

Entwicklung von Objekt orientierten C++ Algorithmen zur b-Quark Identifikation

(Development of methods for b-quark identification in High Energy e^+e^- and
pp collisions)

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der
Rheinisch-Westfälischen Hochschule Aachen zur Erlangung des akademischen
Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von
Diplom Physiker Arno Heister
aus Düren-Birkesdorf

Berichter: Universitätsprofessor Dr. St.Schael
J.W.Rohlf, Professor of Physics

Tag der mündlichen Prüfung 16.November 2007

Diese Dissertation ist auf den Internetseiten der Hochschulbibliothek
verfügbar.

Contents

List of Figures	v
List of Tables	vii
Introduction	1
1 The Standard Model: Basic concepts and their application	3
1.1 Overview	3
1.1.1 Basic concepts	3
1.1.2 Jets	5
1.1.2.1 The Jade jet algorithm	5
1.1.2.2 The inclusive k_{\perp} jet algorithm	6
1.1.2.3 Jet recombination schemes	6
1.1.3 Identification of b-hadrons	7
1.2 Testing the Standard Model: R_b	11
1.2.1 The measurement of R_b at the Z^0 resonance	13
1.2.2 Past and recent R_b measurements	14
1.3 Searching the missing link of the Standard Model: the Higgs	16
1.3.1 Status of the Standard Model Higgs search	16
1.3.2 Standard Model Higgs searches in pp collisions	18
1.3.2.1 Associated Higgs production in the $W^{\pm}H$ channel	19
2 e^+e^- physics	21
2.1 Introduction	21
2.2 The LEP electron-positron collider	23
2.3 The ALEPH experiment at LEP	24
2.4 The investigation of a new b-tag for ALEPH	27
2.4.1 Introduction	27
2.4.2 The benchmark b-tag: the ALEPH lifetime mass tag	28
2.4.3 Event selection and reconstruction	28
2.4.3.1 The hadronic Z^0 -decay and track selection	28
2.4.3.2 Jet reconstruction	29
2.4.3.3 Primary vertex reconstruction	29
2.4.3.4 Secondary vertex reconstruction	31
2.4.3.5 Association of tracks with the secondary vertex	32
2.4.3.6 Significance of the signed 3D impact parameter	33
2.4.4 Description and performance of the new ALEPH b-tag	34
2.4.4.1 Observables used in the new ALEPH b-tag	34
2.4.4.2 Combination of the observables used in the new ALEPH b-tag	35
2.4.5 Summary	38

3	Hadron collider physics	41
3.1	Introduction	41
3.2	The LHC pp collider	42
3.3	The CMS experiment at LHC	44
3.3.1	The magnet	45
3.3.2	The inner tracking system	45
3.3.3	The muon system	48
3.3.4	The calorimeter system	49
3.3.5	The trigger system	51
3.3.6	The luminosity measurement	52
3.4	b-hadron identification in CMS	53
3.4.1	Full simulation of CMS events	53
3.4.2	Charged track reconstruction	54
3.4.2.1	The track reconstruction in the inner tracker	55
3.4.2.2	The muon reconstruction	57
3.4.2.3	The electron reconstruction	58
3.4.3	Vertex reconstruction	60
3.4.3.1	Primary vertex reconstruction	60
3.4.3.2	Secondary Vertex reconstruction	62
3.4.3.3	Association of tracks with the secondary vertex	62
3.4.3.4	Vertex reconstruction from pixel tracks and signal vertex tag	62
3.4.4	Missing energy reconstruction	64
3.4.5	Jet reconstruction	65
3.4.5.1	Jet input: split UE TrackTowers	65
3.4.5.2	Jet algorithm: inclusive k_{\perp} jet algorithm, E_T scheme	67
3.4.5.3	Jet calibration: 2-dimensional detector map	69
3.4.6	b-hadron identification	72
3.4.6.1	Significance of the signed 3D impact parameter	72
3.4.6.2	Observables used in the new combined b-tag and its performance	76
3.4.7	Physics Analysis of WH, $H \rightarrow b\bar{b}$	79
3.4.7.1	L1 and HLT trigger selection	79
3.4.7.2	W mass constraint	79
3.4.7.3	Further constraints to improve the event selection	80
3.4.7.4	Higgs mass reconstruction performance	83
3.4.7.5	Results	83
3.4.7.6	Comparison with the ATLAS Physics TDR result	84
3.4.8	Summary	85
4	Conclusions	87
A	Analytical solution of the double tag method equations	89
B	The object-oriented b-tag framework	91
C	The UE curve for the CMS low-luminosity phase	93
D	The 2-dimensional detector map jet correction function	95
	Bibliography	97

List of Figures

1.1	Topology of a b-hadron decay and the layout of the 3D impact parameter.	8
1.2	Feynman graph at Born level of the weak interaction $e^+e^- \rightarrow f\bar{f}$	11
1.3	Feynman graphs of the two dominant $Z^0 b\bar{b}$ vertex corrections, which introduce the top mass dependence.	12
1.4	Feynman graph of the dominant top quark contribution to the Z^0 vacuum polarization.	12
1.5	R_b analysis results from 1993 till 1997 [CERN97].	15
1.6	Recent R_b measurements and the world average value [EWWG05].	15
1.7	Dependency of $\Delta\chi^2$ versus Higgs boson mass m_H [EWWG06].	17
1.8	The 18 input measurements included in the Standard Model fit [EWWG06].	17
1.9	Cross sections and branching ratios of the Standard Model Higgs [SZ97, DKS98].	18
1.10	Main gluon fusion process incorporating a top quark at Born level.	18
1.11	Feynman graph at Born level of the $W^\pm H$ production.	19
2.1	LEP experimental setup.	22
2.2	ALEPH detector.	24
2.3	ALEPH event display showing a typical $Z^0 \rightarrow b\bar{b}$ decay.	27
2.4	ALEPH lifetime mass tag $P_{H, \text{mass}}$	28
2.5	ALEPH hemisphere primary vertex and b-hadron jet reconstruction performance.	30
2.6	Significance $S_{2nd vtx}$ of reconstructed secondary vertices for ALEPH.	31
2.7	The S_{vtd} observable for tracks which pass the ALEPH QIPBTAG track selection.	33
2.8	ALEPH signed 3D impact parameter significance ($\frac{D}{\sigma}$).	33
2.9	Distributions of P_{jet} , $M_{2nd vtx}$, $\eta_{2nd vtx, track, jet}$ and X_{energy} for ALEPH.	34
2.10	P_{tag} , PM_{tag} and PMR_{tag} b-tag distributions for ALEPH.	36
2.11	$PMRX_{tag}$ b-tag distribution for ALEPH.	37
2.12	Comparison of the b-purity versus b-efficiency for the different b-tags for ALEPH.	37
2.13	Comparison of working points of recent R_b measurements.	38
2.14	R_b versus b-efficiency ϵ_b taken from data only for the $PMRX_{tag}$	39
3.1	LHC experimental setup.	43
3.2	CMS detector.	45
3.3	Layout of the CMS inner tracker.	46
3.4	Layout of the CMS pixel vertex detector.	46
3.5	r - z view of the CMS muon system.	48
3.6	r - z view of the CMS calorimetry geometrical configuration.	49
3.7	r - z view of the CMS ECAL geometrical configuration.	50
3.8	r - z view of the CMS HCAL barrel and endcap region.	51
3.9	CMS track reconstruction performance.	56
3.10	CMS muon reconstruction performance.	58
3.11	CMS electron reconstruction performance.	59
3.12	CMS primary vertex reconstruction performance.	61

3.13	Significance $S_{2nd vtx}$ of reconstructed secondary vertices for CMS.	61
3.14	CMS pixel vertex reconstruction and signal vertex selection for CMS.	63
3.15	Missing energy $E_{x, MET}$ - and $E_{y, MET}$ -resolutions for CMS.	64
3.16	CMS UE curve characteristics versus i - η and η	66
3.17	Jet reconstruction efficiency and fake rate for CMS.	68
3.18	$P_{t, jet}$ -, η_{jet} - and ϕ_{jet} -resolutions of uncorrected jets for CMS.	68
3.19	Profile histograms for uncorrected jets for CMS.	69
3.20	2D detector map correction for CMS.	70
3.21	Profile histograms for corrected jets for CMS.	70
3.22	$P_{t, jet}$ -, η_{jet} - and ϕ_{jet} -resolutions of corrected jets for CMS.	71
3.23	Typical complex LHC event topology for the low-luminosity phase.	72
3.24	Cleaned up LHC event topology to extract the b-hadron lifetime information. . .	73
3.25	η - ϕ distance of reconstructed jets and tracks for CMS.	74
3.26	$D_{3D, track, vtx}$, S_j , D and $\sigma(D)$ for CMS.	74
3.27	Significance of the signed 3D impact parameter $\frac{D}{\sigma}$ for CMS.	75
3.28	Jet lifetime probability tag P_{jet} and its performance for CMS.	76
3.29	Distributions of P_{jet} , $M_{2nd vtx}$, $\eta_{2nd vtx, track, jet}$ and X_{energy} for CMS.	77
3.30	$PMRX_{tag}$ observable and its performance for CMS.	77
3.31	Transverse W mass and E/P distributions for CMS.	80
3.32	Fraction of accepted events in the WH analysis for CMS.	81
3.33	Reconstructed $H \rightarrow b\bar{b}$ mass resolution for CMS.	82
3.34	Reconstructed WH Higgs mass for signal events and main the backgrounds. . . .	83

List of Tables

1.1	Gauge bosons of the Standard Model with corresponding coupling constants. . .	4
1.2	Electroweak quantum numbers Q , I , I_3 and Y for the particles of the SM. . . .	4
2.1	Parameters of LEP's luminous regions [B ⁺ 97].	23
2.2	Main parameters used for the ALEPH secondary vertex reconstruction.	32
3.1	Parameters of the LHC interaction area [LHC95].	44
3.2	Main design parameters of the CMS superconducting solenoid [CMS97c].	45
3.3	Estimated resolution for high P_t (single muon) tracks for the five track parameters: P_t , ϕ , $\cot(\theta)$, d_0 and z_0 [CMS98, CMS05a]	47
3.4	Design goals and the results from full simulation of the global track reconstruction efficiency for muons and charged hadrons (pions) [CMS98, CMS05a].	47
3.5	CMS muon track reconstruction performance [CMS05e].	49
3.6	Contributions to the CMS ECAL energy resolution [CMS97a, CMS06a].	50
3.7	Details about the simulated events used in the CMS study.	53
3.8	Listing of the Monte Carlo events used in the CMS study.	54
3.9	Optimized parameters for track seed finding in the CMS inner tracking [Win02].	55
3.10	Track constraints used for the CMS inner tracking performance [CMS02b]. . . .	57
3.11	Summary of specific track constraints to calculate $\frac{D}{\sigma}$ for CMS.	75
3.12	L1 and HLT single electron and muon trigger [CMS00b, CMS02b].	79
3.13	Details of the event selection criteria for the WH analysis for CMS.	81
3.14	Number of events selected in the WH electron channel for CMS.	82
3.15	Number of events selected in the WH muon channel for CMS.	82
3.16	Comparison of the ATLAS WH analysis result [ATL99a, ATL99b] with the present CMS study.	84
B.1	Member classes of the object-oriented b-tag framework.	91
D.1	Parameter values used in the jet detector correction function D_{map}	95

Introduction

In the past thirty years particle physics has developed rapidly resulting in the formulation of the *Standard Model*, which seems to provide, at least in principle, a microscopic description for all known physical phenomena except gravity.

The Standard Model is not complete, *e.g.* it lacks any explanation for the pattern of particle masses. The *Higgs mechanism* provides a solution to the problem of how particles acquire their masses. It implies the existence of at least one new particle, the Higgs boson H^0 , which has not yet been observed. The Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN) will be switched on in winter 2007. If the Higgs boson exists, the LHC will be able to detect it.

Depending on the mass of the Higgs boson, physicists have a clear idea regarding its experimental signature. For quite low masses ($50 < m_{H^0} < 130$ [GeV])¹ the Higgs will predominantly decay into two b-quarks. The present study describes the investigation of the identification capability of b-quark signatures in the CMS experiment at the LHC. In order to test the b-quark identification methods developed, the data of the ALEPH experiment recorded at the LEP e^+e^- storage ring from 1992 till 1995 are used.

Chapter 1 is divided in three parts. The first part describes basic concepts of the Standard Model and the properties of b-quark decays, which can be utilized to identify (tag) them. The second part describes the foundations of the measurement of R_b at the Z^0 resonance with the ALEPH experiment. The measurement of R_b , which is the partial decay width of the Z^0 boson into $b\bar{b}$ -quark pairs², is used as a benchmark to tune the b-quark identification strategies. The current status of the Higgs boson search and the discovery potential of the CMS experiment for a light Higgs boson decaying into two b-quarks is summarized in the third part of chapter 1.

The measurement of R_b at the Z^0 resonance employing the developed b-tag strategy is described in chapter 2. About 3.8 million events collected by the ALEPH experiment from 1992 to 1995 together with the corresponding simulated data are utilized.

Chapter 3 describes the discovery potential of the CMS experiment for the Higgs produced in association with a W^\pm boson. The study is based on simulated data obtained from the full detector simulation. The analysis investigates the features and the expected performance of the b-tag developed above.

¹The convention $\hbar = c = 1$ is used in the present study.

²If not explicitly mentioned otherwise, $b \equiv \bar{b}$, $c \equiv \bar{c}$, etc. is used.

Chapter 1

The Standard Model: Basic concepts and their application

Contents

1.1 Overview	3
1.1.1 Basic concepts	3
1.1.2 Jets	5
1.1.3 Identification of b-hadrons	7
1.2 Testing the Standard Model: R_b	11
1.2.1 The measurement of R_b at the Z^0 resonance	13
1.2.2 Past and recent R_b measurements	14
1.3 Searching the missing link of the Standard Model: the Higgs	16
1.3.1 Status of the Standard Model Higgs search	16
1.3.2 Standard Model Higgs searches in pp collisions	16

1.1 Overview

The *Standard Model* (SM) is the name given to the summing-up of concepts describing the present known fundamental particles and their interaction [Gla61, Sal68, Wei67]. It is theoretically self-consistent, renormalizable [tH71, tHV72, tH73] and in agreement with all known experimental data [EWWG05, EWWG06]. However what causes the fundamental particles to have masses is not yet well established. The Standard Model has 18 free parameters¹ which need to be fixed by measurement.

An introduction into the basic concepts and some important features of the Standard Model are briefly described in the following.

1.1.1 Basic concepts

The Standard Model of the electromagnetic, the weak and the strong interactions is a quantum field theory based on the gauge group:

$$G_{\text{SM}} = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \quad (1.1)$$

¹The Super-Kamiokande collaboration discovered in 1998 evidence for oscillation of atmospheric neutrinos, which leads to non vanishing neutrino masses [Super-Kamiokande98]. The minimal extension of the Standard Model with non-vanishing neutrino masses leads to additional 7 free parameters.

group	# gauge bosons (fields)	associated symmetry	coupling constant
$SU(3)_C$	8 gluons	color	α_s
$SU(2)_L$	W_1, W_2, W_3	isospin	g
$U(1)_Y$	B	hypercharge	g'

Table 1.1: The gauge bosons of the Standard Model with corresponding coupling constants.

of unitary gauge transformations. $SU(2)_L$ is the non-Abelian left-handed electroweak isospin I group and $U(1)_Y$ is the Abelian hypercharge Y group. Both groups enter into the Glashow-Salam-Weinberg (GSW) theory as the $SU(2)_L \otimes U(1)_Y$ group, which describes the electroweak interactions [Gla61, Sal68, Wei67]. The hypercharge Y is connected with the electric charge Q and the 3rd component of the weak isospin I_3 by the Gell-Mann-Nishijima relation: $Y = 2(Q - I_3)$. The strong interactions are described by the theory of quantum chromodynamics (QCD), which is based on the non-Abelian group $SU(3)_C$. The local invariance of the Standard Model Lagrangian \mathcal{L}_{SM} under the gauge group G_{SM} results in twelve gauge bosons (spin-1 vector fields), which mediate the interactions. A comprehensive introduction into the theoretical framework and the foundations of the Standard Model can be found in [Qui83, AH89, PS95, SSZ00].

The gauge bosons and their corresponding coupling constants are listed in Table 1.1. The W_1 and W_2 gauge bosons can be identified with the experimentally observed W^+ and W^- particles. The experimentally observable neutral gauge bosons Z^0 and A (photon) are correlated to the W_3 and B gauge bosons by means of the weak mixing angle θ_W (Weinberg angle).

The basic constituents of matter, quarks and leptons, are fermions carrying spin-1/2. They are classified in left-handed weak isospin doublets and right-handed weak isospin singlets. In addition quarks are color triplets. Table 1.2 summarizes the electroweak quantum numbers electric charge Q , weak isospin I , 3rd component of the weak isospin I_3 and hypercharge Y for the fundamental particles of the Standard Model. All these particles have been identified experimentally [PDG04].

The strong interactions of quarks and gluons, which both carry color charge, are described by the theory of quantum chromodynamics. The QCD theory implies confinement, namely the experimental fact that free quarks and gluons have never been detected.

Because most of the perturbative calculations in QCD have been performed only to next-to-leading order (NLO), approximation methods (Monte Carlo methods) are used to take into account all orders². The calculations are done in two steps: In the first step (parton shower) the

fermion			Q	I	I_3	Y
$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L$	$\begin{pmatrix} 0 \\ -1 \end{pmatrix}$	$\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} -1 \\ -1 \end{pmatrix}$
$\begin{pmatrix} u \\ d \end{pmatrix}_L$	$\begin{pmatrix} c \\ s \end{pmatrix}_L$	$\begin{pmatrix} t \\ b \end{pmatrix}_L$	$\begin{pmatrix} \frac{2}{3} \\ -\frac{1}{3} \end{pmatrix}$	$\begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}$	$\begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}$
u_R	c_R	t_R	$\frac{2}{3}$	0	0	$\frac{4}{3}$
d_R	s_R	b_R	$-\frac{1}{3}$	0	0	$-\frac{2}{3}$

 Table 1.2: The electroweak quantum numbers electric charge Q , weak isospin I , 3rd component of the weak isospin I_3 and hypercharge Y for the fundamental particles of the Standard Model. The electric charge Q is listed in units of the electron charge $|e|$.

²The perturbatively zeroth (leading) order calculation is named Born approximation.

QCD dynamics are perturbatively calculated until some cut-off. The perturbative calculation is possible, because at small distances the QCD coupling constant α_s becomes small. The running of the coupling constant α_s is a feature of non-Abelian groups and is known as asymptotic freedom. The second step, needed when the value of the coupling constant α_s becomes large, cannot be calculated perturbatively and phenomenological models are used to fragment the quarks and the gluons into colorless hadrons.³

The higher-order QCD calculations involve additional quarks and gluons in the final state, which do not stem from the leading order (LO) interaction. In the parton shower approximation these additional quarks and gluons are classified into objects radiated in the initial-state (ISR) or in the final-state (FSR) of the interaction [ESW96].

1.1.2 Jets

The strong force causes the quarks and gluons to shower into a collection of colorless particles (jets). The energy and momentum of the jets is approximately that of the struck quarks or gluons [Roh94]. The momentum distribution of produced hadrons containing charm (c) or bottom (b) quarks, is approximately described by the Peterson fragmentation function [PSSZ83].

To reconstruct the energy (E) and momentum (\mathbf{p}) of the quarks or gluons a jet algorithm is used to combine the shower products into a single vector (jet). The first jet algorithms utilized simple cones [B⁺77, A⁺83]. The modern jet algorithms used in the present study are the *Jade*, the *Iterative Cone* (IC), the *Midpoint Cone* (MP) and the *inclusive k_\perp* jet algorithm. These utilize different iterative procedures to improve the reconstruction of the initial colored particle.

To compare the reconstructed jet observables, *e.g.* the number of jets, with theoretical calculations, the jet algorithm definition has to be *infrared* and *collinear safe*. An observable is infrared safe if, for any primordial configuration of the colored particles, adding an additional infinitely soft particle does not affect the observable at all. An observable is collinear safe if, for any primordial configuration of the colored particles, replacing one of them with an exact collinear pair of two particles, does not affect the observable at all.

The Jade algorithm is not infrared safe. Still the algorithm is used to investigate the b-hadron flight direction in 2-jet topologies (chapter 2). For such simple jet topologies the Jade algorithm is applicable [Fra94]. The Iterative Cone and the Midpoint Cone jet algorithms are both *almost unsafe*, because for next-to-leading order (NLO) calculations some parameter sets⁴ of these algorithms are safe and for next-to-next-to-leading order (NNLO) calculations all parameter sets are unsafe. The multi-jet topologies, which are investigated in chapter 3, require infrared and collinear safe jet algorithm definitions. The inclusive k_\perp jet algorithm is selected because this algorithm is safe to all orders [Sey98].

A description of the Iterative Cone and Midpoint Cone algorithm can be found in [CMS05k]. The principal jet algorithms utilized, Jade and inclusive k_\perp , are described in detail below.

1.1.2.1 The Jade jet algorithm

The Jade jet cluster algorithm [JADE86] operates on a set of input objects, which can be *e.g.* the 4-vectors of the colorless products of the parton shower. The algorithm calculates for

³For convenience the two separate steps parton shower and hadronization, are described as *shower* in the following.

⁴A parameter set definition of a jet algorithm consists of the algorithm parameters themselves, the complete definition of the jet input objects and the definition of the calibration of final state jets.

each pair of objects i and j the distance y_{ij} defined as:

$$y_{ij} = 2 \cdot \frac{E_i E_j}{E_{\text{vis}}^2} \cdot [1 - \cos(\angle \mathbf{p}_i, \mathbf{p}_j)],$$

where E_i , E_j are the energies and $(\angle \mathbf{p}_i, \mathbf{p}_j)$ is the angle between the momentum vectors \mathbf{p}_i , \mathbf{p}_j of the objects i , j respectively. E_{vis} is the total visible energy of the investigated process, *e.g.* the center-of-mass energy \sqrt{s} . The two objects which have the smallest y_{ij} are merged if [ALEPH00]:

$$\left(\frac{m_{ij}}{E_{\text{vis}}} \right)^2 < y_{\text{cut}}.$$

In other words the algorithm utilizes the invariant mass squared $m_{ij}^2 \simeq (2E_i E_j [1 - \cos(\angle \mathbf{p}_i, \mathbf{p}_j)])$ of two objects i and j to decide if their momentum vectors \mathbf{p}_i and \mathbf{p}_j point into a similar direction.

The iteration stops if the smallest y_{ij} computed is larger than y_{cut} . The final jets are computed from the clumped objects by means of a recombination scheme as described below.

1.1.2.2 The inclusive k_{\perp} jet algorithm

Several implementations of the k_{\perp} jet clustering algorithm [CDW92] exist. In the present study the inclusive k_{\perp} jet algorithm for hadron collisions [ES93] is used and described in the following. The cluster procedure starts again with a list of input objects. For each object i and each pair (i, j) the following distances are calculated:

$$\begin{aligned} d_i &= (E_{T,i})^2 \mathcal{R}^2, \\ d_{ij} &= \min\{E_{T,i}^2, E_{T,j}^2\} R_{ij}^2 \quad \text{with} \quad R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2, \end{aligned}$$

where \mathcal{R} is a dimensionless parameter normally set to 1 as proposed in [H⁺90]. $E_{T,i}$, $E_{T,j}$ are the transverse energies, η_i , η_j are the pseudo-rapidities⁵ and ϕ_i , ϕ_j are the azimuthal angles of two objects i and j respectively.

The algorithm searches for the smallest d_i or d_{ij} in all combinations of members of the input list. If a value of type d_{ij} is the smallest, the corresponding objects i and j are removed from the list of input objects. They are merged by using the E_T recombination scheme as described below and filled as one new individual object into the list of input objects. If a distance of type d_i is the smallest, then the corresponding object i is removed from the list of input objects and filled into the list of final jets.

The procedure is repeated until all objects are included in jets, *i.e.* the list of input objects is empty. The algorithm successively merges into one single jet objects which have a distance $R_{ij} < \mathcal{R}$. It follows that for all final jets i and j : $R_{ij} > \mathcal{R}$.

1.1.2.3 Jet recombination schemes

The Jade and the inclusive k_{\perp} jet algorithm merge in each iteration step input objects into possible final jets. The 4-vectors of the combined objects must be calculated during the clustering. The cone jet algorithms, Iterative Cone and Midpoint Cone, first group the input objects to be merged into proto-jets. The final determination of the jet 4-vectors is done in one step at the end of the jet finding. The principal 4-vector determination (recombination) schemes used are listed below.

⁵The pseudo-rapidity η is defined as $\eta = -\ln(\tan \frac{\theta}{2})$.

The E scheme is a simple 4-vector addition and results in massive jets:

$$E = \sum e^i, \quad p_{x,y,z} = \sum p_{x,y,z}^i. \quad (1.2)$$

The E_T scheme according to the Snowmass convention [H⁺90] for the inclusive k_\perp jet algorithm results in mass-less jets⁶, and is invariant under longitudinal boosts [CDSW93]:

$$E_T = \sum_i E_{T,i}, \quad \eta = \frac{\sum_i E_{T,i} \eta_i}{\sum_i E_{T,i}}, \quad \phi = \frac{\sum_i E_{T,i} \phi_i}{\sum_i E_{T,i}}. \quad (1.3)$$

1.1.3 Identification of b-hadrons

Because b-quarks shower into colorless b-hadrons⁷ (section 1.1.1), only the signatures of these hadrons are experimentally detectable. The identification of weakly decaying b-hadrons is possible by utilizing the differences of these decays compared to other light quark hadron decays. The distinctive properties of b-hadrons, which characterize their decays are [PDG04]:

- The long lifetime of b-hadrons: $\tau \approx 1.6$ ps results in an average high decay length \bar{l} of a few millimeters: *e.g.* For a B^+ -meson with a mass of $m_{B^+} \approx 5$ GeV and a mean energy of 32 GeV the average flight length is $\bar{l} \approx 3$ mm.⁸
- The mass of b-hadrons is on average higher (≈ 5 GeV) than hadrons made from c, s, d or u quarks only.
- The fragmentation function $f(b \rightarrow B)$ of b-quarks is harder compared to *e.g.* the fragmentation function for D-mesons $f(c \rightarrow D)$. It follows that the average energy of a b-hadron is higher compared to *e.g.* the energy of a c-hadron [PSSZ83].

Observables, which are sensitive to the properties listed above, are used in this study to identify b-hadrons.

Approximately 20% of b-quark decays are semi-leptonic and yield prompt leptons, which can also be used to identify the decays. The disadvantage of methods using this specific property is the smaller statistical sample, which results in overall low b-hadron identification efficiencies.

Because almost all hadrons which contain only the light quarks u, d and s behave similarly in the context of b-hadron identification, their contribution to b-hadron identification observables is summarized in a single distribution. To consider the huge amount of gluon-jets in LHC events (section 3) their distribution is separately plotted for these events. It turns out that b-hadron mis-identification is mainly caused by c-hadrons.

Estimation of the performance of b-hadron identification The performance of a particular b-hadron identification (b-tagging) method can be classified by the efficiency to identify b's (eff_b) and by the purity of the b-candidate sample (pur_b):

$$\text{eff}_b := \frac{\# \text{ tagged real b-hadron decays}}{\# \text{ all b-hadron decays}} \quad (1.4)$$

$$\text{pur}_b := \frac{\# \text{ tagged real b-hadron decays}}{\# \text{ all decays tagged as b-hadron decays}} \quad (1.5)$$

⁶For mass-less jets it follows that the transverse jet energy $E_{T,\text{jet}}$ is equal to the transverse jet momentum $P_{t,\text{jet}}$.

⁷In the present study the hadrons are named by their heaviest quark component for convenience: *e.g.* b-quark \Rightarrow b-hadron, c-quark \Rightarrow c-hadron, etc.

⁸The typical mean energy of b-hadrons from a Z^0 decay is 32 GeV [EWWG05]

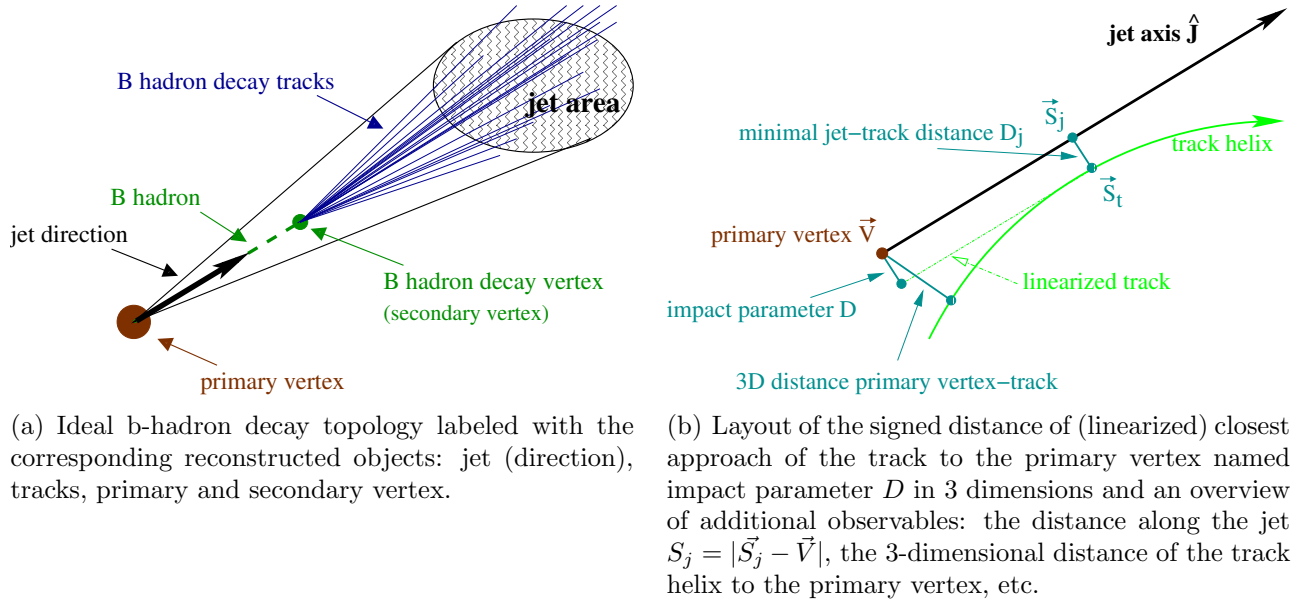


Figure 1.1: The topology of a b-hadron decay (a) and the layout of the 3D impact parameter D (b).

The mis-tagging rate mis_x to identify a hadron type x ($x \neq b$) as a b-hadron can be used instead of the b-purity pur_b . The mis-tagging rate is less sensitive to the composition of the sample under study:⁹

$$\text{mis}_x := \frac{\# \text{ of } x\text{-hadron decays tagged as b-hadron decays}}{\# \text{ all } x\text{-hadron decays}} \quad (1.6)$$

Observables sensitive to b-hadron identification The first step in b-hadron identification is to define the region in which to look for b-hadron decays. For this purpose the production (primary) vertex and the direction of the decaying hadron are determined (Figure 1.1(a)).

The primary vertex is reconstructed by an iterative fitting procedure of charged tracks stemming from this vertex. A jet algorithm (section 1.1.2) or, in case of Z^0 events decaying hadronically into a 2-jet topology, the thrust axis \vec{T} of the event¹⁰ can be used to estimate the flight direction. Objects which are close to the determined direction in η - ϕ space¹¹ are further investigated with a combination of the following observables:

- The long lifetime of b-hadrons leads to a secondary vertex in the direction of the jet. The exponential character of the decay makes it impossible in most of the cases to distinguish this secondary vertex from the primary one, where the b-hadron was produced. The secondary vertex reconstruction strongly depends on the ability of a particular detector to measure the hadron decay products, the charged tracks, with high precision close to the primary event vertex (see chapters 2 and 3). It follows that b-hadron identification

⁹This is the commonly used CMS definition of the mis-tagging rate. It follows that the mis-tagging rate for hadrons of type x is equal to the efficiency to tag these hadrons: $\text{mis}_x \equiv \text{eff}_x$

¹⁰The thrust axis \vec{T} is defined to maximize the ratio $\sum_a |\vec{p}_a \cdot \vec{T}| / \sum_a |\vec{p}_a|$, where $|\vec{T}| = 1$. The index a runs over all final state particles with momentum \vec{p}_a [Sjo94]. For a back-to-back 2-jet event topology the thrust axis has the same direction as the two reconstructed jets.

¹¹The distance of two objects i and j in η - ϕ space is defined as: $R = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}$.

based on explicit secondary vertices will always start with a reduced initial ensemble of objects (*e.g.* jets), namely those for which the reconstruction of a secondary vertex was possible. Consequently the efficiency of these tags is intrinsically lowered.

- An observable which uses the lifetime information of a decay without an explicit reconstruction of the secondary vertex is the jet lifetime probability P_{jet} [Mark II 91, ALEPH92].

For a group of N tracks inside a jet, P_{jet} combines into a single observable the individual probabilities P_{track} of each track ($k = 1, \dots, N$) to be compatible with the primary vertex:

$$P_{\text{jet}} := \prod_{k=1}^N P_{\text{track},k} \cdot \sum_{j=0}^{N-1} \left(-\ln \prod_{k=1}^N P_{\text{track},k} \right)^j / j! \quad (1.7)$$

The probability P_{track} of each track is calculated from the significance $\frac{D}{\sigma}$, where D is the signed distance of (linearized) closest approach of the track to the primary vertex (see Figure 1.1(b)), and σ is the measurement uncertainty on that distance. D is given a positive (negative) sign according to whether the track passes closest to the estimated b-hadron flight path downstream (upstream) of the primary vertex.

If all tracks in a hemisphere are used, the observable is named hemisphere lifetime probability P_{H} . The use of the hemisphere lifetime probability P_{H} for hadronically decaying Z^0 events leads to a slightly increased hadronic Z^0 decay selection efficiency, because in almost all cases two hemispheres can be defined even though jet finding may not have been successful.

The negative tail of the $\frac{D}{\sigma}$ distribution is used to determine the resolution function $\mathcal{R}(x)$. This function is used to calculate the probability P_{track} for each track to stem from the primary vertex:

$$P_{\text{track}}\left(\frac{D}{\sigma}\right) := \int_{-\infty}^{-|\frac{D}{\sigma}|} \mathcal{R}(x) dx \quad (1.8)$$

The resolution function $\mathcal{R}(x)$ can be obtained from real data without any usage of simulated data. The function strongly depends on the complete procedure needed to calculate the significance $\frac{D}{\sigma}$, namely determination of the primary vertex, track reconstruction and jet finding.

Using tracks to detect the possible lifetime signal in spatial dimensions of a few millimeters needs very good track reconstruction for tracks near the beam of the colliding particles. Vertex detectors allow track reconstruction with the required accuracy. A description of the ALEPH vertex detector performance can be found in [C⁺98]. The expected performance of the CMS vertex detector can be found in [CMS03b].

- The effective mass $M_{2nd\,vertex}$ of charged tracks assigned to a secondary vertex is higher for b-hadrons than for c- or uds-hadrons.
- The SLD collaboration¹² pioneered an observable which uses the mass information without an explicit determination of the secondary vertex [SLD97]. This observable was also used by the ALEPH experiment. The ALEPH lifetime mass tag adds up the tracks in a region defined by a jet or a hemisphere in order of decreasing inconsistency with the

¹²The SLD experiment was located at the Stanford Linear Accelerator Center (SLAC) [SLD].

primary vertex until their invariant mass exceeds 1.8 GeV (the mass of a c-hadron). The probability P_{track} of the last track added for which the mass limit was reached is defined as $\mu_H := P_{\text{track}}$ and is used as b-tag observable [ALEPH97].

- The rapidity $\eta_{2\text{nd vtx}, \text{track}, \text{jet}}$ of secondary vertex tracks with respect to the jet direction reflects the different hadron decay topologies. A b-hadron has on average a higher energy than a c-hadron. However, the larger multiplicity of b-hadron decays with respect to c-hadrons leads to smaller track rapidities $\eta_{2\text{nd vtx}, \text{track}, \text{jet}}$ for the b-hadron decay products with respect to c-hadrons. The secondary vertices of uds-hadron decays are mainly the result of badly measured low transverse momentum tracks so that their rapidity distribution is shifted to lower values.
- The fraction of the jet energy X_{energy} carried by charged tracks of a secondary vertex inside this jet is sensitive to differences in the fragmentation function of b- and c-hadrons. The reason is that for jets from b- or c-hadron decays almost all tracks which are assigned to a secondary vertex inside are coming from the b- or c-hadron decay respectively. As described before the secondary vertices in jets from uds-hadron decays are mainly fakes so that the corresponding values of X_{energy} are shifted to lower values.

Combination of observables for b-hadron identification The improved precision of modern particle detectors like ALEPH and CMS, which both have dedicated vertex detectors, as well as improved track and vertex reconstruction methods leads to increased efficiencies in reconstructing secondary vertices (chapter 2 and 3). This means that in these experiments b-tag observables based on explicitly reconstructed secondary vertices can be used without a big loss of efficiency.

A method pioneered by the DELPHI collaboration for the R_b measurement at the Z^0 pole (section 1.2) is based on explicitly reconstructed secondary vertices [DELPHI99]. To further improve b-hadron identification, b-tag observables are combined into one unique observable:

$$y_{\text{LEP1}} := n_c \cdot \prod_i \frac{f_i^c(x_i)}{f_i^b(x_i)} + n_{\text{uds}} \cdot \prod_i \frac{f_i^{\text{uds}}(x_i)}{f_i^b(x_i)} \quad (1.9)$$

The parameters n_c and n_{uds} are the normalized number of jets with an explicitly reconstructed secondary vertex in $Z^0 \rightarrow c\bar{c}$ and $Z^0 \rightarrow \text{uds}$ decays respectively ($n_c + n_{\text{uds}} = 1$). $f_i^b(x_i)$, $f_i^c(x_i)$ and $f_i^{\text{uds}}(x_i)$ are the probability density functions of a b-tag observable x_i in b-, c- and uds-decays respectively. The products in (1.9) run over the chosen b-tag observables for a given jet.

To take into account the huge amount of gluon jets in pp collisions at the LHC, formula 1.9 is revised for the LHC part of the present study:

$$y_{\text{LHC}} := n_c \cdot \prod_i \frac{f_i^c(x_i)}{f_i^b(x_i)} + n_{\text{uds}} \cdot \prod_i \frac{f_i^{\text{uds}}(x_i)}{f_i^b(x_i)} + n_g \cdot \prod_i \frac{f_i^g(x_i)}{f_i^b(x_i)} \quad (1.10)$$

The additional term for the gluon jets containing the probability density function $f_i^g(x_i)$ and the normalization factor n_g leads to $n_c + n_{\text{uds}} + n_g = 1$. The normalization factors are calculated from typical LHC $t\bar{t}$ events (see chapter 3).

1.2 Testing the Standard Model: R_b

Important experimental tests of the Standard Model were performed at the Z^0 pole, by studying the annihilation of an electron e^- and a positron e^+ into a Z^0 boson and the following weak decay of the boson into a fermion f and anti-fermion \bar{f} pair¹³. The process in the lowest order (Born level) is:

$$e^+e^- \rightarrow (\gamma, Z^0) \rightarrow f\bar{f}. \quad (1.11)$$

The corresponding Feynman graph¹⁴ is shown in Figure 1.2. At the centre-of-mass energies close to the Z^0 mass the production cross section is dominated by the Z^0 exchange. The contributions due to the photon exchange and the γZ^0 interference term are of the order 1%. The explicit calculation of the cross section can be found in [Alt89].

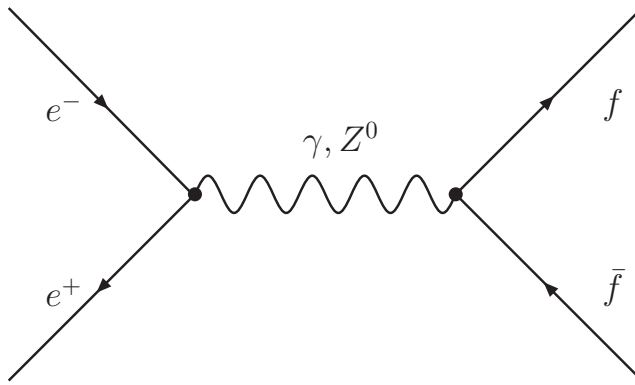


Figure 1.2: The Feynman graph at Born level of the weak interaction $e^+e^- \rightarrow f\bar{f}$.

The LEP and SLD collaborations were able to measure the Z^0 mass with a relative precision of $2.3 \cdot 10^{-5}$. Therefore this mass is one of the most precisely known parameters of the Standard Model. All Z^0 pole results are in good agreement with the Standard Model predictions. The experimental precision was sufficient to probe predictions of the Standard Model at loop level. This access to higher order radiative corrections enabled the indirect measurement of the masses of particles not directly produced, notably the top quark and the Higgs boson [EWVG05].

The measurement of Z^0 boson parameters are in turn used as input for the determination of the Standard Model parameters. One of these observables is the partial decay width R_b defined as:

$$R_b := \frac{\Gamma(Z^0 \rightarrow b\bar{b})}{\Gamma(Z^0 \rightarrow \text{hadrons})}. \quad (1.12)$$

Experimentally accessible is the ratio of cross sections R'_b :

$$R'_b := \frac{\sigma(Z^0 \rightarrow b\bar{b})}{\sigma(Z^0 \rightarrow \text{hadrons})} \quad (1.13)$$

¹³The flavor changing neutral currents (FCNC) have never been detected experimentally. The solution to the problem in the GSW theory was devised by Glashow, Iliopoulos and Maiani [GIM70].

¹⁴For all Feynman graphs the following convention is used: Left-to-right in the diagram represents time; a process begins on the left and ends on the right.

This ratio includes a very small photon propagator contribution of $R_b \approx R'_b + 0.0002$ at the Z^0 centre-of-mass-energy as pointed out before [EWWG05]. In the following this effect is neglected: $R_b \simeq R'_b$.

The uncertainties, which affect all quark flavors equally, cancel in the ratio R_b , *e.g.* to $\mathcal{O}(\frac{\alpha_s}{\pi})$ it is independent of QCD corrections and errors in the measurement of α_s . Because of the high top quark mass m_t only the Feynman diagrams involving a top contribute to the $Z^0 b \bar{b}$ vertex (Figure 1.3) and the vacuum polarization (Figure 1.4). Owing to these corrections R_b is sensitive to m_t and is only weakly sensitive to the Higgs boson, due to cancellations in the ratio [EWWG05, EWWG06]. However R_b isolates the $Z^0 b \bar{b}$ vertex and contributions from new physics could be easily detected [SLD96, Cli96, Hil96].

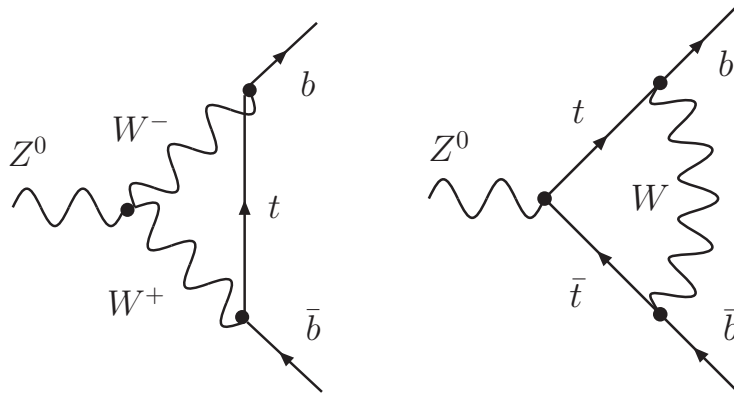


Figure 1.3: The Feynman graphs of the two dominant $Z^0 b \bar{b}$ vertex corrections, which introduce the top mass dependence.

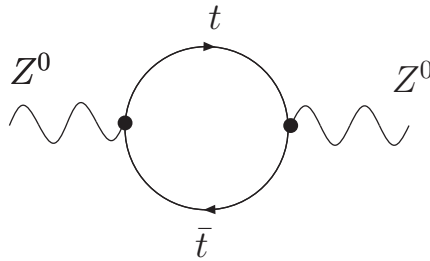


Figure 1.4: The Feynman graph of the dominant top quark contribution to the Z^0 vacuum polarization.

1.2.1 The measurement of R_b at the Z^0 resonance

For the experimental determination of the R_b value at the Z^0 resonance the *double tag method* has commonly been used. This method is designed to minimize the dependence of the measurement from the Monte Carlo simulation and systematic errors coming from imperfect knowledge of the properties of the b-hadrons produced [ALEPH93a].

The double tag method is based on the following assumptions:

- In hadronic Z^0 decays the number of quark flavors is conserved so that quarks are always produced in pairs: *e.g.* b and \bar{b} .
- The flavors of the quarks accompanying the primary b-hadrons are uncorrelated.
- The momenta of the two b-hadrons are quite large, thus dividing each event into two hemispheres will almost always separate the two b-hadrons.

With these assumptions the determination of the b-tag efficiency eff_b and of R_b itself is possible with a minimal use of input from Monte Carlo simulation. In addition the measurement does not depend on the luminosity measurement, because only the number of hadronic events is used.

To apply the double tag method, both hemispheres of a hadronic event, as determined by a plane perpendicular to the thrust axis, are analyzed separately using the observables described in 1.1.3. The determined efficiencies eff_b , eff_c and eff_{uds} can be interpreted as the probability to specify correctly the primary decay quark for one hemisphere to be a b-, c- or uds-quark respectively.

With:

$$R_b = \frac{\Gamma(Z^0 \rightarrow b\bar{b})}{\Gamma(Z^0 \rightarrow \text{hadrons})}, \quad (1.14)$$

$$R_c = \frac{\Gamma(Z^0 \rightarrow c\bar{c})}{\Gamma(Z^0 \rightarrow \text{hadrons})}, \quad (1.15)$$

$$R_{uds} = \frac{\Gamma(Z^0 \rightarrow uds)}{\Gamma(Z^0 \rightarrow \text{hadrons})} \quad (1.16)$$

it follows:

$$R_b + R_c + R_{uds} = 1. \quad (1.17)$$

The fractions F_1 and F_2 of single tagged and double tagged hemispheres respectively are consequently:

$$F_1 = R_b \text{eff}_b + R_c \text{eff}_c + (1 - R_c - R_b) \text{eff}_{uds}, \quad (1.18)$$

$$F_2 = R_b \text{eff}_b^d + R_c \text{eff}_c^d + (1 - R_c - R_b) \text{eff}_{uds}^d, \quad (1.19)$$

where:

$$\text{eff}_x^d := \text{efficiency to tag both hemispheres as quark type } x \text{ decay,} \\ (x = b, c \text{ or } uds).$$

Equation (1.19) can be further evaluated:

$$\begin{aligned}
 F_2 = & R_b \text{eff}_b^2 + R_c \text{eff}_c^2 + (1 - R_c - R_b) \text{eff}_{uds}^2 + \\
 & R_b \lambda_b (\text{eff}_b - \text{eff}_b^2) + R_c \lambda_c (\text{eff}_c - \text{eff}_c^2) + (1 - R_c - R_b) \lambda_{uds} (\text{eff}_{uds} - \text{eff}_{uds}^2)
 \end{aligned} \quad (1.20)$$

where:

$$\lambda_x = \frac{\text{eff}_x^d - \text{eff}_x^2}{\text{eff}_x - \text{eff}_x^2} \quad := \quad \text{hemisphere correlation for quark type x,} \\
 (x = b, c \text{ or } uds).$$

The hemisphere correlations λ_x occur because a tag in one hemisphere influences the efficiency of the tag in the other. For c- and uds-quarks the correlations are negligible in high purity b-tag regions¹⁵. Consequently (1.18) and (1.20) can be expressed as:

$$R_b = \frac{(F_1 - R_c (\text{eff}_c - \text{eff}_{uds}) - \text{eff}_{uds})^2}{F_2 - R_c (\text{eff}_c - \text{eff}_{uds})^2 + \text{eff}_{uds}^2 - 2 F_1 \text{eff}_{uds} - \lambda_b R_b (\text{eff}_b - \text{eff}_b^2)}, \quad (1.21)$$

$$\text{eff}_b = \frac{F_2 - R_c \text{eff}_c (\text{eff}_c - \text{eff}_{uds}) - F_1 \text{eff}_{uds} - \lambda_b R_b (\text{eff}_b - \text{eff}_b^2)}{F_1 - R_c (\text{eff}_c - \text{eff}_{uds}) - \text{eff}_{uds}}. \quad (1.22)$$

The two coupled equations (1.21) and (1.22) are decoupled and solved analytically for R_b and eff_b with MAPLE [Map]. The short MAPLE script is given in Appendix A.

The resulting two solutions for R_b and eff_b have four unknown parameters each: The value of R_c is taken from the actual Standard Model fit [EWWG05], $R_c = 0.1721$, while the other quantities eff_c , eff_{uds} and λ_b are taken from Monte Carlo simulation. The fractions F_1 and F_2 of single tagged and double tagged hemispheres respectively are determined from the data. It turns out that always only one of the two possible solutions of the decoupled equations for R_b and eff_b corresponds to a reasonable physics result.

To obtain the best R_b measurement the statistical and systematic uncertainties have to be as small as possible. One dominant source contributing to the systematic error is the uncertainty of the b-hemisphere correlation λ_b . The value of λ_b itself is normally close to zero and suggests that all effects which contribute are small. However this can happen by chance due to a cancellation of several large single contributions, either positive or negative, in the resulting value of the b-hemisphere correlation λ_b .

In previous analyses it turned out that the main source of the hemisphere correlation was the usage of a common primary vertex, which could be reduced by artificially reconstructing one primary vertex per hemisphere. The decreased primary vertex reconstruction performance and the resulting reduced b-tag performance caused an overall increased statistical error. Nevertheless due to the significantly reduced systematic error the overall combined error of the R_b measurement could be minimized [ALEPH97, DELPHI99].

1.2.2 Past and recent R_b measurements

The first precise measurements of R_b by the LEP experiments and by the SLD collaboration caused particular interest from the beginning of 1993 till 1996. The measurements seemed to be in disagreement with the Standard Model prediction [Hil96, CERN97]. A compilation of the R_b analysis results from 1993 till 1997 is shown in Figure 1.5. The specified years point

¹⁵The correlations are suppressed by the factor: $(R_x \cdot \text{eff}_x) / (R_b \cdot \text{eff}_b)$, $x = c$ or uds

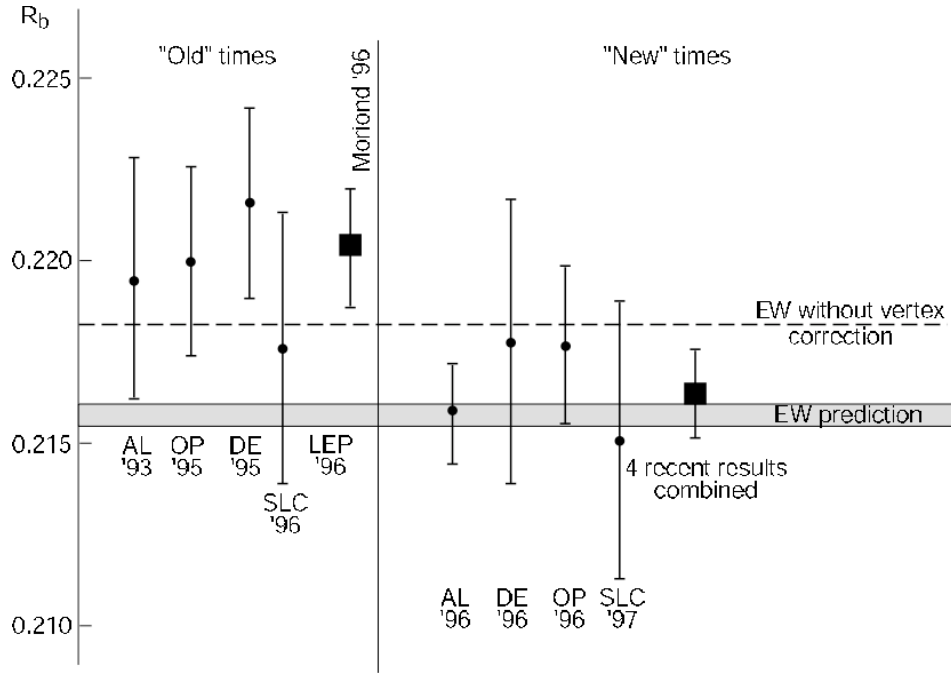


Figure 1.5: The R_b analysis results from 1993 till 1997. Before 1996 the LEP collaborations ALEPH (AL), DELPHI (DE), OPAL (OP) and the SLD collaboration measured *too high* R_b values. The specified years point out when the respective analysis was published. Later results were compatible with the Standard Model (EW) prediction of R_b . The EW prediction without vertex correction corresponds to the value of R'_b , which includes a photon propagator contribution of 0.0003 [CERN97].

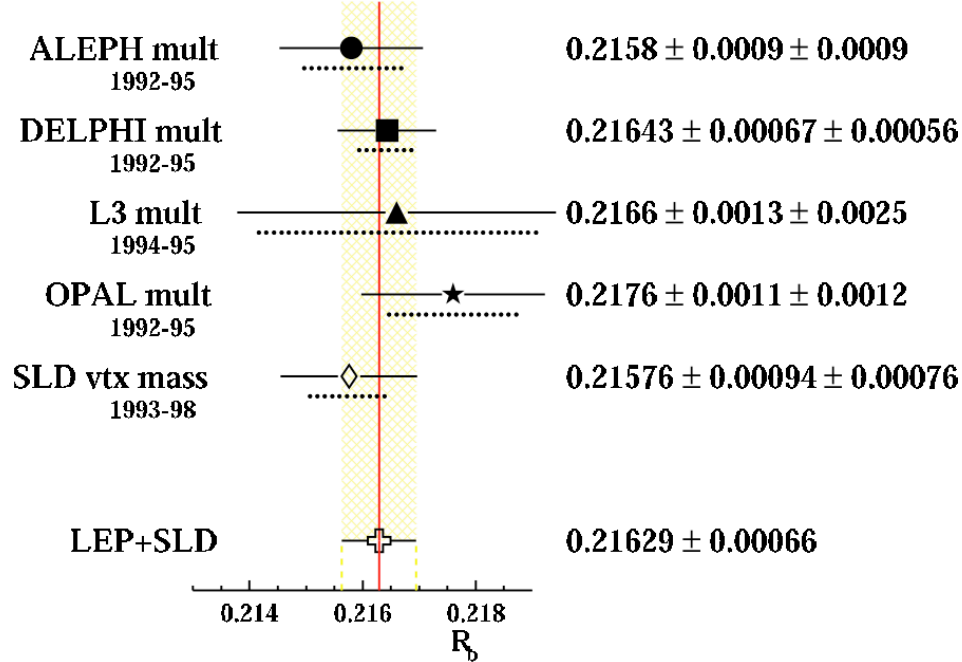


Figure 1.6: Recent R_b measurements and the combined (LEP+SLD) world average value [EWWG05]. The dotted error bars correspond to the statistical error of each single measurement obtained from events during the years listed.

out when the respective analysis was published. After 1996 the collaborations further improved their R_b measurements and the published results became compatible with the Standard Model prediction.

The world average value of $R_b = 0.21629 \pm 0.00066$ remains in agreement with the Standard Model prediction [EWWG05]. Figure 1.6 summarizes the recent R_b measurements of ALEPH, DELPHI, L3, OPAL and SLD and gives an overview of the values used to compute the world average. In this figure the specified years indicate the period of data taking used by the specified analyses. The DELPHI experiment provides the most accurate single measurement.

1.3 Searching the missing link of the Standard Model: the Higgs

The masses of fundamental Standard Model particles can be generated by interactions with a scalar Higgs field [Hig64]. By means of such a field masses are introduced in a consistent way so that the gauge invariance of the SM is not destroyed. The gauge boson and the fermion masses arise from the interaction energies with the ground state of the Higgs field, which is spontaneously broken.

The spontaneous breakdown of the scalar Higgs field symmetry is accompanied by the appearance of mass-less spin-0 Goldstone bosons. The Higgs field of the GSW theory is a complex isodoublet field, corresponding to four real scalar fields which produce four Goldstone bosons. Three of these Goldstone bosons are absorbed and result in massive W^+ , W^- and Z^0 gauge bosons. The A (photon) gauge boson is mass-less. The fermion masses are introduced by Yukawa couplings to the Higgs field ground state [GSW62, Wei67, Wei76].

The fourth Goldstone boson, which is not absorbed, should be experimentally detectable and it is given the name Higgs boson. The mass of the Higgs boson cannot be predicted in the Standard Model. However, internal consistency conditions and extrapolations of the Standard Model to higher energies provide stringent upper and lower bounds. The maximal cut-off value for the Standard Model Higgs boson mass is roughly 1 TeV [SSZ00].

1.3.1 Status of the Standard Model Higgs search

The Standard Model Higgs boson has not been discovered up to now. The ALEPH collaboration at LEP reported an excess compatible with the production of the Standard Model Higgs boson at 115 GeV. The combination of the results from all four LEP collaborations resulted in establishing a lower bound for the Standard Model Higgs mass of 114.4 GeV at 95% confidence level [LHWG03].

LEP has been operational until November 3, 2000. However, the experiments at the Hadron Electron Ring Accelerator (HERA) at DESY and the D0 and CDF experiments at Fermilab's Tevatron are currently running and continuously scanning their data for Higgs Boson signatures [Lam06, Kil06, Tev].

The constraint on the Higgs boson mass m_H obtained from precision electroweak measurements and direct searches performed so far at LEP and by the SLD, CDF, and D0 collaborations, assuming the Standard Model to be the correct theory of nature, is summarized in Figure 1.7. The preferred value corresponds to the minimum of the curve $m_H = 85$ GeV, with an experimental uncertainty of +39 and -28 GeV [EWWG06]. Figure 1.7 serves as a guideline to look for a possible Standard Model Higgs Boson in the mass range of $114 < m_H < 199$ [GeV] [EWWG06]. The input measurements for the Higgs boson mass prediction are listed in figure 1.8.

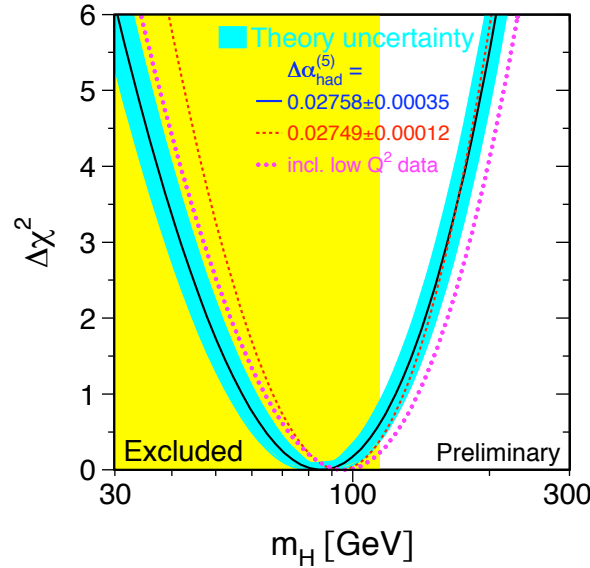


Figure 1.7: The dependency of $\Delta\chi^2$ versus Higgs boson mass m_H [EWWG06]. The solid black line is a result of the combination of the precision electroweak measurements performed at LEP and by the SLD, CDF, and D0, collaborations, assuming the Standard Model to be the correct theory of nature. The blue band represents an estimate of the theoretical error due to missing higher order corrections. The vertical yellow band shows the 95% CL exclusion limit on m_H from direct searches. The red dashed line is the result obtained using a revised evaluation of $\Delta\alpha_{\text{had}}^{(5)}(m_Z^2)$ [dTY05].

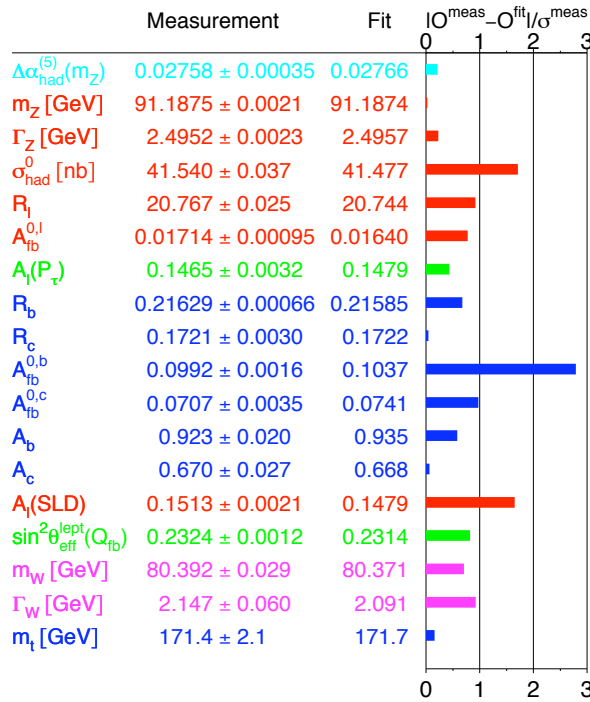
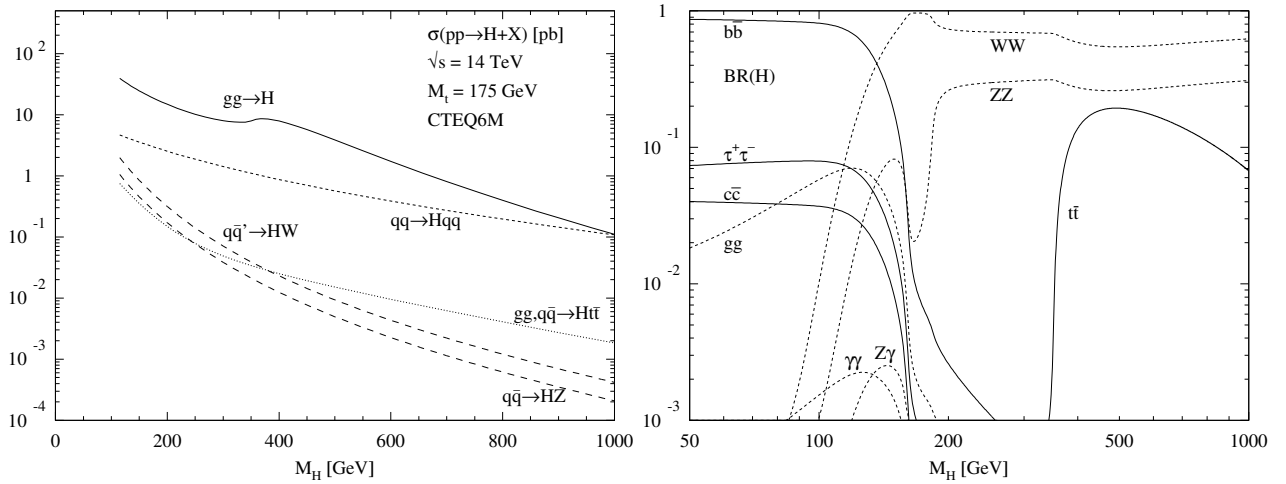


Figure 1.8: Summary of the 18 input measurements included in the Standard Model fit [EWWG06]. The pulls are defined as deviation from the theoretical values in units of experimental 1σ deviations. These measurements were also used for the Higgs boson mass prediction.



(a) The Standard Model Higgs production cross sections in pp collisions for the various production mechanisms as function of the Higgs mass. (b) The branching ratios $BR(H)$ of the dominant decay modes of the Standard Model Higgs.

Figure 1.9: The cross sections of the production mechanisms in pp collisions (a) and the branching ratios of the Standard Model Higgs for different Higgs masses (b) [SZ97, DKS98].

1.3.2 Standard Model Higgs searches in pp collisions

The LHC proton-proton (pp) collider, which will operate from 2007 on at the center-of-mass energy of 14 TeV, will give access to a new energy region [LHC95]. Due to the theoretical stringent upper and lower bounds on the Standard Model Higgs mass, the LHC will give the final answer to the question whether the Standard Model Higgs exists or not [SSZ00].

Figure 1.9(a) shows as a function of the Higgs mass M_H the quite sizeable cross sections for the different Standard Model Higgs production mechanisms in pp collisions at a center-of-mass energy of 14 TeV [SZ97, DKS98]. Over the entire Higgs mass range gluon fusion is the dominant production process (Figure 1.10). About two orders of magnitude lower are the production cross sections of the associated Higgs production channels, where the Higgs is produced together with two top quarks, or with a W^\pm or a Z^0 boson.

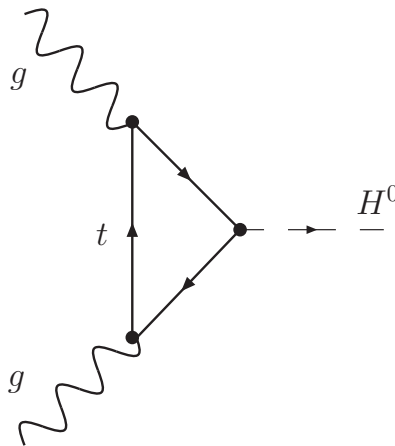


Figure 1.10: The main gluon fusion process incorporating a top quark at Born level.

The branching ratios $\text{BR}(\text{H})$ of the dominant Standard Model Higgs decay modes for different Higgs masses M_{H} are summarized in Figure 1.9(b) [SZ97]. For $M_{\text{H}} < 130 \text{ GeV}$ the Higgs decays predominantly into a $b\bar{b}$ -quark pair. In this mass range the decay rate into $\tau^+\tau^-$ is also sizeable and amounts to 7 %. In the intermediate mass range $120 < M_{\text{H}} < 130 [\text{GeV}]$ the Higgs decay rate of *e.g.* $\text{H}^0 \rightarrow \gamma\gamma$ is quite small but due to the small natural width of $\Gamma_{\text{H}^0} \ll 1 \text{ GeV}$ a possible signal would be clearly visible. For larger Higgs masses M_{H} the decay is mainly through $\text{H}^0 \rightarrow \text{WW}$ and ZZ . At $M_{\text{H}} \geq 2 M_{\text{t}}$ the $\text{t}\bar{\text{t}}$ branching fraction can be almost as large as 20% [SZ97, DKS98].

The Higgs discovery potential of the CMS experiment is summarized in [CMS03a].

1.3.2.1 Associated Higgs production in the W^{\pm}H channel

The Higgs decay into a $b\bar{b}$ -quark pair is experimentally very difficult to detect, due to the huge amount of background events. The associated Higgs production channels can be utilized to increase the signal significance of the $\text{H} \rightarrow b\bar{b}$ signature. The enhanced signal signatures $\text{t}\bar{\text{t}}\text{H}$, W^{\pm}H or Z^0H have lower production cross sections, but are more visible in the presence of backgrounds due to the possibility to preselect events by a leptonic decay signature of the associated Higgs partner(s).

The Feynman graph at Born level for the W^{\pm}H production, where the Higgs boson H^0 decays into a $b\bar{b}$ -quark pair, is illustrated in Figure 1.11. Two quarks $q\bar{q}$ produce an off-shell $\text{W}^{\pm*}$ boson, which radiates the Higgs boson H^0 . The leptonic decays of the W^{\pm} boson into a lepton/neutrino pair enable the extraction of the W^{\pm}H signal from the experimental data.¹⁶

The cross section prediction from next-to-leading order electroweak and QCD calculations can be found in [CDK03] and [B⁺04]. The Higgs-strahlung is one of the theoretically cleanest Higgs production channels at hadron colliders and is important for the determination of the Higgs boson properties and couplings. A clean determination of the H^0WW coupling times the Higgs branching ratio in the W^{\pm}H channel should be possible, if the systematic errors originating from higher-order corrections and structure functions become small [BDH04]. The H^0WW coupling can then be extracted if the Higgs branching ratio has been measured independently in other Higgs production processes.

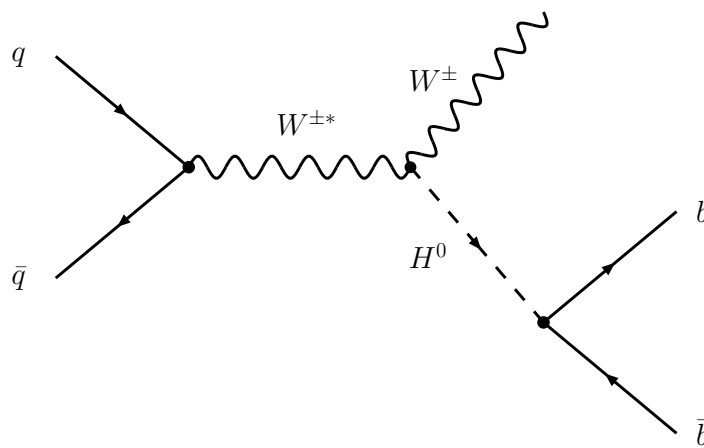


Figure 1.11: The Feynman graph at Born level of the W^{\pm}H production. Two quarks $q\bar{q}$ produce an off-shell $\text{W}^{\pm*}$ boson, which radiates the Higgs boson H^0 . The Higgs boson H^0 decays into a $b\bar{b}$ -quark pair.

¹⁶For convenience the WH production is used as a synonym for the W^{\pm}H production in the following.

Chapter 2

e^+e^- physics

Contents

2.1	Introduction	21
2.2	The LEP electron-positron collider	21
2.3	The ALEPH experiment at LEP	24
2.4	The investigation of a new b-tag for ALEPH	26
2.4.1	Introduction	26
2.4.2	The benchmark b-tag: the ALEPH lifetime mass tag	27
2.4.3	Event selection and reconstruction	28
2.4.4	Description and performance of the new ALEPH b-tag	33
2.4.5	Summary	37

2.1 Introduction

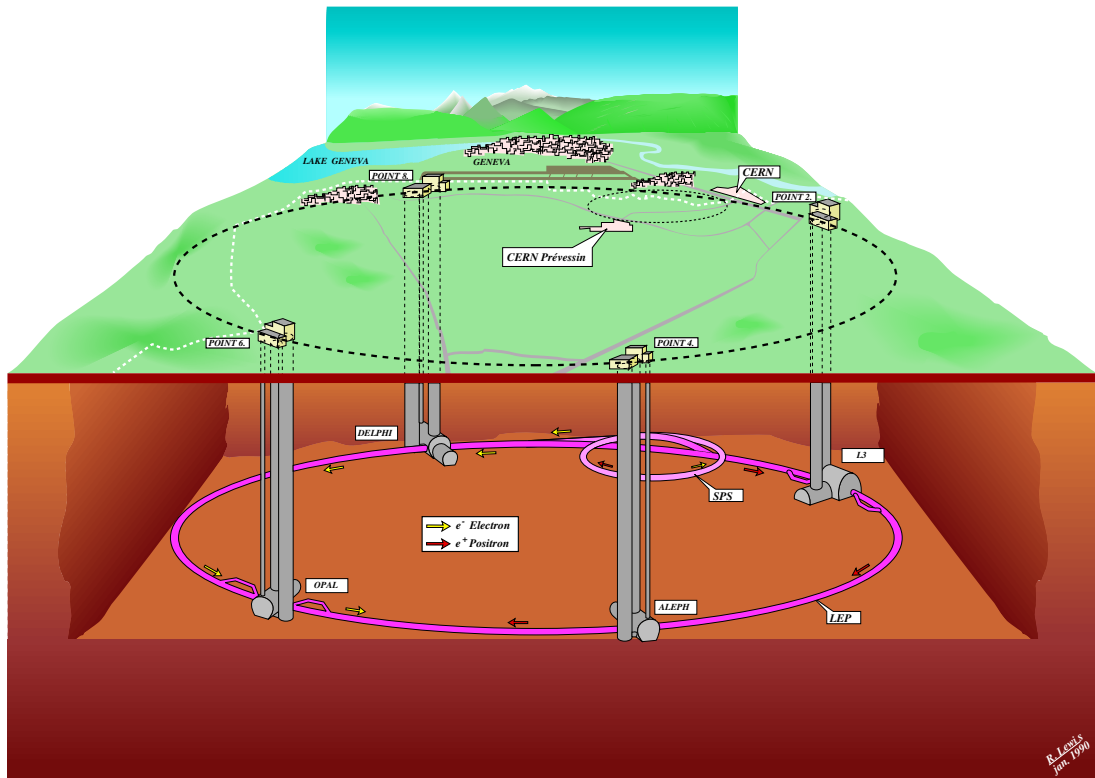
From 1989 to 1995 CERN’s electron-positron collider (LEP) was operated at the centre-of-mass energy of the Z^0 resonance, corresponding to about 91.2 GeV (LEP1 phase). LEP delivered about 16M Z^0 bosons to the four experiments ALEPH, DELPHI, L3 and OPAL during these years.

Some advantages of an e^+e^- -accelerator are:

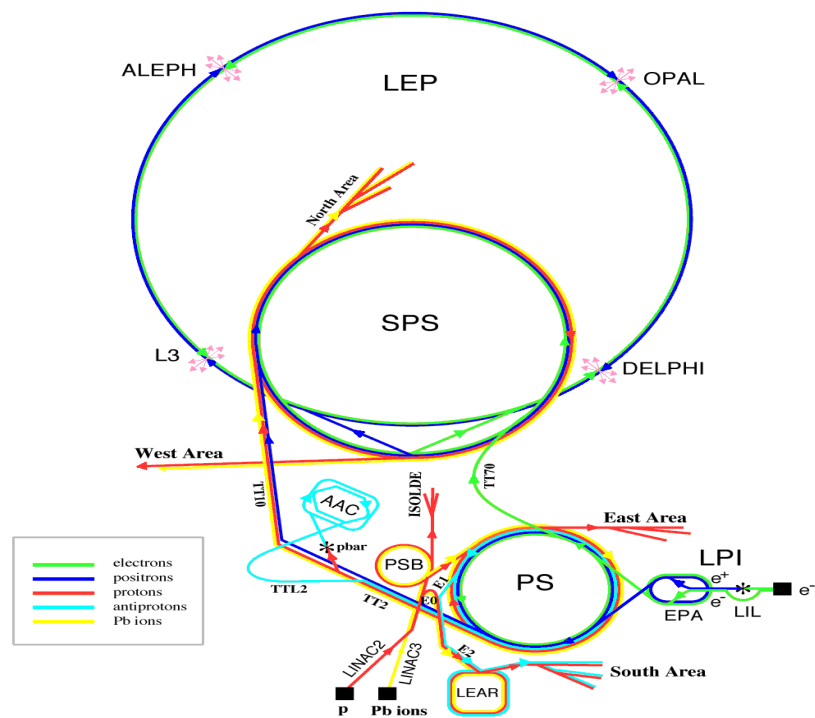
- Colliding beams consist of point-like particles.
- The center-of-mass energy E_{cms} is exactly known: $E_{\text{cms}} = 2 E_{\text{beam}}$
- Very clear signal events mostly without pollution from any backgrounds.

It follows that final states of events from an e^+e^- -collider can be investigated with high precision.

As described in section 1.2, the identification of b-hadrons plays an important role in a high precision measurement of R_b at the Z^0 resonance. This chapter describes the measurement and the new experimental tools in order to improve on previous results. The tools are implemented in C++, based on the ROOT object-oriented analysis framework [BR97].



(a) Overall view of the LEP setup with the four experiments ALEPH, DELPHI, L3 and OPAL.



(b) The LEP storage ring with pre-accelerators.

Figure 2.1: The LEP experimental setup.

2.2 The LEP electron-positron collider

The LEP collider had a circumference of 26.7 km and was located at the Swiss/French border near Geneva in a tunnel which is on average 100 m below the surface (Figure 2.1(a)). The complete data taking period was from 1989 till 2000. The LEP ring consisted of eight arcs alternating with eight straight sections. At four of these straight sections the LEP experiments were situated.

Before the electrons and positrons were injected into the LEP ring, they were accelerated by a chain of pre-accelerators up to an energy of 20 GeV (Figure 2.1(b)). The LEP accelerator itself was a synchrotron operated as storage ring. The acceleration was done by RF cavities, dipole magnets guided the beams through the curved sections and quadrupole and sextupole magnets focused the beams. A detailed description of the complete LEP setup can be found elsewhere [CERN04].

The electrons and positrons were accelerated in bunches. The luminosity \mathcal{L} of an e^+e^- -storage ring is: [Ber02]

$$\mathcal{L} := \frac{N_e^2 n_b f}{4 \pi \sigma_x \sigma_y}, \quad (2.1)$$

where:

$$\begin{aligned} N_e &:= \# \text{ particles in a bunch,} \\ n_b &:= \# \text{ bunches per beam,} \\ f &:= \text{rotation frequency,} \\ 4 \pi \sigma_x \sigma_y &:= \text{transverse beam area.} \end{aligned}$$

LEP reached its design luminosity of $\mathcal{L} = 13 \cdot 10^{30} \text{ cm}^{-2} \text{ s}^{-1}$ in 1991. Later LEP had luminosities of $10^{31} - 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ so that approximately 1000 Z^0 events were recorded every hour by each of the four LEP experiments [EWWG05].

The LEP beam spot constraint Typical values for the size of the luminous region¹ in LEP are listed in Table 2.1. The luminous region had an approximately Gaussian profile and was used by the LEP experiments as constraint for the determination of their interaction region (beam spot constraint).

Coordinate	r.m.s. beam size
x	120-150 μm
y	5 μm
z	7 mm

Table 2.1: Parameters of LEP's luminous regions [B⁺97].

Due to their radial acceleration electrons and positrons in LEP lost energy by emitting synchrotron radiation in the xz -plane (bremsstrahlung). The amount of energy loss in one turn of a particle with mass m and charge q in a ring with radius of curvature ρ is [MS03]:

$$\Delta E = \frac{q^2 \beta^3 \gamma^4}{3 \epsilon_0 \rho}, \quad (2.2)$$

¹The coordinate system is defined at the interaction point with positive x towards the centre of the LEP ring, y vertically upwards and z along the electron beam direction.

where:

$$\begin{aligned}\epsilon_0 &:= \text{permittivity of free space,} \\ \beta &:= \frac{v}{c}, \\ \gamma &:= (1 - \beta^2)^{-\frac{1}{2}}.\end{aligned}$$

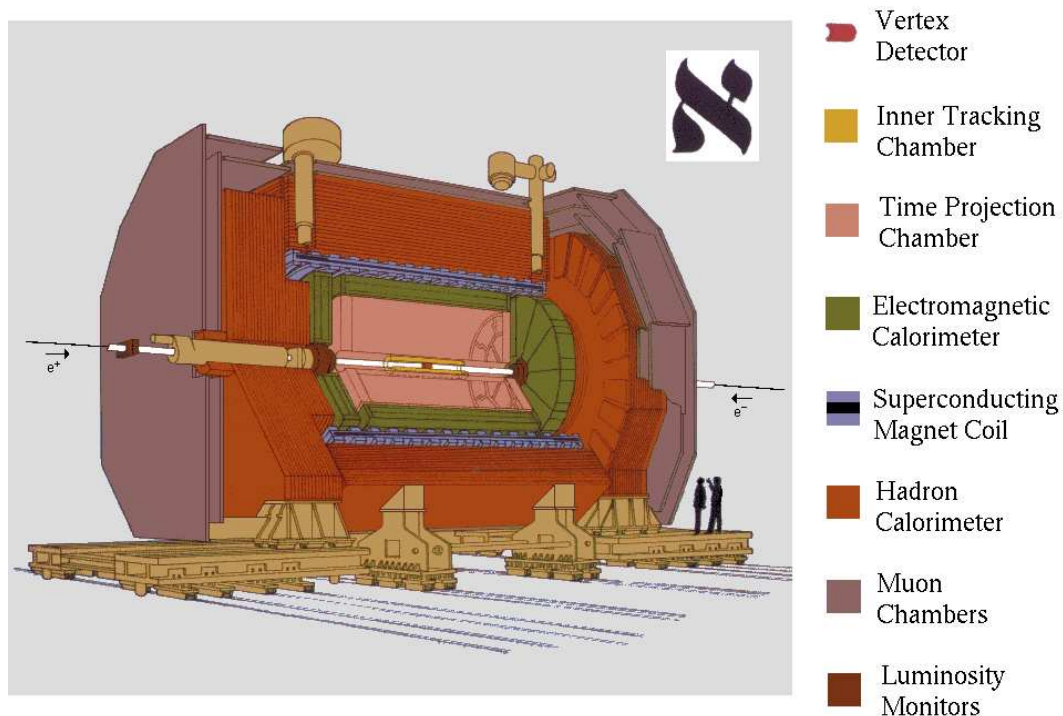
For relativistic particles ($\beta \sim 1$) it follows: $\Delta E \sim 1/m^4$. Consequently for electrons and positrons the losses are severe due to their low masses.

The emission of synchrotron radiation influenced the measurement accuracy of the x -position of the LEP beam, which was decreased in comparison to that of the y -position determination [B⁺97].

2.3 The ALEPH experiment at LEP

The ALEPH detector (Apparatus for LEP Physics) was one of the four multi-purpose detectors at LEP and was located at the experimental area of Point 4 near Echevex (France). The main overall objectives of the ALEPH design were the precise measurement of the parameters of the electroweak Standard Model, testing QCD at large Q^2 and searching for new physics *e.g.* the Higgs boson, etc. [ALEPH90]. The detector had a cylindrical shape with approximately 12 m diameter by 12 m length and consisted of independent and modular sub-detectors arranged in layers around the beam-pipe (radius 5.3 cm), each one specializing in a different task (Figure 2.2).

The ALEPH detector components and their performance are described briefly in the following. A detailed and comprehensive description of the ALEPH detector performance can be found in reference [ALEPH95]. The values given below are taken from this reference.



The ALEPH Detector

Figure 2.2: The ALEPH detector.

The tracking system Charged particles were tracked with three devices inside a superconducting solenoid having an axial field of 1.5 T. The innermost tracking detector was the mini vertex detector (VDET) which was installed in 1991. It consisted of 2 layers of silicon wafers with strip readout in 2 dimensions (radii: ≈ 6.3 cm and 10.8 cm). It was followed by the inner tracking chamber (ITC), a cylindrical multi-wire drift chamber able to provide up to eight precise r - ϕ points per track. Finally the time projection chamber TPC, a cylindrical imaging drift chamber, provided up to 21 three dimensional coordinates of the particle trajectories.

In hadronic Z^0 decays, tracks crossing at least four pad rows in the TPC are reconstructed with an efficiency of 98.6%.

For tracks with two VDET coordinates a transverse momentum resolution of:

$$\frac{\Delta p_t}{p_t} = 6 \cdot 10^{-4} p_t \oplus 0.005 \quad (p_t \text{ in GeV}) \quad (2.3)$$

could be achieved. For these tracks the transverse σ_{d_0} and longitudinal σ_{z_0} impact parameter resolution in hadronic Z^0 events were:

$$\sigma_{d_0} = \sigma_{z_0} = 25 \oplus \frac{95}{p \text{ GeV}^{-1}} \mu\text{m} \quad (p \text{ in GeV}). \quad (2.4)$$

The determination of the mean position of the interaction region (primary vertex) for hadronic events was done with a precision of $\approx 20 \mu\text{m}$ in the horizontal and $\approx 10 \mu\text{m}$ in the vertical direction. The result was obtained for groups of about 75 hadronic events. This accuracy could be achieved by using the additional information from the LEP machine about the size of the luminous region (section 2.2).

The calorimeter system The energy of neutral and charged particles was measured by an electromagnetic calorimeter (ECAL) inside and a hadron calorimeter (HCAL) outside the magnetic coil. Both were sampling calorimeters with different longitudinal segmentation. The averaged granularity of the ECAL read out towers pointing to the nominal interaction point was $0.9^\circ \times 0.9^\circ$. The resulting energy resolution was:

$$\frac{\sigma(E)}{E} = 0.0009 + \frac{0.18}{\sqrt{E}} \quad (E \text{ in GeV}). \quad (2.5)$$

The typical granularity of the projective HCAL towers was $3.7^\circ \times 3.7^\circ$ corresponding to 4×4 ECAL towers. The obtained energy resolution was:

$$\frac{\sigma(E)}{E} = \frac{0.85}{\sqrt{E}} \quad (E \text{ in GeV}). \quad (2.6)$$

The muon system Muons were identified with the tracking capabilities of the HCAL together with the muon chambers (MUON) outside the HCAL. The average muon identification efficiency was 86%. For muons which traveled through both double layers of streamer tubes an accuracy for the track direction measurement of ≈ 10 -15 mrad could be achieved. The track helix of muons was measured by means of the tracking system.

The trigger system and the luminosity measurement The ALEPH trigger was designed to detect events stemming from e^+e^- collisions with very high efficiency. For each physics channel the trigger logic and redundancy allowed trigger efficiencies near 100% to be obtained.

The trigger decision was based on:

1. Total-energy trigger: energy deposits in the electromagnetic calorimeter
2. Electron-track trigger: track segments in the drift chambers with corresponding energy deposits in the electromagnetic calorimeter
3. Muon-track trigger: track segments in the drift chambers with corresponding energy deposits in the hadronic calorimeter
4. Back-to-back trigger: two back-to-back track segments in the drift chambers

The use of the Total-energy and Muon-track trigger to collect hadronic Z^0 events had an efficiency of $(99.99 \pm 0.01)\%$.

The luminosity was measured using Bhabba events. They were triggered with a rate of 2-3 Hz. The overall trigger rate was 4-5 Hz with Z events (at the peak) and 2-photon events contributing about 0.5 Hz each. The remaining contribution stemmed from cosmic rays, noise and beam related background.

Because of the low trigger rate no reduction of the event rate was needed to satisfy the bandwidth limitations of the ALEPH readout.

2.4 The investigation of a new b-tag for ALEPH

2.4.1 Introduction

For this study the ALEPH LEP1 data from 1992 till 1995 (≈ 3.8 million events) together with the corresponding Monte Carlo datasets (≈ 8 Million events) are used.

As reference and benchmark for the investigation of the new b-tag, the ALEPH lifetime mass tag [ALEPH97] for the measurement of R_b with the double tag method is utilized again (section 1.2). This tag uses a combination of the hemisphere lifetime probability P_H and the mass information μ_H (section 1.1.3).

Figure 2.3 shows the ALEPH event display with a typical $Z^0 \rightarrow b\bar{b}$ decay topology as illustrated in Figure 1.1(a). The upper left view shows the two reconstructed b-jets and the ALEPH detector response in the r - ϕ view. The two layers of the ALEPH vertex detector (radii: ≈ 6.3 cm and 10.8 cm) are shown in the upper right view, displaying the reconstructed hit position of tracks from this event. The lower view shows the event details obtained from well reconstructed tracks on a scale of about 1 cm. The assumption is that a b-hadron (\bar{B}_s) decayed into a c-hadron (D_s^+) and an electron (e^-). The produced c-hadron decayed further into lighter hadrons [ALE].

The description of the ALEPH Monte Carlo production and the library of the ALEPH reconstruction tools *ALPHA* can be found in [ALE] and in [ALEPH00]. The ALEPH software is written in the programming language FORTRAN [Ros67]. The reconstructed and preselected ALEPH data are stored in binary PAW NTUPLE files [BCVZ89]. To utilize object oriented C++ methods these files were converted to binary ROOT files [BR97].

The ALEPH reconstruction tool for b-hadron identification QIPBTAG, which provides an optimized track, primary vertex and jet reconstruction for b-hadron identification, is the foundation for the new combined b-tag. The details of event selection and reconstruction are sum-

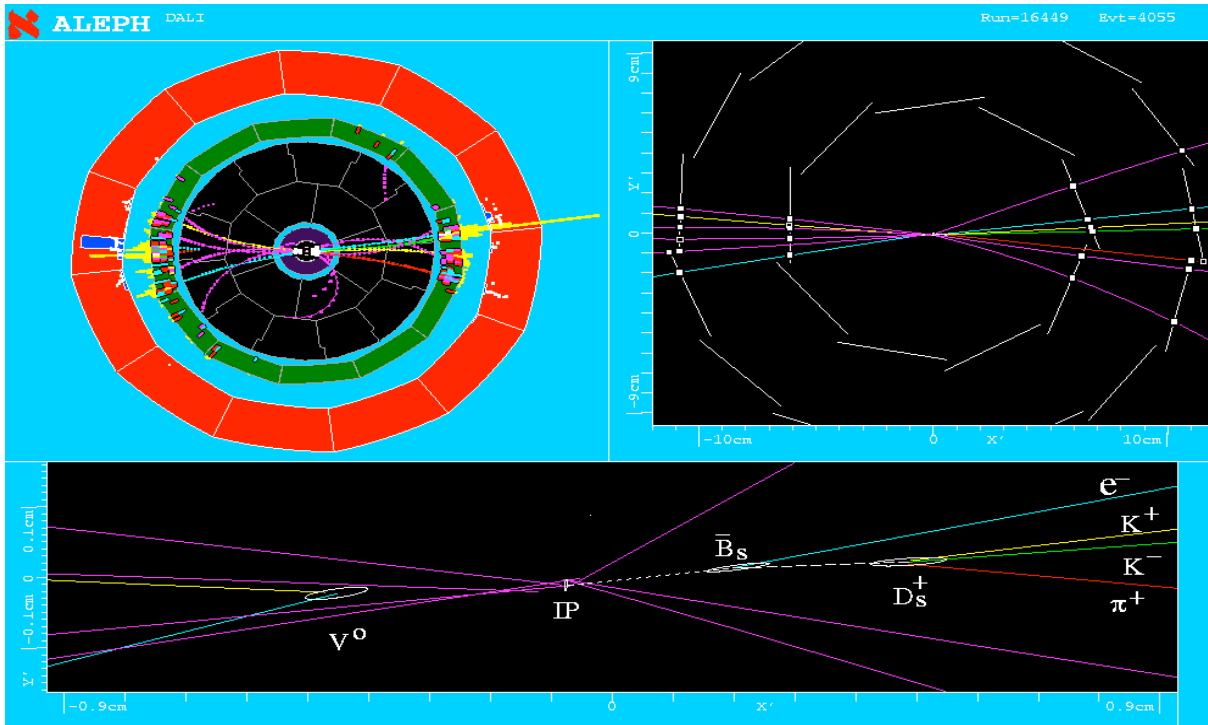


Figure 2.3: ALEPH event display showing a typical $Z^0 \rightarrow b\bar{b}$ decay.

marized in the first part of the following section. The second part describes the new combined b-tag and its performance.

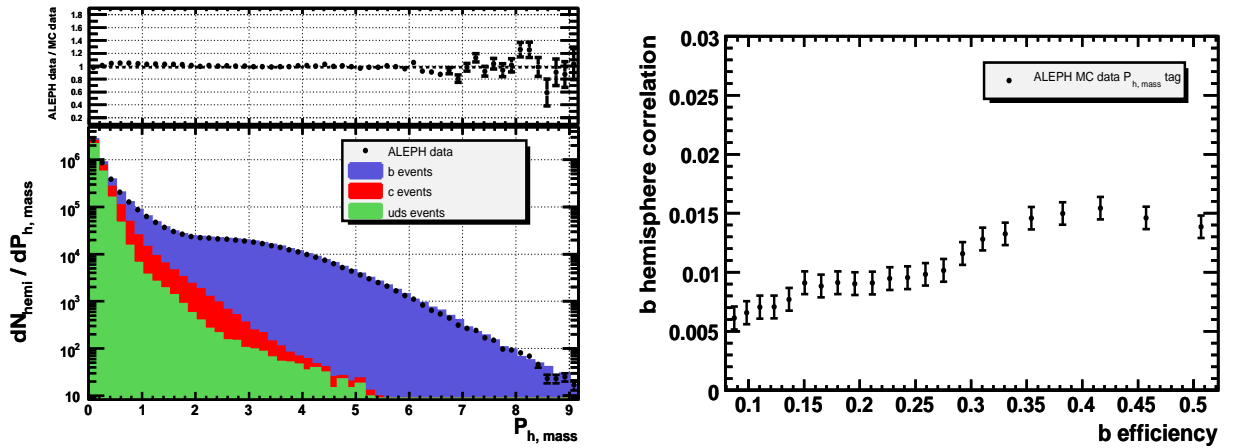
2.4.2 The benchmark b-tag: the ALEPH lifetime mass tag

The full LEP1 data sample was used for this measurement, which was published in 1997. The combined b-tag observable $P_{H,\text{mass}}$ was defined as follows:

$$P_{H,\text{mass}} = 0.7 \log_{10} \mu_H + 0.3 \log_{10} P_H \quad (2.7)$$

The description of the observables μ_H and P_H can be found in section 1.1.3. The final R_b value was measured at $P_{H,\text{mass}} > 1.9$, the point of smallest overall error on R_b . The result was: $R_b = 0.2167 \pm 0.0011(\text{stat}) \pm 0.0013(\text{syst})$ [ALEPH97].

The comparison of the data and the Monte Carlo prediction for $P_{H,\text{mass}}$ separated into the different quark flavors is shown in Figure 2.4(a). The overall agreement within the statistical errors is good, only the very high $P_{H,\text{mass}}$ bins show small discrepancies between the data and the Monte Carlo prediction. The b-hemisphere correlation (see section 1.2), which varies from 0.005 for low b-tag efficiencies to 0.015 for higher efficiencies, is shown in Figure 2.4(b).



(a) The comparison of the ALEPH data with the Monte Carlo prediction for the $P_{H,\text{mass}}$ b-tag observable.

(b) The $P_{H,\text{mass}}$ b-tag hemisphere correlation.

Figure 2.4: The ALEPH lifetime mass tag $P_{H,\text{mass}}$.

2.4.3 Event selection and reconstruction

2.4.3.1 Hadronic Z^0 -decay and track selection

In order to measure R_b , hadronic Z^0 decay modes have to be identified. The separation from non-hadronic Z^0 decay modes is done by using the invariant mass and the total multiplicity of the events [ALEPH95]. Therefore all events have to pass the standard ALEPH hadronic selection based on the observation of at least five good charged particle tracks [ALEPH93b].

In addition the selected events have to satisfy $|\cos \theta| < 0.7$ where θ is the angle between the thrust axis and the beam z -direction. This ensures that tracks from hadronic decays in both hemispheres are well contained inside the acceptance of the silicon vertex detector (VDET).

According to MC simulation, the final hadronic selection efficiency in combination with the additional constraint on the thrust axis is 61.9% with a small contamination of $(0.03 \pm 0.1)\%$, mostly from $Z^0 \rightarrow \tau^+\tau^-$ events [ALEPH97].

The QIPBTAG track selection rejects tracks with no VDET hits to ensure a reasonably good track measurement close to the primary vertex of the events (section 1.1.3). Only tracks which pass at least this selection are used in the following. Further details of the QIPBTAG track selection can be found in [ALEPH00].

2.4.3.2 Jet reconstruction

The flight direction of decaying hadrons is estimated by the QIPBTAG tool by means of the Jade jet algorithm in the E recombination scheme mode (section 1.1.2). Charged tracks and neutral energy clusters are used as jet input. The cut-off value y_{cut} is set to 0.01. Jets are separately searched for in each hemisphere. If there is more than one jet found for one hemisphere, the most energetic (leading) jet is taken. In case the jet finding was not successful for one hemisphere, the thrust axis of the event is used instead in order not to reject such an event [ALEPH00].

The resolutions of the jet variables $P_{t,\text{jet}}$, η_{jet} and ϕ_{jet} for b-hadron jets, where the leading jets are compared to the b-hadron of the respective hemisphere, are shown in Figure 2.5. The mean of the $P_{t,\text{jet}}$ -resolution shown in Figure 2.5(b) is about 10% too high. This means that on average tracks and clusters additional to those from the b-hadron were collected and grouped to a jet 4-vector by the QIPBTAG tool. This is caused *e.g.* by 4-vectors from recovered hemispheres. For such hemispheres the leading jet 4-vector is made from all clusters and tracks of this hemisphere. However, for the determination of the b-hadron flight direction the resolution of the angular coordinates η_{jet} and ϕ_{jet} is more important. As shown in Figures 2.5(d) and 2.5(f) the mean values of these distributions peak at zero with only small deviations: *e.g.* The spread of the ϕ_{jet} -resolution is about 1° .

2.4.3.3 Primary vertex reconstruction

The reconstruction of the primary vertex is done for each event hemisphere separately (section 1.2.1). The standard ALEPH primary vertex reconstruction tool QFNDIP utilized by QIPBTAG is used twice per event to reconstruct a primary vertex for each hemisphere by only using tracks contained in one hemisphere.

To determine the measurement accuracy of the primary vertex position, the reconstructed position is compared to the simulated vertex position for Monte Carlo events. The performance shown in Figure 2.5 is compatible with expectations: The y -coordinate resolution of the hemisphere primary vertex ($\approx 1 \mu\text{m}$) is better than the x -coordinate resolution ($\approx 5 \mu\text{m}$), because of the deterioration of the x -coordinate measurement due to emission of synchrotron radiation (Figures 2.5(a) and 2.5(c)). The latter is in addition a cause of the non-Gaussian tails of the x -coordinate resolution. The non-Gaussian tails of the z -coordinate resolution (Figure 2.5(e)) are caused by material interactions of charged tracks used for the primary vertex determination, *e.g.* with the beam-pipe.

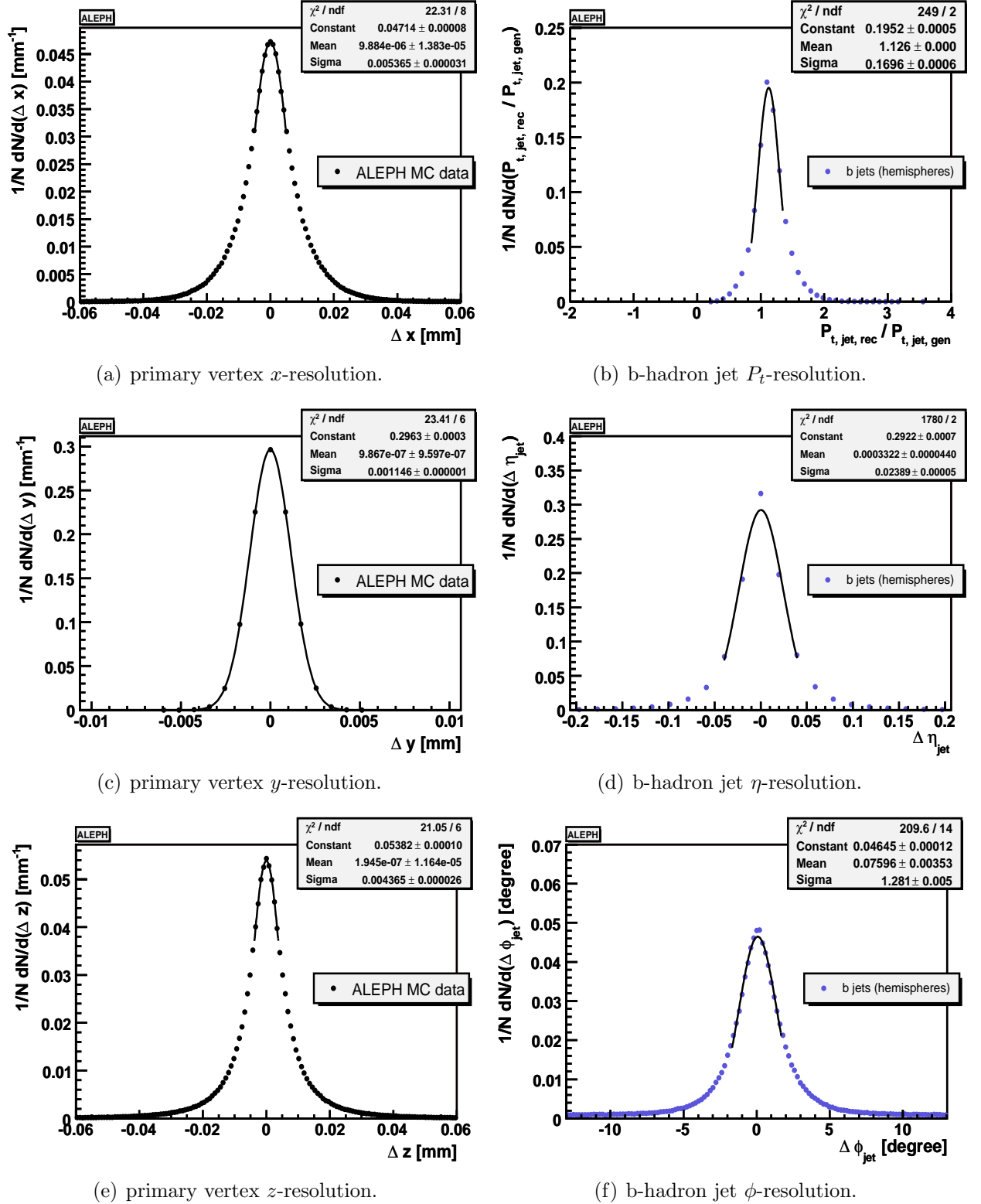


Figure 2.5: The ALEPH hemisphere primary vertex and b-hadron jet reconstruction performance. The primary vertex x -, y - and z -coordinate resolution is shown in the left column (Figures (a), (c) and (e)). The right column shows the resolution of the jet variables $P_{t,\text{jet}}$, η_{jet} and ϕ_{jet} for reconstructed b-hadron jets (Figures (b), (d) and (f)).

2.4.3.4 Secondary vertex reconstruction

The new b-tag presented in the next section is based on reconstructing a secondary vertex. The method used for this purpose is described in [ALEPH01].

Tracks are subjected to a secondary vertex fit based on their distance $\Delta R(\text{track}, \text{jet})$ from the jet axis of the most energetic jet in the hemisphere and on their probability $\mathcal{P}_{\text{b prob, track}}$ to come from a b-hadron decay. $\mathcal{P}_{\text{b prob, track}}$ is determined from a 2-dimensional grid based on the 3D impact parameter significance and the rapidity of the tracks with respect to the jet axis (section 1.1.3). The iterative fitting procedure removes tracks with high χ^2 contributions. Once a vertex is found the procedure is repeated for the remaining tracks. If more than one secondary vertex is found, the vertex with the highest track multiplicity is taken. The resulting secondary vertex finding efficiency is 92% for hemispheres containing a b-hadron [ALEPH01].

To ensure that the reconstructed vertices are not faked, *e.g.* by badly measured tracks due to material interactions, the significance $S_{2\text{nd vtx}}$ of the secondary vertices is loosely constrained: $S_{2\text{nd vtx}} > 4.5$. The significance $S_{2\text{nd vtx}}$ is defined as the 3-dimensional distance between the primary and the secondary vertex divided by the error on this measurement. Figure 2.6 shows the secondary vertex significance $S_{2\text{nd vtx}}$ and the good agreement of the ALEPH data with the Monte Carlo prediction² within the statistical errors. The distribution shows that the additional constraint $S_{2\text{nd vtx}} > 4.5$ increases the probability of the reconstructed secondary vertices to stem from a b-hadron decay. The resulting vertex finding efficiency for hemispheres containing a b-hadron is 54%.

Table 2.2 summarizes the main parameters used for the secondary vertex reconstruction, including the track momentum P_{track} , the 3D impact parameter D and its error $\sigma(D)$, given separately for tracks with one or two VDET hits, the distance along the jet S_j (see Figure 1.1(b)),

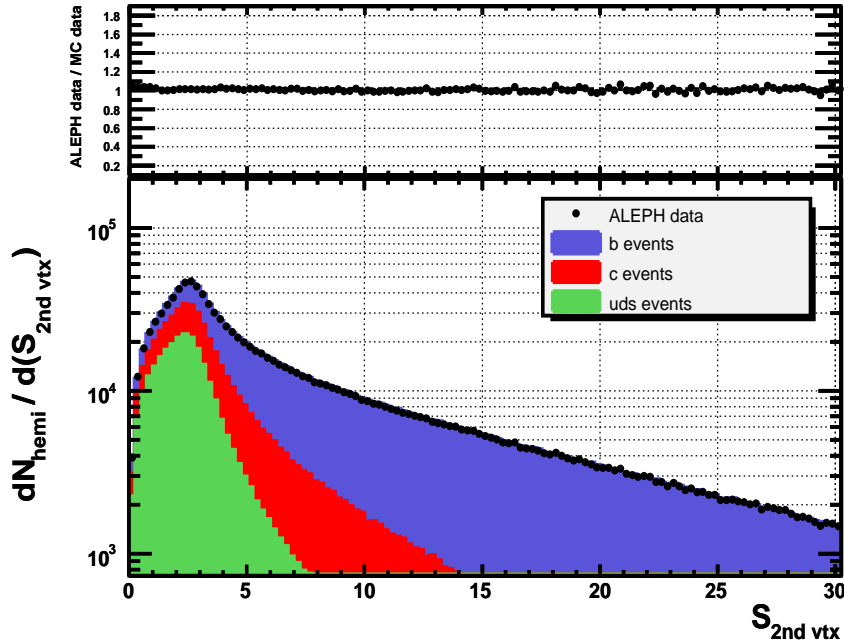


Figure 2.6: The significance $S_{2\text{nd vtx}}$ of reconstructed secondary vertices.

²The Monte Carlo prediction for the different quark flavors b, c and uds is normalized to fit the hadronic Z^0 width. The resulting 3 histograms are added up so that the sum can be compared to the measured ALEPH data. This normalization is used for all distributions showing ALEPH and Monte Carlo data.

etc.

$P_{\text{track}} > 0.4 \text{ GeV}$ $\Delta R(\text{track}, \text{jet}) < 0.7$ $-0.75 < D < 1.00 \text{ [cm]}$ $\sigma(D_{1 \text{ VDET hit}}) < 0.08 \text{ cm}$ $\sigma(D_{2 \text{ VDET hits}}) < 0.04 \text{ cm}$ $S_j > -2 \text{ cm}$ $\mathcal{P}_{\text{b prob, track}} > 0.7$ $\chi^2 < 5$ $S_{2\text{nd vtx}} > 4.5$

Table 2.2: Summary of the main parameters used for the secondary vertex reconstruction.

2.4.3.5 Association of tracks with the secondary vertex

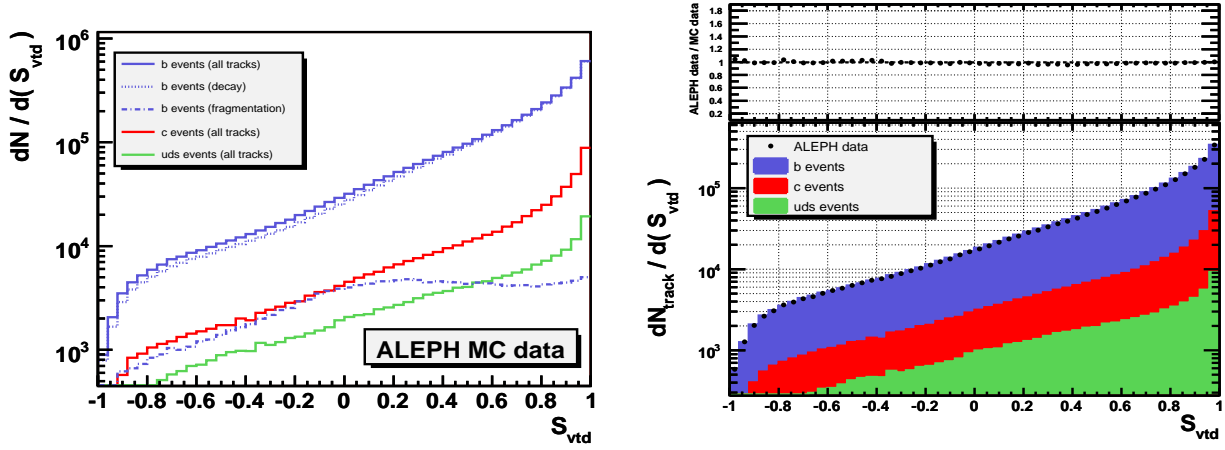
The secondary vertex finding procedure as described before is optimized to determine the secondary vertex position with high precision. To achieve this goal a sample of tracks optimized for this special task is used (table 2.2). Furthermore the iterative procedure removes a fraction of real secondary vertex tracks to improve the χ^2 of the secondary vertex fit. Usually the reconstructed tracks used for the final fit of the secondary vertex candidate are defined as *secondary vertex tracks*.

To recover some of the rejected tracks and to re-consider them for the b-hadron identification the significance of the vertex-track distance S_{vtd} is used to separate the determination of secondary vertex tracks from the secondary vertex reconstruction. The increased number of secondary vertex tracks leads to b-tag observables based on more available information, which in turn leads to an improved b-hadron identification. The S_{vtd} observable is defined as follows:

$$S_{\text{vtd}} := \frac{S_{\text{track}, 1\text{st vtx}} - S_{\text{track}, 2\text{nd vtx}}}{S_{\text{track}, 1\text{st vtx}} + S_{\text{track}, 2\text{nd vtx}}} \quad (2.8)$$

For each track which passed the QIPBTAG track selection and which is contained in the leading jet, the significance of its distance with respect to the primary vertex $S_{\text{track}, 1\text{st vtx}}$ and with respect to the secondary vertex $S_{\text{track}, 2\text{nd vtx}}$ is used to calculate the S_{vtd} observable for this track. The significance of a track with respect to a vertex is defined as the 3-dimensional distance between the track and the vertex divided by the measurement error of this distance. A track is defined to be inside the leading jet if $\Delta R(\text{track}, \text{jet}) < 0.7$.

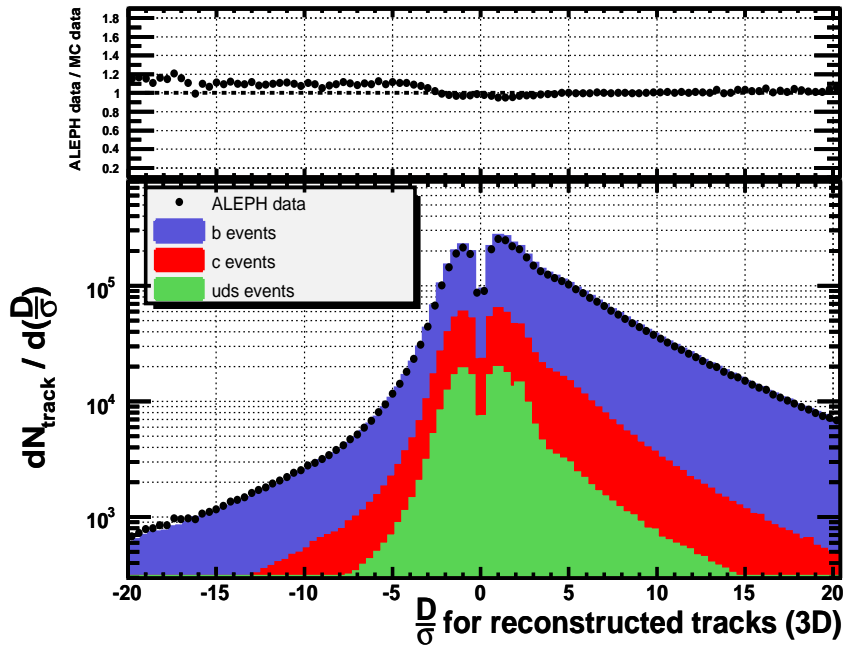
The S_{vtd} observable peaks at -1 for tracks stemming from the primary vertex and at +1 for secondary tracks. The observable for different tracks originating from different production mechanisms is shown in Figure 2.7(a). The discriminating power of the S_{vtd} is clear when comparing the tracks from the b-hadron decay, which are peaked at +1, with the b-hadron fragmentation tracks accompanying the b-hadron, which show an almost flat distribution. It follows that the significance S_{vtd} separates tracks stemming from different vertices reasonably well. The good agreement within the statistical errors of the ALEPH data with the Monte Carlo prediction for the S_{vtd} observable is shown in Figure 2.7(b). Tracks for which $S_{\text{vtd}} > 0$ are defined as secondary vertex tracks.

(a) The significance S_{vtd} for different track types.(b) The significance S_{vtd} for ALEPH data in comparison with the Monte Carlo prediction.Figure 2.7: The S_{vtd} observable for tracks which pass the QIPBTAG track selection.

2.4.3.6 Significance of the signed 3D impact parameter

The significance of the signed 3D impact parameter $\frac{D}{\sigma}$ for reconstructed tracks is used to calculate the jet lifetime probability P_{jet} as described in section 1.1.3. Because the jet lifetime probability P_{jet} is used in the new ALEPH b-tag the description of the ALEPH data by the Monte Carlo prediction for the significance $\frac{D}{\sigma}$ is investigated.

Figure 2.8 shows that the Monte Carlo prediction is in reasonable agreement with the ALEPH data, except for values of $\frac{D}{\sigma} < -2$ where small deviations of 10-15% become visible. The lifetime signal for b-hadron decays (blue) is clearly visible for positive $\frac{D}{\sigma}$.

Figure 2.8: The significance of the 3D impact parameter $\frac{D}{\sigma}$ of reconstructed tracks.

2.4.4 Description and performance of the new ALEPH b-tag

The b-tag observables and their performance presented here are calculated by means of specialized C++ objects embedded in the ROOT analysis framework. Details of the implementation can be found in Appendix B.

2.4.4.1 Observables used in the new ALEPH b-tag

The new ALEPH b-tag developed utilizes for a given jet a combination of b-tag observables as described in section 1.1.3. The observables used are: the jet lifetime probability P_{jet} , the secondary vertex mass $M_{2\text{nd vtx}}$, the rapidity $\eta_{2\text{nd vtx}, \text{track}, \text{jet}}$ and the jet energy fraction X_{energy} of secondary vertex tracks³. Their distributions after requiring a well reconstructed secondary vertex (section 2.4.3.4) are shown in Figures 2.9(a)-(d). All four observables are reasonably

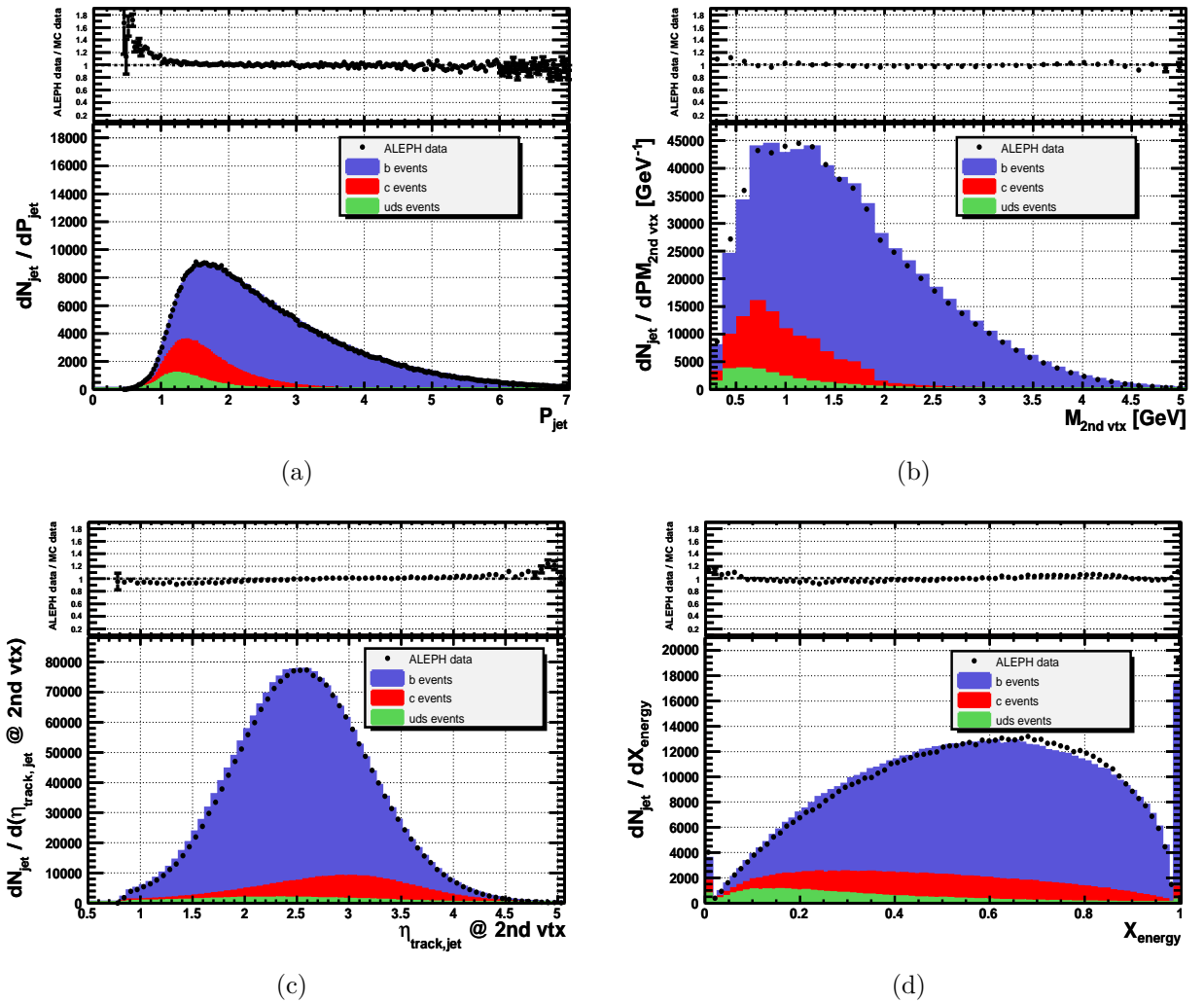


Figure 2.9: The jet lifetime probability P_{jet} (a), the secondary vertex mass $M_{2\text{nd vtx}}$ (b), the rapidity $\eta_{2\text{nd vtx}, \text{track}, \text{jet}}$ of secondary vertex tracks (c) and the jet energy fraction X_{energy} of secondary vertex tracks (d) for the ALEPH data and the Monte Carlo prediction, separated for different quark flavors.

³The non-zero mass of the charged tracks used in the calculations is determined by means of the same methods also utilized by the ALEPH QBMTAG tool [ALEPH00].

well described by the Monte Carlo, however small deviations up to 5% are visible for some bins. Larger deviations up to 10% occur only for bins with low statistics at the edges of the distributions, which are negligible in conjunction with the usage of this observables for the b-hadron identification.

The different response for the different hadron jet types is clearly visible. Due to the requirement of a significant secondary vertex a considerable amount of uds events has already been rejected. By combining all observables it is possible to further improve the purity of the b-hadron jet selection, as described in the next section.

2.4.4.2 Combination of the observables used in the new ALEPH b-tag

By means of equation (1.9) the 4 observables P_{jet} , $M_{2\text{nd vtx}}$, $\eta_{2\text{nd vtx, track, jet}}$ and X_{energy} are combined. For this purpose the corresponding probability density functions of these observables have to be calculated for b-, c- and uds-hadrons separately.

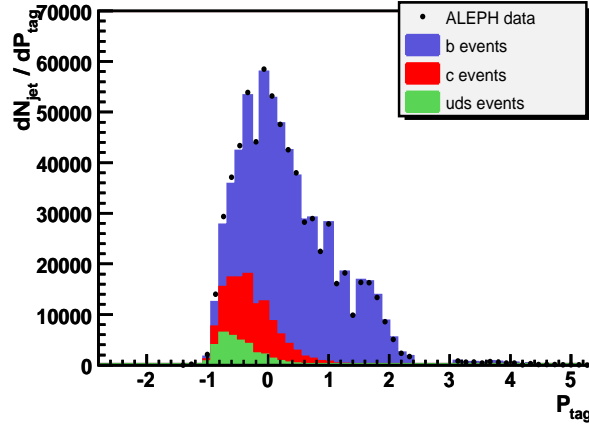
To illustrate the increase of the b-tag performance as each new observable is added to the combination four different tags are defined. All tags are based on having a well reconstructed secondary vertex for the jet:

- The jet lifetime probability b-tag P_{tag} is constructed from the probability density functions of the jet lifetime probability P_{jet} only.
- The PM_{tag} b-tag uses the jet lifetime probability P_{jet} and the secondary vertex mass $M_{2\text{nd vertex}}$.
- The PMR_{tag} b-tag utilizes the jet lifetime probability P_{jet} , the secondary vertex mass $M_{2\text{nd vertex}}$ and the rapidity $\eta_{2\text{nd vtx, track, jet}}$ of secondary vertex tracks.
- The PMRX_{tag} b-tag uses all four observables: the jet lifetime probability P_{jet} , the secondary vertex mass $M_{2\text{nd vertex}}$, the rapidity $\eta_{2\text{nd vtx, track, jet}}$ and the energy fraction X_{energy} of secondary vertex tracks.

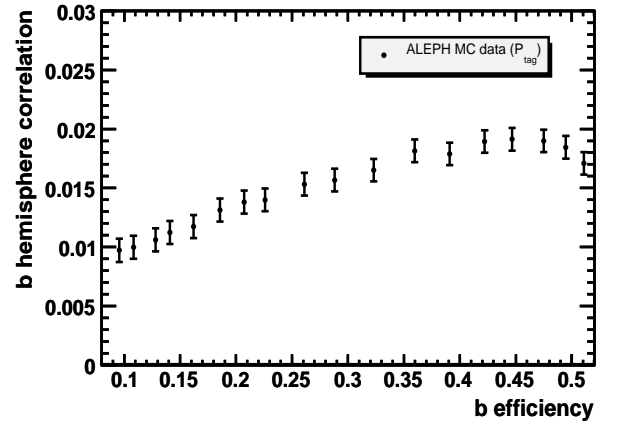
The comparison of the ALEPH data with the Monte Carlo prediction for the four tags is shown in Figures 2.10(a), 2.10(c), 2.10(e) and 2.11(a). All b-tag observables are reasonably well described by the Monte Carlo prediction within the statistical errors, *e.g.* for the PMRX_{tag} b-tag the difference between data and Monte Carlo is only a few per cent, except for the bins with $\text{PMRX}_{\text{tag}} < -2$ where deviations become bigger. However the working point for the R_b measurement would clearly be chosen at considerably higher PMRX_{tag} values.

The b-hemisphere correlation of the different b-tags is shown in Figures 2.10(b), 2.10(d), 2.10(f) and 2.11(b). All correlations are quite small. For the PMRX_{tag} b-tag the correlation lies between 0.005 and 0.015 (Figure 2.11(b)), comparable to that of the reference ALEPH lifetime mass tag (Figure 2.4(b)).

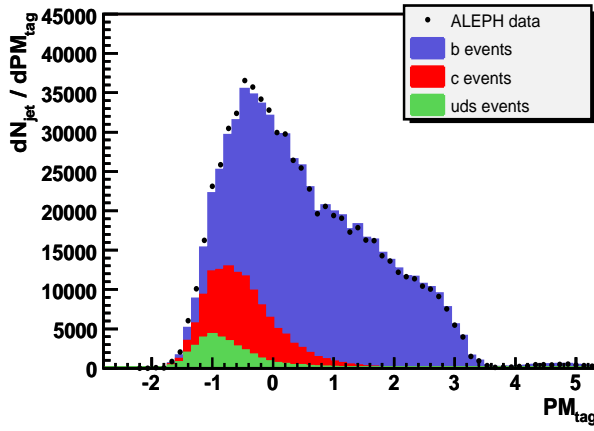
The comparison of b-efficiency versus the corresponding b-purity for the four tags is shown in Figure 2.12. All b-tags have the point (b-efficiency, b-purity) of (54%, 81%) in common because of the initial constraint of the well reconstructed secondary vertex. Starting from the initial point (54%, 81%) the use of more b-tag observables in combination decreases the b-efficiency and at the same time increases the b-purity. For the same b-purity, the b-efficiency increases for each additional observable used in the combination. The PMRX_{tag} has the best performance.



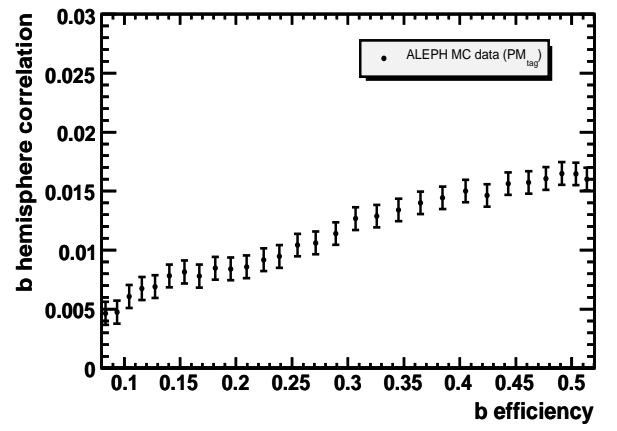
(a)



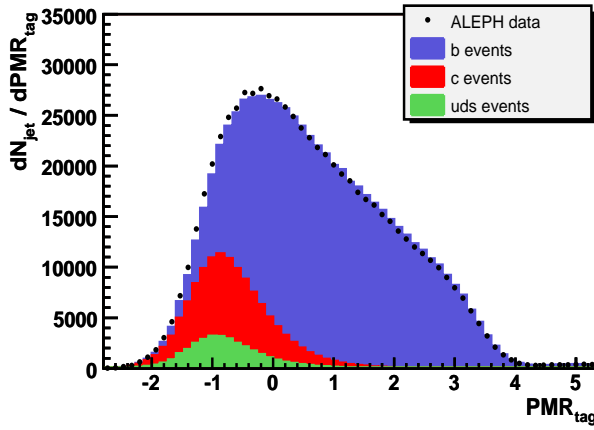
(b)



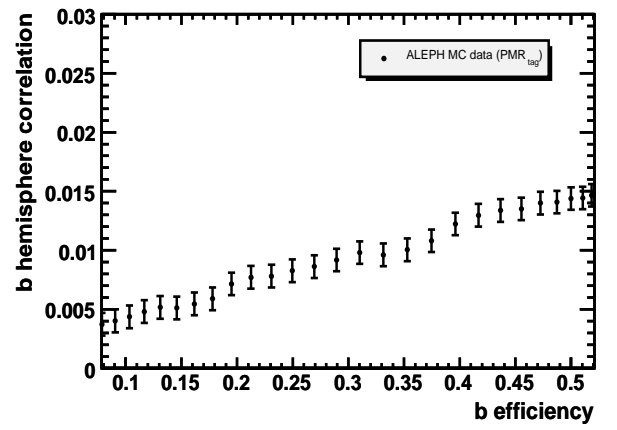
(c)



(d)



(e)



(f)

Figure 2.10: Comparison of the ALEPH data with the Monte Carlo prediction for the P_{tag} , PM_{tag} and PMR_{tag} b-tag observables (left column) and the corresponding b-hemisphere correlations (right column).

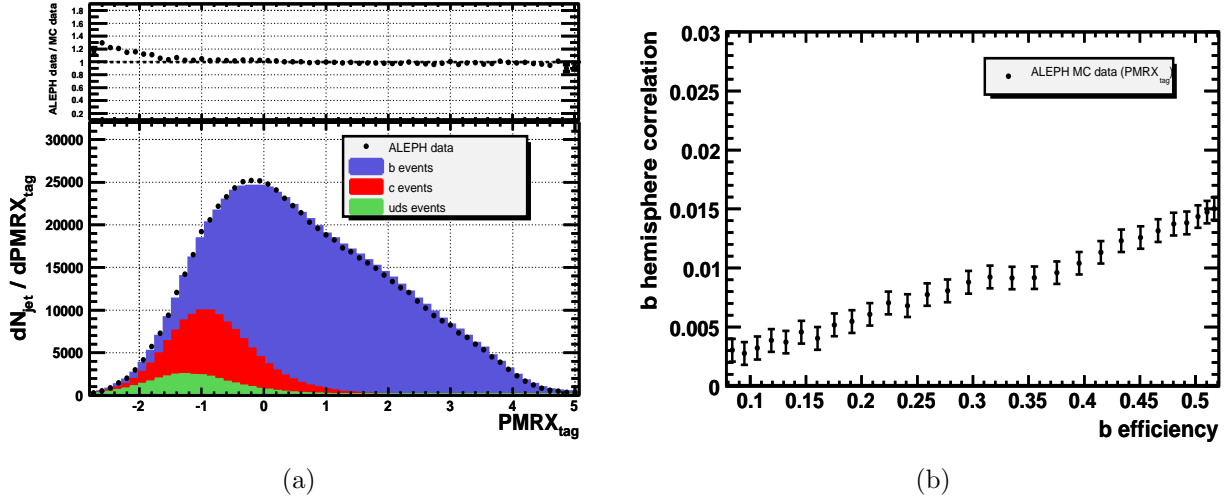


Figure 2.11: Comparison of the ALEPH data with the Monte Carlo prediction for the PMRX_{tag} b-tag observable (a) and the corresponding b-hemisphere correlation (b).

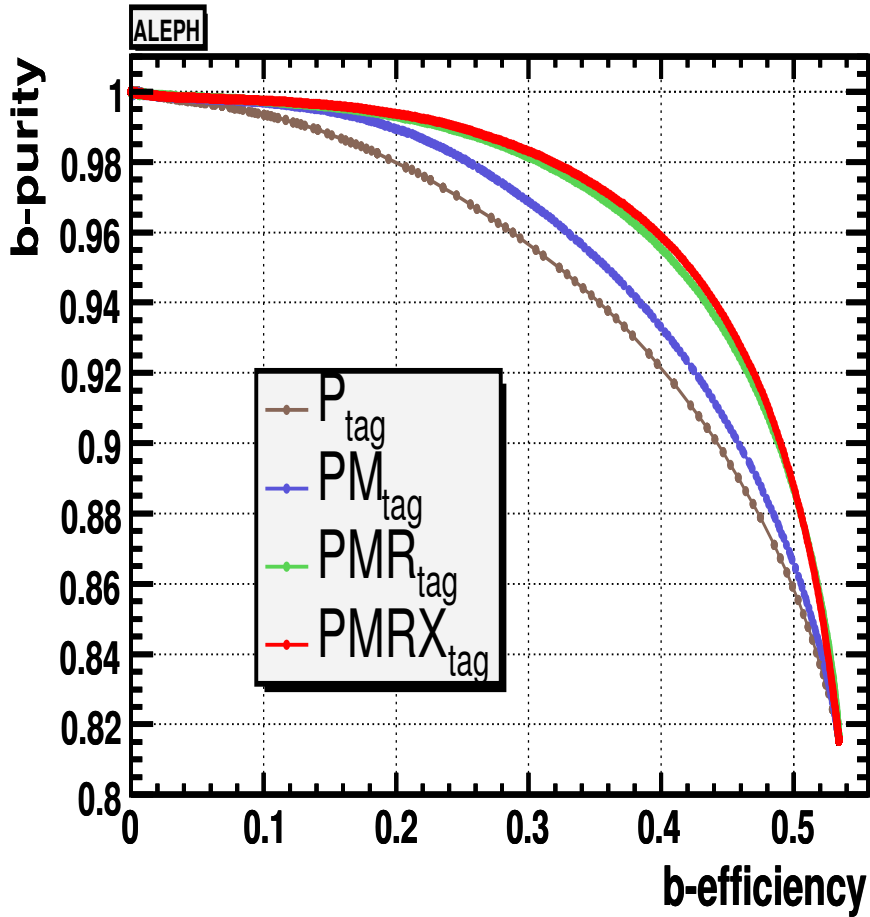


Figure 2.12: Comparison of the b-purity versus b-efficiency for the different b-tags, each employing in combination a different number of observables.

2.4.5 Summary

To put the PMRX_{tag} b-tag into perspective, Figure 2.13 shows a comparison of working points of recent R_b measurements done by the LEP experiments and the SLD collaboration. Experiments featuring good track reconstruction with a vertex detector (ALEPH, DELPHI, SLD) are listed in the high b-efficiency/b-purity region.

The original working point of the ALEPH lifetime mass tag at (b-efficiency, b-purity) of (22%, 98%) [ALEPH97] is slightly off the b-efficiency versus b-purity curve newly computed here for the study of the benchmark b-tag (dark blue line). The identical b-tag definition results in a slightly improved overall performance, because the reconstruction of the ALEPH data, *e.g.* the accuracy of the track reconstruction used in the benchmark study, was improved after the original measurement was published 1997 [Sch].

The SLD detector had the best b-tagging capability. Still the best single R_b measurement (DELPHI mult) was done by the DELPHI collaboration [DELPHI99]. This analysis used the complete LEP1 data with a b-tag operating at (32%, 98%), which was the point of smallest overall error calculated from the statistical and systematical errors. SLD did not have a competitive number of collected events in comparison to the LEP experiments, so that the statistical error downgraded their overall result [SLD96, EWWG05].

The PMRX_{tag} b-tag is able to operate in the same b-efficiency and b-purity region as the

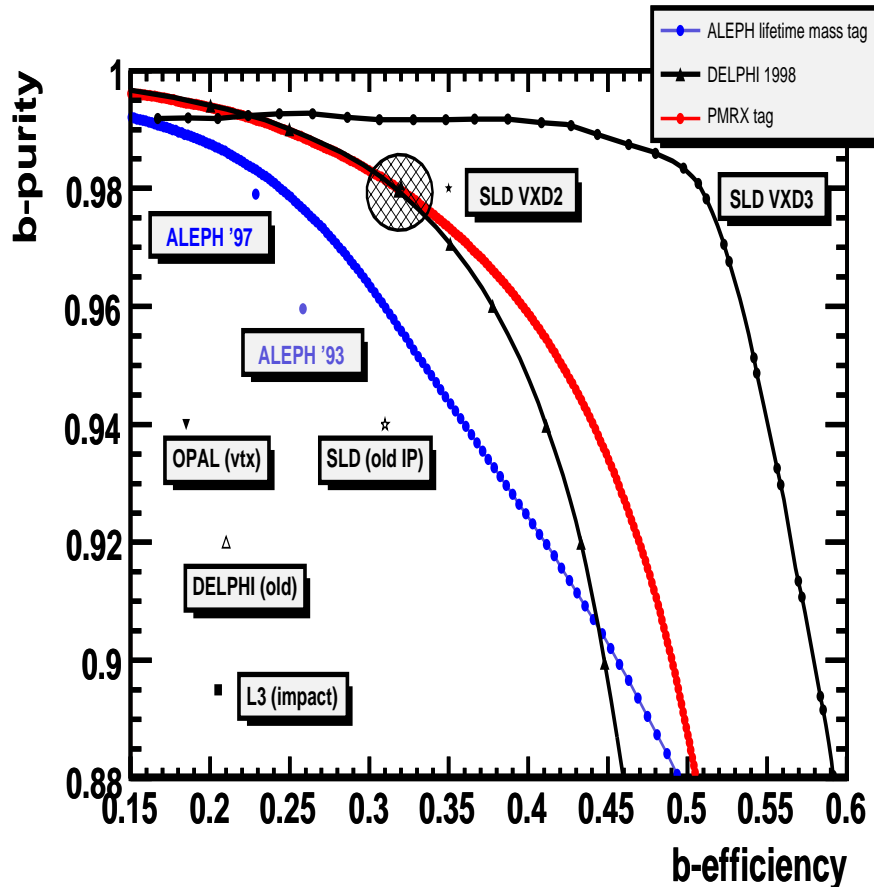


Figure 2.13: Comparison of working points of recent R_b measurements done by the LEP experiments and SLD [ALEPH93a, ALEPH97, EWWG05]. The black circle marks the working point of the best single R_b measurement performed by DELPHI [DELPHI99].

DELPHI mult measurement. In Figure 2.13 this region is marked with a black circle around the working point (32%,98%). It turns out that at the same purity of about 98% the PMRX_{tag} b-tag increases the b-efficiency from 22% to 32% with respect to the working point of the ALEPH analysis performed 1997.

The R_b value versus the b-tag efficiency ϵ_b for the PMRX_{tag} b-tag calculated with the double tag method (section 1.2.1) is shown in Figure 2.14. Within the (statistical) errors the obtained R_b values support the actual world average value of $R_b = 0.21629 \pm 0.00066$ [EWWG05] shown by the red dashed line.

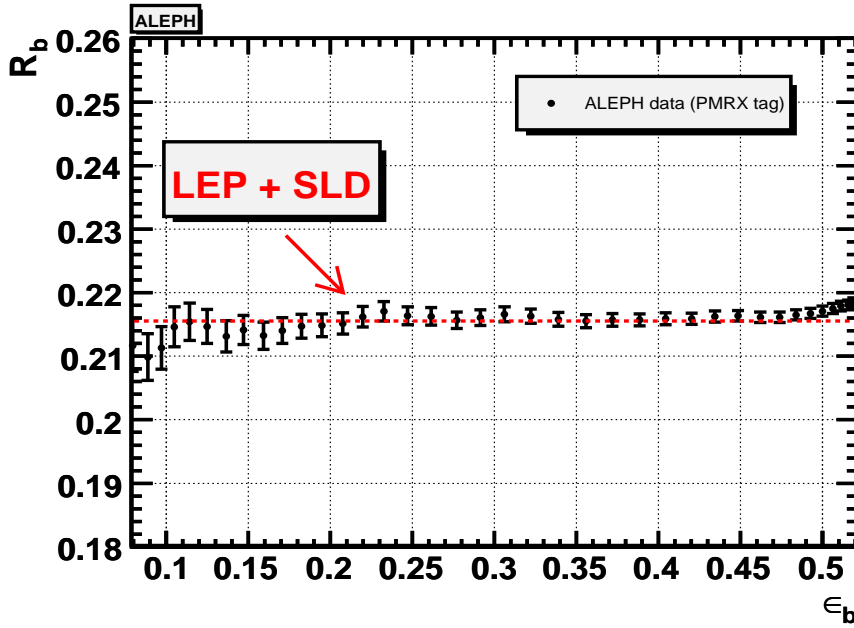


Figure 2.14: The value of R_b as a function of the b-efficiency ϵ_b for the PMRX_{tag} b-tag. The errors shown are statistical errors only. The red dashed line highlights the combined LEP+SLD measurement [EWWG05].

Chapter 3

Hadron collider physics

Contents

3.1	Introduction	39
3.2	The LHC pp collider	39
3.3	The CMS experiment at LHC	42
3.3.1	The magnet	44
3.3.2	The inner tracking system	44
3.3.3	The muon system	45
3.3.4	The calorimeter system	47
3.3.5	The trigger system	50
3.3.6	The luminosity measurement	51
3.4	b-hadron identification in CMS	52
3.4.1	Full simulation of CMS events	52
3.4.2	Charged track reconstruction	53
3.4.3	Vertex reconstruction	58
3.4.4	Missing energy reconstruction	63
3.4.5	Jet reconstruction	63
3.4.6	b-hadron identification	68
3.4.7	Physics Analysis of WH, $H \rightarrow b\bar{b}$	77
3.4.8	Summary	81

3.1 Introduction

Events with b quarks are interesting at the CERN Large Hadron Collider (LHC) because they may provide part of the signature for new physics (*e.g.* $H \rightarrow b\bar{b}$) as well as being a tag for top decays, which are of interest themselves but also present the single largest background in the search for electroweak symmetry-breaking phenomena. The identification of b-hadron jets in the Compact Muon Solenoid (CMS) in the presence of huge backgrounds from pileup events is made possible by the excellent charged-particle tracking capabilities of the inner tracking system. Jets are reconstructed using the electromagnetic and hadron calorimeters and can be further improved by precise momentum measurement of the charged tracks in the 4 T magnetic field.

As noted previously in section 1.3.2 the associated Higgs production can be utilized to increase the signal significance of the $H \rightarrow b\bar{b}$ signature. An investigation of the discovery potential of the WH production for the CMS experiment is described in this chapter.

The analysis is based on the ROOT C++ object-oriented analysis framework [BR97].

3.2 The LHC pp collider

The LHC machine will be installed in the tunnel where LEP was located until 2000 [LHC95]. It will accelerate two beams moving in opposite directions. An overall view of the LHC setup is shown in Figure 3.1(a).

Proton beams will be prepared by CERN's existing accelerator chain (Figure 3.1(b)) before being injected into the LHC, where they reach their final energy of 7 TeV. First operation is planned for winter 2007 [Lam05, LHCa]. The LHC can be operated as a heavy-ion collider, reaching a centre-of-mass energy of 1148 TeV for lead-ion collisions [LHC95]. This offers the possibility to study the physics of strongly interacting matter at extreme energy densities, where the formation of a new phase of matter, the quark-gluon plasma, is expected.

The LHC will exactly follow the geometry of LEP. The number of beam crossing points has been reduced from eight to four, in order to reduce costs and to better optimize the utility insertions containing RF, collimation and beam dump systems. The two high-luminosity interaction regions are located at diametrically opposite straight sections: the ATLAS experiment [ATL] is located at point 1 and the CMS experiment [CMS94] at point 5. ATLAS and CMS are both general purpose experiments that will explore the fundamental nature of matter and the basic forces that shape our universe utilizing the collisions of protons of extraordinarily high energy delivered by the LHC. Two other experiments are located at point 2 (ALICE) and point 8 (LHCb). The ALICE detector is dedicated to study heavy-ion collisions [ALI], whereas LHCb is an optimized detector for B meson physics to study the CP violation [LHCb]. The detectors of the TOTEM experiment will be integrated into the CMS environment. TOTEM will measure the total pp cross section, elastic scattering and diffractive dissociation at the LHC [TOT].

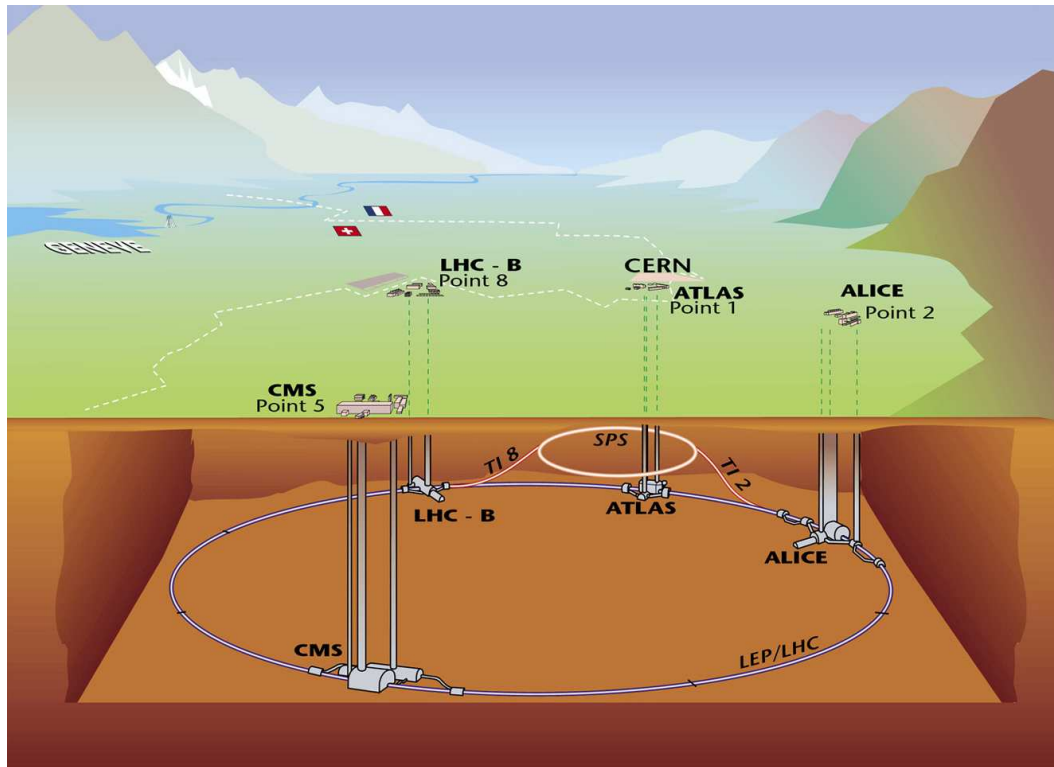
In order to reach high luminosities the LHC beams are made up of a large number of closely spaced bunches (2835 bunches spaced 7.5 m apart). About every 25 ns there is a bunch crossing. The luminosity is given by [LHC95]:

$$\mathcal{L} := \frac{N^2 n_b f \gamma}{4 \pi \epsilon_n \beta^*} F, \quad (3.1)$$

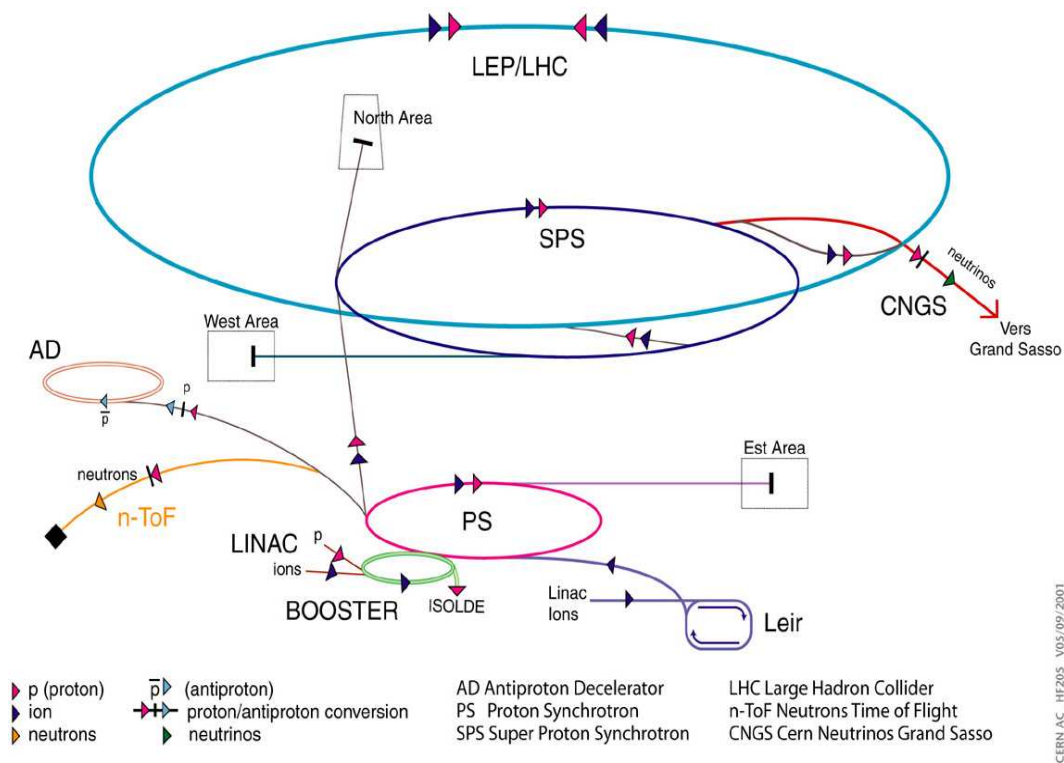
where:

- N := # protons in a bunch,
- n_b := # bunches per beam,
- f := revolution frequency,
- γ := relativistic factor,
- ϵ_n := normalized transverse emittance (has a design value of $3.75 \mu\text{m}$ [CMS06b]),
- β^* := value of the betatron function,
- F := reduction factor caused by the crossing angle of the beams (≈ 0.9).

The design luminosity of $\mathcal{L} = 10^{34} \text{cm}^{-2} \text{s}^{-1}$ leads to $\mathcal{O}(10^9)$ events/s. However during the first full year of physics running, the LHC should reach a peak luminosity of $\mathcal{L} = 2 \cdot 10^{33} \text{cm}^{-2} \text{s}^{-1}$ (low-luminosity phase). The integrated luminosity will likely be about 5fb^{-1} [CMS06b].



(a) Overall view of the LHC setup with the experiments ALICE, ATLAS, CMS and LHCb.



(b) The LHC with pre-accelerators.

Figure 3.1: The LHC experimental setup.

The LHC beam line constraint The expected r.m.s. beam size¹ for the LHC interaction areas at Point 1 (ATLAS) and Point 5 (CMS) is shown in Table 3.1:

Coordinate	r.m.s. beam size
x	15.9 μm
y	15.9 μm

Table 3.1: Parameters of the LHC interaction area [LHC95].

As for the LEP experiments these beam parameters can be used to improve the measurement of the x - and y -position of the primary vertex (beam line constraint).

Since protons are much heavier than electrons, the beam energy loss due to the emission of synchrotron radiation is much lower for LHC than for LEP. Thus, in contrast to LEP, the deterioration of the measurement of the beam x -position is negligible in the LHC case (see section 2.2).

3.3 The CMS experiment at LHC

The CMS experiment is located at Point 5 near Cessy (France). The detector has a configuration very similar to the ALEPH detector (section 2.3). Some important features of the CMS design are:

- Overall diameter: 14.60 m, overall length: 21.60 m)²
- CMS has a large magnetic field of 4 T.
- The hadron calorimeter is inside the magnetic field.
- A preshower detector consisting of two lead/silicon detector layers is placed in front of the endcaps of the electromagnetic calorimeter (ECAL).
- The excellent CMS muon system is able to identify muons and to measure their momentum with high accuracy.
- The trigger system consists of two steps: L1 (level one) and HLT (high level trigger). It is used to reduce the event rate to a level which fills the readout bandwidth with samples of physically pure events

Figure 3.2 shows the layout of the CMS main subdetectors. Charged particles are tracked by a pixel detector located closest to the nominal interaction vertex and by a silicon microstrip tracker. In addition, the iron yoke of the magnet is instrumented with the muon chamber system providing identification and tracking of muon candidates. The energy of neutral and charged particles is measured with electromagnetic and hadron calorimeters, which are located inside the coil. The CMS subdetectors and their (expected) performance are briefly described in the following sections [CMS94].

¹The coordinate system is defined at the interaction point with positive x towards the centre of the LHC ring, y vertically upwards and z along the beam direction towards the Jura mountains as seen from LHC Point 5.

²The very forward calorimeters are excluded.

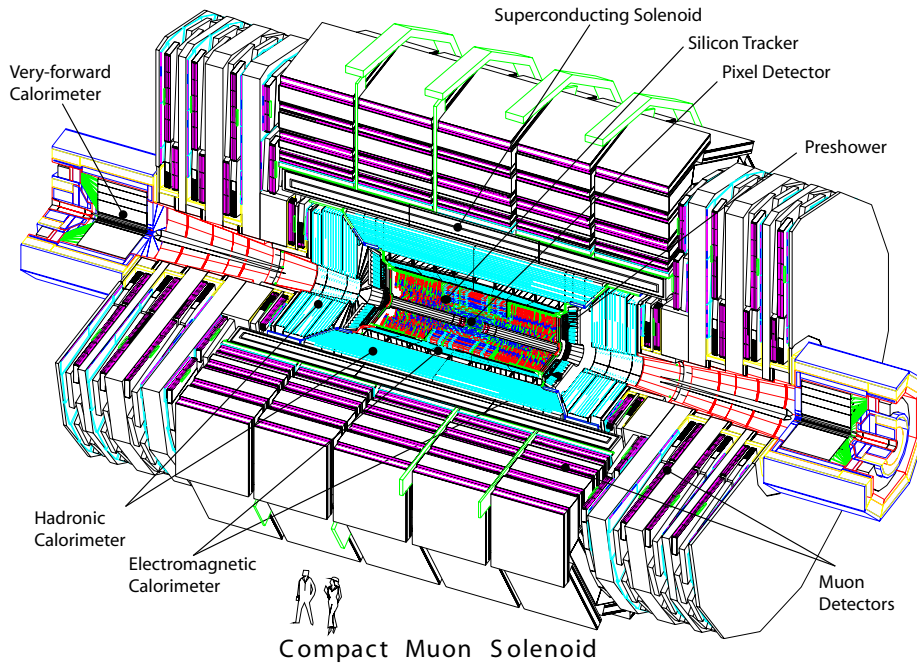


Figure 3.2: The CMS detector.

3.3.1 The magnet

The CMS magnet is a superconducting solenoid with a length of 13 m and inner diameter of 5.9 m, producing an axial field of 4 T. The magnetic flux is returned via a 1.8 m thick saturated iron yoke [CMS97c]. The strong bending power of the magnet allows obtaining a good momentum resolution for charged particles. The main design parameters of the magnet are listed in Table 3.2:

Field	4 T
Inner Bore	5.9 m
Length	12.9 m
Number of Turns	2168
Current	19.5 kA
Stored energy	2.7 GJ
Hoop stress	64 atm

Table 3.2: Main design parameters of the CMS superconducting solenoid [CMS97c].

3.3.2 The inner tracking system

The CMS silicon tracker is situated inside the 4 T magnetic field and consists of silicon pixel and microstrip tracking devices for charged particle detection. The layout of the inner tracker is shown in Figure 3.3. In addition to provide a very good tracking performance, the chosen design seeks to optimize the material budget in order to minimize the deterioration of the electromagnetic calorimeter resolution (photon conversions, nuclear interactions, bremsstrahlung) and the

The microstrip tracker provides 10 to 14 two-dimensional measurements with a resolution of about $30\ \mu\text{m}$.

Physics performance The expected momentum resolution of charged particle tracks has been estimated with the full detector simulation and is summarized in Table 3.3. The performance is consistent with the design requirements.

Observable	P_t [GeV]	Resolution
$\Delta P_t/P_t$	1	0.6-2.0 %
$\Delta P_t/P_t$	10	0.7-2.0 %
$\Delta P_t/P_t$	100	1.5-7.0 %
$\Delta \phi$	1	2.0-5.0 mrad
$\Delta \phi$	10	0.3-0.5 mrad
$\Delta \phi$	100	< 0.2 mrad
$\Delta (\cot(\theta))$	1	$3.0-12.0 \cdot 10^{-3}$
$\Delta (\cot(\theta))$	10	$0.6-3.0 \cdot 10^{-3}$
$\Delta (\cot(\theta))$	100	$0.5-1.5 \cdot 10^{-3}$
Δd_0	1	100.0-200.0 μm
Δd_0	10	20.0-30.0 μm
Δd_0	100	$< 10\ \mu\text{m}$
Δz_0	1	100.0-1000.0 μm
Δz_0	10	50.0-150.0 μm
Δz_0	100	50.0-70.0 μm

Table 3.3: The estimated resolution for high P_t (single muon) tracks for the five track parameters: P_t , ϕ , $\cot(\theta)$, d_0 and z_0 [CMS98, CMS05a].

For tracks inside dense jets, the intended reconstruction efficiency is compared with the global reconstruction efficiency resulting from full detector simulation in Table 3.4. The performance for muons meets the design goals over the full η range but the reconstruction efficiency for charged hadrons (pions) is lower than planned. For the endcap regions ($|\eta| > 1.479$) the track reconstruction efficiency is about 10-20% below the design goal due to the reduced acceptance by the endcap disks especially in the pixel system [CMS05a]. The fraction of tracks which cannot be associated to simulated tracks (fake rate) is well below 1% and depends only weakly on η .

Track source	efficiency	
	Design	Full simulation
muon, $P_t > 1\ \text{GeV}$	$> 98\%$	realized
pion, $P_t = 1\ \text{GeV}$	$\approx 85\%$	$\approx 85\%$, $ \eta < 1.2$ $< 80\%$, $ \eta > 1.2$
pion, $P_t = 10\ \text{GeV}$	$\approx 95\%$	$\approx 90\%$, $ \eta < 1.2$ $< 90\%$, $ \eta > 1.2$

Table 3.4: The design goals and the results from full simulation of the global track reconstruction efficiency for muons and charged hadrons (pions) [CMS98, CMS05a].

3.3.3 The muon system

The iron return yoke of the superconducting magnet is instrumented with muon stations. The muon system utilizes three different technologies to meet the design goals which are [CMS97d]:

- to identify muons
- to measure the muon momentum
- to trigger on muons utilizing redundant measurements

A quadrant of the muon system layout is shown in Figure 3.5. In the barrel region ($0 < \eta < 1.3$) drift tubes (DT) are used. Cathode strip chambers (CSC) are used in the endcap region ($0.9 < \eta < 2.4$). The barrel and the endcap region ($0 < \eta < 2.1$) are both covered by resistive plate chambers (RPC), which serve for triggering purposes. A shortfall of funds has led to the staging of the chambers situated beyond $|\eta| > 1.6$ (Reduced RE system). [CMS06b]

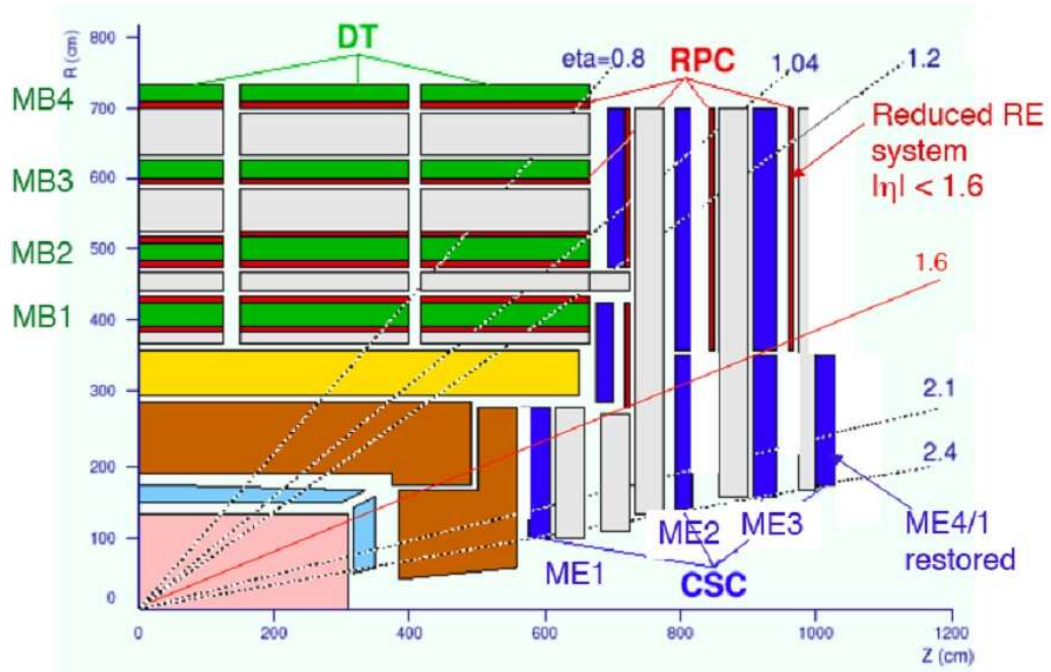


Figure 3.5: The r - z view of one quarter of the CMS muon system. The DT detectors in the barrel and the CSC detectors in the endcap are complemented with RPC detectors. The RPC chambers beyond $|\eta| > 1.6$ are staged (Reduced RE system).

Physics performance The muon reconstruction efficiency inside the detector acceptance is typically 95-99%, except for regions where the muon traverses cracks in the geometry ($|\eta| \approx 0.25, 0.75, 1.2$).

Table 3.5 summarizes the muon reconstruction performance. The P_t -resolution of muons obtained with the muon system only improves significantly (\approx factor 10) by using in addition measurements from the silicon tracker. The improvement in the muon η - and ϕ -resolution when combining the muon system and silicon tracker measurements is even bigger. However, the resolution for low- P_t muons is limited by multiple scattering, mainly in the iron yoke, and for high- P_t muons by the resolution of the muon chamber system [CMS05e].

Observable	P_t [GeV]	Resolution	
		Muon standalone	Muon + silicon tracker
$\Delta P_t/P_t$	< 100	8.0-25.7 %	0.8-3.4 %
$\Delta P_t/P_t$	1000	16.2-2.3 %	5.0-13.0 %
$\Delta \eta$	< 1	< 0.005	< 0.0004
$\Delta \eta$	10	< 0.015	< 0.0006
$\Delta \phi$	< 1	< 10.0 mrad	< 0.2 mrad
$\Delta \phi$	10	< 30.0 mrad	< 0.6 mrad

Table 3.5: The CMS muon track resolution of P_t , η and ϕ for different P_t ranges [CMS05e].

3.3.4 The calorimeter system

The CMS calorimeter system (Figure 3.6) measures the energy and direction of particle jets as well as the missing transverse energy. To achieve this goal an electromagnetic (ECAL) and a hadron (HCAL) calorimeter act in conjunction and provide hermetic coverage of the interaction region. The complete ECAL, the HCAL barrel (HB) and HCAL endcap (HE) calorimeters are located inside the magnet coil. The HCAL outer tail catcher (HO) is situated outside the coil inside the muon system. To extend the hermetic coverage two separate very forward calorimeters (HF) are located downstream of the HE endcaps covering the high η region [CMS97a, CMS97b].

The electromagnetic calorimeter The CMS electromagnetic calorimeter is made of lead tungstate (PbWO_4) crystals and extends to $|\eta| = 3$. Precision energy measurement of electrons and photons is possible for $|\eta| < 2.6$.

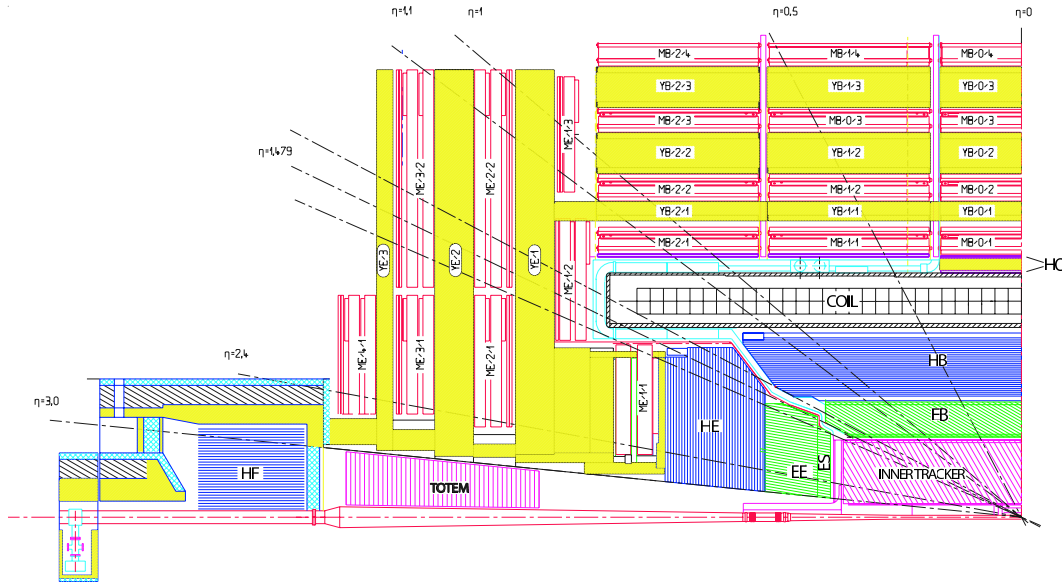


Figure 3.6: The r - z view showing the geometrical configuration of one quarter of CMS calorimetry. The complete ECAL (EB, EE, ES) and 2 parts of the HCAL (HB, HE) are located inside the coil. HF is located downstream of HE, behind the TOTEM experiment. HO is located inside the muon barrel (MB) system. Shown are also the muon endcaps (MB) and the return yoke of the barrel (YB) and the endcap (YE).

The transverse granularity of the crystals in the barrel (EB) is $\Delta\eta \times \Delta\phi = 0.0174 \times 0.0174$. This granularity matches the PbWO_4 Molière radius of 21.9 mm. The small Molière radius results in a reduced area over which the collected energy is summed and so reduces the effect of pileup contributions to the energy measurement. The crystals are grouped into 36 identical super-modules, each covering 20° in ϕ . They are mounted in a quasi-projective geometry with respect to the nominal vertex (the crystal axes are tilted by 3° with respect to the line from the nominal vertex position).

The endcaps (EE) extend over $1.479 < |\eta| < 3.0$. The crystals (front face cross section: $28.6 \times 28.6 \text{ mm}^2$) are grouped in mechanical units of 5×5 (super-crystals) and are mounted off-point from the nominal vertex. They are arranged in a x - y grid (i.e. not in an η - ϕ grid like the EB crystals).

A total thickness of 23 cm (≈ 26 radiation lengths (X_0)) of the crystals in the barrel limits the longitudinal shower leakage of highly energetic electromagnetic showers. A preshower detector (ES) ($3X_0$ of lead) in both endcap regions allows the use of slightly shorter crystals (22 cm) there. The active elements of the preshower are two planes of silicon strip detectors, with a pitch of 1.9 mm, behind the lead absorber disks at depths of $2X_0$ and $3X_0$. The preshower improves the π^0/γ separation in the forward directions.

The ECAL energy resolution can be parametrized as a function of energy by:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2, \quad (3.2)$$

where S is the stochastic term, N the noise and C the constant term. The design values and test beam measurements of a 5×5 crystal array are listed in Table 3.6. The test beam data are in agreement with the design considerations [CMS06a].

Contribution	Value		
	Barrel ($\eta = 0$)	Endcap ($\eta = 2$)	ECAL Test Beam
S	2.7%	5.7%	$2.96 \pm 0.10 \%$
N	0.155 GeV	0.205 GeV	$0.166 \pm 0.340 \text{ GeV}$
C	0.55%	0.55%	$0.32 \pm 0.01 \%$

Table 3.6: The contributions to the CMS ECAL energy resolution. Listed are the design values of a 5×5 crystal array and the measured values of a super-module [CMS97a, CMS06a].

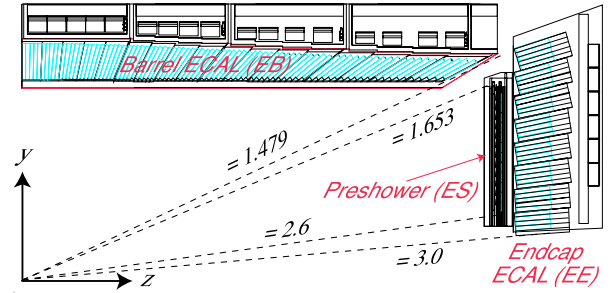


Figure 3.7: The r - z view showing the detailed geometrical configuration of the CMS ECAL, which consists of the EB ($|\eta| < 1.479$) and the EE ($1.479 < \eta < 3.0$). In front of the EE the ES ($1.653 < \eta < 2.6$) is located.

The hadron calorimeters The design of the HCAL barrel (HB) and endcap (HE) sampling calorimeters maximizes the interaction length of material inside the magnet coil in order to minimize the non-Gaussian tails of the energy resolution function. Copper was used as absorber because of its small interaction length. In addition it is non-magnetic and has a fairly low Z in order not to degrade the muon reconstruction. Plastic scintillators act as active material and are read out by embedded wave length shifting fibers. The tile/fiber technology makes

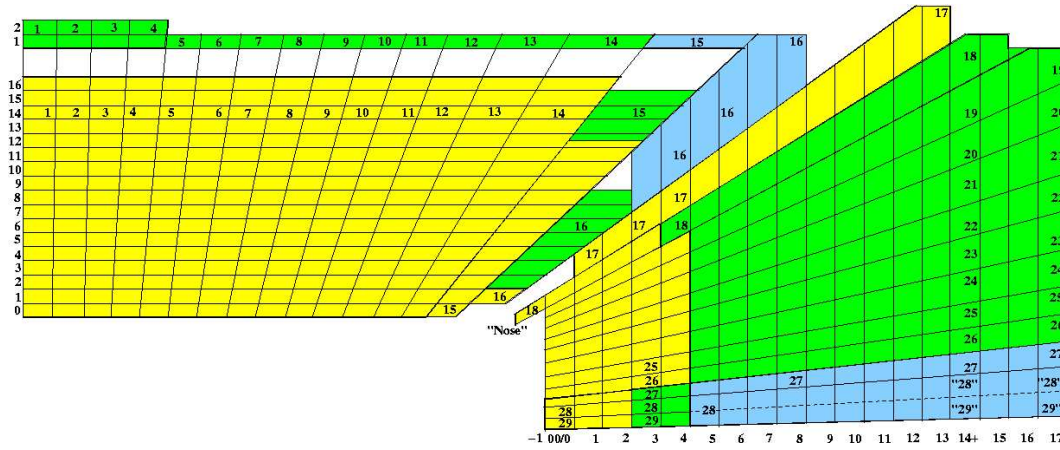


Figure 3.8: The r - z view of one quarter of the CMS HCAL barrel and endcap region showing the final projective readout tower configuration. The HB extends in η from tower index 1 to 16, HO from tower index 1 to 15 and the HE from tower index 17 to 29. The colors correspond to the number of readouts in r , *e.g.* 2 colors (yellow/green) \simeq 2 readouts. In the actual geometry simulation tower 28 and 29 are merged into one big tower 28 [CMS05b].

it possible to avoid non-instrumented cracks. The HCAL readout is organized in projective towers.

The layout of the HCAL barrel and endcap projective towers is shown in Figure 3.8. The 32 HB η -towers have a ϕ -granularity of 5° , which results in a $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ segmentation. The granularity of the HE towers varies from a 5° ϕ -granularity for $\eta < 1.740$ to a 10° geometry for higher η . The outer tail catcher (HO) enhances the amount of absorber for $|\eta| < 1.2$ and reduces the energy leakage, particularly for showers which start deep in HB. It is made of further scintillator fiber layers with the same tower granularity as HB.

The two steel/quartz fiber hadron calorimeters (HF) covering the region $3.0 < \eta < 5.0$ are located at $z = 11$ m from the nominal interaction vertex. The signal originates from Čerenkov light emitted in the quartz fibers. There are 13 towers in η , all with a size of $\Delta\eta = 0.175$, except for the towers at lowest and highest η for which $\Delta\eta = 0.1$ and $\Delta\eta = 0.3$ respectively. The ϕ -granularity of the towers is 10° , except for the highest- η one which has $\Delta\phi = 20^\circ$.

Physics performance Details of the HCAL design together with performance of production modules measured in CERN test beams can be found in [CMS05h], [CMS05i], and [CMS05j].

Jet reconstruction performance as well as the reconstruction of the transverse energy flow has been estimated in [CMS05k], [CMS05g] and will be described in detail in section 3.4.5.

3.3.5 The trigger system

The CMS trigger and data acquisition system (DAQ) is designed to reduce the overall event rate of $\mathcal{O}(10^9)$ Hz at the LHC design luminosity of $\mathcal{L} = 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ by a factor of $\mathcal{O}(10^7)$ to a maximum rate of $\mathcal{O}(10^2)$ Hz to fill the readout bandwidth with samples of physically pure events. The task is performed in two steps: The Level-1 trigger system (L1), which is made of custom electronics, operates on a subset of the data collected for each LHC beam crossing. The resulting event rate of $\mathcal{O}(10^2)$ kHz is further reduced by the High-Level trigger system (HLT) resulting in an event rate of $\mathcal{O}(10^2)$ Hz. The CMS DAQ/HLT is based on a large and complex commercial computer system to avoid built-in architectural or design limitations. Further tasks

of the DAQ are to collect events for calibration purposes, monitor the status of CMS and provide information on what is rejected [CMS00b, CMS02b].

The Level-1 trigger is organized in three major components: the L1 calorimeter trigger, the L1 muon trigger and the L1 global trigger. The trigger decision is based on the best four of each of the following objects: muons, isolated electrons or photons, non-isolated electrons or photons, central jets, forward jets and isolated hadrons or τ -jets. In addition the missing transverse energy, the total transverse energy and eight numbers of jets passing different E_t thresholds are used to select physically interesting events [CMS00b].

The High-Level trigger selects candidate events for specific physics processes. To attain this goal those objects and regions of the detector which are actually needed are reconstructed for events which were accepted by the Level-1 trigger. The HLT tries to reject events as soon as possible. This design leads to the notion of many virtual trigger levels, *e.g.* the calorimeter data is used followed by the tracker pixel data to define a HLT electron candidate. The HLT electron may trigger the partial reconstruction of tracks to search for a b-hadron signature [CMS02b].

The L1/HLT trigger tables used for the present analysis are listed in section 3.4.7.

3.3.6 The luminosity measurement

The normalization of physics analyses will be performed by measuring known cross sections with high precision, *e.g.* the pp total cross section, the W^\pm or Z^0 production rates. The design goal is a systematic accuracy of 5%. However, the real time monitoring of the luminosity will be based on the measurement of high cross section processes to provide statistically precise luminosity information for small samples. The design goal for the real-time luminosity measurement is a 1% statistical accuracy in 0.1s. The general strategies for the luminosity measurement of CMS are summarized in [CMS06b].

3.4 b-hadron identification in CMS

The full simulation, full reconstruction and pre-selection of the CMS data was done by means of the Worldwide LHC Computing Grid [Kno05, CMS05f]. The simulated CMS events were reconstructed and preselected by means of several Grid computing centers worldwide using ORCA³, the CMS reconstruction framework [CMS99a]. The preselected data were then transferred to a local computer farm, where the Condor batch system [Con] was utilized to analyze the events by means of the object-orientated b-tag framework developed for this study (Appendix B).

The CMS ORCA reconstruction framework provides tools to reconstruct the needed information from the recorded events to conclude about their physical origin. It is organized in subsystems specialized in different task: *e.g.* track reconstruction, jet finding, muon reconstruction, etc. In addition the ORCA framework contains the HLT algorithms (section 3.3.5).

The performance of the tools either already available or newly implemented, which are used to study the CMS b-hadron identification performance and the WH Higgs discovery potential, is described in the following.

Since the LHC will start operating from winter 2007 on, only simulated data are used. If not explicitly mentioned otherwise, all distributions demonstrating the detector performance are made by using $t\bar{t}$ events mixed with low-luminosity pileup events.

3.4.1 Full simulation of CMS events

The response to the complex physics processes, which take place in the CMS detector, cannot be calculated analytically. Therefore Monte Carlo methods [HH64, Cow98] are utilized to determine the CMS detector response for the individual physical events.

The simulated event samples used in this study are listed in Table 3.7)⁴. The data were accessed by means of Grid tools at the regional Grid computing centers CERN (Geneva, Switzerland), FNAL (Fermilab, USA) and PIC (Barcelona, Spain) [CMS05f].

Because of the high LHC luminosity (section 3.2), a number of pp collisions occur, for each bunch crossing. Most of them are quite soft interactions and are named *minimum-bias* or *pileup* events. The hard interactions are named *trigger* or *signal* events. For each simulated CMS event a signal event is mixed with a random set of minimum-bias events.

Both the signal and the minimum bias events are produced with the CMS CMKIN FORTRAN program library [CMS96, CMS04d], which interfaces the Monte Carlo generators: *e.g.*

channel	# events	location
WH ($W \rightarrow l\nu$, $H \rightarrow b\bar{b}$)	60000	FNAL
$t\bar{t}$ +jets	2 M	CERN/FNAL
Wbb ($W \rightarrow l\nu$)	1 M	PIC
WW+jets	0.4 M	FNAL
ZW+jets	0.28 M	FNAL
ZZ+jets	0.4 M	FNAL
QCD 2-jets ($0 < P_t < 1800$ GeV)	1 M	FNAL

Table 3.7: Details about the simulated event samples used in this study.

³Object Oriented Reconstruction for CMS Analysis

⁴In tables the 'M' denotes millions of events.

PYTHIA [SLMS03], TopReX [SS02] and CompHEP [CompHEP04]. The full simulation of the detector response is based on the Geant4 C++ toolkit [GEANT403] utilized by the CMS OSCAR⁵ program [CMS99c, CMS06b]. The OSCAR program, written in C++, is able to simulate in detail the electromagnetic and hadronic interactions and the complete detector response for CMS events.

The WH signal and background event generation The WH signal events are generated with the PYTHIA event generator. The Standard Model Higgs mass for the production is fixed to 120 GeV, where the Higgs decays predominantly into $b\bar{b}$ quark pairs (section 1.3.2). To optimize the available computer resources a pre-selection of the WH events is implemented. Events are accepted, if the Higgs decays into a $b\bar{b}$ quark pair and the associated W boson decays leptonically.

All possible triggered WH background events⁶ and minimum-bias events are generated with PYTHIA. The Wbb background events are generated with TopReX. For this channel a pre-selection of events is implemented again, so that only events are accepted for which the W boson decays leptonically.

On average 3.5 minimum-bias events are mixed to each signal event. This is the expected number of events for the LHC low-luminosity phase, which will pileup on top of the signal events [CMS06b].

Details about the used event samples are listed in Table 3.8. The listed cross sections are obtained from up to date next-to-leading order calculations. For the Wbb process only a leading order calculation is available. The predicted NLO correction is very large (\approx a factor of 2.5) [Cam]. The cross section of the WH signal process as described before is computed by means of HDECAY with a top quark mass of $m_{\text{top}} = 175$ GeV [DKS98, Nik]. The cross section for the $t\bar{t}$ production is taken from reference [BCMN98]. The other cross sections are calculated by means of MCFM [CER03, CE, Nik].

channel	cross section [pb]	# events @ 60 fb ⁻¹	reference
WH (W $\rightarrow l\nu$, H $\rightarrow b\bar{b}$)	0.3536	21216	[DKS98, Nik]
$t\bar{t}$ +jets	803.0	48.18 M	[BCMN98]
Wbb (W $\rightarrow l\nu$)	111.0	6.66 M	[CER03, CE, Nik]
WW+jets	114.3	6.86 M	
ZW+jets	49.9	3.00 M	
ZZ+jets	15.3	0.92 M	

Table 3.8: The event samples used in this analysis with the corresponding cross sections. The expected number of events is given for an integrated luminosity of 60 fb⁻¹.

3.4.2 Charged track reconstruction

The charged track reconstruction is based on a combinatorial Kalman filtering (CKF). The recursive method is equivalent to a global least-squares minimization, which is the optimal estimator if the track model is linear and the errors of the track hit measurements are Gaussian. For non-linear models and non-Gaussian errors, the CKF is still the optimal linear estimator [Frü87].

⁵Object Oriented Simulation for CMS Analysis and Reconstruction

⁶The triggered WH background events are named background events for convenience. They are also signal events in contrast to minimum-bias events as described before.

The CMS implementation of the CKF starts from an initial set of track trajectory candidates (seeds), which are recursively propagated from layer to layer (trajectory building), accounting for multiple scattering and energy loss. The large number of mutually exclusive trajectory candidates obtained for each seed is reduced by selecting candidates based on their number of shared hits and their track quality (trajectory cleaning). The final track helix fit of the five track parameters, transverse momentum P_t , ϕ , $\cot\theta$, the transverse d_0 and longitudinal z_0 impact parameter, uses optimal estimates at every hit measurement along the selected trajectories (trajectory smoothing) [CMS05a].

3.4.2.1 The track reconstruction in the inner tracker

The tracking of the charged particles from the first pixel vertex detector layer to the outer layers of the microstrip tracker (*inside-out-tracking*) starts with a seed from each pair of hits in two given layers of the pixel vertex detector, compatible with the beam line constraint and a minimal transverse momentum **ptmin**. The beam line constraint is realized as a cylindrical region, with radius **originRadius**, half length **originHalfLength** and center **originZPosition** around the LHC beam line, from where hits are accepted. This seeding implicitly favors tracks with pixel hits. To ensure an optimal tracking efficiency for b-hadron decay tracks, which contain the b-hadron lifetime information, the CMS default parameters of the seed finding are slightly changed [Win02]. The radius of the cylindrical seed region is increased to 0.3 cm, which matches the b-hadron flight length of ≈ 2.5 mm (section 1.1.3). The minimal transverse momentum of the seed trajectories is reduced to 0.8 GeV, so that a higher number of seeds is utilized. The higher number of seed trajectories takes into account the dense track environment in b-hadron decays. The CMS defaults and the values used are listed in Table 3.9.

seed parameter	CMS default	used value
ptmin [GeV]	0.9	0.8
originRadius [cm]	0.1	0.3
originHalfLength[cm]	15.0	15.0
originZPosition [cm]	0.0	0.0

Table 3.9: The optimized parameters for track seed finding in the CMS inner tracker to efficiently reconstruct b-hadron decay tracks [Win02].

For the determination of the track reconstruction performance, reconstructed tracks are matched to simulated tracks, if they share more than 50% of their hits. Simulated tracks are reconstructible charged tracks inside the silicon tracker volume, which are reconstructed by means of the CKF from simulated hits produced by OSCAR before the CMS detector response is taken into account. A unique matching between reconstructed and simulated hits is provided by OSCAR.

The constraints for the simulated and reconstructed tracks to be accepted are listed in Table 3.10. The constraints ensure that reconstructible simulated tracks from the interaction region are investigated. All reconstructed tracks from inside the inner tracker volume are accepted, if their transverse momentum $p_{t,\min}$ exceeds 0.7 GeV. Tracks with lower transverse momentum cannot be reconstructed with the CKF due to multiple scattering and because they are curling due to the large bending power of the 4 T CMS magnetic field [CMS02b].

The resolution of the five track parameters is shown in Figure 3.9. All distributions have a reasonable mean value and a spread compatible with the expectations listed in Table 3.3.

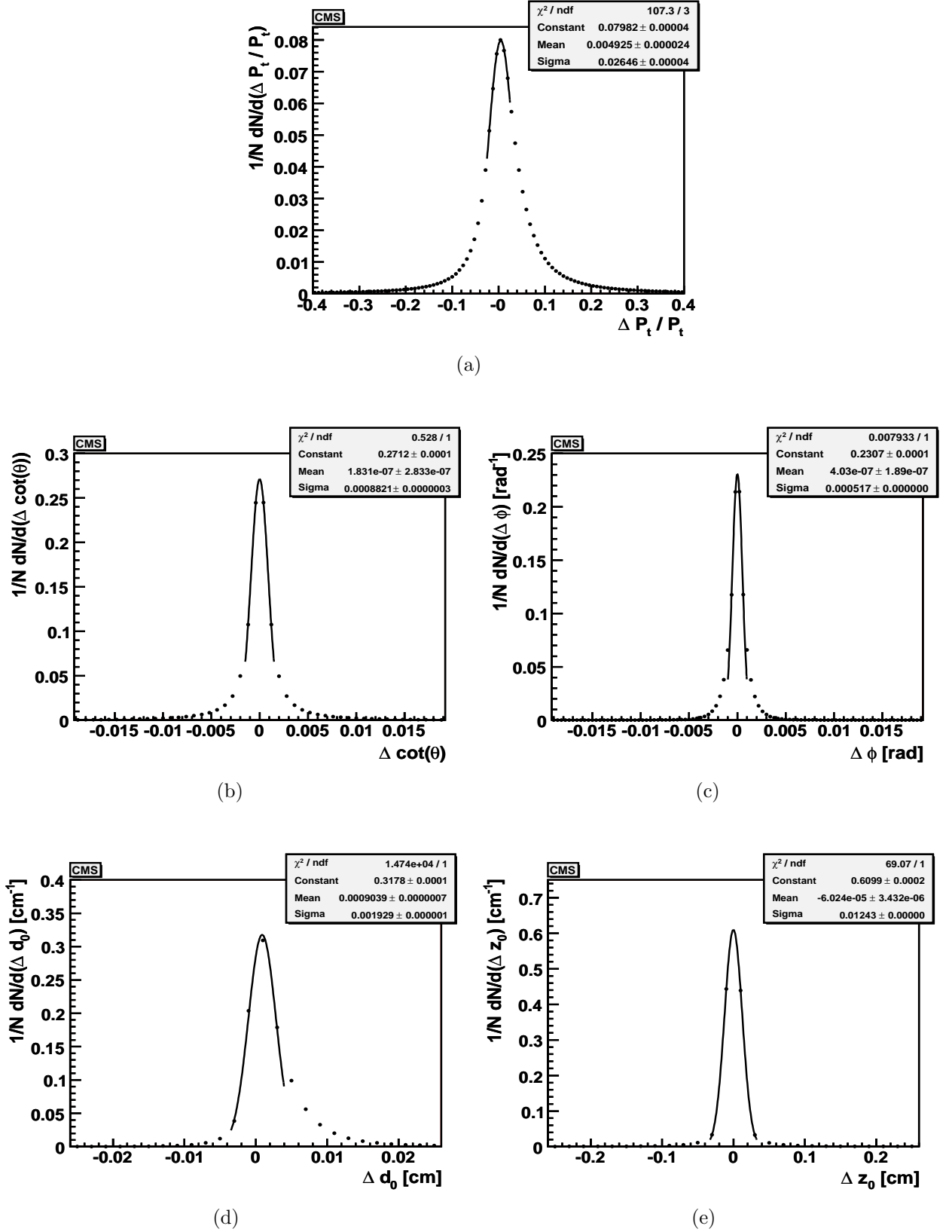


Figure 3.9: CMS track reconstruction performance for the five track parameters transverse momentum P_t , $\cot \theta$, ϕ , transverse impact parameter d_0 and longitudinal impact parameter z_0 .

The non-Gaussian tails visible are caused by the deterioration of reconstructed tracks due to *e.g.* material interactions. In Figure 3.9(d) the non-Gaussian tail is clearly visible for positive values of Δd_0 only, because the frame of reference is chosen in such a way that the reconstructed transverse impact parameter $d_{0,\text{rec}}$ is always positive.

track parameter	simulated tracks	reconstructed tracks
$P_{t,\text{min}}$ [GeV]	0.9	0.7
η_{min}	0.0	0.0
η_{max}	2.5	2.6
$d_{0,\text{max}}$ [cm]	3.0	120.0
$z_{0,\text{max}}$ [cm]	30.0	170.0
minimum # hits	0	8

Table 3.10: The constraints used for the simulated and reconstructed CKF tracks to test the CMS inner tracking performance of tracks, which stem from the nominal interaction region [CMS02b].

3.4.2.2 The muon reconstruction

The global muon track reconstruction (GMT) starts with the reconstruction of muon track candidates in the muon system alone. These muon track segments are used to define a region of interest, a tube, in the inner tracker, where possible additional muon track hits are located. Regional seeds, made from pairs of hits in different tracker layers inside the tube, are propagated opposite to the muon momentum to build muon track candidates in the inner tracker (*outside-in-tracking*). After ambiguity resolution (trajectory cleaning) of the inner tracker muon track candidates, the hits from all muon track candidates selected in the inner tracker and the corresponding muon system hits are used for the muon track fit (trajectory smoothing). To improve the final muon track fit an iterative procedure is utilized to remove hits with a high χ^2 contribution to the fit. Muon track candidates are accepted based on a reasonable χ^2 of the final fit [CMS05e].

Reconstructed muons are used as a component of the jet energy flow input (section 3.4.5) and to reconstruct W bosons which decayed leptonically (section 3.4.7). Therefore the P_t -, η - and ϕ -resolution for muons with $P_t > 2$ GeV are shown in Figure 3.10. The reconstructed muon candidates are matched to generated ones if their η - ϕ distance is smaller than $R = 0.1$)⁷. The presented muon reconstruction performance is in good agreement with the expectations listed in Table 3.5.

⁷The distance of two objects i and j in η - ϕ space is defined as: $R = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}$.

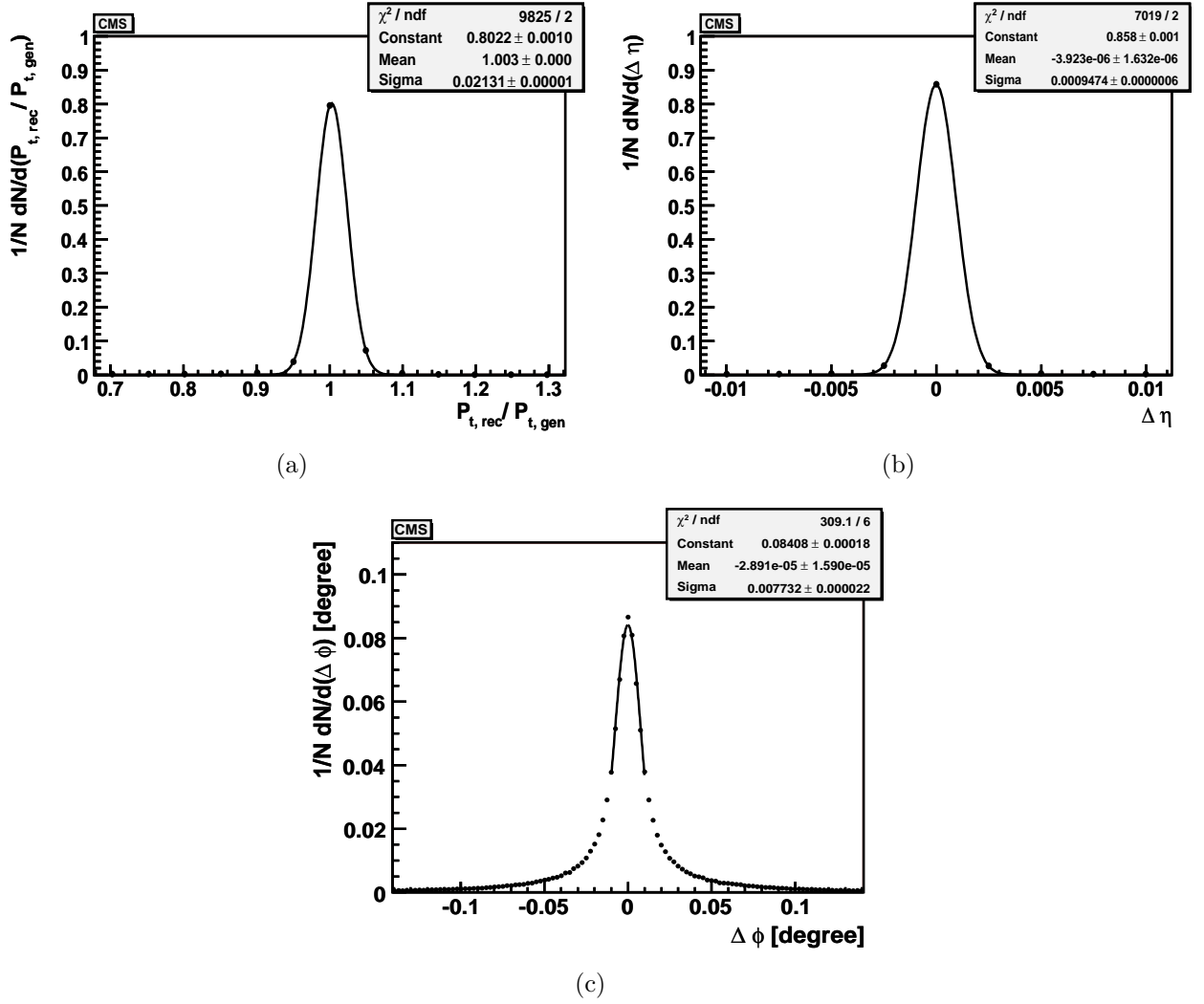


Figure 3.10: CMS muon track resolutions for the transverse momentum P_t , η and ϕ .

3.4.2.3 The electron reconstruction

The reconstructed electron 4-vector is composed of a single reconstructed track emerging from the interaction vertex and a matched electromagnetic super-cluster⁸. Due to the large bending power of the CMS magnetic field (section 3.3.1), the electron reconstruction has to deal with bremsstrahlung of the electrons, which takes place in the r - ϕ plane.

The electromagnetic super-clusters are made by the hybrid algorithm in the ECAL barrel and by the island algorithm in the ECAL endcaps [CMS01a, CMS02b]. Both algorithms try to re-collect the energy deposited in the ϕ direction of the cluster core. The super-clusters are used to determine the pixel seeds for the inside-out-tracking with the combinatorial Kalman filter⁹. A successfully reconstructed electron track candidate matched to a super-cluster consists of at least 3 hits.

Electrons with a transverse momentum $P_t > 26$ GeV are used to reconstruct W bosons that

⁸The described reconstruction works as well for positrons.

⁹The errors of the hit measurement for electron tracks are non-Gaussian, because of the bremsstrahlung. However the non-linear generalization of the Kalman filter, the Gaussian-sum filter (GSF) [AFST05] was not re-implemented for the ORCA version needed to access the simulated data used in the present study.

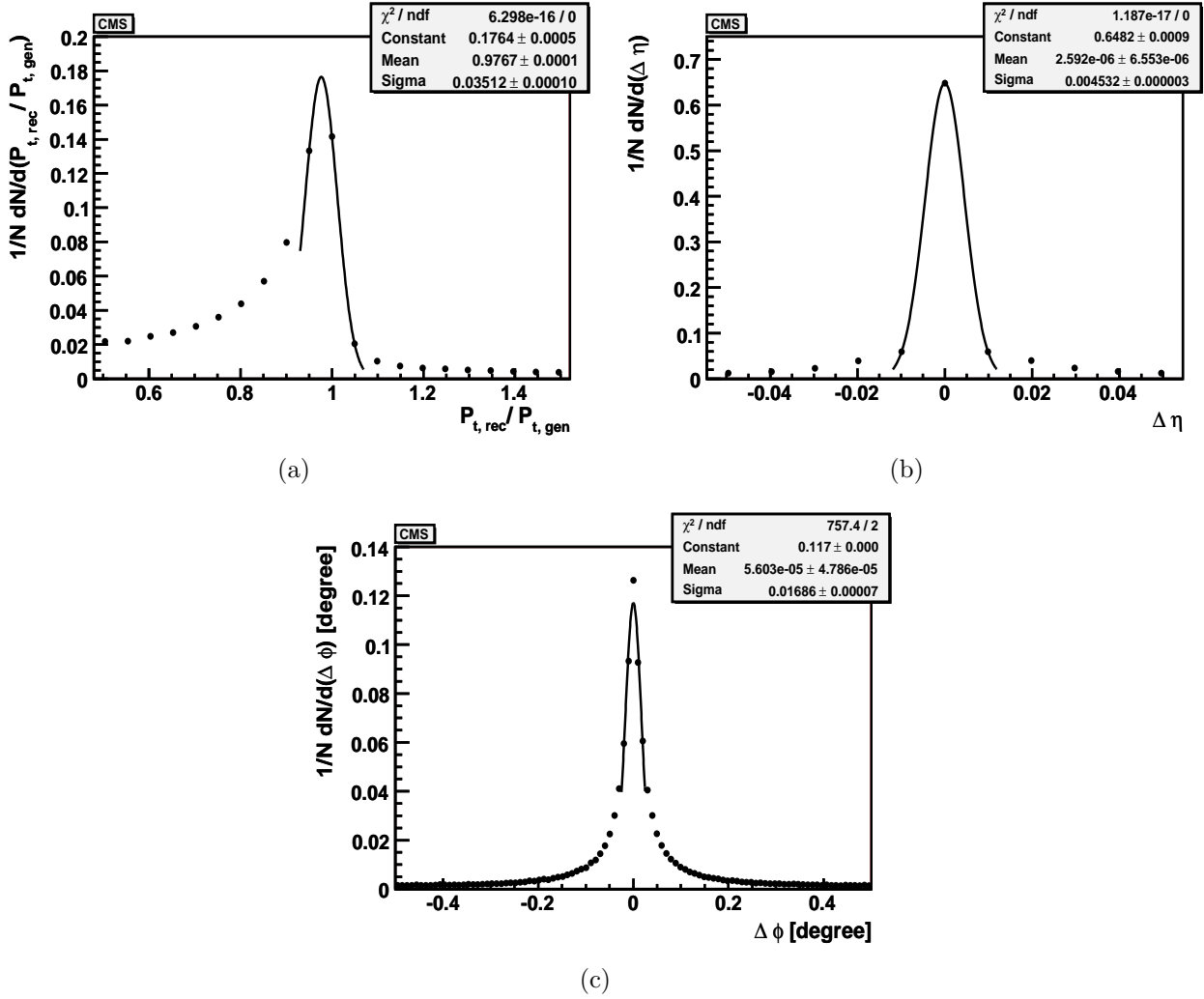


Figure 3.11: CMS electron resolutions of the transverse momentum P_t , η and ϕ .

decayed leptonically (section 3.4.7). For these electrons the P_t -, η - and ϕ -resolutions are shown in Figure 3.11. The reconstructed and generated electrons are matched if they are closer than $R = 0.1$ in η - ϕ space.

The tail on the left of the P_t -resolution in Figure 3.11(a) originates from the energy loss of the electrons due to bremsstrahlung so that a sizable amount of electrons are reconstructed with a P_t lower than the generated one. The r.m.s. resolution for matched candidates in the peak of the distribution is almost comparable to that for muons.

Because of the artificial η - ϕ matching, only electron candidates reconstructed in the direction of the generated electrons enter the resolution plots. It follows that in particular the η - and ϕ -resolutions are biased because reasonably well reconstructed candidates are favored. Despite this the η - and ϕ -resolution of these electrons is roughly one order of magnitude larger than that of muon candidates. The influence of bremsstrahlung is only slightly visible in the non-Gaussian tails of the ϕ -resolution shown in Figure 3.11(c), because electrons with a relatively high transverse momentum are shown. Electron tracks for which $P_t > 26$ GeV are quite stiff so that bremsstrahlung is emitted in the direction of flight mostly. Furthermore the hybrid and island algorithms both are able to compensate the effect of bremsstrahlung partially as described before.

3.4.3 Vertex reconstruction

The reconstruction of a vertex is done in three steps. First a region of interest is defined for the vertex search. Then it is examined whether a group of tracks is compatible with originating from a common vertex. The final step is a vertex fit to determine the vertex position of the selected group of tracks.

The *Principal Vertex Reconstructor* (PVR) chooses vertex candidates by an iterative procedure on reconstructed tracks [CMS02b]. To find the principal vertex the PVR fits a set of tracks to a common vertex and removes tracks with a compatibility less than 5% as long as the χ^2 of the fit is small enough. To find all possible vertex candidates the procedure is applied on the remaining set of incompatible tracks till there are no tracks left. The vertex fitting is done with the Kalman Filter Formalism (KVF), which allows to include a prior position¹⁰ in the fit [CMS03d].

The reconstructed vertices are classified by means of a χ^2_{vtx} criterion formed from the transverse distance of the vertex to the beam line. The cuts used correspond to percentiles of the χ^2 distribution with 2 degrees of freedom. A vertex is considered as *primary* if its transverse distance is compatible with the beam line with a probability larger than 5% ($\chi^2_{vtx} < 5.99$) and as *secondary* if this probability is smaller than 1% ($\chi^2_{vtx} > 9.21$). These are the standard CMS definitions [Spe].

Reconstructed tracks (section 3.4.2) are used for vertex finding if their transverse momentum is reasonably high: $P_{t, \text{track}} > 1.5 \text{ GeV}$.

3.4.3.1 Primary vertex reconstruction

The PVR is modified to perform a primary vertex finding on a set of tracks compatible with coming from the LHC beam line [CMS06b]. The tracks of the obtained primary vertex candidates are fitted by means of the KVF using the beam line constraint in each step of the iterative fitting procedure.

In the OSCAR simulation the beam line is fixed to $x = y = 0$ with an error of $15.9 \mu\text{m}$ given by the LHC beam parameters (section 3.2). These beam parameters are used as a prior position in the KVF fit. The x - and y -position of the beam can eventually be obtained from real LHC data by means of a special track (event) selection excluding any long lived particles, *e.g.* b-hadrons.

The vertex reconstruction described above results in several primary vertex candidates per event, because of additional vertices from pileup events (section 3.4) and because of faked primary vertices, due for example to long lived particles along the beam line. The number of primary vertex candidates per event $N_{1st, vtx}$ is shown in Figure 3.12(a). The efficiency to obtain at least one primary vertex is 98%.

To determine the accuracy of the position measurement for the primary signal vertex the positions of the simulated and reconstructed signal vertices are compared. Therefore the reconstructed primary signal vertex must be associated to the simulated signal vertex. The simulated and reconstructed vertices are associated by means of the simulated and reconstructed tracks assigned to these vertices respectively. Simulated and reconstructed tracks are matched if at least 50% of the hits used to compute these tracks are linked by means of generator information. If at least 50% of the simulated and reconstructed tracks of a vertex pair are matched, the vertices themselves are defined as associated.

¹⁰A prior position defines an additional constraint in the vertex fit.

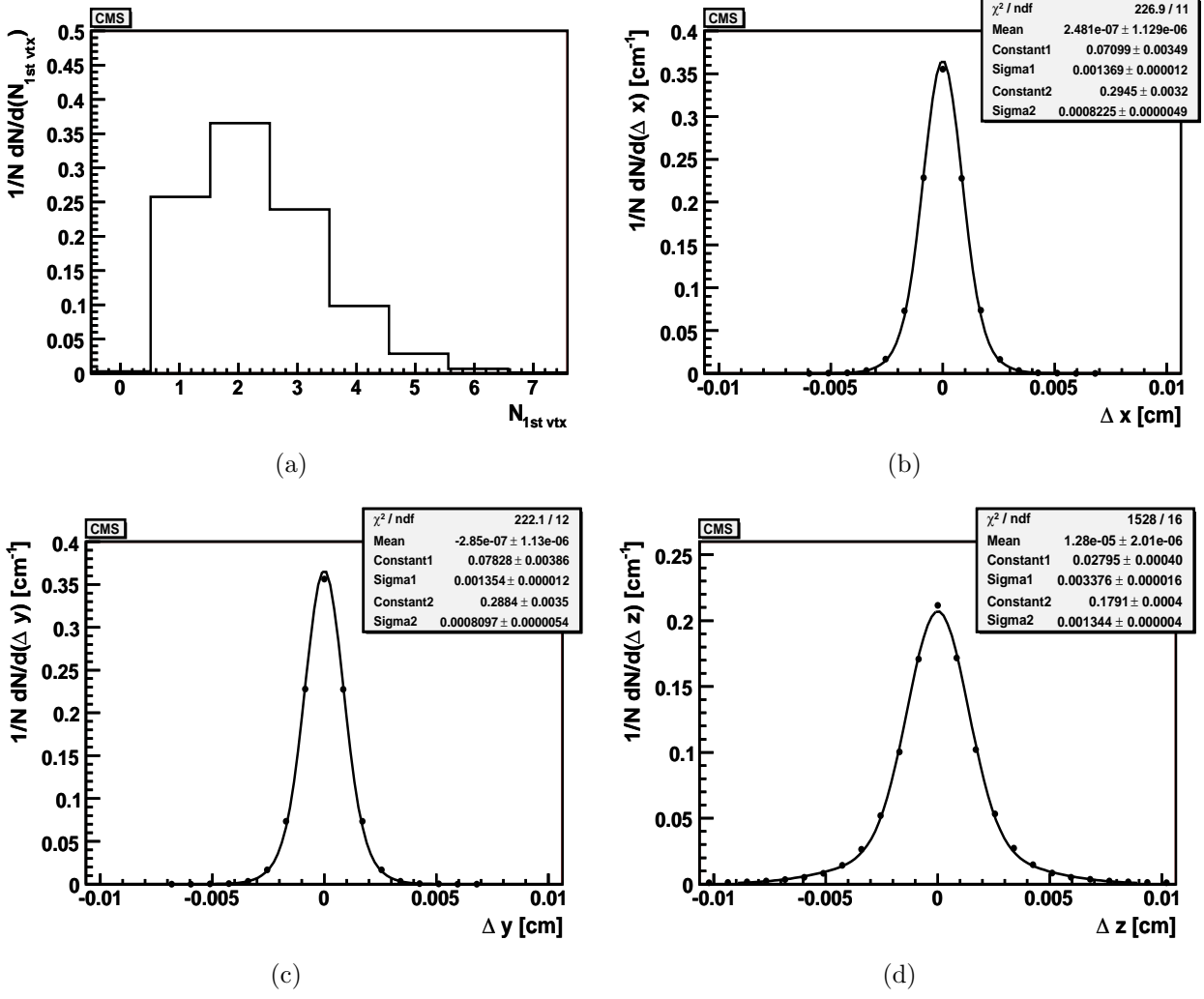


Figure 3.12: The number of reconstructed primary vertex candidates per event (a) and the resolutions of the x -, y - and z -coordinates for the primary signal vertex (Figures (b)-(d)).

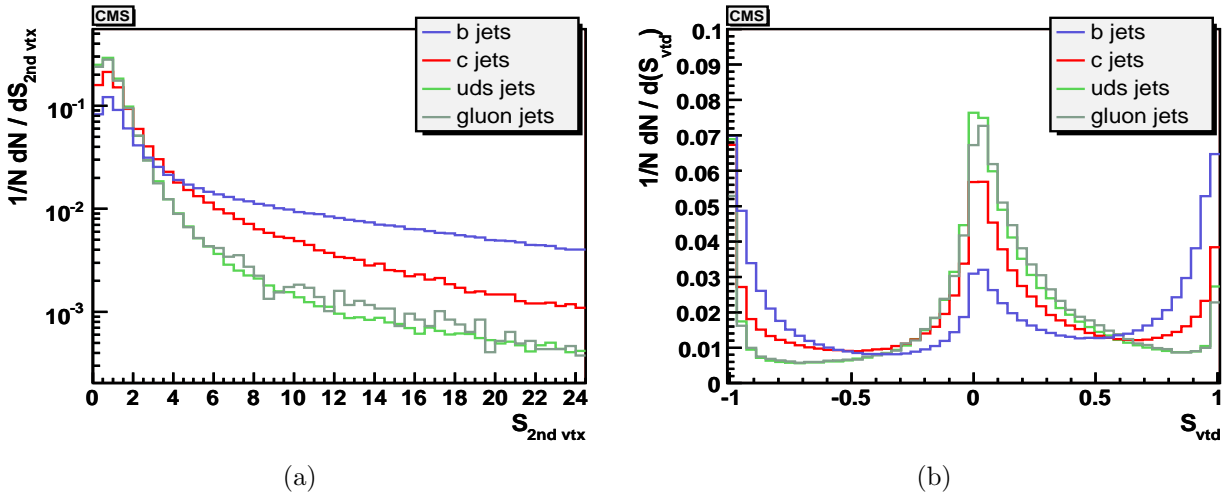


Figure 3.13: The significance $S_{2nd\ vtx}$ of reconstructed secondary vertices (a) and the significance of the vertex-track distance S_{vtd} (b) for different jet flavors.

The measurement accuracy of the signal primary vertex position is shown in Figures 3.12(b)-(d). The distributions are fitted by means of a *Double Gaussian* function DG :

$$DG := Constant1 \cdot \exp\left[-0.5 \left(\frac{x - Mean}{Sigma1}\right)^2\right] + Constant2 \cdot \exp\left[-0.5 \left(\frac{x - Mean}{Sigma2}\right)^2\right] \quad (3.3)$$

All distributions peak reasonably well at zero. The average resolution σ_{comb} is defined as:

$$\sigma_{\text{comb}} = \frac{Constant1 \cdot Sigma1 + Constant2 \cdot Sigma2}{Constant1 + Constant2} \quad (3.4)$$

The average resolution σ_{comb} of the primary vertex z -coordinate is about $16 \mu\text{m}$ compatible with the expected performance (Figure 3.12(d)). In comparison to the standard CMS vertex reconstruction performance listed in the references [CMS03d, CMS06b, CMS05m], the average resolution σ_{comb} of the x - and y -coordinates (Figures 3.12(b) and (c)) improves from about $15 - 20 \mu\text{m}$ to below $10 \mu\text{m}$.

3.4.3.2 Secondary Vertex reconstruction

The secondary vertex reconstruction is done for each reconstructed jet by means of the PVR without any prior position. The jets are reconstructed as described in section 3.4.5. Tracks, which are close to the jets in η - ϕ space, $\Delta R(\text{rec.jet}, \text{rec.track}) < 1.0$, enter into the vertex reconstruction.

If more than one secondary vertex candidate is found for one jet, the one with the highest significance $S_{2\text{nd vtx}}$ is used. Figure 3.13(a) shows the significance of these candidates separately for b-, c-, uds- and gluon-jets.

Secondary vertices are accepted if their significance $S_{2\text{nd vtx}}$ exceeds 6.5. The resulting efficiency for b-hadron jets is 57%. As already observed for ALEPH events (section 2.4.3.4), the $S_{2\text{nd vtx}}$ cut increases the probability of the reconstructed secondary vertices to stem from a b-hadron decay (section 2.4.3.4).

3.4.3.3 Association of tracks with the secondary vertex

The S_{vtd} observable described in section 2.4.3.5 is used again to associate tracks to a reconstructed secondary vertex. Figure 3.13(b) shows the observable for different jet types.

The fraction of tracks from b-hadron jets at $S_{\text{vtd}} = 1$ is higher than for the other jet types, because the probability that these jets contain a real secondary vertex is high and so the probability that the S_{vtd} observable peaks at 1 is increased. The tracks of c-hadron jets are almost equally distributed at $S_{\text{vtd}} = 1$ and -1 . The tracks from uds- and gluon-jets have quite similar distributions and are spread around $S_{\text{vtd}} = 0$, because for these jet types reconstructed secondary vertex candidates are mostly fakes.

Just like for the ALEPH study (section 2.4.3.5) tracks for which $S_{\text{vtd}} > 0$ are defined as secondary vertex tracks.

3.4.3.4 Vertex reconstruction from pixel tracks and signal vertex tag

To select the signal vertex of an event from the list of reconstructed primary vertex candidates (section 3.4.3) the z position of the reconstructed pixel vertex is utilized.

The *Divisive method* is used to perform the primary vertex reconstruction with tracks made from pixel hits only (tracklets) [CMS03b]. This fast and efficient pixel vertex reconstruction is

optimized for HLT purposes and performs a primary vertex finding along the z -axis only with an efficiency close to 100%. The signal pixel vertex is identified (tagged) by means of summing up the P_t of tracklets associated to the vertex candidates respectively. The vertex with the highest sum is taken [CMS02b].

The z -resolution obtained by comparing the reconstructed z -position of the tagged pixel signal vertex with the simulated signal vertex is shown in Figure 3.14(a). As expected the r.m.s. of the z -resolution of about $26 \mu\text{m}$ is higher compared to the primary vertex reconstruction by means of tracks obtained from the full tracker (Figure 3.12(d)).

For events where the signal consists of objects in the final state which have a non negligible lifetime, *e.g.* events WH , $\text{H} \rightarrow \text{b}\bar{\text{b}}$, the high accuracy of the primary vertex reconstruction by means of tracks from the complete inner tracker spoils the determination of the signal primary vertex. In such events a reasonable amount of tracks with high P_t are correctly assigned to true secondary vertices and they do not contribute to the signal vertex tag, which utilizes the P_t of tracks associated to a primary vertex.

Figure 3.14(a) indicates that the tracklets used in the reconstruction of the pixel primary vertex are not as accurately determined as the tracks obtained from the full inner tracker (Figure 3.12(d)). Still it turns out that the signal tag for the pixel vertex using the P_t of tracklets associated to it is more reliable, even in the presence of tracks containing lifetime information, because this information is smeared due to the way the tracklets reconstruction is performed.

To employ the more reliable signal tag for the pixel vertex, the fully reconstructed primary vertex which is closest in z to this signal pixel vertex is defined as the signal primary vertex. In Figure 3.14(b) the z -distance between the true signal pixel and true signal fully reconstructed primary vertices, both identified by choosing the closest ones to the simulated vertex of the hard interaction, is plotted to prove that this association is reasonable (black full line). The small spread of the Gaussian (about $25 \mu\text{m}$) reflects the correlation between the two vertices. Choosing, still on the generator level, another (the next closest in z) primary vertex, not the signal one, results in higher displacements and a very non-Gaussian distribution (red dashed line in Figure 3.14(b)).

The implication from the remarks before is that 98% of the reconstructed events have at

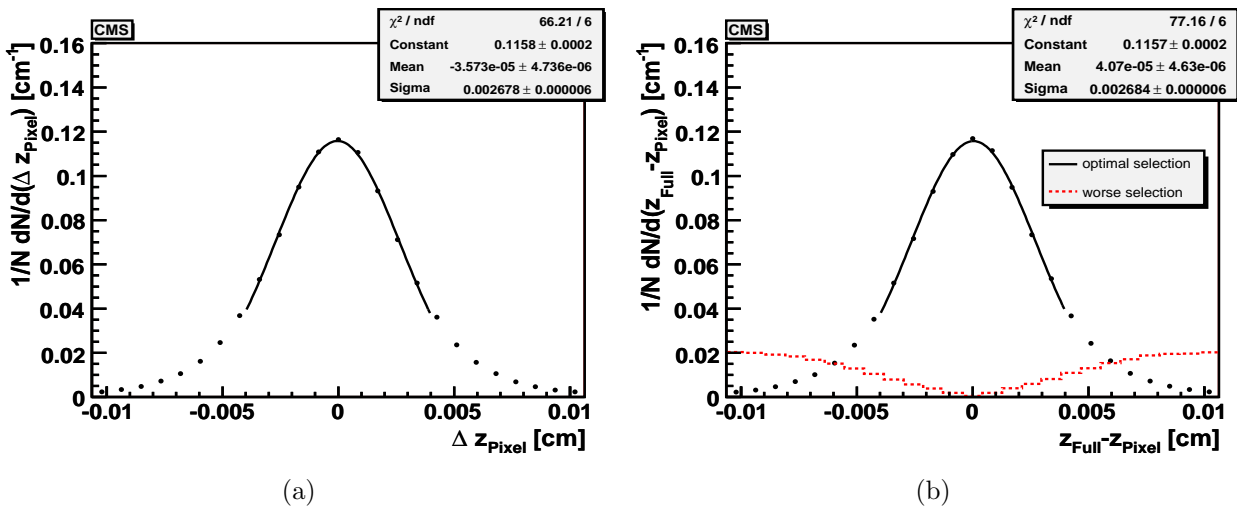


Figure 3.14: The pixel vertex z coordinate resolution (a) and the comparison of the pixel vertex z position with the z position of reconstructed vertices by means of the full tracker (b).

least one pixel and one fully reconstructed primary vertex. For such events the signal vertex tag described above is 100% efficient with a fake rate of well below 1%.

3.4.4 Missing energy reconstruction

The missing transverse energy $E_{T,\text{MET}}$ is reconstructed from a vector sum of the transverse energies $E_{t,\text{tower}}$ of the projective calorimetric towers with muon corrections. The muons are reconstructed as described in section 3.4.2. The transverse energy $E_{t,\text{muon}}$ of the muons is calculated by means of the muon track vector neglecting their mass. The complete formula to calculate $E_{T,\text{MET}}$ is:

$$E_{T,\text{MET}} := -\left(\sum_{\text{towers}} E_{t,\text{tower}} + \sum_{\text{muons}} E_{t,\text{muon}}\right), \quad (3.5)$$

where the first sum runs over all projective towers for which $E_{\text{tower}} > 0.8 \text{ GeV}$ and the second one over all reconstructed muons for which $P_{t,\text{muon}} > 2 \text{ GeV}$.

The resolution of the resulting missing transverse energy separated into x - and y -coordinates is shown in Figures 3.15(a) and (b). The obtained mean values of 0.1-0.3 and the sigma of about 26% are compatible with the expected reconstruction performance for the tested $t\bar{t}$ events mixed with low-luminosity pileup events [CMS05g].

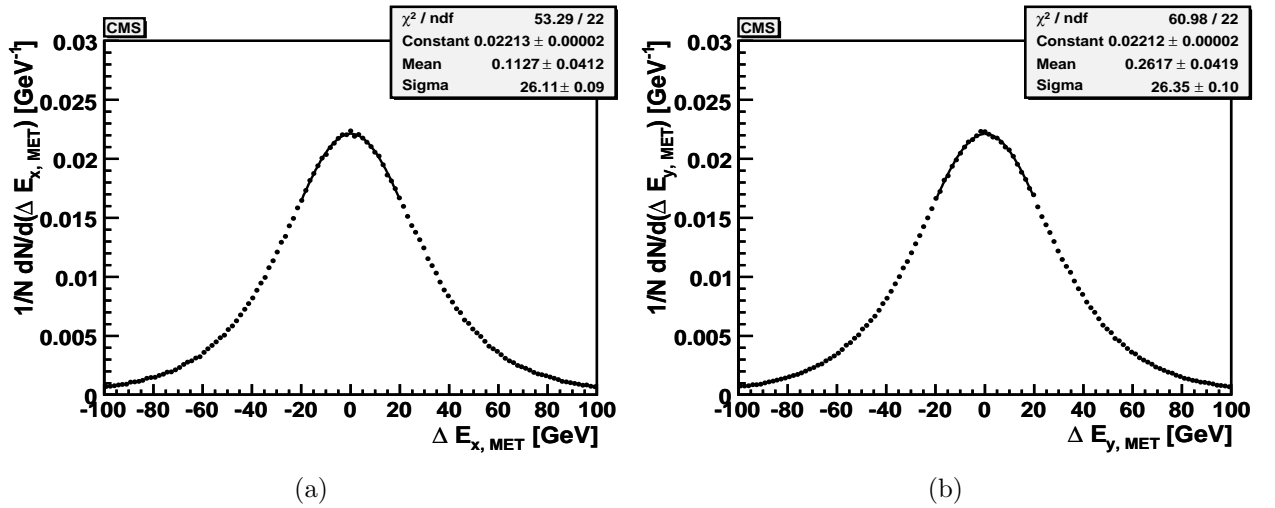


Figure 3.15: The missing energy resolution separated for the x - and y -coordinates $E_{x,\text{MET}}$ (a) and $E_{y,\text{MET}}$ (b).

3.4.5 Jet reconstruction

To study an optimal jet reconstruction the ORCA Jets subsystem had to be revised and additional functionality had to be added. In the following this work is briefly described:

- The MidPoint Cone jet algorithm and the KtJet package [BCCW03], which includes the inclusive k_{\perp} jet algorithm and several recombination schemes, were added. The available jet algorithms, *e.g.* the Iterative Cone algorithm, were revised.
- All reconstructed objects from ORCA, *i.e.* calorimetric towers, inner tracker tracks, muons, etc. needed had to be made available as input objects for jet finding. The combination of these inputs to improve the reconstructed jet 4-vectors was also added (energy flow).
- The possibility of a final state jet correction was introduced, *e.g.* a correction for detector effects.

Since 2003 these modifications were continuously added to the official CMS ORCA reconstruction framework.

The investigation of the jet reconstruction setup used here, consisting of the jet input, the jet algorithm and the final state jet calibration, is described in the next sections.

To ensure that the reconstructed jets are well contained in the inner tracker region of $|\eta| < 2.5$ (section 3.3.2) the reconstructed jet axes are restricted to $|\eta_{\text{jet}}| < 2.4$. The restriction is needed to be able to perform b-hadron identification associating reconstructed tracks from the inner tracker to the jets (section 3.4.6).

3.4.5.1 Jet input: split UE TrackTowers

The energy flow objects *TrackTowers*, which utilize projective calorimetric towers and reconstructed muons, are used as jet input.

The projective calorimetric towers (EcalPlusHcalTowers) consist of the HCAL towers (Figure 3.8) and the geometrically corresponding ECAL modules (Figure 3.7). Only EcalPlusHcalTowers with a minimal energy $E_{\text{tower}} > 0.8 \text{ GeV}$ are used. This energy corresponds to ≈ 3 ADC counts¹¹ above the electronic noise of the readout [Abd].

To consider only EcalPlusHcalTowers, which contain the energy of the signal event, a data driven method (UE method) is devised. The non-signal energy in an event originates from the minimum-bias events and the underlying event (UE). The underlying event energy is defined as everything, except the energy from the signal and the minimum-bias events, *e.g.* beam-beam remnants or semi-hard interactions. The reconstructed jets are used to compute the average unclustered energy $E_{\text{tower, UE}}$ per event. EcalPlusHcalTowers which belong to the jets are removed and the remaining towers are defined to contain the energy of the underlying event.

The average unclustered energy $E_{\text{tower, UE}}$ per event is plotted versus i - η in Figure 3.16(a). The histogram shows the energy of the unclustered EcalPlusHcalTowers versus η separated for each tower index i . The tower index corresponds to the numbering of projective readout towers shown in Figure 3.8. The unclustered energy $E_{\text{tower, UE}}$ in the barrel region (tower index 1~18) shows an almost flat distribution of 1-2 GeV, which increases to about 10 GeV for tower index 26. Due to the strong magnetic field of 4 T, the products of the underlying event are swept to the endcap regions, where the unclustered energy $E_{\text{tower, UE}}$ increases dramatically, *e.g.* charged

¹¹In the CMS simulation OSCAR one ADC count corresponds to $\approx 260 \text{ MeV}$ for the least significant bits (lowest range) [Abd].

particles with a transverse momentum $P_{t,\text{track}} < 0.6 \text{ GeV}$ will never reach the outermost barrel layer of the CMS tracker and could curl on a helical path to the endcap regions [Win02].

Figure 3.16(a) shows the unclustered energy $E_{\text{tower, UE}}$ versus η and the corresponding UE curve obtained from a spline fit. Because the energy per bin in this plot is integrated over more than one tower index, the energy in the barrel region $|\eta| < 1.479$ is slightly enhanced with respect to Figure 3.16(a). The same is true for the higher η regions.

The spikes visible in Figures 3.16(a) and (b) occur either due to the plotting of the non-equidistant EcalPlusHcalTower geometry in equidistant bins of η or because of dramatic changes in the tower geometry, *e.g.* tower 28 is twice as big as the towers located at smaller η (Figure 3.8). To reduce the influence of these big EcalPlusHcalTowers on the jet finding, the towers are split by software into smaller pieces containing equal parts of their original energy [CMS05k]. Because of support structures in front of tower 29, this tower mainly collects energy from the shower leakage of the nearby towers, so that its unclustered energy $E_{\text{tower, UE}}$ is low, as seen in Figure 3.16(a).

The η dependent UE curve in Figure 3.16(b), which is obtained by a spline fit to the $E_{\text{tower, UE}}(\eta)$ distribution, is utilized to reject EcalPlusHcalTowers. Only EcalPlusHcalTowers, which have an energy greater than that given by the curve, are added to the list of jet input 4-vectors. The detailed UE curve parametrization is available in Appendix C. The obtained UE curve depends strongly on the LHC luminosity. It has only a minor dependence on the event type and the jet finding setup used [CMS05k].

The muons, which are reconstructed as described in section 3.4.2, can be simply added to the list of jet input 4-vectors, because they do not deposit any reasonable amount of energy either in the ECAL or in the HCAL [PDG04]. Adding the reconstructed muons slightly improves the reconstructed final state jet quantities, especially for b-hadron jets, for which approximately 10% of the b-quark decays yield prompt muons (section 1.1.3).

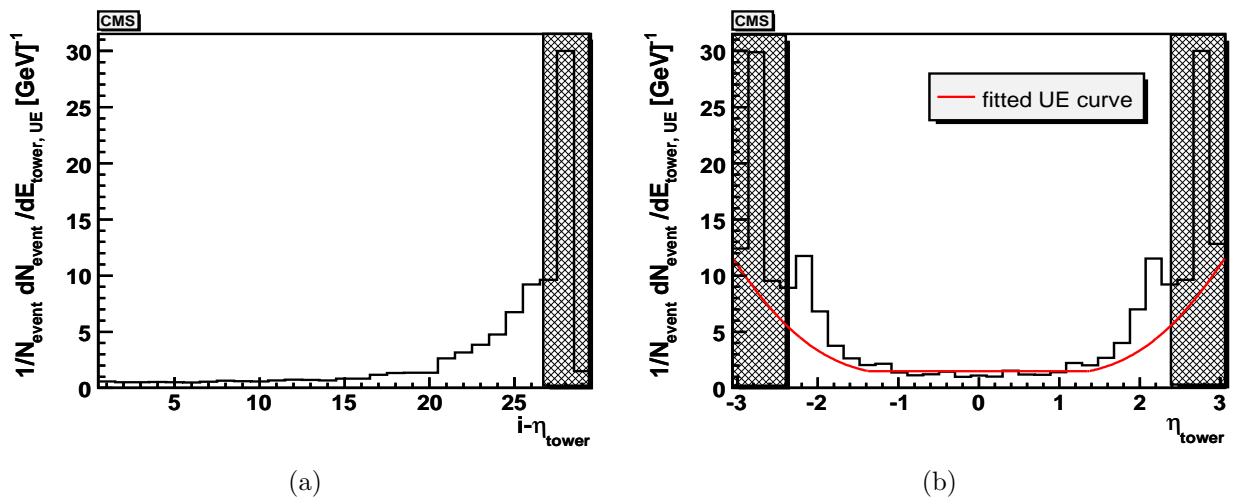


Figure 3.16: The unclustered EcalPlusHcalTower energy per event versus $i-\eta_{\text{tower}}$ (a) and η_{tower} (b). The fitted UE curve (red line) is displayed in (b). The area from where reconstructed jets are accepted $|\eta_{\text{jet}}| < 2.4$ is highlighted. This area corresponds to tower indices 1~27.

3.4.5.2 Jet algorithm: inclusive k_{\perp} jet algorithm, E_T recombination scheme

From a theoretical point of view the inclusive k_{\perp} jet algorithm is favored (section 1.1.2). From the experimental point of view choosing a jet algorithm depends on the efficiency and fake rate of the jet finding, which are defined as:

$$\text{eff}_{\text{jet}} := \frac{\# \text{ matched reconstructed jets}}{\# \text{ simulated jets}}, \quad (3.6)$$

$$\text{fake}_{\text{jet}} := 1 - \frac{\# \text{ matched reconstructed jets}}{\# \text{ reconstructed jets}}, \quad (3.7)$$

$$(3.8)$$

where:

$$\begin{aligned} \# \text{ reconstructed jets} &:= \text{number of jets made from detector objects,} \\ \# \text{ simulated jets} &:= \text{number of jets made from stable generator particles.} \end{aligned}$$

A reconstructed (simulated) jet is matched to a simulated (reconstructed) jet, if they are close in η - ϕ space: $\Delta R(\text{rec.jet}, \text{sim.jet}) < 0.2$. The matching is performed for reconstructed jets which have transverse momentum $P_{t,\text{jet},\text{rec}} > 20$ GeV. The transverse momentum of simulated jets is constraint by $P_{t,\text{jet},\text{sim}} > 10$ GeV in order to reduce the bias of the evaluation of the jet reconstruction performance due to the spread of the reconstructed transverse energy $P_{t,\text{jet},\text{rec}}$ (see Figures 3.18(a) and 3.22(a)).

Figure 3.17 shows the jet reconstruction efficiency and fake rate for the inclusive k_{\perp} jet algorithm using the E_T recombination scheme (see section 1.1.2). The TrackTower input described before is used as jet input. The final state jets are corrected with the 2-dimensional detector map described in the next section. The jet efficiency eff_{jet} reaches 80% and the jet fake rate fake_{jet} 10% for simulated transverse jet energies of $P_{t,\text{jet},\text{sim}} > 100$ GeV. The jet efficiency eff_{jet} saturates at 80% because of the used realistic event sample of $t\bar{t}$ events with low-luminosity pileup. For such dense event topologies the jet reconstruction efficiency never reaches 100%. For quite high $P_{t,\text{sim}}$ the jet efficiency and fake rates are slightly deteriorated, because of the lack of statistics in these bins.

The resolution of the jet observables $P_{t,\text{jet}}$, η_{jet} and ϕ_{jet} for uncorrected final state jets is shown in Figures 3.18(a)-(c). The r.m.s $P_{t,\text{jet}}$ -resolution in Figure 3.18(a) is about 20%, but the mean value of the distribution is 30-40% too low. The overall quite low jet response of the CMS calorimeter is caused by an artificial reduction of the HCAL response to accommodate pion showers starting in the ECAL as well as in the HCAL (the ECAL and HCAL response to pion showers is very different) [Abd]. Furthermore energy losses in uninstrumented regions (cracks) and the loss of charged particles in jets due to the CMS magnetic field are downgrading the detector response. The η_{jet} - and ϕ_{jet} -resolutions peak at 0 with reasonable spreads.

Profile histograms are utilized to further investigate the fraction $P_{t,\text{jet},\text{rec}}/P_{t,\text{jet},\text{sim}}$ of the transverse momentum of matched pairs of reconstructed and simulated jets. For the profile histograms used the markers correspond to the mean value of a Gaussian fit applied to the entries in one bin. The error bars are the standard deviation of the fitted Gaussian.

Figure 3.19(a) shows the $P_{t,\text{jet}}$ profile histogram versus the transverse momentum $P_{t,\text{jet},\text{sim}}$ of simulated jets. The rising response of the CMS calorimeter for increasing $P_{t,\text{jet},\text{sim}}$ leads to increased mean values of $P_{t,\text{jet},\text{rec}}/P_{t,\text{jet},\text{sim}}$ with a slightly reduced spread. The rise of the mean values and the increased spread in bins for which $P_{t,\text{jet},\text{sim}} < 35$ GeV is caused by a threshold effect. For these bins the entries do not follow a Gaussian distribution. The $P_{t,\text{jet}}$ profile histogram versus η_{jet} displayed in Figure 3.19(b) shows an overall flat distribution with

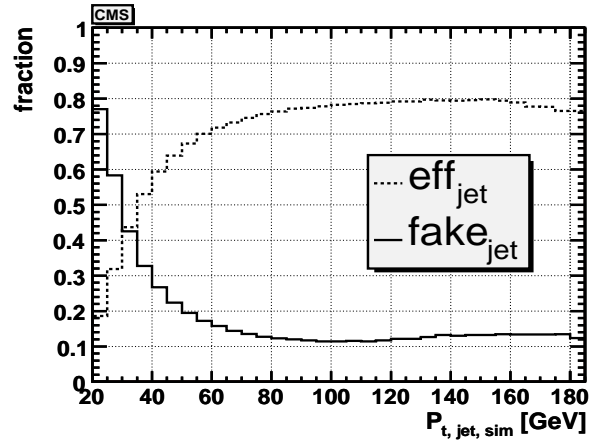
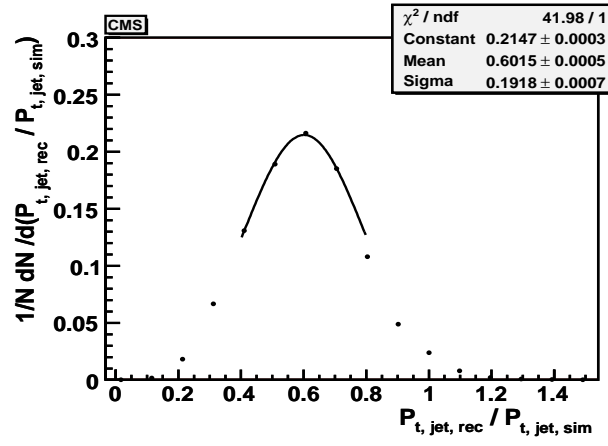
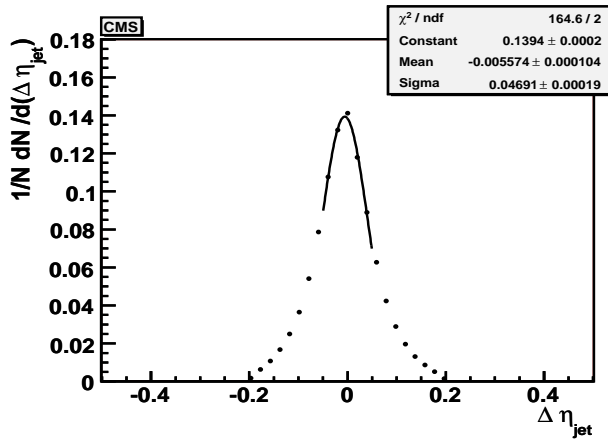


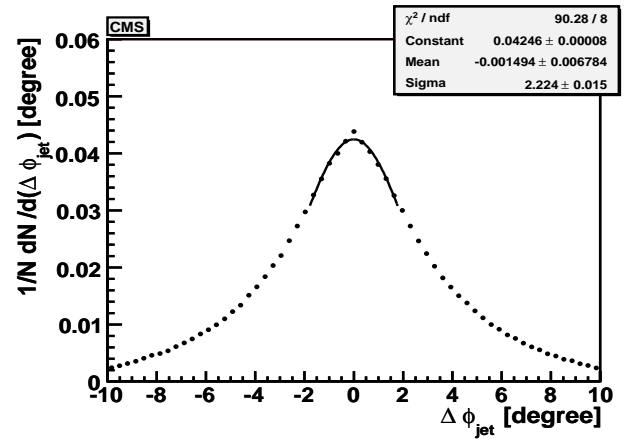
Figure 3.17: The jet reconstruction efficiency eff_{jet} and fake rate fake_{jet} for the inclusive k_{\perp} jet algorithm using the E_T recombination scheme and the TrackTower jet input. The jets are corrected with the 2-dimensional detector map.



(a)



(b)



(c)

Figure 3.18: The $P_{t,\text{jet}}$ -, η_{jet} - and ϕ_{jet} -resolutions of uncorrected jets computed with the inclusive k_{\perp} jet algorithm using the E_T recombination scheme and the TrackTower jet input.

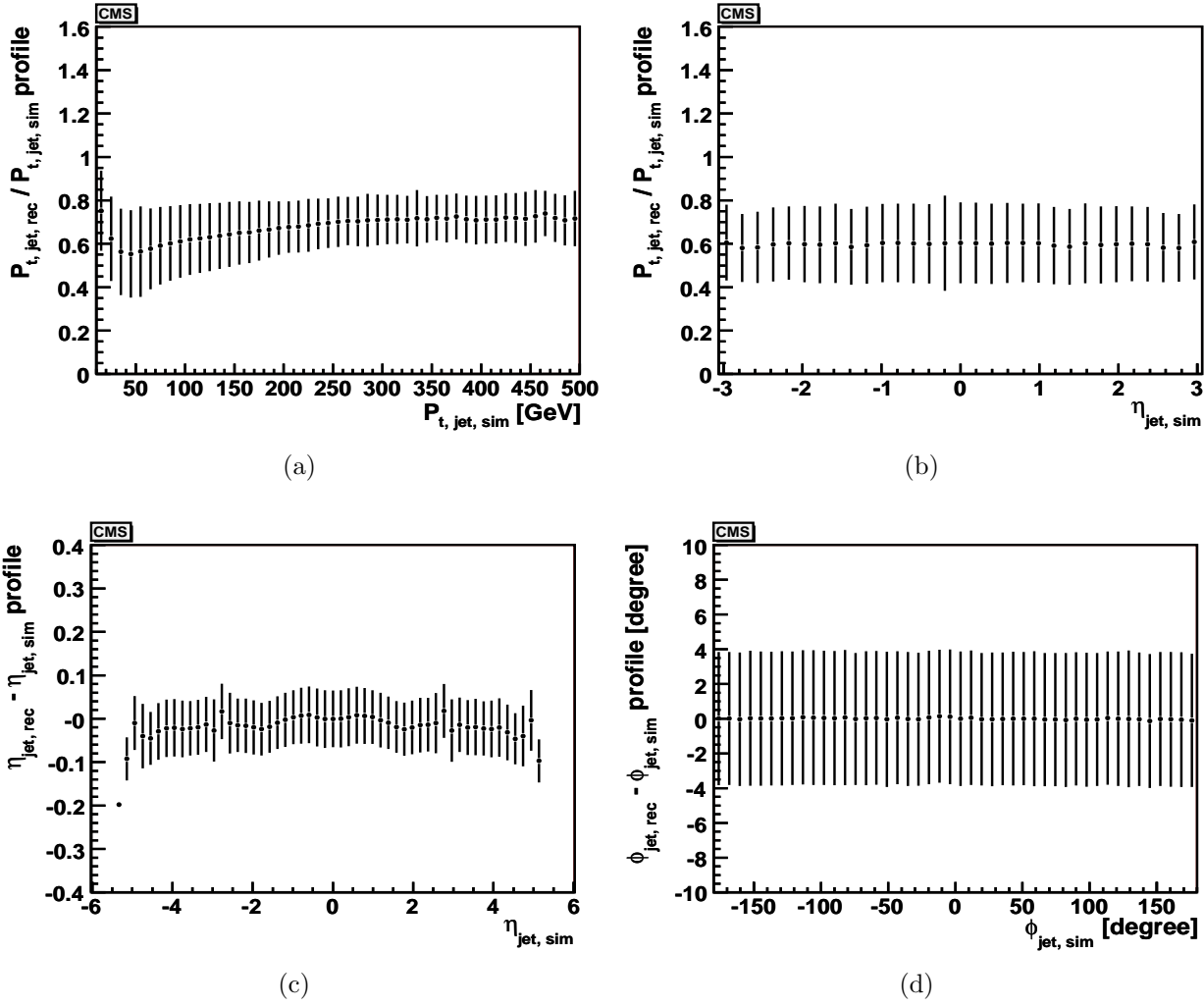


Figure 3.19: Profile histograms for uncorrected jets of the jet observables $P_{t,jet}$, η_{jet} and ϕ_{jet} . The $P_{t,jet}$ profiles are shown versus $P_{t,jet,sim}$ (a) and $\eta_{jet,sim}$ (b).

a mean value of 0.6 and a spread of about 20%. The plot corresponds to the $P_{t,jet}$ -resolution plot shown in Figure 3.18(a) which provides the same information integrated over all η_{jet} . The η_{jet} and ϕ_{jet} profile histograms (Figures 3.19(c) and (d)) have a mean value of 0 and a spread which corresponds to the Figures 3.18(b) and (c), respectively.

All features of the performance of CMS calorimetry and jet reconstruction shown are used as foundation for reference [CMS05k] and therefore in agreement to it.

3.4.5.3 Jet calibration: 2-dimensional detector map

As noted previously the reconstructed jets have a transverse momentum which is on average 30-40% below that of the corresponding simulated jets. To correct the transverse energy (momentum) of the reconstructed jets for the effects described before a 2-dimensional detector map is utilized¹².

The correction depends on the reconstructed transverse energy E_t and the reconstructed $|\eta|$

¹²Jets reconstructed by means of the inclusive k_{\perp} jet algorithm using the E_T recombination scheme are mass-less.

It follows the transverse jet energy $E_{t,jet}$ is equal to the transverse jet momentum $P_{t,jet}$ (see section 1.1.2).

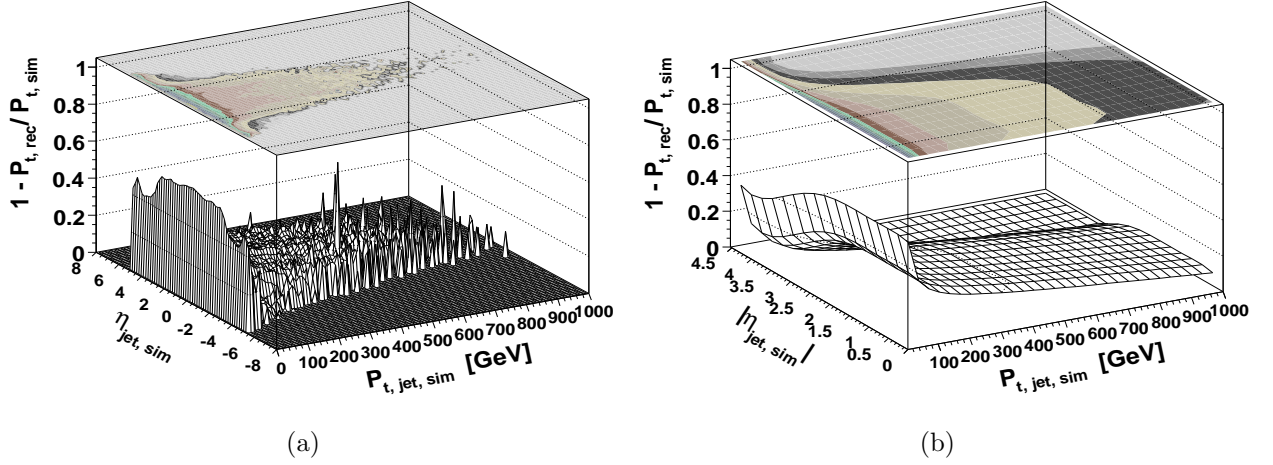


Figure 3.20: The scaled 2-dimensional $P_{t,jet}$ profile histogram (a) which is used to determine the 2-dimensional detector map function $D_{map,A}$ plotted for $\eta \geq 0$ in (b). The function $D_{map,A}$ is symmetric for positive and negative values of η .

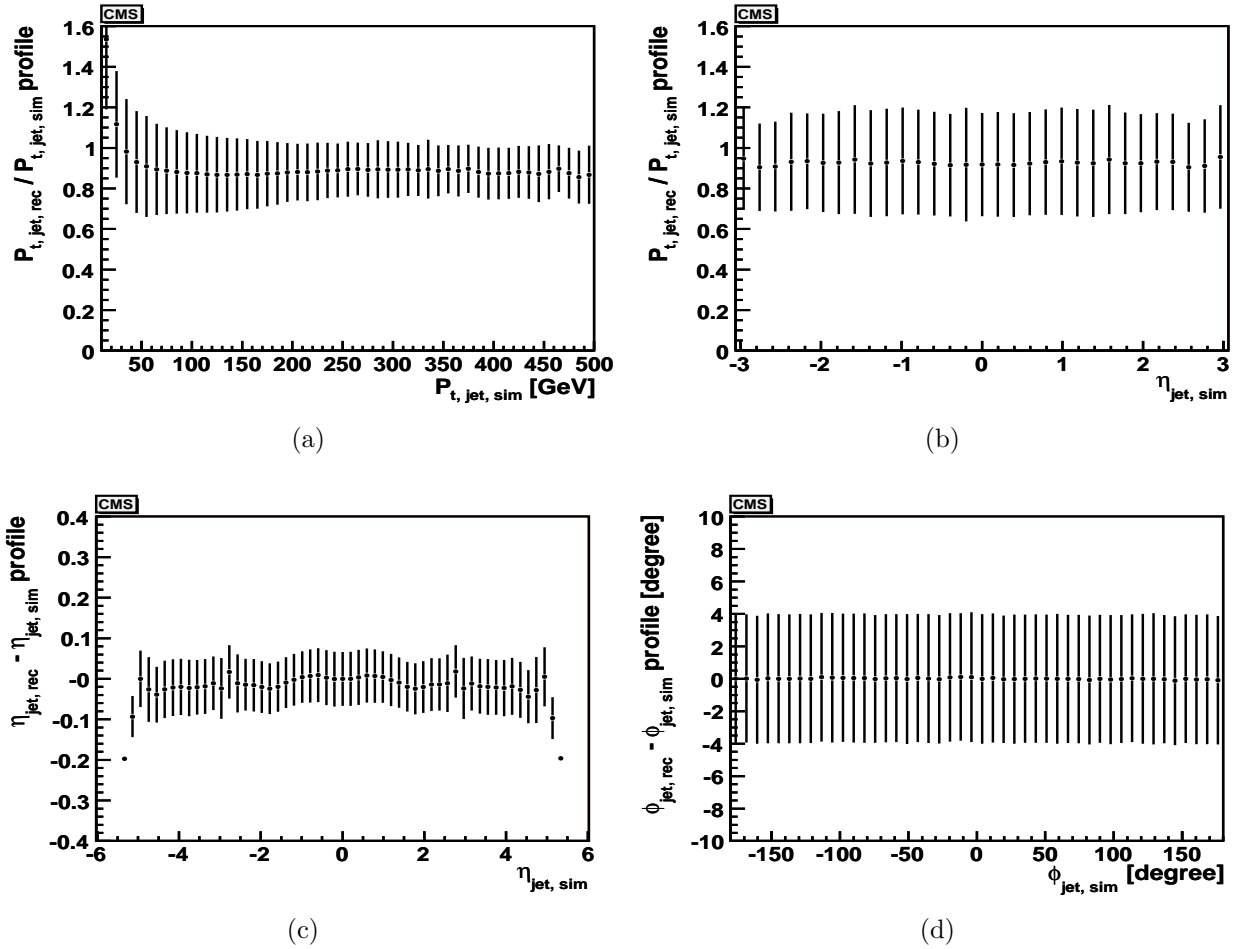


Figure 3.21: Profile histograms for corrected jets of the jet observables $P_{t,jet}$, η_{jet} and ϕ_{jet} . The $P_{t,jet}$ profiles are shown versus $P_{t,jet,sim}$ (a) and $\eta_{jet,sim}$ (b).

of the uncorrected jet. The correction factor E'_t is defined as:

$$E'_t = E_t \cdot (1 + D_{\text{map}}(E_t, |\eta|)), \quad (3.9)$$

where $D_{\text{map}}(E_t, |\eta|)$ is the detector map function. The function is obtained from a 2-dimensional fit to the entries of a scaled 2-dimensional $P_{t,\text{jet}}$ profile histogram showing $1 - P_{t,\text{jet,rec}}/P_{t,\text{jet,sim}}$ versus $P_{t,\text{jet,sim}}$ and $|\eta_{t,\text{jet,sim}}|$, neglecting the threshold effect for lower $P_{t,\text{jet,sim}}$ bins.

In the Figures 3.20(a) and (b) the 2-dimensional $P_{t,\text{jet}}$ profile histogram for uncorrected jets and the corresponding 2-dimensional function D_{map} are shown, respectively. Further details of the correction are described in Appendix D.

The resulting profile histograms for the final state jets after the correction is applied are shown in Figures 3.21(a)-(d). After correction the $P_{t,\text{jet}}$ profile histograms shown in Figures 3.21(a) and (b) have a mean value which is only 10% too low. The resolution of these histograms is still about 20%. The η_{jet} and ϕ_{jet} profile histograms (Figures 3.21(c) and (d)) show a slightly worse resolution with respect to the profiles for the uncorrected jets (Figures 3.19(c) and (d)).

The resolution of the jet observables $P_{t,\text{jet}}$, η_{jet} and ϕ_{jet} is plotted in Figures 3.22(a)-(c) and confirm the observations of the profile histograms of Figures 3.21(a)-(d).

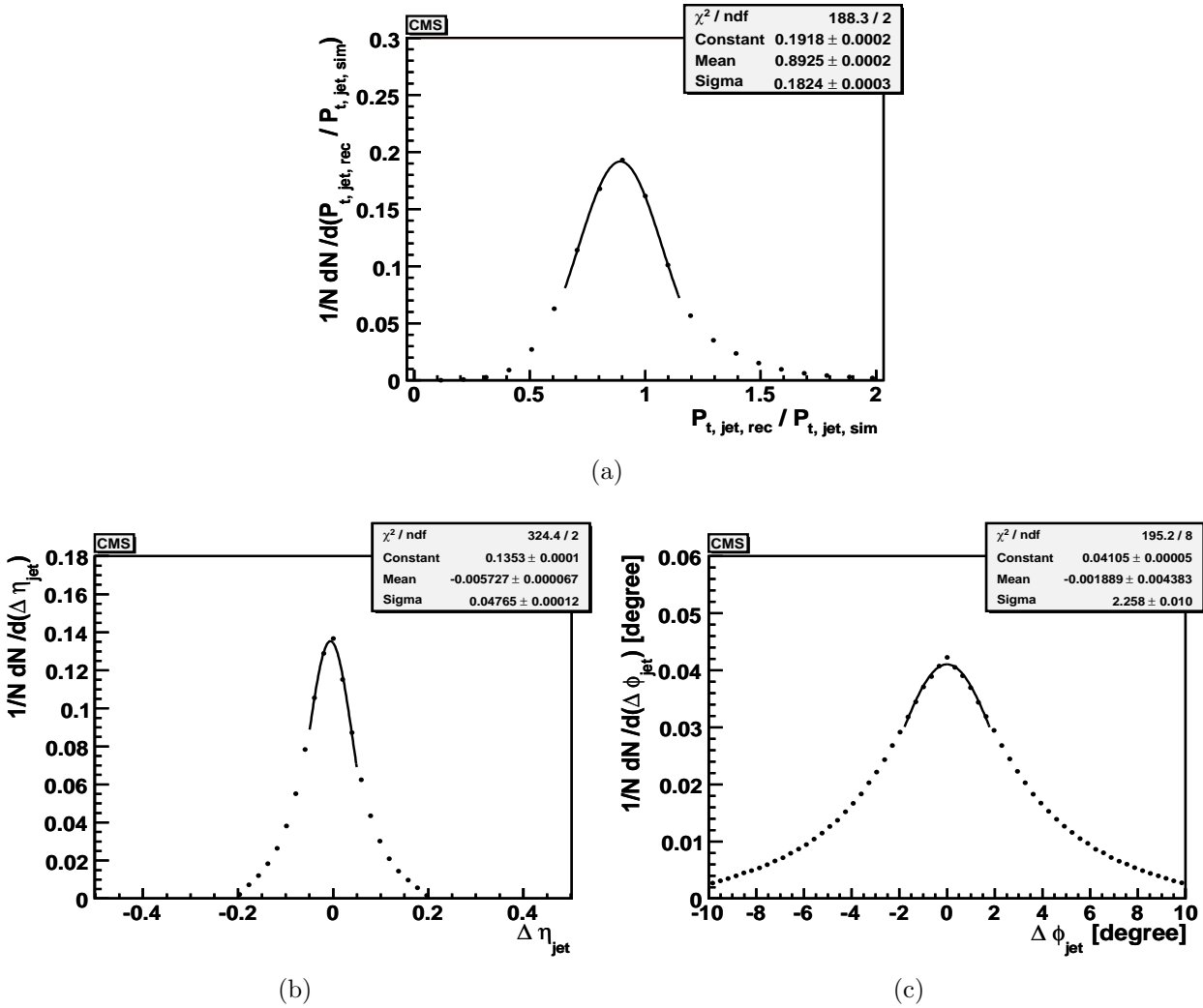


Figure 3.22: The $P_{t,\text{jet}}$ -, η_{jet} - and ϕ_{jet} -resolutions of corrected jets computed with the inclusive k_{\perp} jet algorithm using the E_T recombination scheme and the TrackTower jet input.

The 2-dimensional detector map correction of final state jets enables the usage of reconstructed jets which have a transverse energy greater than about 20 GeV, which is sufficient for this study.

3.4.6 b-hadron identification

The strategy for b-hadron jet identification in the CMS data is comparable to that used for ALEPH (section 2.4.4). Again the observables used in combination are: the jet lifetime probability P_{jet} , the secondary vertex mass $M_{2\text{nd} vtx}$, the rapidity $\eta_{2\text{nd} vtx, track, jet}$ and the jet energy fraction X_{energy} of secondary vertex tracks. For these observables the secondary vertex reconstruction described in section 3.4.3 is used. The significance of the signed 3D impact parameter $\frac{D}{\sigma}$ for reconstructed tracks is required to calculate the jet lifetime probability P_{jet} (section 1.1.3) used in the b-tag.

The first part of this section describes the work done to calculate the track probability P_{track} from the significance of the signed 3D impact parameter $\frac{D}{\sigma}$. P_{track} is then used to calculate the jet lifetime probability P_{jet} . The observables used in the b-tag and their combined b-hadron identification capabilities for CMS are described in the second part of this section.

3.4.6.1 Significance of the signed 3D impact parameter

The most crucial point in computing a reasonable track probability P_{track} from the signed 3D impact parameter $\frac{D}{\sigma}$ is to use only tracks, with a high probability to come directly from the b-hadron decay in the jet under test (primary decay tracks). Figure 3.23 shows a typical complex LHC event topology, where the signal interaction, which consists of 2 jets, is accompanied by two pileup interactions.

To extract the b-hadron lifetime information from the signal b-hadron decay the tracks from the pileup interactions have to be suppressed. In general, several vertices may occur in hadronic

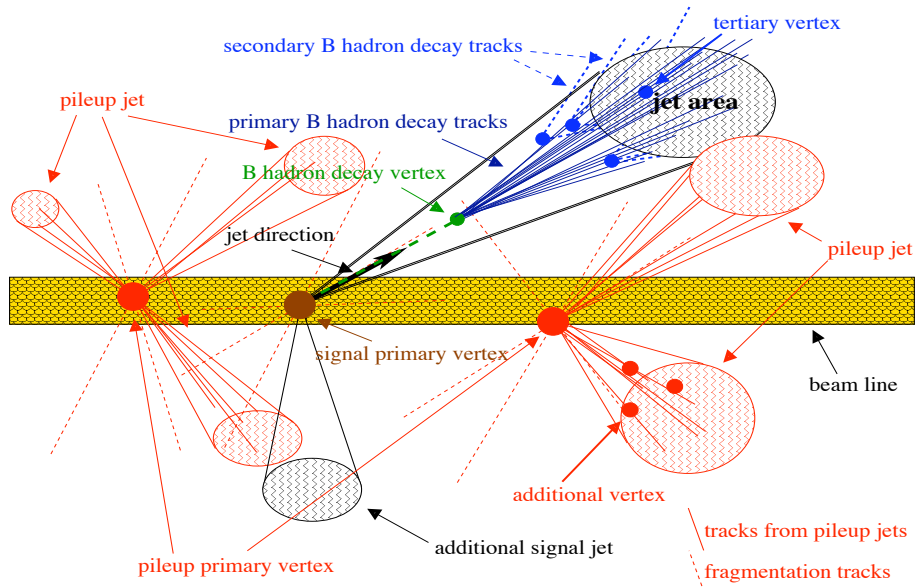


Figure 3.23: A typical complex LHC event topology for the low-luminosity phase. Besides the signal interaction, which consists of 2 jets, additional jets from two pileup interactions are visible in this example.

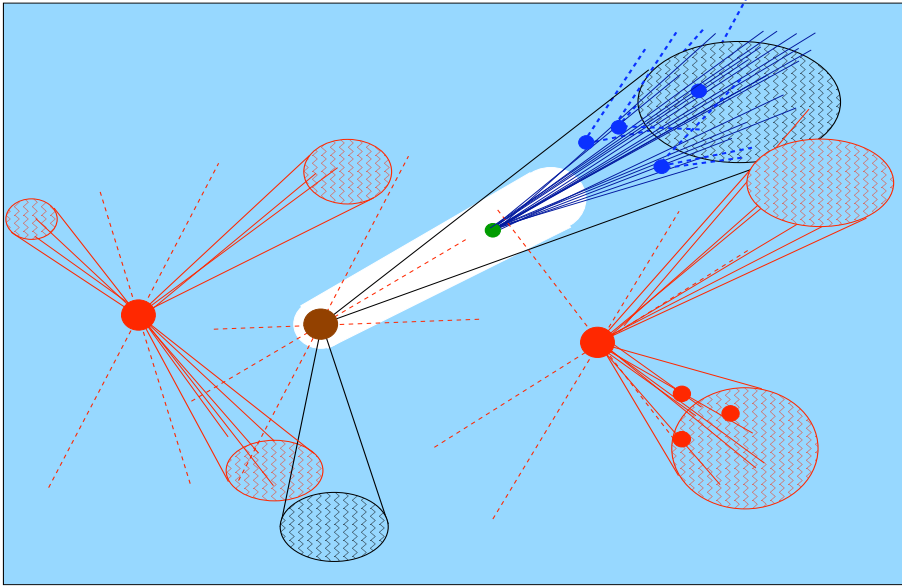


Figure 3.24: The typical complex LHC event topology for the low-luminosity phase which is cleaned up to extract the b-hadron lifetime information. The clean-up is done by means of appropriate constraints to the reconstructed objects of the event.

showers forming jets. In Figure 3.23 additional and tertiary vertices are drawn for two jets. The tracks from these vertices must also be suppressed in order to obtain the proper b-hadron lifetime information. Figure 3.24 illustrates the chosen strategy, which consists of using only tracks originating within a limited tubular region. The criteria that define the shape of the tube, are described in the following.

The signal vertex is reconstructed and tagged as described in sections 3.4.3 and 3.4.3.4. The jet axis is reconstructed by means of the inclusive k_{\perp} jet algorithm described in section 3.4.5. Charged tracks are reconstructed and selected as described in section 3.4.2.1. To ensure a reasonable track measurement accuracy near the signal primary vertex the resulting reconstructed tracks are further constrained by demanding at least 2 pixel hits and a transverse momentum of $P_{t, \text{track}} > 2 \text{ GeV}$.

Reconstructed tracks which are closer than 0.8 to the reconstructed jet axis in η - ϕ space are selected. Figure 3.25 shows the distribution of $\Delta R(\text{track}, \text{jet})$ for b-, c-, uds- and gluon-jets. The cut $\Delta R(\text{track}, \text{jet}) < 0.8$ is chosen to select the core tracks of b-hadron jets.

The 3-dimensional distance of reconstructed tracks to the reconstructed signal primary vertex $D_{3D, \text{track}, \text{vtx}}$ is used to reject tracks from pileup interactions. The full track helix is utilized to calculate this distance which is shown in Figure 3.26(a) for different jet types. Tracks which are too far away from the signal primary vertex do not contain any visible lifetime information so that these tracks are rejected if $D_{3D, \text{track}, \text{vtx}} > 0.5 \text{ cm}$.

To reject possible tertiary or higher order vertices the distance along the jet S_j (see Figure 1.1(b)) is utilized. Figure 3.26(b) shows the distribution of S_j for the different jet types. Again the lifetime signal vanishes for too high values of $|S_j|$ so that only tracks for which $|S_j| < 10 \text{ cm}$ are accepted. The signed 3D impact parameter D and its error $\sigma(D)$ for b-, c-, uds- and gluon-jets are shown in Figures 3.26(c) and (d) respectively. As discussed before, the distribution of the impact parameter D leads to the requirement $|D| < 0.4 \text{ cm}$ for reconstructed tracks to be accepted¹³. To avoid unreasonable impact parameter errors $\sigma(D)$ tracks for which

¹³Negative values for S_j and D are allowed to preserve the negative tail of the signed 3D impact parameter

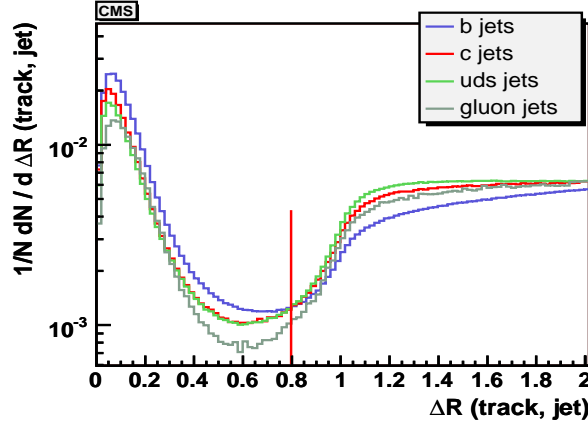
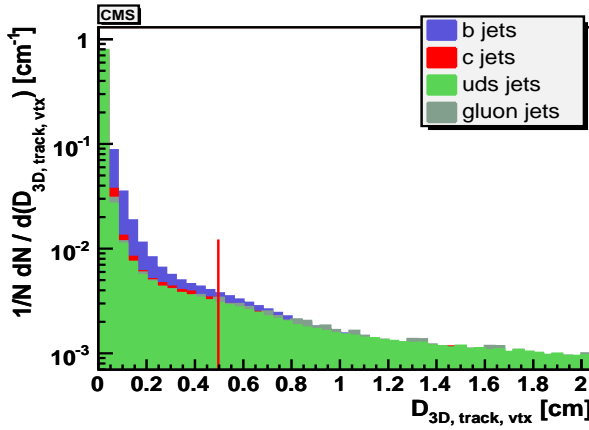
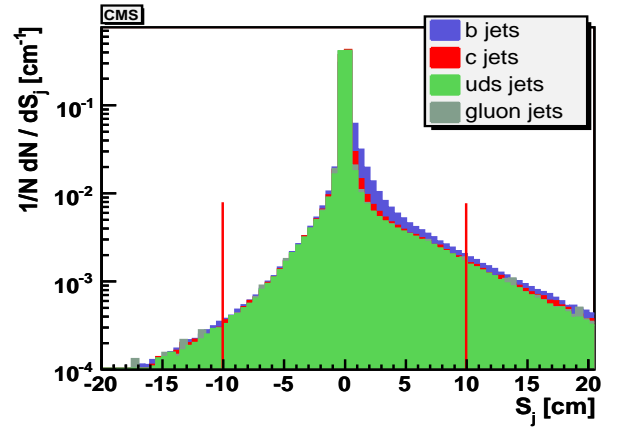


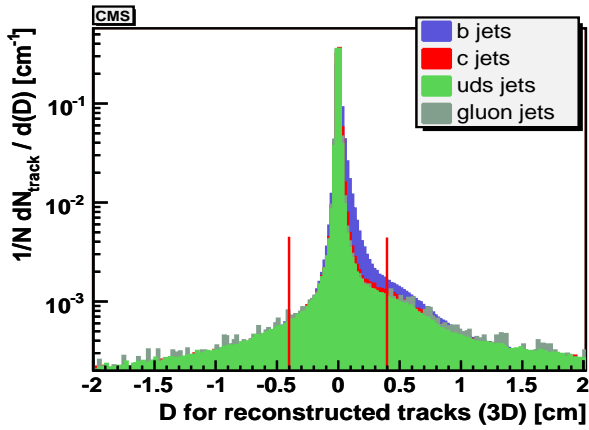
Figure 3.25: The η - ϕ distance $\Delta R(\text{track, jet})$ of reconstructed jets and tracks separately for b-, c-, uds- and gluon-jets. The used cut of $\Delta R(\text{track, jet}) < 0.8$ is highlighted (red vertical line).



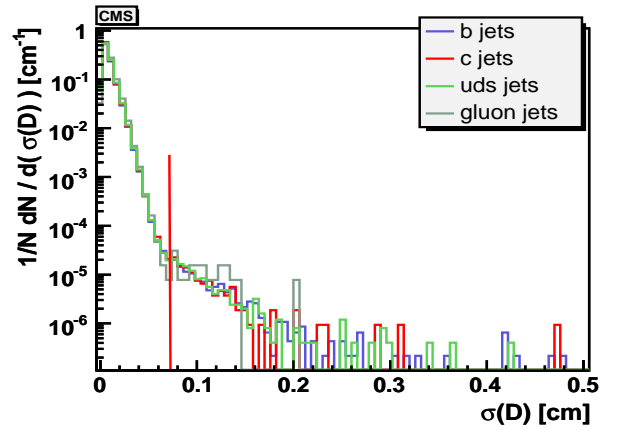
(a)



(b)



(c)



(d)

Figure 3.26: The 3-dimensional distance of reconstructed tracks to the reconstructed signal primary vertex $D_{3D, \text{track}, \text{vtx}}$ (a), the distance along the jet S_j (b), the signed 3D impact parameter D (c) and its error $\sigma(D)$ (d) for b-, c-, uds- and gluon-jets. The cuts applied to this observables are highlighted (vertical red lines).

distribution $\frac{D}{\sigma}$, which is essentially needed to compute the resolution function (see section 1.1.3).

$P_{t, \text{track}} > 2 \text{ GeV}$
$\Delta R(\text{track}, \text{jet}) < 0.8$
$ D_{3D, \text{track}, \text{vtx}} < 0.5 \text{ cm}$
$ S_j < 10 \text{ cm}$
$D < 0.4 \text{ [cm]}$
$\sigma(D) < 0.075 \text{ cm}$

Table 3.11: Summary of specific constraints to define the set of reconstructed tracks from which the signed 3D impact parameter significance $\frac{D}{\sigma}$ is computed.

$\sigma(D) > 0.075 \text{ cm}$ are rejected.

All constraints used to define the limited tube shown in Figure 3.24 are summarized in Table 3.11. The resulting distribution of the significance of the signed 3D impact parameter $\frac{D}{\sigma}$ separated for different jet types is shown in Figure 3.27(a). The b-hadron lifetime signal is clearly visible for positive $\frac{D}{\sigma}$. The negative tails which are quite similar for the different jet types are used in combination to compute the resolution function $\mathcal{R}(x)$ as described in section 1.1.3.

The entries in Figure 3.27(b) have been fitted with the expectation of a Gaussian resolution ($xe^{-x^2/2\sigma^2}$) plus two additional independently exponentially decreasing tail terms ($xe^{-x/\lambda}$). The resulting equation has been normalized with MAPLE [Map] to obtain a properly defined probability density function. The fit is a direct measurement of the resolution function $\mathcal{R}(x)$ of the signed 3D impact parameter significance [ALEPH92] and should be obtained from real data only.

The jet lifetime probability P_{jet} , calculated by means of equations (1.7) and (1.8) is shown in Figure 3.28(a) for different jet types. The distributions show the typical shapes and the b-hadron lifetime is visible for high P_{jet} values. The common decrease of the P_{jet} distributions at lower bins occurs because of the strong constraints imposed on the tracks in the calculation to stem from a b-hadron decay.

The evaluation of the b-hadron identification performance is done by plotting the b-tag

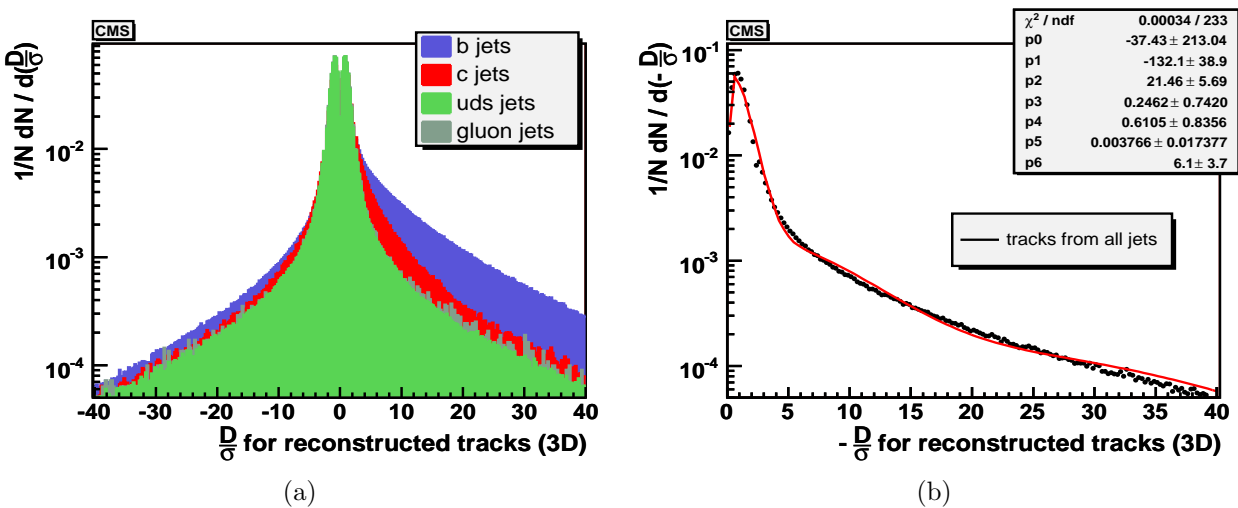


Figure 3.27: The significance of the signed 3D impact parameter $\frac{D}{\sigma}$ for different jet types (a). The b-hadron lifetime signal is clearly visible for positive $\frac{D}{\sigma}$. The negative tail of the $\frac{D}{\sigma}$ distribution and the fit to obtain the resolution function (red line) are shown in (b).

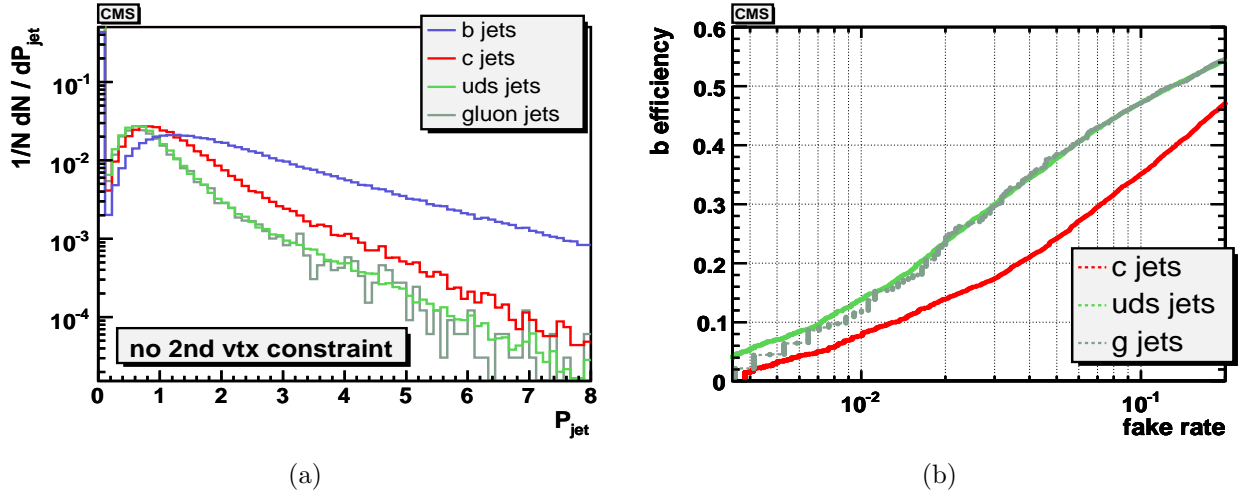


Figure 3.28: The jet lifetime probability tag P_{jet} (a) and its performance (b) for different jet types.

efficiency versus fake rate as defined in section 1.1.3. To obtain the true jet identification by means of the information available from the generator the reconstructed jets are matched to the generator particles. A jet-particle pair is matched if they are closer than $R = 0.2$ in η - ϕ space. No isolation of such pairs is required in order not to hide any problems which could occur, *e.g.* in the jet axis reconstruction.

Figure 3.28(b) shows the reasonable performance of the jet lifetime probability P_{jet} which is comparable to the results described in [CMS02b].

3.4.6.2 Observables used in the new combined b-tag and its performance

The new CMS combined b-tag for a given jet uses the same observables as ALEPH (section 2.4.4.1). The observables used are: the jet lifetime probability P_{jet} , the secondary vertex mass $M_{2nd vtx}$, the rapidity $\eta_{2nd vtx, track, jet}$ and the jet energy fraction X_{energy} of secondary vertex tracks¹⁴. Their distributions after requiring a well reconstructed secondary vertex (section 3.4.3) are shown in Figures 3.29(a)-(d). All four observables show a similar shape as in the ALEPH study (Figures 2.9(a)-(d)) and the different response for the different hadron jet types is clearly visible. Due to the requirement of a significant secondary vertex, a considerable amount of uds- and gluon-jets has already been rejected. The distributions of the observables are quite similar for uds- and gluon-jets, though some differences are visible in the distributions for $\eta_{2nd vtx, track, jet}$ and X_{energy} (Figures 3.29(c) and (d)).

Because of the huge amount of gluon-jets in LHC events (section 1.1.3) the combination of the 4 b-tag observables is done using equation (1.10). This equation employs additional separate weights for the gluon-jet contribution in contrast to equation (1.9) used in the new ALEPH b-tag. The normalization factors n_c , n_{uds} , n_g are also calculated from the used $t\bar{t}$ event sample as outlined in section 1.1.3. As described before this sample includes the contribution of low-luminosity pileup events.

The $PMRX_{tag}$ observable and its performance is shown in Figures 3.30(a) and (b) respectively. The combined observable $PMRX_{tag}$ shows a reasonably good separation of the different

¹⁴The tracks are combined assuming that they have a mass of zero.

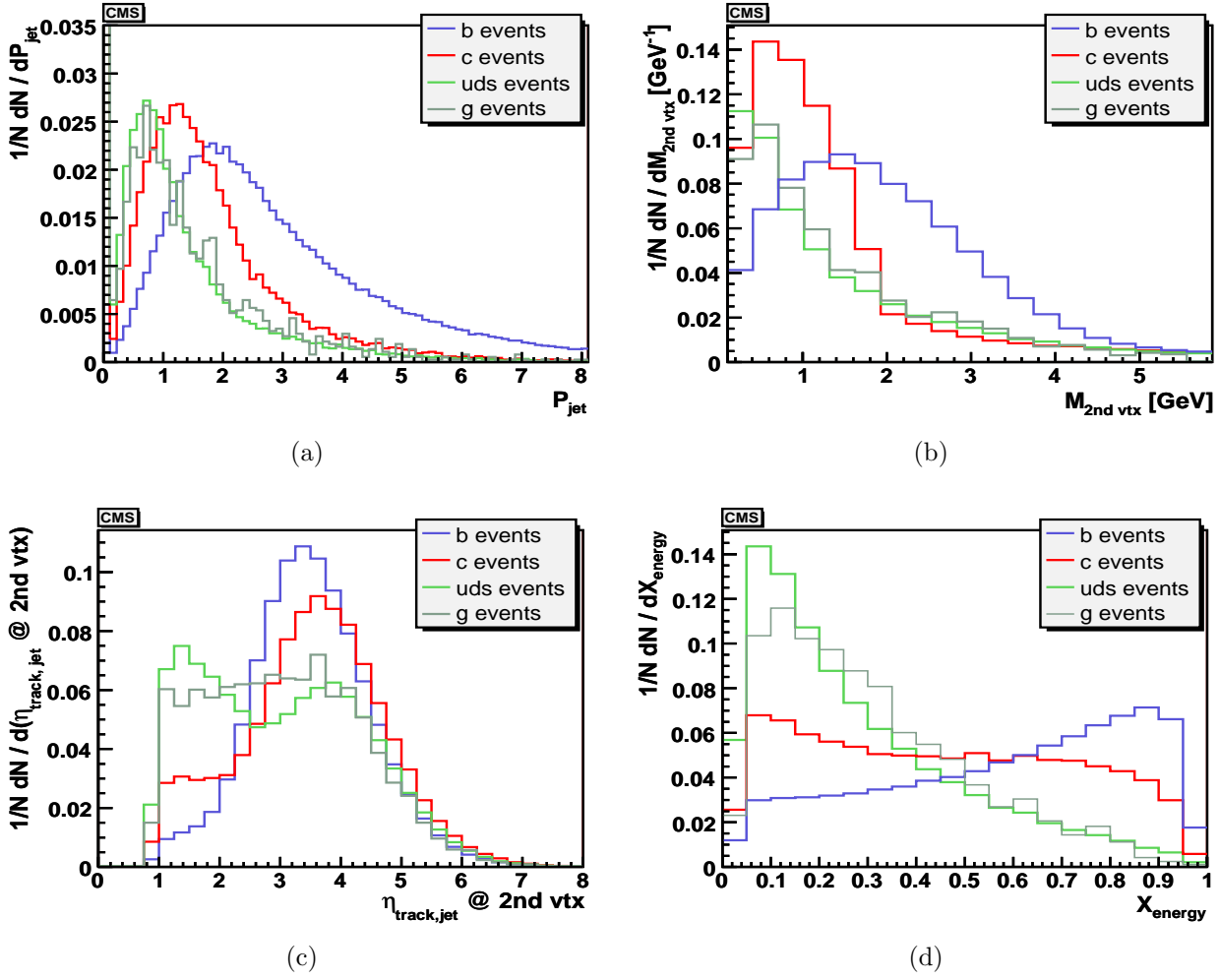


Figure 3.29: The jet lifetime probability P_{jet} (a), the secondary vertex mass $M_{2\text{nd vtx}}$ (b), the rapidity $\eta_{2\text{nd vtx, track, jet}}$ of secondary vertex tracks (c) and the jet energy fraction X_{energy} of secondary vertex tracks (d) for CMS.

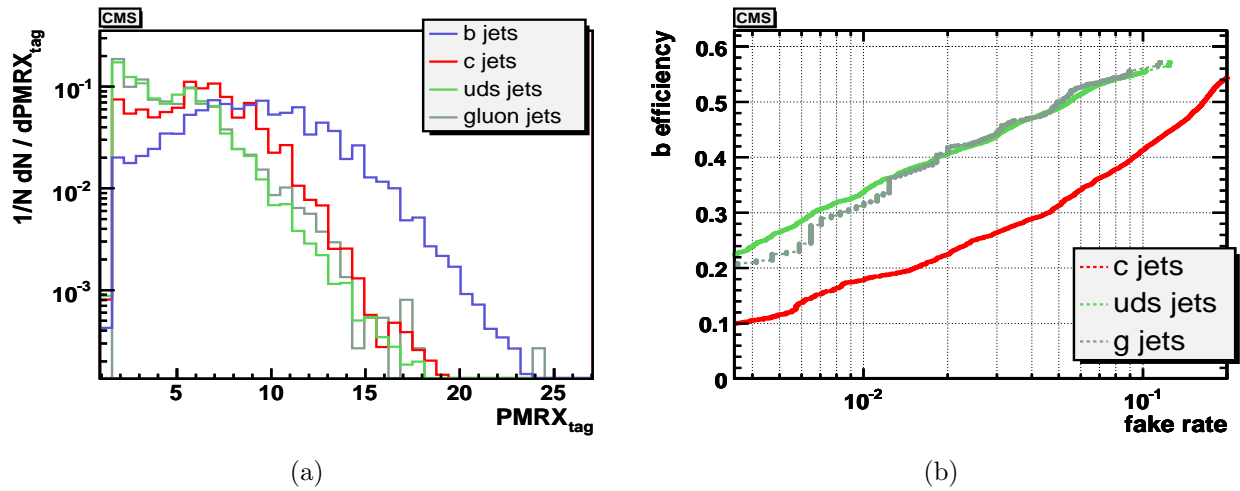


Figure 3.30: The distribution of the combined $PMRX_{\text{tag}}$ observable (a) and its performance (b) for different jet types.

jet types and the b-hadron lifetime information is clearly visible for high $PMRX_{\text{tag}}$ values in Figure 3.30(a).

The b-hadron identification performance by using the $PMRX_{\text{tag}}$ observable (Figure 3.30(b)) is significantly improved with respect to using the jet lifetime probability P_{jet} alone (Figure 3.28(b)). To put the $PMRX_{\text{tag}}$ b-tag into perspective two example comparisons of the performances of these two tags are given:

- **Common fake rate:** At the same fake rate of 1% for c-jets for both tags, the b-jet selection efficiency increased roughly by a factor of two from 10% to 20% by using the $PMRX_{\text{tag}}$. The b-jet efficiency is improved from 15% to 35% by comparing the performance at the same fake rate of 1% for uds- and gluon-jets.
- **Common b-jet selection efficiency:** At a b-jet efficiency of 30% for both tags the c-jet fake rate is lowered from 7% to 5% and the uds-jet fake rate from 3% to 0.7% for the $PMRX_{\text{tag}}$.

3.4.7 Physics Analysis of WH, $H \rightarrow b\bar{b}$

The reconstruction described in the previous sections including the $PMRX_{\text{tag}}$ b-tag is used to estimate the CMS discovery potential for the WH, $H \rightarrow b\bar{b}$ channel in low-luminosity conditions. In section 1.3.2 it was pointed out that the discovery of the unassociated $H \rightarrow b\bar{b}$ final state is almost impossible because of the huge amount of backgrounds. To improve this situation the leptonic decay of the associated Higgs partner is utilized to further constrain the final state. The used cross section for the WH signal process is listed in Table 3.8.

Only W boson decays into final states which contain an electron or muon, respectively, are used in this study¹⁵. The electrons and muons of the WH signal final state are used to define the L1 and HLT trigger conditions (section 3.3.5). Furthermore the reconstructed transverse W mass is used to enhance the fraction of selected signal events.

The backgrounds which are considered in this study, namely $t\bar{t}$, Wbb, ZW, WW and ZZ, have large cross sections (Table 3.8) and consist of final states which look similar compared to the WH signal.

Details about the analysis strategy and the obtained results are described in the following.

3.4.7.1 L1 and HLT trigger selection

The WH final state, which contains one electron or muon from the W boson decay, is recorded by using the single electron and single muon triggers, respectively, for both the L1 and for the HLT trigger. The description of the methods used to reconstruct these leptons for the L1 and HLT decision can be found in [CMS00b] and [CMS02b], respectively. The default L1 and HLT Trigger tables for low-luminosity are used. The thresholds of the single electron and single muon trigger are listed in table 3.12.

Trigger Condition	L1 threshold	HLT threshold
Single muon	$P_t > 14 \text{ GeV}$	$P_t > 19 \text{ GeV}$
Single electron	$E_t > 23 \text{ GeV}$	$E_t > 26 \text{ GeV}$

Table 3.12: L1 and HLT single electron and single muon trigger thresholds for low-luminosity [CMS00b, CMS02b].

3.4.7.2 W mass constraint

The transverse mass m_W^T of the W boson associated with the production of the Higgs is used to further constrain the events which pass the L1 and HLT triggers.

To compute m_W^T from the W decay into a muon and neutrino, the highest P_t muon with a transverse momentum $P_t > 19 \text{ GeV}$ and the reconstructed missing transverse energy of the event is used. The muon and missing transverse energy are reconstructed as described in sections 3.4.2.2 and 3.4.4 respectively. The obtained distribution of the transverse W mass m_W^T is shown in Figure 3.31(a). It is quite broad and peaks around 40 GeV which is 50% lower than the world average W boson mass of $m_W \approx 80 \text{ GeV}$ [EWWG06]. The distribution reflects the loose constraints on the muon selection, the contamination of the missing transverse energy calculation from other sources, *e.g.* additional neutrinos from other interactions in the same event, and the degeneration of the transverse W mass resolution by using only the x and y

¹⁵In the following all final states of the W boson containing an electron, positron, μ^+ or μ^- , respectively, are considered. For convenience they are referred to as electron or muon final states.

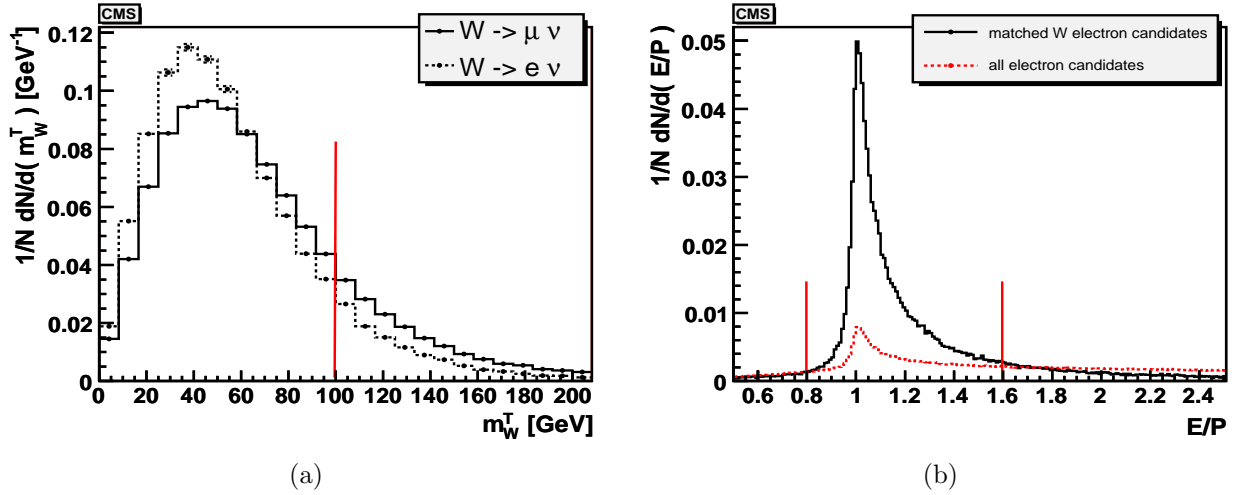


Figure 3.31: The transverse W mass m_W^T reconstructed from the missing transverse energy of the event and from a reconstructed muon or electron, respectively (a). The E/P distribution for true electron candidates and for all electron candidates with $E_t > 26$ GeV is shown in (b). The selection criteria are marked with (red) vertical lines.

coordinates of the reconstructed objects. Nevertheless, these reconstructed W boson candidates can be used to further constrain the final state of the accepted events as described below.

Electron candidates are reconstructed as described in section 3.4.2.3. To reduce the number of fake electrons, the candidates having a transverse energy $E_t > 26$ GeV are further constrained according to their E/P ratio. This ratio is shown in Figure 3.31(b) for all electron candidates and for candidates which are matched to a true generated electron from W decay. The matching is done by requiring $\Delta R(\text{rec.candidate}, \text{gen.electron}) < 0.1$. The transverse mass m_W^T calculated from the highest P_t electron candidate for which $0.8 < E/p < 1.6$ and from the reconstructed missing transverse energy of the event is shown in Figure 3.31(a). The distribution is similar to that obtained for muons. Its peak is slightly shifted to lower values of m_W^T .

Because of the limited performance of the W boson mass reconstruction, all events for which $m_W^T < 100$ GeV are accepted for further analysis. This constraint is used for both decay channels.

3.4.7.3 Further constraints to improve the event selection

The $t\bar{t}$ production has a large cross section at the LHC (table 3.8). It turns out that additional constraints are needed to reduce especially this background, because the fundamental cuts of b-tag and W selection are satisfied.

The t-quark decays predominantly into a b-quark and a W boson [PDG04], so that the probability of additional leptons and jets in $t\bar{t}$ events is increased with respect to WH events. The higher number of leptons and jets is utilized to enhance the selection of WH signal events, *e.g.* by requiring exact one lepton. Also final states of ZW, WW and ZZ events, which contain more than one reconstructible lepton, *e.g.* a $Z \rightarrow ll$, $W \rightarrow l\nu$ final state, can be suppressed in this way. Final states which look almost like the WH decay, *e.g.* $Wb\bar{b}$ events for which the W decayed leptonically, are very difficult to suppress.

All constraints used are listed in table 3.13 separately for the selection of the W decay channel into electrons or muons. The $PMRX_{\text{tag}}$ b-tag requires a well reconstructed secondary vertex

constraint	W decay channel	
	electron	muon
<i>L1 trigger</i>	Single electron	Single muon
<i>HLT</i>	Single electron	Single muon
<i>b-jet</i>	2 tagged b-jets ($PMRX_{\text{tag}}$ b-tag)	
<i>W mass</i>	electron $E_t > 26 \text{ GeV}$, $0.8 < E/p < 1.6$	muon $P_t > 19 \text{ GeV}$
	$m_W^T < 100 \text{ GeV}$	
<i>same lepton veto</i>	no add. electron $E_t > 26 \text{ GeV}$	no add. muon $P_t > 19 \text{ GeV}$
<i>jet veto</i>	no additional jet $P_t > 20 \text{ GeV}$	
<i>wrong lepton veto</i>	no add. muon $P_t > 19 \text{ GeV}$	no add. electron $E_t > 26 \text{ GeV}$

Table 3.13: The details of the event selection criteria for the WH analysis separately for the W decay channel into final states containing an electron or muon.

(see section 3.4.6). This requirement is the minimal constraint to positively identify a jet as b-jet. If more than two jets are identified as b-jet the two jets which have the highest b-tag value are chosen. Jets which have a high $PMRX_{\text{tag}}$ value have a higher probability to stem from a true b-hadron decay, because the fraction of c- and uds-jets positively identified as b-jets decreases for higher $PMRX_{\text{tag}}$ values with respect to the fraction of correctly identified b-jets (Figure 3.30(a)).

Figures 3.32(a) and (b) show the decreasing fractions of accepted signal and background events for each new constraint applied. Clearly the WH signal events but also the Wbb events are favored by both selections. After all constrains are applied the fractions of accepted ZW and WW events are roughly a factor of two lower compared to WH events. The selection efficiency of $t\bar{t}$ and ZZ events can be significantly diminished by almost one order of magnitude by means of the *jet veto* constraint.

Because of the very different cross sections for the WH signal and the various backgrounds the number of finally accepted signal events is small compared to the number of selected background

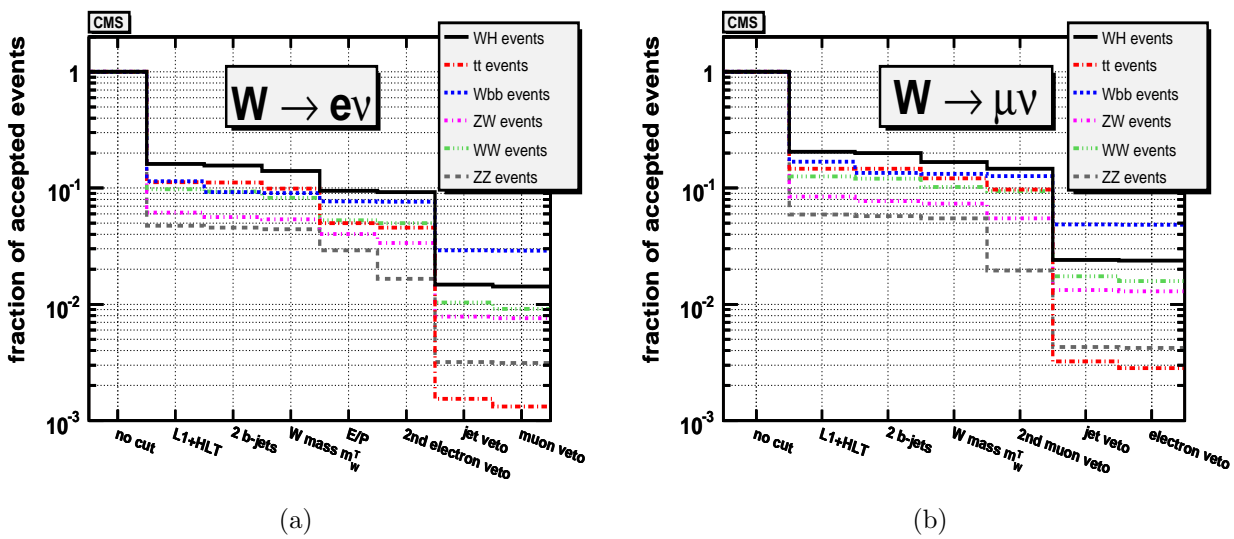


Figure 3.32: The fraction of accepted events in the two W boson decay channels separately for the WH signal and the considered backgrounds.

constraint	electron channel, $L_{\text{int}} = 60 \text{ fb}^{-1}$					
	WH	t \bar{t}	Wbb	ZW	WW	ZZ
<i>no cut</i>	21216	48.18 M	6.66 M	3.00 M	6.86 M	0.92 M
<i>L1 + HLT</i>	3420	5.41 M	762586	184417	669232	43614
<i>b-jet</i>	3322	5.39 M	615581	168358	641034	42059
<i>W mass</i>	2976	4.76 M	604759	161131	565354	40565
<i>E/P</i>	2005	2.40 M	512513	120854	363253	26721
<i>same lepton veto</i>	1953	2.20 M	509874	101046	342002	15241
<i>jet veto</i>	314	74121	193541	23545	71261	2919
<i>wrong lepton veto</i>	303	63694	193069	22769	62611	2865

Table 3.14: The number of events selected in the analysis of the WH electron channel ($L_{\text{int}} = 60 \text{ fb}^{-1}$).

constraint	muon channel, $L_{\text{int}} = 60 \text{ fb}^{-1}$					
	WH	t \bar{t}	Wbb	ZW	WW	ZZ
<i>no cut</i>	21216	48.18 M	6.66 M	3.00 M	6.86 M	0.92 M
<i>L1 + HLT</i>	4367	7.07 M	1.13 M	252660	865884	54336
<i>b-jet</i>	4256	7.06 M	900836	231080	828006	52552
<i>W mass</i>	3548	5.84 M	885003	219483	702222	50256
<i>same lepton veto</i>	3108	4.68 M	842966	164143	647456	17937
<i>jet veto</i>	511	155552	323669	39694	119849	3938
<i>wrong lepton veto</i>	505	136115	322881	38763	108839	3873

Table 3.15: The number of events selected in the analysis of the WH muon channel ($L_{\text{int}} = 60 \text{ fb}^{-1}$).

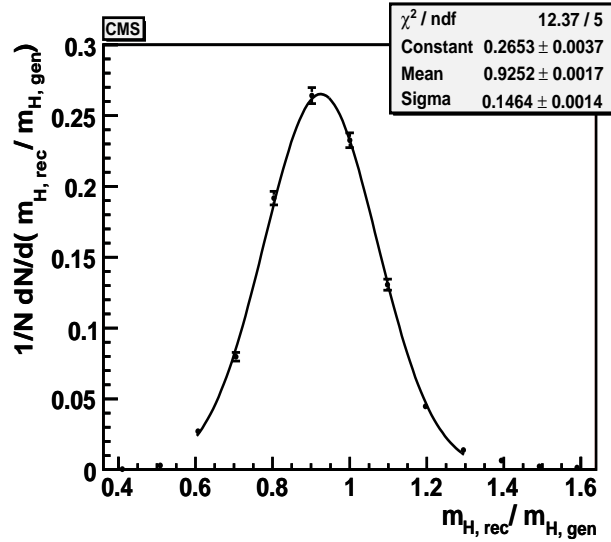


Figure 3.33: The reconstructed $H \rightarrow b\bar{b}$ mass resolution in WH events with low-luminosity pileup.

events. In Table 3.14 and 3.15 the numbers of accepted signal and background events after applying the various selection criteria are listed for an integrated luminosity of 60 fb^{-1} .

In both tables the first row (*no cut*) shows the number of events before any selection criteria is applied (see Table 3.8). About two times more signal events in the muon channel are selected compared to the electron channel. However, the selection efficiency of the listed backgrounds in the muon channel is also increased roughly by a factor of two.

3.4.7.4 Higgs mass reconstruction performance

The Higgs mass resolution, calculated from two jets matched to the true b-quarks from the Higgs decay, is shown in Figure 3.33. The reconstructed jets with $P_t > 20 \text{ GeV}$ are selected from the WH signal event sample by demanding $\Delta R(\text{rec.jet}, \text{b-quark}) < 0.2$.

The r.m.s of the mass resolution is about 15%. The peak is 7.5% too low, compatible with the reduction of about 10% in the jet response, discussed in section 3.4.5.

3.4.7.5 Results

For all events, which pass the selection criteria as described before, the mass m_H of the Higgs candidates is reconstructed by using the two reconstructed jets with the highest probability to originate from a b-hadron decay (section 3.4.7.3). For the identification the $PMRX_{\text{tag}}$ b-tag is used (section 3.4.6).

The Higgs mass m_H distributions of the selected candidates in the electron and muon channel are shown in Figures 3.34(a) and (b), respectively for an integrated luminosity $L_{\text{int}} = 60 \text{ fb}^{-1}$. The number of WH signal events is scaled by 100 and the different background events are added up in both figures.

For both channels the distributions of the WH signal events peak at about 120 GeV as expected from the observed Higgs mass resolution shown in Figure 3.33. Due to the misidentification of jets as Higgs signal decay b-jets by the strategy outlined in section 3.4.7.3 both distributions are broadened. It turns out that the m_H signal shape is different compared to the

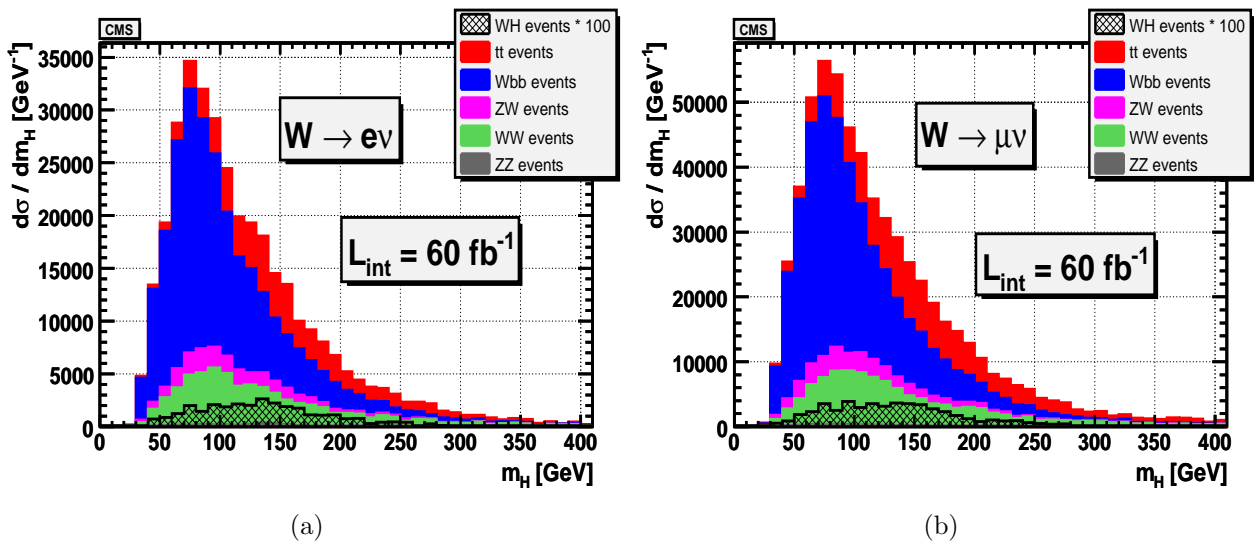


Figure 3.34: The reconstructed Higgs mass candidates for the electron (a) and muon (b) channel separately for signal events and the main backgrounds ($L_{\text{int}} = 60 \text{ fb}^{-1}$).

shown backgrounds. However the signal is almost not visible for the luminosity given due to the large number of background events as shown in Figures 3.34(a) and (b).

3.4.7.6 Comparison with the ATLAS Physics TDR result

The ATLAS collaboration has evaluated its discovery potential for the Standard Model Higgs boson in the WH, $H \rightarrow b\bar{b}$ channel too. The combined result for the electron and muon channel is published in [ATL99a, ATL99b]. An integrated luminosity of $L_{\text{int}} = 30 \text{ fb}^{-1}$ was used.

The analysis has been performed using the ATLAS fast simulation. The presented b-tagging performance and the invariant mass reconstruction of b-jet pairs were validated by means of the ATLAS full detector simulation. The analysis strategy is similar to the one used in the present CMS study. In addition after all selection criteria were applied a mass window of $\pm 2\sigma_{m_H}$ around the nominal Higgs boson mass m_H was used.

Table 3.16 shows a comparison of the predicted event numbers for a Higgs mass of $m_H = 120 \text{ GeV}$ by the ATLAS collaboration with those of the present study¹⁶. For these purposes a $\pm 2\sigma_{m_H}$ mass window is used for the present study too. Furthermore the event numbers are calculated for an integrated luminosity of $L_{\text{int}} = 30 \text{ fb}^{-1}$.

It turns out that in the production cross sections used in the ATLAS study were quite different compared to the up-to-date ones used in this study. Based on the newly available cross sections the obtained number of $t\bar{t}$ +jets and ZW+jets events would be still about three times lower for the ATLAS study. The number of selected Wbb events after applying all selection criteria is much higher in the present CMS study.

A possible explanation of this obvious different result is the usage of the ATLAS fast simulation. The description of the used data sets implies that no realistic event samples were used. In particular the degeneration of the b-tag performance and of the Higgs mass reconstruction due to the low-luminosity pileup events contribution have not been fully considered in the ATLAS study.

channel	ATLAS		present CMS study	
	cross section [pb]	# events	cross section [pb]	# events
WH ($W \rightarrow l\nu$, $H \rightarrow b\bar{b}$)	< 0.4	250	0.3536	185
$t\bar{t}$ +jets	247	3700	803.0	36364
Wbb ($W \rightarrow l\nu$)	70	2000	111.0	119668
WW+jets	—	—	114.3	39224
ZW+jets	0.86	220	49.9	14566
ZZ+jets	—	—	15.3	1660

Table 3.16: The comparison of the ATLAS WH analysis result [ATL99a, ATL99b] with the present CMS study for a simulated Higgs mass of $m_H = 120 \text{ GeV}$. The given event numbers are obtained for an integrated luminosity of $L_{\text{int}} = 30 \text{ fb}^{-1}$ after all selection criteria are applied. A mass window of $\pm 2\sigma_{m_H}$ around the nominal Higgs boson mass m_H is used.

¹⁶In reference [ATL99b] the WH production cross section for a Higgs boson mass of $m_H = 100 \text{ GeV}$ is given only. The ATLAS study did not use any data sets which would be easily comparable to the WW+jets and ZZ+jets event samples of the present CMS study.

3.4.8 Summary

The expected b-hadron identification capability of the CMS detector taken from full simulation was evaluated by means of $t\bar{t}$ events mixed with low-luminosity ($\mathcal{L} = 2 \cdot 10^{33} \text{cm}^{-2} \text{s}^{-1}$) pileup. The performance was determined by evaluating the b-efficiencies and fake rates and by investigating the discovery potential of CMS in the WH ($W \rightarrow l\nu$, $H \rightarrow b\bar{b}$) channel.

To obtain a reasonable estimation of the b-hadron identification capabilities, the track, vertex, and jet finding performance of the CMS reconstruction was determined. This led to revision of the algorithms, in particular those needed for vertex and jet reconstruction. After establishing the functionality of the revised algorithms, the 3-dimensional jet lifetime probability and a new combined b-tag for CMS was implemented. The performance of the $PMRX_{\text{tag}}$ b-tag improves significantly the b-hadron identification with respect to an only jet lifetime probability based tag as described in section 3.4.6.

For the estimation of the CMS discovery potential in the WH channel the reconstruction of electrons, muons and of the missing transverse energy was needed. The obtained reconstruction performance is in agreement with the expectations (sections 3.4.2.2, 3.4.2.3 and 3.4.4).

In the physics analysis of the WH channel (section 3.4.7) it turned out that especially the $t\bar{t}$ +jets and the $Wb\bar{b}$ backgrounds are difficult to handle. The $t\bar{t}$ +jets events have a huge production cross section and the $Wb\bar{b}$ events have a similar decay topology in comparison to the WH signal events. It followed that the WH ($W \rightarrow l\nu$, $H \rightarrow b\bar{b}$) associated Higgs production will not be detectable in the first years of LHC.

Chapter 4

Conclusions

New methods for b-hadron jet identification were implemented for the ALEPH and CMS experiments. The strategies developed for events from e^+e^- collisions recorded by the ALEPH experiment were applied to simulated pp events in the CMS detector resulting in a reasonable b-tag performance.

The pre-selection and analysis of the ALEPH events was performed by means of the Condor High Throughput Computing batch system on a local computer farm. The ALEPH LEP1 data and the corresponding Monte Carlo events were directly accessible on this farm.

The developed b-tag for ALEPH uses a combination of observables sensitive to the differences of b-hadron decays from other light quark hadron decays. The b-tag is based on the the ALEPH reconstruction algorithms for tracks and jets optimized for b-hadron identification. It is provided by the QIPBTAG tool of the ALEPH reconstruction tools library ALPHA. QIPBTAG also provides the calibrated resolution function obtained from the significance of the signed 3-dimensional impact parameter and used this to compute the track probabilities and the jet lifetime probability.

The new combined b-tag for ALEPH was put in perspective by comparison with the latest R_b measurements at the Z^0 resonance. The implemented b-tag is able to operate in the same b-efficiency and b-purity region as the latest DELPHI measurement. For the same b-purity used in the measurement of R_b with the ALEPH lifetime mass tag, the new tag b-efficiency increases by roughly about 10%.

To be able to use a similar combined b-tag for CMS the prerequisites were created. The implementation of a 3-dimensional jet lifetime probability b-tag for CMS events required the optimization of the CMS track, vertex and jet reconstruction. The jet reconstruction was performed by the inclusive k_\perp jet algorithm using a newly developed jet energy flow input (TrackTowers). This jet reconstruction setup is able to provide reasonably determined final state jets in the dense event topologies of the LHC. Because of the occurrence of additional primary vertices in LHC pp collisions, *e.g.* due to pileup interactions, strategies were developed to reconstruct the primary vertices with high accuracy and to reliably identify the signal vertex of the main interaction. The selection of tracks from the primary b-hadron decay vertex was done by implementing constraints to suppress other tracks, *e.g.* from other jets, other primary or tertiary vertices. Having optimized the reconstruction and selection of tracks, vertices and jets as described above, the signed 3-dimensional impact parameter was used to compute the resolution function and the jet lifetime probability. A good performance of the jet lifetime probability b-tag was obtained for typical LHC $t\bar{t}$ events mixed with low-luminosity pileup ($\mathcal{L} = 2 \cdot 10^{33} \text{cm}^{-2} \text{s}^{-1}$).

For the implementation of the combined b-tag in CMS, a search for a secondary vertex was performed inside the reconstructed jets. As for ALEPH, the association of tracks to the secondary vertex was done by comparing the track distance from the secondary to that from

the signal vertex. Because of the huge amount of gluon jets in LHC events the combination of the used observables was revised and a separate weight for the gluon-jets was utilized in the CMS case. The b-hadron jet identification was significantly improved by the combined b-tag with respect to the jet lifetime probability b-tag alone.

The discovery potential of CMS for the WH, $H \rightarrow b\bar{b}$ channel in low-luminosity pileup conditions was investigated by employing the improved and newly developed CMS reconstruction algorithms. The discovery of the Standard Model Higgs boson in this channel is almost impossible in the first years of LHC even though the W mass constraint is utilized, because of the large cross section of the $t\bar{t}$ production and the very similar decay topologies of Wbb events in comparison to the WH production.

The investigation and development of the presented b-hadron jet identification for ALEPH and CMS was done by means of a newly implemented, purpose-built, stand-alone package based on the ROOT object orientated data analysis framework. To decouple the package from the ALEPH reconstruction tools library (ALPHA) and the CMS reconstruction framework (ORCA), the reconstructed events were preselected and stored in a common data format.

The full simulation, full reconstruction and pre-selection of the CMS data was done by means of the Worldwide LHC Computing Grid. The data from the LHC experiments (roughly 15 Peta bytes per year) will be distributed around the globe so that thousands of scientists around the world will be able to access and analyze it.

For the present study, CMS Monte Carlo samples of 2M events were reconstructed dozens of times. The CPU time needed per event is several minutes. Thus, in addition to developing new physics tools and algorithms, the methods used reflect the global way in which analyses of CMS data will be performed.

Appendix A

Analytical solution of the double tag method equations

The two coupled equations (1.21) and (1.22) can be decoupled and solved analytically for R_b and eff_b . By means of the mathematical program package MAPLE [Map], which is able to perform symbolic calculations of mathematical problems, the equations were solved.

The MAPLE input commands to perform this task are listed below.

```
[> restart;
[> coupled_Rb := Rb = (F1-Rc*(Effc-Effuds)-Effuds)^2
(F2-Rc*(Effc-Effuds)^2+Effuds^2-2*F1*Effuds-lambdab*Rb*(Effb-Effb^2));
[> coupled_Effb := Effb = (F2-Rc*Effc*(Effc-Effuds)-F1*Effuds-lambdab*Rb*(Effb-Effb^2))
/(F1-Rc*(Effc-Effuds)-Effuds);
[> result:=allvalues(solve(coupled_Rb,coupled_Effb,Rb,Effb));
[> decoupled_Rb:=op(2,result[1][2]);
[> decoupled_Effb:=op(2,result[1][1]);
```

The ">" sign denotes a MAPLE session input line.

Appendix B

The object-oriented b-tag framework

The development and evaluation of b-tag observables in this study was done by means of a newly developed object-orientated C++ framework embedded in the ROOT object-oriented analysis framework [BR97].

The reconstructed and preselected data of ALEPH and CMS were stored in a common data format and then processed by means of an analysis class. The two analysis classes `TALEPHTreeReader` and `TCMSTreeReader` which are specialized in histogramming the ALEPH and CMS data, respectively, are based on the same skeleton analysis class. Further member classes of the b-tag framework are listed in Table B.1.

The calculation and histogramming of R_b and eff_b calculated by means of the double tag method (see section 1.2.1) is implemented in the specialized class `TALEPHHTagPerformance`, which is based on the class `THJetTagPerformance`.

For the WH analysis the tool class `TSignificance` was developed. Several methods to calculate the signal significance were implemented, *e.g.* the significance S_{CP} , which is the probability obtained from a Poisson distribution of events converted to equivalent number of sigmas of a Gaussian distribution [BEKN05]. S_{CP} is commonly used in CMS.

class name	description
<code>THFlavor</code>	histogram class to plot observables separately for different jet flavors
<code>THFlavor2D</code>	histogram class utilizing 2-dimensional histograms to visualize two observables separately for different jet flavors
<code>THWeight</code>	histogram class to determine flavor dependent weights which are used for the combined b-tag
<code>THJetTagPerformance</code>	histogram class to determine the b-tag performance, <i>e.g.</i> efficiencies, purities and fake rates for different jet types
<code>TJetLifetime</code>	class to calculate the jet lifetime probability
<code>T4tag</code>	class to calculate combined b-tag observables by means of flavor dependent weights
<code>TFit</code>	tool class to fit the resolution function used in the jet lifetime calculation and the weights needed for the combined b-tag
<code>TALEPHTreeReader</code>	analysis class specialized for ALEPH
<code>TCMSTreeReader</code>	analysis class specialized for CMS

Table B.1: The member classes of the object-oriented b-tag framework.

Appendix C

The UE curve for the CMS low-luminosity phase

The 1-parameter UE curve $E_{\text{UE,curve}}(\eta)$ to reject calorimeter towers, which contain the underlying event energy only (section 3.4.5), is presented below. The continuous formula is obtained from a spline fit and it is divided in three parts:

1. $|\eta| < 1.37 : E_{\text{UE,curve}}(\eta) = 1.5 \text{ GeV}$
2. $1.37 < |\eta| < 3.70 :$
$$E_{\text{UE,curve}}(\eta) = 1.26459 - 0.297559 |\eta| - 0.0486882 |\eta|^2 + 5.67154 \cdot 10^{-2} |\eta|^3 + 0.182822 |\eta|^4 - 3.94960 \cdot 10^{-3} |\eta|^5 - 6.24019 \cdot 10^{-3} |\eta|^6 + 8.60195 \cdot 10^{-5} |\eta|^7 \text{ [GeV]}$$
3. $|\eta| > 3.70 : E_{\text{UE,curve}}(\eta) = 1.69 + 4.6 |\eta| \text{ [GeV]}$

The presented UE curve is computed for CMS events, which are simulated for the LHC low-luminosity phase (section 3.4.1). However the data driven UE method can be eventually applied to real CMS events as well [CMS05k].

Appendix D

The 2-dimensional detector map jet correction function

A reconstructed jet is corrected by a two dimensional function D_{map} . The function depends on the transverse energy E_t and the $|\eta|$ of the reconstructed jet and is defined as follows:

$$D_{\text{map}}(E_t, |\eta|) := a + \frac{b}{E_t} + c \cdot \exp\left(\frac{d}{E_t} + e\right) + f \cdot \exp\left(\frac{g}{E_t} + h\right) + i \cdot E_t + j|\eta| + k|\eta|^2 + l|\eta|^3 + m|\eta|^4 + n|\eta|^5 + o|\eta|^6 \quad (\text{D.1})$$

The values of the parameters a, \dots, o are listed in Table D.1. The resulting corrected jet 4-vector (E', p'_x, p'_y, p'_z) is obtained by scaling the uncorrected jet 4-vector (E, p_x, p_y, p_z) with:

$$E'_t = E_t \cdot (1 + D_{\text{map}}(E_t, |\eta|)). \quad (\text{D.2})$$

The final jet 4-vector is computed by:

$$E' = E \cdot \frac{E'_t}{E_t}, \quad p'_i = p_i \cdot \frac{E'_t}{E_t}, \quad i = x, y, z. \quad (\text{D.3})$$

parameter	value
a	$4.92906 \cdot 10^{-02}$
b	13.00409
c	$9.58972 \cdot 10^{-02}$
d	$1.90658 \cdot 10^{-07}$
e	$-2.55468 \cdot 10^{-06}$
f	$5.03254 \cdot 10^{-02}$
g	$6.71619 \cdot 10^{-08}$
h	$-5.56599 \cdot 10^{-06}$
i	$-1.787961 \cdot 10^{-04}$
j	$3.70229 \cdot 10^{-01}$
k	$-6.05095 \cdot 10^{-01}$
l	$4.179091 \cdot 10^{-01}$
m	$-1.4087726 \cdot 10^{-01}$
n	$2.14936 \cdot 10^{-02}$
o	$-1.158319 \cdot 10^{-03}$

Table D.1: The parameter values used in the jet detector correction function D_{map} .

Bibliography

- [A⁺83] G. Arnison et al., *Hadronic jet production at the CERN proton-antiproton collider*, Physics Letters B **132**, 214–222 (November 1983), [Phys.Lett.B132:214,1983](#).
- [Abd] S. Abdouline, private communication.
- [AFST05] W. Adam, R. Frühwirth, A. Strandlie and T. Todorov, *Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC*, J.Phys.G:Nucl.Part.Phys. **G31**(9), N9–N20 (2005), [J.Phys.G:Nucl.Part.Phys.31,N9-N20](#).
- [AH89] I. Aitchison and A. Hey, *Gauge Theories in Particle Physics*, Institute of Physics Publishing, 2nd edition, 1989.
- [ALE] ALEPH home page, <http://aleph.web.cern.ch/aleph/>.
- [ALEPH90] D. Decamp et al. (ALEPH Collaboration), *ALEPH: A DETECTOR FOR ELECTRON - POSITRON ANNIHILATIONS AT LEP*, Nucl. Instrum. Meth. **A294**, 121–178 (1990), [Nucl.Instrum.Meth.A294:121-178,1990](#), [Erratum-ibid.A303:393,1991](#).
- [ALEPH92] D. Brown and M. Frank (ALEPH Collaboration), *Tagging b hadrons using track impact parameters*, (1992), ALEPH 92-135.
- [ALEPH93a] D. Buskulic et al. (ALEPH Collaboration), *A Precise measurement of Gamma ($Z \rightarrow b$ anti- b) / Gamma ($Z \rightarrow$ hadrons)*, Phys. Lett. **B313**, 535–548 (1993), [Phys.Lett.B313:535-548,1993](#).
- [ALEPH93b] D. Buskulic et al. (ALEPH Collaboration), *Update of electroweak parameters from Z decays*, Z.Phys. **C60**, 71–81 (1993), [Z.Phys.C60:71-82,1993](#).
- [ALEPH95] D. Buskulic et al. (ALEPH Collaboration), *Performance of the ALEPH detector at LEP*, Nucl. Instrum. Meth. **A360**, 481–506 (1995), [Nucl.Instrum.Meth.A360:481-506,1995](#).
- [ALEPH97] R. Barate et al. (ALEPH Collaboration), *A measurement of $R(b)$ using a lifetime-mass tag*, Phys. Lett. **B401**, 150–162 (1997), [Phys.Lett.B401:150-162,1997](#).
- [ALEPH00] J. Boucrot et al. (ALEPH Collaboration), *ALPHA: ALEPH Physics Analysis Package (Version ≥ 125)*, (2000), ALEPH 99-087 SOFTWR 99-001.

- [ALEPH01] R. Barate et al. (ALEPH Collaboration), *Investigation of inclusive CP asymmetries in B^0 decays*, Eur. Phys. J. **C20**, 431–443 (2001), [Eur.Phys.J.C20:431-443,2001](#).
- [ALI] ALICE home page, <http://aliceinfo.cern.ch/>.
- [Alt89] Altarelli, G., (ed.), *Z Physics at LEP 1*, (1989), Yellow Report CERN 89-05, Volume 1.
- [ATL] ATLAS home page, <http://atlas.web.cern.ch/>.
- [ATL99a] ATLAS Collaboration, *ATLAS DETECTOR AND PHYSICS PERFORMANCE*, volume 1, [CERN/LHCC 99-14](#), 25 May 1999.
- [ATL99b] ATLAS Collaboration, *ATLAS DETECTOR AND PHYSICS PERFORMANCE*, volume 2, [CERN/LHCC 99-15](#), 25 May 1999.
- [B⁺77] C. Bromberg et al., *Observation of the Production of Jets of Particles at High Transverse Momentum and Comparison with Inclusive Single-Particle Reactions*, Physical Review Letters **38**, 1447–1450 (June 1977), [Phys.Rev.Lett.38:1447,1977](#).
- [B⁺79] C. Bromberg et al., *EXPERIMENTAL TESTS OF QUANTUM CHROMODYNAMICS IN HIGH P(T) JET PRODUCTION IN 200-GeV/c HADRON - PROTON COLLISIONS*, Phys. Rev. Lett. **43**, 565–568 (1979), [Phys.Rev.Lett.43:565-568,1979](#).
- [B⁺97] D. Bourilkov et al., *Beam spot position measurement at the LEP collider*, Nucl. Instrum. Meth. **A394**, 103–114 (1997), [Nucl.Instrum.Meth.A394:103-114,1997](#).
- [B⁺00] G. C. Blazey et al., *Run II jet physics*, (2000), [hep-ex/0005012](#).
- [B⁺04] O. Brein et al., *Precision calculations for associated W H and Z H production at hadron colliders*, (2004), [hep-ph/0402003](#).
- [BCCW03] J. M. Butterworth, J. P. Couchman, B. E. Cox and B. M. Waugh, *Kt-Jet: A C++ implementation of the K(T) clustering algorithm*, Comput. Phys. Commun. **153**, 85–96 (2003), [hep-ph/0210022](#).
- [BCMN98] R. Bonciani, S. Catani, M. L. Mangano and P. Nason, *NLL resummation of the heavy-quark hadroproduction cross-section*, Nucl. Phys. **B529**, 424–450 (1998), [hep-ph/9801375](#).
- [BCVZ89] R. Brun, O. Couet, C. E. Vandoni and P. Zanarini, *PAW: A GENERAL PURPOSE PORTABLE SOFTWARE TOOL FOR DATA ANALYSIS AND PRESENTATION*, Comput. Phys. Commun. **57**, 432–437 (1989), [Comput.Phys.Comm.57:432-437,1989](#).
- [BDH04] O. Brein, A. Djouadi and R. Harlander, *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, Phys. Lett. **B579**, 149–156 (2004), [hep-ph/0307206](#).

- [BEKN05] S. Bitukov, S. Erofeeva, N. Krasnikov and A. Nikitenko, Program for evaluation of the significance, confidence intervals and limits by direct calculation of probabilities, in *PHYSTAT 05 Proceedings Statistical Problems in Particle Physics, Astrophysics and Cosmology*, edited by L. Lyons and M. K. Unel, Imperial College Press, 2005.
- [Ber02] C. Berger, *Elementarteilchenphysik*, Springer Verlag Berlin Heidelberg New York, 2002.
- [Bit02] S. I. Bitukov, *Signal significance in the presence of systematic and statistical uncertainties*, JHEP **09**, 060 (2002), [hep-ph/0207130](#).
- [BK98] S. I. Bitukov and N. V. Krasnikov, *New Physics Discovery Potential in Future Experiments*, Mod.Phys.Let. **A13**(40), 3235–3249 (1998), [physics/9811025](#).
- [BK00] S. I. Bitukov and N. V. Krasnikov, *On the observability of a signal above background*, Nucl.Inst.Meth. **A452**, 518–524 (2000), [Nucl.Instrum.Meth.A452:518-524,2000](#).
- [BK03] S. I. Bitukov and N. V. Krasnikov, *Signal significance in the presence of systematic and statistical uncertainties*, Nucl.Instrum.Meth. **A502**, 795–798 (2003), [Nucl.Instrum.Meth.A502:795-798,2003](#).
- [BK04] S. I. Bitukov and N. V. Krasnikov, *Distinguishability of hypotheses*, Nucl.Instrum.Meth. **A534**, 152–155 (2004), [Nucl.Instrum.Meth.A534:152-155,2004](#).
- [BMSZ04] S. I. Bitukov, V. A. Medvedev, V. V. Smirnova and Y. V. Zernii, *Experimental test of the probability density function of true value of Poisson distribution parameter by single observation of number of events*, Nucl.Instrum.Meth. **A534**, 228–231 (2004), [Nucl.Instrum.Meth.A534:228-231,2004](#).
- [BR97] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389**, 81–86 (1997), [Nucl.Instrum.Meth.A389:81-86,1997](#), [ROOT website: http://root.cern.ch/](#).
- [C⁺98] D. Creanza et al., *Construction and performance of the new ALEPH vertex detector*, Nucl. Phys. Proc. Suppl. **61B**, 201–206 (1998), [Nucl.Phys.Proc.Suppl.61B:201-206,1998](#).
- [Cab63] N. Cabibbo, *UNITARY SYMMETRY AND LEPTONIC DECAYS*, Phys. Rev. Lett. **10**, 531–532 (1963), [Phys.Rev.Lett.10:531-532,1963](#).
- [Cam] J. Campbell, private communication.
- [CDK03] M. L. Ciccolini, S. Dittmaier and M. Kramer, *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*, Phys. Rev. **D68**, 073003 (2003), [hep-ph/0306234](#).

- [CDSW93] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, *Longitudinally invariant $K(t)$ clustering algorithms for hadron hadron collisions*, Nucl. Phys. **B406**, 187–224 (1993), [Nucl.Phys.B406:187-224,1993](#).
- [CDW92] S. Catani, Y. L. Dokshitzer and B. R. Webber, *The K-perpendicular clustering algorithm for jets in deep inelastic scattering and hadron collisions*, Phys. Lett. **B285**, 291–299 (1992), [Phys.Lett.B285:291-299,1992](#).
- [CE] J. Campbell and R. K. Ellis, *MC²FM - Monte Carlo for FeMtobarn processes*, <http://mcfm.fnal.gov/>.
- [CER03] J. Campbell, R. K. Ellis and D. L. Rainwater, *Next-to-leading order QCD predictions for $W + 2\text{jet}$ and $Z + 2\text{jet}$ production at the CERN LHC*, Phys. Rev. **D68**, 094021 (2003), [hep-ph/0308195](#).
- [CERN97] J. Steinberger (CERN Collaboration), *A Brief History Of the Running of R_b* , (2 May 1997), [CERN-OPEN-97-013](#).
- [CERN04] K. Hübner (CERN Collaboration), *Designing and Building LEP*, (2004), [CERN-AB-2004-099-ADM](#).
- [Cli96] J. M. Cline, *$R(b)$ and new physics: A comprehensive analysis*, (1996), [hep-ph/9609327](#).
- [CMS94] CMS Collaboration, *The Compact Muon Solenoid Technical Proposal*, CERN/LHCC 94-38, 15 December 1994.
- [CMS96] L. Taylor (CMS Collaboration), *CMKIN – The CMS Kinematics Interface Package*, (1996), CMS TN-1996/099.
- [CMS97a] CMS Collaboration, *The Electromagnetic Calorimeter Technical Design Report*, CMS TDR 4, CERN/LHCC 97-33, 15 December 1997, CMS TDR 4 Addendum, CERN/LHCC 2002-027.
- [CMS97b] CMS Collaboration, *The Hadron Calorimeter Technical Design Report*, CMS TDR 2, CERN/LHCC 97-31, 20 June 1997.
- [CMS97c] CMS Collaboration, *The Magnet Project Technical Design Report*, CMS TDR 1, CERN/LHCC 97-10, 2 May 1997.
- [CMS97d] CMS Collaboration, *The Muon Project Technical Design Report*, CMS TDR 3, CERN/LHCC 97-32, 15 December 1997.
- [CMS98] CMS Collaboration, *Tracker Technical Design Report*, CMS TDR 5, CERN/LHCC 98-6, 15 April 1998, CMS TDR 5 Addendum 1, CERN/LHCC 2000-016.
- [CMS99a] G. Bagliesi et al. (CMS Collaboration), *Object Oriented Reconstruction for CMS Analysis*, (1999), [CMS IN-1999/001](#).
- [CMS99b] S. I. Bityukov (CMS Collaboration), *Observability and Probability of Discovery in Future Experiments*, (1999), [CMS IN-1999/027](#).

- [CMS99c] M. Schröder (CMS Collaboration), *CMS Detector Simulation Project OSCAR*, (1999), [CMS IN-1999/036](#).
- [CMS00a] S. I. Bityukov and N. V. Krasnikov (CMS Collaboration), *On Observability of Signal Over Background*, Conference Report (2000), [CMS CR-2000/004](#).
- [CMS00b] CMS Collaboration, *The Level-1 Trigger Technical Design Report*, CMS TDR 6.1, [CERN/LHCC 2000-38](#), 15 December 2000.
- [CMS00c] M. Huhkinen (CMS Collaboration), *Impact of beam-pipe geometry on radiation background in CMS*, (21 November 2000), [CMS NOTE 2000/069](#).
- [CMS01a] E. Meschi et al. (CMS Collaboration), *Electron Reconstruction in the CMS Electromagnetic Calorimeter*, (22 June 2001), [CMS NOTE 2001/034](#).
- [CMS01b] A. Nikitenko et al. (CMS Collaboration), *Missing Transverse Energy Measurement with Jet Energy Corrections*, (2001), [CMS NOTE-2001/040](#).
- [CMS02a] S. I. Bityukov (CMS Collaboration), *Uncertainties And Discovery Potential In Planned Experiments*, Conference Report (2002), [CMS CR-2002/005](#).
- [CMS02b] CMS Collaboration, *Data Acquisition and High-Level Trigger Technical Design Report*, CMS TDR 6.2, [CERN/LHCC 2002-26](#), 15 December 2002.
- [CMS03a] S. Abdullin et al. (CMS Collaboration), *Summary of the CMS Potential for the Higgs Boson Discovery*, (2003), [CMS NOTE-2003/033](#).
- [CMS03b] S. Cucciarelli et al. (CMS Collaboration), *Track-Parameter Evaluation and Primary-Vertex Finding with the Pixel Detector*, (4 September 2003), [CMS NOTE-2003/026](#).
- [CMS03c] D. Kotlinski (CMS Collaboration), *Pixel Detector Data Rates And Inefficiencies*, (2003), [CMS IN 2003/004](#).
- [CMS03d] T. Speer et al. (CMS Collaboration), *Vertex Fitting with the Kalman Filter Formalism in the ORCA Reconstruction Program*, (20 February 2003), [CMS IN-2003/008](#).
- [CMS04a] S. Banerjee and S. Banerjee (CMS Collaboration), *Validation of Geant4 Using 1986 CMS HCAL Test Beam Data*, (2004), [CMS NOTE 2004/032](#).
- [CMS04b] S. Cucciarelli and D. Kotlinski (CMS Collaboration), *Pixel Hit Reconstruction*, (1 May 2004), [CMS IN-2004/014](#).

- [CMS04c] V. D. Elvira (CMS Collaboration), *Measurement of the Pion Energy Response and Resolution in the Hadronic Barrel Calorimeter using CMS HCAL Test Beam 2002 Data*, (2004), [CMS NOTE 2004/020](#).
- [CMS04d] V. Karimaki et al. (CMS Collaboration), *User's Guide for CMS Physics Generators Interface CMKIN Version 4*, (2004), [CMS IN-2004/052](#).
- [CMS05a] W. Adam et al. (CMS Collaboration), *Track reconstruction in the CMS tracker*, (25 November 2005), [CMS AN 2005/063](#).
- [CMS05b] S. Banerjee (CMS Collaboration), *Simulation geometry for the CMS HCAL*, (2005), [CMS NOTE 2005/016](#).
- [CMS05c] P. Bartalini, R. Chierici and A. De Roeck (CMS Collaboration), *Guidelines for the Estimation of Theoretical Uncertainties at the LHC*, (2005), [CMS NOTE-2005/013](#).
- [CMS05d] I. Belotelov et al. (CMS Collaboration), *Simulation of Misalignment Scenarios for CMS Tracking Devices*, (2005), [CMS NOTE 2005/036](#).
- [CMS05e] I. Belotelov and N. Neumeister (CMS Collaboration), *Performance of the CMS muon offline reconstruction software*, (11 June 2005), [CMS AN 2005/010](#).
- [CMS05f] CMS Collaboration, *The Computing Project Technical Design Report*, [CERN/LHCC 2005-23](#), 20 June 2005.
- [CMS05g] P. Haifeng et al. (CMS Collaboration), *Missing Transverse Energy in QCD Events at Low Luminosity*, (2005), [CMS AN 2005/007](#).
- [CMS05h] HCAL (CMS Collaboration), *Design, Performance, and Calibration of CMS Hadron-Barrel Calorimeter Wedges*, submitted to Nucl. Instrum. and Methods (2005).
- [CMS05i] HCAL (CMS Collaboration), *Design, Performance, and Calibration of CMS Hadron-Endcap Calorimeter Wedges*, submitted to Nucl. Instrum. and Methods (2005).
- [CMS05j] HCAL (CMS Collaboration), *Design, Performance, and Calibration of CMS Hadron-Forward Calorimeter Wedges*, submitted to Nucl. Instrum. Methods (2005).
- [CMS05k] A. Heister et al. (CMS Collaboration), *Jet Reconstruction and Performance in the CMS Detector*, (2005), [CMS AN 2005/005](#), [Addendum](#), [CMS AN 2005/053](#).
- [CMS05l] G. Q. V. Bartsch (CMS Collaboration), *Expected Signal Observability at Future Experiments*, (2005), [CMS NOTE-2005/004](#).
- [CMS05m] P. Vanlaer (CMS Collaboration), *Track and Vertex Reconstruction in CMS for Key Physics Processes*, Conference Report (17 October 2005), [CMS CR 2005/025](#).

- [CMS06a] R. Brunelière and A. Zabi (CMS Collaboration), *Reconstruction of the signal amplitude of the CMS electromagnetic calorimeter*, (2006), [CMS NOTE 2006/037](#).
- [CMS06b] CMS Collaboration, *Physics Technical Design Report Volume I: Detector Performance and Software*, CMS TDR 8.1, [CERN/LHCC 2006-001](#), 2 February 2006.
- [CMS06c] CMS Collaboration, *Physics Technical Design Report Volume II: Physics Performance*, CMS TDR 8.2, [CERN/LHCC 2006-021](#), 26 June 2006.
- [CMS06d] J. Rohlf and C. Tully (CMS Collaboration), *Recommendations for Jet and Missing Transverse Energy Reconstruction Settings and Systematics Treatment*, (2006), [CMS IN-2006/025](#).
- [CompHEP04] E. Boos et al. (CompHEP Collaboration), *CompHEP 4.4: Automatic computations from Lagrangians to events*, Nucl. Instrum. Meth. **A534**, 250–259 (2004), [hep-ph/0403113](#).
- [Con] Condor High Throughput Computing home page, <http://www.cs.wisc.edu/condor/>.
- [Cow98] G. Cowen, *Statistical Data Analysis*, Oxford University Press, 1998.
- [DELPHI99] P. Abreu et al. (DELPHI Collaboration), *A precise measurement of the partial decay width ratio $R(b)0 = \Gamma(b \text{ anti-}b)/\Gamma(\text{had})$* , Eur. Phys. J. **C10**, 415–442 (1999), [EPHJA,C10,415](#).
- [DKS98] A. Djouadi, J. Kalinowski and M. Spira, *HDECAY: A program for Higgs boson decays in the standard model and its supersymmetric extension*, Comput. Phys. Commun. **108**, 56–74 (1998), [hep-ph/9704448](#).
- [dTY05] J. F. de Troconiz and F. J. Yndurain, *The hadronic contributions to the anomalous magnetic moment of the muon*, Phys. Rev. **D71**, 073008 (2005), [hep-ph/0402285](#).
- [ES93] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. **D48**, 3160–3166 (1993), [hep-ph/9305266](#).
- [ESW96] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and Collider Physics*, Cambridge University Press, 1996.
- [EWWG05] The ALEPH, DELPHI, L3, OPAL, SLD Collaborations, the LEP Electroweak Working Group, the SLD Electroweak and Heavy Flavour Groups (EWWG Collaboration), *Precision Electroweak Measurements on the Z Resonance*, page 302 (2005), [hep-ex/0509008](#).
- [EWWG06] The LEP Collaborations ALEPH, DELPHI, L3, OPAL, and the LEP Electroweak Working Group (EWWG Collaboration), *A combination of preliminary electroweak measurements and constraints on the standard model*, (2006), [hep-ex/0511027](#), Updated for 2006 summer conferences, 11-September-2006, <http://www.cern.ch/LEPEWWG>.

- [Fra94] M. Frank, *Eine Präzisionsmessung des Verzweigungsverhältnisses $\Gamma_{b\bar{b}}/\Gamma_{Had}$ in der e^+e^- -Annihilation auf der Z^0 -Resonanz*, PhD thesis, Max-Planck-Institut für Physik, Werner-Heisenberg-Institut, München, 1994.
- [Frü87] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl.Instrum.Meth. **A262**, 444–450 (1987), [HEPHY-PUB-87-503](#).
- [GEANT403] S. Agostinelli et al. (GEANT4 Collaboration), *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. **A506**, 250–303 (2003), [Nucl.Instrum.Meth.A506:250-303,2003](#).
- [GIM70] S. L. Glashow, J. Iliopoulos and L. Maiani, *WEAK INTERACTIONS WITH LEPTON - HADRON SYMMETRY*, Phys. Rev. **D2**, 1285–1292 (1970), [Phys.Rev.D2:1285-1292,1970](#).
- [Gla61] S. L. Glashow, *PARTIAL SYMMETRIES OF WEAK INTERACTIONS*, Nucl. Phys. **22**, 579–588 (1961), [Nucl.Phys.22:579-588,1961](#).
- [Gol61] J. Goldstone, *FIELD THEORIES WITH 'SUPERCONDUCTOR' SOLUTIONS*, Nuovo Cim. **19**, 154–164 (1961), [Nuovo Cim.19:154-164,1961](#).
- [GSW62] J. Goldstone, A. Salam and S. Weinberg, *BROKEN SYMMETRIES*, Phys. Rev. **127**, 965–970 (1962), [Phys.Rev.127:965-970,1962](#).
- [H⁺90] J. E. Huth et al., *Toward a standardization of jet definitions*, Conference Report (1990), [FERMILAB-CONF-90-249-E](#), Presented at Summer Study on High Energy Physics, Reaearch Directions for the Decade, Snowmass, CO, Jun 25 - Jul 13, 1990.
- [HH64] J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods*, Methuen and Co., Ltd., London, 1964.
- [Hig64] P. W. Higgs, *BROKEN SYMMETRIES AND THE MASSES OF GAUGE BOSONS*, Phys. Rev. Lett. **13**, 508–509 (1964), [Phys.Rev.Lett.13:508-509,1964](#).
- [Hil96] M. D. Hildreth, *The current status of the $R(b)$ and $R(c)$ puzzle*, (1996), [OPAL CR242](#), Prepared for 31st Rencontres de Moriond: Electroweak Interactions and Unified Theories, Les Arcs, France, 16-23 Mar 1996.
- [HM84] F. Halzen and A. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, John Wiley & Sons, 1984.
- [JADE86] W. Bartel et al. (JADE Collaboration), *EXPERIMENTAL STUDIES ON MULTI - JET PRODUCTION IN e^+e^- ANNIHILATION AT PETRA ENERGIES*, Z. Phys. **C33**, 23 (1986), [Z.Phys.C33:23,1986](#).
- [Kil06] B. Kilminster, *Search for low mass Higgs at the Tevatron*, (2006), [hep-ex/0611001](#).

- [KM73] M. Kobayashi and T. Maskawa, *CP violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49**, 652–657 (1973), [Prog.Theor.Phys.49:652-657,1973](#).
- [Kno05] J. Knobloch, editor, *LHC Computing Grid Technical Design Report*, LCG-TDR-001, [CERN-LHCC 2005-24](#), 20 June 2005.
- [Lam05] M. Lamont, *Estimates of Annual Proton Doses in the LHC*, CERN-LHC-Project-Note **375** (2005), [CERN-LHC-Project-Note-375](#).
- [Lam06] S. Lammel, *Search for Higgs and new phenomena at colliders*, (2006), [hep-ex/0601037](#).
- [LHCa] LHC home page, <http://lhc.web.cern.ch/lhc/>.
- [LHCb] LHCb home page, <http://lhcb-public.web.cern.ch/>.
- [LHC95] LHC study group, *The Large Hadron Collider: Conceptual Design*, CERN/AC/95-05(LHC), 20 October 1995.
- [LHWG03] The LEP Collaborations ALEPH, DELPHI, L3, OPAL, and the LEP Working Group for Higgs boson searches (LHWG Collaboration), *Search for the standard model Higgs boson at LEP*, Phys. Lett. **B565**, 61–75 (2003), [hep-ex/0306033](#).
- [Map] Maplesoft home page, <http://maplesoft.com/>.
- [Mark II 91] R. Jacobsen et al. (Mark II Collaboration), *Measurement of the b anti- b fraction in hadronic Z^0 decays with precision vertex detectors*, Phys. Rev. Lett. **67**, 3347–3350 (1991), [Phys.Rev.Lett.67:3347-3350,1991](#).
- [MS03] B. Martin and G. Shaw, *Particle Physics*, John Wiley & Sons, 2nd edition, 2003.
- [Nik] A. Nikitenko, private communication.
- [PDG04] S. Eidelman et al. (PDG Collaboration), *Review of Particle Physics*, Physics Letters B **B592**, 1+ (2004), 2005 and partial update for the 2006 edition available on the [PDG WWW pages](#).
- [PG97] R. B. Palmer and J. C. Gallardo, *Muon muon and other high-energy colliders*, NATO Adv. Study Inst. Ser. B Phys. **365**, 183–272 (1997), [physics/9702017](#).
- [PS95] M. Peskin and D. Schroeder, *An Introduction to Quantum Field Theory*, Westview Press, 1995.
- [PSSZ83] C. Peterson, D. Schlatter, I. Schmitt and P. M. Zerwas, *SCALING VIOLATIONS IN INCLUSIVE e^+e^- ANNIHILATION SPECTRA*, Phys. Rev. **D27**, 105 (1983), [Phys.Rev.D27:105,1983](#).
- [Qui83] C. Quigg, *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions*, Benjamin/Cummings Publishing Company, 1983.

- [Roh94] J. Rohlf, *Modern Physics from α to Z^0* , John Wiley & Sons, 1994.
- [Ros67] S. Rosen, editor, *Programming Systems and Languages*, McGraw Hill, 1967.
- [Sal68] A. Salam, Weak and Electromagnetic Interactions., in *Elementary Particle Theory*, edited by N. Svartholm, page 367, Almquist and Wiksell, 1968.
- [Sch] S. Schael, private communication.
- [Sch99] S. Schael, *B physics at the Z resonance*, Phys. Rept. **313**, 293–366 (1999), [Phys.Rept.313:293-366,1999](#).
- [Sey98] M. H. Seymour, *Jet shapes in hadron collisions: Higher orders, resummation and hadronization*, Nucl. Phys. **B513**, 269–300 (1998), [hep-ph/9707338](#).
- [Sey00] M. H. Seymour, *Jets in hadron collisions*, (2000), [hep-ph/0007051](#).
- [Sjo94] T. Sjostrand, *High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4*, Comput. Phys. Commun. **82**, 74–90 (1994), [Comput.Phys.Commun.82:74-90,1994](#).
- [SLD] SLD home page, <http://www-sld.slac.stanford.edu/>.
- [SLD96] J. A. Snyder (SLD Collaboration), *R_b and R_c ?*, (1996), [CERN-PPE-96-091C](#), Contributed to Moriond Workshop on R(b) and R(c), Les Arcs, France, 24-25 Mar 1996.
- [SLD97] K. Abe et al. (SLD Collaboration), *A Preliminary measurement of $R(b)$ using the upgrade SLD vertex detector*, Conference Report (1997), [SLAC-PUB-7585](#), Submitted to International Europhysics Conference on High- Energy Physics (HEP 97), Jerusalem, Israel, 19-26 Aug 1997.
- [SLMS03] T. Sjostrand, L. Lonnblad, S. Mrenna and P. Skands, *PYTHIA 6.3: Physics and manual*, (2003), [hep-ph/0308153](#).
- [SMS06] T. Sjostrand, S. Mrenna and P. Skands, *PYTHIA 6.4 physics and manual*, (2006), [hep-ph/0603175](#).
- [Spe] T. Speer, private communication.
- [SS02] S. R. Slabospitsky and L. Sonnenschein, *TopReX generator (version 3.25): Short manual*, Comput. Phys. Commun. **148**, 87–102 (2002), [hep-ph/0201292](#).
- [SSZ00] H. Spiesberger, M. Spira and P. M. Zerwas, *The standard model: Physical basis and scattering experiments*, (2000), [hep-ph/0011255](#).
- [Str00] B. Stroustrup, *The C++ Programming Language*, Addison-Wesley, 3rd edition, 2000.

- [Super-Kamiokande98] Y. Fukuda et al. (Super-Kamiokande Collaboration), *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81**, 1562–1567 (1998), [hep-ex/9807003](#).
- [SZ97] M. Spira and P. M. Zerwas, *Electroweak symmetry breaking and Higgs physics*, (1997), [hep-ph/9803257](#), The presented figures are updated for a top quark mass of $M_t = 175$ GeV.
- [Tev] Tevatron Electroweak Working Group, <http://tevewwg.fnal.gov/higgs/>.
- [tH71] G. 't Hooft, *RENORMALIZABLE LAGRANGIANS FOR MASSIVE YANG-MILLS FIELDS*, Nucl. Phys. **B35**, 167–188 (1971), [Nucl.Phys.B35:167-188,1971](#).
- [tH73] G. 't Hooft, *Dimensional regularization and the renormalization group*, Nucl. Phys. **B61**, 455–468 (1973), [Nucl.Phys.B61:455-468,1973](#).
- [tH74] G. 't Hooft, *MAGNETIC MONOPOLES IN UNIFIED GAUGE THEORIES*, Nucl. Phys. **B79**, 276–284 (1974), [Nucl.Phys.B79:276-284,1974](#).
- [tHV72] G. 't Hooft and M. J. G. Veltman, *REGULARIZATION AND RENORMALIZATION OF GAUGE FIELDS*, Nucl. Phys. **B44**, 189–213 (1972), [Nucl.Phys.B44:189-213,1972](#).
- [TOT] TOTEM home page, <http://totem.web.cern.ch/>.
- [Wag05] W. Wagner, *Top quark physics in hadron collisions*, Rept. Prog. Phys. **68**, 2409–2494 (2005), [hep-ph/0507207](#).
- [Wei67] S. Weinberg, *A MODEL OF LEPTONS*, Phys. Rev. Lett. **19**, 1264–1266 (1967), [Phys.Rev.Lett.19:1264-1266,1967](#).
- [Wei76] S. Weinberg, *IMPLICATIONS OF DYNAMICAL SYMMETRY BREAKING*, Phys. Rev. **D13**, 974–996 (1976), [Phys.Rev.D13:974-996,1976](#).
- [Win02] M. Winkler, *A comparative study of track reconstruction methods in the context of CMS physics*, PhD thesis, Vienna University of Technology and CERN, 2002.

Curriculum Vitae

Education

1974	Städt. Kath. Grundschule Birkesdorf
1978	Städtisches Rurtal-Gymnasium Düren
1987	Allgemeine Hochschulreife
1987-1988	Bundeswehr
1988	RWTH Aachen
1992	Vordiplom (pre-diploma) in physics at RWTH Aachen
1999	Diplom (diploma) in physics at the H1 experiment (topic: measurement of multi-jet cross sections)
since 2000	PhD in physics at I.Physikalisches Institut b, RWTH Aachen

Work and teaching experience

1989-1992	Employment at the company ANNOTEXT (programs for lawyers and notaries in Germany, Switzerland and Austria), fields of activity: programmer, teacher, consultant
1992	Employment at GFE (RWTH Aachen), field of activity: operator of an electron microscope
1992-1997	Employment at WZL (RWTH Aachen) / Fraunhofer Institut IPT, field of activity: realization and automatization of an ultrasound scanner for non-destructive measurements of fiber composites (acquired occupational skills in artificial intelligence, fuzzy logic) 1996 Employment at STUDIENKREIS Nachhilfe (firm for extra tuition), field of activity: teaching physics, mathematics, computing science, chemistry, etc.
1999-2000	Employment at the Max-Planck-Institut für Physik (Werner Heisenberg Institut) in Munich, field of activity: Monte Carlo production for the ALEPH experiment
2000-2005	Setup (incl. hardware) and maintaining of a Linux cluster for ALEPH and CMS physics production, reconstruction and analysis activities at I.Physikalisches Institut b, RWTH Aachen
2000-2005	Tutoring of undergraduate students in physics, computer science, statistics and computer algebra at RWTH Aachen; supervising undergraduate students in physical experimentation at RWTH Aachen
2003-2005	Primary developer of the ORCA Jets reconstruction package for the CMS experiment
since 2006	Research Associate at the Boston University

Publications (leading authorship)

Experiment	Year	Title
H1	2001	Three-jet production in deep-inelastic scattering at HERA (hep-ex/0106078)
ALEPH	2006	Internal note: Studies of a new Rb measurement for the ALEPH experiment (in preparation)
CMS	2005	Jet Reconstruction and Performance in the CMS Detector (CMS AN-2005/005)
	2006	Simulation of Misalignment Scenarios for CMS Tracking Devices (CMS AN-2006/008)
	2006	CMS Physics TDR: jets and missing transverse energy, testing robustness of b-tagging algorithms
	2006	C++ Algorithms for b-tagging in CMS (in preparation)
	2006	CMS Higgs discovery potential summary note: WH, $H \rightarrow b\bar{b}$ (in preparation)

Language skills

German(native), English, French
