

UNIVERSITÀ DEGLI STUDI DI MILANO
FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
DOTTORATO DI RICERCA IN
FISICA, ASTROFISICA E FISICA APPLICATA

**THE MEASUREMENT OF THE LORENTZ
ANGLE IN THE BTeV PIXEL DETECTORS:
THE NEW PCI BASED DAQ, THE SETUP AND
THE RESULTS**

Coordinatore: Prof. Rodolfo Bonifacio

Tutore: Dott. Luigi Moroni

Tesi di Dottorato di:

Lorenzo Uplegger

Ciclo XVI

Anno Accademico 2002-2003

Dedicata a mia mamma
che mi ha sempre sostenuto.
Con tanto amore

Lorenzo.

Contents

1	CP Violation in the Standard Model	5
1.1	Introduction	5
1.2	The CKM matrix and the unitarity triangles	6
1.3	CP violation in meson decays	9
1.3.1	Notations and formalism	9
1.4	The three types of CP violation in meson decays	13
1.4.1	CP violation in mixing	13
1.4.2	CP violation in decay	14
1.4.3	CP violation in the interference between decays with and without mixing	15
1.5	Techniques for determining β	15
1.5.1	Results on $\sin 2\beta$	17
1.5.2	Other modes for measuring $\sin(2\beta)$	18
1.6	Techniques for determining α	19
1.6.1	Introduction	19
1.7	Techniques for determining γ	20
1.8	Summary of crucial measurements for CKM physics	22
2	The BTeV experiment at Fermilab	23
2.1	Introduction	23
2.2	Geometry and strategy of the experiment	24
2.2.1	The $b\bar{b}$ production cross-section	24
2.2.2	Characteristics of hadronic b production	26
2.3	Detector description	30
2.4	BTeV analysis magnet	31
2.5	Pixel vertex detector	33
2.5.1	Introduction	33

2.5.2	Overview of the vertex detector	33
2.5.3	Spatial resolution	36
2.5.4	Pattern recognition capability	38
2.5.5	Radiation hardness	38
2.5.6	Readout speed	40
2.5.7	Physics capability	40
2.6	Forward tracking system	42
2.6.1	Introduction	42
2.6.2	Forward silicon tracker	45
2.6.3	Forward straw tracker	48
2.7	Charged particle identification	49
2.7.1	Requirements	49
2.7.2	RICH radiators	51
2.7.3	RICH dimensions	52
2.7.4	Photo-detectors	53
2.8	Electromagnetic calorimeter	54
2.8.1	Introduction	54
2.8.2	Description of the BTeV calorimeter	55
2.8.3	Radiation levels and radiation tolerance	56
2.8.4	Expected resolution	56
2.9	Muon detector	58
2.9.1	Baseline geometry	58
2.10	The BTeV trigger system	62
2.10.1	Trigger overview	62
2.10.2	Level 1 vertex trigger algorithm	64
2.10.3	Level 1 vertex trigger hardware	64
2.10.4	Levels 2/3	65
3	The DAQ for the pixel test-beam at Fermilab	67
3.1	Introduction	67
3.2	Experimental setup	68
3.3	Main characteristics of an FPGA	72
3.4	Mezzanine card	72
3.5	PCI card	74
3.5.1	The CmcConnectors block	75
3.5.2	The ConfigRegisters block	80

3.5.3	The RAMSelect block	84
3.5.4	The plxlocalbusinterface block	85
3.6	Test beam DAQ	86
3.6.1	Producer	88
3.6.2	Consumer	89
3.6.3	Logger	90
3.6.4	Controller	90
3.6.5	Xmonitor	90
4	Description of the DAQ mechanism	93
4.1	Introduction	93
4.2	Read out mechanism	94
4.3	Working mechanism of a single PCI card	95
4.4	Synchronization of the PTA cards	97
4.5	Event builder	100
4.6	Preliminary tests	101
4.7	Benchmarks	101
5	Lorentz angle measurement	103
5.1	Introduction	103
5.2	Physics of silicon detectors	104
5.2.1	Basic semiconductor physics	104
5.2.2	Charge transport	109
5.2.3	Electric field distribution	112
5.3	Lorentz effect	113
5.4	Experimental measurement	115
5.4.1	Setup description	115
5.4.2	Calibration	119
5.4.3	Measurement	120
5.4.4	Data analysis	123
5.4.5	Total Lorentz displacement	138
5.4.6	Comparison with theory	139
5.4.7	Results	141
5.4.8	Comparison with other measurements	142
5.4.9	Prospects	142

List of Figures

1.1	Unitarity Triangle	7
1.2	Up to date constraints to CKM parameters	8
1.3	CP violation in mixing	13
1.4	Decay diagram at the tree level for $B^0 \rightarrow J/\psi K_S$	16
1.5	B_d mixing diagrams	16
1.6	Constraints from $\sin 2\beta$ measurement	18
1.7	Penguin diagram for $B^0 \rightarrow \pi^+\pi^-$	19
1.8	Two diagrams for $\bar{B}_s^0 \rightarrow D_s^\pm K^\mp$	21
1.9	Diagrams for the two interfering processes $B^- \rightarrow K^-\pi^+$	21
2.1	The $b\bar{b}$ cross-section as a function of the rapidity of muons from b decay	26
2.2	The B yield versus η	27
2.3	$\beta\gamma$ of the B versus η	27
2.4	The differential $\delta\phi$ cross-sections for $p_T^\mu > 9$ GeV/c, $ \eta^\mu < 0.6$, $E_T^{\bar{b}} > 10$ GeV, $ \eta^{\bar{b}} < 1.5$ compared with theoretical predictions	28
2.5	Correlation of $b\bar{b}$ production at hadronic colliders	29
2.6	Layout of original BTeV/C0 detector (both arms fully instrumented)	31
2.7	Layout of BTeV phase 1 detector (only one arm instrumented)	32
2.8	Schematic drawing of the pixel detector.	34
2.9	Side view of the vacuum vessel and support structure for the pixel detector	36
2.10	Resolution as a function of the angle of the incident beam for both 2-bit and 8-bit ADC readouts	37
2.11	Multiparticle interaction observed in Fermilab beam-test	39
2.12	Normalized detachment, $L/\sigma(L)$, between the primary vertex and the decay vertex for the decay $B_s^- \rightarrow D_s^- K^+$	40
2.13	Vertex separation and proper time resolution	41
2.14	Occupancies in the first and last stations of straws	44

2.15	Momentum resolution of the forward tracker for b decay products . . .	45
2.16	Sketch of a silicon detector plane	46
2.17	Radiation dose as a function of position in forward silicon tracker station # 1	47
2.18	The momentum distribution of pions in $B_d \rightarrow \pi^+\pi^-$	50
2.19	The momentum distribution of “tagging” kaons	50
2.20	Cherenkov angles for various particle species	51
2.21	BTeV RICH geometry outline	52
2.22	Outline of the important RICH components.	53
2.23	The energies in the PbWO_4 calorimeter (one arm) for an event con- taining two photons from the decay sequence $B^0 \rightarrow \rho^0\pi^0, \pi^0 \rightarrow \gamma\gamma$. .	55
2.24	The maximum dose distribution.	57
2.25	The $\gamma\gamma$ invariant mass for 10 GeV π^0 's incident on the calorimeter . .	57
2.26	Layout of the muon detector geometry shown in elevation view	59
2.27	Sketch of a muon toroid	59
2.28	End and top views of one “plank” of proportional tubes.	60
2.29	Beams-eye view of one muon detector station	61
2.30	BTeV three-level trigger architecture	63
2.31	BTeV Level-1 vertex trigger.	65
3.1	Schematic representation of the pixel planes installed in the thermo- controlled hut	68
3.2	Test-beam electronics schematic	69
3.3	PTA card	70
3.4	PMC card	71
3.5	PMC electronic schematic.	73
3.6	PTA firmware schematic part I.	76
3.7	PTA firmware schematic part II.	77
3.8	Cmconnectors schematic part I.	78
3.9	Cmconnectors schematic part II.	79
3.10	ConfigRegister schematic part I.	82
3.11	ConfigRegister schematic part II.	83
3.12	RAMSelect schematic.	84
3.13	Plxlocalbusinterface schematic.	85
3.14	Global DAQ scheme.	89
3.15	Main panel of the DAQ GUI.	91

3.16	Histogram monitor.	92
4.1	Schematic of the DAQ	94
4.2	Working mechanism of a PTA	96
4.3	Non synchronized data collection	98
4.4	Synchronized data collection	99
4.5	Event builder	101
5.1	Silicon lattice structure	104
5.2	Level splitting in solids	105
5.3	Fermi function in a semiconductor	105
5.4	Extra energy levels for doped silicon	107
5.5	Different dopant impurities for silicon	107
5.6	Depletion region in a P-N Junction	108
5.7	Deformation of the valence and conduction bands at equilibrium. . .	108
5.8	Deformation of the valence and conduction bands with a reverse bias applied.	109
5.9	Mobility vs electric field equated for electrons in silicon according to equation 5.7 at $T = 300^0K$ and $N_{eff} \leq 10^{16}cm^{-3}$	111
5.10	Electric field distribution in a pixel detector	113
5.11	Velocity components under the influence of a magnetic field	114
5.12	Charge release in a silicon detector	115
5.13	Experimental setup to measure the Lorentz angle	116
5.14	Orientation of the pixels in the magnetic field	117
5.15	Example of profile of the charge collected by the pixel detector illu- minated by the LED	118
5.16	Threshold scan example	119
5.17	Example of a calibration curve	120
5.18	Displacement along the X axis	121
5.19	Displacement along Y axis	122
5.20	Mechanical movement along the X axis	124
5.21	Mechanical movement along the Y axis	125
5.22	Lorentz displacement along the X axis	126
5.23	Lorentz displacement along the Y axis	127
5.24	Schematic representation of the difference between the center of grav- ity and a gaussian mean	128
5.25	Distribution of the centers of gravity obtained by data	129

5.26	Distributions of the gaussian mean and the discretized centers of gravity	130
5.27	Schematic representation of the algorithm used to correct data along the Y axis	131
5.28	Comparison between the mechanical movement corrected and non corrected along the X axis	132
5.29	Comparison between the Lorentz displacement corrected and non corrected along the X axis	133
5.30	Correlation between the percentage of the charge collected in one column and the error in the determination of the center of gravity . .	134
5.31	Comparison between the mechanical movement corrected and non corrected along the Y axis	135
5.32	Comparison between the Lorentz displacement corrected and non corrected along the Y axis	136
5.33	Distributions of the statistical errors associated with the measurements, for corrected and non corrected data both in X and Y	137
5.34	Lorentz displacement measured in the detector for the maximum magnetic field applied at different biases	138
5.35	Electric field distribution in the detector.	139
5.36	Displacement expected in μm of the charge carriers along the silicon detector due to the Lorentz force.	140
5.37	Effective Lorentz angle in the detector extrapolated for the 1.5T magnetic field present in BTeV	142

List of Tables

1.1	Measurements of $\sin 2\beta$	17
1.2	Other modes useful for cross-checking $\sin(2\beta)$	19
1.3	$B^0 \rightarrow \rho\pi$ decay modes	20
1.4	Primary modes useful for measuring CP asymmetries for different CKM angles	22
2.1	Some crucial measurements and corresponding detector requirements	25
2.2	BTeV/C0 vertex dipole properties	33
2.3	Pixel vertex detector properties	35
2.4	Properties of the baseline forward silicon tracker (1 arm)	43
2.5	Properties of the baseline forward straw tracker (1 arm)	43
2.6	Properties of PbWO_4	54
2.7	Parameters of the baseline BTeV muon system.	61
4.1	Data transfer rates from the PTA on different architectures	102
5.1	Summary of the Lorentz angle measurements extrapolated at 1.5T . .	141
5.2	Comparison of the Lorentz angle measurements at 1.4T with a 150V bias.	142

Introduction

In the next future, the research in elementary particle physics will be mainly concentrated along two directions: searching for new phenomenology at higher energy and probing the Standard Model with very high precision. In particular, a complete set of measurements on the decays of hadrons containing b and c quarks is necessary to either accurately determine Standard Model parameters or to discover fundamental inconsistencies that could lead to an understanding beyond the model. Among these measurements, of particular interest are CP violation, mixing and rare decays of hadrons containing b or c quarks.

A comprehensive study of these phenomena will be carried out by the BTeV experiment which is expected to run in 2009 at the new Fermilab's Tevatron $C0$ interaction region. Exploiting the large number of b 's and c 's produced at the Tevatron collider, BTeV will make precise measurement of the Standard Model parameters and an exhaustive search for physics beyond the Standard Model.

The detector consists in a two arm spectrometer (only 1 arm will be instrumented in a first phase) which cover the two high rapidity regions of the $p\bar{p}$ collision. The crucial characteristic of the experiment is the trigger system to select on-line events containing heavy quarks. The selection is done for every $p\bar{p}$ collision (at the first level trigger) searching for events with secondary decay vertices significantly detached from the primary interaction vertex. This kind of trigger allows for high efficiency selection of heavy quark events and simultaneously for a powerful rejection of the overwhelming light quark background events, which are produced at a rate about 500 times higher than the b -signal events. This outstanding performance is achieved by exploiting the peculiar characteristic of heavy quark events, i.e. a long mean life and, thus, a long decay length. The trigger system is based on the information provided by a high granularity pixel silicon detector, placed inside a magnetic field, that allows for a very fast and low combinatoric reconstruction of tracks and momenta. The trigger decision is taken on the basis of the number of tracks, above a certain momentum, having an impact parameter, with respect to the

primary reconstructed interaction vertex, with a significance higher than a proper threshold value.

A detailed study of the pixel detector performance is fundamental to design an optimal vertex detector configuration for the BTeV experiment. During my thesis I worked on different aspects related to its characterization and, in particular, the implementation of a completely new DAQ, based on PCI cards and the Lorentz angle measurement in the silicon pixel detectors.

A distinctive feature of the DAQ I contributed to develop is the use of the PCI protocol, a commercially available standard, offering easy to implement, portable and low cost solutions for the development of data acquisition systems. The PCI card used in this project is provided with an FPGA (Field Programmable Gate Array) delivering great flexibility to the system. The DAQ should then be able to accomodate in the same system different detectors to allow studies both in a test-beam environment and in laboratory tests.

The interest in this subject has led me to cover several aspects of the realization of such a system, such as the firmware implementation for the FPGA and the DAQ software which profitably uses the functionalities I implemented in the FPGA.

The test-beam will be carried out in the Meson area at Fermilab using a 120 GeV/c pion beam. The pixel devices under test are located between two stations of four pixel detectors planes each, that provide tracking informations. There will be a total of twelve planes, four under test and eight in the telescope, driven by six PCI cards, each card accomodating two pixel planes. The main goals of the test-beam are, besides the study of the spatial resolution of the detector, also the study of the readout chip in the real BTeV working conditions. There will be no external trigger and the event builder will use the time-stamp information provided by the chip to assemble the events. This is a data driven scheme: data are collected as soon as they are produced in the detector.

The absence of an external trigger posed several problems to the general design of the DAQ because we needed a mechanism to balance the different acquisition rates between the ROCs and the PCI cards and between the cards and the host PC. This problem has been elegantly solved by taking advantage of the ability to program the FPGA to store the data in two memory banks located on the PTA. These two memory banks are crucial for the DAQ design because they are used both to balance the different acquisition rates and to provide a suitable solution to the problem of event building of sparsified events.

Hits are first collected in one of the two available memory buffer present on the

different PCI cards. As soon as one of the memories in the system is full, all boards are synchronously commanded to swap their memories. The one used so far are frozen and immediately read-out to the host computer, while the others are used to continue reading events from the detectors without any data loss. By means of this synchronization mechanism between the PCI cards, we can place, on the host PC, data from many detectors in a memory buffer where hits with the same timestamp are contained in a limited non spread out region. The event builder in such a configuration can be then a simple sorting algorithm with well defined limited boundaries.

The second argument of this thesis is the effect of the magnetic field on the charge collection process inside the pixel detectors. Since the pixel detector will operate in a high magnetic field of 1.5 T, the reconstructed coordinate for a track impinging on a pixel plane should be corrected to account for the apparent displacement due to the Lorentz force. Indeed, the charge released in the detector will experience this force and will be deviated from the original trajectory causing a displacement of the collected signal. The deviation angle is called Lorentz angle and its study is crucial to provide the corrections needed to improve the track reconstruction resolution. Since the Lorentz angle also depends upon the absorbed irradiation doses, different parts of the detector will experience different displacements of the charge collected. A comprehensive study of the Lorentz angle for inhomogeneous irradiation doses is thus needed in order to correctly parametrize the needed corrections to the space-point determination.

The charge displacement induced by the Lorentz force can be best studied using a blue light diode. Indeed, the wavelength ($\lambda = 450 \text{ nm}$) is such that the light is practically absorbed on the surface of the sensor and the drift of the generated charge along the bulk can be easily modeled. The charge displacement has been measured for different values of magnetic field and has been compared with theoretical expectations. Several effects, such as mechanical distortions of the system due to the magnetic field, have been taken into account and are discussed in this thesis. The results I obtained are for non irradiated devices and are in good agreement with theory and with all the previous measurements of ATLAS and CMS. The relatively simple setup will allow me to perform a complete study of this phenomenon as soon as irradiated devices will be available.

Chapter 1

CP Violation in the Standard Model

1.1 Introduction

The subject of CP symmetry and its violation is one of the least tested aspect of the Standard Model. This symmetry violation is an expected consequence of the SM with three quark generations. The CP violation that shows up in a small fraction of weak decays is accomodated simply in the three generations Standard Model Lagrangian.

However, while it is known that CP violation occurs, because it has been observed in K decays [1], it is not yet well known if the description given by the SM is the correct one. Moreover it is unlikely that the SM provides a complete description of CP violation in nature, in fact, there is a great puzzle in cosmology that relates to CP violation, and that is the baryon asymmetry in the Universe [2]. The SM cannot generate a large enough matter–antimatter imbalance to produce the baryon number to entropy ratio observed in the Universe today.

BTeV will provide significant new information on the CP violation. The main source will be measurements of CP violation in various B decays, particularly neutral B decays into final CP eigenstates. A multitude of CP–violating effects are expected in B decays, some of which are cleanly predicted by the Standard Model. Either the relationships between various measurements will be consistent with the Standard Model predictions and fully determine the CKM parameters or there will be a contribution of physics beyond the Standard Model.

BTeV will measure with high precision the parameters of the CKM matrix and thus

will be an important test for the Standard Model.

1.2 The CKM matrix and the unitarity triangles

The quarks are mixtures of base states described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix [3, 4]

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (1.1)$$

The unprimed states are the mass eigenstates, while the primed states denote the weak eigenstates. The V_{ij} 's are complex numbers that can be represented by four independent real quantities. These numbers are fundamental constants of nature that need to be determined from experiments, like any other fundamental constant such as α or G . In the Wolfenstein approximation the matrix is written as [5]

$$V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4) \quad (1.2)$$

This expression is accurate to order λ^3 . The constants λ and A have been measured using semileptonic s and b decays [6]. The phase η allows for CP violation. There are experimental constraints on ρ and η that will be discussed below.

The unitarity of the CKM matrix implies various relations among its elements. Three of them are very useful for understanding the Standard Model predictions for CP violation:

$$V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0, \quad (1.3)$$

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0, \quad (1.4)$$

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \quad (1.5)$$

Each of these three relations requires the sum of three complex quantities to vanish and so can be geometrically represented in the complex plane as a triangle. These are the “unitarity triangles” even if the term “Unitarity Triangle” is reserved for the relation 1.5. In this triangle the angles are all thought to be relatively large. To a good approximation

$$|V_{ud}| \approx |V_{tb}^*| \approx 1, \quad (1.6)$$

which implies

$$\frac{V_{ub}^*}{|V_{cd}V_{cb}^*|} - 1 + \frac{V_{td}}{|V_{cd}V_{cb}^*|} = 0 . \quad (1.7)$$

Since $|V_{cd}V_{cb}^*| = A\lambda^3$, we can define a triangle with sides

$$1 \quad (1.8)$$

$$\left| \frac{V_{td}}{A\lambda^3} \right| = \sqrt{(1-\rho)^2 + \eta^2} = \frac{1}{\lambda} \left| \frac{V_{td}}{V_{ts}} \right| \quad (1.9)$$

$$\left| \frac{V_{ub}^*}{A\lambda^3} \right| = \sqrt{\rho^2 + \eta^2} = \frac{1}{\lambda} \left| \frac{V_{ub}}{V_{cb}} \right|. \quad (1.10)$$

This CKM triangle is depicted in Fig. 1.1.

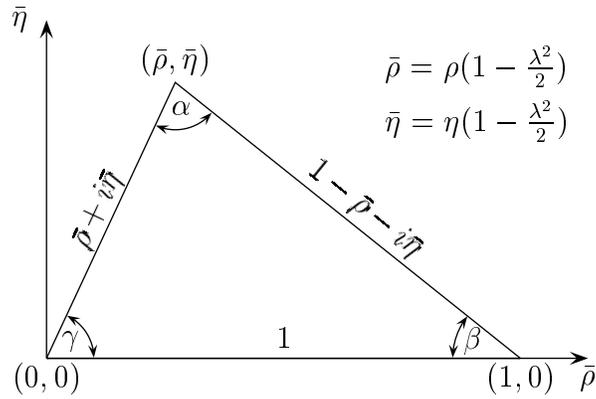


Figure 1.1: Unitarity Triangle

The rescaled unitarity triangle is derived from Eq. (1.5) by (a) choosing a phase convention such that $(V_{cd}V_{cb}^*)$ is real, and (b) dividing the lengths of all sides by $|V_{cd}V_{cb}^*|$. Step (a) aligns one side of the triangle with the real axis, and step (b) makes the length of this side 1. The form of the triangle is unchanged. Two vertices of the rescaled unitarity triangle are thus fixed at $(0,0)$ and $(1,0)$. The coordinates of the remaining vertex correspond to the Wolfenstein parameters (ρ, η) . The three angles of the unitarity triangle are denoted by α, β and γ [7]:

$$\alpha \equiv \arg \left[-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*} \right], \quad \beta \equiv \arg \left[-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*} \right], \quad \gamma \equiv \arg \left[-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} \right]. \quad (1.11)$$

To make predictions for future measurements of CP violating observables, we need to find the allowed ranges for the CKM phases. There are three ways to determine the CKM parameters (see e.g. [8]):

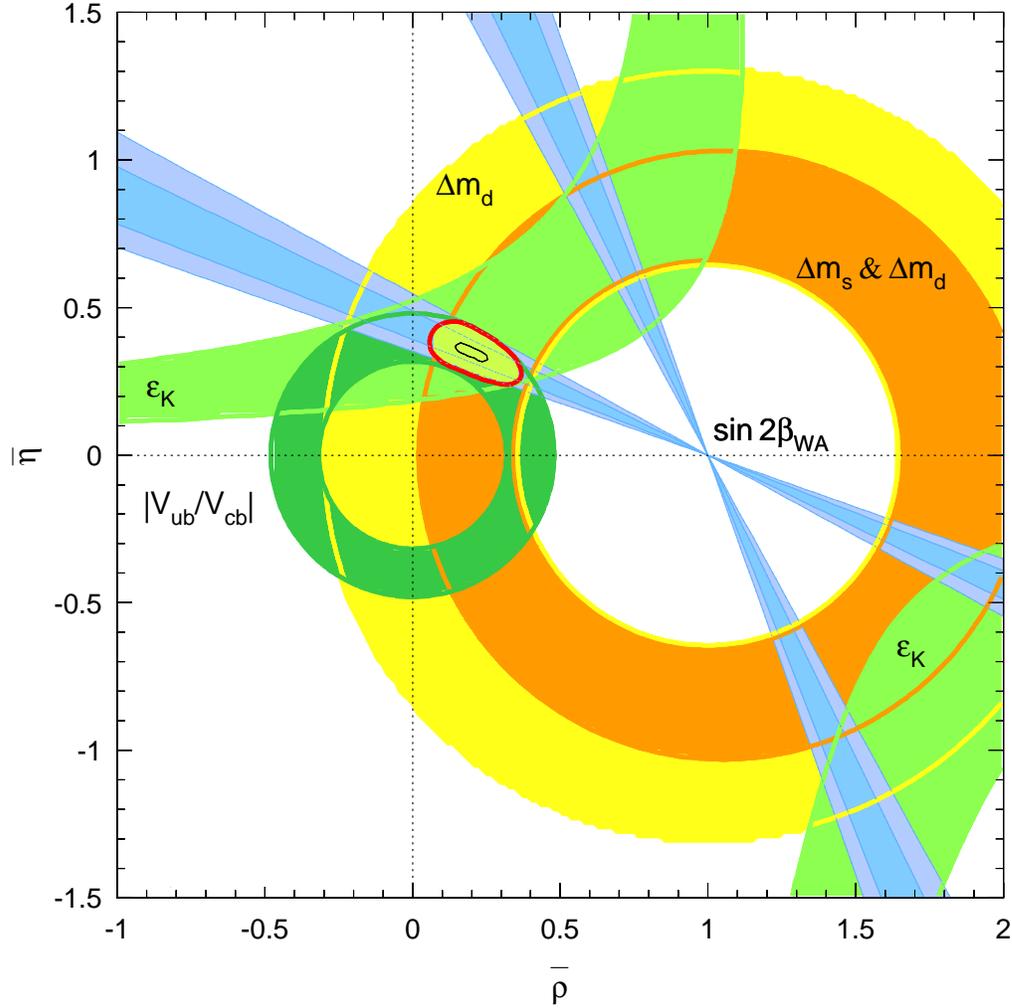


Figure 1.2: Up to date constraints to CKM parameters

- (i) **Direct measurements** are related to SM tree level processes. At present, we have direct measurements of $|V_{ud}|$, $|V_{us}|$, $|V_{ub}|$, $|V_{cd}|$, $|V_{cs}|$, $|V_{cb}|$ and $|V_{tb}|$.
- (ii) **CKM Unitarity** ($V_{CKM}^\dagger V_{CKM} = 1$) relates the various matrix elements. At present, these relations are useful to constrain $|V_{td}|$, $|V_{ts}|$, $|V_{tb}|$ and $|V_{cs}|$.
- (iii) **Indirect measurements** are related to SM loop processes. At present, we constrain in this way $|V_{tb}V_{td}|$ (from Δm_B and Δm_{B_s}) and δ_{CKM} or, equivalently, η (from ϵ_K).

When all available data is taken into account, we find:

$$\lambda = 0.2221 \pm 0.0021, \quad A = 0.827 \pm 0.058, \quad (1.12)$$

$$\rho = 0.23 \pm 0.11, \quad \eta = 0.37 \pm 0.08, \quad (1.13)$$

$$\sin 2\beta = 0.77 \pm 0.08, \quad \sin 2\alpha = -0.21 \pm 0.56, \quad 0.43 \leq \sin^2 \gamma \leq 0.91. \quad (1.14)$$

Of course, there are correlations between the various parameters. The full information can be described by allowed regions in the (ρ, η) as illustrated in Fig. 1.2 or the $(\sin 2\alpha, \sin 2\beta)$ planes (see e.g. [9]).

1.3 CP violation in meson decays

The phenomenology of CP violation in meson decays, in particular for B -meson decays, is illustrated in the following sections. There are three different types of CP violation in meson decays:

- (i) CP violation in mixing, which occurs when the two neutral mass eigenstate admixtures cannot be chosen to be CP-eigenstates;
- (ii) CP violation in decay, which occurs in both charged and neutral decays, when the amplitude for a decay and its CP-conjugate process have different magnitudes;
- (iii) CP violation in the interference of decays with and without mixing, which occurs in decays into final states that are common to B^0 and \bar{B}^0 .

In each case it is useful to identify a particular CP-violating quantity that is independent of phase conventions and discuss the types of processes that depend on this quantity. This will be done in the next section.

1.3.1 Notations and formalism

To define the three types of CP violation in meson decays and to discuss their theoretical calculation and experimental measurement, we first introduce some notations and formalism. We refer specifically to B meson mixing and decays. Our phase convention for the CP transformation law of the neutral B mesons is defined by

$$CP |B^0\rangle = w_B |\bar{B}^0\rangle, \quad CP |\bar{B}^0\rangle = w_B^* |B^0\rangle, \quad (|w_B| = 1). \quad (1.15)$$

Physical observables do not depend on the phase factor w_B . The time evolution of any linear combination of the neutral B -meson flavor eigenstates,

$$a |B^0\rangle + b |\bar{B}^0\rangle \quad (1.16)$$

is governed by the Schrödinger equation,

$$i \frac{d}{dt} \begin{pmatrix} a \\ b \end{pmatrix} = H \begin{pmatrix} a \\ b \end{pmatrix} \equiv \left(M - \frac{i}{2} \Gamma \right) \begin{pmatrix} a \\ b \end{pmatrix}, \quad (1.17)$$

for which M and Γ are 2×2 Hermitian matrices. CPT invariance guarantees $H_{11} = H_{22}$, that is $M_{11} = M_{22}$ and $\Gamma_{11} = \Gamma_{22}$. The off-diagonal terms in these matrices, M_{12} and Γ_{12} , are particularly important in the discussion of mixing and CP violation. M_{12} is the dispersive part of the transition amplitude from B^0 to \bar{B}^0 , while Γ_{12} is the absorptive part of that amplitude. In the Standard Model these contributions arise from the box diagrams with two W exchanges.

The light B_L and heavy B_H mass eigenstates are given by

$$|B_{L,H}\rangle = p |B^0\rangle \pm q |\bar{B}^0\rangle \quad (1.18)$$

and their time evolution is given by

$$\begin{aligned} |B_H(t)\rangle &= e^{-iM_H t} e^{-\Gamma_H t/2} |B_H\rangle, \\ |B_L(t)\rangle &= e^{-iM_L t} e^{-\Gamma_L t/2} |B_L\rangle. \end{aligned} \quad (1.19)$$

The complex coefficients q and p obey the normalization condition $|q|^2 + |p|^2 = 1$. Note that $\arg(q/p^*)$ is just an overall common phase for $|B_L\rangle$ and $|B_H\rangle$ and has no physical significance. The mass difference and the width difference between the physical states are defined as follows:

$$\Delta m \equiv M_H - M_L, \quad \Delta \Gamma \equiv \Gamma_H - \Gamma_L. \quad (1.20)$$

Solving the eigenvalue equation gives

$$(\Delta m)^2 - \frac{1}{4}(\Delta \Gamma)^2 = (4|M_{12}|^2 - |\Gamma_{12}|^2), \quad \Delta m \Delta \Gamma = 4\Re(M_{12}\Gamma_{12}^*). \quad (1.21)$$

$$\frac{q}{p} = -\frac{\Delta m - \frac{i}{2}\Delta \Gamma}{2M_{12} - i\Gamma_{12}} = -\frac{2M_{12}^* - i\Gamma_{12}^*}{\Delta m - \frac{i}{2}\Delta \Gamma}. \quad (1.22)$$

The time evolution of an initially pure B^0 or \bar{B}^0 state is:

$$\begin{aligned} |B_{\text{phys}}^0(t)\rangle &= f_+(t) |B^0\rangle + \frac{q}{p} f_-(t) |\bar{B}^0\rangle, \\ |\bar{B}_{\text{phys}}^0(t)\rangle &= f_+(t) |\bar{B}^0\rangle + \frac{p}{q} f_-(t) |B^0\rangle. \end{aligned} \quad (1.23)$$

where

$$f_{\pm}(t) = \frac{1}{2} \left[e^{-i(M_L - \frac{i}{2}\Gamma_L)t} \pm e^{-i(M_H - \frac{i}{2}\Gamma_H)t} \right] \quad (1.24)$$

In the B system, $|\Gamma_{12}| \ll |M_{12}|$ (see discussion below), and then, to leading order in $|\Gamma_{12}/M_{12}|$ Eqs. (1.21) and (1.22) can be written as

$$\Delta m_b = 2|M_{12}|, \quad \Delta\Gamma_B = 2\Re(M_{12}\Gamma_{12}^*)/|M_{12}|, \quad (1.25)$$

$$\frac{q}{p} = -\frac{M_{12}^*}{|M_{12}|}. \quad (1.26)$$

To discuss CP violation in mixing, it is useful to write Eq. (1.22) to first order in $|\Gamma_{12}/M_{12}|$:

$$\frac{q}{p} = -\frac{M_{12}^*}{|M_{12}|} \left[1 - \frac{1}{2}\Im\left(\frac{\Gamma_{12}}{M_{12}}\right) \right] \quad (1.27)$$

To discuss CP violation in decay, we need to consider decay amplitudes. The CP transformation law for a final state f is

$$CP|f\rangle = w_f|\bar{f}\rangle, \quad CP|\bar{f}\rangle = w_f^*|f\rangle, \quad (|w_f| = 1). \quad (1.28)$$

For a final CP eigenstate $f = \bar{f} = f_{\text{CP}}$, the phase factor w_f is replaced by $\eta_{f_{\text{CP}}} = \pm 1$, the CP eigenvalue of the final state. We define the decay amplitudes A_f and \bar{A}_f according to

$$A_f = \langle f|\mathcal{H}_d|B^0\rangle, \quad \bar{A}_f = \langle f|\mathcal{H}_d|\bar{B}^0\rangle \quad (1.29)$$

where \mathcal{H}_d is the decay Hamiltonian. CP relates A_f and \bar{A}_f . There are two types of phases that may appear in A_f and \bar{A}_f . Complex parameters in any Lagrangian term that contributes to the amplitude will appear in complex conjugate form in the CP-conjugate amplitude. Thus their phases appear in A_f and \bar{A}_f with opposite signs. In the Standard Model these phases occur only in the CKM matrix which is part of the electroweak sector of the theory, hence these are often called “weak phases”. The weak phase of any single term is convention dependent. However the difference between the weak phases in two different terms in A_f is convention independent because the phase rotations of the initial and final states are the same for every term. A second type of phase can appear in scattering or decay amplitudes even when the Lagrangian is real. Such phases do not violate CP and they appear in A_f and \bar{A}_f with the same sign. Their origin is the possible contribution from intermediate on-shell states in the decay process, that is an absorptive part of an amplitude that has contributions from coupled channels. Usually the dominant re-scattering is due to strong interactions and hence the designation “strong phases”

for the phase shifts so induced. Again only the relative strong phases of different terms in a scattering amplitude have physical content, an overall phase rotation of the entire amplitude has no physical consequences. Thus it is useful to write each contribution to A in three parts: its magnitude A_i , its weak phase term $e^{i\phi_i}$ and its strong phase term $e^{i\delta_i}$. Then, if several amplitudes contribute to $B \rightarrow f$ we have

$$\left| \frac{\bar{A}_f}{A_f} \right| = \left| \frac{\sum_i A_i e^{i(\delta_i - \phi_i)}}{\sum_i A_i e^{i(\delta_i + \phi_i)}} \right| \quad (1.30)$$

To discuss CP violation in the interference of decays with and without mixing, we introduce a complex quantity λ_f defined by

$$\lambda_f = \frac{q}{p} \frac{\bar{A}_f}{A_f} \quad (1.31)$$

We further define the CP transformation law for the quark fields in the Hamiltonian (a careful treatment of CP conventions can be found in [10]):

$$q \rightarrow w_q \bar{q}, \quad \bar{q} \rightarrow w_q^* q, \quad (|w_q| = 1). \quad (1.32)$$

The effective Hamiltonian that is relevant to M_{12} is of the form

$$H_{\text{eff}}^{\Delta b=2} \propto e^{+2i\phi_B} [\bar{d}\gamma^\mu(1 - \gamma_5)b]^2 + e^{-2i\phi_B} [\bar{b}\gamma^\mu(1 - \gamma_5)d]^2, \quad (1.33)$$

where $2\phi_B$ is a CP violating (weak) phase. For the B system, where $|\Gamma_{12}| \ll |M_{12}|$, this leads to

$$q/p = w_B w_b^* e^{-2i\phi_B}. \quad (1.34)$$

To understand the phase structure of decay amplitudes, we take as an example the $b \rightarrow q\bar{q}d$ decay ($q = u$ or c). The decay Hamiltonian is of the form

$$\begin{aligned} H_d \propto & e^{+i\phi_f} [\bar{q}\gamma^\mu(1 - \gamma_5)d] [\bar{b}\gamma_\mu(1 - \gamma_5)q] \\ & + e^{-i\phi_f} [\bar{q}\gamma^\mu(1 - \gamma_5)b] [\bar{d}\gamma_\mu(1 - \gamma_5)q] \end{aligned} \quad (1.35)$$

where ϕ_f is the appropriate weak phase. Then

$$\bar{A}_f/A_f = w_f w_B^* w_b w_d^* e^{-2i\phi_f}. \quad (1.36)$$

Eqs. (1.34) and (1.36) together imply that for a final CP eigenstate,

$$\lambda_{f\text{CP}} = \eta_{f\text{CP}} e^{-2i(\phi_B + \phi_f)}. \quad (1.37)$$

1.4 The three types of CP violation in meson decays

Let's study in some detail each type of CP violation.

1.4.1 CP violation in mixing

$$\left| \frac{q}{p} \right| \neq 1 \quad (1.38)$$

This results from the mass eigenstates being different from the CP eigenstates. In fact if $|q/p| = 1$ this term represents a pure phase and due to the freedom of choosing an arbitrary phase in CP transformation we could define:

$$CP |B^0\rangle = \left| \frac{q}{p} \right| |\bar{B}^0\rangle \quad (1.39)$$

With the convention of Eq. (1.39) B_H and B_L become CP eigenstates. But if Eq. (1.38) holds true, than there is no legitimate CP transformation of B^0 and \bar{B}^0 that will make the mass eigenstates CP eigenstates. This type of CP violation requires also a relative phase between M_{12} and Γ_{12} . In fact it could be demonstrated that B_H and B_L can be CP eigenstates if and only if $\Im(M_{12}\Gamma_{12}^*) = 0$, that is when the relative phase between M_{12} and Γ_{12} vanishes. The way it comes about is because there are two ways in which a B^0 can mix into a \bar{B}^0 : one is the direct way (connected with M_{12} , the dispersive part) and the other is via a common decay mode of the two states (connected to Γ_{12} , the absorptive part), as depicted in Fig. 1.3. For

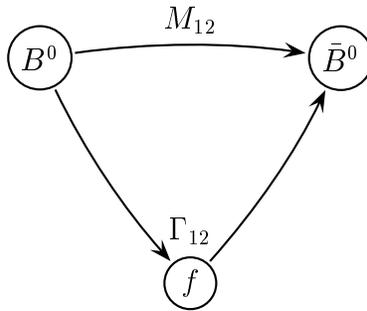


Figure 1.3: CP violation in mixing

the neutral B system, this effect could be observed through the asymmetries in semileptonic decays:

$$a_{\text{SL}} = \frac{\Gamma(\bar{B}_{\text{phys}}^0(t) \rightarrow \ell^+ \nu X) - \Gamma(B_{\text{phys}}^0(t) \rightarrow \ell^- \nu X)}{\Gamma(\bar{B}_{\text{phys}}^0(t) \rightarrow \ell^+ \nu X) + \Gamma(B_{\text{phys}}^0(t) \rightarrow \ell^- \nu X)} \quad (1.40)$$

In terms of q and p ,

$$a_{\text{SL}} = \frac{1 - |q/p|^4}{1 + |q/p|^4}. \quad (1.41)$$

CP violation in mixing has been observed in the neutral K system ($\Re \varepsilon_k \neq 0$). In the neutral B system, the effect is expected to be small, $\lesssim \mathcal{O}(10^{-2})$. The reason is that, model independently, one expects that $a_{\text{SL}} \lesssim \Delta\Gamma_B/\Delta m_B$. The difference in width is produced by decay channels common to B^0 and \bar{B}^0 . The branching ratios for such channels are at or below the level of 10^{-3} . Since various channels contribute with differing signs, one expects that their sum does not exceed the individual level. Hence, we can safely assume that $\Delta\Gamma_B/\Gamma_B = \mathcal{O}(10^{-2})$. On the other hand, it is experimentally known that $\Delta m_B/\Gamma_B \approx 0.7$. To calculate a_{SL} , we use (1.41) and (1.27), and get:

$$a_{\text{SL}} = \Im \left(\frac{\Gamma_{12}}{M_{12}} \right). \quad (1.42)$$

To predict it in a given model, one needs to calculate M_{12} and Γ_{12} . This involves large hadronic uncertainties, in particular in the hadronization models for Γ_{12} .

1.4.2 CP violation in decay

$$\left| \frac{\bar{A}_f}{A_f} \right| \neq 1. \quad (1.43)$$

This appears as a result of interference among various terms in the decay amplitude, and will not occur unless at least two terms have different weak phases and different strong phases. CP asymmetries in charged B decays,

$$a_{f^\pm} = \frac{\Gamma(B^+ \rightarrow f^+) - \Gamma(B^- \rightarrow f^-)}{\Gamma(B^+ \rightarrow f^+) + \Gamma(B^- \rightarrow f^-)}, \quad (1.44)$$

are purely an effect of CP violation in decay. In terms of the decay amplitudes,

$$a_{f^\pm} = \frac{1 - |\bar{A}_{f^-}/A_{f^+}|^2}{1 + |\bar{A}_{f^-}/A_{f^+}|^2}. \quad (1.45)$$

CP violation in decay has been observed in the neutral K system ($\Re \varepsilon_k' \neq 0$). To calculate a_{f^\pm} , we use (1.45) and (1.30). For simplicity, we consider decays with contributions from two weak phases and with $A_2 \ll A_1$. We get:

$$a_{f^\pm} = -2(A_2/A_1) \sin(\delta_2 - \delta_1) \sin(\phi_2 - \phi_1). \quad (1.46)$$

The magnitude and strong phase of any amplitude involve long distance strong interaction physics, and our ability to calculate these from first principles is limited. Thus quantities that depend only on the weak phases are much cleaner than

those that require knowledge of the relative magnitudes or strong phases of various amplitude contributions, such as CP violation in decay.

1.4.3 CP violation in the interference between decays with and without mixing

$$|\lambda_{f_{\text{CP}}}| = 1, \quad \Im(\lambda_{f_{\text{CP}}}) \neq 0. \quad (1.47)$$

Any $\lambda_{f_{\text{CP}}} \neq \pm 1$ is a manifestation of CP violation. The special case (1.47) isolates the effects of interest since both CP violation in mixing, Eq. (1.38), and in decay, Eq. (1.43), lead to $|\lambda_{f_{\text{CP}}}| \neq 1$. For the neutral B system, this effect can be observed by comparing decays into final CP eigenstates of a time-evolving neutral B state that begins at time zero as B^0 to those of the state that begins as \bar{B}^0 :

$$a_{f_{\text{CP}}} = \frac{\Gamma(\bar{B}_{\text{phys}}^0(t) \rightarrow f_{\text{CP}}) - \Gamma(B_{\text{phys}}^0(t) \rightarrow f_{\text{CP}})}{\Gamma(\bar{B}_{\text{phys}}^0(t) \rightarrow f_{\text{CP}}) + \Gamma(B_{\text{phys}}^0(t) \rightarrow f_{\text{CP}})}. \quad (1.48)$$

This time dependent asymmetry is given by

$$a_{f_{\text{CP}}} = \frac{(1 - |\lambda_{f_{\text{CP}}}|^2) \cos(\Delta M t) - 2\Im(\lambda_{f_{\text{CP}}}) \sin(\Delta M t)}{1 + |\lambda_{f_{\text{CP}}}|^2} \quad (1.49)$$

and, for $|\lambda_{f_{\text{CP}}}| = 1$, Eq. (1.49) simplifies considerably to

$$a_{f_{\text{CP}}} = -\Im(\lambda_{f_{\text{CP}}}) \sin(\Delta m_B t) \quad (1.50)$$

CP violation in the interference of decays with and without mixing has been observed for the neutral K system ($\Im \varepsilon_k \neq 0$). It is expected to be an effect of $\mathcal{O}(1)$ in various B decays. For such cases, the contribution from CP violation in mixing is clearly negligible. For decays that are dominated by a single CP violating phase (for example, $B \rightarrow \psi K_S$ and $K_L \rightarrow \pi^0 \nu \bar{\nu}$), so that the contribution from CP violation in decay is also negligible, $a_{f_{\text{CP}}}$ is cleanly interpreted in terms of purely electroweak parameters. Explicitly, $\Im(\lambda_{f_{\text{CP}}})$ gives the difference between the phase of the $B - \bar{B}$ mixing amplitude ($2\phi_B$) and twice the phase of the relevant decay amplitude ($2\phi_f$) (see Eq. (1.37)):

$$\Im \lambda_{f_{\text{CP}}} = -\eta_{f_{\text{CP}}} \sin[2(\phi_b + \phi_f)]. \quad (1.51)$$

1.5 Techniques for determining β

The angle β of the unitarity triangle is defined by:

$$\beta \equiv \arg \left[-\frac{V_{cd} V_{cb}^*}{V_{td} V_{tb}^*} \right], \quad (1.52)$$

The decay $B^0 \rightarrow J/\psi K_S$ is the primary source for measurements of $\sin(2\beta)$. In the common phase convention, CP violation is expected to arise mostly from the mixing, driven by $\text{Im}(q/p)$, while the decay amplitude, $\text{Im}(\bar{A}/A)$, is expected to contribute only a small part. The final state $J/\psi K_S$ is a CP eigenstate and its decay is dominated by only one diagram, shown in Fig. 1.4. For this decay mode

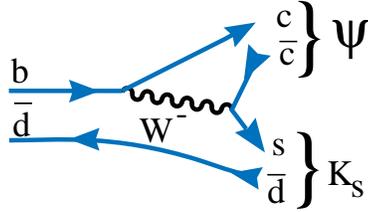


Figure 1.4: Decay diagram at the tree level for $B^0 \rightarrow J/\psi K_S$.

we have

$$\lambda(B_d \rightarrow J/\psi K_S) = - \left(\frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} \right) \left(\frac{V_{cs}^* V_{cb}}{V_{cs} V_{cb}^*} \right) \left(\frac{V_{cd}^* V_{cs}}{V_{cd} V_{cs}^*} \right) \quad (1.53)$$

Where the first term comes from $B_d^0 - \bar{B}_d^0$ mixing, see Fig. 1.5, the second from the ratio $\frac{\bar{A}}{A}$ and the third from the $K^0 - \bar{K}^0$ mixing.

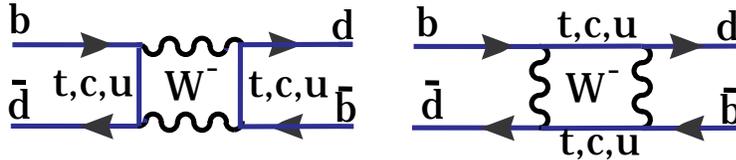


Figure 1.5: The two diagrams for B_d mixing. Although u , c and t quark exchange are all shown, the t quark plays a dominant role, mainly due to his mass, since the amplitude of this process grows with the mass of the exchanged fermion

In this case we do not get a phase from the decay part because

$$\frac{\bar{A}}{A} = \frac{(V_{cb} V_{cs}^*)^2}{|V_{cb} V_{cs}|^2} \quad (1.54)$$

is real. The final state is a state of negative CP , i.e. $CP|J/\psi K_S\rangle = -|J/\psi K_S\rangle$. This introduces an additional minus sign in the result for $\text{Im}\lambda$. Before finishing the discussion of this final state we need to consider in more detail the presence of the

K_S in the final state. Since neutral kaons can mix, we pick up another mixing phase. This term creates a phase given by

$$\left(\frac{q}{p}\right)_K = \frac{(V_{cd}^* V_{cs})^2}{|V_{cd} V_{cs}|^2}, \quad (1.55)$$

which is zero. It is necessary to include this term, however, since there are other formulations of the CKM matrix than Wolfenstein, which have the phase in a different location. It is important that the physics predictions not depend on the CKM convention.

1.5.1 Results on $\sin 2\beta$

The first statistically significant measurements of CP violation in the B system were made recently by BABAR and BELLE [11]. This enormous achievement was accomplished using an asymmetric e^+e^- collider on the $\Upsilon(4S)$. The measurements are listed in Table 1.1, along with other previous indications [12].

Table 1.1: Measurements of $\sin 2\beta$.

Experiment	$\sin 2\beta$
BABAR	$0.59 \pm 0.14 \pm 0.05$
BELLE	$0.99 \pm 0.14 \pm 0.06$
Average	0.79 ± 0.11
CDF	$0.79^{+0.41}_{-0.44}$
ALEPH	$0.84^{+0.82}_{-1.04} \pm 0.16$
OPAL	$3.2^{+1.8}_{-2.0} \pm 0.5$

The average value of 0.79 ± 0.11 is taken from BABAR and BELLE only. These two measurements do differ by a sizeable amount, but the confidence level that they correctly represent the same value is 6%. This value is consistent with what is expected from the other known constraints on ρ and η . We have

$$\bar{\eta} = (1 - \bar{\rho}) \frac{1 \pm \sqrt{1 - \sin^2 2\beta}}{\sin 2\beta}. \quad (1.56)$$

There is a four fold ambiguity in the translation between $\sin 2\beta$ and the linear constraints in the $\rho - \eta$ plane. These occur at β , $\pi/2 - \beta$, $\pi + \beta$ and $3\pi/2 - \beta$. Two of these constraints are shown in Fig. 1.6. The other two can be viewed by extending

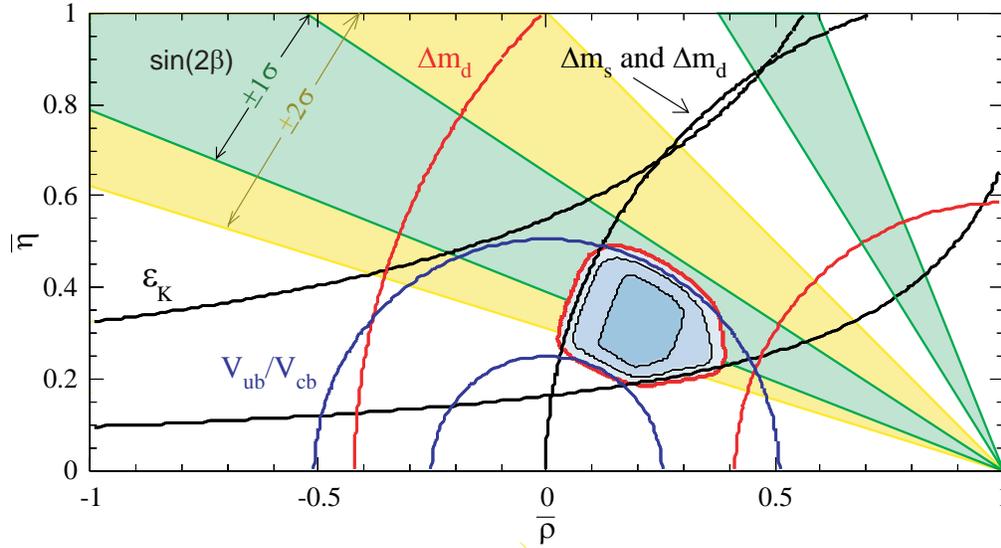


Figure 1.6: Constraints from $\sin 2\beta$ measurement overlaid with other constraints from Hocker *et al* [13]. The inner band is at 1σ while the outer band, shown on one band only, is at 2σ .

these to negative $\bar{\eta}$. We think $\eta > 0$ based only on measurement of ϵ' in the neutral kaon system. This analysis clearly shows that current data are consistent with the Standard Model.

BTeV will improve significantly on the precision of the $\sin(2\beta)$ measurement. Furthermore, it will be able to remove “ambiguities”. These ambiguities can mask the effects of new physics.

1.5.2 Other modes for measuring $\sin(2\beta)$

New physics can add differently to the phases in different decay modes if it contributes differently to the relative decay amplitudes A/\bar{A} . Therefore it is interesting to measure CP violation in redundant modes. For example, the decay $B^0 \rightarrow \phi K_S$ should also measure $\sin(2\beta)$. If it is different than that obtained by $B^0 \rightarrow J/\psi K_S$, that would be a strong indication of new physics [14]. Other interesting modes to check $\sin(2\beta)$ are listed in Table 1.2. The branching ratios listed with errors have been measured [15–17], while those without are theoretical estimates.

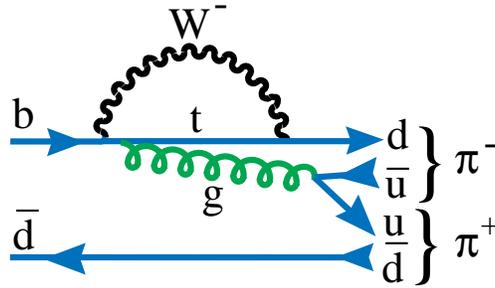
Table 1.2: Other modes useful for cross-checking $\sin(2\beta)$

Decay Mode	Branching Ratio
$B^0 \rightarrow \phi K_S$	$(4.0 \pm 2.1) \times 10^{-6}$
$B^0 \rightarrow D^+ D^-$	$\approx 10^{-3}$
$B^0 \rightarrow D^{*+} D^-$	$\approx 10^{-3}$
$B^0 \rightarrow \eta' K^0,$	$(5.9 \pm 1.9) \times 10^{-5}$
$B^0 \rightarrow J/\psi(\pi^0, \eta \text{ or } \eta')$	$(3.4 \pm 1.6) \times 10^{-5}$

1.6 Techniques for determining α

1.6.1 Introduction

Measuring α is more difficult than measuring β in several respects. First of all, the decay amplitudes are modulated by V_{ub} rather than V_{cb} , making the overall rates small. Secondly, the gluonic penguin rates, see Fig. 1.7, are of the same order causing difficulties in extracting the weak phase angle. The penguin diagrams add a third

Figure 1.7: Penguin diagram for $B^0 \rightarrow \pi^+ \pi^-$.

amplitude to the tree level and mixing amplitudes. It turns out, however, that this complication can be used to remove discrete ambiguities.

The decay $B^0 \rightarrow \pi^+ \pi^-$ has often been cited as a way to measure $\sin(2\alpha)$. However, the penguin pollution mentioned above, makes it difficult to extract the angle.

There is however, a theoretically clean method to determine α . The interference between tree and penguin diagrams can be exploited by measuring the time dependent CP violating effects in the decays $B^0 \rightarrow \rho \pi$ as shown by Snyder and Quinn [18]. There are three such neutral decay modes, listed in Table 1.3, with their respective

Table 1.3: $B^0 \rightarrow \rho\pi$ decay modes

Decay Mode	Decay Amplitudes
$\sqrt{2}A(B^+ \rightarrow \rho^+\pi^0)$	$=S_1 = T^{+0} + 2P_1$
$\sqrt{2}A(B^+ \rightarrow \rho^0\pi^+)$	$=S_2 = T^{0+} - 2P_1$
$A(B^0 \rightarrow \rho^+\pi^-)$	$=S_3 = T^{+-} + P_1 + P_0$
$A(B^0 \rightarrow \rho^-\pi^+)$	$=S_4 = T^{-+} - P_1 + P_0$
$2A(B^0 \rightarrow \rho^0\pi^0)$	$=S_5 = T^{+-} + T^{-+} - T^{+0} - T^{0+} - 2P_0$

penguin, denoted by P , and tree amplitudes, denoted by T^{ij} , where i lists charge of the ρ and j the charge of the π . For the $\rho^0\pi^0$ mode, isospin constraints are used to eliminate T^{00} . The amplitudes for the charged decays are also given.

For the $\rho\pi$ final state, the ρ decay amplitude can be parameterized as

$$f(m, \theta) = \frac{\cos(\theta)\Gamma_\rho}{2(m_\rho - m - i0.5\Gamma_\rho)} \quad , \quad (1.57)$$

where m_ρ is the ρ mass of 0.77 GeV and Γ_ρ the width of 0.15 GeV. θ is the helicity decay angle and the $\cos(\theta)$ dependence arises because the ρ must be fully polarized in this decay which starts with a spin-0 B and ends with a spin-1 ρ and spin-0 π .

The full decay amplitudes for $B^0 \rightarrow \rho\pi \rightarrow \pi^+\pi^-\pi^0$ and the corresponding \bar{B}^0 decay are given by

$$\begin{aligned} A(B^0) &= f^+ S_3 + f^- S_4 + f^0 S_5/2 \\ A(\bar{B}^0) &= f^+ \bar{S}_3 + f^- \bar{S}_4 + f^0 \bar{S}_5/2 \quad , \end{aligned} \quad (1.58)$$

where the superscript on the f indicates the charge of the ρ . The sum over the three neutral B decay amplitudes involves only tree amplitudes; the penguins vanish. The angle between this sum for B^0 decays ($\equiv T$) and the sum for \bar{B}^0 ($\equiv \bar{T}$) is precisely α . Computing the amplitudes gives a series of terms which have both $\sin(\Delta mt)$ and $\cos(\Delta mt)$ time dependences and coefficients which depend on both $\sin(2\alpha)$ and $\cos(2\alpha)$.

1.7 Techniques for determining γ

The angle γ could in principle be measured using a CP eigenstate of B_s decay that was dominated by the $b \rightarrow u$ transition. One such decay that has been suggested

is $B_s \rightarrow \rho^0 K_S$. However, there are the same “penguin pollution” problems as in $B^0 \rightarrow \pi^+ \pi^-$, but they are more difficult to resolve in the vector-pseudoscalar final state, the pseudoscalar-pseudoscalar final state here is $\pi^0 K_S$, which does not have a measurable decay vertex. There are two main methods of measuring γ .

- Using Time-Dependent CP violation in B_s Decays.

The first method uses the decays $B_s \rightarrow D_s^\pm K^\mp$ where a time-dependent CP violation can result from the interference between the direct decays and the mixing-induced decays [19–21]. Figure 1.8 shows the two direct decay processes for \bar{B}_s^0 .

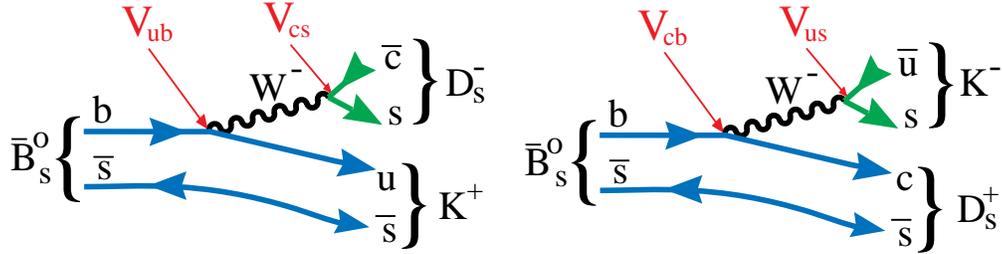


Figure 1.8: Two diagrams for $\bar{B}_s^0 \rightarrow D_s^\pm K^\mp$.

- Using Charged B Decay Rates.

Another method for extracting γ has been proposed by Atwood, Dunietz

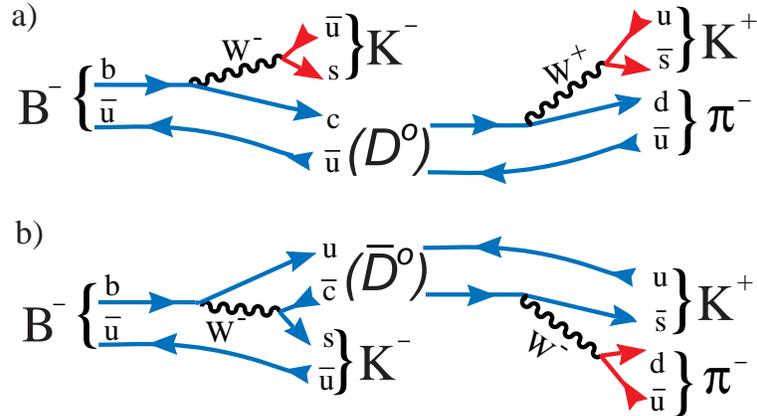


Figure 1.9: Diagrams for the two interfering processes, (a) $B^- \rightarrow D^0 K^-$ (color allowed) followed by $D^0 \rightarrow K^+ \pi^-$ (doubly-Cabibbo suppressed), and (b) $B^- \rightarrow \bar{D}^0 K^-$ (color suppressed) followed by $D^0 \rightarrow K^- \pi^+$ (Cabibbo allowed).

and Soni [22], who refined a suggestion by Gronau and Wyler [23]. A large CP asymmetry can result from the interference of the decays $B^- \rightarrow K^- D^0$, $D^0 \rightarrow f$ and $B^- \rightarrow K^- \bar{D}^0$, $\bar{D}^0 \rightarrow f$, where f is a doubly-Cabibbo suppressed decay of the D^0 , for example $f = K^+ \pi^-$, $K \pi \pi$, see Fig. 1.9. The overall amplitudes for the two decays are expected to be approximately equal in magnitude, note that $B^- \rightarrow K^- \bar{D}^0$ is color-suppressed and $B^- \rightarrow K^- D^0$ is color-allowed. The weak phase difference between them is γ . To observe a CP asymmetry there must also be a non-zero strong phase between the two amplitudes. It is necessary to measure the branching ratio $\mathcal{B}(B^- \rightarrow K^- f)$ for at least 2 different states f in order to determine γ up to discrete ambiguities.

1.8 Summary of crucial measurements for CKM physics

Table 1.4 lists the most important physics quantities and the decay modes that can be used to measure them. After the three angles α , β and γ have been measured, we need to check if they add up to 180° . A discrepancy here would show new physics. New physics can add differently to the phases in different decay modes if it contributes differently to the relative decay amplitudes A/\bar{A} . Therefore it is interesting to measure CP violation in redundant modes. For example, the decay $B^0 \rightarrow \phi K_S$ should also measure $\sin(2\beta)$. If it is different than that obtained by $B^0 \rightarrow J/\psi K_S$, that would be a strong indication of new physics [14].

Table 1.4: Primary modes useful for measuring CP asymmetries for different CKM angles

Physics Quantity	Decay Mode
$\sin(2\alpha)$	$B^0 \rightarrow \rho\pi \rightarrow \pi^+ \pi^- \pi^0$
$\cos(2\alpha)$	$B^0 \rightarrow \rho\pi \rightarrow \pi^+ \pi^- \pi^0$
$\text{sign}(\sin(2\alpha))$	$B^0 \rightarrow \rho\pi$ & $B^0 \rightarrow \pi^+ \pi^-$
$\sin(\gamma)$	$B_s \rightarrow D_s^\pm K^\mp$
$\sin(\gamma)$	$B^- \rightarrow \bar{D}^0 K^-$
$\sin(2\beta)$	$B^0 \rightarrow J/\psi K_S$
$\cos(2\beta)$	$B^0 \rightarrow J/\psi K^0$, $K^0 \rightarrow \pi \ell \nu$
$\cos(2\beta)$	$B^0 \rightarrow J/\psi K^{*0}$ & $B_s \rightarrow J/\psi \phi$

Chapter 2

The BTeV experiment at Fermilab

2.1 Introduction

The BTeV physics involves reconstructing a variety of different decay modes of the B , B_s , and other B hadrons and, in many cases, following their time evolution and tagging the flavor of the parent B at production and at the moment of decay. These decay modes may involve charged hadrons, charged leptons, photons (prompt or from π^0 's), and tertiary vertices from the $b \rightarrow c$ decay chain. In some cases, there are substantial backgrounds from minimum bias (typical hadronic) events, charm decays, or other B decays. In many cases, the branching fractions, including any tertiary decays, are quite small, typically 10^{-5} to 10^{-7} . This, together with the large background of minimum bias events, demands that BTeV be able to reconstruct multi-body final states, with good resolution in invariant mass, and to handle very high rates. In order to carry out the physics program described above, the detector must have the ability to separate decay vertices from the primary interaction vertex and to reconstruct secondary B vertices and daughter charm vertices. This requires a precision vertex detector. It also must be able to measure the time evolution of decays for time-dependent asymmetry studies. The most demanding requirement is to be able to follow the very rapid oscillations of the B_s meson in order to study CP violation. It must have the ability to distinguish pions, kaons, and protons from each other to eliminate kinematic reflections that can contaminate signals and make them difficult to observe. Many key decay modes have π^0 's, γ 's, or particles that decay into them, such as ρ 's or η 's. Leptons, muons and electrons (positrons), appear in many key final states so good lepton identification is also required. Finally, many of the detector properties which are needed to isolate and reconstruct signals

are also needed to perform “flavor tagging”. These translate into a basic set of requirements for the detector, shown in Table 2.1. In the table, there are listed physics topics, a particular decay mode that can be used to study it, and then tabulate the key features necessary to reconstruct the signal and perform flavor tagging where required. It can be seen that in order to carry out this program, the detector must make a complete characterization of the final state particles. A table prepared to address the topic of rare decays would have similar characteristics. It should be clear that a device with these properties, combined with a very powerful and inclusive trigger system for B decays and a high speed data analysis system, can address a very large range of topics.

2.2 Geometry and strategy of the experiment

The physics of hadronic beauty and charm production plays a major role in the design of BTeV. In hadron colliders all B species, B_d , B_u , B_s , b -baryons, and even B_c mesons, are produced at the same time. This allows one to carry out a very large number of interesting studies and to look for unexpected phenomena since the detector is both powerful and flexible, especially in the area of triggering.

2.2.1 The $b\bar{b}$ production cross-section

It is customary to characterize heavy quark production in hadron collisions with two variables, the momentum transverse to the beams, p_t , and the rapidity,

$$y = \frac{1}{2} \ln \left(\frac{E + p_{\parallel}}{E - p_{\parallel}} \right), \quad (2.1)$$

where E is the particle’s energy and p_{\parallel} is its longitudinal momentum. Often, the pseudorapidity η

$$\eta = -\ln(\tan(\theta/2)) \quad , \quad (2.2)$$

where θ is the angle of the particle with respect to the beam direction, is used for the longitudinal variable since this variable is independent of the particle’s mass.

The $p\bar{p}$ production of b quarks has been measured in the Tevatron at a center-of-mass energy of 1.8 TeV in the central rapidity region $|\eta| < 1$ by CDF [24] and D0 [25], and in the forward region $3.2 > y > 2.4$ by D0 [26]. Both CDF and D0 find that the $b\bar{b}$ production cross-section in the central region is underestimated by the Mangano, Nason and Ridolfi (MNR) next-to-leading order QCD calculation [27] by

Table 2.1: Some crucial measurements and corresponding detector requirements. In order to separate signals and background, all studies in BTeV need good primary and secondary vertex resolution, which is equivalent to a requirement on the resolution in proper time, τ , of a small fraction of the B lifetime. The requirement of “superb τ resolution”, referred to in this table, means resolution which is a small fraction of the expected period for B_s mixing and is a much more stringent requirement. The “lepton id” is checked where it is used to extract the signal decay, although it participates in most of the other studies via lepton flavor tagging.

Physics Quantity	Decay Mode	Detector Property				
		Vertex trigger	K/ π separation	γ detection	superb τ resolution	lepton id
$\sin(2\alpha)$	$B^0 \rightarrow \rho\pi \rightarrow \pi^+\pi^-\pi^0$	✓	✓	✓		
$\cos(2\alpha)$	$B^0 \rightarrow \rho\pi \rightarrow \pi^+\pi^-\pi^0$	✓	✓	✓		
$\text{sign}(\sin(2\alpha))$	$B^0 \rightarrow \rho\pi, B^0 \rightarrow \pi^+\pi^-$	✓	✓	✓		
$\sin(\gamma)$	$B_s \rightarrow D_s K^-$	✓	✓		✓	
$\sin(\gamma)$	$B^+ \rightarrow D^0 K^+$	✓	✓			
$\sin(\gamma)$	$B \rightarrow K\pi$	✓	✓	✓		
$\sin(\gamma)$	$B \rightarrow \pi^+\pi^-, B_s \rightarrow K^+K^-$	✓	✓		✓	
$\sin(2\chi)$	$B_s \rightarrow J/\psi\eta', J/\psi\eta$	✓	✓	✓	✓	✓
$\sin(2\beta)$	$B^0 \rightarrow J/\psi K_s$					✓
$\sin(2\beta)$	$B^0 \rightarrow \phi K_s, \eta' K_s, J/\psi\phi$	✓	✓	✓		✓
$\cos(2\beta)$	$B^0 \rightarrow J/\psi K^*, B_s \rightarrow J/\psi\phi$					✓
x_s	$B_s \rightarrow D_s\pi^-$	✓	✓		✓	
$\Delta\Gamma$ for B_s	$B_s \rightarrow J/\psi\eta', K^+K^-, D_s\pi^-$	✓	✓	✓		✓

a factor of approximately two. Since the QCD calculation predicts a cross-section of 50 μb , when integrated over η and p_t , using the data in the central regions leads to a total $b\bar{b}$ production cross-section of 100 μb . The D0 central and forward data are shown in Fig. 2.1.

The fact that the production cross section for states containing b -quarks is $\approx 1/500$ of the total cross section has strong implications for the design of the ex-

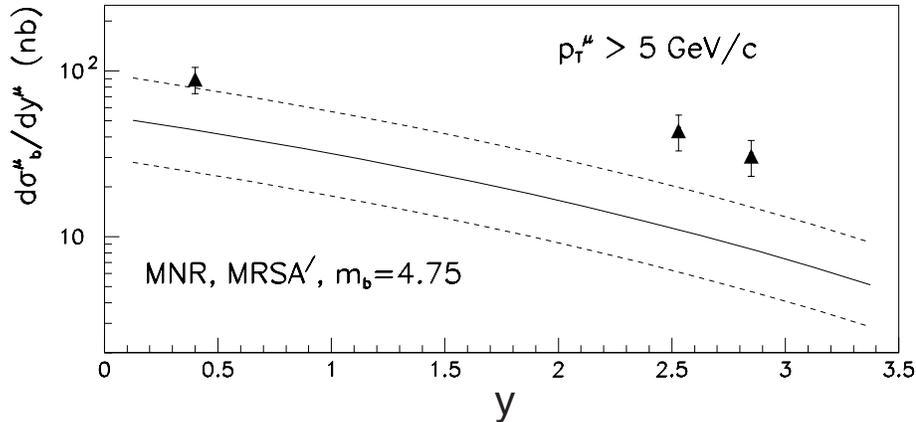


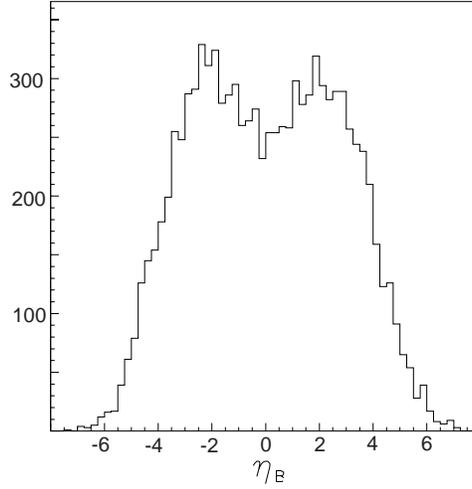
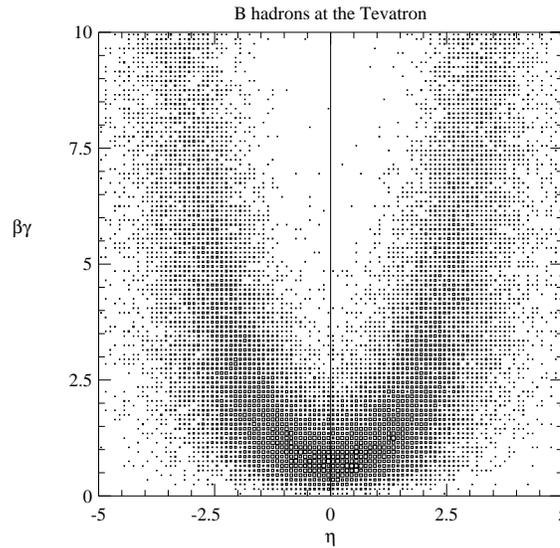
Figure 2.1: The $b\bar{b}$ cross-section as a function of the rapidity of muons from b decay, y^{μ} , measured by D0 for both the forward and central rapidity regions, using muons from b decays with $p_t > 5$ GeV/c. The solid curve is the prediction of the next-to-leading order QCD calculation for a b -quark mass of 4.75 GeV. The dashed curves represent the estimated theoretical 1σ error band.

periment. It means that the experiment must have a very good trigger to reject the very large number of typical interactions which involve only light quarks. It further means that the experiment will have to handle very high particle fluxes, and tolerate very high radiation doses, if it is going to get reasonable samples of the key decay modes it wants to study, especially given the well-known fact that the B decay modes most interesting for CP studies have rather small branching fractions.

2.2.2 Characteristics of hadronic b production

The dominant mechanism for b quark production at the Tevatron is believed to be gluon-gluon fusion. Whenever the two gluons have different Feynman- x , the center of mass of the produced $b-\bar{b}$ pair is boosted along the direction of the higher momentum gluon. According to both simple arguments and detailed QCD calculations, the b 's are produced approximately “uniformly” in η and have a truncated transverse momentum (p_t) spectrum, characterized by a mean value approximately equal to the B mass [28]. The distribution in η is shown in Fig. 2.2.

There is a strong correlation between the B momentum and η . Shown in Fig. 2.3 is the $\beta\gamma$ of the B hadron versus η from the Monte Carlo physics generator Pythia at $\sqrt{s} = 2$ TeV. It can clearly be seen that near η of zero, $\beta\gamma \approx 1$, while at larger values of $|\eta|$, $\beta\gamma$ can easily reach values of 6. This is important because the observed

Figure 2.2: The B yield versus η .Figure 2.3: $\beta\gamma$ of the B versus η .

decay length varies with $\beta\gamma$ and, furthermore, the absolute momenta of the decay products are larger allowing for a suppression of the multiple scattering error.

Since the detector design is somewhat dependent on the Monte Carlo generated b production distributions, it is important to check that the correlations between the b and the \bar{b} are adequately reproduced. Figure 2.4 shows the azimuthal opening angle distribution between a muon from a b quark decay and the \bar{b} jet as measured by CDF [29] and compares it with the MNR next-to-leading order QCD predictions [27].

The MNR model does a good job representing the shape, which shows a strong

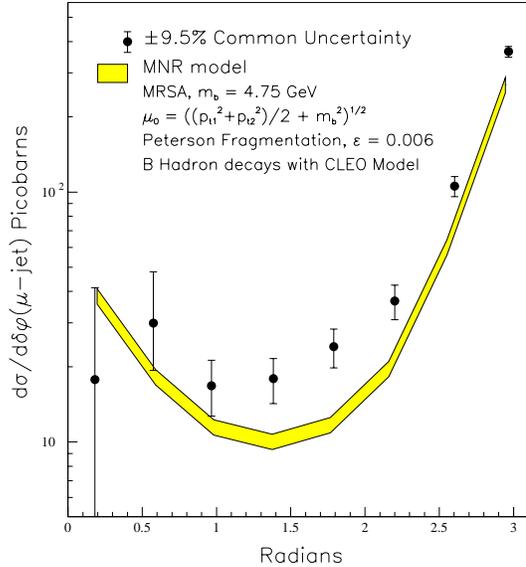


Figure 2.4: The differential $\delta\phi$ cross-sections for $p_T^\mu > 9$ GeV/c, $|\eta^\mu| < 0.6$, $E_T^{\bar{b}} > 10$ GeV, $|\eta^{\bar{b}}| < 1.5$ compared with theoretical predictions. The data points have a common systematic uncertainty of $\pm 9.5\%$. The uncertainty in the theory curve arises from the error on the muonic branching ratio and the uncertainty in the fragmentation model.

back-to-back correlation. The normalization is about a factor of two higher in the data than the theory, which is generally true of CDF and D0 b cross-section measurements.

The “flat” η distribution hides an important correlation of $b\bar{b}$ production at hadronic colliders. In Fig. 2.5 the production angle of the hadron containing the b quark is plotted versus the production angle of the hadron containing the \bar{b} quark. Here zero degrees represents the direction of the incident proton and 180 degrees, the incident anti-proton. There is a very strong correlation in the proton or the anti-proton directions: when the B is forward the \bar{B} is also forward. This correlation between B and \bar{B} production is not present in the central region (near 90 degrees). This is a result of the underlying physics of gluon-gluon collisions described above.

In the forward direction, this correlation is crucial to carry out studies that involve flavor tagging. For many B decay studies that involve mixing, it is necessary to determine the flavor of the signal B hadron at the moment of production. One way to do this is to determine the flavor of the “other” B hadron. Because of the correlated nature of the b -quark production, both B hadrons will be boosted in the same direction and therefore the signal and the tagging decay products will appear

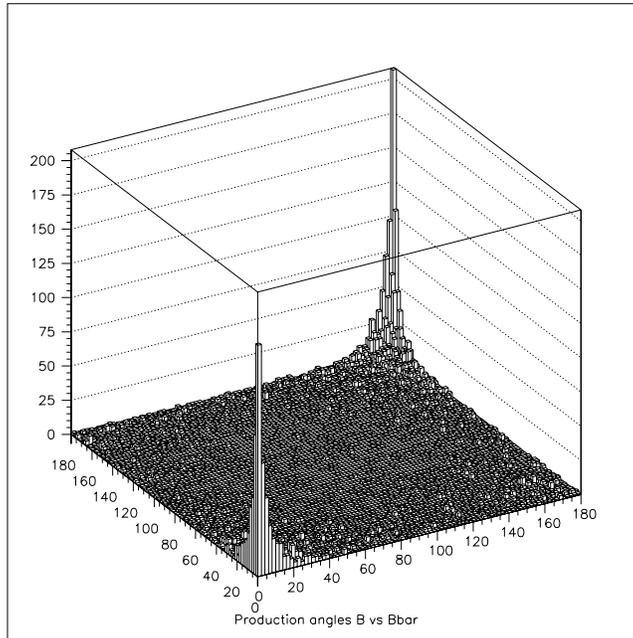


Figure 2.5: The production angle (in degrees) for the hadron containing the b quark plotted versus the production angle for the hadron containing the \bar{b} quark in the same event, from the Pythia Monte Carlo generator.

in the same “arm”.

Thus, the forward direction at the Tevatron has a lot of striking advantages. First of all, there is a large cross-section for the production of correlated $b\bar{b}$ pairs. Secondly, the B hadrons that are formed have relatively large momenta, on average 30 GeV/c, and their decay products are not too badly affected by Multiple Coulomb Scattering. This allows to make precision measurements of their spatial origins, so it is possible to determine if they arise from B hadrons that traveled on the order of several mm prior to their decay. Furthermore the geometry is very natural for certain aspects of detector technology that significantly enhance the physics performance. For example, a Ring Imaging Cherenkov detector using a gas radiator matches the 3-70 GeV/c momentum range for B decay products. The Cherenkov photons can be detected using a relatively small area array of photomultiplier tubes or Hybrid PhotoDiodes(HPDs). Powerful particle identification is essential for high sensitivity b experiments. Another example is the ability to put the silicon pixel vertex detector inside the main beam vacuum. Precision detection of the B decay vertices is crucial for the trigger and in rejecting backgrounds.

Charm production is similar to b production but has a much larger cross section. Current theoretical estimates are that charm is 1-2% of the total $p\bar{p}$ cross-section. The cross section is even more strongly peaked in the forward direction because the average transverse momentum is of the order of only 1.5 GeV/ c . The charm cross section has never been measured because experiments with good acceptance in the central region have very low efficiency for triggering and reconstructing charm. The favorable kinematics in the forward direction gives BTeV a very high efficiency for reconstructing charm.

2.3 Detector description

The BTeV detector will cover the angular region from 10 mr to 300 mr relative to the proton beam and from 10 mr to 300 mr relative to the anti-proton beam. Each angular region is covered by an “arm” of the spectrometer. In the first phase of data taking only one “arm” will be instrumented but will be left room to permit the installation of components on the “uninstrumented side” later on. The single “arm” design is capable of carrying out the full program of measurements albeit on a longer time scale. Figure 2.6 shows the full detector as originally proposed. A schematic of the Phase 1 detector is shown in Fig. 2.7.

The key design features of BTeV include:

- A dipole located on the IR, which permits BTeV to have an effective “two arm” acceptance. In Phase 1, only one arm is instrumented.
- A precision vertex detector based on planar pixel arrays;
- A detached vertex trigger at Level 1 which makes BTeV efficient for most final states, including purely hadronic modes;
- Precision tracking using straw tubes and silicon microstrip detectors, which provide excellent momentum and mass resolution out to 300 mr;
- Excellent particle identification using a Ring Imaging Cherenkov Detector (RICH). The RICH provides hadron identification from 3-70 GeV and lepton identification from 3-20 GeV, out to the full aperture of 300 mr, which is crucial since the muon detector and calorimeter do not cover the full solid angle.

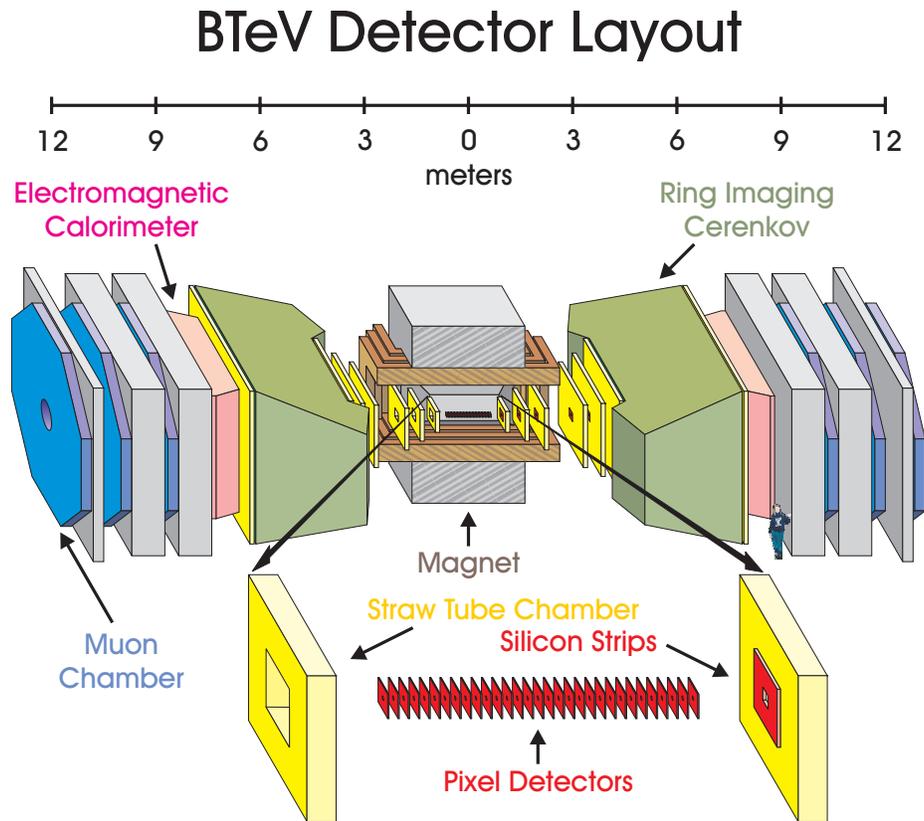


Figure 2.6: Layout of original BTeV/C0 detector (both arms fully instrumented)

- A high quality PbWO_4 electromagnetic calorimeter, covering 200 mr, capable of reconstructing final states with single photons, π^0 's, η 's or η' 's, and identifying electrons;
- Excellent identification of muons out to 200 mr using a dedicated detector consisting of a steel toroid instrumented with proportional tubes. This system has the ability to supply a dimuon trigger;
- A very high speed and high throughput data acquisition system which eliminates the need to tune the experiment to specific final states.

2.4 BTeV analysis magnet

BTeV will use an existing magnet already used for other experiments at Fermilab. Modifications will be done to increase the value of the magnetic field in the central

BTeV Detector Layout

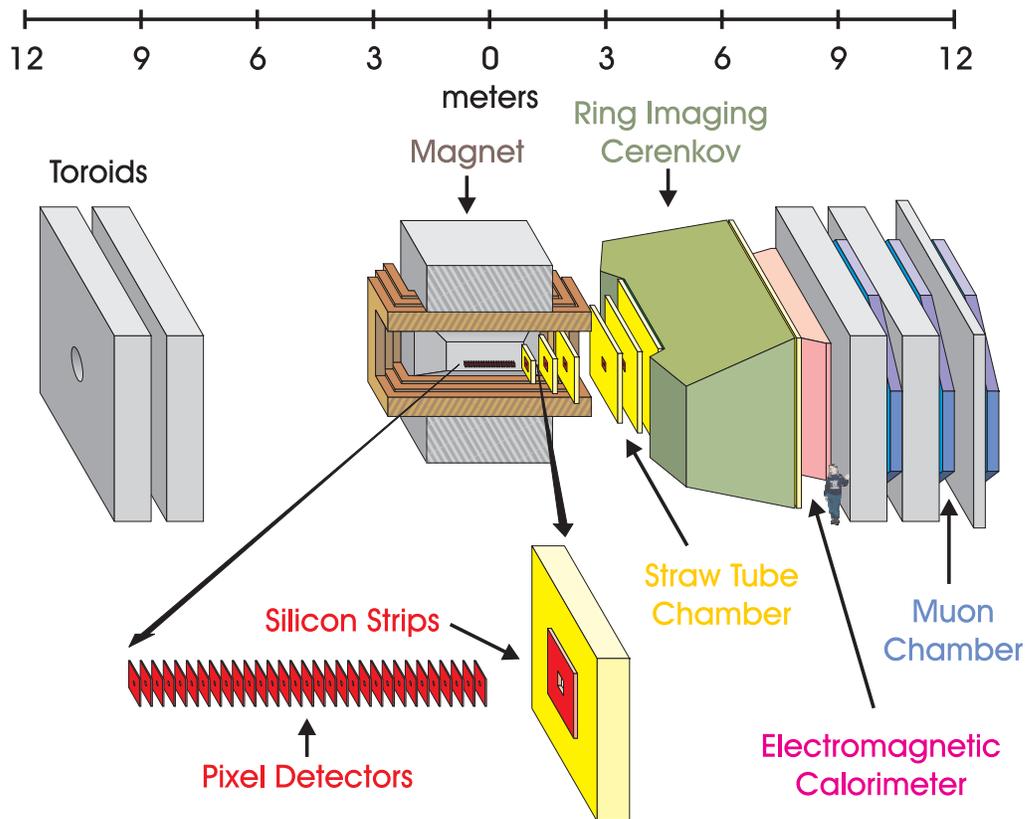


Figure 2.7: Layout of BTeV phase 1 detector (only one arm instrumented)

region from 0.8 to 1.5 Tesla, and an integrated dipole field of 5.0 T-m. The magnet will be oriented so that charged particles are deflected in the vertical plane. The properties of the magnet are listed in Table 2.2. The vertical deflection of the Tevatron beam by the vertex magnet is compensated by two conventional dipoles at each end of the spectrometer.

The magnet is centered on the interaction region thus creating the potential for two forward spectrometers. In this central dipole geometry, there is a strong magnetic field at the vertex detector. Because of the excellent spatial resolution of the vertex detector, it is possible to get a crude measurement of the track momentum using the vertex detector alone. This measurement can be used to reject tracks at the trigger level that are of such low momentum that multiple Coulomb scattering errors make their assignment to a vertex uncertain.

Table 2.2: BTeV/C0 vertex dipole properties

Property	Value	Comment
$\int B \times dl$	5.0 T-m	2.5 T-m on each side of center of IR
Central Field	1.5 Tesla	
Steel Length	3.2 m	
Overall length	5.3 m	
Magnet Vert. aperture	± 0.3 rad	
Magnet Horz. aperture	± 0.3 rad	

2.5 Pixel vertex detector

2.5.1 Introduction

The vertex detector is critical to the success of BTeV. The key goals of the detector are excellent spatial resolution, ease of tracking pattern recognition, radiation hardness, material thinness, and readout of data fast enough for use in the lowest-level BTeV trigger system. The detector design has been guided by these goals, as will be described in the sections below.

2.5.2 Overview of the vertex detector

The baseline vertex detector consists of a regular array of 30 “stations” of “planar” silicon pixel detectors distributed along the interaction region (Fig. 2.8). Each station contains one plane with the narrow pixel dimension vertical, and one with the narrow dimension horizontal. The stations are split, having a top half and a bottom half. The top half-stations are positioned at regular intervals along the beam, and the bottom halves are similarly positioned, but midway between the top stations. This allows for possible overlap of half-planes with a variable-sized, small hole left for the beams to pass through. Table 2.3 summarizes the properties of the pixel detector.

The vertex detector contains nearly twenty-two million rectangular pixels, each $50 \mu\text{m} \times 400 \mu\text{m}$. Each sensor pixel is read out by a dedicated electronics cell. The sensor pixel and the readout cell are connected by a “bump bond.” The basic building block of the detector is a hybrid assembly consisting of a sensor, a number of

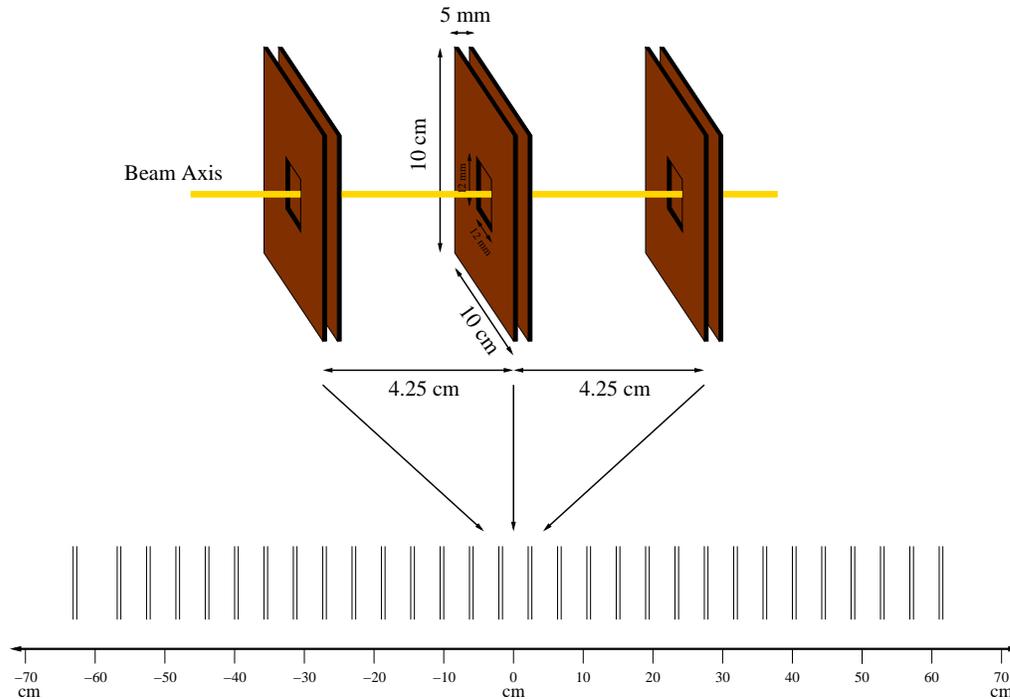


Figure 2.8: Schematic drawing of the pixel detector.

readout chips, and a flexible printed circuit (a high-density interconnect, HDI) which carries I/O signals and power. The sensors are variously sized to accept variable numbers of readout chips to make the required half-plane shape. Each readout chip is “flip-chip” mated to 22 columns of 128 rows of pixels on the sensors, corresponding to 2,816 active channels per readout chip. Each readout chip covers an active area approximately $0.64 \text{ cm} \times 0.92 \text{ cm}$. To avoid any dead space between adjoining readout chips, the pixels on the sensors corresponding to the edge of the readout chip (first and last column) are extended to $600 \mu\text{m}$. These hybrid assemblies are supported by a movable carbon substrate that allows the pixel sensors to be positioned a safe distance away from the beam-line until stable conditions have been established in the Tevatron, at which point they are moved as close to the beam-line as radiation damage considerations will allow. This substrate also provides cooling for the readout electronics.

Figure 2.9 shows a conceptual design for the aluminum vacuum vessel and carbon support structure for the pixel detector. The vessel is a rectangular box with a length of $\sim 165 \text{ cm}$ and a height of $\sim 83 \text{ cm}$. Particles within the 300 mrad acceptance of the spectrometer traverse only the pixel stations and the 0.75 mm thick exit window.

Table 2.3: Pixel vertex detector properties

Property	Value
Pixel size	rectangular: $(50 \times 400) \mu\text{m}$
Outer Plane Dimensions	$(10 \times 10) \text{ cm}$
Central Square Hole (adjustable)	nominal setting: $(12 \times 12) \text{ mm}$
Total Planes	60
Total Stations	30
Pixel Orientations (per station)	one with narrow pixel dimension vertical and the other with narrow dimension horizontal
Separation of Stations	4.25 cm
Sensor Thickness	250 μm
Readout Chip Thickness	200 μm
Total Station Radiation Length (incl. rf shielding)	2.5%
Total Pixels	2.2×10^7
Total Silicon Area	$\approx 0.6 \text{ m}^2$
Readout	analog (almost 3 bits, i.e., 7 thresholds)
Trigger	Signals are used in Level I trigger.
Rate Requirements	Time between beam crossings is 132 ns.
Noise Requirement	desired: $< 10^{-6}$ per channel-crossing required: $< 10^{-5}$ per channel-crossing
Resolution	better than $9\mu\text{m}$
Radiation Tolerance	$> 6 \times 10^{14} \text{ particles/cm}^2$
Power per Pixel	$\sim 60 \mu\text{Watt}$
Operating Temperature	$\sim -5 \text{ }^\circ\text{C}$

The carbon substrate will be attached to a support frame made out of carbon fibers. Its position will be controlled by motors located just outside the vacuum vessel.

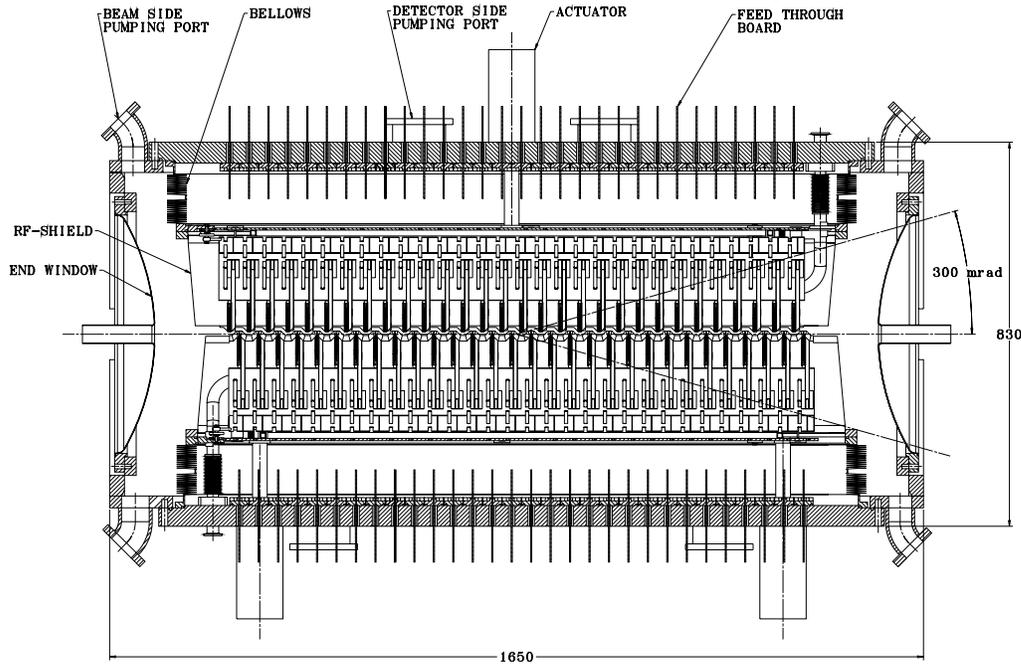


Figure 2.9: Side view of the vacuum vessel and support structure for the pixel detector. The pixel stations are mounted in two halves. Each half sits in a vacuum enclosure separated from the beam by a thin rf shield. Signals are fed through the vacuum vessel via printed circuit boards with high density connectors. Also shown in the figure are actuators and a bellows structure to allow the detectors to be moved in and out of the beam during data-taking and beam refill.

2.5.3 Spatial resolution

BTeV test-beam studies, performed with prototype sensors and readout having pixel sizes of $50 \mu\text{m}$ by $400 \mu\text{m}$, have demonstrated a spatial resolution between 5 and $9 \mu\text{m}$ in the narrow dimension, depending on the track angle of incidence (see Fig. 2.10). The solid line shows the resolution function (Gaussian) used for the Monte Carlo studies. The figure shows both the resolution obtained using 8-bit charge information directly, and also the resolution obtained by degrading the pulse height to 2-bits of information. This result confirms the prediction of the simulations: that excellent resolution can be obtained using charge sharing, even with very coarse digitization. Based on these results it has been decided that the BTeV readout chip will have a 3-bit FADC in each pixel cell. This will provide excellent spatial resolution. In

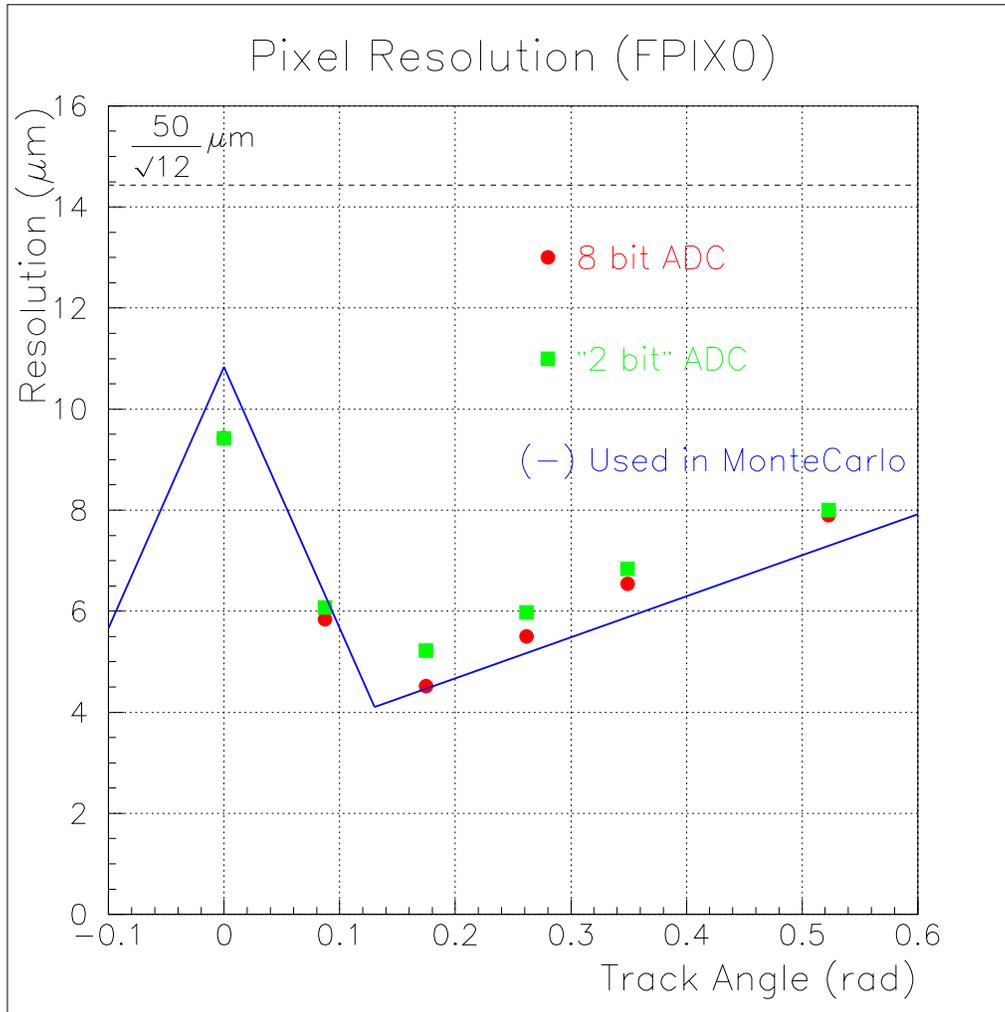


Figure 2.10: Resolution as a function of the angle of the incident beam for both 2-bit and 8-bit ADC readouts. The lines are piecewise linear fits to a simulation of the resolution.

addition, the actual pulse heights may be used to indicate the presence of δ -rays or γ conversions.

The single hit resolution is made possible by the choice of pixel size and a relatively low threshold for readout (approximately 2500 input electrons equivalent compared to about 24000 electrons for a minimum ionizing track at normal incidence for the devices tested). Relatively low dispersion of the thresholds across the chip and low noise in each pixel make the low readout threshold possible. Given the 132 ns beam crossing interval of the Tevatron, time slewing in the chips will

not be a problem. The above performance is for unirradiated devices thus, to minimize effects of radiation damage during the lifetime of the detectors, the pixels will operate at a temperature of about -5 °C. Mounting stability and the necessary pixel alignment, using actual tracks in the final location, will be important to avoid serious degradation of this good resolution.

In order to have also a good track resolution the material in all the components of the system is reduced to the minimum to minimize the multiple scattering. In addition to making the components of the detector proper as thin as possible (see Table 2.3), we allow only a thin membrane between the pixel detector and the beams. Thus, the pixel detector will sit in a secondary vacuum with only a thin aluminum rf shield between that vacuum and the accelerator vacuum. The very close proximity to the interaction region and the spacing between pixel planes is kept to a minimum to reduce the extrapolation distances to vertices, both primary and secondary. All these parameters have been optimized using detailed (MCFast and GEANT) simulations.

2.5.4 Pattern recognition capability

The early choice of pixel technology for the BTeV vertex detector was based, in part, on the space point information that it provides which will help in pattern recognition. Figure 2.11 comes from the 1999 beam-test of BTeV prototype pixel detectors, and shows the power of space points in reconstructing high density tracks. There, an interaction in a carbon target a few mm upstream of the first pixel plane leads to seven tracks reconstructed in much less than 1 cm², a density an order of magnitude more than typical for BTeV.

The pattern recognition capability benefits enormously from the low occupancy, averaging slightly above 1 track per B event in the highest rate readout chip. In addition, the stretching of edge pixels and the shingled mounting (like the lapped wooden siding on a house) of the multichip pixel modules provide complete coverage within the nominal plane acceptance. The regular spacing of planes along the beam also eases the job of the Level 1 trigger.

2.5.5 Radiation hardness

The silicon sensors are based on n+/n/p technology as developed by LHC experiments. The latest readout chips are manufactured with deep sub-micron (0.25 μ m) CMOS technology, an inherently radiation-tolerant process, once enclosed-geometry

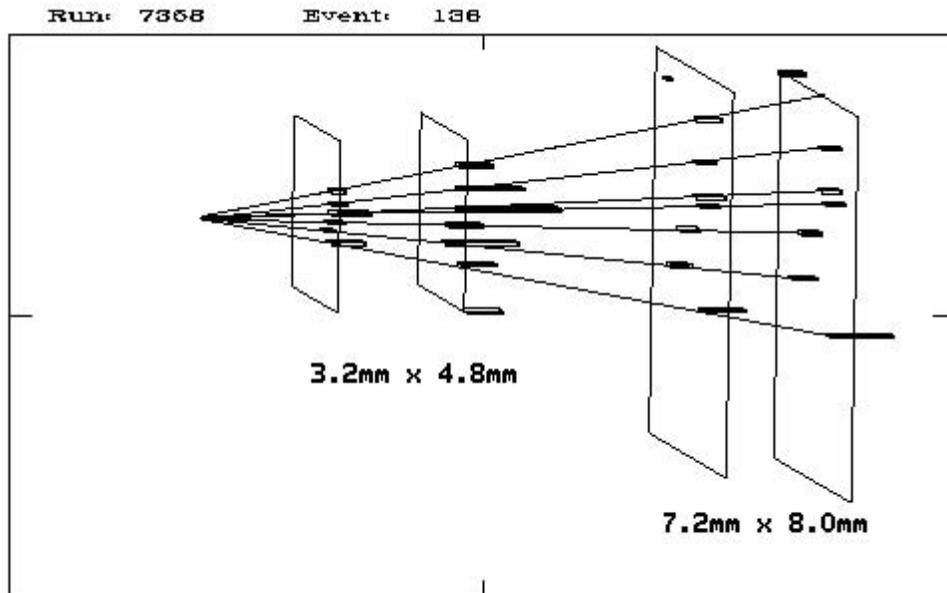


Figure 2.11: Multiparticle interaction observed in Fermilab beam-test. The length of each hit is proportional to the pulse height. The straight lines represent fits to the outgoing tracks.

transistors and appropriate guard ring designs are used. Tests have been made with irradiations up to 0.4×10^{15} 200 MeV protons per cm^2 equivalent for sensors (about 20 MRad) and 0.74×10^{15} 200 MeV protons per cm^2 (equivalent to 43 MRad) for readout chips [30]. These tests show acceptable operation of sensors based on current and capacitance curves vs applied bias voltage, in terms of leakage current, required depletion voltage, and breakdown voltage [31]. The readout chips in deep sub-micron technology appear to be even more radiation-hard. Radiation damage does not seriously affect noise, threshold dispersion, etc. up to these irradiation levels. These irradiation results will be augmented with charge collection and other tests in the test-beam as soon as the beam will be available this year. In addition, the measured rates of single event upset are low enough to be handled easily. No evidence of more serious single event effects has been seen. In addition, the plan is to operate the pixel detector at about -5°C . This will mitigate problems with charge trapping and reverse annealing (the variation in depletion voltage with time).

2.5.6 Readout speed

The pixel readout is data-driven. That is, the readout occurs as soon as data are ready on the readout chip. The token passing from row to row, which is an important part of the potential readout speed, is very fast (0.125 ns per row), and this starts in parallel in all columns. The readout rate allows us to move all the data off chip with negligible loss of data, even if the amount of data is three times that projected for the nominal luminosity of $2 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$. Data output is serialized, but uses a number of parallel readout paths selectable for each readout chip. The data coming off the chip are already highly sparsified, since only pixels above threshold are read out. Sorting out the data and assembling events is done external to the detector in large buffer memories.

2.5.7 Physics capability

Figure 2.12 shows the distribution of $L/\sigma(L)$, which is the normalized detachment between the primary vertex and the B decay vertex, for reconstructed decays $B_s \rightarrow D_s^- K^+$, where, $D_s^- \rightarrow \phi \pi^-$ and $\phi \rightarrow K^+ K^-$. The mean value is 44 standard deviations. Figure 2.13 shows the L-resolution and the proper time resolution for

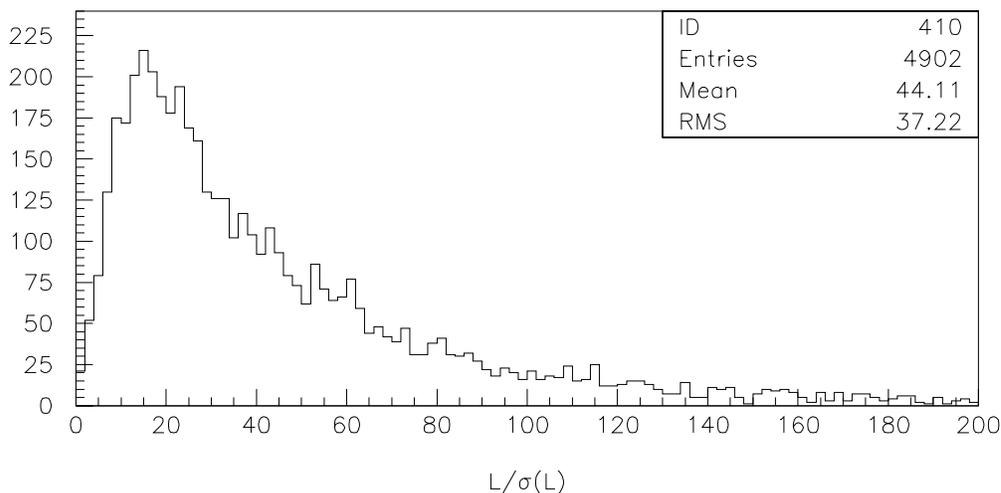


Figure 2.12: Normalized detachment, $L/\sigma(L)$, between the primary vertex and the decay vertex for the decay $B_s \rightarrow D_s^- K^+$.

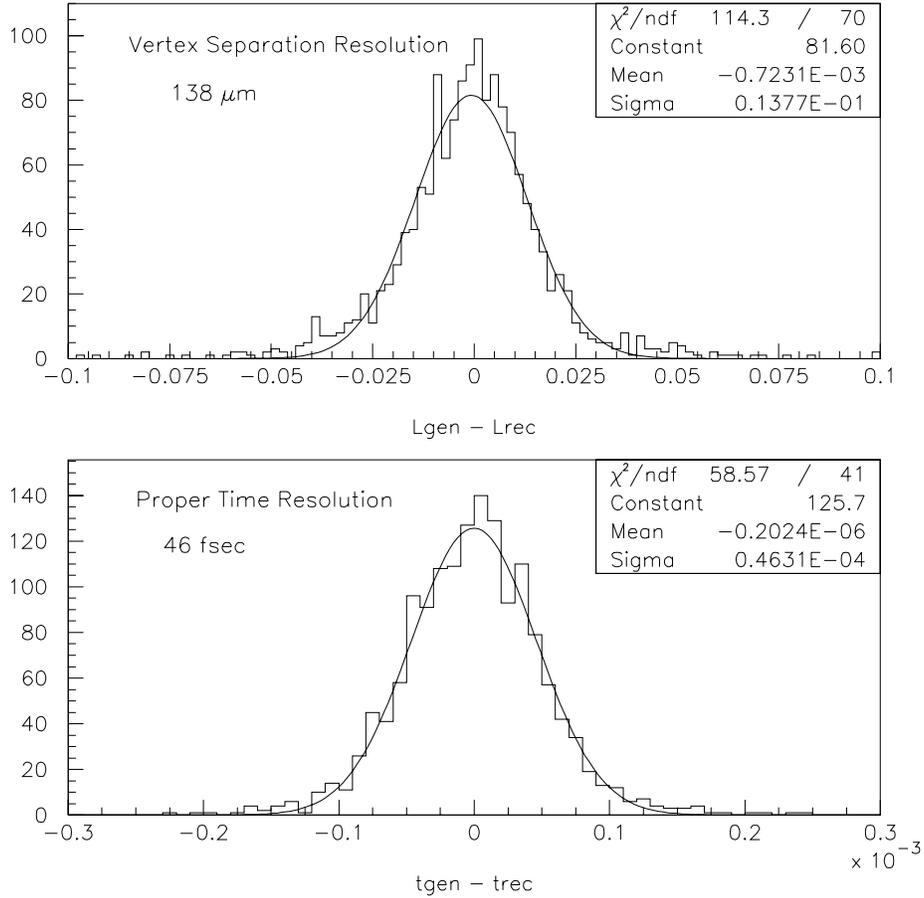


Figure 2.13: Top) The resolution in L, the separation between the primary and secondary vertex. The quantity plotted is the difference between the Monte Carlo generated separation L_{gen} and the reconstructed separation L_{rec} , for the B_s decay. The X- axis is in cm. The L resolution is $138 \mu\text{m}$; and bottom) resolution in proper time. The quantity plotted is the Monte Carlo generated proper time t_{gen} minus the reconstructed proper time, t_{rec} of the B_s decay. The X-axis is picoseconds (10^{-3} nanoseconds). The proper time resolution is 46 fs.

the B_s decay. The resolution in proper time is 46 fs even for this complex multibody decay containing a tertiary vertex (the D_s^- decay). This can be compared with the B_s lifetime of ~ 1500 fs or the B_s mixing period of ~ 400 fs if x_s is about 25. It is clear that the BTeV vertex detector has abundant resolution to carry out detailed

time-dependent analysis even if the B_s were to have a surprisingly high oscillation frequency.

2.6 Forward tracking system

2.6.1 Introduction

The major functions of the forward charged particle tracking system are to provide high precision momentum measurements for tracks found in the pixel system, to reconstruct and measure all parameters for tracks which do not pass through the vertex detector (such as K_s and Λ^0 daughter tracks), and to project tracks into the RICH counters, EM calorimeters, and Muon detectors. Measurements from the forward tracking system are also used on-line in the Level 3 trigger, as explained below.

General description

The baseline forward tracking system consists of 7 stations in one arm, placed transversely to the beam at various distances from the interaction point. Three stations are placed inside the dipole magnet, three stations in the low field region just upstream of the RICH, and one station just downstream of the RICH. The entire system extends over a distance of ~ 7 m and provides polar angle coverage from approximately 10 mr up to 300 mr.

The design of the forward tracking system has been driven by the high density of tracks produced in the forward direction, especially with multiple interactions per crossing. Two different types of detectors are used. Most of the solid angle is instrumented using straw tube drift chambers. Straws have been chosen because they can be used to make large chambers with small cell size, and because they are immune to catastrophic failure of an entire detector from a single wire break. The central region close to the beam is instead instrumented with silicon microstrip detectors because the track density requires a detector with high granularity. Tables 2.4 and 2.5 list all the geometric parameters and the main characteristics of the forward tracker. This forward tracking system configuration has sufficient segmentation to handle the high hit multiplicities that are expected when $b\bar{b}$ events are produced in the forward region along with minimum bias events. Figure 2.14 shows occupancies in the straw tracker predicted by BTeVGeant for the case in which a $b\bar{b}$ event is produced at the design luminosity of $2 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$. The maximum occupancy

Table 2.4: Properties of the baseline forward silicon tracker (1 arm)

Property	Value
Si-sensors	$\sim 7 \times 7 \text{ cm}^2$, <i>p-on-n</i> type
Pitch	100 μm
Thickness	200 μm
Sensor configuration	4 ladders of 4 sensors
Coverage	27cm \times 27cm
Central hole	5.4 cm \times 5.4 cm (7 cm \times 7 cm in last station)
Total stations	7
Z positions (cm)	99, 142, 200, 292, 336, 386, 729
Views per station	3 (X, U, V)
Channels per view	$\sim 5,600$
Total channels	$\sim 127,600$
Readout	sparsified binary

Table 2.5: Properties of the baseline forward straw tracker (1 arm)

Property	Value
Straw size	4 mm diameter
Central hole	27 cm \times 27 cm
Total Stations	7
Z positions (cm)	95, 138, 196, 288, 332, 382, 725
Half size (cm)	27, 41, 61, 88, 102, 116, 204
Views per station	3 (X,U,V)
Layers per view	3
Total number of straws	29,088
Total station thickness	0.6% X_0
Total channels	58,176
Readout	ASD + timing chip (6 bits), sparsified

is 4% in the silicon strip detectors, which have 40 times finer pitch than the straw chambers.

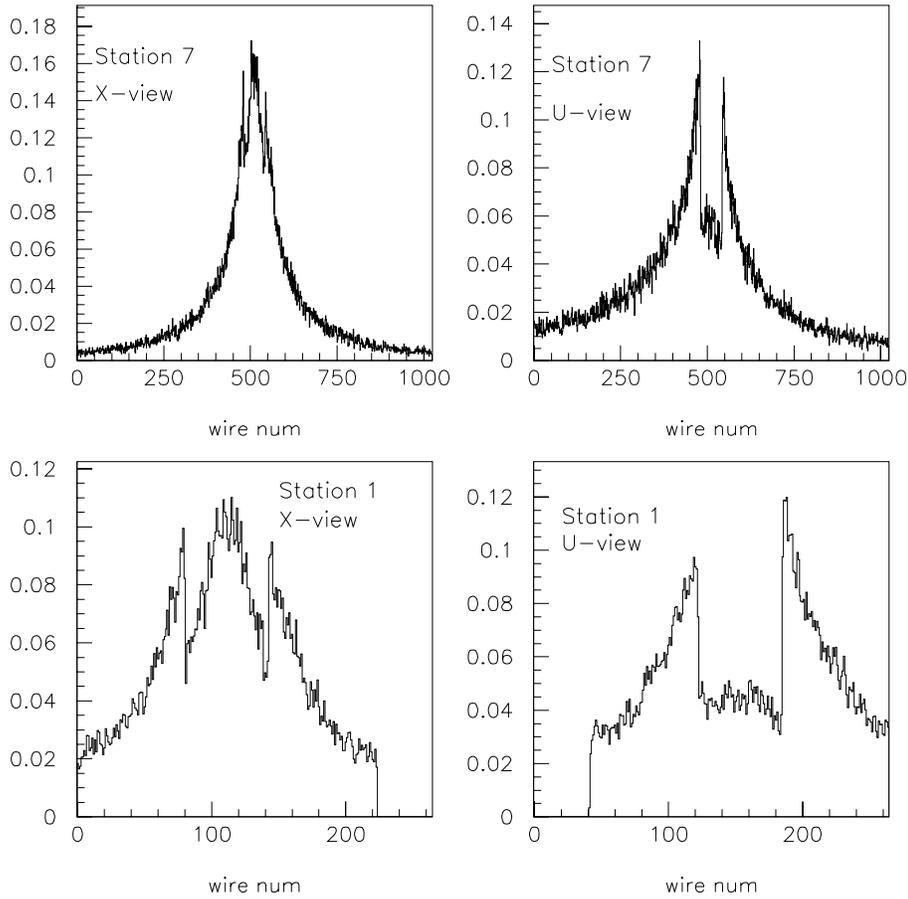


Figure 2.14: Occupancies in the first station of straws, Station 1, and the station just upstream of the EM Calorimeter, Station 7, when a $b\bar{b}$ event is produced at the design luminosity of $2 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$. The two histograms on the left are for X-view straws, while those on the right are for U-view straws. The V-views have identical occupancies to the U-views.

Forward tracker performance

The system just described ensures excellent tracking performance over the full acceptance of the forward spectrometer. Figure 2.15 shows the expected average fractional momentum resolution for b decay products as a function of track momentum and of the track production angle with respect to the beam axis. For these histograms, an effective position resolution of $\sigma_{X,U,V} = 150 \mu\text{m}$ was assumed for each view of the straws and a resolution of $\sigma_{X,U,V} = 29 \mu\text{m}$ assumed for the silicon strip detectors.

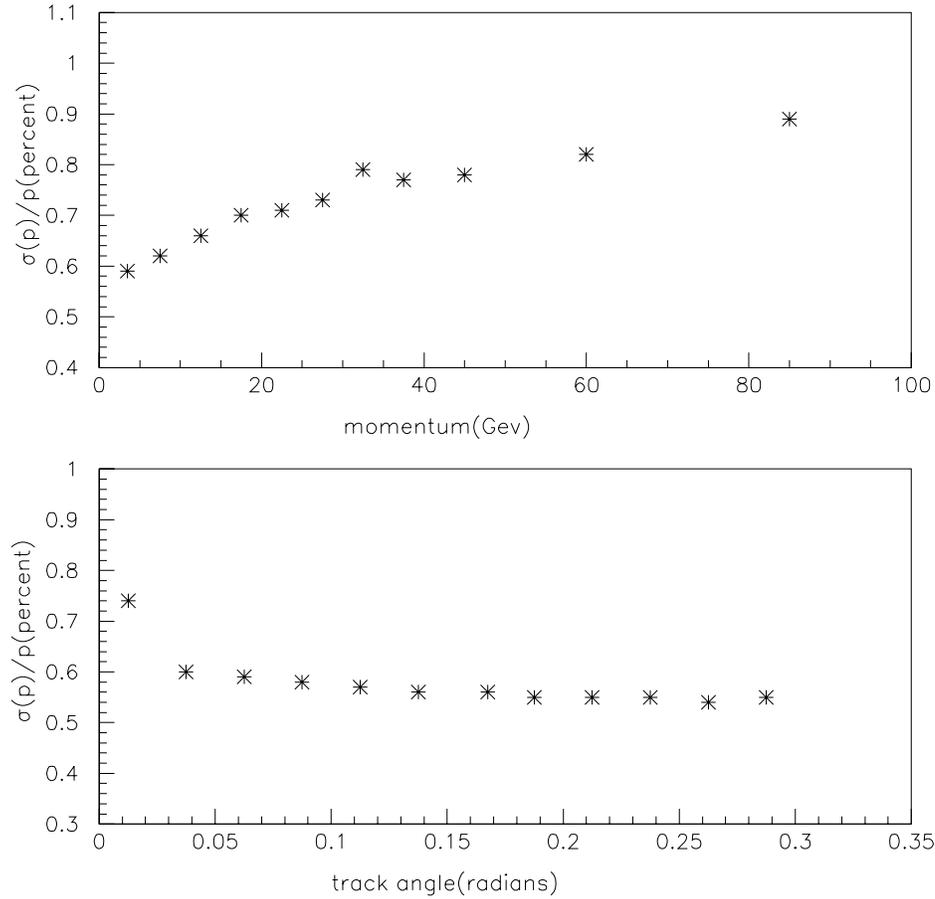


Figure 2.15: Momentum resolution as a function of track momentum (upper plot) and as a function of polar production angle (lower plot) for b decay products.

2.6.2 Forward silicon tracker

Detector description and layout

The strip detector consists of stations with three planes of $200 \mu\text{m}$ thick single-sided silicon microstrip detectors with $100 \mu\text{m}$ pitch. The silicon sensors, which have an area of $7 \times 7 \text{ cm}^2$, are arranged in ladders of 4 daisy-chained sensors each in such a way that four adjacent ladders form a plane as illustrated in Fig. 2.16. The ladders are mounted on a low mass carbon fiber support which is designed to ensure a relative proper alignment among all the elements of the plane.

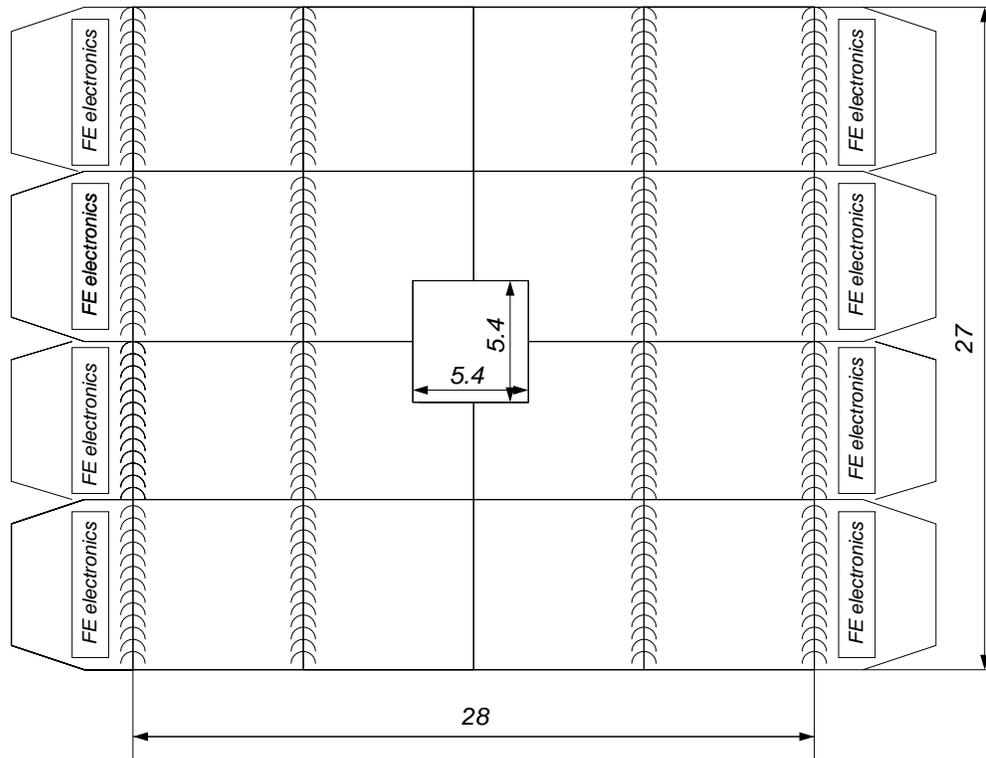


Figure 2.16: Sketch of a silicon detector plane. It consists of 4 ladders of four daisy-chained Si-sensors. The two pairs of sensors on each ladder are read out separately by the front-end electronic chips placed at the two ends of the same ladder. There is some overlap between adjacent ladders to ensure good efficiency over the entire plane. Dimensions are in centimeters.

Three views, called X , U and V , are provided by rotating two of the planes. The two stereo views, U and V , are at $\pm 11.3^\circ$ around the Y bend coordinate. Each plane consists of about 5,600 readout channels; the entire system of 7 stations has about 128,000 channels in total (1 arm).

The Si-sensors are of the standard p -on- n type, with multiple guard rings to allow high voltage operation. The front-end electronics is distributed along the two opposite edges of each plane where it is cooled by a fluid circulating in a duct embedded in the support structure around the periphery of the plane.

The preamplifier chips are AC coupled to the strips by means of capacitors directly integrated on the sensors. Each channel is read out in binary mode providing a $\sigma = 100 \mu\text{m}/\sqrt{12} = 29 \mu\text{m}$ resolution.

Radiation issues

In BTeV the radiation level at the silicon detectors decreases rapidly with increasing distance from the beam. Important radiation damage effects will be confined to a small region closest to the beam line and at the station closest to the interaction region. As shown in Fig. 2.17, the maximum value of the fluency is expected to be $\sim 1.6 \times 10^{13}$ particles/cm²/year, given a luminosity of 2×10^{32} cm⁻² s⁻¹. This is a relatively low radiation dose that can affect at most a minor portion of the detector close to the beam after several years of operation.

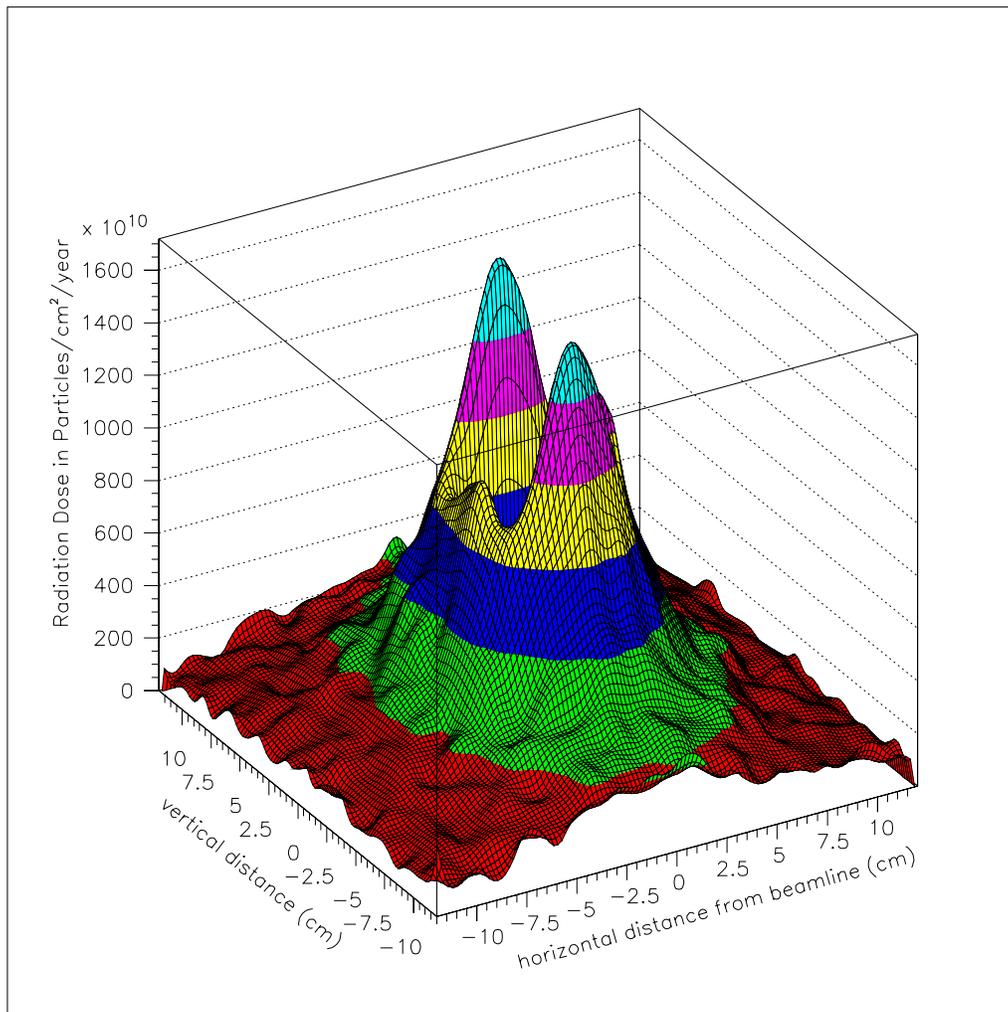


Figure 2.17: Radiation dose as a function of position in forward silicon tracker station # 1. The horizontal magnetic field concentrates more particles above and below the square central beam hole than on either side.

Readout electronics

Even given the low occupancy expected in the Forward Silicon Tracker, the output bandwidth required to read out all hit information from every crossing is higher than is provided by any silicon strip detector (SSD) readout chip. For this reason a new readout chip with very high readout bandwidth has been developed at the universities of Pavia and Bergamo in collaboration with Fermilab. The new preamplifier features an ENC $\sim 800 e^-$ for semi-Gaussian shaping with 60 ns peaking time and a capacitive load at the input of ~ 20 pF, as expected for our longest strips. The binary readout is a simplified version of the readout scheme implemented in the FPIX2 pixel readout chip. The SSD readout chips are designed to interface to the same electronics employed to read out pixel chips. The new readout chip is implemented using $0.25 \mu\text{m}$ CMOS technology, following the radiation tolerant design rules developed for the FPIX2 design.

2.6.3 Forward straw tracker

Detector description and layout

The forward straw tube tracker consists of stations that provide 3 coordinate measurements, X , U and V , where the two stereo views, U and V , are at $\pm 11.3^\circ$ around the Y bend coordinate, same as in the forward silicon detector. With three layers per view, this configuration provides excellent resolution in the bend plane while maintaining a robust ability to reject ghost combinations of hits. It has sufficient redundancy to achieve a high detection efficiency and to resolve the left/right ambiguity a very large fraction of the time. The unit of construction is the “half-view”, itself composed of a number of 48 straw modules. Two half-views fit around the beam-pipe to make up a single view. All the sense wires for the straw cells that do not terminate at the central hole are divided electrically using a small glass capillary. This cuts the occupancy rates in half. In addition, within a 27 cm square region of silicon strips, all straw sense wires are deadened by using two glass capillary beads to isolate the central section of the sense wire. This is done to lower the straw occupancy rates. The time between bunch crossings in the Tevatron will be 132 ns by the time BTeV is operational. This time allows the use of standard gases like Argon-Ethane or Argon-CO₂, which have drift times in the 4mm straws on the order of 60 ns.

Front end electronics and drift time measurement

The straw tube chambers will be instrumented using electronics developed by the University of Pennsylvania [32]. These radiation hard integrated circuits include high gain preamplifiers, pole-zero networks for pulse shaping and ion-tail cancellation, and leading edge discriminators.

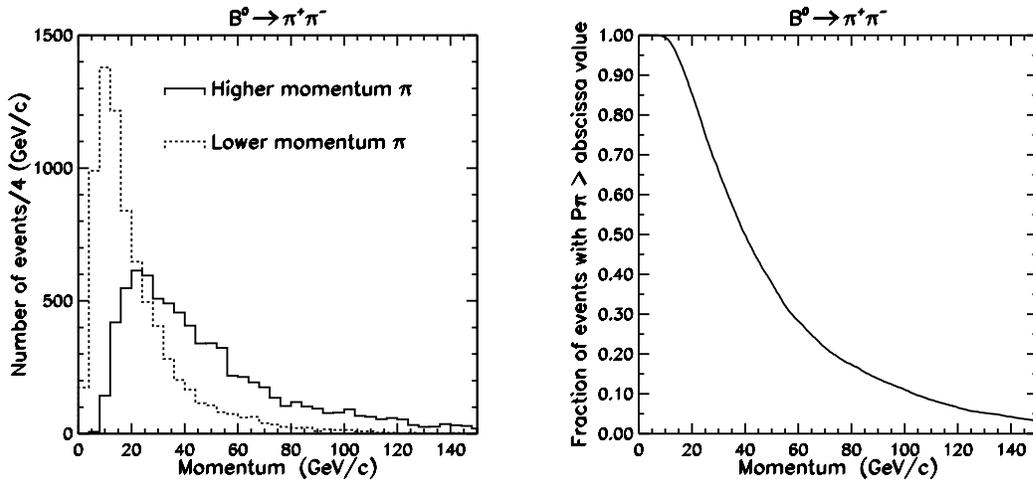
The drift time will be measured using digital TDC's. The information from the straw tracker is digitized and read out for every crossing. A six-bit single-hit TDC, with 1.5 ns wide bins covering 96 ns, will provide a drift distance measurement precision better than 100 μm .

2.7 Charged particle identification

Excellent charged hadron particle identification is a critical component of a heavy quark experiment. Even for a spectrometer with the excellent mass resolution of BTeV, there are kinematic regions where signals from one final state will overlap those of another final state. For example, $B_s \rightarrow D_s K^-$ signal must be distinguished from $B_s \rightarrow D_s \pi^-$ background in order to measure the CKM phase γ . These ambiguities can be eliminated almost entirely by the BTeV RICH detector.

2.7.1 Requirements

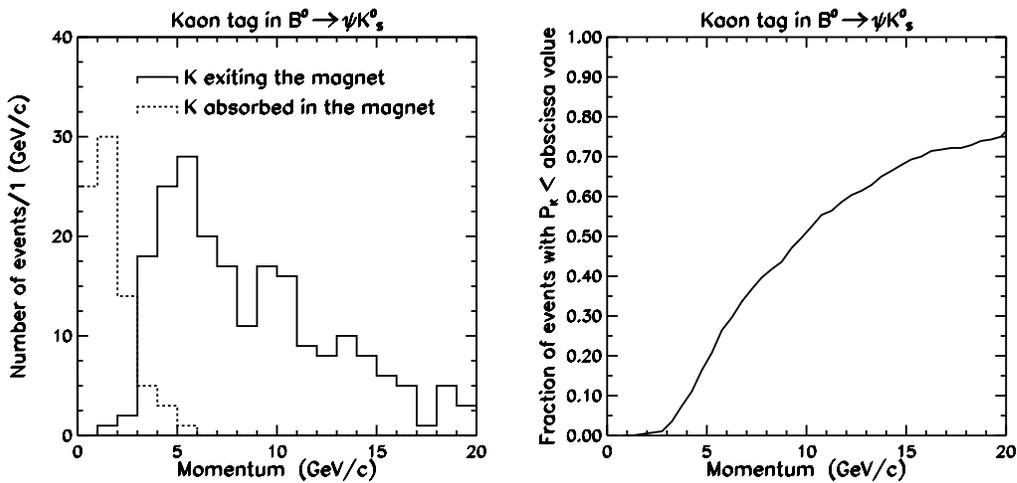
In the design of any particle identification system, the dominant consideration is the momentum range over which efficient separation of the various charged hadron types – π , K , and p – must be provided. In BTeV, the physics goal which sets the upper end of the momentum requirement is the desire to cleanly separate $B_d^0 \rightarrow \pi^+ \pi^-$ from $B_d^0 \rightarrow K^+ \pi^-$ and $B_s^0 \rightarrow K^+ K^-$. These two-body decays produce reasonably high momentum pions and kaons. Figure 2.18 shows the momentum distribution of pions from the decay $B_d^0 \rightarrow \pi^+ \pi^-$ for the case where the two particles are within the spectrometer's acceptance. The low momentum requirement is defined by having high efficiency for "tagging" kaons from generic B decays. Since these kaons come mainly from daughter D -mesons in multibody final state B -decays, they typically have much lower momentum than the particles in two body decays. Figure 2.19 shows the momentum distribution of tagging kaons for the case where the signal particles are within the geometric acceptance of the spectrometer. About 1/5 of the tagging kaons never exit the end of the spectrometer dipole. Almost all kaons exiting the dipole have momenta above 3 GeV. The momentum range required for



(a) Distributions for the lower (dashed line) and higher (solid line) momentum pion in this decay.

(b) Distributions for the higher momentum pion in this decay in integral form, which gives loss of efficiency as a function of the high momentum cut-off of the particle ID.

Figure 2.18: The momentum distribution of pions in $B_d \rightarrow \pi^+ \pi^-$



(a) Distributions for kaons absorbed in (dashed line) and exiting from (solid line) the magnet.

(b) Distributions for kaons exiting from the magnet in integral form, which gives loss of efficiency as a function of the low momentum cut-off of the particle ID.

Figure 2.19: The momentum distribution of “tagging” kaons for the case where the signal particles (ψK_S^0) are within the geometric acceptance of the spectrometer.

the particle identification is then

$$3 \text{ GeV}/c < P_{\text{particle id}} < \sim 70 \text{ GeV}/c \quad (2.3)$$

Finally, kaons and pions from directly produced charm decays have momenta which are not very different from the kaons from B -decays. The range set by the B -physics requirements is a reasonable, if not optimal, choice also for charm physics.

2.7.2 RICH radiators

Because of the large particle momenta there is really only one choice of detector technology: a gaseous ring-imaging Cherenkov counter. Pions and kaons can be separated in the required momentum region with a single gas radiator. The gas used is C_4F_{10} which has an index of refraction of 1.00138 in the visible range. The momentum dependence of the Cherenkov angle for pions, kaons and protons in this gas is shown in Fig. 2.20. Note that below about 9 GeV, no gas can provide K/p separation and that, since kaons are below threshold, the RICH operates in a threshold mode for (K or p) vs. π separation. In order to improve identification of low momentum particles there is also a liquid radiator. The selected liquid, C_5F_{12} ,

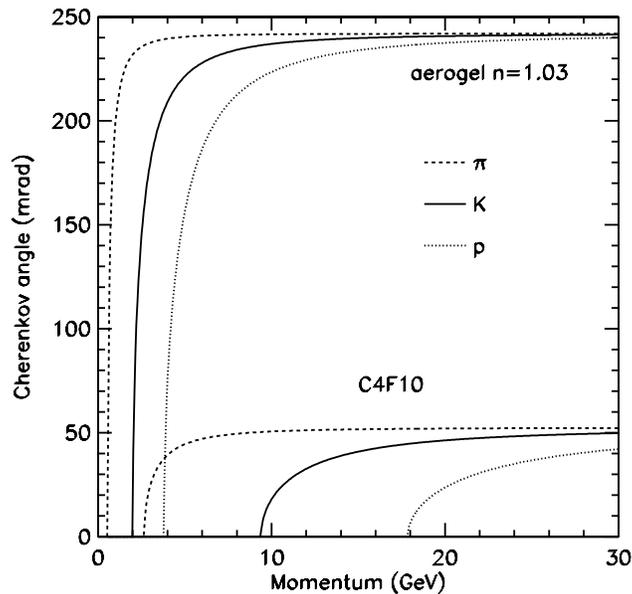


Figure 2.20: Cherenkov angles for various particle species as a function of particle momentum for C_4F_{10} and liquid C_5F_{12} ($n = 1.24$) radiators.

has an index of refraction of 1.24. This produces an intense Cherenkov rings even from a very thin layer of liquid.

2.7.3 RICH dimensions

The RICH detector is located behind the tracking chambers just outside the central dipole magnet, about 4 meters away from the interaction point. The length of the RICH detector must be less than 3 meters to allow sufficient space for the EM calorimeter and the muon system.

The liquid radiator with thickness of about 1 cm will be mounted at the entrance to the RICH vessel. It will cover the entire RICH entrance window, except for a rectangular square around the beam-pipe. The liquid is supported by a 3 mm thick carbon fiber backplane and a 3 mm thick quartz window (for radiation hardness). Spherical mirrors at the end of the gas volume reflect Cherenkov photons, radiated in the C_4F_{10} and focus them into rings at the photo-detection surface.

The geometry of the RICH detector is shown in Fig. 2.21.

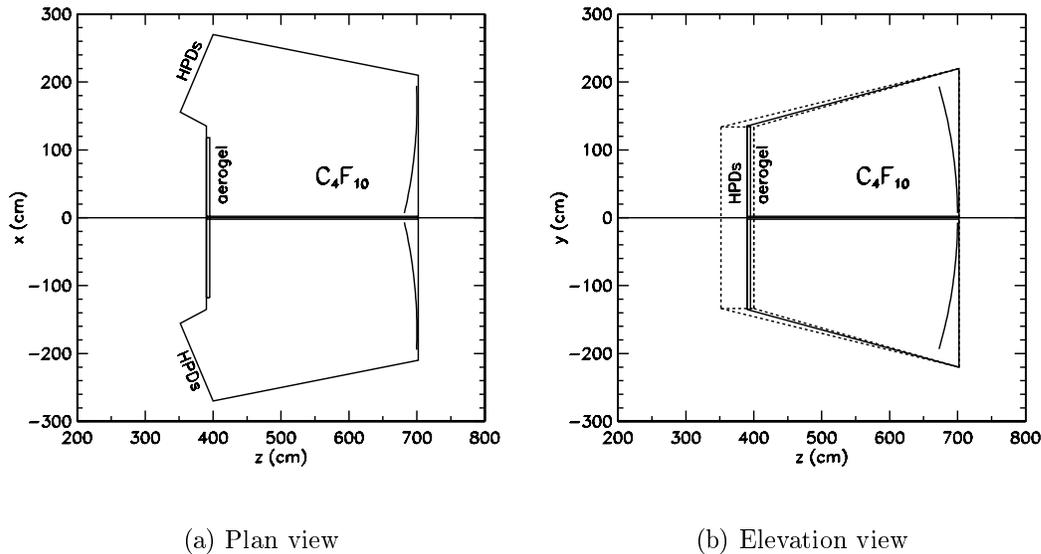


Figure 2.21: BTeV RICH geometry outline. Note the position of the C_5F_{12} on the upstream window and the location of the HPD and photomultiplier sensor planes.

2.7.4 Photo-detectors

Because of the open geometry of the forward spectrometer and the availability of space to install shielding to protect detection elements from the fringe field of the BTeV analysis magnet, arrays of photo-multipliers (PMT) and hybrid photo-diodes (HPD) are used to detect the Cherenkov light emitted in the detector. The photons from the liquid radiator are detected on the PMTs on the sides, the top, and the bottom of the counter and the HPD plane only needs to be sized to detect the photons radiated in the C_4F_{10} gas volume, see Fig. 2.22.

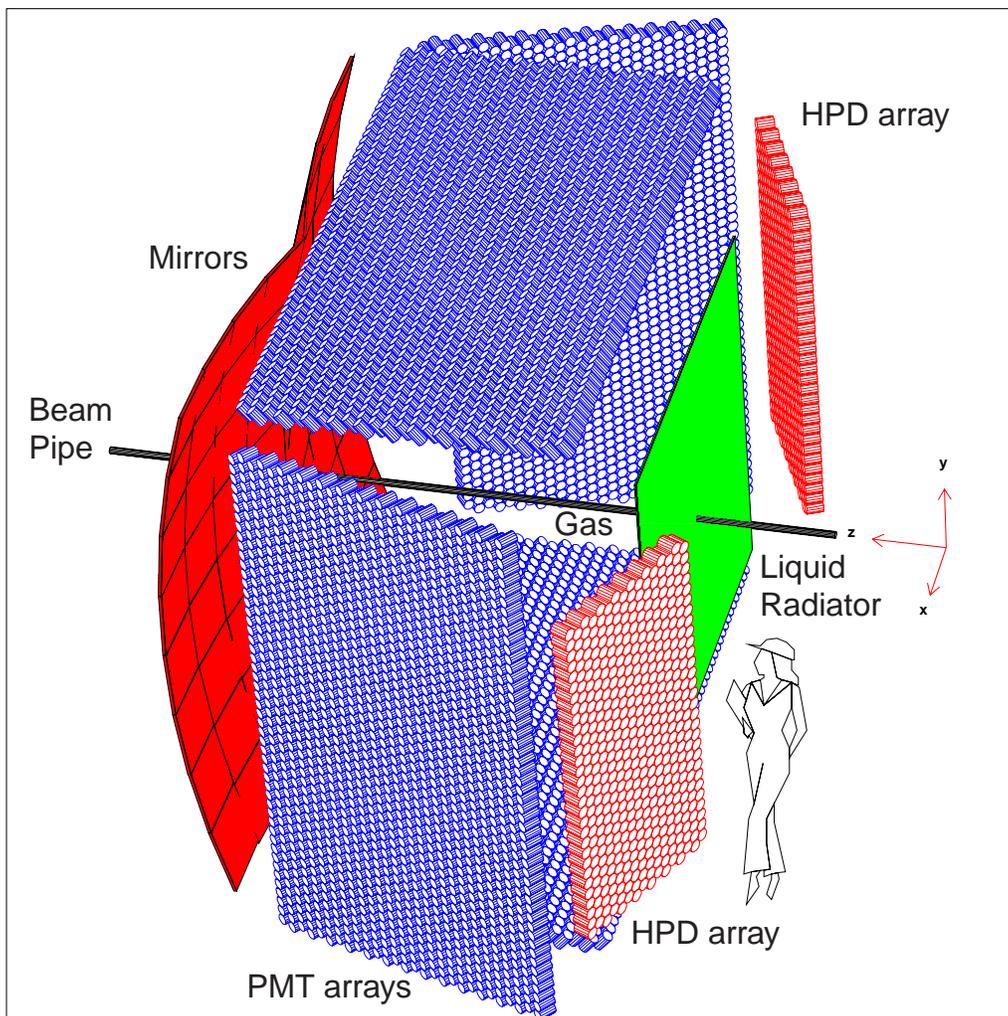


Figure 2.22: Outline of the important RICH components.

2.8 Electromagnetic calorimeter

2.8.1 Introduction

A thorough investigation of B decays requires the ability to study decay modes containing single photons, π^0 's, η 's and electrons. Furthermore, the identification of electrons is useful to reconstruct J/ψ decays and to identify semileptonic decays, both for their intrinsic physics interest and as “flavor tags” for mixing and CP violation studies. Some of the important decay modes for BTeV include: $B^0 \rightarrow (\rho\pi)^0 \rightarrow \pi^+\pi^-\pi^0$, $B_s^0 \rightarrow \psi\eta$, and $\psi\eta'$, semileptonic decays, and $B^0 \rightarrow K^{*0}\gamma$ and $\rho^0\gamma$.

The calorimeter is made of PbWO_4 crystals. This technology has been developed for high energy physics by CMS. The choice of lead tungstate is based on several considerations:

- It satisfies the requirements on energy and spatial resolution. Blocks of appropriate transverse and longitudinal size can be manufactured.
- This material is very resistant to radiation damage, especially when doped with either Nb or La.
- It is fast – crystals deliver 99% of their light output within 100 ns, which is safely less than the bunch crossing time of 132 ns at the Tevatron.

The properties of PbWO_4 which are important for the calorimeter are given in Table 2.6.

Table 2.6: Properties of PbWO_4

Property	Value
Density (g/cm^3)	8.28
Radiation Length (cm)	0.89
Interaction Length (cm)	22.4
Light Decay Time (ns):	5(39%) 15(60%) 100(1%)
Refractive Index	2.30
Maximum of emission (nm)	440
Temperature Coefficient ($\%/^\circ\text{C}$)	-2
Light output/NaI(Tl) (%)	1.3
Light output (pe/MeV into a 2" PMT)	10

2.8.2 Description of the BTeV calorimeter

The crystals are 220 mm long and 28 mm×28 mm in cross section at the rear face. They are slightly tapered to point toward the interaction region to provide a projective geometry. Figure 2.23 shows a representation of the calorimeter, with the crystal hits displayed, for an event generated with GEANT containing a $B^0 \rightarrow \rho^0 \pi^0$ decay. The two photons from the π^0 decay are indicated by the circles. One photon has 19.3 GeV of energy, while the other has 2.4 GeV. The minimum energy displayed per crystal is 10 MeV. This corresponds to the minimum energy crystal that we use

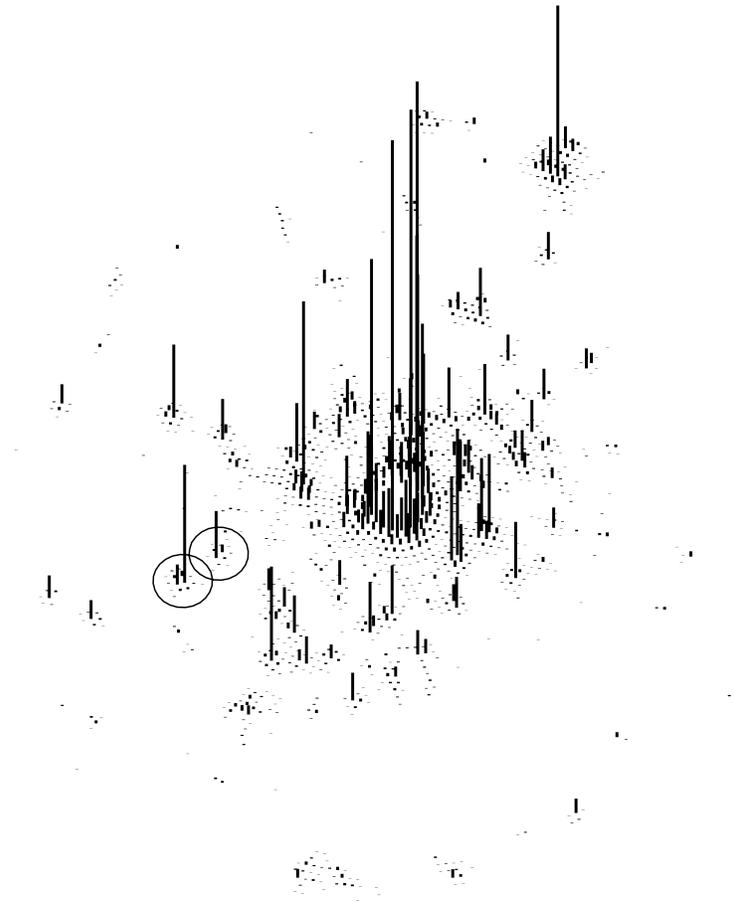


Figure 2.23: The energies in the PbWO_4 calorimeter (one arm) for an event containing two photons from the decay sequence $B^0 \rightarrow \rho^0 \pi^0$, $\pi^0 \rightarrow \gamma\gamma$. The photons of energies 19.3 and 2.4 GeV are surrounded by circles. All energies above 10 MeV are shown, with the height of line proportional to the energy.

in measuring the shower energy. It can be seen even from this one event that there is much more activity near the beam line than further out in radius.

The EMC is far away from the magnetic field thus allowing the use of photomultiplier tubes to read out the calorimeter. The output of the PMT will go to a modified QIE chip [33] located near the photomultiplier base, but outside the region of intense radiation. This chip provides a digitized charge output for each beam crossing. The expected light output is 5000 photoelectrons at 1 GeV. The detector will be housed in a temperature and humidity controlled hut. There will be a dry air environment. Temperature stabilization is necessary because of the thermal coefficient of the PbWO_4 light output. In addition, the gains will be monitored with a light pulsing system based on Light Emitting Diodes.

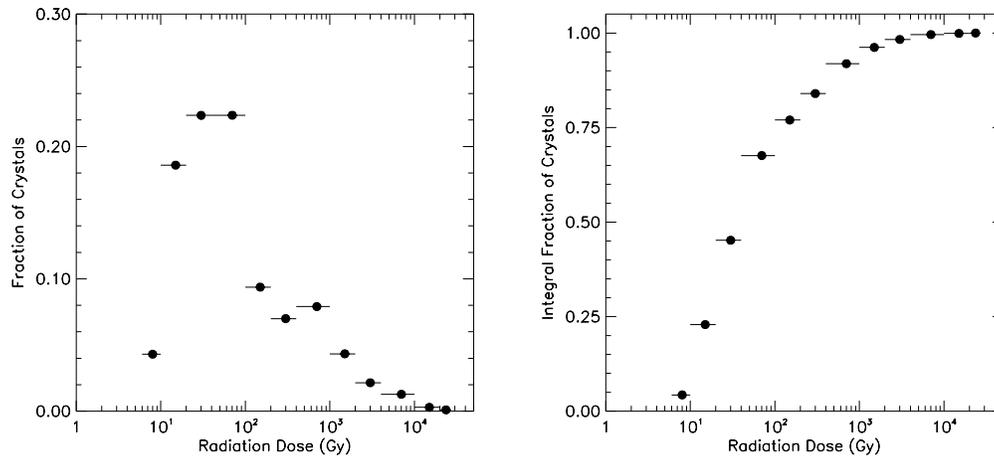
2.8.3 Radiation levels and radiation tolerance

Radiation damage of PWO crystals is a serious issue. Detailed studies [34] reveal that the light transmission of crystals deteriorates due to formation of color centers by radiation, while the scintillation mechanism itself seems unaffected. When a PWO crystal no longer receives radiation, its color centers (semi-stable excited states) disappear, and it recovers from transmission degradation by natural room-temperature annealing. In fact, this annealing goes on even during radiation exposure. Therefore, when crystals are exposed to a constant radiation level, they lose light only up to the point where the rates of radiation damage and natural recovery balance. The integrated dose rates for most crystals are quite modest. In Fig. 2.24 (left) it is shown the dose distribution in the crystals (here dose means the maximum dose in any part of the crystals), for 1 year of running at a luminosity of $2 \times 10^{32} \text{cm}^{-2} \text{s}^{-1}$. It is also shown (right) the cumulative fraction of crystals with doses less than that shown on the horizontal axis. We see that $\sim 90\%$ of the crystals have a yearly accumulated dose of less than 1000 Gy (0.1 Mrad).

2.8.4 Expected resolution

Overall it is expected the energy resolution to be

$$\frac{\sigma_E}{E} = \sqrt{\frac{a^2}{E} + b^2} = \frac{a}{\sqrt{E}} \oplus b, = \frac{1.7\%}{\sqrt{E}} \oplus 0.55\% \quad (2.4)$$



(a) Maximum dose distribution in the crystals for for 1 year of running at a luminosity of $2 \times 10^{32} \text{cm}^{-2} \text{s}^{-1}$.

(b) The integral of the maximum dose distribution.

Figure 2.24: The maximum dose distribution.

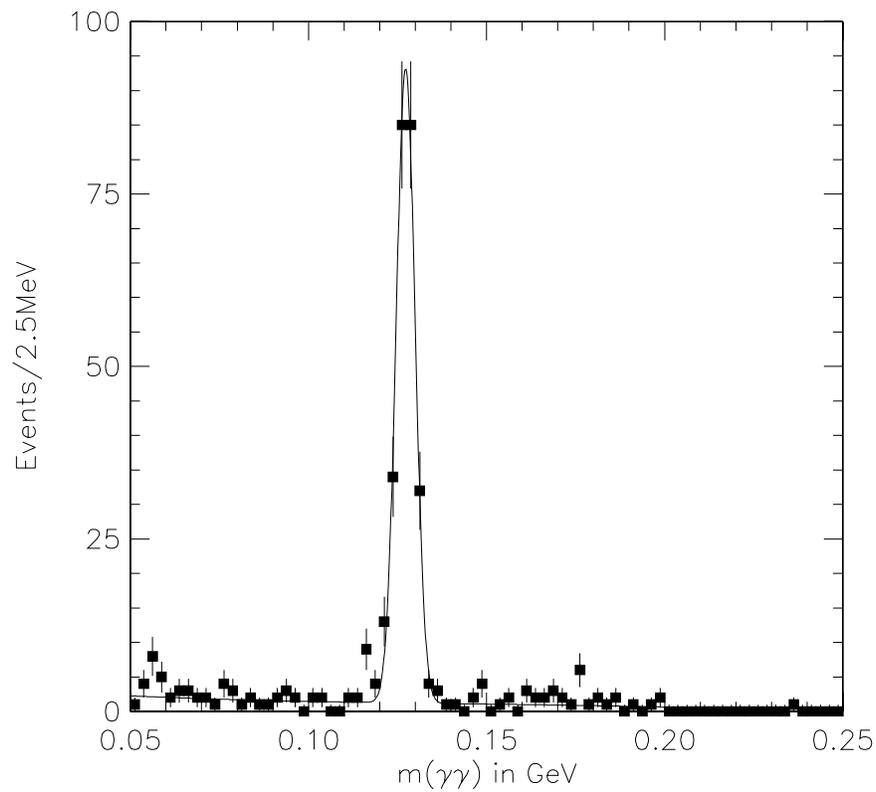


Figure 2.25: The $\gamma\gamma$ invariant mass for 10 GeV π^0 's incident on the calorimeter. The fit is to a Gaussian signal function plus a polynomial background.

The spatial resolution in both directions transverse to the crystal axis is expected to be

$$\sigma_x = \frac{3500 \mu m}{\sqrt{E}} \oplus 200 \mu m . \quad (2.5)$$

Test-beam results are consistent with these expectations. With these single photon resolutions, the π^0 and η mass resolutions are excellent.

Fig. 2.25 shows the invariant $\gamma\gamma$ mass for π^0 's of 10 GeV energy incident on the calorimeter. The energies and positions are simulated by GEANT and a cluster finder algorithm is used to reconstruct the photons. The mass resolution is 2.6 MeV.

2.9 Muon detector

The BTeV muon system has two primary functions:

- J/ψ and Prompt Muon Trigger: besides providing interesting physics (including J/ψ final states of B decays, direct J/ψ production, and semileptonic decays), this trigger performs an important service role by selecting a large enough sample of b events on which the more aggressive and technically challenging vertex trigger can be debugged and its efficiency evaluated.
- Muon Identification: many of the experiment's physics goals (rare decay searches, CP violation studies which require tagging, studies of beauty mixing, searches for charm mixing, etc.) rely on efficient muon identification with excellent background rejection.

The toroidal magnet design combined with fine-grained tracking elements chosen, permits a "stand-alone" trigger: *i.e.* a di-muon trigger based solely on information from the muon detector. In addition, improved background rejection is possible by comparing this measurement with momentum and tracking information from the rest of the spectrometer.

2.9.1 Baseline geometry

The baseline geometry is shown in Fig. 2.26. A cross section of the toroid system is given in Fig. 2.27.

Two toroids, 1 m long with 1.5 T fields, provide the bending power. The muon detectors will be set up in three stations, one between the toroids and two behind the

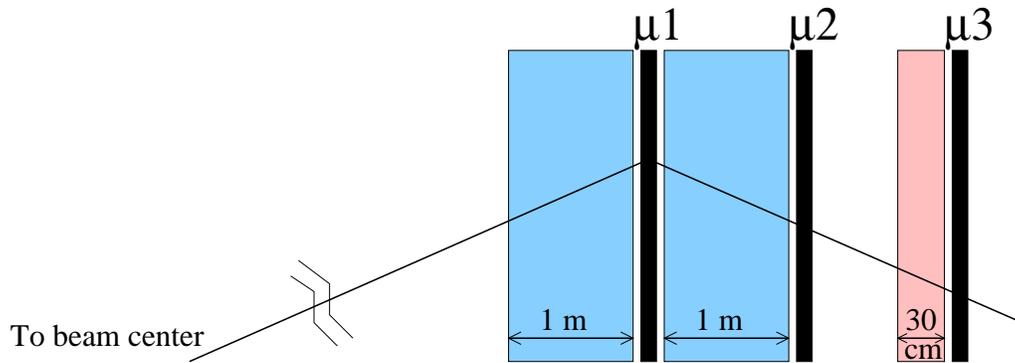


Figure 2.26: Layout of the baseline geometry. The three dark boxes, labeled μ_1 , μ_2 , and μ_3 , represent detector stations with 4 measurement views per station. The two lighter boxes with lengths of 1 m represent magnetized steel toroids, which provide bending power for the muon momentum measurement and which also serve as hadron absorbers. The 30 cm long lighter box is an unmagnetized iron shield. The downstream trajectory is measured by μ_2 and μ_3 . The upstream trajectory is measured using the nominal beam center with possible help from μ_1 .

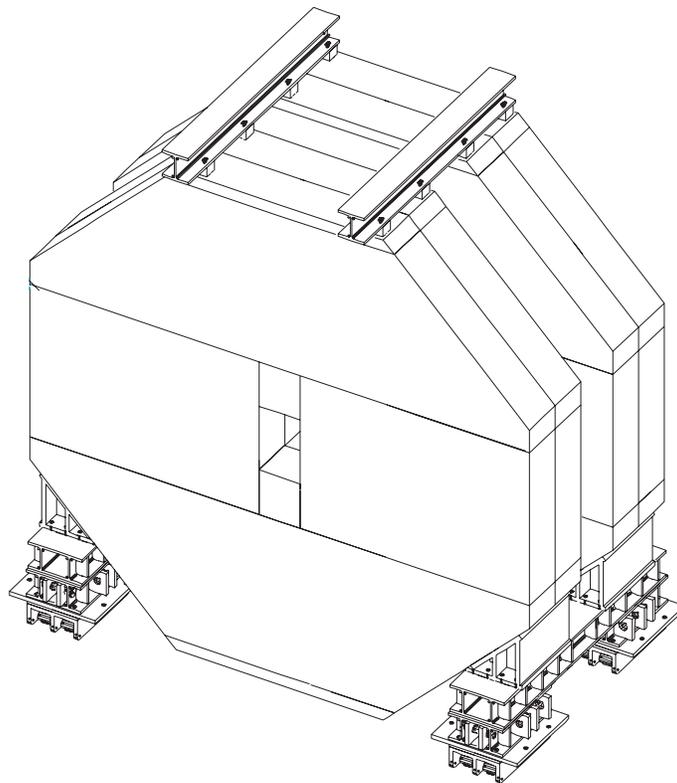


Figure 2.27: Sketch of a muon toroid

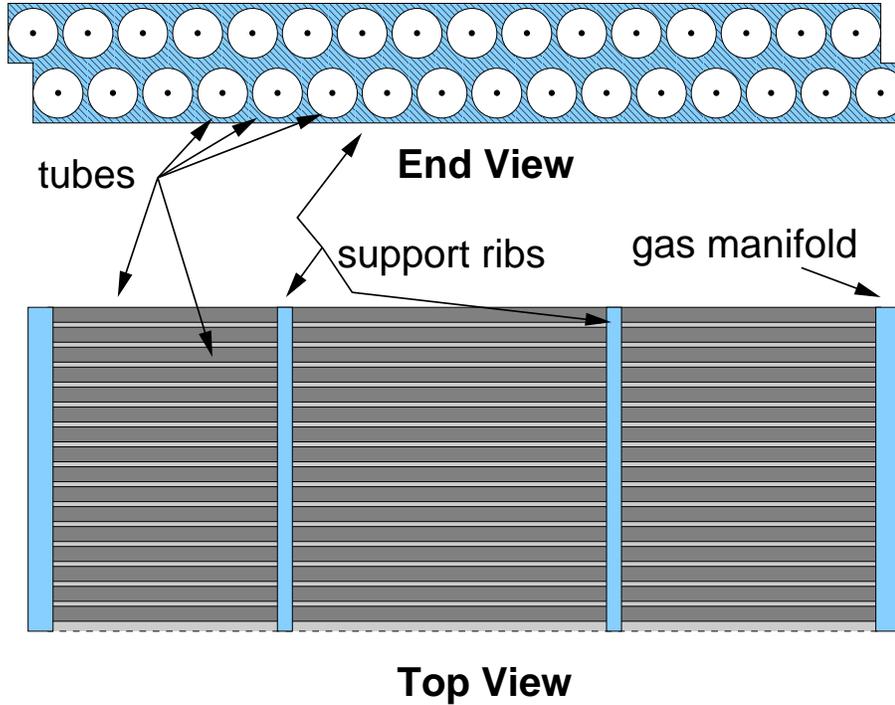


Figure 2.28: End and top views of one “plank” of proportional tubes.

toroids. The momentum can be measured using the two, well shielded, downstream stations and the nominal beam constraint. The station between the two toroids provides a powerful confirming hit to eliminate fake tracks. The basic building block in the construction of a detector station is a “plank” of $3/8$ ” diameter stainless steel proportional tubes as shown in Fig. 2.28. Proportional tubes have been selected as the detector technology because they are robust and have the necessary rate capability. A fast gas (*e.g.* 88% Ar, 10% CF₄, 2% CO₂ ($v_d \approx 9$ cm/ μ s) [35]) will be used so the maximum collection time (drift plus charge integration) for a signal will be less than 60 ns.

To minimize occupancy at small radii and to minimize pattern recognition confusion, each detector station will consist of eight overlapping pie shaped “octants”, as shown in Fig. 2.29a. The four views (r , u , v , and r) in each octant are shown in Fig. 2.29b. The r (radial) view is repeated to provide redundancy for the most important (bend) view and to help reject fake tracks in the trigger. The u and v views are rotated $\pm 22.5^\circ$ from the r view and are used to measure ϕ and to resolve hit ambiguities.

A summary of the baseline BTeV muon system is given in Table 2.7.

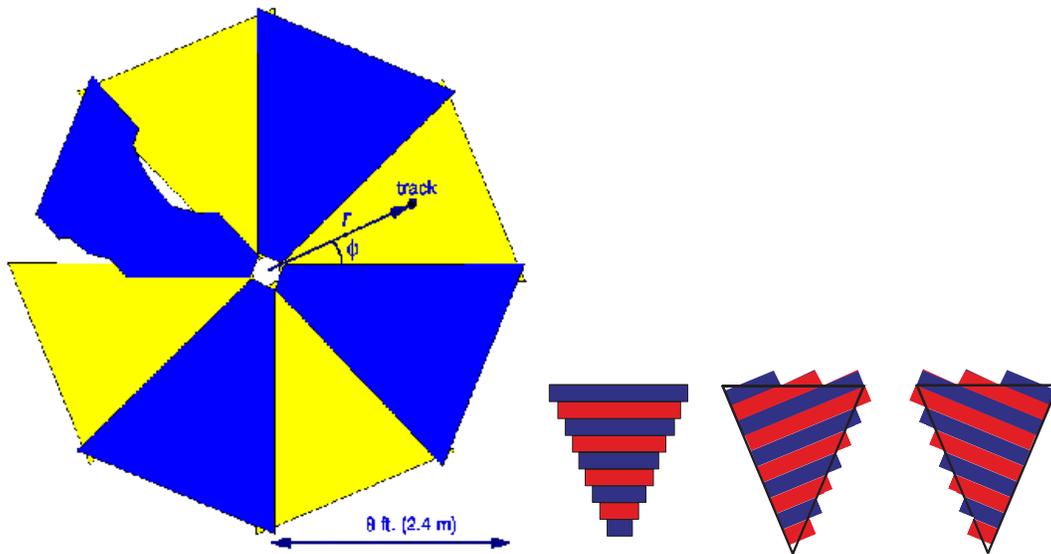


Figure 2.29: Left) Beams-eye view of one muon detector station, which consists of eight overlapping octants arranged in two layers. One octant is cut away in places to show the overlap between adjacent octants. Right) Arrangement of planks to form the four views in an octant (r view is repeated). There will be 12 planks per view (more than shown).

Table 2.7: Parameters of the baseline BTeV muon system.

Radial coverage	38–240 cm
Toroid Z-locations (center)	870, 1010 cm
Average Station Z-location	940, 1080, 1194 cm
Total Length	4 m (includes toroids)
Toroid Length (each)	1 m
Toroidal Fields	1.5 T
Tube cell size	1 cm (diameter)
Wire spacing:	0.5 cm (staggered)
Spatial resolution	1.5 mm
Total channels	36,864 per arm
Momentum resolution	$\sigma_p/p = 19\% \oplus 0.6\% \times p$

2.10 The BTeV trigger system

The trigger system is crucial for the success of BTeV. It finds B events by taking advantage of the main difference between these events and typical hadronic events, i.e. the presence of detached beauty or charm vertices. The trigger detects these vertices by utilizing the superior pattern recognition capabilities of the pixel detector to reconstruct tracks and vertices in the first stage of the trigger, Level 1. This is referred to as the Level 1 vertex trigger, which is the primary trigger for the experiment. In addition to the vertexing capabilities of Level 1, the trigger system includes an independent Level 1 muon trigger that receives data from the muon detector to select J/ψ and prompt muon events. Besides providing interesting physics on its own, the muon trigger is used to calibrate the vertex trigger.

Results from the Level 1 vertex trigger are combined with results from the Level 1 muon trigger in the Global L1 (GL1) trigger, which ultimately selects the beam crossings that pass the first level trigger. Data that survive the selection criteria are assigned to a Level 2/3 processor for Level 2 analysis. Data that survive Level 2 will be analyzed by Level 3 algorithms that decide whether or not the data should be recorded on archival media.

To perform the large number of calculations needed to process and select B events at a rate of 7.6 million beam crossings per second, we require a massively parallel system with several thousand computational elements. These elements include large Field Programmable Gate Arrays (FPGAs), Digital Signal Processors (DSPs), and general-purpose microprocessors. FPGAs are used at the earliest stage of the processing pipeline to perform large numbers of rudimentary calculations that are required for pattern recognition. DSPs offer more programming flexibility than FPGAs, and are used for the Level 1 trigger calculations that entail track and vertex reconstruction. Moreover, the I/O rate capabilities of DSPs are important at Level 1, since we require high bandwidth to get data to the processors. At Levels 2 and 3 the I/O rate requirements are less critical (data rates are lower than at Level 1), and a general-purpose microprocessors will be used for this part of the trigger. The microprocessors provide programming flexibility and significant processing power.

2.10.1 Trigger overview

The trigger system consists of three levels [36]. Each level contributes to the reconstruction of events, and successive levels impose more and more refined selection

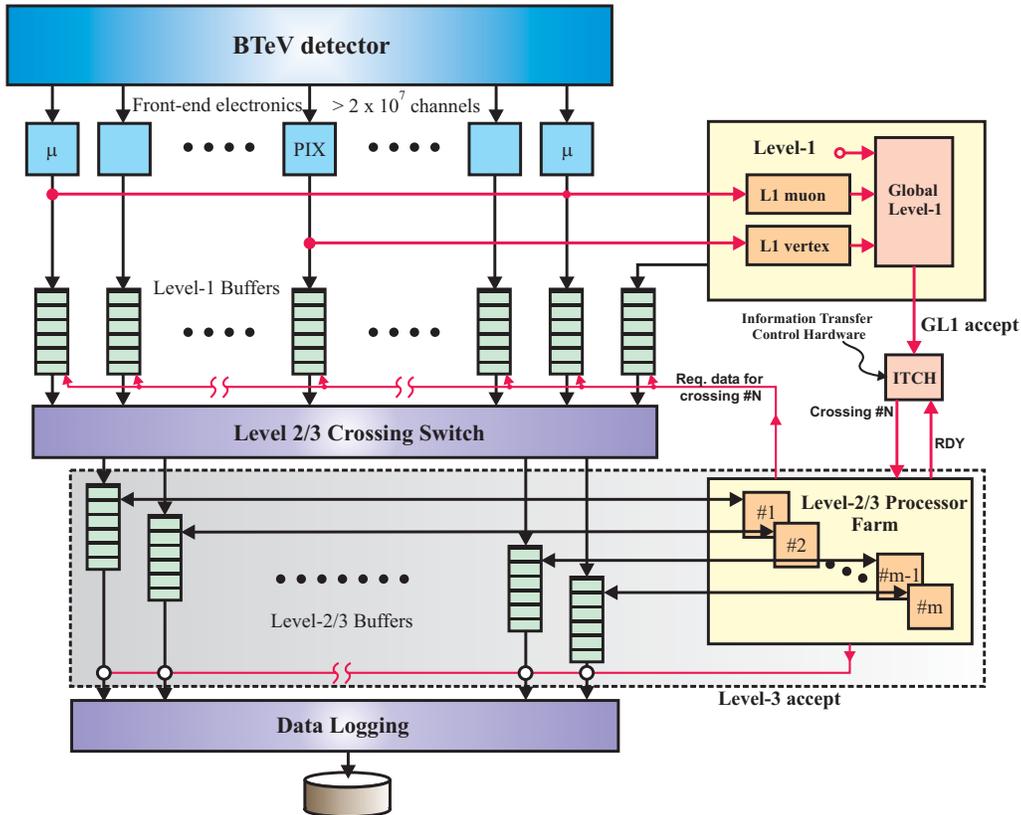


Figure 2.30: BTeV three-level trigger architecture

criteria to select B events and reject light-quark background events. At Level 1 the trigger reduces the beam crossing rate of 7.6 MHz by a factor of 100 while maintaining high efficiency for B decays that can be successfully reconstructed in the spectrometer. The tracks and vertices found at Level 1 are passed to Level 2. At Level 2 we improve the reconstruction of tracks and vertices by reviewing the pixel data used at Level 1, and by including additional pixel hits in the tracks. At Level 3, all of the data for a beam crossing are available and are used to impose the selection criteria for the final trigger decision. The trigger rate is reduced by an additional factor of 20 by Levels 2 and 3. As mentioned previously, BTeV will operate at a luminosity of $2 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$, corresponding to an average of two interactions per beam crossing at a crossing rate of 7.6 MHz. Average event sizes will be ~ 100 KB after zero-suppression of data is performed by front-end detector electronics. Since every beam crossing will be processed, this imposes an extremely high data rate of ~ 800 GB/sec on the experiment. BTeV will employ the three-level hierarchical trigger architecture shown in Fig. 2.30, to handle this high rate. Level 1 will reject

97.5% of all incoming events, thereby reducing the data rate by a factor of ~ 100 . Levels 2 will further reduce the data rate by a factor of ~ 10 . while Level 3 will reduce the number of accepted crossings by at least an additional factor of ~ 2 .

2.10.2 Level 1 vertex trigger algorithm

The first phase of the Level 1 vertex trigger algorithm is the pattern recognition that uses pixel hits to find tracks. This is also referred to as track-segment finding [37]. This phase of the algorithm starts by finding the beginning and ending segments of tracks in two separate regions of the pixel planes, an inner region close to the beam axis and an outer region close to the edge of the pixel planes. The search for the beginning and ending segments of tracks is restricted to these inner and outer regions, respectively. Segments are found using hit clusters from three adjacent pixel stations in the defined regions. Inner segments are required to point back to the beam axis while outer segments are required to project outside pixel plane boundaries. Once these segments are found, they are then matched to form complete tracks in the segment matching stage.

After complete tracks are found, the track and vertex reconstruction phase of the trigger performs calculations to determine the momentum of each track and calculate its transverse distance from the beam axis. Primary vertices are found by looping through all tracks with transverse momenta $p_T \leq 1.2 \text{ GeV}/c$ that appear to originate close to the beam line. Remaining tracks are then tested for their detachment from the primary vertices that were found. The Level 1 vertex trigger selects events if there are at least n tracks in the same arm of the BTeV detector satisfying the following criteria: $p_T^2 \geq 0.25 (\text{GeV}/c)^2$, $b \geq m\sigma$, and $b \leq 2 \text{ mm}$, where b is the impact parameter and n and m are tuned to achieve the desired rejection of minimum-bias events. The Level 1 vertex trigger is able to reject 97.5% of all minimum-bias events while accepting $\sim 60\text{-}70\%$ of the B events that would survive our offline analysis cuts.

2.10.3 Level 1 vertex trigger hardware

A block diagram of the Level 1 vertex trigger is shown in Fig. 2.31. Data from all 30 stations of the pixel detector are sent to FPGA-based pixel processors that group individual pixel hits into clusters. Hit clusters from three neighboring pixel stations are routed to FPGA hardware that finds beginning and ending segments of tracks in the pattern recognition phase of the trigger. Track segments found at this stage

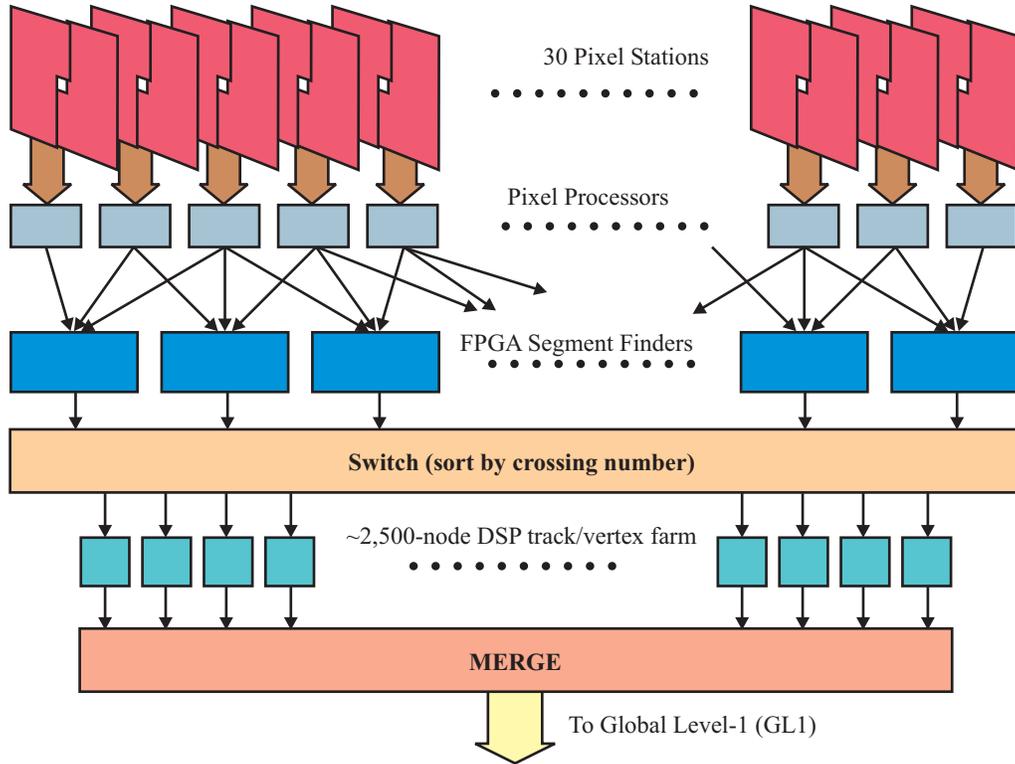


Figure 2.31: BTeV Level-1 vertex trigger.

are sorted by a switch according to their beam crossing number (BCO), and routed to a DSP in the track/vertex farm. This DSP performs segment matching, as well as track and vertex reconstruction.

2.10.4 Levels 2/3

The Level 2 algorithm refines the tracks found at Level 1 by adding pixel clusters from the planes located between the “inner” and “outer” track segments. It then performs a Kalman-filter track fit, and improves the momentum resolution to about 5–10%. One of two requirements must be satisfied to select an event. A secondary vertex must be present, or the collection of detached tracks must satisfy a minimum p_T cut. The result is a joint light-quark rejection of 1000–1 per beam crossing for Levels 1 and 2 combined and $\sim 50\%$ overall efficiency for most B decays of interest. The Level 2/3 trigger is implemented as a farm of commercial processors.

Chapter 3

The DAQ for the pixel test-beam at Fermilab

3.1 Introduction

BTeV has been designed to make a complete set of measurements on the decays of hadrons containing b and c quarks. The pixel detector is crucial for the BTeV program because it is used to trigger heavy quark events by exploiting their peculiar characteristic: a long mean life and thus a long decay length. With its extreme precision and speed it is possible to reconstruct primary and secondary decay vertexes at the first level trigger and thus provide a very high trigger efficiency. The study of the pixel detector is fundamental for the BTeV experiment, therefore test-beams are necessary to test its functionality.

The first test-beam for the pixels was done in 1999 to study the characteristics of the sensor bonded to the first prototype of readout chip, FPIX0, developed at Fermilab. The spatial resolution of the detector has been measured and the results were published [38] and are shown in Fig. 2.10. A new test-beam will be done this year and the final sensors and readout chips, FPIX2, which will be used in BTeV, will be tested.

This time the test-beam purposes include a more comprehensive set of studies. The main goals are, besides the studies of the spatial resolution of the final pixel system, also the study of the readout chip in the real BTeV working condition. There will be no external trigger and the event builder will use the time-stamp information provided by the chip to assemble the events.

To allow this study a completely new DAQ system has been developed and this

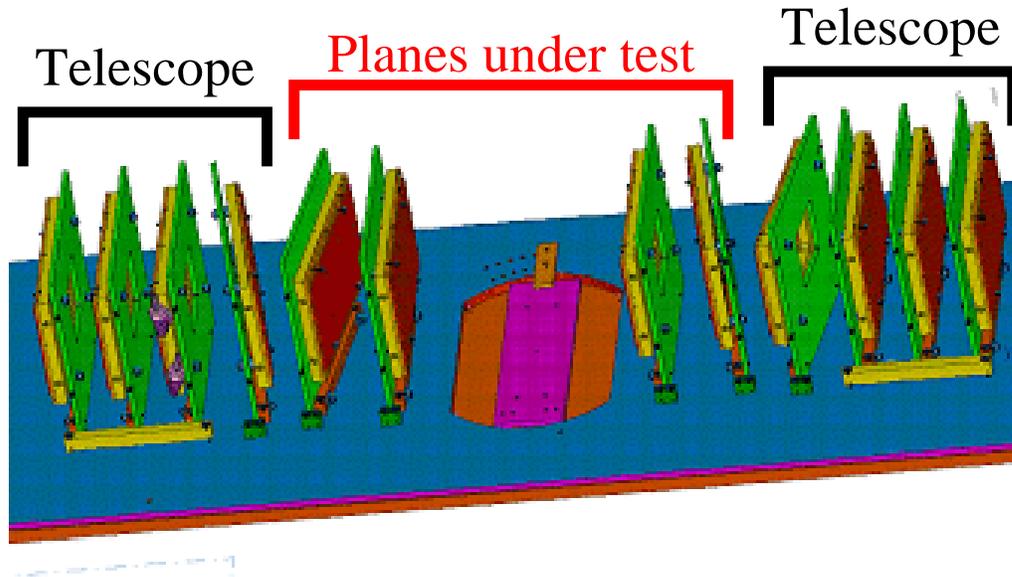


Figure 3.1: Schematic representation of the pixel planes installed in the thermo-controlled hut. The telescope is composed by FPIX1 detectors placed in the first and last four slots. The devices under test (FPIX0, preFPIX2, FPIX2) are positioned in the four central slots and they can be positioned at various angles for spatial resolution studies.

work has been my main task in the last years and is the principal argument of this thesis.

3.2 Experimental setup

The beam test will be carried out at Fermilab in a 120 GeV/c pion beam. The pixel detectors to be tested are hybrid assemblies of several combinations of pixel readout chip prototypes developed at Fermilab, and single chip sensor prototypes. Data will be collected at the MTest beam line located in the Meson Area at Fermilab. Figure 3.1 shows the experimental set-up. The pixel devices under test are located between two stations of four pixel detector planes each that provide tracking information in the x direction, corresponding to the small pixel dimension ($50\mu\text{m}$). The pixel hybrid devices are mounted on printed circuits boards, held inside an aluminum box, where their location is determined by precision machined slots. One of the pixel devices in the telescope and the four devices under test can be positioned at various angles with respect to the beam direction. This allows to measure the

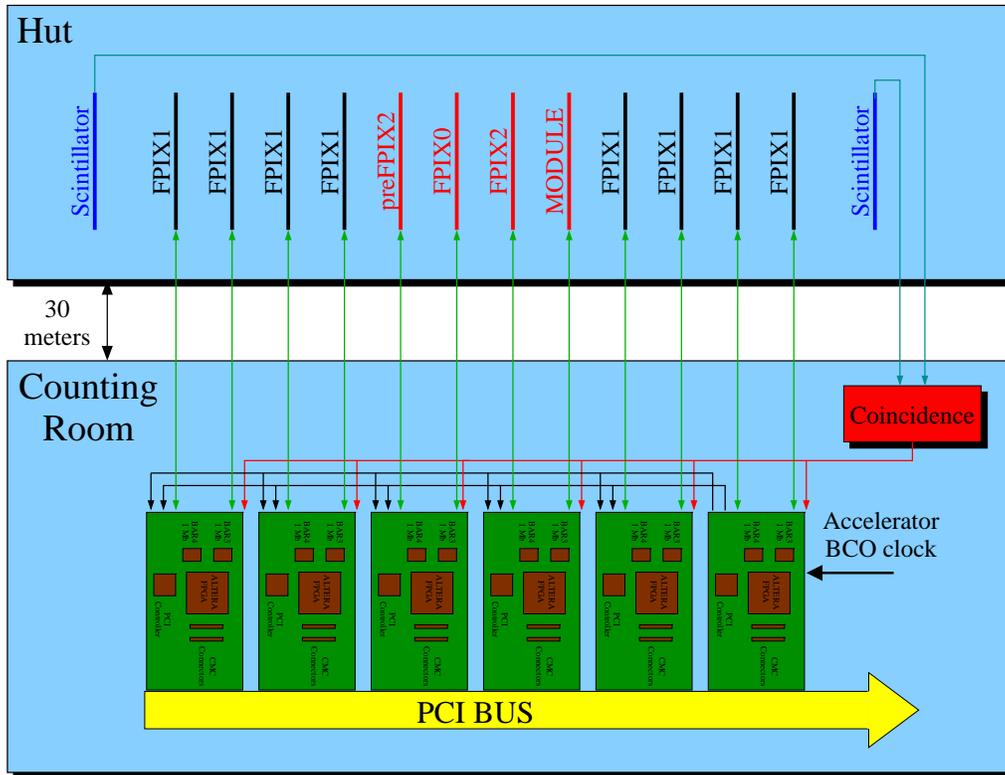


Figure 3.2: Test-beam electronics schematic. The pixel planes inside the hut are connected through 30 meters cables at the PTA/PMC cards in the PCI expander located in the counting room. The twelve planes in the system are connected to the six mated cards, two planes for each card. It is not planned to use the two scintillation counters since the readout chips are data driven.

properties of various pixel prototypes as a function of the pion incident angle. Each telescope plane will consist of an FPIX1 chip bump bonded to a pixel sensor array and mounted on an inner board. The planes under test will consist of an FPIX0 or preFPIX2 or FPIX2 chip bump bonded to a sensor array and mounted on an inner board. One scintillation counter upstream and one scintillation counter downstream will serve as a minimum coincidence trigger configuration, even if it's not planned to use them since the readout chips are data-driven.

The DAQ is no more based on a custom made readout system such as CAMAC, but it's implemented with the common PCI protocol used in standard PC technology, see Fig. 3.2. This allows a high flexibility of the system and a relatively low cost. The PCI card used for the DAQ, PTA (PCI Test Adapter) see Fig. 3.3, has been developed at Fermilab and can send and receive signals from external devices like



Figure 3.3: PTA card. On the left are located the two memory banks of 1 Mbyte each. The Altera FPGA is positioned in the middle. On the right there are the two connectors where the PMC card is attached, through which signals are exchanged between the two mated cards. The component on the bottom is the PCI controller.

the pixel chips under test. The flexibility of this card is given by a programmable device, an FPGA, that allows the user to implement a logic circuit designed for specific purposes. Because of the different pixel planes under test, and the possibility to connect other kind of detectors, another card, PMC (Programmable Mezzanine Card) see Fig. 3.4, has been designed at Fermilab. The task of the PCI card, in fact, is to receive data with a simple protocol, and act as balancing buffer between the production data rate and the PC readout, while the PMC firmware is designed specifically for the kind of detector under test. A PTA card with a mated PMC card will serve to control and read out two planes and are used to control and receive data from the telescope planes and the ones under test. A different PMC firmware has been developed for each type of plane. Firmware in the PTA cards will be common across the entire system. A single PTA/PMC can control up to two planes.

The PTA card has two banks of 1 MByte memory for a total of 2 MBytes. Event data from the pixel planes are stored in the PTA card memory until read out by the PC via the PCI bus. These two memory banks are crucial for the design of the DAQ because they are used both to balance the different acquisition rates of the PCI card and the host PC, and to synchronize the different cards present in the

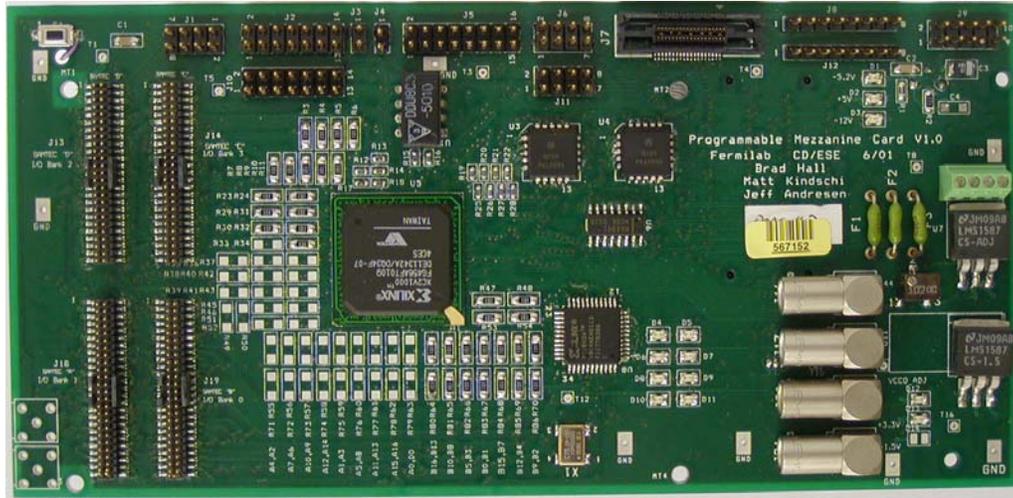


Figure 3.4: PMC card. On the left there are the four connectors used to talk with the pixel planes, two connectors for each plane. In the middle is located the Xilinx FPGA which is responsible of the chip readout control.

expander as we will see in the next chapter.

The general scheme of the DAQ has been implemented having in mind a continuous data flow produced by the pixel planes. The idea is to simulate the real condition of operation of the pixels in the experiment when data are collected and used for the trigger as soon as they are available, in other words the DAQ must be able to build the events in absence of any external trigger which can alert of the presence of a particle interacting with the pixel system.

In order to build an event, the readout chips provide a time-stamp which is a unique mark of the time when a particle interacted with a pixel cell. This time-stamp is an internal counter in the readout chip which is incremented along with the radio frequency of the accelerator which determines the time distribution of the particles in the beam. This counter, the BCO, is then synchronized with the particles bunches in the beam, and this allows to associate a unique time-stamp every time a particle interacts with sensors.

In absence of an external trigger and in a continuous readout mode, hits can pertain to pixel planes at different time intervals, but the presence of a time marker, the time-stamp, relates the data to the precise instant when the particle passed through the detector, thus allowing an event building ¹.

¹An event is the collection of all hits with the same time-stamp. An event-builder is a process able to collect hits together into events

3.3 Main characteristics of an FPGA

Both cards, PTA and PMC, are provided with an FPGA that allows users to implement their own code. This feature gives those cards a high flexibility because it is possible to reprogram their functionalities following the requirements of the system. An FPGA is an integrated circuit that contains many (64 to over 10,000) identical logic cells that can be viewed as standard components. Each logic cell can independently take on any one of a limited set of personalities. The individual cells are interconnected by a matrix of wires and programmable switches. A user's design is implemented by specifying the simple logic function for each cell and selectively closing the switches in the interconnect matrix. Complex designs are created by combining these basic blocks to create the desired circuit. Field Programmable means that the FPGA's function is defined by a user's program rather than by the manufacturer of the device. Depending on the particular device, the program is either 'burned-in' permanently or semi-permanently as part of a board assembly process, or is loaded from an external memory each time the device is powered up. The FPGA has three major configurable elements: configurable logic blocks (CLBs), input/output blocks, and interconnects. The CLBs provide the functional elements for constructing user's logic. The IOBs provide the interface between the package pins and internal signal lines. The programmable interconnect resources provide routing paths to connect the inputs and outputs of the CLBs and IOBs onto the appropriate networks.

3.4 Mezzanine card

The mezzanine cards serve to control and read out the pixel planes. The main duties of the PMC are:

- providing, to the pixel planes, the readout clock and BCO clock, at the exact frequency of operation in BTeV.
- managing the different readout chips
- synchronizing the BCO in the pixel planes and between the PMCs
- extending the pixel chip BCOs
- sending the data to the PCI card

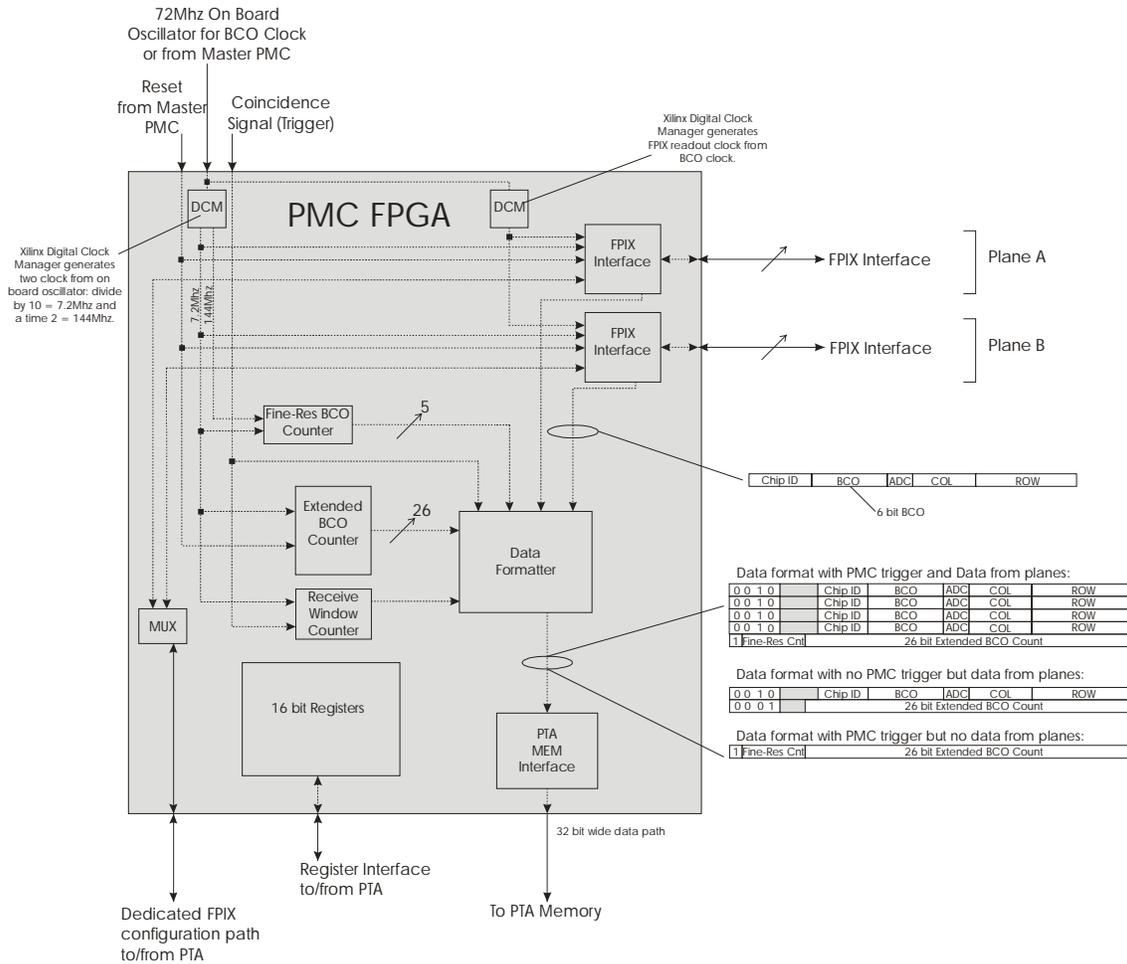


Figure 3.5: PMC electronic schematic.

The readout clock and especially the BCO clock in the pixels are controlled and distributed by the FPGA in the mezzanine card.

It is fundamental that each of the planes in the telescope and planes under test be synchronized and phase aligned with the same BCO clock and BCO count, because when a particle passes through all the detectors the generated hits possess the same time-stamp which is later needed by the event builder. One of the PMCs has been designated the Master PMC. The Master PMC is responsible for distributing the BCO clock and reset signal to all PMCs. The **master accelerator clock** will be available in the counting room for the distribution to each PMC. The master BCO clock is sent to the Master PMC which uses its Digital Clock Manager (DCM) to phase shift the BCO clock with $PERIOD_{BCOCLK}/256$ resolution ($520ps$ with $132ns$ BCO clock). The magnitude of the phase shift is set via a register and is used to

tune the phase of the BCO clock at the hut relative to the arrival of the beam. The Master PMC distributes the BCO clock to the other PMCs in the PCI bus via its Low Skew Clock Distribute resource (see Fig. 3.2). The distributed BCO clock passes through each FPGA on the PMC and is driven via the data cable to each plane in the hut (see Fig. 3.5). With matched cable lengths each plane in the hut should receive phase-aligned BCO clocks.

The BCO counters in the pixel chips have from 6 to 8 bits and with a BCO clock period of $132ns$ there is only $132 \times 2^8 \approx 34\mu s$ of unique time-stamps available. In the DAQ scheme, data are first collected in the two bank of memories on the PTA cards and then read by the PC for the event building. This incredible small time between a wrap around of the BCO counter wouldn't allow an effective event building because there would be for sure hits belonging to different events with the same time-stamp in the same memory buffer before being flushed by the software readout sequence.

In order to allow for a reasonable amount of time to read the data from the memories on the PCI card, before an already used time-stamp value can reappear in them, the Mezzanine card provides a 26 bits Extended BCO counter which is clocked, of course, by the BCO clock. With a 26 bit Extended BCO counter, there are 8.8 seconds of unique Extended BCO counter states, and this time is sufficient to allow an event builder with no ambiguities.

3.5 PCI card

The PCI card is the interface between the Mezzanine, which collects data from the pixel planes, and the PC where the data are built and written to disk for the offline analysis. During my thesis I was in charge of the design of the firmware of the PCI card's FPGA. The design of the firmware is strictly connected with the design of the software which reads the data from the local memory banks on the PCI card and uses the time-stamp information of the hits to build the events. The firmware tasks are to receive data from the mezzanine card and store them in one of the two available memory bank while the PC is reading from the other one. When a memory limit or a timeout, set by the user, are reached the FPGA first swaps automatically the data flow to the other memory bank and then sends an interrupt to the PC which starts reading the data stored in the filled memory. This mechanism allows a continuous read out because while the PC is reading one memory bank, data are flowing in the other bank without stopping the pixel read out.

The software that I used to program the FPGA is the Quartus II software which allows to design logical circuits, to manage the clocks and data flows in the FPGA. The blocks shown in Fig. 3.6 and Fig. 3.7 represent the interfaces between the FPGA and the other components of the PTA and have all been designed with the Quartus II software.

I'll describe in the following sections the data flow inside the FPGA and the firmware implementation that allows the swapping mechanism and the interrupt generation.

3.5.1 The CmcConnectors block

This block, see Fig. 3.8 and Fig. 3.9, is the interface between the Mezzanine and the PTA. Data coming from the mezzanine through a 32 bit connector are stored in a dual clock FIFO. In fact, the clock running inside the FPGA is a 33 MHz clock synchronized with the PCI bus clock, while the clock synchronized with the data, readout clock, is provided by the mezzanine and can run up to 20 MHz. There is therefore a mismatch between the readout clock and the clock used to write and read to the memories and for all the logic elements implemented in the FPGA. The easiest solution is to use a dual clock FIFO: data are buffered at the readout clock frequency and then are sent to the memories at the FPGA and memories frequency. Another clock frequency problem is due to the arbitration between the data coming from the two pixel planes connected at the same mezzanine card. As I previously mentioned, any mezzanine can control up to two different planes. In order to avoid loss of any data the PMC must double the data transfer frequency to the PTA. The pixel planes have a readout clock which is 20 MHz but data are transferred from the mezzanine to the PTA at 40 MHz. This prevents any possible loss of data from the pixel planes to the FIFO, but not from the latter to the memories because the FIFO receives data at a frequency higher than the one at which is flushed out. Even if there is the possibility of losing data, in case the FIFO is completely filled out, this situation is very unlikely to happen because the expected data rate will not be bigger than 100 KHz and even if there will be bunches of data at a very high rate the FIFO is able to store up to 2048 data, enough to avoid any loss. Nevertheless I provided a counter that can tell at least how many data are lost.

Another important feature of this block is the possibility to generate a sequence of number through an internal counter which can simulate the data flow from the mezzanine. During the software development phase, this counter was crucial to test

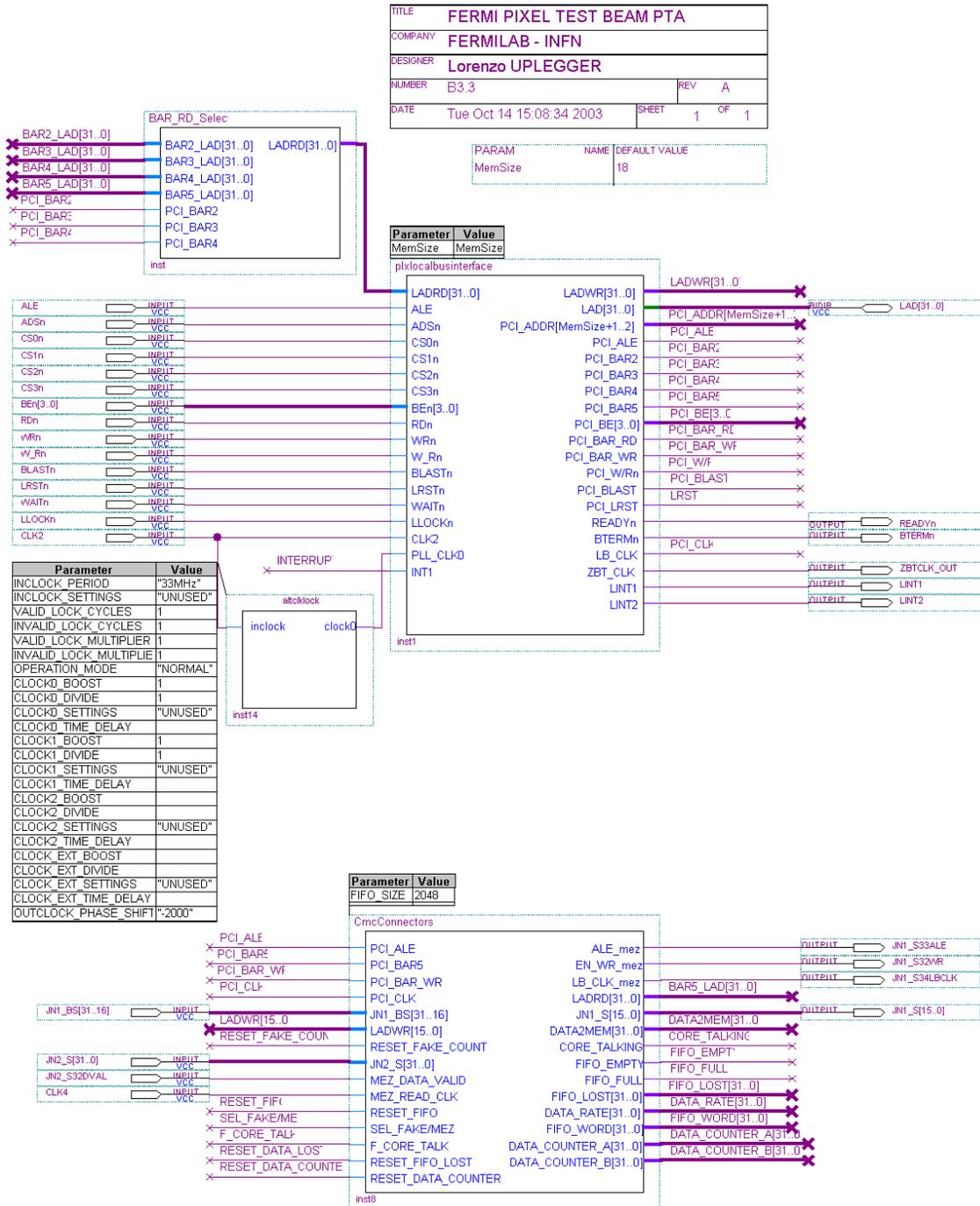


Figure 3.6: PTA firmware schematic part I.

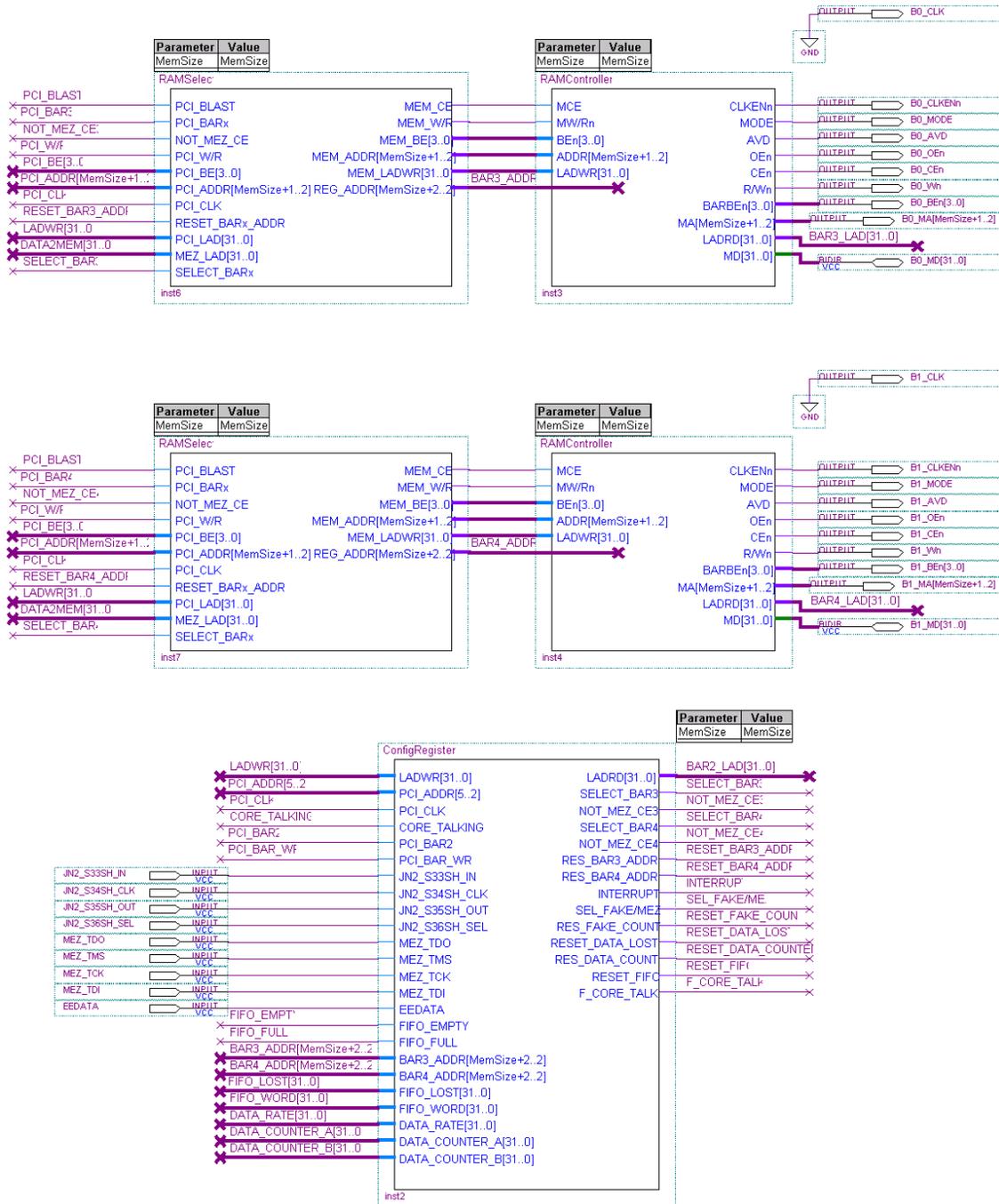


Figure 3.7: PTA firmware schematic part II.

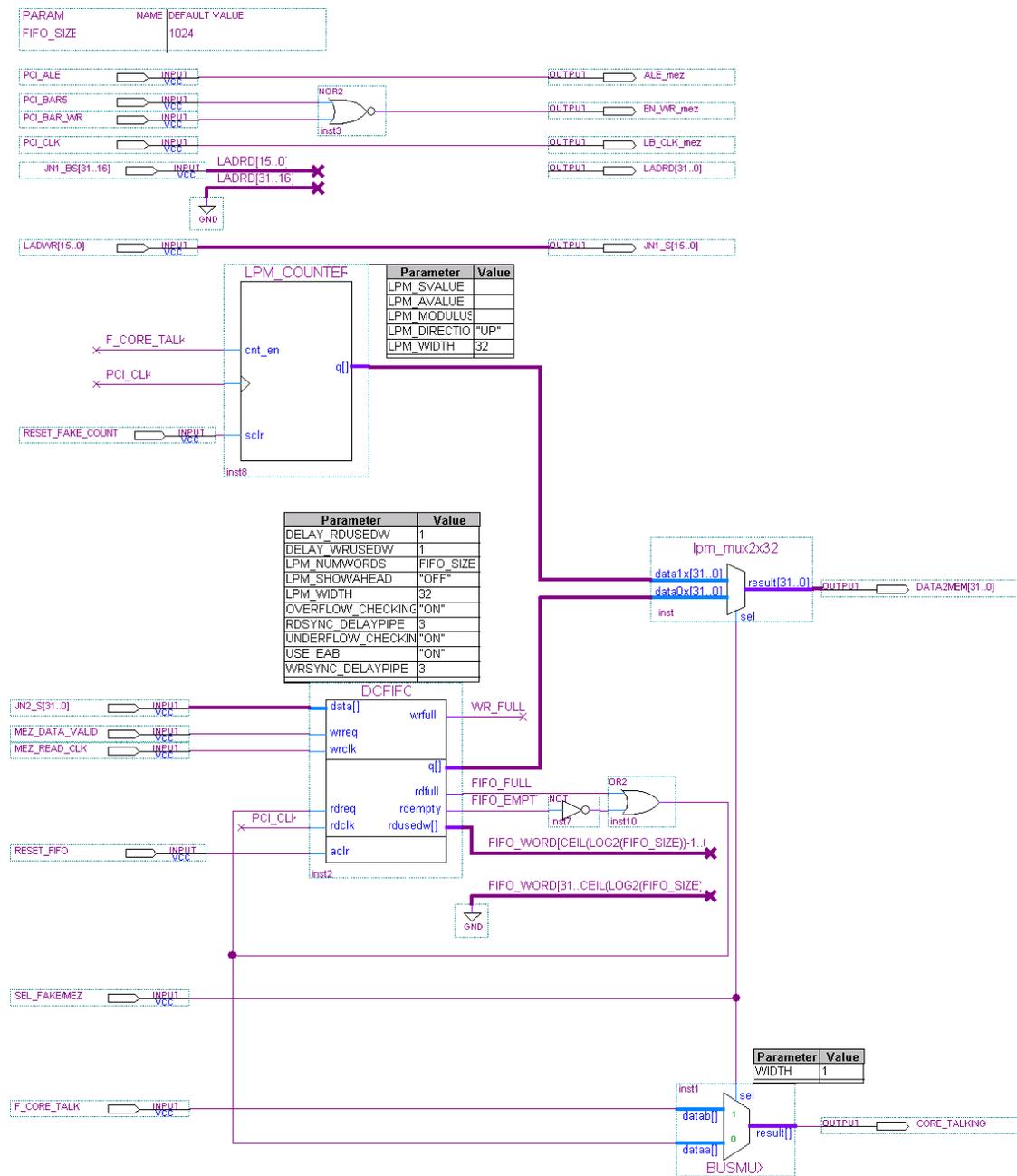


Figure 3.8: Cmconnectors schematic part I.

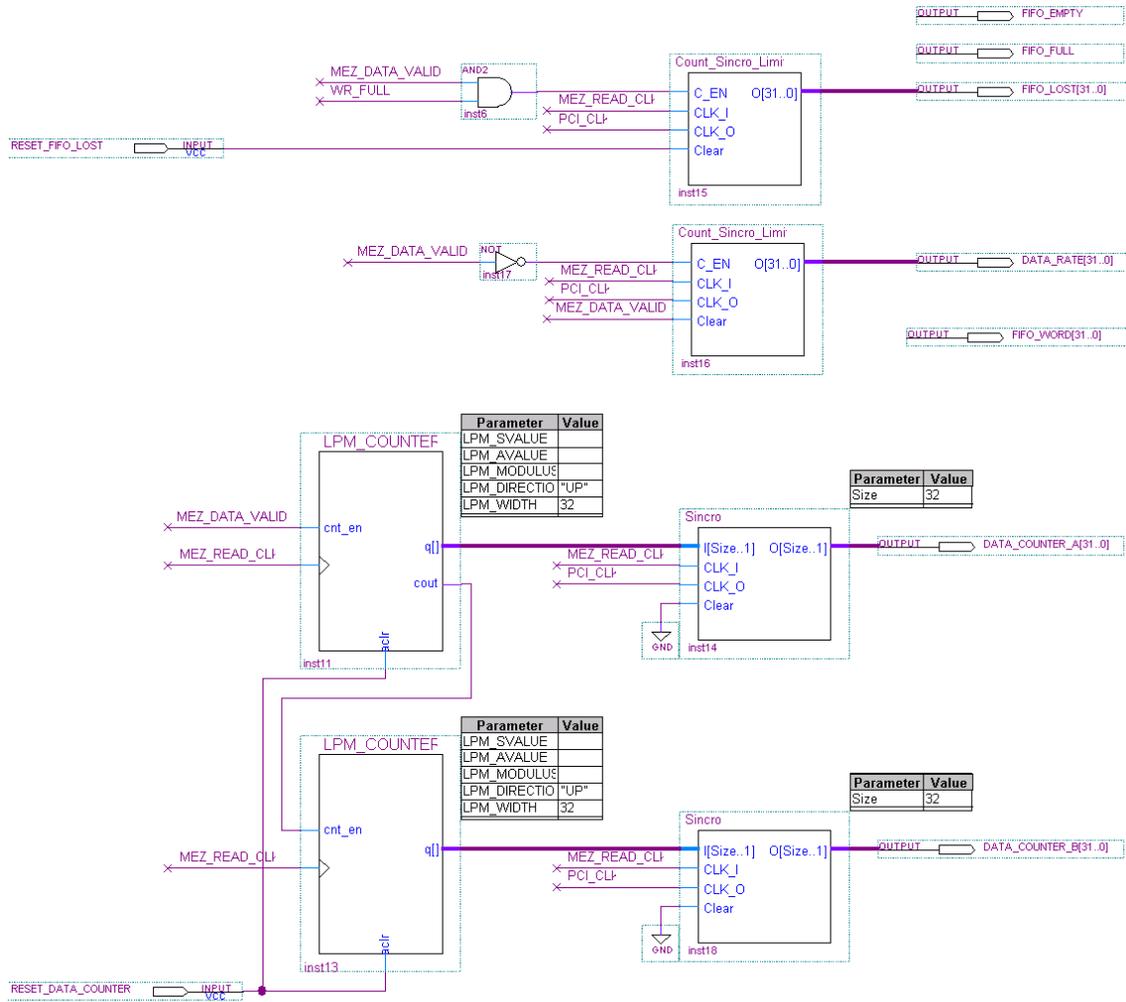


Figure 3.9: Cmconnectors schematic part II.

all the software components without the necessity to connect any detector. This fake data generator provides a sequence of numbers which follows, in the FPGA, the same path of the pixel data from the mezzanine, thus simulating completely a real data acquisition, and can be interpreted as a sequence of time-stamp by a suitable event builder that we implemented and called “fake builder”.

Data from the FIFO go through a connector which is connected to both memory banks. The distribution of the data to one memory bank rather than the other one is managed by the ConfigRegister block.

3.5.2 The ConfigRegisters block

The schematic of the block is shown in Fig. 3.10 and Fig. 3.11. This is the core of the FPGA implemented logic. This block provides the informations to the software through its registers which act also as control logic for the memories swapping.

As described before, the FIFO data bus is connected to both memory banks but only one of them is being filled at any given time. The RAMDirection block manages the data flow. In fact the user can set by software a memory limit. These values, one for each bank, are stored in the RAMLimit blocks and are used by the RAMDirection block to automatically swap the data flow from one bank to the other.

When a memory bank is full, i.e. it reaches the limit set by software, data are immediately redirected to the other memory with no loss of data. This logic is implemented in the RAMDirection block, and even with a continuous data flow all data are saved in either one of the two memories.

Once the data flow has been redirected to the empty memory, an interrupt is send using the INT[0] or INT[1] lines, one line for each memory. When one of those two lines is on, the Interrupt block first sets a register, allowing users to check which of the two memory bank is eventually full, and then set the interrupt line which physically sends the interrupt to the PC. The software receives the interrupt and then can check which memory bank needs to be flushed.

There are three mechanisms that can impose a memory swap:

1. the reaching of the limit set by users in the memory
2. a user-set timeout
3. a software command of swapping

All these mechanisms are implemented with no data loss and the first of them which is satisfied impose the toggle.

The first mechanism has been implemented to limit the number of data stored in the memories before being read by the PC and also to put a threshold reasonably far away the physical end of the memory.

In fact the real bottleneck of the whole readout chain is the data transfer from the PTA to the PC. Since the goal is to be able to read as many events as possible without losses, we provided a fine tuning of the size of the data to transfer, by selecting this memory threshold according to a suitable calibration of the incoming data flux and the readout speed. Fixing the threshold value to a size smaller than the maximum memory available, translates into smaller amounts of data to transfer. By a fine tune of this threshold the best compromise will be chosen to guarantee a well balanced data flux. Another benefit of this approach is related to the possibility of having a large event overflow the memory threshold. Leaving enough physical room above the threshold guarantees that anomalously large events (beam spikes) can be stored without losses.

As mentioned before, the swapping can only occur if the other bank of memory is empty. To provide the FPGA of this feature I provided a register which stores the number of data present in the two memory banks (BAR3_ADDR, BAR4_ADDR). When the software has read all the data, it resets the address counter of the emptied bank and thus allows the firmware to eventually swap again the data flow from the mezzanine.

Since the time-stamp counter has only 26 bits, the time-stamp values are unique only for ≈ 8.8 sec. The time-out induced swap has been implemented because it is extremely important that hits are stored in the memories with a unique time-stamp value for each event. This means that data must be read by the PC before there is a wrap-around of this counter, otherwise there can be hits belonging to different events in the same memory buffer, thus creating ambiguities in the event reconstruction. If the data rate is low and the memory limit is not reached before 8.8 sec the memory buffer must be flushed in any case.

The third swapping mechanism has been implemented to allow, with a simple mechanism, the event building. As I'll explain in the next chapter, we needed a mechanism to synchronize the memory swapping of all the PCI cards. The user is then able to order a toggle of the memories at any given time without waiting for one of the first two conditions, size and time limit, to occur.

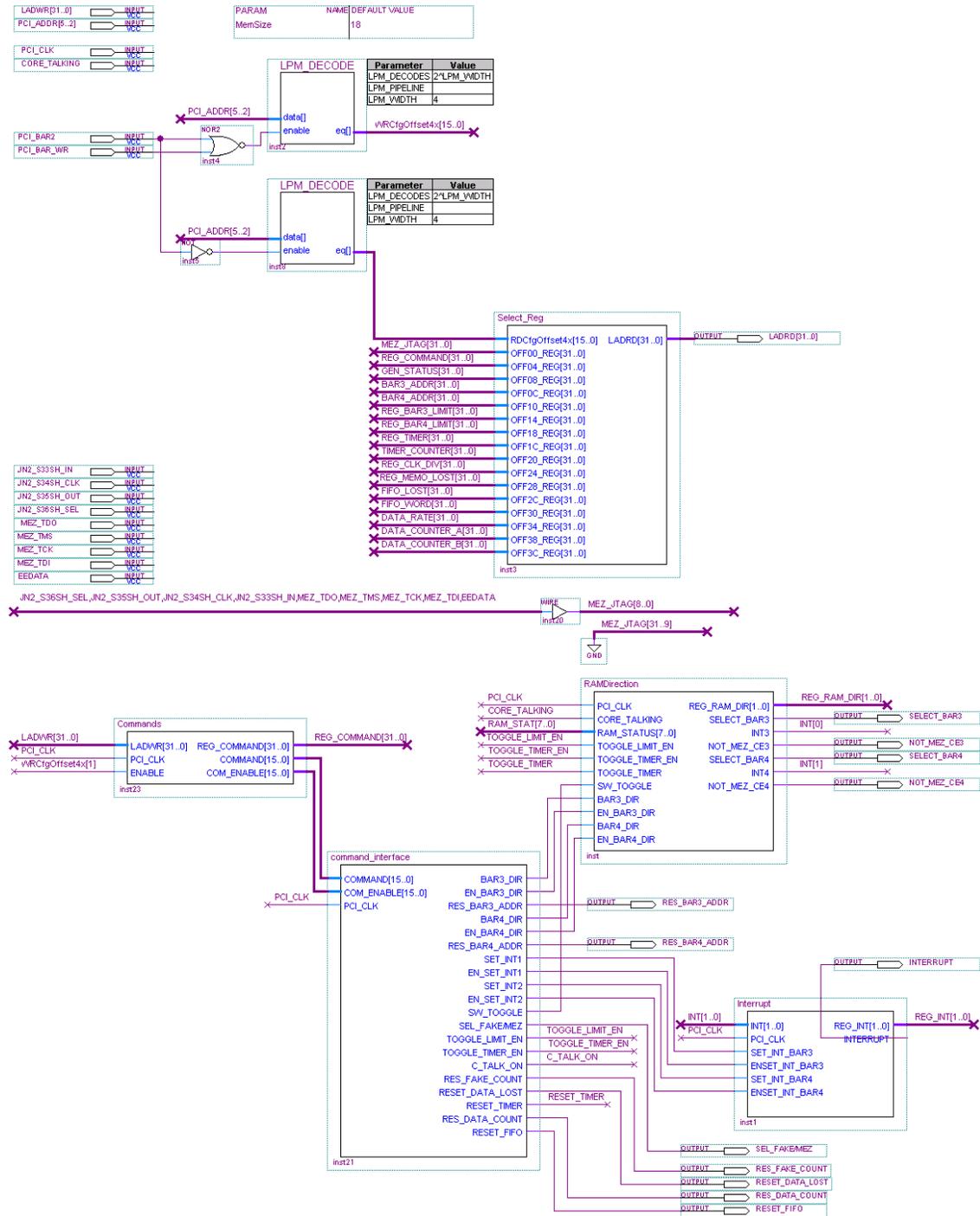


Figure 3.10: ConfigRegister schematic part I.

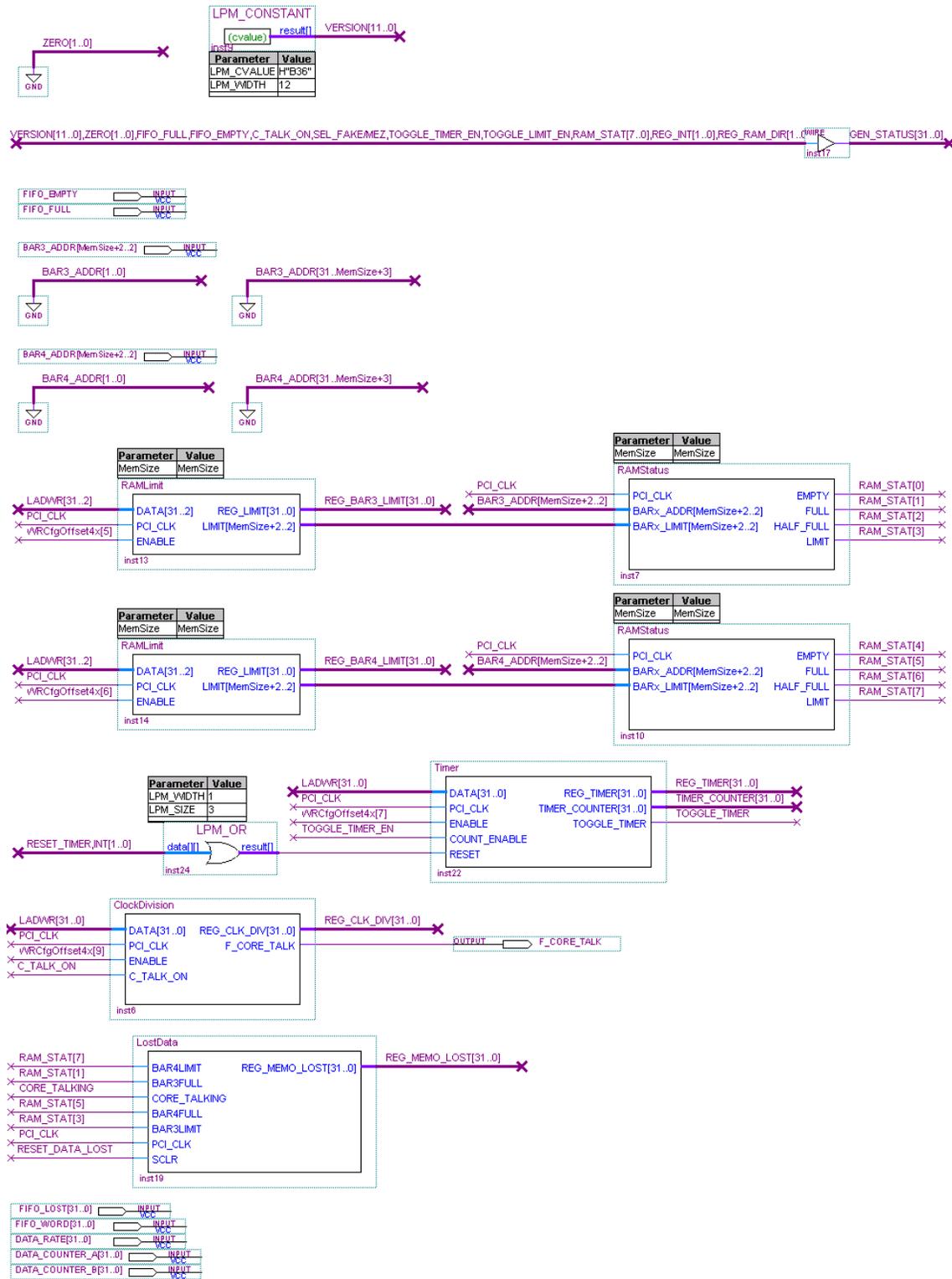


Figure 3.11: ConfigRegister schematic part II.

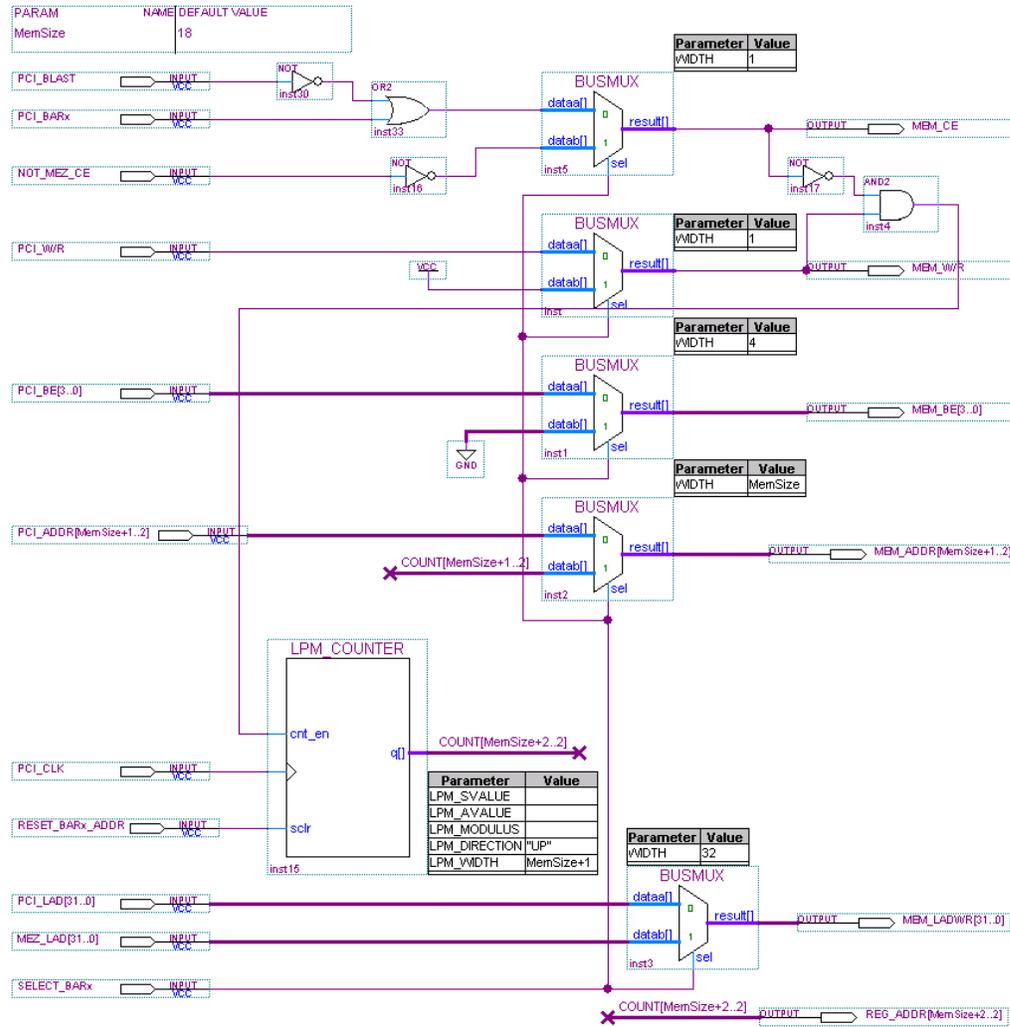


Figure 3.12: RAMSelect schematic.

3.5.3 The RAMSelect block

Data from the mezzanine card are sent alternatively to the memories on the PTA Card. The RAMSelect block manages all the signals needed for read and write operations in the memories and provide also a counter to be used for addressing the data that go in and out the memory. This counter is incremented every time a new data comes from the mezzanine and it is used to check if the memory limit has been reached.

As we can see from Fig 3.12 the multiplexer in front of all the signals that go to the memory, allows data to flow either from the mezzanine or from the PC. The MEM_LADWR is the 32 bits bus forwarding the data to the memory. The

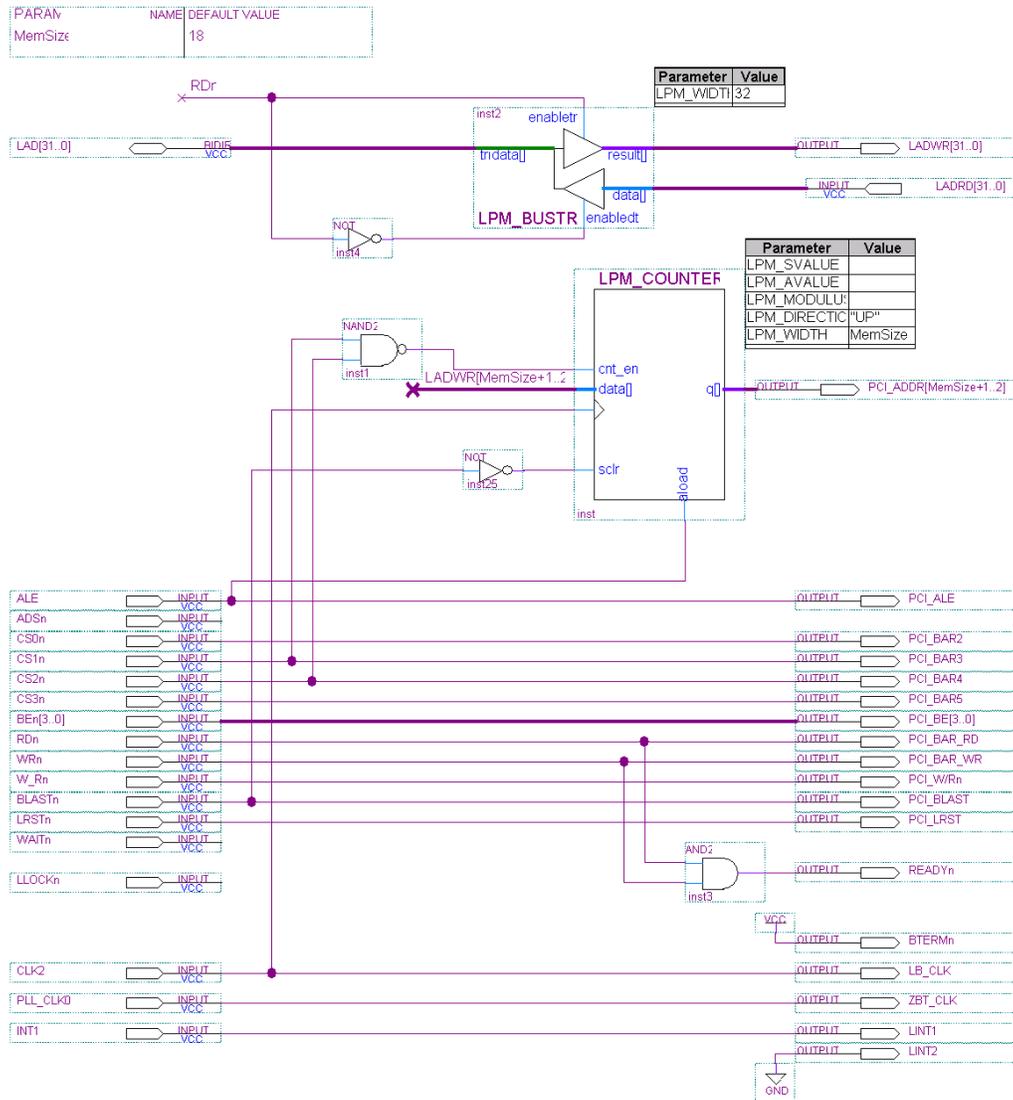


Figure 3.13: Pxllocalbusinterface schematic.

multiplexer allows to read or write from only one device at a time, this means that while the mezzanine is writing, MEZ_LAD selected, the PC cannot access to the memory and vice-versa when the PC is reading or writing, PCLLAD selected, the mezzanine cannot write into the memory.

3.5.4 The pxllocalbusinterface block

The last schematic block (see Fig. 3.13) is the interface between the FPGA and the PCI controller on the PTA. Once the data are stored in the one of the memories

and a swap has occurred, an interrupt is sent to the PC, and the software starts the procedure to flush the filled memory. A software command acts on the FPGA through the signals that enter this block. All the signals in the `plxlocalbusinterface` are used to manage the read and write procedures in the memories and all the logic inside the FPGA. Data from the memories are sent to the PCI bus through the LAD bidirectional bus.

3.6 Test beam DAQ

In the first part of this chapter I described the data flow through the PCI card. To make the project complete, an important part of my thesis has been the design, implementation and testing of a complete DAQ system that takes advantages of the components described before. The DAQ software has been designed to be flexible and easy maintainable in order to have a system which can be used both for testing the pixels in a test bench and at a beam test.

Our goals were to provide a highly modular code, able to accommodate different detectors, with different hardware and software specifications, and robust, in principle able to withstand change of operating system environment, extensive refurbishing and additions of algorithms.

To allow for robustness and modularity the object oriented paradigm has been chosen. Particular care has been taken in the design phase in order to efficiently achieve an optimal decoupling of all components. Furthermore the object oriented modeling allows to implement a software description of the system that mimics the effective hardware hierarchies and relationships; it is possible to define abstract objects functionalities that can be applied to several real object instances.

During the design phase we exploited the basic concepts of object orientation:

- **Classification.** It is the grouping of objects that have a common set of properties and operations. A classification defines a type of object by the set of properties and operations. For example, PTA card objects are defined by the fact that all the objects have properties like “memory bank” and “register” and operation like “read” and “write”. Classification is so useful because it is a means of abstraction, i.e. considering objects at a certain level of detail and ignoring lower level of complexity.
- **Specialization.** It is the identification of a subset of objects within a classification or type that have an additional common set of properties and operations.

A specialization defines a subtype of the original type since objects of the specialization contain the original properties and operations. For example, our event builder collects hits into events using the time-stamp information and we provided different specializations implementing a “fake builder”, which collects “fake data” into events (see Par. 3.5.1) and the “real data builder” which collects hits from the real detectors into events. Those two objects form a subtype of the builder object since they contain all the properties and operations of the builder object, but they also have additional specific properties and operations.

- Polymorphism. It’s the ability of different types of objects to realize the same properties and operations in different ways. Polymorphism is a natural result of the fact that objects of different types may instantiate the same properties and operations in different ways. For example, objects of type detector have the operation “initialize” although there are different ways of initializing different types of detectors. Polymorphism allowed us to define a “virtual detector” and then implement the properties and operations, like “initialization”, in a different way, specific for each detector under test, FPIX0, FPIX1, FPIX2 or any other detector with the same properties and operations.

All the code has then been implemented in C++, an object oriented programming language which possesses all the features described above.

The general DAQ scheme has been designed and implemented around four independent processes:

- Producer
controls the PCI cards and it is responsible for their synchronization. Its main task is to read the data from the cards and store them on the host PC, furthermore it is responsible of the pixel chips initialization.
- Consumer
is used to “consume” the data, and its main task is to read the data previously stored by the Producer and do the event building.
- Logger
handles the error and status messages produced by the first two main processes.
- Controller
sends users commands to the processes to drive the DAQ.

These four processes are completely decoupled and they can share data by means of specialized shared memories (chunks of physical memory in the PC shared by different processes) and communicate through message queues. Figure 3.14 is a schematic representation of the overall architecture. As it is shown the producer is the interface between the PCI cards where the data are first stored in their two memory banks and the PC shared memory where data from all the cards are then stored ready to be used. The consumer, instead, doesn't know anything about how data are generated and collected: it uses data previously stored in the shared memory to build events and to store them on disk. The logger process, instead, collects all the messages from the other processes in a file which can be used by the user to handle all the error situations and to eventually acknowledge the user of the DAQ status. Finally, the controller is needed to send the user's commands to drive the DAQ.

3.6.1 Producer

This process is the only software component dependent upon the characteristics of the hardware system used for the data taking through a particular implementation of a virtual class. Its task is to coordinate the initialization of all the pixel planes under test and the readout sequence of all the cards present in the expander. The begin of a run starts first with the initialization of the detectors. As I mentioned before, the whole hardware system is faithfully represented by virtual classes, thus the main functionalities of each detector are defined by a suitable public interface. When the user sends a begin run command to the producer, it spawns a thread for every single PCI card. Every thread first initializes all the pixel planes connected to the card and then waits for the data. When a hit is above threshold in the detector, its informations are immediately sent to the PTA to be stored in one of the two memories present on the card. When the memory limit is reached the FPGA sends an interrupt which is received by its associated thread and the readout sequence can thus start. Since every card has its own thread, the data collection is done in parallel. All the threads transfer the data read from the PTA into the shared memory, the storage place common to all different processes.

The producer is totally unaware of what happens to those data, its task is to collect them from different cards and store them temporarily together in a place where they can be then used by a subsequent event-builder and data storage mechanism.

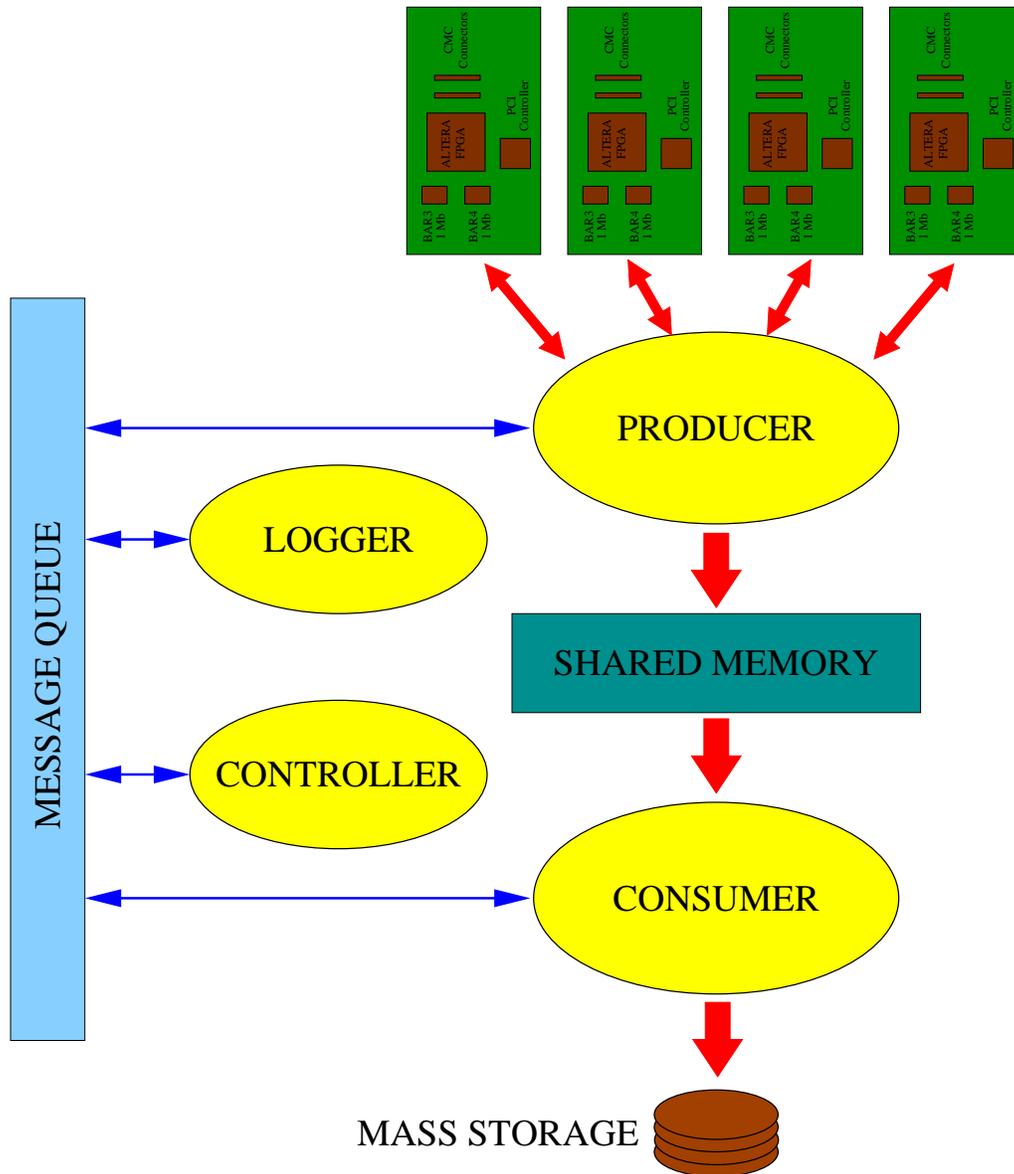


Figure 3.14: Global DAQ scheme.

3.6.2 Consumer

Once the data are stored in the shared memory, the consumer can use them to build the events. Data from different detectors are now stored in a common place and thus hits belonging to the same event are spread in this unique chunk of memory. The design was thought to keep the event-builder algorithm as simple as possible. We will see in the next chapter that the readout threads are synchronized and the data collection is not totally asynchronous. This architecture allows to collect hits

belonging to the same event almost simultaneously, thus storing them in small and well limited parts of the shared memory. The builder is then very simple because hits with equal time-stamp are not spread out over a too large chunk of memory.

The consumer doesn't know anything about the data origin. It only needs the time-stamp provided with every hit in order to build the event no matter what kind of detector or which card has produced it. This architecture completely decouples the place where hits are produced by the place where hits are consumed.

3.6.3 Logger

The logger receives messages from all the other running processes. Any time there is an error or any time a process wants to acknowledge the user of its status a message is sent to the logger. The logger then write those messages in a log file which can be used to keep a record of anything has happened during the run.

3.6.4 Controller

The controller is the interface between the user and the DAQ system. The producer and the consumer are waiting for commands that the user wants to execute. We provided two ways to send commands to those processes:

- a simple command line interface
- a sophisticated graphical interface (GUI) called Uffizi

In Fig. 3.15 it is shown a snapshot of the front-end of the graphical interface. This is a sophisticated interface written in C++ designed with the QT libraries. QT is a multi-platform C++ GUI application framework. It provides application developers with all the functionality needed to build applications with graphical user interfaces. QT is fully object-oriented, easily extensible, and allows true component programming.

This GUI is completely decoupled from the DAQ which can run without Uffizi.

3.6.5 Xmonitor

A tool much needed by a user running the system is a histogram presenter to continuously display significant quantities related to the data flowing in the system, see Fig. 3.16. Such a histogram presenter should not place any additional burden on the

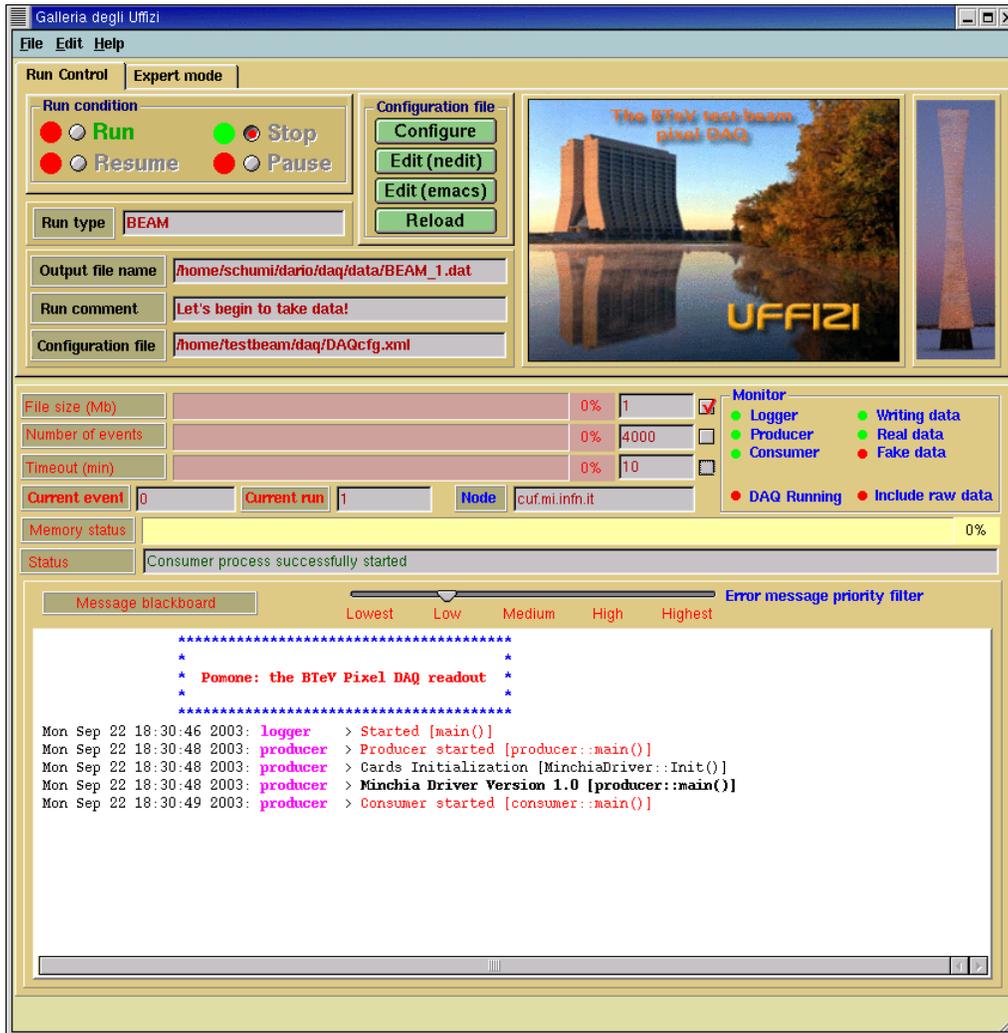


Figure 3.15: Main panel of the DAQ GUI.

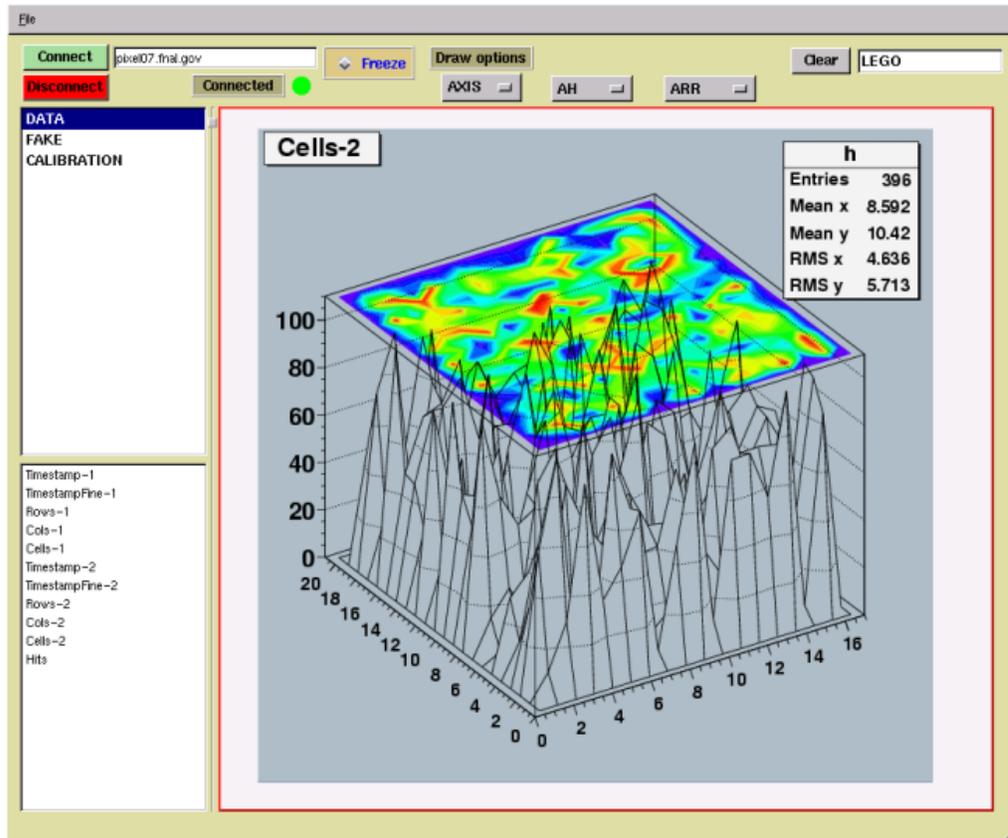


Figure 3.16: Histogram monitor.

host CPU where the producer and consumer are busy taking data. We have thus designed the presenter to receive the histograms as objects through the network (a socket). This tool has been developed in C++ too and uses the QTROOT libraries. ROOT is the analysis tool developed at CERN written in C++ while QTROOT is a merged library between QT and ROOT.

Chapter 4

Description of the DAQ mechanism

4.1 Introduction

The design of the pixel DAQ system was carried out focusing on the real operating conditions of the detectors in the BTeV experiment. At the same time, it allowed me to gain experience in software and hardware technologies that can be profitably used to design the future DAQ for BTeV.

Central to the DAQ is the use of PCI cards which can be easily controlled by custom made software written in C++. The software and firmware provide the card with the functionalities needed by the system to accomplish the sophisticated load-balancing described in Chapter 3, implementing a connection between the data flux provided by the pixel planes to the PTA memories and the readout PC.

A particular emphasis has been placed in the architecture to be able to read out events in a stream-like mode: each pixel element that has pulse-height data, above threshold, sends them to a memory along with space and time coordinates (row, column and time-stamp). Data are continuously read and the event reconstruction is done in a circular memory buffer using the time-stamp information, without any external trigger.

This stream like readout poses several problems for the event building because hits belonging to the same events are sparsified in different memories. All these problems have been elegantly solved and the description of the techniques we have adopted are the argument of this chapter.

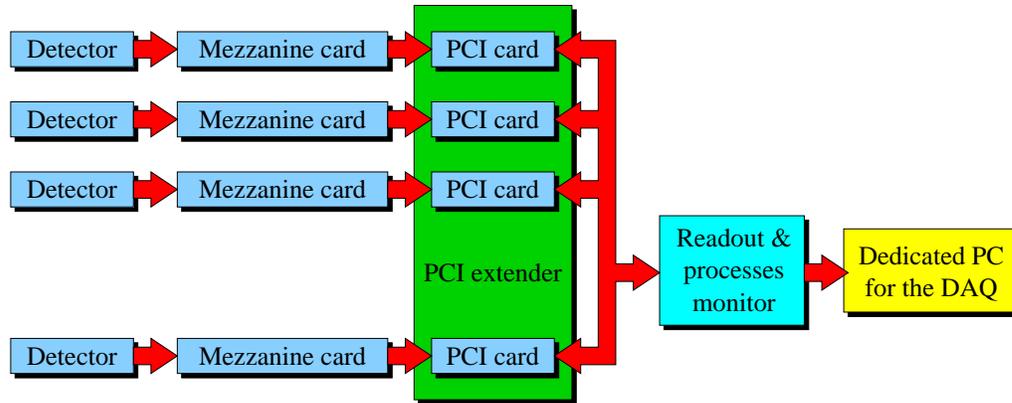


Figure 4.1: Schematic of the DAQ

4.2 Read out mechanism

In Fig. 4.1 is represented the data flow in the pixel test-beam.

In our design each detector is connected to a PMC/PCI board containing an FPGA (for logic control) and two 1 Mbyte memories; several PMC/PCI boards are lodged together on a PCI bus extender and finally connected to a host DAQ PC. Each time a pixel has data above threshold, the address along with pulse height and time-stamp information are sent to the PCI board to be stored on one of the two local memories. The FPGA's are programmed to handle the swapping between these two local memories and synchronization with the external readout process (running on the host DAQ PC) in such a way to smoothly handle a sustained data rate, adequate to the beam test requirements. The central idea in this design, is keeping the event-builder algorithm as simple as possible, since an event, defined by all hits marked by equal time-stamp, is spread out over several PCI boards which can in principle receive data at different rates. In absence of a specifically defined strategy to synchronize the flushing of these memories, this sparse readout scheme generates events spread out over large chunks of memory, making the event builder extremely cumbersome. We have therefore designed an elegant mechanism to restrict the components of an event to be contained in a limited amount of memory, taking advantage of our ability to program the FPGA to generate interrupt signals. The principle of operation of this readout scheme is the following:

- data are received from a detector by the corresponding PCI board in one of its two internal memories
- as soon as any memory in the system is full, all boards are synchronously

commanded to swap their memories. The ones used so far are frozen and immediately read out to the host computer, while the others are used to continue reading events from the detectors without any data loss.

- events are fed to the host computer on a statically allocated shared memory, acting as a compensating buffer to allow for unexpected data rate fluctuations.
- data are then continuously flushed from this memory to disk by a consumer process, building events on the fly.

This is an event-driven scheme: data are collected as soon as they are produced by a detector, and no burden is placed on the DAQ software to generate signals to start a readout chain. This is important, since it allows testing the full functionality of the detector in an environment similar to the one envisaged for the final data taking, where no trigger is used to readout events.

4.3 Working mechanism of a single PCI card

Let's see in detail the data acquisition mechanism that allows to read events in a stream-like mode and still be able to build events. I'll first describe the basic mechanism of operation of a single PCI card. All the PCI cards work in the same manner and we provided an elegant solution to keep the event builder as simple as possible as I will describe in Par. 4.4.

In Fig. 4.2 is represented the basic working mechanism of a single PTA card. Since there is no external trigger data are produced any time there is a pixel above threshold. The data received from the detector are stored in the first available empty memory bank. This readout process continues until either one of the following condition occurs:

- the number of data stored reaches the memory limit selected by the user
- a timeout, selected by the user, occurs

The first condition is used to limit the number of data stored in the memory buffer, while the second is crucial to avoid the presence, in the same buffer, of two hits with equal time-stamp belonging to different events. When one of this two condition occurs, the FPGA firmware redirects the data flow in the other empty memory bank. This memory swap is synchronized in such a way that even if there is a continuous data flux during this procedure all the data are correctly saved in one of

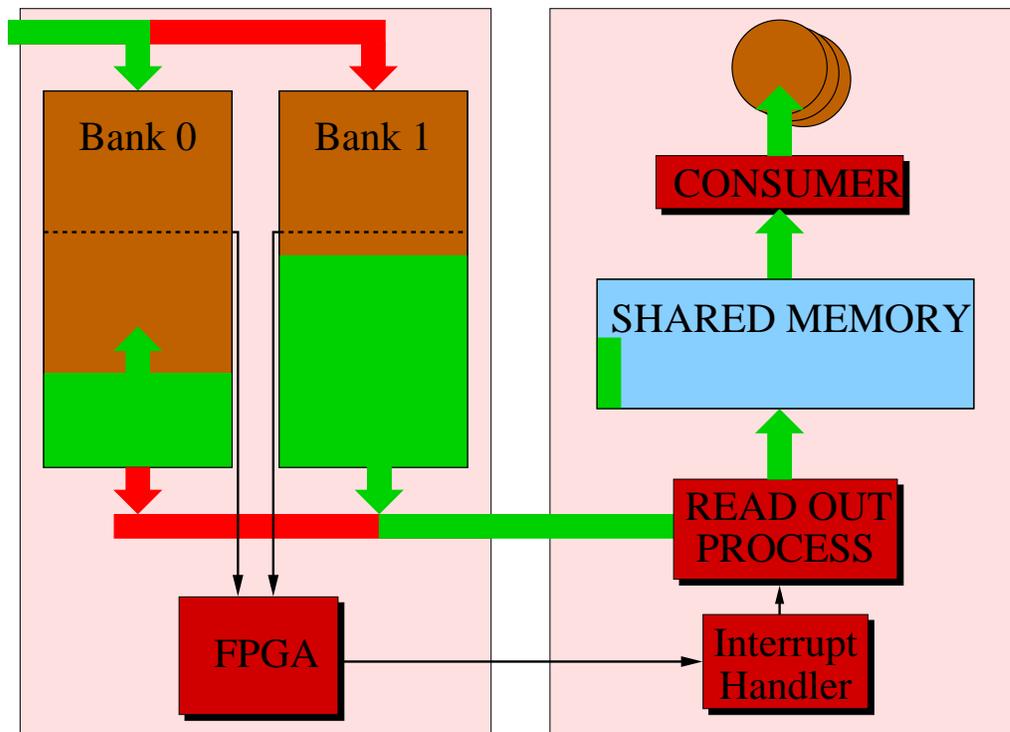


Figure 4.2: Working mechanism of a PTA. Bank 0 is being filled while Bank 1 is flushed by the readout process into the shared memory. The consumer then analyze data and write them to disk.

the two buffers and no data loss is made possible. As soon as there is a memory swap the FPGA sets a register in the ConfigRegister block which is directly connected with the interrupt line on the PCI bus of the host PC. When the software thread associated with this card receives the interrupt, it wakes up and immediately disable the interrupt on the PCI controller of the PTA. It then checks in the register, previously set by the FPGA, which memory bank is “full” and then starts to flush-out it. This procedure doesn’t interfere with the data taking: while one bank is still receiving data the other one is being flushed.

Since the producer and the consumer are different processes on the same PC they need to share a common memory block where data are first collected by the producer and then used by the consumer. Every PTA has its own thread which stores data in the same shared memory. Data are now ready to be flushed by the consumer process which can start the event building. Reconstructed events are then made persistent on a mass storage. This process of periodic memory swap and transfer continues indefinitely until the user decides to stop the data acquisition.

4.4 Synchronization of the PTA cards

The mechanism described in Par. 4.3 is done in parallel for every card since any card has its own thread. In absence of any synchronization mechanism the event building would be a non-trivial task because there are no clear boundaries in time within which to look for an event element. In order to keep the event builder as simple as possible we found an elegant solution by synchronizing the memory swap of all the cards. If each PCI board is read out as soon as one of its banks is full, hits with same time-stamp can end-up anywhere in the readout shared memory. In Fig. 4.3 is shown a typical situation that can happen without this synchronization. If board B has few data to send per event, on average, it will be read out relatively rarely, so hits pertaining to events already written down earlier will be spaced far away in this memory (actually anywhere).

If, on the other hand, we synchronize the swapping (the clock governing this swapping can be set by the board with the first memory reaching the full status), events with same time-stamp will end up not too far away in the shared memory. Pieces of one event will actually be at most in two adjacent buffers corresponding to two consecutive swap cycles, see Fig. 4.4. This synchronized cyclic swapping needed a strategy to be correctly implemented. The first board reaching the full status forces the others to swap (upon appropriate checks that they can do in order to avoid conflicting orders that can enact unwanted multiple swaps).

A complex synchronization among boards has been setup because each board must be somehow knowledgeable about each others readout and swapping status. Each PCI card has its own interrupt-handler process listening for the memory-full signal. As soon as one PCI fills up one of its two 1Mb memories, its interrupt handler gets notified and a series of actions occur: lets examine a typical scenario. Suppose, for instance, that the PCI A is the first being filled up, see Fig. 4.4. The memories of the PCI A get swapped. At this point we would like to swap all other memories together and write their content to a shared memory. The first card that gets filled become the one that commands all others to swap (what we call “the master swapper”). At any given time, though, only one master can be in charge, so we provided a mechanism to establish a hierarchy among boards, because only one board at a time can order swaps, otherwise chaos can ensue. An auxiliary shared memory has been used for this purpose. When an interrupt-handler gets notified that a swap has occurred, it first checks whether another master is already at work, if true, nothing occurs, this board is just passive and does nothing else than flushing

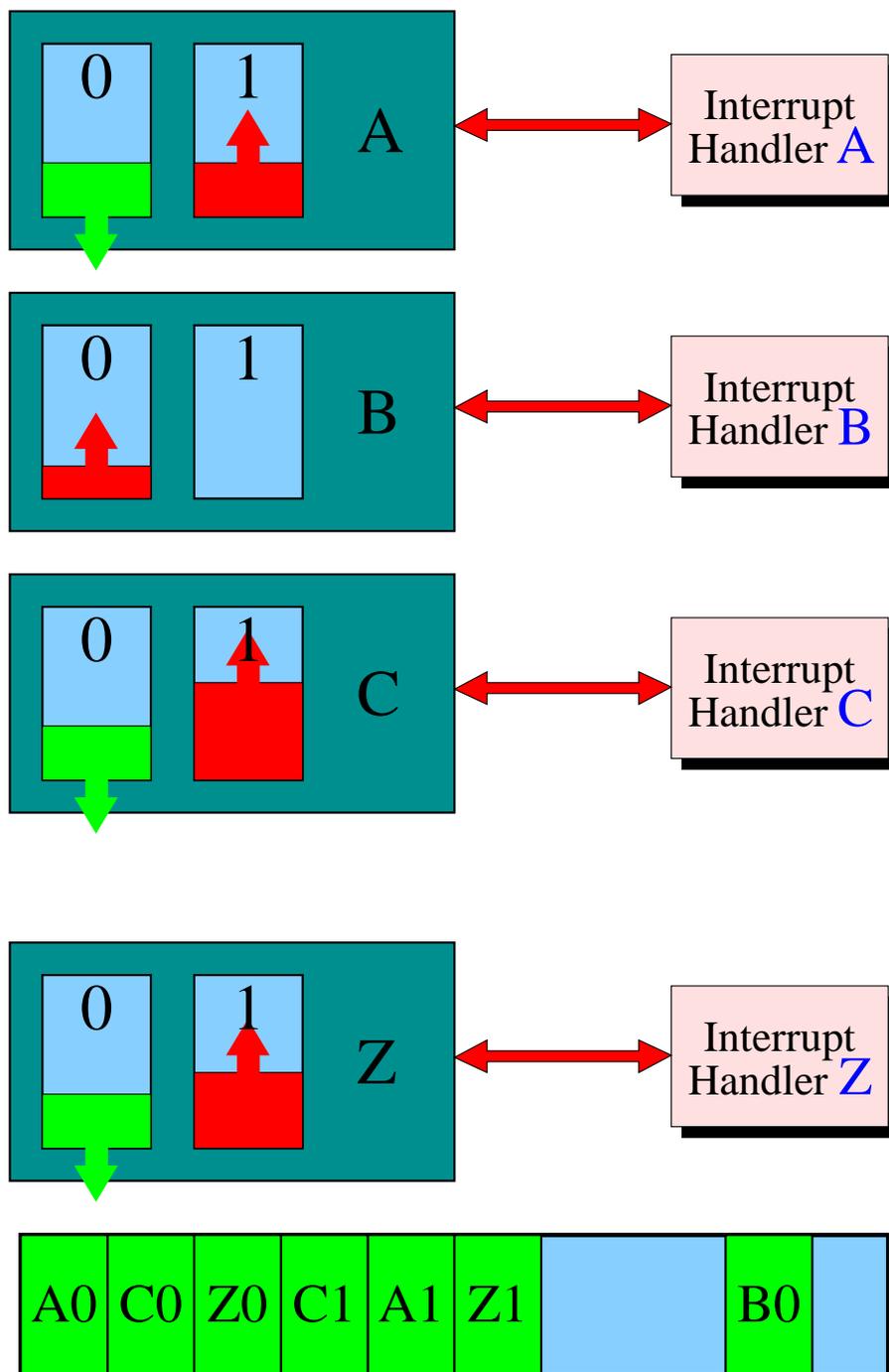


Figure 4.3: In absence of any synchronization mechanism, if board B has few data, on average, it will be readout relatively rarely while, instead, boards A, C and Z are flushed many times. Events with the same time-stamp can thus end up far away in the shared memory making the event builder difficult to implement.

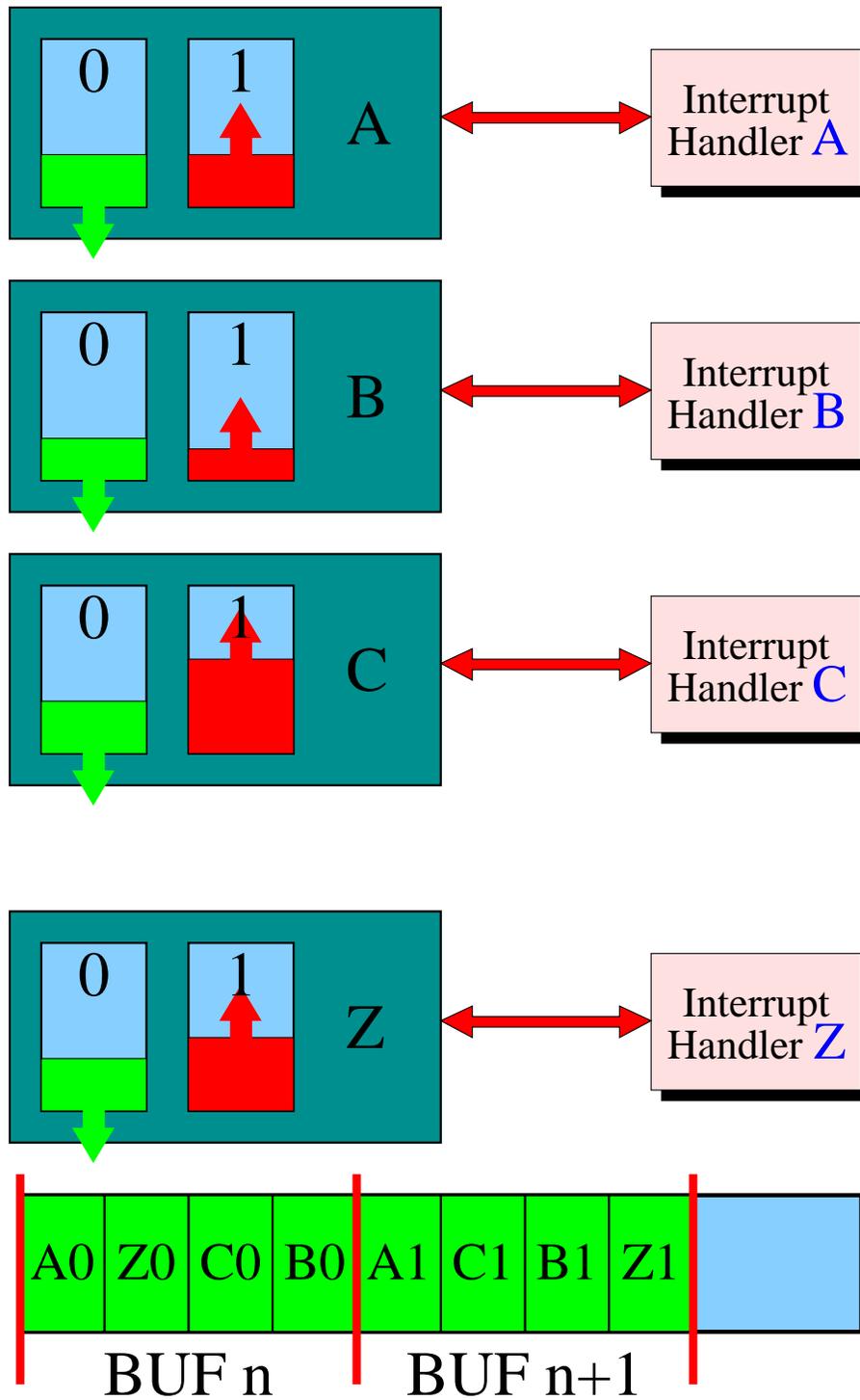


Figure 4.4: With our synchronization mechanism all boards always flush the same memory buffer, here bank 0, on all the PTAs. Events are thus collected in two adjacent buffers at most, BUF n or BUF $n+1$, because an event can be split only when a memory swap occurs.

the memory bank, if false, the auxiliary shared memory is flagged, the PCI A becomes the master swapper and, whether or not each of the other PCI card memories are currently full, it forces them to swap and generate an interrupt. Problem is that interrupt-handler A needs a few CPU cycles to swap all other PCI memories and during this time one of them can become full and swap. In this case I provided on the FPGA a locking mechanism that prevents additional unintended swaps: if the pointer of a bank is not zero, the swap simply does not occur, see Par. 3.5.2. This pointer is reset by the readout process when the last bank has been read out.

This architecture guarantees that events with nearby time-stamps belong to buffers which are contiguous in the readout process, and thereby contiguous in the global shared-memory. As a result the event builder becomes relatively easy to implement. With this readout architecture, events with the same time-stamp are contained within the boundaries of this overall buffer, BUF n , or at least in the next one, BUF $n+1$, but not in BUF $n+2$, making the event-builder a rather trivial implementation of a sorting algorithm.

4.5 Event builder

This synchronization mechanism has been implemented to reduce the complexity of the event builder. In Fig. 4.5 are shown the three stages of the hit-handling in the consumer.

Hits are first collected in the shared memory by the interrupt-handlers of each card. The event builder is implemented with a matrix where the columns are ordered by the time-stamp and the rows are filled by hits with equal time-stamp: an event is thus a column of this matrix. The algorithm checks any data and when a new time-stamp is found it creates a new event (column). Any data which, instead, has a time-stamp already present in the matrix is attached to the right column. This procedure is performed for the BUF n and for the BUF $n+1$. When the analysis of the BUF $n+1$ is completed, it's reasonable to assume that there are no more hits in those events whose first time-stamp was found in BUF n . This assumption is reasonable because events fill the PTA memories at the same instant and thus, with our synchronization mechanism, fill always the same memory bank of all the cards. Hits belonging to the same event are all collected either in the memory bank 0, BUF n , or in bank 1, BUF $n+1$. An event is incomplete before two buffers are analyzed. In fact, during the PTA memory swap, it can be split between two adjacent buffers in the shared memory: the event builder can thus declare an event

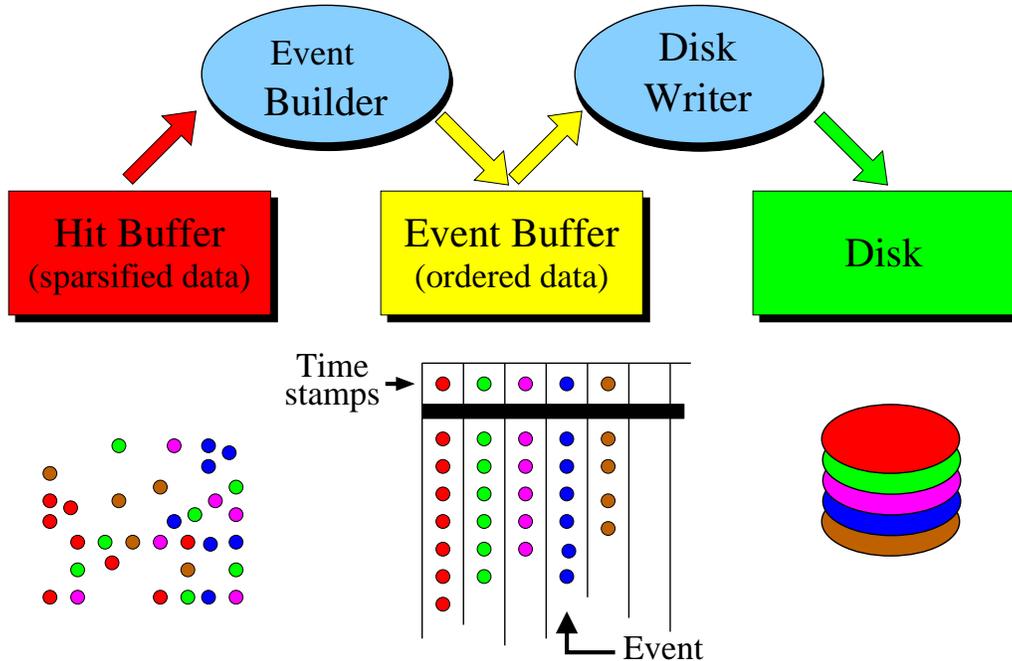


Figure 4.5: Event builder

completely built only after the time-stamp ordering of two buffers. When a column is completed, the consumer write the event to disk freeing the relative buffer column.

4.6 Preliminary tests

The full DAQ system has already been tested in the Meson area where the test-beam will be carried on. All tests have been performed injecting charge in the pixel planes synchronously. Charge has been injected in four pixel detectors connected to two different PMC cards. This minimal system allowed us to test all the features of the DAQ. All hardware and software components worked perfectly and the event builder described above was able to collect hits into events using the extended time-stamp information provided by the mezzanine card.

4.7 Benchmarks

The PLX 9030 PCI Controller on the PTA card we are using does not allow for DMA transfer. It could in principle be used in burst-mode, but this feature is not

Table 4.1: Data transfer rates from the PTA on different architectures

	AMD Athlon (1.3 GHz) and Linux	INTEL PentiumIV (1.8GHz) and Windows 2000
Direct connection to motherboard	6 Mb/sec	4Mb/sec
Connection to motherboard by PCI extender	2.6 Mb/sec	2Mb/sec

available on the INTEL architecture. We are thus forced to run in single word transfer mode, but this seems more than adequate for our beam test needs. In fact we decided to run the DAQ on a AMD processor with Linux OS from the results obtained in Table 4.1. In this configuration we achieve a 2.6 Mbyte/sec data rate with the PCI cards connected to the motherboard through the expander, enough to sustain the data flux from the pixel planes.

Chapter 5

Lorentz angle measurement

5.1 Introduction

The DAQ system I developed with the Milano group is designed to be used for the test-beam at Fermilab. This system has been designed to be extremely flexible: it can be used to test a single pixel detector in a test bench or a full set of detectors in a complex test-beam environment.

The test-beam purpose will be to measure the detector performances (e.g. spatial resolution) in a readout configuration similar to the one envisaged for the experiment, where the detector will provide data at his own pace, independently of an external trigger. This will be the first time where the ROC chip (FPIX) will be used in this configuration on a beam. The pixel detector will operate in a high magnetic field of $1.5\ T$ and the determination of the position of the particles passing through the detector will be deteriorated by the effect of this field on the charge carriers in the silicon. In fact, the charge released in the detector will experience a Lorentz force and will be deviated from the original trajectory resulting in a displacement of the collected signal. The deviation angle is called Lorentz angle and its study has been the subject of part of my thesis and will be the argument of this chapter.

The study of the Lorentz angle is crucial to provide the corrections needed to take into account the displacement of the collected charge induced by the magnetic field in order to improve the track reconstruction resolutions.

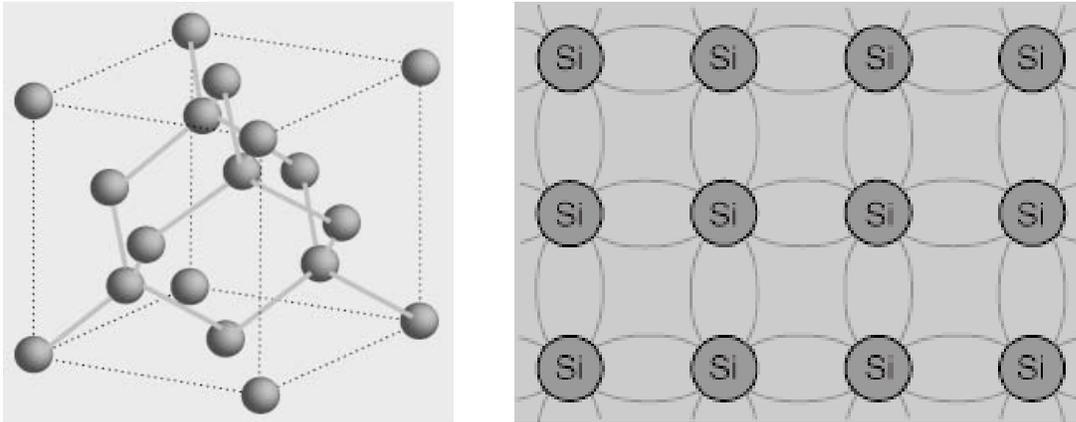


Figure 5.1: Silicon lattice structure

5.2 Physics of silicon detectors

5.2.1 Basic semiconductor physics

In order to understand the physics of semiconductor detectors let's first consider the atomic bonding. Silicon detectors have a characteristic crystal structure. The atoms within the crystal are arranged in a typical periodic order known as lattice. Every symmetric lattice has a representative unit cell. Silicon has a diamond lattice structure as shown in Fig. 5.1 and belong to the cubic-crystal family. This characteristic lattice has its origin in the covalent bonding of the four electrons in the outer orbit where each atom shares its valence electrons with its four neighbors.

For an isolated atom, the electrons of the atom can only have discrete energy levels. When we bring N atoms together to form a crystal, the N -fold degenerate energy level will split into N separate but closely spaced levels due to atomic interactions. This results in a continuous band of energies. The detailed energy band structure of crystalline solids can be calculated in quantum mechanics. Further decrease of the spacing causes the band originating from different discrete levels to lose their identities and merge, forming a single band as shown in Fig. 5.2. When the distance between atoms approaches the equilibrium inter-atomic spacing of the diamond lattice, this band splits again into two bands, see Fig. 5.3. These bands are separated by a region which designated energies that electron in the solid cannot occupy. This region is called forbidden gap, or band gap E_{gap} . The upper band is called conduction band, while the lower band is called the valence band. A full energy band cannot conduct and neither can an empty one. At very low temperatures, the va-

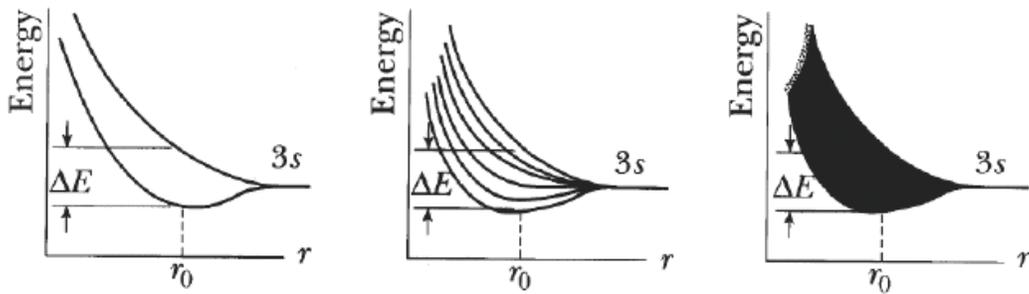


Figure 5.2: Level splitting in solids

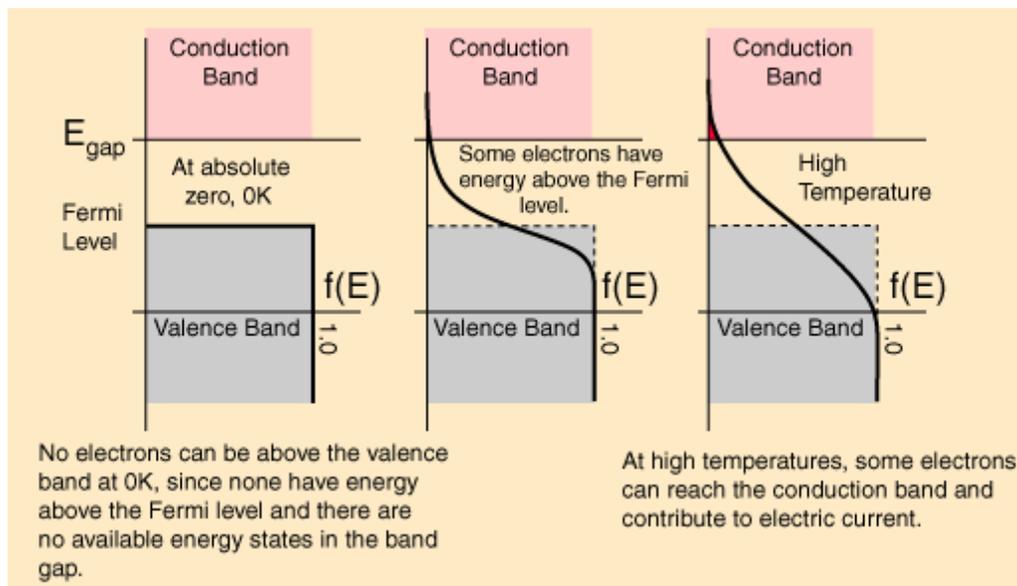


Figure 5.3: Fermi function in a semiconductor

lence band is fully occupied, and the conduction band is completely empty. Under these conditions, no current flows and the semiconductor acts as if it were an insulator. An important parameter in the band theory is the Fermi level, the top of the available electron energy levels at low temperatures. The position of the Fermi level with relation to the conduction band is a crucial factor in determining electrical properties. The Fermi level is the term used to describe the top of the collection of electron energy levels at absolute zero temperature. This concept comes from Fermi-Dirac statistics. Electrons are fermions and by the Pauli exclusion principle cannot exist in identical energy states. So at absolute zero they pack into the lowest available energy states and build up a "Fermi sea" of electron energy states. The

Fermi level is the surface of that sea at absolute zero where no electrons will have enough energy to rise above the surface. For intrinsic semiconductors like silicon, the Fermi level is essentially halfway between the valence and conduction bands. Although no conduction occurs at $0^\circ K$, at higher temperatures a finite number of electrons can reach the conduction band and provide some current. The increase in conductivity with temperature can be modeled in terms of the Fermi function which allows one to calculate the population of the conduction band. The Fermi function gives the probability of occupying an available energy state, but this must be factored by the number of available energy states to determine how many electrons would reach the conduction band. The number of particles per unit volume with energy between E and $E + \Delta E$, $n(E)$, is given by the product of the number of allowed energy states per unit volume, $\rho(E)$, and the probability of occupying that energy range, given by the Fermi function

$$f(E) = \frac{1}{1 + e^{(E-E_F)/kT}} \quad (5.1)$$

where k is the Boltzmann constant, T the absolute temperature in Kelvin and E_F is the Fermi level. Integrating $n(E)$ from the bottom of the conduction band ($E_c = 0$, for simplicity) to the top of the conduction band E_{top} we get the particle density of allowed energy states per unit volume:

$$N = \int_0^{E_{top}} \rho(E)f(E)dE \quad (5.2)$$

Using equation (5.2) the densities of the electrons and of the holes can be calculated

$$n = N_C e^{-\frac{E_C - E_F}{kT}} \quad (5.3)$$

$$p = N_V e^{-\frac{E_F - E_V}{kT}} \quad (5.4)$$

The lowest energy state in the conduction band is called E_C while E_V denotes the highest energy state in the valence band. The effective density of state in the valence band is N_V and N_C is the effective density of states in the conduction band. In case of equilibrium the electron density and the hole density are equal: $n = p = n_i$, with n_i the intrinsic carrier density. For intrinsic silicon the Fermi level is halfway the band gap. This has the consequence, that at room temperature the concentration of free carriers is very low (approximately every 10^{12} atoms are ionized). This number of free charge carriers can be increased if the silicon is doped with impurities. The semiconductor becomes extrinsic and further impurity energy levels are

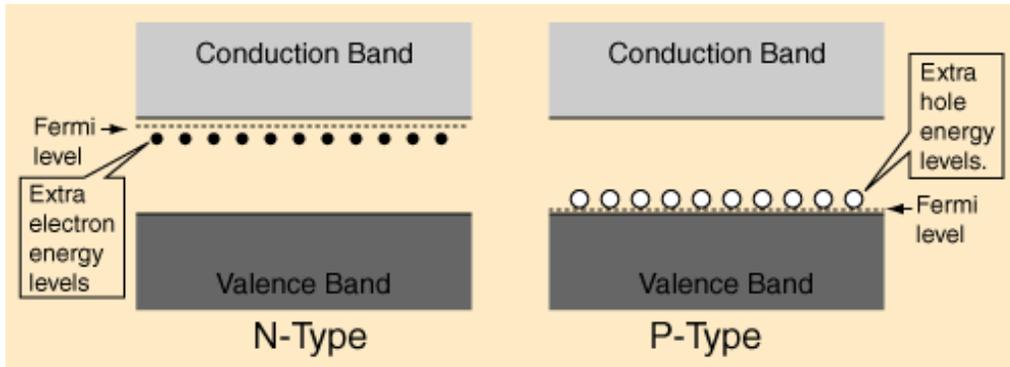


Figure 5.4: The left drawing shows the extra energy levels introduced between the conduction band and valence band in n-type doped silicon. The right drawing shows the extra energy levels for a p-type doped silicon.

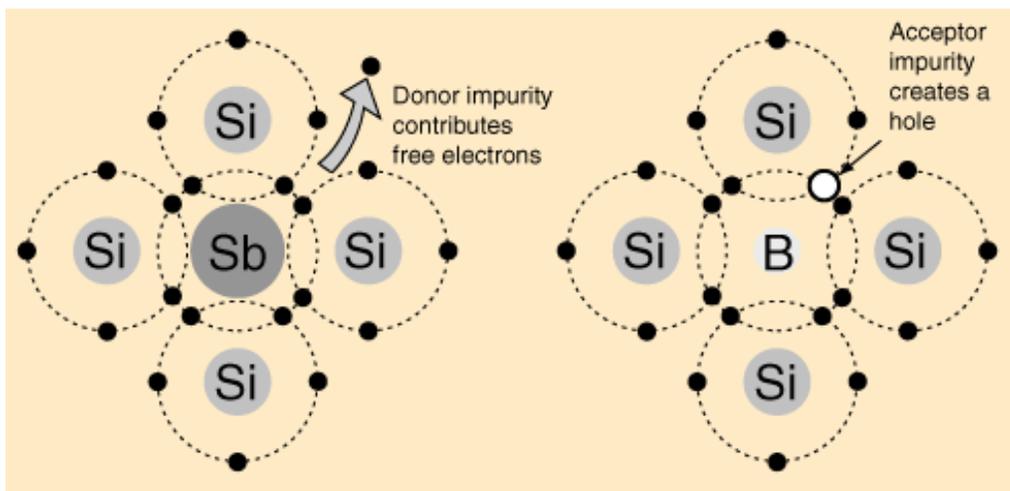


Figure 5.5: Donor (on the left) and acceptor (on the right) dopant impurities.

introduced, see Fig. 5.4. Depending on the number of valence electron the atom bears, see Fig. 5.5, it is called donor (five valence electrons, such as antimony) or acceptor (three valence electrons, e.g. boron). The intrinsic silicon becomes n-type if a considerable number of donors is introduced. If the same is done with acceptors, the silicon becomes p-type. In the p-type region there are holes from the acceptor impurities and in the n-type region there are extra electrons.

Most semiconductor devices, however, incorporate both p- and n-type silicon and it is the junction between these regions that leads to their useful characteristics. When a p-n junction is formed, some of the electrons from the n-region which have reached the conduction band are free to diffuse across the junction and combine with holes,

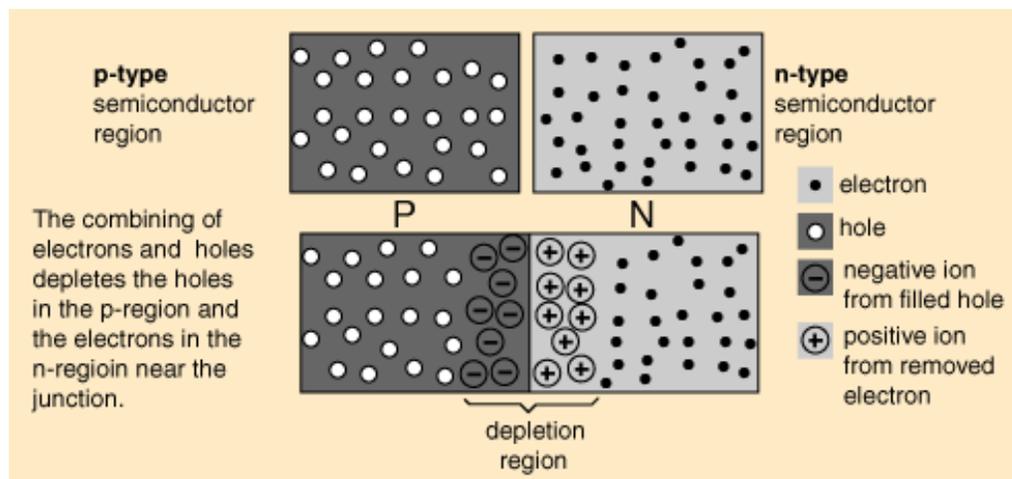


Figure 5.6: Depletion region in a P-N Junction

see Fig. 5.6. Filling a hole makes a negative ion and leaves behind a positive ion on the n-side. A space charge builds up, creating an electric field which prevents any further flow of carriers. The space charge region is called depletion region. The responsible potential is called the *built-in* voltage, V_{bi} , leading to a deformation of the conduction and valence bands as we see in Fig. 5.7. This deformation can be increased by applying an external voltage, V_{bias} , which amplifies the bending and enlarges the depletion region, see Fig. 5.8.

Conventional silicon detectors used in high energy physics are segmented p-n junction diodes that operate upon an external bias voltage. This forms a sensitive depleted region of mobile charge and sets up an electric field that sweeps charge liberated by ionizing particles to the collecting electrodes.

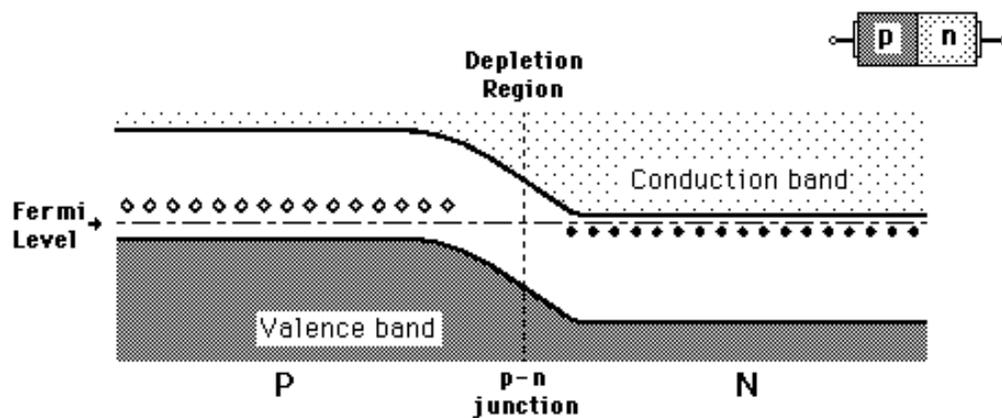


Figure 5.7: Deformation of the valence and conduction bands at equilibrium.

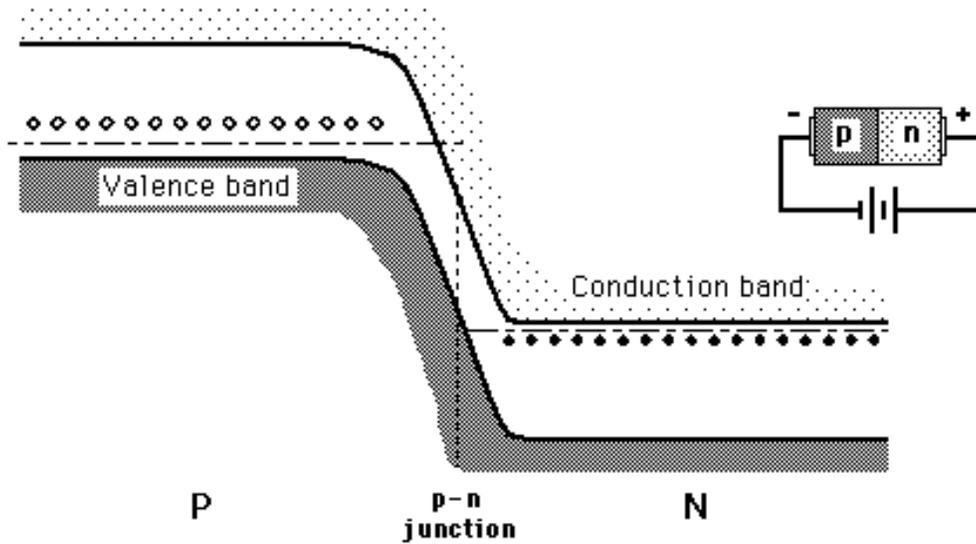


Figure 5.8: Deformation of the valence and conduction bands with a reverse bias applied.

5.2.2 Charge transport

In silicon detectors two ways of motion are distinguished: drift and diffusion. While the latter is not limited to specific particles (also impurities and crystal imperfections can diffuse), drift caused by an external electric field affects only the free charged particles in the semiconductor lattice. If carrier movement into a specific direction is observed, despite of the mechanism which is responsible for this movement, a current can be measured. In semiconductors often external influences such as electric or magnetic fields cause a charge drift into a specific direction.

Drift

If a small electric field is applied to a semiconductor, each carrier will experience a force $\vec{F} = -q\vec{E}$ and will be accelerated along the field. Because of the phonon scatterings with the lattice and with impurities the charge will lose momentum. Macroscopically, a drift with constant velocity is observed. The momentum applied to a carrier between two collisions is given by $\vec{p}_{drift} = -q\vec{E}\tau$ and the momentum gained is $\vec{p}_{drift} = m^*\vec{v}_{drift}$, where m^* is the effective mass of the carrier in the lattice and τ is the mean free time between two scattering events also denoted as relaxation time. The mass of the carriers, m^* , differs from the free carriers mass, m_0 , in fact,

the presence of the periodic potential, due to the atoms in the crystal without the valence electrons, changes the properties of the carriers. The drift velocity is then easily obtained by equating

$$-q\vec{E}\tau = m^*\vec{v}_{drift} \quad (5.5)$$

$$\vec{v}_{drift} = -\left(\frac{q\tau}{m^*}\right)\vec{E} = \mu_{e,h}\vec{E} \quad (5.6)$$

Equation 5.6 states that the drift velocity is proportional to the applied electric field. The proportionality factor is called electron and hole mobility, respectively μ_e and μ_h .

Equation 5.6 implies that a carrier can be accelerated as fast as one likes by increasing the electric field. Unfortunately this is not the case and equation 5.6 is not valid for high electric fields.

In fact, under a high electric field, free charge carriers (electron and holes) gain energy from this field. Since the carriers must remain in the band structure of the semiconductor this energy has to be mediated to the lattice through increase of carrier-phonon scattering. This effect becomes important if the energy gained from the field is of the same order of magnitude as the thermal energy of the carriers (when the drift velocity $\mu \cdot \vec{E}$ becomes greater than the thermal velocity $\approx \sqrt{kT/m^*}$). Since this behavior is similar to an increase of the carrier temperature the carriers are said to be hot (hot electron effect, see e.g. Ferry in [39]). Due to the increased phonon scattering the mobility does not remain constant but decreases with increasing field strength. The expected energy increase goes into heating up the the lattice rather than into the kinetic energy of the carrier. Then, equation 5.6 is no longer valid because the drift velocity does no longer increase linearly with the field. At very high electric fields the velocity almost saturates, increasing very slightly for increasing electric fields. An analytic expression based on an empirical model for the mobility dependence on doping density and electric field strength is derived in reference [40]

$$\mu = \frac{\mu_0}{\left(1 + \left(\frac{E}{E_c}\right)^\beta\right)^{1/\beta}} \quad (5.7)$$

with

$$\mu_0 = \mu_{min} + \frac{\mu_{max} - \mu_{min}}{1 + \left(\frac{N_{eff}}{N_{ref}}\right)^\alpha} \quad (5.8)$$

and

$$E_c = v_m/\mu_0 \quad (5.9)$$

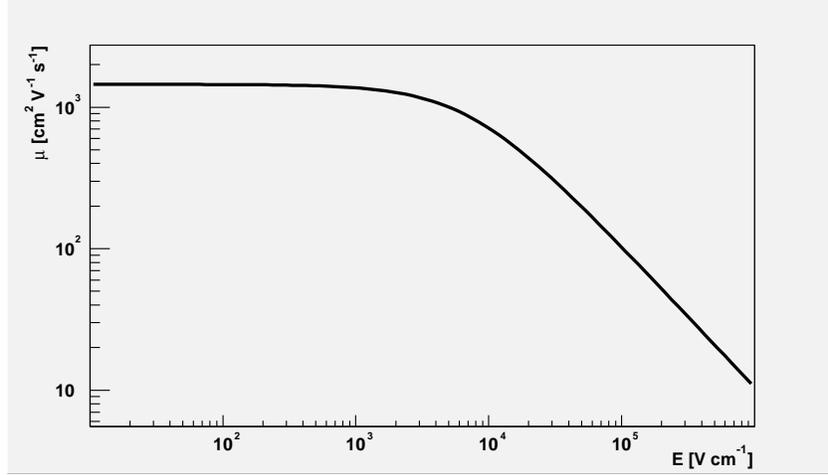


Figure 5.9: Mobility vs electric field equated for electrons in silicon according to equation 5.7 at $T = 300^0K$ and $N_{eff} \leq 10^{16}cm^{-3}$

where N_{ref} and α are normalization parameters, N_{eff} is the doping density and v_m is a constant related to the maximum saturation velocity. If the doping variation is small ($N_{eff} \leq 10^{16}cm^{-3}$) μ_0 can be assumed to be constant and equal to μ_{max} . The three parameters v_m , β and μ_0 have been found empirically [41] and they depend on temperature. The field dependence and the temperature are thus the only remaining scaling parameters. Using Eq. 5.7 the mobilities of electrons and holes can be calculated, as visualized in Fig. 5.9.

Furthermore complexity increases if in addition a magnetic field comes into play. It was shown [42] that the transport of electrons and holes in semiconductors in electric and magnetic fields can be described relatively well by a single classical vector equation for each particle. One parametrization is

$$m^* \cdot \frac{d^2\vec{r}}{dt^2} = q \cdot \frac{\mu_H}{\mu} \cdot \frac{d\vec{r}}{dt} \times \vec{B} + q \cdot \vec{E} + \frac{q}{\mu} \cdot \frac{d\vec{r}}{dt} \quad (5.10)$$

where \vec{r} is the position vector of the electron or hole, \vec{B} the magnetic field and \vec{E} the electric field. The mobility μ is introduced in equation 5.7 and μ_H is the Hall mobility (see Eq. 5.20).

Diffusion

Another transport mechanism well known in the kinetic gas theory is diffusion. If there is a spatial variation of carrier concentration in the semiconductor material, the carriers tend to move from a region of high concentration to a region of low

concentration. This diffusion, first described by Nernst in the kinetic gas theory was extended by Einstein relating the mobility of electrons and holes to the diffusion coefficient D_n

$$D_n = \frac{kT}{q}\mu \quad (5.11)$$

The nearly Gaussian distribution for electrons and holes by diffusion is explained in the theory of Brownian motion. A correlation between the standard deviation and the diffusion coefficient is given by

$$\sigma = \sqrt{2Dt} \quad (5.12)$$

where t is the drift time. Diffusion is also active in the depletion region where drift governs. Using equation 5.12 the spread of a charge cloud can be calculated. In the neutral regions outside the depleted region there is no significant electric field and the carriers move only by diffusion. If electron-hole pairs are generated in the undepleted region there is a finite chance that some of the carriers reach the depletion region before they recombine causing diffusion current.

5.2.3 Electric field distribution

Figure 5.10 shows the electric field distribution on the example of the one-sided abrupt junction with a p^+ -implant in n-bulk material. For high p^+ -doping, the growth of the space charge region into the region of high doping is negligible small compared to the growth toward the n-side. According to the simple development of \vec{E} along the depletion depth it can be approximated with a straight line. From Fig. 5.10 it can be seen that the electric field is maximal for $x = 0$. Taking this into consideration the electric field strength in x can be calculated as

$$E(x) = \frac{2V_{dep}}{d} \left(1 - \frac{x}{d}\right) \quad (5.13)$$

Here it is assumed, that V_{dep} is the depletion voltage used to fully deplete the detector with the thickness d . After the full depletion of the detector the internal electric field is no longer due to the spatial charge in the doped regions so the distribution of the field after depletion is

$$E(x) = \frac{2V_{dep}}{d} \left(1 - \frac{x}{d}\right) + \frac{V_{bias} - V_{dep}}{d} \quad (5.14)$$

where V_{bias} is the bias applied to the sensor. This model satisfies the description of the electric field of simple devices such as junction diodes or surface barrier detectors.

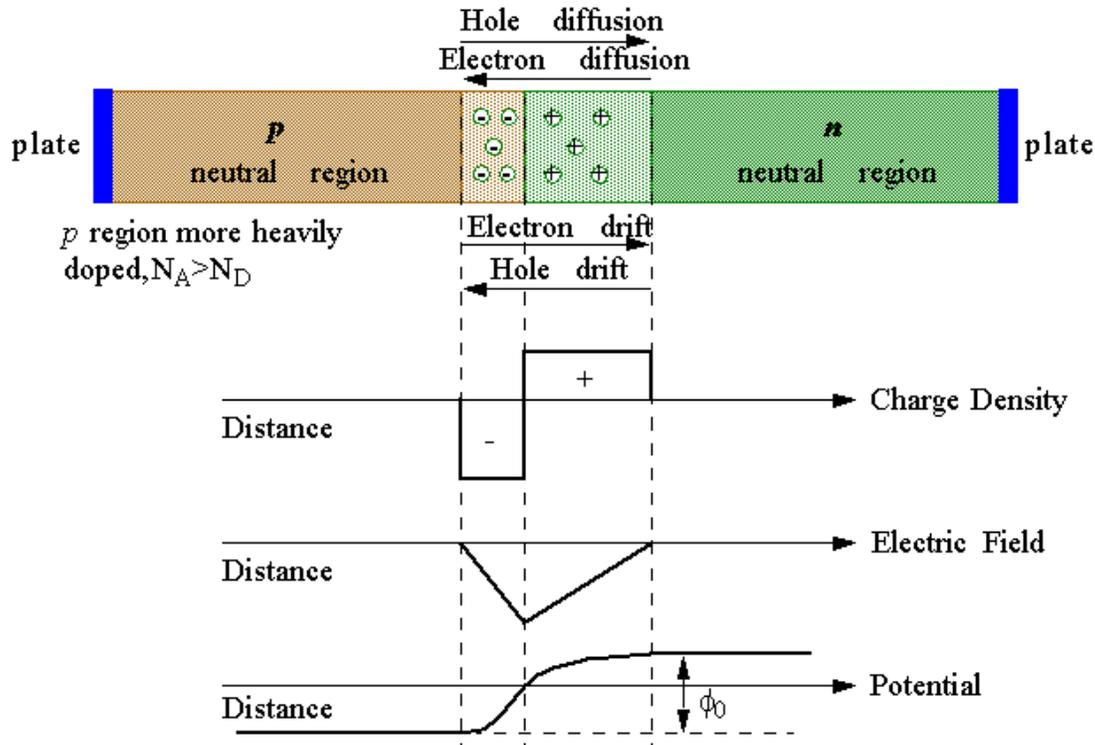


Figure 5.10: Electric field distribution in a pixel detector

In case of more complex devices as for example strip or pixel detectors the field lines are not parallel throughout the entire detector. Near the surface the field lines are focused toward the charge collecting implants. Simulations shows that this bending is dependent of the size of the implants and the gap, between them.

5.3 Lorentz effect

A general equation of motion for charge carriers in an electric field and under the influence of a magnetic field was quoted in equation 5.10.

The Lorentz force, see Fig. 5.11, acts on every charge carrier with a velocity component perpendicular to the magnetic field with

$$\vec{F}_L = -q\vec{v} \times \vec{B} = -qv_z B_y \quad (5.15)$$

This force gives the carriers a new direction, v_L and the angle between the initial and the resulting direction is

$$\tan \Theta_L = \frac{v_x}{v_z} \quad (5.16)$$

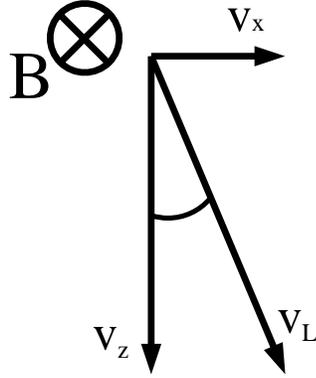


Figure 5.11: Velocity components under the influence of a magnetic field

The value for the velocity along the direction of the electric field is given in Eq. 5.6. Assuming that an electron is deflected its velocity v_x with respect to the magnetic field can be derived from equation 5.15

$$v_x = -\frac{q\mu_e E_z B_y \tau}{m^*} \quad (5.17)$$

Substituting the values in equation 5.16 lead to

$$\tan \Theta_L = -\left(\frac{q\tau}{m^*}\right) B_y \quad (5.18)$$

$$\tan \Theta_L = \mu_H B_y \quad (5.19)$$

Equation 5.19 is the general relation between the applied magnetic field and the Lorentz angle Θ_L . The proportionality factor μ_H is the Hall mobility which is related to the drift mobility via

$$\mu_H^{e,h} = r_H \mu_{e,h} \quad (5.20)$$

The hall factor r_H is a dimensionless value determined to be [43]

$$r_H^e = 1.15 \quad (5.21)$$

$$r_H^h = 0.72 \quad (5.22)$$

for electrons and holes.

5.4 Experimental measurement

5.4.1 Setup description

As described in Par. 5.3, the Lorentz angle Θ_L is defined by:

$$\tan \Theta_L = \frac{v_x}{v_z} \simeq \frac{\Delta x}{d} = \mu_H B = r_H \mu B \quad (5.23)$$

where d corresponds to the drift distance along the electric field and Δx to the shift of the signal position, and in this case we measure an effective Lorentz angle value. The Hall mobility is denoted by μ_H , the conduction mobility by μ . The Hall mobility differs from the conduction mobility by the Hall scattering factor r_H . This factor describes the influence of the magnetic field on the mean free path of carriers of different energy [44]. The Hole scattering factor has a value of ≈ 0.72 for holes and ≈ 1.15 for electrons at room temperature.

When a particle passes through the pixel sensor, it releases the charge all along

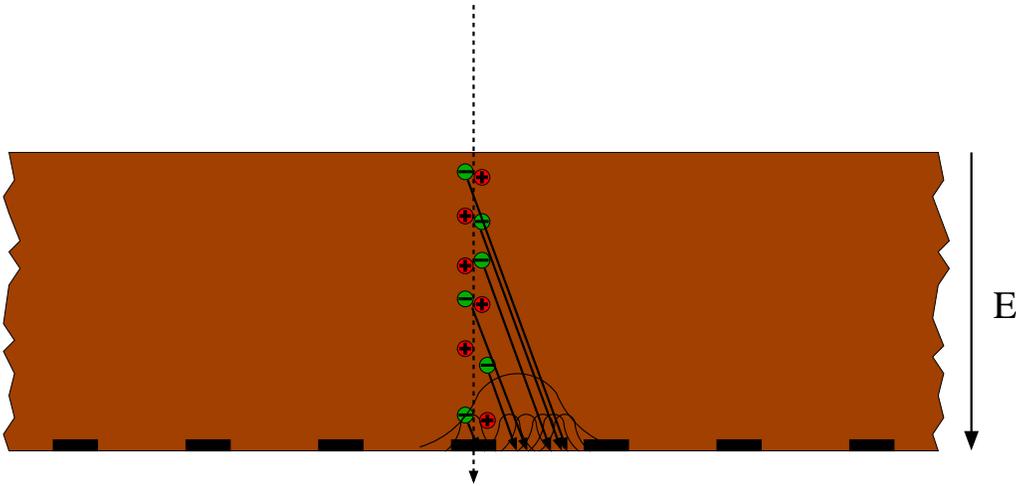


Figure 5.12: The figure shows the charge released by a particle passing through a silicon detector. In presence of a magnetic field the Lorentz force acting on the carriers deviates the trajectory modifying the shape of the cluster.

the whole trajectory. In absence of a magnetic field the charge drifts following the electric field lines of the bias present in the sensor. The signal position doesn't depend by the depth where the charge is generated. When a magnetic field is present, instead, there is also a Lorentz force acting on the carriers which deviates their trajectory. The magnitude of this displacement depends on the depth where the

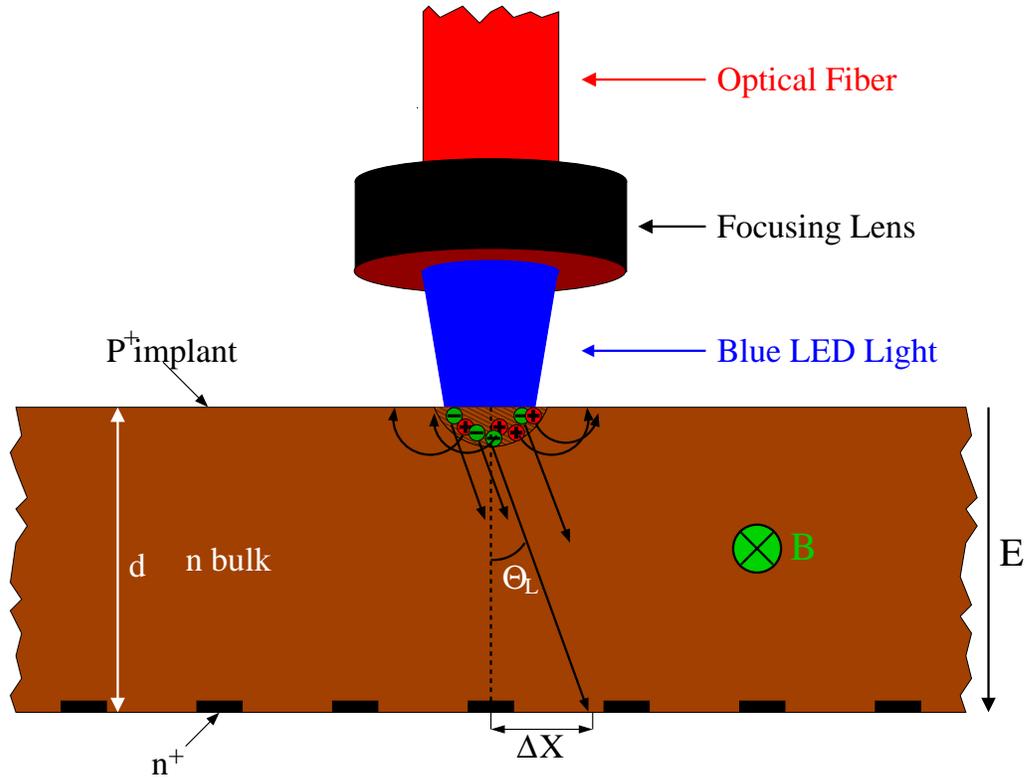


Figure 5.13: The figure shows the principle of operation of the setup to measure the Lorentz angle of electrons. It is equipped with a fiber delivering the blue LED light which has a penetration depth of $\approx 1\mu m$, connected to a focusing lens. The detector is illuminated on the p-side, thus holes are collected immediately while electrons move through the entire detector.

carriers are generated as it is shown in Fig. 5.12. The shape of the cluster in presence of the magnetic field is thus different and a direct measurement of the displacement of the center of gravity of the cluster is not enough to provide the correct estimation of the Lorentz angle because one has to take into account the convolution of the charge production point along the particle trajectory in the detector. An easy way to measure the Lorentz angle, instead, is to produce the charge on the surface. In this case, all carriers are subject to the same displacement, resulting in a translation of the whole cluster with no change in the shape of the charge distribution. The displacement of the center of gravity is thus equal to the Lorentz shift acting on a single carrier.

In this case, the ratio between Δx and d is easy to calculate and it is the ratio

between the movement of the signal position and the detector thickness.

Following these simple considerations an experimental setup has been designed to give a direct measurement of the Lorentz angle. The system is composed by:

- a blue LED. The LED light has a wavelength of $\lambda \approx 450nm$, which has an absorption length in silicon of $\approx 1\mu m$ at room temperature. The LED has been glued to an optical fiber delivering the light to the silicon.
- A focusing lens to focus the LED light on few cells of the detector.
- A mechanical structure supporting both the pixel detector board and the lens where the optical fiber was connected.
- A magnet providing a magnetic field up to $\approx 0.6T$.

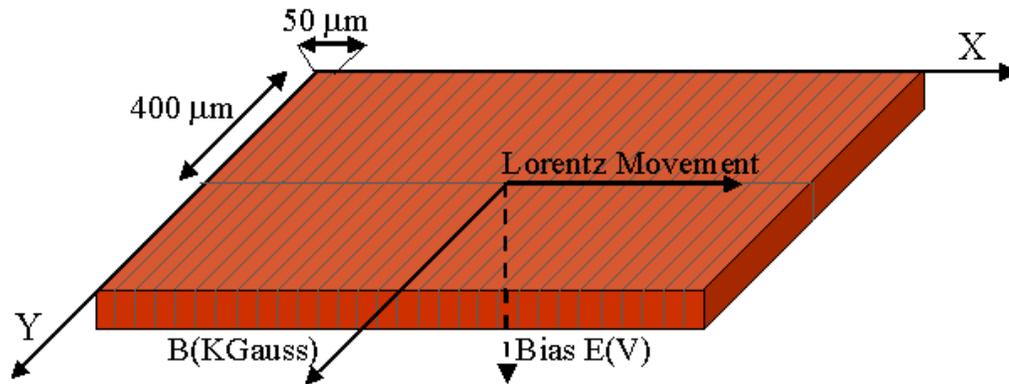


Figure 5.14: Orientation of the pixels in the magnetic field

The experimental hall temperature has always been around $24^{\circ}C$.

The pixel detector is a $n^+/n/p^+$ with the pixel electrodes located on the ohmic side of the device. The sensor was bump-bonded to an FPIX1 readout chip developed at Fermilab. The depletion voltage, measured through the dependence of the leakage current on the reverse bias applied is $45V$. The sensor thickness is about $280 \mu m$.

The metallization, on the p-side, is not uniform but has small holes, coincident with the pixel cells, to let the light pass through. When the detector is illuminated both holes and electrons are produced. Holes are immediately collected on the p-side while electrons move through the silicon to be collected on the other side of the detector connected to the read-out chip. Thus the Lorentz angle measured is essentially due to the movement of electrons, see Fig. 5.13.

The magnetic field was oriented mainly in the y-direction while the electric field E was in the z-direction, see Fig. 5.14. The resulting Lorentz displacement is therefore in the x-direction where the pixel size is $50\mu m$ thus resulting in a better determination of the charge movement.

The signal position is computed using the center of gravity method:

$$\bar{x}(PH) = \frac{\sum PH_i x_i}{\sum PH_i} \quad (5.24)$$

where PH_i is the pulse height of the pixel i and x_i the x coordinate of the middle point of pixel i . An example of a profile obtained measuring the pulse height of every pixel is shown in Fig. 5.15

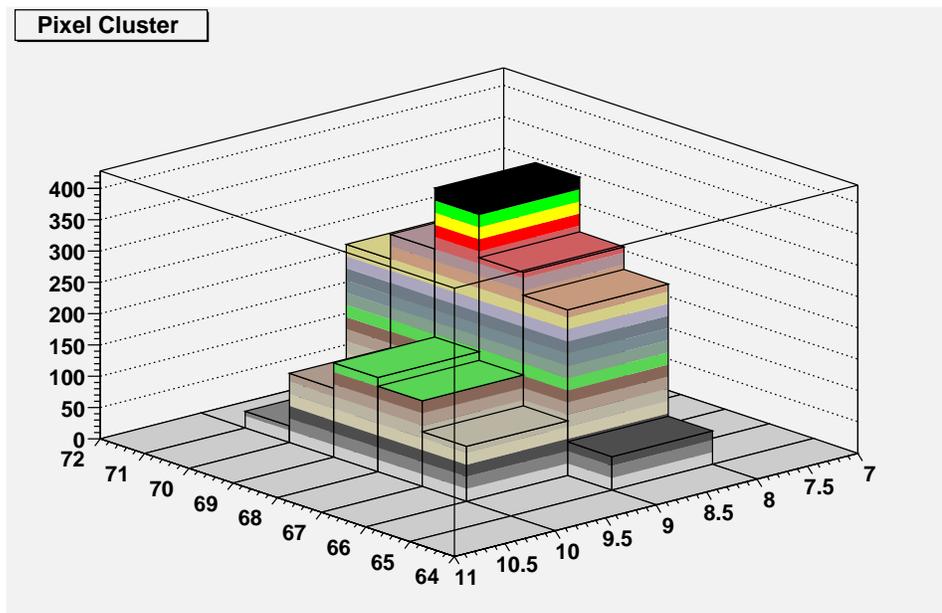


Figure 5.15: Example of profile of the charge collected by the pixel detector illuminated by the LED

5.4.2 Calibration

The pulse height of the pixel cells is obtained with a threshold scan. The FPIX1 read-out chip sets a common threshold to all the pixel cells, by an external voltage. This voltage is provided by an external power supply which can be driven by a GPIB interface. Since every readout channel has a characteristic level of the threshold a calibration must be performed in order to take into account the resulting different response of the pixels.

A relative calibration of the ADC response of each cell on the read-out chip has been performed by injecting charge into individual pixels, sending one hundred pulses to a calibration capacitor for each front-end channel at a fixed value of the pulse voltage in the pulser and at a fixed value of the threshold applied, and then counting the number of responses for that particular threshold. The result is a threshold scan curve which tells at which threshold level the pixel is sensitive for that particular voltage. This information can then be easily converted to the amount of charge needed to have a signal at that particular threshold. An example of threshold scan is given in Fig. 5.16.

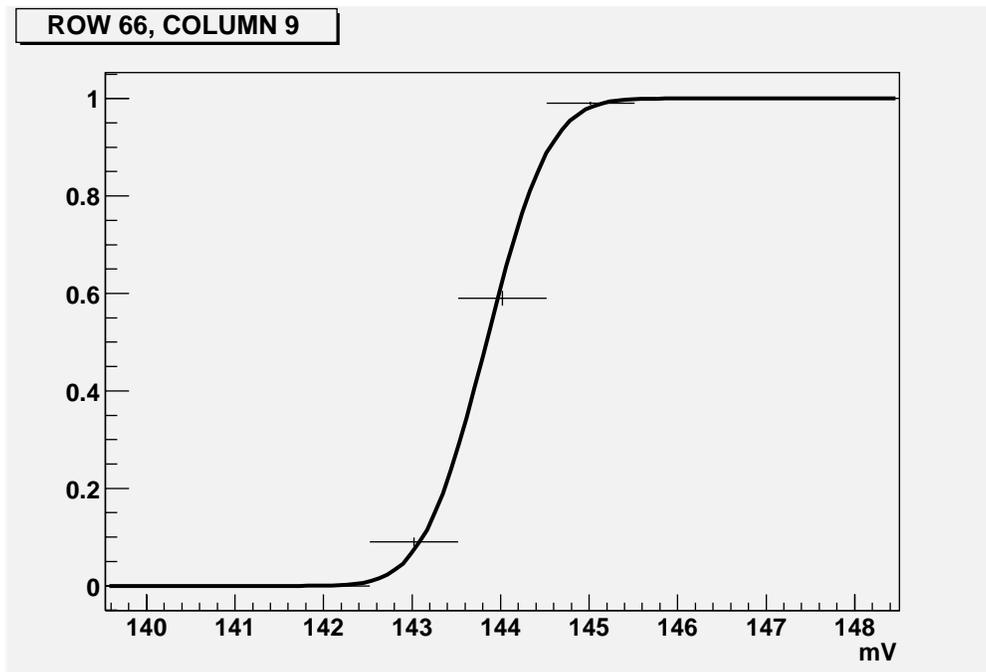


Figure 5.16: The figure shows the threshold scan performed for the cell at row 66 and column 9. The abscissa represents the obtained pulse voltage, expressed in mV, while the ordinate is the ratio between pulses read-out and pulses sent.

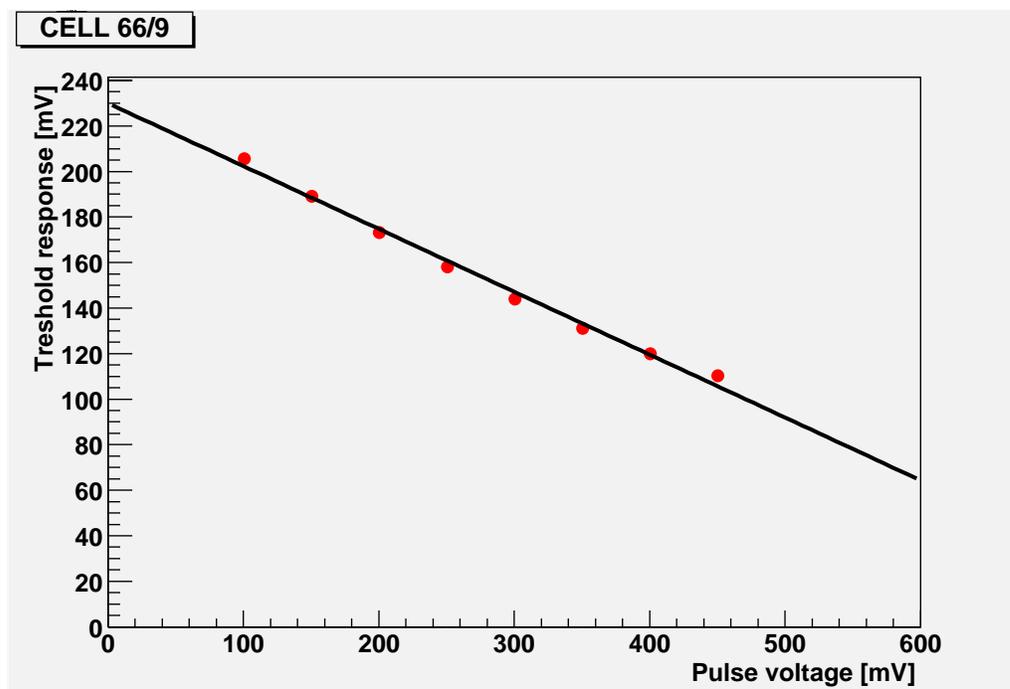


Figure 5.17: This histogram shows the linear relation between the applied pulse voltage (abscissa), and the threshold response (ordinate) for the cell 66/9.

The threshold response is linear with the amount of charge injected. The calibration threshold scan was done at different injecting voltages and then the results fitted with a straight line which provides the proportionality between the threshold response and the charge injected. For every pixel cell these lines provide the conversion factor between the threshold and the amount of charge collected. An example of this linear fit is shown in Fig. 5.17.

Once calibrated the cell response, I could then proceed to evaluate the center of gravity of the charge released in the detector when enlightened by the LED, using Eq. 5.24. If the light source remains fixed, relative to the detector, the center of gravity of the collected charge provides the best available information on the displacement of the charge collection induced by the magnetic field, and thus opens up the possibility of measuring the Lorentz angle, as depicted in Fig. 5.13.

5.4.3 Measurement

In general we would expect the center of gravity of the collected charge to be displaced by an amount linearly proportional to the applied external magnetic field.

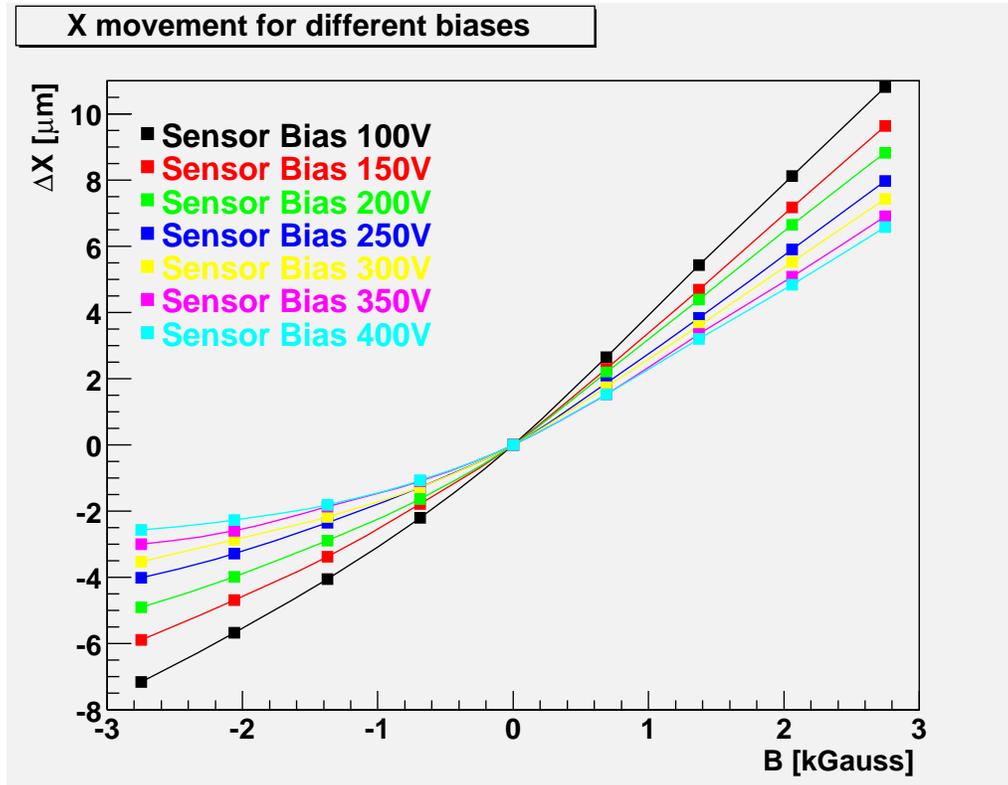


Figure 5.18: Displacement of the cluster center of gravity along the X axis

In particular, for a perfectly aligned sensor (see Fig. 5.14), the displacement should occur along the X axis only. In practice two effects contribute to alter this simple picture:

- the "mechanical" movement of the measurement device, induced by the magnetic field
- the imperfect alignment of the pixel sensor along the magnetic field lines

The resulting measured displacement along the X axis is then the sum of the Lorentz and the mechanical movements, thus resulting in a non linear displacement as shown in Fig. 5.18, while, along the Y axis, there is also a non null movement due to the two contributions described above, as shown in Fig. 5.19.

The experiment setup (described at paragraph 5.4.1) features, in fact, several mechanical components with non negligible sensitivity to a magnetic field, in particular the metallic rod that supports the LED focusing lens. In contrast to the charge carriers in the sensor, which are sensitive to the sign of the applied magnetic field,

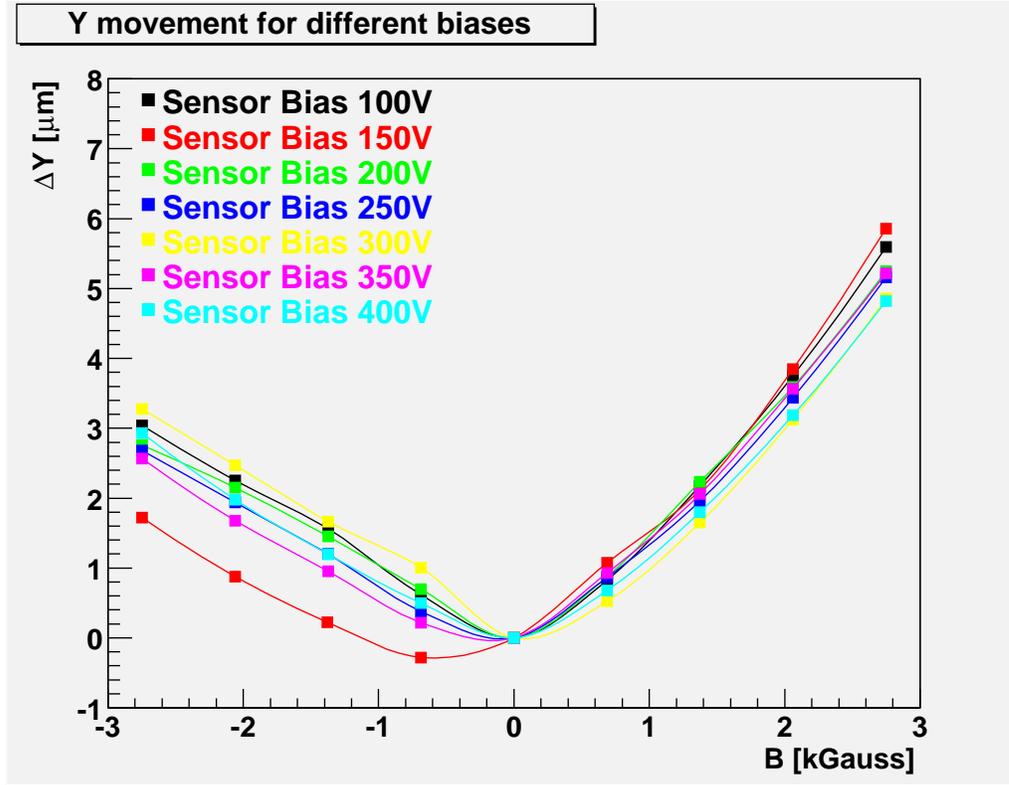


Figure 5.19: Displacement of the cluster center of gravity along the Y axis

the metallic components are only sensitive to the absolute value (modulus) of the field. We therefore expect a two-fold contribution to the displacement of the charge due to those effects combined. But, since the displacement due to the mechanical movement is an even function of the magnetic field (does not change sign reversing the magnetic field sign) while the charge displacement is described by an odd function, it becomes feasible, in principle, to isolate (or cancel) one contribution at time. Taking the semi-difference of the displacement measured at opposite values of the magnetic field allows to cancel the component due to the mechanical movement and thus isolate the pure effect of the Lorentz angle. This is possible assuming non hysteresis effect are present, which I verified to be the case by means of a repeated set of measurements taken at different values of the magnetic field. In summary:

$$\frac{(X_B + X_{-B})}{2} \text{ or } \frac{(Y_B + Y_{-B})}{2} \text{ leaves the pure mechanical movement} \quad (5.25)$$

$$\frac{(X_B - X_{-B})}{2} \text{ or } \frac{(Y_B - Y_{-B})}{2} \text{ leaves the pure Lorentz effect} \quad (5.26)$$

5.4.4 Data analysis

Mechanical movement

I took several data samples under a matrix of different experimental conditions:

- different values of the magnetic field
- different values of the bias voltage

If the hypothesis stated above about the effect on the charge displacement induced in part by a mechanical movement of the supporting structure holds true, one should observe that, once subtracted the Lorentz angle displacement contribution, the remaining displacement should be insensitive to the bias applied to the detector. In order to check this important point, I filled a set of histograms (one for each different applied voltage), see Fig. 5.20 and Fig. 5.21, with the semi-sum of the displacement at opposite values of the magnetic field. This has been done for the whole range of applied fields, from -3 kGauss to +3 kGauss. A combined fit of these plots with a single curve exhibits a good χ^2 (0.07 along the X direction and 1.14 along the Y direction) indicating a substantial insensitivity of the mechanical induced charge displacement from the applied voltage. This strongly supports the hypothesis that the effect of the mechanical-induced displacement can be correctly subtracted.

Lorentz displacement

Figures 5.22 and 5.23 show the results I obtain once the mechanical contribution to the charge displacement has been cancelled. First one should note that, after subtraction, the data nicely follow a linear dependence upon the applied magnetic field. Second, the data along the Y coordinate, where the field should be absent, also exhibit a residual displacement effect (whose linearity is less pronounced): this is due to the non perfect alignment of the sensors with respect to the field lines and to the fact that the metallic rod holding the focusing lens has a non negligible movement component along this coordinate. Along this coordinate, moreover, the sensor dimensions are so large that only two cells contribute to the measurement thus degrading the resolution of the measurement.

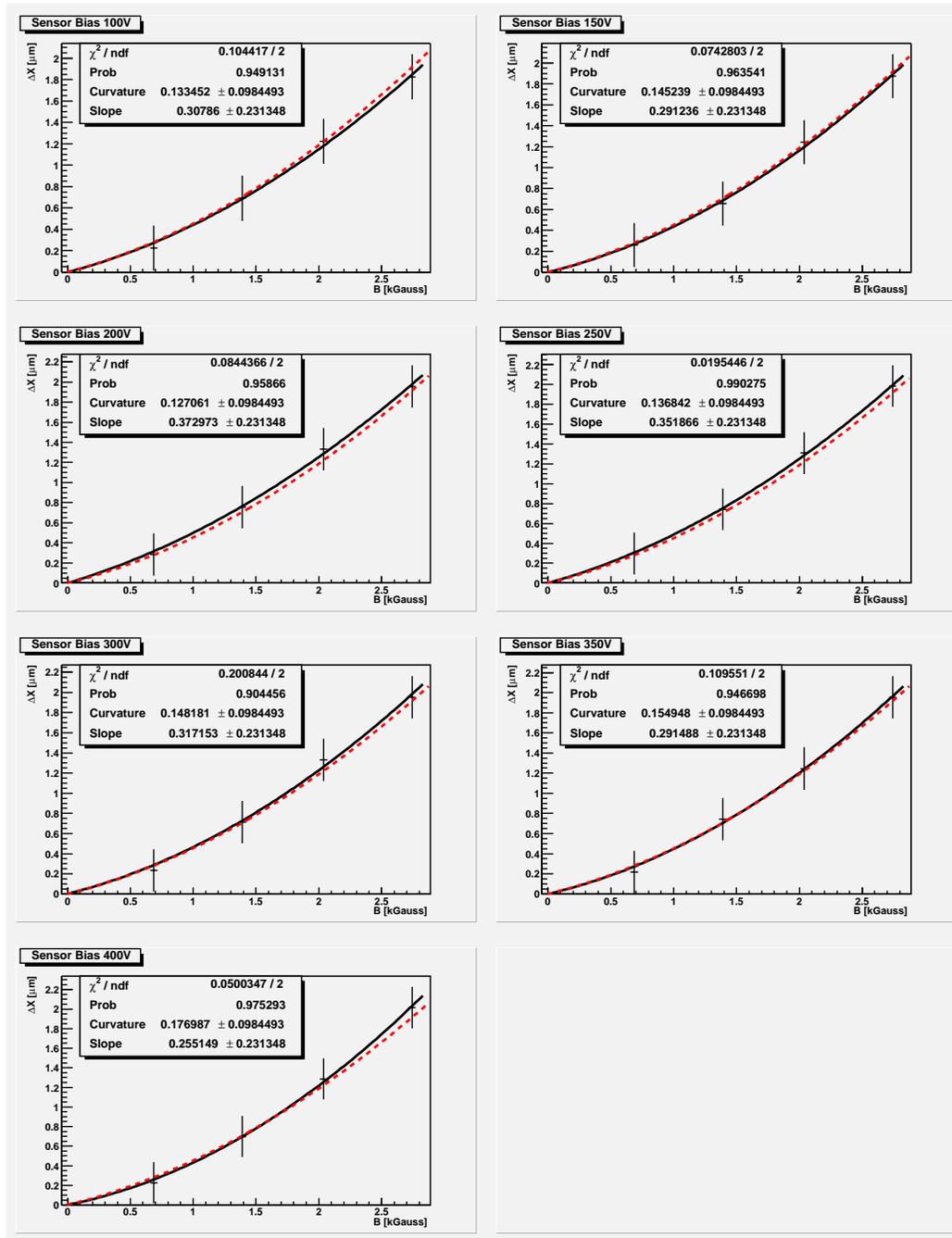


Figure 5.20: Displacement along the X axis obtained with the semi-sum of the measures taken at opposite values of the magnetic field for different biases applied to the sensor, thus effectively removing the Lorentz force contribution. The black lines are the single fits while the red dashed line is the common fit.

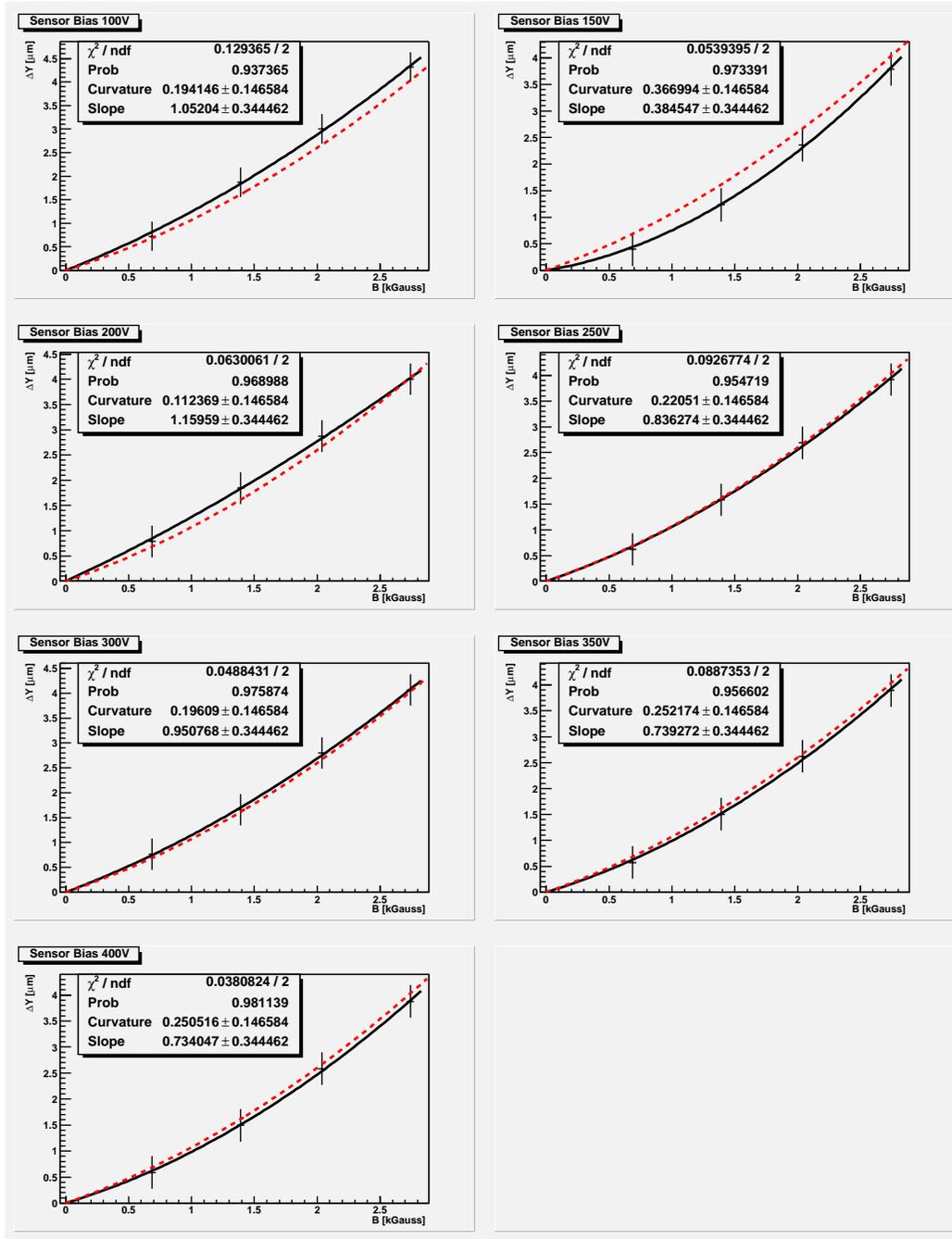


Figure 5.21: Displacement along the Y axis obtained with the semi-sum of the measures taken at opposite values of the magnetic field for different biases applied to the sensor, thus effectively removing the Lorentz force contribution. The black lines are the single fits while the red dashed line is the common fit.

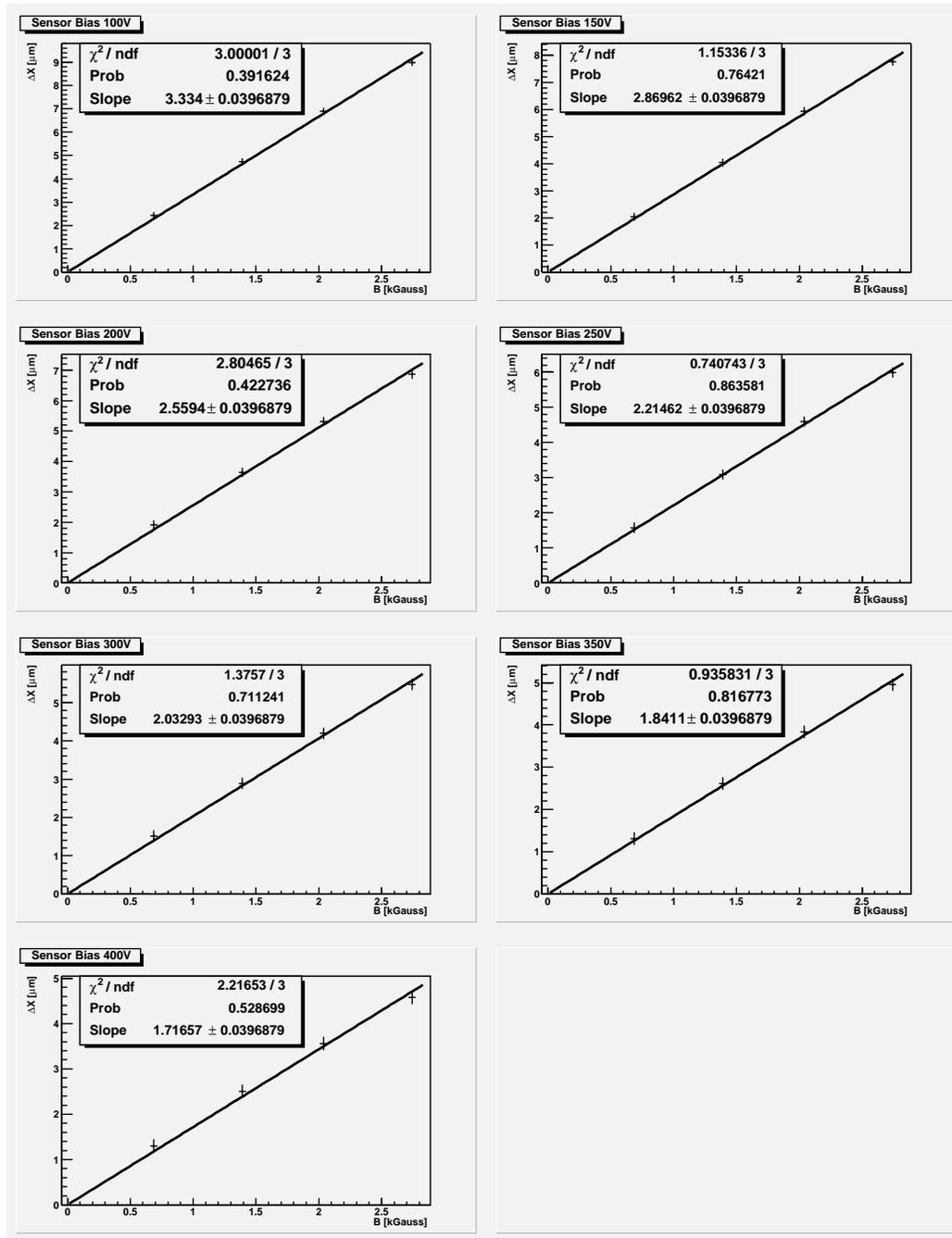


Figure 5.22: Lorentz displacement along the X axis obtained with the semi-difference of the measures taken at opposite values of the magnetic field for different biases applied to the sensor.

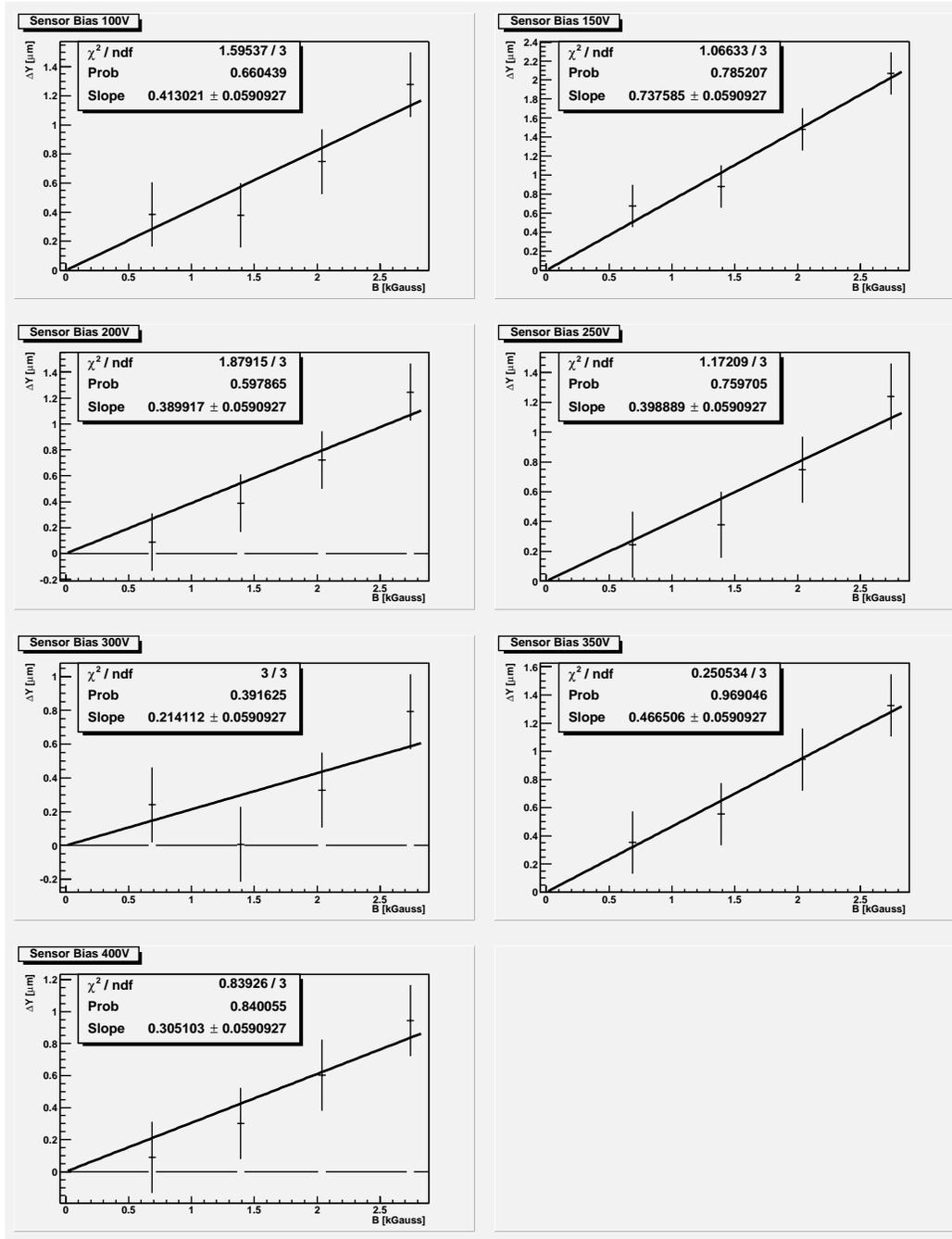


Figure 5.23: Lorentz displacement along the Y axis obtained with the semi-difference of the measures taken at opposite values of the magnetic field for different biases applied to the sensor.

Corrections

Since the Lorentz angle measurement depends upon measuring a charge displacement at different values of a magnetic field, and since each individual position is deduced by computing the charge cluster center of gravity from binned pixel cells, an indetermination is introduced in the measurement by the finite size of these cells. In order to evaluate the effect of this indetermination a detailed MonteCarlo simulation has been performed to compute a correction factor that takes into account this rounding off effect.

In Fig. 5.24 the charge distribution of a single measurement is represented both as raw data and as a superimposed gaussian fit. The true value of the beam spot

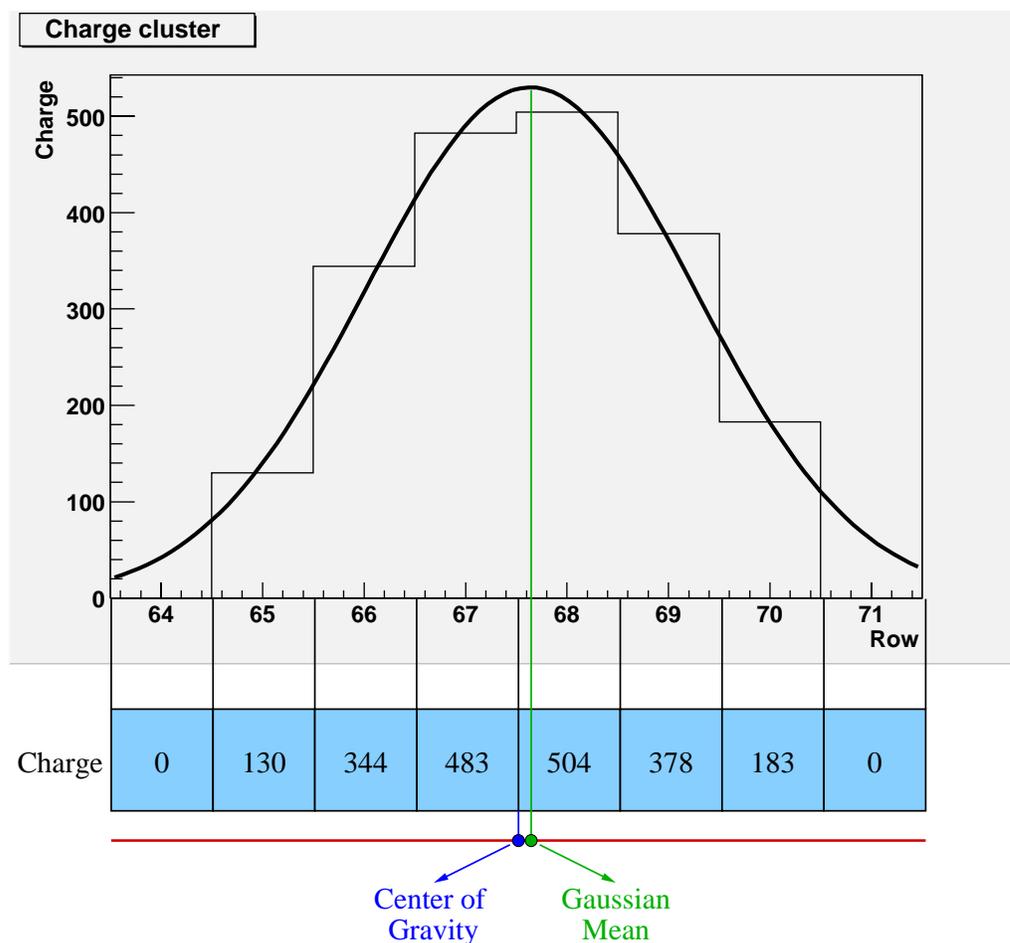


Figure 5.24: When the charge collected is shared only on few pixel cells the center of gravity calculated can be different from the mean value of the distribution.

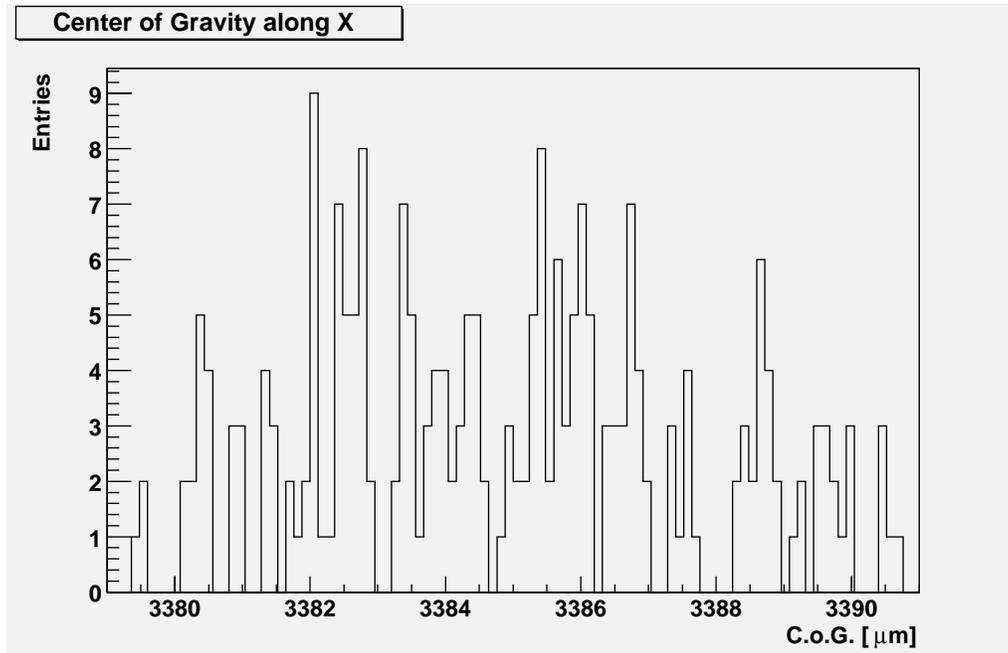


Figure 5.25: Distribution of the centers of gravity obtained by data

center, assuming it could be described by the peak of a perfect gaussian shape, is offset, with respect to the mean value of the charge distribution, by a small quantity introduced by rounding off the continuous charge distribution into finite bins represented by the pixel cells. In order to compute a correction function for this displacement, I did the following study:

1. I computed the mean value of the beam spot from the whole data sample along the X direction. Subsamples at different values of bias and magnetic field were all equally populated. The resulting distribution is shown in figure 5.25. The distribution, almost flat in a limited region around $3386\mu m$, is contained well within the boundaries of the pixel corresponding to raw 67.
2. I then proceeded by generating a sample of 10000 gaussians with fixed width and amplitude taken from the data sample shown in Fig. 5.24 and mean value distributed flat in the range obtained from figure 5.25.
3. For each chosen gaussian, the shape has been superimposed on a grid of pixels at fixed positions, and the fraction of gaussian area overlapping each bin has been computed. The mean value of these redistributed charges has then been computed and plotted in a distribution, see Fig. 5.26a.

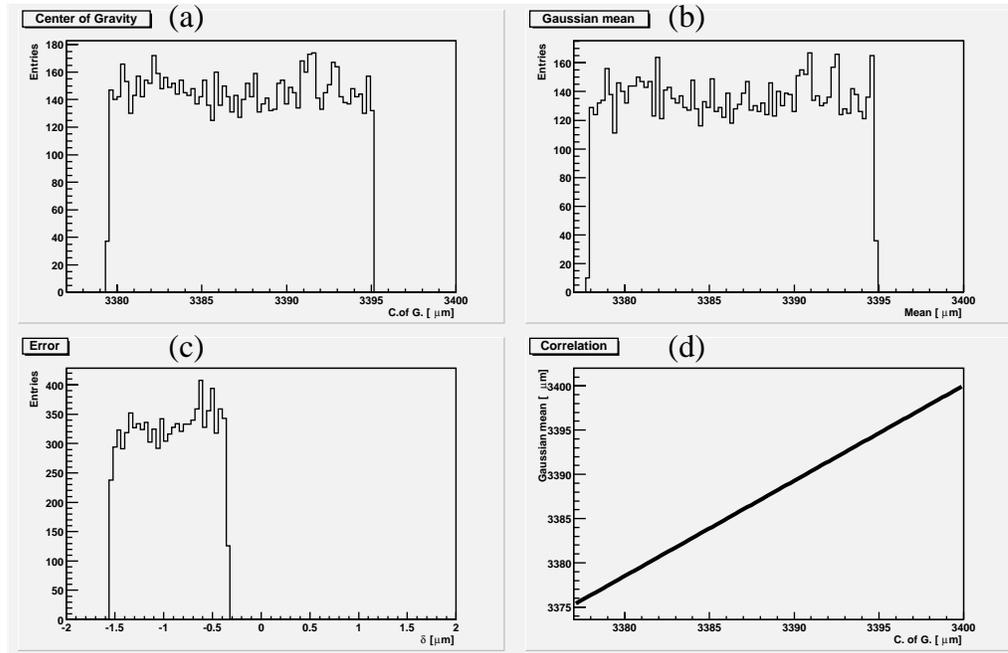


Figure 5.26: The top left histogram is a plot of the center of gravity distribution of the gaussian charge shape spread over six $50\mu\text{m}$ pixel cells. The top right histogram is a plot of the mean value of the gaussian shapes randomly chosen. The bottom left histogram is a plot of the difference between the mean value and the center of gravity of the pixel cells. The bottom right histogram, instead, shows the linear correlation between the center of gravity and the mean value of the gaussian.

4. Since the reference distribution in Fig. 5.25, taken from real data, displays mean values already rounded off from discretized pixels, the true position of the gaussian peak is in this case unknown. The MonteCarlo distribution obtained in point 3 is therefore somewhat different from the mother distribution by an amount presumably equal to the unknown displacement due to rounding off. To adjust for this small value, I iteratively shifted the range of input values to MonteCarlo generation, until the distribution obtained from simulation closely resembles the original data (see figure 5.26b).
5. For each generated gaussian, the difference between the input value (peak position of the gaussian) and the computed mean after discretization is plotted (Fig. 5.26c). The spread turns out of the order of $1\mu\text{m}$.
6. Finally I could obtain the correlation curve between continuous beam spot

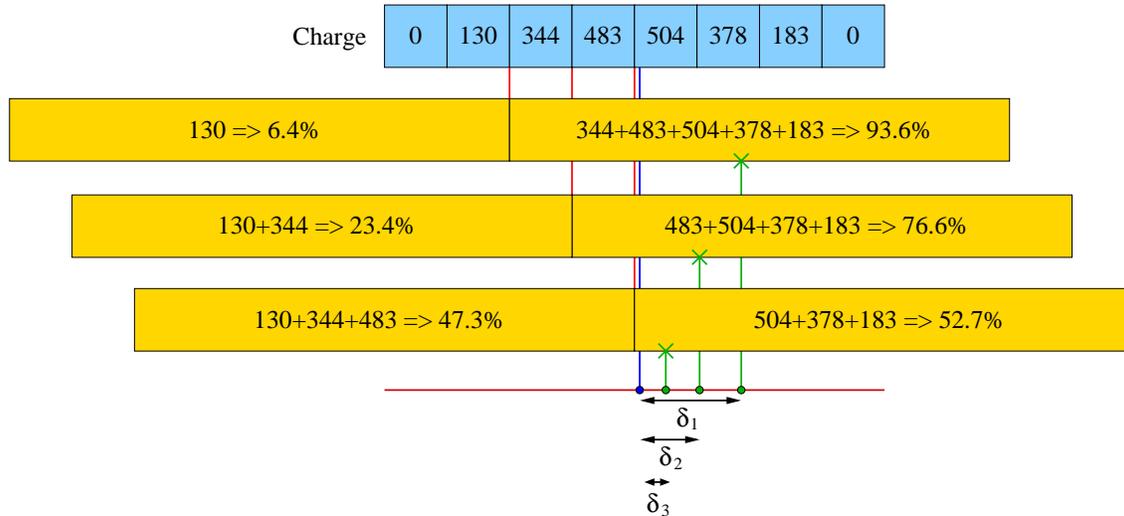


Figure 5.27: The charge of the 6 central pixel cells is distributed on two big cells only and the percentage of the charge collected in those cells is calculated. The mean position calculated with the fine granularity can be different to the one calculated with a coarse distribution. It is then possible to correlate the percentage of the charge with the error in the determination of the position δ .

input coordinates (mini Montecarlo input) and the corresponding computed values from discretized quantities, see Fig. 5.26d

The correction value to be applied to the data is extracted from this curve, used as a look up table to shift input data to corrected values. Once applied the correction, the difference between corrected and uncorrected data is shown in Fig. 5.28 for the mechanical movement and in Fig. 5.29 for the Lorentz displacement. The last figure shows that the correction is linearly proportional to the displacement, being larger for larger displacements. The overall correction turns out to be of the order of 7%.

A similar analysis has been performed for the coordinate parallel to the applied magnetic field. Here the problem is exacerbated by the very coarse granularity of the system: the whole laser beam spot is contained in just two pixel columns, thus discretization is maximal in this case. On the other side the displacement due to the Lorentz angle is negligible in the direction parallel to the magnetic field lines (of the order of less than $1\mu m$), thus the effect of the correction is small anyway. Assuming the shape of the beam spot equal along X and Y (which I verified by rotating the optical fiber by 90 degrees and by repeating the position measurement in this case),

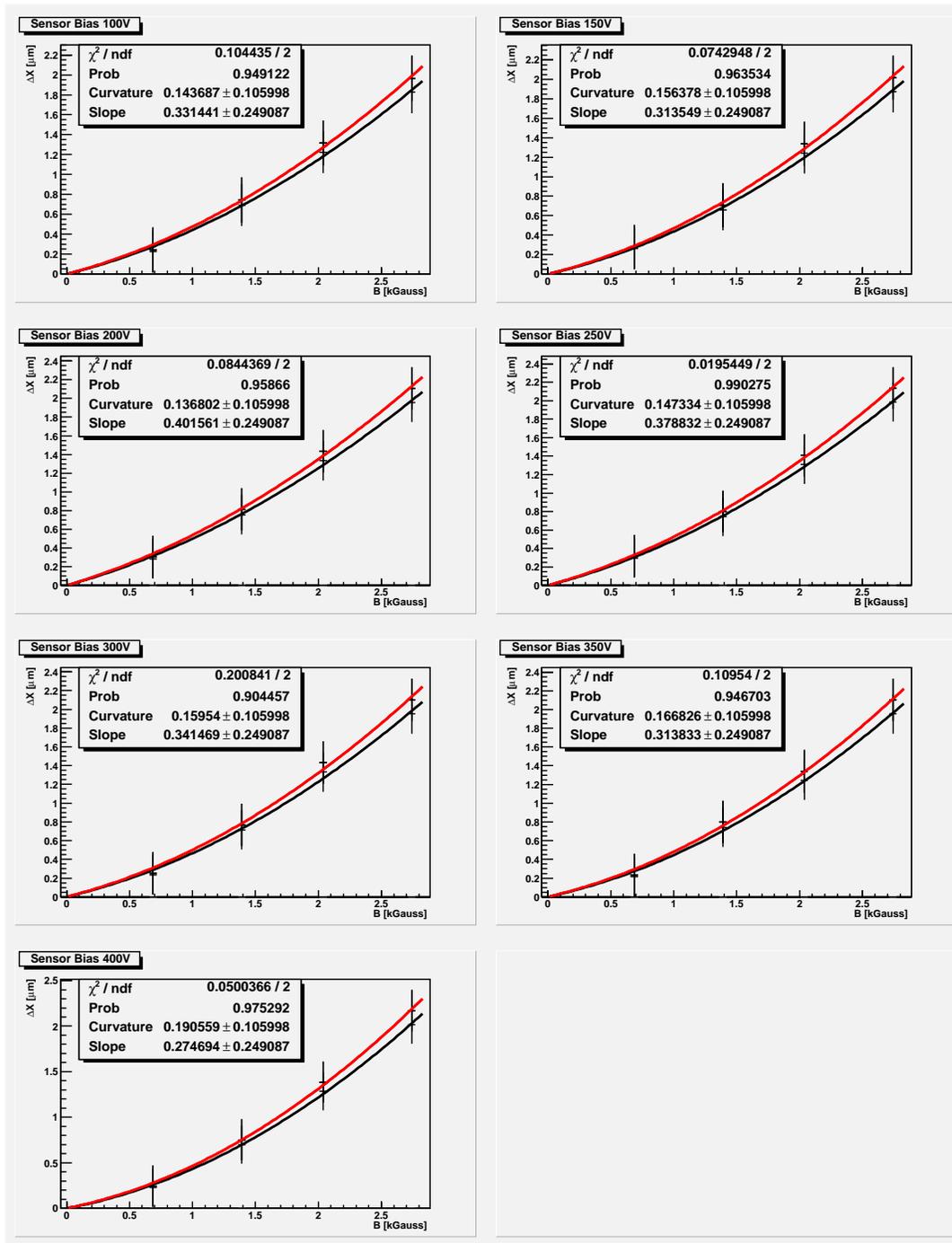


Figure 5.28: Comparison between the mechanical movement corrected (red line) and non corrected (black line) along the X axis

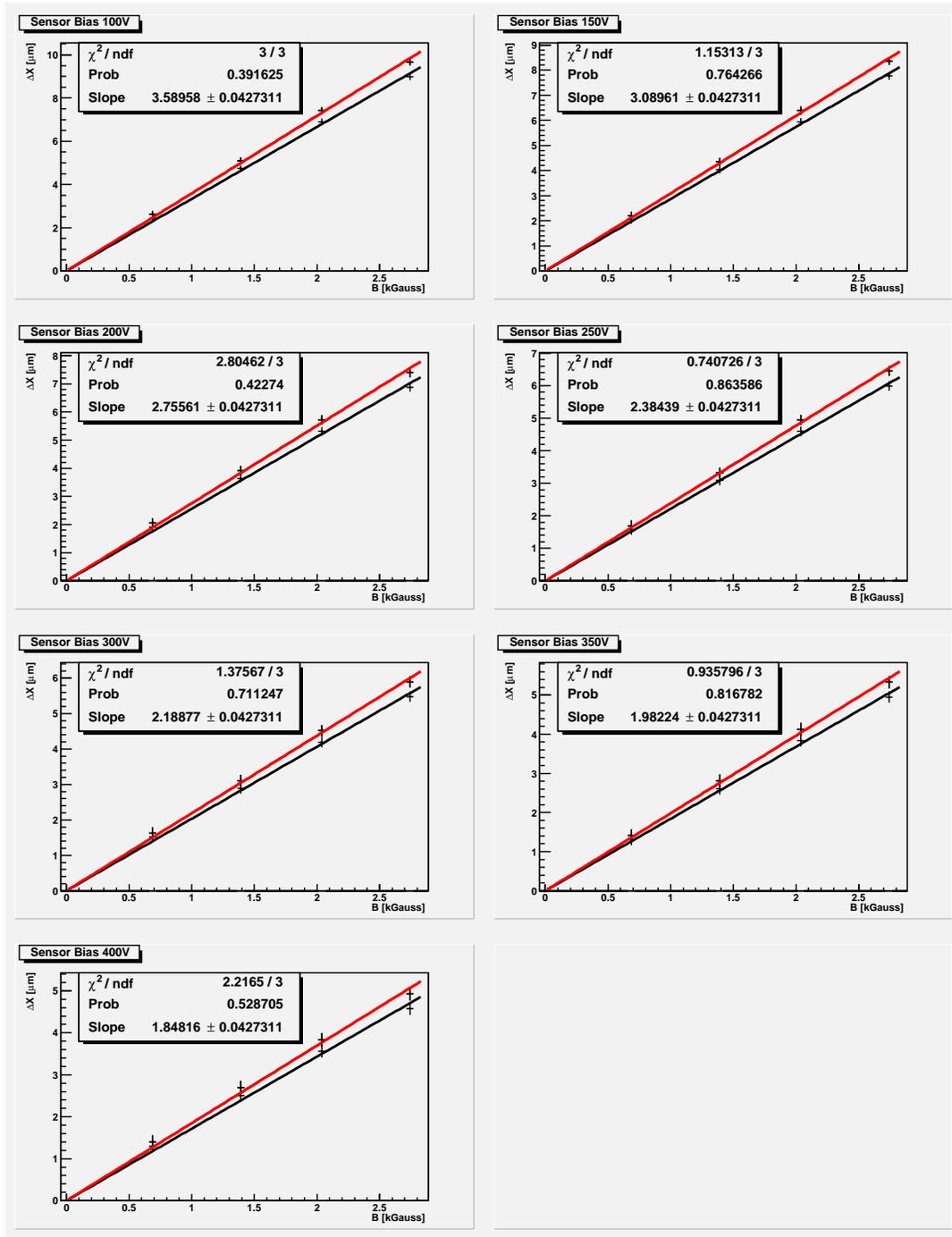


Figure 5.29: Comparison between the Lorentz displacement corrected (red line) and non corrected (black line) along the X axis. The histograms show clearly that the correction is linearly proportional to the displacement.

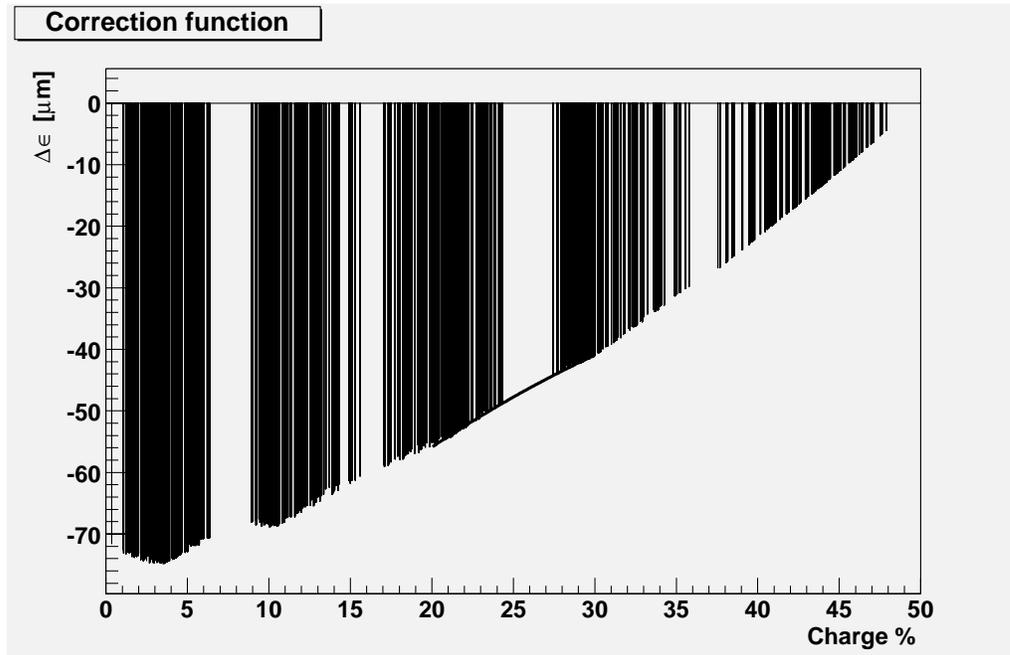


Figure 5.30: The abscissa represent the percentage of the charge in one $400\mu m$ column, while the ordinate is the difference between the center of gravity calculated with charge shared in the two columns and the one calculated distributing the charge over six $50\mu m$ cells.

one can determine the correction to apply to the data along this direction also. Following is a description of the algorithm used in this case which is also depicted in Fig. 5.27:

1. for each data, I redistributed on a grid of two $400\mu m$ cells, representing the two columns illuminated by the LED, the charge measured along the X axis, where the pixel size is $50\mu m$, moving the grid by steps of an integer pixel cell, obtaining for each step the percentage of the charge distributed on the two columns.
2. For each step I calculated the difference between the center of gravity obtained with the fine and the coarse distribution of the charge and then plotted this difference vs. the percentage of the charge collected in one of the two columns, as shown in Fig. 5.30.

The amount of charge released in the first of the two $400\mu m$ columns affected by the laser spot was consistently around 25% of the total. I therefore used a region of $\pm 5\%$

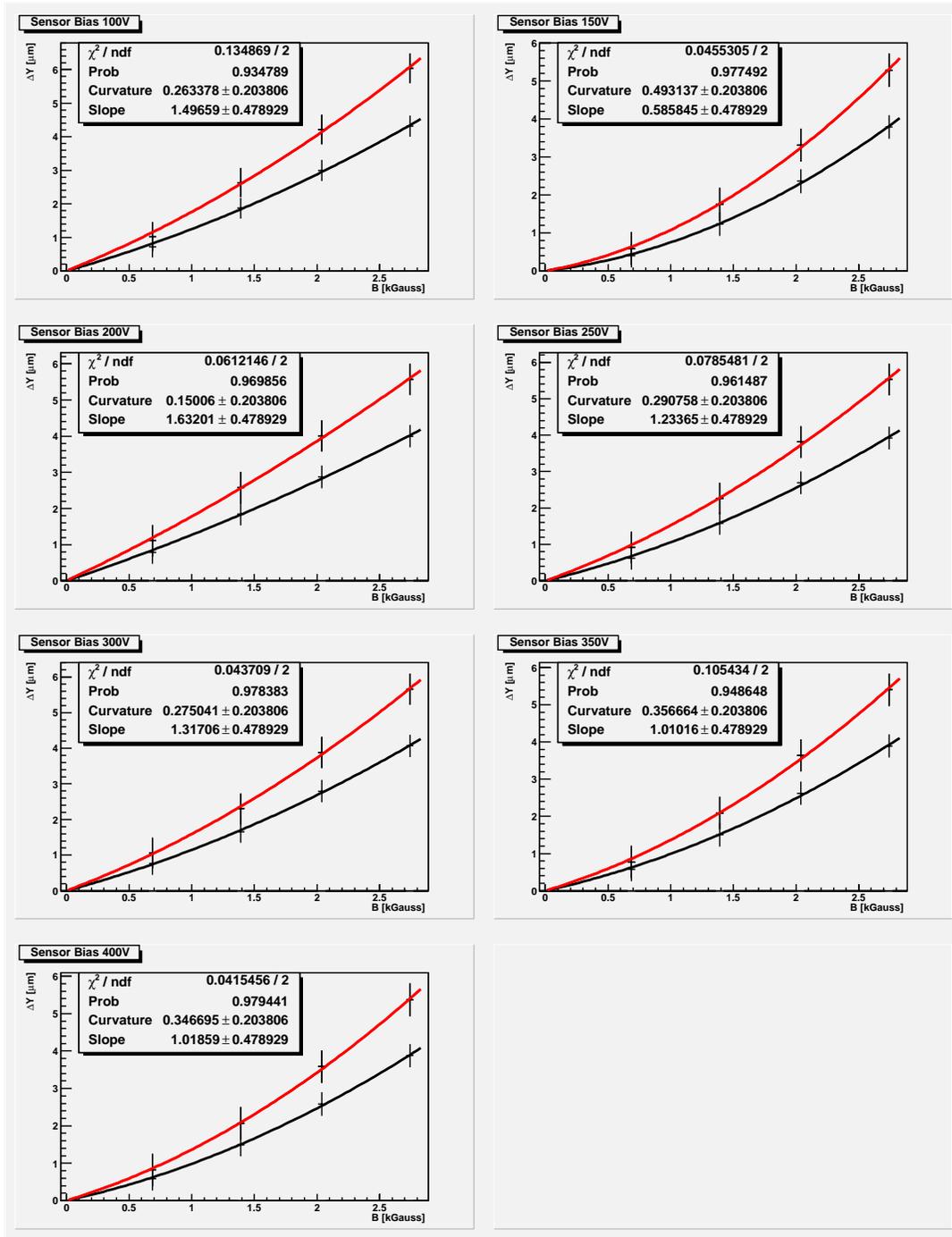


Figure 5.31: Comparison between the mechanical movement corrected (red line) and non corrected (black line) along the Y axis.

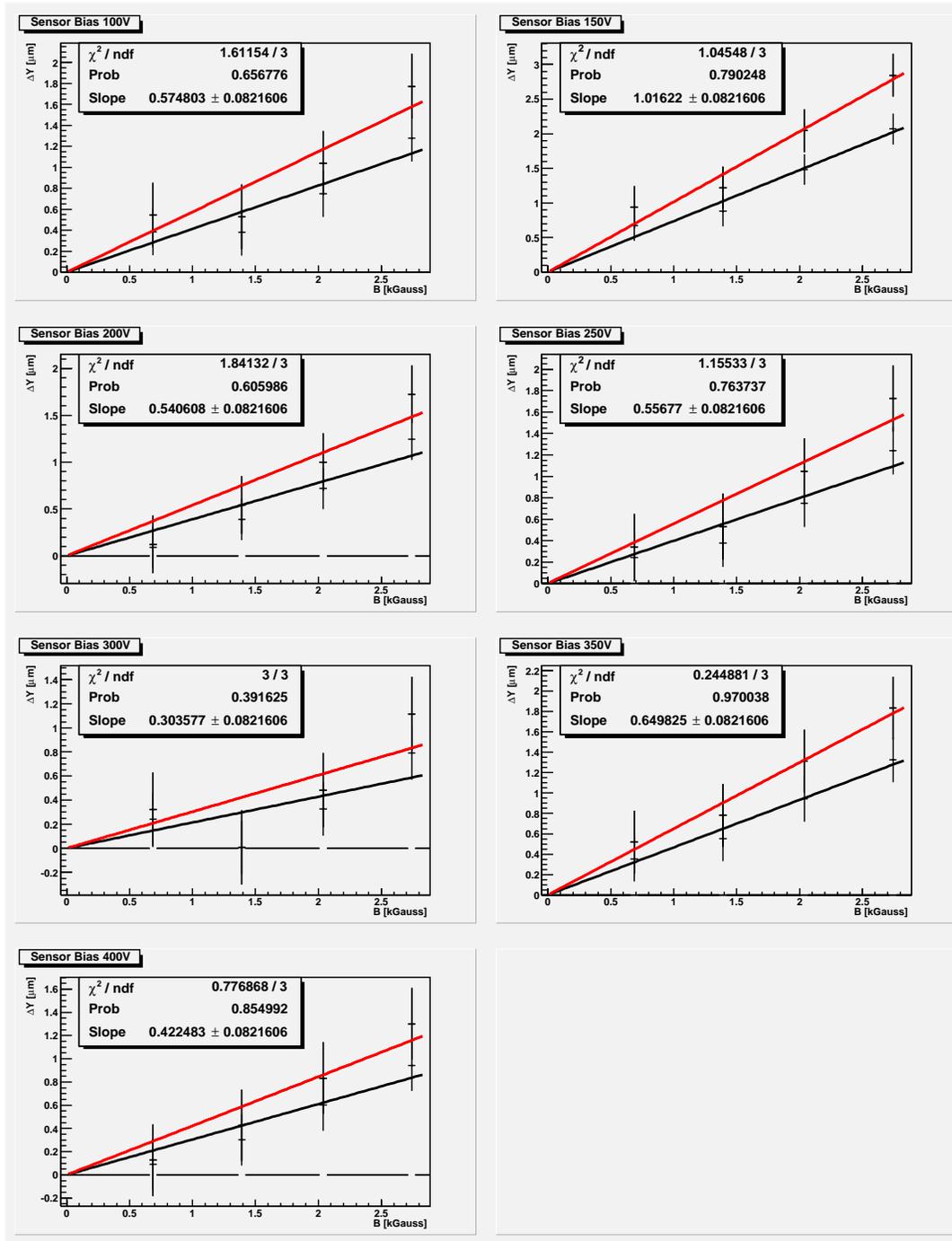


Figure 5.32: Comparison between the Lorentz displacement corrected (red line) and non corrected (black line) along the Y axis. On this axis the correction to the Lorentz displacement is of the order of half μm .

around this value to perform a fit which is then used as the correction function to adjust data along the Y direction. The results obtained applying this correction are shown in Fig. 5.31, for the mechanical movement, and in Fig. 5.32, for the Lorentz displacement.

Error determination

Curves at different voltages exhibit wide chi-square variations taking the mean value of the errors associated to each point, see Fig. 5.33. This is due to the incomplete knowledge of the statistical error associated to a single charge measurement. To improve the determination of this value, I have boosted the error associated to points of the worst fit (100 V for the X direction and 300 V for Y), in order to obtain a $\chi^2 \approx \mathcal{O}(1)$, see Fig 5.29 and Fig. 5.32. These values have then been used as the “standard” error of all the other plots.

Another source of error is the indetermination of the Magnetic field measured with an Hall probe. This indetermination is of the order of the 3%.

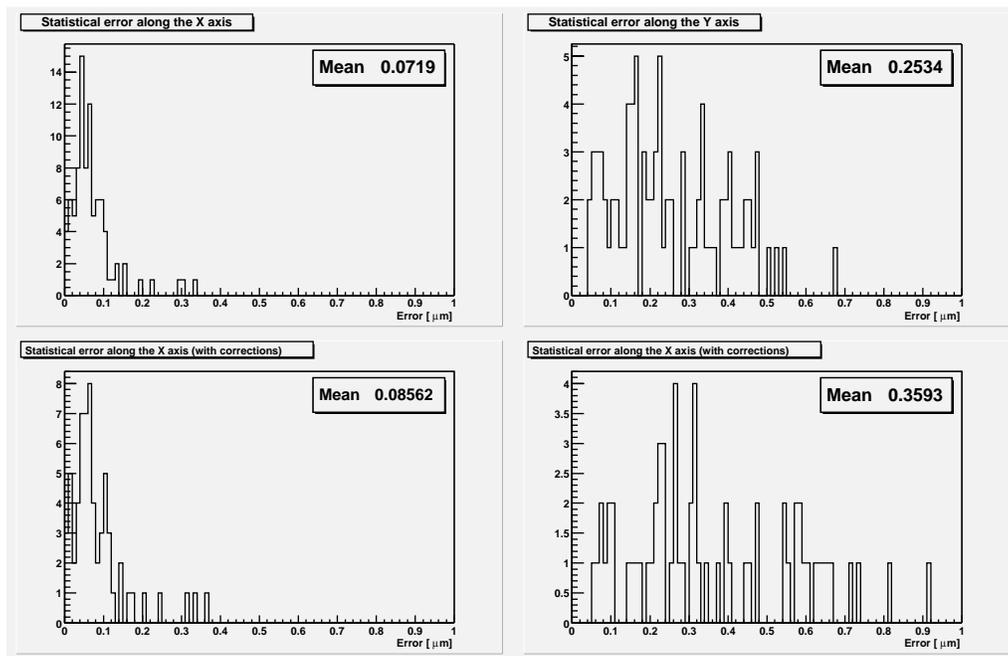


Figure 5.33: Distributions of the statistical errors associated with the measurements, for corrected and non corrected data both in X and Y.

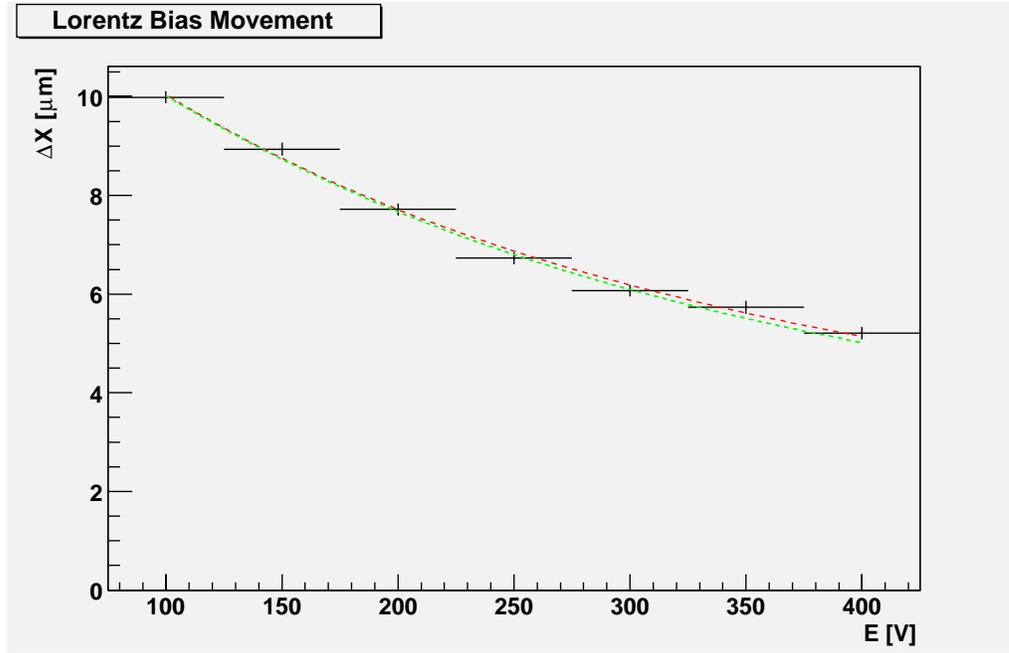


Figure 5.34: Lorentz displacement measured in the detector for the maximum magnetic field applied at different biases. The green line is the expected displacement predicted by theory using Eq. 5.28 while the points are the measured displacements for different biases.

5.4.5 Total Lorentz displacement

In the preceding discussion the displacements in the X and Y direction have been considered separately. The magnetic field is directed along the Y axis so the Lorentz displacement is expected along the X axis only but, as we previously saw, there is also a very small component along Y . This is probably caused by the imperfect alignment of the sensor with respect to the field. The total Lorentz displacement is then the sum in quadrature of the two displacements on the XY plane.

$$\Delta L = \sqrt{x^2 + y^2} = B \times \sqrt{\frac{x^2}{B^2} + \frac{y^2}{B^2}} = B \times \sqrt{m_x^2 + m_y^2} \quad (5.27)$$

The linear fits to the Lorentz displacement in the X (Fig. 5.29) and Y (Fig. 5.32) directions give the slopes used in Eq. 5.27. The associated error is obtained propagating the errors of each measurement in the two perpendicular directions. In Fig. 5.34 it is plotted the Lorentz displacement obtained for the different values of the applied electric field.

5.4.6 Comparison with theory

The setup I used for the measurement allows a straightforward comparison with theory. As we saw at paragraph 5.2.2, the mobility is well described by the empirical formula

$$\mu = \frac{\mu_0}{\left(1 + \left(\frac{E}{E_c}\right)^\beta\right)^{\frac{1}{\beta}}} \quad (5.28)$$

and from Eq. 5.23 is easy to calculate the displacement of a charge

$$\Delta x = \mu_H B d = r_H \mu B d \quad (5.29)$$

The Hall scattering factor is a constant term $r_H = 1.15$ for electrons, B is the value of the applied magnetic field and d is the thickness of the detector. The only non constant term is the mobility which depends upon the applied electric field which varies in the detector due to the space charge distribution. The distribution of the electric field in the detector is shown in Fig. 5.35 assuming a uniform distribution of the charge.

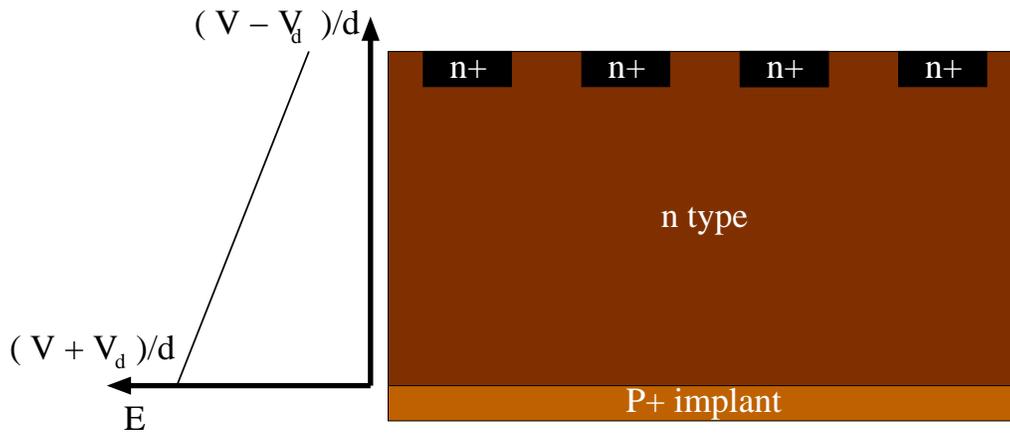


Figure 5.35: Electric field distribution in the detector.

For a charge drifting from the surface of the detector, which in our case corresponds to the whole cluster charge, it is very easy to predict the shifted position to compare with data. The only parameters needed in input are μ_0 , E_c and β . These parameters are determined empirically [41] looking for the best fit to the mobility data using formula 5.28.

These parameters have a temperature dependence which can be well approximated

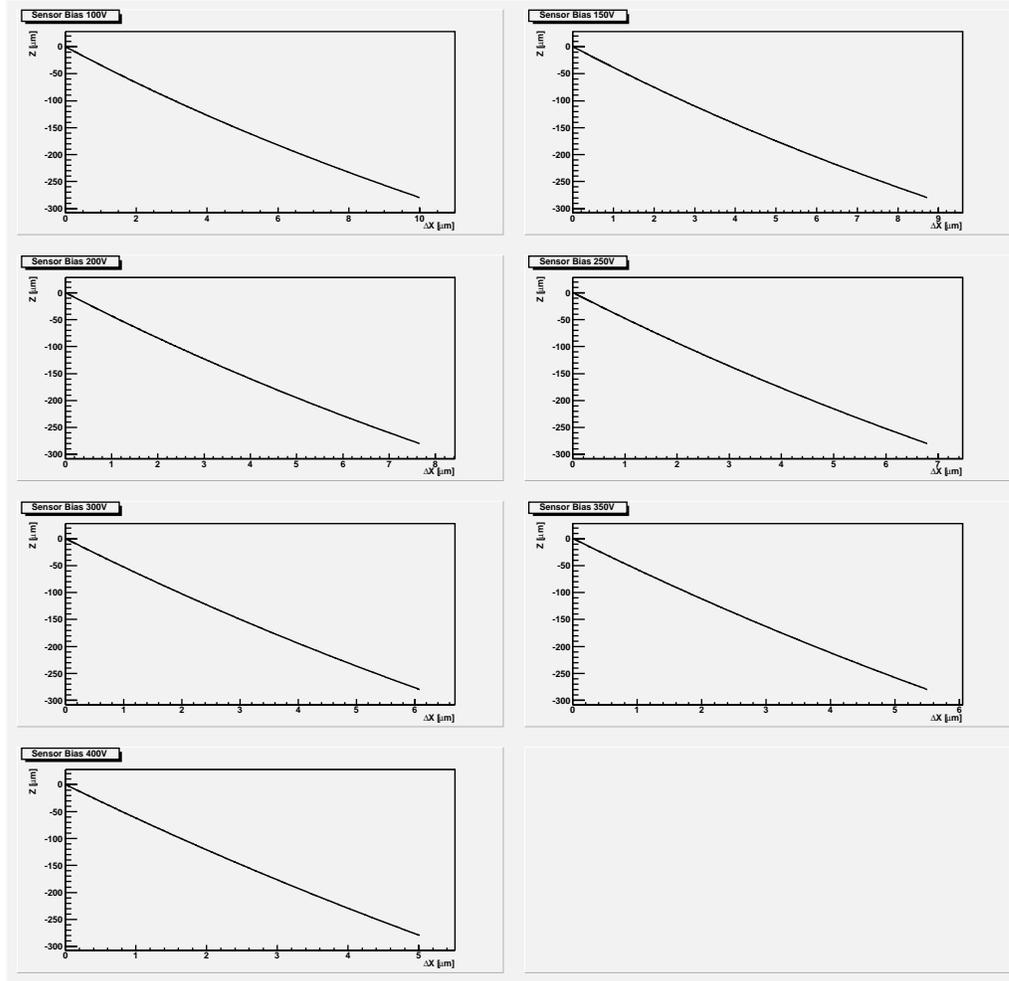


Figure 5.36: Displacement expected in μm of the charge carriers along the silicon detector due to the Lorentz force.

by a power law and for electrons the values quoted in literature are:

$$\mu_0 = 1.42 \cdot 10^9 \cdot T^{-2.42} \text{ cm}^2/V \cdot s \quad (= 1438 \text{ @ } 300^\circ K) \quad (5.30)$$

$$\beta = 2.57 \cdot 10^{-2} \cdot T^{0.66} \quad (= 1.11 \text{ @ } 300^\circ K) \quad (5.31)$$

$$E_c = 1.01 \cdot T^{1.55} \text{ V/cm} \quad (= 6980 \text{ @ } 300^\circ K) \quad (5.32)$$

These relations give just an approximation of the parameters values which are well measured at $T = 300^\circ K$. At room temperature the values which best fit these data available in literature are quoted as: have been found to be:

$$\mu_0 = 1450 \text{ cm}^2/V \cdot s \quad (5.33)$$

$$\beta = 1.30 \quad (5.34)$$

$$E_c = 7240 \text{ V/cm} \quad (5.35)$$

which are slightly different from those predicted by the equations 5.30, 5.31, 5.32. It is now easy to calculate the displacement of the electrons in the detector and the results are plotted in Fig. 5.36. The values obtained are in good agreement with the measured one, as can be seen in Fig. 5.34, where the green line is the displacement predicted with Eq. 5.29.

The red line is the fit to the data using the same equation but with those three parameters free. The values of the parameters obtained from a fit to our data are

$$\mu_0 = 1486 \pm 123 \text{ cm}^2/\text{V} \cdot \text{s} \quad (5.36)$$

$$\beta = 1.19 \pm 0.24 \quad (5.37)$$

$$E_c = 7706 \pm 550 \text{ V/cm} \quad (5.38)$$

The χ^2 is 1.35 for the fit to the data and 1.22 just using the phenomenological model with the parameters found in equations 5.33,5.34,5.35.

5.4.7 Results

In table 5.1 are summarized the results extrapolated at 1.5T, which will be the value of operation of the BTeV magnet. The total error is composed by the statistical error and the indetermination of the magnitude of the magnetic field.

Table 5.1: Summary of the Lorentz angle measurements extrapolated at 1.5T

BIAS	Theory	Measurement	Error
100V	11.18°	11.16°	$\pm 0.14^\circ \pm 0.3^\circ$
150V	9.75°	9.98°	$\pm 0.15^\circ \pm 0.3^\circ$
200V	8.56°	8.62°	$\pm 0.14^\circ \pm 0.3^\circ$
250V	7.59°	7.52°	$\pm 0.14^\circ \pm 0.3^\circ$
300V	6.80°	6.78°	$\pm 0.13^\circ \pm 0.3^\circ$
350V	6.15°	6.40°	$\pm 0.15^\circ \pm 0.3^\circ$
400V	5.60°	5.82°	$\pm 0.14^\circ \pm 0.3^\circ$

In Fig. 5.37 the effective Lorentz angle extrapolated at 1.5T is shown for different values of the biases applied. The green line is the angle predicted by the phenomenological model, while the red line is the fit to the data.

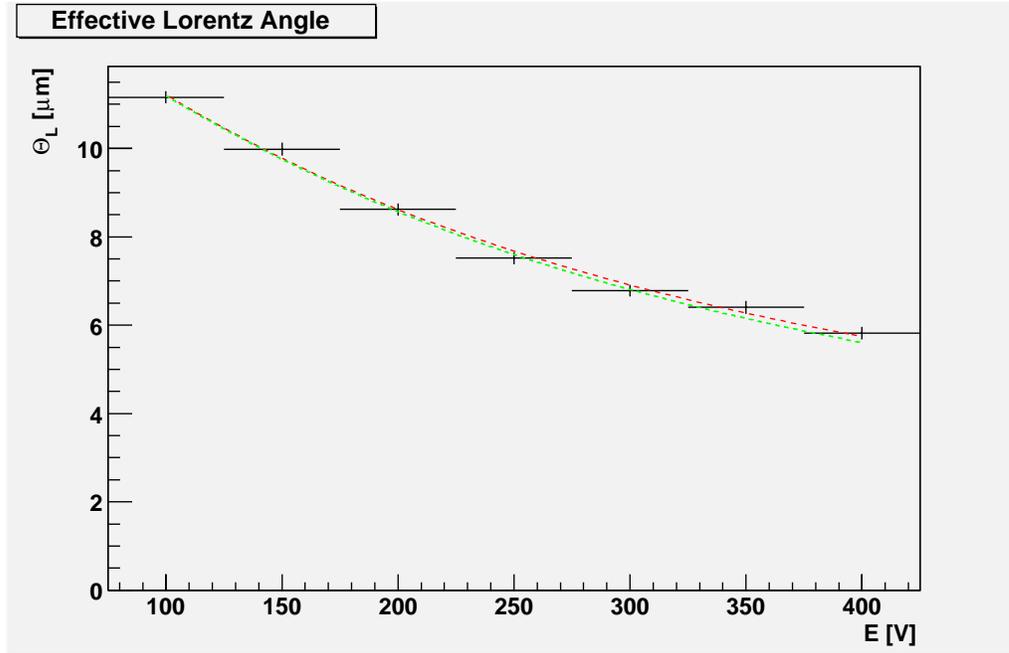


Figure 5.37: Effective Lorentz angle in the detector extrapolated for the 1.5T magnetic field present in BTeV. The green line is the theoretic displacement calculated for different biases using Eq. 5.28, while the red line is the fit to the data.

5.4.8 Comparison with other measurements

The results obtained are in good agreement with other measurements taken by CMS [45] and ATLAS [46]. The comparison for non-irradiated silicon detectors is shown in Tab. 5.2

Table 5.2: Comparison of the Lorentz angle measurements at 1.4T with a 150V bias.

ATLAS	CMS	BTeV
$9.0^\circ \pm 0.4^\circ \pm 0.5^\circ$	9.0°	$9.3^\circ \pm 0.14^\circ \pm 0.3^\circ$

5.4.9 Prospects

The simple experimental setup described in this chapter allows fast and accurate measurements of the Lorentz angle. At the full luminosity at which BTeV plans to operate, the innermost pixel detector will receive an equivalent fluence of $\approx 10^{14}$

minimum ionizing particles/ cm^2 /year. This will lead to radiation damage of the silicon pixel sensors. A study of the Lorentz angle for irradiated devices will be crucial to improve the accuracy of the vertex detector because the radiation damage is not uniform and the Lorentz angle depends upon the radiation damage [47]. The particle-induced lattice displacement throughout the entire crystal affect the mobility and thus the Lorentz angle. Different parts of the detector will then need a different correction. With this setup I will then be able to measure the relation between the Lorentz displacement and the irradiation dose.

Conclusions

During my three years of PHD I have been working in the BTeV experiment. My activity has concentrated on the development and characterization of the BTeV pixel detectors. I developed a completely new DAQ system, which will be used for the next pixel test beam at Fermilab, and, using this system, I measured the Lorentz angle in these detectors.

The originality of the new DAQ stems from the use of the PCI protocol, a standard in PC technology, which gives great advantages in portability, flexibility and costs when compared with other protocols widely used in High Energy Physics such as VME or CAMAC. The present DAQ features a completely data-driven architecture which does not require any external trigger to start the readout sequence. In this way, the pixel detectors will be tested on the beam in the real BTeV working conditions.

I have already demonstrated that both the hardware and software components are working properly by injecting charge with a pulser. The readout chip worked as expected and the DAQ was able to build events using the time-stamp information only, without the presence of any external trigger.

The measurement of the Lorentz angle was performed with a very simple apparatus and the DAQ I developed. The results are consistent with the theoretical expectations and improve the previous measurements obtained by ATLAS and CMS for non-irradiated sensors. The same experimental setup will be used in the next future to study the Lorentz angle in highly non-uniformly irradiated sensors. In this case, the Lorentz angle will depend on the radiation dose absorbed by a particular zone and, consequently, a different correction should be applied in order to reconstruct the right coordinate information. This study will be crucial to understand the performance of the pixel detectors in BTeV.

Bibliography

- [1] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, Phys. Rev. Lett. **13**, 138 (1964).
- [2] A. D. Sakharov, Pisma Zh. Eksp. Teor. Fiz. **5**, 32 (1967).
- [3] N. Cabibbo, Phys. Rev. Lett. **10**, 531 (1963).
- [4] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
- [5] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).
- [6] S. Stone, (1996).
- [7] C. Dib, I. Dunietz, F. J. Gilman, and Y. Nir, Phys. Rev. **D41**, 1522 (1990).
- [8] H. Harari and Y. Nir, Phys. Lett. **B195**, 586 (1987).
- [9] e. Harrison, P. F. and e. Quinn, Helen R., *The BaBar physics book: Physics at an asymmetric B factory*, Papers from Workshop on Physics at an Asymmetric B Factory (BaBar Collaboration Meeting), Rome, Italy, 11-14 Nov 1996, Princeton, NJ, 17-20 Mar 1997, Orsay, France, 16-19 Jun 1997 and Pasadena, CA, 22-24 Sep 1997.
- [10] G. C. Branco, L. Lavoura, and J. P. Silva, *CP violation*, Oxford, UK: Clarendon (1999) 511 p.
- [11] B. Aubert et al., Phys. Rev. **D66**, 032003 (2002).
- [12] K. Hagiwara et al., Physical Review D **66**, 010001+ (2002).
- [13] A. Hocker, H. Lacker, S. Laplace, and F. Le Diberder, Eur. Phys. J. **C21**, 225 (2001).
- [14] Y. Grossman, G. Isidori, and M. P. Worah, Phys. Rev. **D58**, 057504 (1998).

-
- [15] S. J. Richichi et al., Phys. Rev. Lett. **85**, 520 (2000).
- [16] S. E. Kopp, (1998).
- [17] B. Aubert et al., Phys. Rev. Lett. **87**, 151801 (2001).
- [18] A. E. Snyder and H. R. Quinn, Phys. Rev. **D48**, 2139 (1993).
- [19] D.-s. Du, I. Dunietz, and D.-d. Wu, Phys. Rev. **D34**, 3414 (1986).
- [20] R. Aleksan, I. Dunietz, and B. Kayser, Z. Phys. **C54**, 653 (1992).
- [21] R. Aleksan, A. Le Yaouanc, L. Oliver, O. Pene, and J. C. Raynal, Z. Phys. **C67**, 251 (1995).
- [22] D. Atwood, I. Dunietz, and A. Soni, Phys. Rev. Lett. **78**, 3257 (1997).
- [23] M. Gronau and D. Wyler, Phys. Lett. **B265**, 172 (1991).
- [24] F. Abe et al., Phys. Rev. Lett. **75**, 1451 (1995).
- [25] B. Abbott et al., Phys. Lett. **B487**, 264 (2000).
- [26] D. K. Fein, Given at 13th Topical Conference on Hadron Collider Physics, Mumbai, India, 14-20 Jan 1999.
- [27] M. L. Mangano, P. Nason, and G. Ridolfi, Nucl. Phys. **B373**, 295 (1992).
- [28] M. Artuso, HEPSY-94-1.
- [29] F. Abe et al., Phys. Rev. **D53**, 1051 (1996).
- [30] G. Chiodini et al., IEEE Trans. Nucl. Sci. **49**, 2895 (2002).
- [31] M. R. Coluccia et al., IEEE Trans. Nucl. Sci. **49**, 2914 (2002).
- [32] F. M. Newcomer, R. Van Berg, J. Van Der Spiegel, and H. H. Williams, Nucl. Instrum. Meth. **A283**, 806 (1989).
- [33] R. J. Yarema, G. W. Foster, J. Hoff, M. Sarraj, and T. Zimmerman, IEEE Trans. Nucl. Sci. **40**, 750 (1993).
- [34] A. N. Annenkov et al., Radiat. Meas. **29**, 27 (1998).
- [35] J. M. Butler et al., Nucl. Instrum. Meth. **A290**, 122 (1990).

-
- [36] M. H. L. S. Wang, Nucl. Instrum. Meth. **A501**, 214 (2001).
- [37] E. E. Gottschalk, Nucl. Instrum. Meth. **A473**, 167 (2001).
- [38] J. A. Appel et al., Nucl. Instrum. Meth. **A485**, 411 (2002).
- [39] D. K. Ferry, Fundamental aspects of hot electron phenomena, in *Handbook on Semiconductors*, edited by W. Paul, volume 1, North-Holland, Amsterdam, 1982.
- [40] D. M. Caughey and R. E. Thomas, Proc. IEEE **55**, 2192 (1967).
- [41] C. Canali, G. Majni, R. Minder, and G. Ottaviani, IEEE Trans. on Electron Devices **ED-22**, 1045 (1975).
- [42] A. Castoldi, E. Gatti, V. Manzari, and P. Rehak, Nucl. Instrum. Meth. **A399**, 227 (1997).
- [43] Landolt-Bornstein, *Numerical Data and Functional Relationships in Science and Technology, Group III*, volume 17a, Springer, Berlin, 1982.
- [44] R. A. Smith, *Semiconductors*, Cambridge University Press, Cambridge, 1968.
- [45] V. Bartsch et al., Nucl. Instrum. Meth. **A497**, 389 (2003).
- [46] M. Aleppo, Nucl. Instrum. Meth. **A465**, 108 (2000).
- [47] R. Kaufmann and B. Heinrich, CMSIN (1998/3).

Acknowledgements

When I decided to work for three more years with the same group, I didn't know that my life one day would have dramatically changed. Luigi Moroni, Dario Menasce, Daniele Pedrini and Silvano Sala gave me the chance to spend a few months at Fermilab and there I could work with a person who trusted me and gave me so much work, that I was forced to learn, learn and learn. His name is Gabriele Chiodini and he was a boss, the number one, a teacher and especially a friend. He gave me an important job which allowed me to work in collaboration with Stefano Magni. Most of the knowledge I have accumulated in these years comes from this second important person and, only for chronological order, he will be always the boss number two. They have been the best teachers I have ever met and I will remember them forever.

They gave me the chance to be known by the BTeV pixel group (Jeff Appel, Simon Kwan, David Christian, Selcuk Cihangir) which will give me the opportunity to make disasters while working for them.

Two important people lead me to this decision: my mother and Claudia. I think that my mother is tired of seeing me every day and she probably wants to spend her holidays in the coldest place on the Earth, Chicago. She has always given me good advice and I'm sure that she will not be wrong this time either. Claudia, instead, calls me only during the night, maybe because for her it is 6 p.m. or maybe because she knows that if I really want to rest, the easiest solution is to live together at the same meridian altitude. This is the easiest solution to my problems that will become the nightmare for the two best friends of mine: Marco and Claudio. Be ready to receive my calls during the night but also a good bottle of Mescal sometimes.

Luckily, in my future home, I will not have any problem if I want to decorate my floor with marble. I'm sure, in fact, that Brad Hall, a very clever engineer who decided to change completely job, will do that to me for free, hopefully. Thank you in advance for the new floor but mainly for all the help you gave me.

I don't know what the future will give me, but I know that Luigi, Dario, Daniele

and Silvano have always helped me and will always do. Working with Luigi is hard because he always wants the best from you but this is the only way to achieve results.