Evaluation of the LHCb RICH Detectors and a Measurement of the CKM Angle γ

Jonas Rademacker
Merton College, Oxford
LHCb Experiment, CERN

A thesis submitted for the degree of Doctor of Philosophy
at the University of Oxford

Hilary Term, 2001

Evaluation of the LHCb RICH Detectors and a Measurement of the CKM Angle γ

Jonas Rademacker
Merton College, Oxford
LHCb Experiment, CERN

A thesis submitted for the degree of Doctor of Philosophy
at the University of Oxford

Hilary Term, 2001

Abstract

The future LHCb experiment at the Large Hadron Collider (LHC) at CERN will perform precision measurements of $\mathcal{CP}$ violation parameters in the neutral B systems. In this thesis the performance of the LHCb RICH system is evaluated, and the feasibility of a measurement of the $\mathcal{CP}$ violation parameter γ, using the decays $\mathrm{B_d^0} \to \mathrm{D}^{*-}\pi^+$, $\overline{\mathrm{B_d^0}} \to \mathrm{D}^{*-}\pi^+$ and their $\mathcal{CP}$ conjugates, is investigated.

Efficient methods for the reconstruction of $\mathrm{B_d^0} \to \mathrm{D}^{*-}\pi^+$ decays at LHCb are developed. Using the GEANT–based LHCb detector simulation program SICb, the reconstruction efficiencies and signal to background ratios are estimated. It is demonstrated that the decay channel $\mathrm{B_d^0} \to \mathrm{D}^{*-}\pi^+$ can provide a precision in γ of a few degrees after one year of LHCb data taking.

A full–scale prototype of the LHCb RICH 2 detector has been tested in a testbeam at CERN in Summer 1998. It is shown to work according to expectation, in particular in the defining aspects of a RICH detector, the photon yield and the Cherenkov angle resolution. These results demonstrate that the LHCb RICH detectors are well understood and give confidence in the model of the RICH system used in the simulation studies for LHCb detector optimisation. In a testbeam in Summer 1999 a RICH prototype was tested using a 3×3 cluster of Multi–anode Photo Multiplier Tubes equipped with lenses. The tubes were read out within the LHC bunch–crossing interval of 25 ns for the first time. It is demonstrated that the MaPMT performs well and is a suitable photodetector for the LHCb RICH.

Acknowledgements

Amongst the numerous people who have helped me through the $3\frac{1}{2}$ years of this D.Phil., I would like to mention but a few who I am particularly indebted to:

Guy Wilkinson for his competent, professional and work–intensive supervision. For showing a constant interest in my work, and for providing vital input and support whenever needed. Thanks!

Neville Harnew for providing me with the opportunity to work with the LHCb group at Oxford and at CERN, and for looking after the administrative side of my D.Phil.

John Holt, who looked after me during my first year in the Oxford LHCb group.

Sajan Easo, Brinick Simmons, and David Websdale, who it was a pleasure to work and share an office with.

Everybody who participated in the 1998 and 1999 RICH beam tests, which form the basis of a large part of this thesis.

Jim Libby for proof–reading, and even checking the math in section 5.3.2.

Robert Fleischer for answering all my questions regarding $\mathcal{CP}$ violation, B^0–physics and the Standard Model.

I would like to thank all members of the LHCb Collaboration and the Oxford Particle Physics Group for providing a friendly, stimulating and productive working environment.

Contents

List of Figures

List of Tables

Introduction

The future LHCb experiment is designed to perform high precision $\mathcal{CP}$ violation measurements in the B system. It is one of the four experiments at the LHC collider at CERN, which is under construction and scheduled to start operation in 2006.

The research presented in this thesis is concerned with the following aspects of the LHCb project:

- The LHCb RICH system, which provides particle identification. This is essential for obtaining the clean data samples needed for the LHCb physics programme.
- The feasibility of a measurement of the CKM angle γ at LHCb, using the decay $\mathrm{B}^0_\mathrm{d} \to \mathrm{D}^*\pi$.

The structure of this document is as follows:

Chapter 1 introduces the formalism to describe $\mathcal{CP}$ violation in the B^0_d system and reviews the status of $\mathcal{CP}$ violation in the Standard Model. The motivation for precision measurements of $\mathcal{CP}$ violation parameters in the B^0 system with LHCb is given, in particular for the measurement of the CKM angle γ. The channel $\mathrm{B}^0_\mathrm{d} \to \mathrm{D}^*\pi$ is introduced as a theoretically clean way of obtaining γ.

This is followed by an introduction to the LHCb experiment in **chapter 2**. An overview of the main detector components is given; the LHCb RICH is motivated and described in detail. This chapter summarises the LHCb detector design and draws on the following sources: [TP98], [RIC00], [Mag00], [Cal00], [GAM00].

Chapters 3, 4, and 5 present the research performed by the author.

In **chapter 3**, the feasibility of a γ measurement at LHCb with the channel $\mathrm{B}^0_\mathrm{d} \to \mathrm{D}^*\pi$ is investigated. Efficient reconstruction techniques are developed. From event samples generated with the LHCb detector simulation program SICb, signal yields and purites for the $\mathrm{B}^0_\mathrm{d} \to \mathrm{D}^*\pi$ reconstruction are calculated. The results are used to estimate the precision on γ achievable at LHCb.

In **chapters 4 and 5** data from beam tests of RICH prototypes are analysed. Chapter 4 investigates the performance of a full-scale prototype of the RICH 2 detector. The results presented in chapter 5 focus on the photo detector choice for the LHCb RICH. Results from a test of a RICH prototype equipped with a cluster of Multi-anode Photo Multiplier Tubes, read out with high-speed electronics, are presented.

The conclusion in **chapter 6** summarises the results obtained in the previous chapters.

Chapter 1

CP Violation, γ, and $\mathrm{B_d^0} \rightarrow \mathrm{D}^*\pi$

1.1 Overview

In this chapter, the formalism to describe $\mathcal{CP}$ violation in the $\mathrm{B_d^0}$ system is developed, and the theory of $\mathcal{CP}$ violation in the Standard Model is reviewed. More complete accounts of the Standard Model, $\mathcal{CP}$ Violation and the $\mathrm{B_d^0}$ system can be found in the following publications: [PS97, BS00, HQ98].

Firstly the operations of $\mathcal{C}$, $\mathcal{P}$, and $\mathcal{T}$ are introduced in a the context of a Lagrangian field theory. Then $\mathcal{CP}$ violation in the $\mathrm{B_d^0}$ system, and the observables in measurements of time-dependent decay rate asymmetries are described in a model–independent way in section 1.3.

How $\mathcal{CP}$ violation is accommodated in the Standard Model of particle physics is discussed in section 1.4. The close relationship between $\mathcal{CP}$ violation and the generation of fermion masses in the Standard Model is demonstrated, which is the largest source of free parameters in the Standard Model. It will also be shown that $\mathcal{CP}$ violation measurements are particularly suitable to overconstrain the Standard Model and search for New Physics. The description of $\mathcal{CP}$ violation in terms of unitarity triangles is introduced in section 1.4.9. The current experimental status of the $\mathcal{CP}$ violation in the Standard Model will be reviewed. It will become apparent why $\mathcal{CP}$ violation is expected to be large in the $\mathrm{B^0}$ system. The variables of the model-independent description of the $\mathrm{B_d^0}$ system are related to Standard Model model parameters, and the significance of the parameter γ is discussed.

The channel $\mathrm{B_d^0} \rightarrow \mathrm{D}^*\pi$ as a method of measuring γ is introduced in section 1.5. The significance of this measurement in the context of other γ measurements is shown. The parameters used in chapter 3 are introduced, where the feasibility of a γ measurement with $\mathrm{B_d^0} \rightarrow \mathrm{D}^*\pi$ at LHCb is investigated.

1.2 C, P, and T

Two different kinds of symmetry operations can be distinguished, continuous and discrete symmetries. Examples for continuous symmetry operations are rotation in space and translation in time. That those symmetries, and the associated conservation laws for angular momentum and energy, hold exactly is fundamental to our understanding of physics.

On the other hand, while it seems intuitively appealing that the laws of physics should also be invariant under discrete symmetries like mirror reflection, they are not [LY56], and it turns out that it is possible to construct physical theories that "naturally" accommodate this.

$\mathcal{CP}$ is the combined operation of charge conjugation and parity. Charge conjugation, $\mathcal{C}$, is the exchange of a particle with its antiparticle. Parity, $\mathcal{P}$, is a mirror reflection followed by a 180° rotation.

So far, $\mathcal{CP}$ violation has only been unambiguously observed in the kaon system [CCFT64]. It is anticipated however, that $\mathcal{CP}$ violation effects are much larger in the B^0 system. Clear evidence for this is expected soon from measurements at the B–factories and the Tevatron.

$\mathcal{CP}$ is related to the time reversal operator, $\mathcal{T}$, which reverses the direction of motion, via the combined operation $\mathcal{CPT}$. $\mathcal{CPT}$ is an exact symmetry in any local Lagrangian field theory. $\mathcal{CP}$ violation and $\mathcal{T}$ violation are therefore equivalent.

1.2.1 Field Operators and C, P, T

Modern particle physics is described in the language of quantum field theories. Describing $\mathcal{CP}$ violation in this context requires the transformation properties of the relevant fields under the symmetry operations. Below, the transformation properties of spin 0, spin 1, and spin $\frac{1}{2}$ fields under the operations of $\mathcal{C}$, $\mathcal{P}$ and, for completeness, $\mathcal{T}$, are summarised.

Eigenvalues

All of the symmetry operators are their own inverse:

$$\mathcal{C}^2 = 1, \quad \mathcal{P}^2 = 1, \quad \mathcal{T}^2 = 1 \tag{1.1}$$

Therefore the possible eigenvalues are 1 and -1.

Charge Conjugation

The operation of charge conjugation, denoted by $\mathcal{C}$, replaces all particles with their antiparticles. The field operators of spin 1, spin 0, and spin $\frac{1}{2}$ fields transform under $\mathcal{C}$ as follows:

Spin 1 fields

$$\mathcal{C} A_\mu \left(t, \vec{x}\right) \mathcal{C}^{-1} = -A^\mu \left(t, \vec{x}\right) \tag{1.2}$$

Spin 0 fields

$$\mathcal{C} \phi \left(t, \vec{x}\right) \mathcal{C}^{-1} = \phi^\dagger \left(t, \vec{x}\right) \tag{1.3}$$

Spin $\frac{1}{2}$ fields

$$\mathcal{C} \psi \left(t, \vec{x}\right) \mathcal{C}^{-1} = i\gamma^2\gamma^0\, {}^t\bar{\psi} \left(t, \vec{x}\right) \tag{1.4}$$

where the index ^{t} stands for transposition, and γ^i are the Dirac γ matrices.

Parity

Parity, denoted by $\mathcal{P}$, is the operation of replacing all space coordinates x^a with $-x^a$, which is the effect of a mirror reflection followed by a 180° rotation.

The effect of parity on the field operators is given by:
Spin 1 fields

$$\mathcal{P} A_\mu \left(t, \vec{x}\right) \mathcal{P}^{-1} = A^\mu \left(t, -\vec{x}\right) \tag{1.5}$$

Spin 0 fields

$$\mathcal{P} \phi \left(t, \vec{x}\right) \mathcal{P}^{-1} = \phi \left(t, -\vec{x}\right) \tag{1.6}$$

Spin $\frac{1}{2}$ fields

$$\mathcal{P} \psi \left(t, \vec{x}\right) \mathcal{P}^{-1} = \gamma_0 \psi \left(t, -\vec{x}\right) \tag{1.7}$$

This operation reverses the momentum of a particle, but it does not change angular momentum or spin. It therefore transforms left-handed particles, which have their spins pointing opposite to the direction of motion, into right–handed particles, with the spin pointing into the direction of motion.

In section 1.4.1 the notion of left-handed and right-handed fermions will play an important role. We will use the following notation for left and right-handed fermionic fields:

$$\begin{aligned} \psi_L &= \text{left-handed field} \\ \psi_R &= \text{right-handed field} \end{aligned}$$

The operator $\frac{1}{2}\left(1-\gamma_5\right)$ projects out the left-handed component of a Dirac-field:

$$\psi_L \left(t, \vec{x}\right) \equiv \frac{1}{2}\left(1-\gamma_5\right) \psi \left(t, \vec{x}\right). \tag{1.8}$$

Similarly for the right-handed component:

$$\psi_R \left(t, \vec{x}\right) \equiv \frac{1}{2}\left(1+\gamma_5\right) \psi \left(t, \vec{x}\right). \tag{1.9}$$

Using $\{\gamma_5, \gamma_0\} = 0$:

$$\mathcal{P} \psi_L \left(t, \vec{x}\right) \mathcal{P}^{-1} = \gamma_0 \psi_R \left(t, -\vec{x}\right). \tag{1.10}$$

Time Reversal

Time reversal, denoted by $\mathcal{T}$, transforms $t \rightarrow -t$ and thus reverses the direction of motion. It reverses both the momentum, $\vec{p} \rightarrow -\vec{p}$, and the spin, $\vec{s} \rightarrow -\vec{s}$ and therefore does not change the handedness of a particle. The effect on the field operators is:

Spin 1 fields

$$\mathcal{T} A_\mu \left(t, \vec{x}\right) \mathcal{T}^{-1} = A^\mu \left(-t, \vec{x}\right) \tag{1.11}$$

Spin 0 fields

$$\mathcal{P} \phi \left(t, \vec{x}\right) \mathcal{P}^{-1} = \phi \left(-t, \vec{x}\right) \tag{1.12}$$

Spin $\frac{1}{2}$ fields

$$\mathcal{T} \psi \left(t, \vec{x}\right) \mathcal{T}^{-1} = \gamma^1 \gamma^3 \psi \left(-t, \vec{x}\right) \tag{1.13}$$

1.2.2 CP in a Lagrangian Field Theory

In order to be able to see whether a Lagrangian is invariant under $\mathcal{CP}$ or not, we

Table 1.1: Transformation Properties

	Dirac bilinears (fermionic fields)			
	scalar	pseudoscalar	vector	axial vector
$\mathcal{O}$	$\bar{\psi}_i\psi_j$	$i\bar{\psi}_i\gamma^5\psi_j$	$\bar{\psi}_i\gamma^\mu\psi_j$	$\bar{\psi}_i\gamma^\mu\gamma^5\psi_j$
$(\mathcal{CP})\,\mathcal{O}\,(\mathcal{CP})^{-1}$	$\bar{\psi}_j\psi_i$	$-i\bar{\psi}_j\gamma^5\psi_i$	$-\bar{\psi}_i\gamma_\mu\psi_j$	$-\bar{\psi}_i\gamma_\mu\gamma^5\psi_j$

	Bosons and the derivative operator			
	scalar	pseudoscalar	vector boson	derivative
$\mathcal{O}$	H	A	$W^{\pm\mu}$	∂_μ
$(\mathcal{CP})\,\mathcal{O}\,(\mathcal{CP})^{-1}$	H	$-A$	$-W^{\mp}_{\mu}$	∂^μ

need to know how it transforms under this operation. A list of the transformation properties of the relevant building blocks of the Lagrangian is given in table 1.1 (from [HQ98]).

Taking into account the Lorentz invariance and hermiticity of the Lagrangian, the $\mathcal{CP}$ transformation rules in table 1.1 imply that any allowed term built from a combination of the fields and derivatives listed in table 1.1 transforms to its Hermitian conjugate under $\mathcal{CP}$. Coefficients in front of such terms however do not change. Hence $\mathcal{CP}$ is equivalent to complex–conjugating these coefficients, and therefore any Lagrangian containing complex coefficients may violate $\mathcal{CP}$.

It is of course possible that such complex coefficients are merely physically meaningless phases, in which case they can be removed by re-defining the phases of the fields. The Lagrangian of the electroweak theory is an example where such a re-definition is not possible and one complex phase remains. This is the source of $\mathcal{CP}$ violation in the Standard Model and is treated in more detail in section 1.4.7.

1.3 Model–Independent Description of the Neutral B^0–System

1.3.1 Introduction

When $\mathrm{B^0_d}$ mesons are produced at the LHC, they are created in the hadronisation of quarks with definite flavour:

$$\mathrm{B^0_d} = (\bar{b}, d) \qquad \overline{\mathrm{B^0_d}} = (b, \bar{d})$$

However, due to the electroweak interactions these flavour eigenstates are not the mass–eigenstates of the $\mathrm{B^0_d}$ mesons. As a consequence, the flavour eigenstates $\mathrm{B^0_d}$ and $\overline{\mathrm{B^0_d}}$ mix to form the mass and lifetime eigenstates, $\mathrm{B_H}$, and $\mathrm{B_L}$.

$$|\mathrm{B_{H,L}}\rangle = p|\mathrm{B^0}\rangle \mp q|\overline{\mathrm{B^0}}\rangle,$$

The operation of $\mathcal{CP}$ transforms one flavour eigenstate to the other; it is defined up to an arbitrary phase α:

$$\mathcal{CP}|\mathrm{B}^0\rangle = e^{i\alpha}|\overline{\mathrm{B}^0}\rangle$$

In the absence of $\mathcal{CP}$ violation, the mass eigenstates would would be identical to the eigenstates of the $\mathcal{CP}$ operator. Conversely, if $\mathrm{B_H}$ and $\mathrm{B_L}$ are not $\mathcal{CP}$ eigenstates, and if their is no phase convention in which they are, then there is $\mathcal{CP}$ violation in the $\mathrm{B^0_d}$–$\overline{\mathrm{B^0_d}}$ mixing.

$\mathcal{CP}$ violation in the mixing is in fact believed to be negligible in the $\mathrm{B^0_d}$ system. The relevant type of $\mathcal{CP}$ violation for the scope of this thesis, is $\mathcal{CP}$ violation in the interference between decays with and without mixing.

Below, these ideas will be examined in a more formal manner. The parameters describing $\mathcal{CP}$ violation in the $\mathrm{B^0_d}$ system will be related to observables in time dependent decay rates. The formalism developed here will be used later in chapter 3 to describe $\mathcal{CP}$ violation in the decay $\mathrm{B^0_d} \to \mathrm{D}^*\pi$.

In principle, the following discussion is valid for any system of two distinct, $\mathcal{CP}$ conjugate neutral particles, like the K^0 and the $\mathrm{B^0_s}$ system. However, the assumption made below, that the life–time difference between the mass eigenstates is negligible, is valid for $\mathrm{B^0_d}$ system only.

1.3.2 Evolution of the B^0-$\overline{\mathrm{B}^0}$ system

The Schrödinger Equation for a superposition of flavour eigenstates, $a|\mathrm{B}^0\rangle + b|\overline{\mathrm{B}^0}\rangle$, is:

$$i\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} a \\ b \end{pmatrix} = \mathbf{H}\begin{pmatrix} a \\ b \end{pmatrix}. \tag{1.14}$$

This is the Schrödinger Equation restricted to the $|\mathrm{B}^0\rangle - |\overline{\mathrm{B}^0}\rangle$ subspace of state vectors. The full Schrödinger Equation always has a Hermitian Hamiltonian, leaving the length of the state–vector constant in time. But as the system is allowed to leave the $|\mathrm{B}^0\rangle - |\overline{\mathrm{B}^0}\rangle$ subspace by decaying to other particles, $\mathbf{H}$ in equation 1.14 will not be Hermitian. Any matrix can be expressed as the sum of an Hermitian and an anti-Hermitian matrix:

$$\mathbf{H} = \frac{1}{2}\left(\mathbf{H}+\mathbf{H}^\dagger\right) + \frac{1}{2}\left(\mathbf{H}-\mathbf{H}^\dagger\right) = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma}, \tag{1.15}$$

where the Hermitian matrices $\mathbf{M}$ and $\mathbf{\Gamma}$ are defined by:

$$\mathbf{M} \equiv \frac{1}{2}\left(\mathbf{H}+\mathbf{H}^\dagger\right), \quad \mathbf{\Gamma} \equiv i\left(\mathbf{H}-\mathbf{H}^\dagger\right). \tag{1.16}$$

From CPT invariance:

$$\langle \mathrm{B}^0|\mathbf{H}|\mathrm{B}^0\rangle = \langle \overline{\mathrm{B}^0}|\mathbf{H}|\overline{\mathrm{B}^0}\rangle. \tag{1.17}$$

Therefore the diagonal elements of $\mathbf{H}$ are the same and $\mathbf{H}$ can be written as:

$$\mathbf{H} = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{11} \end{pmatrix}. \tag{1.18}$$

As $\mathbf{M}$ and $\mathbf{\Gamma}$ are Hermitian, their off-diagonal elements are complex conjugates of each other:

$$\mathbf{M} = \begin{pmatrix} m_{11} & m_{12} \\ m_{12}^* & m_{11} \end{pmatrix}, \quad \mathbf{\Gamma} = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma_{11} \end{pmatrix}. \tag{1.19}$$

$\mathbf{H}$ has the following eigenvalues:

$$\lambda_{H,L} = h_{11} \pm \sqrt{h_{21} \cdot h_{12}} \tag{1.20}$$

and eigenvectors

$$\vec{v}_{H,L} \equiv \begin{pmatrix} p \\ \mp q \end{pmatrix} \tag{1.21}$$

with

$$\frac{q}{p} = -\sqrt{\frac{h_{21}}{h_{12}}}. \tag{1.22}$$

The diagonalised Hamiltonian $\mathbf{H_d}$ is:

$$\begin{aligned} \mathbf{H_d} &= \begin{pmatrix} H_H & 0 \\ 0 & H_L \end{pmatrix} \\ &= \begin{pmatrix} h_{11} + \sqrt{h_{12}h_{21}} & 0 \\ 0 & h_{11} - \sqrt{h_{12}h_{21}} \end{pmatrix}. \end{aligned} \tag{1.23}$$

The diagonal mass and decay matrices, defined by

$$\mathbf{M_d} = \frac{1}{2}\left(\mathbf{H_d} + \mathbf{H_d^\dagger}\right), \quad \mathbf{\Gamma_d} = i\left(\mathbf{H_d} - \mathbf{H_d^\dagger}\right), \tag{1.24}$$

are:

$$\begin{aligned} \mathbf{M_d} &= \begin{pmatrix} M_H & 0 \\ 0 & M_L \end{pmatrix} \\ &= \begin{pmatrix} \mathrm{Re}(H_H) & 0 \\ 0 & \mathrm{Re}(H_L) \end{pmatrix} \end{aligned} \tag{1.25}$$

$$\begin{aligned} \mathbf{\Gamma_d} &= \begin{pmatrix} \Gamma_H & 0 \\ 0 & \Gamma_L \end{pmatrix} \\ &= \begin{pmatrix} -2\mathrm{Im}(H_H) & 0 \\ 0 & -2\mathrm{Im}(H_L) \end{pmatrix}. \end{aligned} \tag{1.26}$$

The subscripts L and H stand for the "light" and the "heavy" physical B^0-states. The mass- and width difference between those states is:

$$\Delta m = M_H - M_L, \quad \Delta\Gamma = \Gamma_H - \Gamma_L \tag{1.27}$$

which are related to m_{ij} and Γ_{ij} via

$$\begin{aligned} \Delta m - \frac{i}{2}\Delta\Gamma &= 2\sqrt{h_{12}h_{21}} \\ &= 2\sqrt{\left(m_{12} - \frac{i}{2}\Gamma_{12}\right)\left(m_{12}^* - \frac{i}{2}\Gamma_{12}^*\right)}, \end{aligned} \tag{1.28}$$

giving:

$$(\Delta m)^2 - \frac{1}{4}(\Delta\Gamma)^2 = 4\left|M_{12}\right|^2 - \left|\Gamma_{12}\right|^2, \tag{1.29}$$

$$\Delta m \Delta\Gamma = 4\mathrm{Re}(M_{12}\Gamma_{12}^*) \tag{1.30}$$

and

$$\begin{aligned} \frac{q}{p} &= -\sqrt{\frac{h_{21}}{h_{12}}} = -\frac{\sqrt{h_{12}h_{21}}}{h_{12}} \\ &= -\frac{\frac{1}{2}\left(\Delta m - \frac{i}{2}\Delta\Gamma\right)}{m_{12} - \frac{i}{2}\Gamma_{12}} \\ &= -\frac{m_{12}^* - \frac{i}{2}\Gamma_{12}^*}{\frac{1}{2}\left(\Delta m - \frac{i}{2}\Delta\Gamma\right)}. \end{aligned} \tag{1.31}$$

1.3.3 CP Violation in the Mixing

Using a phase convention in which

$$\mathcal{CP}|\mathrm{B}^0\rangle = |\overline{\mathrm{B}^0}\rangle, \tag{1.32}$$

the $\mathcal{CP}$ operator in B^0–$\overline{\mathrm{B}^0}$ space is given by the first Pauli spin matrix:

$$\mathcal{CP} = \sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{1.33}$$

We can write the Hamiltonian, like any other 2×2 matrix, as

$$\mathbf{H} = a^0\mathbf{1} + a^1\sigma^1 + a^2\sigma^2 + a^3\sigma^3 \tag{1.34}$$

where a^i are four complex numbers. Then $\mathcal{CP}$ invariance requires:

$$\left[H_{\mathcal{CP}-\text{conserving}}, \sigma^1\right] = 0. \tag{1.35}$$

With the above definitions:

$$\left[H, \sigma^1\right] = -a^2\sigma^3 + a^3\sigma^2. \tag{1.36}$$

Therefore, for a $\mathcal{CP}$ conserving Hamiltonian, $a^2 = a^3 = 0$, and the most general $\mathcal{CP}$ conserving Hamiltonian can be written as:

$$\mathbf{H}_{\mathcal{CP}-\text{conserving}} = \begin{pmatrix} a^0 & a^1 \\ a^1 & a^0 \end{pmatrix}. \tag{1.37}$$

With equation 1.31 we get therefore:

$$\mathcal{CP}\text{ Conservation} \Rightarrow p = -q \tag{1.38}$$

Changing to a general phase convention with $\mathcal{CP}|\mathrm{B}^0\rangle = e^{i\alpha}|\overline{\mathrm{B}^0}\rangle$ is equivalent to a basis transformation:

$$|\mathrm{B}^0_\mathrm{d}\rangle' = e^{+i\alpha/2}|\mathrm{B}^0_\mathrm{d}\rangle, \quad |\overline{\mathrm{B}^0_\mathrm{d}}\rangle' = e^{-i\alpha/2}|\overline{\mathrm{B}^0_\mathrm{d}}\rangle. \tag{1.39}$$

The column-vectors transform like:

$$\begin{pmatrix} a \\ b \end{pmatrix}' = \begin{pmatrix} e^{-i\alpha/2} & 0 \\ 0 & e^{i\alpha/2} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}, \tag{1.40}$$

and $\mathcal{CP}$ is given by:

$$\begin{pmatrix} e^{-\frac{i}{2}\alpha} & 0 \\ 0 & e^{\frac{i}{2}\alpha} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} e^{\frac{i}{2}\alpha} & 0 \\ 0 & e^{-\frac{i}{2}\alpha} \end{pmatrix} = \begin{pmatrix} 0 & e^{-i\alpha} \\ e^{i\alpha} & 0 \end{pmatrix}. \tag{1.41}$$

Therefore

$$\mathcal{CP}|\mathrm{B}^0\rangle = e^{i\alpha}|\overline{\mathrm{B}^0}\rangle \tag{1.42}$$

as required. In this basis, $\mathbf{H}$ is given by:

$$\mathbf{H} = \begin{pmatrix} h_{11} & h_{12}e^{-i\alpha} \\ h_{21}e^{i\alpha} & h_{11} \end{pmatrix} \tag{1.43}$$

and the ratio q/p transforms like:

$$\frac{q}{p} \longrightarrow \frac{q}{p}e^{i\alpha}. \tag{1.44}$$

So the phase–convention independent statement is:

$$\mathcal{CP}\text{ Conservation} \Rightarrow |p| = |q| \tag{1.45}$$

1.3.4 B^0 decays

In the previous section it was shown that the mass eigenstates for the B^0-system are given by:

$$|\mathrm{B}_{\mathrm{H,L}}\rangle = p|\mathrm{B}^0\rangle \mp q|\overline{\mathrm{B}^0}\rangle, \tag{1.46}$$

with $|p|^2 + |q|^2 = 1$ and $\frac{q}{p} = -\frac{\frac{1}{2}\left(\Delta m - \frac{i}{2}\Delta\Gamma\right)}{m_{12} - \frac{i}{2}\Gamma_{12}}$. Inverting the above equation gives:

$$\begin{aligned} |\mathrm{B}^0\rangle &= +\frac{1}{2p}|\mathrm{B_H}\rangle + \frac{1}{2p}|\mathrm{B_L}\rangle \\ |\overline{\mathrm{B}^0}\rangle &= -\frac{1}{2q}|\mathrm{B_H}\rangle + \frac{1}{2q}|\mathrm{B_L}\rangle. \end{aligned} \tag{1.47}$$

The two B^0_d mesons can in principle have different life times. This is produced by decay channels common to B^0_d and $\overline{\mathrm{B}^0_\mathrm{d}}$. The branching ratios for such channels are of the order of 10^{-3}, or smaller. As various channels contribute with different signs, one expects that their sum does not exceed the individual level, so $\Delta\Gamma \ll \Gamma$, or, more precisely,

$$\frac{\Delta\Gamma}{\Gamma} = \mathcal{O}\left(10^{-2}\right) \tag{1.48}$$

is a rather safe and model-independent assumption. The present measured value of Δm is given by [G+00]:

$$x_d \equiv \frac{\Delta m_{\mathrm{B}^0_\mathrm{d}}}{\Gamma_{\mathrm{B}^0_\mathrm{d}}} = 0.730 \pm 0.029. \tag{1.49}$$

Therefore:

$$\Delta\Gamma \ll \Delta m \tag{1.50}$$

for the B^0_d system. So $\Gamma_{12} \ll m_{12}$, and therefore

$$|p| = |q| \tag{1.51}$$

is a very good approximation. Therefore there is no considerable $\mathcal{CP}$ violation in B^0_d mixing. With $|p| = |q|$ we can write the ratio q/p as:

$$\frac{q}{p} = e^{i\phi_{\mathrm{mix}}} \tag{1.52}$$

where ϕ_{mix} is the mixing phase. Within the Standard Model, the box diagram for B^0_d mixing gives

$$\phi_{\mathrm{mix}} = -2\beta \tag{1.53}$$

as will be shown in section 1.4.13.

The wavefunction of a particle created as a B^0 at time $t = 0$ develops as:

$$\begin{aligned} |\mathrm{B}^0(t)\rangle = &\frac{1}{2p}|\mathrm{B_H}\rangle e^{-i\left(M_H - \frac{i}{2}\Gamma_H\right)t} \\ &+\frac{1}{2p}|\mathrm{B_L}\rangle e^{-i\left(M_L - \frac{i}{2}\Gamma_L\right)t}. \end{aligned} \tag{1.54}$$

With $\Gamma_H \approx \Gamma_L \equiv \Gamma$:

$$\begin{aligned} |\mathrm{B}^0(t)\rangle = &e^{-\frac{\Gamma}{2}t}\frac{1}{2p} \\ &\left\{|\mathrm{B_H}\rangle e^{-iM_Ht} + |\mathrm{B_L}\rangle e^{-iM_Lt}\right\} \\ = &e^{-\frac{\Gamma}{2}t}\frac{1}{2p}\left\{\left(p|\mathrm{B}^0\rangle - q|\overline{\mathrm{B}^0}\rangle\right)e^{-iM_Ht}\right. \\ &\left.+\left(p|\mathrm{B}^0\rangle + q|\overline{\mathrm{B}^0}\rangle\right)e^{-iM_Lt}\right\} \\ = &e^{-\frac{\Gamma}{2}t}\frac{1}{2}\left\{\left(|\mathrm{B}^0\rangle - \frac{q}{p}|\overline{\mathrm{B}^0}\rangle\right)e^{-iM_Ht}\right. \\ &\left.+\left(|\mathrm{B}^0\rangle + \frac{q}{p}|\overline{\mathrm{B}^0}\rangle\right)e^{-iM_Lt}\right\}. \end{aligned} \tag{1.55}$$

The amplitude for the decay to the final state $|f\rangle$ is then:

$$
\begin{aligned}
&\langle f|\mathbf{H}_I|\mathrm{B}^0(t)\rangle \\
&= e^{-\frac{\Gamma}{2}t}\frac{1}{2} \\
&\quad\cdot\left\{\left(\langle f|\mathbf{H}_I|\mathrm{B}^0\rangle - \frac{q}{p}\langle f|\mathbf{H}_I|\overline{\mathrm{B}^0}\rangle\right)e^{-iM_Ht}\right. \\
&\quad\left.+\left(\langle f|\mathbf{H}_I|\mathrm{B}^0\rangle + \frac{q}{p}\langle f|\mathbf{H}_I|\overline{\mathrm{B}^0}\rangle\right)e^{-iM_Lt}\right\}
\end{aligned}
$$

$$
\begin{aligned}
&= e^{-\frac{\Gamma}{2}t}\frac{1}{2}\langle f|\mathbf{H}_I|\mathrm{B}^0\rangle \\
&\quad\cdot\left\{\left(1-\frac{q}{p}\frac{\langle f|\mathbf{H}_I|\overline{\mathrm{B}^0}\rangle}{\langle f|\mathbf{H}_I|\mathrm{B}^0\rangle}\right)e^{-iM_Ht}\right. \\
&\quad\left.+\left(1+\frac{q}{p}\frac{\langle f|\mathbf{H}_I|\overline{\mathrm{B}^0}\rangle}{\langle f|\mathbf{H}_I|\mathrm{B}^0\rangle}\right)e^{-iM_Lt}\right\},
\end{aligned}
\tag{1.56}
$$

where $\mathbf{H}_I$ is the interaction Hamiltonian. Defining

$$
\eta \equiv \frac{q}{p}\frac{\langle f|\mathbf{H}_I|\overline{\mathrm{B}^0}\rangle}{\langle f|\mathbf{H}_I|\mathrm{B}^0\rangle},
\tag{1.57}
$$

$$
\begin{aligned}
\langle f|\mathbf{H}_I|\mathrm{B}^0(t)\rangle &= e^{-\frac{\Gamma}{2}t}\frac{1}{2}\langle f|\mathbf{H}_I|\mathrm{B}^0\rangle \\
&\quad\cdot\left\{(1-\eta)\,e^{-iM_Ht}+(1+\eta)\,e^{-iM_Lt}\right\}.
\end{aligned}
\tag{1.58}
$$

The square of the matrix element, and therefore the decay rate, neglecting the phasespace factor and some constants, is:

$$
\begin{aligned}
&\left|\langle f|\mathbf{H}_I|\mathrm{B}^0(t)\rangle\right|^2 \\
&= e^{-\Gamma t}\frac{1}{4}\left|\langle f|\mathbf{H}_I|\mathrm{B}^0\rangle\right|^2\cdot\left\{2+2\,|\eta|^2\right. \\
&\quad+\left(1-\eta+\eta^*-|\eta|^2\right)e^{-i\Delta mt} \\
&\quad\left.+\left(1+\eta-\eta^*-|\eta|^2\right)e^{+i\Delta mt}\right\} \\
&= e^{-\Gamma t}\frac{1}{2}\left|\langle f|\mathbf{H}_I|\mathrm{B}^0\rangle\right|^2\cdot\left\{\left(1+|\eta|^2\right)\right. \\
&\quad+\left(1-|\eta|^2\right)\cos(\Delta mt) \\
&\quad\left.-2\mathsf{Im}(\eta)\sin(\Delta mt)\right\}.
\end{aligned}
\tag{1.60}
$$

Defining

$$
\overline{\eta} = \frac{p}{q}\frac{\langle\overline{f}|\mathbf{H}_I|\mathrm{B}^0\rangle}{\langle\overline{f}|\mathbf{H}_I|\overline{\mathrm{B}^0}\rangle}
\tag{1.61}
$$

for the CP conjugate final state $\overline{f}$, and making the appropriate substitutions, the matrix element of all four decay rates: $\mathrm{B}^0, \overline{\mathrm{B}^0} \longrightarrow f, \overline{f}$ can be derived. The results are given in equation 1.59, page 10.

The complex parameters η and $\overline{\eta}$ are observable, phase–convention independent quantities that can be extracted from time–dependent decay rate measurements.

1.3.5 CP violation in decays to CP eigenstates

If the final state $|f\rangle$ is a $\mathcal{CP}$ eigenstate, and we assume $\left|\frac{q}{p}\right| = 1$ we get $|\eta| = 1$, and η and $\overline{\eta}$ are complex conjugates of each other. Then the four decay rates in equation 1.59 reduce to two. In the absence of $\mathcal{CP}$ violation, $\eta = 1$ or $\eta = -1$. In the presence of $\mathcal{CP}$ violation, the phase of η is equal to the phase difference between the two decay paths, $\mathrm{B}^0_\mathrm{d} \rightarrow f$ and $\overline{\mathrm{B}^0_\mathrm{d}} \rightarrow f$:

$$
\eta = \pm e^{i\delta_{CP}},
\tag{1.62}
$$

where δ_{CP} is the $\mathcal{CP}$ violating phase. The amplitude of the time–dependent asymme-

Decay rates

$$\begin{aligned}
\left|\langle f|\mathbf{H}_I|\mathrm{B}^0(t)\rangle\right|^2 &= e^{-\Gamma t}\frac{1}{2}\left|\langle f|\mathbf{H}_I|\mathrm{B}^0\rangle\right|^2 \cdot \left\{ \left(1+|\eta|^2\right) + \left(1-|\eta|^2\right)\cos(\Delta m t) - 2\mathsf{Im}(\eta)\sin(\Delta m t) \right\} \\
\left|\langle f|\mathbf{H}_I|\overline{\mathrm{B}^0}(t)\rangle\right|^2 &= e^{-\Gamma t}\frac{1}{2}\left|\langle f|\mathbf{H}_I|\mathrm{B}^0\rangle\right|^2 \cdot \left\{ \left(1+|\eta|^2\right) - \left(1-|\eta|^2\right)\cos(\Delta m t) + 2\mathsf{Im}(\eta)\sin(\Delta m t) \right\} \\
\left|\langle \overline{f}|\mathbf{H}_I|\overline{\mathrm{B}^0}(t)\rangle\right|^2 &= e^{-\Gamma t}\frac{1}{2}\left|\langle \overline{f}|\mathbf{H}_I|\overline{\mathrm{B}^0}\rangle\right|^2 \cdot \left\{ \left(1+|\overline{\eta}|^2\right) + \left(1-|\overline{\eta}|^2\right)\cos(\Delta m t) - 2\mathsf{Im}(\overline{\eta})\sin(\Delta m t) \right\} \\
\left|\langle \overline{f}|\mathbf{H}_I|\mathrm{B}^0(t)\rangle\right|^2 &= e^{-\Gamma t}\frac{1}{2}\left|\langle \overline{f}|\mathbf{H}_I|\overline{\mathrm{B}^0}\rangle\right|^2 \cdot \left\{ \left(1+|\overline{\eta}|^2\right) - \left(1-|\overline{\eta}|^2\right)\cos(\Delta m t) + 2\mathsf{Im}(\overline{\eta})\sin(\Delta m t) \right\}
\end{aligned} \tag{1.59}$$

try of the two decay rates gives the sine of the $\mathcal{CP}$-violating phase:

$$\begin{aligned}
A(\tau) &= \frac{\Gamma\left(\mathrm{B}^0 \to f\right) - \Gamma\left(\overline{\mathrm{B}^0} \to f\right)}{\Gamma\left(\mathrm{B}^0 \to f\right) + \Gamma\left(\overline{\mathrm{B}^0} \to f\right)} \\
&= \mathsf{Im}(\eta)\sin(\Delta m\,\tau) \\
&= \pm\sin(\delta_{\mathrm{CP}})\sin(\Delta m\,\tau)\,.
\end{aligned} \tag{1.63}$$

1.3.6 CP violation in decays to non-CP eigenstates

Even if the final state is not a $\mathcal{CP}$ eigenstate, $\mathcal{CP}$ violation can be measured if both the B^0 and the $\overline{\mathrm{B}^0}$ can decay to the same final state f. Then

$$\begin{aligned}
\eta &= |\eta|\,e^{i\left(\delta_{\mathrm{CP}}+\Delta_{\mathrm{qcd}}\right)} \\
\overline{\eta} &= |\eta|\,e^{i\left(-\delta_{\mathrm{CP}}+\Delta_{\mathrm{qcd}}\right)},
\end{aligned} \tag{1.64}$$

where Δ_{qcd} is a possible, non-$\mathcal{CP}$-violating strong phase. Now all four expressions for the decay rates are needed. The $\mathcal{CP}$-violating effect is in the difference between the asymmetries formed from the decay rates.

$$\begin{aligned}
A_\eta(\tau) &= \frac{\Gamma\left(\mathrm{B}^0 \to f\right) - \Gamma\left(\overline{\mathrm{B}^0} \to f\right)}{\Gamma\left(\mathrm{B}^0 \to f\right) + \Gamma\left(\overline{\mathrm{B}^0} \to f\right)} \\
&= \tfrac{1-|\eta|^2}{1+|\eta|^2}\cos(\Delta m\tau) - \tfrac{2\mathsf{Im}(\eta)}{1+|\eta|^2}\sin(\Delta m\tau)
\end{aligned} \tag{1.65}$$

and

$$\begin{aligned}
A_{\overline{\eta}}(\tau) &= \frac{\Gamma\left(\overline{\mathrm{B}^0} \to \bar{f}\right) - \Gamma\left(\mathrm{B}^0 \to \bar{f}\right)}{\Gamma\left(\overline{\mathrm{B}^0} \to \bar{f}\right) + \Gamma\left(\mathrm{B}^0 \to \bar{f}\right)} \\
&= \tfrac{1-|\eta|^2}{1+|\eta|^2}\cos(\Delta m\tau) - \tfrac{2\mathsf{Im}(\overline{\eta})}{1+|\eta|^2}\sin(\Delta m\tau)\,.
\end{aligned} \tag{1.66}$$

From fits to the amplitudes of the sine and cosine part of the asymmetries, $|\eta|^2$, $\mathsf{Im}(\eta)$ and $\mathsf{Im}(\overline{\eta})$, or equivalently $|\eta|^2$, δ_{CP} and Δ_{qcd}, can be extracted.

1.4 CP Violation in the Standard Model

1.4.1 The Standard Model

The Standard Model of the electroweak and strong interactions is based on the gauge group $SU(3)_C \times SU(2)_L \times U(1)$. The $SU(3)$-colour part of it describes the strong interaction. The $SU(2)_L \times U(1)$ part describes the electroweak interactions. While the $SU(3)$–colour symmetry is directly realised in nature, the $SU(2)_L \times U(1)$ is broken by the Higgs mechanism, giving rise to the massive weak–interaction vector bosons, the $W^{\pm}$ and Z^0. Interactions with the Higgs fields are also responsible for generating fermion masses in a gauge invariant way. One symmetry emerges unbroken from the Higgs mechanism: the $U(1)_{\mathrm{QED}}$ symmetry that describes the electromagnetic interactions, mediated by the massless photon. The electroweak part of the Standard Model is referred to as the Glashow-Salam-Weinberg (GSW) model [Gla61, Sal68, Wei67].

The Standard Model is completely defined by

- the definition of the gauge groups;
- the representation of the gauge groups assigned to the different particles;
- the requirement that the Standard Model should be renormalisable.

The following summarises those Standard Model aspects which are relevant to $\mathcal{CP}$ violation in the B^0_d system.

The particles of the Standard Model with their quantum numbers are listed in tables 1.2 and 1.3, and their masses are given in table 1.4.

Right–handed neutrinos are not listed in table 1.2, although recent evidence for non-zero neutrino masses [F+98] suggests their existence. However, a discussion of the possible consequences of this discovery is beyond the scope of this thesis.

1.4.2 The Strong CP Problem

The most general Lagrangian that fulfils the above requirements of gauge invariance and renormalisability, contains a term that violates $\mathcal{C}$, $\mathcal{P}$ and $\mathcal{CP}$ in the strong interaction. Highly accurate measurements, especially of the neutron electric dipole moment[1], suggest however that the $\mathcal{C}$, $\mathcal{P}$, and $\mathcal{CP}$ violating term in the strong Lagrangian vanishes, leaving only terms that conserve $\mathcal{C}$ and $\mathcal{P}$ separately, or at least that the symmetry violating term is very small. While this is allowed in the Standard Model, it is by no means explained. This is known as the "Strong $\mathcal{CP}$ Problem". For the purpose of this thesis, we will accept that $\mathcal{C}$ and $\mathcal{P}$ are symmetries of the Strong Interaction, and concentrate on $\mathcal{CP}$ violation in the electroweak sector.

---

[1] The current limit is $< 0.63 \cdot 10^{-25} e\mathrm{cm}$ [G+00]

Table 1.2: The particles of the Standard Model with quantum numbers, excluding the gauge bosons which are listed in table 1.3

Fermions			$SU(2)_L$		$U(1)_Y$	$U(1)_{\text{QED}}$
			Weak Isospin		$\frac{1}{2}$Hypercharge	Electric Charge
leptons			T	T^3	Y	$Q = T^3 + Y$
$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$	$\frac{1}{2}$	$\begin{pmatrix} +\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	$-\frac{1}{2}$	$\begin{pmatrix} 0 \\ -1 \end{pmatrix}$
e_R	μ_R	τ_R	0	0	-1	-1
All leptons have 0 colour charge and do not take part in the strong interaction						
quarks						
$\begin{pmatrix} u \\ d \end{pmatrix}_L$	$\begin{pmatrix} c \\ s \end{pmatrix}_L$	$\begin{pmatrix} t \\ b \end{pmatrix}_L$	$\frac{1}{2}$	$\begin{pmatrix} +\frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	$+\frac{1}{6}$	$\begin{pmatrix} +\frac{2}{3} \\ -\frac{1}{3} \end{pmatrix}$
u_R	c_R	t_R	0	0	$+\frac{2}{3}$	$+\frac{2}{3}$
d_R	s_R	b_R	0	0	$-\frac{1}{3}$	$-\frac{1}{3}$
All quarks come in $SU(3)_{\text{QCD}}$-colour triplets and take part in the strong interaction: $u = \left(u^{\text{red}},\ u^{\text{green}},\ u^{\text{blue}}\right)$, $d = \left(d^{\text{red}},\ d^{\text{green}},\ d^{\text{blue}}\right)$, etc.						
Higgs Boson						
h_0			$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	0
The Higgs is colour neutral. It couples to all massive particles.						

Table 1.3: Standard Model Gauge Bosons

Gauge Bosons	$SU(2)_L$		$U(1)_Y$	$U(1)_{\text{QED}}$	$SU(3)_{\text{QCD}}$
	Weak Isospin		$\frac{1}{2}$Hypercharge	Electric Charge	colour charge
	T	T^3	Y	$Q = T^3 + Y$	
Weak: $\begin{pmatrix} W^+ \\ Z^0 \\ W^- \end{pmatrix}$	1	$\begin{pmatrix} +1 \\ 0 \\ -1 \end{pmatrix}$	0	$\begin{pmatrix} +1 \\ 0 \\ -1 \end{pmatrix}$	no colour
QED: photon γ	0	0	0	0	no colour
QCD: 8 gluons g^i	0	0	0	0	$r\bar{g}$, $r\bar{b}$, $g\bar{r}$, $g\bar{b}$, $b\bar{r}$, $b\bar{g}$, $\frac{r\bar{r}-b\bar{b}}{\sqrt{2}}$, $\frac{r\bar{r}+b\bar{b}-2g\bar{g}}{\sqrt{6}}$

Table 1.4: Particle Masses. The precision of all measurements, except for the quark masses, is higher than the number of digits given. All data are taken from [G+00], unless indicated otherwise.

leptons		
ν_e	ν_μ	ν_τ
$< 3\,\mathrm{eV}$ "limit is not without ambiguity" [G+00]	$< 0.19\,\mathrm{MeV}$ (90% C.L.)	$< 18.2\,\mathrm{MeV}$ (95% C.L.)
Super-Kamiokande data are consistent with $\nu_\mu \leftrightarrow \nu_\tau$ oscillations with $5 \cdot 10^{-4} < \Delta\mathrm{m}^2 < 6 \cdot 10^{-3}\,\mathrm{eV}^2$ at 90% C.L. [F+98].		
e	μ	τ
0.5 MeV	0.1 GeV	1.8 GeV
quarks		
u	c	t
$1 - 5$ MeV	$1.15 - 1.35$ GeV	174.3 ± 5.1 GeV
d	s	b
$3 - 9$ MeV	$75 - 170$ MeV	$4.0 - 4.4$ GeV
$m_u/m_d = 0.2 - 0.8$, $m_s/m_d = 17 - 25$		
Higgs h_0		
$> 113.5\,\mathrm{GeV}$ (95% C.L.) [IK00]		
electro-weak gauge bosons		
$W^\pm$	Z^0	γ
80.4 GeV	91.2 GeV	0
QCD gauge bosons g^i		
0		

1.4.3 The GSM Theory of Electroweak Interactions

The GSM theory of weak interactions unifies the electromagnetic and weak interaction. It reconciles gauge invariance of the Lagrangian with the observation of massive vector bosons.

In general, massive gauge bosons lead to non-gauge invariant theories. Glashow, Salam and Weinberg's theory starts with a $SU(2) \times U(1)$ gauge theory with massless gauge bosons. The symmetry is broken via the Higgs mechanism [Hig64b, Hig64a], leading, in a gauge invariant way, to the massive gauge bosons $W^\pm$ and the mediator of the newly predicted neutral currents, the massive Z^0. One massless gauge bosons remains, the photon, which is the gauge boson of the unbroken $U(1)_{\mathrm{QED}}$ symmetry. Crucial for the success of the theory was that it could be shown to be renormalisable [tHV72].

In the GSW theory the masses of the fermions are generated in a gauge invariant and renormalisable way, by Yukawa couplings of the Higgs field to the fermions. It is those Yukawa couplings that lead to $\mathcal{CP}$ violation in the Standard Model.

1.4.4 $SU(2) \times U(1)$ and the Higgs

Massless Gauge Bosons

The covariant derivative for an $SU(2) \times U(1)$ local gauge theory is

$$D_\mu = \partial_\mu - igA^a_\mu T^a - ig'YB_\mu, \tag{1.67}$$

where the T^a are the generators of the $SU(2)$ gauge symmetry. They are evaluated according to the $SU(2)$ representation the derivative acts on. The generator of $U(1)$ is Y, which is just a number. A^a_μ and B_μ are massless $SU(2)$ and $U(1)$ gauge bosons, respectively.

The Higgs Field

To obtain the physically observed massive gauge bosons $W^\pm_\mu$ and Z^0_μ, the $SU(2)\times U(1)$ symmetry is broken via the Higgs mechanism. The Higgs field is a scalar field in the spinor representation of $SU(2)$, with $U(1)$ charge $\frac{1}{2}$:

$$\phi = \begin{pmatrix} H^+ \\ H^0 \end{pmatrix}, \tag{1.68}$$

where H^+ and H^0 are complex fields. The indices $+$ and 0 anticipate the electric charge assignments, to be defined later. For the spinor representation of $SU(2)$, and $U(1)$ charge $\frac{1}{2}$, the covariant derivative is:

$$D_\mu = \partial_\mu - igA^a_\mu \frac{\sigma^a}{2} - i\frac{1}{2}g'B_\mu, \tag{1.69}$$

where σ^a are the Pauli spin matrices.

We can use $SU(2)$ gauge freedom to transform the Higgs field:

$$\begin{pmatrix} H^+ \\ H^0 \end{pmatrix} \to e^{i\xi_a \frac{\sigma^a}{2}} \begin{pmatrix} H^+ \\ H^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h_0 \end{pmatrix} \tag{1.70}$$

where v and h_0 are real numbers. This is called the "unitarity gauge". To break the symmetry, the Higgs acquires a vacuum expectation value:

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.71}$$

The Higgs Boson is the excitation of the Higgs field from the vacuum expectation value and corresponds to h_0 in equation 1.70.

Mass-Eigenstates of Gauge Bosons

The mass terms for the gauge bosons come from the kinetic term of the Lagrangian for the Higgs field:

$$\Delta\mathcal{L} = \left(D_\mu \phi\right)^\dagger \left(D^\mu \phi\right) \tag{1.72}$$

evaluated at the vacuum expectation value of the Higgs. The mass eigenstates of the heavy gauge bosons are:

$$W^{\mu\pm} \equiv \frac{1}{\sqrt{2}}\left(A^1 \mp iA^2\right) \tag{1.73}$$

$$Z^\mu \equiv \frac{1}{\sqrt{g^2 + g'^2}}\left(gA^{3\mu} - g'B^\mu\right). \tag{1.74}$$

One massless gauge boson remains, the photon:

$$A^\mu = \frac{1}{\sqrt{g^2 + g'^2}}\left(g'A^{3\mu} + gB^\mu\right) \tag{1.75}$$

D$_\mu$ in Terms of the Massive Gauge Bosons

Before writing down the covariant derivative in terms of the mass eigenstates of the gauge bosons, it is useful to make the following definitions:

- the charge raising and lowering operators:

$$T^{\pm} \equiv \left(T^1 \pm iT^2\right), \tag{1.76}$$

- the charge quantum number:

$$Q \equiv T^3 + Y \tag{1.77}$$

- the weak mixing angle θ_w:

$$\begin{aligned} c_w \equiv \cos\theta_w &= \frac{g}{\sqrt{g^2+g'^2}} \\ s_w \equiv \sin\theta_w &= \frac{g'}{\sqrt{g^2+g'^2}}, \end{aligned} \tag{1.78}$$

- the electron charge:

$$e = \frac{gg'}{\sqrt{g^2+g'^2}} \tag{1.79}$$

With these definitions, the covariant derivative in terms of the physically observable gauge bosons is given by:

$$\begin{aligned} D_\mu = \partial_\mu &- i\frac{e}{\sqrt{2}s_w}\left(W_\mu^+T^+ + W_\mu^-T^-\right) \\ &- i\frac{e}{s_w c_w}Z_\mu\left(T^3 - s_w^2 Q\right) \\ &- ieA_\mu Q. \end{aligned} \tag{1.80}$$

1.4.5 Fermions

The observation that only left-handed fermions couple to the W boson can be accommodated by choosing two different representations of $SU(2)$ for left and right handed fermions. Left-handed fermions are put into $SU(2)$ doublets with

$$T^3 = \begin{pmatrix} +\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}. \tag{1.81}$$

There are the left-handed leptons with $Y = -\frac{1}{2}$:

$$\begin{aligned} E_L^i &= \begin{pmatrix} \nu^i \\ l^i \end{pmatrix}_L \\ &= \left(\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L\right). \end{aligned} \tag{1.82}$$

The charge $Q = Y + T^3$ has the observed values 0 and 1. Similarly for left-handed quarks

$$\begin{aligned} Q_L^i &= \begin{pmatrix} u^i \\ d^i \end{pmatrix}_L \\ &= \left(\begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L\right) \end{aligned} \tag{1.83}$$

with $Y = +\frac{1}{6}$, $Q = Y + T^3 = +\frac{2}{3}, -\frac{1}{3}$. Right–handed fermions are put into $SU(2)$ singlets, with $T^3 = 0$.

$$\begin{array}{ll} e_R^i = (e_R, \mu_R, \tau_R) & \text{with } Y = Q = -1 \\ u_R^i = (u_R, c_R, t_R) & \text{with } Y = Q = +\frac{2}{3} \\ d_R^i = (d_R, s_R, b_R) & \text{with } Y = Q = -\frac{1}{3} \end{array}.$$

1.4.6 Electroweak Interactions with Fermions

The Lagrangian of electroweak interactions, ignoring fermion masses, follows directly from the above charge-assignments. The kinetic terms are given by:

$$\begin{aligned}\mathcal{L} &= \bar{E}_L \left(i\not{D}\right) E_L + \bar{e}_R \left(i\not{D}\right) e_R \\ &+\bar{Q}_L \left(i\not{D}\right) Q_L + \\ &\bar{u}_R \left(i\not{D}\right) u_R + \bar{d}_R \left(i\not{D}\right) d_R, \end{aligned} \tag{1.84}$$

where D_μ is given by equation 1.67, with T^a and Y evaluated according to the field on which they act, for example for left–handed quarks:

$$T^a = \frac{\sigma^a}{2}, \quad Y = \frac{1}{6} \tag{1.85}$$

or for right–handed electrons:

$$T^a = 0 \quad Y = -1. \tag{1.86}$$

In terms of mass eigenstates of the $W^\pm$ and Z^0 bosons:

$$\begin{aligned}\mathcal{L} &= \bar{E}_L \left(i\not{\partial}\right) E_L + \bar{e}_R \left(i\not{\partial}\right) e_R \\ &+\bar{Q}_L \left(i\not{\partial}\right) Q_L \\ &+\bar{u}_R \left(i\not{\partial}\right) u_R + \bar{d}_R \left(i\not{\partial}\right) d_R \\ &+g\left(W^+_\mu J^{\mu +}_W + W^-_\mu J^{\mu -}_W + Z^0_\mu J^\mu_Z\right) \\ &+eA_\mu J^\mu_{EM} \end{aligned} \tag{1.87}$$

where the W boson currents

$$\begin{aligned} J^{\mu+}_W &= \frac{1}{\sqrt{2}}\left(\bar{\nu}_L\gamma^\mu e_L + \bar{u}_L\gamma^\mu d_L\right) \\ J^{\mu-}_W &= \frac{1}{\sqrt{2}}\left(\bar{e}_L\gamma^\mu \nu_L + \bar{d}_L\gamma^\mu u_L\right) \end{aligned} \tag{1.88}$$

couple the $Y^3 = \frac{1}{2}$ to the $Y^3 = -\frac{1}{2}$ fermions, while the Z currents are all diagonal:

$$\begin{aligned} J^\mu_Z = \frac{1}{c_w}\Bigg\{ &\bar{\nu}_L\gamma^\mu\left(\frac{1}{2}\right)\nu_L \\ &+\bar{e}_L\gamma^\mu\left(-\frac{1}{2}+s^2_w\right)e_L + \bar{e}_R\gamma^\mu\left(s^2_w\right)e_R \\ &+\bar{u}_L\gamma^\mu\left(\frac{1}{2}-\frac{2}{3}s^2_w\right)u_L + \bar{u}_R\gamma^\mu\left(-\frac{2}{3}s^2_w\right)u_R \\ &+\bar{d}_L\gamma^\mu\left(-\frac{1}{2}+\frac{1}{3}s^2_w\right)d_L + \bar{d}_R\gamma^\mu\left(\frac{1}{3}s^2_w\right)d_R\Bigg\}. \end{aligned} \tag{1.89}$$

The terms in the electromagnetic current, are also diagonal and do not distinguish between left and right handed particles:

$$\begin{aligned} J^\mu_{EM} &= \bar{e}\gamma^\mu\left(-1\right)e + \\ &\bar{u}\gamma^\mu\left(+\frac{2}{3}\right)u + \bar{d}\gamma^\mu\left(-\frac{1}{3}\right)d \end{aligned} \tag{1.90}$$

From the expressions for the different currents, we see that we retrieve the experimentally observed $\mathcal{C}$ and $\mathcal{P}$ violation in the weak interactions mediated by the $W^\pm$ and the Z^0, while the electromagnetic current conserves both $\mathcal{C}$ and $\mathcal{P}$ separately. No interaction violates $\mathcal{CP}$ at this stage of the theory.

1.4.7 CP Violation

Fermion Mass Terms

So far in the theory all fermions are massless, in contradiction to observation. Simply adding mass terms to the Lagrangian, such as

$$\Delta\mathcal{L} = -m_e\left(\bar{e}_L e_R + \bar{e}_R e_L\right), \tag{1.91}$$

is not allowed. These terms violate gauge invariance because we had to put the left-handed and the right-handed fermions into different $SU(2)$ representations in order to account for the observed maximal $\mathcal{C}$ and $\mathcal{P}$ violation of the weak interaction.

Gauge invariant mass terms can be generated by couplings to the Higgs field. The most general renormalisable gauge-invariant coupling for the leptonic fields is:

$$\mathcal{L}_m = -\lambda_l^{ij}\bar{E}_L^i\cdot\phi e_R^j + \text{h.c.} \tag{1.92}$$

and for the quark fields:

$$\mathcal{L}_m = -\lambda_d^{ij}\bar{Q}_L^i\cdot\phi d_R^j - \lambda_u^{ij}\epsilon^{ab}\bar{Q}_{La}^i\phi_b^\dagger u_R^j + \text{h.c.} \tag{1.93}$$

where λ_l^{ij}, λ_d^{ij} and λ_u^{ij} are general, complex valued matrices. The second term in equation 1.93 is missing for the leptonic fields due to the absence of right-handed neutrinos.

Below we will show how the quark mass terms in equation 1.93 lead to the possibility of $\mathcal{CP}$ violation in the Standard Model.

Quark Masses

In order to find the Lagrangian in terms of the physical particles, we have to diagonalise the Yukawa couplings in equation 1.93.

Any complex matrix λ can be written as

$$\lambda = UDW^\dagger \tag{1.94}$$

where U and W are unitary matrices, and D is a diagonal matrix with positive eigenvalues. Using the same notation as above, we define the matrices U_u, D_u, W_u and U_d, D_d, W_d by

$$\lambda_u = U_u D_u W_u^\dagger, \quad \lambda_d = U_d D_d W_d^\dagger. \tag{1.95}$$

Changing the basis for right-handed quarks:

$$u_R^i \to W_u^{ij}u_R^j, \quad d_R^i \to W_d^{ij}d_R^j \tag{1.96}$$

eliminates the matrices W_u and W_d from the Higgs couplings in equation 1.93. Since all three right–handed up–type quarks have the same coupling to the gauge fields, as do the three right-handed down–type quarks, W_u and W_d commute with the corresponding covariant derivatives. Therefore

$$\sum_i\left(\bar{u}_R\left(i\not{D}\right)u_R + \bar{d}_R\left(i\not{D}\right)d_R\right) \tag{1.97}$$

is invariant under the transformation 1.96 and W_u and W_d simply disappear from the theory.

Similarly we transform

$$u_L^i \to U_u^{ij}u_L^j, \quad d_L^i \to U_d^{ij}d_L^j. \tag{1.98}$$

This eliminates U_u and U_d from the terms in equation 1.93 that involve the lower component of the Higgs field, and in unitarity gauge this is the only component that is non-zero. In this new basis, only the diagonal elements in equation 1.93 survive. Evaluated at the vacuum expectation value for ϕ:

$$\mathcal{L}_m = -m_u^i \bar{u}_L^i u_R^j - m_d^i \bar{d}_L^i d_R^j \tag{1.99}$$

where m_u^i and m_d^i are the quark masses, given by

$$m_u^i = \frac{1}{\sqrt{2}} D_u^{ii} v, \quad m_d^i = \frac{1}{\sqrt{2}} D_d^{ii} v \tag{1.100}$$

The fact that the quark mass terms are all proportional to v might suggest that they should be of a similar magnitude. In reality they vary over four orders of magnitude, as shown in table 1.4. This is allowed in the Standard Model, since $D_{u,d}^{ii}$ are free parameters, but it is not explained.

The effect of the transformation of the left-handed quark basis cancel as long as only diagonal terms are involved, so U_d and U_u vanish in all the kinetic terms such as $\bar{u}_L \not{\partial} u_L$ and in the electromagnetic and the Z current in equation 1.87. The only place where they remain are the $W^{\pm}$ currents:

$$\bar{u}_L^i \gamma^\mu d_L^i \rightarrow \bar{u}_L^i \gamma^\mu \left(U_u^\dagger U_d\right)^{ij} d_L^j \tag{1.101}$$

and

$$\bar{d}_L \gamma^\mu u_L \rightarrow \bar{d}_L^i \left(U_d^\dagger U_u\right)^{ij} \gamma^\mu u_L^j. \tag{1.102}$$

This leads to the definition of V_{CKM}:

$$V_{\mathrm{CKM}} \equiv \left(U_u^\dagger U_d\right) \tag{1.103}$$

which is the matrix transforming the mass-eigenstates of the down-type quarks to the weak-interaction isospin partners of the up-type quarks (the weak eigenstates). All the Standard Model predicts about this matrix is, that it is unitary; its entries can be complex and therefore violate $\mathcal{CP}$.

One Complex Phase

A general three-dimensional unitary matrix has 9 parameters, which can be taken as 3 real Euler angles plus 6 phases. We can remove phases by re-defining the quark basis:

$$q_L^i \rightarrow e^{i\delta^i} q_L^i \tag{1.104}$$

With three generations of quarks, there are six phases we can change. One over-all phase cancels in equations 1.101 and 1.102, leaving five relative phases that will effect V_{CKM}. We can therefore remove all complex phases from V_{CKM}, except one. This one phase uniquely parametrises $\mathcal{CP}$ violation in the Standard Model.

For the case of two generations, by the same argument, all phases can be rotated away from the mixing matrix, leaving no room for $\mathcal{CP}$ violation. This consideration led Kobayashi and Maskawa in 1973 to the conclusion that a third generation of particles was needed to accommodate $\mathcal{CP}$ violation in the electroweak theory [KM73]. The extension of Cabibbo's concept of a mixing angle to quark mixing in three or more generations, is therefore usually referred to as the Cabibbo–Kobayashi–Maskawa (CKM) matrix, V_{CKM}.

The parametrisation for the CKM matrix favoured by the Particle Data Group [G$^+$00]

$$V_{\text{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} \tag{1.105}$$

with $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$ for generation labels $i, j = (1, 2), (1, 3), (2, 3)$.

is given in equation 1.105, with the three real angles $\theta_{12}, \theta_{13}, \theta_{23}$ and the complex phase δ_{13}.

1.4.8 Structure

It is useful to label the elements of the CKM matrix as follows:

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{1.106}$$

Then the transition amplitude from a d to a u quark is proportional to V_{ud}, the transition amplitude from a u to a d quark to V_{ud}^*, etc.

Experimentally, it is found that the magnitudes of the CKM matrix elements follow a clear structure. In terms of the sine of the Cabibbo angle, $\lambda \equiv \sin\theta_C = 0.22$, the order of magnitude of the CKM matrix elements is:

$$\begin{pmatrix} 1 & \lambda & \lambda^3 \\ \lambda & 1 & \lambda^2 \\ \lambda^3 & \lambda^2 & 1 \end{pmatrix}. \tag{1.107}$$

This observation led Wolfenstein to a parametrisation of the CKM matrix as a power series in the parameter λ, which is, up to $\mathcal{O}(\lambda^3)$, given by [Wol83]:

$$\begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3\left(\rho - i\eta\right) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3\left(1 - \rho - i\eta\right) & -A\lambda^2 & 1 \end{pmatrix}. \tag{1.108}$$

where A, ρ, η are real parameters of order 1. The $\mathcal{CP}$ violating phase is parametrised by η. Up to $\mathcal{O}(\lambda^3)$, the only complex elements in this parametrisation are V_{td} and V_{ub}. In terms of the parametrisation given in equation 1.105, this approximation is given by:

$$\begin{pmatrix} 1 - \frac{1}{2}s_{12}^2 & s_{12} & s_{13}e^{-i\delta_{13}} \\ -s_{12} & 1 - \frac{1}{2}s_{12}^2 & s_{23} \\ s_{12}s_{23} - s_{13}e^{-i\delta_{13}} & -s_{23} & 1 \end{pmatrix}. \tag{1.109}$$

1.4.9 Unitarity Triangles

The requirement that the CKM matrix is unitary:

$$V_{\text{CKM}}V_{\text{CKM}}^{\dagger} = \mathbf{1} \tag{1.110}$$

leads to 9 conditions, which are automatically fulfilled in any of the above parametrisations. Six of them can be expressed in so-called unitarity triangles, which are three complex numbers adding up to zero displayed in the complex plane. Below, the six unitarity relations leading to the six triangles are listed, together with an indication of the length of each side in orders of λ:

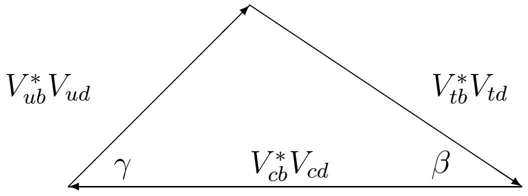
$$\begin{array}{lccccccc}
1) & V^*_{ud}V_{us} & + & V^*_{cd}V_{cs} & + & V^*_{td}V_{ts} & = & 0 \\
 & \mathcal{O}(\lambda) & & \mathcal{O}(\lambda) & & \mathcal{O}(\lambda^5) & & \\
2) & V^*_{cd}V_{ud} & + & V^*_{cs}V_{us} & + & V^*_{cb}V_{ub} & = & 0 \\
 & \mathcal{O}(\lambda) & & \mathcal{O}(\lambda) & & \mathcal{O}(\lambda^5) & & \\
3) & V^*_{us}V_{ub} & + & V^*_{cs}V_{cb} & + & V^*_{ts}V_{tb} & = & 0 \\
 & \mathcal{O}(\lambda^4) & & \mathcal{O}(\lambda^2) & & \mathcal{O}(\lambda^2) & & \\
4) & V^*_{cd}V_{td} & + & V^*_{cs}V_{ts} & + & V^*_{cb}V_{tb} & = & 0 \\
 & \mathcal{O}(\lambda^4) & & \mathcal{O}(\lambda^2) & & \mathcal{O}(\lambda^2) & & \\
5) & V^*_{ud}V_{td} & + & V^*_{us}V_{ts} & + & V^*_{ub}V_{tb} & = & 0 \\
 & \mathcal{O}(\lambda^3) & & \mathcal{O}(\lambda^3) & & \mathcal{O}(\lambda^3) & & \\
6) & V^*_{ub}V_{ud} & + & V^*_{cb}V_{cd} & + & V^*_{tb}V_{td} & = & 0 \\
 & \mathcal{O}(\lambda^3) & & \mathcal{O}(\lambda^3) & & \mathcal{O}(\lambda^3) & &
\end{array} \tag{1.111}$$

If there is no $\mathcal{CP}$ violation, the triangles all degenerate to lines. Describing $\mathcal{CP}$ violation in terms of unitarity triangles has the advantage that changing the parametrisation of the CKM matrix, and hence the phase-convention for the quarks, simply rotates the whole triangle in the complex plane, but leaves the side lengths and the relative angles inside the triangle unchanged. Therefore the unitarity triangles are a convention–independent way of parametrising $\mathcal{CP}$ violation in the Standard Model.

The area of the all unitarity triangles is the same and is the geometric, convention-independent equivalent of the single complex phase in the CKM matrix [Jar85]:

$$\text{Area of all triangles} = \frac{1}{2}\,|J| \tag{1.112}$$

Figure 1.1: *The* Unitary Triangle



with

$$J \sum_{m,n=1}^{3} \epsilon_{ikm}\epsilon_{jln} = \mathsf{Im}\Big(V_{ij}V_{kl}V^*_{il}V^*_{kj}\Big) \tag{1.113}$$
$$\text{for any } i,j,k,l \in \{1,2,3\}$$

1.4.10 *The* Unitarity Triangle and the Angles γ, β

Most of the triangles have very unequal sides. To measure $\mathcal{CP}$ violation in decays related to such triangles, for example in the K^0 system associated to triangle number 1, is very difficult: the two long sides have little relative phase difference and therefore little $\mathcal{CP}$ violating effects; the third side might have a larger phase difference relative to the long sides, but it also corresponds to very small branching fractions.

Only in the last two triangles are all sides of the same order of magnitude. Both are related to observables in decays of neutral B mesons: number 5 to B^0_s decays and number 6 to B^0_d decays. Up to $\mathcal{O}(\lambda^3)$ in the Wolfenstein parametrisation, the two triangles coincide. Therefore triangle number 6 is called *The* Unitary Triangle. This is shown in figure 1.1, where also the angles γ and β are defined. The parameter α that is often found in the literature, corresponds to the third angle in figure 1.1; $\alpha \equiv \pi - \beta - \gamma$. In the Wolfenstein parametrisation up to $\mathcal{O}(\lambda^3)$ only V_{ub} and V_{td} are complex. Dividing each entry of the matrix by its modulus gives:

$$\begin{pmatrix} 1 & 1 & e^{-i\gamma} \\ 1 & 1 & 1 \\ e^{-i\beta} & 1 & 1 \end{pmatrix}. \tag{1.114}$$

This relates the complex phases of the CKM matrix in Wolfenstein parametrisation directly to the convention-independent description in terms of unitarity triangles. Up to $\mathcal{O}(\lambda^3)$, β and γ are the only non-zero phases in the CKM matrix, and only decays involving $b \to u$ or $d \to t$ transitions can violate $\mathcal{CP}$. Both are accessible in the B^0_d system, and both can to be measured at LHCb. How β and γ relate to observables in the B^0_d system will be discussed in section 1.4.13.

1.4.11 Existing Constraints on the CKM Matrix

Measurements Without CP Violation

The magnitudes of eight of the nine entries of the CKM matrix have been measured directly, with varying precision. All numbers are taken from [G⁺00]:

1. $|V_{ud}| = 0.9735 \pm 0.0008$ from nuclear β decay.
2. $|V_{us}| = 0.2196 \pm 0.0023$ from semileptonic kaon and hyperon decays.
3. $|V_{cd}| = 0.224 \pm 0.016$ from charm production by neutrino and antineutrino beams off valence d-quarks.
4. $|V_{cs}| = 1.04 \pm 0.16$ from semileptonic D decays.
5. $|V_{cb}| = 0.0402 \pm 0.0019$ from semileptonic B decays.
6. $|V_{ub}/V_{cb}| = 0.090 \pm 0.025$ from semileptonic B decays.
7. $\frac{|V_{tb}|^2}{|V_{td}|^2+|V_{ts}|^2+|V_{tb}|^2} = 0.99 \pm 0.29$ from semileptonic top decays.
8. $|V_{cs}| = 0.9891 \pm 0.016$ from the ratio of hadronic $\mathrm{W}^\pm$ decays to leptonic decays. Assuming three generations, this measurement tests unitarity and does not constrain $|V_{cs}|$.

For an arbitrary number of generations, but still using unitarity of the expanded matrix, this results in the following 90% confidence limits on the matrix elements:

$$\begin{pmatrix} 0.9722-0.9748 & 0.216-0.223 & 0.002-0.005 & \cdots \\ 0.209-0.228 & 0.959-0.976 & 0.037-0.043 & \cdots \\ 0.00-0.09 & 0.00-0.16 & 0.070-0.993 & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix} \tag{1.115}$$

Assuming three generations, and excluding result number 8:

$$\begin{pmatrix} 0.9742-0.9757 & 0.219-0.226 & 0.002-0.005 \\ 0.219-0.225 & 0.9734-0.9749 & 0.037-0.043 \\ 0.004-0.014 & 0.035-0.043 & 0.9990-0.9993 \end{pmatrix} \tag{1.116}$$

From the measurement of the B^0_d and B^0_s oscillation period, the mass differences between the mass eigenstates in each system are found to be:

$$
\begin{aligned}
\Delta m_{\mathrm{B}^0_\mathrm{d}} &= 0.473 \pm 0.016\,\mathrm{ps}^{-1} && (1.117)\\
\Delta m_{\mathrm{B}^0_\mathrm{s}} &> 14.3\,\mathrm{ps}^{-1} \text{ at } 95\% \text{ C.L.} && (1.118)
\end{aligned}
$$

This yields to the following constraints:

$$|V^*_{tb}V_{td}| = 0.0083 \pm 0.0016 \qquad (1.119)$$

and

$$|V_{td}|\,/\,|V_{ts}| < 0.24. \qquad (1.120)$$

In the case of three generations, $|V_{ts}| \approx |V_{cb}|$ and

$$|V_{td}| < 0.1. \qquad (1.121)$$

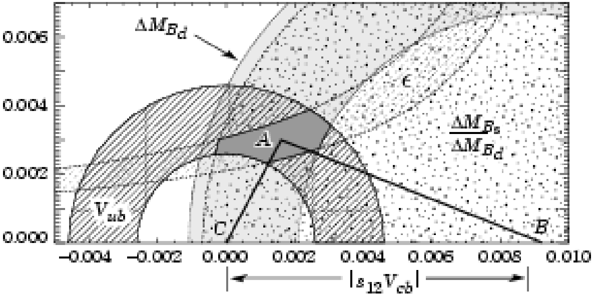
CP Violation Measurements

The unitarity triangle provides an elegant way to combine the constraints from the measurements of the magnitudes of the CKM matrix elements with the results from $\mathcal{CP}$ violation experiments.

The $\mathcal{CP}$ violation parameter ϵ_K from measurements in the K^0 system restricts, in principle, the apex of the unitarity triangle to a hyperbola in the complex plane. Due to theoretical uncertainties in calculating hadronic effects, the interpretation of ϵ_K in terms of CKM parameters is difficult, and the constraints on the unitarity triangle from ϵ_K are weak despite its precise value.

Figure 1.2 shows the constraints on the apex of the unitarity triangle as shown in [G+00] in July 2000. At that time, the best measurement of $\sin(2\beta)$ came from CDF: $\sin(2\beta) = 0.79^{+0.41}_{-0.44}$ [A+00a]. Since then, the first results from the B^0 factories have been released. The current status is:

Figure 1.2: Constraints on the apex of the Unitarity Triangle from the measurements of the sides, and from ϵ_K, taken from [G+00].



$\sin(2\beta) =$

- $0.58^{+0.32}_{-0.34}(\mathrm{stat})^{+0.09}_{-0.10}(\mathrm{sys})$, BELLE [A+01a]
- $0.34 \pm 0.20(\mathrm{stat}) \pm 0.05(\mathrm{sys})$, BaBar[A+01b]

each showing a 1.7σ effect for $\mathcal{CP}$ violation in the B^0_d system.

1.4.12 Status in 2006

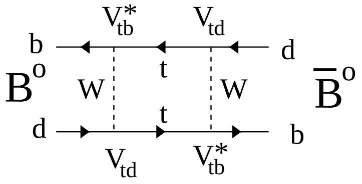
By the year 2006, accurate measurements of $\sin(2\beta)$ are expected from the B^0 factories and the Tevatron. With $500\,\mathrm{fb}^{-1}$ at the $\Upsilon(4S)$ and a similar performance from CDF and DØ, the precision on $\sin(2\beta)$ is expected to be ~ 0.02.

Such a precision is not expected for the measurement of γ. The branching ratios of the relevant decay channels are either very small, typically $\mathcal{O}(10^{-5})$, or $\mathcal{CP}$ violating effects are tiny, as in $B^0_d \to D^*\pi$ with $|\eta| \sim \lambda^2$. Therefore very high statistics are required. The B^0_s system that also allows measurements of γ is not accessible to the B^0-factories if they continue to run at the $\Upsilon(4S)$. While first estimates of γ might be possible, the uncertainties are expected to be too large to give strong constraints on the unitarity triangle.

1.4.13 β and γ in the B^0_d System

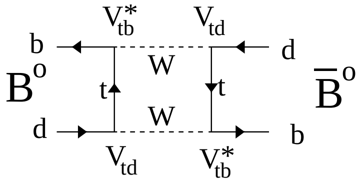
The Angle β

The angle β appears via $V_{td} = |V_{td}|\, e^{-i\beta}$ in the box diagrams for $B^0_d - \overline{B^0_d}$ mixing:



$$\propto e^{-i2\beta} \tag{1.122}$$

and



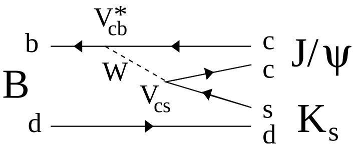
$$\propto e^{-i2\beta}. \tag{1.123}$$

In principle, all up-type quarks (u, c, t) contribute to the internal quark lines; but as the amplitude is proportional to the mass of the internal quark, the top quark contribution dominates entirely.

In terms of the parameters introduced in section 1.3.4:

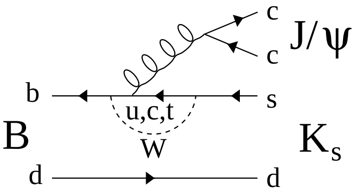
$$\frac{q}{p} = e^{-i2\beta}. \tag{1.124}$$

The decay $B^0_d \to J/\psi K_S$ is often referred to as the "gold–plated decay". It allows a clean extraction of the mixing angle ϕ_{mix}, as both the tree-level diagram:



$$\propto e^{0} \tag{1.125}$$

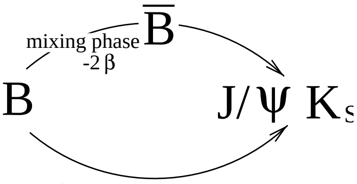
and the penguin contribution to the decay have the same phase:



$$\propto e^{0} \tag{1.126}$$

where the non-zero phase in the penguin contribution due to the $b \to u$ transition can be neglected as it is doubly Cabibbo suppressed relative to the others.

The mixing phase ϕ_{mix} enters as the phase difference between the two interfering decay paths:



$J/\psi K_S$ is a $\mathcal{CP} = -1$ state, and there is no phase in the amplitude $B^0_d \to J/\psi K_S$.

Therefore the parameter η defined in equation 1.57, is given by

$$\eta_{\mathrm{J}/\psi K_S} = \frac{q}{p} \cdot (-1) = -e^{i\phi_{\mathrm{mix}}} = -e^{-i2\beta}. \tag{1.127}$$

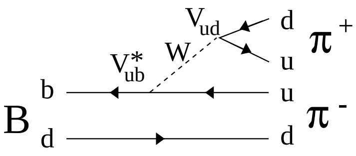
So the time dependent asymmetry is:

$$\begin{aligned} A(\tau) &= \\ &\frac{\Gamma\left(\mathrm{B^0_d} \to \mathrm{J}/\psi\mathrm{K_S}\right) - \Gamma\left(\overline{\mathrm{B^0_d}} \to \mathrm{J}/\psi\mathrm{K_S}\right)}{\Gamma\left(\mathrm{B^0_d} \to \mathrm{J}/\psi\mathrm{K_S}\right) + \Gamma\left(\overline{\mathrm{B^0_d}} \to \mathrm{J}/\psi\mathrm{K_S}\right)} \\ &= -\sin(-2\beta)\sin(\Delta m\tau) \end{aligned} \tag{1.128}$$

where Δm is the mass difference between the two $\mathrm{B^0_d}$ mass eigenstates and τ is the decay eigentime.

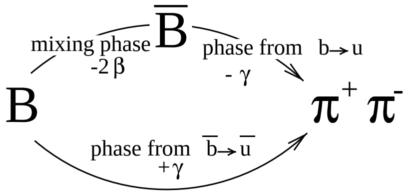
The Angle γ

The angle γ enters whenever a decay involves a $b \to u$ transition, the most prominent example being the decay to the $\mathcal{CP}$ eigenstate $\pi^+\pi^-$:



$$\propto e^{i\gamma} \tag{1.129}$$

Again, the measured quantity is the phase-difference between the two interfering decay paths. In the decay $\mathrm{B^0_d} \to \pi^+\pi^-$ this is, at tree-level, $2(\beta + \gamma)$:



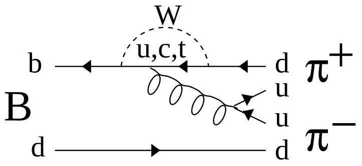
$\pi^+\pi^-$ is a $\mathcal{CP} = +1$ state, and hence η is given by

$$\eta_{\pi\pi} = e^{\phi_{\mathrm{mix}}} \cdot \left(+e^{-i2\gamma}\right) = e^{-i2(\beta+\gamma)}. \tag{1.130}$$

So the time dependent asymmetry is:

$$\begin{aligned} A(\tau) &= \\ &\frac{\Gamma\left(\mathrm{B^0_d} \to \pi^+\pi^-\right) - \Gamma\left(\overline{\mathrm{B^0_d}} \to \pi^+\pi^-\right)}{\Gamma\left(\mathrm{B^0_d} \to \pi^+\pi^-\right) + \Gamma\left(\overline{\mathrm{B^0_d}} \to \pi^+\pi^-\right)} \\ &= \sin(-2(\beta+\gamma))\sin(\Delta m\tau) \end{aligned} \tag{1.131}$$

However, in interpreting the measured phase-difference as $2(\beta+\gamma)$, we have neglected penguin contributions:



(1.132)

but here, in contrast to the case of $\mathrm{B^0_d} \to \mathrm{J}/\psi\mathrm{K_S}$, they do play a significant role. If the relative size of the penguin contributions were known, the result could be interpreted in terms of β and γ, but this is difficult to estimate from first principles. Several techniques have been proposed to measure the penguin contributions [LNQS91, B+00], for example by combining this measurement with results from other, flavour-symmetry related decays [Fle00].

In this thesis though, we will instead look at a decay that does not suffer from penguin contributions and therefore allows a theoretically clean extraction of the Standard Model γ. This decay is $\mathrm{B^0_d} \to \mathrm{D}^{*-}\pi^+$.

1.4.14 Entry Points for New Physics

Most extensions of the Standard Model are at high mass scales ($\gg M_W$). Significant effects at tree–level are therefore unlikely, but may enter as contributions to the higher order processes which play an important role in many $\mathcal{CP}$ violation measurements in the B^0 system. A particular example is B^0_d–$\overline{\mathrm{B}^0_\mathrm{d}}$ mixing which is mediated by a box diagram. Flavour Changing Neutral Currents, which appear in many New Physics models, could affect the mixing phase ϕ_mix measured with $\mathrm{B}^0_\mathrm{d} \rightarrow \mathrm{J}/\psi \mathrm{K_S}$. Generally, as $\mathcal{CP}$ violation is an interference effect, higher order contributions due to New Physics can affect the measured phase-differences. This is true especially for those decays where loop-corrections are already important in the Standard Model. Decays with large penguin contributions are therefore particularly sensitive to New Physics, one example being $\mathrm{B}^0_\mathrm{d} \rightarrow \pi^+\pi^-$.

1.4.15 The B^0_s System

The LHC will not only produce a large number of B^0_d mesons, it will also be a copious source of B^0_s mesons. As the angle γ enters via $b \rightarrow d$ transitions, it is as accessible in the B^0_s system as it is in the B^0_d system, effectively doubling the number of potential decay channels to overconstrain γ. Having access to both the B^0_s and the B^0_d system also provides for a very promising way of constraining penguin contributions by combining the analysis of flavour–symmetry related decays, such as $\mathrm{B}^0_\mathrm{d} \rightarrow \pi^+\pi^-$ and $\mathrm{B}^0_\mathrm{s} \rightarrow \mathrm{K}^+\mathrm{K}^-$ [Fle00].

The equivalent of the 'gold plated' decay $\mathrm{B}^0_\mathrm{d} \rightarrow \mathrm{J}/\psi \mathrm{K_S}$ in the B^0_d system is the channel $\mathrm{B}^0_\mathrm{s} \rightarrow \mathrm{J}/\psi\phi$ in the B^0_s system. It gives a clean measurement of the B^0_s–$\overline{\mathrm{B}^0_\mathrm{s}}$ mixing phase ϕ^s_mix. Up to $\mathcal{O}(\lambda^3)$ in the Wolfenstein parametrisation, this phase is zero. Finding a large mixing phase in the B^0_s system would therefore be a sign of New Physics. Expanding the CKM matrix up to $\mathcal{O}(\lambda^4)$ introduces a small complex phase to V_{ts} of $\mathcal{O}(\lambda^2)$, referred to as $-\delta\gamma$. The B^0_s-$\overline{\mathrm{B}^0_\mathrm{s}}$ mixing phase in the Standard Model is therefore:

$$\phi^s_\mathrm{mix} = -2\delta\gamma \tag{1.133}$$

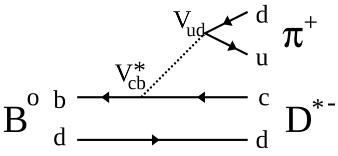
The measurement of γ, $\delta\gamma$ and other parameters in the B^0_s system requires sensitivity to the very rapid B^0_s oscillation period of $< 0.44\,\mathrm{ps}$.

1.5 Extracting the Angle γ from $\mathrm{B}^0_\mathrm{d} \rightarrow \mathrm{D}^*\pi$

With the decay channels $\mathrm{B}^0_\mathrm{d} \rightarrow \mathrm{D}^*\pi$, the angle γ can be measured in a theoretically clean way. The result is easy to interpret in terms of Standard Model parameters as $\mathrm{B}^0_\mathrm{d} \rightarrow \mathrm{D}^*\pi$ does not suffer from penguin contribution. For the same reason, the measurement is expected to be rather unaffected by New Physics. Therefore, the γ measurement from $\mathrm{B}^0_\mathrm{d} \rightarrow \mathrm{D}^*\pi$ provides a benchmark that other γ measurements which are more sensitive to non-Standard Model processes, can be compared against.

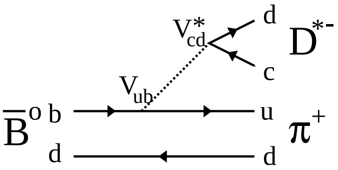
$\mathrm{D}^*\pi$ is not a $\mathcal{CP}$ eigenstate, but as both B^0_d and $\overline{\mathrm{B}^0_\mathrm{d}}$ can decay to the same final state $\mathrm{D}^{*+}\pi^-$, a $\mathcal{CP}$ violation measurement is possible, as shown in section 1.3.6.

The relevant decay amplitudes are:



$$\propto e^0 \qquad (1.134)$$

and



$$\propto e^{-i\gamma}. \qquad (1.135)$$

If we take, using the notation of section 1.3.4, $\mathrm{D}^{*-}\pi^+ \equiv f$ and $\mathrm{D}^{*+}\pi^- \equiv \bar{f}$, then the parameters η and $\overline{\eta}$, defined in equations 1.57 and 1.61, are given by:

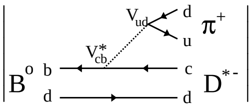
$$\eta = |\eta|\, e^{i\left(\phi_\mathrm{mix} - \gamma + \Delta_\mathrm{qcd}\right)} \qquad (1.136)$$

and

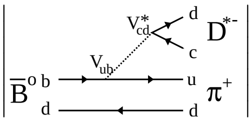
$$\overline{\eta} = |\eta|\, e^{i\left(-(\phi_\mathrm{mix} - \gamma) + \Delta_\mathrm{qcd}\right)} \qquad (1.137)$$

so the $\mathcal{CP}$ violating phase is $\delta_{CP} = \phi_\mathrm{mix} - \gamma$. The mixing-induced phase ϕ_mix will be measured very accurately from the channel $\mathrm{B}^0_\mathrm{d} \to \mathrm{J}/\psi\mathrm{K}_\mathrm{S}$, so γ can be extracted. In the Standard Model, $\phi_\mathrm{mix} = -2\beta$, but even if New Physics enters in the mixing, the ϕ_mix in the $\mathrm{B}^0_\mathrm{d} \to \mathrm{D}^*\pi$ measurement and the ϕ_mix measured in $\mathrm{B}^0_\mathrm{d} \to \mathrm{J}/\psi\mathrm{K}_\mathrm{S}$ are the same quantity, and the γ extraction is unaffected.

As one of the decay amplitudes is doubly Cabibbo suppressed relative to the other:

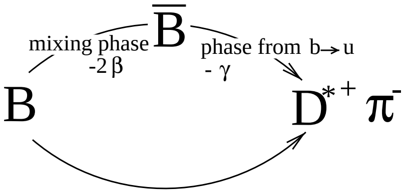


$$\sim G_F \lambda^2 \qquad (1.138)$$



$$\sim G_F \lambda^4 \qquad (1.139)$$

the interference effects between the two decay paths:



are tiny. Therefore, meaningful sensitivity requires high statistics and pure samples.

As shown in section 1.3.6, in principle all unknown quantities, $\left\{|\eta|^2, \mathsf{Im}(\eta), \mathsf{Im}(\overline{\eta})\right\}$, can be extracted from fits to the two asymmetries A_η and $A_{\overline{\eta}}$. This can then be used to calculate $\phi_\mathrm{mix} - \gamma$ and Δ_qcd from

$$\sin(\Delta_\mathrm{qcd} + (\phi_\mathrm{mix} - \gamma)) = \frac{\mathsf{Im}(\eta)}{|\eta|} \qquad (1.140)$$

$$\sin(\Delta_\mathrm{qcd} - (\phi_\mathrm{mix} - \gamma)) = \frac{\mathsf{Im}(\overline{\eta})}{|\eta|} \qquad (1.141)$$

In practice though, the smallness of $|\eta|^2 \sim \lambda^4$ makes fitting it impossible for the number of events expected at LHCb.

Therefore, in the following it is assumed that $|\eta|$ can be obtained from elsewhere with reasonable accuracy. For the purpose of the feasibility study presented in chapter 3, a reasonable value for $|\eta|$ has to be assumed. The simplest assumption regarding the hadronic effects is, that they cancel in the ratio of the ampltidues. Then $|\eta|$ is given by the magnitude of CKM matrix elements:

$$|\eta| = \frac{|V_{ub}|\,|V_{cd}|}{|V_{cb}|\,|V_{ud}|}. \tag{1.142}$$

Using results 1, 3 and 6 in section 1.4.11 gives $|\eta| = 0.021$. This value is assumed throughout this thesis.

1.6 Conclusion

The Standard Model is highly predictive with respect to $\mathcal{CP}$ violation, which it parametrises by a single complex phase δ_{13}. This makes $\mathcal{CP}$ violation measurements particularly suitable to test the Standard Model and search for signs of New Physics. Furthermore, $\mathcal{CP}$ violation and the generation of fermion masses are inextricably linked in the Standard Model via the CKM matrix. Measuring $\mathcal{CP}$ violation parameters might give us insight into the underlying processes of fermion mass generation and flavour mixing, the largest sources of free parameters in the Standard Model.

Although $\mathcal{CP}$ violation has been observed for the first time in 1964, it is still unclear whether the Standard Model explanation of it is correct. The only place where $\mathcal{CP}$ violation has been unambiguously observed is still the neutral K^0 system. Because of theoretical uncertainties in the interpretation of the measured $\mathcal{CP}$ violation parameters in the K^0 system, they do not considerably constrain the Standard Model. However, from these measurements and our current knowledge of the CKM matrix, the Standard Model predicts $\mathcal{CP}$ violation in the B^0 system to be large.

$\mathcal{CP}$ violation can be observed as an interference effect between decay amplitudes that involve CKM matrix elements with different phases. Up to $\mathcal{O}(\lambda^3)$ in the Wolfenstein parametrisation, all CKM matrix elements have zero phase, except for the two smallest entries, $V_{td} = |V_{td}|\,e^{-i\beta}$ and $V_{ub} = |V_{ub}|\,e^{-i\gamma}$. Both are accessible in the $\mathrm{B_d^0}$ system. The phase β appears in the box diagram for $\mathrm{B_d^0}$–$\overline{\mathrm{B_d^0}}$ mixing; this phase will be well measured by the year 2006 by the B^0 factories and the Tevatron. The other phase is γ, on which no strong constraints are expected before LHCb starts data taking.

The phase γ appears in all $\mathrm{B_d^0}$ and $\mathrm{B_s^0}$ decays that involve $b \to u$ transitions. Typical branching fractions for those decays are $\sim 10^{-5}$. Measuring γ in many different decay channels will thoroughly over-constrain the Standard Model parametrisation of $\mathcal{CP}$ violation. Some of these channels are more likely to be affected by New Physics than others, and comparison can reveal New Physics. In order to perform these measurements, high statistics and clean data samples are essential. This requirement on sample purity provides the motivation for the RICH system at LHCb, which is discussed in detail in this thesis.

The decay $\mathrm{B_d^0} \to \mathrm{D}^*\pi$ is an example for a γ sensitive channel that is expected to be robust against New Physics. As it does not suffer from penguin contributions, the measurement is easy to interpret in terms of Standard Model parameters. $B_d^0 \rightarrow D^*\pi$ will therefore provide a benchmark measurement of the Standard Model γ. Comparing this with other measurements of γ will test the Standard Model description of quark mixing and $\mathcal{CP}$ violation.

In chapter 3, the feasibility of a γ measurement with the channel $B_d^0 \rightarrow D^*\pi$ at LHCb is investigated, using a detailed Monte Carlos simulation of the LHCb detector.

Chapter 2

The LHCb experiment

2.1 The LHC

The planned Large Hadron Collider (LHC) at CERN will collide protons at a centre–of–mass energy of 14 TeV at a design luminosity of $\sim 10^{34}\,\mathrm{cm}^{-1}\mathrm{s}^{-1}$. At these energies, direct observation of New Physics at mass–scales of ~ 1 TeV is feasible. The LHC will also be the most copious source of B hadrons in the world.

The accelerator will be housed in the 27 km tunnel that has been built for the LEP experiment, that is presently being dismantled. LHC is scheduled to start data taking in 2006 with a luminosity of $10^{33}\,\mathrm{cm}^{-1}\mathrm{s}^{-1}$ and upgrade to its full luminosity after a few years. The high luminosity will be achieved by filling each of the two rings with 2835 bunches of 10^{11} particles each [Gro95]; the interval between successive bunch crossings at each of the four interaction points is 25 ns.

Four major experiments will take data at the LHC: ALICE, CMS, ATLAS and LHCb. The high luminosity and the nature of hadron collisions result in high radiation levels near the interaction points, requiring radiation–hard detector components. The bunch–crossing rate of 40 MHz imposes high demands on the speed of the detector elements and readout electronics for all experiments.

A diagram of the LHC with the four experiments is shown in figure 2.1.

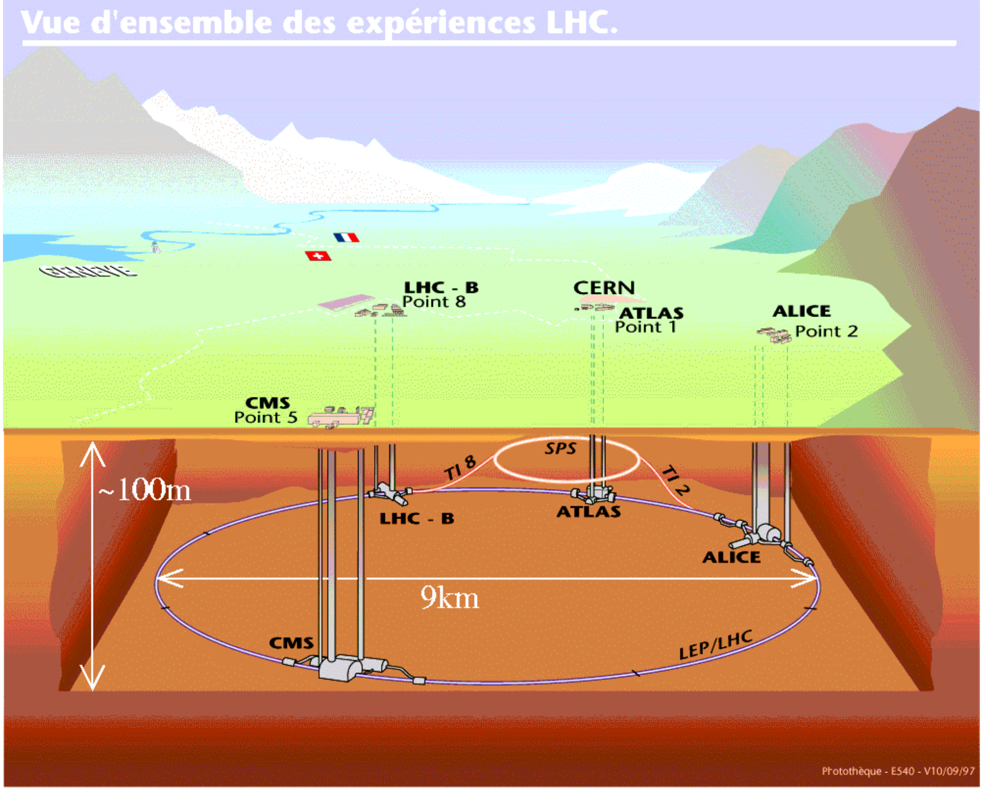
2.2 LHCb

LHCb is specifically designed to make best use of the large number of $b\bar{b}$ pairs produced at the LHC. This large number is a consequence of the huge $b\bar{b}$ cross section for 14 TeV proton–proton collisions of $\sim 500\,\mu$barn [N+00].

The detector is a single arm spectrometer, which is motivated by the kinematics of $b\bar{b}$ production in high energy proton–proton collisions, as discussed below.

The LHCb experiment re–uses the DELPHI pit at interaction point 8, which constrains the total length of the detector to ~ 20 m.

Figure 2.1: LHC and LHCb. The LHC collider is housed in the former LEP tunnel. LHCb is accommodated at interaction point 8.



2.2.1 Bottom Production at LHCb

In proton–proton collisions, $b\bar{b}$ pairs are produced by flavour creation, flavour excitation and gluon splitting [N+00, Bar01]. A typical flavour creation mechanism, gluon fusion, is illustrated in figure 2.2.

At LHC energies, the parton distribution functions of the proton are such that it is most likely that partons with very different momenta interact [Bar01]. This has two major consequences for the design of a dedicated B physics detector:

Figure 2.2: Typical diagram for $b\bar{b}$ production in proton–proton collisions (from [N+00])

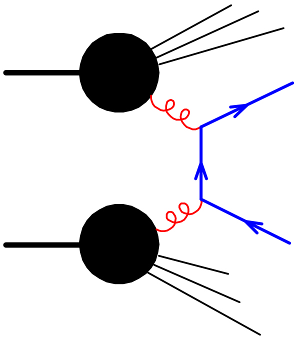
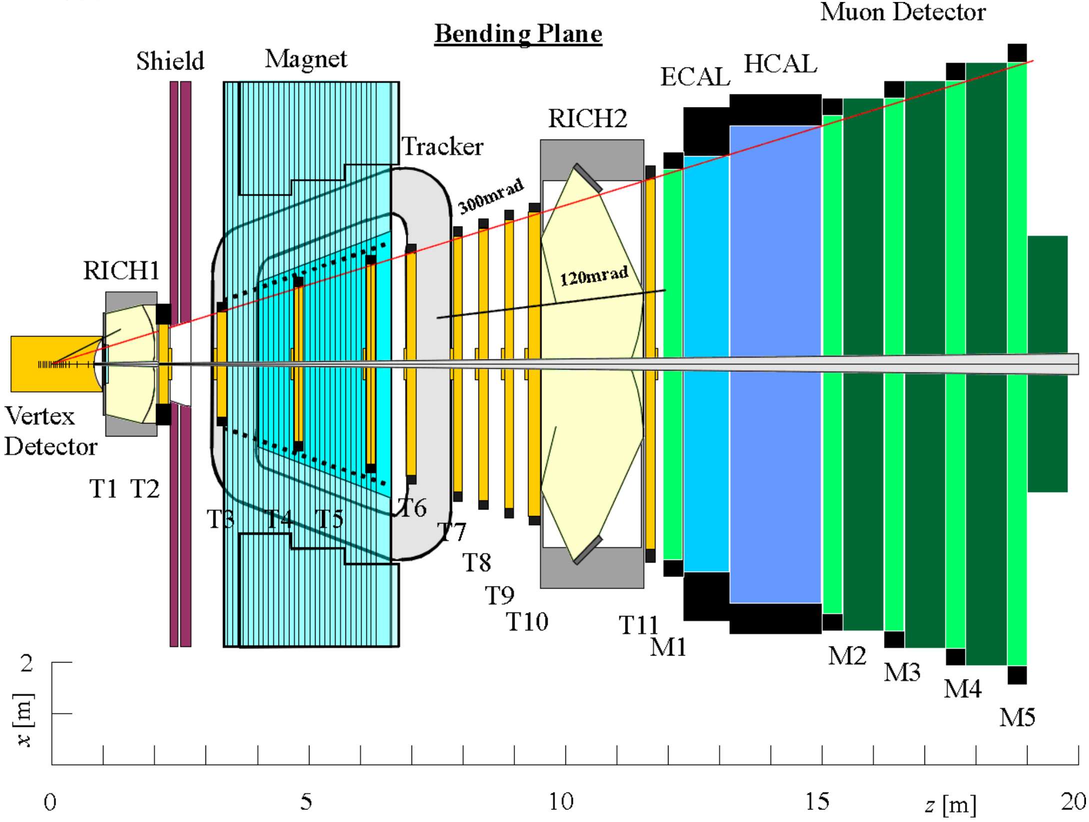


Figure 2.3: The LHCb Detector shown in the bending (x-z) plane. The angular acceptance is 300 mrad.

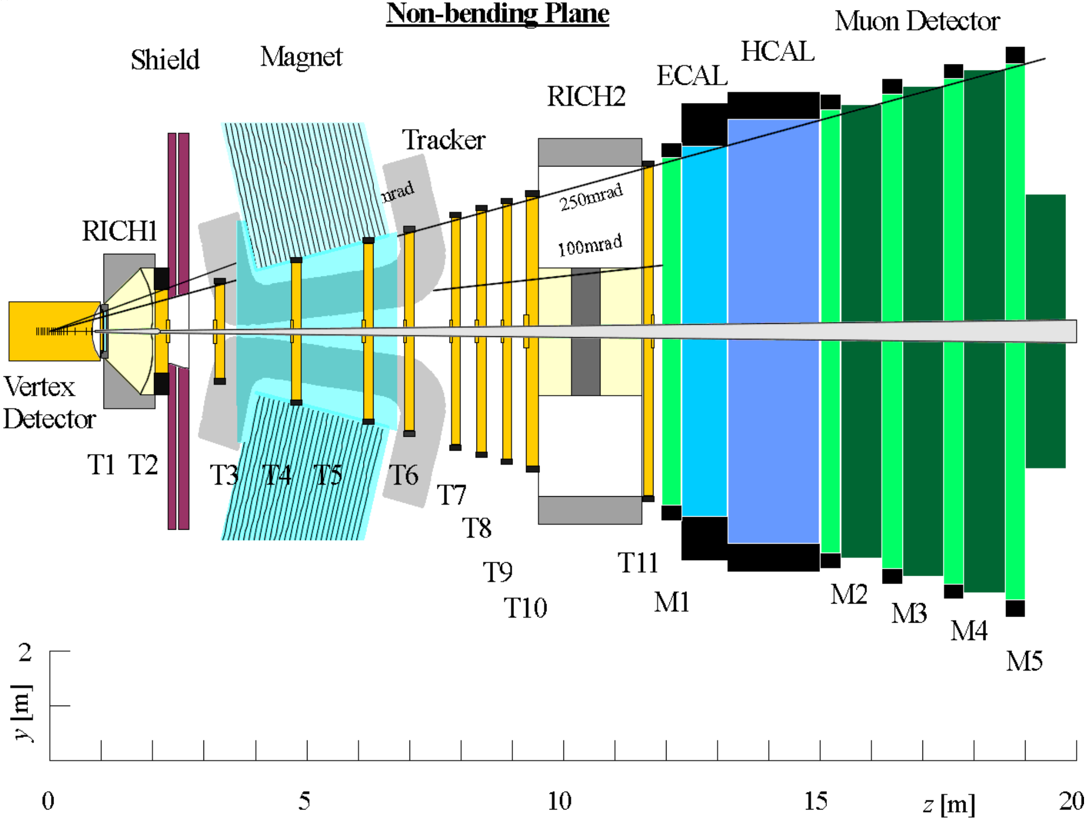


- The B hadrons produced are highly boosted.
- Both, the b and the $\bar{b}$, are predominantly produced in the same forward or backward cone.

Momentum spectra of B_d^0 mesons in $B_d^0 \to \pi^+\pi^-$ decays are shown in figure 2.5. For a lifetime of ~ 1 ps for B hadrons, the high momenta result in typical decay lengths of the order of 1 cm. Long decay lengths are important for the reconstruction of the B^0 decay time, which is needed for the measurement of the time–dependent decay rate asymmetries introduced in chapter 1.

The polar angle distribution of the $b\bar{b}$ pairs produced in 14 TeV proton–proton collisions, simulated with PYTHIA is shown in figure 2.6. The fact that in most cases both of the b quarks are produced within the same narrow forward cone motivates the LHCb detector geometry. Why it is important that *both* b quarks are within the LHCb acceptance will become clear in the following section.

Figure 2.4: The LHCb Detector shown in the non–bending (y-z) plane The angular acceptance is 250 mrad.



2.2.2 B^0–tagging

In order to measure the asymmetries given in equations 1.128 and 1.131, the flavour of the B^0–meson at the time of creation needs to be known. LHCb will predominantly use opposite–side tagging, where the flavour of the reconstructed B^0–meson at creation is estimated from the flavour of the accompanying B–hadron, (the “tagging B”) at decay. It is necessary to assume that the tagging B did not oscillate before decaying. The term “opposite side tagging” has historical roots: as shown above, at LHCb both b quarks are produced in the same forward cone, which makes this tagging method suitable for a single–arm spectrometer.

The flavour of the tagging B is found by finding one particle that is likely to come from a B–decay, but does not come from the reconstructed B^0. This particle is either a lepton with high transverse momentum, or a kaon with high transverse momentum and significant impact parameter with respect to the primary vertex [TP98]. The charge of the lepton gives the charge and therefore the flavour of the b–quark it decayed from.

Figure 2.5: Momentum distribution for $\mathrm{B_d^0}$ mesons in $\mathrm{B_d^0} \rightarrow \pi^+\pi^-$ decays. From [TP98].

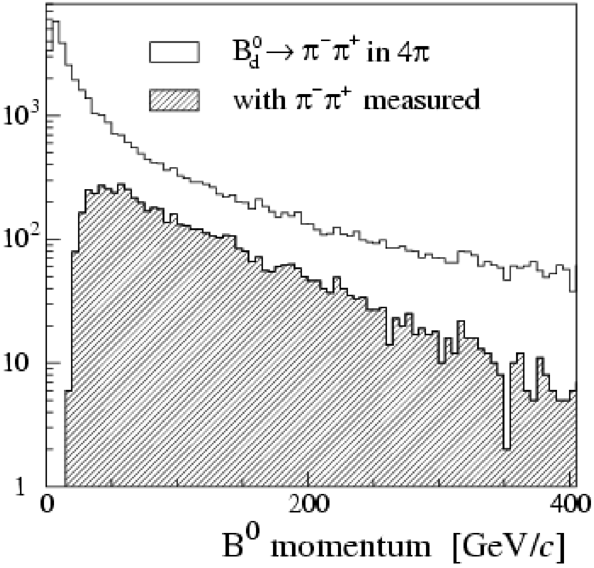
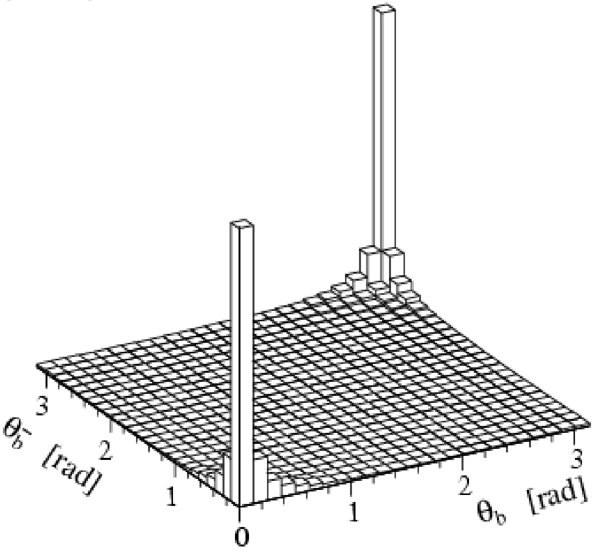


Figure 2.6: Polar angles of hadrons formed from $b\bar{b}$ pairs, calculated by PYTHIA from [TP98].



Conversely, the charge of the tagging kaon is related to the flavour of the b quark via the decay chain $b \rightarrow c \rightarrow s$.

Not all tags will identify the flavour of the reconstructed $\mathrm{B^0}$ meson correctly. One contribution to the mistag is for example the possibility that the tagging–B, which may be a $\mathrm{B_s^0}$ meson, has oscillated before decay. As shown in chapter 3, a wrong tag fraction of ω_{tag} increases the statistical error on the fit parameters in the decay rate asymmetries by a factor of $\frac{1}{(1-2\omega_{\mathrm{tag}})}$. The statistical error is also inversely proportional to the square–root of the number of events, and therefore to $1/\sqrt{\epsilon_{\mathrm{tag}}}$, where ϵ_{tag} is the tagging efficiency.

As a measure the tagging performance it is useful to introduce the “tagging power” P_{tag}:

$$P_{\mathrm{tag}} \equiv \epsilon_{\mathrm{tag}} \left(1 - 2\omega_{\mathrm{tag}}\right)^2 \qquad (2.1)$$

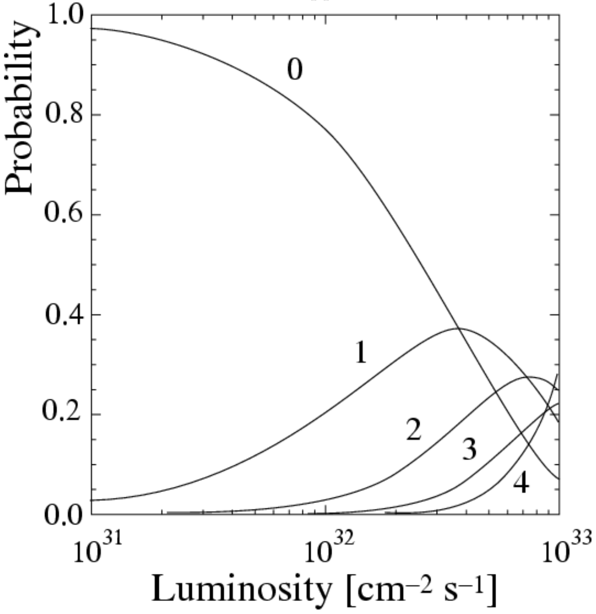
A sample with N_s events, a mistag fraction of ω_{tag}, and a tagging efficiency ϵ_{tag} is statistically equivalent to $P_{\mathrm{tag}} \cdot N_s$ perfectly tagged events.

At LHCb, a tagging efficiency $\epsilon_{\mathrm{tag}} = 40\%$, with a mistag fraction of $\omega_{\mathrm{tag}} = 30\%$ is expected [B+00], which corresponds to $P_{\mathrm{tag}} = 6.4\%$.

2.2.3 Luminosity

At the LHC design luminosity of $\sim 10^{34}\,\mathrm{cm^{-1}s^{-1}}$, most bunch crossings would involve more than one inelastic proton–proton interaction. Such multiple interactions severely complicate the task of $\mathrm{B^0}$–tagging, and of cleanly locating the

Figure 2.7: Probabilities for having $0, 1, 2, 3, 4$ inelastic pp interactions per bunch crossing as a function of the luminosity, assuming an inelastic cross section for pp interactions of $\sigma_{\mathrm{inel\,pp}} = 80\,\mathrm{mb}$ [TP98]



primary and secondary vertices. Both are essential for measuring time dependent decay rate asymmetries. LHCb's pile–up veto counter, which is part of the vertex detector discussed in section 2.2.6, therefore rejects multiple interactions at the trigger level. The luminosity at the LHCb interaction point is reduced by a special defocusing of the beam, in order to optimise it for single interaction events.

Figure 2.7 shows that the maximum number of single interactions can be achieved with a luminosity of about $\sim 4 \cdot 10^{32}\,\mathrm{cm}^{-1}\mathrm{s}^{-1}$, assuming an inelastic pp cross section of 80 mb. The design luminosity of LHCb is $2 \cdot 10^{32}\,\mathrm{cm}^{-1}\mathrm{s}^{-1}$, which results only in a small decrease of single interaction events, while reducing the number of multiple interactions by half. This improves the trigger performance and reduces both radiation damage and occupancy of detectors.

With this luminosity, LHCb expects about $5.6 \cdot 10^{11}$ single–interaction $b\bar{b}$ events passing the pile–up veto per year. Due to its comparably moderate luminosity requirements, LHCb can start its full physics programme from the first day of running.

2.2.4 Detector Overview

Amongst the most important features of the LHCb detector are:

- An excellent proper time resolution to exploit the full B physics potential at the LHC, including measurements in the rapidly oscillating B^0_s system.
- Particle identification by two Ring Imaging CHerenkov (RICH) counters, for clean data samples and flavour tagging with kaons.
- A dedicated B trigger, including high p_t hadron and lifetime triggers for high efficiency.

Figures 2.3 and 2.4 show schematic overviews of the LHCb detector, and also define the orientation of the co–ordinate system. The origin is at the interaction point, the z–axis along the beam line, the x–axis (horizontal) in the bending plane of the magnet, and the y axis (vertical) in the non–bending plane.

LHCb comprises a vertex detector system, which includes the pile–up veto counter; a magnet and a tracking system; two RICH counters; an electromagnetic calorimeter and a hadron calorimeter, and a muon detector. All detector sub–systems, except for the vertex detector and RICH 1, are split into two halves that can be separated horizontally for maintenance and access to the beam pipe.

Below, the different components are described briefly. The following sections are derived from the LHCb Technical Proposal (TP) [TP98], or, where available, the corresponding Technical Design Report (TDR). The TDR was available for the Magnet [Mag00], the Calorimeters [Cal00] and the RICH system [RIC00].

The specifications given in the Technical Proposal provided the basis for the simulation study presented in chapter 3. In case a TDR is available and the specifications differ significantly, this is indicated. As RICH–related studies constitute a large part of the research presented in this thesis, the RICH is described in more detail.

2.2.5 Beam Pipe

The beam pipe consists of three sections. A 1.8 m section around the interaction point with a large diameter of about 120 cm accommodates the vertex detector, placed inside. This is followed by a 1.5 m long conical section inside RICH 1, with a 25 mrad opening angle. The third part, which extends to the end of the detector, is 16 m long and has a 10 mrad opening angle. Current plans foresee a beam pipe made from aluminium, but alternatives are under study, in particular a beryllium–aluminium alloy.

2.2.6 Vertex Detector System

The vertex detector system comprises the vertex detector itself, and the pile–up veto counter. The vertex detector has to

- allow the primary and the secondary vertex to be reconstructed accurately enough for the measurement of time–dependent asymmetries, particularly in the rapidly oscillating B_s^0 system;
- provide information to the high–impact parameter level 1 trigger.

The vertex information needs to be available for the level 1 trigger decision. It is read out within 1 μs.

The pile–up veto counter

- rejects multiple interaction events and provides this information to trigger level 0.

It has to be read out within 25 ns.

The vertex detector

The current vertex detector design consists of 17 stations between $z = -18$ cm and $z = +80$ cm. Each station comprises two silicon discs, positioned perpendicularly to

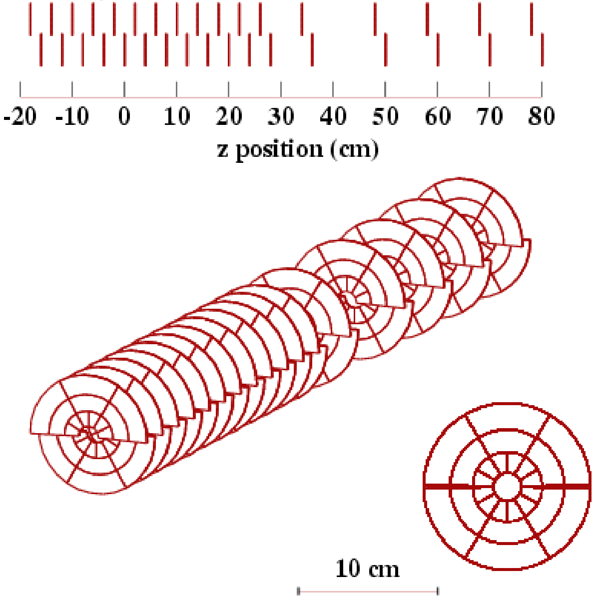


Figure 2.8: Vertex detector layout

the beam; one disc with circular strips, measuring the distance r to the z axis, one with radial strips to measure the azimuthal angle ϕ.

The readout pitch of the r–strips varies from 40 to 80 μm, the pitch for the ϕ strips from 40 to 104 μm. This pitch variation is chosen to provide a low, uniform occupancy of $\sim 0.5\%$, and the best resolution for those hits closest to the interaction point. For a measurement point in a station with a hit in both the r and the ϕ disc, the resolution is between 6 and 10 μm. The impact parameter resolution is about 40 μm for high momentum tracks.

The silicon discs cover radial distances from $r = 6$ cm down to $r = 1$ cm. With the expected radiation levels this close to the interaction point, they are expected to survive for at least one year, when operated at 5°C.

The first 12 discs cover $z = -18$ cm to $z = 26$ cm, placed at 4 cm separations. These are followed by 5 discs further downstream to cover tracks with low polar angles.

With this geometry, every track within the acceptance of the detector, down to $\theta = 15$ mrad, traverses at least three vertex detector stations.

The complete vertex system is located inside an enlarged section of the beam–pipe. The detector stations are separated from the primary vacuum of the LHC machine by a thin wall, that also acts as an RF shield. In order to keep the stresses on the wall low, so that it can be built introducing as little material as possible, the vertex detector stations are placed into a secondary vacuum. The current design foresees each detector station to be enclosed in an individual cap made from 100 μm thick aluminium. As these caps form cavity–like structures, the design requires a dedicated wake field suppressor in the vicinity of the beam.

Including guard rings and the RF shield, the vertex detector reaches down to $r = 4$ mm. In order to avoid the beam being steered into the vertex detector, the whole system will be mechanically retracted from the beam line during filling and tuning periods. To achieve this, the vertex detector discs are split into two halves that can be separated vertically. The two halves are offset by 2 cm in z to achieve an overlap of the sensitive areas.

Pile–up Veto Counter

Two dedicated r–discs act as a pile–up veto, available in time for the level 0 trigger. They are placed upstream of the main vertex detector. Each disc is subdivided into six 60° sectors, each with 300 strips pitched between 120 and 240 mm. The on–detector electronics include a discriminator, from which the binary signals are sent to the nearby vertex–finding processors.

The r coordinates of the outgoing tracks, measured in the two planes, are used to calculate the approximate z position of the primary vertex. If more than one primary vertex is found, the event is rejected. Simulations show that a primary vertex is reconstructed with a resolution of 1 mm along z; with this performance 80% of double interactions are rejected, while 90% of single interactions are retained.

Current Developments

The vertex detector design is still being optimised. The pitch of the circular strips of the r detectors is likely to be increased relative to the TP design. Also, a change in the radial coverage to $0.8\,\mathrm{cm} < r < 4.5\,\mathrm{cm}$, and an increase of the number of stations to 25 is being considered. The design of the RF shielding of the disks is under review [vB+00].

2.2.7 Dipole Magnet

In order to achieve a precision on the momentum measurements of better than half a percent for momenta up to 200 GeV, the magnet has to provide an integrated magnetic field of 4 Tm. The required aperture is ±300 mrad in x and ±250 mrad in y.

For the Technical Proposal [TP98], a window–frame magnet with superconducting bedstead coils and horizontal pole faces had been assumed. This is also the lay–out used for the simulation studies presented in chapter 3.

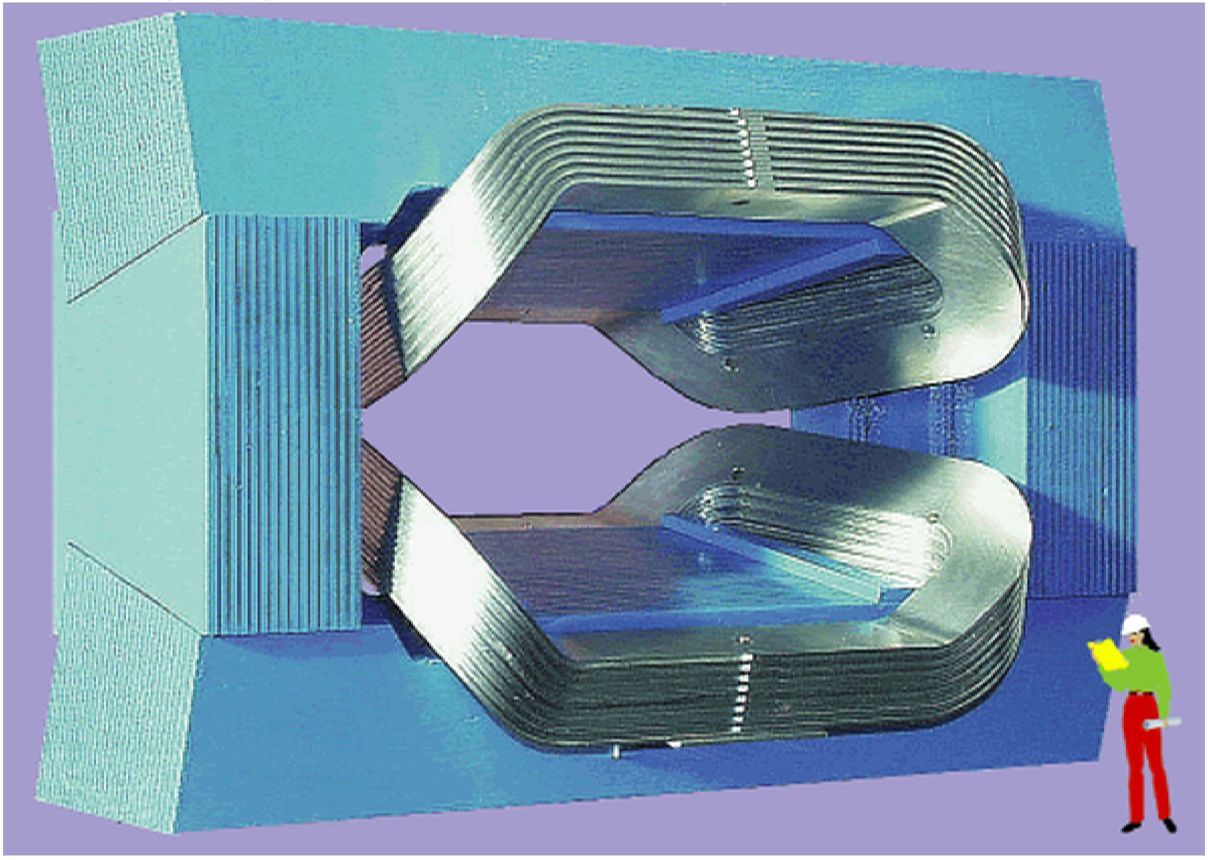
Since then, the design has been changed to a warm magnet. Apart from a faster construction and significant cost savings, the warm magnet design has several advantages, in particular that it permits regular field inversions to reduce systematic errors in $\mathcal{CP}$ violation measurements.

As seen in figure 2.4, in the new design the poles are inclined to follow the acceptance angles. This allows the 4 Tm integrated field to be retained with a power consumption of 4.2 MW. An iron shielding plate protects the vertex detector region and the RICH 1 detector from fringe fields.

A picture of a 1:25 model of the magnet is shown in figure 2.9. The new design is described in detail in [Mag00].

The effect of the new magnet design on the physics performance has been studied in detail and is not detrimental compared to the superconducting design [CFH+99, For99].

Figure 2.9: Photograph of a 1:25 model of the LHCb magnet, with a person drawn to scale (from [Mag00]).



2.2.8 Tracking System

The tracking system is divided into two subsystems, the inner and the outer tracker. The outer tracker covers most of LHCb's acceptance, where the particle flux is below $1.4 \cdot 10^5\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Straw drift chambers will be able to cope with such fluxes. Within a region $|x| < 60\,\mathrm{cm}$ and $|y| < 40\,\mathrm{cm}$, the fluxes approach $3.5 \cdot 10^8\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and a different technology is needed.

LHCb will have 11 tracking stations. The first one is directly after the vertex detector, the last one just after RICH 2. The position of the tracking stations is shown in figures 2.3 and 2.4.

Outer Tracker

The solution foreseen at the time of the Technical Proposal was to use honeycomb drift–chambers, with a resolution of better then $200\,\mu\mathrm{m}$. With a fast CF_4–based drift gas, the signal latency is two bunch–crossing intervals. Since then the design has been changed to straw–chambers with equivalent specifications.

The inner boundary of the Outer Tracker region is defined by requiring an occupancy of less than 10%.

Inner Tracker

Two different options for the Inner Tracker are being considered:

- Microstrip Gas Chambers with GEM
- Silicon Strip Detectors

Microstrip Gas Chambers with GEM: The Microstrip Gas Chamber (MSGC) with Gaseous Electron Multiplier (GEM) is the baseline solution for the inner tracker. MSGC's tend to suffer from sparking which is greatly reduced by the addition of one or more GEMs.

The GEM is a $50\,\mu$m thick kapton foil coated on both sides with copper and perforated with $60\,\mu$m holes at a $140\,\mu$m pitch. The curvature of the conducting surfaces near the holes is very high, so, in conjunction with the small distance between the metal layers, even a moderate voltage difference results in very strong electric fields.

The electrons released by an ionising particle traverse the gas drift region towards the anode, pass through the GEM, where they get multiplied through avalanche processes typically by a factor of 20. To achieve an overall gain of 4000, an MSGC with one GEM can now safely operate at a gain of 200.

The proposed detectors have a 2 mm drift gap, a $50\,\mu$m GEM foil and a 2 mm amplification gap. Glass substrates with a diamond–like coating are used, with a readout pitch of $220\,\mu$m. Amplification with two or more GEMs is under consideration.

In a 1.1 T field the strip hit multiplicity is around 1.5 and the hit resolution is better than $65\,\mu$m. The average occupancy is below 1% and the maximum occupancy 4%, at a minimum distance of 36 mm from the beam. A total sensitive area of $14\,\mathrm{m}^2$ is equipped with about $200,000$ readout channels.

Silicon Strip Detectors: Until the MSGC/GEM detectors are demonstrated to be an entirely feasible technology, Silicon Strip Detectors (SSD's) are kept as a back–up solution. SSD's are well–tested, precise, fast and radiation hard, but unfortunately more expensive.

2.2.9 Calorimeters

The calorimeter system separates hadrons, electrons and photons, and measures their energy and position.

The information provided by the calorimeter system forms the basis of the trigger level 0 decision, which looks for high E_T electrons, photons and hadrons. This task provides the main design constraints as sufficient information for the level 0 decision must be provided and processed within the 25 ns between each bunch crossing.

For the off–line analysis, the calorimeter will play a crucial role in π^0 reconstruction, required for example in the channel $\mathrm{B}_\mathrm{d}^0 \rightarrow \rho\pi$ that is sensitive to γ.

The calorimeter specifications as presented in the Technical Proposal [TP98] form the basis for the simulation study presented in chapter 3. Recently, the calorimeter group have presented their Technical Design Report [Cal00]. The most significant changes are:

- the addition of a scintillator pad detector in front of the preshower;
- a modification of the cell geometry;
- a reduction of the HCAL depth.

Here we will describe the new design, referring to the old specifications where significant changes have taken place.

General Structure

The first element of the calorimeter system is a scintillator pad detector (SPD), that signals charged particles. This is followed by the preshower (PS) detector, which is then followed by 25 radiation length of electromagnetic calorimeter and 5.6 interaction lengths of hadron calorimeter.

Geometry

The acceptance and detector segmentation of the three sub–systems are geometrically matched to facilitate the trigger formation. The inner polar acceptance starts at 30 mrad, and the outer acceptance follows the 300 mrad acceptance in x and 250 mrad acceptance in y of the LHCb spectrometer.

Figure 2.10: SPD/PS, ECAL: lateral segmentation. One quarter of the detector front is shown. The cell sizes are given for the ECAL and reduce by $\sim$ 1.5% for the SPD/PS.

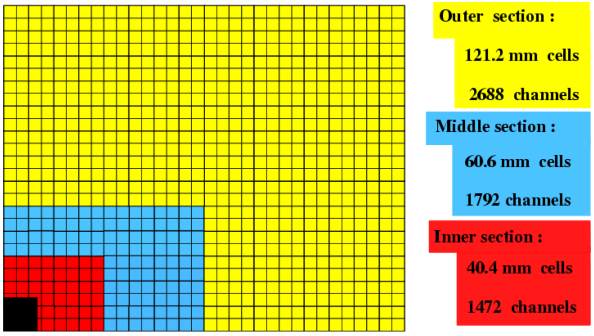
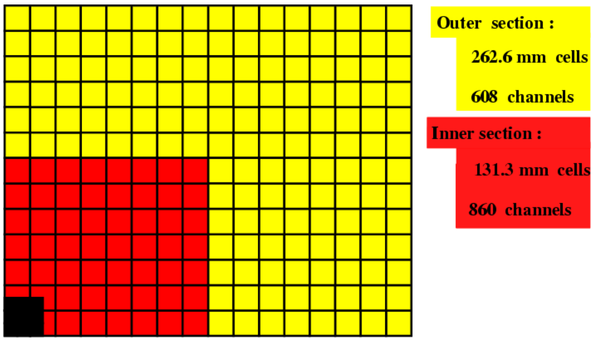
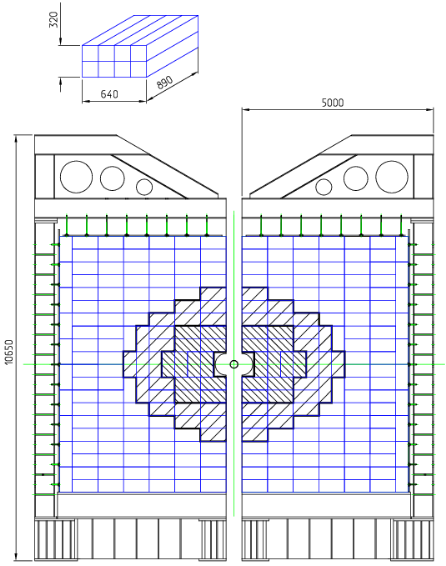


Figure 2.11: HCAL: lateral segmentation. One quarter of the detector front is shown



The lateral segmentation of the calorimeters increases closer to the beam–line, taking into account a variation in hit density of two orders of magnitude. The segmentation of the SPD/PS and the electromagnetic calorimeter is shown in figure 2.10, the

Figure 2.12: ECAL segmentation for TP



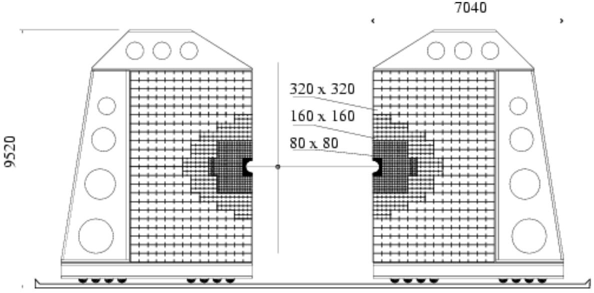
Cell sizes (mm):

inner: 40 × 40

middle: 80 × 80

outer: 160 × 160

Figure 2.13: HCAL segmentation for TP



Cell sizes • inner: 80 × 80 mm^2 • middle: 160 × 160 mm^2 • outer: 320 × 320 mm^2

segmentation of the HCAL in figure 2.11.

The calorimeter design for the TP had a different cell structure, which is shown in figures 2.12 and 2.13. Figure 2.13 also illustrates the retraction of the calorimeter halves about the beam line.

Detector Technologies

Preshower: The first element of the calorimeter system is a scintillator pad detector (SPD), that signals charged particles. This is followed by a 12 mm lead wall and another scintillator pad detector, which together form the preshower (PS) detector. The SPD has been added since the TP.

ECAL: The preshower is followed by an electromagnetic Shashlik calorimeter, which is 25 X_0 thick. Shashlik calorimeters with similar specification to the LHCb design have achieved energy resolutions of

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 1.5\% \quad (E \text{ in GeV})$$

[B$^+$96, B$^+$94]. This performance is achieved using a sampling structure of 2 mm of lead interspersed with 4 mm of scintillator plates, and light collection by wavelength shifting fibres.

HCAL: The structure chosen is an iron/scintillator tile calorimeter, read out by wavelength shifting fibres. The scintillator and iron plates are aligned in the direction of the beam. The sampling structure provides on average 4 mm scintillator thickness for every 16 mm of iron. With an overall material thickness of 1.2 m, or 5.6 λ_I, the expected energy resolution is:

$$\frac{\sigma_E}{E} = \frac{80\%}{\sqrt{E}} \oplus 10\% \quad (E \text{ in GeV}).$$

For the TP an HCAL with 7.3 λ_I was planned, with an energy resolution of

$$\frac{\sigma_E}{E} = \frac{80\%}{\sqrt{E}} \oplus 5\% \quad (E \text{ in GeV}).$$

2.2.10 Muon Detector

The muon detector provides offline muon identification, and information for the trigger level 0, which searches for high p_t muons. It consists of four stations, $M2$ to $M5$, embedded into an iron filter, placed after the calorimeters, and one unshielded station, $M1$, in front of the calorimeters.

Depending on occupancy, either Multigap Resistive Plate chambers (MRPC's) or Cathode Pad Chambers (CPC's) will be used. Most of the coverage will be provided by the cheaper MRPC's. But as their performance begins to decrease significantly at rates above 5 kHz/cm^2, CPC's will be used for station $M1$ and the coverage near the beam line, where the rates are higher.

The shielding in front of the last muon station is listed in table 2.1. In order to reach the last of the five muon stations, a muon must have an energy of at least 6 GeV. Additional shielding immediately behind $M5$ protects it from backsplash from nearby LHC beam elements and from particles emerging from the collider tunnel.

Table 2.1: Muon detector components and shielding in terms of radiation and interaction lengths, listed in the order in which they are encountered by a particle travelling downstream. Numbers are given for the old (TP) and the new calorimeter design.

Detector Element	Composition	Depth	
		z/X_0	z/λ_I
M1	CPC		
ECAL+PS+SPS (new)	Pb/Scint.	27	1.3
ECAL,PS (TP)	*Pb/Scint.*	*28*	*1.3*
HCAL (new)	Fe/Scint.	55	5.6
HCAL (TP)	*Fe/Scint.*	*68*	*7.3*
Shield-1	Fe	17	1.8
M2-5	MRPC/CP		
Shield-2,3,4 (each)	Fe	39	4.2
Total Thickness (new Calo)		216	21
Total Thickness (TP Calo)		*230*	*23*

Recent Developments

Since the Technical Proposal, the muon chamber design has been further optimised. For most parts of the muon detector Multi-Wire Proportional Chambers (MWPC) will be used. The lowest-occupancy region in stations $M3$ $M4$ and $M5$ will be equipped with Resistive Plate Chambers (RPC). The choice of technology for the highest–occupancy regions near the beam pipe in stations $M1$ and $M2$ has still to be made [Sch00].

2.2.11 The Trigger

As B hadrons have a long lifetime and are rather heavy, events with B hadrons can be distinguished from other inelastic pp interactions by the presence of a secondary vertex and particles with high transverse momentum. Trigger levels 0, 1 and 2 select events by these two criteria. The final trigger level, 3, is then used to select the particular B decay topologies of interest.

Level 0

Trigger level 0 triggers on high–$\mathrm{p_t}$ particles and rejects multiple interactions. It has to operate at the LHC bunch–crossing frequency of 40 MHz. Level 0 comprises three high–$\mathrm{p_t}$ triggers, the muon, the hadron and the electron trigger, which use information from the calorimeters and the muon system only. The fourth component of the level 0 trigger is the pile–up veto provided by the vertex detector. Trigger level 0 is designed to achieve a suppression factor of 40.

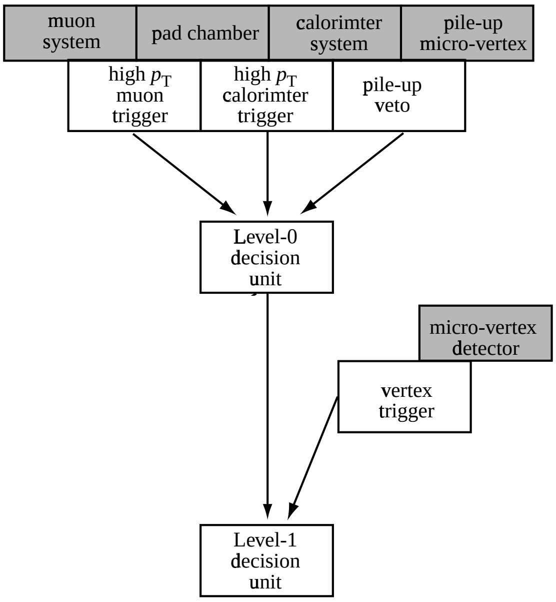
The latency of the level 0 trigger is 128 clock cycles, or 3.2 μs. The requirements of high–speed and radiation hardness demand the use of specialised, purpose-built hardware.

Level 0 Decision Unit

The trigger accepts an event if it has one or more high $\mathrm{p_t}$ candidates above the set thresholds and is flagged as a single interaction by the pile–up veto. The thresholds will be defined to optimise the event yields in the channels of interest. The threshold cuts in the Technical Proposal are:

- $\mathrm{p_t}(\mu) > 1\,\mathrm{GeV}$
- $E_T(e) > 2.34\,\mathrm{GeV}$
- $E_T(\gamma) > 4\,\mathrm{GeV}$
- $E_T(\mathrm{hadron}) > 2.4\,\mathrm{GeV}$

Figure 2.14: Schematic of the LHCb Trigger level 0 and level 1 architecture.



- di–lepton events are accepted with reduced thresholds.

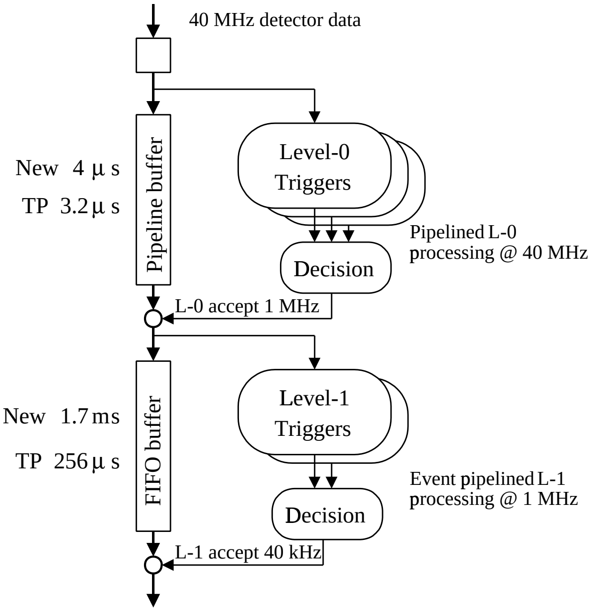
The particle spectra are such that this choice assigns the majority of the available bandwidth to the hadron trigger.

Level 1

Trigger level 1 triggers on detached vertices. Level 1 is designed to suppress the input rate of 1 MHz by a factor of 25. A schematic of the trigger level 0 and level 1 layout, showing the detector components involved at each level, is given in figure 2.14.

The implementation of the level 1 algorithms includes a mixture between specialised hardware and farms of high–performance commercially available proces-

Figure 2.15: Trigger logic of level 0 and level 1



sors. The latency of the level 1 algorithms is variable, but limited to a maximum of 256 μs. To ensure that the level 1 front–end buffer does not overflow, it must be able to store up to 250 events accepted by level 0.

A schematic of the trigger logic is shown in figure 2.15.

Level 0 and 1: Recent Developments

Since the Technical Proposal, the level 0 latency has been increased to 4 μs, and the maximum level 1 latency to 1700 μs.

Level 2

Trigger level 2 uses momentum information from the tracker to eliminate fake secondary vertices which are typically caused by multiple–scattering of low–momentum tracks.

Level 2 is designed to reduce its input rate of 40 kHz by a factor of 8, to 5 kHz.

Level 0, 1 and 2 Performance

The efficiencies of trigger levels 0, 1 and 2 are given in table 2.2 for different decays (from [TP98]).

Level 3

Trigger level 3 performs full and partial event reconstruction to select specific decay channels of interest. It makes use of information from almost all detector components, including possibly the RICH, as illustrated in the flow diagram in figure 2.16. In order to select the decays of interest, it is expected that level 3 will implement a loose sub–set of the final selection cuts. The efficiency of the level 3 trigger is therefore expected to be close to 100% with respect to this selection.

The suppression factor of level 3 is ~ 25, to give a data recording rate of ~ 200 Hz.

2.3 The RICH system

2.3.1 Motivation

LHCb intends to perform high precision measurements on many different B^0 decay

Table 2.2: Trigger efficiencies of Level 0, 1 and 2 for events where the final state is reconstructed offline and its flavour is correctly tagged, for non-specific flavour final states. Level 0, 1 and 2 triggers are applied consecutively.

Decay Mode	Physics	Level-0	Level-1	Level-2
$B^0_d \to \pi^+\pi^-$ + right tag	γ	0.76	0.48	0.83
$B^0_d \to J/\psi(e^+e^-)K_S$ + right tag	β	0.72	0.42	0.81
$B^0_d \to J/\psi(\mu^+\mu^-)K_S$ + right tag	β	0.88	0.50	0.81
$B^0_s \to D^+_s K^-$ + right tag	$\gamma - 2\delta\gamma$	0.54	0.56	0.92
$B^0_d \to \overline{D}^0 K^{*0}$	γ	0.37	0.59	0.95

Figure 2.16: Data flow diagram for Trigger level 3.

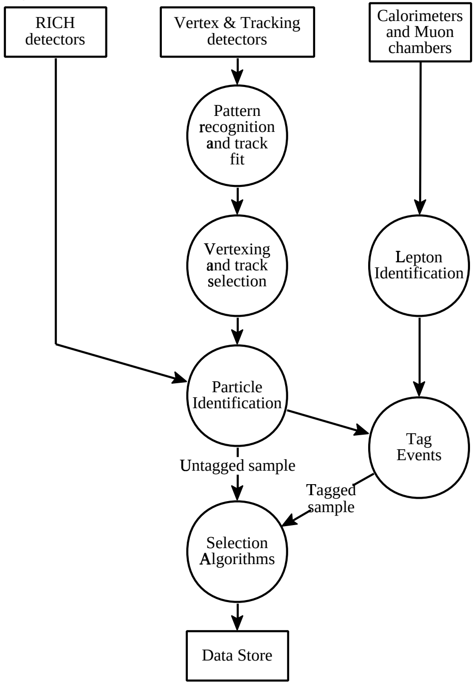


Table 2.3: Some channels that are sensitive to γ, and depend on K/π separation. The ticks indicate the importance of the RICH information for each analysis: $\surd$=significant improvement; $\surd\surd$=essential.

Channel	K/π-sep.
$B^0_d \to D^*\pi$	$\surd$
$B^0_s \to D_s K$	$\surd\surd$
$B^0_d \to \pi\pi$	$\surd\surd$
$B^0_s \to KK$	$\surd\surd$
$B^0_d \to D_d D_d$	$\surd\surd$
$B^0_s \to D_s D_s$	$\surd\surd$

channels. Many interesting decay channels are themselves backgrounds to topologically similar ones. Typically the branching ratios are of the order of $\sim 10^{-5}$. Particle identification and in particular K/π separation is provided by the RICH. This is essential for obtaining the clean samples needed to perform a comprehensive range of high–precision $\mathcal{CP}$ violation measurements.

Figures 2.17, 2.18 and 2.19 show simulated invariant mass distributions for the channels $B^0_d \to \pi^+\pi^-$, $B^0_s \to K^+K^-$ and $B^0_s \to D_sK$ (from [ABF+01]). Also shown are background contributions from various same–topology decays. In all cases, the signal, shown in red, is overwhelmed by the background. The background channels can themselves show $\mathcal{CP}$ violation effects, which could cause uncontrollable systematic uncertainties at these low purities.

Also crucial for the measurement of decay–rate asymmetries is B^0–tagging, which is explained in section 2.2.2. The K/π separation provided by the RICH allows the use of kaons for B^0 tagging. This is equivalent to a $3\frac{1}{2}$ fold increase in statistics compared to the performance without kaon tagging.

Figure 2.20 shows the momentum distribution of (a) pions in $B^0_d \to \pi^+\pi^-$ events, and (b) tagging kaons. This illustrates the need for K/π separation over a wide range of momenta; LHCb seeks K/π separation from

Figure 2.17: $B^0_d \to \pi^+\pi^-$ and same-topology background, without RICH (simulation)

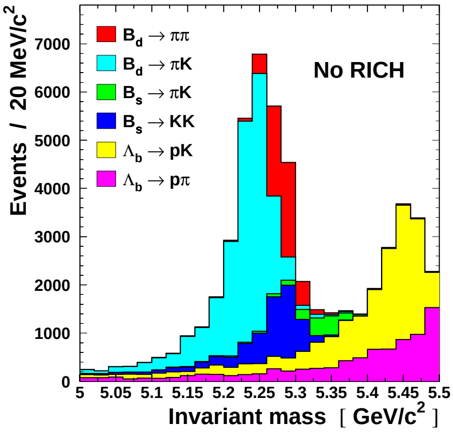


Figure 2.18: $B^0_s \to K^+K^-$ and same-topology background, without RICH (simulation)

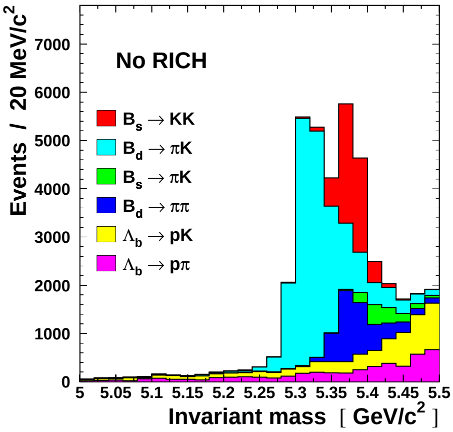
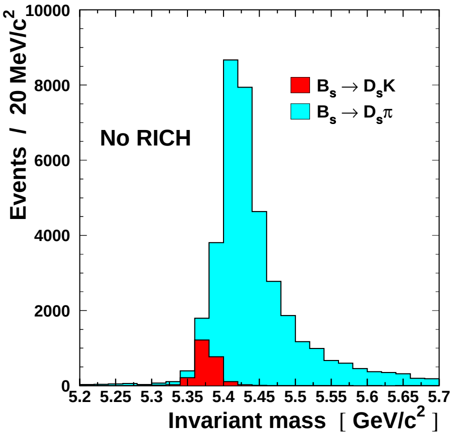


Figure 2.19: $B^0_s \to D_sK$ and same-topology background, without RICH (simulation)



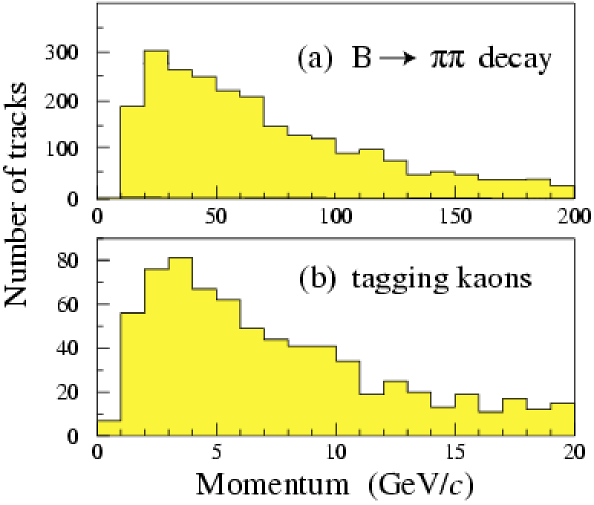


Figure 2.20: Momentum distributions for (a) the highest momentum pion from $B_d^0 \to \pi^+\pi^-$, shown for momenta between 0 and 200 GeV, and (b) tagging kaons between 0 and 20 GeV.

momenta of ~ 1 GeV to beyond 100 GeV.

2.3.2 Cherenkov light

Cherenkov angle, θ_C

When a particle passes through a dielectric and has a speed larger than the speed of light in that substance, a cone of Cherenkov light is emitted. The opening angle θ_C of the cone is related to the speed β of the particle by:

$$\cos\theta_C = \frac{1}{n\beta} = \frac{1}{n}\sqrt{1 + \left(\frac{m}{p}\right)^2} \qquad (2.2)$$

for units with $c = 1$; n is the refractive index of the material, which in general varies with the wavelength of the emitted radiation.

The maximum Cherenkov angle for a given refractive index is:

$$\cos\theta_{\max} = \frac{1}{n}. \qquad (2.3)$$

For $n - 1 \ll 1$:

$$\theta_{\max} \approx \sqrt{2\,(n-1)}. \qquad (2.4)$$

Number of Photons

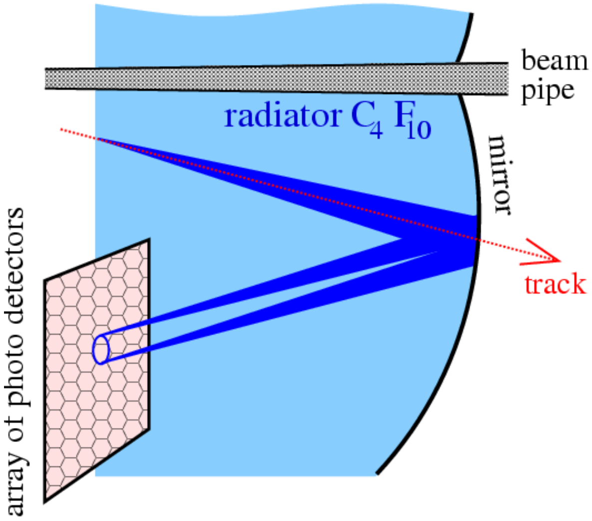
The distribution of photons must, for symmetry reasons, be flat in the azimuthal angle ϕ. The number of photons, N_γ, emitted per unit length of radiator, l, per unit photon–energy $E = \hbar\omega$, is given by [FT37, Gin40]:

$$\begin{aligned} \frac{dN_\gamma}{dl\,dE} &= \frac{\alpha}{\hbar c} Z^2 \left(1 - \frac{1}{(\beta\, n(E))^2}\right) \\ &= \frac{\alpha}{\hbar c} Z^2 \sin^2\theta_C(E) \qquad (2.5) \end{aligned}$$

where Z is the charge of the particle in units of e. Entering the values of the constants in units of cm and eV, and integrating over the length of the particle track, L, we get for a particle with $Z = 1$:

$$\frac{dN_\gamma}{dE} = \frac{370}{\text{eV cm}} \sin^2\theta_C(E) L. \qquad (2.6)$$

Figure 2.21: Principle of operation of a RICH detector



2.3.3 RICH Counters

RICH counters measure the opening angle θ_C of the Cherenkov cone by imaging it onto an array of photodetectors. The theory of RICH detectors is developed in [YS93]. In the LHCb RICH detectors, the Cherenkov cone is focused onto the detector surface with spherical mirrors. The mirrors are slightly tilted to allow the photo detectors to be positioned outside the detector acceptance. The basic principle is illustrated in figure 2.21.

Given the momentum p, the mass m can be calculated, and the particle type identified.

Number of Detected Photons

The detector parameters that determine the number of photons detected can be grouped into "detector response" parameters, and "detector geometry" parameters.

The detector geometry parameters are:

- ϵ_A, the geometric efficiency of the detector. It is the probability that a photon reaching the photodetector plane hits the active area of a photodetector.
- L, the length of the radiator.

The detector response parameters are:

- $T(E)$, the probability that a photon reaches the photo–detector plane. This comprises the transmission of the radiator gas, the mirror reflectivity, and other elements of the detector optics, for example optical filters;
- $Q(E)$, the quantum efficiency of the photodetector. It is the probability that a photon is converted into an electron, as a function of the photon energy;
- ϵ_D, the efficiency with which the photo electron is detected.

The detector response parameters can be summarised in a single figure, N_0, "the figure of merit" for a RICH detector:

$$\mathrm{N}_0 \equiv \frac{370}{\mathrm{eV\ cm}} \int \epsilon_D Q(E) T(E) \mathrm{d}E. \tag{2.7}$$

It is also useful in this context, to define the differential detector response as a function of the photon energy

$$\mathrm{n}_0(E) \equiv \frac{370}{\mathrm{eV\ cm}} \epsilon_D Q(E) T(E). \tag{2.8}$$

Using these definitions together with equation 2.6, the number of photons detected with a Cherenkov detector, $\mathrm{N_{pe}}$, is given by:

$$N_{\mathrm{p.e.}} = L\epsilon_A \int \mathrm{n}_0(E)\sin^2\theta_C(E)\mathrm{d}E. \tag{2.9}$$

For the approximation that $\sin^2\theta_C(E)$ is constant over the relevant energy range:

$$N_{\mathrm{p.e.}} \approx L\epsilon_A \mathrm{N}_0 \sin^2\theta_C. \tag{2.10}$$

Momentum Range

The lower end of the momentum range over which the RICH can operate is defined by the threshold for Cherenkov light emission:

$$\begin{aligned} \frac{1}{n}\sqrt{1+\left(\frac{m}{p}\right)^2} &\leq 1 \\ \Rightarrow \quad p &\geq \frac{m}{\sqrt{n^2-1}}. \end{aligned} \tag{2.11}$$

From equation 2.2, we get for small angles θ_C and for large momenta, $p \gg m$:

$$1 - \frac{1}{2}\theta_C^2 \approx \frac{1}{n}\left(1 + \frac{1}{2}\frac{m^2}{p^2}\right). \tag{2.12}$$

The approximation of small angles is valid for the LHCb RICH system, where the maximum Cherenkov angle is between 32 mrad and 240 mrad (radiator dependent).

The difference in θ_C^2 for two tracks with the same momenta, but different masses, is therefore:

$$\theta_2^2 - \theta_1^2 = -\frac{1}{n}\frac{\Delta(m^2)}{p^2} \tag{2.13}$$

$$\Leftrightarrow \Delta\theta \cdot (\theta_2 + \theta_1) = -\frac{1}{n}\frac{\Delta(m^2)}{p^2} \tag{2.14}$$

where $\Delta\theta = \theta_2 - \theta_1$ and $\Delta(m^2) = m_2^2 - m_1^2$. For large momenta, we can approximate

$$\theta_1 + \theta_2 \approx 2\theta_{\max} \approx 2\sqrt{1 - \frac{1}{n^2}}. \tag{2.15}$$

Hence the difference $\Delta\theta_c$ between the Cherenkov angles of two particles with the same momentum p and different masses is, for $p \gg \Delta(\mathrm{m}^2)$ and $\theta_C \approx \theta_{\max} \ll 1$, given by:

$$\Delta\theta_c \approx -\frac{1}{2\sqrt{n^2-1}} \cdot \frac{\Delta(\mathrm{m}^2)}{p^2}. \tag{2.16}$$

If the Cherenkov angles are Gaussian distributed with a width σ_θ, the number of sigmas of separation between two tracks, n_σ, is given by

$$n_\sigma = \frac{1}{2}\frac{1}{\sigma_\theta\sqrt{n^2-1}} \cdot \frac{\Delta(\mathrm{m}^2)}{p^2}. \tag{2.17}$$

For the limits on the momentum range, we get therefore:

$$\begin{aligned} p_{\mathsf{min}} &\approx \frac{m}{\sqrt{n^2-1}} \\ p_{\mathsf{max}} &\approx \sqrt{\frac{1}{2}\frac{1}{\sigma_\theta\sqrt{n^2-1}} \cdot \frac{\Delta(\mathrm{m}^2)}{n_\sigma}} \end{aligned} \tag{2.18}$$

where, for K/π separation, m is the pion mass, and $\Delta(\mathrm{m}^2)$ the difference between the squares of the kaon and the pion mass. The minimum number of sigmas separation between the two particle hypotheses, n_σ, is typically required to be 3.

Resolution

The width of the Cherenkov angle distribution is determined by

- **Emission point error:** As the spherical mirrors are slightly tilted, the point where a photon hits the detector surface depends on the position along the radiator where it was emitted, which is not known.
- **Chromatic dispersion:** The refractive index, and therefore the Cherenkov angle, depends on the wavelength of the emitted photon, which is not known.
- **Tracking:** To calculate θ_C from the hit position on the detector surface, the track direction needs to be known.
- **Pixelisation:** The photodetector granularity limits the position resolution on the detector plane, and therefore the Cherenkov angle resolution.

A more detailed discussion of the above contributions can be found in chapter 4.

The Cherenkov angle resolution for the complete ring also depends on the

- **Number of photo electrons per track:** The more photo electrons are detected, the more accurate is the Cherenkov angle reconstruction. Assuming the errors on the Cherenkov angle reconstruction for each photon are independent, we get for the Cherenkov angle resolution of the whole ring: $\sigma_\theta \propto \frac{1}{\sqrt{N_{\text{p.e.}}}}$.

The above contributions need to be balanced carefully against each other, and cost. For example, increasing the wave–length range over which the photodetectors are sensitive increases the number of photo electrons per track, but worsens the chromatic dispersion. Improving the photodetector resolution increases the cost in proportion to the number of readout channels. However, the overall performance is not significantly improved once a resolution has been achieved that is on a similar level to the other error contributions.

2.3.4 The LHCb RICH

Two RICH Detectors

As shown in equation 2.18, the upper and the lower limit of the momentum range covered by a RICH detector depend on the refractive index of the radiator, and the Cherenkov angle resolution. In order to provide K/π separation from $\sim 1\,\text{GeV}$ to $> 100\,\text{GeV}$, the LHCb RICH system employs two RICH detectors with three radiators.

The first radiator is silica aerogel with a refractive index of $n = 1.03$. This is suitable for the lowest momentum tracks starting at about 1 GeV. Aerogel is a fine mixture between a solid and a gas. It has a structure that is much smaller than the wavelength of visible light, so its optical properties are a mixture of those of the individual materials and can be fine–tuned.

The useful wavelength range of aerogel is limited by Rayleigh scattering of photons with small wavelength. The transmission of light of wavelength λ through a length L is proportional to

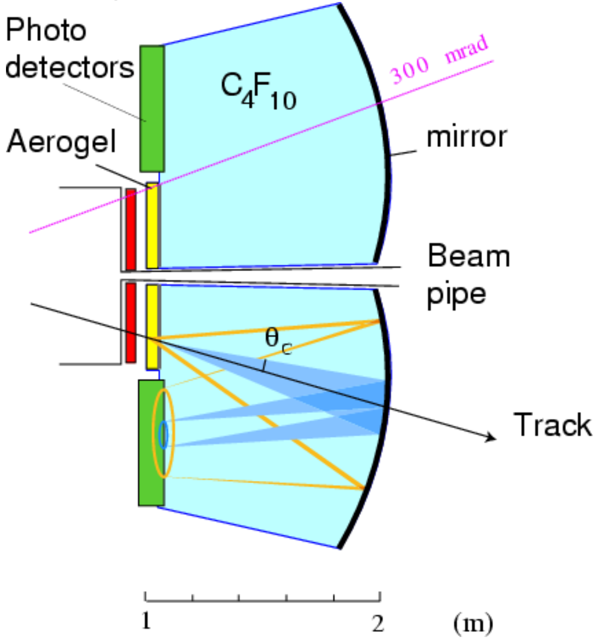
$$e^{-CL/\lambda^4},$$

where C is the clarity coefficient. In the simulation studies for the assessment of the RICH performance, $C = 0.008\,\mu\mathrm{m}^4/\mathrm{cm}$ is assumed. Aerogels with $C = 0.004\,\mu\mathrm{m}^4/\mathrm{cm}$ are however under study, which would result in reduced photon losses due to scattering.

A Mylar filter immediately downstream of the aerogel is used to absorb wavelength below ~ 350 nm in order to reduce the background due to scattered photons. The best possible efficiency for those unscattered photons transmitted by the mylar, extending into the visible spectrum, is an important consideration in the photodetector choice.

The intermediate momentum region is covered by gaseous C_4F_{10} with $n = 1.0014$, providing K/π separation up to ~ 70 GeV. Both radiators are part of the **RICH 1** detector that is positioned close to the interaction region, immediately downstream of the vertex detector. This is to minimise the required photodetector area and to provide a wide acceptance for low-momentum particles. A schematic of RICH 1 is shown in figure 2.22. RICH 1 covers tracks with polar angles between 25 and 300 mrad.

Figure 2.22: RICH 1 Schematic



The particle first traverses 5 cm of aerogel–radiator, and then 83 cm of C_4F_{10} gas. From simulation studies we expect to detect about 7 photons from the aerogel and 33 photons from the C_4F_{10} gas.

Figure 2.23 shows the Cherenkov angle against the momenta of pions and kaons for the two radiators. Figure 2.24 shows the two photo detector planes with (simulated) Cherenkov rings identified by the pattern recognition algorithm. The rings from the aerogel and the gas are clearly visible and distinguishable.

In order to cover momenta up to > 100 GeV and tracks at polar angles below 25 mrad, a second detector, **RICH 2**, is placed further downstream. Its polar angle acceptance starts at 15 mrad. Given the strong correlation between the momentum and the polar angle of the tracks, as demonstrated

Figure 2.23: Cherenkov angle vs momentum for pions and kaons in RICH 1

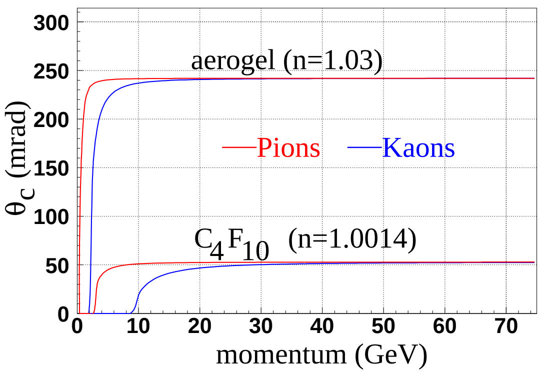


Figure 2.24: Rings in RICH 1

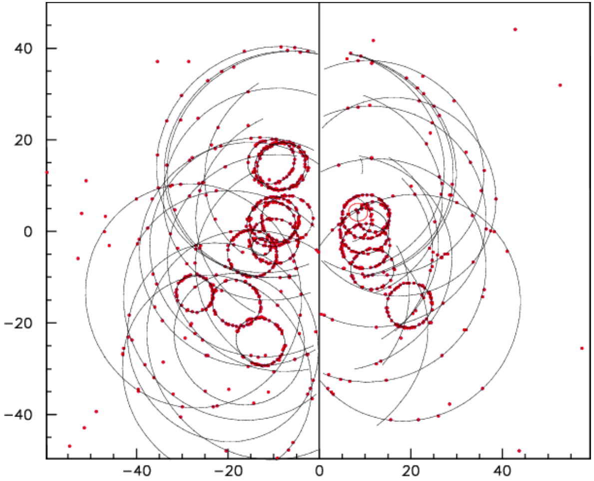
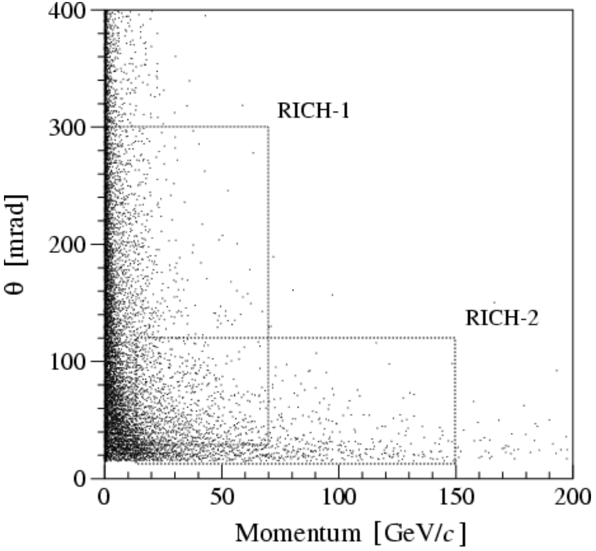


Figure 2.25: Polar angle versus momentum for all tracks in simulated $B_d^0 \rightarrow \pi^+\pi^-$ events. The regions covered by RICH 1 and RICH 2 are indicated by dashed lines.

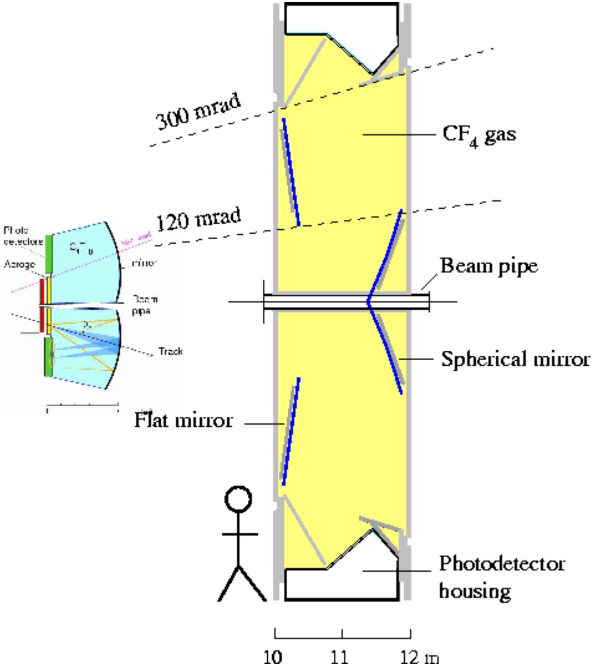


in figure 2.25, it is sufficient for RICH 2 to cover polar angles up 120 mrad.

Equation 2.18 shows that, in order to increase the momentum coverage towards higher momenta, one can either improve the Cherenkov angle resolution, or reduce the refractive index. In RICH 2, both options are pursued. The radiator is CF_4 gas with $n = 1.0005$. To improve the resolution compared with RICH 1, a larger detector is constructed. The focal lengths of the mirror and mirror–detector distance are increased, which gives a larger image and thus reduces the pixelisation error. The small dispersion of CF_4 results in a reduced chromatic error. The smaller Cherenkov angle, and the reduced curvature of the mirror, lead to a reduced emission point error.

A schematic of RICH 2 is shown in figure

Figure 2.26: RICH 2 schematic, with RICH 1 drawn to scale for comparison .



2.26. In order to keep the length of the detector small, the light is reflected from the spherical mirrors first onto two flat mirrors, and from there to the detector plane.

2.3.5 Pattern Recognition

To associate rings in the two RICH detectors with tracks found in the tracking system and to make particle assignments, a global likelihood fit is performed, using all available information from the three radiators simultaneously.

A likelihood function is calculated assuming an approximately Gaussian distribution in the Cherenkov angle distribution for each track, plus background. For each track found in the tracking system, the number and position of photons on the detector surface for a given particle type assumption is calculated and compared with the number found. The combination of particle assignments producing the largest likelihood is the result of the fit.

The log–likelihood function that is maximised is:

$$\ln \mathcal{L} = -\sum_{\text{track } j} \mu_j + \sum_{\text{pixel } i} n_i \ln \left(\sum_{\text{track } j} a_{ij} + b_i \right) , \tag{2.19}$$

where a_{ij} is the expected number of detected photoelectrons from track j in pixel i (under a given set of track particle-type hypotheses), and $\mu_j = \sum_i a_{ij}$ is the expectation for the total number of detected photoelectrons from track j; n_i is the number of photoelectrons falling into pixel i; b_i is the expected background falling in pixel i from sources without a reconstructed track.

2.3.6 Photodetectors

The photodetector for the LHCb RICH must be

- efficient, with sensitivity extending into the visible spectrum;
- fast (1 bunch–crossing every 25 ns);
- precise (pixel size: $\sim 2.5 \times 2.5\,\text{mm}^2$);
- affordable to cover $2.6\,\text{m}^2$;

- radiation hard to exposures up to 3 kRad/year;
- tolerant to magnetic fringe fields of $20-30$ Gauss.

The baseline photodetector choice is the Hybrid Photo Diode, HPD. Multi–anode Photo Multiplier Tubes (MaPMTs) are kept as a back–up solution until all technical challenges of the HPDs have been solved.

HPDs

An HPD consists of a vacuum tube with a light–sensitive photo cathode at the entrance window and a silicon detector at the tube base. Between the photocathode and the silicon detector a high voltage of ~ 20 kV is applied. If a photon hitting the cathode is converted to an electron, this electron is accelerated by the applied electric field onto the Si sensor, where it creates a large number of electron–hole (e–h) pairs. The energy required to create an e–h pair in Si is 3.6 eV, hence the number of e–h pairs created by a 20 kV electron is about $20000/3.6 \approx 5.6 \cdot 10^3$. In practice this number will be slightly smaller, as some energy will be lost in the dead layers of the sensor.

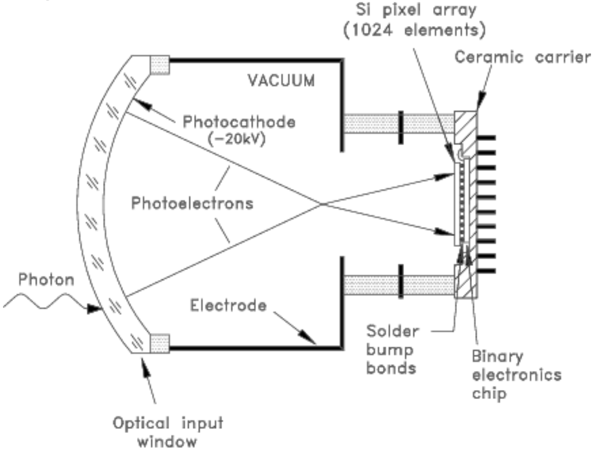
Apart from the high photon detection efficiencies achievable with HPDs, they also allow for a very good signal to noise ratio, as the gain is achieved in a single dissipative process with a very large mean number of e–h pairs. The standard deviation of this distribution is $\sqrt{F \cdot N}$ where F is the Fano Factor [Fan47] and N the number of e–h pairs produced. For electrons incident on a Si target, $F \approx 0.11$ [ABS80]. The relative width of the distribution of the number of created e–h pairs per single photo electron is therefore about $\sqrt{\frac{0.11}{5 \cdot 10^3}} = 0.5\%$. Hence the widths of the peaks in the pulse height spectra shown in figure 2.30 are almost entirely due to electronics noise.

Backscattering

Not all electrons deposit all of their energy in the Si - about 18% are backscattered [Dar75], losing some of their energy before scattering back from the Si to either re–enter at another point, or leave the sensitive region altogether.

In order to estimate the potential background from backscattered electrons which re–enter the Si, we will derive an expression for the probability of a photo electron entering the Si detector in an HPD more than once. For the case of an homogeneous electric field, it is easy to show that the maximum radius over which a backscattered electron can re–enter is $2h$, where h is the distance between the cathode and the Si detector. As shown in [A^+90b], the probability per unit area for a backscattered electron to re–enter at any point within this radius is approximately constant. Therefore, for a Si sensor of diameter d and a tube of height h, the probability that any electron enters the Si twice is $\sim 0.18 \cdot (d/(4h))^2$.

Figure 2.27: Schematic of the Pixel HPD



2.3.7 The Pixel HPD

The HPD for the LHCb RICH detectors is being developed in association with the company DEP[1]. A diagram of the device is shown in figure 2.27. The electrons released from the 75 mm diameter cathode are cross–focused by two electrodes onto the anode, demagnifying the image by a factor of ~ 5.

The Cathode

The cathode is a multi–alkali S20–type deposited on the inside of the 7 mm thick quartz entrance window. The quantum efficiency of the cathode as a function of the photon energy, taken from measurements of HPD prototypes, is shown in figure 2.28. The measurement includes light–losses in the quartz window. The sensitive range of the cathode extends approximately from 2 to 7 eV.

Figure 2.28: Quantum efficiency of the photocathode (including the effect of the quartz window) as a function of photon energy.

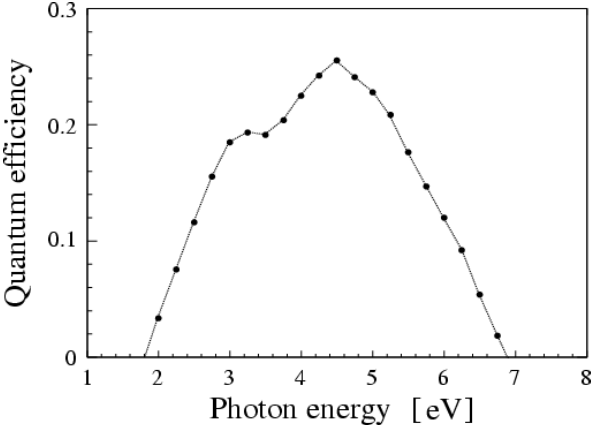
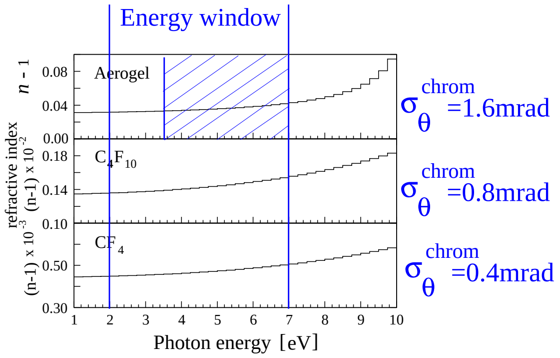


Figure 2.29: Dispersion of radiators.
The plots show $(n-1)$ as a function of the photon energy, indicating the limits set by the sensitivity of the photocathode. The excluded area in the plot for the aerogel is due to the Mylar filter. On the right hand side, the chromatic error for the single–photon resolution is given.



---

[1] Delft Electronische Producten (DEP), The Netherlands

Figure 2.29 illustrates the effect this range has on the chromatic error. The integrated quantum efficiency of the cathode is:

$$\int Q(E)\mathrm{d}E = 0.77\,\mathrm{eV}. \tag{2.20}$$

Electron Optics

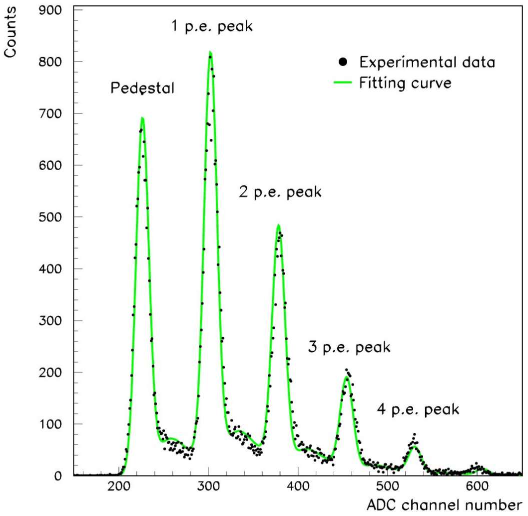
The electron optics cross–focus the photo electrons released from the ø75 mm cathode onto a ø15 mm area on the Si sensor.

Prototype tests have shown that the demagnification is almost perfectly linear across the cathode surface. The spatial resolution of the electron optics is described by the width of the Point Spread Function (PSF). The PSF describes the radial distribution of electrons on the Si surface, generated by a point source on the photo cathode. The width of the PSF is required to be much smaller than the pixel size on the Si sensor, which is $0.5 \times 0.5\,\mathrm{mm}^2$. In prototype tests, the RMS of the PSF was measured to be between 33 and 54 μm, clearly fulfilling this requirement.

The Anode

The anode of the Pixel HPD is a square silicon sensor divided into 32 × 32 pixels of $500 \times 500\,\mu\mathrm{m}^2$ size, bump–bonded to a pixel readout chip. The entire anode–assembly is encapsulated within the tube. With a demagnification factor of 5, the effective pixel size on the tube surface is $2.5 \times 2.5\,\mathrm{mm}^2$.

Figure 2.30: Single–channel pulse height spectrum from full–scale HPD prototype with a 61 pixel Si sensor and external, analogue readout electronics (from [Gys])

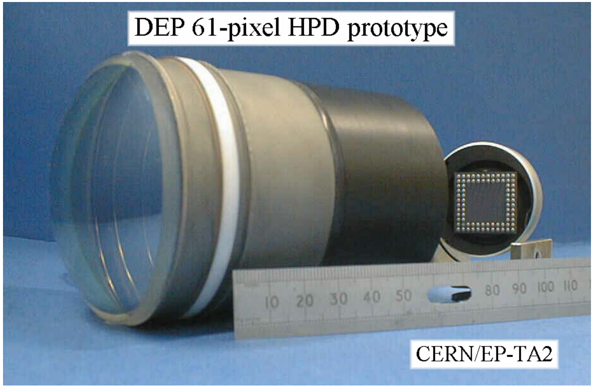


Each logical $500 \times 500\,\mu\mathrm{m}^2$ pixel is segmented into 10 physical $50 \times 500\,\mu\mathrm{m}^2$ diode pixels. Each physical pixel is bump–bonded onto a readout cell with matching dimensions on the front–end pixel readout chip.

Detection Efficiency

Figure 2.30 shows the pulse height spectrum of a full–scale prototype of the tube, however equipped with a 61 pixel silicon sensor and external analogue readout electronics (figure 2.31). The effect of backscattered electrons depositing only a fraction of their energy in the Si is clearly visible in

Figure 2.31: Photograph of full–scale prototype. A mirror behind the device shows the pin–grid array for the readout of the 61 pixel Si sensor (from [Gys]).



the areas between the peaks. Because of the small size of the Si sensor relative to the height of the tube, nearly all backscattered electrons land outside the Si sensor, so that background due to backscattering can be neglected. Including the signal loss due to backscattering, the detection efficiency of photo electrons in the final tube design is expected to be $\sim 90\,\%$.

The Pixel Chip

The pixel chip must be able to read out the Si sensor within the 25 ns between bunch crossings. It provides the digitisation and buffering of the data awaiting the level 0 trigger decision.

After pre–amplification, shaping and discrimination, the signals from the 10 physical pixels are combined in an OR gate to form a single binary signal. The result is stored in one of twenty delay units, awaiting the level 0 decision. After a level 0 accept, the data of each logical pixel are passed to one of four 4–event FIFO buffer, forming together a 16 deep de–randomiser.

The data from each chip are read out in < 900 ns through 32 parallel lines at 40 MHz. With this, the dead time losses in the DAQ are maintained below 1%.

The architecture of a pixel cell is shown in figure 2.32.

The chip is fabricated in $0.25\,\mu$m CMOS technology, using a layout adapted for radiation tolerance. The total power consumption of the chip is ~ 0.5 W.

Occupancy and Binary Readout

Binary readout has the benefits of simplifying the electronics, reducing power consumption, and in particular reducing cost. The physics performance of the RICH will not be significantly affected by the reduced information provided by binary readout, because the occupancy levels per pixel are expected to be low and multiple photo electron events are therefore rare. Simulation studies show that the typical occupancy per pixel is about 1%, and never exceeds 8%. The occupancy per pixel, averaged over each tube, is shown in figure 2.33.

Shielding

Both RICH detectors will be shielded against stray fields from the LHCb dipole magnet. Nevertheless, it is expected that the tubes will have to operate in a resid-

Figure 2.32: Schematic of the Pixel Cell architecture. One of the ten physical pixels making up each logical pixel is shown. The signals from the physical pixels are OR'ed into one binary signal. Two of twenty delay units, and one of four FIFO buffers available for the storage of the binary signal from the logical pixel are shown.

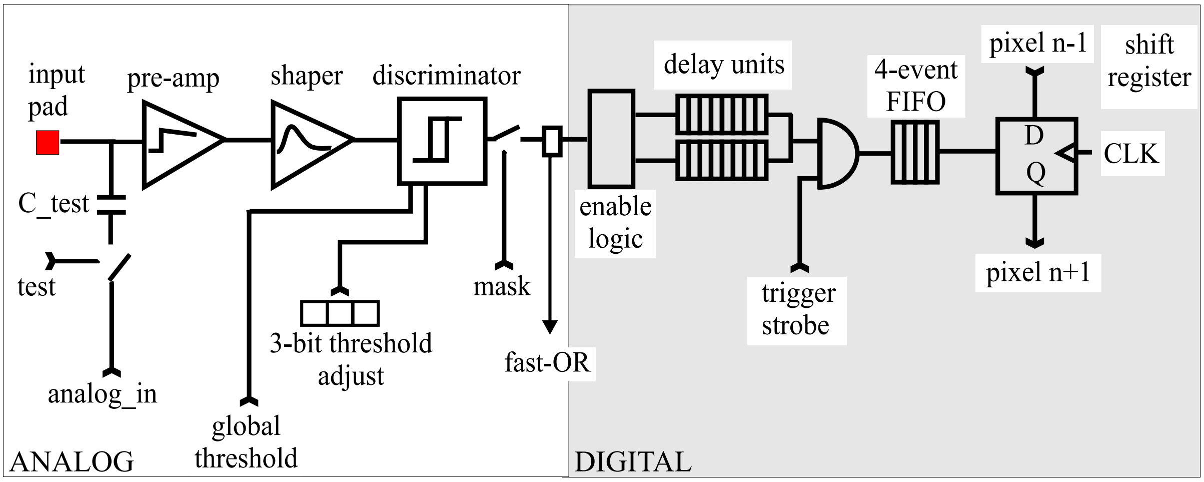
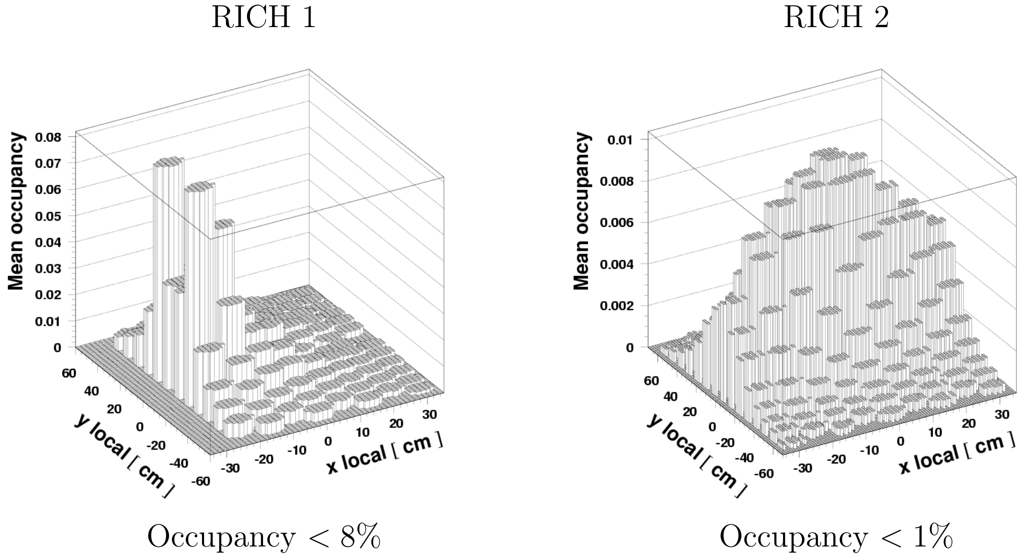


Figure 2.33: Simulated occupancies per pixel in RICH 1 and RICH 2, averaged over one tube. In each case, half the detector plane is shown; the local co–ordinate system is chosen such that beam–pipe is at $x \approx -30\,\mathrm{cm}$ and $y = 0$.



ual field of up to 20 or 30 Gauss. To shield the tubes from these magnetic fields, each will be surrounded by a 0.9 mm thick μ–metal shield. A schematic of the tube with the μ–metal shield is shown in figure 2.35. A prototype of the tube, with full electron optics and fitted μ–metal shield, has been tested in a field of 30 Gauss. The prototype differed from the final design in that it had a phosphor cathode, and a CCD camera instead of the pixel sensor for the photo electron detection. Figure 2.34 shows how a cross on the cathode is imaged on the anode with and without a magnetic field of 30 Gauss, for a vertical field, and for a field along the tube axis. The cross remains on the Si sensor, and the point–spread function is not significantly affected. Hence the magnetic field effects can be corrected off–line.

2.3.8 Readout

The first stage of the RICH readout is performed by the pixel chip encapsulated in each tube. Signals from two tubes are read out by one "Level-0 Adapter Board", mounted on–detector. The adapter board:

- further multiplexes the data,
- drives optical data links to the off–detector (level 1) electronics,
- distributes clock and trigger signals to the front–end chips.

The pixel chip and the adaptor board form together the level 0 electronics. The signals from the adaptor board are passed via optical links to the level 1 readout board, located about ~ 100 m away from the detector in the counting room. This allows the use of non–radiation hard electronics for level 1. The level 1 readout board:

- buffers the data during the level 1 latency,
- performs zero–suppression and de–randomisation,
- removes events that fail trigger level 1,
- provides the interface to the rest of the LHCb detector electronics, the "Timing and Control" (TTC) system and the "Detector Control System" (DCS).
- passes the data on to the DAQ and event–building network.

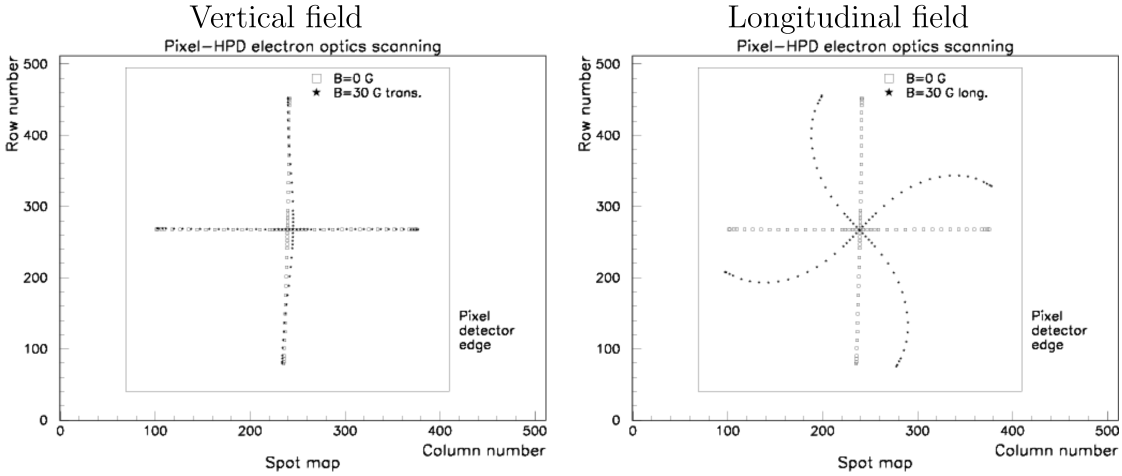
A schematic of the RICH readout architecture is shown in figure 2.36.

Tube Mounting

The tubes will be stacked hexagonally in a metal support structure, with a distance of 87 mm between the tube centres. This is illustrated in figure 2.37. With a packing fraction for hexagonal close packing of 90.7% and an active diameter of 75 mm, this gives an active–area fraction of $\epsilon_A = 67\%$.

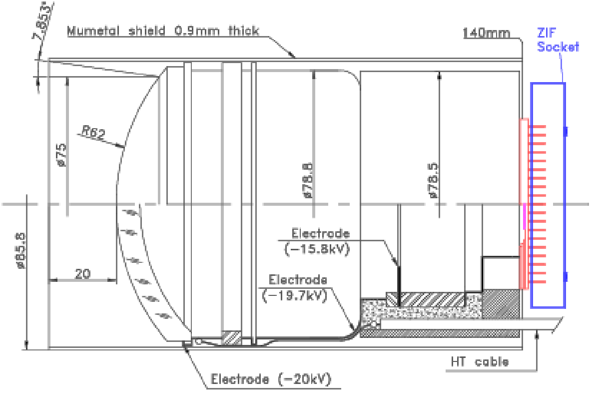
To increase detection efficiency, the photodetectors are slightly tilted relative to the detector plane, so that the cathode windows

Figure 2.34: Prototype test of shielded tube in a magnetic field.
The figures show how a cross on the cathode is imaged on the anode with and without a magnetic field of 30 Gauss. Left for a vertical field, and right for a field along the tube axis. The area of the Si sensor is marked with a box.


Vertical field

Longitudinal field

Figure 2.35: Tube with magnetic shielding



are perpendicular to the incident light. This reduces shadowing effects from the μ–metal shield, and reflection losses. The tilted tube–mounting is illustrated in figure 2.38 for the example of RICH 1.

2.3.9 Mirrors

The mirrors need to cover a total area of $\sim 20\,\mathrm{m}^2$. In RICH 1, the $2.7\,\mathrm{m}^2$ large spherical mirror will be divided into 16 rectangles. The $9.2\,\mathrm{m}^2$ spherical mirror for RICH 2 will be divided into 56 hexagonal mirror segments, and the $7.6\,\mathrm{m}^2$ flat mirrors by 40 squares.

The mirrors will be made from 6 mm glass, coated with 900 nm aluminium, which will be protected by a 200 nm quartz layer. The expected reflectivity is 90%, averaged over the relevant wavelength range.

The mirrors will be aligned first by an in–situ survey of all positions. The precision of this first alignment is expected to be better than 0.5 mrad in RICH 1 and 0.1 mrad in RICH 2. Once LHCb operates, a laser system will be employed to monitor changes in time. To achieve the final precision of bet-

Figure 2.36: A schematic of the LHCb RICH readout architecture

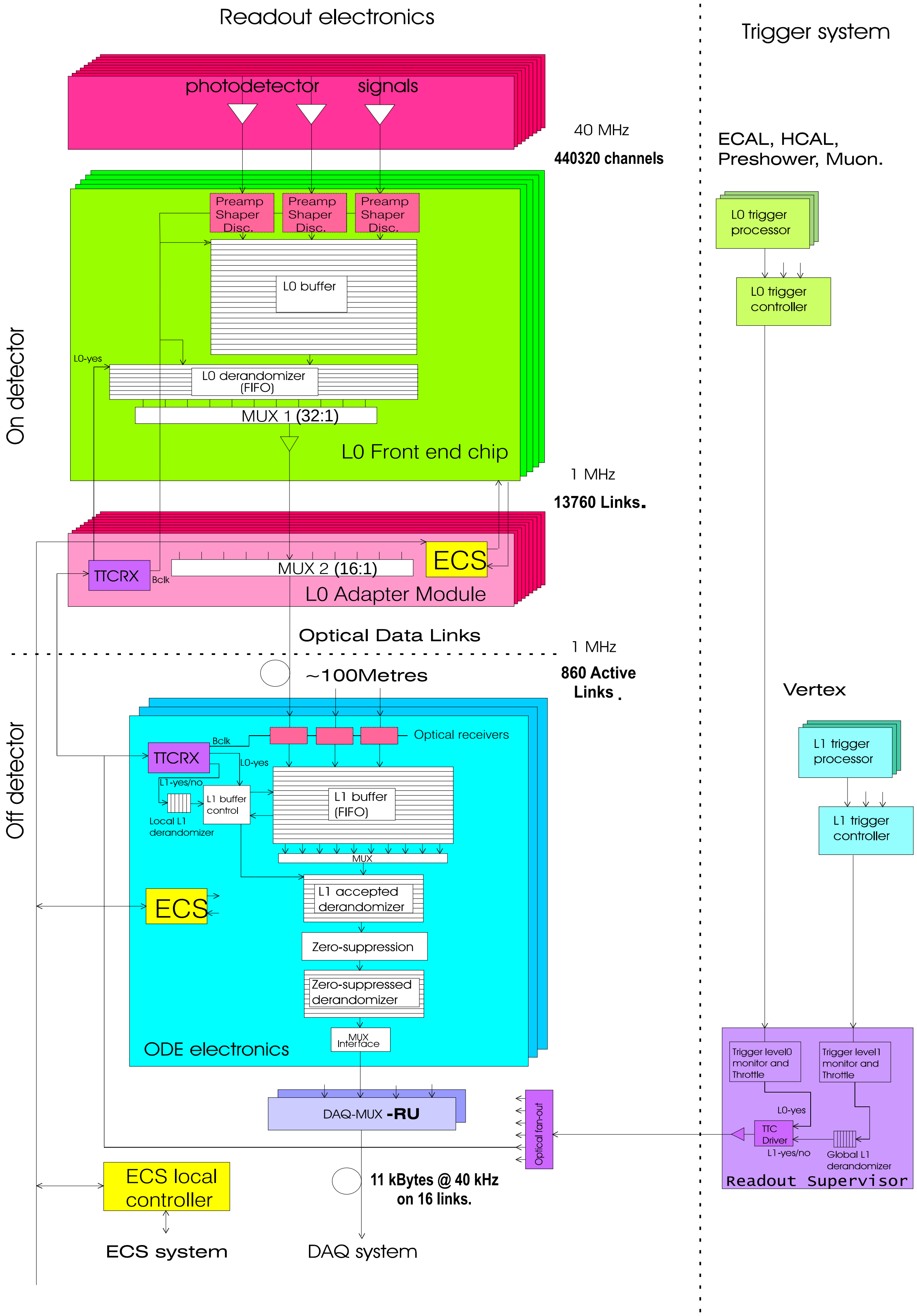


Figure 2.37: Tube Mounting (RICH 1). The tubes, shown with their μ–metal shields, are hexagonally stacked in a metal support structure.

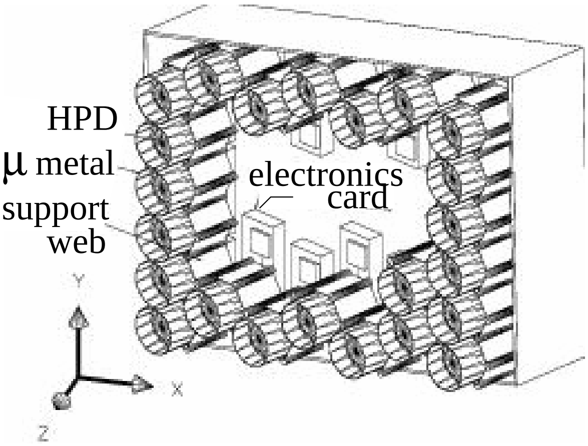


Figure 2.38: Stacked Tubes in RICH 1. Inclining the tubes towards the beam line increases the photon detection efficiency.

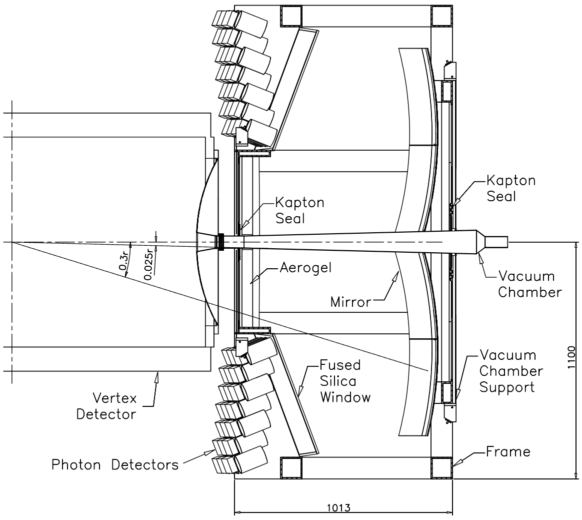


Table 2.4: Contributions (expressed in fractions of a radiation length) to the material in RICH 1 and RICH 2, which fall within the LHCb acceptance.

Item	RICH 1	RICH 2
Entrance window	0.001	0.014
Aerogel	0.033	
Gas radiator	0.024	0.017
Mirror	0.046	0.046
Mirror support	0.030	0.033
Exit window	0.006	0.014
Total (X_0)	0.140	0.124

ter than 0.1 mrad, $\beta = 1$ tracks in the data will be used.

2.3.10 Material Budget

The total material placed in the detector acceptance by the RICH system amounts to $\sim 26\%$ of a radiation length. The contributions from the different RICH components to the material budget are given in table 2.4.

2.3.11 LHCb RICH Performance

Table 2.5 summarises the characteristics of the three radiators of the RICH, and gives the values for the different contributions to the single–photon Cherenkov angle resolution, and the average number of detected photo electrons per ring. The values for σ_θ and $N_{\text{p.e.}}$ are the result of a detailed simulation study [RIC00].

Table 2.5: Characteristics of the 3 radiators

	RICH 2	RICH 1	
	CF_4	C_4F_{10}	Aerogel
Characteristics			
L [cm]	167	85	5
n	1.0005	1.0014	1.03
$\theta_c^{\max}$ [mrad]	32	53	242
$p_{\text{thresh}}(\pi)$ [GeV]	4.4	2.6	0.6
$p_{\text{thresh}}(\text{K})$ [GeV]	15.6	9.3	2.0
θ_c resolution for single photo electrons			
$\sigma_\theta^{\text{emission}}$ [mrad]	0.31	0.74	0.60
$\sigma_\theta^{\text{chromatic}}$ [mrad]	0.42	0.81	1.61
$\sigma_\theta^{\text{pixel}}$ [mrad]	0.18	0.83	0.78
$\sigma_\theta^{\text{track}}$ [mrad]	0.20	0.42	0.26
$\sigma_\theta^{\text{total}}$ [mrad]	0.58	1.45	2.00
Number of photo electrons per ring			
N_{pe}	18.4	32.7	6.6

Assuming for the Cherenkov angle resolution that $\sigma_\theta = \frac{\sigma_{\text{single p.e.}}}{\sqrt{N_{\text{p.e.}}}}$, we get a Cherenkov angle resolution for RICH 2 of 0.135 mrad. From equation 2.18, page 49, we obtain a 3σ K/π separation from ~ 1 GeV up to 93 GeV. A 2σ separation can be achieved up to 114 GeV, and 1σ up to 162 GeV. This assumes that the Cherenkov angle distribution is well described by a simple Gaussian. As shown in chapter 4, this is not always a good approximation. It also assumes that the errors on the Cherenkov angle for each photon are uncorrelated, which is not the case for example for the tracking error. However it might be possible, to reduce the contribution from the tracking error considerably, by refitting Cherenkov rings already identified in the pattern recognition, independent of the tracking information. Such fit methods are currently being investigated.

Monte Carlo studies with a detailed model of the RICH and a full implementation of the pattern recognition algorithm have been performed. The results of the pattern recognition algorithm for the light particles e, μ, K and p are shown in table 2.6. The performance is given in terms of the purity $\mathcal{P}$ of the sample, and the efficiency ε. $\mathcal{P}$ is defined as the fraction of tracks for which a given RICH particle assignment is correct; ε is defined as the fraction of tracks of a given type, that receive the correct particle assignment by the RICH. The table is given for well–reconstructed tracks with $1\,\text{GeV} < p < 150\,\text{GeV}$. The K/$\pi$ separation provided by the RICH is impressive: 95% of pions are identified as pions or lighter particles, and 92% of kaons as kaons or heavier particles.

2.3.12 RICH Physics

The RICH in Event Selection

Figures 2.39, 2.40, and 2.41 illustrate the dramatic improvements in the signal to background ratio that can be achieved by using the RICH, for three channels that measure the angle γ. In all cases the efficiency of the signal selection is 80-90% (from [ABF+01]).

The RICH in B^0 Tagging

In order to decide the flavour of the reconstructed B^0 at its creation, LHCb uses

Table 2.6: Results from the pattern recognition applied to well–reconstructed tracks in triggered and accepted signal events between 1 and 150 GeV/c. Each track gives one entry in the table, and X denotes tracks below threshold in all radiators; the rows give the reconstructed particle type, $\mathcal{P}$ is the purity and ε the efficiency. The sample corresponds to 500k tracks, but has been normalised to 1k. From [RIC00]

	True particle type						
Rec	e	μ	π	K	p	X	$\mathcal{P}$
e	97.4	0.7	24.6	1.4	0.5	3.1	0.76
μ	4.0	8.7	69.5	2.0	0.5	4.9	0.10
π	2.5	1.3	545.7	3.3	0.7	5.1	0.98
K	0.3	0.1	12.7	70.6	4.8	4.3	0.76
p	0.2	0.0	1.7	4.3	35.9	0.0	0.85
X	9.9	0.8	19.8	3.2	0.0	55.6	0.62
ε	0.85	0.76	0.81	0.83	0.85	0.76	

Figure 2.39: $B_d^0 \rightarrow \pi^+\pi^-$ and same–topology background, with and without RICH (simulation)

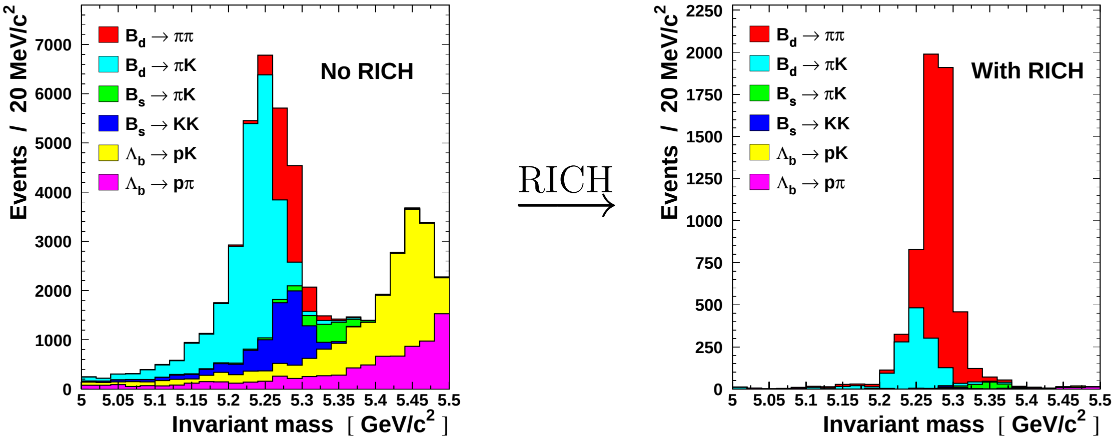


Figure 2.40: $B_s^0 \rightarrow K^+K^-$ and same–topology background, with and without RICH (simulation)

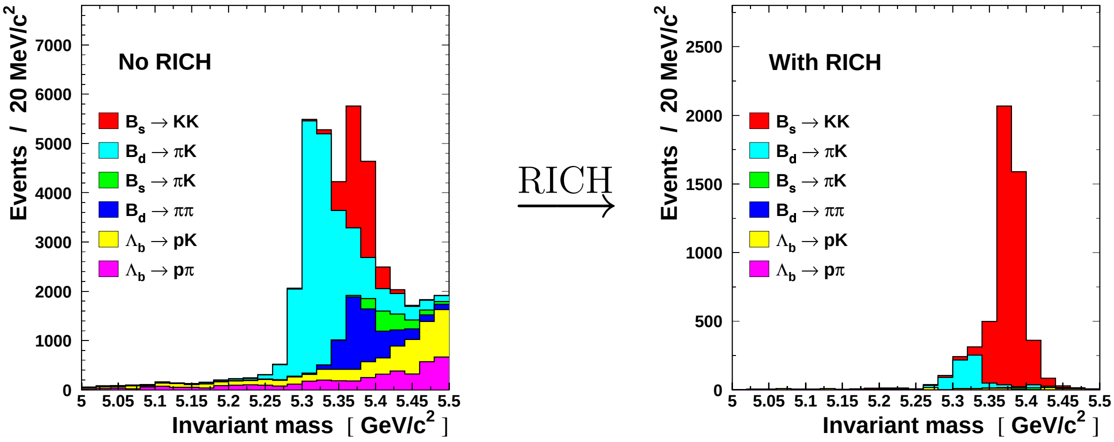


Figure 2.41: $B_s^0 \rightarrow D_sK$ and same–topology background, with and without RICH (simulation)

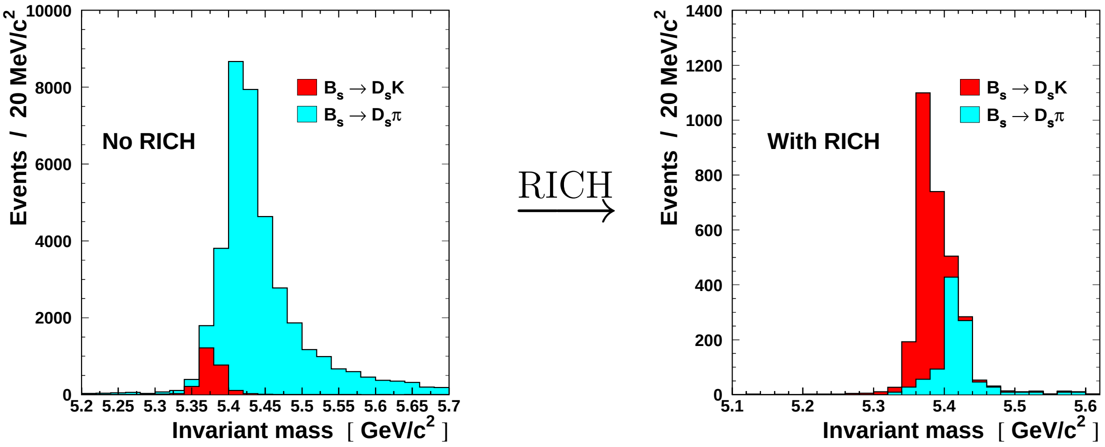
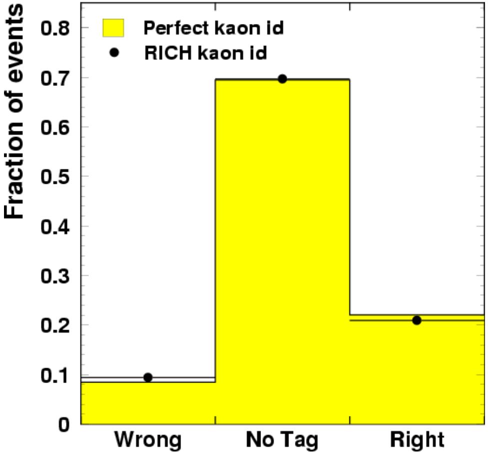


Figure 2.42: Compare Performance with RICH/ with perfect particle identification: with RICH: $P_{\text{tag}} = 4.5\%$ with perfect ID: $P_{\text{tag}} = 6.6\%$



electrons, muons and kaons from the decay of the accompanying B decay (see section 2.2.2).

The kaon tag is only possible with the K/π separation provided by the RICH. Figure 2.42 compares the performance of the kaon tag using the RICH, with the performance if the particle identification were perfect.

The performance parameters of the kaon tag are [RIC00]:

- Efficiency $\epsilon = 31.2\%$
- Mistag rate $\omega = 31.0\%$
- statistically equivalent to $P_{\text{tag}} = \epsilon\,(1-2\omega)^2 = 4.5\%$ perfectly tagged events.

The total tagging power at LHCb is $P_{\text{tag}} \approx 6.4\%$, therefore it can be seen that the kaon contribution dominates.

2.4 Detector Simulation

The results on the $\mathrm{B}^0_\mathrm{d} \to \mathrm{D}^*\pi$ reconstruction efficiencies and purities presented in chapter 3, as well as the physics results taken from [RIC00], [B$^+$00] and [TP98] to illustrate the performance of the RICH, have been obtained using the LHCb detector simulation and reconstruction program SICb [Tsa97]. The relevant versions of SICb for the study presented in chapter 3 are version 112 to 116; the specifications given here refer to those versions. The Monte–Carlo event generator used in SICb is PYTHIA 5.7/JETSET 7.4. The detector response is simulated using the GEANT 3.21 package.

2.4.1 Generating Events

Default settings were used for the physics parameters in PYTHIA. $b\bar{b}$ pairs in hadron collisions are produced by the mechanisms of gluon fusion, gluon splitting and flavour excitation. The contribution of each mechanism to the total cross section is of the same order of magnitude, but the resulting kinematics differ. There are two ways to generate $b\bar{b}$ events in PYTHIA:

1. Using a steering card MSEL=5, a gluon fusion mechanism ($gg \to b\bar{b}$) is simulated. Each event contains at least one $b\bar{b}$ pair.
2. Using a steering card MSEL=1, all QCD 2 $\to$ 2 processes are simulated, and all production mechanisms contribute to the $b\bar{b}$ production. But now only 1% of the events contain a $b\bar{b}$ pair.

The events used in the study presented in chapter 3 were generated using the more efficient method with MSEL=5. Studies performed since indicate that the kinematics obtained with MSEL=1 result in a slightly increased trigger efficiency [CKW98].

In the simulation studies presented in 3, all mesons are treated as scalar particles, including vector mesons such as the D*.

2.4.2 Simulating the Detector

Environment

LHC: SICb contains a simplified model of the LHC beam elements, the LHCb cavern and infrastructure.

Magnet: A detailed field map of the magnetic field of the old, superconducting magnet is provided, including inhomogeneities and stray fields.

Beam Pipe: A simplified model for the beam pipe is used in the simulation. Studies with a more realistic model of an aluminium vacuum pipe performed since indicate that the number of secondary particles generated in the simplified version is rather optimistic [TvH00]. This is one of the motivations for the study of alternative beam pipe designs.

Sub Detectors

Vertex Detector: The simulation of the Vertex Detector includes details such as the energy deposition in the Si wavers and charge sharing between the strips. The support structure and the R.F. shield are included in the simulation.

Tracker: For the honeycomb chambers in the Outer Tracker, the drift time is simulated from the position of the particle trajectory in the drift cell. Hits in the Inner Tracker are generated with a single–hit RMS of 60 μs.

Calorimeters: The energy deposition in the ECAL is simulated using GEANT. Electrons with less then 1 MeV and hadrons with less than 10 MeV are ignored. Above these thresholds, the shower development in the calorimeter materials can be simulated accurately, including energy leakage at the boundaries. The threshold energy for the HCAL is 500 MeV for all particles. Studies using the lower ECAL energy thresholds show that the higher threshold values are a reasonable simplification.

Muon Detector: The iron absorbers and the muon stations are simulated, taking into account low–energy particles produced in secondary interactions when generating the pad–hits.

RICH: For the RICH performance studies performed for the RICH–TDR, which form the basis of the results presented in section 2.3.11, a detailed simulation of Cherenkov radiation and digitisation is used. This is described in detail in [RIC00]. In the SICb versions used for the study presented in chapter 3, only parametrised photo–detector efficiencies and resolutions were available.

Event Reconstruction

Vertexing: Full pattern recognition is implemented for tracks in the vertex detector. The primary vertex is determined using all tracks found in the vertex detector. This includes tracks outside the spectrometer acceptance, without momentum measurement. For such tracks, the strong correlation between polar angle and momentum (figure 2.25) is used to estimate the momentum from the angle. The resolution of the reconstructed primary vertex is $\sigma_{x,y} = 9\,\mu$m and $\sigma_z = 38\,\mu$m for inelastic proton–proton collisions.

Trigger 0 and 1: Full pattern recognition for the level 0 high $\mathrm{p_t}$ trigger and the level 1 vertex trigger is implemented.

Tracking: For the tracking system, pattern recognition is not fully implemented. Instead, Monte Carlo–truth information is used to identify hits in the tracker with tracks. However, preliminary studies indicate a track–seeding efficiency close to 100% [FN00]. The tracks are fitted using a Kalman–filter algorithm, where continuous energy losses and multiple scattering are taken into account.

Particle Identification: Parametrised angular resolutions of measured Cherenkov angles in the three radiators are used to calculate the probabilities for the e, μ, π, K and p hypotheses. For the results from the TDR, presented in section 2.3.11, full pattern recognition was implemented.

Tagging and Trigger Levels 2 and 3: The tagging algorithm was not available in SICb at the time this study was performed; also, trigger levels 2 and 3 were not available. All events passing trigger level 1 are treated as triggered and tagged. Event yields and the physics performance parameters are calculated from the simulation results assuming a tagging efficiency of 40%, a mistag fraction of 30%, and a combined efficiency of trigger levels 2 and 3 of $\sim 92\,\%$ for the decay $\mathrm{B_d^0} \to \mathrm{D}^*\pi$.

2.5 Conclusion

The LHCb detector is designed specifically to make best use of the huge number of $b\bar{b}$ pairs produced at the future LHC collider. Approximately $5.6 \cdot 10^{11}$ $b\bar{b}$ from single interaction events passing the pile–up veto are expected per year.

Crucial for the LHCb physics program is the particle identification provided by the RICH system. The two RICH detectors will provide K/π separation over a large momentum range, starting at momenta of

$\sim 1\,\text{GeV}$. The upper limit is $\sim 100\,\text{GeV}$ for 3σ separation. Useful K/π separation extends beyond 100 GeV.

The LHCb RICH forms the basis of chapters 4 and 5. For the simulation studies presented in chapter 3, a detailed model of the LHCb detector, described in section 2.4, was used.

Chapter 3

The Angle γ from $\mathrm{B}^0_\mathrm{d} \rightarrow \mathrm{D}^*\pi$

3.1 Introduction

The channel $\mathrm{B}^0_\mathrm{d} \rightarrow \mathrm{D}^*\pi$[1] allows a theoretically clean determination of the angle γ, that is expected to be robust against New Physics. It therefore provides a Standard Model benchmark that other γ measurements which are more susceptible to New Physics, such as $\mathrm{B}^0_\mathrm{d} \rightarrow \mathrm{D_d D_d}$ and $\mathrm{B}^0_\mathrm{s} \rightarrow \mathrm{D_s D_s}$, can be compared against. Any deviation between these γ values will signal non–Standard Model higher order processes.

In this chapter a simulation study is presented, investigating the feasibility of a γ measurement at LHCb using $\mathrm{B}^0_\mathrm{d} \rightarrow \mathrm{D}^*\pi$.

The first part of this chapter investigates the event yields and signal to background ratios achievable at LHCb, using LHCb's GEANT–based detector simulation program SICb. Two methods of reconstructing $\mathrm{B}^0_\mathrm{d} \rightarrow \mathrm{D}^*\pi$ events are presented:

- The exclusive method, where the following decay chain

  $$\begin{tikzcd} \mathrm{B}^0_\mathrm{d} \arrow[r] & \mathrm{D}^{*-}\pi^+ \arrow[dr] & & \\ & & \overline{\mathrm{D}^0}\pi^- \arrow[dr] & \\ & & & \mathrm{K}^+\pi^- \end{tikzcd}$$

  is fully reconstructed.

- The inclusive method, where all decays of the $\overline{\mathrm{D}^0}$ are reconstructed simultaneously, as long as they involve at least two charged particles, $\mathrm{X}^+, \mathrm{Y}^-$:

  $$\begin{tikzcd} \mathrm{B}^0_\mathrm{d} \arrow[r] & \mathrm{D}^{*-}\pi^+ \arrow[dr] & & \\ & & \overline{\mathrm{D}^0}\pi^- \arrow[dr] & \\ & & & \mathrm{X}^+\mathrm{Y}^- \ldots \end{tikzcd}$$

  Only the two pions from the B^0_d and the D^* decay vertex are used for the reconstruction of the B^0_d 4–momentum and decay time.

In the second part of this chapter the event yield and signal–to–background results from the first part are used to calculate the expected precision in γ.

The influence that the event yield, the signal to background ratio, and the detector

---

[1] Unless the context requires otherwise, this notation implies all four decay rates: $\mathrm{B}^0_\mathrm{d} \rightarrow \mathrm{D}^{*-}\pi^+$, $\overline{\mathrm{B}^0_\mathrm{d}} \rightarrow \mathrm{D}^{*-}\pi^+$, $\overline{\mathrm{B}^0_\mathrm{d}} \rightarrow \mathrm{D}^{*+}\pi^-$, $\mathrm{B}^0_\mathrm{d} \rightarrow \mathrm{D}^{*+}\pi^-$

parameters have on the precision of the measurement is derived in section 3.5.

3.2 Event Reconstruction

3.2.1 Basic Reconstruction Principles

B Decay

The distinguishing properties of B hadrons are their large mass ($\sim 5\,\mathrm{GeV}$) and their long mean life ($\tau \sim 1\,\mathrm{ps}$). For example in the decay $\mathrm{B}^0_\mathrm{d} \to \mathrm{D}^*\pi$, the 5 GeV B^0_d meson decays to the 3 GeV lighter D^*, and a π that will carry most of these 3 GeV as kinetic energy.

The B^0_d lifetime of $\sim 1.5\,\mathrm{ps}$ corresponds to flight distances of $\sim 1\,\mathrm{cm}$ at LHCb.

To identify B decays, we therefore look for events with detached vertices and high transverse momentum particles. This approach is central to the LHCb trigger algorithm, as explained in chapter 2.

D^* Decay

The D^* decays almost instantaneously, so that the B^0_d-vertex and the D^*-vertex are identical; this is what makes the inclusive reconstruction described in section 3.4 possible.

The mass difference between the D^* and the D^0 is 144.5 MeV, leaving only 5 MeV after subtracting the pion mass, so that the pion will essentially be at rest in the D^* centre-of-mass frame. In the following it is referred to as the "slow pion", π_slow, as opposed to the "fast pion", π_fast, from the $\mathrm{B}^0_\mathrm{d} \to \mathrm{D}^*\pi$ decay. The difference between the reconstructed D^*-mass and $\overline{\mathrm{D}^0}$-mass provides a very powerful cut for the D^*-reconstruction.

Decay Time

The $\mathcal{CP}$ violation parameters are extracted from time-dependent asymmetries. The decay eigentime is given by

$$\tau = \frac{mc^2}{|\vec{p}|\, c \cdot c} l \qquad (3.1)$$

where l is the decay length, obtained from the primary and secondary vertex positions, m is the B^0_d-mass and $\vec{p}$ the 3–momentum.

Extracting the $\mathcal{CP}$ asymmetry with good precision requires that the reconstruction error on the decay time is small compared to the B^0_d-$\overline{\mathrm{B}^0_\mathrm{d}}$ oscillation period of 13.3 ps.

3.3 Exclusive Reconstruction of $\mathrm{B}^0_\mathrm{d} \to \mathrm{D}^*\pi$

3.3.1 Datasets

The LHCb simulation program SICb, described in section 2.4, was used to generate two sets of Monte Carlo data:

Signal Events: 7 000 signal events are generated, where each event contains at least one $B^0_d \to D^*\pi$ decay, with the subsequent decays $D^{*-} \to \overline{D^0}\pi^-$ and $\overline{D^0} \to K^+\pi^-$.

The events are generated such that the momentum of the b quark that forms the B^0_d is within a forward cone of 600 mrad, so that most decays are within the LHCb acceptance. Simulation studies where the b quarks are generated over 4π show that 40% of the events survive a corresponding cut on the b–quark polar angle.

The tagging algorithm was not available for this study. Therefore, all events are treated as if they were tagged. The finite tagging efficiency of 40% is however taken into account when calculating the event yields. The mistag probability of 30% is taken into account when the γ reach is calculated later in this chapter.

With a branching fraction for this decay chain of $6.8 \cdot 10^{-5}$, an efficiency of the cut on the polar angle of the b of 40%, and a tagging efficiency of 40%, the $7k$ generated signal events corresponds to $6.4 \cdot 10^8$ B^0_d events at LHCb. With a probability that a b quark forms a B^0_d meson of 40%, the probability that at least one of the b's forms a B^0_d is 64 %. The event sample corresponds therefore to a total of 10^9 $b\bar{b}$ events or 5 h of LHCb data taking[2].

Generic $b - \bar{b}$ Events: 500 000 $b\bar{b}$ events generated over 4π are used for background estimates. It is expected, that the trigger will highly suppress events without a B hadron decay. Therefore an event set containing only $b\bar{b}$ events is sufficient for this study. Only those of the 500 k $b\bar{b}$ events that pass trigger levels 0 and 1 are written to tape and available for the analysis (43 473 events). Assuming the same tagging efficiency as for the signal events, this corresponds to $1.25 \cdot 10^6$ $b\bar{b}$ events before tagging, or 22 s of LHCb data taking.

3.3.2 Reconstruction and Preselection

Because the flight distance of the D^* is effectively zero, there is only the primary vertex and two decay vertices to reconstruct: the decay vertex of the B^0_d and the decay vertex of the $\overline{D^0}$.

Some loose cuts on the track and vertex quality, the reconstructed particle masses, and the RICH information are applied in the reconstruction:

- Each track is required to be reconstructed from at least 30 measurement points.
- The χ^2 of the track fit is required to be less than 500.
- The χ^2 of the vertex fit is required to be less than 20.
- SICb provides a probability for each track to be of a certain particle type. This is based predominantly on the fast parametrisation of the RICH, with

---

[2] Assuming $5.6 \cdot 10^{11}$ $b\bar{b}$ pairs at LHCb in 10^7 seconds

some information taken into account from the ECAL and the muon chambers. All tracks are required to be 'consistent' (probability > 5%) with the required particle type.

- The reconstructed masses of the D^* and the D^0 are required to be within ±100 MeV of their true value.
- The mass of the reconstructed B^0_d is required to be within ±1 GeV of the true value.

In cases in which two possible B^0_d candidates share one or more tracks, only the one with the better χ^2 on the B^0_d-vertex fit is kept.

3.3.3 Selection Cuts

The selection cuts applied, and their effect, are listed in table 3.1 for signal events, for combinatoric background within signal events, and for the background sample. Events are selected where

1. the transverse momentum (p_t) of the fast pion candidate is larger than 1 GeV;
2. the average p_t of all tracks in the decay is larger than 0.5 GeV;
3. the impact parameter of the fast pion is larger than 50 μm;
4. the decay length of the B^0_d candidate is larger than 0.5 mm;
5. the mass difference between the D^* and the D^0 candidate is between 143.5 and 146.5 MeV;
6. the reconstructed mass of B^0_d candidate is within ±30 MeV of the true B^0_d mass. With the available statistics, no background event passes all cuts. In order to make a meaningful background estimate, therefore, the number of events surviving the wide preselection mass window of ±1 GeV is also recorded.

By far the most powerful cut is the one on the mass difference between the reconstructed D^* and $\overline{D^0}$. Figure 3.1 shows the signal and the background distribution for the mass difference. The plots show all reconstructed and triggered events, without any selection cuts applied.

3.3.4 Event Yield and S/B Estimates

Yield

After all cuts are applied, 140 signal events remain. Assuming $5.6 \cdot 10^{11}$ $b\bar{b}$ events per year at LHCb, and a combined efficiency of 92% for trigger levels 2 and 3[3], this corresponds to an annual yield of (73 ± 6) k triggered, tagged and reconstructed $B^0_d \to D^*\pi$ events. Table 3.2 shows the event yields with and without trigger levels 0 and 1 applied. The combined trigger level 0 and level 1 efficiency is 26%.

---

[3] This assumes the level 2 efficiency for $B^0_s \to D^-_s\pi^+$, given in table 2.2 (page 45), and a 100% efficient level 3.

Table 3.1: Exclusive reconstruction: Selection cuts. For each cut, both the number of events is given that pass the selection if *only* this cut is applied and the number if all cuts *except* this cut are applied. The number of events after the preselection is given in the first row of the table. The number of events passing all cuts is given in the last row.

Cut	Signal		Combinatorical Background		Background	
	only this cut	all except this cut	only this cut	all except this cut	only this cut	all except this cut, with presel. B^0_d-mass cut of ±1 GeV
triggered (level 1) and topology reconstructed	196		208		6058	
(1) $p_t(\pi_{\text{fast}}) > 1\,\text{GeV}$	182	152	131	0	3705	8
(2) $\langle p_t \rangle_{\text{all}} > 0.5\,\text{GeV}$	196	147	171	0	3218	3
(3) $IP > 50\,\mu\text{m}$	169	159	156	0	4305	3
(4) decay length $> 500\,\mu\text{m}$	187	141	175	0	4811	4
(5) B vtx quality: $\chi^2 < 30$	193	140	85	2	1758	48
(6) D^*-D^0 $\Delta m \in [143.5, 146.5]$ MeV	190	144	44	10	121	236
(7) B^0_d-mass within ±30 MeV (only sg)	186	147	47	1	–	–
All cuts	140		0		1	
... narrow mass–window for background					0	

Table 3.2: Trigger Efficiency (exclusive)

	events	efficiency
no trigger	534	
level 0	221	42%
level 1	140	63%
combined efficiency		26%

Background

Applying all cuts except the final mass cut leaves one background event, which has a reconstructed mass of 4.6 GeV. With this one event, the statistical error[4] on the number of background events within the wide mass window is $^{+2.3}_{-0.83}$.

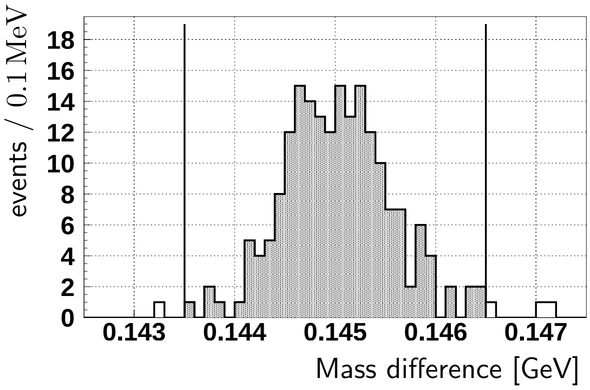
The ratio of background events inside the

---

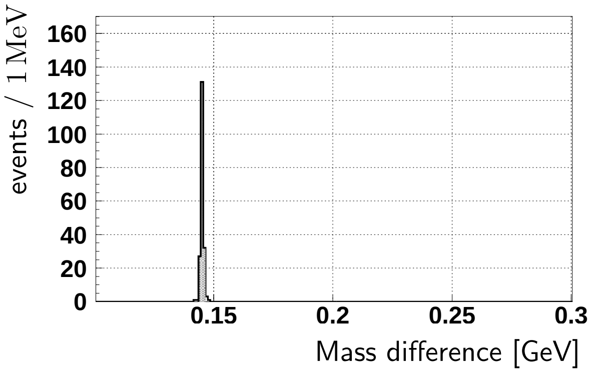
[4] 68.2% central interval confidence limit for a Poisson distribution

Figure 3.1: Mass difference between reconstructed D^* and $\overline{D^0}$ for signal and background.

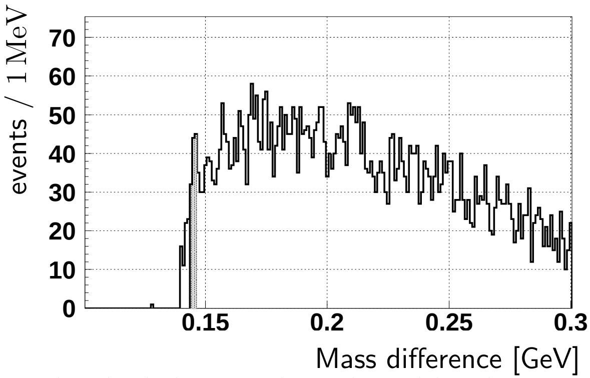
Signal



Signal (scale in x reduced by factor 40)



Background



The shaded interval is selected by the cut

wide mass window to the number inside the narrow mass window is 49, before any cut is applied. Applying all cuts except the preselection requirement on the track qualities of the D^0 decay products, and the cut on the D^*–D^0 mass difference, leaves 305 background events inside the wide mass window and 9 inside the narrow one, giving a ratio of $305/9 = 34$.

Removing instead the preselection requirements on the track qualities of the π_{fast} and π_{slow}, and applying only the cut on the D^*–D^0 mass difference, gives a ratio of $288/9 = 32$.

Assuming that after all cuts are applied, there are 32 times as many background events in the wide mass window as there are in the narrow mass window, leads to an estimate of the signal to background ratio of $5.6^{+26.8}_{-3.9}$. The large uncertainty is due to the statistical fluctuations on the one background event. Neglecting other error contributions, the signal to background ratio is better than 1.4 at the 90% confidence-level.

3.3.5 Mass and Time Resolution

The B^0_d mass distribution for the reconstructed signal events is shown in figure 3.2. The Gaussian fit to the mass-distribution has a width of 13.6 MeV.

An important parameter for the $\mathcal{CP}$ violation measurements is the time resolution. A Gaussian fit to the time resolution, shown in figure 3.3, returns a width of 60 fs. This could be improved further if a constrained

Figure 3.2: Reconstructed B_d^0 mass

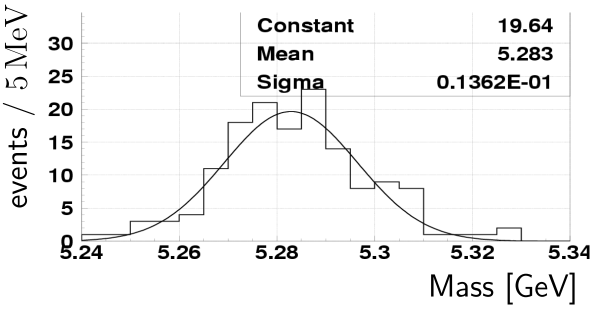
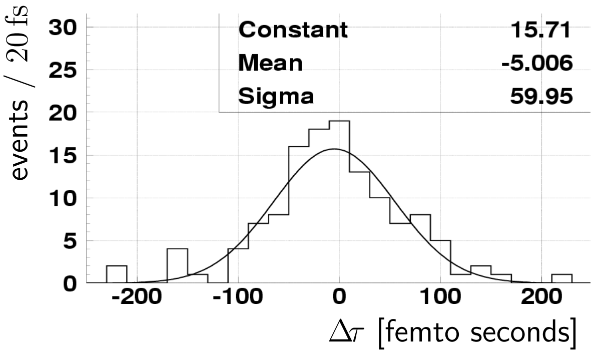


Figure 3.3: Time resolution (exclusive)



fit for the B_d^0 decay vertex reconstruction were to be used. However, 60 fs is 1/200 of the B_d^0-$\overline{B_d^0}$ oscillation period and therefore perfectly adequate for the measurement.

3.3.6 Result Summary

73 000 reconstructed, triggered and tagged exclusive $B_d^0 \to D^*\pi$ decays are expected per year at LHCb. A good signal-to-background ratio, time and mass-resolution can be achieved. A summary of the results is shown in table 3.3.

Table 3.3: $B_d^0 \to D^{*-}\pi^+$ exclusive: Result summary

Signal yield	(73 ± 6) k per year
Signal/Backg	$5.6^{+26.8}_{-3.9}$
	> 1.4 at 90% confidence
mass-resol$^{\underline{n}}$	13.6 MeV
time-resol$^{\underline{n}}$	60 fs

3.4 Inclusive Reconstruction of $B_d^0 \to D^*\pi$

With the inclusive reconstruction method, all decays $B_d^0 \to D^{*-}\pi^+$ with $D^{*-} \to \overline{D^0}\pi^-$ are reconstructed simultaneously, thus increasing the visible branching ratio by a factor of 20. Only the slow and the fast pion are used in the reconstruction of the B_d^0 momentum. Two charged tracks from the $\overline{D^0}$ decay are used for background reduction.

3.4.1 Datasets

For this analysis, a new set of signal data was generated, and the number of $b\bar{b}$ events for the background studies was increased:

1. 20 000 events with at least one $B_d^0 \to D^{*-}\pi^+$ decay each, are generated such that the momentum of the b quark forming the B_d^0 meson lies within a 600 mrad forward cone. As previously, no tagging algorithm is applied, and all events are treated as if they were tagged. The finite tagging efficiency is taken into account when calculating the final event yields. With a branching ratio for this decay of $2.76 \cdot 10^{-3}$,

a probability for a b to be within the 600 mrad forward cone of 40%, and a tagging efficiency of 40%, the sample corresponds to $45.3 \cdot 10^6$ B^0_d events, or $70.8 \cdot 10^6$ $b\bar{b}$ events at LHCb, which is equivalent to 21 min of LHCb operation.

2. The same set of 500 k events as used for the exclusive case, plus 274 k new $b\bar{b}$ events, generated over 4π, are used for background studies. Of those new 274 k events, all are written to tape and are available for the analysis, not only those that pass the trigger. Assuming the same tagging efficiency as for the signal events, this corresponds to a total of $1.94 \cdot 10^6$ $b\bar{b}$ events before tagging, or 35 s of LHCb data taking.

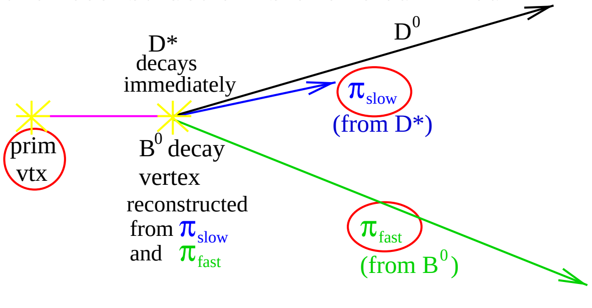
3.4.2 Principle

Below, we will show that this inclusive reconstruction is possible, and that, despite the missing information from the $\overline{\mathrm{D}^0}$ momentum, the full B^0_d 4–momentum can be obtained from the available parameters.

Parameter Counting

The task is to reconstruct the 4–momentum of the B^0_d from the momentum of the fast pion coming from the $\mathrm{B}^0_\mathrm{d} \rightarrow \mathrm{D}^*\pi_\mathrm{fast}$ decay, and the slow pion from the $\mathrm{D}^* \rightarrow \overline{\mathrm{D}^0}\pi_\mathrm{slow}$ decay. The topology of the decay in the laboratory frame of reference is illustrated in figure 3.4.

Figure 3.4: $\mathrm{B}^0_\mathrm{d} \rightarrow \mathrm{D}^*\pi$ decay in space (not drawn to scale). The information used in the reconstruction is encircled in red.

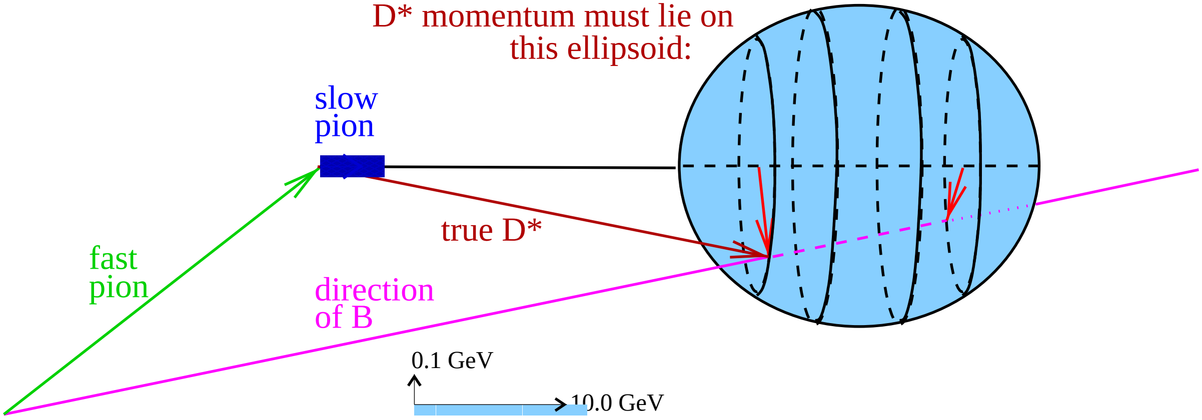


Given the 4–momenta of the two pions, there remain 8 missing parameters, the 4–momenta of the B^0_d and of the D^*. Energy and momentum conservation provide 4 constraints. The D^* mass provides another constraint by relating the energy and momentum of the D^*. The D^0 mass constrains the magnitude of the D^* momentum in the π_slow rest–frame. The remaining two constraints come from the knowledge of the direction of the B^0_d momentum, obtained from the primary vertex and B^0_d-decay vertex position. The B^0_d-decay vertex can be reconstructed from the slow and the fast pion as the flight-distance of the D^* is negligible.

Formulae

In the following the superscript “lab” indicates the laboratory frame and “S” the rest–frame of the slow pion, π_slow. The π_slow rest–frame is chosen rather than the rest frame of the D^*, because it is the π_slow that is detected, and from its 3–momentum and mass we directly get the Lorentz transformation from its rest–frame to the lab–frame:

Figure 3.5: Equation 3.6 corresponds to finding the intercepts between a line and an ellipsoid in momentum space. The vertical scale is enlarged with respect to the horizontal one by a factor of ∼ 20.



$$\begin{pmatrix} E^{\text{lab}} \\ \vec{p}^{\,\text{lab}} \end{pmatrix} = \begin{pmatrix} \frac{E^{\text{lab}}_{\pi}}{m_\pi} & {}^t\!\left(\vec{p}^{\,\text{lab}}_{\pi_s}/m_\pi\right) \\ \frac{\vec{p}^{\,\text{lab}}_{\pi_s}}{m_\pi} & \delta_{ij} + \frac{\left(\frac{E^{\text{lab}}_{\pi_s}}{m_\pi} - 1\right) p^{\text{lab}}_{\pi_s i} p^{\text{lab}}_{\pi_s j}}{\left|\vec{p}^{\,\text{lab}}_{\pi_s}\right|^2} \end{pmatrix} \begin{pmatrix} E^{\,\text{S}} \\ \vec{p}^{\,\text{S}} \end{pmatrix} \tag{3.2}$$

where the spatial entries in the transformation matrix are written in a compact notation with the indices i, j going from 1 to 3. Then the momentum of the D^* in the lab–frame is:

$$\begin{aligned} \vec{p}^{\,\text{lab}}_{D^{*-}} &= \left(\alpha_0 + \Delta\alpha \cos(\theta^S)\right) \vec{p}^{\,\text{lab}}_{\pi_s} \\ &+ \left|\vec{p}^{\,S}_{D^{*-}}\right| \left(1 - \cos^2(\theta^S)\right)^{\frac{1}{2}} \vec{u} \end{aligned} \tag{3.3}$$

where θ^S is the angle of the D^* momentum in the π_s rest–frame with respect to $\vec{p}^{\,\text{lab}}_{\pi_s}$ and $\vec{u}$ is a unit vector perpendicular to $\vec{p}^{\,\text{lab}}_{\pi_s}$. The parameters $\alpha_\circ$ and $\Delta\alpha$ are given by:

$$\alpha_\circ \equiv \frac{E^S_{D^{*-}}}{m_{\pi^-}}, \ \Delta\alpha \equiv \frac{\left|\vec{p}^{\,S}_{D^{*-}}\right|}{m_\pi} \left(1 + \frac{m_\pi^2}{\left|\vec{p}^{\,\text{lab}}_{\pi_s}\right|^2}\right)^{\frac{1}{2}}. \tag{3.4}$$

All quantities on the right hand side of equation 3.3 are known, except for two: θ^{S} and $\vec{u}$. The 3–vector $\vec{u}$ corresponds to only one unknown, as it is defined to have unit length and to be perpendicular to $\vec{p}^{\,\text{lab}}_{\pi_s}$.

The geometrical interpretation of 3.3 is that the momentum vector of the D^* lies on an ellipsoid in momentum space. The ellipsoid is rotationally symmetric around the axis defined by the π_{slow} momentum, and has a major axis of length $\Delta\alpha \left|\vec{p}^{\,\text{lab}}_{\pi_s}\right| \approx 3.9 \left|\vec{p}^{\,\text{lab}}_{\pi_s}\right|$ and a minor axis of length $\left|\vec{p}^{\,S}_{D^{*-}}\right| = 550\,\text{MeV}$.

The direction of the $\mathrm{B}^0_{\mathrm{d}}$, $\vec{V}$, and the momentum of the fast pion, $\vec{p}_{\pi_f}$, are related by:

$$\lambda \vec{V} = \vec{p}_{\pi_f} + \vec{p}_{D^*} \tag{3.5}$$

where λ is unknown and $\lambda\vec{V}$ is the 3–momentum of the $\mathrm{B^0_d}$. Adding this information reduces the number of unknowns from five in equation 3.3, to three unknowns for three equations. Dropping the "lab" superscripts:

$$\begin{aligned} \lambda \vec{V} &= \vec{p}_{\pi_f} + \left(\alpha_\circ + \Delta\alpha \cos(\theta^S)\right) \vec{p}_{\pi_s} \\ &+ \left|\vec{p}^{\,S}_{D^{*-}}\right| \left(1 - \cos^2(\theta^S)\right)^{\frac{1}{2}} \vec{u}. \end{aligned} \tag{3.6}$$

The geometrical interpretation of this equation is illustrated in figure 3.5. The solutions are the intercepts between an ellipsoid of possible D^* momenta, and the line of possible $\mathrm{B^0_d}$ momenta.

In the following we will find the algebraic solution to this problem in terms of the parameters b and $\vec{u}$, where b is defined as

$$b \equiv \cos\left(\theta^S\right). \tag{3.7}$$

We first multiply both sides of equation 3.6 with $\vec{p}_{\pi_{\mathrm{slow}}}$ to get:

$$\begin{aligned} \lambda &= \frac{\alpha_\circ + b\,\Delta\alpha \vec{p}^{\,2}_{\pi_{\mathrm{slow}}} - \vec{p}_{\pi_{\mathrm{fast}}} \cdot \vec{p}_{\pi_{\mathrm{slow}}}}{\vec{V} \cdot \vec{p}_{\pi_{\mathrm{slow}}}} \\ &= b\,A + B, \end{aligned} \tag{3.8}$$

with

$$A \equiv \frac{\Delta\alpha \vec{p}^{\,2}_{\pi_{\mathrm{slow}}}}{\vec{V} \cdot \vec{p}_{\pi_{\mathrm{slow}}}}, B \equiv \frac{\Delta\alpha \vec{p}^{\,2}_{\pi_{\mathrm{slow}}} + \vec{p}_{\pi_{\mathrm{fast}}} \cdot \vec{p}_{\pi_{\mathrm{slow}}}}{\vec{V} \cdot \vec{p}_{\pi_{\mathrm{slow}}}}. \tag{3.9}$$

Inserting this back into equation 3.6 gives:

$$\begin{aligned} &\left(1 - b^2\right)^{\frac{1}{2}} \vec{u} = \\ &\quad \frac{b\,\left(A\vec{V} - \Delta\alpha \vec{p}_{\pi_{\mathrm{slow}}}\right)}{\left|\vec{p}^{\,S}_{D^{*-}}\right|} \\ &+ \quad \frac{B\vec{V} - \alpha_\circ \vec{p}_{\pi_{\mathrm{slow}}} - \vec{p}_{\pi_{\mathrm{fast}}}}{\left|\vec{p}^{\,S}_{D^{*-}}\right|}. \end{aligned} \tag{3.10}$$

Defining

$$\vec{G} \equiv \frac{A\vec{V} - \Delta\alpha \vec{p}_{\pi_{\mathrm{slow}}}}{|\vec{p}^{\,S}_{D^{*-}}|}, \tag{3.11}$$

$$\vec{H} \equiv \frac{B\vec{V} - \alpha_\circ \vec{p}_{\pi_{\mathrm{slow}}} - \vec{p}_{\pi_{\mathrm{fast}}}}{|\vec{p}^{\,S}_{D^{*-}}|}, \tag{3.12}$$

this can be written as

$$\left(1 - b^2\right)^{\frac{1}{2}} \vec{u} = b\,\vec{G} + \vec{H}. \tag{3.13}$$

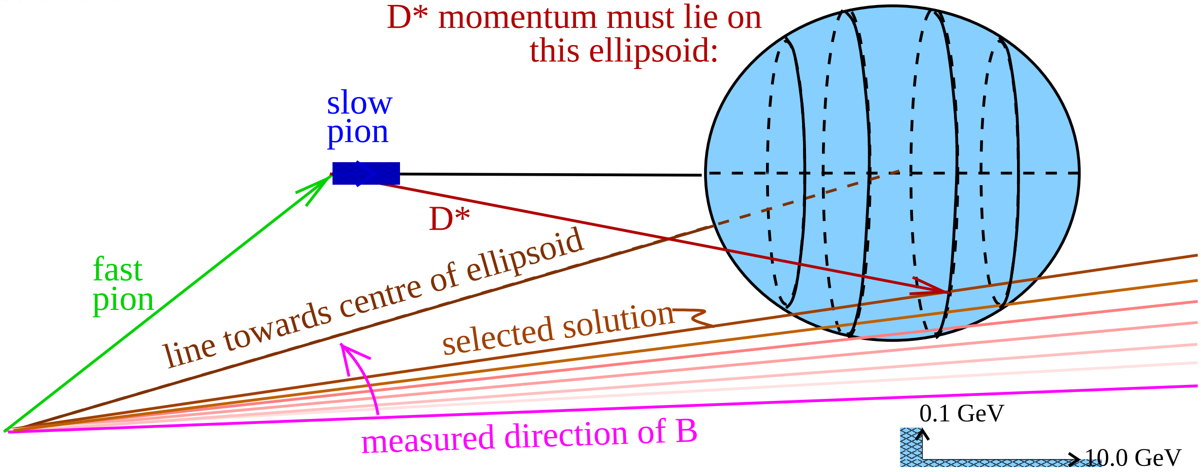
Squaring equation 3.13 gives a quadratic equation in b with the following solutions:

$$b_{1,2} = -\frac{\vec{G} \cdot \vec{H}}{1 + \vec{G}^2} \pm \left(\left(\frac{\vec{G} \cdot \vec{H}}{1 + \vec{G}^2}\right)^2 - \frac{\vec{H}^2 - 1}{1 + \vec{G}^2}\right)^{\frac{1}{2}}. \tag{3.14}$$

Inserting the solution for b back into equation 3.13 gives the solution for the vector $\vec{u}$; inserting both into equation 3.3, gives the 3–momentum of the D^*. Combined with the D^* mass, and the 4–momentum of the fast pion, the full $\mathrm{B^0_d}$ 4–momentum can be reconstructed.

Two solutions: If there are two real solutions to equation 3.14, the solution which

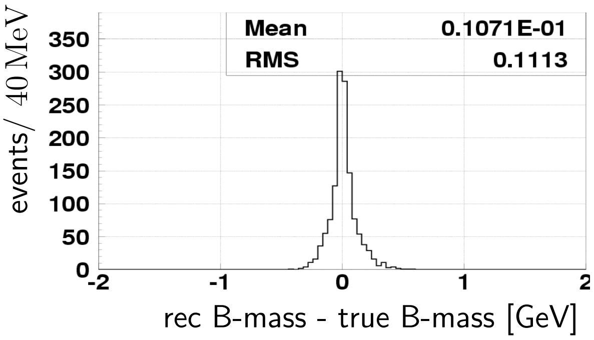
Figure 3.6: Approximating the $\mathrm{B^0_d}$ momentum if equation 3.14 does not have any real solutions.



returns the 'better' value for the $\mathrm{B^0_d}$–mass is chosen. This solution is used for calculating all reconstruction parameters, except for the $\mathrm{B^0_d}$–mass itself. Here, in order to avoid biases in the background estimation, the $\mathrm{B^0_d}$–mass used in the final mass cut is taken as the average of the masses obtained from both solutions. In fact, because $\left|\vec{p}^{\,S}_{D^{*-}}\right|$ is very small, the ellipsoid in figure 3.5 is thin, so the two solutions, and their values for the reconstructed $\mathrm{B^0_d}$-mass, lie very close to each other.

The effect this approximation has on the mass resolution is illustrated in figure 3.7. It shows the reconstructed $\mathrm{B^0_d}$ mass minus the true $\mathrm{B^0_d}$ mass, calculated using Monte Carlo truth information. If always the solution with the better $\mathrm{B^0_d}$ mass had been chosen, the exact $\mathrm{B^0_d}$ mass would have been obtained every time, and this plot would show a δ function at zero. The error introduced by taking the average mass instead is $\sim 0.1\,\mathrm{GeV}$.

Figure 3.7: Monte-Carlo truth for inclusive $\mathrm{B^0_d} \to \mathrm{D}^*\pi$: reconstructed mass - true mass, no cuts applied. The plot shows the effect of using the average $\mathrm{B^0_d}$ mass if there are two solutions, rather then the 'better' one.



No solution: In case equation 3.14 returns no real solution, which means that the measured $\mathrm{B^0_d}$-direction does not go through the ellipsoid, a solution is found by varying the $\mathrm{B^0_d}$-direction as illustrated in figure 3.6. The line indicating the $\mathrm{B^0_d}$-direction is rotated towards the centre of the ellipsoid until it meets the surface of the ellipsoid;

this intercept point is taken as the solution. In the reconstruction routine, a tolerance is placed on how much the corrected B^0_d-direction deviates from the measured one.

3.4.3 Preselection

For convenience, the analysis is performed on a set of preselected events, which are obtained by applying loose cuts on some basic parameters.[5] The preselection program searches for combinations of four charged tracks which meet the listed criteria. These tracks are candidates for the π_fast and the π_slow needed for the B^0_d–reconstruction, and two oppositely charged decay-products from the D^0, in the following referred to as X^+, Y^-. The latter are needed for the cut on the D^*–D^0 mass difference described in section 3.4.6. All combinations of four tracks that satisfy the following preselection criteria are kept:

1. The χ^2 of the B^0_d–vertex reconstruction is less than 25.
2. The transverse momentum of π_fast is larger than 1.5 GeV.
3. The impact parameter significances of π_fast and π_slow with respect to the primary vertex are larger than 2.
4. The B^0_d–vertex is downstream of the primary vertex.
5. The D^0–vertex is downstream of B^0_d–vertex.
6. The impact parameter significance of X^+, Y^- with respect to the primary vertex is larger than 2.
7. The pseudo mass of D^0, calculated from X^+ and $Y-$, is smaller than 2.3 GeV.
8. The D^*–D^0 mass difference, calculated as described below, is less than 0.5 GeV.
9. The impact parameter of π_slow with respect to the D^0–vertex is less than 0.3 mm.
10. The reconstructed B^0_d–mass is less than 20 GeV.

All criteria, except for number 2 and number 9, are also re–applied in the final event selection with tighter cuts. The motivation for cut number 9 is that, because the π_slow and the D^0 momentum in the D^* rest–frame are very small, the D^0 and the π_slow momentum are collinear in the lab–frame to a good approximation.

3.4.4 Event Selection

Due to the reduced information compared to the exclusive case, the inclusive reconstruction is more susceptible to background contamination. This requires hard cuts on the usual B^0_d physics criteria as well as on parameters like the track quality and the vertex χ^2.

Four particle tracks are reconstructed: the two pions necessary for the B^0_d reconstruction, π_fast and π_slow, and two oppositely

---

[5] The program to perform the preselection was kindly supplied by John Holt (Oxford, CERN).

charged tracks from the D^0-decay, X^+, Y^-, used for the cut on the D^*–D^0 mass difference.

The cuts and their effects are summarised in table 3.4, 88.

The two cuts that are specific to the inclusive reconstruction are the cut on the reconstruction parameters, and the cut on the D^*–D^0 mass difference. These are described in the following two sections.

3.4.5 Cutting on Reconstruction Parameters

For the case that there are no real solutions to equation 3.14, the candidate B_d^0 direction is adjusted. This occurs about 40% of the time for well reconstructed signal events. The new direction is required to lie within 2σ of the original direction, where σ is calculated assuming a constant vertex resolution of $\sigma_{x,y} = 9\,\mu\text{m}, \sigma_z = 38\,\mu\text{m}$ for the primary, and $\sigma_{x,y} = 20\,\mu\text{m}, \sigma_z = 250\,\mu\text{m}$ for the secondary vertex. The values for the primary vertex are taken from [TP98]. The values for the secondary vertex are obtained by taking the width of the core Gaussian in fits to the vertex resolution shown in figure 3.8. The plots show the reconstructed vertex position minus the true vertex position separately for the x, y, and z coordinate, and are made with a looser selection in order to increase statistics. The cuts used are indicated with a "$*$" in table 3.4. Here, if more than one combination of tracks passes all cuts in a given event, Monte Carlo truth information is used to select the signal combination, rather than the algorithm used in the final event selection which is described later.

The efficiency of the cut on the correction to the B_d^0 direction is shown in figure 3.9. The top–plot shows the number of signal events as a function of the adjustment of the B_d^0 direction in terms of σ. The bottom plot shows the fraction of events below a given cut. The signal efficiency for a 2σ cut is better than 90%. The signal events selected for these plots are the same as those used in the vertex resolution plots above. For the background sample, the same cuts are applied, but events are selected according to the algorithm described in section 3.4.7 in case more than one event passes all cuts. The cut rejects 94% of the background events.

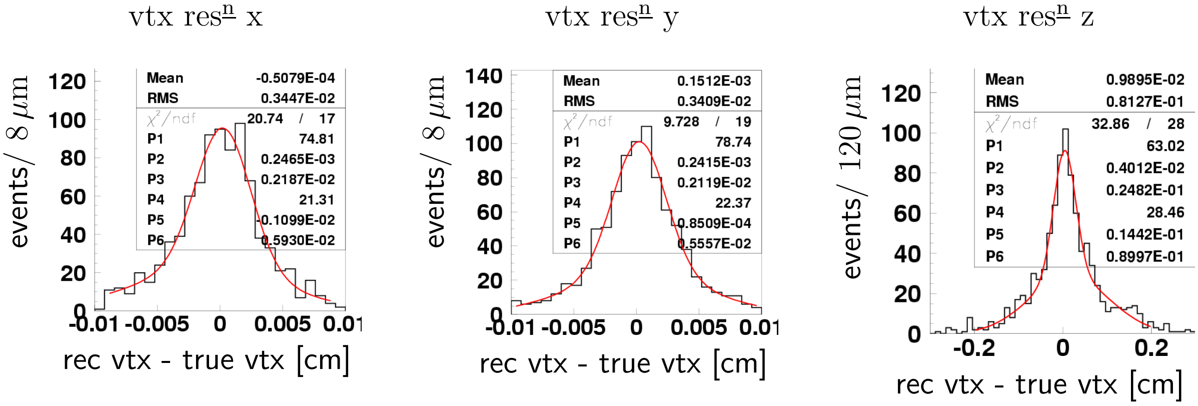
3.4.6 Cutting on the D^*-D^0 Mass Difference

Although the D^0 is not fully reconstructed, the cut on the D^*-D^0 mass difference can still be applied to suppress background, although with reduced efficiency. Two charged tracks from the D^0, X^+, Y^-, are selected. The invariant mass of these two tracks is subtracted from the invariant mass calculated from X^+, Y^- and the slow pion:

$$\begin{aligned}\Delta m \equiv\ & M\left(X^+, Y^-, \pi_{\text{slow}}\right) \\ & - M\left(X^+, Y^-\right) \\ =\ & \sqrt{\left(p_{X^+} + p_{Y^-} + p_{\pi_{\text{slow}}}\right)^2} \\ & - \sqrt{\left(p_{X^+} + p_{Y^-}\right)^2} \qquad (3.15)\end{aligned}$$

In the above equation, p_{X^+} etc are 4–momenta, and p^2 is to be intepreted as $p^\mu p_\mu$.

Figure 3.8: Vertex resolution with double–Gaussian fit. The parameters (P1,P4), (P2,P5), (P3,P6) are the amplitudes, means and widths respectively of the two Gaussians fitted to the distributions.



vtx res$^{\underline{n}}$ x vtx res$^{\underline{n}}$ y vtx res$^{\underline{n}}$ z

If X^+, Y^- constitute all decay products of the D^0, this is simply the D^*-D^0 mass difference. In general though, there will be some missing momentum. In the expression for the mass difference, this missing momentum cancels to some degree, so that this is still a powerful cut for the inclusive case, as shown in figure 3.10. For these plots, as well as in the final selection, the two charged particles are required to form a vertex at least 1σ downstream of the B^0_d vertex, with a vertex χ^2 of less than 4 and an impact parameter significance with respect to the primary vertex of more than 3. The pseudo mass formed from the two tracks is required to be within $[0.4, 1.9]$ GeV.

3.4.7 Cut Summary and Event Yield Estimate

Cut Summary

Below, all cuts are listed with a short description. The effect of the cuts on the number of signal and background events is summarised in table 3.4.

Most cuts applied in the final event selection are either not applied in the preselection, or they are applied with looser requirements. The only exceptions are the following two cuts that are left at their preselection values:

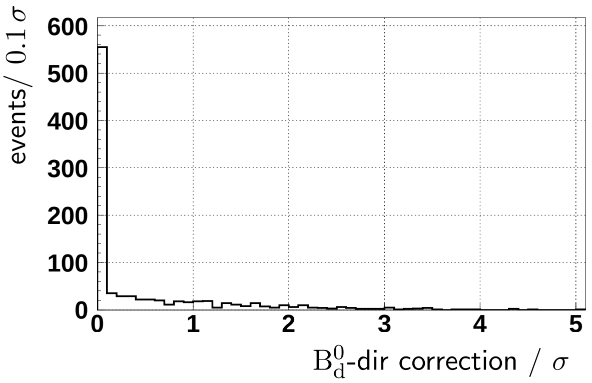
(i) **Transverse Momentum of $\pi_{\text{fast}} > 1.5\,\text{GeV}$:** This cut selects B decays by their large mass.

(ii) **Impact Parameter of π_{slow} with respect to D^0 vertex $< 0.3\,\text{mm}$:** The D^0 and the π_{slow} are nearly at rest in the D^* frame and therefore nearly collinear in the lab–frame. This cut on the vertex formed by X^+, Y^- therefore selects D^0 candidates that are likely to come $D^{*-} \rightarrow \overline{D^0}\pi^-$ decays.

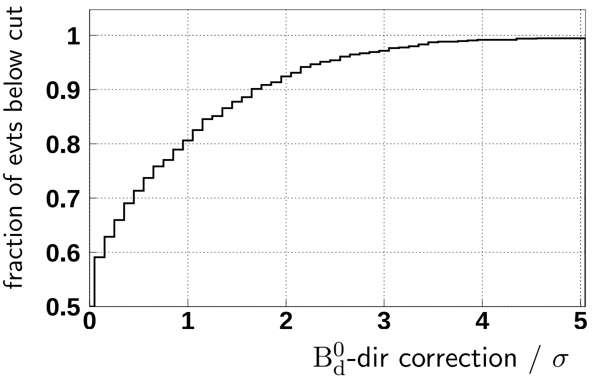
The cuts that are applied in the final event selection are listed below:

Figure 3.9: The cut on the correction of the $\mathrm{B^0_d}$-direction, with reconstruction-quality cuts applied.

Number of signal events



Fraction of signal events below cut.



Fraction of background events below cut.

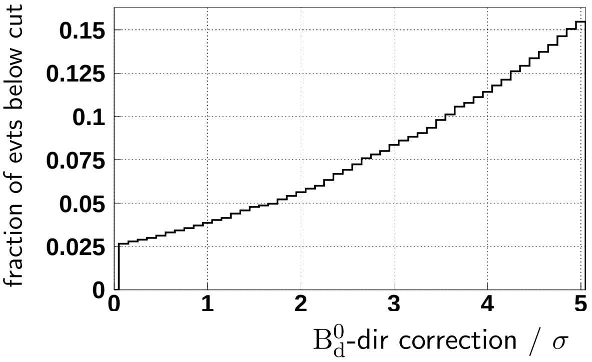
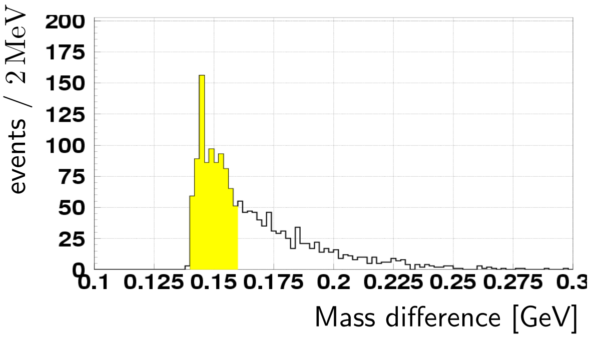
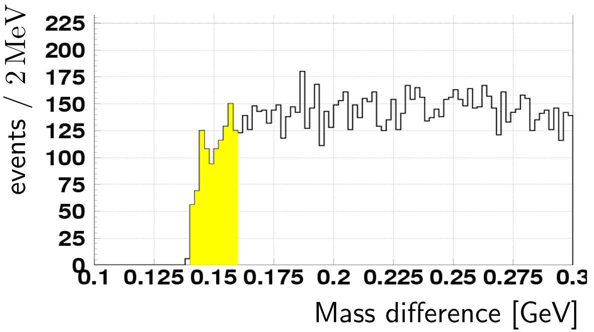


Figure 3.10: Mass difference between reconstructed D^* and $\overline{\mathrm{D^0}}$ for signal and background, in inclusive reconstruction.

Signal



Background



The shaded interval is selected by the cut

1. **Reconstruction cut:** This is the cut on the correction to the $\mathrm{B^0_d}$ direction, necessary in about 40% of the signal events to find a solution to equation 3.14. This cut is described in detail in section 3.4.5.

2. π_{fast} **RICH info:** It is required that the particle identification provided by the RICH pattern recognition is either 'pion' or 'muon'. For momenta above 150 GeV, which are beyond the K/π

separation capabilities of the RICH, tracks with particle assignment ‘kaon’ are also accepted.

3. π_{slow} **RICH info** the same as for π_{fast}, except that ‘kaon’ is never accepted.

4. π_{fast} **momentum** $<$ 200 **GeV** This ensures that the π_{fast} momentum is within the dynamic range of the electromagnetic calorimeter.

5. π_{fast}, π_{slow} **electron–rejection:** Electrons are rejected using a combination of RICH and calorimeter information. As suitable calorimeter reconstruction is not available, the calorimeter performance is approximated using Monte Carlo truth information. For particles within its acceptance, and momenta between 5 and 100 GeV, it is assumed that the calorimeter provides perfect e/hadron discrimination. All electrons within that momentum range within the calorimeter acceptance are rejected. It is assumed that the calorimeter rejects electrons with 90 % efficiency up to 200 GeV.[6] All π_{fast} candidates that would not have a hit in the calorimeter because they are outside its acceptance are rejected, unless the RICH can exclude the possibility that they are electrons (e-probability of $< 5\%$).

6. π_{fast}, π_{slow} **muon–rejection** If either of the two pions has a hit in one of the muon–chamber station 2, 3, 4, or 5, the event is rejected.

7. B^0_d**–vertex** $\chi^2 < 2$**:** This ensures that the π_{fast}–π_{slow} vertex is well measured.

8. **Track** χ^2/dof $<$ **1 of** π_{fast} **and** π_{slow}**:** This ensures that both tracks forming the B^0_d vertex are well measured.

9. **Impact Parameter significance of pions with respect primary vertex** > 4**:** The impact parameter significance is defined as the impact parameter (IP) divided by the error estimate for the impact parameter. Both pions are required to have an impact parameter with respect to the primary vertex of at least 4σ, and are thus compatible with coming from a B^0_d decay.

10. $\mathbf{p_t}$ **of** B^0_d $<$ **60 GeV** A cut on the reconstructed transverse momentum of the B^0_d. This is not a “physics cut”, as both signal and background events tend to have much smaller p_t than 60 GeV. It is applied to eliminate badly reconstructed events and follows the observation that such events can sometimes possess very large values for the p_t of the B^0_d.

11. D^0**–vertex** $\chi^2 < 4$**:** This ensures that the D^0 vertex is well reconstructed.

12. D^0 **vertex downstream of** B^0_d **vertex:** This cut demands that the reconstructed D^0 vertex is at least 1σ downstream of the reconstructed B^0_d vertex, and is thus compatible with being part of the B^0_d decay chain.

---

[6] In the event selection, electrons with momenta $>$ 100 GeV simply pass the cut. If such an event had passed all cuts, it would have been weighted accordingly.

13. **IP-significance of $\mathbf{X^+, Y^-}$:** The impact parameter significance with respect to the primary vertex for each of the two charged tracks from the $\mathrm{D^0}$ is required to be at least **3**. This ensures that X^+ and Y^- are compatible with a B decay.

14. $\mathrm{D^0}$ **mass:** The mass of the partially reconstructed $\mathrm{D^0}$, obtained by adding the 4–momenta of the X^+ and Y^- candidate and forming the invariant mass, is required to be within $[0.4, 1.9]$ **GeV**. The lower limit ensures that the X^+ and Y^- candidates carry a significant part of the $\mathrm{D^0}$ momentum. The upper limit ensures that the candidates are not incompatible with coming from a $\mathrm{D^0}$ decay.

15. $\mathrm{D^*}$–$\mathrm{D^0}$ **Mass difference:** The difference in invariant mass of the combined $X^+, Y^-, \pi_{\mathrm{slow}}$ momentum to the combined X^+, Y^- momentum is required to be within $[140, 160]$ **MeV**. This is a powerful cut for the identification of $\mathrm{D^*}$ decays, as described in detail in section 3.4.6

16. $\mathrm{B_d^0}$**-mass:** The reconstructed $\mathrm{B_d^0}$–mass is required to be within $\pm 0.2\,\mathrm{GeV}$ of the true $\mathrm{B_d^0}$ mass. As, with this mass cut, no background event survives, the mass–window for background events is widened to $\pm 2.0\,\mathrm{GeV}$ to increase statistics.

Cuts 5 and 6 were introduced after it had been found that, without those cuts, semileptonic $\mathrm{B_d^0}$ decays such as $\mathrm{B_d^0} \to \mathrm{D^{*-}}\mu^+\nu_\mu$ constitute the dominant part of the background. Those cuts on the track χ^2/dof appear very tight, but given the distribution found in the simulation, with typical values around 0.6, they have a high efficiency.

Deciding Between Suitable Candidates

In case there is more than one combination of tracks that passes all cuts, a combination is selected according to the following algorithm:

1. Choose the one with the better $\mathrm{B_d^0}$ vertex χ^2.
2. If the two χ^2 differ by less than ±20%, choose the one with the higher p_t of the fast pion.
3. If the above differ by less than ±20%, choose the one with the better $\mathrm{D^0}$ vertex χ^2.
4. If the above differ by less than ±20%, choose the one where the IP significance of π_{slow} with respect to the $\mathrm{D^0}$ decay vertex is smallest.
5. If the above differ by less than ±20%, choose the one where the $\mathrm{D^*}$–$\mathrm{D^0}$ mass difference is closest to its theoretical value.

This selection procedure is only relevant if a highly reduced set of cuts is used. Applying the full set of cuts, except the trigger, to the signal tapes, no event is found where more than one combination of π_{fast} and π_{slow} candidates survives all cuts.

Table 3.4: Inclusive reconstruction: Selection cuts.

Cut	SG–tapes						$b\bar{b}$-tapes					
	only this cut			not this cut			only this cut			not this cut		
	sel. SG	total SG	BG	sel. SG	total SG	BG	sel. SG	total SG	BG	sel. SG	total SG	BG
preselection, including p_t of $\pi_\mathrm{fast}> 1.5\,\mathrm{GeV}$, IP of π_slow wrt D^0 vtx $< 0.3\,\mathrm{mm}$, and trigger 0 and 1	**303**	522	1254				14	21	**16531**			
(1) reconstruction cut (*)	**437**	469	356	**65**	65	0	19	19	**4936**	4	4	**29**
(2) π_fast RICH-info	**268**	471	957	**74**	74	0	13	17	**11437**	4	4	**12**
(3) π_slow RICH-info	**315**	520	1210	**63**	63	0	14	21	**16141**	4	4	**6**
(4) π_fast mom $< 200\,\mathrm{GeV}^*$	**300**	510	1244	**63**	63	0	14	21	**16470**	4	4	**7**
(5) π_fast, π_slow calorimeter (MC-truth)	**280**	468	1140	**69**	69	0	13	20	**14023**	4	4	**13**
(6) π_fast, $\pi_\mathrm{slow}$muon-chamber info	**309**	522	1177	**63**	63	0	14	21	**13792**	4	4	**18**
(7) B-vtx $\chi^2 < 2^*$	**300**	399	1244	**80**	80	0	14	20	**12199**	4	4	**12**
(8) track χ^2/dof < 1 of π_fast and $\pi_\mathrm{slow}{}^*$	**278**	460	1201	**67**	67	0	13	18	**15307**	6	6	**7**
(9) IP signif of pions wrt prim vtx $> 4^*$	**279**	397	737	**73**	73	0	14	16	**9783**	6	6	**11**
(10) z-coord B_d^0-vtx - prim vtx $> 0.15\,\mathrm{cm}^*$	**331**	512	933	**63**	63	0	14	21	**16531**	4	4	**6**
(11) p_t of $\mathrm{B}_\mathrm{d}^0 < 60\,\mathrm{GeV}^*$	**303**	522	1043	**63**	63	0	15	21	**11651**	4	4	**16**
(12) D^0-vtx $\chi^2 < 4$	**283**	457	1094	**73**	73	0	14	20	**14646**	4	4	**9**
(13) z-coord D^0-vtx - B_d^0-vtx $> 1\sigma_\mathrm{dist}$	**302**	506	1222	**63**	63	0	13	20	**16220**	4	4	**6**
(14) IP-signif of X^+, Y^- wrt prim vtx > 3	**321**	487	1001	**63**	63	0	15	21	**14340**	4	4	**6**
(15) D^0 $M(X^+, Y^-) \in [0.4, 1.9]$ GeV	**257**	493	1196	**63**	63	0	11	18	**16115**	4	4	**7**
(16) D^*-D^0 $\Delta m \in [140, 160]$ MeV	**257**	285	285	**102**	102	0	11	11	**4263**	6	6	**23**
(17) B_d^0-mass within $\pm 0.2\,\mathrm{GeV(sg)}/\pm 2.0\,\mathrm{GeV(bg)}$	**299**	315	172	**85**	85	0	21	21	**13624**	4	4	**10**
All cuts (wide mass window for bg)	**63**	63	0				4	4	**6**			
(narrow mass window for bg)							4	4	**0**			

For each cut, the number of signal ("SG") and background ("BG") events surviving is given if either *only* this cut is applied ("only this cut"), or all cuts *except* this cut are applied ("not this cut"). For the SG events, the number of events selected using the algorithm in section 3.4.7 ("sel. SG"), and the total number left in the sample ("total SG") is given. Results are given separately for the set of signal events ("SG–tapes") and the set of generic $b\bar{b}$ events ("$b\bar{b}$-tapes"). Cuts marked with an asterisk (*) are also applied for the vertex, mass and time resolution measurements, but the trigger is not. The reconstruction cut is used for the resolution measurements, but not for the set of vertex resolution measurements that are subsequently used to *define* the reconstruction cut.

Table 3.5: Trigger Efficiency (inclusive)

	events	efficiency
no trigger	189	
level 0	90	48%
level 1	63	70%
combined efficiency		33%

Yields and Trigger Efficiency

After all cuts are applied, 63 signal events remain, corresponding to an annual event yield of 460 ± 60 k events. The efficiencies of trigger levels 0 and 1 are summarised in table 3.5. The four signal events found on the background tapes correspond to $1.1^{+0.8}_{-0.5}$ M events, which is consistent with the above result.

Combined Inclusive and Exclusive Yields

Applying the inclusive reconstruction method to the same signal events that were used for studying the exclusive case yields 30 events.[7] Assuming that all of those would have been selected by the exclusive reconstruction, the combined number from the inclusive and exclusive reconstruction is, rounded to the nearest 10 k, 520 ± 60 k.

3.4.8 Background Estimation

Wide Mass Window

No background events are left after all cuts are applied. In order to obtain enough statistics for a signal-to-background estimate, the B_d^0 mass cut is relaxed for the background events from ±0.2 GeV to ±2 GeV around the true B_d^0 mass. 6 events are found within the wide mass window.

In order to use this result for an estimate of the signal to background ratio within the small mass window, the ratio of background events in the wide mass window to the number in the small mass window needs to be known. The simplest assumption is that a ten times smaller mass window should have ten times less events. As we show below, this is not true due to kinematic reasons.

Background Composition

The following decays constitute the 6 background events that survive the wide mass cut:

1. $B_d^0 \to D^{*-}\pi^0\pi^+$ (−1.5 GeV)
2. $B_d^0 \to D^{*+}\rho^0\pi^-$ (−1.1 GeV)
3. $B_d^0 \to D^{*-}\rho^0\pi^+$ (−1.1 GeV)
4. $B_d^0 \to D^{*-}\omega\pi^+$ (−1.7 GeV)
5. $B_d^0 \to D^{*-}\omega\pi^+$ (−1.7 GeV)
6. $B_d^0 \to D^{*+}D_s^{*-}$ (−1.7 GeV)

---

[7] This number assumes that the efficiency of the cut on the D^0-vertex χ^2 is the same for the exclusive events, although this quantity was not available in the analysis. There are 35 events left without that cut.

where the number in parenthesis is the reconstructed $\mathrm{B^0_d}$ mass minus the true $\mathrm{B^0_d}$ mass. The reconstructed $\mathrm{B^0_d}$ mass is much lower than the true $\mathrm{B^0_d}$ mass in all cases. This is also true for the four background events that survive the other cuts but lie outside the wide mass window – they all have low reconstructed $\mathrm{B^0_d}$ masses.

In all six of the above events the slow pion from the D^* and the two charged particles from the $\overline{\mathrm{D}^0}$ decay are identified correctly. The charged pion from the first generation of the decay, present in all decays except for decay number 6, is taken as the fast pion. In decay number 6, a pion from the decay of the D_s^- from $D_s^{*-} \to D_s^- ...$ is taken as the fast pion.

So the topologies of the surviving background events are very similar to the signal events, except that there is some missing momentum in the $\mathrm{B^0_d}$ reconstruction. This missing momentum is the reason for the systematically low reconstructed $\mathrm{B^0_d}$ masses in the background events. Therefore, scaling the number of background events by a factor of 1/10 to estimate the background in the small mass window would be an over–estimation.

To confirm with higher statistics that the D^* reconstruction is generally successful, the selection is repeated with relaxed cuts on the $\mathrm{B^0_d}$ reconstruction while leaving the D^*–D^0 cuts unchanged. The changes relative to the cuts listed in table 3.4 are:

- Cut 1: reconstruction cut: $< 4\sigma$ instead of $< 2\sigma$
- Cut 7: $\mathrm{B^0_d}$–vertex $\chi^2 < 4$ instead of < 2
- Cut 9: Impact parameter significance of pions > 3 instead of > 4
- Cut 17: no cut on the $\mathrm{B^0_d}$-mass instead of $\pm 2\,\mathrm{GeV}$

43 background events pass those cuts. 8 of them have the slow pion mis-identified, the other 81% are of the same type as the 6 background events that pass the full set of cuts with the wide mass window. Of the decays with mis-identified pion

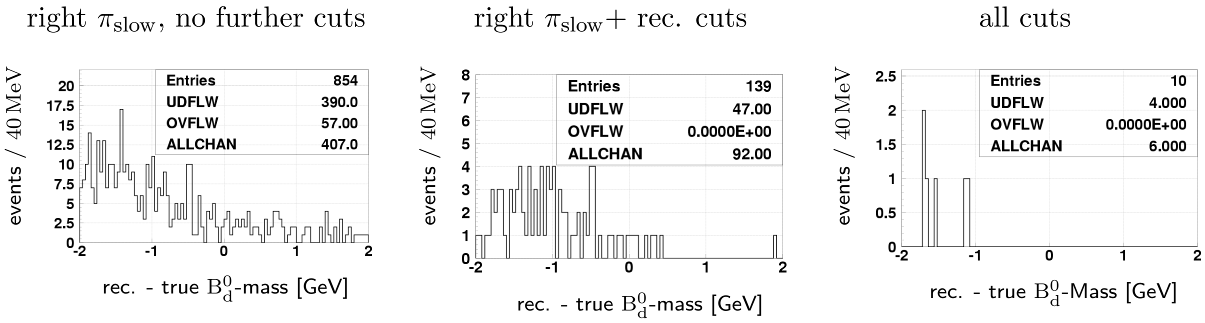
- 2 have a pion from the primary vertex taken as the slow pion.
- 3 are decays of the type $\mathrm{B^0_d} \to \pi +$ many other particles, where the pion from the $\mathrm{B^0_d}$ decay is taken as the slow pion.
- 2 are charged B^+ decays, where a pion from further down the decay chain is taken as the slow pion.
- 1 is a Λ_b decay.

The 8 events with mis-identified pions have a similar mass–distribution as the other events, all lying below the narrow $\pm 0.2\,\mathrm{GeV}$ mass window. Of these, 4 are below the $\pm 2\,\mathrm{GeV}$ mass window.

The Signal to Background Ratio

Taking into account the above observations, the strategy chosen to estimate the fraction of events inside the small mass window is to use Monte-Carlo truth information to select events with a correctly identified π_{slow}, and

Figure 3.11: First two plots: mass–spectra of background events where the slow pion is correctly identified, without and with (relaxed) reconstruction cuts applied. Third plot: mass–spectrum with all cuts applied. For each plot the ratio of events in the wide to the narrow mass window is indicated.



ratio $407 : 25 = 16$ ratio $92 : 6 = 15$ ratio $6 : 0 = \infty$

to loosen all other selection cuts. Without cuts, the ratio of the number of background events in the wide mass window to the number in the small mass window is 16. To make the data selection more realistic, cuts on reconstruction parameters that are important in the mass reconstruction are applied:

- Cut 1: correction $< 4\sigma$ instead of $< 2\sigma$;
- Cut 7: $\mathrm{B_d^0}$–vertex $\chi^2 < 4$ instead of < 2;
- Cut 11: as before: p_t of $\mathrm{B_d^0} < 60\,\mathrm{GeV}$.

Also, Monte-Carlo truth information is used to ensure the right particle type for the fast pion. The trigger is not applied. This results in 92 background events in the wide mass window and 6 in the small one, giving a ratio of 15^{+10}_{-6}. Figure 3.11 shows the mass distributions for events with correctly identified π_{slow} before and after these cuts are applied, and, in the third plot, the six events that survive the standard set of cuts with the wide mass window.

Taking the conservative estimate of a ratio of 15 gives an estimate for the signal to background ratio of $4.4^{+4.4}_{-1.1}$, where the error is calculated as a 68.3% central interval confidence limit using numerical methods. The signal to background ratio is better than 2 at the 90% confidence level. For the confidence limits, only the statistical error on the 6 background events left in the wide mass window and the statistical error in the calculation of the fraction of background events inside the narrow mass window are taken into account.

3.4.9 Mass and Time Resolution

Figures 3.12 and 3.13 show the reconstructed mass and decay time distributions. The lack of events for small decay times is due to the cut on the $\mathrm{B_d^0}$ decay length. Fig-

Figure 3.12: Mass Resolution, with relaxed cuts, keeping those on the reconstruction parameters

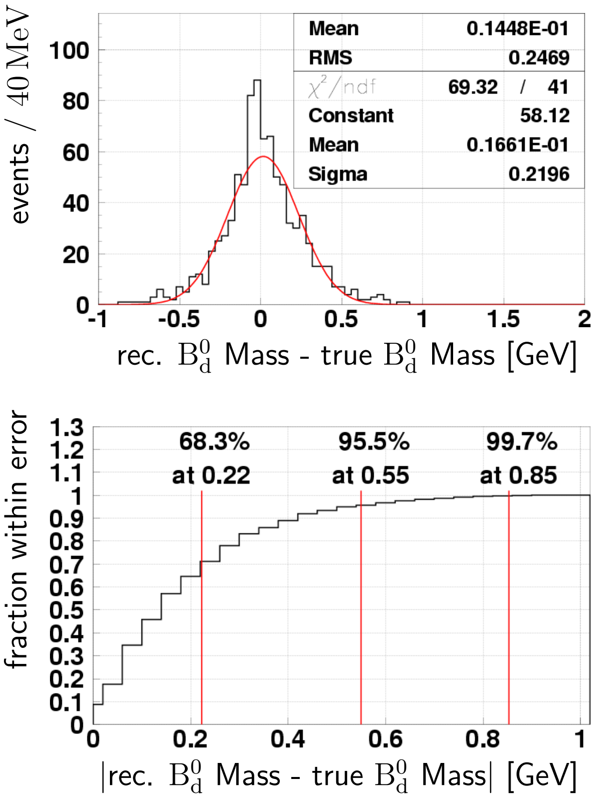
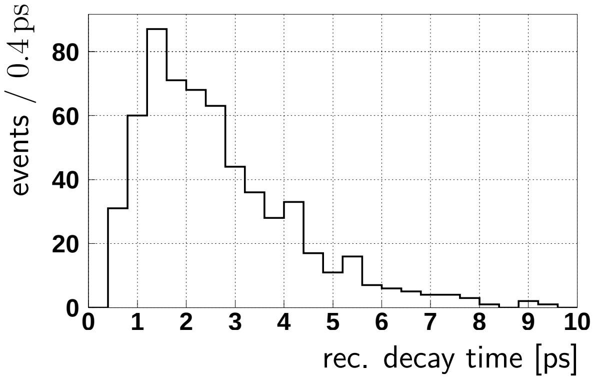


Figure 3.13: Reconstructed decay times with relaxed cuts, keeping those on the reconstruction parameters and B^0_d mass. There is no event below 0.4 ps.



ure 3.16 shows the differences between reconstructed and true values for the decay time. In order to obtain large enough a sample to estimate the mass and the decay time resolution from these plots, only the subset of cuts marked with an asterisk in table 3.4 is applied, including the reconstruction cut. The usual algorithm is used to decide between two combinations of tracks in an event that passes all cuts. The trigger is not applied. For the vertex, momentum and time resolution estimates, the cut on the B^0_d mass is also applied.

A Gaussian fit to the mass distribution returns a width of 220 MeV, as shown in figure 3.12. Because the mass-distribution is not strictly Gaussian, the fraction of events within a certain error on the B^0_d mass is also shown as a function of that error. 68.3 % of events are within in ±220 MeV and 95.5 % are within ±550 MeV.

The time resolution depends directly on the vertex and the momentum resolution. The vertex resolutions, shown in figure 3.14, are consistent with those found earlier, before applying the reconstruction cut and the mass cut. Now the x and y resolution is reasonably well described by a single Gaussian. The momentum resolution is fitted with two Gaussians, the narrower of which has a width of 2 GeV. 68.3 % of events have reconstructed momenta within ±5.7 GeV of the true value.

A Gaussian fit to the time-resolution, shown in figure 3.16, returns a σ of 160 fs. The integrated fraction of events within a given error-range is plotted underneath, showing that 68.3% of events are within

Figure 3.14: Vertex Resolution, with relaxed cuts, keeping those on the reconstruction parameters and B_d^0 mass cut. For the resolution in z, the parameters (P1,P4), (P2,P5), (P3,P6) are the amplitudes, means and widths respectively of the two Gaussians fitted to the distributions.

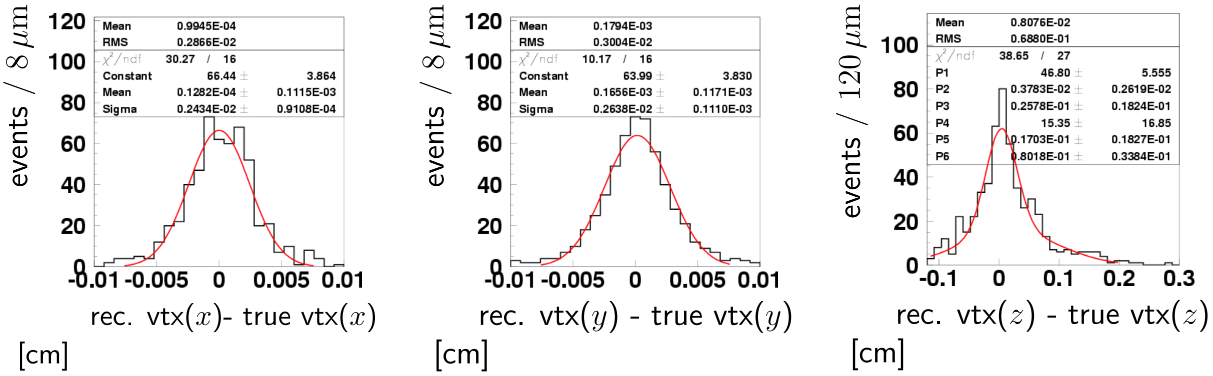


Table 3.6: $B_d^0 \to D^{*-}\pi^+$ inclusive: result summary, and combined inclusive + exclusive event yield.

Signal yield	(460 ± 60) k per year
Combined excl.+incl. yield	(520 ± 60) k per year
Signal/Backg	$4.4^{+4.4}_{-1.1}$
	> 2 at 90% confidence
mass-resol$^{\underline{\text{n}}}$	220 MeV
time-resol$^{\underline{\text{n}}}$	170 fs

±171 fs. Both values are much smaller than the B_d^0-$\overline{B_d^0}$ oscillation period. The effect of the finite time resolution on the γ measurement is therefore negligible.

3.4.10 Result Summary

Using the inclusive reconstruction method, the total event yield can be increased by a factor of ~ 7 compared to the exclusive method, to more than half a million events per year, with good background rejection.

The full B_d^0 4–momentum is reconstructed from the fast and the slow pion. A mass resolution of 220 MeV and time resolution of 170 fs, or 0.013 B_d^0-$\overline{B_d^0}$ oscillation periods, can be achieved, despite the limited information used.

A summary of the main results from this section is given in table 3.6.

3.5 Estimating the γ Sensitivity

In this section, the result of the reconstruction studies are used to estimate the precision with which the CKM angle γ can be extracted from the decay rate asymmetries:

$$A_\eta(\tau) = \frac{\Gamma(B_d^0 \to D^{*-}\pi^+) - \Gamma(\overline{B_d^0} \to D^{*-}\pi^+)}{\Gamma(B_d^0 \to D^{*-}\pi^+) + \Gamma(\overline{B_d^0} \to D^{*-}\pi^+)} \tag{3.17}$$

Figure 3.15: Momentum Resolution, with relaxed cuts, keeping those on the reconstruction parameters and B^0_d mass. The parameters (P1,P4), (P2,P5), (P3,P6) are the amplitudes, means and widths respectively of the two Gaussians fitted to the distributions.

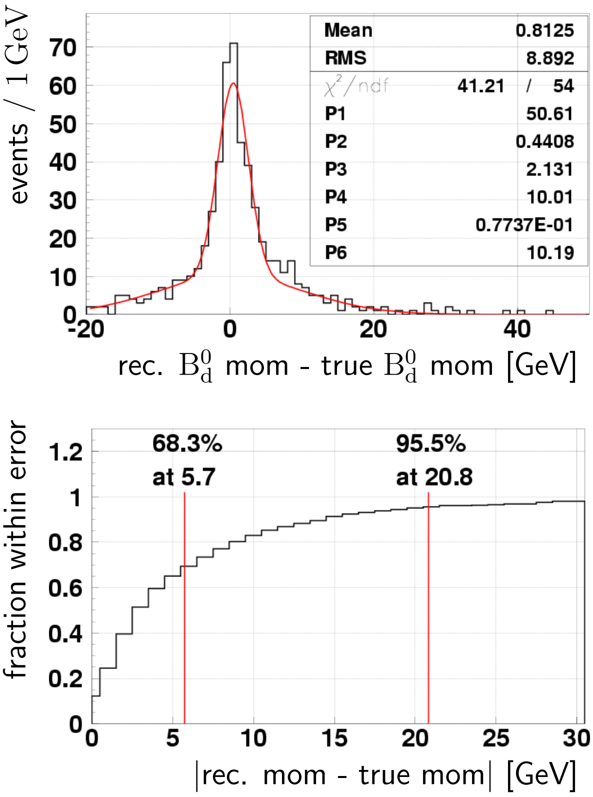
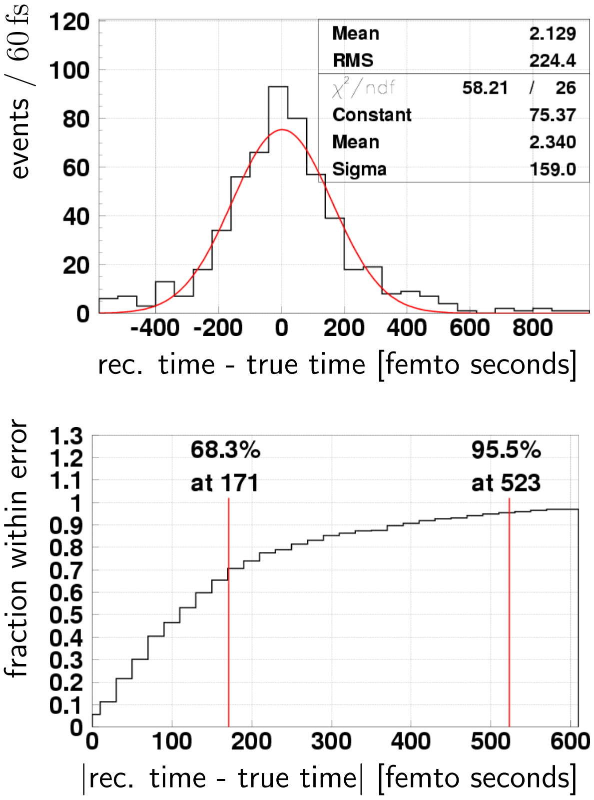


Figure 3.16: Time Resolution, with relaxed cuts. Only the cuts on the reconstruction parameters and B^0_d mass are applied.



and

$$A_{\overline{\eta}}(\tau) = \frac{\Gamma(\overline{B^0_d} \rightarrow D^{*+}\pi^-) - \Gamma(B^0_d \rightarrow D^{*+}\pi^-)}{\Gamma(\overline{B^0_d} \rightarrow D^{*+}\pi^-) + \Gamma(B^0_d \rightarrow D^{*+}\pi^-)} \tag{3.18}$$

as introduced in chapter 1. Note that the asymmetries are defined such that differences in the detection efficiencies of the two final states cancel.

Using a stand–alone program, events are generated according to the measured decay rates of the four decay channels under

The four undiluted decay rates are given apart from a common normalisation factor. The parameters η and $\overline{\eta}$ for the decay $\mathrm{B_d^0} \to \mathrm{D}^*\pi$ are given by:

$$\eta = |\eta|\, e^{i(\Delta_{\mathrm{qcd}}+(\phi_{\mathrm{mix}}-\gamma))} \quad \overline{\eta} = |\eta|\, e^{i(\Delta_{\mathrm{qcd}}-(\phi_{\mathrm{mix}}-\gamma))}.$$

$$\begin{aligned}
&1)\,\Gamma(\mathrm{B_d^0} \to \mathrm{D^{*-}}\pi^+) = R_{\mathrm{D^{*-}}}(\tau) = e^{-\Gamma\tau}\Big\{ \Big(1+|\eta|^2\Big) + \Big(1-|\eta|^2\Big)\cos(\Delta m\tau) - 2\mathsf{Im}(\eta)\sin(\Delta m\tau) \Big\}\\
&2)\,\Gamma(\overline{\mathrm{B_d^0}} \to \mathrm{D^{*-}}\pi^+) = \overline{R}_{\mathrm{D^{*-}}}(\tau) = e^{-\Gamma\tau}\Big\{ \Big(1+|\eta|^2\Big) - \Big(1-|\eta|^2\Big)\cos(\Delta m\tau) + 2\mathsf{Im}(\eta)\sin(\Delta m\tau) \Big\}\\
&3)\,\Gamma(\overline{\mathrm{B_d^0}} \to \mathrm{D^{*+}}\pi^-) = \overline{R}_{\mathrm{D^{*+}}}(\tau) = e^{-\Gamma\tau}\Big\{ \Big(1+|\overline{\eta}|^2\Big) + \Big(1-|\overline{\eta}|^2\Big)\cos(\Delta m\tau) - 2\mathsf{Im}(\overline{\eta})\sin(\Delta m\tau) \Big\}\\
&4)\,\Gamma(\mathrm{B_d^0} \to \mathrm{D^{*+}}\pi^-) = R_{\mathrm{D^{*+}}}(\tau) = e^{-\Gamma\tau}\Big\{ \Big(1+|\overline{\eta}|^2\Big) - \Big(1-|\overline{\eta}|^2\Big)\cos(\Delta m\tau) + 2\mathsf{Im}(\overline{\eta})\sin(\Delta m\tau) \Big\}
\end{aligned} \tag{3.16}$$

study. From log-likelihood fits to the asymmetries, the error on γ is estimated.

The first step in this analysis is to derive the measured decay rates and asymmetries, including detector effects, from the theoretical ones given in sections 1.3.4 and 1.3.6. For ease of reference, and in order to define the notation used in the following sections, the undiluted decay rates are given again in equation 3.16 (page 95).

The asymmetries including the detector effects will be given in terms of $A(\tau)$, which stands for both $A_\eta(\tau)$ and $A_{\overline{\eta}}(\tau)$, before detector effects are taken into account.

3.5.1 Decay Rates, Including Detector Effects

Detector-effects are taken into account by assuming a Gaussian-distributed time resolution of 170 fs, a uniform mistag probability of 30 % and an acceptance function:

$$P_A(\tau) = \max\left(0, \frac{(a\tau)^3}{1+(a\tau)^3} - b\right) \tag{3.19}$$
$$\text{with } a = 0.96\,\mathrm{ps}^{-1}, \quad b = 0.09$$

giving the probability that an event with decay-eigentime τ is recorded. The acceptance function is taken from [TP98]. It takes into account that the trigger, and also the final event selection, rejects events with small decay–lengths. A background to signal ratio of $B/S = 0.2$ is assumed.

Time Resolution

With a true decay rate

$$R_t(\tau) = e^{-\Gamma\tau}\,(a + b\cos(\omega\tau) + c\sin(\omega\tau)) \tag{3.20}$$

the measured decay rate is, taking into account the finite time–resolution and time–

dependent acceptance:

$$R_{A\,\sigma_\tau}(\tau_0) = \int_0^\infty P_A(\tau_0,\tau)\, R_t(\tau) \cdot g_{\tau_0}(\tau_0 - \tau)\, \mathrm{d}\tau. \tag{3.21}$$

Here, τ is the decay eigentime of the B^0_d, and τ_0 is the reconstructed decay eigentime; $P_A(\tau_0,\tau)$ is the acceptance function and $g_{\tau_0}(\tau_0 - \tau)$ the resolution function. In general both might be quite complicated, and numerical methods will be necessary to perform the integral. Here, in order to be able to do the integral analytically, we will assume that the time resolution is described well by a Gaussian, and that the acceptance function is a function of the measured decay time, τ_0, only. Then the expression for the measured decay rate becomes:

$$\begin{aligned} R_{A\,\sigma_\tau}(\tau_0) = P_A(\tau_0) \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma_\tau} e^{-\frac{(\tau_0-\tau)^2}{2\sigma_\tau^2}} \\ e^{-\Gamma\tau}\left(a + b\cos(\omega\tau) + c\sin(\omega\tau)\right) \mathrm{d}\tau. \end{aligned} \tag{3.22}$$

All three parts of the above sum can be solved simultaneously by calculating:

$$F(\tau_0) = \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma_\tau} e^{-\frac{(\tau_0-\tau)^2}{2\sigma_\tau^2}} e^{-\Gamma\tau} e^{i\omega\tau} \mathrm{d}\tau. \tag{3.23}$$

Taking the part independent of τ outside the integral gives:

$$\begin{aligned} F(\tau_0) &= \frac{1}{\sqrt{2\pi}\sigma_\tau} e^{-\frac{1}{2}\left(\frac{\tau_0}{\sigma_\tau}\right)^2} \\ &\quad \int_0^\infty e^{-\frac{1}{2\sigma_\tau^2}\left(\tau^2 - 2\left(\tau_0 + (i\omega-\Gamma)\sigma_\tau^2\right)\tau\right)} \mathrm{d}\tau. \end{aligned} \tag{3.24}$$

Defining

$$z \equiv \tau_0 + (i\omega - \Gamma)\,\sigma_\tau^2 \tag{3.25}$$

and completing the square in the exponent, this becomes:

$$\begin{aligned} F(\tau_0) &= \frac{1}{\sqrt{2\pi}\sigma_\tau} e^{-\frac{1}{2}\left(\frac{\tau_0}{\sigma_\tau}\right)^2} e^{+\frac{1}{2}\left(\frac{z}{\sigma_\tau}\right)^2} \\ &\quad \cdot \int_0^\infty e^{-\frac{(\tau-z)^2}{2\sigma_\tau^2}} \mathrm{d}\tau. \end{aligned} \tag{3.26}$$

In practice, only events with long decay times pass the trigger, and therefore

$$\mathsf{Re}(z) = \tau_0 - \Gamma\sigma_\tau^2 > \text{a few } \sigma_\tau$$

for the relevant values of τ_0. Then the integral from 0 to ∞ can be replaced with an integral from $-\infty$ to ∞, and the solution is:

$$\begin{aligned} F(\tau_0) &= \frac{1}{\sqrt{2\pi}\sigma_\tau} e^{-\frac{1}{2}\left(\frac{\tau_0}{\sigma_\tau}\right)^2} e^{+\frac{1}{2}\left(\frac{z}{\sigma_\tau}\right)^2} \\ &\quad \cdot \int_{-\infty}^\infty e^{-\frac{(\tau-z)^2}{2\sigma_\tau^2}} \mathrm{d}\tau \\ &= e^{-\frac{1}{2}\left(\frac{\tau_0}{\sigma_\tau}\right)^2} e^{+\frac{1}{2}\left(\frac{z}{\sigma_\tau}\right)^2} \\ &= e^{\frac{1}{2\sigma_\tau^2}\left((i\omega-\Gamma)^2\sigma_\tau^4 + 2\tau_0\left((i\omega-\Gamma)\sigma_\tau^2\right)\right)}. \end{aligned} \tag{3.27}$$

Re-ordering gives the final result:

$$\boxed{F(\tau_0) = e^{-\frac{1}{2}\sigma_\tau^2\cdot\left(\Gamma^2+\omega^2\right)} e^{-\Gamma\left(\tau_0-\Gamma\sigma_\tau^2\right)} e^{i\omega\left(\tau_0-\Gamma\sigma_\tau^2\right)} \qquad (3.28)}$$

So the effect of the finite time-resolution on the function

$$R\left(\tau\right) = e^{-\Gamma\tau}\left(a + b\cos(\omega\tau) + c\sin(\omega\tau)\right)$$

can be described by simultaneously scaling the amplitudes a, b, and c, and shifting the parameter τ according to:

$$\begin{aligned} a &\longrightarrow ae^{-\frac{1}{2}\sigma_\tau^2\cdot\Gamma^2} \\ b &\longrightarrow be^{-\frac{1}{2}\sigma_\tau^2\cdot\left(\Gamma^2+\omega^2\right)} \\ c &\longrightarrow ce^{-\frac{1}{2}\sigma_\tau^2\cdot\left(\Gamma^2+\omega^2\right)} \\ \tau &\longrightarrow \tau - \Gamma\sigma_\tau^2. \end{aligned} \qquad (3.29)$$

This transforms the asymmetry to:

$$A_{\sigma_\tau}(\tau) = e^{-\frac{1}{2}\sigma_\tau^2\cdot(\Delta m)^2} A(\tau - \Gamma\sigma_\tau^2) \qquad (3.30)$$

The measured decay rates, taking into account the finite time resolution, but not the acceptance function, are given in equation 3.31 (page 98).

Background

In this study it is assumed that the background is independent of the decay considered; this is to say that at any given time, the number of background events interpreted as $\mathrm{B}_\mathrm{d}^0 \to \mathrm{D}^{*-}\pi^+$ decays is the same as the number interpreted as $\overline{\mathrm{B}_\mathrm{d}^0} \to \mathrm{D}^{*-}\pi^+$ decays. As will be shown in section 3.5.6, this results in a conservative estimate on the statistical precision in γ. For a background to signal ratio of B/S, the measured decay rate for $\mathrm{B}_\mathrm{d}^0 \to \mathrm{D}^{*-}\pi^+$, $R_{\frac{B}{S}\,\mathrm{D}^{*-}}(\tau)$, becomes, in terms of the decay rates without background:

$$R_{\frac{B}{S}\,\mathrm{D}^{*-}}(\tau) = R_{\mathrm{D}^{*-}}(\tau) + \frac{B}{S}\cdot\frac{1}{2}\left(R_{\mathrm{D}^{*-}}(\tau) + \overline{R}_{\mathrm{D}^{*-}}(\tau)\right) \qquad (3.32)$$

and similarly for the other decay rates. The measured asymmetry, $A_{\frac{B}{S}}(\tau)$, is given by:

$$A_{\frac{B}{S}}(\tau) = \frac{1}{1+B/S}A(\tau), \qquad (3.33)$$

where $A(\tau)$ is the asymmetry without background.

Mistag

Including a mistag fraction of ω_tag, the measured decay rate for $\mathrm{B}_\mathrm{d}^0 \to \mathrm{D}^{*-}\pi^+$, $R_{\omega_\mathrm{tag}\,\mathrm{D}^{*-}}(\tau)$, becomes, in terms of the decay rate without mistag:

$$R_{\omega_\mathrm{tag}\,\mathrm{D}^{*-}}(\tau) = \left(1-\omega_\mathrm{tag}\right) R_{\mathrm{D}^{*-}}(\tau) + \omega_\mathrm{tag}\overline{R}_{\mathrm{D}^{*-}}(\tau) \qquad (3.34)$$

and similarly for the other decay rates. The measured asymmetry, $A_{\omega_\mathrm{tag}}(\tau)$, is given by

$$A_{\omega_\mathrm{tag}}(\tau) = \left(1-2\omega_\mathrm{tag}\right) A(\tau) \qquad (3.35)$$

where $A(\tau)$ is the asymmetry for perfect tagging.

Measured decay rates with finite time-resolution, and no other detector effects included. The decay rates are given apart from a common normalisation factor.

$\underline{\mathrm{B_d^0} \to \mathrm{D^{*-}}\pi^+}$:

$$\begin{aligned} R_{\sigma_\tau D^{*-}}(\tau) &= e^{-\Gamma\left(\tau-\Gamma\sigma_\tau^2\right)}\Big\{e^{-\frac{1}{2}\sigma_\tau^2\cdot\Gamma^2}\left(1+|\eta|^2\right) + \\ &\quad e^{-\frac{1}{2}\sigma_\tau^2\cdot\left(\Gamma^2+\omega^2\right)}\left[\left(1-|\eta|^2\right)\cos\Big(\Delta m\left(\tau-\Gamma\sigma_\tau^2\right)\Big) - 2\mathsf{Im}(\eta)\sin\Big(\Delta m\left(\tau-\Gamma\sigma_\tau^2\right)\Big)\right]\Big\} \end{aligned}$$

$\underline{\overline{\mathrm{B_d^0}} \to \mathrm{D^{*-}}\pi^+}$:

$$\begin{aligned} \overline{R}_{\sigma_\tau D^{*-}}(\tau) &= e^{-\Gamma\left(\tau-\Gamma\sigma_\tau^2\right)}\Big\{e^{-\frac{1}{2}\sigma_\tau^2\cdot\Gamma^2}\left(1+|\eta|^2\right) - \\ &\quad e^{-\frac{1}{2}\sigma_\tau^2\cdot\left(\Gamma^2+\omega^2\right)}\left[\left(1-|\eta|^2\right)\cos\Big(\Delta m\left(\tau-\Gamma\sigma_\tau^2\right)\Big) - 2\mathsf{Im}(\eta)\sin\Big(\Delta m\left(\tau-\Gamma\sigma_\tau^2\right)\Big)\right]\Big\} \end{aligned}$$

$\underline{\overline{\mathrm{B_d^0}} \to \mathrm{D^{*+}}\pi^-}$:

$$\begin{aligned} \overline{R}_{\sigma_\tau D^{*+}}(\tau) &= e^{-\Gamma\left(\tau-\Gamma\sigma_\tau^2\right)}\Big\{e^{-\frac{1}{2}\sigma_\tau^2\cdot\Gamma^2}\left(1+|\eta|^2\right) + \\ &\quad e^{-\frac{1}{2}\sigma_\tau^2\cdot\left(\Gamma^2+\omega^2\right)}\left[\left(1-|\eta|^2\right)\cos\Big(\Delta m\left(\tau-\Gamma\sigma_\tau^2\right)\Big) - 2\mathsf{Im}(\overline{\eta})\sin\Big(\Delta m\left(\tau-\Gamma\sigma_\tau^2\right)\Big)\right]\Big\} \end{aligned}$$

$\underline{\mathrm{B_d^0} \to \mathrm{D^{*+}}\pi^-}$:

$$\begin{aligned} R_{\sigma_\tau D^{*+}}(\tau) &= e^{-\Gamma\left(\tau-\Gamma\sigma_\tau^2\right)}\Big\{e^{-\frac{1}{2}\sigma_\tau^2\cdot\Gamma^2}\left(1+|\eta|^2\right) - \\ &\quad e^{-\frac{1}{2}\sigma_\tau^2\cdot\left(\Gamma^2+\omega^2\right)}\left[\left(1-|\eta|^2\right)\cos\Big(\Delta m\left(\tau-\Gamma\sigma_\tau^2\right)\Big) - 2\mathsf{Im}(\overline{\eta})\sin\Big(\Delta m\left(\tau-\Gamma\sigma_\tau^2\right)\Big)\right]\Big\} \end{aligned} \tag{3.31}$$

The Full Expression

The expressions for the measured decay rates as function of the decay eigentime, including all detector effects considered, are given in equations 3.36 (page 100) . The measured asymmetry, $A_m(\tau)$, in terms of the undiluted asummetry $A(\tau)$ is given by:

$$A_m(\tau) = e^{-\frac{1}{2}\sigma_\tau^2\cdot\omega_{\mathrm{tag}}^2}\frac{1-2\omega_{\mathrm{tag}}}{1+B/S}A(\tau-\Gamma\sigma_\tau^2). \tag{3.37}$$

Explicitly written out, this is:

$$\begin{aligned} A_{m,\eta}(\tau) &= D\frac{\left(1-|\eta|^2\right)\cos(\Delta m\,(\tau-\Gamma\sigma_\tau^2))}{2\left(1+|\eta|^2\right)} \\ &+ D\frac{2\mathsf{Im}(\eta)\sin(\Delta m\,(\tau-\Gamma\sigma_\tau^2))}{2\left(1+|\eta|^2\right)} \end{aligned} \tag{3.38}$$

and

$$\begin{aligned} A_{m,\overline{\eta}}(\tau) &= D\frac{\left(1-|\eta|^2\right)\cos(\Delta m\,(\tau-\Gamma\sigma_\tau^2))}{2\left(1+|\eta|^2\right)} \\ &+ D\frac{2\mathsf{Im}(\overline{\eta})\sin(\Delta m\,(\tau-\Gamma\sigma_\tau^2))}{2\left(1+|\eta|^2\right)}, \end{aligned} \tag{3.39}$$

with

$$D \equiv e^{-\frac{1}{2}\sigma_\tau^2\cdot\omega_{\mathrm{tag}}^2}\frac{1-2\omega_{\mathrm{tag}}}{1+B/S}. \tag{3.40}$$

3.5.2 Generating Events

Two methods are used to generate events, one where a series of events is calculated analytically, and one that uses Monte-Carlo techniques.

The preferred method is the analytical approach. Its main advantage is its speed. The analytical method provides an event sample for which the fit always returns the correct parameter values, that were used used in the generation.

The Monte Carlo simulation is used to cross-check the results in two ways. A large number of data samples is generated and fitted. Both the RMS spread in the fit result and the mean of the error estimates for each fit are compared with the results obtained using the default method.

Calculating Events Analytically

In order to generate an event sample that is suitable for an unbinned log–likelihood fit, the measured decay rates given in equation 3.36 need to be translated into four series of decay times, one series for each decay.

For the purpose this discussion, we define the normalised, measured decay rates:

$$r_i(\tau) = \frac{R_{m\,i}(\tau)}{\int R_{m\,i}(\tau)\mathrm{d}\tau} \tag{3.41}$$

where i stands for one of the four channels under study, and $R_{m\,i}(\tau)$ is the measured decay rate for that channel, according to the definitions:

$$R_{m\,1}(\tau) \equiv R_{m\,\mathrm{D}^{*-}}(\tau)$$

Decay rates including all detector effects considered, apart from a common normalisation factor:

$$
\begin{aligned}
R_{m\,\mathrm{D}^{*-}}(\tau) &= P_A(\tau)\cdot\left((1-\omega_{\mathrm{tag}})\,R_{\sigma_\tau \mathrm{D}^{*-}}(\tau) + \omega_{\mathrm{tag}}\overline{R}_{\sigma_\tau \mathrm{D}^{*-}}(\tau) + \frac{B}{S}\cdot\frac{1}{2}\left(R_{\sigma_\tau \mathrm{D}^{*-}}(\tau) + \overline{R}_{\sigma_\tau \mathrm{D}^{*-}}(\tau)\right)\right)\\
\overline{R}_{m\,\mathrm{D}^{*-}}(\tau) &= P_A(\tau)\cdot\left((1-\omega_{\mathrm{tag}})\,\overline{R}_{\sigma_\tau \mathrm{D}^{*-}}(\tau) + \omega_{\mathrm{tag}}R_{\sigma_\tau \mathrm{D}^{*-}}(\tau) + \frac{B}{S}\cdot\frac{1}{2}\left(R_{\sigma_\tau \mathrm{D}^{*-}}(\tau) + \overline{R}_{\sigma_\tau \mathrm{D}^{*-}}(\tau)\right)\right)\\
\overline{R}_{m\,\mathrm{D}^{*+}}(\tau) &= P_A(\tau)\cdot\left((1-\omega_{\mathrm{tag}})\,\overline{R}_{\sigma_\tau \mathrm{D}^{*+}}(\tau) + \omega_{\mathrm{tag}}R_{\sigma_\tau \mathrm{D}^{*+}}(\tau) + \frac{B}{S}\cdot\frac{1}{2}\left(R_{\sigma_\tau \mathrm{D}^{*+}}(\tau) + \overline{R}_{\sigma_\tau \mathrm{D}^{*+}}(\tau)\right)\right)\\
R_{m\,\mathrm{D}^{*+}}(\tau) &= P_A(\tau)\cdot\left((1-\omega_{\mathrm{tag}})\,R_{\sigma_\tau \mathrm{D}^{*+}}(\tau) + \omega_{\mathrm{tag}}\overline{R}_{\sigma_\tau \mathrm{D}^{*+}}(\tau) + \frac{B}{S}\cdot\frac{1}{2}\left(R_{\sigma_\tau \mathrm{D}^{*+}}(\tau) + \overline{R}_{\sigma_\tau \mathrm{D}^{*+}}(\tau)\right)\right)
\end{aligned}
\tag{3.36}
$$

$$
\begin{aligned}
R_{m\,2}(\tau) &\equiv \overline{R}_{m\,\mathrm{D}^{*-}}(\tau)\\
R_{m\,3}(\tau) &\equiv \overline{R}_{m\,\mathrm{D}^{*+}}(\tau)\\
R_{m\,4}(\tau) &\equiv R_{m\,\mathrm{D}^{*+}}(\tau)
\end{aligned}
$$

Principle Given $r_i(\tau)$ and the total number of events to be generated for channel i, N_i, the number ΔN of events between τ and $\tau + \Delta\tau$ is:

$$\Delta N \approx N_i\, r_i(\tau)\Delta\tau. \tag{3.42}$$

Inverting this gives

$$\Delta\tau \approx \frac{\Delta N}{N_i\, r_i(\tau)}. \tag{3.43}$$

The smallest number of events we can generate is 1. Therefore we choose $\Delta N = 1$:

$$\Delta\tau(\tau) \approx \frac{1}{N_i\, r_i(\tau)} \tag{3.44}$$

Thus in principle a suitable event sample of size N_i, distributed according to decay rate $R_{m\,i}(\tau)$, may be generated by starting at $\tau = 0$, and then increasing τ $N_i - 1$ times in steps of $\frac{1}{N_i r_i(\tau)}$, recording one event for each decay time τ.

Starting the iteration at $\tau = 0$ is however not possible for the measured decay rates discussed here, since $r_i(0) = 0$ due to the acceptance function. Therefore, the procedure is started at the value of τ where $r_i(\tau)$ is maximal, and then goes into both the positive and the negative direction direction, until N_i events are generated.

Because $r(\tau) \to 0$ for large τ, the maximum decay time for which events are generated needs to be limited to ensure that $\Delta\tau \approx \frac{\Delta N}{N r(\tau)}$ remains a good approximation. In this analysis, events are generated for a maximum decay time of one B^0_d–$\overline{\mathrm{B}^0_\mathrm{d}}$ oscillation period, which corresponds 8.6 B^0_d lifetimes. Theoretically, restricting the maximum decay time could be problematic, since the acceptance function only affects decays with small decay times. However, this effect is negligible, since the fraction of events beyond 8.6 B^0_d lifetimes is tiny. The agreement of the fit-results using this method with results from the Monte-Carlo study corroborates this.

To increase accuracy, the step size $\Delta\tau$ is calculated about half–way inbetween the generated decay times: For each value of τ, first $\Delta\tau' = \Delta\tau(\tau)$ is calculated, and then τ is increased by $\Delta\tau(\tau + \Delta\tau'/2)$, or decreased by $\Delta\tau(\tau - \Delta\tau'/2)$.

Relative Sample Sizes: In order to generate the correct relative numbers of events for each of the four decay rates, the values of N_i are calculated before the procedure starts, by integrating the expressions for the decay rates, and scaling the result according to the total number of events desired, N_{total}:

$$N_i = N_{\text{total}} \frac{\int\limits_0^{\tau_{\max}} R_{m\,i}(\tau) d\tau}{\int\limits_0^{\tau_{\max}} \sum\limits_{j=1}^{4} R_{mj} \tau d\tau}, \tag{3.45}$$

where $\tau_{\max}$ is the maximum decay time for which events are generated.

Monte-Carlo Generated Events

Standard techniques are used to generate the sample of Monte Carlo events.

Differences Between the Methods

The decay rate distributions the Monte Carlo simulation is based on, and those used for the analytical method, are not identical. While the analytical method uses the decay rates derived in section 3.5.1, the Monte Carlo first generates the decay rates without any detector effects, and then implements the finite time resolution using a Gaussian random number generator. This makes the Monte Carlos simulation independent of the approximation made in section 3.5.1, equation 3.27, where an integration from 0 to ∞ was replaced with an integration from $-\infty$ to ∞. In the section 3.5.1 it was argued that the approximation is very good, as long as the decay times are not too short. But while a fit to the analytical events must be successful because the events are generated with the same function with which they are fitted, a successful fit to the Monte Carlo generated samples validates this argument.

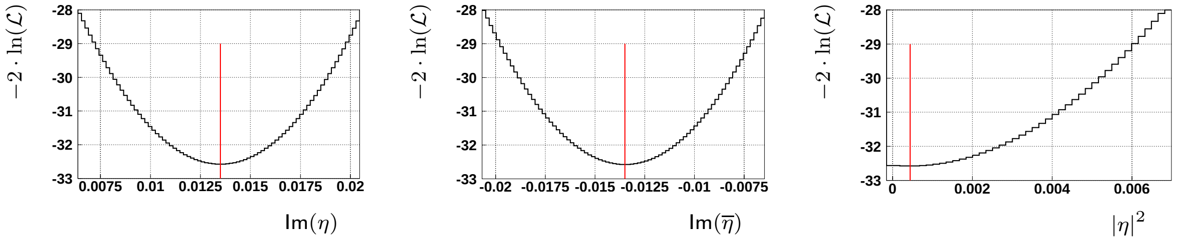
3.5.3 Fitting Events

To estimate the error on γ, the decay time asymmetries $A_{m\,\eta}$ and $A_{m\,\overline{\eta}}$ are fitted to the generated event samples using an unbinned log–likelihood method. The function minimisation and error estimation is performed using the software package MINUIT [Jam94].

The Log–Likelihood Function

The log–likelihood function that is maximised is:

Figure 3.17: Log-likelihood curves for one year of data taking (520 k events), using the analytical method for generating events. For each parameter, the function is shown for values $\pm 2\sigma$ around the fitted value, with the other two parameters fixed at their fit-values. $|\eta|^2$ is forced to be positive. The events were calculated for $(\phi_{\text{mix}} - \gamma) = 40°$ and $\Delta_{\text{qcd}} = 0°$. With $|\eta| = 0.021$, this corresponds to $\mathsf{Im}(\eta) = -\mathsf{Im}(\overline{\eta}) = 0.0134$, and $|\eta|^2 = 4.41 \cdot 10^{-4}$. The red lines indicate the input values.



$$
\begin{aligned}
\ln(\mathcal{L}) \quad = \quad & \sum_{\text{all } B_d^0 \to D^{*-}\pi^+ \text{ evts}} \ln\left(p_\eta(\tau_i)\right) \\
& + \sum_{\text{all } \overline{B_d^0} \to D^{*-}\pi^+ \text{ evts}} \ln\left(1 - p_\eta(\tau_i)\right) \\
& + \sum_{\text{all } \overline{B_d^0} \to D^{*+}\pi^- \text{ evts}} \ln\left(p_{\overline{\eta}}(\tau_i)\right) \\
& + \sum_{\text{all } B_d^0 \to D^{*+}\pi^- \text{ evts}} \ln\left(1 - p_{\overline{\eta}}(\tau_i)\right)
\end{aligned}
\tag{3.46}
$$

where τ_i are the measured decay times. $p_\eta(\tau_i)$ is the probability to find a $B_d^0 \to D^{*-}\pi^+$ event at time τ_i, rather than a $\overline{B_d^0} \to D^{*-}\pi^+$ event, and $p_{\overline{\eta}}(\tau_i)$ is the probability to find a $\overline{B_d^0} \to D^{*+}\pi^-$ event rather than a $B_d^0 \to D^{*+}\pi^-$ event. In terms of the measured asymmetries, p_η and $p_{\overline{\eta}}$ are given by:

$$
p_{\eta,\overline{\eta}}(\tau) = \frac{A_{m,\eta,\overline{\eta}}(\tau) + 1}{2}. \tag{3.47}
$$

The formulae derived in section 3.5.1 for the decay rates and asymmetries in the presence of a finite time resolution are only valid for values of $\tau - \Gamma\sigma_\tau^2$ that are not too small. Therefore, in principle, decay times with $\tau - \Gamma\sigma_\tau^2 < 3\sigma_\tau$ are ignored in the fit. The acceptance function used for generating the events however, removes all events with $\tau < 3.36\sigma_\tau$, so that this restriction does not have any practical consequences.

In figure 3.17 the log-likelihood function is shown for the three fit parameters $\mathsf{Im}(\eta)$, $\mathsf{Im}(\overline{\eta})$ and $|\eta|^2$. The plots are obtained from a fit to the analytically generated data, by varying one parameter and leaving the others fixed at their fit results.

From the third plot it can be seen that the error on $|\eta|^2$ is much larger than $|\eta|^2$ itself. It is impossible, therefore, to fit it with sufficient precision to extract $\sin(\Delta_{\text{qcd}} + (\phi_{\text{mix}} - \gamma)) = \frac{\mathsf{Im}(\eta)}{|\eta|}$ and $\sin(\Delta_{\text{qcd}} - (\phi_{\text{mix}} - \gamma)) = \frac{\mathsf{Im}(\overline{\eta})}{|\eta|}$.

Constraining $|\eta|$

As the value for $|\eta|$ cannot be extracted from the fit, we shall assume that it can be obtained from elsewhere. The precision on the angle γ is estimated for two scenarios: firstly when $|\eta|$ is known exactly, and secondly when $|\eta|$ is known with a Gaussian error of $\frac{\sigma_{|\eta|}}{|\eta|} = 10\%$.

In order to constrain $|\eta|$ in the fit for the case of a finite error on its value, the likelihood function is multiplied with a Gaussian:

$$\mathcal{L} \rightarrow \mathcal{L} \cdot \exp\left(-\frac{1}{2}\frac{\left(|\eta|_{\mathsf{estim}} - |\eta|_{\mathsf{fit}}\right)^2}{\sigma^2_{|\eta|}}\right). \tag{3.48}$$

where $|\eta|_{\mathsf{estim}}$ is the assumed value of $|\eta|$, and $|\eta|_{\mathsf{fit}}$ is the parameter varied in the fit.

Choice of Parameters

The parameters varied in the fit are

$$\left\{|\eta|^2, \mathsf{Im}(\eta), \mathsf{Im}(\overline{\eta})\right\},$$

which are related to the angle γ by

$$\phi_{\text{mix}} - \gamma = \frac{1}{2}\left(\arcsin\left(\frac{\mathsf{Im}(\eta)}{|\eta|}\right) - \arcsin\left(\frac{\mathsf{Im}(\overline{\eta})}{|\eta|}\right)\right). \tag{3.49}$$

$\mathsf{Im}(\eta)$ and $\mathsf{Im}(\overline{\eta})$ are effectively independent of each other. They are in principle correlated by the factor $\left(1 + |\eta|^2\right)$ in the denominator of the amplitude of the sine term in the asymmetry, but since $|\eta|^2 \ll 1$ the two terms are practically independent.

It has been established empirically that the errors in $\mathsf{Im}(\eta)$ and $\mathsf{Im}(\overline{\eta})$ are independent of their actual values. They are also independent of assumptions regarding the error on $|\eta|$, as long as this error does not exceed several hundred percent, and they scale with the square-root of the number of events.

The Error on γ From the Errors on the Fit Parameters

In the following we shall investigate how the errors on $\mathsf{Im}(\eta)$ and $\mathsf{Im}(\overline{\eta})$ translate to an error on the angle γ, and how this depends on the input values for $\phi_{\text{mix}} - \gamma$ and Δ_{qcd}, and the knowledge of the parameter $|\eta|$.

For this study it is useful to make the following definitions:

$$\phi_{\pm} \equiv \Delta_{\text{qcd}} \pm (\phi_{\text{mix}} - \gamma) \tag{3.50}$$

and

$$u_{+} \equiv \sin(\phi_{+}) = \frac{\mathsf{Im}(\eta)}{|\eta|}, \tag{3.51}$$

$$u_{-} \equiv \sin(\phi_{-}) = \frac{\mathsf{Im}(\overline{\eta})}{|\eta|}. \tag{3.52}$$

As long as $|\eta|^2$ is fixed, fitting $\mathsf{Im}(\eta)$ and $\mathsf{Im}(\overline{\eta})$ is equivalent to fitting $u_{\pm}$, and the errors in $u_{\pm}$ and $\mathsf{Im}(\eta)$, $\mathsf{Im}(\overline{\eta})$ are proportional to each other:

$$\sigma_{u_{\pm}} = \frac{1}{|\eta|}\sigma_{\mathsf{Im}(\eta),\mathsf{Im}(\overline{\eta})}. \tag{3.53}$$

In this case u_+ and u_- are uncorrelated. If $|\eta|^2$ is not fixed, the errors on $\sigma_{u_\pm}$ are:

$$\sigma_{u_\pm} = \sqrt{\left(\frac{1}{|\eta|}\sigma_{\mathsf{Im}(\eta),\mathsf{Im}(\overline{\eta})}\right)^2 + \left(u_\pm \frac{\sigma_{|\eta|^2}}{2\,|\eta|^2}\right)^2} \tag{3.54}$$

or in terms of $\sigma_{|\eta|}$ rather than $\sigma_{|\eta|^2}$

$$\sigma_{u_\pm} = \sqrt{\left(\frac{1}{|\eta|}\sigma_{\mathsf{Im}(\eta),\mathsf{Im}(\overline{\eta})}\right)^2 + \left(u_\pm \frac{\sigma_{|\eta|}}{|\eta|}\right)^2}. \tag{3.55}$$

u_+ and u_- are now correlated, with a covariance:

$$\mathrm{cov}(u_+, u_-) = u_+ u_- \left(\frac{\sigma_{|\eta|}}{|\eta|}\right)^2, \tag{3.56}$$

and the errors have a weak dependence on input values. Finally, for the error on γ this translates to:

$$\sigma_\gamma = \frac{1}{2}\Bigg(\left(\frac{\sigma_{u_+}}{\cos(\phi_+)}\right)^2 + \left(\frac{\sigma_{u_-}}{\cos(\phi_-)}\right)^2 \\ -2\tan(\phi_+)\tan(\phi_-)\frac{\sigma_{|\eta|}}{|\eta|}\Bigg)^{\frac{1}{2}}. \tag{3.57}$$

3.5.4 Results

The errors on $\mathsf{Im}(\eta)$ and $\mathsf{Im}(\overline{\eta})$ are:

$$\sigma_{\mathsf{Im}(\eta)} = \sigma_{\mathsf{Im}(\overline{\eta})} = \frac{3.37 \cdot 10^{-3}}{\sqrt{\text{years of data taking}}}. \tag{3.58}$$

For the case of zero–error on the parameter $|\eta|$, this translates directly to:

$$\sigma_{u_\pm} = \frac{0.160}{\sqrt{\mathsf{years\ of\ data\ taking}}}, \tag{3.59}$$

which corresponds to first order, for the case of small angles ϕ_+, ϕ_-, to an error on γ of:

$$\sigma_\gamma \approx \frac{6.5^\circ}{\sqrt{\mathsf{years\ of\ data\ taking}}}. \tag{3.60}$$

For the case of a 10% error on $|\eta|$ we find after one year that the error on $u_\pm$ is

$$0.16 < \sigma_{u_\pm} < 0.19 \tag{3.61}$$

and after 4 years:

$$0.08 < \sigma_{u_\pm} < 0.13 \tag{3.62}$$

and after 16 years:

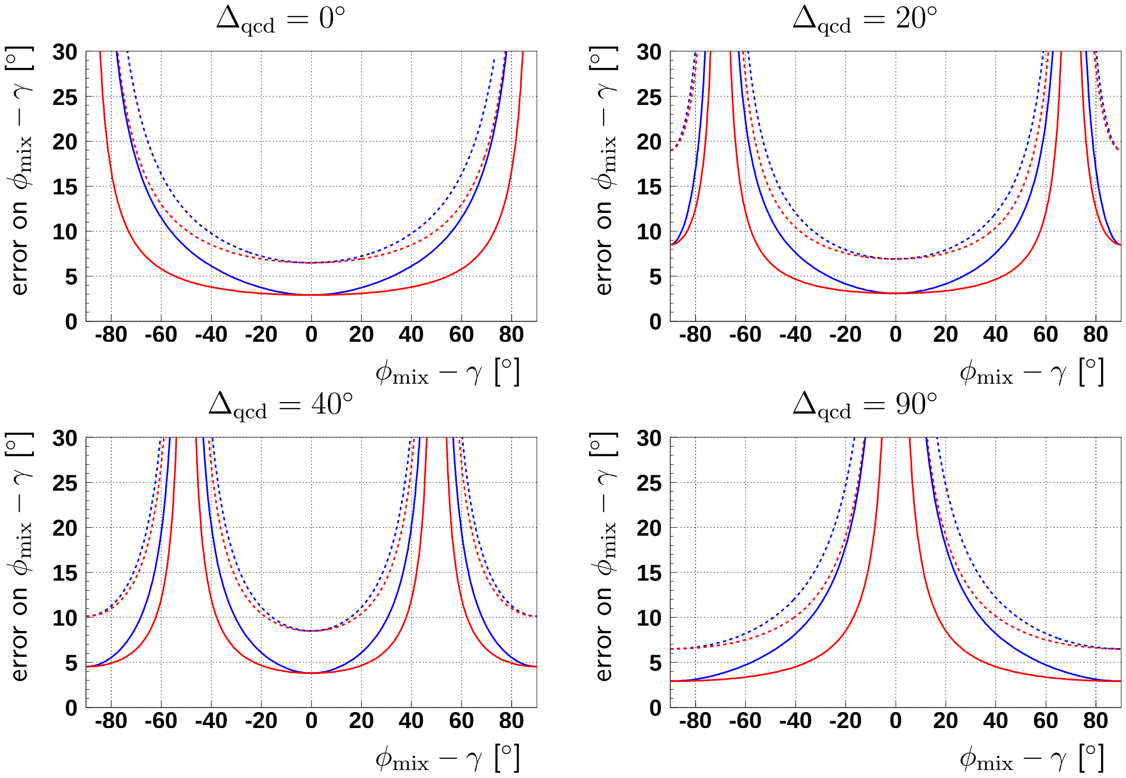
$$0.04 < \sigma_{u_\pm} < 0.11. \tag{3.63}$$

Here the upper bound corresponds to the case that $u_+ = u_- = 1$ and the lower bound to the case that $u_+ = u_- = 0$, which is equivalent in precision to the case when $|\eta|$ is known exactly.

It can be seen that an uncertainty in the knowledge of $|\eta|$ of 10% starts to propagate to an uncertainty on $u_\pm$ after a few years of data taking.

Because of the non-linear transformation between the fitted quantities and the angles, the error in γ is highly dependent on

Figure 3.18: Error on $\phi_{\mathrm{mix}} - \gamma$ after 1 and 5 years of data taking as a function of the value of $\phi_{\mathrm{mix}} - \gamma$, for four different values of Δ_{qcd}. Solid lines: 5 years of data taking, broken lines: 1 year of data taking; red lines: $\frac{\sigma_{|\eta|}}{|\eta|} = 0\%$, blue lines: $\frac{\sigma_{|\eta|}}{|\eta|} = 10\%$. The singularities seen in each plot are unphysical and are discussed in the text.



the values chosen for Δ_{qcd} and $\phi_{\mathrm{mix}} - \gamma$. Figure 3.18 shows the error on $\phi_{\mathrm{mix}} - \gamma$ as a function of the value of $\phi_{\mathrm{mix}} - \gamma$ for four different values of Δ_{qcd}. Results are given for one and five years of LHCb data taking, using equation 3.57 to calculate the values from the result for the error on $\mathsf{Im}(\eta)$. The red lines correspond to the case with zero error on the parameter $|\eta|$ and the blue lines to the case where $\frac{\sigma_{|\eta|}}{|\eta|} = 10\%$; the broken lines indicate the error after one year of data taking, and the solid lines indicate the error after five years.

Near the edges of the parameter space, where $\cos(\phi_{\pm})$ approaches zero, the simple error propagation laws used in equation 3.57 do not give a meaningful answers anymore. Therefore the errors shown are only meaningful for values of $\phi_{\mathrm{mix}} - \gamma$ at a distance from the edge of the parameter space that is larger than the estimated error itself. If it turns out that the values measured in the real data are indeed near the edge of the allowed parameter space an alternative procedure would be required to estimate the uncertainty in γ.

3.5.5 Comparison with Monte Carlo

To cross–check the results, errors are compared between the fits to the analytically generated events and the fits to those generated using Monte Carlo techniques. This is performed for two sets of values for $\phi_{\rm mix} - \gamma$ and $\Delta_{\rm qcd}$:

$$\left.\begin{array}{rcl} \phi_{\rm mix} - \gamma & = & 40^\circ \\ \Delta_{\rm qcd} & = & 0^\circ \end{array}\right\} \Rightarrow \left\{\begin{array}{rcl} u_+ & = & 0.643 \\ u_- & = & 0.643 \end{array}\right.$$

and

$$\left.\begin{array}{rcl} \phi_{\rm mix} - \gamma & = & 40^\circ \\ \Delta_{\rm qcd} & = & 40^\circ \end{array}\right\} \Rightarrow \left\{\begin{array}{rcl} u_+ & = & 0.985 \\ u_- & = & 0 \end{array}\right. .$$

The uncertainty on $|\eta|$ is assumed to be 10%.

For each combination of values for $\phi_{\rm mix} - \gamma$ and $\Delta_{\rm qcd}$, 100 sets of events are generated, each set equivalent to 1 year of LHCb data taking.

Each set of Monte Carlo events is generated with a different value for $|\eta|$. Those values are randomly generated according to a Gaussian distribution of width $\sigma_{|\eta|} = 0.1 \cdot |\eta|$, and mean $\mu_{|\eta|} = 0.021$. When constraining $|\eta|$ in the fit, using equation 3.48, it is always the mean value for $|\eta|$ that is used.

The results for the different methods, presented in table 3.7, show good agreement, giving confidence in the results of section 3.5.4.

3.5.6 Factors that Influence the Statistical Error

It is interesting to establish the dependence that the statistical error, σ_γ, has on certain input parameters and assumptions, in particular the level and distribution of background events.

Detector Effects

The statistical precision of the fit is directly proportional to the amplitude of the asymmetry. The asymmetry in the presence of detector effects is given by.

$$A_{\frac{B}{S},\sigma_\tau}(\tau) = e^{-\frac{1}{2}\sigma_\tau^2\cdot\omega_{\rm tag}^2}\frac{1}{1+B/S}A(\tau - \Gamma\sigma_\tau^2). \tag{3.64}$$

The error is also proportional to $\frac{1}{\sqrt{N_{S+B}}}$, where N_{S+B} is the total number of events. In terms of the number of signal events, N_{S+B} is given by:

$$N_{S+B} = N_S\left(1 + \frac{B}{S}\right) \tag{3.65}$$

where N_S is the number of reconstructed, triggered and tagged signal events. Therefore, the error on γ is proportional to

$$\sigma_\gamma \propto e^{\frac{1}{2}\sigma_\tau^2\cdot(\Delta m)^2}\frac{\sqrt{1+B/S}}{1-2\omega_{\rm tag}}\frac{1}{\sqrt{N_S}}. \tag{3.66}$$

Table 3.7: Fit results and error estimates obtained from fits to 100 sets of Monte Carlo data, compared with the results from the analytical method. Each set of events is equivalent to 1 year of LHCb data taking. The uncertainty in $|\eta|$ is assumed to be 10%. The term "calc. error" refers to the error on $u_\pm$ calculated from the MINUIT error estimate on $\mathsf{Im}(\eta)$ and $\mathsf{Im}(\overline{\eta})$, using equation 3.55.

			MC-generated events			Analytical evts	
$\phi_{\text{mix}}-\gamma$, Δ_{qcd}		input value	mean of 100 fit results	RMS of 100 fit results	mean calc. error from 100 fits	fit result	calc. error
40°, 0°	u_+	0.643	0.646	0.167	0.174	0.643	0.173
	u_-	-0.643	-0.671	0.155	0.174	-0.643	0.173
40°, 40°	u_+	0.985	0.997	0.182	0.190	0.985	0.188
	u_-	0.000	-0.002	0.150	0.161	$8\cdot10^{-5}$	0.160

Dependence on $|\eta|$

As shown above, the errors on $\mathsf{Im}(\eta)$ and $\mathsf{Im}(\overline{\eta})$ are independent of the actual values of $\mathsf{Im}(\eta)$ and $\mathsf{Im}(\overline{\eta})$. γ is related to $\mathsf{Im}(\eta)$ and $\mathsf{Im}(\overline{\eta})$ by

$$\sin(\Delta_{\text{qcd}}+(\phi_{\text{mix}}-\gamma))=\frac{\mathsf{Im}(\eta)}{|\eta|},$$

and

$$\sin(\Delta_{\text{qcd}}-(\phi_{\text{mix}}-\gamma))=\frac{\mathsf{Im}(\overline{\eta})}{|\eta|}.$$

Therefore the error on $\sin(\Delta_{\text{qcd}}\pm(\phi_{\text{mix}}-\gamma))$ scales directly with the value of $1/|\eta|$ that is assumed in the study, and for small errors on $\mathsf{Im}(\eta)$ and $\mathsf{Im}(\overline{\eta})$, so does the error on γ.

Including the dependence on $|\eta|$, the error on γ is proportional to:

$$\sigma_\gamma\propto\frac{1}{|\eta|}e^{\frac{1}{2}\sigma_\tau^2\cdot(\Delta m)^2}\frac{\sqrt{1+B/S}}{1-2\omega_{\text{tag}}}\frac{1}{\sqrt{N_S}}.\qquad(3.67)$$

Background

The simplest model for the background distribution, and that adopted in this chapter, is that the background fraction is independent of the decay time, and of the decay considered. Therefore the number of background events interpreted as $\mathrm{B_d^0}\to\mathrm{D^{*-}}\pi^+$ decays is the same as the number interpreted as $\overline{\mathrm{B_d^0}}\to\mathrm{D^{*-}}\pi^+$ decays, and the number interpreted as $\overline{\mathrm{B_d^0}}\to\mathrm{D^{*+}}\pi^-$ is the same as the number interpreted as $\mathrm{B_d^0}\to\mathrm{D^{*+}}\pi^-$.

Using the symbols X and Y for the two decay rates, the asymmetry without background is given by:

$$A(\tau)=\frac{X(\tau)-Y(\tau)}{X(\tau)+Y(\tau)}\qquad(3.68)$$

which, in the presence of this type of background, transforms to

$$\frac{\left(X+\frac{1}{2}\frac{B}{S}\left(X+Y\right)\right)-\left(Y+\frac{1}{2}\frac{B}{S}\left(X+Y\right)\right)}{\left(1+\frac{B}{S}\right)\left(X+Y\right)}$$

$$= \frac{X(\tau) - Y(\tau)}{\left(1 + \frac{B}{S}\right)\left(X(\tau) + Y(\tau)\right)}.$$

So the background cancels in the numerator, but increases the denominator and hence reduces the amplitude of the asymmetry:

$$A_{\frac{B}{S}}(\tau) = \frac{1}{1 + B/S}\mathcal{A}(\tau). \qquad (3.69)$$

This increases the error on the fit parameters by a factor of $1 + B/S$. Taking into account the extra number of events, which reduces the error again by $\sqrt{1 + B/S}$, the total increase in statistical uncertainty due to a background to signal ratio of B/S is $\sqrt{1 + B/S}$.

This simple model for the background is not very realistic, as it appears that most background will in fact come from $\mathrm{B}^0 \rightarrow \mathrm{D}^{*\pm}$ decays with correctly identified $\mathrm{D}^{*\pm}$. The background will therefore follow the $\mathrm{B}^0_{\mathrm{d}}$ oscillation.

Using the symbols δ_x and δ_y for the part of the background that does not follow the oscillations, the asymmetry in the presence of background becomes:

$$\frac{X(\tau)\left(1 + \frac{B}{S}\right) + \delta_x(\tau) - Y(\tau)\left(1 + \frac{B}{S}\right) - \delta_y(\tau)}{\left(1 + \frac{B}{S}\right)\left(X(\tau) + Y(\tau)\right)}$$

$$= \frac{X(\tau) - Y(\tau)}{X(\tau) + Y(\tau)} + \frac{\delta_x(\tau) - \delta_y(\tau)}{\left(1 + \frac{B}{S}\right)\left(X(\tau) + Y(\tau)\right)}$$

So the asymmetry in the presence of background that is more akin to that found in the event reconstruction is given by:

$$\mathcal{A}(\tau) + \delta\mathcal{A}(\tau), \qquad (3.70)$$

instead of

$$\frac{1}{1 + B/S}\mathcal{A}(\tau) \approx \mathcal{A}(\tau) - \frac{B}{S}\mathcal{A}(\tau), \qquad (3.71)$$

where $\delta\mathcal{A}(\tau)$ is small compared to $\frac{B}{S}\mathcal{A}(\tau)$. The observation that most of the background will come from $\mathrm{B}^0_{\mathrm{d}} \rightarrow \mathrm{D}^{*\pm}$ decays therefore leads to a reduced statistical error, as it leaves the amplitude nearly unchanged.

For the study presented above, the simple background model has been used with a B/S ratio of 0.2. This model gives a conservative estimate of the statistical precision in γ.

3.5.7 Systematic Errors

The results of section 3.5.4 are statistical errors alone. Below we will discuss possible systematic uncertainties, and how they might be controlled.

The Time Resolution

The measured asymmetry can be written as

$$\begin{aligned} A_{m,\eta}(\tau) &= C \cdot \cos\Big(\Delta m \left(\tau - \Gamma\sigma_\tau^2\right)\Big) \\ &\quad + S \cdot \sin\Big(\Delta m \left(\tau - \Gamma\sigma_\tau^2\right)\Big) \qquad (3.72) \end{aligned}$$

The parameters C and S have been introduced as the amplitude of the cosine and the sine term respectively:

$$C = D\frac{1 - |\eta|^2}{2\left(1 + |\eta|^2\right)}$$

$$S = D\frac{\mathsf{Im}(\eta)}{1+|\eta|^2}, \tag{3.73}$$

where D stands for the dilution due to detector effects, as defined in equation 3.40. With these definitions:

$$A_{m,\eta} = C\sqrt{1+\frac{S^2}{C^2}}\cdot\cos\left(\Delta m\left(\tau-\Gamma\sigma_\tau^2\right)-\arctan\left(\frac{S}{C}\right)\right). \tag{3.74}$$

So fitting $\mathsf{Im}(\eta) \approx \frac{1}{2}S/C$ is essentially equivalent to fitting the phase-shift:

$$\Delta\phi \equiv \Delta m\Gamma\sigma_\tau^2 + S/C \tag{3.75}$$

To ignore the finite time-resolution, we would therefore need $\Delta m\Gamma\sigma_\tau^2 \ll 2|\eta|$. Above we found $\sigma_\tau = 170$ fs for the inclusive case. Hence:

$$\frac{\Delta m\Gamma\sigma_\tau^2}{2|\eta|} = 21\%. \tag{3.76}$$

To first order, this corresponds to a bias in $\Delta_{\rm qcd} \pm (\phi_{\rm mix}-\gamma)$ of $\sim 13°$ if the effect were completely ignored. This bias is reduced for the interesting case where $\Delta_{\rm qcd} \ll 1$, and the phase–shift due to the finite time–resolution is the same for $\mathsf{Im}(\eta)$ and $\mathsf{Im}(\overline{\eta})$. Then the bias cancels to first order when calculating

$$\phi_{\rm mix}-\gamma = \frac{1}{2}\left(\arcsin\frac{\mathsf{Im}(\eta)}{|\eta|} - \arcsin\frac{\mathsf{Im}(\overline{\eta})}{|\eta|}\right).$$

The above shows that it will be essential to obtain an accurate description of the time–resolution. It is also important to take into consideration any dependence that the acceptance function might have on the true decay time. This was ignored when deriving the expression for the measured decay rates and would have an effect on the phaseshift, in particular where the acceptance changes rapidly with the decay time.

Background

The background level can be estimated from the sidebands in the distribution of the reconstructed $\mathrm{B_d^0}$ mass. To estimate systematic errors, for example due to possible $\mathcal{CP}$ asymmetries in the background, requires a better knowledge of the background composition than is possible with the limited statistics available.

Mistag

The mistag fraction needs to be known exactly in order extract $\mathsf{Im}(\eta)$ from amplitudes of the asymmetry. It can be measured by fully reconstructing charged B decays. The procedure assumes no correlation between the B used for tagging, and the tagged B. The decay $\mathrm{B_d^0} \to \mathrm{J}/\psi\mathrm{K}^*$ can also be used to extract the mistag fractions [B+00].

It is expected that from the channel $\mathrm{B}^\pm \to \mathrm{J}/\psi\mathrm{K}^\pm$ alone, the mistag fraction can be determined to a relative precision of 10^{-3} within one year [B+00], so the systematic error due to this source can be safely ignored.

These control channels can also be used to determine a possible charge asymmetry in the mistag.

Rate Normalisation

In the expressions for the asymmetries so far, it has been assumed that there are as many $\mathrm{B_d^0}$ produced as $\overline{\mathrm{B_d^0}}$, and that the tagging efficiencies for both mesons are the same. In fact a production asymmetry of the order of a percent or so is anticipated [TP98]. An asymmetry in the detector might also lead to different tagging efficiencies because of the correlation between the detected particle flux and the charge, brought about by the dipole magnet. These detector asymmetry effects will be minimised by switching the magnet polarisation frequently.

Using the symbols ρ and $\bar{\rho}$ for the production rates of $\mathrm{B_d^0}$ and $\overline{\mathrm{B_d^0}}$ mesons respectively, and ϵ and $\bar{\epsilon}$ for their respective tagging efficiencies, the normalisation factor r_B is defined as:

$$r_B(\tau) \equiv \frac{\rho(\tau)\epsilon(\tau)}{\bar{\rho}(\tau)\bar{\epsilon}(\tau)}. \tag{3.77}$$

The asymmetry changes from

$$A(\tau) = \frac{\Gamma(\mathrm{B_d^0} \to \mathrm{D}^{*-}\pi^+) - \Gamma(\overline{\mathrm{B_d^0}} \to \mathrm{D}^{*-}\pi^+)}{\Gamma(\mathrm{B_d^0} \to \mathrm{D}^{*-}\pi^+) + \Gamma(\overline{\mathrm{B_d^0}} \to \mathrm{D}^{*-}\pi^+)} \tag{3.78}$$

to:

$$A_m(\tau) = \frac{\Gamma(\mathrm{B_d^0} \to \mathrm{D}^{*-}\pi^+) - r_B(\tau)\Gamma(\overline{\mathrm{B_d^0}} \to \mathrm{D}^{*-}\pi^+)}{\Gamma(\mathrm{B_d^0} \to \mathrm{D}^{*-}\pi^+) + r_B(\tau)\Gamma(\overline{\mathrm{B_d^0}} \to \mathrm{D}^{*-}\pi^+)}, \tag{3.79}$$

neglecting the other detector effects. For a time–independent r_B, we get:

$$A_m(\tau) = A(\tau) - \frac{1}{2}\left(r_B - 1\right), \tag{3.80}$$

which can in principle be fitted. This would result in an increased statistical error, and relies on the assumption that r_B is time independent. Alternatively, r_B can be extracted from $\mathrm{B_d^0} \to \mathrm{J}/\psi\mathrm{K}^*$ decays, to a precision of about $2 \cdot 10^{-3}$ within a year [TP98, B+00].

Error on η

As $|\eta|$ cannot be extracted from the asymmetries, it must be obtained from elsewhere. One option is to calculate $|\eta|$, which is difficult mainly because of unknown hadronic effects. It might be possible to use experimental data to constrain those effects, for example by using the flavour–symmetry related decay $\mathrm{B_s^0} \to \mathrm{D_s K}$ which has $|\eta| \sim 1$, but this has not yet been investigated.

Any error in the value obtained for $|\eta|$ induces a corresponding error on γ according to expression 3.49.

3.5.8 γ Reach - Result Summary

For 520 000 reconstructed, triggered and tagged $\mathrm{B_d^0} \to \mathrm{D}^*\pi$ events, a wrong–tag fraction of $\omega_{\mathrm{tag}} = 30\%$ and a background to signal ratio of $S/B = 0.2$, the statistical error on the parameters $\mathsf{Im}(\eta)$ and $\mathsf{Im}(\overline{\eta})$ is $3.37 \cdot 10^{-3}$. This error scales with the square–root of the number of events. For $|\eta| = 0.021$, it translates to an uncertainty in γ of $\sim 7°$, depending strongly on the values for $\phi_{\mathrm{mix}} - \gamma$ and Δ_{qcd}.

The extraction of γ from the fit results requires that the parameter $|\eta|$ is known to a reasonable precision. For an error on $|\eta|$ of 10%, the systematic and statistical uncertainties contribute in equal measure to the total uncertainty in γ after a few years of data taking. The method is very sensitive to the exact knowledge of the decay–time resolution function, and also requires that that a dependence of the acceptance function on the true decay time for a given measured decay time is taken into account.

3.6 Conclusion

Two methods of reconstructing the decay $\mathrm{B_d^0} \rightarrow \mathrm{D}^*\pi$ have been presented, the exclusive method, where an entire decay chain is reconstructed, and the inclusive method, where the full $\mathrm{B_d^0}$ 4–momentum is reconstructed using the fast and the slow pion from the $\mathrm{B_d^0}$ and the D^* decay only.

The combined event yield from these two methods is estimated to be about half a million reconstructed, tagged and triggered $\mathrm{B_d^0} \rightarrow \mathrm{D}^*\pi$ events per year. Nearly 90% of the events are provided by the inclusive method alone. Despite the reduced set of information used in the reconstruction, the inclusive method provides a good signal to background ratio of $4.4^{+4.4}_{-1.1}$, and a time resolution of 170 fs, which is small compared with the $\mathrm{B_d^0}$-$\overline{\mathrm{B_d^0}}$ oscillation period. The exclusive method achieves a similar signal to background ratio, and a time resolution of 60 fs.

Given this performance, the angle γ can be measured with a precision of $\sim 7^\circ$ within one year of data taking, depending on the values for $\phi_{\mathrm{mix}} - \gamma$, Δ_{qcd} and $|\eta|$ assumed in the study, and the uncertainty in $|\eta|$. Independent of these assumptions are the results for the parameters $\mathsf{Im}(\eta)$ and $\mathsf{Im}(\overline{\eta})$, which can be fitted with a precision of $3.4 \cdot 10^{-3}$ within one year of data taking.

The γ measurement crucially depends on the knowledge of $|\eta|$, which needs to be taken from elsewhere, and on the exact knowledge of the decay time resolution function.

Chapter 4

Full–Scale RICH 2 Prototype

4.1 Introduction

A full scale prototype of the RICH 2 detector was built, and tested in the summer and autumn of 1998 at the CERN SPS facility. The main focus of this chapter is to report on the performance of this prototype in the two key aspects of a RICH detector, the number of detected photons and the Cherenkov angle resolution. The results are compared with those from a detailed Monte Carlo simulation, in order to establish if our understanding of the prototype is correct and therefore whether the RICH performance studies, using the LHCb detector simulation, are meaningful. First, a brief description of the test–beam setup and the main hardware components is given.

4.2 Experimental Setup

4.2.1 Overview

A full scale prototype of RICH 2 has been built and equipped with commercially available 61–pixel HPDs. The prototype is designed as a vacuum vessel, connected to a gas circulation system to control and monitor the pressure and purity of the $\mathrm{C\,F_4}$ gas, and monitor the temperature. It was placed into the X7 beam line of the CERN SPS facility.

A schematic of the test–beam setup is shown in figure 4.1, together with the definition of the co–ordinate system used in the text.

The beam enters the vessel along a tube of 90 mm diameter, through a 250 μm thick Mylar window. The beam direction is measured using a three–plane Si telescope.

The Cherenkov photons generated in the 1.8 m long gas volume are reflected off a spherical mirror with focal length $f = 4\,\mathrm{m}$, onto the photodetectors placed in the focal plane of the mirror. The mirror axis is tilted by 18° relative to the beam line, and the axis connecting the mirrors and the photo detector plane is at a 36° angle to the beam line.

To trigger the readout electronics, scintillation counters are placed into the beam. To allow testing of the system independent of

Figure 4.1: RICH2 test–beam setup and co–ordinate system

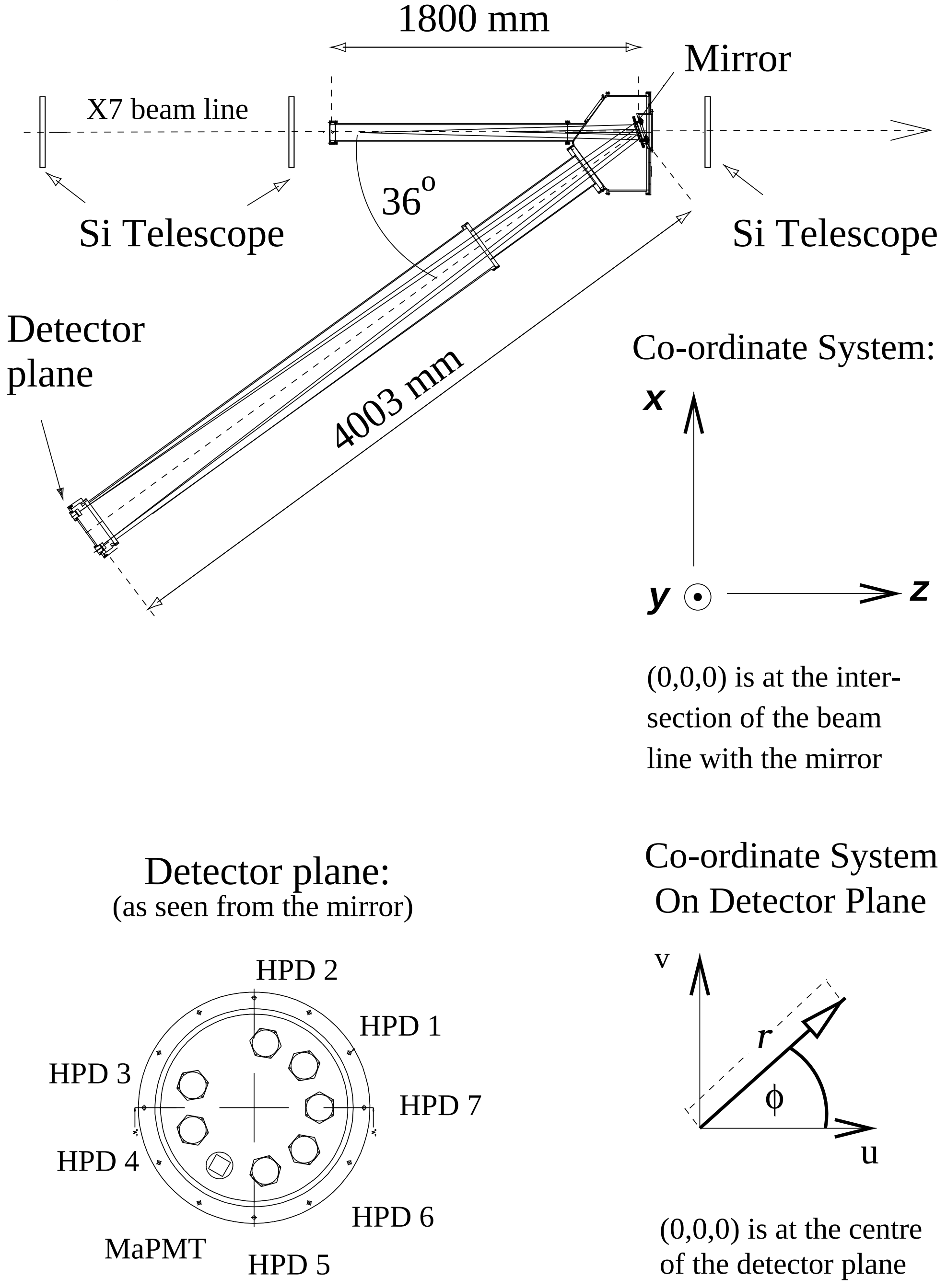
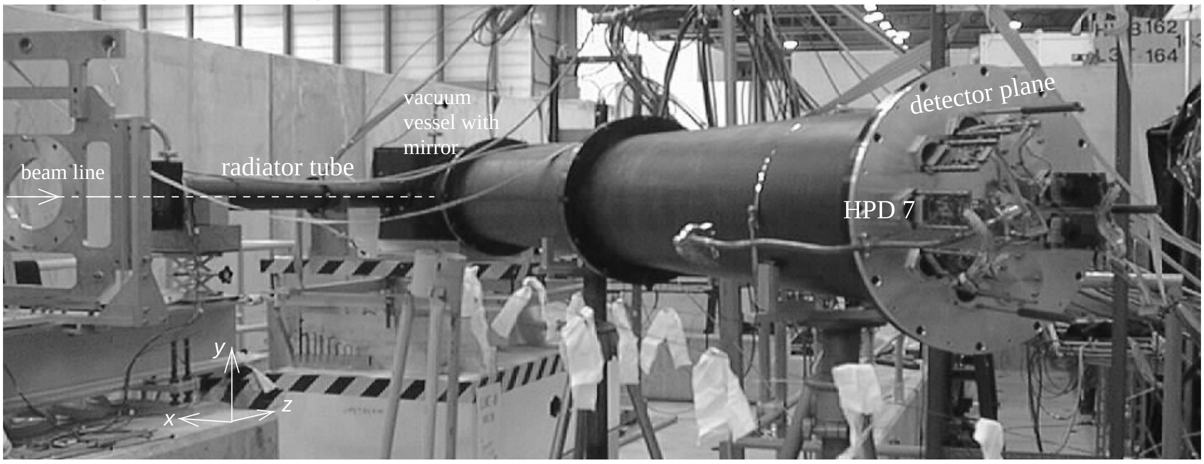


Figure 4.2: Photograph of RICH prototype being set up in the X7 beam area



the beam, an LED is placed inside the vessel near the entrance window of the beam. The LED light source can be synchronised with the readout electronics.

4.2.2 The Beam

The RICH prototype is mounted in the CERN X7 beam line. The beam is set to provide negative particles with momenta from 10 to 150 GeV. The momentum spread of the beam is less than 1%. The beam composition is mainly pions and kaons in the ratio of 9:1.

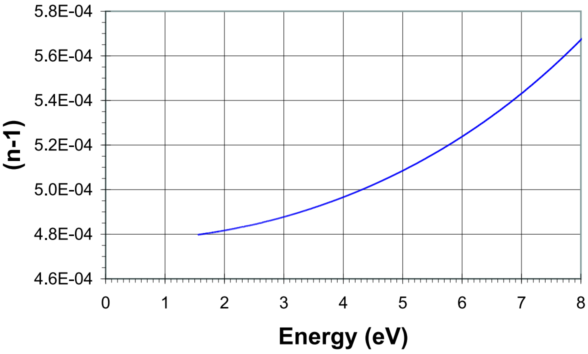
A CEDAR differential Cherenkov counter is placed in the beam, in order to be able to obtain data for a single particle type. A differential Cherenkov counter detects photons within a very narrow Cherenkov angle range, and thus triggers, for a given Cherenkov angle acceptance, on one particular type of particle. The CEDAR identifies the desired particle type with very high purity, but unknown efficiency [BMP^{+}82].

Readout of the detectors is triggered by time correlated signals provided by two pairs of scintillation counters placed 8 m apart along the beam line. The smaller scintillation counter is $20 \times 20\,\text{mm}^2$ in size, which defines the dimensions of the beam.

The beam divergence and position are measured using a silicon telescope consisting of three planes of Si pixel detectors. Two planes are placed upstream of the detector and one downstream, the first and the last plane being 8 m apart. Each Si plane is segmented into 22×22 square pixels with a side length of 1.3 mm.

The aperture of the beam and hence the beam divergence can be varied independently in x and y direction. For the study presented here, data runs with a small beam divergence and wide aperture have been selected. The beam divergence in x and y has been measured to be between 0.12 mrad and 0.18 mrad, as described in section 4.4.

Figure 4.3: Dependence of the refractive index (n) − 1 on the photon energy, for CF_4 at STP



4.2.3 The RICH Prototype

Mechanics

The vacuum vessel is constructed using aluminium alloy[1] plate, folded and welded at the box edges. The box is black–anodised to reduce reflection of stray light. Port–holes are machined into the side and backplate. These can be used, for example, to mount a standard photomultiplier tube to check for light leaks. The ports are sealed using blank plates and O–rings when the box is connected to the pump or the gas system.

A 10 mm thick removable backplate supports the mirror. It is bolted to the box and sealed with a rubber O–ring. The mirror position can be adjusted with three micrometer screws. A movement of a micrometer screw of 1 mm corresponds to a shift of the image on the detector plane of ∼ 8 cm.

Attached to the box are two extensions of the same material as the box. The tube along the beam line provides the RICH 2 radiator length of 1.8 m, and the tube at 36° covers the necessary 4 m distance between mirror and detector plane.

The detector plane provides space for seven HPDs and one MaPMT in a circular arrangement with their centres at a 113 mm radius. The MaPMT will not be discussed here. However, results from a testbeam using only MaPMTs are presented in chapter 5. HPD 1 . . . HPD 7 are positioned at the azimuthal angles 40°, 75°, 165°, 195°, 285°, 320°, 0° respectively. A schematic of the detector plane is shown in figure 4.1.

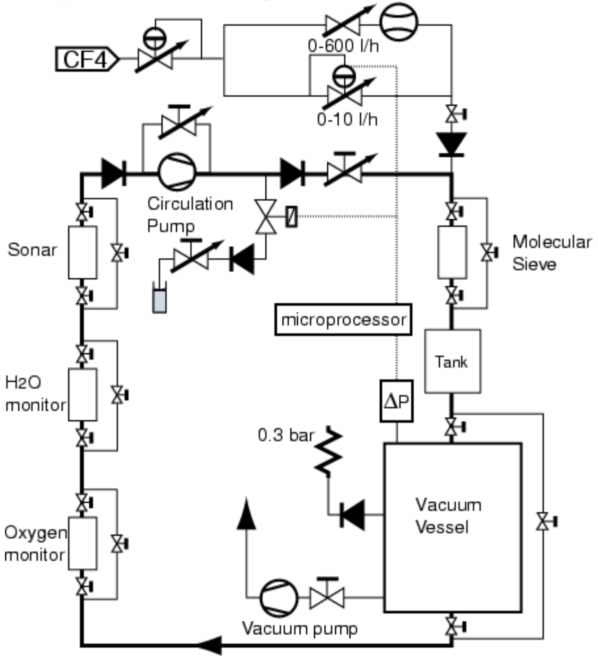
The CF_4 Gas Radiator

For all tests reported here, CF_4 gas is used as the radiator, as in the final RICH 2 design. The pressure and temperature of the gas are constantly monitored. The refractive index of CF_4 as a function of photon energy at STP, as parameterised in [A+90a], is shown in figure 4.3. The data presented here have been taken at 900 mbar and 20°C.

---

[1] BS_HS30

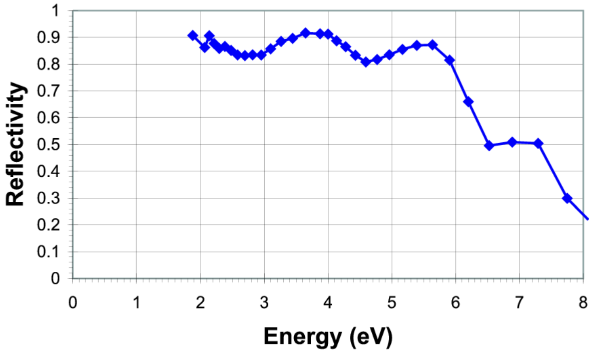
Figure 4.4: CF_4 gas circulation system



The schematic in figure 4.4 shows the gas circulation system which provides the CF_4–gas to the RICH prototype. A molecular sieve is used to remove water vapour. The system uses a microprocessor interface[2] to set and stabilise the gas pressure and to record pressure, temperature and the concentration of water vapour and oxygen throughout the data taking.

The absolute pressure of the CF_4 in the vessel is maintained to within 1 mbar of the required value. The oxygen concentration is always controlled below 0.1% and the water vapour concentration below 100 ppm by volume.

Figure 4.5: Mirror reflectivity as a function of photon energy



The Mirror

The focal length of the mirror is 4003 mm. It is made from glass coated with aluminium. Its reflectivity as a function of photon energy has been measured at CERN and is given in figure 4.5

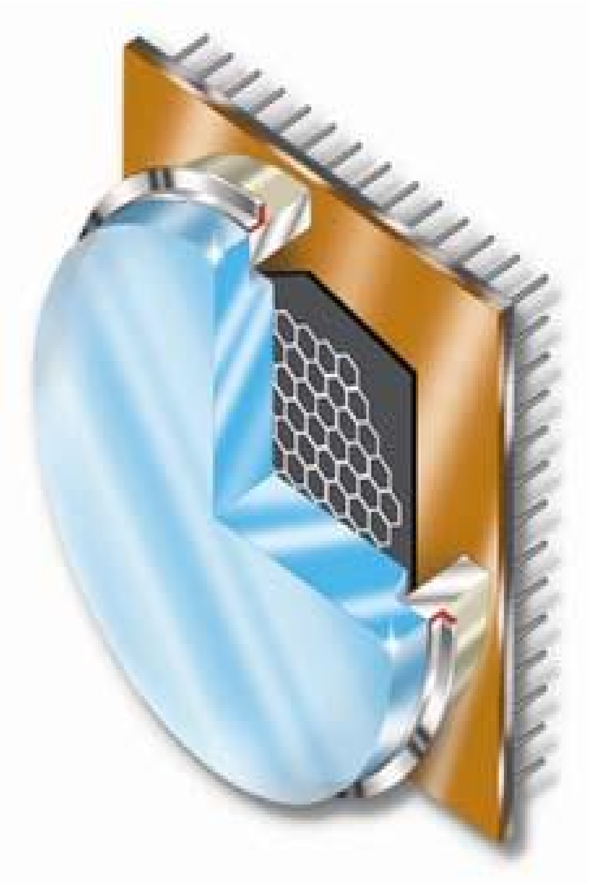
The 61 Pixel HPD

For the RICH Prototype tests the final HPD design described in section 2.3.6 was not yet available. Instead the commercially available “61 Pixel HPD” from DEP is used. A visualisation of the tube is shown in figure 4.6. The photo cathode is an S20 type (semi–transparent $SbNa_2KCs$), deposited on the inside surface of a 2 mm thick fused silica window. The sensitive range of the photo cathode starts at about 2.5 eV and is limited at higher energies by the fused silica at 7.3 eV. The quantum efficiency was measured by DEP and is shown in figure 4.7. The last measurement point is at 6.2 eV, the graph is extrapolated

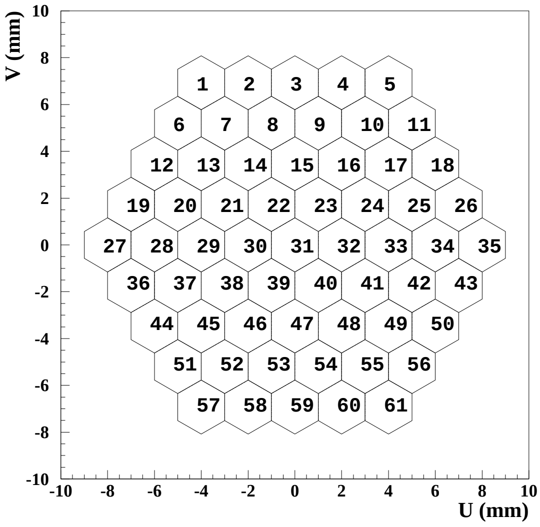
---

[2] Siemens S595U

Figure 4.6: The 61 Pixel HPD

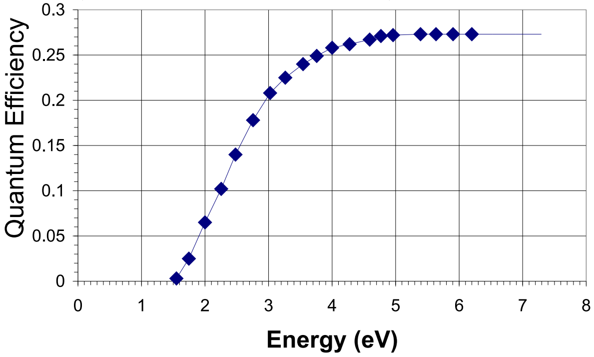


Visualisation of tube



Hexagonal pixel array with the pixel number scheme indicated.

Figure 4.7: Quantum Efficiency of DEP S20 photo cathode as a function of photon energy (averaged over 7 HPDs)



to 7.3 eV, assuming the quantum efficiency stays constant between 6.2 and 7.3 eV. With this, and a cut–off at 7.3 eV, the integrated quantum efficiency is

$$\int Q_{\text{eff}}(E)\,\mathrm{d}E = 1.34\,\text{eV}. \tag{4.1}$$

The Si diode detector is processed on a 300 μm n–type substrate and the p–type implant is segmented as an hexagonal array of 61 hexagonal pixels with pixel dimensions of 2 mm flat to flat, as shown in figure 4.6. The electron optics are such that the image on the pixel array is magnified by 6.1 % with respect to the image on the photo cathode, so the effective pixel size on the cathode surface is 1.88 mm. The effective area on the anode has therefore a diameter of $\sim$ 17 mm.

The photo electrons travel 12 mm from the photo cathode at a potential of -12 kV to the back (ohmic) surface of the silicon at 0 V. This surface implant has been thinned to $\sim$ 100 nm thickness. The diodes are reverse biased at 60 V which ensures depletion. Ohmic contacts to the diode pix-

Figure 4.8: Schematic of 61 pixel HPD

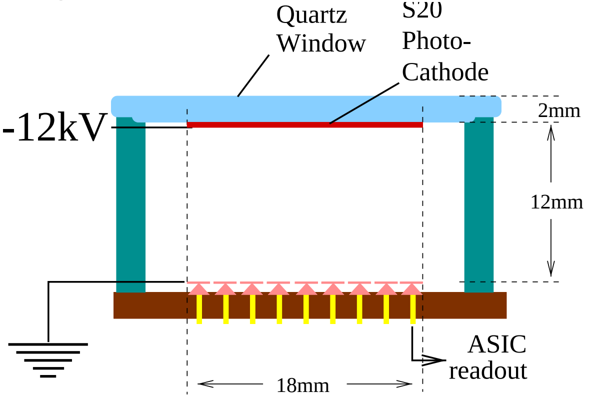
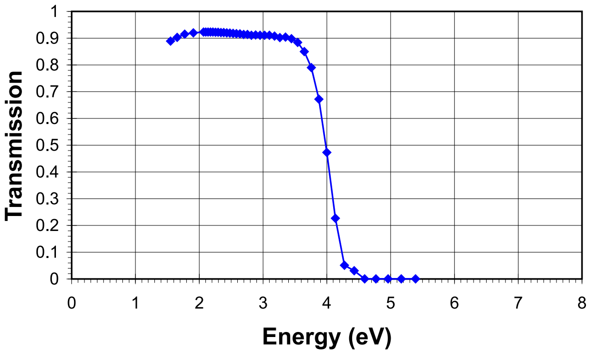


Figure 4.9: Transmission of borosilicate ("Pyrex") filters

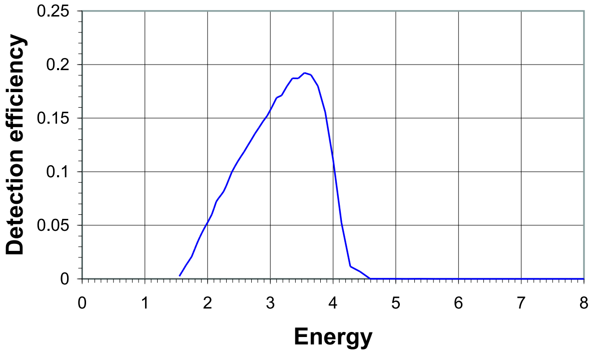


els are provided by a 2.5 mm pitch, 80–pin grid array which feeds through into the vacuum tube via a ceramic carrier, solder bump bonded to the metallised p–type pixels. A schematic of the HPD is shown in figure 4.8.

Filters

Filters are used to limit the wavelength range at higher energies and thus reduce chromatic aberration. The filters are made from borosilicate ("Pyrex"), that cuts off at about 4 eV, or 310 nm. The transmission of these filters as a function of energy is shown in figure 4.9. Data were taken with and without Pyrex filters.

Figure 4.10: Detection efficiency as a function of photon energy



Detector Response N_0

Multiplying the curves 4.5, 4.7 and 4.9 gives the detection efficiency as a function of the photon energy for the case with filter, as shown in figure 4.10. Integrating this, gives the detector response parameter N_0, defined in 2.7:

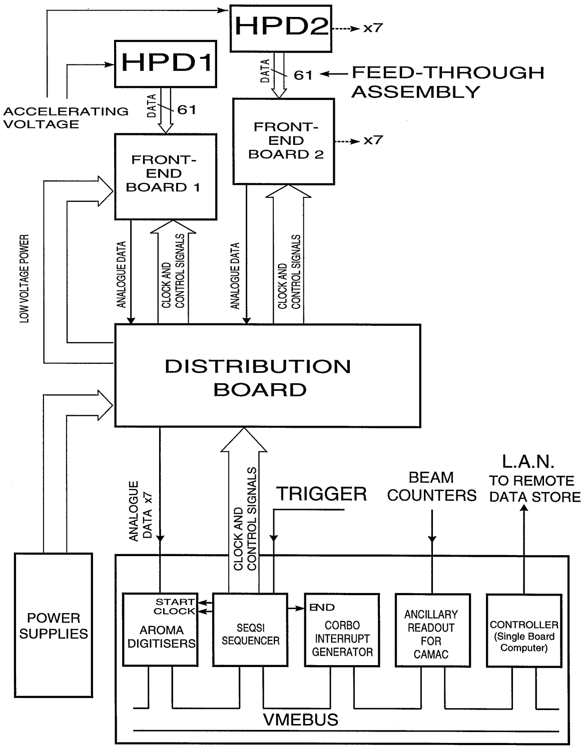
$$N_0 = 115\,\mathrm{cm}^{-1} \tag{4.2}$$

and for the case without filter:

$$N_0 = 358\,\mathrm{cm}^{-1}. \tag{4.3}$$

This neglects the finite efficiency with which the photo electrons generated in the cathode are detected by the Si sensor. In this chapter, all values for N_0 are given for a photo electron detection efficiency of $\epsilon_D = 100\%$. This is motivated by the photon counting method used below, where the signal loss due to undetected photo electrons is corrected for.

Figure 4.11: RICH prototype readout and DAQ

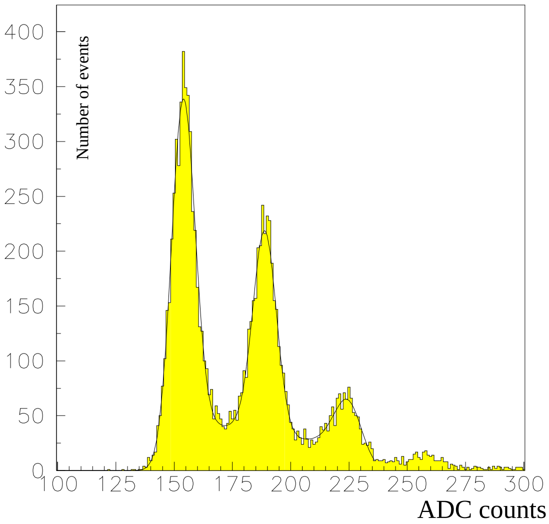


Readout and DAQ

The small single photo electron signals and the large number of pixel connections out of the vacuum envelope make the readout and data acquisition system for the prototype RICH particularly challenging.

A low–noise and slow (compared to the 40 MHz bunch–crossing at LHCb) pre–amplifier is used. Fast readout electronics as described in chapter 5 were not available for this test. Each HPD is connected to a front–end board, on which a 128–channel Viking VA2 pre–amplifier/shaper/multiplexer chip is mounted [MNTW94]. The multiplexed analogue signals from each of the seven front–end boards are transmitted via distribution boards to a VME flash–ADC system[3]. A schematic of the complete readout chain is shown in figure 4.11. A detailed description of the electronics and the data acquisition system can be found in [A+98].

Figure 4.12: Pulse height spectrum from analogue readout of a 61 pixel HPD



An example of the ADC spectrum of a single HPD pixel, obtained with the LED, is shown in figure 4.12. The zero, one two and three photo electron peaks are clearly visible.

---

[3] AROMA: an 8–channel flash–ADC module, designed for the analogue readout of prototype silicon detectors for the CERN ATLAS experiment [ARO]

4.3 Monte Carlo Simulation

The aim of the test–beam was to establish whether the RICH prototype works according to expectation, especially with respect to the photon yield and the Cherenkov angle resolution. To evaluate the prototype, the test–beam results are compared with the results from a detailed purpose–built Monte Carlo simulation. The simulation assumes a particle beam with a beam divergence as measured in the test–beam, typically $\sim 0.125\,\mathrm{mrad}$. The beam profile is assumed to be uniform within the $20 \times 20\,\mathrm{mm}$, as seen in the data. The measurement of the beam divergence and the beam profile are discussed in section 4.4.

Photons are generated with energies between $1.5\,\mathrm{eV}$ and $7.3\,\mathrm{eV}$, according to equation 2.5, at random points along the particle trajectory.

To calculate the refractive index of $\mathrm{C\,F_4}$ as a function of photon energy, the Lorentz–Lorentz equation is used (from [YS93]):

$$\frac{n^2-1}{n^2+2} = a\,f(E) \tag{4.4}$$

with

$$a = 0.3738\,\mathrm{cm}^3 \cdot n_d \tag{4.5}$$

where n_d is the molecular number density of the gas. Here, $\mathrm{C\,F_4}$ gas is treated as an ideal gas, so that

$$n_d = \frac{1}{R_{\mathrm{gas}}} \cdot \frac{P}{T} \tag{4.6}$$

where R_{gas} is the Gas Constant, P the pressure and T the temperature. The function f in equation 4.4 is given by the Sellmeir relation:

$$f(E) = \frac{F_1}{E_1^2 - E^2} + \frac{F_2}{E_2^2 - E^2}. \tag{4.7}$$

From [A$^+$90a], we take $F_1 = 7626.16\,\mathrm{eV}^2$, $E_1 = 20.0\,\mathrm{eV}$, and $F_2 = 0$. Figure 4.3 shows the parametrisation of the refractive index of $\mathrm{C\,F_4}$ as a function of photon energy at STP.

After the Cherenkov angle has been calculated, the photon is ray–traced through a computer model of the prototype. Photons that hit the walls of the vessel or tubes are treated as lost. The measured mirror reflectivity and filter transmission, shown in figures 4.5, and 4.9, are used to calculate the probability that a photon is absorbed at the mirror or the filters. To calculate the probability that a photon reaching the photodetectors is converted to a photo electron, the quantum efficiency as a function of photon energy, averaged over the seven tubes, is used. This is shown in figure 4.7.

Once the photo electron has been produced, its detection efficiency at the Si detector is not simulated. Instead, the data are corrected for signal loss, and the corrected photon counts are compared with the Monte Carlo results.

4.4 Measuring The Beam Divergence

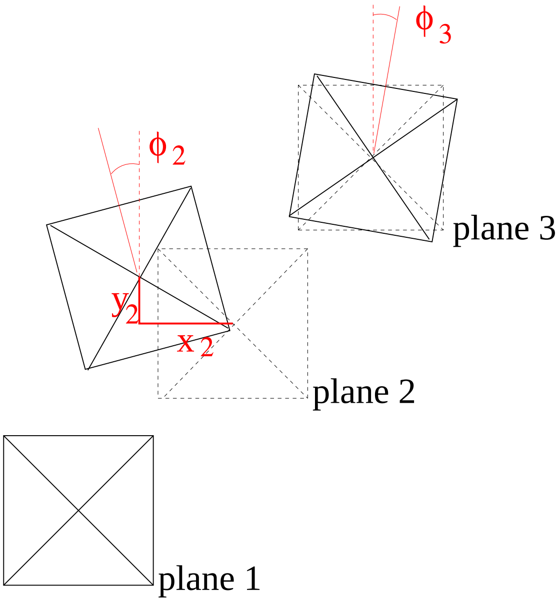
For the Cherenkov angle reconstruction explained later in this chapter, the beam is assumed to go exactly along the z axis. This assumption is not entirely correct and causes an additional spread in the reconstructed Cherenkov angles. In order to be able to reproduce this spread in the Monte Carlo simulation, the beam divergence needs to be known.

The three Si telescope planes are used to measure the beam divergence. For this measurement, events are selected where each of the three Si planes has exactly one hit. As three points always lie in one plane, the task of fitting the particle trajectory can be reduced to finding that plane and then performing a linear regression.

As the positions of the Si planes are not known accurately, their exact relative positions have to be determined from the data. To do so, a fit is performed by varying the assumed x and y position of the central plane as well as the rotation angles of the central and the third plane. The rotation is performed around an axis going through the centre of the plane, parallel to the z–axis. The parameters varied in the fit are shown in figure 4.13.

To find the optimal values for these parameters, the total χ^2 of all the straight–line fits to the Si hits in first 12 k events in the given data set, is minimised. This procedure leads to the results given in table 4.1 for the two runs that are discussed later in this chapter. As the Si telescope position was not changed between the runs, the alignment procedure is expected to return the same results in both cases. As can be seen from the table, the agreement is indeed very good.

Figure 4.13: Variables used in the alignment of the Si telescope. The position of the planes assumed at the start of the fit is given by the broken lines, the true position by the solid lines.



The values for the beam divergence in x and y are between 0.12 and 0.18 mrad, which includes the contribution due to pixelisation of ~ 0.047 mrad.

The hit densities in the three Si planes seen in figure 4.14 show a nearly uniform beam profile within the $20 \times 20\,\mathrm{mm}^2$ acceptance defined by the smaller of the scintillation counters that trigger the read out.

The measured RMS of the beam position

Table 4.1: Results of the beam divergence measurement and Si telescope alignment. The alignment co–ordinates are given in figure 4.13. The results are given for the data discussed in section 4.5.4.

	run with filter	run without filter
x_2	$-0.82\,\mathrm{mm}$	$-0.84\,\mathrm{mm}$
y_2	$0.64\,\mathrm{mm}$	$0.67\,\mathrm{mm}$
ϕ_2	8 mrad	19 mrad
ϕ_3	5 mrad	25 mrad
beam div x	0.135 mrad	0.179 mrad
beam div y	0.120 mrad	0.124 mrad
RMS beam pos. x	5.1 mm	5.0 mm
RMS beam pos. y	5.3 mm	5.3 mm

Figure 4.14: Hit density in the three Si telescope planes for the run with filter. (units are mm)

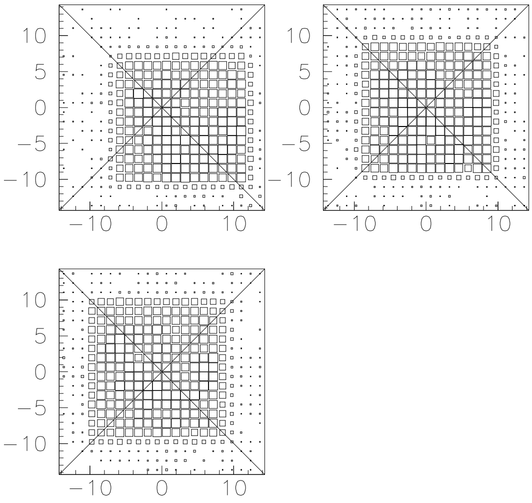
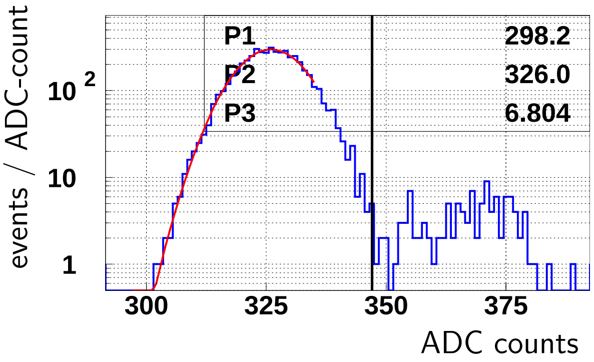


Figure 4.15: ADC spectrum from pixel lying on the Cherenkov ring, after 5 k events, with Pyrex filter. A Gaussian is fit to the pedestal. P1, P2, P3 are amplitude, mean and width respectively, and a line marks the 3σ cut.



in x and y of 5.0 to 5.3 mm is that expected from a nearly flat beam profile. A perfectly flat beam profile would result in an RMS in both coordinates of 5.8 mm for a $20 \times 20\,\mathrm{mm}^2$ acceptance.

4.5 Photon Counting

The number of photons detected determines, together with the Cherenkov angle resolution, the performance of the RICH detector. Thus a key aim of the test–beam was to determine whether the observed photon yield was in accordance with expectation.

The basic photon counting procedure is to define for each pixel a threshold cut and count everything above that cut as one hit. To set the threshold cut, the pedestal of each pixel is fitted with a Gaussian. The cut is set $m\sigma$ above the pedestal mean, where σ is the pedestal width. The default for the analysis presented below is $m = 3$.

As there is some overlap between the pedestal and the signal peak, only the part of the pedestal where this overlap is small, is used in the fit. The interval over which the fit is performed is $[\mu - 5\sigma, \mu + 1.5\sigma]$, where μ is the fitted pedestal mean. For the purpose of defining this interval, a first fit is performed without limits.

4.5.1 Background Estimates

In order to compare the measured Cherenkov photon yield with expectation, hits originating from background sources must first be subtracted.

Three types of background are treated separately:

i) electronic noise;

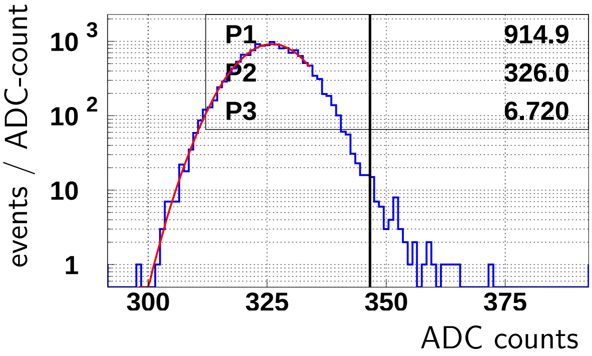
ii) detector noise and other non–Gaussian, beam–independent background;

iii) beam–related background.

Electronic Noise

Assuming that the electronic noise is Gaussian in distribution, the fit to the pedestal is used to estimate the number of hits from this source. The integral of the Gaussian, calculated from the fit result, is multiplied by the fraction of the Gaussian that lies above the cut; for the example of a 3σ cut, this is 0.135%. For pixels without any signal events, so that nearly all events are in the pedestal, the number of background hits per event is therefore $1.35 \cdot 10^{-3}$, if the Gaussian fit describes the pedestal well. In the data, where the average number of hits per pixel per event is $\sim 1\%$, the average number of background hits per pixel per event is found to be $1.34 \cdot 10^{-3}$, in agreement with expectation. The variation between tubes is less than $\frac{1}{2}\%$.

Figure 4.16: ADC spectrum from the same pixel as in figure 4.15, with beam synchronisation switched off, after 15 k events. A Gaussian is fit to the pedestal. P1, P2, P3 are amplitude, mean and width respectively. A line marks the 3σ cut, where σ is taken from the fit to the signal spectrum in figure 4.15. 47 events lie above this line.



Detector Noise and Other Non–Gaussian Beam–Independent Background

The origin of this type of background can be for example thermal electrons in the photo detectors, or photons from stray light entering the vacuum tank.

To estimate this background, pedestal runs are used. These are runs where the synchronisation between beam and readout is switched off. The number of hits in the pedestal run that are above the threshold cut, after subtraction of electronic noise, is taken as the background estimate. Scaled according to the number of events, the background is subtracted pixel by pixel from the photon counts in the beam data. For the pedestal run, the threshold cut is set at $\mu_p + m\sigma_b$, where μ_p is the mean of the pedestal in the pedestal run, while σ_b is the width of the pedestal found in the beam data for the same pixel.

The detector noise varies significantly between the different HPDs. For a 3σ cut, the values found for the run with filters range from $0.7 \cdot 10^{-3}$ hits per event per pixel for HPD 4, to $4.8 \cdot 10^{-3}$ hits per event per pixel for HPD 5; the average value is $1.8 \cdot 10^{-3}$. For the run without filters, the values are $0.6 \cdot 10^{-3}$ for HPD 4, $5.3 \cdot 10^{-3}$ for HPD 5, and $2.0 \cdot 10^{-3}$ for the average value.

The pedestal run chosen for each background estimate is the most recent one prior to the data run of interest. The procedure is repeated with different pedestal runs in order estimate a systematic error.

Beam Related Background

Anything that scatters Cherenkov light randomly can cause such background, such as imperfections in the mirror, dust, etc. Also, the small amount of back–scattered photo electrons that re–enter the Si waver can contribute to the beam–related background. As discussed in section 2.3.6, the probability that a photo electron produces two hits in the Si waver is $\sim 0.18 \cdot (d/(4h))^2$, where h is the height of the tube, d the diameter of the Si sensor, and 0.18 the backscattering fraction. For the case of the 61 pixel HPD, $h = 12\,\text{mm}$ and $d = 18\,\text{mm}$, therefore $\sim 2.5\%$ of all electrons enter the Si twice.

The beam related background is estimated as part of the fit to the Cherenkov angle distribution, as explained in section 4.5.3. The fit is performed after all the other corrections to the data have been applied.

The beam related background is found to be the smallest background contribution, with $\sim 1.8 \cdot 10^{-4}$ hits per event per pixel for the run with pyrex filters, and $3.2 \cdot 10^{-4}$ hits per event per event for the run without filters.

The Total Background

With a 3σ cut, for the data taken with pyrex filers, the total background is $3.3 \cdot 10^{-3}$ hits per event per pixel, or 0.18 hits per event per tube. This corresponds to 42% of all hits. Without pyrex filters, the total background is $3.7 \cdot 10^{-3}$ hits per event per pixel or 0.20 hits per event per tube, which constitutes 21% of all hits. How the background contributions vary for different threshold cuts is shown in tables 4.5 and 4.8, described later.

4.5.2 Signal Loss

Depending on the threshold cut, a certain fraction of signal events will fall below that cut and be lost. In order to estimate this fraction, the output of an HPD pixel is modelled with a function described below. The best fit of this function to the pulse height spectra is used to estimate the signal loss.

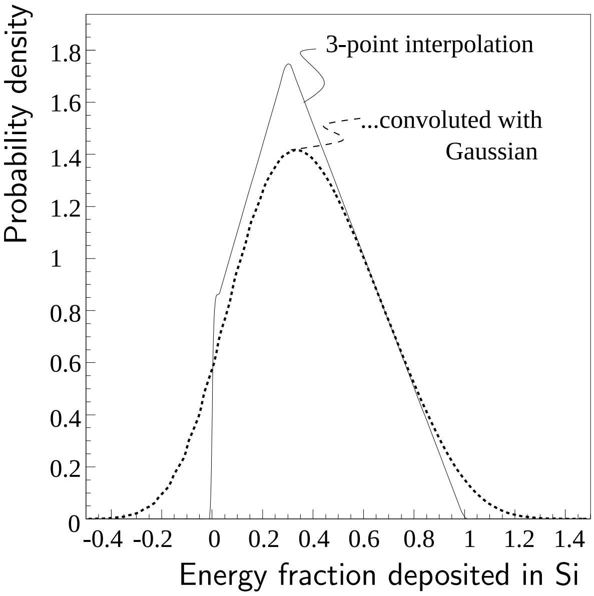
The number of signal events in data taken with Cherenkov photons is too small to fit the shape of the signal distribution. Therefore, the fits are performed on data taken with an LED as the light source.

From fits to the LED spectra, the fraction of single photo electron events lost below the cut is calculated for each pixel. Because of the small number of hits per event, nearly all hits in the Cherenkov data are single photo electrons. However, multiple photo electron events are significant in the LED data. Therefore it is necessary to extract from the LED data the single photo electron loss, rather than the total loss. This information is applied pixel by pixel to correct the photon counts in the Cherenkov data.

Fitting the LED Spectra

The function that is fitted to the LED data consists of a Gaussian for the pedestal and a Gaussian plus a backscattering component for the single and the double photo electron peak. The fit region is restricted from 4σ of the pedestal below its mean to 2σ of the double photo electron peak above the mean of that peak.

Figure 4.17: Three–point approximation to the distribution of the energy fraction deposited by back–scattered electrons in Si. The broken line is the distribution convoluted with a Gaussian, that takes into account electronics effects. For the convoluted curve, the x axis is to be interpreted as “electronics response”, normalised such that the pedestal is at $x = 0$ and the single photo electron peak at $x = 1$.



The back–scattering component: The largest part of the signal loss is due to back–scattering. When electrons are back–scattered, only a fraction of their energy is deposited in the silicon before they re–emerge and possibly enter again elsewhere. The probability that a 12 keV electron is back–scattered from a Si target is $\sim 18\%$ [Dar75]. The distribution of the energy fraction deposited in the Si is modelled from data for electrons with incident energy of 10 keV hitting an aluminium target [Dar75].

As established above, the effect of the re–entering of back–scattered electrons is small. It can therefore safely be ignored in signal loss estimates, as the signal loss itself is only a small correction to the photon count.

A three–point interpolation to the data in [Dar75] is convoluted with a Gaussian, to approximate the pulse height distribution of the back–scattered electrons, as shown in figure 4.17. The Gaussian takes into account the electronic noise and has the same width as the Gaussian that fits the single photo electron peak.

The shape of the backscattering component for the double photon peak is obtained by convoluting two single–photon back–scattering distributions. This is then convoluted with a Gaussian of the same width as the double photo electron peak.

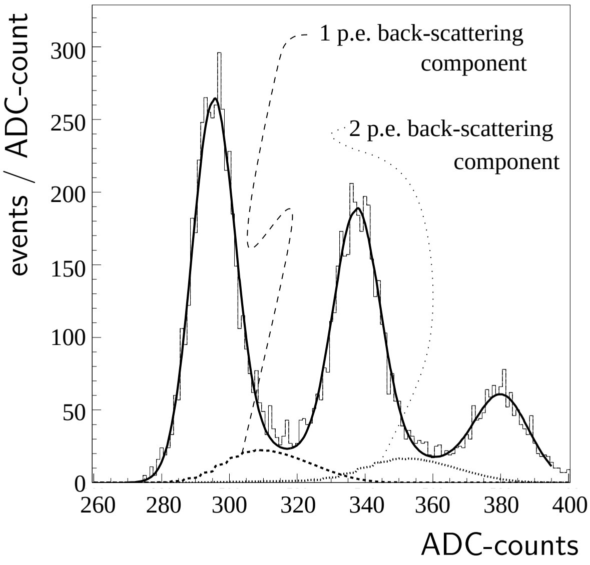
The backscattering components are constrained to 18% of the single photo electron contribution and

$$1-(1-0.18)^2 = 32.76\%$$

of the double photo electron contribution. With these constraints, the back–scattering components do not add any degrees of freedom to the fit, as they are completely determined by the Gaussians that fit the photo electron peaks.

The fit: Figure 4.18 shows a fitted LED spectrum with the backscattering contributions shown as broken lines. On average, for a 3σ cut, the loss of single photo electron events is 12%, where a signal loss of 11% is due to back–scattering alone. Variations in the signal loss estimate due to different assumptions for the back–scattering fraction, and the use of different LED runs, are considered for the systematic error estimates presented in tables 4.4 and 4.7 later in this chapter.

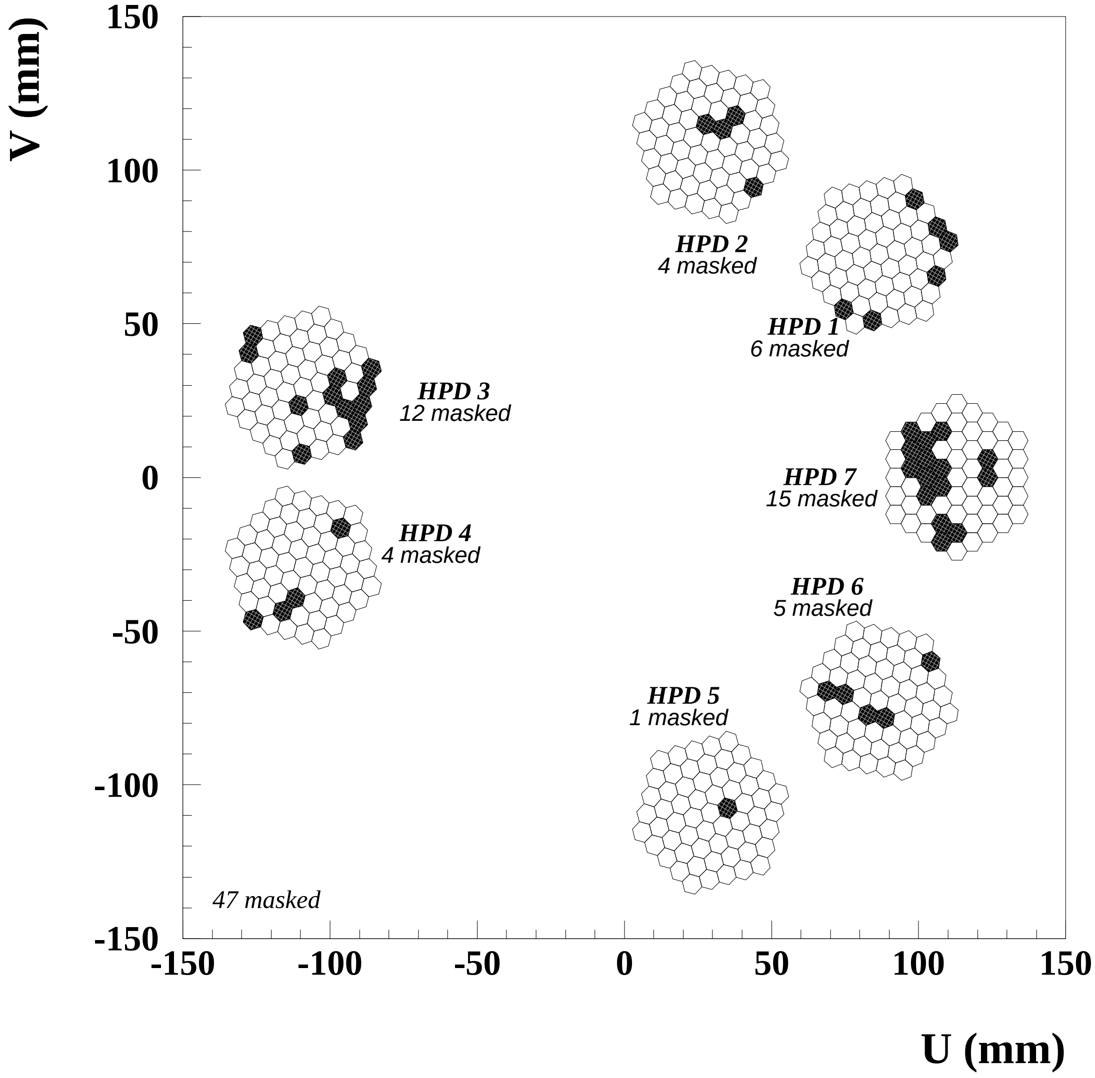
Figure 4.18: LED spectrum with fit. The broken lines are backscattering contributions



Masking Pixels

Certain pixels are either dead or very noisy, or simply the fit to the pedestal does not converge. Such pixels are ignored in the analysis and masked. The same pixels are masked in both data runs presented below, and in the Monte Carlo simulation. The masked pixels, and their position on the de-

Figure 4.19: Masked Pixels:
The HPDs are shown at their position on the detector plane as seen from the mirror. The HPDs themselves have been enlarged by a factor of 2, for clarity.



tector plane, are shown in figure 4.19[4].

Photon Counts

Figure 4.20 shows a colour–coded data display of all the seven HPDs on the detector plane, after the corrections for detector noise and electronic noise, and signal loss. The beam related background has not been subtracted. This, and other parameters, will be estimated from the fit to the Cherenkov angle distribution described below.

In this figure, the number of integrated hits, $\mathrm{N_{hits}}$, is shown for each tube. This counts single photo electron events as well multiple photo electrons events as one hit. The number of photo electrons, $\mathrm{N_{pe}}$, is therefore larger than $\mathrm{N_{hits}}$. As long as the number of photons per pixel per event is small, though, $\mathrm{N_{hits}} \approx \mathrm{N_{pe}}$ is a good approximation. In the data displayed, no pixel has more than $2.5 \cdot 10^{-2}$ hits per event.

4.5.3 Fitting the Cherenkov Angle Distribution

Fitting the Cherenkov angle distribution gives a Monte–Carlo independent estimate of $\mathrm{N_0}$ and an estimate of the resolution of the detector. It also allows the beam related background for the photo electron count to be determined.

The hit distribution on the photo detector plane, integrated over several thousand events, is fitted. A simple model for the Cherenkov angle distribution is used in the fit, essentially a Gaussian plus a flat background, with a few complications explained below. From the distribution in Cherenkov angle space, the expected number of photons in each pixel on the detector surface is calculated.

The fit returns the width of the Cherenkov angle distribution and its mean, the number of background events per pixel, and the number of photo electrons per ring from Cherenkov light that would have been observed for the case that the geometrical coverage were $\epsilon_A = 100\%$. As will become apparent below, the number of photo electrons per ring for $\epsilon_A = 100\%$ emerges “naturally” from this fit method, and allows to calculate $\mathrm{N_0}$ independent of the exact ring position and the complicated geometry in which the photo detectors are arranged on the detector plane.

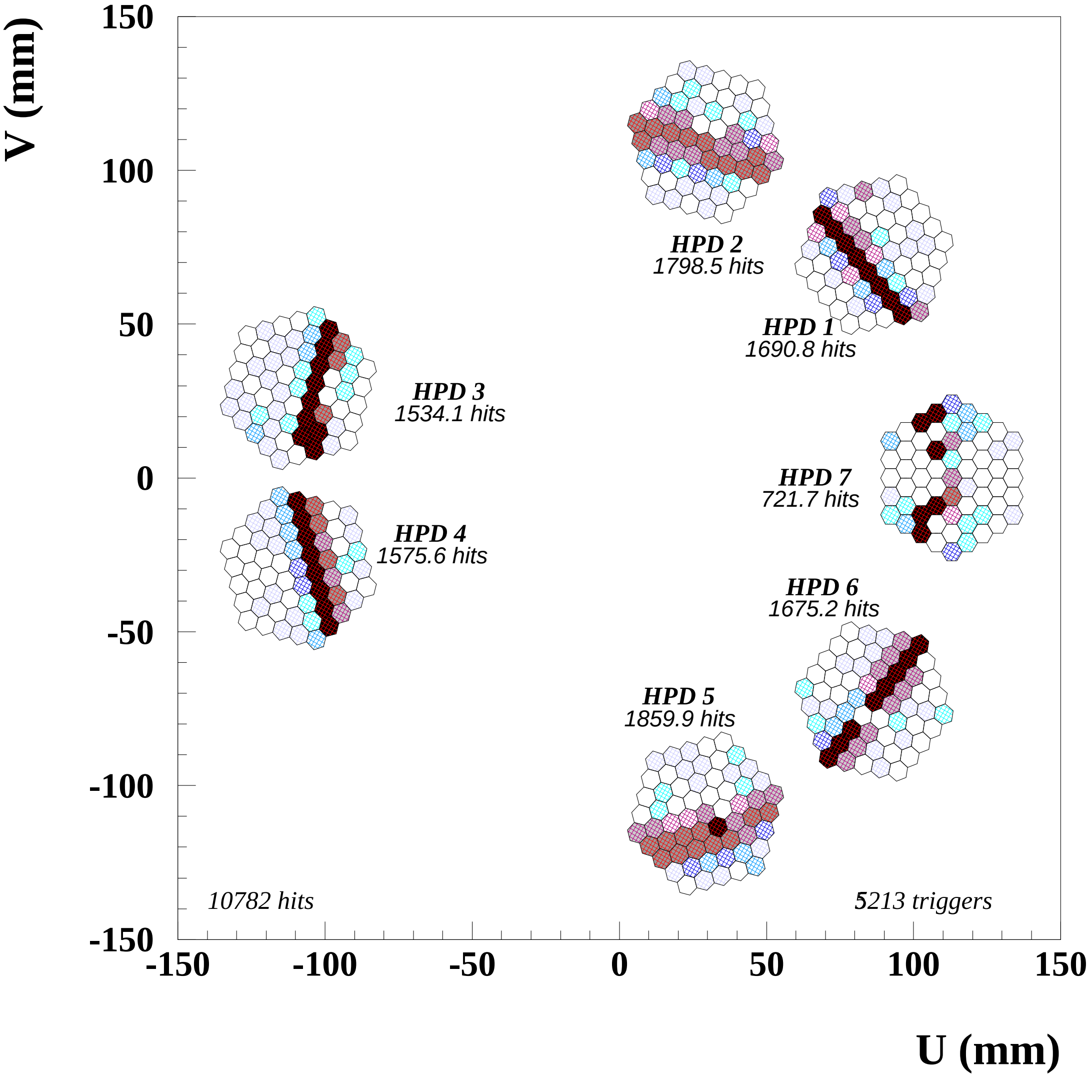
The main advantage of this procedure is, that it naturally solves several problems related to the detector geometry which in other methods, for example the fitting histogramed Cherenkov angle distributions, would have to be corrected for:

- Certain Cherenkov–angles are covered by a larger number of pixels than others and this distorts the Cherenkov angle distribution. This is naturally taken care of in a fit in pixel space.
- Masked pixels do not pose a problem in this fit procedure. They can simply be ignored.

---

[4] The program providing the graphical data display shown in figures 4.19 and 4.20 was kindly supplied by Brinick Simmons (Imperial College).

Figure 4.20: Data Display for run 1415, integrated over 5213 events. The Cherenkov light is generated by 100 GeV kaons passing through 1.8 m of CF_4 at 900 mbar and 20°C. All corrections to the data have been applied, except for the subtraction of beam–related background. The HPDs are displayed in their position on the detector plane as seen from the mirror. For clarity, the size of the HPDs has been increased by a factor of 2.



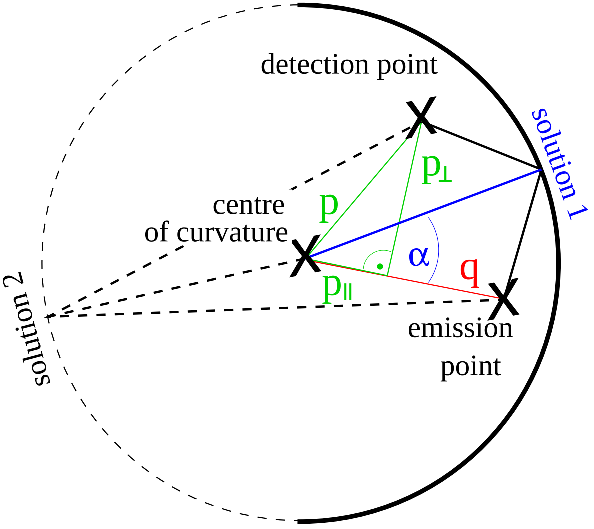
- The expected number of photo electrons per pixel is calculated by integrating the hit density per unit area on the detector plane over the entire pixel surface, instead of only using the value at the pixel centre. The fit result is therefore independent of the pixelisation, and does not include the pixelisation error.
- The ‘figure of merit’, N_0, is defined to be independent of the geometric coverage of the photo detectors. With a geometry as complicated as the one in the test–beam setup, where the geometric coverage crucially depends on the ring position and size, calculating N_0 from the photon counts is difficult. This fit method directly calculates N_0.

Cherenkov Angle Reconstruction and Alignment

Reconstruction: The Cherenkov angle can be reconstructed exactly from the emission point, the detection point, and the position of the centre of curvature (CC) of the spherical mirror. For symmetry reasons, the reflection point on the mirror surface must lie in the same plane as defined by the formerly mentioned three points. The geometry in that plane is illustrated in figure 4.21. The solution for the reflection point and the Cherenkov angle is derived in [FS98].[5] Taking the radius of curvature to be 1, and defining:

- $q \equiv$ distance of emission point from CC

Figure 4.21: Geometry of Cherenkov angle reconstruction.



- $p \equiv$ distance of detection point from CC
- $p_{\|} \equiv$ distance of detection point from CC, parallel to q.
- $p_{\perp} \equiv$ distance of detection point from CC, perpendicular to q.
- $\alpha \equiv$ the polar angle of the solution relative to q,

leads to the following equation for $x \equiv \sin\alpha$,

$$x^4 + ax^3 + bx^2 + cx + d = 0 \qquad (4.8)$$

with

$$a = \frac{-4q^2 p_{\perp}}{(2pq)^2}$$

---

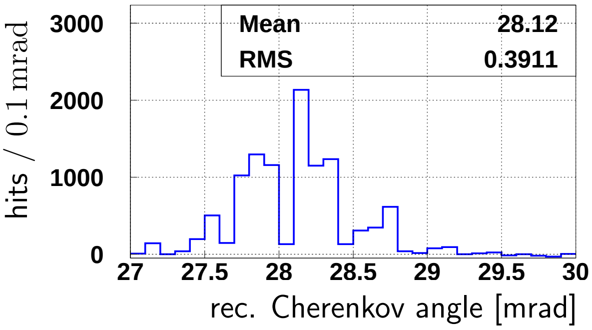
[5] An alternative approach to the one described here can be found in [YS93]

$$\begin{aligned} b &= \frac{p_\perp^2 + \left(p_\parallel + q\right)^2}{(2pq)^2 - 1} \\ c &= \frac{2qp_\perp \left(q - p_\parallel\right)}{(2pq)^2} \\ d &= \frac{p_\perp^2 \left(q^2 - 1\right)}{(2pq)^2} \end{aligned} \tag{4.9}$$

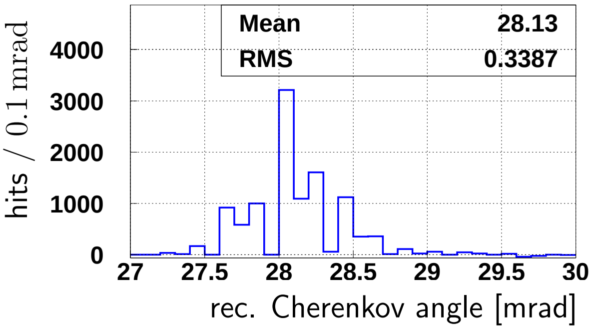
The quartic equation can be solved using standard techniques, given for example in [BS91] or [Lit50]. The CERN program library provides a routine that implements those techniques [Köl94]. The quartic equation returns 2 real solutions, one in the forward direction, and one in the backward direction; the solution in the forward direction is chosen. From $\sin\alpha$ it is easy to find the reflection point on the mirror. Given that, the Cherenkov angle can be calculated from the emission point and the direction of the particle.[6] Neither of the two are known exactly. The emission point is assumed to be at the centre of the radiator, which minimises the associated error. The particles are assumed to travel exactly along the z axis. The uncertainties associated with these assumptions are discussed in section 15 below.

Figure 4.22: Reconstructed Cherenkov angle before and after alignment (5213 events).

Before alignment:



After alignment:



Alignment: An uncertainty in the exact position of the centre of curvature of the mirror introduces an uncertainty in reconstructed the Cherenkov angle. The mirror was positioned by adjusting the micrometer screws until the Cherenkov ring on the event display was well centred. This method of adjustment is accurate to about one pixel size on the detector plane, which corresponds to $\frac{1}{2}$ mrad in Cherenkov angle and $\frac{1}{4}$ mrad in the mirror position. In order to determine the Cherenkov angle resolution precisely, off–line alignment is necessary. The detector is aligned by minimising the overall RMS of the Cherenkov angle distribution as a function of the mirror orientation (the radius of curvature and the distance between mirror and detector are not subject to the alignment procedure). The deviation of the mirror position from its nominal position is found to be 0.007 mrad

---

[6] The FORTRAN implementation of the Cherenkov angle reconstruction used in this thesis was kindly supplied by Roger Forty, CERN

in the $x - z$ plane and -0.15 mrad in the $y - z$ plane. This procedure is very simple, but as it is not a proper likelihood fit, the uncertainty in the result is difficult to estimate. Therefore, the Cherenkov angle resolution measurements are performed in a way that has the least possible requirements on alignment, as will be explained in section below.

The Shape of the Cherenkov Angle Distribution

The shape of the Cherenkov angle distribution, before pixelisation, is determined by three factors

1) chromatic aberration
2) beam divergence
3) emission point error

We will evaluate below the expected Cherenkov angle distribution for the test–beam run described in section 4.5.4, where 100 GeV kaons pass through 1.8 m of CF_4 at 20°C and 900 mbar. The photon energy range is limited by pyrex filters.

The chromatic aberration: The Cherenkov angle distribution due to chromatic aberration can be calculated from the photon detection efficiency as a function of photon energy given in figure 4.10. Using this together with the expression for the number of photo electrons generated per unit energy (equation 2.7), gives the number of photons detected per unit

Figure 4.23: $\frac{\mathrm{dN_{pe}}}{\mathrm{d}E}$ as a function of photon energy, for 100 GeV kaons passing through 1.8 m of CF_4 at 20°C and 900 mbar

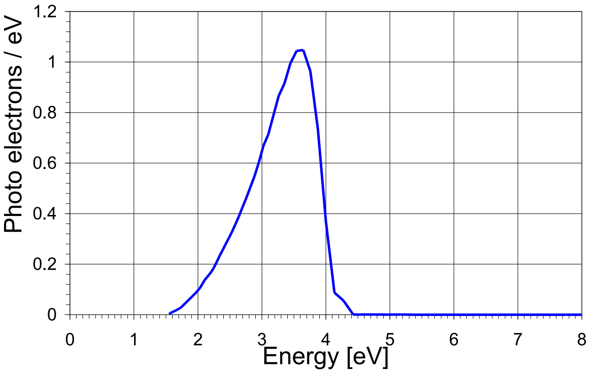
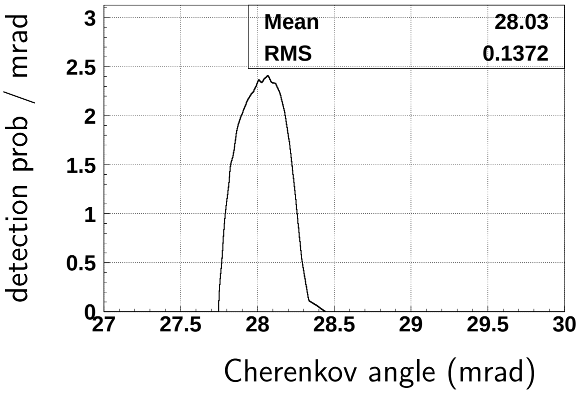


Figure 4.24: Expected contribution of the chromatic aberration to the Cherenkov angle distribution for the test–beam settings described in section 4.5.4.



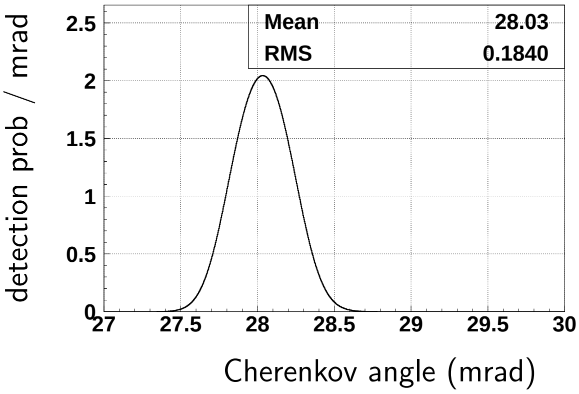


Figure 4.25: Expected Cherenkov angle distribution,including the effects of the chromatic aberration and the beam divergence, for the test–beam settings described in section 4.5.4.

energy. This is given in figure 4.23. For a given particle speed v and dispersion relation $n(E)$, there is a one–to–one relation between the photon energy and the Cherenkov angle, given by:

$$\cos\theta_C = \frac{c}{v \cdot n(E)}$$

Transforming the distribution shown in figure 4.23 into Cherenkov angle space, and normalising it, gives the Cherenkov angle distribution due to the chromatic error shown in figure 4.24. It has a mean of 28.03 mrad and an RMS of 0.14 mrad.

The beam divergence is measured to be 0.135 mrad in x and 0.120 mrad in y.[7] Taking this into account by convoluting the purely chromatic distribution in figure 4.25 with a Gaussian of width 0.125 mrad, gives the Cherenkov angle distribution shown figure 4.25, with a width of 0.18 mrad.

The emission point error arises due to the 18° tilt in the mirror relative to the beam direction, which makes exact knowledge of the emission point necessary to correctly calculate the Cherenkov angle from the hit position on the detector plane.

The emission point error can be separated into three contributions, one from each of the of the uncertainties in x, in y and in z–position. The size of each contribution depends on the azimuthal angle on the detector plane, and is therefore different for each HPD.

To calculate the emission point error, 10 M simulated events are generated with a fixed Cherenkov angle and beam direction, but random emission point within the allowed volume. The length of this volume in z is defined by the length of the radiator, which extends from $z = 0$ at the mirror to $z = -1.8$ m upstream. In x and y it is defined by the smaller of the two scintillation counters that trigger the read–out, which has dimensions of $20 \times 20\,\mathrm{mm}^2$.

The plots in figures 4.26, 4.27, 4.28 show the reconstructed Cherenkov angles versus the azimuthal angle on the detector plane.

The gray–scale indicates the number of detected photons for a given combination of Cherenkov angle and azimuthal angle. The numbers inside the plots are the RMS spreads caused by the given type of uncer-

---

[7] These numbers include the pixelisation error in the beam telescope of ~ 0.47 mrad which must be subtracted off in quadrature in order to obtain the true beam divergence.

Figure 4.26: Emission point errors due to ignorance of x-y position. The numbers inside the plots are the RMS spreads caused by the given type of uncertainty, for each HPD.

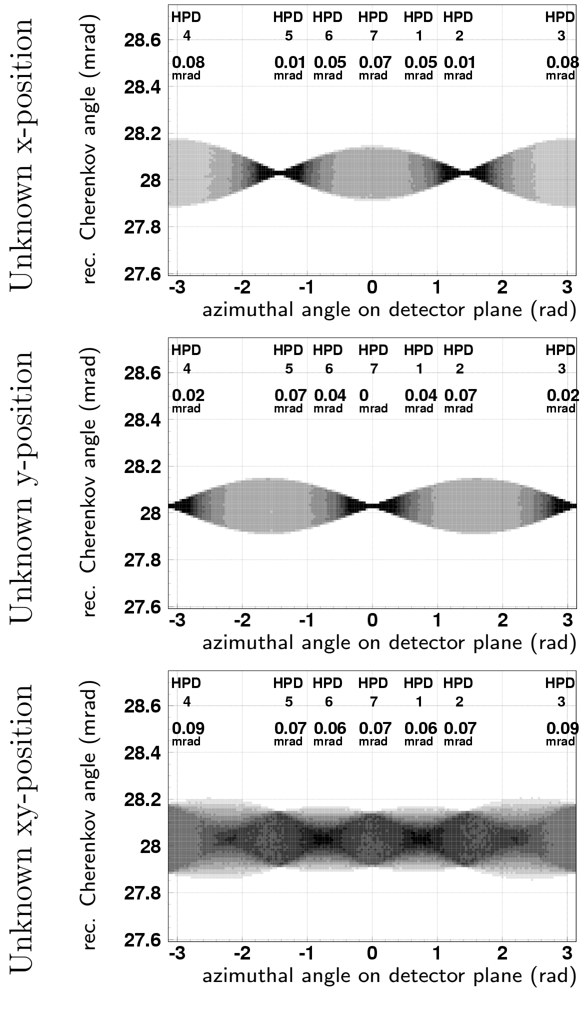


Figure 4.27: Emission point errors due to ignorance of z position.The numbers inside the plots are the RMS spreads due to the emission point error in z, for each HPD.

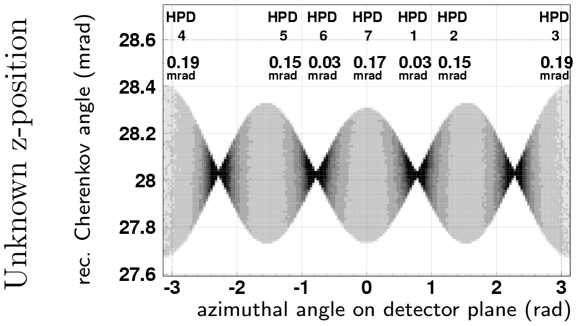
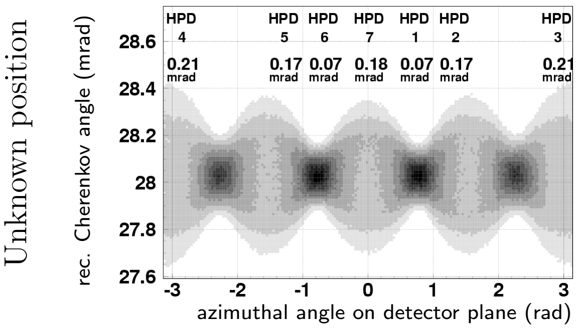


Figure 4.28: Total Emission Point Error. The numbers inside the plots are the RMS spreads due to the total emission point error, for each HPD.



tainty, for each HPD.

For figure 4.26, the Cherenkov angle is reconstructed approximating the x or the y position, or both, by the centre of the radiator. For figure 4.27 only the z position is approximated, and x and y are taken from Monte–Carlo truth. Figure 4.28 shows the total emission point error.

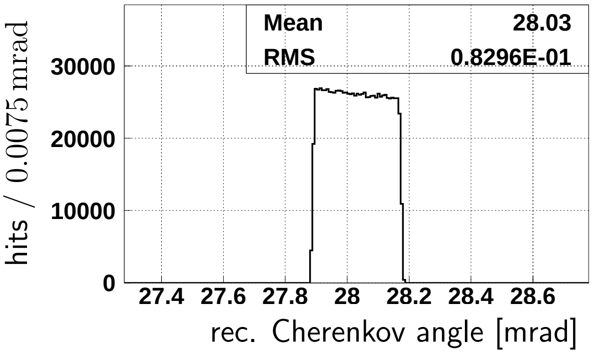
Figures 4.29, 4.31, and 4.32 show the Cherenkov angle distribution from 1 M Monte Carlo events, generated with fixed azimuthal angle $\phi = 165°$. This corresponds to the centre of HPD number 3. For a given azimuthal angle ϕ, the Cherenkov angle distribution due to the emission point uncertainty in only one co–ordinate is well described by a flat distribution between the minimum and the maximum reconstructed Cherenkov angle (in the following referred to as "box–function").

The RMS due to the combined x and y emission point error is about half the size of the combined error due to the chromatic error and beam divergence: 0.06 – 0.09 mrad compared with 0.18 mrad. Also, the difference in RMS between the HPDs due to the combined emission point error in x and y is very small compared to the over all width of the Cherenkov angle distribution. The expected Cherenkov angle distribution at HPD number 3, including the emission point error in x and y, is shown in figure 4.30. The graph is obtained by convoluting the distribution in figure 4.25 with two box–functions of 0.28 mrad and 0.07 mrad width.

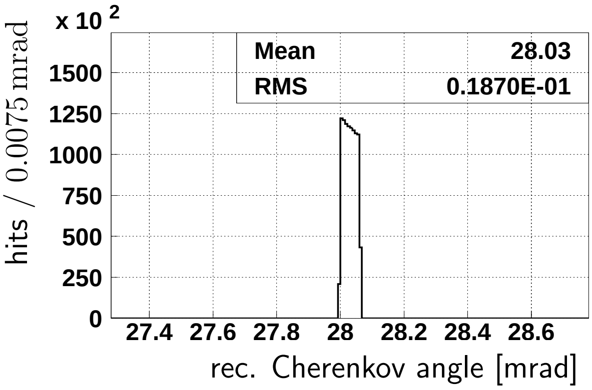
In contrast, figure 4.27 shows that the emission point error due to the uncertainty in z

Figure 4.29: Emission point error at the centre of HPD 3 ($\phi = 165°$) from 1 M simulated events

Error in x:



Error in y:



Error in x and y:

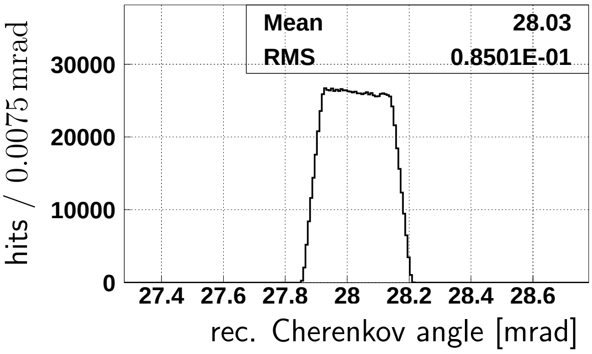


Figure 4.30: Expected Cherenkov angle distribution in HPD 3, including the effects of chromatic error, beam divergence, *and* emission point error in x–y

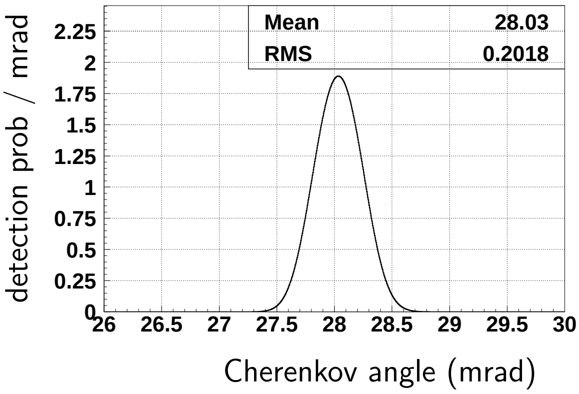


Figure 4.31: Emission point error in z at the centre of HPD 3 from 1 M MC events

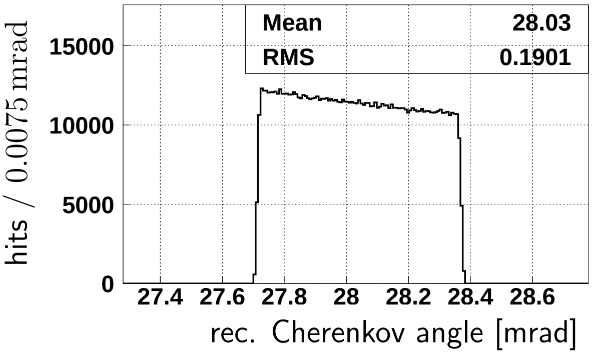


Figure 4.32: Total emission point error at the centre of HPD 3 from 1 M MC events

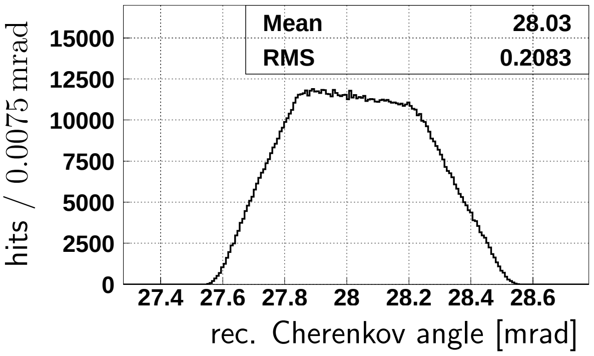
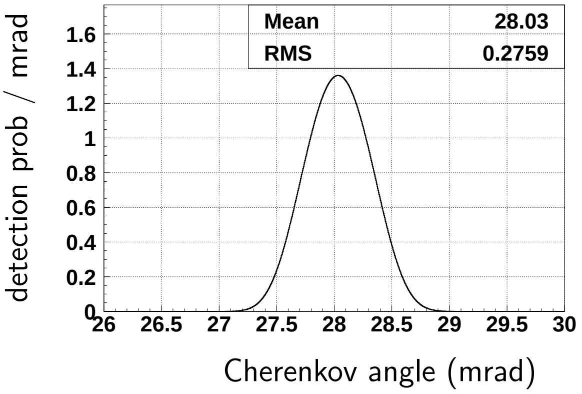


Figure 4.33: Expected Cherenkov angle distribution in HPD 3, with all error contributions.



varies significantly with azimuth. At HPD 1 and HPD 6, the emission point error in z is 0.03 mrad and therefore negligible, while it is 0.19 mrad at HPD 3 and HPD 4, which is larger than the combined error due to chromatic aberration and beam divergence. The total emission point error ranges from 0.07 mrad to 0.21 mrad, as shown in figure 4.28. Figure 4.33 shows the Cherenkov angle distribution in HPD 3, including all error contributions. The graph is obtained by convoluting the distribution in figure 4.25 with three box–functions of appropriate width.

The Fit Function in Cherenkov Angle Space

In order to maintain simplicity and to minimise the number of assumptions, the function chosen to fit the Cherenkov angle distribution is just one Gaussian plus one flat background, with two modifications:

- The mean Cherenkov angle is fitted separately for each tube.
- The emission point error in z is treated separately.

These modifications are now discussed in turn.

Different Mean Cherenkov Angles: A residual mis–alignment of the detector would show up as an increased width of the distribution. Because the azimuthal coverage of the tubes is very small, for each individual tube the effect of a moderate mis-alignment is well described by a shift in the mean of the distribution. Therefore, the mean Cherenkov angle is fitted independently for each of the seven HPDs.

For large mis–alignments, the width of the reconstructed Cherenkov angle distribution in individual tubes could also be affected. The difference in the fit result of the mean Cherenkov angle gives an estimate of the residual alignment. The largest difference between the means of two HPDs is 0.25 mrad. Geometrical considerations show, that this does indeed not have a significant effect on the width of the reconstructed Cherenkov angle distribution for any individual tube. Fitting the mean Cherenkov angle distribution independently for each tube therefore takes sufficient account of possible effects due to misalignment.

Emission point error: As shown above, the emission point error depends on the azimuthal position on the detector plane and is therefore different for each HPD. In every single co–ordinate, the reconstructed Cherenkov angle distribution due to the emission point error is well described by a box–function.

The combined emission point error in x and y, and the differences in that emission point error between the tubes, are small compared to the other error contributions. In the fit it can therefore be accommodated as part of the Gaussian.

The emission point error in z though, varies significantly, reaching values which make it the dominant contribution to the total uncertainty. Therefore, the differences between the tubes cannot be ignored. For tubes with a large emission point error in z, the Gaussian approximation to the Cherenkov angle distribution is bound to be poor. The emission point error is therefore taken into account separately when calculating the hit distribution on the detector surface.

The Gaussian is used only to describe the Cherenkov angle distribution without the emission point error in z. To describe the Cherenkov reconstructed angle distribution at a given azimuthal angle on the detector plane, this Gaussian is convoluted with a box function that describes the emission point error in z at that azimuthal angle. This box function is not part of the fit. The width of the box function can be calculated by comparing the reconstructed Cherenkov angles for the cases that the emission point is assumed to be at the beginning or at the end of the radiator tube.

The normalised probability density for a photon hit as a function of the Cherenkov angle θ and the azimuthal angle ϕ on the detector plane, is given by:

$$P(\theta, \phi) = \frac{1}{2\pi} g(\theta) * b(\theta, \Delta\theta(\phi)), \qquad (4.10)$$

where $g(\theta)$ is a normalised Gaussian:

$$g(\theta) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{(\theta-\mu)^2}{2\sigma^2}}. \qquad (4.11)$$

and "$*$" the convolution operator. $b(\theta, \Delta\theta(\phi))$ represents the box function of width $\Delta\theta$:

$$b(\theta, \Delta\theta) = \left\{ \begin{array}{ll} \frac{1}{\Delta\theta} & \text{if } |\theta| < \frac{\Delta\theta}{2} \\ 0 & \text{otherwise} \end{array} \right\} \qquad (4.12)$$

where $\Delta\theta$ depends on ϕ as shown in figure 4.27.

Only μ and σ are parameters of the fit. As the result of the fit, we expect a Gaussian approximation to the distribution shown in figure 4.30, with a mean of 28.03 mrad and a width of 0.202 mrad.

The Fit Function on the Detector Plane

The density of detected Cherenkov photons per unit area for 100% geometric coverage, as a function of position on the detector plane, is given by:

$$\mathrm{N}_{\mathrm{rg\,pe}} \cdot P(r, \phi) = \mathrm{N}_{\mathrm{rg\,pe}} \cdot P(\theta, \phi) \frac{\mathsf{d}\theta}{r\mathsf{d}r}. \qquad (4.13)$$

$P(r, \phi)$ is the normalised probability density per unit area, to find a photon in the area element $r\,\mathsf{d}\phi\,\mathsf{d}r$ at a distance r from the centre of the detector plane at an azimuthal angle ϕ. $P(\theta, \phi)$ is the convolution of a Gaussian and a box–function defined in equation 4.10. $\mathrm{N}_{\mathrm{rg\,pe}}$ is the total number of photo electrons if the geometric coverage were 100%.

The fit also allows for a flat background, B, so the total number of photo electrons per unit area is given by

$$\frac{dN_{\mathrm{sg+bg}}}{dA}(r, \phi) = \mathrm{N}_{\mathrm{rg\,pe}} \cdot P(r, \phi) + B. \qquad (4.14)$$

From this, the expected number of photo electrons in each pixel is calculated by integrating $\frac{\mathrm{d}N_{\mathrm{sg+bg}}}{\mathrm{d}A}(r,\phi)$ over the pixel surface:

$$N_{\text{pixel fit}} = \int\limits_{\text{pixel surface}} \left(\mathrm{N}_{\mathrm{rg\,pe}} \cdot P(r,\phi) + B\right)\, r\,\mathrm{d}\phi\,\mathrm{d}r \tag{4.15}$$

This integral is evaluated numerically.

The fit parameters are seven mean Cherenkov angles μ_i, one for each HPD, a single width, σ, and the flat background B. Another important result from the fit is the number of photo electrons per ring for 100% geometric efficiency, $\mathrm{N}_{\mathrm{rg\,pe}}$. This not a free parameter in the fit, but determined from the normalisation condition that the total number of photo electrons found in the data be the same as the number of detected photo electrons calculated from the function.

Note that, as before, we approximate $\mathrm{N}_{\mathrm{rg\,hits}} \approx \mathrm{N}_{\mathrm{rg\,pe}}$ in the data analysis.

The fit is performed by minimising:

$$\chi^2 = \sum_{\text{all pixels}} \frac{\left(N_{i\text{ pixel fit}} - N_{i\text{ data}}\right)^2}{\sigma^2_{i\text{ data}}} \tag{4.16}$$

where $N_{i\text{ data}}$ is the number of hits in pixel i found in the data after all correction described in section 4.5 are applied, and $N_{i\text{ pixel fit}}$ is the number of hits in pixel i calculated from the fit parameters, using equation 4.15. The error $\sigma_{i\text{ data}}$ on the photon count in each pixel is calculated from the statistical errors of the different contributions to the photon count, including the background estimates and the signal loss estimates.

Testing the Fit on Monte Carlo Data

As the fit function is a rather simple approximation to the expected distribution of the data, the usual methods to evaluate the reliability of a fit, such as χ^2 probabilities, are not applicable. Therefore the fit has been tested in a series of Monte Carlo experiments: 500 sets of Monte Carlo data, each with 5 k events, have been generated. Table 4.2 shows the mean and the RMS of the 500 fit results for the different parameters. The fit results are compared with the output from the Monte Carlo simulation for 0.5 M generated events, which has a statistical error on the number of photo electrons per ring of less than 0.04%.

The events are generated for 100 GeV kaons passing through 180 cm of $\mathrm{C\,F_4}$ at 900 mbar and 20°C. The beam divergence is taken to be 0.125 mrad in x and y. For the data / Monte Carlo comparison later in this chapter, the Monte Carlo events are generated using the measured beam divergence with slightly different values in x and y as given in table 4.1. For the purpose of comparing the fit results with the Monte Carlo input however, it is useful to generate the events with the same beam divergence in x and y.

The beam profile is assumed to be flat within the $20 \times 20\,\mathrm{mm}^2$ defining the beam dimensions. The centre of curvature of the mirror is shifted relative to its nominal position by -0.007 mrad in x and -0.15 mrad in y, as found in the data. In addition to

Table 4.2: Values obtained from 0.5 M Monte Carlo generated events, compared with the mean fit results from fitting 500 simulated datasets with 5 k events each. N_0 is calculated from the number of hits, not photo electrons, in both, data and Monte Carlo.

	Mean of 0.5 M MC events (not from fit)	mean and RMS of fit results
$N_0[\mathrm{cm}^{-1}]$	112.3	112.8 ±1.9
$N_{\mathrm{rg\,hits}}$	15.88	15.96 ±0.27
μ(mrad)	28.03	28.03 ±0.01
σ(mrad)	0.202	0.198 ±0.006
Background / pixel	$3 \cdot 10^{-3}$	$(2.97 \pm 0.07) \cdot 10^{-3}$
hits/HPD	0.309	0.311 ±0.005

the Cherenkov photons, a flat background is generated with $3 \cdot 10^{-3}$ hits per event per pixel, which corresponds approximately to the total background found in the data.

The width of the Cherenkov angle distribution, excluding the emission point error in z, is 0.202 mrad.

For 0.5 M simulated events, the mean number of photo electrons per ring for 100% geometric coverage is $N_{\mathrm{rg\,ph}} = 16.02$; the value for N_0 calculated from that is $N_0 = 113.3\,\mathrm{cm}^{-1}$. This is slightly smaller than the value given in equation 4.2 in the beginning of this section, because it includes the losses of $\sim 2\%$ due to photons hitting the walls of the radiator tube.

In the data, it is the number of hits per pixel which is counted, rather than the number of photo electrons. The mean number of hit pixels per ring for the same set of simulated events is $N_{\mathrm{rg\,hits}} = 15.88$. Calculating N_0 from $N_{\mathrm{rg\,hits}}$, we find $N_0 = 112.3\,\mathrm{cm}^{-1}$. In the following, all values for N_0 are calculated from $N_{\mathrm{rg\,hits}}$, but as seen here, the difference to calculating it from the N_{pe} is is less than 1%.

In table 4.2 the results after fitting 500 samples each with 5 k events are compared with the input values. In the fit to the data, seven independent values for the means of the Cherenkov angle distribution are fitted. The entry μ in the table refers to the mean of those seven means. The background is given in terms of hits per pixel per event. The last entry, hits/HPD, is the average number of signal hits per HPD per event.

The results show very good agreement between the results calculated by generating 0.5 M Monte Carlo events, and those obtained from the fits. Still, as to be expected given the simplified fit function, the results are not unbiased: For example the statistical error on the mean of the fit results for N_0 is $1.9\,\mathrm{cm}^{-1}/\sqrt{500} = 0.08\,\mathrm{cm}^{-1}$, which is about 6σ away from the correct value. Importantly though, this bias is much smaller than the RMS spread in the fit results, so it does not effect the validity of the fit. The RMS spread on the fit results for 5 k events for each parameter is of the order of 3% or less.

4.5.4 Results

Photon Counts and Resolution for 100 GeV Kaons, using Pyrex filters

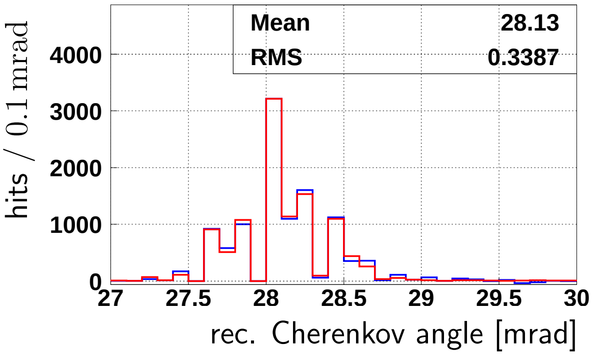
The data were taken with the following configuration:

Pressure	900 mbar
Temperature	20°C
Beam momentum	100 GeV
Particle type	Kaons
Filter	Pyrex
Nº events	5213

The data are corrected for noise and signal loss as described in section 4.5. The fit described in section 4.5.3 is applied to the data and the results are given in table 4.3. The errors given on the data are the systematic error estimate. Also the mean and RMS from fits to 500 MC experiments are given. The fit, the systematic error and the Monte Carlo results, are discussed in detail below.

Figure 4.34 shows the Cherenkov angle distribution of the data and, superimposed, the fit result. While the fit is performed on the detector plane, the plot shows the result in Cherenkov angle space. Because some angles are represented by more pixels on the detector plane than others, the Cherenkov angle distribution is distorted, which explains the small difference between the mean in figure 4.34 and that in table 4.3.

Figure 4.34: Reconstructed Cherenkov angles, shown in blue, with effect of pixelisation clearly visible. Superimposed in red, the fit to the data.



The results from 500 Monte Carlo experiments with the same number of events as the data are also given in table 4.3. The difference in the Monte Carlo program compared to the one used for table 4.2 is that the events are generated using the measured beam divergence, which differs slightly in x and y.

The Monte Carlo events are generated with a mean of $3 \cdot 10^{-3}$ background hits per pixel per event, which corresponds approximately to the total background in the data. The fits to the Monte Carlo generated events return $(2.97 \pm 0.07) \cdot 10^{-3}$ background hits per pixel per event.

Most background contributions in the data though, are subtracted before the fit is performed, leaving only the small beam related background which is given in the table. However, the statistical uncertainty as a consequence of the total background in the data is similar to that in the Monte

Table 4.3: Run with pyrex filter: The results of the fit to the data are compared with the with mean values from 0.5 M simulated events, and the results from 500 fits to Monte Carlo generated data sets containing 5213 events each. The value for the mean Cherenkov angle is the average of the fit result for each HPD.

	Mean of 0.5 M MC events (not from fit)	MC: mean and RMS of fit results	data fit$\pm\sigma_{\mathsf{sys}}$
$N_0[\mathrm{cm}^{-1}]$	112.3	112.5 ±1.7	107.6 ±5.5
$N_{\mathrm{rg\,hits}}$	15.88	15.92 ±0.24	15.26 ±0.6
μ [mrad]	28.035	28.045 ±0.006	28.07 ±0.01
σ [mrad]	0.199 – 0.208	0.198 ±0.006	0.200 ±0.004
Beam-related BG/pixel			$(1.8 \pm 1.4)10^{-4}$
hits/HPD	0.309	0.299 ±0.005	0.284 ±0.015

Carlo generated data sets. A summary of all background contributions found in the data is given in table 4.5.

Systematic Error: In order to estimate a systematic error on the data, the photon counting procedure has been repeated with several different settings compared to the default procedure. The results are given in table 4.4. The systematic error on the following contributions to the final corrected photon count are estimated:

- **Detector noise and other not beam–related background:** The error on this is estimated by using different pedestal runs.
- **Electronic noise:** The pedestal peak in the data is fitted over different intervals.
- **Beam related background:** The fit procedure is altered, allowing a different level of beam related background for each tube.
- **Signal loss** Different LED runs are used for the signal loss estimate, and different assumption regarding the backscatter fraction.
- **Signal and Background correction:** The photon counting is repeated using a different threshold cut. A 4σ cut is chosen for the systematic error estimate, as it still provides a reasonable background level and a reasonable signal loss. The corrected counts for other threshold cuts are given in table 4.5 and also show good agreement.

Adding in quadrature the difference of table entry “default” to all variations considered (excluding d and e), gives conservative estimates for the systematic errors: 5.5 cm^{-1} for N_0, 0.6 for $N_{\mathrm{rg\,hits}}$, 0.01 mrad for the Cherenkov angle mean and 0.004 mrad for the width of the Cherenkov angle distribution.

Table 4.4: Results for different configurations (3σ cut except h)

	configuration	N_0 [cm^{-1}]	$\mathrm{N}_{\mathrm{rg\,hits}}$	μ [mrad]	σ [mrad]	$\frac{\mathrm{bg}}{\mathrm{pix}} \cdot 10^3$	hits/ tube
default		108	15.3	28.075	0.200	0.18	0.284
a	different pedestal run for detector noise	105	14.8	28.066	0.199	0.05	0.275
b	different LED run (1368) for signal loss	110	15.6	28.074	0.200	0.19	0.300
c	assume 20% backscattering	109	15.5	28.075	0.200	0.19	0.288
d	assume 16% backscattering	106	15.1	28.075	0.200	0.18	0.280
e	fit pedestal up to 1.2sigma	108	15.3	28.076	0.202	0.14	0.284
f	fit pedestal up 2.0sigma	108	15.3	28.077	0.202	0.21	0.284
g	allow different bg for each HPD in the fit	105	14.9	28.076	0.199	0.16	0.277
h	4σ cut instead of 3σ	110	15.5	28.075	0.197	0.20	0.289
σ_{sys}		5.5	0.6	0.009	0.004	0.14	0.015

The dominant effect comes from using a different LED run (entry b), which implies that the signal loss estimate is the least certain part of the analysis.

Data / Monte Carlo Comparison: The agreement between data and Monte Carlo is good. The value for N_0 found in the data is about 4% below the expectation from Monte Carlo, which is comfortably within the estimated error margin. The value for the number of photo electrons found per tube is 8% below the Monte Carlo expectation. This value is very sensitive to the exact ring position and size, as it includes all geometric effects, so a larger discrepancy might be expected; nevertheless, it also lies within the estimated errors.

The Cherenkov angle resolution found in the data agrees very well the Monte Carlo prediction. The σ from the Gaussian fit also agrees very well with the RMS of the analytically derived Cherenkov angle distribution shown in figure 4.30 .

Signal and Background for Different Cuts: Table 4.5 shows the various background contributions and the signal loss estimates for different threshold cuts, together with the corrected number of hits/HPD. The corrected photon counts differ relative to the default settings by less than 6%, even for a 2σ cut with 82% background. This confirms the validity of the methods used for the signal loss and background estimates.

Table 4.5: Photo electrons counts, signal loss, and background, for different threshold cuts, with pyrex filter. The background is given relative to the uncorrected photon count. The signal loss is given relative to the corrected photon count.

cut:	2σ	3σ	4σ	5σ
raw hits/tube	**1.61**	**0.430**	**0.303**	**0.240**
detector noise	6%	23%	19%	13%
electronic noise	76%	17%	0.6%	0%
beam related backg.	0%	2.3%	3.6%	4%
total bg	**81%**	**42%**	**23%**	**18%**
loss	**7.8%**	**12%**	**18%**	**31%**
corrected hits/tube	**0.301**	**0.284**	**0.289**	**0.296**

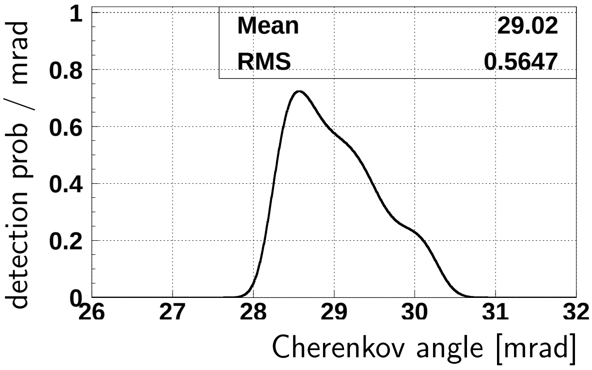
Photon Counts and Resolution for 150 GeV Pions, without Pyrex Filters

The same analysis as above has been repeated for data taken without pyrex filters in front of the photo detectors. The setup is summarised in the following table:

Pressure	900 mbar
Temperature	20°C
Beam momentum	150 GeV
Particle type	Pions
Filter	none
№ events	4461

Due to the absence of the pyrex filters, two complications enter the analysis. Firstly, the Cherenkov angle distribution is dominated by the chromatic error, which has a large skew. This is shown in figures 4.35 and 4.36 for the example of HPD 3. It can be seen that the Gaussian approximation is poor. Secondly, the quantum efficiency of the tubes has not been measured over the entire photon–energy range of interest. For the Monte Carlo simulation, it is assumed that the quantum efficiency stays constant between the last measurement point at 6.2 eV up to the cut–off from the glass window at 7.3 eV, as shown in figure 4.7.

Figure 4.35: Expected Cherenkov angle distribution in HPD 3, without filter, for a beam divergence of 0.125 mrad. The plot includes all contributions to the shape of the distribution, except for the emission point error in z.



The results of applying the fit procedure described in sections 4.5.3 and 4.5.4 are summarised in table 4.6. Figure 4.37 shows the

Figure 4.36: Expected Cherenkov angle distribution in HPD 3, without filter, for a beam divergence of 0.125 mrad. The plot includes all contributions to the shape of the distribution.

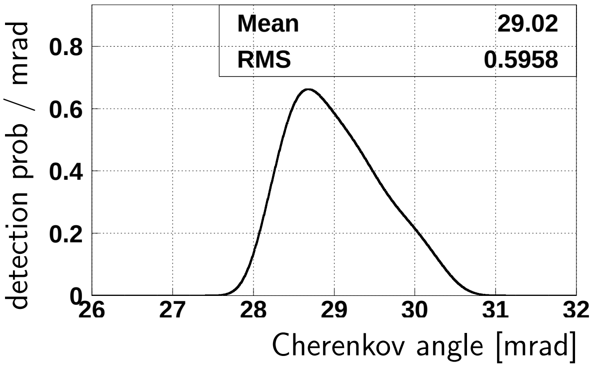
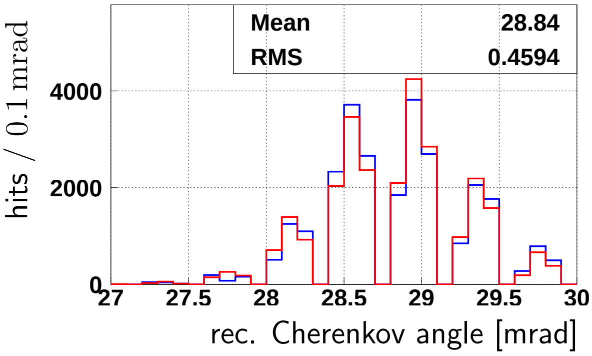


Figure 4.37: Reconstructed Cherenkov angles, shown in blue blue, from data without filter. Superimposed in red, the fit to the data. The effect of the skew in the data distribution is clearly visible



data with the fit superimposed. Comparing this with figure 4.34 (with filter), the effect of the skew is apparent. However, given the simple fit function, the agreement between the fit and the data is good.

Data / Monte Carlo Comparison: Table 4.6 gives the results from the fit to the data as well as the Monte Carlo output from 0.5 M events and the result of fitting 500 Monte Carlo generated event sets.

Despite the assumption that the Cherenkov angle distribution can be described with a simple Gaussian, the fits to the Monte Carlo generated events agree well with the results obtained from generating 0.5 M events. On the other hand, the agreement between data and Monte Carlo is less good. The value for N_0 found in the data is 11% below the MC values, which corresponds to a discrepancy of 3.4σ. The Cherenkov angle distribution in the data is also significantly narrower, and the mean Cherenkov angle is smaller than in the Monte Carlo simulation.

These discrepancies could be explained if the quantum efficiency in the unmeasured range dropped more quickly than has been assumed.

Systematic Errors and Different Cuts: The systematic error is estimated in the same way as for the case with filter; the results are presented in table 4.7. The results in table 4.8 confirm the signal loss and background correction results by comparing the corrected photon counts for different threshold cuts. The agreement is better

Table 4.6: Run without pyrex filter: The results of the fit to the data are compared with the with mean values from 0.5 M simulated events, and the results from 500 fits to Monte Carlo generated data sets containing 4461 events each. The value for the mean Cherenkov angle is the average of the fit result for each HPD. The 500 data sets are generated with an average background of $3 \cdot 10^{-3}$ per pixel per event. The fit returns $(2.5 \pm 0.2) \cdot 10^{-3}$. The table shows only the beam–related background.

	Mean of 0.5 M MC events (not from fit)	MC: mean and RMS of fit results	data fit$\pm\sigma_{\mathsf{sys}}$
N_0 [cm^{-1}]	353.4	359.3 ±7.2	316.5 ±8.0
$N_{\mathrm{rg\,hits}}$	53.4	54.22 ±1.16	47.50 ±1.3
μ(mrad)	28.984	28.961 ±0.033	28.888 ±0.003
σ(mrad)	$0.56 - 0.58$	0.535 ±0.018	0.444 ±0.01
beam-related BG/pixel			$(3.1 \pm 6.0)10^{-4}$
hits/HPD	0.895	0.923 ±0.016	0.843 ±0.021

Table 4.7: Results for different configurations, without Pyrex filter (3σ cut except h)

configuration		N_0 [cm^{-1}]	$N_{\mathrm{rg\,hits}}$	μ [mrad]	σ [mrad]	$\frac{\mathsf{bg}}{\mathsf{pix}} \cdot 10^3$	hits/ tube
default		317	47.5	28.888	0.444	0.31	0.843
a	different pedestal run for detector noise	320	48.2	28.886	0.446	0.44	0.853
b	different LED run (1368) for signal loss	321	48.3	28.889	0.443	0.29	0.856
c	assume 20% backscattering	321	48.2	28.888	0.444	0.32	0.854
d	assume 16% backscattering	312	46.9	28.888	0.444	0.30	0.832
e	fit pedestal up to 1.2sigma	316	47.5	28.888	0.442	0.33	0.843
f	fit pedestal up 2.0sigma	315	47.3	28.886	0.442	0.43	0.839
g	allow different bg for each HPD in the fit	314	47.2	28.889	0.443	0.88	0.837
h	4σ cut instead of 3σ	316	47.4	28.889	0.434	0.20	0.842
σ_{sys}		8.0	1.3	0.003	0.010	0.60	0.021

Table 4.8: Photo electrons counts, signal loss, and background, for different threshold cuts, without pyrex filter. The background is given relative to the uncorrected photon count. The signal loss is given relative to the corrected photon count.

cut:	2σ	3σ	4σ	5σ
raw hits/tube	**2.12**	**0.939**	**0.758**	**0.623**
detector noise	6%	12%	8%	5%
electronic noise	56%	8%	0.2%	0%
beam related backg.	0%	1%	1%	1%
total bg	**63%**	**21%**	**10%**	**6%**
loss	**8.2%**	**13%**	**19%**	**32%**
corrected hits/tube	**0.853**	**0.843**	**0.842**	**0.864**

than 2.5% over the whole range from 2σ to 5σ.

4.6 Conclusions

The full–scale RICH prototype equipped with HPDs has been tested in two different configurations, with and without pyrex filter. The results are corrected for signal loss and background contamination. The good agreement of the corrected photo electron counts for different threshold cuts confirms the correction method. The systematic error has been evaluated by repeating the analysis in eight different ways.

The fit method presented allows an extraction of the RICH performance parameters independent of the complicated geometry in which the photodetectors are arranged on the detector plane. The validity of the fit method has been demonstrated by performing several hundred Monte Carlo experiments, and comparing the fit results with the results from simulating large numbers of events.

The comparison between the data and the detailed Monte Carlo simulation shows very good agreement for the first data set presented, where the HPDs are equipped with pyrex filters. The width of the Cherenkov angle distribution deduced from the fit also agrees very well with distribution derived analytically.

The agreement for the data taken without filter is less good, and a possible explanation for this has been given. For the first data set, the quantum efficiency of the combined filter and HPD is well known, while for the second data set, some part of the quantum efficiency curve had to be extrapolated. It is possible that this extrapolation is not correct, which could be responsible for observed discrepancies.

Overall, the success of the fit method and the data/Monte Carlo comparison show that the test–beam data are well understood. This gives confidence that the simulation of the full RICH system within LHCb are reliable.

Chapter 5

MaPMTs as Photo Detectors for the LHCb RICH

5.1 Introduction

In this chapter, the suitability of the Hamamatsu M64 MaPMT as a photodetector for the LHCb RICH is investigated. Multi-anode Photo Multiplier tubes (MaPMTs) provide a high photon detection efficiency, a good spatial resolution and, compared with HPDs, a large signal, which simplifies the readout.

Single MaPMTs, and a cluster of 3 × 3 MaPMTs in a RICH prototype, have been tested in a test–beam at the CERN SPS facility. The MaPMTs have been equipped with lenses in front of their entrance windows. This is to reduce the losses of photons in the dead areas at the edges of the MaPMT, which corresponds to about 50% of the total area. LHC-speed, pipelined electronics have been used to read out MaPMTs for the first time.

The main aims of the tests were to

- evaluate the performance of the MaPMTs with and without lenses,
- test the performance of the MaPMTs with fast read-out electronics,
- demonstrate that the MaPMT is a viable photodetector solution for the LHCb RICH system.

In the following, the setup and results of the 3 × 3 cluster test will be described in detail. A short account on the tests with single MaPMTs, read out with CAMAC electronics, is given at the end of the chapter.

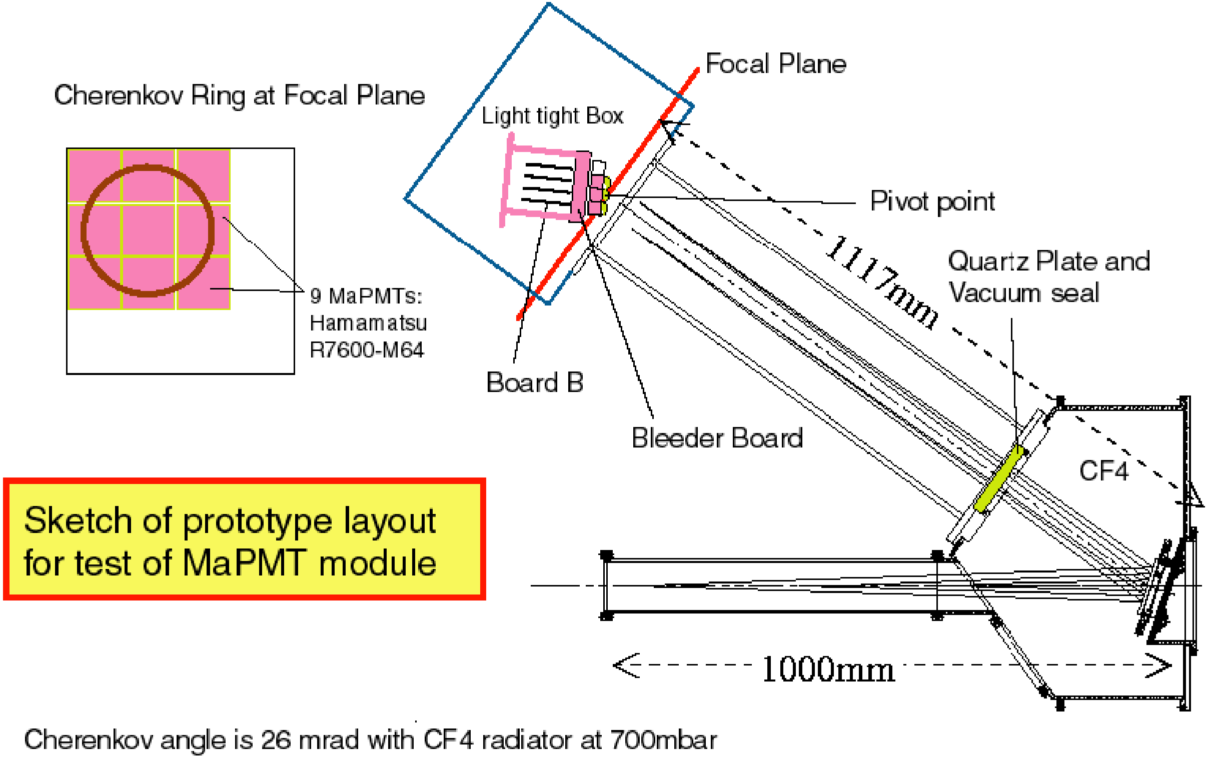
5.2 Experimental Setup

5.2.1 Overview

The test–beam setup used for the MaPMT tests resembles closely that described in chapter 4 for the tests of the full-scale RICH 2 prototype. The major alterations are:

- MaPMTs are used instead of HPDs.

Figure 5.1: Testbeam setup for cluster test



Mounted MaPMT array with electronics

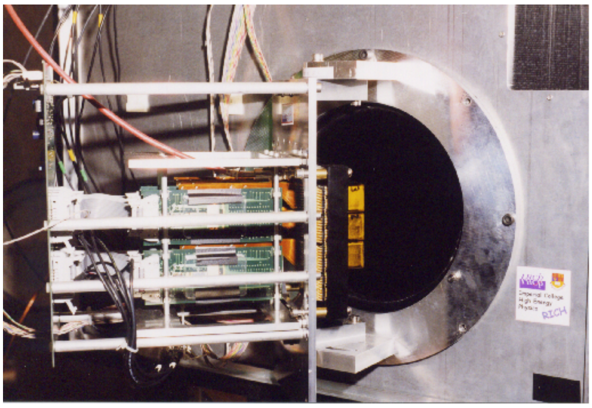
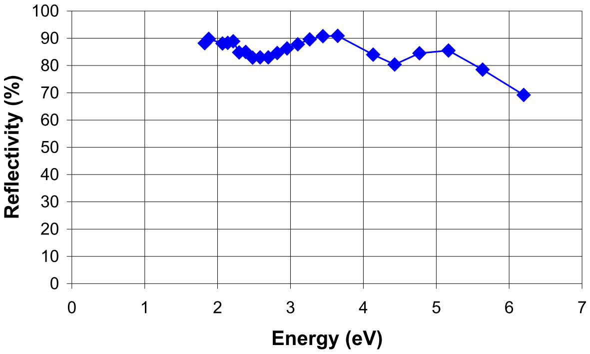


Figure 5.2: Mirror Reflectivity against photon energy, as measured at CERN.



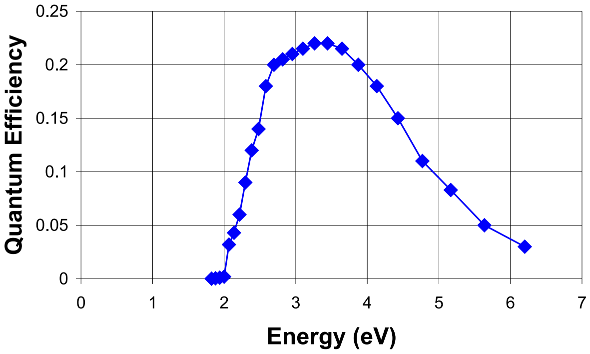
- For the 3×3 cluster test, the electronics are completely new, providing LHC-speed readout.
- In order to fit an entire Cherenkov ring onto the 3×3 cluster, the distance between the mirror and the detector plane, and the focal length of the mirror, have been reduced to 1.1 m.
- The radiator length is reduced to 1 m, as in RICH 1.
- No CEDAR counter is available.

A schematic of the modified RICH prototype geometry is given in figure 5.1. The measured reflectivity of the mirror is between 80% and 90%, as shown in figure 5.2.

5.2.2 MaPMT

The Hamamatsu R7600-03-M64 multi-anode photo multiplier tube (MaPMT) consists of an 8×8 array of square anodes, each with its own metal dynode chain incorporated into a single vacuum tube.

Figure 5.3: Quantum efficiency of the Hamamatsu R7600-03-M64 MaPMT (measured by Hamamatsu)



The geometry of the tube is shown in figure 5.4. Each dynode chain consists of 12 dynodes with a voltage distribution as shown in table 5.1. The data presented in this chapter are taken with a voltage of 1000 V between the cathode and the anode. The photons are converted into photoelectrons in a bialkali photo-cathode deposited on the inside of a 0.8 mm thick UV-glass window. The quantum efficiency of the photo-cathode is shown in figure 5.3. This includes the effect of the UV-glass window, which cuts off at energies above 6.2 eV. The integrated quantum efficiency is:

$$\int Q_{\text{eff}}(E)\mathrm{d}E = 0.57\,\text{eV}. \qquad (5.1)$$

The photoelectrons created in the cathode are electrostatically focused onto the dynode chains. A sketch of the dynode structure is given in figure 5.5. The dynode chain is not symmetric under a rotation of 90°; figure 5.5 therefore defines a vertical and a horizontal direction, for future reference.

Figure 5.4: Hamamatsu R7600-03-M64

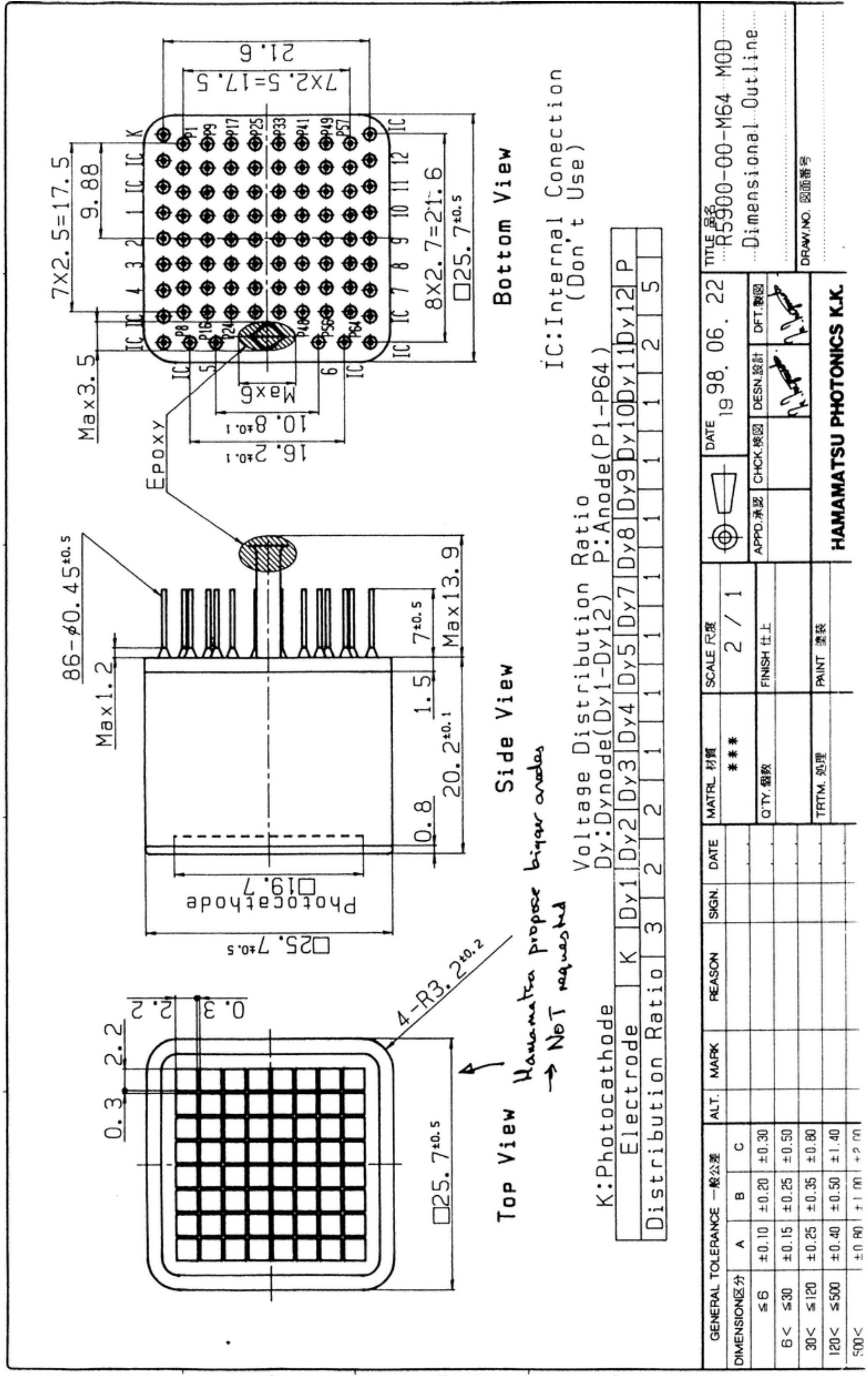


Table 5.1: Voltage distribution in 12–dynode MaPMT, normalised to the voltage between dynodes 3 and 4.

relative voltage		3		2		2		1		1		1	···	1		1		2	
dynode number	Cathode		1		2		3		4		5		···		10		11		12

Figure 5.5: Side-view of a tube, with the cathode at the top and the anode at the bottom, revealing the dynode structure. A horizontal and a vertical direction on the tube–surface is defined relative to the orientation of the dynodes.

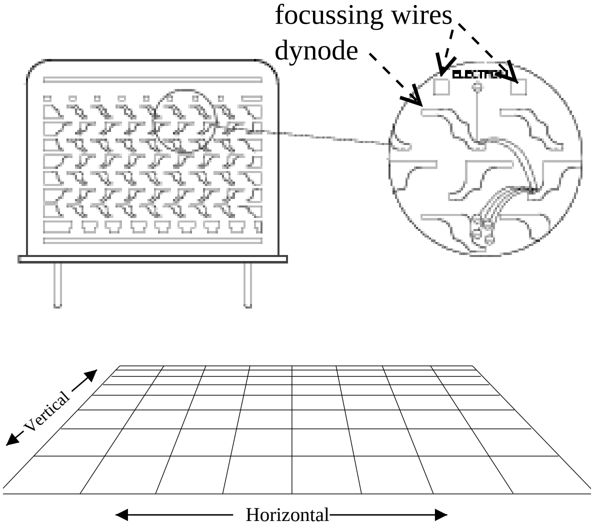
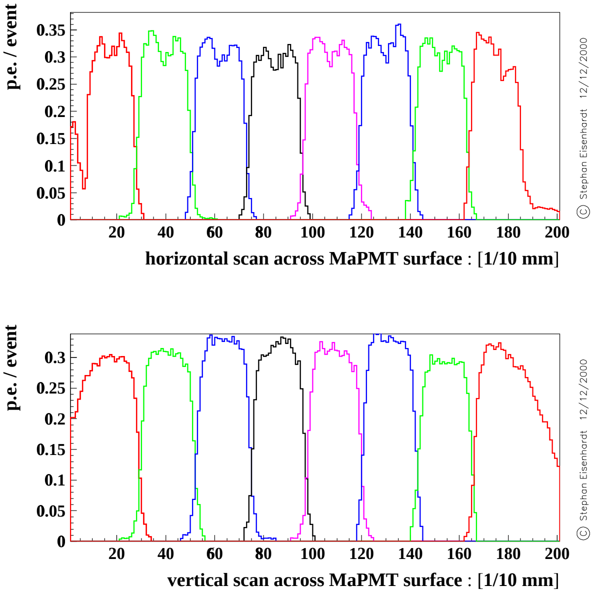


Figure 5.6: LED scan across a pixel row and column, with LED spot-size ∼ 100 μm. The graphs show the number of photoelectrons per event as a function of the position of the light spot. For a description of the scanning facility, see [A+00b]



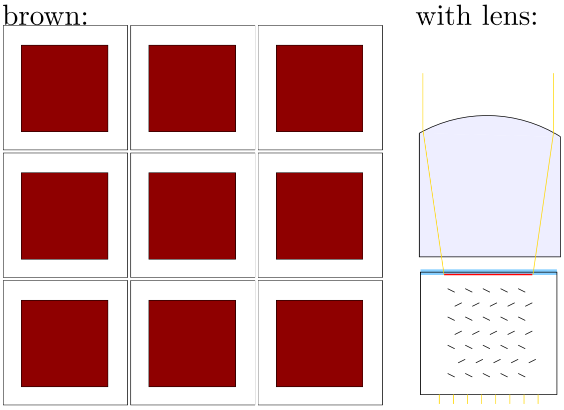
Measurements of the sensitivity of the tube as a function of position are shown in figure 5.6. The MaPMT active area is well modelled by assuming a $2 \times 2\,\mathrm{mm}^2$ pixel size with constant sensitivity, with 0.3 mm gaps between the pixels.

Due to the dead area surrounding the photo-cathode, the active area fraction for the MaPMTs is only about 50%. The geometrical coverage can be increased by placing a lens in front of the tube, as illustrated in figure 5.7. Nine quartz lenses ($n = 1.5$) are used. The lens dimensions are $26 \times 26\,\mathrm{mm}^2$ area, a maximum height of 24 mm and a radius of curvature of 25 mm. There is a ∼ 2 mm gap between the lenses and the entrance window of the MaPMTs. With these specifications, the full aperture of the lenses is focused onto the sensitive area of the MaPMTs, thus doubling the geometric efficiency.

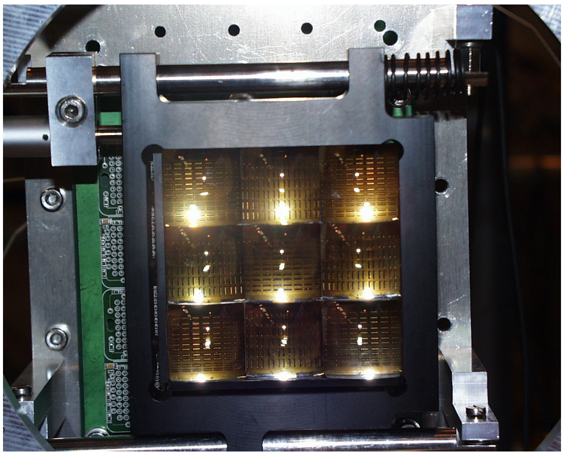
Figure 5.7: Doubling the MaPMT active area using lenses

Schematic of MaPMT array showing the active area in brown:

Schematic of MaPMT with lens:



Photograph of MaPMT array with lenses:



5.2.3 Mounting

The aluminium plate which held the HPDs, has been replaced with a light-tight box inside which the photodetectors are mounted. The mounting for the 3 × 3 array of MaPMTs includes the fast readout electronics. It can be rotated to change the angle of incidence of the Cherenkov photons. A photograph of the MaPMT mounting, with tubes and readout electronics, is shown in figure 5.1.

One readout board provides space for two tubes. Figure 5.8 shows the tubes at their relative positions as seen from the mirror. For clarity, the boxes representing the tubes are not drawn to scale; in reality they nearly touch each other. Tubes that are read out by the same board are grouped together in the figure. The arrows next to the tubes indicate the orientation of the tubes in terms of the horizontal direction, as defined in figure 5.5. The pixel numbering refers to the position of the pixels on the readout board, and is independent of the orientation of the tube. The tube, board and pixel numbering scheme shown in the diagram is used throughout the text.

5.2.4 Electronics

For the beam tests of the 3 × 3 array of MaPMTs, a fast, pipelined electronic read out system is used, which is shown schematically in figure 5.9. In the following, a short description of the main components of that readout system will be given. A more detailed account can be found in [A+00b].

Figure 5.8: Pixel, tube, and board numbering. The tubes are shown in their relative position on the detector plane, as seen from the mirror. Tubes read out by the same board are enclosed in a box. The size of the tubes is not drawn to scale. The arrows next to the tubes indicate the tube orientation in terms of the horizontal direction defined in figure 5.5

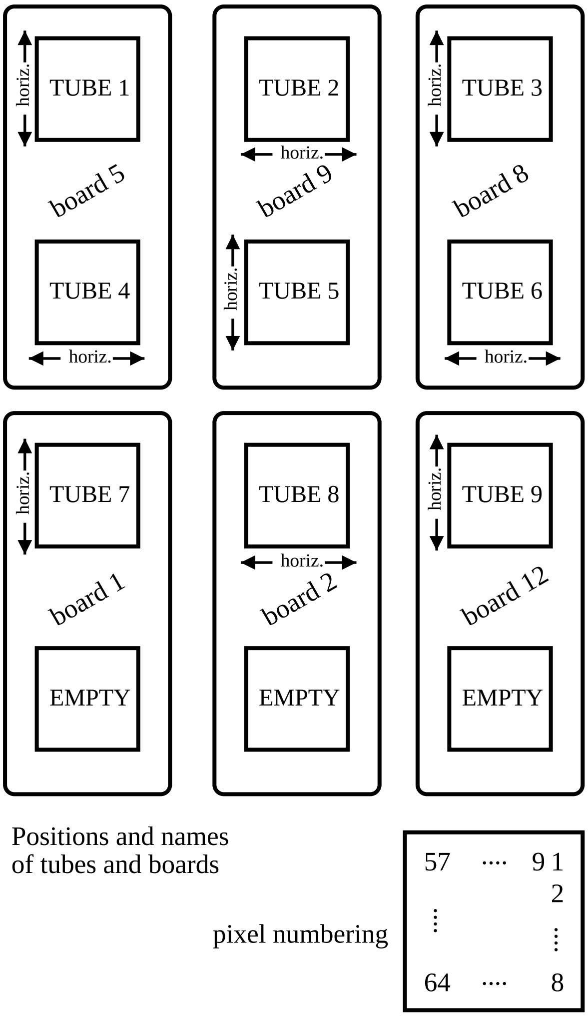
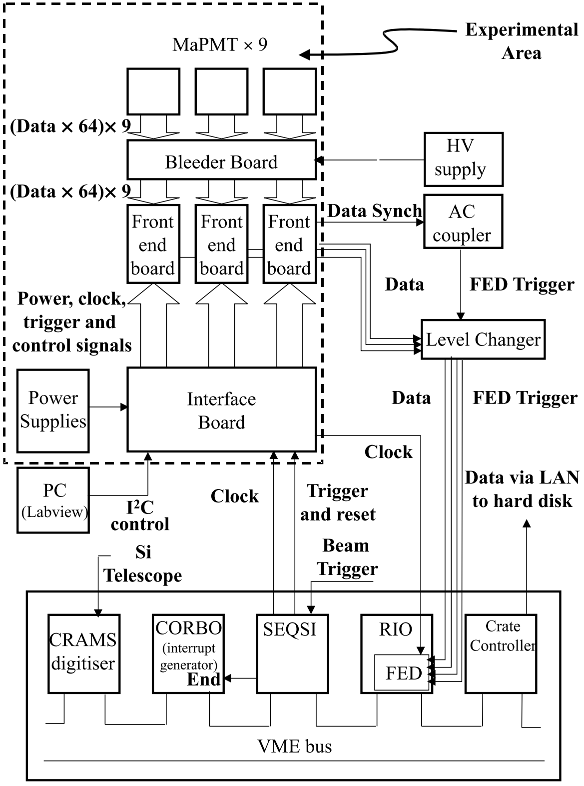


Figure 5.9: Block diagram of electronic readout and data acquisition system for 3 × 3 MaPMT array tests.



Bleeder Board

The bleeder board, shown in figure 5.10, provides the mechanical support and dynode resistor chain network for up to 16 MaPMTs in a 4 × 4 array. The board also adapts the MaPMT anode feed-through pitch of the 1024 data channels to the Pin Grid Array (PGA) pitch of a kapton cable. The bleeder board is constructed as two separate parts: the high voltage distribution board and the pitch adapter board. These are assembled as a single unit using an array of soldered interconnect pins.

Figure 5.10: Bleeder Board (MaPMT side up)

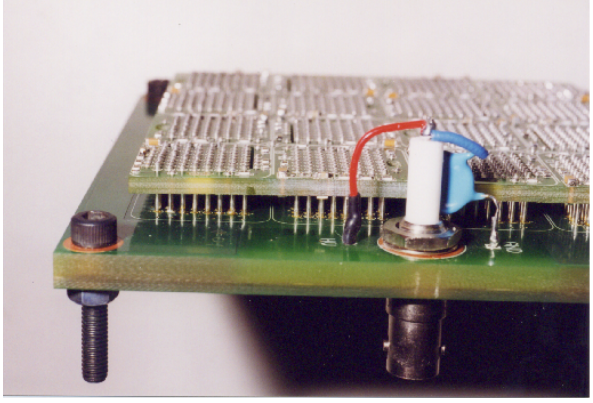
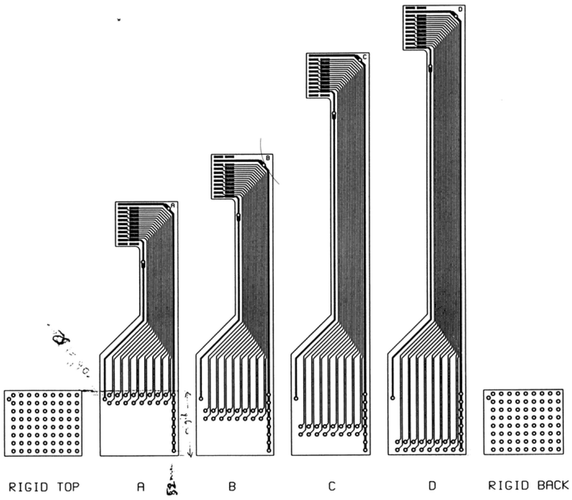


Figure 5.11: 4-layer kapton cable



Kapton Cables

Kapton cables are used to couple the 8 × 8 output PGA for each tube in the backplane of the bleeder board, to the front-end board. The kapton cable from one MaPMT is coupled to two 40-way SAMTEC[1] connectors (64 data channels and 16 ground lines) on the front end board. The cables are flexible printed circuits with the tracks laid on one side of the kapton and a ground plane on the reverse. A cable is composed of four strips of kapton. Each strip connects two columns of MaPMT output to half the channels of one SAMTEC connector. The layout of the kapton strips is shown in figure 5.11.

Front-End Board

The layout of the front-end board is given in figure 5.12. Its main components are the AC-coupler network and the APVm chip. The APV chips are designed for use with silicon strip detectors or micro–strip gas chambers which produce signals approximately ten times smaller than those from a MaPMT. Therefore the signal from the MaPMT has to be attenuated to match the dynamic range of the APVm. This is achieved with an AC coupler made from a ceramic base on which gold tracks are laid. A schematic of the 'tuning fork' design of the AC coupler is shown in figure 5.13.

APVm Chip

The APVm [J+98] is one of the APV series [dFL+96, R+97] of radiation hard chips designed for the readout of the CMS inner detector [CMS98]. The chips have previously been tested in a beam environment for the readout of silicon-strip detectors [dFL+96].

In the readout system for the MaPMT array the APVm, running at the LHC bunch-crossing rate of 40 MHz, samples the 128 input analogue signals every 25 ns. The sig-

---

[1] Cable-to-board SFMC series

Figure 5.12: Schematic of front end board

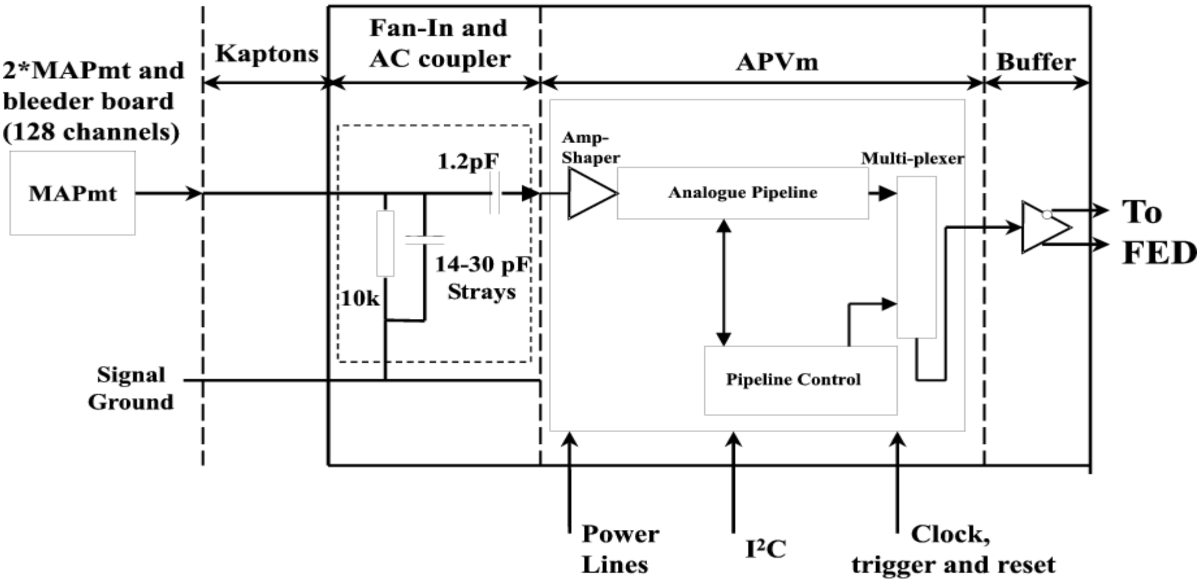
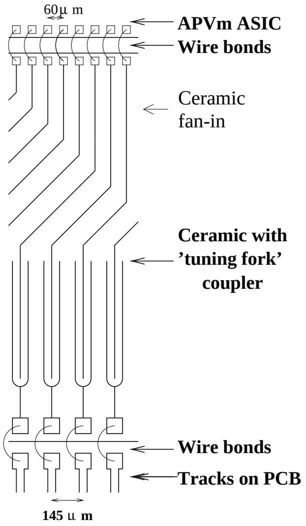


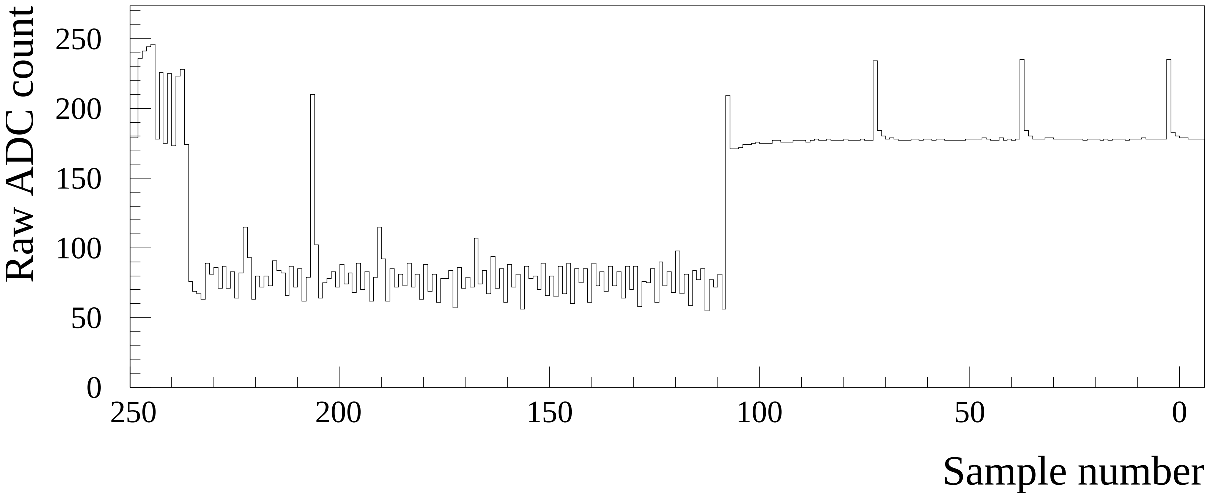
Figure 5.13: The AC coupler



nals are stored in a pipeline structure of 160 cells which allows a maximum first level trigger latency of 3.2 μs. When the events are output by the chip, the 128 signals are multiplexed into one analogue data output. An example of such a data–frame is given in figure 5.14, which shows the 128 signals preceded by 12 samples which contain information set by the chip: 4 bits signal the beginning of the analogue output and whether the chip is in error for an event, 8 bits are used for the pipeline address where the data are stored. The analogue output is cycled out of the chip at half the sampling rate.

The APV chip does not meet the requirements of the LHCb trigger and readout architecture, since it has never been designed to do so. The 1:128 multiplexing of the input signals at 20 MHz is too slow for the LHCb first level trigger rate of 1 MHz and a signal return-to-zero within 25 ns is only possible in the de-convolution mode

Figure 5.14: APVm output: The figure shows the analogue output of the APVm vs. sample number. Time flows from left to right. The earliest part of the output, on the left, gives header information and the pipeline address. This is followed by the data frame, with its lower baseline voltage. The signal-peak is clearly visible near sample number 205. The smaller signals at -1 and ±16 relative to the main signal peak are a consequence of the cross-talk discussed in section 5.6.



of APVm operation, which requires three consecutive samples for the algorithm, and therefore prevents triggering on consecutive events. Chips suitable for LHCb are currently under development [A+97, vB+99].

The readout system used in the testbeam does however provide a test of the suitability of the tubes for integration into the LHCb readout electronics, as the data from the tubes are captured within the 25 ns that are available between bunch crossings at the LHC.

The DAQ System

The main components of the VME–based data acquisition system are shown in figure 5.9. The data are generated from three sources: the nine MaPMTs, the silicon beam telescope and the trigger scintillators. Readout is triggered by the coincidence between the four scintillators in the beam line. Any further beam triggers are gated out for the remainder of the readout cycle. The control of the front-end chip is performed by the SEQSI programmable front-end control module [Mor]. The six APVm chips produce an analogue data output which is digitised using the Front End Digitiser, a prototype module for the readout of the CMS inner tracker [BCH+99, Cou]. The FED is programmed to run in ‘digital scope’ mode, which means that each individual event is read out and no data reduction, such as zero-suppression, is performed.

5.3 Describing PMT Spectra

5.3.1 Introduction

Similar to the case of the HPD spectra, in order to estimate the signal loss, a model of the output of an MaPMT-pixel is required, that can be fitted to the measured pulse height spectra. In order to perform such a fit, an analytical function is needed that can be calculated reasonably quickly by a computer.

Such a function is derived here. In section 5.3.2, an analytical expression is derived that describes the output of a PMT. The derivation is based on the assumption that the number of photoelectrons per event, as well as the number of secondary electrons caused by each primary electron at each stage of the dynode chain, are each described by Poisson distributions. It is then shown how this expression can be adapted to avoid some of the numerical problems arising, so that it can readily be calculated by a computer. A complete numerical recipe is given and a FORTRAN implementation of the program is listed in appendix A. This expression can be used to calculate any “snowball”–like effect described by a series of Poisson distributions.

In section 5.3.3 it is described how the exact expression derived in the first part can be used as the central element of a faster, approximate algorithm, and how the number of parameters can be reduced making reasonable assumptions, so that fitting a large number of spectra in a finite time becomes feasible. This is then adapted to describe the digitised output of laboratory read-out electronics, rather than the number of electrons at the end of a dynode chain.

This approximate function is used in section 5.3.4 to fit Monte Carlo generated pulse-height spectra to demonstrate the validity of the method. Finally it is shown how the fit is used on test–beam data to estimate the loss of single photoelectron events below a threshold cut.

5.3.2 An Analytical Function to Describe the Output of a PMT

An Expression for the Number of Photoelectrons at the End of a Dynode Chain

Assuming the number of photoelectrons produced in the cathode follows a Poisson distribution with mean λ_1, the probability to find k_1 photoelectrons *arriving* at the first dynode is:

$$P(k_1) = e^{-\lambda_1} \frac{\lambda_1^{k_1}}{k_1!}. \qquad (5.2)$$

The probability to find k_2 electrons after the first dynode is the sum over all values for k_1 of the probabilities $P(k_1)$, each multiplied with the probability that the dynode returns k_2 electrons given that k_1 arrive:

$$P(k_2) = \sum_{k_1=0}^{\infty} P(k_1) \cdot P(k_2|k_1). \qquad (5.3)$$

Each of the k_1 electrons produces a Poisson-distributed response from the dynode with mean λ_2 where λ_2 is the gain at the 1st dynode; all k_1 electrons together produce a response distributed according to the convolution of k_1 Poisson distributions, each with mean λ_2. This results in a single Poisson distribution with mean $\lambda_2 \cdot k_1$:

$$P(k_2|k_1) = e^{-\lambda_2 k_1} \frac{(\lambda_2 k_1)^{k_2}}{k_2!}. \tag{5.4}$$

Hence the probability to find k_2 electrons after the first dynode is given by:

$$P(k_2) = \sum_{k_1=0}^{\infty} P(k_1) \cdot e^{-\lambda_2 k_1} \frac{(\lambda_2 k_1)^{k_2}}{k_2!}. \tag{5.5}$$

Inserting the right-hand side of equation 5.2 for $P(k_1)$ yields, after a few manipulations:

$$P(k_2) = e^{-\lambda_1} \frac{\lambda_2^{k_2}}{k_2!} \sum_{k_1=0}^{\infty} \frac{(\lambda_1 e^{-\lambda_2})^{k_1}}{k_1!} k_1^{k_2}. \tag{5.6}$$

Generalising this for $n-1$ dynodes yields:

$$\begin{aligned} P(k_n) \;=\; & e^{-\lambda_1} \frac{\lambda_n^{k_n}}{k_n!} \\ & \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \cdots \sum_{k_{n-1}=0}^{\infty} \\ & \frac{(\lambda_1 e^{-\lambda_2})^{k_1}}{k_1!} \frac{(\lambda_2 e^{-\lambda_3} k_1)^{k_2}}{k_2!} \\ & \frac{(\lambda_3 e^{-\lambda_4} k_2)^{k_3}}{k_3!} \cdots \\ & \frac{(\lambda_{n-2} e^{-\lambda_{n-1}} k_{n-3})^{k_{n-2}}}{k_{n-2}!} \\ & \frac{(\lambda_{n-1} e^{-\lambda_n} k_{n-2})^{k_{n-1}}}{k_{n-1}!} k_{n-1}^{k_n}. \end{aligned} \tag{5.7}$$

Each term in equation 5.7 is of the form of an exponential series, i.e. $\frac{x^k}{k!}$, except for the last term with the summation parameter k_{n-1}, which appears as $\frac{x^k}{k!} k^{k_n}$. Using

$$k_{n-1}^{k_n} = \frac{\mathsf{d}^{k_n}}{\mathsf{d}y^{k_n}} e^{y k_{n-1}} \Big|_{y=0} \tag{5.8}$$

this last term can be written as

$$\begin{aligned} & \frac{(\lambda_{n-1} e^{-\lambda_n} k_{n-2})^{k_{n-1}}}{k_{n-1}!} k_{n-1}^{k_n} \\ = \; & \frac{\mathsf{d}^{k_n}}{\mathsf{d}y^{k_n}} \frac{(\lambda_{n-1} e^{-\lambda_n} k_{n-2} e^{y})^{k_{n-1}}}{k_{n-1}!} \Big|_{y=0} \end{aligned} \tag{5.9}$$

Using equation 5.9, equation 5.7 can be re–written as:

$$\begin{aligned} P(k_n) \;=\; & e^{-\lambda_1} \frac{\lambda_n^{k_n}}{k_n!} \quad \frac{\mathsf{d}^{k_n}}{\mathsf{d}y^{k_n}} \\ & \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \cdots \sum_{k_{n-1}=0}^{\infty} \\ & \frac{(\lambda_1 e^{-\lambda_2})^{k_1}}{k_1!} \frac{(\lambda_2 e^{-\lambda_3} k_1)^{k_2}}{k_2!} \\ & \frac{(\lambda_3 e^{-\lambda_4} k_2)^{k_3}}{k_3!} \cdots \\ & \frac{(\lambda_{n-2} e^{-\lambda_{n-1}} k_{n-3})^{k_{n-2}}}{k_{n-2}!} \\ & \frac{(\lambda_{n-1} e^{-\lambda_n} k_{n-2} e^{y})^{k_{n-1}}}{k_{n-1}!} \Big|_{y=0}. \end{aligned} \tag{5.10}$$

Now each summation can be carried out in turn, starting with that over k_{n-1}:

$$\begin{aligned} & \sum_{k_{n-1}=0}^{\infty} \frac{(\lambda_{n-1} e^{-\lambda_n} k_{n-2} e^{y})^{k_{n-1}}}{k_{n-1}!} \\ = \; & \exp(\lambda_{n-1} e^{-\lambda_n} k_{n-2} e^{y}) \\ = \; & (\exp(\lambda_{n-1} e^{-\lambda_n} e^{y}))^{k_{n-2}}, \end{aligned} \tag{5.11}$$

then over k_{n-2}:

$$\sum_{k_{n-2}=0}^{\infty} \Big(\frac{(\lambda_{n-2} e^{-\lambda_{n-1}} k_{n-3})^{k_{n-2}}}{k_{n-2}!}.$$

$$\begin{aligned}
& \left(e^{(\lambda_{n-1}e^{-\lambda_n e^y})}\right)^{k_{n-2}}\Big) \\
= \; & \exp(\lambda_{n-2}e^{-\lambda_{n-1}}k_{n-3} \cdot e^{(\lambda_{n-1}e^{-\lambda_n e^y})}) \\
= \; & (\exp(\lambda_{n-2}e^{-\lambda_{n-1}} \cdot e^{(\lambda_{n-1}e^{-\lambda_n e^y})}))^{k_{n-3}}
\end{aligned} \tag{5.12}$$

and so on. After performing all these summations, the probability of finding k_n electrons after $n-1$ dynodes, with gains $\lambda_2, \ldots, \lambda_n$, starting off with an average of λ_1 photo electrons arriving at the first dynode, is given by:

$$\boxed{\begin{aligned}
P(k_n) \; = \; & e^{-\lambda_1}\frac{\lambda_n^{k_n}}{k_n!} \frac{\mathsf{d}^{k_n}}{\mathsf{d}y^{k_n}} \\
& \exp(x_1 \exp(x_2 \exp(x_3 \cdots \\
& \exp(x_{n-1}\exp(y)) \cdots)))\Big|_{y=0} \\
\text{with } x_i \; \equiv \; & \lambda_i e^{-\lambda_{i+1}}.
\end{aligned} \tag{5.13}}$$

Calculating $P(k_n)$

In order to calculate $P(k_n)$ it is useful to make the following definitions:

$$\begin{aligned}
f_1 & \equiv e^{x_1 e^{x_2 \cdots e^{x_{n-1}e^y}}} \\
f_2 & \equiv e^{x_2 e^{x_3 \cdots e^{x_{n-1}e^y}}} \\
f_3 & \equiv e^{x_3 e^{x_4 \cdots e^{x_{n-1}e^y}}} \\
& \vdots \\
f_{n-1} & \equiv e^{x_{n-1}e^y} \\
f_n & \equiv e^y.
\end{aligned} \tag{5.14}$$

Equation 5.13 can now be written as:

$$P\left(k_n\right) = e^{-\lambda_1}\frac{\lambda_n^{k_n}}{k_n!} f_1^{(k_n)}(y)\Big|_{y=0}, \tag{5.15}$$

where $f_1^{(k_n)}$ is the $k_n th$ derivative of f_1 with respect to y. With the above definitions, the first derivatives of the functions f_i are given by:

$$\begin{aligned}
f_1' & = f_1 x_1 \quad f_2 x_2 \quad \cdots \quad f_{n-1}x_{n-1} \quad f_n \\
f_2' & = \qquad\quad\; f_2 x_2 \quad \cdots \quad f_{n-1}x_{n-1} \quad f_n \\
f_3' & = \qquad\qquad\qquad\;\; \cdots \quad f_{n-1}x_{n-1} \quad f_n \\
& \vdots \\
f_n' & = \qquad\qquad\qquad\qquad\qquad\qquad\quad f_n.
\end{aligned} \tag{5.16}$$

This gives a recursive formula for the first derivative of f_i:

$$\begin{aligned}
f_i' & = f_i x_i f_{i+1}' \qquad i < n \\
f_n' & = f_n,
\end{aligned} \tag{5.17}$$

which in turn gives a recursive formula for the mth derivative:

$$f_i^{(m)} = \sum_{j=0}^{m-1} \binom{m-1}{j} f_i^{(j)} x_i f_{i+1}^{(m-j)}, \tag{5.18}$$

$$\text{with } f_n^{(j)} = f_n \;\; \forall j \in \mathbb{N}.$$

With this expression, equation 5.15 can finally be calculated, by starting with $f_n(0) = 1$ and calculating $f_i^{(m)}$ subsequently for all values $i = n, n-1, \ldots, 1$ and all values $m = 0, 1, \ldots, k_n$.

Numerical Problems

While the previous section gives a valid algorithm how to calculate $P(k_n)$ using equation 5.15 and the recursive formula 5.18, it turns out that the finite precision of a normal computer will only allow calculations to be performed for rather small values of k_n before some numbers become either too large or too small to be stored straightforwardly in the computer memory. This problem is addressed in the following discussion.

The factor $\frac{\lambda_n^{k_n}}{k_n!}$ For any reasonably large number of dynodes, where the mean number of electrons coming off the last dynode, and therefore the interesting values for k_n, is typically in the thousands or even millions, $e^{-\lambda_1}\frac{\lambda_n^{k_n}}{k_n!}$ quickly becomes very small, while $f_1^{(k_n)}(y)\Big|_{y=0}$ grows to extremely large values. In order to calculate $P(k_n)$ for such values of k_n, it is necessary to absorb the small factor $\frac{\lambda_n^{k_n}}{k_n!}$ into the $f_i^{(m)}$. This can be done by replacing y in equation 5.15 with py and introducing a compensating factor $\left(\frac{1}{p}\right)^{k_n}$:

$$P\left(k_n\right) = e^{-\lambda_1}\frac{\lambda_n^{k_n}}{k_n!}\left(\frac{1}{p}\right)^{k_n}\frac{\mathsf{d}^{k_n}}{\mathsf{d}y^{k_n}}f_1(py)\Big|_{y=0}. \tag{5.19}$$

Choosing p such that $p^{k_n} = \frac{\lambda_n^{k_n}}{k_n!}$ changes equation 5.15 to

$$\begin{aligned} P\left(k_n\right) &= e^{-\lambda_1}\frac{\mathsf{d}^{k_n}}{\mathsf{d}y^{k_n}}f_1(py)\Big|_{y=0} \\ \text{with} \qquad p^{k_n} &= \frac{\lambda_n^{k_n}}{k_n!}. \end{aligned} \tag{5.20}$$

Defining

$$f^{\star(m)}_{k_n,i} \equiv \frac{\mathsf{d}^m}{\mathsf{d}y^m}f_i(p_{k_n}y)\Big|_{y=0} \tag{5.21}$$
$$\text{with} \quad p_{k_n} = \frac{\lambda_n}{(k_n!)^{\frac{1}{k_n}}}$$

gives

$$P\left(k_n\right) = e^{-\lambda_1}f^{\star(k_n)}_{k_n,1}. \tag{5.22}$$

The recursive formula for calculating $f_1^{(k_n)}$ remains essentially the same for $f^{\star(k_n)}_{k_n,1}$:

$$f^{\star(k_n)}_{k_n,i} = \sum_{j=0}^{k_n-1}\binom{k_n-1}{j}f^{\star(j)}_{k_n,i}x_i f^{\star(k_n-j)}_{k_n,i+1} \tag{5.23}$$

$$\text{with } f^{\star(m)}_{k_n,n} = p^m_{k_n} \text{ and } p_{k_n} = \frac{\lambda_n}{(k_n!)^{\frac{1}{k_n}}},$$

with one additional complication. In the original algorithm, when calculating $f_1^{k_n}$ using the recursive formula 5.18, all values for f_i^m with $m < k_n$ calculated in the previous iterations[2] could be used in the recursive formula for the current iteration. Now, for calculating $f^{\star(k_n)}_{k_n,1}$ *all* values $f^{\star(m)}_{k_n,i}$ with $m < k_n$, $i \leq n$ have to be re-calculated at each iteration, because at each iteration the value for p in equation 5.23 changes. To calculate $f^{\star(k_n)}_{k_n,i}$, from equation 5.23, the values for

$$f^{\star(m)}_{k_n,i}, m < k_n$$

are needed. These can be calculated using only the values for $f^{\star(m)}_{k_n-1,i}$ which have been calculated one iteration earlier:

$$f^{\star(m)}_{k_n,i} = f^{\star(m)}_{k_n-1,i}\left(\frac{p_{k_n}}{p_{k_n-1}}\right)^m$$

with

$$\left(\frac{p_{k_n}}{p_{k_n-1}}\right)^{k_n} = \frac{((k_n-1)!)^{1/(k_n-1)}}{k_n}; \tag{5.24}$$

---

[2] where $P\left(0\right), \ldots, P\left(k_n-1\right)$ were calculated

so the values for $f^{\star(k_n)}_{k_n,i}$ need to be stored only for one iteration.

The binomial factor When calculating $f^{\star(k_n)}_{k_n,i}$, using the recursive formula 5.23, the factor $\binom{k_n-1}{j}$ in

$$f^{\star(k_n)}_{k_n,i} = \sum_{j=0}^{k_n-1} \binom{k_n-1}{j} f^{\star(j)}_{k_n,i} x_i f^{\star(k_n-j)}_{k_n,i+1}$$

can get very large for large values of k_n, while the corresponding values for $f^{\star(j)}_{k_n,i} x_i f^{\star(k_n-j)}_{k_n,i+1}$ get very small. To avoid the associated numerical problems, one can define the arrays $u^{(j)}_{k_n,i}$ and $v^{(j)}_{k_n,i}$ that 'absorb' the binomial factor, such that equation 5.23 becomes:

$$f^{\star(k_n)}_{k_n,i} = \sum_{j=0}^{k_n-1} u^{(j)}_{k_n,i} x_i v^{(k_n-j)}_{k_n,i+1}, \tag{5.25}$$

where

$$\begin{aligned} u^{(j)}_{k_n,i} &= \sqrt{\binom{k_n-1}{j}} f^{\star(j)}_{k_n,i} \\ v^{(j)}_{k_n,i} &= \sqrt{\binom{k_n-1}{j-1}} f^{\star(j)}_{k_n,i}. \end{aligned} \tag{5.26}$$

Putting It All Together

Combining the results from the previous two paragraphs, at each iteration k_n, before calculating $f^{\star(k_n)}_{k_n,i}$ using equation 5.25, the values for $u^{(j)}_{k_n,i}$ and $v^{(j)}_{k_n,i}$, $j < k_n$, are calculated from their values in the previous iteration:

$$\begin{aligned} u^{(j)}_{k_n,i} &= \left(\frac{p_{k_n}}{p_{k_n-1}}\right)^j \sqrt{\frac{k_n-1}{k_n-1-j}}\, u^{(j)}_{k_n-1,i} \\ v^{(j)}_{k_n,i} &= \left(\frac{p_{k_n}}{p_{k_n-1}}\right)^j \sqrt{\frac{k_n-1}{k_n-j}}\, v^{(j)}_{k_n-1,i} \\ j &< k_n. \end{aligned} \tag{5.27}$$

These results are then used to calculate:

$$f^{\star(k_n)}_{k_n,i} = \sum_{j=0}^{k_n-1} u^{(j)}_{k_n,i} x_i v^{(k_n-j)}_{k_n,i+1} \tag{5.28}$$

and

$$u^{(k_n)}_{k_n,i} = v^{(k_n)}_{k_n,i} = f^{\star(k_n)}_{k_n,i}, \tag{5.29}$$

starting from

$$u^{(k_n)}_{k_n,n} = v^{(k_n)}_{k_n,n} = f^{\star(k_n)}_{k_n,n} = \frac{\lambda_n^{k_n}}{k_n!} \tag{5.30}$$

and

$$u^{(0)}_{0,i} = v^{(0)}_{0,i} = f^{\star(0)}_{0,i} = f_i, \tag{5.31}$$

where the f_i are defined by equation 5.14.

The Complete Numerical Recipe

Using the above formulae, the problem of calculating the probability distribution of finding k_n electrons at the end of a PMT with $n-1$ dynodes can be solved by a computer. A FORTRAN implementation is listed in appendix A. The program takes as its input the array $\lambda[n]$, with dimension n, which contains the average number of photo electrons arriving at the first dynode $\lambda[1]$ and the gain at each of the $n-1$ dynode, $\lambda[2], \ldots, \lambda[n]$. The program fills the array $P[\mathsf{max}]$ with the probabilities $P[k]$ to find k electrons at the end of the dynode chain for all values $k \leq \mathsf{max}$. The parameter max is also passed to the program.

The values for $u^{(j)}_{k,i}, v^{(j)}_{k,i}$ needed in the recursive formulae, are stored in two two-dimensional arrays, where one dimension is taken by the index $i = 1, \ldots, n$, and the other by the index $j = 0, \ldots, \mathsf{max}$. As the values for $u^{(j)}_{k,i}, v^{(j)}_{k,i}$ are needed only for one value of k at a time, the arrays do not need to be three-dimensional; the values for $u^{(j)}_{k,i}, v^{(j)}_{k,i}$ needed at the iteration calculating $P[k]$ replace those from the previous iteration, $u^{(j)}_{k-1,i}, v^{(j)}_{k-1,i}$.

The steps to calculate $P[k], k = 0, \ldots, \mathsf{max}$ are:

1 Initialise program, test whether input is sensible, for example if the overall gain is larger than 0. Calculate all values for $\left(\frac{p_j}{p_{j-1}}\right)^j, j \leq \mathsf{max}$ and store them in an array $p_{\mathsf{frac}}[j], j = 1, \ldots \mathsf{max}$ for later use.

2 Start with calculating the probability to find zero electrons: $k = 0$

3 Calculate $u^{(0)}_{0,i} = v^{(0)}_{0,i} = f_i$ for $i = n, n-1, \ldots, 1$, as defined by equation 5.14

4 Store the result in the array: $P[0] = e^{-\lambda_1} u^{(0)}_{0,1}$

5 Increment k by 1. If $k > \mathsf{max}$, stop program.

6 Calculate $u^{(k)}_{k,n} = v^{(k)}_{k,n} = \frac{\lambda_n^k}{k!}$

7 Calculate $u^{(j)}_{k,i}, v^{(j)}_{k,i}$ for $j < k$ and $i = n, \ldots, 1$ from $u^{(j)}_{k-1,i}, v^{(j)}_{k-1,i}$ according to equation 5.27, using the values of $p_{\mathsf{frac}}[k]$ calculated in step 1.

8 Calculate $u^{(k)}_{k,i} = v^{(k)}_{k,i}$ for all values of $i < n$ using the recursive formula 5.28. Let the outer loop go over the index i, starting with $i = n-1$ and decrementing it by 1 until $i = 1$, and the inner loop over the summation index j, starting with $j = 0$ and incrementing j by 1 until $j = k-1$.

9 Store result: $P[k] = e^{-\lambda_1} u^{(k)}_{k,1}$

10 Goto step 5

5.3.3 Fitting Spectra

Increasing Speed by Approximating $\mathbf{P(k_n)}$

When fitting PMT-pulse-height spectra, speed is a major problem. The number of operations needed to calculate $P(k_n)$ using the recursive formula in equation 5.18, is

$$N_{\mathsf{steps}} \approx \sum_{i=0}^{k_n} \sum_{j=0}^{i} nj \sim k_n^3, \tag{5.32}$$

which becomes absolutely prohibitive for a typical MaPMT with a gain of $\sim 3 \cdot 10^5$. Therefore, for fitting the spectra, only the exact distribution after the first m dynodes is calculated and then scaled by the gain of the remaining dynodes, $g_{\mathsf{left}} = (g_{m+1} g_{m+2} \cdots g_{n-1})$. When scaling the output of the exact distribution calculated for the first m dynodes, $P_{\mathsf{exact}}(k_{m+1})$, to the final distribution, the result is convoluted with a Gaussian of width σ_{scale}, taking into to account the additional spread in the distribution at each remaining dynode:

$$\sigma_{\mathsf{scale}} = \sqrt{k_{m+1}} \sigma_0 \tag{5.33}$$

with:

$$\sigma_0 = (g_{m+1}g_{m+2}\cdots g_{n-1}) \cdot \left(\frac{1}{g_{m+1}} + \frac{1}{g_{m+1}g_{m+2}} + \cdots + \frac{1}{g_{m+1}\cdots g_{n-1}}\right)^{\frac{1}{2}}. \tag{5.34}$$

So the approximated function, $P_\sim(k_n)$ is

$$P_\sim(k_n) = \sum_{j=0}^{\infty} \frac{1}{\sqrt{2\pi}\sqrt{j}\,\sigma_0} e^{\frac{(j\cdot g_{\mathsf{left}} - k_n)^2}{2\left(\sqrt{j}\,\sigma_0\right)^2}} P(j). \tag{5.35}$$

In practice the sum only needs to be calculated for values of $j \cdot g_{\mathsf{left}}$ that are a few σ_{scale} around k_n.

Reducing the Number of Parameters

$P(k_n)$ depends on n parameters: one for each dynode and one for the number of photoelectrons produced in the cathode. For the case of the 12-dynode MaPMT, there are 13 parameters. It is possible, however, to reduce this number to two:

1. the mean number of photoelectrons produced in the photo cathode
2. the gain at the first dynode.

Using

$$g \propto V^\alpha, \tag{5.36}$$

where V is the voltage difference over which the electron is accelerated, the gain at the other dynodes can be calculated from the gain at the first dynode. The parameter α has values typically between 0.7 and 0.8 [Ham00]; in the following, $\alpha = 0.75$ is used.

Adapting the Function to Fit Measured Data

In practice, spectra are not measured in numbers of photoelectrons, but in ADC counts digitised by the readout electronics. The function describing the spectra needs to relate the ADC counts, k_{adc}, to the number of electrons at the end of the dynode chain, k_n. This requires two parameters: the offset, or pedestal mean, p_0, and the conversion factor, c_n of k_n to ADC counts. The resulting function is convoluted with a Gaussian of width σ to take into account electronics noise:

$$F_{\mathsf{cont}}(k_{\mathsf{adc}}) = \left(\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{k_{\mathsf{adc}}^2}{2\sigma^2}}\right) * \left(P\left(\left(k_{\mathsf{adc}} - p_o\right)/c_n\right) \cdot c_n\right), \tag{5.37}$$

where $*$ is the convolution operator. F_{cont} treats k_{adc} as a continuous variable, with a one-to-one relation between k_{adc} and k_n; in fact the readout electronics deliver only integer-value ADC counts, integrating over the corresponding pulse heights. Thus the final function for describing ADC spectra is:

$$F(k_{\mathsf{adc}}) = \int_{k_{\mathsf{adc}}-0.5}^{k_{\mathsf{adc}}+0.5} F_{\mathsf{cont}}(k'_{\mathsf{adc}})\mathsf{d}k'_{\mathsf{adc}}. \tag{5.38}$$

5.3.4 Example Fits

The fits are performed as binned log-likelihood fits: for each 1-ADC-count wide bin k_{adc}, containing n_i events, the binomial probability of having n_i successes in N_{all} trials is calculated, where N_{all} is the total number of events. The probability of an individual success is given by $F(k_{\mathsf{adc}})$.

The probability distribution for the number of electrons after the fourth dynode is calculated without approximation. Then the function is scaled, approximating the additional spread due to the remaining dynodes with a Gaussian, as described in the previous section.

MC-Generated Spectra

The validity of the method has first been established on Monte Carlo simulated data. The Monte Carlo program simulates the output of a MaPMT pixel. The gain at the first dynode is $g_1 = 5$ and the gains at the other dynodes are calculated from $g \propto V^{\alpha}$ with $\alpha = 0.75$. The values for V are given in table 5.1.

The fit function is applied to two sets of 128 simulations with 10^5 events each, one set with 0.15 photoelectrons per event, one with 3.0 photoelectrons per event. A spectrum from each set is shown in figures 5.15 and 5.16.

The fits are performed varying the gain of only one dynode and calculating the gains at the other dynodes using the same value for α as in the Monte Carlo program that generated the spectrum. The fit results agree very well with the input values, as shown in tables 5.2 and 5.3. To test the sensitivity of the fit result on the exact knowledge of α, the fit to the spectrum in figure 5.15 is repeated assuming different values for this parameter in the fit-function:

Figure 5.15: MC-generated MaPMT ADC-spectrum in black, from 100k events, with $\lambda_1 = 0.15$. The fit is superimposed in blue.

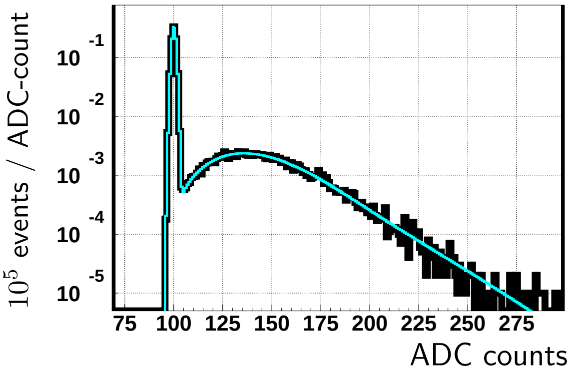


Table 5.2: Monte Carlo input compared with mean and RMS of the results from fits to 128 simulated spectra, with $\lambda_1 = 0.15$

	MC input	Mean fit result ± RMS spread
λ_1	0.150	0.1501±0.0013
g_1	5.000	5.0012±0.058
p_0	100.00	99.999±0.0038
σ	1.0000	1.0004±0.0027
c_n	$3.20 \cdot 10^{-4}$	$(3.23 \pm 0.46) \cdot 10^{-4}$

$\alpha = 0.5$ and $\alpha = 1.0$. The results are given in table 5.4. Another fit was performed that does not use the formula $g \propto V^{\alpha}$. Here it is only assumed, that dynodes with the same accelerating voltage have the same gain. Instead of one gain, three gains need to be fitted, one for each accelerating voltage. The fits are performed using the function minimisation and error analysis package MINUIT [Jam94]. The results from this fit, with error-estimates provided by MINUIT, are given in the last column of table 5.4.

Figure 5.16: MC-generated MaPMT ADC-spectrum in black, from 100k events, with $\lambda_1 = 3$. The fit is superimposed in blue.

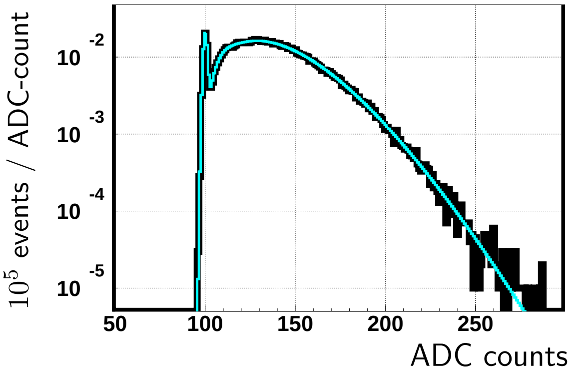


Table 5.3: Monte Carlo input compared with mean and RMS of the results from fits to 128 simulated spectra, with $\lambda_1 = 3$

	MC input	Mean fit result ± RMS spread
λ_1	3.000	3.002±0.022
g_1	5.000	4.985±0.107
p_0	100.000	99.999±0.021
σ	1.000	0.999±0.016
c_n	$6.4 \cdot 10^{-4}$	$(6.45\pm0.17) \cdot 10^{-4}$

Comparing the results for the different assumptions shows, that they have little impact on the fitted value for the number of photo electrons and the gain at the first dynode. Most of the error introduced by an incorrect estimate of the parameter α is absorbed into the ratio of ADC-counts to electrons, c_n, while the values for λ_1 and g_1 come out close to the input values.

Application to Test–Beam Data

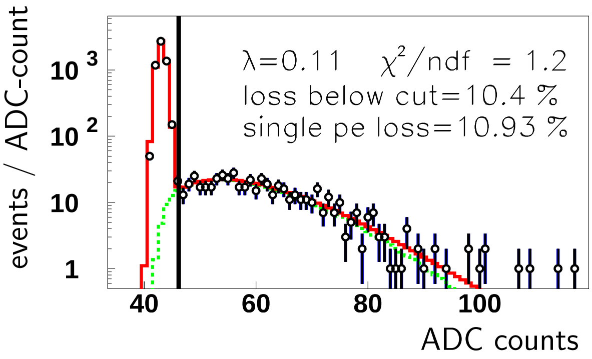
The fit method is applied to spectra obtained in the test–beam, and used to estimate the signal loss at the first dynode and below the threshold cut, given in table 5.7.

An example of such a fit is given in figure 5.17. It shows the pulse height spectrum from $6k$ events in MaPMT 6, pixel 32, run 2610 (no lenses), with the fit superimposed. Also, the single photoelectron contribution is shown. The signal loss figures given inside the plot refer to the signal loss below the threshold cut only and do not include the loss of photo electrons at the first dynode due to the probability of e^{-g_1} that the pri-

Table 5.4: Monte Carlo input compared with fit-result for the MC-generated pulse height spectrum shown in figure 5.15, using different assumptions in the fit.

	MC input	Fit result	Fit result	Fit result	Fit result: 3
	$\alpha = 0.75$	$\alpha = 0.75$	$\alpha = 0.5$	$\alpha = 1$	indep. dyn's
λ_1	0.1500	0.1490	0.1491	0.1489	0.1492±0.0013
g_1	5.00	5.039	4.852	5.291	4.74±0.44
g_2, g_3, g_{12}	$g_1 \cdot (\frac{2}{3})^\alpha = 3.69$				4.51±1.35
$g_4, \ldots, g_{11}$	$g_1 \cdot (\frac{1}{3})^\alpha = 2.19$				1.97±0.21
p_0	100.000	100.000	100.000	100.000	100.000±0.003
σ	1.0000	1.0028	1.0029	1.0027	1.0028±0.0025
c_n	$3.20 \cdot 10^{-4}$	$2.90 \cdot 10^{-4}$	$0.373 \cdot 10^{-4}$	$19.7 \cdot 10^{-4}$	$(4.37 \pm 0.98) \cdot 10^{-4}$

Figure 5.17: Data from 6k events in black, with fit superimposed in red. In green: The single photoelectron contribution. The signal loss refers to the loss below the threshold cut only. The loss at the first dynode is not included.



mary electron produced 0 secondary electrons. The results for the fit parameters, together with the error estimates provided by MINUIT, are given in table 5.5. The fit describes the data well, with a χ^2/ndf of 1.2.[3]

Table 5.5: Result of fit applied to test–beam data

	Fit result
λ_1	0.107±0.005
g_1	3.60±0.20
p_0	43.06±0.01
σ	0.724±0.008
c_n	$(61 \pm 30) \cdot 10^{-4}$

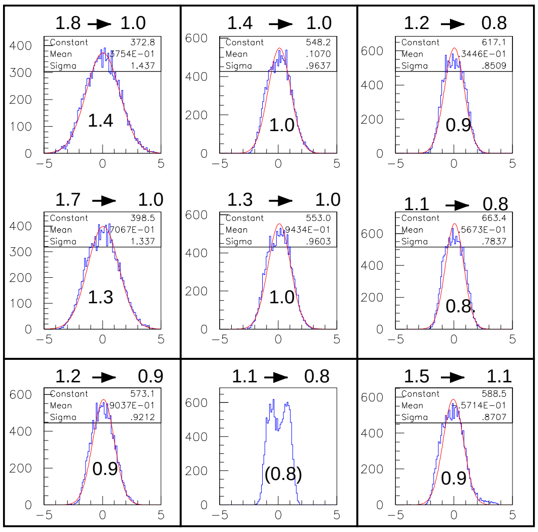
5.4 Preparation of Testbeam Data

Before the data from the testbeam can be used for further analysis, they need to be corrected for electronics effects. Serious problems are caused by cross talk between channels. The identification and correction of cross talk is therefore described in detail. Firstly, the subtraction of common modes is summarised.

---

[3] The fit is performed with the same log-likelihood method that was used for the MC spectra; nevertheless, a χ^2 value is calculated after the fit.

Figure 5.18: Common–mode value for each tube. The histograms are ordered in the same way as the tubes in Figure 5.8. Bold lines indicate shared read–out boards. The numbers above each histogram are the average pedestal width before → after common mode subtraction. The numbers inside the histograms are the σ of the Gaussian fit to the common–mode distribution, except for tube 8 where the RMS is given.



5.4.1 Common Mode

Common mode fluctuations are corrected tube by tube. In a first iteration, the pedestal mean (μ) and width (σ) are found for each pixel. The data are then reprocessed. A pixel with a pulse height below $\mu + 3\sigma$ is classified to be a non-hit pixel. In each event, the average deviation of the non-hit pixels in a tube from their respective pedestal means is subtracted from the pulse heights in all pixels, to correct for common mode fluctuations. Figure 5.18 shows the common mode correction tube by tube for 1.3k events. The common mode behaviour is very similar for tubes on the same read–out board.

5.4.2 Cross Talk

In order identify cross talk, LED runs are used. In these runs, the whole detector surface is illuminated reasonably uniformly, resulting in about 0.05 photoelectrons per pixel per event (above a 5σ cut).

Identifying and Locating Cross Talk

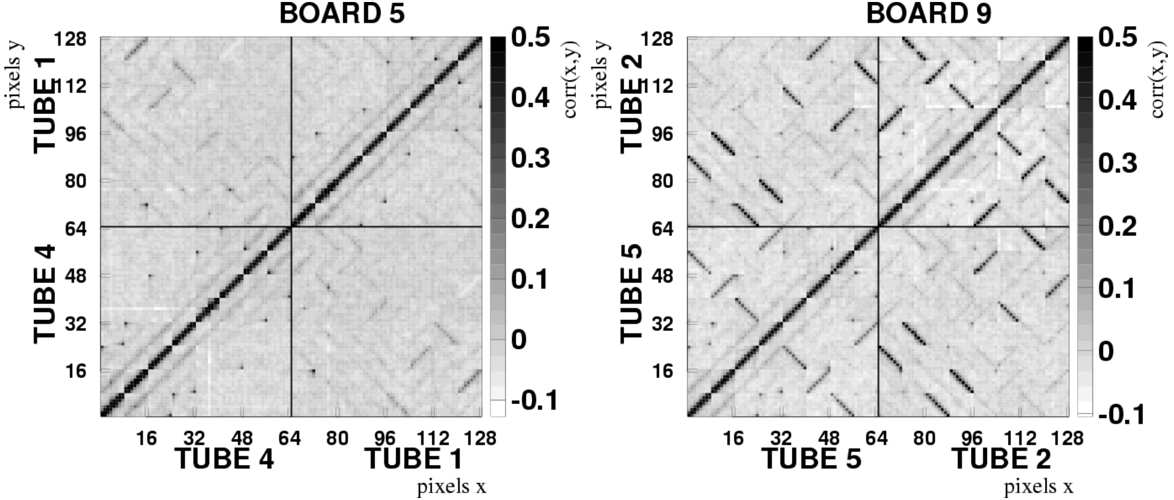
For each pair of pixels x, y, the correlation coefficient between the pulse heights in these pixels, h_x and h_y, is calculated

$$\mathrm{corr}(h_x, h_y) = \frac{\langle h_x \cdot h_y \rangle - \langle h_x \rangle \cdot \langle h_y \rangle}{\sigma_{h_x} \cdot \sigma_{h_y}}$$

where σ_{h_x} and σ_{h_y} are the RMS values of the pulse height distributions in pixels x and y respectively. The results are displayed in two-dimensional plots such as in figure 5.21 for a single tube, or in figure 5.19 for all pixels read out by the same board. The numbers on the axes are the pixel numbers defined in figure 5.8. If two tubes are shown in the same plot, 64 is added to the pixel numbers of one of the tubes.

Significant non–zero correlation is evidence for cross talk. As the correlation coefficient is symmetric by definition, $\mathrm{corr}(h_x, h_y) \equiv \mathrm{corr}(h_y, h_x)$, this method does not allow to identify a possible direction of the cross talk. Therefore, the

Figure 5.19: Correlation coefficients



The plot entries are the correlation coefficients between the pulse heights of the pixel pairs (x,y). To increase readability, the grey-scale has been set to a maximum of 0.5 (no entry off the diagonal has a value larger than 0.43).

probability that a hit in pixel x is caused by a hit in pixel y is calculated:

$$P(x \leftarrow y) = \frac{\text{№ of hits caused by y in x}}{\text{total № of hits in x}}$$

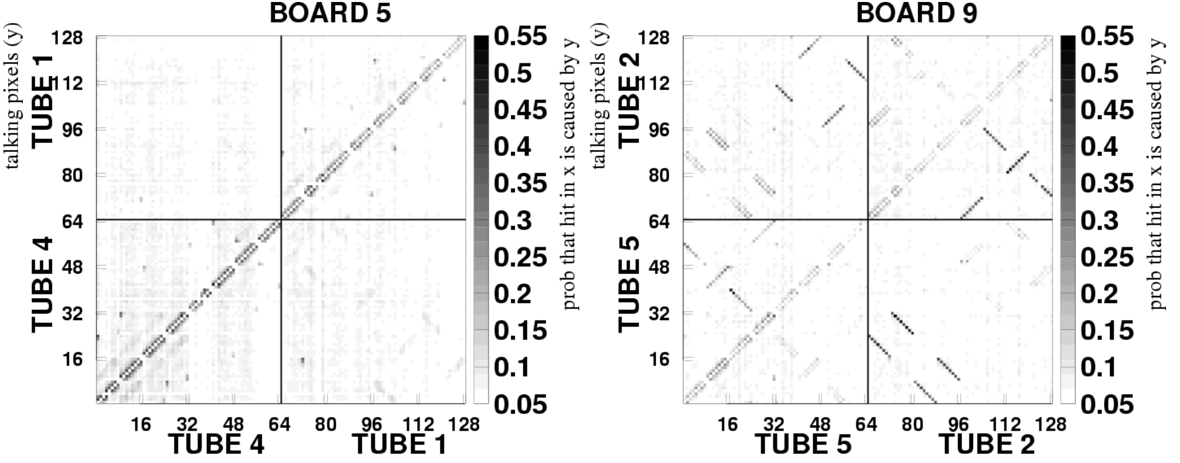
For calculating this probability, as for photon counting, a hit is defined as a pulse height at least 5σ above the pedestal mean; if there are two hits in a given pixel pair, it is assumed that the pixel with the larger pulse height *caused* the hit in the pixel with the smaller pulse height. Sometimes of course, there are two genuine hits (or hits induced by cross talk from other pixels) in a pixel pair. For the LED runs used, this occurs $\sim 5\%$ of the time.

Figure 5.20 shows the cross talk probabilities for two boards with fast readout. The talked–to pixels x are shown on the horizontal axis, and the talking pixels y on the vertical axis. Two classes of cross talk can be identified in those plots:

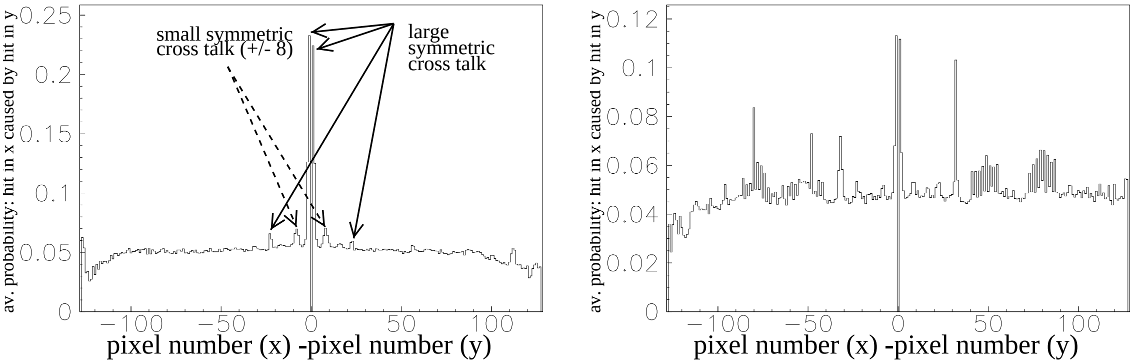
1. Symmetric cross talk, predominantly along $y = x \pm 1$, present in both boards.
2. Asymmetric cross talk, which can be seen on board 9, but not on board 5. The asymmetric cross talk forms irregular patterns, and is as strong between pixels on different tubes, as it is between pixels on the same tube.

Figure 5.21 shows the correlation coefficients for tube 1, once read out with the fast readout, and once with a CAMAC system which does not utilise the fast electronics. No cross talk is observed when the tube is read out with the CAMAC system. This demonstrates that cross talk seen in the fast read out is entirely due to electronics effect.

Figure 5.20: cross talk probabilities

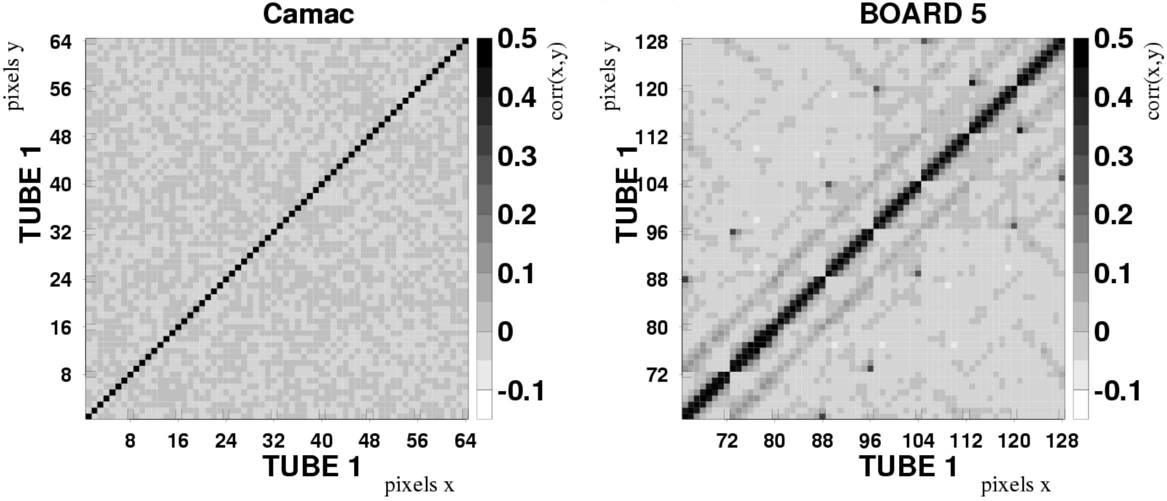


The plot entries are the probabilities that a hit in pixel x on the horizontal is "caused by" a hit in pixel y on the vertical axis. The cross talk in board 5 is symmetric. The pair of tubes read out by board 9 show, in addition to that, asymmetric cross talk within the tubes and also between the tubes.

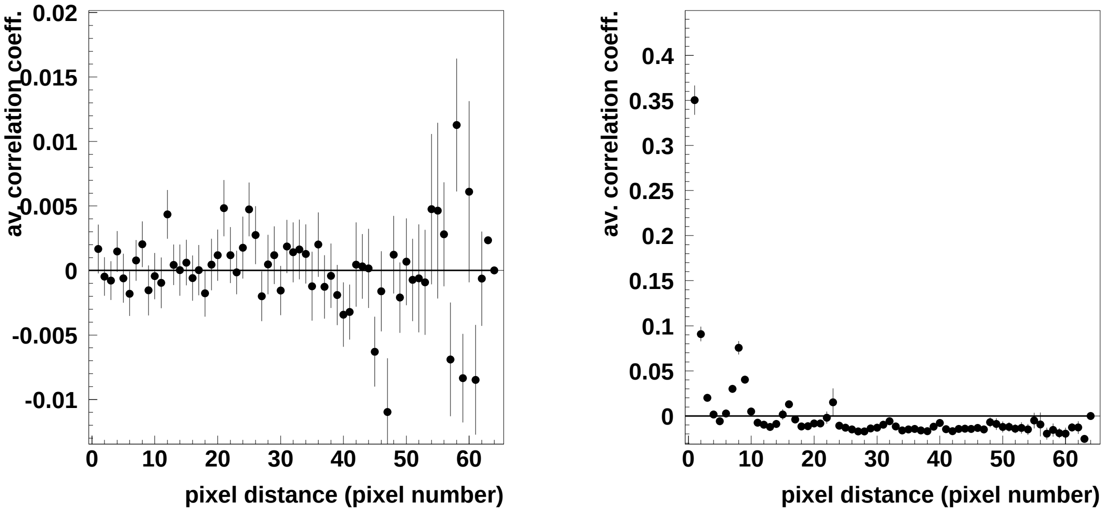


Average cross talk probability as a function of pixel-distance in terms of pixel-number

Figure 5.21: Correlation coefficients between the pulse heights of pixels within one tube for CAMAC readout (left) and fast readout (right).

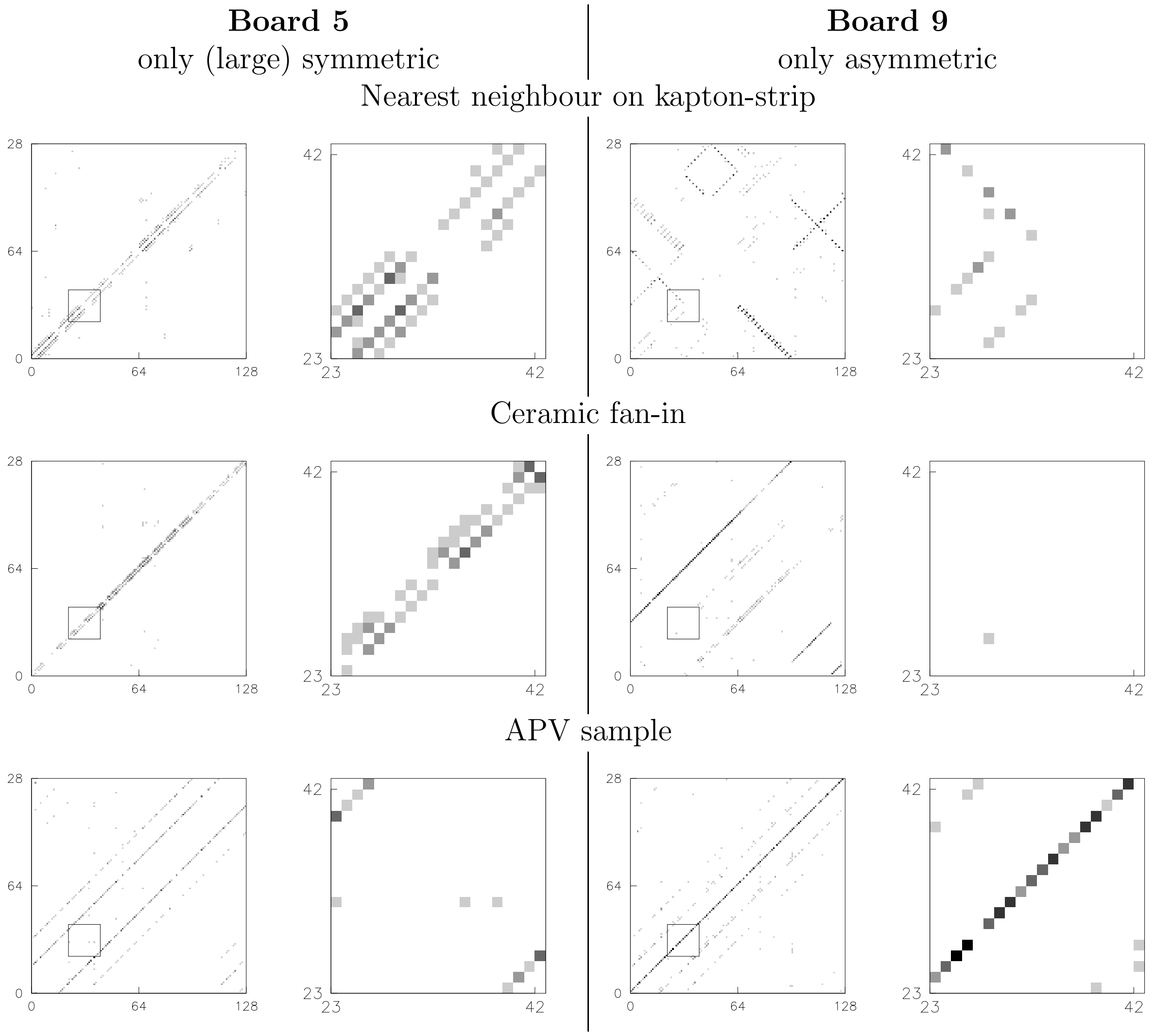


Correlation coefficients for all pixel pairs. For better readability, the scale on the 2-D plots has been reduced to a maximum of 0.5.



Mean correlation coefficients as a function of the pixel distance in terms of pixel numbers.

Figure 5.22: Re–numbering channels:
The plots show cross talk probabilities for two boards, for different channel numberings. The plots for board 5 show only large symmetric cross talk. The small symmetric cross talk (type 1a) is also present, but is not visible on this scale. Board 9 exhibits both symmetric and asymmetric cross talk, although in this plot, the former contribution is suppressed so that only the large asymmetric cross talk can be seen. The plots shows that the large symmetric cross talk is between neighbouring channels in the ceramic fan–in, while the asymmetric cross talk is between channels with consecutive APV sample numbers. No cross talk is seen between neighbouring channels on the same kapton strip.



In order to identify the cross talk sources, plots as in figure 5.20 are reproduced with changed pixel numberings. The pixels are re-numbered such that physically neighbouring pixels in some part of the readout chain are assigned consecutive numbers, so that they become neighbours on the plot. If the cross talk shows up along $y = x \pm 1$ for one of these numbering schemes, then there is cross talk between channels that are physical neighbours in the corresponding part of the readout chain. This points towards the source of the cross talk. The procedure is illustrated in figure 5.22 for three numbering schemes. The first row of plots the figure 5.22 shows no cross talk between neighbouring channels on the same kapton strip. The second one reveals the origin of the large symmetric cross talk, and the third one, of the asymmetric cross talk.

The following numbering schemes reveal cross talk:

- APV input number. This corresponds to the channel ordering in the ceramic coupler.
- APV sample number. This corresponds to channels that are neighbours in the APV analogue data frame, as shown in figure 5.14, page 160. It is the origin of the asymmetric cross talk.
- Pixel number, as defined in figure 5.8. This corresponds to the actual pixel-position on the readout board, independent of the orientation of the tube.

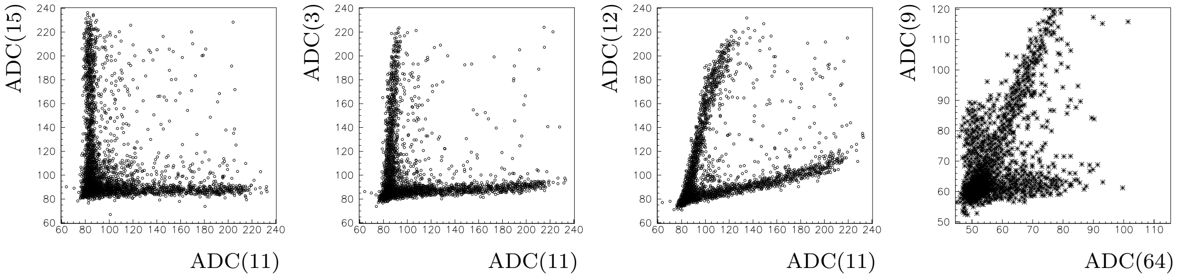
The different types of cross talk, with the sources and respective strengths, are listed in table 5.6, and are discussed below. In the table, the strength of each type is indicated by the average ratio of the cross talk pulse height to the pulse height of the signal that induces the cross talk.

Table 5.6: Four types of cross talk. The last column indicates the strength of the cross talk. It gives the average ratio of the cross talk pulse height to the pulse height of the signal that induced the cross talk.

	Source	Relationship	Ratio
Symmetric cross talk			
1a	AC–coupler	adjacent channels	0.15
1b	Kapton cables and PGA	± 8 pixel numbers	0.011
Asymmetric cross talk			
2a	APVm ASIC	Output sample $y \rightarrow y - 1$	0.33
2b	APVm ASIC	Output sample $y \rightarrow y + 2$	0.034

Scatter plots relating the pulse heights in two pixels are shown in figure 5.23, for pixel pairs exhibiting different types of cross talk.

Symmetric Cross Talk The symmetric cross talk is present in all boards to a similar level. The large symmetric cross talk (type 1a in table 5.6) is seen between nearest neighbours in the ceramic fan in. It is responsible for all large cross talk seen in figure 5.19 in board 5. This type of cross talk can be seen mainly along $y = x \pm 1$, but also in the series of cross talking pixel pairs scattered between $y = x \pm 8$ and $y = x \pm 32$.



No cross talk between pixels 11 and 15 of tube 4

Small Symm. cross talk (±8) between pixels 11 and 3 of tube 4

Large Symm. cross talk, between pixels 11 and 12 of tube 4

Asymm. cross talk pixels 64 of tube 2 talks to pixel 9 of tube 2

Figure 5.23: Scatter plots showing the ADC counts in pixel x vs ADC counts in pixel y for different pairs of pixels, showing either no, small symmetric, large symmetric or asymmetric cross talk.

It is also responsible for cross talk between tubes, for example the strong cross talk between pixel 8 in tube 1 and pixel 16 in tube 4 that can be seen figure 5.19.

The small symmetric cross talk (type 1b) is seen between pixels with pixel numbers x and $x \pm 8$. Possible locations for the small symmetric cross talk are either the pin–grid–array, or the kapton cables. In the kapton cables it would not be between neighbouring lines on the same kapton strip, but between lines on different layers of the kapton cable. This is possible, because a small area of the back of each kapton strip is not covered by a metal ground line.

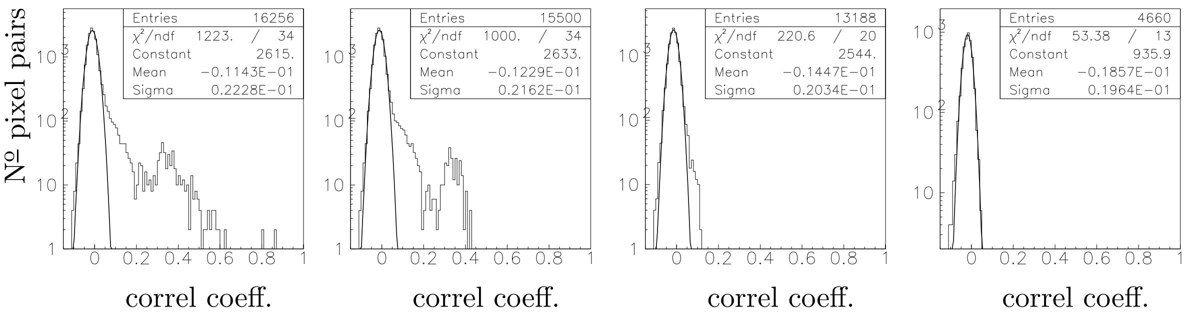
It is important to note that, although the small symmetric cross talk between pixel numbers x and $x \pm 8$ also corresponds to nearest neighbours in the tubes themselves, its cause cannot lie within the tube. This has been demonstrated in figure 5.21. This important conclusion has been corroborated in laboratory-tests carried out using the LED scanning facilities at Oxford and Edinburgh University [A^{+}00b].

Asymmetric Cross Talk Large Asymmetric cross talk (type 2a) is between nearest neighbours in the APV output data frame. The pixel with APV sample number y induces cross talk into the pixel with APV sample number $y - 1$, but no signal is induced by $y - 1$ into y. It also has the unexpected property that there is no next–to–nearest neighbour cross talk. If the pixels x, y, and z are related by asymmetric cross talk in the following way

$$x \xleftarrow{\text{talks to}} y \xleftarrow{\text{talks to}} z,$$

none of the signal induced by pixel z into pixel y is passed on to pixel x. This is the reason why the scatter plot in figure 5.23 for the asymmetric cross talk, looks as if the cross talk were sometimes “switched off”: the hits in pixel y (number 9) that do not

Figure 5.24: Correlation coefficients for different subsets of pixels-pairs (board 9)


all pixel-pairs

entries related by asymmetric cross talk removed

... and entries related by large symm. cross talk removed

... and entries related by small symm. cross talk removed

induce a signal into pixel x (number 64) are those induced by asymmetric cross talk from a third pixel.

The strength of the asymmetric cross talk varies greatly between boards, and in fact within boards. The numbers given in table 5.6 are for the worst cases of board 9 and board 12.

The small asymmetric cross talk (2b) is strongest in board 2, which corresponds to the ratio given in table 5.6. To a smaller extent, it can also be seen in board 5, where it is responsible for the small, but non–zero correlation coefficients between pixels in different tubes, that can be identified in figure 5.19.

Identifying All Cross Talk: To ensure that all cross talk is found, the correlation coefficients of all pixel–pairs in a given board are histogramed. This is repeated, step by step removing entries from pixel pairs that are related by a known type of cross talk. This includes also pixel pairs, that are indirectly related. How far combinations of different types of cross talk are considered when removing entries from the histograms, depends on the strengths of the types of cross talk involved. The most indirect relationship considered is that of next–to–next–to–nearest neighbours.

When all cross talk is removed, the correlation coefficients should be symmetrically centred around zero. Figure 5.24 shows four plots of correlation coefficients between the channel–pairs on board 9, with a Gaussian fitted to each distribution. From the left to the right, more and more types of cross talk are considered. The last plot shows a good fit of a Gaussian to the data.

The fact that the mean of the Gaussian fit is 1σ below zero means that a hit in one pixel slightly reduces the pulse heights in all other pixels; the same effect can be seen in figure 5.21.

No pixel pair has a correlation coefficient above the range described by the Gaussian fit. This shows, that all cross talk is found, that is large enough to be observable with the methods used.

Cross Talk Conclusion

The cross talk seen in the testbeam data is very large; for the asymmetric cross talk seen, typically 33% of the signal seen in a channel is induced to its cross-talk partner; for the large symmetric cross talk, this number is 15%. Each channel is directly related to 4 or 5 other channels by significant cross talk.

All cross talk can be related to parts of the readout system chain. Note that the readout electronics described here are not intended to be used in the final design.

As far as the suitability of the MaPMT for the LHCb RICH is concerned, the cross talk study confirms the good performance of the tubes. For the analysis of the testbeam data though, it will be essential to take the effects of the cross talk into account.

5.5 Monte Carlo Simulation

The same Monte Carlo simulation as for the HPD testbeam is used here, with a few modifications that adapt it to the different setup.

- Lenses and the quartz window have been added.
- The quantum efficiency and the mirror reflectivity are changed to those measured for the MaPMT and the 1117 mm focal length mirror. Those are shown in figures 5.3 and 5.2.
- The different detector geometry, including the detector–plane with the MaPMT cluster, is taken into account.

The reflection losses at the lens surfaces and the quartz window are calculated assuming normal incidence. Photons hitting the sides of the lenses are treated as lost. The MaPMT pixels are modelled as $2 \times 2\,\mathrm{mm}^2$ pixels with uniform efficiency, separated by 3 mm gaps with zero efficiency. The beam divergence in the simulation is 0.125 mrad in x and in y.

A difference with respect to the HPD analysis is that the signal loss estimates are not added back into the data, but are instead subtracted from the Monte Carlo results.

5.6 Photon Counting

5.6.1 Principle

Photon counting is performed pixel by pixel. The pedestal mean (μ) and width (σ) are found from the common-mode corrected pulse height spectra. A threshold cut is defined at $\mu + 5\sigma$; any ADC count above that threshold is considered a single hit.

The most important differences to the photon counting procedure described in chapter 4 are:

- The data need to be corrected for cross talk.
- $N_{hits} \approx N_{pe}$ is not a good approximation anymore, since the number of photoelectrons per pixel per event is larger. N_{pe} is therefore calculated from N_{hits}.
- The photon counts in this chapter refer to the number of photo electrons above the threshold cut, uncorrected for signal loss.
- The background estimate is much simplified. Electronics noise is negligible with a 5σ cut. Pedestal runs show very small beam–independent background. As most pixels in each tube lie outside the Cherenkov ring, the remaining background can be estimated tube by tube by averaging over the photon counts in those pixels, without the need for a fit.

5.6.2 The Number of Photoelectrons.

While in chapter 4, the emphasis lay more on the performance of the RICH detector as a whole, rather than that of the commercial HPDs, the focus of this test–beam evaluation the photodetector system, comprising the lenses, the readout electronics, and the tubes themselves. It is therefore more appropriate, to quote the photon counts uncorrected for signal loss, as this is the parameter of interest in assessing the suitability of the photodetectors for LHCb. Signal loss is however taken into account in the data–Monte Carlo comparison by correcting the Monte Carlo simulation using signal loss estimates obtained from the data.

In this analysis, the number of photoelectrons per event per pixel is too large for the effect of double photoelectron hits to be ignored. Therefore, the approximation $N_{hits} \approx N_{pe}$ is not used and the number of photoelectrons is calculated from the number of hits as described below.

Notation and Definitions

The following notation is used throughout the text (all numbers refer to photon counts in a single pixel):

$\lambda_{\text{p.e.}}$	Mean number of photoelectrons per event generated in the cathode.
$\lambda_{5\sigma}$	Mean number of photoelectrons/event detected above a 5σ cut.
$n_{5\sigma}$	Mean number of hits/event detected above a 5σ cut, counting single and multiple photoelectrons as a single hit
$\lambda'_{5\sigma}$	The approximation to $\lambda_{5\sigma}$ used throughout the text.

As shown in appendix B, the parameter $\lambda'_{5\sigma}$:

$$\lambda'_{5\sigma} \equiv -\ln(1 - n_{5\sigma}) \qquad (5.39)$$

gives a very good approximation to $\lambda_{5\sigma}$, the mean number of photoelectrons detected above a 5σ cut. It has the advantage that it can directly be calculated from the observed mean number of hits above the cut, $n_{5\sigma}$, independent of signal loss estimates. At the same time, the error introduced is very small, less than $\sim 3\%$ for any given pixel in this analysis.

In the following, all photon counts will be given in terms of $\lambda'_{5\sigma}$, including those calculated from the Monte Carlo simulation. If the integrated number of photoelectrons for a pixel is given, rather than the mean number per event, it is calculated by multiplying $\lambda'_{5\sigma}$ with the number of events N_{events}:

$$N_{5\sigma\ \text{pe}} = N_{\text{events}} \lambda'_{5\sigma} \tag{5.40}$$

5.6.3 Correcting cross talk

For photon counting in the 3×3 array the presence of cross talk has to be corrected for. Two methods are used to achieve this:

a) Pixel Hit Correction

Removing Cross Talk: Event by event, each pixel hit is only retained if none of the associated cross–talk partners, identified by the method described in section 5.4.2, has a higher pulse height. A cross talk partner to pixel x is here defined as a pixel that can induce cross talk to pixel x.

Recovering Losses: Losses due to the rejection of genuine double hits in the cross talk correction are recovered at the end of the procedure. For this, the probability for each pixel to lose a hit in the above correction procedure due to a genuine hit in one of its cross talk partners is calculated. This is then used, pixel by pixel, to correct the photon counts.

To calculate this probability, it is assumed that, when the two cross talk related pixels x and y have a genuine photon hit, the probability that the pulse height in pixel x, h_x, is greater than the pulse height h_y in its cross talk partner y, is $p(h_y > h_x) = \frac{1}{2}$.

With this, the probability that pixel x loses a hit in the cross talk correction procedure due to a genuine hit in one of its cross talk partners y_i, is:

$$P(x\ \text{loses hit}) = 1 - \prod_{\text{all xtalk partners}} \left(1 - \frac{1}{2} n_{5\sigma,i}\right) \tag{5.41}$$

where $n_{5\sigma,i}$ are the mean number of hits per event in pixels i. The product goes over all cross–talk partners of x, identified as described in section 5.4.2. The result of the product–term in the above equation is the total probability that none of the cross talk partners has a hit with a larger pulse height than x; subtracting that from 1 gives the probability that any of them has a larger pulse height.

To perform this calculation, the true values for $n_{5\sigma,i}$ are needed, but only the values before the losses are recovered are known. Therefore, an iterative method is used, in which the corrected values for $n_{5\sigma,i}$ calculated in one iteration are the input for the probability calculation in the next iteration. This iteration process converges rather quickly and is stopped when the relative difference in the photon counts between two successive iterations is less than 10^{-6}.

For the Cherenkov data presented below, this recovery procedure adds 9% to the photon counts, averaged over all pixels.

Errors: The pixel hit correction method only requires a complete list of pixels that are related by cross talk. It is independent of the particular type and intensity of the cross talk, and is therefore unaffected by the associated uncertainties.

However, there is a small uncertainty in the resulting photon counts associated with assumption that $p(h_x > h_y) = \frac{1}{2}$ for all pixel pairs. The combined photon count for each pixel pair related by symmetric cross talk is rather insensitive to which value of $p(h_x > h_y)$ is chosen, as a higher count in pixel x is balanced by a correspondingly lower count in pixel y, because $p(h_x > h_y) = 1 - p(h_y > h_x)$. However, this cancellation does not take place for the case of asymmetric cross talk. For asymmetric cross talk, an error in $p(h_x > h_y)$ of Δp will result in an error in the corrected value for $n_{5\sigma,x}$ of

$$\sim \Delta p \cdot n_{5\sigma,x} \cdot n_{5\sigma,y}. \tag{5.42}$$

The difference in the photon counts obtained using $p(h_x > h_y) = 0$ and $p(h_x > h_y) = 1$ for the asymmetric cross talk, compared with $p(h_x > h_y) = \frac{1}{2}$, is less than 1% for the Cherenkov ring, while it is $\sim$ 6% for LED data. The small value for the Cherenkov data is due to the fact that the pixels related by asymmetric cross talk are not spatially related on the detector plane, so that in most cases, at least one of the pixels x, y is off the ring and $n_{5\sigma,x} n_{5\sigma,y}$ is a small number.

b) Pulse height Correction:

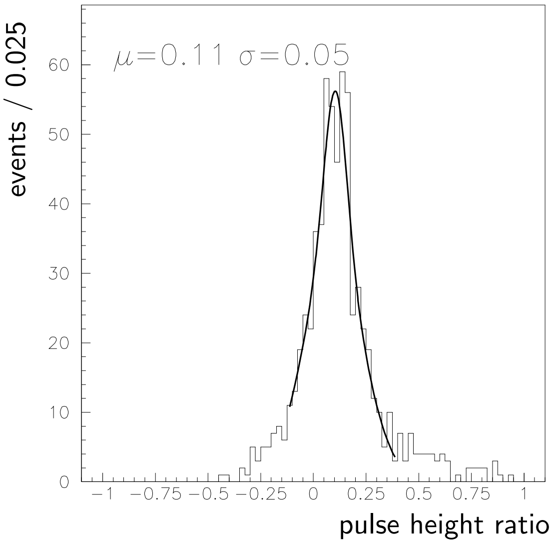
The pulse height in each pixel in each event is corrected for cross talk. This produces the corrected pulse height spectra used later for fits to the spectra, and the signal loss estimates. This method allows event–by–event cross talk correction, while method (a) relies on statistical methods to recover lost hits.

To correct the pulse height in a given pixel x, a fraction of the pulse height found in the same event in each of its cross talk partners is subtracted from the pulse height in x.

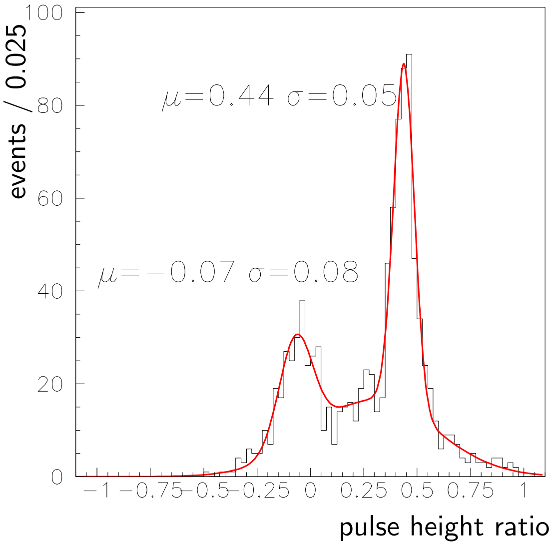
To establish what fraction of the pulse height of a pixel is typically induced into each of the pixels it talks to, LED data runs are used. For each pixel pair related by cross talk, the ratio of the pulse height in the talked–to pixel to the pulse height in the talking pixel is histogramed, as shown for two examples in figure 5.25. One pixel pair shown is related by symmetric cross talk, and one by asymmetric cross talk. For the symmetric cross talk, two such plots exist for each related pixel pair x, y, one where x is taken as the talking pixel, and one where y is taken as the talking pixel. Only those events are used for the plots where there is a hit in the talking pixel, and where the pulse height in the talking pixel is larger than that in the talked–to pixel.

Figure 5.25: Pulse height ratios: The pulse height in the talked–to pixel is divided by the pulse height in the talking pixel.

a) symmetric xtalk



b) asymmetric xtalk



For the plots shown in figure 5.25, fits to the peaks in the histograms show a mean cross talk ratio of 0.11 and 0.44 for the symmetric and the asymmetric case respectively. The histogram for the asymmetric cross talk shows an additional peak around zero. This is because there is no next-to-nearest neighbour contribution to asymmetric cross talk: the peak around zero comes from events that were induced themselves into the talking pixel by asymmetric cross talk from a third pixel.

To obtain the mean cross talk ratios used in the pulse height correction one Gaussian is fitted to each peak, plus one broader Gaussian to take into account the background from genuine double hits. Only the mean and width from the fits to the peaks are given in figure 5.25.

Since there is no next–to–nearest neighbour cross talk in the case of the asymmetric cross talk, this adds a complication to the cross talk correction. Assuming that the pixels x, y, and z are related by asymmetric cross talk in the following way:

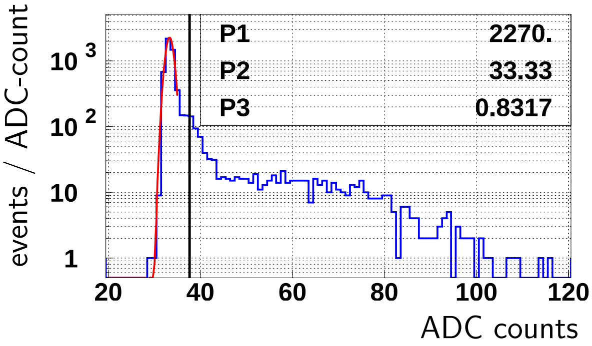
$$x \xleftarrow{\text{talks to}} y \xleftarrow{\text{talks to}} z,$$

it is necessary to know whether the hit in pixel y is itself induced by asymmetric cross talk from z, or not. If this is the case, is no subtraction from the pulse height in x is required. This is decided by comparing the pulse height in pixel y with that in z. If the ratio of the pulse heights in y and z is close to the typical ratio for cross talk induced by z into y, the pulse height in x is not corrected.

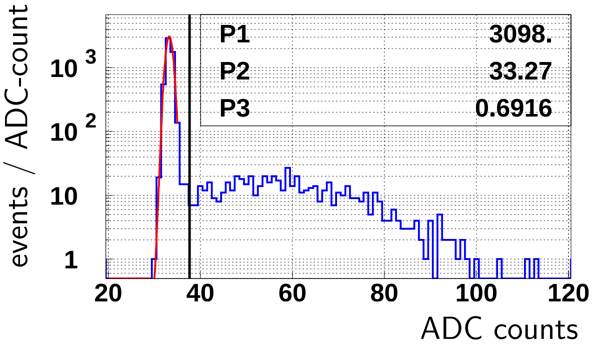
Figure 5.26 shows a pulse height spectrum before and after the event-by-event cross talk correction is applied.

Figure 5.26: Spectra before and after cross talk correction with method (b). The line indicates the 5σ cut. P1, P2, P3 are the amplitude, mean and sigma of the fit to the pedestal respectively. The threshold cut is defined using the results from the fit to the pedestal before the cross talk correction, and remains at the same place after the cross talk correction.

Before cross talk correction



After cross talk correction



Comparing the Methods

The two methods agree up to $\sim 7\%$ as seen in Table 5.8 (page 194). The largest difference (20%) is in tube 4, where the event by event cross talk correction is not completely successful. It is the tube that has by far the largest signal, and therefore requires also the largest pulse-height correction in method (b), while the pedestal width, and therefore the threshold cut, is the same as for the other tubes. This is illustrated in figure 5.27, that shows a pulse height spectrum from a pixel in tube 4 before and after cross talk correction.

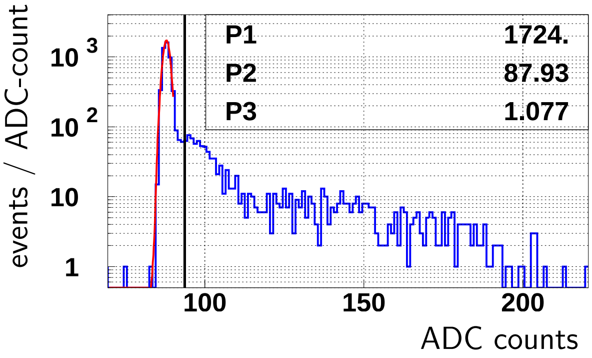
Because method (a) is robust against such problems, it has been chosen as the default method, against which the Monte Carlo results will be compared. Method (b) is still needed, as it provides the corrected pulse height spectra that can be fitted for the signal loss estimates. It also provides a useful systematic cross check.

5.6.4 Fitting Spectra

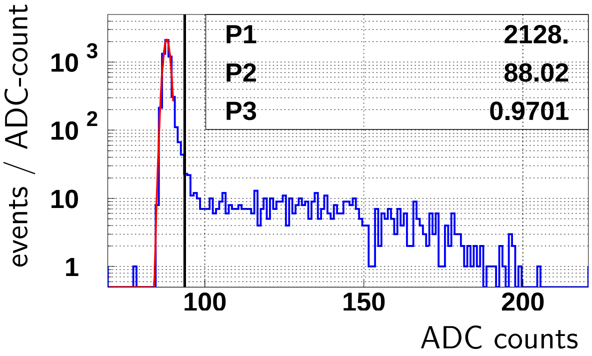
The pulse height correction method removes the cross talk distortion from the pulse height spectra, and therefore allows us to fit them with the function described in section 5.3.2. Successful fits to spectra of one pixel from each tube on the Cherenkov ring are shown figure 5.28, together with the

Figure 5.27: Spectra before and after cross talk correction with method (b), in tube 4. The parameters P1, P2, P3 are defined as in figure 5.26. The signal from tube 4, and in particular the cross talk, is much larger, and method (b) is not so reliable.

Before cross talk correction (pixel 35 in tube 4)



After cross talk correction (pixel 35 in tube 4)



fit results for the gain at the first dynode and the mean number of photoelectrons. Not all fits are successful however. An example for a spectrum that cannot be fitted, because it is still distorted after the pulse height correction, is shown in figure 5.27.

In order to minimise the sensitivity to residual cross talk distortions, only Cherenkov data are used for the fits. Results for one pixel from each tube on the Cherenkov ring are shown in figure 5.28.

The values obtained for the gain at the first dynode are compatible with that presented in [RIC00]. That the spectra can be described by the fit function indicates, that the full dynamical range of the MaPMT output can be captured within the 25 ns available for the readout.

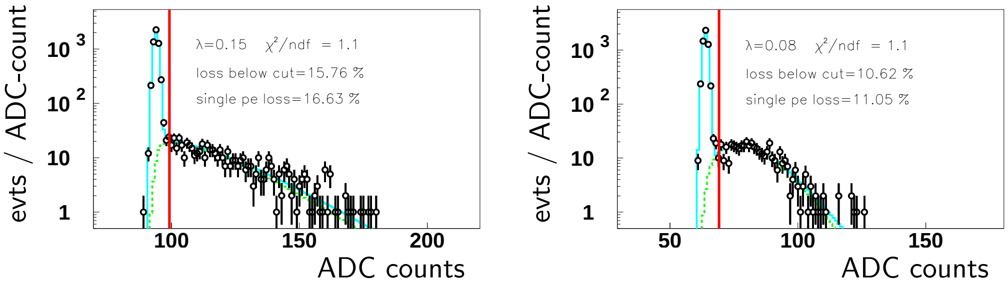
Signal Loss

From the fits to the spectra, the fraction of photoelectrons that remain undetected can be estimated.

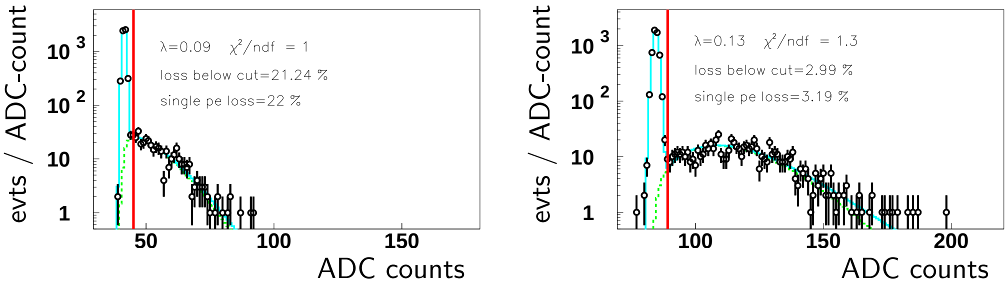
There are two sources of signal loss: the flow of photoelectrons terminates at some stage in the dynode chain, due to the probability of e^{-g_i} that an electron arriving at dynode i with gain g_i produces 0 secondaries. This signal loss is dominated by the probability that a single photoelectron arriving at the first dynode does not produce any secondaries, e^{-g_1}. The second source of signal loss is the loss due the pulse height being below the threshold cut. g_1 is a parameter of the fit, and the number of hits lost below the threshold cut can also be esti-

Figure 5.28: Fitted pulse height spectra for each tube on the Cherenkov ring. From 6000 events taking Cherenkov data without lenses. Shown are successful fits to spectra from pixels on the ring. The signal loss figures inside the plots do not include the loss at the first dynode. (From run number 2610).

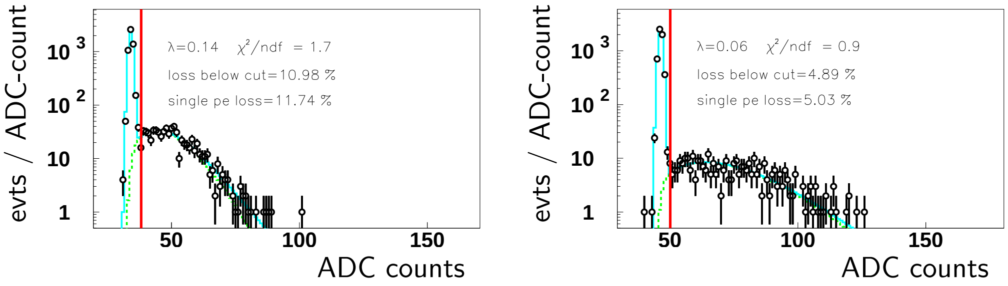
Tube 1 pixel 6, $\lambda_{\text{p.e.}} = 0.15$, $g_1 = 2.3$ Tube 2, pixel 27, $\lambda_{\text{p.e.}} = 0.08$, $g_1 = 4.1$



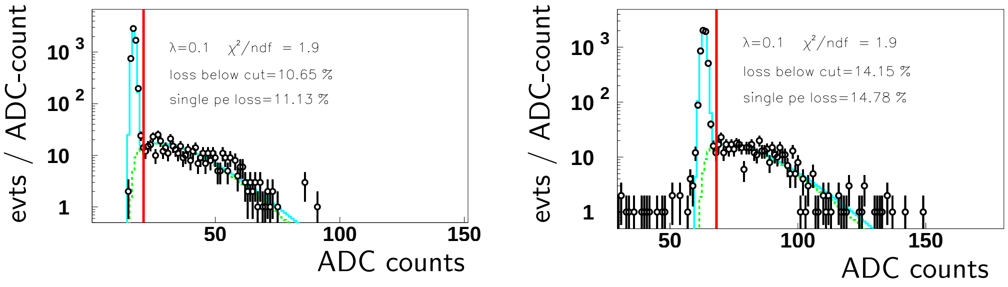
Tube 3 pixel 55, $\lambda_{\text{p.e.}} = 0.09$, $g_1 = 2.6$ Tube 4, pixel 34, $\lambda_{\text{p.e.}} = 0.13$, $g_1 = 4.4$



Tube 6 pixel 25, $\lambda_{\text{p.e.}} = 0.14$, $g_1 = 3.8$ Tube 7, pixel 9, $\lambda_{\text{p.e.}} = 0.06$, $g_1 = 3.7$



Tube 8 pixel 14, $\lambda_{\text{p.e.}} = 0.10$, $g_1 = 3.1$ Tube 9, pixel 41, $\lambda_{\text{p.e.}} = 0.10$, $g_1 = 3.0$



mated from the shape of the signal distribution that can be obtained from the fit. Both types of signal loss can therefore be estimated separately.

In order to obtain a figure that is independent of the number of photoelectrons per event in the particular run that is used for the estimates, it is the loss of single photoelectron events that is calculated, rather than the total loss. The loss of multiple photo electron events can be neglected.

A sufficient number of hits is needed to determine the signal shape, therefore only spectra from pixels with at least 0.05 photoelectrons per event are selected. Only fits with a χ^2/ndf of less than 2 are used in the signal loss estimates. The signal loss estimates from all successful fits are averaged tube by tube.

Table 5.7 shows the result of the signal loss estimates. These signal loss estimates are used to correct the simulation. The uncertainty in the signal loss introduces an error on the Monte Carlo result for the photon count per ring of ∼ 3%. This estimate is obtained assuming the same number of hits in all tubes except tube 5, and calculating the error on the mean total loss for each tube by dividing the RMS of the fit-results by the square–root of the number of pixels that contribute to that result.

Table 5.7: Signal loss estimates. Shown are the mean and the RMS of all successful fits to pulse height spectra from pixels with more then 0.05 p.e./event, from two runs, one with and one without lens.

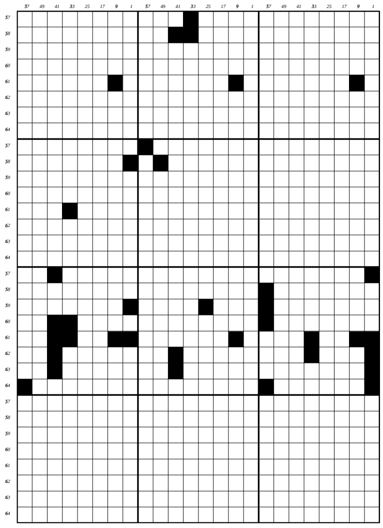
Tube	№ of pixels	single p.e. loss below cut (%)	Gain 1st dyn	Total single p.e. loss (%)
1	3	8 ± 4	3.2 ± 0.5	13 ± 7
2	5	12 ± 5	3.9 ± 0.4	14 ± 6
3	3	14 ± 2	3.2 ± 0.4	18 ± 4
4	2	5 ± 2	4.0 ± 0.4	7 ± 2
5	0			
6	10	7 ± 3	3.8 ± 0.8	10 ± 7
7	4	14 ± 4	3.2 ± 0.4	17 ± 6
8	14	11 ± 3	3.4 ± 0.5	15 ± 5
9	2	11 ± 3	3.7 ± 0.7	14 ± 5

5.6.5 Further Corrections

Dead and Noisy Pixels

An LED run is used to identify dead channels. Laboratory scans of the tubes show that these dead channels are in the electronic read–out and not the tubes. Also, noisy channels which contain a very large number of hits after the cross talk correction are identified; these are mostly due to a failed fit to the pedestal, giving too small a pedestal mean. Both dead and noisy channels are ignored in the photon counts, in the data as well as in the Monte-Carlo simulation. Altogether 37 channels are removed; only 5 and 7 of these correspond to pixels on the Cherenkov ring with and without lenses, respectively.

Figure 5.29: Masked Pixels.The plot shows the pixels at their position on the detector plane. The gaps between the active areas of the tubes are not shown. All channels that are read out are displayed, including those in the bottom 8 rows that are not connected to a photodetector.



Background

The sources of background were described in chapter 4. Two of the three types of background described there, are negligible in this analysis:

- For a 5σ cut, electronic noise is negligible
- Pedestal runs show, that non-Gaussian, beam–independent background is very small.

The remaining background is estimated assuming a flat distribution on the detector surface and averaging, tube by tube, over all pixels that lie outside the Cherenkov ring.

5.6.6 Photon Counts in Data and Monte Carlo

Figures 5.31 and 5.33 show the photon counts for each pixel, integrated over 6000 events, for the case with and without lenses respectively, for cross talk correction method (a). Figures 5.32 and 5.34 show the data display for the corresponding Monte Carlo simulation, for ten times as many events. The data display before cross talk correction for the run with lenses is shown in figure 5.30.

A complete summary of photon counting results for the 3×3 array is given in table 5.8. The photoelectron counts and background estimates for the case with and without lenses are given, in the case with lenses for both cross talk correction methods.

The photon counts in all figures are given in terms of

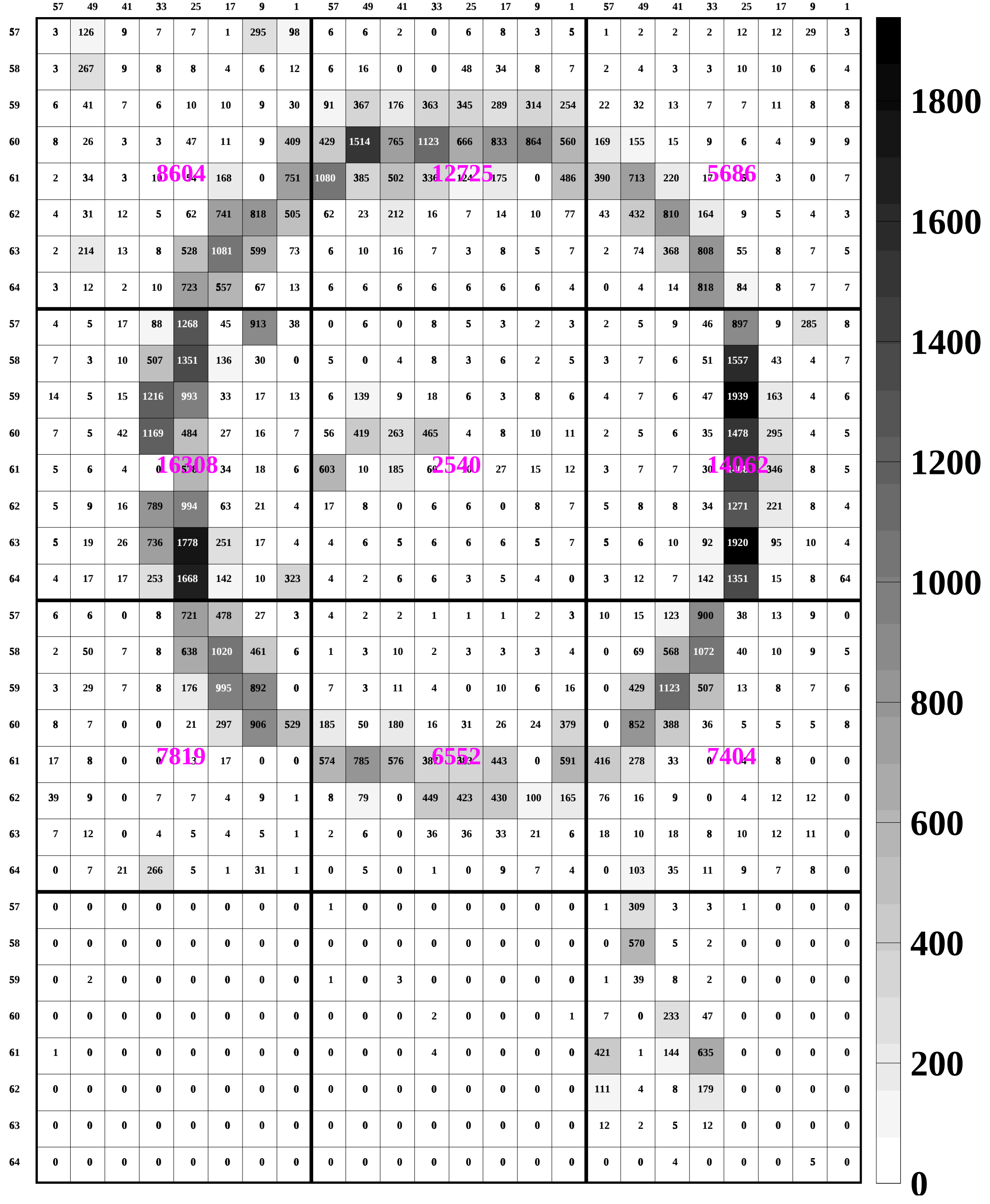
$$N_{5\sigma\ \mathrm{pe}} = N_{\mathrm{events}} \lambda'_{5\sigma}$$

as defined in equations 5.39 and 5.40, i.e. the number of photoelectrons above the 5σ cut, corrected for multiple photoelectron events. The results are compared to the results from the Monte Carlo simulation. The same set of pixels is masked in data and the simulation.

5.6.7 Error Estimates

The large number of events (6000 per run) leads to the statistical errors being negligible. Systematic errors have been estimated

Figure 5.30: Photon counts, integrated over 6k events, for the run with lenses, **without cross talk correction**. The total photon count is 88 489. The pixels are shown in their position on the detector plane, as seen from the mirror. The bottom 8 rows of pixels correspond to channels that are not connected to a photodetector. The total photon count given in this caption includes those rows, while the final results given in table 5.8 do not. The photon counts per pixel, and, in larger figures, per tube are given. The gaps between the active areas of the tubes are not shown. The run number is 2634



	57	49	41	33	25	17	9	1	57	49	41	33	25	17	9	1	57	49	41	33	25	17	9	1
57	3	126	9	7	7	1	295	98	6	6	2	0	6	8	3	5	1	2	2	2	12	12	29	3
58	3	267	9	8	8	4	6	12	6	16	0	0	48	34	8	7	2	4	3	3	10	10	6	4
59	6	41	7	6	10	10	9	30	91	367	176	363	345	289	314	254	22	32	13	7	7	11	8	8
60	8	26	3	3	47	11	9	409	429	1514	765	1123	666	833	864	560	169	155	15	9	6	4	9	9
61	2	34	3	[illegible]	[illegible]	168	0	751	1080	385	502	336	[illegible]	175	0	486	390	713	220	[illegible]	[illegible]	3	0	7
62	4	31	12	5	62	741	818	505	62	23	212	16	7	14	10	77	43	432	810	164	9	5	4	3
63	2	214	13	8	528	1081	599	73	6	10	16	7	3	8	5	7	2	74	368	808	55	8	7	5
64	3	12	2	10	723	557	67	13	6	6	6	6	6	6	6	4	0	4	14	818	84	8	7	7
57	4	5	17	88	1268	45	913	38	0	6	0	8	5	3	2	3	2	5	9	46	897	9	285	8
58	7	3	10	507	1351	136	30	0	5	0	4	8	3	6	2	5	3	7	6	51	1557	43	4	7
59	14	5	15	1216	993	33	17	13	6	139	9	18	6	3	8	6	4	7	6	47	1939	163	4	6
60	7	5	42	1169	484	27	16	7	56	419	263	465	4	8	10	11	2	5	6	35	1478	295	4	5
61	5	6	4	[illegible]	[illegible]	34	18	6	603	10	185	[illegible]	[illegible]	27	15	12	3	7	7	[illegible]	[illegible]	346	8	5
62	5	9	16	789	994	63	21	4	17	8	0	6	6	0	8	7	5	8	8	34	1271	221	8	4
63	5	19	26	736	1778	251	17	4	4	6	5	6	6	6	5	7	5	6	10	92	1920	95	10	4
64	4	17	17	253	1668	142	10	323	4	2	6	6	3	5	4	0	3	12	7	142	1351	15	8	64
57	6	6	0	8	721	478	27	3	4	2	2	1	1	1	2	3	10	15	123	900	38	13	9	0
58	2	50	7	8	638	1020	461	6	1	3	10	2	3	3	3	4	0	69	568	1072	40	10	9	5
59	3	29	7	8	176	995	892	0	7	3	11	4	0	10	6	16	0	429	1123	507	13	8	7	6
60	8	7	0	0	21	297	906	529	185	50	180	16	31	26	24	379	0	852	388	36	5	5	5	8
61	17	8	0	[illegible]	[illegible]	17	0	0	574	785	576	[illegible]	[illegible]	443	0	591	416	278	33	[illegible]	[illegible]	8	0	0
62	39	9	0	7	7	4	9	1	8	79	0	449	423	430	100	165	76	16	9	0	4	12	12	0
63	7	12	0	4	5	4	5	1	2	6	0	36	36	33	21	6	18	10	18	8	10	12	11	0
64	0	7	21	266	5	1	31	1	0	5	0	1	0	9	7	4	0	103	35	11	9	7	8	0
57	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	309	3	3	1	0	0	0
58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	570	5	2	0	0	0	0
59	0	2	0	0	0	0	0	0	1	0	3	0	0	0	0	0	1	39	8	2	0	0	0	0
60	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	1	7	0	233	47	0	0	0	0
61	1	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	421	1	144	635	0	0	0	0
62	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	111	4	8	179	0	0	0	0
63	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12	2	5	12	0	0	0	0
64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	5	0

Figure 5.31: Photon counts, integrated over 6k events, for a run with lenses, with cross talk correction (a). The photon count is 40 668. The pixels are shown in their position on the detector plane, as seen from the mirror. The bottom 8 rows of pixels correspond to channels that are not connected to a photodetector. The total photo count given in this caption includes those rows, while the final results given in table 5.8 do not. The photon counts per pixel are given, and, in larger figures, the photon counts per tube. The run number is 2634.

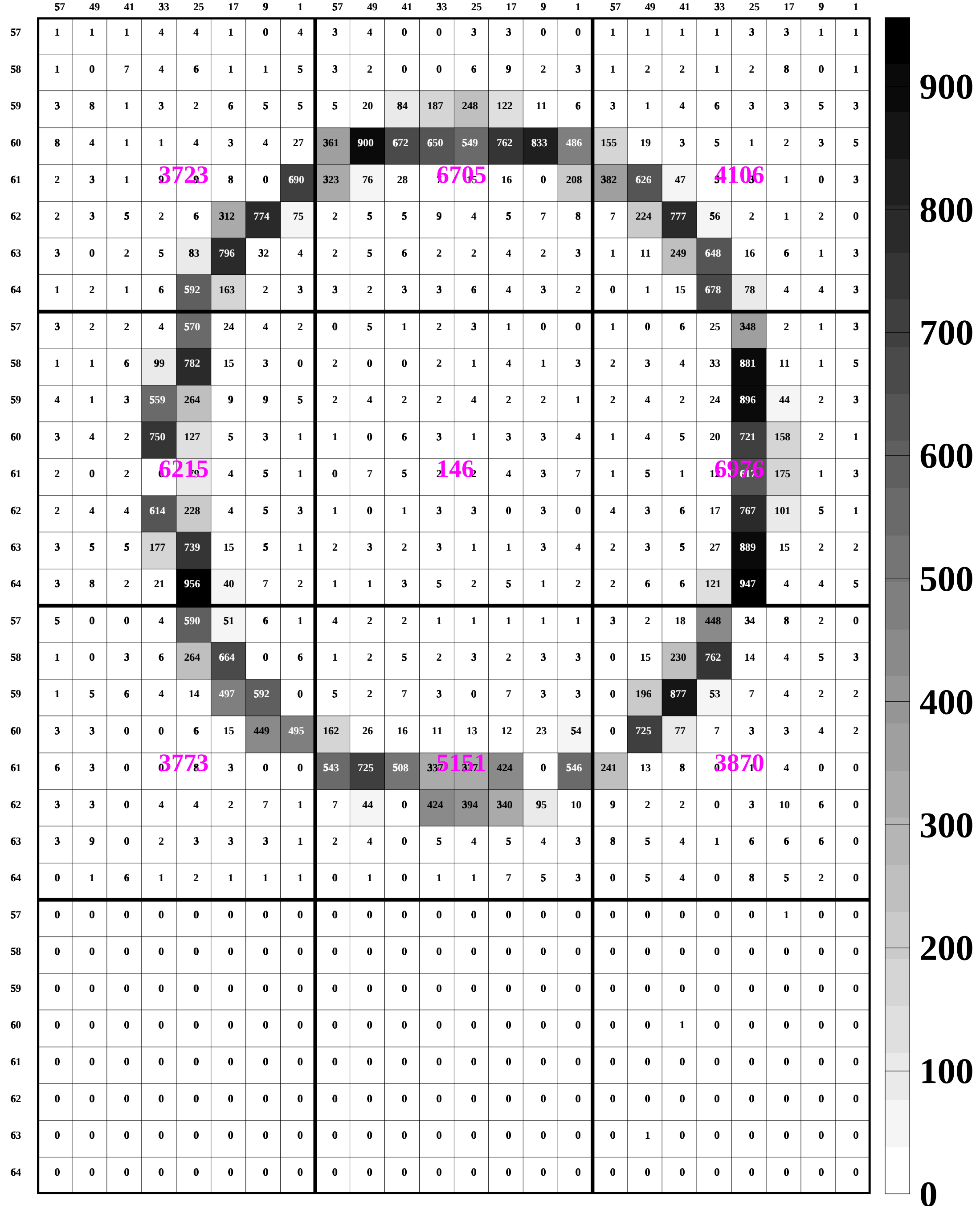
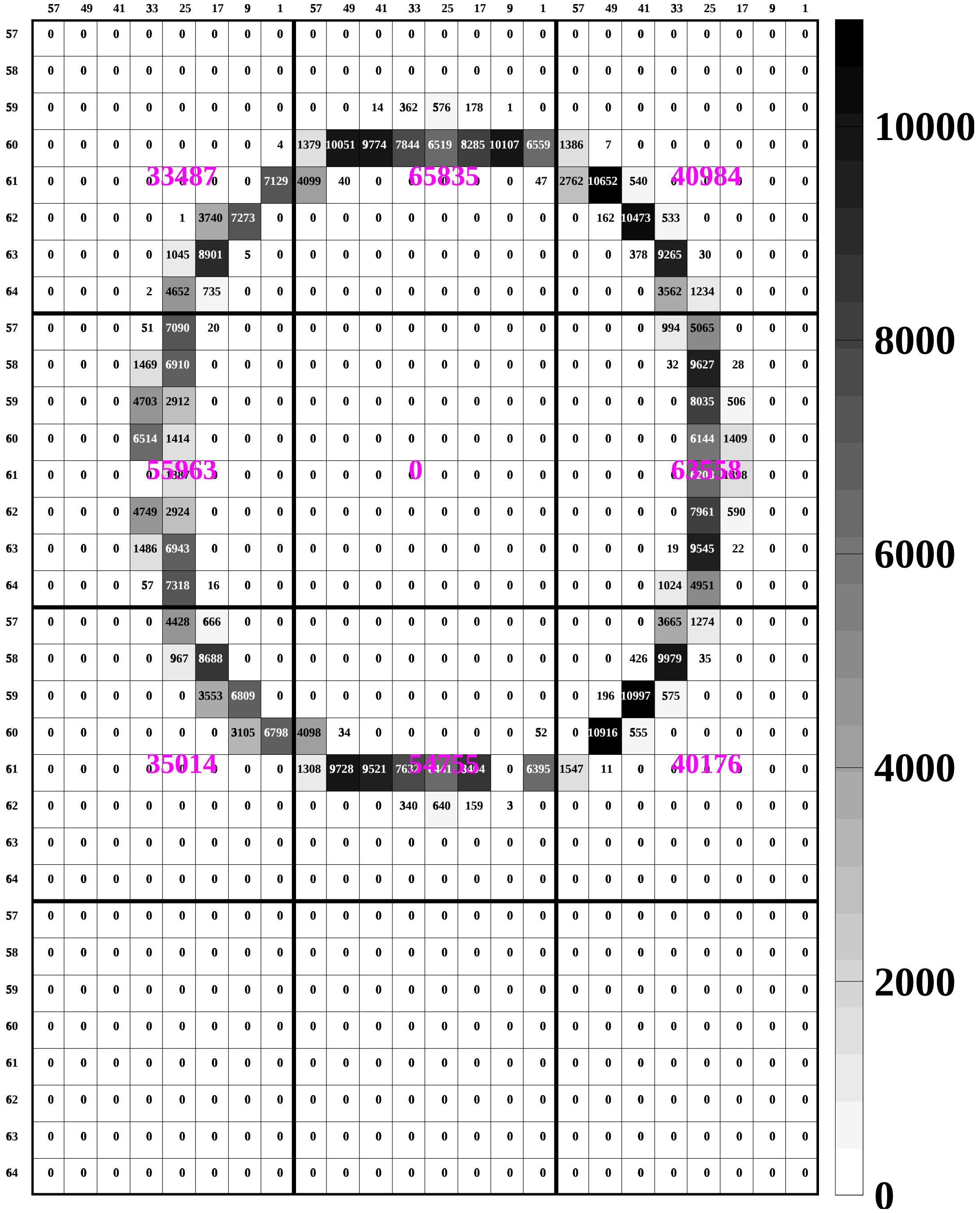
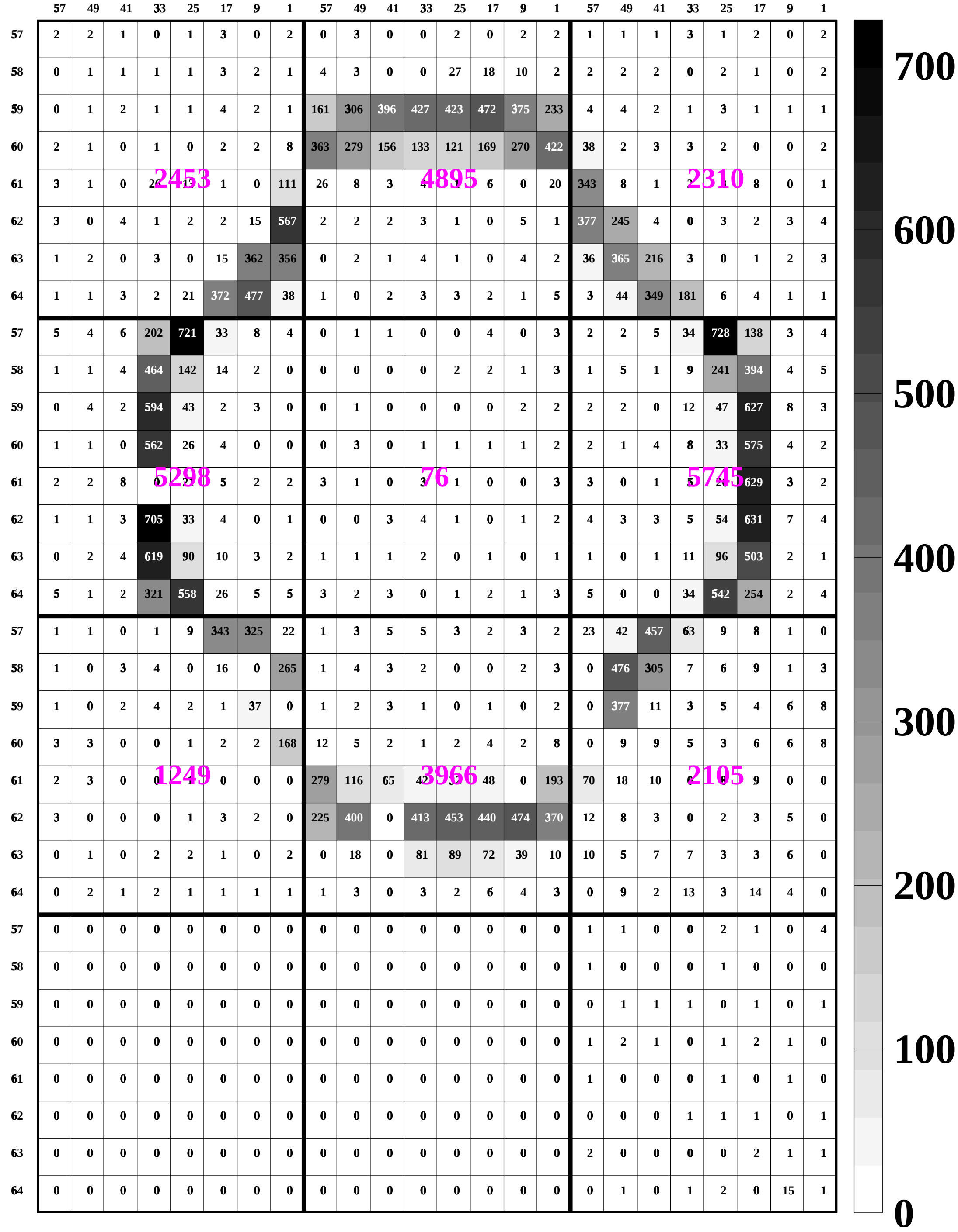


Figure 5.32: MC-simulated photon counts, for the case with lenses, integrated over 60k events, ten times as many as for the data. The total photon count is 389 772. The pixels are shown in their position on the detector plane, as seen from the mirror. The photon counts per pixel are given, and, in larger figures, the photon counts per tube.



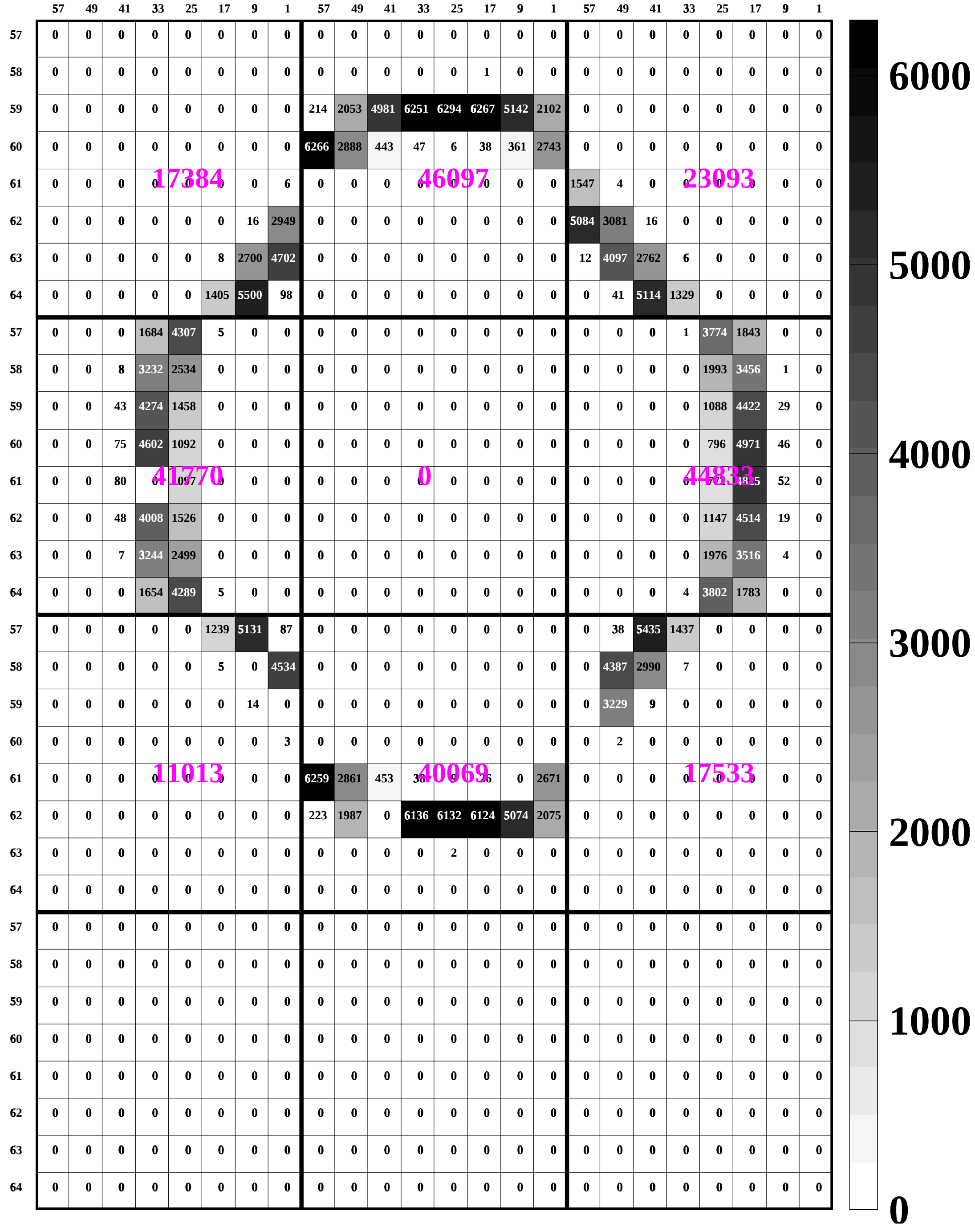
	57	49	41	33	25	17	9	1	57	49	41	33	25	17	9	1	57	49	41	33	25	17	9	1
57	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
59	0	0	0	0	0	0	0	0	0	0	14	362	576	178	1	0	0	0	0	0	0	0	0	0
60	0	0	0	0	0	0	0	4	1379	10051	9774	7844	6519	8285	10107	6559	1386	7	0	0	0	0	0	0
61	0	0	0	0	0	0	0	7129	4099	40	0	0	0	0	0	47	2762	10652	540	0	0	0	0	0
62	0	0	0	0	1	3740	7273	0	0	0	0	0	0	0	0	0	0	162	10473	533	0	0	0	0
63	0	0	0	0	1045	8901	5	0	0	0	0	0	0	0	0	0	0	0	378	9265	30	0	0	0
64	0	0	0	2	4652	735	0	0	0	0	0	0	0	0	0	0	0	0	0	3562	1234	0	0	0
57	0	0	0	51	7090	20	0	0	0	0	0	0	0	0	0	0	0	0	0	994	5065	0	0	0
58	0	0	0	1469	6910	0	0	0	0	0	0	0	0	0	0	0	0	0	0	32	9627	28	0	0
59	0	0	0	4703	2912	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8035	506	0	0
60	0	0	0	6514	1414	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6144	1409	0	0
61	0	0	0	0	1387	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6208	1398	0	0
62	0	0	0	4749	2924	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7961	590	0	0
63	0	0	0	1486	6943	0	0	0	0	0	0	0	0	0	0	0	0	0	0	19	9545	22	0	0
64	0	0	0	57	7318	16	0	0	0	0	0	0	0	0	0	0	0	0	0	1024	4951	0	0	0
57	0	0	0	0	4428	666	0	0	0	0	0	0	0	0	0	0	0	0	0	3665	1274	0	0	0
58	0	0	0	0	967	8688	0	0	0	0	0	0	0	0	0	0	0	0	426	9979	35	0	0	0
59	0	0	0	0	0	3553	6809	0	0	0	0	0	0	0	0	0	0	196	10997	575	0	0	0	0
60	0	0	0	0	0	0	3105	6798	4098	34	0	0	0	0	0	52	0	10916	555	0	0	0	0	0
61	0	0	0	0	0	0	0	0	1308	9728	9521	7632	6441	8404	0	6395	1547	11	0	0	0	0	0	0
62	0	0	0	0	0	0	0	0	0	0	0	340	640	159	3	0	0	0	0	0	0	0	0	0
63	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
57	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
59	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
62	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
63	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 5.33: Photon counts, integrated over 6k events, for the run without lenses, with cross talk correction (a). The total photon count is 28 153. The pixels are shown in their position on the detector plane, as seen from the mirror. The bottom 8 rows of pixels correspond to channels that are not connected to a photodetector. The total photon count given in this caption includes those rows, while the final results given in table 5.8 do not. The photon counts per pixel, and, in larger figures, per tube are given. The gaps between the active areas of the tubes are not shown. The run number is 2610.



	57	49	41	33	25	17	9	1	57	49	41	33	25	17	9	1	57	49	41	33	25	17	9	1
57	2	2	1	0	1	3	0	2	0	3	0	0	2	0	2	2	1	1	1	3	1	2	0	2
58	0	1	1	1	1	3	2	1	4	3	0	0	27	18	10	2	2	2	2	0	2	1	0	2
59	0	1	2	1	1	4	2	1	161	306	396	427	423	472	375	233	4	4	2	1	3	1	1	1
60	2	1	0	1	0	2	2	8	363	279	156	133	121	169	270	422	38	2	3	3	2	0	0	2
61	3	1	0	26	[illegible]	1	0	111	26	8	3	[illegible]	[illegible]	6	0	20	343	8	1	[illegible]	[illegible]	8	0	1
62	3	0	4	1	2	2	15	567	2	2	2	3	1	0	5	1	377	245	4	0	3	2	3	4
63	1	2	0	3	0	15	362	356	0	2	1	4	1	0	4	2	36	365	216	3	0	1	2	3
64	1	1	3	2	21	372	477	38	1	0	2	3	3	2	1	5	3	44	349	181	6	4	1	1
57	5	4	6	202	721	33	8	4	0	1	1	0	0	4	0	3	2	2	5	34	728	138	3	4
58	1	1	4	464	142	14	2	0	0	0	0	0	2	2	1	3	1	5	1	9	241	394	4	5
59	0	4	2	594	43	2	3	0	0	1	0	0	0	0	2	2	2	2	0	12	47	627	8	3
60	1	1	0	562	26	4	0	0	0	3	0	1	1	1	1	2	2	1	4	8	33	575	4	2
61	2	2	8	0	21	5	2	2	3	1	0	3	1	0	0	3	3	0	1	5	[illegible]	629	3	2
62	1	1	3	705	33	4	0	1	0	0	3	4	1	0	1	2	4	3	3	5	54	631	7	4
63	0	2	4	619	90	10	3	2	1	1	1	2	0	1	0	1	1	0	1	11	96	503	2	1
64	5	1	2	321	558	26	5	5	3	2	3	0	1	2	1	3	5	0	0	34	542	254	2	4
57	1	1	0	1	9	343	325	22	1	3	5	5	3	2	3	2	23	42	457	63	9	8	1	0
58	1	0	3	4	0	16	0	265	1	4	3	2	0	0	2	3	0	476	305	7	6	9	1	3
59	1	0	2	4	2	1	37	0	1	2	3	1	0	1	0	2	0	377	11	3	5	4	6	8
60	3	3	0	0	1	2	2	168	12	5	2	1	2	4	2	8	0	9	9	5	3	6	6	8
61	2	3	0	0	[illegible]	0	0	0	279	116	65	42	[illegible]	48	0	193	70	18	10	0	8	9	0	0
62	3	0	0	0	1	3	2	0	225	400	0	413	453	440	474	370	12	8	3	0	2	3	5	0
63	0	1	0	2	2	1	0	2	0	18	0	81	89	72	39	10	10	5	7	7	3	3	6	0
64	0	2	1	2	1	1	1	1	1	3	0	3	2	6	4	3	0	9	2	13	3	14	4	0
57	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	2	1	0	4
58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0
59	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	1
60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2	1	0	1	2	1	0
61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0
62	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1
63	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	2	1	1
64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	2	0	15	1

Figure 5.34: MC-simulated photon counts, for the case without lenses, integrated over 60k events, ten times as many as for the data. The total photon count is 241 792. The pixels are shown in their position on the detector plane, as seen from the mirror. The photon counts per pixel and, in larger figures, per tube are given. The gaps between the active areas of the tubes are not shown.



	57	49	41	33	25	17	9	1	57	49	41	33	25	17	9	1	57	49	41	33	25	17	9	1
57	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
58	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
59	0	0	0	0	0	0	0	0	214	2053	4981	6251	6294	6267	5142	2102	0	0	0	0	0	0	0	0
60	0	0	0	0	0	0	0	0	6266	2888	443	47	6	38	361	2743	0	0	0	0	0	0	0	0
61	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	1547	4	0	0	0	0	0	0
62	0	0	0	0	0	0	16	2949	0	0	0	0	0	0	0	0	5084	3081	16	0	0	0	0	0
63	0	0	0	0	0	8	2700	4702	0	0	0	0	0	0	0	0	12	4097	2762	6	0	0	0	0
64	0	0	0	0	0	1405	5500	98	0	0	0	0	0	0	0	0	0	41	5114	1329	0	0	0	0
57	0	0	0	1684	4307	5	0	0	0	0	0	0	0	0	0	0	0	0	0	1	3774	1843	0	0
58	0	0	8	3232	2534	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1993	3456	1	0
59	0	0	43	4274	1458	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1088	4422	29	0
60	0	0	75	4602	1092	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	796	4971	46	0
61	0	0	80	0	1097	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	772	4825	52	0
62	0	0	48	4008	1526	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1147	4514	19	0
63	0	0	7	3244	2499	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1976	3516	4	0
64	0	0	0	1654	4289	5	0	0	0	0	0	0	0	0	0	0	0	0	0	4	3802	1783	0	0
57	0	0	0	0	0	1239	5131	87	0	0	0	0	0	0	0	0	0	38	5435	1437	0	0	0	0
58	0	0	0	0	0	5	0	4534	0	0	0	0	0	0	0	0	0	4387	2990	7	0	0	0	0
59	0	0	0	0	0	0	14	0	0	0	0	0	0	0	0	0	0	3229	9	0	0	0	0	0
60	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0
61	0	0	0	0	0	0	0	0	6259	2861	453	38	9	26	0	2671	0	0	0	0	0	0	0	0
62	0	0	0	0	0	0	0	0	223	1987	0	6136	6132	6124	5074	2075	0	0	0	0	0	0	0	0
63	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0
64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
57	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
58	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
59	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
62	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
63	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
64	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 5.8: Photon counting results. Shown are the displays integrated over $6k$ events, and the number of photoelectrons per event above 5σ. The results are shown per tube, and integrated over all tubes, for all data, for the background estimate, for the background-corrected data, and the Monte Carlo. All photon counts, including the Monte Carlo results, are calculated according to equations 5.39 and 5.40

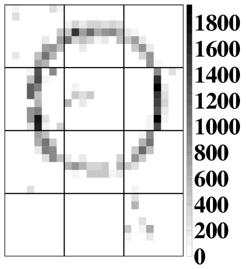
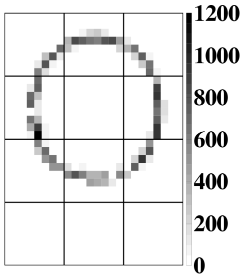
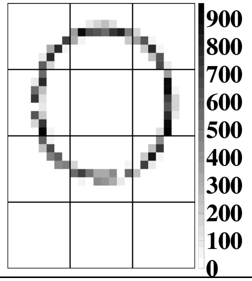
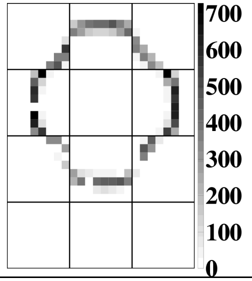
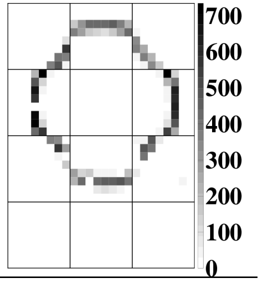
	With lenses No correction	With lenses method (a)	With lenses method (b)	Without lenses method (a)	Without lenses method (b)
data dis-play					
Data	1.43 2.12 0.95 2.73 0.42 2.34 1.30 1.09 1.23 total 13.61	0.62 1.12 0.68 1.04 0.02 1.16 0.63 0.86 0.65 total 6.78	0.65 1.14 0.68 1.25 0.03 1.19 0.68 0.93 0.70 total 7.25	0.41 0.82 0.39 0.88 0.01 0.96 0.21 0.66 0.35 total 4.69	0.41 0.82 0.38 0.96 0.01 0.96 0.31 0.70 0.42 total 4.97
Background estimate		0.03 0.03 0.03 0.03 0.02 0.03 0.03 0.02 0.03 total bg 0.26	0.03 0.03 0.03 0.05 0.03 0.03 0.03 0.02 0.05 total bg 0.29	0.02 0.02 0.02 0.02 0.01 0.03 0.01 0.03 0.05 total bg 0.21	0.02 0.02 0.03 0.03 0.01 0.03 0.01 0.03 0.10 total bg 0.28
Data, background-corrected		0.59 1.09 0.66 1.00 0.00 1.13 0.60 0.83 0.61 total sg 6.51	0.62 1.11 0.65 1.20 0.00 1.16 0.66 0.90 0.66 total sg 6.96	0.39 0.80 0.37 0.86 0.00 0.93 0.20 0.64 0.30 total sg 4.49	0.39 0.80 0.35 0.92 0.00 0.93 0.30 0.67 0.33 total sg 4.69
MC		0.56 1.10 0.68 0.93 0.00 1.06 0.58 0.91 0.67 total MC 6.49		0.29 0.77 0.38 0.70 0.00 0.75 0.18 0.67 0.29 total MC 4.03	

Table 5.9: Systematic Cross Check: Comparing background-corrected photon counts for different data runs, with the same specifications.

With lenses

run 2634

0.59	1.09	0.66
1.00	0.00	1.13
0.60	0.83	0.61

total sg 6.52

run 2662

0.57	1.06	0.65
1.01	0.00	1.09
0.63	0.79	0.64

total sg 6.44

Without lenses

run 2631

0.39	0.83	0.37
0.85	0.00	0.92
0.20	0.66	0.26

total sg 4.48

run 2610

0.39	0.80	0.37
0.86	0.00	0.93
0.20	0.64	0.30

total sg 4.49

from several sources. The comparison between the two different cross talk correction methods leads to an error of 4.5%, where tube 4 has been ignored, as cross talk method (b) fails here for known reasons. The error due to the assumption of $p(x > y) = \frac{1}{2}$ is estimated by trying different values and is found to be $\sim 1\%$. Comparing the photon counts for the same setups but different runs gives an error estimate of $\sim 1.2\%$, as seen in table 5.9. All these error contributions yield a total systematic error on the photon count for the whole ring of 5%.

With this systematic error, the agreement between data and Monte Carlo for the case with lenses is very good. For the case without lenses, data and Monte Carlo still show reasonable agreement, with the Monte Carlo estimating a photon yield of 10% less than found in the data.

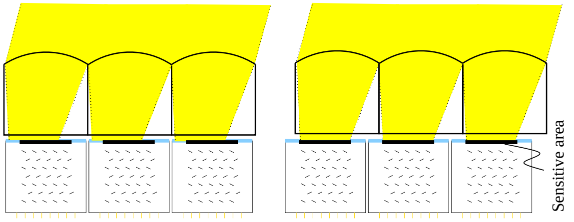


Figure 5.35: Gaining photons in a tilted lens array by shifting the lenses relative to the MaPMTs.

5.6.8 Varying the Angle of Incidence

The final LHCb RICH design may have a geometry where the Cherenkov light is not normally incident on the photo–detector plane. Therefore, the performance of the array with lenses is tested for the case where the photons do not arrive at normal incidence, but at an angle α relative to the normal on the detector plane. LED runs are used for these tests.

Here, the data are not corrected for multiple photoelectron events, and all results are given in terms of pixel hits, i.e. events above 5σ, after cross talk correction. For a homogeneous illumination with ~ 0.06 hits per pixel, there is a difference of $\sim 6\%$ between the number of hits and the number of photo electrons. This is adequate for the purpose of this study, where only relative photon counts are considered.

Tilting the array by an angle α leads to a reduction in effective area; the cross–section of the cathode relative to the photon direction is proportional to $\cos\alpha$. The number of detected photons is also reduced by in-

Table 5.10: Hits/event/tube at different angles, with the MaPMT shifted to recover as much of the image as possible. Observed values are compared with values expected from the 0° case.

	0°	20°		30°	
	Observed	Expected	Observed	Expected	Observed
Black	4.05	3.68	3.55	2.93	2.98
Diffuse	3.64	3.31	3.42	2.64	2.98

creased reflection losses and a shift in the image produced by the lens array on the MaPMT cathode. As long as the image is not outside the lens, which has opaque sides, the loss due to the shift can be completely recovered by displacing the lens array relative to the MaPMTs, as illustrated in Figure 5.35.

Figure 5.36 shows the results in terms of the average number of hits per pixel, excluding dead pixels, depending on which column the pixel is in. The column numbering is such that columns 1 and 8 correspond to the left and right respectively in figure 5.35. Some of the lenses have blackened sides, while others have diffuse sides separated with white paper. Those lenses with diffuse sides at the left or right edge of the 3×3 array are ignored in this analysis. Table 5.10 shows the results for the case that the tubes are displaced with respect to the lenses such that as much of the image is recovered as possible ("ideal shift"). The results are given separately for lenses with blackened sides and lenses with diffuse sides. These measurments are further compared with the expectation from the 0° case, considering all the contributions to the loss of photons mentioned above.

For the 20° case, the loss is dominated by the reduction in effective area, while for the 30° case, a significant fraction of the lens image is outside the lens. The disagreement between expected and measured values for 30° with diffuse sides is interpreted as some of the light hitting the side of the lens being reflected back onto the photo–cathode surface. This light is considered lost when calculating the expected value.

5.6.9 Single MaPMT

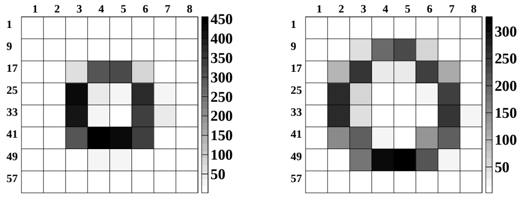
Initially, single tubes read–out with the CAMAC based electronics were tested to evaluate the performance of MaPMTs and, for the first time, their performance with the lenses.

The readout chain of CAMAC amplifiers and ADCs that is used is very similar to that of the HPD tests. Some adaptations for the increased signal size had to be made. This is described in detail in [A+98].

Air is used as a radiator, at atmospheric pressure ($n - 1 = 3.3 \cdot 10^{-4}$), producing ring-segments on the tube's surface, and at very low pressure (49 mbar, $n - 1 = 1.6 \cdot 10^{-6}$), such that a whole Cherenkov ring from a 120 GeV pion fits on a single tube. Several configurations were tested: with and without lenses, with and without a pyrex filter in front of the tube and with and without the quartz window that separates the radiator gas from the photodetec-

Table 5.11: Single MaPMT photon counts. Given are the number of photo electrons per event above a 5σ cut, in terms of the parameter $\lambda'_{5\sigma}$, integrated over the tube. The background is estimated by averaging over pixels that are clearly outside the ring.

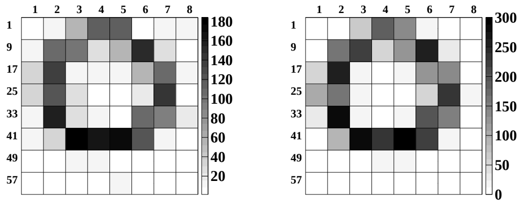
Full ring, tube 1, with/without lenses


With lens (15k events) | Without lens (15k events)

Backg.	0.022	0.014
Signal	0.259	0.274

Ratio without/with: 1.06 (expect: 1.09)

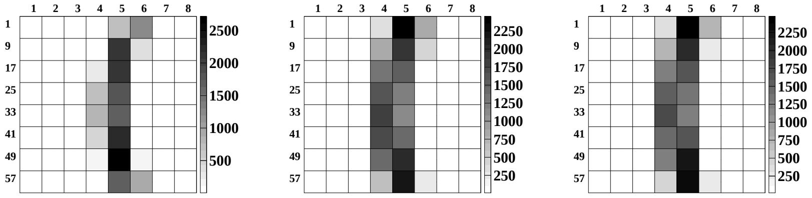
Full ring, tube 2, with/without pyrex filter


With filter (15.7k events) | without filter (15.4k events)

Backg.	0.016	0.016
Signal	0.16	0.26

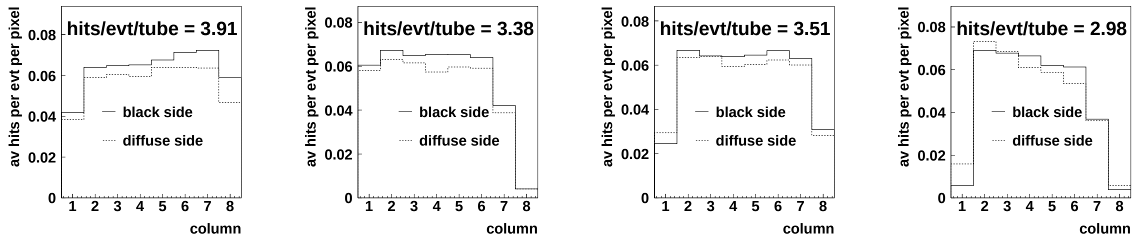
Ratio with/without: 0.62 (expect 0.56)

Ring segment, tube 1, with/without lens, with/without quartz window


with lens & window (20.7k events) | no lens, with window (33.8k events) | no lens, no window (30.4k events)

Backg.	0.0239	0.0269	0.0560
Signal	1.00	0.753	0.816
ratio	1.33	$\equiv 1$	1.08
expect	1.31	$\equiv 1$	1.09

Figure 5.36: Average number of hits per pixel depending on left-right position for light incident from the right (column 8) at different angles. Results are shown from four tubes with blackened sides, and two tubes with diffuse sides.



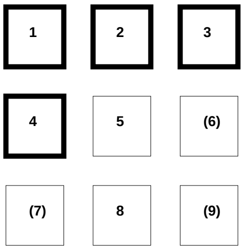
0°

20°
shift = 2.5 mm
(too small)

20°
shift = 4.5 mm
(ideal shift)

30°
shift = 5.8 mm
(ideal shift)

Figure 5.37: Arrangement of lenses with black and diffuse sides



Lenses 1 to 4 have black, 5 to 9 have diffuse sides (separated with white paper). Lenses 6, 7, 9 are excluded from the angle-of-incidence analysis.

tors. The results are given in Table 5.11.

The expected values for the ratios of the photon counts with and without lens are calculated from purely geometrical considerations, and the assumption of 8% light loss due to reflection at the lens surfaces.

The with/without ratio for the quartz window also assumes 4% light loss at each surface; for the pyrex filter, the expected value given is the ratio of the integrated photon yields calculated from the measured quantum efficiency, the transmission of the pyrex filters and the reflectivity of the mirror, as shown in figures 5.3, 4.9 and 5.2 respectively.

The single MaPMT tests show, independent from Monte Carlo simulations, that the optical properties of the quartz window, the pyrex filter, and the lenses, are in accordance with expectation.

5.7 Conclusion

The test–beam results have shown that the MaPMTs, singly and within an array, perform largely according to expectation. The lens–system was tested with Cherenkov light, and with LED light coming from different directions. The lens system performed well, and doubled the active area of the tubes at the cost of ∼ 8% reflection loss. This loss could be avoided in a final design with optical contact between the lenses and the tubes.

For the first time in a test–beam environment, photodetectors suitable for the LHCb RICH have been read out within the LHC-bunch-crossing interval of 25 ns. The pipelined readout system, with a data-capture rate of 20 MHz, performed well, apart from problems associated with cross talk.

The origin of the cross talk could be traced to specific parts of the readout chain, none of which are intended to be used in the final detector design. Future MaPMT tests would have readout electronics specifically designed to fulfil the LHCb-RICH requirements. The tubes themselves are cross-talk free.

The photon counting results, using two different methods for cross talk correction, agree well with each other, and with expectation from Monte Carlo simulation.

The success of these test–beam studies played an important part in the LHCb-RICH Photodetector Panel's conclusion in November 2000 that the MaPMT is a viable and safe photodetector option for the LHCb RICH, and should be kept as a back-up solution in case problems arise in the final stages of the development of the less expensive Pixel HPD.

Chapter 6

Conclusion

When LHCb starts data taking in 2006, $\mathcal{CP}$ violation in the B_d^0 system is expected to be a well-established phenomenon, and the B_d^0 mixing phase ϕ_mix will be precisely measured. Evidence for New Physics is in principle possible, for example in the case that ϕ_mix turns out to be near zero. However, meaningful constraints on the Standard Model from γ measurements are not expected.

LHCb will make use of the huge number of $b\bar{b}$ pairs created at the LHC to perform a comprehensive B–physics programme, that includes precision measurements of γ in a variety of decay channels. One of the key–features that facilitate this programme is the LHCb–RICH system.

An important aspect in the construction of the RICH is the choice of suitable photo detectors, that have to be precise, sensitive, and fast. One photo detector option, the commercially available MaPMT, has been shown to fulfil the LHCb RICH requirements. A full-scale RICH 2 prototype using HPD technology was tested in the CERN X7 beam line. It worked according to expectation, in particular in the two key-aspects that define the performance of a RICH detector, the photon yield and the Cherenkov angle resolution.

Table 6.1: Selected channels that are sensitive to γ. It is indicated whether they are likely to be affected by New Physics (NP) or not (SM). A estimate of the error on γ after one year of data taking is given. (From this thesis, [B+00] and [TP98]).

Channel	SM	NP	σ_γ after 1 year
$\mathrm{B}_\mathrm{d}^0 \rightarrow \mathrm{D}^{*\pm}\pi^{\mp}$	✓		$\sim 7^\circ$
$\mathrm{B}_\mathrm{s}^0 \rightarrow \mathrm{D}_\mathrm{s}^{\pm}\mathrm{K}^{\mp}$	✓		$3^\circ - 13^\circ$
$\mathrm{B}_\mathrm{d}^0 \rightarrow \rho\pi$		✓	$\sim 3^\circ$
$\mathrm{B}_\mathrm{d}^0 \rightarrow \pi^+\pi^-$		✓	$4^\circ - 8^\circ$
$\mathrm{B}_\mathrm{s}^0 \rightarrow \mathrm{K}^+\mathrm{K}^-$			
$\mathrm{B}_\mathrm{d}^0 \rightarrow \mathrm{D}_\mathrm{d}^+\mathrm{D}_\mathrm{d}^-$		✓	$\sim 3^\circ$
$\mathrm{B}_\mathrm{s}^0 \rightarrow \mathrm{D}_\mathrm{s}^+\mathrm{D}_\mathrm{s}^-$			

These results give confidence that the model of the RICH used in the simulation studies for the RICH Technical Design Report is accurate, and the results are meaningful. These studies demonstrate useful K/π separation for momenta ranging from $\sim 1\,\mathrm{GeV}$ to beyond $100\,\mathrm{GeV}$.

Table 6.1 gives a few examples for decay channels that measure the CKM angle γ. Most of the measurements listed in this table would be impossible without the K/π separation provided by the RICH. The figures given in the table are approximate values for the expected precision on γ after one year of data taking, and are subject to different assumptions for certain physics parameters, especially on the frequency of the flavour–oscillation the $\mathrm{B_s^0}$ system.

The last five channels shown in the table are expected to have strong penguin contributions and are therefore susceptible to New Physics. For the channels $\mathrm{B_d^0} \rightarrow \mathrm{D_d D_d}$ and $\mathrm{B_s^0} \rightarrow \mathrm{D_s D_s}$ the angle γ actually enters only via the penguin contributions.

In complete contrast to this are the channels $\mathrm{B_d^0} \rightarrow \mathrm{D}^{*\pm}\pi^{\mp}$ and $\mathrm{B_s^0} \rightarrow \mathrm{D_s^{\pm} K^{\mp}}$ where γ enters at tree level, and there are no penguin contributions. When compared with the former cases this will provide a very sensitive test of New Physics. It has been demonstrated in this thesis that enough $\mathrm{B_d^0} \rightarrow \mathrm{D}^{*\pm}\pi^{\mp}$ decays can be reconstructed at LHCb to measure γ to a precision of a few degrees.

Further improvements in the results are expected from extending the analysis to similar channels, for example $\mathrm{B_d^0} \rightarrow \mathrm{D^* a_1}$, which has a higher branching fraction than $\mathrm{B_d^0} \rightarrow \mathrm{D}^*\pi$. The same inclusive reconstruction method can be used. The extraction of γ is equally possible, albeit slightly more involved since it requires an angular analysis to de-convolute the different parities of the two–vector final state.

Appendix A

FORTRAN Routine to Calculate $P(k_n)$

```
*



      SUBROUTINE DYNODE_CHAIN(OUT, MAX, LAMBDA, DYNODES)
      IMPLICIT NONE
*     This program takes as its input the maximum number of electrons at
*     the end of the dynode chain, for which it should calculate P(k_n),
*     MAX, the average number of photo-electrons hitting the first
*     dynode, LAMBDA(1), the gains at each dynode, LAMBDA(2),
*     ... LAMBDA(DYNODES) and the dimension of the array LAMBDA:
*     DYNODES. It calls the routine MAKE_P_RATIO, which is listed at the
*     end of this file.
*
*     The output is put into the array OUT(MAX), where the probability
*     to find k_n < MAX electrons at the end of the dynode chain is
*     given by OUT(k_n).
*
*     Written by Jonas Rademacker.
*
      INTEGER MAX, DYNODES
      DOUBLE PRECISION OUT(0:MAX), LAMBDA(DYNODES)

      INTEGER ABS_MAX, MAX_DYN
      PARAMETER(ABS_MAX=50001,MAX_DYN=13)

      INTEGER IX,IY,M,I, K, J

*     To avoid having to define a limit on the number k_n that can be
*     calculated, one could create these arrays outside the program and
*     pass them on.
      DOUBLE PRECISION F(1:MAX_DYN) ! corrsponds to f^{star} in the text
      DOUBLE PRECISION U(1:MAX_DYN,0:ABS_MAX),V(1:MAX_DYN,0:ABS_MAX)
      DOUBLE PRECISION X(1:MAX_DYN)

      DOUBLE PRECISION FASTNULL
      PARAMETER (FASTNULL=1.d-300)

      DOUBLE PRECISION MEAN

      DOUBLE PRECISION P_ratio(ABS_MAX), F_FACTOR, U_FACTOR, V_FACTOR

      INTEGER MAX_OLD
      SAVE MAX_OLD
      DATA MAX_OLD/-9999/

      SAVE P_ratio

*     -- Some initialisations and tests --
      DO IX=1,MIN(ABS_MAX,MAX),+1
         OUT(IX)=0.d0
      ENDDO
      IF(ABS_MAX.LT.MAX)THEN
         RETURN
      ENDIF
      MEAN = 1.D0
      DO IX=1,DYNODES,+1
         MEAN = MEAN*LAMBDA(IX)
      ENDDO
      IF(MEAN.LE.0.d0)THEN
         OUT(0)=1.d0
         RETURN
      ENDIF

*     -- make and save the factors P_ratio(k)=(p_{k}/p_{k-1})^{k} --
      IF(MAX.GT.MAX_OLD)THEN
         MAX_OLD=MAX
         CALL MAKE_P_RATIO(P_ratio,MAX)
      ENDIF

*     -- Calculate the probability to see zero electrons (k_n=0) --
      F(DYNODES)=1.d0
      U(DYNODES,0)=F(DYNODES)
      V(DYNODES,0)=F(DYNODES)
      DO IX=DYNODES-1,1,-1
         X(IX)   = LAMBDA(IX)*DEXP(-LAMBDA(IX+1))
         F(IX)   = DEXP(X(IX)*F(IX+1))
         U(IX,0) = F(IX)
         V(IX,0) = F(IX)
      ENDDO
      OUT(0)=DEXP(-LAMBDA(1))*F(1) !   <---- save the result

*     -- Calculate the probabilities for k_n=1,...,MAX electrons --
      DO K=1,MAX,+1
*     .  calculate f_n
         IF(F(DYNODES).LT.FASTNULL)THEN
            F(DYNODES)=0.d0
         ELSE
            F(DYNODES)=F(DYNODES) * LAMBDA(DYNODES)/DBLE(K)
         ENDIF
         U(DYNODES,K)=F(DYNODES)
         V(DYNODES,K)=F(DYNODES)

*     .  re-calculate U and V from previous iteration:
         DO J=0,K-1,+1
            F_FACTOR=P_ratio(K)**(DBLE(J)/DBLE(K))
            IF(K-1-J.GT.0)THEN
               U_FACTOR=DSQRT(DBLE(K-1)/DBLE(K-1-J))*
     &              F_FACTOR
            ELSE
               U_FACTOR=F_FACTOR
            ENDIF
            V_FACTOR=DSQRT((DBLE(K-1)/DBLE(K-J)))*
     &           F_FACTOR

            DO I=DYNODES,1,-1
               U(I,J)=U(I,J)*U_FACTOR
               V(I,J)=V(I,J)*V_FACTOR
            ENDDO
         ENDDO

*     .  apply the recursive formula to get f^{k}_i
```

```
         DO I=DYNODES-1, 1, -1
            F(I)=0.d0
            DO J=0,K-1
               F(I)=F(I)+U(I,K-1-J)*X(I)*V(I+1,J+1)
            ENDDO
            U(I,K)=F(I)
            V(I,K)=F(I)
         ENDDO
*     .  calculate P(k):
         OUT(K)=DEXP(-LAMBDA(1))*F(1) !   <---- save the result
      ENDDO

      RETURN
      END

*------------------------------------------------------------------
      SUBROUTINE MAKE_P_RATIO(P_ratio,MAX)
      IMPLICIT NONE
      INTEGER MAX
      DOUBLE PRECISION P_ratio(MAX)

      INTEGER N
      DOUBLE PRECISION NFAC

      DOUBLE PRECISION PI, E
      PARAMETER(PI=3.1415927d0, E=2.718281828d0)

      INTEGER APPROX_FROM
      PARAMETER(APPROX_FROM=25)

      NFAC=1.D0
      P_ratio(1)=1.d0
      DO N=2,MIN(APPROX_FROM-1,MAX),+1
         NFAC=NFAC*DBLE(N-1)
         P_ratio(N)=(NFAC**(1.d0/DBLE(N-1)))/DBLE(N)
      ENDDO

      DO N=APPROX_FROM,MAX,+1
         P_ratio(N)=
     &        (2.D0*PI*DBLE(N-1))**(1.D0/(2.D0*DBLE(N-1)))*
     &        DBLE(N-1)/(E*DBLE(N))*
     &        (1.d0+1.d0/DBLE(12*(N-1))+
     &        1.d0/DBLE(288*(N-1)**2)
     &        )**(1.D0/DBLE(N-1))
      ENDDO

      RETURN
      END

*----------------------------------------------------------------
*
```

Appendix B

Approximating $\lambda_{5\sigma}$ with $\lambda'_{5\sigma}$

In chapter 5, the parameter $\lambda'_{5\sigma}$ was introduced as an approximation to the mean number of photoelectrons per event in a given pixel, detected above a 5σ threshold cut,$\lambda_{5\sigma}$.

To evaluate this approximation, it is useful to make the following definitions (all numbers refer to photon counts in a single pixel):

$\lambda_{\mathsf{p.e.}}$	Mean number of photoelectrons per event generated in the cathode.
n_{hits}	Mean number of hits per event, counting single and multiple photo electrons as a single hit and assuming 100% photoelectron detection efficiency. From Poisson statistics, this is related to $\lambda_{\mathsf{p.e.}}$ by $\lambda_{\mathsf{p.e.}} = -\ln(1 - n_{\mathrm{hits}})$
$\lambda_{5\sigma}$	Mean number of photoelectrons/event detected above a 5σ cut.
$n_{5\sigma}$	Mean number of hits/event detected above a 5σ cut.
$\lambda'_{5\sigma}$	The approximation to $\lambda_{5\sigma}$ used throughout the text, defined by $$\lambda'_{5\sigma} \equiv -\ln(1 - n_{5\sigma}) \tag{B.1}$$
l_1	Loss of single photoelectrons: The fraction of single photo electrons generated in the photo cathode, that are not detected.

If the number of photoelectrons that are detected above a 5σ cut were Poisson distributed, $\lambda_{5\sigma}$ could be calculated from the mean number of hits above 5σ, $n_{5\sigma}$, and would be given by $\lambda'_{5\sigma} = -\ln(1 - n_{5\sigma})$. However, the distribution is not exactly Poissonian, because the detection probabilities of two or more photoelectrons generated in the same event are not uncorrelated.

Their pulse heights combine to one large pulse height, and the probability that this is below the threshold cut is smaller than the probability that two individual photo electrons both remain undetected. Therefore, the probability of single photo electron losses needs to be known in order to calculate $\lambda_{5\sigma}$.

Assuming that only single photoelectron events are lost below the threshold cut, the number of hits above the cut is given by

$$n_{5\sigma} = 1 - e^{-\lambda_{\mathsf{p.e.}}} - l_1 \lambda_{\mathsf{p.e.}} e^{-\lambda_{\mathsf{p.e.}}}, \tag{B.2}$$

and the mean number of photoelectrons above the cut is

$$\lambda_{5\sigma} = \lambda_{\mathsf{p.e.}} - l_1 \lambda_{\mathsf{p.e.}} e^{-\lambda_{\mathsf{p.e.}}}. \tag{B.3}$$

Expanding equation B.1 in terms of $n_{5\sigma}$, inserting equation B.2 for $n_{5\sigma}$ and comparing the result with equation B.3, shows that $\lambda'_{5\sigma}$ under-estimates $\lambda_{5\sigma}$ by $\sim (1 - \frac{1}{2}l_1) l_1 \lambda^2_{\mathsf{p.e.}}$. For the values of $\lambda_{\mathsf{p.e.}}$ and l_1 found in the test–beam, this corresponds to less than 3% of the photon count in any given pixel for the data presented in chapter 5.

It is interesting to note that $\lambda'_{5\sigma}$ is an even better approximation to another useful parameter, the mean number of photoelectrons above the cut, for the limit that all hits are single photoelectron events:

$$\lambda^{\mathsf{single}}_{5\sigma} = (1 - l_1)\lambda_{5\sigma}. \tag{B.4}$$

The mean number of photoelectrons above the cut, $\lambda_{5\sigma}$, still has some weak dependence on the particular RICH prototype setup in which the photo detectors are tested, because the signal loss is different for multiple hits than it is for single hits. For example a shorter distance between mirror and detector–plane increases the fraction of multiple hits. On the other hand, $\lambda^{\mathsf{single}}_{5\sigma}$ is truly independent of the fraction of multiple hits. The parameter $\lambda'_{5\sigma}$ over-estimates $\lambda^{\mathsf{single}}_{5\sigma}$ by $\sim \frac{1}{2}(l_1 \lambda_{\mathsf{p.e.}})^2$, which is less than 0.5% of the photon count in any given pixel, for the values of $\lambda_{\mathsf{p.e.}}$ and l_1 found in the data presented in chapter 5.

Bibliography

[A^+90a] R. Abjean et al. Refractive index of carbon tetraflouride in the 300-140 nm wavelength range. *Nucl. Instrum. Meth.*, **A292**:593, 1990.

[A^+90b] C. Angelini et al. Sources of noise in high resolution tracking with scintillating fibers. *Nucl. Instrum. Meth.*, **A289**:356, 1990.

[A^+97] F. Anghinolfi et al. Scta – a rad-hard bicmos analogue readout ascic for the atlas semiconductor tracker. *IEEE Trans. Nucl. Sci.*, **44**:298, 1997.

[A^+98] E. Albrecht et al. First observation of Cherenkov ring images using hybrid photon detectors. *Nucl. Instrum. Meth.*, **A411**:249, 1998. Reports on RICH testbeam 1997.

[A^+00a] T. Affolder et al. A measurement of $\sin(2\beta)$ from $B^0 \rightarrow J/\psi K^0_s$ with the CDF detector. *Phys. Rev.*, **D61**:072005, 2000. [hep-ex/9909003].

[A^+00b] E. Albrecht et al. Performance of a multianode photo multiplier cluster equipped with lenses. August 2000. LHCb 2000-83 RICH. To be submitted to Nucl. Instrum. Meth.

[A^+01a] A. Abashian et al. Measurement of the CP violation parameter $\sin(2\phi_1)$ in B^0_d meson decays. *KEK preprint 2000-142, Belle preprint 2001-1*, January 2001. [hep-ex/0011090], submitted to Phys. Rev. Lett.

[A^+01b] B. Aubert et al. Measurement of CP-violating asymmetries in B^0 decays to CP eigenstates. *SLAC-PUB-8777*, January 2001. [hep-ex/0102030] Accepted for publication in Phys. Rev. Lett. (Feb 21, 2001).

[ABF^+01] M. Adinolfi, N. Brook, R. Forty, M. John, J. Libby, B. Simmons, and G. Wilkinson. A simulation study of the LHCb RICH performance. 2001. LHCb 2000-066 RICH.

[ABS80] C. Alig, S. Bloom, and C. W. Struck. Scattering by ionization and phonon emission in semiconductors. *Phys. Rev.*, **B22**:5565, December 1980.

[ARO] *AROMA manual.* AROMA: an 8-channel flash-ADC module, designed for the analogue readout of prototype silicon detectors for the CERN ATLAS experiment (unpublished).

[B+94] J. Badier et al. Shashlik calorimeter: Beam test results. *Nucl. Instrum. Meth.*, **A348**:74, 1994.

[B+96] A. Bazilevsky et al. Performance of the Phenix EM Calorimeter. *IEEE Trans. Nucl. Sci.*, **43**:1491, 1996.

[B+00] P. Ball et al. B decays at the LHC. In *1999 CERN Workshop on Standard Model Physics (and more) at the LHC, CERN, Geneva, Switzerland, 25 - 26 May 1999: Proceedings*, edited by G. G. Altarelli and M. L. Mangano, 2000. CERN 2000-004.

[Bar01] P. Bartalini. Private communications, 2001.

[BCH+99] S. A. Baird, J. A. Coughland, J. Hartley, W. J. Heynes, and T. Parthipan. A PMC based ADC card for CMS tracker readout. In *Proceedings of the Fifth Workshop on Electronics for LHC Experiments, Snowmass, Colorado, USA, September 1999*, p. 357, October 1999. CERN/LHCC/99-33.

[BMP+82] C. Bovet, R. Maleyran, L. Piemontese, A. Placci, and M. Placidi. The CEDAR counters for particle indentificaton in the SPS secondary beams: A description and an operation manual. December 1982. CERN 82-13.

[BS91] I. Bronstein and K. Semendjajew. *Taschenbuch der Mathematik.* B. G. Teubner Verlagsgesellschaft Stuttgart · Leibzig and Verlag Nauka, Moskau, 1991. ISBN 2-8154-2000-8.

[BS00] I. I. Bigi and A. I. Sander. *CP Violation.* Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2000. ISBN 0 521 44349 0.

[Cal00] *LHCb Calorimeters, Technical Design Report*, September 2000. CERN/LHCC/2000-0036.

[CCFT64] J. Christenson, J. Cronin, V. Fitch, and R. Turlay. Evidence for the 2 π decay of the K_2^0 meson. *Phys. Rev. Lett.*, **13**:138, 1964.

[CFH+99] P. Colrain, R. Forty, R. Hierck, A. Jacholkowska, M. Merk, B. Schmidt, and G. Wilkinson. Presentations during the LHCb week in Amsterdam, 1999, 1999. LHCb 99-052.

[CKW98] S. Conetti, I. Korolko, and G. Wilkinson. Performance of the LHCb calorimeter triggers. Technical report, LHCb, October 1998. LHCb 98-14 TRIG.

[CMS98] *The Tracker System Project Technical Design Report*, 1998. CERN/LHCC/98-6.

[Cou] J. Coughlan. *CMS Front–End Driver PMC, User Manual.*

[Dar75] E. H. Darlington. Backscattering of 10-100 keV electrons from thick targets. *Journal of Physics D: Appl. Phys.*, **8**:85, 1975.

[dFL$^+$96] M. D. M. de Fez-Lazo et al. *Nucl. Instrum. Meth.*, **A382**:533, 1996.

[F$^+$98] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, **81**:15, 1998.

[Fan47] U. Fano. Ionization yield of rations. II. The fluctuations of the number of ions. *Phys. Rev.*, **72**:26, 1947.

[Fle00] R. Fleischer. Constraining penguin contributions and the CKM angle γ through $\mathrm{B}_\mathrm{d}^0 \rightarrow \pi^+\pi^-$. *Eur. Phys. J.*, **C16**:87, 2000.

[FN00] R. Forty and M. Needham. Presentation during the LHCb week May 2000, May 2000. http://lhcbw3.cern.ch/forthcomingevents/2000/html/MayWeek/Transparencies/needham.pdf.

[For99] R. Forty. Influence of the new magnet design on tagging. 1999. LHCb 99-010.

[FS98] R. Forty and O. Schneider. RICH pattern recognition. April 1998. LHCB/98-040.

[FT37] I. M. Frank and I. E. Tamm. *Doklady AN SSSR*, **14**:107, 1937.

[G$^+$00] D. Groom et al. The review of particle physics. *The European Physical Journal*, **15**, July 2000. Available on the PDG WWW pages (URL: http://pdg.lbl.gov/).

[GAM00] G. G. Altarelli and M. L. Mangano, editors. *1999 CERN Workshop on Standard Model Physics (and more) at the LHC, CERN, Geneva, Switzerland, 25 - 26 May 1999: Proceedings*, 2000. CERN 2000-004.

[Gin40] V. L. Ginzburg. *Zh. Eksp. Teor. Fiz*, **10**:589, 608, 1940.

[Gla61] S. L. Glashow. Partial symmetries of weak interactions. *Nucl. Phys.*, **22**:579, 1961.

[Gro95] T. L. S. Group. The Large Hadron Collider - conceptual design, October 1995. CERN/AC/90-05(LHC).

[Gys] T. Gys. The LHCb–RICH pixel hybrid photon detectors. http://tilde-gys.home.cern.ch/~gys/LHCb/PixelHPDs.htm status: January 8, 2001.

[Ham00] *Hamamatsu Book on Photo Multipliers*, September 2000.

[Hig64a] P. W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, **13**:508, 1964.

[Hig64b] P. W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys. Lett.*, **12**:132, 1964.

[HQ98] P. Harrison and H. Quinn, editors. *The* BaBar *Physics Book*. BaBar, October 1998. SLAC-R-504.

[IK00] P. Igo-Kemenes. Status of Higgs boson searches. November 2000. Talk given at the LEP Seminar in November 3, 2000.

[J$^+$98] L. L. Jones et al. A 128 channel analogue pipeline chip for MSGC read-out at LHC. In *Proceedings of the Fourth Workshop on Electronics for LHC Experiments, Rome, September 1998*, p. 185, October 1998. CERN/LHCC/98-36.

[Jam94] F. James. *MINUIT Function Minimization and Error Analysis. Reference Manual. Version 94.1*, March 1994. CERN Program Library Long Writeup D506.

[Jar85] C. Jarlskog. Commutator of the quark mass matrices in the standard electroweak model and a measure of maximal CP violation. *Phys. Rev. Lett.*, **55**:1039, 1985.

[KM73] M. Kobayashi and T. Maskawa. CP violation in the renormalizable theory of weak interaction. *Progress in Theoretical Physics*, **49**:652, 1973.

[Köl94] K. S. Kölbig. Cernlib routine C208: Roots of a quartic equation, 1994. User Entry Names: RRTEQ4, DRTEQ4. Library: MATHLIB. Number: C208. Documented at http://wwwinfo.cern.ch/asdoc/shortwrupsdir/c208/top.html.

[Lit50] D. E. Littlewood. *A University Algebra*. Heinmann, London, 1950.

[LNQS91] H. J. Lipkin, Y. Nir, H. R. Quinn, and A. Snyder. Penguin trapping with isospin analysis and CP asymmetries in B decays. *Phys. Rev.*, **D44**:1454, 1991.

[LY56] T. D. Lee and C. N. Yang. Question of parity conservation in weak interactions. *Phys. Rev.*, **104**:254, 1956.

[Mag00] *LHCb Magnet, Technical Design Report*, 2000. CERN/LHCC/2000-007.

[MNTW94] S. Masciocchi, E. Nygård, O. Toker, and P. Weilhammer. VIKING, a CMOS low noise monolithic 128 channel frontend for Si-strip detector readouot. *Nucl. Instrum. Meth.*, **A340**:572, 1994.

[Mor] M. Morrisey. *SEQSI Manual.* unpublished.

[N$^+$00] P. Nason et al. Bottom production. In *1999 CERN Workshop on Standard Model Physics (and more) at the LHC, CERN, Geneva, Switzerland, 25 - 26 May 1999: Proceedings*, edited by G. G. Altarelli and M. L. Mangano, 2000. CERN-2000-004.

[PS97] M. E. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory.* Addison-Wesley Publishing Company, 4 edition, 1997. ISBN 0-201-50397-2.

[R$^+$97] M. Raymond et al. The APV6 readout chip for CMS microstrip detectors. In *Proceedings of the Third Workshop on Electronics for LHC Experiments, London, 1997*, p. 167, October 1997. CERN/LHCC/97-60.

[RIC00] *LHCb RICH, Technical Design Report*, September 2000. CERN/LHCC/2000-0037.

[Sal68] A. Salam. Weak and electromagnetic interactions. In *Elementary Particle Theory, Proceedings af The Nobel Symposium Held 1968 at Lerum, Sweden*, edited by N. Svartholm, p. 367. Amkvist and Wiksell, Stockholm, 1968. Reprinted in *Lichtenberg, D. B. (ed.), Rosen, S. P. (ed.): Developments in the Quark Theory of Hadrons, Vol. 1*, 160-170, and in *Lai, C. H. (ed.): Gauge Theory of Weak and Electromagnetic Interactions*, 188-198, Also in *Rosner, J.L. (ed.): New particles* 29-39, Also in *Ali, A. (ed.) et al.: Selected papers*.

[Sch00] B. Schmid. The LHCb muon system, June 2000. Talk at 2nd Int. Symp. ‘LHC Physics and Detectors’ JINR Dubna 28-30 June 2000. Available on the www: http://lhcb-muon.web.cern.ch/lhcb-muon/html/Dubna_Muon.pdf.

[tHV72] G. ’t Hooft and M. Veltman. Regularization and renormalization of gauge fields. *Nucl. Phys.*, **B44**:189, 1972.

[TP98] *LHCb Technical Proposal*, February 1998. CERN/LHCC/98-4.

[Tsa97] A. Tsaregorodtsev. *SICb User's Guide - GEANT3-based Simulation Package for the LHCb Experiment*, June 1997. Available on the www: http://lhcb-comp.web.cern.ch/lhcb-comp/sicb/html/sicbug.html.

[TvH00] V. Talanov and G. von Holtey. Flux evaluation for the first realistic LHCb vacuum chamber. March 2000. LHCb 2000-014. Internal note with restricted access.

[vB$^+$99] N. van Bakel et al. Design of a prototype frontend and bias generator for a new readout chip for LHCb. In *Proceedings of the Third Workshop on Electronics for LHC Experiments, London, September 1999*, p. 167, October 1999. CERN/LHCC/99-33.

[vB$^+$00] N. van Bakel et al. VELO geometry optimisation. October 2000. LHCb 2000-90 VELO. Internal LHCb Note with restricted access.

[Wei67] S. Weinberg. A model of leptons. *Phys. Rev. Lett.*, **19**:1264, 1967.

[Wol83] L. Wolfenstein. Parametrization of the Kobayashi-Maskawa matrix. *Phys. Rev. Lett.*, **51**:1945, 1983.

[YS93] T. Ypsilantis and J. Seguinot. Theory of ring imaging Cherenkov counters. *Nucl. Instrum. Meth.*, **A 343**:30, 1993.