Determination of the semi-leptonic branching fraction of neutrino-induced charm hadrons using nuclear emulsion

Bart Van de Vyver
Promotor: Prof. Jacques Lemonne
Co-promotor: Dr. Pierre Vilain

April 4, 2002

Vakgroep Natuurkunde
Faculteit van de Wetenschappen
Vrije Universiteit Brussel

Proefschrift ingediend met het oog
op het behalen van de Wettelijke
Graad van Doctor in de Wetenschappen

CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

The common man marvels at uncommon things,
the wise man marvels at common things.
Confucius.

To claim that nuclear emulsion has a long and respectable history would be an understatement. The first recorded use of photographic plates in physics research was Henri Becquerel's accidental discovery of radioactivity in 1896. Ever since, photographic emulsions have played an important role in both nuclear and particle physics. Their heyday should probably be situated in the two decades after World War II, when they allowed the discovery of the pion as well as systematic studies of *strange* particles produced by cosmic rays. In essence, the physico-chemical processes responsible for the formation of an image in nuclear emulsions are not different from those in ordinary photography. Photographic films consist of an emulsion of silver halide crystals in a matrix material which is chiefly gelatin. As the shutter of the camera opens, photons with a wavelength in the visible spectrum are focused by the objective on the film. When encountering a silver halide crystal, their interaction may render the latter developable.

The prime difference for nuclear emulsions, also called nuclear track emulsions, is that they record not an outside scene in two dimensions but rather the passage of charged particles through the emulsion. Along their path, charged particles produce ionization which in turn renders the crystals developable. After development, this leads to a trail of minute silver grains, yielding a true three-dimensional image. The spatial resolution that can be obtained is determined by the size of the silver halide crystals, which varies between a tenth of a micron and one micron. The analysis of nuclear emulsion is performed by visually inspecting the charged particle tracks with a light microscope. All the research referred to above relied on the painstaking efforts of dedicated scanners. As the need for larger data samples grew, this became increasingly cumbersome and nuclear emulsion was gradually replaced by bubble chambers and electronic detectors even though neither of these could rival its exceptional spatial resolution.

Hybrid experiments try to combine the best of both worlds by employing nuclear emulsion in the region where the beam particles interact, followed by downstream electronic detectors which serve to guide the scanning and to aid in the reconstruction of kinematic quantities as well as in the particle identification. Even so, the time and effort required for the scanning ultimately limit the number of events that can be analyzed. To overcome this limitation, several attempts have been made to at least partially automate the process. The most successful to date have been the efforts of the Nagoya FKEN laboratory, which pioneered semi-automatic scanning techniques for the E531 experiment at Fermilab. In this case, the scanner is partially relieved of the burden of positioning the plate under the microscope and

noting down the result of what he sees. The image is read out on a video camera, reference marks are recognized electronically, and the microscope table is automatically directed to the appropriate position. The image is projected on a video monitor, and the scanner's measurements are recorded electronically. The real breakthrough was the advent of fully automatic scanning at the end of the eighties, significantly enhancing the physics reach of the E531 experiment. In this case, not only reference marks but also the tracks themselves are recognized electronically by dedicated hardware. For this purpose, it is more convenient for the plates to be oriented perpendicular to the beam. As the objective and camera move along the optical axis, they acquire successive images in depth, permitting the three-dimensional reconstruction of the particle tracks.

The E531 experiment performed a search for charm particles produced in neutrino interactions. As a by-product, the same data sample was used to set an upper limit on $\nu_\mu \rightarrow \nu_\tau$ neutrino oscillations, exploiting the capability of the experiment to identify the charged tau lepton that would be produced in a tau neutrino charged current interaction, and the absence of any such event. Neutrino oscillations are the process whereby the neutrino flavour – electron, mu, or tau — is different between the point of production and the point of interaction of the neutrino. They are possible only if neutrinos are massive, and if the mass eigenstates are different from the flavour, or weak interaction, eigenstates. The production and interaction of the neutrino probe the flavour eigenstate, whereas the propagation through space probes the mass eigenstate. In the Standard Model of particle physics, neutrinos are massless and no mixing between flavours can occur. In spite of its tremendous success in describing the constituents of matter and the interactions between them, the theoretical arguments to believe that the Standard Model is at least incomplete are compelling. Neutrino oscillations are one of several possible avenues to explore physics beyond the Standard Model. The phenomenology of neutrino mixing is governed by the squared difference between mass eigenvalues and by a unitary mixing matrix relating the mass and flavour states. If the oscillation can be described in terms of two families, then the mixing matrix can be summarized by a single number, the so called mixing angle.

The search for $\nu_\mu \rightarrow \nu_\tau$ neutrino oscillations acquired a renewed urgency at the end of the eighties when theoretical and cosmological arguments combined with experimental data indicated that the region of small mixing angles and moderate neutrino masses was of particular interest. The CHORUS experiment was proposed specifically to search for $\nu_\mu \rightarrow \nu_\tau$ oscillations through the appearance of ν_τ in a ν_μ beam, aiming to explore the domain of small mixing angles down to $\sin^2 2\theta_{\mu\tau} \sim 3 \times 10^{-4}$ for mass parameters $\Delta m^2 \sim 1$ eV^2. To detect the tiny signal of ν_τ interactions from the overwhelming background of ν_μ interactions, the experiment deployed nuclear emulsion on an unprecedented scale.

Today, it is becoming ever less likely that CHORUS would see a signal for neutrino oscillations. Over the past two decades, a range of experiments has been designed and operated, exploring different regions in the parameter space of mixing angles and squared mass differences. Several of these have not observed the phenomenon, excluding the region of parameters to which they were sensitive. On the other hand, the claims for observation from some of the other experiments are now firmly established. The long-standing solar neutrino problem has been addressed by a wide range of experiments, covering different parts of the energy spectrum, including the possibility to observe neutral current interactions in the case of the Sudbury Neutrino Observatory (SNO) experiment in Canada. The available data

strongly favour an interpretation in terms of neutrino oscillations, with little or no room for an astrophysical explanation. The second evidence for neutrino oscillations comes from the study of atmospheric neutrinos by the Kamiokande and Super-Kamiokande experiments in Japan. The key signal in these experiment is the variation in the muon neutrino to electron neutrino rate as a function of the incidence angle, corresponding to the length traversed through the earth. The different measurements performed in the Super-Kamiokande experiment, each of them a variation on the same theme, can all be interpreted in terms of $\nu_\mu \rightarrow \nu_\tau$ oscillations with maximal mixing and a mass squared difference of the order of 10^{-3} eV^2. The first data from the KEK to Kamioka (K2K) experiment tentatively confirms this interpretation, using an accelerator neutrino beam. Combining the data from solar neutrino and atmospheric neutrino experiments, a picture emerges of neutrino mixing between all three active neutrinos, with large to maximal mixing angles and mass squared differences at the 10^{-3} eV^2 level for atmospheric neutrinos and at the 10^{-5} eV^2 level or lower for solar neutrinos. Under such a scenario, the prospects for an observation of neutrino oscillation in the CHORUS experiment are bleak. However, it is worth mentioning that the Liquid Scintillator Neutrino Detector (LSND) at Los Alamos has reported evidence for $\nu_\mu \rightarrow \nu_e$ and $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ neutrino oscillations with small mixing, $\sin^2 2\theta_{e\mu} \sim 10^{-3}$, and a mass squared difference of the order of 1 eV^2. If confirmed, this would be difficult to reconcile with the interpretation of the solar and atmospheric neutrino data given above. Even though this would not significantly alter the prospects for an observation in the CHORUS experiment, it would make an exclusion limit theoretically more valuable. The LSND claim is being investigated by the Mini-BooNE experiment at Fermilab, which has recently started to take data.

When the CHORUS experiment was proposed, it was believed that $\sim$ 40,000 interactions could be automatically scanned in two years, in an 800 kg active emulsion target. The downstream electronic detectors were responsible for the prediction of the track positions at the exit face of the emulsion and for the reconstruction of kinematic quantities used to enhance the ν_τ fraction in the sample of events to be scanned. As it turns out, the CHORUS experiment itself provided a further boost to the development of automatic scanning techniques and the speed has improved by an order of magnitude every three years over the past decade. So far, $\sim$ 150,000 interactions have been automatically located and more are expected to follow. Ultimately, the sample of located neutrino interactions will be larger than planned by a factor of up to six. Originally, the scanning procedures were specifically optimized for the search of the tau lepton in ν_τ interactions. However, thanks to the increasing scanning capacity, the algorithms have gradually become more inclusive and a wide range of physics can now be addressed. In this work, the emphasis will be on the production of charmed hadrons in neutrino interactions. The charmed hadrons have a lifetime similar to that of the tau lepton, allowing their identification in emulsion through the observation of a decay topology at tens to hundreds of microns from the neutrino interaction point.

At energies sufficiently above the charm mass threshold, up to 10 % of charged current neutrino nucleon interactions give rise to the production of a charm quark, in turn leading to a charm hadron after fragmentation. The nucleon consists of up and down quarks, held together by gluons and accompanied by a sea of virtual quark-antiquark pairs of strange quarks, as well as up and down quarks. The strange and charm quark belong to the second generation, and scattering through the weak interaction off a strange quark typically leads to the production of a charm quark. However, just as the flavour eigenstates for neutrinos do not coincide with the mass eigenstates, so do the weak interaction eigenstates of quarks

differ from the mass eigenstates. In this context, the corresponding mixing matrix is referred to as the Cabibbo-Kobayashi-Maskawa matrix, usually abbreviated to the CKM matrix. By convention, it is expressed as operating on the down type quarks:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad .$$

As the modulus of the elements V_{ub}, V_{cb}, V_{td}, and V_{ts} is less than 0.04, the 2×2 submatrix for the first two generations is approximately unitary and can be written in terms of a single mixing angle, referred to as the Cabibbo angle. The modulus of the elements V_{ud} and V_{us} has been measured at the percent and permille level, respectively. Through unitarity, these measurements translate into a range for $|V_{cd}|$ from 0.219 to 0.225 at 90 % CL. As a consequence, charm quarks can be produced in neutrino interactions via the Cabibbo-suppressed transition $d \to c$, in addition to the Cabibbo-enhanced transition $s \to c$. The relative strength of the former process is proportional to $|V_{cd}|^2/|V_{cs}|^2 \sim 0.05$, but it is enhanced by the greater abundance of down quarks in the nucleon, compared to sea quarks. In total, the two processes give roughly equal contributions. The charm quark mass has been measured to lie in the range from 1.15 to 1.35 GeV/c^2, comparable to typical momentum transfers in neutrino interactions for the energy spectrum of the CHORUS beam. As a consequence, the charm mass threshold suppresses the cross section for charm production by a factor of about two.

In summary, the study of neutrino-induced charm production potentially provides valuable insights in three areas: the CKM matrix elements $|V_{cd}|$ and $|V_{cs}|$, the strange content of the nucleon, and the charm mass. However, the small cross section of neutrinos inevitably imposes the need for massive targets. Traditionally, this was most often realized in massive detectors consisting of a coarse calorimeter followed by an iron toroid spectrometer to measure sign and momentum of the muon. In such detectors, the direct detection of charm particles, with lifetimes of the order of 10^{-12} s and, correspondingly, path lengths of hundreds of microns, is impossible. Instead, the charm production process has been studied indirectly through the analysis of so-called opposite sign dimuon events. In such events, the first muon originates at the neutrino interaction vertex, the second one from the leptonic or semi-leptonic decay of the charm hadron. The average semi-leptonic branching fraction for the various charm hadrons that are produced, weighted by their production fractions, amounts to about 10 %. Unfortunately, the semi-leptonic branching fraction B_μ has been measured only once, based on the limited statistics of the E531 experiment: 121 events containing a charmed particle, leading to a statistical error of 10 % on the value for B_μ. This immediately translates into the uncertainty on the measurements of $|V_{cd}|$ and $|V_{cs}|$ from a study of opposite sign dimuon events. For the determination of $|V_{cs}|$, other methods are preferred anyhow. However, for measuring $|V_{cd}|$ neutrino-induced charm production does provide the most stringent experimental constraint.

The subject of this work is a determination of B_μ using the CHORUS sample of neutrino-induced charm events. So far, about 50,000 neutrino interactions have been fully analyzed for the presence of a charmed particle, an order of magnitude increase compared to the E531 experiment. Out of these, 1055 events are selected as candidate charm events using algorithms for which the selection purity was measured to be better than 90 %. In the near future, the full sample of 150,000 located neutrino interactions will be analyzed in the same manner

and further event location is underway in parallel. It is expected that the CHORUS experiment will eventually collect a sample of 3,000 to 4,000 neutrino interactions with a charm particle. For specific studies, samples of even higher purity can be obtained by performing manual checks. For the determination of the average semi-leptonic branching fraction, it is sufficient to statistically subtract the 10 % background contribution.

Chapter 1 will describe the theoretical framework describing neutrino-induced charm production. It will go into somewhat more detail than what is strictly required for the study of the semi-leptonic branching fraction, but the hope is that it may serve as a starting point for more advanced studies using the CHORUS data. The neutrino beam and the electronic detectors in the CHORUS experiment will be briefly reviewed in chapter 2. Far more attention will be devoted to the emulsion target and the automatic scanning technique in chapter 3, the centerpiece of this work. Automatic scanning involves computer-steered precision tables, microscope optics with high resolution, an image sensor, custom hardware for real-time image processing and computer infrastructure to store the results, possibly after on-line reconstruction. After an introduction to the fundamental concepts underlying each of these components, the microscopes at CERN will be used as an illustration of how they all fit together. The chapter concludes with a description of the procedures for alignment between plates and for the location of the neutrino interaction, referred to as the *scanback*. Once the interaction is located, the primary vertex is analysed using the net scan technique, in particular to search for the presence of any secondary vertices. The net scan technique is based on a volume scan in an area of 1.5×1.5 mm^2 around the estimated vertex position in eight consecutive plates – one plate upstream from the vertex plate, the vertex plate itself, and six plates downstream from the vertex plate. The scanning picks up all tracks within an angular acceptance of 400 mrad. The reconstruction of this type of data will be the subject of chapter 4, demonstrating how rigorous error propagation allows a systematic treatment, minimizing the number of arbitrary assumptions needed in the selection of decay topologies. This immediately leads to a sample of candidate charm events, for which the analysis is described in chapter 5. Manual checks were performed on a subset of the selected sample to evaluate the selection purity. The efficiency of the selection was determined on the basis of Monte Carlo simulation. Through the muon identification in the electronic detectors, events where the charm particle decays semi-leptonically into a muon are isolated from the overall sample of selected events. The number of such events corresponds to a measurement of the average semileptonic branching ratio. To conclude, the impact of this measurement on the determination of the CKM element $|V_{cd}|$ from opposite sign dimuon analysis will be discussed.

1. NEUTRINO-INDUCED CHARM PRODUCTION

1.1 Introduction

This chapter describes the theoretical framework of neutrino-induced charm production on nucleons. First, deeply inelastic neutrino interactions are treated in terms of structure functions, scalar coefficients indicating the relative importance of different contributions to the cross section. This formalism is a posteriori justified in the parton model of the nucleon, with its underlying QCD dynamics. Additional effects, such as Fermi motion and radiative corrections, are indicated without further elaboration. A more comprehensive discussion of the QCD description of deep inelastic scattering can be found in [1].

The most striking difference between the inclusive charged current cross section and the semi-inclusive charm production cross section is a threshold behaviour associated with the charm quark mass. One possible approach to take this effect into account is the 'slow-prescaling' prescription. We describe this prescription, pointing out its limitations and providing an alternative, formulated in the helicity formalism. This formalism is also used to incorporate a target mass correction, effects related to the non-zero mass of the nucleon on which the neutrino is scattered.

After this discussion of the ingredients – structure functions and threshold effects due to the charm mass – we will be able to formulate the entire cross section, at zeroth order in the strong coupling constant α_s. An extension of the formalism to higher orders of perturbation theory will require a discussion of the DGLAP evolution equations, as well as the additional Feynman diagrams at the parton level. Furthermore, the factorization theorem which links structure function, partonic subprocess and fragmentation will no longer be a simple multiplication of three factors. Rather, it will have to be formulated as a convolution in an appropriate integration variable.

Finally, this theoretical discussion will be complemented by a brief review of experimental results on neutrino-induced production of charmed particles. Most of these are from the observation of opposite-sign dimuon events in large calorimeter experiments. However, the normalization of their results critically depends on a single measurement in the E531 hybrid emulsion experiment.

The CHORUS experiment uses a beam of muon-neutrinos with small contaminations of anti-muon-neutrinos, electron-neutrinos and anti-electron-neutrinos. In the remainder of this chapter, any mention of neutrinos will refer to muon-neutrinos, unless explicitly stated otherwise.

1.2 The inclusive cross section for neutrino scattering

In this section, we will show how the most general formulation of neutrino nucleon scattering naturally leads to a description of the nucleon in terms of structure functions. The

derivation here will be relatively cursory, a more comprehensive treatment can be found in several introductory texts on quantum field theory [2, 3, 4].

The master equation for scattering cross sections in quantum field theory is based on Fermi's golden rule and reads

$$d\sigma \quad = \quad d\Gamma \, \frac{|\mathcal{M}|^2}{\Phi} \quad , \tag{1.1}$$

expressing the differential cross section $d\sigma$ in terms of three factors:

- $d\Gamma$: the density of final states per incident particle, a measure for the phase space volume of all final states considered;
- $|\mathcal{M}|^2$: the spin averaged square of the invariant amplitude describing the interaction dynamics, a measure for the strength of the process considered;
- Φ: the incident particle flux, a measure for the normalization of the initial state phase space volume.

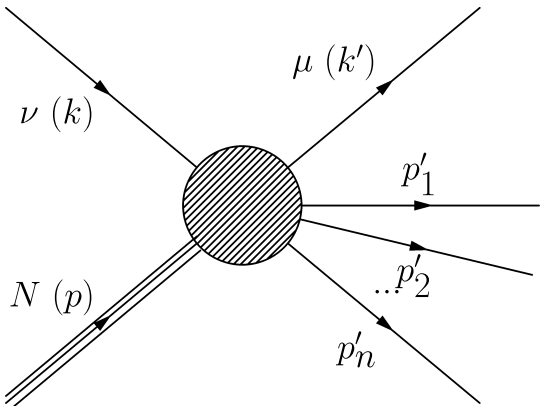


Fig. 1.1: Schematic representation of the charged current scattering off a nucleon, the final state being composed of a muon and n hadrons, each carrying four-momentum p'_i (i=1,2,..,n).

The interaction we wish to describe is the charged current scattering of neutrinos off a nucleon, schematically shown in figure 1.1. The final state consists of a muon and a hadronic shower, the details of which will not concern us since all hadronic final states are integrated over. Schematically, we consider the process

$$\nu(k) \ + \ N(p) \quad \rightarrow \quad \mu(k') \ + \ X(p') \ , \tag{1.2}$$

where k, p, k' and p' are the four-momenta we will consider. As indicated, X corresponds to a set of n final state hadrons, each of which carries four-momentum p'_j. The four-momentum of the hadronic final state is simply

$$p' \quad = \quad \sum_{j=1}^{n} p'_j \ . \tag{1.3}$$

The kinematics of this process is most conviently described in terms of the following Lorentz-invariant quantities:

$$Q^2 = -q^2 \qquad \text{the square of the four-momentum transfer,} \tag{1.4}$$

$$W^2 = p'^2 \qquad \text{the invariant mass squared of the hadronic final state,} \tag{1.5}$$

$$x = \frac{Q^2}{2p.q} \qquad \text{the Bjorken scaling variable, and} \tag{1.6}$$

$$y = \frac{p.q}{p.k} \qquad \text{the inelasticity,} \tag{1.7}$$

where $q = k - k' = p' - p$ is the four-momentum transfer.

The incident particle flux Φ is most easily expressed using the triangle function, defined as

$$\Delta(a,b,c) \equiv \sqrt{a^2 + b^2 + c^2 - 2(a.b + a.c + b.c)} \,. \tag{1.8}$$

In the scattering off a nucleon with mass M_N of a neutrino of mass m_ν and energy E_ν (corresponding to a centre-of-mass energy $\sqrt{s}$), the flux factor is given by

$$\Phi \quad = \quad 2\,\Delta(s, m_\nu^2, M_N^2) \quad = \quad 4\,M_N\,E_\nu \,, \tag{1.9}$$

where the first equality is general, the second specialized to the Lorentz frame in which the nucleon is at rest. Furthermore, we have neglected m_ν with respect to E_ν which is an extremely good approximation at all energies of interest.

The final state phase space is the product of all possible configurations for the muon and all possible configurations for the hadronic final state. Assuming the scattering is incoherent, the density of final states can be written as

$$d\Gamma = S \sum_n (2\pi)^4 \delta^4 \left(k - k' + p - \sum_{j=1}^{n} p'_j \right) \frac{d^3\vec{k'}}{(2\pi)^3 2E_\mu} \prod_{j=1}^{n} \frac{d^3\vec{p'}_j}{(2\pi)^3 2E'_j} \,, \tag{1.10}$$

with the sum running over final state multiplicities n and S a combinatorial factor, the product of $1/j!$ for each group of j identical particles in the final state. In the sum, the first factor expresses four-momentum conservation, the second the muon phase space volume and the third the phase space volume for the hadronic final state.

The physics of the actual interaction is embodied in the matrix element $\mathcal{M}$. In the Standard Model, charged current interactions are mediated through the exchange of a W boson, as shown in figure 1.2. Furthermore, the Standard Model considers the leptons to be fundamental, pointlike particles and their quantum numbers are fixed by the representation in the electroweak gauge group $SU(2)_L \otimes U(1)_Y$. The nucleon is not a fundamental entity, and we keep the hadronic current in its general form. We thus write the matrix element for a given spin state of the nucleon as

$$\mathcal{M} = \sqrt{2} G_F \underbrace{\bar{\mu}(k')\gamma^\alpha(1-\gamma_5)\nu(k)}\, L^\alpha \frac{1}{1 + Q^2/M_W^2} \underbrace{\langle X | J_\alpha | p, \sigma \rangle}\, W_\alpha \,. \tag{1.11}$$

where G_F is the Fermi constant, giving the strength of weak interactions, the first factor is the leptonic current, the second the W boson propagator and the third the hadronic current. The hadronic current indicates the strength for the transition from a nucleon with four-momentum p and helicity σ to a final state X, through the coupling to a W boson.

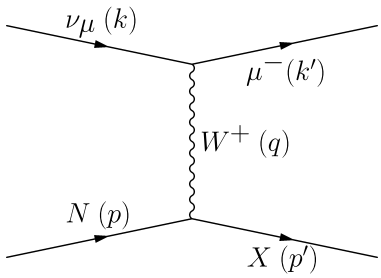


Fig. 1.2: The Feynman diagram for neutrino-nucleon scattering, with the interaction mediated by a W, one of the weak interaction bosons.

Squaring this matrix element, we obtain a covariant product of the leptonic tensor $L_{\alpha\beta}$ with a hadronic tensor $W_{\alpha\beta}$. The leptonic tensor can be written out explicitly:

$$\begin{aligned} L_{\alpha\beta} &= L_\alpha^\dagger . L_\beta & (1.12)\\ &= \bar{\nu}(k)\gamma_\alpha(1-\gamma_5)\mu(k').\bar{\mu}(k')\gamma_\beta(1-\gamma_5)\nu(k) & (1.13)\\ &= 8\left(k_\alpha k'_\beta + k_\beta k'_\alpha - g_{\alpha\beta}k.k' \mp i\epsilon_{\alpha\beta\gamma\delta}k^\gamma k'^\delta\right)\,, & (1.14)\end{aligned}$$

where the plus (minus) sign in the last term refers to neutrinos (antineutrinos). The hadronic tensor on the other hand incorporates all the details of the structure of the nucleon in terms of its constituent quarks and gluons. Without sacrificing any generality, we can perform a tensor decomposition, assuming only Lorentz invariance. The most general rank two tensor that can be built using the Lorentz vectors involved, is

$$\begin{aligned} W^{\alpha\beta} = & -g^{\alpha\beta}W_1 + \frac{p^\alpha p^\beta}{M^2}W_2 - \frac{i\epsilon^{\alpha\beta\gamma\delta}p_\gamma q_\delta}{2M^2}W_3 + \frac{q^\alpha q^\beta}{M^2}W_4 \\ & + \frac{p^\alpha q^\beta + p^\beta q^\alpha}{M^2}W_5 + \frac{i(p^\alpha q^\beta - p^\beta q^\alpha)}{2M^2}W_6\,, \end{aligned} \tag{1.15}$$

with the W_i Lorentz scalars depending only on Lorentz-invariant quantities such as q^2, $p.q$, the type of target off which the neutrino is scattered, whether a neutrino or an antineutrino is scattered, etc. The W_i are called the structure functions of a particular target, probed in a particular scattering reaction.

If we assume the charged lepton mass to be negligible, then the terms proportional to q^α will vanish after contraction with $L_{\alpha\beta}$. As this assumption is certainly satisfied for the case of a muon in processes with a Q^2 value of several GeV^2, we will drop the terms $W_{4,5,6}$ in the following. Note that this does not mean that the numerical value of $W_{4,5,6}$ is small, but merely that it is of the same order as $W_{1,2,3}$. After contraction, these terms acquire a factor m_l (*i.e.* m_μ for the case considered here) and can therefore be neglected with respect to the terms $W_{1,2,3}$.

To facilitate the physical interpretation of the remaining three W_i terms, a specific Lorentz frame needs to be considered. In particular, we will choose the coordinate system in which the nucleon is at rest, the incoming neutrino travels along the Z axis, and the outgoing muon makes an angle θ_μ with respect to the neutrino direction. In this coordinate system, we

find the following explicit expression for the Lorentz-invariant quantities defined in equations 1.4-1.7:

$$Q^2 = 2E_\nu(E_\mu - p_\mu \cos\theta_\mu) - m_\mu^2 \approx 4E_\nu p_\mu \sin^2\frac{\theta_\mu}{2} \quad , \tag{1.16}$$

$$W^2 = M_N^2 - Q^2 + 2M_N E_{had} \quad , \tag{1.17}$$

$$x = \frac{Q^2}{2M_N E_{had}} \quad , \tag{1.18}$$

$$y = \frac{E_{had}}{E_\nu} \quad , \tag{1.19}$$

where M_N is the nucleon mass, E_ν the energy of the incoming neutrino, and E_μ the energy of the outgoing muon. The quantity E_{had} is defined as $E_\nu - E_\mu$ and can be considered the energy of the hadronic final state X. The approximation for Q^2 is valid if $Q^2 \gg m_\mu^2$. The cross section can now be expressed as a function of the muon angle and energy:

$$\frac{d^2\sigma^{\overset{(-)}{\nu}N}}{d\Omega\, dE_\mu} = \frac{G_F^2 E_\mu^2}{2\pi^2\left(1+\frac{Q^2}{M_W^2}\right)^2}\left(2\sin^2\frac{\theta_\mu}{2}W_1^{\overset{(-)}{\nu}} + \cos^2\frac{\theta_\mu}{2}W_2^{\overset{(-)}{\nu}} \pm \frac{E_\nu+E_\mu}{M_N}\sin^2\frac{\theta_\mu}{2}W_3^{\overset{(-)}{\nu}}\right), \tag{1.20}$$

where the sign of the last term is $+$ for neutrinos, $-$ for antineutrinos and $d\Omega = 2\pi\sin\theta_\mu d\theta_\mu$. Alternatively, we can express the cross section as a function of the Lorentz-invariant quantities x and Q^2:

$$\frac{d^2\sigma^{\overset{(-)}{\nu}N}}{dx\, dQ^2} = \frac{G_F^2}{2\pi x\left(1+\frac{Q^2}{M_W^2}\right)^2}\left(y^2 x M_N W_1^{\overset{(-)}{\nu}} + \left(1-y-\frac{Mxy}{2E_\nu}\right)\nu W_2^{\overset{(-)}{\nu}} \pm y\left(1-\frac{y}{2}\right)x\nu W_3^{\overset{(-)}{\nu}}\right), \tag{1.21}$$

using the same convention for the sign of the last term. The terminology of deep inelastic scattering is largely determined by conventions, firmly established after three decades of theoretical and experimental work.

This concludes our discussion of the inclusive cross section for neutrino scattering off a nucleon. So far, all the arguments are entirely general and the only assumptions made are Lorentz invariance and the smallness of the charged lepton mass with respect to other scales in the process. This last assumption can easily be dropped by carrying the terms $W_{4,5,6}$ through in the entire calculation. However, this is only necessary when considering the scattering of tau-neutrinos; examples of such calculations can be found in [5, 6, 7].

1.3 The parton model

Consciously, we have kept the discussion in the previous section as general as possible and have not attempted to ascribe any underlying meaning to the structure functions W_i. They simply follow from the tensor decomposition of the hadronic current. The interpretation of their value and functional dependencies will require a model for the hadron. First of all, we

consider the 'structure functions' for a scalar, pointlike target of mass m:

$$2mw_1^{point}(\nu, Q^2) = \frac{Q^2}{2m\nu}\delta\left(1 - \frac{Q^2}{2m\nu}\right), \tag{1.22}$$

$$\nu w_2^{point}(\nu, Q^2) = \delta\left(1 - \frac{Q^2}{2m\nu}\right). \tag{1.23}$$

These 'structure functions' depend on a single, dimensionless quantity $\frac{2m\nu}{Q^2}$, expressing our assumption that the target has no substructure.

If we now assume the hadron to be built of pointlike constituents, then it should be possible to express the W_i for the hadron in terms of the corresponding w_i of its constituents. The only remaining dependency will be in the spin structure of the constituents: pointlike fermions and pointlike scalars have different 'structure functions'. Furthermore, as we reach larger Q^2 values, these constituents are individually resolved and the structure functions of the hadron come to depend on a single, dimensionless quantity as in the case of a pointlike target. To see this more clearly, we introduce the scaling structure functions

$$2xF_1(x, Q^2) = 2xM_N W_1(\nu, Q^2), \tag{1.24}$$

$$F_2(x, Q^2) = \nu W_2(\nu, Q^2), \tag{1.25}$$

$$xF_3(x, Q^2) = x\nu W_3(\nu, Q^2), \tag{1.26}$$

which merely reduce the W_i to dimensionless quantities F_i. In the limit of large Q^2, if the scattering is off individual, pointlike objects, we expect the scaling structure functions to depend on a single, dimensionless quantity, a property referred to as Bjorken scaling [8]. Historically, the experimental observation of scaling behaviour in the electron-nucleon scattering experiments at SLAC [9, 10] lead to the development of the parton model, since it implied pointlike entities inside the nucleon [11]. Several decades earlier, Rutherford's legendary experiments scattering α-particles off a gold foil had prompted him to a similar conclusion, revealing the substructure of the atom. Bjorken scaling can be formally expressed as

$$\lim_{\substack{x\ fixed \\ Q^2\to\infty}} F_i\left(x, Q^2\right) = F_i\left(x\right). \tag{1.27}$$

Strictly speaking, this relation only holds in the infinite momentum frame, a coordinate system in which all longitudinal momenta are boosted to infinity. In this reference frame, all interactions between quarks happen at timescales much longer than that of the scattering and the scattering off different partons can be treated incoherently. The other assumption for scaling to hold is that the hadron composition in terms of partons, is independent of the scale at which the hadron is probed. We will see later that this is no longer true once the scale dependence of the strong interactions is taken into account. For the time being, we will work in the infinite momentum frame and assume Bjorken scaling to hold exactly. The partons are taken to be on their mass shell, moving collinearly with the nucleon, and carrying a fraction $\xi = \frac{p_i}{p}$ of the nucleon's momentum.

In terms of these scaling structure functions, the differential cross section with respect to the Lorentz-invariant scaling variables x and y reads

$$\frac{d\sigma^{\nu(\bar{\nu})}}{dx\,dy} = \frac{G_F^2 M_N E_\nu}{\pi(1 + Q^2/M_W^2)^2}\left[\frac{y^2}{2}2xF_1 + \left(1 - y - \frac{M_N xy}{2E_\nu}\right)F_2 \pm y\left(1 - \frac{y}{2}\right)xF_3\right]. \tag{1.28}$$

We further assume that the partons building up the nucleon are spin 0 or spin 1/2 particles. Since the neutrino is a spin 1/2 particle, the total spin in the scattering process will be 0 or 1 for spin 1/2 partons and 1/2 for spin 0 partons. Because of the V-A structure of the weak current, scattering off spin 1/2 particles will select the left-handed component of particles, or the right-handed component of antiparticles, as illustrated in figure 1.3. The different allowed configurations only differ in the angular behaviour and we can summarize all possibilities as follows:

$$total\ spin\ 0 \quad : \quad \frac{d\sigma^{\nu q}}{d\cos\theta^*} = \frac{d\sigma^{\bar{\nu}\bar{q}}}{d\cos\theta^*} = \frac{G_F^2 M_N E_\nu}{\pi\left(1+Q^2/M_W^2\right)^2}\,, \tag{1.29}$$

$$total\ spin\ 1 \quad : \quad \frac{d\sigma^{\bar{\nu}q}}{d\cos\theta^*} = \frac{d\sigma^{\nu\bar{q}}}{d\cos\theta^*} = \frac{G_F^2 M_N E_\nu}{\pi\left(1+Q^2/M_W^2\right)^2}\left(\frac{1+\cos\theta^*}{2}\right)^2\,, \tag{1.30}$$

$$total\ spin\ \frac{1}{2} \quad : \quad \frac{d\sigma^{\nu k}}{d\cos\theta^*} = \frac{d\sigma^{\bar{\nu}k}}{d\cos\theta^*} = \frac{G_F^2 M_N E_\nu}{\pi\left(1+Q^2/M_W^2\right)^2} 2\left(\cos\frac{\theta^*}{2}\right)^2\,. \tag{1.31}$$

In these expressions, θ^* is the angle of the muon with respect to the incoming neutrino direction in the neutrino-parton centre-of-mass frame. A priori, one would expect the neutrino to scatter only off spin 1/2 particles, the quarks. However, any quark with transverse momentum has a component which in the infinite momentum frame appears to have spin 0 nature. In equations 1.29-1.31, q and $\bar{q}$ to states with spin 1/2 whereas k refers to states with zero spin. This convention will be kept in the following, with p used to denote any type of parton.

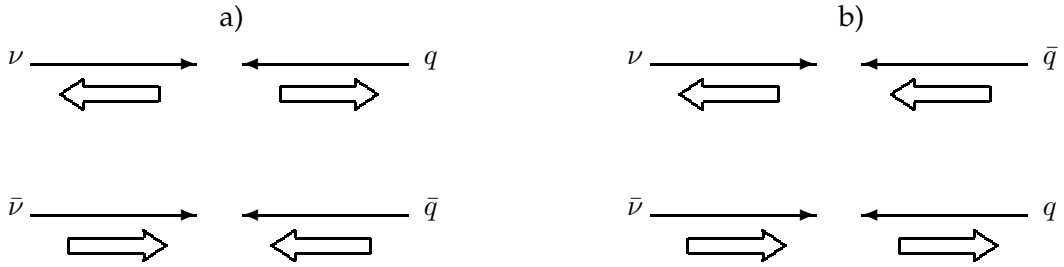


Fig. 1.3: Possible helicity configurations in the scattering between neutrinos or antineutrinos and quarks or antiquarks. a) corresponds to the case where both spins are opposite yielding total spin 0, realized for neutrino-quark or antineutrino-antiquark scattering. b) corresponds to the case where both spins are parallel yielding total spin 1, realized for neutrino-antiquark or antineutrino-quark scattering.

So far, all considerations have been entirely general, and would apply equally to any composite objects with constituents of either spin 0 or spin 1/2 nature. To make the connection between the expressions for the neutrino-parton scattering cross section and the structure functions appearing in the neutrino-hadron scattering cross section, we now introduce parton density functions. We define $p_H(x)dx$ to be the probability of finding in a hadron H a parton of type p carrying a fraction x to $x+dx$ of the hadron's momentum. Strictly speaking, these are not probability density functions but instead have the following normalization for

a proton:

$$\int_0^1 u(x)dx \;=\; 2 \;\text{ and }\; \int_0^1 d(x)dx \;=\; 1 \;,$$

expressing the fact that the proton is composed of two valence up quarks and one valence down quarks. For the neutron, the normalization is analogous, with the up and down quarks reversed. In terms of these parton density functions, the total cross section for (anti-)neutrino scattering off a target hadron T can be formulated as an incoherent sum over the cross section off different types of partons:

$$\frac{d^2\sigma^{\nu T}}{dx\,dy} = \frac{G_F^2 xs}{\pi\left(1+Q^2/M_W^2\right)^2}\left[q_T^{\nu}(x) + (1-y)^2\bar{q}_T^{\nu}(x) + 2(1-y)k_T^{\nu}(x)\right] \;, \tag{1.32}$$

$$\frac{d^2\sigma^{\bar{\nu} T}}{dx\,dy} = \frac{G_F^2 xs}{\pi\left(1+Q^2/M_W^2\right)^2}\left[q_T^{\bar{\nu}}(x) + (1-y)^2\bar{q}_T^{\bar{\nu}}(x) + 2(1-y)k_T^{\bar{\nu}}(x)\right] \;, \tag{1.33}$$

where we have been careful to indicate that the composition of the hadron T as seen in neutrino scattering – q_T^{ν}, $\bar{q}_T^{\nu}$ and k_T^{ν} – may be different from the composition as seen in antineutrino scattering – $q^{\bar{\nu}}$, $\bar{q}^{\bar{\nu}}$ and $k^{\bar{\nu}}$. Comparing 1.28 with 1.32 and 1.33, we can now interpret the scaling structure functions in terms of parton density functions:

$$2xF_1^{\nu(\bar{\nu})T} = 2\left[xq_T^{\nu(\bar{\nu})}(x) + x\bar{q}_T^{\nu(\bar{\nu})}(x)\right] \;, \tag{1.34}$$

$$F_2^{\nu(\bar{\nu})T} = 2\left[xq_T^{\nu(\bar{\nu})}(x) + x\bar{q}_T^{\nu(\bar{\nu})}(x) + 2xk_T^{\nu(\bar{\nu})}(x)\right] \;, \tag{1.35}$$

$$xF_3^{\nu(\bar{\nu})T} = 2\left[xq_T^{\nu(\bar{\nu})}(x) - x\bar{q}_T^{\nu(\bar{\nu})}(x)\right] \;. \tag{1.36}$$

Assuming that the target content of virtual bottom and top quarks is negligible, we can further specify $q_T^{\nu(\bar{\nu})}$ and $\bar{q}_T^{\nu(\bar{\nu})}$ using the Standard Model couplings of $W^{\pm}$ to the various quark flavours:

$$q_T^{\nu}(x) = d_T(x) + s_T(x) \qquad \bar{q}_T^{\nu}(x) = \bar{u}_T(x) + \bar{c}_T(x) \tag{1.37}$$

$$q_T^{\bar{\nu}}(x) = u_T(x) + c_T(x) \qquad \bar{q}_T^{\bar{\nu}}(x) = \bar{d}_T(x) + \bar{s}_T(x) \;. \tag{1.38}$$

In the following, only target protons or target neutrons will be considered, and we will take the proton as reference, dropping the T subscript. With this definition, the assumption of strong isospin invariance can be expressed as follows:

$$d(x) = d_p(x) = u_n(x) \qquad u(x) = u_p(x) = d_n(x) \tag{1.39}$$

$$\bar{d}(x) = \bar{d}_p(x) = \bar{u}_n(x) \qquad \bar{u}(x) = \bar{u}_p(x) = \bar{d}_n(x). \tag{1.40}$$

A priori, one expects the distributions of strangeness and charm to be identical between protons and neutrons, and we can make the following identifications:

$$s(x) = s_p(x) = s_n(x) \qquad c(x) = c_p(x) = c_n(x) \tag{1.41}$$

$$\bar{s}(x) = \bar{s}_p(x) = \bar{s}_n(x) \qquad \bar{c}(x) = \bar{c}_p(x) = \bar{c}_n(x) \;. \tag{1.42}$$

The fact that nucleons carry no net strangeness or charm can be expressed as

$$\int_0^1 s(x)\,dx = \int_0^1 \bar{s}(x)\,dx \qquad \int_0^1 c(x)\,dx = \int_0^1 \bar{c}(x)\,dx \;. \tag{1.43}$$

If their distributions are also taken to be equal, then this leads to the more stringent conditions

$$s(x) \;=\; \bar{s}(x) \qquad c(x) \;=\; \bar{c}(x)\,. \tag{1.44}$$

Using the above definitions and assumptions, we can now write the quark density functions as seen in (anti-)neutrino scattering off an isoscalar nucleon, defined as the average over protons and neutrons in equal contributions ($N = \frac{n+p}{2}$):

$$q_N^{\nu}(x) = \frac{1}{2}\left[u(x) + d(x) + 2s(x)\right] \qquad \bar{q}_N^{\nu}(x) = \frac{1}{2}\left[\bar{u}(x) + \bar{d}(x) + 2\bar{c}(x)\right] \tag{1.45}$$

$$q_N^{\bar{\nu}}(x) = \frac{1}{2}\left[u(x) + d(x) + 2c(x)\right] \qquad \bar{q}_N^{\bar{\nu}}(x) = \frac{1}{2}\left[\bar{u}(x) + \bar{d}(x) + 2\bar{s}(x)\right] \tag{1.46}$$

Finally, we define the valence quark distributions for u and d quarks:

$$u_V(x) = u(x) - \bar{u}(x) \qquad and \qquad d_V(x) = d(x) - \bar{d}(x)\,. \tag{1.47}$$

Rewritten in terms of quark flavour distribution functions, the scaling structure functions become

$$2xF_1^{\nu(\bar{\nu})N}(x) \;= x\left[u(x) + \bar{u}(x) + d(x) + \bar{d}(x) + s(x) + \bar{s}(x) + c(x) + \bar{c}(x)\right]\,, \tag{1.48}$$

$$F_2^{\nu(\bar{\nu})N}(x) \;= 2xF_1^{\nu(\bar{\nu})N} + 2x2k^{\nu(\bar{\nu})}(x)\,, \tag{1.49}$$

$$xF_3^{\nu N}(x) \;= x\left[u_V(x) + d_V(x) + 2s(x) - 2c(x)\right]\,, \tag{1.50}$$

$$xF_3^{\bar{\nu}N}(x) \;= x\left[u_V(x) + d_V(x) - 2s(x) + 2c(x)\right]\,. \tag{1.51}$$

For the case where the nucleon is entirely composed of spin $1/2$ constituents with vanishing transverse momentum, relation 1.49 reduces to $F_2 = 2xF_1$, a property known as the Callan-Gross relation [12]. Historically, the experimental observation of the Callan-Gross rule allowed the conclusion that the pointlike particles inside the nucleon had to be spin $1/2$ fermions.

1.4 Additional effects

As we will see later, the great success of the structure function formalism is its predictive power. Given a description of the scattering process at the parton level, the experimentally accessible structure functions can be inverted to quark density functions. The quark density functions in turn are independent of the scattering process: they can be compared between experiments using different probes. To apply this procedure, two effects have to be taken into account: nuclear effects and electroweak radiative corrections.

1.4.1 Nuclear effects

Structure functions defined for an isoscalar nucleon, a hypothetical entity corresponding to the average of a free neutron and a free proton, are particularly useful for the discussion of the interaction of neutrinos with nuclei. A large number of lepton-nucleon scattering experiments are performed using nuclear targets since they offer the best ratio of mass per unit volume, a consideration which is particularly important for neutrino experiments. When extracting the structure functions from data on nuclear targets, the following effects need to be taken into account.

Non-isoscalarity. Relatively few materials have equal amounts of protons and neutrons. Given the ratio of the number of protons to the number of neutrons, or the effective ratio in the case of composite materials, the measured cross sections can be reweighted to obtain the result for an isoscalar target.

Fermi motion. The quark density functions of nucleons are altered when the nucleon is bound inside a nucleus: the nucleons will have non-zero momentum, as well as nucleon-nucleon correlations, all related to the nuclear binding. These effects mainly influence the high x region and tables summarizing the effect can be found in the literature (see *e.g.* [13, 14]). It turns out that the result, a larger structure function for bound nucleons compared to that for free nucleons, is relatively independent of the nuclear inputs used.

Shadowing. For $x < 0.05 - 0.10$, the structure function for a nucleon bound in a nucleus is smaller than that for a free nucleon. For $x \sim 0.1 - 02$, it is a few percent larger. This can be explained by assuming that low x partons spread over a large longitudinal distance because of the uncertainty principle. Partons of different nucleons may thus overlap in space and fuse, reducing the density of low momentum partons and increasing that of higher momentum ones. The inner nucleons are, so to speak, shadowed by the surface ones.

The EMC effect. Between the low x region characterized by shadowing and the high x region characterized by Fermi motion, the ratio of the structure function for a bound nucleon to that for a free nucleon, decreases until it reaches a minimum around $x \sim 0.6$ and then begins to rise again. This behaviour was first observed in the European Muon Collaboration [15] and has since been dubbed the EMC effect. The theoretical models can broadly be classified into two categories. The first is based on conventional nuclear physics: because of the nuclear potential, the effective nucleon mass is reduced, which implies a shift of x to higher values and thus a softening of the quark distribution function. This shift is often accompanied by an increased density of virtual pions. The second class of models invokes an increase of the quark confinement size in nuclear matter, translating into a reduction of the quark Fermi momentum because of the uncertainty principle.

An excellent overview of the nuclear effects in structure functions, covering both experimental results and their theoretical interpretation, can be found in [16].

1.4.2 Electroweak radiative corrections

Radiative corrections for the strong interaction are implicitly taken into account for both the initial and the final state by the definition of structure functions and quark fragmentation, respectively. Radiative corrections for the electroweak interaction should be taken into account explicitly since they lead to a smearing of the kinematic variables, and affect the experimental acceptance. Calculations of these effects are available in the literature (see *e.g.* [17]), often accompanied by FORTRAN code that can be used for their implementation. Figure 1.4 shows examples of radiative corrections that can be incorporated by multiplying the cross section with appropriate weights, depending on the kinematic variables. The dominant effect is that the measured muon energy is lower than the muon energy considered in the calculation, due to the energy carried away by quasi-collinear photons.

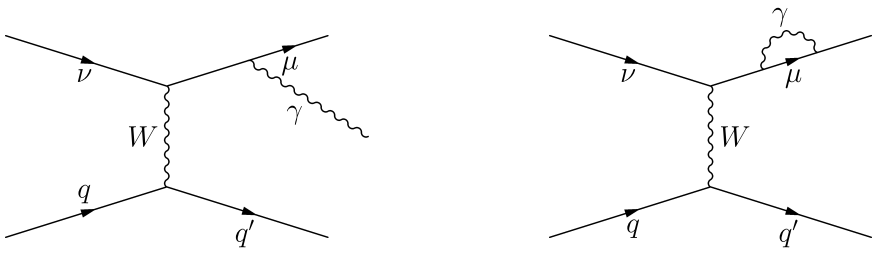


Fig. 1.4: Examples of electroweak radiative corrections to neutrino-quark scattering.

1.5 QCD and the factorization theorem

We have seen how the Bjorken scaling property of nucleon structure functions is explained in the parton model by assuming pointlike, spin 1/2 constituents that build up the nucleon. These constituents are the spin 1/2 quarks, originally postulated by Gell-Mann and Zweig to explain hadron spectroscopy [18, 19, 20]. The fundamental theory that describes the interactions between quarks is QCD, shorthand for quantum chromodynamics, the theory of strong interactions. The charge quantum number associated to QCD is colour, and the force carriers are massless gluons. The group structure of QCD is SU(3), a non-Abelian group. For non-Abelian theories, the gauge bosons carry charge and interact among themselves. The basic vertices of QCD are given in figure 1.5. The strength of strong interactions is given by the strong coupling constant α_s.

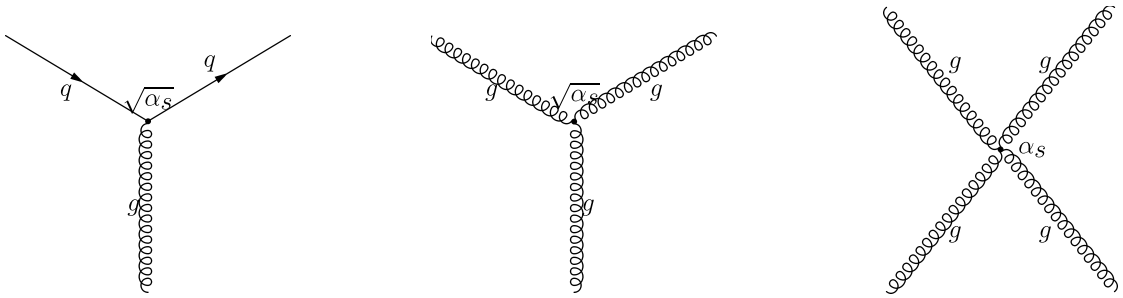


Fig. 1.5: The fundamental vertices in QCD, the theory of strong interactions.

The two most striking features of QCD are confinement and asymptotic freedom, and both are somehow related to the running of the coupling constant, a feature of all quantum field theories.

Confinement. All observed particles are colourless: as two colour charges move apart, their interaction grows strong enough to create colour-anticolour pairs from the vacuum. These colour charges recombine with the original charges to form colourless objects in a process called hadronization. Thus, confinement is related to the rising of the coupling constant as the distance between charges becomes large, or equivalently as the energy scale Q^2 goes to zero.

Asymptotic freedom. The interaction between quarks asymptotically tends to zero as the distance between quarks goes to zero. As a consequence, quarks in the nucleon can

be considered free particles. Thus, asymptotic freedom is related to the falling of the coupling constant as the distance between charges becomes small, or equivalently as the energy scale Q^2 goes to infinity[21].

The running of the coupling constant results from the renormalization of the theory, the procedure which removes infinities arising from calculations in perturbation theory by a redefinition of the parameters at each order. Grossly simplified, renormalization consists in cutting off infinite integrals at some scale μ_R, characteristic for the process being considered. In the case of neutrino-induced charm production, the two characteristic scales μ_R^2 are the momentum transfer squared Q^2 and the charm quark mass squared m_c^2. Various renormalization procedures differ in their choice of cut-off for each of the divergent integrals in the calculations. For example, the modified minimal subtraction scheme $\overline{MS}$ defines μ_R^2 to be the same for every divergent integral.

Obviously, observable quantities can depend neither on the renormalization scheme nor on the renormalization scale. Formally, the fact that any observable $\mathcal{M}$ is independent of the renormalization procedure can be expressed as

$$\mu \frac{d\mathcal{M}}{d\mu} = 0 \,, \tag{1.52}$$

the 'renormalization group equation'. In the mathematical physics literature on quantum field theories, it is often referred to as the Callan-Symanzik equation. The renormalization group is the set of all possible choices of renormalization. It is important to note that renormalization has to be performed at every order in perturbation theory, since the contributions to infinite quantities are different in each order. As an example, we give the renormalization group equation for the coupling constant α_s in the $\overline{MS}$ renormalization scheme:

$$\mu \frac{d\alpha_s(\mu^2)}{d\mu} = -\frac{\beta_1}{2\pi}\alpha_s^2(\mu^2) - \frac{\beta_2}{8\pi^2}\alpha_s^3(\mu^2) - \mathcal{O}\left(\alpha_s^4(\mu^2)\right) \,, \tag{1.53}$$
$$\text{with} \quad \beta_1 = (11N_c - 2n_f)/3 \qquad \text{and} \quad \beta_2 = (102N_c - 38n_f)/3 \,.$$

In this equation, n_f is the number of quark flavours and N_c the number of colours. The positive contributions in both terms arise from diagrams with gluon loops, the negative contributions from diagrams with quark loops. Different contributions are shown in figures 1.6, 1.7 and 1.8. To solve equation 1.53, it is necessary to introduce an integration

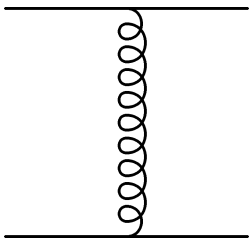


Fig. 1.6: The lowest order potential in any field theory.

constant $\alpha_s(\mu_0^2)$, the strong coupling strength at some reference momentum scale μ_0^2. This

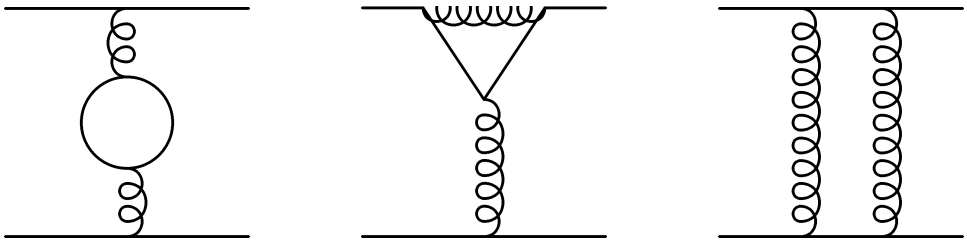


Fig. 1.7: Corrections to the potential in abelian, as well as non-abelian, field theory. The fermion loop diagram gives a negative contribution since the charge is effectively screened.

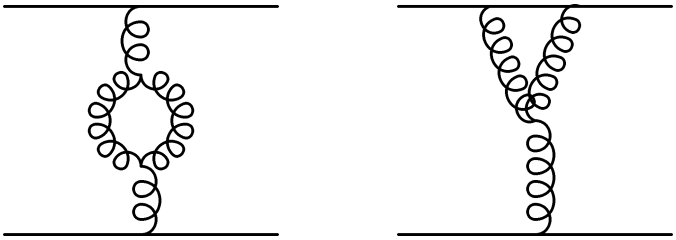


Fig. 1.8: Non-abelian corrections to the field theory potential. The gluon loop diagram gives a positive contribution since the charge has leaked out.

integration constant is the only fundamental constant in the theory and needs to be determined from experiment. Once it is fixed, the entire dynamics can be derived from the group structure of SU(3). To lowest order, the solution of equation 1.53 reads

$$\alpha_s(\mu^2) \;=\; \frac{\alpha_s(\mu_0^2)}{1+\frac{\beta_1}{4\pi}\alpha_s(\mu_0^2)\log\frac{\mu^2}{\mu_0^2}}\;. \tag{1.54}$$

Instead of $\alpha_s(\mu_0^2)$, one can introduce a dimensional parameter Λ, defined at lowest order as

$$\Lambda \;=\; \mu_0 \exp\frac{-2\pi}{\beta_1\alpha_s(\mu_0^2)}\;. \tag{1.55}$$

Keeping both terms in equation 1.53 and rewriting it in terms of Λ, we find

$$\frac{\alpha_s(\mu^2)}{4\pi} \;=\; \frac{1}{\beta_1\log\frac{\mu^2}{\Lambda^2}}\left[1-\frac{\beta_2\log\left(\log\frac{\mu^2}{\Lambda^2}\right)}{\beta_1^2\log\frac{\mu^2}{\Lambda^2}}+\mathcal{O}\left(\frac{1}{\log^2\frac{\mu^2}{\Lambda^2}}\right)\right]\;. \tag{1.56}$$

At momentum scales μ larger than Λ, the coupling constant becomes small and perturbation theory can be safely applied. On the other hand, for momentum scales smaller than Λ, the coupling constant becomes of order one and one has to resort to other techniques to perform calculations.

Returning to structure functions, QCD provides an explanation for the fact that Bjorken scaling is only approximate. Varying the scale Q^2 at which the nucleon is probed, the quark

distributions show a slight change. This can be attributed to the strong interaction giving rise to a 'sea' of virtual quark-antiquark pairs and gluons. Effectively, the gluon field of the quarks appears to have further substructure and this substructure is resolved as Q^2 increases. The dynamics of these scaling violations can be entirely predicted in QCD using the DGLAP equations: a set of integro-differential equations derived independently by Dokshitzer [22], Gribov and Lipatov [23, 24] and Altarelli and Parisi [25]. This set of equations is most easily written in terms of splitting functions $P_{ij}(\frac{x}{y}, t)$, defined as the probability for a parton of type j carrying a momentum fraction y to be resolved as a parton of type i carrying a momentum fraction x. Obviously, the splitting functions depend on the scale $t = \log Q^2$ at which the nucleon is probed. Some of the diagrams that contribute to the splitting functions are shown in figures 1.9 and 1.10, at leading order and next-to-leading order, respectively.

Fig. 1.9: Some of the diagrams contributing to the LO splitting functions in the DGLAP equations.

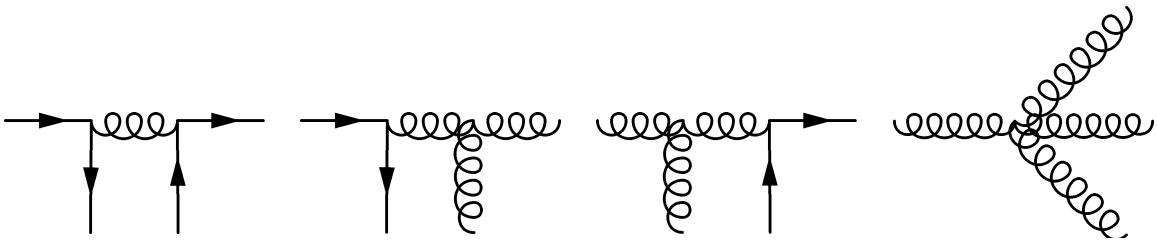


Fig. 1.10: Some of the diagrams contributing to the NLO splitting functions in the DGLAP equations.

To formulate the DGLAP equations, we define the following combinations of parton density functions:

$$\text{singlet} \quad : \quad q_S(x,t) = \sum_i \left[q_i(x,t) + \bar{q}_i(x,t)\right] \quad , \tag{1.57}$$

$$\text{non-singlet} \quad : \quad q_{NS}(x,t) = \sum_i \left[q_i(x,t) - \bar{q}_i(x,t)\right] \quad . \tag{1.58}$$

The terminology singlet and non-singlet refers to the colour content of these objects. In the $SU(3)$ space for colour, the former are scalar objects with a zero value for the colour quantum number. The latter are states in a multiplet, and the projection of their colour vector on a given axis may be non-zero. Using these definitions, as well as the gluon distribution $g(x,t)$

and the splitting functions, the DGLAP equations can be written as

$$\frac{dq_{NS}(x,t)}{dt} = \frac{\alpha_s(t)}{2\pi}\int_x^1 \frac{dy}{y} P_{qq}\left(\frac{x}{y},t\right)\; q_{NS}(x,t)\;, \tag{1.59}$$

$$\frac{dq_S(x,t)}{dt} = \frac{\alpha_s(t)}{2\pi}\int_x^1 \frac{dy}{y}\left[P_{qq}\left(\frac{x}{y},t\right)\; q_S(y,t) + 2n_f P_{qg}\left(\frac{x}{y},t\right)\; g(y,t)\right]\;, \tag{1.60}$$

$$\frac{dg(x,t)}{dt} = \frac{\alpha_s(t)}{2\pi}\int_x^1 \frac{dy}{y}\left[P_{gq}\left(\frac{x}{y},t\right)\; q_S(y,t) + P_{gg}\left(\frac{x}{y},t\right)\; g(y,t)\right]\;. \tag{1.61}$$

To make the connection between these equations and the scaling structure functions F_i defined before, we rewrite the structure functions in terms of the singlet and non-singlet quark distributions at the renormalization scale μ^2, introducing hard scattering coefficients $\mathcal{C}_i$:

$$2xF_1(x,\mu^2) = x\int_x^1 dy\left[\mathcal{C}_1^q\left(\frac{x}{y},\mu^2\right)\; q_S(y,\mu^2) + \mathcal{C}_1^g\left(\frac{x}{y},\mu^2\right)\; g(y,\mu^2)\right]\;, \tag{1.62}$$

$$F_2(x,\mu^2) = x\int_x^1 dy\left[\mathcal{C}_2^q\left(\frac{x}{y},\mu^2\right)\; q_S(y,\mu^2) + \mathcal{C}_2^g\left(\frac{x}{y},\mu^2\right)\; g(y,\mu^2)\right]\;, \tag{1.63}$$

$$xF_3(x,\mu^2) = x\int_x^1 dy\,\mathcal{C}_3^q\left(\frac{x}{y},\mu^2\right)\; q_{NS}(y,\mu^2)\;. \tag{1.64}$$

If we were to switch off the strong interaction, the hard scattering coefficients reduce to

$$\mathcal{C}_i^q \propto \delta(1-\frac{x}{y}) \qquad \text{and} \qquad \mathcal{C}_i^g = 0\;, \tag{1.65}$$

which implies

$$2xF_1(x,\mu^2) = F_2(x,\mu^2) = x\,q_S(x,\mu^2) \qquad \text{and} \qquad F_3(x,\mu^2) = x\,q_{NS}(x,\mu^2)\;. \tag{1.66}$$

In this formulation, we have implicitly used the 'factorization theorem', one of the theoretical foundations for QCD phenomenology. It relates the matrix element $W^{\alpha\beta}$ for the hadronic process to the matrix element $\omega_a^{\alpha\beta}$ for the partonic subprocess involving parton a. For lepton-nucleon scattering, the factorization theorem can be formulated as follows:

$$W^{\alpha\beta}(q,p) = \sum_a \int_\xi^1 \frac{d\xi'}{\xi'} f_a^N(\xi',\mu_F)\,\hat{\omega}_a^{\alpha\beta}(q,\xi' p,\mu_F,\alpha_s(\mu_R)) \tag{1.67}$$

$$= \sum_a f_a^N(\xi',\mu_F)\;\otimes\;\hat{\omega}_a^{\alpha\beta}(q,\xi' p,\mu_F,\alpha_s(\mu_R))\;, \tag{1.68}$$

where the sum runs over parton species a, μ_F and μ_R are the factorization and renormalization scales, respectively, and ξ is the momentum fraction of parton a in hadron N. Thus, the factorization theorem states that the parton distribution f_a^N and the hard scattering process $\hat{\omega}_a$ are independent. The hard scattering process at the parton level is dominated by ultraviolet effects and can be reliably calculated in perturbation theory, which yields the coefficients $\hat{\omega}_a$ for each parton species. The parton distributions are dominated by infrared effects and cannot be calculated in QCD perturbation theory. To extract them from experiment, the experimentally measurable hadronic matrix elements W are unfolded using the partonic matrix elements $\hat{\omega}$ as a 'kernel'. It is precisely this distinction between IR and UV effects which fixes the factorization scale μ_F: any effects at a scale below μ_F^2 are absorbed in the

parton distribution functions, whereas any effects at a scale above μ_F^2 enter the calculation of the hard scattering coefficients. Conventionally, μ_F and μ_R are taken to be equal.

Even though the theory of strong interactions predicts the evolution of structure functions as a function of Q^2, their shape at some reference scale remains to be fixed by experiment. Taking into account the evolution, results from different experiments at different energy scales can be reliably combined. Furthermore, the factorization theorem tells us that the parton distribution functions and the scattering process are independent, so experiments using different probes (neutrinos, electrons and muons) can be combined as well. Such combinations have been performed, and several compilations of experimental data are available in the literature. The most widely used are those from the MRS [26], CTEQ[1] [27] and GRV [28] groups. Furthermore, the PDFLIB library [29, 30] provides a convenient interface to access the parton distribution functions of different groups through a common FORTRAN interface.

1.6 The charm mass threshold and target mass effects

So far, the discussion of structure functions was limited to the inclusive case of scattering off light partons, and we have assumed throughout that any masses in the process can be neglected: these potentially include the lepton masses, the quark masses and the mass of the nucleon itself. Obviously, this approximation is no longer valid for the production of charm quarks, since m_c is of the order of 1.5 GeV/c^2 and the kinematic range we will consider corresponds to Q^2 values of the order of 10 GeV^2.

The first theoretical attempts to include the effect of the finite charm quark mass in neutrino-induced production date back to 1976 [31, 32]. This formalism is referred to as 'slow rescaling' and consists in a redefinition of the Bjorken scaling variable x. In particular, the prescription is to replace x by a slow rescaling variable z defined as

$$z \;=\; \frac{-q^2+m_c^2}{2p\cdot q} \;=\; x\left(\frac{-q^2+m_c^2}{-q^2}\right) \;=\; x+\frac{m_c^2}{2M_N E_\nu y}\,. \tag{1.69}$$

Experimental cross sections can be calculated assuming that the parton distributions are probed at z, whereas the experimentally observed x is actually lower. Their difference is the kinetic energy used to excite the charm quark onto its mass shell. Even though this model explains the qualitative behaviour of charm production, the threshold in the neutrino energy and the suppression of large x and small y, it also shows serious shortcomings. First of all, the slow rescaling prescription is only valid in the simplest version of the parton model and cannot be extended to higher orders of QCD perturbation theory. Furthermore, it does not take into account the effect of the finite mass of the nucleon. Since the nucleon mass is also of the order of 1 GeV/c^2, it is not justified to account for the charm mass only.

A more recent calculation [33, 34] offers a description of both effects by using a suitable reference frame, taking a reference axis along the direction of the W boson. This procedure is referred to as the 'helicity formalism' since this choice of basis provides a natural separation of the structure functions for different W boson helicities. To fix this coordinate system, we define polarization vectors for a W boson with four-momentum q in the target nucleon rest

[1] Coordinated Theoretical-Experimental project on QCD.

frame as

$$\epsilon_L(q) = \frac{1}{\sqrt{2}}(0,1,i,0)\,, \tag{1.70}$$

$$\epsilon_R(q) = \frac{1}{\sqrt{2}}(0,1,-i,0)\,, \tag{1.71}$$

$$\epsilon_l(q) = \frac{1}{\sqrt{-q^2}}(\sqrt{-q^2+\nu^2},0,0,\nu)\,, \tag{1.72}$$

corresponding to left-handed, right-handed and longitudinally polarized bosons, respectively. The longitudinally polarized case is often referred to as 'scalar'. With these definitions, we can replace the tensor decomposition of equation 1.15 with a helicity decomposition:

$$W^\alpha_\beta = \sum_{\lambda=L,R,l} \epsilon^*_\lambda(p,q)^\alpha \epsilon_\lambda(p,q)_\beta F_\lambda(x,q^2)\,, \tag{1.73}$$

where the dependence of ϵ on the coordinate system is indicated through the argument p, the four-momentum of the target nucleon. The helicity decomposition can be inverted to obtain a definition of the helicity structure functions F_λ:

$$F_{\lambda=L,R,l} = \epsilon^*_\lambda(p,q)_\alpha\, W^\alpha_\beta\, \epsilon_\lambda(p,q)^\beta\,. \tag{1.74}$$

The correspondence between helicity structure functions and scaling structure functions can be worked out to obtain

$$F_L = F_1 + \frac{1}{2}\sqrt{1+\frac{Q^2}{\nu^2}}\,F_3\,, \tag{1.75}$$

$$F_R = F_1 - \frac{1}{2}\sqrt{1+\frac{Q^2}{\nu^2}}\,F_3\,, \tag{1.76}$$

$$F_l = -F_1 + \left(1+\frac{Q^2}{\nu^2}\right)\frac{1}{2x}\,F_2\,. \tag{1.77}$$

Note that the zero target mass limit $M^2/Q^2 \to 0$ is equivalent to $Q^2/\nu^2 = 4x^2M^2/Q^2 \to 0$, which implies $F_L = F_1 + F_3/2$, $F_R = F_1 - F_3/2$ and $F_l = -F_1 + F_2/(2x)$. To discuss the cross section in terms of these helicity structure functions, we return to the matrix element defined in equation 1.11 and rewrite it using helicity indices n, m instead of Lorentz indices α, β:

$$\mathcal{M} = \sqrt{2}\,G_F\,j^n(Q^2)\,\frac{d^1(\psi)^m_n}{1+Q^2/M_W^2}\,\langle X|J_m|p,\sigma\rangle\,. \tag{1.78}$$

In this expression, j and J are the leptonic and hadronic currents, respectively, and $d^1(\psi)$ is a spin-1 rotation matrix specifying the relative orientation of the leptonic and hadronic vertex, as illustrated in figure 1.11. Considering the 'rest frame' of the exchanged W boson, the coordinate axis can be taken parallel to the hadron momenta or to the lepton momenta, and the hyperbolic cosine $\cosh\psi$ describes precisely the transformation between the two. Written out, the $d^1(\psi)$ matrix is

$$d^1(\psi) = \begin{pmatrix} \frac{1+\cosh\psi}{2} & \frac{-\sinh\psi}{\sqrt{2}} & \frac{1-\cosh\psi}{2} \\ \frac{-\sinh\psi}{\sqrt{2}} & \cosh\psi & \frac{\sinh\psi}{\sqrt{2}} \\ \frac{1-\cosh\psi}{2} & \frac{\sinh\psi}{\sqrt{2}} & \frac{1+\cosh\psi}{2} \end{pmatrix}\,. \tag{1.79}$$

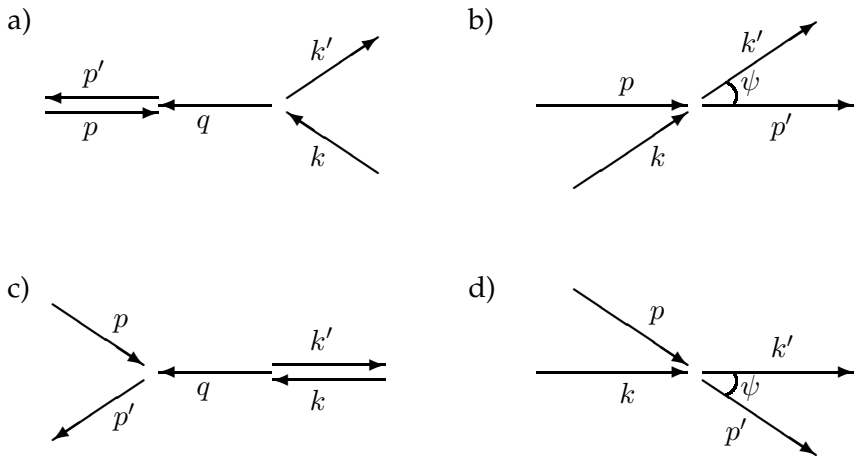


Fig. 1.11: The coordinate systems used in the helicity formalism. In a) the reference axis is taken to be in the plane defined by the W and hadron momenta. In c) the reference axis is taken to be in the plane defined by the W and lepton momenta. After a Lorentz boost to the frame where the W is at rest, a) transforms to b) whereas c) transforms to d).

The hyperbolic cosine itself can be obtained from the formulae

$$\cosh\psi = \frac{2\cdot(k+k')}{\Delta(-Q^2,p^2,p'^2)} \tag{1.80}$$

$$= \frac{E_\nu + E_\mu}{\sqrt{Q^2+\nu^2}} = \frac{\eta^2 M_N^2 - Q^2 + 2\eta(s-M_N^2)}{\eta^2 M_N^2 + Q^2} \tag{1.81}$$

$$\overset{M_N\to 0}{\to} \frac{2-y}{y}\,, \tag{1.82}$$

where the first equality is general, the remaining relations are given in the laboratory reference frame. In these expressions, η is a generalization of the Bjorken scaling variable x for the case where the target nucleon mass is taken into account. It is defined through the implicit equation

$$2\,q\cdot p = \frac{Q^2}{\eta} - \eta\, M_N^2\,. \tag{1.83}$$

To show the correspondence between x and η more clearly, equation 1.83 can be rewritten as

$$\frac{1}{x} = \frac{1}{\eta} - \eta\,\frac{M_N^2}{Q^2}\,. \tag{1.84}$$

It is clear that x and η become identical in the limit $M_N^2/Q^2 \to 0$. The variable η is sometimes referred to as the Nachtmann variable [35]; it also applies to the inclusive neutrino cross section in the absence of heavy quarks.

In equation 1.28, we showed the decomposition of the inclusive cross section for deeply inelastic scattering in terms of the scaling structure functions F_i. A similar derivation gives this cross section in terms of the helicity structure functions:

$$\frac{d\sigma^{\nu(\bar{\nu})}}{dxdy} = \frac{yQ^2}{4\pi(1+Q^2/M_W^2)^2}\left[(1+\cosh^2\psi)\frac{F_R+F_L}{2} + \sinh^2\psi F_l \mp \cosh\psi(F_R - F_L)\right] , \tag{1.85}$$

where the plus (minus) sign in the last term refers to the case of neutrinos (antineutrinos).

To understand what we have gained by these transformations, we have to return to the factorization theorem, equation 1.68. The point to note is that the tensor decomposition of $W^{\alpha\beta}$ and the tensor decomposition of $\hat{w}^{\alpha\beta}$ no longer coincide when masses are introduced. The hadronic tensor $W^{\alpha\beta}$ is decomposed in terms of linear combinations of the four-momentum transfer q and the hadronic four-momentum p, whereas the partonic tensor $\hat{w}^{\alpha\beta}$ is decomposed in terms of linear combinations of the four-momentum transfer q and the partonic four-momentum $\xi' p$. Neglecting both the nucleon mass and all quark masses involved, there is a direct correspondence between W_i and $\hat{w}_i$: the hadronic structure functions are obtained by simple multiplication of $\hat{w}_i$ with the parton structure functions, summing over all partons. This situation changes dramatically as masses are introduced since the p and $\xi' p$ will no longer be parallel. Therefore, the W_i will be linear combinations of all $\hat{w}_j$, including terms for which $j \neq i$. In general, the partonic four-momentum can be written as $\hat{p}^\mu = \zeta_p p^\mu + \zeta_q q^\mu$. The zero-mass limit is then recovered through the replacements $\zeta_p \to \xi'$ and $\zeta_q \to 0$.

On the other hand, in the helicity formalism, tensors are decomposed with respect to the W boson polarization vectors, introduced in equations 1.70–1.72. Analogous to the hadron decomposition given in 1.74, we can define parton helicity structure functions

$$\hat{w}_{\lambda=L,R,l} \quad = \quad \epsilon^*_\lambda(\hat{p},q)_\alpha \; \hat{w}^\alpha_\beta \; \epsilon_\lambda(\hat{p},q)^\beta \; . \tag{1.86}$$

The simplification of the helicity approach now follows from the fact that the two sets of polarization vectors $\epsilon_\lambda(\hat{p},q)$ and $\epsilon_\lambda(p,q)$, as opposed to the two sets of Lorentz invariants in $\hat{p}$ and p, are identical and remain identical when masses are introduced: the direction of the polarization axis is given by three-momenta, not four-momenta. In the infinite momentum frame, the three-momenta $\hat{p}$ and p are collinear by construction. Thus, the helicity formalism retains the one-to-one correspondence between hadronic and partonic structure functions, even when quark and nucleon masses are introduced.

Furthermore, because of the chiral couplings of electroweak theory, the parton-level helicity amplitudes ω_λ exhibit a simple symmetry and structure. In particular, for the W-exchange process encountered in charged current neutrino scattering, only left-handed chiral couplings are involved. Working out the parton helicity structure functions for this process, we obtain

$$\omega_R = |V_{12}|^2 \; \frac{Q^2 + m_1^2 + m_2^2 - \Delta(-Q^2, m_1^2, m_2^2)}{\Delta(-Q^2, m_1^2, m_2^2)} \; \delta\left(\frac{\xi}{\chi} - 1\right) , \tag{1.87}$$

$$\omega_L = |V_{12}|^2 \; \frac{Q^2 + m_1^2 + m_2^2 + \Delta(-Q^2, m_1^2, m_2^2)}{\Delta(-Q^2, m_1^2, m_2^2)} \; \delta\left(\frac{\xi}{\chi} - 1\right) , \tag{1.88}$$

$$\omega_l = |V_{12}|^2 \; \frac{\frac{(m_2^2 - m_1^2)^2}{Q^2} + m_1^2 + m_2^2}{\Delta(-Q^2, m_1^2, m_2^2)} \; \delta\left(\frac{\xi}{\chi} - 1\right) , \tag{1.89}$$

where m_1 and m_2 are the masses for the initial state and final state quarks, respectively and V_{12} is the CKM matrix element for the transition between quark flavours 1 and 2. Finally, χ is a further generalization of the Bjorken scaling variable x, now for the inclusion of quark masses. In terms of the previously introduced Nachtmann variable η, it is defined as

$$\chi \;=\; \eta\,\frac{Q^2+m_2^2-m_1^2+\Delta(-Q^2,m_1^2,m_2^2)}{2Q^2}\;. \tag{1.90}$$

To close the discussion of quark and nucleon mass effects, we show the relation between χ and the slow-rescaling variable z mentioned in the beginning of this section. The variable χ takes into account the effects of the initial state quark mass m_1, the final state quark mass m_2 and the target nucleon mass M_N. On the other hand, z only accounts for the final state quark mass m_2, but is recovered through the obvious approximations:

$$\chi \overset{m_1\to 0}{\to} \eta\left(1+\frac{m_2^2}{Q^2}\right) \overset{M_N\to 0}{\to} x\left(1+\frac{m_2^2}{Q^2}\right) \;=\; z\;. \tag{1.91}$$

1.7 Parton-level contributions to charm production

In the previous section, we have seen how quark and target nucleon masses can be introduced in the inclusive cross section for neutrino-nucleon scattering to prepare the discussion of the semi-inclusive charm production cross section. The leading order contribution to this process comes from the Born-level diagram of figure 1.2, specialized to the case where the final state contains a charm quark. Neglecting any intrinsic charm component in the nucleon, the only contributions are from the excitation of d or s quarks. The former is Cabibbo-suppressed, whereas the latter is suppressed by the smallness of the strange sea in the nucleon. Both contributions are schematically shown in figure 1.12. The sum of both is often referred to as 'flavour excitation' to distinguish it from 'flavour creation' which will be discussed below.

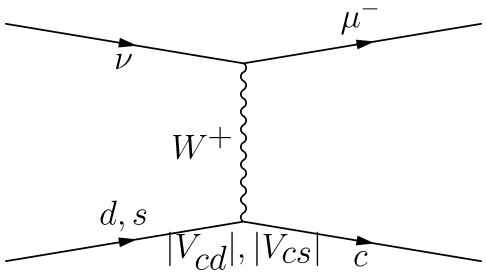


Fig. 1.12: Feynman diagram of the Born-level contributions to charm production in parton language.

To simplify the discussion, we factor out the leptonic part of

$$\nu(k)+N(p)\to\mu(k')+X(p')\;, \tag{1.92}$$

shown in figure 1.2, and restrict our attention to the process

$$W(q)+N(p)\to c(p_c)+X(p_X)\;, \tag{1.93}$$

where X no longer refers to the inclusive hadronic final state, but rather to this final state excluding the charm quark. In this process, due to the competing effects of Cabibbo suppression and strange sea suppression, the contributions of d and s quarks turn out to be almost equal. Since the strange sea contribution is large in spite of the smallness of the corresponding quark distribution $s(x)$, contributions from other processes must be considered as well. In particular, any gluon-initiated contribution will be strongly enhanced by the size of the gluon distribution $g(x)$ which is an order of magnitude larger than $s(x)$. This may be large enough to compensate for an additional factor of α_s, a consequence of the fact that there is no zeroth order diagram initiated by a gluon.

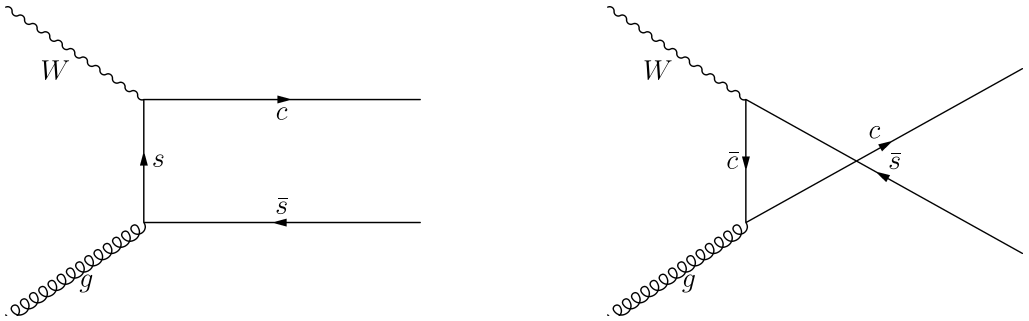


Fig. 1.13: Feynman diagrams for the boson-gluon fusion process.

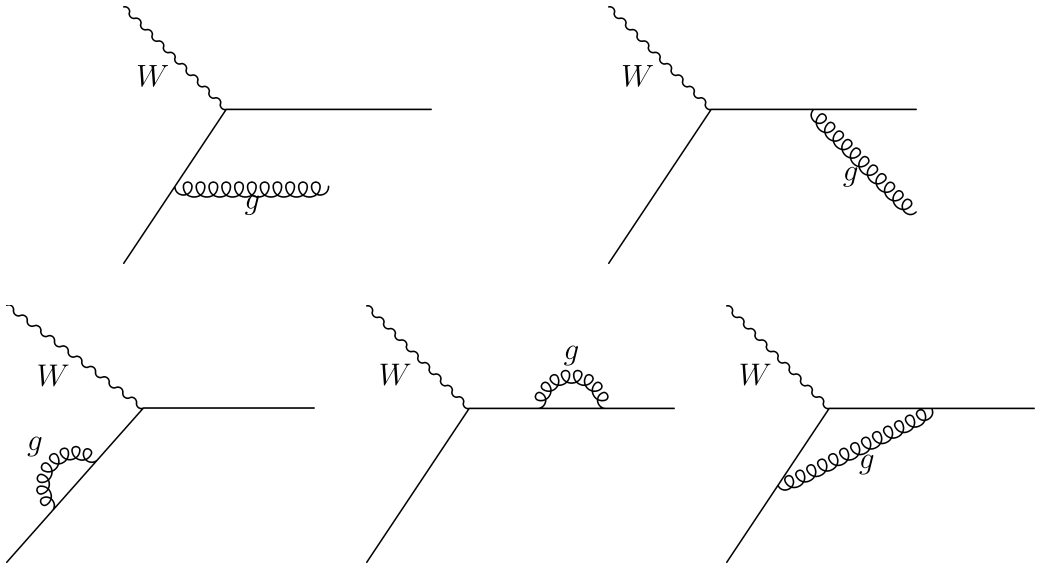


Fig. 1.14: Feynman diagrams for radiative gluon processes.

Higher order terms can be divided into two categories: boson-gluon fusion, shown in figure 1.13, and radiative-gluon processes, shown in figure 1.14. Boson-gluon fusion is often referred to as 'flavour creation' and corresponds to the process

$$W + g \rightarrow c + q' \,. \tag{1.94}$$

On the other hand, the radiative-gluon processes correspond to

$$W + q \rightarrow c\,(+g)\,. \tag{1.95}$$

For the second class, the argument of a large underlying parton distribution does not hold, since they are either initiated from a sea quark or Cabibbo-suppressed. Furthermore, any configuration with quasi-collinear gluons is implicitly taken into account in the DGLAP evolution of the quark distributions, as discussed in section 1.5. In summary, the latter diagrams can be safely neglected.

For the boson-gluon fusion process, the situation is slightly more complicated. On the one hand, the underlying parton distribution is very large, especially at small x. On the other hand, the leading log DGLAP evolution accounts for the logarithmic terms in the gluon splitting, corresponding to configurations where the quark is quasi-collinear with the gluon. In the $\overline{MS}$ scheme, the perturbative splitting function for a gluon to give rise to a quark is

$${}^{1}f_g^q(\xi,\mu) \;=\; \frac{\alpha_s(\mu)}{2\pi}\,\underbrace{\frac{1}{2}(1-2\xi+2\xi^2)}_{P_{gq}(\xi)}\log\frac{\mu^2}{m_q^2}\,, \tag{1.96}$$

where ξ is the fraction of the gluon's momentum carried by the quark q and μ is the renormalization scale. As expected, there is one power of α_s in this expression, picked up in the $gq\bar{q}$ vertex. The question whether the boson-gluon fusion process is entirely accounted by the DGLAP evolution equations, or should also be included in the hard scattering calculation, critically depends on the scales in the process. If all quarks are massless and if there are no transverse momenta in the initial state, then only the quasi-collinear configuration is kinematically allowed. However, for charm production near threshold, the charm mass introduces a relatively small scale in the hard scattering. Quark masses and transverse momenta can no longer be neglected, and the full boson-gluon fusion contribution needs to be taken into account. It is clear though that in the collinear region, care needs to be taken to avoid double counting between the boson-gluon fusion in the hard scattering and gluon splitting in the DGLAP evolution.

To calculate the boson-gluon fusion process, together with the zeroth order term of figure 1.12, we make the distinction between the raw scattering coefficient ω_q and the finite scattering coefficient $\hat{\omega}_q$, where q labels the parton species. The former corresponds to the full calculation of all Feynman diagrams of a free quark contributing to this interaction. Indicating the number of powers of α_s in the superscript, it can be written as

$$\omega_q \;=\; {}^{0}\omega_q \;+\; {}^{1}\omega_q \;+\; higher\ orders\,, \tag{1.97}$$

but turns out to be divergent, due to IR contributions arising from the emission of quasi-collinear gluons. This kind of IR contributions can be resummed, as in the DGLAP equations, to obtain a finite scattering coefficient $\hat{\omega}_q$. In doing so, we must account for the fact these contributions can actually transform between parton species, and thus we arrive at a 'factorization theorem' at parton level:

$$\omega_q = f_q^{q'} \otimes \hat{\omega}_{q'}\,, \tag{1.98}$$

introducing the splitting function $f_q^{q'}$ as the probability for a parton of type q to be resolved as a parton of type q'. Using this definition and keeping terms up to one power of α_s, the

raw scattering coefficient can be written as

$$\omega_q = \underbrace{{}^0f_q^{q'} \otimes {}^0\hat{\omega}_{q'}}_{\mathcal{O}(\alpha_s^0)} + \underbrace{{}^0f_q^{q'} \otimes {}^1\hat{\omega}_{q'} + {}^1f_q^{q'} \otimes {}^0\hat{\omega}_{q'}}_{\mathcal{O}(\alpha_s^1)} + \underbrace{higher\ orders}_{\mathcal{O}(\alpha_s^2)} . \tag{1.99}$$

The zeroth order splitting function ${}^0f_q^{q'}$ contains no powers of α_s and therefore cannot induce any transitions:

$${}^0f_q^{q'} = \delta_q^{q'} \quad \Longrightarrow \quad {}^0\omega_q = {}^0\hat{\omega}_q . \tag{1.100}$$

Similarly, we can express the first order contribution to the raw scattering coefficient:

$${}^1\omega_q = {}^1\hat{\omega}_q + {}^1f_q^{q'} \otimes {}^0\hat{\omega}_{q'} . \tag{1.101}$$

To show the implications of the last equation, we rewrite it for the specific case of boson-gluon fusion:

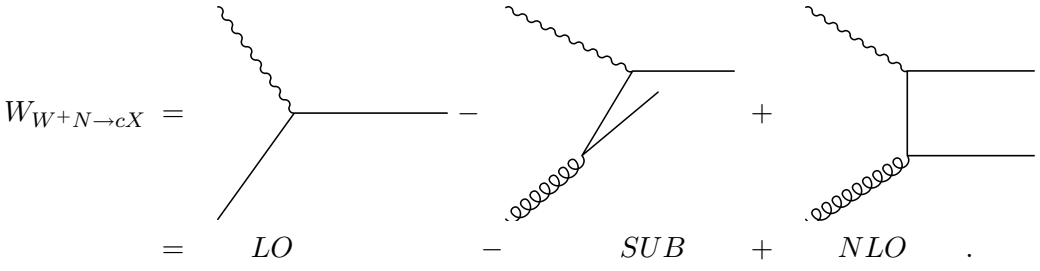
$${}^1\hat{\omega}_{W^+g\to c\bar{q}'} = {}^1\omega_{W^+g\to c\bar{q}'} - {}^1f_g^{q'} \otimes {}^0\omega_{W^+q'\to c\bar{q}'} . \tag{1.102}$$

The second term accounts for the fact that some part of the gluon fusion diagram of figure 1.13, is already implicitly contained in the quark-initiated diagram of figure 1.12. We will refer to this term as the subtraction term, whereas the first term will be referred to as the next-to-leading order term. The latter is somewhat of a misnomer, since it still refers to a tree-level diagram, albeit containing a power of α_s. Nevertheless, this is the convention most often found in the literature.

Combining these results, we can define an effective structure function for neutrino-induced charm production as:

$$W_{W^+N\to cX} = f_q^N \otimes {}^0\omega_{W^+q\to cX} - f_g^N \otimes {}^1f_g^{q_i} \otimes {}^0\omega_{W^+q_i\to cX} + f_g^N \otimes {}^1\omega_{W^+g\to c\bar{q}_i X} , \tag{1.103}$$

or schematically:



1.8 Current experimental status

The study of neutrino-induced charm production poses contradictory requirements. On the one hand, the small neutrino cross section calls for the use of very heavy targets, most often realized in massive detectors consisting of a coarse calorimeter followed by an iron toroid spectrometer to measure sign and momentum of the muon. On the other hand, the small lifetime of charmed hadrons inhibits their detection in these large calorimeter experiments. The only case in which traditional neutrino experiments are sensitive to charm production

is the semileptonic decay of the charmed hadron into a muon. This class of events has a characteristic signature consisting of two opposite sign muons in the final state, from the processes

$$\begin{aligned} \nu_\mu \,+\, d,s \quad &\longrightarrow \quad \mu^- + c \,+\, X \\ &\qquad\quad \downarrow \\ &\qquad\quad D \,\rightarrow\, \mu^+ \,+\, \nu_\mu \,+\, X' \end{aligned} \tag{1.104}$$

and

$$\begin{aligned} \bar{\nu}_\mu \,+\, \bar{d},\bar{s} \quad &\longrightarrow \quad \mu^+ + \bar{c} \,+\, X \\ &\qquad\quad \downarrow \\ &\qquad\quad D \,\rightarrow\, \mu^- \,+\, \bar{\nu}_\mu \,+\, X' \,. \end{aligned} \tag{1.105}$$

The first observation of such events in 1975 [36, 37], a few months after the discovery of the J/ψ [38, 39], provided corroborative evidence for a new quantum number: charm [40]. The existence of a fourth quark was first postulated by Bjorken and Glashow in 1964 [41], speculating that there ought to be four quarks (d,u,s,c) in analogy to the four leptons (e,ν_e,μ,ν_μ) known at the time. Later, Glashow, Iliopoulos, and Maiani offered a more compelling reason, now known as the GIM mechanism [42]. In 1963, Cabibbo had suggested [43] that the quark couplings to the W differed from the leptonic couplings by a factor of $\cos\theta_C$ for $u \rightarrow d$ and by a factor of $\sin\theta_C$ for $u \rightarrow s$. Given a small Cabibbo angle θ_C, this explains the smaller rate for strangeness-changing processes compared to strangeness-conserving processes. In spite of its success in correlating dozens of decay rates and cross sections, the Cabibbo theory had problems of its own, for instance a large amplitude for the $K_0 \rightarrow \mu^+\mu^-$ decay via a box diagram involving a u quark. The GIM mechanism disposed of these problems by postulating a charm quark whose couplings to d and s quarks carry factors of $-\sin\theta_C$ and $\cos\theta_C$, respectively. The ensuing cancellation diagrams rid the Cabibbo theory of its anomalous predictions. In 1973, Kobayashi and Maskawa extended the Cabibbo-GIM theory of quark generations with off-diagonal charged weak couplings to three families [44], introducing the Kobayashi-Maskawa matrix we have encountered in the introduction. This 3×3 unitary matrix is often referred to as the Cabibbo-Kobayashi-Maskawa or CKM matrix to stress the connection to the Cabibbo angle describing the case of two quark generations. Interestingly, the third generation was postulated before any quark or lepton of the third generation was discovered. Instead, the motivation was a desire to explain CP violation within the Cabibbo-GIM scheme. This requires a complex number in the *rotation* matrix, but with only two generations such a term can always be eliminated by a suitable redefinition of the quark phases.

A detailed study of opposite sign dimuon events only became possible with the advent of high intensity neutrino beams, both at CERN and at Fermilab. So far, three experiments have made a detailed study of this process, with a statistics of several thousand events in each of the experiments. The first investigation was performed in the CDHS experiment [45], later followed by the CCFR experiment [46] at Fermilab and the CHARMII [47, 48] and NOMAD [49] experiments at CERN. Each of these analyses only considers the leading-order charm excitation process. A second analysis conducted by the CCFR collaboration [50, 51] also takes into account the boson-gluon fusion process. Folllowing the authors' convention, the first and second CCFR analyses are referred to as leading-order (LO) and next-to-leading

order (NLO), respectively. Table 1.1 compares the number of events in each of the experiments. A similar study has been carried out using neutrino interactions in the CHORUS calorimeter [52], but this analysis is not yet completed.

	ν-induced	$\bar{\nu}$-induced
CDHS [45]	11041	3684
CCFR(LO) [46]	5044	1062
CCFR(NLO) [50, 51]	5030	1060
CHARMII [47, 48]	4111	871
NOMAD [49]	2714	115

Tab. 1.1: Comparison of the data samples in published analyses of dimuon events in neutrino experiments.

Each of the experiments has performed an analysis of charm production in terms of structure functions, allowing to fit the physical parameters that enter this cross section: the size and shape of the strange sea, the charm mass and the CKM matrix elements $|V_{cd}|$ and $|V_{cs}|$. Typically, the measured quantity is $B_\mu|V_{cd}|^2$, where B_μ is the average semi-leptonic branching ratio defined below. To eliminate the dependence on the CKM matrix element $|V_{cs}|$, two approaches are commonly used. The first simply consists in treating its value, based on the unitarity of the CKM matrix or on the measurements in other experiments, as an input parameter. The second is based on the comparison between the dimuon rate observed in a neutrino beam with the dimuon rate observed in an antineutrino beam. For the former, the relevant quark transitions are $d \to c$ and $s \to c$, for the latter $\bar{d} \to \bar{c}$ and $\bar{s} \to \bar{c}$. With the assumption that the quark sea is symmetric, equation 1.43, subtracting the antineutrino-induced rate from the neutrino-induced rate effectively removes the contribution from sea quarks, yielding an expression which only depends on the density function for valence d quarks and on the CKM matrix element $|V_{cd}|$.

The two fundamental difficulties are the level of background and the normalization. The major source of background is caused by the decay in flight into muons of pions and kaons in the hadronic shower of inclusive, charged current neutrino events. Its size can be estimated from the rate of like-sign dimuon events and is found to be 10 to 20 % of the opposite-sign dimuon event rate. Furthermore, it is significantly larger in the case of antineutrino beam. The problem of normalization is related to the requirement of a semi-leptonic decay for the charmed particle. One typically uses an average semi-leptonic *branching ratio* for decay into a muon, defined as the probability for a charmed hadron, produced in neutrino scattering, to give rise to a muon in the final state:

$$B_\mu \quad = \quad \sum_{D_i} f_{D_i}\, BR(D_i \to \mu X)\,, \tag{1.106}$$

where the sum runs over all charmed hadrons D_i = D^0, D^+, D_s^+, Λ_c,... The production fraction f_{D_i} for a hadron D_i is defined as the probability that a charmed quark, produced in neutrino scattering, hadronizes into D_i, with the obvious normalization

$$\sum_{D_i} f_{D_i} \quad = \quad 1\,. \tag{1.107}$$

The semi-leptonic branching ratios entering into the definition of B_μ, equation 1.106, have been determined to a reasonable degree of accuracy in e^+e^- experiments, both at the $c\bar{c}$ threshold and at higher energies. Unfortunately, this is not true for the production fractions f_{D_i}. By their very nature, they can only be measured in neutrino experiments able to identify the charmed hadron produced. So far, only one such experiment has been conducted: the E531 experiment [53] at Fermilab, with a design very similar to the CHORUS experiment. To tag and identify charmed particles produced in neutrino interactions, it used an emulsion target followed by a series of electronic detectors providing both kinematic reconstruction and particle identification. From a study of 122 charmed-particle decays among 3855 located neutrino interactions, the E531 collaboration has extracted the total rate for each of the charmed hadrons [54]. The experiment also presents the dependence on the kinematic variables [55]: Bjorken x and y characterizing the neutrino scattering, and the fragmentation variable z describing the hadronization process. This sample of 122 events continues to be the only reference for the production fractions f_{D_i} and has recently been reanalyzed in this respect [56].

2. THE CHORUS BEAM AND DETECTOR

2.1 Introduction

The CHORUS experiment was proposed [57, 58] primarily to search for $\nu_\mu \rightarrow \nu_\tau$ oscillations through the appearance of ν_τ in a ν_μ beam, aiming to explore the domain of small mixing angles down to $\sin 2\theta_{\mu\tau} \sim 3 \times 10^{-4}$ for mass parameters $\Delta m^2 \sim 1eV^2$. This represents an order of magnitude improvement over the previous generation of experiments [59]. The requirement for the neutrino beam is to provide as many neutrinos as possible, at an average energy well above the threshold for ν_τ charged current interactions. For the detector, the major novelty is the large-scale deployment of nuclear emulsion: an active target of 800 kg with a spatial resolution high enough to reveal not only the decay products of the tau lepton, but also the parent track itself.

The CHORUS detector is schematically shown in figure 2.1. It comprises an active target of nuclear emulsion, a scintillating fibre tracker, a hadron spectrometer, a high resolution calorimeter and a muon spectrometer. The emulsion target, described in more detail in the next chapter, serves both as neutrino target and as detector, offering sub-micron resolution for the track and vertex reconstruction in the vertex region. The major drawback is the absence of time information: any charged particle traversing the emulsion during the exposure will leave a track.

The overriding aim in designing the experiment consisted in reaching a sensitivity to $\nu_\mu \rightarrow \nu_\tau$ neutrino oscillations at least one order of magnitude better than what had been achieved in the E531 experiment. Based on this number, and given the available scanning capacity, the requirements on the electronic detector performance were expressed in terms of the enhancement of ν_τ interactions in the sample of events to be analyzed in the emulsion. At the time of the CHORUS proposal, it was estimated that a total of $\sim$ 40,000 events could be scanned in two years using semi-automatic techniques, provided electronic detectors would predict the impact point and slope of tracks at the exit face of the emulsion within at most a few hundred μm and a few mrad, respectively. The required resolution and two-track-separation has been achieved using scintillating fibre trackers. A magnetic field in the emulsion target region was excluded not only because of the opto-electronic readout of the fibre trackers, but also because it would have severely complicated the matching between electronic trackers and emulsion.

A sample of $\sim$ 40,000 neutrino interactions does not by itself allow to reach the sensitivity aimed for, especially in the absence of any particle identification or charge determination. The main purpose of the downstream detectors is precisely to enrich the sample of events sent to scanning, giving preference to events which are more likely to be ν_τ charged current interactions, from the overwhelming background of plain ν_μ interactions.

For the decay mode $\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$, a negative muon would originate at the secondary vertex, whereas events with a muon attached to the primary vertex are of no interest to the

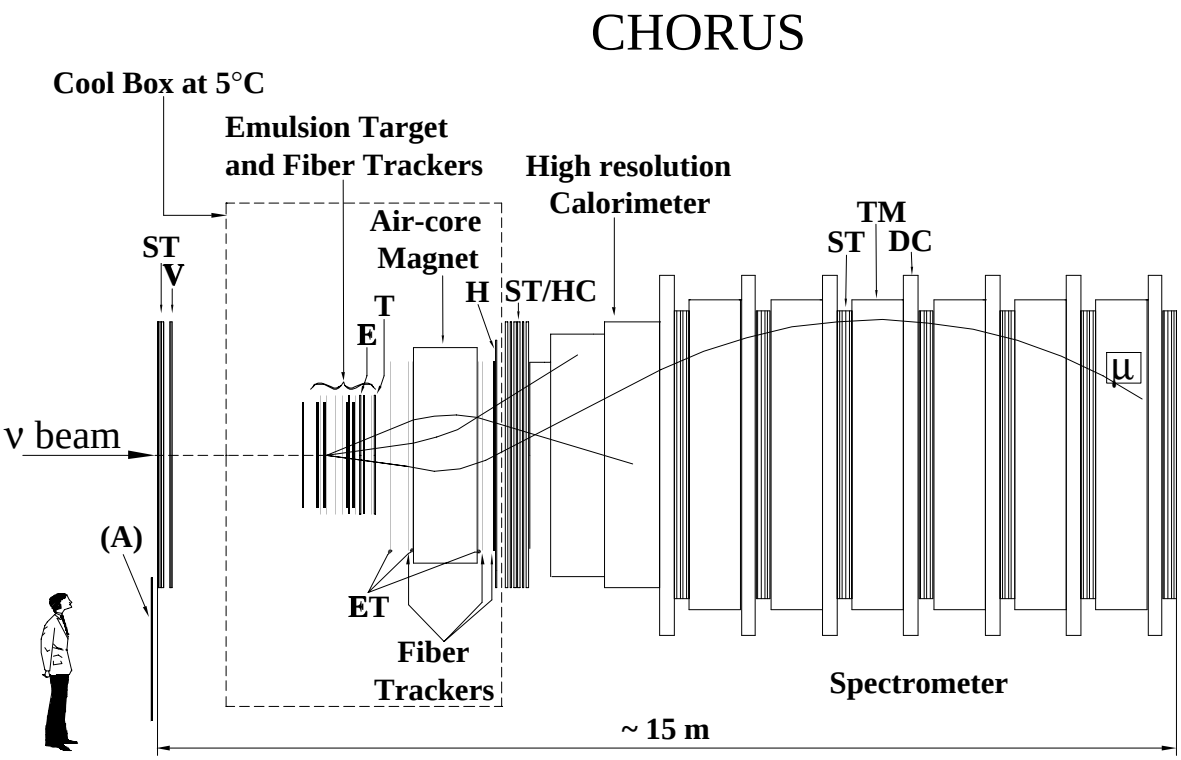


Fig. 2.1: Sketch of the CHORUS detector: fibre trackers in the emulsion target region (section 2.3.1); the hexagonal magnet spectrometer (section 2.3.2) containing fibre trackers, honeycomb chambers (HC) and emulsion trackers (ET); the high-resolution calorimeter (section 2.3.3); the muon spectrometer (section 2.3.4) containing streamer tubes (ST), toroidal magnets (TM) and drift chambers (DC); finally the trigger system (section 2.3.5) composed of a veto plane (V), an anticounter plane (A), an emulsion plane (E), a timing plane (T) and a trigger hodoscope (H).

oscillation search. An important distinction between the reaction $\nu_\tau N \to \tau^- X$ followed by $\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$ or $\tau^- \to \pi^- \nu_\tau$ and the reaction $\nu_\mu N \to \mu^- X, \nu_\mu X$ lies in the missing transverse energy carried away by neutrinos and the characteristic correlation of the missing energy vector with the energy vector of the hadronic final state X. For hadronic decay modes of the tau, neutrino-induced charm production where the primary muon has not been identified, constitutes an important background. But for the dominant beam component of ν_μ, the charm hadron would be positive. The branching ratio and decay detection efficiency for the muonic decay of the τ lepton are 17.4 % and 39 %, respectively. For the single-prong hadronic decay $\tau \to h^-(n\pi^0)\nu_\tau$, the corresponding values are 49.5 % and 11 %.

In summary, the event properties which can be exploited to enhance the signal of ν_τ interactions compared to the background of plain ν_μ interactions are the presence of a muon, the charge of hadrons, and the energy vector of the hadronic shower. The magnetized iron spectrometer reliably identifies muons and measures their charge and momentum. The scintillating fibre trackers before and after an air-core hexagonal magnet allow the charge determination of hadrons, perturbing them only slighly so as not to compromise their calorimetric

energy measurement. The high resolution calorimeter measures not only the shower energy, but also its direction.

We will now briefly describe the neutrino beam and each of the electronic detectors. The discussion will be fairly terse, referring instead to the physics literature whenever possible. A more detailed description of the entire apparatus can be found in [60].

2.2 The neutrino beam

Since the 1970s, high-energy neutrino beams have been available both at CERN and at Fermilab, produced from the decay of π and K mesons originating in the interaction of high-energy protons with a target. Broadly speaking, they can be subdivided in narrow-band and wide-band beams (NBB or WBB). The former include momentum selection of the parent mesons, leading to neutrino beams with a well-defined energy spectrum. The latter are designed to focus a wide range of meson momenta to reach the largest possible neutrino beam intensity. Focusing with wide momentum acceptance is achieved by the horn-reflector system, a doublet of conically shaped, toroidal magnets first proposed by S. van der Meer (and subsequently used not only in neutrino beams but also in the CERN antiproton accelerator, see [61]). Depending on the polarity of horn and reflector, either positively or negatively charged particles will be focused, with the particles of opposite charge defocused. Positive mesons will give rise to neutrinos through the decays $\pi^+, K^+ \to \mu^+\nu_\mu$, negative mesons to antineutrinos through the charge conjugate decay.

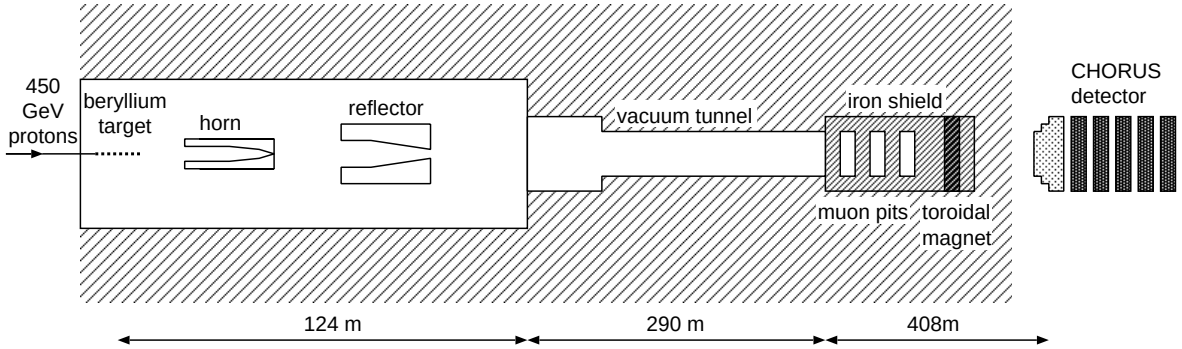


Fig. 2.2: Sketch of the WANF neutrino beamline.

For the CHORUS oscillation search, the West Area Neutrino Facility (WANF) [62] at CERN, schematically shown in figure 2.2, provided a beam of predominantly muon neutrinos. The Super Proton Synchrotron (SPS) accelerates protons from the Proton Synchrotron (PS) to an energy of 450 GeV, in a cycle repeated every 14.4 s. Once accelerated, the SPS protons are delivered to the neutrino facility in two extractions, also called spills, each lasting about 6 ms and separated by 2.7 s. For the experiments, a longer spill is preferable because it reduces the dead time as well as the effect of beam muons overlaid on neutrino events. On the other hand, the high currents in the focusing magnets can only be sustained during a relatively short pulse.

In each spill, up to $1.8 \cdot 10^{13}$ protons impinge on the neutrino production target composed of a succession of 11 beryllium rods of 10 cm length and 3 mm diameter, separated by 9 mm

gaps [63]. This particular geometry not only minimizes the number of reinteractions in the target, it also reduces the thermal and mechanical stresses to which the target is exposed. The horn and reflector are located 20 m and 90 m downstream of the production target. They both consist of a 7 m long aluminium cylinder, with a conically shaped inner surface. The outer diameter of horn and reflector is 40 cm and 80 cm, respectively. Both are operated with short pulses (8 ms) of high current (100 kA). The space between horn and reflector and downstream of the reflector is occupied by helium-filled tubes to minimize the number of interactions with air. Two collimator windows, one behind the production target, the other between horn and reflector, serve to reduce the number of mesons which are not focused. A 2 mm thick titanium window, situated 124 m downstream of the production target, provides the entrance to a 290 m long tunnel of low vacuum (0.5 Torr). The diameter of the tunnel is 220 cm for the first 30 m and 120 cm for the remaining 260 m.

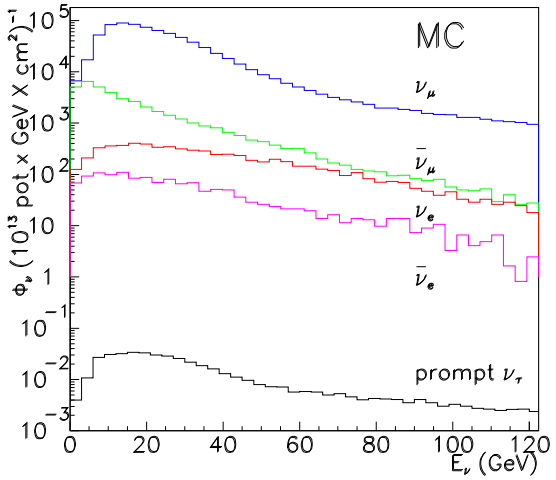


Fig. 2.3: Energy spectra (flux rates) of the different components of the neutrino beam intercepted by the CHORUS emulsions as predicted by a full Monte Carlo simulation of the beamline.

Positively charged pions and kaons account for respectively 90 % and 9 % of all charged particles reaching the vacuum tunnel. Both are unstable, with a respective flight length of 55.9 m/GeV and 7.52 m/GeV, and decay predominantly into a positive muon and a muon-neutrino, with branching ratios of 99.99 % and 63.5 %. Mesons which did not decay, protons which did not interact in the target, and hadrons or charged leptons from the meson decays are all absorbed downstream of the decay tunnel in shielding composed of 225 m of iron and 144 m of earth. A toroidal magnet sweeps the positive particles towards the center to make the iron shielding more effective. The neutrino beam emerging behind the shielding is composed of predominantly ν_μ, with a ∼5.6% contamination of $\bar{\nu}_\mu$, ∼0.7% of ν_e and ∼0.17% of $\bar{\nu}_e$. The electron-neutrinos come primarily from the K_{e3} decay of positive kaons, whereas the electron-antineutrinos are due to negative mesons which were not sufficiently defocused nor absorbed in the collimators and to the K_{e3} decay of K_L^0. The ν_τ content of the beam

originates from D_s production in the beryllium target, followed by the $D_s \rightarrow \tau\nu_\tau$ decay. It was calculated to be below the level at which even a single event could be detected over the lifetime of the experiment [64].

Figure 2.3 show the energy spectra of each component indicating also their relative abundance. The rates shown are based on simulation but have been validated from neutrino data, most comprehensively in [65]. The simulation of the beam has been further refined using the results of a dedicated experiment which studied the production of pions and kaons by 450 GeV protons hitting a beryllium target [66]. The proton intensity in each spill is counted by beam current transformers just upstream from the production target, and the alignment of the beamline, as well as constraints on the beam composition, are derived from a measurement of the muon flux in the shielding, by a series of Si counters.

2.3 Electronic components of the CHORUS detector

2.3.1 Scintillating fibre trackers

For each of the four successive emulsion target modules placed along the beam, one emulsion sheet with good angular resolution is installed immediately downstream from the target, labeled SS in figure 2.4. A second one, labeled CS in the figure, is located 38 mm further downstream to allow the tracks to separate. The second sheet is used as interface to the electronic detectors. To limit the density of background tracks, the CS or *changeable sheet* is exchanged more frequently than the target, typically every few months. The purpose of the electronic trackers is to reconstruct the event and predict the impact point and slope of tracks in these interface sheets, with an accuracy good enough for the scanning to uniquely identify this track among those integrated over several weeks or months. To reduce extrapolation errors, the tracker system must also offer good two-track separation such that it can be located immediately downstream from the emulsion.

These requirements were met using a scintillating fibre tracker [67, 68] consisting of eight modules, interleaved between the four emulsion stacks. Each module consists of four planes, measuring two transverse coordinates Y,Z and two stereo projections $Y^{\pm}$,$Z^{\pm}$ rotated by $\pm 8^{\circ}$ relative to Y,Z. The geometry of half of the target region is shown in figure 2.4, the other half is identical and located immediately downstream from the first. In turn, each plane contains seven layers of $\oslash$ 500 μm fibres in a staggered geometry. The far ends are polished and sputtered with aluminium, obtaining a reflectivity of about 80 % and a 6.0 m *effective* attenuation length instead of the 2.2 m *bare* attenuation length. They are read out by 40 optoelectronic systems, capable of measuring individual photo-electrons with spatial information. Each optoelectronic chain is composed of three electrostatic image intensifiers, a micro-channel plate (MCP) and a CCD camera. The net gain for the entire assembly lies between 10^4 and 10^5 and the average quantum efficiency (QE) is 18 %, determined by the QE of the first photocathode. With the overall demagnification factor of 0.11, the 16×23 μm^2 pixels on the sensor correspond to 145×208 μm^2 at the input window. Two parameters characterize the spatial resolution of the readout system: the *spot size* due to the intrinsic resolution of the chain and the *spot displacement* caused mainly by the focusing of the first image intensifier. Their standard deviations at the input window are 136 μm and 89 μm, respectively. The CCD contains an image zone and a memory zone, with a transfer time from one to the other of 125 μs. The actual readout takes about 20 ms, limiting the number

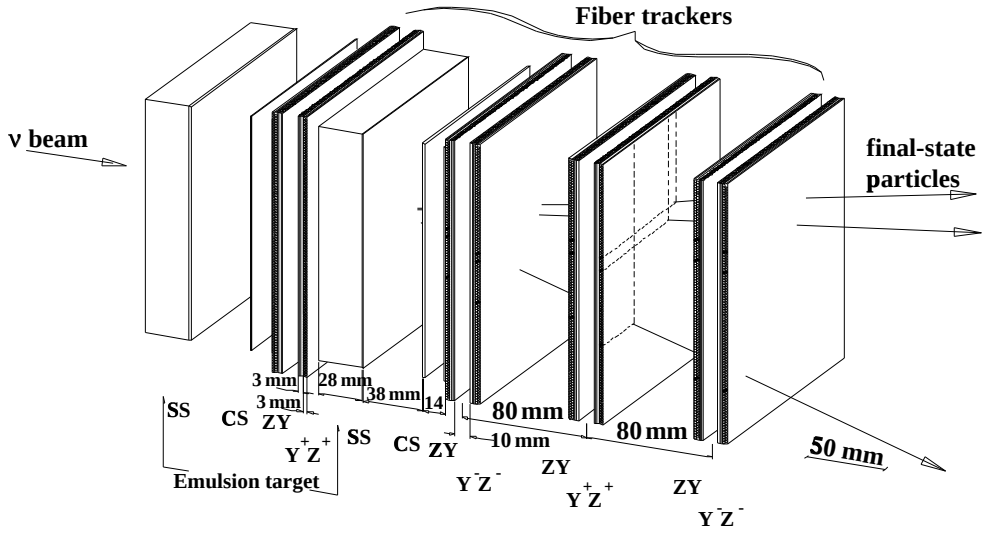


Fig. 2.4: Geometry of the trackers in the emulsion target region; the labels Y, Z, $Y^{\pm}$, $Z^{\pm}$ denote the different orientations of the tracker planes. The complete target region consists of two such arrangements placed behind each other along the X axis.

of events that can be recorded per spill to two.

For the seven layers per plane, the Poisson mean of the hit density ranges between 5 and 7 hits at the far and near ends of the detector, respectively. Using the data from beam muons, the inefficiency per plane — the probability not to record any signal for a minimum ionizing particle passing through — was measured to be about $2 \cdot 10^{-3}$, consistent with the expectation from Poisson statistics based on the mean hit density. The tracker resolution is most easily expressed through its prediction accuracy, defined as the width of the distribution describing the deviations of the target tracker predictions from the tracks found in the emulsion sheets. Corrected for the intrinsic angular resolution in the emulsion scanning, the prediction accuracies are $\sim 150\mu m$ in position and $\sim 2mrad$ in slope.

2.3.2 Hexagonal magnet spectrometer

For hadronic tau decays, the rejection of background from charm production depends critically on the ability to determine the charge of hadrons. A second source of background are the so-called white kinks: elastic or inelastic hadron interactions which show no visible recoil nor delta rays or Auger electrons and which in the emulsion cannot be distinguished from a decay. To reduce this type of background, the measurement of hadron momenta must be precise enough to reject decay topologies with a low-momentum secondary or corresponding to small transverse momentum. The selection of tau candidates includes lower limits of 1 GeV/c and 250 MeV/c for the secondary momentum and transverse momentum, respectively. The spectrometer to measure the charge and momentum of hadrons is subject to several external constraints. It should be short enough not to compromise the acceptance

of the downstream detectors and the amount of material should be small enough not to spoil the calorimeter measurement. Furthermore, no stray field from the magnet can be tolerated as that would exclude the use of electrostatic image intensifiers and unnecessarily complicate tracking in the target region. Finally, the spectrometer must match the large dimensions of the neutrino target.

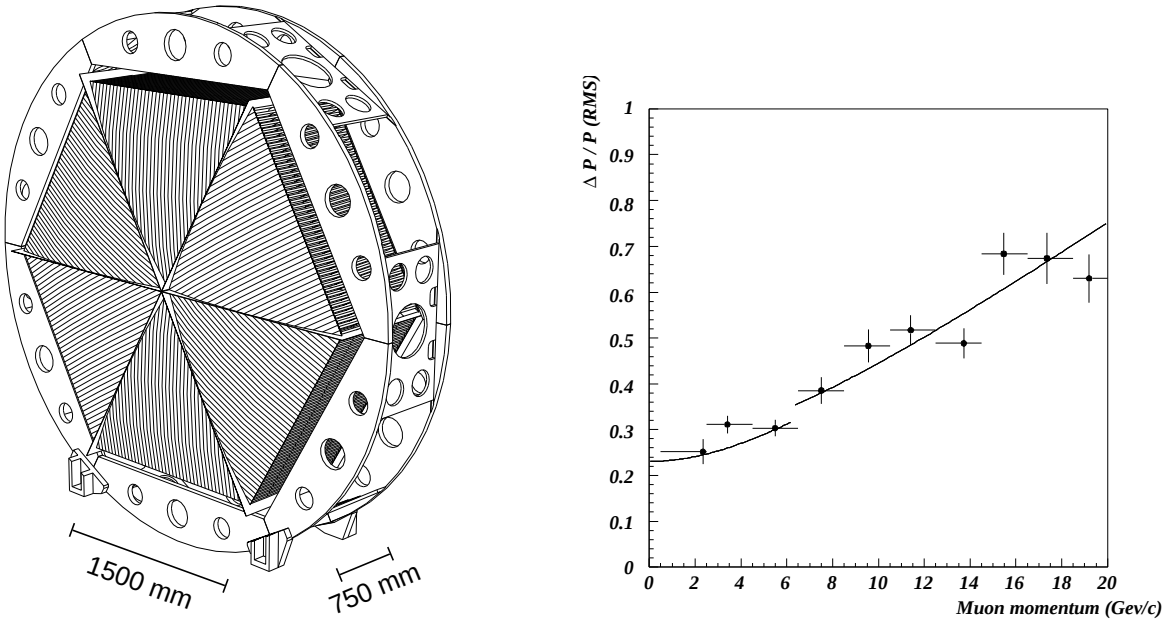


Fig. 2.5: Left panel: drawing of the hexagonal air-core magnet. Right panel: resolution on the momentum measured by the magnet trackers for particles traversing the hexagonal magnet, the momentum measurement uses only diamond tracker information and the resolution was determined by comparing, in neutrino events, the momenta measured for muons which are also reconstructed in the muon spectrometer (see section 2.3.4).

The hexagonal spectrometer, consisting of an air-core magnet and scintillating fibre trackers, fulfills these requirements. The hexagonal toroidal magnet [69] has an external radius of 1.5 m and an axial length of .75 m. It is composed of regular triangular sectors, with equidistant windings running orthogonal to the polygon side for the front and back faces and parallel to the axis for the other planes, as shown in the left panel of figure 2.5. This arrangement results in a homogeneous field in each sector, parallel to the outer side of the polygon and constant over the entire volume of the sector, with a negligibly small free-field region along the central axis. The conductors are made of 2.5 mm Al alloy, corresponding to less than 4 % of a radiation length. Only the diagonal planes represent material all along the axial depth, resulting in dead space for the momentum reconstruction. These spokes have an azimuthal thickness of 11 mm and are crossed by about 15 % of all particles with momenta between 2.5 GeV/c and 10 GeV/c originating in the emulsion target. The 3200 A magnet current is pulsed in accordance with the neutrino beam and generates a 0.12 T field with a flat top of 12 ms, enclosing the 6 ms duration of the neutrino ejection. Effective shielding limits the stray field to $\sim$ 1.5 G at 1 cm distance from the magnet surface.

The 0.12 T field over 0.75 m along the axis corresponds to a 2.5 mm sagitta for a par-

ticle with 1 GeV/c momentum. The sagitta is measured using scintillating fibre trackers with the same components as the trackers in the target region. There are three tracker modules, one upstream and two downstream from the magnet, read out by 18 optoelectronic systems. Each module is hexagonal and consists of two planes, in turn composed of three adjacent *diamond-shaped* paddles measuring one of three coordinates at 120° from each other. The second plane measures the projection at 60° from the first. The magnet and diamond tracker (DT) share a six-fold symmetry such that one of the two coordinates measured in each module is in the bending plane. Broadly speaking, the momentum reconstruction algorithm uses the accurately known track parameters measured in the target region, together with the space point in the upstream diamond tracker module, to define a range of hypotheses downstream from the magnet, corresponding to a range of momenta. If hits in the two downstream modules can be matched to the track hypothesis, this immediately leads to a momentum determination. The resolution obtained is

$$\frac{\Delta p}{p} = 22\ \% + 3.5\ \% \frac{p}{\mathrm{GeV}/c}\ ,$$

with the first term due to multiple scattering and the second term reflecting the measurement accuracy. The right panel of figure 2.5 shows the width of the difference between the hadron spectrometer and muon spectrometer measurements for penetrating tracks; the line corresponds to the two resolutions added in quadrature and is dominated by the hadron spectrometer resolution.

The hexagonal spectrometer resolution is sufficient for the CHORUS analysis, but unfortunately the momentum cannot be measured for a large number of tracks: the four planes downstream of the magnet are quite often insufficient to resolve ambiguities, especially for the many events where an electromagnetic or hadronic shower develops already in the target region. To increase the momentum reconstruction efficiency, an additional set of tracker planes was inserted after the first two years of data taking, not only improving the redundancy in the overall system but also, crucially, allowing for stand-alone tracking downstream of the magnet. The new tracker had to cover the hexagonal magnet area of 5.8 m^2, had to fit in the 21 cm gap between the last fibre plane and the calorimeter and could not introduce much material, as this might cause additional showering. Honeycomb trackers [70] fulfill those requirements and three planes were succesfully installed and operated during the last two years of data taking [71]. They essentially consist of a set of drift tubes, hexagonal honeycomb cells, made from earthed, conductive foils with central wires at high voltage. The basic element is a 1 $\times$ 2.7 m^2 monolayer with 216 cells of 1 m length. A 1 m wide and 75 μm thick foil of conductive plastic is folded to a half hexagonal structure with a pitch of 12.7 mm and point-welded to a second, symmetrically folded foil to form hexagonal cells. At the center of each cell, a 30 μm thick anode wire is strung. Three monolayers are mounted side by side to make a single plane of three meter long drift tubes with a total area of 2.7$\times$3 m^2. Six of these planes are subsequently glued on top of each other, staggered so as to minimize the acceptance loss from the dead space of 1.2 mm width in between honeycomb cells. Three such modules, put together under maximum stereo angles 0°, 60° and 120°, form the CHORUS honeycomb chamber (HC). The information from this tracker has been included in the momentum reconstruction software, and the design aim of 200 μm single cell resolution has been achieved.

The hadron spectrometer, using the hits from both diamond tracker planes and honeycomb chambers, can reliably determine the charge for hadrons with momenta up to 10 GeV/c,

or about 44 % of the charged secondaries from single-prong hadronic tau decays in CHORUS. To extend this range up to 30 GeV/c, adding a further 26 % of the single-prong hadronic tau decays, required a tracker with very high position resolution, given the relatively modest field integral in the hexagonal magnet. Over such a large surface, the only technique which offers sufficient position resolution at an acceptable cost is nuclear emulsion. For the last two years of data taking, an emulsion tracker (ET) was installed, consisting of the three modules indicated in figure 2.1: one immediately downstream from the target tracker and two on both sides of the hexagonal magnet, in between the magnet surface and the diamond trackers [72]. The scanning time involved excludes the systematic momentum measurement of all tracks from all neutrino interactions, but an ET measurement is certainly possible for all secondaries from decay topologies, not only for tau candidates but also for the large sample of charm events. In each of the three planes, tracks are searched for in an area of 1 mm^2; requiring a track to be found in all three planes leads to an efficiency of about 50 %. Replaced once during each year of data taking, the emulsion sheets integrate up to 60 tracks/mm^2 within a 300 mrad angular acceptance. After comparison of the slope measured in emulsion to the prediction from the electronic detectors, and requiring the track to be found in all three planes, a background track is selected for about 2 % of the tracks in the beam direction, and significantly less than that for tracks in other directions. The position resolution is dominated by the alignment precision and amounts to $\sigma \sim 30\mu m$, corresponding to a momentum resolution of

$$\frac{\Delta p}{p} = 22\ \% + 1\ \% \ \frac{p}{\mathrm{GeV}/c} \ .$$

To monitor relative movements of detector elements, a number of RASNIK systems (Relative Alignment System of NIKhef) has been installed [73]. These consist of a light source with a coded mask, a lens and a CCD camera, each of them mounted onto detector elements or mechanical support structures. Their relative movement results in a change in the image, with transverse motion shifting the image and movement along the axis changing its size. Four such systems were installed into the CHORUS detector, mounted on diamond tracker and honeycomb chamber planes. Apart from day-night effects and cycles corresponding to the ventilation system, an unanticipated effect was also observed: a shift in the diamond tracker plane which could be attributed to the air-core magnet resting at a lower temperature when it is not pulsed.

2.3.3 High resolution calorimeter

Since the presence of a neutrino in the tau decay results in an unbalance of the visible transverse momentum, the number of events to be scanned in the emulsion can in principle be reduced through a selection on the basis of kinematic variables. This reduction is only possible if the calorimeter measures the hadronic shower with sufficient energy and direction resolution. To facilitate the matching between trajectories reconstructed in the target tracker and in the muon spectrometer, the calorimeter must also have the capability to track through-going muons.

These requirements are met by a *spaghetti* calorimeter interleaved with streamer tube planes. A spaghetti calorimeter consists of $\oslash$ 1 mm scintillating fibres embedded into a lead matrix, a technique developed by the LAA-SPACAL collaboration [74]. The volume ratio of 4:1 between lead and scintillator assures both compensation and good sampling.

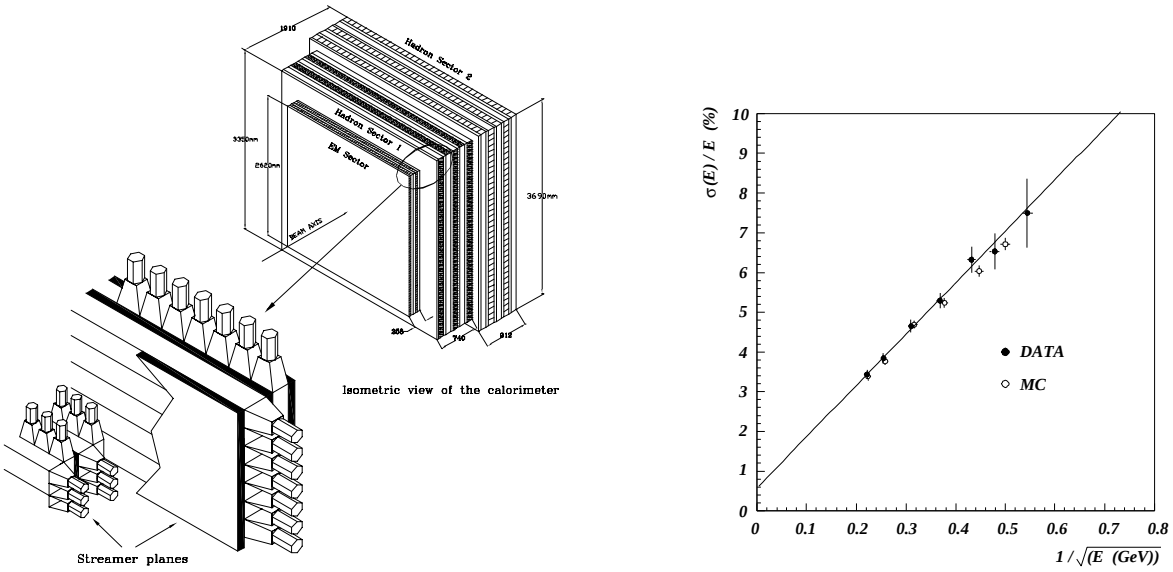


Fig. 2.6: Left panel: schematic view of the calorimeter, indicating the three sections and the different granularity of the electromagnetic and hadronic parts. Right panel: the hadronic energy resolution of the calorimeter, the full dots are data points from an exposure to pion test beams of well-defined energy.

In this context, compensation refers to the ratio between the response for electromagnetic showers and that of hadronic showers. As the fraction of energy absorbed in the passive material, in this case lead, is larger for hadrons, the measured response is lower for hadronic showers. This effect can be mitigated by maximizing the surface of active material covering a given volume of passive material, with the suitable volume ratio to be determined for each geometrical setup. For perfect compensation, the ratio in the energy response would be one; in the CHORUS calorimeter, the deviation from perfect compensation is less than 10 %.

The CHORUS calorimeter [75] differs from those conceived previously in the orientation of the fibres: to achieve the angular resolution, they are placed perpendicular to the beam direction. The size of the neutrino beam requires modules of considerable length; to reduce the problems generated by light attenuation, the fibres are read out at both ends. The geometric mean of the pulseheight in the two signals yields a value for the module response which is to good approximation independent of the hit position.

Section	EM	HAD1	HAD2
Depth	$21.5X_0$, $0.78\lambda_I$	$55.2X_0$, $2.0\lambda_I$	$67.1X_0$, $2.44\lambda_I$
Number of planes	2H + 2V	3H + 2V	2H + 3V
Area	262×262 cm^2	335×335 cm^2	369×369 cm^2
Readout module	4×4 cm^2	8×8 cm^2	10×10 cm^2

Tab. 2.1: Characteristics of the calorimeter sections.

The calorimeter consists of three sections with decreasing granularity, schematically shown in the left panel of figure 2.6. Table 2.1 indicates the dimensions of each of the sections, show-

ing also the orientation of the planes and the increasing granularity of the readout. The first section, called EM, measures the electromagnetic component of the shower. The other two sections, called HAD1 and HAD2, complete the measurement of the hadronic component. The HAD1 section is similar in construction to EM but is read out in larger modules and has larger dimensions, in both the transverse and longitudinal directions. The HAD2 section does not use scintillating fibres but instead a more conventional sandwich structure, with five 1.6 cm thick lead plates and five 0.4 cm thick scintillator strips forming a module.

To efficiently track through-going muons, 22 planes of limited streamer tubes (see section 2.3.4) are inserted between the calorimeter planes, as indicated in the left panel of figure 2.6. They are arranged in pairs with horizontal and vertical wires, read out in digital mode.

The total calorimeter thickness is 5.2 interaction lengths λ_I (144 radiation lengths X_0), sufficient to contain on average 99 % of the shower induced by a 5 GeV pion. For about 90 % of the hadrons from neutrino interactions in CHORUS, the energy is less than 5 GeV. To determine the absolute energy calibration, the calorimeter response was studied using test beam of electrons and pions with known energies in the range from 2.5 to 10 GeV and from 3 to 20 GeV, respectively [76]. Both the linearity as a function of energy and the uniformity as a function of position were found to be very good. The energy resolution for electrons amounts to

$$\frac{\sigma(E)}{E} = \frac{(13.8 \pm 0.9)\%}{\sqrt{E/GeV}} + (-0.2 \pm 0.4)\% \ ,$$

while for pions it is

$$\frac{\sigma(E)}{E} = \frac{(32.3 \pm 2.4)\%}{\sqrt{E/GeV}} + (1.4 \pm 0.7)\% \ ,$$

as illustrated in the right panel of figure 2.6. The angular resolution for a 10 GeV hadronic shower is about 120 mrad (FWHM). The constant term in the resolution function is related to instrumental effects, including electronic noise and broken modules, but dominated by intercalibration. As hadronic showers have larger coverage both laterally and longitudinally, they are more affected by the intercalibration between modules within a section and between sections.

2.3.4 Muon spectrometer

The calorimeter represents 5.2 hadronic interaction lengths, filtering nearly all particles produced by neutrino interactions in the target region except muons with momentum higher than $\sim$ 1.5 GeV/c. This provides their identification as muons, and the role of the muon spectrometer is to measure their trajectory, momentum and charge. Just about every electronic detector in high-energy neutrino beams included muon tracking devices, and most of the components for the CHORUS spectrometer could be recuperated from previous experiments. Apart from measuring muons, the spectrometer must also offer some calorimetric ability to measure the energy leakage from hadronic showers.

The CHORUS spectrometer consists of magnetized iron, instrumented with scintillators, drift chambers and streamer tube chambers, all of them indicated in the left panel of figure 2.7. The six toroidal magnets in the CHORUS spectrometer were originally used as

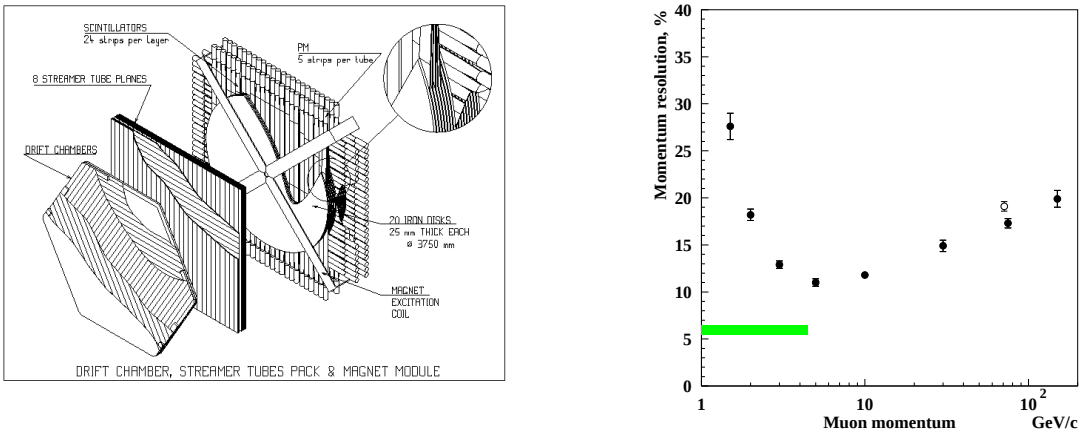


Fig. 2.7: Left panel: one of the six sections in the CHORUS spectrometer, each one composed of a drift chamber with three planes, eight planes of streamer tubes and an iron magnet with four planes of scintillators. Right panel: muon spectrometer momentum resolution as a function of incident momentum; the full dots refer to simulation results, the open dot to test beam muons at 71 GeV/c; the shaded band corresponds to muons stopping in the spectrometer, with their momentum determined by range.

calorimeter in the upgraded CDHS detector [77] and later in the CHARM II end-system. Each magnet is constructed from twenty 2.5 cm thick iron disks with an outer diameter of 375 cm, interleaved with 0.5 cm thick scintillator planes inserted into 0.6 cm gaps between the disks. The integral of the field, nearly symmetric azimuthally and varying by 25 % radially, is approximately 0.85 T·m per magnet. Negative muons are bent towards the center of the magnet.

Seven hexagonal drift chambers [78], developed for the original CDHS detector, are mounted in front of, between or behind the six magnets. Each chamber consists of three planes at a relative orientation of 60°, each one containing 62 sense wires with a diameter of 40 μm, spaced by 6 cm. The hit resolution obtained from a drift time measurement is about 1 mm and the efficiency per plane is better than 99 %. For 83 % of the forward going tracks, left-right ambiguities can be resolved using only the chamber.

Each of the drift chambers is followed by eight planes of limited streamer tubes, previously used in the CHARM II calorimeter [79]. The planes, covering an active area of 367×367 cm^2, consist of 352 square-shaped limited streamer tubes with an inner area of 9×9 mm^2 and an average wire spacing of 10.5 mm, followed by 18 mm wide cathode pick up strips with 21 mm spacing, orthogonal to the wire direction. Using a drift-time measurement, the resolution achieved for the wire hits is about 800 μm, whereas the hit resolution of the strip readout is limited to 2.4 mm (rms) due to a relatively high noise rate. For the wire readout, the hit efficiency per plane for beam muons is (90±2)%; for the strip readout, the effective[1] efficiency is (80±2)%. Consecutive planes have alternating horizontal and vertical wire orientations so that a stack of streamer tubes provides up to eight independent measurements in each of the two projections. On average, a forward going muon produces 3.6

[1] Only clusters with two or more neighbouring strips having signals above the threshold are used for subsequent analysis.

wire hits and 3.2 usable strip hits per projection. The same streamer tubes, but with only a digital readout, are also used for coarse tracking elsewhere in the experiment: 4 planes in front of the veto scintillators, 6 planes between the trigger hodoscopes and the calorimeter (in August 1996 replaced by the honeycomb chamber, see section 2.3.2) and 22 planes interleaved between the calorimeter planes.

Finally, the twenty layers of scintillator strips embedded in each magnet are grouped in four planes of five consecutive layers, two planes with vertical and two planes with horizontal orientation. These are used for triggering purposes, for the detection of energy leakage from the calorimeter and for a range measurement of muons with up to 5 GeV/c at the entrance of the spectrometer. The energy measurement of the shower tails is primarily of interest for studies of neutrino nucleon structure functions where the calorimeter is used as massive target.

The momentum resolution of the muon spectrometer has been evaluated using a detailed simulation, the results of which are shown in the right panel of figure 2.7. The absolute calibration was performed using negative 75 GeV/c test beam muons, corresponding to 71 GeV/c at the entrance of the spectrometer due to ionization losses in the upstream detectors and beam elements. These also provide a validation of the resolution determined from simulation, indicated by the open dot. For stopping muons, the charge is determined from the curvature of the trajectory, whereas the momentum is determined from their range, resulting in a momentum resolution of 6 %, estimated from the comparison with the nominal momentum for simulated muons. The efficiencies of the trigger signals provided by the spectrometer scintillators are close to 100 %. The spectrometer energy resolution for shower tails above 1 GeV is (110-150%)/$\sqrt{E(GeV)}$. Relative alignment of the spectrometer components is performed with 100 GeV/c test beam muons recorded with the magnetic field off. Alignment relative to the rest of the CHORUS detector is done by reconstructing high-momentum muons in the neutrino beam traversing the entire detector.

2.3.5 Trigger and data acquisition

The purpose of the trigger is to select neutrino interactions in the emulsion target and to synchronize the detector data taking with the time structure of the neutrino beam. A particular constraint is the fact that the optoelectronic system can only read out two events per neutrino spill. The system must also be flexible enough to accomodate other trigger types, besides those for the emulsion events. In particular, in each spill about a hundred neutrino interactions will occur in the calorimeter and spectrometer and a fraction of those is used in specific analyses, for instance the measurement of neutrino nucleon structure functions or the study of charm production through the dimuon signature. Once an event has been selected by the trigger system, the data acquisition (DAQ) is responsible for reading out the data stored in the front-end electronics, building the event — combining the data recorded by all subdetectors involved — and storing the data after validation. The DAQ should also allow for the continuous monitoring of the detector performance, for the logging of configuration changes or problems and for the interfacing to the user.

The trigger system [80] consists of a number of scintillator planes, marked A (anti-counter), V (veto), E (emulsion), T (timing) and H (hodoscope) in figure 2.1, each one made of two staggered planes. The E system covers an area of 150×148 cm^2, slightly larger than the 144×144 cm^2 area of the emulsion target. The T system is located behind the last fibre tracker plane and oriented horizontally, just as the H system. The 10 cm width of T and H scintil-

lator strips provides sufficient angular definition to separate forward tracks originating in the target from cosmic rays and particles from interactions in the floor as both of these make larger angles with respect to the beam direction. The V system is installed 2 m upstream of the target region, far enough to avoid vetoes due to backscattered particles from neutrino interactions in the target region by timing. The precision of 2 ns at FWHM per trigger scintillator in the T and V systems is sufficient to discriminate between forward and backward going tracks. The A system vetoes events originating in the concrete floor upstream of the detector; it is situated behind a 10 cm iron wall to absorb soft backscattered particles. The $1.5 \cdot 10^{-3}$ veto inefficiency provides sufficient rejection against the muon flux during neutrino beam of about 20/m^2 for every 10^{13} protons on target. The neutrino trigger for interactions in the target region is defined by a hit coincidence in E, T and H planes with no activity in the veto systems V and A. The combination of T and H strips requires consistency with a particle trajectory with $|\tan\theta| < 0.2$. Finally, at least one calorimeter plane or one spectrometer magnet scintillator must be hit.

The readout of all subdetectors is done in VME and CAMAC crates, controlled by a total of 35 VME processors running the OS-9 real-time kernel and communicating through a local VICbus with a bandwidth of 10 Mbytes/s. The DAQ back-end consists of 7 workstations and 2 VME boards running a standard Unix environment. Each of the OS-9 DAQ processes has been constructed to run as a finite state machine. The sets of possible states, the sets of allowed state transitions and the sets of actions to be performed on these state transitions are defined in Concurrent Hierarchical State Machines (CHSM, see [81]), an extra programming language on top of C++. Communication among the front-end OS-9 processes and the back-end Unix processes is based on a general-purpose, data-driven messaging system referred to as the dispatcher. All tasks for monitoring, logging, histogramming and interfacing to the user are implemented on the Unix back-end, using a combination of object-oriented programs, shell scripts and Tcl/Tk scripts for the graphical user interface parts. The DAQ is described in considerable detail in [82].

The measured dead time of the trigger system is less than 10 % per effective[2] spill of 4 ms and 10^{13} protons on target; about half of the dead time can be attributed to the optoelectronic system. For ν_μ charged current events passing the offline selection criteria, a trigger efficiency of 99 % has been measured. The observed rate of neutrino triggers is 0.5 events per 10^{13} protons on target, which implies an effective triggerable mass of 1700 kg, in good agreement with simulations. The fraction of those interactions originating in the emulsion target is $\sim$ 45 %. In the years 1994-1997, about $2.3 \cdot 10^6$ events were collected of which approximately $7.9 \cdot 10^5$ charged currents originating in the emulsion target. The data acquisition turned out to be both flexible and robust. With minor modifications, it has been used in a different experiment; for the 1998 data taking without the optoelectronic system, it has even been run unattended, maintaining an efficiency higher than 99 %.

[2] The effective neutrino spill length is defined as the length of an ideal, square-shaped neutrino spill which would give the same effective instantaneous rate.

3. NUCLEAR EMULSION AND AUTOMATIC SCANNING

Introduction

To claim that nuclear emulsion [1] has a long and respectable history would be an understatement. In 1896, Henri Becquerel placed uranium salts near to photographic plates wrapped in opaque paper and found that the plates nevertheless became fogged. From this seminal discovery, for which he shared the 1903 Nobel prize with Marie and Pierre Curie, emerged an entire new discipline of physics, now referred to as nuclear physics. In 1947, fifty years later but again using photographic plates, Cecil Powell and his collaborators made the first observation of the pion and was able to characterize the difference between pions and muons. He was duly awarded the 1950 Nobel prize for *his development of the photographic method of studying nuclear processes and his discoveries regarding mesons made with this method*. Again, this discovery proved crucial in the development of the discipline that has since become known as particle physics, or rather inaccurately as high-energy physics.

For just over a decade after the discovery of the pion, emulsion remained one of the prime tools in experimental particle physics. After that, studies of cosmic rays were largely replaced by detectors at particle accelerators and visual observation of a handful of events made way for systematic investigations of hundreds, occasionally thousands, of events recorded electronically or with the bubble chamber technique. Only for very few experiments did the benefit from emulsion's unrivalled spatial resolution outweigh the drawbacks connected to the manual scanning. Recently, this situation has reversed, largely due to two interrelated developments. First, the combination of nuclear emulsion with electronic detectors in so-called hybrid experiments offers the best of both worlds: triggering and kinematic measurements in the electronic detectors, high-resolution track and vertex information in the emulsion. Second, the development of automatic scanning techniques opens up the possibility to analyze several hundred thousand events in a reasonable time frame, and largely without laborious manual interventions.

In this chapter, we will describe the use of nuclear emulsion in the CHORUS experiment and the techniques applied in the automatic scanning of these emulsions. Most, if not all, of the issues involved are largely technical in nature. To better appreciate the different choices made in the case of the CHORUS experiment, the description of this specific case will be preceded by more general considerations. In section 3.1, we look at photographic emulsion itself as a particle detector: the properties of photographic emulsion for the recording of charged particle tracks, the advantage of hybrid experiments and the role of emulsion in

[1] Several authors insist on the denomination 'nuclear track emulsion' or 'nuclear research emulsion' to highlight the fact that the emulsion is no more or less nuclear than any other piece of matter. We believe this is sufficiently obvious and throughout will adopt 'nuclear emulsion' or 'photographic emulsion'.

those experiments, and the geometrical structure of the CHORUS emulsion target.

The rest of this chapter is entirely devoted to the discussion of automatic scanning. Typically, this involves computer-steered precision tables, microscope optics with high resolution, an image sensor, custom hardware for real-time image processing and computer infrastructure to store the results, possibly after on-line reconstruction. Figure 3.1 provides a rough sketch of the automatic scanning technique, which will in the following be elaborated in great detail. The emulsion plates have been exposed perpendicular to the neutrino beam and high-energy tracks are mostly in the forward direction. An optical system with high numerical aperture delivers an image corresponding to a thin slice of emulsion. Moving the objective along the optical axis, one obtains a sequence of images which are read out by a CCD camera. Dedicated digital signal processing (DSP) hardware is responsible for the recognition of grains within individual images. In turn, tracking algorithms on general purpose processors combine series of aligned grains into tracks, which are subsequently stored in a database for further, off-line analysis.

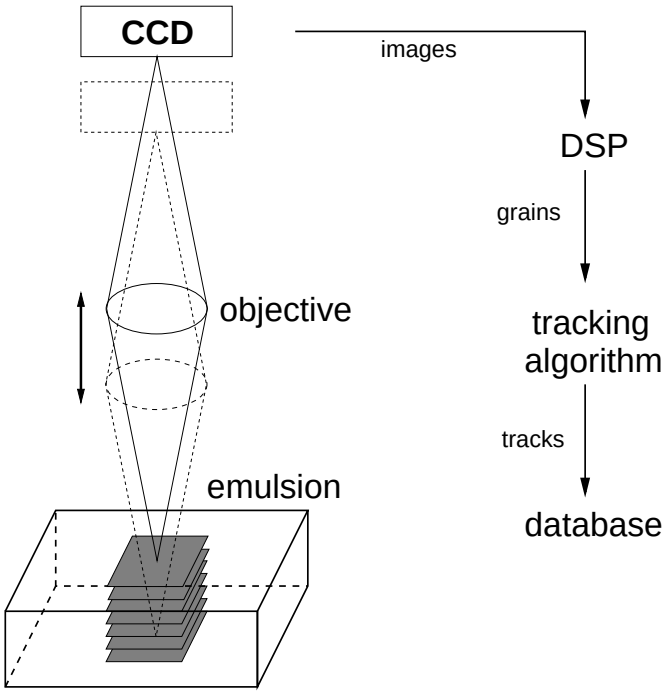


Fig. 3.1: Schematic view of automatic emulsion scanning in the case where the emulsion plate has been exposed perpendicular to the incoming particle beam.

A somewhat atypical part in an environment of particle physics is the microscope optics, described in section 3.2. The treatment of electronic signals is a problem which is encountered in just about any other experiment and only the type of our data is slightly less common: two-dimensional images from a camera rather than signals from a great many channels. However, it turns out that the algorithms and custom hardware developed for the multimedia industry can be advantageously applied to our case as well. These are generally referred to as digital signal processing (DSP), introduced in section 3.3. After these general considerations, section 3.4 will review the particular implementation for the CERN

microscopes: the properties of the optical system and the implementation of a digital filter algorithm, but also the infrastructure in the laboratory, the mechanics of the scanning tables, the imaging sensor used, and the computing hardware. Finally, in section 3.5, we will discuss how all of this hardware is put to use: how developed emulsion grains are recognized in individual images, how these grains are combined to tracks, and how these tracks enter into the CHORUS scanning procedures.

3.1 Nuclear emulsion

Silver halide emulsion of the type used for registering the tracks of charged particles consists of about equal parts by volume of halide crystals, a few tenths of a micron in diameter, and a matrix material which is chiefly gelatin. An ionizing particle on encountering a crystal may render it developable. After development, followed by fixing and washing to remove the undeveloped crystals, the gelatin is transparent. With a microscope the paths of charged particles that penetrated the emulsion are visible as trails of minute silver grains. A true three-dimensional image is produced. The paths of particles, outlined by silver, literally exist in space. Arguably, photographic emulsion was the first solid state detector. The above paragraph was taken literally from the first volume of Walter Barkas' 1963 classic 'Nuclear Research Emulsions' [83] but remains entirely valid to this day. As a matter of fact, this textbook remains the most comprehensive reference on the use of photographic emulsion in particle physics. The fact that it is no longer in print and available only from very few libraries demonstrates the extent to which these techniques had disappeared from the mainstream little more than a decade ago [2].

3.1.1 Image formation and photographic development

The composition of nuclear emulsion and the physical processes underlying the formation of a latent image and its development are largely identical to those for photographic emulsion. The chief differences are the larger thickness of the plates used for nuclear or particle physics research — up to several hundred μm as opposed to few μm in the case of photographic film — and a higher density of silver halide crystals. For photographic films, the diameter of the grains varies between a fraction of a micron and several micron. For nuclear track emulsion, it is consistently below 1 micron and can be as small as 0.1 micron. The physical processes which are merely outlined here are described in more detail in [84].

The latent image is produced when photons are absorbed to excite electrons from the valence to the conduction band, leaving behind positive holes. The electrons move freely until they encounter a local potential energy minimum where they can be trapped. Possible traps include defects in iodine-doped crystals, colloidal silver, and silver sulfide. The negative charge of the captured electron will attract interstitial silver ions and particles of silver could be formed. As little as three silver atoms constitute a latent image speck, and the silver collected in these specks is referred to as photolytic silver. The threshold energy of excitation from the valence band to the lowest unoccupied levels in the conduction band is about 2.5 eV in the case of silver bromide. The conduction bands of metallic silver lie about 1 eV below those of the silver in silver bromide and a silver speck can therefore act as a trap for electrons. Nevertheless silver atoms, either isolated or as aggregates, are unstable in the

[2] In the same vein, the fact that it is missing from most libraries where it does appear in the catalog is testimony to the recent revival of emulsion techniques.

crystal lattice and will collect only on external or internal surfaces, with a strong dependence on the availability of sensitizing substances in the surrounding medium or of crystal defects.

The quantum efficiency is defined as the number of free atoms of photolytic silver produced per photon above the excitation threshold. The maximum value approaches unity, rising with temperature but dependent also on the duration and intermittent character of the exposure. For ordinary photographic exposures to light of high intensity or short duration, the quantum efficiency falls due to the larger probability for recombination of electrons and positive holes. In the region where the quantum efficiency is constant, or equivalently where the amount of photolytic silver is proportional to the intensity and duration of the exposure, the reciprocity condition is said to be fulfilled. Reciprocity is observed down to exposure times of 0.01 s which immediately implies that the migration time, the time between the excitation of an electron and its condensation in a latent image speck, is of the order of 10^{-3} s.

In the case of charged particles traversing the crystal, the above processes are triggered by photons deriving from the particle's ionization. Since a single charged particle is involved, the exposure time will be no longer than 10^{-15} s, and the corresponding reciprocity failure greatly reduces the quantum efficiency. The quantum efficiency, defined as the number of photolytic silver atoms divided by the number of incident photons, is the appropriate measure of the sensitivity for the case of photographic exposure. But in the case of charged particles, where a large fraction of the photons that make up the ionization field have energies far above the excitation threshold, the measure of sensitivity should consider the efficiency in the conversion of the particle's energy loss into the production of conduction electrons. Especially for small crystals, a relatively small fraction of the energy loss effectively contributes to the formation of the latent image.

The production of a developable silver halide crystal requires only 30 eV of energy divided among a few photons. In contrast, the production of a developable grain along a particle track consumes one to two orders of magnitude more energy in the form of ionization, due to the combined effect of reciprocity failure and the higher energy per photon. The required ionization energie lies in the range from 500 eV to 3 keV, with the precise value depending on the sensitivity of the particular type of nuclear emulsion. Numerous substances are added to the gelatin to produce a increased photographic sensitivity: additional sulfur on top of the natural content in gelatin, dye molecules adsorbed to the silver crystal, or small concentrations of gold atoms. The sensitization procedures have a direct impact on the performance and applicability of nuclear emulsion. As a result, manufacturers are largely unwilling to publish the precise composition of their products. The binding potential for latent image specks, especially if they are composed of only a few atoms, is relatively shallow and with time, the latent image of a particle track gradually fades until the emulsion is processed. For specific applications, this might be an advantage as it can be used to eradicate old tracks from the emulsion. However, for long exposures, particular care must be taken that fading does not set in for the tracks under study. As illustrated in figure 3.2, the rate of fading increases with temperature and humidity; it is larger for small grain sizes. An important contribution is oxidation and fading is slower if the emulsion is deployed in an atmosphere of nitrogen or in vacuum.

Development is the process by which the latent image is converted into a stable image. The latent image specks act as germs for the formation of metallic silver spheres, whereas the silver halide crystals that are not part of the latent image are dissociated into their respective ions which are evacuated from the emulsion in the fixing stage. The various procedures

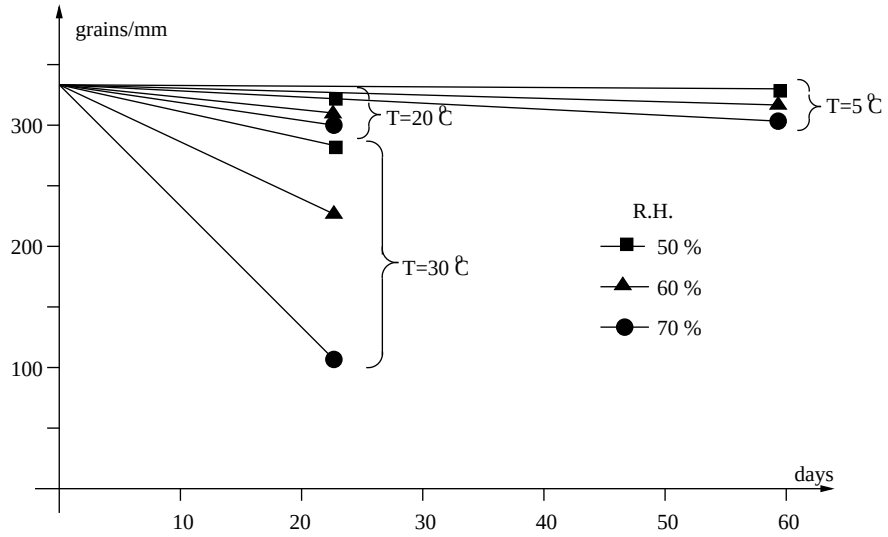


Fig. 3.2: Fading of the latent image for the Fuji ET-7B emulsions used in the CHORUS experiment. Fading is characterized by the number of developed grains per mm of track length as a function of the time elapsed between the exposure of the emulsion and its development. The development procedure is identical for all batches, but after exposure the various batches are stored in environments with different temperature or relative humidity (R.H.). At a temperature of 5° C, fading is negligible but strongly dependent on the relative humidity for long periods of time.

are again very similar to those used in photographic development, but the larger thickness of nuclear emulsion plates leads to greatly increased soaking times to allow the developing and fixing agents to penetrate into the emulsion. Furthermore, to increase the ion mobility, the emulsion is swollen by the addition of glycerin. Generally speaking, the developer is a reducing agent: due to its lower electron affinity as compare to the positive silver ions in the silver halide crystal, it donates one of its electrons to silver ions which in turn accumulate around image specks of metallic silver causing these to grow. The development rate is determined by the developer's reduction potential relative to the one for silver ions, usually strongly dependent on the pH of the medium. Various substances, for instance alkalis or buffers to stabilize the pH, are usually included in the developing bath to stabilize the development rate. In parallel to the chemical development based on a reduction reaction, physical development may be used. Physical developers contain silver or perhaps other metal ions that directly deposit on the latent image centers. Typically, nuclear emulsion work relies on combined chemical and physical developers. For instance, sulfite and bromide are solvents of the silver halide, forming complexes from where the silver ions are reduced to metallic silver precipitating in the gelatin and plating out on the silver grains. It is worth noting that the development process is based on a delicate balance between different reaction rates, the formation of metallic silver aided by the presence of latent image centers. Sufficiently prolonged action of a strong reducing agent develops all the silver halide micro-crystals, including those that a priori are not part of the latent image.

3.1.2 Composition of nuclear emulsion

The silver halide is usually in the form of silver bromide crystals with lattice defects induced by a small amount of iodine. As described above, these defects play a role in the formation of the latent image. The density of the silver halide is of the order of 6.5 g/cm^3. The grain diameter depends on the type of emulsion, within the range from 0.05 μm to 0.5 μm. However, this merely indicates the mean value and the actual size of individual grains is usually well described by a normal distribution, with a width about 10 % to 20 % of the mean value. Generally, the sensitivity goes up with increasing grain size but sensitization procedures can make up for this. Emulsion recently developed by Fuji yields well over fifty developable grains per 100 μm track length for grain sizes as small as 0.1 μm. To improve the precision and efficiency of track measurements, one strives not only for a uniform grain size, but also for a good distribution of the grains within the emulsion volume.

Photographic gelatin is made from clippings of calf hides, ears and certain other pieces or from pig skin and bone. The value of gelatin as a photographic emulsion matrix stems from a unique combination of properties. It shows exceptional mechanical strength, it is stable, components of it are photographically sensitizing, it permits penetration of solutions, it keeps grains dispersed to avoid clumping and to ensure a uniform distribution throughout the volume, it is strongly adsorbed to the silver halide crystals and it allows the growth of a grain during emulsion manufacture provided that ammonia or soluble halide is present to act as a solvent for the silver halide. Unfortunately, gelatin is particularly hygroscopic, causing it to swell strongly with increasing humidity and to distort on drying. Furthermore, it is an animal product, not completely under control with respect to its composition and subject to damage by microorganisms. This latter point remains an issue after development. In particular, care must be taken in the storage and handling to prevent mould from developing, especially acute because an environment with relative humidity between 50 % and 70 % is required to prevent the emulsion from drying out which would result in cracks, fissures and the occasional detachment of the emulsion from its support plate. So far, attempts to use synthetic materials such as polyvinyl alcohol instead of the animal gelatin have met with relatively little success.

Table 3.1 indicates the chemical composition for two types of emulsion, along with their density. As previously mentioned, manufacturers are generally reluctant to reveal the precise composition and the knowledge is at best partial. For instance, the values reported for Fuji ET-7A only add up to a total density of 3.55 g/cm^3, or about 95 % of the actual density. As a matter of fact, the manufacturer reports small concentrations of heavy elements as well, still only adding up to about 97 %. We have specifically chosen to display the composition of Fuji ET-7A gel, first introduced in 1947, because the emulsion gel used in the CHORUS experiment, Fuji ET-7B, is a direct successor to Fuji ET-7A. Meanwhile, a gel of type Fuji ET-7C with smaller grains, more uniform in size and more uniformly distributed in the volume, has been successfully used in a pion test beam at CERN and in the DONUT experiment [85] at Fermilab which has been the first to directly observe tau neutrino interactions.

The "standard" emulsion is *defined* to have a density of 3.815 g/cm^3, with its composition derived from the detailed figures available for Ilford G.5 emulsion at 60 % which has approximately this density and used to be one of the most common types in the heyday of emulsion experiments. It should be stressed that the composition of any type of emulsion is strongly dependent on its environment, in particular due to the hygroscopic nature of gelatin. For instance, if emulsion is placed in vacuum, the water gradually diffuses out

Element	concentration (g/cm^3)	
	"Standard"	Fuji ET-7A (65 % R.H.)
Ag	1.8088	1.74
Br	1.3319	1.26
I	0.0119	0.049
C	0.2757	0.278
H	0.0538	0.034
O	0.2522	0.161
N	0.0737	0.010
S	0.0072	0.015
	3.815	3.75

Tab. 3.1: Chemical composition of nuclear research emulsions, indicating for the most abundant components their concentration in g/cm^3. The bottom line indicates the total density of the emulsion at the time of exposure. The types of emulsion that appear in the table are described in the text.

of it, and its mean density increases. Also more intricate interaction with its environment may alter the composition and density of nuclear emulsion. We will henceforth assume the CHORUS emulsion to be sufficiently well described by "standard" emulsion.

3.1.3 Contents and characteristics of emulsion images

When viewed through an optical microscope using transillumination, the emulsion image appears as a number of tiny black spots on a more or less uniformly white background. The white background simply reflects the fact that the gelatine and the mechanical support of the emulsion plate are transparent in most of the optical spectrum. The metallic silver spheres, resulting from the development of those silver halide crystals that form the latent image, are reflecting. The black spots are the shadows of those silver spheres that lie within the depth of focus being viewed. Ideally, all of them mark the passage of a charged particle and tracks appear as a trail of aligned points, possibly curved if the emulsion was exposed in a magnetic field. Tracks parallel to the emulsion surface are seen within a field of view. For tracks perpendicular to the surface, typically only one or two grains fall within the depth of focus. Moving the focus continuously up and down, subsequent grains appear one after another and one can mentally picture the track in three dimensions. Depending on the type of emulsion, a minimum ionizing particle gives rise to 300 or more grains per mm traversed. The metallic spheres that are produced during development tend to be larger, by up to a factor of two, than the corresponding silver halide crystals, but this growth is largely uniform. When interpreting the barycenter of a grain as a position measurement for the track, the resolution is given by the crystal diameter in the undeveloped emulsion, ranging from 0.05 μm to 0.5 μm. Such a sensitivity and resolution are unrivalled, but should not be taken at face value. First, the large majority of grains in emulsion are not part of the tracks one is interested in, even more so for long exposures. Instead, they are randomly developed grains which constitute *fog*. Second, the intrinsic resolution is only meaningful to the extent that measurements can be made in the same reference frame where the emulsion has

been exposed. Shrinkage during development as well as distortions make this well nigh impossible.

Fog. As mentioned in section 3.1.1, the development process is always differential: the grains containing latent images caused by an ionizing particle are chemically reduced more rapidly than the unaffected grains. All will be developed, however, under the sufficiently prolonged action of a sufficiently active reducing agent. Some such random grains are always present and constitute fog. They do not differ in any discernible way from the track grains. Low energy electrons due to natural levels of radioactivity will produce single-grain tracks further contributing to fog. Various effects lead to a larger fog density at the surface: because of the imperfect penetration of chemical compounds, the surface layer is usually most strongly developed, it may be somewhat light-struck, it may have been affected by external chemical agents, and it is subject to abrasion often leading to dense streaks of developed grains.

Shrinkage. The silver halide crystals occupy about half of the emulsion volume, but only a tiny fraction of them contributes to the latent image. The remainder are removed during processing leading to a substantial reduction in the emulsion thickness, the lateral area is unaffected because it is determined by the size of the mechanical support plate. The shrinkage factor, defined as the emulsion thickness at the time of exposure divided by the thickness at the time of scanning, is typically about two. After development the thickness may still vary with time, in particular if equilibrium with respect to the water content of the atmosphere has not been attained. During the scanning for instance, the top surface is typically insulated by a thin layer of immersion oil whereas the bottom surface is kept under vacuum leading to dessication. Under such conditions, the shrinkage factor may vary along the depth.

Distortions. Emulsion is mechanically very strong but essentially soft and pliable, subject to mechanical alterations. The source of distortion most commonly cited is the release of tensions existing at the time of exposure. During the processing, the emulsion is swollen to facilitate the mobility of chemical agents. This causes the tensions built up at the time of pouring to relax, so that even if no new distortions were introduced, the release of the old ones will have the same effects as the introduction of new ones. New distortions are introduced as well because the emulsion tends to swell in all directions whereas the one surface mounted to the supporting plate is constrained to retain its dimensions, leading to further shearing stresses.

Figure 3.3 schematically represents the combined effect of shrinkage and distortion. The distortion shown in this figure can be described by a lateral displacement which behaves quadratically as a function of depth. This C-shaped distortion is by no means the only possible type, but it does cover the most typical cases and anyhow may serve as a suitable approximate description for more complicated types of distortion. It has been confirmed by measurements on the deviation of high-energy tracks and can also be understood intuitively. The stresses introduced may stem from temperature gradients during the drying after processing, from swelling during the development, from the gravitational field parallel to the emulsion surface when it is hung to dry after development, etcetera. All of these have in common that the stress vector will be approximately parallel to the surface. However, the surface mounted on the support plate must at all times retain its dimensions, so that one

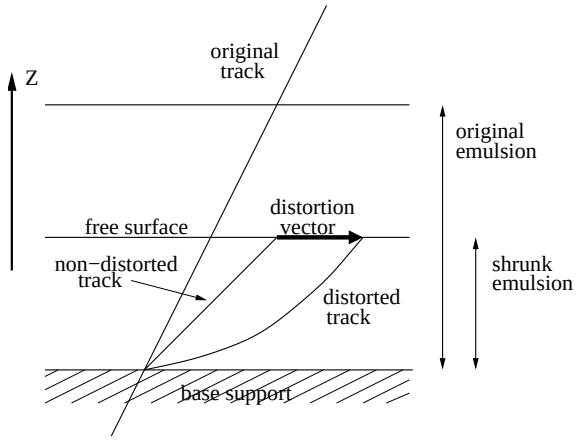


Fig. 3.3: Distortion on one side of an emulsion plate, the mechanical support plate is located at the bottom of the image. With respect to the original track, the distorted track displays the combined effect of longitudinal shrinkage and a C-shaped distortion.

would expect the stress vector to grow as one moves away from the constrained surface. The external surface is not constrained and down to a certain depth will show more or less the same shear. This can be worked out in detail and indeed leads to a quadratic profile. The net result is that the position at the support plate is free of distortion, whereas the slope of the track is most accurately measured at the surface, where it is only affected by the shrinkage factor. Alternatively, using double-sided emulsion coated on a thick plastic support plate, a precise direction is defined by the two points at the support plate, an idea first suggested in [86] and used for the CHORUS interface sheets which are described in section 3.1.5.

3.1.4 Hybrid experiments

Nevertheless it should be pointed out that emulsion has no great advantage when one is looking for a particle of predicted properties. It is of greatest use for discovering utterly new things, the anomalous behavior of which often can be recognized from a single event. As examples, W.H.Barkas quotes the discoveries of π^+, π^-, π^0, K^+, K^-, and $\bar{\Lambda}$, all made using nuclear emulsion. Again, this remains very much true today. A recent example of how telling even a single event may be is the observation by the CHORUS experiment of neutrino-induced diffractive production of D_s^* followed by the decay chain $D_s^* \to D_s^+\gamma$, $D_s^+ \to \tau^+\nu_\tau$ and $\tau^+ \to \mu^+\nu_\mu\bar{\nu}_\tau$, reported in [87]. Even more recently, the DONUT collaboration published evidence for the first direct observation of tau neutrino interactions, based on four events observed in emulsion [85]. While this may not classify as utterly new things, it does demonstrate the type of detailed studies which can be undertaken using nuclear emulsion. However, occasionally the problem of discovering utterly new things is that of convincing the scientific community, especially if the theoretical framework is lacking at the time of finding. A sad but illustrative example is the observation of charmed particles by K.Niu and co-workers in 1971, two years before B.Richter and S.Ting made the discovery for which they were awarded with a Nobel prize only three years later. A posteriori, it seems clear that the event reported in [88] shows the production of a pair of charmed particles. From the emulsion only, it was possible to mea-

sure both the mass and the flight length of both particles but nevertheless this remained an isolated event which received hardly any further attention. The reason for that may well be the lack of additional information that might corroborate this event.

Precisely because they provide such additional information, hybrid experiments have led to a recent revival of the emulsion technique. In these experiments, the nuclear emulsion target or vertex detector is followed by conventional, electronic detectors to provide measurement of energies and momenta, as well as particle identification, all of these to a degree of precision that cannot be achieved using emulsion alone [3]. Electronic tracking devices immediately downstream of the emulsion permit the connection of tracks found in emulsion to those found in the electronic detectors. This provides a time stamp, not available from the emulsion which simply integrates all charged particle tracks traversing in the period between pouring and development. In fact, one typically proceeds in the other direction: for tracks reconstructed in the electronic detectors, emulsion scanning starts from their predicted impact point in the emulsion. With such a technique, only a minute fraction of the emulsion volume must be inspected leading to an enormous increase in the number of events that can be analyzed for a given scanning capacity. In fact, as early as 1965 an experiment along those lines was performed in the PS neutrino beam at CERN and found seven neutrino interactions [89, 90]. Nevertheless, peering through a microscope remained a laborious procedure and in the era of large samples collected at ever more intense particle accelerators, the emulsion technique was no match for the electronic detectors which could be analyzed automatically.

The first large-scale hybrid experiment was WA17 at CERN, exposing 31.5 l of emulsion in front of BEBC, the Big European Bubble Chamber, and locating 169 neutrino charged current interactions among which eight candidates for charmed particle production and decay were found [91]. The breakthrough came only with the development of first semi-automatic, later fully automatic scanning techniques. Previously, emulsion plates were typically oriented parallel to the beam direction such that a long track section could be seen within a field of view. When using the emulsion as a tracking device rather than as a visual detector, the plates were oriented perpendicular to the beam. In the E531 experiment [53] at Fermilab, 58.6 l of emulsion were exposed to a neutrino beam produced by 350 GeV and 400 GeV protons. To evaluate the advantages and drawbacks of different scanning procedures, part of the emulsion was oriented parallel to the beam, part of it perpendicular to the beam. Among 3886 located neutrino interactions, 121 charmed particles were identified, permitting a measurement of the production cross section [54] as well as their lifetime and decay modes [55]. In addition, this experiment could be considered the direct predecessor of the CHORUS experiment, having performed a search for $\nu_\mu \rightarrow \nu_\tau$ oscillation in the same kinematic region [59]. The WA75 experiment [92] at CERN, drawing together the groups which had participated in the WA17 and E531 experiments, aimed to search for beauty particles among the interactions of 350 GeV negative pions in an 80 l emulsion target. Out of about 10,000 located interactions, one showed a clear example of beauty pair production and the subsequent decay into charmed particles. With the pioneering work of the Nagoya laboratory, fully automatic scanning became a reality [93] and truly systematic studies became possible. The E653 experiment [94] at Fermilab performed a study of hadronically produced

[3] Strictly speaking, emulsion can provide both particle identification and kinematic measurements to a high degree of precision. The use of the emulsion cloud chamber (ECC) technique in the OPERA experiment is a case in point. However, that does require massive amounts of emulsion and a scanning capacity that even with the most advanced techniques can only treat a limited number of events.

heavy flavour states using a 71 l emulsion target, exposed to an 800 GeV proton beam and a 600 GeV pion beam. In total, more than 50,000 interactions were located; the proton exposure yielded 146 identified charm events [95], the pion exposure yielded nine hadronically produced beauty pairs [96]. A review of hybrid experiments, focusing in particular on neutrino experiments, is given in [97].

3.1.5 The CHORUS emulsion target

The development, described in the previous section, towards ever more massive emulsion targets in which an ever greater number of events could be located with the aid of automatic scanning culminated in the CHORUS experiment. More than 400 l of nuclear emulsion served as the active target for a high-energy neutrino beam, in two exposures of two years each, the first one in 1994 and 1995, the second one in 1996 and 1997. To date, this represents the largest emulsion target ever built, and operated during an exceptionally long exposure. A detailed description of this target, as well as of the CERN pouring and processing facilities specifically upgraded for this purpose, can be found in [98]. Here, we will highlight only the main features, with particular attention for those aspects that are relevant to the automatic scanning described below.

The main goal of the CHORUS experiment is a search for $\nu_\mu \rightarrow \nu_\tau$ oscillations through the appearance channel, exploring oscillation probabilities down to 10^{-4}. To collect the required number of events, considering also the small neutrino interaction cross section, clearly requires a heavy target. On the other hand, if the scanning is to be guided by the reconstruction of tracks in electronic detectors downstream of the emulsion target, it should be sufficiently thin. Otherwise secondary activity, in particular the development of electromagnetic showers, will greatly complicate the task of the electronic tracking detectors. The effective thickness, from the point of view of the electronic tracking detectors, was reduced by segmenting the target and interleaving emulsion modules with tracker planes. The target mass of slightly less than 800 kg is built up of four stacks, each with a thickness of 2.9 cm and covering an area of 1.4×1.4 m^2. The first and third stack are followed by one hyperplane of scintillating fibre trackers, the second and last stack by three hyperplanes and each hyperplane consists of four planes to provide measurements in the two orthogonal directions and two stereo projections tilted by 8° with respect to the orthogonal directions, as shown in figure 2.4.

To facilitate the emulsion handling during pouring, development and scanning, but also to improve accuracy, each stack is laterally divided into eight modules with an area of 0.71×0.36 m^2. In turn, the 2.9 cm total thickness for each module corresponds to 36 plates, piled up and kept under vacuum to maintain their relative position throughout the exposure. Each plate consists of two 350 μm layers of emulsion, glued on either side of a 90 μm plastic base which acts as mechanical support. The support plate is referred to as the base.

The emulsion technique is unique in its ability to reconstruct the topology at an exceptional degree of detail. However, the unavoidable distortions discussed in section 3.1.3 might preclude precise angular measurements, especially if these need to be interpreted in the coordinate system corresponding to the time of exposure. On the other hand, the large lever arm of the fibre tracker permits the measurement of track slopes with errors that can be as small as 1 mrad. Compared to the 200 μm error for the reconstructed transverse position at the emulsion surface, the reconstructed slope is more distinctive with respect to the background of emulsion tracks accumulated during the two years of exposure. To achieve a

similar slope resolution, whilst at the same time maintaining the μm resolution in position typical of emulsion, emulsion interface sheets were introduced. These consist of two layers of emulsion, no more than 100 μm in thickness, coated on both sides of an 800 μm plastic base. Combining the sub-μm resolution for the two track points at the base, neither of them affected by distortion, and the 800 μm lever arm, one obtains the required resolution for the track slope. To further simplify the task of distinguishing tracks predicted by the electronic detectors from the background tracks accumulated during the exposure of the emulsion target, the interface sheets were exchanged more often than the target plates. One interface sheet, the "special sheet" (SS) was placed just downstream of the 36 target plates and part of the same vacuum package. Two more sheets, the "changeable sheets" (CS) were placed 40 and 50 mm further downstream to be as close as possible to the fibre tracker planes. The target plates in each module were used for two consecutive years, whereas the special sheet was replaced at the end of the first year. The changeable sheets were replaced at regular intervals throughout the year. Experience has shown that this was an unnecessary precaution: seven sets were used in 1994, three in 1995, two in 1996 and a single one in 1997.

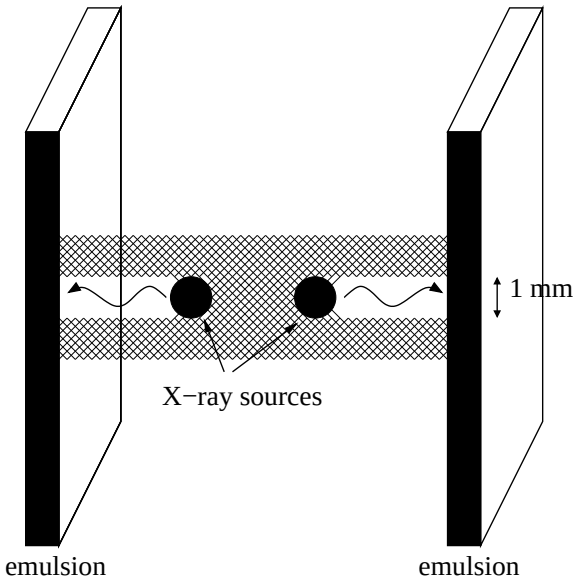


Fig. 3.4: Schematic representation of an X-ray gun contained in the honeycomb spacer between two interface emulsion sheets.

The intrinsic emulsion resolution is meaningful only to the extent that a precise intercalibration can be performed between subsequent emulsion plates and between the emulsion target as a whole and the electronic detectors. Furthermore, throughout the exposure the entire setup must be mechanically stable within a few microns over distances up to tens of centimetres. Each module constitutes a vacuum package, pressing the 36 target plates and the special sheet plate together and maintaining their relative position. The entire target region — the emulsion target, the fibre trackers, the hadronic spectrometer magnet described in section 2.3.2, and the trigger planes described in section 2.3.5 — is placed in a cool box where the temperature is stabilized at 5.0 $\pm$ 0.5 °C and continuously monitored. The changeable sheets are mounted on frames where beta-ray sources can be embedded, directed towards the fibres such that a measurement of their position can be used as a first indication of the

position of the plates in the experiment. Further, more precise positioning is achieved using alignment tracks located on the basis of this first approximation. Honeycomb spacers separate the target module from the changeable sheets. Included in these spacers are ^{55}Fe X-ray sources to mark corresponding points on adjacent sheets, as indicated in figure 3.4. Each interface plate is marked by 15 such sources of 6 keV photons, which traverse the paper surrounding the emulsion and induce a dense, black spot 1 mm in diameter at the emulsion surface after development. The center of such X-ray marks can be determined to within better than 50 μm. Finally, to intercalibrate the target plates in a coordinate system where the large-scale distortions can be corrected for, a grid of reference points is printed on these before processing them. The grid consists of small black dots 50 μm in diameter, spaced by 19.5 mm in two orthogonal directions; in the following, these will be referred to as fiducial marks. To produce them, the emulsion is placed over a black film with transparent dots and exposed to a flash of white light shone from below, through the film.

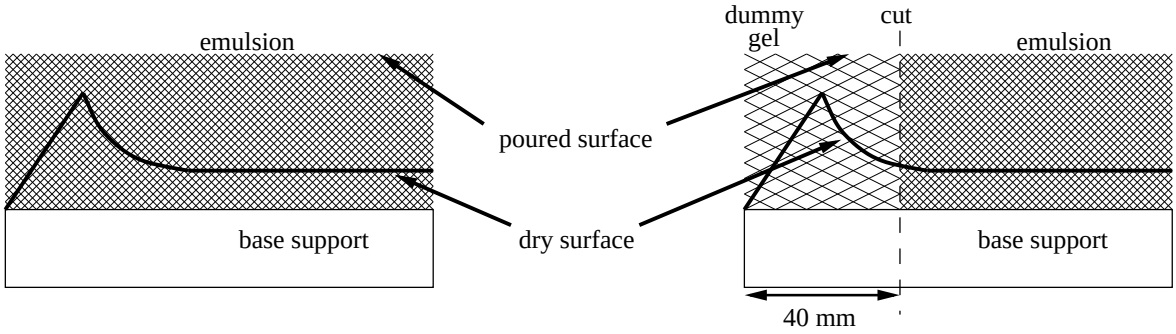


Fig. 3.5: The reduction of edge distortions using a frame of dummy gel.

To keep the background low and to have the possibility of producing sheets during the run, the entire set of emulsion was poured at CERN just before exposure and processed immediately afterwards, again at CERN. For this purpose, the existing emulsion processing facilities, built for the WA75 experiment [92], were entirely refurbished. The emulsion gel is Fuji ET-7B at a concentration of 80 %, poured on 90 μm thick tri-acetate cellulose (TAC) base plates for the target sheets, 800 μm thick Plexiglas plates for the interface sheets. After pouring, the gel is about ten times thicker than the final emulsion due to the water content, 90 % in volume, which is removed at the drying stage. Drying is performed slowly, at 80 % relative humidity, to minimize the accumulation of stresses that will later show up in the form of distortions, as explained in section 3.1.3. The borders dry first and are most severely affected. To reduce heavy distortion at the edges, a 4 cm wide frame of dummy gel, containing inexpensive pure gelatine and silicate, is poured all around after the actual gel before drying. After drying, the dummy gel is cut away as indicated in figure 3.5.

The two most critical aspects of a very long exposure are the fading of tracks from neutrino events and the accumulation of background, both due to fog and to passing tracks not related to any neutrino event. Fading, the erasing of the latent image with time, can be significantly reduced by lowering the temperature. At 5 °C, corresponding to the exposure conditions, the rate of fading is negligible for an exposure of up to two years, as indicated in figure 3.2. The fog, random-developed grains dispersed in the emulsion, increased with time but only rarely exceeded 5 grains per $10 \times 10 \times 10$ μm^3 in the developed emulsion. Background tracks may be due to stray particles from nearby particle beams, to cosmic rays

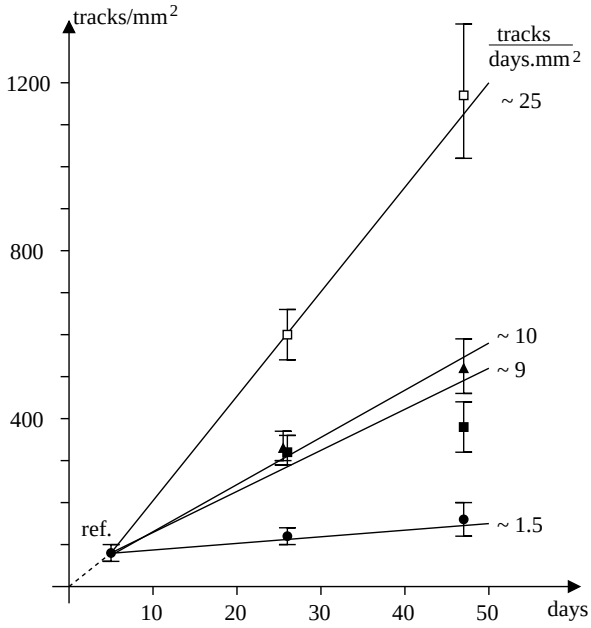


Fig. 3.6: Density of background tracks due to cosmic rays as a function of exposure time. The upper two curves correspond to unshielded emulsion, either at ground level (open squares) or at the ALEPH pit (triangles). The lower two curves represent the cases of interest to the experiment: shielded emulsion at ground level as it was the case during the exposure (full squares), or shielded emulsion at the ALEPH pit 150 m underground were the plates were stored during the winter shutdown between two consecutive exposures (circles).

or to ambient radioactivity. During the 1994 and 1995 exposures, muons from the nearby SPS West Area test beams X7 and X9 represented a significant background, but concentrated in a relative narrow angular region situated at -8 mrad in the vertical plane and at -77 mrad and -120 mrad in the horizontal plane for the X7 and X9 beams, respectively. These tracks have been used for the alignment of emulsion plates, but an angular cut of 80 mrad around this particular direction was imposed in the analysis of neutrino events. For the 1996 and 1997 exposures, the background level was negligible owing to the different settings of these beams. For cosmic ray particles, the direction is peaked around the vertical with an intensity falling of as $I(\theta) \sim \cos^2\theta$, with a total flux over a vertical surface of $\sim 90/\mathrm{m}^2$ s. Considering the fact that only horizontal tracks are of interest to the experiment, this does not lead to a significant background. However, an excessive pile-up might as in the case of fog compromise the track visibility. To minimize this effect, the stacks were stored underground during the six month shutdown between the exposures of two consecutive years. At this location, the pit of the ALEPH experiment at CERN, the cosmic ray flux is reduced to less than 0.1 % of its value at ground level. As a further precaution, the emulsion was protected by a cooled iron shield. The ensuing reduction in the number of cosmic ray tracks is shown in figure 3.6. Finally, ambient radioactivity contributes essentially through low-energy gamma rays, leading through Compton- or photo-electrons to sparse grains or heavily scattered near-minimum ionizing tracks. Again, this cannot lead to fake tracks but might decrease the visibility. To limit the

effect, the target area was protected by ≥ 8 cm thick iron plates on the floor and on the walls around it.

3.2 Microscope optics

The key to unlock the emulsion's exceptional spatial resolution is the appropriate optical system. However, at the outset, we would like to stress that the purpose of the optical system is not to determine the properties of grains. The shape and appearance of individual grains are determined by chemistry and optics; for the reconstruction and measurement of charged particle tracks, the only relevant parameter for each grain is its position. In the CHORUS experiment, the grain dimension does not play any role since the tracks of interest are all at the minimum of ionization. The optical system should yield an image of sufficiently high contrast such that grains can readily be identified and of sufficiently high resolution, both transversely and axially, such that their position can be accurately determined. For automatic scanning, the image is captured with a camera placed in the primary image plane of the objective, an eyepiece is generally not used.

Within the CERN scanning group, Piero Zucchelli was the first to investigate how the optical quality affects the overall performance of a nuclear emulsion scanning system [99]. He also noted that the optical conditions are very much unlike those encountered in the majority of optical microscopy applications. Light microscopes are generally used to generate a diffraction limited, high magnification image of a small object. In contrast, the image of nuclear emulsion — because the grain size is comparable to the wavelength of visible light — is essentially in the diffractive regime, the magnification in the range from $20\times$ to $60\times$ is relatively small and the object, in the sense of field of view, is comparably large. Another fairly unusual requirement is the need for a three-dimensional image across a large depth. The thickness of emulsion plates can be as large as 1 mm and the optical system must be able to visualize thin slices at any depth.

In this section, we will introduce the basics of light microscopy for the specific case of a conventional optical system used in emulsion readout. First, the intrinsic limits on the resolution are reviewed, both in the transverse plane and along the axis. After that, optical aberrations in a monochromatic system are described. Finally, we discuss the principle of Köhler illumination. These topics are sufficient to appreciate the particular features of the optical system, described in section 3.4.3, which was specifically designed for the microscopes at CERN. A very detailed and complete introduction to the field of microscopy, both concerning the underlying theoretical principles and down-to-earth problems in the daily usage of microscopes, can be found in [100]. For a more concise introduction, with specific emphasis on some of the recent developments, we refer to [101].

3.2.1 Transverse resolution and numerical aperture

After development, the average diameter of grains in the CHORUS emulsion is about 0.8 μm. For any wavelength in the visible spectrum, the image of a grain will be a diffraction pattern produced by a sphere of metallic silver. Therefore, the approximations of geometrical optics, so-called ray tracing, cannot be used and one must account for the wave nature of light. The discussion is slightly simplified by the fact that the cross sections for absorption, diffusion and dispersion of photons are essentially flat across the visible spectrum, both for silver

spheres of 0.8 μm diameter and for gelatin. This is confirmed by the gray appearance of emulsion plates. However, for wavelengths below 380 μm, the transmission drops rapidly and emulsion can no longer be considered transparent in the ultraviolet or the UV edge of the visible spectrum.

Presumably, the refraction of light by lenses or water filled glass spheres was known in ancient times. Nevertheless, the earliest writings describing optical experiments using segments of glass spheres date from around the year 1000 AD, by Ibn al-Haitam. In 1267, Roger Bacon pointed out that such spherical segments may be used as magnifying glasses. The microscope itself was invented in Holland in the early seventeenth century, possibly by Hans and Zacharias Jansen from Middelburg, possibly by C.J.Drebbel in Alkmaar. Around the same time, Galilei introduced the telescope. For several centuries, both instruments remained essentially unchanged apart from mechanical improvements, largely because a detailed understanding of the underlying optical principles was lacking. Only in the early nineteenth century did J.von Fraunhofer perform systematic studies in that direction. A comprehensive theory of image formation was finally formulated by E.Abbe in 1886, based on a series of beautiful experiments. It immediately lead to the development of homogeneous oil immersion and of apochromatic objectives, objectives which are simultaneously corrected for three different wavelengths. Since then, the mathematical problem of image formation accounting for the wave nature of light has been analytically solved for a number of idealized cases. To describe more realistic situations, one typically uses a combination of geometrical optics, ideas borrowed from problems that have been analytically solved, and numerical solutions.

The Abbe principle states that a minimum of two successive orders of diffracted light need to be captured for a particular spacing to be resolved. It follows that the resolution limit for an optical system with opening angle σ is given by

$$b = \frac{\lambda}{\sin\sigma} \ .$$

In a medium with refractive index n, the wavelength λ is modified by a factor 1/n and the resolution limit becomes

$$b = \frac{\lambda}{n\sin\sigma} = \frac{\lambda}{\mathrm{NA}_{obj}} \ .$$

The combination n $\sin\sigma$ is called the numerical aperture; it is the single most important quantity to characterize the resolution of an optical system. The above considerations apply to the case of illumination parallel to the optical axis. In general however, the aperture of the illumination system is smaller than that of the imaging system. With NA_{cond} defined as NA_{obj} to express the largest inclination of incident light [4], the resolution limit for the combined system becomes

$$b = \frac{\lambda}{\mathrm{NA}_{obj} + \mathrm{NA}_{cond}} \quad \text{if} \quad \mathrm{NA}_{cond} \leq \mathrm{NA}_{obj} \ .$$

3.2.2 Axial resolution and depth of focus

For an infinitely small point of light imaged by a perfect objective lens, the image in the intermediate image plane consists of a Fraunhofer diffraction pattern formed by the exit

[4] The notation NA_{cond} refers to the fact that the microscope illumination system typically contains a condenser lens; this will be described in more detail in 3.2.4.

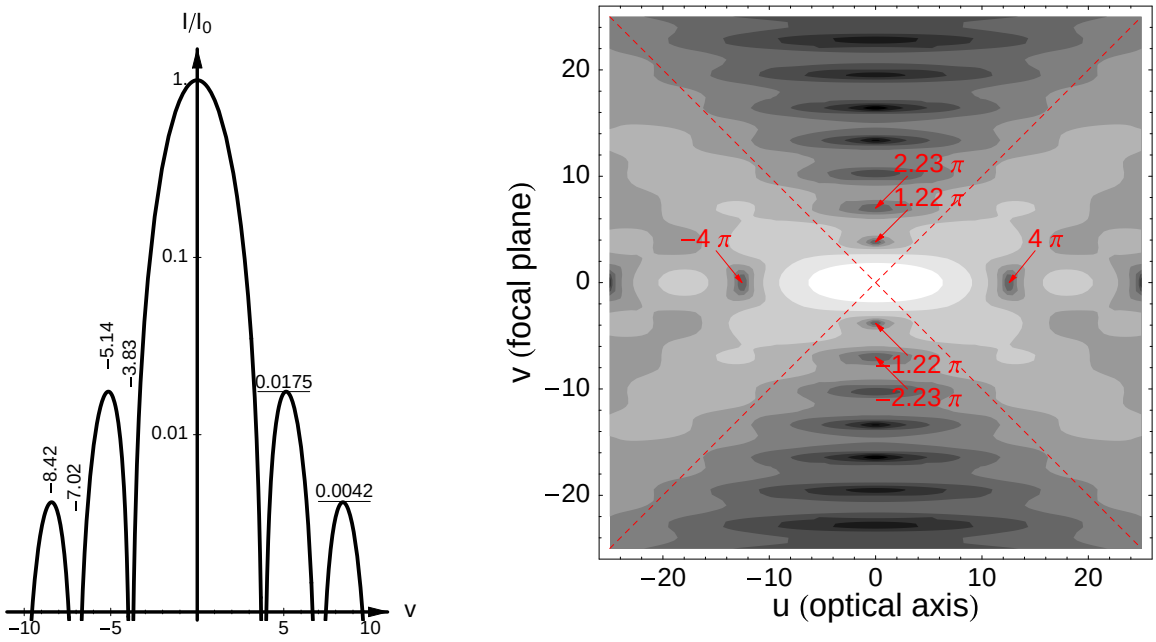


Fig. 3.7: Left panel: Airy diffraction pattern. The intensity for the image in the intermediate image plane of an infinitely small point of light in the object plane, as a function of the transverse distance from the optical axis expressed in units of NA/λ, relative to the intensity on the optical axis. Right panel: Intensity point spread function. Representation of isophots, contours of equal intensity, for the three-dimensional image of a point source of light near the focal plane in object space. The coordinates u and v measure the distance from the focal point in image space: along the optical axis and in the focal plane, respectively. The units for u and v are described in the text. The isophots are drawn assuming a log scale.

pupil of the objective lens from which spherical waves converge to the focal point. The image which results is called the Airy diffraction image. The intensity pattern as a function of v, a measure of the diffraction angle, is found to be

$$I(v) \;=\; I_0 \left(\frac{2J_1(v)}{v} \right)^2 \quad ,$$

where J_1 is the first order Bessel function. In terms of the numerical aperture of the system NA, the magnification M and the distance r_i from the optical axis in the image plane v is given by

$$v \;=\; 2\pi \, \frac{\mathrm{NA}}{M\,\lambda} \, r_i \quad .$$

As the distance in the image plane r_i is equal to the distance in the object plane r_0 times the magnification M, the definition of v can be rewritten as

$$v \;=\; 2\pi \, \frac{\mathrm{NA}}{\lambda} \, r_0 \quad .$$

The left panel of figure 3.7 shows the intensity distribution as a function of v in units of NA/λ. The problem is rotationally symmetric so this intensity distribution holds for any direction perpendicular to the optical axis. The central peak corresponds to a central bright disk in the diffraction image and is known as the Airy disk. The first minimum at $\pm$ 3.83 defines the radius of the Airy disk; in object space units it is given by

$$r_{Airy} \;=\; 0.61\,\frac{\lambda}{\mathrm{NA}} \quad .$$

When there exist two equally bright, self-luminous points of light separated by a small distance d in object space, their diffraction images lie side-by-side in the image plane. As d becomes smaller so that the first minimum of one diffraction image overlaps with the central maximum of the neighbouring diffraction image, their sum measured along the axis joining the two maxima still contains a dip of 26.5 % of the peak intensities. The condition $d \geq r_{Airy}$ is called the Rayleigh criterion. Once d becomes less than this distance, the two diffraction images reach a stage where instead of a small dip, their sum shows a single, broad peak. This happens at $d = 0.78 r_{Airy}$ and is called the Sparrow criterion. These considerations for two point sources of light apply equally well to two absorbing dots, assuming that they were illuminated incoherently. In general, the illumination will be at least partially coherent and the numerical aperture of the illuminating system will enter the above expressions.

The two-dimensional Airy pattern that is formed in the image plane of a point object is actually a cross section of a three-dimensional pattern that extends along the optical axis of the microscope. Analytical expressions to describe this pattern were first calculated by Linfoot and Wolf in 1952 [102]. In addition to the dimensionless variable v defined above to express the distance from the optical axis in the focal plane, we introduce a dimensionless variable u to express the distance from the focal plane along the optical axis as

$$u \;=\; 2\pi\,\frac{\mathrm{NA}^2}{M^2\,\lambda}\,z_i \quad ,$$

where z_i is the axial distance perpendicular to the image plane, related to the axial distance in object space via

$$z_i \;=\; z_0\,\frac{M^2}{n} \quad .$$

In terms of u and v, the intensity distribution is, for $u < v$, given by

$$I(u,v) \;=\; I_0 \left(\frac{2}{u}\right)^2 \left[U_1^2(u,v) + U_2^2(u,v)\right]$$

and, for $u > v$, by

$$I(u,v) \;=\; I_0 \left(\frac{2}{u}\right)^2 \left[1 + V_0^2(u,v) + V_1^2(u,v) - 2V_0(u,v)\cos\alpha - 2V_1(u,v)\sin\alpha\right] \quad .$$

In these expressions, I_0 is the maximum intensity, α is a dimensionless variable defined as

$$\alpha \;=\; \frac{1}{2}\left(u + \frac{v^2}{u}\right)$$

and U_k, V_k are series expansions called the Lommel functions

$$U_k(u,v) \;=\; \sum_{s=0}^{\infty}(-1)^s \left(\frac{u}{v}\right)^{k+2s} J_{k+2s}(v) \;\;,$$

$$V_k(u,v) \;=\; \sum_{s=0}^{\infty}(-1)^s \left(\frac{v}{u}\right)^{k+2s} J_{k+2s}(v) \;\;.$$

The resulting intensity pattern is shown in the right panel of figure 3.7 as isophots in the meridional section defined by u and v. In the focal plane, for $u = 0$, we recognize the Airy pattern with its first minima at $\pm 0.61\ 2\pi = \pm 3.83$ and $\pm 1.12\ 2\pi = \pm 7.02$. Along the optical axis, the first minimum occurs at $u = \pm 4\pi$ or, if expressed as a distance in object space, at

$$z_{min} \;=\; 2\,\frac{\lambda\, n}{\mathrm{NA}^2} \;\;.$$

This corresponds to the distance by which we have to raise the microscope objective in order to change from exact focus of a small pinhole to the first intensity minimum in the center of the observed diffraction pattern. It can be interpreted as a measure for the axial resolution, seen to be significantly worse than the transverse resolution. Furthermore, the ratio of axial to lateral resolution, $z_{min}/r_{Airy} \;=\; 3.28\, n/\mathrm{NA}$, depends on the numerical aperture of the objective. Closely related to the axial resolution is the depth of focus, the axial depth on both sides of the image plane within which the image remains acceptably sharp. It is usually defined as $1/4$ of the axial distance between the first minima above and below focus of the diffraction image of a small pinhole. In the image plane, this distance is equal to

$$z \;=\; \frac{M^2\,\lambda}{\mathrm{NA}^2} \;\;.$$

The depth of focus defined like this is the diffraction-limited, or physical, depth of focus. However, there is a second and sometimes dominating contribution to the total depth of focus due to the lateral resolution of the detector used to capture the image. This part is called the geometrical depth of focus and is given by

$$D \;=\; \frac{M}{\mathrm{NA}}\, e \;\;,$$

where e is the smallest distance resolved by a detector placed in the image plane of an objective with magnification M and numerical aperture NA. The counterpart of depth of focus, defined in image space, in object space is the depth of field: the depth that appears to be in focus within the image, without readjustment of the microscope focus. Using the fact that axial distances in object space are a factor $\mathrm{n/M^2}$ smaller than corresponding distances in image space, the total depth of field is found to be

$$d_{tot} \;=\; \frac{\lambda\, n}{\mathrm{NA}^2} + \frac{n}{M\,\mathrm{NA}}\, e \;\;.$$

Again, one should note that the above considerations for the depth of field and for the intensity distribution in the three-dimensional diffraction pattern all refer to the case of incoherently illuminated, or emitting, point sources, or equivalently to the case where $\mathrm{NA}_{cond} \geq \mathrm{NA}_{obj}$. In general, the depth of field increases by up to a factor of two as the coherence of illumination increases, or equivalently as $\mathrm{NA}_{cond} \to 0$.

3.2.3 Optical aberrations

The Airy diffraction pattern discussed in the previous section correctly accounts for the wave nature of light but for essentially all cases of interest, it still represents an idealization. In particular, it assumes that all lenses that form an optical system are infinitely thin and it only describes the image for a point on the optical axis. In realistic situations, neither of these conditions is fulfilled. The object points correspond to various opening angles, and are imaged by ray bundles for which the main rays are inclined with respect to the optical axis. The impact points of a large number of rays involved in forming the image, on a plane perpendicular to the optical axis, are represented by a spot diagram. Imaging errors, also called optical aberrations, describe the difference between the spot diagram of an actual system and the ideal Airy diffraction pattern. There are two types of image errors: monochromatic errors, occurring for monochromatic light on diffracting surfaces or independent of the wavelength on reflecting surface, and chromatic errors or colour errors, determined by the dispersion of lens materials. Dispersion, the variation of the refractive index as a function of the wavelength, affects all materials but may be more or less pronounced depending on the electromagnetic properties of each material. As nuclear emulsion does not contain any colour information, the scanning typically employs light of a single wavelength and we will restrict ourselves to the discussion of monochromatic aberrations. Furthermore, we will assume that the optical system is well aligned, in particular that the optical axis of the illumination system is identical to that of the imaging system.

The classical treatment of optical aberrations was first formulated by the mathematician Seidel and is referred to as Seidel's imaging error theory. It is based on replacing the trigonometric functions which appear in the expressions for diffraction and reflection by their Taylor series up to third order. From this description emerge five types of monochromatic aberrations:

- opening error or spherical aberration,
- coma,
- astigmatism,
- field curvature, and
- deformation.

For a bundle of rays, incident on a convergent lens and parallel to the optical axis, the rays furthest from the axis will converge nearest to the lens. This effect is schematically shown in figure 3.8. It is called opening error or spherical aberration, referring to the fact that it is often related to the sphericity of lens surfaces. The caustic surface or focal surface is the envelope of the outgoing ray bundle; the caustic line is the intersection of the caustic surface and the meridional plane. Figure 3.8 shows the projection in the meridional plane and the caustic line is indicated by the bold line. For a particularly large opening error the longitudinal position where the bundle is at its most narrow, indicated as A in figure 3.8, can be considered the effective image position. For a smaller opening error, the optimal adjustment is the one for which the central disk is brightest, indicated as B in figure 3.8. Depending on whether rays further from the axis converge before or after the focal point, a lens is said to be under-corrected or over-corrected. In general, convergent lenses are under-corrected

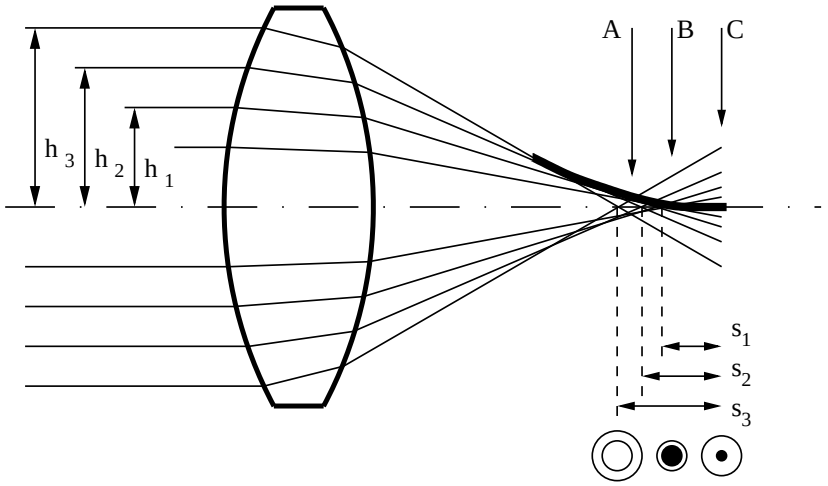


Fig. 3.8: Spherical aberration of a convergent lens. To each incident height h_i corresponds a longitudinal aberration s_i, the distance of the point where the ray crosses the optical axis from the focal point of the lens. The distance s_i is a signed quantity, negative for under-corrected spherical aberration which is shown here, positive for over-corrected spherical aberration. The longitudinal positions marked A, B, and C correspond to the point where the ray bundle is at its most narrow, the optimal image position, and the focal point, respectively. The bottom part of the figure shows the diffraction pattern of a star test for out-of-focus, optimal and in-focus adjustment.

whereas divergent lenses are over-corrected. Through an appropriate combination of convergent and divergent lenses, the opening error of the overall system can be eliminated, or at least minimized, for a certain opening angle or region of angles.

Of particular interest to immersion microscopy are aplanatic lenses. For homogeneous immersion, they are free of any opening error for arbitrarily large opening angles. This is schematically illustrated in figure 3.9, where the outgoing rays are seen to converge to, or rather diverge from, a single point forming a virtual image at the negative focal length of s'. The aplanatic condition is $n.s = n'.s'$, where n and n' are the refractive indices of the immersion oil and of the material behind the lens, respectively. The radius of curvature is given by

$$r = \frac{ns}{n+n'} = \frac{n's'}{n+n'} \ .$$

The magnification by a diffracting surface is given by

$$M = \frac{s'/n'}{s/n} \ ,$$

and can be used to express the numerical aperture for the outgoing rays in terms of the numerical aperture in the object plane as

$$\mathrm{NA}' = n' \sin\sigma' = n \sin\sigma \frac{1}{M} = n \sin\sigma \frac{n'^2}{n^2} = \mathrm{NA}\frac{1}{n^2} \ .$$

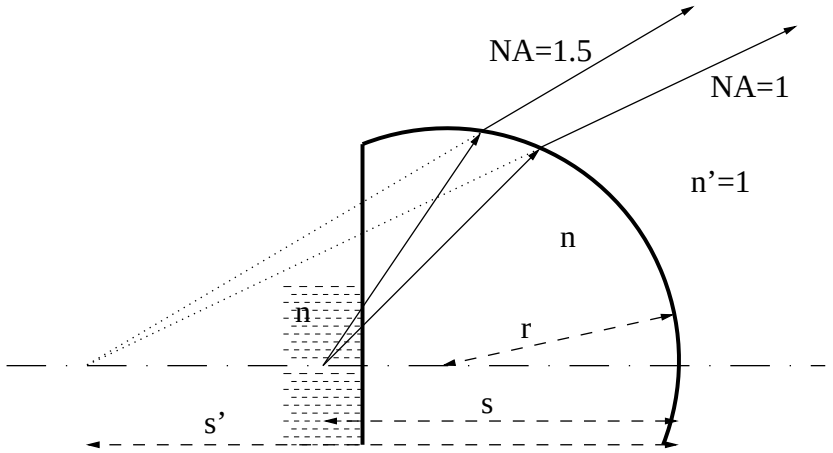


Fig. 3.9: Aplanatic front end of a homogeneous immersion objective. The lens has the same refractive index n as the immersion oil. The curvature radius is r, the focal length for the object plane is s, the focal length of the virtual image is s'. The material behind the lens is assumed to be air, with a refractive index $n' = 1$.

In the last equality, we have taken the material behind the lens to be air with a refractive index $n' = 1$. This immediately implies that an aplanatic lens reduces the numerical aperture by a factor of $1/n^2$. For instance, assuming an immersion oil refractive index of 1.51, to reach a numerical aperture of 1.05 in the object plane requires the lenses after the aplanatic front lens to be corrected up to a numerical aperture of only 0.46.

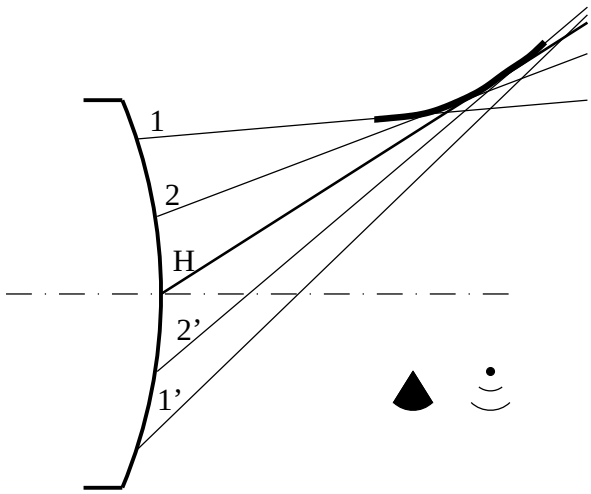


Fig. 3.10: Inner coma aberration, shown in the meridional plane. Indicated are the primary ray H as well as several secondary rays 1, 2, 1′, 2′. The caustic line is shown in bold. The bottom part shows what the image of a point source, for instance in a star test, would look like, first for a geometrical optics treatment of coma, then taking into account wave optical effects showing up as the characteristic diffraction pattern.

Coma is an asymmetric image error, affecting only points that lie off the optical axis. It arises when the intersection of different meridional rays, all belonging to the same opening region, does not lie on the primary ray, as illustrated in figure 3.10. Meridional rays are those that lie in the plane containing the primary ray as well as the optica axis. Coma manifests itself in a shadow for the image of a point source in a plane perpendicular to the optical axis. The shadow is radially oriented with respect to the center of the image. For inner coma, the shadow is directed towards the image center; for outer coma, it is directed outwards. Any lens which fulfills Abbe's sine condition

$$\frac{n \sin \sigma}{n' \sin \sigma'} = \frac{y'}{y} = M$$

is free of coma. For instance the aplanatic lenses mentioned previously fulfill the sine condition and therefore do not show any coma aberrations.

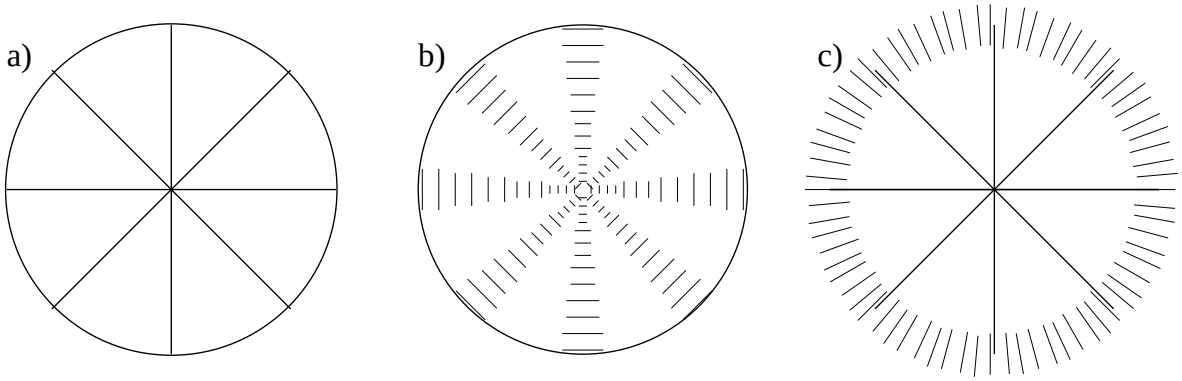


Fig. 3.11: Illustration of astigmatism. The image of a spoke wheel, shown in a, formed by an optical system affected by astigmatism appears as indicated in b and c if it is observed in the meridional or sagittal plane, respectively. Note that the thin lines in this figure are meant to represent a blurred image at somewhat lower relative intensity, merely for the purpose of this drawing. They do not appear as sharp lines in the actual image.

Apart from coma, the image of points that lie off the optical axis may also suffer from astigmatism, even for small opening angles. Astigmatism occurs if the diffraction by the lens is different for meridional rays and for sagittal rays. Sagittal rays are those that lie in the plane containing the primary ray and perpendicular to the meridional plane. As a result, the intersection of the meridional ray bundle with the primary ray will not coincide with the intersection of the sagittal ray bundle. The corresponding longitudinal positions are called the meridional, or tangential, image position and the sagittal image position, respectively. The effect of astigmatism is most strikingly illustrated in the case of a spoke wheel: as shown in figure 3.11, depending on the longitudinal position where the image is observed, either radial or tangential features appear sharp. In the case of astigmatism, points appear as sharp at different longitudinal positions depending on their distance from the optical axis in the object plane, longitudinal positions which furthermore are different for radial or tangential features. The longitudinal difference between the plane where a point appears sharp and the plane erected perpendicular to the optical axis at the position where a central point appears

sharp can be plotted to quantitatively express the amount of astigmatism. Typically, one draws a third curve to indicate the Petzval surface, the surface where point-like objects with neither tangential nor radial features appear sharp.

Once astigmatism has been entirely corrected for, or equivalently once the meridional and sagittal surfaces coincide with the Petzval surface, the image may still suffer from image curvature: regardless of their radial or transverse feature, the longitudinal position where the image of an object appears as sharp may vary depending on the distance of the object from the optical axis in the object plane. As a result, if the focus has been adjusted for the central part of the image, features towards the edges will appear out of focus. It is interesting to note that out of sagittal image curvature, meridional image curvature, and astigmatism, only two are independent. The Petzval coefficient characterizes the curvature of any optical element: the sign indicates whether the image plane is bent inwards or outwards whereas the absolute value indicates the importance of the effect. For a combined optical system, the Petzval coefficients can be summed to give the overall Petzval sum. Typically, elements with negative and positive Petzval coefficients are combined to yield an overall zero Petzval sum, or equivalently a system free of both field curvature and astigmatism. Well known examples of flat field objectives are the planachromats and planapochromats.

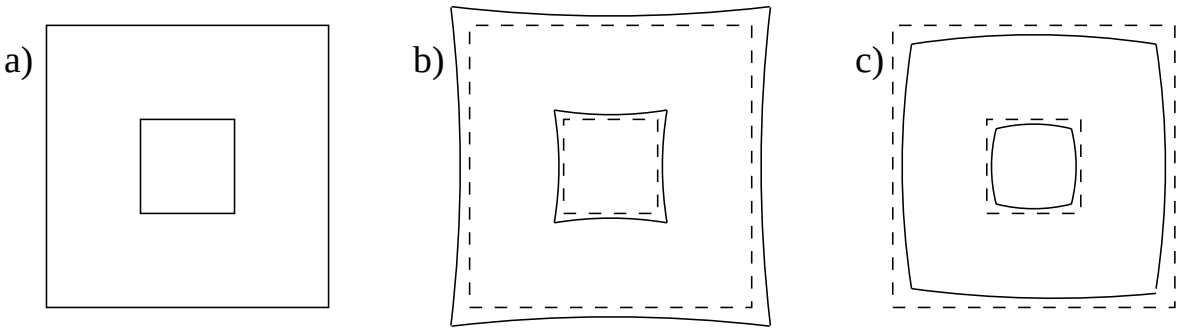


Fig. 3.12: Illustration of deformation. The image of two centred squares, shown in a, formed by an optical system affected by image deformation appears as indicated in b and c if the magnification as a function of the distance from the centre increases or decreases, respectively.

The last monochromatic aberration in the Seidel theory to third order is image deformation, due to a variation in the magnification depending on the distance from the optical axis in the object plane. As a result, the image no longer correctly reflects the shape of objects. If the magnification increases as a function of the distance from the axis, a square will appear as a cushion. Conversely, if the magnification decreases, a square will appear as a barrel. Both of these cases are indicated in figure 3.12. Of all monochromatic aberrations, image deformation is in many applications the least serious because it does not affect the image brightness or contrast. However, in our case where precise length measurements are to be performed out to the edges of the image, no image deformation can be tolerated.

3.2.4 Köhler illumination

Depending on the setup of the illumination system, microscopic observations can be divided into two broad categories, either bright-field or dark-field. In bright-field microscopy, the illuminating light reaches the objective directly, passing through the specimen and affected by the presence of microscopic objects which absorb or reflect the incident light. The resulting image can be thought of as being composed of shadows on a bright background. In dark-field microscopy, the illuminating light cannot reach the objective directly and the image consists entirely of scattered or otherwise deviated light. Consequently, the microscopic objects appear as bright spots on a dark background.

For bright-field microscopy, one further distinguishes transmission illumination and reflection illumination. In the first case, light source and condenser are on one side of the specimen, detector and objective on the other side. The light traverses the object and image contrast typically arises from absorption. In the second case, the image is formed by the reflected or backscattered light of the specimen which is illuminated from the same side where the objective is located. Bright-field reflection illumination requires the entire sample to be reflective to some degree, and is used primarily for opaque and thick samples, typical applications include metallurgy or the semiconductor industry.

In the CERN microscope group, we have at several occasions considered the use of dark-field microscopy for the scanning of nuclear emulsions, but so far have not pursued this any further. Here, we will restrict ourselves to a description of the illumination for bright-field microscopes, more specifically bright-field transmission illumination. A possible setup for this mode is the so-called critical illumination, whereby the light source and the condenser are arranged such that the light source is projected into the object plane. Typically, one obtains high brightness and high contrast, but at the expense of uneven illumination and of a numerical aperture limited by the size of the light source. Furthermore, the image of the light source is overlaid on the image of the specimen, with the features of the lamp disturbing observation of the object under study. In 1893, A.Köhler introduced a different procedure to realize bright-field transmission illumination and it has since become by far the most common. It results in a particularly homogeneous illumination of the specimen and it offers the possibility to vary both numerical aperture and field of view of the illumination system, independently from each other. The advantage of homogeneous illumination is obvious. Generally speaking, the numerical aperture of the illumination system should not be larger than that of the imaging system as this would not increase the resolution. Lowering the numerical aperture typically increases the contrast and in specific circumstances, this might be more important than achieving the highest possible resolution. Similarly, the illuminated field should not be larger than the field of view of the objective as the additional light will not contribute to image formation, but may adversely affect it through light reflected on surfaces, flanges or dust in the light path.

A setup for Köhler illumination is schematically shown in figure 3.13. The essential components are an appropriate light source, a collector lens or lens system situated close to the light source, a field diaphragm or field stop, an aperture diaphragm or aperture stop, and a condenser lens or lens system situated close to the specimen. The collector projects a magnified image of the light source in the aperture diaphragm, placed in the front focal plane of the condenser.

The top part of figure 3.13 shows the ray paths for a point on the optical axis in the field iris, situated close to the collector and very evenly illuminated. The condenser projects an

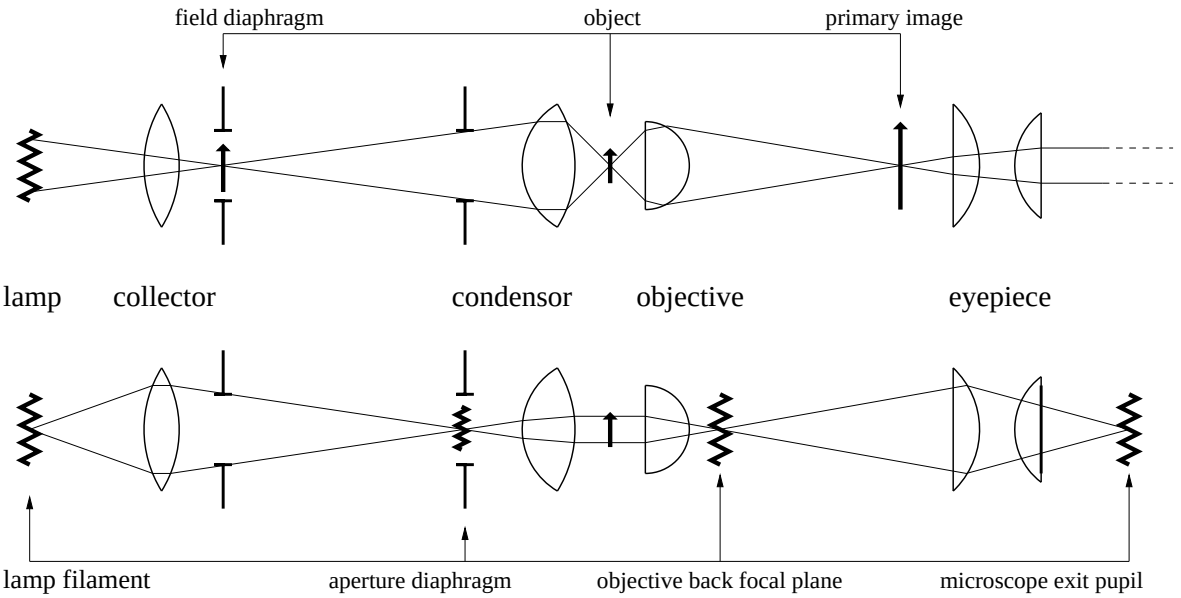


Fig. 3.13: Ray paths in a transmitted light microscope adjusted for Köhler illumination. For the same setup consisting of lamp, collector, condenser, objective and eyepiece, the top part shows the planes conjugate to the object plane whereas the bottom part shows the planes conjugate to the lamp filament. Also indicated are the field diaphragm and aperture diaphragm which can be used to limit the field of view and the numerical aperture, respectively.

image of the field stop in the object plane, which determines the size of the illuminated field. The objective projects the object and the field stop in the primary image plane. The bottom part of figure 3.13 shows the ray paths for a point on the optical axis in the light source. In general, the image of the light source covers the aperture iris when it is fully open. Taken together, the condenser and the objective project the aperture stop and therefore the image of the light source in the rear focal plane of the objective, also called the exit pupil of the objective.

From these ray paths, it is clear that the field diaphragm and aperture diaphragm determine the illuminated field and the numerical aperture of the illumination system, respectively. The field iris should be adjusted such that the illuminated field is no larger than the field used, as given by the size of the image detector. The aperture iris should be adjusted such that the numerical aperture of the illumination system is no larger than the numerical aperture of the imaging system. In fact, the theoretical limit for maximum resolution is achieved for $\mathrm{NA}_{cond} = 1.5\mathrm{NA}_{obj}$ but is not of any practical significance for high-NA objectives. In nuclear emulsion scanning, typically using a dry condenser in conjunction with a high-NA oil immersion objective, the numerical aperture of the condenser should always be maximized, combined with electronic contrast enhancement applied to the image from the detector. For visual observations, the higher contrast obtained at lower numerical aperture may occasionally be advantageous. The Köhler procedure allows for a homogeneous illumination of the object field, even for inhomogeneous light sources such as incandescent lamps or arc lamps. If necessary, the homogeneity can be further improved by inserting mat plates in the light path close to the collector, either between lamp and collector or between collector

and field iris, to destroy any structure in the image of the light source. Of course, the use of mat plates does reduce the image brightness.

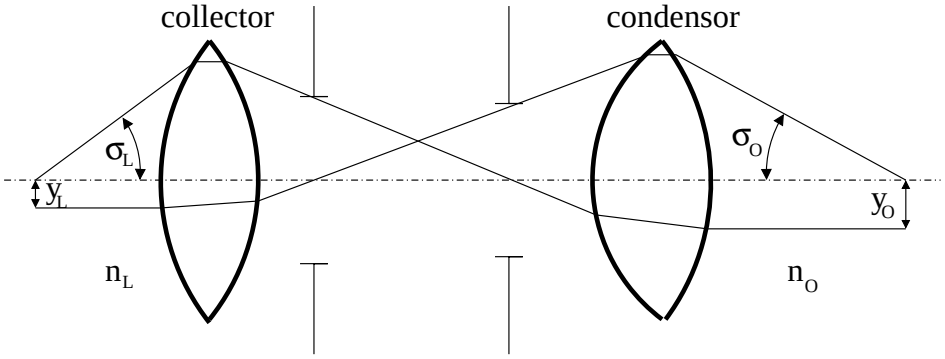


Fig. 3.14: Ray paths between the light source and the object plane in the case of Köhler illumination. The variables determining the image brightness in the field of view are indicated: the transverse size y of the lamp (subscript L) and of the object (subscript O), the largest angle σ accepted by the collector and condenser lenses, and the refractive index of the respective media.

In summary, the Köhler procedure achieves even illumination of the object plane as well as the possibility to independently vary the size of the illuminated field and the numerical aperture of the illumination system. Compared to critical illumination, the most serious drawback is a significantly lower image brightness for an equivalent intensity of the light source. It is worth considering this point a bit further as it becomes all the more important in automatic scanning where one would like to minimize the exposure time and maximize the field of view. Inevitably, smaller exposure time or larger field of view will both further reduce the image brightness. Figure 3.14 shows the quantities which determine the image brightness in the field of view, given the intensity of the light source. These variables satisfy

$$y_L n_L \sin\sigma_L \geq y_O n_O \sin\sigma_O \ ,$$

where the inequality indicates that varying the field iris or aperture iris can only reduce the right hand side. Up to a factor of π, the square of the above expression is equal to the optical flux Λ:

$$\Lambda \ = \ \pi S_L (n_L \sin\sigma_L)^2 \ = \ \pi S_O (n_O \sin\sigma_O)^2 \ ,$$

restricting ourselves to the equality, and introducing the surface S which is proportional to the square of the transverse size y. The optical flux characterizes the transparency of an optical system for radiant intensity, assuming any losses are due to the limited acceptance rather than to absorption in the optical elements. The image brightness E is proportional to the product of optical flux Λ and light current B in the lamp, per unit surface at the detector position. Accounting for the fact that the surface in the image plane differs from the surface in the object plane by a factor of M^2, we find that

$$E \ = \ \pi\, B\, \frac{(n\sin\sigma)^2}{M^2} \ = \ \pi\, B\, \frac{\mathrm{NA}^2}{M^2} \ ,$$

where we have implicitly assumed that the numerical aperture of the objective is at least as large as that of the illumination system. The signal recorded by the image detector will further be proportional to the exposure or integration time, and we finally find that

$$\text{signal} \propto B \frac{\text{NA}^2\, t}{M^2} \quad .$$

For a given lamp and detector, the integration time needed to obtain a fixed signal is proportional to the square of the magnification and inversely proportional to the square of the numerical aperture, provided that the entire area of the light source contributes to the numerical aperture and that the numerical aperture of the collector lens matches the field of view.

At this point, we will interrupt the discussion of microscope optics and instead turn the attention to the electronic treatment of the image, and more specifically to the algorithms for the recognition of grains. Once the principles of digital signal processing have been laid out, the specific implementation in the CERN microscopes will be covered systematically, including the discussion of the optical system in section 3.4.3 which will refer to the general principles that have been outlined here.

3.3 Digital signal processing

Digital signal processing (DSP) is a field of engineering concerned with the digital representation of signals and the use of digital processors to analyze, modify or extract information from signals. In this context, a signal is defined as any variable that carries or contains some kind of information; it includes speech, biomedical signals, sound and music, video and image, and radar signals. The theoretical framework of DSP builds on concepts and techniques in analogue signal processing and time series analysis. The general principles have been laid out several decades ago, but recently the field has known explosive growth due to two interrelated developments. First, with the continuously improving performance of computing hardware, several high bandwidth applications for which previously only analogue methods were suitable can now be handled using digital methods. Second, as the price of computing hardware has come down, multimedia applications have entered the mass consumer market, providing a further boost to new developments. A very practical introduction to the field is given in [103], which contains references to the extensive literature on the subject.

Within the CERN scanning group, it was Piero Zucchelli who first suggested to apply these techniques to the pattern recognition in nuclear emulsion [99]. However, as we will show below, there is no great mystery behind DSP and ideas do not change because the buzzwords do. Ever since its conception, the Nagoya track selector has used a digital filter to enhance the grain features in the microscope image of nuclear emulsion. Nevertheless, there are two specific difference in the implementation at CERN and in Nagoya. At CERN, the CCD readout is digital, so there is no need for an analogue-to-digital conversion before processing the image; in contrast, the CCD camera in Nagoya uses analogue signals. More important though is the fact that the digital filter is hardwired in the Nagoya track selector, through algorithm-specific chips in previous implementations or via a field programmable gate array (FPGA) in the latest version; in contrast, at CERN the algorithm has been programmed on a general-purpose DSP processor of the type described in section 3.3.2. The

application of DSP in the scanning at CERN will be described in more detail in 3.4.5. In the following two sections, we merely define digital filters and introduce some of the techniques used in hardware optimized for digital signal processing. Some of the definitions will be glossed over, the interested reader is referred to appendix 5.7.2 for a more mathematical treatment.

3.3.1 Digital filters

A filter is a system or network that selectively changes the wave-shape characteristics of a signal in a desired manner, affecting its amplitude-frequency or phase-frequency characteristics or both. Filters are generally used to improve the quality of a signal by reducing or removing noise or to extract information from signals by enhancing properties related to the information required. A digital filter is a mathematical algorithm implemented in hardware, software, or both that operates on a digital input sequence to produce a digital output sequence to act as a filter.

The main advantages of finite-impulse response (FIR) filters are that they can have an exactly linear phase response, avoiding any phase distortion, and that they are always stable, especially when realized non-recursively. In contrast, infinite-impulse response (IIR) filters suffer from a nonlinear phase response, especially at the band edges, and are sensitive to instabilities. Instabilities are typically due to finite word-length effects which include coefficient quantization, arithmetic roundoff errors and overflows. Nevertheless, applications requiring a sharp cutoff and high throughput will often be implemented using IIR filters because compared to FIR filters they show a sharper cutoff for the same number of coefficients. Furthermore, the design of IIR filters may start from the properties of known analogue filters, such as Butterworth, Chebyshev, or elliptic filters and transform these into equivalent IIR filters. For FIR filters, no analogue counterparts exist.

3.3.2 Hardware for digital signal processing

A broad classification of DSP processors distinguishes general purpose and special purpose devices. Special purpose hardware may be further subdivided in hardware designed for efficient execution of specific algorithms, such as digital filters or fast Fourier transforms, and hardware designed for specific applications, such as telecommunications or image processing. An example of algorithm-specific digital signal processors are the digital filter chips that were used in previous versions of the Nagoya track selector. Here, we will restrict our attention to general purpose processors as only these are used in the image processing of the microscopes at CERN.

Standard processors are generally based on a von Neumann architecture: the same memory space contains both data and program instructions and the instructions are executed sequentially, with units of the processor not involved in the current instruction idly waiting until control is passed on to them. Such a structure is not particularly suitable for DSP, where the algorithms typically involve repetitive but essentially simple arithmetic operations, such as multiply, add or memory access, and heavy data flow through the processor. Architectures suitable for real-time DSP are generally characterized by multiple bus structures with separate memory space for data and program instructions, by extensive I/O facilities to pass data to and from external devices including the possibility of direct memory access (DMA), and by arithmetic units for specific logical and arithmetic operations. Compared to other

algorithms, DSP operations typically have very few branchings, offering extensive scope to improve the performance by parallelization. The techniques used include

- a Harvard architecture,
- pipelining,
- fast, dedicated hardware multiplier/accumulator,
- special instructions dedicated to DSP,
- replication, and
- on-chip memory,

and are all intended to facilitate the use of parallelization.

In contrast to a von Neumann architecture, in a Harvard architecture the program and data memories lie in two separate spaces. This implies that the fetching of the next instruction can overlap the execution of the current instruction. The TMS320 family of processors used in the CERN microscopes is based on a modified Harvard architecture: separate program and data memory spaces are still maintained, but communication between the two memory spaces is possible, unlike in the strict Harvard architecture.

Pipelining is a technique which allows two or more operations to overlap during execution; a task is broken down into a number of distinct subtasks which are then overlapped. The simplest example is that of a processor which at the ith cycle is simultaneously fetching the ith instruction, decoding the $(i-1)$th instruction, and executing the $(i-2)$th instruction. A more advanced use of pipelining is possible for processors with multiple arithmetic or logical units which can operate simultaneously. For such processors, DSP algorithms can often be formulated such that two or more distinct instructions are executed simultaneously; of course, care must be taken that the result of an instruction is not used before it is available, especially for instructions which take more than one clock cycle.

The basic arithmetic operations in DSP are multiplications and additions. More specifically, the fundamental DSP building blocks are convolution, correlation, filtering, and discrete transforms; all of these require the calculation of finite series where each term is the product of two factors. In standard microprocessors, multiplication is particularly time consuming and the same holds for additions if floating point arithmetic is used. To allow real-time applications, essentially all digital signal processors have a fast, dedicated hardware multiplier-accumulator (MAC) using fixed or floating point arithmetic. Often, an instruction which combines multiplication of two terms and accumulation in a running series is explicitly provided.

Apart from multiply-accumulate, digital signal processors typically provide other special instructions as well. For instance, to accommodate the data shifts or delays to make room for new data samples in digital filtering or correlation, instructions may be provided to copy a data sample to the next higher memory address during the same cycle where it is being fetched from memory or operated on. Other examples of special instructions are those used to reduce the overhead in instruction loops. Instead of the traditional procedure of incrementing a counter and comparing it with the value that indicates the end of the loop, on DSP processors one often finds a specific instruction which allows the next instruction to be repeated any number of times.

Replication involves using two or more basic units of the same type. The replicated units may be arithmetic logical units (ALU), multipliers, or memory units within a single processor; or they may be entire DSP processors which taken together with other identical processors form a larger configuration. In the latter case, the different processors typically share certain resources such as external memory or a bus; the coordination of the various processors is typically taken care of by a specific processor of a different type. Both types of replication are encountered in the DSP hardware used in the CERN microscopes: the individual processors contain several identical units and the DSP boards contain several identical processors.

In most cases, DSP chips operate so fast that slow memories are unable to keep up. On top of that, to get the most out of the benefits from special DSP instructions, zero-overhead loops, or pipelining generally requires a large number of registers and fast, internal memory. To satisfy these requirements, there are usually several types of memory: on-chip program cache which may be used to hold often repeated sections of a program or numerical constants, on-chip data cache to hold the section of data on which the processor is working, and random access memory (RAM) to hold the entire data set, accessible either from a single processor or shared by multiple processors with appropriate synchronization between them.

3.4 The microscope hardware at CERN

Originally, the emulsion scanning in CHORUS relied entirely on the Italian and Japanese groups in the collaboration, especially the FKEN laboratory at Nagoya, Japan. After the first years of data taking, it soon became clear that even for a hybrid experiment, the emulsion work continues to be the central most important aspect affecting the physics results. In view of that, the CERN group decided in the summer of 1996 to set up a laboratory for automatic scanning, first and foremost to get a better understanding of the technique. The obvious first step was to simply copy the system which at that point was operational in Nagoya. The core of the system was entirely retained: a NIKON motorized table with digitized movements and microscope illumination unit, a Tiyoda 50$\times$ objective, and custom hardware for image processing and track recognition. This custom hardware is referred to as the track selector (TS) and expects a standard video signal to supply the images at successive depths corresponding to 512$\times$480 pixels at a 30 Hz frame rate. For less critical components based on Japanese standards, we developed alternative implementations based on off-the-shelf hardware of European or American make: an EEV CCD camera with analogue readout instead of the original Sony CCD camera, a PMC motor controller instead of the DOMS interface, an Olivetti PC instead of the Japanese computer. Thanks to extensive help from our Japanese colleagues, and to the excellent technical staff at CERN, this first system was operational in less than a year [104]. The original Japanese system is described in detail in [105].

In the ensuing years, the aim of the project gradually widened towards a substantial contribution to the CHORUS scanning capacity and towards the design of a more modular system, flexible enough to accommodate the requirements of future emulsion experiments. Each of the hardware components was redesigned, preferring whenever possible commercial solutions relying on software algorithms over custom solutions relying on hardwired algorithms. For the scanning table and the optical system, the requirements are too specific to be met by commercially available components; they were designed in industry according to our specifications and in close collaboration with the CERN technical staff. The image

sensor consists of a commercially available, state of the art CCD chip with associated read-out electronics designed and built at CERN. The image processing relies on standard DSP boards programmed partly in C and partly in Assembler to reach optimal performance in the digital filtering operation itself. Track recognition is entirely done in software. Preliminary results of this new system, at that time still under development, were reported in [106]. Meanwhile, there have been further upgrades and since early 2000, three identical systems are fully operational at CERN. These are the systems that will be described here, occasionally referring to the first system to indicate the differences.

3.4.1 Layout of the laboratory

The scanning facility at CERN occupies a total area of about 50 m^2 spread over three adjacent rooms. The first is the clean room where the microscopes themselves are located. To ensure stable scanning conditions and to guarantee safe long term storage of the emulsion, the environment is precisely controlled and monitored. The temperature is kept at $21 \pm 1\,^{\circ}C$, the relative humidity at 70 ± 5 %. A small overpressure is maintained to prevent dust from entering and accumulating in the room, with a dust carpet at the entrance for the same purpose. Nevertheless, clean room is a bit of a misnomer and should not be taken all too literally. Once mounted in the plastic frames which are described below, the emulsions are generally stored like this as well. For this purpose, several aluminium boxes have been built to hold the frames at a small distance from each other. The dimensions of these boxes were chosen to fit neatly in the empty spaces between microscopes. Finally, the emulsion room contains the emulsion preparation table, where emulsions are cleaned, mounted in their frames and marked. The second room, separated from the clean room by a glass wall is the usual working area. Here are located the computers connected to each of the microscopes, with a second screen to show the microscope image during data taking. In the design of the scanning systems, the aim has been to minimize the need for work in the clean room. Apart from hardware maintenance and plate changing, all routine work can be carried out in the working area. Besides the data acquisition systems, it contains several other computers used for software development or for interactive monitoring tasks. The third room is separate from the other two and hardly ever needs to be accessed: it essentially contains bare computing capacity, mostly without keyboards or displays. The computing services provided include the dispatcher host, the database server, and the Linux PC farm for on-line and off-line tracking. These will be described below.

3.4.2 Mechanics of the scanning table

There is some variation in the dimensions of CHORUS emulsion sheets, for instance the interface sheets are a few cm larger than the target sheets to cover the entire acceptance. In any case, the largest sheet measures about 73×37 cm^2. On the original NIKON stage, with transverse dimensions of 34×34 cm^2, the plates had to be mounted twice. Not only did this double the number of manual interventions, it also reduced the fiducial volume because even in two iterations, the entire surface of the plates could not be covered. This simple observation was the main reason to develop a new table with a working range large enough to cover the entire surface. Following the example in Nagoya, we further opted for micro-step motors instead of the DC servo motors on the NIKON stage.

According to these specifications, a large surface covered in the three axes by micro-step motors, the German company MICOS AG designed a new table. It features a usable range of 80×40 cm^2, a vertical stroke of 20 cm, and micro-step motors with 40,000 steps per turn. The moving parts are mounted on a granite gantry and a table consisting of honeycomb boards between steel plates, mechanically isolated from the floor by air cushions in the four feet. Depending on the axis, the screw pitch is either 2 or 4 mm, corresponding to a step size of 0.05 or 0.1 μm. The maximum speed that can be attained is 80 mm/s, but to avoid wear on the screws it is limited in practice to 30 mm/s. The acceleration varies between 6.4 and 640 mm/s^2. For the horizontal axes, the highest values cannot be used because they lead to oscillations; in any case, the speed at which the emulsion is moved from one transverse position affects the overall scanning time only marginally, especially because the points which need to be visited during an acquisition for one plate are ordered according to their position. In contrast, for the vertical axis, the maximum acceleration is always used. Images are taken during the vertical motion and for them to be equidistant, the motion must be at the required speed over the entire range of images. For lower acceleration values, the motion would have to start from a position further away from the region of interest.

The positions of all three axes are continuously monitored using linear encoders with a precision of 0.2 μm. However, the spacing between opaque and transparent regions on the ruler is 8 μm; the additional factor of 40 is obtained from the electronic combination of the sinusoidal signals from two photo-detectors. As a consequence, reproducibility is only guaranteed if a given position is always reached from the same direction. This is accounted for in the software by separating each move into two parts. First, the table moves to a position at a fixed distance from the actual target position. Then, the remaining distance is covered in a second displacement. The fixed distance is referred to as the backlash correction vector.

The controller, driving the motors and reading the position from the encoders, communicates with the data acquisition computer through an RS-232 serial interface. Nominally, it supports a baud rate of 19200 bps, but in practice it is operated at 9600 bps. At the higher baud rate, we have experienced several problems the most conspicuous of which is the loss of digits resulting in unpredictable behaviour. The ASCII protocol further limits the communication speed, together with the fact that several commands are not well adapted to our application and unnecessarily wasteful of bandwidth. For instance, the position can only be reported for all three axes together using eight digits per axis. In our case, only the vertical position needs to be read during data taking. Recently, MICOS has developed a new controller which overcomes many, if not all, of these shortcomings. It will feature an Ethernet-based binary interface, as well as an extended command set. As soon as it is commercially available, the microscopes at CERN will be equipped with this new controller.

The part which moves in the two horizontal directions supports an aluminium frame holding a glass plate below which the illumination system is located. The glass plate is ordinary window glass, but carefully chosen not to show any small air bubble. The transverse dimensions are 96×56 cm^2 for the aluminium frame and 79.5×45 cm^2 for the glass plate. Over such a large surface supported only at the edges, some degree of sag is unavoidable, and can only be reduced through the use of a sufficiently thick plate. On the other hand, too thick a plate would not only imply excessive weight for the motors to carry, it would also complicate the design of the illumination system. The 4 mm thickness represents an adequate compromise between these conflicting requirements. The centre of the plate was measured to lie up to 200 μm lower than the edges, in agreement with a simple calculation. Immediately around the glass plate runs a vacuum line in the aluminium frame, followed

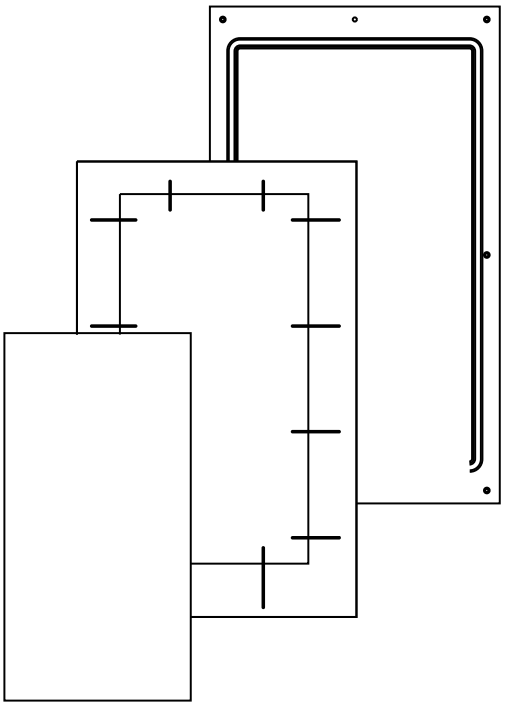


Fig. 3.15: Mounting of the emulsion plates on the scanning table. From top to bottom, the sketch indicates the emulsion plate, the plastic frame and the aluminium framed glass plate. On the plastic frame, the small lines perpendicular to the sides represent the channels connecting to the vacuum line. On the aluminium frame, the innermost rectangle indicates the edges of the glass plate, the second rectangle indicates the position of the O-ring.

by an O-ring to minimize vacuum leaks. The emulsion plates are mounted in plastic frames which close off the vacuum lines on top in such a way that the vacuum volume is essentially located between the emulsion and the glass plate, keeping the emulsion plate stretched flat and keeping it in position for periods up to several hours, corresponding to the data taking for an entire plate. Figure 3.15 shows the entire setup with emulsion on top, mounted in plastic frames shown in the middle and the aluminium frame below, supporting the glass plate surrounded by a vacuum line and an O-ring. The plastic frame consists of two parts. The lower part contains 3 mm wide channels which connect to the vacuum system, these are perpendicular to the outer side and extend from between the vacuum line and the O-ring to the emulsion area. The upper part is essentially a window corresponding to the size of the emulsion plus somewhat less than a mm on either side. Glued together, these two plates constitute a frame. The emulsion fits into the upper part and is connected to it using adhesive tape along the edges.

A single vacuum system serves all three microscopes, with valves at several positions to isolate microscopes that are not being used from the rest of the system. A 40 l vacuum tank is located in the clean room, the vacuum pump itself is situated outside of the building because

of the associated noise and mechanical vibrations. The requirements on the quality of the vacuum are fairly loose, it merely serves to keep the emulsion in position during scanning and this is easily attained for any vacuum better than 50 mbar. On the other hand, as the framed emulsion sheets are not particularly well sealed, the vacuum system must tolerate significant leaks. The diaphragm vacuum pump used satisfies these requirements, evacuating up to 3.3 m^3/min to maintain the pressure in the tank at 2 mbar. Another problem is the immersion oil which, through leaks in the emulsion frame, may enter the vacuum system. To prevent it from reaching the vacuum tank and ultimately the pump, oil traps are situated along the lines from the vacuum tank to each microscope. They can collect up to 500 ml of immersion oil.

3.4.3 The optical system

The intrinsic resolution of nuclear emulsion is related to the grain diameter before development, about 0.4 μm for the emulsions used in CHORUS. First and foremost, the optical resolution should provide a good match to the grain diameter, not only in the transverse direction but also axially. The discussion in sections 3.2.1 and 3.2.2 clearly indicates that high resolution requires large numerical aperture, typically achieved using an oil immersion objective. On the other hand, the free working distance (FWD) of the objective must be sufficient to travel through the entire thickness of the emulsion plate, up to about 1 mm. The need for a high-NA objective with a long free working distance is specific to nuclear emulsion work and for obvious reasons, the commercially available objectives used in biology or metallurgy provide either one or the other, but never both. In the 1970s, Tiyoda designed an objective specifically intended for emulsion work, on request from and in collaboration with Prof.Niu at Nagoya University. The result was a 50$\times$ oil immersion lens with a numerical aperture of 0.85 and a 1.2 mm free working distance. It is corrected up to a diameter of about 200 μm when operating with monochromatic light at 550 nm, the green light which offers the most comfortable viewing conditions, in a medium with a refractive index of 1.51, the refractive index of nuclear emulsion. Ever since, this has been the objective of choice in automatic emulsion scanning, both at the Nagoya FKEN laboratory and in the other Japanese and European scanning laboratories. On the NIKON stage, it is used in conjunction with a standard, dry condenser with a numerical aperture of 0.6, illuminated by a halogen lamp with electronically adjustable intensity and supplemented by a green filter to select the right wavelength. An additional filter blocks the infrared light to which CCD cameras are particularly sensitive.

For this configuration, the expression for the total depth of field given in section 3.2.2 evaluates to 2.3 μm, the sum of a 1.7 μm physical depth of field and a 0.6 μm geometrical depth of field for the 14 μm wide pixels of a typical CCD camera. The theoretical expression for the depth of field is especially interesting to recognize the dependence on numerical aperture, wavelength of the light, or magnification. But it does represent a theoretical abstraction, not immediately applicable to a practical case. In particular, it still refers to a point-like object whereas a developed grain has a 0.8 μm diameter. On top of that, the nominal values of numerical aperture for optical components refer to ideal conditions; in practice, the optical axes for the condenser and the objective may be misaligned and scattering off grains above and below the focal plane deteriorates the optical conditions. To include all such effects, the theoretical depth of field is replaced by an empirical measure. The empirical depth of field is defined as half the axial distance over which the objective is moved between the appearance

and disappearance of a single grain. A grain is said to have appeared if it is picked up by the grain recognition algorithm, essentially a threshold cut on the filtered image. The numerical value of this empirical depth of field is not unique, in particular because a tighter grain selection will more effectively remove out of focus grains. However, this definition has two very important advantages: it is intimately related to axial resolution under actual scanning conditions and it offers an objective evaluation of the optical conditions, particularly useful when tuning the system or when comparing two systems. The empirical depth of field on the NIKON system is about 2.6 μm, measured under favourable conditions and using a fairly selective filter.

The combination of the Tiyoda objective with a high-NA condenser is well adapted for emulsion work, but also represents a compromise between the requirements of manual and automatic scanning. The field of view roughly corresponds to the size which can be inspected instantaneously; a larger field is of little use in manual scanning but can significantly reduce the number of moves when covering a large surface, provided that the image from a wide field camera is analyzed automatically. As mentioned before, a 550 nm wavelength is optimal for comfortable viewing; working at a shorter wavelength, the resolution can be slightly improved, provided that the image is recorded on an automatic sensor not suffering from fatigue or distress. Finally, the numerical aperture in the range 0.7 to 0.8 represents a compromise between resolution and contrast; at higher numerical aperture, the transverse and especially the axial resolution can be significantly improved, provided that any corresponding loss in contrast is corrected for using electronic contrast enhancement. According to our specifications and in close collaboration with the CERN technical staff, Jenoptik designed a new optical system, entirely geared towards fully automatic emulsion scanning. To satisfy the stringent optical requirements, the imaging system and the illumination system were designed together so as to reach an optimal match between the two. This also provided us with the opportunity to accommodate a variable attenuator to compensate for variations in the emulsion transparency and a shutter synchronized with the CCD exposure to further improve the axial resolution as images are taken during vertical motion.

At very high numerical apertures, an optical system becomes particularly sensitive to the thickness and refractive index of all materials in the light path from the object plane to the objective surface, and to a lesser extent from the condenser surface to the object plane. For the CHORUS emulsion, the most significant variation is between target sheets and interface sheets due to the different refractive index for the emulsion gelatin and the plastic support plate. After development, target sheets have about 200 μm of emulsion coated on both sides of a 90 μm thick tri-acetate cellulose (TAC) support plate, whereas the emulsion thickness for interface sheets is only about 50 μm, glued onto an 800 μm acryl plate. As a result, the optical conditions for target plates are dominated by the emulsion, with a refractive index of 1.539 and those for interface plates by the plastic support, with a refractive index of 1.4915. The immersion oil can be chosen to match each configuration, thereby ensuring homogeneous immersion. The optical system must of course deliver optimal performance for both target and interface sheets.

All of the requirements mentioned above were met by the optical system depicted in figure 3.16 [107]. The imaging system is based on a vario, or variable oil immersion, objective lens with infinity focus. The numerical aperture is 1.05, the diameter of the field of view 0.5 mm, and the free working distance 1.3 mm. It can accommodate a variable refractive index between the object and the front surface of the objective lens, within the range $1.49 < n < 1.54$. The working wavelength is 436 nm, corresponding to the mercury g-line, with a

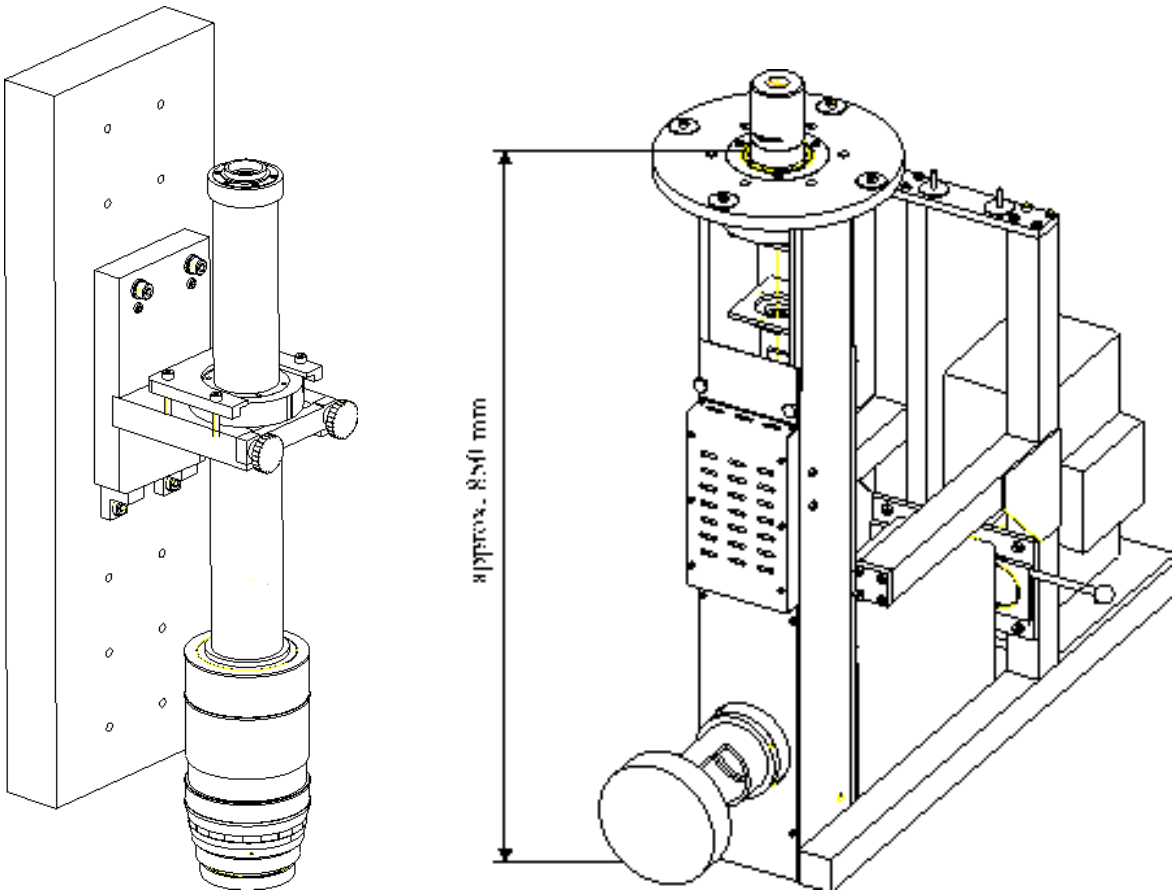


Fig. 3.16: Left panel: the imaging system, measuring approximately 48 cm from top to bottom. It consists of an objective lens followed by a more narrow tube lens, the ensemble attached to the plate which moves in the vertical direction. Right panel: the illumination system, measuring approximately 85 cm from top to bottom. Further details are given in the text.

bandwidth of 10 nm. To allow for variable magnification, the objective focuses at infinity and is followed by an exchangeable tube lens which forms a real image at the camera position. There are three such tube lenses, with magnifications of 40×, 60×, and 80×. The image field size and the overall length from the object plane to the image plane is identical for all three: 20 mm and 488 mm respectively. The objective contains eleven lens elements, organized in seven groups. The optical aberrations caused by a change in refractive index, in particular spherical aberration, are compensated by shifting a lens group inside the lens system. This is a manual procedure, performed between data taking for interface or target plates. Any decentering during the longitudinal shifting would lead to asymmetrical aberrations such as coma. Hence, the accuracy of the mechanical components meets stringent requirements: the tolerance is 2 μm over the 1.5 mm length for shifting. The field curvature is independent of depth and is below 1 μ up to the very edge of the field of view. At the highest numerical aperture for the illumination system described below, the theoretical depth of field is 1.1 μm, with roughly equal contributions from the physical and the geometrical depth of field. Using

the same grain recognition algorithm as for the NIKON system, the empirical depth of field is about 1.2 μm, an improvement by more than a factor of two.

The illumination system is a microscope illuminator of Köhler type with integrated shutter and attenuator, shown in the right panel of figure 3.16. The figure indicates how the illumination unit is attached to the table; it literally hangs from the scanning table such that the microscope remains mechanically isolated from the floor. The box in the back of the figure represents the housing of the lamp, a 200 W mercury arc lamp of the type HBO. At first, the light path is horizontal passing through a diffuser plate, a collector lens and a filter package. The diffuser plate improves the illumination homogeneity but at the expense of a loss in image brightness; for emulsion which is not sufficiently transparent, it can be removed as indicated by the handle in the figure. The collector lens has a numerical aperture of 0.6 and is followed by a spectral filter package which isolates the g-line from the Hg spectrum at this early stage to protect the rest of the system from heating. The appropriate wavelength is further selected by a diagonal mirror directing the light upwards. Light of other wavelengths passes unaltered and impinges on a screen along the horizontal axis, the cylinder in the figure. This screen is in a plane which is conjugate to the object plane, to serve as an aid during the adjustment of the lamp position. Along the vertical light path are another filter, a liquid crystal (LC) unit, the field stop, the aperture stop and finally the condenser lens. The second filter narrowly selects the g-line, with a bandwidth of $\pm$4 nm. At the position of the ventilation element in the figure are the LC elements which make up the variable attenuator and the light shutter. The function of field stop and aperture stop has been elaborated in section 3.2.4; in the figure, their adjustment handles can just be discerned. The condenser has a numerical aperture of 0.9, It is based on a lens with an aplanatic lower surface and a concentric upper surface, a dry variant of the aplanatic lens discussed in section 3.2.3. As the condenser system works at fixed focus, the illuminated field must be homogeneous not only over a diameter of 0.5 mm, but also vertically across the 1 mm height over which the objective travels. To prevent vignetting, the appearance of dark areas at the edges, the illumination diameter in the nominal object plane is 1.4 mm.

The LC-unit consists of two combined LC-cells: the LC variable attenuator (ALM) and the LC light shutter (ALM-Sh). The ALM compensates for variations in the emulsion transparency. It consists of a twisted nematic liquid crystal display (TN-LCD). The LC glass plates carry a transparent indium tin oxide (ITO) electrode which is covered by the rubbed polyimide. The nematic LC molecules align parallel to the buffing direction and are inclined to the glass plate. The rubbing direction of the upper and lower glass plates are perpendicular and hence the molecules perform a 90° twist over the thickness of the LC layer. Polarized light follows the 90° twist on its way through the LC material. Under the application of an electric field, the liquid crystal molecules change their orientation resulting in a different adjustment of the polarization state for a light beam. By mounting the LC modulator between crossed polarizers, amplitude modulation of the light throughput is achieved. The ALM is driven by a DC signal in the range 0 to 5 V. Maximum voltage corresponds to an attenuation by a factor of 30. The DC signal is delivered by one of the analogue outputs of the CerealBox[5] connected to each microscope. This device, communicating with the PC via an RS-232 serial interface line, provides 8 analogue output channels, 8 analogue input channels and 24 digital input/output channels. The update rate is 30 Hz and an analogue signal is generated or read with 12 bit precision. An analogue output channel may provide a DC signal in the

[5] CerealBox LV824-H, manufactured by BG SYSTEMS INC., Palo Alto, CA 94306, USA.

range 0-4095 mV, while an analogue input channel may sample a DC signal in the range 0-5000 mV. The 24 digital signals are grouped in 3 sets of 8, each of which must be configured entirely as input or output. Apart from supplying the DC signal for the ALM, the CerealBox is used to read a number of sensors to monitor the environmental conditions.

The ALM-Sh, in design similar to the ALM, acts as a shutter, synchronized with the exposure of the camera. The nematic LC is 90° twisted between the glass plates, and the light polarized by the ALM follows the twist of the ALM-Sh on its way through the LC layer. At the end of its path the light passes a polarizer which is crossed to the analyzer of the ALM. In this orientation state, the light passes the second LC cell. If a sufficiently high voltage is applied, full deformation of the LC layer is reached, and the light can no longer pass the last polarizer, attenuated by a factor similar to the one achieved in the ALM. The shutter is activated by a Transistor-Transistor Logic (TTL) signal. The start of the CCD camera exposure is very well defined; in contrast, after the actual exposure the camera remains sensitive while it is readout, shifting the lines across the image zone during 1.1 ms. In the absence of a shutter, the image would have an ill-defined longitudinal position due to the motion during exposure and additional smearing due to the integration during readout. The ALM-Sh has a switching time from maximum to minimum transmission of 0.3 ms, well below the typical exposure time of 5 ms. The time from minimum to maximum transmission is between 6 ms and 8 ms, sufficient for frame rates up to 100 Hz. The TTL signal for the shutter activation is provided by the camera interface card described below.

The use of an oil immersion objective is imperative to attain the required optical resolution, but represents a significant disadvantage in itself. During the automatic scanning of large surfaces, the space between the objective and the emulsion surface must at all times be filled with immersion oil. Simply covering the entire surface with a 2 mm layer of oil is not only costly, it also complicates the changing of the plates between acquisitions and leads to an accumulation of dust and dirt. To overcome this, the technical staff at CERN and in Brussels designed the oil containment system which is schematically illustrated in figure 3.17. It concentrates the oil in the small volume below the objective leaving the emulsion sheet to move freely underneath. A rubber ring rests on the emulsion surface, kept in place by the weight of the lead ring which it supports. In turn, the lead ring fits around the objective with a Teflon protection between the two. The oil in the volume defined by the rubber ring is continuously refreshed by a peristaltic pump, using two sets of plastic tubes connected to a reservoir. The inlet tube dips into the oil essentially down to the emulsion surface, whereas the outlet tube is situated at a height of up to 3 mm from the surface, defining the volume of oil. To prevent overflow, the diameter of the outlet tube is significantly larger than that of the inlet tube. The oil is recycled after filtering, guaranteeing good optical conditions whilst limiting the consumption of oil to less than 100 ml per day, to be compared with 2 l per day if the emulsion surface were simply to be covered with oil for each acquisition.

3.4.4 The digital camera

To exploit the large field of view available with the new optical system, the original 512×480 analogue charge coupled device (CCD) camera was replaced by a 1024×1024 digital camera, sometimes referred to as a Megapixel camera. A concise description of the working principle for CCD cameras, indicating the pros and cons of different types, can be found in [108]. The image sensor is a Thomson THX7888A chip, containing 1024×1024 square pixels with a 14 μm side. It is read out by two 15 MHz amplifiers, each of them coupled to a 12 bit ADC.

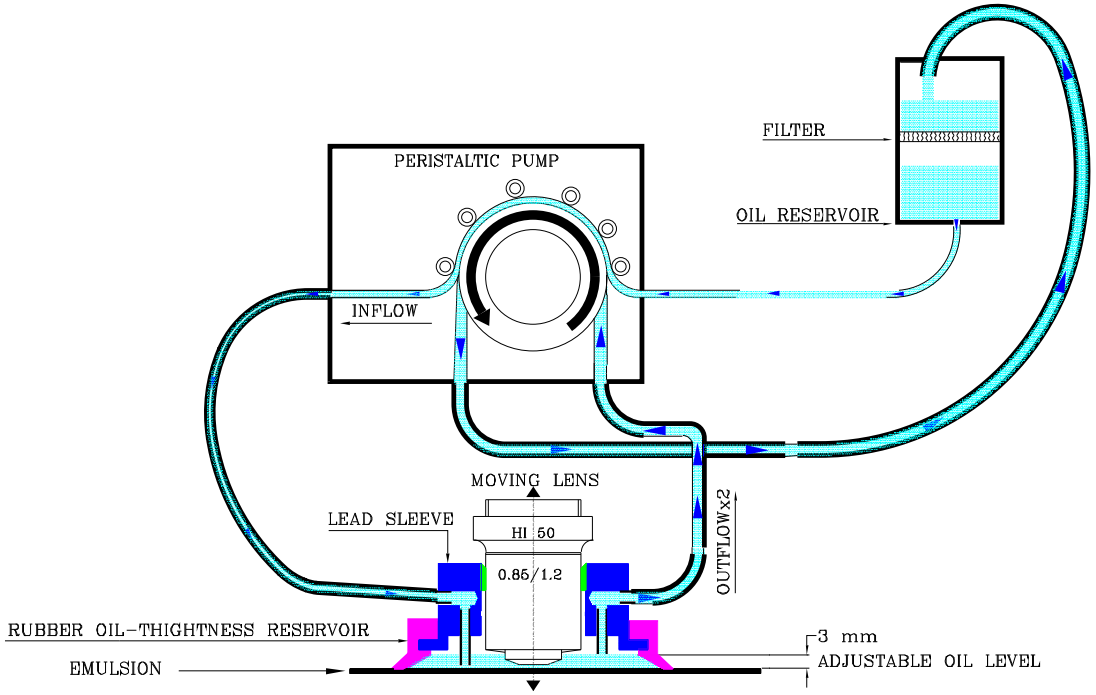


Fig. 3.17: The oil immersion system. The peristaltic pump simultaneously drives a constant supply of oil to the inlet, shown to the left of the objective and reaching the level of the emulsion. At the same time, the oil is pumped away out of the outlet, shown to the right of the objective and at about 3 mm above the emulsion surface.

The corresponding frame rate is 30 Hz, or 15 Hz if only one of the amplifiers is used. When both are used, one reads the left half of the image from left to right, the other one the right half from right to left. This chip is commercially available, but had not yet been incorporated in any camera when we started using it. The electronics — to steer the camera from the DSP board described in the next section and to format the readout signal for further processing — were designed and built at CERN: an I/O driver card situated immediately next to the image sensor and an interface card situated next to the data acquisition computer.

The readout of 8 out of the 12 available bits offers sufficient dynamic range for the grain recognition algorithm. Which of the bits are used can be configured on the I/O driver card. In particular, the low quantum efficiency of this sensor for the blue light of the mercury g-line dictates the use of bits 1 to 8, where 0 is taken to be the least significant bit. At 15 Hz frame rate the data is transferred in 8 bit words; at 30 Hz in 16 bit words with each amplifier output occupying 8 bits. Before processing or display, the corresponding pixels on each line must be reordered, a procedure referred to as byte reordering. Apart from the two sets of 12 bit data signals, the THX7888A also provides frame, line and pixel synchronization signals and has a set of digital input channels which are used to set the camera mode and to trigger the exposure. The exposure time itself is defined by the length of a digital pulse.

Over the 15 m from the interface card to the I/O driver card, control signals back and forth are transmitted using RS-422. For data signals, Low Voltage Differential Signal (LVDS) was preferred because it is faster and shows less slew and skew than RS-422.

3.4.5 The DSP boards

Two DSP boards, connected to the PCI bus of the data taking PC, are responsible for camera control and real-time image processing. Their operation is described in considerable detail in [109]. We will merely list their characteristics and describe their function in the data acquisition.

The first board, referred to as the C80, is the PCI/C81 board built by Loughborough Sound Images (LSI), now Blue Wave Systems (BWS) and based on the Texas Instruments (TI) TMS320C80 processor. Using a Digital Video Module (DVM), piggy backed onto the DSP board and connected to the camera interface card, the C80 is responsible for camera control and image readout. One of the control signals triggers the interface card to generate the TTL signal for the shutter in the optical system. Before the introduction of the second DSP board, the C80 also performed byte reordering and digital filtering. In the present scheme, it remains in charge of the data taking, mostly because of its extensive I/O facilities both to read the raw images from the DVM and to transfer the filtered images to the host PC.

The C80 is a single chip multi-processor, operating at 50 MHz. It has a floating-point 32-bit RISC Master Processor (MP) and four fixed-point 32-bit Parallel Processors (PP). The MP has on-chip data and instruction caches of 4 kb each. It has a 32-bit instruction bus and a 64-bit bus for other on-chip memory accesses. Each PP has an on-chip instruction cache and can access the 4 kb on-chip data and parameter RAMs. They have a 64-bit instruction word with operations performed simultaneously or independently in a Multiple Instruction Multiple Data (MIMD) configuration. In a single clock cycle, each PP can perform up to two address operations, one multiplication, and one general purpose operation in the Arithmetic and Logical Unit (ALU). Eight 40-bit wide registers are available to the ALU and the multiplier unit; four 32-bit registers are associated to each of the two address units; finally, three more registers allow for as many hardware controlled loops. The memory provided on the PCI/C81 board consists of 32 Mb of Synchronous Dynamic RAM (SDRAM) and 512 kb of Flash Memory. A PCI First-In First-Out (FIFO) allows for Direct Memory Access (DMA) transfers at about 100 Mb/s between the C80 and other PCI devices, specifically to the second DSP and to the host PC processor. An I/O FIFO transfers the images captured with the CCD camera from the DVM module to the DSP memory at around 400 Mb/s. Communication and synchronization with the PC is achieved with signals and messages, with a signal generated and received within 50 μs.

The second board, referred to as the C620, is the PCI/C6600 board also built by BWS but based on two TI TMS320C6201 fixed point processors, operating at 200 MHz. In terms of effective clock speed, the TMS320C80 and TMS320C6201 processors are equivalent, operating at 4×50 MHz and 200 MHz, respectively. However, each of the C620 processors can perform up to eight instructions per cycle, compared to four in the case of a C80 PP. Thanks to more internal memory as well as a larger number and greater flexibility of the registers, the digital filter we use could indeed be implemented in half the number of cycles on the C620 processor [110]. Combined with another factor of two, simply from the fact that there are two processors on a C620 board, the overall improvement in speed is a factor of four. All image processing is now done on the C620 board used as a slave for the C80 master which

remains in charge of the communication to the camera and to the host processor and of the synchronization between all processes and transfers.

The core of the C620 processor is based on Texas Instruments' VelociTI Very Long Instruction Word (VLIW) architecture. Each processor contains two sets of functional units, each of those equipped with a multiplier, three ALUs, and 16 general purpose registers. In total, up to eight 32-bit instructions can be performed during a clock cycle, encoded in 256-bit wide packet fetches to or from the 64 kb internal program memory. There are a further 64 kb of internal data memory and on-chip peripherals including a Host Port Interface (HPI), an External Memory Interface (EMIF), and a four channel DMA controller. Each processor has access to 16 Mb of SDRAM whereas two 1 Mb banks of shared Static RAM (SRAM) are accessible from both processors via two local buses. A PCI interface chip is connected to the first bus, so that the two processors and one of the shared memory banks are visible in the PCI address space. This allows other PCI devices to perform master/slave accesses on the internal memory of the C620 processors and on one of the two shared memory banks. The two processors can interact with each other via interrupts or via single word messages. Interrupts in particular allow the synchronization of tasks and the control of the overall data flow. Finally, the board has an I/O port and a PCI Mezzanine Card (PMC) site, both connected to the internal buses. In the future, digital data input via the PMC site should remove the need for the C80 board.

3.5 Data acquisition and on-line analysis

In the previous section, we have covered the various hardware components that make up the emulsion scanning system at CERN. We will now turn our attention to the procedures for CHORUS scanning, realized with the hardware described. In section 3.5.1, an overview of the data flow is given, indicating the direction and type of all signals. After that, the actual reconstruction is described. Broadly speaking, this consists of two steps. First, the grains in a single image are extracted by clustering the pixels above a certain threshold in an image that has been filtered to enhance grain features, a procedure discussed in section 3.5.2. Second, the grains from a tower of images serve as input to the track finding algorithm described in section 3.5.3. This sequence — from images via grains to tracks — is common to essentially all procedures in CHORUS scanning, with relatively minor differences depending on the purpose of reconstruction or the type of emulsion sheets. In the CHORUS experiment, scanning always starts from an event reconstructed in the electronic detectors. Out of this event, one or more tracks are selected, typically those which can be most reliably reconstructed. Their impact point on the most downstream sheet — the changeable sheet, see section 3.1.5 — serves as a prediction to the scanning, covering an area and angular acceptance which reflect the combined errors of the electronic detector reconstruction and the emulsion measurement. If a candidate track is found in the changeable sheet, it is followed upstream in the consecutive plates, a procedure referred to as scan-back. During the scan-back, only tracks with the predicted direction are of interest and this restriction greatly reduces the combinatorics in the track finding. For thin sheets, both sides are scanned and the direction measurement is made across the base, as described in section 3.5.4. In contrast, the scan-back in thick sheets uses only a relatively thin layer at the upstream surface; this procedure is discussed in section 3.5.5. Obviously, all of this assumes that consecutive plates have been aligned with respect to one another. Data taking for alignment is similar to data taking for scan-back; the

alignment algorithms are sketched in section 3.5.6. At some point, the scan-back will come to an end: the scan-back track is no longer found in the next plate. This might be due to random inefficiency, but especially if the track is not found in two consecutive plates, the most likely explanation is that it originates in the plate where it was first missed. This plate is then called the vertex plate and vertex analysis, described in 3.5.7, is performed around the position where the scan-back track disappeared in the vertex plate.

3.5.1 Overview of the data flow

Figure 3.18 depicts the overall data flow and is composed of three parts: the microscope hardware, the host computer, and the network. Most of the signals have been mentioned already in the corresponding hardware sections, and our purpose here is mainly to collect this information, and prepare for the description of algorithms which follows.

The serial ports COM1 and COM2 are both used. The first communicates with the CerealBox mentioned in section 3.4.3. An outgoing signal sets the value of the DAC channel connected to the light attenuator. Incoming signals not indicated in the figure are related to the monitoring of environmental parameters, they correspond to the readings of ADC channels on the CerealBox connected to sensors of camera temperature, oil flow, and relative humidity. The second serial port provides the communication with the stage controller, exchanging ASCII messages to issue motion commands to the stage or to read the current stage position. Furthermore, settings such as the maximum speed, maximum acceleration, limits, and so on can all be read or modified via this interface. Not indicated in the figure is a joystick which is located at the microscope and which connects directly to the stage controller. The host computer may enable or disable joystick control at any time, and other motion commands can still be issued while the joysticks are enabled.

The central device in the acquisition and processing of image data is the C80 DSP, in figure 3.18 indicated as PCIC81 and occupying one of the long PCI slots of the host computer. Control signals are converted to the RS-422 standard in the interface card and distributed to the camera and the shutter. Signals to the camera are used to toggle between the 15 Hz and 30 Hz modes and to define the exposure, with signals from the camera marking the frame, line and pixel synchronization. The signal to the shutter defines the shutter opening time and is synchronized to the camera exposure. The camera output is packed into 8-bit or 16-bit words on the I/O card and then transferred to the interface card as LVDS signals. From the interface card, the raw data is sent to a Matrox digital module and to the DVM. The Matrox digital module occupies an ISA slot in the host computer; it passes the raw image data to a Matrox Pulsar occupying a long PCI slot. The Matrox Pulsar board is a frame grabber card with an integrated Matrox Millenium VGA display adapter. It connects to the second display which displays any image the camera has taken, typically continuously for as long as the acquisition program is running, even when it is in the idle state. The same raw image data from the interface card is also transferred to the DVM, via which it reaches the C80 FIFO. At that stage, there are three possible scenarios depending on the state of the acquisition program. First, the program may be idle. The camera continues to be triggered at any time but the images are used only to refresh the display on the second monitor; as far as the DSP board is concerned, the data is discarded. Second, the program may be acquiring raw data. Such images are used for instance in the off-line manual scanning program, an animation tool allowing for the manual inspection of previously stored images. The C80 board stores the raw images in its own memory, from where they are read by the host processor using

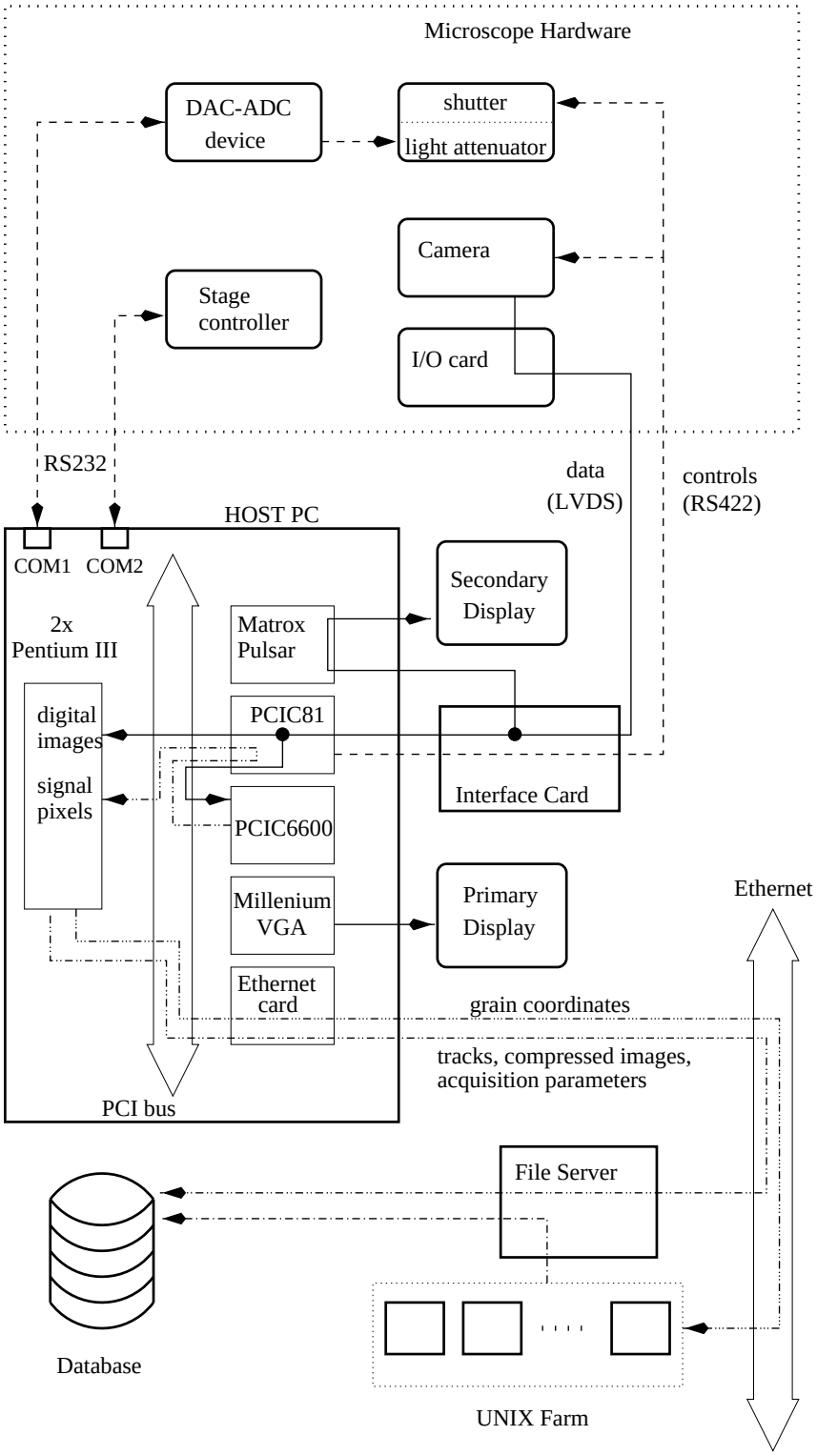


Fig. 3.18: Data flow in the automatic scanning system. Explanations are given in the text.

DMA transfers. Third, the program may be acquiring grain data. In this case, the C80 writes the raw data into the C620 memory and sends to one of the two C620 processors a signal that an image is waiting to be analyzed. The C620 applies a digital filter to the image and then binarizes the output according to a previously set threshold. Once the C80 receives a signal that the packed output is ready, it reads the result using another DMA transfer and makes it available to the host computer.

The host computer can access both filtered data and raw data from the C80 memory using DMA transfers. The type of data, as well as the number of consecutive images, has been specified previously in the command from the host computer which trigger the acquisition of a series of images. At present, raw images are not stored in the Objectivity/DB database. Rather, they are written as plain files on the file server, either in uncompressed TIFF or in compressed JPEG format. In contrast, the packed output of filtering followed by binarization is further processed. First of all, clustering is performed to combine adjacent pixels into grains. The grains may be written directly to the database, with track finding to be performed later on a different machine. Alternatively, they may be the input to the on-line track finding algorithm, typically restricted to a limited angular acceptance. In this case, the parameters of tracks found are written to the database, as well as those grains that are part of or close to reconstructed tracks. Regardless of whether grain coordinates or track parameters are stored in the database, the data is accompanied by any information necessary to characterize the acquisition, for instance microscope settings, calibrations, positions of reference points, or emulsion thickness. If grain coordinates have been stored, the track finding is run on a Unix farm, more than twenty PCs running Linux. Track finding jobs run continuously, waiting for data to be processed. They read their input, grain coordinates with the information defining the coordinate system, from the database and write their output, reconstructed tracks, to the database.

Not indicated in the figure is a messaging system based on the dispatcher [111]. The dispatcher server is a program written in C and C++, running on one of the Linux PCs in the farm. It allows processes, in this context referred to as dispatcher clients, running on the same or other machines to send messages consisting of a tag and a message body as well as to register for all messages with one or more specific tags. The server keeps track of registered clients and stores messages until all interested parties have read them. Technically, the system is implemented using the TCP/IP protocol for message transfers and the use of shared memory on the server. The acquisition program on each microscope PC is a dispatcher client, as well as each of the track finding processes running anywhere in the Linux PC farm. An implementation of the dispatcher client is also available in Java, and this one is used in a number of user interface and monitoring tasks. The microscope acquisition program and the on-line track finding program have no user interfaces; instead, they are implemented as state machines with transitions between states triggered by dispatcher messages. These messages correspond to commands issued in the independent user interface programs. The acquisition and track finding programs also send messages themselves to indicate the state they are in and to report on their progress. The Java monitoring tasks are registered to receive these messages and report the acquisition status in a number of graphical panels. Third, the acquisition program upon writing grain coordinates in the database for further processing, signals this to the track finding programs via a dispatcher message.

The configuration including a dispatcher messaging system has at least three advantages. First, the acquisition program gains in robustness by decoupling its user interface: the control panel sends only messages that make sense given the acquisition state, and the

acquisition program is not affected by crashes of the control panel. Second, both control and monitoring are not tied to a specific machine but can be performed remotely. Finally, this system of messages between the various tasks allows for a seamless integration of the overall setup preserving the independence of each component. Requirements for computing power can be balanced across a number of machines, met in each case by the most suitable hardware platform or operating system.

3.5.2 Real-time processing of single images

The first step in the reconstruction of emulsion data is the extraction of grain coordinates from the microscope image. We have already hinted at this in the context of digital signal processing in sections 3.3 and 3.4.5 and will restrict ourselves here to a brief summary. A very complete and detailed discussion can be found in [109].

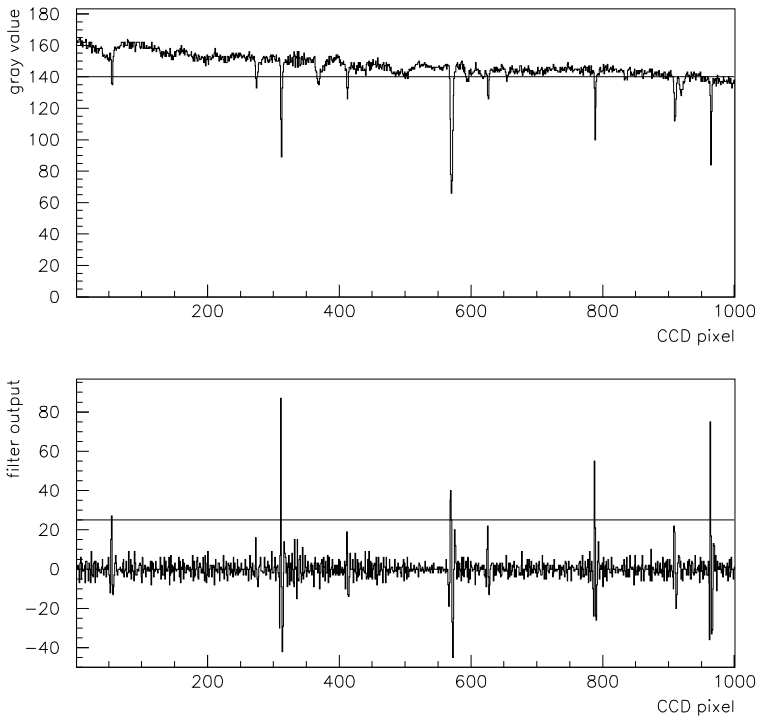


Fig. 3.19: Top panel: gray values for one line of the CCD camera corresponding to an emulsion image. Lower values indicate darker points; the horizontal line corresponds to a possible threshold cut to select black pixels as those with a lower gray value. Bottom panel: the gray value output of the digital filter applied to the CCD line shown in the top panel. The horizontal line corresponds to a possible threshold cut to select pixels that belong to grains as those with a higher gray value.

The top part of figure 3.19 shows an example of the CCD output for an emulsion image, indicating the gray values for all pixels on a single line. A number of features are readily

apparent.

Small gradient. This is an artifact of the readout process in CCD cameras. At equal brightness, the pixels furthest away from the amplifier still have a slightly smaller gray value due to the charge transfer efficiency. Overlaid on top of this, there may be a gradient due to uneven illumination or transparency of the emulsion.

Spikiness. Across the line, the signal varies by a few counts from one pixel to the next. This is the combined effect of electronic noise in the amplifier, more pronounced if less significant bits are used, of differences in the quantum efficiency of individual pixels, and of minor differences in brightness, more pronounced for smaller exposure times.

Broad dips. At several positions, for instance around CCD pixel 370 and CCD pixel 500, the gray value is significantly lower over a range of twenty to fifty pixels. This corresponds to dirt in the processed emulsion such as clusters of fog grains or to the shadow of grains which are out of focus and yield a broad but shallow signal.

Sharp minima. At several well-defined positions, among others at CCD pixels 315 and 570, there is a very pronounced drop in the gray level, extending for only a few pixels. These correspond to the small, black dots indicating the presence of a grain in the focal plane.

The sharp minima constitute the signal we are interested in. The figure further indicates that a simple threshold cut, for instance selecting all pixels with a gray value below 140, is insufficient to reliably identify grains in focus. An intolerable level of background would result from the overall gradient as well as the broad dips. Grains in focus have a characteristic darkness and size and a fixed threshold only exploits one of these. As we have extensively argued before, the application of a digital filter can enhance certain frequencies or, equivalently, features of a certain size. Assuming a grain diameter after development of 0.8 μm, a 40$\times$ magnification, and a 14 μm pixel side, the image of a grain in focus will occupy a more or less circular area with a diameter of two to three pixels. Actually, the signal is slightly broader than the naive estimate because of the diffraction pattern discussed in section 3.2.1, even more so for grains situated marginally above or below the focal plane.

Roughly speaking, a drop in brightness over two or three pixels corresponds to frequencies around 0.25 pixel^{-1} and 0.17 pixel^{-1}, respectively. By comparing the tracking results for different filters and thresholds, we found that a high-pass filter with a cutoff frequency of 0.15 pixel$^{-}1$ gives the best background rejection for a given signal efficiency. Because of the small variation in grain sizes and the small range of frequencies contained in the grain image, the phase response is not an issue in our case. If the filtered image is somewhat displaced with respect to the original image, this is absorbed in the mapping between coordinates on the CCD camera and on the microscope table, provided the displacement is the same for all grains. In contrast, the requirements on processing speed are fairly stringent: images must be treated in real time, corresponding to a throughput of 15 Mb/s or 30 Mb/s, depending on the camera mode. As indicated in section 3.3.1, IIR filters are to be preferred when speed is the overriding requirement because they achieve the same strength and slope with less coefficients, compared with FIR filters. For our purposes, a three-pole IIR filter is sufficient. The actual calculation of filter structure and coefficients was left to the `mkfilter` program, applying the Butterworth method. The result is the following three-pole, high-pass IIR filter

$$y_n \;=\; (x_{n+2} - x_{n-1}) + 3\,(x_n - x_{n+1}) + c_3\,y_{n-3} + c_2\,y_{n-2} + c_1\,y_{n-1} \quad ,$$

with coefficients

$$c_1 = 1.1619\,,\ c_2 = -0.6959\,,\ \text{and}\ c_3 = 0.1378\ .$$

The bottom part of figure 3.19 shows the same CCD line as the top part, but after the application of this filter. The pixels which belong to grains are reliably detected by requiring a filter output larger than 25. The actual efficiency depends on the emulsion and optical conditions, varying between 80 % and 95 % per grain. This particular filter has been implemented in Assembler for both the C80 and the C620 boards. On a C80 parallel processor, it is executed in nine cycles. On one of the two C620 processors, this decreases to six cycles. Furthermore, it includes the byte reordering for 16-bit input from the camera and the application of a threshold. A counter keeps track of the number of pixels above threshold, the output is a binary image where each byte corresponds to eight pixels. It should be stressed that such performances fully exploit the architecture of the DSP board and require considerable work. One such example of a specific, highly optimized implementation of digital filtering code on the C80 and C620 boards is described in [110], using an only slightly different filter.

The result is a binarized image containing those pixels for which the filter output exceeds a certain threshold. Ideally, these are the pixels that belong to grains and only those. In data volume, this image is eight times smaller than the original one. In passing, we also note that the time for all steps up to here is independent of the contents of the image; this greatly simplifies any synchronization issues. The acquisition program on the host computer reads the binarized image from the C80 memory using a DMA transfer and first of all performs clustering, where clustering means combining pixels to form grains. The clustering algorithm is particularly simple: any adjacent pixels above threshold are assumed to belong to the same grain. In this context, adjacent might be horizontal or vertical, as well as diagonal. The grain coordinate is given by the barycenter of all pixels that belong to it, and its size by the number of pixels. In general, clusters that contain only a single pixel or that contain more than twenty pixels are rejected as these are not likely to be genuine *in-focus* grains.

3.5.3 On-line tracking for a tower of images

Once the grains have been identified in single images, tracks may be searched for in the grains from a sequence of images. On average, an image taken in CHORUS emulsion contains about 5,000 grains and a typical acquisition sequence is composed of a tower of about 20 CCD images in the same emulsion plate. Clearly, the most important requirement on the track finding algorithm is to efficiently handle the combinatorics involved. Section 3.5.3 introduces a multi-dimensional binary tree as a possible means to efficiently search through $\mathcal{O}\left(10^5\right)$ objects for those within a specific range of positions. The actual algorithm, discussed in section 3.5.3, is loosely based on ideas from minimum spanning tree algorithms, implemented recursively and relying on binary searches to reach an acceptable performance. The problem of combinatorics is greatly simplified if the track search can be restricted to a relatively limited angular acceptance, for instance during scan-back where one looks for a track with a particular direction. Section 3.5.3 describes an algorithm to select regions in the tower of images that might contain a track with the predicted direction. For a more extensive discussion of the algorithms outlined here, we refer to [112, 113].

The multi-dimensional binary tree

A recurring problem in track finding is that of collecting all hits within a specific volume. The typical example is a possible track section consisting of a handful of more or less aligned points: these points define an approximate direction and if the track hypothesis is correct, more hits should exist along the extrapolation of this direction. A linear search corresponds to simply looping over all objects and checking whether they lie in the appropriate volume. Obviously, the time this takes scales linearly with the number of objects. Alternatively, the collection of objects may be sorted to permit a binary search, for which the time scales only logarithmically. There is some overhead involved in first sorting the objects, but the effect becomes less important as more searches are performed among more points.

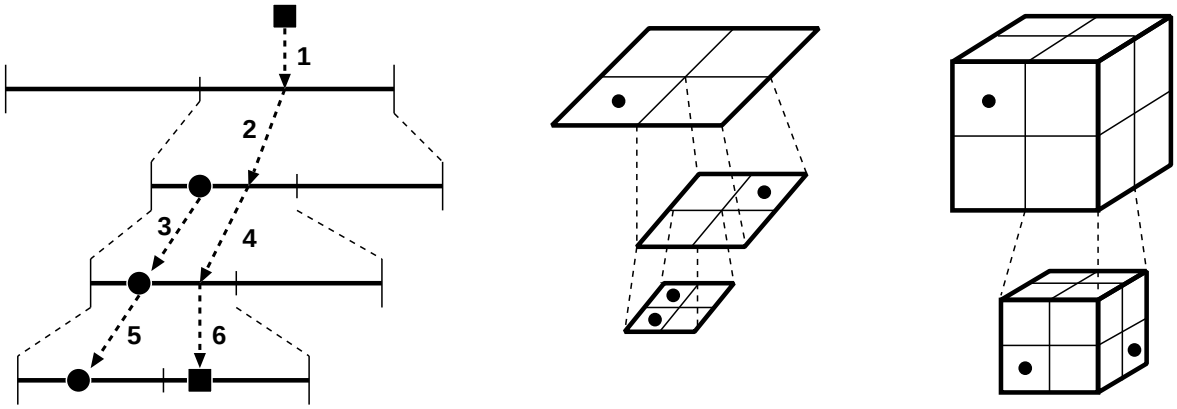


Fig. 3.20: Schematic representation of a binary tree in one, two, and three dimensions. The algorithm for inserting new points is illustrated in the one-dimensional case. The square indicates the element that is to be inserted, the circle an element that is already present, and the numbers the sequence of events. If the new element belongs to a sub-volume that is already occupied, this volume is split and both elements are moved down; this is repeated as long as the elements occupy the same sub-volume.

In track finding problems, the objects are not simply numbers with a definite order but rather coordinates of points in a plane or in space where the concept of order is less obvious. A possible solution consists in generalizing the concept of a binary tree to more dimensions, as indicated in figure 3.20 for two or three dimensions. Instead of storing a collection of objects in a linear array-like structure, they are organized as a hierarchical tree-like structure. Starting from the trunk, a series of nodes successively divide the interval in halves until the interval is small enough for it to contain only a single object, in this context referred to as a leaf of the tree. A search in this collection involves the navigation through successively smaller branches until one reaches the leaves that lie within the desired volume. A similar approach can be used for two or three dimensions: instead of splitting a line segment in two segments, a rectangle may be split in four smaller rectangles or a cube may be split in eight smaller cubes. In practice, this has been implemented for an arbitrary dimension with the possibility to store any type of object for which an ordering operator of the appropriate dimension has been defined.

It can be shown that the time it takes to traverse a tree of dimension D is proportional to ${}^{2^D}\log N$, where N is the number of objects. This scaling behaviour was empirically verified, at least for the case where the search volume is smaller than or of the order of the average distance between objects. To compare the performance to that of a linear search, we consider the case of 10^5 three-dimensional objects. For simplicity, we assume that the number of searches is equal to the number of objects, corresponding for instance to the case where one is interested in finding close neighbours to each object. On a 500 MHz Pentium III processor, a single linear search takes 160 ns $\times N$, or about 1600 s to perform 10^5 searches among 10^5 objects. On the same processor, it takes 350 ns $\times \log(N!)$ to build the tree and 900 ns $\times \log N$ to perform a single binary search. Taking into account the time to build the tree, 10^5 searches among 10^5 objects requires a total of 0.9 s, an improvement by more than three orders of magnitude.

The recursive track finding algorithm

It is instructive to think of the problem of track finding in terms of graph theory, more specifically in terms of the similarity to the well known travelling salesman problem [114, 115]. The aim of track finding is to identify those hits that are aligned, or more generally that are compatible with a certain track hypothesis, possibly different from a straight line due to the effects of a magnetic field, multiple Coulomb scattering, or distortion in emulsion. If the degree of compatibility with a track hypothesis is expressed as a single number and this number interpreted as the distance parameter in graph theory, then track finding is mathematically similar to finding the shortest possible network of direct links. Apart from the similarities between the two problems, there are two important differences as well. First, grains on a single track are connected but there should be no connections between grains on different tracks. Even more strikingly, an overwhelming majority of grains does not belong to any track at all, or at least not to any track of interest. For instance, only between 10^{-2} and 10^{-3} of the grains in the CHORUS emulsion belong to forward-going tracks of high energy. As a consequence, the track finding algorithm described here is only loosely related to a minimum spanning tree.

The input to the algorithm are the positions of all grains in a certain volume, typically corresponding to a tower of between 20 and 50 images, each covering an area of 350×350 μm^2. In a first step, a network of references is built from each grain to its close neighbours, these connections are called links. In a second step, this network is searched for straight routes, referred to as segments, corresponding to the tracks of charged particles. In this step, any link is tried as a starting seed for a segment and the position and direction of a segment is used as prediction to look for additional grains. Due to distortion, the track may not be entirely straight and this is accounted for by using only the last portion of a segment to define its direction. The third and last step checks which segments are compatible with tracks, more specifically selecting those that are sufficiently long, sufficiently straight and containing a sufficient density of grains.

The first step is relatively straightforward. To each grain corresponds a certain volume defined by the required angular acceptance and the average distance between grains on a track, taking into account the grain finding efficiency. To speed up the search for all neighbours in this volume, the grains are all sorted in a three-dimensional binary tree.

The second step, so called segment growing, is implemented recursively; it is schematically illustrated in figure 3.21. Black dots indicate grains, the lines connecting these are the

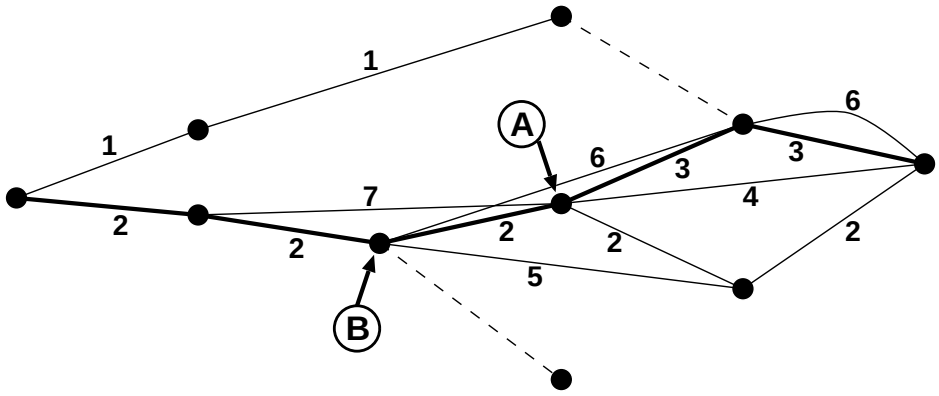


Fig. 3.21: Illustration of the recursive segment growing procedure, not drawn to scale. Explanations are given in the text.

links built in the first step. The numbers indicate the sequence of events in the algorithm. Any grain which has only forward going links, for instance the left most grain in the figure, is used as a starting point. Each of its links acts as the seed for a segment. The position and direction of a segment define a prediction, which in turn translates into an acceptance criterion for further links. At each successive grain, all links are tried which fulfill the acceptance criterion, defined below. This recursive procedure continues until the last grain has no more forward links that are accepted. For instance, the dashed link at the end of segment 1 is no longer accepted because it is not compatiblem with the direction of the previous two links. The algorithm then retreats up to the last grain where another link may still be tried, in this case the second one on the leftmost grain. Segment 2 continues up to the grain which has no more forward links. At point A, the algorithm has two further choices resulting in segments 3 and 4, both of them variants of 2. This continues until all possibilities have been tried. The acceptance criterion can be freely varied and we have experimented with several possibilities: a cone, a cylinder, or a combination of the two. The cone is most suitable if the uncertainty on the direction is dominant, a cylinder if it is the dispersion of the grains from an ideally straight line. The opening angle of the cone, and to a lesser extent the radius of the cylinder, are reduced as the number of grains in the segment grows. However, bearing in mind that tracks are only straight for relatively short sections due to distortion, there is an upper limit to the number of links that can be used to define the prediction. Empirically, the best results were obtained using a cylinder with a radius varying from 1.8 μm for a segment defined by a single link down to 1.4 μm for a segment containing five or more grains, and using no more than the last ten grains to define the prediction. However, these numbers are only meant to indicate the order of magnitude; they vary considerably depending on the type of emulsion, the distance between images and the requirements in terms of efficiency and background rejection.

When multiple links fulfill the acceptance criterion, all of them are followed, resulting in multiple branches at the end of the segment. This leads to a tree of segments, indicated in figure 3.21, with the stem at the starting grain. The branches of the tree then get pruned, working from the leaves to the trunk. In practice, decisions which branch to retain are made by comparing branches pairwise. The segment containing more grains is favoured. In the point marked B for example, segment 2 is preferred over segments 5 and 6 because it contains

more grains. If the number of grains is equal, preference is given to the one that is most compatible with a track hypothesis. In CHORUS emulsion, the track hypothesis is a straight line and compatibility is expressed by the sum of the distances for all grains from a straight line fitted to all of them. In the point marked A for example, segment 3 is favoured over segment 2 because it is more compatible with a straight line. After pruning, a single segment remains, indicated in the figure by a thick line, and following segment 2 up to point A, then segment 3.

At the end of the second step, we are left with a number of segments that link grains, with each grain belonging to at most one segment. Out of all possible segments, those that are more likely to correspond to tracks have been selected. However, that does not yet mean that they do indeed correspond to tracks. The final selection is applied in the third step, and essentially consists in requiring a minimum length.

Track finding with restricted angular acceptance

On a 500 MHz Pentium III processor, the track finding algorithm described in the previous section takes anything between 5 s and 30 s for a tower of 25 images covering an area of 350×350 μm^2 in which tracks up to 400 mrad are searched for. The actual time varies considerably, depending on the number of grains and the acceptances used. Considering the fact that the input consists of $\mathcal{O}\left(10^5\right)$ grains, this is remarkably efficient. Nevertheless, it does not quite match the data taking time. At a 30 Hz frame rate, the 25 images are taken in less than a second. Even accounting for the time during which the vertical acceleration and deceleration take place, as well as the time to move transversely to the next position, a tower of 25 images is typically acquired every 3 s. As a consequence, applying the track finding algorithm on-line on the computer controlling the microscope introduces considerable overhead, with an expensive device, the microscope, left idly waiting for a cheap device, the computer, to perform its task. One solution consists in storing the grains rather than reconstructed tracks and leaving track finding up to a PC farm while the data taking PC acquires new grain data. This approach is used for the vertex analysis which will be described in section 3.5.7. The alternative solution relies on the restriction of track finding to a limited angular acceptance. This is for instance applicable to the scan-back in CHORUS, with the scanning driven by a prediction in both position and direction. The basic idea is to apply a loose selection on the input data, the set of grains, to identify regions that might contain a track under a specific direction. The algorithm described in the previous section is then applied to these regions only, resulting in considerable time savings thanks to the reduced combinatorics.

The basic ideas underlying the selection of regions that might contain a track under a particular direction are inherited from the Nagoya track selector algorithm. The track selector algorithm can be cursorily summarized as 'shifting and summing'. The input are binarized images, corresponding in our case to the DSP output on which the clustering algorithm is applied. First, each of the binarized images is shifted by the reverse of the displacement corresponding to the predicted direction. Then, the summed image is obtained by counting for each pixel the number of shifted images in which this pixel is above threshold. If a track under the predicted direction is present, it will show up in the sum as a distinct peak above a flat background reflecting the average grain density. The appeal of such an algorithm lies in its computational simplicity, making it particularly suitable for a hardware implementation, requiring little more than shifts and summations. However, it does have a significant

drawback in that the signal quickly becomes less significant for distorted tracks or tracks that have a slightly different direction. The problems due to distortion are alleviated by expanding the input signal: the binarized images contain not only pixels above threshold after filtering, but also their neighbours up to a distance of three pixels. To cover tracks under slightly different directions, the algorithm is simply repeated for each point in a grid of angles to cover the required angular acceptance. Barring the possibility of parallel processing, time consumption is proportional to the square of the angular acceptance.

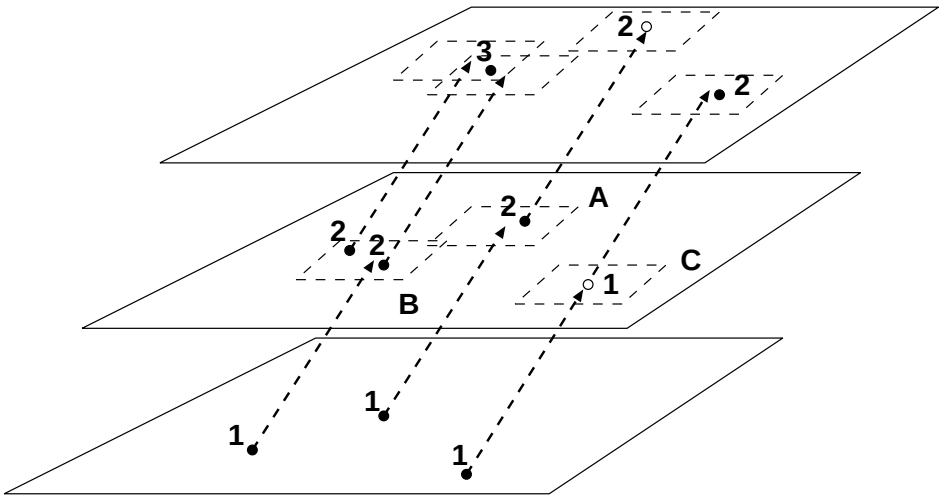


Fig. 3.22: Illustration of the grain summing procedure for the track trigger. Filled dots indicate the positions of actual grains, open dots indicate the extrapolated position of a promoted grain. The dashed arrow correspond to the predicted direction, the dashed squares to the angular acceptance. The number next to each actual or promoted grain is the sum value. The three possible cases A, B, and C are described in the text.

Our grain summing algorithm, illustrated in figure 3.22, is conceptually similar but uses the grain coordinates as input rather than binarized images of pixels above threshold. The implementation in software offers some additional flexibility to deal with the problems of distortion and angular acceptance. The grains are first of all organized in two-dimensional binary trees, one for each layer. In the first layer, a sum value of one is assigned to each grain. Each of these is extrapolated to the next layer along the predicted direction and used to define an area proportional to the angular acceptance. If this area contains a grain on the next layer, its sum value is taken to be that of the grain that was extrapolated incremented by one, for example the point marked A in the figure. If the area contains more than one grain, this is applied to all of them, as indicated by point B. Due to the finite grain density and the grain finding efficiency, a track does not necessarily contain a grain in each layer. To account for this, grains for which no candidate is found in the area on the next layer, are promoted: a virtual grain is created at the expected position and is given the same sum value as the one from which it originates. This is the case for the point marked C. This algorithm is repeated for all images in the sequence. The sum value for all grains in the last layer, including those promoted from previous layers, is similar to the sum in the track selector algorithm. Throughout this procedure, the information on the position of grains

in preceding layers is discarded; as a result, small curvatures due to distortion are easily accommodated. The angular acceptance is increased by enlarging the area in the next layer, resulting in a time consumption that grows by somewhat less than the square of the angular acceptance. Owing in particular to the speed performance of the binary tree, sum values covering an angular acceptance of 30 mrad for 25 images with an area of 350×350 μm^2 are obtained within 0.5 s to 2 s, depending on the number of grains.

By itself, the grain summing algorithm is too crude to reliably identify tracks: efficiencies above 98 % are only reached for small thresholds on the sum value, leading to intolerable levels of background. To confirm the presence of a track, we resort to the algorithm described in the previous section. For each grain above the summation threshold in the last layer, an inclined cylinder is erected, with a width proportional to the angular acceptance and a direction equal to the predicted one. For 25 images, the region defined by such a cylinder contains fifty to a hundred grains. With an angular acceptance of 30 mrad and a summation threshold ensuring at least 98 % efficiency, there may be up to hundred such regions. As a result, the track finding algorithm needs to consider at most $\mathcal{O}\left(10^4\right)$ grains instead of the original $\mathcal{O}\left(10^5\right)$. Even more importantly, the combinatorics is restricted to regions containing no more than hundred grains. Under these conditions, track finding is achieved in less than 0.5 s, even when using a relatively large acceptance to reach high efficiencies.

3.5.4 Scanning procedure for thin sheets

In the previous two sections, we have seen how tracks can be found in a tower of images. After applying a digital filter, a binarized image is formed using the pixels above threshold, which in turn are clustered to grains. These grains form the input of the track finding algorithm, restricted to particular regions if only a limited angular acceptance needs to be covered. We now turn our attention to the data taking procedure which defines the tower of images in the first place, in particular for scan-back data where the position and direction of a track are predicted. This section discusses the case of thin sheets, applicable to the changeable and special sheets in the CHORUS experiment; the next section considers thick sheets, applicable to the CHORUS target sheets.

The CHORUS thin sheets consist of an $\sim$ 800 μm plastic base, coated on both sides with a $\sim$ 100 μm layer of nuclear emulsion. For the 1996 run, the emulsion thickness was reduced to 70 μm. Due to the shrinkage discussed in section 3.1.3, the emulsion thickness is reduced during development by a factor of about two. The empirical depth of field is 1.2 μm, slightly larger if the grain finding algorithm is somewhat relaxed. As a result, a given image corresponds to a slice in depth with a thickness of at least 2.4 μm and no more than 20 independent images can be taken when covering the emulsion thickness. Under unfavourable conditions, even 20 images will show some redundancy. On the other hand, the background rejection of the track finding algorithm improves with the number of independent images. For the grain densities prevailing in CHORUS emulsion, 15 images are a strict minimum to reliably detect tracks and the background is reduced to zero only for more than 25 images. In summary, optimal background conditions are only reached if the entire emulsion thickness is covered, with a spacing between images equal to twice the empirical depth of field.

The left panel of figure 3.23 shows how this is achieved. The thick lines show the two emulsion surfaces: below the interface between the plastic support and the emulsion, above the external surface of the sheet. Initially, the objective focal plane is situated inside the base, at a distance of five to ten microns from the emulsion, indicated by the thin, solid

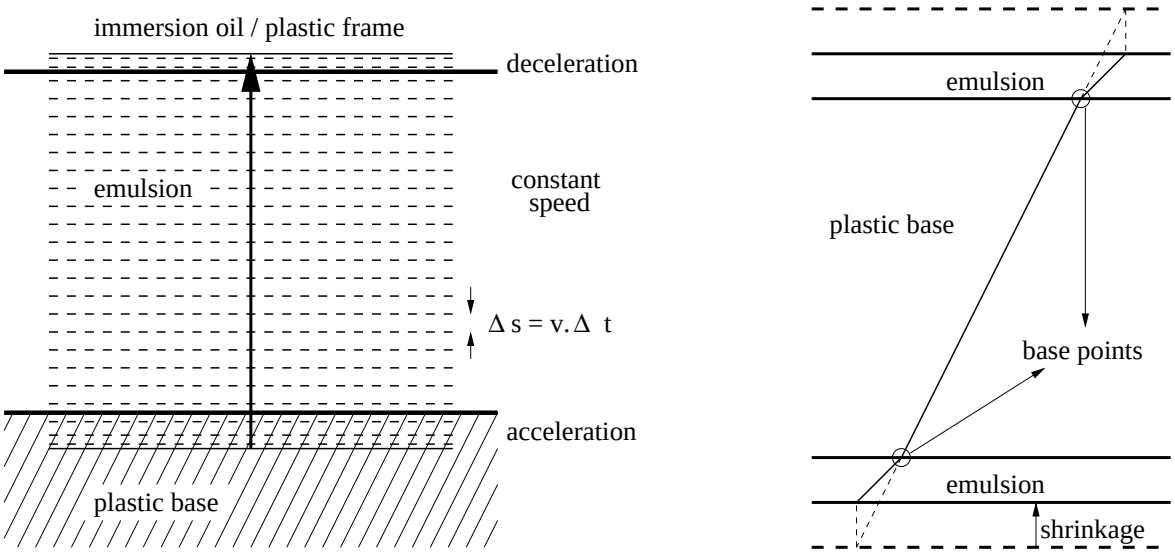


Fig. 3.23: Data taking in thin sheets. Left panel: the emulsion thickness of ~ 50 μm after development is entirely covered, moving from the base surface to the external surface at constant speed while taking images. Right panel: the direction of reconstructed tracks is determined from a base measurement.

line in the figure. The stage is instructed to move at constant speed along the vertical axis towards the external surface of the emulsion plate. The speed is chosen such that, given the camera frame rate and the assumed emulsion thickness, 20 images would cover the entire thickness. The acceleration is set to the maximum allowed value. Immediately after the instruction to the stage controller, a command is issued to the DSP to process the next 25 images, producing the binarized image of pixels above threshold for each plane. The dashed lines indicate the vertical position of each image, initially at a lower spacing because the stage is still accelerating to reach its target speed. As soon as all images have been acquired, the stage is instructed to stop at a position indicated by the solid line; it will come to a halt a bit further, again shown as a solid line. The last two images are not analyzed by the DSP. As mentioned in section 3.5.2, the DSP algorithms not only produce the binarized image of pixels above threshold, they also count the number of such pixels in each image. In addition, the DSP keeps track of the time at which each image is taken. These times are correlated with the times at which the stage position is read, corrected for any delays involved. The number of pixels above threshold determines whether images were taken inside or outside of the emulsion. The acquisition is considered successful if both the first two and the last two images are taken outside of the emulsion. This condition ensures that the stage has reached a constant speed when entering the emulsion layer and that the entire thickness is used. If the number of empty images before and after differs from two by more than one, the initial position and the speed are adjusted and this procedure is repeated until the acquisition is successful. The adjusted values are typically valid for neighbouring views as well and data usually need to be re-taken only if the transverse distance from one acquisition to the next is larger than a few mm. This is the case for instance from one prediction to the next and is due to the sagging of the microscope glass plate, as well as variations in the emulsion

or base thickness. The left panel of figure 3.23 shows the scanning of the emulsion side facing upwards, with the microscope moving up from the base to the immersion oil between objective and emulsion sheet. For the emulsion side facing downwards, the microscope moves down from the base to the plastic emulsion frame, to which the emulsion is pressed by vacuum.

For CHORUS changeable sheets, the prediction is given by tracks reconstructed in the fibre tracker discussed in section 2.3.1. The errors on reconstructed track parameters vary from 180 μm to 300 μm for the transverse coordinates and from 1.5 mrad to 4 mrad for the slopes. All of these numbers refer to one of the two projections; the smaller errors typically correspond to high-energy muons reconstructed in the muon spectrometer whereas the upper range is characteristic of hadrons and electrons with energies below 2 GeV. A single microscope view of 350×350 μm^2 is insufficient to cover a three sigma range. Instead, the required area is covered using adjacent views. The overlap between views is proportional to the track angle, such that at least one view contains the track over the entire emulsion thickness. Far more selective than the prediction for a track's position is the prediction of its direction which must be matched in precision by the slope measurement in the emulsion. From the images taken on one emulsion side as described above, the slope can only be determined with an error of 8 mrad or worse due to distortion, the unknown shrinkage factor, and the small lever arm. In contrast, the base measurement depicted in the right panel of figure 3.23 yields a slope error of less than 2 mrad: the position of the two base points, not affected by distortion or shrinkage, has an error of less than 1 μm while the thickness of the base provides a lever arm of about 800 μm.

In summary, an area of 3×3 to 5×5 views needs to be scanned on both sides to account for the position error of the prediction and to reach a comparable slope resolution. Reproducibility of the stage position is only guaranteed on the condition that any target position is always reached from the same direction, as discussed in section 3.4.2. Data taking across the emulsion layer always proceeds from the plastic base towards the outer surface. Taken together, these points immediately lead to the scanning sequence depicted in figure 3.24. The transverse position of the predicted track is used as the center for a rectangular area of overlapping views, the overlaps are not indicated in the figure. To minimize the number of additional moves related to the backlash correction, the correction is set to the distance from one view to the next in both directions and the order in which the views are covered follows a diagonal pattern. At first, the emulsion side facing upwards is scanned: always traversing the emulsion from the base to the immersion oil and moving from the end point of one acquisition to the starting point of the next in a single move for views 3, 5, 6, and 8 or two moves for views 2, 4, 7, and 9. The track parameters and nearby grains are stored for all tracks found in any of the views on this side. The procedure is then repeated for an area on the side facing downward. The center of the area on the other side is displaced by the amount expected for the predicted slope. Once all views have been scanned on both sides, matching pairs are searched for in the two sets of tracks, requiring agreement within errors for both position and slope. For these pairs, a base measurement is used to determine the track parameters and all the results are stored in the database: parameters from the base measurement, track parameters and nearby grains on either side, and references to further information for this acquisition. Further information includes the hardware settings, coordinate transformations from pixel space in the CCD camera to micron space on the stage and from coordinates on the stage to coordinates on the emulsion plate, the predicted position and direction, and so on.

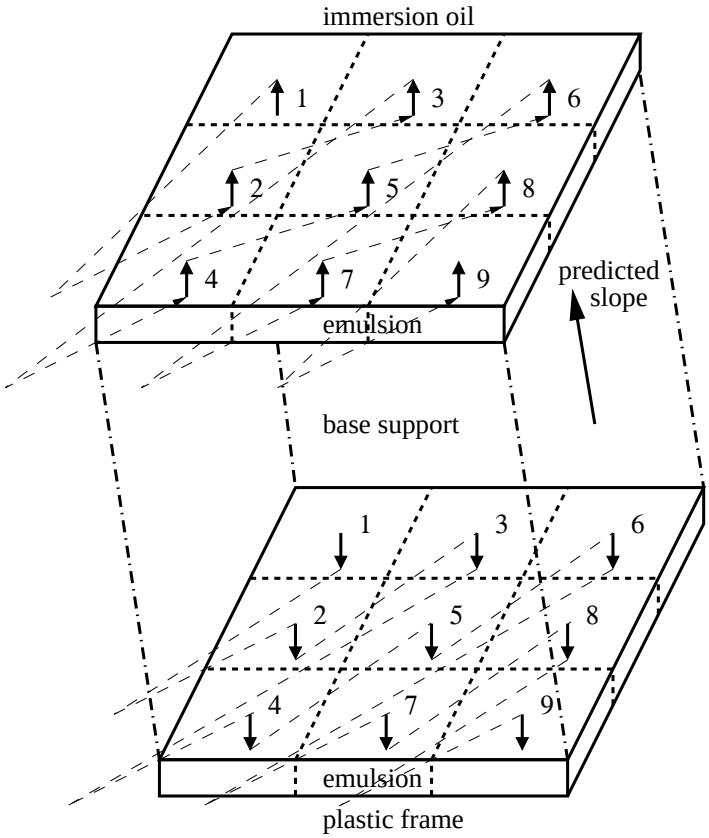


Fig. 3.24: Double-sided data taking for multiple views. Explanations are given in the text.

The track parameters of the prediction are given in the coordinate system used by the electronic detector reconstruction. The alignment of the emulsion sheets with respect to this reference frame will be discussed in section 3.5.6. However, the position readily available during the scanning is only that of the stage and the emulsion frame does not have a definite position on the glass plate. During a data taking run, the emulsion is kept in place by the vacuum but it will typically be at a different position from one run to the next. The X-ray marks, discussed in section 3.1.5, provide a reproducible coordinate system. The 1 mm diameter black spots on the emulsion surface are scanned whenever the emulsion plate has been taken off the microscope table, or if they have not been measured for over 24 hours, in which case the vacuum may not have held them firmly in place. The spot itself is larger than the field of view, but a determination of the edge at four positions yields a measurement of the center with a reproducibility between 20 and 50 μm, depending on the sharpness of the edge. The comparison between nominal and measured position for the four X-ray marks at the corners of the emulsion plate defines the mapping between coordinates measured on the microscope table and coordinates with respect to these fixed reference points, accounting for a translation and rotation of the emulsion plate with respect to the stage.

For CHORUS special sheets, the predicted position and direction are not given by the values reconstructed in the electronic detectors but rather by the extrapolation of tracks found in the corresponding changeable sheet. The distance between the two is 3.2 cm, leading to

a $\sim$ 60 μm extrapolation error for a 1.8 mrad error on the base measurement of the track slope. The error due to the relative alignment between changeable and special sheet is of the order of 50 μm. Adding the extrapolation and alignment errors in quadrature, the error on the transverse position of a special sheet prediction is $\sim$ 80 μm, in each of the two projections. As a result, a three sigma error can be covered with 2×2 views, in exceptional cases a single view. Apart from this difference, the scanning of special sheets is identical to that of changeable sheets.

3.5.5 Scanning procedure for thick sheets

The choice of data taking procedure in thin sheets is driven by the relatively small emulsion thickness available, the coverage of a sufficiently large area, and the need for a base measurement to reach the required angular resolution. In thick sheets, the CHORUS target plates, the conditions are very much the opposite. After development, the emulsion layer on each side has a thickness of at least 150 μm, easily accomodating 50 independent images as long as the empirical depth of field is below 1.5 μm. The prediction is based on the track found in the previous plate, situated immediately downstream from the current one. For reasons which will become clear in section 3.5.7, tracks are searched for at the upstream surface of target plates. As the plates are stacked together under vacuum during the exposure, the distance between the upstream surfaces of consecutive plates is equal to the plate thickness, slightly less than 800 μm. For a 1.5 mrad error on the slope, the extrapolation error is less than 1.5 μm. The position error due to the alignment of consecutive plates depends on the number of tracks that can be used for alignment, but never becomes larger than 5 μm. Combining the extrapolation and alignment errors, a three sigma region around the predicted position occupies but a small fraction of the field of view. Finally, a base measurement is not particularly useful in the target sheets. In changeable and special sheets, a precise angular measurement is necessary to reliably select the correct track. For target sheets, the window in position is sufficiently narrow for the correct track to be identified even with an angular resolution as large as 30 mrad. Anyhow, the 80 μm thickness of the plastic base precludes the accurate determination of the slope using a base measurement.

The scanning of thick sheets uses a single tower of 25 images, with a spacing of 3 μm and starting just below the upstream surface. The plates are put on the microscope with the upstream side facing upwards. A layer of 5 to 10 μm at the surface is excluded because it shows a larger fog density, with scratches and dirt further reducing the visibility. The region at the external surface, rather than at the plastic base, is preferred because it is less affected by angular distortion, as shown in figure 3.3. More specifically, the track finding algorithm can assume that tracks are well described by straight lines. The side facing upwards is preferred because of the better optical conditions: the image quality of the other side is significantly worse due to the light scattering off grains between the object plane and the objective.

The spatial resolution required in the scanning of target sheets can only be reached locally: due to large-scale distortion, transverse distances of more than a few mm differ by several tens of microns before and after development. The grid of fiducials, discussed in section 3.1.5, provides reference points every 1.95 cm, with a position known before development and therefore affected by distortion in the same way as the tracks. Instead of the coordinate transformation based on X-ray marks at the four corners of the plate, the scanning of thick sheets uses a coordinate transformation based on three or four nearby fiducial marks. The degrees of freedom, rotation, translation, and transverse shrinkage, are the same

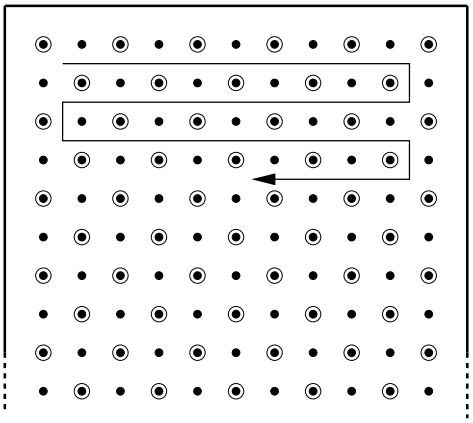


Fig. 3.25: The order in which predictions are scanned in CHORUS target plates.

as for thin sheets, but are allowed to vary across the plate. The central dot of each fiducial has a diameter of 50 μm and its center can be determined with an error of about 1 μm, slightly worse if the emulsion is dirty or the edge of the fiducial badly defined. In the scanning of thin sheets, the X-ray marks are measured at the start of the acquisition; after that, the transformation parameters are used up to 24 hours. For thick sheets, the micron precision on the fiducial position is guaranteed only for a limited time as the plate will inevitably move by several microns over a period of hours, in spite of the vacuum. Therefore, the measurement of tracks must immediately follow that of the fiducials which are used in the determination of the coordinate system. To avoid remeasuring the same fiducial during a single acquisition, the predictions are sorted according to their position on the plate, such that the microscope describe a snake pattern following the rows of fiducials, as indicated in figure 3.25. For each track prediction, the three closest fiducials are scanned but without rescanning any that were used already for previous predictions. From the figure, it is clear that the time elapsed between the track measurement and the fiducial measurement will correspond to at most the acquisition of all predictions along two rows of fiducials, which is never larger than 15 min.

The angular resolution is less of a concern. Because the extrapolation from one plate to the next uses the track angle measured in special sheet, the angles measured in the bulk plates are merely needed to select the correct track, from an area determined by the 3 μm position error. For this purpose, the 8 mrad angular resolution is sufficient. It is limited by distortion and by local variations in the emulsion thickness and shrinkage.

3.5.6 Alignment

The X-ray marks in thin sheets and the fiducial marks in thick sheets define a coordinate system on the plate, independent of the way the emulsion is mounted on the microscope table. From one data taking run to the next, possibly on a different microscope, the X-ray and fiducial marks define a reference frame that is reproducible to within ~ 30 μm and ~ 1 μm, respectively. Apart from this mapping between so-called stage coordinates and nominal coordinates, it is also necessary to relate the nominal coordinate system on one plate to that on the next, either upstream or downstream, or more generally to the external coordinate system defined by the electronic detector reconstruction. This is referred to as the mapping

between nominal coordinates and calibrated coordinates.

Changeable and special sheets

As mentioned in section 3.1.5, the position of the X-ray marks has been determined with respect to the electronic detectors in special runs where the X-ray sources were replaced with β sources, of which the emitted electrons are detected by the fibre trackers. This provides an initial position for each X-ray mark on the downstream surface of a changeable sheet, corresponding to a radioactive source in the honeycomb spacer immediately downstream of the sheet. The honeycomb spacers separating special and changeable sheet contain X-ray sources as well, creating marks at the same position on the upstream face of the changeable sheet and the downstream face of the special sheet. As a result, a changeable sheet has a set of fiducials on both surfaces, one related to the electronic detectors through the runs with β sources, the other related to the set of marks on the special sheet.

After the first changeable sheet period, the relative position for the two sets has been determined and since then, only the set of X-ray marks present on both CS and SS is used. Their nominal position provides the starting point in changeable sheet scanning of predictions from the electronic detector, but on any given sheet the transverse displacement of the actual marks may be as large as 0.5 mm. To determine this offset, a subset of hundred predictions for high-energy muons is scanned first, using a scanning window enlarged by 1 mm. After the determination of the transverse displacement based on a hundred predictions, another hundred predictions with slopes larger than 200 mrad is scanned using a window enlarged by 0.5 mm. The large angle predictions are used to determine the longitudinal position of the plate. The scanning window for all remaining predictions is simply given by three times the error on the track parameters reconstructed in the electronic detectors. Predictions are broadly classified according to quality — preferring muons over hadrons and electrons, and track angles below 200 mrad over those above 200 mrad — and scanned in that order. At the end of each category, the data is used to refine the calibration. Degrees of freedom are translation, rotation, transverse shrinkage, and longitudinal position.

After the CS scanning of all predictions, the final calibration is performed using all muon tracks that have been found. It parametrizes the offsets between prediction and candidate in the four variables, two transverse coordinates and two transverse slopes, as a second-degree polynomial in the two transverse position. The maximal difference with respect to the original calibration is of the order of 50 μm and 0.5 mrad in each projection. This calibration allows the inclusion of non-linear effects such as a bending of the changeable sheet with respect to the honeycomb spacer. This calibration is performed a posteriori because it profits from the largest possible statistics and anyhow does not significantly affect the position or width of the scanning window. However, it does allow for a more selective extrapolation to the special sheet. The selection of candidates for each prediction is based on the chi-squared matching probability in the four relevant variables, taking into account the second-order calibration as well as off-diagonal elements in the four-dimensional error matrix. For predictions that are found, the average number of candidates varies between 1.1 and 1.5, depending on the distance from the center of the beam and especially the angular distance from the beam direction. The fraction of predictions for which at least one candidate is found, a number often but wrongly referred to as the changeable sheet scanning efficiency, varies between 45 % for low-energy hadrons or electrons and 85 % for high-energy, small-angle muons. The inefficiency is due to fake tracks in the electronic detector reconstruction

and underestimation of the track parameter errors, either in the detector reconstruction or in the scanning, as well as genuine track finding inefficiency.

The alignment, selection and extrapolation procedures for special sheets are almost identical to those for changeable sheets. Out of the tracks extrapolated from changeable sheet, more than 80 % is found on special sheet as well.

Target sheets

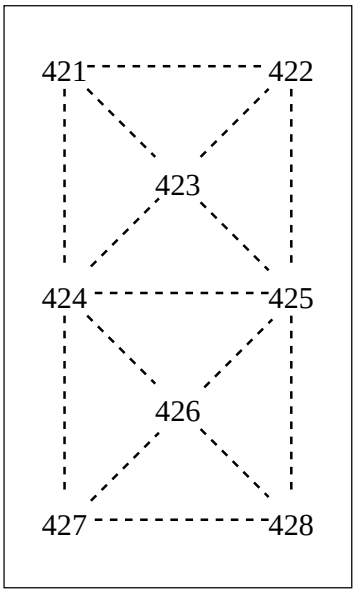


Fig. 3.26: Position of the eight track maps on the plate, each with its identifier. The dashed line corresponds to regions using the same alignment parameters, determined from the alignment of the three closest track maps.

In target sheets, the nominal reference frame is locally defined by the grid of fiducial marks. As explained in section 3.1.5, this grid is printed before development and therefore affected by distortion in the same way as tracks accumulated during the exposure. However, contrary to the X-ray marks in changeable and special sheets, the grid of fiducials does not by itself relate to any external coordinate system. The plates are positioned one after another on the film with the pattern of fiducials which ends up being slightly shifted or rotated from one plate to the next. As a result, the grid of fiducials on consecutive plates connects relative positions for two points on each plate, but not their absolute position. To determine the absolute position, or equivalently the position of the fiducial grid with respect to the tracks accumulated during the exposure, we use eight so-called track maps. A track map is an area of 1.5×1.5 mm^2 over which tracks in the beam direction are searched for with an angular acceptance of 50 mrad. This corresponds to between 800 and 1500 tracks per map, depending on the distance from the beam center. The positions of the track maps are shown in figure 3.26. Track maps are taken on the special sheet, as well as on each of the target sheets. The tracks found in the special sheet act as an initial reference, and the pattern of tracks on consecutive sheets is matched against this one. This results in a transverse and

longitudinal offset for each map on consecutive pairs of plates. In turn, these offsets are equivalent to the difference between nominal and measured positions of X-ray marks on thin sheets: for an arbitrary point, the offsets for the three closest maps are used to define the translation and rotation between consecutive plates, or more accurately of the grids on consecutive plates with respect to the tracks accumulated during exposure.

Once the grid of fiducials on consecutive plates has been related using track maps, further calibration is similar to that in changeable and special sheets. The scan-back tracks themselves are used to determine the relative position of consecutive plates, taking into account transverse displacement, rotation, longitudinal distance, and transverse shrinkage. As mentioned before, the track slopes measured in target sheets are used to select the correct track, but extrapolation from one plate to the next always uses the slope measured in special sheet. Instead of the single calibration used on thin sheets, thick sheets are subdivided in four, six, or eight rectangular regions which are calibrated independently. This is necessary to reach the 3 μm position resolution which is required. The number of regions depends on the track density and is chosen such that each region contains at least 150 predictions.

3.5.7 Vertex analysis

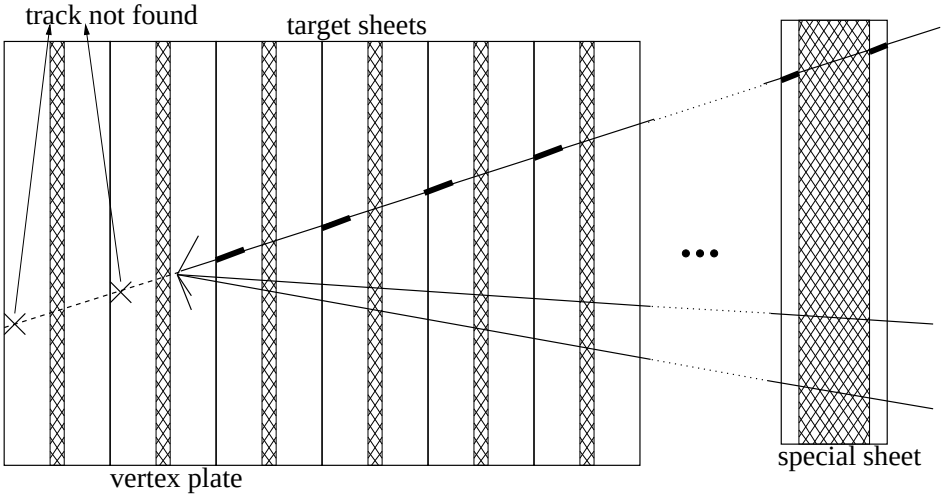


Fig. 3.27: Schematic representation of the scan-back procedure. The track is followed up, from downstream to upstream in each consecutive plate.

The data taking procedures described in sections 3.5.4 and 3.5.5 refer to the scan-back in thin sheets and thick sheets, respectively. The initial prediction is a track reconstructed in the electronic detectors and is searched for in the changeable sheet. Any candidates are then searched for in the special sheet, and then followed upstream in each consecutive plate, as illustrated in figure 3.27. In each of the target plates, the track is searched for at the upstream surface. As soon as it is missed in two consecutive plates, this constitutes a vertex trigger and the first plate where the track was missed is called the vertex plate. There are three possible causes for the track to be missed.

Inefficiency. The track may exist at the upstream surface of the vertex plate, but missed in the scanning due to track finding inefficiency. On top of that, because of alignment

problems or because of a wrong measurement of a nearby fiducial, the data may be taken at the wrong position, again resulting in the loss of the track. The total inefficiency per plate varies between 6 % and 10 %, depending on the track angle, the distance from the edge of the plate, the quality of nearby fiducials, and the number of tracks available for alignment. For a vertex trigger which requires the track to be missed in two consecutive plates, the inefficiency drops below 1 %.

Primary vertex. The scan-back track may originate from a neutrino vertex in the vertex plate.

Secondary vertex. The scan-back track may lead to a secondary vertex in the vertex plate, with the neutrino vertex further upstream. In this case, the parent of the scan-back track is either neutral or makes an angular difference to the scan-back track larger than the angular acceptance.

The aim of the vertex analysis is precisely to distinguish between these three cases. For interesting events, in particular those which have characteristics compatible with a secondary vertex, manual checks may still follow. The fraction of events which needs to be manually checked depends on the type of analysis, varying from 10 % in the case of the oscillation search to 3 % in the case of a high purity charm selection. For the majority of events though, the vertex analysis is the last emulsion scanning procedure.

Contrary to the scan-back procedures, the vertex analysis can no longer be restricted to the angle of the scan-back track and the algorithm of section 3.5.3 is no longer applicable. This immediately implies that the track finding can no longer be performed on the computer controlling the microscope as this would dramatically slow down the data taking. Instead, as indicated in section 3.5.1, the grains are stored directly in the database, including the information necessary to position them in any appropriate coordinate system. The recursive track finding algorithm of section 3.5.3 is applied in a farm of PC's, either on-line if signalled by a dispatcher message from the microscope or in batch mode at a later point in time.

Two types of data are taken: volume data and net scan data. Volume data consists of a single view at the position where the scan-back track was expected in the vertex plate, across the entire thickness on both emulsion sides. The main advantage of this procedure is the fact that the track is known to exist on the downstream surface, simply because it has been found on the upstream surface of the plate immediately downstream from the vertex plate. Using the reconstruction across the entire thickness of the plate, it is possible to determine the fate of this track. In particular, if the loss of the track was due to inefficiency, it is easily recovered using volume data: track finding starting from the downstream surface can accomodate any type of distortion and if the track is not found at the downstream surface, this provides a strong indication that the position of the view was wrong because of alignment problems or a wrong measurement for a nearby fiducial. Conversely, if the track is found to originate in the vertex plate, the presence of a parent track can be immediately detected. In summary, there are four possible outcomes for the vertex analysis using volume data.

- The track is found and reaches the upstream surface. The loss during scan-back was due to inefficiency and the track parameters determined from volume data are extrapolated to the plate upstream to continue scan-back.

- The track is not found at the downstream surface. Typically, entire regions are affected, for instance with a common fiducial for the local transformation between stage and

nominal coordinates, and recovery requires a manual intervention.

- The track is found at the downstream surface, stops in the vertex plate, and is attached to a charged primary. The track parameters of the primary are extrapolated to the plate upstream and scan-back continues using this track to find the neutrino vertex.
- The track is found at the downstream surface, stops in the vertex plate, and is not attached to a charged primary. The vertex is classified as the primary neutrino vertex. Further analysis is done using net scan data.

Net scan data taking was introduced by the Nagoya FKEN laboratory to fully exploit the capabilities of the Ultra Track Selector (UTS). The UTS relies on the original track selector algorithm, but applies it in massively parallel fashion to reach large angular acceptance, up to 400 mrad in both projections, without affecting the data taking time. Using 16 images from a 120 Hz CCD camera of 512×480 pixels together with the Tiyoda 50× objective described in section 3.4.3, a field of view of 150×120 μm^2 is covered in 0.3 s, including UTS track finding up to 400 mrad. The performance of the CERN microscopes is similar, but achieved very differently: the time taken per view is larger by a factor of ten, but the area covered per view is also larger by about a factor of ten. The track finding result is not available immediately after data taking, but this is not required either. On the other hand, because the grain data is stored, track finding can be repeated using different parameters or algorithms would the need arise.

In any case, either system is well geared to the scanning of large areas at an angular acceptance up to 400 mrad, or more specifically for net scan data taking, where NET might stand for Nagoya Emulsion Tracker or simply to refer to the fact that a certain area is covered using a net of adjacent microscope views. A layer at the surface is covered over a certain area in a number of consecutive plates, picking up all tracks with angles less than 400 mrad, regardless of any prediction. In CHORUS, the net scan volume covers an area of 1.5×1.5 mm^2, centered on the position in the vertex plate where the scan-back track has stopped, and extends over eight plates, one upstream from the vertex plate, the vertex plate itself, and six plates downstream from the vertex plate. Obviously, the large majority of reconstructed tracks are completely independent of the event under study, and have simply been accumulated during the two year exposure of the CHORUS target sheets. Nevertheless, using the information from the electronic detectors, it is possibly to select those trajectories that belong to the event and this allows for a particularly detailed track and vertex reconstruction of the neutrino event. Essentially any decay topology, including decays of neutral particles, can be studied using net scan data, limited only by the fiducial volume defined by the scanning window and by the measurement precision. The reconstruction of net scan data and the selection of secondary vertices will be the subject of the next chapter.

4. EVENT RECONSTRUCTION IN EMULSION

4.1 The net scan procedure

In the previous chapter, we have seen how a track reconstructed in the electronic detectors is used to locate the vertex: the track parameters define a position and direction on the changeable sheet, from where the track is searched for in one plate after the other, moving upstream. As soon as it is not found in two consecutive plates, the first of these is referred to as the vertex plate because this is presumably the plate where the neutrino interacted and produced the track that was followed. For each event where a muon is reconstructed in the electronic detectors, the muon track parameters provide the initial prediction for the vertex location procedure. Once the vertex plate has been identified from the so called *scanback* track, additional scanning is performed to further analyze the event. The primary interest concerns the identification of short-lived particles, for the oscillation search as well as for studies of charmed particles produced in neutrino interactions. In the course of the CHORUS experiment, various scanning strategies have been adopted for the vertex analysis, each of them with their own advantages and drawbacks. All of them were ultimately limited by the scanning capacity available: each procedure was meant to reach the best physics performance for a given volume scanned at a given angular acceptance.

In parallel, development of the scanning techniques themselves continued, specifically to increase the speed of automatic scanning. Ultimately, the point has been reached where a large volume around the assumed vertex position can be scanned at wide angular acceptance for each event, rendering all previous procedures obsolete. In the following, we will be concerned exclusively with this type of scanning, referred to as net scan data taking [116]. Figure 4.1 shows the fiducial volume, indicating the track which was used in the vertex location. A surface of 1.5×1.5 mm^2 is covered in an area scan on eight consecutive plates: the plate immediately upstream from the vertex plate, the vertex plate itself, and the six plates downstream from the vertex plate.

As indicated, the scanback track will typically have been found on the upstream side of all plates downstream from the vertex plate, and the first question to be asked concerns the fate of the scanback track, indicated by the question mark in the figure. As mentioned before, the track may have been missed in two consecutive plates due to inefficiency or it may have been a background track picked up at a previous plate. If the neutrino vertex can be confirmed, the second question concerns the presence of short-lived particles, produced at the neutrino vertex and decaying within the first couple of millimeters. For each of the plates in figure 4.1, the area shown is scanned on the upstream side, starting from 10 μm to 20 μm below the surface and covering a depth of around 100 μm. This is indicated by the arrows at the top of the figure. For the most upstream plate where the scanback track has still been found, the area is centered on that position. On the other plates, the scanning area is centered across the same position extrapolated along the neutrino beam direction, which

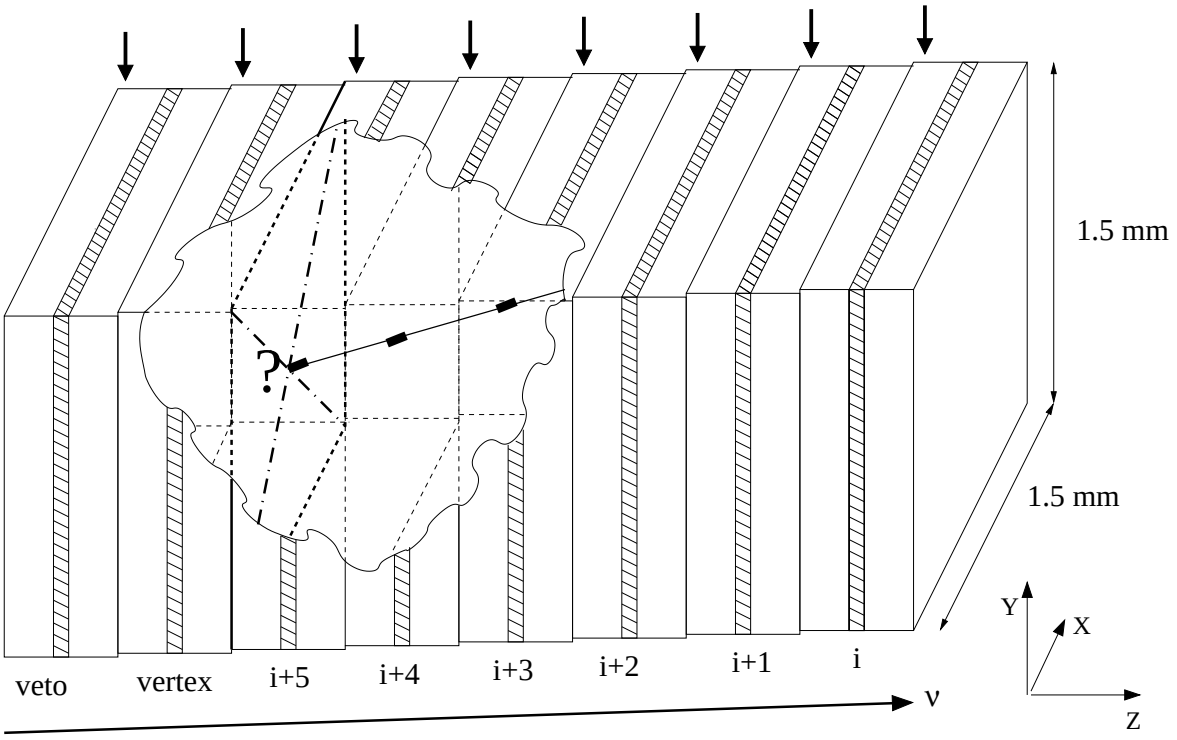


Fig. 4.1: The fiducial volume for net scan data taking. The hatched area shows the 80 μm plastic base for each plate, further details are given in the text.

has an upward slope of 42 mrad. The angular acceptance of the scanning is 400 mrad with respect to the Z axis.

The most upstream plate is called the veto plate because it serves to reject passing tracks. For a transverse area of 1.5×1.5 mm^2 exposed during two years, there are of the order of a thousand straight tracks with slopes less than 400 mrad. With the exception of the muons from nearby test beams discussed in section 3.1.5, the majority of these are particles produced in neutrino interactions further upstream, both inside and outside the CHORUS detector. Any track which is seen at the upstream surface of the veto plate, *i.e.* upstream of the primary vertex, is considered passing and cannot possibly belong to the event under consideration. However, the passing tracks do provide a sufficiently large sample to determine the precise alignment from one plate to the next.

Net scan analysis treats the emulsion very much like a tracking detector rather than as a visual detector. In electronic detector language, the plates would be called 'layers' and the segments in these plates would be called 'hits'. However, the analogy is not complete and there are at least three important differences:

- at the longitudinal position of the plate, each segment measures both transverse coordinates, in contrast to electronic trackers that are usually projective, measuring only one of the two transverse coordinates;

- even though the transverse position of each segment contains more information than the transverse slope, the slope information is by no means irrelevant;

- given the exceptional spatial resolution in emulsion and its high density, the effect of multiple scattering from one plate to the next is prominent: for a 1 GeV/c track, the displacement[1] induced by multiple scattering in a single plate is ten times larger than the measurement error.

The first two points clearly are advantages with respect to conventional detectors. The last point is a drawback from the point of view of track fitting, even though it does in principle offer the possibility of estimating the momentum even in the absence of a magnetic field. In any case, any track fitting algorithm should exploit all available information while accounting for the effects of multiple Coulomb scattering (MCS) on the optimum track parameters as well as on their covariance matrix.

It is worth noting that the importance of the multiple scattering contribution is intimately related to the exceptional spatial resolution in emulsion. In absolute terms, the effect is fairly small: net scan analysis considers lengths in emulsion of at most 10 mm, corresponding to a scattering-induced transverse displacement of 30 μm for a 1 GeV/c track. This is to be compared with a position measurement error of the order of 0.3 μm.

Each track is measured at most once per plate, always at the upstream surface, and the geometric layout of each plate is identical. To allow for the treatment of the emulsion as a homogeneous medium, an effective radiation length is defined based based on the radiation lengths of the base and emulsion materials and on the fraction of the plate thickness they represent. Energy loss, about 5 MeV for a minimum ionizing particle traversing 10 mm in emulsion, will be implicitly neglected.

In the event reconstruction, our aim will be to exploit the similarities between the emulsion and other tracking detectors by applying what might be called textbook techniques. The ground work for this consists in specifying an adequate model for the measurement errors, economical in the number of parameters but sufficient to describe all relevant features of the data. Section 4.2 will show that such a model can indeed be found. As we have argued above, the errors due to multiple scattering are at least as important as the measurement errors and section 4.3 will be devoted to the derivation of the covariance matrix for multiple scattering, specific to the geometry considered here but without a single numerical approximation. Once the errors for the input data are correctly described, analytical expressions can be found for just about every quantity of interest. Section 4.4 shows how the segments are combined to a track, section 4.5 how the probability can be formulated that a given emulsion track corresponds to a track reconstructed in the electronic detectors, and section 4.6 how tracks can be combined to form a vertex.

Once we have the tracks, attached to vertices and matched to the electronic detectors, the event can be analyzed for the presence of short-lived particles. One particularly interesting quantity in this respect is the impact parameter or distance of closest approach, to be discussed in section 4.7. Throughout the text, the calculations will be illustrated using the actual data but formulated as generally as possible. Only at the end is the actual reconstruction described in more detail, in section 4.8 for what concerns the assignment of emulsion tracks to vertices and to detector tracks, in section 4.9 for what concerns the selection of short-lived particles.

A significant part of the work described here was triggered by the improvements obtained from a rigorous application of textbook track fitting to the CHORUS electronic tracking detectors, described in [117]. A particularly exhaustive reference in this context is [118].

[1] shorthand for 'the width of the distribution of displacement ..'

The basic formalism to incorporate multiple scattering in the error matrix of a track fit was derived nearly forty years ago [119] and was comprehensively reviewed in [120]. Beyond this basic formalism, more recent work on the subject has generally been concerned with computational efficiency [121] or the effects of a magnetic field [122, 123] or both [124]. For net scan data, the basic formalism is sufficient and we merely extend it to incorporate the case where each hit contains not only position but also slope information.

All of the algorithms described were implemented in C++ for the CHORUS net scan data. The majority was initially written as Mathematica functions, especially convenient in the development phase thanks to its built-in support for matrix algebra and statistical distributions as well as its visualization capabilities [125].

4.2 Description of the measurement errors

The coordinate of a segment has four components, two for the transverse position and two for the transverse slope, all given at a specific longitudinal position. This can be written as

$$\vec{x}_z = \begin{pmatrix} x \\ y \\ a_x \\ a_y \end{pmatrix}_z \quad , \tag{4.1}$$

with the subscript z indicating the longitudinal position and $a_{x,y}$ the slopes with respect to the longitudinal axis, defined as

$$a_x \;=\; \frac{dx}{dz} \;\text{ and }\; a_y \;=\; \frac{dy}{dz} \;. \tag{4.2}$$

The longitudinal position is treated as a parameter rather than as a variable and no error is assigned to it. Implicitly, the measurement error on the longitudinal position will be absorbed in the error assigned to the transverse position. In turn, this leads to a correlation between the two transverse projections.

The error contributions from the longitudinal and the transverse measurement can be disentangled by transforming the segment residuals from the orthogonal axes X and Y to the orthogonal axes R and L, which stand for radial and lateral, respectively. The radial axis is defined as the projection of the track direction onto the X,Y plane; the lateral axis is taken perpendicular to both R and Z. This transformation corresponds to a rotation of the X,Y plane over the azimuth angle of the track, or formally

$$\begin{pmatrix} r \\ l \end{pmatrix} = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{with} \quad \phi = \arctan\frac{a_y}{a_x} \quad . \tag{4.3}$$

Any error in the longitudinal position translates into an error in the radial position r without affecting the lateral position l. The data show no correlation between r and l and the covariance matrix is taken to be diagonal. Using the transformation matrix T for the rotation over the azimuth angle, we find the covariance matrix for x and y:

$$C_{x,y} \;=\; T.C_{r,l}.T^T \;=\; \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_l^2 \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \tag{4.4}$$

$$= \begin{pmatrix} \cos^2\phi\,\sigma_r^2 + \sin^2\phi\,\sigma_l^2 & \sin\phi\cos\phi\,(\sigma_r^2 - \sigma_l^2) \\ \sin\phi\cos\phi\,(\sigma_r^2 - \sigma_l^2) & \sin^2\phi\,\sigma_r^2 + \cos^2\phi\,\sigma_l^2 \end{pmatrix} \quad . \tag{4.5}$$

The correlation coefficient

$$\rho = \frac{c_{xy}}{\sigma_x \sigma_y} \tag{4.6}$$

between the X and Y projections depends on the azimuth angle and is proportional to the difference between the radial and lateral variances. It is zero for $\phi = -\pi/2, 0, \pi/2, \pi$ as $\cos\phi = 0$ or $\sin\phi = 0$. The extrema of $(\sigma_r^2 - \sigma_l^2)/(\sigma_r^2 + \sigma_l^2)$ and $(\sigma_l^2 - \sigma_r^2)/(\sigma_r^2 + \sigma_l^2)$ are reached for $\phi = -3\pi/4, \pi/4$ and $\phi = -\pi/4, 3\pi/4$ corresponding to $\sin\phi = \cos\phi = \pm 1/\sqrt{2}$ and $\sin\phi = -\cos\phi = \pm 1/\sqrt{2}$, respectively. Figure 4.2 shows that the expected behaviour is indeed observed in the data.

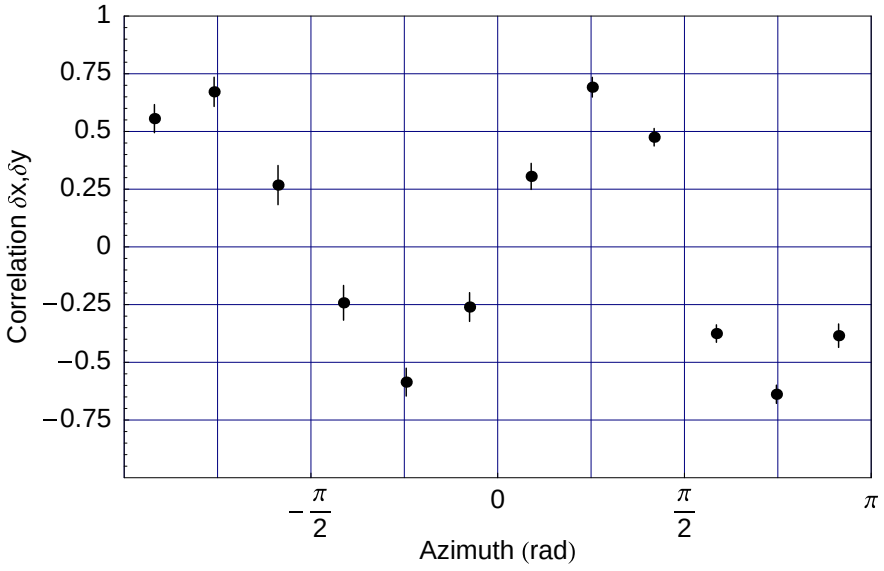


Fig. 4.2: Correlation coefficient ρ between the X and Y projections for the residual of a segment with respect to a track, as a function of the track azimuth. The vertical lines extend over $\pm 1/\sqrt{N}$, where N is the number of segments used to determine the correlation in each bin of ϕ. The data sample is described in the text.

All figures in this section correspond to a sample of 3042 tracks selected from the data of modules 221 and 260. Each of these tracks has a segment in six consecutive plates and is matched to a track in the electronic detector with a momentum, measured in the electronic detectors, of at least 2 GeV/c. The study of the segment measurement errors is done using the third and fourth segment on the track as the error on track parameters determined from interpolation is smaller than the error on track parameters determined from extrapolation. The residual is defined as the difference between the measured coordinates of the segment and the track parameters determined from a fit of the remaining five segments. The condition that the probability for the fit using the five remaining segments exceed 0.1 % excludes 47 of the 6084 segments that can be used. For properly estimated errors, the width of the residual distribution will be equal to the sum of squares for the segment measurement error and the fit error. The track sample is chosen precisely to minimize the contribution to the residual from the fit. The track fit itself as well as the matching between emulsion tracks and

electronic detector tracks will be described later in this chapter and do of course depend on the model for the segment measurements which will be elaborated in this section.

The fact that the error on the radial position contains contributions from both the transverse (x,y) and the longitudinal (z) measurement immediately implies a dependence on the absolute value of the track slope $a = \sqrt{a_x^2 + a_y^2}$. In particular, one expects for the radial error a constant term from the corresponding transverse measurement and a term proportional to a from the longitudinal measurement. Again, this prediction is borne out by the data, as illustrated in figure 4.3. The error on the lateral position is independent of the track direction whereas the error on the radial position is independent of the track azimuth but increases linearly as a function of the track slope. For tracks parallel to the Z axis, any error in the longitudinal measurement should leave the transverse position unaffected and the radial error indeed becomes equal to the lateral error as $a \to 0$.

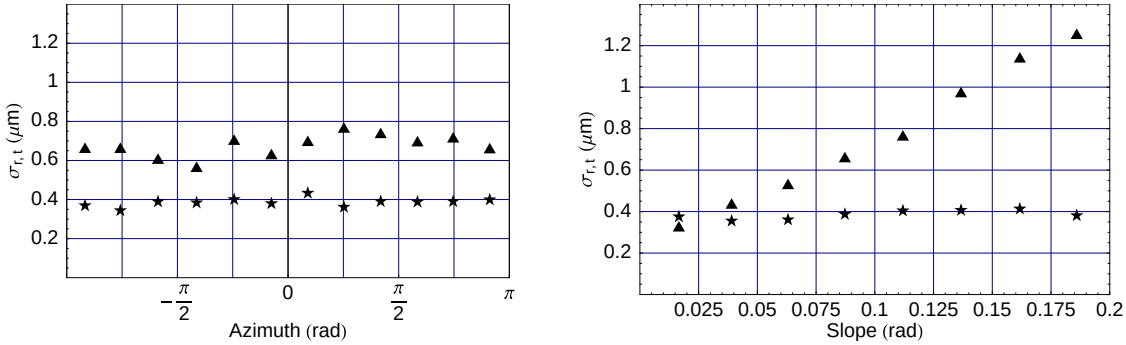


Fig. 4.3: Left panel: width of the distribution of radial (triangles) and lateral (stars) position residuals, as a function of the track azimuth angle ϕ. Right panel: width of the distribution of radial (triangles) and lateral (stars) position residuals, as a function of the track slope $a = \sqrt{a_x^2 + a_y^2}$. The data sample is described in the text.

In summary, we are led to a description for the segment measurement errors in terms of a constant lateral error σ_l and a radial error σ_r, linear in the track slope a, with no correlation between them. The covariance matrix for the transverse position, defined in equation 4.4, becomes

$$C_{x,y} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} \sigma_l^2 + a^2\sigma_r'^2 & 0 \\ 0 & \sigma_l^2 \end{pmatrix} \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} , \tag{4.7}$$

where $\sigma_r^2 = \sigma_l^2 + a^2\sigma_r'^2$ uses the fact that the constant term is equal for the radial and transverse direction.

The above considerations for the measurement errors on the transverse position apply equally to the errors on the slopes with respect to the Z axis. Again, we find that rotation over the track azimuth ϕ transforms the slopes a_x, a_y with correlated errors to slopes a_r, a_l with uncorrelated errors. As in the case for the position, the error on a_l is constant whereas the error on a_r depends linearly on the track slope a. None of this is particularly surprising if one considers that the slope measurement corresponds simply to position measurements on successive images taken along the Z direction.

However, one more point deserves particular attention: the possible correlation between the transverse position and the transverse slope in the same projection. In the fit of a straight line to a series of points, one obtains an offset and a slope. The offset can be given at any position but only for the center of the sequence of points are the errors on the offset and on the slope independent. The track selector algorithm is based on the sum for a sequence of zero-suppressed images, after shifting each image by an amount proportional to the distance in depth from the first image. To cover the angular acceptance, this procedure is repeated on the same sequence of images for a large number of shifts. A peak in any of the sums corresponds to a track, traversing the first image at the position of the peak with transverse slopes proportional to the shift in the last image. This determination of the position and slope for a track, in each of the transverse directions, is mathematically equivalent to a straight line fit with offset and slope given at the first point and one expects to find a correlation between the errors for the segment position and slope. Net scan data are taken on the upstream side of the emulsion, with the first image slightly below the surface and subsequent images in depth at a spacing of several μm. As the Z direction runs from upstream to downstream, the correlation between the position and the slope in the first, or most upstream, image is negative. Figure 4.4 shows this correlation in both the radial and the lateral direction. The correlation is significantly smaller in the radial direction because the error on the radial position contains a significant contribution from the error on the longitudinal position which is not correlated to the slope measurement. For perpendicular tracks, the correlation is identical in both directions.

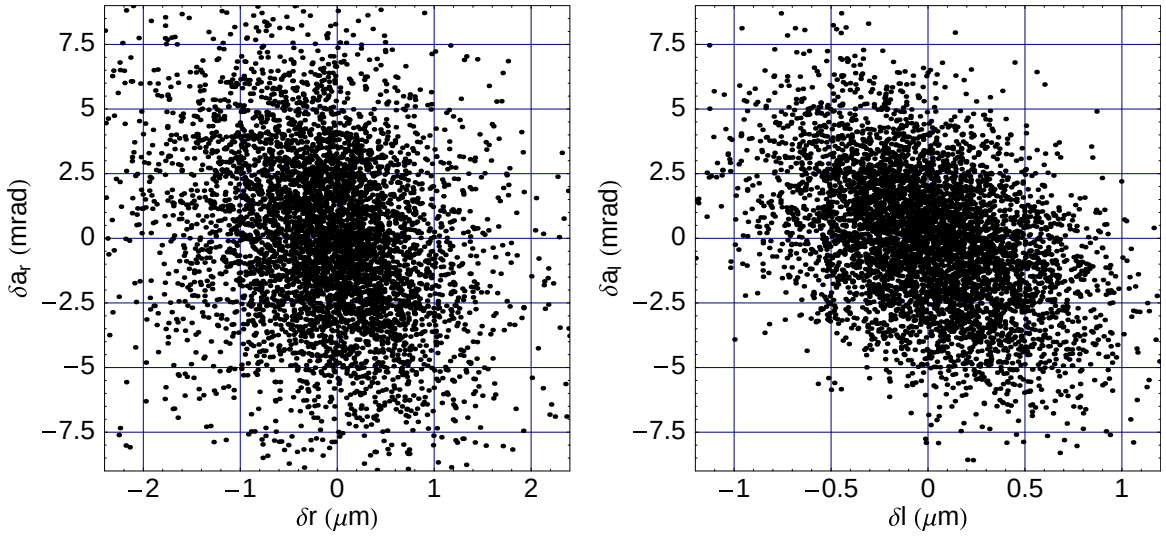


Fig. 4.4: Left panel: correlation between the residuals for the radial position r and the radial slope a_r. The correlation coefficient for the data shown is -27%; for tracks with a slope smaller than 50 mrad, it increases to -42%. Right panel: correlation between the residuals for the lateral position l and the lateral slope a_l. The correlation coefficient for the data shown is -43%. The data sample is described in the text.

Combining all of the above, the covariance matrix for the four-dimensional coordinate

Module	σ_l (μm)	σ'_r (μm/rad)	σ_{a_l} (mrad)	σ'_{a_r} (mrad/rad)	ρ_r	ρ_l
210	0.37	5.8	2.9	24	-.48	-.19
211	0.35	6.2	3.1	21	-.46	-.15
220	0.51	4.4	2.7	21	-.33	-.22
221	0.32	4.8	2.7	23	-.47	-.28
250	0.35	6.0	3.0	21	-.48	-.18
251	0.51	4.5	3.0	22	-.33	-.14
260	0.32	5.7	2.7	20	-.45	-.11
261	0.32	4.1	2.6	24	-.45	-.36

Tab. 4.1: Numerical values for the constants in the model for the segment measurement errors, determined from the data in each of the halfmodules of stack 2.

(x, y, a_x, a_y) of a segment can be written as

$$T^T \cdot \left[\begin{pmatrix} \sigma_l^2 & 0 & \rho_l\sigma_l\sigma_{a_l} & 0 \\ & \sigma_l^2 & 0 & \rho_l\sigma_l\sigma_{a_l} \\ & & \sigma_{a_l}^2 & 0 \\ & & & \sigma_{a_l}^2 \end{pmatrix} + a^2 \begin{pmatrix} \sigma_r'^2 & 0 & \rho_r\sigma'_r\sigma'_{a_r} & 0 \\ 0 & 0 & 0 & 0 \\ & & \sigma_{a_r}'^2 & 0 \\ & & & 0 \end{pmatrix} \right] \cdot T \quad , \tag{4.8}$$

$$\text{with} \quad T = \begin{pmatrix} c_\phi & s_\phi & 0 & 0 \\ -s_\phi & c_\phi & 0 & 0 \\ 0 & 0 & c_\phi & s_\phi \\ 0 & 0 & -s_\phi & c_\phi \end{pmatrix} \quad \text{and} \quad \begin{array}{rcl} c_\phi & = & \cos\phi = a_x/a \\ s_\phi & = & \sin\phi = a_y/a \\ a & = & \sqrt{a_x^2 + a_y^2} \end{array} \quad . \tag{4.9}$$

In total, the error model contains six numerical constants which need to be extracted from the data: three for the constant term that appears in both the radial and the lateral direction and three for the term in the radial direction proportional to the track slope. In each case, the three parameters are a position error, a slope error, and the correlation between them. These constants depend on the characteristics of the emulsion as well as on the hardware performance of the microscope. Generally, net scan data taking for all plates of a single halfmodule is performed on a single microscope. Hence, it is only natural to determine the parameters of the error model for each halfmodule, but applying the same set to all data in that module. This determination essentially follows the line of reasoning shown in this section, but for two minor differences. First, the available number of tracks may be up to a factor of four smaller. The 6037 segments entering the illustrations in this chapter are obtained by combining the data from two, central halfmodules and the event density in modules away from the beam center can be up to a factor of two smaller. Second, when determining the numerical constants from residual distributions, adjustment must be made for the error from the fit. This has been glossed over in the text but is readily implemented by subtracting the mean fit error in quadrature from the width of the residual distribution to obtain the segment measurement error. To illustrate the size of the measurement errors and the variation between modules, table 4.1 shows the numerical values of the eight constants for the eight halfmodules of stack 2.

4.3 The effect of multiple Coulomb scattering

A charged particle, traversing a thickness l in a material of radiation length X_0 will experience both a deviation and a displacement due to the combined effect of a large number of small-angle scatters. At least for small angles, the deviation is roughly Gaussian with a width θ_0 for the projected angular distribution of

$$\theta_0 = \frac{13.6\text{MeV}}{\beta cp} Z\sqrt{l/X_0}\,[1 + 0.038\log(l/X_0)]\ , \tag{4.10}$$

where p, βc and Z are the momentum, velocity and charge number of the traversing particle. To separate the part which characterizes the material and the particle from the part related to the geometry of the problem, we define the scattering angle per unit length as

$$\sigma_0 = \theta_0/\sqrt{l} = \frac{13.6\text{MeV}}{\beta cp} Z\sqrt{1/X_0}\,, \tag{4.11}$$

where the logarithmic term has been neglected. This term contributes less than 10 % in the region $0.072\,X_0 < l < 14\,X_0$. For our purposes, it can be safely neglected: for small values of l the effect of multiple scattering is anyhow small, values of l larger than $1\,X_0$ never occur in the analysis of net scan data.

A priori, we have no information about the particle and we will simply assume that it has momentum 1 GeV/c, is relativistic ($\beta = 1$) and carries charge number 1. The last two conditions are certainly fulfilled for all cases of interest. As for the assumed value of the momentum, this translates into an implicit selection of high-momentum tracks. Tracks with momenta lower than 1 GeV/c experience more scattering than we assume, and the probabilities assigned to them will correspondingly be too low. For any emulsion tracks that are matched to tracks in the electronic detectors with a momentum measurement, the measured value will be used as momentum hypothesis in the track fit in emulsion provided that the measured momentum is larger than 1 GeV/c. The effect of this will be discussed at the end of section 4.4.

Within the same approximations as for the deviation, the displacement in a direction perpendicular to the track is Gaussian with a width of $1/\sqrt{3}l\theta_0$. Because part of the displacement is induced by the deviation, there exists a correlation between the two: the correlation coefficient is $\sqrt{3}/2$. For the purposes of error propagation, it is more convenient to work with independent Gaussian random variables (r_1, r_2) with mean zero and variance one, and then set

$$\begin{aligned}
\text{displacement } \delta x &= r_1 l\theta_0/\sqrt{12} + r_2 l\theta_0/2 && (4.12)\\
&= r_1 l^{3/2}\sigma_0/\sqrt{12} + r_2 l^{3/2}\sigma_0/2 && (4.13)\\
\text{deviation } \delta a_x &= r_2\theta_0 && (4.14)\\
&= r_2 l^{1/2}\sigma_0\ . && (4.15)
\end{aligned}$$

Strictly speaking, the deviation is an angle rather than a slope but the small angle approximation is certainly justified in this case. From these expressions, the covariance matrix due to multiple scattering for the position and slope of a track in a direction perpendicular to the track is immediately found as

$$C_{x,ax} = \sigma_0^2 \begin{pmatrix} l^3/3 & l^2/2 \\ l^2/2 & l \end{pmatrix}\ . \tag{4.16}$$

For orthogonal directions in the plane perpendicular to the track, the covariance matrices are identical and there exists no correlation between the two directions. However, as mentioned in the section about the measurement errors for a segment, we wish to treat the longitudinal position as a parameter rather than as a variable. Because the covariance matrix given above refers to a direction perpendicular to the track rather than to a direction perpendicular to the longitudinal axis, it is not directly applicable to our problem.

To find the multiple scattering covariance matrix for the directions perpendicular to the longitudinal axis, we first consider the slopes (a_x, a_y). In a second step, we will treat the positions (x, y). The unit vector pointing along the direction of the undisturbed track is

$$\vec{t} = \begin{pmatrix} a_x/s \\ a_y/s \\ 1/s \end{pmatrix} \quad \text{with} \quad s = \sqrt{1 + a_x^2 + a_y^2} = \sqrt{1 + a^2} \ . \tag{4.17}$$

In terms of two uncorrelated multiple scattering angles σ_1 and σ_2, the disturbed track direction is given by

$$\vec{t'} = \vec{t} + \sigma_1 \vec{t_1} + \sigma_2 \vec{t_2} \ , \tag{4.18}$$

where $\vec{t_1}$ and $\vec{t_2}$ are orthogonal unit vectors in the plane perpendicular to $\vec{t}$, for instance

$$\vec{t_1} = \begin{pmatrix} -a_y/a \\ a_x/a \\ 0 \end{pmatrix} \text{ and } \vec{t_2} = \begin{pmatrix} -a_x/(as) \\ -a_y/(as) \\ a/s \end{pmatrix} \ . \tag{4.19}$$

The Jacobian of $\vec{t'}$ with respect to $\vec{\sigma}$ is given by

$$\mathcal{J}(\vec{\sigma} \to \vec{t}) = \frac{\partial \vec{t}}{\partial \vec{\sigma}} = \begin{pmatrix} -a_y/a & -a_x/(as) \\ a_x/a & -a_y/(as) \\ 0 & a/s \end{pmatrix} \tag{4.20}$$

while the covariance matrix for $\vec{\sigma}$ is simply

$$C_{\vec{\sigma}} = \begin{pmatrix} \sigma_0^2 l & 0 \\ 0 & \sigma_0^2 l \end{pmatrix} \ . \tag{4.21}$$

Error propagation then leads to the covariance matrix for $\vec{t'}$ as

$$C_{\vec{t}} = \mathcal{J}_{\vec{\sigma}\to\vec{t}} C_{\vec{\sigma}} \mathcal{J}^T_{\vec{\sigma}\to\vec{t}} = \sigma_0^2 \frac{l}{s^2} \begin{pmatrix} 1 + a_y^2 & -a_x a_y & -a_x \\ -a_x a_y & 1 + a_x^2 & -a_y \\ -a_x & -a_y & a^2 \end{pmatrix} \ . \tag{4.22}$$

The transformation of the unit direction vector (t_x, t_y, t_z) back to the vector of slopes $(a_x = t_x/t_z, a_y = t_y/t_z)$ is described by the Jacobian matrix

$$\mathcal{J}(\vec{t} \to \vec{a}) = \frac{\partial \vec{a}}{\partial \vec{t}} = \begin{pmatrix} s & 0 & -a_x s \\ 0 & s & -a_y s \end{pmatrix} \ , \tag{4.23}$$

where the normalization of the direction vector $t_x^2 + t_y^2 + t_z^2 = t_z^2 s^2 = 1$ was used. Finally, we find the covariance matrix for the slopes with respect to the longitudinal axis as

$$C_{\vec{a}} = \mathcal{J}_{\vec{t}\to\vec{a}} C_{\vec{t}} \mathcal{J}^T_{\vec{t}\to\vec{a}} = \sigma_0^2 l s^2 \begin{pmatrix} 1 + a_x^2 & a_x a_y \\ a_x a_y & 1 + a_y^2 \end{pmatrix} \ . \tag{4.24}$$

A similar procedure can be applied to find the covariance matrix for the displacement with respect to the longitudinal axis. With the origin of the coordinate system at the track position at $z = z_0$, the position of the undisturbed track at $z = z_1$ is

$$\vec{t} = \begin{pmatrix} a_x z_1 \\ a_y z_1 \\ z_1 \end{pmatrix} \quad . \tag{4.25}$$

Using the same set of orthogonal unit vectors in the plane perpendicular to the track, the disturbed position can be written as

$$\vec{t'} = \vec{t} + \sigma_1 \vec{t}_1 + \sigma_2 \vec{t}_2 = \begin{pmatrix} a_x z_1 \\ a_y z_1 \\ z_1 \end{pmatrix} + \sigma_1 \begin{pmatrix} -a_y/a \\ a_x/a \\ 0 \end{pmatrix} + \sigma_2 \begin{pmatrix} -a_x/(as) \\ -a_y/(as) \\ a/s \end{pmatrix} \quad , \tag{4.26}$$

where σ_1 and σ_2 are now two uncorrelated scattering offsets with covariance matrix

$$C_{\vec{\sigma}} = \begin{pmatrix} \sigma_0^2 l^3/3 & 0 \\ 0 & \sigma_0^2 l^3/3 \end{pmatrix} \quad . \tag{4.27}$$

The disturbed position $\vec{t'}$ no longer lies in the plane $z = z_1$ but it can be extrapolated back to this plane along the track direction by subtracting $\sigma_2 a/s$ times the direction vector. Using the fact that $a/s + 1/(as) = s/a$, the disturbed position in the plane $z = z_1$ can be written as

$$\vec{t'} = \begin{pmatrix} a_x z_1 \\ a_y z_1 \\ z_1 \end{pmatrix} + \sigma_1 \begin{pmatrix} -a_y/a \\ a_x/a \\ 0 \end{pmatrix} + \sigma_2 \begin{pmatrix} -a_x s/a \\ -a_y s/a \\ 0 \end{pmatrix} \quad , \tag{4.28}$$

where the Jacobian for the transformation from $\vec{\sigma}$ to (t_x, t_y) can be read off immediately:

$$\mathcal{J}(\vec{\sigma} \to \vec{t}) = \frac{\partial \vec{t}}{\partial \vec{\sigma}} = \begin{pmatrix} -a_y/a & -a_x s/a \\ a_x/a & -a_y s/a \end{pmatrix} \quad . \tag{4.29}$$

Error propagation then gives the covariance matrix for the transverse positions $x = t_x$ and $y = t_y$:

$$C_{\vec{t}} = \mathcal{J}_{\vec{\sigma}\to\vec{t}} C_{\vec{\sigma}} \mathcal{J}^T_{\vec{\sigma}\to\vec{t}} = \sigma_0^2 l^3/3 \begin{pmatrix} 1 + a_x^2 & a_x a_y \\ a_x a_y & 1 + a_y^2 \end{pmatrix} \quad . \tag{4.30}$$

The derivation for the multiple scattering covariance matrix for the four coordinates of a track, two transverse positions and two transverse slopes, is largely similar to that for the positions and slopes separately but needs to be carried out explicitly to find the correlation between the position and the slope of a track. The scattering contribution to the transverse position and the direction vector at $z = z_1$ can be rewritten as

$$\begin{pmatrix} x' \\ y' \\ t'_x \\ t'_y \\ t'_z \end{pmatrix} = \begin{pmatrix} a_x z_1 \\ a_y z_1 \\ \frac{1}{s}\frac{a_x}{a} \\ \frac{1}{s}\frac{a_y}{a} \\ \frac{1}{s} \end{pmatrix} + \sigma_1 \begin{pmatrix} -\frac{a_y}{a} \\ \frac{a_x}{a} \\ 0 \\ 0 \\ 0 \end{pmatrix} + \sigma_2 \begin{pmatrix} -s\frac{a_x}{a} \\ -s\frac{a_y}{a} \\ 0 \\ 0 \\ 0 \end{pmatrix} + \sigma_3 \begin{pmatrix} -\frac{l}{2}\frac{a_y}{a} \\ \frac{l}{2}\frac{a_x}{a} \\ -\frac{a_y}{a} \\ \frac{a_x}{a} \\ 0 \end{pmatrix} + \sigma_4 \begin{pmatrix} -\frac{l\,s}{2}\frac{a_x}{a} \\ -\frac{l\,s}{2}\frac{a_y}{a} \\ -\frac{1}{s}\frac{a_x}{a} \\ -\frac{1}{s}\frac{a_y}{a} \\ \frac{a}{s} \end{pmatrix} \quad , \tag{4.31}$$

where σ_1 and σ_2 are two uncorrelated scattering offsets for the part independent of the deviation and with a width $\sigma_0 l^{3/2}/\sqrt{12}$ whereas σ_3 and σ_4 are two uncorrelated scattering angles with a width of $\sigma_0 l^{1/2}$, as seen from equations 4.13 and 4.15. Error propagation from the covariance matrix for $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3, \sigma_4)$ to the covariance matrix for the track parameters $\vec{x} = (x, y, a_x, a_y)$ consists of a transformation from $\vec{\sigma}$ to $\vec{t} = (x, y, t_x, t_y, t_z)$ using the Jacobian

$$\mathcal{J}(\vec{\sigma} \to \vec{t}) = \frac{\partial \vec{t}}{\partial \vec{\sigma}} = \begin{pmatrix} -a_y/a & -a_x s/a & -\frac{l}{2} a_y/a & -\frac{l}{2} a_x s/a \\ a_x/a & -a_y s/a & \frac{l}{2} a_x/a & -\frac{l}{2} a_y s/a \\ 0 & 0 & -a_y/a & -a_x/(as) \\ 0 & 0 & a_x/a & -a_y/(as) \\ 0 & 0 & 0 & a/s \end{pmatrix} , \tag{4.32}$$

followed by a transformation from $\vec{t}$ to $\vec{x}$ with Jacobian

$$\mathcal{J}(\vec{t} \to \vec{x}) = \frac{\partial \vec{x}}{\partial \vec{t}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & s & 0 & -a_x/s \\ 0 & 0 & 0 & s & -a_y/s \end{pmatrix} . \tag{4.33}$$

Inserting the covariance matrix for $\vec{\sigma}$

$$C_{\vec{\sigma}} = \sigma_0^2 l \begin{pmatrix} l^2/12 & 0 & 0 & 0 \\ 0 & l^2/12 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} , \tag{4.34}$$

we find that the covariance matrix for $\vec{x}$ is

$$C_{\vec{x}} = \mathcal{J}_{\vec{t}\to\vec{x}} \mathcal{J}_{\vec{\sigma}\to\vec{t}} C_{\vec{\sigma}} \mathcal{J}^T_{\vec{\sigma}\to\vec{t}} \mathcal{J}^T_{\vec{t}\to\vec{x}} \tag{4.35}$$

$$= \sigma_0^2 l \begin{pmatrix} (1+a_x^2)\, l^2/3 & a_x a_y\, l^2/3 & (1+a_x^2)\, ls/2 & a_x a_y\, ls/2 \\ & (1+a_y^2)\, l^2/3 & a_x a_y\, ls/2 & (1+a_y^2)\, ls/2 \\ & & (1+a_x^2)\, s^2 & a_x a_y\, s^2 \\ & & & (1+a_y^2)\, s^2 \end{pmatrix} \tag{4.36}$$

$$= \sigma_0^2 z s^3 \begin{pmatrix} (1+a_x^2)\, z^2/3 & a_x a_y\, z^2/3 & (1+a_x^2)\, z/2 & a_x a_y\, z/2 \\ & (1+a_y^2)\, z^2/3 & a_x a_y\, z/2 & (1+a_y^2)\, z/2 \\ & & 1+a_x^2 & a_x a_y \\ & & & 1+a_y^2 \end{pmatrix} , \tag{4.37}$$

where the last equation uses the fact that the distance traversed l can be written in terms of the distance z along the longitudinal axis as $l = zs = z\sqrt{1 + a_x^2 + a_y^2}$.

4.4 Track fit

The result from the `ecfsal` alignment algorithm and the `ecvtxa` reconstruction algorithm comes in the form of tracks composed of segments, track pairs consisting of two tracks, and vertices built on the basis of one or more track pairs. These programs primarily deal with the combinatorics of the track and vertex finding problem, but leave the track and

vertex fitting unaddressed. Obviously, determining which tracks form a pair or build a vertex presupposes a determination of the track parameters, but a fairly rudimentary fit is sufficient for this purpose. The fit algorithms are limited in particular by the need to minimize computation time as they are applied to all of the tracks; on average, there are several thousand tracks in one net scan acquisition for one event. After track finding in `ecfsal` and `ecvtxa`, any tracks that leave the fiducial volume at the upstream end are removed, reducing the number of tracks by a factor of about ten. Hence, it is conceivable to perform a more computationally involved fit for each of the remaining tracks. The main gain is reliable errors on the track parameters. This greatly simplifies all subsequent steps in the reconstruction: almost any question can be reframed in terms of a χ^2 probability. Typical questions include whether a track in the emulsion is compatible in angle with a track in the electronic detectors, whether two tracks intersect, or whether a track shows a non-zero impact parameter with respect to a vertex.

Starting from the expressions for the covariance matrix of the segment measurement errors, equation 4.8, and the covariance matrix for multiple scattering, equation 4.37, it is straightforward to derive a track fit which uses all available information in a statistically correct manner. The problem can be formulated as follows: given a track composed of segments with measured coordinates $\vec{x}_i = (x, y, a_x, a_y)_i$ at n longitudinal positions $z_{i=1,..,n}$, determine the best estimate for the track parameters $\vec{x}_0 = (x, y, a_x, a_y)_0$ at z_0 and the covariance matrix for this estimate.

It is worth stressing that the choice of z_0, the longitudinal position at which the track parameters are determined, affects the outcome in a non-trivial way. Because of the multiple scattering, segments at a longitudinal position z_i will contribute more if z_i is closer to z_0. Consequently, the choice of z_0 depends on the question that is to be addressed. For a vertex fit, it will be the best-guess value for the longitudinal position of the vertex, typically situated in the plate immediately upstream from the most upstream segment on the track. In contrast, for matching to a track in the electronic detectors, it will be the most downstream plate in the emulsion module, the position at which the parameters for a track in the electronic detectors are given.

To simplify the notation, we will consider the two-dimensional problem of finding parameters $\vec{x}_0 = (x, a_x)$ at z_0 for a track with measured coordinates $\vec{x}_i = (x, a_x)_i$ at $z_{i=1,..,n}$. Obviously, the fit in the (z, y) plane is entirely analogous to that in the (z, x) plane. The expressions given below are easily generalized to the three-dimensional case, provided one introduces the full covariance matrix for the segment measurement errors as well as the correlations between the multiple scattering in the X and Y directions, proportional to $a_x a_y$. As a further simplification, the factors $1 + a_x^2$ and $s = \sqrt{1 + a_x^2 + a_y^2}$ will be omitted in the covariance matrix for multiple scattering in the (z, x) plane.

Figure 4.5 illustrates the problem for the case of two segments. The measured position and slope in the (z, x) plane are (x_1, a_1) at z_1 and (x_2, a_2) at z_2. Typically, we are interested in the track parameters at a longitudinal position z_0 that lies either upstream or downstream from all segments: $z_0 < z_{1,2}$ or $z_{1,2} < z_0$. This immediately implies that the contribution to the covariance matrix from scattering between z_0 and z_1 is correlated with the contribution from scattering between z_0 and z_2. For definiteness we choose $z_0 < z_1 < z_2$. The deviation and displacement at z_2 due to scattering now receives two contributions: the first corresponding to the longitudinal distance $Z_1 = z_1 - z_0$, the second corresponding to the longitudinal distance $Z_2 = z_2 - z_1$, as shown in figure 4.5. Strictly speaking, the traversed

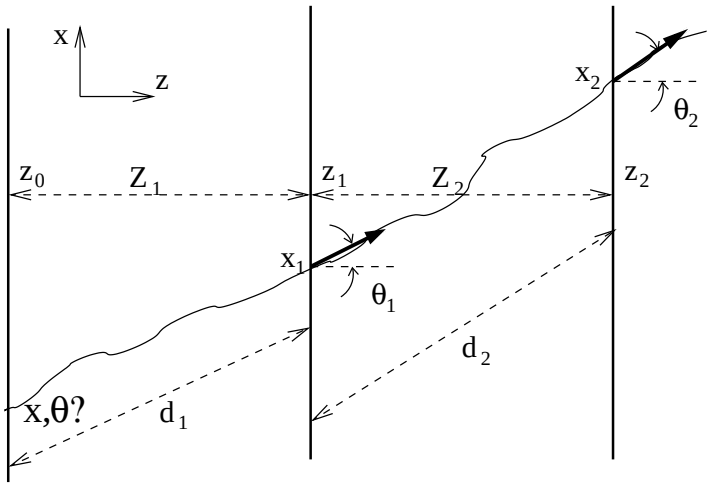


Fig. 4.5: At longitudinal position z_0, best-fit values are determined for the impact point x and transverse slope $a_x = \tan\theta$ of a track. The track consists of two segments: one with a measured impact point x_1 and slope $a_1 = \tan\theta_1$ at longitudinal position z_1 and one with a measured impact point x_2 and slope $a_2 = \tan\theta_2$ at longitudinal position z_2.

distances d_1 and d_2 differ from the longitudinal distances Z_1 and Z_2 by a factor of $\sqrt{1+a_x^2}$ or $s = \sqrt{1+a_x^2+a_y^2}$ in the three-dimensional case.

In the absence of a magnetic field, the track model is linear in the parameters $\vec{x}_0 = (x_0, a_0)$ and simply states that the estimated position and slope at z_0 are $x_i = x_0 + (z_i - z_0)a_0$ and $a_i = a_0$. In matrix notation, this can be written as $\vec{y} = A\vec{x}_0$, where

$$A = \begin{pmatrix} 1 & z_1 - z_0 \\ 0 & 1 \\ 1 & z_2 - z_0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & Z_1 \\ 0 & 1 \\ 1 & Z_1 + Z_2 \\ 0 & 1 \end{pmatrix} \tag{4.38}$$

are the model coefficients and

$$\vec{y} = \begin{pmatrix} x_1 \\ a_1 \\ x_2 \\ a_2 \end{pmatrix} \tag{4.39}$$

is the vector of measurements. For this model, the χ^2 is defined as

$$\chi^2 = (\vec{y} - A\vec{x}_0)^T V^{-1} (\vec{y} - A\vec{x}_0) \ , \tag{4.40}$$

which can be minimized (see *e.g.* section 7.2 in [126]) to yield a linear least squares estimator

$$\hat{\vec{x}}_0 = \left(A^T V^{-1} A\right)^{-1} A^T V^{-1} \vec{y} \tag{4.41}$$

with covariance matrix

$$U = \left(A^T V^{-1} A\right)^{-1} \quad . \tag{4.42}$$

The covariance matrix V for this problem is composed of two terms: the first to characterize the segment measurement errors, the second to describe the multiple scattering between z_0 and $z_{1,2}$. Using the fact that the measurement errors are independent, the corresponding term in the covariance matrix is

$$V_{meas} = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_a & 0 & 0 \\ & \sigma_a^2 & 0 & 0 \\ & & \sigma_x^2 & \rho\sigma_x\sigma_a \\ & & & \sigma_a^2 \end{pmatrix} , \tag{4.43}$$

where the errors σ_x and σ_a on the transverse position and slope in the X direction, as well as their correlation ρ are derived from the error model of equation 4.8. To find the multiple scattering term in the covariance matrix, we note that the deviation at z_2 is simply the sum of the deviations in Z_1 and Z_2 whereas the displacement at z_2 contains the sum of the displacements as well as an additional contribution from the deviation at z_1 projected on Z_2, or symbolically

$$\delta x_2 = \delta x_2' + \delta x_1 + Z_2\,\delta a_1 \tag{4.44}$$

$$\delta a_2 = \delta a_2' + \delta a_1 \ , \tag{4.45}$$

where $\delta x_2'$ and $\delta a_2'$ correspond to the scattering between z_1 and z_2 which is by definition independent of the scattering between z_0 and z_1. The covariance matrix for multiple scattering is found by applying the transformation with Jacobian

$$\mathcal{J} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & Z_2 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \tag{4.46}$$

to the block-diagonal covariance matrix for the independent scattering contributions

$$\begin{pmatrix} \sigma_{x_1}^2 & \rho\,\sigma_{x_1}\sigma_{a_1} & 0 & 0 \\ & \sigma_{a_1}^2 & 0 & 0 \\ & & \sigma_{x_2}^2 & \rho\,\sigma_{x_2}\sigma_{a_2} \\ & & & \sigma_{a_2}^2 \end{pmatrix} , \tag{4.47}$$

where $\sigma_{x_{1,2}}$ and $\sigma_{a_{1,2}}$ are the scattering offsets and angles for $Z_{1,2}$ whereas $\rho = \sqrt{3}/2$ is the correlation between a scattering offset and the corresponding scattering angle. Using the identities $\sigma_a = \sigma_x\sqrt{3}/Z$ and $\sigma_{x_1}^2/Z_1^3 = \sigma_{x_2}^2/Z_2^3$, this leads to the covariance matrix

$$V_{mcs} = \begin{pmatrix} \sigma_{x_1}^2 & \rho\,\sigma_{x_1}\sigma_{a_1} & \sigma_{x_1}^2 + Z_2\rho\,\sigma_{x_1}\sigma_{a_1} & \rho\,\sigma_{x_1}\sigma_{a_1} \\ & \sigma_{a_1}^2 & \rho\,\sigma_{x_1}\sigma_{a_1} + Z_2\sigma_{a_1}^2 & \sigma_{a_1}^2 \\ & & \sigma_{x_1}^2(Z_1+Z_2)^3/Z_1^3 & \rho\,\sigma_{x_1}\sigma_{a_1}(Z_1+Z_2)^2/Z_1^2 \\ & & & \sigma_{a_1}(Z_1+Z_2)/Z_1 \end{pmatrix} \tag{4.48}$$

to describe the multiple scattering. The actual implementation is done most easily by noting that the covariance between segments i and segments j can be written entirely in terms of $Z_{ij} = z_j - z_i$ and the covariance matrix elements for segment i as

$$\mathrm{Cov}(i,j) \;=\; \mathrm{Cov}(i,i) + \begin{pmatrix} Z_{ij}\mathrm{Cov}(x_i,a_i) & 0 \\ Z_{ij}\mathrm{Cov}(a_i,a_i) & 0 \end{pmatrix} \;=\; \mathrm{Cov}(i,i) \begin{pmatrix} 1 & 0 \\ Z_{ij} & 1 \end{pmatrix} \;, \tag{4.49}$$

for $|z_j - z_0| > |z_i - z_0|$.

To recapitulate, equations 4.41 and equation 4.42 provide a linear least squares estimator and its covariance matrix for the track parameters at z_0, based on measurements (x_1, a_1) and (x_2, a_2) at longitudinal positions z_1 and z_2. In these expressions, the covariance matrix V is the sum of the measurement errors given in 4.43 and the multiple scattering errors given in 4.48. The measurement errors are block-diagonal and their individual elements can be determined from 4.8. For the multiple scattering errors, the submatrices on the diagonal are given in 4.37 whereas the off-diagonal part is defined through the relation 4.49. All of this applies equally well to the three-dimensional case provided the correlations between the two directions are taken into account.

The expressions derived above for the case of two segments are easily generalized to a larger number of segments n. The longitudinal position at which the optimal track parameters are to be determined may lie upstream or downstream from all segments, and in rare cases somewhere in between. Out of the n segments, the j that lie upstream from z_0 and the k that lie downstream can be renumbered such that

$$z_{-j} - z_0 < z_{-j+1} - z_0 < \ldots < z_{-1} - z_0 < z_0 < z_1 - z_0 < \ldots < z_{k-1} - z_0 < z_k - z_0 \;, \tag{4.50}$$

with $j+k = n$. The scattering between z_0 and z_{-j} is independent from the scattering between z_0 and z_k. The $4n \times 4n$ covariance matrix can be considered as a 2×2 matrix with the $4j \times 4j$ and $4k \times 4k$ submatrices on the diagonal constructed as above and the $4j \times 4k$ and $4k \times 4j$ submatrices off the diagonal containing only zero entries.

The quality of the overall fit can be judged by calculating the χ^2 as in equation 4.40. It should be distributed as a χ^2 for $4n-4$ degrees of freedom. If one decides not to use the slope information of the segments, all expressions derived can still be used provided one removes in the covariance matrix and in the matrix describing the linear model all columns and rows related to them. The χ^2 then corresponds to $2n - 4$ degrees of freedom. By evaluating the contribution to the χ^2 from each segment separately, possible outliers may be detected as those that give anomalously large contributions. The fit probability is defined as one minus the cumulative distribution for a χ^2 with the appropriate number of degrees of freedom. The χ^2 cumulative distribution function is evaluated using the `gammp` algorithm from [127].

Figure 4.6 shows the distribution of the track fit χ^2 probability for an unbiased sample of tracks that are matched to the electronic detectors. Ideally, this distribution should be flat and as such it provides a powerful test of all hypotheses entering into the track model, in particular the description of the segment measurement errors and of the effect of multiple Coulomb scattering between measurements. In the case of muons, shown on the left, the only deviation from the expected distribution is a slight excess at low probabilities, due to the presence of background segments as well as to non-gaussian tails in the distribution of measurement errors. The latter are caused by a variety of effects, the most important of which are emulsion distortion and problems with the plate to plate alignment. It has

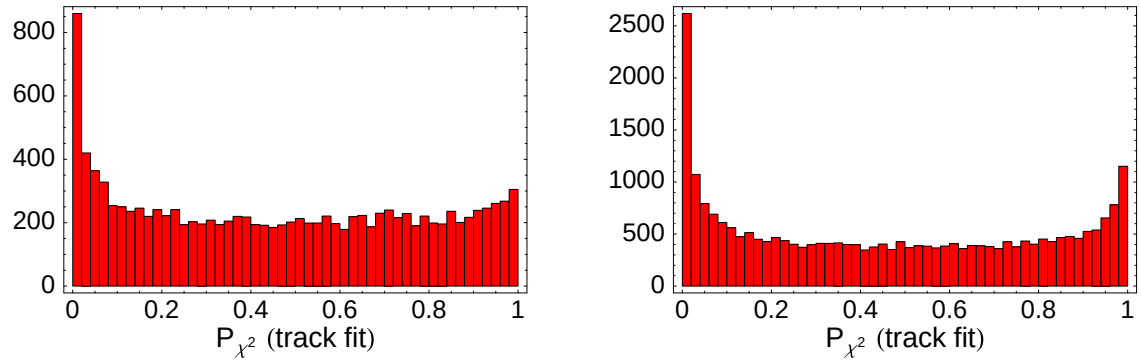


Fig. 4.6: The χ^2 probability of the track fit including the slope information for tracks that are matched to the electronic detectors. Left panel: tracks which are identified as muons in the electronic detectors. Right panel: tracks which are not identified as muons in the electronic detectors.

been checked that the size of the excess does not depend on the track slope, the number of segments, or the momentum measured in the electronic detectors.

In the case of hadrons, shown on the right, the excess at low probabilities is slightly more pronounced, and a further excess exists at high probabilities. At this point, we recall the assumption made in section 4.3 concerning the momentum. In the absence of any information, the momentum is fixed at a value of 1 GeV/c, arbitrary but for the fact that it is an acceptable threshold for secondaries of a neutrino interaction at energies typical in the CHORUS experiment. If a track can be matched to the electronic detectors, and if the momentum of the corresponding track could be measured in the electronic detectors, the measured value is used without taking into account any error. For muons, this procedure is entirely adequate. Muon momenta up to 5 GeV/c are measured by range, corresponding to a resolution of 6 %. For larger momenta, the effect of the measurement error on the multiple scattering is anyhow small. All of this contrasts sharply with the case of hadrons. For these, momenta can only be measured in the hexagonal spectrometer. This leads to a limited acceptance because of the magnet spokes and an altogether worse momentum resolution compared to that of the muon spectrometer.

Of the tracks shown in the right hand side of figure 4.6, 20 % do not have a momentum measurement. This includes tracks passing through the magnet spokes as well as tracks for which the hadron spectrometer reconstruction failed. For a further 6 %, the measured momentum is less than 1 MeV, an artefact of the momentum reconstruction algorithm. In general, one expects that hadrons for which no momentum is measured or for which the measured value is less than 1 GeV/c are by and large tracks with momenta higher than 1 GeV/c. With the assumption of a 1 GeV/c momentum, the multiple scattering of such tracks will be overestimated and this will lead in turn to an overestimation of the fit probability. Indeed, as shown in the left panel of figure 4.7, the excess at high probabilities is almost entirely due to those tracks where no momentum information is available or where it is expected to be of poor quality.

For measured momenta above 1 GeV/c, the track fit χ^2 probability for hadrons is similar to that for muons, as shown in the right panel of figure 4.7. The range between 2 GeV/c and 5 GeV/c corresponds to typical momenta for hadrons in neutrino interactions at CHORUS

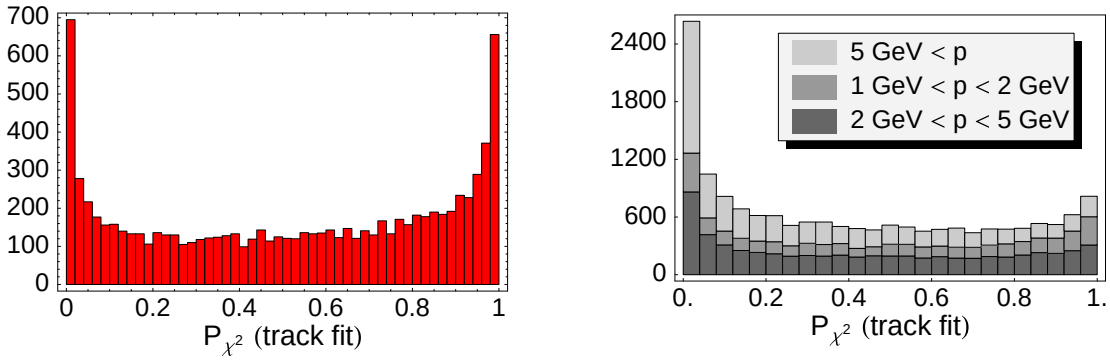


Fig. 4.7: The track fit χ^2 probability for tracks that are not identified as muons in the electronic detectors. Left panel: tracks for which the momentum could not be measured in the hexagonal spectrometer or for which the measured value lies below 1 GeV/c. In both cases, the track fit assumes a momentum value of 1 GeV/c. Right panel: tracks for which the momentum measured in the hexagonal spectrometer lies above 1 GeV/c.

energies. For lower measured values, the distribution tends to higher probabilities indicating an underestimate of the momentum. For higher measured values, the peak at low probabilities is most pronounced, corresponding to a possible overestimate of the momentum.

In summary, the model of section 4.2 for the segment measurement errors and the description in section 4.3 of the effect of multiple Coulomb scattering result in a track fit which adequately describes the data. The only effects not taken into account that do seem to play a role are non-gaussian measurement errors related to emulsion distortion or plate to plate alignment and the error on the momentum measurement for hadrons.

4.5 Track matching to the electronic detectors

The problem of identifying those tracks in the emulsion that actually belong to the neutrino event which gave rise to a trigger in the electronic detectors becomes a straightforward matter if the covariance matrix is available for the parameters of tracks both in the emulsion and in the electronic detectors. One merely needs to evaluate both sets of parameters at a common longitudinal position, build the difference between the two and compare this difference to the respective errors added in quadrature. It turns out the position information is essentially meaningless. The error on the transverse position of a track in the electronic detectors is at least 180 μm and in many cases considerably larger. The transverse area of the net scan fiducial volume is typically $\pm$ 750 μm^2, or barely $\pm 3\sigma$ when expressed in terms of the error for a detector track. The electronic detector tracks are selected to be compatible with the primary vertex as defined in the electronic detector reconstruction, corresponding to a window of one to two millimetres. For the emulsion tracks, only the slope information is left as a criterion to find candidates for each of the detector tracks. To make a meaningful comparison, two issues need to be addressed. First, the parameters, as well as their error, of the track in emulsion need to be evaluated at the downstream face of the emulsion stack, the

position at which track parameters and their errors are given by the electronic detector reconstruction. Second, the parameters of emulsion tracks and electronic detector tracks must obviously be given in the same reference frame before any meaningful comparison can be made.

The net scan fiducial volume consists of a transverse region on eight consecutive plates out of the thirty-six plates that make up a stack. To obtain the track parameters, and more specifically their error, in a reference plane at the downstream face of the stack, the multiple scattering between the net scan fiducial volume and the downstream face must be taken into account. As argued in section 4.4, all of the available information is optimally used only when the track fit itself is performed at the longitudinal position where the parameters will be used. However, this procedure becomes numerically unstable when the scattering between the reference plane and the most downstream measurement is large compared to the measurement errors. If all measurements lie to one side of the reference plane and if the contribution of the measurement errors to the covariance matrix becomes negligible compared to the contribution of the multiple scattering errors, then all measurements are fully correlated and the fit becomes degenerate. To avoid this problem, the distance between the first measurement and the reference plane is limited such that the scattering error never exceeds three times the measurement error. For low momentum tracks measured in the upstream part of the stack, this precludes a fit at the downstream face. In that case, track parameters and their errors are first determined at the most downstream position allowed and then extrapolated to the downstream face, taking into account the multiple scattering across this distance.

As for the reference frame, one should bear in mind that the `ecfsal` alignment algorithm performs only internal alignment of a net scan acquisition, essentially fixing the relative positions of subsequent plates in an acquisition as well as systematic distortion effects on the slope of segments in each plate. The algorithm relies on the large number of passing tracks that are not related to the event itself. It does not in any way refer to the electronic detector tracks, if only because there are too few in a single event to provide any useful constraint. In the absence of any external information, the internal alignment would be completely insensitive to a shift of all slopes by the same amount because the effect cancels against a transverse shift of each plate proportional to the distance from the first. Similarly, a scaling of all slopes by the same factor can be cancelled by changing the longitudinal distance between plates. To somewhat alleviate this problem, the `ecfsal` algorithm starts from the alignment parameters for an entire halfplate used in the vertex location and allows for relatively minor deviations to accomodate local distortions or variations in the shrinkage factor. However, the coordinate system used in the vertex location is similar, but not identical, to the reference frame for the electronic detector reconstruction. In fact, it treats the special sheet measurements as a reference. This implies that any systematic misalignment between the special sheet and the electronic detectors will translate into a misalignment between net scan data and electronic data. However, this also points towards a possible solution: for a given module, all events will be affected by the same misalignment and its parameters can consequently be determined from the average over all tracks in all events. This provides ample statistics to find the three alignment parameters: constant offsets Δa_x and Δa_y for the slope in each of the two projections, and a scaling factor s affecting both projections. In the reference frame for net scan data, the transverse slopes of a track in the electronic detectors

Module	Δa_x (mrad)	Δa_y (mrad)	s	σ_{a_x} (mrad)	σ_{a_y} (mrad)
210	5.2	-0.6	1.024	3.0	4.4
211	2.0	-2.2	1.023	2.4	1.4
220	4.7	-6.6	1.021	4.2	3.7
221	0.6	-1.9	1.022	6.2	4.4
250	4.3	-5.5	1.010	3.0	2.8
251	1.7	0.9	1.011	7.5	5.5
260	5.1	-2.6	1.007	2.7	2.6
261	-1.5	-0.6	1.012	5.0	5.0

Tab. 4.2: Numerical values for the constants in the description of misalignment between net scan reconstruction and electronic detector reconstruction, determined from the data in each of the halfmodules of stack 2.

become

$$\begin{pmatrix} a_x^{ele} \\ a_y^{ele} \end{pmatrix}' = \begin{pmatrix} s & 0 \\ 0 & s \end{pmatrix} \begin{pmatrix} a_x^{ele} \\ a_y^{ele} \end{pmatrix} + \begin{pmatrix} \Delta a_x \\ \Delta a_y \end{pmatrix} , \tag{4.51}$$

where (a_x^{ele}, a_y^{ele}) is the slope in the reference frame of the electronic detector reconstruction. This correction addresses the overall misalignment for a module but the sheer fact that the `ecfsal` algorithm is at all necessary means there are further effects which vary from one event to the next. The small number of event-related tracks precludes a correction of these additional effects and they can only be accounted for as a source of error. The size of this error may be different between the two projections. It is determined in each module by considering the distribution of the difference between emulsion slopes and detector slopes, accounting for the contribution of their respective fit errors. In total, there are five parameters per module describing the effect of misalignment between track slopes from net scan reconstruction and track slopes from the electronic detector reconstruction: constant offsets for each of the two projections, a scaling factor, and error contributions of remaining misalignment in each of the two projections. Table 4.2 shows numerical values for each of these in the eight halfmodules of stack 2.

We now have all the ingredients to formulate a criterion for a track in the emulsion to be matched to a track in the electronic detectors. The emulsion track parameters and their errors can be given at the downstream face of the emulsion stack, the longitudinal position where track parameters are given by the electronic detector reconstruction. The direction of a track in the electronic detectors can be transformed to the same reference frame as the net scan reconstruction, up to an error described by a 2×2 covariance matrix with $\sigma_{a_x}^2$ and σ_{a_y} on the diagonal and zero elsewhere. Apart from the covariance matrix for misalignment, there is a 2×2 covariance matrix V^{emu} for the slope (a_x^{emu}, a_y^{emu}) of the emulsion track and a 2×2 covariance matrix V^{ele} for the slope (a_x^{ele}, a_y^{ele}) of the detector track. These are obviously independent and their sum

$$V_{\delta a} = V^{emu} + V^{ele} + \begin{pmatrix} \sigma_{a_x}^2 & 0 \\ 0 & \sigma_{a_y}^2 \end{pmatrix} \tag{4.52}$$

is the covariance matrix of the angular difference

$$\delta a \;=\; \begin{pmatrix} \delta a_x \\ \delta a_y \end{pmatrix} \;=\; \begin{pmatrix} a_x^{emu} \\ a_y^{emu} \end{pmatrix} - \begin{pmatrix} a_x^{ele} \\ a_y^{ele} \end{pmatrix} \;. \tag{4.53}$$

For a sample of matched tracks, the quantity

$$\chi^2 \;=\; \delta a^T \, V_{\delta a}^{-1} \, \delta a \tag{4.54}$$

will be distributed as a χ^2 with two degrees of freedom. Or conversely, the slopes of an emulsion track and a detector track may be considered compatible if the corresponding χ^2 probability exceeds 0.1 %.

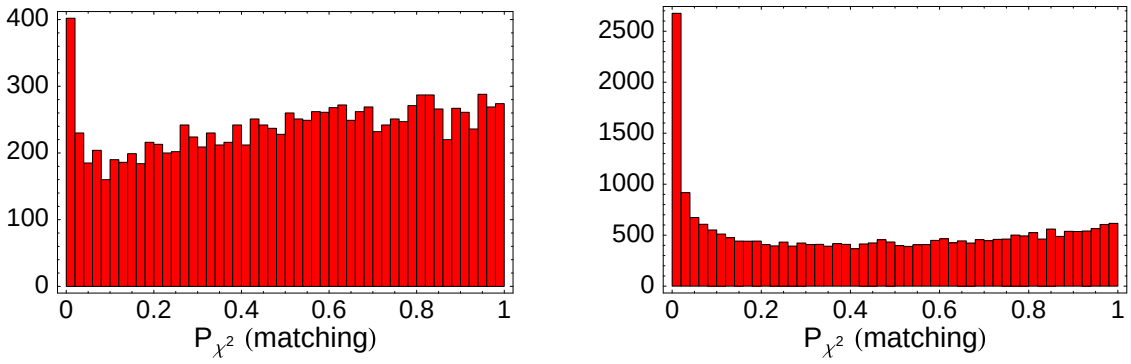


Fig. 4.8: The χ^2 probability corresponding to the difference between the track slope measured in emulsion and the track slope measured in the electronic detectors. Left panel: tracks which are identified as muons in the electronic detectors. Right panel: tracks which are not identified as muons in the electronic detectors.

Figure 4.8 shows the distribution of this probability for a sample of tracks that are considered matched. For tracks that are identified as muons, shown on the left, the observed distribution agrees well with the expectation. The mild excess at low probabilities is partly due to fake matches, but the effect on the purity is at a level below 2 %. For tracks not identified as muons, the excess is more pronounced, reflecting the larger number of tracks resulting in a greater degree of ambiguity as well as the lower quality of the momentum measurement for hadrons.

The reconstruction algorithms consider multiple matches for a single detector track: whichever emulsion track leads to a χ^2 matching probability larger than 0.1 % is treated as a candidate match. Among multiple candidates, one is selected not on the basis of the matching probability but on the basis of other criteria, described in section 4.8.4. The most important such criterion is to give preference to emulsion tracks that are attached to a vertex for which other attached tracks are matched to the electronic detectors. After this ranking, the only remaining background is due to tracks in the electronic detectors for which a genuine match does not exist in emulsion, either because it is not reconstructed in the emulsion or because the electronic track itself is fake. Such cases lead to random associations to emulsion tracks that do not belong to the event. To prevent such tracks from contributing to the background for a selection of decay topologies, the criteria on isolated secondary tracks are considerably tighter than for secondary vertices, This will be discussed in section 4.9.

To conclude this section on the matching between emulsion and detector tracks, figure 4.9 shows the actual value of the slope difference in one of the two projections, referring to the same samples as figure 4.8. For tracks identified as muons, the distribution is well described by a single gaussian with a width of 3.7 mrad. For hadrons, the single gaussian with a width of 6.4 mrad does not cover the tails of the distribution.

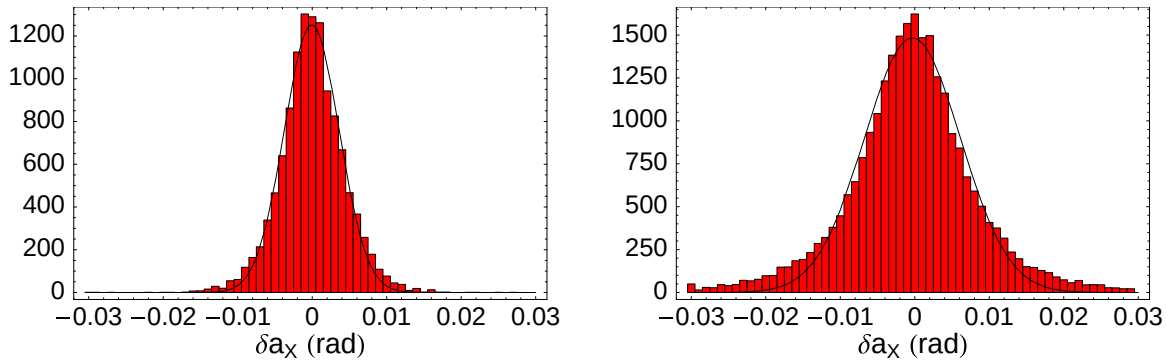


Fig. 4.9: Difference between the emulsion measurement and the detector measurement of the track slope $a_X = dX/dZ$ at the downstream face of the stack for emulsion tracks matched to the electronic detectors. Left panel: tracks which are identified as muons in the electronic detectors. Right panel: tracks which are not identified as muons in the electronic detectors.

4.6 Vertex fit

Fitting the common vertex for a set of tracks generally involves the minimization of a quantity

$$\chi^2 = \sum_i (\Delta\vec{q}_i)^T \, W_i \, \Delta\vec{q}_i \ , \tag{4.55}$$

where the sum runs over all tracks, $\Delta\vec{q}_i$ is the impact vector of the track with respect to the vertex and W_i are the covariance matrices for the impact vector. The covariance matrix for the impact vector must be calculated from the covariance matrix for the track parameters. To this end, one requires the matrix describing the transformation from track parameters to impact vector components. The partial derivatives of the transformation matrix with respect to the coordinates of the vertex build the Jacobian matrix entering the error propagation from track parameters to impact vector. This problem is not linear and its solution generally relies on iterative techniques [120, 128]. The situation is further complicated when effects of multiple scattering are introduced, usually done through suitable approximations [122]. Instead of non-iterative techniques, linear filtering techniques based on a Kalman filter have been developed for the vertex fitting problem [129].

Given the spatial resolution and track density in net scan data, the assignment of tracks to a vertex is relatively straightforward. Furthermore, the distance along a track between the vertex and the first measurement is at most the thickness of a plate or slightly less than

800 μm. This implies that the point of closest approach between any pair of tracks is relatively accurate. In the absence of magnetic field and given the point of closest approach, the expression 4.55 can be recast in linear form, as shown in [130], based on the perigee parametrization of [131].

As a first step, the point of closest approach is determined using the nominal track parameters, *i.e.* the parameters without their error. The track parameters are then reevaluated at the corresponding longitudinal position to fully account for multiple scattering between the supposed vertex and the first measurement. This reevaluation may be done through an explicit refit of the track parameters at the appropriate longitudinal position, as shown in section 4.4. Alternatively, the track parameters and their covariance matrix can simply be extrapolated if the fit has been performed already at a position that lies upstream from all segments on the track. If $\vec{x}_0 = (x, y, a_x, a_y)_0$ are the track parameters at $z = z_0$ and V_0 the corresponding covariance matrix, then the parameters at $z = z_1$ are $\vec{x}_1 = \vec{x}_0 + (z_1 - z_0)(a_x, a_y, 0, 0)$ and have covariance matrix

$$V_1 \;=\; \begin{pmatrix} 1 & 0 & z_1 - z_0 & 0 \\ 0 & 1 & 0 & z_1 - z_0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} V_0 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ z_1 - z_0 & 0 & 1 & 0 \\ 0 & z_1 - z_0 & 0 & 1 \end{pmatrix} \;+\; V_{mcs}(z_1 - z_0) \;\;, \tag{4.56}$$

where $V_{mcs}(z_1 - z_0)$ is the covariance matrix for multiple scattering across a longitudinal distance $z_1 - z_0$. However, refitting the track parameters altogether is preferable as it uses all the available information, contrary to simple extrapolation which leads to larger errors in the transverse position. In either case, it is clear that the uncertainty on the slope translates into an uncertainty on the transverse coordinate evaluated at the longitudinal position of the point of closest approach. Because of the fair agreement between the point of closest approach determined from the nominal track parameters and the point of closest approach corresponding to the vertex fit which accounts for the errors, we argue that the uncertainty on the slope may be neglected in the vertex fit itself. In fact, it turns out that the longitudinal positions before and after the vertex fit agree to within 10 μm, corresponding to less than 0.01 μm transversely for a 1 mrad error on the slope.

So far, we have characterized tracks through their impact position (x, y) and slopes (a_x, a_y) in a reference plane at longitudinal position z_0. The perigee parametrization consists in choosing a coordinate system where one of the axes is parallel to the track. As the tracks in the net scan data show relatively modest slopes with respect to the Z axis, it seems natural to transform to a coordinate system where the $Z_\parallel$ axis is parallel to the track. As for the other two axes, $Y_\perp$ lies in the XY plane and is normal to the trajectory, $X_\perp$ is defined such that $(X_\perp, Y_\perp, Z_\parallel)$ forms a right handed coordinate system. This transformation can be achieved through two rotations. First, the (X, Y) axes are rotated through ϕ, the azimuth angle of the track direction, around the Z axis to yield $(X', Y_\perp)$. Then, the (X', Z) axes are rotated by θ, the zenith angle of the track direction, around the $Y_\perp$ axis to yield $(X_\perp, Z_\parallel)$. These rotations are shown in figure 4.10. Formally, the transformation from $\vec{r} = (x, y, z)$ to $\vec{r}\,' = (x_\perp, y_\perp, z_\parallel)$ is defined as

$$\vec{r}\,' \;=\; R(\phi, \theta)\, \vec{r} \;=\; R_Y(\theta)\, R_Z(\phi)\, \vec{r} \;\;, \tag{4.57}$$

where $R_Z(\phi)$ is the rotation by ϕ around the Z axis and $R_Y(\theta)$ is the rotation by θ around the

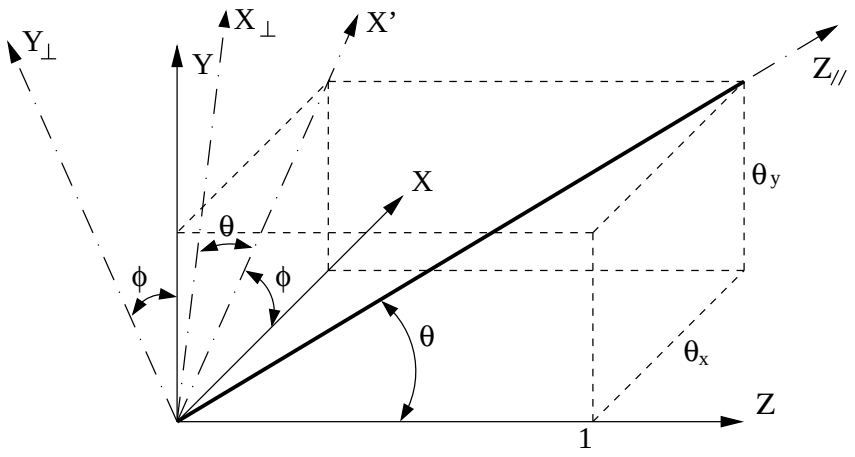


Fig. 4.10: Transformation between the global coordinate system (X, Y, Z) and a coordinate system $(X_\perp, Y_\perp, Z_\parallel)$ for which the $Z_\parallel$ axis is parallel to a track with direction vector $(a_x, a_y, 1)$.

$Y_\perp$ axis. The transformation matrix $R(\phi, \theta)$ can be explicitly written out as

$$R(\phi,\theta) = \begin{pmatrix} \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0 \\ \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \end{pmatrix} \quad . \tag{4.58}$$

In terms of the track direction vector $(a_x, a_y, 1)$, the azimuth angle ϕ and the zenith angle θ are given by

$$\phi = \arctan\left(\frac{a_y}{a_x}\right) \quad , \quad \theta = \arctan\left(\sqrt{a_x^2 + a_y^2}\right) \quad . \tag{4.59}$$

After transformation to the $(X_\perp, Y_\perp, Z_\parallel)$ coordinate system, the third component of the impact vector $\Delta\vec{q}_i$, introduced in equation 4.55, is zero by construction. This allows us to reformulate the original problem in terms of $\vec{q}_\perp = (x_\perp, y_\perp)$, the two-dimensional vector normal to the track, which is calculated as

$$\vec{q}_\perp \;=\; Q(\phi,\theta)\;\Delta\vec{q} \;, \tag{4.60}$$

where Q is a 2×3 matrix formed by the first and second row of the matrix $R(\phi, \theta)$. So far, we have merely performed a coordinate transformation, without a single approximation. However, to recast the original problem in terms of $\vec{q}_\perp$ requires the partial derivatives of the transformation $R(\phi, \theta)$ with respect to the components of the vertex position $\vec{r}_v = (x_v, y_v, z_v)$. The result is again a non-linear expression. The problem can be greatly simplified by assuming the matrix $Q(\phi, \theta)$ to be constant, discarding the errors on the track slope. Provided the reference point on the track is the point of closest approach to the vertex, this approximation is entirely justified. Under this assumption, the χ^2 function simplifies to

$$\chi^2 \;=\; \sum_i \left(\vec{r}_v - \vec{r}_i\right)^T \, Q_i^T V_i^{-1} Q_i \, \left(\vec{r}_v - \vec{r}_i\right) \;, \tag{4.61}$$

where $\vec{r}_v$ is the vertex position, $\vec{r}_i$ is the point of closest approach on track i, Q_i is the partial transformation matrix introduced in equation 4.60 and V_i is the covariance matrix for the track impact point at the longitudinal position z_i:

$$V_i = \begin{pmatrix} Cov(x,x) & Cov(x,y) \\ Cov(x,y) & Cov(y,y) \end{pmatrix} . \tag{4.62}$$

It is clear that the χ^2 function implicitly discards the minimization of $z_v - z_i$, assuming this difference to be zero by construction. Again, this is justified to the extent that the longitudinal position z_i where the above construction is performed is a sufficiently good approximation to the point of closest approach. Insofar as Q_i and V_i are independent of $\vec{r}_v$, the χ^2 function is a linear function of the parameters (x_v, y_v, z_v). With the definition $W_i = Q_i^T V_i^{-1} Q_i$, the solution for the vertex coordinates $\vec{r}_v$ becomes

$$\vec{r}_v = \left(\sum_i W_i \right)^{-1} \sum_i W_i \vec{r}_i \tag{4.63}$$

and its covariance matrix is given by

$$V = \left(\sum_i W_i \right)^{-1} . \tag{4.64}$$

For a set of n tracks, the χ^2 obtained should be distributed as a χ^2 with $2n - 3$ degrees of freedom. To estimate whether a particular track belongs to a vertex, its contribution to the overall χ^2 of the fit can be isolated.

As a criterion for a set of tracks to form a vertex, we require that the χ^2 of a vertex fit exceeds 0.1 %. In a first step, a relatively large acceptance is used to find sets of tracks that may originate from a common vertex. Then, the tracks with the largest contribution to the vertex fit χ^2 are successively removed until the fit probability exceeds 0.1 %. The left hand side of figure 4.11 shows the distribution of vertex fit χ^2 probabilities for vertices in which at least one of the attached tracks is matched to the electronic detectors. The right hand side refers to the same sample, showing only the region of probabilities below 10 %. The selection of secondary vertices, described in the following sections, is meant to achieve high purity and may under certain conditions assign tracks to the primary vertex even though the vertex fit probability would seem to indicate they are incompatible. This is the origin of the tail in the probability distribution below 0.1 %. Part of this tail reflects an inefficiency in the selection of secondaries, the remainder is due to wrong track parameters or an underestimate of their errors.

Apart from the tail at low vertex fit probabilities, the left hand side of figure 4.11 shows a further excess at high probabilities. Since the vertex fit is based on propagation of the fit errors for the parameters of each of the attached tracks, it is reasonable to expect that an overestimate of the track fit probability will lead to an overestimate of the vertex fit probability. To test this hypothesis, one defines the overall track fit probability for all tracks attached to the vertex. As the different track fits are all independent, the sum of the different χ^2 values should be distributed as a χ^2 itself with the number of degrees of freedom equal to the sum over the different tracks of their respective number of degrees of freedom. The corresponding probability is shown on the horizontal axis of figure 4.12, where the vertical axis

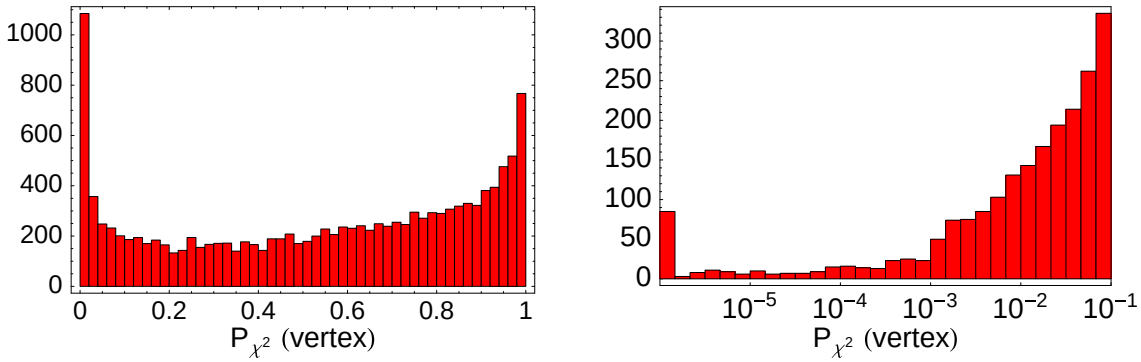


Fig. 4.11: Left panel: the vertex fit χ^2 probability defined in the text for vertices where at least one of the tracks is matched to the electronic detectors. Right panel: the same quantity for the same sample of vertices, but showing only the vertices with a probability less than 10 %.

refers to the vertex fit probability itself. Indeed, high track fit probabilities are correlated with high vertex fit probabilities, explaining the entire excess at high vertex fit probabilities. As described in section 4.4, overestimated track fit probabilities are largely due to the assumptions about the momentum of hadrons. Figure 4.12 also indicates, at least qualitatively, that part of the tail at low vertex fit probability is due to an underestimate of the track fit probability.

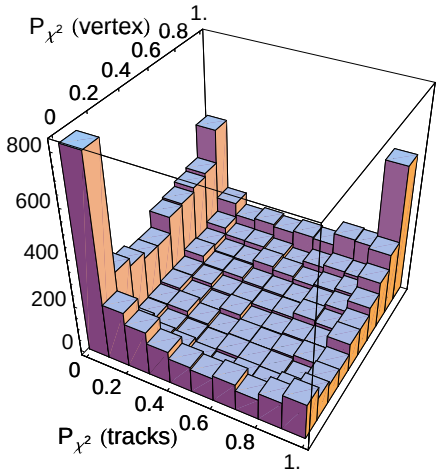


Fig. 4.12: The dependence of the vertex fit quality on the quality of the input track parameters. The horizontal axis shows the combined probability for the fits of all tracks included in the vertex, the vertical axis shows the vertex fit probability.

Figure 4.13 delivers striking testimony to the exceptional spatial resolution in emulsion. It shows the error on the vertex position in each of the three projections, defined as the square root of the diagonal elements of the 3×3 covariance matrix. The left hand side refers to the errors σ_x and σ_y on the transverse position, the right hand side to the error σ_z on the longitu-

dinal position. Because of the angular acceptance of 400 mrad and because most of the tracks anyhow have slopes well below 200 mrad, the error on the longitudinal vertex position is about an order of magnitude larger than the error on the transverse vertex position.

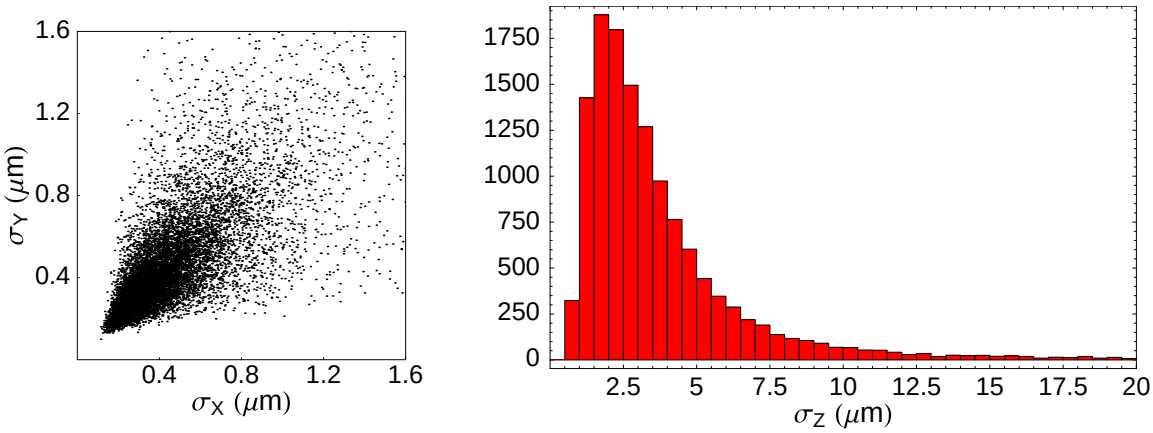


Fig. 4.13: Left panel: the error on the transverse vertex position in the two directions X and Y, defined as the square root of the first and second elements on the diagonal of the 3×3 covariance matrix from the fit. Right panel: the error on the longitudinal vertex position, defined as the square root of the third element on the diagonal of the 3×3 covariance matrix from the fit.

Finally, the distribution of the vertex position itself is shown in figure 4.14. The left hand side indicates at which longitudinal position within the plate the vertices are situated. The vertex depth is defined as the longitudinal distance between the vertex position and the position at which segment measurements are given in the plate immediately downstream from the vertex plate. The most striking feature in the distribution of vertex depth is the dip at around 400 μm, extending across about 80 μm. This corresponds precisely to the position and thickness of the plastic support. The plastic support density is 1.03 g/cm^3 compared to 3.82 g/cm^3 for the emulsion itself, leading to a lower number of neutrino interactions per unit of volume in the plastic base. On both sides of the plastic support, there is an emulsion layer with a thickness of about 350 μm and one might naively expect a symmetric distribution for the vertex position on both sides of the dip. Instead, the downstream part is some 100 μm wider than the upstream part. This is due to the definition of the vertex plate: because the scanning is performed slightly below the upstream surface of each plate, vertices at this surface are considered to lie in the next plate. On top of that, the scanning itself covers a certain depth and tracks may still be found even if they do not reach the upstream edge of the scanning volume. Vertices at the upstream part of the scanning volume itself account for the occurrence of negative vertex depths, up to about -50 μm.

The right hand side of figure 4.14 shows the plate number for the vertex plate; the 36 plates are numbered from downstream to upstream. The general downward trend can be attributed to the combination of two effects. First, the electronic detector reconstruction is less efficient for neutrino interactions upstream in the emulsion stack because there is a larger risk that a shower will develop before the first tracker plane is reached. Second, the vertex location itself is less efficient because tracks originating in the upstream plates have to be followed across a larger number of plates. Any further variation in the distribution

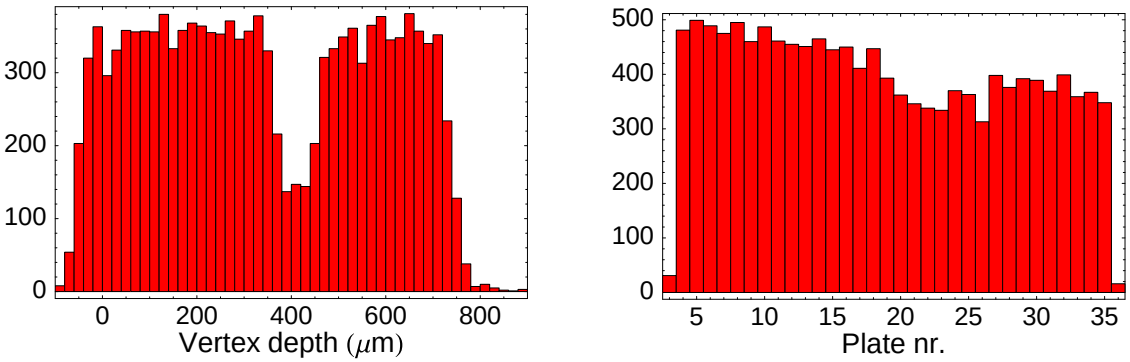


Fig. 4.14: Left panel: the longitudinal distance between the vertex position and the position of the most upstream segment on any of the tracks attached to the vertex. Right panel: the number of vertices in each plate, where the plate numbering runs from the most downstream to the most upstream.

reflects emulsion conditions, or simply the fact that net scan data taking is not yet finished for all plates at this point in time. Vertices in the three most downstream plates cannot be reconstructed on the basis of the net scan data only because the number of segments that can be measured on each track is at most two. Figure 4.14 includes vertices in the most upstream plate, but these are not considered in the search for secondary topologies because the absence of an additional plate to veto passing tracks leads to an unacceptable level of background.

4.7 Distance of closest approach

One obvious prerequisite for a selection of short-lived particles is a means to decide whether or not a given track is compatible with the primary vertex, or whether two reconstructed vertices are equal within errors and both correspond to the primary vertex. In general, there are three cases to be considered: two vertices, a vertex and a track, and two tracks. The first category is the easiest. To determine whether two reconstructed vertices may in fact be one and the same, it is sufficient to consider their distance in space, build the corresponding 3×3 covariance matrix as the sum of the covariance matrices for the respective vertices and calculate the χ^2 probability for the distance to be compatible with zero. The second category will be discussed in this section, the third one in the next.

Alternatively, instead of treating these particular cases, one could simply try to build a vertex using all tracks involved: those attached to each of the two vertices, those attached to the vertex plus the additional one, or the two tracks. However, the advantage of the approach described here is that it leads to an explicit value for the distance, and not just the probability for this distance to be compatible with zero. On top of that, at least for the case of a track and a vertex, the two approaches are not entirely equivalent: if the probability for a given vertex is particularly high, it may still be acceptable once another track is wrongly added to it.

4.7.1 Impact parameter between a track and a vertex

Given a vertex position and the corresponding 3×3 covariance matrix as well as track parameters and the corresponding 4×4 covariance matrix, we need to find the probability for the track to originate from the vertex. One solution would consist in simply performing a new vertex fit including the additional track and judge on the basis of the vertex fit probability. Nevertheless, it is often useful to address the question without resorting to a refit of the vertex; for instance because it leads to an expression for the numerical value of the impact parameter, defined as the distance of closest approach between the track and the vertex. For track parameters (x_0, y_0, a_x, a_y) given at a longitudinal position z_0, $\vec{t} = (x_0, y_0, z_0)$ is a position on the track and $\vec{a} = (a_x, a_y, 1)$ is a vector along the track direction. Using $\vec{t}$ and $\vec{a}$ together with the vertex position $\vec{v} = (v_x, v_y, v_z)$, the vector from the vertex to the point of closest approach on the track is

$$\vec{d} \;=\; \vec{v} - \vec{t} \;-\; \frac{\left(\vec{v} - \vec{t}\right).\vec{a}}{|\vec{a}|^2}\,\vec{a} \;. \tag{4.65}$$

The distance of closest approach d is nothing but the length of $\vec{d}$. To find the probability for this distance to be compatible with zero, one might derive the covariance matrix for $\vec{d}$ by propagating the covariance matrices of $\vec{v}$, $\vec{t}$, and $\vec{a}$. However, it is immediately clear that this will lead to a singular matrix as there are only two, not three, degrees of freedom: by construction, $\vec{d}$ is perpendicular to $\vec{a}$ which fixes one component. Instead of $\vec{d}$ we need the two-dimensional projection of $\vec{d}$ in the plane perpendicular to the track direction. This is precisely the transformation of equation 4.57, used for the vertex fit in section 4.6. Again, the errors related to the track direction can be neglected provided that the track fit is performed at the longitudinal position of $\vec{v}+\vec{d}$, the point of closest approach. Referring to equation 4.65, in this case the term proportional to $\vec{a}$ becomes zero.

The projection of $\vec{d}$ on the plane perpendicular to the track direction defines the vector $\vec{d}_\perp$ with components $\vec{d}_{x\perp}$ and $\vec{d}_{y\perp}$. The 2×2 covariance matrix of $\vec{d}_\perp$ is

$$V_{\vec{d}_\perp} \;=\; Q\, V_{\vec{d}} Q^T \;, \tag{4.66}$$

where Q is defined as in equation 4.60. Neglecting the slope errors, the 3×3 covariance matrix $V_{\vec{d}}$ for the distance of closest approach in space is nothing but the sum of the covariance matrix for the vertex position and the covariance matrix for the X and Y position of the track at the point of closest approach, extended with a third row and column of zeroes to reflect the fact that the longitudinal position is treated as a parameter, not a variable. For tracks originating from the vertex, the quantity

$$\chi^2 = \vec{d}_\perp^{\,T}\; V_{\vec{d}_\perp}^{-1}\; \vec{d}_\perp \tag{4.67}$$

will be distributed as a χ^2 with two degrees of freedom. The corresponding χ^2 probability can be used as a criterion to decide whether a track is compatible with a given vertex. One obvious test of the above is to consider reconstructed vertices, refit them after having removed one of their tracks, and then calculate the probability for the removed track to be compatible with the vertex position that has been determined without it. This results in the probability distribution shown in figure 4.15. The selection based on this quantity will be

described below. As expected, the distribution closely resembles that of the vertex fit probability, shown in figure 4.11. Figure 4.16 explicitly shows the relationship between the two quantities.

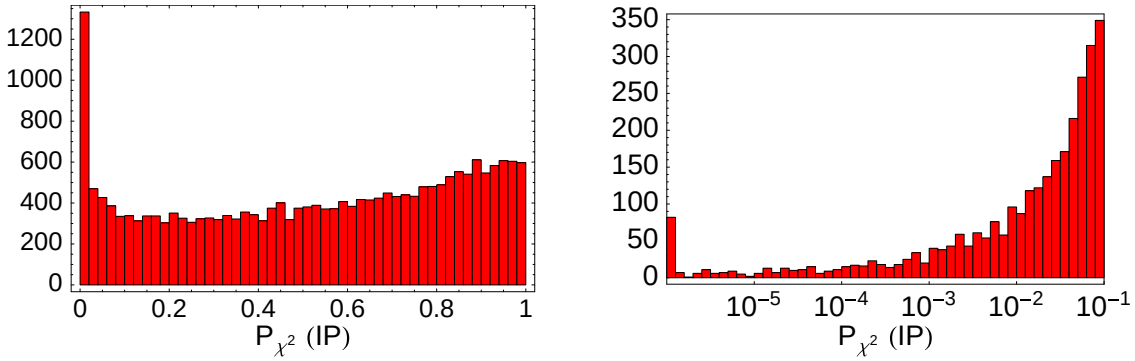


Fig. 4.15: The probability for a track to be compatible with a vertex as defined in the text, evaluated for events containing a vertex with three or more tracks that are matched to the electronic detectors. One by one, each of the tracks is taken out from the vertex fit and the probability is evaluated for the distance to the vertex to be compatible with zero. On the right hand side, the first bin indicates the number of entries with a value outside the range shown.

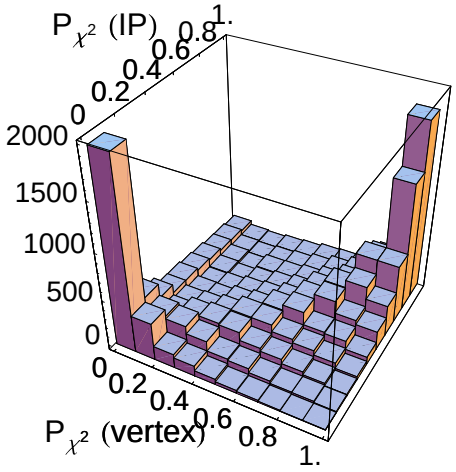


Fig. 4.16: The horizontal axis shows the vertex fit probability when the track under consideration is included in the fit. The vertical axis shows the probability that the distance of closest approach between the track under consideration and the vertex is compatible with zero, excluding the track from the vertex fit.

Instead of the probability for a non-zero impact parameter, one may also consider the numerical value of the impact parameter as a selection variable. This is especially meaningful if the results of an automatic selection of secondary vertices is to be confronted to the results of a manual selection. Small values for the impact parameter correspond to small changes in

the track direction which may well go undetected manually. The left hand side of figure 4.17 shows the length of $\vec{d}_\perp$ for the same sample as figure 4.15. The right hand side shows the two components of $\vec{d}_\perp$. Note that the asymmetry is a mere artefact related to the choice of axes in the plane perpendicular to the track direction.

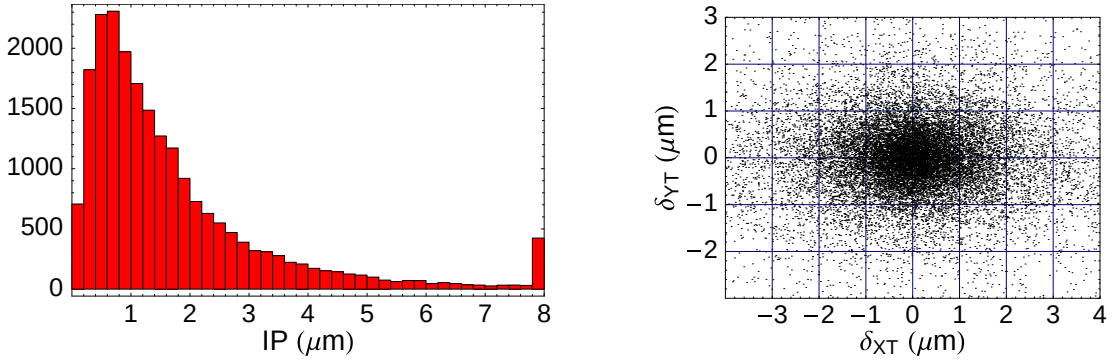


Fig. 4.17: Left panel: distance of closest approach between a track and a vertex, for tracks that are matched to the electronic detectors and which were attached to the vertex in the standard reconstruction. Right panel: distance of closest approach between a track and a vertex, broken down into the transverse components defined in the text.

4.7.2 Impact parameter between two tracks

To find the probability for two tracks to intersect, the reasoning follows very much the same lines as in the case of the probability for a track to be compatible with a vertex. Again, we start from the distance of closest approach between the two tracks. For a track with parameters $(x_0^1, y_0^1, a_x^1, a_y^1)$ at z^1 and a track with parameters $(x_0^2, y_0^2, a_x^2, a_y^2)$ at z^2, the vector between their respective points of closest approach is

$$\vec{d} \;=\; \frac{\left(\vec{a}^1 \times \vec{a}^2\right) \;.\; \left(\vec{t}^1 - \vec{t}^2\right)}{\left|\vec{a}^1 \times \vec{a}^2\right|^2}\; \vec{a}^1 \times \vec{a}^2 \;\;, \tag{4.68}$$

where $\vec{t}^i = (x_0^i, y_0^i, z_0^i)$ is defined as a position on the track and $\vec{a}^i = (a_x^i, a_y^i, 1)$ as the track direction vector. Whereas the distance of closest approach between a track and a vertex has two degrees of freedom, in this case we have only a single one. By construction, the vector $\vec{d}$ is perpendicular to both $\vec{a}^1$ and $\vec{a}^2$ and this completely fixes the direction of $\vec{d}$, leaving only the length as a degree of freedom. It is convenient to define a unit vector in this direction as

$$\vec{n} \;=\; \frac{\left(\vec{a}^1 \times \vec{a}^2\right)}{\left|\vec{a}^1 \times \vec{a}^2\right|} \;\;. \tag{4.69}$$

This allows us to rewrite $\vec{d}$ as $d\;\vec{n}$ where d is nothing but the length of d or the distance of closest approach. Conversely, the only degree of freedom in this problem can be written as

$$d \;=\; \vec{d}.\vec{n} \;\;. \tag{4.70}$$

Noting that the dot product is a projection operator, we see that the 1×1 covariance matrix V for d is equal to $\vec{n}^T\, V_{\vec{d}}\, \vec{n}$. In fact, this is a special case of the expression

$$V_{\vec{a}.\vec{b}} \;=\; \vec{a}^T V_{\vec{b}} \vec{a} \;+\; \vec{b}^T V_{\vec{a}} \vec{b} \tag{4.71}$$

for the covariance matrix of the dot product. Above, we have implicitly assumed the covariance matrix for $\vec{n}$ to be zero. Since $\vec{n}$ is defined entirely in terms of the track directions, its covariance matrix can be neglected provided that the fit of both tracks is performed at their respective points of closest approach. Under this condition, equation 4.68 reduces to $\vec{d} = \vec{t}_1 - \vec{t}_2$ and the corresponding covariance matrix is simply the sum of the 2×2 covariance matrices for (x_0^1, y_0^1) and (x_0^2, y_0^2) extended with a third row and column of zeroes to reflect the fact that z is treated as a parameter, not a variable. Instead of this extension, the third component of $\vec{n}$ can be dropped to define the two-dimensional vector $\vec{n}_{xy} = (n_x, n_y)$. Combining all of the above, for pairs of intersecting tracks, the quantity

$$\chi^2 \;=\; \frac{d^2}{\vec{n}_{xy}^T\, (V_1 + V_2)\, \vec{n}_{xy}} \tag{4.72}$$

is distributed as a χ^2 with one degree of freedom, where V_i are the covariance matrices for (x_0^i, y_0^i) under the condition that the z_0^i correspond to the points of closest approach. The corresponding χ^2 probability provides a criterion to decide whether two tracks intersect. The test that can be performed here is similar to the one for the distance of closest approach between a track and a vertex, but now considering reconstructed vertices with two tracks instead of vertices with at least three tracks. One obtains the probability distribution shown in figure 4.18.

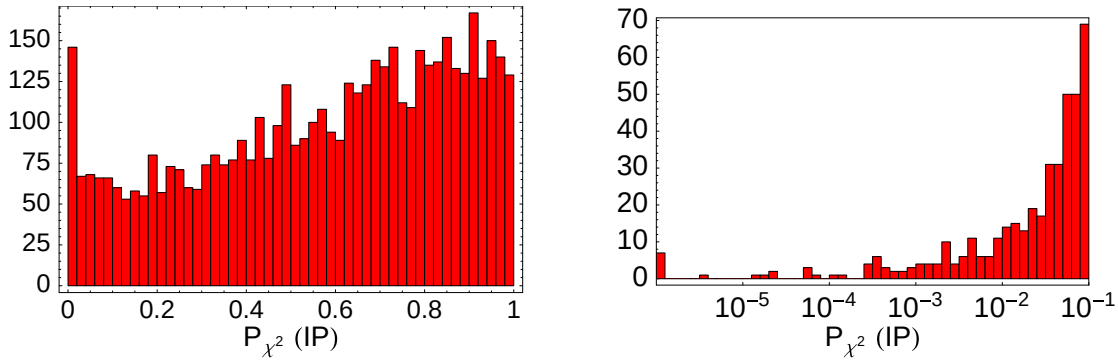


Fig. 4.18: The probability for two tracks to intersect as defined in the text, evaluated for all events for which the vertex contains two tracks that are matched to the electronic detectors. On the right hand side, the first bin indicates the number of entries with a value outside the range shown.

And again as in the case for the impact parameter of a track to a vertex, we can also consider the distance itself between two tracks. The left hand side of figure 4.19 shows the distribution of this distance for the sample of track pairs of figure 4.18. The right hand side shows the error on the distance of closest approach, defined as

$$\sigma_d \;=\; \sqrt{\vec{n}_{xy}^T\, (V_1 + V_2)\, \vec{n}_{xy}} \;\; . \tag{4.73}$$

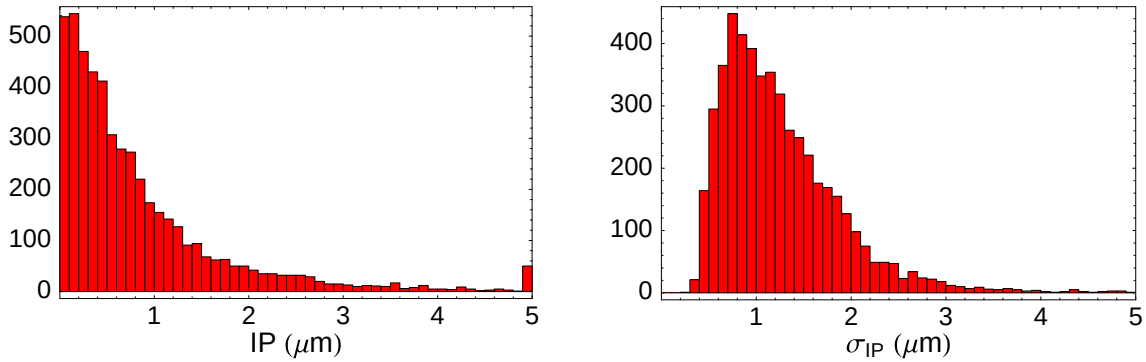


Fig. 4.19: Left panel: distribution of the distance of closest approach between the two tracks matched to the electronic detectors for all events for which the vertex contains two matched tracks. The last bin indicates the number of entries with a value outside the range shown. Right panel: the error on the distance of closest approaches based on the propagation of the track parameter errors, as described in the text.

4.8 Event reconstruction

We now have all the necessary ingredients to tackle the actual event reconstruction. The basic building blocks are the model for segment measurement errors of section 4.2 and the description for multiple scattering of section 4.3. A sequence of segments can be combined to form a track, for which the track parameters and their errors are obtained, as well as fit probability to judge the quality of the track and reject outlier segments if appropriate. Likewise, a collection of tracks may originate from a vertex, for which the position is found together with its covariance matrix, and again a fit probability. After alignment of the two coordinate systems, the direction of tracks in emulsion can be directly compared with that of tracks from the electronic detector reconstruction to select the tracks that belong to the event, once more using a χ^2 probability. Finally, we have seen how the probability can be formulated for a track to be consistent with a given vertex or to intersect a different track, based on the distance of closest approach. Most of the above is entirely general and in no way related to the CHORUS setup as far as the formal expressions are concerned. However, the above does not necessarily lead to a unique reconstruction for the event. For instance, two emulsion tracks may both have an acceptable probability to be matched to the same track in the electronic detectors or a single track may have an acceptable probability to be attached to two different vertices. To resolve such ambiguities, we need to introduce additional information based on what we know a typical event should look like. Inevitably, this type of information will be less general than the analytical expressions for the track fit probability or such like. Another problem that has not been addressed so far is the choice of numerical values. For instance, a procedure rejecting vertices with a fit probability below 1 % will lead to different results from a procedure which accepts vertices with a fit probability down to 0.1 %. Here, the choice will depend on our preference: any cut value will lead to a different trade-off between efficiency and purity. The purpose of this section is precisely to address those points. Essentially, we will go through the sequence of steps in the reconstruction algorithm and indicate the numerical value of cuts if any, the ranking used to resolve ambiguities, etcetera.

4.8.1 Track fitting and matching

The starting point for the event reconstruction is the output of the `ecvtxa` program: tracks composed of segments and attached to vertices. However, as pointed out before, the track finding in `ecvtxa` relies on a relatively crude track fit and may well have accepted or rejected segments even though their residual would lie respectively outside or inside of the acceptance once a more detailed model for the measurement errors and multiple scattering is used. The more time-consuming fit performed here is possible only because it is limited to stopping tracks. Stopping tracks are defined as those which, according to the `ecvtxa` reconstruction, do not exit from the fiducial volume on the upstream side. First, a refit is performed for each of the stopping tracks, removing outlier segments until the track fit probability exceeds 0.1 %. Then, the segment raw data are accessed in an attempt to find additional segments on plates for which no segment has yet been assigned to the track, both between the most upstream and downstream segment of the track and extending the track in the upstream and downstream directions for up to two plates or until the end of the fiducial volume is reached. If, after this step, any track is found to exit from the fiducial volume on the upstream side, it is rejected. Obviously, tracks for which only a single segment remains are removed as well. After the refit, an attempt is made to match the stopping tracks to tracks in the electronic detectors, requiring the matching probability to exceed 0.1 %. At this stage, multiple candidates in the emulsion for the same detector track or multiple candidates in the electronic detectors for the same emulsion track are retained. To profit from the momentum information for matched tracks, all of these are again refit, once more removing outlier segments and searching for additional segments in the raw data. The access to the raw data requires a fast lookup algorithm in the 4-dimensional space of transverse position and slope. A multi-dimensional binary tree [112] is particularly suitable for this purpose.

4.8.2 Vertex fitting

Vertex fitting again starts from the vertices found by the `ecvtxa` algorithm. Each vertex is refit, based on the track parameters and their error from the track refit, to determine the position and covariance matrix. Outlier tracks are removed until the vertex fit probability exceeds 0.1 %. If this condition is not yet fulfilled even when only two tracks are left, the vertex is removed altogether. No attempt is made to attach additional tracks to the vertex. But, as will be shown below, further tracks may at a later stage be assigned to an existing vertex. Ambiguities in the sense of a single track originating from multiple vertices are allowed. It is worth pointing out that the `ecvtxa` algorithm allows only for tracks to originate from a vertex, not to lead to a vertex, *i.e.* the most upstream segment of a track attached to a vertex is required to lie downstream from the vertex position, up to a distance related to the depth covered in the scanning of a given plate. Two track vertices with an opening angle less than 30 mrad are removed; typically, these are e^+e^- pairs which may or may not belong to the event but which in any case are of no interest to the selection of short-lived particles. Vertices that lie upstream of the fiducial volume are removed as well. These may arise either from chance crossings between stopping tracks or involve tracks that are not stopping but for which the most upstream segment was not found. Either way, they again are of no interest in the following.

The requirement for tracks to stop in the fiducial volume removes one type of tracks that do not belong to the event itself. Another source are short tracks: low energy tracks,

random associations between segments of different passing tracks, or parts of tracks from a different year of data taking if the alignment for a pair of plates is sufficiently similar for the two years. Systematically rejecting any short track, for instance any track with less than three segments, would lead to an unacceptable loss in efficiency. On the other hand, tracks attached to a vertex are clearly more likely to be genuine than tracks which are not. Only for isolated, short tracks do we need a further rejection. First of all, for tracks of only two segments, the track fit probability is required to exceed 1 % rather than 0.1 %. Secondly, an isolated track can only be used as the match for an electronic detector track if its efficiency exceeds 70 %. In this context, the efficiency is defined as the ratio of the number of segments to the number of plates for which a segment could be found.

4.8.3 The scanback track and vertex

A very simple observation which has not been used so far is the fact that net scan data are not taken at some arbitrary position. In the transverse direction, it is centered around the position of the most upstream measurement of the track which was used in the vertex location. Longitudinally, the most upstream plate for net scan data taking lies two plates beyond the plate where the so-called scanback track was last measured. Net scan data taking itself is independent of the vertex location, but the scanback track should obviously still be there. Because of the way the data are organized, it is cumbersome to access the vertex location data. Instead, we try to infer from the net scan data which track may have been the scanback track. This also leaves the possibility to profit from the most recent version of the electronic detector reconstruction rather than having to restrict oneself to the version which provided the initial prediction for vertex location.

Any matched track for which the parameters point to within 250 μm from the center of the vertex plate is a candidate scanback track. If no such track exists, the entire event is rejected. This typically happens whenever a background track was picked up during the vertex location, for instance in the plate where the scanback track reached the primary vertex. Alternatively, there may have been multiple candidates on a given plate, leading to multiple net scan acquisitions for the same event. The procedure followed in these cases will be mentioned below. If for a given acquisition a candidate scanback track exists, the event is considered to be "found" in that acquisition. If there is more than one candidate scanback track, a single one is selected based on a series of ranking rules.

- A track identified as a muon in the electronic detectors is preferred over any track that is not. In fact, this corresponds precisely to the choice of prediction for vertex location: in events with a muon, it is always the muon which is used to locate the vertex.

- If the matching probability for a given candidate exceeds 1 % whereas that for another candidate does not, the former one is preferred. For reasons of efficiency, a strict cut on the matching probability is applied only at 0.1 %. Whenever there is ambiguity, a stricter cut seems appropriate.

- A track that is attached to a vertex with a given multiplicity is preferred over isolated tracks as well as tracks attached to a vertex with lower multiplicity. Two types of multiplicity are considered: the number of matched tracks and the total number of tracks. Vertices are first compared on their matched multiplicity, then on their overall multiplicity.

- If all of the above still leaves some ambiguity, the track attached to the vertex closest to the center of the fiducial volume is preferred or the track passing itself closest to the center if only isolated tracks remain.

Admittedly, some of these are arbitrary but one should keep in mind that the second ranking is only applied if the first one leaves any ambiguity, the third one if the second leaves ambiguity, and so on. It turns out that in the large majority of cases, the first criterion settles the issue. It has also been tested that variations of the above criteria do not alter the result in any significant way: the preference for tracks identified as muons is decisive, not only because the scanback track is always a muon but also because the quality of both the track fit and the matching probability is higher than for other tracks. The reasons for this have been discussed at length in sections 4.4 and 4.5. Once a single emulsion track matched to a single detector track has been singled out as the scanback track, any other matches involving the same emulsion or detector track are removed. If the scanback track is attached to one or more vertices, a scanback vertex can now be defined as well. The selection among multiple vertices is based on the multiplicity, with preference given to the one with the highest matched multiplicity or the highest overall multiplicity if the number of matched tracks attached to both vertices is equal. If this leaves any ambiguity, the most upstream vertex is chosen; the scanback track, at least if it is a muon, is expected to reach the primary vertex even though it may be compatible with a secondary vertex. For any detector track which is matched to an emulsion track attached to the scanback vertex, all other matches are removed.

4.8.4 Ambiguities

Now that a scanback track and vertex have been identified, we can attempt to use this information to resolve ambiguities that were left over in the attachment of tracks to vertices and in the matching between the emulsion and the electronic detectors. First of all, matches of different detector tracks to the same emulsion track are removed.

- If a match exists for which the probability is higher than 1 %, it is preferred over any for which this is not the case.
- To make sure that a maximum number of electronic detector tracks gets to be matched to emulsion tracks, preference is given to the match corresponding to the detector track with the smallest number of candidate matches.
- If everything else fails, we simply select the detector track for which the matching probability is highest.

Next, we consider matching ambiguities in the other direction: different emulsion tracks matched to the same detector track.

- Matches with a probability higher than 1 % are preferred over those where the probability is less and exceeds only 0.1 %.
- Any match to an emulsion track with only two segments is rejected if the same detector track can be matched to a track with more than two segments. As discussed before, various sources of background lead to short tracks, a problem that is most severe for tracks composed of only two segments.

- Preference is given to emulsion tracks which are attached to a vertex, comparing first the number of matched tracks, then the overall number of tracks between vertices.
- If everything else fails, the emulsion track for which the matching probability is highest is selected.

Finally, ambiguities in the attachment of tracks to vertices are addressed, *i.e.* tracks that are attached to more than one vertex.

- If the same track is attached to a vertex for which the fit probability is higher than 1 % and to a vertex for which it is not, it is removed from the latter.
- If a vertex to which a track is attached is situated in the plate immediately upstream from the most upstream segment on the track, it is preferred over any vertex for which this is not the case.
- First the matched multiplicity, then the overall multiplicity are compared between vertices to select the one with higher multiplicity.
- If everything else fails, we simply select the vertex for which the fit probability is highest.

This concludes the event reconstruction. The result is a set of tracks which are not passing, *i.e.* do not exit upstream from the fiducial volume. One or more of these tracks are matched to the electronic detectors, with the requirement that this matching is unique in both directions. One of these matched tracks has been singled out as the scanback track, possibly attached to a vertex referred to as the scanback vertex. Likewise, any of the other tracks may also be attached to a vertex, again in a unique way.

4.9 Selection of secondaries

Once the reconstruction described in the previous section has been performed, we can try to identify events in which a short-lived particle has been produced. In essence, this should be relatively simple: if all matched tracks are compatible with a single vertex, there is no indication of a short-lived particle; otherwise, there is. In passing, we note that this immediately excludes any secondary vertices for which none of the tracks is matched to the electronic detectors. But this of course assumes that any matched tracks do indeed belong to the event which cannot be taken for granted. On average, there are of the order of one hundred stopping tracks in any net scan acquisition and no more than a few of these are part of the event. Clearly, in spite of the efforts described above to define the matching probability as tightly as possible and to resolve ambiguities in a way which should enhance correct matches, some background will remain. This is the type of background to be addressed in the selection of secondaries, using rules about what a genuine event should look like, formulated in a manner as general as possible. Our primary interest is the production of charmed particles in charged current interactions. This somewhat simplifies the selection since, for charged current interactions, the scanback track will be the muon. Hence, we can assume that the scanback vertex is also the primary vertex or at least that the scanback track leads to the neutrino interaction, whether or not a corresponding vertex has been reconstructed. In the following, the term scanback track will be used interchangeably with primary muon,

just as scanback vertex and primary vertex will be. A vertex may not be reconstructed either because only the muon falls within the scanning angular acceptance, for instance in quasi-elastic events, or because of inefficiencies in the track and vertex finding.

Any track which belongs to the event but does not lead to the same vertex as the primary muon is a candidate secondary from either the decay of a short-lived particle or the interaction of a low-energy hadron. To profit from the reconstruction of the scanback vertex if one exists, the algorithms are slightly different between the case of a scanback vertex and the case of an isolated scanback track.

4.9.1 Scanback vertex

Each of the tracks matched to the electronic detectors but not attached to the scanback vertex is considered one by one. If the corresponding detector track has an angular difference of less than 50 mrad to a track attached to the primary vertex and not yet matched to any other detector track, the match of the candidate secondary track is removed and replaced by a match to the track attached to the primary vertex. The right panel of figure 4.9 shows that the angular difference in projection can be as large as 30 mrad, leading to tails in the angular difference in space $\sqrt{\delta a_x^2 + \delta a_y^2}$ that may extend up to 50 mrad. The fact that a track originates from the same vertex as the primary muon is a sufficiently strong indication that it belongs to the event to consider it as a match even though the matching probability itself is less than 0.1 %. As mentioned before, the matching probability will be underestimated for any tracks where the momentum hypothesis is too large, a problem which affects all tracks, and specifically hadrons, with a momentum of less than 1 GeV/c or a measured momentum considerably larger than the true value. Some of these cases are recovered here using the additional information that a track is attached to the primary vertex. These are responsible for the tail at probabilities below 0.1 % in figure 4.8.

Secondary track attached to a vertex

If the candidate secondary track is attached to a vertex, this vertex is considered a candidate secondary vertex. An angular acceptance on the parent is imposed by requiring the line connecting primary vertex and candidate secondary vertex to make an angle of at most 400 mrad to the longitudinal axis. The covariance matrix for the distance in space between the two vertex positions is simply the sum of their respective covariance matrices. If both vertices are in fact one and the same, their distance will be compatible with zero which can be formulated in terms of a probability noting that the quantity

$$\chi^2 \;=\; (\vec{v}_{1ry} - \vec{v}_{2ry})^T \;\; (V_{1ry} + V_{2ry})^{-1} \;\; (\vec{v}_{1ry} - \vec{v}_{2ry}) \tag{4.74}$$

is distributed as a χ^2 with three degrees of freedom for compatible vertices. If the corresponding probability is larger than 1 % or if the Euclidean length of $\vec{v}_{1ry} - \vec{v}_{2ry}$ is less than 20 μm, all tracks attached to the candidate secondary vertex are attached to the primary vertex and the other vertex is removed altogether. The cut on the distance itself serves two purposes. First, it excludes cases where the non-zero difference results from an underestimate of the track parameter errors, for instance if the momentum hypothesis is too big. Second, it excludes cases where a genuine secondary vertex could anyhow not be detected visually. As we will see later, manual checks provide one of the most powerful means to

validate the selection of secondaries. Whenever the effect on the efficiency is negligible, the selection focuses on those events for which such checks are possible.

If only one of the tracks attached to the secondary vertex is matched to the electronic detectors, we need to consider the possibility that the track belongs to the event but forms a chance vertex with background tracks. In this case, the distance of closest approach between the candidate secondary track and the primary vertex is evaluated using the procedure described in section 4.7.1. If the probability for this distance to be zero exceeds 10^{-4} or if the distance itself is smaller than 5 μm, the track is attached to the primary vertex, and the candidate secondary vertex is no longer of any interest. The cut on the distance is again related to the possibility of underestimated errors as well as the requirements of manual checks. The reconstructed track parameters may also be affected by the presence of a segment which does not belong to the track. In that case, the track fit probability tends to be lower and the cut on the distance is relaxed to 10 μm for tracks with a fit probability below 1 %. Finally, if the candidate secondary vertex has two tracks of which only one is matched and if both tracks have only two segments, the vertex as well as the attached tracks are rejected. As discussed before, various sources of background lead to short tracks. In particular, if on a given pair of plates, the alignment of the two years is sufficiently similar, the hundreds of passing tracks from the other year will all end up as stopping tracks of two segments. This explosion in the number of short tracks on a particular pair of plates greatly increases the risk of crossing tracks to be reconstructed as a vertex, with one of them associated by chance to a track in the electronic detectors.

A candidate secondary vertex which passes all of the above criteria is accepted. Unless it lies in the same plate as the primary vertex, a search is also performed to find the charged parent if it exists. If the two vertices are in adjacent plates, the raw data are accessed to check whether any segment is compatible with both vertices. Compatibility is defined on the basis of a probability based on the segment measurement errors described in section 4.2 and accounting for the extrapolation to the respective vertices. If there is more than one plate between the two vertices, any track attached to the primary vertex, not matched to the electronic detectors, and having no segment downstream from the secondary vertex, is considered as a possible parent. The probability for the distance of closest approach is used to decide whether it leads to the secondary vertex.

Secondary track isolated

We now consider a candidate secondary track not attached to any vertex. Because the scanning acceptance is limited to 400 mrad, the same acceptance can be applied to the parent direction by considering the line connecting the primary vertex to the most upstream segment on the track. This provides an effective rejection of random associations between detector tracks and emulsion tracks not related to the event.

Section 4.7.1 described the distance of closest approach of a track to a vertex as well as the probability for this distance to be compatible with zero. If this probability exceeds 10^{-4} or if the distance itself is smaller than 5 μm, the track is attached to the primary vertex. As before, the cut on the distance is relaxed to 10 μm for tracks with a fit probability less than 1 %.

The probability that a given track matched to the electronic detectors is a background track is obviously higher for isolated tracks than for tracks attached to a vertex. Fortunately, this is compensated by the fact that if the track is the only charged daughter of a short-

lived particle produced in the event, the parent itself should be charged and hence visible. Inevitably, this requirement leads to some inefficiency for neutral particles where one of the two daughters is outside of the scanning acceptance or has been lost because of scanning and track finding efficiencies. Isolated tracks stopping in the vertex plate or in the plate downstream from the vertex plate are accepted. Not only do they typically have up to five or six segments, the angular acceptance also confines them to a relatively narrow region. In contrast, tracks stopping in plates further downstream are shorter and thus more affected by background and the 400 mrad cone from the primary vertex covers the entire scanning region after only three plates. Unless there is a track attached to the primary vertex, not matched to the electronic detectors, and with its most downstream segment upstream of the most upstream segment on the candidate secondary track, the candidate secondary is rejected. For completeness, we mention that the raw data are accessed to find the parent for a track stopping in the plate downstream from the vertex plate. Again, χ^2 probabilities are used to select those segments which intersect the daughter track and reach the primary vertex. Even if no parent segment is found, the track is still selected as a secondary track to avoid the inefficiencies mentioned above. On the other hand, if the angular difference between the candidate parent and the candidate daughter is less than 30 mrad, the segment is added to the track and the track in turn to the primary vertex. This procedure is again meant to take into account the tails in the measurement errors and to remove events which cannot be confirmed visually.

4.9.2 Scanback track

The procedure to select secondary tracks in an event where the scanback track is not attached to any vertex is entirely similar to the case where the primary vertex has been identified. All other matched tracks are considered one by one, again with a slightly different treatment for secondary tracks attached to a vertex and isolated secondary tracks.

Secondary track attached to a vertex

Assuming that the neutrino interaction is situated somewhere along the scanback track, no more than one plate upstream from the last segment on the track, the possible directions of the parent connecting the candidate secondary and the primary vertex are calculated. If none of these is less than 400 mrad, the candidate secondary is rejected. Next, it is checked whether the scanback track might originate itself from the candidate secondary vertex, using the same conditions as in section 4.9.1 but substituting primary track and secondary vertex for secondary track and primary vertex. If this is the case, the scanback track is attached to the candidate secondary vertex, henceforth considered the primary vertex and any further selection relies on the procedures described in section 4.9.1.

As in section 4.9.1, special attention must be given to those candidate secondary vertices for which only one of the attached tracks is matched to the electronic detectors. For reasons described above, any such vertex with two tracks that both have only two segments, is systematically rejected. Otherwise, the distance of closest approach between the matched track and the scanback track is considered, defined using the expressions of section 4.7.2. If the probability for this distance to be compatible with zero exceeds 1 %, or if the distance itself is smaller than 5 μm, a primary vertex is built using the two tracks and any further selection relies on the procedures of section 4.9.1.

Any vertex passing the above criteria is selected as secondary vertex. The raw data are accessed to find any segments that intersect the scanback track and lead to the secondary vertex. If one is found, it is treated as the parent and used to build a vertex with the scanback track. Again, this brings us to the case of section 4.9.1.

Secondary track isolated

To define an angular acceptance for the parent, all directions are considered which connect the most upstream segment of the candidate secondary track to a point on the scanback track up to one plate upstream from its last segment. If none of these is less than 400 mrad, the candidate secondary is rejected. Next, a check is performed on the distance of closest approach between the two tracks. If the probability for this distance to be compatible with zero is larger than 1 %, or if the distance itself is less than 5 μm, a primary vertex is built, and any remaining matched tracks are treated on the basis of the criteria in section 4.9.1. In section 4.9.1, the requirement of a parent track was used to further reduce the background of isolated tracks, especially for tracks stopping several plates upstream of the vertex plate. In this case, the primary track itself is isolated as well so there are no candidate parent tracks. A different, but admittedly poor, solution to the same problem consists in simply restricting the fiducial volume beyond the 400 mrad angular acceptance. In particular, if the most upstream segments of the two tracks are more than one plate apart, the candidate secondary is rejected. If they are one plate apart, it is rejected unless the distance of closest approach is smaller than 200 μm. If both stop in the same plate, the distance of closest approach must be smaller than 100 μm.

This concludes the reconstruction of emulsion events, as well as the selection of short-lived particles. We had previously alluded to the possibility of multiple net scan acquisitions for the same event. Each of these is reconstructed as described here. After the reconstruction, the best event is selected simply on the basis of the number of matched tracks. The main origin of multiple acquisitions are multiple candidates on one of the plates during the vertex location. Whatever is the origin of the background track, it is most unlikely to lead to a vertex for which other tracks happen to match to the electronic detectors as well. For those few cases where two or more acquisitions contained several matched tracks, they always turned out to have an almost identical fiducial volume so it anyhow made no difference which one was chosen.

The results obtained with this reconstruction for the purpose of finding charmed particles produced in charged-current neutrino interactions, will be the subject of the next chapter.

5. THE SEMI-LEPTONIC BRANCHING FRACTION

Introduction

We have now come full circle. The first chapter described the theoretical framework for neutrino-induced charm production, showing how the study of this semi-inclusive process is sensitive to the strangeness content of the nucleon, the running mass of the charm quark, and the Cabibbo-Kobayashi-Maskawa matrix elements $|V_{cd}|$ and $|V_{cs}|$. In so-called dimuon events, a first muon originates at the primary neutrino vertex, a second one stems from the decay of the charm hadron. To date, the only measurement of neutrino-induced charm production including both hadronic and semi-leptonic charm decays was performed by the E531 experiment [54], relying on nuclear emulsion for the detection of short-lived particles [53]. For completeness, we should mention that the NOMAD experiment did measure D^{*+} production by neutrinos, relying on the excellent kinematic capabilities of the NOMAD detector to select events where the $D^{*+} \rightarrow D^0 + \pi^+$ decay is followed by $D^0 \rightarrow K^- + \pi^+$ [132].

The CHORUS experiment's capabilities for charm detection are similar to those of the E531 experiment, in a sample almost two orders of magnitude larger. Large statistics emulsion experiments have become possible thanks to the ever increasing speed of automatic scanning, culminating in the net scan technique, which has been described in the previous chapter. The reconstruction of net scan data leads to a reasonably pure selection of short-lived particles, simply by requiring the existence of a second vertex or an isolated track not compatible with the primary vertex. And here we get back to neutrino-induced charm production, the most abundant source of short-lived particles in the CHORUS experiment.

Applied to the CHORUS sample of net scan events, the reconstruction and selection algorithms define a set of candidate events for charmed particles produced in charged current neutrino interactions. Section 5.1 will discuss the existing net scan sample. As the scanning has not yet been completed, the sample should be considered as a snapshot, causing some distinct problems for normalization. The performance of the charm selection is best characterized by two numbers: the purity and the efficiency. The purity is the fraction of selected events that indeed contain a charmed particle, and can be reliably determined through manual scanning of the events. Section 5.2 describes the result of manual scanning for a subset of the events, showing also how a tighter selection achieves significantly higher purity with an acceptable loss of efficiency. The efficiency is the fraction of events containing a charmed particle which are selected, and can be estimated only on the basis of simulation. Section 5.3 describes the event simulation, which relies on a combination of generated events of a particular type with existing events to provide realistic background conditions. Section 5.4 compares the simulated events with the actual data in a number of variables. Overall, the agreement is very good, except for the distribution of multiplicity at the secondary vertex, where the simulation underestimates the number of single prong decays. Several hypotheses are explored, none of them being entirely satisfactory. We will

however concentrate on the determination of the semi-leptonic branching ratio of charmed particles, which depends only on relative efficiencies, and large variations in the simulation hypotheses have a fairly small effect on this determination. Section 5.5 wraps up these issues and presents the selection efficiency for each charm species, including the case where the decay is muonic. The remaining ingredient to determine the semi-leptonic branching ratio is the identification of the secondary muon, described in section 5.6. Finally, section 5.7 presents the actual determination and compares our result with the indirect measurements from other experiments.

5.1 The data sample

Even though the CHORUS experiment has collected emulsion data for four years, current analyses all concentrate on the Run II data corresponding to the 1996 and 1997 exposures. In several respects, the Run I data corresponding to the 1994 and 1995 exposures are less interesting than the Run II data:

- the statistics represents less than a third of the total because of a lower neutrino beam intensity, combined with lower data taking efficiency in 1994,
- up to a quarter of the emulsion modules is damaged and cannot adequately be scanned, partly because some of the emulsion gel used was NIKFI gel suffering from more severe fading compared to FUJI gel,
- the momentum measurement for hadrons cannot profit from the honeycomb chambers and emulsion trackers which were installed for Run II,
- because of frequent changes in the scanning algorithms for vertex location, the sample of located events is particularly inhomogeneous.

As a result, when net scan data taking started at the Nagoya FKEN laboratory in the summer of 1999, it was decided to give priority to located events from Run II. Other scanning laboratories are now performing this type of vertex analysis as well, but have generally started later. At the time of writing, the accumulated statistics of net scan data collected at the Nagoya FKEN laboratory is an order of magnitude higher than that of all other scanning laboratories combined. To avoid the problems related to the combination of samples with different characteristics, we will in the following forgo any data from other laboratories. This is unfortunate, especially for what concerns the data taken with the CERN microscopes as the present work mostly describes their setup. Nevertheless, it should be pointed out that the reconstruction of Nagoya net scan data described in this work has benefited tremendously from developments initiated in the CERN scanning lab.

Unfortunately, not even the Nagoya data set should be considered truly homogeneous.

- Vertex analysis is currently being performed for all events previously located with the scanback technique. Over time, both the hardware and the algorithms used for vertex location have changed.
- The hardware and algorithms used in the vertex analysis were still being refined during the first months of net scan data taking, affecting specific modules.

- A data quality check is performed for the vertex analysis data, based on the alignment between consecutive plates. Acquisitions for which this check fails are later repeated. Given that the current data set is essentially a snapshot of an ongoing data acquisition process, some modules have gone through more iterations than others.

These inhomogeneities are clearly reflected in table 5.1, describing the data sample which will be used in all of the following.

This sample consists entirely of Run II events, *i.e.* from the beam exposures in 1996 and 1997, for which the vertex location was performed previously, rejecting any events where the reconstructed muon momentum is larger than 30 GeV/c. Net scan data were all taken in Nagoya up to July 17, 2001, for a bit less than half of the sixteen halfmodules in each stack. The first row in the table shows the number of events for which the primary vertex in the electronic detector reconstruction is situated in the fiducial volume of the relevant halfmodules. The second row shows the fraction of these events for which one or more tracks reconstructed in the target tracker are matched to a muon, identified either in the muon spectrometer or as a minimum ionizing particle in the calorimeter. The remainder consists of neutral current events as well as charged current events where the primary muon was not identified, either because the track is outside of the acceptance or because of inefficiencies in the muon identification itself or in the matching between the muon spectrometer and the target tracker.

The third row shows the number of events for which at least one set of net scan data satisfies the data quality criteria. The loss, compared to the number of events with one or more identified muons, is primarily due to the vertex location efficiency of 45 %. In the following, events for which net scan data exist will be referred to as *located*. An additional loss is related to events which have been located but for which no net scan data has yet been taken, or at least none which satisfied the data quality check. If at any stage during the vertex location, there are multiple candidates, then both will be followed and vertex analysis will be performed independently for their respective 'vertex' plates. This is the origin of multiple data sets, indicated in the fourth row. To reduce the background, the scanning procedures for changeable and special sheets described in section 3.5.4 were changed considerably between the 1996 and 1997 data. The result is reflected in a lower fraction of multiple data sets, but unfortunately also in a lower vertex location efficiency. However, it should be noted that the 1997 data is more severely affected by hardware problems during net scan data taking since an acquisition for a given module first treats all of the 1996 events, then all of the 1997 events. Over time, this difference should become less important as the scanning proceeds, until a complete data set exists for both years.

The one but last line shows the number of events for which, in at least one of the acquisitions, one or more reconstructed tracks in emulsion are matched to a track in the electronic detectors. The loss, compared to the number of events for which net scan data exists, has two components. First, the scanback track may have been a background track, picked up during the vertex location but not related to the event. In that case, the event will be rejected on the basis of net scan data where the angular resolution is of the order of 1 mrad in each projection, to be compared with 6 mrad for the scanback data. Second, none of the tracks belonging to the event may have been reconstructed and matched due to inefficiencies. This affects in particular events with low multiplicity. If at least one emulsion track is matched to a track in the electronic detectors, the event is referred to as *found*.

1996 data taking	stack 1 7 halfmodules		stack 2 8 halfmodules		stack 3 6 halfmodules		stack 4 7 halfmodules	
Predicted in emulsion	29952		36932		25170		33869	
Identified muon(s)	22428	74.9 %	24636	66.7 %	18770	74.6 %	24596	72.6 %
Net scan data ($p_\mu < 30$ GeV/c)	8277	36.9 %	9648	39.2 %	8284	44.1 %	10998	44.7 %
(multiple sets)	(1527	18.5 %)	(1519	15.7 %)	(2411	29.1 %)	(2864	26.0 %)
Matched track(s)	6603	79.8 %	8126	84.2 %	5905	71.3 %	8097	73.6 %
Matched muon(s)	5922	89.7 %	7256	89.3 %	4910	83.1 %	6904	85.3 %
vs. identified muon(s)		26.4 %		29.5 %		26.2 %		28.1 %
1997 data taking	stack 5 7 halfmodules		stack 6 7 halfmodules		stack 7 8 halfmodules		stack 8 8 halfmodules	
Predicted in emulsion	31392		35194		38824		36932	
Identified muon(s)	23011	73.3 %	23759	66.7 %	26589	68.5 %	28515	66.8 %
Net scan data ($p_\mu < 30$ GeV/c)	6435	28.0 %	6112	25.7 %	7942	29.9 %	7828	27.5 %
(multiple sets)	(392	6.1 %)	(216	3.5 %)	(485	6.1 %)	(841	10.7 %)
Matched track(s)	5887	91.5 %	5660	92.6 %	7229	91.0 %	7048	90.0 %
Matched muon(s)	5718	97.1 %	5437	96.1 %	7035	97.3 %	6782	96.2 %
vs. identified muon(s)		24.8 %		22.9 %		26.5 %		23.8 %

Tab. 5.1: Statistics of the net scan data collected up to July 17, 2001 in the Nagoya FKEN laboratory for modules where the scanback procedure imposed an upper limit of 30 GeV/c on the measured muon momentum. Each row corresponds to a subset of the previous row, with the fraction indicated. The individual items are described in the text.

	module 18		module 37		module 38		module 67		module 88	
Reconstructed in emulsion	7209		9761		7810		11485		9520	
Identified muon(s)	5192	72 %	6842	70 %	5405	69 %	7899	69 %	6271	66 %
Net scan data	1093	21 %	1775	26 %	1171	22 %	3068	39 %	939	15 %
(multiple sets)	(82	8 %)	(188	11 %)	(102	9 %)	(390	13 %)	(81	9 %)
Matched track(s)	872	80 %	1447	82 %	948	81 %	2695	88 %	681	73 %
Matched muon(s)	827	95 %	1359	94 %	907	96 %	2483	92 %	632	93 %
vs. identified muon(s)		16 %		20 %		17 %		31 %		10 %

Tab. 5.2: As in table 5.1, for the modules where the scanback procedure was performed for all events with an identified muon without imposing a momentum cut of 30 GeV/c.

	1996 data				1997 data			
	with p_μ cut		without p_μ cut		with p_μ cut		without p_μ cut	
Matched muon(s)	24992		3093		24972		3115	
Selected	612	2.4±0.1 %	64	2.1±0.3 %	647	2.6±0.1 %	119	3.8±0.4 %
Selected dimuon	57	9.3±1.2 %	6	9.4±4.0 %	59	9.1±1.2 %	6	5.0±2.1 %

Tab. 5.3: Number of events which show an indication for the decay of a short-lived particle, through the presence of a second vertex or a track matched to the electronic detectors but not compatible with the primary vertex. The last line shows the number of selected events with two or more identified muons in the electronic detectors. The errors are statistical only, and assume Poisson distributions for the number of events in each category.

Finally, for this analysis, we consider only events where one or more emulsion tracks are matched to the tracks identified as muons in the electronic detectors. This further reduces the background related to wrong tracks in the vertex location combined with a spurious match between an emulsion track and a detector track. As shown in figure 4.9, the angular difference between tracks in the emulsion and in the electronic detectors is almost a factor two wider for hadrons than for muons because of the softer spectrum leading to more important multiple scattering between the net scan fiducial volume and the downstream face of the stack. Furthermore, the requirement of a matched muon leads to a more homogeneous data sample, since the reconstruction of net scan data relies on the muon to resolve any ambiguities, as explained in section 4.8.4. As indicated in table 5.1, the differences between efficiencies from one module to the next mostly cancel when considering the fraction of reconstructed events with one or more identified muons for which the muon is matched to the corresponding net scan data. The average fraction of reconstructed events in the fiducial volume which can be used for analysis is 27.6 % for 1996 and 24.5 % for 1997, with the difference between the two years accounted for by the tighter selection in changeable sheet and special sheet scanning combined with the larger effect from incomplete net scan data taking. In total, the data up to July 17, 2001 contain 49,964 events with a track in emulsion matched to a muon in the electronic detectors, 24,992 events from 1996 and 24,972 from 1997.

The net scan data in table 5.1 is entirely based on events for which vertex location was performed in the framework of the oscillation analysis. An event with an identified muon can only be a candidate ν_τ charged current interaction if the muon does not originate at the primary vertex, but is the decay product of the tau lepton with the decay signalled by the change in track direction, the so called kink. The kink finding efficiency is a function of the kink angle and decreases rapidly below 30 mrad. This region corresponds to high momenta for the tau lepton parent and muon daughter. Hence, muons with a reconstructed momentum above 30 GeV/c offer little scope for the oscillation search and such events were not considered in the vertex location to reduce the scanning load. Unfortunately, this runs counter to the requirements of a search for neutrino-induced charm events, for which the cross section increases as a function of the neutrino energy because of the charm mass threshold.

Recently, some effort has been made to locate events without requiring the muon momentum to be below 30 GeV/c. The data of this type, collected up to 17 July 2001 in five modules, is summarized in table 5.2. The various fractions fluctuate even more than in table 5.1, simply reflecting the very incomplete state of this data taking. In total, there are 6,208 located events for which at least one track in at least one set of net scan data can be matched to a muon in the electronic detectors. In principle, a sample without the muon momentum cut could serve to estimate the effect of this cut on the determination of the semi-leptonic branching ratio. Unfortunately, the current statistics of twelve selected dimuon events is too small to make such a comparison worthwhile. Instead, we will impose the cut on the muon momentum a posteriori and combine this sample with the samples described in table 5.1.

The simplest possible selection of charmed particles consists in requiring the presence of a secondary vertex in the result of the event reconstruction, described in the previous chapter. In this context, both primary and secondary vertex may also be a single track matched to the electronic detectors. For genuine vertices, at least one track must be matched to the electronic detectors. The result of such a selection, for events with an identified muon at the primary vertex, is shown in table 5.3. In total, there are 1442 candidate events for charged current neutrino interactions involving the production of a short-lived particle. Out of these, 128 events or 8.7 % contain two muons in the electronic detectors.

5.2 Manual checks

The sample of table 5.3 provides an initial selection for neutrino-induced charm production. Our purpose will now be to further refine this selection and to determine the purity, *i.e.* the fraction of genuine charm events among the selected events. To this end, we rely on what is commonly referred to as manual scanning, the visual inspection of the emulsion event under the microscope and in particular of the track or tracks that have been reconstructed as secondaries. Since net scan data are always taken on the upstream side of the plate, the most upstream segment measured for a track originating in plate n will be in plate $n+1$. The position of this segment, combined with the alignment information, is used to direct the microscope to the appropriate position on plate $n+1$. The emulsion image is displayed on a video screen with a graphics overlay indicating the track direction to serve as a visual aid to locate the track. Once the correct track has been visually identified, it is followed upstream to determine its origin. For genuine decay topologies or secondary interactions, the tracks that have been reconstructed as candidate secondaries will lead to a secondary vertex or show a kink before reaching the primary vertex. For background due to reconstruction artefacts, they will pass through the entire plate or lead to the primary vertex without showing any distinct kink point.

Broadly speaking, there are two types of background: genuine secondary vertices not due to charm, and fake secondary vertices due to shortcomings in the reconstruction. The former include secondary interactions of hadrons produced at the neutrino vertex, as well as decays of strange particles. These will be discussed later. The latter can be referred to as reconstruction artefacts and consist of five categories.

Primary tracks. The reconstructed impact parameter for a primary track with respect to the remaining tracks may be incompatible with zero due to mismeasurement of the track parameters or non-gaussian errors. For the same reasons, two or more primary tracks may form a reconstructed vertex displaced from the primary vertex.

Electron-positron pairs. Because of the narrow opening angle, an electron-positron pair typically demonstrates itself as a single dark track for one or two plates until the two trajectories are sufficiently separated so that they can be distinguished. If the reconstruction builds only a single track, it may match a track in the electronic detectors and have a non-zero impact parameter to the primary vertex. Arguably, these are genuine secondary vertices as well. They are here classified as reconstruction artefacts to reflect the fact that a pair consisting of two reconstructed tracks would be rejected on the basis of its small opening angle.

Passing tracks. Due to scanning and track reconstruction inefficiencies, a passing track may appear to be stopping. If the angle happens to be compatible with a track in the electronic detectors not yet matched to a track attached to the primary vertex, such a passing track will be selected as a candidate for a secondary track.

Fake vertex. A pair of low energy tracks or passing tracks reconstructed as stopping may appear to form a vertex when their extrapolation is compatible within errors. Such cases will be selected if one of the attached tracks has an angle compatible with an electronic detector track, as in the case of passing tracks.

Matched muon(s)	13178	
Selected	341	2.6±0.1 %
Not checked	14	4.1±1.1 %
Checked	327	95.9±1.1 %
confirmed decays	264	80.7±2.2 %
secondary interactions	29	8.9±1.6 %
background	34	10.4±1.7 %
primary	13	
electron pair	6	
passing	8	
fake vertex	3	
unrelated	4	

Tab. 5.4: Events from stacks 1 and 2 selected as candidates for neutrino charged current interactions involving the production of a short-lived particle, with the breakdown of background based on manual scanning. The different categories are defined in the text. The errors are statistical only and assume Poisson distributions for the number of events in each category.

Unrelated. Finally, there is a category of so-called unrelated topologies, linked to the fact that nuclear emulsion has no time discrimination. Within the angular acceptance of the event under study, there may be a second neutrino interaction, or the decay or secondary interaction of a particle originating further upstream. Any such event topology will be selected if at least one of the daughter tracks has an angle compatible with an electronic detector tracks, as in the case of passing tracks.

The primary task of manual scanning consists in confirming genuine decay topologies and rejecting secondary interactions or reconstruction artefacts. From the point of view of a selection based on the net scan data, decays and secondary interactions must both be considered as signal as the distinction is based on the presence of heavily ionizing nuclear fragments, Auger electrons, or blobs at the secondary vertex. None of these features manifests itself in the net scan data which only measures tracks near the minimum of ionization at the upstream surface of each plate. Only image data taken across the full thickness or indeed manual checks permit the separation of decays and secondary interactions. On the other hand, background due to reconstruction artefacts clearly points to the shortcomings in the automatic analysis of net scan data. In that sense, manual checks offer the possibility to further improve the automatic selection, to the extent that signal and background events show different characteristics in automatically reconstructed variables. Table 5.4 summarizes the results from manual checks on 341 events selected out of a sample of 13,178 events with an identified muon, corresponding to stacks 1 and 2 from the 1996 data. Apart from the categories described above, there is a further category of 14 *not checked* events. It is an unfortunate consequence of the fact that manual checks require physical access to the emulsion, which for this sample is stored in Nagoya. During my stay at the Nagoya FKEN laboratory, I performed manual scanning for all selected events which had not previously been checked. In the meantime, the correction of small mistakes in the reconstruction software and the tuning of parameters have slightly altered the selection introducing some events not

present in the original sample. Simply considering all of these as background is the most conservative assumption, but not an entirely unreasonable one since their selection depends on slight variations in the input parameters. Furthermore, as will be shown below, all of the not checked events have certain characteristics that are typical of background events.

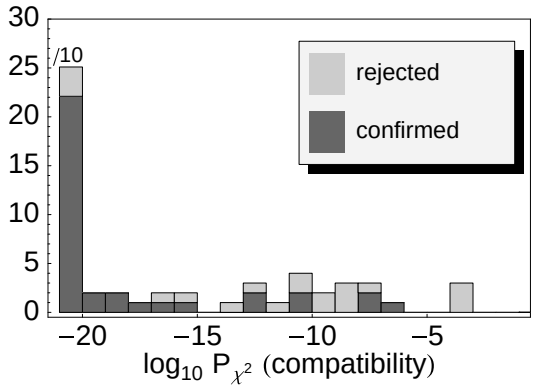


Fig. 5.1: Distribution of the χ^2 probability for the track or vertex reconstructed as secondary to be compatible with the track or vertex reconstructed as primary for selected events without a reconstructed charged parent track. Manually confirmed events include genuine decay topologies as well as secondary interactions. The leftmost bin includes all χ^2 probabilities below 10^{-20} and has been scaled down by a factor of ten.

The essential requirement in the selection described in the previous chapter is for the distance between two vertices, or the impact parameter between a vertex and an isolated track or between two isolated tracks, to be incompatible with zero, through a χ^2 probability criterion. Figure 5.1 shows the χ^2 probability for the distance measure, separating the manually confirmed and manually rejected events. Events in the category *primary* would ideally show a non-zero probability, but it is clear from the figure that a tail is present extending down to probabilities of 10^{-10} and even 10^{-15}. It is related to the combination of non-gaussian contributions to the segment measurement errors due to strong local distortion, the inclusion of outlier segments in the track fit or outlier tracks in the vertex fit, and wrong estimates of the multiple scattering contribution due to mismeasurement of the particle momentum. On the other hand, genuine decay topologies and secondary interactions are concentrated at probabilities below 10^{-20}, a direct consequence of the exceptional spatial resolution in nuclear emulsion.

Figure 5.2 considers only those events where the candidate secondary is an isolated track, 77 events out of a total of 317 for which the probability for primary and secondary to be compatible is less than 10^{-15}.. In this case, if the low probability for the distance measure to be compatible with zero is due to the inclusion of outlier segments or to an underestimate of the multiple scattering errors, then one should also expect the track fit to show a small probability. Indeed, the left panel shows that the track fit probabilities are peaked towards small values for rejected events. In fact, the track fit probability also offers some rejection power for passing tracks. Passing tracks are reconstructed as stopping if no segment is found in either of the next two plates upstream. This may be due to scanning efficiency

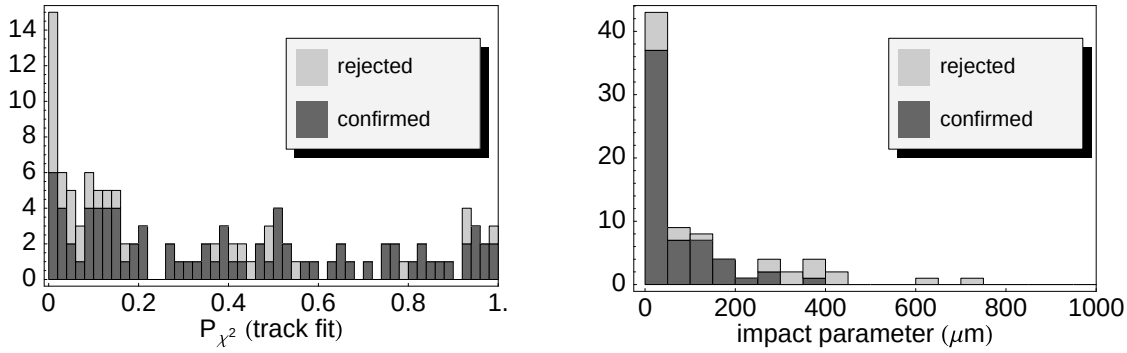


Fig. 5.2: Selected events with an isolated track as reconstructed secondary for which manual checks were performed to separate genuine decay topologies and secondary interactions from background due to reconstruction artefacts. Left panel: distribution of the χ^2 probability for the track fit of the isolated track. Right panel: impact parameter of the isolated track with respect to the primary vertex, or to the primary track if it is isolated as well.

but also to a wrong hypothesis because of a bad track fit, reflected in a low probability. For confirmed events, the probability distribution is comparable to the one shown in figure 4.6, as expected.

Since passing tracks are unrelated to the event under study, they are randomly distributed in the fiducial volume of the net scan acquisition. This is partly dealt with using the angular acceptance criteria described in section 4.9. However, for tracks stopping two plates downstream from the primary vertex, $\pm$ 400 mrad still corresponds to typically $\pm$ 450 μm in each projection, or about 30 % of the entire scanning volume. This applies even more to large angle tracks. Such cases lead to a large impact parameter, and indeed the right panel of figure 5.2 shows several rejected events with an impact parameter larger than 300 μm, compared to a marginal fraction of the confirmed events.

The above considerations all hint towards criteria to reduce the background from the categories *primary* and *passing*. Electron pairs can to some extent be considered a subcategory of *primary*. This leaves the fake vertices to be dealt with. One obvious requirement is for the secondary vertex to lie downstream from the primary vertex. As seen in figure 5.3, some events are selected for which this is not the case and the majority of them is rejected manually. In a handful of confirmed events, the secondary vertex lies upstream from the primary vertex due to wrong assignment of tracks to vertices or to a wrong identification of primary and secondary vertex. A second criterion for fake vertices is based on the vertex depth, defined as the longitudinal distance between the reconstructed vertex position and the most upstream segment on any of the tracks attached to the vertex. For genuine vertices, the two or more tracks are all expected to have a segment in the plate immediately downstream from the vertex plate, corresponding to a vertex depth in the range from -100 μm to 800 μm. The lower limit relates to the depth covered in the data taking and to the distance between the data taking volume and the physical surface of the plate. The upper limit is nothing but the plate thickness. A vertex depth smaller than -100 μm is unphysical, whereas a vertex depth larger than 800 μm is only possible if the most upstream segment is missed

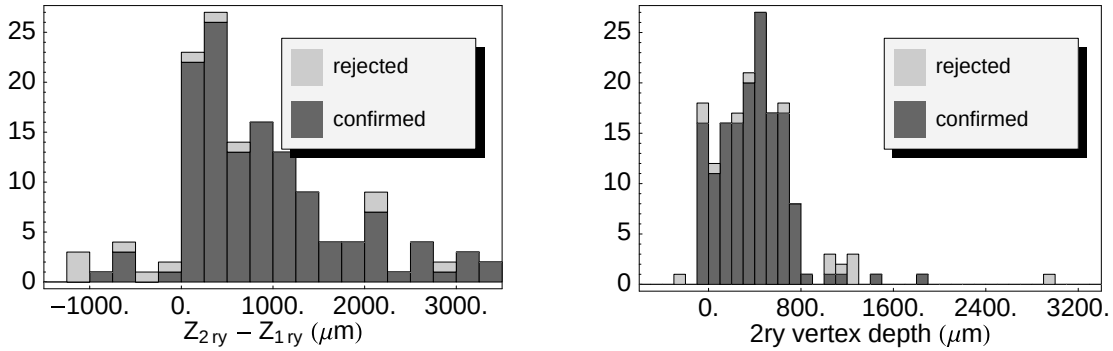


Fig. 5.3: Selected events with two reconstructed vertices for which manual checks were performed to separate genuine decay topologies and secondary interactions from background due to reconstruction artefacts. Left panel: distribution of the longitudinal distance between the reconstructed position of the two vertices. Right panel: distribution of the longitudinal distance between the reconstructed position of the secondary vertex and the most upstream segment in any of the attached tracks.

on all tracks attached to the vertex. Rather more likely, it may be due to fake vertices, seen in the right panel of figure 5.3 to be the main cause.

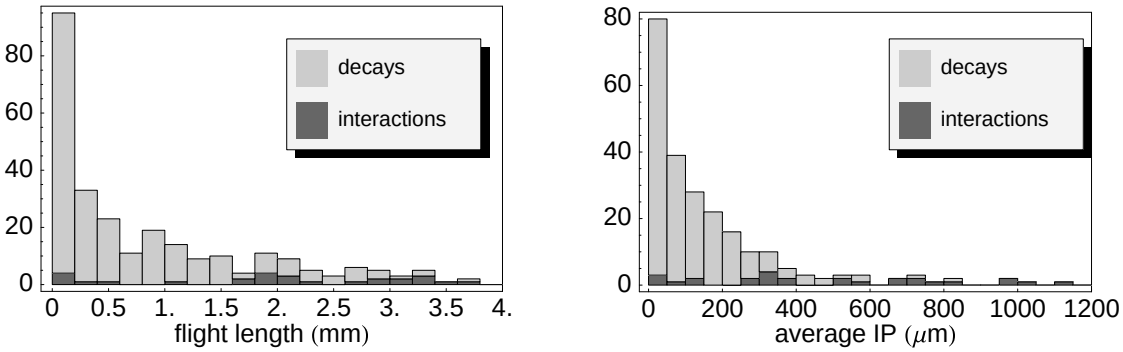


Fig. 5.4: Flight length distribution for manually confirmed events without any nuclear activity, presumable decays, and with nuclear activity, indicating a secondary interaction. Left panel: the distance between the primary and secondary vertex if the position of both can be reconstructed. Right panel: the average impact parameter as defined in the text for events where the position of the primary vertex can be reconstructed.

So far, to evaluate the quality of the selection based on reconstruction of the net scan data, we have lumped together secondary interactions with the actual signal of decay topologies. From the point of view of a charm analysis, however, these should be considered as a source of background. Apart from the nuclear activity at the interaction vertex, there are two further characteristics which distinguish interactions from decays.

First, because of charge conservation, the number of daughter tracks is odd for the decay

of a charged particle and even for that of a neutral particle. For secondary interactions, this is not necessarily the case. However, due to the angular acceptance of the scanning and to the track finding inefficiency, there is a significant difference between the reconstructed and the true multiplicity at the secondary vertex. Furthermore, if the primary and secondary vertex are situated in the same plate, the automatic reconstruction is unable to determine whether the parent is charged or neutral. Hence, this type of topological information, however useful it may be in the context of manually checked events, cannot be exploited based on the automatic reconstruction alone.

The second difference is related to the flight length distribution. For secondary interactions, it is flat except for variations in the selection efficiency as a function of flight length. For decays, this distribution falls off exponentially reflecting the proper time distribution of the different charm species convoluted with their respective momentum spectra. Based on the net scan data, the flight length cannot be reconstructed for events where the primary or secondary are isolated tracks since the track origin may lie anywhere between the scanning region on two consecutive plates. This affects in particular events where a secondary vertex with a single outgoing track lies in the same plate as the primary vertex. The left panel of figure 5.4 shows the flight length distribution for those events where it can be reconstructed. Taking into account the implicit rejection of events with short flight length, the above prediction is at least qualitatively borne out by the data. However, the flight length does not by itself permit a rejection of secondary interactions without a corresponding loss in the signal of charm decays with large flight lengths.

On the other hand, since the mean proper decay length of all charmed particles is 315 μm or less, decay lengths larger than 2 mm necessarily correspond to high momenta and a γ factor of about five to ten. In turn, a large Lorentz boost of the parent particle leads for the daughter particles to small emission angles in the laboratory frame. For secondary interactions, no such correlation between flight length and emission angle exists. These considerations can be formalized by noting that the impact parameter of a secondary track with respect to the primary vertex is nothing but the product of the flight length of the parent and the sine of the emission angle of the secondary track with respect to the parent direction. Since the flight length is proportional to γ and the emission angle is inversely proportional to γ, the impact parameter is independent of the parent momentum and directly reflects the proper decay length up to a constant angular factor. For secondary interactions, a large flight length simply results in a large impact parameter. The right panel of figure 5.4 shows the distribution of the impact parameter where it can be defined, namely events where the primary is not an isolated track. If the secondary multiplicity is larger than one, the average of the impact parameter over all daughter tracks is used. As expected, the impact parameter shows a better separation between decay topologies and secondary interactions than the geometrical flight length alone.

In figures 5.1, 5.2, 5.3, and 5.4 the separation between signal and background is based on manual checks, but the quantities shown are all calculated in the automatic reconstruction. Because of the different distribution of signal and background in each of these variables, they can serve to improve the selection purity of 76.9 % + 8.9 % = 85.8 % quoted in table 5.4. To this end, the following cuts are applied:

- for events without a reconstructed charged parent, a probability for the distance measure to be compatible with zero smaller than 10^{-15} (figure 5.1),
- for isolated secondary tracks, a track fit probability larger than 2 % and an impact

	all		'loose' cuts		'tight' cuts	
Selected	341		294		262	
confirmed decays	264	77 %	245	83 %	236	90 %
secondary interactions	29	9 %	28	10 %	14	5 %
background	48	14 %	21	7 %	12	5 %
not checked	14		0		0	
primary	13		5		3	
electron pair	6		5		3	
passing	8		6		2	
fake vertex	3		1		1	
unrelated	4		4		3	

Tab. 5.5: As in table 5.4, including the results after additional cuts. The sets of 'loose' and 'tight' cuts are described in the text.

parameter smaller than 300 μm (figure 5.2),

- for secondary vertices, a vertex depth in the range from -100 μm to 1000 μm and a positive longitudinal distance from the primary vertex (figure 5.3),
- for events where it can be defined, an impact parameter or average impact parameter smaller than 500 μm.

In table 5.5, these are referred to as the 'tight' cuts. Their application results in a reduction of the background by a factor of four and the number of secondary interactions by a factor of two, while keeping 90 % of the events with a confirmed decay topology. None of the events which have not been checked pass these additional criteria. Hence, the large majority of them, if not all, are most likely to be background events. In any case, if they are not, this would only affect the relative efficiency of the selection including the 'tight' cuts with respect to the original selection, but not the purity of the selection after application of the 'tight' cuts. For the 'loose' cuts, the criteria on isolated secondary tracks are relaxed to a fit probability of at least 1 % and an impact parameter smaller than 500 μm, and no selection is applied on the average impact parameter. However, as shown in the table, the effect on the efficiency is marginal whereas the background rejection is severely compromised. In the following, only the 'tight' cuts will be used. The remaining background of 12 events corresponds to less than 0.1 % of the original sample of 13,168 events with an identified muon.

5.3 Event simulation in emulsion

The effects to be included in a simulation of the emulsion response are legion. First of all, the CHORUS emulsions are exposed for two years, integrating passing tracks from cosmic ray particles, nearby beams, and neutrino interactions upstream. Then, the plates are subjected to misalignment with respect to each other and significant distortions are introduced locally during development. Development itself is a physico-chemical process where minor variations can measurably affect the fog density as well as the grain density on minimum ionizing tracks. Once processed, the plates are scanned on microscope tables with slight variations in

the resolution of the mechanics as well as in the response of the optical system. Only then do we reach the stage where further steps are amenable to simulation: online segment finding by the track selector hardware, and offline reconstruction by the algorithms described in the previous chapter.

Fortunately, a global simulation can be avoided by using actual data and simulating the presence of events of the type considered , amongst this background of real data. Net scan data taking at arbitrary positions in the emulsion stack by definition provides the most realistic background conditions. Starting from the true track parameters of all charged particles in a simulated neutrino interaction, the following steps are taken to combine an 'empty' acquisition with a simulated event of the appropriate type.

Propagation. Each charged track is propagated from the point of production through the net scan fiducial volume to obtain the position and direction at the upstream surface of each plate. The propagation is based on the model for multiple scattering elaborated in section 4.3, considering the true momentum of the track.

Smearing. The measured position and direction of each segment is obtained by smearing the propagated position and direction according to the error model described in section 4.2. This accounts for the emulsion distortion as well as for the mechanical accuracy of the microscope table. Differences between microscopes are included by virtue of the fact that the parameters of the error model are independently determined for each module, and all plates for a given module are always scanned on a single microscope table. Non-gaussian tails in the measurement errors, even though they are known to be present, are not simulated.

Hardware efficiency. The track selector algorithm generates a so called pulseheight, equivalent to the number of frames out of sixteen for which a grain is found. Offline, the threshold used in all of the reconstruction algorithms is 12 whereas the online threshold is only 10. Using the large number of background tracks, the pulseheight distribution has been parametrized as a function of the radial and transverse angle of the track, independently for each emulsion module. The simulated pulseheight is based on this parametrization, implicitly accounting for the hardware efficiency.

Alignment. As mentioned before, the alignment from one plate to the next in the net scan fiducial volume is based on the passing tracks not related to the event and in the case of simulation provided by the 'empty' acquisition. The inverse of the alignment transformation is applied on the simulated segments to obtain segments in the same coordinate system as the raw data from the 'empty' acquisition.

The result consists of a data set that can be treated in exactly the same manner as the real data, but now including an event with known physical characteristics.

Samples of neutrino interactions for a beam spectrum simulated by GBEAM [133, 134] are produced by the JETTA event generator [135], which simulates deep-inelastic neutrino interactions, including deep-inelastic charm production. The JETTA generator is derived from JETSET 7.4 [136], combined with LEPTO 6.1 [137]. This event generator is explicitly restricted to the deep-inelastic regime by a cut of 2 GeV/c^2 on W, the invariant mass of the hadronic system. The quasielastic regime, including the production of resonances, is described independently by the RESQUE event generator [138]. However, for the present analysis, no samples have been generated with RESQUE as it does not include quasi-elastic

charm production. Furthermore, events with less than 4 GeV measured in the calorimeter will be explicitly discarded from the data sample. The JETTA event generator offers a description of deep-inelastic processes which is sufficient for our purposes, even though some of the effects discussed in chapter 1 are not taken into account. Out of the nuclear effects, described in section 1.4.1, non-isoscalarity is accounted for in each of the materials where the neutrino can interact, Fermi motion is modelled with a simple Fermi distribution for the transverse momentum of nucleons in the nucleus, shadowing and the EMC effect are entirely discarded. As far as the electroweak radiative corrections of section 1.4.2 are concerned, these are included for the hadronic part, but not for the final state radiation connected to the primary muon. Instead of the slow rescaling prescription introduced in section 1.6, the effect of the charm quark mass is accounted for by a kinematical cut. This provides a good approximation, but leads to a small shift in the x distribution and does not include the target mass effect. Finally, the partonic subprocess is described at leading order in the strong coupling constant α_s, discarding the contributions from the boson-gluon fusion and radiative gluon processes of section 1.7.

The sample of neutrino interactions generated with JETTA is processed by the EFICASS detector simulation for the CHORUS electronic detectors, based on GEANT 3.21 [139]. The result consists of raw data, formally equivalent to the raw data collected by the detectors, and it is processed by the CHORAL reconstruction program.

At this point, it is worth pointing out two limitations of the simulation as it stands now. First, the set of input events was generated in 1999, and since then a number of mistakes has been uncovered in the various programs involved. At the generator level, the only significant change is related to the choice of structure functions and a more accurate beam spectrum. For the present analysis, both of these have a negligible effect. At the detector simulation level, a considerable number of parameters have been tuned based on comparisons with the data. Meanwhile, the CHORAL reconstruction has been superseded by the CHANT package. An unfortunate consequence of this has been that any sets of simulated data generated after 1999 can no longer be treated by CHORAL because of incompatibilities in the data format. On the other hand, the present analysis is entirely based on real data reconstructed using CHORAL so any use of more recent sets of simulated data would break the symmetry between simulated and real data.

A second problem is related to the choice of 'empty' boxes. With the data taking presently still incomplete, priority is obviously given to the real data rather than to the scanning of arbitrary positions in the emulsion volume. Nevertheless, some of the data taken for a different purpose can actually be used as if it were simply taken at an arbitrary position. As mentioned before, the vertex signal consists of a scanback track not found in two consecutive plates. However, in some cases this is due to inefficiencies. More specifically, if there are multiple candidates at any stage during the scanback, both will be followed. Of course, only one of them will actually lead to the neutrino interaction vertex. Any others are typically passing tracks falling within the angular acceptance but eventually lost again. Once the corresponding net scan data is analyzed, no track will be found since the matching acceptances are considerably tighter. In that sense, any acquisitions where no track is matched can be considered 'empty'. But there is a caveat here: an actual event may be present but without any track matched. This might occur as a result of inefficiencies, for instance due to bad emulsion conditions. If an event is present at the center of the fiducial volume, it is likely to cause confusion with the simulated event overlaid at the same position. This problem is overcome by applying all data quality checks on empty events to ensure that none of them

is affected by emulsion conditions or alignment problems.

5.4 Comparison between data and simulation

Manual checks of selected events allow the determination of the selection purity, *i.e.* the fraction of selected events which show a secondary topology, with or without nuclear activity. To fully characterize the selection, we also need the selection efficiency, *i.e.* the fraction of events with a charmed particle which are selected. This can only be determined from the simulation. However, it is worth stressing that the measurement of the semi-leptonic branching ratio of charm requires only the relative efficiencies. The charm quark produced at the neutrino charged current interaction immediately hadronizes into any of the charmed mesons or baryons. Except for D^0, D^+, D_s^+, and Λ_c, all of the charmed hadrons decay via the strong interaction to one of these four, for instance $D^{*0} \rightarrow D^0\pi^0$ or $\Sigma_c^{++} \rightarrow \Lambda_c\pi^+$. In turn, the D^0, D^+, D_s^+, or Λ_c produced at the primary vertex will travel a finite distance in the emulsion before decaying via the weak interaction. Experimentally, the strongly decaying charm hadrons have no effect other than altering the production fractions of the metastable charm hadrons via their branching fractions. Any secondary muons are produced at the weak decay vertex. If the efficiencies to select decays for any of the four charmed hadrons were equal, and independent of the presence of a muon at the secondary vertex, then their absolute level would not affect in any way the semi-leptonic branching ratio. The measurement consists simply of counting the number of decays with and without a secondary muon. Instead, if the selection efficiency depends on the type of hadron, then the composition of the selected sample will be different from that of all charmed hadrons produced. Since the composition determines the weight to be used in averaging the semi-leptonic branching fraction of each hadron to obtain the average semi-leptonic branching fraction of charmed hadrons produced in neutrino charged current interactions, any distortion of the composition leads to a distortion of the semi-leptonic branching fraction.

Since the selection is entirely based on the identification of a secondary vertex in the emulsion, it depends solely on the topology of the event, essentially the flight length of the charmed hadron and the number of charged tracks at the decay. Hence, the primary requirement on the simulation is for it to correctly reproduce the flight length distribution and the secondary vertex multiplicity.

For this comparison, three samples will be used: the actual data, the simulated data, and the background estimation. The selection of events for scanning and the scanback procedure are not explicitly simulated. However, the following cuts are applied, both on the data and on the simulation:

- primary vertex not reconstructed in the most upstream plate,
- primary vertex not reconstructed in the four most downstream plates,
- muon momentum smaller than 30 GeV/c,
- muon slope with respect to the horizontal axis smaller than 400 mrad, and
- measured energy in the calorimeter larger than 4 GeV.

The first two criteria serve to avoid any complications that would arise from variations in the net scan fiducial volume depending on the position of the event in the emulsion stack.

The third and fourth are implicitly present in the data, but need to be applied to the simulation sample to account for the fact that only such events are selected for scanning. The final criterion serves to remove, but only to some extent, contributions from quasi-elastic and diffractive processes since these are not present in the sample of simulated data. The effect of these will be further discussed below. After this selection, the data sample contains 1055 selected events of which 244 are manually checked as described in section 5.2. Out of the 244 events, 23 have been identified as background, with secondary interactions and reconstruction artefacts each accounting for about half of the total. For the following, we define

$$N^{selected} = 1055 \ , \tag{5.1}$$

and

$$purity = \frac{221}{244} = 90.6 \pm 1.9\,\% \ . \tag{5.2}$$

In the absence of a background simulation, the measured background events can serve as an approximation of the expected background, after scaling by a factor of 1055/244. The sample of simulated events, which passes the 'tight' cuts of section 5.2 as well as the selection applied here, contains 5356 events. It is normalized to the data by a factor of 90.6 % to account for the selection purity and by a factor of 1055/5356 to account for the relative size of the sample. In all of the following, the error bars on the data are only statistical and use Poisson statistics, even for bins where the small number of entries would dictate the use of binomial statistics. The simulated sample and the background estimation are normalized to the data and given without errors. For the simulation, the statistical error is a factor of two smaller than for the data. For the background estimation, the statistical error is large but the background contribution is not particularly severe in any single region of the parameter space.

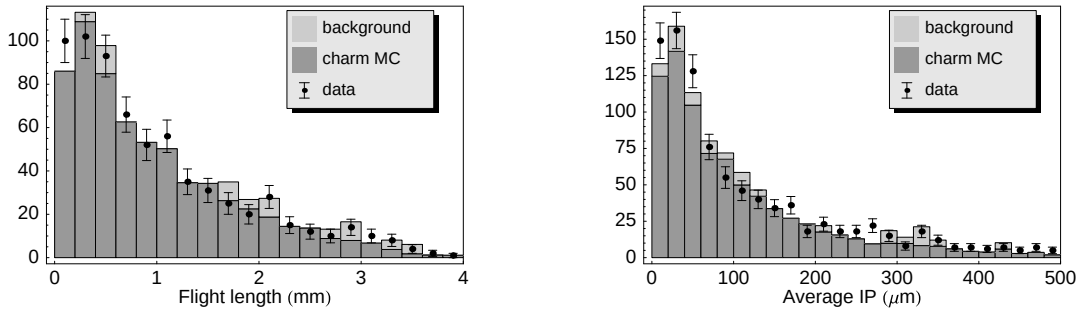
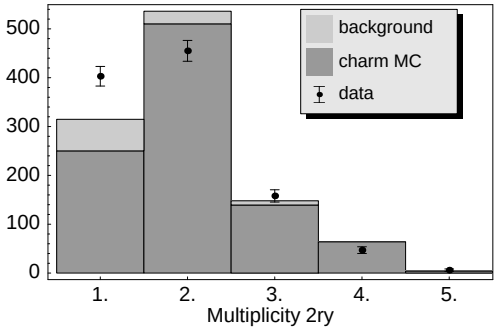


Fig. 5.5: Left panel: flight length distribution for events where it can be calculated using reconstructed quantities. Right panel: the average impact parameter of secondary tracks to the primary vertex for events where the primary track is not isolated. The samples are described in the text.

The first quantity to be compared between the data and the Monte Carlo simulation is the flight length. The distribution is shown in the left panel of figure 5.5. As explained before, it can only be calculated if the primary multiplicity is larger than one and if the secondary is either a vertex or an isolated track with a charged parent. Except for a marginal

excess in data in the first bin, the agreement is good. The right panel of figure 5.5 shows the distribution of the average impact parameter, a better measure of the proper decay length than the geometrical flight length. It can be calculated whenever the primary multiplicity is larger than one, including the case where the secondary is an isolated track. Here, the excess for small values is slightly more pronounced, but the overall agreement remains satisfactory. A possible explanation for the excess would be that the simulation underestimates the selection efficiency for events with a small average impact parameter and correspondingly small proper decay length. Alternatively, the data might contain larger contributions than predicted by the simulation for the shortest-lived charm hadron Λ_c.



	data	MC + BG
2-prong/1-prong	1.16±0.08	1.71±0.11
2-prong/3-prong	2.88±0.27	3.74±0.20
4-prong/1-prong	0.12±0.02	0.21±0.02
4-prong/3-prong	0.30±0.05	0.45±0.03

Fig. 5.6: Distribution of the reconstructed multiplicity for the secondary. The samples are described in the text, MC + BG is the sum of the charm Monte Carlo and the background estimation.

The second quantity to be compared is the reconstructed multiplicity for the secondary, shown in figure 5.6. Here we note a considerable discrepancy: compared to the data, the simulation predicts too large a ratio of even to odd multiplicity. If the reconstructed multiplicity is a reliable estimate of the true multiplicity, then this would suggest that the data contains a larger contribution from the charged charmed hadron – D^+, D_s^+, and Λ_c as opposed to the neutral D^0 – than predicted by the simulation.

This hypothesis can be tested independently by considering events where the primary and secondary are separated by at least one plate. In those cases, the reconstruction algorithms include an explicit search for the presence of a charged parent. Figure 5.7 compares the flight length distribution for events with and without a reconstructed charged parent. Again, the simulation seems to underestimate the fraction of charged hadrons.

Any possible explanation for the discrepancy in the distribution of reconstructed secondary multiplicity falls in one of the following generic categories.

Event generator. The assumptions about the fragmentation fractions are based on scarce data and could for instance lead to an overestimate of neutral charm hadrons. Similarly, the topological branching ratios for some of the charmed hadrons are not well known. The most striking is the case of the Λ_c, for which the measured decay modes account for a mere 45 % of the total decay width [140].

Track finding efficiency. As the excess of odd multiplicities in the data is most pronounced for single prongs compared to two prongs, any overestimate of the track finding effi-

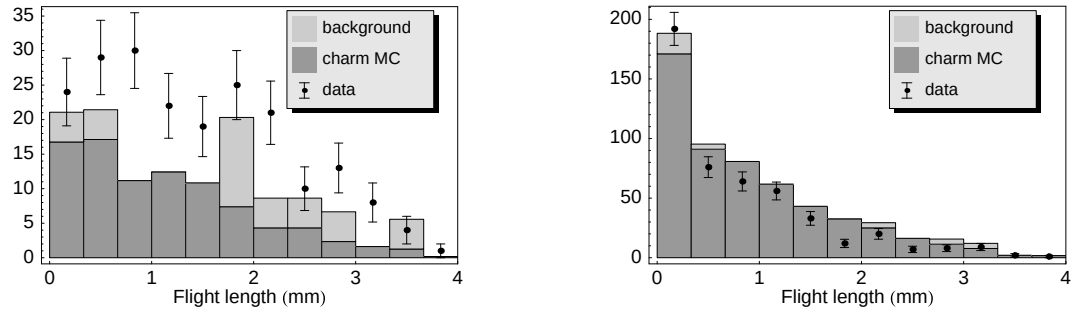


Fig. 5.7: Left panel: flight length distribution for events with a reconstructed charged parent. Right panel: flight length distribution for events where the primary and secondary are separated by at least one plate, but no track is found connecting the two.

ciency in the simulation would have a similar effect, biasing the simulation towards higher reconstructed multiplicity.

Background estimate. The background is concentrated at the lowest secondary multiplicities. Any underestimate of the background would therefore lead to an underestimate in the simulation for the lowest secondary multiplicities.

The background estimate is based on the data itself and does not include the contribution from so-called white star kinks: secondary interactions without any visible nuclear activity. However, these account for no more than 10 % of all secondary interactions [141]. Above, we have seen that secondary interactions with nuclear activity make up only 5 % of the selected events. Hence, the last explanation can be discarded as unlikely. Before discussing a possible source of error at the event generator in section 5.4.4, we will investigate in sections 5.4.1, 5.4.2, and 5.4.3 whether the simulation can be expected to correctly describe the track finding efficiency.

5.4.1 Description of the measurement errors

As described in the previous chapter, the entire selection boils down to error propagation from the segments via tracks and vertices to the distance between two vertices, between a track and a vertex, or between two tracks. This relies critically on the description of the segment measurement errors, most succinctly summarized in the distribution of the track and vertex fit probabilities. Figure 5.8 shows this distribution for the primary: on the left, the χ^2 probability of the vertex fit for events where the primary multiplicity is larger than one, on the right, the track fit probability of the track fit for an isolated primary track. In both cases, the agreement between data and Monte Carlo is good, except for a moderate excess in the data at high vertex fit probability. Figure 5.9 shows the same distributions as figure 5.8 but now for the case of the secondary. Here, the excess at large vertex fit probabilities is slightly more pronounced. Furthermore, the simulation predicts a peak at low track fit probabilities which is not present in the data. This discrepancy is not understood but does not by itself compromise the overall agreement between data and simulation. It should be stressed that the track and vertex fit probabilities are particularly sensitive to the errors. To find agreement

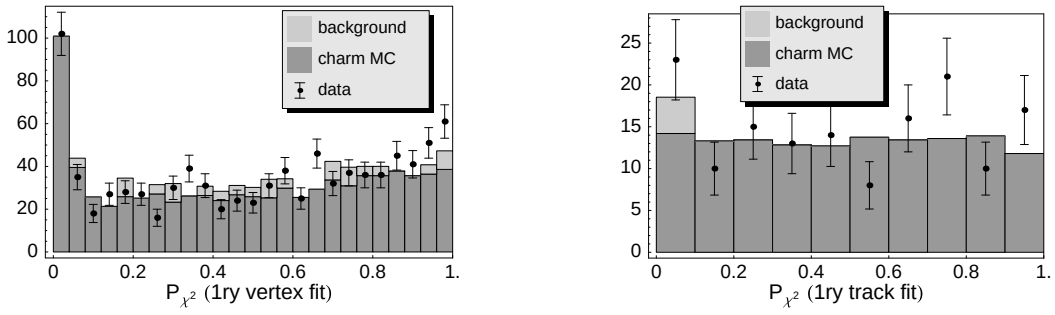


Fig. 5.8: Left panel: distribution of the χ^2 fit probability for the primary vertex in events where there is more than one reconstructed track at the primary. Right panel: distribution of the χ^2 fit probability for the primary track in events where it is isolated. The samples are described in the text.

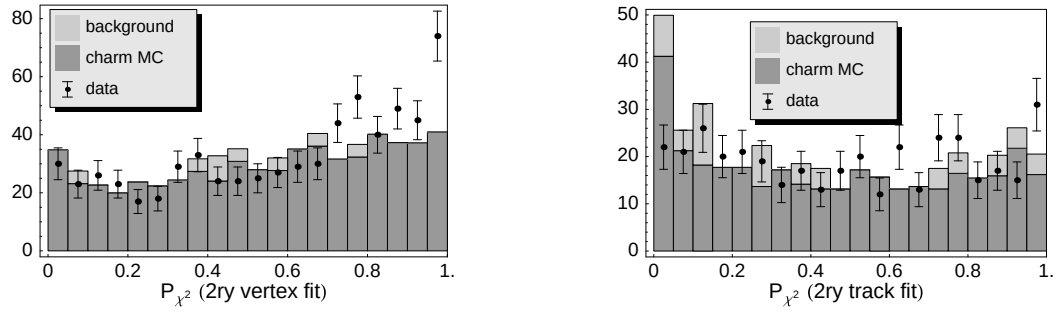


Fig. 5.9: Left panel: distribution of the χ^2 fit probability for the secondary vertex in events where there is more than one reconstructed track at the secondary. Right panel: distribution of the χ^2 fit probability for the secondary track in events where it is isolated. The samples are described in the text.

between data and simulation at this level, both for primary and secondary, strongly suggests that the description of the measurement errors is sufficiently accurate.

5.4.2 Description of the segment finding efficiency

The track finding efficiency is a function of only two parameters: the segment measurement errors and the hardware efficiency to find a segment. For the measurement errors, figures 5.8 and 5.9 can be taken as evidence that the simulation describes the data sufficiently well. For the segment finding efficiency, variations are reflected in the number of reconstructed segments on a track traversing a given number of plates in the net scan fiducial volume. Figure 5.10 shows the number of segments on each track, summed over all tracks attached to a vertex, the primary in the left panel and the secondary in the right panel. The net scan fiducial volume is such that one expects six segments on a primary track in the beam direction. This number is reduced for tracks where one or more segments are lost due to the segment finding efficiency, but also for tracks at large angle exiting from the fiducial volume on the

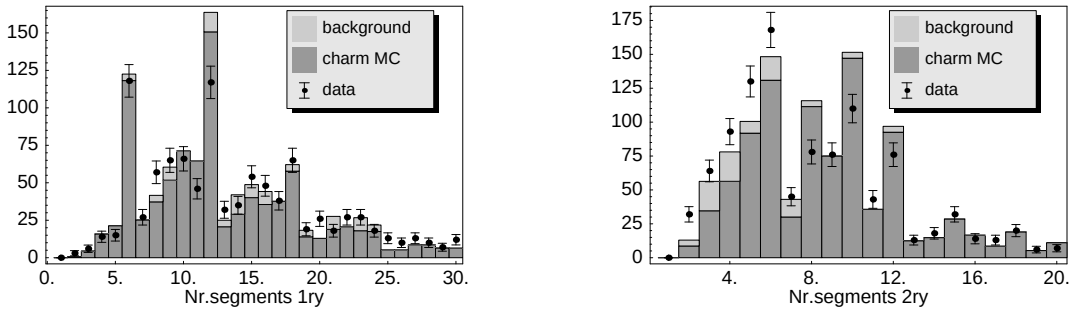


Fig. 5.10: Left panel: sum of the number of segments over all tracks attached to the primary. Right panel: sum of the number of segments over all tracks attached to the secondary. The samples are described in the text.

side and for tracks from a secondary vertex in a plate downstream from the primary vertex plate. For primary vertices with only one or two tracks, each typically has all six segments resulting in the peaks for six or twelve segments. As the multiplicity increases, it becomes increasingly likely for one or more tracks to be at large angle or to have missing segments due to inefficiencies, gradually broadening out the distribution. Nevertheless, for all track multiplicities, the trend in the number of segments is well reproduced by the simulation. The same holds for the secondary vertex as well, but in that case is obviously affected by the excess and deficit in the number of one prong and two prong secondaries, respectively. We conclude that the segment finding efficiency is well described in the simulation and hence expect the track finding efficiency to be correctly reproduced as well.

5.4.3 Migration of multiplicity between the true and reconstructed value

Instead of ascertaining the simulation of the two parameters which determine the track finding efficiency, it is possible to study the track finding efficiency directly in the data. For that purpose, we use a sample of 393 events for which the decay topology has been confirmed manually. For these events, the real multiplicity is known as the error on a manual count of the number of secondary tracks is negligible. Comparing the true and reconstructed multiplicity on an event by event basis allows a determination of the migration based on the data itself. For simulated data, the true multiplicity is known and again the migration between true and reconstructed multiplicity can be determined.

Comparing tables 5.6 and 5.7, the migration in multiplicity between the true and the reconstructed value is large, but in good agreement between the data and the simulation. Arguably, some discrepancy exists for the migration at true multiplicities of four or more. Even though this hints at an overestimate of the track reconstruction efficiency in the simulation, the effect is too small to significantly affect the migration at low multiplicities.

Considering the agreement between data and simulation for the track and vertex fit probabilities, for the number of segments on tracks, and for the migration of the multiplicity from the true to the reconstructed value, it seems unlikely that the excess for odd decay multiplicities would be due to an inadequate description of the emulsion in the simulation. Instead, it leads one to believe that the event generator does not offer an adequate description of the processes under study. This conclusion is further strengthened by the large discrepancy in

True mult.	Fraction	Reconstructed multiplicity				
		1	2	3	4	5
1	19.1±2.0%	92±3%	5±3%	3±2%		
2	39.4±2.5%	24±3%	73±4%	3±1%		
3	27.5±2.3%	18±4%	44±5%	39±5%		
4	12.5±1.7%	4±3%	25±6%	43±7%	27±6%	2±2%
5	1.3±0.6%		20±18%	40±22%	40±22%	

Tab. 5.6: Migration in the secondary multiplicity between the true value, as obtained from manual scanning, and the reconstructed value. The first two columns show the true multiplicity and the corresponding fraction of the total. The remaining columns show, for each true multiplicity, the fraction with a given reconstructed multiplicity. All errors are statistical only, the sample contains 393 manually checked events.

True mult.	Fraction	Reconstructed multiplicity				
		1	2	3	4	5
1	6.7±0.4%	94.9±1.2%	3.5±1.0%	1.0±0.6%		
2	49.1±0.7%	24.7±0.9%	74.4±0.9%	0.5±0.2%		
3	18.8±0.6%	20.1±1.4%	42.4±1.7%	37.1±1.6%		
4	14.6±0.5%	4.6±0.8%	20.0±1.5%	37.6±1.9%	37.9±1.2%	
5	1.7±0.2%	8.8±3.2%	15.0±4.0%	32.5±5.2%	25.0±4.9%	18.8±4.4%

Tab. 5.7: Migration in the secondary multiplicity between the true value, as given in the simulation, and the reconstructed value. The first two columns show the true multiplicity and the corresponding fraction of the total. The remaining columns show, for each true multiplicity, the fraction with a given reconstructed multiplicity. All errors are statistical only, the sample contains 4643 simulated events which have been selected.

the second column between tables 5.6 and 5.7. The excess of single prong decays is clearly present in the data from manual scanning and cannot in this case be explained by artefacts of the reconstruction. Of course, it might still be that the simulation somehow underestimates the selection efficiency for single prong decays. We will come back to this point at the end of the next section.

5.4.4 Contribution of quasi-elastic and diffractive processes

The simulation includes only the deep-inelastic production of charmed particles, for which the cross section corresponds to about 4 % of the neutrino charged current cross section. However, other sources of charm production should not be neglected, in particular quasi-elastic and diffractive processes. The quasi-elastic process produces only the charmed baryons Λ_c^+, Σ_c^+, Σ_c^{++}, and excited states of these. All the charmed baryons decay via the strong interaction to Λ_c^+. The diffractive process can lead to either D_s^+ or D^+, as well as to the respective excited states. However, the diffractive production of D^+ is Cabibbo-suppressed. As a consequence, the quasi-elastic and diffractive processes can be expected to increase the fraction

of Λ_c and D_s^+ in the total, both of them charged and with shorter lifetimes than D^+.

Only one measurement exists of the neutrino quasi-elastic charm production cross section, based on three events in the E531 experiment [54]. Compared to the charged current cross section, it corresponds to a rate of

$$\frac{\sigma^{QE}_{\Lambda_c}}{\sigma_{CC}} = 0.3^{+0.3}_{-0.2}\,\% \ . \tag{5.3}$$

This is in good agreement with the average value of various theoretical predictions [142, 143, 144] using different approaches,

$$\frac{\sigma^{QE}_{\Lambda_c}}{\sigma_{CC}} = 0.3\,\% \ , \tag{5.4}$$

even though it should be noted that this excludes those calculations predicting a yield one order of magnitude higher [145, 146, 147, 148]. In view of a study of the Λ_c^+ absolute branching ratios in the CHORUS experiment [149], the theoretical description of the different models has been incorporated in a dedicated event generator for quasi-elastic Λ_c production, called QEGEN.

The neutrino diffractive charm production has been studied in several experiments. Six candidates for either D_s or D_s^* with subsequent decay into $\phi\pi^+$ from the combined data of the WA21, WA25, WA59, and E180 bubble chamber experiments [150] lead to a diffractive production rate of

$$\frac{\sigma^{diff}_{D_s/D_s^*}}{\sigma_{CC}} = 0.29 \pm 0.10\,\% \ . \tag{5.5}$$

More recently, the NuTeV experiment at Fermilab measured a rate of

$$\frac{\sigma^{diff}_{D_s/D_s^*}}{\sigma_{CC}} = 0.32 \pm 0.06\,\% \ , \tag{5.6}$$

based on the analysis of dimuon events with low hadronic energy [151]. Finally, the CHORUS observation [87] of a single event compatible with diffractive D_s^{*+} production and subsequent decay $D_s^{*+} \to D_s^+ \to \tau^+ \to \mu^+$ translates into an upper limit for the production cross section of [152]

$$\frac{\sigma^{diff}_{D_s^*}}{\sigma_{CC}} \leq 0.46\,\% \text{ at } 90\% CL \ . \tag{5.7}$$

The two theoretical models [153, 154] which correctly describe the experimental data have both been incorporated in a dedicated event generator for diffractive D_s and D_s^* production, called ASTRA [155].

Assuming that the cross sections for both diffractive and quasi-elastic charm production are each about 0.3 % of the charged current cross section, then the sum corresponds to 15 % of the deep-inelastic charm production process. It is not our aim here to measure the contribution of these processes, but merely to understand at least qualitatively whether they could account for the discrepancies seen in the flight length and secondary multiplicity

distributions. Compared to the deep-inelastic production of charmed particles, the quasi-elastic and diffractive production are both characterized by relatively lower hadronic energy and smaller emission angles for the primary muon with respect to the neutrino direction. Figure 5.11 compares the distribution of these two quantities in the data and the simulation including only deep-inelastic processes. In both cases, the data show an excess in the region where quasi-elastic and diffractive charm production are expected, as seen in figure 5.12, based on a sample of events produced by the ASTRA and QEGEN event generators. This simulation uses the CHORUS beam spectrum, and the same simulation of the electronic detector response as used for the simulation of deep-inelastic production. However, it is worth noting that the discrepancy seen in the left panel of figure 5.11 could also be explained by a known problem with the CHORUS calorimeter simulation. For all types of events, the simulated distribution of calorimeter energy is systematically biased towards higher energies than observed in the data [156].

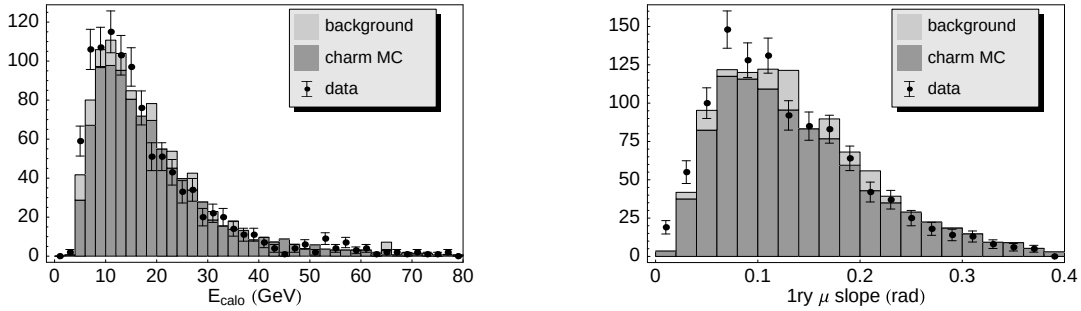


Fig. 5.11: Left panel: distribution of the energy measured in the calorimeter. Right panel: distribution of the slope of the primary muon with respect to the neutrino direction. The samples are described in the text.

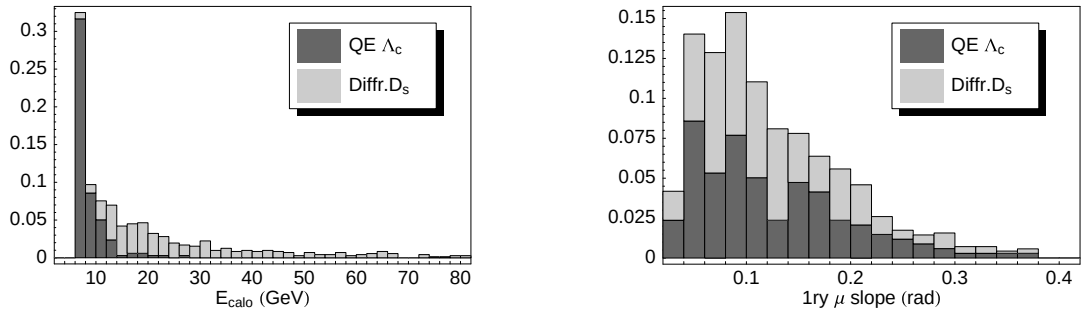


Fig. 5.12: Left panel: distribution of the energy measured in the calorimeter for a sample of events simulated according to quasi-elastic charm production and to diffractive charm production. Right panel: distribution of the slope of the primary muon with respect to the neutrino direction for the same sample.

However, the events have not been processed with the emulsion simulation and reconstruction and hence have not been corrected for any biases introduced by the charm selec-

tion. To estimate the effect of including these processes, we modified the fragmentation fractions in the simulated sample of deep-inelastic charm production, increasing the contribution of D_s and Λ_c and correspondingly decreasing the fraction of D^0 and D^+. It turns out that the distributions of reconstructed multiplicity at the secondary vertex as well as average impact parameter or flight length are essentially unaltered. The reason is clear from figure 5.13: D^0 decays account for more than half of the events with a reconstructed secondary multiplicity of one and of the events with a small average impact parameter. This is a result of the combined effect of a large D^0 fragmentation fraction, significant migration in the reconstructed multiplicity from two to one, and higher selection efficiency for two prong decays compared to one prong decays. Even if the contributions of D_s and Λ_c are doubled, the net effect on the distributions of figure 5.13 is marginal since the D^0 and D^+ contributions are scaled down to readjust the overall normalization.

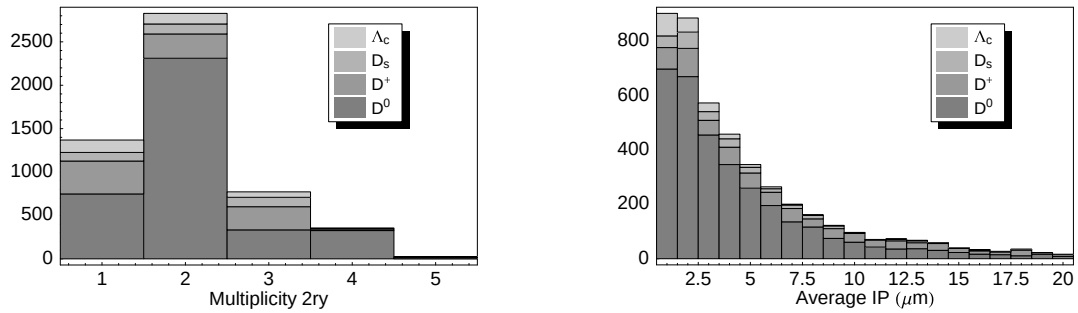


Fig. 5.13: Left panel: reconstructed multiplicity at the secondary vertex for the selected events in the simulated sample. Right panel: average impact parameter for selected events in the simulated sample where it can be calculated.

In summary, it is important to consistently include the diffractive and quasi-elastic processes in the Monte Carlo description of these data, but they cannot explain all of the discrepancies observed. However, for the present analysis, the simulation affects the final result only through the selection efficiencies which will be discussed in the next section. Furthermore, the result is only weakly dependent on the relative efficiencies. For instance, as will be shown in section 5.7, a change by a factor of two in the selection efficiency for single prong decays leads to a ten percent change in the measured value for the semi-leptonic branching fraction, comparable to the statistical error or to the systematic error related to the description of the muon identification. Given the generally good agreement between data and simulation, a factor of two difference in the single prong efficiency can be considered an extreme assumption. Figure 5.14 shows the effect of such a change, combined with an enhancement of D_s and Λ_c, on the distributions for the reconstructed secondary multiplicity and the average impact parameter. These are to be compared with figures 5.5 and 5.6.

At the end of the previous section, we pointed out that the excess of single prong decays is clearly present for the manually determined multiplicity as well, in that case unaffected by reconstruction artefacts even though still implicitly relying on the charm selection. As pointed out in the context of tables 5.6 and 5.7, the manually determined multiplicity exhibits a clear excess of single prong decays, just as the reconstructed multiplicity. This is shown graphically in the left panel of figure 5.15, indicating the contribution of the four

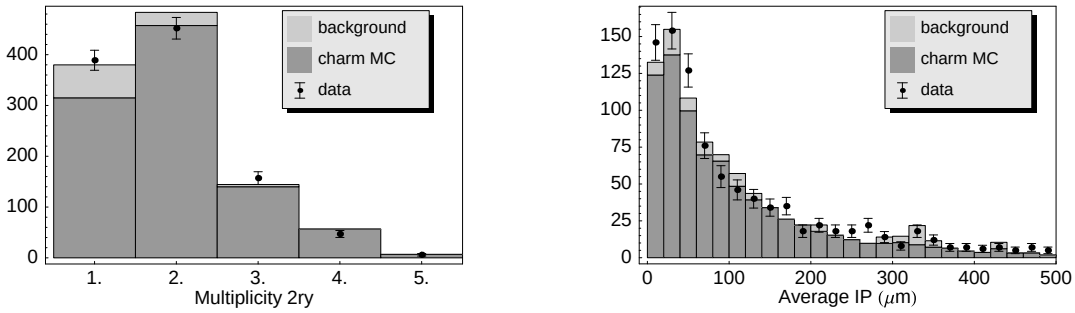


Fig. 5.14: Left panel: reconstructed multiplicity at the secondary vertex for the selected events in the simulated sample. Right panel: average impact parameter for selected events in the simulated sample where it can be calculated. In the simulated sample, the D_s and Λ_c contribution as well as the single prong selection efficiency have been enhanced as described in the text.

metastable charmed hadrons. The discrepancy is too large to be explained solely by the additional production of D_s and Λ_c through the diffractive and quasi-elastic processes, respectively. However, if on top of that, one assumes that the single prong selection efficiency is underestimated by a factor of two in the simulation, then the agreement becomes satisfactory, as seen in the right panel of figure 5.15.

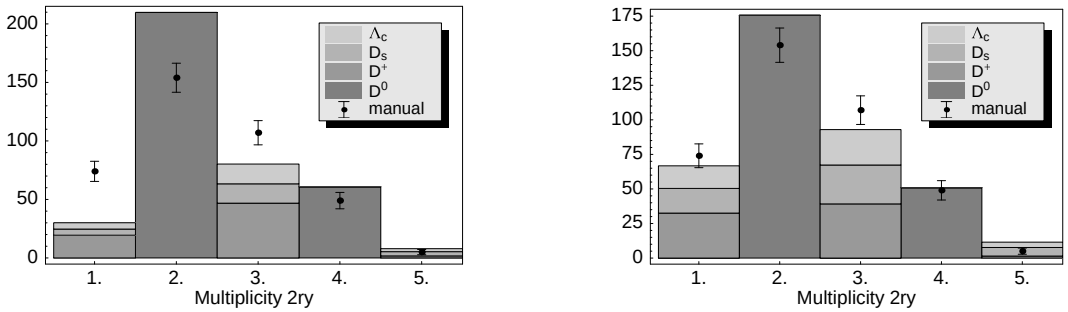


Fig. 5.15: True multiplicity at the secondary vertex. The points with error bars, statistical only, refer to the sample of manually checked events where the multiplicity is measured manually. The stacked histogram refers to the selected events in the simulated sample, where the multiplicity is taken at the event generator level. Left panel: the unweighted set of simulated events. Right panel: the set of simulated events, reweighted to enhance the D_s and Λ_c contributions and to increase the selection efficiency for single prong decays.

5.5 The charm selection efficiency

The sample of charm events is defined by the selection. Out of these, a subsample is identified as dimuon events using the muon identification described in the next section. Then,

the ratio of the number of events in these two samples is approximately equal to the semi-leptonic branching ratio, the number of charm quarks produced in charged current neutrino interactions which after fragmentation and charm hadron decay give rise to a second muon in the final state. Here, the approximation is in the implicit assumption that the selection efficiency is independent of the charm hadron species and independent of whether or not the decay is muonic. The correction for this effect requires knowledge of eight efficiencies, one for each of the four weakly decaying charmed hadron species and one for each of those four decaying muonically. Before getting to this, we will briefly review the efficiency as a function of different variables, primarily to ascertain that the behaviour can be qualitatively understood.

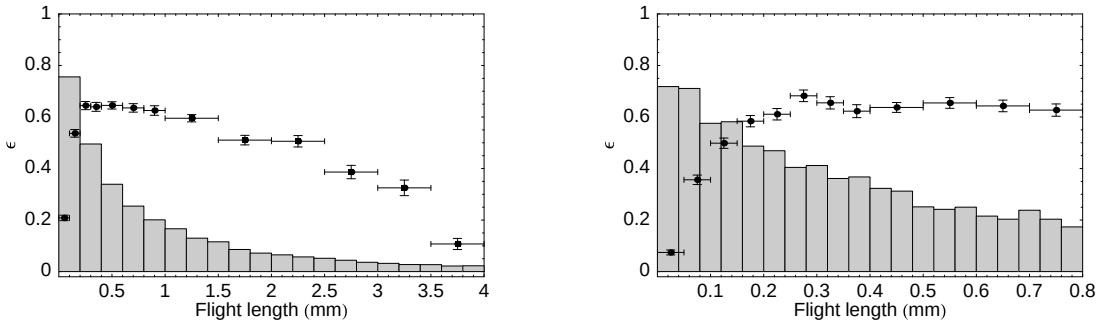


Fig. 5.16: The dots with error bars indicate the charm selection efficiency, determined from the simulation, as a function of the flight length of the charmed hadron. The vertical errors are statistical only. The histograms show the flight length distribution before selection.

Figure 5.16 shows the selection efficiency as a function of the charm hadron flight length, superimposed on the flight length distribution for the simulated sample. The efficiency of a selection based entirely on the geometric characteristics of the event is obviously zero unless the charm hadron travels at least some distance. For short flight lengths, the selection efficiency slowly rises until it reaches 65 % for flight lengths of 250 μm or more. At approximately 1.5 mm, the thickness of two plates, the efficiency drops by 10 % because isolated tracks without a charged parent are not accepted if there are more than two plates between their most upstream segment and the primary vertex. Finally, the efficiency starts dropping further at 2.5 mm as the decay point gets to lie so close to the edge of the fiducial volume, either longitudinally or sideways, that no secondary tracks can any longer be reconstructed. As a matter of fact, for the simulated sample, 19 % of the D^+, 5 % of the D^0, and 6 % of the D_s decay outside of the fiducial volume altogether. This does not occur for any of the 986 Λ_c events in the simulation.

The left panel of figure 5.17 shows the selection efficiency as a function of the neutrino energy. The variation is mostly due to the correlation between the neutrino energy and the momentum of the charmed hadron, with a very strong variation of the efficiency as a function of the latter as shown in the right panel. In turn, this is partly due to the correlation between the flight length and the charm momentum. On top of that, there is an efficiency penalty for small charm momenta due to the large angle of both the charm particle and its decay products, leading to smaller track reconstruction efficiencies. Conversely, for large momenta the decay angles become too small to be detected.

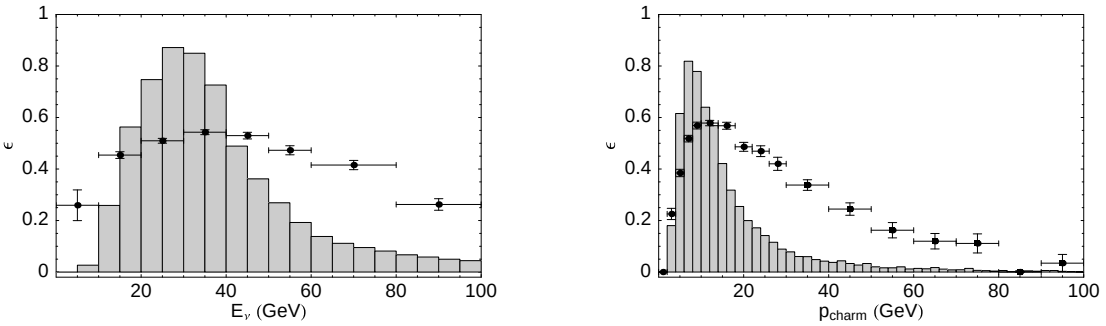


Fig. 5.17: The dots with error bars indicate the charm selection efficiency, determined from the simulation, as a function of the true neutrino energy in the left panel and as a function of the true momentum of the charmed hadron in the right panel. The vertical errors are statistical only. The histograms show the neutrino energy and charm hadron momentum distributions before selection.

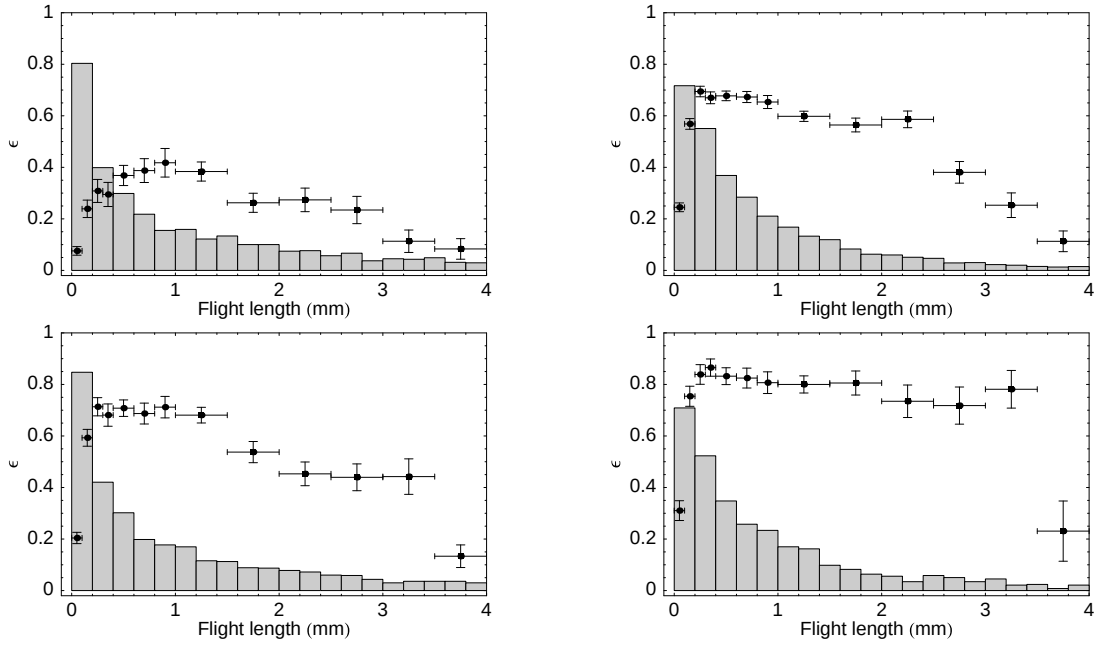


Fig. 5.18: The dots with error bars indicate the charm selection efficiency, determined from the simulation, as a function of the geometrical flight length. Left: for charged charm particles, decaying into a single charged particle (above) or into three charged particles (below). Right: for neutral charm particles, decaying into two charged particles (above) or four charged particles (below). The vertical errors are statistical only. The histograms show the flight distribution before selection for the respective samples.

Except for the charm flight length, the largest effect on the selection efficiency stems from the number of charged tracks at the decay vertex. The most extreme example are the 4 % of

D^0 which decay into neutral particles. For these, the selection efficiency is obviously zero. As the selection requires the reconstruction of at least one of the charged decay particles, the efficiency is expected to rise as the number of charged decay products increases. Figure 5.18 shows the selection efficiency as a function of flight length for a true number of charged particles at the secondary vertex of one, two, three and four. Singling out the region where the flight length lies between 500 μm and 1000 μm, the average selection efficiency is 39±3 %, 67±1 %, 70±3 %, or 82±2 % for one-prong, two-prong, three-prong, and four-prong decays, respectively.

True mult.	Selection efficiency
1	29.7±1.1%
2	59.9±0.6%
3	57.3±1.0%
4	75.7±1.2%
5	65.5±3.4%
6	83±15%

Tab. 5.8: The charm selection efficiency, determined from the simulation, as a function of the true charged multiplicity at the secondary vertex. The errors are statistical only.

The left-hand side of the figure corresponds to charged charm particles for which the drop in efficiency at about two plates is clearly visible. For neutral charm particles, on the right, there is obviously no such effect. Table 5.8 shows the average efficiency as a function of the number of charged tracks at the secondary vertex, integrated over the flight length distribution. Except for the general trend of increasing efficiency as the multiplicity increases, the efficiency is higher for neutral charm particles than for charged charm particles. This is simply related to the different flight length distribution. For charged charm, the short lifetimes of D_s and Λ_c lead to a concentration at flight lengths of less than 250 μm while the relatively long lifetime of D^+ leads to the loss of a sizeable fraction due to the limited scanning volume.

This brings us to the selection efficiencies for each of the four weakly decaying charm hadron species, shown in table 5.9. Compared to D^+, the enhancement of D^0 has been discussed above. For Λ_c, the efficiency is slightly lower because of the short flight length. For D_s, it is higher largely because of the 11 % five-prong decays, compared to 2 % for D^+. The rightmost column shows the selection efficiency for each of the four species, but restricted to semi-leptonic decays into a muon. In essence, the correction to the semi-leptonic branching ratio for charm produced in neutrino charged current interactions will be driven by the differences between the second and third column in table 5.9. By and large, such differences are small. Nevertheless, it should be kept in mind that all of this is based exclusively on simulation and that there exists a significant disagreement between data and simulation for the distribution of the reconstructed multiplicity at the secondary vertex. A possible cause discussed above is an underestimate of the selection efficiency for single-prong decays. An increase in the single-prong efficiency by a factor of two leads to the efficiencies per species indicated between parentheses.

Species	Selection efficiency			
	$D_i \to any$		$D_i \to \mu X$	
D^+	36.9±1.0%	(47.5%)	41±2%	(58%)
D^0	54.1±0.6%		50±2%	
D_s^+	44.0±1.8%	(53.2%)	45±6%	(53%)
Λ_c	33.5±1.5%	(40.6%)	29±7%	(53%)

Tab. 5.9: The charm selection efficiency, determined from the simulation, for each of the four metastable charmed hadrons which are produced. The third column indicates the selection efficiency for the case where the charmed hadron decays semi-leptonically into a muon. The errors are statistical only. The numbers between parentheses are based on the assumption that the single-prong efficiency is underestimated by a factor of two in the simulation.

5.6 Selection of muonic charm decays

Muons are identified in the downstream detectors, the calorimeter and the muon spectrometer. Any charged particle reaching the muon spectrometer after having traversed 5.2 interaction lengths in the calorimeter, can safely be assumed to be a muon. However, this requirement leads to a loss in angular acceptance and to a threshold for the muon momentum between 1.6 GeV/c and 2.5 GeV/c, depending on the track angle. An attempt is made to identify also muons exiting sideways from the calorimeter or stopping in the calorimeter based on the presence of a minimum ionizing particle in the calorimeter. There exist two independent algorithms for this purpose. The first, CLUNET, is based on a neural network and performs the overall calorimeter reconstruction. The second, CATRAS, is a track finding algorithm using the information of the streamer tube planes that are inserted between calorimeter modules, as described in section 2.3.3. Each of the two is applied independently for the two projections, leading to a total of four possible tags for muons in the calorimeter. Finally, an overall muon flag is defined as

$$\mu \text{ flag } = 5 * SPEC + CLUNET_y + CLUNET_z + CATRAS_y + CATRAS_z \ , \qquad (5.8)$$

where $SPEC$, $CLUNET_i$, and $CATRAS_i$ are one for muons identified in the muon spectrometer, by the CLUNET algorithm in projection i, and by the CATRAS algorithm in projection i, respectively, and zero otherwise. This yields a number between zero and nine, with higher values indicating a more reliable identification and values of five or more for muons identified in the muon spectrometer. Because of the background for the calorimeter algorithms, the minimum requirement is for a muon flag of three or more.

Figure 5.19 compares the distribution of this muon flag for the data and the simulation, using the same samples as in section 5.4. If more than one muon is present in the event, the one for which the muon flag is higher is defined as the primary one, or the one with the higher momentum if two have the same flag. The disagreement between data and Monte Carlo simulation for a muon flag of eight or nine indicates that the simulation overestimates the efficiency of the CATRAS and CLUNET algorithms. However, the sum of the last two bins is described well. This is also true for the same quantity for the secondary muon in events where one is present, as shown in the right panel of figure 5.19. Because of the lower

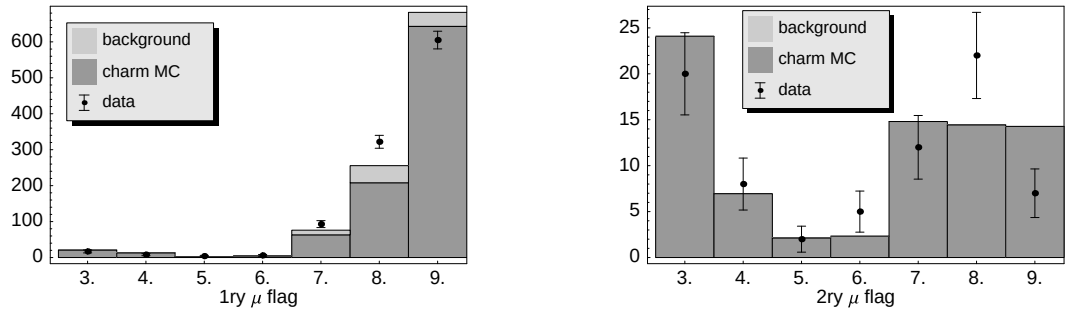


Fig. 5.19: Distribution of the muon flag defined in equation 5.8. Left panel: the primary muon, as defined in the text. Right panel: the secondary muon if more than one is present in the event. The samples are described in the text. No background event was found among the sample of manually checked dimuon events.

momentum spectrum for the secondary muon, the calorimeter muons with flag values three or four play a much larger role than in the case of the primary muon. Finally, it is worth noting that none of the 26 background events contain a second identified muon.

To tag the muonic charm decays among the sample of events selected for showing a decay topology, we rely on the presence of a second muon in the event. There are several reasons for not explicitly requiring the second muon to be matched to a secondary track in the emulsion reconstruction.

- The muon identification is based on the downstream detectors, whereas the matching between the emulsion and the electronic detector relies on the target tracker system. Hence, explicitly requiring an emulsion track to be matched to a detector track depends on the matching between the target tracker and the downstream detectors, with considerable uncertainty arising from the extrapolation of the track parameters through the calorimeter.

- Based on the manual checks, the purity of the charm selection is estimated to be more than 90 %. On the other hand, the assignment of secondary tracks to the secondary vertex has not been optimized and a genuine charm event would still be selected even if only one of the secondary tracks is identified as such.

The Monte Carlo sample of 5356 selected events contains 500 events where the charm hadron decays semi-leptonically into a muon. In 272 out of these 500 events, a second muon is identified with a μ flag > 3, corresponding to a muon identification efficiency of 272/500 = 54 %. However, the requirement that three out of four calorimeter muon identification algorithms give a positive result is relatively loose and leads to an additional 171 selected events being wrongly identified as dimuons. Hence, the muon identification can be said to have a purity of 272/443 = 61 %, defined within the sample of events selected as charm. Phrased differently, the fraction of single muon events wrongly identified as dimuon events is 171/4856 = 3.5 %. The purity can of course be increased by requiring a higher minimum muon flag for identification, but only at the expense of a lower efficiency, as seen in the table of figure 5.20. The reason for the low efficiency is immediately obvious when comparing the momentum spectrum of secondary muons from charm decay with the muon

μ flag	purity	efficiency	data	corrected
3	61±2%	54±2%	76	86±10
4	71±3%	43±2%	56	90±13
5	70±3%	38±2%	48	89±14
6	73±3%	37±2%	46	89±14
7	74±3%	36±2%	41	84±14
8	82±3%	26±2%	29	90±18
9	86±4%	14±2%	7	44±17

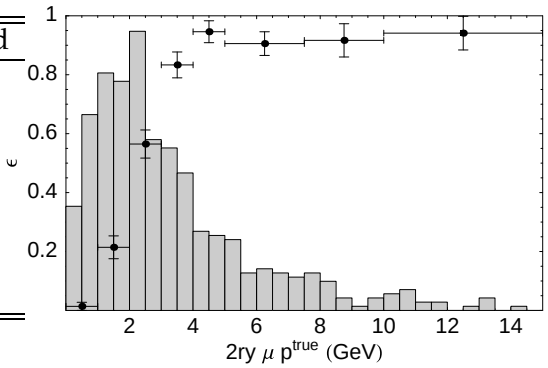


Fig. 5.20: The purity and efficiency of the secondary muon selection determined from simulation as a function of the minimum value of the muon flag required for identification. The errors are statistical only. This leads to a correction on the number of events with a secondary muon: from 'data' to 'corrected'. The figure to the right shows the efficiency of the muon identification as a function of the true muon momentum, with the histogram indicating the spectrum of secondary muon momenta in selected events for the simulation.

identification efficiency as a function of muon momentum, both shown in the right panel of figure 5.20. Here, the efficiency is shown based on a muon identification requiring μ flag ≥ 3.

Starting from the observed number of events identified as dimuons, the real number of dimuon events can be estimated by multiplying with the purity and dividing by the efficiency. The table of figure 5.20 shows that the estimated number of dimuon events in the data is stable with respect to changes in the definition for the muon identification for minimum flags from three to eight. As could be expected from the discrepancy observed in the right panel of figure 5.19, the estimate does differ significantly when a muon flag of nine is required for identification. The average corrected number for a minimum flag from three to eight is

$$N_{2\mu}^{selected} \;=\; 88 \;\pm\; 10\,(stat.) \;\pm\; 8\,(syst.) \;\;, \tag{5.9}$$

where the statistical error is based on the largest sample, *i.e.* requiring a muon flag of three or more. The systematic error is estimated from the variation in the corrected number as the minimum muon flag for identification is varied between three and eight.

For illustration purposes, figure 5.21 shows the distributions of the average impact parameter and of the reconstructed multiplicity at the secondary vertex for the sample of selected events classified as dimuons. These are entirely analogous to the right panel of figure 5.5 and the left panel of figure 5.6, respectively. As far as any conclusions can be drawn from only 76 events, there is no indication to suspect systematic differences between the selected events classified as dimuons and the overall sample of selected events.

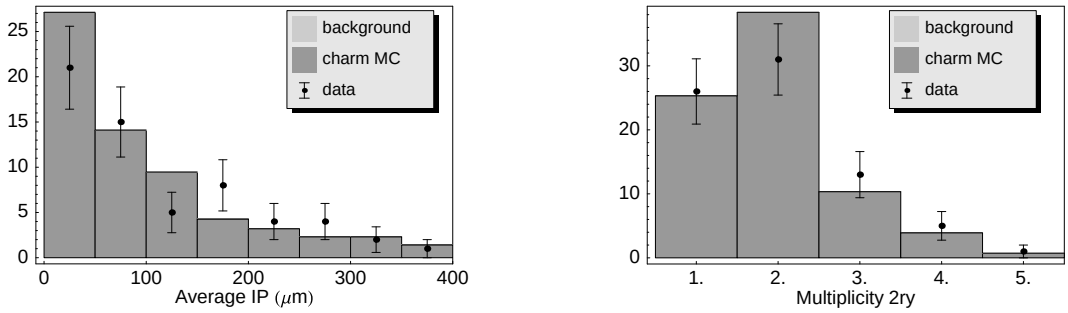


Fig. 5.21: Left panel: distribution for the average impact parameter for selected events classified as dimuons where the primary vertex position can be reconstructed. Right panel: distribution of the reconstructed multiplicity at the primary vertex for selected events classified as dimuons. The samples are described in section 5.4 and the muon identification requires a muon flag of three or more. No background event was found among the sample of manually checked dimuon events.

5.7 Determination of the semi-leptonic branching ratio

We now have all the ingredients to determine the semi-leptonic branching ratio for charm quarks produced in neutrino charged current interactions, defined as

$$B_\mu \;=\; \sum_{D_i} f_{D_i}\, BR(D_i \to \mu X) \;\;, \tag{5.10}$$

where the sum runs over the four metastable charmed hadron species which are produced $D_i = D^+, D^0, D_s$, and Λ_c. The f_{D_i} are the corresponding charm fragmentation fractions. First of all, the overall normalization is given by the number of selected events from equation 5.1 corrected for the selection purity of equation 5.2:

$$N^{selected} \;=\; 956 \,\pm\, 35 \;\;, \tag{5.11}$$

where the statistical error is based on the propagation of the errors on the number of manually checked background events, the number of manually confirmed events, and the number of selected events, taking into account the respective correlations. Combined with the number of selected events where the charmed particle decays semi-leptonically into a muon, estimated in equation 5.9, this yields the fraction of selected events for which the decay is semi-leptonic:

$$\frac{N_{2\mu}^{selected}}{N^{selected}} \;=\; 9.2 \,\pm\, 0.9\,(stat.) \,\pm\, 0.8\,(syst.)\,\% \;\;, \tag{5.12}$$

where the statistical error is again based on the propagation of the statistical error for the various samples involved with the correlations among them. The systematic error derives from the uncertainty on the Monte Carlo description of the muon identification algorithms.

Different selection efficiencies for the various species or for their semi-leptonic decay modes leads to a correction factor between the fraction of selected events corresponding

to semi-leptonic decays into a muon and the semi-leptonic branching ratio which can be written as

$$R = \frac{\sum_{D_i} \epsilon_{D_i} f_{D_i}}{\sum_{D_i} \epsilon_{D_i^\mu} f_{D_i}} \quad . \tag{5.13}$$

In this expression, the sum runs over the four species, the f_{D_i} are the fragmentation fractions, ϵ_{D_i} the corresponding selection efficiencies, and $\epsilon_{D_i^\mu}$ the selection efficiencies when restricted to semi-leptonic decays into a muon. The selection efficiencies are derived from the simulation and compiled in table 5.9.

Species	JETTA		E531	
			$E_{vis} > 5$ GeV	$E_{vis} > 20$ GeV
D^+	22.5±0.3 %	(19.6 %)	16±4 %	20±5 %
D^0	61.2±0.3 %	(53.2 %)	53±5 %	56±5 %
D_s	7.2±0.2 %	(12.8 %)	13±4 %	11±4 %
Λ_c	9.2±0.2 %	(14.5 %)	17±4 %	11±4 %

Tab. 5.10: The fragmentation fractions for charm quarks produced in charged current neutrino interactions after any strong or electromagnetic decays. The different columns are described in the text.

As for the fragmentation fractions, the experimental knowledge is rather poor. Table 5.10 lists three possible estimates. The column labeled JETTA is based on the sample of simulated events, the error is statistical only. The numbers between parentheses correspond to enhanced contributions of D_s and Λ_c to approximate the effect of including diffractive and quasi-elastic processes, assuming for each of the two a cross section equal to 7.5 % of the deep-inelastic cross section. The statistical error for the corrected fractions is identical to that for the uncorrected fractions. The systematic error is difficult to assess, but amounts to at least 5 % given the uncertainty concerning the cross section of the different charm production processes.

The column labeled E531 is the only direct experimental measurement of these quantities, performed by the E531 emulsion experiment but reanalyzed by T.Bolton [56]. Because of the different energy dependence of the deep-inelastic, diffractive, and quasi-elastic charm production processes, the fractions vary as a function of energy. In the present analysis, there is no explicit minimum for the visible energy. Nevertheless, the 4 GeV minimum for the energy measured in the calorimeter translates into an effective minimum visible energy of about 10 GeV once the primary muon energy has been included. It is interesting to note that the experimental data is better described by the simulation after inclusion of quasi-elastic and diffractive processes, inspite of the rough estimates on which our choice for the relative normalization is based.

To actually measure the fragmentation fractions, the CHORUS experiment does not have the same kinematic capabilities as the E531 experiment but this drawback should be compensated for by the much larger statistics. So far, only the fragmentation fraction for D^0 has been measured [157, 158]. Based on a subsample of 535 manually confirmed charm events, the result is

$$f_{D^0} = \frac{\sigma_{D^0}}{\sigma_{charm}} = 53 \pm 11\,\% \quad . \tag{5.14}$$

A first attempt has been made to fit the fragmentation fractions of the charged charm hadrons [159], but it is not expected to give a competitive result until the complete data set will be available.

The ratio of D^+ to D^0 is the least sensitive to the details of the charm production and fragmentation process and can be qualitatively understood as follows. First, the probabilities to pick up a u or d quark during the fragmentation process are roughly equal. Second, the ratio of vector to pseudoscalar meson production is about three, simply from counting the number of spin states. This leads to relative abundances of D^+:D^0:D^{*+}:D^{*0} = 1:1:3:3. All of the D^{*0} decay via the strong or electromagnetic interaction to D^0, accompanied by either a neutral pion or a photon. On the other hand, only 32 % of the D^{*+} decay strongly or electromagnetically to D^+ with the remainder decaying to $D^0\pi^+$. As a consequence, after the fast decays, one expects relative abundances of D^+:D^0 = 1:3.

Combining the selection efficiencies of table 5.9 with the fragmentation fractions of table 5.10, we can estimate the correction factor R defined in equation 5.13 under a variety of assumptions. The four estimates of the fragmentation fractions – the simulation of deep-inelastic charm production at face value, the simulation after adding further D_s and Λ_c contributions, and the reanalysis of the E531 experimental result with a lower cut on visible energy at 5 GeV or at 20 GeV — lead to values for R of 1.042, 1.043, 1.049, and 1.042, respectively. However, as pointed out before, the disagreement between data and Monte Carlo in the distribution of reconstructed multiplicity at the secondary vertex points towards an additional source of systematic error. The disagreement is reduced considerably by arbitrarily assuming that the single-prong selection efficiency is underestimated by a factor of two. This corresponds to the efficiencies per species given between parentheses in table 5.9. Under this assumption, the correction factors become 0.981, 0.968, 0.969, and 0.978 for the four estimates of the fragmentation fractions. In the absence of any argument in favour or against any one of the sets of assumptions, we will simply use the average value. The crucial point, however, is the relatively mild variation inspite of the large range covered by the different assumptions. Given the generally good agreement between the data and the simulation shown in section 5.4, the range in the correction factor can be considered a conservative estimate for the systematic error on this number. Hence, we find that

$$R \;=\; 1.01 \;\pm\; 0.05 \quad . \tag{5.15}$$

Combining equations 5.12 and 5.15 yields the semi-leptonic branching fraction

$$B_\mu \;=\; \frac{N_{2\mu}^{selected}}{N^{selected}} \;\times\; \frac{\sum_{D_i} \epsilon_{D_i} f_{D_i}}{\sum_{D_i} \epsilon_{D_i^\mu} f_{D_i}} \;=\; 9.3 \;\pm\; 0.9\,(stat.) \;\pm\; 0.9\,(syst.)\;\% \quad , \tag{5.16}$$

where the systematic error includes the uncertainty on the Monte Carlo description of the muon identification, as well as the uncertainty on the correction factor related to the fragmentation fractions and the selection efficiencies for different species and different topologies. However, the second turns out to give a negligible contribution. Likewise, we expect all other sources of systematic error which have not been explicitly estimated to be negligible compared to the present statistical error and the present uncertainty on the muon identification.

5.7.1 Comparison with other experiments

The current work describes the first direct measurement of the semi-leptonic branching ratio for charm produced in charged current neutrino interactions. However, it can be compared to two types of indirect measurements.

The first is based on the analysis of opposite-sign dimuon production in neutrino interactions. This yields a measurement of the product of B_μ with the Cabibbo-Kobayashi-Maskawa matrix element $|V_{cd}|$ squared, as described in section 1.8. This can be transformed into an estimate for B_μ by taking the value for $|V_{cd}|$ derived from the measurement of other CKM matrix elements and the assumption of a unitary CKM matrix. This procedure leads to the estimates shown in table 5.11. In the table, $< E_\nu >$ is the average energy of the neutrino beam by flux. Both the central value and the error of our result are comparable to those for the NOMAD and CHARM II experiments, both of them performed in the WANF neutrino beam. However, the beam setup was slightly modified for the CHORUS and NOMAD experiments compared to the CHARM II experiment. The different fiducial volume in the NOMAD experiment, compared to the CHORUS emulsion target, effectively results in a somewhat different neutrino spectrum. The dependence of the semi-leptonic branching fraction on the neutrino energy will be further discussed in the next subsection.

Experiment	$< E_\nu >$	B_μ
CDHS [45]	20.0 GeV	8.4±1.4 %
NOMAD [49]	23.6 GeV	$9.5\pm0.7^{+1.4}_{-1.3}$ %
CHARM II [47]	23.6 GeV	9.1±0.7±0.7 %
CCFR [50]	140 GeV	10.9±0.8±0.6 %

Tab. 5.11: The semi-leptonic branching ratio based on measurements of opposite-sign dimuon production, assuming unitarity of the CKM matrix.

Decay mode	PDG 2001	JETTA
$D^+ \to lX$	17.2±1.9 %	15.9±0.8 %
$D^0 \to lX$	6.8±0.3 %	8.1±0.3 %
$D_s \to lX$	8.0±6.0 %	2.6±0.3 %
$\Lambda_c \to lX$	4.5±1.7 %	2.0±0.3 %

Tab. 5.12: Measurements of the semi-leptonic branching fractions for metastable charmed hadron species. The column labeled PDG 2001 refers to the 2001 update of the Review of Particle Physics produced by the Particle Data Group [160]. The column labeled JETTA is based on simply counting the number of muonic decays in the simulated sample, and the error is only statistical.

The second existing measurement for B_μ is independent of the theoretical prejudice about unitarity of the CKM matrix, but combines data from different experiments. On the one hand, the sample of 122 neutrino-induced charm events from the E531 experiment is analyzed to yield the charm fragmentation fractions. On the other hand, the semi-leptonic branching ratios for each charm hadron species are extracted from the measurements per-

formed by experiments at e^+e^- colliders. Table 5.12 compares the current knowledge of these branching ratios with the assumptions of our Monte Carlo generator. In each case, lepton universality is assumed to interpret a measured branching fraction for inclusive electronic decays as that for inclusive muonic decays. This procedure yields

$$B_\mu = 8.3 \pm 0.6_{CF} \pm 0.8_{BR}\,\% \quad , \tag{5.17}$$

when using the E531 fragmentation fractions measured for $E_{vis} > 20$ GeV and

$$B_\mu = 7.7 \pm 0.5_{CF} \pm 0.8_{BR}\,\% \quad , \tag{5.18}$$

for $E_{vis} > 5$ GeV. In both cases, the first error is related to the charm fragmentation fractions and the second to the charmed hadron semi-leptonic branching ratios.

5.7.2 The Cabibbo-Kobayashi-Maskawa matrix element $|V_{cd}|$

The values for B_μ in table 5.11 are based on a measurement of $B_\mu|V_{cd}|^2$, substituting for $|V_{cd}|^2$ the value based on unitarity of the CKM matrix. Alternatively, we can substitute our value for the semi-leptonic branching ratio, equation 5.16, to obtain a measurement of the CKM matrix element, independent of unitarity of the CKM matrix. Unfortunately, this is not quite as straightforward as it may seem, due to the energy dependence of the semi-leptonic branching ratio. At energies sufficiently large for threshold effects to be neglected, the cross section for quasi-elastic charm production is independent of the neutrino energy, whereas the cross section for deep-inelastic charm production is proportional to the neutrino energy. As a consequence, the fractional contribution of the quasi-elastic process and therefore the fragmentation fraction of Λ_c is most important at moderate energies. Of the four metastable charmed hadrons, Λ_c has the smallest semi-leptonic branching ratio. Hence, B_μ can be expected to rise as a function of the neutrino energy. The values reported in table 5.11 are in agreement with this prediction.

Given the currently available statistics, the energy dependence of B_μ can as yet be determined only approximately in the CHORUS experiment. The equivalent of equation 5.16 for subsamples corresponding to a range in visible energy lead to the result shown in figure 5.22. The errors shown are only statistical but these are anyhow dominant given that there are only a few dozen selected events with a second muon in each energy bin. Furthermore, the systematic error stems primarily from the muon identification. Hence, it is strongly correlated from bin to bin and does not affect the trend clearly present in the figure.

Due to the energy dependence, the value for B_μ measured in one experiment, with a given neutrino beam and a given acceptance as a function of neutrino energy, cannot be applied directly to determine $|V_{cd}|$ from the measurement of $B_\mu|V_{cd}|^2$ performed by a different experiment, in a different neutrino beam and a different energy dependence for the acceptance. The CCFR experiment has extracted a value for $|V_{cd}|$ by combining its study of opposite-sign dimuon events with a determination of B_μ based on the E531 data, considering only events with a visible energy above 30 GeV [56, 50]. This approach has been adopted by the Particle Data Group [160] as well, even though it is then applied on the average over $B_\mu|V_{cd}|^2$ measurements from different experiments. For the sake of comparison, we have extracted B_μ for the subsample of events with visible energy larger than 30 GeV. This yields

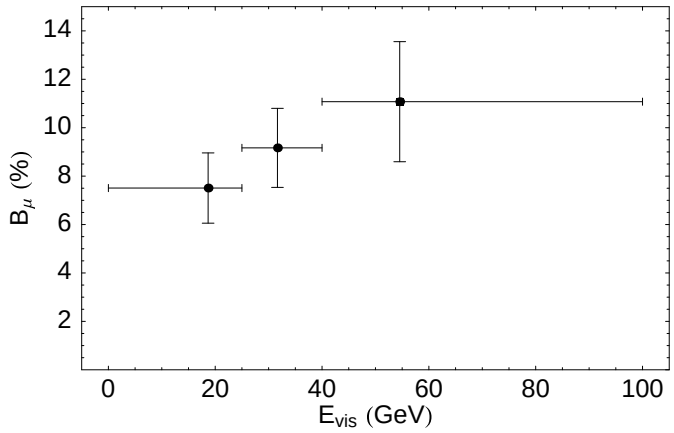


Fig. 5.22: The average semi-leptonic branching fraction as a function of the visible energy. The horizontal bars indicate the width of the energy bin. The vertical bars are drawn at the mean value of the energy in each bin, and their height indicates the statistical error on B_μ.

a value of

$$B_\mu \;=\; 10.2 \;\pm\; 1.6\,(stat.) \;\pm\; 1.0\,(syst.)\,\% \;\; . \tag{5.19}$$

Adding the statistical and systematic errors in quadrature, and combining with the average $B_\mu|V_{cd}|^2$ used by the Particle Data Group,

$$B_\mu|V_{cd}|^2 \;=\; 0.49 \pm 0.05 \times 10^{-2} \quad , \tag{5.20}$$

we find that

$$|V_{cd}| \;=\; 0.219 \;\pm\; 0.022 \;\; , \tag{5.21}$$

to be compared with $|V_{cd}| \;=\; 0.224 \pm 0.016$, the value currently quoted in the Review of Particle Properties and based on the determination of B_μ from E531 data. If unitarity is assumed, the 90 % confidence interval for $|V_{cd}|$ stretches from 0.219 to 0.225.

Given the expected increase in statistics, and improved algorithms for the muon identification, the CHORUS experiment hopes to reduce both the statistical and the systematic error on the measurement of B_μ by a factor of two. In that case, the error on $|V_{cd}|$ will come to be dominated by the extrapolation of the E531 and CHORUS results for B_μ to the region of neutrino energies for which $B_\mu|V_{cd}|$ is measured in the dimuon experiments. Further reducing the error will require the fully inclusive measurement of charm production including the hadronic decay modes, such that the dimuon results and their dependence on B_μ can be avoided altogether. Within the next one or two years, the CHORUS experiment will have collected sufficient net scan data to perform this type of fully inclusive measurement.

CONCLUSIONS AND OUTLOOK

Thanks to the considerable progress in automatic scanning techniques over the past few years, the CHORUS experiment has been able to enlarge the scope of its physics programme. On top of the ongoing search for $\nu_\mu \rightarrow \nu_\tau$ oscillation, the experiment has embarked on a systematic investigation of the production of charmed particles by neutrinos. When the scanning for the CHORUS experiment started in 1994, the most advanced automatic scanning systems were capable of covering a 150×120 μm^2 area in $\sim$5 s, with an angular acceptance of ±30 mrad around the predicted track direction. However impressive this may have seemed at the time, it did impose significant constraints on the physics programme that could be pursued. In essence, each of the analyses was based on the idea of selecting an appropriate track for event location and investigating whether this particular track showed any indication of a decay topology before reaching the vertex. For instance, in events where a muon was identified in the electronic detectors, the scanning procedures were optimized to distinguish the case where the muon track originated from the primary vertex and the case where it was a decay product of a tau lepton. Since 1994, and in no small part driven by the wish to increase the sensitivity to neutrino oscillations in the CHORUS experiment, there has been tremendous progress in the performance of the automatic scanning. Today, the most advanced systems cover a 150×120 μm^2 area in $\sim$0.3 s, picking up all tracks with slopes less than 400 mrad with respect to the direction perpendicular to the emulsion plate. Effectively, this corresponds to an improvement in the speed by more than three orders of magnitude. And yet faster systems are currently being developed.

The above numbers all refer to the microscopes at the Nagoya FKEN laboratory. For the CERN microscopes, the performance is comparable even though the development has followed a slightly different path. First and foremost, the choice was made to implement both grain finding and track finding algorithms in software, in contrast to the in-house development of dedicated hardware. As far as the grain finding is concerned, *i.e.* the pattern recognition applied on individual images, the aim has been to fully exploit the presence of a huge commercial market for processors tailored specifically for image processing. A digital filter was designed to recognize grains in an emulsion image and implemented on a commercially available digital signal processor. As a consequence, only the electronics for the interface between a state of the art CCD camera and a state of the art DSP board had to be developed in-house. Since the first implementation, both the camera and the DSP board have been replaced as the respective technologies evolved, without affecting the rest of the system.

For the track finding, extensive use was made of object-oriented technologies to build algorithms that are not only performant, but also generic and flexible. The core track finding algorithm is independent of any predicted track direction, covering slopes up to 400 mrad. Nevertheless, when a prediction is available, for instance during the vertex location, a fast

but crude track finding algorithm acts as a first selection. This defines a number of regions, which are subsequently treated by the more general core algorithm. An advantage of this approach based on software run on a general purpose processor, is that it effectively decouples the track finding from the data acquisition, ensuring the scalability of the system. Whenever the data rate from the camera – or rather, the rate at which the grain finding algorithm produces data – exceeds the capacity of the track finding algorithm run on the data acquisition PC, the grain data is sent to a farm of computers. This allows the track finding to be distributed over several processors, leaving the data acquisition PC free to collect microscope images at the highest rate allowed by the hardware itself.

To profit as much as possible from this approach based on software rather than hardware, an objective with a large field of view as well as excellent optical properties was developed, in collaboration with industry. At $40\times$ magnification, a single microscope view from this system covers an area of 350×350 μm. Even though the track finding can in principle make use of a prediction for the track direction, the overall system is clearly optimized for the case where a large area is covered at the full angular acceptance of ±400 mrad. Under these conditions, a single 350×350 μm^2 view is treated in $\sim$4 s, comparable to the Nagoya system when expressed per unit of area. For the CERN system, the speed is limited by the 30 Hz frame rate of the CCD camera and the mechanical response of the microscope table. The former is being addressed with the introduction of a CMOS camera which will run at 80 Hz. As far as the latter is concerned, new controls have recently been installed on the microscope table.

At this point, it is worth mentioning that the scanning laboratory at CERN – in addition to its contribution to the development of the automatic scanning technique in general – has played a specific role within the CHORUS collaboration. Simply because of its vicinity to the electronic detector, it has greatly facilitated the exchange between emulsion experts and detector experts, an essential prerequisite for the success of any hybrid experiment. For instance the development of new tracking algorithms, and more specifically the choice of tracks to be used in the vertex location, benefited from the possibility of quickly trying out different possibilities with immediate feedback. In general, care has been taken to keep the analysis software for the data from the CERN scanning as generic as possible. This effort paid off in particular with the advent of the first net scan data from the Nagoya laboratory. Very easily, many of the software components that had been developed at CERN were adapted to treat this type of data, levering the pioneering work performed in Nagoya.

The novel feature of this analysis is the net scan technique. A volume scan in eight consecutive emulsion plates over an area of 1.5×1.5 mm^2 around the vertex position, covering an angular acceptance of 400 mrad, permits the detection of both charged and neutral short-lived particles through the reconstruction of their decay point. The net scan data is sufficiently similar to the data from conventional, electronic tracking detectors for the algorithms developed in that context to be applicable here. In particular, provided the measurement errors on the individual track segments and the effect of multiple Coulomb scattering are correctly described, the entire track and vertex reconstruction can be phrased as a problem of error propagation.

In this work, it has been shown that a rigorous application of these techniques leads to a selection of secondary decay topologies which is both efficient and pure. By manually inspecting the selected events, it was validated that more than 90 % are genuine decays. On

the basis of Monte Carlo simulation, the efficiency for the selection of charmed particles was found to be of the order of 50 %, depending on the flight length of the charmed hadron and on the charged multiplicity at the decay vertex. So far, the net scan data of about 50,000 neutrino interactions have been analyzed using these reconstruction and selection algorithms, resulting in a sample of 1055 candidate events for charm production. For 88 out of the 1055 candidate events, a secondary muon is identified in the electronic detectors, after correcting for the efficiency and purity of the muon identification algorithms. Taking into account the purity of the sample, as well as the relative selection efficiency for charm events with and without a secondary muon, this leads to a direct measurement of the average semi-leptonic branching fraction of

$$B_\mu \;=\; 9.3 \;\pm\; 0.9\,(stat.) \;\pm\; 0.9\,(syst.)\;\%\;\;,$$

for charmed particles produced by neutrino interactions in the CHORUS experiment. To first order, the quantity B_μ fixes the overall normalization for opposite sign dimuon events, such that their analysis can in turn be used to extract the CKM matrix elements $|V_{cd}|$ and $|V_{cs}|$. This is particularly important for $|V_{cd}|$ since neutrino-induced charm production provides the most stringent experimental constraint.

Unfortunately, only under certain conditions can the value of B_μ measured in one experiment be used to interpret the results from an analysis of opposite sign dimuon experiments in a different experiment. More specifically, it should be stressed that the average semi-leptonic branching fraction is not a fundamental parameter but rather an empirical definition of a quantity which depends on the neutrino spectrum. At low neutrino energies, quasi-elastic charm production accounts for a larger fraction of the overall neutrino-induced charm production cross section. In the quasi-elastic process, only Λ_c are produced for which the semileptonic branching fraction is lower than for the D^0, D^+, and D_s mesons produced in deep-inelastic interactions. As a consequence, B_μ is a function rising with neutrino energy. The CHORUS experiment has been the first where this rise could be observed in a single experiment. However, given that the relative contributions from the various processes are only poorly determined, the extrapolation from one energy range to another is affected by large uncertainties.

The present analysis convincingly shows the presence of deep-inelastic and diffractive processes, but does not attempt to quantitatively assess their contribution. Likewise, there are indications that the multiplicity at the secondary vertex is not adequately described in the Monte Carlo simulation, due at least in part to relatively poor knowledge of the branching fractions for some of the charmed hadrons. Nevertheless, neither the uncertainty on the contribution from the different production processes nor the uncertainty on the branching fractions from the different decay modes and associated selection efficiencies have a significant impact on the systematic error on the present measurement of the average semi-leptonic branching fraction. Instead, the dominant systematic error is related to the algorithms for muon identification in the electronic detectors. The overall systematic error and the statistical error are of similar magnitude, both situated at the 10 % level. In the coming year, the CHORUS experiment is expected to collect a sample of up to 4,000 identified charm events, an increase by a factor of four with respect to the current sample. In parallel, work has started on the development of more rigorous muon identification algorithms, with the aim of improving the identification efficiency and of reducing the systematic error on the efficiency by a factor of at least two.

At present, the aspect of the CHORUS charm data that is least understood is the charged multiplicity at the secondary vertex. Several hypotheses have been formulated, but further study is clearly required to understand the origin of the discrepancy between the data and the Monte Carlo simulation. The solution of this problem will almost directly lead to the measurement of the fragmentation fractions for charmed hadrons produced in neutrino interactions, based on a sample more than an order of magnitude larger than the existing measurement from the E531 experiment. Somewhat further into the future, the expected 3,000 to 4,000 charm events in the CHORUS experiment constitute a sample comparable in size to the existing samples of opposite sign dimuon events. As a consequence, it will be possible to determine the ratio of the charm production cross section to the total charged current cross section at similar precision, but unaffected by the average semi-leptonic branching fraction. In spite of the absence of a large sample of antineutrino-induced charm events, such a measurement will very likely lead to the world's most precise determination of the CKM element $|V_{cd}|$. One of the key features of the present analyis is the fact that it does not require manual checks to be performed on the entire data sample. Hopefully, this will allow it to remain a cornerstone of the CHORUS analysis as it unfolds over time.

APPENDIX

A. DIGITAL SIGNAL PROCESSING: A PRIMER

A.1 Definitions in DSP

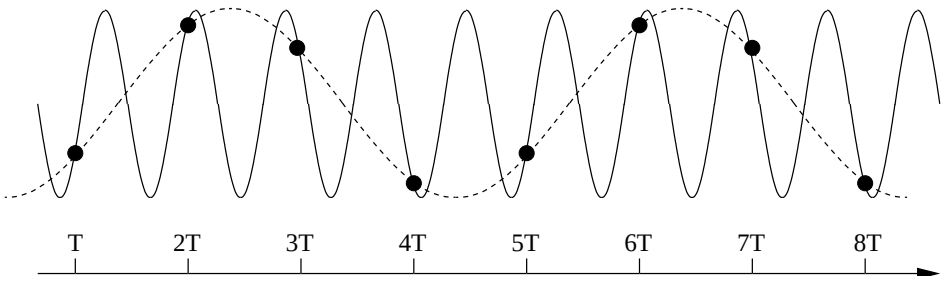


Fig. A.1: An example of aliasing in the time domain. The solid line corresponds to the actual signal, the dashed line to the sampled signal corresponding to instants T, 2T, 3T,... where the sampling frequency is too small.

Quite often, the input to a DSP algorithm is an analogue signal which has been digitized in an analogue-to-digital converter (ADC). This conversion involves sampling of the signal at regular intervals and quantization of the corresponding value. The sampling theorem states that if f_{max} is the highest frequency component of interest, then the signal should be sampled at the rate of at least 2 f_{max} for the samples to describe the signal completely: $F_s \geq 2f_{max}$, where F_s is the sampling frequency. If the sampling frequency is less, then any high frequencies will be folded over or aliased into the desired frequency band, as indicated in figure A.1. This is most intuitive in the frequency domain representation of figure A.2. The spectrum of the sampled signal is the same as that of the original analogue signal, but repeats at multiples of the sampling frequency F_S. The higher order components are referred to as image frequencies. If the sampling frequency F_S is too small, overlap or aliasing will occur about the point F_N, equal to half the sampling frequency. This frequency point is called the folding frequency or Nyquist frequency.

Quantization consists in the assignment of the sampled signal to one of 2^B values, where B is the number of bits used to store the result. This leads to an error of approximately one-half of the least significant bit, assuming that the value is rounded. Any errors due to this effect are referred to as finite word-length effects.

A discrete signal has values defined only at discrete values of time or some other appropriate variable. In our application, the appropriate variable will be the position on the CCD camera line. In general, the amplitude of a discrete-time signal may have discrete values or may be continuous. In our case, the amplitude takes on discrete values as well, correspond-

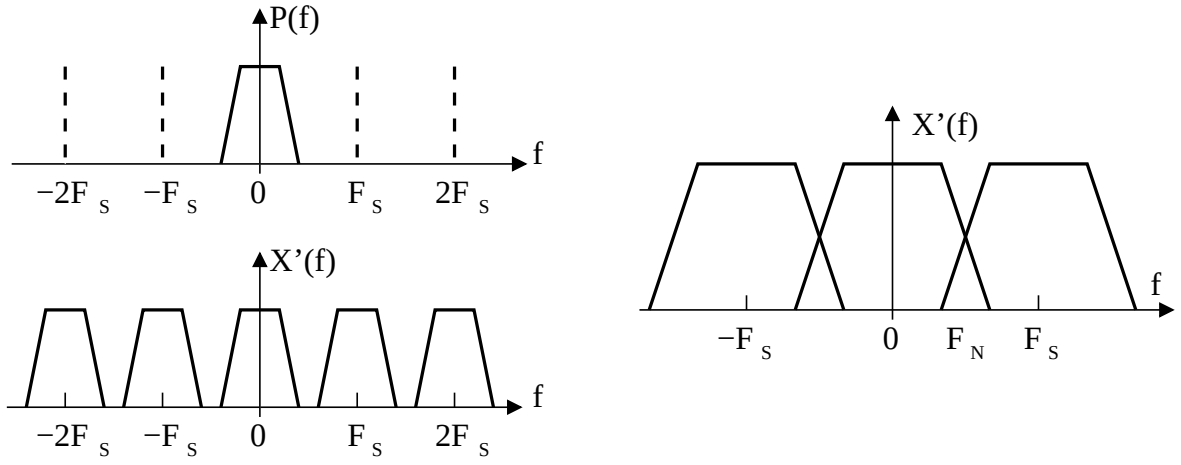


Fig. A.2: Left panel: frequency domain representation of the sampling process. If a signal with frequency spectrum $P(f)$ is sampled at a frequency F_S, the spectrum of the sampled signal $X'(f)$ repeats at multiples of the sampling frequency. Right panel: spectrum of an under-sampled signal, showing the fold-over region due to aliasing. The Nyquist frequency is indicated as F_N.

ing to gray levels of the CCD camera, which has a digital readout. Such a signal is referred to as discrete time, discrete amplitude.

A discrete time system is a mathematical algorithm that takes an input sequence $x(n)$ and produces an output sequence $y(n)$. A system is linear if it obeys the superposition principle, *i.e.* if the response to two or more inputs is equal to the sum of the responses to each input acting separately in the absence of all other inputs, or formally if

$$a_1\, x_1(n) + a_2\, x_2(n) \;\rightarrow\; a_1\, y_1(n) + a_2\, y_2(n) \;\; .$$

A system is time invariant, also called shift invariant, if the output is independent of the time at which the input is applied, or formally if

$$x(n-k) \;\rightarrow\; y(n-k) \;\; .$$

Of particular importance are linear, time invariant systems (LTI); for instance digital filters are of this type. For an LTI system, the relationship between input and output can be expressed as a convolution sum:

$$y(n) \;=\; \sum_{k=-\infty}^{+\infty} h(k)\, x(n-k) \;\; ,$$

where $h(k)$ is called the impulse response of the system. The system is stable if

$$\sum_{k=-\infty}^{+\infty} |h(k)| \;<\; \infty \;\; ,$$

which is satisfied if either $h(k)$ is of finite duration or $h(k)$ tends to zero for large values of k. A system is causal if it produces an output only when there is an input or, equivalently, if $h(k) = 0$ for negative values of k.

A.2 The z-transform

The z-transform originally appeared in probability theory, where de Moivre introduced it in 1730 as the generating function. It is the single most important tool in the analysis and design of digital systems, where it has a role similar to the Laplace transform for analogue systems. For a sequence $x(n)$, valid for all n, the z-transform is defined as the function

$$X(z) \;=\; \sum_{n=-\infty}^{+\infty} x(n)\, z^{-n} \;\;,$$

with a complex argument z. For causal systems, where $x(n)$ is zero for all negative values of n, one uses the one-sided z-transform

$$X(z) \;=\; \sum_{n=0}^{\infty} x(n)\, z^{-n} \;\;.$$

The z-transforms of common sequences are available in closed form, usually given in the form of tables similar to those used for differentiation and integration, as well as Laplace transforms. The poles and zeros of $X(z)$ are the values of z for which $X(z) = \infty$ and $X(z) = 0$, respectively.

The region of convergence (ROC) is the region where the z-transform converges, *i.e.* where $X(z)$ is finite. For a causal sequence of finite duration, the ROC is the entire complex plane except $z = 0$; for a causal sequence of infinite duration, it is the part of the complex plane outside of a circle bounded by the radius of the pole with the largest radius. The ROC for a stable, causal system always encloses the circle of unit radius; as we will see below, this property implies that such systems have a well-defined frequency response. A system with one or more poles on the unit circle, and not coincident with a zero to cancel them, is unstable because of the unavoidable effects of using a finite precision. Typically, the design characteristics of a digital filter are given by its frequency and phase response which can be expressed using its z-transform For an actual implementation in the time domain, it is then necessary to recover the corresponding discrete-time sequence. Formally, we can define the inverse z-transform (IZT) through

$$x(n) \;=\; Z^{-1}[X(z)] \;\;,$$

where $X(z)$ is the z-transform of $x(n)$ and Z^{-1} is the symbol used for the inverse z-transform. If the z-transform for a causal system is expressed as a power series

$$X(z) \;=\; \sum_{n=0}^{\infty} x(n)\, z^{-n} \;=\; x(0) + x(1)\, z^{-1} + x(2)\, z^{-2} + x(3)\, z^{-3} + \mathcal{O}(z^{-4}) \;\;,$$

then the values of $x(n)$ are obtained directly by inspection. In practice, $X(z)$ is more often expressed as a ratio of polynomials of finite order

$$X(z) \;=\; \frac{a_0 + a_1\, z^{-1} + a_2\, z^{-2} + ... + a_N\, z^{-N}}{b_0 + b_1\, z^{-1} + b_2\, z^{-2} + ... + b_M\, z^{-M}} \quad .$$

The methods most commonly applied to obtain the IZT from this form are the power series method, the partial fraction expansion method, and the Cauchy residue method [161]. The power series expansion method essentially corresponds to long division of the polynomials. The partial fraction expansion method expands the z-transform into a sum of simple partial fractions, then taking the IZT for each from tables and summing these to give the overall IZT. Finally, the residue method uses the Cauchy theorem which states that

$$x(n) \;=\; \frac{1}{2\pi i} \oint_C z^{n-1}\, X(z)\, dz$$

is equal to the sum of the residues of $z^{n-1}\, X(z)$ at all poles inside C, the path of integration enclosing all the poles of $X(z)$.

The z-transform has a number of interesting properties.

Linearity. If the z-transforms of $x_1(n)$ and $x_2(n)$ are $X_1(z)$ and $X_2(z)$, respectively, then the z-transform of $a\, x_1(n) + b\, x_2(n)$ is $a\, X_1(z) + b\, X_2(z)$.

Delays or shifts. If the z-transform of $x(n)$ is $X(z)$, then the z-transform of $x(n-m)$ is given by $z^{-m}\, X(z)$.

Differentiation. If the z-transform of $x(n)$ is $X(z)$, then the z-transform of $n\, x(n)$ is given by $-z\, \frac{\mathrm{d}X(z)}{\mathrm{d}z}$.

Convolution. The z-transform of the impulse response $h(k)$ for an LTI system is called the system transfer function $H(z)$. To the input $x(n)$ with z-transform $X(z)$ corresponds an output $y(n) = \sum_{k=\infty}^{+\infty} h(k)\, x(n-k)$ for which the z-transform is given by $Y(z) = H(z)X(z)$. In other words, convolution in the impulse domain corresponds to multiplication in the z-domain. For a system with N zeros and M poles, the system transfer function might be written as

$$H(z) \;=\; K\, \frac{(z-z_1)\,(z-z_2)...(z-z_N)}{(z-p_1)\,(z-p_2)...(z-p_M)} \quad ,$$

where K is the gain factor, z_i the i-th zero and p_i the i-th pole.

Relation to the Laplace transform. If we call $s = d + i\,\omega$ the complex Laplace variable, then the variable for the z-transform is given by $z = e^{sT} = e^{dT}\, e^{i\omega T}$ where T is the sampling interval. As illustrated in figure A.3, frequency points in the s-plane, *i.e.* points for which $d = 0$, are related to points on the z-plane unit circle by $\arg z = \omega T = 2\pi \frac{f}{F_S} = 2\pi \frac{\omega}{\omega_S}$, where F_S and ω_S are the sampling frequency in units of s^{-1} and $rad\, s^{-1}$, respectively. The fact that a single point on the unit circle in the z-plane corresponds to different frequencies, separated by multiples of the sampling frequency, again illustrates the existence of image frequencies discussed above.

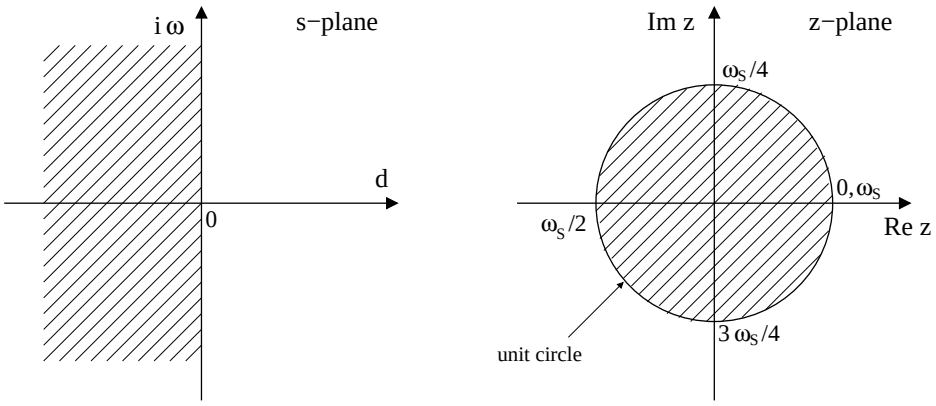


Fig. A.3: The Laplace transform mapping the s-plane to the z-plane. The iω-axis in the s-plane maps onto the unit circle in the z-plane, the left-hand side maps to the interior of the unit circle, as indicated by the shaded area, and the right-hand side maps to the exterior of the unit circle.

The frequency response of a system is given by its Fourier transform which can be found by evaluating the z-transform around the unit circle:

$$H\left(e^{i\omega T}\right) \;=\; \sum_{n=-\infty}^{+\infty} h(n)\, z^{-n}\big|_{z=e^{i\omega T}} \;=\; \sum_{n=-\infty}^{+\infty} h(n)\, e^{-in\omega T} \quad .$$

The frequency response $H(e^{i\omega T})$ is a complex quantity, its modulus corresponding to the magnitude response, its argument to the phase response. If the coefficients a_k, b_k of the transfer function, written as a ratio of polynomials, are all real, then the magnitude response will be symmetrical about the Nyquist frequency $F_N = F_S/2$, the phase response will be anti-symmetrical about the Nyquist frequency, and the frequency response will be periodic with a period of ω_S. The impulse response of a discrete-time system is defined as the IZT of the system transfer function $H(z)$:

$$h(k) \;=\; Z^{-1}\left[H(z)\right] \quad \text{with} \quad k = 0, 1, \ldots \;\; ,$$

corresponding to the response of the system to a unit impulse $u(n)$ with $u(n) = 1$ for $n = 0$ and $u(n) = 0$ for $n \neq 0$.

A.3 The mathematics of digital filters

Digital filters are linear, time-invariant systems characterized by a convolution sum. For causal systems, one distinguishes infinite impulse response (IIR) and finite impulse response (FIR) systems depending on their length:

$$\text{IIR} \;:\quad y(n) \;=\; \sum_{k=0}^{\infty} h(k)\, x(n-k) \;\; ,$$

$$\text{FIR} \; : \quad y(n) \; = \; \sum_{k=0}^{N} h(k)\, x(n-k) \;\; .$$

For the sake of implementation, the impulse response is more conveniently expressed by a recursive relation known as the difference equation:

$$y(n) \; = \; \sum_{k=0}^{N} a_k\, x(n-k) - \sum_{k=1}^{M} b_k\, y(n-k) \;\; ,$$

relating the output sample $y(n)$ to the input sample $x(n)$ through the system coefficients a_k, b_k and the previous outputs $y(n-k)$. It follows that the z-transform of the output is

$$Y(z) \; = \; \sum_{k=0}^{N} a_k\, z^{-k}\, X(z) - \sum_{k=1}^{M} b_k\, z^{-k}\, Y(z) \;\; ,$$

which immediately leads to an expression for the z-domain transfer function:

$$H(z) \; = \; \frac{Y(z)}{X(z)} \; = \; \frac{\sum_{k=0}^{N} a_k\, z^{-k}}{1 + \sum_{k=1}^{M} b_k\, z^{-}k} \;\; .$$

For FIR systems all b_k are equal to zero and the difference equation and system transfer function simplify to

$$y(n) \; = \; \sum_{k=0}^{N} a_k\, x(n-k) \quad \text{and} \quad H(z) \; = \; \frac{Y(z)}{X(z)} \; = \; \sum_{k=0}^{N} a_k\, z^{-k} \;\; ,$$

respectively. In such systems, the impulse response $h(k)$ is given by the coefficients a_k and its length is finite. For IIR systems on the other hand there is at least one $b_k \neq 0$ corresponding to at least one non-zero pole. From the difference equation, it is clear that IIR filters are feedback systems whereas FIR filters are not.

BIBLIOGRAPHY

[1] G. Sterman et al. Handbook of perturbative QCD. *Reviews of Modern Physics*, 67(1):157–248, 1995.

[2] D.J. Griffiths. *Introduction to elementary particles*. John Wiley and Sons Inc, 1987.

[3] C. Itzykson and J.-B. Zuber. *Quantum field theory*. McGraw Hill Inc, 1980.

[4] F. Mandl and G. Shaw. *Quantum field theory*. John Wiley and Sons Ltd, 1984.

[5] C.H. Llewellyn Smith. A lower bound of heavy-lepton production by neutrinos. *Nuclear Physics B*, 56:325–332, 1973.

[6] C.H. Albright and C. Jarlskog. Neutrino production of M^+ and E^+ heavy leptons (I). *Nuclear Physics B*, 84:467–492, 1975.

[7] C.H. Albright, C. Jarlskog, and L. Wolfenstein. Neutrino production of M^+ and E^+ heavy leptons (II). *Nuclear Physics B*, 84:493–502, 1975.

[8] J.D. Bjorken. Asymptotic sum rules at infinite momentum. *Physical Review*, 179(5):1547–1553, 1969.

[9] E.D. Bloom et al. High-energy inelastic $e-p$ scattering at 6° and 10°. *Physical Review Letters*, 23(16):930–934, 1969.

[10] M. Breidenbach et al. Observed behavior of highly inelastic electron-proton scattering. *Physical Review Letters*, 23(16):935–939, 1969.

[11] R.P. Feynman. Very high-energy collisions of hadrons. *Physical Review Letters*, 23(24):1415–1417, 1969.

[12] C.G Callan and D.J. Gross. High-energy electroproduction and the constituents of the electric current. *Physical Review Letters*, 22(4):156–159, 1969.

[13] A. Bodek and J.L. Ritchie. Fermi-motion effects in deep-inelastic lepton scattering from nuclear targets. *Physical Review D*, 23(5):1070–1091, 1981.

[14] A. Bodek and J.L. Ritchie. Further studies of Fermi-motion effects in lepton scattering from nuclear targets. *Physical Review D*, 24(5):1400–1402, 1981.

[15] European Muon Collaboration, J.-J. Aubert, et al. The ratio of the nucleon structure functions F_2^N for iron and deuterium. *Physics Letters B*, 123:275–278, 1983.

[16] M. Arneodo. Nuclear effects in structure functions. *Physics Reports*, 240:301–393, 1994.

[17] D.Y. Bardin and V.A. Dokuchaeva. On the radiative corrections to the neutrino deep inelastic scattering. JINR, Dubna. JINR-E2-86-260-mc, April 1986.

[18] M. Gell-Mann. A schematic model of baryons and mesons. *Physics Letters*, 8(3):214–215, 1964.

[19] G. Zweig. An SU3 model for strong interaction symmetry and its breaking. CERN Preprint TH 401, 1964.

[20] G. Zweig. An SU3 model for strong interaction symmetry and its breaking. CERN Preprint TH 412, 1964.

[21] D.J. Gross and F. Wilczek. Ultraviolet behaviour of non-abelian gauge theories. *Physical Review Letters*, 30(26):1343–1346, 1973.

[22] Y.L. Dokshitzer. Calculation of structure functions of deep-inelastic scattering and e^+e^- annihilation by perturbation theory in quantum chromodynamics. *Soviet Physics JETP*, 46(4):641–653, 1977.

[23] V.N. Gribov and L.N. Lipatov. Deep inelastic ep scattering in perturbation theory. *Soviet Journal of Nuclear Physics*, 15(4):438–450, 1972.

[24] V.N. Gribov and L.N. Lipatov. e^+e^--pair annihilation and deep inelastic ep scattering in perturbation theory. *Soviet Journal of Nuclear Physics*, 15(6):675–684, 1972.

[25] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. *Nuclear Physics B*, 126:298–318, 1977.

[26] E.W.N. Glover, A.D. Martin, R.G. Roberts, and W.J. Stirling. Can partons describe the CDF jet data ? *Physics Letters B*, 381:353–358, 1996.

[27] H.L. Lai et al. Global QCD analysis of parton structure of the nucleon : CTEQ5 parton distributions. hep-ph/9903282, August 1999.

[28] M. Glück, E. Reya, and A. Vogt. Dynamical parton distributions of the proton and small-x physics. *Zeitschrift für Physik C*, 67:433–448, 1995.

[29] H. Plothow-Besch. *PDFLIB: Structure functions and α_s calculations*. CERN-PPE, 1991. User's manual – Version 1.00, W5051, PDFLIB.

[30] H. Plothow-Besch. *PDFLIB: Nucleon, pion and photon parton density functions and α_s calculations*. CERN-PPE, 1997. User's manual – Version 7.09, W5051, PDFLIB.

[31] R.M. Barnett. Evidence for new quarks and new currents. *Physical Review Letters*, 36(20):1163–1166, 1976.

[32] R.M. Barnett. Evidence in neutrino scattering for right-handed currents associated with heavy quarks. *Physical Review D*, 14(1):70–79, 1976.

[33] M.A.G. Aivazis, F.I. Olness, and W. Tung. Leptoproduction of heavy quarks. I. General formalism and kinematics of charged current and neutral current production processes. *Physical Review D*, 50(5):3085–3101, 1994.

[34] M.A.G. Aivazis, J.C. Collins, F.I. Olness, and W. Tung. Leptoproduction of heavy quarks. II. A unified QCD formulation of charged and neutral current processes from fixed-target to collider energies. *Physical Review D*, 50(5):3102–3118, 1994.

[35] O. Nachtmann. Positivity constraints for anomalous dimensions. *Nuclear Physics B*, 63:237–247, 1973.

[36] A. Benvenuti et al. Observation of new-particle production by high-energy neutrinos and antineutrinos. *Physical Review Letters*, 34(7):419–422, 1975.

[37] A. Benvenuti et al. Further observation of dimuon production by neutrinos. *Physical Review Letters*, 35(18):1199–1202, 1975.

[38] J.J. Aubert et al. Experimental observation of a heavy particle J. *Physical Review Letters*, 33:1404–1406, 1974.

[39] J.-E. Augustin et al. Discovery of a narrow resonance in e^+e^- annihilation. *Physical Review Letters*, 33:1406–1408, 1974.

[40] A. Benvenuti et al. Characteristics of dimuons as evidence for a new quantum number. *Physical Review Letters*, 35(18):1203–1206, 1975.

[41] J. D. Bjorken and S. L. Glashow. Elementary particles and SU(4). *Physics Letters*, 11:255–257, 1964.

[42] S.L. Glashow, J. Iliopoulos, and L. Maiani. Weak interactions with lepton-hadron symmetry. *Physical Review D*, 2(7):1285, 1970.

[43] N. Cabibbo. Unitary symmetry and leptonic decays. *Physical Review Letters*, 10(12):531–533, 1963.

[44] M. Kobayashi and T. Maskawa. CP-violation in the renormalizable theory of weak interaction. *Progress of Theoretical Physics*, 49(2):652–657, 1973.

[45] H. Abramowicz et al. Experimental study of opposite-sign dimuons produced in neutrino and antineutrino interactions. *Zeitschrift für Physik C*, 15:19–31, 1982.

[46] CCFR Collaboration, S.A. Rabinowitz, et al. Measurement of the strange sea distribution using neutrino charm production. *Physical Review Letters*, 70(2):134–137, 1993.

[47] CHARMII Collaboration, P. Vilain, et al. Leading-order QCD analysis of neutrino-induced dimuon events. *European Physical Journal C*, 11:19–34, 1999.

[48] V. Lemaitre. *Production inclusive de particules charmées par interaction neutrino-nucléon dans l'expérience CHARM II*. PhD thesis, Université Catholique de Louvain-la-Neuve, 1995.

[49] NOMAD Collaboration, P. Astier, et al. Neutrino production of opposite sign dimuons in the NOMAD experiment. *Physics Letters B*, 486:35–48, 2000.

[50] CCFR Collaboration, A.O. Bazarko, et al. Determination of the strange quark content of the nucleon from a next-to-leading order QCD analysis of neutrino charm production. *Zeitschrift für Physik C*, 65:189–198, 1995.

[51] A.O. Bazarko. *Determination of the strange quark distribution from a next-to-leading order QCD analysis of neutrino and antineutrino production of charm*. PhD thesis, Columbia University, 1994.

[52] K. van der Poel. *Neutrino induced charm production in the CHORUS calorimeter*. PhD thesis, Katholieke Universiteit Nijmegen, 1999.

[53] N. Ushida et al. Experimental details on lifetime measurements of neutrino-produced charmed particles in a tagged emulsion spectrometer. *Nuclear Instruments and Methods*, 224:50–64, 1984.

[54] E531 Collaboration, N. Ushida, et al. Cross sections for neutrino production of charmed particles. *Physics Letters B*, 206(2):375–379, 1988.

[55] E531 Collaboration, N. Ushida, et al. Production characteristics of charmed particles in neutrino interactions. *Physics Letters B*, 206(2):380–384, 1988.

[56] T. Bolton. Determining the CKM parameter V_{cd} from νN charm production. hep-ex/9708014, August 1997.

[57] CHORUS Coll., N. Armenise, et al. A new search for $\nu_\mu - \nu_\tau$ oscillation. CERN-SPSC/90-42.

[58] CHORUS Coll., M. de Jong, et al. A new search for $\nu_\mu - \nu_\tau$ oscillation. CERN-PPE/93-131.

[59] E531 Collaboration, N. Ushida, et al. Limits to $\nu_\mu, \nu_e \rightarrow \nu_\tau$ oscillations and $\nu_\mu, \nu_e \rightarrow \tau^-$ direct coupling. *Physical Review Letters*, 57(23):2897, 1986.

[60] E. Eskut, CHORUS Collaboration, et al. The CHORUS experiment to search for $\nu_\mu \rightarrow \nu_\tau$ oscillation. *Nuclear Instruments and Methods A*, 401:7–44, 1997.

[61] S. van der Meer. Improved collection of secondaries from a long, narrow target by a horn doublet. CERN/PS/AA/80-4, June 1980.

[62] G. Acquistapace et al. The West area neutrino facility for CHORUS and NOMAD experiments: 1994 - 1997 operation. CERN-ECP-95-014.

[63] S. Peraire, M. Ross, and J.M. Zazula. The SPS target station for CHORUS and NOMAD neutrino experiments. Talk given at 5th European Particle Accelerator Conference (EPAC 96), Sitges, Spain, 10-14 June 1996.

[64] B. Van de Vyver. Prompt ν_τ background in wide band nu_μ beams. *Nuclear Instruments and Methods A*, 385:91–99, 1997.

[65] R.G.C. Oldeman. *Measurement of differential cross-sections and structure functions in neutrino and anti-neutrino scattering on lead*. PhD thesis, Universiteit van Amsterdam, 2000.

[66] NA56/SPY Collaboration, G. Ambrosini, et al. Measurement of charged particle production from 450 gev/c protons on beryllium. *European Physics Journal C*, 10:605–627, 1999.

[67] S. Aoki et al. Scintillating fiber trackers with optoelectronic readout for the CHORUS neutrino experiment. *Nuclear Instruments and Methods A*, 344:143, 1994.

[68] CHORUS Collaboration, P. Annis, et al. Performance and calibration of the CHORUS scintillating fiber tracker and opto-electronics readout system. *Nuclear Instruments and Methods A*, 367:367–371, 1995.

[69] Bergsma F. et al. The hexagonal toroidal air-core magnet of the CHORUS detector. *Nuclear Instruments and Methods A*, 357:243–248, 1995.

[70] H. van der Graaf. The honeycomb strip chamber: the application in LHC/SSC experiments and the test results of a prototype. *Nuclear Instruments and Methods A*, 307:220–230, 1991.

[71] J.W.E. Uiterwijk et al. The CHORUS honeycomb tracker and its bitstream electronics. *Nuclear Instruments and Methods A*, 409:682–686, 1998.

[72] CHORUS Collaboration, S. Aoki, et al. The large area emulsion tracker of the CHORUS experiment. CERN-EP-2001-074, submitted to Nuclear Instruments and Methods, October 2001.

[73] J. Dupraz, D. Saltzberg, and G. Van Beek. The optical alignment monitoring system of CHORUS (RASNIK). *Nuclear Instruments and Methods A*, 388:173–179, 1997.

[74] D. Acosta et al. Electron, pion and multiparticle detection with a lead/scintillating-fiber calorimeter. *Nuclear Instruments and Methods A*, 308:481–508, 1991.

[75] S. Buontempo et al. Construction and test of calorimeter modules for the CHORUS experiment, 1994.

[76] E. Di Capua et al. Response to electrons and pions of the calorimeter for the CHORUS experiment. *Nuclear Instruments and Methods A*, 378:221–232, 1996.

[77] CDHS Collaboration, A. Blondel, et al. Electroweak parameters from a high statistics neutrino nucleon scattering experiment. *Zeitschrift für Physik C*, 45:361–379, 1990.

[78] G. Marel et al. Large planar drift chambers. *Nuclear Instruments and Methods*, 141:43–56, 1977.

[79] CHARM II Collaboration, J.P. De Wulf, et al. Test results of the streamer-tube system of the CHARM II neutrino detector. *Nuclear Instruments and Methods A*, 252:443–449, 1986.

[80] M.G. van Beuzekom et al. The trigger system of the CHORUS experiment. *Nuclear Instruments and Methods A*, 427:587–606, 1999.

[81] P.J. Lucas. An object-oriented language system for implementing concurrent, hierarchical, finite-state machines. Master's thesis, University of Illinois at Urbana-Champaign, 1993.

[82] A. Artamonov et al. The data acquisition system of the CHORUS experiment. *Nuclear Instruments and Methods A*, 479:412–438, 2002.

[83] W.H. Barkas. *Nuclear research emulsions. I. Techniques and theory*. Academic Press, 1963.

[84] J.W. Mitchell and N.F. Mott. The nature and formation of the photographic latent image. *The Philosophical Magazine*, 8:1149–1170, 1957.

[85] DONUT Collaboration, K. Kodama, et al. Observation of tau neutrino interactions. *Physics Letters B*, 504:218–224, 2001.

[86] A.J. Apostolakis and J.V. Major. The determination of linear distortion in nuclear emulsion. *British Journal of Applied Physics*, 8:9–11, 1957.

[87] CHORUS Collaboration, P. Annis, et al. Observation of neutrino induced diffractive D_s^{*+} production and subsequent decay $D_s^{*+} \to D_s^+ \to \tau^+ \to \mu^+$. *Physics Letters B*, 435:458–464, 1998.

[88] K. Niu, E. Mikumo, and Y. Maeda. A possible decay in flight of a new type particle. *Progress in Theoretical Physics*, 46:1644–1646, 1971.

[89] B.G. Duff, D.A. Garbutt, and W.T. Toner. The use of nuclear emulsion in association with spark chambers. *Nuclear Instruments and Methods*, 29:351–355, 1964.

[90] E.H.S. Burhop et al. The location and analysis of neutrino interactions in photographic emulsion. *Nuovo Cimento*, 39:1037–1044, 1965.

[91] D. Allasia et al. Investigation of the decay of charmed particles produced in neutrino interactions. *Nuclear Physics B*, 176:13–36, 1980.

[92] S. Aoki et al. A hybrid experiment to search for beauty particles. *Nuclear Instruments and Methods A*, 274:64–78, 1989.

[93] S. Aoki et al. Fully automated emulsion analysis system. *Nuclear Instruments and Methods B*, 51:466–472, 1990.

[94] K. Kodama et al. Hybrid emulsion spectrometer for the detection of hadronically produced heavy flavor states. *Nuclear Instruments and Methods A*, 289:146–167, 1990.

[95] E653 Collaboration, K. Kodama, et al. Charm meson production in 800 gev/c proton-emulsion interactions. *Physics Letters B*, 263:573–578, 1991.

[96] E653 Collaboration, K. Kodama, et al. Measurement of beauty hadron pair production in 600 gev/c π^- emulsion interactions. *Physics Letters B*, 303:359–367, 1993.

[97] K. Niwa. Neutrino oscillations in emulsion experiment. In M. Fukugita and A. Suzuki, editors, *Physics and astrophysics of neutrinos*, pages 520–558. Springer Verlag, 1994.

[98] S. Aoki et al. Nuclear emulsions in a large, hybrid experiment (CHORUS) to search for $\nu_\mu \to \nu_\tau$ oscillations. *Nuclear Instruments and Methods A*, 447:361–376, 2000.

[99] P. Zucchelli. On microscope images, filtering and compression. CHORUS note 97027, February 1998.

[100] H. Beyer and H. Riesenberg, editors. *Handbuch der Mikroskopie*. VEB Verlag Technik Berlin, third edition, 1988.

[101] S. Inoué and R. Oldenbourg. Microscopes. In Bass [162], chapter 17.

[102] E.H. Linfoot and E. Wolf. Diffraction images in systems with an annular aperture. *Proceedings of the Physical Society*, 66:145–149, 1953.

[103] E.C. Ifeachor and B.W. Jervis. *Digital Signal Processing*. Addison-Wesley, 1993, reprinted 1997.

[104] B. Van de Vyver. Scanning at CERN. CHORUS note 97012, June 1997.

[105] T. Nakano. PhD thesis, Nagoya University, 1997.

[106] A. Di Bartolomeo et al. Status of the automatic microscopes at CERN. CHORUS note 98006, June 1998.

[107] K. Winkler et al. Optical microscope for nuclear emulsion readout – system design and results in application. In *Proceedings SPIE 44th annual meeting, Denver, Colorado*, 1999. paper 3779-43.

[108] T.J. Tredwell. Visible array detectors. In Bass [162], chapter 22.

[109] I.M. Papadopoulos. *Detector Techniques for the Application of Topological and Kinematical Criteria for a $\nu_\mu \rightarrow \nu_\tau$ Oscillation Search*. PhD thesis, Physics Department, University of Athens, Greece, 2001.

[110] M. Chizhov. On an implementation of the filter algorithm for the DSPs TMS320C80 and TMS320C62x. CHORUS note 98009, July 1998.

[111] R. Gurin and A. Maslennikov. Controlhost: Package for distributed data handling. Talk given at HEPiX, http://www-hep.fzu.cz/computing/HEPiX/HEPiX95/talks/conhost.ps, June 1995.

[112] J.W.E. Uiterwijk, J. Panman, and B. Van de Vyver. Fast tracking algorithms for large volume emulsion scanning. In preparation.

[113] J.W.E. Uiterwijk. *Title to be defined*. PhD thesis, Technische Universiteit Twente, 2002. In preparation.

[114] R.C. Prim. Shortest connection networks and some generalizations. *Bell System Technical Journal*, pages 1389–1401, November 1957.

[115] E.W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1:269–271, 1959.

[116] S. Aoki. Tracking and vertex reconstruction in modern emulsion experiments. *Nuclear Instruments and Methods A*, 473:192–196, 2001.

[117] M. de Jong. Treatment of multiple Coulomb scattering in CHORAL. CHORUS internal note 99004, June 1999.

[118] R. Frühwirth, M. Regler, R.K. Bock, H. Grote, and D. Notz. *Data analysis techniques for high-energy physics*. Cambridge University Press, second edition, 2000.

[119] R.L. Gluckstern. Uncertainties in track momentum and direction, due to multiple scattering and measurement errors. *Nuclear Instruments and Methods*, 24:381–389, 1963.

[120] H. Eichinger and M. Regler. Review of track-fitting methods in counter experiments. CERN Yellow Report 81-06, 1981.

[121] P. Billoir. Track fitting with multiple scattering: a new method. *Nuclear Instruments and Methods*, 225:352–366, 1984.

[122] D.H. Saxon. Three-dimensional track and vertex fitting in chambers with stereo wires. *Nuclear Instruments and Methods A*, 234:258–266, 1985.

[123] G. Lutz. Optimum track fitting in the presence of multiple scattering. *Nuclear Instruments and Methods A*, 273:349–361, 1988.

[124] V. Karimäki. Effective circle fitting for particle trajectories. *Nuclear Instruments and Methods A*, 305:187–191, 1991.

[125] S. Wolfram. *The Mathematica book*. Cambridge University Press, fourth edition, 1999. Electronically available at http://www.wolfram.com.

[126] G. Cowan. *Statistical data analysis*. Clarendon Press, 1998.

[127] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical recipes in C*. Cambridge University Press, second edition, 1992. Electronically available at http://www.ulib.org/webRoot/Books/Numerical_Recipes.

[128] P. Billoir. Fast and flexible vertex fit. Collège de France, LPC 85-01, 1985.

[129] R. Frühwirth. Application of Kalman filtering to track and vertex fitting. *Nuclear Instruments and Methods A*, 262:444–450, 1987.

[130] V. Karimäki. Effective vertex fitting for particle physics experiments. HIP-1997-77/EXP, 1997.

[131] P. Billoir and S. Qian. Fast vertex fitting with a local parametrization of tracks. *Nuclear Instruments and Methods A*, 311:139–150, 1992.

[132] NOMAD Collaboration, P. Astier, et al. Study of D^{*+} production in ν_μ charged current interactions in the NOMAD experiment. CERN-EP/2001-083, submitted to Physics Letters B, November 2001.

[133] S. Sorrentino. GBEAM, the neutrino beam simulation. CHORUS internal note 98005, May 1998.

[134] R. Oldeman. Measurement of the beam flux from charged-current neutrino interactions in the CHORUS calorimeter. CHORUS internal note 98001, March 1998.

[135] P. Zucchelli. *Calorimetric techniques for the kinematical selection of events in CHORUS*. PhD thesis, Università di Ferrara, 1995.

[136] T. Sjöstrand. High-energy-physics event generation with PYTHIA 5.7 and JETSET 7.4. *Comput.Phys.Commun.*, 82:74–90, 1994.

[137] G. Ingelman. LEPTO version 6.1: the Lund Monte Carlo for deep inelastic lepton-nucleon scattering. In W. Buchmüller and G. Ingelman, editors, *Workshop on Physics at HERA*, pages 1366–1394, May 1992.

[138] S. Ricciardi. *Ricerca di interazioni quasi-elastiche di ν_τ nell'esperimento CHORUS*. PhD thesis, Università di Ferrara, 1996.

[139] *GEANT 3.21*. CERN program library long writeup W5013.

[140] P. Migliozzi. Private communication.

[141] A. Satta. *Analysis of the $\nu_\mu \rightarrow \nu_\tau$ oscillation search in the CHORUS experiment*. PhD thesis, Università di Roma 'La Sapienza', 2001.

[142] J. Finjord and F. Ravndal. Neutrino-production of charmed baryons. *Physics Letters B*, 58:61, 1975.

[143] A. Amer, M.B. Gavela, A. Le Yaouanc, and L. Oliver. Problem of charmed baryons production rate by neutrino. *Physics Letters B*, 81:48, 1979.

[144] S.G. Kovalenko. Quasielastic neutrino production of charmed baryons from the point of view of local duality. *Soviet Journal of Nuclear Physics*, 52:934–936, 1990.

[145] R.E. Shrock and B.W. Lee. Estimates of charm production in exclusive neutrino reactions. *Physical Review D*, 13:2539–2550, 1976.

[146] C. Avilez, T. Kobayashi, and J.G. Körner. Weak production of charmed baryon resonances. *Physics Letters B*, 66:149, 1977.

[147] C. Avilez, T. Kobayashi, and J.G. Körner. Quark-model calculation of charmed-baryon production by neutrinos. *Physical Review D*, 17:709–724, 1978.

[148] C. Avilez and T. Kobayashi. Comment on the neutrino production of the ground state charmed baryons. *Physical Review D*, 19:3448–3449, 1979.

[149] P. Migliozzi, G. D'Ambrosio, G. Miele, and P. Santorelli. A direct evaluation of the Λ_c^+ absolute branching ratios: a new approach. *Physics Letters B*, 462:217–224, 1999.

[150] A. E. Asratian et al. Diffractive production of charmed strange mesons by neutrinos and anti-neutrinos. *Zeitschrift für Physik C*, 58:55–60, 1993.

[151] NuTeV Collaboration, T. Adams, et al. Evidence for diffractive charm production in ν_μ Fe and $\bar{\nu}_\mu$ Fe scattering at the Fermilab Tevatron. *Physical Review D*, 61:092001, 2000.

[152] O. Melzer. *Study of charm production by neutrinos in nuclear emulsion*. PhD thesis, Universiteit van Amsterdam, 2001.

[153] M.K. Gaillard, S.A. Jackson, and D.V. Nanopoulos. Diffractive elastic neutrino-production of vector mesons. *Nuclear Physics B*, 102:326, 1976.

[154] D. Rein and L.M. Sehgal. Coherent pi0 production in neutrino reactions. *Nuclear Physics B*, 223:29, 1983.

[155] O. Melzer. Die neutrinoinduzierte diffraktive produktion von D_s-Mesonen und der Zerfall $D_s \to \tau\nu_\tau$. Master's thesis, Westfälischen Wilhems-Universität Münster, January 1997. CERN-THESIS-98-012.

[156] P. Righini. *Decadimento $\tau \to 3\pi\nu_\tau$ per la ricerca delle oscillazioni $\nu_\mu \to \nu_\tau$ in CHORUS*. PhD thesis, Università di Cagliari, 2001. CERN-THESIS-99-04.

[157] A.M. Güler. *D^0 production rate measurement in neutrino interactions and a limit on muon neutrino to tau neutrino oscillations*. PhD thesis, Middle East Technical University, Ankara, 2000.

[158] CHORUS Collaboration, A. Kayis-Topaksu, et al. Measurement of D^0 production in neutrino charged-current interactions. CERN-EP/2002-005, accepted for publication in Physics Letters B, January 2002.

[159] F.R. Spada. *Neutrino production of charmed particles in the CHORUS experiment*. PhD thesis, Università di Roma 'La Sapienza', 2001.

[160] Particle Data Group, D.E. Groom, et al. Review of particle physics. *European Physical Journal C*, 15:1–878, 2000.

[161] E.I. Jury. *Theory and application of the z-transform method*. Krieger, 1964, reprinted 1973.

[162] M. Bass, editor. *Handbook of optics. Volume II. Devices, measurements and properties*. Optical society of America, second edition, 1994.

Acknowledgements

According to the Oxford Advanced Learner's Dictionary of Current English, to *acknowledge* is to *confess; admit the truth, existence or reality of*, but also to *express thanks for*. Having reached the last page of my thesis, it is my pleasure to indulge in both. First, to confess that none of the preceding is entirely mine, to admit that it was only possible thanks to the efforts of many, too numerous to be done justice with just a couple of lines. Second, to express thanks for the advice and encouragement I have received over the past five years, without which I would never have been in a position to write these lines.

The gentle but firm hand of Prof. Jacques Lemonne has guided me throughout, leaving me sufficient freeway to pursue my own interests while reminding me regularly enough not to be led astray. The fact that Jacques did not himself participate in the CHORUS experiment has undoubtedly been to my advantage: this offered me the opportunity of regular feedback from a safe distance, uncluttered by the mundane concerns of a running experiment.

I am equally indebted to Dr. Pierre Vilain, whose hand was all but gentle. At some point, I invited him to exert more pressure on me but was kindly reminded that this was entirely up to myself. Nevertheless, he somehow did, simply by asking the right questions at the right time, urging me forward to offer all I could muster.

The centerpiece of this thesis concerns the CERN microscope project, for which I spent several years permanently at CERN. During this time, and right until today, Dr. Jaap Panman created the environment in which I felt I could flourish. Yes, it has been tough, and yes the pressure was never abating. But it was also a whole lot of fun.

Looking back, the start of my thesis now seems ages ago and over the years, many people have contributed to this work. When I arrived at CERN, the CHORUS experiment was in the midst of data taking and the people that operated made sure that I immediately felt home. In particular Maarten de Jong, Piero Zucchelli, David Saltzberg, Rui Ferreira, and Jürgen Brunner helped me a lot to get started. Muriel Vander Donckt tried her very best to turn me into an 'expert' of the opto-electronic readout. I doubt whether I ever deserved the label but certainly not for any failings in her explanations.

It was great to find that both Aimo Bülte and Oliver Melzer, with whom I had previously been a summer student at CERN, were also starting a Ph.D. at CERN. The discussions with them, as well as with Rolf Oldeman and Marumi Kado, were certainly among the most stimulating moments.

More or less around the time I started, Jaap Panman had suggested the CERN group set up its own emulsion scanning system, an activity coordinated by Henry Wong. Even though the system was originally intended as a mere copy of the Japanese system, it still took the efforts of several people for more than half a year to get it up and running: Mihail Chizhov adapted the Japanese device drivers for European computers, Beverley Friend set up the CCD camera and frame grabber card as well as the motion control, and Jean-Pierre Dupraz made sure with Daniel Bourillot that all the mechanics behaved properly, building custom pieces whenever we needed to them.

Once the basics were in place, the Nagoya group took an active interest in helping us out with the more sophisticated aspects. Prof. Kimio Niwa encouraged us all along, and

Nakano-san, Kozaki-san, and Toshito-san all spent a couple of days at CERN, each of them contributing to a specific aspect.

After the first year, the scope of the project was considerably enhanced and we decided to develop an entirely new scanning system. Based on the experience we had gained with the *clone* of the Japanese system, one by one all of the components were reviewed and new designs made. The development of new computer-controlled tables and new optics was done by the MICOS and Jenoptik companies, respectively. At all times, the collaboration with their engineers and technical staff has been both interesting and pleasant. At CERN, development mostly centered on the camera readout and the software for the system. Ioannis Papadopoulos joined Beverley Friend and Mihail Chizhov to adapt subsequent generations of digital signal processors for our purposes, obtaining performances that impressed even the suppliers of the hardware. For the software running on the data acquisition computers, Johan Uiterwijk came up with ever more sophisticated algorithms, reaching unprecedented performances with both elegant and robust code. Mathieu Doucet complemented the system with environmental controls, and Helge Meinhard and Maarten Litmaath helped us keep a small farm of computers up and running. Guy Van Beek, together with Jean-Pierre Dupraz and Daniel Bourillot, designed and built the device to contain the immersion oil, a device first suggested by Piero Zucchelli.

As the system matured, it became possible for several people to actually use it for physics analyses, further spurring the development and actively participating in this development. Franca Cassol and Erhan Pesen performed a measurement of white star kinks using emulsions that had been exposed in a test beam, Stefania Ricciardi showed how emulsion plates in a magnetic field could be deployed as a spectrometer, and Oliver Melzer developed the tools for the manual scanning of special events. I also wish to thank the summer students who spent some time with us, and in particular Sergey Kalinin who remained in CHORUS and took a very active role in the development of the microscopes, and is now participating in the charm physics analysis.

After those exciting but demanding years at CERN, it was almost a relief to return to Brussels. Somehow, the department had not changed, but many of the people had and it was a pleasure to not only find the familiar faces of Gaston Wilquet, Laurent Favart, Muriel Vander Donckt, Roel Heremans, and Pascal Vanlaer, but also to get to know Olivier Devroede, Caroline Collard, and Xavier Janssen. I also wish to thank Vincent Lemaître for his encouragement and help with the analysis of charm physics, as well as Rudy Frühwirth for his patience in explaining the finer details of track and vertex fitting in a dense medium.

The Nagoya FKEN laboratory was equally welcoming during the relatively short periods I spent with them. Sato-san, Nakano-san, Komatsu-san, Toshito-san, Naotaka-san, and Narita-san all contributed to making those stays not only interesting and helpful, but also enjoyable. Particularly pleasant were the discussions with Prof. Kimio Niwa, who managed time and again to trigger my imagination and taught me everything I could ever hope to know about nuclear emulsions. If the culture shock was just about manageable, then only thanks to the presence of some other Europeans in the Nagoya laboratory, in particular Francesca Spada and Murat Güler who helped with the analysis as well.

Obviously, I forgot many. But then again, not really. All of you are vividly in my mind, and I know you know that...