UNIVERSITE PIERRE & MARIE CURIE
LA SCIENCE A PARIS

Thèse de doctorat de l'Université Paris 6

spécialité
Champs Particules Matière

présentée par
Eli BEN-HAIM

pour obtenir le grade de
Docteur de l'Université Paris 6

La fonction de fragmentation du quark b, du LEP au TeVatron

The b Quark Fragmentation Function, From LEP to TeVatron

soutenue le 21 décembre 2004 devant le jury composé de :

M. Jean-Eudes AUGUSTIN	Président
M. Matteo CACCIARI	Examinateur
M. Lorenzo FOA	Rapporteur
M. Henry FRISCH	Examinateur
M. Patrick ROUDEAU	Directeur de thèse
M. André ROUGÉ	Rapporteur
Mme Aurore SAVOY-NAVARRO	Directrice de thèse
M. Jean-Bernard ZUBER	Examinateur

A ma mère

Contents

La fonction de fragmentation du quark b, du LEP au TeVatron **9**
Introduction . . . 9
Théorie . . . 9
Le cadre expérimental . . . 12
Analyse de DELPHI . . . 13
Extraction de la partie non perturbative . . . 16
Fragmentation des quarks *b* mesurée dans CDF . . . 20

Annexe : chapitres en Anglais **25**

Abstract **27**

1 Introduction **29**

2 Theory of Bottom Production, Fragmentation and Decay **31**
2.1 Overview: The Life Story of a Bottom Quark . . . 31
2.2 Bottom Quark Production in The Hard Process . . . 33
2.2.1 Bottom Quark Production at LEP . . . 33
2.2.2 Bottom Quark Production at the TeVatron . . . 34
2.3 Theoretical Aspects of *b* Fragmentation . . . 37
2.3.1 Definitions of Fragmentation Functions . . . 37
2.3.2 Perturbative and Non-perturbative Parts . . . 38
2.3.3 Perturbative QCD . . . 39
2.3.3.1 Theoretical QCD Calculations . . . 39
2.3.3.2 Parton Showers in Monte Carlo Generators . . . 41
2.3.4 Non-perturbative QCD . . . 44
2.3.4.1 Hadronization in Monte Carlo Generators . . . 44
Independent Hadronization . . . 45
Cluster Hadronization . . . 45
String Hadronization . . . 45
Baryon Production . . . 49
2.3.4.2 Phenomenological Hadronization Models . . . 50
The Peterson Model . . . 50
The Collins-Spiller Model . . . 51
The Kartvelishvili Model . . . 52

The Lund Symmetric Fragmentation Function 52
The Bowler Model 53
2.4 Excited States 53
2.5 B-hadron Production Rates 55
2.6 B Decays 56

3 Experimental Framework I- The LEP Collider and the DELPHI Experiment 57
3.1 The Large Electron Positron Collider 57
3.2 The DELPHI Experiment 59
3.3 Tracking Detectors 61
3.3.1 The Vertex Detector 61
3.3.2 The Inner Detector 62
3.3.3 The Time Projection Chamber 63
3.3.4 The Outer Detector 63
3.4 Other Detectors 65
3.4.1 Ring Imaging Cherenkov Detectors 65
3.4.2 Electromagnetic and Hadron Calorimeters 65
3.4.3 Scintillators 66
3.4.4 Muon Chambers 66
3.5 Particle Identification and Reconstruction 67
3.5.1 Track Reconstruction 67
3.5.1.1 Primary Vertex Reconstruction 68
3.5.1.2 Impact Parameter Reconstruction 68
3.5.2 Hadron Identification 68
3.5.3 Lepton Identification 70
3.6 DELPHI Monte-Carlo Simulation 70
3.7 Data Reprocessing 71

4 Experimental Framework II- The TeVatron Collider and the CDF Experiment 73
4.1 TeVatron - the Source of $p\overline{p}$ Collisions 73
4.2 The CDF-II Detector 75
4.3 Standard Definitions in CDF-II 76
4.4 Tracking Systems 78
4.4.1 Silicon Tracking Detectors 79
4.4.2 Central Outer Tracker 80
4.4.3 Pattern Recognition Algorithms 81
4.4.4 Momentum Scale 83
4.5 Time of Flight 85
4.6 Calorimeters 85
4.7 Muon Systems 86
4.8 Triggering 87
4.8.1 Level 1 Trigger 89
4.8.2 Level 2 Trigger 90
4.8.3 Level 3 Trigger 91
4.9 Luminosity Measurement 92

5 B Fragmentation at DELPHI **95**
5.1 General Event Selection and Jet Energy Measurement . . . 95
5.1.1 Data/Monte-Carlo comparison and adjustments . . . 97
5.1.1.1 Accuracy of track reconstruction . . . 97
5.1.1.2 Efficiency and track energy distribution . . . 97
5.1.2 Jet energy reconstruction . . . 99
5.2 B-Energy Reconstruction . . . 102
5.3 Selection of B Candidates . . . 104
5.4 Measurement of the B-Fragmentation Distribution . . . 105
5.4.1 Fit Results on Real Data Events . . . 107
5.4.2 Fit Results on Simulated Events . . . 109
5.5 Systematic Uncertainties . . . 109
5.5.1 Real Data and Simulation Tuning . . . 110
5.5.1.1 Energy Calibration . . . 110
5.5.1.2 Level of the Non-b Background . . . 111
5.5.1.3 Track Energy and Multiplicity Tuning . . . 112
5.5.1.4 Jet Multiplicity . . . 112
5.5.1.5 Summary . . . 113
5.5.2 Physics Parameters . . . 113
5.5.2.1 b-Hadron Lifetimes . . . 114
5.5.2.2 B^{**} Production Rate . . . 115
5.5.2.3 b-Hadron Charged Multiplicity . . . 115
5.5.2.4 $g \rightarrow b\bar{b}$ Rate . . . 115
5.5.2.5 Summary . . . 116
5.5.3 Parameters Used in the Analysis . . . 116
5.5.3.1 Parametrization of the Weight Function . . . 116
5.5.3.2 b-Tagging Selection . . . 116
5.5.3.3 Jet Clustering Parameter Value . . . 116
5.5.3.4 Level of Ambiguous Energy . . . 117
5.5.3.5 Secondary Vertex Charged Multiplicity . . . 118
5.5.3.6 Summary . . . 118
5.6 Comparison with Other Experiments . . . 118

6 Extraction of the x-Dependence of the Non-perturbative QCD Component **123**
6.1 Introduction . . . 123
6.2 Extracting the x-Dependence of the Non-perturbative QCD Component . . . 125
6.3 x-Dependence Measurement of the Non-perturbative QCD Component . . . 127
6.3.1 The Perturbative QCD Component is Provided by a Generator . . . 127
6.3.2 The Perturbative QCD Component is Obtained by an Analytic Computation Based on QCD . . . 128
6.4 Results Interpretation . . . 132
6.4.1 Comparison with Models . . . 132
6.4.2 Proposal for a New Parametrization . . . 134
6.5 Checks . . . 137
6.5.1 The Use of a Fitted Parametrization . . . 137

6.5.2 The Effect of Parametrization . . . 137
6.5.3 Number of Degrees of Freedom . . . 137
6.5.4 Using a Different Tuning of the Monte Carlo . . . 139
6.6 Combination of Fragmentation Distributions from All Experiments . . . 140
6.7 Comparison of Results for All Experiments . . . 141
6.8 Thoughts about Fitting Moments of Fragmentation Functions . . . 141
6.9 Charm Fragmentation . . . 146
6.10 Conclusions . . . 149

7 B Fragmentation and Related Studies at CDF 151
7.1 Introduction . . . 151
7.2 Data Sample . . . 153
7.2.1 Reconstruction of $B^{\pm} \to J/\psi K^{\pm}$ . . . 153
7.2.2 Subtracting the Backgrounds in the Data . . . 155
7.3 Monte Carlo Samples . . . 158
7.3.1 General Description . . . 158
7.3.2 PYTHIA Parameters . . . 160
7.4 Outline of the Analysis Method . . . 161
7.5 Preliminary Monte Carlo Studies . . . 164
7.6 Data and Monte Carlo Comparisons . . . 166
7.6.1 Comparisons with msel=5 Samples . . . 166
7.6.2 Comparisons with an msel=1 Sample . . . 172
7.7 A Method of Fitting the Fragmentation Function Parameters . . . 179
7.8 An Estimate of the *b* Production Cross Section . . . 188
7.8.1 Evaluation of Efficiency . . . 188
7.8.2 The Inclusive *b* Quark Production Cross Section . . . 189
7.8.3 Statistical Error Estimation . . . 192
7.8.4 Systematic Error Estimation . . . 192
7.8.4.1 Luminosity . . . 192
7.8.4.2 Branching Ratios and Production Fraction . . . 193
7.8.4.3 Trigger and Reconstruction Efficiencies . . . 193
7.8.5 Comparison with Other Measurements and with Theoretical Predictions 195

8 Conclusion 199

A The Mellin Transformation 201

B Fitting Histograms of Singular Error Matrices 205

Acknowledgments 209

Bibliography 211

List of Figures 217

List of Tables 223

La fonction de fragmentation du quark b, du LEP au TeVatron

Introduction

L'idée de base de cette thèse est la mesure de la fonction de fragmentation du quark *b* dans un environnement de collisionneur e^+e^-, le LEP, à une énergie dans le centre de masse qui est celle du pic du Z^0 et sa transposition dans le cadre d'un collisionneur hadronique, le TeVatron, à une énergie dans le centre de masse de 1,96 TeV. Pour cela on utilise le cadre de l'expérience DELPHI au LEP et celui de l'expérience CDF au TeVatron. Ces deux détecteurs sont particulièrement bien adaptés pour réaliser cette mesure. Tous deux bénéficient du nécessaire ensemble de sous-détecteurs de traces et d'identification de particules très performants. De plus CDF s'est équipé d'un système de déclenchement novateur, basé sur l'information du trajectographe gazeux et du microvertex, qui lui permet de tirer le meilleur profit des hauts taux de production de *b* en collisionneur hadronique. L'analyse développée ici extrait les composantes QCD perturbative et QCD non perturbative de la fonction de fragmentation dans le contexte e^+e^-. En effet, dans ce cadre, le processus QCD est bien compris, ce qui permet d'extraire la partie non perturbative directement à partir des données et non pas, comme habituellement, par comparaison avec les prédictions de divers modèles de fragmentation introduits dans le Monte Carlo et en ajustant les paramètres pour faire correspondre au mieux données simulées et réelles. On peut directement utiliser la partie non perturbative de la fonction de fragmentation ainsi trouvée à l'environnement d'un collisionneur hadronique, pourvu que l'on se place dans le même cadre de travail et que l'on tienne correctement compte de tous les processus QCD de production du b. L'impact de cette étude pour d'autres mesures telle que, par exemple, celle des oscillations B_s ou celle de la section efficace de production du *b* au TeVatron est présenté en conclusion de ce travail.

Théorie

La vie du quark *b* peut se résumer par les différentes étapes qu'il traverse à partir de sa création jusqu'à ce qu'il se matérialise, c'est à dire soit détectable. Considérons pour cela, tout d'abord, le cas de sa création lors de l'annihilation d'un électron et d'un positron en un vecteur boson intermédiaire γ ou Z^0, qui se désintègre à son tour en une paire de quarks *b*, c'est un processus de diffusion électro-faible. Les états initiaux dans cette réaction, nous voulons parler de l'électron et du positron, irradient des photons ce qui réduit d'autant l'énergie disponible pour

le processus fort. L'importance réelle de cet effet est en pratique très faible lorsque l'on effectue les mesures au voisinage du pôle du Z^0. Une fois produits, les quarks b et $\bar{b}$ irradient à leur tour des gluons par interaction forte. Ce processus est décrit par la Théorie Chromodynamique Quantique (QCD) et peut être estimé par les calculs perturbatifs à cause de la haute échelle d'énergie $Q^2 >> \Lambda^2_{QCD}$ qui implique que le paramètre $\alpha_S << 1$. Cette partie du processus que l'on étudie, est définie comme *la composante QCD perturbative de la fragmentation* et sera appelée ainsi tout au long de cette thèse. Elle peut être obtenue par les calculs théoriques de QCD ou à l'aide de générateurs Monte Carlo.

Une fois que les quarks b et $\bar{b}$ se séparent, l'échelle d'énergie du processus diminue et l'interaction de couleur entre les deux quarks devient plus forte. Quand l'échelle d'énergie devient équivalente à Λ_{QCD} le processus considéré entre en régime non perturbatif. Grâce aux interactions entre gluons irradiés, il se crée une région caractérisée par une densité en énergie de plus en plus élevée. À un certain point, le potentiel d'énergie croît de telle sorte qu'il devient suffisant pour permettre la création, à partir du vide, d'une autre paire de quarks $q\,\bar{q}$. Les quarks de cette paire se propagent à leur tour en répétant le même processus et ainsi de suite jusqu'à ce que le système crée des réseaux (clusters) de quarks et de gluons de bas moment interne et de couleur nulle. Il en résulte que le couplage de couleur entre partons à l'intérieur de ces clusters les transforme en hadrons. Cette partie du processus qui comprend la phase dite d'hadronisation sera appelée dans toute la suite *la partie QCD non perturbative de la fragmentation*. Cette partie, comme son nom l'indique, n'est pas calculable. Elle est décrite par des modèles phénoménologiques d'hadronisation. Cette thèse propose une méthode pour extraire le comportement de la partie non perturbative directement des données et indépendamment des hypothèses de tout modèle hadronique.

Finalement les produits du processus d'hadronisation forment à leur tour deux jets de particules qui se déplacent plus ou moins dans la même direction que les quarks b et $\bar{b}$ qui ont été produits dans l'interaction d'origine. Parmi les particules dans ces jets il y a deux hadrons B qui contiennent les quarks b initiaux. S'ils sont produits au pôle du Z^0, ces hadrons B transportent environ 70% de l'énergie des quarks initiaux. Le restant de l'énergie est distribué entre toutes les autres particules dans le jet.

Nous pouvons faire ici deux remarques. La première est que le quark top qui est le plus lourd des quarks, avec une masse beaucoup plus grande que Λ_{QCD}, se désintègre par processus faible, avant hadronisation. La deuxième remarque est que pour les quarks légers (c'est à dire dont la masse est plus faible que Λ_{QCD}), contrairement aux quarks lourds, le quark initial ne peut être identifié dans les états hadroniques finaux. La possibilité de pouvoir suivre les quarks b depuis leur production jusqu'à leur désintégration, fournit une sonde unique pour pouvoir explorer les phénomènes qui interviennent tout au long de ce processus. C'est une motivation essentielle des études de fragmentation du quark b. Pour les quarks charmés, le couplage des gluons aux paires $c\,\bar{c}$, ne peut-être négligé aux énergies de LEP, ce qui dilue un peu le lien entre le quark initialement produit et le hadron lourd reconstruit.

Les analyses expérimentales présentées dans cette thèse sont basées sur l'étude des hadrons B qui se désintègrent par interaction faible. Néanmoins la plupart des hadrons B, à l'issue du processus d'hadronisation, sont des états excités: des mésons B^* ou B^{**}. Ces états se désintègrent par interaction forte ou électromagnétique en hadrons B plus légers et en d'autres particules comme les pions, les kaons et les photons. Dues aux échelles en temps très courtes des interactions forte et électromagnétique, la désintégration de ces états excités se produit

au voisinage du point d'interaction primaire. La production de ces états est d'ailleurs partie intégrante de ce que l'on appelle: la fragmentation.

Par contre, les hadrons B qui se désintègrent par processus faible traversent une distance relativement longue (environ 3 mm au pôle du Z^0), avant de se désintégrer. Cette distance est due au temps de vie du hadron B produit (environ 1.6 ps) et à ce que ce dernier emporte, en moyenne, une large fraction de l'énergie du quark *b*. Comme les produits de désintégration des hadrons B ont une multiplicité moyenne de cinq traces chargées environ, il est possible d'étiqueter expérimentalement les hadrons B, par leur temps de vie, en se basant sur des mesures précises des trajectoires des traces chargées dans le voisinage de la région d'interaction des faisceaux.

Dans un collisionneur e^+e^-, l'état initial est très bien défini et constitue donc un excellent laboratoire pour ces études de QCD. Le fait que l'énergie des faisceaux soit une bonne approximation de l'énergie des quarks initiaux fait que ces machines sont vraiment appropriées pour les études de fragmentation. Par contre, dans un collisionneur hadronique, l'énergie des quarks *b* est inconnue et varie d'un événement à l'autre. Il y a donc une différence cruciale dans les études sur la fragmentation entre ces deux environnements expérimentaux. La source dominante de quarks *b* au TeVatron provient de l'interaction forte. La contribution à la section efficace totale de production des *b* par les processus électrofaibles: $W \to c\bar{b}$ ou $Z \to b\bar{b}$ est petite et peut même être négligée en général. Il y a trois mécanismes qui contribuent à la production QCD d'une paire de quarks *b* et $\bar{b}$. Ce sont: la création de saveur ou processus direct d'annihilation quark-antiquark (ou deux gluons), l'excitation de saveur et le "splitting" de gluons. Cette thèse illustre leur importance relative.

Considérons maintenant les aspects théoriques de la fragmentation du *b*. La fonction de fragmentation du quark *b* est la fonction de densité de probabilité de la fraction d'énergie du quark *b* qui sera emportée par le hadron B. Une autre variable communément utilisée dans les Monte Carlo est *z*, définie comme étant le rapport entre la somme de l'énergie et de l'impulsion du hadron, évaluées le long de la direction du quark, et de la somme correspondante pour le quark *b*. Ce rapport est calculé en se plaçant dans le référentiel du centre de masse du sous-processus qui s'hadronise. Un point délicat est la définition de la frontière entre partie perturbative et non perturbative. Elle est en quelque sorte un peu arbitraire. La composante non perturbative décrit l'hadronisation, mais on doit prendre garde qu'elle contient aussi tous les termes qui n'ont pas été pris en considération dans la partie perturbative. Le fait que la composante non perturbative de la fonction de fragmentation dépende fortement de la composante perturbative a été souvent négligé. Ceci a pu amener à des conclusions erronées sur le non accord entre prédictions de QCD et résultats expérimentaux sur la section efficace du processus de production $b\,\bar{b}$.

C'est le Théorème de Factorisation de QCD qui permet de séparer les parties perturbatives et non-perturbatives. Il a comme conséquence importante, pour l'analyse présentée dans cette thèse, que la partie non perturbative ne dépend pas de l'état initial et donc de l'environnement correspondant au collisionneur e^+e^- ou hadronique.

Plusieurs approches perturbatives ont été considérées dans l'étude présentée ici. Tout d'abord l'approche QCD théorique qui donne un développement en série en puissance de α_S de la fonction de fragmentation qui s'arrête au "leading log" (LL), au "next to leading order" (NLL) ou bien à l'approximation NLL avec "Dressed Gluon Exponentiation". L'ensemble de ces différents calculs ont été effectués au départ de cette thèse et comparés avec ceux de leurs auteurs.

Ces calculs détaillés sont nécessités par l'approche nouvelle employée dans cette thèse afin d'extraire la fonction de fragmentation (voir le détail dans le chapitre 6 de l'appendice). Une autre approche consiste à extraire la partie perturbative de la fonction de fragmentation à partir des générateurs Monte Carlo.

Cette physique doit être complétée par une composante non calculable: la partie non perturbative qui décrit l'hadronisation. D'habitude elle est décrite par un modèle phénoménologique qui est inclus dans le générateur d'événements. Il est d'usage d'utiliser ces Monte Carlo pour ajuster les paramètres des modèles afin de les faire correspondre au mieux aux données. Il existe plusieurs modèles pour décrire la physique de l'hadronisation. Ils utilisent une hadronisation indépendante ou une modélisation basée sur des cordes ("strings") ou bien des clusters. Un certain nombre de ces modèles, parmi les plus utilisés (Peterson, Collins-Spiller, Kartvelishvili, Lund ou Bowler), sont décrits dans le Chapitre 2 de l'appendice à cette thèse. La partie non-perturbative de la fonction de fragmentation extraite avec notre approche sera comparée à celles obtenues avec ces modèles.

Le cadre expérimental

Cette thèse s'est déroulée dans deux cadres expérimentaux. L'un est celui de l'expérience DELPHI auprès du collisionneur e^+e^- LEP au CERN qui s'est arrêté en Novembre 2000, et les données utilisées pour ce travail sont celles recueillies au pic du Z^0, c'est à dire à l'énergie dans le centre de masse de 91 GeV, entre 1992 et 1995. L'autre cadre expérimental est celui de l'expérience CDF auprès du collisionneur proton antiproton, le TeVatron, au Fermi National Laboratory aux USA dans la phase du Run II. Les données considérées pour cette partie sont celles qui ont été enregistrées jusqu'à Août 2004, à l'énergie dans le centre de masse de 1,96 TeV.

Dans l'une et l'autre expérience une partie cruciale du système de détection est constituée par l'ensemble des détecteurs de traces que ce soient les microvertex à microbandes de Silicium ou les détecteurs centraux gazeux: la TPC (Time Projection Chamber) dans le cas de DELPHI ou le COT (Central Outer Tracker), chambre avec cellules tiltées, pour CDF. Tous deux donnent une excellente résolution pour la mesure des impulsions des particules chargées. De plus, grâce à la mesure du dE/dx, ils permettent d'identifier les particules. Des détecteurs de microvertex très performants dans les deux cas permettent de reconnaître des vertex secondaires et de mesurer les paramètres d'impact des traces chargées ce qui autorise l'étiquetage des particules issues de la désintégration des hadrons-B. Dans le cas de DELPHI, le détecteur RICH a joué un rôle pour l'identification des particules chargées. À CDF, un détecteur de temps de vol performant fournit un moyen d'identification des particules très précieux, et pour la physique du B en particulier. La grande différence entre ces deux expériences est naturellement due à leur environnement respectif. Dans un collisionneur hadronique le haut taux d'événements nécessite des systèmes de déclenchement très sophistiqués. Ce n'est pas le cas dans une machine e^+e^-. CDF a été le premier détecteur, dans un collisionneur hadronique, à installer un microvertex performant ceci depuis le Run I en 1992. Pour le Run II, la grande innovation a été d'inclure, au niveau premier du déclenchement, l'information du COT et, au niveau 2, celle du microvertex. C'est une première au niveau mondial qui permet de tirer plein profit des hauts taux de production de b produits par interaction forte dans les collisionneurs hadroniques. Un déclenchement

de niveau 1 (XFT = eXtremly Fast Tracker), utilisant l'information des couches axiales du détecteur de traces gazeux, le COT, dans le barrel central, permet de reconstruire en temps réel les traces et de déterminer, avec une très bonne précision, leur impulsion transverse (lorsqu'elle est supérieure à 1,5 GeV/c). Un déclenchement au deuxième niveau part des traces reconstruites par le XFT dans le COT et les relie à celles qu'il détermine alors dans le microvertex, constitué de 5 couches double face de microbandes en Silicium. Il peut ainsi déterminer les traces venant d'un vertex secondaire (produites par désintégration de hadrons charmés ou beaux) et calculer les paramètres d'impact correspondants. C'est un outil qui révolutionne la manière de faire la Physique du B ou du charme dans les collisionneurs hadroniques. L'ensemble des trajectographes, le système de déclenchements associés et les outils d'identification des particules de l'expérience CDF en font un outil de mesure particulièrement adéquat pour l'étude de la fragmentation qui est l'objet de cette thèse. Plus de détails sur les environnements expérimentaux de DELPHI et de CDF peuvent être trouvés dans les chapitres 3 et 4 de l'appendice à cette thèse ainsi que dans les références qui y sont reportées.

Analyse de DELPHI

Des mesures de la fonction de fragmentation du quark *b*, au pôle du Z^0, existent et ont été publiées par d'autres expériences e^+e^-.

La mesure du spectre de l'énergie du hadron beau, effectuée dans cette thèse, utilise une méthode inclusive, comme la plupart des expériences précédentes. Seule l'expérience ALEPH a reconstruit des désintégrations semileptoniques des mésons beaux pour faire ce type d'études.

Les événements retenus pour cette analyse sont issus d'un lot enrichi en beauté. Cet enrichissement est obtenu en s'assurant de la présence de traces chargées ayant un paramètre d'impact important par rapport au point d'interaction des faisceaux. Ceci est réalisé par un algorithme (standard dans DELPHI) qui combine les mesures précises de ces traces au niveau du détecteur de vertex en silicium. Pour les événements situés dans l'acceptance de ce détecteur, on sélectionne ainsi les événements contenant de la beauté avec une efficacité voisine de 70% et la pureté de l'échantillon est voisine de 90%.

Il convient alors de distinguer les traces issues de la désintégration des hadrons beaux de celles provenant de l'hadronisation des quarks *b* initiaux. Pour cela, dans un même jet, on détecte la présence de traces chargées ayant un paramètre d'impact significatif par rapport au vertex primaire de l'événement. Ces traces sont utilisées pour former un vertex secondaire. Les traces chargées restantes, dans le jet, sont assignées au vertex primaire ou bien au vertex secondaire en fonction de leurs paramètres d'impact respectifs par rapport à ces deux positions. Les traces neutres présentes dans le jet sont réparties entre ces deux vertex suivant la valeur de leur rapidité, mesurée par rapport à l'axe du jet. Les traces les plus "rapides" sont associées au vertex secondaire comme l'indique la simulation des événements. On s'assure aussi que le système de traces ainsi formé, représentant les traces issues de la désintégration d'un hadron B, a une masse inférieure à une valeur maximale de 5,2 GeV, avant d'ajouter une nouvelle trace.

Certaines particules ne sont pas reconstruites pour des raisons physiques (s'il s'agit de neutrinos) ou bien liées à l'efficacité du détecteur ou bien parce qu'elles interagissent dans la matière de l'appareillage.

L'estimateur de l'énergie du hadron B est obtenu de la manière suivante. Nous effectuons

un ajustement de l'ensemble des paramètres cinématiques des traces en tenant compte de leurs erreurs de mesure, en autorisant pour chaque jet la présence d'un 4-vecteur manquant et en imposant la conservation de l'énergie-impulsion. Ceci permet notamment de calibrer l'énergie des jets à partir de l'énergie du faisceau. L'énergie du hadron B est estimée en soustrayant, à l'énergie du jet celle des traces qui n'ont pas été attribuées au B. Il s'agit d'une méthode originale par rapport à celles utilisées dans les autres expériences qui utilisent les traces attribuées au B pour construire un estimateur de son énergie.

À partir du programme MINUIT, on détermine une fonction de pondération à appliquer aux événements simulés de manière à ce qu'ils reproduisent la distribution en énergie du hadron B mesurée dans les données réelles, après avoir franchi les mêmes étapes de l'analyse. Les résultats de ces ajustements sont donnés sur la Figure 1.

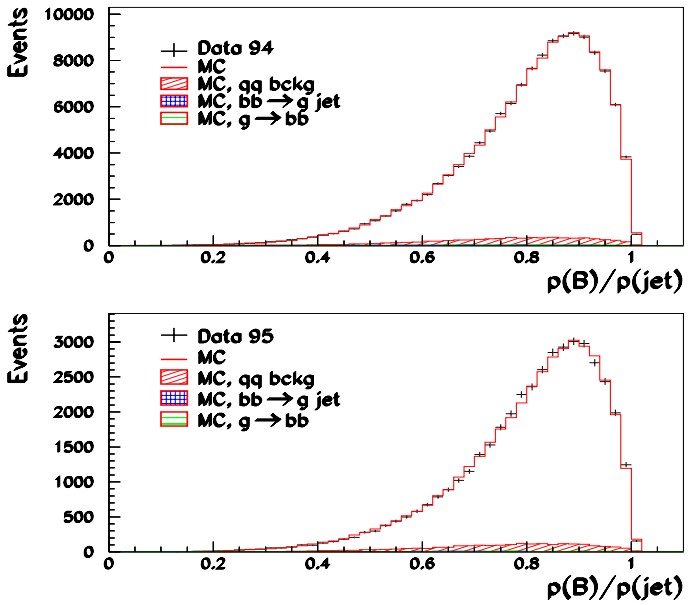

Figure 1: *Distributions fittées, en fonction de $x_p^{rec} = p_B^{rec}/p_{jet}$, pour les événements sélectionnés. En haut: échantillon 1994. En bas: échantillon 1995.*

Cette fonction de pondération est obtenue en fonction de la valeur de la variable z qui définit la valeur de la fonction de fragmentation non-perturbative utilisée dans le générateur d'événements simulés. La Figure 2 montre la fonction ainsi mesurée, qui a un comportement assez différent de celui de la fonction de Peterson utilisée initialement dans la simulation.

La distribution en énergie du hadron B, corrigée de l'ensemble des effets liés à la sélection et à la méthode de mesure, est obtenue en appliquant la fonction de pondération ainsi trouvée aux événements simulés avant toute sélection.

L'analyse est basée sur une statistique de 134000 candidats extraits des données enregistrées en 1994 et de 42000 candidats en 1995. Les incertitudes sont dominées par les effets systématiques et principalement par ceux liés à la calibration en énergie. L'importance de ce type d'incertitudes a été évalué en divisant le lot des données en sous-échantillons sur lesquels

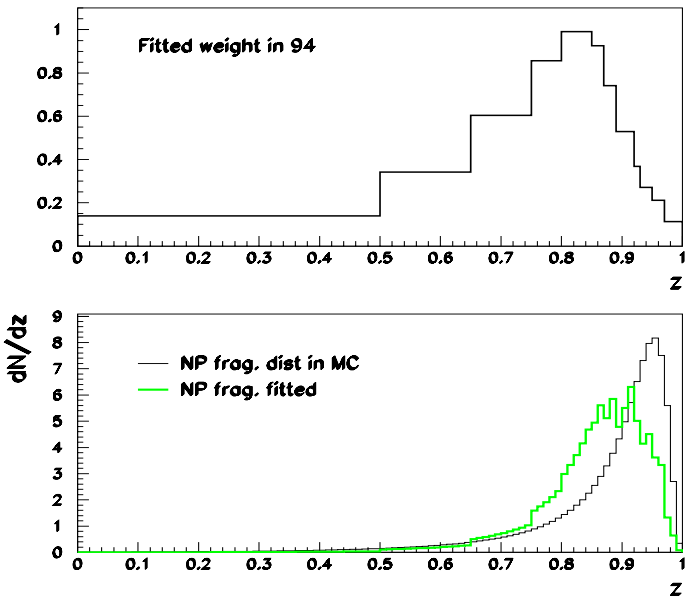

Figure 2: *Distributions, en fonction de la variable z, obtenues pour les événements sélectionnés en 1994. En haut: Distribution fittée des poids. En bas: Comparaison entre la distribution de Peterson utilisée dans le générateur et la distribution fittée sur les données.*

l'analyse est répétée. Une incertitude liée à la mesure de l'énergie du B est déduite de la dispersion des résultats obtenus. Plusieurs autres sources possibles d'incertitudes systématiques ont été analysées. Elles ont été regroupées en trois catégories: celles liées à l'accord entre les données réelles et la simulation, celles venant d'incertitudes sur les paramètres définissant les propriétés des mésons B et celles liées aux sélections faites pour l'analyse. Mise à part l'incertitude sur la mesure de l'énergie du B, l'autre erreur dominante provient de la modélisation de la fonction de pondération. Cette dernière a été estimée en comparant les distributions initiales et reconstruites pour l'énergie des hadrons B dans les données simulées.

Les mesures de la fonction de fragmentation du quark *b* ont été fournies sous forme de valeurs intégrées dans neuf intervalles, pour lesquelles les matrices d'erreur statistique et systématique ont été données. Ces informations sont utilisées, dans le chapitre suivant pour extraire analytiquement la composante non-perturbative de la fonction de fragmentation du quark *b*.

Ces résultats sont comparés, sur la Figure 3, aux mesures obtenues par les autres groupes. Les mesures effectuées dans cette thèse sont très voisines de celles de SLD. La fonction de fragmentation trouvée est "moins dure" que celle obtenue par ALEPH. La comparaison, à ce niveau, fait intervenir l'ensemble du spectre en énergie des hadrons B, nous verrons dans le chapitre suivant comment se comparent les composantes non-perturbatives correspondant à ces différentes mesures.

Les valeurs moyennes de ces distributions, $< x_E >$, sont comparées dans la Table 1, avec les autres mesures similaires effectuées à l'énergie du Z^0.

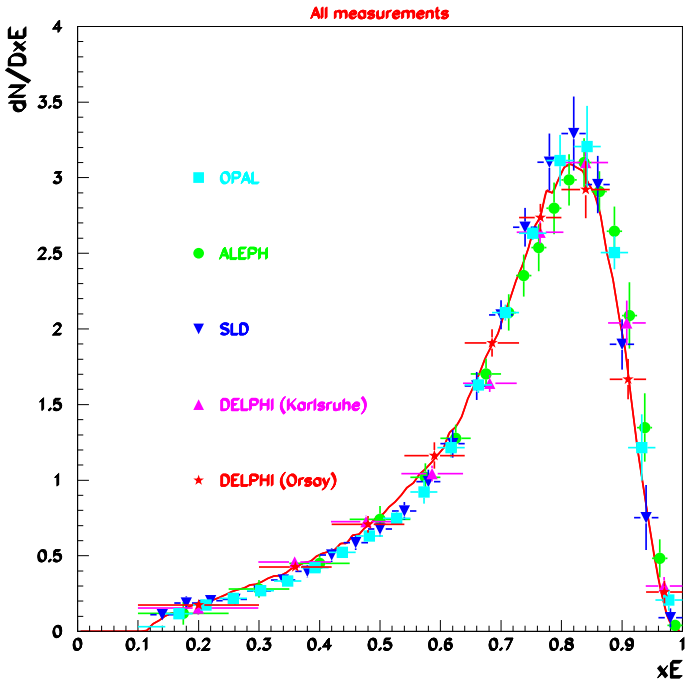

Figure 3: *Comparaison entre les distributions de la variable x_E mesurées par ALEPH, DELPHI, OPAL et SLD.*

Extraction de la partie non perturbative

Nous venons de présenter dans la section précédente la mesure de la fonction de fragmentation du quark *b* dans le cadre de l'expérience DELPHI au LEP. Cette distribution est définie comme celle de la fraction de l'énergie du faisceau emportée par un hadron B qui se désintègre par interaction faible et qui est produit dans le processus $e^+e^- \rightarrow b\bar{b}$ au pôle ou au voisinage du pôle du Z^0. Les expériences ALEPH et OPAL au LEP et SLD au SLC (SLAC) ont également fait ces mesures, ce qui permet d'avoir un ensemble de résultats précis qui renforcent le résultat final.

Nous avons précédemment discuté comment, dans le cadre des collisionneurs e^+e^-, la frag-

Experiment	$< x_E >$	err. stat.	err. syst.
Cette thèse	0.704	0.001	0.008
ALEPH	0.716	0.006	0.006
DELPHI (Karl.)	0.715	0.001	0.005
OPAL	0.719	0.002	0.004
SLD	0.709	0.003	0.004

Table 1: *Différentes mesures de $< x_E >$ à l'énergie du Z^0.*

mentation peut être vue comme résultant des trois étapes suivantes: l'interaction primaire, c'est à dire ici le processus d'annihilation $e^+e^- \to b\bar{b}$, une description par la QCD perturbative de l'émission des gluons par les quarks et une partie QCD non perturbative qui inclut tous les mécanismes qui interviennent entre la production des quarks b et leur matérialisation en hadron B se désintégrant en mode faible. La composante perturbative s'obtient soit par un calcul analytique soit à l'aide de générateurs Monte Carlo. La composante non perturbative est généralement issue d'une paramétrisation selon une modélisation. La comparaison avec les résultats expérimentaux se fait alors en combinant les deux composantes, suivant l'expression:

$$\mathcal{D}_{predicted}(x) = \int_0^1 \mathcal{D}_{pert.}(z) \times \mathcal{D}^{model}_{non-pert.}(\frac{x}{z})\,\frac{dz}{z} \tag{1}$$

pour évaluer la dépendence en x attendue. La variable x est par exemple égale à $x_E = E_B/E_{faisc.}$ qui varie entre $x_{min} = 2m_B/\sqrt{s}$ et l'unité. Dans la suite nous utiliserons:

$$x = p_B/p_b = \sqrt{x_E^2 - x_{min}^2}/\sqrt{1 - x_{min}^2} \tag{2}$$

qui varie entre 0 et 1.

Les paramètres d'un modèle particulier sont ajustés en comparant la dépendance en x de la distribution de la fragmentation du quark b mesurée, avec celle attendue. Les différentes expériences au LEP et à SLC ont déjà fait ces comparaisons en utilisant, pour la partie perturbative, les calculs par Monte Carlo avec comme générateurs JETSET ou HERWIG qui incluent des développements en gerbes de partons. Avec la précision actuelle des mesures, il est démontré que la plupart des modèles disponibles pour la partie non perturbative, sont incapables de reproduire correctement les résultats expérimentaux. Ceux qui s'en approchent le plus sont les modèles de Lund et de Bowler. Ces calculs ont été répétés et les résultats ont été rassemblés dans la Table 6.1 et la Figure 6.7.

L'application de la transformée de Mellin à l'expression 1 permet d'obtenir la distribution résultante sous forme d'un produit de deux distributions:

$$\tilde{\mathcal{D}}(N) = \tilde{\mathcal{D}}_{pert.}(N) \times \tilde{\mathcal{D}}_{non-pert}(N) \tag{3}$$

où $\tilde{\mathcal{D}}(N) = \int_0^\infty x^{N-1}\mathcal{D}(x)dx$ est le moment d'ordre N de la fonction $\mathcal{D}(x)$. À partir du calcul des moments de la fonction de fragmentation mesurée, et de ceux de la partie perturbative, il est possible d'obtenir, de manière indépendante d'un modèle, les moments de la composante non perturbative. Ceci est une démarche utilisée par plusieurs théoriciens travaillant dans ce domaine.

L'originalité du travail présenté dans cette thèse a été de calculer la distribution en x de la fonction non-perturbative en appliquant la transformée de Mellin inverse. On peut ainsi obtenir, point par point, une distribution qui peut alors être utilisée dans un générateur ou bien comparée à différents modèles.

La méthode utilisée pour extraire la dépendance en x de la composante non perturbative ne dépend pas d'une fonction de fragmentation mesurée particulière. Dans cette étude, nous avons appliqué au départ cette méthode aux résultats de l'expérience ALEPH, et c'est donc la principale mesure à laquelle il est fait référence dans cette section. Les fonctions de fragmentation mesurées par les autres expériences sont également présentées dans cette thèse.

Nous référons à la section 6.2 de l'appendice à cette thèse pour la description détaillée de la méthode utilisée pour extraire la composante QCD non perturbative. En particulier, dans la section 6.3 de cet appendice, on trouvera l'extraction faite à partir de plusieurs évaluations de la composante QCD perturbative utilisant:

- le générateur JETSET, ajusté avec les données de l'expérience DELPHI, pour tenir compte du développement en gerbes des partons;
- un calcul analytique basé sur la théorie QCD à l'ordre NLL ("Next to Leading Log");
- un calcul analytique basé sur la théorie QCD à l'ordre NLL ("Next to Leading Log") avec DGE ("Dressed Gluon Exponentiation).

Les distributions de la composante QCD non-perturbative de la fonction de fragmentation du quark *b*, obtenues à partir des mesures effectuées dans DELPHI, à l'occasion de cette thèse, et en utilisant deux approches pour évaluer la composante QCD perturbative, sont représentées sur les Figures 4 et 5.

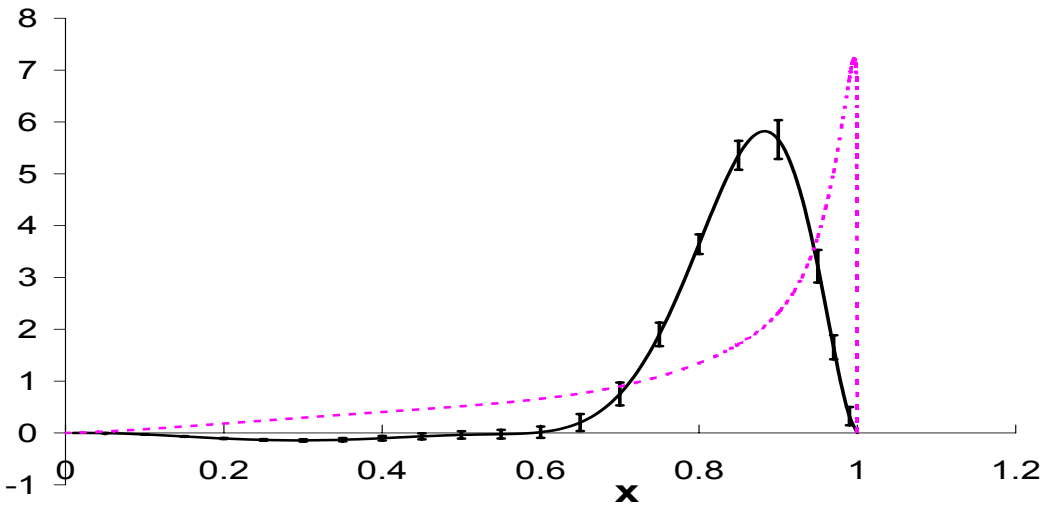

Figure 4: *La distribution de la composante QCD non-perturbative de la fonction de fragmentation du quark b, obtenue à partir des mesures effectuées dans DELPHI, à l'occasion de cette thèse, et en utilisant le générateur JETSET pour évaluer la composante QCD perturbative.*

Dans la section 6.4 de l'appendice à cette thèse, on trouvera une discussion détaillée sur les résultats obtenus avec notre méthode. Une paramétrisation de la composante QCD non perturbative y est donnée, et dans la section 6.5 de l'appendice à cette thèse, sont présentées différentes vérifications faites pour tester la robustesse de la méthode proposée.

Les principales conclusions de cette étude sont les suivantes:

- la dépendance en x de la composante QCD non-perturbative de la fonction de fragmentation a été extraite d'une manière indépendante de tout modèle de physique hadronique. Elle dépend étroitement de la façon dont est évaluée la composante QCD perturbative. La distribution est remarquablement différente de celles prédites par différents modèles.

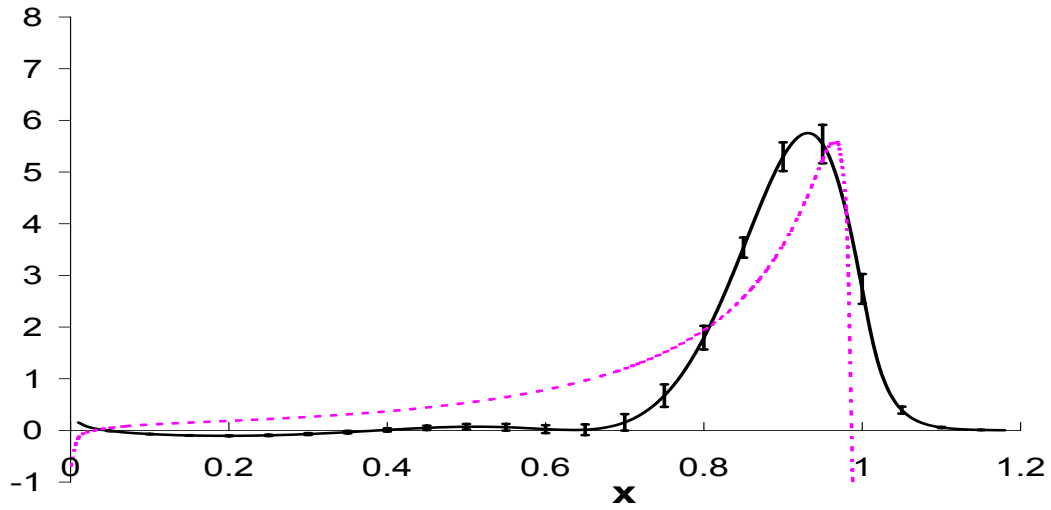

Figure 5: *La distribution de la composante QCD non-perturbative de la fonction de fragmentation du quark b, obtenue à partir des mesures effectuées dans DELPHI, à l'occasion de cette thèse, et en utilisant un calcul analytique basé sur QCD à l'ordre NLL pour évaluer la composante QCD perturbative.*

- Pour $x < 0.6$, cette distribution est compatible avec zéro, ce qui est une indication que la plus grande part de la radiation de gluons est bien reproduite par la composante QCD perturbative évaluée soit en utilisant le Monte Carlo JETSET pour le développement en gerbes de partons, soit par un calcul analytique. Comme la distribution QCD non perturbative est évaluée pour toute valeur de la variable *x*, on peut vérifier si elle reste physique sur tout l'intervalle [0,1] quand est utilisé un générateur Monte Carlo qui fournit la partie perturbative. L'évidence pour des régions non physiques indiquerait que la simulation ou les mesures ne sont pas correctes. Une telle évidence n'a pas été trouvée dans cette analyse.

- Pour $x > 0.6$, la distribution obtenue a une forme semblable à celle prédite par les modèles de Lund symétrique ou de Bowler, quand on prend la partie QCD perturbative donnée par JETSET. Si l'on prend la composante perturbative calculée par Cacciari et Catani, on trouve que, à cause du comportement analytique de la composante QCD perturbative, la distribution de la composante QCD non perturbative doit être prolongée au-delà de $x = 1$. Le comportement en x de la composante non perturbative pour $x > 1$, est déterminé par l'existence possible d'une valeur nulle de $D_{pert}(N)$, pour $N > 0$. Si la partie perturbative a des aspects non physiques, il n'est donc pas justifié de la combiner avec un modèle physique donné. Nous avons trouvé une approche pour résoudre ce problème et nous avons proposé une paramétrisation de la distribution ainsi obtenue.

En se basant sur le théorème de factorisation de QCD, la composante non perturbative ainsi extraite est considérée être valable dans un environnement autre que celui du processus d'annihilation e^+e^-, tant que la partie QCD perturbative est évaluée dans le même cadre, c'est à dire par calcul analytique de QCD ou par un générateur Monte Carlo donné et en utilisant les même valeurs pour les paramètres qui interviennent dans cette évaluation que ce soient: m_b^{pole},

$\Lambda^{(5)}_{QCD}$, ou bien des valeurs ajustées pour les paramètres du générateur.

Dans la section suivante on présente une étude sur la fragmentation du quark *b* dans l'expérience CDF en utilisant comme cadre de travail celui fourni par le générateur Monte Carlo PHYTIA.

Fragmentation des quarks *b* mesurée dans CDF

Le travail expliqué dans les sections précédentes est maintenant appliqué à l'analyse de la production des hadrons B dans le cadre de l'expérience CDF au TeVatron. Ces données vont être comparées aux distributions attendues par le générateur PYTHIA.

La production de quarks *b* lors de collisions $p\overline{p}$ fait intervenir des diagrammes de QCD plus variés qu'en annihilation e^+e^-. L'étude présentée ici est exploratoire car les données ont été enregistrées récemment et la simulation des événements a demandé plusieurs itérations avant d'être réaliste. La possibilité de pondérer des événements simulés, afin de modifier la partie non-perturbative de la fonction de fragmentation a été introduite récemment et les premières données simulées ayant cette facilité sont en cours de production.

Les caractéristiques cinématiques des hadrons B dépendent des contributions respectives des différents mécanismes de production et de la forme de la composante non-perturbative de la fonction de fragmentation. Des différences peuvent exister aussi par rapport aux données simulées si les efficacités du détecteur ne sont pas correctement prises en compte. Afin d'essayer de distinguer ces différents effets trois types d'objets ont été analysés: les hadrons B, les traces chargées qui les accompagnent dans le jet et les traces chargées présentes en-dehors de ce jet. Ces dernières sont principalement utilisées pour vérifier le réalisme de la simulation du détecteur, pour les traces chargées.

Pour faire cette analyse dans CDF, on utilise l'échantillon des données du canal de production exclusif du $B^{\pm} \longrightarrow J/\psi K^{\pm}$. Avec la luminosité intégrée jusqu'à présent, c'est à dire environ 350 pb^{-1}, plusieurs milliers (environ 6000) désintégrations ont été reconstruites. L'avantage de ce canal exclusif, par rapport au canal inclusif, $B \longrightarrow J/\psi X$ analysé par un autre groupe, est que la quadri-impulsion du méson B est bien mesurée, ceci en prenant avantage de toutes les spécificités du détecteur CDF II que nous avons précédemment décrites. Ayant reconstruit l'ensemble des particules issues de la désintégration du méson B il est aussi possible d'étudier les traces qui l'accompagnent, sans avoir de risque de confusion.

L'analyse comporte les aspects suivants:

- à partir d'une simulation comprenant une sous-classe des mécanismes contribuant à la production de la beauté au TeVatron, la sensibilité de différentes observables à la forme de la composante non-perturbative de la fonction de fragmentation a été évaluée. L'information contenue dans les traces accompagnant le B dans son jet sont souvent plus sensibles à ces différences que la distribution en impulsion transverse du hadron-B.

- les différences observées entre les mesures et les résultats de ce type de simulation qui n'inclut que les mécanismes d'annihilation $q\overline{q}$ et la fusion gg sont cependant trop grandes (et incompatibles entre elles) pour pouvoir être corrigées par un changement dans la modélisation du mécanisme d'hadronisation. Les autres mécanismes contribuant à la production de *b*, dans PYTHIA et expliquées dans la partie théorique de cette thèse, sont indispensables.

- les données simulées contenant l'ensemble de ces mécanismes sont trouvées en assez bon accord avec les distributions mesurées. La sensibilité de ces distributions, à l'importance des différents mécanismes, a été étudiée. Les données ont été simulées avec la composante utilisée, par défaut, à CDF pour la partie non-perturbative de la fonction de fragmentation. Cette distribution est moins piquée à hautes valeurs que la fonction extraite des données de LEP. Une procédure a été mise en place afin de pouvoir étudier l'effet d'un changement de modèle à partir d'un nouveau lot d'événements simulés.

Une partie de ces aspects est illustrée par les Figures 6 et 7, qui montrent des comparaisons entre données et simulation, contenant tous les mécanismes de production du quark *b*. La première figure montre la distribution de l'impulsion transverse au faisceau du méson-B. Sur cette figure, les contributions des différents mécanismes de production du B ont été séparés. La deuxième figure montre la distribution de l'impulsion transverse au faisceau des traces dans un cône défini autour du B.

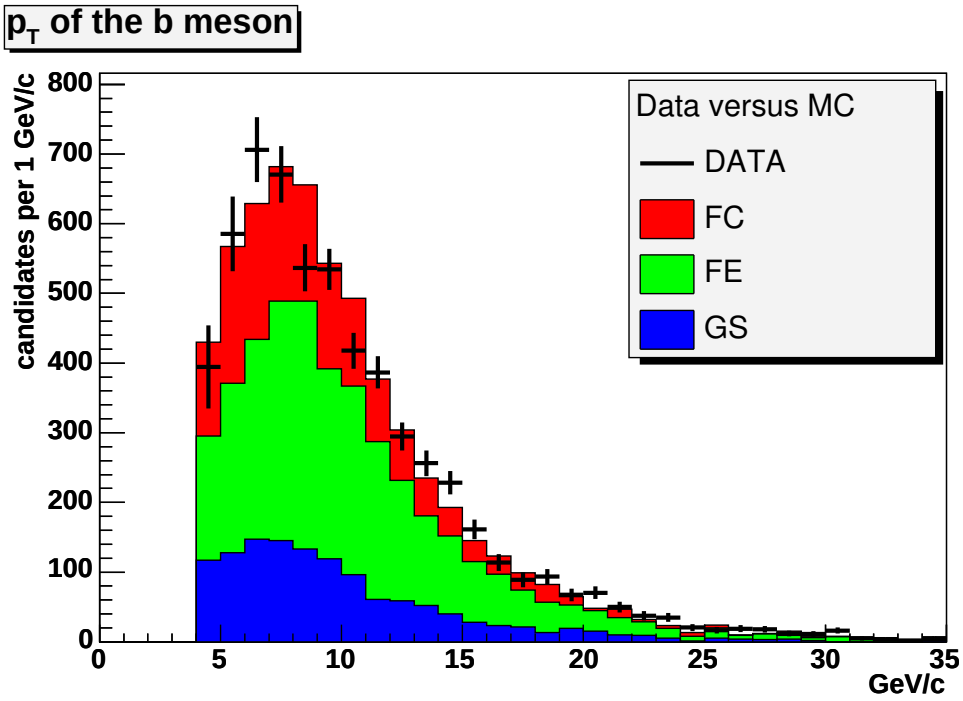

Figure 6: *Comparaison entre données et simulation, contenant tous mécanismes de production du quark b. La distribution de l'impulsion transverse au faisceau du méson-B. Les contributions des différents mécanismes de production du B sont montrés séparément les une des autres.*

On espère pouvoir ainsi déterminer à la fois l'importance des différents mécanismes de production de *b*, présents dans la simulation, et valider l'appliquabilité de la partie non-perturbative de la fonction de fragmentation, mesurée à LEP, à l'environnement du TeVatron. L'accord entre données et Monte Carlo est un ingrédient crucial pour la mesure de la fraction de mésons B qui proviennent des désintégrations: $B^{**} \longrightarrow B\pi$. Dans ce type d'analyses, PYTHIA est nécessaire pour aider à contraindre les spectres de masse du bruit de fond des spectres $B^{\pm}\pi^{\mp}$ (avec les

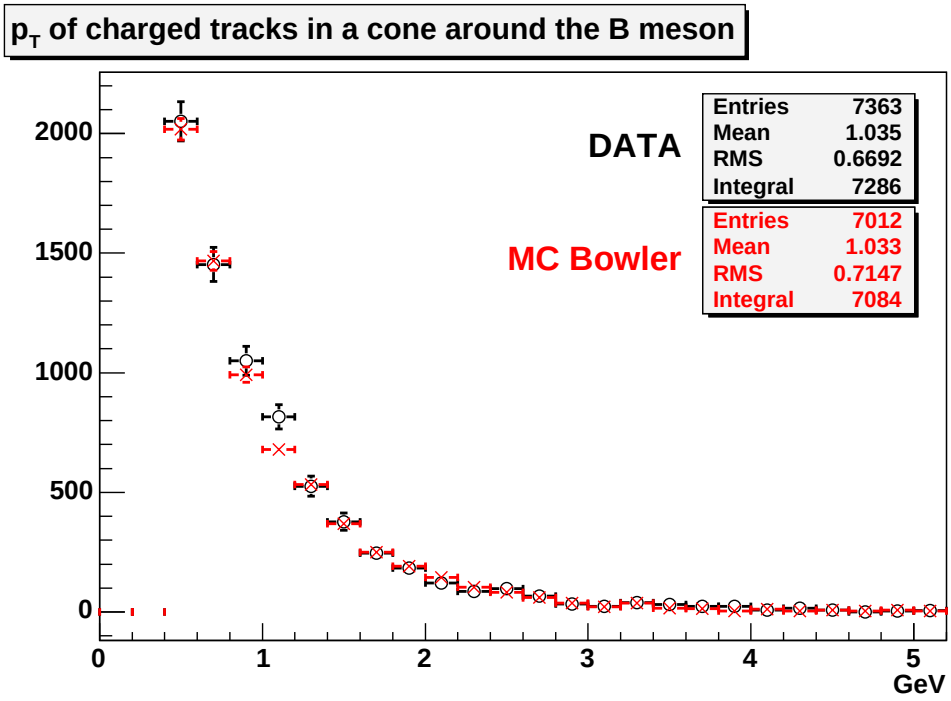

Figure 7: *Comparaison entre données et simulation, contenant tous mécanismes de production du quark b. La distribution de l'impulsion transverse au faisceau des traces dans un cône défini autour du B*

signes de charge corrects: "right sign" RS) et ceux avec les signes de charge incorrects (WS = "wrong sign") c'est à dire $B^{\pm}\pi^{\pm}$.

Cependant la réelle importance de réaliser l'accord entre les données et le Monte Carlo PYTHIA est qu'ainsi on confirmera que PYTHIA ainsi ajusté peut reproduire le comportement cinématique et les corrélations entre les mésons B et les traces qui l'accompagnent. Si un tel accord peut être obtenu de manière univoque, CDF pourrait utiliser PYTHIA pour estimer la dilution dans l'étiquetage dit du Kaon de même signe ("Same Sign Kaon Tagging" , SSKT) dans les études sur le B_s. Ceci constitue un prerequisit pour obtenir une limite sur la fréquence des oscillations $B_s^0 B_s^0$, c'est à dire la mesure de Δm_s.

Nous avons obtenu enfin une valeur de la section efficace de production des hadrons B, pour deux régions en pseudo-rapidité et fourni la variation de cette section efficace en fonction de l'impulsion transverse du hadron B. Nous avons trouvé [1]:

$$\begin{aligned}\sigma(p\bar{p} \to bX,\, p_T(B) > 4\text{GeV/c},\, |\eta| < 1.0) &= 19.0 \pm 0.8\,(\text{stat.}) \pm 1.6\,(\text{syst.})\,\mu\text{b} \\ \sigma(p\bar{p} \to bX,\, p_T(B) > 4\text{GeV/c},\, |\eta| < 0.6) &= 11.9 \pm 0.6\,(\text{stat.}) \pm 1.0\,(\text{syst.})\,\mu\text{b}\end{aligned} \quad (4)$$

La section efficace pour $p_T(B) > 0$ a été estimée comme:

$$\begin{aligned}\sigma(p\bar{p} \to bX,\, |\eta| < 1.0) &= 27.5 \pm 1.2\,(stat.) \pm 2.3\,(syst.)\,\mu b \\ \sigma(p\bar{p} \to bX,\, |\eta| < 0.6) &= 16.9 \pm 0.8\,(stat.) \pm 1.4\,(syst.)\,\mu b\end{aligned} \quad (5)$$

La Figure 8 montre la section efficace différentielle de production de quark *b*.

[1] $\sigma(p\bar{p} \to bX) = \sigma(p\bar{p} \to BX) + \sigma(p\bar{p} \to \bar{B}X)$

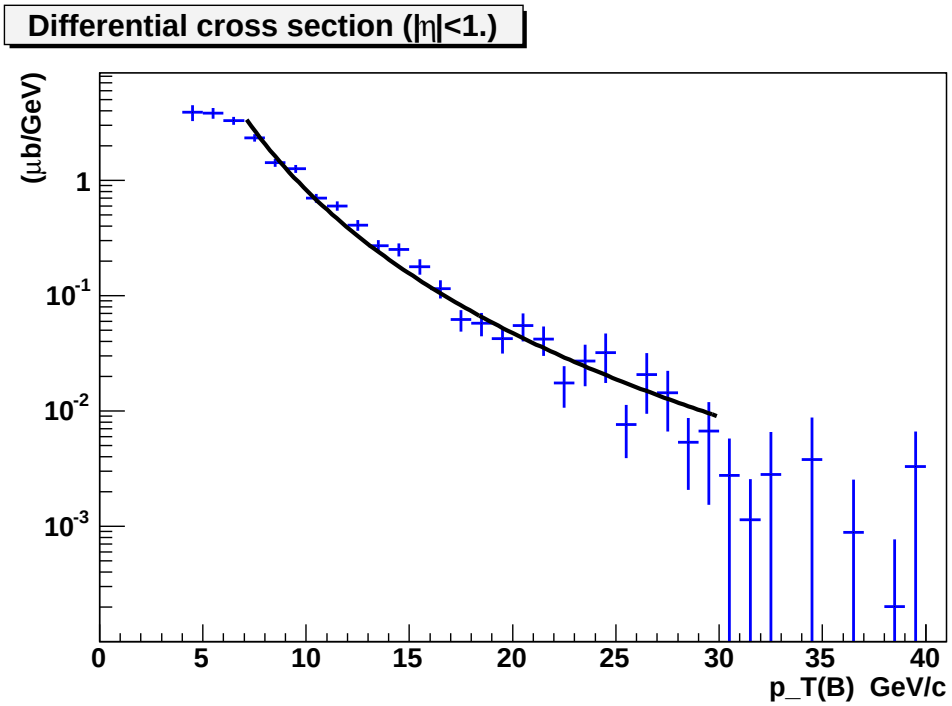

Figure 8: *La section efficace différentielle de production du quark b avec pseudo-rapidité inférieure à 1.*

Annexe : chapitres en Anglais

Abstract

The *b* quark fragmentation distribution has been measured, using data registered by the DELPHI experiment at the Z pole, in the years 1994-1995. The measurement made use of 176000 inclusively reconstructed *B* meson candidates. The uncertainties of this measurement are dominated by systematic effects, the principal ones being related to the energy calibration. The distribution has been established in a nine bin histogram. Its mean value has been found to be $< x_E >= 0.704 \pm 0.001\,(stat.) \pm 0.008\,(syst.)$

Using this measurement, and other available analyses of the *b*-quark fragmentation distribution in e^+e^- collisions, the non-perturbative QCD component of the distribution has been extracted independently of any hadronic physics modeling. This distribution depends only on the way the perturbative QCD component has been defined. When the perturbative QCD component is taken from a parton shower Monte-Carlo, the non-perturbative QCD component is rather similar with those obtained from the Lund or Bowler models. When the perturbative QCD component is the result of an analytic NLL computation, the non-perturbative QCD component has to be extended in a non-physical region and thus cannot be described by any hadronic modeling. In the two examples, used to characterize these two situations, which are studied at present, it happens that the extracted non-perturbative QCD distribution has the same shape, being simply translated to higher-*x* values in the second approach, illustrating the ability of the analytic perturbative QCD approach to account for softer gluon radiation than with a parton shower generator.

Using all the available analyses of the *b*-quark fragmentation distribution in e^+e^- collisions, together with the result from DELPHI presented in this thesis, a combined world average *b* fragmentation distribution has been obtained. Its mean value has been found to be $< x_E >= 0.714 \pm 0.002$.

An analysis of the *B* hadron production at CDF is ongoing. It makes use of ~ 6000 $B^{\pm}$ candidates, from $333 pb^{-1}$ of data recorded by the CDF experiment, fully reconstructed in the decay channel $B^{\pm} \to J/\psi K^{\pm}$. Characteristics of *B* mesons and accompanying tracks have been examined in the perspective of understanding the effect of fragmentation. These studies, done in the framework of the PYTHIA event generator, also involve the contributions from different $b\bar{b}$ production mechanisms. Distributions from a fully reconstructed Monte Carlo sample have been compared to data, and the agreement has been found to be reasonable. The analysis is ongoing, and the goal is to fit the fragmentation function parameters and/or the relative contributions from different production mechanisms to improve the agreement between data and Monte Carlo.

A measurement of the *b* quark production cross section has been obtained using the same

data. The analysis is still under way, and therefore the result is preliminary. It has been found:

$$\begin{aligned}\sigma(p\overline{p} \to bX,\, p_T(B) > 4\text{GeV/c},\, |\eta| < 1.0) &= 19.0 \pm 0.8\,(\text{stat.}) \pm 1.6\,(\text{syst.})\,\mu\text{b} \\ \sigma(p\overline{p} \to bX,\, p_T(B) > 4\text{GeV/c},\, |\eta| < 0.6) &= 11.9 \pm 0.6\,(\text{stat.}) \pm 1.0\,(\text{syst.})\,\mu\text{b}\end{aligned} \tag{6}$$

Chapter 1

Introduction

The research work described in this thesis consists in providing a detailed study of the fragmentation of b-quarks in view of controlling the non-perturbative aspects of this mechanism.

Data registered at LEP, by the DELPHI collaboration, at the Z pole offer a clean environment to study the distribution of the fraction of the beam energy taken by B hadrons. This distribution is defined as the fragmentation function of b-quarks. With these data it is possible to separate the perturbative QCD processes, mainly gluon radiation, from the non-perturbative QCD part. In this purpose, the fragmentation function was measured in DELPHI, using a large sample of inclusively reconstructed B mesons.

The determination of the non-perturbative QCD component is important in the framework of the factorization theorem of QCD. The decoupling of low-Q^2 from high-Q^2 processes allows us to transport the non-perturbative distribution measured in e^+e^- collisions to other environments, such as the one prevailing at the $p\overline{p}$ collider. In Chapter 2, production characteristics of B-hadrons at the two colliders are explained.

Chapters 3 and 4 provide information on the two experimental environments: DELPHI at LEP and CDF at the TeVatron.

The measurement of the b quark fragmentation distribution in DELPHI is explained in Chapter 5.

From measured b quark fragmentation distributions, and using a perturbative QCD computation, theorists have evaluated moments of the Mellin-transformed non-perturbative QCD component. In the thesis, this approach has been extended, so as to provide directly the dependence of the non-perturbative component versus the energy fraction taken by the B hadron. In this form, such a function can be used in event generators without the need of any hadronic models. Distributions provided in this thesis are hadronic model independent. All these aspects are explained in Chapter 6.

The energy distribution of B hadrons as measured in CDF is considered in Chapter 7. $B^{\pm}$ hadrons decaying into $J/\psi K^{\pm}$ with the $J/\psi \to \mu^+\mu^-$ have been used in this purpose. The sensitivity of the distribution to different non-perturbative QCD fragmentation modeling is studied. At the $p\overline{p}$ collider, one has also to consider the importance of different perturbative QCD mechanisms which have been included in the event generator. Their various contributions have been evaluated. Finally, the effect of the hadronization modeling on the characteristics of the tracks acompanying the B hadron has been studied. Experimental control of these effects is of importance in the study of tagging capabilities for B^0B^0 oscillation measurements.

Chapter 2

Theory of Bottom Production, Fragmentation and Decay

This chapter will be dedicated to the theoretical aspects of the bottom quark production, fragmentation and decay, in the two experimental environments considered in this work: e^+e^- and $p\overline{p}$ colliders. A special attention will be given to the theory of b fragmentation.

2.1 Overview: The Life Story of a Bottom Quark

A scheme of the process that a b quark goes through, from its production until the B hadron decays to the final states particles that hit the detector, is shown in Figure 2.1. The first stage is the hard scattering process, where the b and $\overline{b}$ quarks are produced. In Figure 2.1 this process is taken as the e^+e^- annihilation to an intermediate vector boson γ/Z^0 that finally decays to a $b\overline{b}$ pair. Initial state electrons may radiate virtual photons before annihilating. This initial state radiation (ISR) reduces the energy available for the hard process. In a hadronic collider the actors of the hard scattering change to quarks and gluons.

Once the b and $\overline{b}$ quarks are produced, they radiate secondary gluons by strong interaction. This process is described by QCD and can be calculated perturbatively, due to the high virtuality scale $Q^2 \gg {\Lambda_{QCD}}^2$ that implies $\alpha_S \ll 1$. In the following, this part of the process will be referred to as the *perturbative QCD part of the fragmentation*. It is calculable either by theoretical QCD computations or using Monte Carlo generators.

When the b and $\overline{b}$ quarks separate, the energy scale diminishes and the color interaction between the two quarks becomes stronger. When the energy scale becomes $\sim {\Lambda_{QCD}}^2$ the process enters the non-perturbative regime. Through the self interaction of the radiated gluons, a region of higher and higher energy density is created between the b and the $\overline{b}$. At some point, the increasing potential energy is sufficient to create from the vacuum another $q\overline{q}$ pair despite the penalty of providing the extra $q\overline{q}$ mass. The outgoing quark and antiquark continue on their way, and the process repeats until the system creates clusters of quarks and gluons, having zero net color and low internal momentum. Consequently, the color coupling between partons inside the clusters is very strong and turns them into hadrons. This part of the process, that includes the hadronization phase will be referred to as the *non-perturbative QCD part of the*

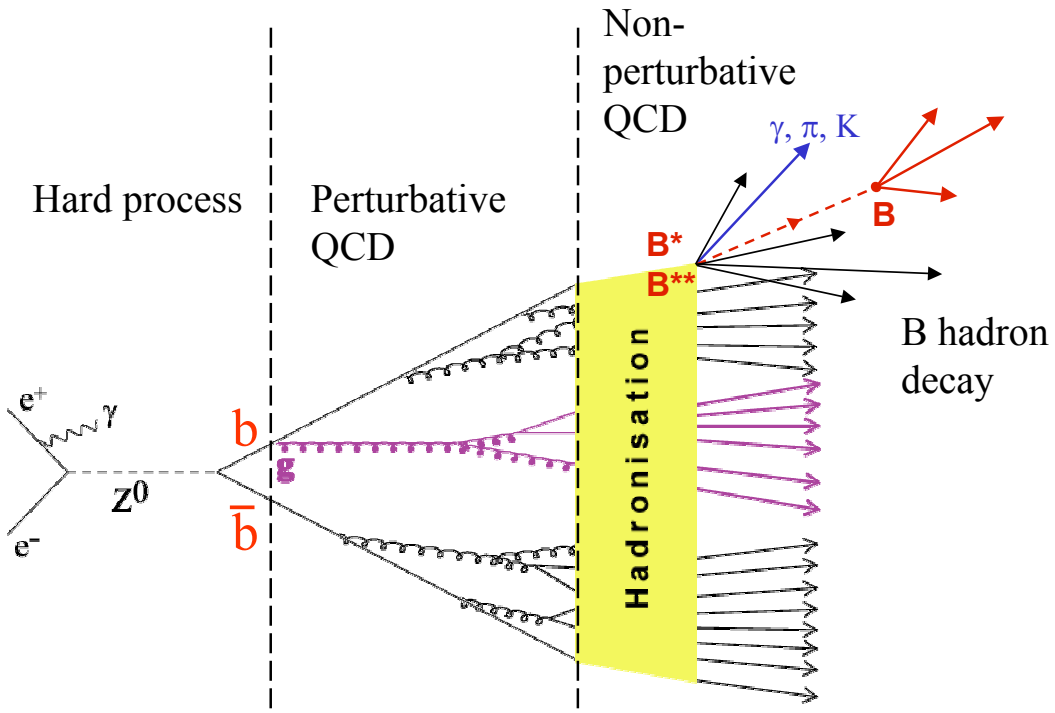

Figure 2.1: *A scheme of the b production, fragmentation and decay.*

fragmentation [1]. The non-perturbative part, as simply indicated by its name, is not calculable. It is usually described by phenomenological hadronization models. In this work, a method is proposed to extract the behavior of the non-perturbative part directly from data independently of any model assumption.

The products of the hadronization process form two jets of particles traveling more or less in the direction of the b and $\bar{b}$ quarks that originally came out of the hard process. Among the particles in the jets there are two B hadrons that contain the original b quarks. When produced at the Z^0 pole, those B hadrons carry $\sim 70\%$ of the original quark's energy. The remaining energy is distributed among the other particles in the jet.

An important difference between light and heavy [2] quark fragmentation is that, as mentioned above, in the second case one identifies in the final state's hadrons the heavy quarks that have been produced in the hard process. The ability to follow those quarks from their production until they decay provides a tool to probe the physics processes that intervene all along the way. This is an essential motivation of b quark fragmentation studies.

The experimental analyses presented in this work are based on B hadrons decaying by weak

[1] The terms of *fragmentation* and *hadronization* are sometimes mixed up. The "non-perturbative component" in the terminology adapted in this work is sometimes referred to as the "fragmentation function" in the literature.

[2] For a quark to be heavy, its mass has to be larger than Λ_{QCD} . Nevertheless, the top quark is too heavy and decays weakly before hadronizing. Therefore in the context of fragmentation, bottom and charm can be considered as heavy quarks.

interaction. Nevertheless, most of the B hadrons outgoing from the hadronization process are excited states (B^* and B^{**} mesons) that decay by strong or electromagnetic interaction to weakly decaying B hadrons [3] and other particles like pions, kaons and photons. Due to the short time scales of the strong and electromagnetic interactions, the decay of the excited states occurs in a place that is not experimentally distinguishable from the primary vertex, where the hard process took place. The decay products of the excited B hadrons are therefore difficult to discriminate from the particles created in the fragmentation process.

Contrarily to the excited states, the weakly decaying B hadrons travel along a measurable distance before they decay [4]. This distance ($\sim 3mm$ at the Z^0 pole) is due to the lifetime of the weakly decaying hadrons, that is $\sim 1.6ps$, and to the hard fragmentation of the b quarks.

Since B decays have a mean charged multiplicity of ~ 5, it is possible to tag B hadrons using a lifetime tag, based on accurate measurements of charged particle trajectories in the vicinity of the beam interaction region.

2.2 Bottom Quark Production in The Hard Process

2.2.1 Bottom Quark Production at LEP

In electron positron collisions, b quarks are produced from e^+e^- annihilation to a Z^0 or a virtual photon γ^*, with a subsequent decay of the intermediate state into a $b\bar{b}$ pair. The corresponding diagrams, at tree level, are presented in figure 2.2. In this experimental environment the initial state is very well defined and therefore it is an excellent laboratory for QCD studies. The fact that in this type of collider the beam's energy is a good approximation of the b quark's energy, makes e^+e^- machines appropriate for fragmentation studies.

Using the electroweak Lagrangian of the standard model, at the tree level (Born approximation), the cross section is found to be [1]:

$$\sigma(e^+e^- \to \gamma, Z^0 \to f\bar{f}) = N_c \frac{4\pi\alpha^2}{3s} \times$$
$$\left[\underbrace{q_e^2 q_f^2}_{\gamma \text{ exchange}} + \underbrace{(v_e^2 + a_e^2)(v_f^2 + a_f^2)|\chi|}_{Z^0 \text{ exchange}} + \underbrace{2 q_e q_f v_e v_f \Re(\chi)}_{\text{interference term}} \right] \tag{2.1}$$

In this expression N_c is the number of colors for a respective fermion (1 for leptons, 3 for quarks). $\sqrt{s}$ is center of mass energy. q_e and q_f ate the electric charges in natural units of the electron and the final state fermion, respectively. v_e and a_e (v_f and a_f) are the weak vector and axial vector couplings of the electron (final state fermion). χ is the Breit-Wigner parametrization of the Z^0 resonance:

$$\chi(s) = \frac{1}{4\sin^2\theta_w \cos^2\theta_w} \times \frac{s}{s - M_Z{}^2 + iM_Z\Gamma_Z} \tag{2.2}$$

M_Z and Γ_Z are the mass and the width of the Z^0 boson. θ_w is the weak mixing angle $sin^2\theta_w = 1 - \frac{M_W{}^2}{M_Z{}^2}$, where M_W is the mass of the $W^\pm$ bosons.

[3] Namely, $B^\pm$, B_d^0 and B_s^0 mesons as well as b-baryons.

[4] The flight distance of a particle is defined as $L = \gamma\beta c\tau$, where τ is its proper lifetime.

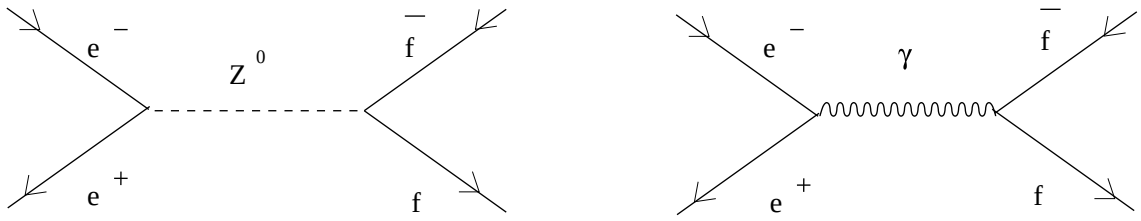

Figure 2.2: *Fermion pair production at LEP.*

As mentioned above, Equation (2.1) gives the cross section at tree level. In order to account for radiative processes, weak, electromagnetic and QCD corrections should be included. In particular, the symmetric shape of the Breit-Wigner parametrization of the Z^0 resonance given in Equation (2.2) is distorted by the initial state radiation. When photons are emitted from the initial state particles, the effective center of mass energy diminishes. Because of the strong energy dependence of the annihilation cross section around the Z^0 resonance, the cross section for initial state radiation is strongly enhanced above the Z^0 resonance, but suppressed on or just below M_Z. The part of this thesis that concerns the DELPHI analysis is based on data registered at the Z^0 pole, where the center of mass energy is $\sqrt{s} \approx M_Z$. Under these conditions, the contribution of the initial state radiation is small, and therefore the beam energy can be taken as a good approximation of the *b* quark's energy right after the hard process and before it radiates any gluons. When $\sqrt{s} = M_Z$ the interference term of Equation (2.1) vanishes and the γ exchange term becomes negligible comparing to the Z^0 exchange term. At the Z^0 pole the *b* production cross section is of $\sim$ 6 nb. Hadronic events account for about 70 % of the total production rate, and among those, the fraction of $b\bar{b}$ events is $\sim$ 22% [5]. This is in fact one of the advantages of the Z^0 region in B physics studies: B hadrons are copiously produced [6].

2.2.2 Bottom Quark Production at the TeVatron

Unlike in e^+e^- collisions near the Z^0 pole, in a hadronic collider the *b* quark's energy is unknown and varies from one event to another. This makes a crucial difference between fragmentation analyses in the two different experimental environments.

The dominant source of *b* quarks at the TeVatron is QCD. Contributions to the total *b* quark cross section from electroweak processes as $W^+ \to c\bar{b}$ or $Z \to b\bar{b}$ are small, and can generally be neglected. The main production mechanism is the gluon-gluon fusion process $g + g \to b + \bar{b}$. A more detailed description of bottom quark QCD production mechanisms is given later in this section.

The $b\bar{b}$ production cross section can be predicted using perturbative QCD. Theoretical QCD calculations exist at the next-to-leading-order (NLO) accuracy, and account for all terms to a fixed order α_S^3 [2, 3]. Other computations perform the resummation of the logarithms of p_T/m_b, with next-to-leading logarithmic accuracy (NLL), and the matching with the fixed-order, exact

[5] The fraction of hadronic events in which a $b\bar{b}$ pair is produced from gluon splitting was measured to be $\sim 3\,10^{-3}$. The contribution from this process to the *b* production near the Z^0 pole is therefore negligible.

[6] In the intermediate energy region ("continuum"), where the γ exchange term of Equation (2.1) is dominant, the cross section scales with the energy available in the center of mass (squared), being of the order of 30 pb at 30 GeV and of about 10 pb at 60 GeV. In this energy range the fraction of bb events is $\sim$ 9%.

NLO calculation for massive quarks (FONLL) [4]. In general, these calculations are effected by an uncertainty of up to $20-30\%$ due to the choice of renormalization and factorization scales (μ_R, μ_F) and by additional uncertainties due to the choice of parton distribution functions (PDFs) and of the value of the *b* quark mass.

In the past few years these QCD predictions were thought to underestimate the measured bottom quark production cross section by a factor that laid between 2 and 3. Recent theoretical studies [5, 6] accompanied by new available experimental measurements [7, 8] show that the discrepancy is much smaller, and that in fact it is not significant. A part of the previously presumed discrepancy disappears by taking into account correctly the *b* fragmentation function measured by the e^+e^- experiments.

Another way to predict the $b\bar{b}$ production cross section is to use the parton shower approach, implemented in Monte Carlo generators such as PYTHIA and HERWIG. This model is not exact to any order in perturbation theory. It is based on leading-order matrix elements and incorporates higher order effects by using a probabilistic model for initial and final state radiation. This probabilistic approach captures the leading-log characteristics of multiple-parton emission. Parton shower Monte Carlo generators will be discussed in Section 2.3.3.2.

There are three processes contributing in QCD $b\bar{b}$ production, at the lowest order of diagrams [7]:

- **Flavor Creation** (also known as direct production). In this category, either a quark and an antiquark or two gluons from the beam particles interact to give a $b\bar{b}$ pair in the final state. Both *b* quarks are outcoming from the hard scattering.

- **Flavor Excitation** In this case a *b* quark from the sea of one of the beam particles is scattered through a hard interaction with a parton from the other beam particle. In this mechanism, only one *b* quark participates in the hard scattering.

- **Gluon Splitting**, This is the case where the $b\bar{b}$ pair is created from a gluon after the hard scattering, as a part of the fragmentation process of the event. In this category none of the *b* quarks participate in the hard scattering.

Figure 2.3 shows some of the lowest order Feynman diagrams characteristic of each of these three categories.

When using a Monte Carlo generator, one can classify the bottom production mechanism in the event in a more detailed way. This classification scheme, that is mentioned in [9] [8], is based on the identification of the incoming and outgoing partons in the hardest 2-to-2 parton scattering, and on the number of *b* quarks in the final state:

0. *Non*-$b\bar{b}$ *event*: These events do not have any *b* quarks in the final state, regardless of the details of the hard scattering.

1. *Gluon fusion* ($g+g \rightarrow b+\bar{b}$): In these events, two gluons participate in the hard scattering to create a $b\bar{b}$ pair.

[7] One can only classify events in that way in the context of Monte Carlo generators, where the diagrams involved in the process are completely determined. In reality when interference effects take place, this classification is not needed. These processes are naturally a part of a NLO calculation, without any need to be treated separately.

[8] This scheme was originally suggested by R. D. Field.

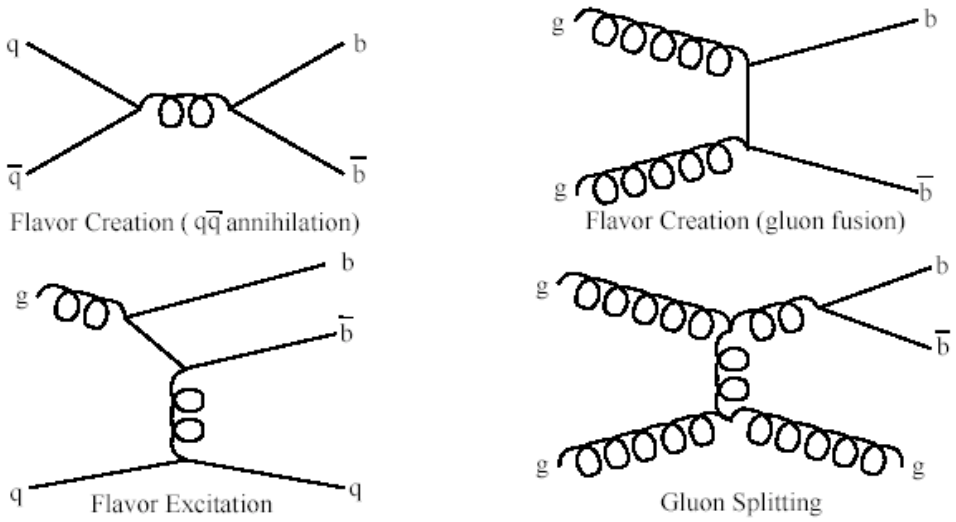

Figure 2.3: *The lowest order contributions to b production in the TeVatron.*

2. *Quark Annihilation* ($q+\overline{q} \rightarrow b+\overline{b}$): In these events, a quark and an antiquark annihilate in the hard scattering to produce a $b\overline{b}$ pair. This process includes $b+\overline{b} \rightarrow b+\overline{b}$.

3. *Gluon-Initiated Flavor Excitation* ($g+b \rightarrow g+b$ or $g+\overline{b} \rightarrow g+\overline{b}$): In these events, the hardest scattering involves a gluon and a *b* quark in the initial state going to a gluon and a *b* quark in the final state. To preserve bottom quantum number, there is a second *b* quark in the event of the opposite flavor as part of the additional activity in the event.

4. *Quark-Initiated Flavor Excitation* ($q+b \rightarrow q+b$, $\overline{q}+b \rightarrow \overline{q}+b$, $q+\overline{b} \rightarrow q+\overline{b}$, $\overline{q}+\overline{b} \rightarrow \overline{q}+\overline{b}$, where q is any quark except b). The hardest parton scattering in these events contains one non-*b* quark and one *b* quark in both the initial state and the final state. Like the process above, there is a second *b* quark generated by the rest of the activity in the event.

5. *Bottom-Quark-Initiated Flavor Excitation* ($b+b \rightarrow b+b$ or $\overline{b}+\overline{b} \rightarrow \overline{b}+\overline{b}$). In these events, both incoming and outgoing partons in the hardest scattering are *b* quarks of the same flavor. The opposite flavor case is part of quark annihilation above. In this class of event, there will always be four *b* quarks in the final state: two from the hard scattering, and two from the rest of the activity in the event.

6. *Gluon Splitting* (two gluon final state): The *b* quarks in this class of events do not come

from the hard scattering. The two outgoing partons from the hardest scattering in this event are both gluons.

7. *Gluon Splitting* (one gluon final state): The b quarks in this class of events do not come from the hard scattering. One of the two outgoing partons from the hard scattering is a gluon. The other is a quark of some flavor other than bottom.

8. *Gluon Splitting* (no gluon final state): In this class of events, the b quark does not participate in the hardest scattering in the event. Neither of the outgoing partons from the hardest scattering are gluons.

9. *Error*: This category is reserved for situations that should not arise from normal QCD processes, like events with only one b quark or events containing two b quarks but no $\bar{b}$ quarks. These events should only result from bugs in the generator program or generation of non-QCD events.

Events from classes *1* and *2* correspond to *flavor creation*. Classes *3*, *4*, and *5* combine to yield *flavor excitation*, while *gluon splitting* is made up from events from categories *6*, *7*, and *8*.

A study of the contribution to the b production from each one of this mechanisms will be shown in Chapter 7

The TeVatron RunIIa is currently the most copious source of B hadrons due to the large $b\bar{b}$ production cross section. At the center of mass energy of $\sqrt{s}$ = 1.96 TeV and selecting events in the central region, this number is reported to be [7]:

$$\sigma(p\bar{p} \to \bar{b}X, |y_b| < 1) = 29.4\ ^{+6.2}_{-5.4}\ _{stat+syst}\ \mu\text{b}\ , \tag{2.3}$$

where $|y_b|$ is the b quark's rapidity. The differential b-production cross section depends on the rapidity and on the transverse momentum.

2.3 Theoretical Aspects of b Fragmentation

The b quark fragmentation function is the probability density distribution for the fraction of the b quark energy taken by the B hadron.

In this section the theoretical aspects of fragmentation function will be reviewed.

2.3.1 Definitions of Fragmentation Functions

In this section, all the variables with the subscript b refer to the b quark and those with the subscript B to the B hadron. E stands for the energy, p for the three-momentum and m for the mass.

The b fragmentation function is a probability density function of a variable x:

$$\mathcal{D}(x) \equiv \frac{1}{\sigma_{tot}} \frac{d\sigma}{dx} \tag{2.4}$$

In general, x relates the momentum-energy of the b quark and of the B hadron. In the following measurements, only weakly decaying B hadrons have been considered [9]. There are a few common choices for this variable x. One of them:

$$x_E = \frac{E_B}{E_b} \tag{2.5}$$

is the energy taken by the weakly decaying b hadron with respect to the energy of the b quark right after it is produced in the hard process, before it radiated any gluons. This variable is appropriate to e^+e^- collisions, because both the nominator and the denominator are experimentally observable. In particular, as explained in 2.2.1 the denominator is simply the beam energy. Therefore:

$$x_E = \frac{2E_B}{\sqrt{s}} = \frac{E_B}{E_{beam}} \tag{2.6}$$

Another definition uses the ratio between three momenta, assuming that $m_B = m_b$:

$$x_p = \frac{p_B}{p_b} = \frac{\sqrt{x_E^2 - x_{min}^2}}{\sqrt{1 - x_{min}^2}} \tag{2.7}$$

where $x_{min} = \frac{2m_B}{\sqrt{s}}$ is the minimal value of x_E, due to the mass of the B hadron. At a center of mass energy of 91.2GeV this value is $x_{min} \cong 0.12$. The minimum value of x_p is 0.

Another possible choice for the fragmentation variable is z, sometimes called the light-cone momentum fraction, defined as:

$$z = \frac{(E + p_L)_B}{(E + p)_b} \tag{2.8}$$

Where p_L is the momentum of the formed hadron along the direction of the quark. This variable is invariant against Lorentz boosts along the jet direction.

z is commonly used in Monte Carlo generators. In this framework the "quark" in the definition of z is understood to be the parton which has already radiated gluons, just before the generator routines that create a B hadron are called. When defined in this way, z describes only the non-perturbative part of fragmentation, it is not an observable quantity, and has a meaning only in the context of a parton shower Monte Carlo generator. Besides, in the Monte Carlo generators which use the string model for hadronization (see Section 2.3.4.1), this variable is defined in the reference system of a string stretched between the b quark and a gluon, an antiquark or a diquark.

2.3.2 Perturbative and Non-perturbative Parts

From the simplified description of the b fragmentation process in Section 2.1, one can understand that the frontier between the perturbative and non-perturbative parts is somehow arbitrary. The non-perturbative component describes the hadronization, but one has to be aware of the fact that it also includes all the terms that have not been taken into account in the perturbative component.

[9] This variable may be also defined for the leading B hadron, i.e the one directly created by the fragmentation process, that is the most often an excited B state. This is not the definition adopted in this work.

The fact that the non-perturbative component of the fragmentation function depends strongly on the perturbative one has been sometimes neglected.

The possibility to separate between the perturbative and non-perturbative parts, while still missing the *Factorization theorem* of QCD [10], is based on the fact that dynamical processes taking place on well separated physical scales are quantum-mechanically incoherent. This allows us to treat different subprocesses independently of each other. Applying factorization to b fragmentation allows us to fold the perturbative and non-perturbative components using the appropriate convolution product to get a prediction for the observed fragmentation function [10]:

$$\mathcal{D}_{observed}(x) = \int_0^\infty \mathcal{D}_{pert.}(z) \times \mathcal{D}_{non-pert.}(\frac{x}{z})\ \frac{dz}{z} \tag{2.9}$$

where x is the chosen fragmentation variable.

A consequence of the factorization theorem is that the non-perturbative component, that describes the physics at a scale lower than the "factorization scale" μ_F, does not depend on the initial state.

A clear demonstration of the dependence of the non-perturbative QCD component on the perturbative one is presented in Chapter 6, where the non-perturbative QCD component is extracted from the measured fragmentation function for a few perturbative QCD approaches.

2.3.3 Perturbative QCD

Two approaches to predict the perturbative QCD component of the fragmentation function are considered in the next sections. The first one is based on theoretical QCD calculations, and the other uses Monte Carlo generators. The second approach is usually adopted by experimentalists.

2.3.3.1 Theoretical QCD Calculations

The full expression of the perturbative QCD calculation at order α_S is:

$$\begin{aligned}\tfrac{1}{\sigma}\tfrac{d\sigma}{dx} = \delta(1-x) + \tfrac{\alpha_S(Q^2)}{2\pi}\Bigg\{ C_F + C_F\Bigg[\ln\tfrac{Q^2}{m^2}\left(\tfrac{1+x^2}{1-x}\right)_+ \\ + 2\tfrac{1+x^2}{1-x}\ln x - \left(\tfrac{\ln(1-x)}{1-x}\right)_+ (1+x^2) + \tfrac{1}{2}\left(\tfrac{1}{1-x}\right)_+ (x^2-6x-2) \\ + \left(\tfrac{2}{3}\pi^2 - \tfrac{5}{2}\right)\delta(1-x)\Bigg]\Bigg\} + O\left(\tfrac{m}{Q}\right)\end{aligned} \tag{2.10}$$

where Q is the center of mass energy, m the quark's mass and $C_F = 4/3$ the color factor. The "plus" distribution, $\left(\frac{1}{1-x}\right)_+$ is defined so that its integral with any sufficiently smooth distribution f is:

$$\int_0^1 dx \frac{f(x)}{(1-x)|_+} = \int_0^1 dx \frac{f(x)-f(1)}{1-x} \tag{2.11}$$

[10] The physical distributions vanish for $x > 1$. The necessity to integrate to infinity will be clarified in Chapter 6

and

$$\frac{1}{(1-x)|_+} = \frac{1}{1-x} \tag{2.12}$$

for $0 \leq x < 1$.

In Equation (2.10), terms containing $\ln \frac{Q^2}{m^2}$ and $\frac{1}{1-x}$ are large. They are related to collinear and soft gluon radiation, respectively.

The perturbative QCD expansion of the fragmentation as a power series in α_S contains these large logarithmic terms, due to soft and collinear gluons, at all orders in α_S . One class of such terms is:

$$T_{n,m} = \alpha_S^n \ln^m(1-x)/(1-x) \tag{2.13}$$

with $m \leq 2n-1$. The contribution from soft gluons is particularly problematic in the so called Sudakov region, where $x \to 1$. According to the Bloch-Nordsieck theorem [10], the singularities from soft gluons cancel between real and virtual diagrams, as shown in Figure 2.4. When $x \to 1$

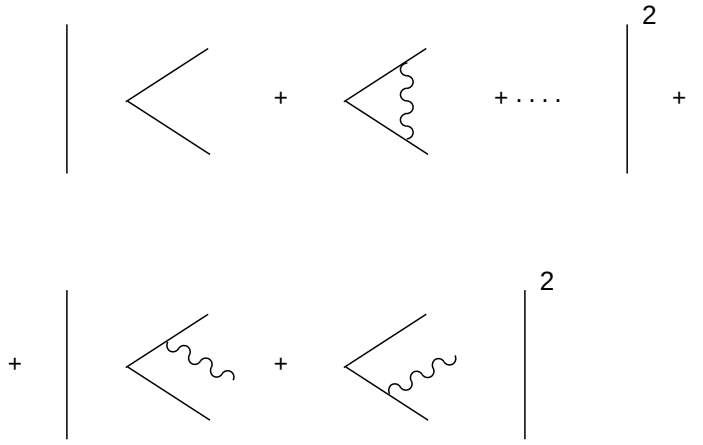

Figure 2.4: *Cancellation of real gluon radiation and virtual gluon exchange.*

and the phase-space for gluon radiation is suppressed, this cancellation of real gluon radiation and virtual gluon exchange is incomplete and the theoretical calculation breaks down. Under these conditions, any finite order perturbative calculation can-not provide reliable predictions. A resummation of the large logarithmic terms to all orders in α_S of the perturbative expansion is then necessary and solves partially that problem. The higher order resummation of Logarithmic terms is evaluated, the more reliable the computation is, and the closer it goes to $x = 1$ before breaking down. A LL (Leading Log) resummation takes into account the class of largest logarithmic contribution at all orders in α_S (i.e $\sum_{n=1}^{\infty} T_{n,2n-1}$, where $T_{n,m}$ is defined in Equation (2.13)). A Next to Leading Log (NLL) resummation, sums up also the next class of logarithms, $\sum_{n=1}^{\infty} T_{n,2n-2}$ and so on.

A fixed-order calculation of the fragmentation function at order α_S^2 in e^+e^- annihilation was done in [11]. This result does not include the resummation of logarithmic classes as described above, but does include correctly all terms up to order α_S^2, including terms without any logarithmic enhancement. A LL resummation formula has been obtained in [12] and a NLL resummation has been performed in [13, 14]. An approach that also retains NLL accuracy terms, the Dressed Gluon Exponentiation (DGE), has been presented in [15].

In the case of heavy quark fragmentation, the large mass of the quark regulates collinear singularities and suppresses one class of logarithmic divergences. Yet, the logarithmic contributions $\alpha_S ln(Q^2/m^2)$ of collinear origin become large when the center of mass energy Q is much larger than the heavy quark's mass m. This is the case at LEP energies, and therefore also these terms need to be resummed to all orders in α_S. This resummation has been performed in [12] up to NLL level, using Altarelli-Parisi evolution of a perturbatively calculable initial condition [11]. Unlike terms related to soft gluon radiation, the resummation of terms in $\alpha_S ln(Q^2/m^2)$ is important in all regions of the x spectrum.

In Chapter 6 the extraction of the x-dependence of the non-perturbative QCD component of the fragmentation function is presented. The extraction has been done for a few perturbative QCD approaches. In this context, we will mainly refer to the NLL computation from [14]. Moreover, the present thesis work started by carrying out this QCD calculation step by step. The results were compared with those of the authors of [14] and found to be identical. The detailed computation, beyond its main results that are presented in [14], was needed for the analysis presented in Chapter 6.

The computation in [14] resums to all orders leading and next-to-leading logarithmic terms in the perturbative expansion at all orders in α_S [12].

The Next to Leading Log with Dressed Gluon Exponentiation (NLL+DGE) computation of [15] adds two ingredients to the former:

- It also resums in an approximate way *all* the subleading logarithmic terms, i.e. $\alpha_S^n log^{n-1}(N)$, $\alpha_S^n log^{n-2}(N)$,... etc. This approximation, known as the large- β_0 approximation, includes only the logarithms generated by the fermion loops in an emitted gluon. These loops can be calculated by using a "massive gluon" as a computational tool.

- The series generated by the above procedure is divergent, like all infinite series in QCD. Therefore a regularization has to be introduced. The functional form of the uncertainty related to the regularization comes in the form of power corrections, with a specific pattern. Hence, by regularizing the perturbative series one can gather some pieces of information about the behavior of the power corrections.

2.3.3.2 Parton Showers in Monte Carlo Generators

The useful and easily understandable "parton shower" picture for the perturbative part of fragmentation is appropriate for integration in computer Monte Carlo generators. In this model, partons are created by successive elementary branching processes $q \to qg$, $g \to gg$, and $g \to q\bar{q}$, where interference between the individual branchings is neglected. It is shown here that this probabilistic approach captures the leading-log characteristics of multiple-parton emission.

The cross section of the perturbative part of the fragmentation process, that is, the evolution from an initial parton to the final multi-parton system, may be calculated in the *Leading Logarithm* (LL) Approximation, where only the leading logarithmic terms at all orders in α_S are

[11] This initial condition is the fragmentation function at the scale of the heavy quark mass

[12] In effect this is done in the moments space by resuming the logarithmic terms of type $\alpha_S^n log^{n+1}(N)$ and $\alpha_S^n log^n(N)$ in the exponent

kept in the perturbative expansion:

$$\sigma \cong \sum_n a_n \left(\alpha_S \ln \frac{Q^2}{{\Lambda_{QCD}}^2} \right)^n \tag{2.14}$$

The expression for α_S is:

$$\alpha_S \left(Q^2\right) = \frac{1}{\beta_0 \ln \left(\frac{Q^2}{\Lambda_{QCD}} \right)} \tag{2.15}$$

For the LL approach to be valid, the sum in Equation (2.14) must dominate the asymptotic behavior of the cross section for a large virtuality scale Q^2. This requires that $\alpha_S(Q^2)$ be much smaller than unity and that $\alpha_S(Q^2) ln(Q^2/\Lambda_{QCD}) \approx 1$. Both requirements are fulfilled according to Equation (2.15).

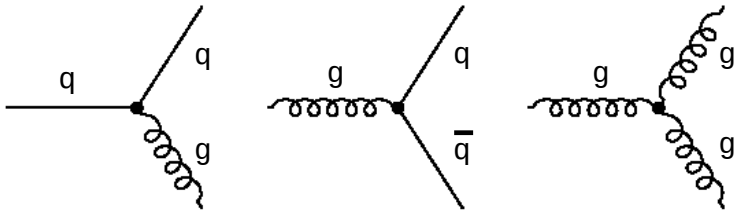

Figure 2.5: *Elementary processes contributing to the parton shower model.*

The LL approximation corresponds to a selection of diagrams according to the topology of the subprocesses rather than the number of vertices, like a typical fixed order α_S calculation. Therefore it is well suited to describe final states with a large number of high energy partons. Gribov and Lipatov show [16] that the LL approximation suppresses diagrams with internal loops, and that only simple branchings of one parton into two partons, as in Figure 2.5, contribute to the leading logarithms. This means that there is no contribution from interference between individual branchings. The combination of these elementary steps and the lack of interference lead to tree-like diagrams like those in Figure 2.5.

This approach simplifies significantly the calculation and simulation of complex parton configurations. It allows a simple iterative modeling of event evolution in parton shower Monte Carlo generators, where the appearance of branchings to partons of specific types and energies can be described by classic probabilities. The branching probabilities are expected to fulfill some basic requirements. The probability should of course be proportional to $\alpha_S(Q^2)$. The remaining contributions to the branching probability should be scale-invariant, i.e. only depend on the types of the final state partons b and c in a process $a \rightarrow bc$, and on the fractions z and $(1-z)$ of the energy associated with partons b and c, $0 < z < 1$. This leads to the expression

$$Q^2 \frac{d^2 P_{a \rightarrow bc}}{dz dQ^2} = \frac{\alpha_S(Q^2)}{2\pi} P_{a \rightarrow bc}(z) \tag{2.16}$$

Here, $P_{a \rightarrow bc}(z)$ is the scale-independent parton splitting function for the process $a \rightarrow bc$. Integration of Equation (2.16) over the full z range leads to the Altarelli-Parisi type equation [17]

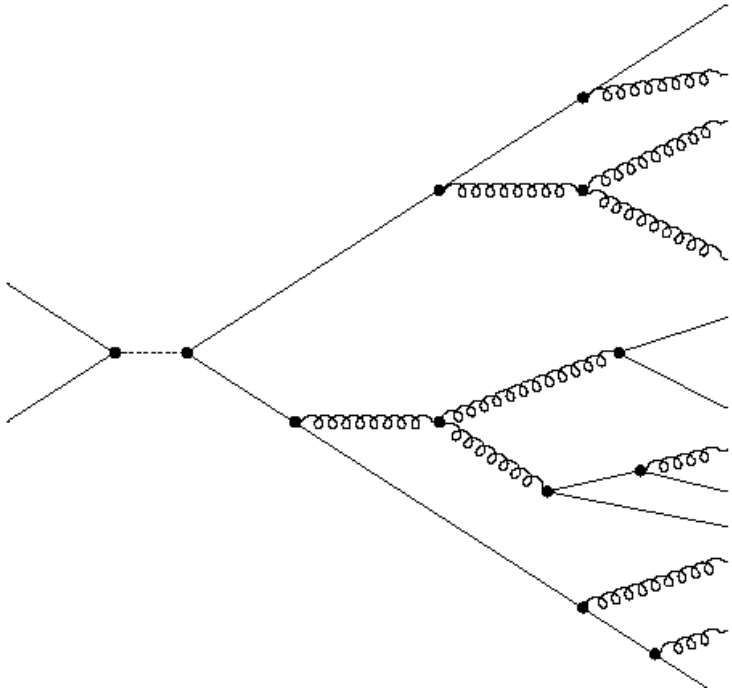

Figure 2.6: *A scheme of a parton shower process.*

$$Q^2 \frac{d^2 P_{a\to bc}}{dQ^2} = \frac{\alpha_S(Q^2)}{2\pi} \int P_{a\to bc}(z) dz \tag{2.17}$$

Parton splitting functions are obtained as solutions to the above equation. For QCD, the so-called Altarelli-Parisi splitting kernels [17] in the spin-averaged, lowest order approximation are

$$\begin{aligned} P_{q\to qg} &= C_F \left(\tfrac{1+z^2}{1-z}\right) \\ P_{g\to q\bar{q}} &= T_R \left(z^2 + (1-z)^2\right) \\ P_{g\to gg} &= 2C_A \left(\tfrac{1-z}{z} + \tfrac{z}{1-z} + z(1-z)\right) \end{aligned} \tag{2.18}$$

Where $C_F = 4/3$ and $C_A = 3$ are the color factors, and $T_R = n_f/2$, n_f being the number of flavors.

As expected, the kernels for symmetric final states are invariant under the exchange between z and $1-z$.

Parton shower Monte Carlo models are usually based on the probability that a given parton branches during a given decrease of the virtuality scale Q^2. This probability can be calculated by summing the splitting functions for all possible final states and integrating over the considered Q^2 range. Splitting is then applied randomly according to these probabilities. The parton shower model is only valid for large energy scales; once all partons have been tracked down to an energy Q_0 of the order of 1GeV, parton shower generators usually stop the iterative shower evolution and proceed to the non-perturbative hadronization of the shower partons. Q_0, also called the virtuality cutoff, and Λ_{QCD} are the main parameters of a parton shower model. This cut-off

limits the allowed range of z in splittings, so as to avoid the singular regions corresponding to excessive production of very soft gluons.

In analytical QCD computations, the kernals in Equation (2.18) contain terms in $\delta(1-z)$, and are used with the "plus" distribution $\frac{1}{(1-z)|_+}$. These terms ensure the flavor and energy conservation. The corresponding problem is solved trivially in Monte Carlo programs, where the shower evolution is traced in detail, and flavor and four-momentum are conserved at each branching. Indeed, this is reflected by the need to introduce the virtuality cut-off.

Parton shower models are very successful in modeling the distribution of secondary partons which are comparably soft and almost collinear to the primary partons, but their results for additional high energy partons with large transverse momentum (leading to three or more clearly separated jets) is not optimal.

Refinements to the simplest parton shower described in this section were introduced by taking interference between individual branching into account. It was established that the impact of interference can be described easily as suppression of certain regions of phase-space during branching, such that the opening angle of a branching is always smaller than the opening angle of the previous branching. This effect is taken into account in Monte Carlo generators as JETSET [18], PYTHIA [19], and HERWIG [20].

2.3.4 Non-perturbative QCD

The non-perturbative part of the fragmentation bridges the gap between the perturbative calculable part and the physical fragmentation function measured by an experiment. It is therefore an important ingredient for comparisons between the predictions of physical observables and their measured values. As has already been mentioned, this part of the fragmentation describes the physics at a low energy scale of $\sim \Lambda_{QCD}$, where the value of α_S is large.

During the non-perturbative part of the fragmentation, the production of hadrons from the generated partons takes place. Usually, phenomenological schemes are used to model the carry-over of parton momenta and flavor to the hadrons. It is also possible to extract the non-perturbative component directly, without using any model assumption. This is done by theorists for moments of the fragmentation distribution. In Chapter 6, a method will be proposed to perform this extraction also for the x-dependence of the fragmentation function.

2.3.4.1 Hadronization in Monte Carlo Generators

The computation of the hadronization process in Monte Carlo programs is not based on first principals of QCD, but on phenomenological models. The most popular models of hadronization are the string model, implemented in the JETSET [18] and PYTHIA [19] Monte Carlo event generators, and the cluster hadronization of the HERWIG Monte Carlo event generator [20]. The earliest model, the independent hadronization model is not used in this thesis but has great historical importance. The analysis presented in this work relies only on PYTHIA and JETSET. Particular attention will therefore be given to the string hadronization model in the discussion below. The hadronization models differ in how gluons created during the perturbative part are treated; how (if at all) different partons interact during the hadronization; and the number, type and momentum distribution of the created hadrons. The independent and string models use phenomenological parametrizations for the fragmentation variable. These

parametrizations, that are sometimes referred to as "fragmentation functions", will be reviewed in Section 2.3.4.2.

Independent Hadronization This is the simplest scheme for generating hadron distributions from those of partons.In this model, first proposed by Field and Feynman [21], partons are supposed to fragment independently from one another. Each quark in the system after the showering process is combined with an antiquark from a $q\bar{q}$ pair created out of the vacuum to give a "first generation" meson with energy fraction z of the initial quark. The leftover quark, with energy fraction $1-z$, is fragmented in the same way, over and over, until the leftover energy falls below a certain threshold. The variable z is distributed according to a given probability density function, which is the "hadronization model" (see Section 2.3.4.2). The most significant drawback of this model is the fact that it is not able to simulate interference effects. This model is still used to simulate events with two well separated jets, where interference factors are less important. The independent fragmentation can also be used by PYTHIA [19] as a non-default option.

Cluster Hadronization Cluster fragmentation is used in the HERWIG Monte Carlo generator [20]. In this model, assuming a local compensation of color, based on the pre-confinement property of perturbative QCD [22], the remaining gluons at the end of the parton shower evolution are split into quark-antiquark pairs. Color singlet clusters of typical mass of a couple of GeV are then formed from quark and antiquark of color-connected splittings. These clusters decay directly into two hadrons unless they are either too heavy (then they decay into two clusters) or too light, in which case a cluster decays into a single hadron, requiring a small rearrangement of energy and momentum with neighboring clusters. The decay of a cluster into two hadrons is assumed to be isotropic in the rest frame of the cluster unless a perturbative-formed quark is involved. A decay channel is chosen based on the phase-space probability, the density of states, and the spin degeneracy of the hadrons. Cluster fragmentation has a compact description with few parameters, due to the phase-space dominance in the hadron formation. A scheme of the cluster fragmentation is shown in Figure 2.7.

String Hadronization A large class of models describes hadronization through the image of color flux tubes, or strings, each spanned between two quarks created during the perturbative part of fragmentation. Usually the motion of a string is described by classical, relativistic dynamics. The treatment of gluons seems more natural in the string picture than in the cluster or independent hadronization schemes: perturbative gluons are simply incorporated into the strings connecting two nearby quarks. They appear as a kink in the string. This complicates the dynamics, but does not lead to fundamentally different effects. Gluons will therefore be neglected in the following considerations. A scheme of the string hadronization is shown in Figure 2.8.

The strength between two color sources, with large mutual distance, is expected to be independent of the distance due to the self-interaction of gluons. The energy density, or string tension κ, is constant along the string, assuming the transverse extension of the string is constant over its full length and much smaller than the longitudinal extension. The string is estimated to have a diameter of ~ 1fm and a tension of $\kappa \approx 1\text{GeV}/\text{fm}$. As the quarks fly apart and the string

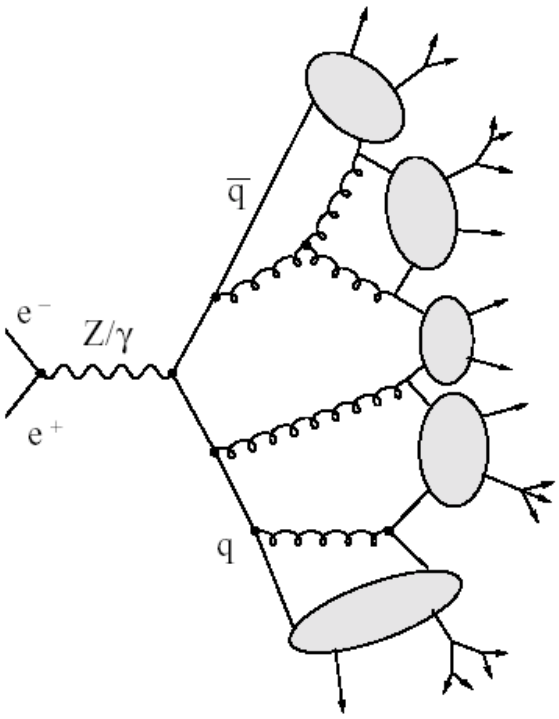

Figure 2.7: *A scheme of the cluster hadronization model.*

is stretched, the energy stored in the string increases according to $E(d) \propto \kappa d$, where d is the distance between the quark and the antiquark situated at both ends of the string, i.e. the length of the string. If a virtual $q\bar{q}$ pair fluctuates out of the vacuum, somewhere along the string, and if this pair has the same color as the endpoint quarks of the string, the color field is locally compensated and the string breaks into two pieces. This process is repeated until the remaining energy of the individual string segments is no longer sufficient to transform another virtual $q\bar{q}$ pair into a real one.

Because the string is assumed to be uniform along its length, the probability of a $q\bar{q}$ pair creation occurring on the string per unit length and per unit time is a constant P_0. To obtain the probability that the string breaks due to this pair creation at a given point, the history of this point has to be taken into account: the probability of a string break is proportional to the probability that no previous break occurred in its backward light cone. Had there been any previous break within the light cone, the considered space-time point would not be on the string, since the endpoints of the newly created sub-strings move apart at the speed of light [13] (assuming massless endpoint quarks). The probability for no string break in the backward light cone is given by the Wilson area law [23]

$$\frac{d\mathcal{P}}{dA} = P_0 e^{-P_0 A} \tag{2.19}$$

with A being the space-time area in the backward light cone. Essentially this area law states that evolution of long string pieces over long periods of time without decay is exponentially suppressed.

[13] This has the immediate consequence that all string breaks are causally disconnected.

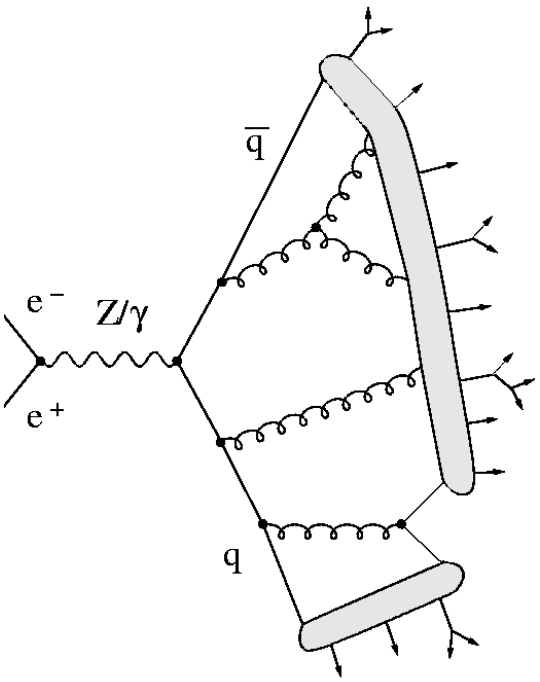

Figure 2.8: *A scheme of the string hadronization model.*

The area in the backward light-cone can be calculated in two-dimensional light-cone variables, in the center of mass system of quark and antiquark, with the string extended along the x axis:

$$A = \left| \left(\frac{\Delta t - \Delta x}{\sqrt{2}} \right) \left(\frac{\Delta t + \Delta x}{\sqrt{2}} \right) \right| = \frac{1}{2} \left| \Delta t^2 - \Delta x^2 \right| = \frac{1}{2\kappa^2} \left| p_x^2 - E^2 \right| = \frac{m_t^2}{2\kappa^2} \tag{2.20}$$

Thus the invariant transverse mass $m_t = \sqrt{m^2 + p_t^2}$ of a string piece is proportional to the area swept over by the backward light cone. Hence, the area law can be used to deduce the transverse mass distribution of sub-strings:

$$\frac{d\mathcal{P}}{dm_t^2} = \frac{d\mathcal{P}}{dA} \frac{dA}{dm_t^2} = P_0 e^{-P_0 A} \frac{1}{2\kappa^2} = b e^{-b m_t^2} , \quad b = \frac{P_0}{2\kappa^2} \tag{2.21}$$

It should be pointed out again that these results are only valid for massless quarks. Bowler [24] has performed a modification of this calculation which takes into account a non-zero mass of the endpoint quarks of a string. This is done by replacing the straight world line of the quarks by curved lines.

The first string model was proposed by Artru and Mennessier in 1974 [25], based on the formalism reproduced above. The problem encountered in this model was that iterative application of string splitting by the area law lead to an infinite series of ever smaller sub-strings. To avoid this problem, Artru and Mennessier introduced a mass cut-off m_0, below which further string fission is prohibited. They end up with a finite set of hadronic states with a continuous mass spectrum, which strongly resembles the clusters described above. These cluster-like resonances decay into two usual on-shell hadrons.

Rather than using a strict cut-off at m_0, the CalTech II string hadronization scheme [26] suppresses string splitting to unphysically small mass sub-strings by smoothly decreasing the probability of further string breaks. This probability is parametrized as

$$P_{break} = \begin{cases} 1 - e^{-P_0(m-m_0)^2/\kappa^2} & m > m_0 \\ 0 & m \leq m_0 \end{cases} \tag{2.22}$$

If, with the above probability, a further string break takes place, its space-time coordinates are found according to the area law. If no further string break takes place, the remaining sub-string decays into two hadrons as in the cluster scheme. If the cut occurs close to one end of the string, one of the two sub-strings can have an invariant mass below the kinematic threshold $m_1 < m_0$ for production of two hadrons. These very light sub-strings are transformed into a single hadron which is generally not on-shell. The momentum is then corrected by moving the string break along the string until the decay into a single hadron is kinematically allowed. Gottschalk and Morris have shown that the extent to which the string break point has to be moved is smaller than the transverse extension of the string [26], and that this procedure is therefore acceptable.

Hadron mass shell constraints and uniform splitting probability are irreconcilable, as evident from the above considerations. The Artru-Mennessier and CalTech II models rely on the uniform splitting probability and are thus forced to include cluster-like transitions from string fragments to hadrons.

Probably the most well-known and widely used string hadronization model, invented by the Lund group [27] and implemented in the Jetset [18] and Pythia [19] Monte Carlo event generators, abandons the concept of a necessarily constant $q\bar{q}$ pair creation probability over the full length of the string. Strings are broken exactly where necessary to ensure the creation of on-shell hadrons. Since all string breaks are causally disconnected, they can be treated in any possible order. The Lund scheme makes use of this property to establish a simple iterative algorithm: hadronization is started at the outer ends of the string, and the creation of hadrons is continued towards the inside until the invariant mass of the remaining string drops below a threshold value. This remainder is then split into two hadrons, once again, in analogy to a cluster decay.

The creation of $q\bar{q}$ pairs along the string in the Lund model takes the quark masses into account [14]. This has important consequences: $q\bar{q}$ pairs are expected to be created locally, i.e. at one space point. However, in order to transfer a virtual $q\bar{q}$ pair into mass-shell particles, energy has to be taken from the string. For quarks of a mass m_q, and a transverse momentum p_t with respect to the string, the minimum required energy is $E_{min} = 2\sqrt{m_q^2 + p_t^2} = 2m_t$. Here it is assumed that both members of the $q\bar{q}$ pair have a transverse momentum of the same size, but opposite direction. Otherwise the string would acquire a transverse excitation, which is assumed not to happen. The required energy can only be obtained by a spatial separation of quark and antiquark along the string of $\Delta x_{\min} = E_{\min}/\kappa = 2m_t/\kappa$. The string fission is thus suppressed by the probability of a tunneling process of a locally created $q\bar{q}$ pair to spatially separated on-shell quarks. The suppression factor was found to be [28]

$$e^{-\frac{\pi}{\kappa}\left(m_q^2+p_t^2\right)} = e^{-\frac{\pi}{\kappa}m_q^2} \times e^{-\frac{\pi}{\kappa}p_t^2} \tag{2.23}$$

[14] Unfortunately it is not quite clear which quark masses should be used. The Lund model uses so-called kinematical masses for hard processes like quark pair creation, and constituent masses fitted to the hadron spectra to estimate masses of resonances that have not yet been found.

The first term provides for an automatic suppression production during string hadronization. The ratio of flavor production rates is estimated to be $P_u : P_d : P_s : P_c = 1 : 1 : 0.3 : 10^{-11}$ using kinematical quark masses. The second term in the tunneling probability corresponds to a suppression of large transverse momenta. Interestingly, the transverse momentum distribution is predicted to be independent of the quark flavor. A further prediction is that baryon production (requiring two quarks and two anti-quarks) is less likely than meson production, in accordance with observations.

Choice of flavor and transverse momentum for $q\bar{q}$ pairs is thus strictly regulated in the Lund string scheme. Hence, the quark content and transverse momentum of hadrons are also well constrained. Further quantum numbers, e.g. angular momentum, relative spin orientation, etc., are chosen based on rather general symmetry assumptions, or according to measured production ratios. What remains is the longitudinal hadron momentum with respect to the string axis. The choice of a longitudinal momentum is usually expressed in terms of the fraction of the string's light cone momentum to be taken by the hadron, i.e. the variable z introduced in Section 2.3.1. The probability distribution $f(z)$ of momentum fractions assigned to hadrons can be chosen arbitrarily. Each choice corresponds to a specific distribution of cut points on the string, which does not necessarily comply with Wilson's area law, or a constant breaking probability over the full string length. The most common parametrizations of f(z) will be discussed later in this chapter.

A short summary of the Lund recipe amounts to the following: after the quark flavors for a hadron have been chosen, the next step is to select the remaining quantum numbers (e.g. spin, orbital angular momentum) for the hadron. Finally, the longitudinal momentum is determined by the fragmentation function. This function in the Lund model is independently normalized for each hadron species that is selected for the next hadronization step:

$$\int_0^1 f(z)dz = 1 \tag{2.24}$$

Since the hadron mass usually enters as a parameter in the fragmentation function, there is in fact a different fragmentation function for each type of hadron, $f(z) \rightarrow f_h(z)$.

Baryon Production A common problem arising in all models is the production of baryons. The primary partons are confined to hadrons by creating additional $q\bar{q}$ pairs and combining the additionally created (anti-)quarks to mesons. The formation of baryons is a more complicated process which is mainly modeled using two different scenarios. One approach is to create a certain amount of diquark-antidiquark pairs in the same manner as for quark-antiquark pairs [29]. These can then combine with additional single quarks to form hadrons. The other approach, the so-called popcorn model [30], features a step-by-step baryon production: the creation of $q\bar{q}$ pairs whose colors do not match their neighboring quarks keeps up a color field in which additional $q\bar{q}$ pairs can emerge before mesons are formed. Diquark production is applicable to all hadronization models described above. The popcorn model is mainly used in the string model.

2.3.4.2 Phenomenological Hadronization Models

There is a number of different phenomenological models that tend to parametrize the energy fraction taken by the b hadron with respect to the energy of the b quark. These models correspond to different functions of the fragmentation variable z and they usually depend on one or two parameters. The models are usually used to describe the non perturbative part of fragmentation, and therefore the quark's energy is understood as the energy *after the perturbative phase*. In the framework of Monte Carlo generators, the last statement is synonym to *after the parton showering process*. The use of the independent or string hadronization models requires the use of one of the phenomenological parametrizations described below. In particular, one must use them in the JETSET [18] and PYTHIA[19] Monte Carlo generators that have been used in the present analysis.

It is a common practice to fit the parameters, those models depend on, when measuring the b fragmentation function in e^+e^- collisions. This has been done by ALEPH [31] , OPAL [32] and SLD [33] as a part of their b fragmentation analyses, and also by DELPHI [34], previously to the analysis presented in this work.

Phenomenological hadronization models have sometimes been folded with theoretical QCD computations of the perturbative component of the fragmentation function. It will be shown in Chapter 6 that this approach is problematic.

Some of these functions, derived from phenomenological ideas, have been contributed by Peterson *et al.* [35], Collins and Spiller [36], and Kartvelishvili *et al.* [37]. Models that are built on more basic ideas of the string fragmentation process have been suggested by Andersson *et al.* [38] and Bowler [24]. In the sections below, where these models are presented, the fragmentation variable is noted as z to be consistent with the original publications.

The Peterson Model A simple phenomenological parameterization of the spectra of heavy hadrons was presented in 1983 by Peterson *et al* [35]. This model relies on simple kinematical arguments.

The basic idea behind this model is that when a light quark $\overline{q}$ combines with a heavy quark Q to form a hadron $H = Q\overline{q}$, the heavy quark decelerates only slightly. Thus, Q and $Q\overline{q}$ should carry almost the same energy; this effect is expected to dominate over more subtle dynamical details.

Therefore the gross feature in the amplitude $\mathcal{M}$ of the transition $Q \to Q\overline{q} + q$ is supposed to be given by the energetic gain in the process:

$$\mathcal{M} \propto \frac{1}{\Delta E} = \frac{1}{E_H + E_q - E_Q} \tag{2.25}$$

this energy difference between the initial and the final state can be written as

$$\Delta E = \sqrt{m_H^2 + z^2 p^2} + \sqrt{m_q^2 + (1-z)^2 p^2} - \sqrt{m_Q^2 + p^2} \tag{2.26}$$

Where p is the three-momentum of the heavy quark just before it hadronizes. m_H, m_Q and m_q are the masses of the heavy hadron, the heavy quark and the light quark respectively. When the hadron takes a fraction z of the quark momentum, the rest of the energy is left to the light

quark. Since the hadron mass is, as a first approximation, not much different from the heavy quark mass, the last expression can be rewritten as:

$$\Delta E = \sqrt{m_Q^2 + z^2 p^2} + \sqrt{m_q^2 + (1-z)^2 p^2} - \sqrt{m_Q^2 + p^2} \tag{2.27}$$

Moreover, for high energy accelerators where the assumption $p/m_Q \gg 1$ is justified, the expression can be approximated by:

$$\Delta E \simeq \frac{m_Q^2}{2p}\left(\frac{1}{z} + \frac{m_q^2/m_Q^2}{1-z} - 1\right) \propto 1 - \frac{1}{z} - \frac{\varepsilon}{1-z} \tag{2.28}$$

Where $\varepsilon = \frac{m_q^2}{m_Q^2}$. The value of ε can not be calculated exactly, because the light quark mass in the equation is just a representation of the non-perturbative strong interaction scale. An estimate for this scale is the constituent mass of light quarks in a ρ meson [35]. Using this value gives an estimate of ε for the case of b quark fragmentation $\varepsilon_b \approx \frac{(m_\rho/2)^2}{m_{B^0}^2} \approx 5\,10^{-3}$ [15]. The actual fragmentation function is supposed to be proportional to the squared amplitude $\mathcal{M}^2$ of Equation (2.25), completed by a phase space factor of $1/z$. Therefore the following parametrization is obtained:

$$D_Q^H(z) = N(\varepsilon)\frac{1}{z\left(1 - \frac{1}{z} - \frac{\varepsilon}{1-z}\right)^2} \tag{2.29}$$

Where $N(\varepsilon)$ is an analytically calculable normalization factor.

The Collins-Spiller Model Collins and Spiller noted in 1985 that the Peterson parametrization shows the behavior $\propto (1-z)^2$ when z approaches unity [36]. This is in contradiction with the $\propto (1-z)$ behavior expected for mesons from dimensional counting arguments by Brodsky *et al.* [39]. For the same reason they stated that the Peterson parametrization cannot fulfill the reciprocity relation [40] which connects exclusive heavy meson H structure functions $F_H^Q(z)$ to exclusive fragmentation functions $D_Q^H(z)$ at high z and expects, at the limit $z \to 1$, that $F_H^Q(z) \approx D_Q^H(z)$. This may be understood qualitatively by the argument that at large z in both cases there is a heavy quark that dominates the properties of a hadron, and a low energy remainder. The probability, to create from a quark a hadron with a fraction z of the quark energy, is therefore expected to be close to the probability to extract a quark with an energy fraction z from a hadron.

Collins and Spiller have suggested a fragmentation parametrization that ensures the requirements from reciprocity and dimensional counting and gives the wanted asymptotic behavior of $(1-z)$ when $z \to 1$. Their model is based on their own prediction of heavy meson structure functions, which is calculated by perturbative QCD, taking into account off-shell mass effects. The fragmentation parametrization obtained in [36] for heavy quarks is:

$$D_Q^H(z) = N(\varepsilon)\left(\frac{1-z}{z} + \frac{2-z}{1-z}\varepsilon\right)\left(1+z^2\right)\frac{1}{\left(1 - \frac{1}{z} - \frac{\varepsilon}{1-z}\right)^2} \tag{2.30}$$

Where $N(\varepsilon)$ is an analytically calculable normalization factor.

[15] In practice ε is treated as a free parameter.

The Kartvelishvili Model An approach similar to the one by Collins and Spiller has been proposed by Kartvelishvili *et al.* already in 1978 [37]. It also makes use of the reciprocity relation to derive a heavy quark fragmentation function from the corresponding heavy meson structure function.

Kartvelishvili *et al.* derive the fragmentation parametrization:

$$D_Q^H(z) = (\alpha+1)(\alpha+2)z^\alpha(1-z) \tag{2.31}$$

and predict that, for *b* fragmentation, $\alpha \approx 9$.

The Lund Symmetric Fragmentation Function According to the string fragmentation model, hadrons are created iteratively at each end of the string. The *z* distribution determines the actual positions of breaks on the string. Andersson *et al.* have imposed the requirement that hadronization from either end of the string, whether it is a quark or an antiquark, should lead to the same distribution of string breaks. They have shown [38] that this symmetry condition, along with additional assumptions like no transverse string excitations and a constant string tension, is sufficient to constrain the functional form of the fragmentation function to

$$D_Q(z) = N\frac{1}{z}z^{a_\alpha}\left(\frac{1-z}{z}\right)^{a_\beta}\exp(-\frac{bm_t}{z}) \tag{2.32}$$

where m_t is the transverse mass of the meson created in the hadronization process. *b* is a universal parameter. a_α and a_β are parameters for the quark flavor before and after the current hadronization step. Usually one assumes $a_\alpha = a_\beta = a$ for all flavors. This assumption is not based on firm theoretical grounds, but on the fact that there are no experimental results contradicting it. It leads to a simplification of the fragmentation function as

$$D_Q(z) = N\frac{1}{z}(1-z)^a\exp(-\frac{bm_t}{z}) \tag{2.33}$$

It has been mentioned previously that dimensional counting arguments determine the behavior of the fragmentation function for $z \to 1$. The Lund symmetric function behaves like $(1-z)^a$ in the large *z* region. It depends thus on the parameter choice whether this theoretically motivated boundary condition mentioned above is fulfilled or not.

The parameters *a* and *b* do not only impact the energy spectrum of hadrons produced in the fragmentation spectrum, but also their rapidity distribution and the multiplicity of hadrons in a jet. These aspects will not be discussed further in this thesis, although they provide means for independent cross-checks of parameter fits to the hadron energy spectrum.

In Chapter 6, the Lund model is used to fit the non perturbative component of the fragmentation function. It has been verified that the transverse momentum of a heavy hadron obtained in the string fragmentation process is negligible when compared to the mass of the *b* quark, and therefore the product bm_t^2 has been used as one fit parameter. The fragmentation analysis of OPAL [32] has also followed this convention. This remark applies also to the Bowler model described below.

The Bowler Model Bowler [24] refers back to the original string model of Artru and Mennessier [25] and modifies the light cone area calculation to prove that the inclusion of quark masses at the endpoints of a string leads to the Lund symmetric function, corrected by an additional factor $z^{-bm_Q^2}$ where m_Q is the transverse mass of the endpoint quark:

$$D_Q(z) = N\frac{1}{z^{1+bm_Q^2}}(1-z)^a \exp(-\frac{bm_t}{z}) \tag{2.34}$$

For heavy quarks, the transverse mass of the heavy quark is approximately equal to the transverse mass of the heavy hadron, $m_Q \approx m_t$:

$$D_Q(z) = N\frac{1}{z^{1+bm_t^2}}(1-z)^a \exp(-\frac{bm_t}{z}) \tag{2.35}$$

The main modifications of the Bowler model with respect to the original Artru-Mennessier picture is the replacement of straight lines in space-time (massless particles) by hyperbolae (massive particles), modifying the area in the history of light-cones to be considered for string fission. An additional change is the replacement of the continuous mass spectrum of Artru-Mennessier string decays by a discrete spectrum. This is done by restricting possible string break points in space-time to discrete hyperbolae corresponding to different invariant masses. Only those space-time points in the light cone history of a possible string break are taken into account that coincide with one such hyperbola [41]. This leads to another modification of the area decay law. The advantage of this approach is that in this way, the uniform probability for string breaks is ensured despite the mass shell constraints, whereas the Lund symmetric approach had to give up uniformity.

The ansatz by Bowler is quite different from the symmetry constraint approach by the Lund group, and yet the results are remarkably similar. Although this might be surprising at first glance, it is not unexpected. Almost all conditions imposed by the Lund group are intrinsic to the basic Artru-Mennessier model. The Lund group starts from an iterative string fission ansatz and tries to ensure that during iteration the symmetries and properties of the basic model are accounted for. Bowler, on the other hand, starts from the basic model properties and arrives at a parametrization which is also usable for iterative treatment of the string hadronization.

2.4 Excited States

At the last step of the hadronization process are produced b-hadron resonances of various masses. The frontier remains unclear between low mass clusters and high mass resonant systems; usually well defined states are simulated only for first excited states, namely the B^* and B^{**} mesons in-case of *b*-flavored mesons.

As the B^* mass differs by only 46 MeV/c^2 from the weakly decaying B meson mass, because of spin counting, it is expected that prompt B^* production is three times larger than the prompt production of the peseudo-scalar states. This has been verified by experiments running at the Z^0 pole.

B^{**} is the generic name for the first orbital excitations (L=1) of $b\bar{q}$ bound states. The determination of B^{**} production rates in jets is not so precisely determined. This is partly because

J_q	J^P	B^{**} state	Decay Mode	Width
$1/2$	0^+	B_0^*	$B\pi$	Broad ($L_p = 0$, S-wave)
$1/2$	1^+	B_1^*	$B^*\pi$	Broad ($L_p = 0$, S-wave)
$3/2$	1^+	B_1	$B^*\pi$	Narrow ($L_p = 2$, D-wave)
$3/2$	2^+	B_2^*	$B^*\pi, B\pi$	Narrow ($L_p = 2$, D-wave)

Table 2.1: *Properties of the four* $B_{u,d}^{**}$ *states.*

one expects four such states and two of them are supposed to be broad. In addition experimental signals sit above a large combinatorial background whose shape depends on details of the *b*-fragmentation.

In the resulting theoretical framework of Heavy Quark Effective Theory (HQET) [42], the spin of the b quark ($\vec{S}_b$) in a hadron is decoupled from the total angular momentum of the light spectator system ($\vec{J}_q = \vec{S}_q \oplus \vec{L}$). Each energy level of the $b\bar{q}$-system thus consists of a pair of states labeled by J_q and the *total spin* $\vec{J} = \vec{J}_q \oplus \vec{S}_b$, i.e. by $J = J_q \pm 1/2$. In the heavy quark limit the two states are degenerate in mass and have the same width.

For orbitally excited states with $L = 1$, combining $J_q = 1/2$ and $3/2$ with $S_b = \pm 1/2$ gives *four* B^{**} meson states with $J^P = 0^+,\ 1^+,\ 1^+$ and 2^+. Parity conservation limits the orbital angular momentum (L_p), present in their 2-body decay modes, to be even valued. Not considering multi-pion decay channels, leads to the decay modes listed in Table 2.1 for the $B_{u,d}^{**}$ states. A similar pattern is expected for B_s^{**} states although, isospin conservation demands that - if the B_s^{**} mass is above threshold - these states decay to $B^{(*)}K$.

Members of the $J_q = 3/2$ doublet are expected to be narrow, since only $L_p = 2$ (D-wave) transitions are allowed. The other states can decay in an S-wave and are expected to be broad. The broad state with $J = 1$ will be denoted as B_1^* in order not to get confused with the corresponding narrow state.

The widths of the broad states are expected to lie between 200 and 300 MeV/c^2 [43, 44].

Properties of orbitally excited charm mesons are expected to follow the same pattern with corrections to the heavy quark limit given by the ratio between the c and the b quark mass, m_b/m_c, i.e. approximately three [42, 45, 46]. This scaling is expected to apply for all mass splittings which cancel in the heavy quark mass limit. It is well verified for the doublets comprising the 0^- and 1^- states: $m_{B^*} - m_B \simeq (m_{D^*} - m_D)/3$.

Evidence for narrow $B_{u,d}^{**}$ states first emerged in analyses at LEP in 1995 in which a charged pion produced at the primary event vertex was combined with an inclusively reconstructed B meson [47, 48].

Table 2.2 summarizes measurements of narrow $B_{u,d}^{**}$ states. To enable a comparison to be made, published numbers have been adjusted, where possible, to be valid for a common set of input parameters.

Results in Table 2.2 show some agreement for the measured masses but there is a spread of values in rate measurements which needs to be understood. In this purpose, the crucial issue is the control of the combinatorial background contribution which formed the major source of systematic error in the inclusive measurements, listed in Table 2.2. This has been the main line followed by a new DELPHI measurement [49] in which depleted and enriched samples in B^{**} candidates are used to control, using real data events, the shape of the combinatorial

	$m(\mathrm{B}^{**}_{\mathrm{u,d}})$ [MeV/c^2]	$\sigma(\mathrm{B}^{**}_{\mathrm{u,d}})/\sigma_{\mathrm{b}}$
OPAL incl.	5712 ± 11	0.21 ± 0.05
DELPHI incl.	$5732 \pm 5 \pm 20$	$0.27 \pm 0.02 \pm 0.06$
ALEPH incl. (all)	$5734 \pm 3 \pm 16$	$0.214 \pm 0.012 \pm 0.045^{+0.030}_{-0.045}$
ALEPH incl.(peak)	$5734 \pm 3 \pm 16$	$0.144 \pm 0.008 \pm 0.030$
L3 incl.	$\mathrm{B}^*_2 : 5768 \pm 5 \pm 6$	$0.32 \pm 0.03 \pm 0.06$
CDF semi-excl.	$\mathrm{B}_1 : 5710 \pm 20$	$0.22 \pm 0.05 \pm 0.02$
ALEPH excl.	$\mathrm{B}^*_2 : 5739^{+8+6}_{-11-4}$	$0.24 \pm 0.07^{+0.05} - 0.04$

Table 2.2: *Current world results for the mass and production rate of* $\mathrm{B}^{**}_{\mathrm{u,d}}$ *states.*

background. With this technique, they obtain:

$$\frac{\sigma(\mathrm{B}^{**}_{\mathrm{u,d}}) \cdot \mathrm{BR}(\mathrm{B}^{**}_{\mathrm{u,d}} \to \mathrm{B}^{(*)}\pi)}{\sigma_{\mathrm{b}}} = 0.157 \pm 0.011(\mathrm{stat})^{+0.022}_{-0.014}(\mathrm{syst}) \pm 0.019(\mathrm{model}),$$

and the narrow $\mathrm{B}^{**}_{\mathrm{u,d}}$ production rate is about 60 % of this value.

2.5 B-hadron Production Rates

Production rates of the different weakly decaying B-hadron particles have been accurately determined using measurements from LEP and TeVatron experiments. These determinations use also results from $\mathrm{B} - \overline{\mathrm{B}}$ mixing obtained at *b*-factories and at the previous facilities.

Several types of data have been combined. They comprise "direct" measurements as the production rates of B-hadrons decaying into a specific final state of known branching fraction (semileptonic decays or exclusive channel). A more inclusive approach was also followed by DELPHI using neural networks to separate charged from neutral b-hadrons. "Indirect" measurements have been used also as the mixing probability χ. For *b*-hadron jets, produced at high energy, this probability corresponds to the average of the contributions from $\mathrm{B}^0_{\mathrm{d}}$ and $\mathrm{B}^0_{\mathrm{s}}$ mesons:

$$\overline{\chi} = f_d \chi_d + f_s \chi_s \tag{2.36}$$

in which f_q are the fractions of $\overline{b}q$ mesons in the jet and $\chi_{d,s}$ are the oscillation probabilities for $\mathrm{B}^0_{\mathrm{d}}$ and $\mathrm{B}^0_{\mathrm{s}}$ meson respectively.

These oscillation probabilities can be obtained by integrating the corresponding time distribution for mixed states.

$$\chi_q = \frac{1}{2} \frac{\left(\Delta m_q \tau_{B_q}\right)^2}{1 + \left(\Delta m_q \tau_{B_q}\right)^2} \tag{2.37}$$

$\Delta m_d \tau_{B_d}$ have been accurately determined, present results being dominated by those from *b*-factories. For $\mathrm{B}^0_{\mathrm{s}}$ mesons, the lifetime has been measured with 4% accuracy and a limit exists on the value of Δm_s ($\geq 14.5\ ps^{-1}$ at 95 % CL). This last value implies that $\chi_s > 0.49885$ [50]. Having measured $\overline{\chi}$ in jets at high energy colliders, Equation 2.36 provides a constraint on the

values of $f_{d,s}$. Another constraint on these rates comes from the expected equality of B^0_d and B^+ production fractions. Such an equality is expected as *u* and *d* quarks are almost identical relative to strong interactions (isospin invariance). In the last steps of the hadronization process one thus expects similar rates for the creation of $\bar{b}d$ and $\bar{b}u$ systems, which are mainly excited states decaying into B^0_d or B^+ particles, unless there is isospin violation in the decay of low mass resonances. These violations were observed for *c*-hadrons, were D^{*0} decays only into D^0 and, as a result, production fractions of D^0 and D^+ in *c*-jets were rather different. For *b*-hadrons, all B^* mesons decay electromagnetically and no asymmetry is introduced between B^0_d and B^+ production fractions. Also, masses of B^{**} states are sufficiently above their decay threshold so that no asymmetry is expected. As a final constraint, it is expected that the *b*-quark is contained within one of the possibly produced weakly decaying states:

$$f_d + f_u + f_s + f_{b-baryon} = 1 \tag{2.38}$$

Combining all these informations, the Heavy Flavor Averaging Group obtains:

$$f_d = f_u = 0.397 \pm 0.010,\ f_s = 0.107 \pm 0.011 \text{ and } f_{b-baryon} = 0.099 \pm 0.017. \tag{2.39}$$

2.6 B Decays

In a $p\bar{p}$ environment the *B* decay products are situated in events where the mean multiplicity is much larger than the one at the Z^0 pole. Furthermore the ratio $\sigma_{b\bar{b}}/\sigma_{tot}$ is of the order of a few permil. As a consequence, only specific channels e.g. with fully reconstructed final states, or semileptonic decays, can be studied with a reasonable signal-to-background ratio.

Chapter 3

Experimental Framework I- The LEP Collider and the DELPHI Experiment

3.1 The Large Electron Positron Collider

The Large Electron Positron Collider LEP was a circular e^+e^- accelerator constructed at the CERN laboratory in Geneva (Switzerland). It operated from 1989 for 11 years and was, with a circumference of 27 km, the largest collider designed to that point. It consisted of eight 2.9 km long arcs and eight 0.43 km straight sections housing about 3400 dipole bending magnets, 820 quadrupoles, 500 sextupoles and 700 correcting magnets. Table 3.1 summarizes the main properties of the LEP collider. An extensive review can be found in [51].

At four points of the LEP circumference the e^+ and e^- beams crossed. In those interaction points the four experiments, namely ALEPH, DELPHI, L3 and OPAL, were placed to record the particles produced in the collisions. Figure 3.1 shows the LEP ring and the four experiments localization.

The possibility of creating new heavy particles in a collider depends on the center of mass energy provided in the collision by the accelerated beams. For an e^+e^- symmetric machine like LEP, this is $\sqrt{s} = 2E$, where E is the beam energy. From 1989 to 1995 LEP operated at the energy of the Z resonance, $\sqrt{s} = 91.2$ GeV. This phase was named the LEP-I period and allowed for a wide study of Z production and decays. From 1995 onward, the energy of the beams was increased by progressively adding super-conducting cavities, reaching the maximum energy of about 210 GeV in the year 2000. This phase was aimed at studying W^+W^- pairs, produced above $\sqrt{s} = 161$ GeV, and to search for new particles such as the Higgs boson and supersymmetric particles. This phase was called the LEP-II period.

One of the main challenges in circular e^+e^- accelerators is to compensate for the energy loss due to the synchrotron radiation. Photons are emitted by electrons along their flight direction due to the transverse acceleration. Since the energy loss by synchrotron radiation is proportional to the beam energy raised at the fourth power, it can reach large values and has to be compensated for by the accelerating cavities. In addition, the dissipated power has to be absorbed by the ring components. At LEP, the synchrotron radiation was about 125 MeV/turn when running at the Z energy and reached about 3400 MeV/turn in the LEP-II period.

Apart from the mass of the system of created particles, given by $\sqrt{s}$, another important characteristic in a collider is the number of produced events. This is determined by the luminosity

$\mathcal{L}$. For a given physical process with a cross section σ, the total number of events produced per unit of time is:

$$\frac{dN}{dt} = \mathcal{L}\sigma. \tag{3.1}$$

For two beams of similar dimensions, the luminosity can be expressed as:

$$\mathcal{L} = \frac{k_b N_{e+} N_{e-} f_{rev}}{4\pi\sigma_x\sigma_y} \tag{3.2}$$

where k_b is the number of particle bunches per beam, N_{e+} and N_{e-} are the number of positrons and electrons per bunch respectively, f_{rev} is the revolution frequency and $\sigma_x\sigma_y$ is the transversal section of the beams. Figure 3.2 shows the time integrated luminosity [1] reached by the LEP machine during its entire operation period.

Circumference	26.658 km
Mean radius of curvature	4242.893 m
Depth	80-130 m
Interaction points	4
Experiments	ALEPH, DELPHI L3 and OPAL
Injection Energy	22 GeV
Maximum energy reached at the c.m.s. (LEP-II)	208.8 GeV
Beam size	$\sigma_x = 100 - 300\,\mu$m $\sigma_y = 2 - 6\,\mu$m
Number of Bunches	2 - 4
Particles per bunch	$\sim 10^{11}$
Revolution frequency	11.2 kHz
Synchrotron radiation	$\sim 125 - 3400$ MeV/turn
Luminosity	$\sim 1 - 10 \times 10^{31}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$

Table 3.1: *Properties of the LEP e^+e^- collider. The operation mode depends on the LEP period.*

The data used in this work was recorded by the DELPHI detector during the years 1992 to 1995 corresponding to the LEP-I period. The integrated luminosity reached by LEP in that period was $179\,\mathrm{pb}^{-1}$.

[1] The luminosity is usually expressed as an integrated quantity in units of inverse barns ($1\,\mathrm{barn} \equiv 10^{-28}\,\mathrm{m}^2$).

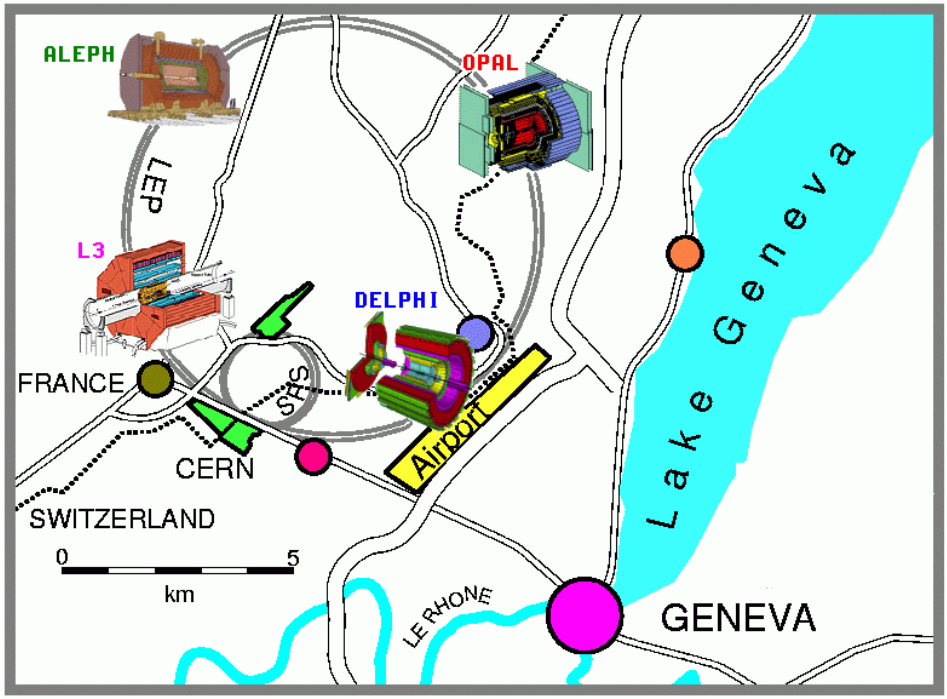

Figure 3.1: *LEP collider and the placement of the four experiments.*

3.2 The DELPHI Experiment

DELPHI (**DE**tector with **L**epton, **P**hoton and **H**adron **I**dentification) was designed and installed at LEP with the aim of providing high precision and granularity, and very effective particle identification, thanks to the implementation of a Ring Imaging Cherenkov detector. An advanced silicon detector also allowed very precise tracking and vertex determination.

The DELPHI detector is described in reference [52]. This section only seeks to summarize its main characteristics and sub-detectors. Furthermore, this document will refer only to the performance of DELPHI during the LEP-I period. It should be noted that many of these characteristics changed and improved during the years.

The DELPHI ensemble consisted of a cylindrical section in the central part called the "barrel" and two end-caps covering the "forward" regions. Figure 3.3 shows the layout of the barrel and of one end-cap. The DELPHI standard coordinate system is given by the z axis along the electron beam direction, the x axis pointing towards the center of LEP and the y axis upwards. Cylindrical coordinates (R, θ, ϕ) usually are used, ϕ and R being the azimuthal angle and radius in the x-y plane, respectively, and θ the polar angle with respect to the z axis. The super-conducting solenoid was 7.4 m long with 5.2 m inner diameter. It provided a highly uniform magnetic field of 1.23 T, corresponding to a current of 5000 A, parallel to the z axis through the central part of the barrel. The super-conducting cable consisted of 17 wires made of 300 Nb-Ti filaments embedded in copper and cooled by liquid helium at 4.5 K. The main purpose of the super-conducting solenoid was to curve the trajectory of charged particles allowing their momentum measurement and identification (see Section 3.5.1).

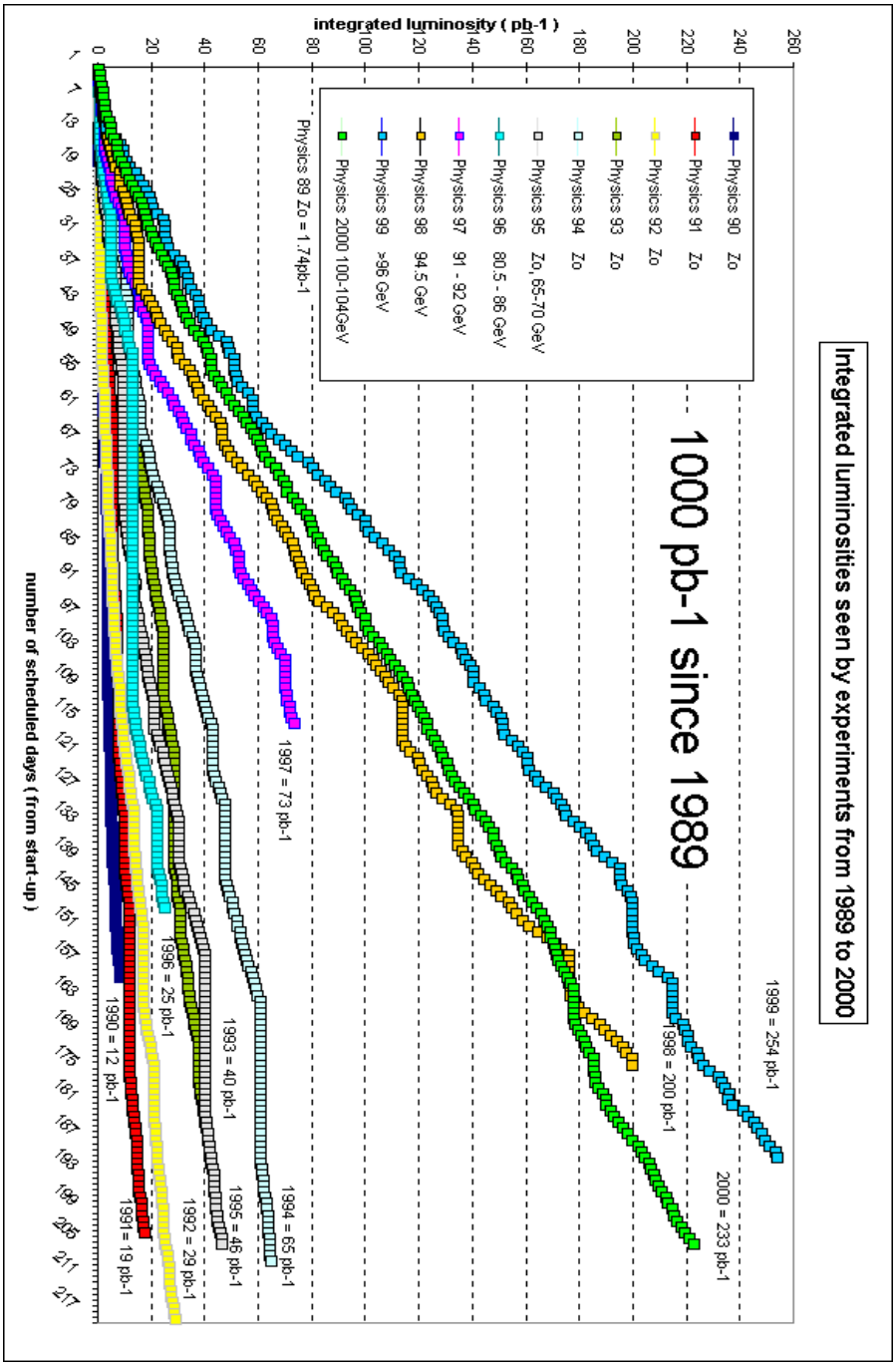

Figure 3.2: *Luminosity reached by the LEP collider during its operation period.*

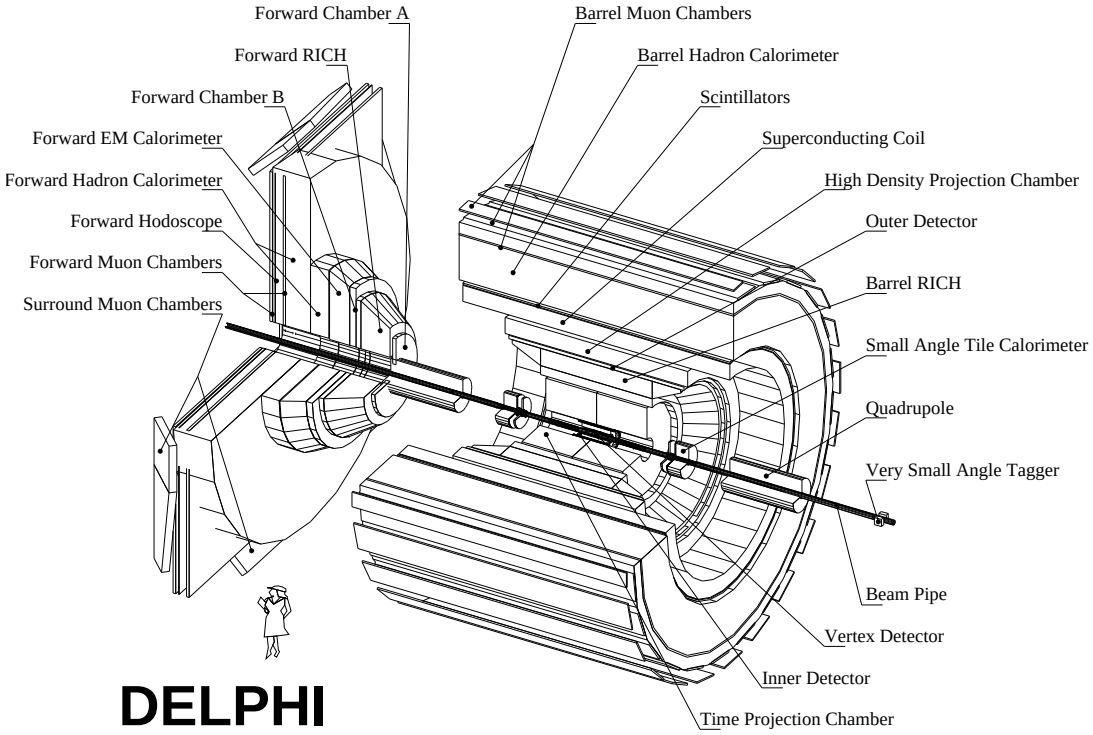

Figure 3.3: *Schematic layout of the DELPHI detector.*

3.3 Tracking Detectors

The tracking system provided the trajectories of the particles by evaluating their momenta and impact parameters. Detectors in this system were: the Vertex Detector (VD), the Inner Detector (ID), the Time Projection Chamber (TPC) and the Outer Detector (OD), in the barrel region; and the tracking chambers (FCA and FCB) in the forward region. They were the most important detectors for analyses involving B decays as we will see later.

3.3.1 The Vertex Detector

This detector provided a very precise measurement of the passage points of charged particles near the interaction point. This allowed for the determination of both primary and secondary vertices and the track impact parameter, providing a track reconstruction with a track resolution of about 300 μm. It is essential in the study of B hadrons coming from Z decays since their mean decay length is about 3 mm from the primary vertex.

The VD detector consisted of three coaxial cylindrical layers of silicon strip detectors in the barrel region, at average radii of 6.3 cm, 9 cm, and 10.9 cm around the beam pipe. Each layer was formed by 24 overlapping (about 10° in ϕ) modules of 23.6 cm length containing four detector plates each (see Figure 3.4). The polar angle coverage for charged particles hitting all three layers was $44^\circ \leq \theta \leq 136^\circ$. The readout pitch was 50 μm in the Rϕ plane perpendicular to the beam direction. In April 1994, the inner and outer layers were equipped with double-sided silicon detectors. The new detectors increased the polar angle coverage to $24^\circ \leq \theta \leq 155^\circ$.

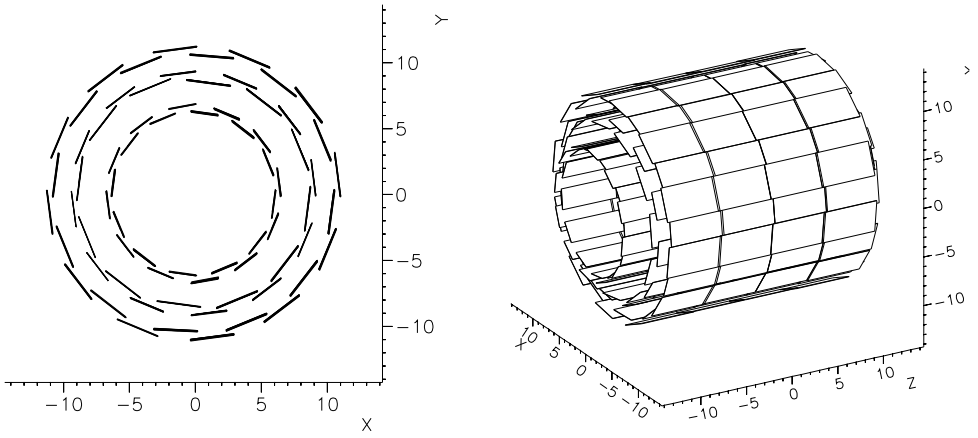

Figure 3.4: *Schematic cross-sections of the DELPHI Vertex Detector (1994-1995 configuration).*

The alignment of the VD was necessary for the track reconstruction. A first mechanical alignment used laser located between the VD and ID detectors determining the position of each module. After that, dimuon $Z \rightarrow \mu\bar{\mu}$ events and hadron tracks were used in order to calibrate the detector. Tracks in the overlaps were taken into account to refine the Rϕ rotations and translations of the modules in a layer. Overlap and dimuon tracks were also used for the z alignment.

The intrinsic resolution for a single hit of the detector was about 8 μm in the Rϕ plane, and about 9 μm for the z coordinate in case of tracks perpendicular to the modules.

3.3.2 The Inner Detector

The Inner Detector was placed between the VD and the TPC. It was responsible for extrapolating the reconstructed track elements from the TPC to the vertex detector. It also contributed to the first and second levels of the DELPHI trigger, reducing the rate due to cosmic and beam gas events.

The ID consisted of two parts:

- The drift chamber (called Jet-Chamber, at 11.8 cm < R < 22.3 cm) divided into 24 azimuthal sectors of 15 degrees in ϕ. Each sector consisted of 24 sense wires arranged parallel to the beam direction, which provided about 24 Rϕ points per track. The polar angle coverage was $30^\circ \leq \theta \leq 150^\circ$ until the beginning of 1995. From 1995 onward a new longer ID was installed, increasing the θ coverage to $15^\circ \leq \theta \leq 165^\circ$. The single wire resolution was of the order of 90 μm in Rϕ. The resolution of one track element was about 40 μm in Rϕ. The 2-track separation was about 1 mm.

- The trigger layers: until 1995 five cylindrical MWPC (Multi-Wire Proportional Chamber) layers were arranged surrounding the drift chamber. Each layer consisted of 192 sense wires spaced by about 8 mm and with 192 circular cathode strips providing Rz information for the track elements. The Rϕ measurement was mainly used in the trigger. It also resolved the left-right ambiguity in the drift chamber. Since 1995, the MWPC layers were

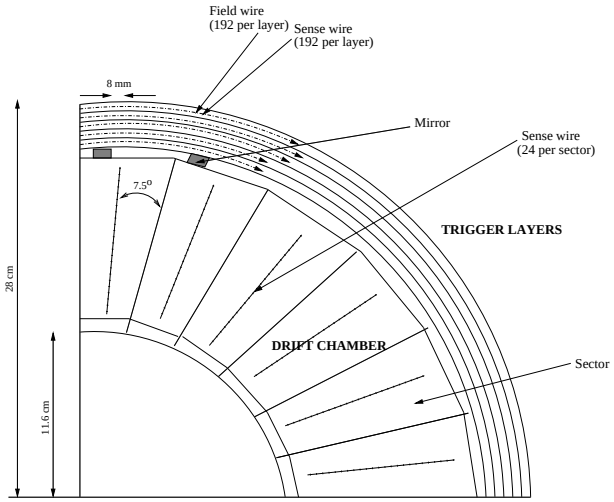

Figure 3.5: *Scheme of DELPHI Inner Detector.*

replaced by 5 cylindrical layers of straw tube detectors, each layer comprising 192 tubes. They had the same function as the old MWPC trigger layers but there was no longer a z measurement. For the old MWPC layers, the z precision from a single layer, for an isolated track, varied from 0.5 mm to 1 mm depending on θ.

Figure 3.5 shows a scheme of the DELPHI Inner Detector.

3.3.3 The Time Projection Chamber

The Time Projection Chamber, placed between the Inner and Outer Detectors, was the central tracking device in DELPHI. Along with the VD, it contributed to the track reconstruction, providing a precise measurement of the particle momenta. The TPC was a gas-filled (80% argon and 20% methane) cylinder, 340 cm long and with 120 cm radius, divided into 6 azimuthal sectors each with 192 sense wires and 16 circular pad rows (see Figures 3.6 and 3.7). Particles traversing the TPC ionized the gas. Free electrons drifted towards the end-caps by means of a 150 V/cm electric field parallel to the magnetic field of DELPHI. This electric field was created by a high tension plate (-20 kV) which separated the cylinder in two regions. The charge of the avalanche created by the drift electrons was collected on the two end plates of the TPC by means of anode wires and pad rows. This charge provided information on the charged particle energy loss per unit of length dE/dx. The z coordinate was calculated by measuring the drift time of the electrons and knowing their drift velocity. The TPC provided 16 space points per particle trajectory at 40 cm < R < 110 cm radii and between polar angles $39^\circ \leq \theta \leq 141^\circ$. The pad rows gave information about the ionization position in the Rϕ plane with a resolution of 250 μm per point. The z coordinate resolution was about 900 μm.

3.3.4 The Outer Detector

The Outer Detector complemented the TPC, improving the momentum determination for charged particles and complementing the geometrical acceptance by covering the dead zones

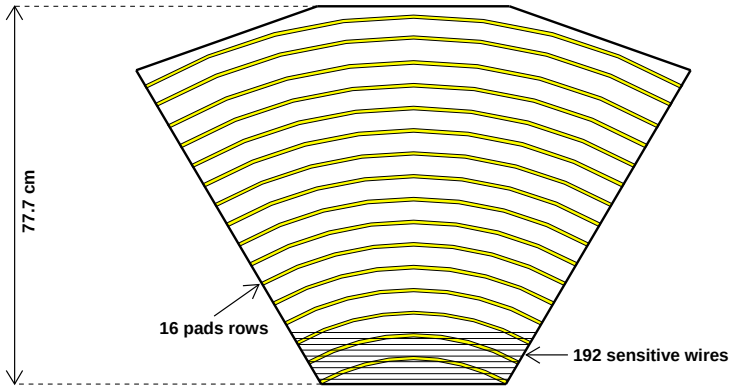

Figure 3.6: *Schematic view of the sense wires and pads rows in one sector of the TPC.*

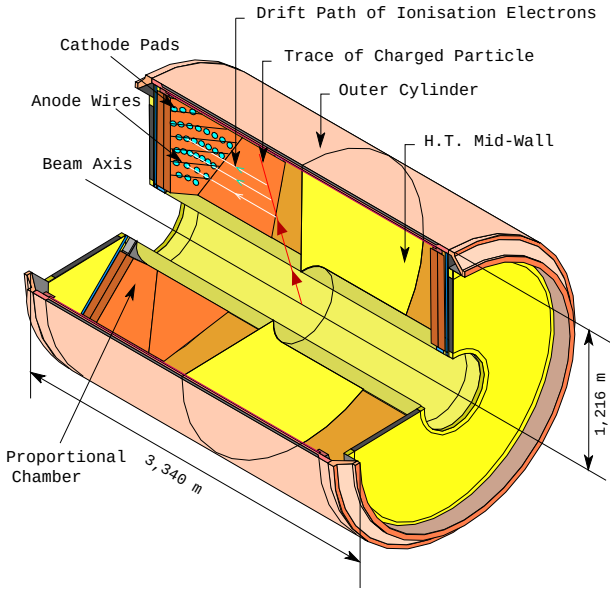

Figure 3.7: *Scheme of the DELPHI Time Projection Chamber.*

of the TPC. It consisted of 3500 drift tubes, operating in the limited *streamer* mode. The tubes were arranged in 5 layers allocated between radii of 197 cm and 206 cm. The tubes, of 4.7 m length, were oriented in the beam direction. The OD detected tracks within the range of $42^\circ \leq \theta \leq 138^\circ$, with a resolution of 100 μm in the Rϕ plane independent of the drift distance. Three layers read the z coordinate by timing the signals at the end of the anode wires with a resolution of 5 cm.

In addition to these tracking detectors, the Forward chambers FCA and FCB formed part of the tracking system covering the forward and backward regions of DELPHI. They were drift chambers located at both end-caps, which contributed to the track reconstruction.

3.4 Other Detectors

3.4.1 Ring Imaging Cherenkov Detectors

The Ring Imaging Cherenkov detectors (BRICH in the barrel and FRICH in the forward regions) were decisive in providing a good particle identification. The barrel RICH was located between the TPC and the OD detectors (123 cm < R < 197 cm), covering polar angles in the region $40^\circ \leq \theta \leq 140^\circ$. It consisted of 48 sectors, each containing a liquid radiator (C_6F_{14}), one drift tube with multi-wire proportional chambers at the extremes, a gas radiator (C_5F_{12} at 40°C), and six mirrors which focused the produced Cherenkov photons. The liquid radiator (refractive index $n = 1.283$) was used for particle identification in the momentum range 0.7-8 GeV/c, whereas the gas radiator ($n = 1.00172$) was used for particles in the 2.5-25 GeV/c range. The angular precision per track (obtained from $Z \rightarrow \mu^+\mu^-$ events) varied between 5.2 mrad in the liquid radiator and 1.5 mrad in the gas radiator. Figure 3.8 shows the working principle of the RICH detector.

3.4.2 Electromagnetic and Hadron Calorimeters

Electron and photon identification was provided in DELPHI primarily by the electromagnetic calorimetry system. The main electromagnetic calorimeters were the High Projection Chamber HPC in the barrel region, and the Forward ElectroMagnetic Calorimeter FEMC in the forward ones. Two additional very forward calorimeters (the Small angle TIle Calorimeter, STIC [2], and the Very Small Angle Tagger, VSAT) were placed for luminosity monitoring through $e^+e^- \rightarrow e^+e^-$ Bhabha scattering events. The HPC was made of 144 modules each consisted of a small time projection chamber with layers of lead wires in the gas volume used as conversion material. The total absorber thickness corresponded to $18/sin\theta$ radiation lengths. The angular resolution measured for high energy photons was about 2 and 1 mrad in ϕ and θ respectively, and the energy resolution could be parametrized as $\frac{\sigma(E)}{E} = \frac{32\%}{\sqrt{E(GeV)}} \oplus 4.3\%$. The FEMC, placed in the forward regions, consisted of two arrays with 4532 Cherenkov lead glass blocks. The energy resolution in this calorimeter could be parametrized as $\frac{\sigma(E)}{E} = \frac{12\%}{\sqrt{E(GeV)}} \oplus \frac{11\%}{E(GeV)} \oplus 3\%$.

[2] It replaced the Small Angle Tagger, SAT, in 1994.

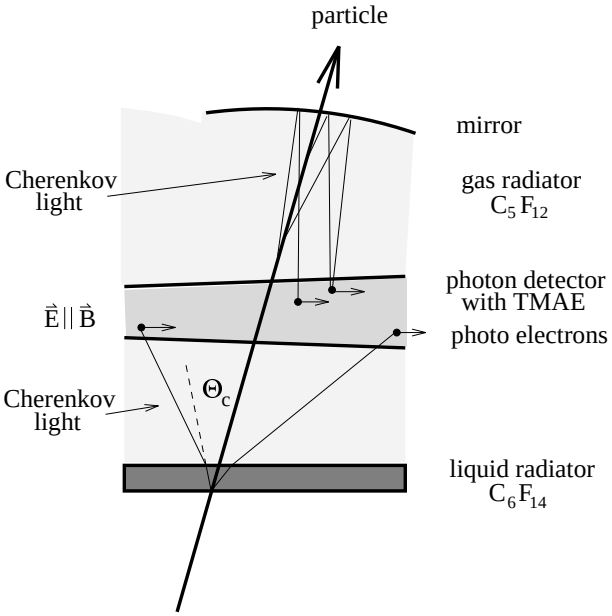

Figure 3.8: *Operation principle of the RICH detector.*

The hadron calorimeter in DELPHI was the HCAL, which provided energy measurements of charged and neutral hadrons. It consisted of a sampling gas detector (more than 19000 streamer tubes) incorporated in the iron magnet yoke. Its energy resolution, given by pions from τ decays, was found to be $\frac{\sigma(E)}{E} = \frac{1.12}{\sqrt{E\,(GeV)}} \oplus 0.21$.

3.4.3 Scintillators

The Time of Flight (TOF) detector in the barrel region (HOF in the end caps) consisted of a single layer of plastic scintillators which gave a rapid signal (1.2 ns) used in the DELPHI trigger. In addition to providing a way to distinguish between dimuon and cosmic events, it also contributed to the alignment and calibration of the DELPHI subdetectors.

3.4.4 Muon Chambers

Three layers of drift chambers in the barrel (MUB), operating in the proportional mode, and of two layers in the endcaps (MUF) [3] formed the muon system. The inner layer was embedded in the iron of the HCAL, and the other two layers were the most external detectors of DELPHI. In this way, muons with momenta larger than 2 GeV were the only charged particles that could traverse the lead and iron of both calorimeters essentially unaffected by them and be detected by the muon chambers. For dimuon events, the spatial resolution was 2 mm in Rϕ and 8 cm in z.

[3] In 1994 a layer of Surrounding Muon Chambers (SMC) based on limited streamer tubes was installed on the endcap's periphery to fill the gap between barrel and forward regions.

Detector	# measurements in Rϕ	Resolution per point	# measurements in z	Resolution per point
VD	3	8 μm	2 (from 1994)	10 μm
ID (jet)	24	90 μm	-	-
ID (trigger)	5	500 μm	5	500 μm
TPC	16	250 μm	192	900 μm
OD	5	100 μm	3	5 cm

Table 3.2: *Number of measured points and resolution given by each tracking detector.*

3.5 Particle Identification and Reconstruction

The information given by the DELPHI detectors was combined in order to identify the detected particles. This information mainly relied on the track reconstruction, the energy loss of traversing particles and on the lepton identification.

3.5.1 Track Reconstruction

The trajectory of a particle with momentum p and charge e in a constant magnetic field $\vec{B}$ is a helix with curvature radius ρ and pitch angle λ. Tracking detectors determined the momentum from the curvature of the track in the DELPHI magnetic field according to:

$$\rho[m] = \frac{\cos\lambda \; p[\mathrm{GeV/c}]}{0.3 \, B[\mathrm{T}]} \tag{3.3}$$

Each DELPHI sub-detector provided measured points along the particle trajectory which allow the track reconstruction (see Table 3.2). DELANA was the DELPHI standard tracking reconstruction program [53]. It applied detector calibrations, performed track finding and finally calculated particle track parameters. DELANA started from the information of a track segment seen in the TPC and hits registered in the VD [4] then the track was extrapolated outwards and inwards to the OD and ID detectors. In the dead zones of the TPC, track elements from the ID and OD were linked together. Track elements, obtained in this way, were passed through a track fitting processor. From the calorimeters, clusters of energy were associated to charged particle tracks. Neutral tracks were obtained from the remaining energy clusters. Hits in the Muon Chambers were associated to charged tracks. RICH information was used on the fitted final tracks.

The resulting events were stored on Data Summary Tapes (DST). The DSTs are accessible from PHDST [54] and SKELANA [55] packages.

[4] The Vertex Detector was included in the track fitting software in 1996, for the reprocessing of the data (see Section 3.7.)

3.5.1.1 Primary Vertex Reconstruction

The accurate primary vertex [5] determination is essential to know if tracks in an event correspond to the first particles created in the collision, or to their decay products. Its reconstruction is done by fitting several selected tracks, constrained by the beamspot position (the point of the beam crossing). The position of the beamspot is very accurate (a few μm) but to use it as a constraint one must consider the intrinsic transverse size of the beams ($\sigma_x = 100 - 300\,\mu$m and $\sigma_y = 2 - 6\,\mu$m). The selected tracks are required to have at least 2 VD hits, and a rejection procedure is applied to avoid bias due to wrong hit associations or to tracks coming from secondary vertices. The position of the primary vertex, $\vec{P}^{PV}$, is obtained by minimising the χ^2 function:

$$\chi^2(\vec{P}^{PV}) = \sum_i \frac{\vec{d}_i^2}{\sigma^2_{\vec{d}_i}} + \left(\frac{(P_x^{bs} - P_x^{PV})^2}{\sigma^2_{P_x^{bs}}} + \frac{(P_y^{bs} - P_y^{PV})^2}{\sigma^2_{P_y^{bs}}} \right) \tag{3.4}$$

Here, $\vec{d}_i$ ($\sigma_{\vec{d}_i}$ its error) is the distance of closest approach of each selected track i to the primary vertex position. This quantity is called track impact parameter. $P^{bs}_{x,y}$ and $\sigma_{P^{bs}_{x,y}}$ are the x and y coordinates of the beamspot position and its error, respectively.

3.5.1.2 Impact Parameter Reconstruction

The impact parameter $\vec{d}$ is the minimal distance between the trajectory of a track and a determined decay vertex, and it is used to evaluate the probability that the track is coming from this vertex. It is usually defined with respect to the primary vertex of the event. The Rϕ and Rz components are evaluated separately. Figure 3.9 schematically shows how the two impact parameters, $d_{R\phi}$ and d_{Rz}, are defined.

The sign of the impact parameter is defined with respect to the jet direction and it is shown in Figure 3.10. It is positive if the vector joining the primary vertex to the point of closest approach of the track is less than 90° from the direction of the jet to which the track belongs.

A quantity which is often used is the track significance. It is defined as the ratio of the impact parameter value to its error, $S = d/\sigma_d$ and since two impact parameters are considered, also the significances are defined in the Rϕ and Rz planes. The sign of the significance is taken to be the same as the impact parameter sign.

3.5.2 Hadron Identification

Charged hadrons, with sufficiently long lifetime to traverse a large part of the DELPHI detector (such as kaons, pions and protons), are identified by means of the TPC and RICH detectors. The energy loss, dE/dx, measurement in the TPC provides information on the charge, mass and momentum of the particle according to the Bethe-Bloch equation [56]. The emission angle θ_c of the Cherenkov light cone created in the RICH detector gives information on the mass m and momentum p of the particle according to $\cos\theta_c = \frac{1}{n}\sqrt{1 + \frac{m^2}{p^2}}$, n being the refractive index of the dielectric medium. By combining these pieces of information through several techniques, and using the registered data by other detectors, the mass, charge and momentum of the particles are determined.

[5] Also called main vertex.

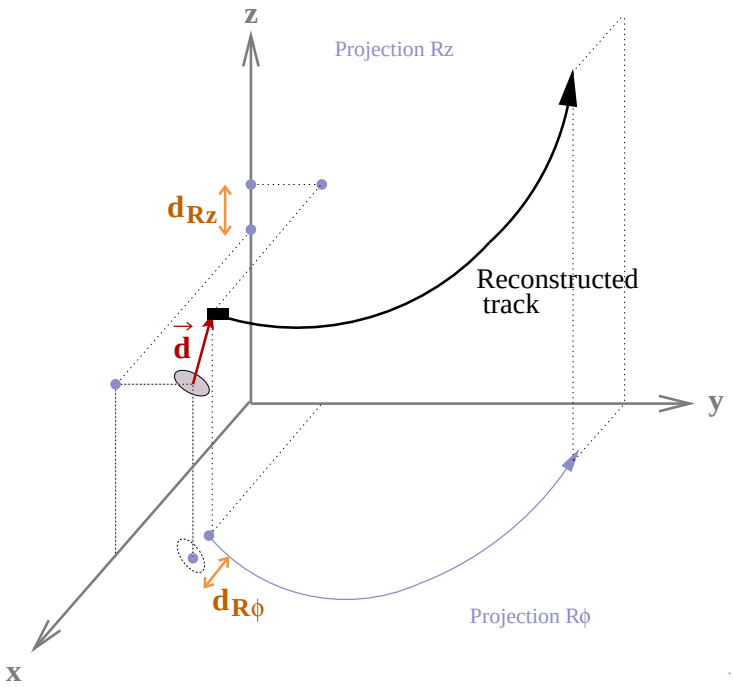

Figure 3.9: *Impact parameter definition for the Rϕ and Rz planes. d is the minimal distance, in space between the primary and secondary vertices.*

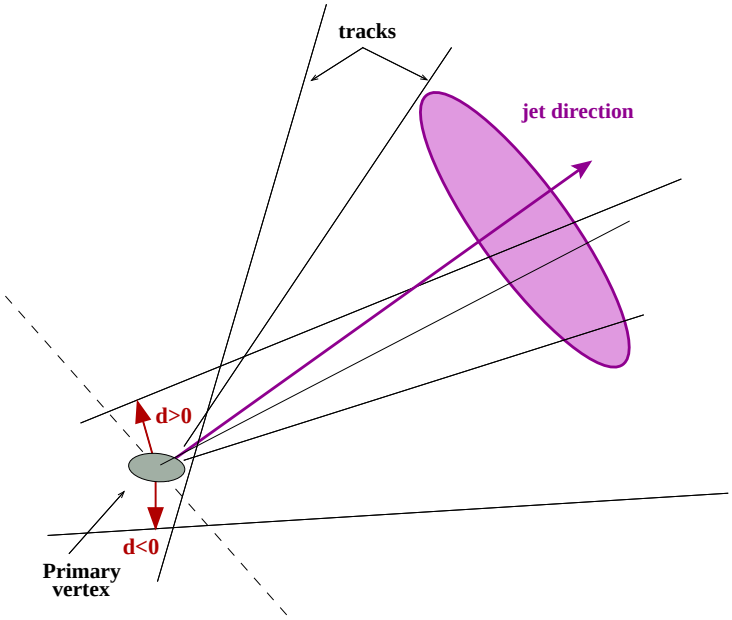

Figure 3.10: *Impact parameter sign definition.*

	Tag	*Loose*	*Standard*	*Tight*
e^-	Efficiency (%)	80	50	45
	Misid. Prob. (%)	~ 1.6	~ 0.4	~ 0.2
μ^-	Efficiency (%)	95	86	76
	Misid. Prob. (%)	$\sim 2.$	~ 0.7	~ 0.4

Table 3.3: *Lepton tag characteristics.*

3.5.3 Lepton Identification

Electrons are identified using the dE/dx measurement in the TPC and the energy deposition in electromagnetic calorimeters. The difficulty in this identification lies in the fact that electrons also interact in front of the calorimeters. Probabilities from calorimetric measurements and tracking detectors are combined in a neural network in order to achieve an electron hypothesis.

Muons can be separated from hadrons since most of the hadrons are stopped in the iron of the hadronic calorimeter. Muons with momenta above 2 GeV pass the hadron calorimeter and reach the muon chambers, where they are detected. Tracking information joined to hits in the muon chambers provide the muon identification. Detailed explanations of the DELPHI lepton identification can be found in [52, 57].

Several cuts are defined to label the detected leptons with a given efficiency and purity. These cuts give rise to three categories for leptons, the *loose*, *standard* and *tight* criteria [6]. The probability of misidentifying a hadron as a muon or an electron has been studied on data by using anti-*b* tagged samples. Electrons from photon conversions, mainly produced in the outer ID wall and in the inner TPC frame, can be rejected by reconstructing the conversion vertex. In this way, about 80% of them have been removed with negligible loss of signal. In Table 3.3 the usual efficiencies and misidentification probabilities are given separately for electrons and muons.

3.6 DELPHI Monte-Carlo Simulation

The standard DELPHI simulation software package is called DELSIM [58]. In DELSIM, the simulation of the physics processes can be carried out using several generators, for which different parameters are tuned according to the aimed study.

After particles have been generated, they are entered in the DELPHI detector simulation package. This simulation includes the specific properties of each subdetector. It provides the necessary information to record hits in the detector components and to treat the simulated particles as if they were real. The simulated response of the detectors is then processed by DELANA to produce DST tapes by following the same procedure as for real data.

[6] For muons another tag corresponding to a *v*ery loose criteria, is also available.

3.7 Data Reprocessing

In 1996 and 1997 the data collected by DELPHI in the LEP-I period was reprocessed. In addition to the best knowledge of the detector calibrations and alignment, a New Barrel Tracking (NBT) software was employed, which comprised three different reconstruction algorithms. The main feature of this new code was that the Vertex Detector was incorporated leading to a reduction of the systematics in the track determination, the main cause of uncertainty in several measurements [59]. Events with heavy quarks largely benefit from this fact because, having complex topologies, several of their tracks were lost. Apart for improvements in track reconstruction techniques, energy measurements and particle identification were also improved, resulting in a general higher quality of the data.

Chapter 4

Experimental Framework II- The TeVatron Collider and the CDF Experiment

Fermilab's TeVatron collider represents the high energy frontier in particle physics. It is currently the source of the highest energy proton - antiproton ($p\overline{p}$) collisions. The collisions occur at two points on an underground ring, which has a radius of exactly 1 km. At these collision points are two detectors: the Collider Detector at Fermilab (CDF-II) and D0. This analysis uses data collected by the CDF-II experiment.

Between 1997 and 2001, both the accelerator complex and the collider detectors underwent major upgrades, mainly aimed at increasing the luminosity of the accelerator, and gathering data samples of 2 fb^{-1} or more. The upgraded machine accelerates 36 bunches of protons and anti-protons, whereas the previous version of the accelerator operated with only 6. This has allowed the time between bunch crossings to be decreased from 3.5 μs for the previous version to 396 ns for the current collider.

The new configuration required detector upgrades at CDF-II to ensure a maximum response time shorter than the time between beam crossings. In the following pages, we describe how the proton and anti-proton beams are produced, accelerated to their final center of mass energy of 1.96 TeV, and collided. We then describe the components used to identify and measure properties of the particles produced in the collision.

4.1 TeVatron - the Source of $p\overline{p}$ Collisions

To create the world's most powerful particle beams, Fermilab uses a series of accelerators. The diagram in Figure 4.1 shows the paths taken by protons and antiprotons from initial acceleration to collision in the TeVatron.

The Cockcroft-Walton[60] pre-accelerator provides the first stage of acceleration. Inside this device, hydrogen gas is ionized to create H^- ions, which are accelerated to 750 keV of kinetic energy. Next, the H^- ions enter a linear accelerator (Linac)[61], approximately 500 feet long, where they are accelerated to 400 MeV. The acceleration in the Linac is done by a series of "kicks" from RF cavities. The oscillating electric field of the RF cavities groups the ions into

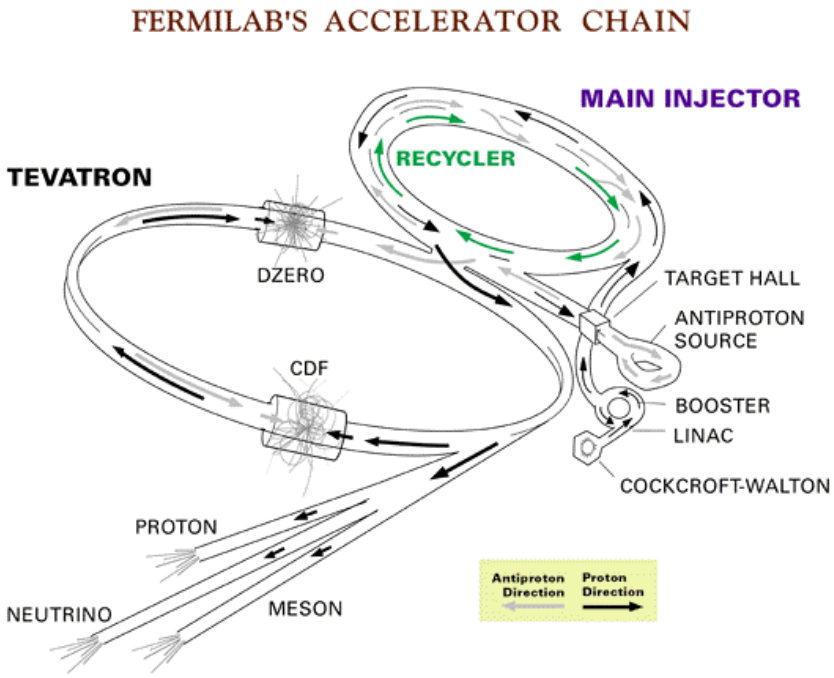

Figure 4.1: *Layout of the Fermilab accelerator complex.*

bunches.

The 400 MeV H^- ions are then injected into the Booster, a circular synchrotron[61] 74.5 m in diameter. A carbon foil strips the electrons from the H^- ions at injection, leaving bare protons. The intensity of the proton beam is increased by injecting new protons into the same orbit as the circulating ones. The protons are accelerated from 400 MeV to 8 GeV by a series of "kicks" applied by RF cavities. Each turn around the Booster, the protons accrue about 500 keV of kinetic energy.

Protons are extracted from the Booster into the Main Injector [62], which operates at 53 MHz. It has four functions. It accelerates protons from 8 GeV to 150 GeV before injection into the TeVatron, it produces 120 GeV protons, which are used for anti-proton production, it receives anti-protons from the Antiproton Source and accelerates them to 150 GeV for injection into the TeVatron, and finally, it injects protons and antiprotons into the TeVatron.

The Main Injector replaced the Main Ring accelerator which was situated in the TeVatron tunnel. The Injector is capable of containing larger proton currents than its predecessor, which results in a higher rate of anti-proton production. The Main Injector tunnel also houses the Antiproton Recycler. Not all antiprotons in a given store are used up by the collisions. Recycling the unused antiprotons and reusing them in the next store would significantly reduce the stacking time. The task of the Antiproton Recycler is to receive antiprotons from a TeVatron store, cool them and re-integrate them into the stack, so that they can be used in the next store.

To produce anti-protons, 120 GeV protons from the Main Injector are directed into a nickel target. In the collisions, about 20 antiprotons are produced per one million protons, with a mean kinetic energy of 8 GeV. The anti-protons are focused by a lithium lens and separated from other particle species by a pulsed magnet.

Before the anti-protons can be used in the narrow beams needed in the collider, the differ-

ences in kinetic energy between the different particles need to be reduced. Since this process reduces the spread of the kinetic energy spectrum of the beam, it is referred to as "cooling" the beam. New batches of anti-protons are initially cooled in the Debuncher synchrotron, collected and further cooled using stochastic cooling [63] in the 8 GeV Accumulator synchrotron. The principle of stochastic cooling is to sample a particles motion with a pickup sensor and correct its trajectory later with a kicker magnet. In reality, the pickup sensor samples the average motion of particles in the beam and corrects for the average. Integrated over a long period of time, this manifests itself as a damping force applied onto individual particles which evens out their kinetic energies. It takes between 10 and 20 hours to build up a "stack" of anti-protons which is then used in collisions in the TeVatron. Anti-proton availability is the most limiting factor for attaining high luminosities, assuming there are no technical problems with the accelerator (assuming, for example, perfect transfer efficiencies between accelerator subsystems)[61, 62].

Roughly once a day, the stacked anti-protons (36 bunches of about 3×10^{10} anti-protons per bunch) are injected back into the Main Injector. They are accelerated to 150 GeV together with 36 bunches of roughly 3×10^{11} protons. Both the protons and anti-protons are transferred to the TeVatron.

The TeVatron is the last stage of Fermilab's accelerator chain. It receives 150 GeV protons and anti-protons from the Main Injector and accelerates them to 980 GeV. The protons and antiprotons circle the TeVatron in opposite directions. The beams are brought to collision at two "collision points", B0 and D0. The two collider detectors, the Collider Detector at Fermilab (CDF-II) and D0 are built around the respective collision points.

The luminosity of collisions can be expressed as:

$$\mathcal{L} = \frac{f N_B N_p N_{\overline{p}}}{2\pi(\sigma_p^2 + \sigma_{\overline{p}}^2)} F\left(\frac{\sigma_l}{\beta^*}\right) \tag{4.1}$$

where f is the revolution frequency, N_B is the number of bunches, $N_{p/\overline{p}}$ are the number of protons/anti-protons per bunch, and $\sigma_{p/\overline{p}}$ are the rms beam sizes at the interaction point. F is a form factor which corrects for the bunch shape and depends on the ratio of σ_l, the bunch length to β^*, the beta function, at the interaction point. The beta function is a measure of the beam width, and is proportional to the beam's x and y extent in phase space.

Table 4.1 shows a comparison of Run I and design Run II[62] accelerator parameters. Figure 4.2 shows peak luminosities for stores used in this analysis.

4.2 The CDF-II Detector

The CDF-II detector [64] is a substantial upgrade of the original CDF-II detector [65]. It is located at the B0 collision point of the TeVatron collider. The detector is designed to detect and measure properties of particles emanating from $p\overline{p}$ collisions. The design of the detector is not geared toward one particular physics measurement, but rather optimized toward extracting as a number of different properties about all particle species created in the $p\overline{p}$ collision. Such particle detectors are often called multi-purpose detectors.

A diagram of the CDF-II detector is shown in Figure 4.3. A quadrant of the detector is cut out to expose the different subdetectors. The detector subsystems can be grouped as follows. The innermost system is the integrated tracking system. The tracking system is barrel-shaped

Parameter	Run I	Run II
number of bunches (N_B)	6	36
bunch length [m]	0.6	0.37
bunch spacing [ns]	3500	396
protons/bunch (N_p)	2.3×10^{11}	2.7×10^{11}
anti-protons/bunch ($N_{\overline{p}}$)	5.5×10^{10}	3.0×10^{10}
total anti-protons	3.3×10^{11}	1.1×10^{11}
$\beta^*(cm)$	35	35
interactions/crossing	2.5	2.3
integrated luminosity [pb^{-1}]	112	450
peak luminosity [$\mathrm{cm}^{-2}\mathrm{s}^{-1}$]	2×10^{31}	1.2×10^{32}

Table 4.1: *Accelerator parameters for Run I and Run II configurations.*

and consists of cylindrical subsystems which are concentric with the beam. It is designed to detect charged particles, measure their momenta and displacements from the point of collision (primary interaction vertex). The tracking system is surrounded by the Time of Flight system, designed to provide particle identification for low-momentum charged particles. Both the tracking and Time of Flight systems are placed inside a superconducting coil, which generates a 1.4 T solenoidal magnetic field. The coil is surrounded by calorimetry systems, which measure the energy of particles that shower when interacting with matter. The calorimetry systems are surrounded by muon detector systems. When interacting with matter, muons act as "minimally ionizing particles" - they only deposit small amounts of ionization energy in the material. Therefore, they are able to penetrate both the tracking and calorimeter systems. The integrated material of the tracking system, TOF, solenoid and calorimetry systems serves as a particle filter. Particles which penetrate through all that material are mostly muons, and they are detected by leaving tracks in the muon detection system, located outside of the calorimeter.

The most important parts of the detector for this analysis are the tracking system and the trigger, and these will be described in detail in the following sections. The description of the remaining systems will be brief. More detailed information on these systems can be found in the Technical Design Reports of the CDF-II [65, 64].

4.3 Standard Definitions in CDF-II

Because of its barrel-like detector shape, CDF-II uses a cylindrical coordinate system (r, φ, z) with the origin at the center of the detector and the z axis along the nominal direction of the proton beam. The y axis points upwards. Since the coordinate system is right-handed, this also defines the direction of the x axis, which is away from the center of the ring. Particles moving through a homogenic solenoidal magnetic field follow helical trajectories. Reconstructed particle trajectories are referred to as "tracks". The plane perpendicular to the beam is referred to as the "transverse plane", and the transverse momentum of the track is referred to as p_T. As opposed to $e^+ - e^-$ collisions, in $p - \overline{p}$ collisions, all of the center of mass energy of the $p - \overline{p}$ system is not absorbed in the collision. The colliding partons inside the proton carry only a

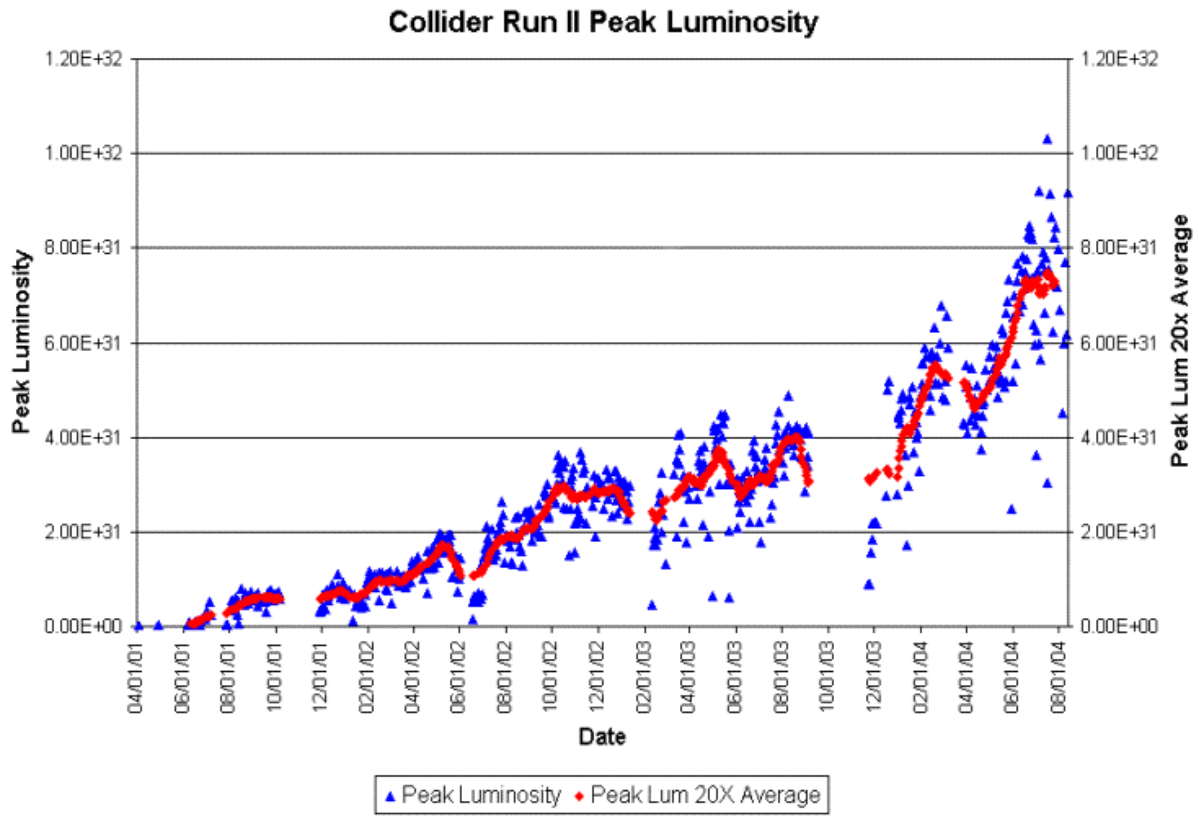

Figure 4.2: *Peak luminosities for stores collided between April 2001 and August 2004.*

fraction of the kinetic energy of the proton. As a result, the center of mass system of the parton collisions is boosted along the beam direction (the "longitudinal" direction) by an unknown amount. Quantities defined in the transverse plane are conserved in the collisions. For instance, the sum of all transverse momenta of particles in the collisions is zero ($\sum \vec{p_T} = 0$).

To uniquely parametrize a helix in three dimensions, five parameters are needed. The CDF-II coordinate system chooses three of these parameters to describe a position, and two more to describe the momentum vector at that position. The three parameters which describe a position describe the point of closest approach of the helix to the beam line. These parameters are d_0, φ_0, and z_0, which are the ρ, φ and z cylindrical coordinates of the point of closest approach of the helix to the beam. The momentum vector is described by the track curvature (C) and the angle of the momentum in the $r-z$ plane ($\cot\theta$). From the track curvature we can calculate the transverse momentum. The curvature is signed so that the charge of the particle matches the charge of the curvature. From $\cot\theta$, we can calculate $p_z = p_t \times \cot\theta$. At any given point of the helix, the track momentum is a tangent to the helix. This basically means that the angle φ_0 implicitly defines the direction of the transverse momentum vector at the point of closest approach p_T.

The impact parameter (d_0) of a track is another signed variable; its absolute value corresponds to the distance of closest approach of the track to the beamline. The sign of d_0 is taken

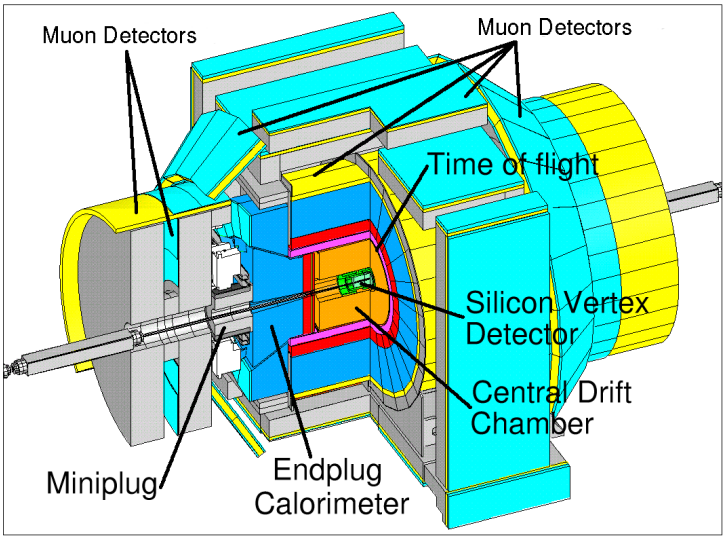

Figure 4.3: *The CDF-II detector with quadrant cut to expose the different subdetectors.*

to be that of $\hat{p} \times \hat{d} \cdot \hat{z}$, where $\hat{p}$, $\hat{d}$ and $\hat{z}$ are unit vectors in the direction of p_T, d_0 and z, respectively. An alternate variable that describes the angle between the z axis and the momentum of the particle is pseudorapidity (η) which is defined as:

$$\eta \equiv -\ln \tan(\theta/2) \tag{4.2}$$

For decaying particles, we often define the displacement L_{xy},

$$L_{xy} = \vec{d} \cdot \hat{p_T} \tag{4.3}$$

where $\vec{d}$ is the displacement of the decay vertex in the transverse plane, and $\hat{p_T}$ is the unit vector in the direction of $\vec{p_T}$.

4.4 Tracking Systems

The detector has a cylindrical tracking system immersed in a 1.4-T solenoidal magnetic field for the measurement of charged-particle momenta. We will describe this system starting from the devices closest to the beam and moving outwards. The innermost tracking device is a silicon strip vertex detector, which consists of three subdetectors. A layer of silicon sensors, called Layer 00 (L00) [66], is installed directly onto the beryllium vacuum beam pipe, at a radius of 1.7 cm from the beam. The beam pipe is made of beryllium because this metal has the best mechanical qualities, yet lowest nuclear interaction cross section of all materials.

The layer of silicon on the beam pipe is followed by five concentric layers of silicon sensors (SVXII) [67] located at radii between 2.5 and 10.6 cm. The Intermediate Silicon Layers (ISL) [68] are the outermost silicon subdetector systems, consisting of one layer at a radius of 22 cm in the central region and two layers at radii 20 and 28 cm in the forward regions. Surrounding the silicon detector is the Central Outer Tracker (COT) [69], a 3.1-m-long cylindrical open-cell drift chamber covering radii from 40 to 137 cm.

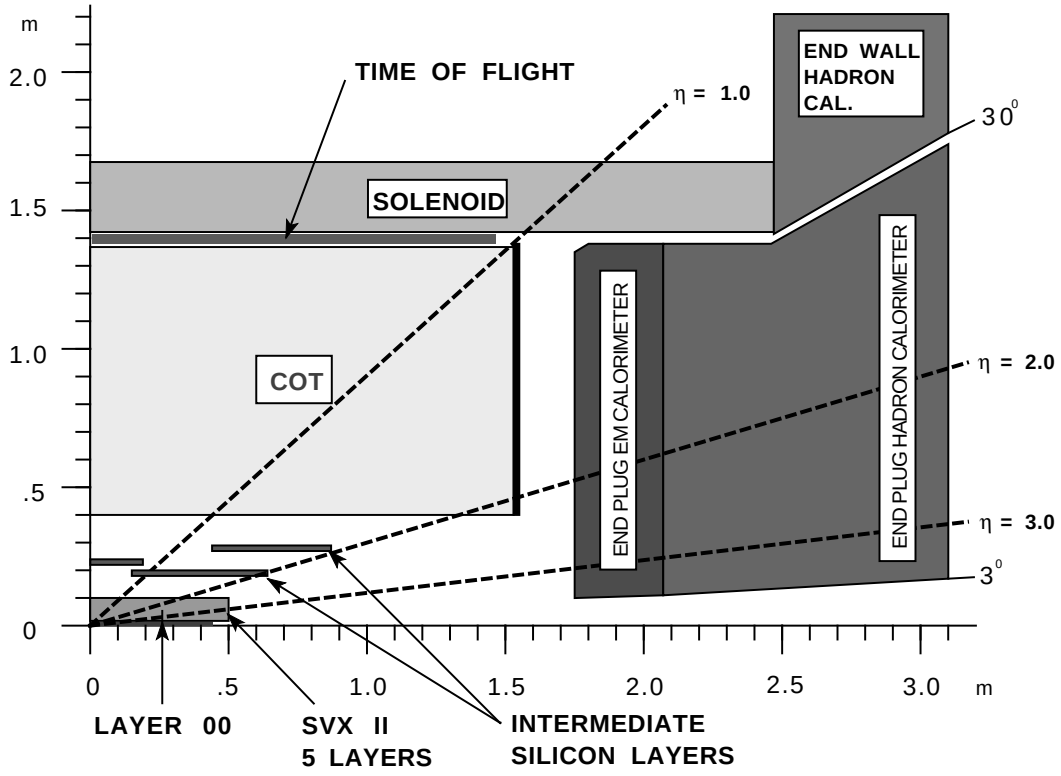

Figure 4.4: *A diagram of the CDF-II tracker layout showing the different subdetector systems.*

4.4.1 Silicon Tracking Detectors

Silicon tracking detectors are used to obtain precise position measurements of the path of a charged particle. A silicon tracking detector is fundamentally a reverse-biased p-n junction. When a charged particle passes through the detector material, it causes ionization. In the case of a semi-conductor material, this means that electron-hole pairs will be produced. Electrons drift towards the anode, and holes drift toward the cathode, where the charge is gathered. The amount of charge is, to first order, proportional to the path length traversed in the detector material by the charged particle.

By segmenting the p or n side of the junction into "strips" and reading out the charge deposition separately on every strip, we obtain sensitivity to the position of the charged particle. All the CDF-II silicon tracking detectors are implemented as micro-strip detectors. The typical distance between two strips is about 60 μm. Charge deposition from a single particle passing through the silicon sensor will be read out on one or more strips. This charge deposition is called a "cluster". There are two types of microstrip detectors: single and double-sided. In single-sided detectors, only one (p) side of the junction is segmented into strips. Double-sided detectors have both sides of the junction segmented into strips. The benefit of double-sided detectors is that while one (p) side has strips parallel to the z direction, providing $r-\varphi$ position measurements, the other (n) side can have strips at an angle (stereo angle) with respect to the z

direction, which will give z position information.

The innermost layer, L00, is made of single-sided silicon sensors which only provide $r-\varphi$ measurements. The SVX-II and ISL are made of double-sided silicon sensors. As shown in Table 4.2, the SVX-II layers have different stereo angles. Two layers have a small (1.2 $^\circ$) stereo angle and three have a 90° stereo angle. The ISL detector provides small-angle (1.2°) stereo information.

Property	Layer 0	Layer 1	Layer 2	Layer 3	Layer 4
number of φ strips	256	384	640	768	869
number of Z strips	256	576	640	512	869
stereo angle	90°	90 $^\circ$	+1.2°	90°	-1.2°
φ strip pitch	60 μm	62 μm	60 μm	60 μm	65 μm
Z strip pitch	141 μm	125.5 μm	60 μm	141 μm	65 μm
active width (mm)	15.30	23.75	38.34	46.02	58.18
active length (mm)	72.43	72.43	72.38	72.43	72.43

Table 4.2: *Relevant parameters for the layout of the sensors of different SVX-II layers.*

Four silicon sensors are stacked length-wise into a “ladder” structure which is 29 cm long. The readout electronics are mounted onto the ends of the ladders. The ladders are organized in an approximately cylindrical configuration, creating “barrels”. A SVX-II barrel is segmented into 12 wedges, each covering approximately 30° in φ with a small overlap at the edges, allowing for many Silicon hits per track. There are three SVX-II barrels, adjacent to each other along the z axis, covering the nominal interaction point in the center of the CDF-II detector. The coverage of the silicon detector subsystems is shown in Figure 4.5. The silicon tracking system is used in stand-alone mode to provide an extension of tracking down to 2.8 in pseudo-rapidity.

4.4.2 Central Outer Tracker

The COT drift chamber provides accurate information in the $r-\varphi$ plane for the measurement of transverse momentum, and substantially less accurate information in the $r-z$ plane for the measurement of the z component of the momentum, p_z. The COT contains 96 sense wire layers, which are radially grouped into eight “superlayers”, as inferred from the end plate section shown in Figure 4.6. Each superlayer is divided in φ into “supercells”, and each supercell has 12 sense wires and a maximum drift distance that is approximately the same for all superlayers. Therefore, the number of supercells in a given superlayer scales approximately with the radius of the superlayer. The entire COT contains 30,240 sense wires. Approximately half the wires run along the z direction (“axial”). The other half are strung at a small angle (2°) with respect to the z direction (“stereo”).

The active volume of the COT begins at a radius of 43 cm from the nominal beamline and extends out to a radius of 133 cm. The chamber is 310 cm long. Particles originating from the interaction point which have $|\eta| < 1$ pass through all 8 superlayers of the COT. Particles which have $|\eta| < 1.3$ pass through 4 or more superlayers.

The supercell layout, shown in Figure 4.7 for superlayer 2, consists of a wire plane containing sense and potential (for field shaping) wires and a field (or cathode) sheet on either side.

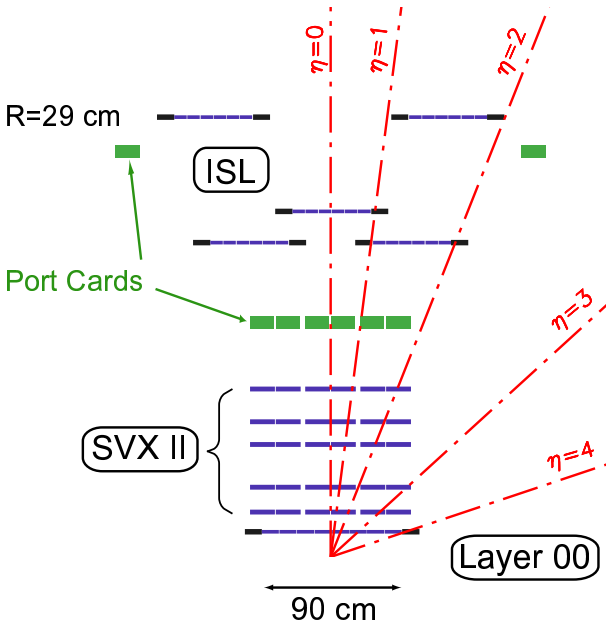

Figure 4.5: *Coverage of the different silicon subdetector systems projected into the $r-z$ plane. The r and z axes have different scales.*

Both the sense and potential wires are 40 μm diameter gold plated Tungsten. The field sheet is 6.35 μm thick Mylar with vapor-deposited gold on both sides. Each field sheet is shared with the neighboring supercell.

The COT is filled with an Argon-Ethane-CF4 (50:35:15) gas mixture. The mixture is chosen to have a constant drift velocity across the cell width. When a charged particle passes through, the gas is ionized. Electrons drift towards the sense wires. The electric field in a cylindrical system grows exponentially as its radius decreases. As a result, the electric field very close to the sense wire is large, resulting in an avalanche discharge when the charge drifts close to the wire surface. This effect provides a gain of $\sim 10^4$. The maximum electron drift time is approximately 100 ns. Due to the magnetic field that the COT is immersed in, electrons drift at a Lorentz angle of $\sim 35^\circ$. The supercell is tilted by 35° with respect to the radial direction to compensate for this effect.

Signals on the sense wires are processed by the ASDQ (Amplifier, Shaper, Discriminator with charge encoding) chip, which provides input protection, amplification, pulse shaping, baseline restoration, discrimination and charge measurement [70]. The charge measurement is encoded in the width of the discriminator output pulse, and is used for particle identification by measuring the ionization along the trail of the charge particle (dE/dx). The pulse is sent through ~ 35 ft of micro-coaxial cable, via repeater cards to Time to Digital Converter (TDC) boards in the collision hall. Hit times are later processed by pattern recognition (tracking) software to form helical tracks.

4.4.3 Pattern Recognition Algorithms

As explained in the previous sections, charged particles leave small charge depositions as they pass through the tracking system. By following, or "tracking" these depositions, pattern

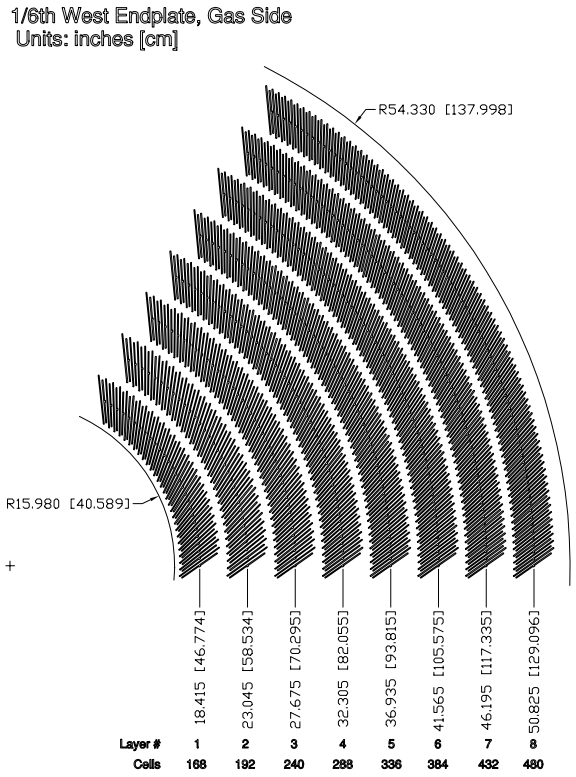

Figure 4.6: *Layout of wire planes on a COT endplate.*

recognition algorithms can reconstruct the charged particle track.

There are several pattern recognition algorithms used to reconstruct tracks in the CDF-II tracking system. Most of the tracks are reconstructed using "Outside-In" algorithms which we will describe here. The name of this group of algorithms suggest that the track is followed from the outside of the tracking system inward.

The track is first reconstructed using only COT information. The COT electronics report hit time and integrated charge for every wire in an event. The hit time corresponds to the time that an avalanche occurred at a sense wire. The hit time can be interpreted as the drift time of the charge in the gas, but first it has to be corrected for time of flight. The hit timing resolution is of the order of a few ns; this roughly corresponds to the average spread in collision times. It is assumed that the collision times always happen at the same time in a cycle during a store. An average of collision times is done for many previous events and this is used as the event collision time. Hit times corrected for the collision time are interpreted as drift times and used in pattern recognition. To perform the final track fit, an additional time of flight correction is performed assuming massless particles.

The helical track, when projected into the two dimensional $r-\varphi$ plane, is a circle. This simplifies pattern recognition, so the first step of pattern recognition in the COT looks for circular paths in radial superlayers of the COT. Super-cells in the radial superlayers are searched for sets of 4 or more hits that can be fit to a straight line. These sets are called "segments". The straight-line fit for a segment gives sufficient information to extrapolate rough measurements of curvature and φ_0. Once segments are found, there are two approaches to track finding.

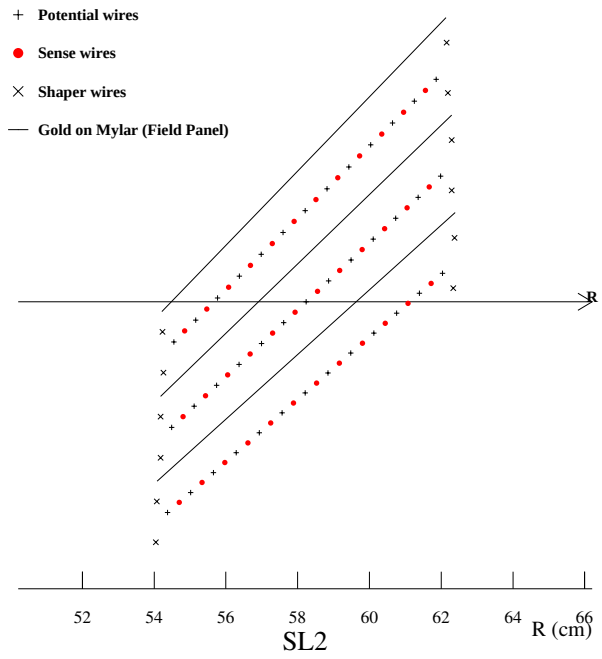

Figure 4.7: *Layout of wires in a COT supercell.*

One approach is to link together segments for which the measurements of curvature and φ_0 are consistent. The other approach is to improve the curvature and φ_0 measurement of a segment reconstructed in superlayer 8 by constraining its circular fit to the beamline, and then adding hits which are consistent with this path. Once a circular path is found in the $r-\varphi$ plane, segments and hits in the stereo superlayers are added by their proximity to the circular fit. This results in a three-dimensional track fit. Typically, if one algorithm fails to reconstruct a track, the other algorithm will not. This results in a high track reconstruction efficiency ($\sim 95\%$) in the COT for tracks which pass through all 8 superlayers ($p_T \geq 400$ MeV/c^2). The track reconstruction efficiency mostly depends on how many tracks there are to be reconstructed in the event. If there are many tracks present close to each other, hits from one track can shadow hits from the other track, resulting in efficiency loss.

Once a track is reconstructed in the COT, it is extrapolated into the SVX-II. Based on the estimated errors on the track parameters, a three-dimensional "road" is formed around the extrapolated track. Starting from the outermost layer, and working inward, silicon clusters found inside the road are added to the track. As a cluster gets added, the road gets narrowed according to the knowledge of the updated track parameters. Reducing the width of the road reduces the chance of adding a wrong hit to the track, and also reduces computation time. In the first pass of this algorithm, $r-\varphi$ clusters are added. In the second pass, clusters with stereo information are added to the track.

4.4.4 Momentum Scale

As the charged particle traverses through the tracker material, it loses energy. For a track that passes through the entire SVX-II volume, the amount of energy loss is roughly 9 MeV.

The value is roughly constant, regardless of the momentum of the particle. In the reconstructed distribution of invariant mass of $J/\psi \rightarrow \mu^+\mu^-$ decays, this effect will be more noticeable for low-momentum J/ψ decays than high-momentum decays. Figure 4.8 illustrates this effect. We use the momentum-dependence of the $\mu^+\mu^-$ invariant mass to calibrate the momentum scale of our detector. The J/ψ mass has to be invariant of transverse momentum and match with the world average [71] value if the momentum scale is correctly calibrated. Our calibration procedure follows two steps. First, the momentum dependence of the J/ψ mass is removed by correctly accounting for the energy loss in the tracker material, and then the overall shift of the J/ψ mass is removed by correcting the value of the magnetic field used in the conversion of curvature into transverse momentum.

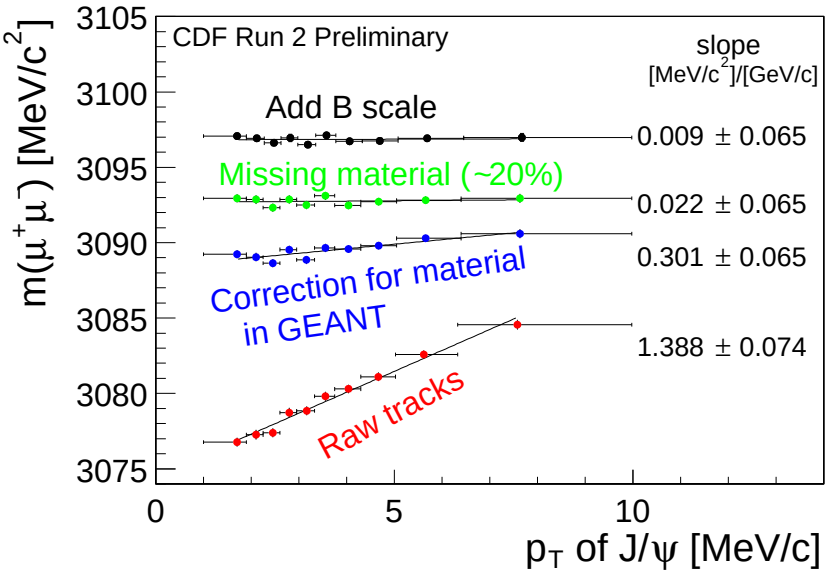

Figure 4.8: *Dependence of the reconstructed invariant mass of $J/\psi \rightarrow \mu^+\mu^-$ decays on the transverse momentum of the J/ψ.*

There are two types of material in the SVX-II tracker. The silicon sensors are read out and therefore called active material. Everything else in the silicon tracker (readout chips, cards, cables, cooling pipes) is passive material. The energy loss in the active material of the tracking system is taken into account by mapping out the material in the GEANT [72] description of our detector. The passive material in the detector description is not complete, so some energy loss is unaccounted for by this method. An additional layer of material is added to the detector description, to correct for the missing material on average. By tuning the amount of missing material, the momentum dependence of the J/ψ mass is removed. The remaining discrepancy with respect to the PDG average is corrected for by scaling the magnetic field. Because of the implementation of this procedure, we can not use it to measure the J/ψ mass, but the results of the calibration process (the amount of missing material and the corresponding magnetic field) can be used to correct the momentum scale in any other measurement.

4.5 Time of Flight

Outside the tracking system, still inside the superconducting magnetic coil, CDF-II has a Time of Flight (TOF) [73] system. The TOF system is designed to distinguish low momentum pions, kaons and protons by measuring the time it takes these particles to travel from the primary vertex of the $p\overline{p}$ collision to the TOF system. The system consists of 216 bars of scintillating material, roughly 300 cm in length and with a cross-section of 4×4 cm. The bars are arranged into a barrel around the COT outer cylinder. They are surrounded by superconducting solenoid on the outside. Particles passing through the scintillating material of the bars, deposit energy causing small flashes of visible light. This light is detected by photomultiplier (PMT) tubes which are attached at the end of each bar. The signal from the photomultiplier tube is processed by a pre-amplifier circuit mounted directly onto the tube. The amplified signal is sent via a twisted pair to the readout electronics in the collision hall. The readout electronics perform both time and amplitude digitization of the signal. The TDC information is a digitization of the time when the signal pulse reaches a fixed discriminator threshold. This time depends on the amplitude of the pulse, since a large pulse crosses the threshold earlier (time walk). The digitization of the pulse amplitude is needed to correct for this effect. After correcting for time walk effects, the timing resolution of the TOF system is currently about 110 ps for particles crossing the bar exactly in front of one of the photomultiplier tubes. The timing resolution varies with displacement from the photomultiplier tube. Large pulses give better timing resolution, and light attenuates while traveling through the scintillator material. Therefore, particles passing through the bar near the photomultiplier tube have better timing resolution than those which are farther away.

4.6 Calorimeters

The main effort of the Run II upgrade of the CDF-II calorimeter system dealt with upgrading the electronics to handle the faster bunch crossings. The active detector parts were taken over from Run I without modification. Since this analysis does not use calorimetry information, this system will be described briefly. A detailed description can be found in the CDF-II Technical Design Report [65].

The CDF-II calorimeter has a “projective tower” geometry. This means that it is segmented in η and φ “towers” that point to the interaction region. The coverage of the calorimetry system is 2π in φ and $|\eta| < 4.2$ in pseudo-rapidity. The calorimeter system is divided into three regions: central, plug and forward. Corresponding to these regions, the subsystems will have one of the letters C, P and F in their acronym. Each calorimeter tower consists of an electromagnetic shower counter followed by a hadron calorimeter. This allows for comparison of the electromagnetic and hadronic energies deposited in each tower, and therefore separation of electrons and photons from hadrons.

There are two subdetectors for the electromagnetic calorimeter: CEM and PEM. These correspond to the central and plug regions of $|\eta|$, respectively. The CEM uses lead sheets interspersed with scintillators as the active detector medium. The PEM uses proportional chambers. The hadron calorimeters in the central region are the central (CHA) and endwall (WHA). The plug region is covered by the PHA calorimeter. The CHA and WHA are composed of alternat-

System	$\|\eta\|$ coverage	Energy Resolution	Thickness
CEM	$\|\eta\| < 1.1$	$13.5\%/\sqrt{E_T} \oplus 3\%$	18 X_0
PEM	$1.1 < \|\eta\| < 2.4$	$28\%/\sqrt{E_T} \oplus 2\%$	18-21 X_0
CHA	$\|\eta\| < 0.9$	$50\%/\sqrt{E_T} \oplus 3\%$	4.5 λ_0
WHA	$0.7 < \|\eta\| < 1.3$	$75\%/\sqrt{E_T} \oplus 4\%$	4.5 λ_0
PHA	$1.3 < \|\eta\| < 2.4$	$90\%/\sqrt{E_T} \oplus 4\%$	5.7 λ_0

Table 4.3: *Pseudorapidity coverage, energy resolution and thickness for the different calorimeter subdetectors of the CDF-II experiment. The $\oplus$ symbol means that the constant term is added in quadrature to the resolution. λ_0 signifies interaction lengths and X_0 radiation lengths.*

ing layers of iron and scintillator. The PHA subdetector is made of alternating layers of iron and gas proportional chambers. The pseudorapidity coverage, resolutions and thickness for the different electromagnetic and hadron calorimeters is given in Table 4.3.

4.7 Muon Systems

Muons are particles which interact with matter only by ionization. For energies relevant to this experiment, they do not cause showers in the electromagnetic or hadronic calorimeters. As a result, if a muon is created in the collision and has enough momentum, it will pass through the calorimeter with minimal interaction with the material inside. Therefore, the calorimeter can be considered as a "filter" which retains particles that shower when interacting with matter and muons, and passes muons, which do not. Muon detection systems are therefore placed radially outside the calorimeters.

The CDF-II detector has four muon systems: the Central Muon Detector (CMU) [74], Central Muon Upgrade Detector (CMP) [75], Central Muon Extension Detector (CMX) [76], and the Intermediate Muon Detector (IMU). The CMU and CMP detectors are made of drift cells, and the CMX detector is made of drift cells and scintillation counters, which are used to reject background based on timing information. Using the timing information from the drift cells of the muon systems, short tracks (called "stubs") are reconstructed. Tracks reconstructed in the COT are extrapolated to the muon systems. Based on the projected track trajectory in the muon system, the estimated errors on the tracking parameters and the position of the muon stub, a χ^2 value of the track-stub match is computed. To ensure good quality of muons, an upper limit is placed on the value of χ^2_φ, the χ^2 of the track-stub match in the φ coordinate.

Most of the particles that pass through the calorimeter without showering are muons, but it is also possible for pions or kaons to survive the passage. These particles can then fake muon signals in the muon chambers. Typically, these fake rates are at the percent level, as seen in Figure 4.9 for the CMU and CMP detectors combined. The Figure shows the rate at which charged pions and kaons fake muon signals in the muon systems. The difference between K^+ and K^- rates comes from the different cross section for interaction of these two mesons with the calorimeter material. The different interaction cross section for these two mesons comes from their quark content. In the K^+, the strange quark is the anti-quark.

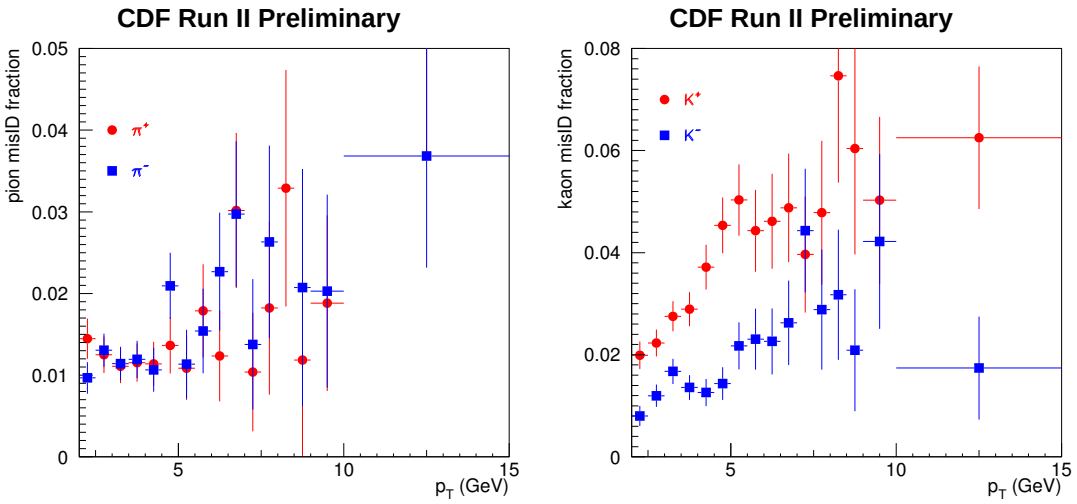

Figure 4.9: *Rate of kaon and pion tracks faking muon signals in the CDF-II detector. Roughly 1% of all pions(left) and $2-4\%$ of all kaons (right) will fake a muon signal.*

4.8 Triggering

Triggering systems are necessary because it is not physically possible to store information about every single $p\overline{p}$ collision. Collisions happen roughly at a rate of 2.5 MHz, and the readout of the full detector produces an event roughly the size of 250 kB. There is no medium available which is capable of recording data this quickly, nor would it be practical to analyze all this data later on. The trigger system is a pre-filter, which reduces data rates and volumes to manageable levels, according to all foreseen physics prescriptions.

The CDF-II triggering system is designed based on three conditions. The first condition is that the trigger be deadtimeless. This means that the trigger system has to be quick enough to make a decision for every single event, before the next event occurs. The second condition is imposed by the TeVatron upgrade for Run II, and it is the minimum expected time between collisions, 132ns. The last condition is that the data logging system can write about 30-50 events per second to tape, because of limited resources. In short, the trigger has to be fast enough to analyze every collision, and it has to figure out which 50 of 2.5 million events it should save in a given second. This is achieved by staging trigger decisions in three levels, as shown in Figure 4.10.

Each level of the trigger is given a certain amount of time to reach a decision about accepting or rejecting an event. By increasing the time allowed for triggering at different levels of the trigger, the complexity of reconstruction tasks can be increased at every level. At the first level of the trigger, only very rough and quick pattern recognition and filtering algorithms are used. In order to do this in time, the Level 1 and Level 2 triggering mechanisms are implemented with custom electronics. The third level of the trigger is implemented with a PC farm with about 300 CPU's. Using each CPU as an event buffer allows for nearly one second to be allocated for the trigger decision. As a result, nearly offline quality of event reconstruction is available at the third level of triggering. The Level 3 rejection rate is about 10, resulting in 30 events/sec being accepted by the Level 3 trigger and written to tape.

The delay necessary to make a trigger decision is achieved by storing detector pre-readout information in a storage pipeline. At Level 1, for every TeVatron clock cycle, the event is moved up one slot in the pipeline. By the time it reaches the end of the pipeline, the trigger will have reached a decision whether to accept or reject this event. If the event is accepted, its information will be sent to the higher level of the trigger. Otherwise, the event is simply ignored. Since the Level 1 buffer has 42 slots, the time allocated for making a trigger decision is about 5 μs. The rejection factor after Level 1 is about 150, so the Level 1 accept rate is below 50 kHz. At Level 2, there are 4 event buffers available. This allows for 20 μs for the trigger decision. The Level 2 rejection factor is again around 150, and the accept rate is around 300 Hz.

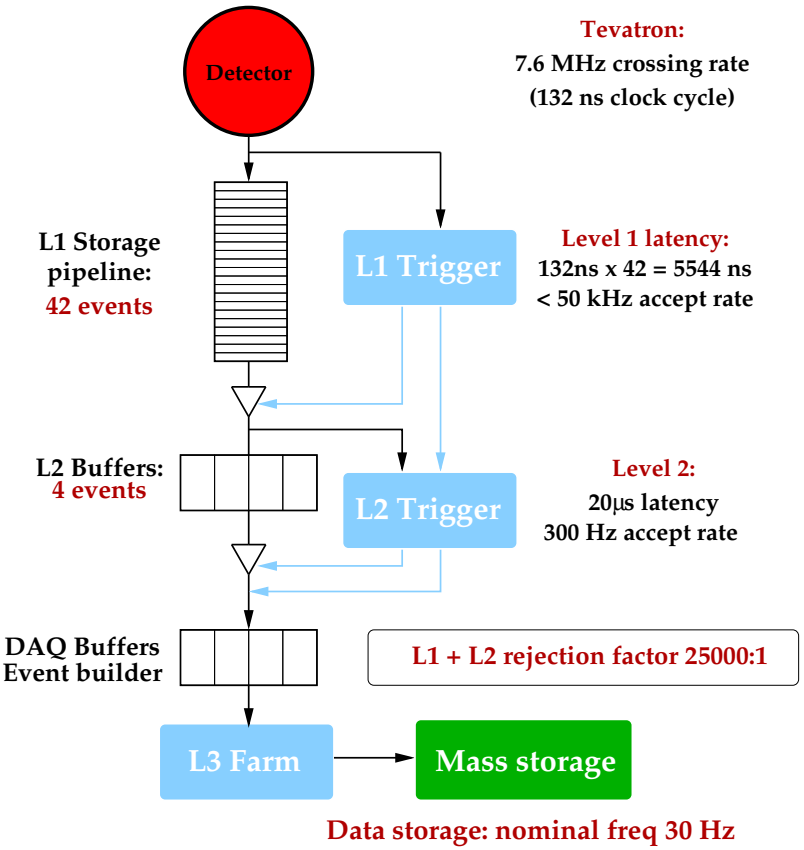

Figure 4.10: *Diagram of the CDF-II trigger system.*

A set of requirements that an event has to fulfill at Level 1, Level 2 and Level 3 constitutes a trigger path. Requiring that an event be accepted through a well defined trigger path unfortunately eliminates volunteer events. A volunteer event is an event which passed a higher level (L2, L3) trigger requirement but did not pass the preceding lower level (L1, L1/L2) trigger requirement. The CDF-II trigger system implements about 100 trigger paths. An event will be accepted if it passes the requirements of any one of these paths. The trigger path used in this analysis is the "Two Track" trigger path, which is shown in Figure 4.11, and which we will describe in detail here. The trigger path is optimized for finding charm and bottom mesons that decay in hadronic final states. The strategy of the trigger path is as follows. At Level 1, rough measurements of track momenta are available. By cutting on track momenta and angles, most of the inelastic background will be rejected. At Level 2, the additional time available for

reconstruction is utilized to use SVX-II information and obtain better impact parameter measurements of the tracks. Requiring non-zero impact parameters of tracks will require that they come from decays of long-lived particles: charmed and bottom mesons.

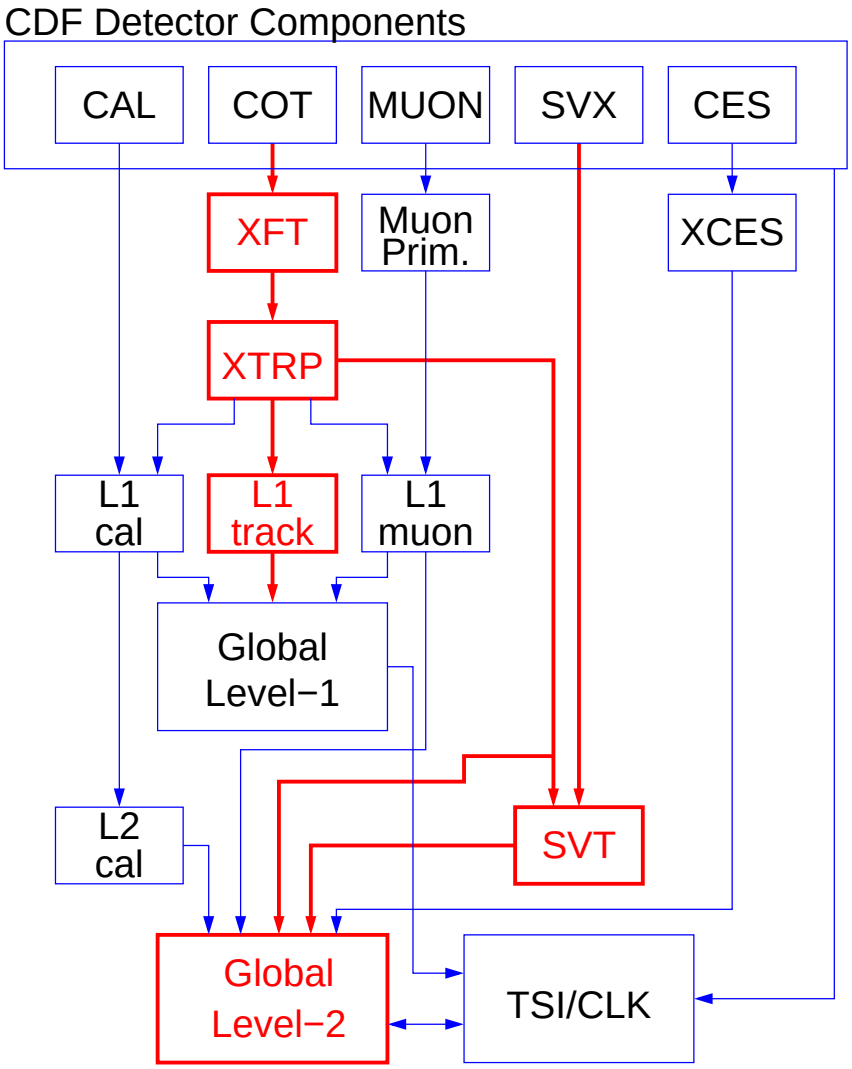

Figure 4.11: *Diagram of the different trigger paths at Level 1 and 2. The data flow for the track trigger paths is outlined in red.*

4.8.1 Level 1 Trigger

The Level 1 trigger decision is based on the information from the eXtremely Fast Tracker (XFT) [77]. This device examines the hit information of the COT in wedges of 15 degrees. It reports the measurement of the track p_T and φ_6, the angle of the transverse momentum at the sixth superlayer of the COT, which is located 106 cm radially from the beamline. Based on pre-loaded patterns of COT hits, it is capable of recognizing track segments for tracks with $p_T > 1.5$ GeV/c^2 in 15° wedges of the COT. Two tracks are reported from a given 15° wedge: the two tracks which are closest to the left and right boundary of a given wedge. As mentioned

in Section 4.4.3, based on the hit information from a single COT cell, assuming that the track comes from the beamline, a rough measurement of the track p_T and φ_6 is obtained. This is the information that the XFT device determines per track. An event is accepted at Level 1 if two tracks are found in the event such that they have opposite charge, both tracks have $p_T > 2$ GeV/c, the scalar sum of transverse momenta $p_{T1} + p_{T2} > 5.5$ GeV/c and the φ separation between the tracks at superlayer 6 is $|\Delta\varphi_6| < 135^\circ$.

4.8.2 Level 2 Trigger

At Level 2, rough tracking information from the XFT is combined with SVX-II cluster information by the Silicon Vertex Tracker (SVT) [78]. The goal of the second level of the trigger is to obtain a precise measurement of the track d_0, and improved measurements of p_T and φ_0.

Figure 4.12 shows the principle of SVT operation. As mentioned in Section 4.4.1, the SVX-II is segmented into 12 wedges in φ and three mechanical barrels in z. The SVT makes use of this symmetry and does tracking separately for each wedge and barrel. Tracks which cross wedge and barrel boundaries are only reconstructed under certain circumstances. An SVT track starts with a two dimensional XFT "seed". The XFT measurement is extrapolated into the SVX-II, forming a "road". Clusters of charge on the inner four $r-\varphi$ layers of the given wedge have to be found inside this road. The silicon cluster information and the XFT segment information are fed into a linearized fitter which returns the measurements of p_t, φ_0 and d_0 for the track.

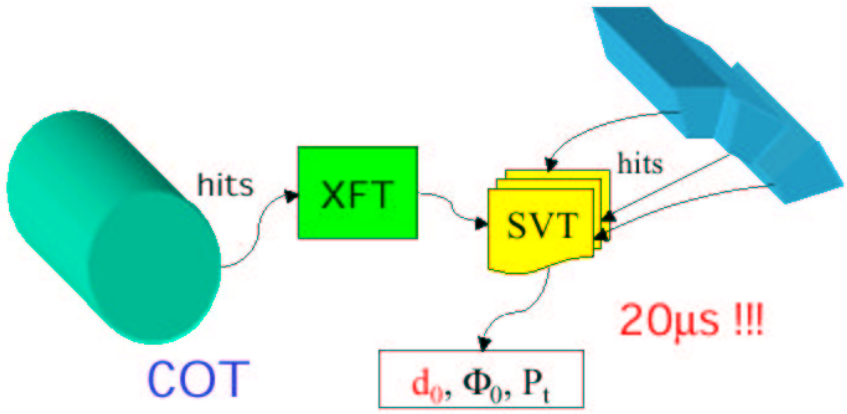

Figure 4.12: *SVT principle of operation.*

As shown in Figure 4.13, the track impact parameter resolution is about 35 μm for tracks with $p_T > 2$ GeV/c. The width of the Gaussian fit for the distribution of track impact parameters in Figure 4.13 is 47 μm. This width is a combination of the intrinsic impact parameter resolution of the SVT measurement, and the transverse intensity profile of the interaction region. The region profile is roughly circular in the transverse plane and can be approximated by a Gaussian distribution with $\sigma \sim 35$ μm. The intrinsic SVT resolution is obtained by subtracting the beamline width from the width of the d_0 distribution in quadrature. The Level 1 trigger conditions are confirmed with the improved measurements of p_T and φ_0. An event passes Level 2 selection if there is a track pair reconstructed in the SVT such that the tracks have opposite

charge, each track has $p_T > 2.0\text{GeV}/c$ and $120\mu\text{m} < |d_0| < 1\text{mm}$. The vertex of the track pair has to have $L_{xy} > 200\mu\text{m}$ with respect to the beamline.

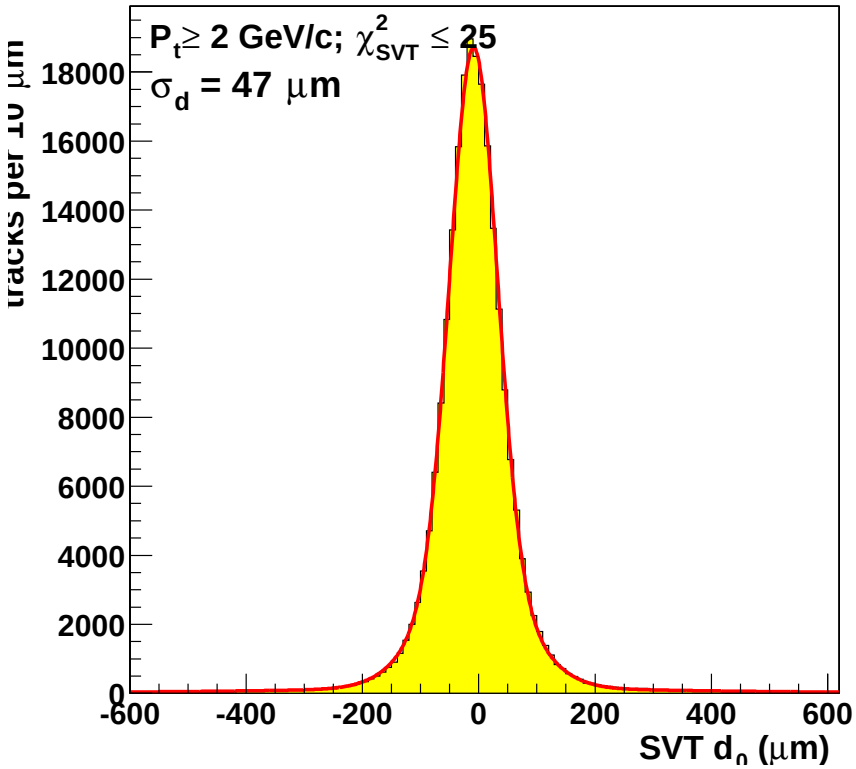

Figure 4.13: *SVT impact parameter resolution.*

4.8.3 Level 3 Trigger

The third level of the trigger system is implemented as a PC farm. Every CPU in the farm provides a processing slot for one event. With roughly 300 CPUs, and a input rate of roughly 300 Hz, this allocates approximately 1 second to do event reconstruction and reach a trigger decision.

Figure 4.14 shows the implementation of the Level-3 farm. The detector readout from the Level 2 buffers is received via an Asynchronous Transfer Mode (ATM) switch and distributed to 16 “converter” node PC’s , shown in Figure 4.14 in light blue. The main task of these nodes is to assemble all the pieces of the same event as they are delivered from different subdetector systems through the ATM switch. The event is then passed via an Ethernet connection to a “processor” node, which there are about 150 in the farm and are shown in Figure 4.14 in green. Each processor node is a separate dual-processor PC. Each of the two CPU’s on the node process a single event at a time. The Level 3 decision is based on near-final quality reconstruction performed by a “filter” executable. If the executable decides to accept an event, it is then passed to the “output” nodes of the farm, which are shown in Figure 4.14 in yellow. These nodes send the event onward to the Consumer Server / Data Logger (CSL) system for storage first on disk, and later on tape.

For most of the data used for this analysis, full COT tracking was being used to reconstruct tracks. The measurements of p_T, z_0, φ_0 and $cot\theta$ from the COT are combined with the d_0 measurement from the SVT to create a further improved track. The Level 1 and Level 2 trigger conditions (including the requirement on the two-track vertex L_{xy}) are repeated at Level 3 using improved track measurements. For later data, which is not used in this analysis, full SVX-II tracking is available, and the trigger conditions are repeated using a combined COT/SVX-II fit of the track helices.

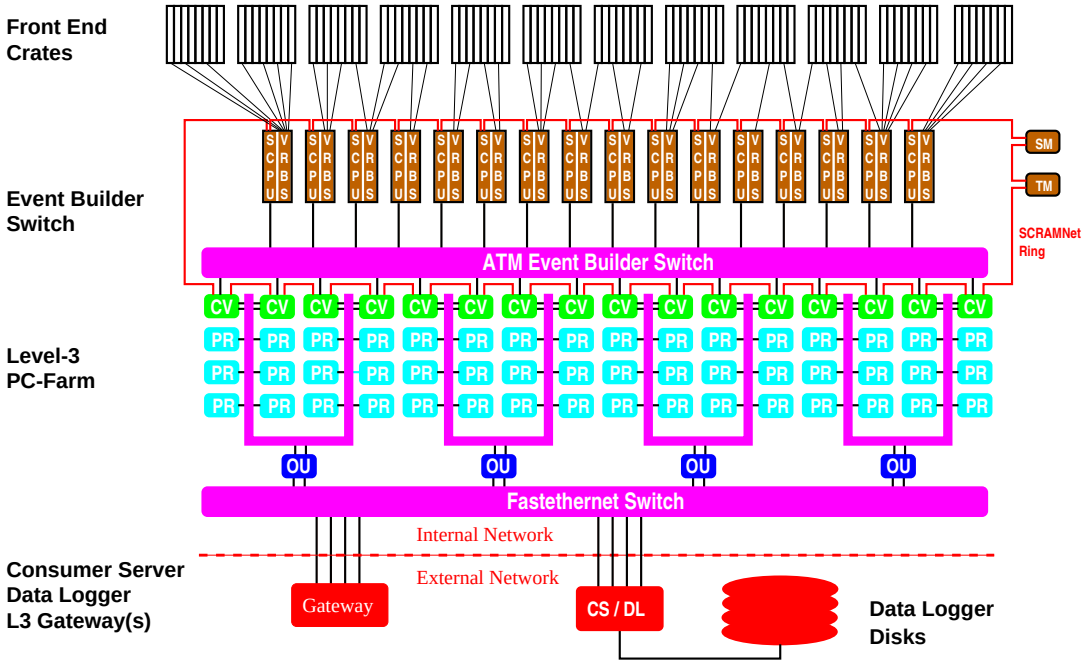

Figure 4.14: *Event building and Level 3 operating principle: data from the front end crates is prepared by Scanner CPU's (SCPU) and fed into the ATM switch (purple). On the other side of the switch, converter nodes (green) assemble events and pass them to processor nodes. Accepted events are passed to output nodes (dark blue) which send them to the Consumer Server and data logging systems (red).*

4.9 Luminosity Measurement

At hadron collider experiments the beam luminosity can be measured using the process of inelastic $p\overline{p}$ scattering. It has a large cross section, $\sigma_{in} \sim 60\,mb$. The rate of inelastic $p\overline{p}$ interactions is given by:

$$\mu \cdot f_{BC} = \sigma_{in} \cdot L \tag{4.4}$$

where L is the instantaneous luminosity, f_{BC} is the rate of bunch crossings in the TeVatron and μ is the average number of $p\overline{p}$ interactions per bunch crossing.

To detect inelastic $p\overline{p}$ events efficiently, a dedicated detector at small angles, operating at high rate and occupancy, is required. The Cherenkov Luminosity Counters (CLC) are being used by CDF to measure the TeVatron luminosity. The CLC is designed to measure μ accurately

(within a few percent) all the way up to the high luminosity regime $L \sim 2 \times 10^{32} cm^{-2} s - 1$. The CLC modules and the luminosity measurement method are described in detail in [79].

There are two CLC modules in the CDF detector, installed at small angles in the proton (East) and anti-proton (West) directions with rapidity coverage between 3.75 and 4.75. Each module consists of 48 thin, long, gas-filled, Cherenkov counters. The counters are arranged around the beam-pipe in three concentric layers, with 16 counters each, and pointing to the center of the interaction region. The cones in two outer layers are about 180 cm long and the inner layer counters (closer to the beam pipe) have the length of 110 cm. The Cherenkov light is detected with fast, 2.5 cm diameter, photomultiplier tubes. The tubes have a concave-convex, 1 mm thick, quartz window for efficient collection of the ultra-violet part of Cherenkov spectra and operate at a gain of 2×10^5.

The counters are mounted inside a thin pressure vessel made of aluminum and filled with isobutane. The Cherenkov angle is 3.1^o and the momentum threshold for light emission is $9.3 MeV/c$ for electrons and $2.6 GeV/c$ for pions.

The number of $p\overline{p}$ interactions in a bunch crossing follows Poisson statistics with mean μ, where the probability of empty crossings is given by $P_0(\mu) = e^{-\mu}$. An empty crossing is observed when there are fewer than two tubes with signals above threshold in either module of the CLC. The measured fraction of empty bunch crossings is corrected for the CLC acceptance and the value of μ is calculated. The measured value of μ is combined with the inelastic $p\overline{p}$ cross section to determine the instantaneous luminosity using Equation (4.4).

The CLC is one of the upgrades of the CDF Run II detector. It is designed to provide an improved measurement of the luminosity with respect to the device used in Run I. The luminosity measured by the CLC is used to monitor the TeVatron's performance.

Chapter 5

B Fragmentation at DELPHI

5.1 General Event Selection and Jet Energy Measurement

Hadronic Z decays have been selected by applying the following conditions:

- $|\cos(\theta_{thrust})| < 0.95$, where θ_{thrust} is the angle of the event thrust axis relative to the beam direction,
- at least 15 tracks, charged and neutrals, reconstructed,
- track quality cuts: neutrals must have a momentum larger than 500 MeV/c, charged tracks must have a momentum larger than 100 MeV/c, a length longer than 20 cm and an estimated uncertainty on the momentum smaller than the momentum value itself. Momentum of charged tracks is limited to the value of 35 GeV/c. In real data, 18% of VD-only tracks are removed, randomly, so that their abundance becomes similar to the simulation. VD-only tracks have been reconstructed using only the hits remaining unassociated to other track trajectories, in the layers of the vertex detector.

As explained in Section 2.1, charged particles from *b*-hadron decays can be separated from other charged hadrons using their positive impact parameter measured relative to the event main vertex. For a hadronic event resulting from the hadronization of light quarks, charged track impact parameters are expected to be compatible with the beam interaction position [1]. Using the probability distribution of negative significances, $f(S)$, as measured in data and supposed to be symmetric $f(S) = f(-S)$, one determines the probability, $P(S^0)$, for a track to have a significance $|S| > S^0$:

$$P(S^0) = \int_{S^0}^{\infty} f(S)dS \tag{5.1}$$

The values of $P(S^0)$ are expected to have a flat distribution for tracks originating from the primary vertex, whereas this distribution is peaked at low values for those having large positive significances. This approach can be generalized for a set of N tracks. A *b*-tagging package has been developed in DELPHI which includes the information from track impact parameters and

[1] Hadrons interacting in the material surrounding the interaction point or originating from the weak decay of low energy pions or kaons can have non-zero impact parameters. Usually most of them can be recognized and eliminated by considering tracks associated with a minimum number of hits in the silicon vertex detector.

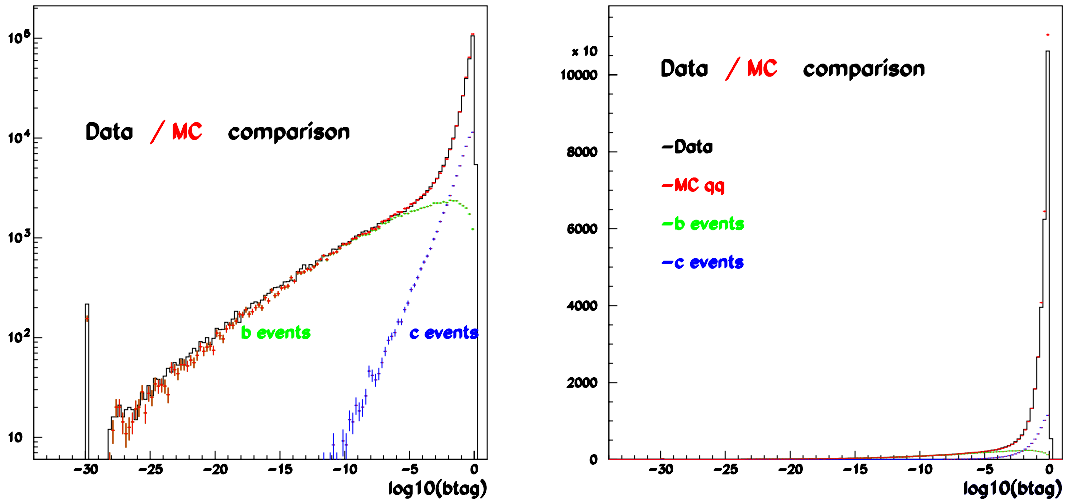

Figure 5.1: *Distribution of* $-\log_{10}\left(P_{btag}\right)$ *in log- (left) and lin- (right) scale.*

other variables having a different behavior for b and non-b events [80]. It provides a quantity P_{btag} which can be used to select event samples with different level of purity in b-events.

In Figure 5.1 distributions of this quantity have been compared for real and simulated events. In the following, samples of hadronic events depleted in b flavor have been selected by a cut on the b-tagging probability ($P_{btag} \geq 10$ %) evaluated for the whole event whereas b-enriched samples have been retained using $P_{btag} \leq 10^{-3}$. In Table 5.1, the fraction of selected events in data and simulation, the expected fraction of non-b events and the efficiency for b-events are given. According to these values, it is possible to isolate samples of hadronic events, containing less than 10% contamination from non-b events, with an efficiency higher than 50% for those originating from b-quarks. Remaining differences between real and simulated events will be included in the evaluation of systematics.

cut on $P_(btag)$	$< 10^{-3}$	$< 10^{-4}$	$< 10^{-5}$	$< 10^{-6}$	$< 10^{-8}$	$< 10^{-10}$
Data, fract. of sel. events (%)	17.6	14.3	11.8	9.9	6.8	4.5
MC, fract. of sel. events (%)	17.2	14.0	11.5	9.5	6.4	4.2
MC, b-purity (%)	88.7	93.5	96.1	97.6	99.0	99.9
MC, b-efficiency (%)	69.4	59.3	50.1	41.9	28.6	19.1

Table 5.1: *Variation of the selected event sample composition and efficiency for b events, versus the cut on the btag-variable.*

5.1.1 Data/Monte-Carlo comparison and adjustments

A few differences between real and simulated events have been identified and corrected to some extent.

5.1.1.1 Accuracy of track reconstruction

As the reconstruction accuracy for charged tracks depends on the type of attached sub-detectors, and as more of such detectors are associated, in the simulation, as compared to data, some of these detectors have been removed. This was done by rescaling the values of measurement errors and by smearing the corresponding track parameter values. These changes depend on the type of the removed sub-detector and were determined using the simulation.

Changes applied on the 1994 events sample are given in Table 5.2.

Simulated events			
attached sub-detectors	absolute fraction	kept sub-detectors	relative fraction
VD-ID-TPC-OD	29.8%	ID-TPC-OD	1.5%
		VD-TPC	5.5%
		VD-TPC-OD	3.0%
VD-ID	3.6%	VD	6.0%
VD-ID-TPC	31.4%	ID-TPC	1.0%
Real events			
attached sub-detectors	absolute fraction	kept sub-detectors	relative fraction
VD-TPC	4.7%	TPC	10.%
ID-TPC	4.2%	TPC	5.%

Table 5.2: *Fractions of tracks adjusted in simulated and real events so that the fractions of attached subdetectors are similar for the two samples.*

An additional smearing corresponding to 20% of the measurement uncertainty has been applied on simulated tracks to account for the fact that weakly decaying particle reconstructed signals (as those corresponding to the D^0 or D^+) have a width which is larger in real data by this amount.

5.1.1.2 Efficiency and track energy distribution

After these corrections have been applied, track energy distributions have been compared in real and simulated events. These distributions have been normalized to the number of hadronic events. In Figure 5.2 ratios of these distributions (Data/MC) have been displayed versus $\log_{10}(p)$ separately for charged and neutral tracks. These distributions are given for *b*-depleted and *b*-enriched samples. They show similarities at low momentum. For *b*-enriched samples, one can notice that there is a variation versus the assumed fragmentation model for *b*-hadrons. Plots have been provided for the default fragmentation function used in the DELPHI simulation (a Peterson distribution with parameter equal to 2.29 10^{-3}), for the fitted distribution as obtained on data, for a softer Peterson distribution (with parameter equal to 6 10^{-3}) and for the distribution extracted by the Karlsruhe group analysis [34].

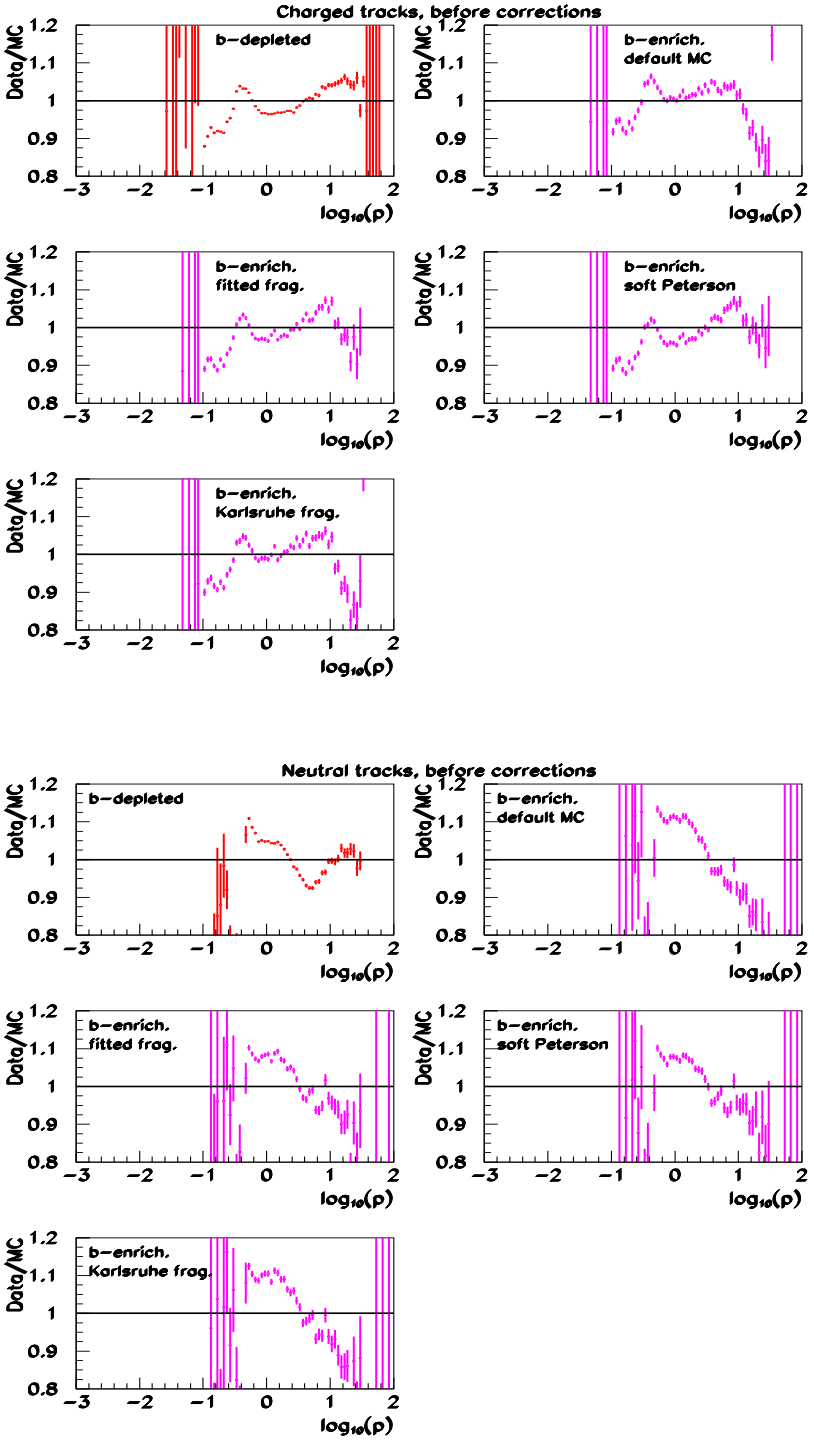

Figure 5.2: *Ratio Data/MC for charged (top) and neutral (bottom) track momentum distributions ($\log_{10} p$). Plots in upper left corners correspond to b-depleted samples. The others are obtained using b-enriched samples and different models for the b-fragmentation distribution.*

To match Data/MC distributions, the correction consists in removing tracks alternatively in Data or in the simulation, depending if the studied ratio is larger or lower than unity. Such corrections have been determined, separately, for *b*-depleted and *b*-enriched samples.

To avoid a possible bias induced by a correlation between the assumed shape of the fragmentation function and the applied correction, the correction has been evaluated iteratively using, as input in the determination of the correction, the fragmentation distribution measured at the previous step. In practice one iteration was used as the observed absolute variation, between second and first step, on the fitted $< x_E >$ value was of the order of 10^{-3}.

Track momentum distributions obtained after having applied these corrections are given in Figure 5.3.

5.1.2 Jet energy reconstruction

Jets are reconstructed using the Lund LUCLUS algorithm with the *d join* parameter value set to 5.0 GeV/c.

A first evaluation of the jet energies is obtained using the jet directions, energies, masses and imposing total energy-momentum conservation for the whole event.

If the missing energy, in a jet, is larger than 1 GeV, a four-vector is added to the jet. Its direction is taken to be the same as the jet direction and the missing momentum is evaluated assuming that the missing particle mass is zero.

Energy momentum conservation is then applied again to the whole event and track parameters (for charged, neutral and eventually missing) are fitted. It has been assumed, analyzing simulated events, that the relative uncertainty on the missing energy is 20% and that uncertainties on angles of the missing particle are 50 mrad (it has been checked that similar results are obtained assuming 50% relative error on the energy and 100 mrad error on the angles). If after the fit, 4-vectors corresponding to missing particles have an energy lower than 1 GeV they are removed and a new fit is done.

After this procedure, 4-vectors of charged and neutral tracks have been fitted and possible new 4-vectors corresponding to missing energy in each jet have been obtained. Energy-momentum conservation for the whole event is exactly imposed, introducing a non-zero mass for the missing-energy 4-vectors. Jets are reevaluated using this set of tracks and applying the same LUCLUS algorithm.

In Figure 5.4 have been compared the jet energy distributions, keeping the three highest energy jets in each event, obtained in data and in the simulation after the energy-momentum constrained fit.

In Figures 5.5 and 5.6 have been compared the fitted charged, neutral and missing energy fractions for *b*-depleted and *b*-enriched events selected in real and simulated events. These fractions are also compared in Table 5.3. Relative differences are at the level of a few 10^{-3}.

A comparison has been made also between the average multiplicities of charged and neutral tracks and of their variance, results are given in Table 5.4.

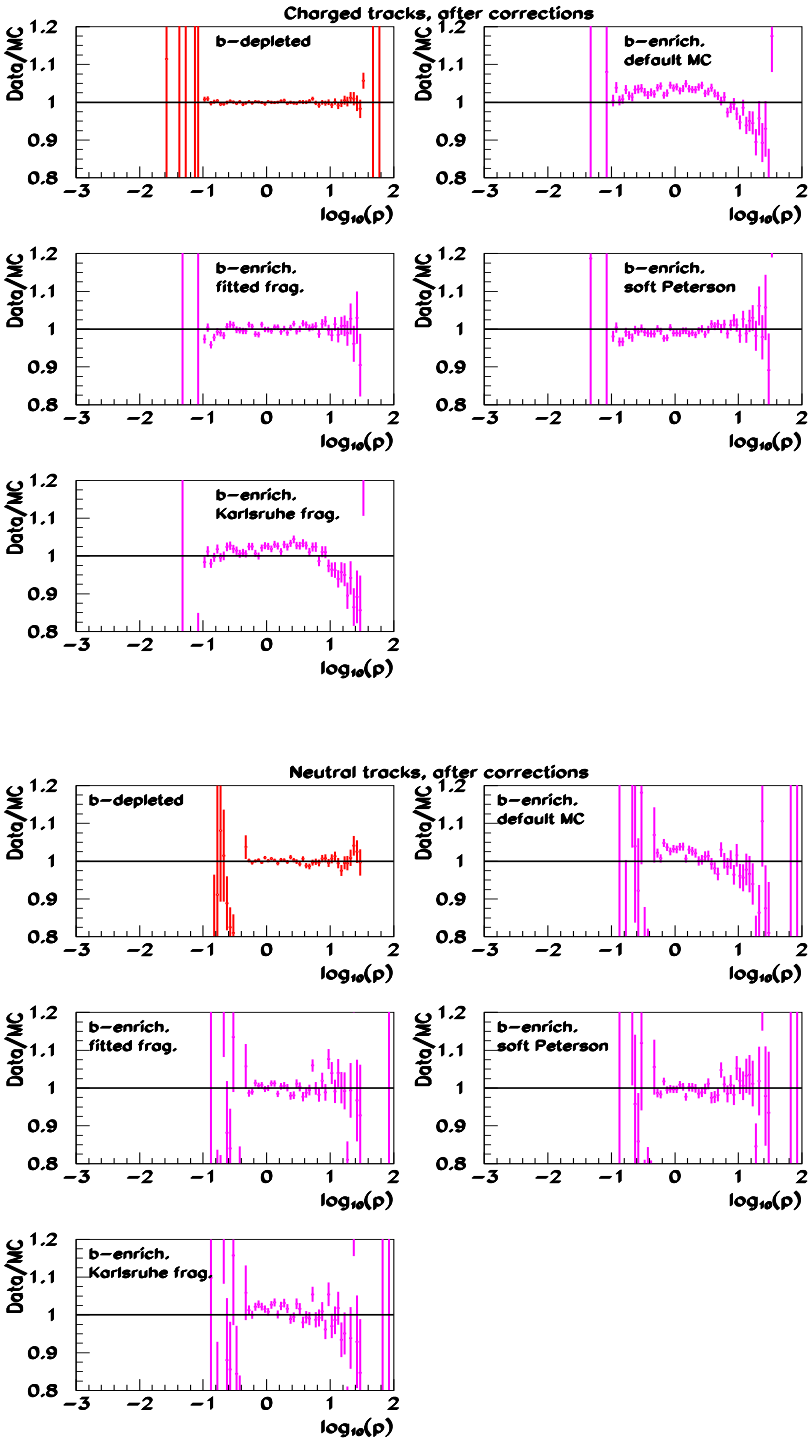

Figure 5.3: *Ratio Data/MC for charged (top) and neutral (bottom) track momentum distributions ($\log_{10} p$) after having applied corrections.*

b-depleted events				
Sample	Ech + Eneu	Ech	Eneu	Emiss
Data	0.8644	0.5759	0.2879	0.1363
MC	0.8676	0.5778	0.2893	0.1323
(Data-MC)/MC	-0.0037	-0.0033	-0.0048	$+0.030$
b-enriched events				
Sample	Ech + Eneu	Ech	Eneu	Emiss
Data	0.8423	0.5891	0.2528	0.1589
MC	0.8423	0.5885	0.2535	0.1579
(Data-MC)/MC	0.0000	$+0.0010$	-0.0027	$+0.0063$

Table 5.3: *Fitted fractions of charged (Ech) and neutral (Eneu) energies, and of their sum, reconstructed in b-depleted and b-enriched event samples. The missing energy fitted fraction is also given.*

b-depleted events		
Sample	charged	neutrals
Data	22.93 (7.94)	10.47 (3.84)
MC	22.96 (7.62)	10.56 (3.82)
b-enriched events		
Sample	charged	neutrals
Data	25.34 (7.60)	10.95 (3.80)
MC default	24.74 (7.45)	10.88 (3.82)
MC fitted	25.16 (7.48)	10.96 (3.83)

Table 5.4: *Charged and neutral track multiplicities (variance) measured in data and simulation, after corrections.*

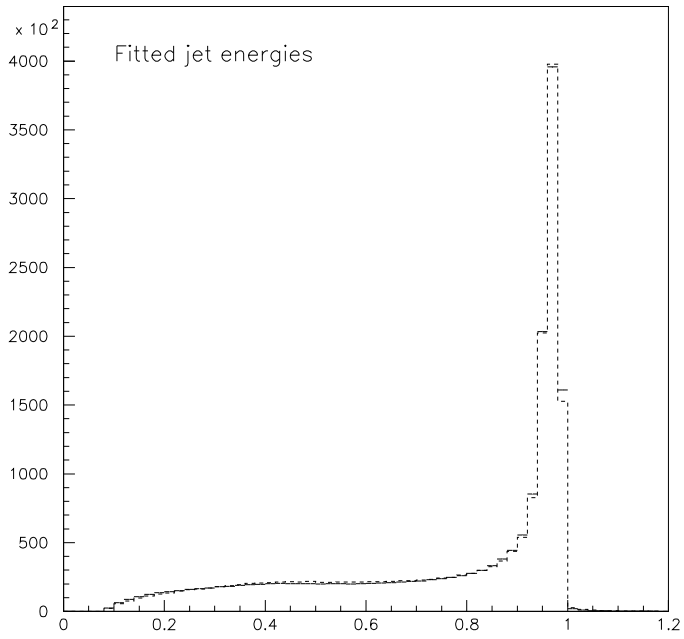

Figure 5.4: *Comparison, for a depleted b sample, between data (points with error bars) and simulation (dashed histogram) for the fitted jet energy, normalized to the beam energy. The three higher energy jets are kept in each event.*

5.2 B-Energy Reconstruction

Each jet, pointing through the detector barrel region defined by $\left|\cos\theta_{jet}\right| < 0.75$, is considered in turn and charged tracks belonging to the jet are used. It is required that these tracks have at least two VD hits associated in Rϕ and a minimum positive impact parameter larger than $\sqrt{3}\sigma$ relative to the main vertex of the event. A secondary vertex is then reconstructed. Tracks with a too large contribution to the χ^2 are removed from the fit in an iterative way. It is required that at least 3 tracks, with the Z coordinate measured in the VD, remain and that the distance between the secondary and the primary vertex, projected along the jet direction, is larger than 500 μm. The reconstructed mass must not exceed the B-mass (all particles are assumed to be pions). If not, tracks ordered by increasing values of their rapidity, measured relative to the jet axis, are eliminated in turn. If the reconstructed mass is smaller than the B-mass, tracks belonging to the same jet, ordered by decreasing rapidity values, are added. For charged tracks, offsets relative to the primary and secondary vertex are also examined and it is required that the track has a larger probability to originate from the secondary vertex.

The B-momentum is obtained by subtracting, from the fitted jet momentum, the momentum of the bunch of tracks, from the jet, that have not been assigned to the B-candidate.

In Figure 5.7 is given the distribution of the difference between the reconstructed and the simulated *b*-momentum, divided by the simulated value.

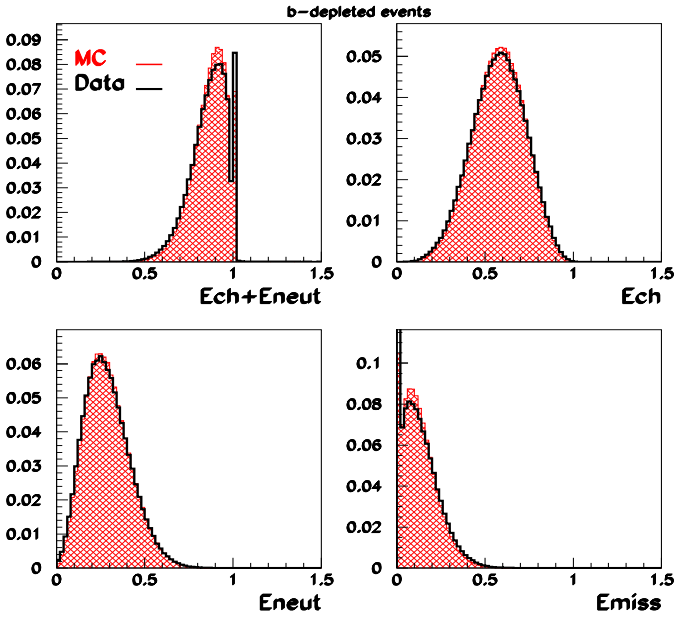

Figure 5.5: *Comparison, using a depleted b sample, between data (points with error bars) and simulation (dashed histogram) for charged, neutral and missing energy measurement after applying the constrained on total energy-momentum conservation.*

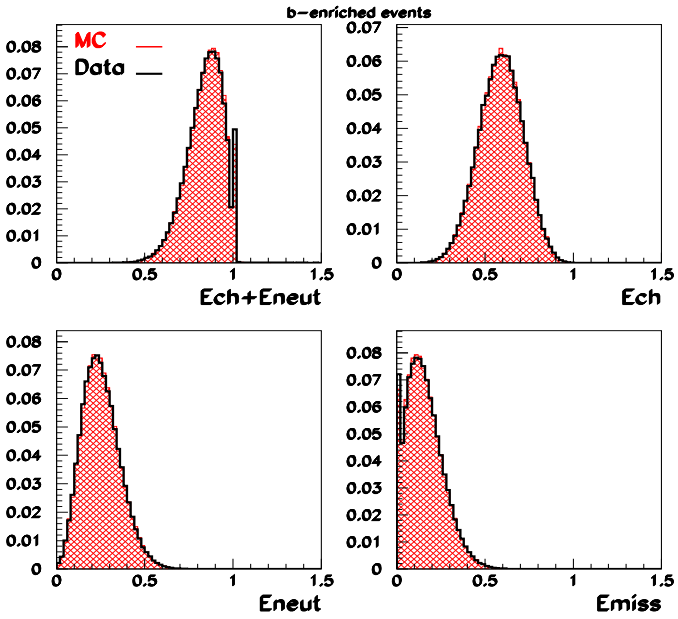

Figure 5.6: *Comparison, using a b-enriched sample, between data (points with error bars) and simulation (dashed histogram) for charged, neutral and missing energy measurement after applying the constrained on total energy-momentum conservation.*

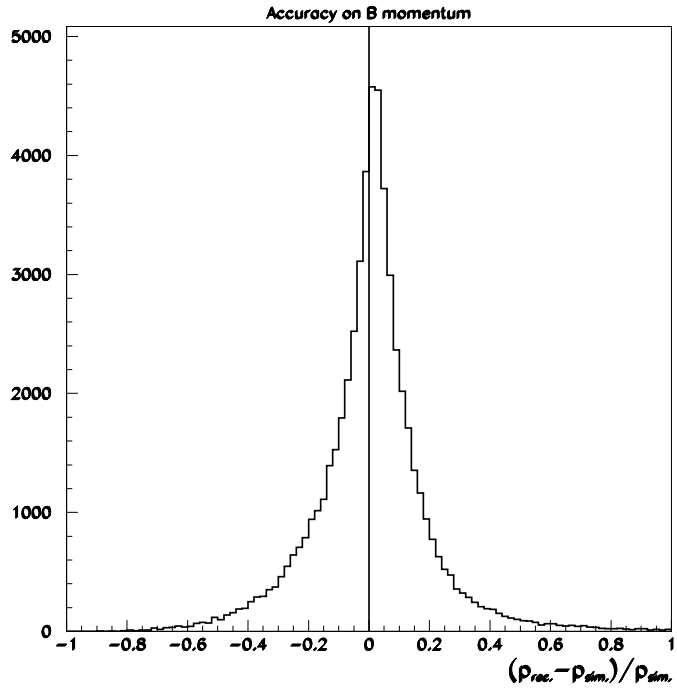

Figure 5.7: $(p_{rec} - p_{sim})/p_{sim}$ *for B-hadrons.*

5.3 Selection of B Candidates

For the b-fragmentation measurement the following, more restrictive, cuts have been applied:

- $|\cos\theta_{thrust}| < 0.7$
- *b*-tagging probability for the total event lower than 0.001
- at least three charged tracks at the secondary vertex
- the sum of the jet neutral energy and of the charged energy, for tracks which are simultaneously compatible with the primary and secondary vertices, smaller than 20 GeV

According to the simulation this corresponds to an average efficiency for the signal of 19% (Figure 5.8) and a contamination of 5%. There are 134282 and 42364 events selected in the 1994 and 1995 samples respectively.

Distributions for the fraction of the beam energy taken by the *b*-hadron momentum ($x^{rec}_{p,\,beam} = p^{rec.}_B / p_{beam}$) are compared, in Figure 5.9, with expectations from the simulation. Distributions agree for *b*-depleted events and show a clear difference for events from the *b*-enriched sample. In the following, the needed transformation of the non-perturbative QCD distribution, used in the simulation, has been determined, such that these two distributions agree.

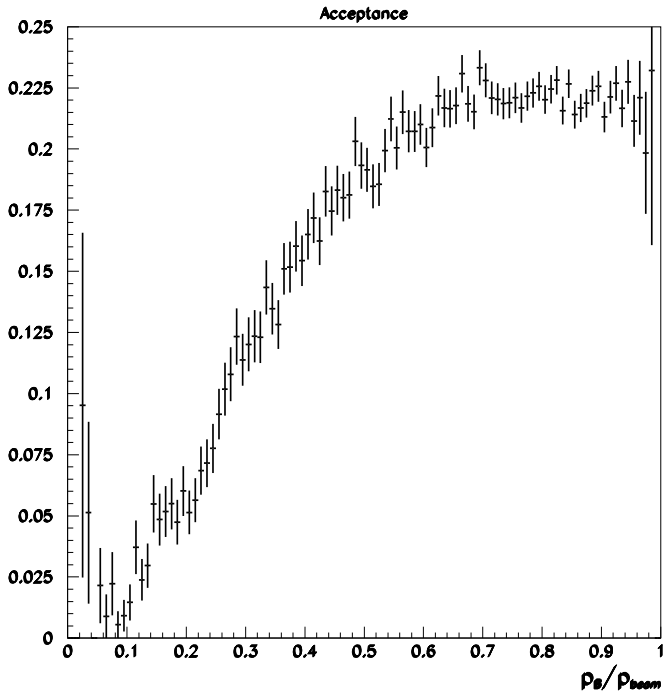

Figure 5.8: *Fitted acceptance for signal events versus* $x_p = p_B/E_{beam}$.

5.4 Measurement of the B-Fragmentation Distribution

The binned distribution of the reconstructed $x_p^{rec} = p_B^{rec.}/p_{jet}$ variable has been fitted in MINUIT by minimizing a χ^2 which includes effects from the Monte-Carlo statistics and the weighting procedure.

In each bin the number of measured events is compared with an estimated number obtained in the following way:

- contributions from background events are taken from the $q\bar{q}$ simulation. They comprise three components: non-b jets in non-$b\bar{b}$ events, non-b jets in $b\bar{b}$ events and b jets from gluon splitting. These events have been normalized to the number of hadronic events registered in data. Gluon splitting candidates have been multiplied by 1.5 to correct for the difference between simulated and measured rates at LEP. For simulated events, the fractions of these components are respectively equal to 5.2%, 0.45% and 0.24% of the analyzed events.

- the number of signal events is obtained by weighting $b\bar{b}$ simulated events. The weight contains several components which have been determined to correct the values of parameters, used in the simulation, so that they agree with corresponding measured quantities as: lifetimes, B charged-track multiplicity and B^{**} fraction in jets. A weight, whose parameters are fitted, is also applied for each value of the simulated z variable. For the analysis described in this thesis, this distribution consists of a histogram having a non-uniform

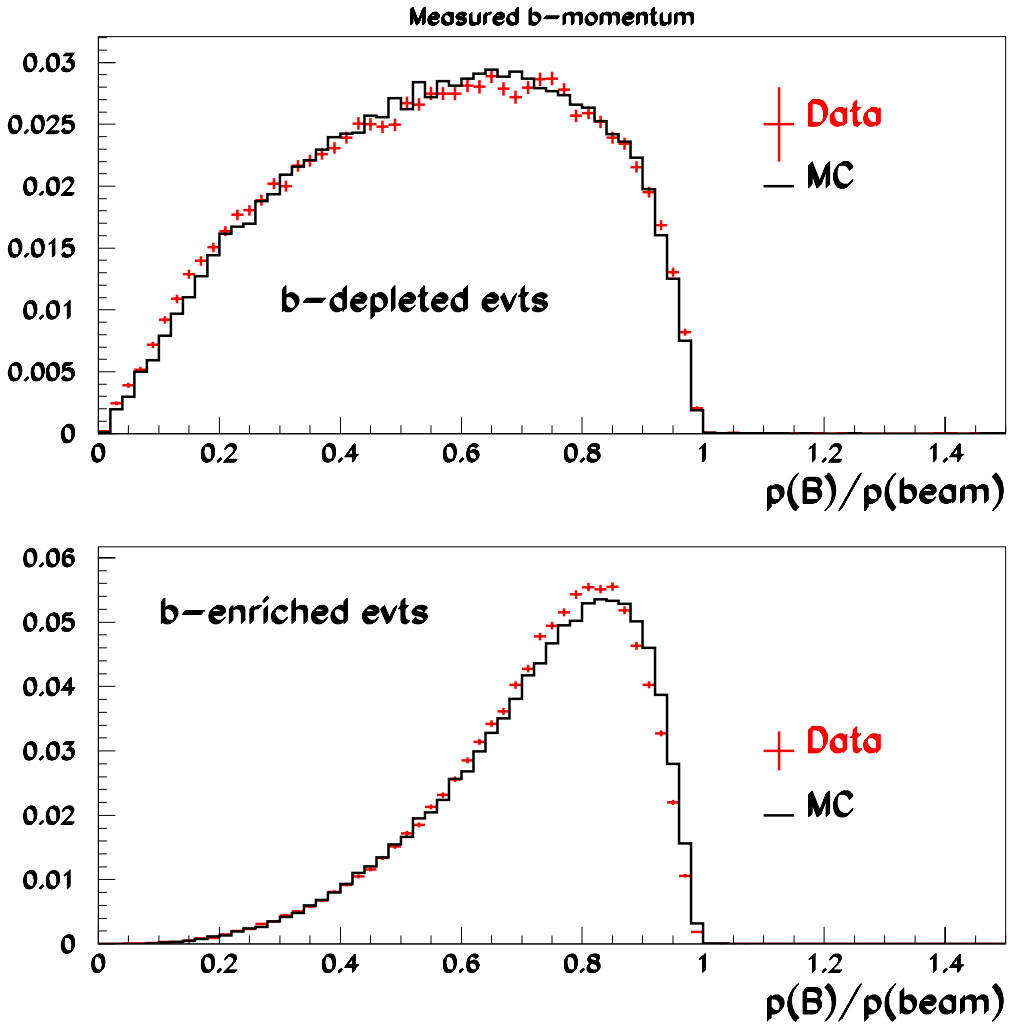

Figure 5.9: *Comparison between the measured x_p^{rec} distributions obtained in data (points with error bars) and in the MC $q\bar{q}$ simulation (histogram). Top: depleted b-sample; Bottom: enriched b-sample.*

binning.

- The normalization of *b*-events is taken as a free parameter.

To prevent oscillations between the contents of nearby bins of the weight histogram, a regularization term is included in the χ^2:

$$\chi^2 = ... + curv \times [2 \times bin(i) - bin(i-1) - bin(i+1)]^2 \tag{5.2}$$

where *curv* is a parameter whose value ($curv = 1$) has been determined empirically using simulated events.

Distributions corrected for all effects are then obtained, using corresponding generated distributions from simulated events, before any selection criteria, and by applying the weight distribution fitted on real events, which depends on the value of the *z* variable transmitted for each simulated *b*-hadron. Statistical uncertainties on these distributions have been obtained using the full covariance matrix of fitted parameters. Toy experiments are generated, each corresponding to a given set of values for the parameters and these values have been distributed according to the error matrix of the fit.

It has been also tried to fit the measured distribution using, for the signal, the generated x_p distribution multiplied by the acceptance function (Figure 5.8) and convoluted with a resolution function. This resolution function was obtained using $b\bar{b}$ simulated events. For each value of the simulated x_p a distribution for the reconstructed value, x_p^{rec}, was evaluated and this distribution was normalized to unity. It was then realized that the resolution function strongly depends on the assumed shape for the fragmentation distribution and thus the resolution function needs to be determined at each iteration of the fit. This approach was thus abandoned.

5.4.1 Fit Results on Real Data Events

Results of the fit on real data events are illustrated on Figure 5.10.

The fitted weight distribution obtained with the 1994 sample is given in Figure 5.11. This figure shows also the *z* distribution as favored by the data. It is rather different from the Peterson distribution which was used in the simulation, and given on the same figure.

As explained previously, using the fitted *z*-weight distribution, it is possible to construct any distribution relative to *b*-hadrons, before having applied selection cuts. This is illustrated on Figure 5.12 for the variable $x_E = E_B/E_{beam}$. Average values of these distributions are equal to $< x_E >_{rec.}^{data} = 0.6982 \pm 0.0010$ using the 1994 sample and $< x_E >_{rec.}^{data} = 0.7089 \pm 0.0016$ using the 1995 sample. The index *rec*. indicates that values have been reconstructed using the procedure explained previously. The difference between these two values indicates that the measurement is mainly limited by systematics. In the following, the central value $< x_E >_{rec.}^{data} = 0.704 \pm 0.001$ will be used which corresponds to the simple average between the two previous values, assuming that they have a similar weight.

To extract analytically, the non-perturbative QCD component of the fragmentation function, integrals of the fitted $x_{E\ rec.}^{data}$ distribution over (nine) specified intervals have been used. In each bin, the quantity

$$\frac{\Delta F}{\Delta x_E} = \frac{1}{x_E^{max} - x_E^{min}} \int_{x_E^{min}}^{x_E^{max}} \frac{dF}{dx_E}\, dx_E \tag{5.3}$$

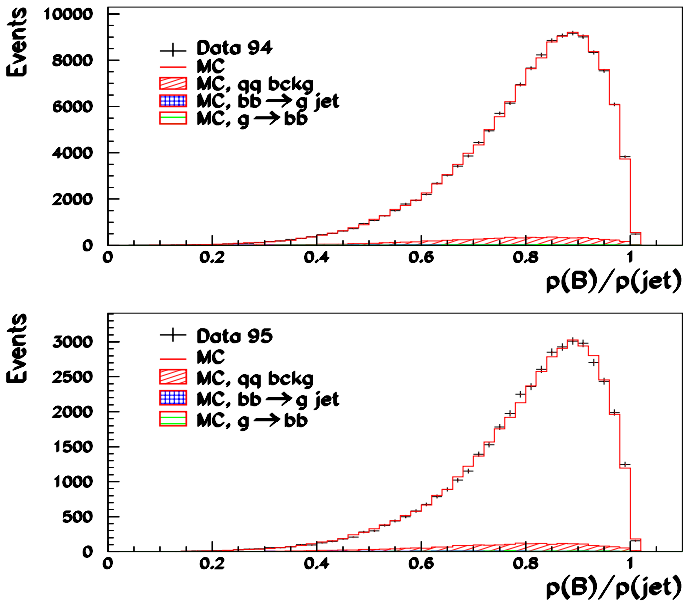

Figure 5.10: *Fitted $x_p^{rec} = p_B^{rec}/p_{jet}$ distributions on selected events. Top: 1994 sample. Bottom: 1995 sample.*

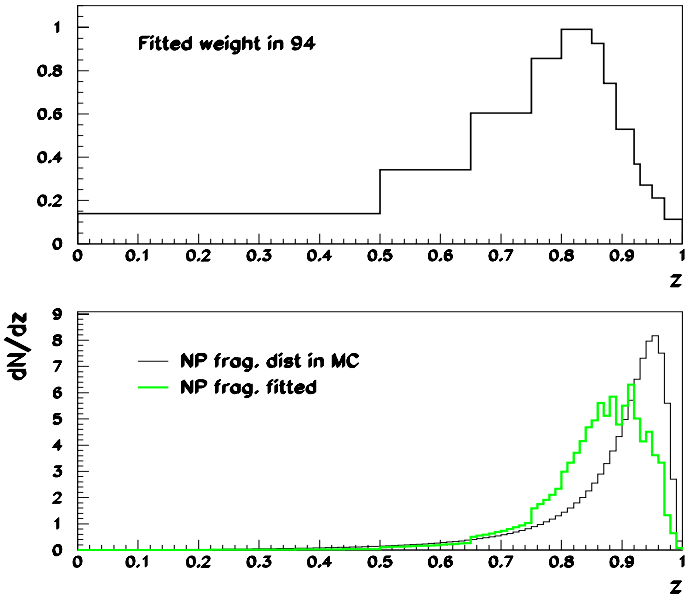

Figure 5.11: *Fitted z distribution on 1994 selected events. Top: Distribution of the fitted weights. Bottom: Comparison between the initial Peterson distribution, used in the simulation generator, and the corresponding distribution favored by data events.*

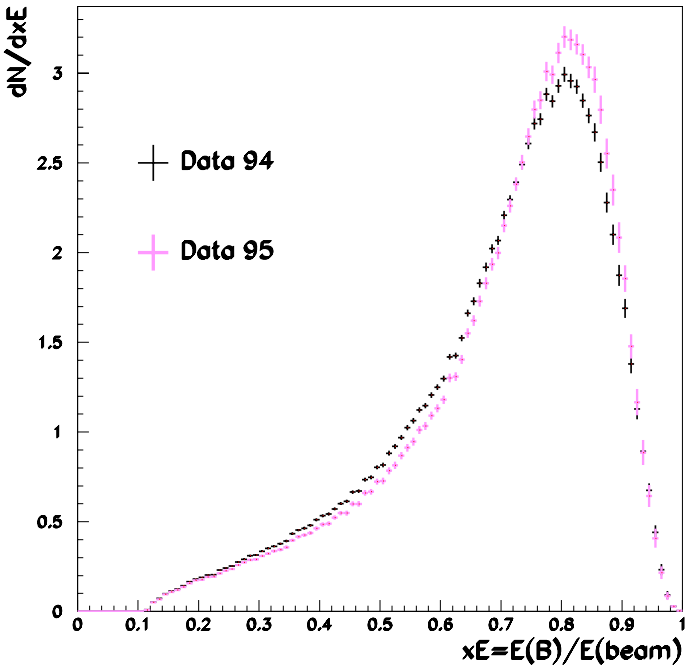

Figure 5.12: *Comparison between x_E distributions obtained using the 1994 and 1995 events samples.*

is evaluated and the corresponding statistical error matrix is obtained using the approach explained in the introduction of this section. Values x_E^{max} and x_E^{min} are the bin limits.

5.4.2 Fit Results on Simulated Events

Results on simulated $q\bar{q}$ events are illustrated in Figure 5.13.

Generated events correspond to the average value $< x_E >^{sim.}_{gen.} = 0.7057$ and have been reconstructed at $< x_E >^{sim.}_{rec.} = 0.7060$. The quoted values have been corrected for the effect of the beam radiation which corresponds to an increase of 0.0015. This average value depends on the choice for the value of the curvature parameter introduced in the χ^2 expression (see Equation 5.2), as given in Table 5.6. Variations are at the level of ± 0.002 on simulated events and can be even smaller on real data events.

In the following analysis the value $curv = 1$ has been used. In Figure 5.13 the simulated and fitted distributions have been compared and the observed difference has been included in the systematics.

5.5 Systematic Uncertainties

Several sources of uncertainties have been identified; including:

bin low edge	0.10	0.30	0.42	0.54	0.64	0.73	0.80	0.88	0.94
$\frac{\Delta F}{\Delta x_E}$	0.1745	0.4263	0.7083	1.1631	1.9072	2.7371	2.9209	1.6685	0.2601
Error matrix									
	0.0027	0.9933	0.8251	0.4431	-0.0507	-0.1364	0.2224	0.2808	-0.0394
	0.9933	0.0076	0.8648	0.4868	-0.0777	-0.1992	0.1800	0.2498	-0.0322
	0.8251	0.8648	0.0117	0.8281	0.0833	-0.3005	0.0354	0.3134	-0.0336
	0.4431	0.4868	0.8281	0.0171	0.4971	-0.1817	-0.1430	0.2946	0.0028
	-0.0507	-0.0777	0.0833	0.4971	0.0247	0.5337	-0.1735	-0.0375	0.1562
	-0.1364	-0.1992	-0.3005	-0.1816	0.5337	0.0398	0.2600	-0.4136	0.2342
	0.2224	0.1800	0.0354	-0.1430	-0.1735	0.2600	0.040	-0.1068	-0.0824
	0.2808	0.2498	0.3134	0.2946	-0.0375	-0.4136	-0.1068	0.0429	-0.4019
	-0.0394	-0.0322	-0.0336	0.0028	0.1562	0.2342	-0.0824	-0.4019	0.0204

Table 5.5: *Integral and corresponding statistical error matrix of the fitted fragmentation distribution, over specified intervals. Diagonal elements of the error matrix (σ_i) correspond to evaluated uncertainties, correlation coefficients (ρ_{ij}) are given in the non-diagonal elements. Coefficients of the full statistical error matrix can be obtained from these informations; diagonal elements being equal to σ_i^2 and non-diagonal to $\rho_{ij}\ \sigma_i\ \sigma_j$.*

curv value	0.02	0.2	1.0	5.0
$< x_E >_{rec.}^{sim.}$	0.7042 ± 0.0011	0.7050 ± 0.0008	0.7060 ± 0.0008	0.7076 ± 0.0006
$< x_E >_{rec.}^{data}$	0.6986 ± 0.0015	0.6986 ± 0.0012	0.6982 ± 0.0010	0.6971 ± 0.0008

Table 5.6: *Variation of the fitted average value of the beam energy taken by a b-hadron with the value taken for the curv parameter.*

- differences between the detector behavior and its simulation;
- uncertainties coming from those attached to physics parameters;
- uncertainties related to the stability of results versus the values of selection criteria used in the analysis.

For each source of systematic uncertainty, the variations observed on the average value of the x_E distribution ($< x_E >$) and on the value of $\frac{\Delta F}{\Delta x_E}$, in each of the considered bin in x_E, have been evaluated. Denoting as Y, these various quantities, the variations:

$$\delta Y = Y(modified\ parameter\ value) - Y(nominal\ analysis) \tag{5.4}$$

have been determined.

5.5.1 Real Data and Simulation Tuning

5.5.1.1 Energy Calibration

The analysis uses the beam energy as a constraint in a global fit of charged and neutral tracks 4-momentum such that the total energy and momentum of the event is conserved.

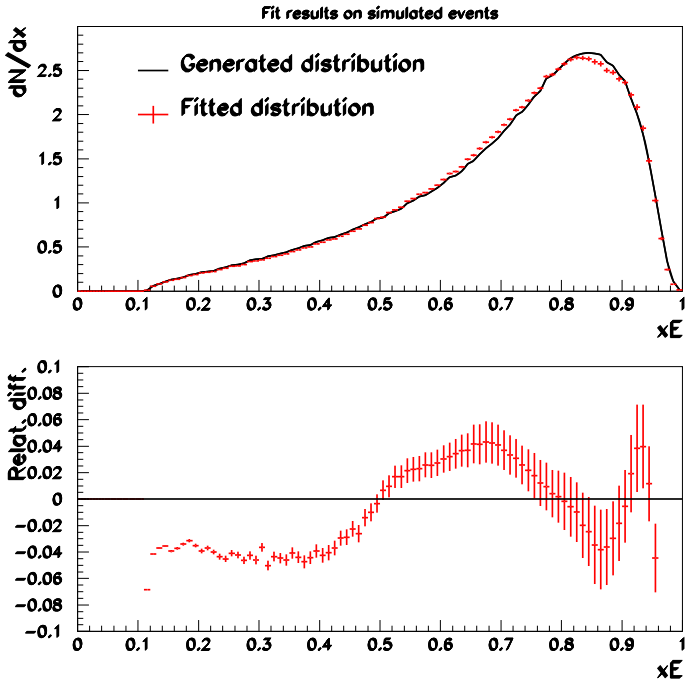

Figure 5.13: *Comparison between generated and fitted x_E distributions. Top: The generated (full line) and fitted (points with error bars) distributions are compared. Bottom: The relative difference between the fitted and simulated distributions is given.*

Remaining uncertainties have been evaluated in the following way. As the difference between the values of $< x_E >$ obtained using 1994 and 1995 samples is larger than the corresponding statistical accuracy, these samples can be split and observed differences $< x_E >_{rec.}^{sample\ i} - < x_E >_{rec.}^{data}$ are displayed in Figure 5.14. From the dispersion of these values, a systematic uncertainty is evaluated (supposed to be of similar importance for each measurement) such that the χ^2/NDF is equal to unity. This uncertainty amounts to 0.0064.

Corresponding variations on $\frac{\Delta F}{\Delta x_E}$ are given in Table 5.7. They have been obtained by comparing, in each bin, results using 1994 and 1995 data and measured differences have been scaled such that they correspond to a variation of ± 0.0064 on $< x_E >$.

5.5.1.2 Level of the Non-*b* Background

In the analyzed sample, with the cut $P(btag) \leq 10^{-3}$, the estimated fraction of non-*b* candidates amounts to 5.2%. In Table 5.1, it was observed that the fraction of selected events is a few % (relative) higher in real data. As this effect remains in samples of high purity in *b* events, its main origin comes, most probably, from a difference in efficiency between real and simulated *b*-events. A possible underestimate of the selection efficiency to non-*b* events, in the simulation, amounts then to 10% (relative), at maximum. The effect of a $\pm 20\%$ variation on the non-*b* background level has been evaluated; it gives $\delta < x_E >_{rec.}^{data} = \pm 0.0012$. Corresponding

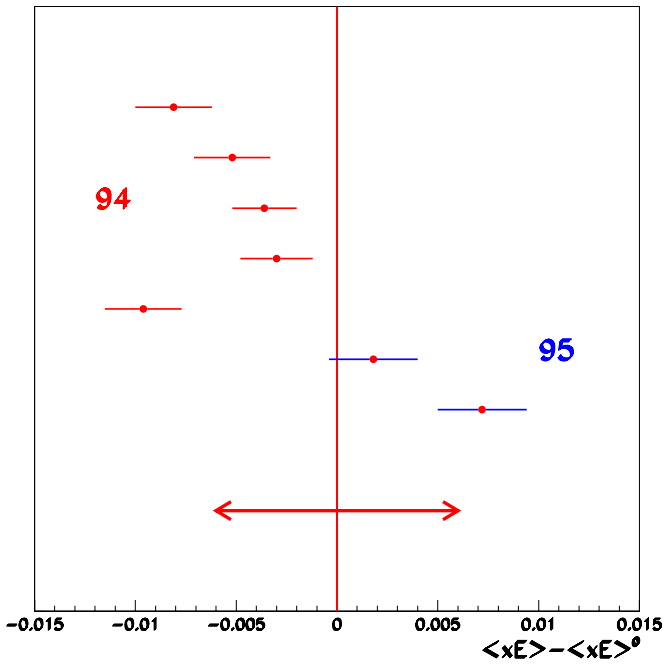

Figure 5.14: *Differences between the fitted x_E values, relative to the average, for different data samples.*

variations on $\frac{\Delta F}{\Delta x_E}$ are given in Table 5.7.

5.5.1.3 Track Energy and Multiplicity Tuning

Corrections applied on charged and neutral track distributions have been described in Section 5.1.1.2. They induce a rather small variation on $< x_E >$ of +0.0017. The corresponding systematic uncertainty has been evaluated taking the effect of half this correction. Corresponding variations on $\frac{\Delta F}{\Delta x_E}$ are given in Table 5.7.

5.5.1.4 Jet Multiplicity

Measured jet multiplicities are different in data and simulated events as displayed in Figure 5.15.

Simulated events have been weighted accordingly so that the two distributions become in agreement. The systematic uncertainty has been evaluated to be one third of the previous correction, from the observed differences between the two events samples in Figure 5.15. It corresponds to $\delta < x_E >^{data}_{rec.} = -0.0001$. Corresponding variations on $\frac{\Delta F}{\Delta x_E}$ are given in Table 5.7.

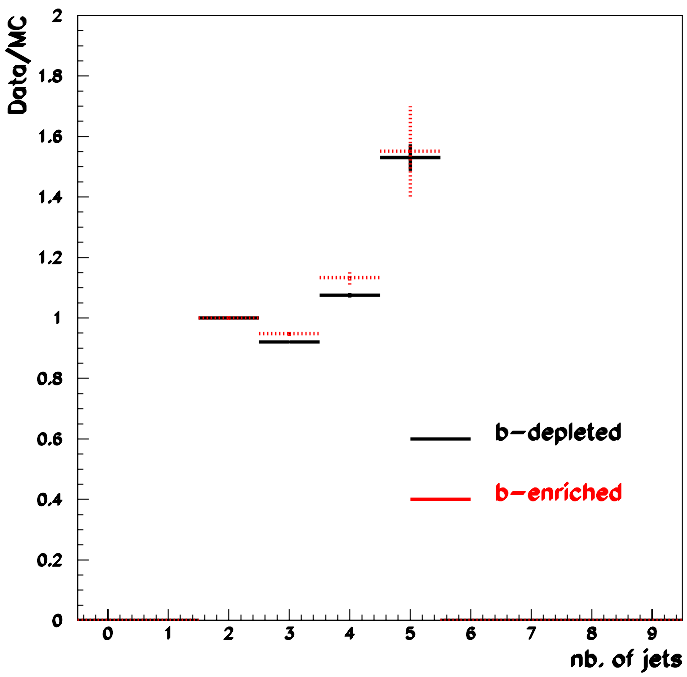

Figure 5.15: *Ratio between number of events having a given jet multiplicity in data and in the simulation. For 2-jet events this ratio has been set to unity. Events depleted and enriched in b hadrons have been compared.*

5.5.1.5 Summary

Systematics related to uncertainties on the tuning of the simulation are summarized in Table 5.7 and are illustrated in figure 5.16.

bin	1	2	3	4	5	6	7	8	9
Energy calibration	0.0063	0.0227	0.0465	0.0695	0.0446	0.0596	0.1546	0.0753	0.0088
non-*b* background	0.0016	0.0055	0.0096	0.0114	0.0015	0.0170	0.0233	0.0073	0.0012
Track distribution tuning	0.0005	0.0004	0.0020	0.0071	0.0206	0.0208	0.0181	0.0403	0.0057
Jet multiplicity	0.0003	0.0006	0.0004	0.0026	0.0034	0.0008	0.0034	0.0035	0.0007

Table 5.7: *Systematic uncertainties, in each bin of x_E, related to the tuning of the simulation.*

5.5.2 Physics Parameters

Variations of the values of parameters, that govern decay properties or production characteristics of B-hadrons, have been considered.

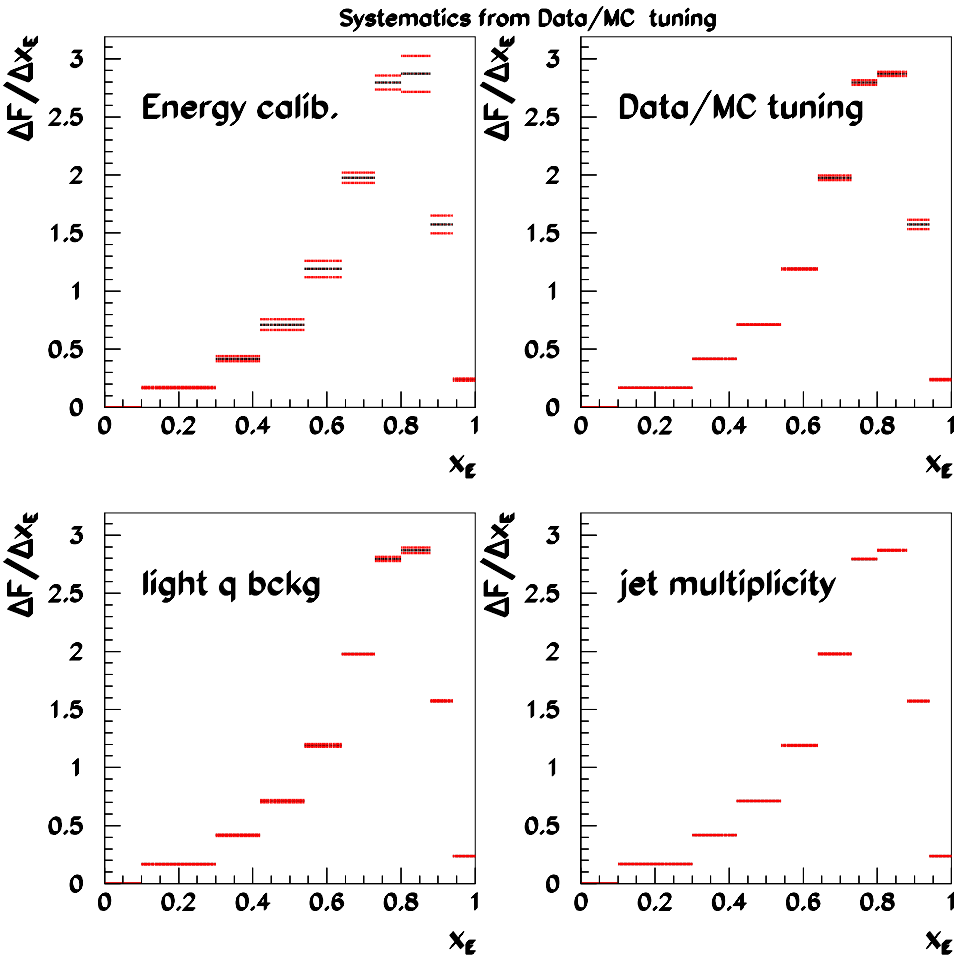

Figure 5.16: *Systematic uncertainties originating from uncertainties in the tuning of the simulation.*

5.5.2.1 *b*-Hadron Lifetimes

Simulated events have been generated using the same lifetime value of $\tau_0 = 1.6$ *ps*. Events have been weighted so that each type of *b*-hadron is distributed according to its corresponding lifetime, as given in PDG. The applied weight is:

$$w_{time}(t) = \frac{\tau_0}{\tau_B} \exp -t \left(\frac{1}{\tau_B} - \frac{1}{\tau_0} \right) \tag{5.5}$$

where τ_B is the measured B-hadron lifetime and t is its decay proper time, in the studied event. Taking, as systematics, the variation induced by this correction, the variation on $< x_E >^{data}_{rec.}$ is equal to -0.0005. Corresponding variations on $\frac{\Delta F}{\Delta x_E}$ are given in Table 5.8.

5.5.2.2 B^{**} Production Rate

In the simulation, the B^{**} production rate, in a b-quark jet amounts to $p_0(B^{**}) = 32\%$. Recent measurements [81] favor a lower value of $(16 \pm 1 \pm 2 \pm 2)\%$. A weight, $w_{B^{**}}$, is applied on B-hadrons which originate from a B^{**} decays. It has been defined such that:

$$p(B^{**}) = \frac{w_{B^{**}} p_0(B^{**})}{w_{B^{**}} p_0(B^{**}) + (1 - p_0(B^{**}))}. \tag{5.6}$$

The central value $p(B^{**}) = 20\%$ has been used in the following and the corresponding variation relative to $p_0(B^{**})$ as been taken as a systematic.

The variation on $< x_E >^{data}_{rec.}$ is equal to -0.0009. Corresponding variations on $\frac{\Delta F}{\Delta x_E}$ are given in Table 5.8.

5.5.2.3 b-Hadron Charged Multiplicity

Differences between simulated and measured average charged track multiplicities in b-hadron decays amount to 0.06.

$$n^{PDG}_{ch}(B) = 4.97 \pm 0.03 \pm 0.06,\ n^{sim.}_{ch}(B) = 4.91 \tag{5.7}$$

This difference has been corrected by weighting events using a weight which has a linear variation with the actual b-hadron charged track multiplicity in a given event. The simulated multiplicity distribution has been fitted with a Gaussian of standard deviation (σ_{nch}) equal to 2.03 charged tracks. Probability values, for a given charged multiplicity i, P_i, have been transformed into:

$$P^T_i = P_i \left[1 + \beta(< n > - i)\right]. \tag{5.8}$$

The value of β is obtained by expressing that the new average multiplicity, computed using P^T_i, is equal to $n^{PDG}_{ch}(B)$. Then:

$$\beta = \frac{n^{sim.}_{ch} - n^{PDG}_{ch}}{\sigma^2_{nch}}. \tag{5.9}$$

The corresponding systematic uncertainty has been evaluated by considering an uncertainty of ± 0.1 charged track.

The variation on $< x_E >^{data}_{rec.}$ is equal to ∓ 0.0008 and corresponding variations on $\frac{\Delta F}{\Delta x_E}$ are given in Table 5.8.

5.5.2.4 $g \to b\bar{b}$ Rate

The rate for b-hadron production originating from gluon coupling to $b\bar{b}$ pairs has been measured by LEP experiments and found to be larger than the rate used in the simulation by a factor 1.5. The corresponding systematic uncertainty has been evaluated, considering the uncertainty, of 30%, obtained by DELPHI on this quantity [82]. The variation on $< x_E >^{data}_{rec.}$ is equal to -0.0001 and corresponding variations on $\frac{\Delta F}{\Delta x_E}$ are given in Table 5.8.

5.5.2.5 Summary

Systematics related to uncertainties on the values of physics parameters governing B-hadron decay properties and production characteristics are summarized in Table 5.8 and are illustrated in Figure 5.17.

bin	1	2	3	4	5	6	7	8	9
b-hadron lifetime	0.0006	0.0020	0.0038	0.0054	0.0033	0.0041	0.0103	0.0078	0.0011
B^{**} rate	0.0030	0.0054	0.0027	0.0031	0.0043	0.0260	0.0207	0.0045	0.0129
Charged track multiplicity	0.0001	0.0006	0.0020	0.0051	0.0089	0.0061	0.0066	0.0178	0.0079
$g \to b\bar{b}$ rate	0.0002	0.0005	0.0009	0.0012	0.0003	0.0017	0.0026	0.0009	0.0007

Table 5.8: *Systematic uncertainties, in each bin of x_E, related to uncertainties on physics parameters governing B-hadron decay and production characteristics.*

5.5.3 Parameters Used in the Analysis

5.5.3.1 Parametrization of the Weight Function

The fitted weight function consists of 11 bins in z whose content is fitted. Choices for the bin definition and bin number can induce a systematic on the extracted distribution. This has been studied by comparing the generated and fitted distributions in simulated events (see Section 5.4.2). The observed difference corresponds to a variation on $< x_E >_{rec.}^{data}$ equal to $+0.0003$; corresponding variations on $\frac{\Delta F}{\Delta x_E}$ are given in Table 5.12.

5.5.3.2 b-Tagging Selection

The stability of the fitted value, $< x_E >_{rec.}^{data}$, has been studied for different selections on the value of the $P(btag)$ variable. Results are summarized in Table 5.9.

cut on $P(btag)$	$< 10^{-3}$	$< 10^{-4}$	$< 10^{-6}$	$< 10^{-10}$
nb of events	134282	120082	93220	49533
$< x_E >_{rec.}^{data}$	0.6982 ± 0.0010	0.6989 ± 0.0011	0.6992 ± 0.0012	0.6972 ± 0.0016

Table 5.9: *Variation of the fitted average value of the beam energy taken by a b-hadron with the cut on $P(btag)$.*

Results are stable within ± 0.001. For the corresponding systematic evaluation, half the differences obtained using cuts at 10^{-4} and 10^{-10} have been used. Corresponding variations on $\frac{\Delta F}{\Delta x_E}$ are given in Table 5.12.

5.5.3.3 Jet Clustering Parameter Value

Hadronic jets have been reconstructed using the LUCLUS algorithm with the value of the parameter defining the jets, djoin=5 GeV. Sensitivity of present results on the value of this parameter has been studied by redoing the measurements using djoin=10 GeV. The variation on $< x_E >_{rec.}^{data}$ is equal to $+0.0002$; corresponding variations on $\frac{\Delta F}{\Delta x_E}$ are given in Table 5.12.

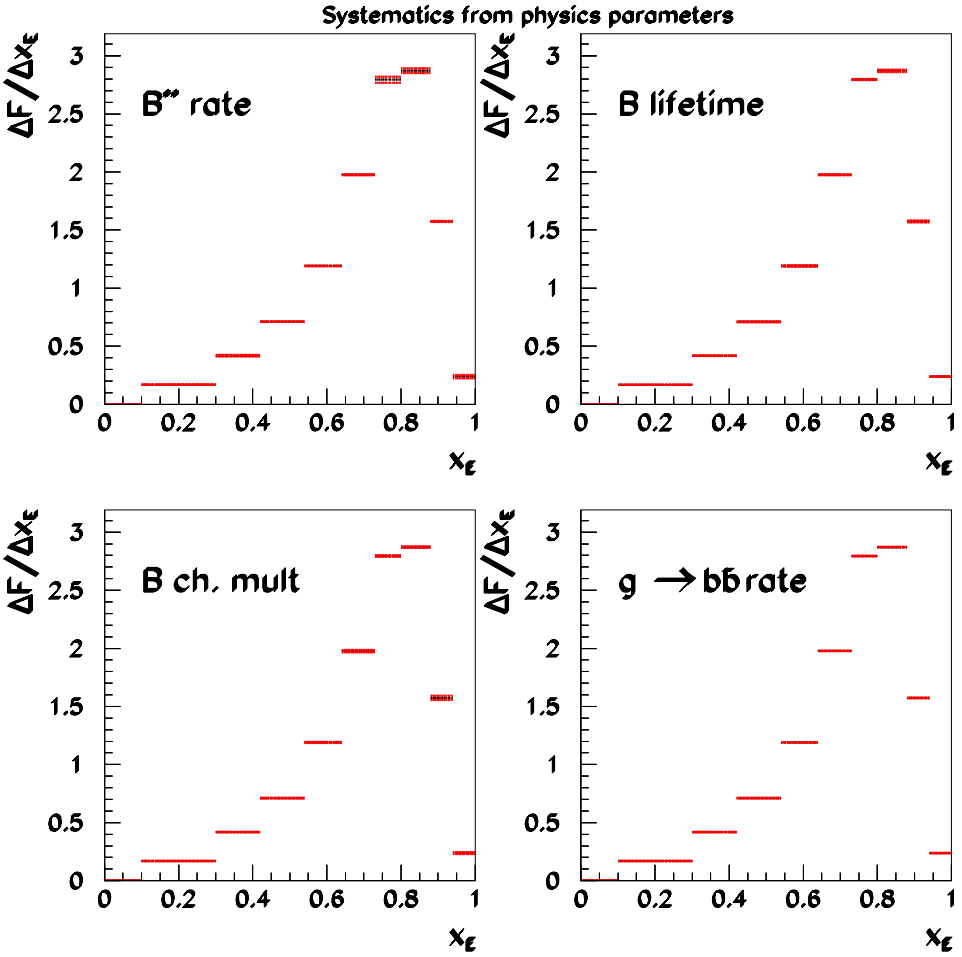

Figure 5.17: *Systematic uncertainties originating from uncertainties on physics parameters governing B-hadron decay and production characteristics.*

5.5.3.4 Level of Ambiguous Energy

In a jet, there are charged particles which can be compatible simultaneously with the primary and the secondary vertex. Neutral particles cannot be either attached with confidence to one of the two vertices. In the analysis, events have been selected requiring that the total "ambiguous" energy is lower than 20 GeV. Results stability has been studied, in Table 5.10, by changing the value for this selection, Taking the differences observed for selections at 10 and 20 GeV, as an evaluation for the corresponding systematic, the variation on $< x_E >^{data}_{rec.}$ is equal to -0.0034 and corresponding variations on $\frac{\Delta F}{\Delta x_E}$ are given in Table 5.12.

cut on amb. energy	< 20GeV	< 15GeV	< 10GeV
nb of events	134282	102784	62029
$< x_E >^{data}_{rec.}$	0.6982 ± 0.0010	0.6977 ± 0.0011	0.6947 ± 0.0014

Table 5.10: *Variation of the fitted average value of the beam energy taken by a b-hadron with the cut on "ambiguous" energy value.*

5.5.3.5 Secondary Vertex Charged Multiplicity

Events have been selected, requiring at least three charged particles at the candidate B decay vertex. In Table 5.11, results have been reported for different values of the minimum charged track multiplicity.

B-vert. mult.	≥ 3	≥ 4	≥ 5	≥ 6
nb of events	134282	80277	41065	17597
$< x_E >^{data}_{rec.}$	0.6982 ± 0.0010	0.6987 ± 0.0011	0.6994 ± 0.0016	0.6977 ± 0.0024

Table 5.11: *Variation of the fitted average value of the beam energy taken by a b-hadron with the minimal charged track multiplicity at the B decay vertex.*

Taking the differences observed for selections with at least three and five charged particles, as an evaluation for the corresponding systematic, the variation on $< x_E >^{data}_{rec.}$ is equal to $+0.0012$ and corresponding variations on $\frac{\Delta F}{\Delta x_E}$ are given in Table 5.12.

5.5.3.6 Summary

Systematics related to uncertainties on the values of physics parameters governing B-hadron decay properties and production characteristics are summarized in Table 5.12 and are illustrated in Figure 5.17.

bin	1	2	3	4	5	6	7	8	9
fitting procedure	0.0077	0.0213	0.0053	0.0343	0.0656	0.0225	0.0553	0.0078	0.0372
btag selection	0.0012	0.0041	0.0078	0.0105	0.0013	0.0187	0.0205	0.0002	0.0017
jet clustering parameter	0.0012	0.0028	0.0019	0.0036	0.0193	0.0277	0.0148	0.0587	0.0246
ambiguous energy	0.0036	0.0125	0.0233	0.0320	0.0186	0.0209	0.0578	0.0601	0.0034
B-vert mult.	0.0013	0.0039	0.0062	0.0074	0.0026	0.0026	0.0126	0.0027	0.0517

Table 5.12: *Systematic uncertainties, in each bin of x_E, related to the values and procedures used in the analysis.*

5.6 Comparison with Other Experiments

The present measurement of $< x_E >$ is compared, in Table 5.13, with the other results obtained at the Z energy, by ALEPH [31], DELPHI(Karlsruhe) [34], OPAL [32] and SLD [33] collaborations.

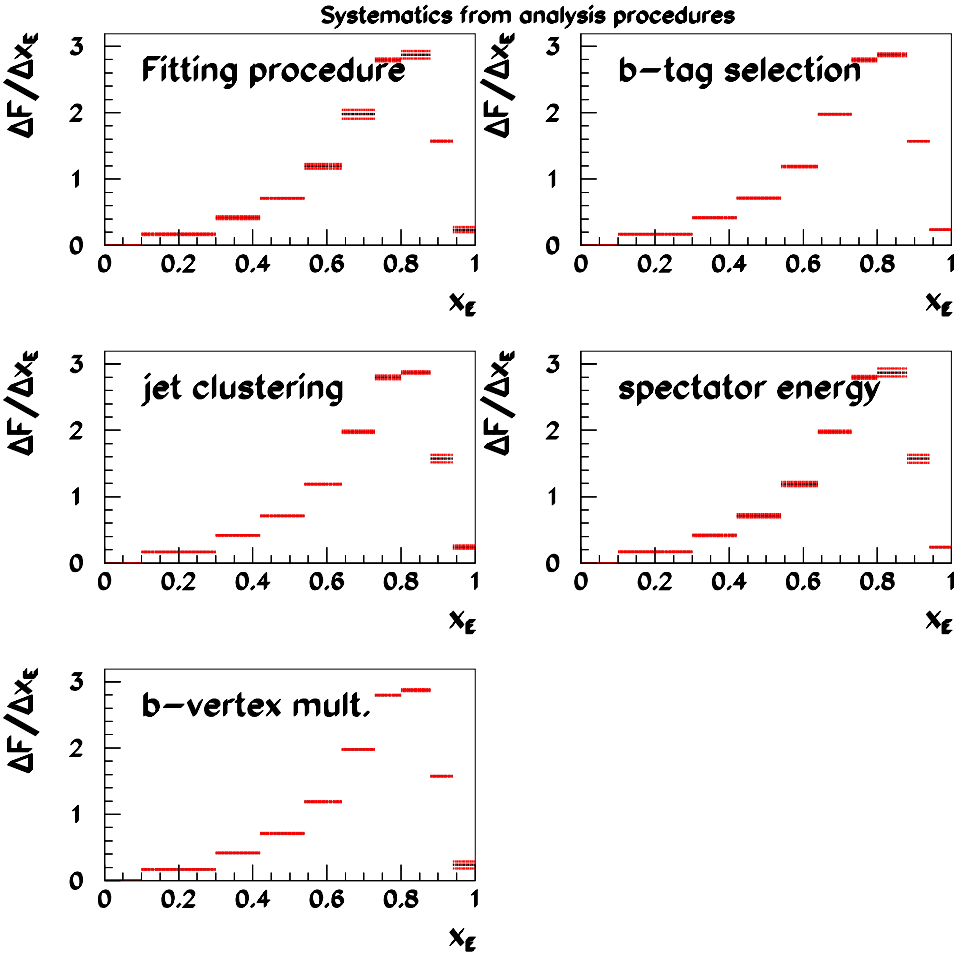

Figure 5.18: *Systematic uncertainties related to the stability of results versus the values of criteria used in the analysis.*

Corresponding x_E distributions have been compared, in Figure 5.19. The total error matrix of present measurements is given in Table 5.14.

The present measurement is quite compatible with the SLD result, over all the distribution. The non-perturbative QCD components of the *b*-fragmentation distribution, extracted from these various measurements are compared in Section 6.7. The obtained non-perturbative distribution is quite similar with the distributions obtained from SLD and OPAL measurements and somewhat softer, but compatible within uncertainty, as compared with the one from ALEPH. It is softer than the non-perturbative distribution corresponding to the measurement from DELPHI (Karlsruhe).

The ALEPH measurement is using semileptonic decays of B_d^0 and B^+ mesons. The other experiments are sensitive to decay modes of all weakly decaying *b*-hadrons. In Figure 5.20, x_E distributions have been compared, corresponding respectively to non-strange *b*-mesons and to

Experiment	$< x_E >$	Stat. err.	Syst. err.
This thesis	0.704	0.001	0.008
ALEPH	0.716	0.006	0.006
DELPHI (Karl.)	0.715	0.001	0.005
OPAL	0.719	0.002	0.004
SLD	0.709	0.003	0.004

Table 5.13: *Different measurements of $< x_E >$ at the Z energy.*

bin low edge	0.10	0.30	0.42	0.54	0.64	0.73	0.80	0.88	0.94
Error matrix									
	0.0120	0.9838	0.7497	0.3210	-0.1965	-0.6306	-0.3905	-0.4510	0.2314
	0.9838	0.0353	0.8389	0.4514	-0.0670	-0.7036	-0.5027	-0.5522	0.2310
	0.7497	0.8389	0.0549	0.8618	0.4436	-0.7594	-0.8533	-0.6982	0.0257
	0.3210	0.4514	0.8618	0.0869	0.8182	-0.5620	-0.9465	-0.6619	-0.1696
	-0.1965	-0.0670	0.4436	0.8182	0.0912	-0.0527	-0.7606	-0.5351	-0.2814
	-0.6306	-0.7036	-0.7594	-0.5620	-0.0527	0.0903	0.5406	0.2474	-0.1434
	-0.3905	-0.5027	-0.8533	-0.9465	-0.7606	0.5406	0.1878	0.6990	0.0090
	-0.4510	-0.5522	-0.6982	-0.6619	-0.5351	0.2474	0.6990	0.1332	-0.2096
	0.2314	0.2310	0.0257	-0.1696	-0.2814	-0.1434	0.0090	-0.2096	0.0750

Table 5.14: *Total error matrix of the fitted fragmentation distribution, over specified intervals. Diagonal elements of the error matrix (σ_i) correspond to evaluated uncertainties, correlation coefficients (ρ_{ij}) are given in the non-diagonal elements. Coefficients of the full error matrix can be obtained from these informations; diagonal elements being equal to σ_i^2 and non-diagonal to $\rho_{ij}\ \sigma_i\ \sigma_j$.*

the sum of B_s^0 and *b*-baryons components. As the latter have a larger mass, their corresponding fragmentation distribution is peaked at higher values. Taking into account the various production rates of *b*-hadrons in jets, one obtains:

$$< x_E >_{all\ b-had.} = < x_E >_{B^0,\ B^+} + 0.0011 \tag{5.10}$$

It is thus expected that the ALEPH measurement is sligtly lower in $< x_E >$ as compared with the other measurements. In practice, considering uncertainties attached to these measurements, this difference is non-significant.

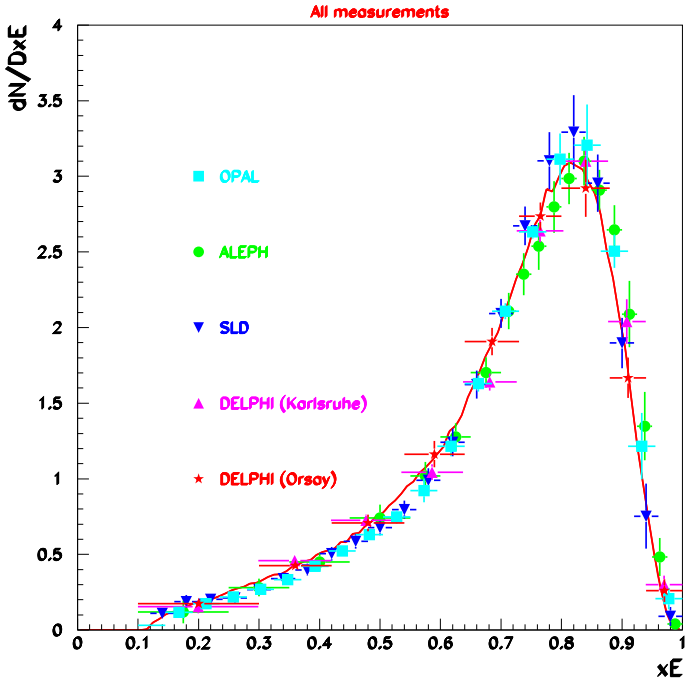

Figure 5.19: *Comparison between the measured x_E distribution and corresponding results obtained by ALEPH, DELPHI, OPAL and SLD.*

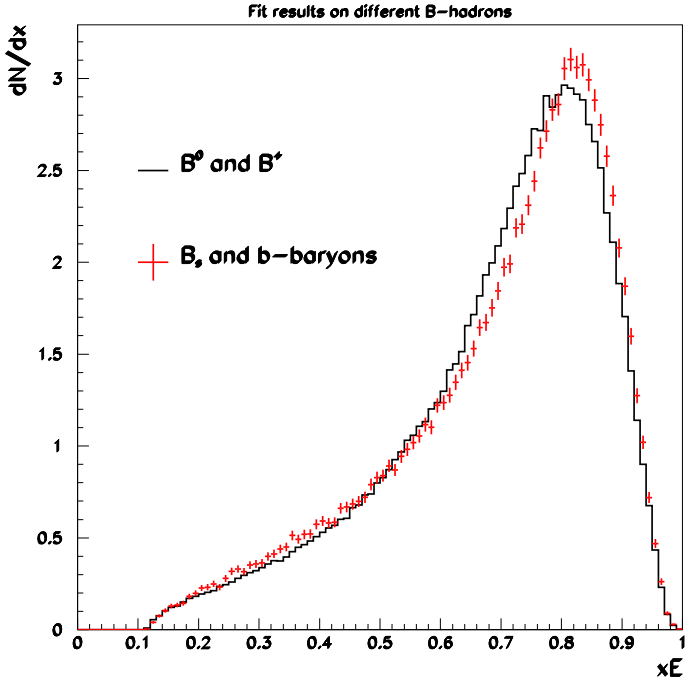

Figure 5.20: *Comparison between expected x_E distributions for B_d^0 and B^+ mesons and for B_s^0 and b-baryons.*

Chapter 6

Extraction of the x-Dependence of the Non-perturbative QCD Component

6.1 Introduction

In Chapter 5 a measurement of the *b*-quark fragmentation distribution has been presented. It has been defined as the fraction of the beam energy taken by a weakly decaying *B*-hadron in $e^+e^- \rightarrow b\bar{b}$ events registered at, or near, the Z pole. Similar measurements have been obtained by the ALEPH [31], DELPHI(Karlsruhe) [34], OPAL [32] and SLD [33] collaborations.

As discussed in Chapter 2, this distribution is generally viewed as resulting from three components: the primary interaction (e^+e^- annihilation into a $b\bar{b}$ pair in the present study), a perturbative QCD description of gluon emission by the quarks and a non-perturbative QCD component which incorporates all mechanisms at work to bridge the gap between the previous phase and the production of weakly decaying *B*-mesons. The perturbative QCD component can be obtained using analytic expressions or Monte Carlo generators. The non-perturbative QCD component is usually parameterized phenomenologically via a model.

To compare with experimental results, one must fold both components to evaluate the expected *x*-dependence:

$$\mathcal{D}_{predicted}(x) = \int_0^1 \mathcal{D}_{pert.}(z) \times \mathcal{D}^{model}_{non-pert.}(\frac{x}{z})\,\frac{dz}{z} \tag{6.1}$$

In the present chapter, the fragmentation variable is chosen as x_p (see Section 2.3.1). In order to simplify the notation, The index p will be omitted all along this chapter and therefore x_p will be referred to as x. The experimental results presented in this thesis work, as well as the other measurements of the fragmentation function [31, 32, 33] have been originally obtained in terms of the variable $x_E = \frac{2E_B}{\sqrt{s}}$. One can easily change this variable to x by using

$$x = \frac{\sqrt{x_E^2 - x_{min}^2}}{\sqrt{1 - x_{min}^2}} \tag{6.2}$$

where $x_{min} = \frac{2m_B}{\sqrt{s}}$ is the minimal value of x_E. The final and the perturbative components are defined over the $[0,1]$ interval. As explained in the following, the non-perturbative distribution

must be evaluated for $x > 1$, if the perturbative component is non-physical. The parameters of the model are then fitted by comparing the measured and predicted x-dependence of the b-quark fragmentation distribution. Such comparisons have already been made by the different experiments using, for the perturbative component, expectations from generators such as the JETSET or HERWIG parton shower Monte-Carlo. It has been shown, with present measurement accuracy, that most existing models of the non-perturbative part are unable to give a reasonable fit to the data [31, 34, 32, 33]. Best results have been obtained with the Lund and Bowler models [38, 24].

In the following, a method is presented to extract the non-perturbative QCD component of the fragmentation function directly from data, independently of any hadronic model assumption. This distribution can also be compared with models to learn about the non-perturbative QCD transformation of b-quarks into B-hadrons. It can then be used in another environment than e^+e^- annihilation, as long as the same parameters and methods are taken for the evaluation of the perturbative QCD component. Consistency checks, on the matching between the measured and predicted b-fragmentation distribution, can be defined which provide information on the determination of the perturbative QCD component itself.

The method to extract the x dependence of the non-perturbative component does not depend on a particular measured fragmentation function. In this thesis work, the method has been originally applied to the result from ALEPH [31], and therefore this is the main measurement that will be referred to in this chapter. The extracted non-perturbative components for all the other available measured b fragmentation functions, including the measurement done in Chapter 5, are presented in Section 6.5 below.

In Section 6.2, the method used to extract the non-perturbative QCD component is presented. In Section 6.3, the extraction is performed for two determinations of the perturbative QCD component using:

- the JETSET 7.3 generator [18], tuned on DELPHI data [83], running in the parton shower mode,
- an analytic computation based on QCD at NLL order [14],
- an analytic computation based on Dressed Gluon Exponentiation (DGE), that also retains NLL accuracy terms [15].

In Section 6.4 these results are discussed and insights obtained with the present method are explained. A parametrization for the non-perturbative QCD component is proposed. In Section 6.5 different verifications are presented that have been performed to check the robustness of the proposed method. Studies of different subjects related to the extraction method are presented. In Section 6.8 the differences between fitting fragmentation functions and their moments are discussed. Finally, in Section 6.9 remarks about the non-perturbative QCD component of the charm fragmentation function are presented.

6.2 Extracting the x-Dependence of the Non-perturbative QCD Component

The method is based on the use of the Mellin transformation which is appropriate when dealing with integral equations as given in (6.1). More details about this integral transformation and necessary formulas for the computations given in the present chapter are quoted in Appendix A. The Mellin transformation of the expression for $\mathcal{D}(x)$ is:

$$\tilde{\mathcal{D}}(N) = \int_0^\infty dx\, x^{N-1}\, \mathcal{D}(x) \tag{6.3}$$

where N is a complex variable. For integer values of $N \geq 2$, the values of $\tilde{\mathcal{D}}(N)$ correspond to the moments of the initial x distribution [1]. For physical processes, x is restricted to be within the $[0,1]$ interval. The interest in using Mellin transformed expressions is that Equation (6.1) becomes a simple product:

$$\tilde{\mathcal{D}}(N) = \tilde{\mathcal{D}}_{pert.}(N) \times \tilde{\mathcal{D}}_{non-pert.}(N) \tag{6.4}$$

Having computed, in the N-space, distributions of the measured and perturbative QCD components, the non-perturbative distribution, $\tilde{\mathcal{D}}_{non-pert.}(N)$ is obtained from Equation (6.4). Applying the inverse Mellin transformation on this distribution one gets $\mathcal{D}_{non-pert.}(x)$ without any need for a model input:

$$\mathcal{D}_{non-pert.}(x) = \frac{1}{2\pi i} \oint dN\, \frac{\tilde{\mathcal{D}}_{meas.}(N)}{\tilde{\mathcal{D}}_{pert.}(N)}\, x^{-N} \tag{6.5}$$

in which the integral runs over a contour in the complex N-plane. The integration contour is taken as two symmetric straight half-lines, one in the upper half and the other in the lower half of the complex plane. The angle of the lines, relative to the real axis, is larger or smaller than 90 degrees for x values smaller or larger than unity, respectively. These lines are taken to originate from $N = (1.01, 0)$. The contour is supposed to be closed by an arc situated at infinity in the negative and positive directions of the real axis, for the two cases respectively. It has been verified that the result is independent of a definite choice for the contour in terms of the slope of the lines and of the value of the arc radius. The result is also independent of the choice for the position of the origin of the lines, on the real axis, as long as the contour encloses the singularities of the expression to be integrated and stays away from the Landau pole present in $\tilde{\mathcal{D}}_{pert.}(N)$ which is discussed in the following. In practice, the Mellin transformed distribution of present measurements, $\tilde{\mathcal{D}}_{meas.}(N)$, has been obtained after having adjusted an analytic expression to the measured distribution in x, and by applying the Mellin transformation on this fitted function. The motivation to proceed in that way rather than calculating the Mellin transformation of the original histograms will be explained in Section 6.5. The following expression, which depends on five parameters and gives a good description of the measurements (see Figure 6.1-left), has been used.

$$D(x) = p_0 \times \left[p_1 x^{p_2} (1-x)^{p3} + (1-p_1) x^{p_4} (1-x)^{p5} \right] \tag{6.6}$$

[1] By definition $\tilde{\mathcal{D}}(1)$ $(= 1)$ corresponds to the normalization of $\mathcal{D}(x)$.

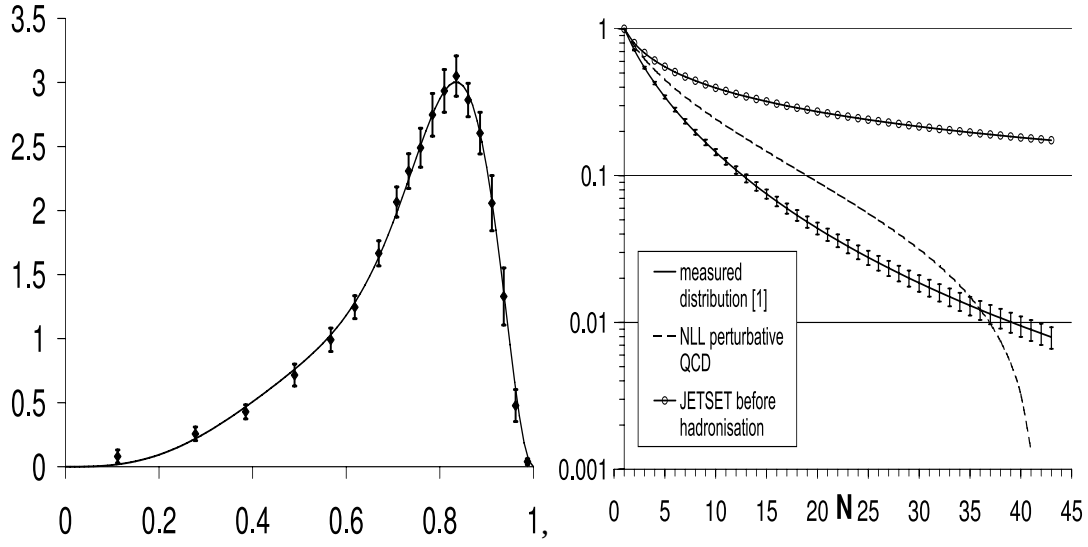

Figure 6.1: *Left: Comparison between the measured (points with error bars) b-fragmentation distribution and the fitted parametrization using Equation* (6.6)*. Right: Moments of the measured (full line) x distribution, of the perturbative QCD component [14](dashed line) and of the generated distribution obtained in JETSET before hadronization (full line with circles). Data from [31] have been used.*

where p_0 is a normalization coefficient. Values of the parameters have been obtained by comparing, in each bin, the measured bin content with the integral of $D(x)$ over the bin. In order to check the effect of a certain choice of parametrization, the whole procedure has been done replacing the expression of Equation (6.6) by another function: a cubic spline, with five intervals between $0 < x < 1$, continuous up to the second derivative, normalized to 1, and forced to be 0 at $x = 0$ and $x = 1$. This function also depends on five parameters. The results obtained with the two parametrizations have been found to be similar. The comparison is presented in Section 6.5.

Measurements of the *b*-fragmentation distribution, in [31], have been published in a binned form, after unfolding of the experimental energy resolution. Values in the bins are correlated and, as the bin width is smaller than the resolution, the error matrix is singular. Only positive eigenvalues of this matrix have been considered and the seven largest eigenvalues have been used. A detailed explanation of this fit method is presented in Appendix B.

The distribution of moments obtained with data from [31], and computed using the fitted distribution corresponding to Equation (6.6), is given in Figure 6.1-Right. The N representation of the fit function of Equation (6.6) is analytically calculable, as explained in Appendix A. The expression is:

$$\tilde{\mathcal{D}}(N) = p_0 \left[p_1 \frac{\Gamma(p_2+N)}{\Gamma(p_2+p_3+N+1)} + (1-p_1) \frac{\Gamma(p_4+N)}{\Gamma(p_4+p_5+N+1)} \right] \tag{6.7}$$

Quoted uncertainties, in Figure 6.1-Right, correspond to actual measurements and are highly correlated. They have been obtained by propagating uncertainties corresponding to the covariance matrix of the $p_{1,..,5}$ fitted parameters. When computing the moments, very similar results

are obtained, for $N < 10$, using directly the measured x-binned distribution. For higher N values, effects induced by the variation of the distribution within a bin, as expressed by Equation (6.6), have to be included. More details about this argument will be given in Section 6.5.

The Mellin transformed distribution of the JETSET perturbative QCD component has been obtained in a similar way, whereas the NLL QCD perturbative component is computed directly as a function of N in [14]. At large values of N, this last distribution is equal to zero for $N = N_0 \simeq 41.7$ and has a Landau pole situated at $N_L \simeq 44$. Values for N_0 and N_L depend on the exact values assumed for the other parameters entering into the computation; see Section 6.3.2 where values of these parameters have been listed.

6.3 x-Dependence Measurement of the Non-perturbative QCD Component

The x distribution of the non-perturbative QCD component extracted in this way depends on the measurements and also on the procedures adopted to compute the perturbative QCD component. In the following, two approaches have been considered. The first one is generally adopted by experimentalists whereas the second is more frequent for theorists.

6.3.1 The Perturbative QCD Component is Provided by a Generator

A detailed explanation of parton shower Monte Carlo generators and string hadronization is given in Section 2.3.

The JETSET 7.3 Monte Carlo generator, with values of the parameters tuned on DELPHI data registered at the Z pole has been used [2]. Events have been produced using the parton shower option of the generator and the b-quark energy is extracted, after radiation of gluons, just before calling the routines to create a B-hadron that takes a fraction z of the available string energy. $z = \frac{E^B + p_L^B}{E^b + p_L^b}$ is the boost-invariant fraction of the b-jet energy taken by the weakly decaying B meson. This variable is defined for a string stretched between the b-quark and a gluon, an anti-quark or a diquark. In the present analysis only distributions in terms of the x variable have been used as no string model has been considered. The x distribution for b-quarks, after gluon radiation, is displayed in Figure 6.3. It has to be complemented by a δ-function at $x = 1$ which contains $\sim 4\%$ of all events. In this peak, B-hadrons carry all the energy of the b-quark as no gluon has been radiated.

Applying the method explained in Section 6.2, the corresponding non-perturbative QCD component has been extracted, and is displayed also in Figure 6.3. Above $x = 1$, it is compatible with zero, as expected.

The quoted error bar, for a given value of x, has been obtained by evaluating the values of $\mathcal{D}_{non-pert.}(x)$ for different shapes of the b-quark fragmentation distribution which are obtained by varying parameters $p_{1,..,5}$ according to their measured error matrix. This matrix has been obtained by propagating the errors of the measured distribution to the fitted parameters. The program that generated sets of variables $p_{1,..,5}$ took into account also the asymmetrical individual uncertainty for each parameter. The resulting curves, that have been used to estimate the errors

[2] It has been verified that the ALEPH tuning [84] of this generator gives similar results. See also Section 6.5.

on the extracted non-perturbative component are shown in Figure 6.2, together with the curve that represents the central values of $p_{1,..,5}$. The extraction procedure has been applied on each one of the generated curves, for selected points in x. The error bar represents, for a given value of x, the standard deviation of the computation results for the different curves.

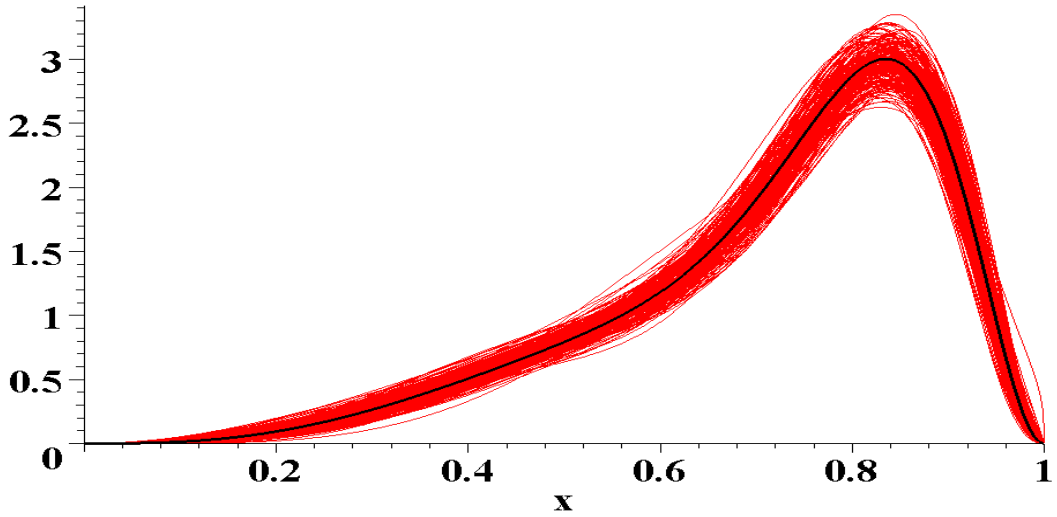

Figure 6.2: *Curves that have been used for the error bar generation (red), together with the central curve that represents the central values of the parameters $p_{1,..,5}$ (black).*

6.3.2 The Perturbative QCD Component is Obtained by an Analytic Computation Based on QCD

The perturbative QCD fragmentation function is evaluated according to the approach presented in [14]. This next-to-leading log (NLL) accuracy calculation for the inclusive b-quark production cross section in e^+e^- annihilation, generalizes previous calculations by resumming the contribution from soft gluon radiation (which plays an important role at large x) to all perturbative orders and to NLL accuracy. These computations are done directly in the N-space. As explained in Section 2.3.3.1, soft gluon radiation contributes to the logarithm of the fragmentation function large logarithmic terms of the type $(\log N)^p$, in the exponent, with $p \leq n+1$. These terms appear at all perturbative orders n in α_s. In the calculation at NLL accuracy [14], the two largest terms, corresponding to $p = n+1$ and n, have been resummed at all perturbative orders. The calculation is expected to be reliable when N is not too large (typically less than 20). To obtain distributions for the variable x from results in moment space, one should apply the inverse Mellin transformation, that consists in integrating over a contour in N (Section 6.2 and Appendix A). When x gets closer to 1, large values of N contribute and thus the perturbative fragmentation distribution is not reliable in these regions. This behavior affects also values of the distribution at lower x as moments of this distribution are fixed. In addition to the breakdown of the theory for large values of N, uncertainties attached to the determination of the theoretical perturbative QCD component are related to the definition of the scales entering into the computation. This component also depends on two parameters: the b-quark pole mass (m_b) and

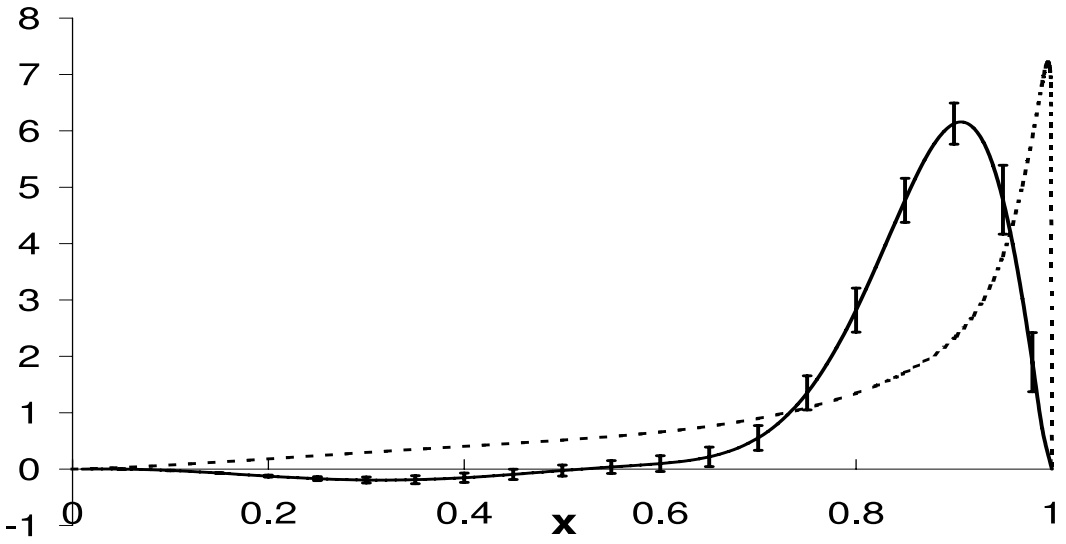

Figure 6.3: *x-dependence of the perturbative (dotted line) and non-perturbative (full line) QCD components of the measured [31] b-fragmentation distribution. These curves are obtained by interpolating corresponding values determined at a large number of points in the x-variable. Quoted error bars correspond to measurement uncertainties and are correlated for different x-values. The perturbative QCD component is extracted from the JETSET 7.3 Monte Carlo generator. The dotted curve has to be complemented by a* δ*-function containing 4% of the events, located at* $x = 1$.

$\Lambda^{(5)}_{QCD}$, that have been taken as $m_b^{pole} = (4.75 \pm 0.25)$ GeV/c^2 and $\Lambda^{(5)}_{QCD} = (0.226 \pm 0.025)$ GeV. Scale and parameter depending variations of the moments of the perturbative QCD component are given in Figure 6.4. These variations are fully correlated versus N.

The extracted non-perturbative component is given in Figure 6.5. Its shape depends on the same quantities as those used to evaluate the perturbative distribution, and thus similar variations appear, as drawn also in the Figure.

It has to be noted that the data description in terms of a product of two QCD components, perturbative and non-perturbative, is not directly affected by uncertainties attached to the determination of the perturbative component. This is because the non-perturbative component, as determined in the present approach, compensates for a given choice of method or of parameter values.

To obtain the complete expected x-distribution of B-hadrons from the NLL theoretical calculation, one has to be able to evaluate the integral given in Equation (6.1). When x becomes close to 1 the distribution is sensitive to the large N breakdown of the theory. Consequently, the high-x ($x > 0.96$) behavior of the perturbative QCD component has been studied. As mentioned before, this region corresponds to high-N values where the perturbative approach fails. As a result, the high-x behavior of the distribution is non-physical; it oscillates. To have a numerical control of the distribution in this region, it has been decided to take into account x values which are below a given maximum value, x_{max}, above which the distribution is assumed to be equal to zero. Moments of this truncated distribution show a small discrepancy when compared with moments of the full distribution. This difference has an almost linear dependence with N. To

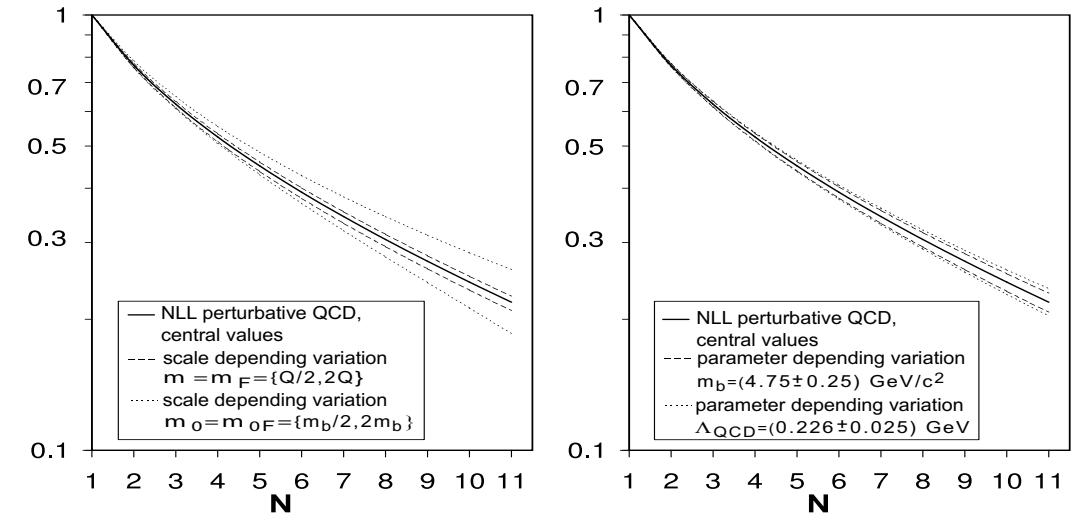

Figure 6.4: *Variations of the calculated perturbative QCD component [14], depending on the renormalization scales (μ, μ_0), the factorisation scales (μ_F, μ_{0F}), the b-quark mass (m_b) and $\Lambda^{(5)}_{QCD}$. The full lines are corresponding to the central values ($\mu = \mu_F = Q = 91.2$ GeV, $\mu_0 = \mu_{0F} = m_b$, $m_b = 4.75$ GeV/c^2 and $\Lambda^{(5)}_{QCD} = 0.226$ GeV).*

correct for this effect, x_{max} is chosen such that the difference between moments is a constant value (the slope in N being close to zero at this point). This difference can then be corrected by adding simply a δ-function at $x = 1$, so that the total distribution is normalized to 1. A typical value for x_{max} is 0.997 and the δ component corresponds to 5% of the distribution. We stress that the truncation at x_{max} and the added δ-function do not contribute in the determination of the non-perturbative component using Equation (6.5). But this procedure is necessary for checking that the extracted non-perturbative component in the x-space, when convoluted with the perturbative distribution, effectively reproduces the measurements and also for testing hadronic models given in the x-space.

Unlike the perturbative QCD component which was defined in [14] within the $[0,1]$ interval, the non-perturbative component has to be extended in the region $x > 1$. This "non-physical" behavior comes from the zero of $\tilde{\mathcal{D}}_{pert.}(N)$ for $N = N_0$ which gives a pole in the expression to be integrated in Equation (6.5). Using properties of integrals in the complex plane, it can be shown that, for $x > 1$, the non-perturbative QCD distribution can be well approximated by x^{-N_0}.

Errors bars, given in Figure 6.5, have been obtained using the same procedure as explained in Section 6.3.1.

The extraction method has been also applied to the NLL+DGE computation of [15]. As explained in Section 2.3.3.1 this approach is different from the former one by the fact that it adds an approximate resummation of all the subleading logarithmic terms. When applying the extraction method on the NLL+DGE calculation, the inverse Mellin transformation of the non perturbative QCD component becomes numerically complicated when x is large. The extraction has been therefore performed only until $x = 0.96$.

In Figure 6.6 the x-dependence of the non-perturbative QCD component is presented for

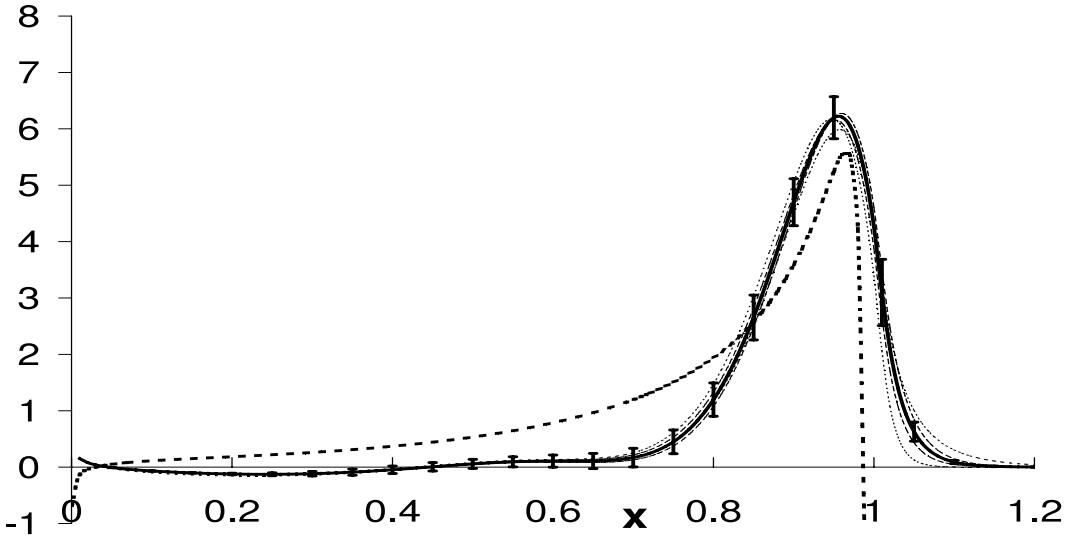

Figure 6.5: *x-dependence of the perturbative (dotted line) and non-perturbative (full line) QCD components of the measured b-fragmentation distribution. These curves are obtained by interpolating corresponding values determined at a large number of points in the x-variable. The perturbative QCD component is given by the analytic computation of [14]. The thin lines on both sides of the non-perturbative distribution correspond to* $\mu_0 = \mu_{0F} = \{m_b/2, 2m_b\}$ *(dotted lines) and* $\Lambda^{(5)}_{QCD} = (0.226 \pm 0.025)$ GeV *(dashed lines). Variations induced by the other parameters,* $\mu = \mu_F = \{Q/2, 2Q\}$ *and* $m_b = (4.75 \pm 0.25)$ GeV/c^2 *are smaller. In addition, quoted error bars correspond to measurement uncertainties and are correlated for different x-values. The perturbative QCD dotted curve has to be complemented by a* δ*-function containing 5% of the events, located at* $x = 1$.

all the three QCD approaches treated in this chapter. From this figure it is clear that when the accuracy of the QCD approach increases, the non-perturbative component is shifted towards larger x values. This reflects the fact that the more logarithmic divergences are resummed, the higher becomes the x value up to which the perturbative QCD computation is reliable. When the perturbative QCD component is "more reliable" it needs to be complemented by a smaller non-perturbative one. It is also clear from this figure that for $x < 0.6$ all of the three approaches give a vanishing non-perturbative component. In that region [3], gluon radiation is well accounted for, even in the Monte Carlo generator's LL approach. None of the perturbative QCD approaches examined above needs to be complemented by a non-perturbative component in the low x region. The non-perturbative distributions obtained for perturbative components from a Monte Carlo generator or from a NLL QCD computation are similar in shape. This is not the case for the NLL+DGE QCD computation.

[3] The low x region corresponds typically to hard gluon radiation

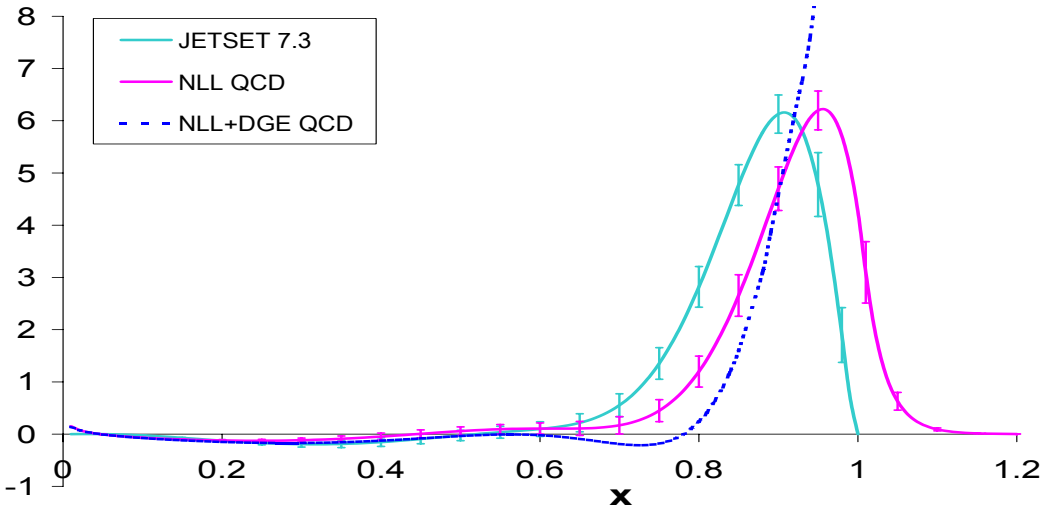

Figure 6.6: *Comparison between the not-perturbative component obtained for the three perturbative QCD approaches treated in this chapter: JETSET, a NLL theoretical QCD computation [14] and a theoretical NLL+DGE computation [14]*

6.4 Results Interpretation

The x-dependence of the non-perturbative QCD component, obtained in this way, does not depend on any non-perturbative QCD model assumption but its shape is tightly related to the procedures used to evaluate the perturbative component, and thus, the two distributions have to be used jointly.

6.4.1 Comparison with Models

Non-perturbative components of the b-quark fragmentation distribution, taken from models, have been folded with a perturbative QCD component obtained from a Monte Carlo generator and compared with measurements by the different collaborations [31, 34, 32, 33]. The Lund and Bowler models were favored in these comparisons. In Figure 6.7 the directly extracted non-perturbative components are compared with distributions taken from models [37, 35, 36, 38, 24] whose parameters have been fitted on data from [31]. Results have been obtained by comparing, in each bin, the measured bin content with the integral, over the bin, of the folded expression for $\mathcal{D}_{predicted}(x)$. They are thus independent of the analytic expression given in Equation (6.6). These results are summarized in Table 6.1. It must be noted that parameters given in this Table, when the perturbative QCD component is taken from JETSET, may differ from those quoted in original publications as an analytic computation is done in the present study, using Equation (6.1), whereas a string model is used in the former; other sources of difference can originate from the exact values of the parameters used to run JETSET and from the definition of x which varies between 0 and 1 in our case. Numerically, if one compares present results with those obtained by OPAL [32], in their own analysis, for the values of the parameters of the Lund and Bowler models, differences are situated well within quoted uncertainties.

When the perturbative QCD component is taken from the analytic NLL calculation the Lund

Model	JETSET		NLL Pert. QCD	
	param.	χ^2/NDF	param.	χ^2/NDF
Kartvelishvili [37] $x^{\varepsilon_b}(1-x)$	$\varepsilon_b = 12.3 \pm 0.7 \pm 0.4$	$35/6$	$\varepsilon_b = 14.6 \pm 0.6$	$60/6$
Peterson [35] $\frac{1}{x}\left(1-\frac{1}{x}-\frac{\varepsilon_b}{1-x}\right)^{-2}$	$\varepsilon_b = (4.1^{+0.4\ +0.2}_{-0.3\ -0.1}) \times 10^{-3}$	$47/6$	$\varepsilon_b = (2.8 \pm 0.2) \times 10^{-4}$	$82/6$
C.S [36] $\left(\frac{1-x}{x}+\frac{\varepsilon_b(2-x)}{1-x}\right)(1+x^2)\left(1-\frac{1}{x}-\frac{\varepsilon_b}{1-x}\right)^{-2}$	$\varepsilon_b = (3.3 \pm 0.5^{+0.4}_{-0.9}) \times 10^{-3}$	$117/6$	$\varepsilon_b = (1.1 \pm 0.3) \times 10^{-4}$	$37/6$
Lund [38] $\frac{1}{x}(1-x)^a \exp\left(-\frac{bm_{b\perp}^2}{x}\right)$	$a = 1.68 \pm 0.18^{+0.09}_{-0.06}$ $bm_{b\perp}^2 = 15.6 \pm 1.2^{+0.5}_{-0.4}$	$7/5$	$a = 0.00 \pm 0.04$ $bm_{b\perp}^2 = 9.3 \pm 0.8$	$11/5$
Bowler [24] $\frac{1}{x^{1+bm_{b\perp}^2}}(1-x)^a \exp\left(-\frac{bm_{b\perp}^2}{x}\right)$	$a = 0.89 \pm 0.11 \pm 0.10$ $bm_{b\perp}^2 = 75. \pm 9.^{+.5}_{-9.}$	$18/5$	$a = 0.00 \pm 0.02$ $bm_{b\perp}^2 = 60. \pm 6.$	$8/5$

Table 6.1: *Values of the parameters and of the χ^2/NDF obtained when fitting results from Equation* (6.1)*, obtained for different models of the non-perturbative QCD component, to the measured b-fragmentation distribution. The two situations corresponding, respectively, to the perturbative QCD component taken from JETSET or from [14] have been distinguished. The Lund and Bowler models have been simplified by assuming that the transverse mass of the b-quark, $m_{b\perp}$, is a constant. The last quoted uncertainty, in the second column corresponds to the variation induced by selecting, in the fitting procedure, between five and nine eigenvalues of the measured error matrix. Such an uncertainty is not given in column four as corresponding fits of models are only indicative there.*

and Bowler models give also the best χ^2. Values for the parameter a in these models are even compatible with zero corresponding to a behavior in $1/z^{\alpha}\ \exp(-A/z)$, to accommodate the non-zero value of the non-perturbative QCD component at $x = 1$.

But, as these models have no contribution above $x = 1$, their folding with the perturbative QCD component cannot compensate for the non-physical behavior of the latter. The folded distribution oscillates at large x-values (see Figure 6.8-Left). In particular, the predicted value in the last measured x-bin, which is found to be in reasonable agreement with the measurements after the fitting procedure when using the Lund and Bowler models, results from a large cancellation between a negative and a positive contribution within that bin. This is more clearly seen when considering the moments of the overall distribution, which are given by Equation (6.4). For moments of order N, the weight x^{N-1} introduces a variation within the x-bin size, which was not accounted for in the previous fit in x, and effects are amplified mainly at large N values which correspond to the high x region. This is illustrated in Figure 6.8-Right.

This study indicates that all models have to be discarded when folded with the NLL perturbative QCD fragmentation distribution, if one wishes to correctly describe the full measured x spectrum. The goodness of the fit in x, as measured by the corresponding χ^2 value, does not reflect all the information because it was not required that the folded distribution remains physical (positive) over the $[0,1]$ interval. This folding procedure has thus to be considered only as an exercise and the non-perturbative QCD distribution has to be extracted from data.

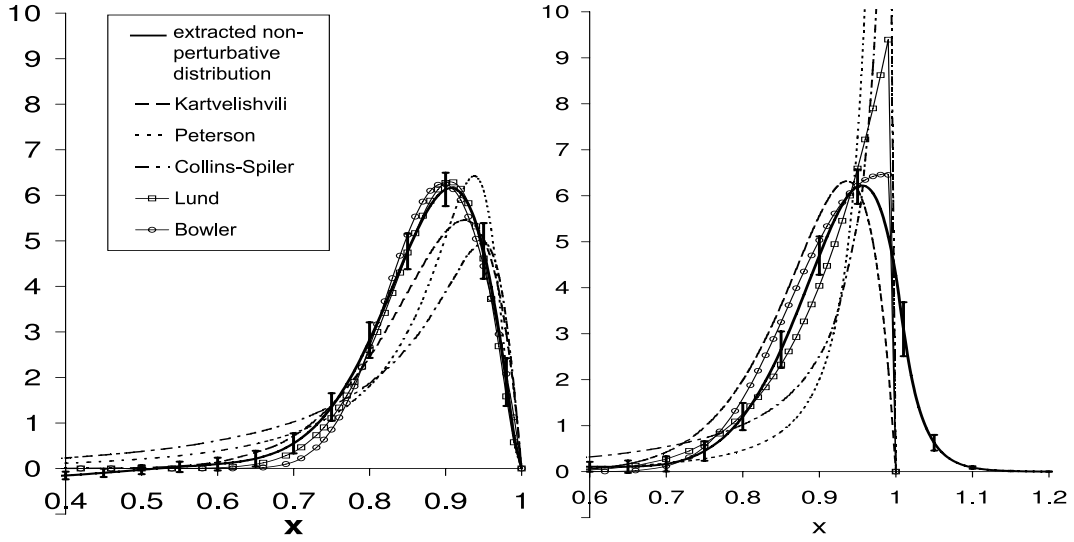

Figure 6.7: *Comparison between the directly extracted non-perturbative component (thick full line) and the model fits on data taken from [31]. Left: the perturbative QCD component is taken from JETSET. Right: the theoretical perturbative QCD component[14] is used.*

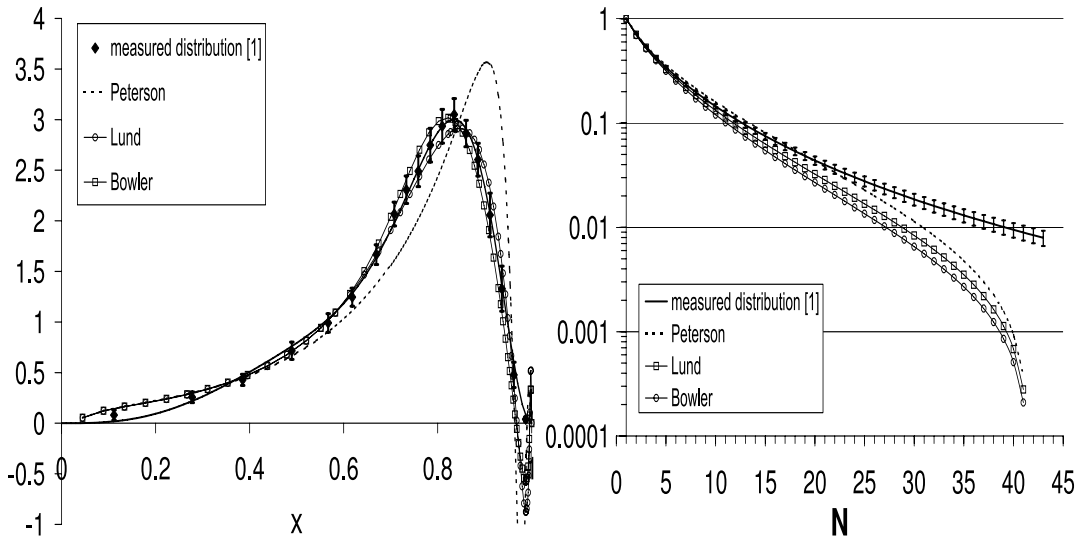

Figure 6.8: *Comparison between the measured and fitted x-distributions using different models. Left: the measured binned-distribution in x is compared with fitted results. Right: moments of the corresponding distributions are compared.*

6.4.2 Proposal for a New Parametrization

As explained in Section 6.2, the non-perturbative QCD component of the *b*-fragmentation distribution has been extracted independently of any hadronic model assumption but it depends on the modeling of the perturbative QCD component.

When a Monte-Carlo generator is used to obtain the perturbative component, it can be ver-

ified that the non-perturbative component, extracted in this way, has a physical behavior for all x-values. In Figure 6.3, below $x = 0.6$, $(5^{+3}_{-2})\%$ of the integrated distribution is negative. A larger deviation would have indicated some inconsistency between experimental measurements and gluon radiation, as implemented in the generator.

Such a test cannot be made, a priori, when the perturbative QCD component is taken from an analytic computation, as this distribution is already unphysical in some regions. The non-perturbative extracted distribution, as given in Figure 6.5, is expected to compensate precisely for these effects. It can be noted that, also in this case, the distribution is compatible with zero below $x = 0.6$. This shows that the perturbative QCD evaluation of hard gluon radiation is in agreement with the measurements. The small spike, close to $x = 0$, related to the multiplicity problem [85] in the perturbative evaluation, has no numerical effect in practice.

To provide an analytic expression, which agrees better with the extracted point-to-point non-perturbative QCD distribution, the following function has been used:

$$\begin{aligned}
\mathcal{D}_{non-pert.}(x) &= N(\mathbf{p}) \exp -\frac{(x-x_0)^2}{2\sigma_-^2} \text{ for } x < x_0; \\
&= N(\mathbf{p}) \exp -\frac{(x-x_0)^2}{2\sigma_+^2} \text{ for } x_0 < x < 1; \\
&= N(\mathbf{p}) \exp -\frac{(1-x_0)^2}{2\sigma_+^2} x^{-N_0} \text{ for } x > 1;\ \mathbf{p} = (x_0, \sigma_-, \sigma_+)
\end{aligned} \tag{6.8}$$

$N(\mathbf{p})$ is a normalisation factor such that the integral of the expression, given in Equation (6.8), between 0 and ∞ is equal to unity. This distribution corresponds, for $0 < x < 1$, to Gaussian distributions with different standard deviations when x is situated on either sides of x_0. As explained in Section 6.2, the behaviour of the non-perturbative distribution for $x > 1$ is related to the presence of a zero in $\mathcal{D}_{pert.}(N)$ located at $N = N_0$. When the perturbative distribution is taken from a Monte-Carlo, there is not such a zero and it has been considered that $\mathcal{D}_{non-pert.}(x) = 0$ when $x > 1$. When the perturbative distribution is taken from theory, the value of N_0 has been fixed to 41.7. The value for N_0 depends a lot on the central values for the parameters and for the scales, adopted in the perturbative QCD evaluation:

$$N_0 = 41.7^{-5.8}_{+4.5}(\Lambda^{(5)}_{QCD}) \pm 2.5(m_b)^{+34.6}_{-21.7}(\mu_0 = \mu_{0F}) \tag{6.9}$$

N_0 is independent on the value assumed for the other scale $\mu = \mu_F$.

The parameter values, obtained when the perturbative QCD component is taken either from JETSET or from theory, are given in Table 6.2. It can be noted that the shape of the non-perturbative QCD distribution is similar in the two cases, the maximum being displaced to an higher value in the latter. The correlation matrices for the fitted parameters are given in Table 6.3.

Comparison, in x- and in moment-space, between the measured and fitted distributions, using the perturbative QCD component of [14] are given in Figure 6.9. For moments higher than 40, the fitted moments strongly deviate from the measurements as one is approaching the Landau pole and the formalism is no longer valid.

	x_0	σ_-	σ_+	χ^2/NDF
JETSET	0.925 ± 0.003	0.094 ± 0.005	0.031 ± 0.004	$6/4$
NLL pert. QCD	0.961 ± 0.006	0.091 ± 0.006	0.038 ± 0.006	$2/4$

Table 6.2: *Values for the fitted parameters of the non-perturbative QCD component corresponding to JETSET and NLL perturbative QCD using data from [31].*

JETSET				NLL pert. QCD			
	x_0	σ_-	σ_+		x_0	σ_-	σ_+
x_0	1	0.23	−0.73	x_0	1	0.18	−0.83
σ_-	0.23	1	−0.41	σ_-	0.18	1	−0.27
σ_+	−0.73	−0.41	1	σ_+	−0.83	−0.27	1

Table 6.3: *The correlation matrices on the fitted parameters of the proposed parametrization.*

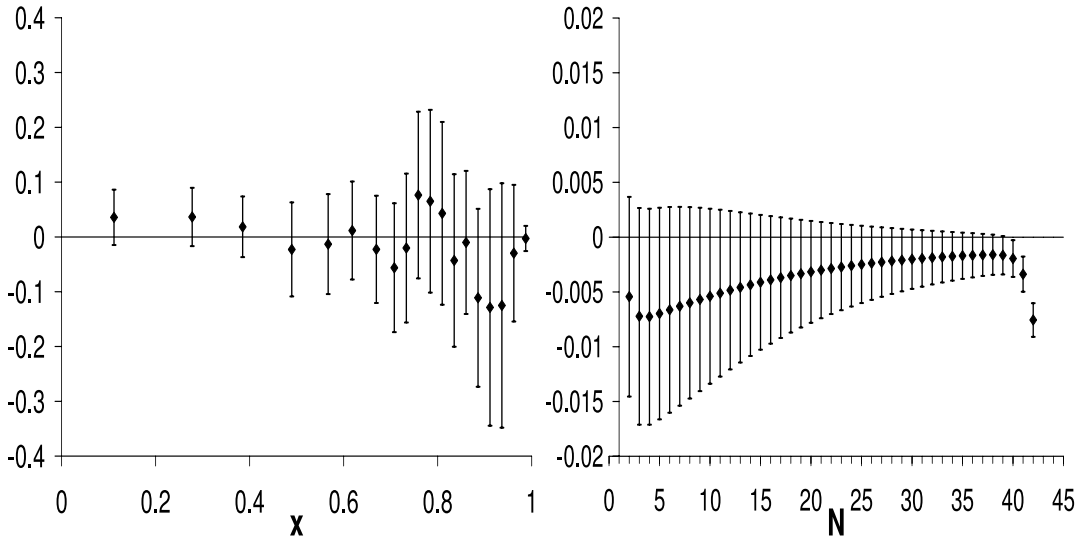

Figure 6.9: *Comparison between the fitted and measured b-quark fragmentation distributions. Left: Differences between the fitted and the measured distributions in each x-bin are shown and the total error bars, relative to the measurements, are displayed. The fitted results correspond to the averaged values obtained from Equation* (6.1) *over the bin. Right: Differences between the fitted and measured moments of the b-fragmentation distribution. Fitted moments result from the product given in Equation* (6.4) *in which* $\tilde{\mathcal{D}}_{non-pert.}(N)$ *are the moments of the fitted non-perturbative QCD distribution corresponding to Equation* (6.8) *and* $\tilde{\mathcal{D}}_{pert.}(N)$ *corresponds to the analytic computation [14].*

6.5 Checks

In the present chapter, a few checks have been mentioned. They have been designed in order to verify the robustness of the extracted non-perturbative component under different changes in the extraction method. Some of these checks are given here with some details and illustrated with figures.

6.5.1 The Use of a Fitted Parametrization

As explained in Section 6.2, the Mellin transformed distribution of the measured fragmentation function has been obtained after having adjusted an analytic expression to the measured distribution in x, and by applying the Mellin transformation on this fitted function. The reason for this choice is that high order moments are sensitive to the shape of the distribution, and are therefore influenced by the variation of the distribution within a bin. This issue is crucial when applying the inverse Mellin transformation to extract the non-perturbative QCD component. The integration contours sometimes need to reach values of large $|N|$ in order to enable evaluation of the inversion integral [4]. The effect of binning on moments of the measured b fragmentation function has been estimated using a "toy", that allowed to isolate the problem. The parametrization of Equation (6.6) that has been fitted to the fragmentation function from ALEPH has been considered as the real physical fragmentation function for this check. Integrals of the fitted function have been calculated within the 19 bins chosen by ALEPH. Then, moments have been obtained for the smooth function and for its binned version. The error on these moments have been estimated by propagating the errors of the fitted parameters. Figure 6.10 presents the difference between moments of the smooth and the binned distributions divided by the error, as a function of the moment order. With the 19 bins chosen by ALEPH for their fragmentation analysis, the effect of binning is of 1 σ at a moment of order ~ 50. The same check has been done for 10 bins in x, and for that case a 1 σ difference is observed already at a moment of order 10.

6.5.2 The Effect of Parametrization

The possible effect of a particular choice of parametrization has been also checked. The extraction of the non-perturbative component has been performed replacing the expression of Equation (6.6) by another function: a cubic spline, with five intervals between $0 < x < 1$, continuous up to the second derivative, normalized to 1, and forced to be 0 at $x = 0$ and $x = 1$. This function also depends on five parameters. The computation of the Mellin transformation of the spline is presented in Appendix A. The results obtained with the two parametrizations have been found similar. These results for the fragmentation function from ALEPH, when the perturbative QCD component has been taken from JETSET are presented in Figure 6.11.

6.5.3 Number of Degrees of Freedom

As mentioned in Section 6.2, the fragmentation distribution from ALEPH, that has been used in the present chapter, has a singular error matrix. In order to fit a parametrization to

[4] The inversion integral converges typically for $|N| \sim 1000$ for large x

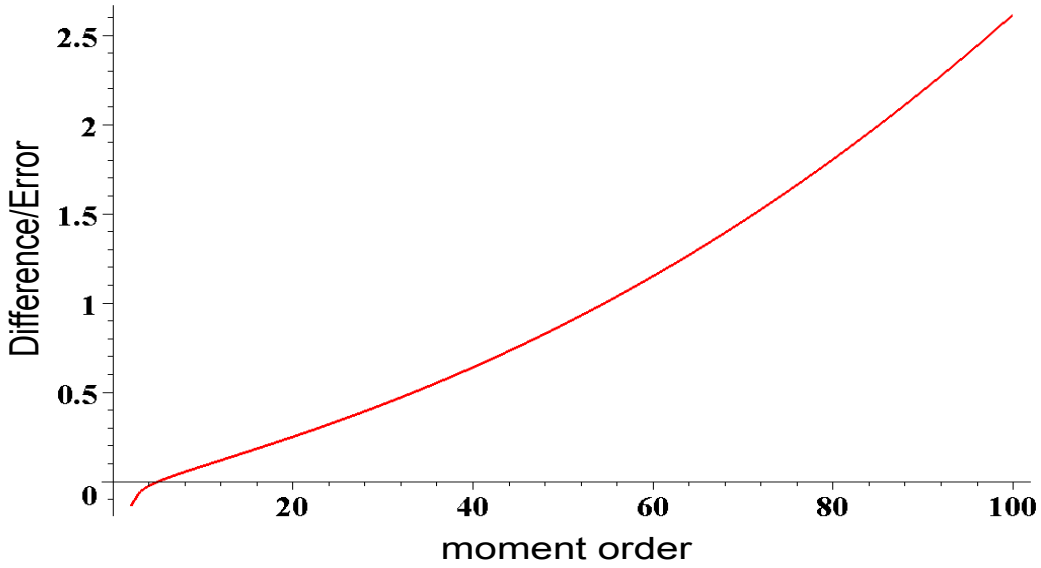

Figure 6.10: *The difference between moments of a smooth fragmentation distribution and its binned version divided by the error as a function of the moment order.*

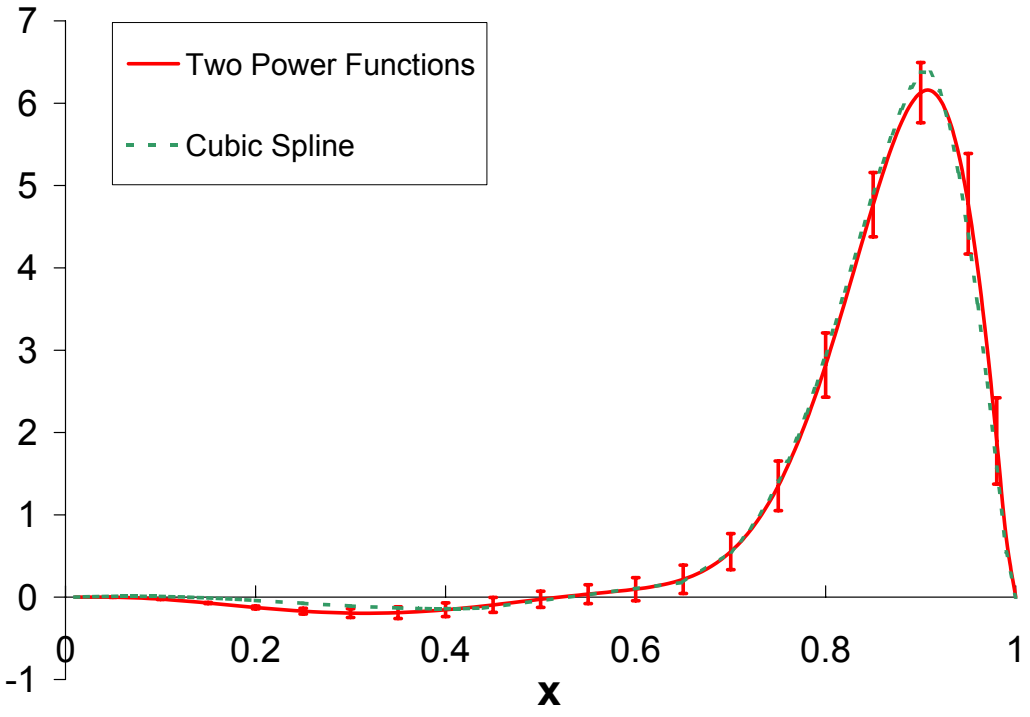

Figure 6.11: *The extracted non-perturbative component of the fragmentation function from ALEPH, obtained using two different parametrizations: the power function of Equation* (6.6) *and a cubic spline. The perturbative QCD component has been taken from JETSET.*

the distribution, only the seven largest eigenvalues of the error matrix have been used. The fit method is explained in Appendix B. The effect of the number of eigenvalues that have been taken into account is evaluated as a systematic error of the model fits in Section 6.4.1.

Here, in order to illustrate the effect of the number of eigenvalues considered in the fit, the fragmentation function measured by OPAL [32] is used. The error matrix of this distribution is also singular and has 14 positive eigenvalues out of 20. The extraction of the non-perturbative QCD component has been performed using the largest positive eigenvalues, with a different number of eigenvalues at each time. The results for 6, 7 and 8 eigenvalues are presented in Figure 6.12. In the same figure, is also presented the result obtained by using only the diagonal elements of the error matrix, neglecting the correlations between bins. The parameters obtained for all these cases were found to be compatible. The perturbative QCD component in these computations has been taken from JETSET.

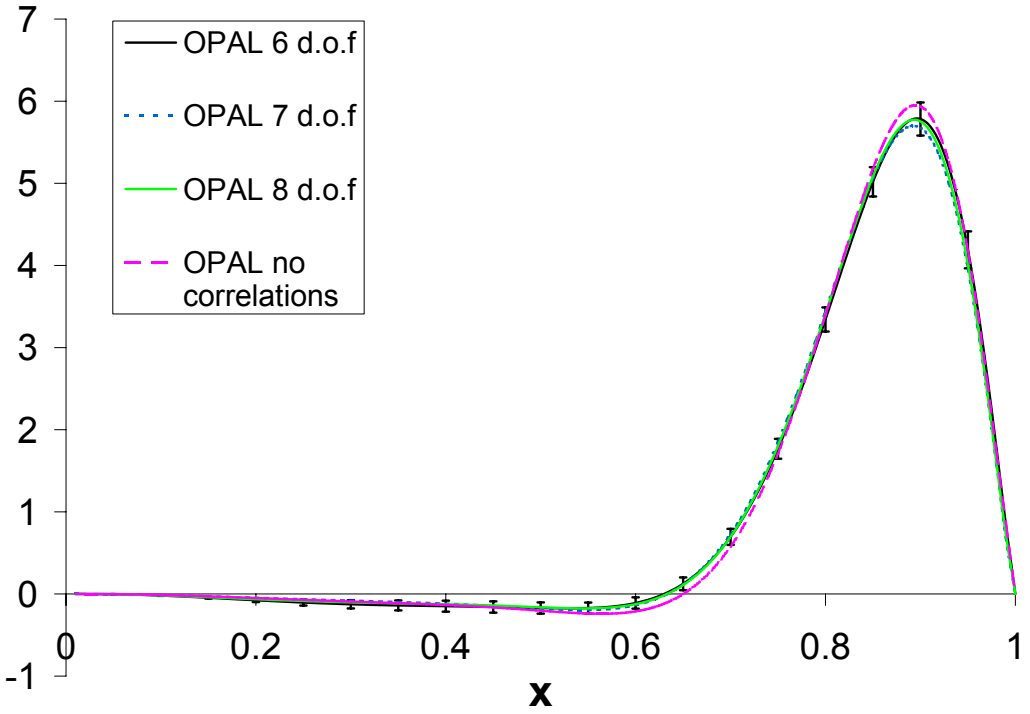

Figure 6.12: *The non-perturbative component of the fragmentation function from OPAL [32], obtained with a different number of eigenvalues in the fit of a parametrization to the measured function. The result for the case where correlations between bins have been completely neglected is also presented.*

6.5.4 Using a Different Tuning of the Monte Carlo

In this chapter, when taking the perturbative QCD component from JETSET 7.3, the generator's parameters tuned on DELPHI data have been used. In order to check the effect of this particular tuning of the generator on the extracted non-perturbative component, the procedure has been also applied for JETSET 7.3 with the ALEPH tuning [84]. The extracted

non-perturbative distributions obtained in the two cases are presented in Figure 6.13; results are hardly distinguishable.

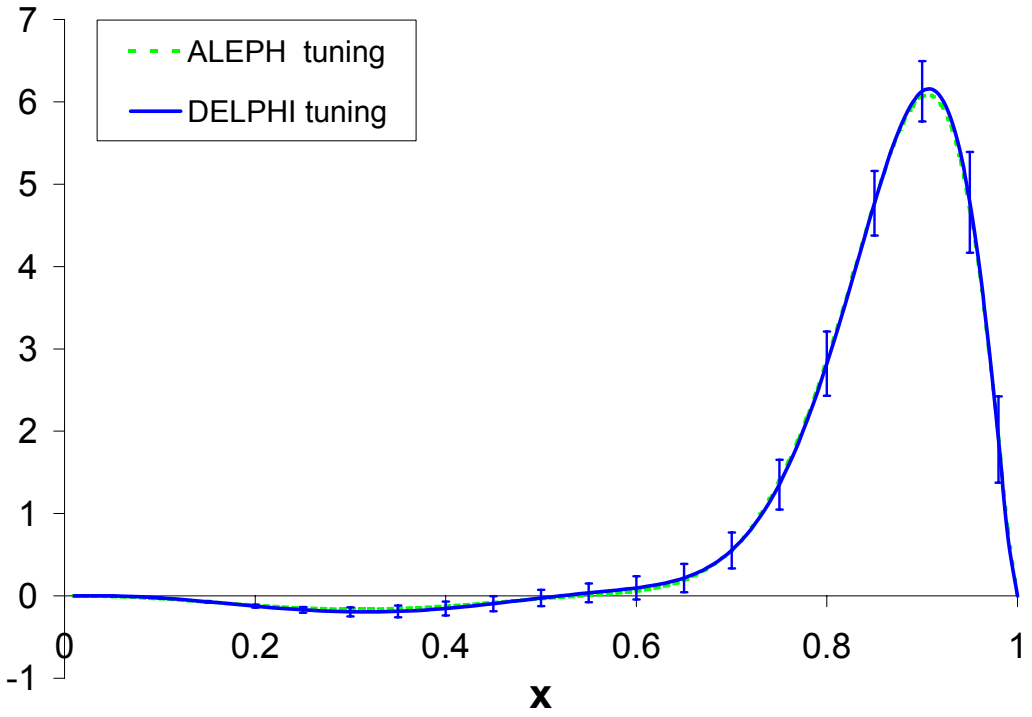

Figure 6.13: *The non-perturbative component for jetset with aleph tuning.*

6.6 Combination of Fragmentation Distributions from All Experiments

The fragmentation distributions from ALEPH, OPAL and SLD [31, 32, 33], together with the result from DELPHI presented in this thesis, have been combined to give a world average *b* fragmentation distribution. Each measurement is given with a different choice of binning and has a different number of effective degrees of freedom. In order to combine the results, a global fit has been done, using the smooth parametrization of equation (6.6). The χ^2, minimized in the fit, is the sum of χ^2 for the different experiments, computed as explained in Section 6.2. The number of degrees of freedom for ALEPH, OPAL and DELPHI has been chosen to yield a χ^2 per degree of freedom close to unity, following the prescription of Appendix B. For SLD, the diagonal error matrix has been used, as the full error matrix was not detailed in [33]. The fit has been done for the variable x_E.

The fitted parameters $p_1..p_5$, that represent the world average fragmentation distribution,

have been found to be:

$$\begin{aligned} p_1 &= 11.29 \pm .41 \\ p_2 &= 2.374^{+0.091}_{-0.089} \\ p_3 &= 2.101^{+0.082}_{-0.073} \\ p_4 &= 1.24^{+0.19}_{-0.16} \\ p_5 &= 0.630^{+0.027}_{-0.028} \end{aligned} \tag{6.10}$$

The full error matrix on $p_1..p_5$ is given in Table 6.4.

$$\begin{array}{c} \\ p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \end{array} \begin{pmatrix} p_1 & p_2 & p_3 & p_4 & p_5 \\ .16615 & .034469 & -.021158 & -.044472 & .0063418 \\ .034469 & .0082353 & -.0046156 & -.010354 & .0012780 \\ -.021158 & -.0046156 & .0067565 & .015063 & -.0020707 \\ -.044472 & -.010354 & .015063 & .034758 & -.0047225 \\ .0063418 & .0012780 & -.0020707 & -.0047225 & .00073397 \end{pmatrix}$$

Table 6.4: *The error matrix on the fitted parameters of the combined fragmentation distribution.*

The moments of the combined distribution are given in Table 6.5. They have been calculated using the analytic expression of Equation (6.7) with the fitted values of the parameters $p_1..p_5$. The error have been propagated from the error matrix on the parameters.

6.7 Comparison of Results for All Experiments

The non-perturbative QCD component has been extracted for all available measured *b* fragmentation functions from e^+e^- colliders [31, 34, 32, 33], including the result from DELPHI presented in this thesis, and the combined fragmentation function of Section 6.6. The extraction has been done for the two main perturbative QCD approaches considered in this chapter: JETSET 7.3 parton shower Monte Carlo and the theoretical NLL QCD computation from [14]. Comparisons of the results for the two perturbative QCD approaches are shown in Figures 6.14, 6.15 and 6.16. The distribution measured in the present thesis work is quite similar with those obtained from SLD and OPAL measurements and is somewhat softer, but compatible within uncertainty, compared with the one from ALEPH.

6.8 Thoughts about Fitting Moments of Fragmentation Functions

Moments of distributions are naturally strongly correlated. In the case of fragmentation functions, where bins of the measured distributions of the *x* variable are highly correlated themselves, correlations between moments are even higher. The typical value of correlation between two consecutive moments of a measured fragmentation function is $1-\varepsilon$ when ε is $\sim 0.1\%$.

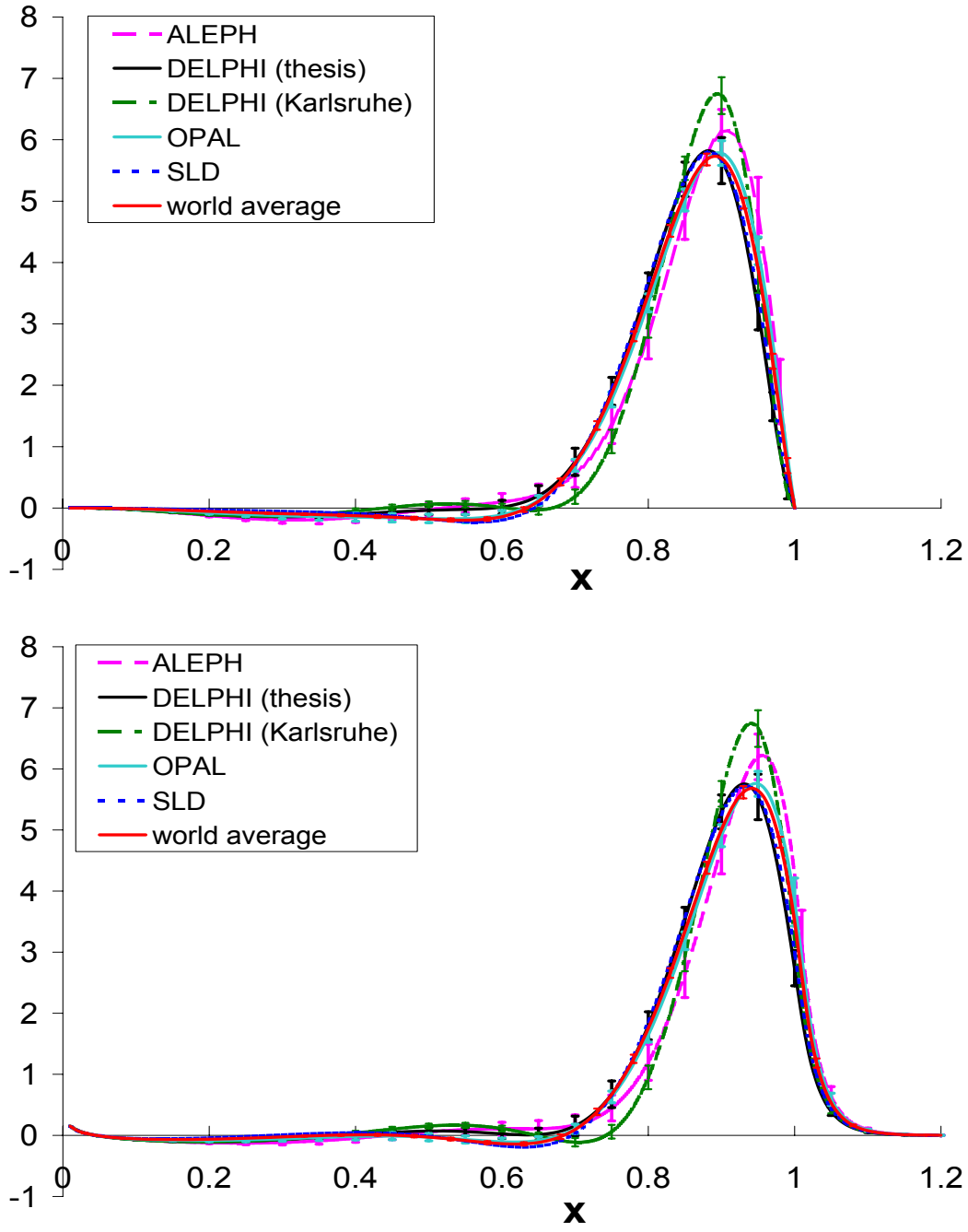

Figure 6.14: *Comparison of the extracted non-perturbative QCD component of the b quark fragmentation function for ALEPH [31], DELPHI (the result presented in this thesis), DELPHI- Karlsruhe [34], OPAL [32], SLD [33] and the combined fragmentation distribution. Top: the perturbative QCD component has been taken from JETSET 7.3. Bottom: the perturbative QCD component has been taken from NLL QCD [14].*

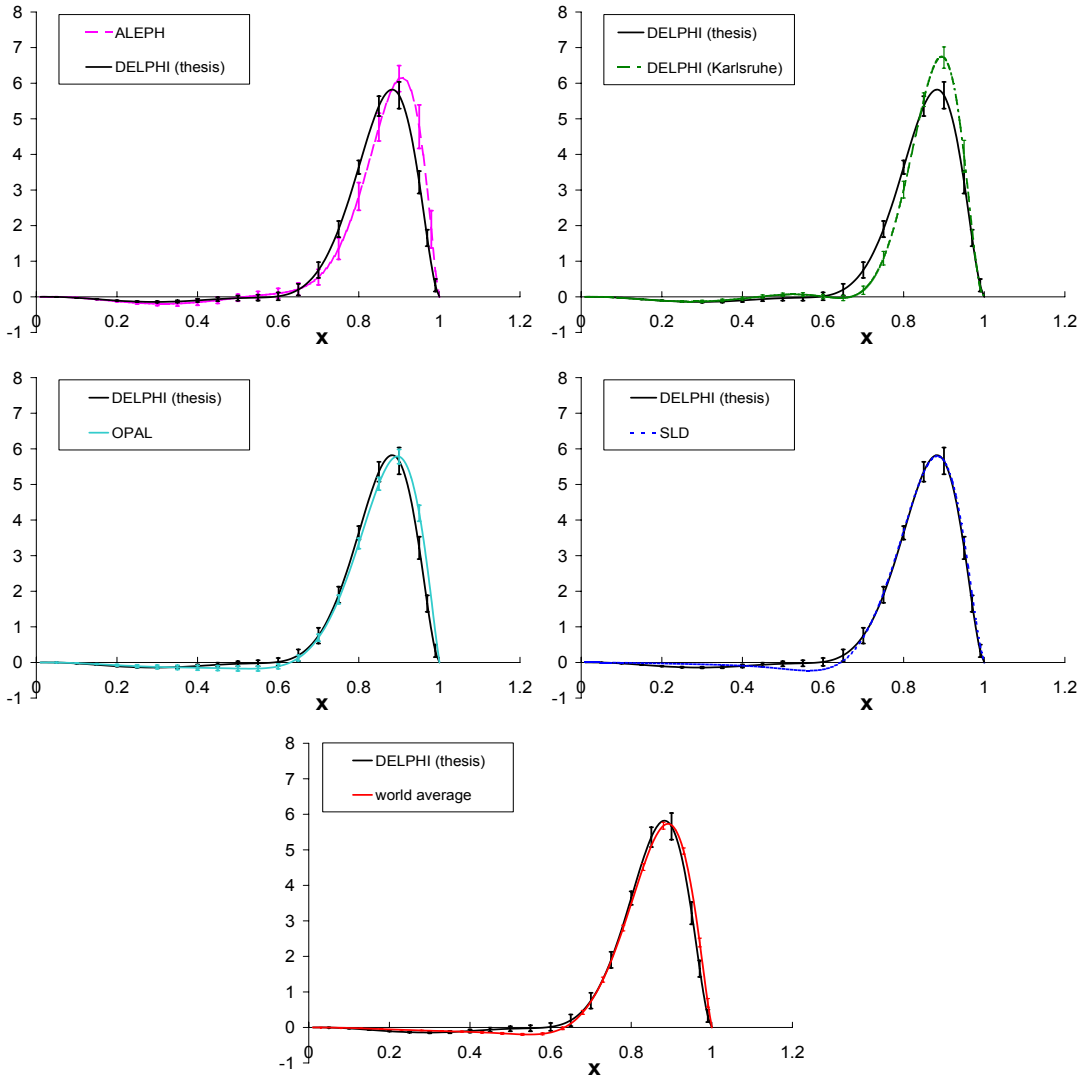

Figure 6.15: *The extracted non-perturbative QCD component of the b quark fragmentation function from the present thesis work, compared separately to the results from ALEPH [31], DELPHI- Karlsruhe [34], OPAL [32] SLD [33] and the combined fragmentation distribution. The perturbative QCD component has been taken from JETSET 7.3.*

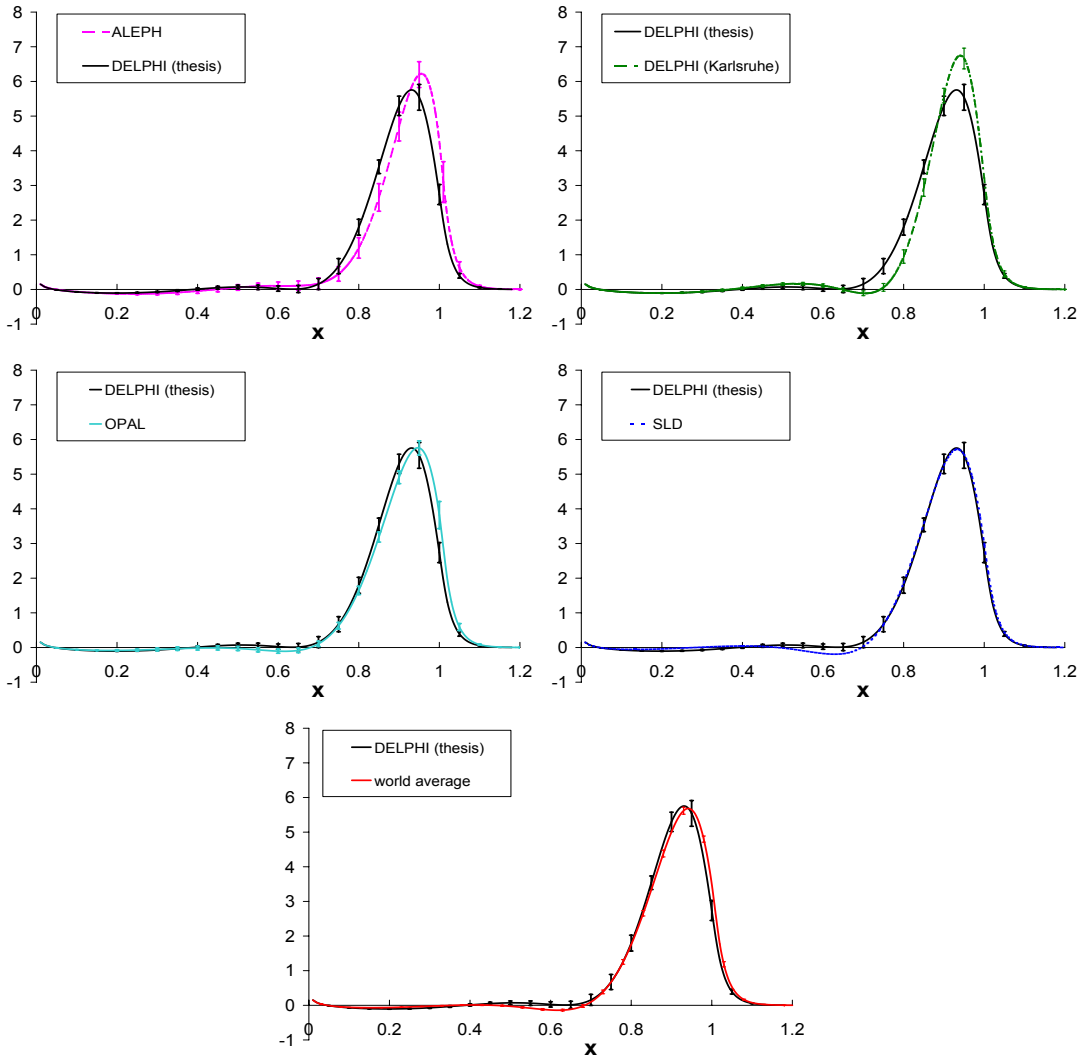

Figure 6.16: *The extracted non-perturbative QCD component of the b quark fragmentation function from the present thesis work, compared separately to the results from ALEPH [31], DELPHI- Karlsruhe [34], OPAL [32] SLD [33] and the combined fragmentation distribution. The perturbative QCD component has been taken from NLL QCD [14].*

N	moment	error		N	moment	error
				21	.0355	.0008
2	.7140	.0016		22	.0322	.0007
3	.5391	.0020		23	.0292	.0007
4	.4213	.0020		24	.0267	.0007
5	.3372	.0019		25	.0244	.0006
6	.2749	.0018		26	.0223	.0006
7	.2273	.0017		27	.0205	.0006
8	.1903	.0016		28	.0189	.0005
9	.1609	.0015		29	.0174	.0005
10	.1373	.0014		30	.0161	.0005
11	.1181	.0013		31	.0149	.0005
12	.1023	.0012		32	.0138	.0005
13	.0892	.0012		33	.0129	.0004
14	.0783	.0011		34	.0120	.0004
15	.0690	.0010		35	.0112	.0004
16	.0611	.0010		36	.0104	.0004
17	.0544	.0009		37	.0097	.0004
18	.0486	.0009		38	.0091	.0004
19	.0436	.0008		39	.0086	.0004
20	.0393	.0008		40	.0080	.0003

Table 6.5: *Moments of the combined fragmentation function for x_E. $N = 2$ is corresponding to the mean value of the distribution. The moments have been calculated using the analytic expression of Equation* (6.7)*. Errors have been propagated using the error matrix of Table 6.4*

For example, the correlation between the moments of order 2 and 3 of ALEPH's fragmentation function [31] is $\simeq 0.998$.

Fitting such highly correlated information leads to numerical problems due to singular error matrices. Correlations contain information about the measured distribution, and ignoring them is dangerous and might lead to wrong conclusions. An example to illustrate this point, is the fact that if one neglects the correlations between moments, the Peterson model may give a reasonable fit with $\chi^2/d.o.f \sim 1$ to the first moments of the non perturbative component of the fragmentation function; on the other hand, if correlations are taken into account the fit indicates that this model is excluded. This result is consistent with the observation presented in Section 6.4.1 of the present chapter.

Each moment contains a weighted information from the measured bins. Typically, a high order moment reflects mainly the information from large x. This is illustrated in Figure 6.17 which presents the contribution from each measured bin of the fragmentation function to the moment of order N as a function of N. It is clear from that figure that the information measured in the x-space is distributed among moments (not necessarily low order ones).

Moments of fragmentation functions are important for phenomenological studies given that the theoretical QCD computations are originally performed in the N space. The problem of fitting a parametrization to these moments is complicated and should be handled with care.

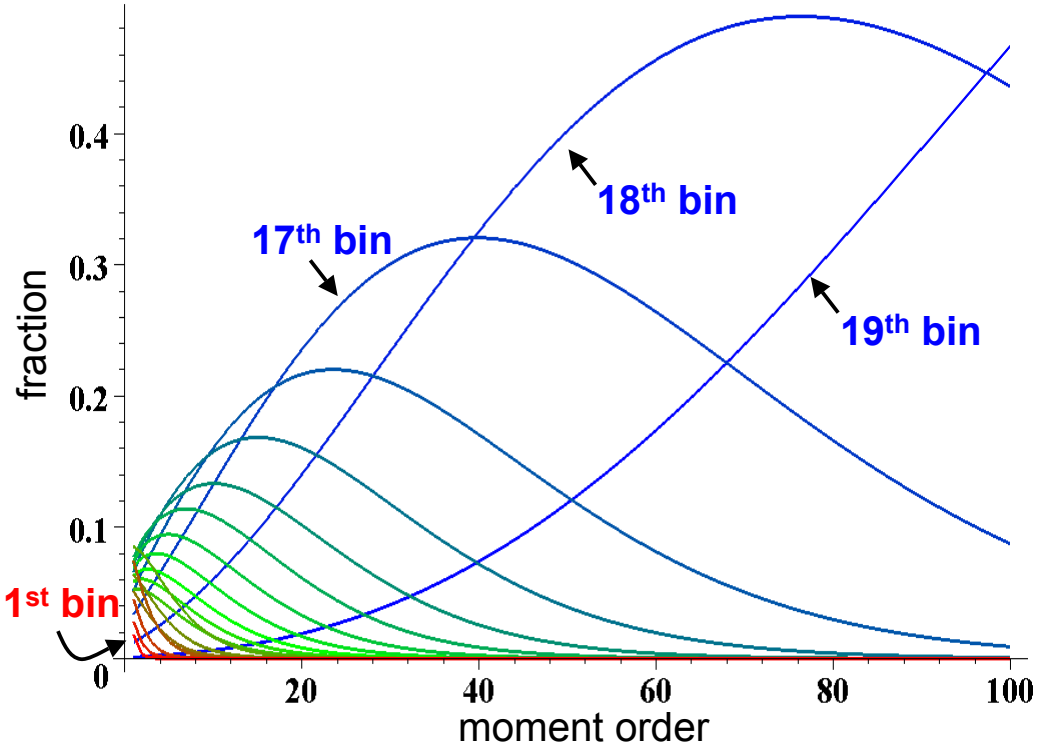

Figure 6.17: *The fraction of the moment of N'th order contributed by the each bin of the measured fragmentation function . This illustration is based on the result from ALEPH [31]. The red curves on the small N side of the figure correspond to the smallest x bins, and the blue curves that mainly contribute to moments of large order N correspond the largest x bins.*

Fitting moments of a distribution or fitting the distribution itself are not necessarily equivalent and may sometimes bring to different conclusions. The prescription for fitting highly correlated data with a singular correlation matrix presented in Appendix B may be appropriate to fits of moments.

6.9 Charm Fragmentation

The charm quark, that has a pole mass of ~ 1.5GeV, may also be considered as a heavy quark. The NLL QCD computation of [14] may be applied to charm fragmentation, using the appropriate mass. A measurement of the c quark fragmentation function has been obtained recently by the CLEO collaboration [86], for D mesons in e^+e^- annihilation at 10.52GeV. The extraction of the non-perturbative QCD component may therefore be tried for the charm case as well.

The measured spectrum from [86] of the scaled momentum (x_p), for D^0 mesons, is shown in Figure 6.18. The histogram is normalized to unity. In the same figure, is presented the parametrization of Equation (6.6), fitted to the histogram. The extraction was done, applying the procedure explained in the present Chapter, using the NLL QCD computation from [14],

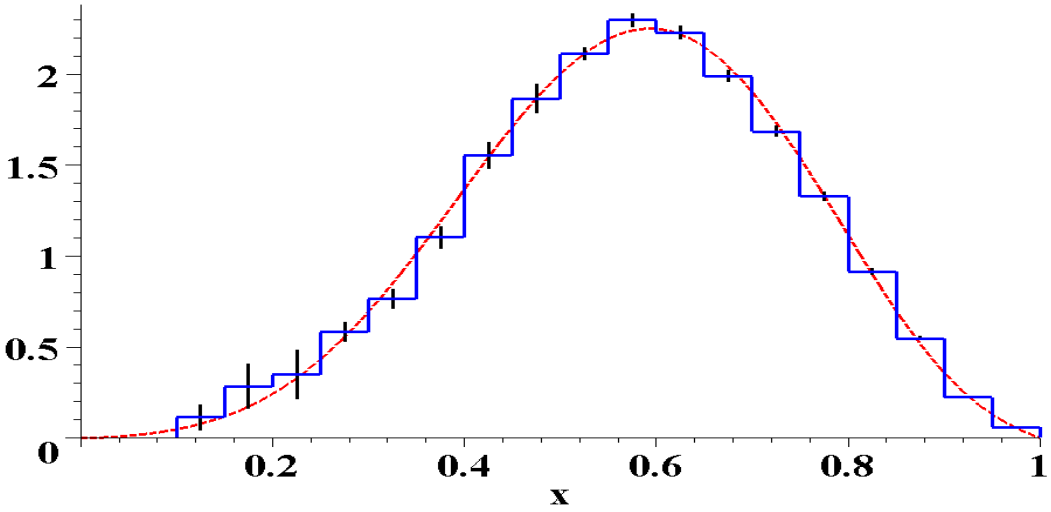

Figure 6.18: *Normalized spectrum of the scaled momentum, x_p of D^0 mesons in electron-positron annihilation as measured by CLEO [86]. The dashed curve is the parametrization of Equation* (6.6)*, fit to the histogram.*

with a quark pole mass of 1.5GeV and a center of mass energy of 10.52GeV. The perturbative and the non-perturbative components are shown together in Figure 6.19. The non-perturbative component alone is shown in Figure 6.20.

The perturbative components in the two cases considered here: bottom quark at LEP and charm quark at CLEO, are not directly comparable, because of their dependence on the center of mass energy. The non-perturbative QCD components, on the other hand, are supposed to be independent of the initial state, and may be compared. The fact that the non-perturbative component for the charm is softer than the one for bottom is the expected behavior, considering the masses of the b and c quarks. In the case of the b quark, the non-perturbative component is found to be 0 for $x < 0.6$, indicating that gluon radiation is well accounted for, in that region, by the perturbative QCD component. This is not valid for the charm quark, for which the non-perturbative QCD component does not vanish even at small x values. The non-perturbative QCD component of the c quark fragmentation has to be extended beyond $x_p = 1$, when used with the NLL QCD computation, similarly to the case of the b.

In [86], a fit of the non-perturbative component of the fragmentation function has been performed in the framework of the JETSET event generator. The fit was done to reproduce the measured D^0 meson spectrum by the Monte Carlo, and adopted the Lund model. The parameters of the Lund Symmetric Fragmentation function were found to be $a = 0.178 \pm 0.007$ and $b = 0.393 \pm 0.006$. The function is shown in Figure 6.20, taking the transverse mass as the D^0 meson mass. This fitted function, as the non-perturbtive component corresponding to the NLL QCD computation, contributes at low x values and confirms the conclusion about the insufficiency of the perturbative QCD component in that region. The full extraction procedure proposed in the present chapter has not been applied, taking the perturbative component form JETSET. Performing this extraction, and comparing the results with the fitted non-perturbative component from [86], requires the use of the same parameters of the event generator as used for the fit.

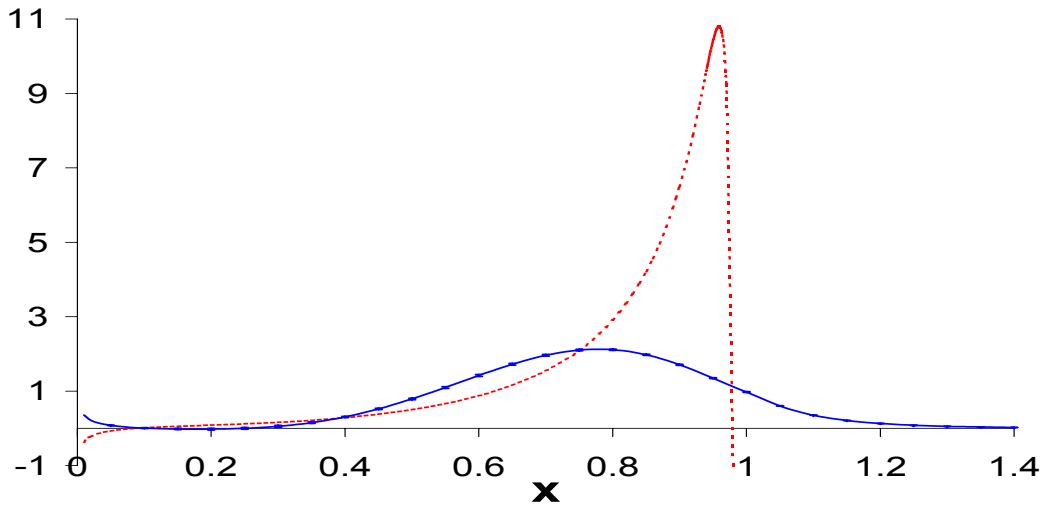

Figure 6.19: *x-dependence of the perturbative (red dotted line) and non-perturbative (full blue line) QCD components of the measured charm quark fragmentation distribution. These curves are obtained by interpolating corresponding values determined at a large number of points in the x-variable. The perturbative QCD component is given by the analytic computation of [14].*

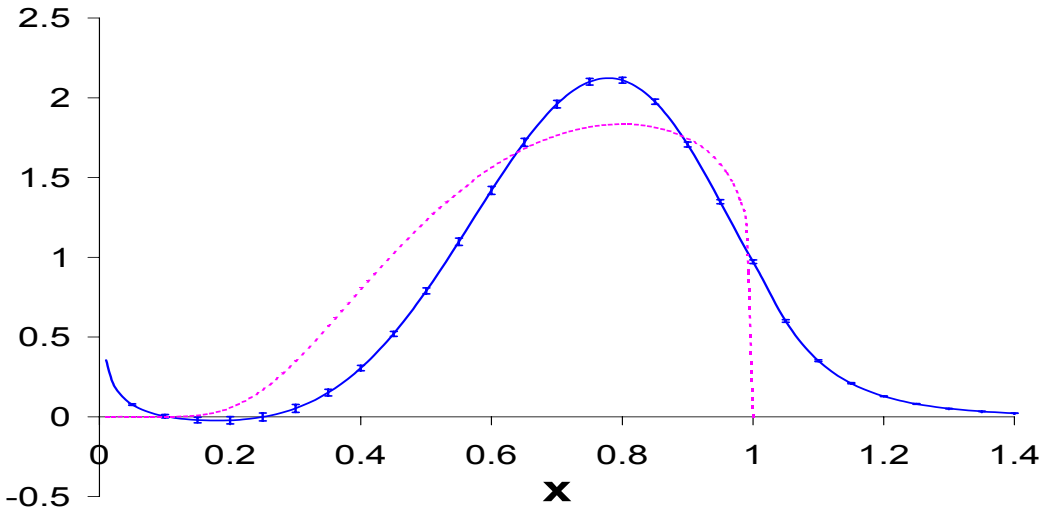

Figure 6.20: *x-dependence of the non-perturbative QCD components of the measured charm quark fragmentation distribution. The solid blue curve has been obtained by the extraction procedure, using the perturbative QCD component given by the analytic computation of [14]. The red dotted curve described the Lund Fragmentation function, with the parameters fitted in [86] to reproduce the D^0 spectrum: $a = 0.178$ and $b = 0.393$. The transverse mass in the Lund model was assumed here to be the mass of the D^0 meson.*

The non-perturbative component of the charm fragmentation function, as in the case of the bottom, is an ingredient for the theoretical prediction of the charm cross section at hadronic colliders. This cross section is now experimentally measurable at CDF, thanks to the silicon vertex trigger. CDF has published its first results, using $6pb^{-1}$ of data [87].

6.10 Conclusions

The measured *b*-quark fragmentation distribution has been analyzed in terms of its perturbative and non-perturbative QCD components.

The *x*-dependence of the fragmentation distribution has been extracted in a way which is independent of any model for non-perturbative hadronic physics. It depends closely on the way the perturbative QCD component has been evaluated. The obtained distribution differs markedly from those expected from various models.

Below $x = 0.6$, this distribution is compatible with zero indicating that most of hard gluon radiation is well accounted by the perturbative QCD component evaluated using the LUND parton shower Monte-Carlo or computed analytically.

As the non-perturbative QCD distribution is evaluated for any given value of the *x*-variable it can be verified if it remains physical over the interval $[0,1]$ when used with a Monte-Carlo generator which provides the perturbative component. The evidence for unphysical regions would indicate that the simulation or the measurements are incorrect. There is not such an evidence in the present analysis.

Above $x = 0.6$, the obtained distribution is similar in shape with those expected from the Lund symmetric [38] or Bowler [24] models, when the perturbative QCD component is taken from JETSET. When the perturbative QCD component is taken from the analytic result of [14] it has been found that, because of the analytic behaviour of the perturbative QCD component, the non-perturbative QCD distribution must be extended above $x = 1$. The *x*-behaviour of the non-perturbative component, for $x > 1$, is determined by the possible existence of a zero in $\tilde{\mathcal{D}}_{pert.}(N)$, for $N > 0$. When the perturbative component has non-physical aspects, it is thus not justified to fold it with any given physical model. An approach has been proposed to solve this problem and a parametrization of the obtained distribution has been provided.

The non-perturbative component, extracted in this way, is expected to be valid in a different environment than e^+e^- annihilation, as long as the perturbative QCD part is evaluated within the same framework (analytic QCD computation or a given Monte Carlo generator), and using the same values for the parameters entering into this evaluation as m_b^{pole}, $\Lambda_{QCD}^{(5)}$ or generator tuned quantities. In the next chapter a fragmentation analysis in the CDF experiment is presented, where it is proposed to check the validity of the fragmentation function from LEP in the framework of the PYTHIA Monte Carlo generator.

Chapter 7

B Fragmentation and Related Studies at CDF

7.1 Introduction

As explained in Section 2.2, in an e^+e^- collider at the Z^0 pole b quarks are produced from e^+e^- annihilation to a Z^0 which then decays to a $b\bar{b}$ pair. The hard process is well defined and the energies of its initial and final states are known. Under these conditions, the perturbative QCD component of the b fragmentation function is well understood. On the other hand, in a hadronic collider there are several diagrams that contribute significantly to b quark production and the center of mass energy of the hard process is not determined. Perturbative QCD processes are thus more complex in the second experimental environment.

The non-perturbative QCD component of the fragmentation process, as explained in Section 2.3.2 and in Chapter 6, may be obtained from studies in e^+e^- collider experiments [31, 34, 32, 33, 88]. This part of the fragmentation should not depend on the initial state and therefore the distribution of the non-perturbative QCD component obtained by other experiments may be, at least as a first approximation, implemented in CDF studies. The non-perturbative QCD component depends strongly on the perturbative QCD approach that has been used for its extraction or fit. In the present analysis the PYTHIA event generator [19] is employed, and therefore the non-perturbative component obtained from e^+e^- experiments that used PYTHIA may be implemented. In [32, 33, 88], the Lund Symmetric Fragmentation Function [38] and the Bowler model [24] are favored. Part of the present study uses the Lund model , with the parameters obtained in Chapter 6 . The non-perturbative component is implemented in PYTHIA as the "fragmentation function", and is set by certain parameters and switches of the generator.

The observables related to the B meson and tracks in its vicinity are influenced not only by the perturbative and non-perturbative components of the fragmentation but also by the mechanism through which the b quarks were produced in the hard process. The model predictions depend on the interplay between these three factors. The separation of fragmentation effects from production mechanisms is a non-trivial issue. However, this separation may be done by studying the influence of the production mechanisms on observables that are not sensitive to the fragmentation effects, such as the distributions of tracks whose directions are well separated from the B meson's direction.

Another issue closely related to the fragmentation is the production rate of B^{**} mesons in the b jet. These B excited states decay by strong interaction to light and B mesons at the primary vertex. A measurement of the B^{**} production rate at CDF is under way. For the present analysis, values of a recent measurement from DELPHI [49] have been used.

In the present analysis the exclusive channel of fully reconstructed $\mathrm{B}^{\pm} \to \mathrm{J}/\psi\mathrm{K}^{\pm}$ is used. With the current data collected by CDF, this channel provides a sample of a few thousand B mesons. The primary advantage of this channel over inclusive ones such as $B \to J/\psi X$ is that the 4-momentum of the B meson is well measured.

The aims of the present analysis are to

- study the influence of the fragmentation function and of the different b production mechanisms on the properties of the B meson and of the tracks in its vicinity. Tracks may be classified according to their origin: Hadronization (HA), which correspond to tracks in the b jet and Pile-Up events (PU).
- demonstrate the agreement between data and PYTHIA in a number of variables related to fragmentation, or to tune PYTHIA until there is agreement.
- measure the b quark production cross section using the exclusive decay channel $\mathrm{B}^{\pm} \to \mathrm{J}/\psi\mathrm{K}^{\pm}$.
- study the fragmentation function of B mesons in $p\bar{p}$ collisions at TeVatron energies in the environment of PYTHIA.

The analysis is still in progress, and therefore not all of the above goals have been achieved. In this chapter, the procedure and current state is presented.

The agreement between data and the PYTHIA Monte Carlo generator is a crucial ingredient in the measurement of the fraction of B mesons originating from $B^{**} \to B\pi$ decays. In that analysis, PYTHIA is needed to help understand the signal and background distributions of the mass spectrum of $\mathrm{B}^{\pm}\pi^{\mp}$ ("right sign" or RS) and $B^{\pm}\pi^{\pm}$ ("wrong sign" or WS) hadronization.

However, the main importance of achieving agreement between the data and PYTHIA at CDF is in confirming that PYTHIA can reproduce the kinematic behavior and correlations between B mesons and tracks produced around them. If such agreement can be unequivocally established, CDF could use PYTHIA to estimate the dilution of the Same Side Kaon Tagging (SSKT) method in B_s decays. Obtaining the dilution of SSKT is a requisite ingredient of setting the limit on the $B_s^0\bar{B}_s^0$ oscillation frequency, Δm_s.

When completed, this analysis will provide both tuning of the PYTHIA generator and an interesting measurement by itself. The final outcome will be the first measurement of the b fragmentation function in a hadronic collider.

In section 7.2 the data sample used for the analysis is described. Fiducial, trigger and analysis cuts are detailed and the background subtraction from data is explained. General discussion of the Monte Carlo samples used for this analysis is given in Section 7.3. The fragmentation analysis, still in progress, is outlined in Section 7.4. The following sections, 7.5 and 7.6, describe studies related to the fragmentation analysis. They have been done by comparing different types of Monte Carlo samples between them, and to data. A method to fit the fragmentation function parameters in PYTHIA is proposed in Section 7.7. An ongoing analysis to measure the b quark cross section is presented in section 7.8

7.2 Data Sample

The present analysis is based on events collected by the CDF Run-II detector up to August 2004. The exclusive $B^{\pm} \to J/\psi K^{\pm}$ decay channel is used to assemble a sample of ~ 6000 B meson candidates. These decays are reconstructed from the compressed di-muon trigger data stream and comprises 333 ± 20 pb^{-1} of the total offline integrated luminosity (with 1333 good runs) [1].

7.2.1 Reconstruction of $B^{\pm} \to J/\psi K^{\pm}$

The di-muon trigger requires two tracks with $p_T > 1.5$ GeV/c which match to the stubs in the muon chambers. A maximum opening angle in the transverse plane of $\Delta\phi < 135^o$ between the two tracks is also enforced at the trigger level. The muons are constrained to pass through a common point. Pairs of oppositely charged muons are then combined to form a J/ψ. At this level the invariant mass of $\mu^+\mu^-$ pair must be between 2.9 and 3.3 GeV/c^2. Additionally, each of the reconstructed tracks are required to have at least 10 axial and 10 stereo hits in the drift chamber, with $|\eta| < 2.0$; where η is the pseudo-rapidity, and at least two hits in the silicon detector.

The kaon candidates are tracks with $p_T > 1.0$ GeV/c that are consistent with the J/ψ decay point. The $\mu^+\mu^-$ invariant mass is then constrained to the J/ψ PDG value before the determination of the $\mu^+\mu^-K$ decay point. The transverse momentum of the combined system must satisfy $p_T(\mu^+\mu^-K) > 4.0$ GeV/c, and the invariant mass of the $\mu^+\mu^-K$ triplet must lie between 4.9 and 5.7 GeV/c^2.

For the $B^{\pm} \to J/\psi K^{\pm}$ decay channel, analysis cuts on the candidates have already been optimized during the Same Side Tagging studies [89]. A full optimization based on $S/\sqrt{S+B}$ was performed, and the resulting analysis cuts are listed in Table 7.1. Analysis cuts on the other tracks in the event were chosen to ensure that the track was prompt and associated with the B.

For the B meson, the quantity χ^2_{xy} is the χ^2 value for the fit of the $\mu^+\mu^-K$ decay point in the transverse (xy) plane of the CDF detector. L_{xy} is defined as the distance in the transverse plane from the point where the $p\bar{p}$ interacted to the $\mu^+\mu^-K$ decay point. Requiring a large $L_{xy}/\sigma(L_{xy})$ ensures that the reconstructed B meson has traveled out from the interaction point by a distance consistent with the B lifetime.

The impact parameter d_0 measures the distance of closest approach of a track to the interaction point if that track were projected back beyond its production point. Requiring a small $d_0/\sigma(d_0)$ ensures that the track is prompt, as it is consistent with the interaction point. R is defined as a distance in the η-ϕ plane. Requiring tracks within a cone of opening angle ΔR around the B meson includes mainly tracks associated with the hadronization of the B meson.

The $\mu^+\mu^-K$ invariant mass is shown in Figure 7.1. The flat sidebands to the left and to the right of this peak can be used as samples of pure combinatorial background, which may then be subtracted from the various distributions produced with the data in the signal region.

Figure 7.2 shows the breakdown of the $B^+ \to J/\psi K^+$ signal, from a full simulation of the inclusive b-hadron decays into $J/\psi X$. The invariant mass distribution in Figure 7.2 describes the physics backgrounds for $B^+ \to J/\psi K^+$ and illustrates why the left sideband can only reach

[1] In the future, other B decay channels will be added in order to increase the sample size

Candidate	Cut Value	Units
$\mu^{\pm}$	$p_T > 1.5$	GeV/c
J/ψ	$\lvert m(J/\psi) - 3096.88\rvert < 80$	MeV/c^2
K	$p_T > 1.0$	GeV/c
B	$p_T > 4.0$ $\chi^2_{xy} < 15.0$ $L_{xy}/\sigma(L_{xy}) > 4.5$	GeV/c
track	$d_0/\sigma(d_0) < 3.5$ $\Delta R(B, track) < 0.7$ $\lvert\Delta z(B, track)\rvert < 5.0$ $p_T > 0.4$	 cm GeV/c

Table 7.1: *Analysis cuts for the decay $B^{\pm} \to J/\psi K^{\pm}$ ($J/\psi \to \mu^+\mu^-$) along with requirements made on tracks in the vicinity of the B.*

down to $m_{J/\psi K^+} = 5.17$ GeV/c^2. The region below 5.17 GeV/c^2 also includes physics decays not present under the signal peak. In addition, the Cabibbo-suppressed decay $B^+ \to J/\psi\pi^+$ shows up as a shoulder on the right side of the peak, contributing both to the peak and to the right sideband. It corresponds to $\sim 4\%$ of the $B^+ \to J/\psi K^+$ sample, as predicted from the ratio of branching ratios for these two decay modes [71] [2]. This contribution is not taken into account in the sideband subtraction. Therefore, in the fit shown in Figure 7.1, the $B^+ \to J/\psi K^+$ signal is modeled with two Gaussians of different widths with the means constrained to be the same. The $B^+ \to J/\psi\pi^+$ component is offset from $B^+ \to J/\psi K^+$, and its size is fixed to 4% of the area of the two signal Gaussians. The combinatorial background is modeled with a linear function. A sum of these probability density functions is fit to data in the region between 5.17 and 5.66 GeV. The result of this fit is given in Table 7.2.

Fit parameter	Value	Error
Mean(signal)	5.2792	0.0003
σ(core)	0.0101	0.0011
Norm(core + tail)	60.1	1.8
σ(tail)	0.020	0.004
Mean($B^+ \to J/\psi\pi^+$)	5.33	0 (fixed)
Norm($B^+ \to J/\psi\pi^+$)	4% of $B^+ \to J/\psi K^+$	0 (fixed)
σ($B^+ \to J/\psi\pi^+$)	0.072	0.058
Comb. bkg. constant	2735	205
Comb. bkg. slope	-336	37

Table 7.2: *The result of the fit shown in Figure 7.1.*

[2] The quoted branching ratios $BR(B^+ \to J/\psi\pi^+)$ and $BR(B^+ \to J/\psi K^+)$ in [71] are $(4.0 \pm 0.5) \cdot 10^{-5}$ and $(1.00 \pm 0.04) \cdot 10^{-3}$ respectively

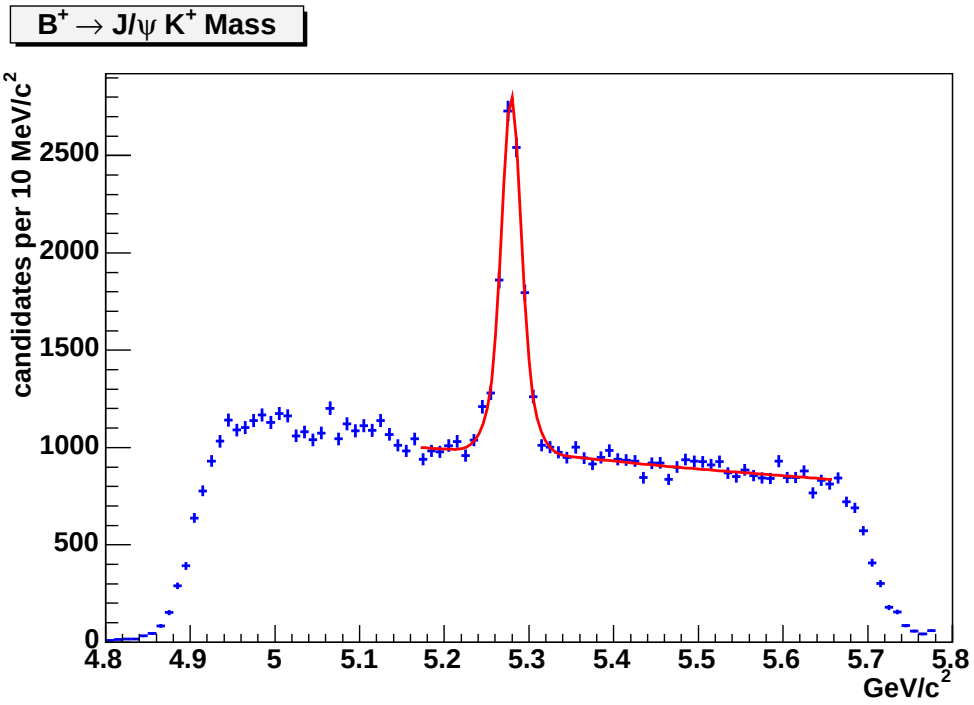

Figure 7.1: *The invariant mass of $B^{\pm} \to J/\psi K^{\pm}$, with the fit overlaid. The signal consists of the two signal Gaussians for $B^{\pm} \to J/\psi K^{\pm}$, and another one for the Cabibbo-suppressed decay mode $B^{\pm} \to J/\psi \pi^{\pm}$, which is fixed to 4% of $J/\psi K^{\pm}$. The combinatorial background is modeled as a linear function.*

7.2.2 Subtracting the Backgrounds in the Data

Two kinds of distributions have to be considered when subtracting background contributions: those relative to the B meson, and those relative to the accompanying tracks. In distributions corresponding to B meson production, characteristic contributions from background events, such as the $p_T(B)$ distribution, have been obtained using sidebands of the B mass distribution. Signal events $P_T(B)$ distributions are produced by subtracting, after proper normalization, the $p_T(B)$ distribution corresponding to sideband events from the one obtained by selecting events in the signal region. When performing the sideband subtraction, the signal region is defined as $m(J/\psi \mathrm{K}^+) \in [5.2391, 5.31935]$, which corresponds to $\mathrm{Mean(signal)} \pm 2\sigma(tail)$. The low sideband is defined as $m(J/\psi \mathrm{K}^+) \in [5.17, 5.21904]$, where the high boundary is taken as $\mathrm{Mean(signal)} - 3\sigma(tail)$. The high sideband is defined as $m(J/\psi \mathrm{K}^+) \in [5.33941, 5.66]$, where the lower edge is taken as $\mathrm{Mean}(\mathrm{B}^+ \to \mathrm{J}/\psi\pi^+) + \sigma(\mathrm{B}^+ \to \mathrm{J}/\psi\pi^+)$.

However, the distributions for tracks suffer from an additional background, which comes from tracks originating in other $p\bar{p}$ interactions. At the instantaneous luminosities which are typical of the analyzed sample, about one event in four has more than one primary interaction. Such additional primary interactions are defined as *'Pile-Up' events (PU)*. Fortunately, at the TeVatron the luminous detector region is rather elongated; the beam spot has a $\sigma_z \sim 27$ cm. Since the Z_0 resolution of the CDF silicon tracker is a few millimeters at worst (without hits in the layers of the silicon detector with 90^o strips), it is relatively easy to remove the majority of the PU tracks by applying a cut on the absolute value of the distance between the Z coor-

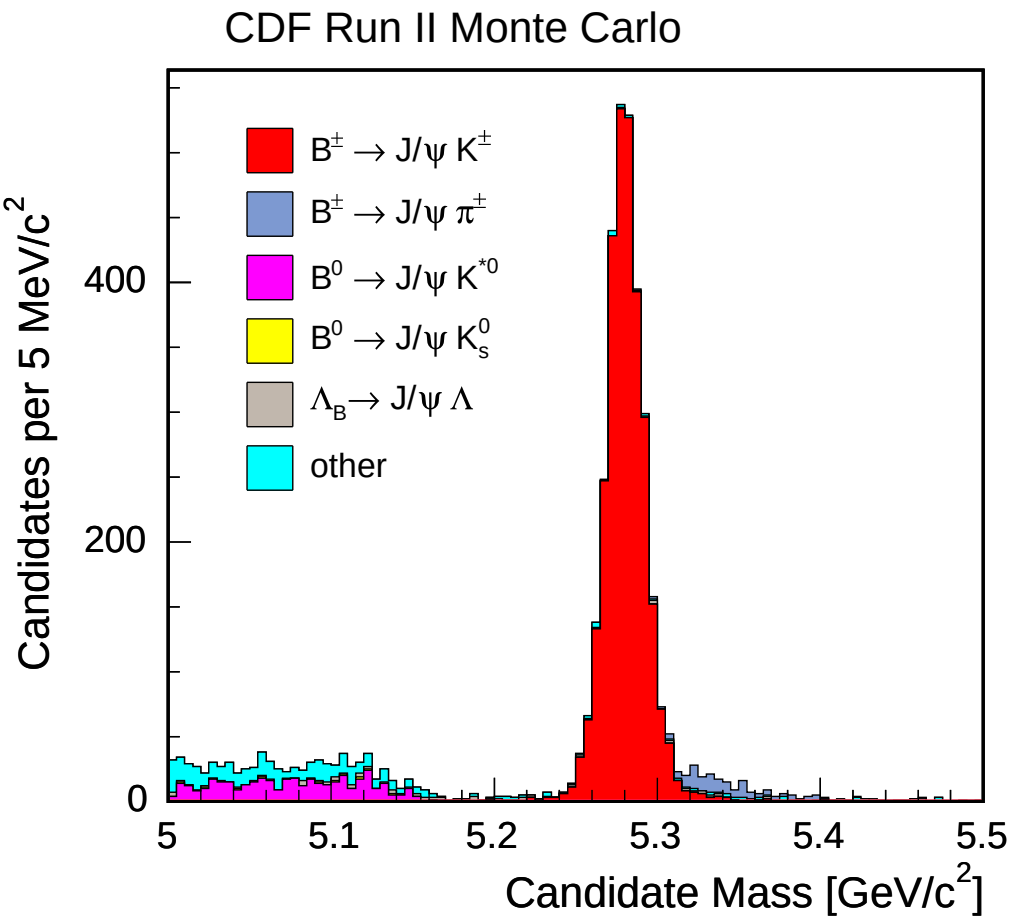

Figure 7.2: *Different contributions to the* $\mathrm{B}^+ \rightarrow \mathrm{J}/\psi\mathrm{K}^+$ *signal, from an inclusive* $H_b \rightarrow J/\psi X$ *Monte Carlo simulation (where* H_b *is a b-flavored hadron).*

dinate of the B^+ vertex and the Z_0 of the surrounding tracks, $\Delta z \equiv Z(\mathrm{B}^+) - \mathrm{Z}_0(\mathrm{track})$. In the present analysis this cut is set at $|\Delta z| < 5$ cm. An example of the distribution of Δz is shown in Figure 7.3.

However, despite the $|\Delta z| < 5$ cm cut, some amount of PU tracks 'leak' into the signal region. That contribution must be subtracted, and the "Δz" sidebands are used as a representative sample of pure background. The sideband subtraction proceeds as usual, with the sidebands defined as $(-60,-20)$ and $(20,60)$ centimeters from the signal peak at $\Delta z = 0$.

Figure 7.4 shows the effect of removing the contribution of PU events by subtracting the Δz sidebands. This effect is not large, since the ratio of the primary interaction signal and the Pile-Up background in the signal region is $\sim 31:1$.

The process of removing the Δz sidebands becomes somewhat convoluted as it involves the $m(J/\psi\mathrm{K}^+)$ sidebands as well. The regions of pure background (referred to as 'sidebands') and the regions with enhanced signal ('signal regions') must be clearly defined. The tracks in this analysis come from four sources:

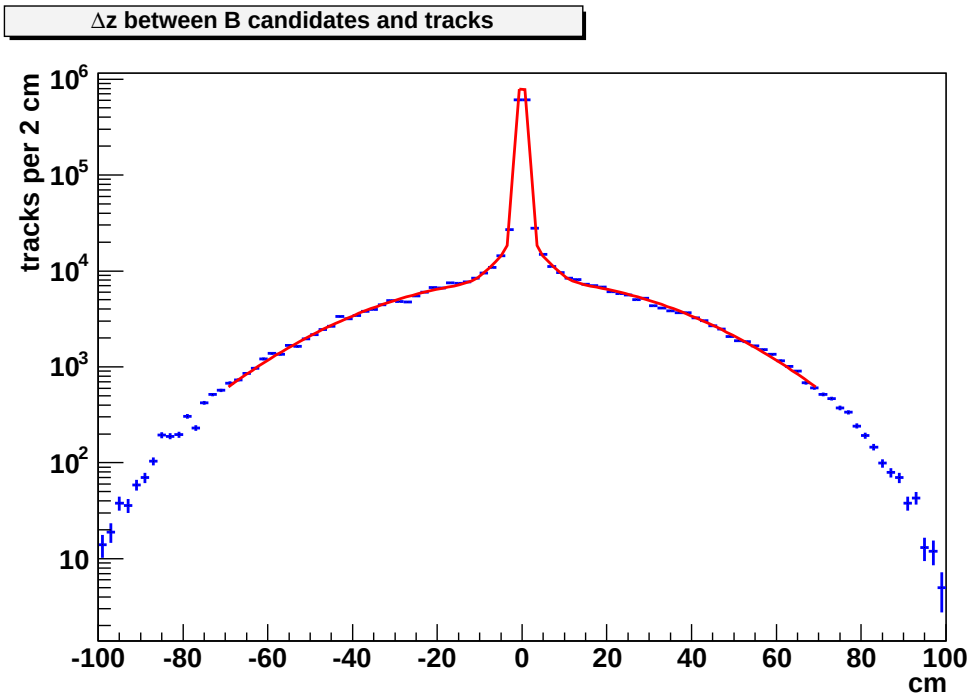

Figure 7.3: *The distribution of* $\Delta z \equiv Z(B^+) - Z_0(\text{track})$ *where* $Z(B^+)$ *is the Z coordinate of the* B^+ *meson's decay point, and* $Z_0(\text{track})$ *is the Z coordinate of the point of the closest approach. The* Δz *consists of two narrow Gaussians to describe the signal and a broad Gaussian to describe the contribution of tracks from Pile-Up events.*

- **Type I:** Pure signal. Emitted tracks accompanying the B^+ meson in the event.
- **Type II:** B^+ signal but Δz sidebands. These are tracks from Pile-Up interactions from the events when a real B^+ meson was produced.
- **Type III:** Δz signal but B^+ sidebands. These are tracks produced along with fake B^+ mesons (*i.e.*, combinatorial background). These are likely hadronization or decay tracks from true *b*-hadron, but not $B^+ \to J/\psi K^+$.
- **Type IV:** Sidebands in both B^+ mass and Δz. These are PU events in combinatorial background events.

For simplicity, let us confine the discussion to the yields in the four relevant regions of the two-dimensional ($m(J/\psi K^+)$,Δz) space: (SR_m,SR_z), (SR_m,SB_z), (SB_m,SR_z), (SB_m,SB_z), where SR stands for 'Signal Region' and SB for 'Sideband', as illustrated by Figure 7.5. An identical argument can be made for each bin of any distribution, and thus the prescription below is valid for an arbitrary distribution, including those shown below.

- (SB_m,SB_z) contains only Type IV tracks; $N_{(SB_m,SB_z)} = N_{IV}$
- (SB_m,SR_z) contains Type III and Type IV tracks; $N_{(SB_m,SR_z)} = N_{III} + N_{IV}$
- (SR_m,SB_z) contains Type II and Type IV tracks; $N_{(SR_m,SB_z)} = N_{II} + N_{IV}$

Fit parameter	Value ± Error
Mean narrow	$(-12.5 \pm 8.9) \times 10^{-4}$
Norm 1	$(978.2 \pm 2.4) \times 10^{+3}$
σ 1	$(994.2 \pm 2.6) \times 10^{-3}$
Norm 2	$(111.3 \pm 2.7) \times 10^{+2}$
σ 2	$(479.5 \pm 7.3) \times 10^{-2}$
Mean wide	0.0 (fixed)
Norm wide	$(796.4 \pm 3.0) \times 10^{+1}$
σ wide	$(3062.4 \pm 6.5) \times 10^{-2}$

Table 7.3: *The result of the fit shown in Figure 7.3.*

- (SR_m,SR_z) contains all four types of tracks; $N_{(SR_m,SR_z)} = N_I + N_{II} + N_{III} + N_{IV}$

There are two equivalent ways to proceed:

1. Use all four regions simultaneously. Naively, one would simply subtract (SB_m,SR_z) and (SR_m,SB_z) from (SR_m,SR_z) to obtain the pure signal, but that would in fact produce $(N_I + N_{II} + N_{III} + N_{IV}) - (N_{III} + N_{IV}) - (N_{II} + N_{IV}) = N_I - N_{IV}$. Therefore, it is necessary to add $N_{IV} = N_{(SB_m,SB_z)}$ to get the true signal:

$$N_I = N_{(SR_m,SR_z)} - N_{(SB_m,SR_z)} - N_{(SR_m,SB_z)} + N_{(SB_m,SB_z)} \tag{7.1}$$

2. Subtract one type of background first. This is equivalent to rearranging Equation 7.1 into two groups:

$$N_I = (N_{(SR_m,SR_z)} - N_{(SR_m,SB_z)}) - (N_{(SB_m,SR_z)} - N_{(SB_m,SB_z)})$$

   That is, first subtract the Δz background from the mass signal region and the mass sidebands, and then subtract the Δz-subtracted mass sidebands from the Δz-subtracted signal region (or vice versa).

The plots in this chapter are produced using the first method, while the second is used as a cross-check (as the results should clearly be identical).

Charged tracks from PU events correspond to about 4% of the candidate tracks accompanying the *b*-hadron, as shown in Figure 7.4.

7.3 Monte Carlo Samples

7.3.1 General Description

The PYTHIA Monte Carlo samples are generated in the CDF analysis framework, and involve the successive use of the following steps (performed by different executables):

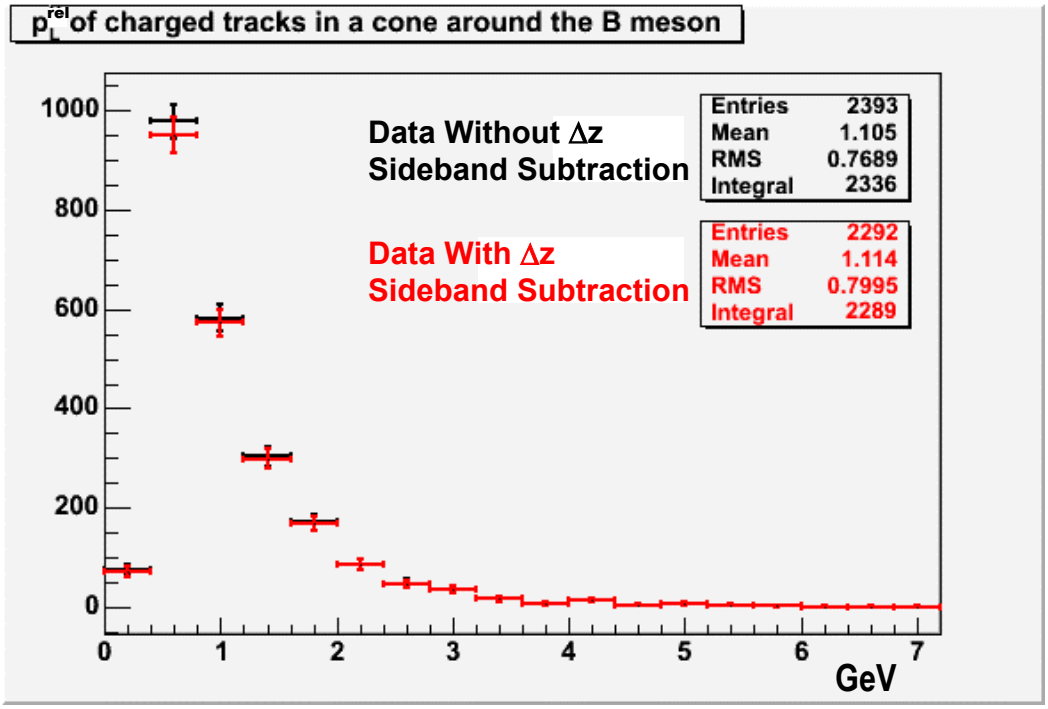

Figure 7.4: *The effect of the removal of the pile-up events by subtracting the Δz sidebands on p_L^{rel}, the component of the track momentum parallel to the B meson.*

- **Generation phase (cdfGen):** This phase starts with PYTHIA, which creates an event. Two types of PYTHIA generation are considered, one in which only a $b\bar{b}$ pair is created from lowest order QCD diagrams, and another in which all processes are used and $b\bar{b}$ are created as only one of several possibilities. Both of these options are described below, in Section 7.3.2.

  Also in the generation phase, the **EvtGen** program runs to decay the B mesons. In this analysis the decays $B^+ \to J/\psi K^+$ and $J/\psi \to \mu^+\mu^-$ are forced, when a B^+ meson has been generated (B^- decays are not forced).

- **Detector simulation (cdfSim):** This phase runs a standard simulation of the CDF detector (from the release 5.3.3). The CDF detector simulation operates at the level of hits for all detector components except the calorimetry, where the shower evolution is computationally prohibitive. However, the tracking, especially the hits in the Silicon Detector, are simulated at a very detailed level, and involve the strip-to-strip variations in performance as well as the generation of random noise throughout the detector. The output of the cdfSim looks like the output from the CDF data acquisition system.

- **Trigger simulation (TrigSim++):** The detector-like information is then fed into a detailed simulation of the trigger system. Since the $B^+ \to J/\psi K^+$ channel is reconstructed in the J/ψ data stream, only the di-muon triggers are included in the TrigSim++ for these

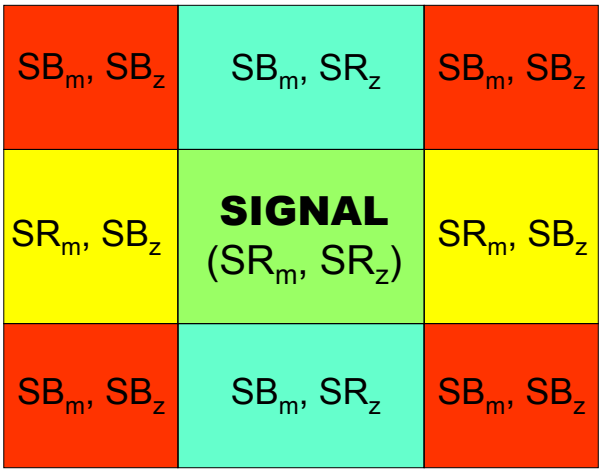

Figure 7.5: *Regions of signal and sidebands for $m(J/\psi K^+)$ and Deltaz*

samples. The TrigSim++ runs as a filter, and transmits only those Monte Carlo events which would pass the real trigger system.

- **Event reconstruction (ProductionExe):** Next the simulated events which pass the di-muon triggers are processed with the standard CDF production executable. At this stage, the hits in the muon chambers (CMU, CMP, and CMX) are reconstructed and linked into muon stubs. The hits in the Central Outer Tracker (COT) are reconstructed and linked into COT tracks. The COT tracks are then extrapolated and matched with the muon stubs. The other tracks are also extrapolated into the Silicon Detector where the Silicon hits are attached to these tracks.

- $B^+ \rightarrow J/\psi K^+$ **analysis (UFind):** Finally, the decay is reconstructed using exactly the same executable used to reconstruct the $B^+ \rightarrow J/\psi K^+$ decay mode in data. The combinatorial reconstruction algorithm is described in Section 7.2.

7.3.2 PYTHIA Parameters

Two types of PYTHIA event samples are used in this analysis, differing in the choice of the "msel" parameter of the generator. This parameter determines which hard scattering processes are generated [3]. With msel=5, only those with outgoing b and $\bar{b}$ quarks are created. These kinds of processes are referred to as 'Flavor Creation (FC)'. The choice msel=1 generates instead a generic hard scattering that results in light jets; however, this process occasionally results in the creation of $b\bar{b}$ pair, via any of the possible mechanisms ('Flavor Creation', 'Flavor Excitation (FE)', or 'Gluon Splitting (GS)'). These production mechanisms are explained in detail in Section 2.2.2 . In both cases an event with a B meson is required at the end of PYTHIA, and therefore most of the events generated with msel=1 are discarded. For this reason, msel=1 takes significantly more time to be generated.

[3] The hard scattering process in PYTHIA is defined as the highest energy primary interaction of two partons in the colliding beam particles.

In the present analysis the msel=1 sample has been used as a model of the complete contribution from all the production mechanisms of the *b* quark. In future, the contribution from FC alone will be replaced by a sample generated with msel=5. The reason is that when msel is set to 1, PYTHIA uses a massless quark approximation to calculate the matrix elements. Thus, *b* quarks are *also* treated as massless in the hard scattering process. In the final state, the *b* quark 3-momentum is rescaled along its direction to make the kinematics consistent with a massive *b* quark; the energy is left unchanged. On the other hand, when msel is set to 5 PYTHIA uses massive matrix elements to generate the two *b* quarks outgoing from the hard scattering process.

In an msel=1 sample a cut-off on the p_T of the hard scattering must be imposed. This is achieved by setting the variable CKIN(3). The usual setting in CDF is CKIN(3)=5 GeV. Setting it to 0 in an msel=5 sample is safe, because the *b* quark is taken as massive and the mass regulates the cross section. The importance of setting CKIN(3) to 0 for msel=5 PYTHIA can be seen comparing Figures 7.6 and 7.14. In both figures, comparisons are done between the p_T spectrum of the *B* meson between data and msel=5 Monte Carlo. In the former, the Monte Carlo sample has been generated with CKIN(3)=0, and in the later with CKIN(3)=5. It is clear that the rise of the p_T spectrum of the *B* meson depends on the cut-off defined by CKIN(3). This cut off has no reason to be set, and therefore in all the data and msel=5 Monte Carlo comparisons in this document, apart from Figure 7.6, CKIN(3) has been set to 0.

The PYTHIA input parameters for the samples generated for this analysis are briefly detailed below:

- a set of parameters that have been tuned in CDF for the underlying event. In the collaboration, this set is referred to as "Tune A" [4] [90];
- a B^{**} rate of 20%, as quoted by recent measurement from DELPHI [49]);
- the default values for B^{**} masses and widths have been replaced by recent measurements [71];
- samples with three different fragmentation functions (Bowler, soft Peterson and hard Peterson) for msel=5 PYTHIA setting. These three functions are shown in Figure 7.7 and described in more detail in Section 7.4;
- the default Bowler fragmentation function has been used in the msel=1 sample;
- CKIN(3)=0 in the msel=5 samples.

7.4 Outline of the Analysis Method

The first step of the analysis required defining observables related to the *B* meson and the tracks in its vicinity which were sensitive to the fragmentation function. The behavior of these variables for three primary fragmentation functions has been studied in three different Monte

[4] The underlying event is defined as all the activity in the event that arises from the hadronization of the debris of the $p\bar{p}$ interaction, *i.e.* of the constituent partons of the incident proton and antiproton which do not participate in the hard scattering. The Underlying Event is almost completely uncorrelated with the outgoing *b* and $\bar{b}$ quarks.

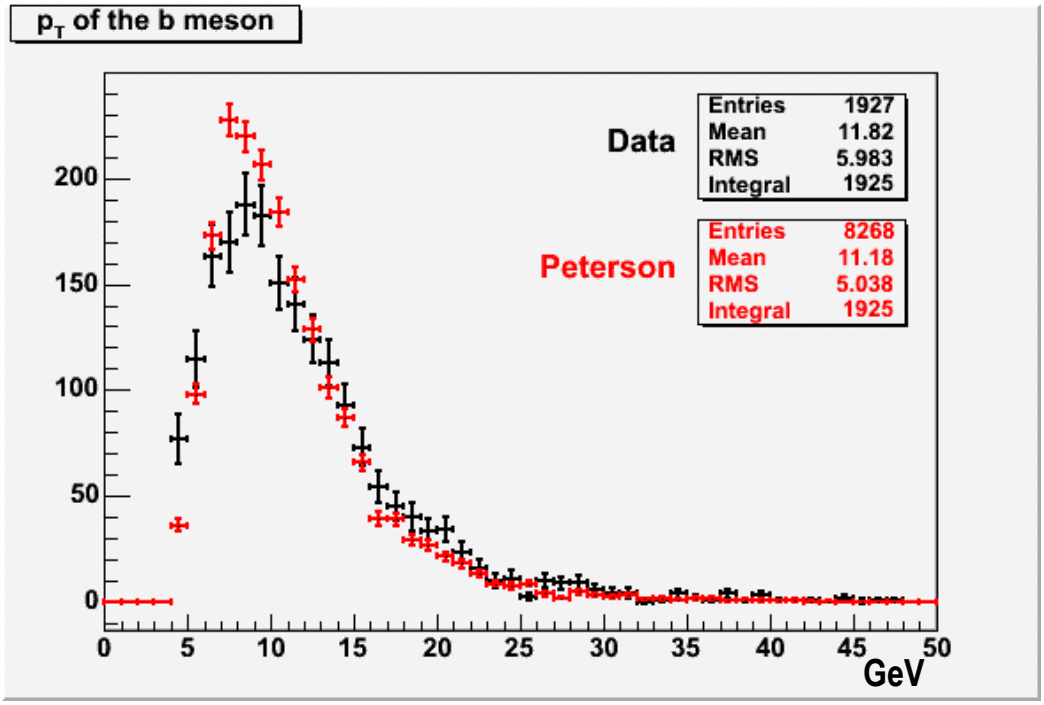

Figure 7.6: *$B^+ \to J/\psi K^+$ data (black) and PYTHIA with msel=5 (red) and CKIN(3)=5. With this value of CKIN(3), events are missed when $p_T(B)$ is ~ 5 GeV/c or below. All other msel=5 plots in this text have been produced with CKIN(3) set to 0. The histograms are normalized to the same number of B mesons.*

Carlo samples with the msel=5 PYTHIA setting. These studies will be detailed in 7.6.1. Tracks, not attached to the studied b-hadron, corresponding to $\Delta R > 1$ have been used as a control sample to verify the quality of the agreement between simulated and real events.

The fragmentation functions referred to in this document are the following:

- the Bowler model [24] with the parameters $a = 0.3$, $b = 0.58$, and $r_Q = 1$. This model with the specified choice of parameters is the default fragmentation function in PYTHIA.

- the Lund Symmetric Fragmentation Function [38] with the parameters $a = 1.68$, $b = 0.55$. This model with these parameters was shown in Chapter 6 to accurately fit the LEP data. The Lund model with a compatible choice of parameters has been favored by other e^+e^- collider fragmentation analyses [32, 33] .

- the Peterson model [35] with two choices of the parameter: $\varepsilon = 0.006$ and $\varepsilon = 0.002$, referred to in the following as "soft Peterson" and "hard Peterson", respectively. This modeling of the fragmentation function is often used by experimentalists. It has been shown in Chapter 6 and also by [32, 33] that it does not describe well e^+e^- data. A motivation for using this model in the present study will be given in Section 7.7.

A comparison between the Peterson, Lund, and Bowler hadronization models with the parameters detailed above is shown in Figure 7.7 [5]. Unless explicitly specified, when the Bowler, Lund, and Peterson models are mentioned in the following, the values of the parameters specified above are implied. The Peterson models with $\varepsilon = 0.006$ and $\varepsilon = 0.002$ are referred to as soft and hard Peterson, respectively.

The fit of the fragmentation function in PYTHIA has not yet been implemented. However, the method to proceed has been suggested and the necessary tools have been provided. This method will be detailed in Section 7.7.

Comparisons between the distributions of track variables in the data and in two types of Monte Carlo samples, the first with msel=5 and the second with msel=1, have also been done. This step clearly demonstrated that Monte Carlo msel=5 samples do not give an acceptable description of the data, as shown in Section 7.6.1. Therefore, the contribution of the different *b* quark production mechanisms to the different observables must be studied using an msel=1 sample.

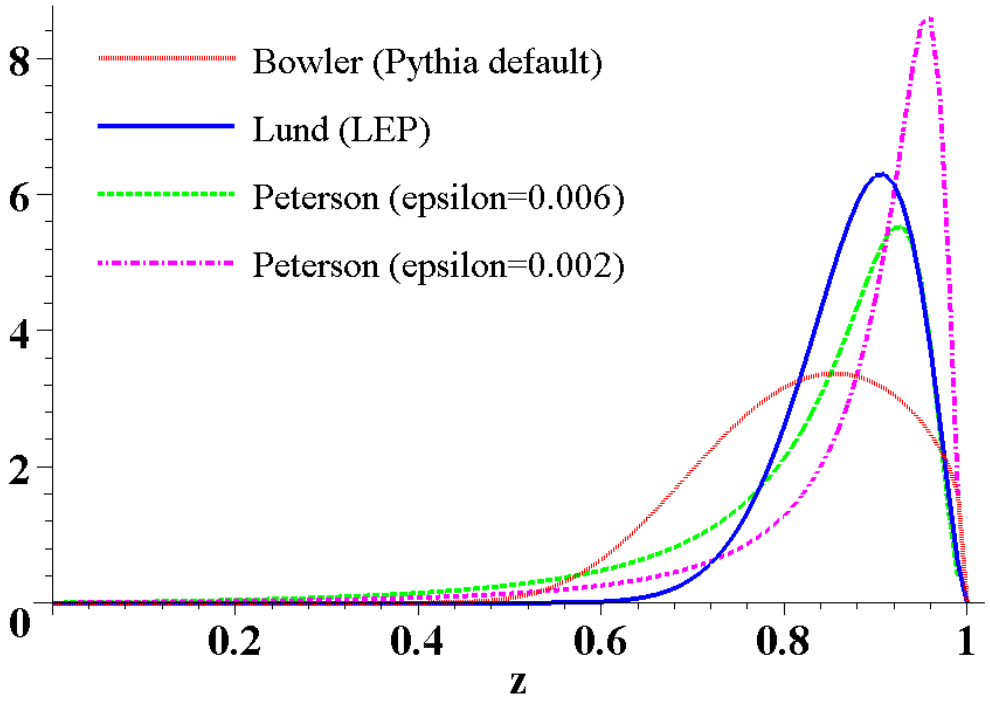

Figure 7.7: *Three hadronization models: Peterson [35] with* $\varepsilon = 0.006$ *(soft) and* $\varepsilon = 0.002$ *(hard), Bowler [24] with* ($a = 0.3$, $b = 0.58$, $r_Q = 1$) *(the default in PYTHIA), and Lund [38] with* ($a = 1.68$ *;* $b = 0.55$).

[5] The Lund and Bowler models have been plotted taking the transverse mass as $m_T = 5.3\ GeV$.

7.5 Preliminary Monte Carlo Studies

The first step of the present analysis was to study in the Monte Carlo the influence of the fragmentation function on different variables related to the B meson and to tracks in its vicinity. One aim of this comparison was to examine qualitatively the effects of modifying the fragmentation function. Additionally, the study was necessary in order to select the variables which are most sensitive to the fragmentation. These variables will be used later to tune the fragmentation function in the Monte Carlo.

The present studies were done on four Monte Carlo samples generated by PYTHIA, fully simulated and analyzed as data. The first sample was generated with the Bowler fragmentation function, the second one with soft (s) Peterson, the third one with hard (h) Peterson and the fourth with Lund.

In PYTHIA, when setting the hadronization model to Lund or to Bowler, the same parameters are used for the fragmentation of light and heavy quarks. The direct use of the rather hard Lund fragmentation function, that has been extracted from e^+e^- measurements for the b quark, might therefore lead to wrong conclusions. It is possible to use this function directly when considering distributions of the B meson itself, but not ones of tracks in the event. However, in near future, it is planned to use the Lund fragmentation function by reweighting events, as explained in Section 7.7. This procedure ensures that the fragmentation function is applied only to b quarks, and that light quark fragmentation remains unaffected. Events generated with the Bowler model correspond to the default option for the PYTHIA simulation.

Unlike the Lund and Bowler fragmentation, the Peterson model in PYTHIA applies only to heavy quarks. This is the reason why it is used for the comparisons in this section.

All the samples were generated with msel set to 5. As explained in Section 7.6.1, a Monte Carlo sample generated with msel=5 fails to describe the data. Initial studies were performed on this type of sample because it is faster to generate and the effect of the fragmentation function is still expected to be similar to that seen in a full msel=1 sample. 16 million events were generated for each of the fragmentation functions,and thus all four samples have the same luminosity. They are directly comparable to each other without any need for a scale factor.

In order to quantify the sensitivity of a variable to the fragmentation function, a χ^2 test has been done. Histograms of a given variable were compared for the three fragmentation functions and the defined χ^2 is

$$\chi^2 = \sum_{bins} \frac{(N_i^1 - N_i^2)^2}{(\sigma_i^1)^2 + (\sigma_i^2)^2} \tag{7.2}$$

where N_i^1 and N_i^2 are the bin contents for the first and the second cases and σ_i^2, σ_i^2 are the corresponding errors. Only bins that contained entries for at least one of the samples have been counted. The χ^2 corresponding to the comparison between the hard Peterson and the default Bowler fragmentation functions are given in Table 7.4. Comparison plots for a choice of the examined variables are shown in Figures 7.8 through 7.12. Both in the table and in the plots, tracks have been considered in a cone of opening $\Delta R = 0.7$ in the η-ϕ plane around the B meson.

p_T, for B mesons and tracks, is the transverse momentum, defined as $|\vec{p} \times \hat{z}|$ is a unit vector pointing to the z direction (the beam axis).The sums over transverse momenta (p_T) and energies refer to all tracks present in the cone defined around the B meson. p_T^{rel} and p_L^{rel} are the transverse and longitudinal components, respectively, of the track's momentum relative to the B meson.

Variable	Number of bins (N)	$\frac{\chi^2}{N}$ (full sample)
p_T of the B meson	46	1.9
p_T of tracks	54	3.3
$\sum p_T$ of tracks	55	2.0
$\frac{E_B - \sum E_{tracks}}{E_B + \sum E_{tracks}}$	38	3.7
$p_T^{rel} = \frac{\lvert\vec{p}_{track} \times \vec{p}_B\rvert}{\lvert\vec{p}_B\rvert}$	28	5.4
$p_L^{rel} = \frac{\vec{p}_{track} \cdot \vec{p}_B}{\lvert\vec{p}_B\rvert}$	59	2.6
$p_T^{leading\,momentum}$	30	2.2
$p_T^{leading\,direction}$	29	2.1
Number of tracks in the cone	11	15.3

Table 7.4: *A χ^2 calculated to compare two Monte Carlo samples, with the Bowler and the hard Peterson fragmentation functions, for a choice of variables. A larger number indicates better sensitivity for the corresponding variable. p_T, for B mesons and tracks, is the transverse momentum, defined as $|\vec{p} \times \hat{z}|$, where $\hat{z}$ is a unit vector pointing to the z direction (the beam axis).*

$p_T^{leading\,momentum}$ is the largest transverse momentum of a track in the cone and $p_T^{leading\,direction}$ refers to the track whose direction is the closest to the B meson's direction.

An interesting result of this comparison is that other variables have better sensitivity to discriminate between hadronization models than the p_T spectrum of the B meson.

The ability to discriminate between different models is also a function of the number of events. When the sample is smaller, the discrimination becomes more difficult. The Monte Carlo samples used for the present comparisons are roughly ten times larger than the actual data sample.

The total B hadron production cross section is expected to be independent of the choice of a given fragmentation distribution. Nevertheless, as measured events have to satisfy selection criteria, some variations on the number of accepted events can be observed. The global variation of the measured B production cross section, expected when considering different fragmentation models, can be derived from the number of reconstructed B mesons. It was found to be:

- 2.4% (Bowler and hard Peterson).
- 4.0% (Bowler and soft Peterson).
- 3.6% (Bowler and Lund).
- 6.5% (soft Peterson and hard Peterson).
- 6.2% (soft Peterson and Lund).
- 0.3% (hard Peterson and Lund).

Those variations arise from the fact, that different shapes of fragmentation functions result in migration of events between the observed and non observed parts of the signal, due to analysis cuts. While the global variation remains relatively small, between 0.3% and 6.5%, the cross

section may vary by $\sim 20\%$ in some regions of the p_T spectrum of the B meson. This is illustrated in Figure 7.13. Similar variations of 20% in the cross section, obtained with the soft and hard Peterson fragmentation function, in the region of $p_T(B) \sim 20$GeV have been reported in [5].

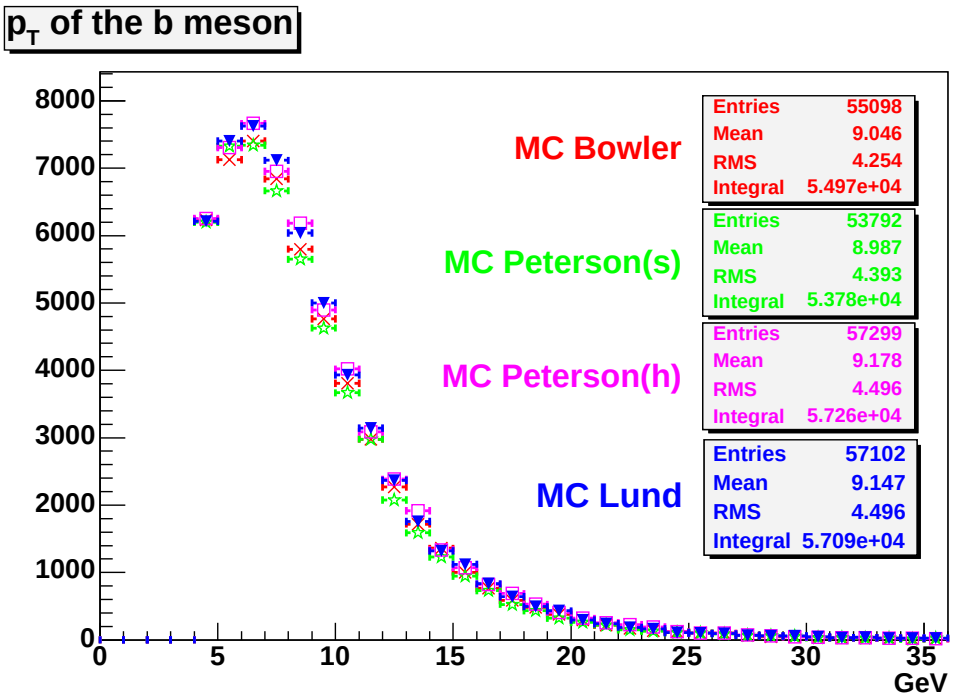

Figure 7.8: *The p_T spectrum of the B meson in four Monte Carlo samples, generated with different fragmentation functions. The Lund model is not shown in corresponding plots for tracks.*

7.6 Data and Monte Carlo Comparisons

7.6.1 Comparisons with msel=5 Samples

An msel=5 PYTHIA sample is simple to generate. Only b and $\bar{b}$ quarks are produced by the hard scattering simulation, and therefore no event has to be discarded. Nevertheless, the studies in this section demonstrate that an msel=5 sample is not able to correctly describe the data, and that in particular, it cannot be used for fragmentation studies.

Comparisons of the data to a msel=5 samples generated with the Bowler and soft Peterson functions are presented in Figures 7.14 through 7.16. These figures show the p_T spectrum of the B meson ($p_T(B)$), the p_T of tracks in a cone of opening $\Delta R = 0.7$ around the B meson, and the η distribution for tracks which have $\Delta R > 1$. (control tracks). All the histograms, in this section, are normalized to the number of B mesons

From Figure 7.14 it is clear that the msel=5 PYTHIA generation creates a $p_T(B)$ distribution which is *too soft, i.e.,* the non-perturbative fragmentation function needed to reproduce the measured distribution would be peaked very close to 1. Indeed, in the preliminary studies of

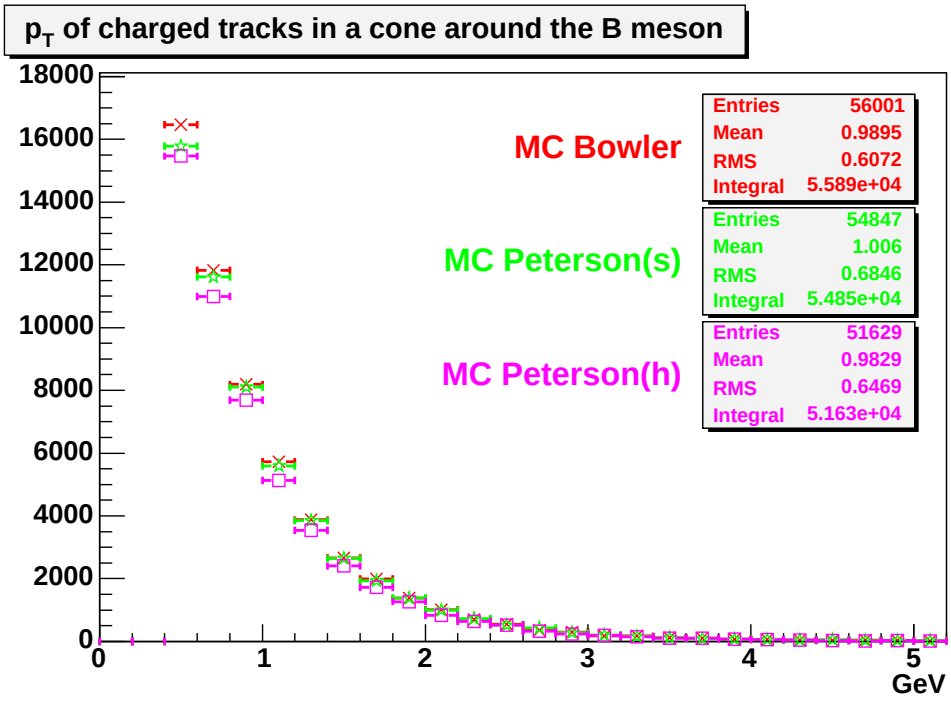

Figure 7.9: *The p_T spectrum of tracks in a cone around the B meson in three Monte Carlo samples, generated with different fragmentation functions.*

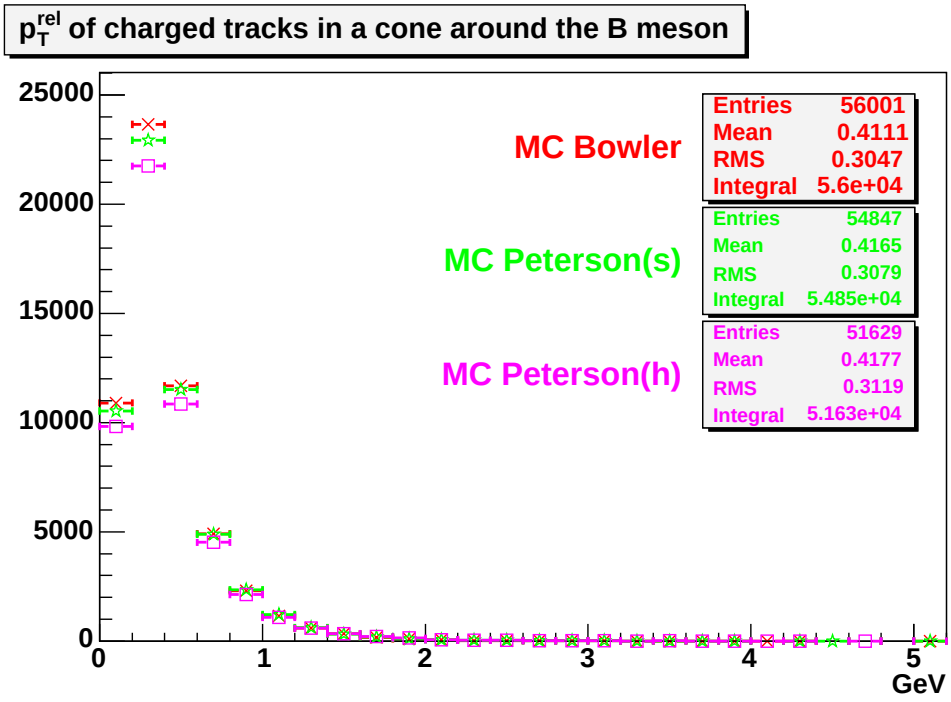

Figure 7.10: *The component of tracks momentum transverse to the B meson, for tracks in a cone around it in three Monte Carlo samples, generated with different fragmentation functions.*

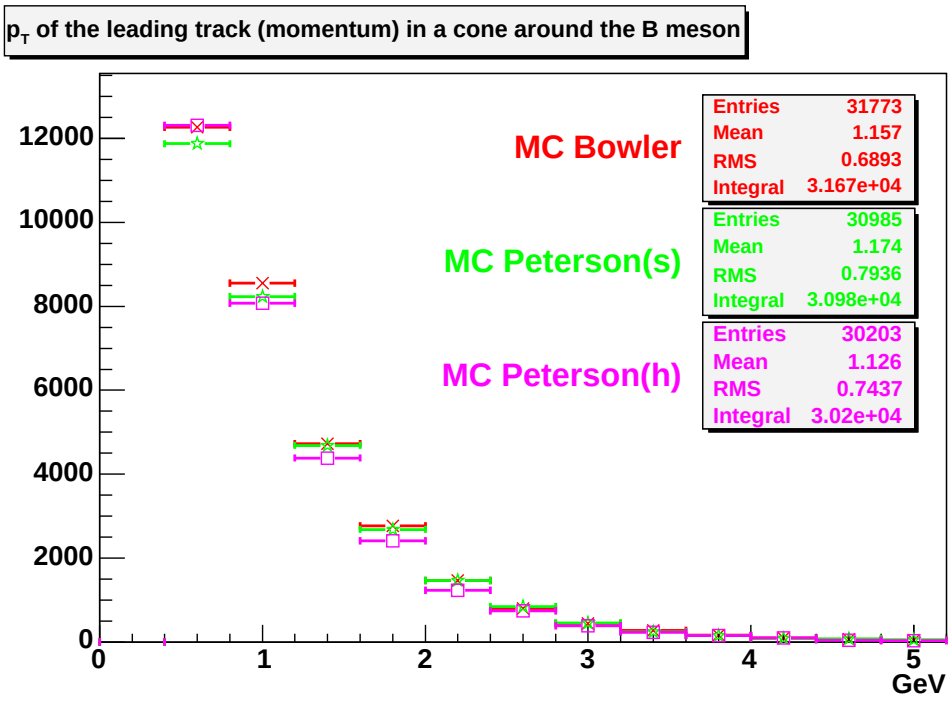

Figure 7.11: *The p_T spectrum of the highest p_T track in a cone around the B meson in three Monte Carlo samples, generated with different fragmentation functions.*

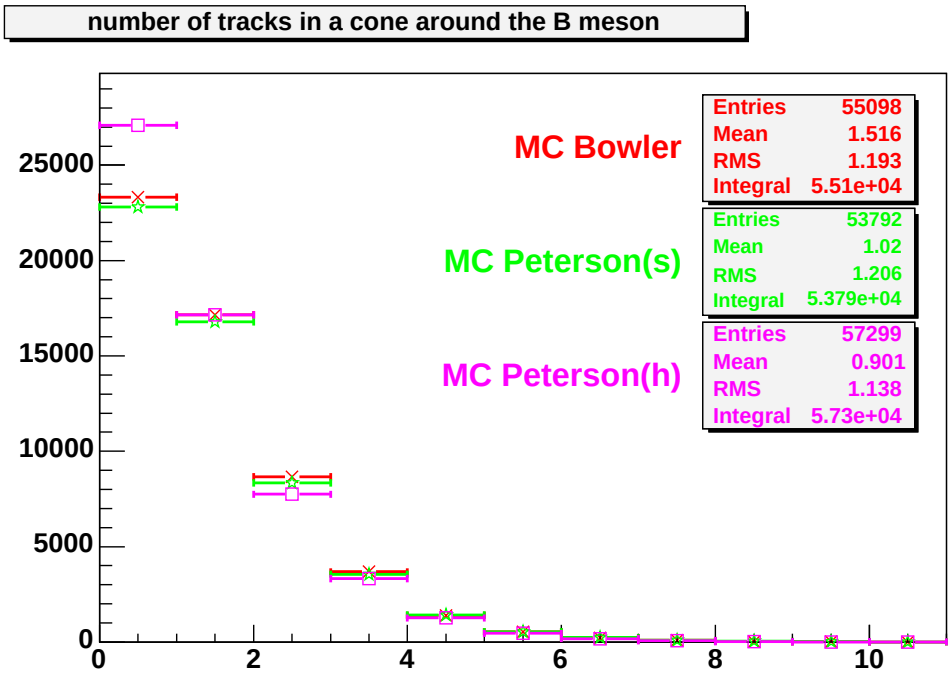

Figure 7.12: *The multiplicity of tracks in a cone of opening $\Delta R = 0.7$ around the B meson, in three Monte Carlo samples, generated with different fragmentation functions.*

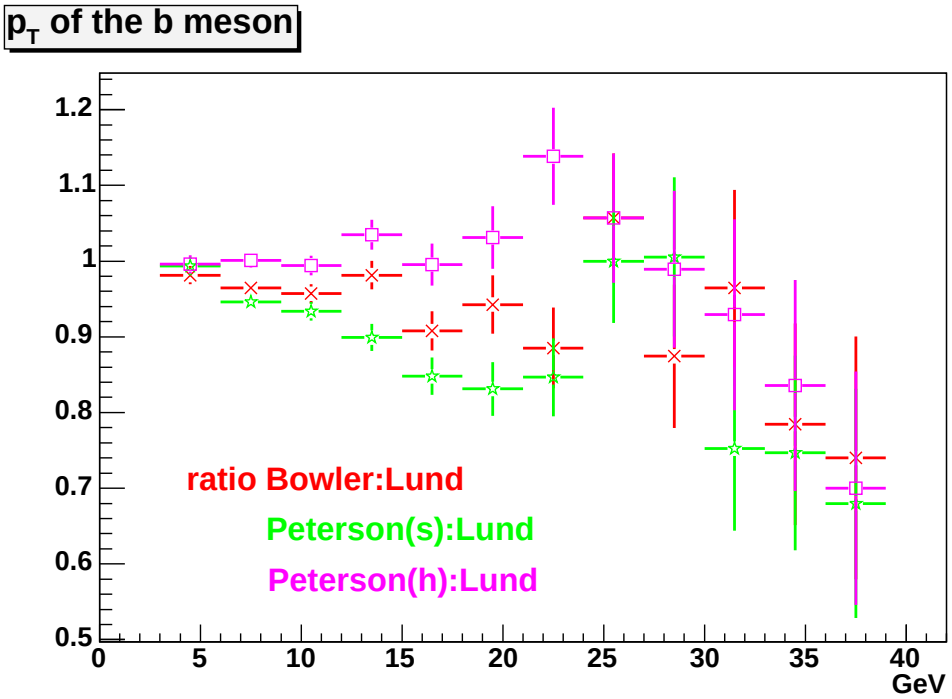

Figure 7.13: *Variation of the B production cross section, expected when considering different fragmentation models. Ratios of the production rate expected in bins of the p_T distribution, are presented for Bowler, soft Peterson and hard Peterson fragmentation functions, with respect to the prediction obtained with the Lund fragmentation function.*

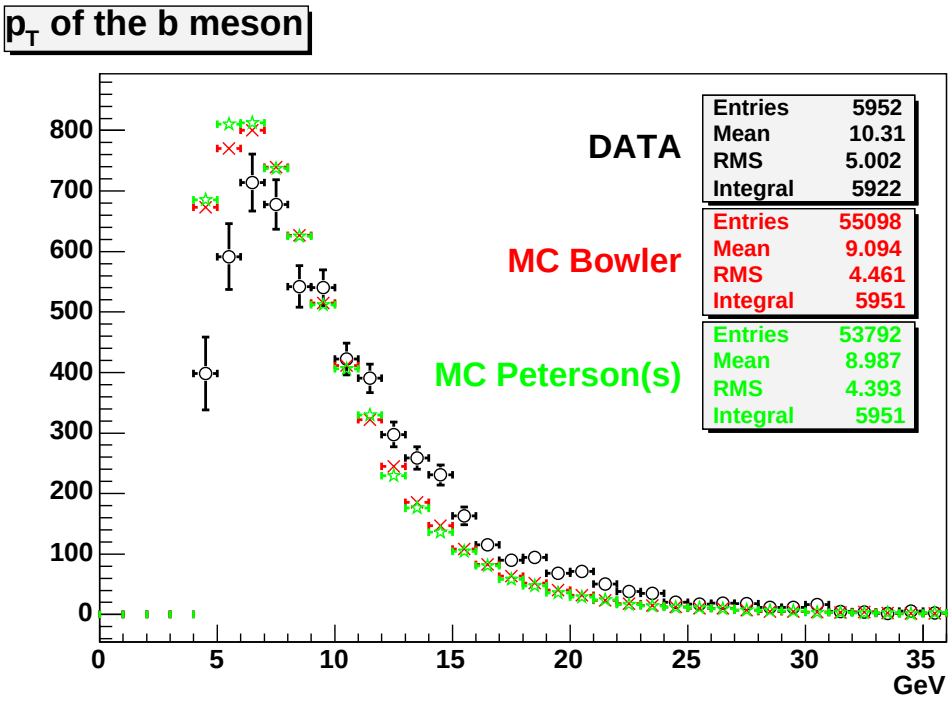

Figure 7.14: *The p_T spectrum of the B. Comparison between the data and two Monte Carlo msel=5 samples with the Bowler and Peterson fragmentation functions. Histograms are normalized to the number of B mesons.*

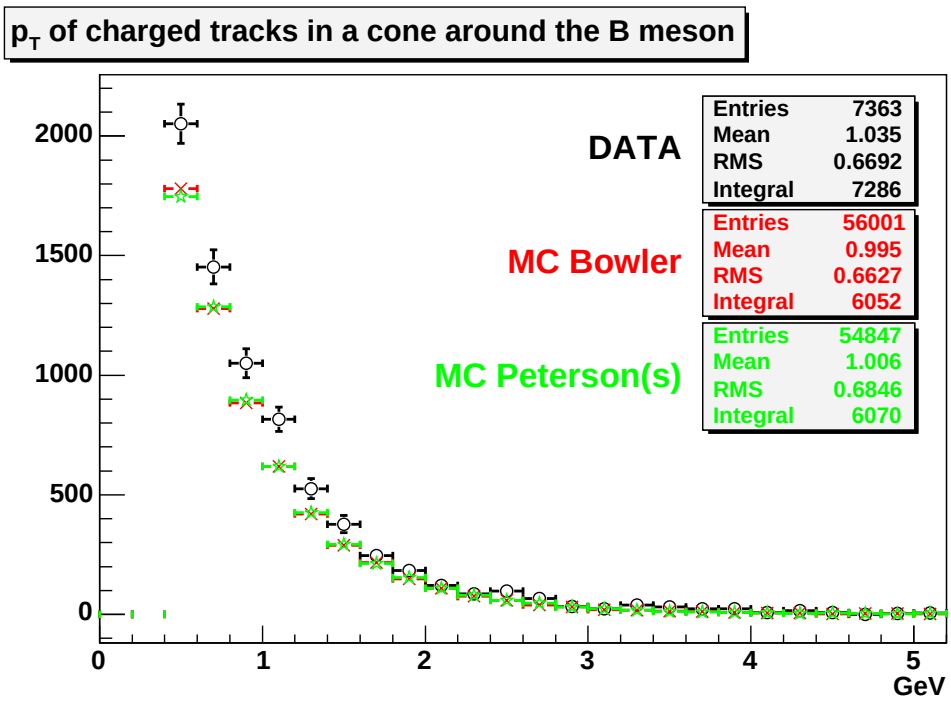

Figure 7.15: *The p_T spectrum of tracks in a cone around the B meson. Comparison between the data and two Monte Carlo (msel=5) samples with the Bowler and Peterson fragmentation functions. Histograms are normalized to the number of B mesons.*

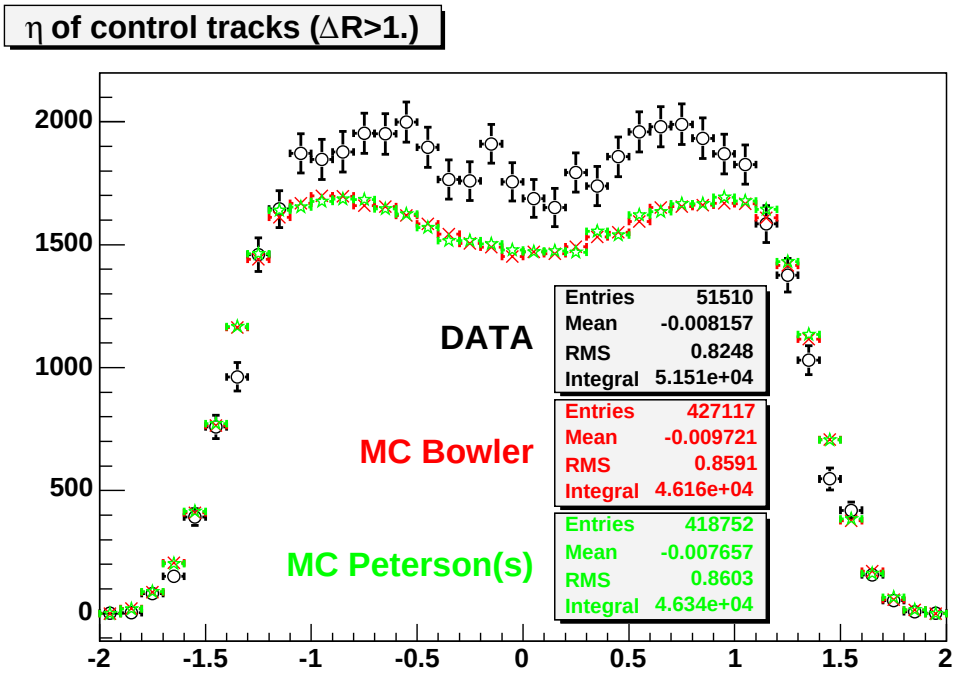

Figure 7.16: *The η spectrum of tracks in a cone with $\Delta R > 1$ (control tracks). Comparison between the data and two Monte Carlo (msel=5) samples with the Bowler and Peterson fragmentation functions. Histograms are normalized to the number of B mesons.*

Section 7.5, one can see that the B meson's spectrum is not very sensitive to the fragmentation function. This indicates that the data-MC discrepancy in $p_T(B)$ distribution is difficult to correct with fragmentation alone.

In addition, such a fragmentation function would not be able to simultaneously accommodate the distributions of the momenta of the tracks in the cone around the B meson, which requires a softer fragmentation function in order to agree with the data. This can be understood qualitatively in the distributions for tracks in 7.5, as for example Figure 7.15, which shows the fact that tracks in the data are harder than in the Monte Carlo.

This study shows clearly that Monte Carlo samples generated with msel=5 fail to describe the data.

7.6.2 Comparisons with an msel=1 Sample

In contrast to a msel=5 PYTHIA generation, where $b\bar{b}$ quarks are created from lowest order diagrams, all other QCD b production mechanisms are simulated in a msel=1 PYTHIA job. Adding the other bb creation mechanisms improves the agreement between data and MC in the p_T of the B distribution. This can be seen in Figure 7.17. All the histograms in the present section are normalized with respect to the same number of B mesons.

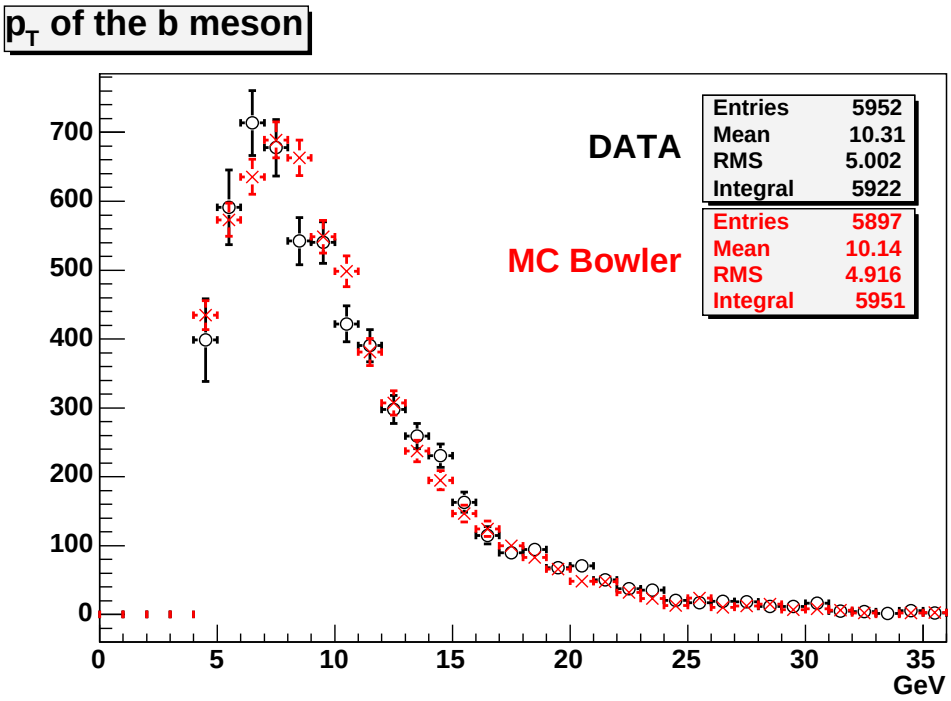

Figure 7.17: *With the msel=1 PYTHIA setting, there is good agreement between data and Monte Carlo for the p_T of the B meson. Histograms are normalized to the number of B mesons.*

The following plots show that, unlike the msel=5 samples, with msel=1 the data and Monte Carlo agreement is acceptable simultaneously for the p_T spectrum of the B meson (Figure 7.17),

the tracks in the vicinity of the *B* meson (Figures 7.18- 7.25), and for control tracks (Figures 7.26- 7.29).

Figure 7.29 indicates a loss of efficiency for track reconstruction in data as compared with the Monte Carlo. for cross-section measurements, only the region $|\eta| < 1$ has been used.

Figure 7.30 shows a map of the track activity in the event in data and for the msel=1 Monte Carlo sample. These figures show the distance of tracks from the *B* meson in the η-ϕ plane ($\Delta\eta$ -$\Delta\phi$). The activity of tracks in the vicinity of the *B* meson, which is present in the center of the plane, is clear. Also the tracks accompanying the opposite side *B* meson are visible.

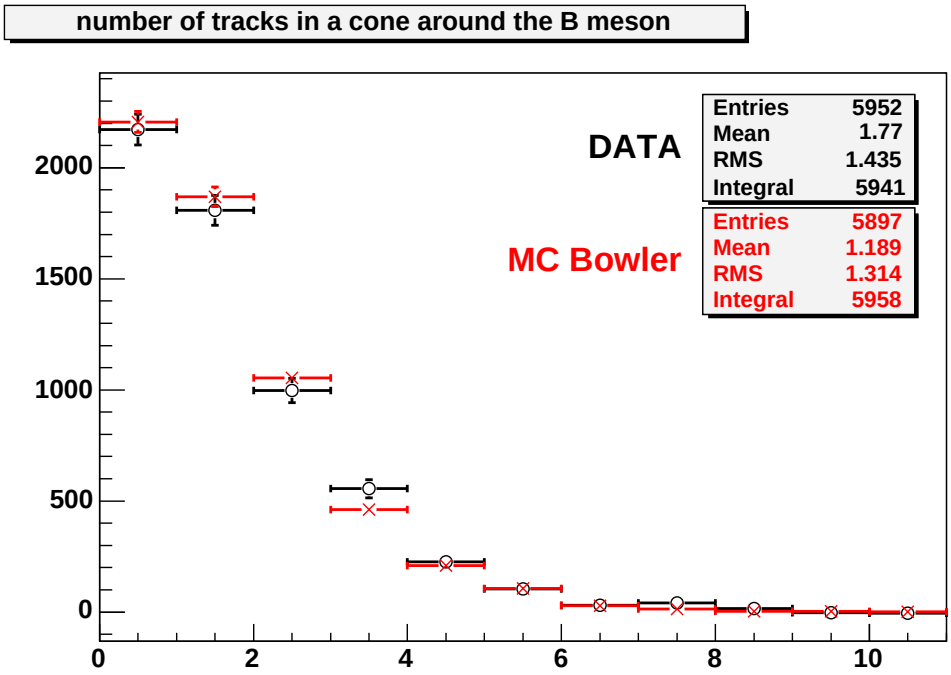

Figure 7.18: *The multiplicity of tracks within a cone of $\Delta R < 0.7$ around the B meson in the msel=1 PYTHIA sample. Histograms are normalized to the number of B mesons.*

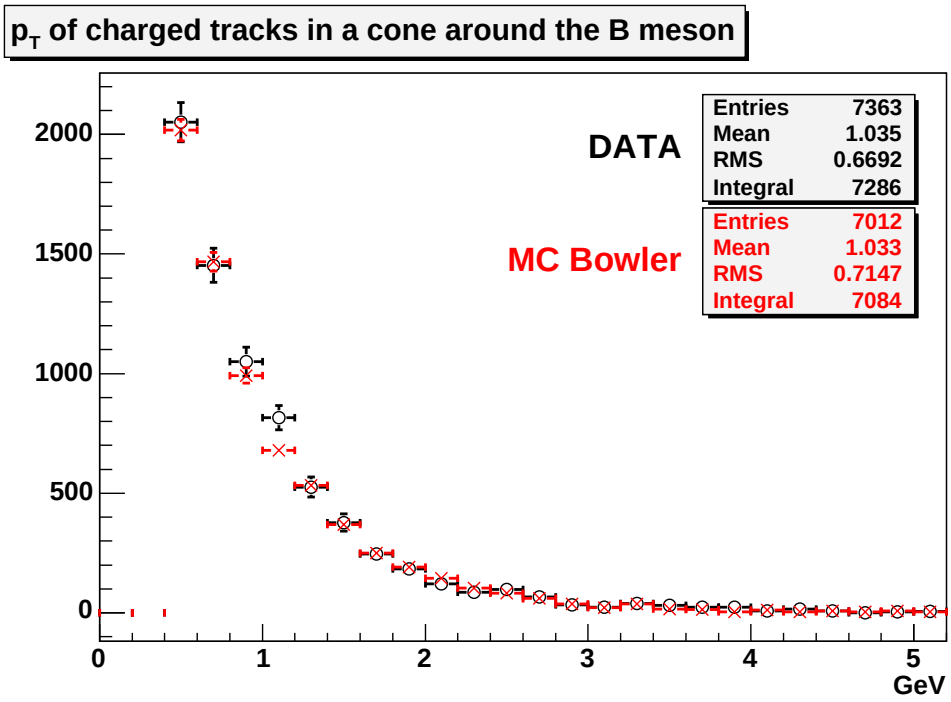

Figure 7.19: *The transverse momentum p_T of all tracks within a cone of $\Delta R < 0.7$ around the B meson in the msel=1 PYTHIA sample. Histograms are normalized to the number of B mesons.*

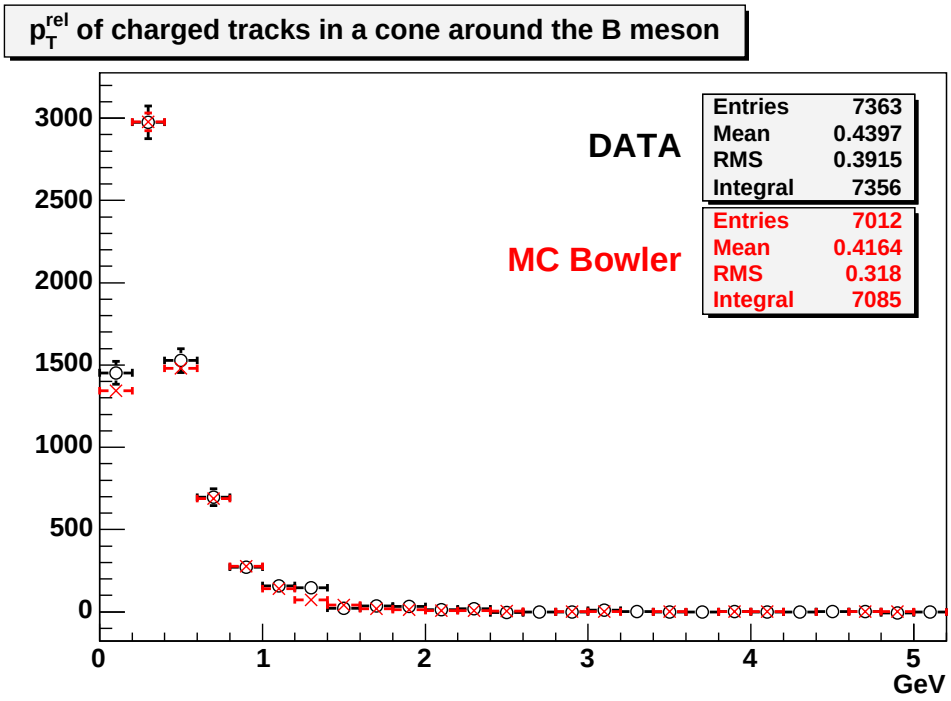

Figure 7.20: *p_T^{rel} of tracks within a cone of $\Delta R < 0.7$ around the B meson in the msel=1 PYTHIA sample. Histograms are normalized to the number of B mesons.*

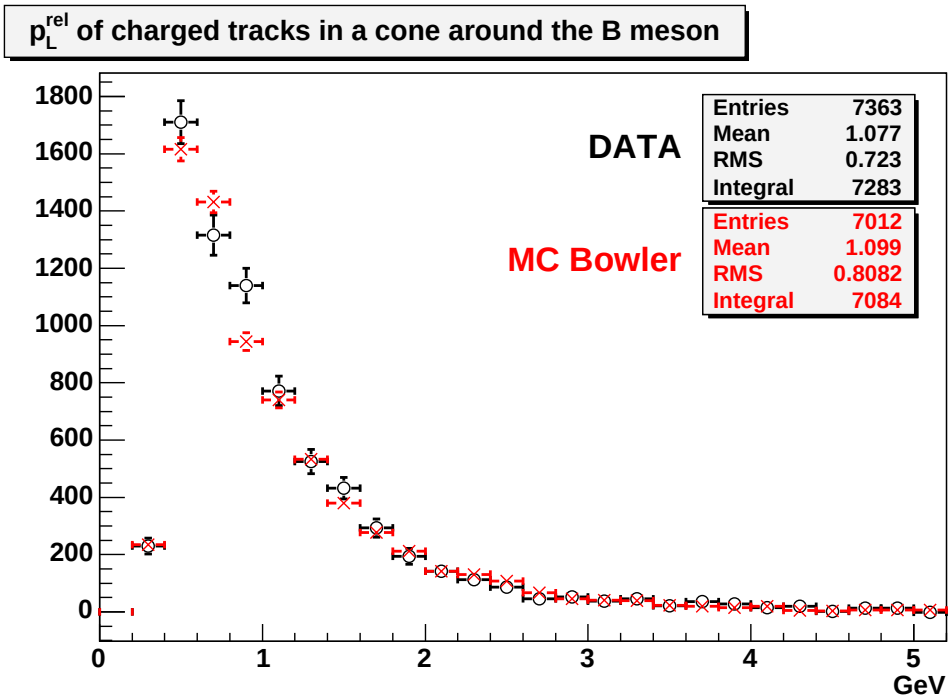

Figure 7.21: *p_L^{rel} of tracks within a cone of $\Delta R < 0.7$ around the B meson in the msel=1 PYTHIA sample. Histograms are normalized to the number of B mesons.*

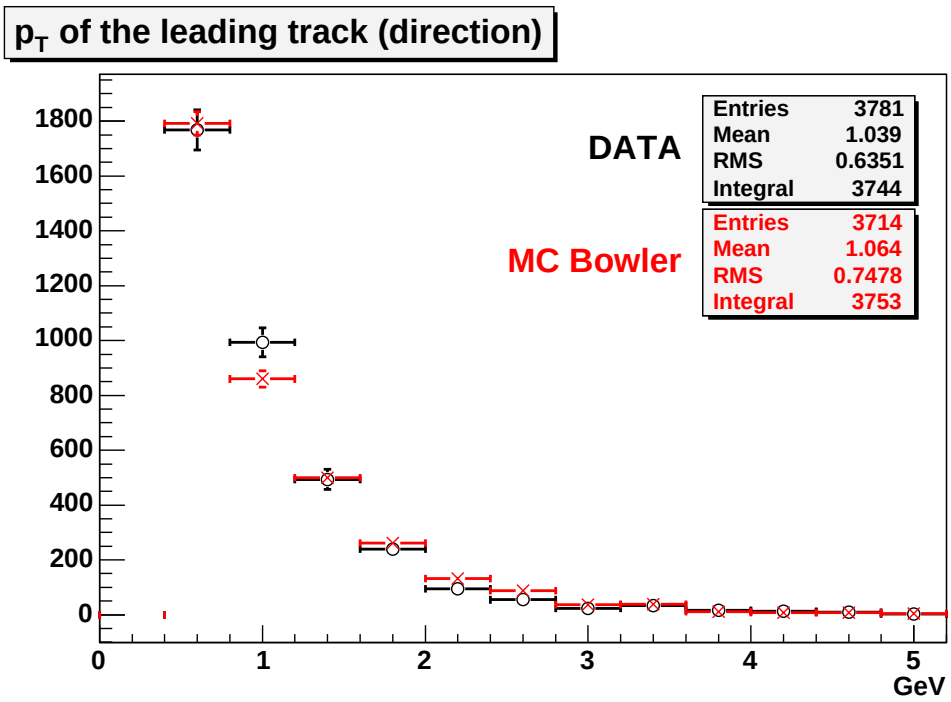

Figure 7.22: *The transverse momentum p_T for the leading directional track in a cone of $\Delta R < 0.7$ around the B meson in the msel=1 PYTHIA sample. Histograms are normalized to the number of B mesons.*

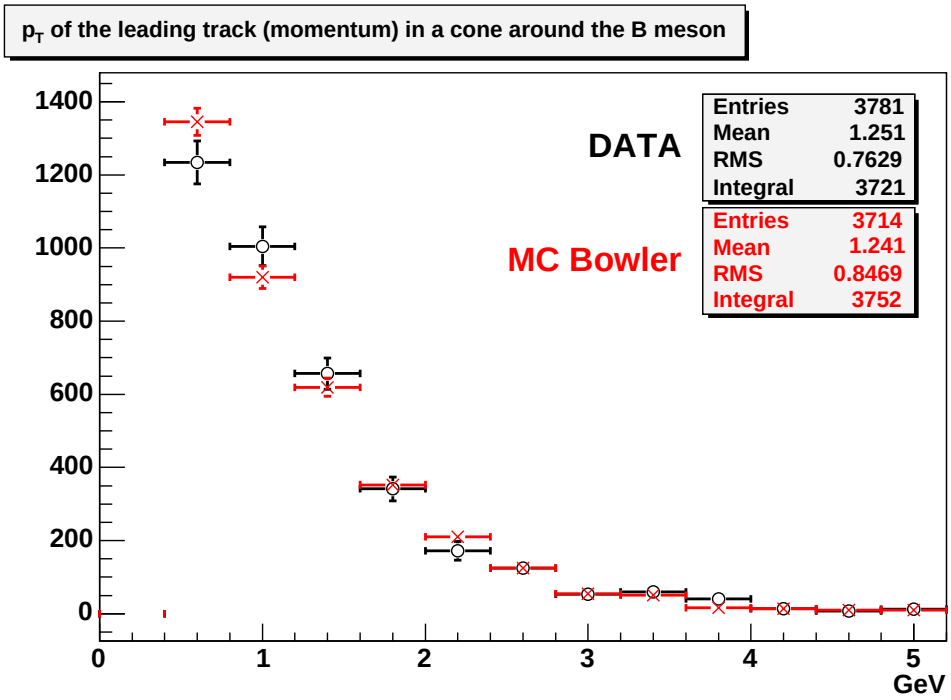

Figure 7.23: *The transverse momentum p_T for the leading momentum track in a cone of $\Delta R < 0.7$ around the B meson in the msel=1 PYTHIA sample. Histograms are normalized to the number of B mesons.*

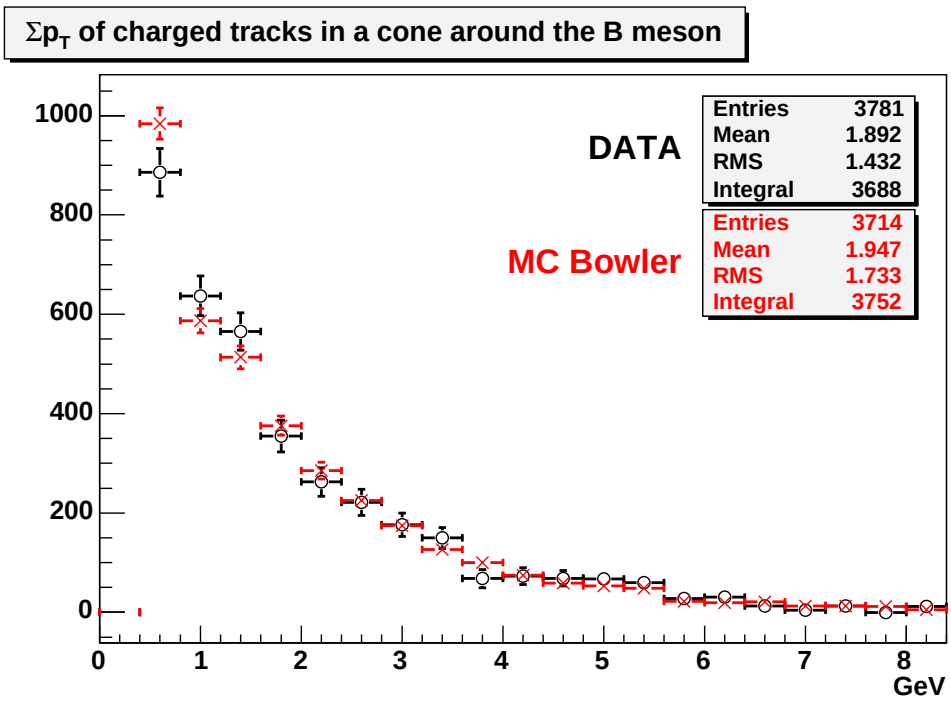

Figure 7.24: *The sum of transverse momenta of tracks in a cone of $\Delta R < 0.7$ around the B meson in the msel=1 PYTHIA sample. Histograms are normalized to the number of B mesons.*

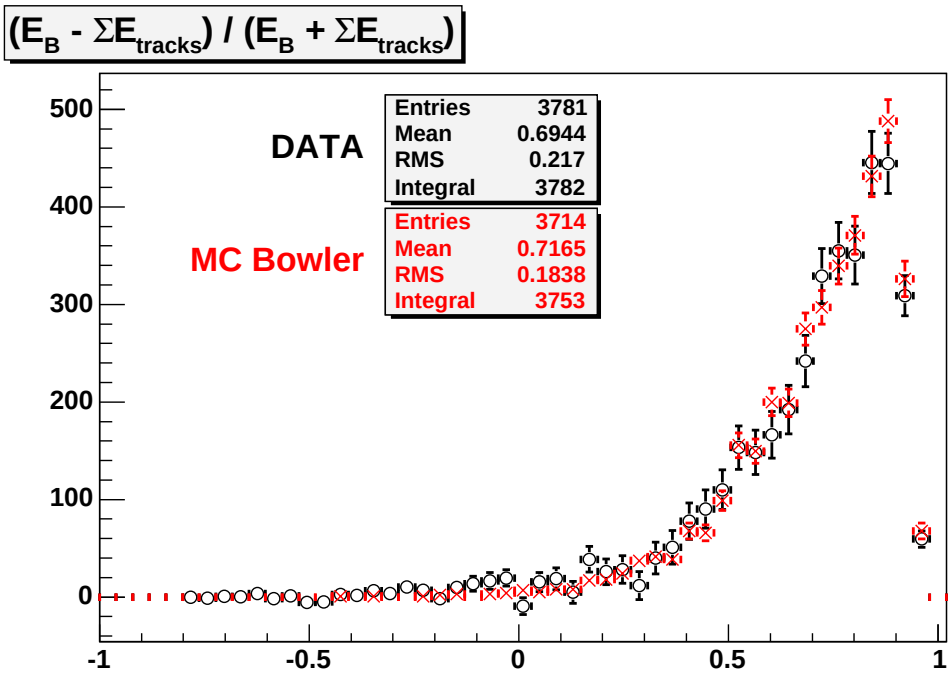

Figure 7.25: *The distribution of the variable* $\frac{E_B - \sum E_{tracks}}{E_B + \sum E_{tracks}}$*, where the sum is over tracks in in a cone of* $\Delta R < 0.7$ *around the B meson in the msel=1 PYTHIA sample. Histograms are normalized to the number of B mesons.*

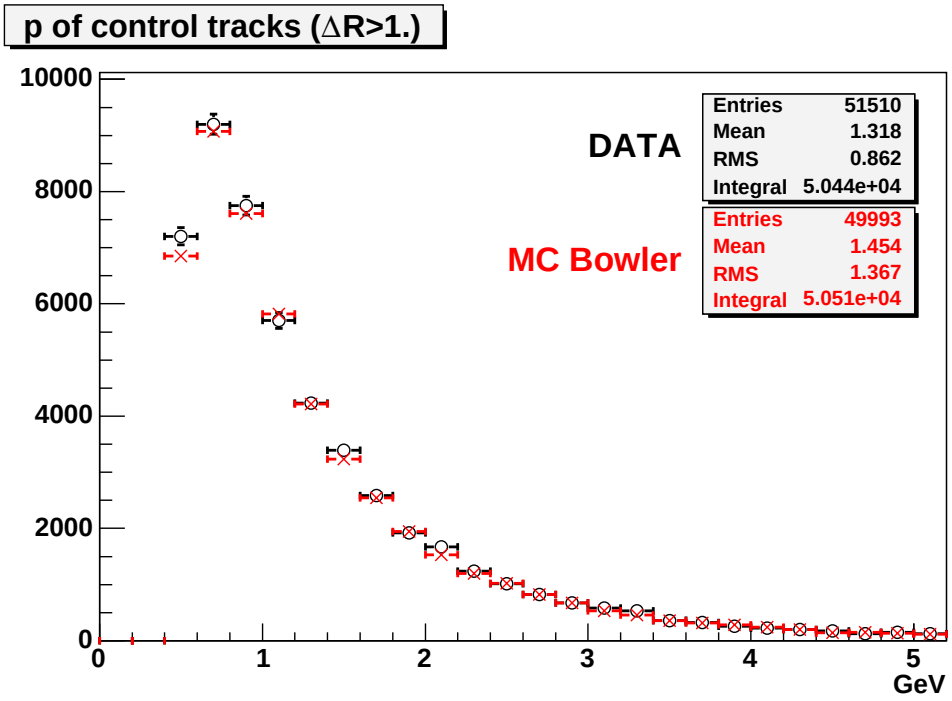

Figure 7.26: *Plot of the 3-momentum for control tracks in the msel=1 PYTHIA sample. Histograms are normalized to the number of B mesons.*

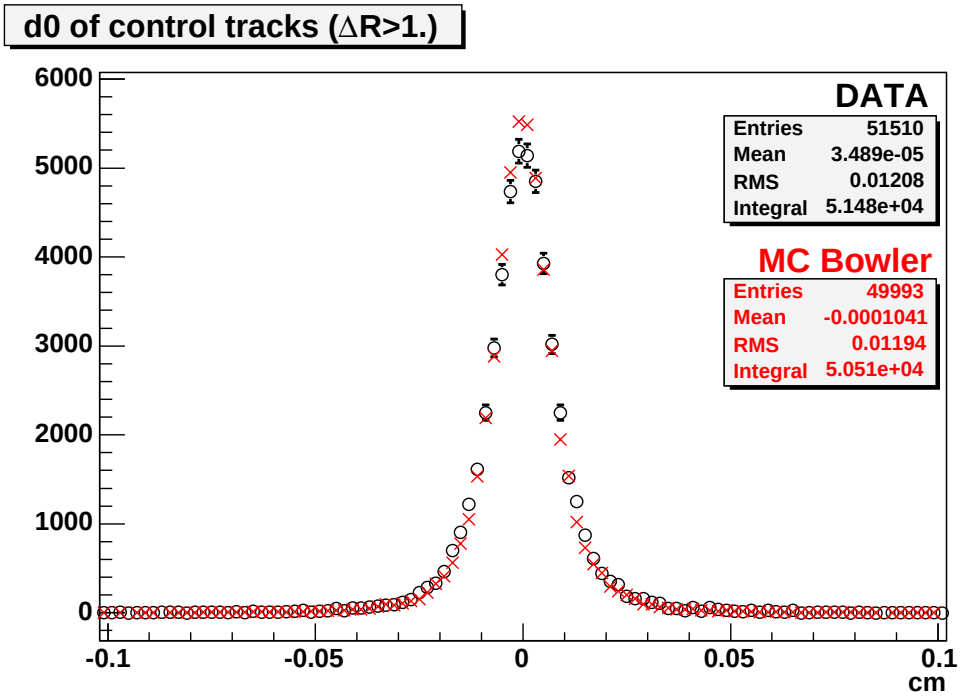

Figure 7.27: *Plot of the impact parameter d_0 for control tracks in the msel=1 PYTHIA sample. Histograms are normalized to the number of B mesons.*

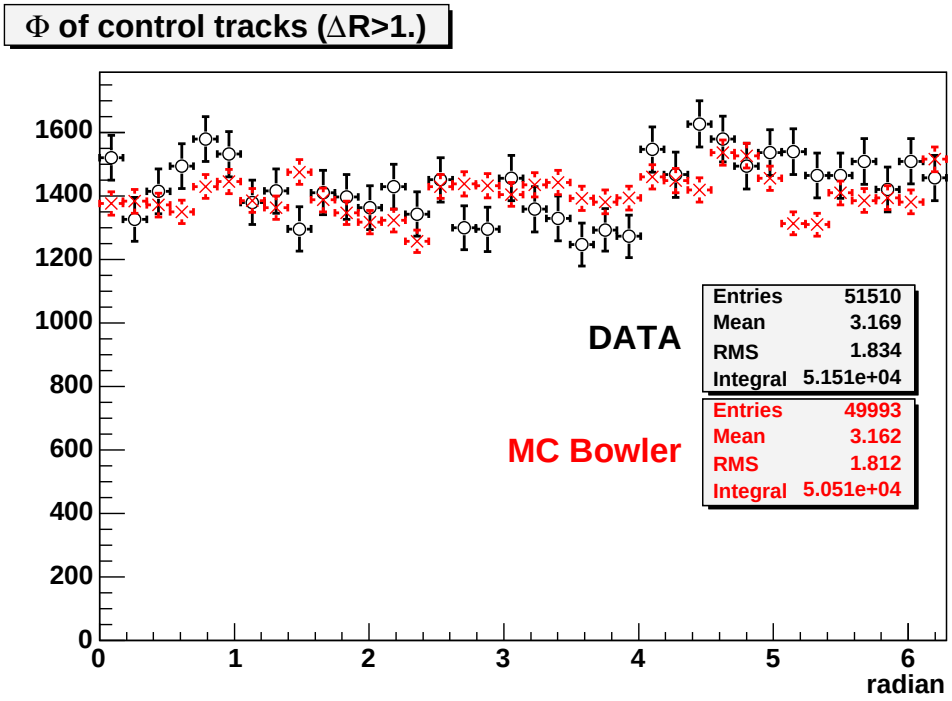

Figure 7.28: *Plot of the angular variable ϕ for control tracks in the msel=1 PYTHIA sample. Histograms are normalized to the number of B mesons.*

The b quark production mechanisms in each event of the Monte Carlo sample have been classified as explained in Section 2.2.2 . Figures 7.31 through 7.34 show the expected contributions from each of the b production mechanisms (FC, FE, GS) for a few distributions. In Figure 7.31 the subclasses are also given as a digit from 1 to 8, following the prescription of Section 2.2.2 . The angular distributions in $\Delta\eta$ and $\Delta\Phi$ of tracks with respect to the B meson are shown in Figure 7.34. The distributions of $p_T(B)$ and p_T of tracks in a cone around the B meson, differ between the three production mechanisms. This is illustrated by Figures 7.35 and 7.36, where histograms for all the three contributions are normalized to unity. A similar comparison has been done for the angular distributions of tracks, $\Delta\eta$ and $\Delta\Phi$, in Figure 7.37. For these two variables the shapes of the contributions from FC, FE and GS are similar.

The approximate ratio of contributions from the different mechanisms is $FC : FE : GS \sim 1 : 1.9 : 0.8$. The FC, which is the only mechanism generated when msel is set to 5, accounts only for $\sim 27\%$ of the events.

7.7 A Method of Fitting the Fragmentation Function Parameters

In this section a method is proposed to fit the fragmentation function in PYTHIA to achieve the best agreement with the data without any need of model input. The proposed method allows to proceed with a single Monte Carlo sample, avoiding the generation of multiple large samples to scan the fragmentation function's parameter space. This procedure is similar to the one applied in Chapter 5.

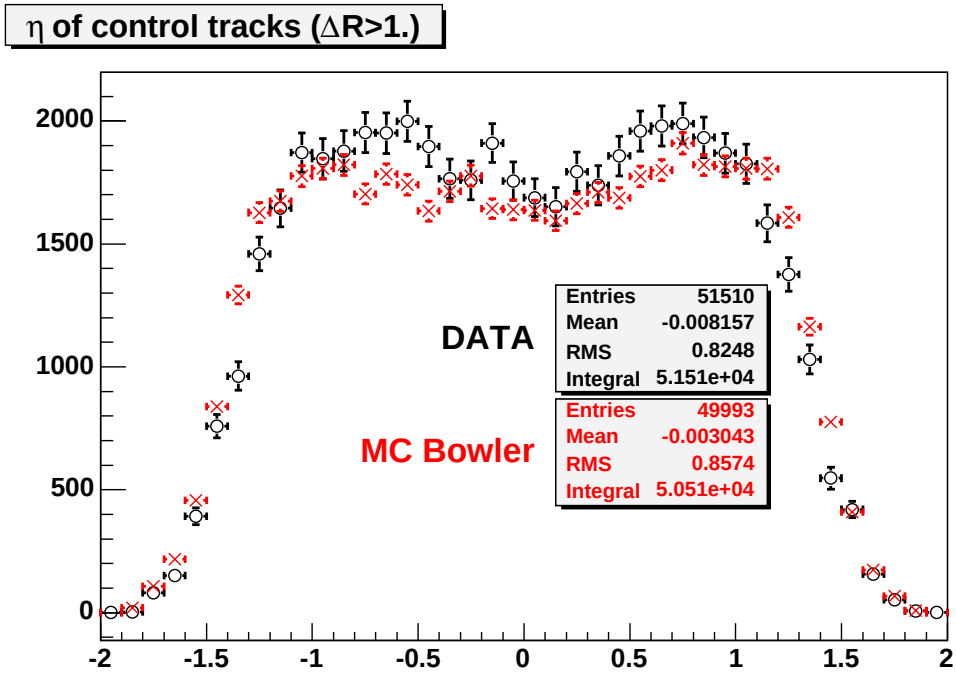

Figure 7.29: *Plot of the pseudo-rapidity variable* η *for the control tracks in the msel=1 PYTHIA sample. Histograms are normalized to the number of B mesons.*

As explained above, the resulting physics distributions are not solely influenced by the fragmentation function, which describes the non-perturbative part of the process. The observables are also influenced by the physics of the hard scattering process, and the perturbative QCD part of the fragmentation. Fitting the fragmentation function by the proposed method will be valid within the choice of PYTHIA parameters that has been used in the Monte Carlo sample used for the fit. Therefore, before proceeding with the fragmentation fit, the other PYTHIA parameters should be tuned to the best data and Monte Carlo agreement for physics observables that are less sensitive to the fragmentation function.

The events generated by PYTHIA are passed through the full simulation of the detector and the trigger system as described in Section 7.3.1. The application of the method supposes that this simulation phase is under control. Distributions of tracks in the control region allow to validate this step.

The idea behind the proposed method is to reweight the Monte Carlo sample with respect to the light cone variable z. Given that the fragmentation function is the distribution used to generate this variable in the Monte Carlo, reweighting the sample with respect to z allows direct modification of the fragmentation function of the sample.

In order to apply the method, it is necessary to generate a Monte Carlo sample and keep the fragmentation information that intervenes in the string model. This information consists of the variable z that has been used and of the primary B hadron it has been applied for. The identification of the primary B hadron is given by the line number that refers to it in the PYTHIA event record . Each primary hadron from a b or c quark fragmentation in the event has its own value of z. This information is not kept by default in the CDF framework. As a part of the

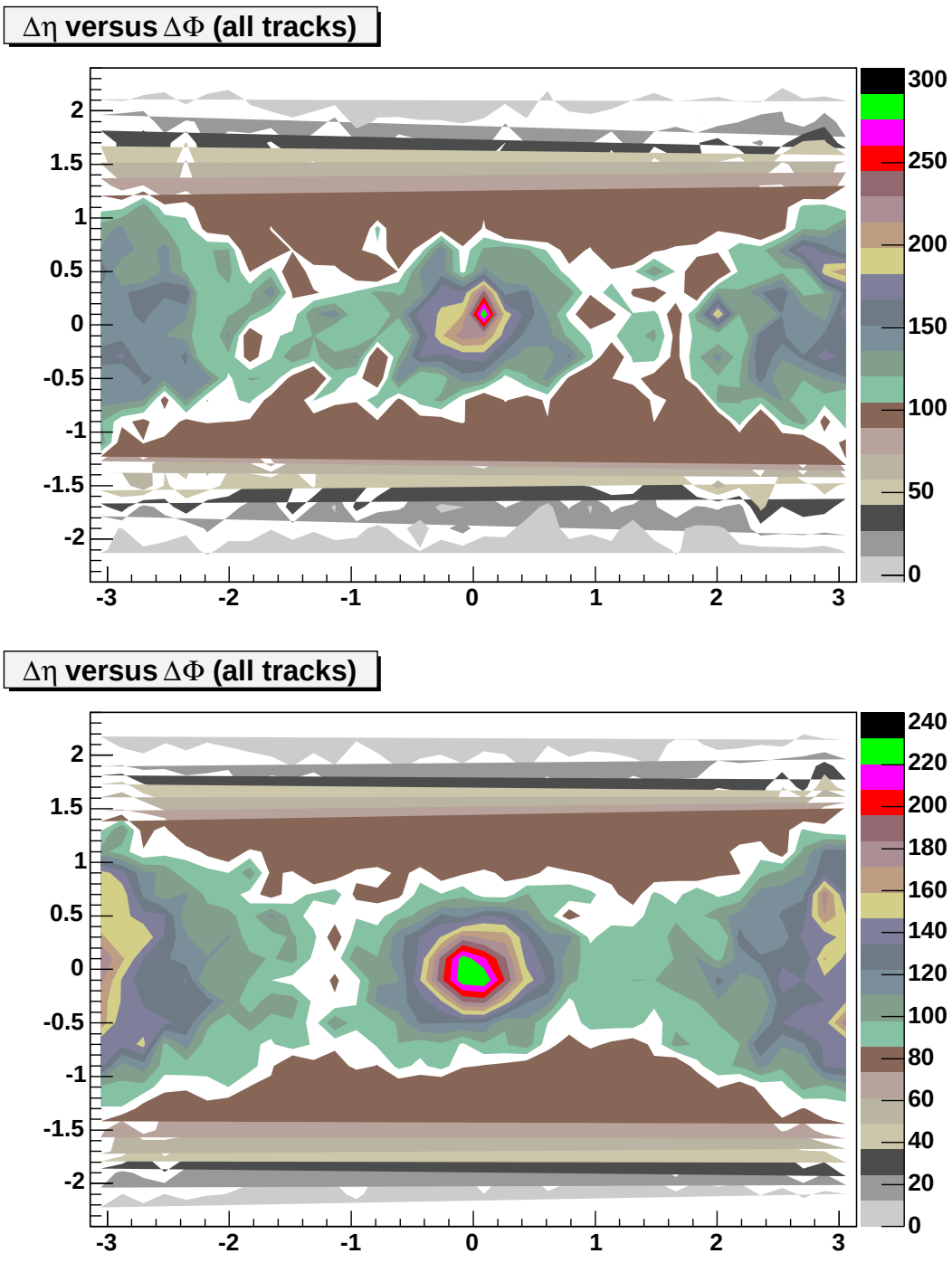

Figure 7.30: *The track activity in the Δη-Δϕ plane around the B meson in data (top) and the Monte Carlo msel=1 sample (bottom). The activity in the vicinity of the opposite side B meson is also visible.*

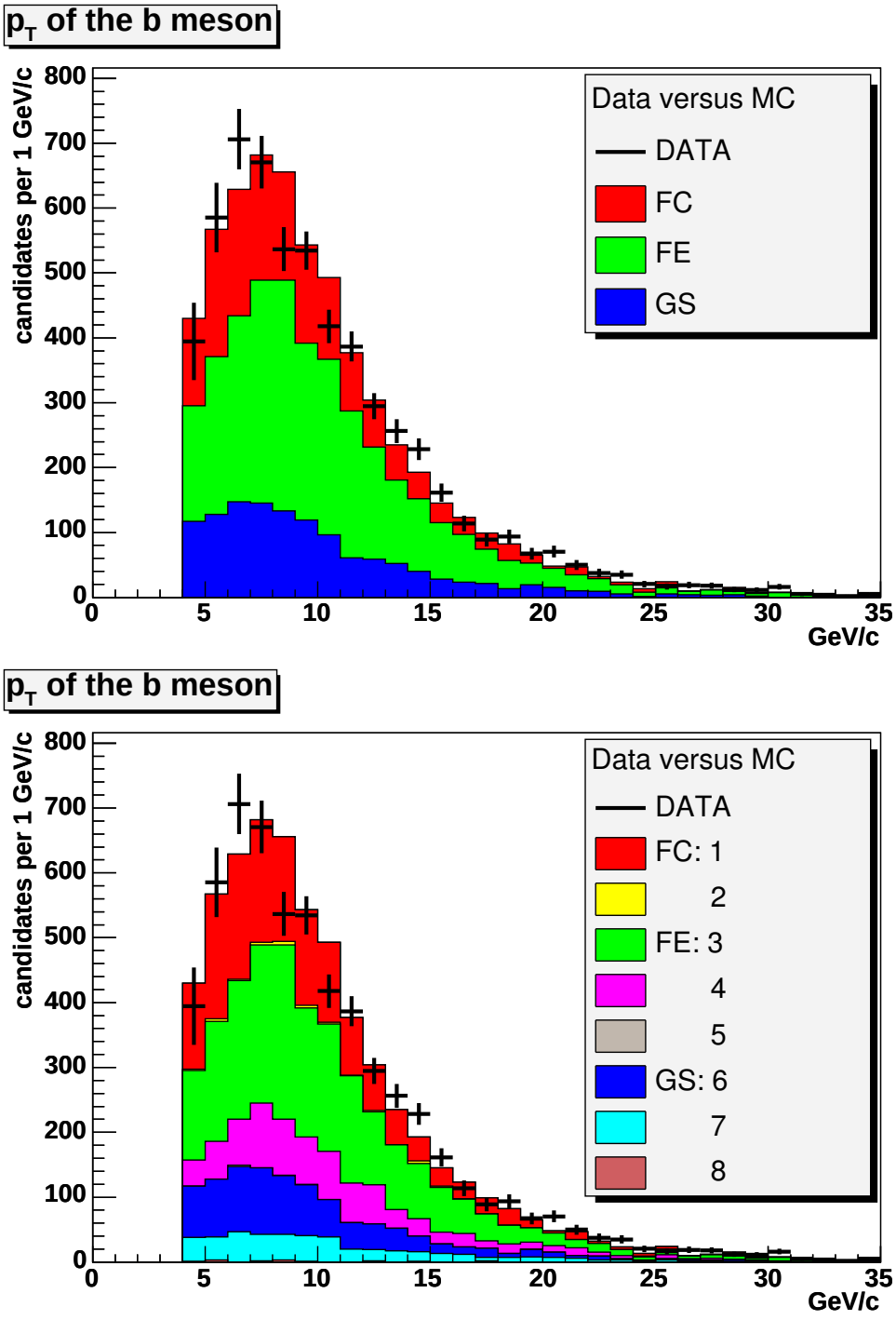

Figure 7.31: *Data and Monte Carlo comparison for $p_T(B)$. The upper plot details the different contributions in the Monte Carlo sample from the FC, FE and GS b quark production mechanisms. The lower plot details the contributions for the classes of processes explained in Section 2.2.2. Histograms are normalized to the number of B mesons.*

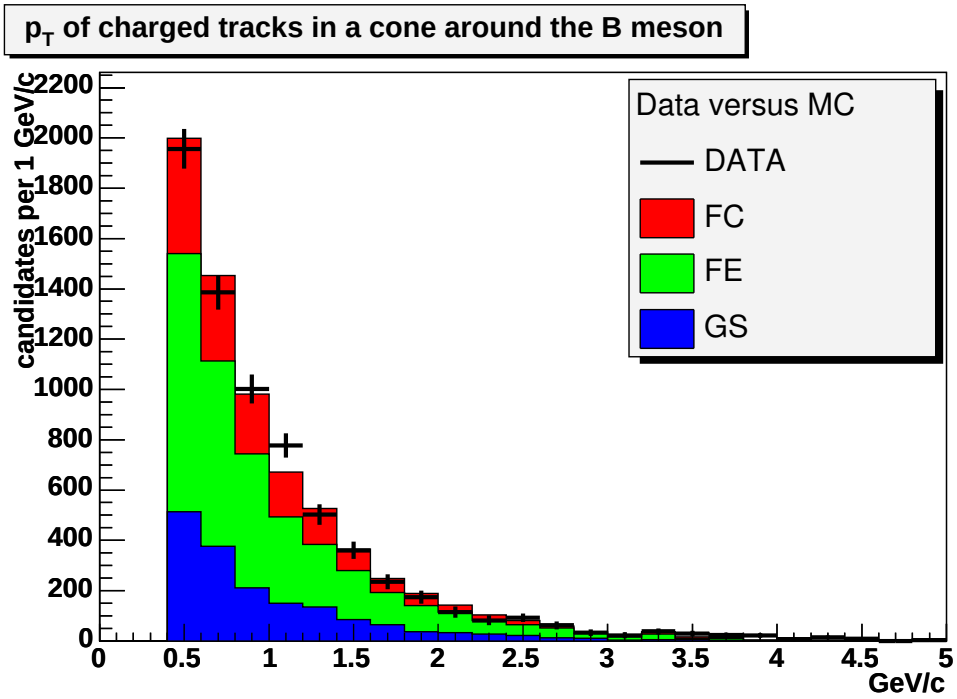

Figure 7.32: *Data and Monte Carlo comparison for p_T of tracks in a cone around the B meson. Contributions in the Monte Carlo sample from the FC, FE and GS b quark production mechanisms are detailed. Histograms are normalized to the number of B mesons.*

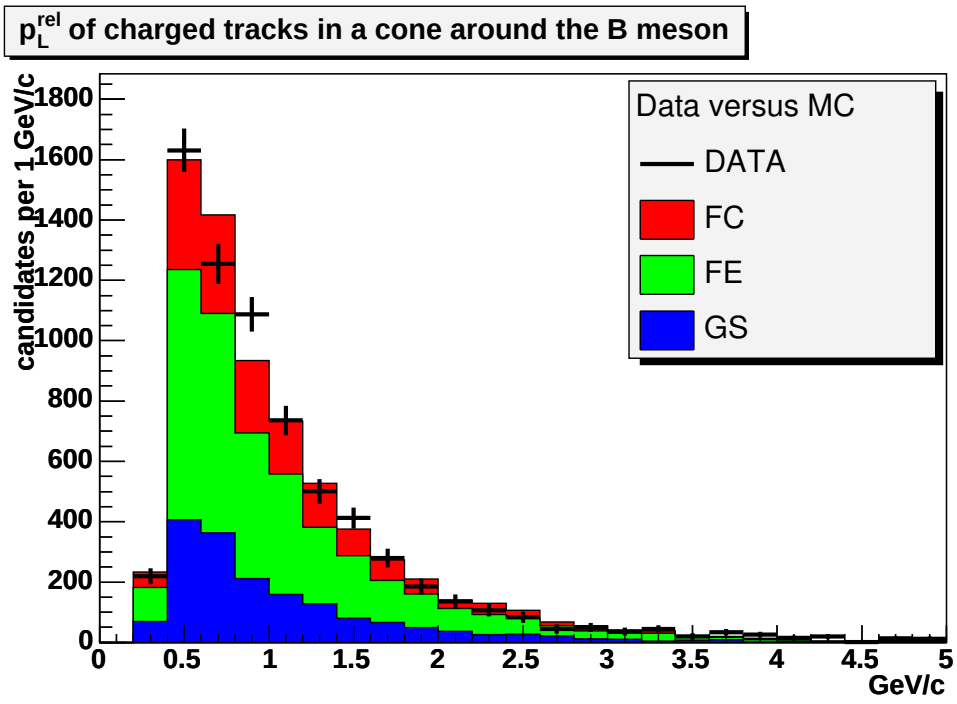

Figure 7.33: *Data and Monte Carlo comparison for p_L^{rel} of tracks in a cone around the B meson. Contributions in the Monte Carlo sample from the FC, FE and GS b quark production mechanisms are detailed. Histograms are normalized to the number of B mesons.*

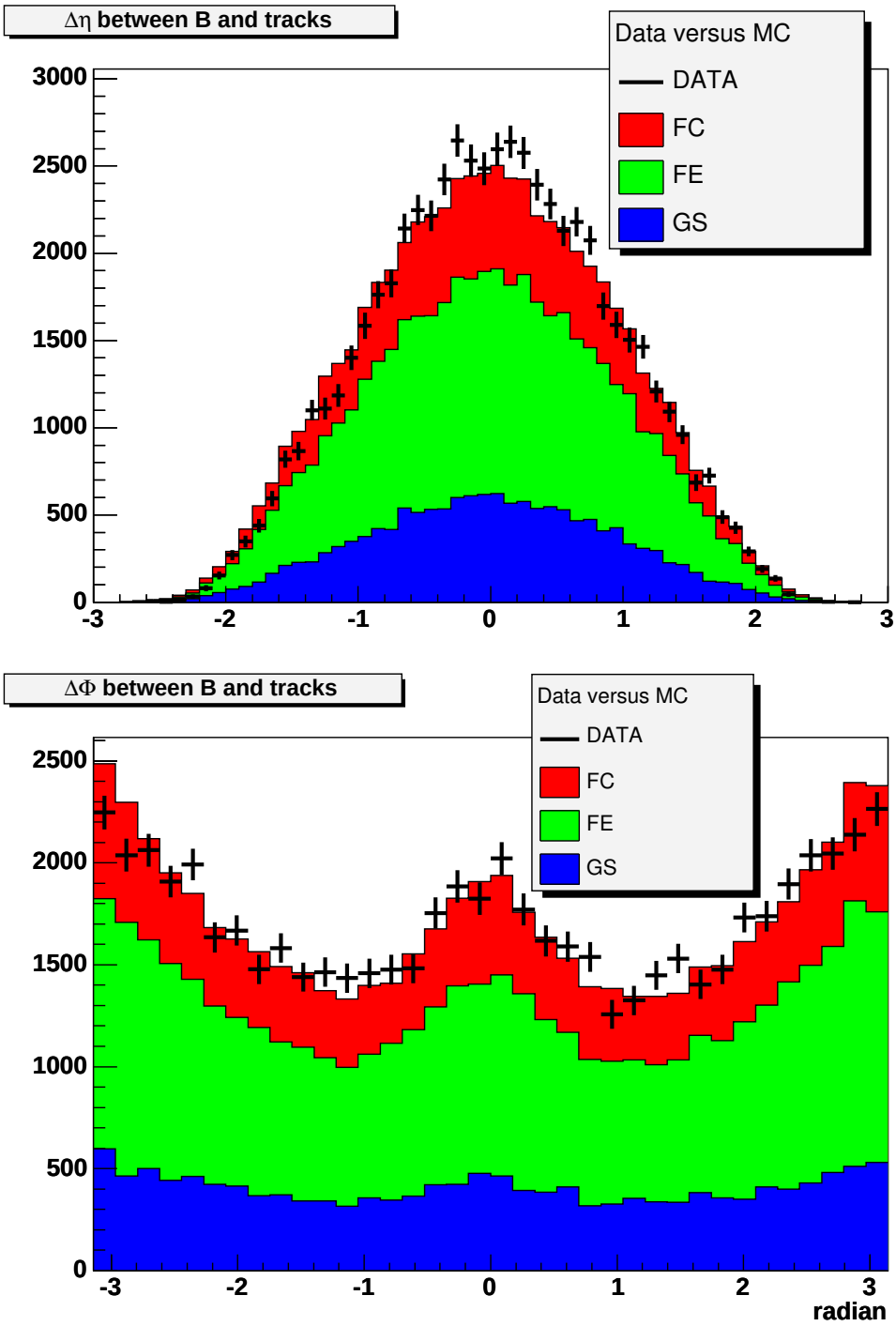

Figure 7.34: *Data and Monte Carlo comparisons for the angular distributions of all the tracks in the event with respect to the B meson direction. The upper plot shows the distribution for* $\Delta\eta$*, and the lower for* $\Delta\Phi$*. Contributions in the Monte Carlo sample from the FC, FE and GS b quark production mechanisms are detailed. Histograms are normalized to the number of B mesons.*

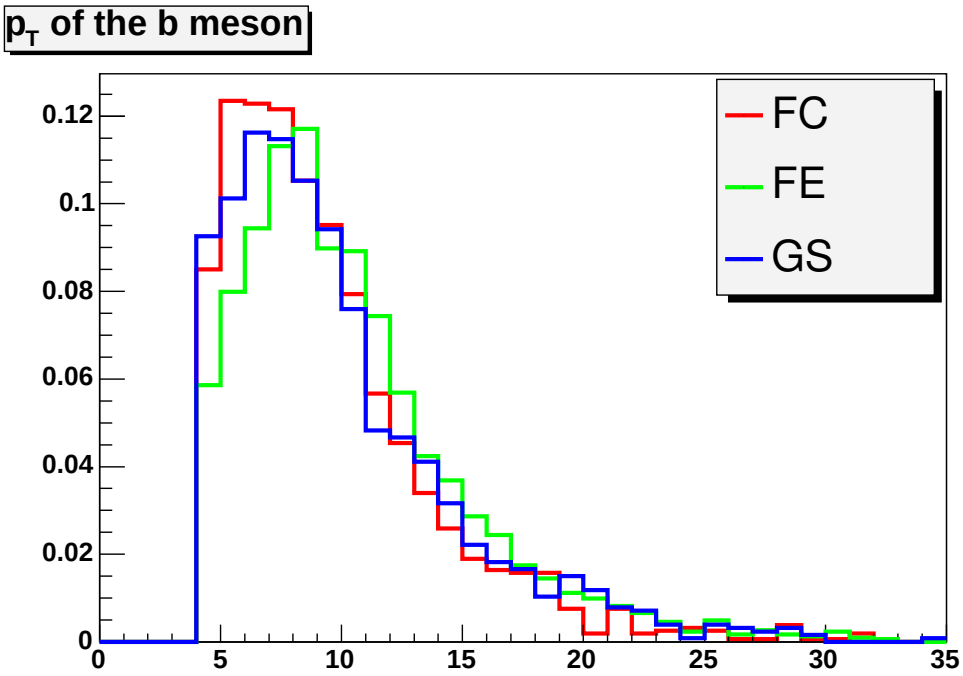

Figure 7.35: *Contributions in the Monte Carlo sample from the FC, FE and GS to the p_T distribution of the B meson. Histograms are normalized to unity. Shapes of the distributions differ between the mechanisms.*

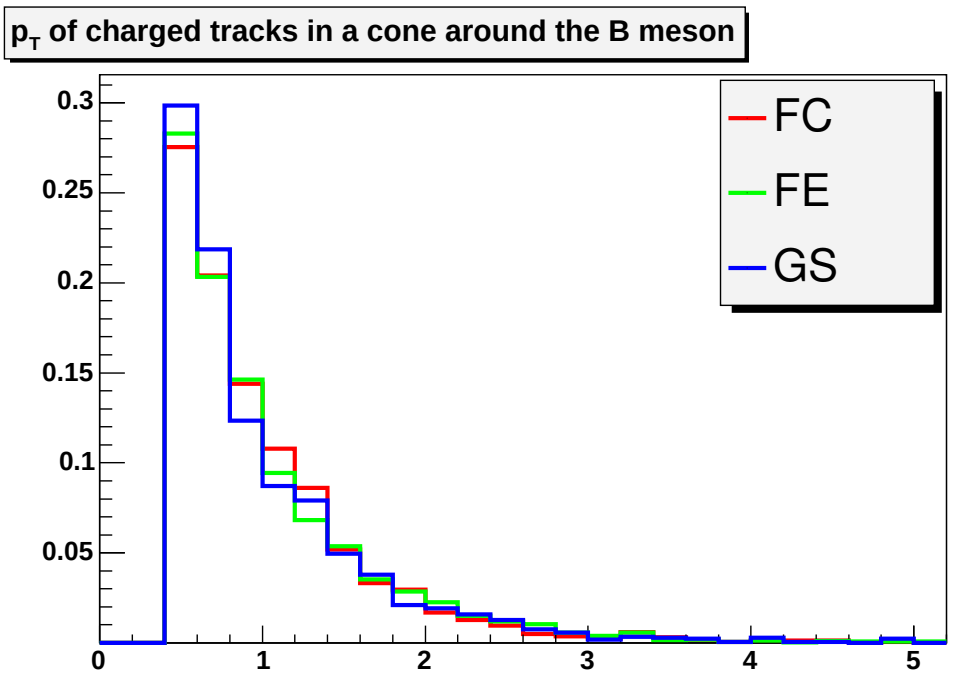

Figure 7.36: *Contributions in the Monte Carlo sample from the FC, FE and GS to the p_T distribution of tracks in a cone around the B meson. Histograms are normalized to unity. Shapes of the distributions differ between the mechanisms.*

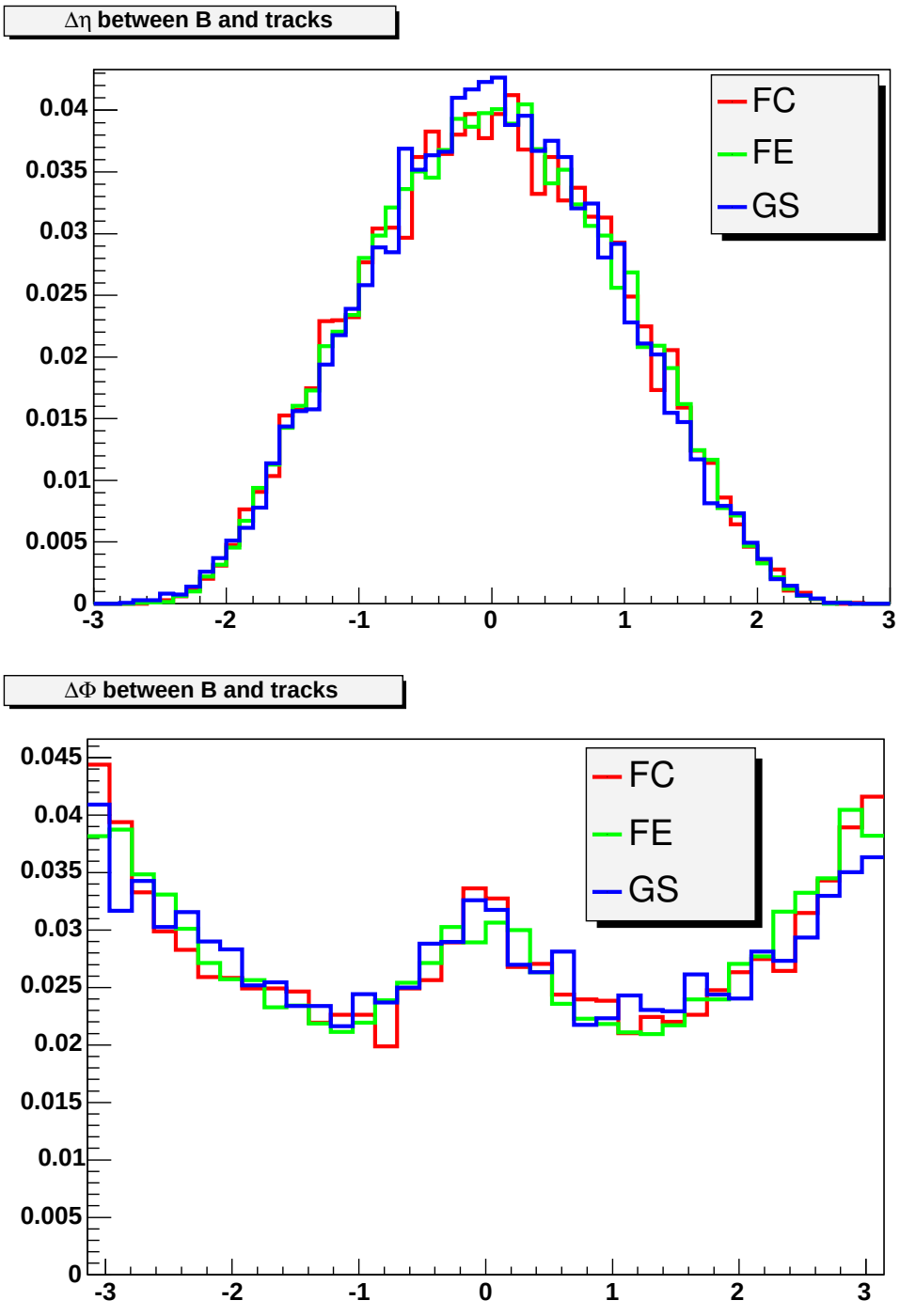

Figure 7.37: *Contributions in the Monte Carlo sample from the FC, FE and GS to the* $\Delta\eta$ *(top) and* $\Delta\phi$ *(bottom) distribution of all tracks in the event. Histograms are normalized to unity. Shapes of the distributions from the different mechanisms are similar.*

present work, a tool has been created to retrieve the necessary fragmentation information from the PYTHIA common blocks and to save it to disk at the time the sample is generated.

It is proposed to generate the Monte Carlo sample for this study using the Peterson fragmentation function with $\varepsilon = 0.006$. This choice has two reasons:

- In Pythia, the Peterson model does not affect light quark fragmentation. The Parameters used by the Lund and Bowler models in the generator, on the other hand, apply also for light quarks.
- The Peterson model, unlike the Lund and Bowler models, has a contribution not only around the peak region but has a low z tail. By reweighting events with respect to z, the tail can be canceled with a vanishing value of the weight function. Conversely, a tail cannot be created for the Lund or the Bowler model by multiplication with a weight function. A function with a low z tail is therefore necessary in order to favor or disfavor the contribution in the low z region of the fragmentation function.
- As a first approximation, the fragmentation function in CDF is expected to be similar to the one extracted in e^+e^- experiments. The Peterson parameter value of $\varepsilon = 0.006$ has been chosen because with this value the Peterson model's peak is close to that of the Lund function extracted from LEP.

These two last points are illustrated in Figure 7.7.

The z distribution of the B mesons in the reconstructed Monte Carlo sample is not identical to the distribution in the generator. One reason is that the original fragmentation function applies to the primary hadron, and not to the resulting $B^{\pm}$ mesons. Another reason is the influence of detector effects and of the analysis selection. However, both distributions can be corrected by the same weight function.

The Monte Carlo sample necessary for this study is currently being generated in CDF. The method will be applied to this sample when it is ready.

7.8 An Estimate of the b Production Cross Section

Using the data sample of fully reconstructed $B^{\pm}$ mesons, a measurement of the b quark production cross section is performed. The analysis is still under way, and therefore the results presented in this section are preliminary, and do not include a complete study of systematic uncertainties.

A similar measurement has been already done in CDF Run-II, using 37 pb^{-1} of data, in the inclusive decay channel $H_b \to J/\psi X$ [7], where H_b denotes B hadrons that decay to J/ψ.

7.8.1 Evaluation of Efficiency

The selection criteria and analysis procedure have been applied to the fully simulated msel=1 Monte Carlo sample described in 7.3. Using the distributions for $B^{\pm}$ mesons in the Monte Carlo generator before any detector and reconstruction effects, the total detection and analysis efficiency of $B^{\pm}$ has been evaluated as a function of $p_T(B)$. The evaluation has been done over

two different regions of the pseudo-rapidity: $|\eta| < 1$ and $|\eta| < 0.6$. The efficiencies for the two cases are presented in Figure 7.38.

7.8.2 The Inclusive b Quark Production Cross Section

The differential cross section for $p\overline{p} \to bX$ in each of the considered intervals in pseudo-rapidity has been computed for each $p_T(B)$ interval (i) as:

$$\frac{d\sigma_i}{dp_T} = \frac{1}{A}\frac{1}{\mathcal{L}}\frac{1}{\Delta p_T}\frac{N_i^B}{\varepsilon_i} \tag{7.3}$$

where $\mathcal{L}$ is the luminosity of the data set, N_i^B is the corresponding number of $B^{\pm}$ mesons, Δp_T is the interval width and ε_i the efficiency. A is the product of branching ratios and B^+ production fraction [71]:

$$\begin{aligned} A = BR(B^+ \to J/\psi K^+) \times BR(J/\psi \to \mu^+\mu^-) \times f(\bar{b} \to B^+) = \\ (1.00 \pm 0.04) \cdot 10^{-3} \times (5.88 \pm 0.10) \cdot 10^{-2} \times (39.7 \pm 1.0) \cdot 10^{-2} \end{aligned} \tag{7.4}$$

This cross-section has to be divided by two if one is interested in the $b\bar{b}$ production cross-section. In the following, the notation $\sigma(p\overline{p} \to bX)$ has been used to represent the sum: $\sigma(p\overline{p} \to bX) + \sigma(p\overline{p} \to \bar{b}X)$.

The measured differential cross section is shown in Figures 7.39 and 7.40, for $|\eta| < 1$ and $|\eta| < 0.6$, respectively. The distributions have been fit with the function $\frac{p_1}{(p_T)^{p_2}}$ in the region $7 < p_T(B) < 30\text{GeV/c}$. p_2 was found to be 4.13 ± 0.08 for $|\eta| < 1$ and 4.18 ± 0.09 for $|\eta| < 0.6$. In both cases, the p_T dependence of this distribution is compatible with the theoretical prediction of [91, 92] [6].

The total cross section has been evaluated by summing the contributions from all intervals in $p_T(B)$:

$$\begin{aligned} \sigma(p\overline{p} \to bX,\, p_T(B) > 4\text{GeV/c},\, |\eta| < 1.0) = 19.0 \pm 0.8\,(\text{stat.}) \pm 1.6\,(\text{syst.})\,\mu\text{b} \\ \sigma(p\overline{p} \to bX,\, p_T(B) > 4\text{GeV/c},\, |\eta| < 0.6) = 11.9 \pm 0.6\,(\text{stat.}) \pm 1.0\,(\text{syst.})\,\mu\text{b} \end{aligned} \tag{7.5}$$

The evaluation of uncertainties is explained in Sections 7.8.3 and 7.8.4.

The cross sections for $p_T(B) > 0$ are estimated using the distribution of $B^{\pm}$ mesons in the Monte Carlo sample before detector simulation and analysis procedure. The ratio of $B^{\pm}$ mesons in the region $p_T(B) > 4\text{GeV/c}$ over the total number of $B^{\pm}$ mesons was found to be 0.691 for $|\eta| < 1$ and 0.704 for $|\eta| < 0.6$. Therefore:

$$\begin{aligned} \sigma(p\overline{p} \to bX,\, |\eta| < 1.0) = 27.5 \pm 1.2\,(stat.) \pm 2.3\,(syst.)\,\mu b \\ \sigma(p\overline{p} \to bX,\, |\eta| < 0.6) = 16.9 \pm 0.8\,(stat.) \pm 1.4\,(syst.)\,\mu b \end{aligned} \tag{7.6}$$

An additional error has to be assigned to the extrapolated cross section for $p_T(B) > 0$, originating from the uncertainty on the shape of the B hadron spectrum in the Monte Carlo sample. A part of this uncertainty is due to the fragmentation function, and to the relative contributions

[6] The predicted behavior of the cross section is $d\hat{\sigma}_b/d\hat{p}_T \simeq A/\hat{p}_T^a$, Where a is expected to be approximately 4-5.

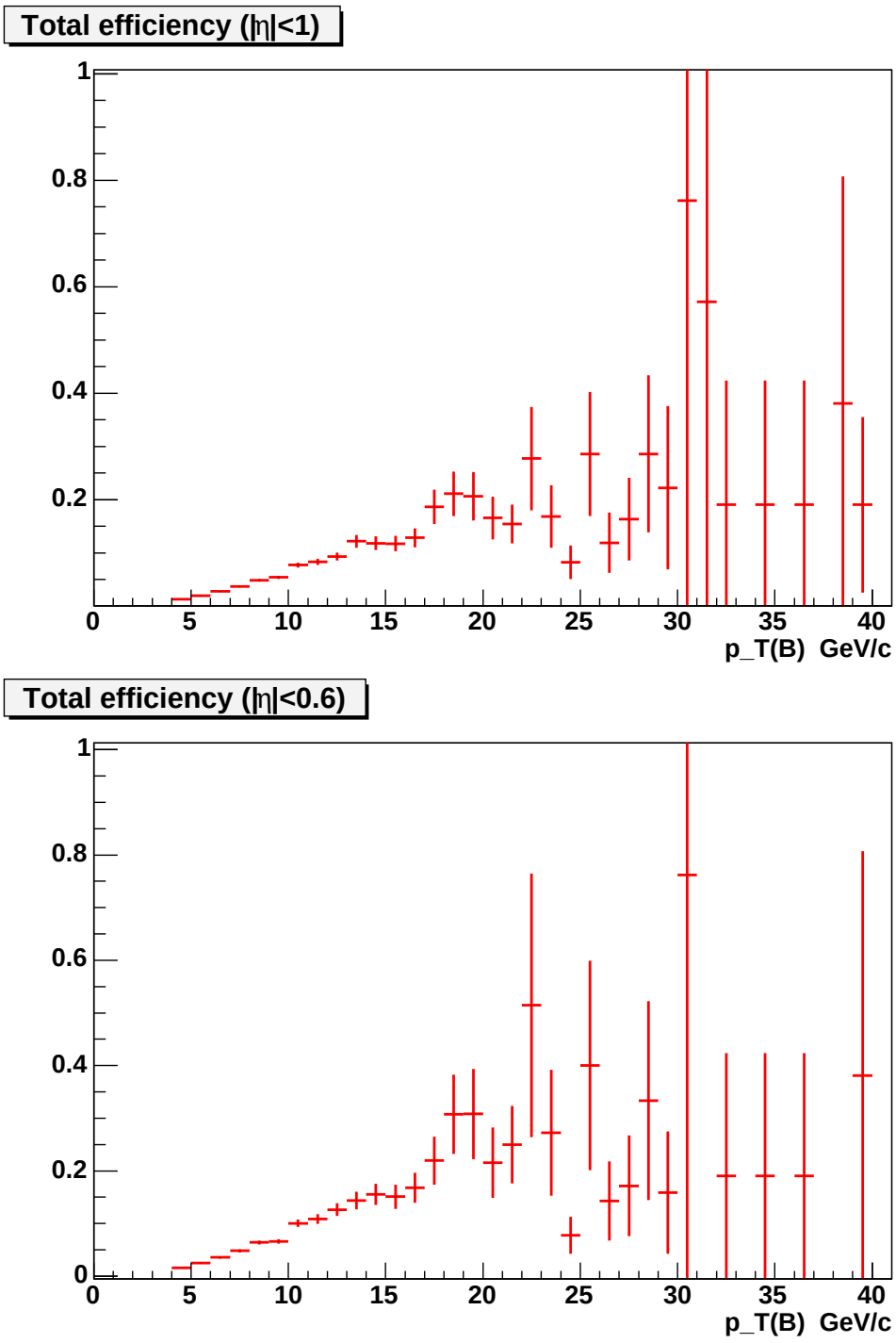

Figure 7.38: *Total efficiency of detection and analysis effects for the* $B^{\pm}$ *mesons, in two regions of pseudo-rapidity:* $|\eta| < 1$ *(top) and* $|\eta| < 0.6$ *(bottom)*

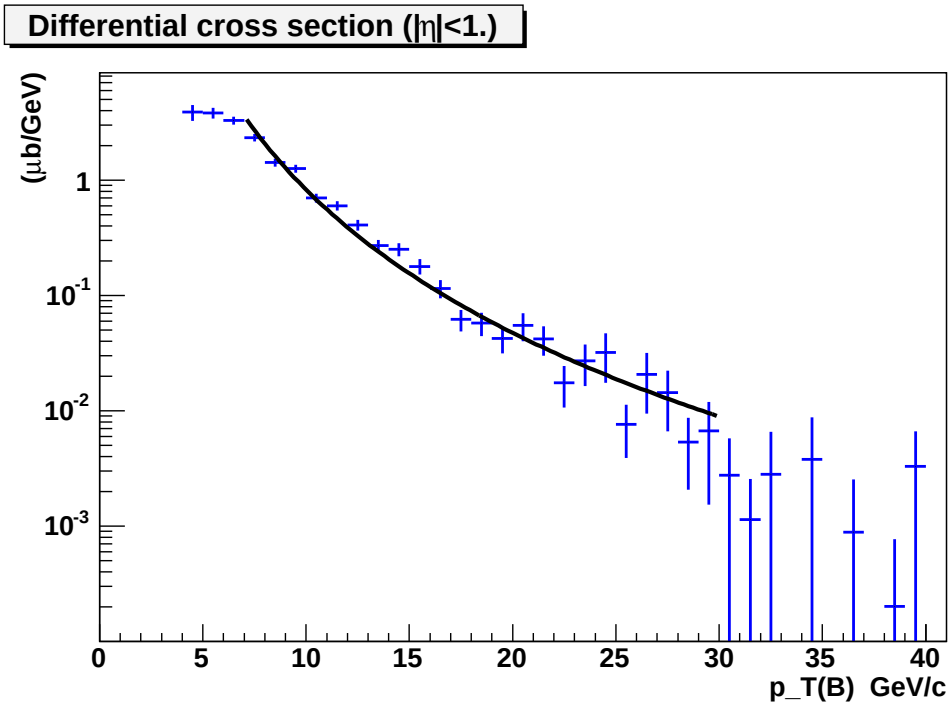

Figure 7.39: *The differential cross section for* $|\eta| < 1$.

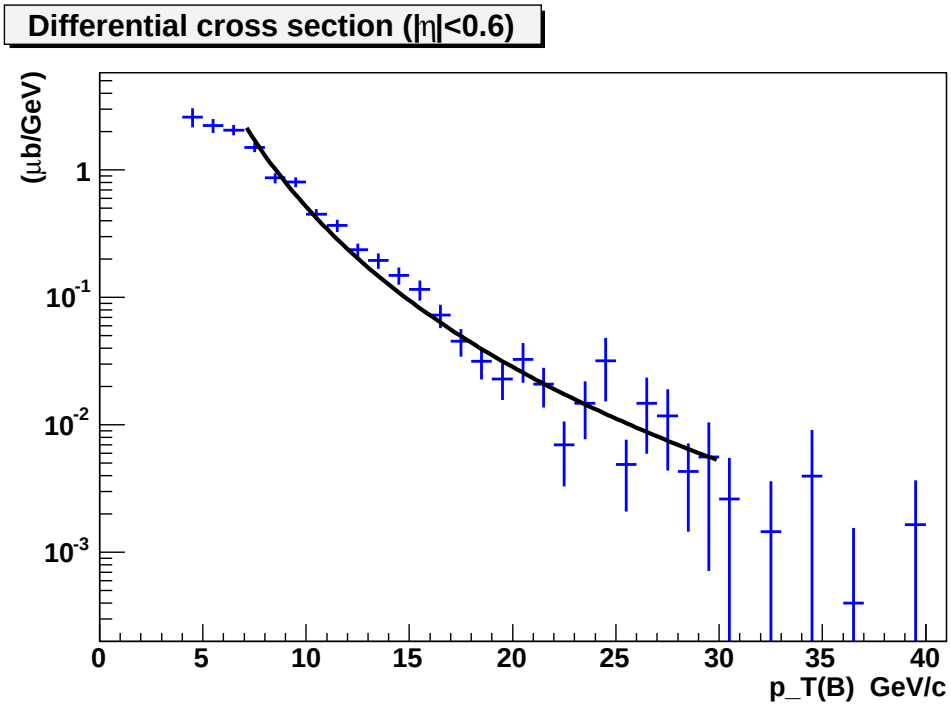

Figure 7.40: *The differential cross section for* $|\eta| < 0.6$.

from the different *b* production mechanisms. This uncertainty could be better estimated after the fragmentation fit, proposed in the present chapter, is accomplished. In order to estimate this additional source of uncertainty, the ratio of $B^{\pm}$ mesons in the region $p_T(B) > 4\text{GeV/c}$ over the total number of $B^{\pm}$ mesons has been compared between the Monte Carlo sample used for the present analysis and the theoretical prediction of [6], given in terms of pseudo rapidity [93]. This prediction uses the Kartvelishvili hadronization model, with the parameter $\alpha = 29.1$, and yields the ratio of 0.682, for $|\eta| < 0.6$. The systematic effect of extrapolation is therefore predicted to be of a few percent.

7.8.3 Statistical Error Estimation

All along the analysis process, when doing operations on histograms, as sideband subtraction, efficiency calculation etc., the statistical errors on bin contents have been transmitted. The efficiency corrected yields in each interval of the differential cross section, include all these effects. The uncertainty on the determination of the selection efficiency, coming from the statistical uncertainty due to the finite Monte Carlo statistics, has been accounted for by this procedure.

The statistical errors on the total cross sections have been obtained by adding in quadrature the statistical errors on the bins of the differential cross section.

7.8.4 Systematic Error Estimation

Systematic uncertainties from several sources have been evaluated. These sources of uncertainty have been assumed to be uncorrelated, and therefore their respective contributions have been added in quadrature to obtain the total systematic uncertainty. The study of systematic errors given here is preliminary, and is following the lines of a similar study done for the inclusive cross section measurement [7].

7.8.4.1 Luminosity

The method to measure the instantaneous luminosity in CDF, using the CLC modules, is explained in Section 4.9. These modules monitor the average number of inelastic $p\overline{p}$ interactions in each bunch crossing, which permits to derive the instantaneous luminosity using Equation (4.4). That information is read out by the main CDF data acquisition and stored in the event files.

The integrated luminosity is obtained by retrieving this information in all the files in the used data sample, and integrating the instantaneous luminosity for each file. The numbers are then summed up to give the total luminosity of the sample.

The systematic error on the luminosity measurement is dominated by three main contributions [79]:

- uncertainty on the $p\overline{p}$ cross section ($\sim 4\%$)
- the CLC acceptance ($\sim 4\%$)
- the time dependence of the CLC acceptance ($\sim 2\%$)

Adding these contributions in quadrature yields the total relative uncertainty of 6% on luminosity. From Equation (7.3) it is clear that the same relative uncertainty is induced on the differential and total cross sections.

As mentioned before, the data sample used by the present analysis corresponds to a total integrated luminosity of $333 \pm 20 pb^{-1}$.

7.8.4.2 Branching Ratios and Production Fraction

The quoted uncertainties on the branching ratios and B^+ production fraction in Equation (7.4) yield a relative error of 5.0% on the value of the product *A*. The same relative uncertainty is induced on the differential and total cross sections.

7.8.4.3 Trigger and Reconstruction Efficiencies

In the present section, the systematic effects from trigger and reconstruction efficiencies are examined. In Section 7.8.1, the msel=1 Monte Carlo sample has been used to evaluate the total efficiency. The uncertainty on the individual trigger and reconstruction efficiencies, and possible differences in their behavior in data and in the Monte Carlo sample, may therefore affect the cross section measurement.

Figures 7.41, 7.42 and 7.43 show comparisons between data and Monte Carlo of η distributions for the *B* mesons, and for the J/ψ and kaon candidates used to reconstruct them. From these plots, no significant difference between the reconstruction efficiency in data and Monte Carlo can be seen, in the region $|\eta| < 1$, considered for the cross section measurement.

The efficiency of the Level 1 dimuon trigger, used in the present analysis, has been estimated by [7] to have a constant uncertainty of $\pm 1.5\%$ over all bins in p_T of the J/ψ candidate. The total muon reconstruction efficiency, which comprises the tracking, the muon system, and matching between the muon track and "stub" [7] has been also estimated in the same reference. It was found to have an uncertainty of $\sim 2.5\%$. The track-stub matching efficiency depends weakly on p_T of the muon. Its variations between bins, of order 0.2% have been neglected here. The separate uncertainty on tracking efficiency, for muons of $p_T > 1.5$GeV, was found to be less than 1%. The same estimate has been adopted here for the kaon candidates, which are required by the present analysis to have $p_T > 1$GeV.

The sources of systematic uncertainty and their contributions are detailed in Table 7.5.

Source	Relative Systematic uncertainty
Luminosity	6%
Branching ratios	5%
L1 trigger efficiency	1.5%
Reconstruction of muons	2.5%
Reconstruction of kaons	1%
Total	8.4%

Table 7.5: *Summary of the systematic uncertainties in the* $b\bar{b}$ *production cross section measurement.*

[7] A stub is a segment resulting from a least square fit to hits in the muon drift cells.

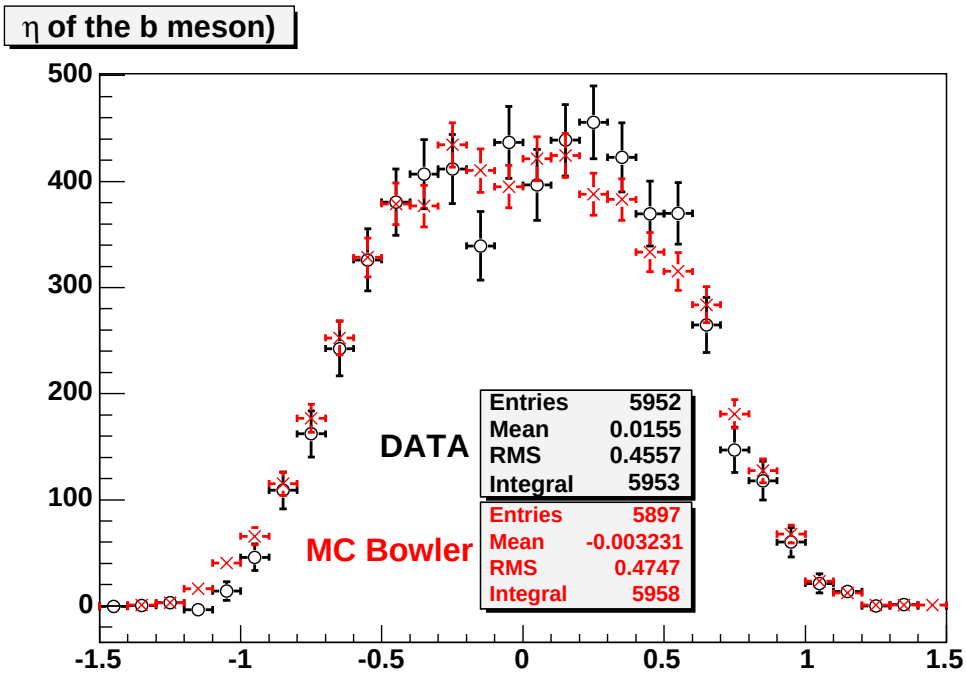

Figure 7.41: η *distributions of the reconstructed B meson, in data and in the Monte Carlo sample, used to calculate the efficiency. Histograms are normalized to the number of B mesons.*

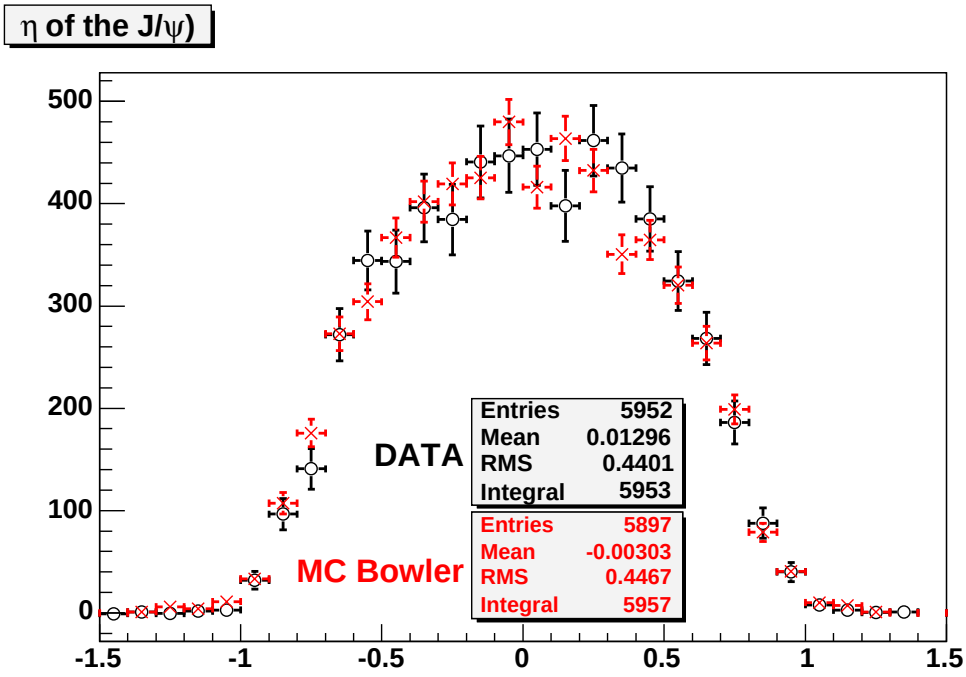

Figure 7.42: η *distributions of the* J/ψ *candidates, in data and in the Monte Carlo sample, used to calculate the efficiency. Histograms are normalized to the number of B mesons.*

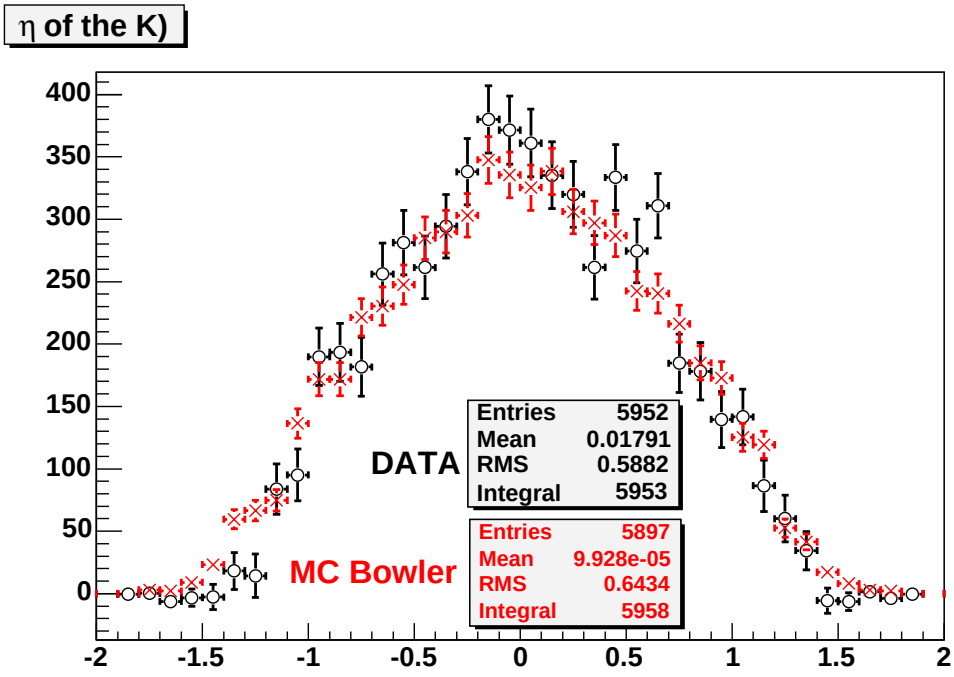

Figure 7.43: η *distributions of the kaon candidates, in data and in the Monte Carlo sample, used to calculate the efficiency. Histograms are normalized to the number of B mesons.*

7.8.5 Comparison with Other Measurements and with Theoretical Predictions

Using the same notations, an analysis of $H_b \to J/\psi X$, done in CDF has obtained:

$$\sigma(p\overline{p} \to bX, |y| < 0.6) = 35.2 \pm 0.8\,(stat.)\,^{+5.0}_{-4.6}\,(syst.)\,\mu b \tag{7.7}$$

in which y is the rapidity of the B Hadron. Using this variable, instead of the pseudo-rapidity, η, to select B hadrons we have measured:

$$\sigma(p\overline{p} \to bX, p_T(B) > 4\text{GeV/c}, |\text{y}| < 0.6) = 14.2 \pm 0.7\,(\text{stat.}) \pm 1.2\,(\text{syst.})\,\mu\text{b} \tag{7.8}$$

Results obtained with a cut on y are not equivalent to the ones with a cut on η. To illustrate this difference, Figures 7.44 and 7.45 show the p_T spectrum of B mesons that passed each one of these cuts, and the y and η spectra of the B meson with no cuts, respectively. Both figures have been obtained for generated B mesons, from the Monte Carlo sample used for the analysis. It is clear that the cut on $|y|$ selects more B mesons. The difference is particularly visible in the low p_T region, where the mass of the B meson ensures a flat distribution of y as a function of p_T.

The two analyses have been compared in Figure 7.46, which gives the variation of the differential cross-section versus the transverse momentum of the B hadron.

The b quark cross section in $p\overline{p}$ collisions at 1.96 TeV has been evaluated, using recent theoretical developments [6]. The evaluation includes a resummation of the logarithms of p_T/m_b, to

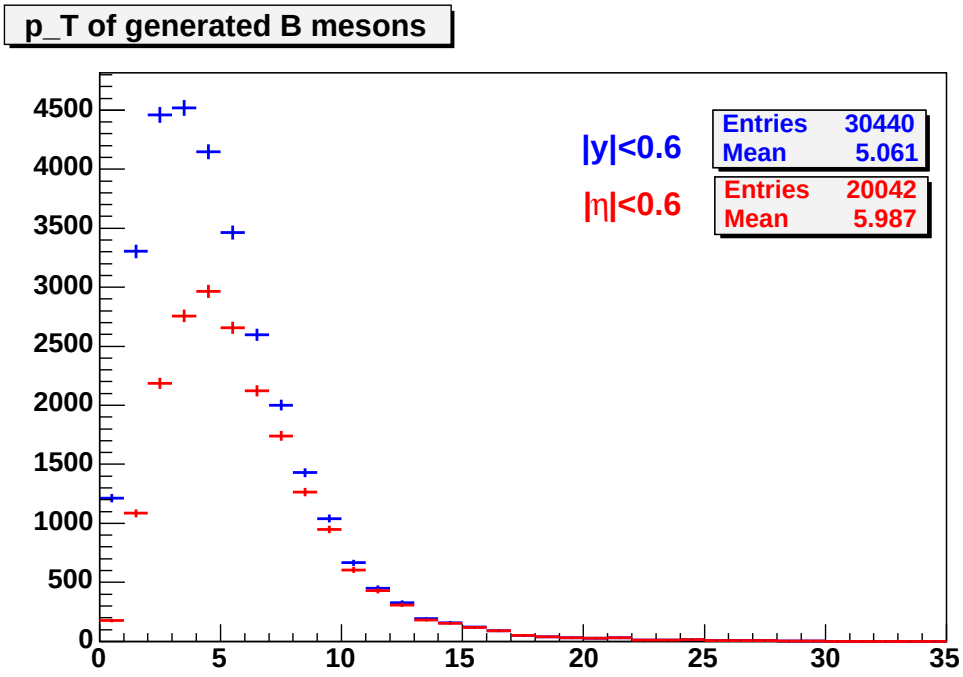

Figure 7.44: *p_T spectrum of B mesons that passed the $|y| < 0.6$ cut (blue) and the $|\eta| < 0.6$ cut (red), at the generator level of the Monte Carlo sample used for the present analysis.*

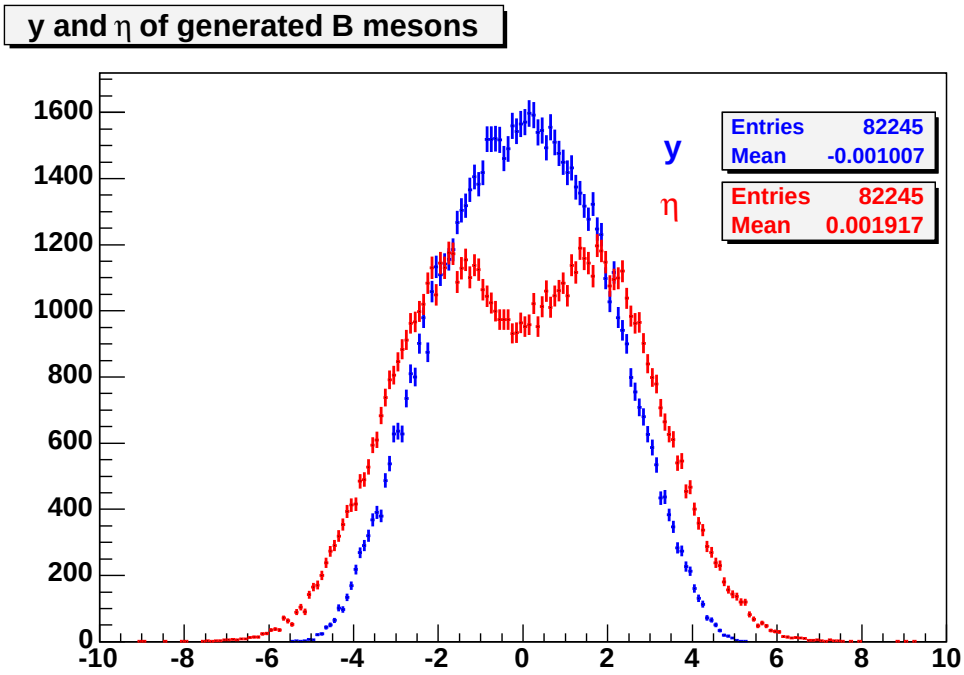

Figure 7.45: *Spectra of y (blue) and η (red) for B mesons at the generator level of the Monte Carlo sample used for the present analysis. No cuts have been applied.*

NLL accuracy, and the matching with the fixed order, exact NLO calculation for massive quark. This computation is referred to as FONLL. In order to transform this prediction, obtained at the *b* quark level, to the *B* hadron level, the Kartvelishvili fragmentation function, with the parameter $\alpha = 29.1$ has been used in [6]. The FONLL prediction has a few theoretical uncertainties, originating from the choice of scales, quark mass and the Parton Distribution Function (PDF). A comparison between the measurement, done in the present thesis work, and FONLL is presented in Figure 7.46. The curve for FONLL is shown without theoretical uncertainties.

Measurements obtained by CDF using the RunIb statistics [94], which correspond to about 400 exclusive decays $B^{\pm} \to J/\psi K^{\pm}$, have been also compared with present measurements in Figure 7.47 (note that the selection criterion is now $|y| < 1$). Present results correspond to a rate which is about 40% lower.

From present studies one can conclude that the disappearance of the excess of *b*-events in data from RunIb, as compared with QCD expectations, available at that time, comes from:

- a decrease of the experimental cross section by about 40%,
- an increase of theoretical expectations (different PDF, resummation etc.)

The effect coming from the choice of the fragmentation distribution is quite marginal. From Figure 7.13, the difference observed by changing the Peterson distribution, with parameter $\varepsilon = 0.006$, by the fragmentation distribution favored at LEP, varies from less than 1% at $P_t = 4$ GeV to about 17% at $P_t = 20$ GeV.

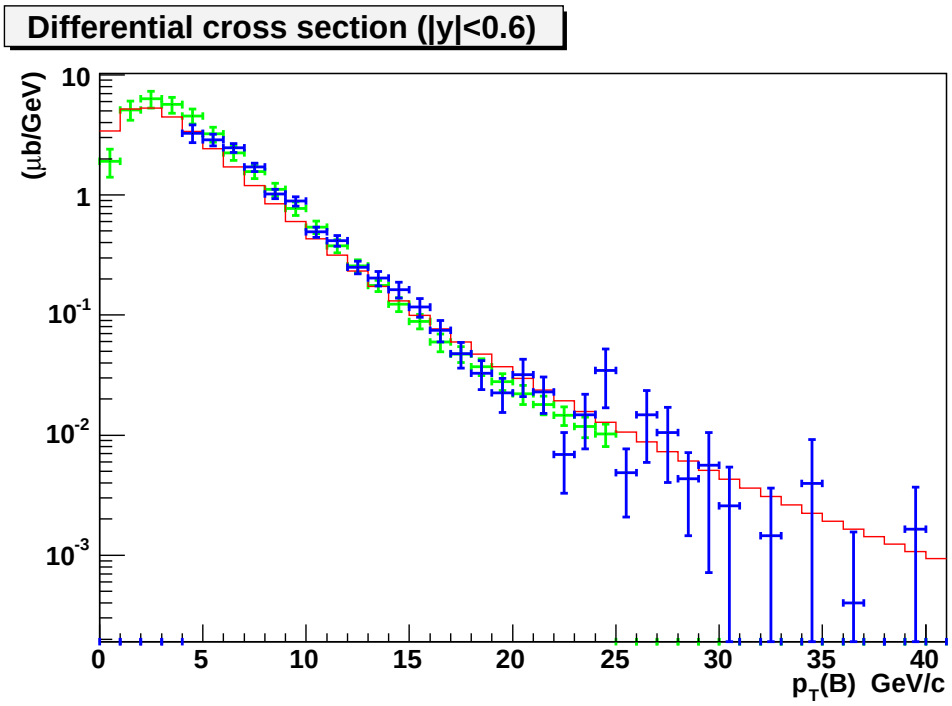

Figure 7.46: *The measured differential cross section for $|y| < 0.6$ (blue histogram with error bars), and the central value of the FONLL theoretical prediction [6] (red solid line). Measurements obtained by CDF [7] using $H_b \to J/\psi X$ final states have been also displayed (green histogram with error bars).*

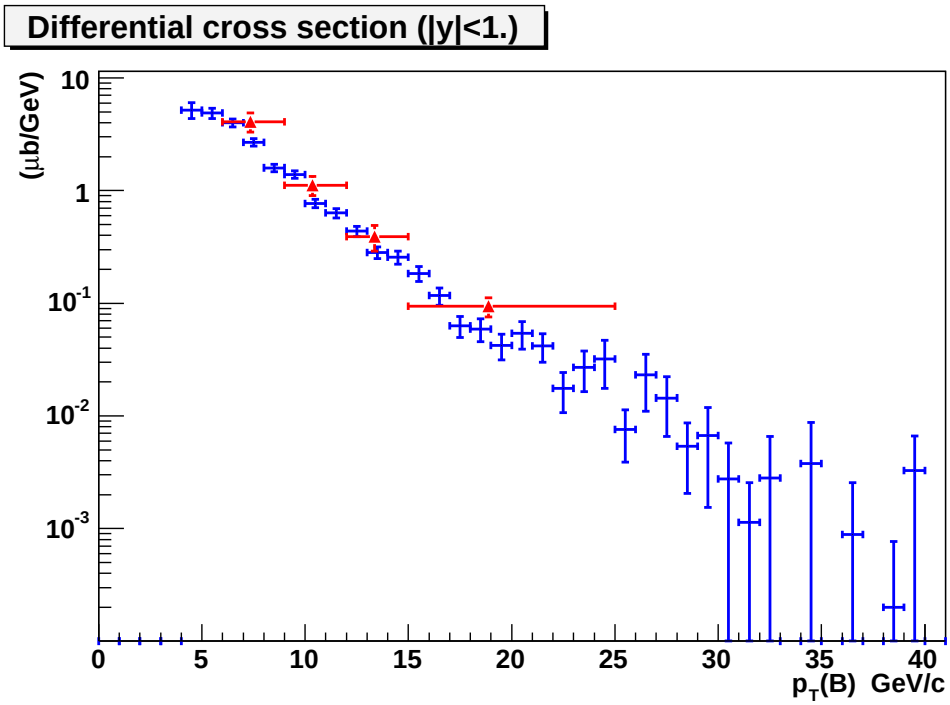

Figure 7.47: *The measured differential cross section for $|y| < 1$. obtained in the present thesis (blue histogram), compared to the similar measurement from CDF RunI [94] (red histogram).*

Chapter 8

Conclusion

The b quark fragmentation distribution has been measured, using data registered by the DELPHI experiment at the Z pole, in the years 1994-1995. The measurement made use of 176000 inclusively reconstructed B meson candidates. The errors of this measurement are dominated by systematic effects, the main one being related to the energy calibration. The distribution has been established in a nine bin histogram. Its mean value has been found to be $< x_E >= 0.704 \pm 0.001\,(stat.) \pm 0.008\,(syst.)$

The fragmentation distributions from ALEPH, OPAL and SLD, together with the result from DELPHI presented in this thesis, have been combined to give a world average b fragmentation distribution. Its mean value has been found to be $< x_E >= 0.714 \pm 0.002$

The measurement of the scaled energy (x) spectrum of the B hadron may be obtained by folding a perturbative and a non-perturbative QCD components. The x dependence of the non-perturbative part, usually described by a hadronisation model, is extracted in the present work directly from experimental data, and therefore it does not depend on a hadronization model. On the other hand, the extracted non-perturbative QCD component depends closely on the way the perturbative QCD component has been evaluated. In this thesis, the extraction method has been applied, using the perturbative component taken from the JETSET Monte Carlo event generator and using a theoretical NLL QCD calculation. The obtained distributions differ markedly from those expected from various hadronization models. The non-perturbative QCD component corresponding to a Monte Carlo generator has been found to be similar to the Lund and Bowler models. When used with the NLL QCD computation, the non-perturbative component has to be extended to the region $x > 1$.

Below $x = 0.6$, this distribution is compatible with zero indicating that most of gluon radiation is well accounted by the perturbative QCD component evaluated using the LUND parton shower Monte-Carlo or computed analytically.

The non-perturbative component, extracted in this way, is expected to be valid in a different environment than e^+e^- annihilation, as long as the perturbative QCD part is evaluated within the same framework (analytic QCD computation or a given Monte Carlo generator), and using the same values for the parameters entering into this evaluation as m_b^{pole}, $\Lambda_{QCD}^{(5)}$ or generator tuned quantities.

A comparison has been made between the non-perturbative QCD components, corresponding to the fragmentation functions measured by ALEPH, OPAL, SLD and in the present work. The distribution obtained with this thesis is quite similar with those obtained from SLD and

OPAL measurements and somewhat softer, but compatible within uncertainty, as compared with the one from ALEPH.

The conclusions from the previous parts of the work have been applied to the analysis of the *B* hadron production at CDF. Distributions of exclusively reconstructed *B* mesons, in the decay channel $B^{\pm} \to J/\psi K^{\pm}$, and for accompanying tracks has been studied. This analysis uses a sample of ~ 6000 $B^{\pm}$ candidates, from $333 pb^{-1}$ of data registered by the CDF experiment.

Effects of the fragmentation function on these distributions, in the framework of the PYTHIA event generator, have been examined, together with the effect of different $b\bar{b}$ production mechanisms in the generator. The distributions in a fully reconstructed Monte Carlo sample have been compared to data, and the agreement has been found to be reasonable. The analysis is ongoing, and the goal will be to fit the fragmentation function parameters and/or the relative contributions from different production mechanisms to improve the agreement between data and Monte Carlo.

A measurement of the *b* quark production cross section has been established using the same data. The analysis is still under way, and therefore the result is preliminary. It has been found:

$$\begin{aligned}
\sigma(p\bar{p} \to bX,\, p_T(B) > 4\text{GeV/c},\, |\eta| < 1.0) &= 19.0 \pm 0.8\,(\text{stat.}) \pm 1.6\,(\text{syst.})\,\mu\text{b} \\
\sigma(p\bar{p} \to bX,\, p_T(B) > 4\text{GeV/c},\, |\eta| < 0.6) &= 11.9 \pm 0.6\,(\text{stat.}) \pm 1.0\,(\text{syst.})\,\mu\text{b}
\end{aligned} \tag{8.1}$$

and

$$\begin{aligned}
\sigma(p\bar{p} \to bX,\, p_T(B) > 4\text{GeV/c},\, |\text{y}| < 1.0) &= 22.8 \pm 1.1\,(\text{stat.}) \pm 1.9\,(\text{syst.})\,\mu\text{b} \\
\sigma(p\bar{p} \to bX,\, p_T(B) > 4\text{GeV/c},\, |\text{y}| < 0.6) &= 14.2 \pm 0.7\,(\text{stat.}) \pm 1.2\,(\text{syst.})\,\mu\text{b}
\end{aligned} \tag{8.2}$$

These measurements are compatible with another recent analysis done at CDF, using a different *B* hadron decay channel, and they are somewhat lower than results from RunIb.

At present, there is no evidence for an excess of *B* hadrons in data collected at CDF, relative to what is expected from QCD.

Appendix A

The Mellin Transformation

A complete description of the Mellin transformation and its properties may be found in [95]. In the present Appendix, the aim is only to describe a few computations involving this transformation used in this document.

The use of the Mellin transformation is appropriate for computing integral equations encountered in QCD. In particular, it is useful when dealing with fragmentation functions; the Mellin transformation of the folding product of the perturbative and non-perturbative distributions is the simple product of the two Mellin-transformed distributions. Namely, the product:

$$\mathcal{D}_{predicted}(x) = \int_0^\infty \mathcal{D}_{pert.}(z) \times \mathcal{D}_{non-pert.}(\frac{x}{z})\,\frac{dz}{z} \tag{A.1}$$

for the x variable, becomes, after applying the Mellin Transformation:

$$\tilde{\mathcal{D}}_{predicted}(N) = \tilde{\mathcal{D}}_{pert.}(N) \times \tilde{\mathcal{D}}_{non-pert.}(N) \tag{A.2}$$

The Mellin transformation of a distribution $\mathcal{D}(x)$ is defined as:

$$\tilde{\mathcal{D}}(N) = \int_0^\infty dx\, x^{N-1}\, \mathcal{D}(x) \tag{A.3}$$

where N is a complex variable. For integer values of $N \geq 2$, the values of $\tilde{\mathcal{D}}(N)$ correspond to the moments of the initial x distribution [1].

The N-space representation, $\tilde{\mathcal{D}}(N)$, can be transformed back to the x space, applying the inverse Mellin transformation:

$$\mathcal{D}(x) = \frac{1}{2\pi i} \oint dN\, \tilde{\mathcal{D}}(N) x^{-N} \tag{A.4}$$

in which the integral runs over a closed contour in the complex N-plane. The integration contour can be taken as two half-lines, symmetric with respect to the x-axis (one in the upper half and the other in the lower half of the complex plane) and joining at their intersection point with this axis. The contour is closed by a circular arc at infinity. In practice, the integration is done only over the line in the upper half of the complex plane (see Figure A.1). Indeed, the fact that $\mathcal{D}(x)$ is a real function implies that the imaginary parts of the integral (A.4), compensates between

[1] By definition $\tilde{\mathcal{D}}(1) = 1$ reflects the normalization of $\mathcal{D}(x)$.

the upper and lower halves of the complex plane. The real parts, on the other hand, give equal contributions for the two lines of integration. The integral is over the variable z, along the line. When the angle Φ is larger than $\pi/2$, the contour is referred to as L_-, the other case is referred to as L_+. The choice of L_- or L_+ depends on the function and on the x value, so as the contribution of the circular arc vanishes at infinity (see below).

The inversion formula is therefore calculated as follows:

$$\frac{1}{\pi}\int_0^\infty Im\left(e^{i\phi}x^{-c-ze^{i\phi}}\mathcal{D}\left(c+ze^{i\phi}\right)\right)dz \tag{A.5}$$

where c is the intersection point of the contour with the real axis.

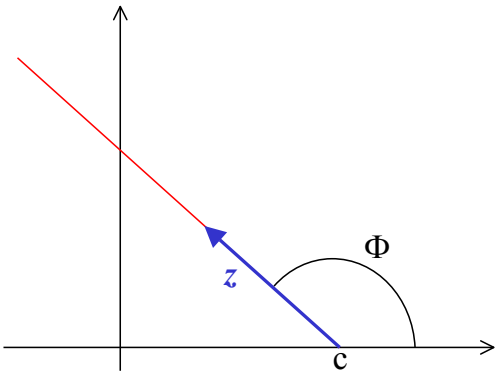

Figure A.1: *Integration contour for the inverse Mellin transformation.*

In Chapter 6, the parametrization:

$$\mathcal{D}(x) = p_0 \times \left[p_1 x^{p_2}(1-x)^{p3} + (1-p_1)x^{p_4}(1-x)^{p5}\right] \tag{A.6}$$

where p_0 is a normalization coefficient, has been used to parametrize the measured fragmentation functions. Besides being able to fit the experimentally measured fragmentation functions for bottom and charm quarks, this parametrization has an analytically calculable Mellin transformation. For the function $(1-x)^\beta$, the Mellin transformation is:

$$\frac{\Gamma(\beta+1)\Gamma(N)}{\Gamma(\beta+N+1)} \tag{A.7}$$

It is straightforward to check the following property:

$$x^\alpha \mathcal{D}(x) \rightarrow \tilde{\mathcal{D}}(N+\alpha) \tag{A.8}$$

(where $\tilde{\mathcal{D}}(N)$ denotes as above the transform of $\mathcal{D}(x)$). This simple property allows to calculate the Mellin transformation of $x^\alpha(1-x)^\beta$:

$$\frac{\Gamma(\beta+1)\Gamma(N+\alpha)}{\Gamma(\alpha+\beta+N+1)} \tag{A.9}$$

Equation (A.6) is therefore transformed as:

$$\tilde{\mathcal{D}}(N) = p_0 \left[p_1 \frac{\Gamma(p_2+N)}{\Gamma(p_2+p_3+N+1)} + (1-p_1) \frac{\Gamma(p_4+N)}{\Gamma(p_4+p_5+N+1)} \right] \tag{A.10}$$

Fitting a fragmentation function with the parametrization of (A.6), one gets directly the moment-space representation, by using (A.10) with the fitted parameters.

The exact Mellin transformation of the Kartvelishvili parametrization is calculated in a similar way as above. For the Peterson and Collin-Spiller models, only the values of moments, corresponding to positive integer values of N, are analytically calculable. Moments haven't been calculated analytically for the Lund and Bowler models.

In Chapter 6, the Mellin transformation of functions defined piecewise over $0 < x < 1$ has also been used. Namely, histograms and a fitted cubic spline have been considered to model the measured fragmentation function. Here, to illustrate the problem of Mellin inversion of a function defined separately for different intervals, polynomials of order J, defined piecewise for $0 < x < 1$ are considered. For simplicity, the discussion is done for two intervals, as shown in Figure A.2. It can be then easily generalized to a larger number of intervals.

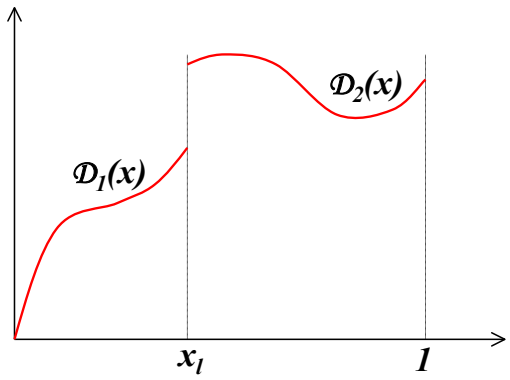

Figure A.2: Piecewise defined function for $0 < x < 1$. In each interval the function is taken as a polynomial of order J.

The distribution $\mathcal{D}(x)$ is defined as:

$$\mathcal{D}(x) = \begin{cases} \sum_{j=0}^{J} A_j x^j = D_1(x) & 0 < x < x_l \\ \sum_{j=0}^{J} B_j x^j = D_2(x) & x_l < x < 1 \end{cases} \tag{A.11}$$

The Mellin transformation of $\mathcal{D}(x)$ is obtained by two simple integrals, and rearrangement of the terms:

$$\begin{aligned} \tilde{\mathcal{D}}(N) &= \int_0^{x_l} dx x^{N-1} D_1(x) + \int_{x_l}^{1} dx x^{N-1} D_1(x) = \\ &= \sum_{j=0}^{J} \left[\frac{A_j - B_j}{N+j} x^{j+N} + \frac{B_j}{N+j} \right] \end{aligned} \tag{A.12}$$

In order to take the inverse Mellin transformation, and get back to the polynomials in the x space, it is necessary to consider separately two contributions to the integral for $\mathcal{D}(x)$:

$$\mathcal{D}(x) = I_1(x) + I_2(x) \tag{A.13}$$

where:

$$\begin{aligned} I_1(x) &= \sum_{j=0}^{J} \oint dN \left(\frac{x}{x_l}\right)^{-N} x_l^j \left(\frac{A_j - B_j}{N+j}\right) \\ I_2(x) &= \sum_{j=0}^{J} \oint dN\, x^{-N} \left(\frac{B_j}{N+j}\right) \end{aligned} \tag{A.14}$$

This separation is necessary, because different contours (namely either L_- or L_+) must be considered for I_1 and I_2, with respect to the value of x. In order to get a convergent (and computable) integral, the contribution on the circular arc of the terms $\left(\frac{x}{x_l}\right)^{-N}$ and x^{-N} in I_1 and I_2, respectively, should vanish at infinity. Therefore, the contours must be taken as detailed in Table A.1.

	$0 < x < x_l$	$x_l < x < 1$	$x > 1$
I_1	L_-	L_+	L_+
I_2	L_-	L_-	L_+

Table A.1: *Integration contours for the inverse Mellin transformation for a function defined piecewise in two intervals over* $0 < x < 1$.

For the present case, when $x > 1$, the integration yields 0, because there are no poles in the right side of the plane. No poles are therefore included in the L_+ contour for both integrals I_1 and I_2.

When the number of intervals is larger than two, a similar but more cumbersome reasoning, involving the splitting of the inversion formula for $\mathcal{D}(x)$ and an appropriate choice of the contour, has to be done.

On the contrary, what has been presented for two intervals trivially reduced to the case when the function is perfectly smooth (indefinitely differentiable) on $0 < x < 1$; then only one integral intervenes and the integration contour is taken as L_+ for $x < 1$ and L_- for $x > 1$.

Appendix B

Fitting Histograms of Singular Error Matrices

Measurements of the b-fragmentation distribution have been published by ALEPH [31], OPAL [32] and SLD [33] in a binned form, after unfolding the experimental energy resolution. Values in the bins are correlated and, as the bin width is smaller than the resolution the error matrix is singular. The effective number of degrees of freedom in these analyses is smaller than the number of bins in the published distributions. Using directly the published results to fit a smooth function to the distributions is impossible because the fit requires the inversion of the error matrix. Doing such a fit was necessary for the extraction of the non-perturbative component of the fragmentation function as explained in Chapter 6, and therefore, a method had to be employed to overcome this issue. This method consisted in reducing the number of degrees of freedom of the error matrix by taking its eigenvalues and considering only the few largest ones.

Let V be the covariance matrix of the N bins of a histogram. v_i and λ_i ($1 \leq i \leq N$) are the eigenvectors of V and their corresponding eigenvalues. v_i are column vectors such that $|\mathrm{v}_i| = 1$. T is a matrix whose columns are the vectors v_i.

$$T = \left(\begin{bmatrix} \\ \mathrm{v}_1 \\ \\ \end{bmatrix} \dots \begin{bmatrix} \\ \mathrm{v}_N \\ \\ \end{bmatrix} \right). \tag{B.1}$$

T is in fact the unitary matrix that may be used to transform V to the basis of its eigenvectors, where it is diagonal. In this basis the matrix elements of V are simply its eigenvalues.

$$V' = T^t \cdot V \cdot T = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{pmatrix}. \tag{B.2}$$

A vector δ can also be transformed by T:

$$\delta' = T^t \cdot \delta \tag{B.3}$$

The aim is to fit the binned distributions with a smooth function $f(x;\mathbf{a})$, where $\mathbf{a}$ is a vector of parameters. Hence, each bin content has been compared to the integral of the function

within the bin boundaries. The vector of differences between the bin contents and the function predictions is:

$$\delta = \begin{pmatrix} (bin\ content)_1 - \int\limits_{bin_1} f(x;\mathbf{a})dx \\ \vdots \\ (bin\ content)_N - \int\limits_{bin_N} f(x;\mathbf{a})dx \end{pmatrix}. \tag{B.4}$$

The χ^2 to minimize is calculated using V and δ. In the original basis it is expressed in matrix notation as:

$$\chi^2 = \delta^t \cdot V^{-1} \cdot \delta \tag{B.5}$$

In the basis of eigenvectors, where the correlations are vanishing it may be written as the simple sum:

$$\chi^2 = \sum_{i=1}^{N} \frac{(\delta_i')^2}{\lambda_i} \tag{B.6}$$

The χ^2 is a scalar, and therefore it does not depend on the basis. However, this is easy to verify by using the unitarity of the transformation matrix T.

The effective number of degrees of freedom of the analysis is the number of eigenvalues of v that are significantly positive. Therefore we want to cut off all the negative eigenvalues and those that are positive but not significantly different than 0. If the eigenvalues $\lambda_{1..N}$ are sorted in an descending order, and if we take the n largest eigenvalues, the χ^2 to minimize becomes:

$$\chi^2 = \sum_{i=1}^{n<N} \frac{(\delta_i')^2}{\lambda_i} \tag{B.7}$$

A special attention has been paid to the following point. In some cases we had to transform the error matrix. This was necessary in order to move from an error on the mean value in a bin as published by OPAL [32] to the error on the bin content that has been used for the fit. A transformation was also needed as a part of the variable change $x_E \longrightarrow x_p$ as explained in Section 6.1. If one calculates the full rigorous χ^2 like in Equations (B.5) or (B.5) it is strictly equivalent to transform the error matrix before or after calculating the eigenvalues. On the other hand, once degrees of freedom are cut from the problem as described above, one can't go back to the original situation by applying a linear operator. Therefore all the transformations have to be done on the eigenvalues of the original published error matrix, before cutting some of them off.

Knowing which one of the positive eigenvalues is compatible with 0 consists in knowing the error on each eigenvalue. This requires the knowledge of the error on each matrix element of V. In the present case we did not have this information, and therefore had to determine the effective number of degrees of freedom approximately. In fact, we took the maximum number of eigenvalues that yield a χ^2 per degree of freedom ≈ 1 in the fit. This is based on the argument that if the fitted function agrees with the actual values one expects the $\chi^2/d.o.f$ to be ≈ 1, because it is the mean value of the χ^2 distribution. This provides an indication that the errors are neither underestimated nor overestimated [96]. This choice ensures that the errors on the fitted histogram's bins are propagated correctly to the fitted parameters.

As illustrated in Figure B.1 for the case of OPAL, beyond a certain number of eigenvalues, the χ^2 of the fit increases steeply due to the smallness of the eigenvalues of V that represent underestimated errors in the diagonal base [1].

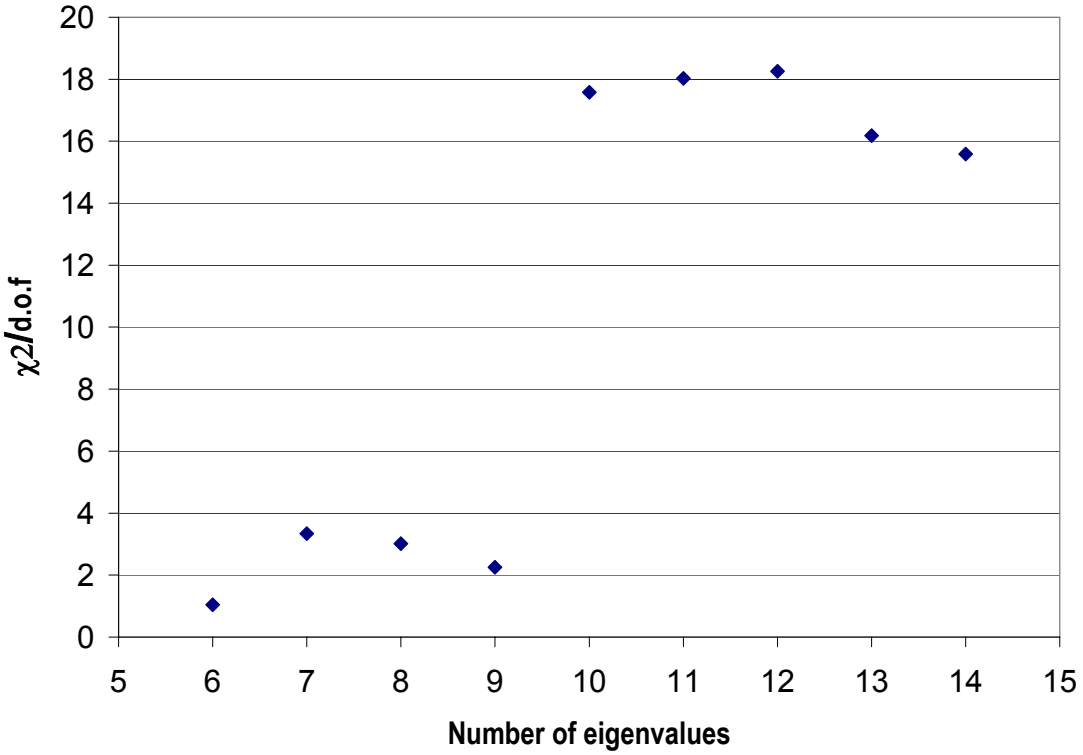

Figure B.1: *$\chi^2/d.o.f$ of the fit as a function of number of eigenvalues for OPAL's fragmentation function.*

The number of effective number of degrees of freedom has been estimated to be 7-8 for ALEPH's fragmentation function, where the number of bins is 19 and the error matrix has 13 positive eigenvalues. In the case of OPAL, where the number of the number of bins is 20, and the error matrix has 14 positive eigenvalues, the number of effective degrees of freedom has been estimated to be 6. However, it has been verified that using a larger number of degrees of freedom the parameters of the fitted function do not change significantly, as shown in Section 6.5.

[1] As mentioned in Section 6.2, the function used to fit the histograms is the one in (6.6), that has 5 parameters. The number of degrees of freedom is therefore $n - 5$, where n is the number of eigenvalues.

Acknowledgments

I would like to begin by thanking my two co-advisors, Patrick Roudeau and Aurore Savoy-Navarro. Patrick initially introduced me to the subject of this dissertation, and his guidance continues to be invaluable for my work.

I would also like to thank the directors of both LAL and LPNHE for hosting me in their laboratories. I'm particularly grateful to Jean-Eudes Augustin, director of LPNHE, for the interest he showed in my research and for his advice throughout the course of my work.

When I arrived at CDF in 2002, Henry Frisch warmly welcomed me to the collaboration by introducing me to my fellow researchers and sharing his experience with me. Henry's opinion and advice continue to be important for me. Many fellow CDF researchers generously explained the technical aspects of the experiment and computing tools and also helped me set up the framework for my analysis. Countless brainstorming sessions with some of these colleagues helped advance my analysis. Among these people I would like to thank Franco Bedeschi, Mary Bishay, Stefano Giagu, Guillelmo Gomez-Ceballos, Andreas Korn, Ilya Kravachenko, Nuno Leonardo (who kindly took the time to introduce me to the CDF software), Christophe Paus, Jonatan Piedra and Marjory Shapiro. My close collaboration with Petar Maksimovic and Jennifer Pursley was not only a pleasure, but was essential to the success of my analysis. Their cooperation and understanding during the writing of this dissertation went above and beyond the call of duty. So many colleagues in CDF were helpful for the success of my research that I cannot thank them all by name, but I deeply appreciate their knowledge and cooperative spirit.

My CDF-LPNHE colleagues Stéphane Tourneur and Antonio Sidoti with whom I shared an office were sources of mutual support and encouragement.

My collaboration with members of the other DELPHI institution also working on b- fragmentation, the group from the Karlsruhe Institute for Experimental Nuclear Physics, and especially Christian Weiser, has been a very productive one. I had the pleasure of completing my research alongside my fellow graduate student Arantza Oyanguren Campos who was always available to help when necessary.

Theoretical works on b-fragmentation were crucial to the success of my analysis. Experts and authors of theoretical articles in the domain were very helpful and cooperative. I would like to thank Stefano Catani and Michel Fontannaz for their help at the beginning of my work. I would also like to thank Yuri Dokshitzer, Einan Gardi and Gevin Salam for many fruitful discussions. The collaboration with Matteo Cacciari was an enormous pleasure. His help in understanding the physical and technical aspects of b-fragmentation theory played a key role in my research.

During the preparation of my DEA diploma ("Champs, Particules Matière"), I had the good fortune to take Achille Stocchi and Fabian Zomer's course on particle physics. Their enthusiastic introduction to the subject was very important as I began my career. As an active DELPHI

group member at LAL, Achille was permanently involved in my analysis, and his advice was essential to its success. Fabian became my godparent at LAL.

Initially I came to physics through the interface between biology and physics, a very different domain from particle physics. During my master's degree (maîtrise and DEA) I worked with Annick Lesne and Jean-Marc Victor of the Laboratoire du Physique Théorique des Liquides (LPTL) of Jussieu on modelling the structure and mechanical properties of the Chromatine fiber. It is thanks to Annick and Jean-Marc that I realized I did, in fact, want to do research in physics. Because ultimately I was more drawn to particle physics I chose not to pursue the Chromatine fiber research, but the work I did with Annick and Jean-Marc remains a meaningful chapter in my experience as a physicist.

I would like to thank the referees of this dissertation, Lorenzo Foa and André Rougé, as well as the members of the jury, Jean-Eudes Augustin, Matteo Cacciari, Henry Frisch, and Jean-Bernard Zuber.

This work was supported by EEC RTN contract HPRNCT-00292-2002. My co-advisor Aurore Savoy-Navarro played a key role in finding this funding which made it possible to complete my research. I began my thesis before the RTN was established. I was able to benefit from the financial support of CERN, LAL and LPNHE, for which I would like to thank Philippe Charpentier (CERN), Serge Jullian (LAL) and Jean-Eudes Augustin (LPNHE).

Finally, this dissertation would not have been able to be completed without the support, love and understanding of my family and friends, to whom I owe deep thanks. I especially want to thank Roland Bost and Naomi Davidson, who know how grateful I am.

Bibliography

[1] B. Foster, *Electron-Positron annihilation physics*, Ed. Adam Hilger (1990), Bristol, New York.

[2] P. Nason, S. Dawson, and R. K. Ellis, Nucl. Phys. **B303** (1988) 607.

[3] P. Nason, S. Dawson, and R. K. Ellis, Nucl. Phys. **B327** (1989) 49, and erratum-ibid. B**335** (1989) 260.

[4] M. Cacciari, M. Greco and P. Nason, JHEP **9805**, 007 (1998).

[5] M. Cacciari and P. Nason, Phys. Rev. Lett. **89** (2002) 122003 [arXiv:hep-ph/0204025].

[6] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, JHEP **0407** (2004) 033 [arXiv:hep-ph/0312132].

[7] CDF Note 6285;
C. Chen, [CDF Collaboration], presentation at "Beauty 2003"
`http://www-cdf.fnal.gov/physics/new/bottom/030904.blessed-bxsec-jpsi`;
M. Bishai [CDF Collaboration], presentation at Fermilab, Dec 5. 2003, `http://www-cdf.fnal.gov/∼bishai/papers/wandc.pdf`;
CDF Note 7037.

[8] S. Chekanov *et al.*, ZEUS Collaboration [arXiv:hep-ex/0312057].

[9] K. Lannon, K. Pitts, "Bottom Quark Production Using PYTHIA and HERWIG", CDF Note 6253 (2003).

[10] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and Collider Physics* (Cambridge University Press, Cambridge, 1996) and references therein.

[11] P. Nason and C. Oleari, Nucl. Phys. **B565** (2000) 245.

[12] B. Mele and P. Nason, Nucl. Phys. **B361** (1991) 626.

[13] Y. L. Dokshitzer, V. A. Khoze and S. I. Troian, Phys. Rev. **D53** (1996) 89.

[14] M. Cacciari and S. Catani, Nucl. Phys. **B617** (2001) 253.

[15] M. Cacciari and E. Gardi, Nucl. Phys. B **664**, (2003) 299.

[16] V. N. Gribov, L. N. Lipatov, Sov. J. Nucl. Phys. **15** 438 (1972);
V. N. Gribov, L. N. Lipatov, Sov. J. Nucl. Phys. **15** 675 (1972).

[17] G. Altarelli, G. Parisi, Nucl. Phys. **B126** 298 (1997).

[18] T.Sjöstrand, Comp. Phys. Comm. **82** (1994) 74.

[19] T.Sjöstrand *et al.*, Comp. Phys. Comm. **135** 238 (2001).

[20] G. Marchesini *et al.*, Comp. Phys. Comm. **67**, 465 (1992);
G. Corcella *et al.*, JHEP **0101**, 010 (2001).

[21] R. D .Field and R. P. Feynman, Nucl. Phys. **B136** (1978) 1.

[22] D. Amati and G. Veneziano, Phys. Lett. **B83**, 87 (1979).

[23] K. G. Wilson, Phys. Rev. **D10** 2445 (1974).

[24] M.G. Bowler, Z. Phys. C **11** (1981) 169.

[25] X. Artru, G. Mennessier, Nucl. Phys. **B70** 93 (1974).

[26] T. D. Gottschalk, D. A. Morris, Nucl. Phys. **B288** 729 (1987.

[27] B. Andersson, G. Gustafson, C. Peterson, Z. Phys. **C1** 105 (1979);
B. Andersson, G. Gustafson, G. Ingelman, T. Sjöstrand, Phys. Rept. **97** 31 (1983).

[28] A. Casher, H. Neuberger, S. Nussinov, Phys. Rev. **D20** 179 (1979).

[29] B. Andersson, G. Gustafson, T. Sjöstrand, Nucl. Phys. **B197** 45 (1982) ;
T. Meyer, Z. Phys. **C12** 77 (1982).

[30] B. Andersson, G. Gustafson, T. Sjöstrand, Phys. Scripta **574** (1985).

[31] A. Heister *et al.*, ALEPH Collaboration, Phys. Lett. **B512** (2001) 30.

[32] G. Abbiendi *et al.*, OPAL Collaboration, Eur.Phys.J. **C29** N4 (2003) 463.

[33] K. Abe *et al.*, SLD Collaboration, Phys. Rev. **D65** (2002) 092006, Erratum-ibid. **D66** (2002) 079905.

[34] G. Barker *et al.*, DELPHI Collaboration, Contributed Paper for ICHEP 2002 (Amsterdam), DELPHI 2002-069 CONF 603.

[35] C. Peterson, D. Schlatter, I. Schmitt, P.M. Zerwas, Phys. Rev. **D27** (1983) 105.

[36] P. Collins, T. Spiller, J. Phys. **G11** (1985) 1289.

[37] V.G. Kartvelishvili, A.K. Likehoded, V.A. Petrov, Phys. Lett. **B78** (1978) 615.

[38] B. Andersson, G. Gustafson, B. Soderberg, Z. Phys. C **20** (1983) 317.

[39] S.J. Brodsky, G.R. Farrar, Phys. Rev. Lett. **31** (1973) 1153;
D. Sivers, S.J. Brodsky, R. Blankenbecler, Phys. Rept. **C23** (1976) 1.

[40] S.D. Drell, D.J. Levy, T.-M. Yan, Phys. Rev. **D1** (1970) 1617;
V.N. Gribov, L.N. Lipatov, Sov. J. Nucl. Phys. **15** (1972) 675;
R. Gatto, P. Menotti, I. Vendramin, Phys. Rev. **D7**(1973) 2524.

[41] M.G. Bowler, Z. Phys. C **22** (1984) 155.

[42] N. Isgur, M. B. Wise, Phys. Lett. **B232** (1989) 113;
N. Isgur, M. B. Wise, Phys. Lett. **B237** (1990) 527;
M. B. Wise: 'Heavy Flavor Theory: Overview'. In: Proceedings of the 16th International Symposium on Lepton and Photon Interactions, 1993, Ithaca, NY, pp 253-273, ed by P. Drell, D. Rubin, AIP Conference Proceedings 302;
N. Isgur, M. B. Wise: 'Heavy Quark Symmetry'. In: *B Decays*, 2nd edition, ed. S. Stone (World Scientific, 1994), 231;
M. Neubert: 'B Decays and the Heavy-Quark Expansion'. In: *Heavy Flavours II*, Advanced Series on Directions in High Energy Physics - Vol. 15, ed by A. J. Buras, M. Lindner (World Scientific, 1998), 239.

[43] S. Godfrey and R. Kokoski, Phys. Rev. **D 43** (1991) 1679.

[44] N. Isgur, Phys. Rev. **D 57** (1998) 4041.

[45] D. Ebert, V. O. Galkin and R. N. Faustov, Phys. Rev. **D 57** (1998) 5663;
Erratum-ibid. **D 59** (1999) 019902.

[46] A. F. Falk, M. E. Peskin, Phys. Rev. **D 49** (1994) 3320.

[47] ALEPH Collab., D. Buskulic *et al.*, Z. Phys. **C 69** (1996) 393.

[48] OPAL Collab., R. Akers *et al.*, Z. Phys. **C 66** (1995) 19.

[49] DELPHI Collab., Z. Albrecht *et al.*, DELPHI 2004-025 CONF 700, submitted to ICHEP04 (Beijing).

[50] `http://www.slac.stanford.edu/xorg/hfag/`.

[51] LEP design report, 1983, CERN-LEP/TH/83-29,
vol. 1 *The LEP injector chain*;
LEP design report, 1984, CERN-LEP/TH/84-01,
vol. 2 *The LEP main ring*.

[52] P. Abreu *et al.*, DELPHI Collaboration, Nucl. Instr. and Meth. **A378** (1995) 57;
P.A. Aarnio *et al.*, DELPHI Collaboration, Nucl. Instr. and Meth. **A303** (1991) 233.

[53] DELPHI Collaboration, *DELANA User's Guide*, DELPHI 89-44 PROG 137.

[54] V. Perevoztchikov and N. Smirnov, DELPHI Collaboration, DELPHI 92-118 PROG 189 Rev. 3, *PHDST Package Description*,
`http://nsmirnov.home.cern.ch/nsmirnov/phdst.html`.

[55] N. Smirnov and Tz. Spassoff, DELPHI Collaboration, *SKELANA User's Guide*, http://cossutti.home.cern.ch/cossutti/skelana.html.

[56] P. Rice-Evans, 1974, "Spark, Streamer Proportional and Drift Chambers", The Richelieu Press.

[57] C.J. Kreuter, *Electron Identification using a Neural Network*, DELPHI 96-169 PHYS 658;
F. Stichelbaut, G. Wilkinson, *Performance of Muon Identification in DELPHI for the 93 and 94 Data*, DELPHI 95-140 PHYS 565;
G. Wilkinson, *Improvements to the Muon Identification in the 94C2 Short DST Production*, DELPHI 97-37 PHYS 690;
K.D. Brand, I. Roncagliolo, F. Simonetto, *Electron Identification for Electro-Weak b,c Physics*, DELPHI note 96-23 PHYS 598;
P. Abreu *et al.*, DELPHI Collaboration. Eur. Phys. J **C20** (2001) 455.

[58] DELPHI Collaboration, *DELSIM, DELPHI Event Generation and Detector Simulation - Reference Manual*, DELPHI Note 89-68 PROG 143.

[59] D. Wicke, DELPHI Collaboration, DELPHI 98-163 PROG 236 TRACK 92.

[60] G. Aubrecht *et al.*, "A Teachers Guide to the Nuclear Science Wall Chart", Contemporary Physics Education Project, 2003.
http://www.lbl.gov/abc/wallchart/teachersguide/pdf/Chap11.pdf.

[61] C. W. Schmidt, "The Fermilab 400-MeV Linac upgrade"
FERMILAB-CONF-93-111 (1993).

[62] Fermilab Beams Division, "Run II Handbook",
http://www-bd.fnal.gov/runII/index.html.

[63] J. Marriner, "Stochastic Cooling Overview",
FERMILAB-CONF-03-158(2003).

[64] R. Blair *et al.*, "The CDF-II detector: Technical design report",
FERMILAB-PUB-96/390-E (1996).

[65] F. Abe *et al.*, Nucl. Inst. and Meth. Phys. Res, **271A**, 387 (1988);
FERMILAB-PUB-94/024-E (1994).

[66] T. K. Nelson *et al.*,
FERMILAB-CONF-01/357-E.

[67] A. Sill *et al.*, Nucl. Instrum. Meth., **A447**, 1–8 (2000).

[68] T. Affolder *et al.*, Nucl. Instrum. Meth., **A485** , 6–9 (2002).

[69] K. T. Pitts *et al.*,
FERMILAB-CONF-96-443-E.

[70] R. Wagner *et al.*, "CDF Central Outer Tracker", CDF Note 6267 (2003).

[71] Particle Data Group, S. Eidelman *et al.*, Phys. Lett. B **592**, 1 (2004).

[72] R. Brun, K. Hakelberg, M. Hansroul, and J.C. Lasalle, CERN-DD-78-2-REV; CERN-DD-78-2.

[73] G. Gomez for the TOF group, "The CDF-II Time of Flight Detector", CDF Note 6258 (2003).

[74] P. Schlabach, CDF Central Muon Detector, `http://www-cdf.fnal.gov/internal/upgrades/muon/cmu.html`.

[75] P. Schlabach, CDF Central Muon Detector, `http://www-cdf.fnal.gov/internal/upgrades/muon/cmp.html`.

[76] P. Schlabach, CDF Central Muon Detector, `http://www-cdf.fnal.gov/internal/upgrades/muon/cmx.html`.

[77] E. J. Thomson *et al.*, IEEE Trans. Nucl. Sci **49**, 1063 (2002).

[78] W. Ashmanskas *et al.*, Nucl. Instrum Methods Phys Res. A **447**, 218 (2000), W. Ashmanskas *et al.*, FERMILAB-CONF-02/035-E, A. Bardi *et al.*, Nucl. Instrum Methods Phys Res. A **485**, 6 (2002).

[79] D. Acosta *et al.*, Nucl. Inst. Methods. A **497**, 57, (2002).

[80] J. Abdallah *et al.*, DELPHI Collaboration, Eur. Phys. J. **C32** (2004) 185 [arXiv:hep-ex/0311003].

[81] Z. Albrecht *et al.*, DELPHI Collaboration, Contributed Paper for ICHEP 2004 (Beijing), DELPHI 2004-025 CONF 700.

[82] P. Abreu *et al.*, DELPHI Collaboration, Phys. Lett. **B462** (1999) 425.

[83] DELPHI Collaboration P. Abreu *et al.*, Zeit. Phys. **C73** (1996) 11.

[84] R Barate *et al.* Phys Reports **294**, (1998) 1 (CERN PPE 96-186); Gerald Rudolph, private communication.

[85] A. Bassetto, M. Ciafaloni and G. Marchesini, Phys. Rep. **100** (1983) 201.

[86] CLEO Collaboration, R. A. Briere *et al.*, Phys. Rev. D **62**, (2000) 072003.

[87] P. J. Bussey for the CDF Collaboration [arXiv:hep-ex/0408020]; CDF Collaboration, D. Acosta *et al.*, Phys. Rev. Lett. **91** (2003) 241804.

[88] E. Ben-Haim, P. Bambade, P. Roudeau, A. Savoy-Navarro and A. Stocchi, Phys. Lett. B **580** (2004) 108.

[89] CDF Note 7184 (2004).

[90] R. Field, CDF note 6004, The Underlying Event in Run 2 (2003).

[91] S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, Adv. Ser. Direct. High Energy Phys. **15** (1998) 609 [arXiv:hep-ph/9702287].

[92] P. Nason *et al.*, Report of the 1999 CERN Workshop on SM physics (and more) and the LHC [arXiv:hep-ph/0003142].

[93] Matteo Cacciari, private communication.

[94] D. Acosta *et al.*, Phys. Rev. **D65**, (2002) 052005.

[95] A. P. Prudnikov, Yu. A. Brychkov and O. I. Marichev, *Integrals and Series- vol. 3: more special functions*, (Gordon and Breach Science Publishers, 1990).

[96] R. J. Barlow, *Statistics*, (Wiley, 1989).

List of Figures

1 Distributions fittées de l'énergie reconstruite du B, pour les événements sélectionnés par l'analyse de DELPHI. . . . 14
2 Distributions, en fonction de la variable *z*, obtenues pour les événements sélectionnés en 1994, par l'analyse de DELPHI. . . . 15
3 Comparaison entre les distributions de la variable x_E mesurées par ALEPH, DELPHI, OPAL et SLD. . . . 16
4 La distribution de la composante QCD non-perturbative de la fonction de fragmentation du quark *b*, en utilisant le générateur JETSET pour évaluer la composante QCD perturbative. . . . 18
5 La distribution de la composante QCD non-perturbative de la fonction de fragmentation du quark *b*, en utilisant un calcul analytique basé sur QCD à l'ordre NLL pour évaluer la composante QCD perturbative. . . . 19
6 Comparaison entre données et simulation, contenant tous mécanismes de production du quark *b*. La distribution de l'impulsion transverse au faisceau du méson-B. Les contributions des différents mécanismes de production du B sont montrés séparément les une des autres. . . . 21
7 Comparaison entre données et simulation, contenant tous mécanismes de production du quark *b*. La distribution de l'impulsion transverse au faisceau des traces dans un cône défini autour du B . . . 22
8 La section efficace différentielle de production du quark *b* avec pseudo-rapidité inférieure à 1. . . . 23

2.1 A scheme of the b production, fragmentation and decay. . . . 32
2.2 Fermion pair production at LEP. . . . 34
2.3 The lowest order contributions to b production in the TeVatron. . . . 36
2.4 Cancellation of real gluon radiation and virtual gluon exchange. . . . 40
2.5 Elementary processes contributing to the parton shower model. . . . 42
2.6 A scheme of a parton shower process. . . . 43
2.7 A scheme of the cluster hadronization model. . . . 46
2.8 A scheme of the string hadronization model. . . . 47

3.1 LEP collider. . . . 59
3.2 LEP luminosity. . . . 60
3.3 Schematic layout of the DELPHI detector. . . . 61
3.4 Schematic cross-sections of the DELPHI Vertex Detector. . . . 62
3.5 Scheme of the DELPHI Inner Detector. . . . 63

3.6 Schematic view of the sense wires and pads rows of the DELPHI TPC. 64
3.7 Scheme of the DELPHI Time Projection Chamber. 64
3.8 Operation principle of the RICH detector. 66
3.9 Impact parameter definition. 69
3.10 Impact parameter sign definition. 69

4.1 Layout of the Fermilab accelerator complex. 74
4.2 Peak luminosities for stores collided between April 2001 and August 2004. . . 77
4.3 The CDF-II detector with quadrant cut to expose the different subdetectors. . . 78
4.4 A diagram of the CDF-II tracker layout showing the different subdetector systems. 79
4.5 Coverage of the different silicon subdetector systems projected into the $r-z$ plane. The r and z axes have different scales. 81
4.6 Layout of wire planes on a COT endplate. 82
4.7 Layout of wires in a COT supercell. 83
4.8 Dependence of the reconstructed invariant mass of $J/\psi \to \mu^+\mu^-$ decays on the transverse momentum of the J/ψ. 84
4.9 Rate of kaon and pion tracks faking muon signals in the CDF-II detector. 87
4.10 Diagram of the CDF-II trigger system. 88
4.11 Diagram of the different trigger paths at Level 1 and 2. 89
4.12 SVT principle of operation. 90
4.13 SVT impact parameter resolution. 91
4.14 Principle of Event Building and Level 3 Filtering. 92

5.1 Distribution of $-\log_{10}\left(P_{btag}\right)$. 96
5.2 Ratio Data/MC for charged and neutral track momentum distributions before applying corrections. 98
5.3 Ratio Data/MC for charged and neutral track momentum distributions after applying corrections. 100
5.4 Comparison, for a depleted b sample, between data and simulation for the fitted jet energy. 102
5.5 Comparison, using a depleted b sample, between data and simulation for charged, neutral and missing energy measurement. 103
5.6 Comparison, using a b-enriched sample, between data and simulation for charged, neutral and missing energy measurement. 103
5.7 $(p_{rec}-p_{sim})/p_{sim}$ for B-hadrons. 104
5.8 Fitted acceptance for signal events versus $x_p = p_B/E_{beam}$. 105
5.9 Comparison between the measured x_p^{rec} distributions obtained in data and in the MC $q\bar{q}$ simulation. 106
5.10 Fitted $x_p^{rec} = p_B^{rec}/p_{jet}$ distributions on selected events. 108
5.11 Fitted z distribution on 1994 selected events. 108
5.12 Comparison between x_E distributions obtained using the 1994 and 1995 events samples. 109
5.13 Comparison between generated and fitted x_E distributions. 111
5.14 Differences between the fitted x_E values, relative to the average, for different data samples. 112

5.15 Ratio between number of events having a given jet multiplicity in data and in the simulation. . . . 113
5.16 Systematic uncertainties originating from uncertainties in the tuning of the simulation. . . . 114
5.17 Systematic uncertainties originating from uncertainties on physics parameters governing B-hadron decay and production characteristics. . . . 117
5.18 Systematic uncertainties related to the stability of results versus the values of criteria used in the analysis. . . . 119
5.19 Comparison between the measured x_E distribution and corresponding results obtained by ALEPH, DELPHI, OPAL and SLD. . . . 121
5.20 Comparison between expected x_E distributions for B_d^0 and B^+ mesons and for B_s^0 and *b*-baryons. . . . 122

6.1 x and N representations of the measured fragmentation function. Moments of the perturbative QCD component. . . . 126
6.2 Curves used to generate error bars for the non-perturbative QCD component. . 128
6.3 The extracted non-perturbative QCD component when the perturbative one is taken from JETSET 7.3. . . . 129
6.4 Moments of the fragmentation function from NLL QCD and their dependence on scales. . . . 130
6.5 The extracted non-perturbative QCD component when the perturbative one is taken from a theoretical NLL QCD calculation. . . . 131
6.6 Comparison between the not-perturbative components obtained for three different QCD Approaches. . . . 132
6.7 Comparison of the fitted fragmentation models and the extracted non-perturbative component. . . . 134
6.8 The measured fragmentation function comparing to predictions from models and from the directly extracted function (NLL QCD). . . . 134
6.9 Comparison between the fitted and measured *b*-quark fragmentation distributions. 136
6.10 Binning effect on the moments of a distribution. . . . 138
6.11 The extracted function for two different parametrizations of the measured fragmentation function. . . . 138
6.12 The non-perturbative component obtained with a different number of eigenvalues. 139
6.13 The non-perturbative component for jetset with aleph tuning. . . . 140
6.14 Comparison of the extracted non-perturbative QCD component for different experiments. . . . 142
6.15 Comparison of the extracted non-perturbative QCD component from the present thesis with the one of other experiments. Perturbative component from JETSET. 143
6.16 Comparison of the extracted non-perturbative QCD component from the present thesis with the one of other experiments. Perturbative component from NLL QCD. 144
6.17 Contribution to moments of the fragmentation function from each bin of the measured distribution. . . . 146
6.18 Normalized spectrum of the scaled momentum x_p of D^0 mesons measured by CLEO. . . . 147

6.19 x-dependence of the perturbative and non-perturbative QCD components of the measured charm quark fragmentation distribution. 148
6.20 x-dependence of the non-perturbative QCD component of the measured charm quark fragmentation distribution. 148

7.1 The invariant mass of $B^{\pm} \to J/\psi K^{\pm}$, with the fit overlaid. 155
7.2 Different contributions to the $B^{+} \to J/\psi K^{+}$ signal, from an inclusive $H_b \to J/\psi X$ Monte Carlo simulation. 156
7.3 The distribution of Δ for tracks. 157
7.4 The effect of the removal of the pile-up events by subtracting the Δz sidebands on p_L^{rel}. 159
7.5 Regions of signal and sidebands for $m(J/\psi K^{+})$ and $Deltaz$ 160
7.6 $B^{+} \to J/\psi K^{+}$ data and PYTHIA with msel=5 and CKIN(3)=5. 162
7.7 The Peterson, Lund and Bowler hadronization models. 163
7.8 The p_T spectrum of the B meson in four Monte Carlo samples, generated with different fragmentation functions. 166
7.9 The p_T spectrum of tracks in a cone around the B meson in three Monte Carlo samples, generated with different fragmentation functions. 167
7.10 The component of tracks momentum transverse to the B meson, for tracks in a cone around it in three Monte Carlo samples, generated with different fragmentation functions. 167
7.11 The p_T spectrum of the highest p_T track in a cone around the B meson in three Monte Carlo samples, generated with different fragmentation functions. 168
7.12 The multiplicity of tracks in a cone of opening $\Delta R = 0.7$ around the B meson, in three Monte Carlo samples, generated with different fragmentation functions. 168
7.13 Variation of the B production cross section, expected when considering different fragmentation models. 169
7.14 The p_T spectrum of the B. Comparison between the data and two Monte Carlo msel=5 samples with the Bowler and Peterson fragmentation functions. Histograms are normalized to the number of B mesons. 170
7.15 The p_T spectrum of tracks in a cone around the B meson. Comparison between the data and two Monte Carlo (msel=5) samples with the Bowler and Peterson fragmentation functions. 170
7.16 The η spectrum of tracks in a cone with $\Delta R > 1$ (control tracks). Comparison between the data and two Monte Carlo (msel=5) samples with the Bowler and Peterson fragmentation functions. 171
7.17 With the msel=1 PYTHIA setting, there is good agreement between data and Monte Carlo for the p_T of the B meson. 172
7.18 The multiplicity of tracks within a cone of $\Delta R < 0.7$ around the B meson in the msel=1 PYTHIA sample. 173
7.19 The transverse momentum p_T of all tracks within a cone of $\Delta R < 0.7$ around the B meson in the msel=1 PYTHIA sample. 174
7.20 p_T^{rel} of tracks within a cone of $\Delta R < 0.7$ around the B meson in the msel=1 PYTHIA sample. 174

7.21 p_L^{rel} of tracks within a cone of $\Delta R < 0.7$ around the B meson in the msel=1 PYTHIA sample. 175
7.22 The transverse momentum p_T for the leading directional track in a cone of $\Delta R < 0.7$ around the B meson in the msel=1 PYTHIA sample. 175
7.23 The transverse momentum p_T for the leading momentum track in a cone of $\Delta R < 0.7$ around the B meson in the msel=1 PYTHIA sample. 176
7.24 The sum of transverse momenta of tracks in a cone of $\Delta R < 0.7$ around the B meson in the msel=1 PYTHIA sample. 176
7.25 The distribution of the variable $\frac{E_B - \sum E_{tracks}}{E_B + \sum E_{tracks}}$, where the sum is over tracks in in a cone of $\Delta R < 0.7$ around the B meson in the msel=1 PYTHIA sample. 177
7.26 Plot of the 3-momentum for control tracks in the msel=1 PYTHIA sample. . . . 177
7.27 Plot of the impact parameter d_0 for control tracks in the msel=1 PYTHIA sample. 178
7.28 Plot of the angular variable ϕ for control tracks in the msel=1 PYTHIA sample. 179
7.29 Plot of the pseudo-rapidity variable η for the control tracks in the msel=1 PYTHIA sample. 180
7.30 The track activity in the $\Delta\eta$-$\Delta\phi$ plane around the B meson in the data and Monte Carlo msel=1 sample. 181
7.31 Data and Monte Carlo comparison for $p_T(B)$, with different contributions in the Monte Carlo sample from the FC, FE and GS b quark production mechanisms. 182
7.32 Data and Monte Carlo comparison for p_T of tracks in a cone around the B meson. Contributions in the Monte Carlo sample from the FC, FE and GS b quark production mechanisms are detailed. 183
7.33 Data and Monte Carlo comparison for p_L^{rel} of tracks in a cone around the B meson. Contributions in the Monte Carlo sample from the FC, FE and GS b quark production mechanisms are detailed. 184
7.34 Data and Monte Carlo comparisons for the angular distributions of all the tracks in the event with respect to the B meson direction. Contributions in the Monte Carlo sample from the FC, FE and GS b quark production mechanisms are detailed. 185
7.35 Contributions in the Monte Carlo sample from the FC, FE and GS to the p_T distribution of the B meson, normalized to unity. 186
7.36 Contributions in the Monte Carlo sample from the FC, FE and GS to the p_T distribution of tracks in a cone around the B meson, normalized to unity. 186
7.37 Contributions in the Monte Carlo sample from the FC, FE and GS to the $\Delta\eta$ and $\Delta\phi$ distribution of all tracks in the event. Histograms are normalized to unity. . 187
7.38 Total efficiency of detection and analysis effects for the $B^{\pm}$ mesons, in two regions of pseudo-rapidity. 190
7.39 The differential cross section for $|\eta| < 1$. 191
7.40 The differential cross section for $|\eta| < 0.6$. 191
7.41 η distributions of the reconstructed B meson, in data and in the Monte Carlo sample, used to calculate the efficiency. 194
7.42 η distributions of the J/ψ candidates, in data and in the Monte Carlo sample, used to calculate the efficiency. 194
7.43 η distributions of the kaon candidates, in data and in the Monte Carlo sample, used to calculate the efficiency. 195

7.44 p_T spectra for generated B mesons after y and η cuts. 196
7.45 Spectra of y and η for generated B mesons. 196
7.46 The measured differential cross section for $|y| < 0.6$ and the FONLL theoretical prediction 197
7.47 The measured differential cross section for $|y| < 1.$, compared to the measurement from CDF RunI. 198

A.1 Integration contour for the inverse Mellin transformation. 202
A.2 Piecewise defined function for $0 < x < 1$. 203

B.1 $\chi^2/d.o.f$ of the fit as a function of number of eigenvalues for OPAL's fragmentation function. 207

List of Tables

1 Différentes mesures de $< x_E >$ à l'énergie du Z^0. 16

2.1 Properties of the four $B^{**}_{u,d}$ states. 54
2.2 Current world results for the mass and production rate of $B^{**}_{u,d}$ states. 55

3.1 Properties of the LEP e^+e^- collider. 58
3.2 Measured points and resolution of the DELPHI tracking system. 67
3.3 Lepton tag characteristics. 70

4.1 Accelerator parameters for Run I and Run II configurations. 76
4.2 Relevant parameters for the layout of the sensors of different SVX-II layers. . . 80
4.3 Pseudorapidity coverage, energy resolution and thickness for the different calorimeter subdetectors. 86

5.1 Variation of the selected event sample composition and efficiency for *b* events, versus the cut on the *btag*-variable. 96
5.2 Fractions of tracks adjusted in simulated and real events. 97
5.3 Fitted fractions of charged and neutral energies, reconstructed in *b*-depleted and *b*-enriched event samples. 101
5.4 Charged and neutral track multiplicities measured in data and simulation, after corrections. 101
5.5 Integral and corresponding statistical error matrix of the fitted fragmentation distribution, over specified intervals. 110
5.6 Variation of the fitted average value of the beam energy taken by a *b*-hadron with the value taken for the *curv* parameter. 110
5.7 Systematic uncertainties, in each bin of x_E, related to the tuning of the simulation. 113
5.8 Systematic uncertainties, in each bin of x_E, related to uncertainties on physics parameters governing B-hadron decay and production characteristics. 116
5.9 Variation of the fitted average value of the beam energy taken by a *b*-hadron with the cut on $P(btag)$. 116
5.10 Variation of the fitted average value of the beam energy taken by a *b*-hadron with the cut on "ambiguous" energy value. 118
5.11 Variation of the fitted average value of the beam energy taken by a *b*-hadron with the minimal charged track multiplicity at the B decay vertex. 118
5.12 Systematic uncertainties, in each bin of x_E, related to the values and procedures used in the analysis. 118
5.13 Different measurements of $< x_E >$ at the Z energy. 120

5.14 Total error matrix of the fitted fragmentation distribution, over specified intervals. 120

6.1 Fit results for fragmentation models parameters. 133
6.2 the fitted parameters of the proposed parametrization. 136
6.3 The correlation matrices on the fitted parameters of the proposed parametrization. 136
6.4 The error matrix on the fitted parameters of the combined fragmentation distribution. 141
6.5 Moments of the combined fragmentation function for x_E. 145

7.1 Analysis cuts for the decay $B^{\pm} \to J/\psi K^{\pm}$ ($J/\psi \to \mu^+\mu^-$) along with requirements made on tracks in the vicinity of the B. 154
7.2 The result of the B invariant mass fit. 154
7.3 The result of the Δz fit for tracks. 158
7.4 A χ^2 calculated to compare two Monte Carlo samples, with the Bowler and the hard Peterson fragmentation functions, for a choice of variables. 165
7.5 Summary of the systematic uncertainties in the b production cross section measurement. 193

A.1 Integration contours for a function defined piecewise. 204