Charged particle tracks are curved rather than straight, making the reconstruction and analysis programs more complex and computer-time consuming.

9.3.3.2 Calorimeters

A particle entering a dense medium will interact to produce secondary particles, which themselves interact producing further particles, generating a *shower*. If the material is thick enough, most of the energy of the original particle will be absorbed and appear as ionization or excitation in the medium. A calorimeter uses this principle to measure the energy of the original particle. This is particularly useful for neutral particles whose energy cannot be determined otherwise.

The apparatus of NA14 includes three electromagnetic calorimeters covering between them the angular acceptance: OLGA covers the forward region of 0−80mrad in the laboratory frame, ILSA covers the 80−150mrad region, corresponding to about 90^0 in the centre-of-mass frame and the Crown covers the centre-of-mass backward region (up to 275mrad in the laboratory frame). All three have the following structure:

1. Charged veto: discriminates between charged particles and neutrals, permitting the rejection of charged particles for the analysis of photons;
2. Preconverter: initiates the showering of photons and electrons, thus allowing the measurement of their position by a position detector. It also assists in the rejection of neutral hadrons, since hadronic showers develop more slowly than electromagnetic;
3. Position detector: determines the position of the shower in the calorimeter. It consists of two scintillator hodoscopes with orthogonal planes of thin fingers. The problem of ambiguities of such a system is helped by the segmentation of the calorimeter proper, and by comparison of pulse heights in the two projections.
4. Calorimeter proper; The Crown and OLGA use a homogeneous active component, lead glass, and detect the Cerenkov light caused by relativistic particles in the shower. ILSA has distinct active and passive components in the form of alternate planes of scintillator and lead.

Details of the calorimeters are given in Table 16.

Table 16: Electromagnetic calorimeter characteristics

	OLGA	ILSA	CROWN
Distance from target [m]	15·5	13·5	2·35
Angular acceptance [mrad]	0-80	80-150	80-275 V 150-275 H
Construction	Pb-glass	Pb/scint.	Pb-glass
Cell size [cm^2]	14×14	25×25*	9×9
Number of radiation lengths	18·5	18	14·8
Minimum energy measured [GeV]	1·5	1·5	1
CHARGE VETO	Scintillator	Scintillator	Wire chamber
PRECONVERTER	Active	Active	Passive
Construction	Pb glass	Pb-scint.	Pb
Number of radiation lengths	3	4·5	5
POSITION DETECTOR	Scintillator	Scintillator	Scintillator
Width of fingers [cm]	1·5	1·5	0·8

* Cell size in each projection 25 × 100 cm^2.

9.3.3.3 Muon filter

Muon identification relies on their ability to penetrate matter with little loss of energy. For this purpose, an iron wall 5.4m thick, 20m downstream of the target, is utilized to stop all particles but muons. These are detected using three scintillator hodoscopes, H1H, H3H and H3V whose parameters are given in Table 12.

9.4 Trigger

The NA14 trigger is designed to select hadronic interactions of high energy beam photons. It is not a very selective trigger (sometimes referred to as *'minimum bias'*). It comprises two *levels*: the *pre-trigger*, and the *final trigger*. This two-level structure permits a quick identification of potentially interesting events in the pretrigger without having to go through the (time consuming) full trigger calculation for every event. Thus, dead time is minimised. The pretrigger 'locks' the information from every detector until the decision for the full trigger is available. If the decision is positive, data from all detectors are read by the data acquisition computer; otherwise, a fast clear is sent to all detectors which are then ready for the next event. For a schematic of the trigger requirement, refer to Figure 37.

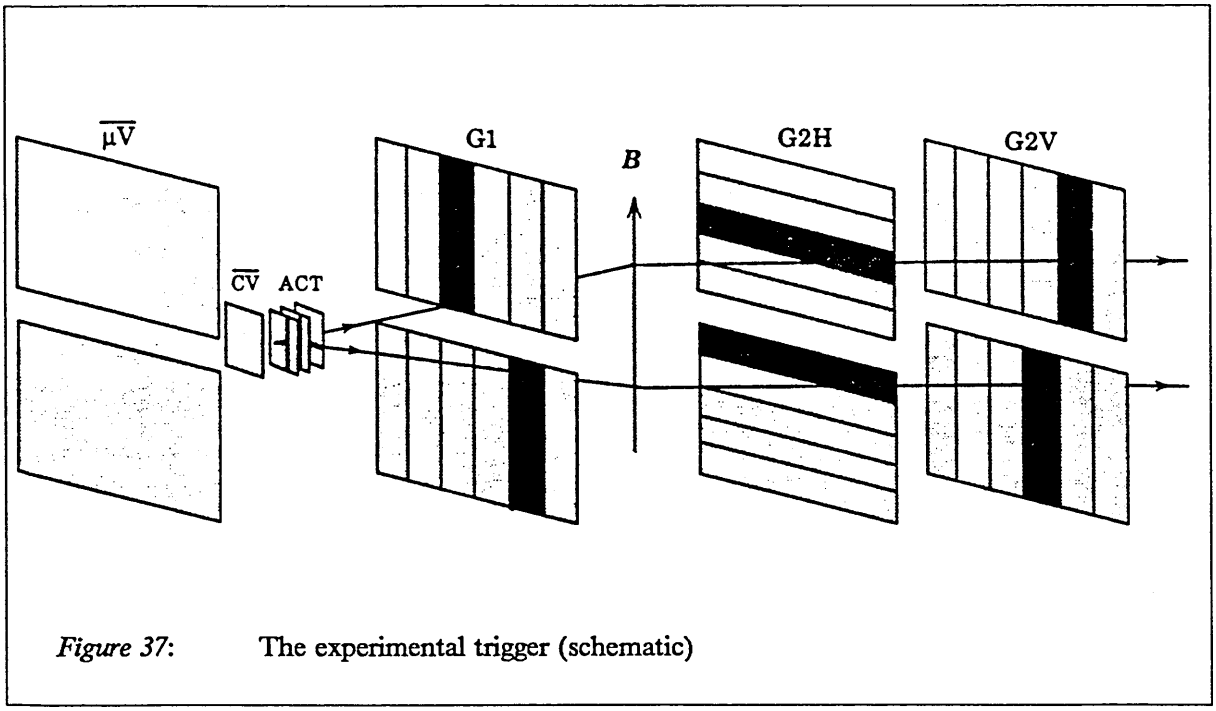


Figure 37: The experimental trigger (schematic)

9.4.1 Pretrigger

The pretrigger is the coincidence of

$$BEAM \wedge ACT \wedge HODO \wedge \overline{PRE}$$

where $\wedge$ implies logical AND. The components have the following meaning:

BEAM is a signal that a photon of more than about 40GeV has reached the target. The BTC counters are used for this task as well as CV, the charge veto counter that resides in front of the active target and is used to reject charged particles interacting in the experimental target.

ACT is the active target condition of more than 2.5minI deposited in at least one of the central 15 vertical strips of each of planes 28 and 29, and at least one of the central 13 horizontal strips of plane 31 of the target. This condition is applied to trigger on events where a hadronic interaction has occurred, as opposed to electromagnetic pair production. The number of active target strips in the trigger is chosen to match the trigger rate to the data acquisition system dead time.

HODO is a combination of signals from the scintillator hodoscopes of the apparatus, designed to increase the hadronic content of the trigger. In more detail:

$$HODO = (\overline{\mu V} \wedge G1 \wedge G2V \wedge G2H)_{top} \wedge G2H_{bot}) \vee ((\overline{\mu V} \wedge G1 \wedge G2V \wedge G2H)_{bot} \wedge G2H_{top})$$

where $\vee$ implies logical OR. G1, G2V and G2H are the hodoscopes described in section 3.2.1 and μV is the muon veto. The *top* and *bot* subscripts refer to the halves of the detectors above and below the central gap.

PRE indicates that a previous pretrigger has already occurred but has not been dealt with yet.

9.4.2 Final trigger

The final trigger differs only slightly from the pretrigger. It is the following:

$$TRIG = PT \wedge DARM$$

where PT is the pretrigger and DARM ('Double arm') a more restrictive requirement for the hodoscopes than HODO. It is designed to select events with a track in both the upper and lower halves of the apparatus. This enhances the hadronic content of the trigger, since electromagnetic pairs are swept in a narrow horizontal band. In more detail, the requirement is:

$$DARM = (\overline{\mu V} \wedge G1 \wedge G2V \wedge G2H)_{top} \wedge (\overline{\mu V} \wedge G1 \wedge G2V \wedge G2H)_{bot}$$

In the 2 second burst of the experimental beam, there were typically 220 pretriggers and about 100 final triggers, out of which about 65 events were written to tape. The dead time of the data acqui-

sition system thus represented a significant source of inefficiency during data taking. From a visual inspection of event displays, the trigger was found to be about 95% hadronic. Monte Carlo studies indicate that it enriches the charm content of events by a factor of 2, compared with all hadronic interactions (see chapter 11).

Chapter 10

DATA PROCESSING

10.1 Overview

NA14/2 registered 17 million events during the main data taking periods of June 1985 and summer/autumn 1986. These 'rawdata' events are written to tape and consist of *digitized* information: wire addresses, pulse heights and so on. To analyse these events it is necessary to reconstruct the kinematic variables of the tracks that gave rise to these digitisings, identify the particles involved and finally reconstruct the event topology. It is only then that physics analysis may proceed. The above task of event reconstruction is performed by the reconstruction program, the process of passing rawdata through it usually referred to as *production.*

The reconstruction program in NA14 is called TRIDENT. It takes on average about 8 seconds of CPU time per event (IBM 168-equivalent units). Processing all data through it would therefore require about 3.7 years of CPU time (168 units), a rather substantial amount of computing power. A way to circumvent this problem, is to use what is known as a filtering scheme: A filtering program is basically a 'stripped down' version of the reconstruction program, less accurate but having the advantage of being much faster. NA14/2 in its original proposal [52] suggested the usage of a variety of filtering schemes as a basis for its physics analysis. These schemes were independent from one another and not completely orthogonal to each other, each aiming at different or complementary physics topics. This arrangement proved invaluable during the early stages of analysis program development – the computer time efficiency of those filters provided charm enriched samples quickly and the overlap of the filters permitted analysis program comparisons to be made. However, the use of a filter meant that typically about one in two charm events registered on tape was thrown away. This, together with

the need for a study to estimate biases of the various filters, suggested the direct processing of part of the data through the full reconstruction program.

The processing of a substantial amount of the total statistics though the full reconstruction program was made possible with the 3081/e processor pilot project farm at CERN. The use of this facility made NA14/2 one of the first experiments to gain a high level of expertise on emulator farms. The major part of this chapter will therefore be devoted to discussing the NA14/2 'emulator project'.

10.2 The production program

TRIDENT is the program responsible for charged track reconstruction in NA14, using the MWPCs and the microstrips. It has its origins in, and bears the same name as, the Omega spectrometer reconstruction program [53].

Track reconstruction in TRIDENT starts from the MWPC information and then moves to the microstrips, trying to associate hits with tracks already found. The hits in the wire chambers are associated to form space points using groups of planes that are adjacent along the beam direction but have wires at different orientations, such as a (Y,Z,U,V) arrangement. The space points are then associated to form 'roads' in the spectrometer. In the field-free regions of the apparatus the tracks are straight lines and fits are done in the various stacks of chambers. The association of the straight-line segments is done by extrapolation through the magnetic fields (for which there are detailed maps) using the Runge-Kutta method. Microstrip hits are then matched with tracks from the rest of the spectrometer. About 90% of the well measured tracks (found both before and after the Goliath magnet) are matched with the microstrips. The momentum resolution for microstrip matched tracks ranges from $\Delta p/p \approx 0.002p\ GeV^{-1}c$ for tracks traversing only the AEG magnet to $\Delta p/p \approx 4\times10^{-4}p\ GeV^{-1}c$ for tracks traversing both the AEG and Goliath magnets. For tracks seen by the spectrometer but not matched in the microstrips the above values are $\Delta p/p \approx 0.006p\ GeV^{-1}c$ and $\Delta p/p \approx 5\times10^{-4}p\ GeV^{-1}c$ respectively.

10.3 Preproduction and filtering schemes

A preproduction program, TYPHON, was especially developed for NA14. The raw data were processed through it immediately following acquisition (on the 1986 runs) and served as a way of monitoring the apparatus performance. It uses a subset (45 out of 73) of the MWPC planes and performs the propagation of the tracks through the magnetic fields using a thin-lens type formula. It has a much lower momentum resolution than TRIDENT, but is about 20 times faster (typical momentum resolution for TYPHON is $\Delta p/p \approx 0.05p\ GeV^{-1}c$).

Four major filtering schemes were applied to part or the whole of the statistics. A brief account for each one is given below:

Microstrip Filter (I): [47] A fast pattern recognition algorithm using microstrip hit and active target information only (no MWPC information is used in this filter). Events are selected if consistent with a secondary vertex hypothesis. This filter selected 12% of triggers. However, unlike the rest of the filtering schemes this filter was only applied to a small part of the total statistics, about 1.2 million events. It was used at an early stage in the experiment, before TYPHON data were available.

Microstrip Filter (II): [54] Similar in conception with filter (I), but using some MWPC information (made available from TYPHON processing) as well. It selects events if found to be inconsistent with single vertex topology. This filter retains 16% of the triggers with good charm efficiency according to preliminary Monte Carlo studies. The bulk of the NA14 D^0 and D^+ statistics come from this filter.

Clean Active Target Filter: [40] This filter utilizes only active target information to filter off coherent (clean active target) events which exhibit a jump in multiplicity. It retains 7% of all triggers. Of the selected events, about half of them are hadronic and are mainly low multiplicity events. This sample was used for the study of coherent charm event production; also, the cleanliness of the data selected with this filter allows the reconstruction of decay modes with big combinatorial possibilities (such as $D^0 \rightarrow K\pi\pi\pi$) above a low background [55].

Double Kaon Filter: [56] This selects events with two or more kaons or protons; the momentum information comes from TYPHON and particle identification from the INDRA Cerenkov counter.

This filter retains 4.5% of events. It was mainly used to study charm decays with two kaons in the final state ($D_s \rightarrow \phi\pi$ for instance) and the charmed baryon decay $\Lambda_c \rightarrow pK\pi$.

10.4 Direct rawdata processing

The two-stage (filtering-production) approach provides a manageable sample of charm enriched events at a reasonable timescale. However, although invaluable during the early stages of program development and computer CPU time efficient, it is not without its disadvantages. These are basically the following:

- Part of the statistics is thrown away due to the imperfect charm retaining efficiency of the filters.
- These filters introduce biases which are not always easy to estimate and which have to be well understood and accounted for in any lifetime or cross section calculation.

By fully processing part of the rawdata events without the intervention of any filters, NA14 could obtain a large sample of fully reconstructed unbiased events, and hence be able to calculate the various filter efficiencies and biases. Since no filter efficiency is involved, one has the extra bonus of reconstructing more charm particles from a given rawdata sample. The factor that one gains, as we shall discuss later on, is quite significant. This need for raw data production for NA14 was satisfied by the 3081/E emulator farm. Basically the main advantage of this approach is the very high throughput compared with other more 'conventional' computing schemes together with an excellent performance to price ratio.

10.5 The 3081/E emulator farm

10.5.1 Emulator farms

An *emulator* is a unit with some processing power, a means of communicating with the outside world (an input/output [I/O] port) and some memory, which *emulates,* i.e. for the same program gives the same results as, a mainframe computer. It is usually a specialized, 'stripped down' version of the computer it emulates, not being able to perform all the tasks of the emulated machine, but at only a

fraction of its cost. The emulator farm used by NA14/2 is based on the 3081/E emulator [57], a powerful processor which emulates an IBM system 370 CPU. It is of modular architecture, consisting of 5 execution units plus memory and interface cards, all on separate boards. The execution unit cards include a control and register card, an integer arithmetic card and floating point arithmetic add/subtract, multiply and divide cards. It has separated program and data memory configurable in 0.5 Mbyte units giving a maximum of 7 Mbytes of memory per processor. It emulates a VM/370 system performing all FORTRAN and ASSEMBLER instructions – with the exception of I/O operations – giving bit-by-bit identical results. Its performance is about 1 IBM 370/168 unit, or about 4mips (million instructions per second).

It is a reduced instruction set machine, running its own 3081/E microcode. The IBM source code has to, therefore, pass a second compilation stage known as the *translation* stage before the program is ready to run on the 3081/E. Since the 3081/E itself cannot perform I/O operations, some external intelligence is needed to allow the emulator to communicate with the outside world. This is known as the host computer, and it is usually a small IBM mainframe computer running the VM 370 system. The reason that a VM 370 system is usually chosen is because it makes the exchange of library routines, etc, as well as the bit-by-bit verification much easier, since the 3081/E uses VM/370 routines and the results of the two systems are identical.

Since an emulator is an independent unit, many of those units can be connected together to the same host machine, giving rise to an emulator 'farm'. The host computer in this case is in charge of all I/O operations together with the supervision of individual processors.

10.5.2 Emulator farms in high energy physics

High Energy physics has a growing need for computing power. The major proportion of this power is needed for running Monte Carlo and reconstruction programs on large data samples. The rest of the analysis requires relatively less computer resources. This need for computing power for reconstruction programs is one of the design aims of the 3081/E farm.

A typical high energy physics reconstruction program has some very distinct features that make it different from other 'typical' computing applications:

- The program loops many times over the same sequence of routines, processing events which have no interconnection between them. (an event does not need any variables derived from a different event to be reconstructed)
- It usually has three very distinct parts:

1. the reading of a raw data event from mass storage (usually tape).
2. The processing of the event.
3. The writing of the reconstructed event out to tape.

From those three parts, the first and last consist of mainly I/O operations with minimum computation, whereas the second part is mainly CPU intensive, with little in terms of I/O operations. Therefore, such a program can easily be split in two parts: a CPU intensive part, and its I/O counterpart. The I/O intensive part can run on a normal machine such as the host computer of an emulator farm, whereas the CPU intensive part can run on a 'stripped down' version of the same machine, i.e. an emulator. Moreover, since the computation of each event is independent of the calculation of any other event, different events can run simultaneously on different emulators. In this way the emulator farm is running as a parallel machine, the parallelism being at a very 'coarse' level, namely at the event level.

10.5.3 The 3081/E farm at CERN

The farm where the NA14 production took place is the 3081/E pilot project farm at CERN. Throughout most of the production it utilized 5 to 12 emulators all connected to a 4361 IBM computer, serving as a host, through a VME to IBM Channel Interface (VICI) capable of transfer rates of up to 3Mb per second. Two 6250 bpi tape drives were connected to the 4361 and, some time after the NA14 production started, two 3480 cartridge drives were installed. NA14 rawdata were stored on 6250 bpi tapes but the output of the reconstruction program was written on 38K cartridges as soon as the 3480 drives were in operation. The reasons that cartridges were preferred to normal 6250 bpi tapes were on the one hand their better storage capacity to price ratio and on the other, the much quicker

read/write time of the 3480s, (although for a program like TRIDENT, with the tape writing time only representing a small fraction (about 15%) of the host time, the overall increase in speed was marginal).

10.6 Running on the 3081/E farm

10.6.1 Program preparation for running on the 3081/E farm

To make a program run on a 3081/E farm requires a series of modifications. Since the exercise for TRIDENT was in many ways original some of the details of the modification stage, although of a technical nature, will be discussed in this section.

To modify a reconstruction program like TRIDENT for emulator production the following steps are necessary:

- Ensure that the program compiles on standard FORTRAN 77. Note that certain errors that give warning messages of the FORTRAN compiler will not be accepted later on at the translation stage; for instance, unreachable code only gets a warning message from the standard FORTRAN compiler, but gives a fatal error by the translator. Also note that one has to use the IL(DIM) option of the compiler if the program uses variable length arrays.
- Split the program in two parts: a CPU intensive part that will run on the emulators, and an I/O intensive part that will run on the host. Translate the emulator part (ie pass the program through a second compilation stage to convert it to 3081/E microcode)
- Decide which common blocks are to be transferred to the emulator at the initialization stage (these are mainly the database constants) and which common blocks need to be transferred to and from the emulator for every event (mainly the blank common containing the event bank structure)
- Test that there is bit-by-bit agreement between emulator and host results. If this is not achieved, it should be due to one or more of the following reasons:

1. Not all the right common blocks are transmitted to/from the emulators; this can easily be checked by transmitting all common blocks for every event and see if the problem disappears.

2. There is an uninitialised variable in the reconstruction program. Then, the value of this variable at the beginning of an event is taken as the value it possessed at the end of the previous event. Therefore, the event reconstruction results would depend on the order that the events appear, and, since on the emulators one does not necessarily get the events in the same order as on the host, discrepancies can arise. To correct, follow the standard bug fixing procedure on the host.

None of these steps require any special skills or knowledge – they can easily be performed by anyone with reasonable knowledge of FORTRAN and are quite straightforward for any reasonably well written and documented reconstruction program.

In the case of TRIDENT there were also a few points where special care had to be taken. These are discussed in more detail in the next section.

10.6.2 Main changes to TRIDENT

TRIDENT is a typical HEP reconstruction program using ZBOOK [58] for dynamic memory management and the ZEP routines (which is part of the ZBOOK package) for reading and writing data. The ZEP routines use the EPIO package which in turn uses IOPACK routines for I/O operations. When TRIDENT was first introduced, it was not foreseen that it was going to be run in an emulator farm configuration, therefore no specific effort was made to have the I/O intensive part and the CPU intensive part well separated. More specifically, the routine that reads an event in also does the CPU-intensive ZBOOK bank creation. Therefore, for maximum efficiency, the program had to be split just after the call to ZEPIN, the ZBOOK subroutine that reads an event, in a subroutine nesting three levels inside the main program. Therefore, three routines had to be split in two parts and were put respectively in the host and the emulator part of the program. All local variables were transmitted from the host part to the emulator part in COMMON blocks. This decreased the host overhead time, the time spent on the host per event, from 22% to 7% of the average time to process a complete event. This decrease in host overhead time is very important for a program like TRIDENT (where the host time represents a sizable proportion of the total event time) if a farm with a large number of emulators is to be used efficiently.

There were two more necessary changes to TRIDENT for emulator usage. The first comes from the fact that the ZBOOK structure resides on different physical memory locations on the host and the emulator(s). Therefore, the pointer to the master ZBOOK bank had to be restored to its correct value every time a host-emulator data transfer occurred. This single variable is the only one that needs to be changed explicitly; all other pointers in ZBOOK are calculated relative to it.

The other modification has to do with the EPIO package – the package used for reading events from tape and writing the reconstructed event out. The input and output EPIO buffers are part of the ZBOOK structure but, unlike the rest of the banks which are transferred to and from the emulators, these must always remain on the host, since event input and output is done serially by the host. Therefore, care must be taken not to overwrite the EPIO buffers after reading an event back from an emulator – in the case of TRIDENT this was done by copying those buffers to another array in the host and restoring them to their correct value after the transfer back of each event from an emulator.

10.6.3 Software debugging

The requirement for bit-by-bit identical results, apart from ensuring consistency of emulator and non-emulator program results, is also a powerful way for discovering possible bugs in the code of the source program. The reason why a problem undetected in the normal version becomes apparent when comparing emulator to non-emulator runs is that in the emulator the memory allocation is different both in the 'spatial' sense (due to the fact that on the emulator the program and data memories are physically different units whereas on the host data and program are interleaved in the same physical memory space) and in the 'temporal' sense (since event processing order in a multiprocessor farm is generally not the same as on the host). In the case of TRIDENT, an uninitialised variable problem was found when comparing emulator and non-emulator results; a variable in some cases was not initialized properly giving small differences when bit-by-bit comparison was performed. The reason that this bug was not discovered in normal TRIDENT running is because the final value of that variable depended on an iterative method and in most cases convergence was obtained even with the wrong initial value. In any case, this problem only affected a small percentage of tracks.

10.6.4 Hardware debugging

Testing a piece of hardware like a 3081/E emulator is not an easy task and there are no magic recipes for a perfect hardware testing program. However, the deterministic nature of these machines allows us to use the comparison of an emulator run to some 'reference' run to check the state of a subset of the hardware of the processors involved. For that reason, a comparison program was written to be used as a hardware test version of TRIDENT. It runs on the whole or part of a sample of 80 NA14 rawdata events stored on disk to minimize access time. It then compares the results of an emulator run with those of a run on the host computer, and reports on any differences. Thus, TRIDENT has joined the other programs already existing for this purpose, bridging the gap between long computation programs and short, I/O bound tests already existing. The TRIDENT hardware testing program has since helped in isolating a problem due to a malfunctioning memory chip on one of the processors; The problem did not manifest itself when running other test programs, since it so happened that they were never addressing the problematic memory location.

10.7 Performance

10.7.1 Timing tests

TRIDENT was run in a variety of configurations out of which we will mention only the most important ones: Two versions of TRIDENT were run on the emulator farm: V1.10 and V1.11, an improved version with more accurate track reconstruction which however took about 50% more CPU time per event. In addition, V1.10 was run in two different code-splitting configurations: a simple one, (A), and the one with optimized splitting of the code as already described (B). Host overhead, defined in this context as the difference of the total time per event minus the emulator execution time over the event total time, varied from a non negligible 22% for V1.10/A to an impressive 5% for V1.11. The results of timing runs can be seen in Table 17. For those runs, 6250bpi tapes were used as input and output storage media.

Table 17: Breakdown of average event

TRIDENT version	V1.10/A	V1.10/B	V1.11
Time (s)			
Host input	0.87	0.12	0.12
Load Emulator	0.10	0.10	0.10
Emulator execution	5.02	5.41	7.46
Read back emulator	0.10	0.10	0.10
Host output	0.07	0.06	0.06
Total per event (s)	6.15	5.79	7.84
Host overhead (%)	22	7	5
Program efficiency (%)	78	93	95

10.7.2 Program efficiency

To quantify a program's suitability to run on an emulator farm, we can define the program efficiency as the ratio of the emulator execution time over the total time per event. The less I/O and host execution time compared with the mean event time, the more efficient the program is and therefore the more suitable for multiprocessor emulator farm running[19]. This program efficiency, as defined above, determines the maximum *effective* number of processors[20] of a program/farm configuration. This corresponds to the situation where the system is totally saturated by host computations and I/O operations and is equal to the reciprocal of the total host overhead. Adding processors to a farm with $[N_{eff}]_{max}$ processors does not increase throughput, it merely increases the average time of a processor being in an idle state waiting to be served by the host. In the case of TRIDENT, program efficiency as can be seen in Table 17 was of the order of 95% for version V1.11, giving a value of $[N_{eff}]_{max} = 20$. Efficiency of this order made the utilization of a 12 processor farm sensible since, even with this number of processors the host was busy (performing computations or I/O operations) only about 50% of

---

[19] From the definitions of host overhead and progam efficiency it clearly follows that: (program efficiency + host overhead) = 1

[20] The effective number of processors, N_{eff}, is defined as the total time taken by a 1 processor configuration to run a certain number of events, over the total time taken by an N processor configuration

the time.

10.7.2.1 Comparison with the computing centre 3090/200

One of the raw data tapes [BN3220] that was run on a 12-processor farm was also run on the computing centre's IBM 3090/200. Only part of the tape was processed due to the 720 minute upper limit for a batch job at CERNVM – sufficient for processing only a fraction of a 20K event NA14 tape. On the 3090 the CPU time per event was 0.956 seconds, implying 4.79 seconds of normalized (168 equivalent) CPU time. The real time per event was 5.3 seconds (the test was performed during the early hours of the morning when computer load is smallest – the figures suggest that five other jobs of similar priority were being processed at the same time) The same tape on the farm took an average of 0.530 seconds (real time) per event[21]. This implies a ratio of 3090/200 CPU time over 3081/E farm real time of 1.8 to 1 in this configuration and a ratio of 3090/200 'typical best' real time over 3081/E farm real time of 10 to 1.

10.7.3 Throughput

NA14 rawdata tapes contain about 20K events each. Since TRIDENT creates a lot of extra output, one input rawdata tape is split in two TRIDENT output tapes. In total, about 350 rawdata tapes were processed, corresponding to an output of about 400 tapes and 300 cartridges. The data processed cover the whole of the September 1986 run and a small part of the July 1986 run. Table 18 shows detailed information on input and output tape numbers for the emulator TRIDENT production. When 12 processors were used, a typical 21000 event job took a record low time of about 230 minutes to be completed (real time per event was about .65sec). At that rate, the farm was capable of processing 0.6M events per week (assuming that the farm would be used for production during evenings and weekends, while being available for development work during the day).

---

[21] This tape had a somewhat lower mean time per event than a typical NA14 rawdata tape

Table 18: NA14 raw data tapes processed through the 3081/E farm

version	period	raw data tapes	output tapes	# input tapes
V1.10	Jul 86	BN2738 - BC2753	BN0001 - BN0022	11
V1.10	Sep 86	BN3035 - BC3102	BN0023 - BN0158	68
V1.11	Sep 86	BN3103 - BN3227	BN0159 - BN0400	121
V1.11	Sep 86	BN3231 - BN3370	BC0001 - BC0272	136
V1.11	Jul 86	BN3019 - BN3034	BC0273 - BC0302	15
total number of tapes processed				351

10.8 Tests and comparisons using the emulator-reconstructed events

Two charm decay channels have been analysed using the new TRIDENT (V1.11) emulator data: The $D^0 \rightarrow K\pi$ and $D^+ \rightarrow K\pi\pi$ decays. The total number of events analyzed is 4.3M. Special runs were excluded from this analysis (corresponding to about 300K events). The fraction of TRIDENT output tapes not analysed due to tape errors or due to a problem during the original TRIDENT processing was about 10%. The motivation for this work as we have already mentioned was twofold: Firstly to evaluate the performance and efficiency of the filtering schemes (in this case the microstrip filter II) used in the 'traditional' analysis chain, and also to provide a bias-free sample for cross section and life-time measurements. In this section we shall present the results of the filter efficiency tests as well as some indication of the bias-free nature of the emulator data. [This section inevitably uses concepts regarding the physics analysis of the data, developed in chapter 11. The reader is, therefore, kindly asked to refer to that chapter for a detailed account of the analysis and vertex reconstruction package used, as well as of the cuts and their definition.]

10.8.1 Comparison with the microstrip filter II analysis chain

The $K\pi$ spectrum from those 4.3M rawdata events at different vertex separation cuts[22] can be seen in Figure 46. If we impose a 3σ cut on the separation between the primary and secondary vertices, we get about 330 D^0 events at a signal to noise ratio (S/N) of approximately 1:2. Figure 38 shows the D^0 spectrum when a D^* cut is imposed: The mass of the D^0, combined with a pion in the event, is reconstructed and only events compatible with the decay $D^* \rightarrow D^0\pi$ are kept. This additional requirement greatly improves signal to noise ratio for a given vertex separation cut. At a 2σ cut in the distance of primary to secondary vertex there are about 80 D^0 events with small background. Figure 47 shows the $K\pi\pi$ spectrum. When requesting a minimum vertex separation cut of 4σ we get about 360 events at a S/N ratio which is again about 1:2.

Therefore, the NA14 reconstruction chain, used as described above, is capable of producing about 160 reconstructed D^0 and charged D events per million rawdata events at a reasonable signal to background ratio in the $K\pi$ and $K\pi\pi$ channels. This corresponds to about 1300 $D^0 \rightarrow K\pi$ and about 1400 $D^+ \rightarrow K\pi\pi$ reconstructed events from the full statistics of 17M triggers.

The comparison with the microstrip filter II and its complete analysis chain is not straightforward since the signal to background ratio at the same σ cut is not the same as that of the emulator data, although the same analysis program is used in both cases. This is due to the fact that the microstrip filter, being a geometrical filter, rejects events with small values of $N\sigma$. If we impose a similar signal to background ratio, for the D^0, the 2σ microstrip filter spectra correspond to the 4σ emulator spectra (S/N is about 1:1.4 in both cases). Then the number of reconstructed D^0 per million rawdata events obtained from the emulator approach is about 75, whereas for the filter analysis chain it is about 35. For the charged D, the 4σ microstrips filter II analysis spectra correspond to 6σ for the emulator data (S/N is again about 1:1.4) and the numbers of reconstructed charged D events per million rawdata events for the emulator and filter analysis are, respectively, 67 and 34. Comparing the two approaches at the same sigma cuts, for the D^0 at 2σ we get 88 and 35 for the emulator and microstrip filter II

---

[22] The vertex separation cut (also referred to as the $N\sigma$ cut) is defined as the distance between primary and secondary vertices, expressed in terms of their combined position uncertainty, σ, on this separation.

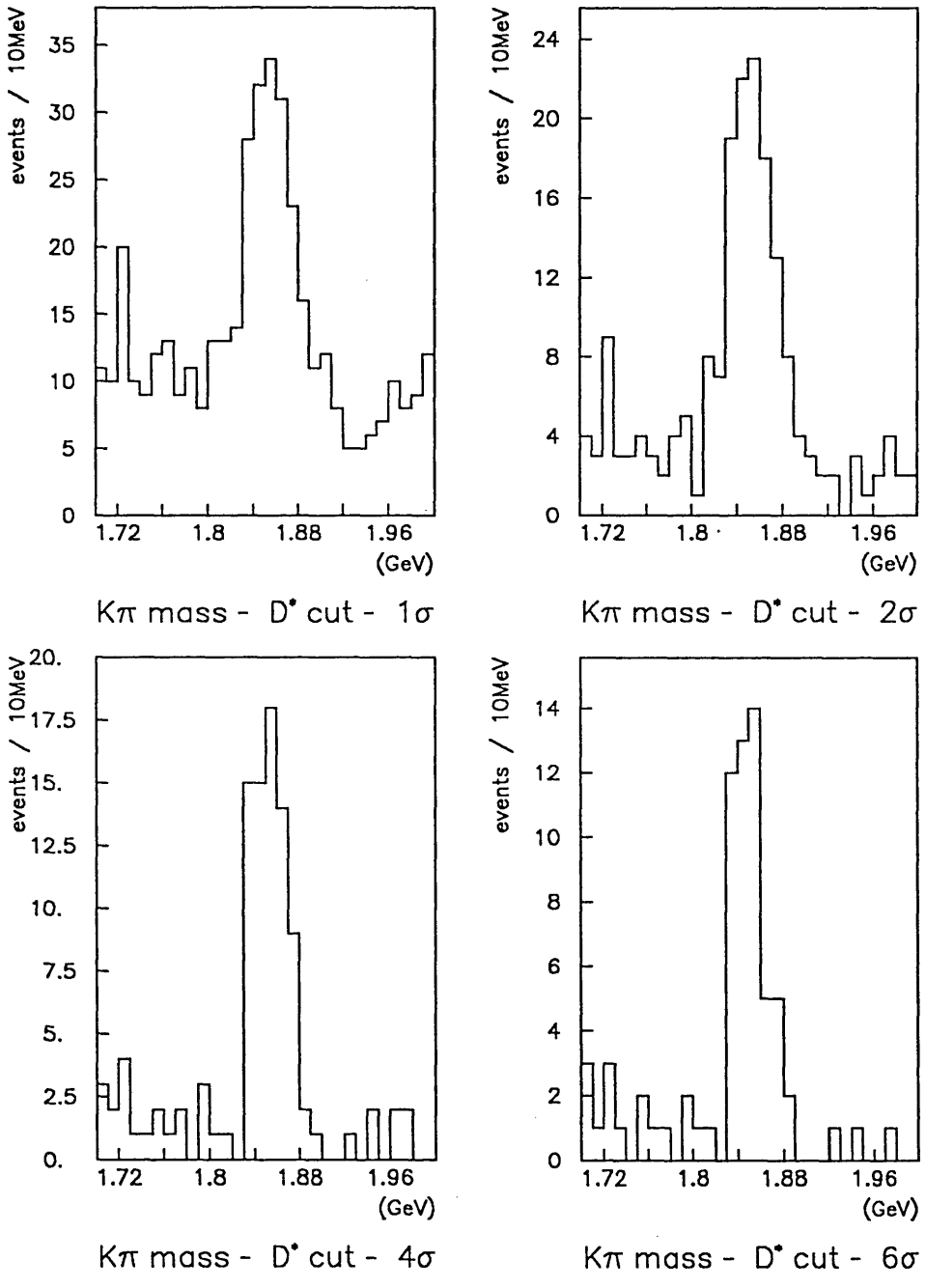


Figure 38: Kπ spectrum using a D* mass cut

approaches respectively, whereas the values for the charged D at 4σ are 84 to 34. The results are summarized in Table 19. The overall efficiency of the microstrips filter II analysis chain for the D^0 is

therefore around 40 − 50% and for the charged D somewhat lower, at around 30 − 40% depending on the definition of equivalent data sets. This is in agreement with the value predicted in the NA14/2 proposal [52]. Nevertheless, the above figures demonstrate that a gain of at least a factor of two in the number of reconstructed charm events is achieved using the emulator compared with the filtered sample, thus showing the superiority of the emulator approach in this respect.

Table 19: comparison of emulator and microstrip filter analysis yields

comparison	sample	emulator	filter	efficiency
same σ cut	D^0 at 2σ total	380	590	
	D^0 at 2σ / 10^6 ev.	88	35	40%
	D^+ at 4σ total	320	440	
	D^+ at 4σ / 10^6 ev.	84	26	31%
same S/N ratio	D^0 at 1:1.3 total	310	590	
	D^0 at 1:1.3/10^6 ev.	75	35	47%
	D^+ at 1:1.4 total	290	440	
	D^+ at 1:1.3/10^6 ev.	67	26	39%

10.8.2 Lifetime measurements

To demonstrate the bias-free nature of the emulator data, the lifetime of the charged and neutral D mesons was measured from the $K\pi$ and $K\pi\pi$ samples using no acceptance correction (apart from fiducial volume considerations). Details on the maximum likelihood method used can be found in reference [59]. We obtain:

$$\tau_{D^0} = 0.436 \pm 0.034 \pm 0.030 \qquad (N\sigma > 4)$$

$$\tau_{D^+} = 0.966 \pm 0.067 \pm 0.090 \qquad (N\sigma > 6)$$

The first error quoted is statistical and the second systematic. The lifetime obtained does not depend on the choice of σ cut. These values are in good agreement with world average values and the exercise gives an indication that no biases are introduced with the emulator processing. For the microstrip filter data, an acceptance correction has to be introduced. This acceptance correction, arising from the geometrical condition used to select events in the filtering scheme, needs to be determined from Monte

Carlo and increases the systematic uncertainties. Its effect is marginal on the measurement of the D^0 lifetime but affects the D^+ to a greater **extent** due to its longer lifetime. Therefore, even though the filtered data have higher statistics, the overall error of the lifetime of the D^+ is smaller for the emulator data [59].

Chapter 11

MEASUREMENT OF THE CHARM PHOTOPRODUCTION CROSS SECTION

11.1 Overview

In this chapter we shall discuss in some detail the physics analysis leading to a measurement of the cross-section for the photoproduction of charm using NA14 data. More specifically we shall look into the photoproduction cross-section (and its energy dependence) of neutral and charged D mesons decaying through the channels $D^0 \to K\pi$ and $D^+ \to K\pi\pi$[23].

The cross-section measurement at a given photon energy is performed by comparing the number of normal hadronic events produced in the experimental target (whose cross-section is well known) to the number of charmed particles created. At our disposal we have a certain number of events that have passed the experimental trigger requirement and which are written onto tape. Some of these events contain tagging information which can assist in estimating the energy of the photon that produced the event. These rawdata events are passed through an analysis chain to reveal a D meson signal. Since some of the D mesons in the original rawdata sample do not pass all of the analysis cuts, an *analysis efficiency correction* needs to be applied to estimate the number of D mesons recorded on tape. The analysis efficiency as a function of a certain quantity is thus defined as the ratio of the number of events passing the whole chain of reconstruction and analysis over the number of events registered on tape per interval of this quantity. Having obtained this number, we can then estimate the

---

[23] Throughout this work, $D^0 \to K\pi$ is used as an abbreviation of the decay processes $D^0 \to K^-\pi^+$ and its charge conjugate $\overline{D}^0 \to K^+\pi^-$; $D^+ \to K\pi\pi$ is used as abbreviation of the $D^+ \to K^-\pi^+\pi^+$ and $D^- \to K^+\pi^-\pi^-$ processes. Also, throughout this study the symbol for a particle is taken to include the corresponding antiparticle, unless explicitly stated otherwise.

number of D mesons produced at the experimental target if we apply a *trigger efficiency correction.* This is similarly defined as the ratio of events passing our trigger requirements over the number of events interacting at the experimental target. For calculating the photon energy from the tagging information, we shall also need a *'tagging' correction.* All these corrections are obtained using Monte Carlo simulation. The only experimental information we have at our disposal is events that have been written to tape after the tagging and trigger steps. Therefore for the tagging and triggering corrections we need to rely heavily on simulation and on a few assumptions which we submit to consistency checks, having little feedback from real data. For the analysis corrections, which are more complicated both in terms of physics processes and in terms of hardware complexity (i.e. detector acceptance, etc.) we have all the event information stored on tape to assist us with cross checks and tuning of Monte Carlo parameters.

This chapter is organised as follows: first the tagging scheme and its associated Monte Carlo will be discussed in some detail. This is a key item in this discussion and something which has been especially developed for this study. This will be followed by a short discussion on the triggering and analysis efficiencies. The next section will present the physics analysis used in obtaining a D signal followed by the cross section analysis and final results on the absolute value of the charm photoproduction cross-section and its energy dependence.

11.2 Tagging

11.2.1 Overview

To review the tagging setup, we follow the history of a typical electron that has just been created in the converter of the second stage of the NA14 beam creation process (see section 9.2.1): it passes through a series of bending magnets known as the chicane magnets and its energy can be measured by the tagging system using the position hodoscopes. It will then hit the radiator target and undergo acceleration in the electric field of the target nuclei creating bremsstrahlung photons, which may then interact in the experimental target giving rise to a hadronic interaction (an 'event'). Assuming that an

event occurs, it may trigger the NA14 apparatus and consequently be written onto tape. The electron, after traversing the radiator target, will pass through the downstream series of bending tagging magnets, its trajectory being recorded if it falls within the geometrical acceptance of the position hodoscopes. The difference of the electron energies before and after the radiator (provided that these can be calculated), which is equal to the sum of energies of all radiated photons, constitutes the *tagging answer.* Approximately 20% of all recorded events have an unambiguous tagging answer associated to them (a single track reconstructed before and after the radiator).For a further 10% of the data, more than one track has been reconstructed in the tagging position hodoscopes, and therefore no unambiguous solution exists. However, for those events, in some cases timing information from the BTC counters is successfully used to pick up the most likely solution, or, in the absence of BTC counter information, the χ^2 of the fits of the two solutions are compared, and the one with the better χ^2 probability retained [60]. Thus, if all solutions are to be taken into account, the overall tagging efficiency for the period we are considering is about 30%.

A study [61] of unambiguous and ambiguous tagging solutions has shown that both are equally reliable: the study compared the tagging answer with the total energy seen in an event, and found no differences in the distributions of ambiguous and unambiguous tagging solutions. In the analysis to follow we shall, therefore, not distinguish between the two.

11.2.2 The bremsstrahlung process

The NA14 beam is a bremsstrahlung beam; bremsstrahlung is a process in which electrons emit photons when accelerated in the presence of the electric field of a nucleus. It can be regarded as a classical, continuous process with the emitted photon energy given by the formula:

$$\rho(k)dkdT = (dk/k)(4/3-4/3y+y^2)dT, \quad (y=k/E) \qquad (1)$$

where $\rho(k)dk$ is the number of photons in the energy range dk after an electron of energy E has passed through a target of thickness dT radiation lengths [62]. To avoid the infra-red infinity implied by eq.(1) we have to introduce a minimum cut off energy, δk, for the created photon. The probability of

emitting a photon above this minimum cut off energy can thus be calculated from (1) by integrating with respect to dk from this cut off, δk, to E, the total electron energy, where the normalisation is absorbed in the definition of T:

$$\int_{\delta k}^{E} \rho(k)dkdT = (4/3ln(E/\delta k) - 4/3(E-\delta k) - 1/2(\delta k/E)^2 + 1/2)dT \qquad (2)$$

11.2.3 Reference tagging distributions and radiation target width

11.2.3.1 Upstream and downstream electron spectra

Relevant information provided by the tagging consists of the upstream (incoming) and downstream (outgoing) electron energy spectra.

Due to computation time constraints, the tagging analysis was not run on all analysed real data. To provide a reference for the tagging answer spectrum for hadronic events and the upstream and downstream electron spectra, 13 raw data tapes, evenly distributed from the period analysed, were randomly selected and processed through the tagging analysis. The distributions can be seen in Figure 39 and represent the 'reference' histograms for the tagging Monte Carlo study.

The total number of events from these 13 tapes is 120,000, representing 2.8% of all analyzed data. The statistics is high enough to introduce negligible uncertainties in the analysis to follow. The mean incoming electron energy is approximately 150GeV and the spread of the distribution approximately 20GeV. The outgoing electron energy spectrum has a 'hole' around 50GeV due to hardware problems during the main period we are considering (the September 86 run): it comes from the fact that some of the central BTC fingers were not operational during that period. (A hit in the BTC is one of the trigger conditions for our physics runs. Outgoing electrons passing through the dead BTC modules do not trigger the apparatus and, hence, even if the photon(s) produced were to interact, the event is not registered to tape.) The problem is less pronounced but still exists in other runs of this analysis (corresponding to the end of the July 86 run).

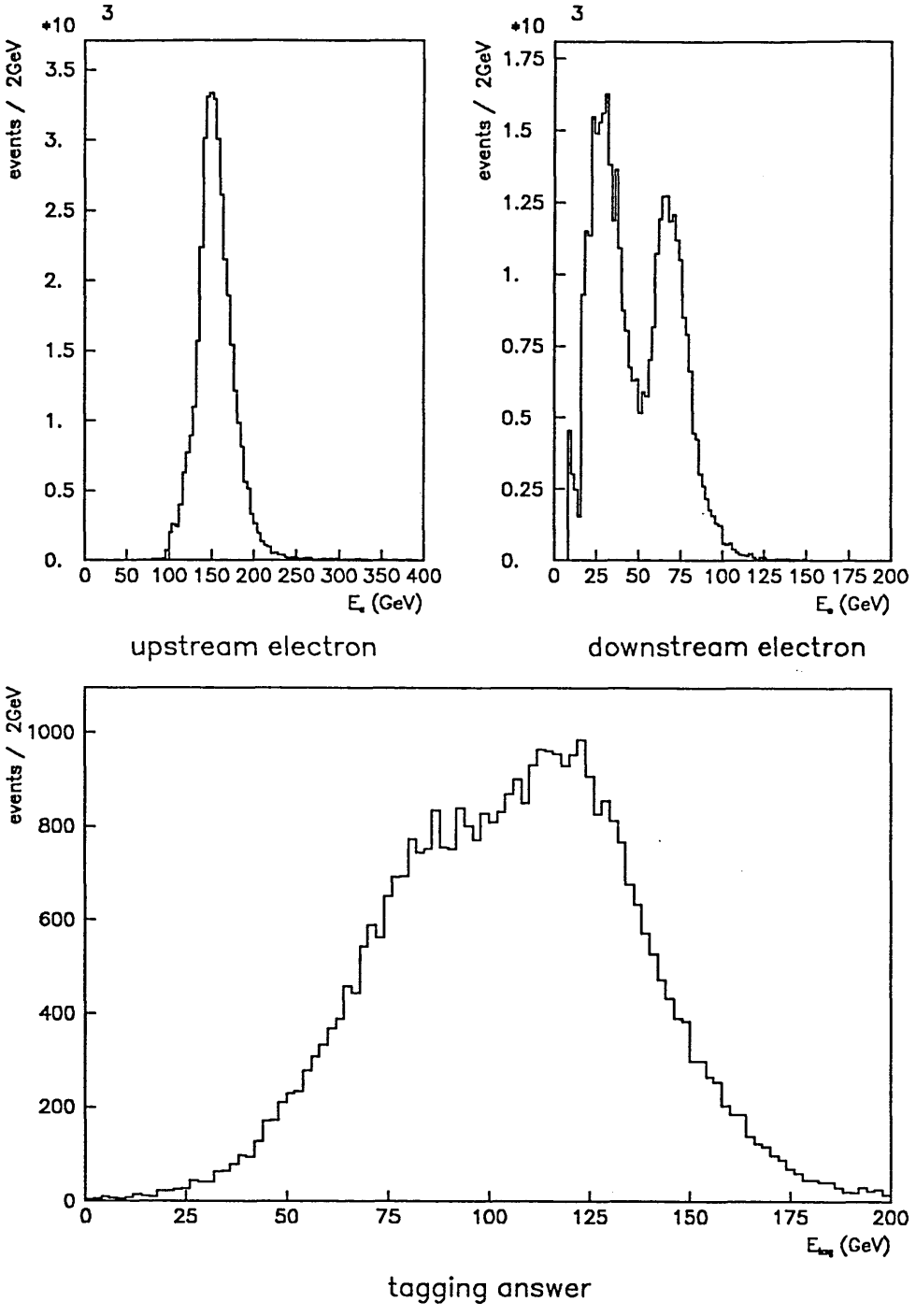


Figure 39: Tagging histograms from rawdata events

11.2.3.2 Radiation target width

The amount of material that constitutes our radiator is an important quantity that needs to be known for an accurate simulation of our tagging scheme. The radiator target consists of a lead sheet of thickness of about 0.5mm, corresponding to about 10% radiation lengths; this can be taken as a lower bound in the effective total radiator thickness. The surrounding materials also contribute an estimated 2.5% of a radiation length. For the analysis to follow, the value of 0.125 radiation lengths was used. The uncertainty on this value was taken to be 25%, and taken into account as one of the possible sources of systematic errors; its effect on the cross section measurement, as will be shown later in this chapter, is small.

11.2.4 Simulation

11.2.4.1 The need for a tagging simulation

In the NA14 tagging scheme, the tagging answer given is just the difference in the measured energy of the electron before and after the radiator target. This can naively be taken as the interacting photon energy under the implicit assumption that all the electron's energy was given to a single, high energy photon, or, at least, that the energy of the other radiated photons was a negligible fraction of that of the most energetic photon. This interpretation could be further justified by the fact that the radiator used is only a fraction of a radiation length thick and therefore the probability of more than one high energy photon being created is small. However, as we shall see later, in a substantial number of cases of radiated photons this (naive) interpretation is not correct. As an example, for a tagging answer between 100 and 120GeV, more than 20% of the time the real photon energy that gave rise to an event differed by more than 10GeV from the tagging answer. For higher tagging answers, this value is even bigger; for the 140 to 160GeV region, more than 26% of the time the tagging answer differs by 10GeV or more from the real photon energy.

Thus the tagging answer does not necessarily correspond to the energy of the photon that produced the interaction at the experimental target; we need to apply a 'tagging correction' to the tagging answer to transform it to the real photon energy. This correction cannot be applied at the individual

event level, since no additional information is available from the experimental apparatus. However, a correction can be applied globally on a given tagging answer spectrum to transform it to the interacting photon energy spectrum.

This correction is supplied by a Monte Carlo simulation of the tagging scheme; initial and final state distributions can be supplied from real data, and the physics underlying the photon beam creation, namely the bremsstrahlung process, which is well understood, is simulated. Such a simulation also provides consistency checks for the hardware and software of the tagging system.

11.2.4.2 Philosophy of the tagging simulation

The tagging simulation is a physics simulation as opposed to a complete detector simulation. It generates electrons according to a distribution, passes them through a certain amount of radiator material and records the energies of the photons created. It can then account for outgoing electron acceptance, trigger acceptance for the generated photons and, if necessary, cross section variations with energy for the interacting photons (however the hadronic cross section in the energy range we are considering is effectively constant). The exact mechanism of how all these acceptances are taken into account is discussed in the next section.

11.2.4.3 Simulation procedure

In this simulation program electrons are taken through a material in small steps; more specifically, electrons are taken through 125 slabs of material each being 10^{-3} radiation lengths thick, resulting in an overall radiator thickness of 12.5% radiation lengths. At each step, the probability of emitting a photon (above a minimum cut-off energy) is calculated according to eq. (2) and a decision on whether a photon is to be emitted or not in that slab is taken according to that probability. If a photon is emitted its energy is chosen according to eq.(1). Note that this probability is energy dependent and, therefore, needs to be recalculated as the electron loses energy while traversing the radiator. The cut off energy was chosen to be 100MeV, sufficiently low to give reliable answers given the accuracy of the tagging system (about 3GeV) and the average incoming electron energy (about 150GeV). After the electron has traversed the radiator, all the emitted photons are individually treated. The program

decides if they are to interact in the experimental target (according to a cross section variation with energy function) and, if so, if they are to be seen by the experimental trigger (according to a trigger acceptance function). Since we are simulating normal hadronic events, there is no cross section variation with energy as we have already mentioned, therefore all photons have equal probabilities of interacting. The trigger acceptance as a function of the energy of the interacting photon is taken from the main NA14 detector simulation (see section 11.3.1).

The simulation generates an incoming electron spectrum according to our reference incoming electron spectrum taken from hadronic rawdata events (see Figure 39), thus implicitly assuming that the initial spectrum of the incident electron beam has not been modified by tagging and trigger requirements. This assumption is justified since the incident electron energy spectrum is not very broad and small changes in the incident electron energy spectrum do not significantly modify the energy profile of the generated photons. The simulation checks performed (see next section) justify this point, as will be seen later.

The next piece of information at our disposal is the final electron energy as measured by the tagging system. Here the situation is somewhat more complicated, since the final electron energy is directly related to the energy of the photon(s) produced, and the triggering efficiency depends on the incident photon energy. Thus, no 'direct' comparison between tagging Monte Carlo and real data can be performed until a trigger acceptance correction is introduced in the tagging Monte Carlo. Therefore, we have to rely upon a Monte Carlo simulation of our trigger. This does not increase uncertainties significantly since our triggering scheme is relatively simple and the physics processes (hadronic event photoproduction) are well understood. Also, some cross-check is possible through small samples of special data collected in between normal runs, when the hadronic trigger conditions were relaxed.

Having obtained a trigger efficiency from our Monte Carlo (Figure 43) we could, in principle, check the tagging Monte Carlo against real data as follows: we could start with the original incident electron energy spectrum, simulate the bremsstrahlung of photons, see which of the photons will finally be seen by our trigger, and check the Monte Carlo-generated final electron energy (with trigger

acceptance taken into account) against real data. Had the two distributions been similar we would gain confidence in our simulation Monte Carlo. However, such a comparison is not possible for the data period we are considering (run numbers 5826 to 6347) due to hardware problems (discussed in 11.2.3) creating a 'hole' in the final electron energy distribution, thus making impossible to formulate the outgoing electron acceptance in a satisfactory way. The way out for obtaining the correct outgoing electron spectrum was to take the 'reference' outgoing electron spectrum into account *a posteriori* to weight individual events in the Monte Carlo so that the spectrum of outgoing electrons resembles the real one. Thus, agreement between Monte Carlo and real data was achieved but, of course, comparison between the two is meaningless.

11.2.4.4 Tagging Monte Carlo checks

Figure 40 shows the normalised incoming and outgoing electron spectra as well as the tagging answer obtained by the tagging Monte Carlo simulation for normal hadronic events (solid line). Also shown, superimposed, are the reference distributions, also normalised (dashed line). Figure 41 shows the ratio of Monte Carlo to reference distributions of Figure 40. The vertical scale has arbitrary units. This ratio should be flat for good agreement between Monte Carlo and real data; the agreement is good as expected from the *a posteriori* choice of the outgoing electron spectrum. However, the fact that the ratio of the incoming electron spectrum between Monte Carlo and real data is clearly falling with energy, shows that our assumption that the trigger and outgoing electron acceptances do not modify the incoming electron spectrum significantly, is not entirely correct. This is particularly the case for the high energy electron tail, where most of the electrons cannot lose enough energy to fall within the outgoing electron acceptance (which is effectively zero above 100GeV). Nevertheless the simulation is not a bad approximation of the real data for the energy range we shall be considering (40 to 160GeV) and the overall agreement demonstrates that both the simulation and the tagging are self-consistent.

It should be stressed at this point that the comparison between real data and Monte Carlo is only performed to establish the self-consistency of the scheme. Good agreement between real data and

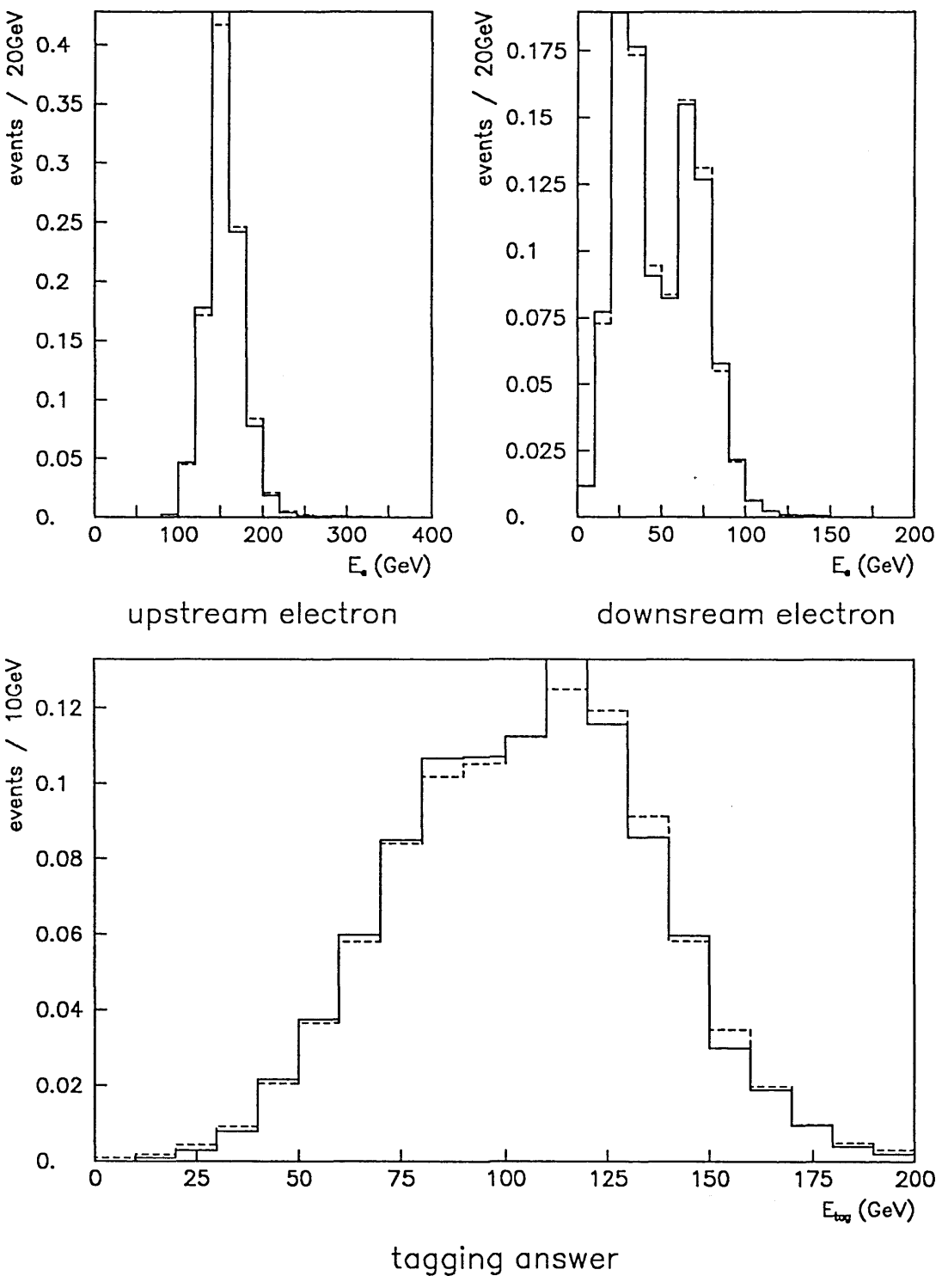


Figure 40: Real and simulated tagging data

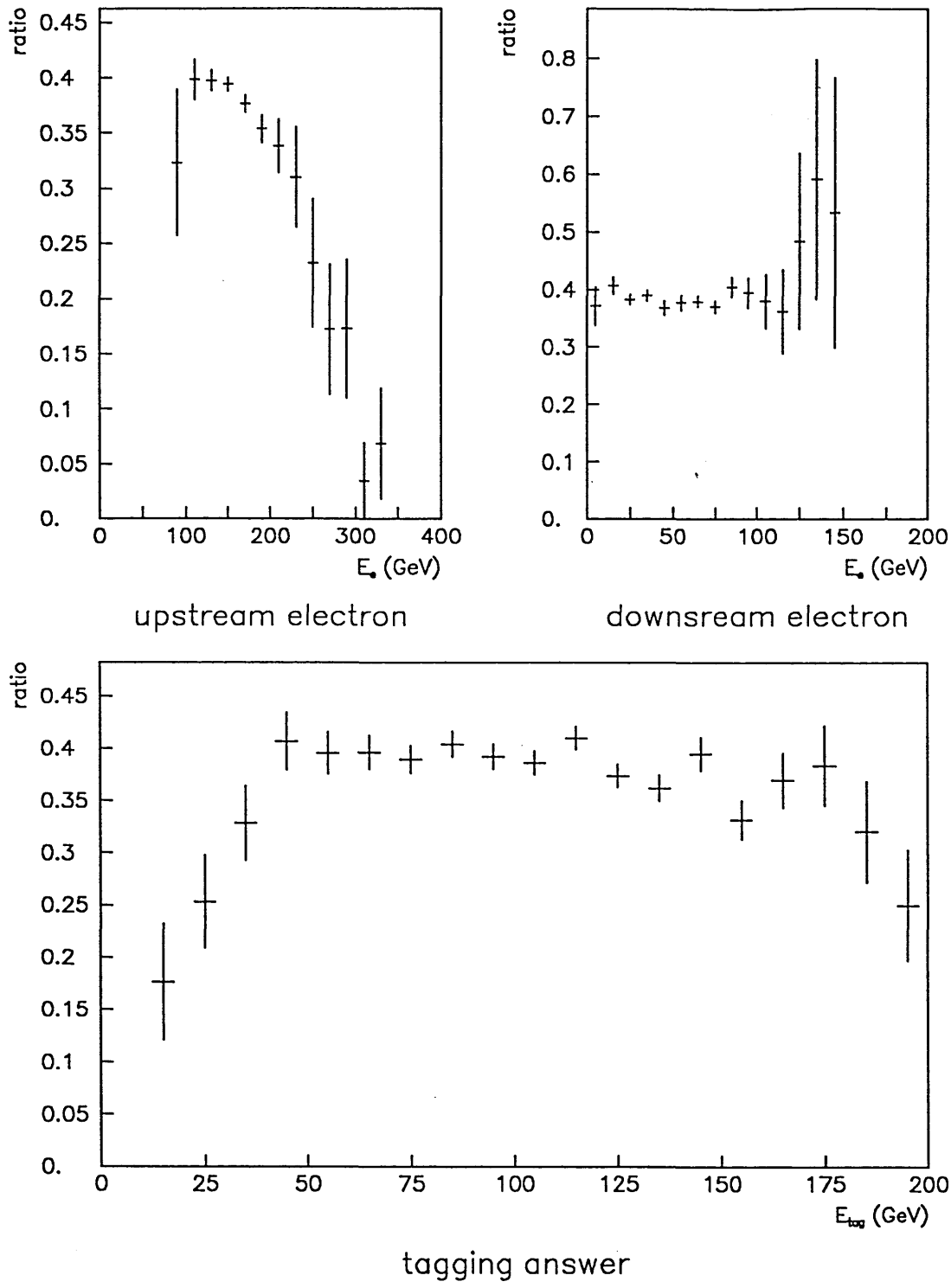


Figure 41: Comparison between real and simulated tagging data

Monte Carlo (meaning that all acceptances have been correctly taken into account) is not needed for what we are aiming to use this simulation for: the derivation of a scheme to convert tagging answer to incident photon energy. This, to a good approximation, only depends on the physical process involved (electron bremsstrahlung over a given amount of material) and not on the incident or outgoing electron spectra; nor does it depend on the trigger acceptance curve: in the discussion to follow, we shall apply the tagging correction obtained with the normal hadronic event trigger acceptance to our charm sample (where the trigger acceptance curve is different).

11.2.4.5 Tagging Monte Carlo results

The reason for introducing the tagging Monte Carlo is to obtain a means of transforming a tagging answer distribution to a photon energy distribution. If we assume that our tagging and real photon spectrum are discreet functions that can take the form of an n-dimensional vector (where n is the maximum energy we are considering over our granularity) then this transformation function can have the form of a an (n×n) matrix. The elements of this matrix are actually the entries (suitably normalised) of a two-dimensional histogram of tagging answer versus photon energy which can be obtained from our Monte Carlo simulation. For the range 0 to 200GeV, and with a granularity of 10GeV, the transformation matrix elements are listed in Table 20. The amount of radiator material is, again, 12.5%.

As a consistency check, the tagging answer as obtained from the Monte Carlo was transformed using the transformation matrix coefficients shown in Table 20 to its equivalent photon energy spectrum. This was checked against the actual Monte Carlo-generated photon energy spectrum. The results, shown as a ratio of the two distributions can be seen in Figure 42. The errors shown correspond to the statistical uncertainties of the Monte Carlo generated spectra had the two spectra been uncorrelated. This is clearly not the case as can easily be deduced by looking at this figure.

Table 20: Tagging answer to photon energy transformation matrix coefficients

	0	10.	20.	30.	40.	50.	60.	70.	80.	90.	100.	110.	120.	130.	140.	150.	160.	170.	180.	190
0.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
10.	0.000	1.000	0.098	0.078	0.013	0.011	0.014	0.008	0.006	0.009	0.005	0.008	0.006	0.006	0.003	0.007	0.000	0.000	0.000	0.000
20.	0.000	0.000	0.902	0.096	0.035	0.022	0.015	0.023	0.008	0.009	0.010	0.006	0.009	0.008	0.007	0.007	0.004	0.007	0.017	0.000
30.	0.000	0.000	0.000	0.826	0.127	0.064	0.031	0.037	0.024	0.017	0.020	0.013	0.015	0.017	0.020	0.009	0.004	0.014	0.000	0.000
40.	0.000	0.000	0.000	0.000	0.826	0.157	0.071	0.041	0.037	0.031	0.032	0.020	0.018	0.020	0.017	0.013	0.007	0.014	0.000	0.000
50.	0.000	0.000	0.000	0.000	0.000	0.746	0.172	0.075	0.048	0.039	0.035	0.024	0.030	0.015	0.018	0.031	0.021	0.021	0.000	0.000
60.	0.000	0.000	0.000	0.000	0.000	0.000	0.698	0.134	0.050	0.035	0.032	0.027	0.027	0.011	0.020	0.011	0.018	0.007	0.000	0.000
70.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.682	0.156	0.064	0.038	0.037	0.028	0.019	0.028	0.042	0.007	0.028	0.017	0.034
80.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.670	0.160	0.074	0.041	0.037	0.028	0.024	0.027	0.021	0.021	0.000	0.034
90.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.634	0.154	0.066	0.042	0.031	0.026	0.013	0.021	0.021	0.000	0.000
100.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.600	0.121	0.068	0.039	0.039	0.029	0.028	0.014	0.034	0.000
110.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.637	0.147	0.068	0.056	0.031	0.035	0.007	0.017	0.000
120.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.572	0.158	0.062	0.033	0.039	0.014	0.000	0.034
130.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.580	0.123	0.044	0.032	0.028	0.017	0.000
140.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.557	0.133	0.049	0.077	0.017	0.103
150.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.570	0.109	0.035	0.068	0.000
160.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.606	0.106	0.102	0.069
170.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.585	0.102	0.138
180.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.610	0.069
190.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.517

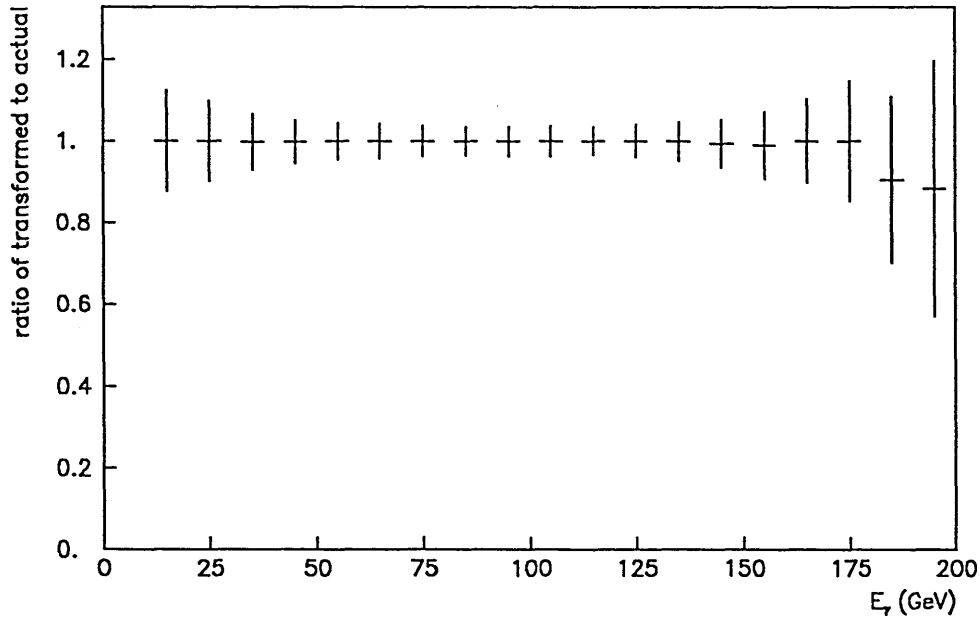


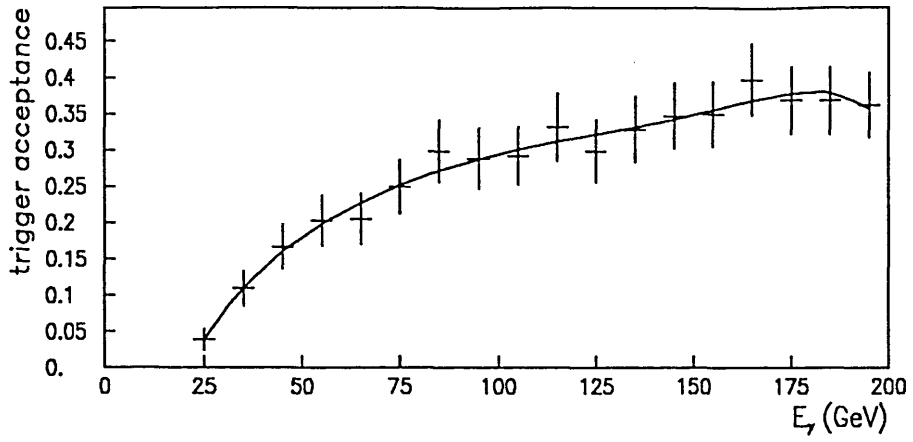
Figure 42: Ratio of transformed to actual photon energy distribution

11.3 Acceptances

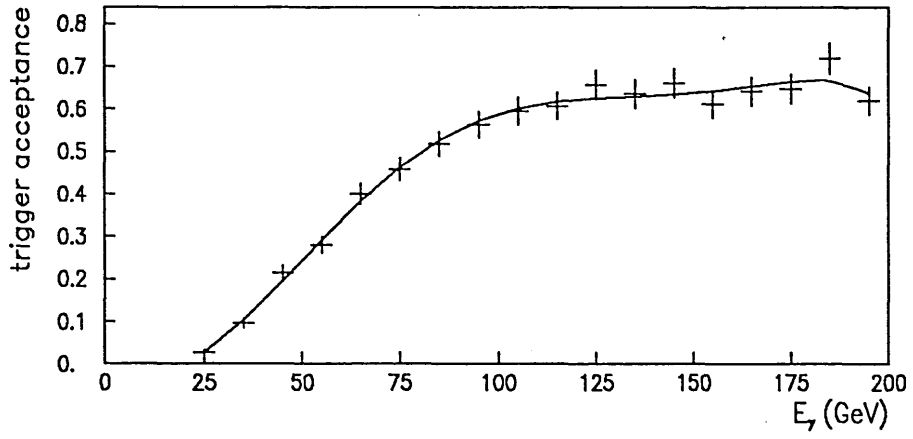
11.3.1 Trigger acceptance

The trigger acceptances were obtained from the NA14 detector simulation (discussed in chapter 8). The charm and non-charm (hadronic) event trigger acceptance was calculated as a function of the incident photon energy by generating 10,000 events in each category with an incident photon energy ranging from 20 to 200GeV and recording the percentage of events passing the full trigger condition. The result is a curve that rises with energy, as the average longitudinal momentum, p_L, of the generated charged particles increases and, therefore, more particles penetrate Goliath, and then flattens out to about 37% for the normal hadronic event case, and to 67% for the charm case, as it can be seen from Figure 43. The parametrization chosen for the fitting curves is that of a high order polynomial but is only included to guide the eye. In the analysis to follow the trigger acceptance corrections were performed using the raw distributions instead of the equivalent parametrised functions to avoid possible systematic effects when comparing the normal and charm event spectra. Figure 44 shows the ratio of the charm trigger acceptance to the normal hadronic event acceptance. The solid line is the ratio of the corresponding fitted functions. For high energies, the ratio between the trigger acceptances of normal and charm events is roughly constant with a charm enrichment factor of about 2, as originally predicted by the NA14 proposal [52].

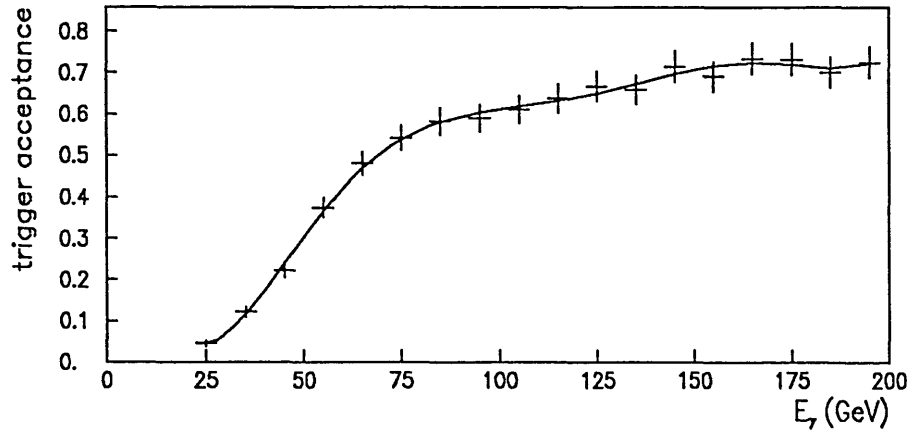
A check of the Monte Carlo-derived trigger acceptance with real data is possible (in the case of normal hadronic events). Data exist which were taken with a fixed energy pion beam and without any trigger condition (only the requirement that a SPS spill should be present). These data are useful in providing an absolute means for checking the Monte Carlo-calculated trigger efficiency at this fixed energy. However, statistics are limited and the energy dependance cannot be checked since the pion beam is monochromatic. A study of some of those non-trigger data taken with a 90GeV pion beam has been made [63], and the trigger efficiencies measured were found to be consistent with what the Monte Carlo predicts at that energy.



trigger acceptance - normal events



trigger acceptance - Kπ events



trigger acceptance - K$\pi\pi$ events

Figure 43: Trigger acceptance for hadronic and charm events

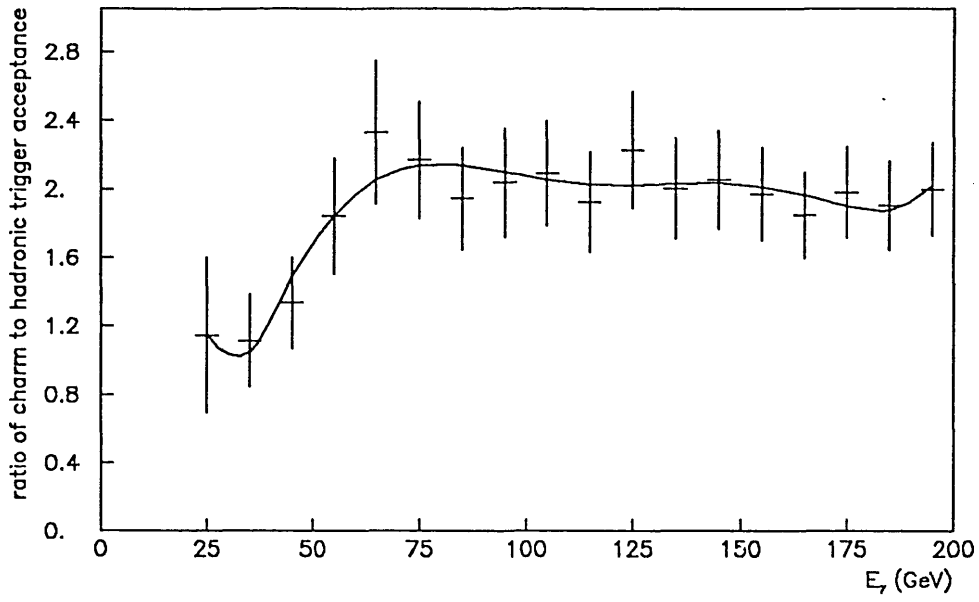


Figure 44: Ratio of trigger acceptances of charm / normal events

11.3.2 Trigger hadronicity

Another important feature of the trigger is its *hadronicity*, that is the percentage of triggered events that are not electromagnetic. The trigger of NA14 is designed to select hadronic events, and minimum contamination from electromagnetic events is expected. A study [63] has estimated the hadronicity of the trigger by visually scanning a number of events recorded with the standard trigger condition. The trigger was found to be 95% (±3%) hadronic.

11.3.3 Analysis acceptance

The analysis acceptance is defined as the combined acceptance of the whole analysis chain: this comprises the event reconstruction and the main analysis (including all selection cuts to obtain our charm signal). It also includes the 'smearing' effect due to the finite resolution of the detectors.

Unlike the trigger acceptance where the obvious parametrisation is versus the energy of the interacting photon, here we have many possible parametrisation variables. One solution is to again

parametrise as a function of the energy of the event generating photon. However, the analysis acceptance is mainly dependent on the D meson kinematics, therefore it is more appropriate to parametrise in terms of the D meson momentum.

This analysis efficiency is derived using the main NA14 simulation program and is defined as the ratio of selected events after reconstruction and after all analysis cuts per (measured) D momentum interval over the number of events generated per (generated) D momentum interval. The detector resolution is taken into account by selecting the *measured* D momentum for the one case and the *generated* D momentum for the other. The effect of double counting (an event with two different combinations in the right mass region) is also taken into account by this method.

Again, 10,000 events were generated for the neutral and the charged D respectively, forcing their decay modes to $D^0 \rightarrow K\pi$ and $D^+ \rightarrow K\pi\pi$. These were processed through the reconstruction program and the same analysis programs used to obtain our real event signals. Then the number of D events in the specified 'signal region' per D momentum interval was divided by the number of generated D events. This ratio is the analysis efficiency for that momentum interval.

The analysis acceptance curves estimated thus can be seen in (a) and (b) of Figure 45, for the case of the D^0 and the D^+ respectively. For the case of D^0 the vertex separation cut (see next section) was 3σ and for the D^+, 4σ. The fitted curves are from the respective parametrised functions, a high order polynomial in both cases with good associated χ^2 probability. It is the parametrised functions rather than the raw distributions that have been used for the actual analysis acceptance correction. This is to reduce the error arising from the statistical fluctuations in the raw analysis acceptance curve. We can safely do that since a) the acceptance curve is expected to be smooth and b) we will not have to compare distributions corrected with two parametrised acceptance curves and the consequent risk of systematic errors (unlike the case for the trigger acceptance, where we do have to compare hadronic and charm events). The reason for the asymmetry of the two distributions at high D momenta comes from a cut in the analysis program of any charged particle that appears to have a momentum higher than 70GeV/c. (For such stiff tracks the detector momentum resolution is bad.) This cut introduces a sharp cutoff in the D^0 case of 140GeV/c, whereas the corresponding cutoff for the D^+ is too high at

210GeV/c.

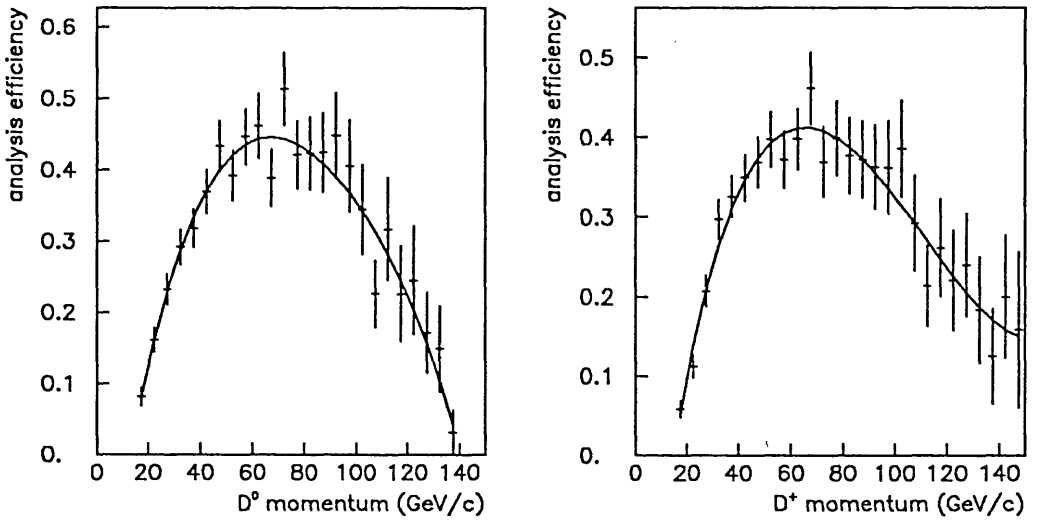


Figure 45: D^0 and D^+ analysis efficiency as a function of the D momentum

11.4 Main analysis

11.4.1 Data sample

The results presented here are based on 4.3 million fully reconstructed rawdata events taken from the emulator-processed sample. These data correspond to run numbers 5239 to 6325 from the September 86 data taking period (20th of September 1986 to 18th of October 1986) and run numbers 4784 to 4873 from the July 86 period (28th to 30th of July 1986). Emulator data were chosen because they are free of biases that would have been introduced if a filtering scheme had been incorporated (see 10.8.2). The decay channels analysed are the $D^0 \rightarrow K\pi$ and the $D^+ \rightarrow K\pi\pi$ channels. These were chosen because of the good statistics we can obtain due to their high branching ratio and good acceptance in our spectrometer.

11.4.2 The vertex and analysis package

The main strength of the NA14 spectrometer in isolating charmed decays lies in the very good spatial resolution of the microstrip telescope, since the most characteristic signature for charm is tracks offset from the main vertex. This is due to the fact that charmed particles decay weakly, therefore the typical charm lifetime is long enough (of the order of $10^{-13}-10^{-12}$ seconds) to allow the charmed particle to travel some distance before decaying. Resolving this offset, which is of the order of a few tens of microns in Y and Z projections and a few hundreds of microns along the X axis (the beam axis), requires a very precise tracking device such as a microstrip chamber. For this reason, vertex finding (for the primary and secondary vertices) plays a key role in the analysis for obtaining a charm signal. Information from other parts of the detector is essential for obtaining particle identities and momenta, but it is the cut on the distance between secondary and primary vertices that is crucial for increasing the charm signal to noise ratio. The analysis and vertex reconstruction package used for this study was developed by colleagues at LAL, Orsay, and is described in detail in references [64] and [54]. Its general philosophy and main features are noted below. The analysis uses the tracks reconstructed by NA14 TRIDENT; it loops over all tracks selecting the right combination of K's and π's (oppositely charged K and π for the D^0, similarly charged π's and oppositely charged K for the D^+) and computes the distance of minimum approach with its corresponding error. If the χ^2 probability of the hypothesis that these tracks form a vertex is better than 1% then this combination is taken as a candidate D decay. The position of the point of minimum approach in space constitutes the secondary vertex position. The corresponding D meson momentum vector is then computed and the search for primary vertex begins: Each remaining track (excluding the D daughter tracks) is taken in turn and its minimum distance of approach from the D track is computed. This is done for all available tracks and the overall χ^2 is computed; if the χ^2 probability is less than 1%, the track contributing the highest partial χ^2 is rejected (the D track is not considered for this rejection) and the process is repeated until there are either less than 2 tracks left, in which case the event is rejected, or a probability of more than 1% is achieved; in the second case, the position in space to be taken as the position of the primary vertex is the point where the sum of squared distances of all tracks remaining in the fit is

minimum.

Given a primary and a secondary vertex with their corresponding errors, the vertex separation can be presented in terms of the combined error on primary and secondary vertices, Nσ. Cutting on this quantity, provided that the errors have been calculated correctly, produces the best effect on background rejection with the smallest loss of good events.

Kaons and pions are distinguished with the help of the second Cerencov counter, INDRA, operating in threshold mode and being able to discriminate kaons from pions between 5.7GeV, the pion threshold, and 20.3GeV, the kaon threshold. However the requirement for K identification significantly reduces our statistics, since particles with momenta greater than 20GeV/c are rejected. For this reason, the definition of an 'extended' kaon was adopted, a particle above the kaon threshold in the Cerenkov counter which can be either a K or a π. These 'extended' kaons were treated as kaon candidates, increasing the D candidate events. The effect of this generalised kaon adoption increases our signal significantly, requiring a slightly higher Nσ cut to obtain a similar signal to noise (S/N) ratio as in the identified kaon case as shown in [47]. Throughout this analysis the definition of a Kaon follows the above extended definition. A further cut, as already mentioned, limits the maximum momentum of any track to 70GeV/c. This cut was introduced since the spectrometer momentum resolution deteriorates at such high momenta.

11.4.3 D^0 and D^+ mass spectra

The mass spectra of D^0 and D^+ for various Nσ cuts can be seen in Figure 46 and Figure 47 respectively.

To estimate the amount of background and to obtain the reconstructed D mass and width (resulting from the detector resolution), the histograms corresponding to the Nσ cuts used for the rest of the analysis were fitted with a function parametrised as follows: a polynomial for the background and a gaussian with unconstrained mean and width for the signal. When fitting the background the mass range where the signal was expected to be was excluded from the fit. To avoid bias, the 'signal area' was defined as the region having the world average D mass as mean and a width derived from the

detector resolution. The expected detector resolution was determined from the NA14 Monte Carlo simulation. The polynomial order for the background parametrisation was varied and the effect on the χ^2 of the fit and on the number of the background events in the signal region tabulated. For the D^+ the resulting χ^2 fit was best with a linear background variation, therefore the S/N ratio obtained using that parametrisation was used. For the D^0, the S/N ratio used was obtained from a fifth order polynomial parametrisation to the background, that gave the best fit. However, the S/N ratio was compatible with that obtained using a linear background fit. Intermediate order polynomials gave a poorer S/N ratio, but also a worse χ^2 probability.

The signal to noise ratio obtained from the fits is 1:2.1 for the D^+ (cut at 4σ) and 1:1.5 for the D^0 (cut at 3σ) case. The uncertainty in estimating the S/N ratio will be taken into account in the overall systematic error. Figure 48 shows the fitted D^0 and D^+ signals used for this analysis. The background parametrisation is linear for these plots. The D masses obtained from our fit, which did not vary significantly with different background parametrisation were $1.853\pm.014$GeV and $1.861\pm.010$GeV for D^0 and D^+ respectively, compatible with the world average values of $1.8645\pm.0006$GeV and $1.8693\pm.0006$GeV [44].

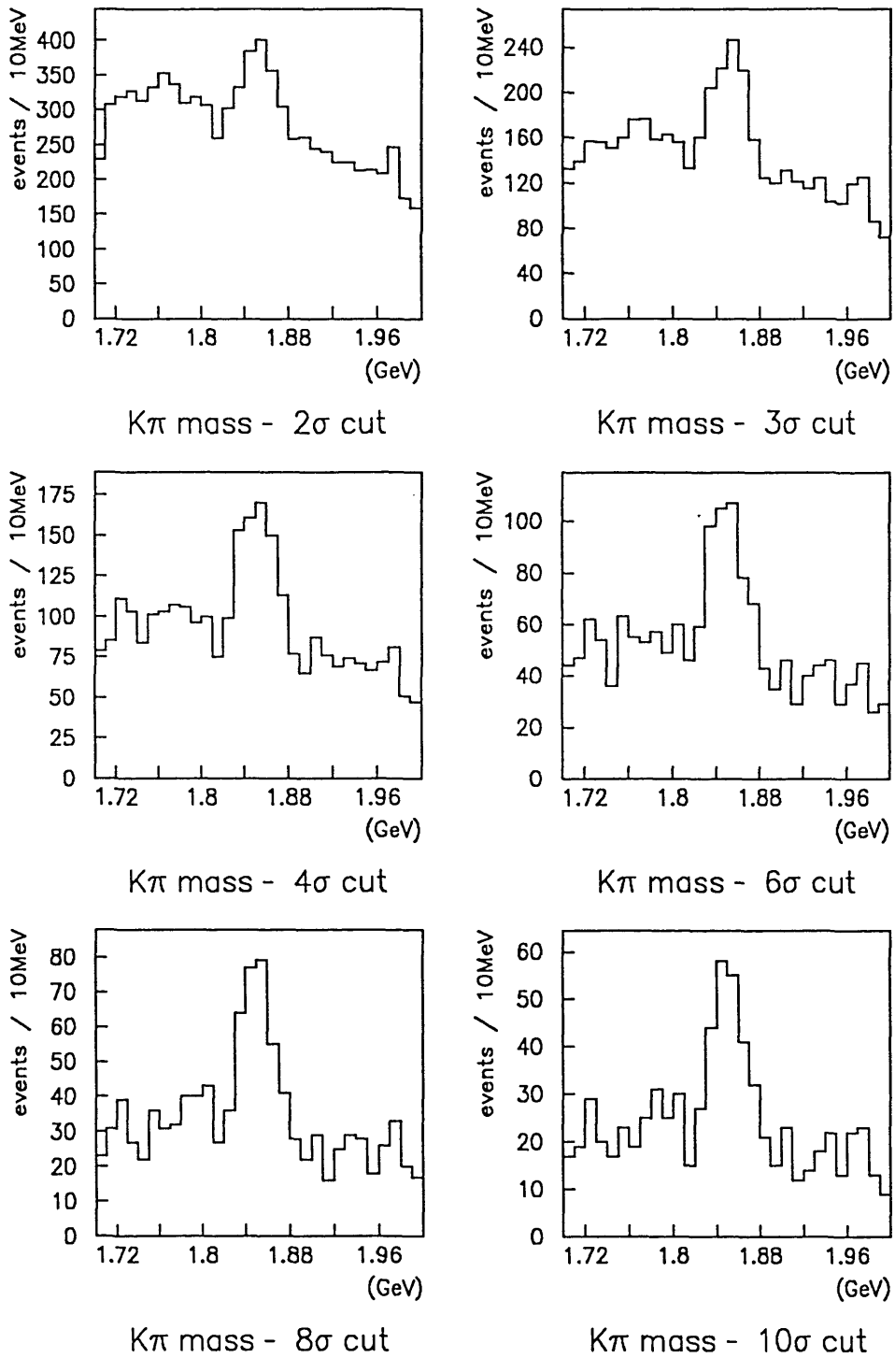


Figure 46: $D^0 \rightarrow K\pi$ mass spectra obtained for various $N\sigma$ cuts

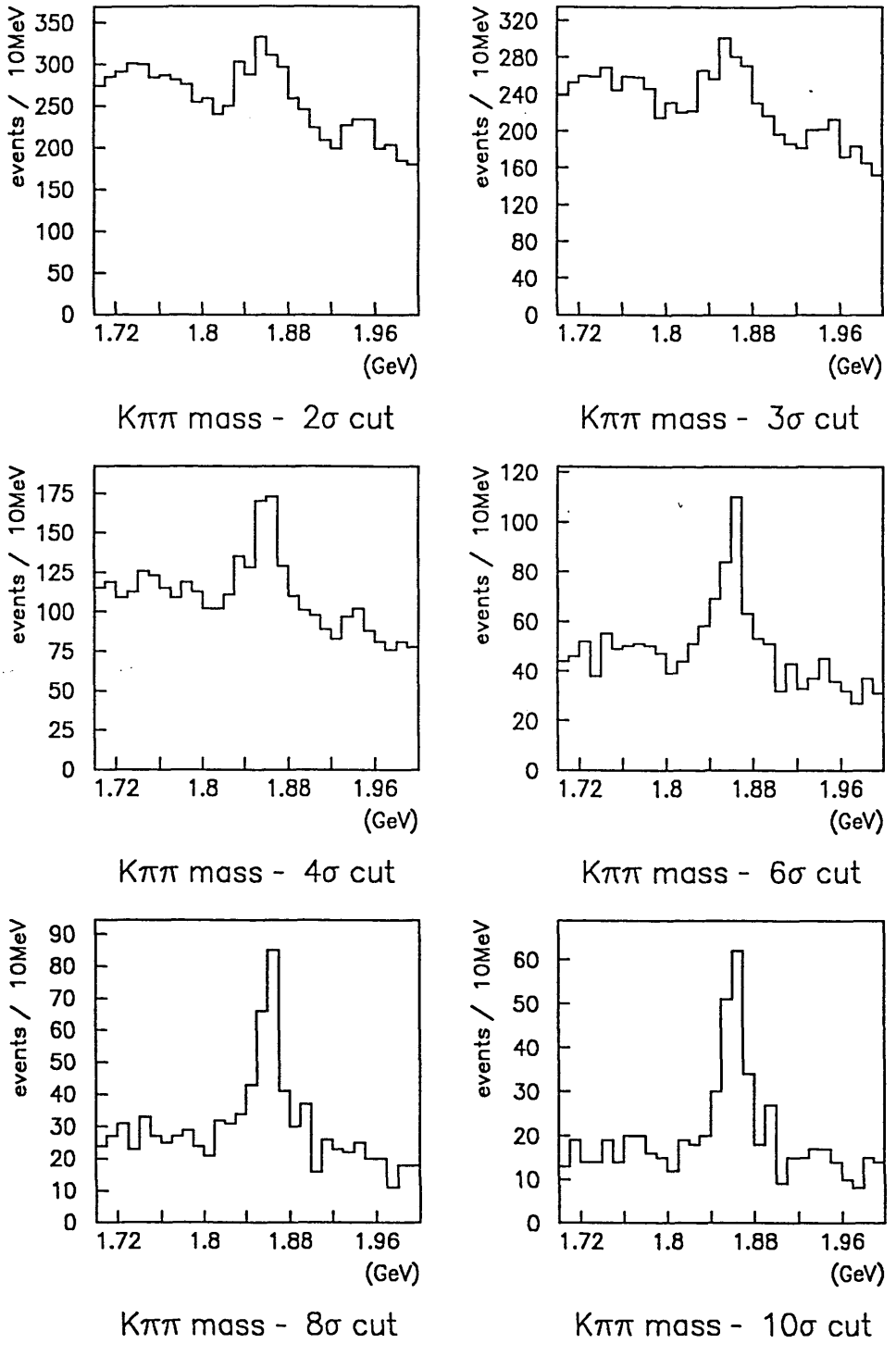


Figure 47: $D^+ \rightarrow K\pi\pi$ mass spectra obtained for various Nσ cuts

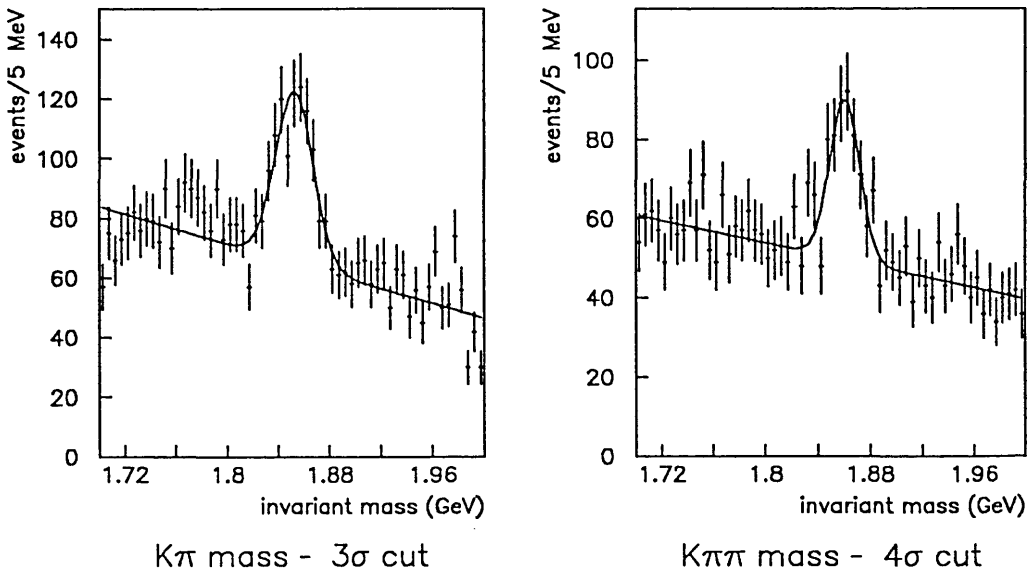


Figure 48: D^0 and D^+ fitted signals

11.4.3.1 Choice of Nσ cut for the cross section analysis

The Nσ cut determines the number of events in our signal region and the signal to noise ratio. The choice of the Nσ cut depends on two conflicting parameters: on the total statistics that can be used (the smaller the Nσ cut the better), and on the S/N ratio (deteriorating with decreasing Nσ cut). Deciding on the best Nσ cut is not simple and depends on how different the signal and background distributions with respect to the measured quantity are. This decision can only be made after different Nσ cuts are used for the same analysis, and the one producing the smallest (fractional) errors chosen.

The step in the analysis where the choice of the Nσ cut used reflects in the overall errors is the step of background subtraction. On calculating the overall error after this step for various Nσ cuts we found an increase in the overall errors at big Nσ cuts and a rather flat behaviour of the errors for small Nσ cuts. We finally chose the Nσ cut which gives the best S/N ratio without deteriorating the overall errors. This corresponds to a vertex separation cut of 3σ for the D^0 case, giving a S/N ratio of about 1:1.5, and 4σ for the D^+ analysis giving a S/N ratio of 1:2.1.

In the case of D^0 we can enhance our signal by imposing a D^* cut: the candidate D^0 is checked against the hypothesis that it was a product of a $D^* \rightarrow D^0\pi$ decay. In this way, a much cleaner signal is obtained for the events that pass this cut, therefore a much lower Nσ cut is needed for the same S/N ratio (see Figure 38). There are about 100 such events in the sample at an Nσ cut of 0.25 and a S/N ratio close to 1:1.5; about 70 of them survive the 3σ cut, which is the cut used in the D^0 analysis. The gain obtained by treating the D^* cut events at a lower Nσ cut separately would therefore be only about 10% of the total D^0 statistics (30 events in a total of 300). This was not considered enough to justify separate treatment.

The signal region was defined taking into account the reconstructed D^0 and D^+ masses (1.853 and 1.859 GeV respectively) and our detector resolution, slightly worse in the three body decay of the D^+, and were chosen to be, for the case of D^0 1.830 to 1.875 GeV and for the case of D^+ 1.830 to 1.885 GeV.

11.4.4 Raw tagging answer spectra

The tagging answer spectrum associated with events in the signal region as well as the energy spectrum of the background (taken from the wings of the distribution) for both D^0 and D^+ can be seen in Figure 49. The average tagging efficiency for the selected events is 32%. We have 323 D^0 events with a tagging answer and mass between 1.83 and 1.875 GeV (out of the total of 969 events in the same mass region) and 246 D^+ events with a tagging answer and a mass between 1.83 and 1.885 GeV (out of the total of 802 events in the same mass region).

Note that the shape of the background is very similar to the shape of the signal. This could suggest that our background events at the Nσ cuts used have the same cross section variation with energy as charm events if we assume that the acceptance of the background events as a function of energy has a similar shape to the acceptance of charm events. If this is correct then the similarity in shape suggests that the background at this level consists mainly of charm events that have been partly reconstructed. However, this is only a speculation since the acceptance for non-charm events as a function of energy has not been studied.

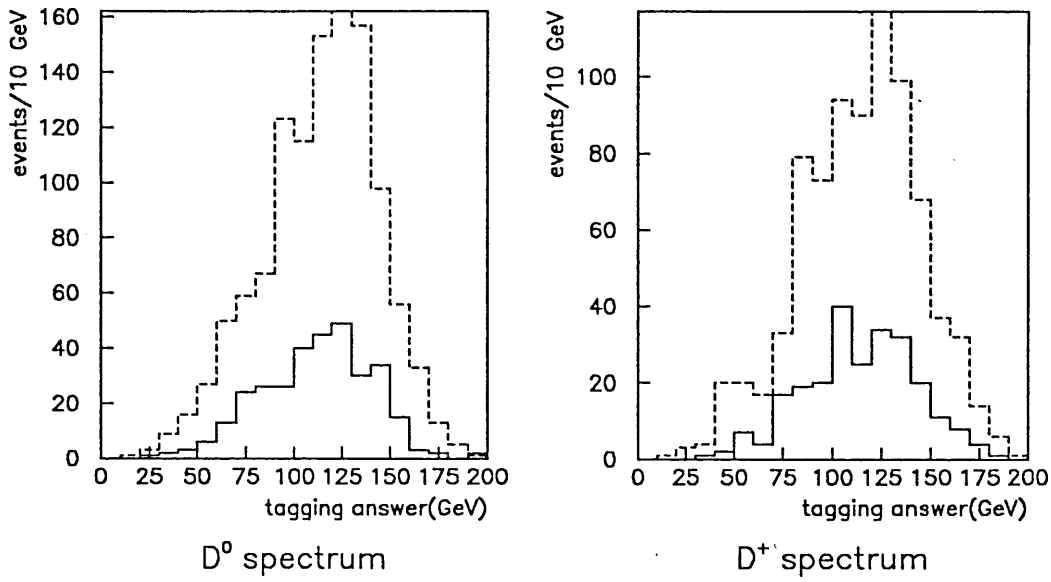


Figure 49: $D^0 \rightarrow K\pi$ and $D^+ \rightarrow K\pi\pi$ tagging answer spectrum (raw). Background events from the wings of the mass plot appear as the dashed curve. The solid line is from events in the signal region.

11.4.5 Background subtraction

Background subtraction is performed under the assumption that background events in the signal region of the mass distribution have an equivalent tagging answer spectrum shape and D momentum spectrum shape to background events taken from the wings of the mass distribution. We can then use the spectra obtained from the wings of the mass distribution to model the spectra of the background events in the signal region. Knowing the signal to background ratio in our signal region, we can subtract the background to determine the pure signal spectrum. If $T(E_T,p_D)$ and $B(E_T,p_D)$ are the spectra of the events in the signal region and the background region respectively as a function of the tagging answer, E_T, and the D momentum, p_D, and n is the signal to background ratio then the pure signal spectrum, $S(E_T,p_D)$ is given by:

$$S(E_T,p_D) = T(E_T,p_D) - \frac{1}{n+1}B(E_T,p_D)$$

Note that we need to express our spectra in the two dimensional space of tagging answer and D momentum since we shall need the D momentum information for the analysis acceptance correction. Figure 49 shows the projection to tagging answer space of $T(E_T p_D)$ (solid line) and $B(E_T p_D)$ (dashed line). Figure 50 shows the result of the background subtraction operation, $S(E_T p_D)$, also projected in the tagging answer space.

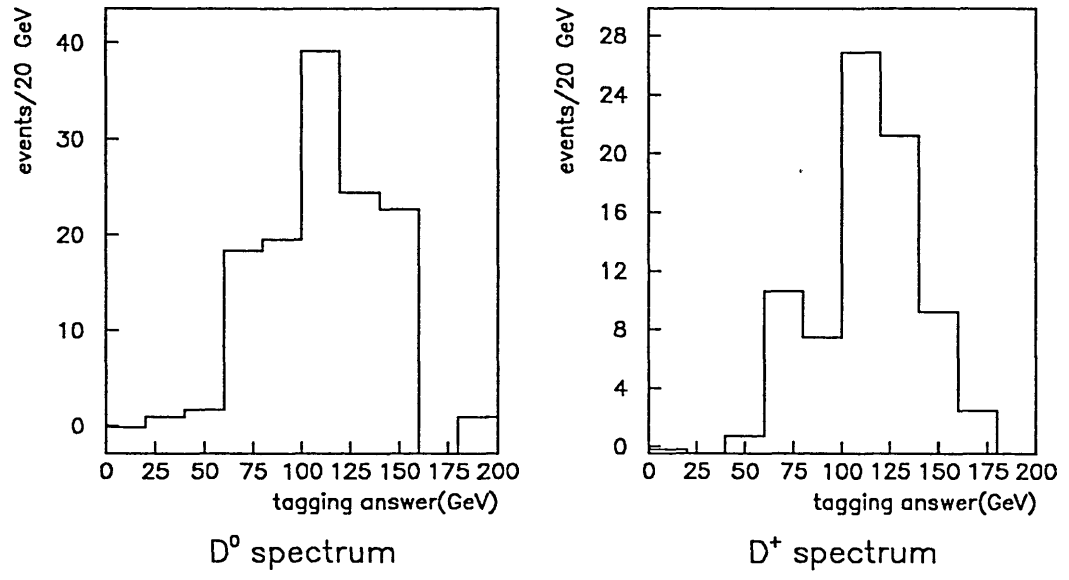


Figure 50: $D^0 \rightarrow K\pi$ and $D^+ \rightarrow K\pi\pi$ tagging answer spectrum (background subtracted)

11.4.6 Analysis acceptance correction

Having obtained the pure signal spectrum in the two dimensional space of the tagging answer and D momentum, we can now proceed to correct for the analysis acceptance (discussed in section 4.3.3) simply by dividing the pure signal spectrum by the analysis efficiency (Figure 45) in the D momentum space.

Since for very low D momenta (below 15GeV/c) the analysis acceptance is very small, the correction factor for the low statistics in that region is large. This yields bins with a large number of events and correspondingly large errors. If errors are taken into account properly these events should not bias

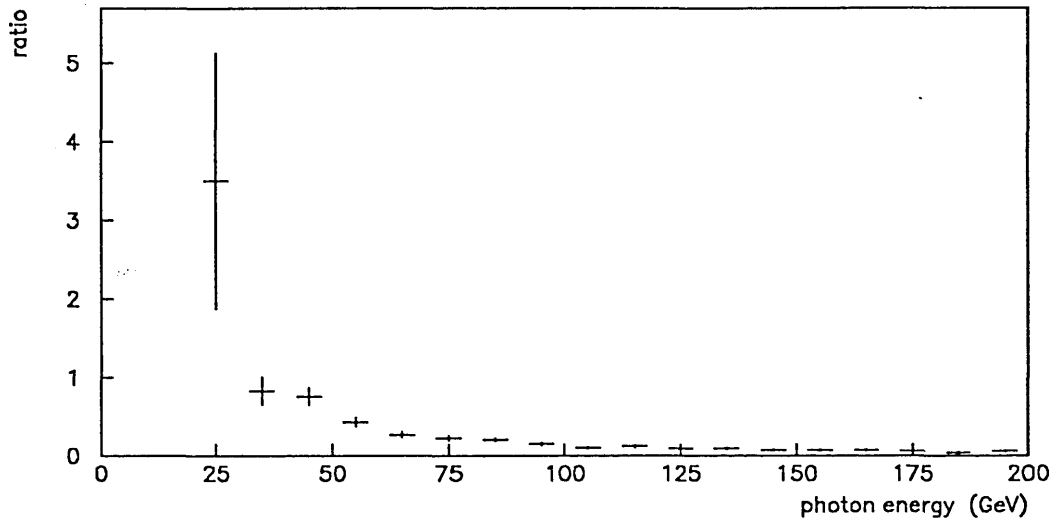


Figure 51: Ratio of events with D momentum less than 15GeV/c to events with D momentum greater than 15GeV/c versus incident photon energy

our analysis. However, to estimate the systematic error of our Monte Carlo simulation for this momentum region would not be easy and in any case the small number of events with such low D momenta add little information to our analysis. For that reason, a cut on the momentum of the D of 15GeV/c was applied, thus avoiding the region where statistical uncertainties on the signal and systematic uncertainties on the acceptance correction are high. The effect of this cut was accounted for in the following way: From the NA14 simulation Monte Carlo the ratio of events per energy interval with D momentum less than 15GeV/c was plotted versus the incident photon energy; this can be seen in Figure 51; This histogram was taken into account after the analysis correction to compensate for the loss of events due to the 15GeV/c cut in the tagging answer space. For the energy range relevant for this analysis (40 to 160GeV) the correction is of the order of 15%.

Note that this procedure is not entirely correct: the correction factor is calculated versus the incident photon energy spectrum whereas our data reside in the tagging answer space. However, the effect represents a small perturbation to the already low (15%) correction, therefore we can safely make this

simplification. The effect of the analysis correction in the D momentum space can be seen in Figure 52; the solid line represents the Tagging answer spectrum after the analysis acceptance correction, whereas the dashed line is the spectrum before the analysis acceptance correction.

The pure signal spectrum (in Tagging space this time), corrected for the effect of the 15GeV/c cut, is shown in Figure 53.

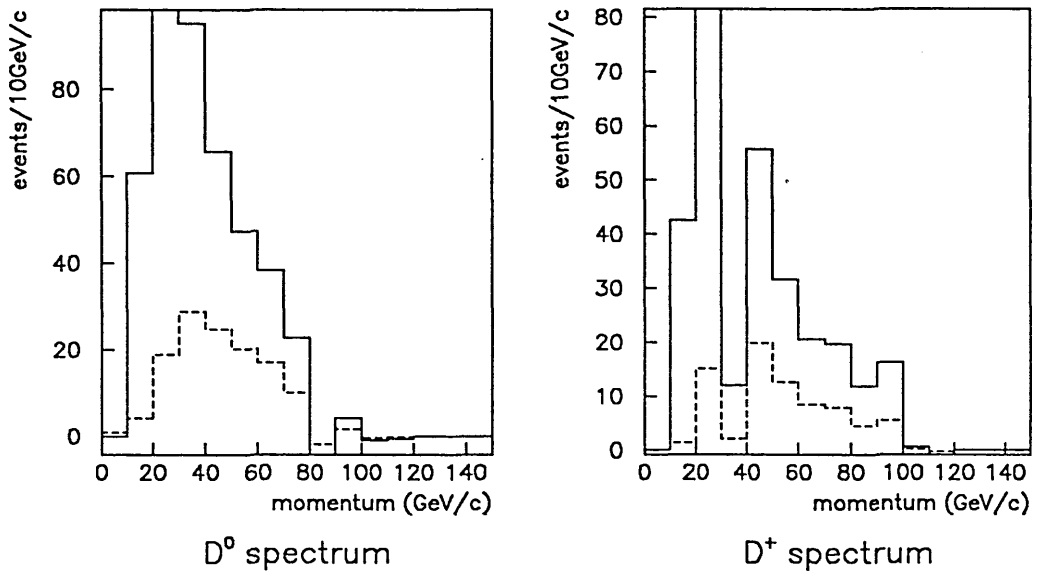


Figure 52: D^0 and D^+ momentum spectra before (dashed curve) and after (solid curve) the analysis correction

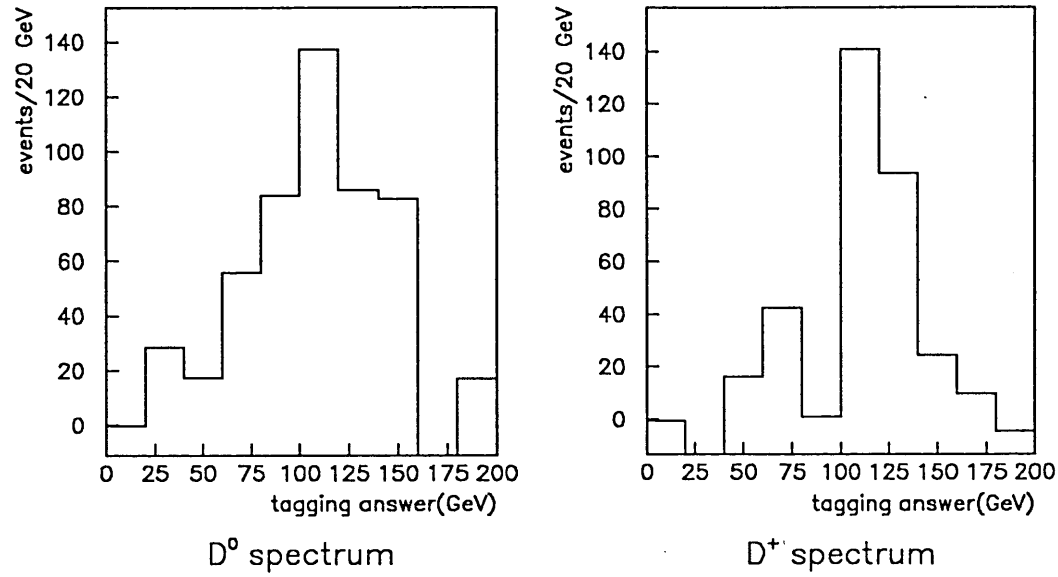


Figure 53: Analysis corrected D^0 and D^+ spectra in tagging answer space

11.4.7 Tagging correction

Having corrected for the analysis acceptance, we now have the spectrum of D events that would have been obtained had our detector and analysis chain been 100% efficient in finding every event that was recorded on tape. However, this spectrum appears in the Tagging answer space which is not very useful since the tagging answer is not related directly to the incident photon energy. We can, however, transform those spectra to the incident photon energy space if we apply the tagging answer to photon energy transformation discussed in section 11.2.4.4. The real photon energy spectra obtained thus can be seen in Figure 54. Also, the equivalent histogram (with the tagging correction) representing the energy spectrum of normal hadronic events can be seen in Figure 55.

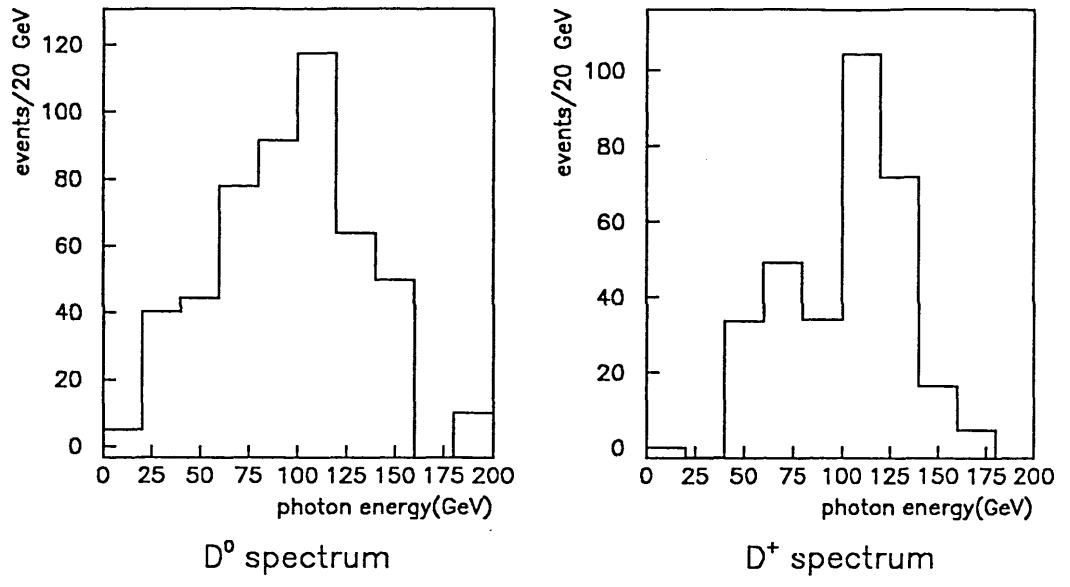


Figure 54: D^0 and D^+ spectra transformed from tagging answer to incident photon energy

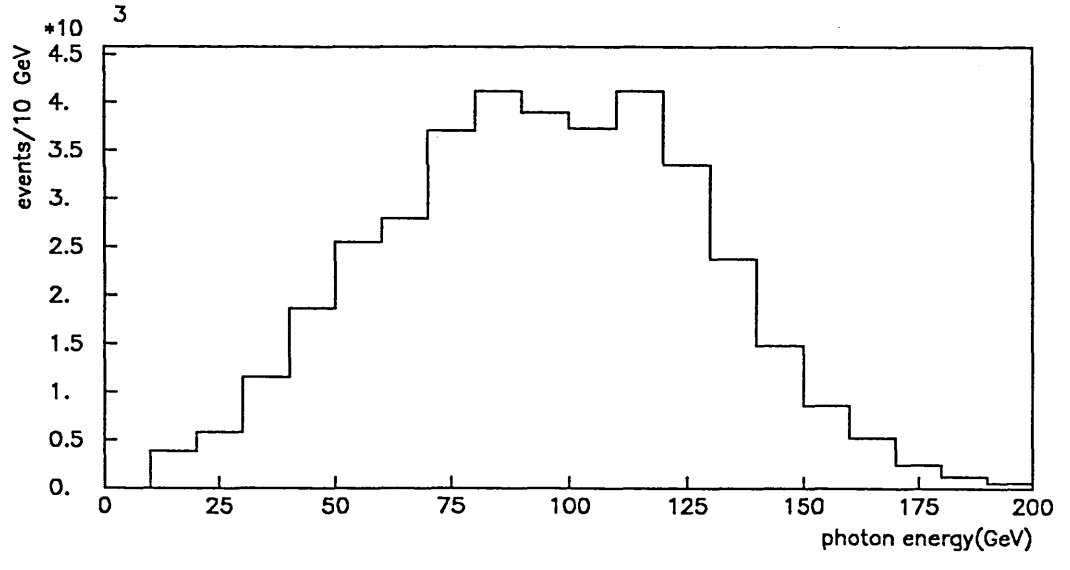


Figure 55: Incident photon energy spectrum for hadronic events transformed from tagging answer space

11.4.8 Trigger acceptance correction

The last step for obtaining the photon energy spectrum of D^0 and D^+ is to apply the trigger acceptance correction discussed in section 4.3.2. The results of this operation can be seen in Figure 56; the spectra shown represent the total number of $D^0 \rightarrow K\pi$ and $D^+ \rightarrow K\pi\pi$ events generated at the experimental target during the run considered (during which 4.3M rawdata events were recorded). These spectra will be used to derive the relative cross section variation with respect to energy by comparing them with the equivalent histogram of normal hadronic events. This equivalent rawdata spectrum (transformed to photon energy spectrum and corrected for trigger acceptance) taken from the reference tagging histograms can be seen in Figure 57. There is an additional correction that needs to be applied to this spectrum: the hadronicity of our trigger is taken to be 95% (see 11.3.2), and we assume this to be constant with energy, therefore we reduce the number of events in each bin by 5%, the amount of electromagnetic contamination. (There is effectively no electromagnetic contamination to our charm sample.) All trigger acceptances are defined above 20GeV, therefore no data are plotted below that value.

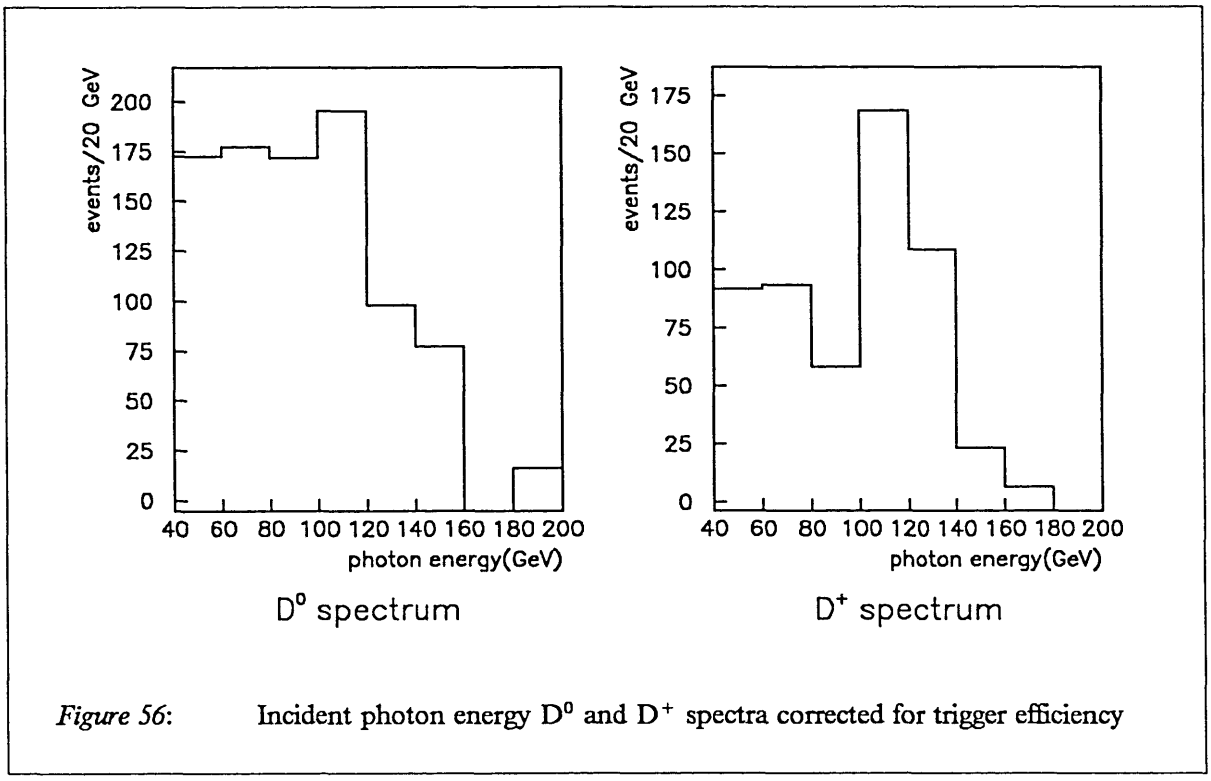


Figure 56: Incident photon energy D^0 and D^+ spectra corrected for trigger efficiency

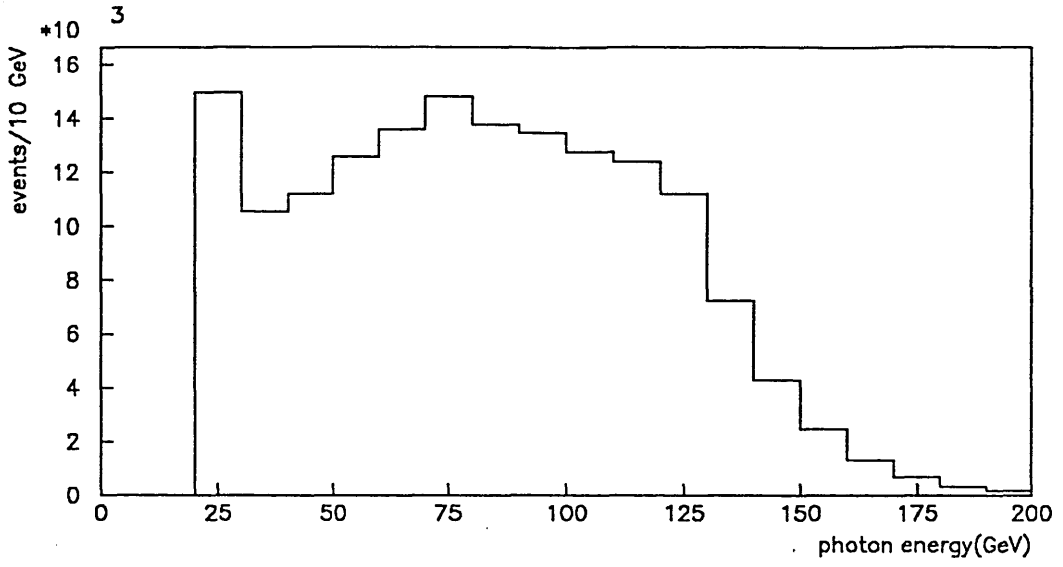


Figure 57: Incident photon energy spectrum corrected for trigger efficiency for normal events

11.5 Absolute cross section measurement.

Having obtained a photon energy spectrum for charm as well as for normal hadronic events corrected for all acceptances we can proceed to measure the charm photoproduction cross section. This can be achieved for a specific charmed particle decay mode of known branching ratio, provided we can estimate the contribution of that specific particle to the total charm production. Then, by comparing the number of events obtained for that charmed particle with the number of events obtained for a process with a known cross section, we can calculate the charm cross section at the energy range we are considering. The process with a known cross section is the photoproduction of hadronic events, and the charmed particles considered are the D^0 and D^+.

The photon energy range chosen for this study is 40 to 160GeV. Although we do get events below 40 and above 160GeV, the available statistics are very low. Also, for the low energy part of the spectrum the uncertainties in the calculation of the various acceptances are high.

The charm photoproduction cross section of *any* charmed particle (denoted here by "D") – and, hence, the D^0 and D^+ which are of relevance in this analysis – is obtained from the following relation:

$$\sigma_c = \sigma_h \left(\frac{N_c}{N_h} \right) \left(\frac{A_{Si}^{\alpha_h}}{A_{Si}^{\alpha_c}} \right) \qquad (3a)$$

$$where\ N_c = N_D \frac{1}{Br_D} \frac{1}{P_D} \qquad (3b)$$

where:

σ_c is the required charm cross section

σ_h is the total hadronic photoproduction cross section

N_h is the number of hadronic events generated at the experimental target

N_D is the number of events with a charmed particle of a given species that were produced and decayed in the channel under consideration

Br_D is the branching ratio of decay channel considered

P_D is the mean number of charmed particles of a given species produced in one charm event ($\sum_D P_D = 2$)

A_{Si} is the atomic number of Silicon, the target material (= 28.1)

α_h is the A dependance of the cross section for hadronic events (= 0.920±0.002)

α_c is the A dependance of the cross section for charm events (= 0.94±0.02±0.03)

The total hadronic photoproduction cross section for the energy range considered is taken from reference [65]. More specifically, the total hadronic photoproduction cross section is 115μb (with a very small uncertainty, less than 1%). However, from that value we have to subtract the diffractive component, where our apparatus is insensitive (due to the trigger requirement of a track both above and below the median plane of the detector). This diffractive component is mainly due to diffractive ρ^0 and ω production (plus a small contribution from non-resonant $\pi^+\pi^-$). The cross section of the reaction $\gamma p \rightarrow p\pi^+\pi^-$ (which includes resonant ρ^0 and non-resonant $\pi^+\pi^-$) is 11$\pm$1μb whereas the cross section from the diffractive reaction $\gamma p \rightarrow p\omega$ is 1.2$\pm$0.3μb. Therefore the total non-diffractive hadronic cross section is taken to be 103$\pm$3μb. The uncertainty of 3% represents the systematic uncertainty of our knowledge of the diffractive part of the cross section and the acceptance of our trigger for such diffractive events.

N_h and N_D are taken from the trigger corrected photon energy spectra seen in Figure 56 and Figure 57. The branching ratios of $D^0 \rightarrow K\pi$ and $D^+ \rightarrow K\pi\pi$ were taken to be 4.20%$\pm$.04%$\pm$.04% and 9.1%$\pm$1.3%$\pm$.04% respectively [66].

The D^0 and D^+ contributions to the total cross section are calculated to be:

$$\sigma_{D^0} = \sigma_c\left(\frac{P_{D^0}}{2}\right) = 233\pm31\pm43nb$$

and

$$\sigma_{D^+} = \sigma_c\left(\frac{P_{D^+}}{2}\right) = 68\pm15\pm13nb$$

where the first error is statistical, calculated from the two contributing statistical errors (of N_c and N_h) added in quadrature, and the second error is systematic taken as the quadratic sum of the (independent) systematic contributions. The sources of errors are discussed at some length in the last section of this chapter.

The ratio of $\frac{D^+}{D^0}$ is thus:

$$\frac{D^+}{D^0} = 0.29\pm0.07\pm0.08$$

This is compatible with the prediction from the LUND model and γg fusion of our Monte Carlo of 0.41 ± 0.03 (see chapter 8).

To derive the total charm photoproduction cross section we need to estimate the contribution to the cross section of the remaining charmed particles, mainly the D_s and the Λ_c. For the contribution of the D_s we use the ratio of D_s/D derived from our Monte Carlo ($=0.38\pm0.15$, see Figure 26). However, for the Λ_c, there is a discrepancy between the predicted and the observed ratio. Λ_c production seems to be about three times higher than expected: the ratio of observed to predicted production ratios is $(\Lambda_c/D^0)_{exp}/(\Lambda_c/D^0)_{Lund} = 3.2\pm1.0$ and $(\Lambda_c/D^+)_{exp}/(\Lambda_c/D^+)_{Lund} = 3.0\pm0.8$ determined using a branching fraction $Br(\Lambda_c\rightarrow pK\pi) = 0.05$ [67]. To derive a number for the total Λ_c production we need to extrapolate over the low energy range where our apparatus is insensitive. The accuracy of this extrapolation is limited by the measured level of $D/\overline{D}$ asymmetry. If we assume a value for the branching fraction of $\Lambda_c\rightarrow pK\pi = 5\%$, we estimate a ratio of $(\Lambda_c+\overline{\Lambda_c})/(D+\overline{D})=0.15\pm0.06$.

We have therefore used the D_s/D ratio derived from our Monte Carlo and the *observed* Λ_c/D ratio to estimate sum of the contributions of the D^0 and the D^+ to the total charm cross section to be 0.78 ± 0.07. Therefore the total charm cross section is calculated to be:

$$\sigma_{c_{total}} = 393\pm46\pm84nb$$

Again the errors correspond to statistical and systematic contributions for the first and the second quoted value respectively, with the statistical error taken as the quadratic sum of the two statistical partial cross section errors and the systematic taken as the sum of the two respective errors since the analysis chain for both signals is the same and the systematic effects, therefore, are clearly not independent. This last value we shall use for the normalisation of the relative cross section which is discussed in the next section.

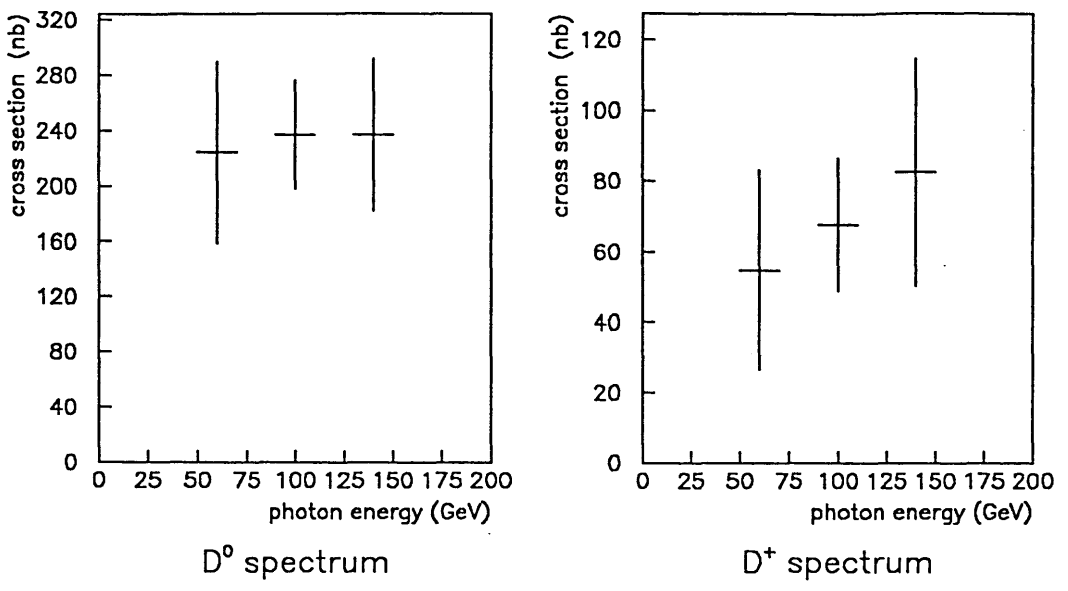


Figure 58: Photoproduction cross section going into D^0 and D^+

11.6 Calculation of the cross section variation with energy

To estimate the variation of the charm cross section with energy we proceed as follows: The D^0 and D^+ cross section variations with energy are compared with the normal hadronic event cross section which has a known energy dependance – in fact it is nearly constant at NA14 energies. More precisely if we follow the parametrisation of [44], the hadronic cross section increases by less than 2% in the energy range we are considering; it increases from 114.2μb at 60GeV to 115.5μb at 100GeV and to 116.1μb at 140GeV. Three energy intervals are used for this study, the granularity being dictated by the available statistics: 40 – 80GeV, 80 – 120GeV and 120 – 160 GeV.

By considering the intervals stated above for the cases of the D^0 and the D^+, we divide the histograms of Figure 56 and Figure 57, taking into account that the overall cross section as measured in the previous section is 233$\pm$31$\pm$43nb for the D^0 and 68$\pm$15$\pm$13nb for the D^+. The cross section variation obtained for the three energy intervals considered is then given in Figure 58. The errors shown are statistical. The systematic errors (discussed in the last section of this chapter) are common to all

three energy bands considered, with the exception of the systematic error due to the uncertainty in the radiation thickness of the radiation target and, to a small extent, the systematic error due to the uncertainty in the signal to noise ratio. Therefore the systematic error can be regarded as an overall uncertainty in the vertical scale (which is about 21% in the D^0 case and 22% in the D^+ case – see Table 22). Figure 59, finally, combines the D^0 and D^+ statistics in the same way as in the absolute cross section case, to give the overall variation of charm photoproduction cross section in the energy range considered, 40 to 160GeV. The error shown on the figure is again the statistical error. The uncertainty in the vertical scale reflects the systematic error in this measurement and is equal to 21%. The exact numbers for the cross section and its errors can be seen in Table 21.

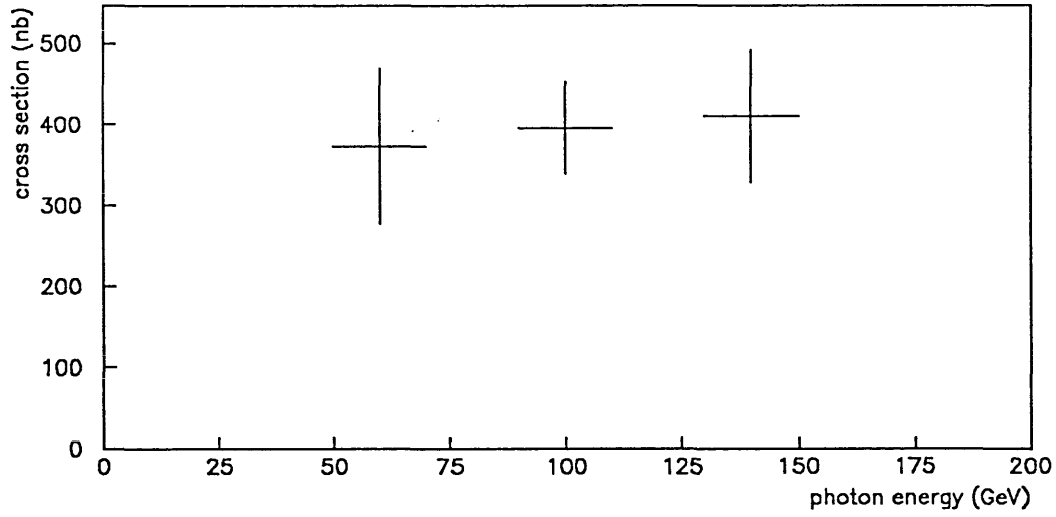


Figure 59: Photoproduction cross section variation with energy

Table 21: Cross section variation with energy

40- 80GeV:	σ_c = 373 ±95 ±78 nb
80-120GeV:	σ_c = 395 ±56 ±82 nb
120-160GeV:	σ_c = 410 ±81 ±87 nb

11.7 Comparison with theory

As we have discussed in chapter 8, the cross section variation with energy is sensitive to the charmed quark mass. This variation has been calculated to second order in QCD [29] and can be seen in Figure 21 on page 89, Figure 22 on page 90 and Figure 23 on page 90 for the three different charmed quark masses of 1.2GeV, 1.5GeV and 1.8GeV respectively. The theoretical uncertainty, which is calculated quite conservatively, is quite large but we can take the 'best estimate' value (corresponding to the middle curve in the three figures) to compare to our experimental measurements. These 'best estimate' curves, corresponding to the values of Λ = 260MeV and μ^2 = 10GeV^2, can be seen, together with the measurements obtained from this study, in Figure 60 on page 180. In this figure the statistical and systematic errors of the cross section measurement have been combined, so that there is no additional uncertainty in the vertical scale. As it can be seen, charmed quark masses of 1.2GeV or lower can be excluded and the data favours a charmed quark mass between $m_c = 1.5GeV$ and $m_c = 1.8GeV$. However, the level of the theoretical uncertainties does not allow for a more precise determination of m_c. The NA14/2 points are in good agreement with the rest of published measurements (see Figure 21 on page 89 for instance).

We should also note that another quantity which is sensitive to the charmed quark mass is the p_T distribution. This distribution was used by NA14/2 to derive an estimate for the charmed quark mass of $m_c = 1.58 \pm 0.08$ [68]. The cross section measurement obtained here is in good agreement with the above value for m_c.

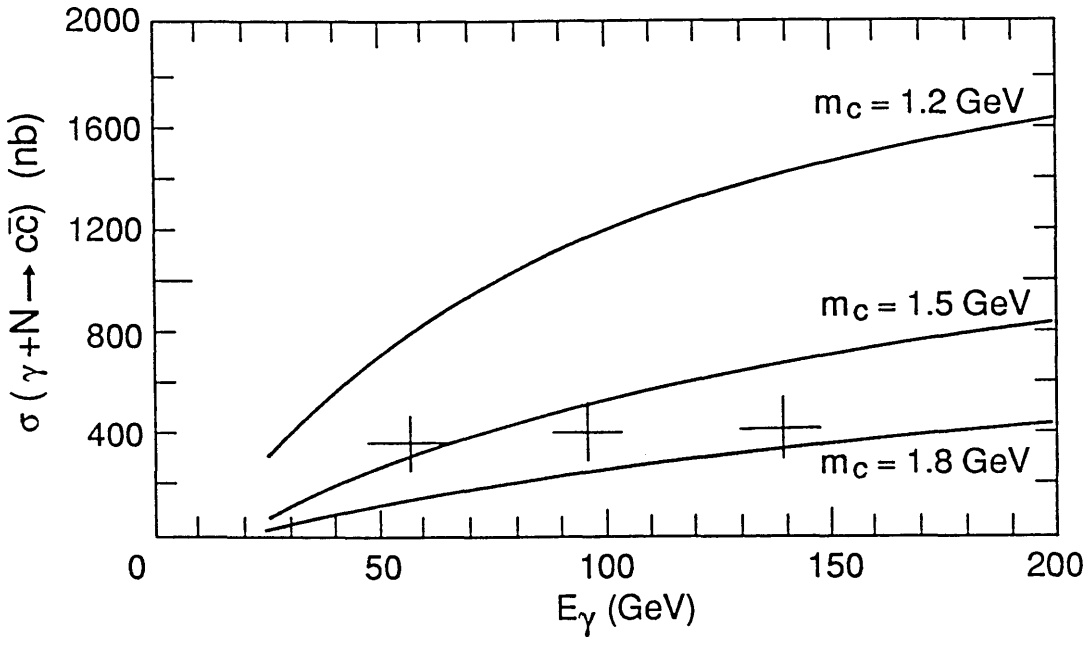


Figure 60: Comparison of cross section measurement with theory

11.8 Statistical and systematic error calculation

The statistical and systematic errors quoted in the previous section were calculated as follows:

For the statistical error we start with an uncertainty on the total number of events per energy interval for the signal region distribution and for the background region distribution, equal to the square root of the entries in that particular energy interval. To derive the error following the background subtraction we start from the definition of the error associated to a function, f, of two variables: f≡f(x,y):

$$(\sigma(f))^2 = \left(\frac{\partial f}{\partial x}\right)^2 \sigma_{xx}^2 + \left(\frac{\partial f}{\partial y}\right)^2 \sigma_{yy}^2 + 2\frac{\partial f}{\partial x}\frac{\partial f}{\partial y}\sigma_{xy}^2$$

$$\text{where } \sigma_{pq}^2 = \frac{1}{N}\sum_{i=1}^{N}(p_i - \bar{p})(q_i - \bar{q})$$

In our case the function, f is in fact the pure signal, S_i in a specific bin i related to the total signal, T_i, and the background, B_i, of the equivalent bins by the formula

$$S_i = T_i - B_i n\alpha$$

where n is the relative B to T normalisation, $n = \sum \frac{T_i}{B_i}$, and α is the overall background to total ratio related to the signal to noise ratio, α', by $\alpha(\alpha' + 1) = 1$. α is an external quantity given by the fit to the mass plot of the D. Now, according to the error formula and taking into account that no cross term exists since the two variables are unrelated we derive:

$$\sigma_S^2 = \sigma_{T_i}^2 + n^2\alpha^2\sigma_{B_i}^2$$

but $\sigma_T = \sqrt{T}$ the total number of events in that bin and $\sigma_B = \sqrt{B}$. The above formula is valid if we assume we know α, the overall background fraction with infinite precision. (Since we do not, the uncertainty in calculating α will be taken into account in our systematic error calculation.) Hence the overall error from the background subtraction step per bin is given by:

$$\sigma_{S_i}^2 = T_i + n^2\alpha^2 B_i$$

The analysis correction does not introduce an additional statistical error since it has been parametrised; therefore the associated uncertainty will be regarded as a systematic uncertainty; the same holds for the tagging to photon energy space transformation. The trigger correction has not been parametrised as such, ie no function was fitted on the Monte Carlo points, therefore it does contribute to the statistical error (by a small amount) as well as to the systematic uncertainty. The statistical error introduced this way is much smaller than the systematic uncertainty of the trigger acceptance correction. Another factor contributing in the increase of the statistical error is the uncertainty in the reference (hadronic) tagging answer histogram. By far the largest contributing factor in the whole analysis is the statistical error resulting from the background subtraction step; the rest of the statistical error contributions, in any case, could be reduced by running more events through the simulation or the tagging program.

Table 22: Factors contributing to cross section systematic error (in %)

factor	energy range (GeV)		
	40-80	80-120	120-160
$D^0 \rightarrow K\pi$			
S/N ratio - signal	±5.9	±4.8	±6.1
S/N ratio - fit	±0.5	±0.5	±0.5
analysis acceptance	±5	±5	±5
trigger acceptance	±5	±5	±5
radiator thickness	−2.3	+0.5	+1.6
	+2.8	−0.2	−1.9
hadronic cross section	±3	±3	±3
branching ratio	±13.5	±13.5	±13.5
fraction to total charm	±10	±10	±10
A dependence	−7	−7	−7
	+10	+10	+10
overall	+22.0	+21.6	+21.9
	−20.7	−20.3	−20.7
$D^+ \rightarrow K\pi\pi$			
S/N ratio - signal	±8.6	±7.9	±7.5
S/N ratio - fit	±0.5	±0.5	±0.5
analysis acceptance	±5	±5	±5
trigger acceptance	±5	±5	±5
radiator thickness	−0	+1.2	+0
	+1.6	−1.6	−2.5
hadronic cross section	±1	±1	±1
branching ratio	±14.3	±14.3	±14.3
fraction to total charm	±10	±10	±10
A dependence	−7	−7	−7
	+10	+10	+10
overall	+23.0	+22.7	+22.6
	−21.8	−21.6	−21.6

For the systematic error estimation we have taken the sources of uncertainties listed below, their effect on cross section being added in quadrature (since all these errors are uncorrelated) to give the total systematic error. These systematic error contributions, expressed as a percentage of the cross section in the three energy intervals considered, can be seen in Table 22.

- The variation of the overall signal to noise ratio: The signal to noise ratio has been calculated from a background fit of the D mass spectrum; the fit variables were therefore changed by one standard deviation and the result on the overall cross section noted. This was repeated for all the fit variables taking into account that the fit variables are usually highly correlated. This results in a negligible uncertainty. A more substantial uncertainty comes when we consider the statistical error of the determination of the signal to noise ratio. This depends on the overall D statistics (and not just on the subsample with a tagging answer) and, although it is the same for all three energy bands, since it is only applied to the background which is subtracted, affects energy bands with smaller statistics more strongly.
- The analysis acceptance parametrisation uncertainty: again we applied the same technique of varying the fit parameters by one standard deviation as in the signal to noise ratio case. However, the variation of the fit parameters give a negligible 1% variation in the final cross section. There is of course also the error due to the uncertainties in the Monte Carlo that derived the acceptance curve which was taken as 5%.
- Trigger acceptance parametrisation uncertainty: again a 5% error was introduced.
- The uncertainty in the thickness of effective radiator in the beam creation. This uncertainty was taken into account by varying the radiator thickness by 2.5% radiation lengths, obtaining a new tagging to photon spectrum transformation matrix and noting its effect on the cross section measurement. The effect of changing the radiator material thickness is to redistribute the tagging answer space events; its effect is, however, rather small.
- The uncertainty on the total hadronic cross section is taken to be 3%. This represents the uncertainty of the amount of diffractive cross section and of the acceptance of our trigger to diffractive events.
- The branching ratio uncertainty is taken from [66].
- The uncertainty in the A dependance comes primarily from the A dependance of the charm cross section. The error in the exponent is quoted to be $\pm 3.5\%$ ($0.94 \pm 0.2 \pm 0.3$) [37] leading to an asymmetric error in the A dependence of -7% to $+10\%$. The asymmetry in this error

comes from the fact that the A dependence of charm is constrained not to exceed the overall hadronic dependence ($=0.920\pm.002$)

- The final contribution comes from the uncertainty in the estimation of the mean number of D events per charm events; this depends on a variety of assumptions on the hadronisation scheme (which are justified from the agreement between Monte Carlo predictions and real data measurements on asymmetries and charmed particle ratios [68]), and an estimate of its overall uncertainty is 10% [69].

Conclusions

A search for Magnetic monopoles

Monopoles have recently become indispensable in many gauge theories, which endow them with extraordinarily high masses (of the order of 10^{16} GeV).

As a result of their high mass, monopoles could only be created at the very early stages of the universe, but the predictions of monopole abundance from Big Bang theories vary by enormous amounts. Astronomical observations, on the other hand, place stringent limits on their abundance. Widely accepted as the most reliable of these limits is the Parker bound ($< 6 \times 10^{-16} cm^{-2} s^{-1} sr^{-1}$). This makes monopoles rare objects indeed.

Monopole detectors have predominantly used either induction or ionization. The latter depend on assumptions about the monopole characteristics (such as their velocities and mass) whereas the former are only sensitive to magnetic charge, the only quantity of the monopole that it is assured. All inductive monopole detectors to date have used superconductivity as the means to achieving clear monopole signals. A lot of the recent work on superconductive detectors was triggered by a report of a monopole candidate event (in 1982) by such a superconductive detector of $10cm^2$ area.

The Imperial College detector is a superconductive type monopole detector. With the novel idea of the Window Frame loop, a loop strongly coupled to the superconducting shield of the detector, it has an active area of $1{,}800cm^2$, the biggest of its kind when it started collecting data in September 1984. The detector's design philosophy was to maximise sensitive area, but at the expense of the redundancy of information if an event occurred (such as two loops operating in coincidence, or generating a unique signal from the passage of a monopole). Therefore, it was intended to be a 'null' detector that would produce a useful upper bound on the monopole flux if it registered no monopole candidate events, but if an event was seen it would be difficult to be sure that it had no spurious cause.

The detector was operated for a year between September 84 and September 85. In terms of sensitivity and reliability, it has largely fulfilled its design specification, not a trivial achievement given the novel techniques that the detector incorporated. Internally generated mechanical shocks, caused by differential thermal expansion as the cryogenic helium level falls, pose a major design problem. Considerable care is also needed with the thermal environment of the SQUID sensors that were used to pick up any monopole signals.

The detector observed four events with no obvious cause, all of them in the large (window frame) detector loop. On closer examination, three of those events were found to be associated with a preceding mechanical shock, as far as 150 seconds prior to the event.

The remaining event has no explanation in terms of instrumental effects and it has survived close scrutiny from all aspects. However, the lack of coincidence and signal uniqueness in the detector design weakens our position to claim it as the genuine passage of a monopole.

A measurement of charmed particle photoproduction

The photoproduction of charm has several attractive features: it can be described in the framework of QCD due to the high mass of the charmed quark compared to the hadronisation scale of QCD, which makes perturbative expansion valid even for low particle momenta; there is only one structure function involved, that of the gluon, even in second order QCD calculations; finally the hadronisation process can be described simply in the framework of the Dual Parton Model.

Charmed particles decay weakly with lifetimes of the order of $10^{-12} - 10^{-13}$ seconds. This corresponds to a typical distance travelled by the charmed particle before decay of a few millimetres. A significant aid for the detection of charm is, therefore, a high precision tracking device, capable of resolving the vertex structure of an event involving charmed particles.

NA14/2, a fixed target experiment at CERN's SPS has obtained a large statistics sample of charmed event decays from a total of 17 million triggers taken over the period 1985-86. A silicon microstrip vertex detector, developed by NA14/2, was of critical importance for the enrichment of

charm in the data.

The high statistics of data collected required a substantial amount of computing power for event reconstruction and analysis. Emulator farms, like the 3081/E farm at CERN, a parallel processor facility that delivers considerable computing power, are useful tools for processing large amounts of High Energy Physics data. NA14/2 made use of this facility to process a significant fraction of its total statistics.

The complementary energy range of NA14/2 compared with other photoproduction experiments suggests the measurement of quantities depending on the incident photon energy. One of these measurements that contributes significantly to the world statistics is the measurement of the charm photoproduction cross section.

To measure such a quantity, one needs a thorough understanding of the detector and the underlying physics processes. NA14 has developed a Monte Carlo program that uses QCD photon-gluon fusion for the photoproduction process, the Dual Parton Model for the hadronisation scheme and the Lund model for fragmentation. This Monte Carlo has provided good agreement with the measured quantities.

Two channels have been investigated for the extraction of charm cross section: the $D^0 \rightarrow D\pi$ and $D^+ \rightarrow K\pi\pi$ channels. The respective cross sections obtained for the energy range $40-160$GeV are:

$$\sigma_{D^0} = 233 \pm 31 \pm 43 nb \quad \text{and} \quad \sigma_{D^+} = 68 \pm 15 \pm 13 nb.$$

The total charm photoproduction cross section estimated using the above two contributions and the Monte Carlo simulation is:

$$\sigma_{c_{total}} = 393 \pm 46 \pm 84 nb.$$

The energy dependence of the above cross sections has also been calculated. This calculation, when compared to the second order QCD predictions for the energy range considered, disfavours light charmed quark masses ($m_c = 1.2 GeV$ or below) and is in good agreement with the results of other experiments.

Acknowledgements

Oh, these acknowledgements! Seven years of one's life is a long time. There are a lot of people and a lot of things I would like to thank; some of them are going to be left out, inevitably, but they need not protest; I ve loved them all.

Firstly, I am deeply indebted to my high school, Athens College. It is a world-class establishment. Special thanks go to my physics teacher, Takis Papachristou, for stimulating my interest in Physics, and my A level teacher, Thanos Assimakis, for developing it.

My supervisor, David Websdale, was always very helpful and provided untiring tuition of my work. I would also like to thank my colleagues at the Imperial College Monopole Detector and especially Chris Guy for hours of stimulating conversation.

It is a great pleasure to thank everybody at NA14 for making it the best experiment in the world. Everybody in the collaboration deserves special thanks for their contribution (each one in their own way) to the success of the experiment. People that proved particularly helpful to me in my thesis work were Robert Barate, Daniel Treille, Patrick Roudeau and Tony Duane.

I would like to thank CERN for the excellent world-leading facility they provide.

This period has been quite intensive socially, and I would like to thank all the friends I have made and all the places and things I have enjoyed all these years. Special thanks go to Roger (Joe) Forty, Themos Tsikas, Dimitris Xenakis, Cromwell 194, Nick and Elika, Nick Manaras, the Band, my brother Vassilis, the City and Guilds Motor Club, Lyn, Mitch Wayne, the St. Pauli, Theodoris Geralis, Paris Sphicas, Andreas Gougas, the Goose, la Fete du Lac, the Farmhouse, Lils, Marco Cattaneo, PJD, Goodwood, the Spit, Whoofy, route de Napoleon and j,p,g&r. I would also like to thank the French Duane and the Imperial College traffic wardens for making my life exciting.

Financial assistance has been provided by my parents, the SERC, the IN2P3, and CERN.

Finally, I would like to express heartfelt thanks to my parents for their overwhelming support, support which has extended far beyond my university and postgraduate years. Without them, this thesis would simply not have been.

Bibliography

1. P.A.M. Dirac, Proc. R. Soc. London **A133** (1931) 60.
2. G. 't Hooft, Nucl. Phys. **B29** (1974) 276;
   A.M. Polyakov, Pis'ma Zh. Eksp. Teor. Fiz. **20** (1974) 430 [JETP Lett. **20** (1974)].
3. J. Preskill, *Magnetic Monopoles in Particle Physics Today*, proc. Inner Space/Outer Space, Fermilab, 1984.
4. S. Errede, *The Current Status of Monopole Search experiments* proc. ICOBAN '84 Conference, Park City, Utah, 1984.
5. M.S. Turner, *Monopoles, Cosmology and Astrophysics – Update 1985*, First Winter Conference, Aspen, 1985.
6. T.W.B. Kibble, J. Phys. **A9** (1976) 1387.
7. I. Wasserman, Cornell University Report, Ithaka, NY, 1984.
8. P. Langacker, et al., Phys. Rev. Lett. **45** (1980) 1.
9. A.H. Guth, Phys. Rev. **D23** (1981) 347.
10. K.A. Olive, and D. Seckel, proc. Monopole '83, J.L. Stone, Ed. (Plenum Press, New York, NY,1984).
11. G. Lazarides, and D. Shafi in *The Very Early Universe*, G. Gibbons, S. Hawking and S. Siklos, Eds. (Cambridge Univ. Press, Cambridge, England, 1983).
12. M.S. Turner, Phys. Lett. **115B** (1982) 95.
13. M.S. Turner, E.N. Parker and T. Bogdan, Phys. Rev. **D26** (1982) 1296.
14. S. Dimopoulos, S. Glashow, E. Purcell and F. Wilczek, Nature **298** (1982) 824.
15. K. Freese, and M.S. Turner, Phys. Lett. **123B** (1983) 293.

16. S. Ahlen, in *Magnetic Monopoles*, R.A. Carrigan and W.P. Trower, eds. (Plenum Press, New York, NY, 1983).

17. E.N. Parker, Astrophys. J. **160** (1980) 383.

18. E.E. Salpender, S. Shapiro and I. Wasserman, Phys. Rev. Lett. **49** (1982) 1114.

19. Y. Rephaeli and M.S. Turner, Phys. Lett. **121B** (1983) 115.

20. A.S. Goldhaber in *Magnetic monopoles*, R.A. Carrigan and W.P. Trower, eds. (Plenum Press, New York, NY,1983).

21. B. Cabrera, Phys. Rev. Lett. **48** (1982) 1378.

22. T. Van Duzer and C.W. Turner, *Principles of Superconductive Devices and Circuits*, (Elsevier, 1981).

23. M.J. Price, *The detection of cosmic magnetic monopoles using a room temperature coil*, preprint CERN/EF 83-2 (1983).

24. C.N. Guy, proc. Monopole '83, J.L. Stone, Ed. (Plenum press, New York, NY, 1984).

25. C.N. Guy and J G Park, J. Phys. **D17** (1985) 871.

26. A.D. Caplin et al., J.Phys. **E20** (1987) 850.

27. A.D. Caplin et al., Nature **321** (1985) 402.

28. C.N. Guy, Nature **325** (1985) 436.

29. R.K. Ellis and P. Nason, Nucl. Phys. **B312** (1989) 551.

30. M. Fontanaz, B. Pire and D. Schiff, Z. Phys. **C11** (1981) 211.

31. L.M. Jones and H.W. Wyld, Phys.Rev. **D17** (1978) 759.

32. M. Jonker et al., Phys. Lett. **B99** (1981) 265.

33. S.J. Brodsky, R. Blankenbecler and D. Sivers, Phys. Rep. **C23** (1976) 1.

34. J.V. Allaby et al, Phys. Lett. **B197** (1987) 281.

35. M. Diemoz, F. Feroni, E. Longo and G. Martinelli, preprint CERN-TH-4751/87 (1987).

36. J.C. Anjos et al., Phys. Rev. Lett. **62** (1988) 513.

37. M.D. Sokoloff et al., (the E691 collab.) Phys. Rev. Lett. **57** (1986) 3003.

38. D.O. Caldwell et al., Phys. Rev. Lett. **42** (1979) 553.

39. A. Capella and J. Tran Thanh Van, Phys. Lett. **B93** (1980) 146.

40. M. Primout,Thèse de doctorat d'Etat, Univ. de Paris-Sud (1987).

41. M.P. Alvarez et al, preprint CERN-EP/88-148 (1988).

42. T. Sjostrand, Lund preprint LUTP 85-10 (1985).

43. K. Abe et al. (SHF), Phys. Rev. **D33** (1986) 1.

44. M. Aquilar et al., Phys. Lett. **B204** (1988) 1.

45. M.P. Alvarez et al., Phys.Lett. **B246** (1990) 261.

46. R. Forty, PhD Thesis, Imperial College London, RALT-065 (1988).

47. M. Cattaneo, PhD Thesis, Imperial College London, RALT-048 (1987).

48. R. Barate et al., Nucl. Instrum. Methods **A235** (1985) 235.

49. G. Barber et al., Nucl. Instrum. Methods **A253** (1987) 530.

50. P. Roudeau, Private communication.

51. J. Dixon, PhD Thesis, Imperial College London.

52. R. Barate et al., *A program for heavy flavour production*, proposal CERN/SPSC/82-73 (1982).

53. J.C. Lassalle et al, *TRIDENT Track Finding and Vertex Identification for the Omega Particle Detector System*, CERN DD/EE/79-2 (1979).

54. C. Kraft, Thèse de doctorat, Univ. de Paris-Sud (1987).

55. R. Barate et al.,preprint CERN-EP/87-211(1987).

56. C. Magneville,Thèse de doctorat d'Etat, Univ. de Paris-Sud (1988).

57. P.F. Kunz et al, *The 3081/E processor*,CERN DD/83-3 (1983).

58. *CERN Program Library*, G. Benassi, Ed. (CERN, 1988).

59. M.P. Alvarez et al., Z. Phys **C47** (1990) 54.

60. G. Wormser, Thèse d'Etat, LAL Orsay, LAL 84-45 (1984).

61. M. Wayne, Private communication.

62. Yung-Su Tsai, Rev. Mod. Phys. **46** 4 (1974).

63. D.M. Websdale, private communication.

64. P. Druet, Thèse d'Etat, LAL Orsay, 1988.

65. *Total Cross sections for Reactions of High Energy Particles*, H. Shopper, Ed. (Landolt-Bornstein, New Series, Vol 12b, 1987) p. No. 345.

66. J. Adler et al., Phys. Rev. Lett. **60** (1988) 89.

67. M.P. Alvarez et al., Phys. Lett. **B246** (1990) 256.

68. M.P. Alvarez et al., *Study of Charm Photoproduction Mechanisms*, Z. Phys. C (to be published).

69. P. Roudeau, private communication.

Appendix A

WF signal amplitude probability: sides contribution

For a WF detector arrangement, it is possible to calculate analytically the signal amplitude probability distribution for tracks that intersect the sides of the detector but not its end-caps. This is done by considering the solid angle between two small areas ds and ds' and then integrating over all possible combinations of their positions.

Consider a cylinder of length H and radius R. The solid angle subtended by an infinitesimal area $ds = Rdzd\theta$ at a height z from a point P′ separated by an angle θ is

$$d\Omega = \frac{Asin(\theta/2)}{2(z^2+A^2)^{3/2}} Rdzd\theta \quad where \quad A = 2Rsin(\theta/2)$$

For an isotropic monopole flux of $Icm^{-2}sr^{-1}s^{-1}$ the number of monopoles entering through a small area $ds' = Rdz'd\theta'$ centered at P′ and leaving through ds is

$$N(\theta)dsds' = Id\Omega \frac{Asin(\theta/2)}{(z^2+A^2)^{3/2}} Ds'$$

so

$$N(\theta)dsds' = \frac{IA^4}{4(z^2+A^2)^2} d\theta d\theta' dzdz'$$

Integrating over dz, dz′ and dθ′ we get the number density probability distribution $N(\theta)d(\theta)$

$$N(\theta)d\theta = \pi RHIsin(\theta/2)\tan^{-1}\left(\frac{H}{2Rsin(\theta/2)}\right)d\theta$$

Now, the signal from a monopole passing through ds and ds′ is given by

$$S = \phi_0(\theta/\pi)$$

Therefore $P(S)d(S)$, the signal probability distribution, is given by

$$P(S)d(S) = c\frac{\pi}{\phi_0}N(S\pi/\phi_0)dS$$

where c is a normalization factor.

Appendix B

WF: 4π-averaged area calculation

The sensitive area of the detector is a cylinder of, say, length H and radius R. Its projected area at an angle θ to the horizontal is

$$A(\theta) = (2HR + \pi R^2)\cos(\theta)$$

independent of the azimuthal angle ϕ, since the detector is symmetric with respect to ϕ. The 4π averaged area, therefore, is

$$A = \frac{\left(\int 2LR\cos\theta\, d\phi \cos\theta\, d\theta + \int \pi R^2 d\phi \cos\theta\, d\theta\right)}{4\pi}$$

Thus

$$A = \frac{\pi(RH + R^2)}{2}$$

The actual detector dimensions are

$$L = 958cm \quad R = 114cm$$

So its physical 4π-averaged area is

$$A_{ph} = 1976cm^2$$

Obviously the *sensitive* 4π-averaged area will be less than that, its exact value depending on the threshold used and the WF signal amplitude probability distribution.