LAL 14-296

UNIVERSITÉ PARIS-SUD

École Doctorale: Particules, Noyaux et Cosmos - ED 517
Laboratoire de l'Accélérateur Linéaire - UMR 8607
Centre Scientifique d'Orsay, Bâtiment 200 - BP 34, 91898 Orsay CEDEX - France
Discipline: Physique
Spécialité: Physique des particules

THÈSE DE DOCTORAT

soutenue le 19/09/2014

par

Olga KOCHEBINA

Study of Rare Charm Decays with the LHCb Detector at CERN

Étude de désintégrations charmées rares avec le détecteur LHCb au CERN

Co-directeur de thèse:	Benoît VIAUD	Chercheur (LAL)
Directrice de thèse:	Marie-Hélène SCHUNE	Directrice de recherche (LAL)

Composition du jury:

Président du jury:	Achille STOCCHI	Professeur (LAL, PSUD-XI)
Rapporteurs:	Vera LUTH	Professeur (SLAC)
	Guy WILKINSON	Professeur (Oxford University)
Examinateurs:	Svjetlana FAJFER	Professeur (Institute Jozef Stefan)
	Urs LANGENEGGER	Chercheur (Paul Scherrer Inst.)

Contents

Synthèse **11**

Introduction **21**

1 Theoretical overview **25**
1.1 The Standard Model . . . 25
1.1.1 Short overview . . . 25
1.1.2 The symmetries . . . 29
1.1.3 The SM Lagrangian . . . 30
1.1.4 Higgs mechanism . . . 32
1.1.5 Neutral currents and the GIM mechanism . . . 34
1.1.6 The CKM matrix . . . 36
1.1.7 CP violation . . . 38
1.1.8 Unanswered questions of the Standard model . . . 40
1.2 Rare charm decays as probes of New Physics models . . . 41
1.2.1 Introduction to charm physics . . . 42
1.2.2 General theoretical approach . . . 43
1.2.3 Effective Hamiltonian for the $D \to h(h')\mu^+\mu^-$ decays . . . 46
1.2.4 Overview of New Physics models . . . 49

2 The LHCb detector at the LHC **55**
2.1 General overview . . . 56
2.2 Tracking . . . 57
2.2.1 VELO . . . 57
2.2.2 Tracking stations . . . 58
2.3 Dipole magnet . . . 60
2.4 Particle identification . . . 60
2.4.1 Ring Imaging Cherenkov detectors . . . 60
2.4.2 Calorimeter system . . . 62
2.4.3 Muon chambers . . . 64
2.5 Trigger . . . 65
2.5.1 L0 . . . 65
2.5.2 HLT . . . 66
2.5.3 Categories of the trigger decisions: TIS, TOS, Dec . . . 67
2.6 LHCb data processing, stripping . . . 68
2.7 Upgrade . . . 68

2.7.1 Trigger . . . 69
2.7.2 Tracking system . . . 70
2.7.3 Particle identification . . . 71
2.8 Monte Carlo simulations . . . 71
2.9 Important algorithms for data analyses of rare charm decays . . . 71
2.9.1 Track reconstruction . . . 71
2.9.2 Particle identification . . . 72
2.10 MC deficiencies and performance of the LHCb detector . . . 74
2.10.1 Performance of the detector for rare charm decays . . . 75
2.11 Variables used in LHCb data analyses . . . 80

3 Search for $D^+_{(s)} \to \pi^+\mu^+\mu^-$ and $D^+_{(s)} \to \pi^-\mu^+\mu^+$ decays 87
3.1 Overview of the search for $D^+_{(s)} \to \pi^+\mu^+\mu^-$ and $D^+_{(s)} \to \pi^-\mu^+\mu^+$ with 1.0 fb^{-1} at $\sqrt{s} = 7$ TeV . . . 87
3.2 Systematic uncertainties . . . 91

4 Data preprocessing for $D^+ \to h^+h^{(\prime)-}\mu^+\mu^-$ decays analyses 99
4.1 Reoptimization of HLT2 lines for 2012 data taking period . . . 99
4.1.1 HLT2 rare charm decays lines at LHCb . . . 99
4.1.2 Trigger re-optimization . . . 100
4.1.3 The HLT2 selection for $D^0 \to h^+h^-h^+h^-$ decay . . . 102
4.2 Stripping offline preselection . . . 107

5 Measurement of the partial branching ratio of the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ decay with 2 fb^{-1} of pp data collected by LHCb in 2012 109
5.1 Motivation . . . 109
5.2 Analysis strategy . . . 111
5.3 Data and MC samples used in the analysis . . . 112
5.4 Selection . . . 113
5.4.1 Trigger . . . 114
5.4.2 Offline selection . . . 114
5.4.3 Multiple candidates treatment . . . 116
5.5 Efficiency ratio determination . . . 117
5.5.1 Data driven efficiency: muon identification . . . 117
5.5.2 Pion inefficiency: decays in flight and nuclear interactions in detector's material. . . . 121
5.5.3 L0 trigger efficiency . . . 122
5.5.4 Higher Level Trigger efficiency . . . 124
5.5.5 Data/MC comparisons . . . 125
5.5.6 Re-evaluation of the efficiencies based on a Smeared MC . . . 144
5.5.7 Model dependence . . . 144
5.5.8 Systematic uncertainty due to the track reconstruction . . . 147
5.5.9 Systematic uncertainty due to the kaon identification . . . 148
5.6 Additional specific backgrounds . . . 149
5.6.1 Mechanisms to produce additional backgrounds . . . 150
5.6.2 Quantification of the highest background yields . . . 151

5.7 Fit for the measurement of the branching fraction . . . 153
5.7.1 Global structure . . . 153
5.7.2 Parametrisation of peaking structures . . . 154
5.7.3 Parametrisation of the combinatorial background . . . 159
5.7.4 Parametrisation of the double misidentification background from $D^0 \to K^-\pi^+\pi^+\pi^+$ . . . 161
5.7.5 Validation of the fit model on data control sample . . . 170
5.7.6 Unblinded fit . . . 172
5.8 Branching fraction uncertainty . . . 172
5.8.1 Uncertainties related to the signal extraction . . . 172
5.8.2 Uncertainty related to the L0 efficiency ratio . . . 176
5.8.3 Uncertainty summary . . . 179
5.9 Conclusion . . . 180

6 Study of $D^0 \to h^+h^{(\prime)-}\mu^+\mu^-$ decays reconstructed with a D^{*+}-tag 181
6.1 Introduction to multivariate analysis techniques . . . 181
6.1.1 Artificial neural networks . . . 182
6.1.2 Boosted decision trees . . . 182
6.1.3 BDT with TMVA . . . 185
6.1.4 Application of MVA techniques: offline selection of the $D^{+*} \to \pi^+_{slow}(D^0 \to K^+\pi^-\mu^+\mu^-)$ decays . . . 186
6.1.5 Results . . . 195
6.2 Sensitivity to branching ratios and asymmetries in $D^0 \to h^-h^{(\prime)+}\mu^+\mu^-$ decays . . . 198
6.2.1 Sensitivity to the total and partial branching ratios . . . 198
6.2.2 Sensitivity to asymmetries . . . 200

7 Test of the upgraded calorimeter front-end electronics 203
7.1 Front-End Board . . . 204
7.2 Test beam setup . . . 207
7.3 Test beam results . . . 208
7.3.1 Noise . . . 208
7.3.2 Linearity . . . 209
7.3.3 Integrator plateau . . . 210
7.3.4 Spill over . . . 211
7.3.5 Reflection . . . 211
7.4 Conclusions and consequences . . . 212

8 Conclusion 215

A BDT. Input variable distributions, Correlation matrices and variable ranking for $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ mode. 217

B Additional plots related to Section 5.7 221
B.1 Tests with Double Crystal Ball function . . . 221
B.2 Toy Study . . . 225

C The comparison of $ProbNNmu$ distributions in data and MC. 227

D Checks of binnings for PIDCalib 229
D.1 $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ vs $D^0 \to K^-\pi^+\pi^+\pi^-$ in MC 236

E Distributions of variables used for BDT training in data and MC. `L0HadronTOS||L0MuonTOS` trigger selection. 239
E.1 $D^0 \to K^-\pi^+\pi^+\pi^-$ 239
E.2 Tests of BDT distributions for compatibility of data/MC on 20% $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ sample. 243
E.3 Tests of efficiency of BDT selection on `L0HadronTOS||L0MuonTOS` sample. 247

F Kinematical distributions in data and MC for $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ selected with `L0HadronTOS||L0MuonTOS` and `L0HadronTIS` (nominal). 249

G Scans for sensitivity studies 253

Résumé

Les désintégrations charmées rares interviennent principalement via des courants neutres changeant la saveur (FCNC). Le Modèle Standard (SM) n'autorise les courants qu'au niveau des boucles. Dans les désintégrations du charme, les FCNC sont sujets à une très efficace suppression de GIM. Des processus très rares sont donc à considérer. Ils sont d'excellents outils pour la recherche Nouvelle Physique (NP) au-delà du SM. Les particules de NP pourraient devenir détectables en écartant des observables telles que des rapports de branchement et des asymétries CP et angulaires prévisions de le SM.

Le sujet principal de cette thèse est la mesure du rapport de branchement $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$. Il sera précieux en particulier en tant que mode de normalisation lors de l'étude de toutes les désintégrations $D^0 \to h^\pm h^{(\prime)\mp}\mu^+\mu^-$: $D^0 \to K^-\pi^+\mu^+\mu^-$, $D^0 \to \pi^+\pi^-\mu^+\mu^-$, $D^0 \to K^+K^-\mu^+\mu^-$ et $D^0 \to K^+\pi^-\mu^+\mu^-$. En utilisant $2\,\text{fb}^{-1}$ de données collectées par LHCb en 2012, nous mesurons:

$$\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)) = (4.37 \pm 0.12_{stat.} \pm 0.53_{syst.}) \times 10^{-6}.$$

C'est la toute première mesure de ce mode.

Nous avons également étudié la sensibilité qu'attendra LHCb dans les modes $D^0 \to h^\pm h^{(\prime)\mp}\mu^+\mu^-$ pour la mesure de rapport des branchement totaux et partiels, et pour celle d'asymétries, avec les échantillons de données qui seront collectés d'ici 10 ans.

Par ailleurs, nous avons déterminé les incertitudes systématiques touchant les recherches de désintégrations à 3 corps, $D^+_{(s)} \to \pi^+\mu^+\mu^-$ et $D^+_{(s)} \to \pi^-\mu^+\mu^+$, effectués par LHCb dans les données recueillies en 2011 ($1\ \text{fb}^{-1}$).

Enfin, les tests effectués sue les prototypes des cartes d'électronique embarquée qui assureront la lecture des calorimètres de l'expérience le LHCb mise à jour sont présenté dans cette thèse.

Les mots clés: Modèle Standard, LHC, expérience le LHCb, désintégrations charmées, désintégrations charmées rares, FCNC, Nouvelle Physique, LHCb Upgrade

Abstract

Rare charm decays proceed mostly through the $c \to u$ Flavor Changing Neutral Current (FCNC), which is possible only at loop level in the Standard Model (SM). In charmed decays, FCNCs are subject to a very efficient GIM suppression leading to very rare processes. Consequently, rare charm decays are good tools to probe New Physics (NP) beyond the SM. NP particles could become detectable by causing observables such as branching ratios and CP or angular asymmetries to deviate from the SM predictions.

The main subject of this thesis is the measurement of the branching ratio of the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$, the partial branching ratio of the of the $D^0 \to K^-\pi^+\mu^+\mu^-$ mode, in the region of the $m(\mu^+\mu^-)$ dominated by the ρ and ω resonances: $675\ \mathrm{MeV}/c^2 < m(\mu^+\mu^-) < 875\ \mathrm{MeV}/c^2$.

. It will be precious in the future, in particular as a normalisation mode in the study of all: $D^0 \to h^\pm h^{(\prime)\mp}\mu^+\mu^-$ decays $D^0 \to K^-\pi^+\mu^+\mu^-$, $D^0 \to \pi^+\pi^-\mu^+\mu^-$, $D^0 \to K^+K^-\mu^+\mu^-$ and $D^0 \to K^+\pi^-\mu^+\mu^-$. Using $2\,\mathrm{fb}^{-1}$ of 2012 LHCb data we find:

$$\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)) = (4.37 \pm 0.12_{stat.} \pm 0.53_{sys.}) \times 10^{-6}.$$

This is the first measurement of this mode.

We also determined sensitivities to total and partial branching fractions and asymmetries in $D^0 \to h^\pm h^{(\prime)\mp}\mu^+\mu^-$ decays with future LHCb datasets.

In addition, the systematic uncertainties affecting the searches for the 3-body decays, $D^+_{(s)} \to \pi^+\mu^+\mu^-$ and $D^+_{(s)} \to \pi^-\mu^+\mu^+$, carried out by LHCb based on the data collected in 2011 ($1\ \mathrm{fb}^{-1}$).

Finally, the results of the tests of front-end electronic board for the Upgrade of LHCb are presented.

Key words: Standard Model, New Physics, LHC, LHCb, Charm, Rare Charm Decays, Flavour Changing Neutral Current (FCNC), LHCb Upgrade

Synthèse

Le sujet principal de cette thèse est l'étude de désintégrations charmées rares en quatre corps, deux hadrons et deux muons $D^0 \to h^\pm h^{(\prime)\mp}\mu^+\mu^-$. Il s'agit de $D^0 \to K^-\pi^+\mu^+\mu^-$, $D^0 \to \pi^+\pi^-\mu^+\mu^-$, $D^0 \to K^+K^-\mu^+\mu^-$ et $D^0 \to K^+\pi^-\mu^+\mu^-$.

À courte distance, les désintégrations charmées rares procèdent principalement via des courants neutres changeant la saveur c en u, qui sont possibles seulement au niveau des boucles dans le Modèle Standard (SM). Elles sont également très supprimées par le mécanisme de GIM. Par conséquent, les désintégrations charmées rares sont d'excellents outils pour la recherche de Nouvelle Physique (NP). Ainsi, de nouvelles particules qui contribueraient aux diagrammes en boucles, deviendraient détectables grâce à des rapports de branchement et des asymétries plus élevés que predit par le SM.

Cependant, il n'est pas facile d'accéder à la physique à courte distance dans ces désintégrations. En effet, leurs rapports de branchement totales sont dominés par des contributions à longue distance (LD) dues à des états intermédiaires résonnants, comme $D \to XV'(\to \mu^+\mu^-)$. Une facon possible de les tester est de mesurer les rapports de branchement partiels, c'est-à-dire dans les régions extrêmes du spectre $m(\mu^+\mu^-)$, où les contributions des résonances sont moins importantes. Une autre possibilité est de mesurer des asymétries, comme les asymétries dites T-odd ou avant-arrière. Dans ce cas, même les régions dominées par les résonances (la physique de LD) peuvent être utilisées. Néanmoins, des contraintes théoriques et expérimentales indiquent que les rapports de branchement à courte distance ne doivent pas être supérieurs à 10^{-8}, et que les asymétries ne doivent pas dépasser quelques pourcents. Pour aider à comprendre l'intérêt des études des désintégrations charmées rares, une courte introduction théorique et un panorama des prévisions des modèles de NP sont présentés dans le Chapitre 1.

L'expérience LHCb présentée dans le Chapitre 2 est un laboratoire idéal pour les études d'effets supprimés dans les désintégrations charmées rares. Grâce à un excellent système de déclenchement, la reconstruction de vertex et la performance générale du détecteur, des tests de précision du SM dans le secteur des saveurs lourdes sont possibles. Le programme central de physique de LHCb porte sur les mesures d'asymétries de CP- et angulaires ($B_s^0 \to J/\psi\phi$ [1], $B \to K\mu^+\mu^-$ [2–4], $B \to DK^{(*)}$ [5, 6]), ainsi que les recherches de la Nouvelle Physique dans les désintégrations rares ($B_s^0 \to \mu^+\mu^-$ [7], $B \to K\mu^+\mu^-$ [8], $B^0 \to K^*\gamma$ [9] etc.). La physique du charme est également devenue une partie importante du programme. Ce dernier inclut la recherche de violation de CP [10], de mélange [11] dans les désintégrations $D \to h^+h'^-$ et des désintégrations rares $D^0 \to \mu^+\mu^-$ [12], $D^+ \to \pi^+\mu^+\mu^-$ [13], $D^0 \to hh\mu^+\mu^-$ [14].

Le détecteur LHCb [15] est un spectromètre à un bras avec une couverture angulaire

en avant comprise entre 10 mrad et 250 mrad dans le plan non-courbé et entre 10 mrad et 300 mrad dans le plan courbe, ceci correspond à un couverture de $2 < \eta < 5$ en pseudorapidité. Une vue schématique du détecteur est présentée Figure 1. Le géométrie du détecteur est motivée par le fait que les paires $b\bar{b}$ et $c\bar{c}$ sont produites dans des collisions pp à très petits angles dans les régions avant ou arrière. Ces paires sont en effet produites par des gluons virtuels hautement boostés qui viennent de la diffusion inélastique profonde des pp.

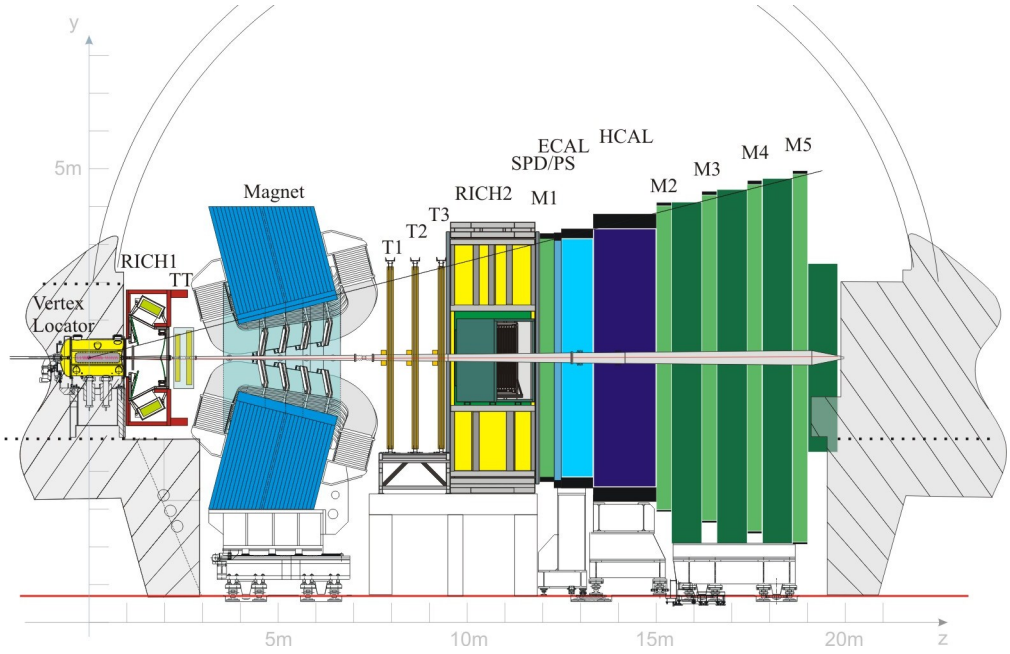


Figure 1: Vue schématique du détecteur LHCb.

Récemment plusieurs désintégrations charmées rares ont été recherchées dans LHCb. Les limites suivantes ont été obtenues :

- $\mathcal{B}(D^0 \to \mu^+\mu^-) < 7.6 \times 10^{-9}$ @ 95% CL [12].
- $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-) < 8.3 \times 10^{-8}$ @ 95% CL [13].
- $\mathcal{B}(D^0 \to \pi^+\pi^-\mu^+\mu^-) < 6.7 \times 10^{-7}$ @ 95% CL [14].

Ces mesures sont des ordres de grandeurs meilleures que les limites précédentes obtenues par les Flavour Factories, Tevatron ou l'expérience E971.

Grâce aux données aquisent pendant le Run I, une amélioration est aussi attendue dans le cas des désintégrations $D^0 \to h^\pm h^{(\prime)\mp}\mu^+\mu^-$, qui constituent le cœur du travail effectué au cours de cette thèse. Les limites actuelles obtenues par la collaboration E791 [16], sont :

$$\mathcal{B}(D^0 \to K^+K^-\mu^+\mu^-) < 3.3 \times 10^{-5} \text{ @ 90\% CL}$$
$$\mathcal{B}(D^0 \to K^-\pi^+\mu^+\mu^-) < 3.59 \times 10^{-4} \text{ @ 90\% CL}$$

Les asymétries CP et angulaires semblent également être utiles dans ces modes, pour lesquels un plus grand nombre d'observables est disponible grâce à la multiplicité d'états finaux (dans des asymétries T-odd et Forward-Backward). Cependant, les rapports de branchement attendus pour ces désintégrations à quatre corps sont *a priori* trop faibles pour espérer mesurer des asymétries avec une grande précision avec l'échantillon Run I de LHCb. Notre premier objectif est donc de mesurer les rapports de branchement totaux pour prédire ce que sera le potentiel de ces mesures dans le future, c'est à dire le Run II et l'Upgrade. Nous prévoyons également de mesurer ces rapports de branchement dans les régions hautes et basses du spectre $m(\mu^+\mu^-)$, pour également contraindre des modèles de NP.

La Chapitre 3 présente la recherche de désintégrations non-résonantes $D^+_{(s)} \to \pi^+\mu^+\mu^-$ et $D^+_{(s)} \to \pi^-\mu^+\mu^+$ à la quelles j'ai participé. Un échantillon de données est utilisé qui correspond à des collisions proton-proton à une luminosité intégrée de 1.0 fb^{-1}, enregistré à $\sqrt{s} = 7$ TeV par l'expérience LHCb en 2011. Les résultats ont été publiés dans [13].

Les désintégrations $D^+_{(s)} \to \pi^+\mu^+\mu^-$ sont recherchées dans cinq régions du spectre $m(\mu^+\mu^-)$ (Figure 2), pour séparer les régions sensibles à la Nouvelle Physique de celles où les contributions des résonances dominent.

J'étais en charge de l'évaluation de l'incertitude systématique de l'efficacité de la sélection. Cette incertitude est due au fait que la simulation Monte Carlo (MC) ne décrit pas parfaitement les données. La principale variable ($BDTCat$) utilisée dans la sélection combine les variables principales cinématiques et topologiques décrivant le signal. Les données et les simulations MC ont été comparées pour les distributions de ces variables ainsi que l'efficacité de la coupure sur $BDTCat$. La comparaison de l'efficacité obtenue pour le MC et les données est présentée sur la Figure 3. Elle a été utilisée dans la détermination de l'incertitude systématique.

Le chapitre suivant (Chapitre 4) détalle une série de *lignes*, que j'ai crées, dédiées pour les désintégrations en trois et quatre corps. L'analyse d'un mode de désintégration à LHCb n'est pas possible sans une sélection de déclenchement de haut niveau (HLT) adapté.

Le défi du système de déclenchement HLT est qu'un très grand nombre de modes doivent être sélectionnés pour les analyses. Ceci signifie que le déclenchement doit être efficace pour la sélection du signal mais aussi respecter les conditions suivantes :

- avoir *un taux de rétention* inférieur à la bande passante est limitée;
- avoir *un temps de calcul* inférieur au temps de traitement des données en moyenne par événement.

Ces aspects sont difficiles à réunir pour des désintégrations à plusieurs corps comme $D \to h(h')\mu\mu$, qui souffrent d'un bruit de fond combinatoire élevé et qui ne peuvent pas utiliser la détection d'une particule à la haute impulsion transversale comme d'autres modes.

La condition de temps de calcul est particulièrement difficile à satisfaire car trois ou quatre traces à sélectionner dans l'état final impliquent un grand nombre de combinaisons à traiter. Pour éviter les coupures strictes au niveau des traces affectant toutes les

traces de l'état final et causant une perte importante de l'efficacité, une stratégie en deux étapes est adoptée :

- dans un premier temps uniquement les événements avec une paire de muons sont présélectionnés. Grâce au petit nombre de candidats contenant une paire de muons par rapport au nombre de candidats contenant un ou plusieurs hadrons, une grande partie de bruit de fond est supprimée directement à cette étape là. La dernière est commune aux lignes sélectionnant les désintégrations en trois et quatre corps, $D \to h\mu\mu$ et $D \to hh'\mu\mu$.

- dans un second temps, les hadrons sont combinés pour former les candidats D. À cette étape, des sélections supplémentaires sont appliquées.

Le sélection de déclenchement de haut niveau n'est pas très stricte. Ainsi des lignes de *stripping* sont gérées centralement sur tout l'échantillon de données pour faire la présélection après laquelle l'échantillon peut être accédé régulièrement par un analyste. J'ai écrit et maintenu ces lignes pour les modes charmes rares en trois et quatre corps. Il s'agit d'une sélection similaire à la sélection finale. Elle également décrit dans le Chapitre 4.

Les lignes de déclenchement et de stripping mentionnées ci-dessus sont adaptées pour étudies les mésons D^0 de deux facons différentes :

- la première méthode utilise les D^0 produits dans les désintégrations $D^* \to D^0\pi^+$. Le étiquetage D^* donne le saveur du D^0 et aide beaucoup à réduire le bruit de fond combinatoire. Il est bien adapté pour la recherche de désintégrations très rares.

- la deuxième méthode utilise tous les D^0 produits directement dans les collisions proton-proton, quel que soit le mode de production. Ces analyses *non-étiquetées* sont moins propres que celles étiquetées, mais bénéficient d'une augmentation du taux de production [17]. La stratégie est bien adaptée au cas de désintégrations $D^0 \to K^-\pi^+\mu^+\mu^-$, où le rapport de branchement total devrait être de l'ordre de 10^{-6}, ce qui est suffisamment élevé pour garantir une pureté du signal acceptable même en présence de bruit de fond plus élevé.

Les rapports de branchement (ou leur limite supérieure) des désintégrations $D^0 \to h^\pm h^{(\prime)\mp}\mu^+\mu^-$ seront déterminés par la première facon dans les mois suivant cette thèse. Comme indiqué précédemment, cette thèse port principalement sur la mesure du rapport de branchement de désintégrations $D^0 \to K^-\pi^+\mu^+\mu^-$ dans la région de résonances ρ/ω du spectre $m(\mu^+\mu^-)$. C'est fait en utilisant les désintégrations non-étiquetées. Le mode $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ doit être mesuré car il servira de mode de normalisation pour tous les désintégrations $D^0 \to h^\pm h^{(\prime)\mp}\mu^+\mu^-$. Elle partage beaucoup de caractéristiques avec ces canaux. Cette normalisation réduit une grande partie des incertitudes systématiques liées à la reconstruction et à la sélection. La mesure de $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$ est également un laboratoire qui prépare les mesures

d'étiquetées D^*. En particulier, elle a permis le développement de méthodes pour l'étude des bruits de fond piqués qui sont présents dans les mesures de tous les modes. Ces bruits de fond affectent ces mesures car les désintégrations $D^0 \to K^-\pi^+\pi^+\pi^-$, $D^0 \to K^-K^+\pi^+\pi^-$, ou $D^0 \to \pi^+\pi^-\pi^+\pi^-$ sont généralement de 4 à 6 ordres de grandeur plus fréquentes que les modes $D^0 \to h^\pm h^{(\prime)\mp}\mu^+\mu^-$. Ils polluent quand deux pions dans l'état final sont identifiés comme des muons.

La mesure de $\mathcal{B}(D^0 \to K^-\pi^+\mu^+\mu^-)$ est décrite dans le Chapitre 5. Nous avons effectué la première mesure du rapport de branchement partiel pour la désintégration $D^0 \to K^-\pi^+\mu^+\mu^-$ dans la région de résonances ρ/ω définie comme $m(\mu^+\mu^-) \in [675; 875]$ MeV/c^2. Cette mesure a été effectuée sur des données correspondant à une luminosité intégrée de $2\,\text{fb}^{-1}$ collecté à $\sqrt{s} = 8$ TeV par LHCb en 2012.

Nous nous sommes concentrés sur la région ρ/ω pour deux raisons. Tout d'abord, le reste du spectre $m(\mu^+\mu^-)$ peut contenir de la NP, il ne peut donc pas être utilisé pour la normalisation car cette dernière servira justement à son étude. Deuxièmement, l'efficacité de la sélection varie beaucoup dans l'espace de phases, qui dans le cas d'une désintégration à quatre corps est décrit par 5 variables indépendantes (voir la Section 1.2.4). Un très large échantillon de MC n'étant pas disponible, seulement quelques bins par variable peuvent être utilisées pour déterminer cette efficacité. Ceci provoque une dépendance du modèle de désintégrations supposées dans la génération des événements dans la simulation. Aucun modèle fiable n'est connu à ce jour pour la désintégration $D^0 \to K^-\pi^+\mu^+\mu^-$. La réduction de l'espace de phase utilisé dans l'analyse permet de réduire l'impact de la variation de cette efficacité.

Le rapport de branchement dans la fenêtre ρ/ω devrait être $\mathcal{O}(10^{-6})$ [18], cette valeur est accessible par LHCb avec l'échantillon des données 2012. Pour augmenter la sensibilité de notre mesure, nous sélectionnons les candidats D^0 non-étiquetés, c'est à dire sans spécifier leur mode de production.

La mesure peut donc être résumée par l'équation ci-dessous :

$$\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)) = \mathcal{B}(D^0 \to K^-\pi^+\pi^+\pi^-) \times \frac{N_{signal}}{N_{norm}} \times \frac{\epsilon_{norm}}{\epsilon_{signal}} \quad (1)$$

où N (respectivement ϵ) est le nombre d'événements (respectivement l'efficacité totale) du signal ou de la normalisation. Le mode $D^0 \to K^-\pi^+\pi^+\pi^-$ est choisi comme un mode de normalisation car son rapport de branchement est haut et connu précisément, $\mathcal{B} = (8.287 \pm 0.043_{stat.} \pm 0.200_{syst.})$ % [19]. Un ingrédient essentiel de cette analyse est une mesure de rapport des efficacités entre $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ et $D^0 \to K^-\pi^+\pi^+\pi^-$. Grâce à sa ressemblance (topologique, cinématique, multiplicité des traces chargées) le rapport des efficacités devrait être proche de 1, par ailleurs, la plupart des incertitudes systématiques devraient être annulées.

Toutefois, il y a une différence essentielle car le mode de normalisation est un état final purement hadronique alors que le signal comporte une paire de muons. Cette différence implique essentiellement deux incertitude systématiques dues à :

- l'efficacité d'identification des muons;
- l'efficacité de la première étape du système de déclenchement (L0);

Ces deux point affectent le rapport des efficacités de l'Équation 5.3 car une partie des incertitudes systématiques ne peut pas être annulée.

Il est possible de limiter la première incertitude. Pour la désintégration $D^0 \to K^-\pi^+\pi^+\pi^-$, il faudrait utiliser une première étape du système de déclenchement (L0) adaptée pour sélectionner les hadrons (`L0HadronTOS`). Alors que pour la désintégration $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ il faudrait utiliser le système déclenchement adapté aux muons (`L0MuonTOS`). Cette différence est *a priori* la plus gênante. Ces deux lignes de déclenchement utilisent les informations de sous-détecteurs différents (HCAL, ECAL et des stations à muons) et avec des seuils différents pour les impulsions transversales. Par conséquent, les efficacités sont différentes pour les modes hadroniques et muoniques et ne s'annulent pas dans le rapport des efficacités (Équation 5.3). On sait par ailleurs que ces distributions ne sont pas bien décrites par la simulation. Donc, pour réduire cette différence entre les canaux nous avons besoin que ces deux modes soient déclenchés d'une même manière qui ne dépende pas de la désintégration. Nous avons ainsi choisis `L0HadronTIS` qui respecte cette condition puisqu'il est déclenché par un hadron qui ne vient pas de l'état final de $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ ni de celui de $D^0 \to K^-\pi^+\pi^+\pi^-$, mais du reste de l'événement. Cette approche a une efficacité plus faible mais est plus robuste pour annuler les effets systématiques.

La détermination des efficacités liées à l'identification des muons utilise une méthode basée sur les données qui est présentée dans la Section 5.5.1.

Une autre point important de cette analyse et présenté dans la Section 5.4.2 est la sélection multi-variable. L'outil utilisé est appelé *Boosted Decision Tree (BDT)*, son principe est expliqué dans la thèse.

Le bruit de fond piqué est causé lorsque deux poins de la désintégration $D^0 \to K^-\pi^+\pi^+\pi^-$ sont identifiés comme muons.Il pique environ $20 MeV/c^2$ en-dessous du pique de signal D^0 dans la distribution de masse invariante. Une variable d'identification des muons, $ProbNNmu$ [1], est utilisée pour réduire ce bruit de fond. La contamination résiduelle est comptée dans l'ajustement du maximum de vraisemblance de l'échantillon de signal utilisé pour obtenir $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$. La forme de ce bruit de fond n'est pas triviale à déterminer car les pions peuvent se désintégrer en vol en $\mu^+\nu_\mu$ ceci provoque des queues qui ne peuvent être étudiées précisément dans le MC en raison de la rareté de la double misidentification $\pi \to \mu$.

L'analyse a été réalisée en aveugle donc plusieurs échantillons de contrôle sont utilisés pour guider sa conception. Une partie de l'échantillon de signal (20 % de l'échantillon de données) est utilisé pour optimiser la sélection. Cet échantillon permet de trouver le meilleur rapport signal sur bruit en fonction de la coupure sur des variables telles que BDT et $ProbNNmu$. Un échantillon obtenu avec la même sélection que l'échantillon de signal mais où la conditions `L0HadronTIS` est remplacée par la conditions `L0HadronTOS ou L0MuonTOS` qui a une efficacité plut haut, est utilisé pour un certain nombre de contrôles croisés. La moitié de l'échantillon des désintégrations $D^0 \to K^-\pi^+\pi^+\pi^-$ est également utilisée en tant que mode de contrôle.

L'ajustement final des données est présenté Figure 4, il est décrit dans la Section 5.7.

Le rapport de branchement partiel est alors obtenu :

$$\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)) = (4.37 \pm 0.12_{stat.} \pm 0.53_{syst.}) \times 10^{-6},$$

où la première erreur est l'erreur statistique et la seconde systématique.

[1] Voir la Section 2.9.2

J'ai également effectué une partie importante de l'analyse de D^*-étiqueté dans les désintégrations $D^0 \to h^\pm h^{(\prime)\mp}\mu^+\mu^-$ où $D^{*+} \to D^0\pi^+$. Les résultats sont présentes dans la Chapitre 6.

Nous visons à rechercher et éventuellement à mesurer certaines propriétés de désintégrations charmées rares $D^0 \to K^-\pi^+\mu^+\mu^-$, $D^0 \to K^+\pi^-\mu^+\mu^-$ et $D^0 \to K^+K^-\mu^+\mu^-$. Parce que leurs rapports de branchement non-résonants sont très petits $\mathcal{O}(10^{-9})$, il est crucial de développer une sélection performante. Une des variables principale contre le bruit de fond combinatoire est: $\Delta m = m(D^{*+}) - m(D^0)$. Une analyse multivariée (MVA) est également un bon outil. Nous avons considéré plusieurs MVA, qui sont détaillés dans la Section 6.1, pour finalement choisir la plus optimale. Dans la Section 6.2, pour la MVA choisie nous donnons la prédiction des sensibilités des rapports de branchement et des asymétries qui doivent être atteints par LHCb pour les Run I, Run II et enfin pour l'ère du Upgrade de LHCb, qui va commencer en 2019.

J'ai également contribué à l'Upgrade de LHCb. En effet, j'ai participé aux tests en faisceau du premier prototype de Front-End Board (FEB) des calorimètres. Ceci est décrit dans le Chapitre 7. Les deux calorimètres utilisent la même électronique et cette dernière doit être remplacée.

Pendant la période d'Upgrade les calorimètres électromagnétiques (ECAL) et hadroniques (HCAL) contribueront à l'identification des particules et au Low Level Trigger (LLT), le nouveau système de déclenchement qui filtrera les événements de données brutes. Pour une sélection par le LLT efficace, la lecture de l'ECAL et du HCAL doivent être effectuées à 40 MHz.

Le premier prototype de la FEB a été testé en faisceau au CERN en Novembre 2012 dans des conditions proches des celles de l'expérience.

Les mesures du bruit, de la résolution, de la linéarité, du spill-over et de la taille du plateau du signal intégré montrent que le comportement de la FEB se situe dans les spécifications, en exception pour la largeur de l'impulsion du signal, tel que révélé par les tests. Pour éliminer ce problème, il a été proposé de mettre en œuvre un filtre pôle-zéro (PZ) qui permet de réduire la queue de la distribution et de satisfaire les spécifications. Les filtres pôle-zéro sont installés après l'amplificateur et avant les intégrateurs. Ils sont constitués de deux résistances et d'un condensateur. Il supprime certaines fréquences de l'impulsion et améliore de cette manière la forme du signal. Les simulations indiquent, en effet, que la forme de signal devrait être considérablement améliorée avec le filtre pôle-zéro. La forme du signal obtenue à partir des simulations avec le PZ est présentée dans la Figure 5 ainsi que la forme sans le filtre et la forme correspondant au FEB utilisé actuellement.

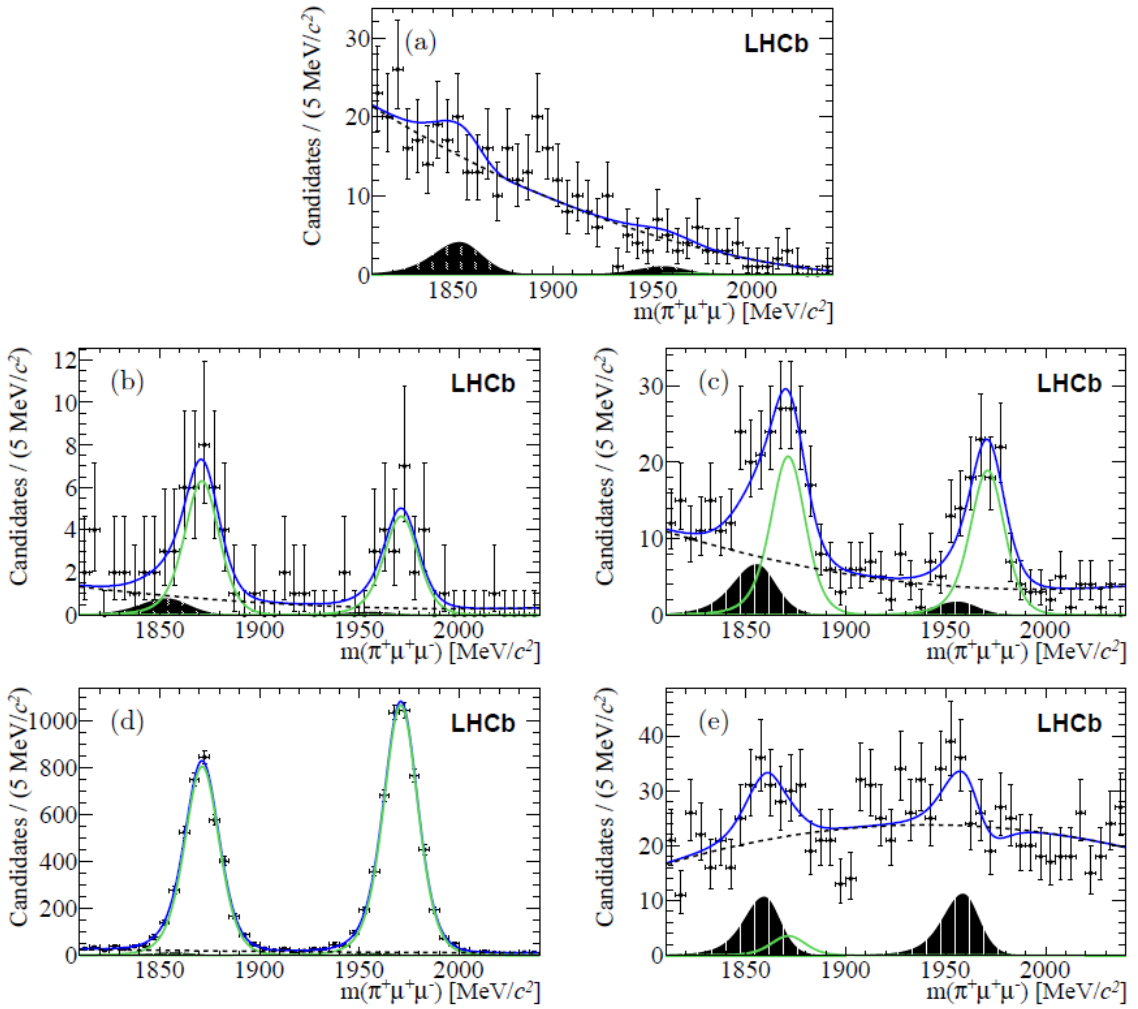


Figure 2: Distributions de la masse invariante de $D^+_{(s)} \to \pi^+\mu^+\mu^-$ de candidats dans les cinq régions définies par la masse de la paire $\mu^+\mu^-$, $m(\mu^+\mu^-)$. (a) pour 250 MeV/c^2 $< m(\mu^+\mu^-) <$ 525 MeV/c^2 (low-$m(\mu^+\mu^-)$), (b) pour 525 MeV/c^2 $< m(\mu^+\mu^-) <$ 565 MeV/c^2 (η), (c) pour 565 MeV/c^2 $< m(\mu^+\mu^-) <$ 850 MeV/c^2 (ρ/ω), (d) pour 850 MeV/c^2 $< m(\mu^+\mu^-) <$ 1250 MeV/c^2 (ϕ) et (e) pour 1250 MeV/c^2 $< m(\mu^+\mu^-) <$ 2000 MeV/c^2 (haut-$m(\mu^+\mu^-)$). Les données sont représentées par des points noirs et leur ajustement (PDF) par la courbe bleue foncé. Les composantes de l'ajustement sont également présentées : pour le signal (vert clair), pour le bruit de fond piqué (zone solide) et pour le bruit de fond non-piqué (pointillés).

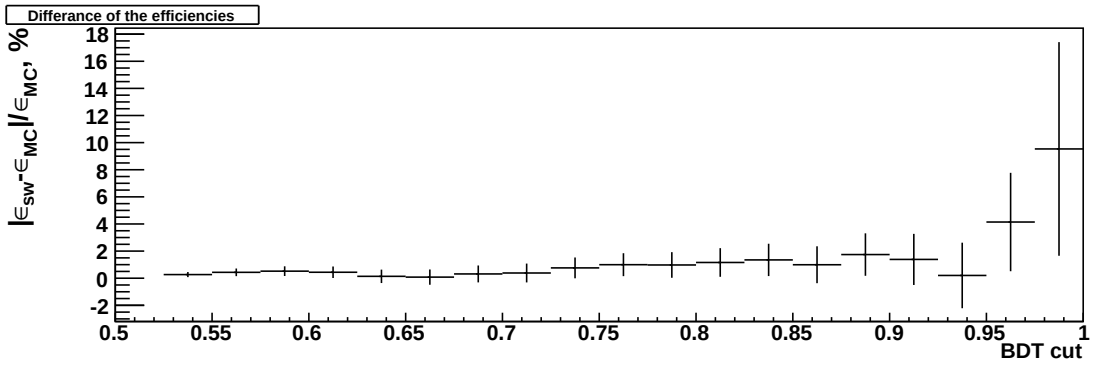


Figure 3: Différence relative de l'efficacité dans les données et le MC en fonction de la valeur de la coupure sur $BDTCat$. Ici, pour la désintégration $D^+ \to \pi^+\phi(\to \mu^+\mu^-)$. La coupure utilisée dans l'analyse est $BDTCat \geq 0.9$.

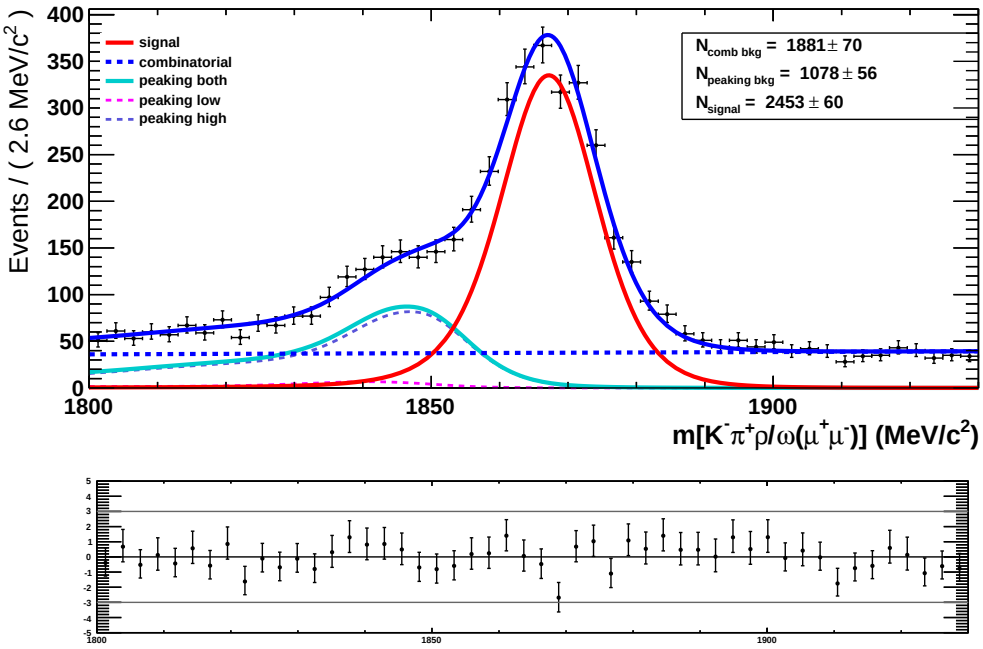


Figure 4: Distribution de la masse invariante du $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ dans l'échantillon de données, sélectionné avec `L0HadronTIS`. La courbe rouge solide représente le signal, celle en pointillés bleus le combinatoire, celle en cyan solide le bruit de fond piqué, les courbes violeté et magenta sont les composants de "highPT" et "lowPT" de la forme du bruit de fond piqué. Les paramètres des ajustements peuvent être trouvés dans le Tableau 5.14 du Chapitre 5.

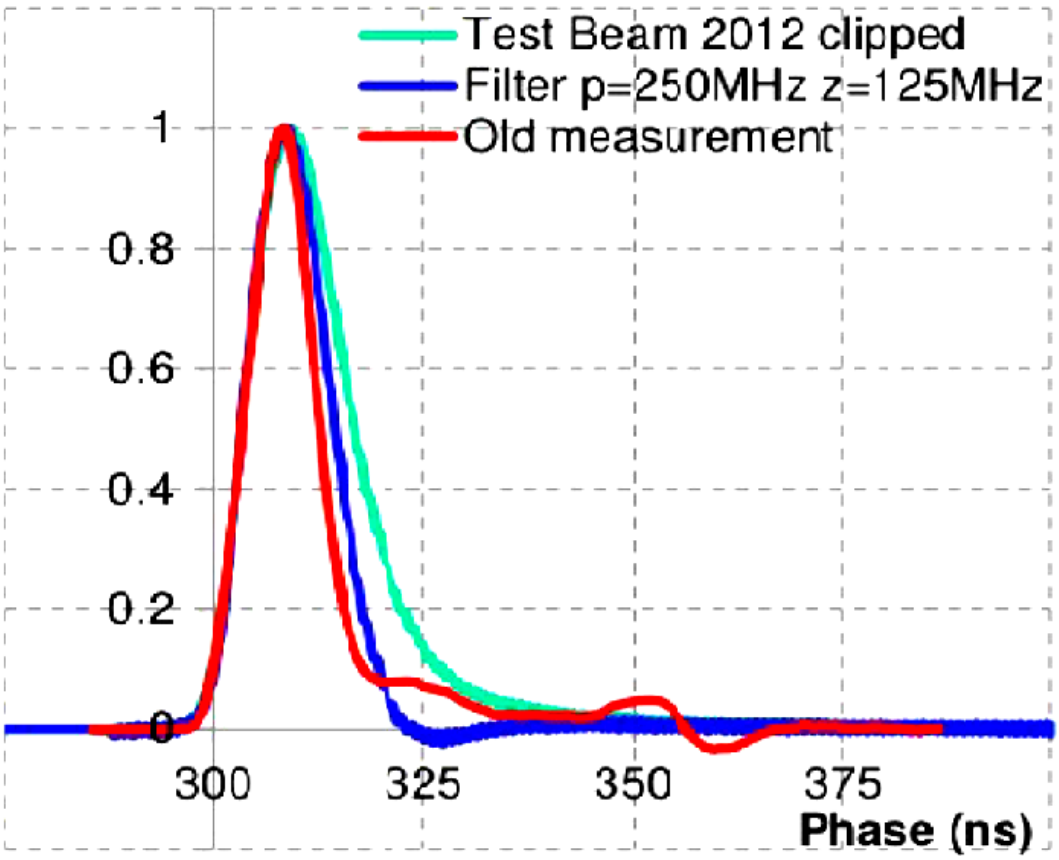


Figure 5: Forme du signal obtenue pendant les test en faisceau avec la nouvelle FEB (cyan), des mesures avec la FEB actuellement utilisée par LHCb (rouge) et avec le filtre à pôle zéro (bleu) simulé après les tests en faisceau.

Introduction

The main subject of this thesis is the study of four-body rare charm decays into two hadrons and two muons $D^0 \to h^{\pm}h^{(\prime)\mp}\mu^+\mu^-$. This comprises $D^0 \to K^-\pi^+\mu^+\mu^-$, $D^0 \to \pi^+\pi^-\mu^+\mu^-$, $D^0 \to K^+K^-\mu^+\mu^-$ and $D^0 \to K^+\pi^-\mu^+\mu^-$.

At short distance (SD), rare charm decays proceed mostly through the $c \to u$ Flavor Changing Neutral Current (FCNC), which is possible only at loop diagram level in the Standard Model (SM). Moreover, they are highly suppressed by the GIM mechanism. Consequently, rare charm decays are good probes of New Physics (NP). New particles can contribute in loop diagrams and become detectable by enhancing branching ratios with respect to the SM predictions.

However, it is not easy to access SD physics in these decays. Indeed, their total branching ratios are dominated by long distance (LD) contributions due to resonant intermediate states, like $D \to XV'(\to \mu^+\mu^-)$. A possible way to probe them is to measure partial branching ratios in extreme regions of the $m(\mu^+\mu^-)$ spectrum, where the resonant contributions are minimal. Another possibility is to search for sizeable asymmetries, like T-odd or Forward-Backward asymmetries. In this case, even the resonant regions dominated by LD physics can be used. However, it is known from theoretical and experimental constraints that SD branching ratios should not be higher 10^{-8}, and that asymmetries should not exceed a few %. A short theoretical introduction and a panorama of NP predictions are presented in Chapter 1 to help understand the interest of rare charm decays.

The LHCb experiment (Chapter 2) gives a unique opportunity to study such small effects in rare charm decays. Indeed, the charm production cross section is large in proton-proton collisions at a centre-of-mass energy of $\sqrt{s}$=7 or 8 TeV.

Several rare charm decays have been searched for at LHCb recently. The following upper limits have been obtained:

- $\mathcal{B}(D^0 \to \mu^+\mu^-) < 7.6 \times 10^{-9}$ @ 95% CL [12].
- $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-) < 8.3 \times 10^{-8}$ @ 95% CL [13].
- $\mathcal{B}(D^0 \to \pi^+\pi^-\mu^+\mu^-) < 6.7 \times 10^{-7}$ @ 95% CL [14].

These are about two orders of magnitude better than the previous limits obtained by the Flavour Factories, Tevatron or E971 experiments. My participation in [13] is described in Chapter 3.

We expect the same improvement with the LHCb Run I dataset in the case of the $D^0 \to h^\pm h^{(\prime)\mp}\mu^+\mu^-$ decays, which constitute the heart of the work performed during this thesis. Present limits, due to E791 [16], are

$$\mathcal{B}(D^0 \to K^+K^-\mu^+\mu^-) < 3.3 \times 10^{-5} \text{ (at 90\% C.L.)}$$
$$\mathcal{B}(D^0 \to K^-\pi^+\mu^+\mu^-) < 3.59 \times 10^{-4} \text{ (at 90\% C.L.)}$$

Asymmetries also seem to be useful probes in these modes, where more observables are available thanks to the final states multiplicity (T-odd, Forward-Backward asymmetries). However, the expected branching ratios of the four body decays are *a priori* too low to hope for a precise asymmetry measurement with LHCb Run I dataset. Our first goal is therefore to measure the total branching ratios in order to predict what will be the potential for such measurements with future LHCb datasets (from Run II and Upgrade). We also plan to measure these branching ratios in high and low regions of the $m(\mu^+\mu^-)$ spectrum, in order to also constrain New Physics models.

This thesis presents my contributions to the $D^0 \to h^\pm h^{(\prime)\mp}\mu^+\mu^-$ program outlined in the previous paragraphs.

The analysis of a given decay mode at LHCb is not possible without an adapted High Level Trigger selection. I designed a series of dedicated *lines* for the three and four body decays. This is described in Chapter 4. The amount of data recorded by LHCb after triggering is large. Hence offline *Stripping* lines are centrally run on the full dataset to pre-select the sample to which an individual analyst can access routinely. I wrote and maintained such lines for three and four body rare charm modes. This involves a selection already close to the final offline selection. This is also described in Chapter 4.

The Trigger and Stripping lines mentioned above are adapted to the two most common ways to study D^0 decays:

- by using D^0's produced in $D^* \to D^0\pi^+$ decays. The D^*-tag provides the flavour of the D^0 and helps a lot to reduce the background. It is well suited for searches for very rare decays.

- by using any D^0 directly produced at the proton-proton collision vertex, whatever its exact production mode. These *untagged* analyses are less clean than the tagged ones, but benefit from a higher production rate [17]. The strategy is well suited to the case of $D^0 \to K^-\pi^+\mu^+\mu^-$, where total branching ratio, expected to be a few 10^{-6}, is large enough to guarantee an acceptable signal purity even with higher backgrounds.

The $D^0 \to h^\pm h^{(\prime)\mp}\mu^+\mu^-$ branching ratios (or the upper limits) that will be determined in the months following this thesis will employ D^*-tagged analyses. Before that, we measure the branching ratio of the $D^0 \to K^-\pi^+\mu^+\mu^-$ decay in the ρ/ω region of the m$(\mu^+\mu^-)$ spectrum. This untagged measurement is the main work presented in this thesis. This mode had to be measured since it will serve as a normalization mode for

all the other $D^0 \to h^{\pm}h^{(\prime)\mp}\mu^+\mu^-$ branching ratios. This mode shares a lot of features with the other $D^0 \to h^{\pm}h^{(\prime)\mp}\mu^+\mu^-$ channels. This cancels the greatest part of the reconstruction-related systematic uncertainties. This measurement is also a laboratory that prepares the D^*-tagged measurements. In particular, it allowed to design the treatment of the peaking backgrounds that affect all the modes. They are among the main difficulties affecting these analyses since $D^0 \to K^-\pi^+\pi^+\pi^-$, $D^0 \to K^-K^+\pi^+\pi^-$, or $D^0 \to \pi^+\pi^-\pi^+\pi^-$ decays are typically 4 to 6 orders of magnitude more frequent than the signal modes they pollute when two of their final states pions are misidentified as muons. The untagged measurement of $\mathcal{B}(D^0 \to K^-\pi^+\mu^+\mu^-)$ is described in Chapter 5.

I also carried out an important part of the D^*-tagged analyses of the $D^0 \to h^{\pm}h^{(\prime)\mp}\mu^+\mu^-$ modes. Besides the trigger and stripping lines mentioned above, I designed the rest of the offline analysis. I applied and compared several Multivariate approaches for that purpose. This is described in Chapter 6, where I also study the potential of the future LHCb datasets for the measurement of branching ratios and asymmetries. This includes the Upgrade of LHCb.

My work is connected to the Upgrade of LHCb in a second way: I participated in the beam test of the first prototype of the calorimeter system Front-End Board. This is described in Chapter 7.

Chapter 1

Theoretical overview

In this chapter a theoretical overview is presented. Currently the Standard Model (SM) synthesizes our understanding of elementary particles and fundamental interactions, electromagnetic, strong and weak, except for the gravitation. The SM is a gauge quantum field theory. Its success is outstanding: it is able to account for all the phenomena observed at accelerator-based experiments.

1.1 The Standard Model

1.1.1 Short overview

Everything in the Universe is made of a few basic building blocks, fundamental particles, governed by four fundamental forces. To date, these particles and interactions are best understood in the framework of the Standard Model of particle physics.

Since the early 20th century, a wide range of particles have been observed. Science attempted to classify them in a similar manner as was done for chemical elements in the periodic Mendeleev's table. The fundamental particles in the SM are 6 quarks, 6 leptons, 4 gauge bosons, one Higgs boson (Figure 1.1) and their anti-particles.

Each fundamental particle has a set of quantum numbers (Table 1.1), characterizing their interactions. All subatomic particles, composed and elementary, can be classified by their spin and corresponding statistics (Figure 1.2). Particles holding half-integer spins, fermions, are characterized by Fermi-Dirac statistics and follow the Pauli exclusion principle, unlike the bosons, particles with integer spins, that obey Bose-Einstein statistics. Among elementary particles, the roles of fermions and bosons are different. Fermions are the elements of the matter.

The fundamental fermions are classified as quarks and leptons. The latter carry no colour charge and they are not involved in the strong interaction while quarks do. Quarks cannot be observed individually but they are variety of observable particles, hadrons - mesons (made of a quark and an anti-quark) and baryons (three-particle combination of quarks and antiquarks).

Particle	m, MeV/c^2	J	Q	L	B	S	C	B'	T
u	2.3	1/2	+2/3	0	1/3	0	0	0	0
d	4.8	1/2	-1/3	0	1/3	0	0	0	0
s	95	1/2	-1/3	0	1/3	-1	0	0	0
c	$1.3 \cdot 10^3$	1/2	+2/3	0	1/3	0	+1	0	0
b	$4.18 \cdot 10^3$	1/2	-1/3	0	1/3	0	0	-1	0
t	$173 \cdot 10^3$	1/2	+2/3	0	1/3	0	0	0	+1
e	0.51	1/2	-1	1	0	0	0	0	0
ν_e	$<$2.2 eV/c^2	1/2	0	1	0	0	0	0	0
μ	105.66	1/2	-1	1	0	0	0	0	0
ν_μ	$<$170 keV/c^2	1/2	0	1	0	0	0	0	0
τ	1776.82	1/2	-1	1	0	0	0	0	0
ν_τ	$<$15.5	1/2	0	1	0	0	0	0	0
g	0	1	0	0	0	0	0	0	0
γ	0	1	0	0	0	0	0	0	0
$W^\pm$	$80.3 \cdot 10^3$	1	0	0	0	0	0	0	0
Z^0	$91.2 \cdot 10^3$	1	0	0	0	0	0	0	0
H	$125.9 \cdot 10^3$	0	0	0	0	0	0	0	0

Table 1.1: Main quantum numbers for fundamental particles. In the table m - mass, J - spin, Q - electric charge, L - lepton number, B - baryon number, S - strangeness, C - charm, B' - beauty, T - topness.

Both quarks and leptons are classified in three families or generations:

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix} \quad \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix} \quad \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix} \\ \begin{pmatrix} u \\ d \end{pmatrix} \quad \begin{pmatrix} c \\ s \end{pmatrix} \quad \begin{pmatrix} t \\ b \end{pmatrix} \tag{1.1}$$

this derives from the fact that at first order weak interaction couples only quarks and leptons of the same family. The SM accounts for this with a gauge symmetry that acts on doublets, like those presented above (Equation 1.1). An intriguing pattern also exists here: a mass hierarchy between families. The $(u,\ d,\ , e^-, \nu_e)$ are lighter than $(s,\ c,\ , \mu^-, \nu_\mu)$, which are far lighter than the third family. Such behaviour is not understood in the framework of the SM.

Moreover, fermions are classified by their electric charge as well. Quarks are of two types: up-type quarks, u, c and t, with charge $\frac{2}{3}$ and down-type quarks, d, s and b, with charge $-\frac{1}{3}$. Leptons are also classified by electric charge: charged e, μ and τ and neutral ν_e, ν_μ and ν_τ. Neutrinos have an important typical property that they have been observed only in one helicity state: left - for neutrino and right - for anti-neutrino.

Fundamental bosons carry interaction between fermions. At present, the existence of four fundamental forces are known:

- Gravitational;

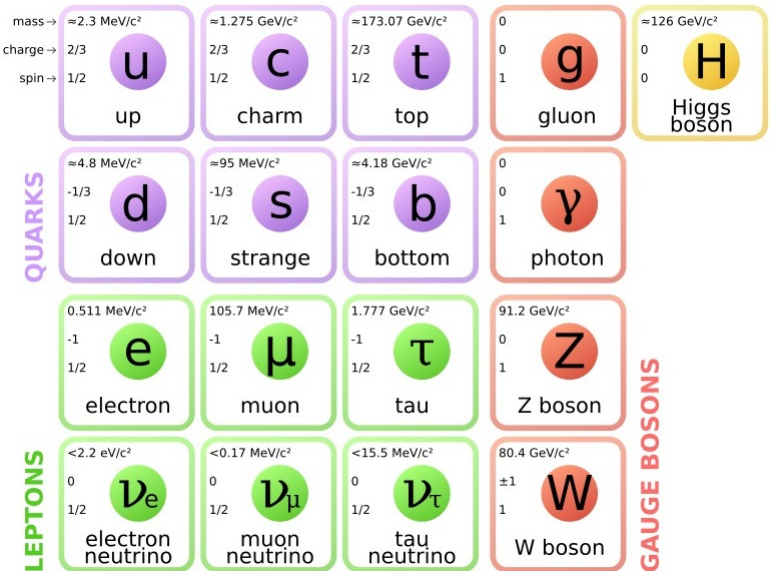


Figure 1.1: The Standard Model fundamental particles.

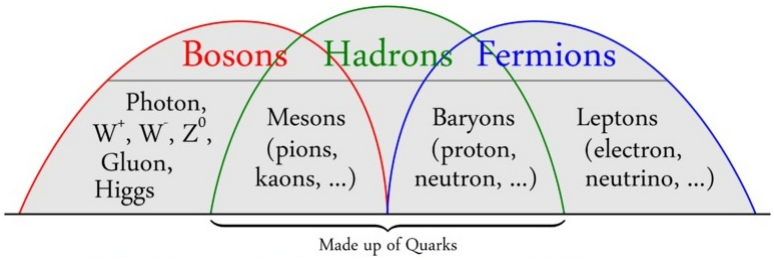


Figure 1.2: The Standard Model fundamental particles classification.

- Electromagnetic;
- Strong;
- Weak.

The electromagnetic force manifests itself, for example, in an interaction of two electron shells and electron-nucleon, making possible chemical reactions. It acts between all objects with an electric charge by exchanging massless and electrically neutral photon. The relativistic quantum field theory describing the electromagnetic force is called quantum electrodynamics or QED.

The strong interaction is observable in two regimes: on a larger scale (about 1 to 3 fm) - attraction between nucleons in a nucleus and on a smaller scale (the radius of a nucleon) - interaction between quarks inside of a nucleon. Quarks and gluons, participating in the strong interaction, are the only fundamental particles which carry non-vanishing colour charge. The interaction and behaviour between them are described by Quantum Chromodynamics (QCD). The strength of interaction is governed by the strong coupling constant. It has two important specific properties:

- it increases in long distances, compared to size of the nucleon, causing *confinement*.

It implies that the quarks are never observed individually. At the same time uncoloured combination of quarks, hadrons, are not affected.

- it decreases at short distances causing *asymptotic freedom*. This means that the interaction between quarks becomes arbitrarily small as the distance between the particles decreases. In other words, in the asymptotic limit $r \to 0$ the particles cease to interact and can be considered as free.

The weak interaction is responsible for nuclear beta decays and other similar processes. It affects all particles, quarks and leptons. One can distinguish two types of weak interactions: through neutral and charged currents. Indeed, transmitters of the interaction are $W^{\pm}$ and Z^0 vector bosons. Since they are not massless particles, the uncertainty principle dictates a finite interaction range, $\sim 10^{-18}$ m.

A similar approach applies to the description of all interactions in the quantum relativistic theory: (i) fermions subject to a given interaction must possess a "charge", conserved during the interaction; (ii) the exchange of energy, momentum and charge proceed thought the intermediate boson between two interacting fermions (Figure 1.3). The main difference between fundamental interactions is their strength and range, summarized in Table 1.2.

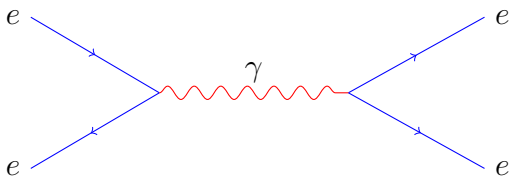


Figure 1.3: Feynman diagram illustrating the interaction between two charged fermions (e) though a boson (γ) exchange.

At the scale of particle physics the gravitational interaction is neglected as interacting objects have too small masses. The Standard Model unifies the three other fundamental forces: electromagnetic, weak and strong.

interaction	mediating particles	charge	relative strength	range (m)
Strong	gluon	colour	1	10^{-15} ($\sim d_{nucleus}$)
Electromagnetic	photon	electric charge	$\frac{1}{137}$	∞
Weak	$W^{\pm}$, Z^0	flavour	10^{-6}	10^{-18} ($\sim 0.1\%\ d_{proton}$)
Gravitation	graviton(?)	mass	6×10^{-39}	∞

Table 1.2: The known fundamental interactions and their typical charges, transmitting particles, relative strengths and ranges.

1.1.2 The symmetries

Symmetries play a fundamental role in particle physics. The invariance or non-invariance of physics laws under a given transformation helps to characterize them and dictates the explicit form of the mathematical objects (for example, Lagrangian $\mathcal{L}$) describing them theoretically. A symmetry is a property of a physical system or mathematical object to be "preserved" under some transformation (rotation, translation, gauge etc.). Mathematically it is represented by a unitary operator U: $|\psi_i'\rangle = U|\psi_i\rangle$, where $|\psi_i'\rangle$ is the initial state of the system transformed from an initial state $|\psi_i\rangle$ by the operator U. These two initial states ψ_i' and ψ_i evolving in time become final states ψ_f' and ψ_f. Such a time evolution is described by the Hamiltonian H. In a presence of a symmetry: $|\langle\psi_f|H|\psi_i\rangle|^2 = |\langle\psi_f'|H|\psi_i'\rangle|^2$. In this case, H and U commute.

As stated by Nöether's theorem [20], for any transformation under which the Lagrangian of a system is invariant, there is a corresponding conservation law. This way rotational invariance leads to angular momentum conservation, translational invariance to linear momentum conservation, time translation invariance to energy conservation, etc.

The electric charge conservation in electromagnetic interaction guided the design of the corresponding Lagrangian, which must be invariant under local gauge transformation. This criterion allows to determine this Lagrangian entirely.

If a transformation is continuous, the conservation law is additive and the sum of quantum numbers is conserved in a reaction. In the Table 1.3 such quantum numbers and conservation laws for different interactions are presented.

Quantum number	Strong	E/m	Weak
Electric charge, Q	+	+	+
Energy, E	+	+	+
Momentum, $\vec{p}$	+	+	+
Spin, $\vec{J}$	+	+	+
Baryon number, B	+	+	+
Lepton numbers, L_e, L_μ, L_τ	+	+	+
Strange, s	+	+	−
Charm, c	+	+	−
Beauty, b	+	+	−
Top, t	+	+	−
Strong isospin, I	+	−	−
Third projection of strong isospin, I_3	+	+	−

Table 1.3: Quantum numbers for different interactions, "+" means that the quantum number is conserved in the interaction, "−" - is not conserved.

Invariance under discrete transformations also leads to corresponding quantum numbers. There are three specific discrete symmetries important for the SM:

- Parity, denoted by P, sends $(\overrightarrow{x}, t) \to (-\overrightarrow{x}, t)$, reversing the handedness of space. Thus, the direction of motion $\overrightarrow{p} \to -\overrightarrow{p}$ and helicity $h = \frac{\overrightarrow{p}\cdot\overrightarrow{s}}{|\overrightarrow{p}|}$, where $\overrightarrow{s}$ is a spin, are also reversed.

- Time reversal, denoted by T, sends $(\overrightarrow{x}, t) \to (\overrightarrow{x}, -t)$, interchanging the forward and backward light-cones. In other words, an initial and final states of the system are reversed under the T transformation: $A + B \to C + D \Rightarrow C + D \to A + B$. Time reversal changes the direction of motion $\overrightarrow{p} \to -\overrightarrow{p}$ and spin $\overrightarrow{s} \to -\overrightarrow{s}$, leaving helicity unchanged.

- Charge conjugation, denoted by C. It is nonspacetime operation, under which particles and antiparticles are interchanged, leaving unchanged momenta $\overrightarrow{p}$, spins $\overrightarrow{s}$, masses.

Until 1956 these transformations were believed to be symmetries of nature. The discovery of P and C, as well as combined CP, violation [21–24] revealed the nontrivial behaviour of the weak interaction, which consequently carries important information about the form of the SM Lagrangian.

Of fundamental importance is the CPT combination, which is an exact symmetry [25] in any Lorentz-invariant quantum field theory with a hermitian Hamiltonian. Invariance under CPT leads to the fact that any process must go with the same probability as its transformed process, i.e. where all particles switched to their antiparticles, spin projections has an opposite sign and final and initial states are swapped. This implies that masses, life times and absolute values of electric charge and magnetic moment of particles and antiparticles must be the same.

All interactions, except the weak, are invariant under all discrete and combined CP transformations (see Table 1.4).

Quantum number/characteristic	Strong	E/m	Weak
Parity, P	+	+	−
Charge parity, C	+	+	−
Time reversal invariance, T	+	+	−
Combined parity, CP	+	+	−
CPT invariance	+	+	+

Table 1.4: Quantum numbers/characteristics and conservation laws for different interactions.

1.1.3 The SM Lagrangian

The observation of decays such as $\pi^+ \to \mu^+\nu_\mu$, i.e. $u\bar{d} \to \mu^+\nu_\mu$, that show the coupling of two different quarks and two different leptons, suggests that the gauge symmetry here should act on doublets. This is unlike in electromagnetism, where particles couple only to themselves, their antiparticles and bosons, where a singlet description is enough. The $SU(2)$ symmetry is well suited to deal with doublets.

However, neutral currents are also observed, therefore the weakly interaction particles must also appear in singlets. We would also like to incorporate the electromagnetic interaction to weak theory [26–28]. This is possible thanks to $U(1)$ gauge group with the associated quanta, weak hypercharge $Y = 2(Q - T_3)$, where Q is electric charge and T_3 is the third component of weak isospin.

Another crucial observation is the existence of only left (right) handed particles (antiparticles) in charge current processes, inferred from the angular distributions of the decay products. As a consequence we have the theory in which $SU(2)$ group acts only on doublets of left-handed fields. The gauge group used to build the electroweak interaction is therefore $SU(2)_L \times U(1)_Y$. Built this way, the theory accounts for the maximal Charge and Parity violations observed experimentally in the $\pi^+ \to \mu^+\nu_\mu$ decays, that occur only with a left-handed neutrino. Its P-conjugate with a right-handed ν is never observed. The C transformation of the process leads to a decay with a left-handed $\bar{\nu}$, which again does not exist in nature. The maximal violation of P and C came as a surprise and no fundamental physical principals predicted this.

Baryon states consisting of three fermions with the same quantum numbers are possible in nature, which contradicts to Pauli law. To avoid this problem, a new quantum number was introduced, colour. A colour triplet should be constructed in the theory. It should be invariant under $SU(3)_C$ gauge transformation. The basic $SU(3)_C$ state is the triplet with three colour components labelled as red, green and blue. All observed mesons are composed of a quark and an anti-quark with opposite colour (e.g., red and anti-red) and baryons composed of three particles, quarks and antiquarks, with three different colours yielding a "white" state. All leptons are colour singlets.

The Standard Model incorporates these three interactions. Thus, $U(1)_Y \times SU(2)_L \times SU(3)_C$ gauge invariance is needed. The full Lagrangian of SM can be presented as:

$$\begin{aligned} \mathcal{L} = \sum_{\substack{gauge \\ bosons}} -\tfrac{1}{4} F_{\mu\nu}F^{\mu\nu} + \sum_{fermions} \bar{\psi}(\gamma^\mu D_\mu)\psi + \\ +\mathcal{L}_{Yukawa} + \\ +(D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi) \end{aligned} \quad , \tag{1.2}$$

where ψ are fermion fields, Φ is the Higgs doublet (see Section 1.1.4), $F_{\mu\nu}$ are gauge bosons fields, $\mathcal{L}_{Yukawa}$ is Yukawa component of the Lagrangian (see Section 1.1.4). The first and second terms are gauge boson and fermion kinematic terms, the third is the fermion mass term and the fourth is the Higgs boson field term. These two last lines in Equation 1.2 are responsible for the generation of masses, flavour and CP violation (see Section 1.1.4).

The covariant derivative, D_μ, is defined this way:

$$D_\mu = \partial_\mu + ig_1 B_\mu Y + ig_2 \frac{\overrightarrow{\tau}}{2} \cdot \overrightarrow{W^\mu} + ig_3 \frac{\lambda_a}{2} G_a^\mu, \tag{1.3}$$

where g_1, g_2 and g_3 are the electroweak and strong coupling constants, and B^μ, W^μ and G^μ are the gauge boson fields. The coefficients are Y, the weak hypercharge, which is different for left and right handed fermions, $\overrightarrow{\tau}$, the Pauli matrices, and λ_a, the 8 generators of the $SU(3)$ colour group, which correspond to 8 gluons. The covariant derivative is built in order to ensure the invariance of the Lagrangian under the $SU(3) \times SU(2) \times U(1)$ transformation. It also introduces the gauge bosons and their couplings to the fermions.

Experimentally all known quark and leptons are observed to be either electroweak $SU(2)$ right-handed singlets or left-handed doublets

$$\psi_R = \{l_R^-,\ u_{R\alpha},\ d_{R\alpha}\}, \qquad \psi_L = \left\{ \begin{pmatrix} \nu_l \\ l^- \end{pmatrix}_L,\ \begin{pmatrix} u_\alpha \\ d_\alpha \end{pmatrix}_L \right\}. \tag{1.4}$$

The l denotes a lepton, and u and d denote quarks of up- or down-types. An additional index α is needed to describe quark transformation in the $SU(3)$ colour space.

1.1.4 Higgs mechanism

The $\mathcal{L}$ (Equation 1.2) is the result of the gauge invariance. If the mass of the fermions and bosons is introduced by simply adding term of the form $m\bar{\psi}\psi$, the gauge invariance is lost because $m\bar{\psi_R}\psi_L$ or $m\bar{\psi_L}\psi_R$ term appears, where left- and right-handed fermions are not transformed the same way by $SU(2)_L \times U(1)_Y$. This can be solved by assuming that the physics observed is the result of a spontaneously broken $SU(2)_L \times U(1)_Y$ symmetry. This means that the gauge symmetry is valid for the Lagrangian in Equation 1.2, but not for the ground state of the system.

To demonstrate this we consider the potential from Equation 1.2. The most general, non-trivial, renormalizable potential should has the form

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \frac{1}{4}\lambda(\Phi^\dagger \Phi)^2 \tag{1.5}$$

To get the minimum of the potential $V(\Phi)$ one needs to solve:

$$\Phi^\dagger \Phi(\mu^2 + \frac{1}{4}\lambda(\Phi^\dagger \Phi)^2) = 0. \tag{1.6}$$

Two possible solutions exist:

- $\mu^2 > 0$ gives the trivial solution $\Phi^\dagger\Phi$=0
- $\mu^2 < 0$ provides:

$$\Phi^\dagger \Phi = \frac{-\mu^2}{2\lambda} = \frac{v^2}{2} \tag{1.7}$$

The Higgs field is a SU(2) doublet composed by two complex scalar fields:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{1.8}$$

where

$$\phi^+ = \frac{\phi_1 + i\phi_2}{\sqrt{2}} \quad \phi^0 = \frac{\phi_3 + i\phi_4}{\sqrt{2}}. \tag{1.9}$$

Let's choose the situation where the vacuum state, ϕ_0, is:

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.10}$$

and the Higgs field doublet becomes:

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ v + \phi_3 + i\phi_4 \end{pmatrix} \tag{1.11}$$

When ϕ gets a vacuum expectation value the Lagrangian contains extra terms from the covariant derivative for the electroweak interaction:

$$\phi_0^\dagger (ig_1 B_\mu \frac{Y}{2} + ig_2 \frac{\overrightarrow{\tau}}{2} \cdot \overrightarrow{W^\mu})^\dagger (ig_1 B_\mu \frac{Y}{2} + ig_2 \frac{\overrightarrow{\tau}}{2} \cdot \overrightarrow{W^\mu}) \phi_0. \tag{1.12}$$

The final contribution to the Lagrangian is:

$$\left(\frac{v}{2\sqrt{2}}\right)^2 \left(2g_2^2 W_\mu^+ W^{\mu-} + (g_1^2 + g_2^2) Z_\mu Z^\mu + \frac{g_2 g_1 - g_1 g_2}{\sqrt{g_1^2 + g_2^2}} A_\mu A^\mu\right), \tag{1.13}$$

where

$$W_\mu^\pm = \frac{1}{\sqrt{2}} \left(W_\mu^1 \mp i W_\mu^2\right), \; Z_\mu = \frac{1}{\sqrt{g_1^2 + g_2^2}} (g_2 W_\mu^3 - g_1 B_\mu), \; A_\mu = \frac{1}{\sqrt{g_1^2 + g_2^2}} (g_2 W_\mu^3 + g_1 B_\mu). \tag{1.14}$$

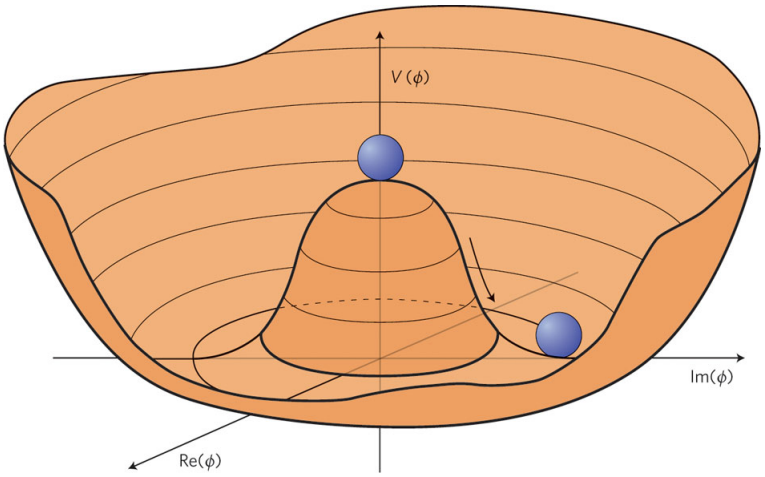


Figure 1.4: Higgs potential.

This way the masses of gauge bosons are:

$$M_W = \frac{1}{2} v g_2, \; M_Z = \frac{1}{2} v \sqrt{g_1^2 + g_2^2}, \; M_\gamma = 0. \tag{1.15}$$

Thus, the spontaneously breaking symmetry brings masses for $W^\pm$ and Z^0 bosons and leaves the photon massless.

Via a gauge transformation, one can fix the gauge in order to reduce the Higgs field to:

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \tag{1.16}$$

This way, only one field corresponding to a physical particle is left. Out of the four fields in Equation 1.11 (ϕ_1, ϕ_2, ϕ_3 and ϕ_4) three are absorbed in the additional degrees of

freedom that are necessary to describe massive gauge bosons. The $H(x)$ field represents a scalar boson, called the Higgs, corresponding to the Higgs-Brout-Englert boson [29–32].

Substituting the ϕ_0 definition of the from Equation 1.16 into Yukawa term of Lagrangian 1.2:

$$\mathcal{L}_{Yukawa} = \sum_{fermions} \{-g_Y^u[\bar{\psi}_L\tilde{\Phi}\psi_R^u + \bar{\psi}_R^u\tilde{\Phi}^\dagger\psi_L] - g_Y^d[\bar{\psi}_L\Phi\psi_R^d + \bar{\psi}_R^d\Phi^\dagger\psi_L]\}, \tag{1.17}$$

where the constants g_Y^i correspond to the coupling of Higgs fields (Φ, $\tilde{\Phi} = i\tau^2\Phi^*$) with fermion fields ψ_L, ψ_R, i.e. left-handed doublet and right-handed singlet (see Equations 1.4). One obtains the terms that generate the fermion masses:

$$M_{ferm} = \frac{1}{2}vg_Y^f, \tag{1.18}$$

where g_Y^f is an arbitrary Higgs-fermion coupling constant, unpredicted by the theory.

In 2012 the Higgs boson was observed by the CMS and ATLAS Collaborations at CERN [33, 34], which was a triumph of the Standard Model.

Inserting the interaction terms of Lagrangian 1.2 allows to write explicitly the interaction between gauge bosons and fermions. For example the electromagnetic term is:

$$\mathcal{L}_{em} = -\frac{g_1g_2}{\sqrt{g_1^2+g_2^2}}\left(-\bar{l}\gamma_\mu l + \frac{2}{3}\bar{q}^u\gamma_\mu q^u - \frac{1}{3}\bar{q}^d\gamma_\mu q^d\right)A^\mu, \tag{1.19}$$

where $q = \frac{g_1g_2}{\sqrt{g_1^2+g_2^2}}$ is elementary electric charge, l, q^u and q^d are wave functions of leptons and quarks type up and down correspondingly.

Taking in account definition $\psi_L = \frac{1}{2}(1-\gamma^5)\psi$ and $\psi_R = \frac{1}{2}(1+\gamma^5)\psi$ one can write two types of terms for charged and neutral currents of the weak interaction:

$$\mathcal{L}_{CC} = -\frac{g_2}{\sqrt{2}}\left(\bar{\psi}^u\gamma_\mu\frac{1-\gamma_5}{2}\psi^dW_\mu^+ + \bar{\psi}^d\gamma_\mu\frac{1-\gamma_5}{2}\psi^uW_\mu^-\right), \tag{1.20}$$

$$\mathcal{L}_{NC} = -\frac{1}{2\sqrt{g_1^2+g_2^2}}\bar{\psi}\gamma_\mu\frac{g_V - g_A\gamma_5}{2}\psi Z^\mu, \tag{1.21}$$

where ψ are fermion fields, g_V^f and g_A^f vector and axial couplings, that depend on the flavour of a fermion.

1.1.5 Neutral currents and the GIM mechanism

At the time when the SM was designed, only 3 quarks, u, d and s were known. The $SU(2) \times U(1)$ gauge symmetry organised them in a doublet $\left\{\begin{pmatrix}u\\d\end{pmatrix}\right\}$ and a singlet $\{s\}$ [35, 36]. This approach excludes the interactions of u and d quarks with s quark. However, $K^+ \to \mu^+\nu_\mu$ decay with strangeness violation was observed. This means that

u and d quarks have to couple with s quark. It was proposed that in the case of the d quark the state that undergoes the weak interaction is not the observable state, i.e. the mass eigenstate. These states are related like:

$$\begin{pmatrix} u \\ d' \end{pmatrix} = \begin{pmatrix} u \\ \cos\theta_C d + \sin\theta_C s \end{pmatrix}. \tag{1.22}$$

Hence the observed, or mass, eigenstates are not the interacting, or flavour eigenstates, and the transformation between these two eigenstates is done by a matrix [37]:

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos\theta_C & \sin\theta_C \\ -\sin\theta_C & \cos\theta_C \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}, \tag{1.23}$$

where $\theta_C \sim 13°$ is the Cabibbo angle. In this manner a coupling is possible between the u and s quarks.

However, this also implied the existence of a flavour changing neutral current (FCNC) coupling the d and s quarks, mediated by a neutral boson. This was excluded by experimental data: the decay $K^0 \to \mu^+\mu^-$ is very suppressed compared to $K^+ \to \mu^+\nu_\mu$. This FCNC disappears if all quarks belong to doublets. The Glashow-Iliopoulos-Maiani (GIM) [38] proposed the existence of the fourth c quark, forming with the s quark the second generation.

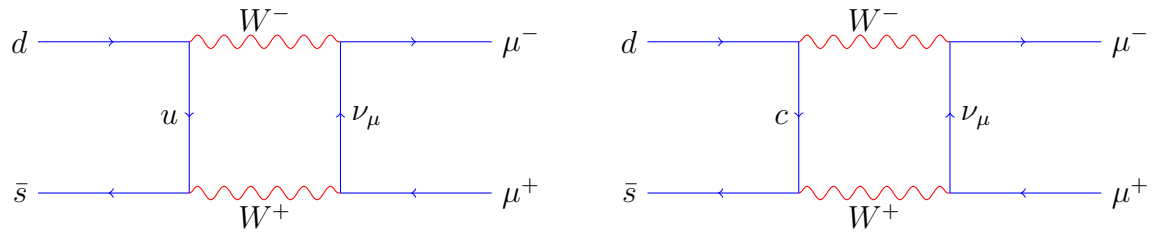


Figure 1.5: Feynman diagrams illustrating different contributions in amplitudes and GIM suppression mitigation in down-type quark (a) and up-type quark (b) sectors.

Inserting the matrix relation 1.23 in the Neutral current part of the Lagrangian (Equation 1.21), the impossibility of FCNC's appears clearly:

$$J^\mu_{neutral} = (\bar{u}\ \bar{c}\ \bar{t})\gamma^\mu(a_u - \frac{\gamma^5}{2}) \begin{pmatrix} u \\ c \\ t \end{pmatrix} - (\bar{d'}\ \bar{s'}\ \bar{b'})\gamma^\mu(a_d - \frac{\gamma^5}{2}) \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}, \tag{1.24}$$

where $a_u = \frac{1}{2} - \frac{4}{3}\sin^2\theta_w$ and $a_d = \frac{1}{2} - \frac{2}{3}\sin^2\theta_w$ are coefficients depending on the flavour of quark with the Weinberg angle, θ_w, measured experimentally $\sin^2\theta_w = 0.23$.

For the three generations the transformation similar to Equation 1.23 looks like:

$$J^{\mu}_{neutral} = (\bar{u}\ \bar{c}\ \bar{t})\gamma^{\mu}(a_u - \tfrac{\gamma^5}{2}) \begin{pmatrix} u \\ c \\ t \end{pmatrix} - (\bar{d}\ \bar{s}\ \bar{b})V^{+}_{CKM}\gamma^{\mu}(a_d - \tfrac{\gamma^5}{2})V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} =$$

$$(\bar{u}\ \bar{c}\ \bar{t})\gamma^{\mu}(a_u - \tfrac{\gamma^5}{2}) \begin{pmatrix} u \\ c \\ t \end{pmatrix} - (\bar{d}\ \bar{s}\ \bar{b})\gamma^{\mu}(a_d - \tfrac{\gamma^5}{2})V^{+}_{CKM}V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \tag{1.25}$$

$$(\bar{u}\ \bar{c}\ \bar{t})\gamma^{\mu}(a_u - \tfrac{\gamma^5}{2}) \begin{pmatrix} u \\ c \\ t \end{pmatrix} - (\bar{d}\ \bar{s}\ \bar{b})\gamma^{\mu}(a_d - \tfrac{\gamma^5}{2}) \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

where V_{CKM} is a unitary matrix, that will be discussed in Section 1.1.6. Thus the neutral quark currents contain only diagonal terms like $(\bar{d}d)$ and $(\bar{u}u)$. All non-diagonal elements, neutral currents with changing of flavour, like $(\bar{s}d)$ or $(\bar{u}c)$ are absent. Thus, in the SM flavour changing neutral currents are impossible at tree level.

Also the charm quark provides a second loop diagram (Figure 1.5(b)) with a coupling of opposite sign that cancel contribution with u quark (Figure 1.5(a)). In fact, were the mass of the charmed quark equal to the mass of the up quark the two diagrams would exactly cancel. For unequal masses, the result must be proportional to the difference $m_c^2 - m_u^2$. This suppression occurs both in beauty and charm meson decays. However, it far more pronounced in the latter case. This is one of the key features of charm physics.

1.1.6 The CKM matrix

The Yukawa sector, from which the fermion masses arise, involves terms of the type $m_{ij}\bar{\psi}_i\psi_j$. Nothing imposes that these couplings m_{ij} be diagonal. However, the corresponding Flavour Changing interactions are not observed experimentally: their occur only in charged current interaction. This is solved by assuming that the interacting states are flavour eigenstates, while the observed particles are mass eigenstates. To write the Lagrangian in terms of the physical (=mass) eigenstates, the corresponding matrices m_{ij} should be diagonalized. This is achieved with unitary matrices. The transformation relating the interaction and mass eigenstates involves:

$$J^{\mu}_{ch} = (\bar{u'}\ \bar{c'}\ \bar{t'})\gamma^{\mu}\frac{1-\gamma_5}{2} \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} \quad \Rightarrow \quad J^{\mu}_{ch} = (\bar{u}\ \bar{c}\ \bar{t})\gamma^{\mu}\frac{1-\gamma_5}{2}V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \tag{1.26}$$

where the V_{CKM} is called the Cabibbo-Kobayashi-Maskawa (CKM) matrix, or quark mixing matrix [39]:

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{1.27}$$

that performs the rotation between mass and flavour eigenstates:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.28}$$

As a result, the charged current $W^{\pm}$ interactions couple to the physical up- and down-type quarks with the couplings given by the CKM matrix with an important property: each CKM matrix element determines the strength with which a quark from one family couples to the quarks from other families. Quarks from the same family couples with a highest $V_{ii}^{CKM} \sim 1$, while the coupling between first and the second families are proportional to ~ 0.23 and between second and third are proportional to ~ 0.04. The smallest couplings are between the first and the third families and go as ~ 0.004. This hierarchy is illustrated in Figure 1.6 and the magnitudes of all nine CKM elements are [40]:

$$V_{CKM} = \begin{pmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.0412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{pmatrix}, \tag{1.29}$$

which also can be written in terms of the parameters λ, A, $\bar{\rho}$ and $\bar{\eta}$ [41]:

$$V_{CKM} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{1.30}$$

This non trivial hierarchy of the couplings is not explained by the SM. It suggests the existence of a more fundamental physics where this pattern make sense, i.e. that the study of the Flavour sector is a possible gate to NP.

Unitarity of the matrix implies $\sum_i V_{ij}V_{ik}^* = \delta_{jk}$ and $\sum_j V_{ij}V_{kj}^* = \delta_{ik}$. The six vanishing combinations can be represented as triangles in a complex plane. The most studied unitarity triangle arises from

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \tag{1.31}$$

If one divides each element by $V_{cd}V_{cb}^*$ the triangle as presented in Figure 1.7 is obtained. The openness of this triangle, i.e. large angles β and γ, predicts large CP asymmetries. The unitarity relations are a key feature of the SM. They imply relations between observables (branching ratios, CP asymmetries) that have a priori no reason to exist in theories built in a different way. Therefore looking for violations of these relations is a way to seek NP. Consequently, many of the angles and sides of the unitarity triangle are measured with high precision. The results of the precise measurements of CP asymmetries, mixing, semileptonic and rare decays are superimposed in Figure 1.8. One observes consistency between the measurements, which is a major success of the SM.

One important feature of the CKM matrix is the existence of two complex elements (see Equation 1.30), which are responsible for CP violation, as discussed in the next section.

A $N \times N$ complex matrix depends on N^2 real parts and N^2 imaginary parts. If this matrix is unitary, then N^2 unitarity relations reduce the number of independent parameters. Moreover, $2 \times N - 1$ of these can be absorbed in a re-definition of the quark fields overall phase. Therefore, the final number of independent parameters is $(N-1)^2$. Among them, $N(N-1)/2$ are real mixing angles. The remaining ones are $(N-1)(N-2)/2$ complex phases. Thus, in case of only $N = 2$ generations exist, the

2×2 matrix has only one mixing angle parameter and no complex phase. When a third generation is added, one complex phase remains, that will allow the SM to account for CP violation phenomena.

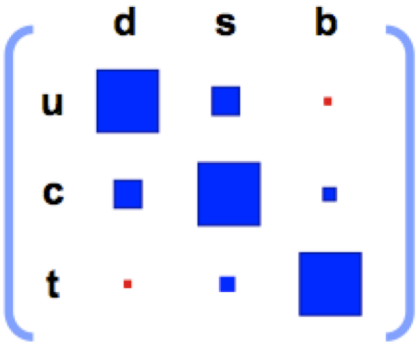


Figure 1.6: Illustration of CKM matrix elements. Size of squares corresponds to the CKM values, i.e. strength of certain quark weak transitions.

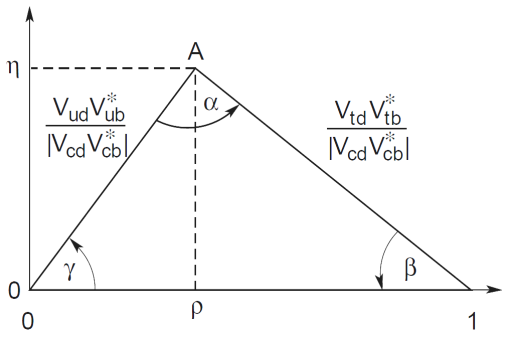


Figure 1.7: The illustration of the unitarity triangle.

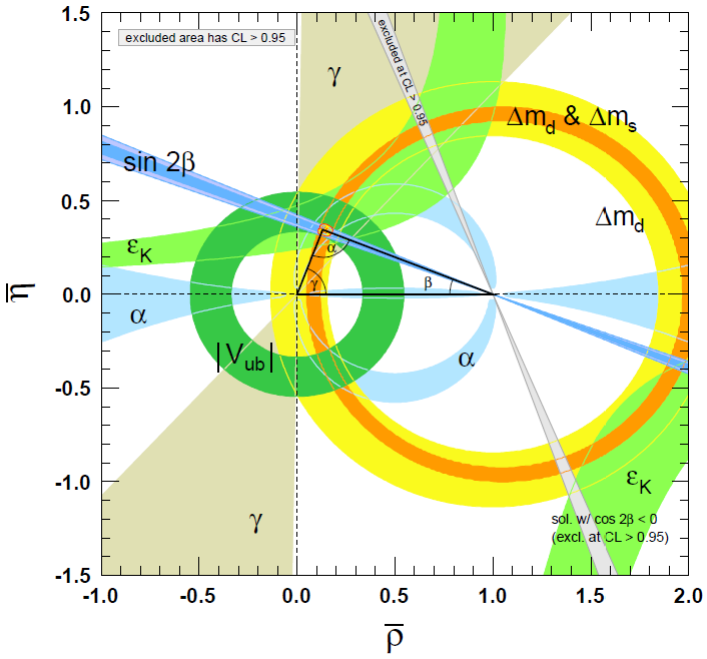


Figure 1.8: Global fit of the CKM Unitary Triangle (in black), basing on the available experimental measurements on the CKM parameters (coloured regions).

1.1.7 CP violation

CP violation arises from the complex phases in the CKM matrix. To illustrate this, recall the CP transformations:

$$\bar{\psi}_i \psi_j \to \bar{\psi}_j \psi_i, \quad \bar{\psi}_i \gamma^\mu W_\mu (1 - \gamma_5) \psi_j \to \bar{\psi}_j \gamma^\mu W_\mu (1 - \gamma_5) \psi_i. \tag{1.32}$$

In the charged current sector of the Lagrangian

$$-\frac{g_2}{\sqrt{2}}\left[V_{ij}^{CKM}\bar{\psi}_i^u\gamma^\mu W^{+\mu}\frac{1-\gamma_5}{2}\psi_j^d + V_{ij}^{*CKM}\bar{\psi}_j^d\gamma^\mu W^{-\mu}\frac{1-\gamma_5}{2}\psi_i^u\right], \tag{1.33}$$

the CP operation interchanges the two terms except for V_{ij}^{*CKM} and V_{ij}^{CKM}. Thus, the SM Lagrangian is not invariant under CP since some of the V_{ij}'s are complex. In other words, it describes the CP-violation.

One possible way to violate CP involves decay amplitudes:

$$A_f = \langle f|\mathcal{H}|D\rangle, \quad \bar{A}_{\bar{f}} = \langle \bar{f}|\mathcal{H}|\bar{D}\rangle, \tag{1.34}$$

that are not the same, i.e. $A_f \neq \bar{A}_{\bar{f}}$. In Equation 1.34 D and f are initial meson and final multi-particle states, $\bar{D}$ and $\bar{f}$ are their CP-conjugates and $\mathcal{H}$ is the Hamiltonian governing weak interactions. For such a violation to happen, several decay amplitudes contributing to $A_i = |A_i|e^{i\phi}$ are necessary. Each amplitude carries a CP-odd weak phase, ϕ_i, originating from complex coupling constants, and a CP-even strong phase, δ_i. As the strong interaction is invariant under CP transformation, the strong phases are conserved with CP conjugation.

Therefore, we have amplitudes:

$$A_f = \sum |A_i| e^{i(\delta_i+\phi_i)}, \quad \bar{A}_{\bar{f}} = \sum |A_j| e^{i(\delta_j-\phi_j)}. \tag{1.35}$$

and the CP asymmetry $\mathcal{A}_{CP}$ is:

$$\mathcal{A}_{CP} = \frac{|A_f|^2 - |\bar{A}_{\bar{f}}|^2}{|A_f|^2 + |\bar{A}_{\bar{f}}|^2} \sim \sum_{ij} |A_i A_j| \sin(\phi_i - \phi_j)\sin(\delta_i - \delta_j). \tag{1.36}$$

From Equation 1.36 one can see that CP violation in decay amplitudes requires the presence of different weak and strong phases among the various contributing amplitudes. This can be illustrated with the example of $B^0 \to K^+\pi^-$ decays, where the CP asymmetry has been observed to be -0.098 ± 0.012 [40]. The first amplitude contributing to this decay is a tree diagram and the second amplitude is a penguin-diagram, as shown in Figure 1.9. It is clear from the figure that there is a weak phase difference between the tree ($\sim V_{ub}^* V_{us} \cong A\lambda^3(\rho + i\eta)\lambda$) and penguin amplitude ($\sim V_{tb}^* V_{ts} \cong -A\lambda^2$), and in general a different strong phase is expected. Hence it is not surprising that experiment measures a non-zero asymmetry. This type of CP-violation is called *direct*.

The example above shows that CP violation arises from the interference between amplitudes. At least two amplitudes must contribute significantly to the decay. The mixing of neutral B, D or K mesons provides an additional amplitude that makes it possible to meet this condition.

The description of neutral meson mixing is based on the Schrödinger equation presented below, which involves an effective Hamiltonian. It is non-hermitian in order to not only account for the mixing, but also for the decay of the system. This equation is:

$$i\frac{d}{dt}\begin{pmatrix} D^0(t) \\ \bar{D}^0(t)\end{pmatrix} = \mathcal{H}_{eff}\begin{pmatrix} D^0(t) \\ \bar{D}^0(t)\end{pmatrix}, \tag{1.37}$$

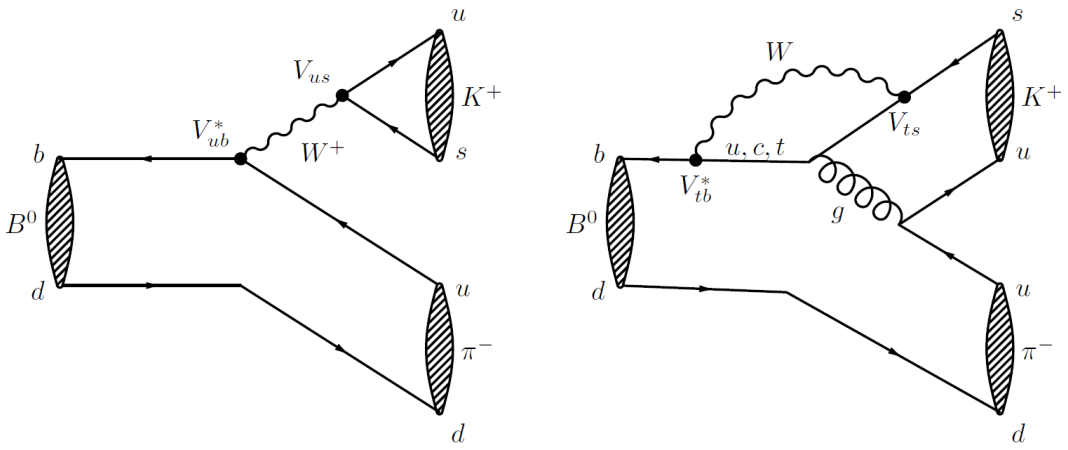


Figure 1.9: The two interfering diagrams of the decay $B^0 \to K^+\pi^-$.

with

$$\mathcal{H}_{eff} = \mathcal{M} - \frac{i}{2}\mathbf{\Gamma} = \begin{pmatrix} M_{11} & M_{12} \\ M_{12}^* & M_{22} \end{pmatrix} - \frac{i}{2}\begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma_{22} \end{pmatrix}, \tag{1.38}$$

where $\mathcal{M}$ and $\mathbf{\Gamma}$ are hermitian matrices. The mixing allows for an interference between the direct decay of a meson to its final state and the decay after it has oscillated into it associated anti-meson. Large CP-violating phenomena are possible thanks to this phenomenon, called CP violation in the interplay between decays with and without mixing. Note that the mixing amplitude can also be a source of *indirect* CP violation in that sense that it can carry a weak phase.

However, as introduced as it is in the SM, the magnitude of CP violation is not enough to explain of matter-antimatter asymmetry in the Universe, therefore some New Physics beyond the Standard Model is required to explain additional sources of CP violation. Nevertheless, no significant NP effects have been measured so far. Thus, searches for the NP in flavour sector must be continued.

1.1.8 Unanswered questions of the Standard model

The SM is a very successful theory, that is able to describe all the phenomena observed by experiments. Some of its predictions have received spectacular experimental confirmations (existence of neutral and charged gauge bosons, existence of the Higgs bosons, etc.). However, a certain number of questions remains unanswered in this framework, among which:

- Dark matter and dark energy. Cosmological observations indicate that the particle described in the SM can account for only 4.5% of the total energy in the Universe. All the rest is dark matter (about 22.5%) and dark energy. A unknown form of energy present in all space and responsible for the acceleration of the Universe expansion. Attempts to explain dark energy in terms of the vacuum energy of the Standard Model lead to a mismatch of 120 orders of magnitude.

- Matter-antimatter asymmetry. The Universe mainly consists of matter. The deficit of antimatter requires, among other conditions, the existence of CP violation in particle interactions. The magnitude of the CP violating phenomena predicted by the SM does not suffice, however, to explain its quasi absence. Moreover, some consider CP violation is included in the SM in an *ad hoc* way, and would prefer it to derive from a more physical principle. This could be the indication that the SM description is the effective description of a more fundamental physics.

- Fermion masses and Flavour Violation. The very large differences between the fermion masses and the quark flavour violating couplings are not explained in the SM.

- Hierarchy problem. One aspect of this problem concerns the mass of the Higgs boson. It receives quantum corrections that far larger than its observed mass. This means that the bare mass parameter of the Higgs must be fine tuned so as to cancel the vast bulk of the quantum corrections, over many orders of magnitudes. This level of fine tuning is deemed unnatural by many theorists.

- Strong CP problem. Theoretically it is possible to have violation of CP symmetry in the strong interaction sector also. Experimentally, however, no such asymmetry has been found, implying that the coefficient of this term is very close to zero. This fine tuning is also considered unnatural.

- Non-unification of couplings. The gauge couplings of the SU(3), SU(2) and U(1) groups are independent quantities. Due to higher-order corrections, each of these is actually a function of the typical energy scale Q relevant to the process. In many grand unified theories the three gauge couplings are predicted to meet at some high energy unification. However, this unification does not occur when the couplings extrapolated using the SM model expression.

- Number of parameters. The Standard Model contains 26 numerical parameters unknown originally, that could be fixed only experimentally; at least 20 of these parameters are related to the flavour physics. In electroweak theory nothing dictates the values of the interaction couplings and masses.

- Gravitation. The SM does not include gravitational interaction and it is incompatible with general relativity, the most successful gravitation theory up to date.

All these problems that are unsolved by the Standard Model indirectly indicate that some NP beyond it should exist.

1.2 Rare charm decays as probes of New Physics models

The main subject of this thesis is searches for rare charm decays, $D^+ \to \pi^+\mu^+\mu^-$ and $D^+ \to h^+h^{(\prime)-}\mu^+\mu^-$, where $h,\ h'$ can be a pion or kaon. Such decays proceed mostly

through the $c \to u\ \mu^+\mu^-$ Flavour Changing Neutral Current(FCNC) process, which is possible only at loop diagram level in the Standard Model. For rare charm decays, unlike for B decays, there is no Standard Model heavy particle, such as the top quark, that can enter the loop diagram to mitigate the GIM suppression. Consequently, rare charm decays are good tools to probe to New Physics beyond the SM.

1.2.1 Introduction to charm physics

The charm sector plays a very important role in the understanding of the particle physics picture. Historically, it was predicted by GIM mechanism, which was needed to explain the suppression of the FCNC. The discovery of J/ψ meson [42,43], i.e. $c\bar{c}$ bound state, in 1974 proved the existence of the fourth quark. This discovery finally convinced the physics community of the quark model's validity. The c quark hence played a crucial role in definition of the structure of the SM and in its construction. After the c quark discovery the principal modern form of the SM Lagrangian was defined. This and the realisation of the difference between mass and flavour eigenstates plus the observation of CP violation led to the prediction of the third family.

However, in later years charm physics come to be considered as less promising than B or K physics due to large uncertainties in theoretical calculations. Resonance contributions in D decays are dominant and these cannot be described in perturbative QCD. The effective field theory also cannot be easily constructed at the charm scale, unlike in the case of B mesons, since b quarks are much heavy than Λ_{QCD} so that the perturbative QCD methods can be applied by neglecting the higher order terms in the $1/m_b$ expansion. The D meson system also differs from the K meson system, where perturbative series at scale $(m_s)^n$ are possible since it can be assumed that $m_s \to 0$.

Another complication comes from the rich resonance structure of D decays. Even for dominant modes high multiplicity final states occur. This adds even more complexity into theoretical calculations, making any precise predictions to be difficult.

Moreover, CP violation is expected to be absent in D decays in the SM, as only quarks from the first two families mainly participate. At the same time the period of the oscillations $D^0 - \overline{D}^0$ is long, which again reduces any CP violation effects. Even rare decays are very rare in the charm sector as the mitigation of GIM suppression is not of the same order of magnitude as in case of decays involving down-type quarks. For the first one d, s and b quarks enter in the loop, for the second - s, c and t. The mass difference between down-type quarks involved in a loop is smaller than between up-type quarks, which makes them to be more suppressed. Too small possible CP violation effects and too small branching fractions for rare decays reduced the attraction of charm physics in comparison to B and K physics

Nevertheless, all these drawbacks bring the advantage of small SM backgrounds, thus, possible NP effects may be seen relatively easily. Deviations from the SM predictions in CP violation and branching fractions ($\sim 10^{-8} - 10^{-9}$) would be signs of the physics beyond the SM. The high multiplicity of the final states of the D decays bring many observables and makes possible studies in the n-dimensional Dalitz space and angular analyses (T-odd, Forward-Backward asymmetries). Any deviations, even if they are very small or limited to a particular part of the phase space could be considered as

good candidates for the appearance of NP. The domination of long distance effects may in fact be used in observing CP violation effects as the strong phase difference $\Delta\delta_i$ can become large.

Moreover, most of the interesting effects are possible for high branching fractions modes, such as $D^0 \to K^-\pi^+\pi^+\pi^-$, $D^0 \to K^+K^-$ and $D^0 \to K^-\pi^+$, and they have large production cross sections at modern accelerators. This allows for the precise experimental tests of the SM in charm sector. The semileptonic D decays, $D^0 \to K^-l^+\nu_l$ and $D^0 \to \pi^-l^+\nu_l$, also can be very useful for testing lattice QCD methods, as they work at the scale of low energy hadron, K^- and π^-, representing a larger fraction of the decay. In addition, rare charm decays bring an unique opportunity to study FCNCs in up-type sector, where NP effects could differ from down-type sector.

Consequently, charm decays have a high potential for discovering effects beyond the SM and new features of the SM itself, like CP violation.

1.2.2 General theoretical approach

Heavy hadron decays are difficult to describe theoretically. The main difficulty in the theoretical treatment of heavy mesons decays stems from the fact that the weakly decaying quark undergoes simultaneously effects from the strong interaction. Also many different scales are involved in the decay process, from below 1 GeV to the mass of the heaviest virtual particles that might have an impact on the decay process (> 100 GeV). The weakly interacting quarks are not free fields: they are bound within hadrons. The strong interaction responsible for the bound states involves energy transfers of the order Λ_{QCD}. No perturbative description is possible at this energy scale. Furthermore, the charm quark mass, being not too far from the light vector resonances, leads to the dominance of processes such as $D^0 \to VV' \to hh'\mu^+\mu^-$ over the decay $D^0 \to hh'\mu^+\mu^-$, which emphasizes the importance of long distance ($\mathcal{LD}$) physics. In addition, even for the short distance part, traditional perturbative calculations are difficult, due to *large logarithms* that spoil the convergence of the perturbative series even where α_s would be small enough.

The most common approach to describe the heavy flavour physics is based on the *Operator Product Expansion* (OPE), which allows an effective theory to be built that has the following structure:

$$\mathcal{H}_{eff} = \sum_i C_i \cdot O_i, \tag{1.39}$$

and leads to amplitudes of this form:

$$\mathcal{A}_{eff} = \sum_i C_i(\mu) \cdot \langle O_i(\mu) \rangle. \tag{1.40}$$

The C_i are the Wilson coefficients and the O_i are operators, that comprise the field of the decaying quarks and those of the other particles involved in decays described by $\mathcal{H}_{eff}$. There is one O_i for each type of quark interaction.

Current-Current operators O_1 and O_2:

$$O_1^d = (\bar{u}_L^\alpha \gamma_\mu d_L^\beta)(\bar{d}_L^\beta \gamma^\mu c_L^\alpha), \qquad O_2^d = (\bar{u}_L^\alpha \gamma_\mu d_L^\alpha)(\bar{d}_L^\beta \gamma^\mu c_L^\beta), \tag{1.41}$$

$$O_1^s = (\bar{u}_L^\alpha \gamma_\mu s_L^\beta)(\bar{s}_L^\beta \gamma^\mu c_L^\alpha), \qquad O_2^s = (\bar{u}_L^\alpha \gamma_\mu s_L^\alpha)(\bar{s}_L^\beta \gamma^\mu c_L^\beta), \tag{1.42}$$

where $q_L{=}P_L q$ and $P_{L,R} = \frac{1}{2}(1 \pm \gamma_5)$ is the chirality operator, α, β are colour indices. The full theory diagrams which are translated into O_1 and O_2 are shown in Figure 1.10. Such tree-level amplitudes in general dominate the decays to which they contribute. Note that O_1 does not exist at the leading order in QCD. It appears only with the inclusion of the QCD corrections. It is therefore suppressed by α_s by about a factor 2 or 3 with respect to O_2. In practice, both operators differ only by their colour structure (the α and β indices).

QCD penguin operators O_3 to O_6:

$$O_3 = (\bar{u}_L^\alpha \gamma_\mu c_L^\alpha) \sum_q (\bar{q}_L^\beta \gamma^\mu q_L^\beta), \qquad O_4 = (\bar{u}_L^\alpha \gamma_\mu c_L^\beta) \sum_q (\bar{q}_L^\beta \gamma^\mu q_L^\alpha), \tag{1.43}$$

$$O_5 = (\bar{u}_L^\alpha \gamma_\mu c_L^\alpha) \sum_q (\bar{q}_R^\beta \gamma^\mu q_R^\beta), \qquad O_6 = (\bar{u}_L^\alpha \gamma_\mu c_L^\beta) \sum_q (\bar{q}_R^\beta \gamma^\mu q_R^\alpha), \tag{1.44}$$

The full theory diagrams of these operators are shown in Figure 1.11. They correspond to loop amplitudes. This, in general, give rise to a $\sim 1/(16\pi^2)$ suppression factor with respect to current-current operators. However, the GIM suppression is far more severe in the case of c decays (see later).

Magnetic and Chromomagnetic penguin operators O_7 and O_8:

$$O_7 = \frac{e}{16\pi^2} m_c (\bar{u}_L \sigma_{\mu\nu} c_R) F^{\mu\nu}, \qquad O_8 = \frac{g_s}{16\pi^2} m_c (\bar{u}_L \sigma_{\mu\nu} T^a c_R) G_a^{\mu\nu}. \tag{1.45}$$

The full theory diagrams of these operators are shown in Figure 1.12.

Semileptonic penguin operator O_9 and box operator O_{10}:

$$O_9 = \frac{e^2}{16\pi^2} (\bar{u}_L \gamma_\mu c_L)(\bar{l}\gamma^\mu l), \qquad O_{10} = \frac{e^2}{16\pi^2} (\bar{u}_L \gamma_\mu c_L)(\bar{l}\gamma^\mu \gamma_5 l). \tag{1.46}$$

Together with O_7, they govern the non-resonant contribution to $D \to h^+(h^{(\prime)-})\mu^+\mu^-$. The full theory diagrams from which they arise are shown in Figure 1.13.

QCD corrections play an important role in flavour physics alongside pure electroweak interactions. The QCD correction introduces the *renormalization scale* μ, which separated the short distance ($\mathcal{SD}$) and long distance ($\mathcal{LD}$) physics. The $\mathcal{SD}$ contributions ($p > \mu$) are described in the *Wilson Coefficients* C_i, while $\mathcal{LD}$ contributions ($p < \mu$), including the hadronisation of quarks into hadrons, are encompassed in the matrix elements of the O_i operators.

This $\mathcal{SD}$-$\mathcal{LD}$ separation is an important feature of this technique. If the full theory would be used, i.e. the usual electroweak and QCD Lagrangians, QCD corrections would lead to a non-convergent series. With this approach, the perturbative high energy contributions are cleanly separated from the non-perturbative ones. This is

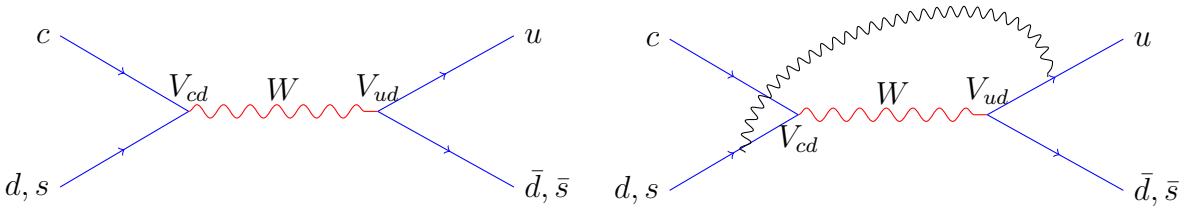


Figure 1.10: Feynman diagrams for operators O_1 and O_2 in OPE.

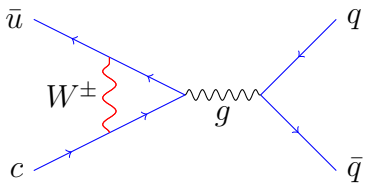


Figure 1.11: Feynman diagrams for operators O_3-O_6 in OPE.

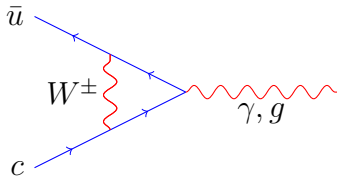
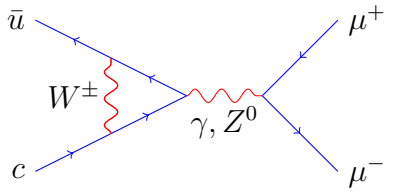


Figure 1.12: Feynman diagrams for operators O_7-O_8 in OPE.



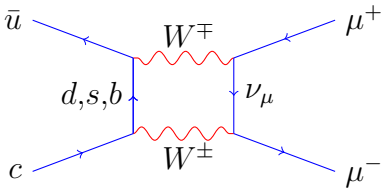


Figure 1.13: Feynman diagrams for operators O_9 and O_{10} in OPE.

particularly important for some observables, like asymmetries, defined such that the $\mathcal{LD}$ contributions cancel. Indeed, the latter dominate the theoretical uncertainty evaluated with the help of non-perturbative, often model-dependent, techniques, such as the factorization approximation. This assumes that a given $H \to M_1 M_2$ decay of hadron, H, into two mesons, M_1 and M_2, can be factorized as:

$$\langle M_1 M_2 | \bar{q}_2 \gamma_\mu q_2 \bar{q}_1 \gamma^\mu h | H \rangle = \langle M_2 | \bar{q}_2 \gamma_\mu q_2 | 0 \rangle \langle M_1 | \bar{q}_1 \gamma^\mu h | H \rangle. \tag{1.47}$$

During the decay process a colour singlet $\bar{q}_2 q_2$ pair is created. It has enough energy to be uninfluenced by the other quarks. The corresponding matrix element $\langle M_2 | \bar{q}_2 \gamma_\mu q_2 | 0 \rangle$ depends essentially on the M_2 decay constant, f_{M_2}, measured in leptonic decays of M_2. The other matrix element, $\langle M_1 | \bar{q}_1 \gamma^\mu h | H \rangle$, can be evaluated with the non perturbative techniques such as Lattice QCD or QCD sum rules (see [44]). It is also possible to use decays which differ from the decay of interest only by the nature of the final state quarks: if they are measured experimentally, one can then derive their matrix elements and relate them to the $\langle O_i \rangle$ by using the symmetries of the strong interaction (strong isospin, SU(3)).

The effective Hamiltonian must be consistent with the full theory. A *matching* between full and effective approaches is done at a scale $\mu \sim M_W$, where both theories are valid. It is performed by equating the contributions to the amplitudes in both theories and adjusting the Wilson coefficients, that contain the information about the influence of the heavy particles (Z^0, $W^{\pm}$, t, b), *integrated out* and not explicitly present in the effective operators.

The scale μ should be chosen of the order of the mass of the decaying particle, i.e. m_c in the cases of charm decays. The C_i at $\mu = m_c$ are derived from C_i at $\mu = M_W$ using Renormalization Group Improved (RGI) perturbation theory. This technique is similar to the well-known running of $\alpha_s(\mu)$, which determines the intensity of the strong interaction as a function of the energy available in the process. In the case of RGI, the divergences affecting the operators can be only cancelled by counter-terms from other operators. Under renormalization different operators mix and the coefficients $C_i(m_c)$ become a linear combination of $C_i(M_W)$. Thus, some operators are enhanced with respect to the mere electroweak amplitudes, when the QCD corrections are included.

1.2.3 Effective Hamiltonian for the $D \to h(h')\mu^+\mu^-$ decays

The effective Hamiltonian describing the processes with $\Delta C = 1$, $\Delta S = 0$, such as $D^0 \to \pi^+\pi^-\mu^+\mu^-$, $D^0 \to K^+K^-\mu^+\mu^-$ or $D^+ \to \pi^+\mu^+\mu^-$ decays, at the scale $\mu \sim M_W$ is:

$$\begin{aligned} H_{eff}^{\Delta C=1,\Delta S=0} \quad = \quad & -\frac{G_F}{\sqrt{2}}\left[V_{cd}^*V_{ud}\left(\sum_{i=1,2} C_iO_i^d + \sum_{i=3,\ldots,10} C_iO_i\right)\right. \\ & +V_{cs}^*V_{us}\left(\sum_{i=1,2} C_iO_i^s + \sum_{i=3,\ldots,10} C_iO_i\right) \\ & \left. + V_{cb}^*V_{ub}\left(\sum_{i=1,2} C_iO_i^b + \sum_{i=3,\ldots,10} C_iO_i\right)\right] \end{aligned} \tag{1.48}$$

This form can be simplified as the masses of the d, s and b quarks propagating in loop diagrams can be neglected compared to M_W. Thus, all Wilson Coefficients have the same values and, therefore, they are cancelled due to the unitarity of the CKM matrix:

$$H_{eff}^{\Delta C=1,\Delta S=0} = -\frac{G_F}{\sqrt{2}}\left[V_{cd}^*V_{ud}\sum_{i=1,2} C_i(O_i^d - O_i^b) + V_{cs}^*V_{us}\sum_{i=1,2} C_i(O_i^s - O_i^b)\right] \tag{1.49}$$

The ideal GIM suppression pointed out above is mitigated by the inclusion of the QCD corrections. At the scale $\mu \sim m_c$ the b-quark is integrated out and the operator mixing makes some of the C_i more sizeable. The effective Hamiltonian now reads:

$$\begin{aligned} H_{eff}^{\Delta C=1,\Delta S=0} \quad = \quad & -\frac{G_F}{\sqrt{2}}\left[V_{cd}^*V_{ud}\left(\sum_{i=1,2} C_iO_i^d + \sum_{i=3,\ldots,10} C_iO_i\right) + \right. \\ & \left. V_{cs}^*V_{us}\left(\sum_{i=1,2} C_iO_i^s + \sum_{i=3,\ldots,10} C_iO_i\right)\right]. \end{aligned} \tag{1.50}$$

From the unitarity of the CKM matrix it becomes:

$$H_{eff}^{\Delta C=1,\Delta S=0} = -\frac{G_F}{\sqrt{2}}\left[V_{cd}^* V_{ud}\sum_{i=1,2} C_i O_i^d + V_{cs}^* V_{us}\sum_{i=1,2} C_i O_i^s - V_{cb}^* V_{ub}\sum_{i=3,...,10} C_i O_i\right], \tag{1.51}$$

Note that the penguin loop contributions are still very suppressed since $V_{cb}^* V_{ub} \sim 10^{-5}$.

For $D \to h(h)\mu^+\mu^-$ decays the *current-current* operators $O_{1,2}^{(q)}$ govern the resonance contributions to rare charm decays, i.e. the modes like $D \to h(h)V(\to \mu^+\mu^-)$, also called the long distance (LD) contribution. This is illustrated in Figure 1.14. Such contributions dominate the rare D decay amplitudes. Depending on the mode, the total branching ratio typically hovers around 10^{-6}.

Precise non perturbative calculations are extremely difficult to obtain here. However, estimations are possible thanks to the Vector Meson Dominance approach (VMD) [45], where $D \to h(h^{(')})\mu^+\mu^-$ decays proceed predominantly via resonances, V/V', that couple to a virtual photon $V' \to \gamma^* \to \mu^+\mu^-$. Schematically, this approach allows the amplitude to be broken into a series of two body amplitudes, that are easier to manage. For instance, the $D^0 \to V(\to hh^{(')})V'(\to \mu^+\mu^-)$ amplitude is the product of amplitudes like $\langle VV^{'}|\mathcal{H}|D^0\rangle$, $\langle hh^{'}|\mathcal{H}|V\rangle$, $\langle\gamma^*|\mathcal{H}_{V\gamma}|V^{'}\rangle$ at each point of the (m_{hh}, m_{ll}) space and for each value of the final state's angular momentum [18]. Resonance decays are described by a Breit-Wigner propagator. One can calculate $\langle VV^{'}|\mathcal{H}|D^0\rangle$ with the help of the current-current operators of the effective Hamiltonian presented above, $\langle VV^{'}|O_{1,2}|D^0\rangle$. They are based on measured $V^{(')}$ decay constants and on $D^0 \to V^{(')}$ form factors that have been partly determined from measurements, and are also the subject of non perturbative QCD methods (Lattice, Light Cone Sum Rules, etc...). Experiments also provide information on $\langle hh^{'}|\mathcal{H}|V\rangle$ and $\langle\gamma^*|\mathcal{H}_{V\gamma}|V'\rangle$.

In the case of $D^0 \to h^{\pm}h^{(')\mp}\mu^+\mu^-$ decays, a second type of LD contribution exists: two body $D^0 \to K^{\mp}\pi^{\pm}$, $D^0 \to \pi^+\pi^-$ and $D^0 \to K^+K^-$ decays followed by the radiation of a photon. This contribution is easy to evaluate, and is found to be two orders of magnitude below the resonant contributions [18, 46].

The *QCD penguin* operators, $O_{3,...,6}$ can also contribute (Figure 1.15). However, those most sensitive to NP contributions correspond to the non-resonant transitions that are governed by the *magnetic* and *semileptonic* penguin operators, $O_{7\gamma}$, O_9 and O_{10} (Figure 1.16). These contributions are also called short distance (SD) contributions.

Operator mixing plays an important role for three last operators, $O_{7\gamma}$, O_9 and O_{10}. Here two-loop QCD corrections lead to a large enhancement:

- The C_7 coefficient receives the largest contribution from C_2. However, this contribution does not dominate in $\mathcal{H}_{eff}$.
- The C_9 also gets contributions from C_2. This is possible since the four-quark operator O_2 can produce a similar diagram to the electroweak diagram generating O_9, by contracting the two s or d legs in a loop and radiating a photon. This contribution to C_9 remains the dominant one after the strong interaction is included, although the latter causes further mixing. C_9 is at least an order of magnitude higher that C_7 since its mixing with O_2 starts at tree level.

- The C_{10} does not obtain additional contributions as O_{10} do not mix with the other operators due to its chirality. It is not affected by the renormalization group running, and therefore remains very suppressed: $C_{10}(m_c) = C_{10}(m_W) \sim \frac{m_s^2}{m_W^2}$.

The considerations above are discussed in [47–50]. Evaluations of the non-resonant contribution to various decays are performed in the framework of the SM. In [47], the non-resonant inclusive rate is quantified: $BF^{SD}(D \to X_u e^+ e^-) \simeq 3.7 \times 10^{-9}$. It should be slightly lower for the muonic final state. For individual exclusive three or four body modes, we therefore do not expect branching ratios above 10^{-9}, which is three orders of magnitude less than resonant contributions. Examples of LD calculations and of their interplay with SD can be found in [18, 46].

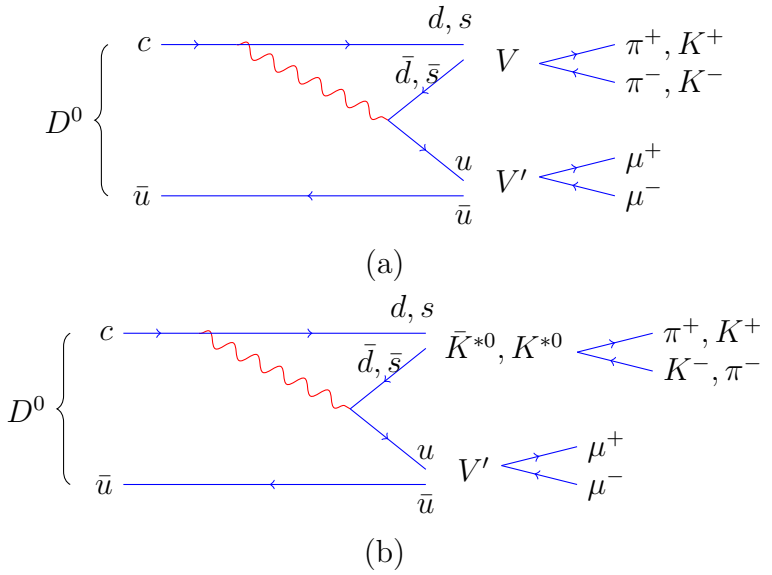


Figure 1.14: The resonance contributions to rare charm decays $D \to h(h')\mu^+\mu^-$ governed by current-current operators $O_{1,2}^{(q)}$.

For the Cabibbo favoured $D^0 \to K^-\pi^+\mu^+\mu^-$ and doubly Cabibbo suppressed $D^0 \to K^+\pi^-\mu^+\mu^-$ decays with $\Delta C = 1$ and $\Delta S = 1$ it is necessary to introduce an additional term into the Hamiltonian (Figure 1.14(b)):

$$H_{eff}^{\Delta C=1,\Delta S=1} = -\frac{G_F}{\sqrt{2}}\left[V_{cs}^* V_{ud} \sum_{i=1,2} C_i O_i^{sd} + V_{cd}^* V_{us} \sum_{i=1,2} C_i O_i^{ds}\right], \tag{1.52}$$

where two new sets of current-current operators appear:

$$O_1^{sd} = (\bar{u}_L^\alpha \gamma_\mu d_L^\beta)(\bar{s}_L^\beta \gamma^\mu c_L^\alpha), \qquad O_2^{sd} = (\bar{u}_L^\alpha \gamma_\mu d_L^\alpha)(\bar{s}_L^\beta \gamma^\mu c_L^\beta), \tag{1.53}$$

$$O_1^{ds} = (\bar{u}_L^\alpha \gamma_\mu s_L^\beta)(\bar{d}_L^\beta \gamma^\mu c_L^\alpha), \qquad O_2^{ds} = (\bar{u}_L^\alpha \gamma_\mu s_L^\alpha)(\bar{d}_L^\beta \gamma^\mu c_L^\beta). \tag{1.54}$$

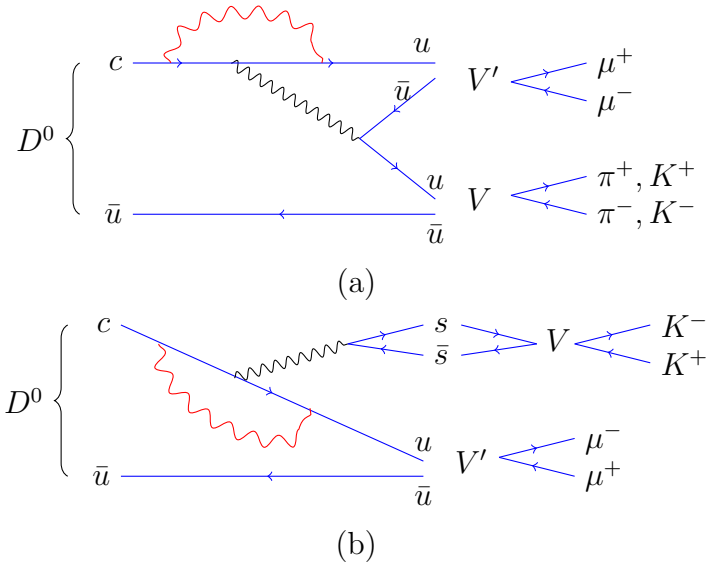


Figure 1.15: The QCD loop contributions, $O_{3,\ldots,6}$, to rare charm decays $D \to h(h')\mu^+\mu^-$.

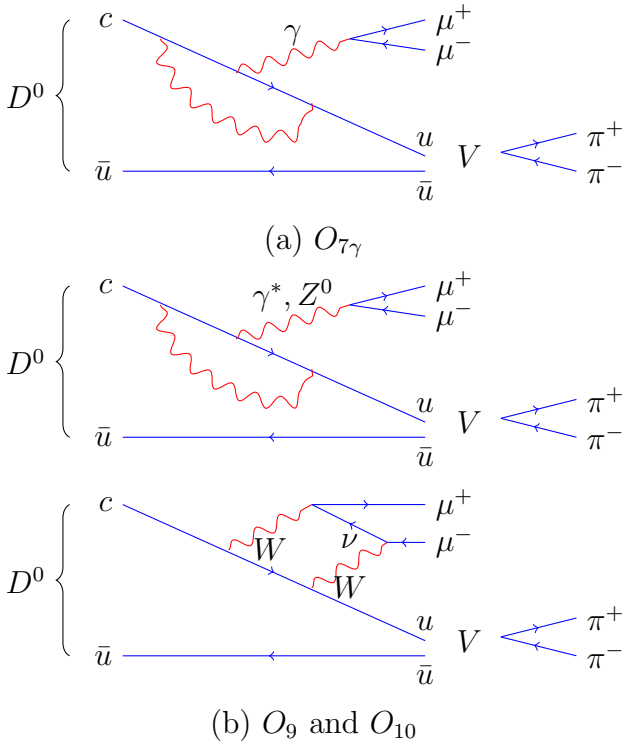


Figure 1.16: The non-resonant contributions to $D \to h(h')\mu^+\mu^-$ amplitudes.

1.2.4 Overview of New Physics models

The contributions of the electromagnetic and dilepton operators, $O_{7,9,10}$, are very suppressed in the SM. Furthermore, it is natural to expect NP in these operators since they are loop dominated. The effectiveness of the GIM mechanism in charm

decays suggests that NP models where tree-level FCNCs are possible should play a leading role in rare charm decays. Several such models are presented below in Table 1.5. However, models involving loop amplitudes with internal new particles can also be important. Examples of such models are listed in Table 1.6. In this section, a panorama of theoretical studies relevant to $D \to h(h')\mu^+\mu^-$ decays is presented. We do not give a detailed description of these models and the calculations carried out to predict their impact.

We present in this section estimates of affected experimental observables. There are two main types:

- **Branching ratios** of non-resonant contributions. The rare D decays branching ratios are totally dominated by resonant contributions, like $D^0 \to V\gamma^*$. Therefore, experimentalist consider partial branching ratios $\Delta\mathcal{B}$ in extreme regions of the $m(\mu^+\mu^-)$ spectrum, where these contributions are minimal. This is illustrated in Figure 1.17, which shows the differential $D^+ \to \pi^+\mu^+\mu^-$ branching ratio as predicted by the SM and a NP model. This is also the approach adopted for the measurements based on the LHCb 2011 and 2012 datasets (see Chapters 3, 5 and 6).

  In a presence of NP one can expect any of the three operators of SD $c \to u\mu^+\mu^-$ transitions to be modified. In most of the models the relative enhancement is large for C_{10}. The enhancement is moderate for C_9 and C_7. Due to the experimental constraints from FCNC charm decays, the recent trend is to expect only a limited modification from NP: the SD branching ratios of the three or four-body decays should hardly reach 10^{-8}.

- **Asymmetries**. Measurable NP effects are also possible in asymmetries, that provide an additional way to access SD amplitudes: via the interference between the amplitudes from O_7, O_9 and O_{10} with the LD ones. Asymmetries depend on their relative size and on the relative strong and weak phases.

  As considered in Section 1.1.7 a **CP asymmetry** can be generated in rare charm decays by NP in a presence of large weak phases. For example, in some NP models direct CP asymmetries are possible due to phases acquired by C_7 and C_9 [46, 51].

  In the SM, the weak interaction couples more to the left-handed fermions. This has a profound impact on the angular distributions observed in meson decays. NP does not necessarily respect this non-trivial chiral structure. It can, for instance, enhance the contribution of O_{10}, with a different chirality than the other operators, like those generating the LD contributions. This can also be caused by the contribution of new operators like Q'_9 of Q'_{10}, of opposite handedness to their SM *alter egos*. Any change in angular distributions could be a sign of the presence of NP presence. In particular, it could cause, for instance, sizeable **Forward-Backward asymmetries**, $\mathcal{A}_{FB}$:

$$\mathcal{A}_{FB} = \langle sgn(\cos\theta_l)\rangle = \frac{1}{\Gamma}\left[\int_0^1 d\cos\theta_l \frac{d\Gamma}{d\cos\theta_l} - \int_{-1}^0 d\cos\theta_l \frac{d\Gamma}{d\cos\theta_l}\right], \qquad (1.55)$$

where θ_l - the angle between the positive lepton and the D meson (Figure 1.18). Indeed, $\cos\theta_l = -\frac{\hat{\mathbf{n}}\cdot\mathbf{k}_+}{|\mathbf{k}_+|}$, where $\hat{\mathbf{n}}$ is the direction of the dihadron system as seen from the D^0 rest frame and $\mathbf{k}_+$ is the l^+ momentum in the dilepton CM system. There are also two others angles defined in Figure 1.18: $\cos\theta_h = -\frac{\hat{\mathbf{n}}\cdot\mathbf{p_l}}{|\mathbf{p_l}|}$, with $\mathbf{p_l}$ - the h_1 momentum in the dihadron CM system, and $\cos\phi = -\frac{\mathbf{p_l}^\perp\cdot\mathbf{k}_+^\perp}{|\mathbf{p_l}^\perp||\mathbf{k}_+^\perp|}$, where $\mathbf{p_l}^\perp$ and $\mathbf{k}_+^\perp$ are the components of $\mathbf{p_l}$ and $\mathbf{k}_+$ perpendicular to $\hat{\mathbf{n}}$. Thus, a four body decay can be described in term of five variables: two invariant masses, m_{hh} and m_{ll}, and three angles, θ_h, θ_l and ϕ.

For four body decays such as $D^0 \to VV'$ a **T-odd asymmetry** can be singled out by an angular asymmetry in the diplane angle ϕ (Figure 1.18):

$$\mathcal{A}_{T_{odd}} = \langle sgn(\sin\phi\cos\phi)\rangle = \frac{1}{\Gamma}\int_0^{2\pi}\frac{d\Gamma}{d\phi}d\phi^{*1}. \tag{1.56}$$

The final state of $D^0 \to VV'$ decays is a superposition of several angular momentum eigenstates. The interference between the corresponding amplitudes can generate T-odd asymmetries if their relative weak phase is not zero. According to [18], this can be brought by the O_1, O_2 and QCD penguin operators.

At the same time the LD branching ratios for the $D^0 \to h^+h^{(\prime)-}\mu^+\mu^-$ are predicted to be [18]:

- $\mathcal{B}(D^0 \to K^-\pi^+\mu^+\mu^-) = 6.8\times10^{-6}$, dominated by the $\bar{K}^*(\to K^-\pi^+)\rho(\to \mu^+\mu^-)$ and $\bar{K}^*(\to K^-\pi^+)\omega(\to \mu^+\mu^-)$ resonant intermediate states.
- $\mathcal{B}(D^0 \to \pi^+\pi^-\mu^+\mu^-) = 1.4\times10^{-6}$, dominated by the $\rho(\to \pi^+\pi^-)\rho(\to \mu^+\mu^-)$ and $\rho(\to \pi^+\pi^-)\phi(\to \mu^+\mu^-)$ resonant modes.
- $\mathcal{B}(D^0 \to K^+K^-\mu^+\mu^-) = 1.2\times10^{-7}$, dominated by the $\phi(\to K^+K^-)\rho(\to \mu^+\mu^-)$ resonant mode.
- $\mathcal{B}(D^0 \to K^+\pi^-\mu^+\mu^-) = 1.9\times10^{-8}$, dominated by the $K^*(\to K^+\pi^-)\rho(\to \mu^+\mu^-)$ and $K^*(\to K^+\pi^-)\omega(\to \mu^+\mu^-)$ resonant modes.

Possible effects predicted on the branching ratios or on the asymmetries from different NP models are reviewed in Tables 1.5 and 1.6. The highest effect on $\mathcal{B}$ is predicted by Minimal Supersymmetric Model with R-parity violation for $D^+ \to \pi^+\mu^+\mu^-$: $\mathcal{B} = 6.5\times10^{-6}$ [52]. Several models, as Randall-Sundrum model with a warped extra dimension [51], Littlest Higgs Model with T-parity [47] and some generic models with additional weak phases [46] or with mediating Z [18], predict enhancement of some of the asymmetries. In most cases, $\mathcal{O}(1\%)$ effects are possible. In some cases, 5% or more is possible. It is noted that in the papers we quote, predictions above 1% are not rare. However, some of them (in particular in [18] and [46]) rely on NP models parameters that have been tuned to reproduce the large experimental value of Δa_{cp} [53]. In more recent measurements of Δa_{cp}, the largest possible deviation from 0 is smaller. Consequently, we expect that these asymmetries will be smaller if the studies that predict them are updated.

[1] $\int_0^{2\pi} d\phi^* \equiv \left[\int_0^{\pi/2} - \int_{\pi/2}^{\pi} + \int_{\pi}^{3\pi/2} - \int_{3\pi/2}^{2\pi}\right]$

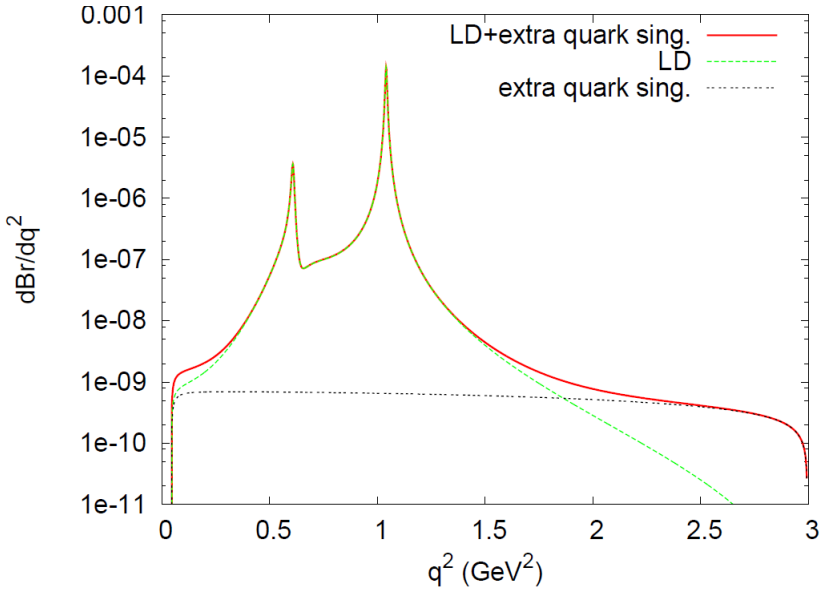


Figure 1.17: Distributions of the maximal branching ratios in the model with extra quark singlet [52] for the decay modes $D^+ \to \pi^+\mu^+\mu^-$. Full line represents the combined LD and SD contributions.

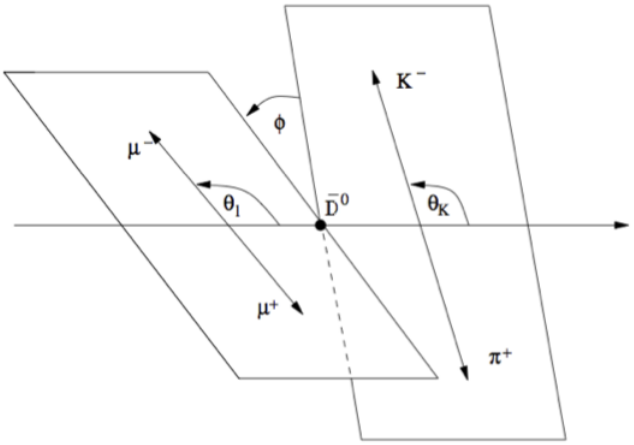


Figure 1.18: Illustration of the angles θ_l and ϕ, used in the definition of $\mathcal{A}_{FB}$ and $\mathcal{A}_{T_{odd}}$

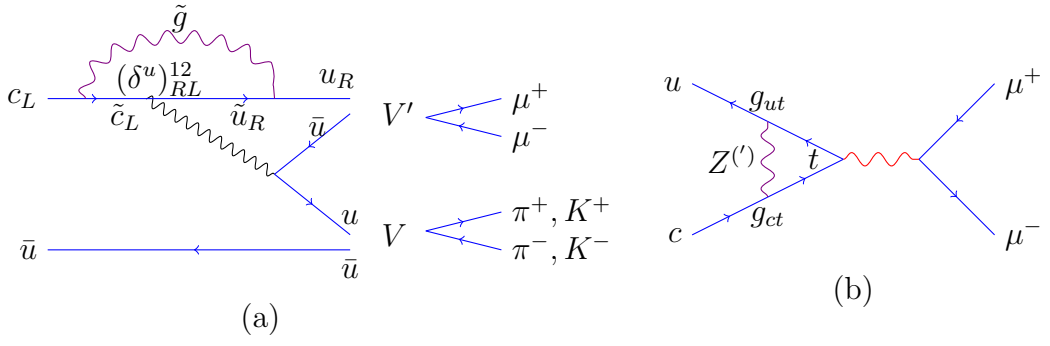


Figure 1.19: Examples of New Physics diagrams.

Model name	Main characteristics	Affected observables
Minimal Supersymmetric Model with R-parity violation (MSSM$\not{R}$)	$c \to u\mu^+\mu^-$ current is possible at tree level via down-type squark. It has a very large impact on C_9 and C_{10}, quantified in [49, 50], and updated in [52] in the light of the constraints brought in 2007 by the discovery of the D^0 mixing. In the light of recent constraints from $K \to \pi\nu\nu$ decays and charm decays, there is now little hope to get sizeable contributions from this kind of NP to the decays we are interested in. See for instance [54].	For $D^+ \to \pi^+\mu^+\mu^-$: $\mathcal{B} = 6.5 \times 10^{-6}$ [52], recently measured $\mathcal{B} < 7.3 \times 10^{-8}$@90%$CL$ [13]. New constraints from $D^0 \to \mu^+\mu^-$: $\mathcal{B} = 2 \times 10^{-8}$ [55]
Extra up-like quark singlet	New quark doublet or singlet, extended CKM matrix. The FCNC possible at tree level with cuZ coupling. A study of their impact can be found in [56], with an update in [52]. Large effects on C_9 and C_{10} were predicted there.	For $D^+ \to \pi^+\mu^+\mu^-$: $\mathcal{B} = 1.6 \times 10^{-9}$ [52]. For $D^0 \to \rho^0 l^+ l^-$: $\mathcal{A}_{FB} \sim$ few %
Littlest Higgs Model	Particular version of models with cuZ coupling, where the Higgs boson is a pseudo-Nambu-Goldstone boson of spontaneously broken global symmetry. It contains a new massive gauge boson and a new up-like quark $\tilde{t}$. Weak currents are modified, CKM is extended to be 4×3 [56, 57]. The model modifies coefficients C_9 and C_{10}. In particular, C_{10}, while $\simeq 0$ in the SM becomes of the order of C_9 [56].	For $D^+ \to \pi^+\mu^+\mu^-$: $\mathcal{B} = 8.0 \times 10^{-11}$ [57]; For $D^0 \to \rho^0\mu^+\mu^-$: $\mathcal{A}_{FB} \sim \mathcal{O}(10^{-3})$ [57]
Leptoquark model	Carrying both lepton and baryon numbers, new bosons can couple to a lepton and a quark [58].	For $D^+ \to \pi^+\mu^+\mu^-$: $\mathcal{B} = 9.4 \times 10^{-8}$ [58]
Randall-Sundrum model with a warped extra dimension	New gauge bosons appear, that mediate flavour violation. It brings a small contribution to C_9, which at the most could be comparable to the SM value (for some marginal values of the models parameters). On the other hand, as in other models, the tiny C_{10} is enhanced by several orders of magnitude [51].	For $D \to X_u\mu^+\mu^-$: $\mathcal{A}_{FB}$, $\mathcal{A}_{CP} \sim$ few % $\mathcal{A}_{FB}^{CP}$[a] $> \mathcal{O}(10\%)$ [51]

Table 1.5: Overview of the NP theoretical models that have FCNC at tree level. The estimates of affected observables are presented as well.

[a] Difference between $\mathcal{A}_{FB}$ in $\bar{D} \to X_u\mu^+\mu^-$ and $D \to X_u\mu^+\mu^-$

Model name	Main characteristics	Affected observables
Minimal Supersymmetric Model with R-parity conservation (MSSM R)	New sources of flavour symmetry breaking. In the mass insertion approach, off-diagonal elements in the squark mass matrix yield flavour changing couplings $(\delta^q_{i,j})_{H,H'}$. They allow squarks of flavour i, of helicity H and type q (up or down), to turn spontaneously into a squarks of flavour j, of helicity H' and type q. Loop amplitudes as that in Figure 1.19(a) are then possible. They enhance C_7, C_8 and C_9. This is discussed in [49, 59–61].	For $D^0 \to \rho^0\mu^+\mu^-$: $\mathcal{B} \simeq 1.3 \times 10^{-6}$ [49]
Littlest Higgs Model with T-parity (LHT)	LH Model with additional T-parity. Enhancement of the C_9 and C_7 is very small. The main effect is in fact on C_{10}, which is enhanced by orders of magnitude.	For $D \to X_u\mu^+\mu^-$: $\mathcal{A}_{FB} \sim \mathcal{O}(0.5\%)$, $\mathcal{A}_{FB}^{CP}$ up to $\mathcal{O}(10\%)$ [47]
Generic models with generated weak phases	Models that generate weak phases acquired by C_7 and C_9 without sensitive impact on C_{10} [46].	For $D^+ \to \pi(\mu^+\mu^-)_\phi$: $\mathcal{A}_{CP} \sim \mathcal{O}(1\%$-$10\%)$ [46]
Generic Z-mediated models	Loop amplitudes with an internal $Z^{(\prime)}$ and an internal top quark (Figure 1.19(b)) can bring C_9 and C_{10} up to $O(1)$, if the couplings they involve are tuned to reproduced the measured value of ΔA_{CP} [18].	For $D^+ \to h^+h^{(\prime)-}\mu^+\mu^-$ [18]: $\mathcal{A}_{T_{odd}}$ up to 8%, $\mathcal{A}_{FB}$ up to 3%

Table 1.6: Overview of the NP theoretical models that have FCNC at loop level. The estimates of affected observables are presented as well

Chapter 2

The LHCb detector at the LHC

The LHCb experiment is one of four large experiments installed at the Large Hadron Collider (LHC) at CERN. It is dedicated the study of CP violation and rare decays in the bottom and charm sectors.

The LHC is a two ring collider, producing proton-proton (pp) and heavy ion collisions. It is located near Geneva. The main ring of the LHC, almost 27 km in circumference, runs in an underground tunnel, at an average depth of about 100 m. The main goal of the LHC was the search for Higgs boson, but not only: searches for phenomena beyond the SM, precision tests of the SM, the measurement of top quark mass, studies of quark-gluon plasma are also essential elements in its program.

For proton beams the nominal centre-of-mass energy is 14 TeV. They are obtained from helium atoms. To access such high energy they are accelerated in several steps (see Figure 2.1): firstly, the protons are accelerated by the linear accelerator (LINAC2), they are then injected to the PS Booster. They are subsequently accelerated by the Proton Synchrotron(PS) followed by the Super Proton Synchrotron (SPS) before being injected into the LHC, where they are accelerated until they reach the desired energy.

Four large experiments and three smaller ones benefit from the high performance of the LHC: ATLAS [62], CMS [63], ALICE [64], LHCb [15] and TOTEM, LHCf, MoEDAL. ATLAS and CMS were designed mainly to search and to study the properties of the Higgs boson and to search for New Physics particles. The ALICE experiment focuses on the physics of strongly interacting matter at extreme energy densities, testing the existence and properties of the quark-gluon plasma. The TOTEM (Total Cross Section, Elastic Scattering and Diffraction Dissociation) aims at total proton-proton cross section measurements, at studies of elastic scattering and of diffractive dissociation processes. The LHCf (LHC-forward) is designed for measurements related to astroparticle physics and could help to explain the origin of ultra-high-energy cosmic rays. The purpose of MoEDAL (Monopole and Exotics Detector At the LHC) is to directly search for the Magnetic Monopole or Dyon and other highly ionizing Stable (or pseudo-stable) Massive Particles (SMPs) at the LHC.

The primary goal of the LHCb detector is to look for indirect evidences of New Physics in CP violation and rare decays of charm and beauty hadrons. An overview of the LHCb detector is presented in this chapter.

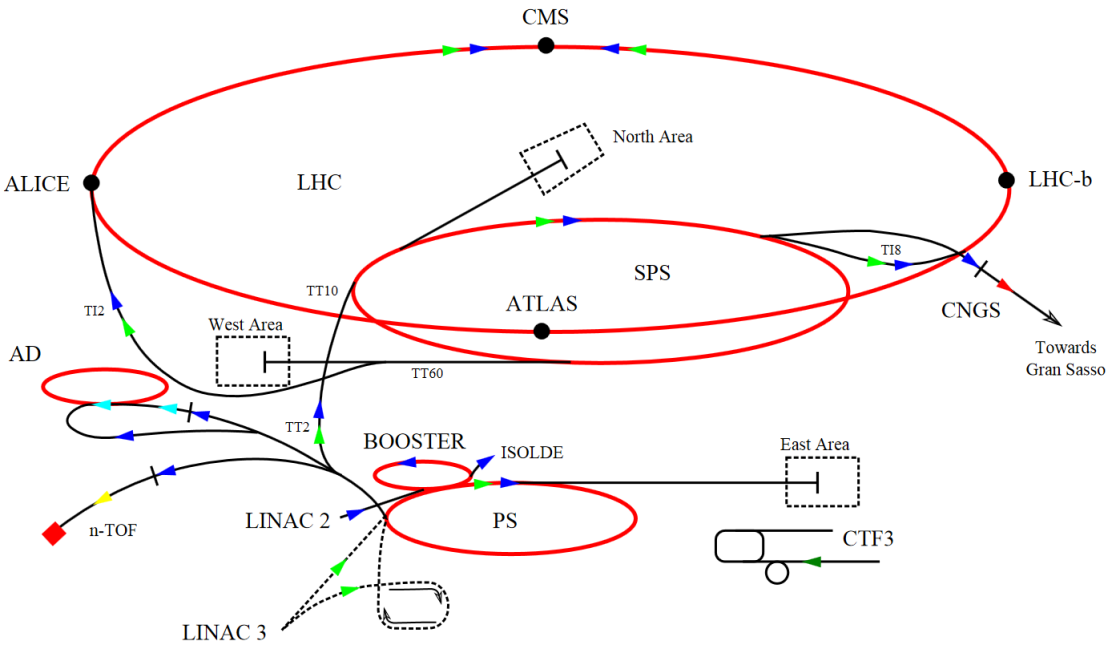


Figure 2.1: The CERN accelerator complex.

2.1 General overview

During more than two years of running LHCb has so far recorded a luminosity of about $1.1\,\text{fb}^{-1}$ at centre-of-mass energy $\sqrt{s} = 7$ TeV (2010-2011) and more than $2\,\text{fb}^{-1}$ at $\sqrt{s} = 8$ TeV (2012). During this time 26×10^{10} of $b\bar{b}$ and 59×10^{11} $c\bar{c}$ pairs were produced within the LHCb acceptance.

Thanks to an excellent trigger, vertex reconstruction and the general performance of the detector, precision tests of the SM in the heavy flavour sector are possible. The core parts of the LHCb physics program are: measurements of CP- and angular asymmetries asymmetries ($B_s^0 \to J/\psi\phi$ [1], $B \to K\mu^+\mu^-$ [2–4], $B \to DK^{(*)}$ [5, 6]), New Physics searches in rare decays ($B_s^0 \to \mu^+\mu^-$ [7], $B \to K\mu^+\mu^-$ [8], $B^0 \to K^*\gamma$ [9] etc.). Charm physics has also become an important part of the physics program: this includes searches for CP-violation [10], the mixing [11] in $D \to h^+h'^-$ and for rare decays $D^0 \to \mu^+\mu^-$ [12], $D^+ \to \pi^+\mu^+\mu^-$ [13], $D^0 \to hh\mu^+\mu^-$ [14].

The LHCb detector [15] is a single-arm spectrometer with a forward angular coverage [10 mrad, 250 mrad] in the non-bending plane and [10 mrad, 300 mrad] in the bending plane, which corresponds to a $2 < \eta < 5$ pseudorapidity range. A schematic view of the detector is presented in Figure 2.2. This forward detector geometry is justified by the fact that $b\bar{b}$ and $c\bar{c}$ quark pairs are produced in pp collisions at very low angles, i.e. in the forward or backward regions. Indeed, $b\bar{b}$ and $c\bar{c}$ pairs are produced by highly boosted virtual gluons from pp deep inelastic scattering.

The pp collisions' primary vertex (PV) is reconstructed by VErtex LOcator (VELO) (Section 2.2.1). The produced particles go through the first RInging CHerenkov detector (RICH-1) (Section 2.4.1) and the first pair of tracking stations, named the Tracker

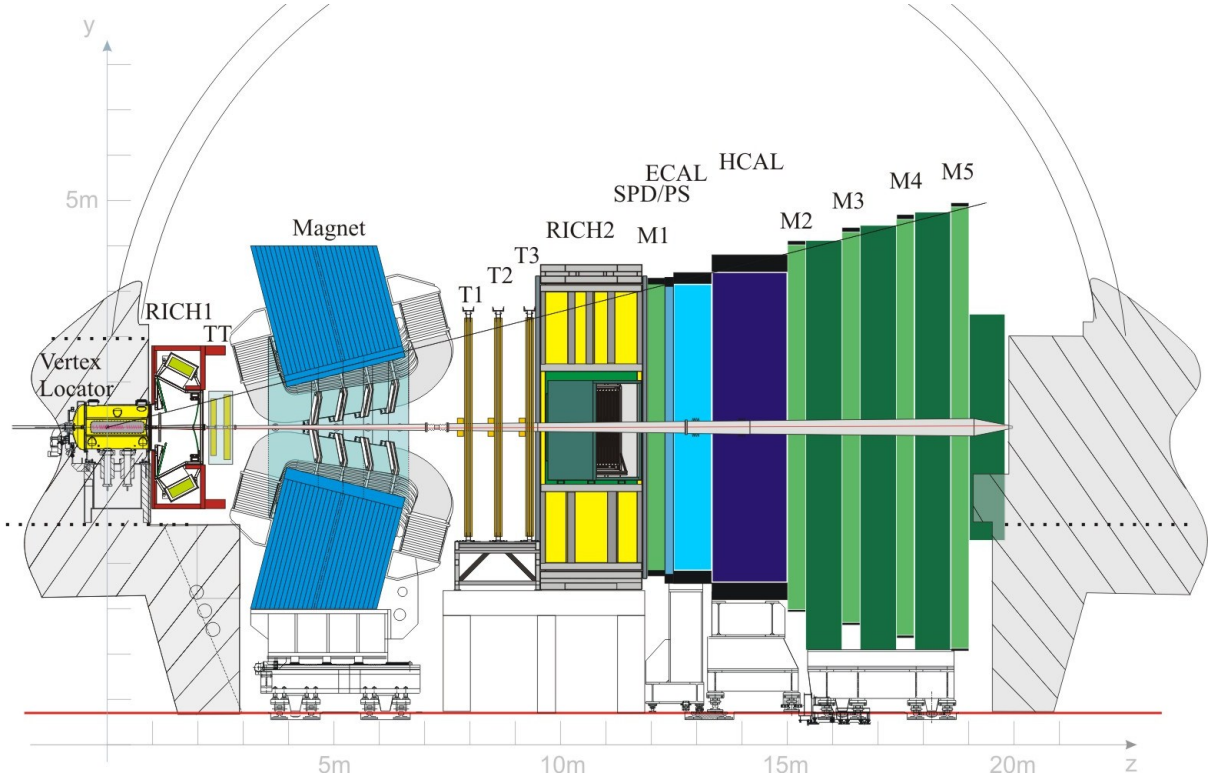


Figure 2.2: View of the LHCb detector.

Turicensis (TT) (Section 2.2.2). The tracks of charged particles are bent by a magnetic field (Section 2.3) before they reach tracking stations T1, T2, T3 (Section 2.2.2), RICH2 (Section 2.4.1), the calorimeter system (SPD, PS, ECAL, HCAL presented in Section 2.4.2) and the muon stations M1, M2, M3, M4, M5 (Section 2.4.3).

2.2 Tracking

The LHCb tracking system consists of five subsystems: the VErtex LOcator (VELO), the Tracker Turicensis (TT) and the three T1-T3 tracking stations. The VELO and TT are silicon strip detectors, T1-T3 stations use silicon microstrips in their inner part (close to beam pipe) and straw-tubes in outer region.

2.2.1 VELO

The VELO is dedicated to the reconstruction of displaced vertices. Hadrons containing a b or a c quark are long lived (typical lifetime for B mesons is $\sim 10^{-12}$ s and D mesons - $0.5 - 1 \times 10^{-12}$ s). Thus their decay vertices are displaced by ~ 1 cm from the primary vertex where they are produced. This important feature of b and c hadrons decays is used for their selection at trigger level and by offline analyses.

The VELO comprises 21 stations of silicon modules, each composed of 2,048 silicon strips, perpendicular to the beam pipe. Two types of modules are used: R-sensors used for the measurement of radial distances to the beam line and ϕ-sensors giving the

azimuthal coordinates of traversing charged particles. The third spatial coordinate is derived from the sensor position along the beam line. The sensors are closer to the beam than the aperture allowed by the LHC during injection. That is why the VELO stations consist of two retractile parts (Figure 2.3). During injection or unstable beams these parts are separated by 6 cm. In stable beam conditions they overlap, covering the full acceptance. The full system is contained in a vessel that maintains vacuum and separation from the beam vacuum, thanks to a thin aluminium sheet.

Two additional modules are installed upstream of the VELO sensors (pile-up veto system). These provide information about the positions of vertexes as well as the total backward charged track multiplicity. The system is used to measure the backward charged track multiplicity and to detect multiple interaction events.

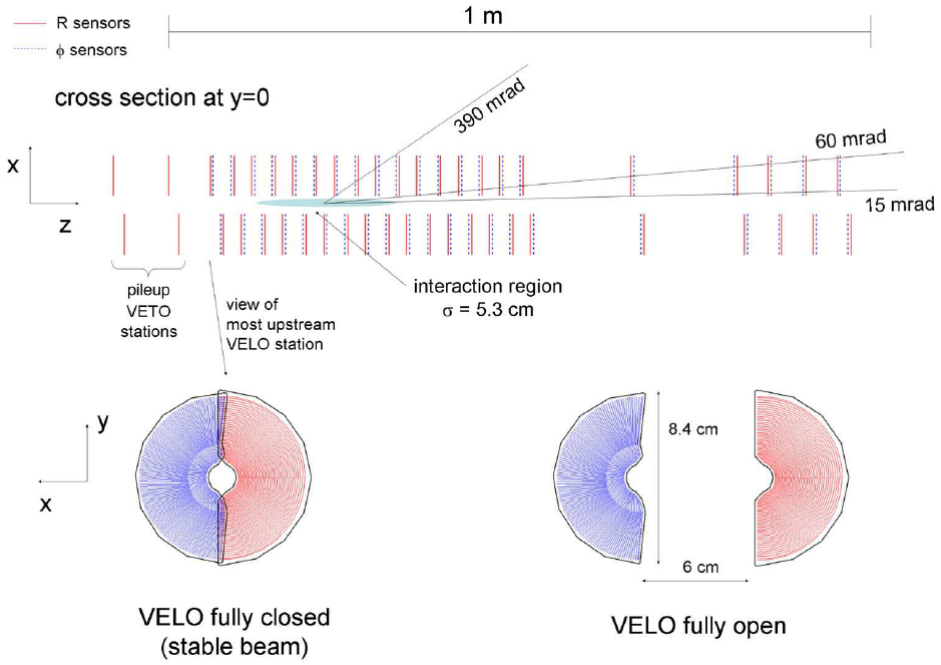


Figure 2.3: Scheme of the VELO detector layout. The front face of the first modules is also illustrated in both the closed and open positions. The two pile-up veto stations are located upstream of the VELO sensors. The R-sensors are red and ϕ-sensors are blue.

The VELO is the most precise sub-detector of LHCb. The size of the smallest (largest) pitches is 40 (102) µm and 38 (97) µm in the R-sensors and the ϕ-sensors, respectively. The hit resolution is $\sim 4\,\mu m$, the best vertex resolution in transverse plane is $\sim 13\,\mu m$ and $\sim 70\,\mu m$ in the z-direction. The best impact parameter resolution is $\sim 13\,\mu m$ for tracks with high transverse momentum.

2.2.2 Tracking stations

The role of the tracking stations is to reconstruct the trajectories of charged particles and measure their momentum. There are two types of tracking stations: silicon microstrip trackers and straw-tube drift chambers. The TT and the inner part of tracking stations T1, T2, T3 are of the first type, the outer part is of the second type. Information from

the tracking system also provides the fast estimation of particles momenta used by the High Level Trigger.

The Silicon Tracker: Tracker Turicensis and Inner Tracker

The Silicon Tracker (ST) comprises two detectors: the Tracker Turicensis (TT) and the Inner Tracker (IT). Both of them use silicon microstrip sensors with a strip pitch of about 200 µm providing single hit spacial resolution of 50 µm. Each station has four detection layers in an arrangement *(x-u-v-x)*. Strips at the first and the last layers are vertical, while for u and v they are rotated by a stereo angle of $-5°$ and $+5°$ respectively. Such a pattern permits the reconstruction all of three spatial coordinates. In Figure 2.4 one can find an example view of an ST station, TT. Different readout strip lengths are chosen for different regions of the detector to keep maximum strip occupancies at the level of a few percent while minimizing the number of readout channels.

The TT system is located upstream of the magnet. Its size is 150 cm × 130 cm. It provides a full coverage of LHCb's acceptance. There are two types of detection modules with sensors split into either two or three readout sections (Figure 2.5). In the first configuration, one readout section connects four or three sensors. In the second type of module there are sections with four, two and one sensors. The latter is used in the region close to beam pipe as particle flow is higher here.

The IT systems are located downstream of the magnet. They cover the inner region of each of the three tracking stations T1, T2 and T3 (Figure 2.7), where the occupancy is higher than in the outer region, and higher resolution is required. Each 240 cm × 40 cm IT system contains four detector boxes surrounding the beam pipe. They are installed with overlaps to avoid gaps and to help alignment (Figure 2.6).

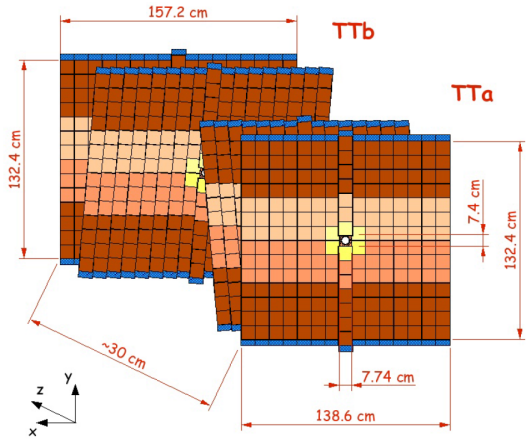


Figure 2.4: View of a TT detector module with three readout sections.

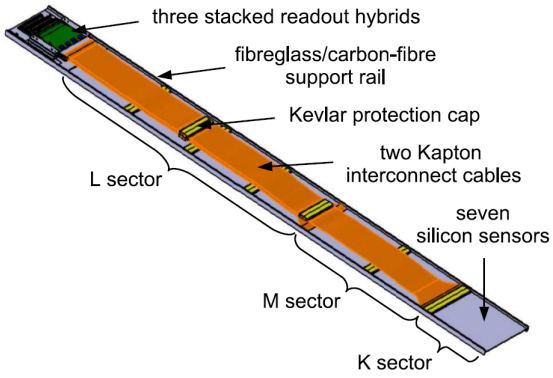


Figure 2.5: View of a TT detector module with three readout sections.

Outer Tracker

The LHCb Outer Tracker (OT) is a drift-time detector. It is designed as an array of individual, gas-tight straw-tube modules. Each module has two monolayers of drift-tubes with diameters of 4.9 mm. The tubes are filled with a mixture of CO_2 (30%) and

Argon (70%) which provides a drift time below 50 ns and sufficient drift-coordinate resolution (200 μm).

As in the case of the ST, the OT stations have four layers, oriented in an (x-u-v-x) geometry (see previous subsection) for each T1, T2 and T3 (see Figure 2.7).

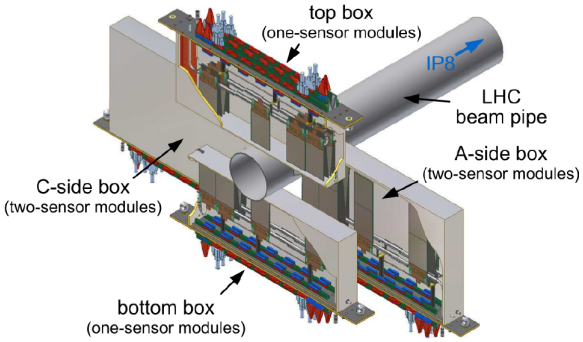


Figure 2.6: The IT station layout.

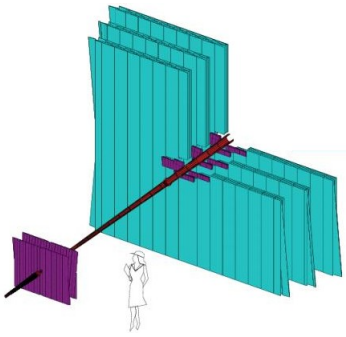


Figure 2.7: Scheme of the TT, T1, T2 and T3 stations. Cyan colour corresponds to the OT, violet - the ST.

2.3 Dipole magnet

The dipole magnet of LHCb is a warm magnet with integrated field of 4 Tm. It has two identical coils of conical saddle shape, which are placed mirror-symmetrically to each other in the magnet yoke (Figure 2.8). The non-uniformities of the field are at the order of 1%. The magnetic field bends the charged particle trajectories depending on their momenta. The measurement of their curvature by the tracking system yields the momentum of the charged particles. The polarity of the magnet can be reversed in order to study systematic effects due to possible left-right detection asymmetries. The magnetic field is vertical so two configurations are possible: $MagUp$ and $MagDown$. The region subject to this field extends essentially from $z = 2.5$ m to $z = 7.95$ m (Figure 2.9).

2.4 Particle identification

Particle identification is fundamental for LHCb data analyses. Many decay modes have the same multiplicity but different final state particles. Thus, one must be able to distinguish between e, γ, μ, π, K and p. There are four subsystems for this task: two Ring Imaging Cherenkov detectors (RICH1 and RICH2), electromagnetic and hadron calorimeters (ECAL and HCAL) and five muon stations.

2.4.1 Ring Imaging Cherenkov detectors

Kaon-pion separation is an essential requirement for many LHCb physics studies. Two Ring Imaging Cherenkov detectors (RICH1 and RICH2) are used for this purpose. A

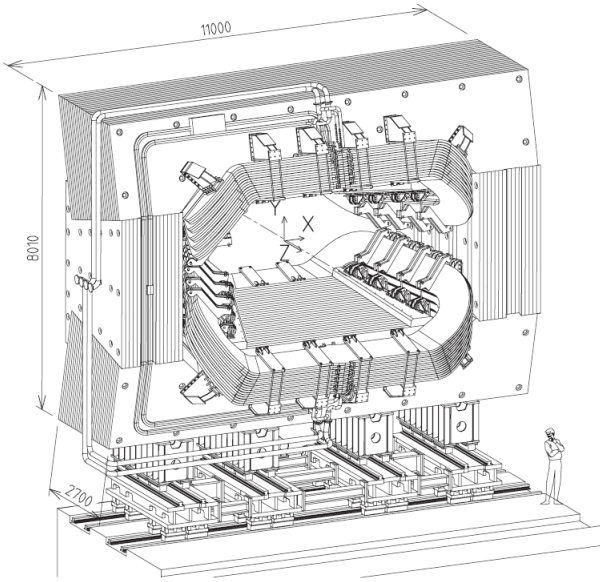


Figure 2.8: Scheme of the LHCb dipole magnet. The interaction point lies behind the magnet.

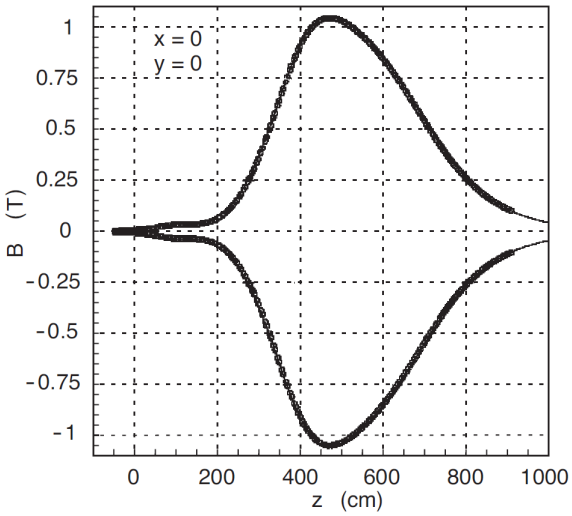


Figure 2.9: Magnetic field along the z axis.

charged particle passing thought a medium faster than speed of light in this medium produces a cone of Cherenkov light. The opening angle of the cone depends on the medium and the velocity of the passing particle. By combining this information with momentum derived from the tracking system one can deduce the mass of the particle, i.e. identify it.

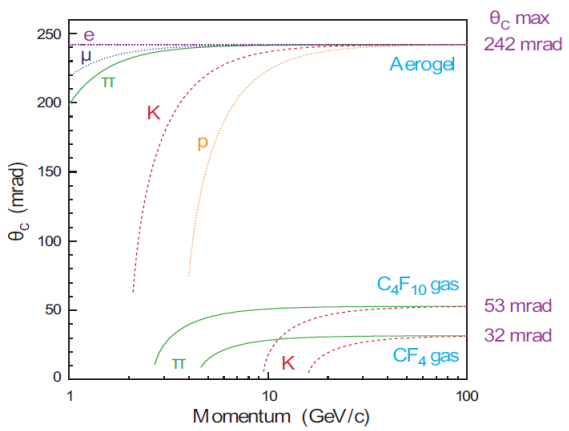


Figure 2.10: Cherenkov angle versus particle momentum for the RICH radiators.

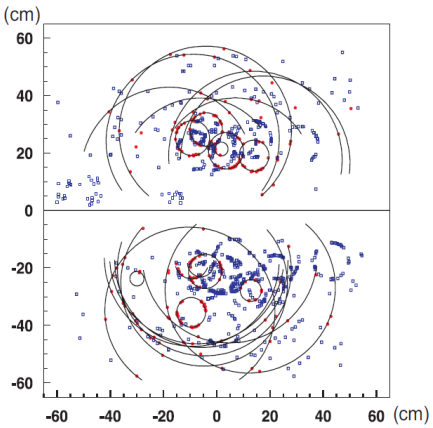


Figure 2.11: Display of a typical LHCb event in RICH1.

The RICH1 and RICH2 cover the full momentum range of typical B and D decay products (Figure 2.10). The RICH1, installed just after the VELO, illustrated in Figure 2.12, covers the low momentum range $\sim 1 - 60$ GeV/c. It uses aerogel and C_4F_{10} gas as radiators. The RICH1 has a wide angular acceptance: from ± 25 mrad to ± 300 mrad (horizontal) and ± 250 mrad (vertical). On the other hand the RICH2

(Figure 2.13) extends LHCb's particle identification capabilities to higher momentum range: from $\sim$ 10 GeV/c to 100 GeV/c and beyond. Radiator material is CF_4. The RICH2 has a limited angular acceptance from $\pm$15 mrad to $\pm$120 mrad (horizontal) and $\pm$100 mrad (vertical).

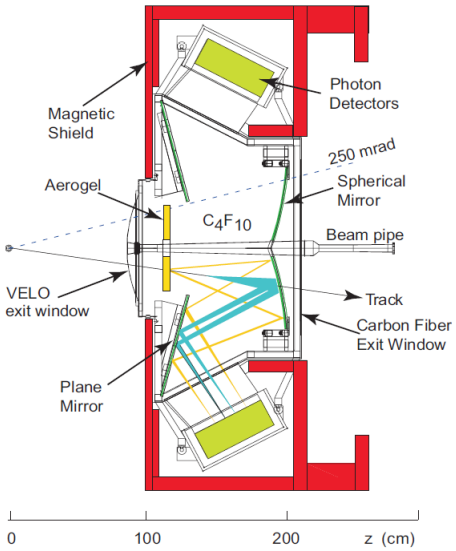


Figure 2.12: Side view schematic layout of the RICH1 detector.

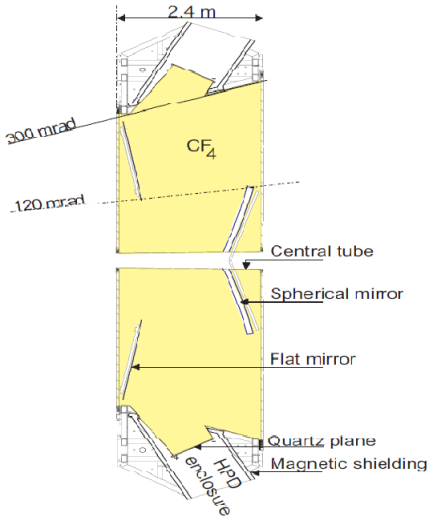


Figure 2.13: Side view schematic layout of the RICH2 detector.

Both RICH1 and RICH2 use a combination of spherical and flat mirrors to reflect the Cherenkov light out of the spectrometer acceptance, where a matrix of Hybrid Photon Detectors (HPD) is located. The system detects Cherenkov rings (Figure 2.11).

2.4.2 Calorimeter system

The LHCb calorimeter system, placed downstream of RICH2, consists of four subsystems: a Scintillator Pad Detector (SPD), PreShower detector (PS), separated by a lead converter, and Electromagnetic (ECAL) and Hadronic (HCAL) Calorimeters (Figure 2.14). All four subdetectors work with "shashlik" or "tails" technology: they are sampling devices using scintillator material inter-layered with absorber. Scintillation light is transmitted to PhotoMultiplier tubes (PMT) by WaveLength-Shifting (WLS) fibres. The PS, SPD and ECAL use lead as an absorber and HCAL uses iron.

The PS and SPD is a matrix of cells, which are a 15 mm thick scintillators. On the other hand, the ECAL is composes of shashlik cells, where one cell is composed of 66 alternating layers if 2 mm thick lead and 4 mm thick scintillator, is illustrated in Figure 2.15. The HCAL employs a non-typical structure where the scintillating tiles are arranged parallel to the beam pipe as shown in Figure 2.15. The HCAL uses 1 cm thick tiles of iron alternated with scintillator layers of 3 mm thick.

A particle passing through the calorimeter absorber material creates a shower. The minimum ionizing particles produced in this shower cause the production of light in the scintillators. The total quantity of light depends on the number of the particles, thus on the total energy E in the shower. Besides the energy, the position is also measured.

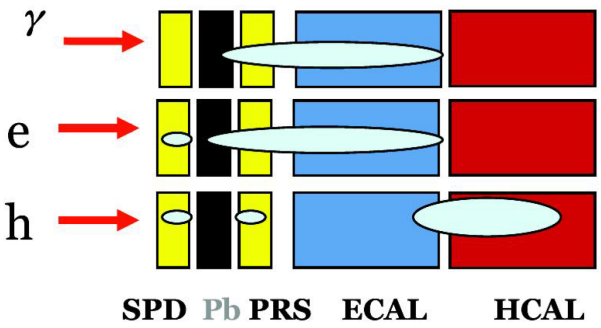


Figure 2.14: Scheme of the procedure of the photon, electron and hadron separation with the LHCb Calorimeter stations.

A schematic indicating the separation between electrons, photons and hadrons is shown in Figure 2.14. It can be explained as following:

- photons do not leave hits in the SPD station but will start to interact in the lead converter and create electromagnetic shower in PRS and ECAL;
- electrons have the same behaviour as photons at the level of PRS+ECAL but leave hits in the SPD;
- hadrons leave part of their energy in the SPD, PRS and almost nothing in the ECAL. Essentially all the energy is absorbed by the HCAL.

All calorimeter sections are separated into lateral segments as the hit density is higher in the region close to the beam pipe. The PS, SPD and ECAL are divided into three sectors, the HCAL into two, as presented in Figure 2.17.

The high granularity provides a good energy resolution: $\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{s}} \oplus 1\%$, where E is in GeV, for the ECAL and $\frac{\sigma_E}{E} = \frac{69\%}{\sqrt{s}} \oplus 9\%$ for the HCAL. These resolutions are sufficient to use the calorimeter in the L0 trigger system. The front-end electronics of the systems are designed to be as fast as possible in order to contribute to the trigger.

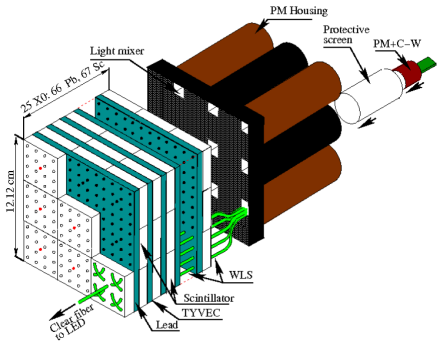


Figure 2.15: Illustration of ECAL cell stricture.

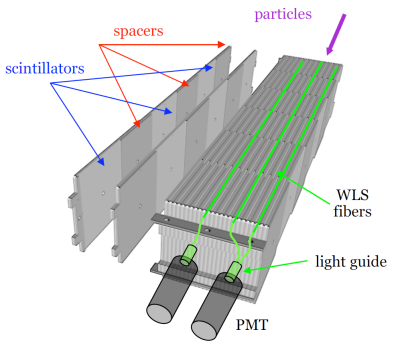


Figure 2.16: Illustration of HCAL cell stricture.

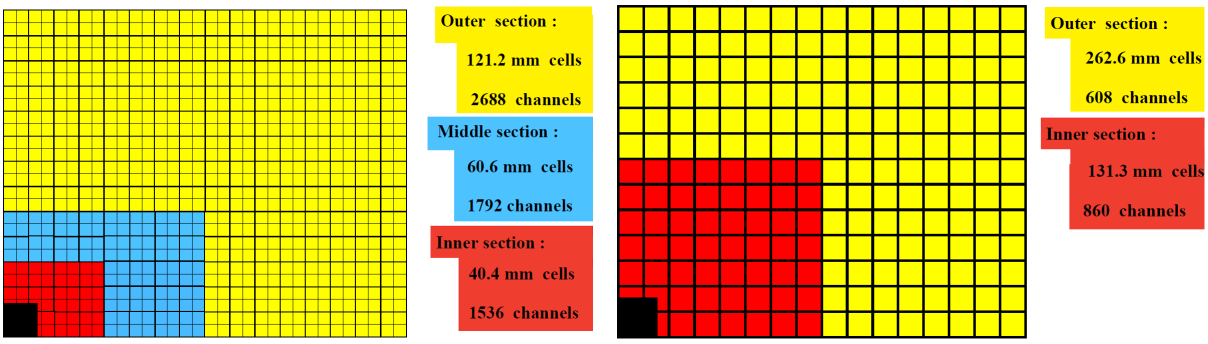


Figure 2.17: Schematic view of cell segmentation in SPD, PS and ECAL (left) and HCAL (right) for the one firth part of the detector.

2.4.3 Muon chambers

The LHCb muon system is used for the muon-based line of the L0 trigger (Section 2.5) and offline to separate muons from hadrons. This is a very important task for rare charm decays with a pair of muons in the final state, since purely hadronic decays with the same multiplicity and topology are produced at least 10^4 times more often.

There are five muon chambers: M1 is located upstream of the SPD and M2-M5 are downstream of the HCAL. The last four stations are interleaved with 80 cm thick iron absorbers. Muons transverse typically 3 to 5 stations, unlike hadrons, which are absorbed either by the calorimeter of the absorbers. Trajectories from the tracking stations are extrapolated to the muon system and if there are hits found in the muon stations, such tracks are considered to be muon candidates.

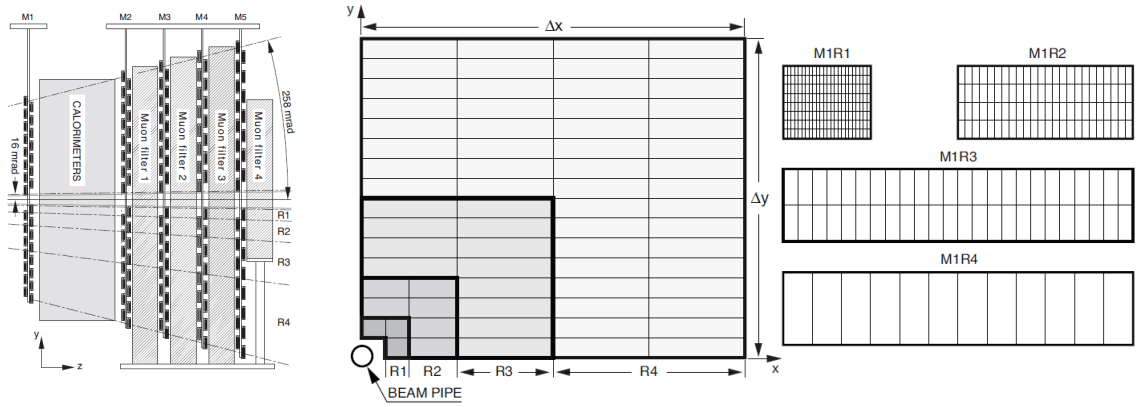


Figure 2.18: Overview of the LHCb muon stations (left) and their segmentation (right).

As in the case of the calorimeters there are also regions with different granularity (Figure 2.18) R1, R2, R3, R4 with a segmentation scale in the ratio 1:2:4:8. With this geometry, the particle flux is expected to be the roughly the same over the four regions of a given station. All muon stations use the technology of Multi-Wire Proportional Chambers (MWPC) except for the inner region R1 of the station M1, where the triple-GEM is used.

The Muon stations allow muons to be reconstructed with a 20% of p_t resolution, sufficient for the L0 trigger.

2.5 Trigger

The bunch crossing frequency at the LHC is 40 MHz. The rate of visible interactions is typically 10 MHz. At the same time, only 4-5 kHz of data can be saved to a disk. Necessary selection is the task of the trigger system. It aims at an efficient selection of interesting events for physics analyses, while rejecting uninteresting events as strongly as possible. The scheme of the LHCb trigger system in 2011-2012 is presented in Figure 2.19.

There are two levels of trigger at LHCb: a hardware Level-0 (L0) trigger and a software trigger, called the High Level Trigger (HLT), which are presented in this section.

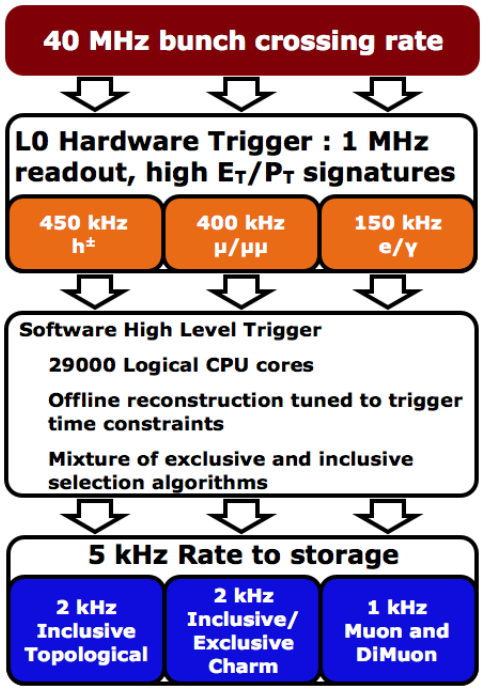


Figure 2.19: Scheme of the LHCb trigger system during 2011-2012 data taking period.

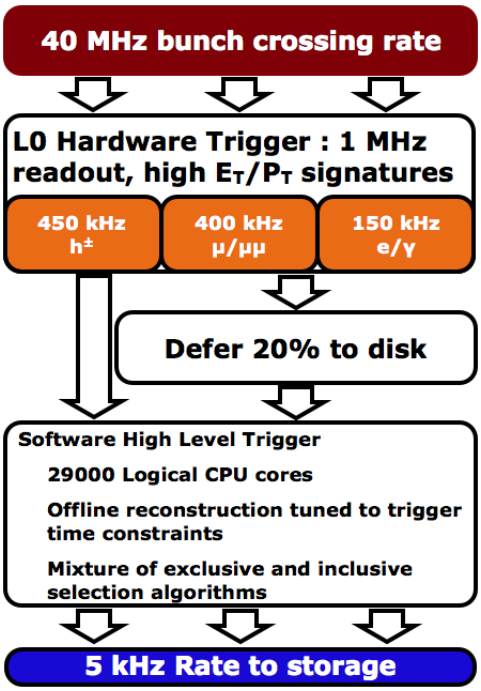


Figure 2.20: Scheme of the LHCb trigger system with deferring used during 2012 data taking period.

2.5.1 L0

The Level-0 (L0) is a hardware trigger based on custom made front-end electronics. It operates synchronously with 40 MHz bunch crossing frequency. The L0 reduces the rate to 1 MHz, at which the entire detector can be read out.

The trigger exploits the fact that heavy B and D mesons produce secondary particles with large transverse momentum (p_t) and transverse energy (E_t). The L0 trigger attempts to reconstruct the highest E_t hadron, electron or photon in the calorimeter systems and the two highest p_t muons in the muon chambers. Thus, the trigger uses information from calorimeter and muon stations.

Both the Calorimeter Trigger system uses all calorimeter stations, i.e. the SPD, PS, ECAL and HCAL, forms clusters by adding 2×2 cells and computes the transverse energy of these clusters. The trigger selects the cluster with the highest E_t in each of the 32 cells region. If the candidate has a transverse energy of a cluster in HCAL and ECAL above some threshold value, it is saved as a "hadron" or "electromagnetic" object. For the muons two types of L0 objects are possible: "muon" with a threshold on the transverse momentum and "dimuon" with a requirement placed on the sum of the transverse momenta of the two muons. In the muon chambers the track finding algorithm looks for hits in all five chambers. The L0 muon trigger selects the two muons with the highest transverse momenta for each quadrant of the muon detector.

Thus, in the L0 trigger there are several types of objects, e, γ/π^0, h, μ and $(\mu\mu)$, that can fire for the event to be triggered on. The corresponding L0 trigger lines are `L0Electron`, `L0Photon`, `L0Hadron`, `L0Muon` and `L0Dimuon`. In the analyses discussed in this thesis `L0Muon` and `L0Hadron` are used. In 2012 data taking period the `L0Muon` line selects muon objects with a $p_t > 1.76$ GeV/c and the `L0Hadron` lines has a threshold at $E_t > 3.5 - 3.74$ GeV/c.

The total number of SPD hits provides a measure of the charged track multiplicity in the event in order to remove events that are too big to be used in offline analyses, it can also be used to reduce the L0 trigger rate.

2.5.2 HLT

Events passing the L0 are processed by the High Level Trigger (HLT). This is a software trigger that processes events asynchronously to the bunch crossing on the Event Filter Farms (EFF) where a C++ application runs. The HLT trigger uses full event data and reduces the L0 trigger rate to 4-5 kHZ after which the data are stored to disk.

HLT1

The first step of the software trigger, HLT1, performs a partial event reconstruction using the information from the VELO, the Tracking Stations and the Muon system. Depending on the type of the L0 decision the event is passed to different HLT1 selection lines. The general strategy is to refine the p_t measurement of the L0 trigger by matching the corresponding L0 objects to track segments in the Tracking Stations or in the VELO. For the confirmed L0 candidate, the track impact parameter with respect to the primary vertex can be determined and used to reject events. Secondary vertices are reconstructed using additional tracks and cuts on the vertex properties and the two track invariant mass are applied to reject background. The HLT1 lines are divided in three categories according to the L0 object that is used as its input: muon, hadron and electromagnetic.

For the muon trigger a fast muon identification is performed as well by matching

VELO tracks with muon chamber hits. For example, the `HLT1MuonTrack` algorithm works in the following way: using as seed a confirmed `L0Muon` candidate and additional tracks in tracking stations and VELO, a secondary vertex is reconstructed. A decision is taken by cutting on the flight direction, the mass of the muon+track system and either on the invariant mass or on the impact parameter.

The generic HLT1 line commonly used is `HLT1TrackAllL0`, executed for all events accepted by L0 and satisfies criteria: displaced from every reconstructed PV in the event (impact parameter with respect to each PV > 0.1 mm), and that has a relatively large estimated p_t ($p_t > 1.7$ GeV/c). The information from VELO, TT and three tracking stations are combined to find candidates. The line is designed to select decays which are significantly displaced from a PV. The HLT1 reduces the rate to 30 kHz.

HLT2

The rate after the HLT1 trigger is low enough to perform quasi offline-like full event reconstruction at the level of the HLT2 trigger. VELO tracks are built using a seeding algorithm, and their measured momentum from the trackers is required to be $p >$ 5 GeV/c and $p_t > 0.5$ GeV/c per track. Moreover, muon and electron identifications are possible by matching tracks in muon stations and ECAL.

Thus, at the level of the HLT2 it is possible to apply event selection criteria corresponding to those of the offline analyses. In practice, the HLT2 contains a series of *lines*, each of them selecting a particular (group of) decay(s), more or less, inclusively. The used criteria are in general quite close to the ones used by the subsequent offline selections. It is particularly true for the exclusive selections or lines selecting channels with a lower priority. The bandwidth granted to each line is the main constraint in their design, and is modulated according to LHCb physics priorities. This flexibility is one of the strong points of this highly performant trigger. Dedicated lines were developed for the analyses described in this thesis; they are discussed in Section 4.1.

During 2012 data taking period, the so-called "deferred HLT2", was also used (Figure 2.20). 20% of the L0 accepted events were temporarily saved on the unused EFF local disks and processed during the next period between LHC fills. This procedure allowed tracks to be reconstructed and used with $p > 3$ GeV/c and $p_t > 0.3$ GeV/c instead of $p > 5$ GeV/c and $p_t > 0.5$ GeV/c as was the case in 2011. Thus, the performance of the charm decays selections is enhanced significantly.

2.5.3 Categories of the trigger decisions: TIS, TOS, Dec

At any level of the trigger an event can be classified in three non-exclusive categories:

- Trigger On Signal (TOS): Events which are triggered on the signal decay independently of the presence of the rest of the event. The TOS criterion is satisfied if there exists at least one trigger object all of whose tracks have overlap with the signal.

- Trigger Independent of Signal (TIS): Events which are triggered independently of the presence of the signal. In order for an event to be TIS, there must exist at least one trigger object which does not have any overlap with the signal. TIS

events are trigger unbiased except for correlations between the signal decay and the rest of the event.

- Trigger Decision (Dec): Events which are triggered either by signal trigger (TOS) or by the trigger independent of the signal (TIS) without separating these two categories.

2.6 LHCb data processing, stripping

The different LHCb applications are built in a software framework called Gaudi [65] (Figure 2.21). Data recorded by the LHCb detector is reconstructed with the Brunel [66] application and stored in "data summary tape" (dst) files for physics analysis. The stripping is a first, loose selection which is applied on these files, dividing the data in different groups or streams depending on the type of analysis for which the set of cuts is conceived. Doing this first selection centrally notably reduces the processing time for analysis. In each stream, different stripping lines are written that select candidates for specific decays on which exhaustive physics analyses will be performed. The selected candidates from one stream are stored in output files, ready to be used by analysts. Examples of stripping streams are "calibration", "charm", "Bhadron", "semileptonic" and so on. Stripping selections are revisited and broadened from time to time and new processings and stripping versions are released and run periodically, for new and older data.

D hadron candidates selected by the stripping for physics analysis are retrieved by running the DaVinci application on the stripped data files. The decay of interest is specified inside a Python script using LHCb specific classes, and a ROOT tree [67] is produced that contains as leaves the properties of each one of the particles in the decay chain. The physics analyses are then performed on these ROOT files.

2.7 Upgrade

One of the most important feature of the Upgrade of LHCb [68, 69] is its new trigger strategy. The front-end electronics will be upgraded to allow the event readout at the LHC clock rate, and send this information to an entirely software trigger (HLT). This means that data acquisition and event building will be done at the full rate of 40MHz. The ability of a software trigger to treat complex events allows the luminosity to be increased. We expect to run at a luminosity of 2×10^{33} cm^{-2}s^{-1}.

Ideally then, the L0 hardware trigger would no longer be used. It is one of the main limitations of the current system since it has to apply hard E_t cuts to hadronic decays, causing an important loss in efficiency. A software version of the current L0, called Low Level Trigger (LLT), will be implemented in the event building stage. It will allow to reduce the input rate to the HLT trigger during the commissioning phase of LHCb upgrade running or in the early phase of the data taking, if not enough CPUs are available in the PC trigger farm.

The other main changes to the detectors are:

- a 40MHz readout electronics for all subsystems;

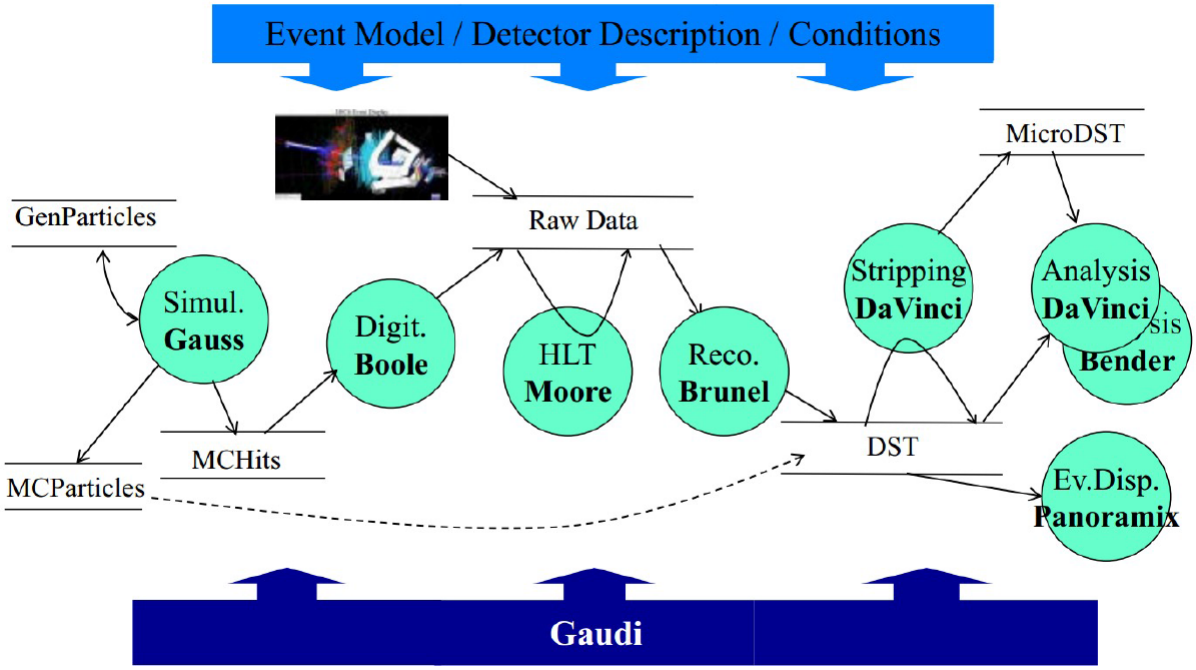


Figure 2.21: LHCb software applications.

- the VELO, tracking stations will be upgraded completely;
- the PS, SPD sub-detectors, M1 muon station and aerogel in RICH1 will be removed.

2.7.1 Trigger

The Low Level Trigger (LLT) will reduce the event rate to 15-30 MHz before the HLT stage. Here a full track reconstruction is performed as well as particle identification. Finally, offline like selections are applied. The rate is expected to be reduced until 20-100 kHz (Figure 2.22).

At a luminosity of 10^{33} cm^{-2}s^{-1} the data taking with the current L0 trigger is inefficient, because the current p_t-triggered hardware readout will be saturated. It has been shown in 2010 that with 1MHz bandwidth it is impossible to profit from the luminosity larger than $\sim 2-3\times10^{33}$ cm^{-2}s^{-1}. The new Low Level Trigger is developed to perform a selection enriched in interesting events by using looser p_t cuts. As the PreShower and Scintillating Pad detectors will be removed, the LLT will not distinguish between photons and electrons nor have a veto on hadron-induced showers. The LLT will be based on the hardware used for the existing LHCb L0 trigger: the Calorimeter and Muon systems. The ECAL and HCAL detectors are read out by new front-end boards equipped to handle the new architecture but with a separate data path for the trigger. The goal is to build ECAL and HCAL clusters 2×2 cells and select the ones with the highest transverse energy in the 32 cells region in software, in PC doing the event building. Since the SPD is removed the occupancy of the detector will be

controlled directly by HCAL and ECAL. The muon trigger will use the same hardware as the L0 muon trigger, with the distinction that the first muon station (M1) will no longer be present, hence the momentum will be determined by M2 and M3.

A more flexible and efficient HLT will be modified in order to process the higher input data. The rate at which events are accepted by the trigger for permanent storage will increase from the current 5kHz to an estimated 20kHz. The combination of a more efficient software trigger and the increased rate of data collection is estimated to increase the annual yield of many charm decay modes by an order of magnitude.

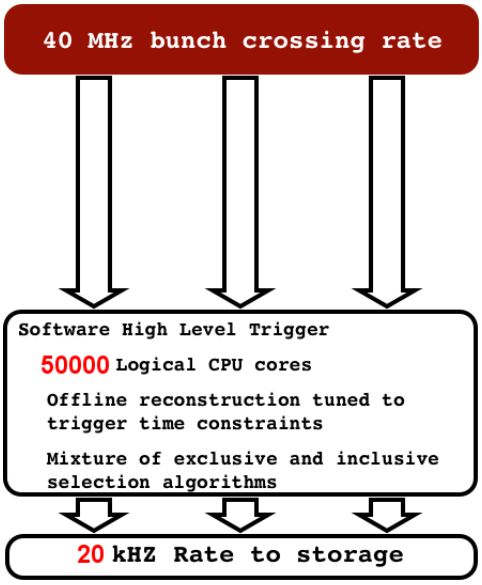


Figure 2.22: Scheme of the LHCb trigger system after upgrade.

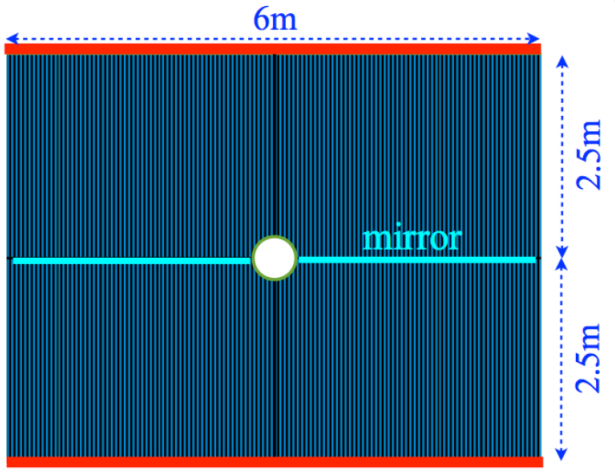


Figure 2.23: Layout of scintillating fibre detector(SciFi). The SiPMs are in red.

2.7.2 Tracking system

Due to increase of occupancy of the detector the tracking system has to be changed. The VELO will be completely replaced by new hybrid pixel detector, VELO-II [70]. The sensors will be changed to more thin ones, the system will be placed closer to the beam and new front-end electronic will be used.

The TT station will be replaced by a silicon strip detector, the Upstream Tracker (UT) [71]. As before it will have four layers (x-u-v-x) and finer segmentation in the regions around the beam pipe, but closer to it. New readout chip is needed as well.

The T1, T2, T3 stations will be replaced by systems of scintillating fibre detectors (SciFi) [71], illustrated in Figure 2.23. The fibres are 2.5 m long with a diameter of 250 μm. They are read out by silicon photo-multipliers, SiPMs, which are placed outside of the LHCb acceptance.

2.7.3 Particle identification

The particle identification systems, RICHs, Calorimeters and Muon stations, will be upgraded as well.

In the RICH1 the aerogel layer will be removed, as the track multiplicity will be too high, and the optics will be modified in order to increase ring size. The Hybrid Photon detectors in RICH1 and RICH2 will be replaced by multianode photomultipliers (MaPMT).

In the Calorimeter system the PS and SPD will be removed, as the occupancy will be too high and the hardware trigger will not be used any more. In the HCAL and ECAL the gain of the PMT will be reduced to lower ageing effects. It will be compensated by increasing the gain in the electronics, the redesign is mainly driven by the 40 MHz readout. The new front-end boards will be discussed in more details in Section 7. Some of the inner modules of HCAL and ECAL will be exchanged as well.

In the Muon system the first muon station, M1, will be removed. This is planned for the same reasons as the PS and SPD removal. Additional shielding will be installed in front of the second station, M2. The readout of the muon stations will be also changed.

2.8 Monte Carlo simulations

Simulated signal events are used in the analysis to evaluate efficiencies and to develop the offline selections. For the signal simulation, pp collisions are generated using Pythia 6.4 [72] and 8 [72] with a specific LHCb configuration [73]. Decays of hadronic particles are described by EvtGen [74]. The interaction of the generated particles with the detector and its response are implemented using the GEANT4 toolkit [75] as described in Ref. [76].

2.9 Important algorithms for data analyses of rare charm decays

In any data analyses the reconstruction and identification of the particles are important steps.

2.9.1 Track reconstruction

It comprises three main stages:

1. Pattern recognition of the detector signals, the so-called “hits”, produced typically by charged particles passing through the detecting layers.

2. Identification of the best possible trajectory, fitted with a Kalman filter [77]. This takes into account the energy loss and multiple scattering in the detector components. The tracks are fitted in one subsystem of the detector, for example VELO, and the algorithm looks for hits in the field of interest (FoI) in other subsystems. The size of FoI depends on $1/p_t$ of the passing particle.

3. Removal of duplicated tracks.

Several types of tracks are defined (Figure 2.24): VELO- and T-tracks are tracks reconstructed only in the VELO or T1, T2, T3. The VELO-tracks are used as a starting point for the track fitting and the T-tracks are used to "seed" the reconstruction of "long tracks". Long tracks are the tracks that are reconstructed in the whole detector with hits both in the VELO and in the T-stations. Such tracks have the most precise momentum measurements. It is also possible to define upstream tracks, reconstructed by hits only in the VELO and TT stations (usually they correspond to low momentum particles that are bent out of the detector acceptance by magnetic field), and downstream tracks that are reconstructed only by TT and T1, T2, T3 stations. Such tracks correspond to charged decays products of long lived particles, such as K_S^0 mesons.

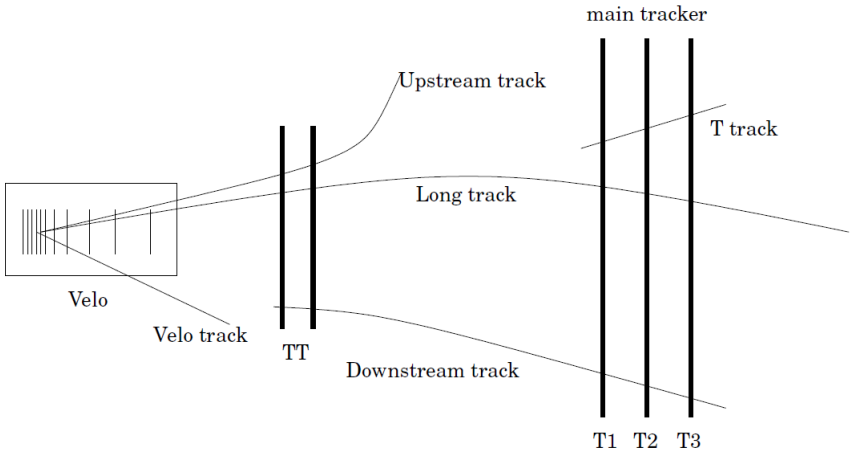


Figure 2.24: Schematic view of the track types in the LHCb tracking system.

2.9.2 Particle identification

Particle identification (PID) algorithms combine information coming from the RICH's, calorimeters and muon stations. In case of LHCb the identification of hadrons the PS, SPD and HCAL are used and the separation between π and K is done by the RICH's. The e and γ detection the ECAL+PS+SPD are involved. The muon stations plays the most important role in the identification of muons.

For each track the available PID information is combined into a log-likelihood difference,

$$\Delta LL_{h/tr} = \ln L(h) - \ln L(tr) = \ln\left(\frac{L(h)}{L(tr)}\right) \tag{2.1}$$

expressing the difference in log-likelihood between track tr and a given particle hypothesis h (e, γ, K, π, p and μ). The global likelihood hypothesis for each particle type is formed multiplying the different subdetector contributions:

$$\begin{aligned} L(K) &= L^{\mathrm{RICH}}(K) \cdot L^{\mathrm{CALO}}(\text{non } e) \cdot L^{\mathrm{MUON}}(\text{non } \mu), \\ L(\pi) &= L^{\mathrm{RICH}}(\pi) \cdot L^{\mathrm{CALO}}(\text{non } e) \cdot L^{\mathrm{MUON}}(\text{non } \mu), \\ L(\mu) &= L^{\mathrm{RICH}}(\mu) \cdot L^{\mathrm{CALO}}(\text{non } e) \cdot L^{\mathrm{MUON}}(\mu) \end{aligned} \tag{2.2}$$

A more detailed description of K, π and μ identification is presented in this section as these types of particles are mainly used in the analyses explored in this thesis.

Hadron identification

For the hadron identification the RICH system is mostly used. The reconstructed Cherenkov rings are compared to the pattern expected under a given particle hypothesis. The $PIDK$ variable used in analyses is a combined delta-log-likelihood for the kaon hypothesis with respect to the pion:

$$PIDK = \Delta LL_{K/\pi} = \ln\left(\frac{L(K)}{L(\pi)}\right), \tag{2.3}$$

where $L(K)$ and $L(\pi)$ are taken from Equation 2.2.

As shown in the Figure 2.25 the performance of separation π and K depends on momentum of the particles. It also depends on the pseudorapidity of the track. The average overall efficiency for the $\Delta LL_{K/\pi} > 0$ is $\sim 95\%$ with pion misidentification efficiency at the level of $\sim 10\%$.

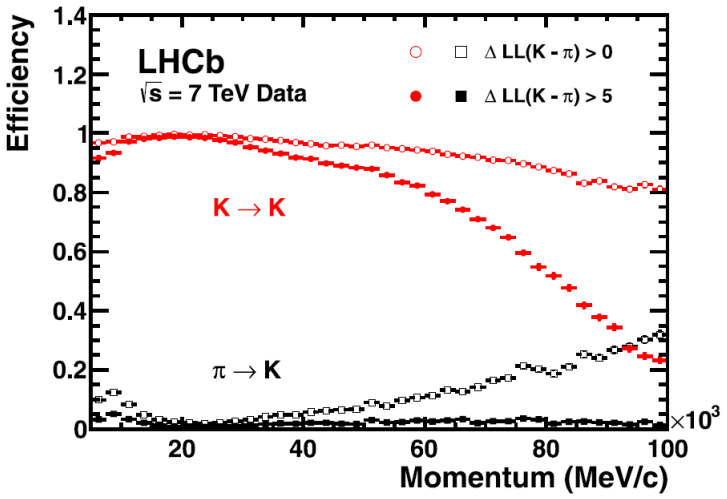


Figure 2.25: Kaon identification efficiency and pion misidentification rate measured on data as a function of track momentum. Two different $\Delta LL_{K/\pi}$ requirements have been imposed on the samples, resulting in the open and filled marker distributions, respectively.

Muon identification

To distinguish muons from other particles the muon station information is used. This is matched with the information about reconstructed tracks in the tracking system and combined to form a delta-log-likelihood. Hits in the muon stations are searched for in some Field of Interest (FoI) around the track extrapolation.

The following four variables are used for muon identification in offline analysis:

- *isMuon*. This is binary variable that is assigned a true value if extrapolated tracks have at least one hit in the FoI in a number of stations, depending on the particle momentum. Requirements on the number of stations are presented in Table 2.1

Momentum range	Muon stations
3 GeV/c < p < 6 GeV/c	M2+M3
6 GeV/c < p < 10 GeV/c	M2+M3+(M4 or M5)
p > 10 GeV/c	M2+M3+M4+M5

Table 2.1: Muon stations (from 1 to 5) required to trigger the muon decision as a function of momentum range.

- *nShared* is defined as the number of additional tracks in the event which are selected by the $IsMuon$ requirement due to at least one shared hit with the current track. This was designed to reduce the fraction of misidentification due to nearby muons while keeping a high efficiency for true muons.

- $PIDmu$ is the combined delta-log-likelihood for the muon hypothesis with respect to the pion, $PIDmu = \Delta LL_{(\mu/\pi)} = \ln(L(\mu)/L(\pi))$. The likelihoods $L(\mu)$ are computed as the cumulative probability distributions of the average squared distance significance, D^2:

$$D^2 = \frac{1}{N}\sum_{i=0}^{N-1}\left\{\left(\frac{x_{closest,i} - x_{track}}{pad_x}\right)^2 + \left(\frac{y_{closest,i} - y_{track}}{pad_y}\right)^2\right\}, \qquad (2.4)$$

  where the index i runs over the fired stations, $(x_{closest,i}, y_{closest,i})$ are the coordinates of the closest hit to the track extrapolation for each station and $pad_{x,y}$ correspond to one half of the pad sizes in the x, y directions.

- $ProbNNmu$ is a multivariate identification variable, i.e. an artificial neural network is used. $ProbNNmu$ is obtained from a combination of muon ΔLL's, $isMuon$, $nShared$, $inMuonAcc$ [1] and $isMuonLoose$ [2] variables.

2.10 MC deficiencies and performance of the LHCb detector

A certain number of quantities are not described precisely enough by the MC simulations. The most relevant to our analyses are:

[1] Equals to 1 if the track is in Muon acceptance.

[2] The variable similar to $isMuon$ but is equal "true" if there is at least one hit in FoI in at least two stations among M2, M3, M4 for 3 GeV/c < p < 6 GeV/c tracks and at least one hit in one of three stations among M2, M3, M4, M5 for p > 6 GeV/c.

- The L0 trigger efficiencies are consequently determined using real data control samples. The "TISTOS" method is used: that efficiency of the TOS selection is calculated in TIS events. One selects TIS events, that at first order do not depend on the signal mode, and count the fraction of TOS events among them. The efficiency is then: $\epsilon = N_{(TIS\&TOS)}/N_{(TIS)}$. Performance of L0 trigger lines [78] as a function of transverse momentum p_t is shown in Figure 2.26.

- All identification variables ($PIDK$, $PIDmu$, $ProbNNmu$ etc.). In this case we account for the fact that efficiency depends mainly on kinematics of the particles. We use control samples $D^0 \to K^-\pi^+$, $J/\psi \to \mu^+\mu^-$, $\Lambda \to p\pi$ etc. to measure the efficiency as a function of momentum (p), pseudorapidity (η) and detector occupancy ($nTracks$).

  Performance on the data sample of muon identification and probability of misidentification $K \to \mu$, $\pi \to \mu$ and $p \to \mu$ are shown in Figure 2.27 [79]. The muon identification efficiency was also observed to be robust against the variation of detector occupancies and presents a weak dependence on momentum and transverse momentum. A total muon efficiency is at the level of 93%, the hadron misidentification probabilities are below 0.6%.

  Hadron identification was discussed already in Section 2.9.2.

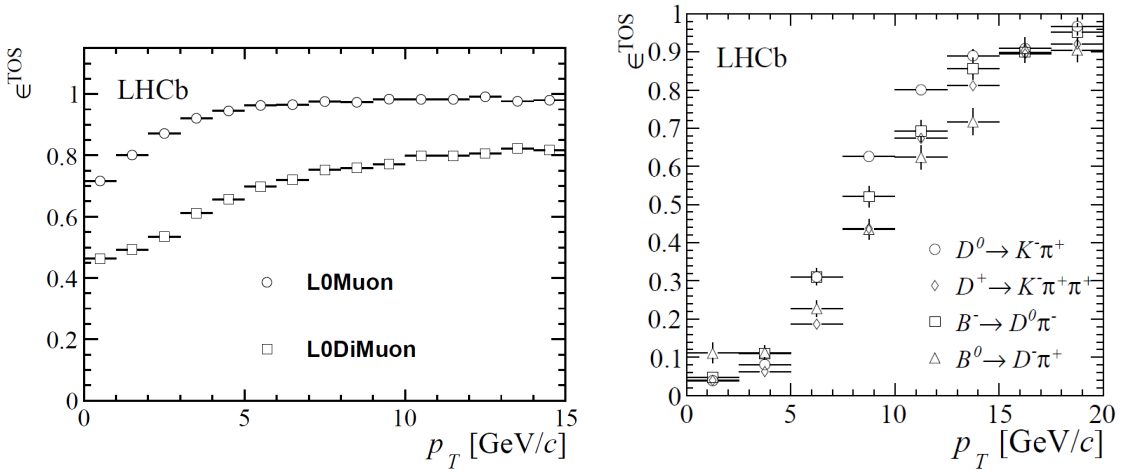
- Tracking. Detector performance on track reconstruction in VELO and T stations is presented in Figure 2.28 with a comparison to efficiencies in MC [80]. One can see that the reconstruction efficiency of LHCb is at the level of 98%, however, there is a discrepancy in data/MC efficiencies. This effect is corrected using the data control same, $J/\psi \to \mu^+\mu^-$. From Figure 2.28 it is seen that the tracking efficiency depends on kinematics, more precisely on momentum and pseudorapidity of decay products (p, η). Calibration sample provides the correction tables as function of p and η.

2.10.1 Performance of the detector for rare charm decays

Thanks to the excellent performance of the LHCb detector the study of rare charm decays is possible. High momentum resolution is possible due to the tracking stations and dipole magnet. This is illustrated in Figure 2.29 where one can see that $\frac{\delta p}{p}$ for charged final state tracks is $\sim 0.5\%$ for both high and low momentum. The resolution on transverse momentum $\frac{\delta p_t}{p_t}$ varies from $\sim 0.59\%$ for tracks with $p_t > 2000$ MeV/c to $\sim 0.85\%$ for tracks with $p_t < 2000$ MeV/c.

The resolution on the position of a secondary vertex coming from a B or D decay, $\frac{\delta z}{z}$, is about 4%. This performance is possible thanks to high spacial resolution of the VELO system.

Particle identification is performed by the Calorimeters, RICH's and Muon stations of LHCb. In Figure 2.30 one can see the good separation for kaon and pion hypothesis. The average overall efficiency for the $\Delta LL_{K/\pi} > 0$ is $\sim 95\%$ with a pion misidentification efficiency at the level of $\sim 1\%$.



(a) `L0Muon` (b) `L0Hadron`

Figure 2.26: The efficiency ϵ^{TOS} [78] of `L0Muon` for $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$ and `L0Hadron` for $B^0 \to D^-\pi^+$, $B^- \to D^0\pi^-$, $D^0 \to K^-\pi^+$ and $D^+ \to K^-\pi^+\pi^+$ are shown as a function of p_t.

Muon identification, which is very important for rare charm decays, is illustrated in Figure 2.31, where a muon identification variable is plotted. The red component represents the distribution for pion hypothesis and blue that for the muon hypothesis. Agreed separation of these two curves is observed. The efficiency of the cut $ProbNNmu > 0.3$ on pions is about 0.5%, while on muons it is $\sim 93\%$.

The invariant mass resolution for $D^0 \to K^-\pi^+\mu^+\mu^-$ candidates is about 7 MeV/c^2 as shown in Figure 2.32, where a fit with a double crystal ball function on the D^0 mass is presented.

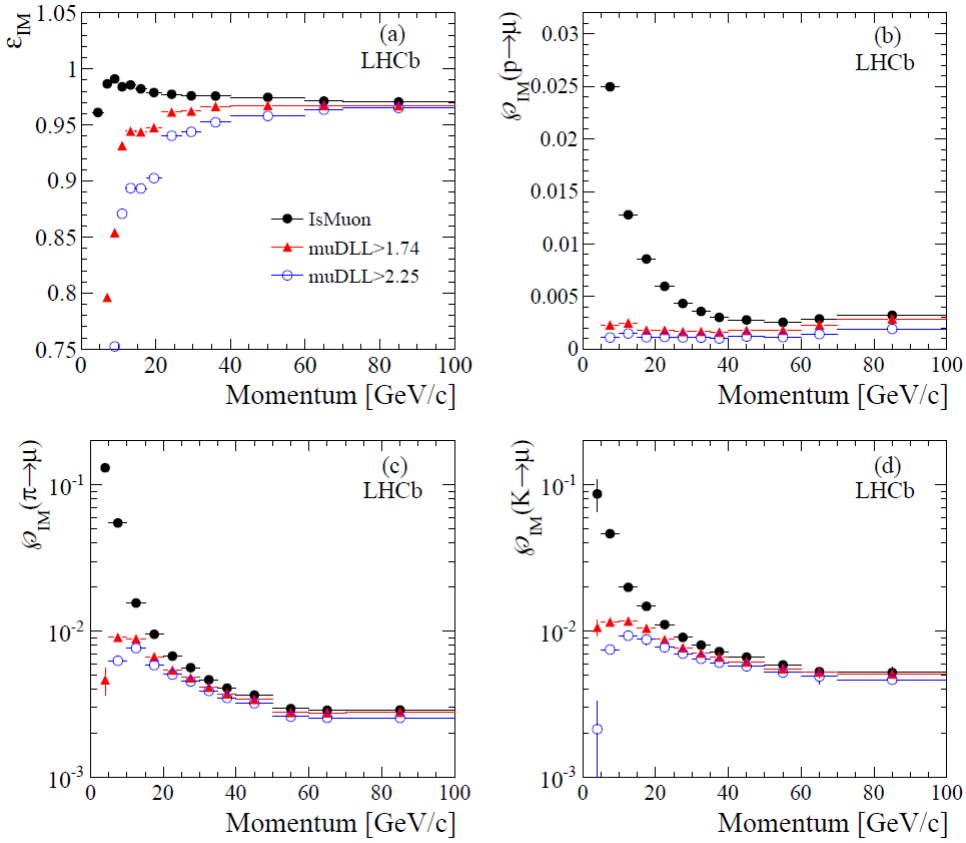


Figure 2.27: Muon efficiency [79] (a) and misidentification probabilities for protons (b), pions (c) and kaons (d) as a function of the particle momentum for the *IsMuon* requirement alone (black solid circles) and with the additional cuts $PIDmu > 1.74$ (red triangles) and $PIDmu > 2.25$ (blue open circles).

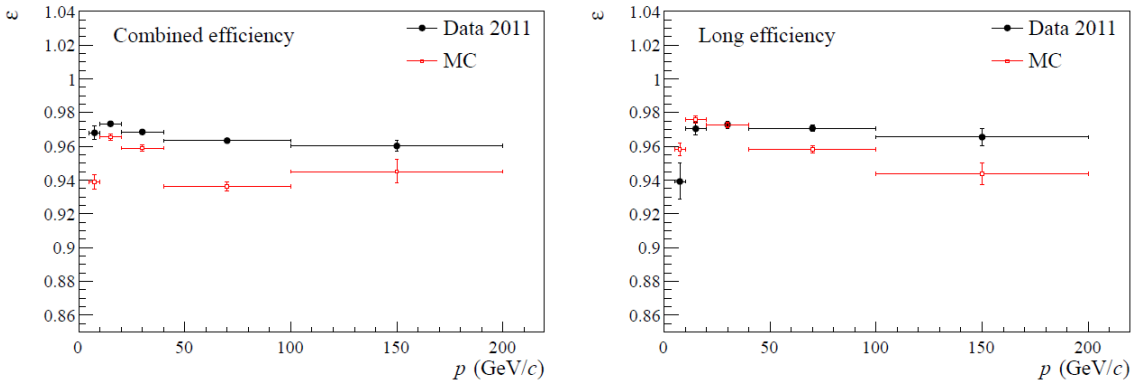


Figure 2.28: Tracking efficiency for the 2011 data and weighted MC simulation for the VELO method (left) and T station method (right) as a function of the momentum, p. The error bars indicate the statistical uncertainty [80].

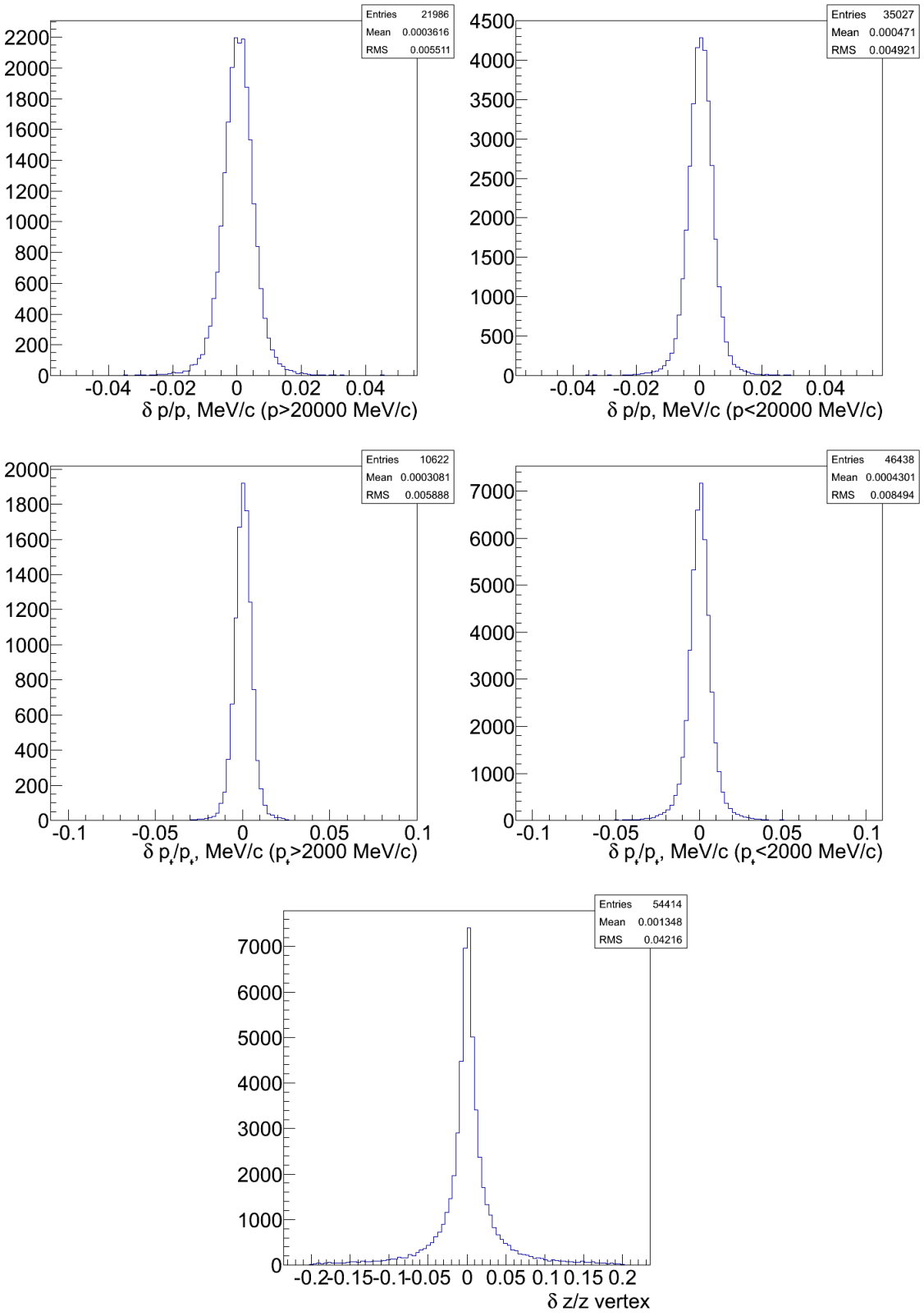


Figure 2.29: Resolution on p, p_t and position of secondary vertex for rare charm decays.

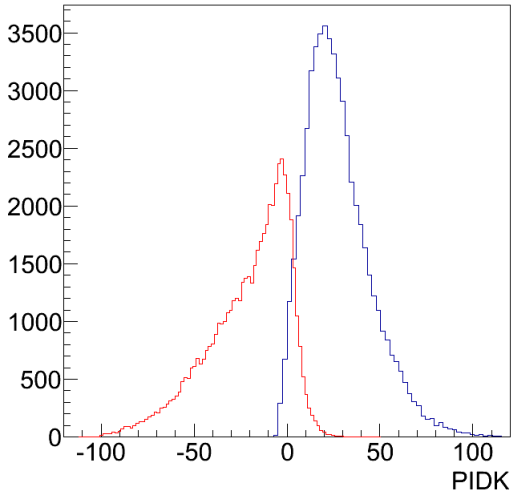


Figure 2.30: Hadron identification, delta-log-likelihood with respect to the pion. Red curve corresponds to pion hypothesis ($\Delta LL_{\pi/\pi}$) and blue curve corresponds to kaon hypothesis ($\Delta LL_{K/\pi}$).

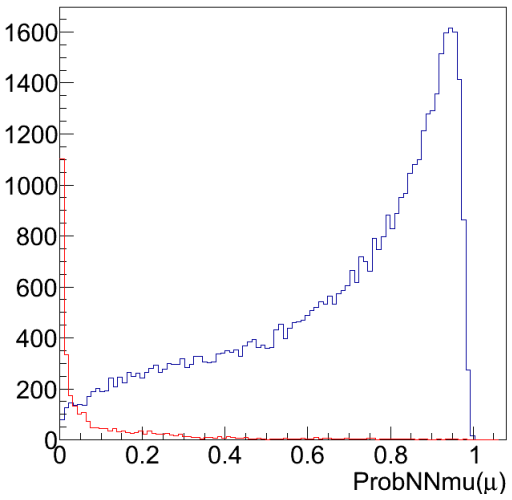


Figure 2.31: Muon identification variable distribution in blue for muon hypothesis and in red for pion hypothesis.

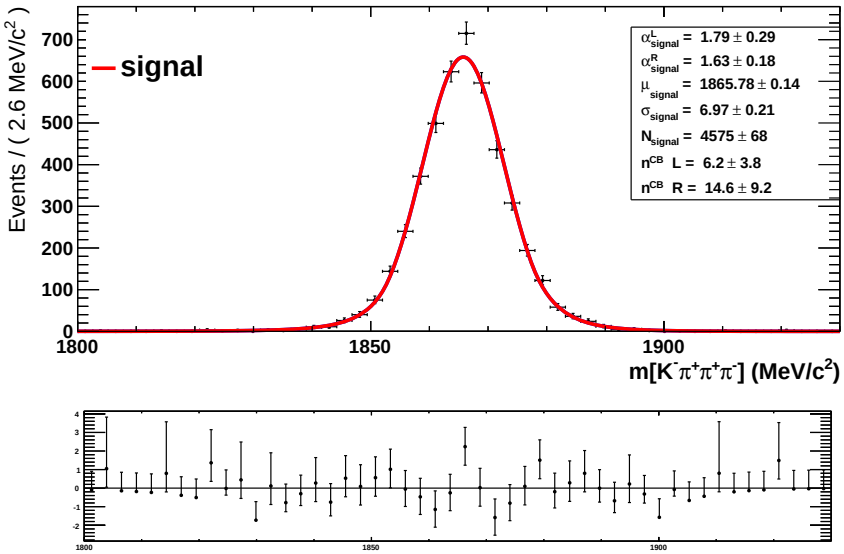


Figure 2.32: $D^0 \to K^+\pi^-\mu^+\mu^-$ invariant mass for illustration of mass resolution.

2.11 Variables used in LHCb data analyses

The discriminating variables used in data analyses are based on kinematic and topological characteristics of the decays, such as displaced vertexes, quality and of final state particles tracks, transverse and total momenta (Figure 2.33). The variables most commonly used in rare charm decay trigger lines are listed in Table 2.2.

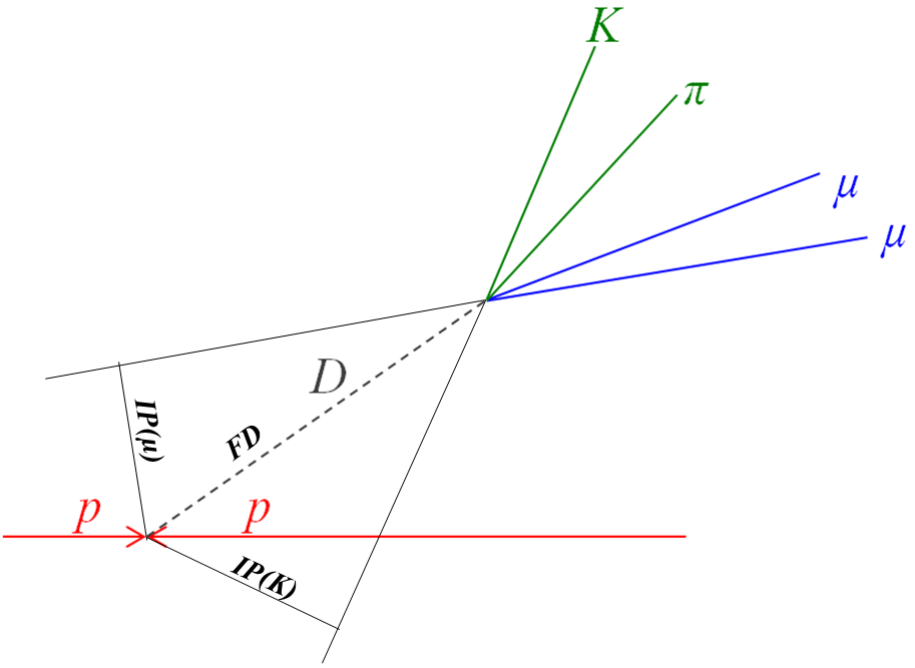


Figure 2.33: Schematic view of D meson decay at LHCb.

Variable	Meaning
p_t	Transverse momentum (of D meson or final state particles)
p	Total momentum (of D meson or final state particles)
Track χ^2	Quality of the track reconstruction
IP	Impact parameter, the shortest distance between a track and the primary vertex
IP χ^2	The χ^2 of the impact parameter of the particle with respect to a vertex
FD	Flight Distance, distance between primary and secondary vertexes
FD χ^2	Squared Flying Distance significance, $(FD/\sigma FD)^2$, relative uncertainty on impact parameter flight distance
VTX χ^2	Quality of the vertex reconstruction
DOCA	Distance of Closest Approach of tracks between two tracks
MAXDOCA	Maximum DOCA among all the pairs of tracks that can be made out of the decay products
MINDOCA	Minimum DOCA among all the pairs of tracks that can be made out of the decay products
DIRA	Cosine of the angle between the D meson momentum and the line joining the primary vertex and the D decay vertex
PIDK	Hadron particle identification variable (see Section 2.9.2)
PIDmu, ProbNNmu	Muon particle identification variable (see Section 2.9.2)
isMuon, nShared	Muon particle identification variable (see Section 2.9.2)
nPV	Number of primary vertices
p_t^{asy}	the p_t asymmetry isolation variable defined as $\left(\frac{p_t(D)-(\sum \vec{p})_t}{p_t(D)+\sum \vec{p})_t)}\right)$, where $p_t(D)$ is the p_t of the D meson and $\sum \vec{p}_t$ is the transverse component of the sum of the momenta of all charged particles found within a cone around the candidate, excluding the signal tracks
τ	Lifetime
DTF χ^2	DecayTreeFitter χ^2, χ^2 of fitting the whole decay tree of the particle

Table 2.2: Typical variables, used in rare charm decays analyses at LHCb.

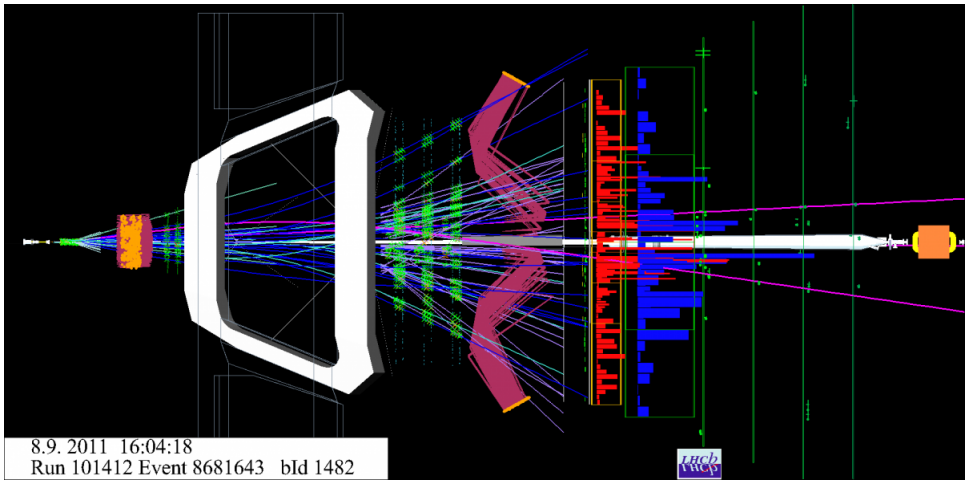


Figure 2.34: Illustration of an event in the LHCb detector. This event contains a $B_s \to \mu^+\mu^-$ candidate, which decay products are visible thanks to the two tracks isolated in the muon system.

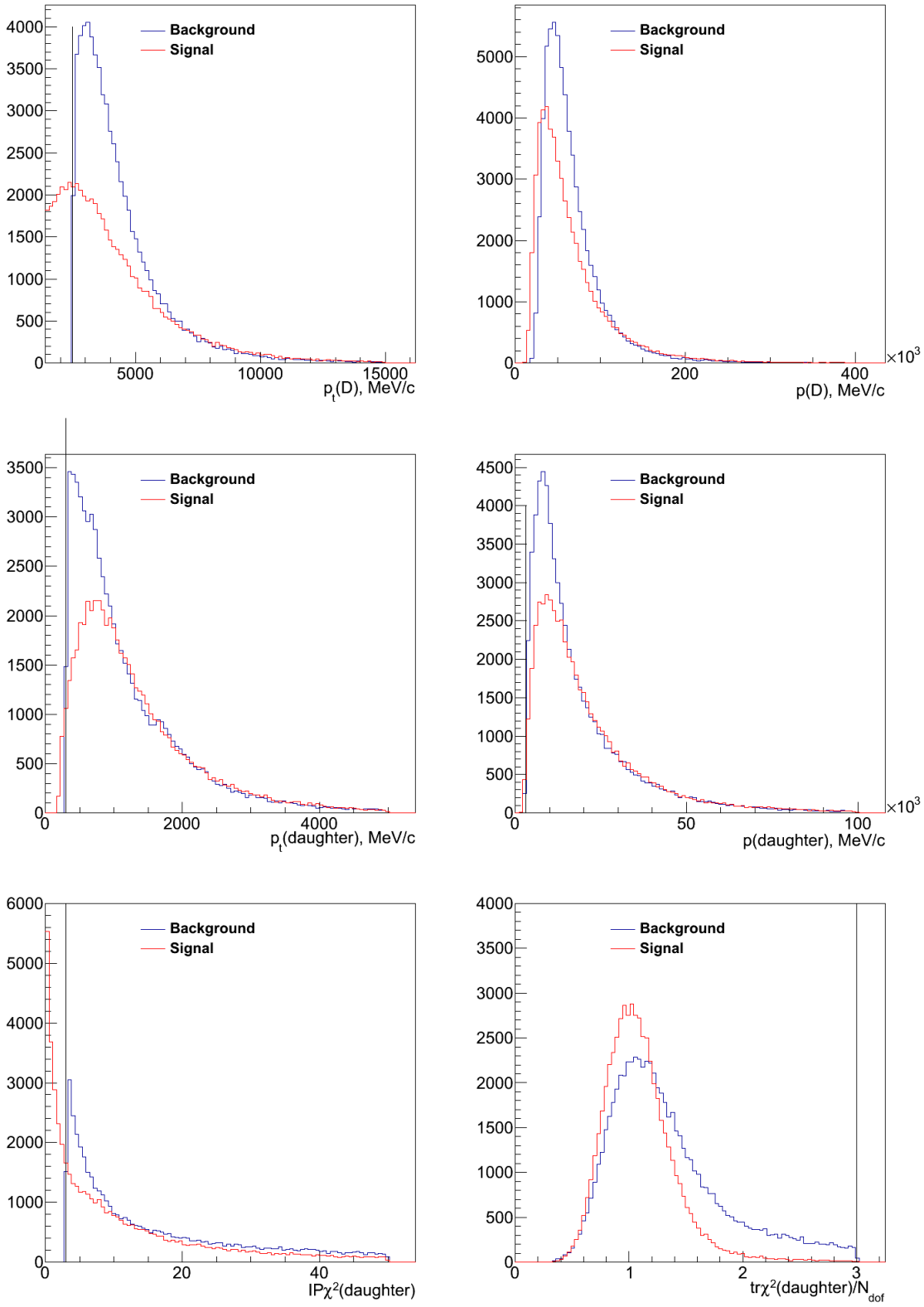


Figure 2.35: Variables used in rare charm decays analyses. Red curves correspond to signal sample (LHCb MC2012, loosened stripping), blue curves correspond to background sample (upper side-band ($m(D^0) > 1890$ MeV/c^2) in 2012 data sample after stripping preselection). Black vertical lines represent the cuts at stripping level ($D^0 \to hh\mu^+\mu^-$ mode in 2012).

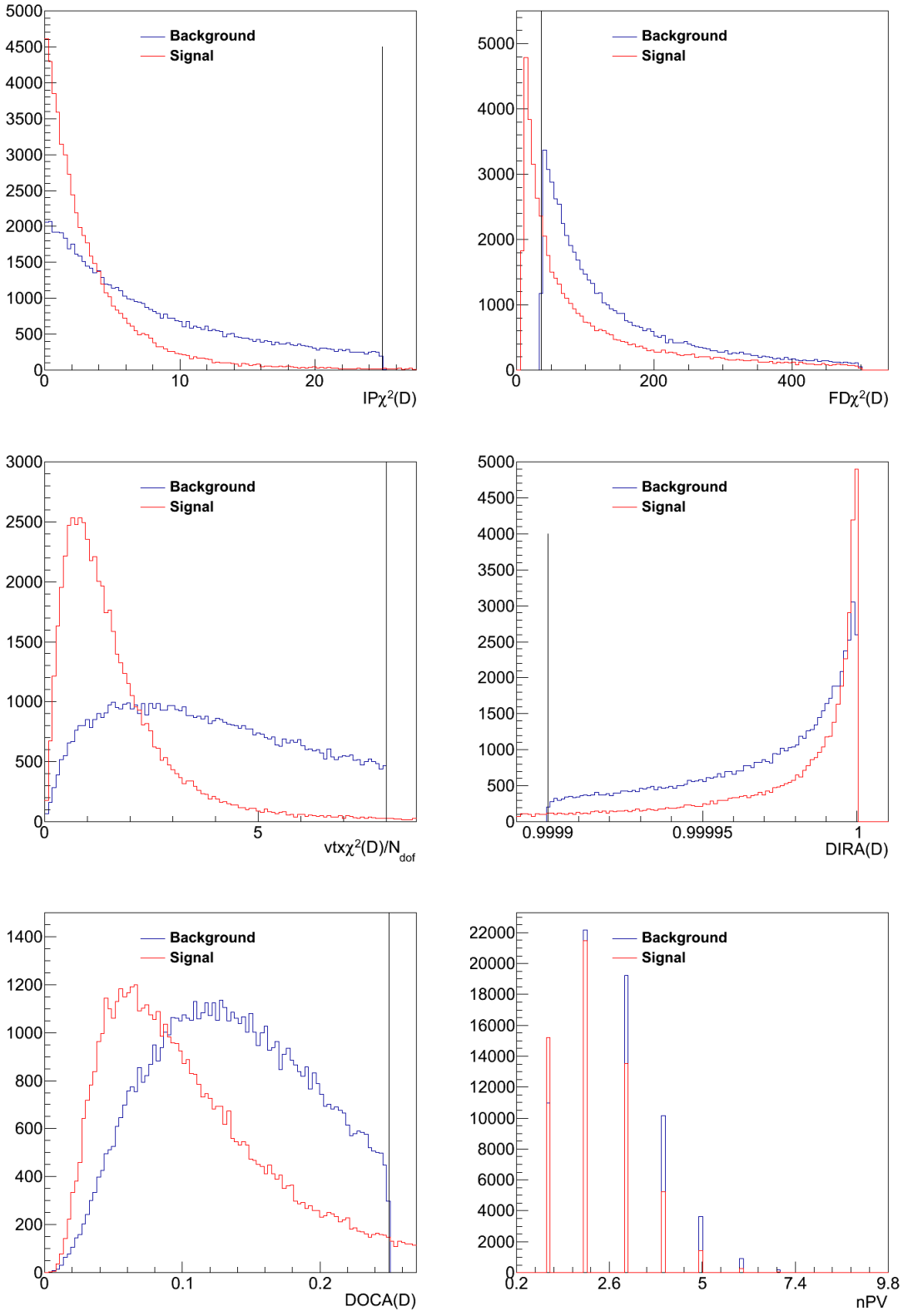


Figure 2.36: Variables used in rare charm decays analyses. Red curves correspond to signal sample (LHCb MC2012, loosened stripping), blue curves correspond to background sample (upper side-band ($m(D^0) > 1890$ MeV/c^2) in 2012 data sample after stripping preselection). Black vertical lines represent the cuts at stripping level ($D^0 \to hh\mu^+\mu^-$ mode in 2012).

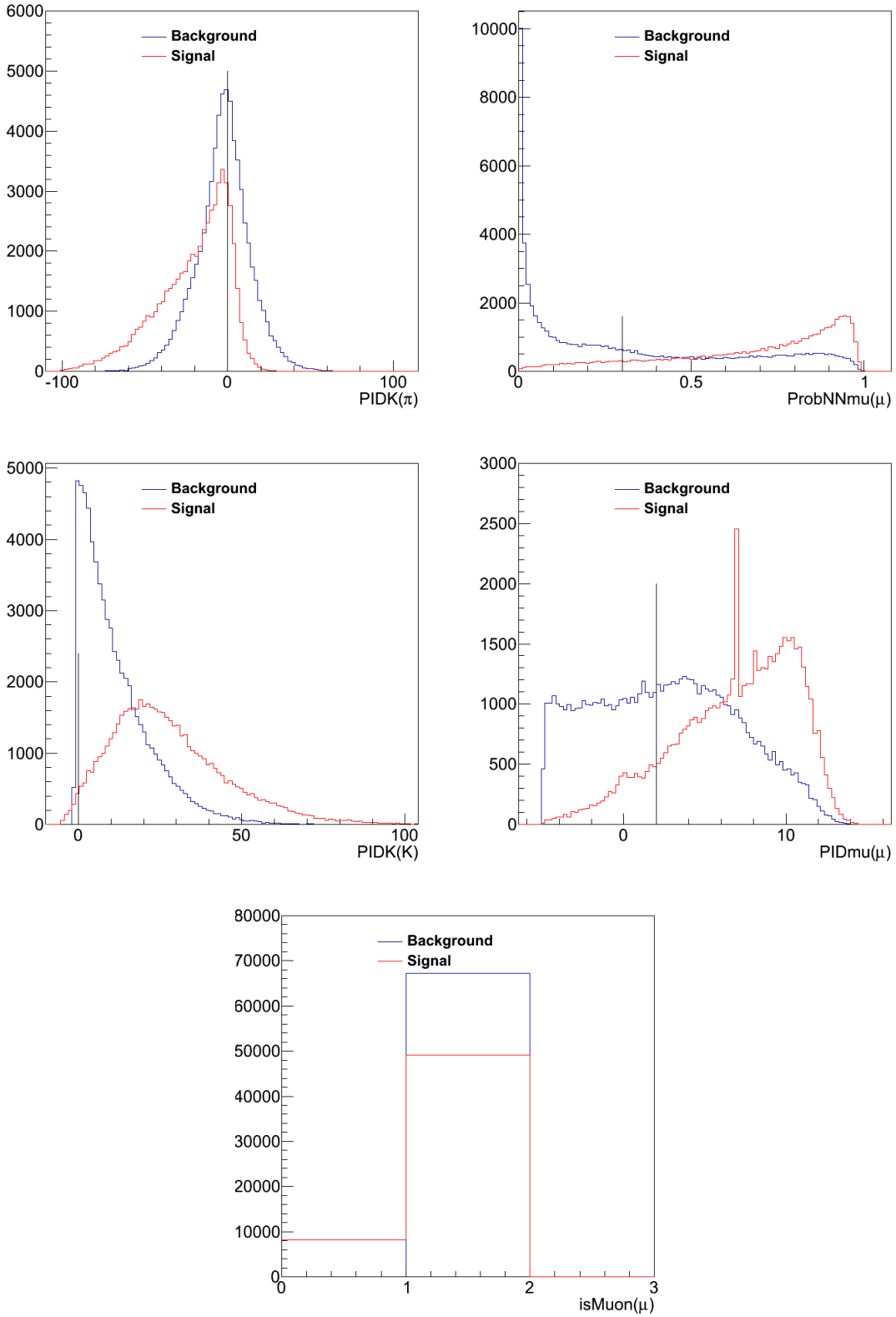


Figure 2.37: Variables used in rare charm decays analyses. Red curves correspond to signal sample (LHCb MC2012, loosened stripping), blue curves correspond to background sample (upper side-band ($m(D^0) > 1890$ MeV/c^2) in 2012 data sample after stripping preselection). Black vertical lines represent the cuts at stripping level ($D^0 \to hh\mu^+\mu^-$ mode in 2012).

Chapter 3

Search for $D^+_{(s)} \to \pi^+\mu^+\mu^-$ and $D^+_{(s)} \to \pi^-\mu^+\mu^+$ decays

The decays $D^+_{(s)} \to \pi^+\mu^+\mu^-$ proceed via FCNC, which are very suppressed in the SM (see Section 1.2). The most stringent limit on non-resonant contribution before our search for these decays was $\mathcal{B}(D^+ \to \pi^+\ \mu^+\mu^-\ ) < 3.9 \times 10^{-6}$ (90% CL) by the D0 collaboration [81].

Lepton number violating (LNV) processes such as $D^+ \to \pi^-\mu^+\mu^+$ are forbidden in the SM, because they may only occur through lepton mixing facilitated by a non-SM particle such as a Majorana neutrino [82]. The most stringent limits on the analysed decays at 90 CL are $\mathcal{B}(D^+ \to \pi^-\mu^+\mu^+) < 2\times10^{-6}$ and $\mathcal{B}(D^+_{(s)} \to \pi^-\mu^+\mu^+) < 1.4\times10^{-5}$ set by the BaBar collaboration [83].

In 2012, I participated in the search for non-resonant $D^+_{(s)} \to \pi^+\mu^+\mu^-$ and $D^+_{(s)} \to \pi^-\mu^+\mu^+$ decays using a proton-proton collision data sample corresponding to an integrated luminosity of 1.0 fb^{-1}, recorded at $\sqrt{s} = 7$ TeV by the LHCb experiment in 2011. The results were published in [13]. I was in charge of the evaluation of the systematic uncertainty related to the selection efficiency. In Section 3.1 of this chapter a brief description of this analysis and of its result are given. In Section 3.2 the work I carried out is presented in more detail.

3.1 Overview of the search for $D^+_{(s)} \to \pi^+\mu^+\mu^-$ and $D^+_{(s)} \to \pi^-\mu^+\mu^+$ with 1.0 fb^{-1} at $\sqrt{s} = 7$ TeV

The $D^+_{(s)} \to \pi^+\mu^+\mu^-$ are searched for in five regions of $m(\mu^+\mu^-)$ spectrum, in order to separate the regions sensitive to NP from those where resonant contributions dominate the short distance branching ratio. The definition of these regions can be found in Table 3.2.

The $D^+_{(s)} \to \pi^+\phi(\to \mu^+\mu^-)$ decay with $\mathcal{B} = (1.60 \pm 0.13) \times\ 10^{-6}$ [40] for D^+ decays and $\mathcal{B} = (4.5 \pm 0.4)\%$ [40] for D^+_s are used as a normalization mode. Therefore, in each

$m(\mu^+\mu^-)$ region, we set a limit on:

$$\mathcal{B}(D^+_{(s)} \to \pi^+\mu^+\mu^-) = \frac{N_{D^+_{(s)}\to\pi^+\mu^+\mu^-}}{N_{D^+_{(s)}\to\pi^+\phi(\to\mu^+\mu^-)}} \times \frac{\epsilon_{D^+_{(s)}\to\pi^+\phi(\to\mu^+\mu^-)}}{\epsilon_{D^+_{(s)}\to\pi^+\mu^+\mu^-}} \times \mathcal{B}(D^+_{(s)} \to \pi^+\phi(\to \mu^+\mu^-)) \tag{3.1}$$

The N are the fitted yields, ϵ are the efficiencies, including the geometrical acceptance of the detector, track reconstruction, muon identification, selection and trigger efficiencies.

The $D^+_{(s)} \to \pi^+\phi(\to \mu^+\mu^-)$ decays are also used as a control mode for the signal shape extraction, data/simulation comparison and other cross-checks.

Background rejection is a crucial element in any search for rare decays. In order to increase the combinatorial background rejection a Boosted Decision Tree (BDT) is used. The most dangerous peaking background is expected to be from the $D^+_{(s)} \to \pi^+\pi^+\pi^-$ with two pions misidentified as a muons. The branching ratios of these decays are at least five orders of magnitude above that of the signal: $\mathcal{B}(D^+ \to \pi^+\pi^+\pi^-) = (3.18 \pm 0.18) \times 10^{-3}$ and $\mathcal{B}(D^+_s \to \pi^+\pi^+\pi^-) = (1.10 \pm 0.06)\%$ [40]. These backgrounds are controlled with the help of muon identification criteria.

The first step of the selection is trigger. The following details the trigger chain, where candidates must pass the following criteria, applied for each sample and region,

- `L0Muon||L0DiMuon`[1]
- `Hlt1TrackMuon`[2]`||Hlt1DiMuonLowMass`[3] (high-$m(\mu^+\mu^-)$[4] opposite-sign decay only);
- `Hlt2CharmSemilepD2HMuMu`[5]`||Hlt2DiMuonDetached`[6] (high-$m(\mu^+\mu^-)$ opposite-sign decay only).

The next step is a stripping preselection. The corresponding cuts are shown in Table 3.1.

An offline selection is then applied to the remaining candidates. The BDT is trained on a simulated sample of $D^\pm \to \pi^\pm\mu^+\mu^-$ events as the signal, and 36.5 pb^{-1} of sideband data extracted from the sample collected in 2010 as the background. The inputs to this BDT are the following variables (definition of the variables can be found in Section 2.11):

- D candidate: χ^2 of the impact parameter (IPχ^2), end vertex (EVχ^2), and flight distance (FDχ^2).
- D candidate: Cosine of the direction angle (DIRA) and maximum distance of closest approach (MAXDOCA).
- D candidate and 3 tracks: p and p_t.

[1] Level-0 trigger selects muon and dimuon signal candidates

[2] Track selected by L0-Muon trigger.

[3] This line accepts dimuon vertexes with a mass above 1 GeV/c^2 where both tracks have a significant $IP\chi^2$ to the primary vertex which minimizes the impact parameter (IP) ($IP\chi^2 > 3$)

[4] $m(\mu^+\mu^-)$ is taken above 1 GeV/c^2

[5] Dedicated High Level trigger line for the 3 body rare charm decays

[6] Main detached dimuon trigger for low masses

Cut		**Value**
D decay vertex χ^2_{NDOF}	$<$	5
D maximum distance of closest approach χ^2	$<$	0.25 mm
D DIRA	$>$	0.9999
D impact parameter χ^2	$<$	30
$\pi\mu\mu$ invariant mass from PDG value	$<$	200 MeV
$\pi\mu\mu$ invariant mass	$>$	1763 MeV
π^+ p	$>$	3000 MeV
π^+ p_t	$>$	500 MeV
π^+ Track χ^2_{NDOF}	$<$	8
π^+ Minimum Impact Parameter χ^2	$>$	4
$\mu^\pm$ p	$>$	2000 MeV
$\mu^\pm$ p_t	$>$	300 MeV
$\mu^\pm$ Track χ^2_{NDOF}	$<$	8
$\mu^\pm$ Minimum Impact Parameter χ^2	$>$	6
$\mu\mu$ invariant mass	$>$	250 MeV

Table 3.1: Cuts applied in the stripping for selecting $D^+ \to \pi^+\mu^+\mu^-$.

- 3 tracks: χ^2 of the impact parameter (IPχ^2).
- Event: isolation variable defined as the p_t asymmetry ($\frac{\mathrm{p_t(mother)-p_t(cone)}}{\mathrm{p_t(mother)+p_t(cone)}}$) inside a cone of radius 1.5 ($p_t^{asy}(1.50)$). Due to this variable containing a delta-function at unity - when only the reconstructed mother is in the cone - the algorithms treat this circumstance differently from when there are other tracks from the event present in the cone.

The optimal offline selection and PID cuts are discovered by estimating the significance, defined as $\frac{\mathrm{s}}{\sqrt{\mathrm{s+b}}}$, where s and b are signal and background yields, in the non-resonant regions. The $D^+ \to \pi^+\phi(\to \mu^+\mu^-)$ mode, for which a large sample is available, is used as a proxy to the signal. Its branching fraction is rescaled to 10^{-9} to estimate s. This value is a plausible value of the branching ratio in this region. The background b is obtained by extrapolating the combinatorial background fitted in the ϕ region, with the help of a phase-space coefficient relating this region of the $m(\mu^+\mu^-)$ invariant mass spectrum with the non-resonant regions. The chosen cut is $(BDT > 0.9; PID > 0.1)$.

Several additional cuts are applied:

- The pion candidate must have $PIDmu < 0$[7].
- Hadronic backgrounds with kaons are further reduced requiring the pion candidate to have $PIDK < 0$[8].

[7]Delta-log likelihood between muon and pion hypothesis (Section 2.9.2)
[8]Delta-log likelihood between kaon and pion hypothesis (Section 2.9.2)

- The two muon candidates are forbidden to share hits in the muon stations by requiring each to satisfy $N_{shared} = 0$.
- No multiple candidates are permitted. The best candidate is chosen based on vertex χ^2 ($\sim 1/10^3$ events).
- The muons must have satisfied the $IsMuon$ requirement, described in Section 2.9.2.

The shapes and yields of the signal and background contributions are determined using a binned maximum likelihood fit to the invariant mass distributions of the $D^+_{(s)} \to \pi^+\mu^+\mu^-$ candidates. A simultaneous fit to several samples is performed. Firstly, the peak shape parameters are fitted assuming that they are the same in every $m(\mu^+\mu^-)$ region. This allows to determine them from the ϕ region. Secondly, the shape and yield of the $D^+ \to \pi^+\pi^+\pi^-$ peaking background decays present in the signal sample are constrained by fitting simultaneously a $D^+ \to \pi^+\pi^+\pi^-$ control sample based on the same selection as the signal, with loose muon identification criteria. The ratio between the $D^+ \to \pi^+\pi^+\pi^-$ yield in this sample and in the signal sample is assumed to be the same in all $m(\mu^+\mu^-)$ regions.

The $D^+_{(s)} \to \pi^+\mu^+\mu^-$ signal is described by the function:

$$f(x) \propto \exp\left(\frac{-(x-\mu)^2}{2\sigma^2 + (x-\mu)^2\alpha_{L,R}}\right), \tag{3.2}$$

which is a Gaussian-like peak of mean μ, width σ and where $\alpha_L(x < \mu)$ and $\alpha_R(x > \mu)$ parametrise the tails.

A second-order polynomial function is used to describe the PDF of all other combinatorial or partially reconstructed backgrounds that vary smoothly across the fit range. The coefficients of the polynomial are permitted to vary independently in each bin.

The invariant mass spectra together with the results are shown in Figure 3.1 and Table 3.2. No significant excess of candidates is seen in any of the NP search regions (low and high $m(\mu^+\mu^-)$).

Trigger conditions	Bin	$m(\mu^+\mu^-)$, MeV/c^2	D^+ yield	D^+_s yield
Triggers without $m(\mu\mu) > 1.0$ GeV/c^2	low-$m(\mu^+\mu^-)$	250-525	-3 ± 11	1 ± 6
	η	525-565	29 ± 7	22 ± 5
	ρ/ω	565-850	96 ± 15	87 ± 12
	ϕ	850-1250	2745 ± 67	3855 ± 86
All triggers	ϕ	850-1250	3683 ± 90	4857 ± 90
	high-$m(\mu^+\mu^-)$	1250-2000	16 ± 16	-17 ± 16

Table 3.2: Signal yields for the $D^+_{(s)} \to \pi^+\mu^+\mu^-$ fits. The ϕ region yields differ due to the different trigger conditions.

Upper limits on the $D^+_{(s)} \to \pi^+\mu^+\mu^-$ branching fractions are determined using the CL_s method [84] and observed distribution of the signal confidence level, CL_s, as a

function of the branching fraction in each $m(\mu^+\mu^-)$ bin. Total branching fractions are found using the same method and by considering the fraction of simulated signal candidates in each $m(\mu^+\mu^-)$ bin. The simulated signal assumes a phase-space model for the non-resonant decays. The upper limits at 90% and 95% CL and the p-values ($1 - \mathrm{CL}_b$) for the background-only hypothesis are shown in Table 3.3. Upper limits were also set on lepton flavour violation process, $\mathcal{B}(D^+ \to \pi^-\mu^+\mu^+) < 2.2(2.5) \times 10^{-8}$ and $\mathcal{B}(D^+_s \to \pi^-\mu^+\mu^+) < 1.2(1.4) \times 10^{-7}$ at the 90% (95%) CL limits. In these upper limits the systematic uncertainty discussed later in Section 3.2 is taken in account.

The obtained upper limits on $\mathcal{B}(D^+_{(s)} \to \pi^+\mu^+\mu^-)$ and $\mathcal{B}(D^+_{(s)} \to \pi^-\mu^+\mu^+)$ are almost 50 times better than the previous ones but they are still orders of magnitude above the SM predictions and BF enhancements from possible NP are not excluded yet. These limits are two orders of magnitude below than the predictions with MSSM$\not{R}$ [52, 55] but this model even more constrained from new upper limits on $D^0 \to \mu^+\mu^-$ decays [12].

Decay	Bin	90% [$\times 10^{-8}$]	95% [$\times 10^{-8}$]	p-value
$D^+ \to \pi^+\mu^+\mu^-$	low-$m(\mu^+\mu^-)$	2.0	2.5	0.74
	high-$m(\mu^+\mu^-)$	2.6	2.9	0.42
	Total	7.3	8.3	0.42
$D^+_s \to \pi^+\mu^+\mu^-$	low-$m(\mu^+\mu^-)$	6.9	7.7	0.78
	high-$m(\mu^+\mu^-)$	16.0	18.6	0.41
	Total	41.0	47.7	0.42

Table 3.3: Upper limits in each $m(\mu^+\mu^-)$ bin and total branching fractions at the 90% and 95% CL and p-values for the background-only hypothesis. The systematic uncertainty discussed later in Section 3.2 is taken in account in these upper limits.

3.2 Systematic uncertainties

Systematic uncertainties in the calculation of the signal branching fractions arise due to imperfect knowledge of the control mode branching fraction, the efficiency ratio and the yield ratio.

A systematic uncertainty of the order 10% accompanies the branching fraction of the control mode $D^+_{(s)} \to \pi^+\phi(\to \mu^+\mu^-)$ and is the dominant source of the systematic uncertainty on the branching fraction measurement.

A systematic uncertainty affecting the efficiency ratio is due to the geometrical acceptance of the detector, which depends on the angular distributions of the final state particles, and thus on the decay model. By default, signal decays are simulated with a phase-space model. A conservative 1% uncertainty is determined by recalculating the acceptance assuming a flat $m(\mu^+\mu^-)$ distribution. For that purpose, we wrote a reweighting algorithm applying weights across the $m(\pi^+\mu^-)^2 : m(\mu^+\mu^-)^2$ Dalitz plane to obtain a flat $m(\mu^+\mu^-)$ distribution in samples generated with the phase space decay model.

The uncertainties on the tracking and particle identification corrections also affect the efficiency ratio and involve statistical components due to the size of the data samples and systematic uncertainties inherent in the techniques employed to determine the

corrections. The corrections depend upon the choice of control sample, the selection and trigger requirements applied to this sample, and the precise definition of the probe tracks. We use the binning to weight the efficiency as a function of the momentum, pseudorapidity and multiplicity. The binning is varied to evaluate the uncertainty. The uncertainty in the choice of phase space model is accounted for by comparing the efficiency corrections in the extreme bins of the $m(\mu^+\mu^-)$ distributions. In total, the uncertainty due to particle reconstruction and identification is found to be 4.2% across all bins.

The trigger requirements imposed to select the signal are varied in order to test the imperfect simulation of the online reconstruction and 3% uncertainty is deduced.

Stripping and BDT efficiency

The quantities used by these two steps of the selection are described in Section 3.1. Most of them should be well described by the Monte Carlo simulation. A notable exception is the impact parameter of the final states tracks (IP, $IP\chi^2$). A way to account for that is to reprocess the MC samples used in this analysis, smearing the tracks' parameters in such a way that the distribution of their impact parameter matches that observed in real data [85]. The smearing process replaces (x, y), the coordinates of the point of each reconstructed track which is the closest to the beam line, by $(x + \delta x, y + \delta y)$, where δx and δy are drawn from two gaussians. The sigmas of these gaussians are chosen in order obtain the same distribution in data and MC for the tracks that come from the primary vertex: this distribution is a measure or the resolution on this impact parameter. Among the variables used in our selections, the most affected are those that combine several tracks: the D's $IP\chi^2$ and vertex χ^2, and the distance of closest approach between the D's tracks ($MAXDOCA$).

The amount of the smearing is tuned for better matching to our samples. The analysis can be repeated based on these samples. In particular, the discriminant variable provided by the BDT can be re-computed. The difference in efficiency between the standard and smeared samples provides the corresponding systematic uncertainty. Figures 3.2 and 3.4 show that the smearing improves the data/MC agreement.

The variation in the efficiency ratio with respect to the default simulation is done. No trend is seen as a function of $m(\mu^+\mu^-)$. For each mode, we computed the average variation between the smeared and unsmeared case, and combined it with its uncertainty. The largest value is found in the case of $D_s^+ \to \pi^-\mu^+\mu^+$: 4%. We quote this as the corresponding systematic uncertainty for all modes and regions. This should be conservative since a part of this value is probably due to statistics.

An additional cross-check is performed by comparing the distribution of the discriminating variable provided by the BDT ($BDTCat$) in data and MC. For that purpose, the $D^+ \to \pi^+\phi(\to \mu^+\mu^-)$ and $D_s^+ \to \pi^+\phi(\to \mu^+\mu^-)$ peaks observed in real data can be used.

The distributions with subtracted combinatorial background can be subsequently be used for data/MC comparisons. Figure 3.6 compares the subtracted data distributions of $BDTCat$ with those found in MC samples. We can see a clear discrepancy in the high $BDTCat$ region. It is due to one of the most important variables used by the

BDT: the isolation variable $p_t^{asy}(1.50)$. At the same time $p_t^{asy}(1.50)$ depends strongly upon the total number of tracks in the event (the higher the number of tracks, the lesser the likelihood for the D to be isolated). It is a well known fact that the Monte Carlo simulation underestimates this quantity. Forcing the MC to be closer to the real data by cutting out events with number of tracks per event $nTracks < 110$ improves the data/MC agreement concerning $p_t^{asy}(1.50)$. Simultaneously, the data/MC discrepancy in $BDTCat$ reduces to a low level (Figures 3.7). The relative data/MC difference in the efficiency of the cut used by the analysis ($BDTCat > 0.9$) is $\sim 2-3\%$. This is shown in Figure 3.5. The quality of data/MC agreement in the efficiency ratios should be at least of that level. It should not be affected by the discrepancy in number of tracks ($nTracks$) since this variable, at first order, does not vary as a function of $m(\mu^+\mu^-)$ (see Figure 3.3).

The sources of uncertainty discussed so far are given in Table 3.4. Final uncertainty on the efficiency ratio arises due to the finite size of the simulated samples. It is calculated separately in each $m(\mu^+\mu^-)$ bin. These contributions are included in the systematic uncertainties shown in Table 3.5.

Source	Uncertainty (%)
Geometric acceptance	1.0
Track reconstruction and particle identification	4.2
Stripping and BDT efficiency	4.0
Trigger efficiency	3.0
$\mathcal{B}(D^+_{(s)} \to \pi^+\phi(\to \mu^+\mu^-))$ uncertainty	8.1 (10.9)

Table 3.4: Relative systematic uncertainties averaged over all bins and decay modes for the control mode branching fraction and efficiency ratio. The number in parentheses refers to the D_s^+ decay.

Bin	$D^+_{(s)} \to \pi^+\mu^+\mu^-$ (%)
low-$m(\mu^+\mu^-)$	11.8 (16.9)
high-$m(\mu^+\mu^-)$	11.2 (15.5)

Table 3.5: Total systematic uncertainty in each $m(\mu^+\mu^-)$ and $m(\pi^-\mu^+)$ bin with the uncertainty on the control mode branching fraction, the efficiency ratio and the statistical uncertainty stemming from the size of the simulated samples added in quadrature. The numbers in parentheses refer to the D_s^+ decay

The systematic uncertainties affecting the yield ratio are taken into account when the branching fraction limits are calculated. The shapes of the signal peaks are assumed to be the same in all $m(\mu^+\mu^-)$ bins. A 10% variation of the width of the Gaussian-like function, seen in simulation, is taken into account for variation across the bins. In each bin, the shape of the $D^+_{(s)} \to \pi^+\pi^+\pi^-$ peaking background is taken from a simultaneous fit to a larger sample to which looser muon identification $PIDmu$ criteria is applied. As simulation shows the shape is altered by a $PIDmu$ requirement. A variation in the peaking background's fitted width equal to 20% is applied as a systematic uncertainty.

The pion-to-muon misidentification rate is assumed to be the same in all bins. Simulation suggests that a systematic variation of 20% in this quantity is conservative.

Contributions to the yield ratio final systematic uncertainty are found to increase the upper limit on the branching fraction by around 10%.

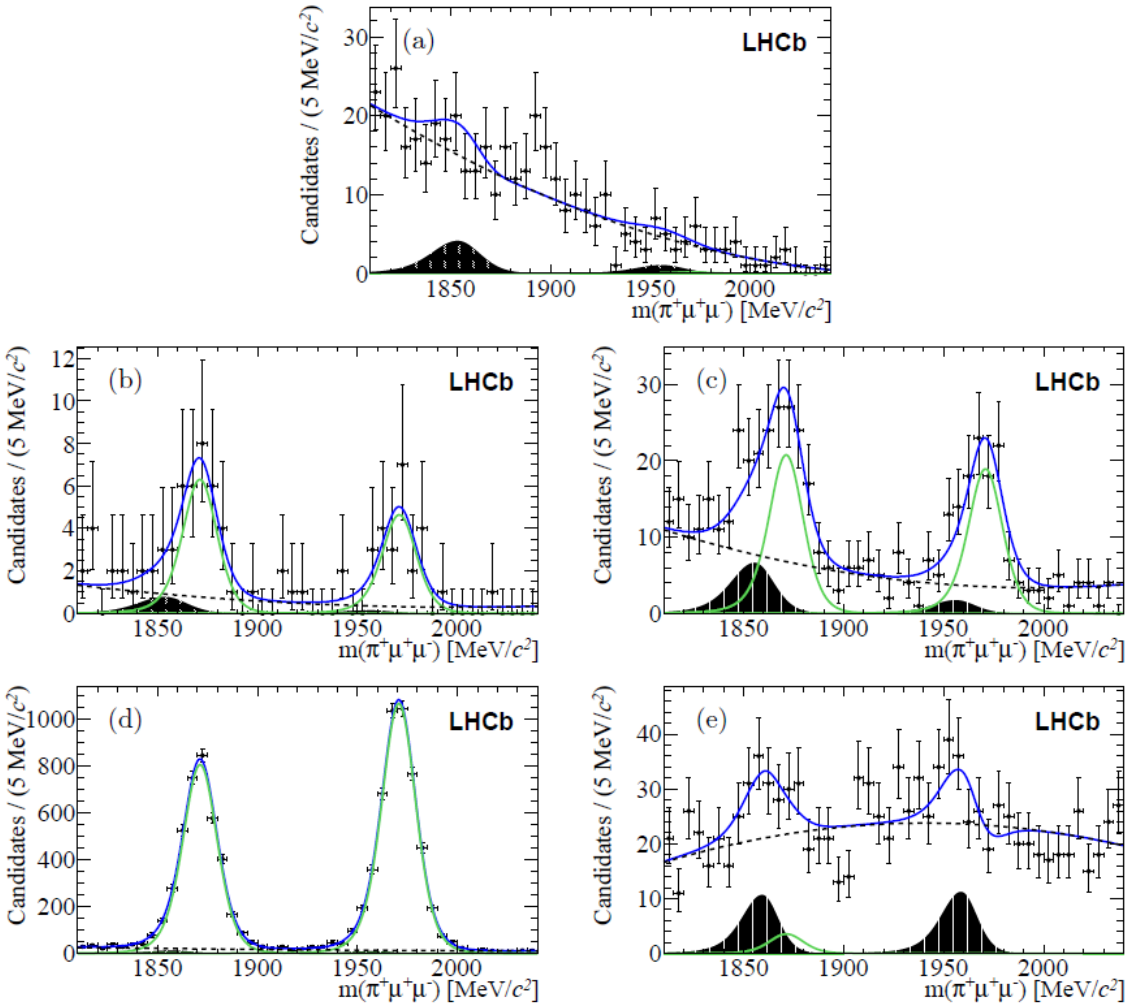


Figure 3.1: Invariant mass distributions for $D^+_{(s)} \to \pi^+\mu^+\mu^-$ candidates in the five $m(\mu^+\mu^-)$ bins. Shown are the (a) low-$m(\mu^+\mu^-)$, (b) η, (c) ρ/ω, (d) ϕ (including trigger lines with $m(\mu^+\mu^-) > 1.0$ GeV/c^2), and (e) high-$m(\mu^+\mu^-)$ regions. The data are shown as points (black) and the total PDF (dark blue line) is overlaid. The components of the fit are also shown: the signal (light green line), the peaking background (solid area) and the non-peaking background (dashed line).

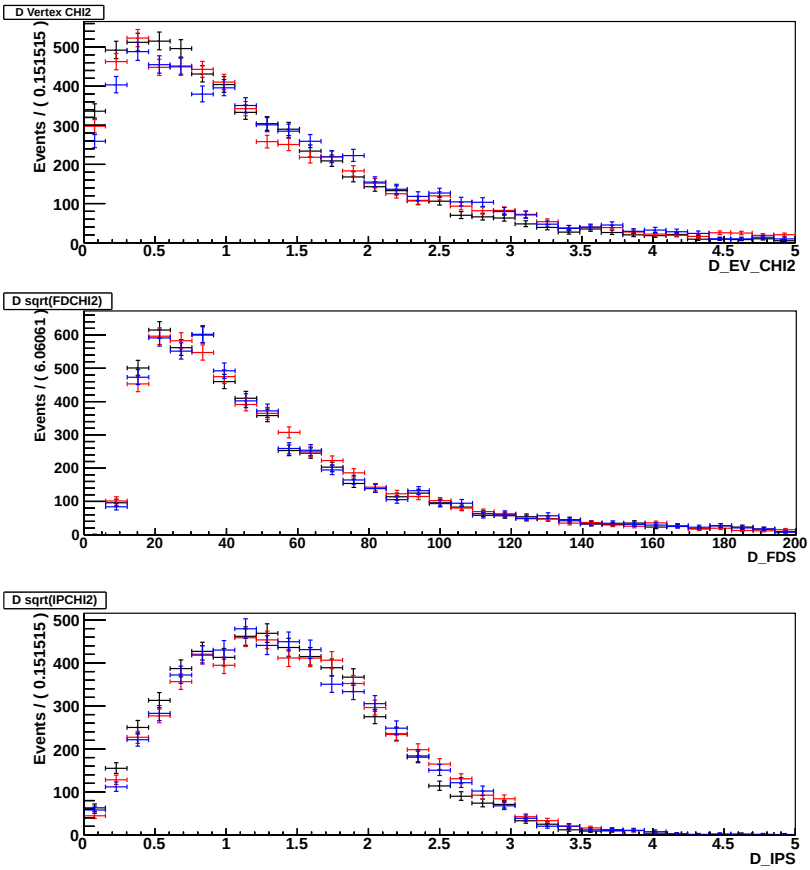


Figure 3.2: Distributions of the D vertex χ^2 (top), flying distance significance (middle) and impact parameter significance with respect to the primary vertex. The subtracted $D^+ \to \pi^+\phi(\to \mu^+\mu^-)$ peak found in real data (red) is compared to the default MC (black) and the smeared MC (blue).

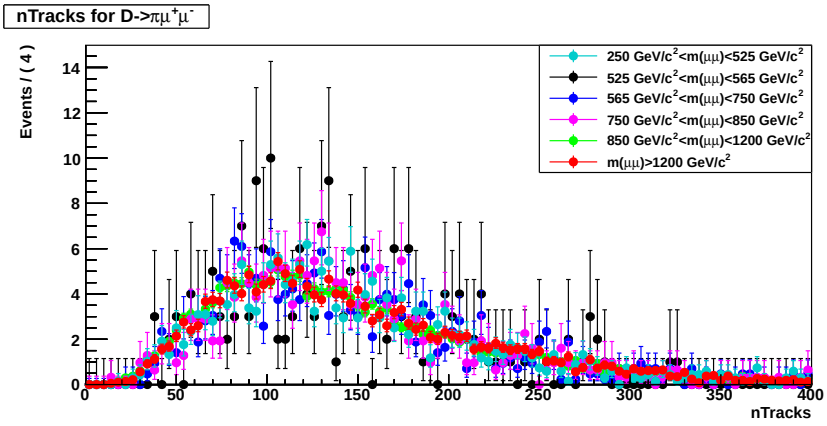
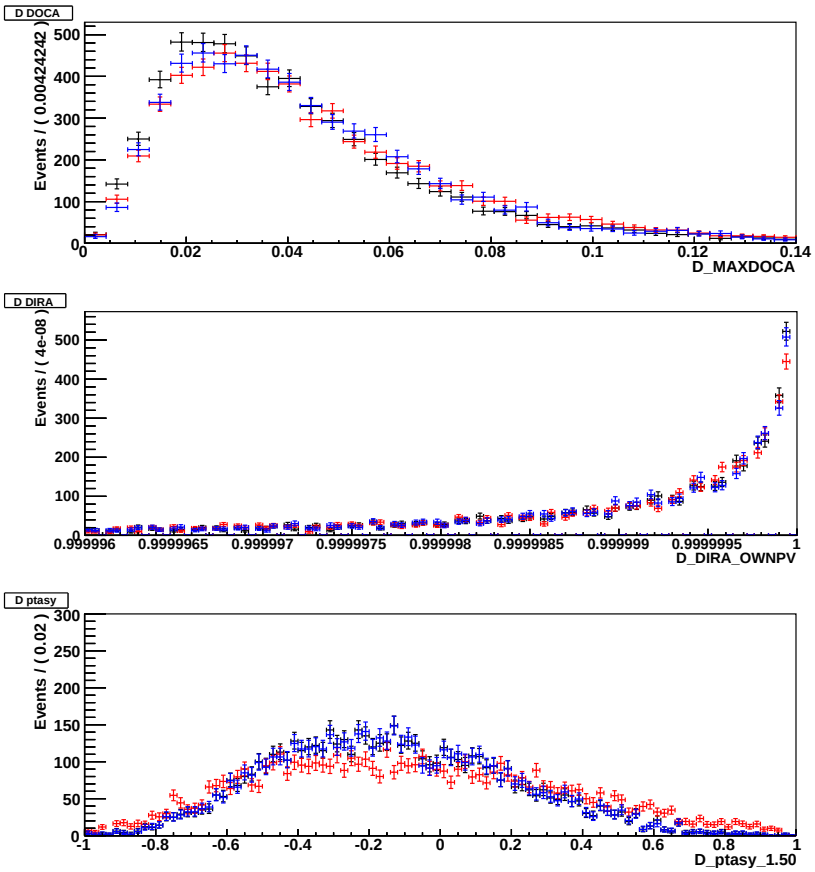
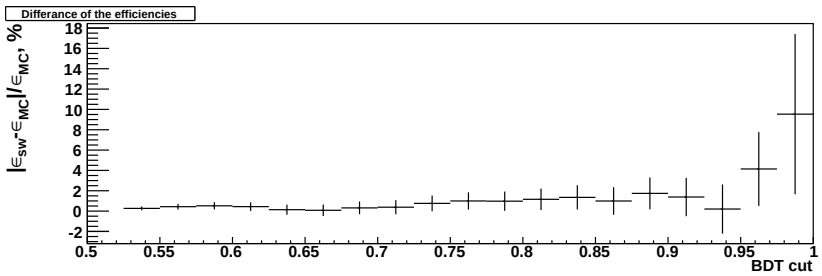


Figure 3.3: Number of charged tracks ($nTracks$) distribution in the data in different regions of invariant mass squared of the dimuons.



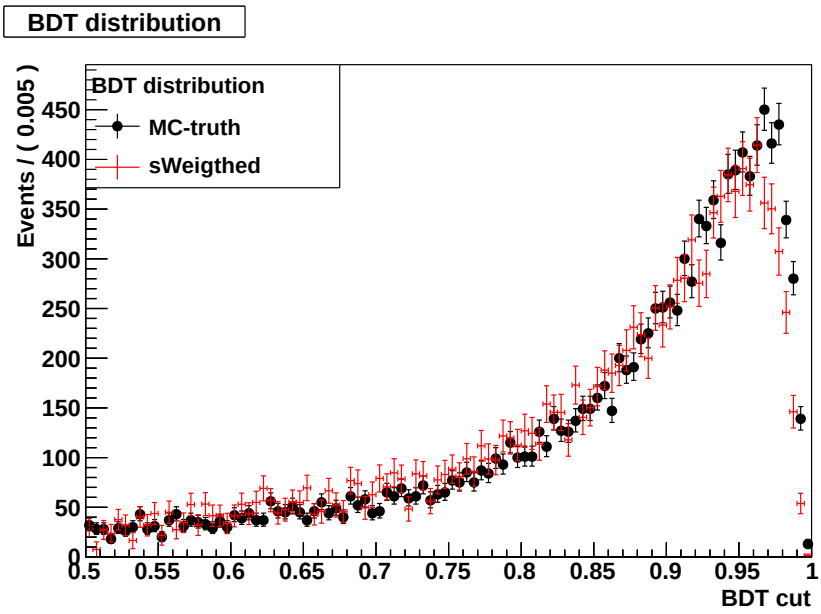
(a)

Figure 3.4: Distributions of the D maximum Distance of Closest Approach (DOCA) (top), Cosine of the angle between the D meson momentum and reconstructed line of flight (DIRA) (middle) and of the isolation variable $p_t^{asy}(1.50)$ (bottom). The subtracted $D^+ \to \pi^+\phi(\to \mu^+\mu^-)$ peak found in real data (red) is compared to the default MC (black) and the smeared MC (blue).



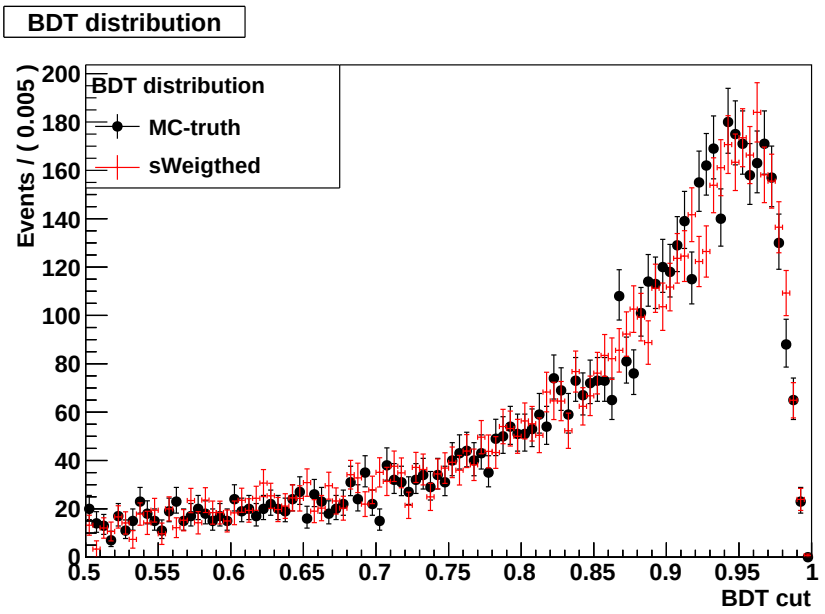
(a)

Figure 3.5: Relative data/MC difference in the efficiency as a function of the value of the cut on $BDTCat$. Here this is shown for $D^+ \to \pi^+\phi(\to \mu^+\mu^-)$. The cut used in the analysis is $BDTCat \geq 0.9$.



(a)

Figure 3.6: Data/MC comparison of the *BDTCat* distribution: After background subtraction using the sPlot technique, the distribution of $D^+ \to \pi^+\phi(\to \mu^+\mu^-)$ events in the data (red), is superimposed to the distribution found in Monte Carlo for such decays (black).



(a)

Figure 3.7: Data/MC comparison of the *BDTCat* distribution after applied requirement on number of tracks in per event $nTracks > 110$. The background subtraction allows to superimpose the distribution of $D^+ \to \pi^+\phi(\to \mu^+\mu^-)$ decays from the data (red) to the distribution found in Monte Carlo for such decays (black).

Chapter 4

Data preprocessing for $D^+ \to h^+h^{(\prime)-}\mu^+\mu^-$ decays analyses

4.1 Reoptimization of HLT2 lines for 2012 data taking period

The L0 output rate is at the level of ~ 1 MHz, but only about 5 kHz could be recorded for offline analysis in 2012. This means that the data must be reduced by the High Level Trigger (HLT) at least by factor 200. This reduction is achieved in two stages, HLT1 and HLT2. The LHCb trigger system is described in more detail in Section 2.5. This section concentrates on HLT2 lines designed to select the $D^+ \to \pi^+\mu^+\mu^-$, $D_s^+ \to \pi^+\mu^+\mu^-$, $D^+ \to K^+\mu^+\mu^-$, $D_s^+ \to K^+\mu^+\mu^-$, $D^+ \to \pi^-\mu^+\mu^+$, $D_s^+ \to \pi^-\mu^+\mu^+$, $D^+ \to K^-\mu^+\mu^+$, $D_s^+ \to K^-\mu^+\mu^+$ and $D^0 \to \pi^+\pi^-\mu^+\mu^-$, $D^0 \to K^+K^-\mu^+\mu^-$, $D^0 \to K^\pm\pi^\mp\mu^+\mu^-$ decays.

4.1.1 HLT2 rare charm decays lines at LHCb

The HLT2 is based on a full event reconstruction, using algorithms that are only a slightly simplified version of those run offline. With resolutions and efficiencies approaching these of the offline, almost all the variables usually involved in a sophisticated selection can already be used at this stage. One can therefore offer a dedicated *line* to every (group of) channel(s) of interest. In practice, the HLT2 comprises dozens of lines, performing more or less inclusive selections, that can all benefit from an independent optimisation.

The challenge for the HLT is that there is a very large number of modes that are to be selected for analysis. This means that besides the signal efficiency and purity the requirements on online trigger are:

- *Retention rate* as the bandwidth is limited;
- *CPU timing* as the time to process an event online is limited.

These aspects are challenging in case of multibody decays such as $D \to h(h')\mu\mu$, which suffer from a large combinatorial background and cannot rely on the signature of high p_t tracks as much as other modes.

The CPU requirement is particularly difficult to satisfy: three or four tracks in the final state mean a lot of combinations to process. In order to avoid hard cuts at track level (i.e. before the $(h)h\mu\mu \to D^0$ combination), affecting all the final state tracks and causing an important loss in efficiency, a two-stage strategy is adopted:

- at the first stage only events with pair of muons are preselected. Thanks to the smaller number of muon candidates with respect to hadron candidates a large part of the background is removed directly at this stage.
- at the second stage, hadrons are added and D candidates are reconstructed. At this stage again selections on final state particles and D meson are applied.

The first stage is common for 3-body, $D \to h\mu\mu$, and 4-body, $D \to hh'\mu\mu$, decay lines.

Historically, this approach originates from $D \to hhh$, where pair of hadrons is selected at the first stage.

4.1.2 Trigger re-optimization

In 2011, the lines devoted to the $D \to h(h')\mu\mu$ modes were designed before the analyses actually started. Therefore, the optimal way to select such modes was not known. For this reason, cuts from existing line $D \to hhh$, closest topologically and kinematically, were adopted as the starting point. Having two muons in the final state means an easier background rejection than in the case of a purely hadronic state since muon chamber identification is available whereas the information from RICH necessary for $K - \pi$ separation is not. Thus, it was possible to loosen cuts on most of the variables, like flying distance χ^2, impact parameter χ^2, vertex χ^2 (definition of the variables can be found in Section 2.11). Loosening the cuts related to the D lifetime was particularly important since these lines were also used to select $D_s^+ \to h\mu\mu$, and the lifetime of the D_s^+ is almost twice smaller than that of the D^+. The cuts were tightened in the case of the four-body lines because of higher background due to the higher multiplicity mode. The cuts obtained this way are shown in the first column of Table 4.1.

During 2012 data taking, the conditions for HLT2 were changed. Thanks to the "deferred HLT2" (see Figure 2.20 in Section 2.5.2) and the increase of the output bandwidth it became possible to release cuts on momenta of final state particles from $p > 5$ GeV/c and $p_t > 0.5$ GeV/c to $p > 3$ GeV/c and $p_t > 0.3$ GeV/c at the HLT2 tracking stage. These cuts were particularly damaging in 2011 to multibody hadronic decays since tracks from charm mesons are typically softer than B decay products. This is a major problem for the Dalitz-like analyses using D-multibody hadronic decays, where the loss of entire regions of the phase space (such as the corners of the Dalitz planes) is due to these momenta cuts. In 2011, a partial solution of this problem was the two-stage line concept. Only in the first stage (looking for detached pairs of hadron) the cuts above were applied. The tracks added during the second stage were chosen among a list of tracks that had been completed by a "second loop" tracking recovering softer tracks ($p > 3$ GeV/c, $p_t > 0.3$ GeV/c). This second loop was run only for events passing the first stage.

Lower p and p_t cuts mean more background. For that reason, it was necessary to re-optimize the HLT2 lines. The efficiency of the lines was evaluated with respect to

MC signal events selected by a loose offline selection (Table 4.2). In particular, the p and p_t cuts were loosened, in order to measure the benefit of the softer HLT2 tracking.

The re-optimized lines had to respect the following criteria: the CPU had to be similar to 2011 and the same number of events should be retained from the data control sample. In practice, the optimization of rare charm decays HLT2 lines is a trade-off between signal efficiency and the two constraints on timing and retention rate. The latter requirement has to respect the bandwidth allocated to a given line, according the LHCb's physics priorities. It was checked on a sample of 905000 real data events, selected by an open trigger during the previous data-taking campaign. In practice, the complete set of lines was allowed to retain only of the order of 1 event from this sample. The total CPU consumption was not allowed to be more than a ~ 0.5 s of CPU, when running offline on an `lxplus 2.8 GHz Xeon` machine. These constraints had to be satisfied before the new lines were included in HLT2.

To make the rate and efficiency results easier to interpret new lines were written for each mode:

- instead of generic $\underline{D^+ \to h^\pm\mu^+\mu^\mp}$ line:
  $D^+ \to \pi^+\mu^+\mu^-$,
  $D_s^+ \to \pi^+\mu^+\mu^-$,
  $D^+ \to K^+\mu^+\mu^-$,
  $D_s^+ \to K^+\mu^+\mu^-$,
  $D^+ \to \pi^-\mu^+\mu^+$,
  $D_s^+ \to \pi^-\mu^+\mu^+$,
  $D^+ \to K^-\mu^+\mu^+$,
  $D_s^+ \to K^-\mu^+\mu^+$

- instead of generic $\underline{D^0 \to h^+h^-\mu^+\mu^-}$ line:
  $D^0 \to \pi^+\pi^-\mu^+\mu^-$
  $D^0 \to K^+K^-\mu^+\mu^-$
  $D^0 \to K^\pm\pi^\mp\mu^+\mu^-$.

The cuts in Table 4.3 is a result of the scan for the best signal efficiency. Going from tight to loose momenta cuts makes the retention, timing and efficiency increase by almost a factor two. Additional cuts are required in order to come back to the initial timing and reasonable retention rate, keeping the efficiency to be as high as possible. Thus, some cuts, like track quality and Direction Angle (DIRA), are tightened. For 4-body decays there are additional cuts, such as vertex χ^2, impact parameter χ^2, flight distance χ^2, that must be harsher, as in high multiplicity decays backgrounds tend to increase faster than the signal yield when loosening cuts. The cuts changed in 2012 are presented in the second column of Table 4.1 in comparison to the 2011 level of cuts shown in the first column. The unchanged cuts between 2011 and 2012 are not repeated in the second column.

The gains of all lines can be found in the first column of Table 4.3. One can notice the better improvement for the modes with a kaon. As a kaon is heavier than pion, the larger part of energy of the D meson is recuperated by a kaon, so the rest of the energy shared between muons is smaller. This means that muons, in general, have

lower momenta in the decays with a K (Figure 4.1). Thus, the loosening of the criteria on momenta has a larger impact for these decays. In the case of 4-body decays this effect is more pronounced as there are pairs of kaons in the final state that makes the muons softer. It was checked that for three trigger lines, $D^0 \to K^+K^-\mu^+\mu^-$, $D^0 \to \pi^+\pi^-\mu^+\mu^-$ and $D^0 \to K^-\pi^+\mu^+\mu^-$, this behaviour stays similar between the online-offline selections.

The gain from the loosening of momenta cuts in 2012 trigger configuration without any additional cuts applied is about 50-60%. However, the fully optimized offline selections tighten the cuts, thus the gain in efficiency is less than this in the end. Indeed, the efficiency gain obtained offline after a nominal stripping preselection (shown in the second column of Table 4.3) is about 10% for $D \to h\mu\mu$ modes and 20-30% for $D \to hh\mu\mu$ decays.

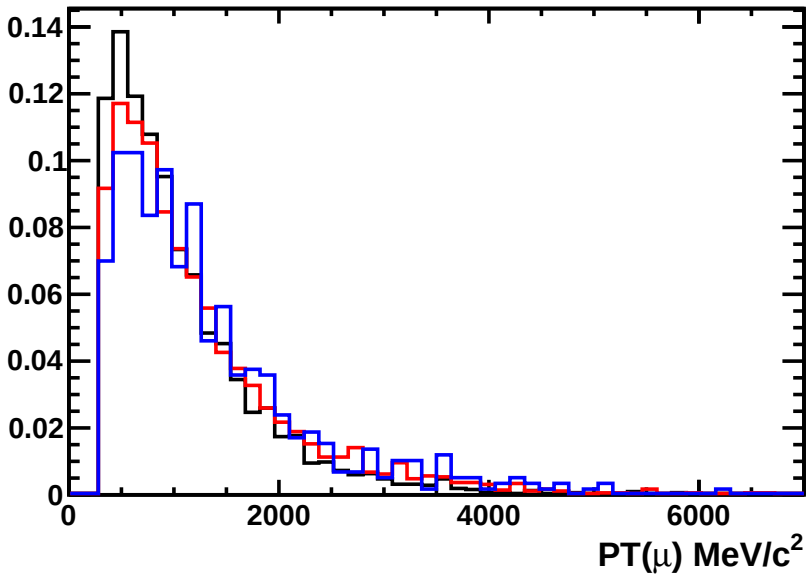


Figure 4.1: Transverse momentum distributions of a muon coming from $D^0 \to K^+K^-\mu^+\mu^-$ (black), $D^0 \to K^+\pi^-\mu^+\mu^-$ (red) and $D^0 \to \pi^+\pi^-\mu^+\mu^-$ (blue) in MC to illustrate that muons are softer in $D^0 \to K^+K^-\mu^+\mu^-$ decay than in $D^0 \to K^+\pi^-\mu^+\mu^-$ and $D^0 \to \pi^+\pi^-\mu^+\mu^-$.

4.1.3 The HLT2 selection for $D^0 \to h^+h^-h^+h^-$ decay

The HLT2 lines for $D^0 \to h^+h^-h^+h^-$ decay are similar to $D^+ \to h^+h^{(\prime)-}\mu^+\mu^-$ lines. The optimization of these lines was not part of my work. However, HLT2 lines for purely hadronic modes are important for the analysis presented in Section 5, especially for the $D^0 \to K^-\pi^+\pi^+\pi^-$ mode used as a normalization and control mode. Thus, the HLT2 strategy and 2012 level of cuts used to select candidates are presented here.

As in case of $D^0 \to h^+h^-h^+h^-$ lines a two stage strategy is used:

- find a hadron pair;

- add two hadrons in selected events.

The two hadrons added at the second stage can be soft thanks to the second loop tracking. The written lines are rather inclusive and select candidates: $D^0 \to \pi^+\pi^-\pi^+\pi^-$, $D^0 \to K^+\pi^-\pi^+\pi^-$, $D^0 \to K^+K^-\pi^+\pi^-$ and $D^0 \to K^+K^-K^+\pi^-$. The cut values used during the 2012 data taking period are listed in Table 4.4. It was kept to be as close as possible to the HLT2 selection of $D^+ \to h^+h^{(\prime)-}\mu^+\mu^-$ decays.

Cuts 1st stage		2011	2012
p_t	$>$	500. MeV/c	300. MeV/c
p	$>$	5000. MeV/c	3000. MeV/c
Impact parameter (IP) χ^2	$>$	2.	
Track χ^2/n_{dof}	$<$	5.	4.
MIN distance of closest approach (DOCA)	$<$	0.1 mm	
Flight distance (FD) χ^2	$>$	20.	9.
Cuts 2ed stage for $D \to h\mu\mu$		2011	2012
p_t	$>$	300. MeV/c	
p	$>$	3000. MeV/c	
$\sum IP\chi^2$	$>$	15.	
Track χ^2/n_{dof}	$<$	5.	4.
D MIN DOCA	$<$	0.1 mm	
D MAX DOCA	$<$	0.25mm	
$\sum p_t$	$>$	1500. MeV/c	500. MeV/c
D vertex χ^2/n_{dof}	$<$	20.	
D direction angle (DIRA)	$>$	0.9998	0.9999
D IP χ^2	$<$	36.	
D FD χ^2	$>$	20.	
Cuts 2ed stage for $D \to hh'\mu\mu$		2011	2012
p_t	$>$	300. MeV/c	
p	$>$	3000. MeV/c	
MAX $IP\chi^2$	$>$	12.	
Track χ^2/n_{dof}	$<$	5.	4.
MIN DOCA	$<$	0.1 mm	
MAX DOCA	$<$	0.2 mm	
$\sum p_t$	$>$	2500. MeV/c	3000. MeV/c
D vertex χ^2/n_{dof}	$<$	20.	15.
D DIRA	$>$	0.9998	0.9999
D IP χ^2	$<$	36.	25.
D FD χ^2	$>$	25.	36.

Table 4.1: The final set of cuts used in HLT2 in 2012 (the second column) in comparison to the 2011 level of cuts (the first column). The unchanged cuts between 2011 and 2012 are not repeated in the second column.

$D \to h\mu\mu$ line		Cut
$p_t(h)$	$>$	300. MeV/c
$p(h)$	$>$	2000. MeV/c
$p_t(\mu)$	$>$	300. MeV/c
$p(\mu)$	$>$	3000. MeV/c
PIDK[a]	$>$	-1.
Track χ^2/n_{dof}	$<$	8.
D vertex χ^2/n_{dof}	$<$	5.
D DIRA[b]	$>$	0.9999
D IP[c] χ^2	$<$	60.
$D \to hh\mu\mu$ line		Cut
$p_t(h)$	$>$	300. MeV/c
$p(h)$	$>$	3000. MeV/c
$p_t(\mu)$	$>$	300. MeV/c
$p(\mu)$	$>$	3000. MeV/c
Max IP[c] χ^2	$>$	20.
PIDK[a]	$>$	-1.
Track χ^2/n_{dof}	$<$	8.
D vertex χ^2/n_{dof}	$<$	5.
D MAX DOCA[d]	$<$	0.25 mm
D DIRA[b]	$>$	0.9999
D IP[c] χ^2	$<$	30.
D FD[e] χ^2	$>$	9.

Table 4.2: The level of offline selection (stripping) cuts applied on Monte Carlo samples to efficiency tests.

[a]Particle Identification variable. Delta-log likelihood between kaon and pion hypotheses for a given track

[b]Direction angle

[c]Impact parameter

[d]Distance of Closest Approach

[e]Flight distance

Line	Gain, %
$D^+ \to \pi^+\mu^+\mu^-$	19.2 ± 0.9
$D_s^+ \to \pi^+\mu^+\mu^-$	21.8 ± 1.2
$D^+ \to \pi^-\mu^+\mu^+$	18.9 ± 1.1
$D_s^+ \to \pi^-\mu^+\mu^+$	16.4 ± 1.0
$D^+ \to K^+\mu^+\mu^-$	23.5 ± 1.3
$D_s^+ \to K^+\mu^+\mu^-$	18.8 ± 1.1
$D^+ \to K^-\mu^+\mu^+$	22.8 ± 1.2
$D_s^+ \to K^-\mu^+\mu^+$	24.6 ± 1.5
$D^0 \to K^+K^-\mu^+\mu^-$	24.5 ± 2.2
$D^0 \to \pi^+\pi^-\mu^+\mu^-$	11.2 ± 0.7
$D^0 \to K^+\pi^-\mu^+\mu^-$	19.6 ± 2.3

Table 4.3: The gain in efficiency of rare charm HLT2 lines between configurations in 2011 and 2012 data taking periods.

Cut		Value
Cuts 1st stage		
p_t	$>$	300. MeV/c
p	$>$	3000. MeV/c
Impact parameter (IP) χ^2	$>$	2.8
Track χ^2/n_{dof}	$<$	3.
MIN distance of closest approach (DOCA)	$<$	0.1 mm
Cuts 2ed stage		
p_t	$>$	300. MeV/c
p	$>$	3000. MeV/c
Impact parameter (IP) χ^2	$>$	2.8
Track χ^2/n_{dof}	$<$	3.
D MIN Distance of Closest Approach (DOCA)	$<$	0.1 mm
D MAX DOCA	$<$	0.25 mm
$\sum p_t$	$>$	1500. MeV/c
D vertex χ^2/n_{dof}	$<$	15.
D Direction angle (DIRA)	$>$	0.9999
D IP χ^2	$<$	50.
D Flight Distance (FD) χ^2	$>$	36.
p_t on π_{slow}	$>$	300. MeV/c
p on π_{slow}	$>$	3000. MeV/c

Table 4.4: The final set of cuts used in HLT2 in 2012 for purely charm hadronic modes $D^0 \to h^+h^-h^+h^-$.

4.2 Stripping offline preselection

The saved data are pre-processed offline for the analyses. This is done by the so-called "stripping", which is a loose selection, specialized for each interesting physics decay mode. This selection is also optimized to fulfil general LHCb requirements on timing and retention rate. At the same time this selection is kept as loose as possible.

Stripping lines, selecting signal $D^+ \to h^+h^{(\prime)-}\mu^+\mu^-$ and normalization/control modes $D^0 \to h^+h^-h^+h^-$, were developed in order to keep them as similar as possible. The level of cuts are presented in Table 4.5, where one may pay attention to the prescale for $D^0 \to h^+h^-h^+h^-$ lines. Table 4.6 presents the preselection for the $D^+ \to h^+h^{(\prime)-}\mu^+\mu^-$ and $D^0 \to h^+h^-h^+h^-$, where D^0 comes from $D^{*+} \to D^0\pi^+_{slow}$.

Cut		Sign $D^+ \to h^+h^{(\prime)-}\mu^+\mu^-$	Norm/Ctrl $D^0 \to h^+h^-h^+h^-$
p	$>$	3000. MeV/c	3000. MeV/c
p_t	$>$	300. MeV/c	300. MeV/c
Min IP[a] χ^2	$>$	2.	3.
Track χ^2/n_{dof}	$<$	3.	3.
μ: PIDmu	$>$	-5.	-
K: PIDK	$>$	-1.	-1.
GhostProb[b]	$<$	-	0.5
On D^0			
DOCA[c]	$<$	0.25 mm	0.25 mm
p_t	$>$	2500. MeV/c	2500. MeV/c
Max daugh. IP[b] χ^2	$>$	15.	15.
Fight distance χ^2	$>$	36.	36.
Vertex χ^2	$<$	8.	8.
DIRA[d]	$>$	0.9999	0.9999
D IP[b] χ^2	$<$	25.	25.
mass window	$=$	$\pm$80.	$\pm$80.
Dimuon Mass	$>$	250.	250.
Prescale	$=$	1.	0.03

Table 4.5: Cuts and prescales applied in the stripping selection for $D^+ \to h^+h^{(\prime)-}\mu^+\mu^-$ and $D^0 \to h^+h^-h^+h^-$.

[a] Impact parameter

[b] Ghost Probability variable, a Neural Net output variable, combining various tracks variables in order to distinguish the true and the ghost tracks that could be associated with a reconstructible MC particle.

[c] Distance of Closest Approach

[d] Direction angle

Cut		$D^+ \to h^+h^{(\prime)-}\mu^+\mu^-$	$D^0 \to h^+h^-h^+h^-$
On D^{*+}			
DOCA[a]	$<$	0.3 mm	0.3 mm
p_t	$>$	2000. MeV/c	2000. MeV/c
Vertex χ^2	$<$	20.	20.
On π^+_{slow}			
p_t	$>$	120. MeV/c	120. MeV/c
On daughters, K, π, μ			
p	$>$	3000. MeV/c	3000. MeV/c
p_t	$>$	300. MeV/c	300. MeV/c
IP[b]χ^2	$>$	3.	3.
Track χ^2/n_{dof}	$<$	3.	3.
$isMuon(\mu)$	$=$	True	-
On D^0			
DOCA[a]	$<$	0.3 mm	0.3 mm
p_t	$>$	2000. MeV/c	2000. MeV/c
Vertex χ^2	$<$	20.	20.
Max daughters $IP\chi^2$	$>$	9.	9.
Fight distance χ^2	$>$	30.	30.
D IP[b] χ^2	$<$	36.	36.
DIRA[c]	$>$	0.9998	0.9998
mass window	$=$	$\pm$100.	$\pm$100.
Δm lower	$>$	-8. MeV/c	-8. MeV/c
Δm upper	$<$	18. MeV/c	18. MeV/c
Prescale	$=$	1.	0.01

Table 4.6: Cuts and prescales applied in the stripping selection for $D^+ \to h^+h^{(\prime)-}\mu^+\mu^-$ and $D^0 \to h^+h^-h^+h^-$ coming from D^{*+}.

[a]Distance of Closest Approach
[b]Impact parameter
[c]Direction angle

Chapter 5

Measurement of the partial branching ratio of the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ decay with 2 fb^{-1} of pp data collected by LHCb in 2012

5.1 Motivation

The measurement of $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$ is required for the future analyses, where this mode will be used as normalization.

In the published LHCb analysis based on 1 fb^{-1} [14] searches for $D^0 \to \pi^+\pi^-\mu^+\mu^-$, the $D^0 \to \pi^+\pi^-\phi(\to \mu^+\mu^-)$ mode was used. However, the branching ratio of this mode has never been measured directly. Thus, in [14], it was derived from the amplitude analysis of $D^0 \to K^+K^-\pi^+\pi^-$ by CLEO-c [86]:

$$\mathcal{B}(D^0 \to \pi^+\pi^-\phi(\to \mu^+\mu^-)) = \mathcal{B}(D^0 \to \pi^+\pi^-\phi(\to K^+K^-)) \times \frac{\mathcal{B}(\phi \to \mu^+\mu^-)}{\mathcal{B}(\phi \to K^+K^-)}, \quad (5.1)$$

where $\mathcal{B}(D^0 \to \pi^+\pi^-\phi(\to K^+K^-))$ is derived from the total $D^0 \to K^+K^-\pi^+\pi^-$ branching fraction [40] and from the total $D^0 \to \pi^+\pi^-\phi$ fit fraction found in [86]:

$$\mathcal{B}(D^0 \to \pi^+\pi^-\phi(\to K^+K^-)) = \frac{\int d\Omega \left|\sum_{k=1}^{N_{\phi\pi^+\pi^-}} |a_k| e^{i\phi_k} \mathcal{A}_k(\Omega)\right|^2}{\int d\Omega \left|\sum_{j=1}^{N} |a_j| e^{i\phi_j} \mathcal{A}_j(\Omega)\right|^2} \times \mathcal{B}(D^0 \to K^+K^-\pi^+\pi^-) \quad (5.2)$$

with $k \in \{(\phi\rho)_{S-wave}, (\phi\rho)_{D-wave}, \phi(\pi^+\pi^-)_{S-wave}, ...\}$.

We decided to use another mode, $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$, for the following reasons:

- The above approach assumes that the decay of the ϕ is independent of the rest of the decay process, and that the relative amplitudes and phases of the various

$\phi\pi^+\pi^-$ components are the same if measured in the $\mu^+\mu^-\pi^+\pi^-$ final state. This is not strictly true. The formalism used by the amplitude analysis assumes that one can describe the $D^0 \to K^+K^-\pi^+\pi^-$ decay with a rather limited number of components, individually described by a simple function (e.g. Breit-Wigner), and by their relative phases. This neglects the final state interactions between the kaons and pions. Ideally, each component should be described by a long sum of contributions in order to describe all configuration of the possible final state interactions. Consequently, the result of this approximation depends on the final state interactions. The same amplitude analysis applied to the ϕ components of $D^0 \to \pi^+\pi^-\mu^+\mu^-$ is therefore not guaranteed to yield the same result as with $D^0 \to \pi^+\pi^-K^+K^-$. The possible difference is expected to be small due to of the narrowness of the ϕ. However, there's no easy way to quantify precisely this difference. For that reason, the main results in [14] were ratios of branching ratios, relative to $\mathcal{B}(D^0 \to \pi^+\pi^-\phi(\to \mu^+\mu^-))$, rather than absolute branching ratios.

- The evaluation of $\mathcal{B}(D^0 \to \pi^+\pi^-\phi(\to \mu^+\mu^-))$ with the approach above is limited by a 20% uncertainty. It stems from the variation in the result between all the alternative amplitude models considered in [86]. This was absolutely sufficient in the case of [14], where the goal was to set upper limits. Even with this uncertainty, these limits were better by two orders of magnitude with respect to previous studies. However, the updated study of $D^0 \to h^+h^{(\prime)-}\mu^+\mu^-$ decays with 3 fb^{-1} of LHCb data aims at measuring not only partial branching ratios in regions of $m(\mu^+\mu^-)$, but also total branching ratios including the resonant contributions. In some cases, a signal will be found and we will measure the branching ratio, not set an upper limit. A better known normalization branching fraction is preferable in this case.

- The fit fractions in [86] are very different from the previous values found in [40], where the results of complicated amplitude analyses are combined. Thus, independent additional information is therefore valuable.

- The $D^0 \to \pi^+\pi^-\phi(\to \mu^+\mu^-)$ component of the $D^0 \to \pi^+\pi^-\mu^+\mu^-$ decay interferes with the $D^0 \to \pi^+\pi^-\rho(\to \mu^+\mu^-)$ component. Taking $\mathcal{B}(D^0 \to \pi^+\pi^-\rho)$ again from the amplitude analysis of an hadronic final state [40], and scanning all possible relative phases, we determined the maximal effect on the total $D^0 \to \pi^+\pi^-\mu^+\mu^-$ branching ratio in the ϕ region could reach 10%. This is an additional source of uncertainty.

- One of the uncertainties in the measurement of $D^0 \to K^+\pi^-\mu^+\mu^-$ and $D^0 \to K^+K^-\mu^+\mu^-$ relative to $\mathcal{B}(D^0 \to \pi^+\pi^-\phi(\to \mu^+\mu^-))$ stems from the uncertainty on the hadron identification efficiency, which cannot cancel in the efficiency ratio since the final states are different. Using $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ as a normalisation mode would improve that.

- We believe $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$ can be measured with a precision better than 15% using LHCb 2012 data. Even better precision can be foreseen using the data that will be collected in Run II. This data would benefit of a better

understood detector and larger control samples to reduce systematic uncertainties, which dominate the above precision.

5.2 Analysis strategy

The aim of this analysis is to measure $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$, i.e. the partial branching fraction of $D^0 \to K^-\pi^+\mu^+\mu^-$ in the region of the $m(\mu^+\mu^-)$ spectrum: [675;875] MeV/c^2. We use the data collected by LHCb in 2012 (~ 2 fb^{-1} of pp collisions at $\sqrt{s} = 8$ TeV).

We focus on the ρ/ω region defined above for two reasons. First, this keeps blinded the rest of the spectrum, in particular the non resonant regions that can be used for NP searches. Second, the selection efficiency varies a lot across the phase space, which in the case of a four body decay is described by 5 independent variables (see Section 1.2.4). Unless a very large MC sample is available, only a few bins per variable can be used to determine this efficiency. This causes a dependence upon the decay model assumed when generating MC events. No reliable model is known so far for the decay $D^0 \to K^-\pi^+\mu^+\mu^-$. Reducing the portion of phase space used in the analysis reduces the impact of this efficiency variation.

The branching ratio in the ρ/ω window is expected to be $\mathcal{O}(10^{-6})$ [18], which is accessible for LHCb with the 2012 data sample. To further improve the statistics, we select prompt untagged D^0 decays, that are not required to come from the decay of a B hadron nor from that of a D^*.

The measurement can be summarized by the equation below:

$$\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)) = \mathcal{B}(D^0 \to K^-\pi^+\pi^+\pi^-) \times \frac{N_{signal}}{N_{norm}} \times \frac{\epsilon_{norm}}{\epsilon_{signal}} \tag{5.3}$$

The $D^0 \to K^-\pi^+\pi^+\pi^-$ mode is chosen as a normalization mode. Its branching ratio is large and precisely known, $\mathcal{B} = (8.287 \pm 0.043_{stat.} \pm 0.200_{syst.})$ % [19]. A crucial ingredient in this measurement is to measure the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ to $D^0 \to K^-\pi^+\pi^+\pi^-$ efficiency ratio. Thanks to its similarity (topology, kinematics, charge tracks multiplicity) to the signal, it is expected to be close to 1, and so systematic uncertainties will be cancelled to a large extent.

However, a crucial difference comes from the fact that the normalization mode has a purely hadronic final state while the signal involves a pair of muons. Thus some of systematic uncertainties cannot cancel in the efficiency ratio of Equation 5.3. These are essentially the muon identification efficiency and L0 trigger efficiency if the most "natural" and efficient lines are used, i.e. `L0HadronTOS`[1] in the case of $D^0 \to K^-\pi^+\pi^+\pi^-$ and `L0MuonTOS`[2] in the case of $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$. The later difference is *a priori* the most troublesome. These two trigger lines use information from different subdetectors (HCAL, ECAL and muon stations) and with different transverse momenta thresholds. Therefore, their efficiencies are different for modes with hadrons or muons in the final state and do not cancel in the efficiency ratio of Equation 5.3. Moreover, they are known to be imperfectly described by the MC. In order to reduce difference from L0

[1] The line where L0 selects hadron objects from the searched decay (Section 2.5.1)

[2] The line where L0 selects muon objects from the searched decay (Section 2.5.1 and Section 2.5.3).

between two channels we therefore require both modes to be triggered in the same way, using the hadron trigger but in a manner that does not depend on the signal decay, `L0HadronTIS`, which means that L0 was fired by a hadron that does not originate from the final state of $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ or $D^0 \to K^-\pi^+\pi^+\pi^-$, but from the rest of the event. Although this approach has lower efficiency it is systematically more robust.

The determination of the muon and hadron identification efficiencies does not rely on simulations. We use a data driven method (see Section 5.5.1).

The offline selection is performed by a dedicated stripping lines followed by a Boosted Decision Tree.

A large peaking background is caused by hadronic $D^0 \to K^-\pi^+\pi^+\pi^-$ decays with a double misidentification of the π as a μ. It peaks $\sim 20 MeV/c^2$ below the D^0 signal peak in the invariant mass distribution. A muon identification variable, $ProbNNmu$ [3], is used to reduce this background. The residual contamination is accommodated in the binned likelihood fit to the signal sample that we use to measure the signal and derive $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$. The shape of this background is not trivial to determine: pion decays in flight $\pi^+ \to \mu^+\nu_\mu$ causes tails which cannot be studied precisely in MC due to the rareness of the double $\pi \to \mu$ misidentification.

The analysis was designed as blind. Several control samples are used to guide its design. A part of the signal sample (20% of the data sample) is used to optimize the selection: the significance of the signal peak is optimized as a function of the cut on the BDT score and on $ProbNNmu$. A sample obtained with the same selection as the signal sample apart from the `L0HadronTIS` requirement which is replaced by the more efficient criteria `L0HadronTOS||L0MuonTOS` in order to yield a large sample of $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ decays without unblinding the analysis, is used for a number of cross-checks. Half the sample of $D^0 \to K^-\pi^+\pi^+\pi^-$ decays is also used as a control mode.

5.3 Data and MC samples used in the analysis

We use the data collected by LHCb in 2012 (~ 2 fb^{-1} of pp collisions at $\sqrt{s} = 8$ TeV). No dedicated HLT2 lines existed in 2011. The sample is split into three sub-samples:

- 2%: upper sideband, $m(D^0) > 1890.$ MeV/c^2, is used for BDT training and testing as a background input sample;
- 20%: used for finding an optimal selection, checking the fit model (see Section 5.7.5), for data/MC comparisons and systematic studies (see Section 5.5.5) etc.
- 78%: used for the measurement of $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$.

As far as the normalisation mode is concerned, we use 50% of 2012 data for the measurement and the rest as a control mode. Dedicated HLT2 lines and preselection (stripping) lines have been used (see Section 5.4). They were designed with rather loose cuts in order to be as close as possible to the lines selecting $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$.

[3] See Section 2.9.2

In order to keep their rates affordable, these lines were prescaled. The prescale factor equals 0.1 for the HLT2 line and 0.03 for the stripping line.

We use Monte Carlo samples, generated using two versions:

- "old MC", produced in 2012[4]. The $D^0 \to K^-\pi^+\pi^+\pi^-$ mode was described by an incoherent sum of the fit fractions found in [40].
- "new MC", produced in late 2013[5], when a better knowledge of detector's conditions, for example, calibration, in 2012 was obtained. The $D^0 \to K^-\pi^+\pi^+\pi^-$ modes are generated using a package (MINT [87]), describing the resonance structures, studied in data.

For $D^0 \to K^-\pi^+\mu^+\mu^-$ the decay model is an incoherent sum of six decay: $D^0 \to K^{*0}\rho$, $D^0 \to K^{*0}\omega$, $D^0 \to K^{*0}\mu^+\mu^-$, $D^0 \to K^-\pi^+\rho$, $D^0 \to K^-\pi^+\omega$, $D^0 \to K^-\pi^+\mu^+\mu^-$ (non res). This was inspired by [18] and by observations in control modes. The "Old MC" $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ sample is used only for BDT training and testing. The "New MC" samples are use for the estimation of the efficiency ratios.

A summary of the used samples is presented in Table 5.1.

Mode	Statistics
"old MC" $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$	3 M
"new MC" $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$	4 M
"old MC" $D^0 \to K^-\pi^+\pi^+\pi^-$	2 M
"new MC" filtered $D^0 \to K^-\pi^+\pi^+\pi^-$	2 M
data $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ MagUp	1034 pb^{-1}
data $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ MagDown	1027 pb^{-1}
data $D^0 \to K^-\pi^+\pi^+\pi^-$ MagUp	1012 pb^{-1}
data $D^0 \to K^-\pi^+\pi^+\pi^-$ MagDown	1027 pb^{-1}

Table 5.1: List of used samples

5.4 Selection

The signal selection comprises several steps. The first one is a trigger selection. It is presented in Section 5.4.1, followed by the stripping (Section 4.2). On top of the stripping and the trigger we apply multivariate analysis techniques to reduce the amount of combinatorial background and the muon identification criteria are applied against peaking background.

[4]For LHCb experts: Sim06b, Reco 14, Beam4000GeV-MayJune2012-MagUp-Nu2.5-EmNoCuts
[5]For LHCb experts: Sim08, Reco14, Beam4000GeV-2012-MagUp-Nu2.5-Pythia8

5.4.1 Trigger

The main features of this analysis are dictated by the difference between the semileptonic final state of the signal and the purely hadronic final state of the normalization mode. The most efficient L0 trigger criterion for $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ is `L0HadronTOS||L0MuonTOS`. However, $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ events would be triggered primarily due to the `L0Muon` line, as opposed to $D^0 \to K^-\pi^+\pi^+\pi^-$, which would be dominated by `L0Hadron`. This is due to the difference in the nature of their decay products. These two lines are based on different subdetectors, and apply different cuts on the p_t of the objects they search for (see Section 2.5.1). The systematic uncertainty on the efficiency with which each mode is selected is therefore different, and does not cancel in the efficiency ratio. Because this uncertainty is difficult to determine, we decided to use criteria that ensures similar efficiency and systematics: both modes are required to satisfy `L0HadronTIS`, i.e. the L0 trigger is fired by particles not from the signal decay. Indeed, the rest of the event should be similar at first order for the signal and normalization mode, giving a reasonably unbiased sample. The `L0HadronTIS` requirement is at least five times less efficient than `L0HadronTOS||L0MuonTOS` but the measurement is still possible thanks to the high branching ratios and charm production cross-section.

For the HLT1 we use the same line for both modes, and lines that are similar to a large extent in the case of HLT2. This is summarized below:

- L0: `D Hadron TIS`, i.e. an event must be fired by a particle with $E_t > 3.5 - 3.74$ GeV/c not from the signal decay
- HLT1: `D TrackAllL0 TOS`, i.e. selection of the reconstructible tracks passed L0 selection and with impact parameter with respect to primary vertex $IP > 0.1$ mm and $p_t > 1.7$ GeV/c.
- HLT2 dedicated line for signal: `D CharmSemilepD02KPiMuMu TOS`
- HLT2 dedicated line for normalization: `D CharmHadD02HHHH_K3pi TOS` (with a 0.1 prescale)

The detail of the cuts applied by the HLT2 lines is available in Section 4.1 and the definition of `TOS` and `TIS` can be found in Section 2.5.3.

Table 5.2 presents the fractions of stripped events that pass the main trigger lines.

The preselection (stripping) selection is presented in Table 4.5 of Section 4.2.

5.4.2 Offline selection

On the top of the stripping selection we apply additional offline cuts. We use a Multivariate Analysis (MVA) selection against combinatorial background and a muon particle identification variable, $ProbNNmu$[6], to minimise peaking background from hadronic modes with $\pi - \mu$ misidentification.

In this analysis we use a Boosted Decision Tree (BDT) for the first step. The BDT is adapted from the tagged analysis (see Section 6.1.4) without including the variables

[6]Definition of the variables can be found in Section 2.9.2.

Trigger line	Fractions of stripped events selected by each trigger line, %
L0Global TOS	91.7
L0Muon TOS	84.6
L0Muon TIS	10.6
L0Hadron TOS	13.9
L0Hadron TIS	20.7
Hlt1Global TOS	91.3
Hlt1TrackAllL0 TOS	56.7
Hlt2Global TOS	94.5
Hlt2CharmSemilepD02KPiMuMu TOS	75.2
Hlt2CharmHadD02HHHH_K3pi TOS	6.0

Table 5.2: Fraction of stripped events in which certain trigger lines that matter in this analysis are fired. *Global* lines are given for a comparison. All events accepted by the signal stripping line are considered, including background events.

on π^+_{slow} and D^{*+}, which are not defined for prompt D^0 candidates. The resulting Receiver Operating Characteristic (ROC) curve and overtraining checks are presented in Figures 5.1 and 5.2.

After this, as in Section 6.1.4, we look for the best combination of values for the cuts on (BDT;$ProbNNmu$) in order to maximise the signal significance. The result is consistent with what we find when minimizing the uncertainty on the fitted yield.

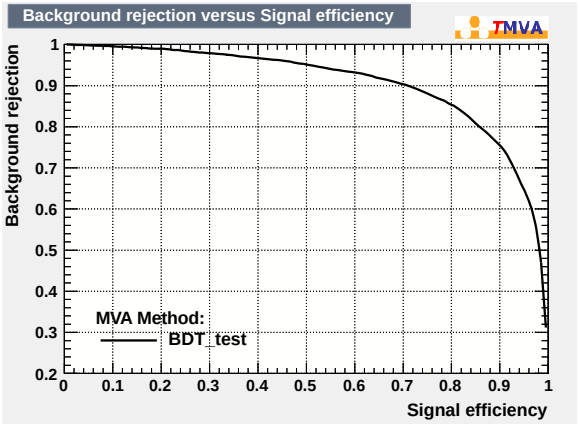


Figure 5.1: Background rejection vs. signal efficiency (Receiver Operating Characteristic (ROC) curve) for the trained Boosted Decision Tree (BDT).

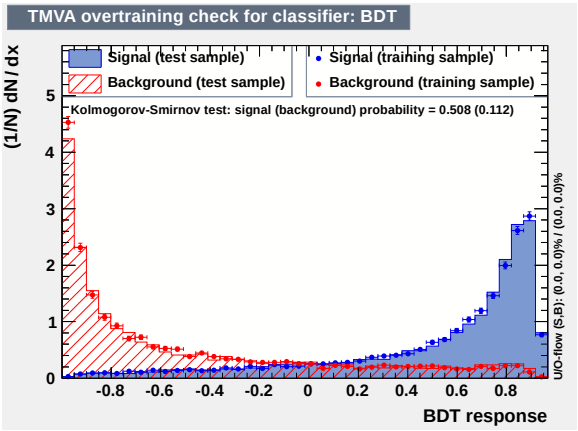


Figure 5.2: Signal (blue) and background (red) classification for the test and training samples.

In order to choose the cut on the BDT output variable, BDT, and muon identification selection criterion, $ProbNNmu$, we use 20% of the data signal sample which are not used in the final fit (see Section 5.7). We make a scan over various cuts on BDT and $ProbNNmu$ applied on both muon candidates. For each value of the cuts we obtain a signal yield N_{signal}, the combinatorial background level, N_{comb}, and the peaking

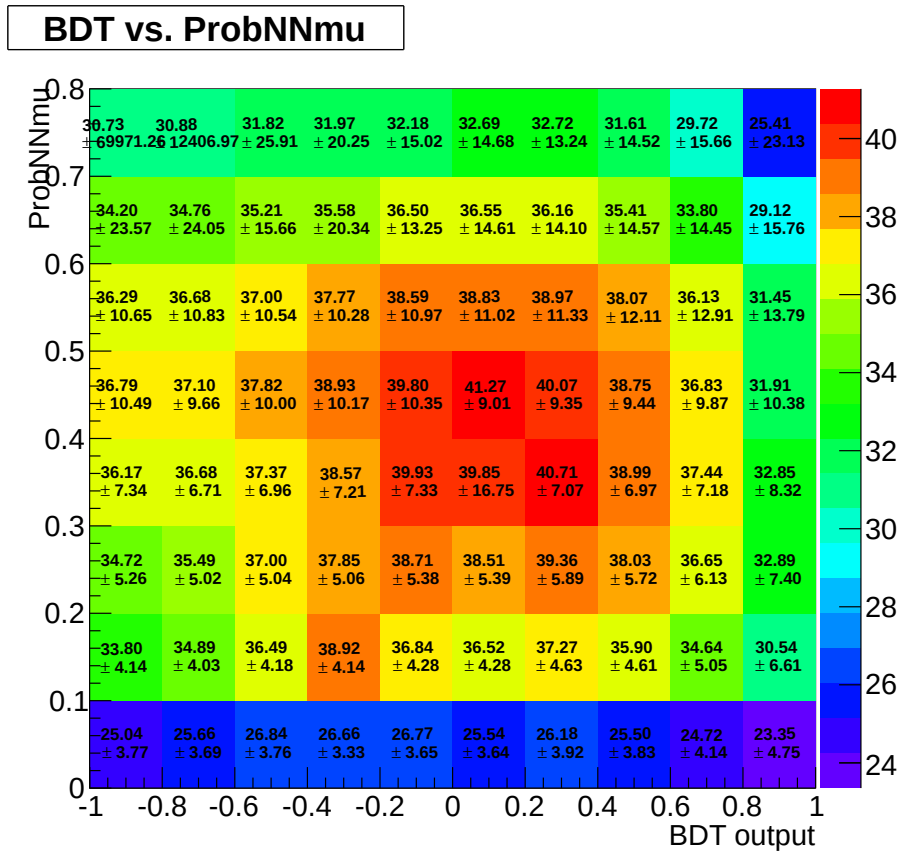


Figure 5.3: Optimization of the cut values of BDT and $ProbNNmu$ cuts.

contribution under D^0 mass peak, $N_{peaking}$. For each step the signal significance $N_{signal}/\sqrt{N_{signal} + N_{comb} + N_{peaking}}$ is computed, where all the yields are extracted from the signal box and scaled to the full sample, i.e. to the sample eventually used for the measurement (78% of the 2 fb^{-1} sample). The scan is shown in Figure 5.3. We observe a plateau in the significance. We chose this combination of cuts: ($BDT > 0.2; ProbNNmu > 0.3$). We chose the low $ProbNNmu$ side of the plateau for the following reason: the harder the cut on $ProbNNmu$, the larger the difference between the signal and normalisation mode, and the larger the systematic uncertainty in the efficiency ratio.

5.4.3 Multiple candidates treatment

The same tracks can be used several times to construct more than one candidate in each event. Moreover, it is possible to have different candidates constructed from the same tracks used in a different order. For example, tracks t_1, t_2, t_3, t_4 can give two candidates $D^0 \to K^-_{t_1}\pi^+_{\mathbf{t_2}}\mu^+_{\mathbf{t_3}}\mu^-_{t_4}$ and $D^0 \to K^-_{t_1}\pi^+_{\mathbf{t_3}}\mu^+_{\mathbf{t_2}}\mu^-_{t_4}$, where both t_2 and t_3 are seen as π^+ or μ^+. If such D^0 candidates have similar masses, they could introduce a multiple counting in the signal peak region, which would bias the yield extraction.

In order to avoid the multiple counting we search for such candidates, i.e. composed of the same tracks with similar masses, and keep randomly only one such candidate per

event. It has an effect of less than 1% on the fitted yield.

5.5 Efficiency ratio determination

The ratio between the efficiency with which the signal and normalization modes are selected is one of the main ingredients in the measurement of $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$.

The determination of $\epsilon_{D^0\to K^-\pi^+\rho/\omega(\to\mu^+\mu^-)}$ and $\epsilon_{D^0\to K^-\pi^+\pi^+\pi^-}$ is based mainly on MC.

The simulation is known to describe the physics of the decays and the performance of the detector in an imperfect manner. This difficulty is particularly acute when it comes to the L0 trigger and the identification of the muons. This is of crucial importance in this analysis because the final states of the signal and normalisation modes differ, the latter involving pions where the former involve muons. The data-driven determination of the systematic uncertainty on the L0 trigger efficiency is described in Section 5.5.3. The determination of the muon identification efficiency is based on data. This is presented in Section 5.5.1.

The accuracy with which the MC reproduces the distributions of all the other variables used in the analysis (e.g. (transverse) momenta, impact parameters, etc.) has a direct impact on the determination of the selection efficiency. This is the *raison d'être* of Section 5.5.5. The main part of this section focuses on the BDT. It combines most of these variables. It is therefore a natural choice to address the matter in a synthetic way.

To evaluate the impact of data/MC discrepancies on the cuts applied at stripping level (where no data/MC comparison can be done), the MC samples are re-processed in order to smear some of the track parameters. The LHCb impact parameter smearing procedure is used in Section 5.5.6.

The MC samples used to determine the selection efficiency are presented in Section 5.3. The efficiency of all the selection criteria is derived from MC. One exception is the efficiency of the muon identification criteria (Section 5.5.1), which is determined from real data. Table 5.4 presents the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ efficiencies at the main stages of the selection, as well as their ratio.

5.5.1 Data driven efficiency: muon identification

Identification criteria are applied to the final state muons for $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ but are not applied in the case of $D^0 \to K^-\pi^+\pi^+\pi^-$. Their efficiency does not cancel in the efficiency ratio and must be studied with special care. The Monte Carlo simulation might not reproduce with enough accuracy the distribution of the estimators used to identify muons: $isMuon$, binary muon identification variable, and $ProbNNmu$[7]. This is why the efficiency of these criteria is evaluated with the help of $J/\psi \to \mu^+\mu^-$ decays isolated in real data. This is done in practice by using the $PIDCalib$ [88] package.

The unbiased sample of muons from $J/\psi \to \mu^+\mu^-$ events is obtained using a *tag-and-probe* method. This means that the J/ψ candidates are selected with kinematical requirements and muon identification criteria applied only on one of the muons. The

[7] Definitions given in Section 2.9.2

second muon remains unbiased by any identification cut and is used for the muon identification efficiency calculations.

The efficiency strongly depends on the kinematics of the muon tracks and the multiplicity of the considered event. The $PIDCalib$ package provides a data-based measurement in regions of momentum p, pseudorapidity η and of $nTracks$, the event track multiplicity. The precise definition of these regions is important. The binning should not be too coarse as in this case we lose sensitivity to the variation of efficiency inside a bin. It should not be too fine either, as then we would not have enough statistics inside a bin. The efficiencies found in each region by $PIDCalib$ are used to weight signal MC events that have been selected without any muon identification criteria. The ratio of the weighted sum of these events to the total number of events in the MC sample provides the efficiency.

In order to define the finest reasonable binning, we proceed as follows:

1. Check the variation of the efficiency as a function of muon momentum in 6 bins of equal size in pseudorapidity η.

2. Merge bins in momentum p in order to increase the statistics inside a bin, especially at high momentum regions of the spectrum. No information on the efficiency variation seems to be lost when the number of bins is reduced to 13.

3. Define 6 bins in $nTracks$ with the similar population and check the stability of the efficiency distributions again.

One can find an illustration of these binnings in Appendix D. However, as can be seen in Figure 5.4, which shows the distribution of the MC signal events, as well as that of the calibration muons used by $PIDCalib$, superimposed to the binning, some bins at low-P and high-η contain no or very few calibration muons. In this case, no usable $PIDCalib$ weights exist. In order to avoid this problem, our finest binning, called "fine", is actually a composition of two binnings: MC events found in a bin where no usable $PIDCalib$ weight is available are reassigned to the corresponding region of a coarser binning - $(p;\eta;nTracks) = (9;3;6)$ - in which a weight can be be calculated. This is illustrated in Figure 5.4, where the violet area represents regions where the original $(p;\eta;nTracks) = (13;6;6)$ binning is applied while in orange regions the $(p;\eta;nTracks) = (9;3;6)$ binning is used. In order to study the dependence of the efficiency upon the binning choice, we also defined a "coarse" binning: $(p;\eta;nTracks) = (5;1;1)$. As will be shown later, the efficiency obtained with the latter differs from that obtained with the "fine" binning by only a few percent. This suggests that one intermediate binning between the "fine" and "coarse" ones will lie in a region where the dependence upon the binning definition is sufficiently low. Consequently, our "default" binning is $(p;\eta;nTracks) = (9;3;6)$.

Table 5.3 presents the efficiencies for each binning when $IsMuon$ and $ProbNNmu$ are applied to one muon only and when it is applied on both muons simultaneously. Variation of efficiencies between "fine" and "default" binnings is less than 2.5%. This can be taken as the corresponding systematic uncertainty.

The efficiency computed with $PIDCalib$, $(75.7 \pm 0.4)\%$, is comparable with the one from MC, $(71.6 \pm 1.9)\%$, presented in the last line of Table 5.3.

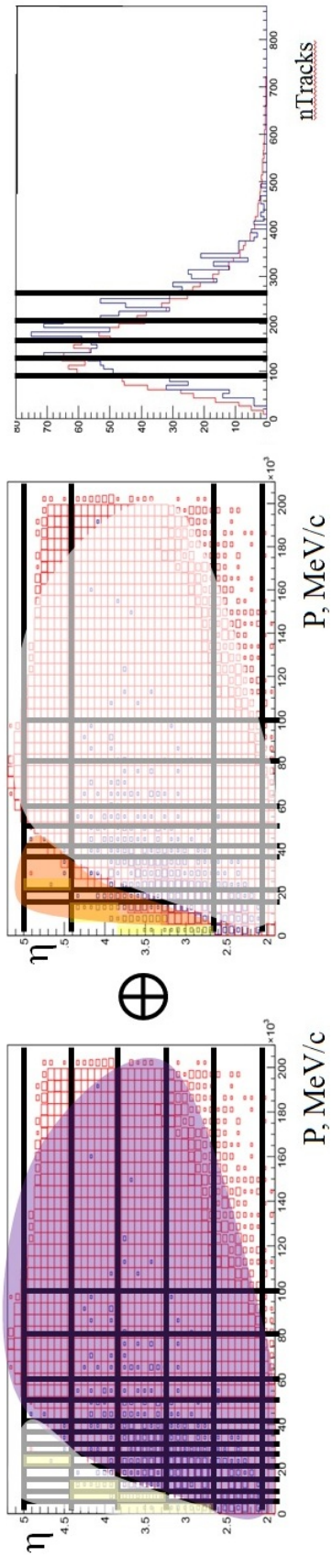


Figure 5.4: Definition of "fine" binning: Calibration sample (red), MC signal sample (blue). Violate area is the region where "fine" binning is used, orange area represents where "default" binning is used (see text for detailed explanation).

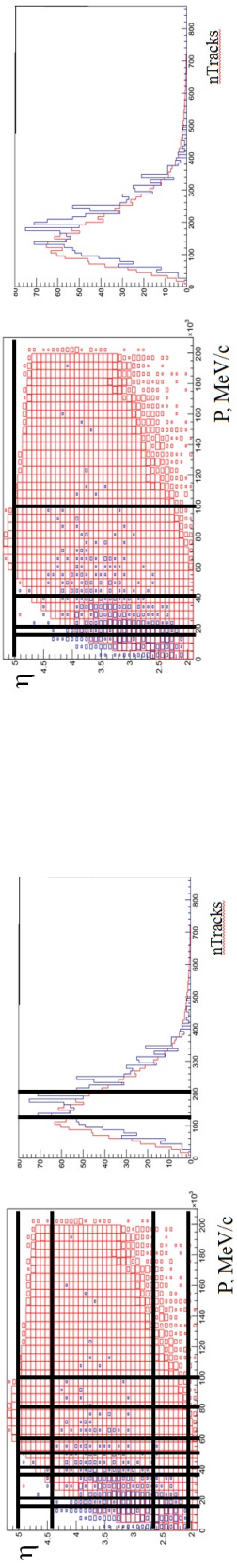


Figure 5.5: Definition of "default" binning: Calibration sample (red), MC signal sample (blue)

Figure 5.6: Definition of "coarse" binning: Calibration sample (red), MC signal sample (blue)

Name	Binning in $(P;\eta;nTracks)$	$\epsilon(\mu^+)$	$\epsilon(\mu^-)$	$\epsilon(\mu^+ \times \mu^-)$
Fine	(13/9;6/3;6)	85.7 ± 0.6	86.0 ± 0.7	73.9 ± 0.8
Default	(9;3;6)	86.7 ± 0.3	87.3 ± 0.3	75.7 ± 0.4
Coarse	(5;1;1)	87.0 ± 0.2	87.0 ± 0.2	75.7 ± 0.2
MC	-	84.8 ± 1.5	84.9 ± 1.5	71.6 ± 1.9

Table 5.3: Efficiency in % of $ProbNNmu > 0.3$ cut for each binning for μ^+ and μ^- separately and simultaneously

The $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ efficiency finally used in the branching fraction measurement is computed as follows:

- The efficiency of a looser selection that applies no muon identification criteria and also ignores HLT2 is determined from the simulation.
- This efficiency is multiplied by the efficiency determined with $PIDCalib$: $\epsilon = 75.7 \pm 0.4\%$. (see Table 5.3).
- The efficiency is further multiplied by the efficiency of the HLT2 requirement, determined on MC events on which all the selection criteria, including the muon identification criteria, have been explicitly imposed.

The final efficiencies and efficiency ratios are shown in Table 5.4

Selection criteria	$\epsilon_{D^0\to K^-\pi^+\rho/\omega(\to\mu^+\mu^-)}$ (%)	$\epsilon_{D^0\to K^-\pi^+\pi^+\pi^-}$ (%)	$\frac{\epsilon_{D^0\to K^-\pi^+\rho/\omega(\to\mu^+\mu^-)}}{\epsilon_{D^0\to K^-\pi^+\pi^+\pi^-}}$
Acceptance	19.13 ± 0.10	19.85 ± 0.08	0.964 ± 0.006
Reco., Strip., Presel.	0.228 ± 0.002	0.165 ± 0.003	1.38 ± 0.02
BDT >0.2	0.190 ± 0.002	0.142 ± 0.002	1.34 ± 0.02
L0 TIS	0.0407 ± 0.0006	0.0306 ± 0.0005	1.33 ± 0.03
HLT1 TOS	0.0171 ± 0.0004	0.0134 ± 0.0002	1.27 ± 0.04
$IsMuon$ and $ProbNN(\mu^+\|\mu^-) > 0.3$, from PIDCalib,	0.0130 ± 0.0003	0.0134 ± 0.0002	0.96 ± 0.03
HLT2 TOS	0.0103 ± 0.0003	0.0100 ± 0.0002	1.04 ± 0.03

Table 5.4: $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ cumulative selection efficiencies, and their ratio, after the data driven determination of the muon identification efficiency

5.5.2 Pion inefficiency: decays in flight and nuclear interactions in detector's material.

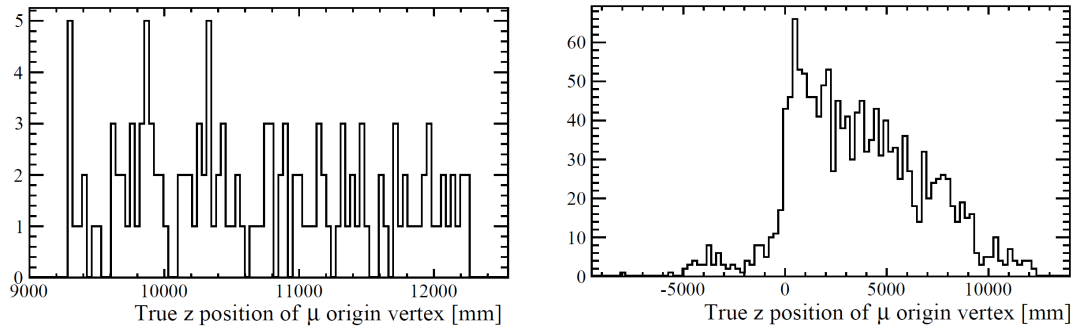
The similarity of the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ final states (see also Section 5.5.5) suggests that the selection efficiency should essentially be the same for both modes. Thus, the efficiency ratio is expected to be close to one. In Table 5.4, one can see this rule is violated twice. One case is obvious: the muon identification criteria are applied only to the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ mode. This creates a $\sim 25\%$ difference in efficiency between $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$. One more case is the 38% difference in the "reconstruction+stripping[8]+preselection" efficiency. This is essentially explained by the fact that the fraction of pions which are "reconstructible" is lower that the same fraction for muons. This is due to decays in flight and nuclear interactions in the material of the detector. The LHCb tracking algorithm reconstructs tracks that leave a minimum number of hits, in some fields of interest opened up in the three tracking stations, T1, T2 and T3. If a particle decayed before the last tracking station or interacted with the detector material it is not reconstructible by the standard algorithm. Using our MC samples, we determined the efficiency with which at generator level $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ decays are selected when one requires all the decay products to be "reconstructible", and "reconstructed", derived from the tracking algorithm. The result is in Table 5.5. We focus on events which satisfy the stripping requirements on momenta and transverse momenta, applied at generation level. One can notice the efficiency ratio of "reconstructible" decays is 1.47 ± 0.11, and that efficiency ratio of "reconstructed" decays is 1.40 ± 0.12. This is consistent with the result in Table 5.4.

Criteria	$\epsilon_{D^0\to K^-\pi^+\rho/\omega(\to\mu^+\mu^-)}$ (%)	$\epsilon_{D^0\to K^-\pi^+\pi^+\pi^-}$ (%)	$\frac{\epsilon_{D^0\to K^-\pi^+\rho/\omega(\to\mu^+\mu^-)}}{\epsilon_{D^0\to K^-\pi^+\pi^+\pi^-}}$
Reconstructible	0.045 ± 0.002	0.031 ± 0.002	1.47 ± 0.11
Reconstructed	0.040 ± 0.002	0.028 ± 0.001	1.40 ± 0.12

Table 5.5: Efficiency for all the decay products of $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ to be simultaneously "reconstructible" and "reconstructed", and corresponding efficiency ratios.

To obtain evidence that the difference in the "reconstructible" efficiency is partly due to pion decays in flight, we used the MC truth information to isolate decays such as $D^0 \to K^-\pi^+\pi^+(\pi^- \to \mu^-\nu)$. Figure 5.7 shows the z-position of the true creation vertex of the muon of this decay chain when the corresponding pion is "reconstructible", or not. From this figure one can see that majority of non-reconstructible events has a vertex of $\pi^- \to \mu^-\nu_\mu$ before the 9000 mm, i.e. before the last tracking station, thus, cannot be reconstructed by the tracking algorithm. The effect of nuclear interactions is quantified in [89].

[8]LHCb centralized loose preselection



(a) Reconstructible $D^0 \to K^-\pi^+\pi^+(\pi^- \to \mu^-\nu)$

(b) Non-reconstructible $D^0 \to K^-\pi^+\pi^+(\pi^- \to \mu^-\nu)$

Figure 5.7: Z-position of the creation vertex of the muon from $D^0 \to K^-\pi^+\pi^+\pi^-(\to \mu^-\nu)$, when the corresponding pion is "reconstructible" (a), or not (b).

5.5.3 L0 trigger efficiency

Due to the nature of their decay products, the D^0 candidates are most efficiently selected when one requires them to explicitly trigger the `L0Muon` line (`D_L0MuonTOS`) in the case of the signal and the `L0Hadron` line (`D_L0HadronTOS`) in the case of the normalisation mode.

However, the efficiency of these criteria is not re-produced accurately by the MC simulation. The associated systematic uncertainty would essentially cancel in the $\epsilon_{D^0\to K^-\pi^+\rho/\omega(\to\mu^+\mu^-)}$ / $\epsilon_{D^0\to K^-\pi^+\pi^+\pi^-}$ efficiency ratio, if both the `L0Muon` and `L0Hadron` line applied similar criteria, based on the same sub-detectors. This is not the case. Consequently, this efficiency ratio cannot be determined precisely.

The solution we chose is to require that the rest of the event triggers L0, rather than the D^0 candidate. For that purpose, in both the signal and normalisation samples, we select candidates that satisfy the `D_L0HadronTIS` requirement.

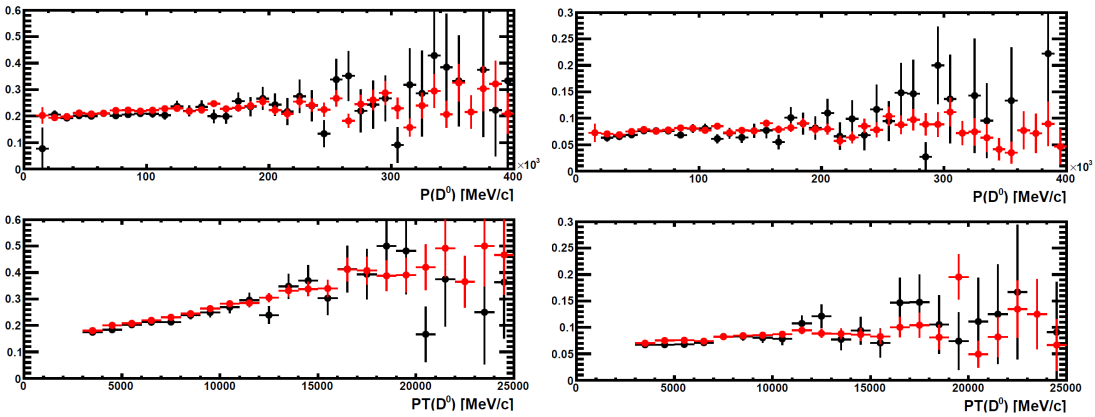
This choice was driven by the fact that in events containing a $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ decay, the rest of the event should be very similar to what it would be, had the $D^0 \to K^-\pi^+\pi^+\pi^-$ decay been produced instead. Indeed, the D^0 production mechanism and features influences the production of the other particles in the event. However, this does not depend on the D^0 decay mode. The kinematics of these particles is correlated with that of the D^0, which depends in principle on the way it decays, since two different final states sometimes require differing selections. However, the features of the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ final states are very similar, which allowed both modes to be selected in a very similar manner. For that reason, we expect the efficiency of the `D_L0HadronTIS` criteria to be very similar in $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ events.

A departure from this expected behaviour is possible if the particles from the rest of the event share hits in some sub-detectors with the D^0 decay products. For instance, one of clusters they cause in the hadronic calorimeter might overlap with one due to the D^0 decay products. This artificially increases the energy of the cluster and the probability for the L0 hadron line to fire the trigger. This would happen more often in the case of $D^0 \to K^-\pi^+\pi^+\pi^-$ because there are more hadrons in the final state.

To test the assumption according to which the `D_L0HadronTIS` efficiency is the same in $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ events, we compare the efficiency ratio with that obtained with the `D_L0MuonTIS` criteria. The sub-detectors used in this case are different. Thus, if the effect described in the previous paragraph actually exists, the `D_L0HadronTIS` and `D_L0MuonTIS` efficiency ratio should differ.

Figure 5.8 shows the efficiency of the `D_L0HadronTIS` requirement as a function of $p_t(D^0)$ and $p(D^0)$, when applied on events already selected by the pre-selection and the BDT. We notice that the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ efficiencies agree within a few percent. The same conclusion holds when the `D_L0MuonTIS` criteria are applied (Figure 5.8).

Figure 5.9 compares the $\epsilon_{D^0\to K^-\pi^+\rho/\omega(\to\mu^+\mu^-)}$ / $\epsilon_{D^0\to K^-\pi^+\pi^+\pi^-}$ efficiency ratios the criteria `D_L0HadronTIS` are applied to the ratio obtained with `D_L0MuonTIS`, as a function of $p_t(D^0)$ and $p(D^0)$. Both ratios are consistent with each other everywhere, and close to 1. It suggests that the efficiency ratio obtained with the MC simulation is correct.



(a) `D_L0HadronTIS` (b) `D_L0MuonTIS`

Figure 5.8: Efficiency of two L0 requirements as a function of $p(D^0)$ (top) and $p_t(D^0)$ (bottom), determined in MC. The $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ signal (black), is compared with $D^0 \to K^-\pi^+\pi^+\pi^-$ (red).

In Table 5.6, we show these efficiency ratios in the whole phase space. The difference between the `D_L0HadronTIS` and the `D_L0MuonTIS` efficiency ratios is below a percent.

The observations made from MC can be cross-checked on real data. After unblinding we compare the branching fractions obtained with `D_L0HadronTIS` criterion with those obtained with `D_L0MuonTIS` and `D_L0MuonTIS || D_L0HadronTIS` criteria. Before unblinding it was controlled in the following way: we compared how the signal yield varies in data when changing the `D_L0HadronTIS` criteria for `D_L0MuonTIS`. We compared this to the variation observed in MC.

In practice, we tested the ratio of ratios of yields in MC and data

$$\xi = \frac{N^{signal}_{L0MuonTIS}/N^{signal}_{L0HadronTIS}}{N^{norm}_{L0MuonTIS}/N^{norm}_{L0HadronTIS}}$$

at the level of nominal cuts. We used the 20% data signal sample to determine $N^{signal}_{L0HadronTIS}$ scaled to the full statistic (78%) and the full sample to de-

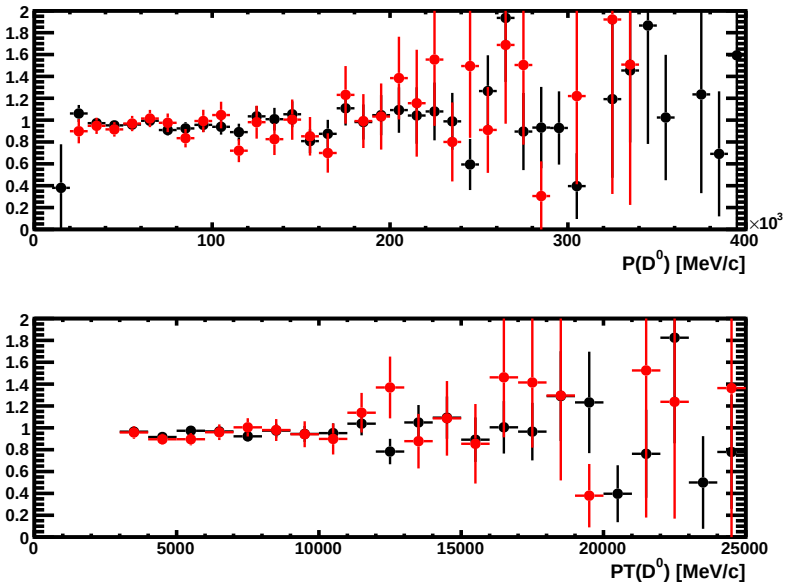


Figure 5.9: Comparison of the L0 efficiency ratio, $\epsilon_{D^0\to K^-\pi^+\rho/\omega(\to\mu^+\mu^-)}$ / $\epsilon_{D^0\to K^-\pi^+\pi^+\pi^-}$, for two criteria: `D_L0HadronTIS` (black) and `D_L0MuonTIS` (red).

termine $N^{signal}_{L0MuonTIS}$. To obtain $N^{norm}_{L0HadronTIS}$ and $N^{norm}_{L0MuonTIS}$ we used the full $D^0 \to K^-\pi^+\pi^+\pi^-$ data sample.

The results are presented in Table 5.7. One can see that MC and data agree within error bars. The uncertainty on the comparison is 8%.

5.5.4 Higher Level Trigger efficiency

The signal and normalisation mode selections both use the same HLT1 requirement: `D_Hlt1TrackAllL0DecisionTOS`, generic line to select decays which are significantly displaced from a primary vertex (Section 2.5.2). The similarity of the kinematics of the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ modes ensures a similar efficiency. Specific lines are used as far as the HLT2 is concerned: the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ selection retains candidates that are TOS on the dedicated line to select $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$, `Hlt2CharmSemilepD02KPiMuMuDecisionTOS`, while the $D^0 \to K^-\pi^+\pi^+\pi^-$ selection requires the candidates to be selected also by dedicated line, `D_Hlt2CharmHadD02HHHH_K3piDecisionTOS`.

To evaluate the uncertainty on the HLT efficiency ratio due to the finite accuracy with which it is described by the MC, we apply the same approach as in [13] and in [14]. To simulate the impact this effect, one can modify the cuts applied by the trigger lines. One way to do it is to require the selected events to be "Dec", i.e. triggered either by signal trigger (TOS) or by the trigger independent of the signal (TIS) without separating these two categories, rather than "TOS", i.e. fired by one of the signal tracks. This is equivalent to a looser selection, since not all the signal decay products are required to satisfy the line's cuts. The variation of the individual efficiencies between the default trigger requirements and this new configuration is large: 20%. However, the impact on the efficiency ratio is moderate: 4%.

Selection criteria	$\epsilon_{D^0\to K^-\pi^+\rho/\omega(\to\mu^+\mu^-)}$ (%)	$\epsilon_{D^0\to K^-\pi^+\pi^+\pi^-}$ (%)	$\frac{\epsilon_{D^0\to K^-\pi^+\rho/\omega(\to\mu^+\mu^-)}}{\epsilon_{D^0\to K^-\pi^+\pi^+\pi^-}}$
`L0HadronTIS`	20.9±0.3	21.5±0.1	0.975±0.02
`L0MuonTIS`	7.4±0.2	7.7±0.1	0.967±0.03

Table 5.6: L0 trigger efficiencies and efficiency ratio, according to the MC simulation. The trigger is applied after all the other selection criteria have been applied, apart from the HLT.

Name	$N_{L0MuonTIS}$	$N_{L0HadronTIS}$	$N_{L0MuonTIS}/N_{L0HadronTIS}$	ξ
Signal in data	1691 ± 67	3371 ± 162	0.498 ± 0.028	0.954 ± 0.055
Norm. in data	88556 ± 471	169752 ± 903	0.522 ± 0.004	
Signal in MC	460 ± 21	1331 ± 36	0.346 ± 0.018	0.979 ± 0.055
Norm. in MC	3379 ± 58	9568 ± 98	0.353 ± 0.007	

Table 5.7: Yields and yields ratios for signal and normalization modes in data and MC for the L0 cross-check.

An alternative method is to emulate the trigger lines based on online reconstructed quantities, the difference observed in MC between the trigger and offline reconstructions being a proxy for the difference in trigger reconstructed quantities between data and MC. When emulating the HLT1 trigger, the individual efficiencies double, but the efficiency ratio varies only by 4%.

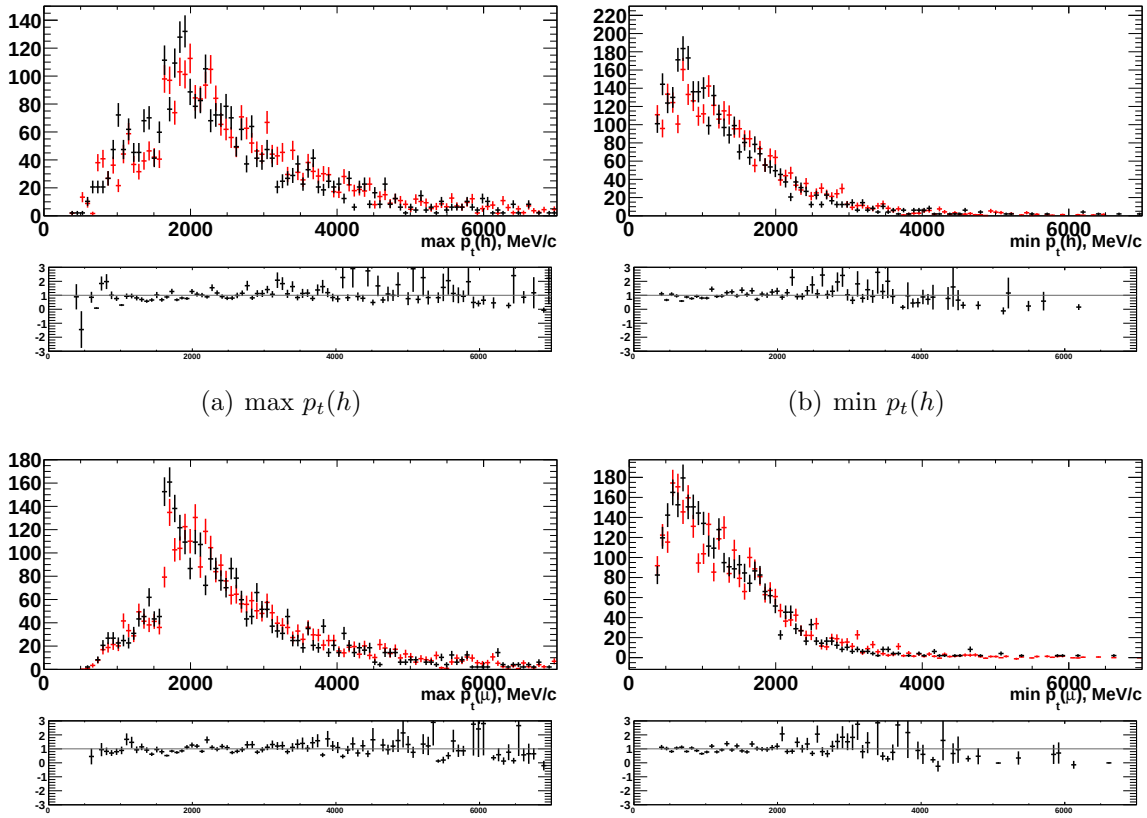
Based on these tests, we quote a systematic uncertainty of 4% on the HLT efficiency ratio.

5.5.5 Data/MC comparisons

Although better described by the simulation than those related to the L0 and the particle identification, the other quantities used in the selection are also subject to data/MC discrepancies whose impact on the efficiency ratio must be quantified.

Figures 5.10-5.12 show the data/MC comparison of these variables in the case of the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ signal. The *sPlot* technique is applied to the full sample (98%). This allows the signal distributions to be isolated. The nominal selection is applied here. Figures 5.13-5.15 show also the equivalent comparisons in the case of $D^0 \to K^-\pi^+\pi^+\pi^-$. Although not perfect, the agreement between data and MC seems sufficient to not cause a large bias on the efficiency ratio determined from MC.

These comparisons are repeated based on samples obtained with a modified selection, where the `L0HadronTIS` requirement is replaced by `L0HadronTOS||L0MuonTOS`. This improves the statistical precision of the study, and provides the same test as the comparisons performed with the nominal selection, provided no important data/MC



(a) max $p_t(h)$ (b) min $p_t(h)$

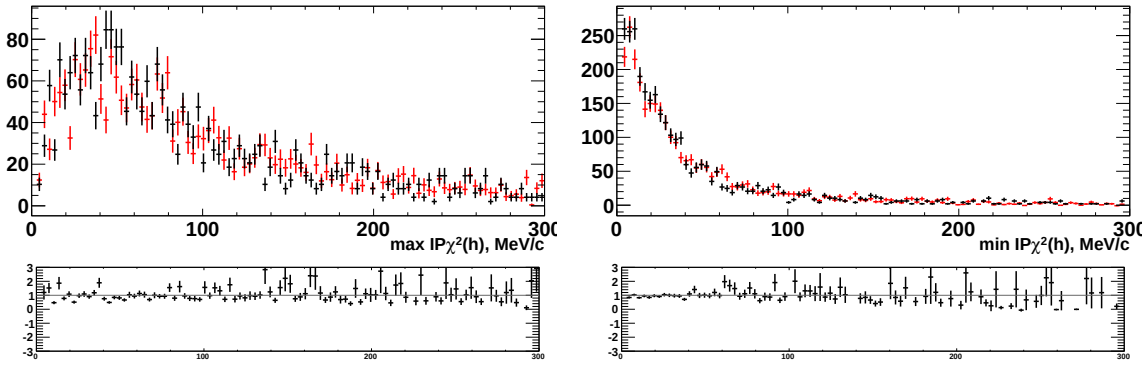
(c) max $p_t(\mu)$ (d) max $p_t(\mu)$

Figure 5.10: Momentum distributions of daughters tracks in data (red) and MC (black) in $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ with nominal selection (`L0HadronTIS`). The pull distributions show the ratio between these two sets of points.

discrepancy lies in the events removed at low transverse momentum by the L0 TOS requirements. Figures 5.16-5.18 for $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ (and Figures E.1-E.3 for $D^0 \to K^-\pi^+\pi^+\pi^-$) confirm the conclusions drawn with the nominal selection. These samples kinematically differ a bit from the nominal ones (see Appendix F), however any data-MC discrepancies are expected to be similar in magnitude. Distributions for all variables used in the BDT can be found in Appendix F. There are larger discrepancies of distributions between $D^0 \to K^-\pi^+\pi^+\pi^-$ and $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ selected with `L0HadronTOS||L0MuonTOS` trigger than with `L0HadronTIS`. This is explained by the fact that $D^0 \to K^-\pi^+\pi^+\pi^-$ events are selected mainly by `L0HadronTOS` while signal events fire `L0MuonTOS` in the majority of cases. These lines involve different p_t cuts.

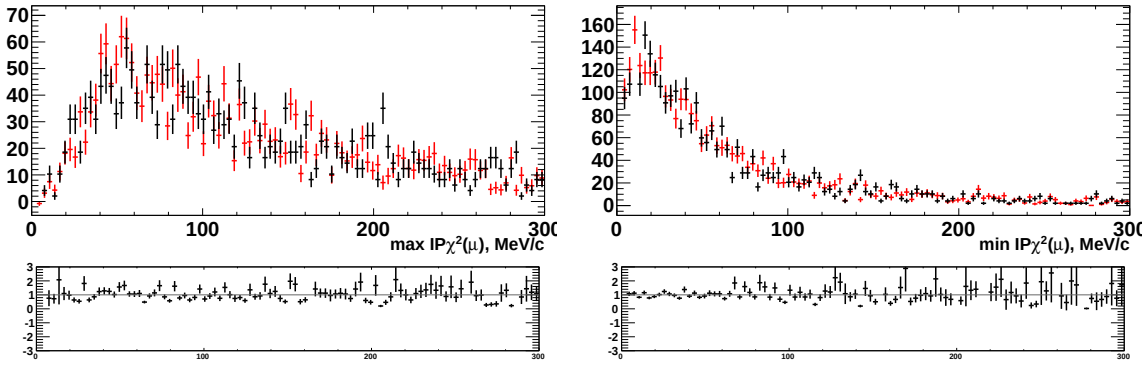
The discrepancies in the main variable distributions are also checked between $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ in MC at generation level (Figure 5.19) and just after the stripping cuts applied (Figures 5.20-5.21). In these figures one can see a confirmation of the hypothesis about similarity these two modes. The agreement is still satisfactory after the full selection is applied (Figures 5.22-5.24). The same distributions but in MC are presented in Appendix D.1.

These figures confirm that both modes are similar to a large extent. This brings



(a) max $IP\chi^2(h)$

(b) min $IP\chi^2(h)$



(c) max $IP\chi^2(\mu)$

(d) max $IP\chi^2(\mu)$

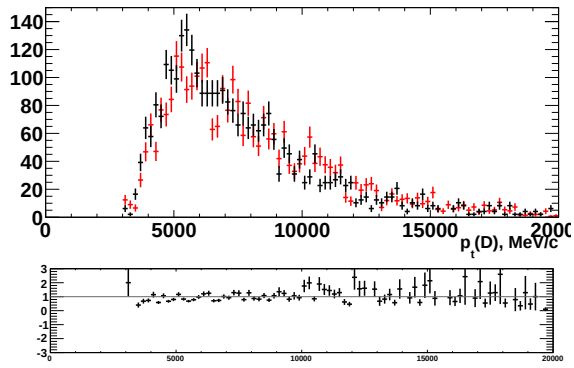
Figure 5.11: Impact parameter χ^2 distributions of daughters tracks in data (red) and MC (black) in $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ with nominal selection (`L0HadronTIS`). The pull distributions show the ratio between these two sets of points.

confidence in the assertion that the systematic uncertainties due to the data/MC limited discrepancies observed above will mostly cancel in the efficiency ratio, with one possible exception, which is uncertainties from the model used in MC, which is studied in Section 5.5.7 and taken in account.

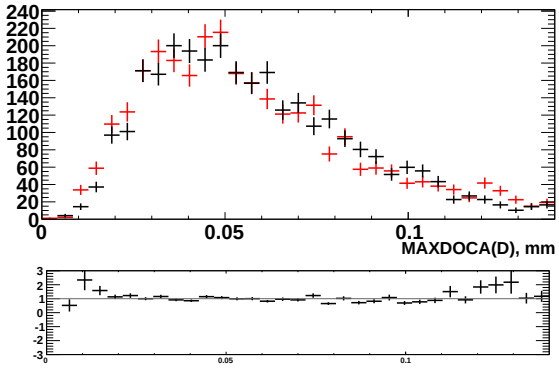
The discriminating variable produced by the BDT is a combination of all the variables studied above, and one of the main cuts in the selection. The uncertainty on the efficiency of this cut is consequently a good quantification of the effect of the data/MC discrepancies studied in this section, as well as one of the systematics we have to evaluate.

The BDT distributions for both modes are presented in Figure 5.25. One can see good agreement between data and MC points. This is confirmed by the efficiency ratio presented in Figure 5.26. The relative data/MC difference in the efficiency ratio at the level of the cut used by the analysis ($BDT > 0.2$) is less than 1%, which shows good data/MC agreement at this level. In order to quote a safe systematic uncertainty, we use the typical data/MC difference in Figure 5.26(b), in the region of the cut on

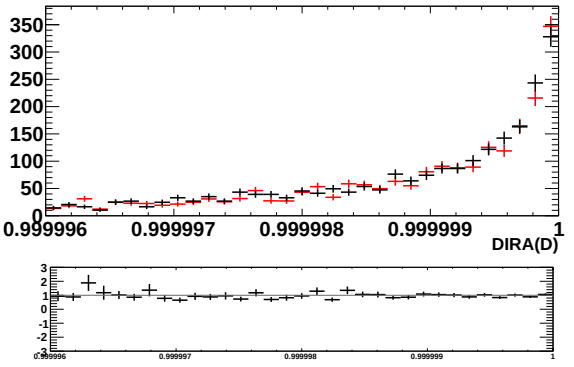
the BDT score where the systematic uncertainty is supposed to be the largest (i.e. at high values). This amounts to a 2% systematic uncertainty. Moreover, this systematic uncertainty is completed in next section by comparing smeared and unsmeared MC with data.



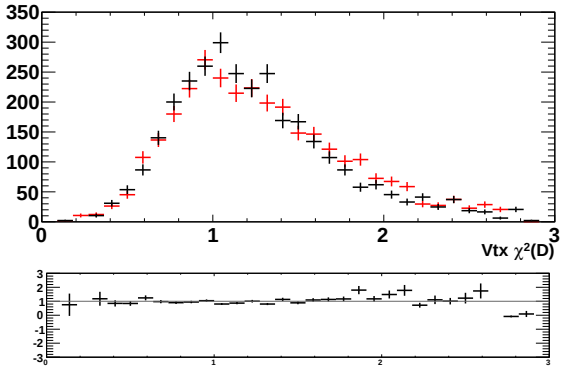
(a) $p_t(D^0)$



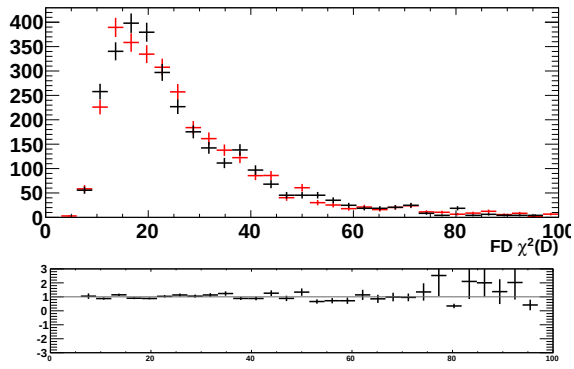
(b) $DOCA(D^0)$



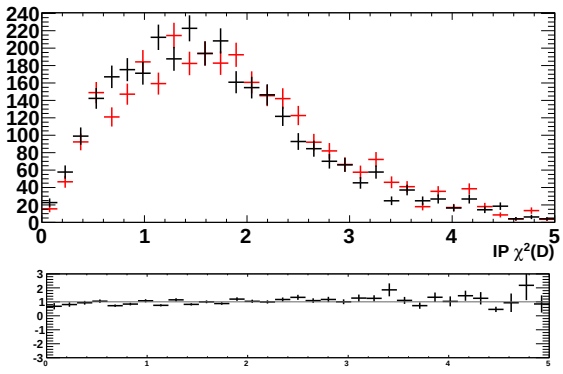
(c) $DIRA(D^0)$



(d) $Vtx\chi^2(D^0)$

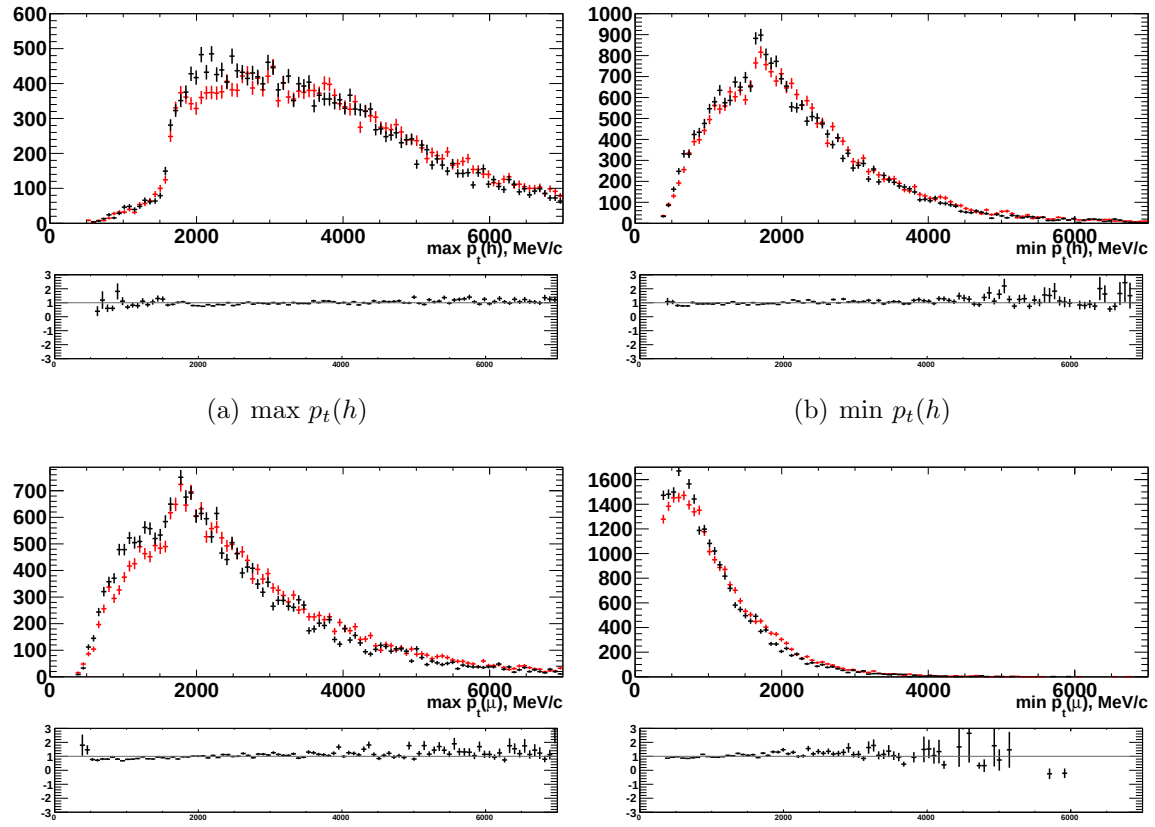


(e) $FD\chi^2(D^0)$



(f) $IP\chi^2$

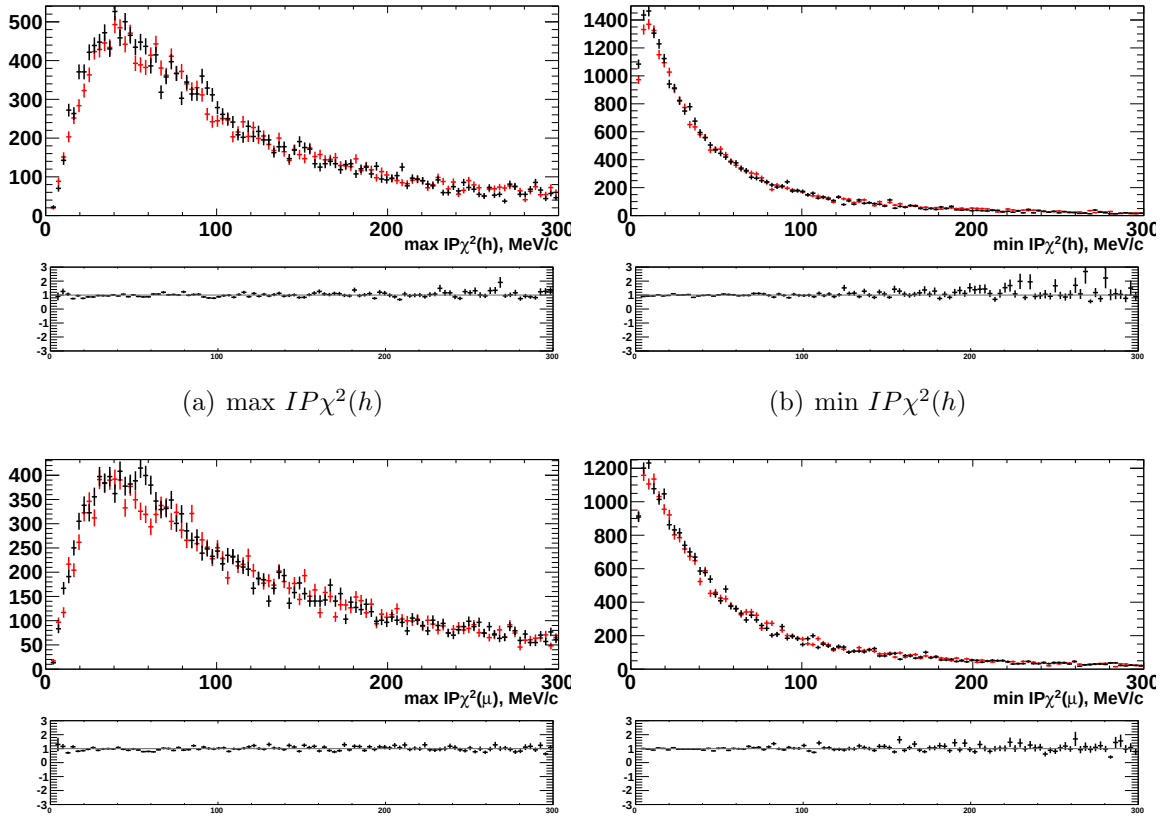
Figure 5.12: Momentum, Distance of Closest Approach (DOCA), Direction Angle (DIRA), Vertex χ^2, Flight distance χ^2 and impact parameter IPχ^2 of D^0 in data (red) and MC (black) in $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ with nominal selection (`L0HadronTIS`). The pull distributions show the ratio between these two sets of points.



(a) max $p_t(h)$ (b) min $p_t(h)$

(c) max $p_t(\mu)$ (d) max $p_t(\mu)$

Figure 5.13: Momentum distributions of daughters tracks in data (red) and MC (black) in $D^0 \to K^-\pi^+\pi^+\pi^-$ with nominal selection (`L0HadronTIS`). The pull distributions show the ratio between these two sets of points.



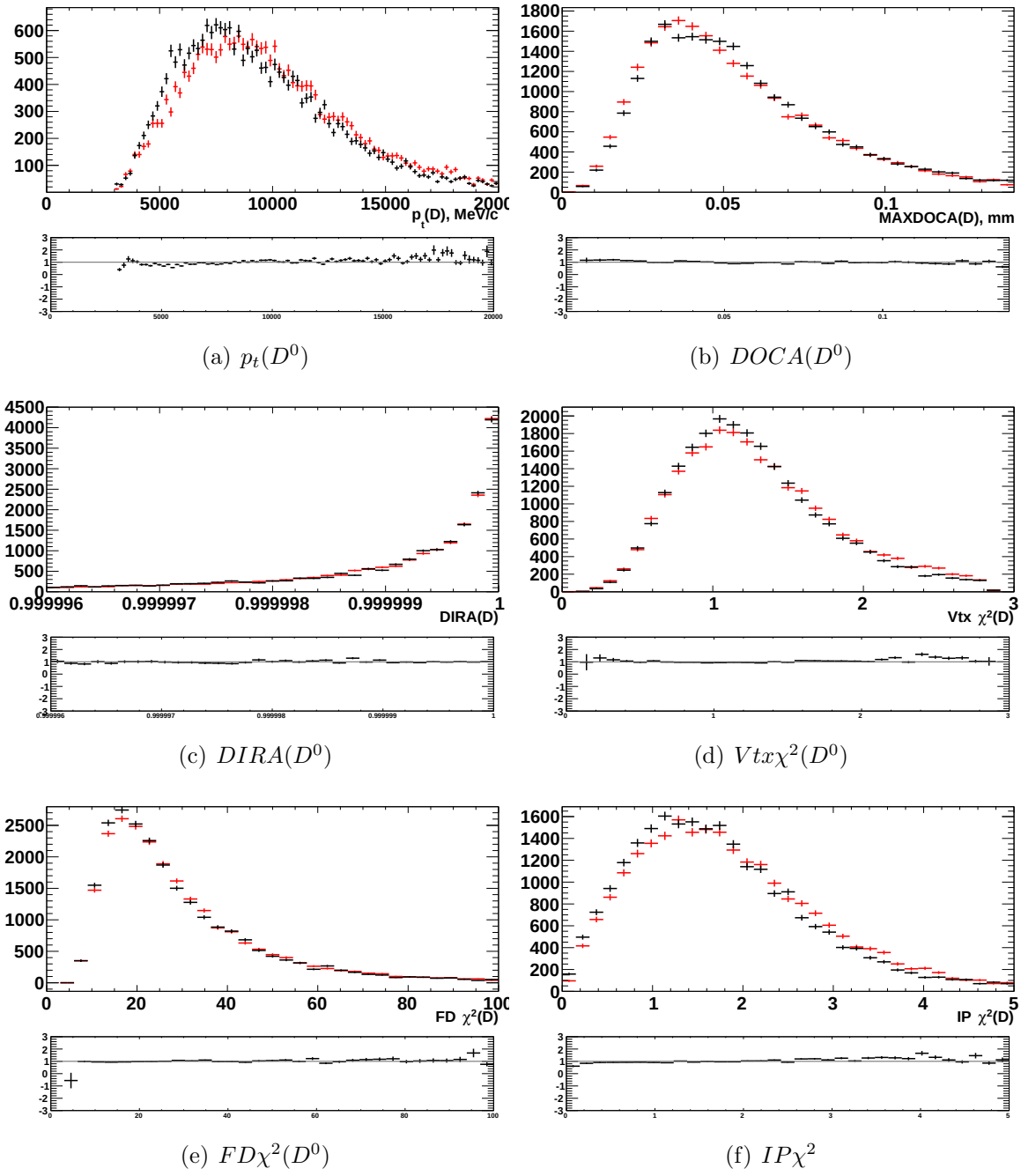
(a) max $IP\chi^2(h)$

(b) min $IP\chi^2(h)$

(c) max $IP\chi^2(\mu)$

(d) max $IP\chi^2(\mu)$

Figure 5.14: Impact parameter χ^2 distributions of daughters tracks in data (red) and MC (black) in $D^0 \to K^-\pi^+\pi^+\pi^-$ with nominal selection (`L0HadronTIS`). The pull distributions show the ratio between these two sets of points.

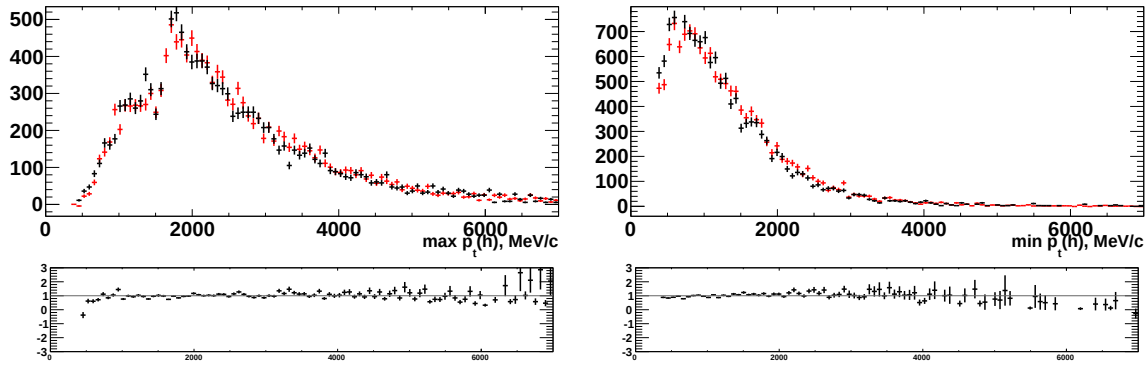


(a) $p_t(D^0)$ (b) $DOCA(D^0)$

(c) $DIRA(D^0)$ (d) $Vtx\chi^2(D^0)$

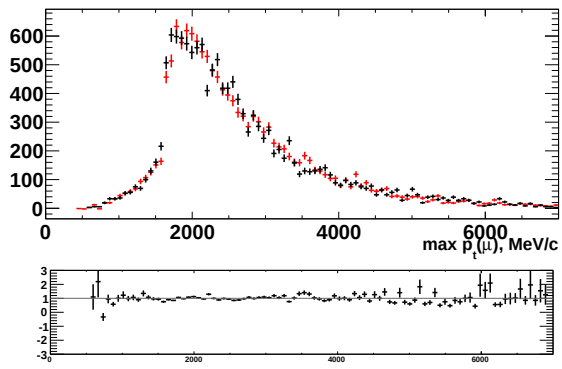
(e) $FD\chi^2(D^0)$ (f) $IP\chi^2$

Figure 5.15: Momentum, Distance of Closest Approach (DOCA), Direction Angle (DIRA), Vertex χ^2, Flight distance χ^2 and impact parameter IPχ^2 of D^0 in data (red) and MC (black) in $D^0 \to K^-\pi^+\pi^+\pi^-$ with nominal selection (`L0HadronTIS`). The pull distributions show the ratio between these two sets of points.

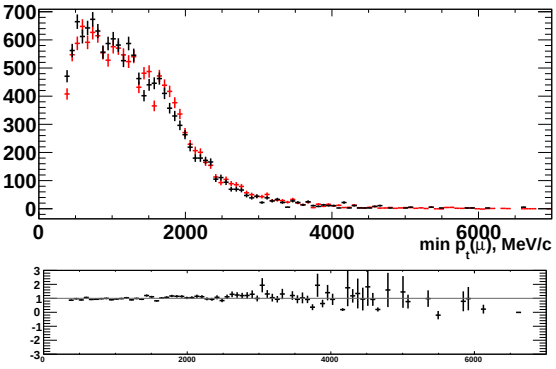


(a) max $p_t(h)$

(b) min $p_t(h)$

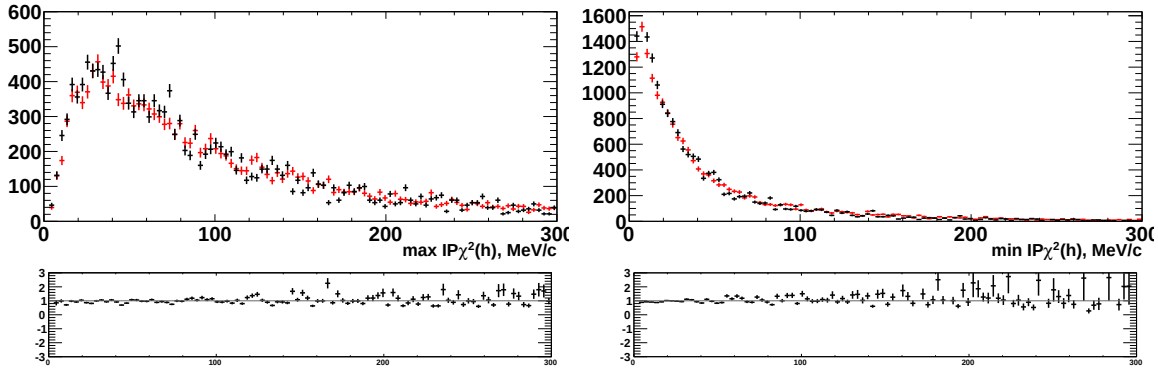


(c) max $p_t(\mu)$



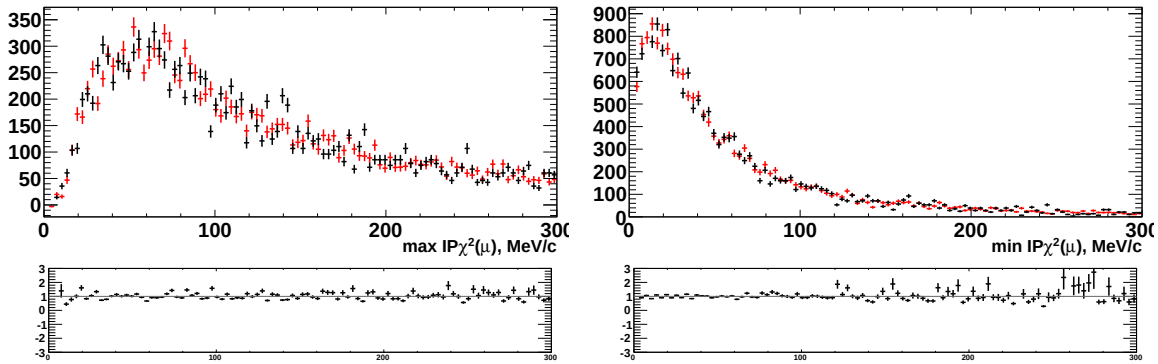
(d) max $p_t(\mu)$

Figure 5.16: Momentum distributions of daughters tracks in data (red) and MC (black) in $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ with nominal selection (`L0HadronTOS||L0MuonTOS`). The pull distributions show the ratio between these two sets of points.



(a) max $IP\chi^2(h)$

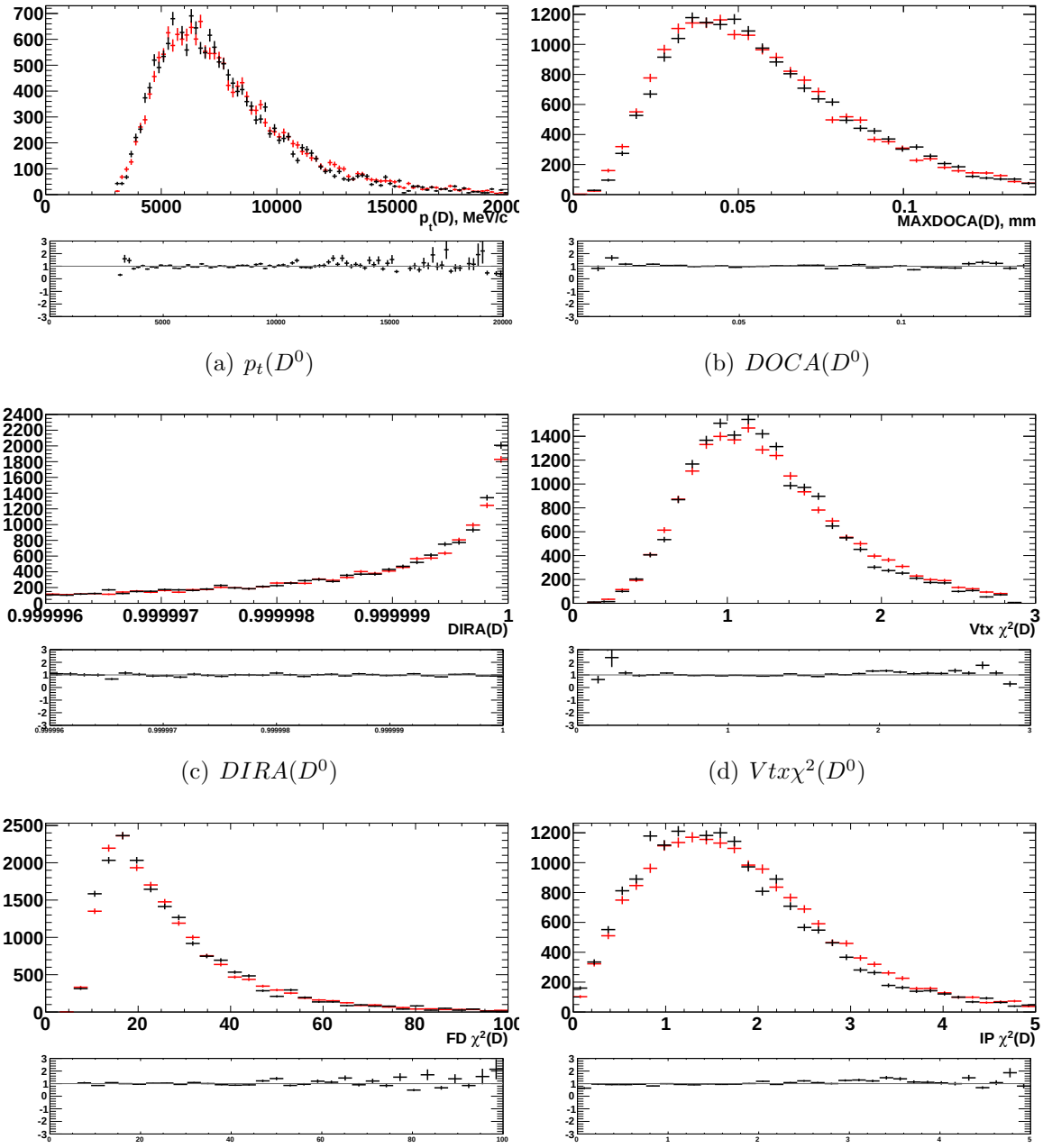
(b) min $IP\chi^2(h)$



(c) max $IP\chi^2(\mu)$

(d) max $IP\chi^2(\mu)$

Figure 5.17: Impact parameter χ^2 distributions of daughters tracks in data (red) and MC (black) in $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ with nominal selection (`L0HadronTOS||L0MuonTOS`). The pull distributions show the ratio between these two sets of points.



(a) $p_t(D^0)$ (b) $DOCA(D^0)$

(c) $DIRA(D^0)$ (d) $Vtx\chi^2(D^0)$

(e) $FD\chi^2(D^0)$ (f) $IP\chi^2$

Figure 5.18: Momentum, Distance of Closest Approach (DOCA), Direction Angle (DIRA), Vertex χ^2, Flight distance χ^2 and impact parameter IPχ^2 of D^0 in data (red) and MC (black) in $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ with nominal selection (`L0HadronTOS||L0MuonTOS`). The pull distributions show the ratio between these two sets of points.

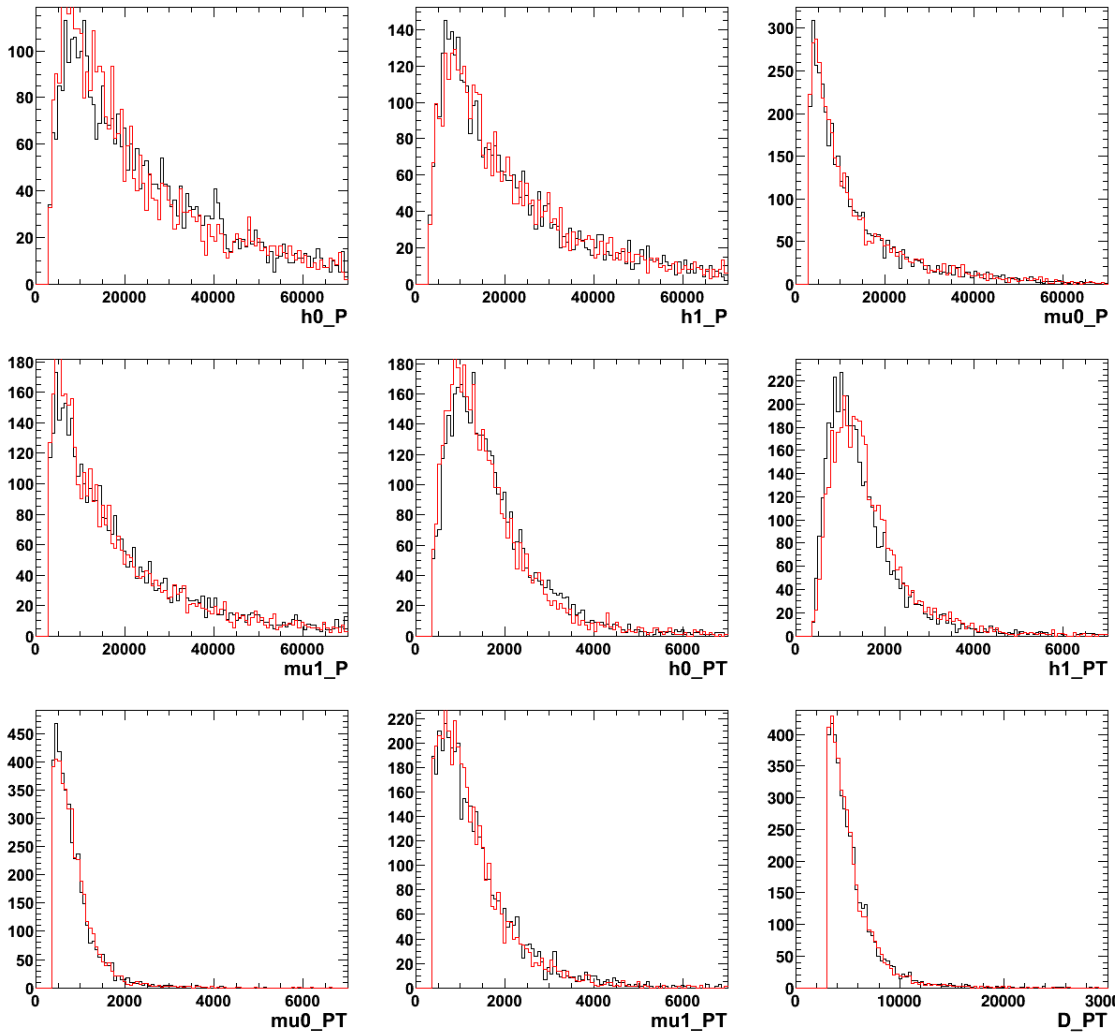


Figure 5.19: Momentum and transverse momentum distributions of daughters tracks and D^0 candidate in MC for signal (red) and normalization (black).

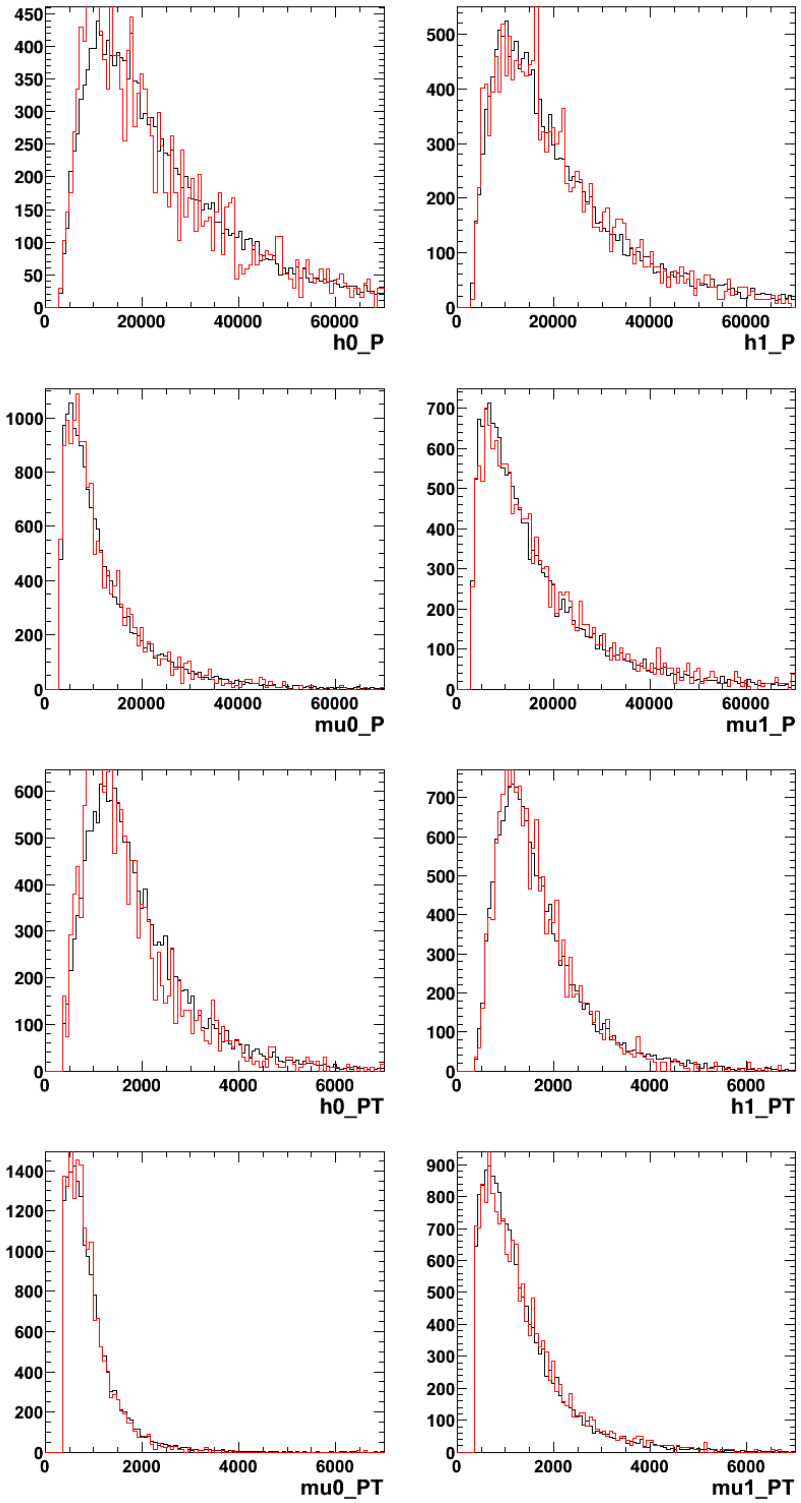


Figure 5.20: Momentum distributions of daughters tracks in MC for signal (red) and normalization (black).

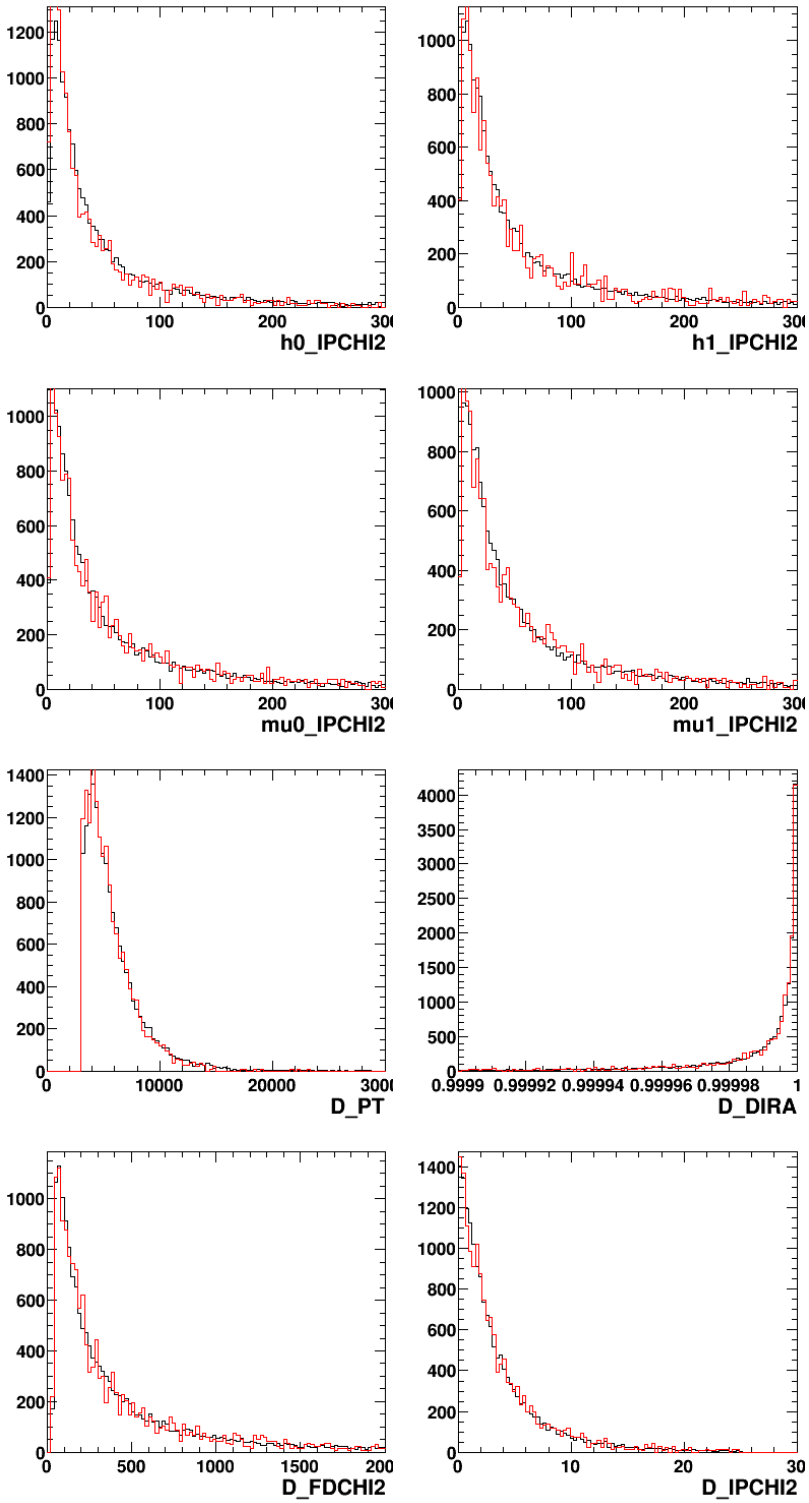
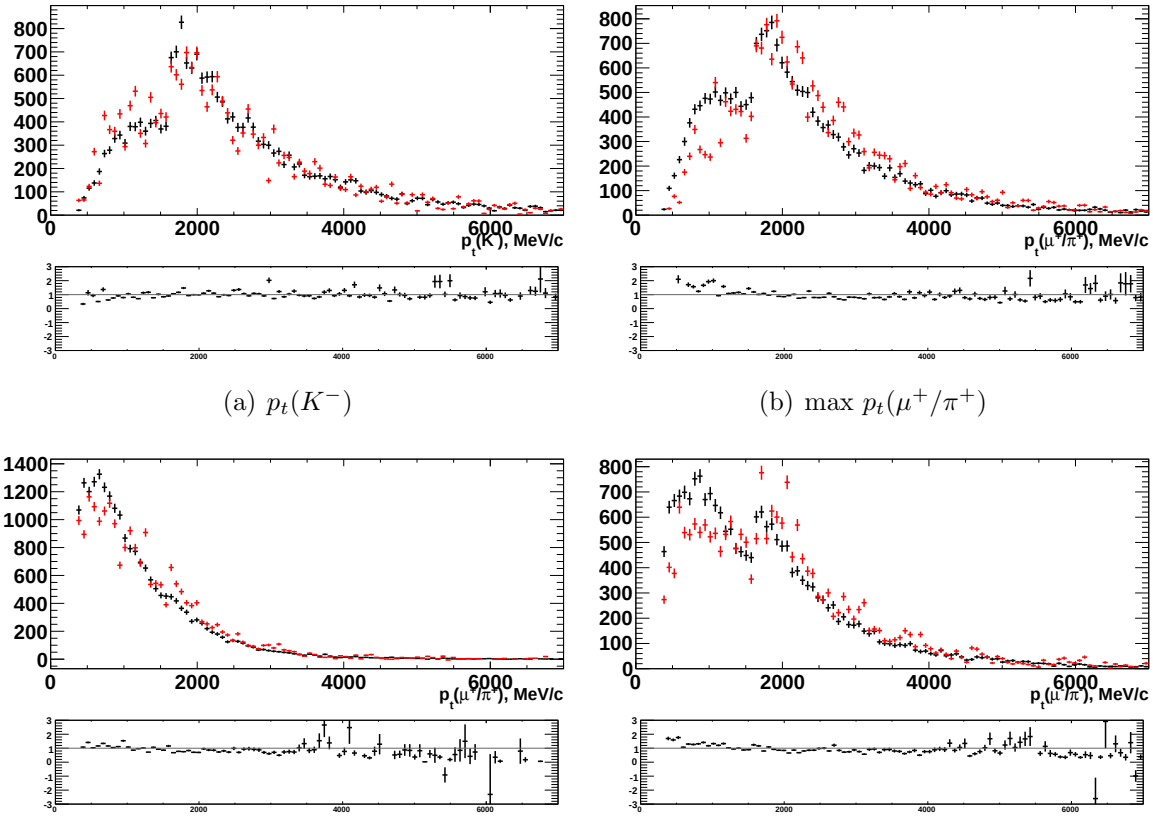


Figure 5.21: Impact parameter χ^2 distributions of daughters tracks and distributions of D^0 candidate in MC for signal (red) and normalization (black).



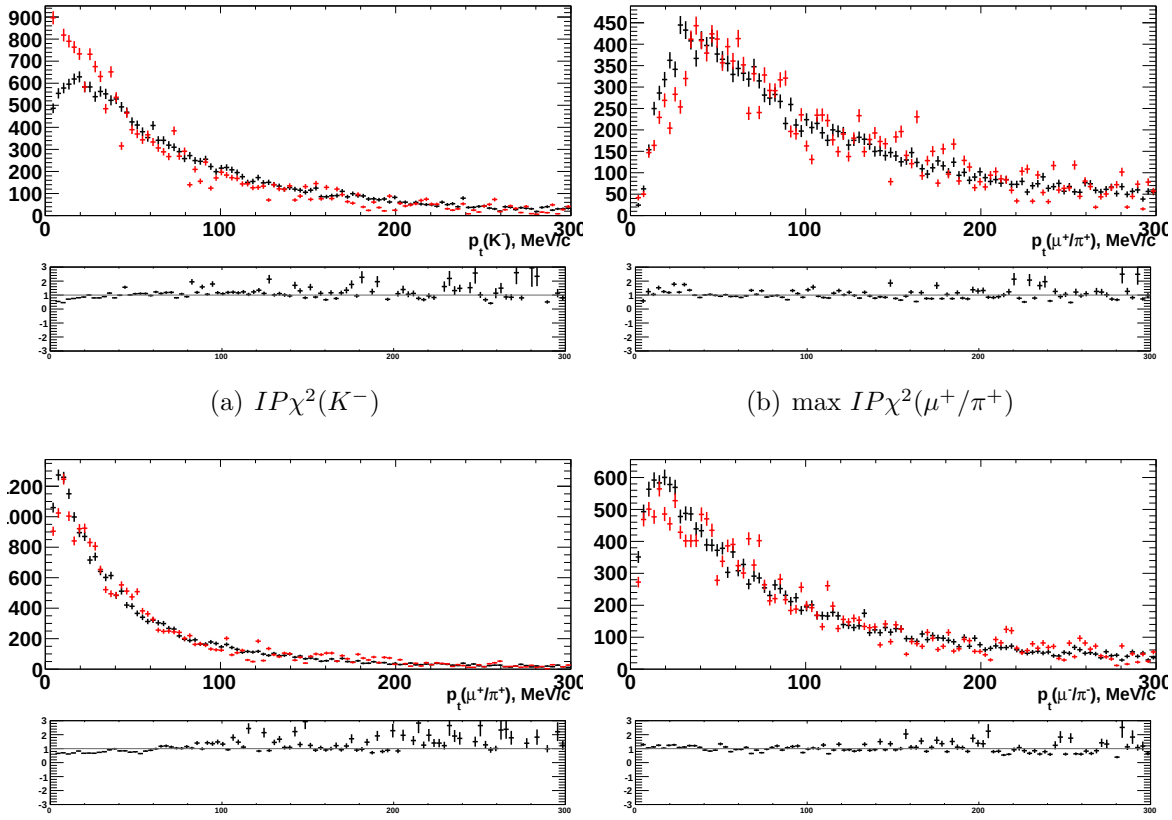
(a) $p_t(K^-)$

(b) max $p_t(\mu^+/\pi^+)$

(c) min $p_t(\mu^+/\pi^+)$

(d) $p_t(\mu^-/\pi^-)$

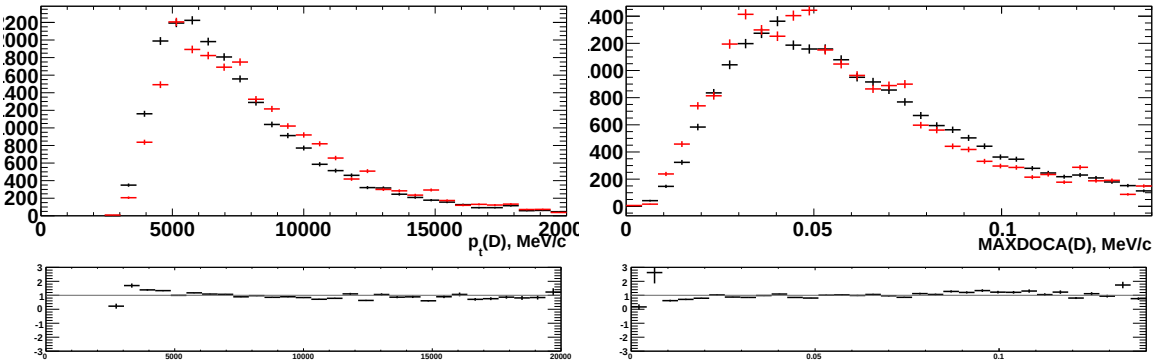
Figure 5.22: Transverse momentum distributions of daughters tracks in data for signal (red) and normalization (black). The pull distributions show the ratio between these two sets of points.



(a) $IP\chi^2(K^-)$ (b) max $IP\chi^2(\mu^+/\pi^+)$

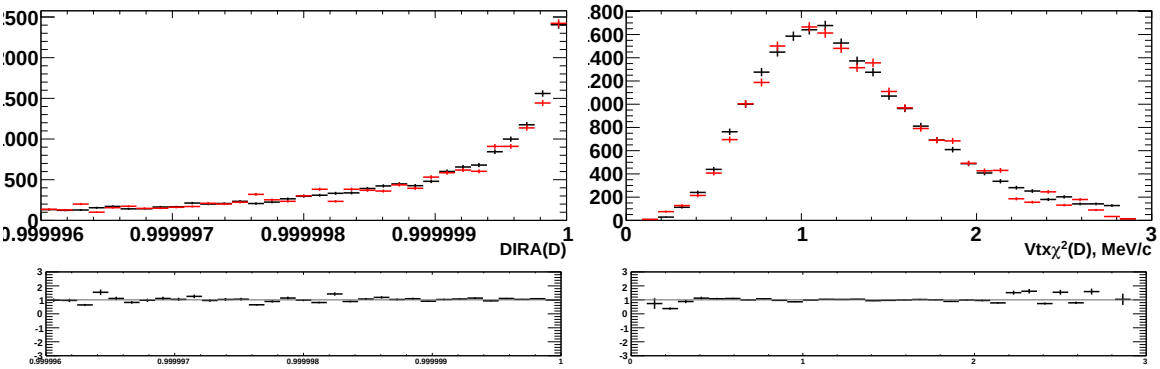
(c) min $IP\chi^2(\mu^+/\pi^+)$ (d) $IP\chi^2(\mu^-/\pi^-)$

Figure 5.23: IPχ^2 distributions of daughters tracks in data for signal (red) and normalization (black). The pull distributions show the ratio between these two sets of points.



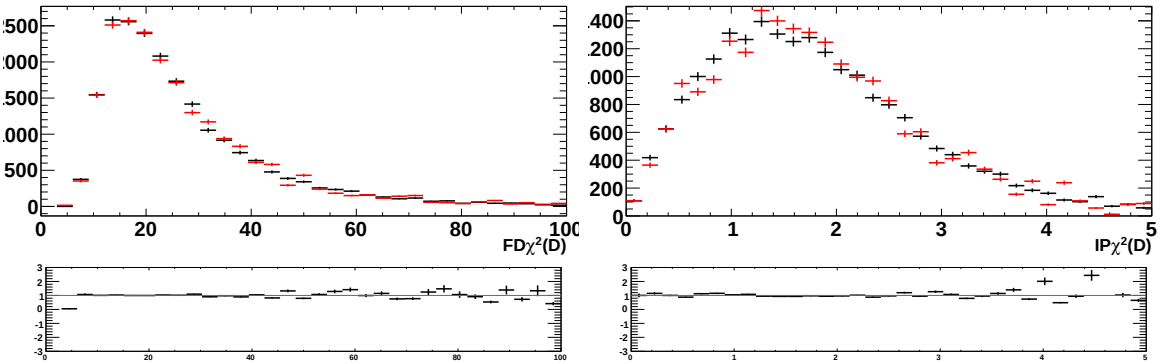
(a) $p_t(D^0)$

(b) $DOCA(D^0)$



(c) $DIRA(D^0)$

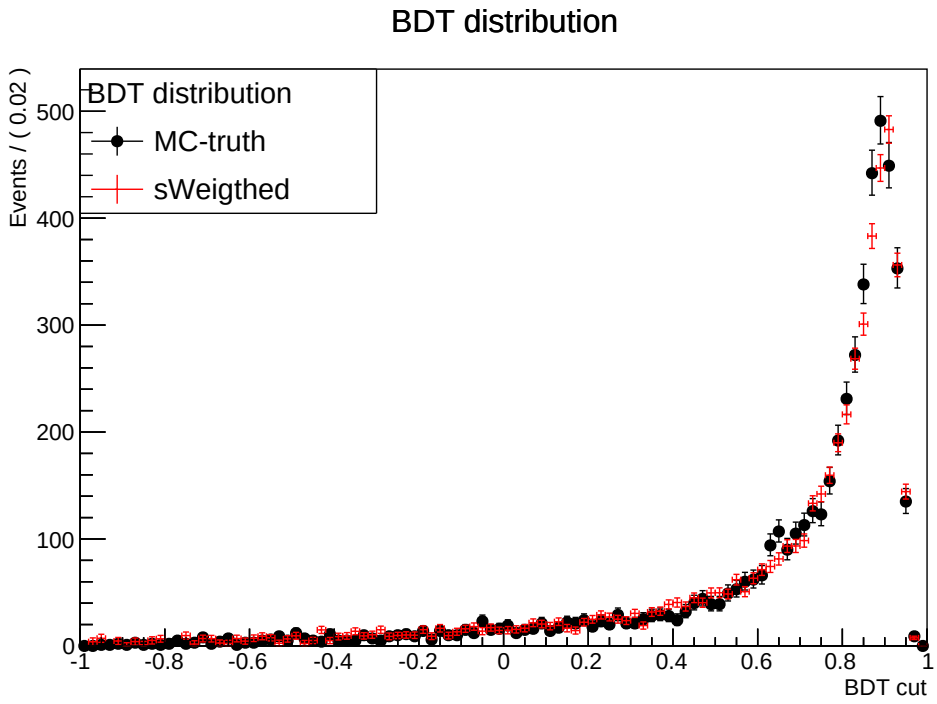
(d) $Vtx\chi^2(D^0)$



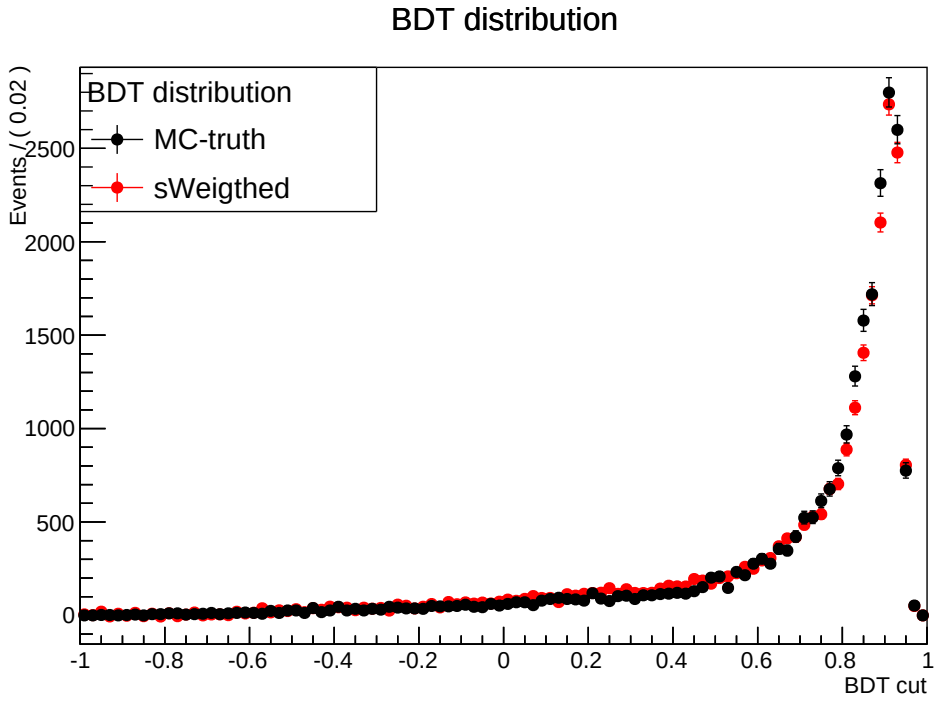
(e) $FD\chi^2(D^0)$

(f) $IP\chi^2$

Figure 5.24: Momentum, DOCA, DIRA, Vertex χ^2, Flight distance χ^2 and IPχ^2 of D^0 in data for signal (red) and normalization (black). The pull distributions show the ratio between these two sets of points.

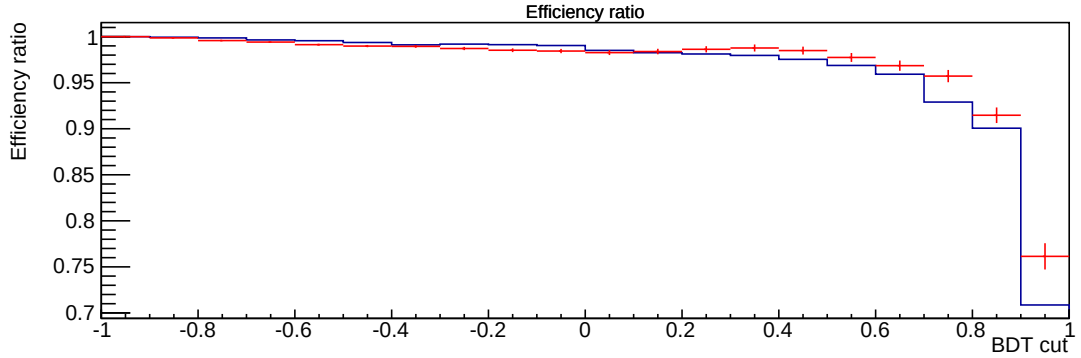


(a) $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ sample

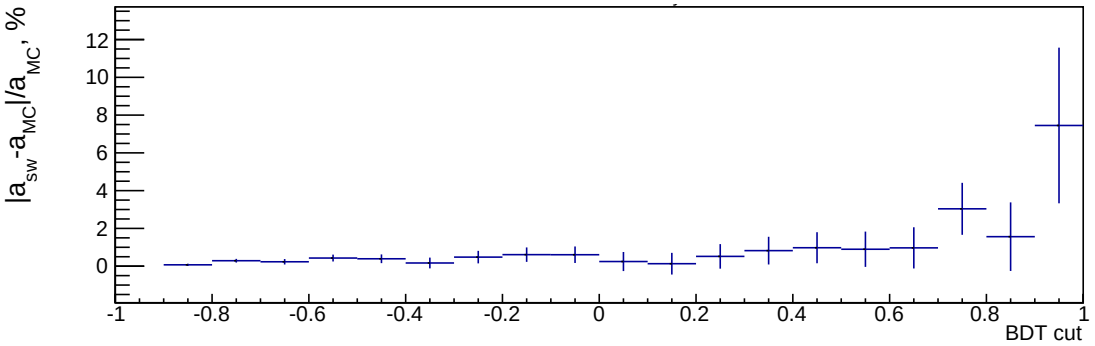


(b) $D^0 \to K^-\pi^+\pi^+\pi^-$ sample

Figure 5.25: BDT distributions in data (red) and MC (black) for the signal and normalization samples selected with `L0HadronTOS||L0MuonTOS`.



(a) Efficiency ratio in data (red), MC (blue)



(b) Difference of efficiency ratio in MC and data

Figure 5.26: Efficiency ratio in data and MC (top) and relative difference between these two ratios (bottom). This based on the samples selected with the `L0HadronTOS||L0MuonTOS` criteria. The cut used in the analysis is $BDT > 0.2$.

5.5.6 Re-evaluation of the efficiencies based on a Smeared MC

The study in Section 5.5.5 allows the impact of data-MC discrepancies on the efficiency of the selection criteria applied after the stripping to be constrained. If the variables used in this pre-selection are also affected by discrepancies, in a region that is cut away by the stripping, we have no way to study it in data. We use again the same approach that was used by analyses [13] and [14]: we re-evaluate the efficiencies based on a smeared Monte-Carlo.

The smearing actually worsens the data/MC agreement in the case of variables like vertex χ_2 and DOCA. This has been seen by many other analyses. If we used the nominal amount of smearing, it would also exaggerate the correction to the D impact parameter χ^2 and D flight distance χ^2. To not overestimate the corresponding systematic uncertainty, we choose the amount of smearing by 0.5.

Table 5.8 compares the final efficiency obtained with the default and smeared simulation. The variations due to the smearing are very limited. We conclude a systematic uncertainty of 1.5% can be quoted to account for this effect.

Selection criteria	$\epsilon_{D^0\to K^-\pi^+\rho/\omega(\to\mu^+\mu^-)}$ (%)	$\epsilon_{D^0\to K^-\pi^+\pi^+\pi^-}$ (%)	$\frac{\epsilon_{D^0\to K^-\pi^+\rho/\omega(\to\mu^+\mu^-)}}{\epsilon_{D^0\to K^-\pi^+\pi^+\pi^-}}$
Default MC	0.0099 ± 0.0003	0.0010 ± 0.0002	0.988 ± 0.033
Smeared MC	0.0096 ± 0.0003	0.0096 ± 0.0002	1.003 ± 0.034
Difference (%)	3%	4%	1.5%

Table 5.8: $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ selection efficiencies, and their ratio, obtained with the default and smeared MC, and the relative difference between them.

5.5.7 Model dependence

The kinematics of four body decays can be described by five independent variables: $m(h^+h^-)$, $m(\mu^+\mu^-)$, ϕ, θ_h and θ_l (see Section 1.2.4 for definitions). If the decay model assumed in MC does not predict the distributions of these variables correctly, and if the selection efficiency varies as a function of them, the efficiency determination will be biased.

Within the available data and MC statistics, it is difficult to do a precise data/MC comparison in five dimensions. We carry this comparison out only for the $m(K^-\pi^+)$ and $m(\mu^+\mu^-)$ masses. Figure 5.27 shows *sPlots*[9] comparing the signal's distributions

[9]With subtracted background

in data with MC distributions. Here, the `L0HadronTOS||L0MuonTOS` selection is used in order to have sufficient statistics. In Figure 5.27(a), we show the $m(K^-\pi^+)$ distribution. One can see two or three contributions: one from the $K^{*0}(892)^0$, one from another resonance, and probably one non resonant contribution. Superimposed on this figure are the $K^*(892)^0$ and non-resonant contributions present in MC, as well as their sum. There is no other resonance in MC, where the $K^*(892)^0$ and non-resonant components are given the same branching ratio. We see that the MC does not describe the data perfectly, in particular at high $m(K^-\pi^+)$. The same observations can be made on Figure 5.27(b), which displays the $m(\mu^+\mu^-)$ distribution in data and MC. The main difference in this case can be found on the right of the ω peak, where the MC overestimates the true population. We checked that the efficiency is similar in this region. This difference could be due, for instance, to the interference between the ρ and the ω.

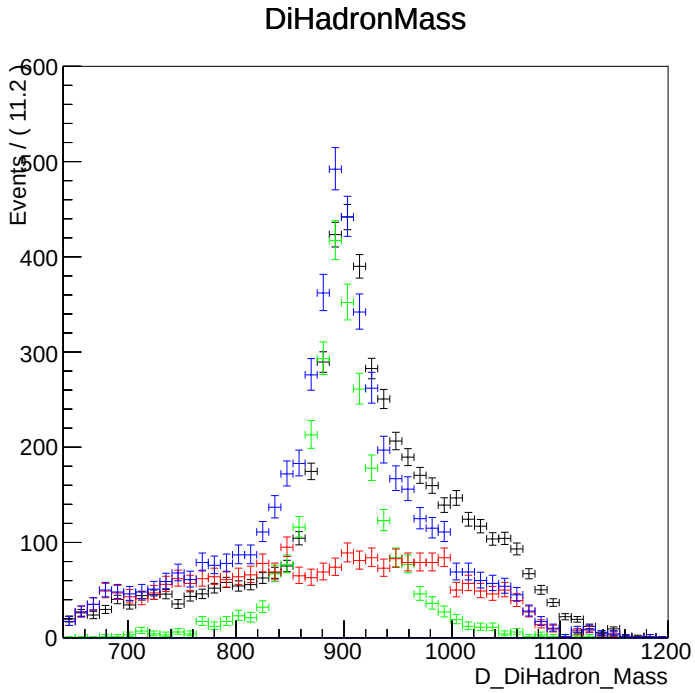
The data/MC discrepancies are also quantified across the $m(K^-\pi^+)$:$m(\mu^+\mu^-)$ plane. Seven regions have been defined:

- Region 1: $m(\mu^+\mu^-) > 800\ MeV/c^2$
- Region 2: $m(\mu^+\mu^-) < 770\ MeV/c^2$ and $m(K^-\pi^+) < 840\ MeV/c^2$
- Region 3: $m(\mu^+\mu^-) < 770\ MeV/c^2$ and $m(K^-\pi^+) > 960\ MeV/c^2$
- Region 4: $m(\mu^+\mu^-) < 770\ MeV/c^2$ and $840 < m(K^-\pi^+) < 960\ MeV/c^2$
- Region 5: $770 < m(\mu^+\mu^-) < 800\ MeV/c^2$ and $m(K^-\pi^+) < 840\ MeV/c^2$
- Region 6: $770 < m(\mu^+\mu^-) < 800\ MeV/c^2$ and $m(K^-\pi^+) > 960\ MeV/c^2$
- Region 7: $770 < m(\mu^+\mu^-) < 800\ MeV/c^2$ and $840 < m(K^-\pi^+) < 960\ MeV/c^2$

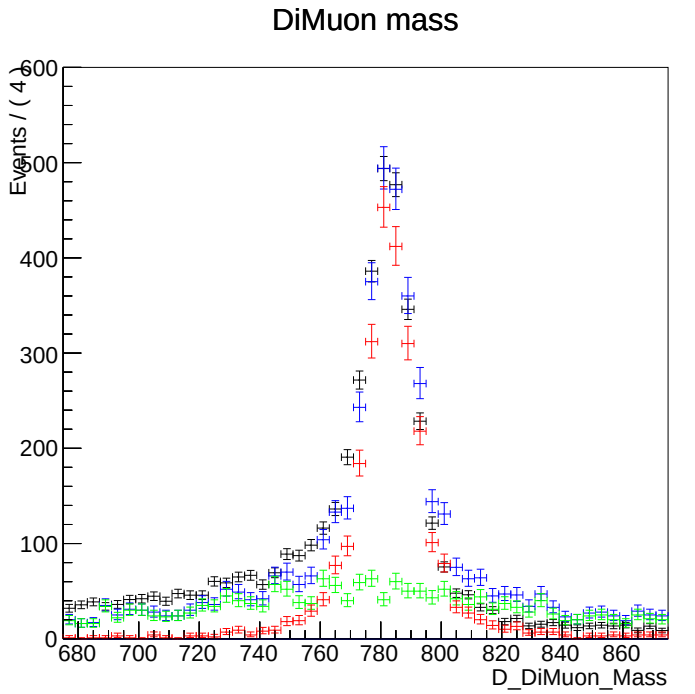
Figure 5.28 shows the fraction of signal events in each of these regions in MC and data (obtained by fitting to the data sample in each region). The ratios between these fractions in data and MC is used to reweight the MC efficiency. One can see in Figure 5.29 that the MC efficiency is quite stable across the $m(K^-\pi^+)$:$m(\mu^+\mu^-)$, which suggests a low model dependence. The total MC efficiency of the signal selection using `L0HadronTOS||L0MuonTOS` is $(0.034 \pm 0.005)\%$. It is re-calculated as the sum of the efficiencies in Figure 5.29, weighted by the ratios between the fractions of events in data and MC (Figure 5.28). The variation with respect to the default is $0.7 \pm 0.1\%$.

The exercise above accounts for data/MC differences as a function of $m(K^-\pi^+)$ and $m(\mu^+\mu^-)$ only. Discrepancies in the angular distributions could also have an impact. We assume that the differences between the true angular distributions and those generated with the decay model used in MC are not larger than the differences between the angular distributions of the $D \to VV$ components and those of the non-resonant modes. Based on this assumption, we compare the efficiency in Region 7, which is dominated by $D \to VV$ with the efficiency elsewhere. We find a difference of 1%.

The $D^0 \to K^-\pi^+\pi^+\pi^-$ efficiency is re-evaluated using a MC sample where the decay model provided by the MINT package, where resonant structure is well described, is replaced by an incoherent sum of the resonances involved in the decay, according to [40]. The efficiency ratio varies by 2% with respect with the default one.



(a) $m(K^-\pi^+)$



(b) $m(\mu^+\mu^-)$

Figure 5.27: Dihadron (a) and dimuon (b) mass distributions. Black crosses are sPlots obtained from the signal data sample selected with the `L0HadronTOS||L0MuonTOS` criterion. In (a), the green distribution is the contribution of the $K^*(892)^0$ component in MC, the red distribution shows the non resonant component, while the total distribution to be compared to the data is the blue one. In (b), the ρ and ω components are in green and red.

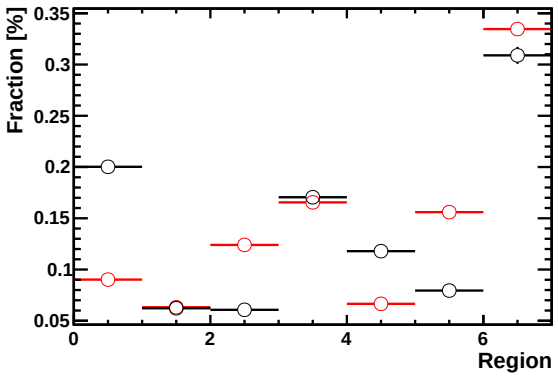


Figure 5.28: Fractions of signal events lying in the seven $m(K^-\pi^+)$:$m(\mu^+\mu^-)$ regions, in data (red) and MC (black).

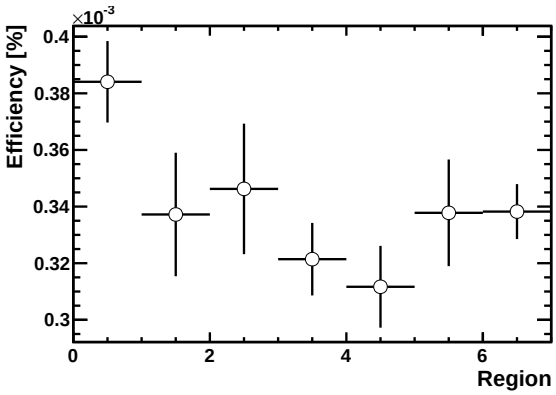


Figure 5.29: Monte Carlo efficiency in the seven $m(K^-\pi^+)$:$m(\mu^+\mu^-)$ regions. This applies to the selection using the `L0HadronTOS||L0MuonTOS` criterion.

We combine this variation with the two observed in the case of $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$. The total systematic uncertainty due to the decay models used in MC is 2.5%.

5.5.8 Systematic uncertainty due to the track reconstruction

The systematic uncertainty on the efficiency of the tracking algorithm is described in [90]. A clean sample of $J/\psi \to \mu^+\mu^-$ decays is isolated in real data using a "tag and probe" approach. The efficiency measured with the "probe" muon is compared with that observed in MC. Corrections factors are derived from this comparison. The efficiency of the tracking algorithm depends on the momentum and rapidity of the particle. The correction factors are therefore evaluated in regions of the momentum-pseudorapidity $(p;\eta)$ two-dimensional space. The MC total efficiencies are re-computed using these factors. Each $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ or $D^0 \to K^-\pi^+\pi^+\pi^-$ candidate is re-weighted by a product of four such correction factors: one per track, chosen according to the p and η of this track. This efficiency also depends on the track multiplicity in the event. Therefore, the re-weighting above is applied to MC events that have already been re-weighted against the track multiplicity in order to match real data.

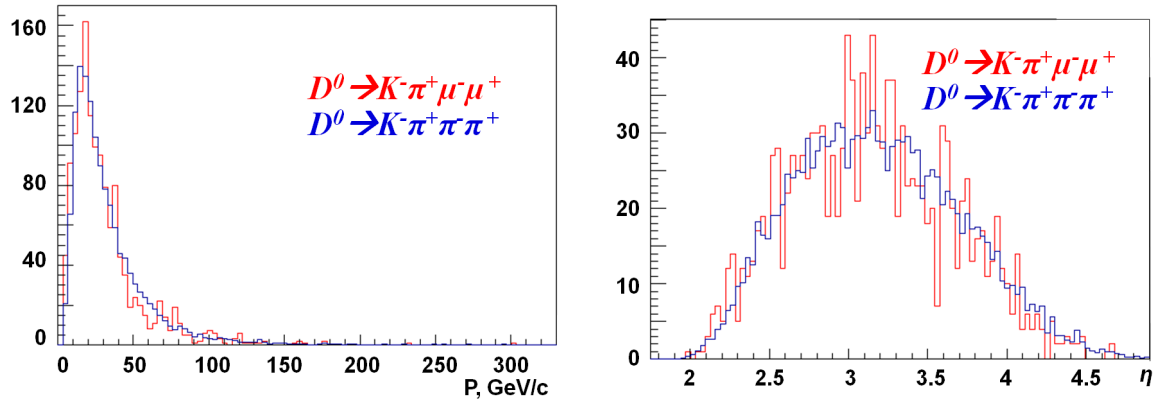
These corrections modify the total $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ MC efficiency by 4.5± 0.8%. The track by track contributions are the following: 0.5% for the kaon, 1.1% for each muon, and 1.6% for the pion. This follows the kinematics of these tracks since the data/MC discrepancy is larger at low momentum and the pion is in average the softest track in $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and the kaon the hardest. The corrections are very close the ones above in the case of $D^0 \to K^-\pi^+\pi^+\pi^-$. This is due to the fact both modes have similar kinematics. The efficiency ratio is modified by only 0.04±0.02 %.

The correction on the efficiency ratio described above appears negligible. However, it concerns particles that have reached the last stations of the tracking system. This is the efficiency of the tracking *algorithm*. It does not account for the fact that a part

of the hadrons are not reconstructed because of nuclear interactions in the detector's material. Had both $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ the same decay products, this effect would essentially vanish in the efficiency ratio. However, one has to account for the difference between a pair of muons and a pair of pions. According to [90], the uncertainties on the hadronic interactions' cross sections and on the material budget yields an uncertainty of 1.4% per pion. Therefore, the uncertainty on the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ branching ratio due to the track reconstruction is 2.8%.

5.5.9 Systematic uncertainty due to the kaon identification

The requirement on the kaon identification, $PIDK > -1.0$, is applied at the preselection or stripping level to the K^- from the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ decays. To a good accuracy, one can assume that $PIDK$ varies only as a function of the kaon kinematics and of the multiplicity of the event. In practice, the efficiency of this criteria is studied as a function of the kaon's p, η and of $nTracks$, the number of tracks in the event. Both the signal and normalization modes have very similar kinematics. This is true, in particular, in case of the kaons. Thus, this efficiency is expected to be very similar in both cases. The distributions of momentum p and pseudorapidity η of kaon candidates shown in Figure 5.30 confirm this at the MC level.

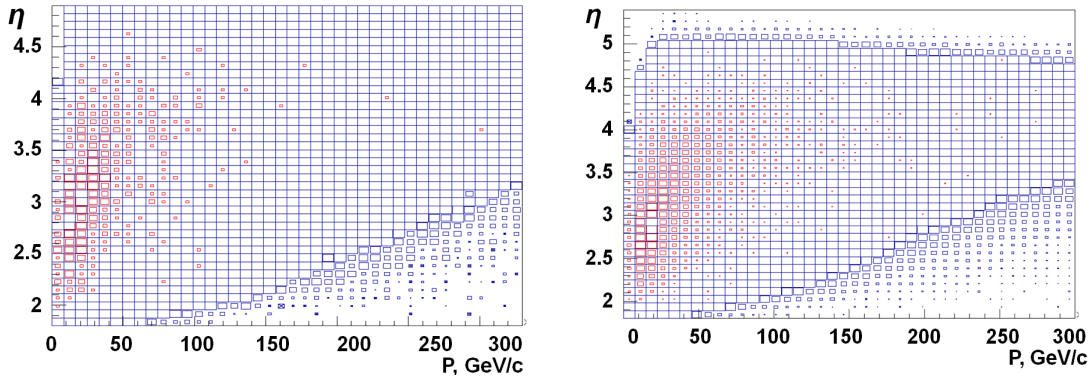


(a) Momentum distribution

(b) Pseudorapidity distribution

Figure 5.30: Kaon momentum and pseudorapidity distributions in MC for $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$(red) and $D^0 \to K^-\pi^+\pi^+\pi^-$(blue). All the cuts of the selection described in Section 5.4 are applied except for the cut on $PIDK$ of the K^-.

We evaluate the ratio of the efficiencies of the $PIDK > -1.0$ criterion in real data using the same approach as for the $ProbNNmu$ efficiencies (Section 5.5.1). A sample of $D^0 \to K^-\pi^+$ events, selected without using any particle identification criterion on the kaon, provides a large sample of unbiased calibration kaons. It is used to obtain efficiency tables as a function of $(p;\eta;nTracks)$. Several binnings are used (see below). We have checked that all the bins are populated by sufficient statistics of calibration kaons. This can be seen in Figure 5.31, which compares $(p;\eta)$ distributions for $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ decays with one for calibration decay. As a baseline we use the default binning with $(18;4;4)$ bins in $(p;\eta;nTracks)$ provided by the package and two finer binnings (twice and four times finned than



(a) $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ (b) $D^0 \to K^-\pi^+\pi^+\pi^-$

Figure 5.31: Distribution of the K^- from $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)/D^0 \to K^-\pi^+\pi^+\pi^-$(red) and calibration $D^0 \to K^-\pi^+$(blue) in the (p, η) plane.

Binning	N_{bins} $(p;\eta;nTracks)$	$\epsilon_{D^0\to K^-\pi^+\rho/\omega(\to\mu^+\mu^-)}$	$\epsilon_{D^0\to K^-\pi^+\pi^+\pi^-}$	$\frac{\epsilon_{D^0\to K^-\pi^+\rho/\omega(\to\mu^+\mu^-)}}{\epsilon_{D^0\to K^-\pi^+\pi^+\pi^-}}$
Very fine	(72;16;16)	98.8±0.3	98.89±0.05	0.999±0.002
Fine	(36;8;8)	98.1±0.2	97.96±0.07	1.001±0.002
Default	(18;4;4)	98.0±0.1	97.82±0.06	1.002±0.002
Coarse	(9;2;2)	97.8±0.1	97.71±0.05	1.001±0.002
Very coarse	(4;1;1)	97.3±0.1	97.15±0.03	1.001±0.001
MC	-	99.3±0.4	98.7±0.2	1.006±0.005

Table 5.9: Kaon efficiencies from PIDCalib.

default) and two coarser (twice and four times coarser). The obtained efficiencies and efficiency ratios, which are very close to 1, are presented in Table 5.9.

The largest difference between efficiency ratios over the binnings is $\sim 0.25\%$, while the largest difference between efficiency ratios in data and MC is $\sim 0.7\%$, which is rounded up to 1% and accounted as a systematic uncertainty die to the kaon identification.

5.6 Additional specific backgrounds

The two main backgrounds to the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ decay are the combinatorial background and the peaking background due to $D^0 \to K^-\pi^+\pi^+\pi^-$ decays. There are discussed in Sections 5.7.3 and 5.7.4.

In this section, we consider the possibility of additional significant specific backgrounds.

5.6.1 Mechanisms to produce additional backgrounds

Five mechanisms have been considered to identify the potential sources for such backgrounds. We focussed on background in which the D-mass distribution displays a peak or an endpoint situated in the fit region (see Section 5.7). Indeed, above a certain yield, it might be difficult for the fit components treating the combinatorial or the specific background to absorb these additional backgrounds, if their shape differs significantly from that used for the combinatorics.

Heavier charm hadron 5 or 6-body decays

Losing one or two pions in the decay of a heavy charm hadron can lower the invariant mass of the remaining decay products into the fit region. This effect is accentuated by $p \to \pi, \mu$ or K misidentification. Even if too much mass is lost this way, additional misidentifications can yet again bring the candidate's invariant mass back in the fit region (e.g.: $\pi \to K$).

Examples of such decays are:

- $D^* \to \pi^+ D^0 (K^-\pi^+\mu^+\mu^-, K^-\pi^+\pi^+\pi^-)$ with the slow pion replacing one of the pions from the D^0 decay.
- $\Sigma_c^0(2455) \to \Lambda_c(\to pK^-\pi^+)\pi^-$.
- $\Sigma_c^0(2520) \to \Lambda_c(\to pK^-\pi^+)\pi^-$.
- $\Sigma_c^0(2800) \to \Lambda_c(\to pK^-\pi^+)\pi^-$.
- $\Lambda_c^+(2595) \to \Lambda_c(\to pK^-\pi^+)\pi^-\pi^+$.
- $\Lambda_c^+(2765) \to \Lambda_c(\to pK^-\pi^+)\pi^-\pi^+$.
- $\Lambda_c^+(2880) \to \Lambda_c(\to pK^-\pi^+)\pi^-\pi^+$.

Semileptonic decays

A semileptonic D decays with four charged tracks in the final state can be reconstructed as a signal decay. The loss of the neutrino will cause the mass of these candidates to follow a distribution with an endpoint at the mass of the decaying particle. This endpoint will be shifted if there are more than four tracks in the final state. It could however be found in the fit region if, for instance, one of the π's used to build the candidate is mis-identified as a K.

A priori, the most dangerous semileptonic decays is $D^0 \to K^-\pi^+\pi^-\mu^+\nu_\mu$.

Radiative decays

The most dangerous potential backgrounds involving a photon are a priori $D^0 \to K^-\pi^+\eta'(\to \pi^+\pi^-\gamma)$ and $D^0 \to K^-\pi^+\eta(\to \pi^+\pi^-\gamma)$. Indeed, once the photon is lost, the decays can be selected as easily as $D^0 \to K^-\pi^+\pi^+\pi^-$. Although their branching ratio should guarantee a far lower yield, it's not obvious a priori that these background will actually be negligible. We quantify this yield in Section 5.6.2.

D^0 and $D^+_{(s)}$ 5 or 6 body decays

Signal candidates can be made from such decays if one or two pions are lost. For a fraction of them, the $m_{K^-\pi^+\mu^+\mu^-}$ invariant lies in the fit region due to the $\pi \to K$ misidentification it takes to form such a candidate.

We list below the modes with the highest production rate relative to $D^0 \to K^-\pi^+\pi^+\pi^-$ (combining branching ratios from [40] and cross sections from [17]):

- $D^0 \to \pi^+\pi^-\pi^+\pi^-\pi^0$ ($\mathcal{R}^{prod}_{D^0\to\pi^+\pi^-\pi^+\pi^-\pi^0/K3\pi} \sim 1/20$).
- $D^+ \to \pi^+\pi^-\pi^+\pi^-\pi^+$ ($\mathcal{R}^{prod}_{D^+\to\pi^+\pi^-\pi^+\pi^-\pi^+/K3\pi} \sim 1/125$).
- $D^+_s \to \pi^+\pi^+\pi^+\pi^-\pi^-\pi^0$ ($\mathcal{R}^{prod}_{D^+_s\to\pi^+\pi^+\pi^+\pi^-\pi^-\pi^0/K3\pi} \sim 1/13$).

The production rates of these modes are at least an order of magnitude below that of $D^0 \to K^-\pi^+\pi^+\pi^-$. The ratio of efficiencies should be more than 10 times lower. Indeed, to form a signal candidate, the decays above need one additional misidentification: $\pi \to K$. Moreover, cuts on the direction angle ($DIRA$) or the impact parameter χ^2. of the D^0 should be less efficient than when they are applied to $D^0 \to K^-\pi^+\pi^+\pi^-$. Consequently, we expect less than 10 such event in the signal sample. We consider this source of background to be negligible.

τ decays

The only τ decay with a final state that is a potential source of background is $\tau \to 3\pi^-2\pi^+\nu$. Losing the neutrino and one of the pions leaves four tracks which invariant mass can be consistent with the D^0's provided one pion is misidentified as a kaon. Any other more complicated scenario should yield a lower efficiency.

The τ production cross section in the acceptance of the detector is of the order of 80 μb [91]. Moreover, $\mathcal{B}(\tau \to 3\pi^-2\pi^+\nu_\tau) = (1.02 \pm 0.04) \times 10^{-3}$ [40]. This means this decay is about 800 times less frequent than $D^0 \to K^-\pi^+\pi^+\pi^-$. We conclude it should provided a negligible contribution to the signal sample.

5.6.2 Quantification of the highest background yields

The decays for which we expect a priori the largest yields have been studied using generation level MC production. The following cuts, which are used by the selection described in Section 5.4, can be easily applied on generator level quantities:

- All the tracks used to form a signal candidate have to be in the acceptance of the detector [10].
- All the tracks used to form a signal candidate have to satisfy $p_t > 350$ MeV/c^2 and $p > 3000$ MeV/c^2.
- The signal candidate has to satisfy $p_t(K^-\pi^+\mu^+\mu^-) > 3000$ MeV/c^2.

[10] For LHCb experts: the same criteria as "DaughterInLHCb" at Gauss level

- The invariant mass of the dimuon pair involved by the candidate must fulfil $675 < m_{\mu\mu} < 875$ MeV/c^2.
- The cosine of the angle between the momentum of the D^0 candidate and the line joining the primary and secondary vertex has to satisfy the direction angle criterion, $DIRA > 0.9999$.
- The distance between the primary vertex and the decay vertex of the D^0 candidate has to satisfy the flight distance criterion, $FD > 1$ mm. Although this cut is not explicitly applied by the signal selection (see Section 5.4), we observe in full MC $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ samples all truth-matched candidates happen to satisfy it.

Decay mode	$\mathcal{B}$ [40]	Misidentification	Yield
$D^0 \to K^-\pi^+\pi^-\mu^+\nu_\mu$	$2.8^{+1.4}_{-1.1} \times 10^{-4}$	$\nu_\mu \to$ lost; $\pi^- \to \mu^-$	10 ± 5
$D^* \to \pi^+_{slow} D^0(K^-\pi^+\mu^+\mu^-, K^-\pi^+\pi^+\pi^-)$	$(\sim 6.7 \times 10^{-7}, 5.4 \pm 0.1\%)$	$\pi^+_{slow} \to \pi^+$; $\pi^+ \to$ lost	15 ± 2
$D^0 \to K^-\pi^+\eta'(\to \pi^+\pi^-\gamma)$	$(2.2 \pm 0.6) \times 10^{-3}$	$\gamma \to$lost; $\pi \to \mu$	< 1
$\Sigma_c^0(2455) \to \Lambda_c(\to pK^-\pi^+)\pi^-$	$(5.0 \pm 1.3)\%$	8 combinations of $[p, K^-, \pi^+, \pi^-] \to [K^-\pi^+\mu^+\mu^-]$	negligible
$\Lambda_c^+(2595) \to \Lambda_c(\to pK^-\pi^+)\pi^-\pi^+$	$(3.4 \pm 1.0)\%$	$\pi^+ \to$lost; 8 combinations of $[p, K^-, \pi^+, \pi^-] \to [K^-\pi^+\mu^+\mu^-]$	negligible

Table 5.10: List of the studied specific backgrounds.

We applied these cuts to each of the generator level samples produced above to determine the efficiency with which they select each of the potential backgrounds, relative to the efficiency with which the same cuts select $D^0 \to K^-\pi^+\pi^+\pi^-$ decays, for which a generator level sample was also produced. Combined with the evaluation of the production rate ratio and with the $D^0 \to K^-\pi^+\pi^+\pi^-$ yield observed in the real data signal sample, this allows the yields of such backgrounds in this sample to be evaluated or constrained.

A slightly different procedure is applied in the case of $D^* \to \pi^+_{slow} D^0(K^-\pi^+\mu^+\mu^-, K^-\pi^+\pi^+\pi^-)$, where a full MC sample happened to be available.

The result of the studies of potentially dangerous modes is present in Table 5.10. The non negligible yields may be expected from $D^0 \to K^-\pi^+\pi^-\mu^+\nu_\mu$ (10 ± 5 events) and $D^* \to \pi^+_{slow} D^0(K^-\pi^+\mu^+\mu^-, K^-\pi^+\pi^+\pi^-)$ (15 ± 2 events). It was studied that these decays can not create peaking structures in the mass fit window and are consequently

absorbed by the combinatorial background contribution in the fit described in next section (Section 5.7). Moreover, the bias it can cause to the signal extraction is dominated by the other uncertainties affecting this procedure, such as the description of the tails of the $D^0 \to K^-\pi^+\pi^+\pi^-$ peaking background. We do not perform further treatment of this issue.

5.7 Fit for the measurement of the branching fraction

To measure $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$, we perform a simultaneous D^0 mass fit to two samples obtained by applying the signal and $D^0 \to K^-\pi^+\pi^+\pi^-$ selections described in Section 5.4. The first is the signal sample based on 78% of the 2012 data set. The second one is the normalization sample based on 50% of the data set. We describe this procedure in this section.

5.7.1 Global structure

The fit model assumes that the signal sample is made of three types of events:

- The signal, i. e. true $D^0 \to K^-\pi^+\mu^+\mu^-$ events.
- The peaking background, or "punch through" background, made of $D^0 \to K^-\pi^+\pi^+\pi^-$ decays that satisfied the selection criteria due to the misidentification of two pions as muons.
- The combinatorial background, which forms a pedestal below the two other event types.

A simpler composition is assumed for the normalization sample: a large peak of $D^0 \to K^-\pi^+\pi^+\pi^-$ candidates on top of a combinatorial background. No specific background is considered in this case.

Table 5.11 gives an overview of the fit model. Each of the three contributions listed above is described by a dedicated probability density function (PDF). As will be described in more detail in Sections 5.7.2, 5.7.3 and 5.7.4, both the signal peak and the $D^0 \to K^-\pi^+\pi^+\pi^-$ peak in the normalization sample are described using the "Cruijff" PDF [92]. The combinatorial background is described by a first order Chebyshev polynomial. The $D^0 \to K^-\pi^+\pi^+\pi^-$ peaking background is accounted for with the help of non-parametric kernel estimation functions, RooKeysPDFs. The fit is simultaneous in the sense that some of the parameters of the Cruijff function are assumed to be the same for $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$, namely the width σ of the peaks, as well as the parameter describing the small tail extending above the mass peak: α_R. The choice of a common σ is dictated by the presence below the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ mass peak of the peak due to doubly misidentified $D^0 \to K^-\pi^+\pi^+\pi^-$ decays. They are separated by only ~ 20 MeV/c^2. This makes the determination of the signal's σ difficult if only the statistics of the signal sample can

be used. The peaking background also overlaps with the left tail of the signal peak ($\alpha_L^{K\pi\mu\mu}$). This parameter is consequently also difficult to determine. Here, we can't assume this tail is the same as in the normalisation mode, since the final state pions from $D^0 \to K^-\pi^+\pi^+\pi^-$ sometimes decay in flight ($\pi \to \mu\nu_\mu$), causing a larger tail on the left of the $m_{K^-\pi^+\pi^+\pi^-}$ peak. For that reason, $\alpha_L^{K\pi\mu\mu}$ is not allowed to float in the fit. It is fixed at the value found in the MC signal sample (Table 5.12).

The signal yield measured by the fit is automatically translated into a branching ratio by the fitter, that is provided with the efficiency ratio determined in Section 5.5 and with the $D^0 \to K^-\pi^+\pi^+\pi^-$ branching ratio.

Components	$D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$	$D^0 \to K^-\pi^+\pi^+\pi^-$
Signal Peak	$m_{D^0}^{K\pi\mu\mu}$, $[\sigma]$, $(\alpha_L^{K\pi\mu\mu})$, $[\alpha_R]$, $\mathcal{B} = \frac{N^{K\pi\mu\mu}}{N^{K\pi\pi\pi}} \frac{\epsilon^{K\pi\pi\pi}}{\epsilon^{K\pi\mu\mu}} \mathcal{B}_{D\to K\pi\pi\pi}$	$m_{D^0}^{K\pi\pi\pi}$, $[\sigma]$, $\alpha_L^{K\pi\pi\pi}$, $[\alpha_R]$ $N^{K\pi\pi\pi}$
Peaking Background	$N_{K\pi\pi\pi-misID}^{K\pi\mu\mu}$, $f^{low\ P_T}$	
Combinatorial Background	$a_{Chebyshev}^{K\pi\mu\mu}$, $N_{combi}^{K\pi\mu\mu}$	$a_{Chebyshev}^{K\pi\pi\pi}$, $N_{combi}^{K\pi\pi\pi}$

Table 5.11: Overview of the fit model. Parameters between square brackets are fitted simultaneously to the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ samples. The one between parentheses is fixed to the value found by the fit to the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ MC sample.

5.7.2 Parametrisation of peaking structures

The Cruijff function is used to describe both the shape of the signal peak and that of $D^0 \to K^-\pi^+\pi^+\pi^-$ in the normalisation sample. It is defined as follows:

$$f(m) \propto exp\left(\frac{-(m-m_{D^0})^2}{2\sigma^2 + (m-m_{D^0})^2 \times \alpha_{L/R}}\right). \tag{5.4}$$

This is a Gaussian-like distribution of mean m_{D^0} and width σ, where α_L ($m < m_{D^0}$) and α_R ($m > m_{D^0}$) parametrise the tails.

We checked this parametrisation is indeed adequate by fitting three samples, to which the nominal signal and normalization selections have been applied:

- MC signal sample, retaining only truth-matched candidates (Figure 5.33).

- MC $D^0 \to K^-\pi^+\pi^+\pi^-$ sample, retaining only truth-matched candidates (Figure 5.32).
- $D^0 \to K^-\pi^+\pi^+\pi^-$ data sample (Figure 5.35).

Figures 5.33, 5.32 and 5.35 show that this parametrisation can provide a satisfactory description of signal and $D^0 \to K^-\pi^+\pi^+\pi^-$ distributions.

A fourth sample was also used: the signal sample where the `L0HadronTIS` criteria is replaced by `L0HadronTOS||L0MuonTOS` (Figures 5.34 and 5.36). Although the kinematics of the signal events in these two samples is slightly different (see Appendix F), the impact on the shape of the signal fit seems be limited, as can be seen by comparing Figures 5.33 and 5.34. Thus, the sets of parameter values provided by these latter fits allow alternative fits to be performed, as cross checks, or to evaluate systematic uncertainties.

Table 5.12 presents the values of the fitted parameters for each of these fits.

Cruijff Parameters	Signal	Normalisation	Signal L0HadronTOS||L0MuonTOS
Monte Carlo			
m_{D^0}, MeV/c^2	1866.2 ± 0.2	1865.49 ± 0.08	1865.9 ± 0.1
σ, MeV/c^2	6.5 ± 0.2	6.15 ± 0.07	6.4 ± 0.1
α_L	0.11 ± 0.01	0.124 ± 0.004	0.102 ± 0.006
α_R	0.09 ± 0.01	0.107 ± 0.004	0.096 ± 0.007
Data			
m_{D^0}, MeV/c^2		1866.60 ± 0.03	1866.5 ± 0.2
σ, MeV/c^2		6.76 ± 0.04	7.2 ± 0.2
α_L		0.132 ± 0.004	0.04 ± 0.03
α_R		0.096 ± 0.004	0.058 ± 0.003

Table 5.12: Values of the "Cruijff" function shape parameters, obtained by fits to several data and MC samples, chosen to validate the choice of this PDF. The shape of the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ peaks in these samples is assumed to be close to that in the samples used for the measurement of $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$.

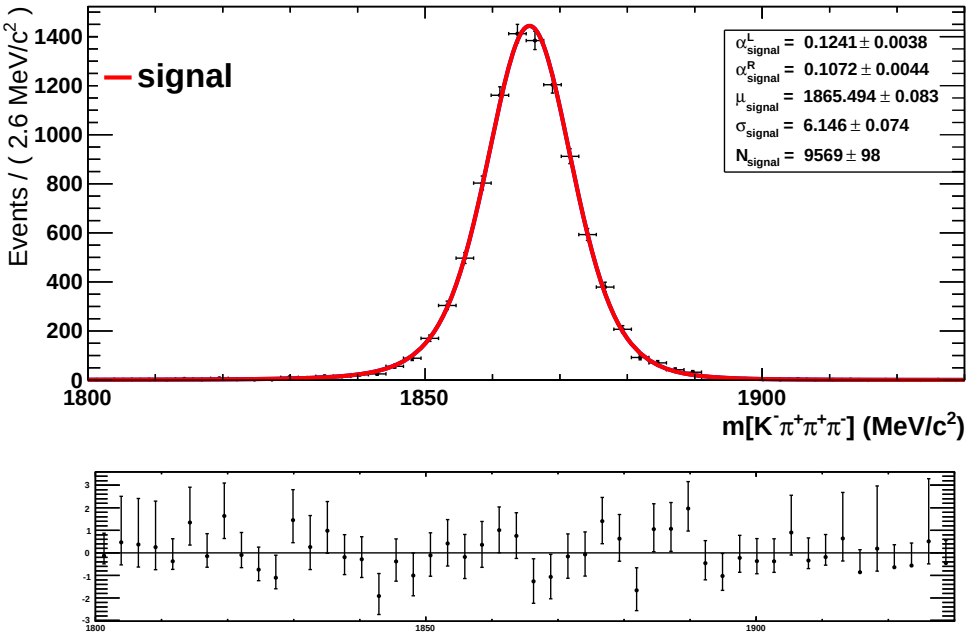


Figure 5.32: Fit to the $D^0 \to K^-\pi^+\pi^+\pi^-$ MC sample with the Cruijff function.

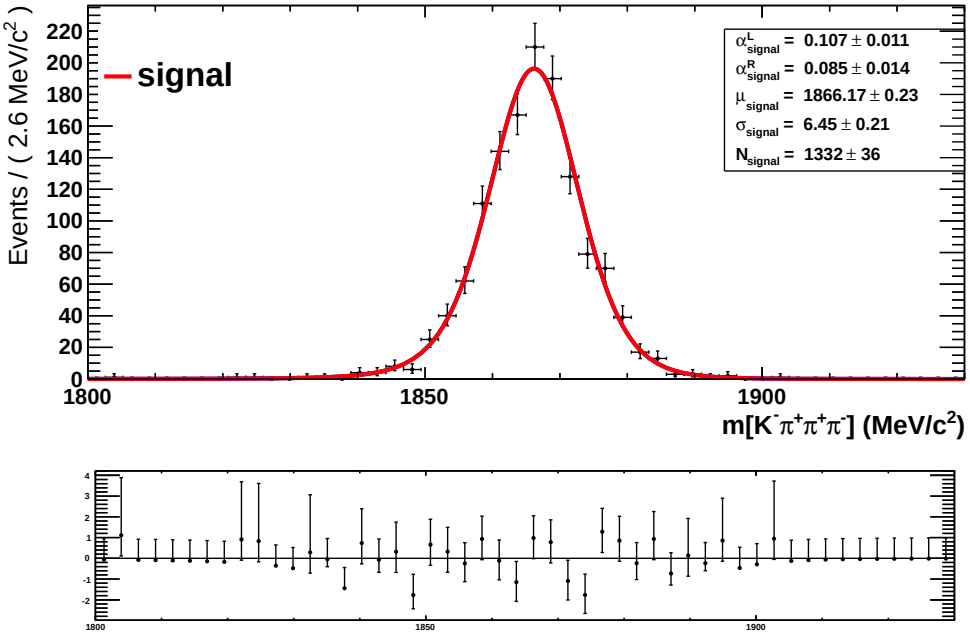


Figure 5.33: Fit to the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ MC sample, selected with `L0HadronTIS` requirement, with the Cruijff function.

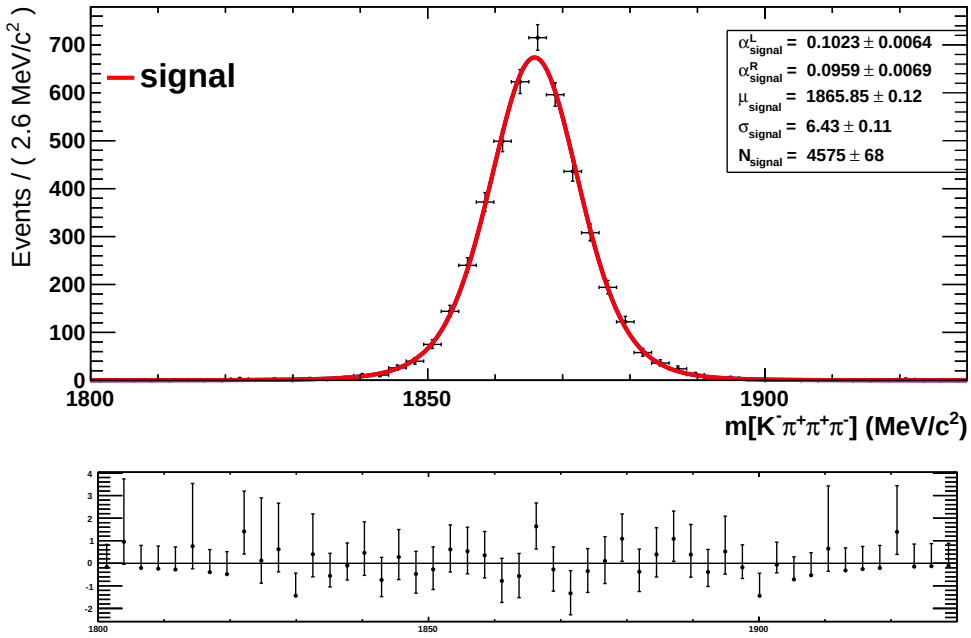


Figure 5.34: Fit to the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ MC sample, selected with `L0HadronTOS||L0MuonTOS` requirement, with the Cruijff function.

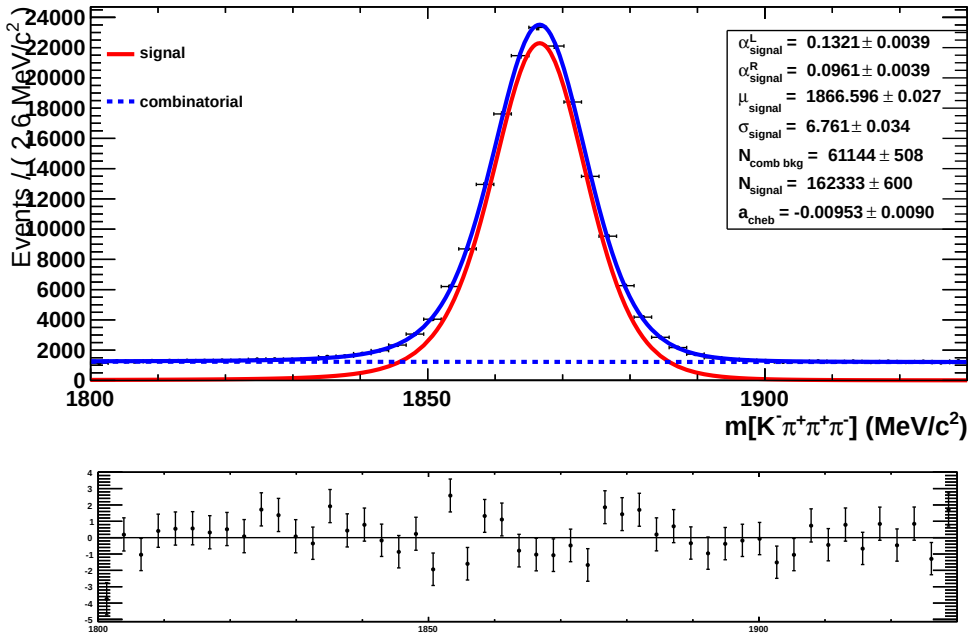


Figure 5.35: Fit to the $D^0 \to K^-\pi^+\pi^+\pi^-$ data sample, using the Cruijff function to parametrize the peak. The treatment of the non-peaking component is described in Section 5.7.3.

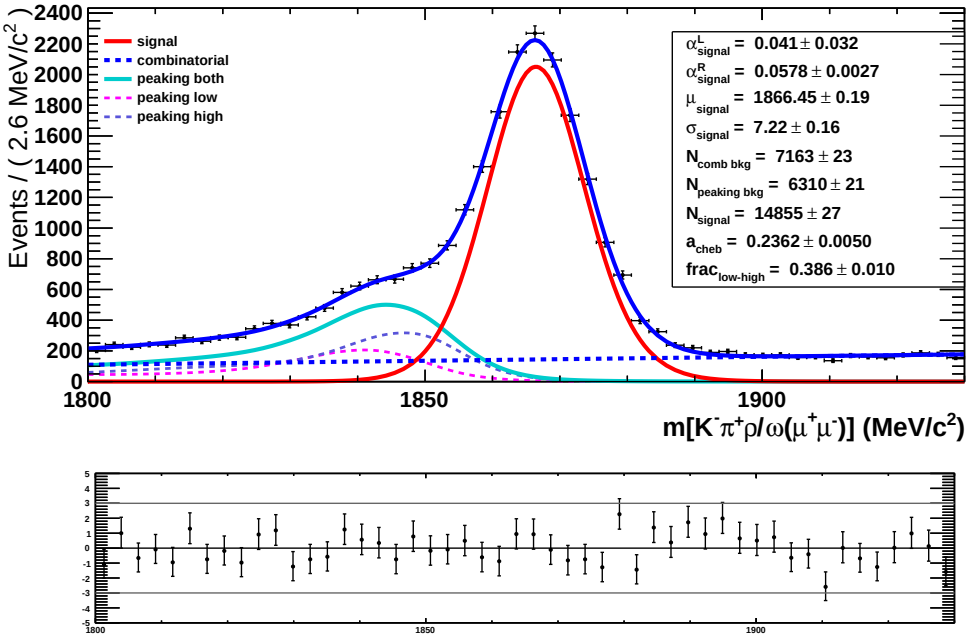


Figure 5.36: Fit to the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ data sample, selected with `L0HadronTOS||L0MuonTOS` requirement, using the Cruijff function to parametrize the signal peak. The other contributions to the total PDF are described in detail in sections 5.7.3 and 5.7.4.

5.7.3 Parametrisation of the combinatorial background

Both the fits to the signal and normalisation samples describe the combinatorial background with a polynomial of order one. Fits displayed in Figures 5.36 and 5.35 validate this choice.

In the case of the $D^0 \to K^-\pi^+\mu^+\mu^-$ sample, however, the combinatorial background present in the left sideband ($1800 < m_{K^-\pi^+\mu^+\mu^-} < 1840$ MeV/c^2) cannot be observed directly, due to the presence of doubly misidentified $D^0 \to K^-\pi^+\pi^+\pi^-$ events. As an additional cross-check that a linear description of this background is adequate, we *generated* signal candidates which features are similar to that of the combinatorial candidates, following the method described in the next paragraphs.

The signal selection (Section 5.4) has to be tight in order to control the high background inherent to a four-body decay. Consequently, most of the background candidates should actually involve a kaon coming from a D meson. Only its association with pions or muons is random. Based on this picture, it should be possible to generate a typical combinatorial background candidate by:

- taking the signal candidate in the current event,
- removing randomly one pion or one muon from the decay chain,
- replacing this particle by a pion or a muon chosen randomly in the decay chain of a signal candidate chosen in one of the previous events.

We also generated candidates by replacing two or three particles instead of only one. The distributions of the three sets of candidates obtained this way are summed up to form the distribution we need to study.

This procedure was tested in $D^0 \to K^-\pi^+\pi^+\pi^-$ events. Figure 5.37 shows the generated combinatorial distribution superimposed on the usual $m_{K^-\pi^+\pi^+\pi^-}$ one. The generated distribution matches well the real one. We also tested this method on the $D^0 \to K^-\pi^+\mu^+\mu^-$ sample selected with the `L0HadronTOS||L0MuonTOS` requirement. The result can be seen in Figure 5.38. When applied to the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ signal sample used to measure $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$, this method yields Figure 5.39. A Chebyshev polynomial of order 1 allows a good fit to the distributions in Figures 5.37, 5.38 and 5.39.

This combinatorial background generation will also be used to cross check the description of the doubly misidentified $D^0 \to K^-\pi^+\pi^+\pi^-$ background. This is subject is covered in detail in Section 5.7.4. One of the difficulties is to determine the yield of such decays where one or two pions decayed in flight ($\pi \to \mu\nu_\mu$) before the end of the tracking system, relative to that of other misidentified $D^0 \to K^-\pi^+\pi^+\pi^-$ decays. It's easier to fit this quantity in control data when the Chebyshev polynomial parameter is fixed to an a priori value.

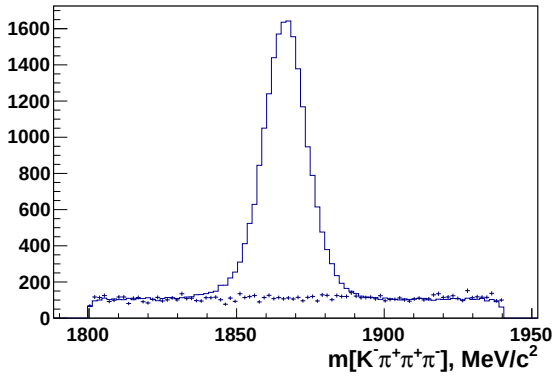


Figure 5.37: $D^0 \to K^-\pi^+\pi^+\pi^-$ invariant mass distribution and generated combinatorial background (crosses).

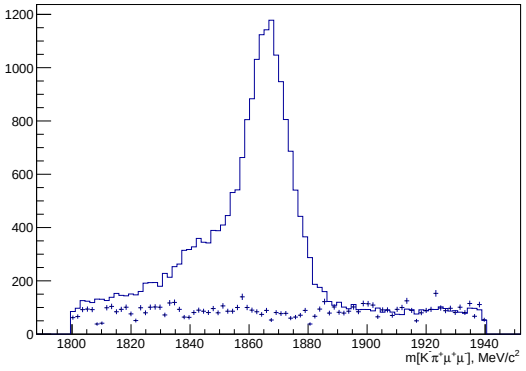


Figure 5.38: Invariant mass distribution and generated combinatorial background (crosses) for the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ selected with `L0HadronTOS||L0MuonTOS` requirement.

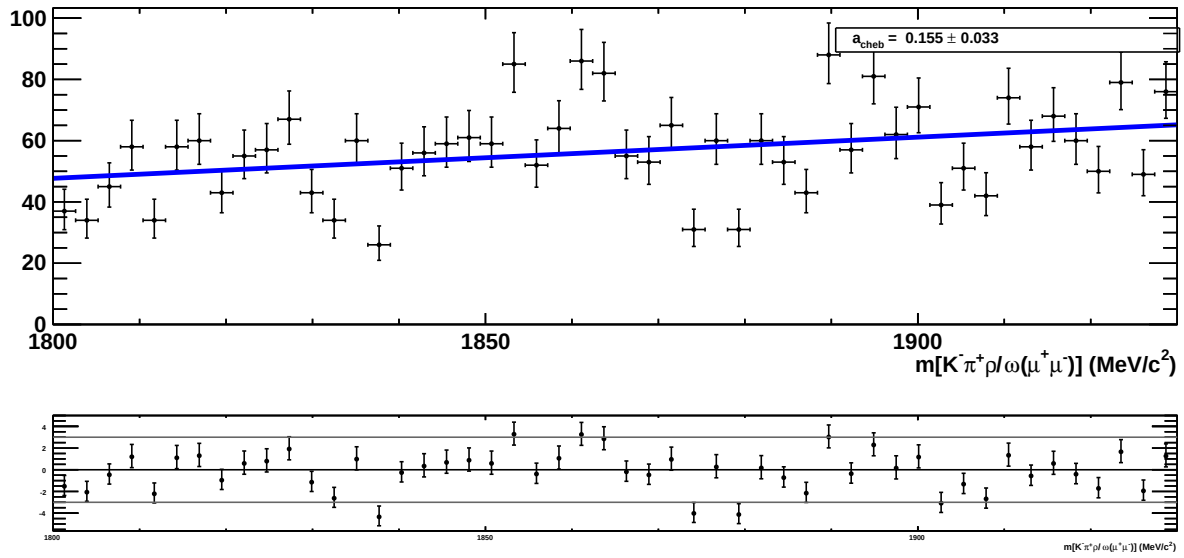


Figure 5.39: Generated combinatorial background produced with the nominal $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ signal sample.

5.7.4 Parametrisation of the double misidentification background from $D^0 \to K^-\pi^+\pi^+\pi^+$

The most problematic background to $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ is $D^0 \to K^-\pi^+\pi^+\pi^-$ where two pions are misidentified as muons. Its branching fraction is four orders of magnitude larger than what we expect for the signal: $\mathcal{B} = (8.287 \pm 0.043_{stat.} \pm 0.200_{syst.})$ % [19]. The efficiency of the muon identification criteria applied to a pair of pions is about 0.25×10^{-4}, while the other selection criteria should have the same efficiency as in the case of the signal, since both channels are similar kinematically and topologically. Consequently, the number of $D^0 \to K^-\pi^+\pi^+\pi^-$ events that we expect to find in the signal sample is of the same order as the signal peak itself. Both peaks are separated by only ~ 20 MeV/c^2, i.e. about twice their width. For that reason, unless the shape of this peaking background is well determined, it can make the determination of the signal yield imprecise and biased, given the statistics available in the signal sample. It is difficult, in particular, to determine simultaneously the width of both peaks. The solution we chose is to use non-parametric kernel estimation functions (RooKeysPDF) to model the shape of $D^0 \to K^-\pi^+\pi^+\pi^-$ background events. We calibrated them using a real data control sample, as well as the $D^0 \to K^-\pi^+\pi^+\pi^-$ MC sample, as is described in detail in this section.

In the final state of $D^0 \to K^-\pi^+\pi^+\pi^-$ there are two pions of the same sign, π^+. Thus two combinations of double misidentification are possible, which we distinguish according to the relative transverse momenta of these two pions:

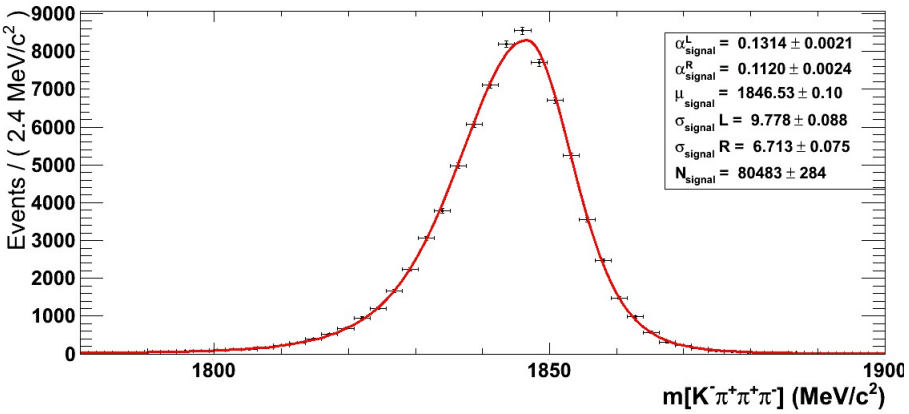
- $D^0 \to K^-(\pi^+_{high} \to \mu^+)\pi^+(\pi^- \to \mu^-)$ which is called here "high p_t"
- $D^0 \to K^-\pi^+(\pi^+_{low} \to \mu^+)(\pi^- \to \mu^-)$ which is called here "low p_t".

The shape of invariant mass distribution depends a lot on the kinematics of the misidentified pion. This can be seen in Figure 5.40. This is why we treat both cases separately: we use two separate PDF's to model these backgrounds.

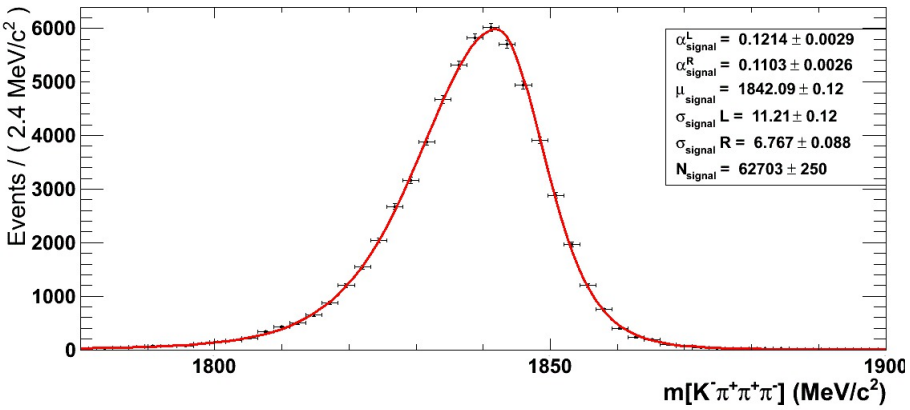
Final state pions from $D^0 \to K^-\pi^+\pi^+\pi^-$ often decay in flight to $\mu^+\nu_\mu$. Two categories of misidentification have to be considered: pions that do not decay in flight or only decay downstream of the tracking system, and pions that decay in flight before they have reached the last station of the tracking system. Thus, for double misidentification, four possible configurations have to be treated:

1. neither the π^+ nor the π^- decays in flight before the end of the tracking system;
2. π^+ does, π^- does not;
3. π^- does, π^+ does not;
4. both do.

In the first category above, only the mass hypothesis between π and μ is swapped (see Figure 5.40) to predict the shape of the mass peak from a control sample containing a peak of $D^0 \to K^-\pi^+\pi^+\pi^-$ decays. In the three other cases, one has to account for the fact that the loss of a neutrino leads to an underestimate of the momentum measured



(a) High p_t



(b) Low p_t

Figure 5.40: Invariant mass distributions of $D^0 \to K^-\pi^+\pi^+\pi^-$ candidates in MC, with swapped mass hypothesis $\pi^+ \to \mu^+$ and $\pi^- \to \mu^-$ in two possible configurations: among the decay products of $D^0 \to K^-\pi^+\pi^+\pi^-$, either the "higher p_t" π^+, or the "lower p_t" π^+ is misidentified.

by the tracking system. This creates a tail to the left of the main peak, extending all over the lower mass sideband of the fit region. This tail is hard to distinguish from the combinatorial background. Consequently, the yields of peaking and combinatorial backgrounds measured by the fit are correlated, causing large uncertainties, biases and fit instability in the absence of a precise knowledge of the shape of these tails.

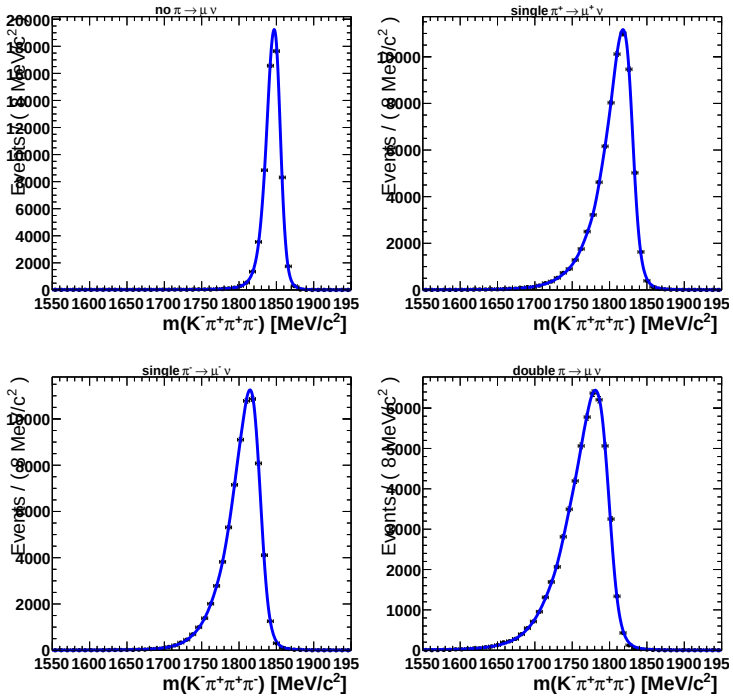
The $D^0 \to K^-\pi^+\pi^+\pi^-$ MC sample (see Section 5.3) is large: more than 150000 events are available after stripping. However, given the very low $\pi \to \mu$ misidentification rate, this is not enough to characterize the tails described above. However, it contains several thousand events in which one of the three pions decayed in flight. This "decayed pions" sample is used to simulate the effect of the neutrino loss on the reconstructed momentum of a pion. The following steps are followed to obtain the shape of the tail:

- A large $D^0 \to K^-\pi^+\pi^+\pi^-$ sample is built, by applying the full selection to the MC sample mentioned above, except for a muon identification, and by removing events in which one of the pions is identified as a pion which decayed in flight.
- One or two pions are chosen from among the decay products (depending on the configuration for which we want to determine the tail).
- Based on its (their) electric charge, p_x, p_y and p_z, this (these) pion(s) is (are) matched to one (some) of the pions from the "decayed pions" sample. The momentum of the pion(s) under consideration is shifted by some δp_x, δp_y and δp_z equal to the shift observed between the true and reconstructed quantities of the reference pion(s) found in the "decayed pions" sample. This reproduced in average the effect of the lost neutrino.
- The $m_{D^0}^{K\pi\mu\mu}$ invariant mass is recalculated based on this new kinematics.
- The distributions obtained this way are transformed into non-parametric kernel estimation functions RooKeysPDFs.

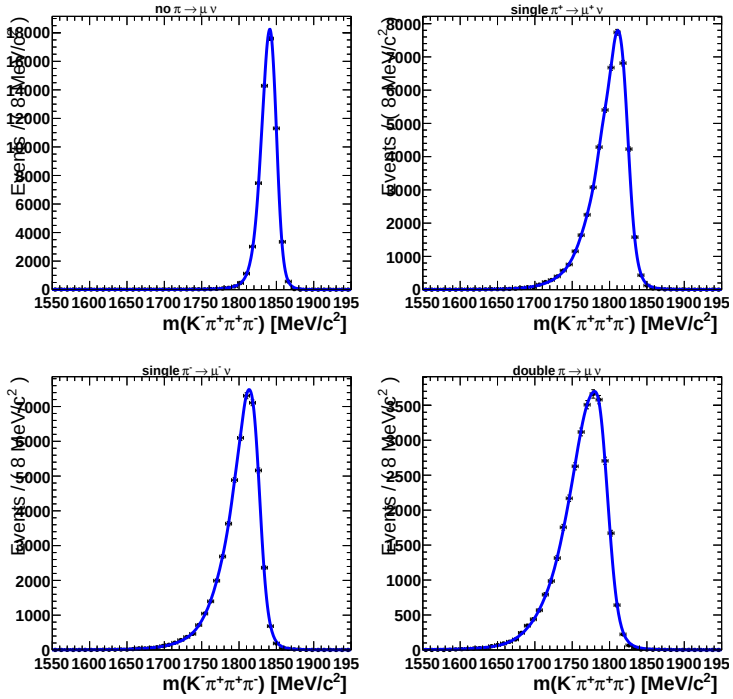
This way we obtain four shapes (one for each of the decay in flight configurations outlines above) for each of the two possible choices of misidentified π^+, "high p_t" (Figure 5.41(a)) and "low p_t" (Figure 5.41(b)). Thus we have 4+4 shapes. We use non-parametric kernel estimation function (RooKeysPDF), as stated above. In the case with no decays in flight before the end of the tracking system, the shape is obtained by isolating correctly reconstructed decay in a $D^0 \to K^-\pi^+\pi^+\pi^-$ data control sample and by re-computing the invariant mass of the decay products assuming wrongly that the mass of two pions equals the mass of a muon.

The next step is to evaluate relative contribution of each of the 4 peaking background components:

- f^0, the contribution of events with no decay in flight before the end of the tracking system. It is fixed to 1. The other contributions are calculated with respect to this one.
- f^{1,π^+} and f^{1,π^-} the contribution of events in which one π^+ or the π^- decays before the end of the tracking system. Both fractions are supposed to be equal. It can indeed be seen on Figure 5.41) that their shapes are too similar to be separated.



(a) High p_t



(b) Low p_t

Figure 5.41: RooKeysPDFs describing the invariant mass of the $D^0 \to K^-\pi^+\pi^+\pi^-$ peaking background. Both the "high p_t" (top) and "low p_t" (bottom) components receive four contributions: no decay in flight (top left), π^+ decay in flight (top right), π^- decay in flight (bottom left), both misidentified pions decayed in flight (bottom right). These shapes are obtained by recalculating the mass of candidates after the kinematics of one or two pions has been modified to reproduce the effect of a decay in flight (see the detailed description in the text).

- f^2, the contribution of events where both a π^+ and the π^- decay in flight before the end of the tracking system.

To determine the fractions outlined above, we perform a fit to a large statistics control sample: the signal sample where the `L0HadronTIS` is replaced by `L0HadronTOS||L0MuonTOS`. As $\pi \to \mu$ misidentification depends significantly on the momentum of the final state particles, we have checked the similarity of the muons p and p_t distributions (Figure 5.42) in the `L0HadronTOS||L0MuonTOS` and `L0HadronTIS` samples. The momentum distributions correspond to invariant mass region $m(D^0) \in [1800; 1840]$ MeV/c^2, where the peaking background contributes the most. Figure 5.42 shows enough similarity for the relative fractions obtained with a fit on the `L0HadronTOS||L0MuonTOS` sample to be used in the final fit. The `L0HadronTOS||L0MuonTOS` $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ sample is split in two subsamples corresponding to the "high p_t" and "low p_t" misidentification: the first subsample contains events where $p_t(\mu^+) > p_t(\pi^+)$, the second subsample contains the rest of the events. The fit is performed on these two samples separately. This can be visualized on Figures 5.43. The resulting values of f^0, $f^{1,\pi^{+/-}}$ and f^2 are shown in Table 5.13.

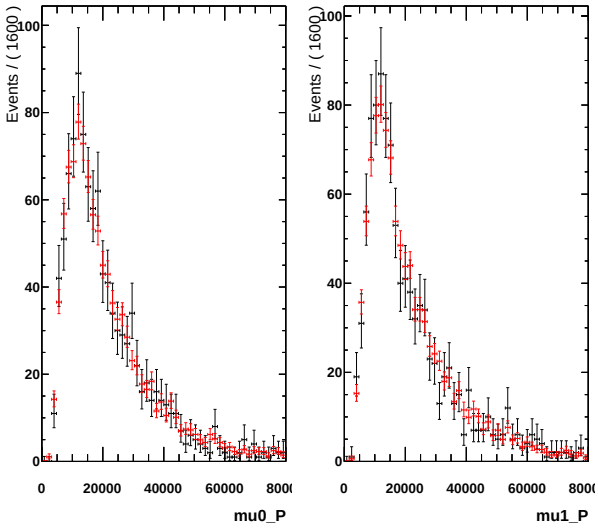
The fit is repeated with the parameter of the first order Chebyshev polynomial describing the combinatorial background, α_{cheb}, fixed to the value found by the fit to the *generated* combinatorial background presented in Section 5.7.3 $\alpha_{cheb} = 0.155 \pm 0.033$. This can be visualized in Figure 5.44. The resulting values of f^0, $f^{1,\pi^{+/-}}$ and f^2 are very close to the default ones.

The final shapes for the peaking background are presented in Figure 5.45.

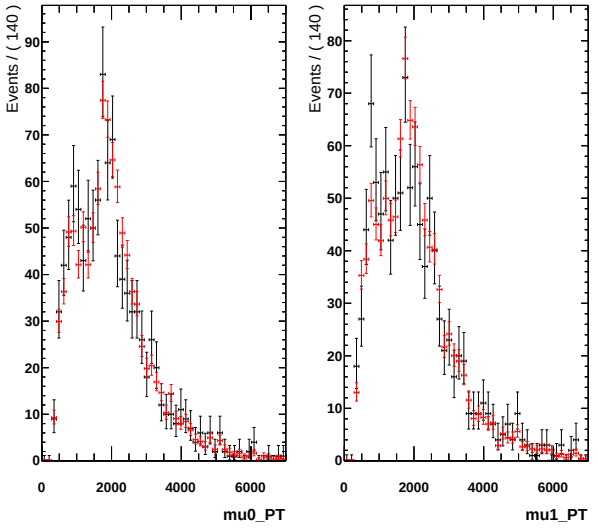
	f^0	$f^{1,\pi^+} = f^{1,\pi^-}$	f^2	α_{cheb}
High p_t	1.0	0.16 ± 0.03	0.000 ± 0.008	floats
Low p_t	1.0	0.09 ± 0.05	0.03 ± 0.03	floats
High p_t	1.0	0.16 ± 0.01	0.000 ± 0.008	fixed
Low p_t	1.0	0.08 ± 0.02	0.03 ± 0.03	fixed

Table 5.13: Relative fractions of different peaking background components shown in Figure 5.45. The coefficient of the first order polynomial describing the combinatorial background either floats or is fixed to the value fitted to the *generated* background.

For the final fit measuring $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$ we do not divide sample on "low p_t" and "high p_t" subsamples. The peaking $D^0 \to K^-\pi^+\pi^+\pi^-$ background is consequently described by the sum of the RooKeysPDFs "low p_t" and "high p_t" shown on Figure 5.45. The fraction of "low p_t" contributing is a free parameter in the fit, $f^{low\ P_T}$.

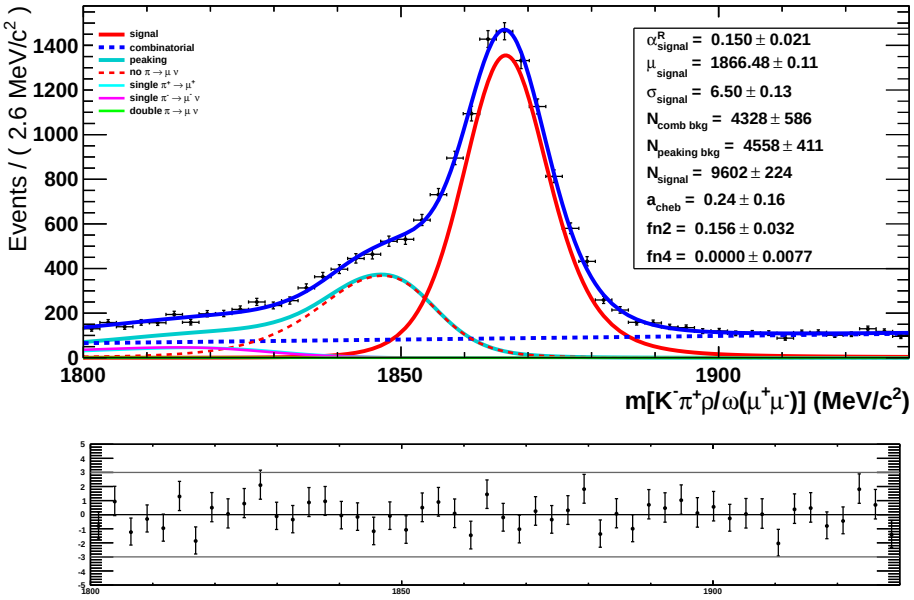


(a) Momentum distributions

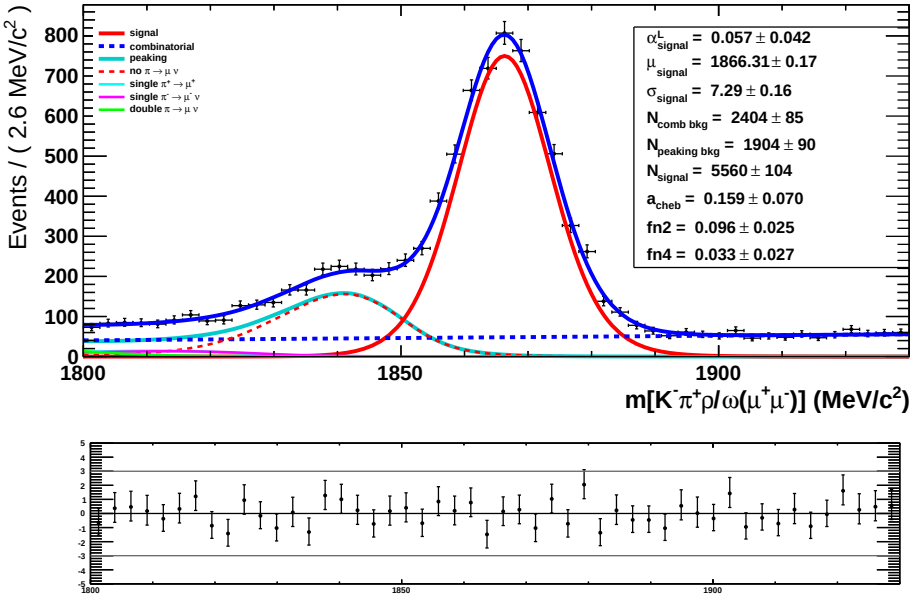


(b) Transverse momentum distributions

Figure 5.42: Momentum p and transverse momentum p_t distributions of the two muons involved in $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ candidates. Two samples are compared: the nominal (`L0HadronTIS`) signal sample (black) and the sample requiring `L0HadronTOS||L0MuonTOS` (red). The nominal selection is applied. The comparison is restricted to the left mass sideband ($1800 - 1840$ MeV/c^2): most of the misidentified $D^0 \to K^-\pi^+\pi^+\pi^-$ background events can be found in this region, including the vast majority of the candidates for which one or two pions decayed in flight.

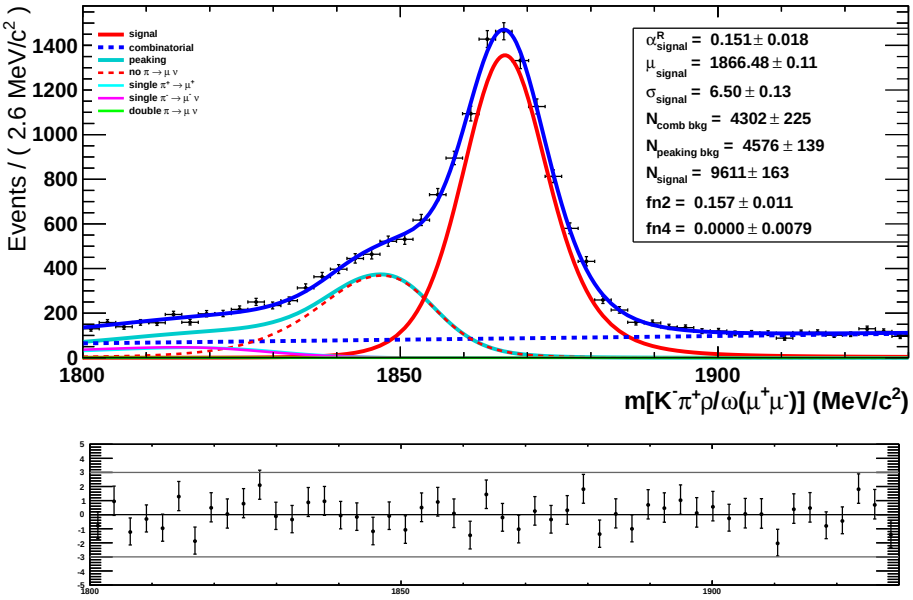


(a) High p_t

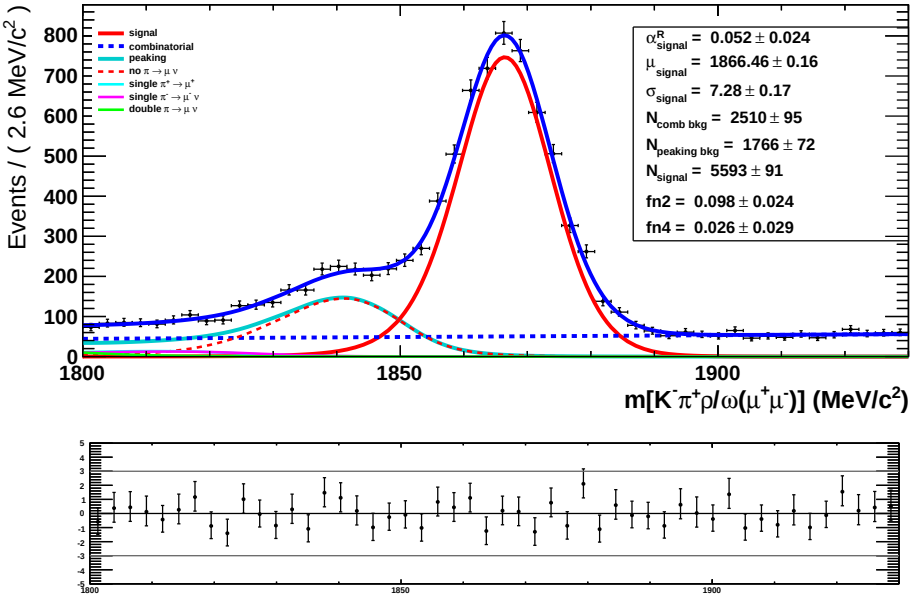


(b) Low p_t

Figure 5.43: Fit to the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ sample selected with `L0HadronTOS||L0MuonTOS` requirement. It is performed in order to validate the description of the $D^0 \to K^-\pi^+\pi^+\pi^-$ peaking background based on the RooKeysPDFs which determination is described in the text, as well as in order to determine the relative fractions of the four contributions this background is made of. The components of the total PDF are: the signal (red solid line), the combinatorial background (blue dashed line) and the various components of the peaking background. These components are with: no pion decayed in flight (red dashed line), one π^+ decayed in flight (solid cyan line), one π^- decayed in flight (magenta solid line), both pions decayed in flight (solid green line). The coefficient of the first order Chebyshev polynomial describing the combinatorial background is free to float.

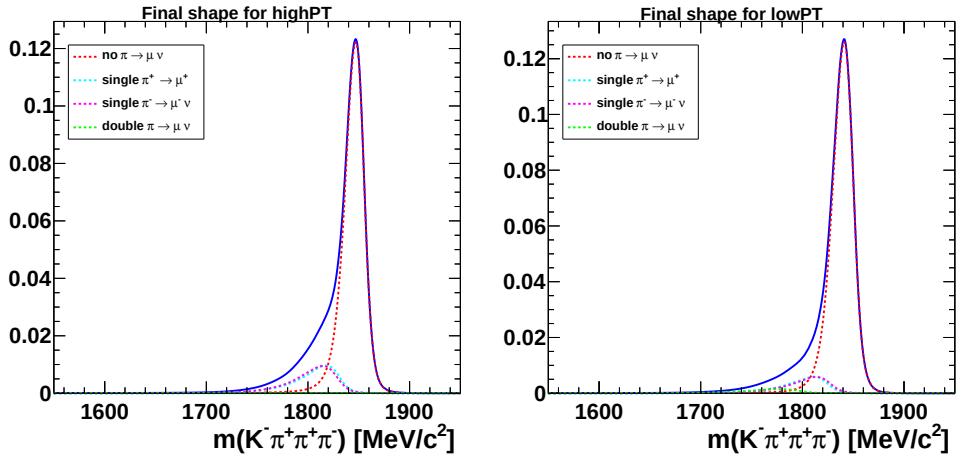


(a) High p_t



(b) Low p_t

Figure 5.44: Fit to the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ sample selected with `L0HadronTOS||L0MuonTOS` requirement. It is performed in order to validate the description of the $D^0 \to K^-\pi^+\pi^+\pi^-$ peaking background based on the RooKeysPDFs which determination is described in the text, as well as in order to determine the relative fractions of the four contributions this background is made of. The components of the total PDF are: the signal (red solid), the combinatorial background (blue dashed) and the various components of the peaking background. These components are with: no pion decayed in flight (red dashed), one π^+ decayed in flight (solid cyan), one π^- decayed in flight (magenta solid), both pions decayed in flight (solid green). The coefficient of the first order Chebyshev polynomial describing the combinatorial background is fixed to the value found by a fit to the *generated* combinatorial background (see Section 5.7.3).



(a) High p_t (b) Low p_t

Figure 5.45: Final RooKeysPDF shapes with parameters from Figure 5.43 (coefficient of the Chebyshev polynomial describing the combinatorial background floating). Red line - component with no decays in flight, cyan and magenta - components with one decay in flight, green - component with two decays in flight.

5.7.5 Validation of the fit model on data control sample

To validate the fit model described in the previous sections, we perform a fit to the sample introduced in Section 5.4 to optimise the selection (20% of the 2012 luminosity). In Figure 5.46 one can see the data seems to be correctly described by the fit. The same is true when the fit is performed to the `L0HadronTOS||L0MuonTOS` sample (Figure 5.47).

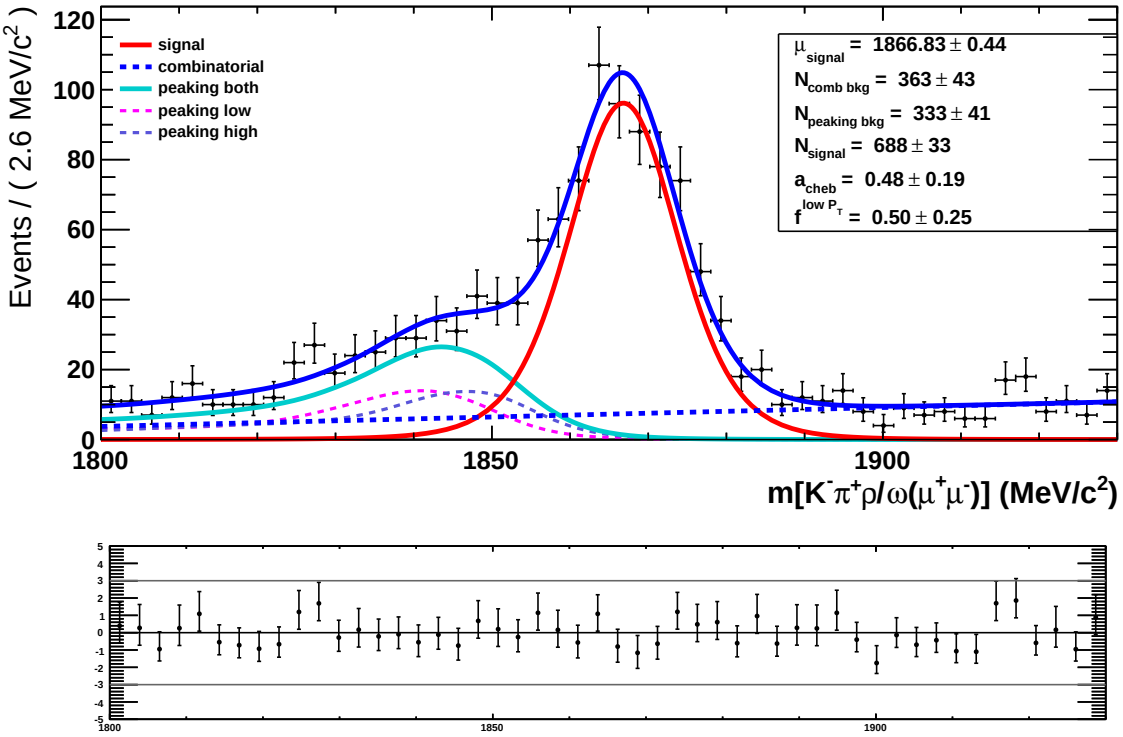


Figure 5.46: Validation fit to the sample used in Section 5.4.2 to optimize the selection. The fit model built from Section 5.7.1 to Section 5.7.4 is employed. The components of the total PDF (blue solid line) are: the signal (red solid line), the combinatorial background (blue dashed line) and the "high p_t" (violet dashed line) and "low p_t" (magenta dashed line) components of the peaking background. The coefficient of the first order Chebyshev polynomial describing the combinatorial background is free to float.

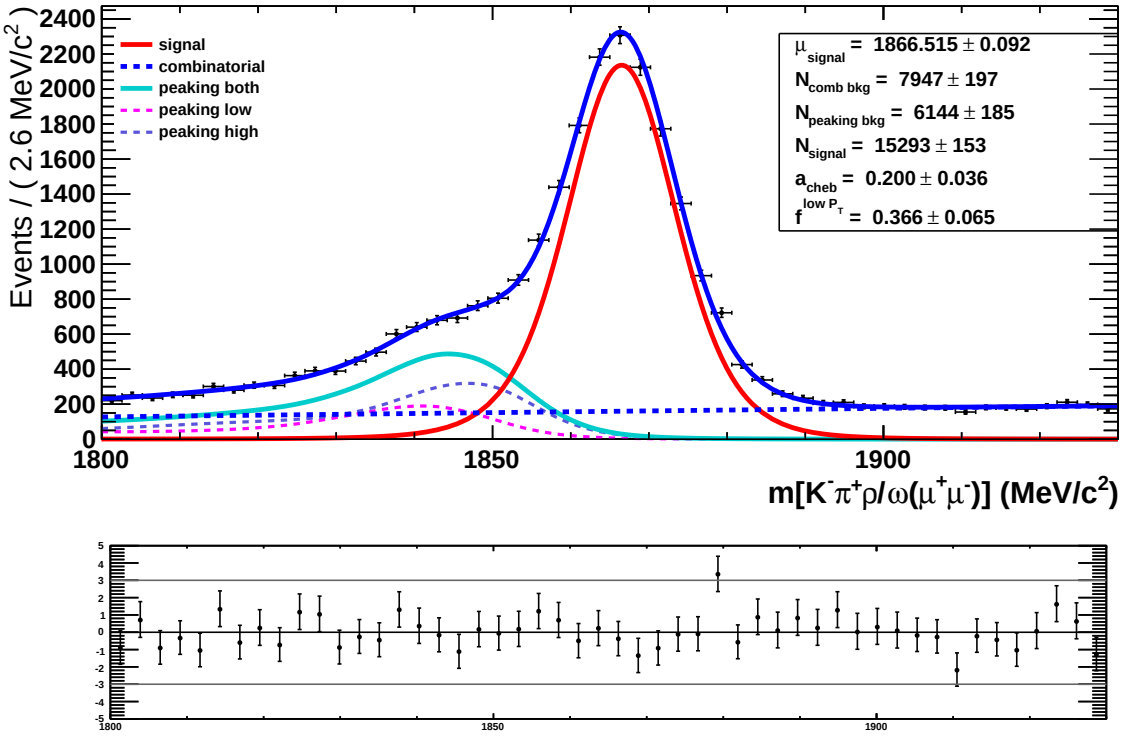


Figure 5.47: Fit to the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ data sample, selected with `L0HadronTOS||L0MuonTOS` requirement. The fit model built from Section 5.7.1 to Section 5.7.4 is employed. This is a way to validate it. The components of the total PDF (blue solid line) are: the signal (red solid line), the combinatorial background (blue dashed line) and the "high p_t" (violet dashed line) and "low p_t" (magenta dashed line) components of the peaking background. The coefficient of the first order Chebyshev polynomial describing the combinatorial background is free to float.

5.7.6 Unblinded fit

After the unblinding and the fitting we obtain the following branching ratio:

$$\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)) = (4.37 \pm 0.12_{stat.}) \times 10^{-6},$$

where we show only the statistical uncertainty returned by the fitter. The systematic uncertainties are discussed in Section 5.8.3.

The fit converges and describes the data well (Figure 5.48). In particular the pull distributions shows satisfactory behaviour. The parameters of the final fit are shown in Table 5.14.

Parameters	Fitted values
m_{D^0}, MeV/c^2	1867.2 ± 0.3
σ, MeV/c^2	6.78 ± 0.05
α_L	0.136 ± 0.006
α_R	0.097 ± 0.006
α_{cheb}	0.04±0.07
f^{lowP_T}	0.07±0.17
N_{signal}	2453±60
$N_{peakingbkg}$	1078±56
$N_{combbkg}$	1881±70

Table 5.14: Values of the shape parameters, obtained by the fit on data $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$, selected with `L0HadronTIS` requirement (Figure 5.48).

5.8 Branching fraction uncertainty

5.8.1 Uncertainties related to the signal extraction

"Toy studies" have been carried out to study the performance of the fit and evaluate some of the systematic uncertainties.

The value of the shape parameters and background yields assumed to generate the toys are similar to those found by the fit to the data in Section 5.7.6.

Figure 5.49 shows the distribution of the signal branching fraction returned by the fitter, as well as the pull distribution. It shows that in its default configuration, the fitter is not biased and that errors computed by the fitter are accurate. This toy study also gives an estimation of the expected statistical uncertainty on $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$:

$$\sigma_{fit} = 3\%, \tag{5.5}$$

which is in agreement with the value obtained in data $\sigma_{fit} = 2.8\%$ (Section 5.7.6).

Pull distributions for the other fitted parameters are shown in Appendix B.

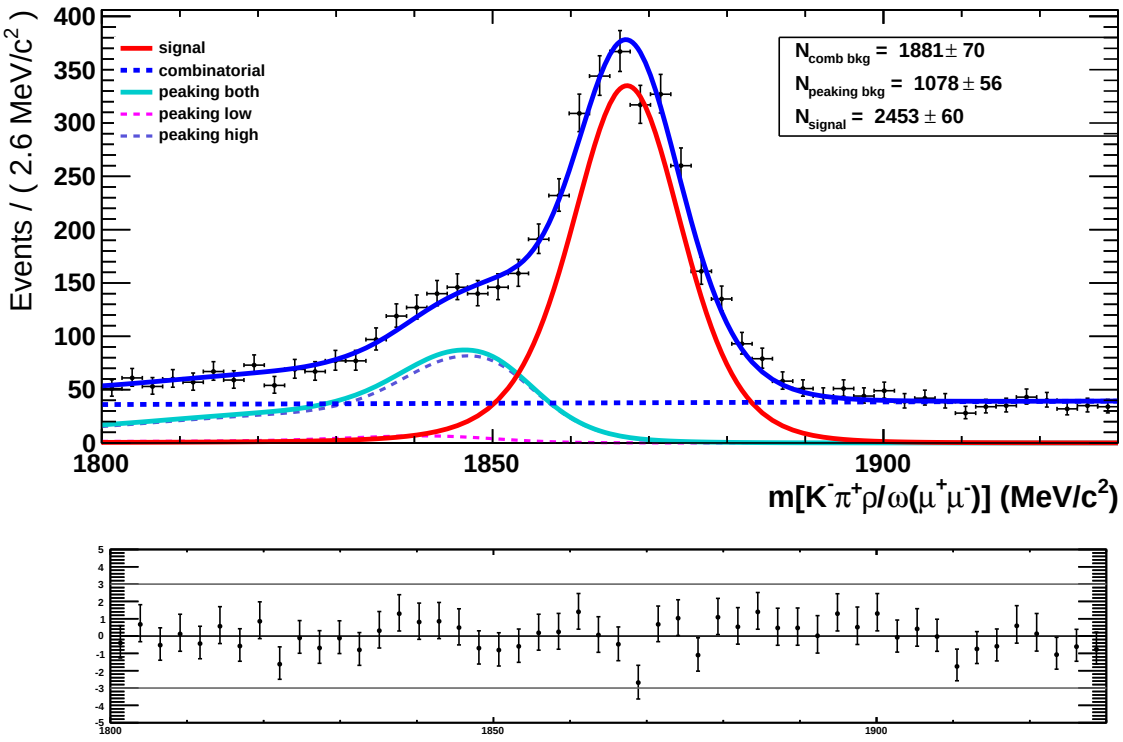
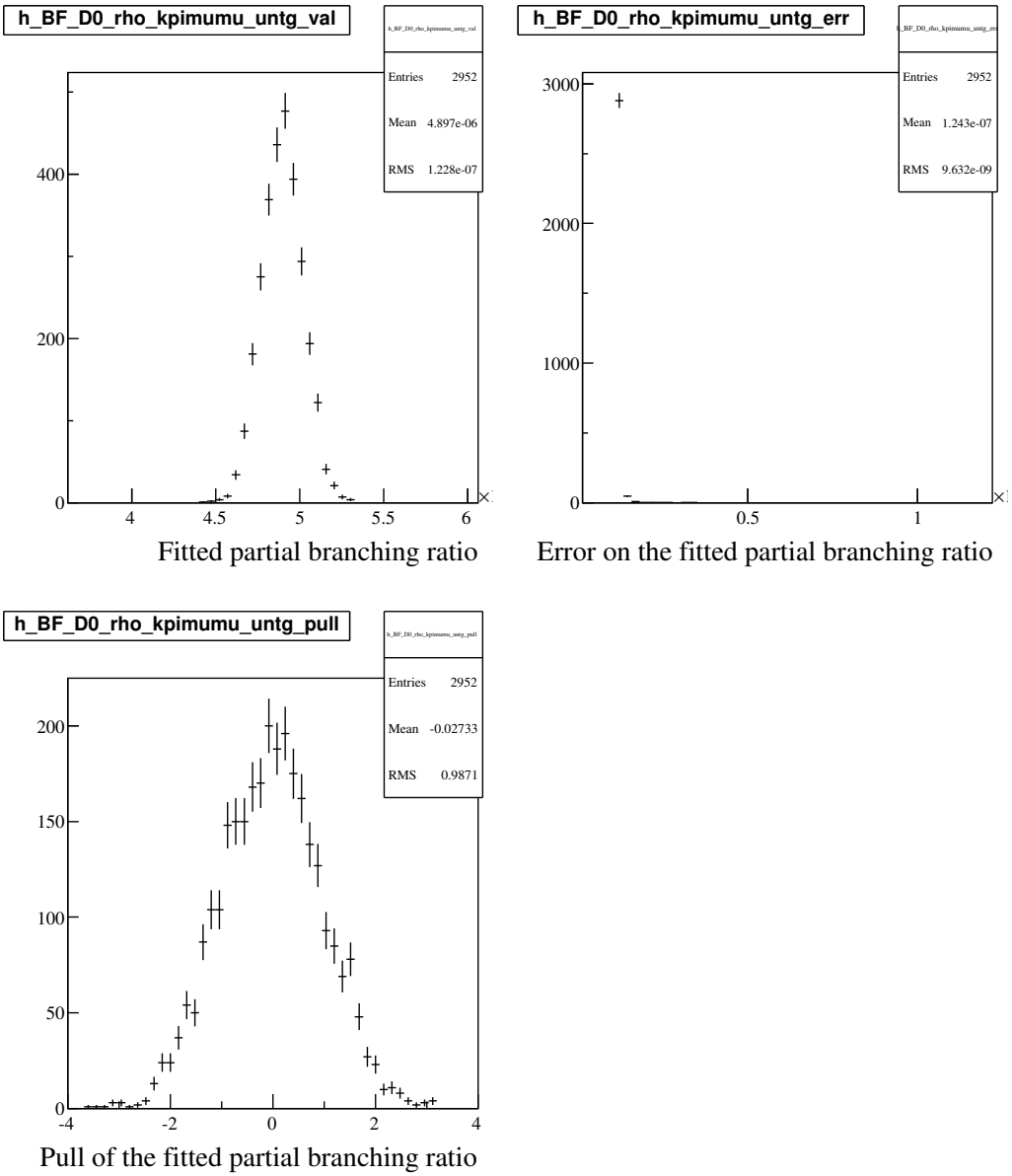


Figure 5.48: Invariant mass distribution in unblind $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ data sample, selected with `L0HadronTIS`. Red solid line - signal, blue dashed - combinatorial, cyan solid - peaking background, violet and magenta dashed - components of "highPT" and "lowPT" peaking forms. Fit parameters can be found in Table 5.14.



Fitted partial branching ratio

Error on the fitted partial branching ratio

Pull of the fitted partial branching ratio

Figure 5.49: Distribution of the values of $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$ obtained by fitting toy experiments representative of the signal data sample used for the actual measurement (top left), distribution of the uncertainty on $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$ determined by the fitter (top right) and corresponding pull distribution.

The choice of the fit model and of the fixed parameters values are a source of systematic uncertainty. These uncertainties can be evaluated with toy experiments. We performed a certain number of toys studies, listed below:

- One delicate aspect in the determination of the shape of the doubly misidentified $D^0 \to K^-\pi^+\pi^+\pi^-$ decays is the shape of the tails due to the pion decay in flight, as well as their relative contribution with respect to those which decay products reach the end of the tracking system. We generate toys assuming the tails determined in Section 5.7.4 and used by default by the fitter, and fit these samples with a model assuming they do not exist at all. The distribution of the fitted signal branching fraction and corresponding pull distribution are shown Figure B.6. It displays a bias of the order of 3%. We take this value as the corresponding systematic uncertainty.

- The default fit model assumes the same right tail (α_R) for the signal and the $D^0 \to K^-\pi^+\pi^+\pi^-$ normalization mode, while $\alpha_L^{K\pi\mu\mu}$ is fixed to the value found in MC (see Sections 5.7.1 and 5.7.2).

  To evaluate the systematic uncertainty due to these choices, we generate toys with the default fit model, and fit them with the tail parameters values fixed to the values found by fitting the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ sample selected with the `L0HadronTOS||L0MuonTOS` requirement (Figure 5.36 and Section 5.7.2). A bias of the order of 4% is found which is assigned as the corresponding uncertainty.

- The treatment of the signal's high mass tail is further tested by generating toy experiment according to the default model, and by fitting it with the value of α_R fixed to what was found by the fit to the $D^0 \to K^-\pi^+\mu^+\mu^-$ MC sample. A bias of the order of 0.5% is expected, negligible compared with the previous estimation of the uncertainty due to the signal's tail. No additional systematic uncertainty is assigned.

- The default fit model uses the Cruijff function to describe the $m_{K^-\pi^+\mu^+\mu^-}$ distribution of the signal and of $D^0 \to K^-\pi^+\pi^+\pi^-$ decays in the normalisation sample. This is in principle an imperfect description, leading to a bias on the fitted branching ratio. We perform a toy study to evaluate the effect of changing a Cruijff for a Double Crystal Ball. We generate toy samples based on the default model, and fit them with Double Crystal Ball functions. The left tail parameters of the Crystal Ball used for the signal peak are taken from a fit to the signal MC sample. The right tail is taken again from the normalisation mode. The signal width is also assumed equal to that of the normalisation peak. This study reveals a 1% bias.

- As explained in Section 5.7.4, the main component of the background from doubly misidentified $D^0 \to K^-\pi^+\pi^+\pi^-$ decays is made of decays whose products reach the end of the tracking system before any pion decays in flight. Its shape in $m_{K^-\pi^+\mu^+\mu^-}$ can easily be obtained by isolating correctly reconstructed decays in a $D^0 \to K^-\pi^+\pi^+\pi^-$ data control sample and by re-computing the invariant

mass of the decay products assuming wrongly that the mass of two pions equals the mass of a muon. This is how this component is described in the default fit model. It is also possible to use a MC sample instead of the data control sample mentioned above. We generate toy experiments based on the default fit model, and fit to them an alternative fit model using the PDF derived from MC. We conclude that this effect is negligible.

Table 5.15 shows the various systematic uncertainties related to the signal fit. The total systematic uncertainty is:

$$\sigma_{fit} = 5.6\% \tag{5.6}$$

Combining this with the statistical uncertainty, we obtain the following total fit uncertainty:

$$\sigma_{fit} = 6.3\%. \tag{5.7}$$

Effect	Uncertainty from toys (%)	Uncertainty from unblinded fit(%)
Peaking Background tails	3%	3.3%
Signal Tails	4%	4.3%
Alternative PDF	1%	1.3%
Total	5%	5.6%
Statistics	3%	2.8%
Fit uncertainty	6%	6.3%

Table 5.15: List of the uncertainties related to the signal extraction.

5.8.2 Uncertainty related to the L0 efficiency ratio

An additional cross check was performed after unblinding. We compared the branching ratios of $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ obtained in three data sets, selected with different L0 trigger requirements:

- `L0HadronTIS`, trigger fired by hadrons from non-signal decay (Figure 5.48)
- `L0MuonTIS`, trigger fired by muons from non-signal decay (Figure 5.50)
- `L0HadronTIS||L0MuonTIS` (Figure 5.51)

Here we investigate the potential difference in the L0 efficiency between the signal and normalization mode.

The obtained partial branching ratios are presented in Table 5.16. The differences between obtained values are also shown in last two lines of the table. The results are consistent with each other with one possible exception for $\mathcal{B}^{L0HadronTIS} - \mathcal{B}^{L0HadronTIS\ ||\ L0MuonTIS} = (0.17 \pm 0.08) \times 10^{-6}$, where the consistency is at the level of 2σ.

We propose the uncertainty on the difference between `L0HadronTIS` and `L0MuonTIS` as a conservative systematic uncertainty on the L0 efficiency ratio: 7.7%, which is consistent with uncertainty obtained in Section 5.5.3 from the cross check of these L0 trigger efficiencies.

	`L0HadronTIS`	`L0MuonTIS`	`L0MuonTIS|| HadronTIS`
$\mathcal{B}$	$(4.37 \pm 0.19) \times 10^{-6}$	$(4.66 \pm 0.28) \times 10^{-6}$	$(4.54 \pm 0.17) \times 10^{-6}$
$\mathcal{B} - \mathcal{B}^{L0HadronTIS}$	-	$(0.29 \pm 0.34) \times 10^{-6}$	$(0.17 \pm 0.08) \times 10^{-6}$
$\mathcal{B} - \mathcal{B}^{L0MuonTIS}$	$(0.29 \pm 0.34) \times 10^{-6}$	-	$(0.12 \pm 0.23) \times 10^{-6}$

Table 5.16: Partial branching ratios of $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ obtained in different data samples. Differences between the obtained results are shown in the last two lines. The error bars take into account only the statistical uncertainties from data and MC.

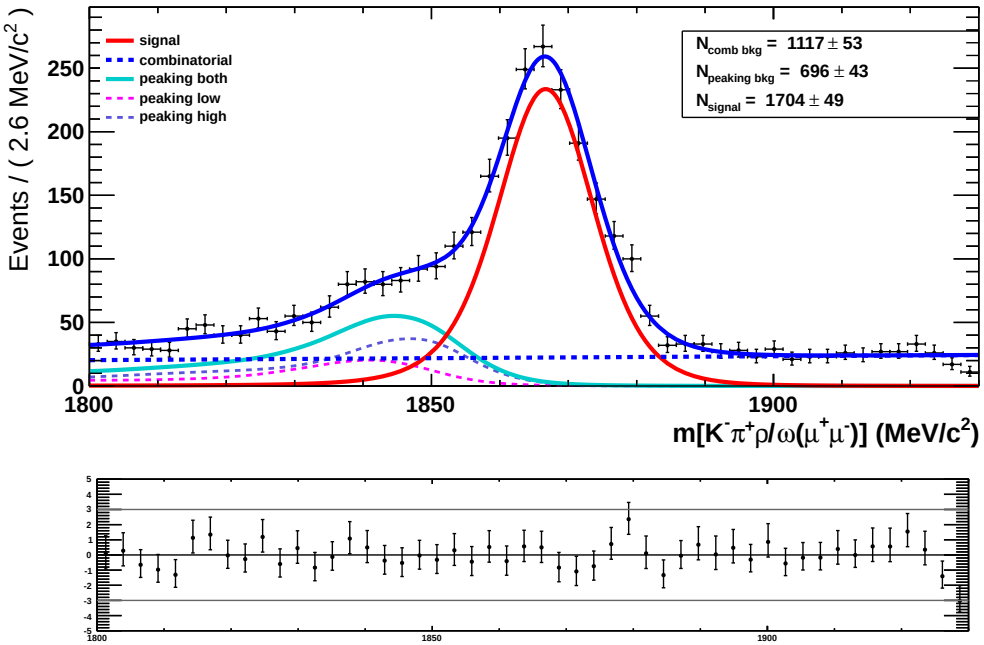


Figure 5.50: Invariant mass distribution in unblinded $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ data sample, selected with `L0MuonTIS`. Red solid line - signal, blue dashed - combinatorial, cyan solid - peaking background, violet and magenta dashed - components of "highPT" and "lowPT" peaking forms. The fit parameters can be found in Table 5.14.

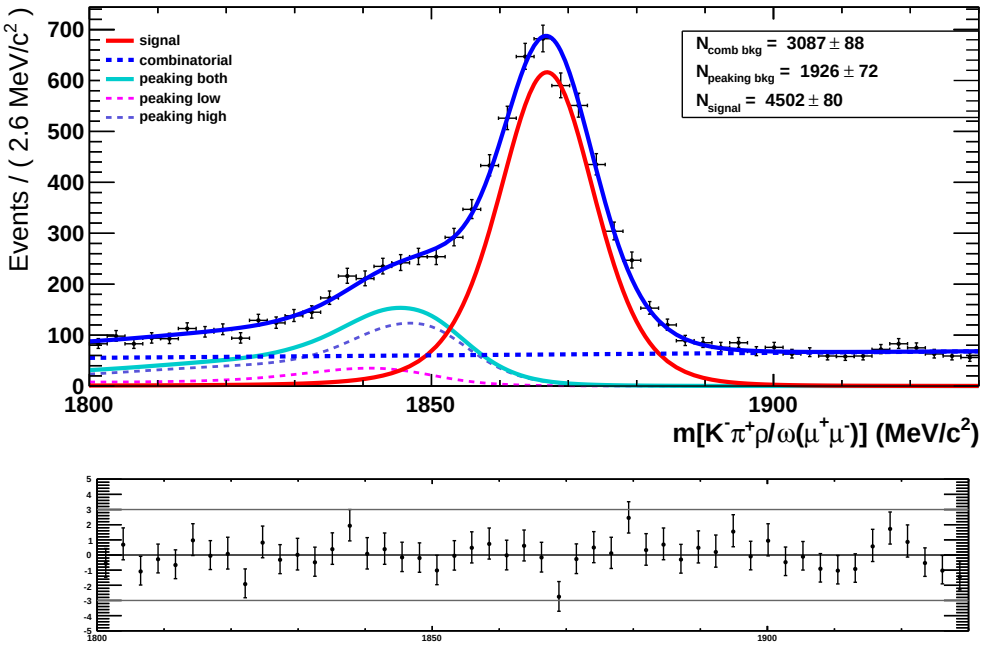


Figure 5.51: Invariant mass distribution in the unblinded $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ data sample, selected with `L0HadronTIS||L0MuonTIS`. Red solid line - signal, blue dashed - combinatorial, cyan solid - peaking background, violet and magenta dashed - components of "highPT" and "lowPT" peaking forms. The fit parameters can be found in Table 5.14.

5.8.3 Uncertainty summary

In the previous sections we evaluate the systematic uncertainties from the offline selection (Section 5.5.6), the trigger requirements L0 (Section 5.8.2), the High Level Trigger (HLT) selection (Section 5.5.4), the muon identification criteria (Section 5.5.1), the tracking efficiency 5.5.8, the kaon identification efficiency (Section 5.5.9) and on the decay model used in MC (Section 5.5.7). Table 5.17 summaries these uncertainties together with the statistical errors.

They are combined with the expected statistical uncertainty. The total uncertainty on $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$ is 12.5%.

Type	Relative uncertainty	Obtained in Section
Efficiency ratio		
Tracking	2.8 %	5.5.8
Offline selection	2.5 %	5.5.6
L0	7.7 %	5.8.2
HLT	4.0 %	5.5.4
Muon identification	2.5 %	5.5.1
K identification	1.0 %	5.5.9
MC sample stat	3.0 %	
Generation models	2.5 %	5.5.7
Signal extraction (fit)		
Peak shapes	4.3 %	
$D^0 \to K^-\pi^+\pi^+\pi^-$ peaking		5.8.1
background shape	3.3 %	
Alternative PDF	1.3 %	
σ^{tot}_{syst}	11.9 %	
σ_{stat}	2.8 %	
$\mathcal{B}(D^0 \to K^-\pi^+\pi^+\pi^-)$	2.5 %	
Total	12.5 %	

Table 5.17: Summary table of the systematic uncertainties.

5.9 Conclusion

We performed the first measurement of the partial branching ratio of the $D^0 \to K^-\pi^+\mu^+\mu^-$ decay in the ρ/ω region, defined as $m(\mu^+\mu^-) \in [675; 875]$ MeV/c^2. This measurement was carried out on data corresponding to an integrated luminosity of $2\,\text{fb}^{-1}$ collected at $\sqrt{s} = 8$ TeV by LHCb in 2012. The obtained partial branching ratio is:

$$\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)) = (4.37 \pm 0.12_{stat.} \pm 0.53_{syst.}) \times 10^{-6},$$

where the first error is statistical and the second systematic.

Chapter 6

Study of $D^0 \to h^+h^{(\prime)-}\mu^+\mu^-$ decays reconstructed with a D^{*+}-tag

We aim to search for and eventually measure certain properties of $D^0 \to K^-\pi^+\mu^+\mu^-$, $D^0 \to K^+\pi^-\mu^+\mu^-$, $D^0 \to K^+K^-\mu^+\mu^-$ rare charm decays. Because their non-resonant branching fractions are very small $\mathcal{O}(10^{-9})$, it is crucial to develop a peformant selection. We use $D^0 \to h^+h^{(\prime)-}\mu^+\mu^-$ reconstructed with D^{*+}-tag, i.e. $D^{*+} \to D^0(\to h^+h^{(\prime)-}\mu^+\mu^-)\pi^+$. This provides powerful variable against the combinatorial background, $\Delta m = m(D^{*+}) - m(D^0)$. Multivariate Analysis (MVA) techniques are also good tools for this purpose. They are discussed in more details in Section 6.1, where we consider several MVA's, in order to select the best one. In Section 6.2, for a chosen MVA we evaluate the sensitivities on branching ratios and asymmetries that should be reached by LHCb during Run I, Run II and finally the era of the LHCb Upgrade, due to begin operation in 2019.

6.1 Introduction to multivariate analysis techniques

Searches for rare charm decays require an excellent signal-background separation with a maximal signal significance. Such separation can be obtained from specific kinematical and topological features of the signal decay, which are represented by a set of discriminating variables, discussed in Section 2.11. Signal-background separation is usually a complicated task involving many variables. The simplest approach is to apply cuts on each of these variables one by one. However, if the pattern of the distributions are non-trivial or they correlate, such rectangular cuts do not give the most optimal selection as it shown in Figure 6.1, where an illustration of rectangular cut procedure with comparison to multivariate algorithms, for example, Artificial Neural Networks (ANN) and Boosted Decision Trees (BDT) [93], is given.

Multivariate methods (MVA) maximise the signal-background separation, by making optimal use of all the available information, including the correlation between the many discriminating variables. These variables are treated as an ensemble rather than as individually and independently. Instead of applying rectangular cuts, *classifiers*, i.e. algorithms that implement classification between signal and background, are built

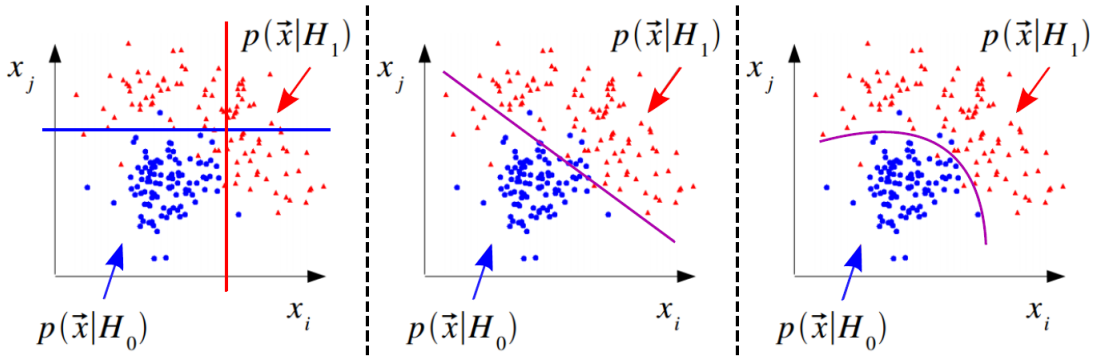


Figure 6.1: Illustration of rectangular cuts (on the left) and multivariate algorithms (on the middle and right), for two variables x_i and x_j. Red and blue points correspond to signal and background samples [93].

by combining the variables into one single output variable. One can view the MVA as a mapping from a multi-dimensional space of input variables $x_1; ...; x_{n_{var}}$ into a one-dimensional spaces.

6.1.1 Artificial neural networks

An example of an MVA is an artificial neural network (ANN) (Figure 6.2). ANNs are classification models inspired by an animal's central nervous systems. An ANN is a set of nodes, or *neurons*, connected between each other. The neurons can be presented as computational units. Each neuron calculates a linear combinations of input variables, or a given function of these combinations. The result becomes an input to the neurons of the next layer, where each neuron again computes a combination of the outputs of the neurons in the previous layer. Interaction between the neurons occurs through these connections. The learning process consists of adjustment of the weights, i.e. numerical parameters, for the links between the different neurons in order to obtain output rate for the right answer that is as high as possible.

6.1.2 Boosted decision trees

Another widely used multivariate method is the Boosted Decision Tree (BDT). A *decision tree* is a binary tree-structured classifier similar to the one sketched in Figure 6.3. The training of a decision tree is the process that defines the variable and the splitting criteria for each node. The training starts with the root node and a sequence of binary splits using the discriminating variables x_i is applied to the data. At each node, the split of the training sample is determined by finding the variable and corresponding cut value that provides the best separation between signal and background. This means that at every split a fixed number of cuts on each of the variables is tested to choose the most discriminant one, according to a certain criterion. The growing of the tree stops once it has reached the minimum number of events in a leaf or maximal number of nodes, specified in the configuration. The nodes are classified as signal or background

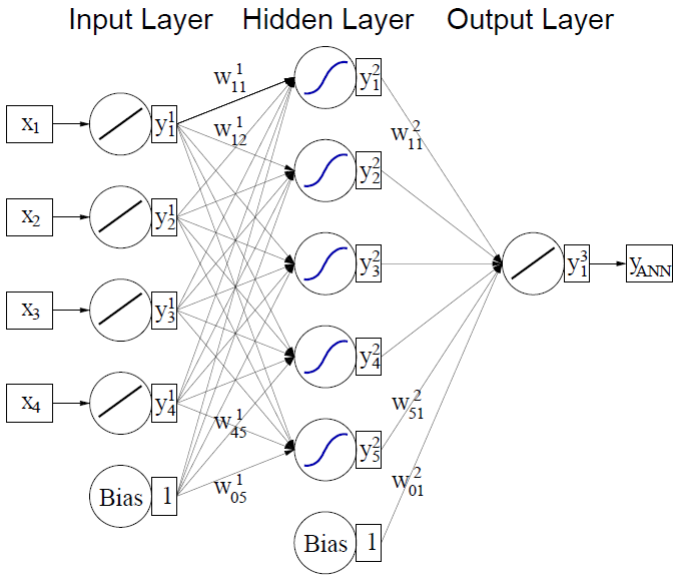


Figure 6.2: Illustration of Artificial neural networks learning process [93].

according to the class the majority of events belongs to. There are several possibilities to define the evaluation of the signal-background fractions. The one commonly used is the *Gini index*, defined as $p \cdot (1-p)$, where p is the purity $p = S/(S+B)$, with S the number of signal events and B the number of background events in the new leaf after the considered splitting.

The *boosting* is a combination of several decision trees, called *weak classifiers*, i.e. with very little discriminative power. In this manner one obtains a more performant classifier. The set of trees is built by giving a higher weight to misclassified events in the decision trees. After each iteration, misclassified samples are weighted more heavily. Thus, the severity of misclassifying them in the following iterations is increased. Any boosting method at each iteration requires the training of a new weak learner on the weighted sample.

The boost weight is derived from the misclassification rate, ϵ_m, of the previous tree at the iteration m:

$$\alpha_m = \frac{1}{2} \log \frac{1-\epsilon_m}{\epsilon_m}. \tag{6.1}$$

The weights of the entire event sample are then renormalised such that the sum of weights remains constant. The boosting can be represented as:

$$F(\mathrm{x}) = \sum_{m=1}^{M} \alpha_m f_m(\mathrm{x}), \tag{6.2}$$

where $F(\mathrm{x})$ is a weighted sum of parametrised functions $f_m(\mathrm{x})$, weak classifiers. The boosting adjusts the parameters to minimize the difference between the model response $F(\mathrm{x})$ and the true value y. This difference is measured by a *loss function*, $L(F, y)$.

There are many methods to perform boosting with different loss functions. Here we give only two the most commonly used examples:

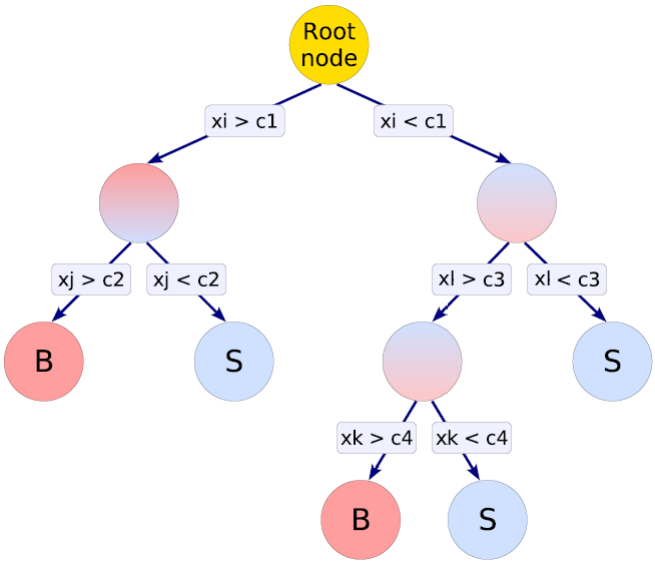


Figure 6.3: Schematic view of a decision tree

- *Adaptive Boost* (*AdaBoost*). The AdaBoost [94] is described by an exponential loss function,

$$L(F, y) = e^{-F(\mathrm{x})y}.$$

  This means that the updated weights are $w_n^{m+1} = w_n^m e^{-\alpha_m y_n h_m(\mathrm{x}_n)}/Z_m$, where m is a current iteration, Z_m is a normalization constant. This exponential loss has the shortcoming that it lacks robustness in presence of outliers or mislabelled data points. The performance of AdaBoost therefore is expected to degrade in noisy settings.

- *Gradient Boost* (*GradBoost*) [95] uses the binomial log-likelihood loss

$$L(F, y) = \ln\left(1 + e^{-2F(\mathrm{x})y}\right)$$

  for boosting. As the boosting weights corresponding to this loss function cannot be obtained in a straightforward manner, as in case of AdaBoost, one has to find the best gradient descent step to perform the minimisation. This is done by calculating the current gradient of the loss function and adjusting it to match the mean value of the gradient. Iterating this procedure yields the desired set of decision trees minimising the loss function.

When using multivariate analysis techniques one should keep in mind that an overtuning or *overtraining* is possible. This means that at certain point an algorithm starts to be optimized on statistical fluctuations and not on the general pattern of the data. In order to check overtraining the BDT is applied on a independent sample, unused during the training. The final distributions for signal and background, obtained in the training and testing samples, are compared. Another test is possible using the error function calculation, as it is shown in Figure 6.4. In this figure a learning error, i.e. misclassification error, is plotted for the BDT with two nodes ("Tree-2") in blue and

one node, i.e. rectangular cuts ("Cuts"), in red. The error are illustrated for testing and training. One can see that in the test sample the error starts to increase, while the error for the training is still decreasing. This is also a sign of overtraining.

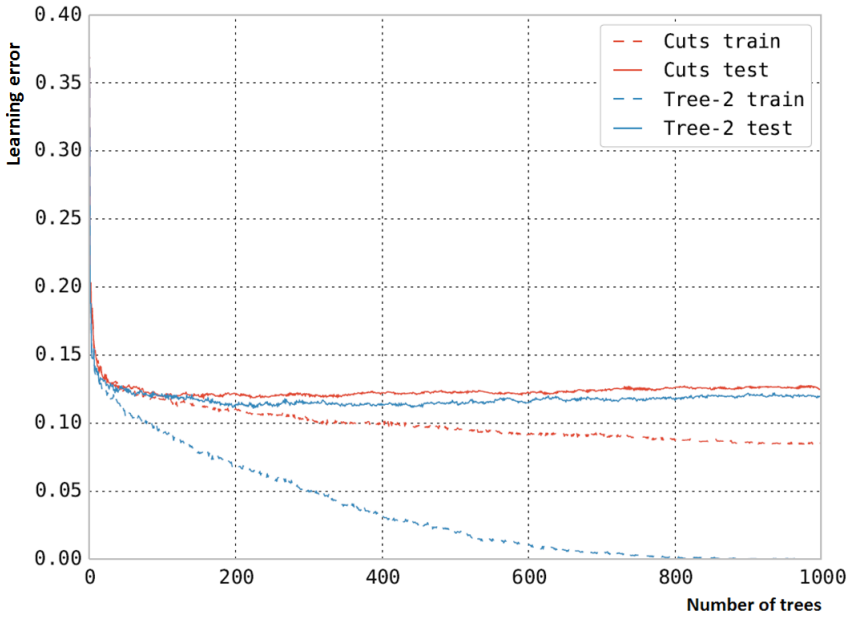


Figure 6.4: Learning errors for the BDT with two nodes, "Tree-2", (blue) and one node, i.e. rectangular cuts, "Cuts", (red). Curves for the training are presented by the dashed line, while the solid line corresponds to tests.

6.1.3 BDT with TMVA

The Toolkit for Multivariate Analysis (TMVA) [93] provides a ROOT-integrated environment for the processing, parallel evaluation and application of multivariate classification. All multivariate techniques in TMVA make use of training events, for which the desired output is known, to determine the mapping function that describes a decision boundary. The TMVA is specifically designed for the needs of high-energy physics (HEP) applications.

The main parameters of the BDT are:

- Training parameters, i.e. maximal size of an individual tree (depth), number of trees, boosting parameters for definition of a loss function $L(F, y)$ etc.
- Input variables. It is essential to use variables, characterizing the signal, that are distributed differently for signal and background (Figure 2.35-2.37). If the correlations between input variables are non-linear, the MVA will benefit in performance.

- Categories. Category methods allow sub-regions of phase space to be specified, which define disjoint sub-populations of the training sample, thus improving the modelling of the training data.

The performance of a BDT or any MVA method can be represented by Receiver Operating Characteristic (ROC) diagrams, which in their standard form show background rejection vs. signal efficiency for various cuts on the classifier produced by the MVA.

6.1.4 Application of MVA techniques: offline selection of the $D^{+*} \to \pi^+_{slow}(D^0 \to K^+\pi^-\mu^+\mu^-)$ decays

Our goal is to design an optimal selection for the $D^0 \to K^-\pi^+\mu^+\mu^-$, $D^0 \to K^+K^-\mu^+\mu^-$, $D^0 \to \pi^+\pi^-\mu^+\mu^-$ and $D^0 \to K^+\pi^-\mu^+\mu^-$ modes. For this purpose, we have tested ten MVA approaches using the TMVA package [93] plus one from the MultiBoost package [96], developed by the AppStat group of Laboratoire de l'Accélérateur Linéaire.

The decay we consider as the most interesting is $D^0 \to K^+\pi^-\mu^+\mu^-$, the so-called wrong sign (WS) mode, since the kaon has the opposite sign to that expected in a Cabibbo favoured decay. Of the modes considered, it is expected to be the most suppressed decay in the framework of the SM, and thus the most sensitive to NP among decays studied in this thesis. For that reason, it plays a central role in this study. We aim to measure the total branching ratios of these modes, including the resonant regions, which might be large enough to be measurable. Therefore we maximise here the signal significance.

Tests of each MVA are performed in two steps: first, training+testing of the MVA; and second, the scan over the MVA score, the BDT, vs. muon identification variable, $ProbNNmu$[1], in order to find the optimal cut on these two variables, that provides the maximal significance. This procedure is described in the following sections.

Data and MC samples

For the testing and training of the different MVAs we use three samples:

- Monte Carlo sample[2] of 2M $D^0 \to K^+\pi^-\mu^+\mu^-$ generated events. After the preselection before the MVA training, presented in next subsection, we have $\sim$ 3600 events for training and testing, where this sample is used as a signal reference;

- Data sample of $D^0 \to K^-\pi^+\mu^+\mu^-$ decays. It is kinematically very similar to $D^0 \to K^+\pi^-\mu^+\mu^-$ and has higher branching ratio, as a signal peak can be fitted in 2012 data. This mode is supposed to be less interesting, thus it is less crucial to sacrifice part of this data for the training+testing of MVA. We use 50% of 2012 data sample for $D^0 \to K^-\pi^+\mu^+\mu^-$, also called the right sign (RS) mode:

  - 10% on the sample is used to provide a background reference for the training and testing of the MVA. We select a two dimensional side band in

[1] See definition in Section 2.9.2

[2] For LHCb experts: Sim06b, Reco 14, Beam4000GeV-MayJune2012-MagUp-Nu2.5-EmNoCuts

$(m(D^0); \Delta m = m(D^{+*}) - m(D^0))$ as it is shown in Figure 6.5. We have checked that the distributions of events passing this selection is similar between $D^0 \to K^-\pi^+\mu^+\mu^-$ and $D^0 \to K^+\pi^-\mu^+\mu^-$ modes. This sample contains about 6000 events after the preselection, described in the next subsection;

– 40% of the sample is used to find a point of maximal significance (see one of the next subsections).

The signal and background samples are separated into two halves for the training and the testing of the BDT.

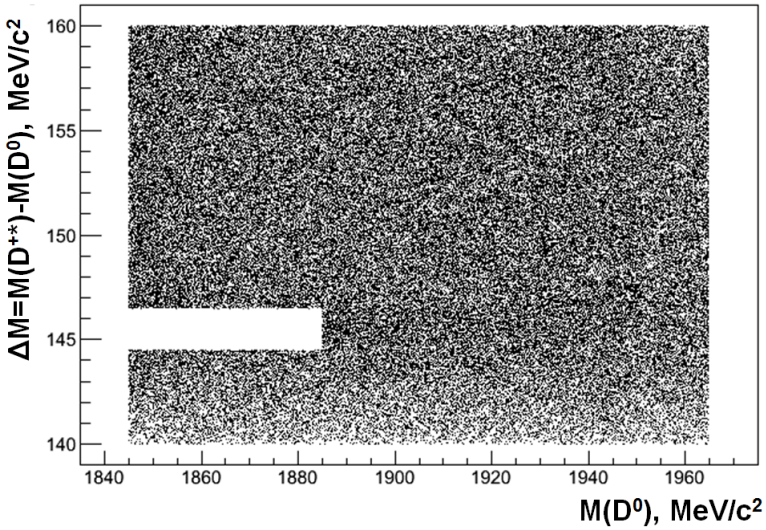


Figure 6.5: Illustration of the sideband used for the training of different MVA methods.

Preselection

A preselection is used to both the signal and background samples. The following cuts are applied at each stage:

- L0: `L0MuonDecision_TOS||L0HadronDecision_TOS`. These triggers are fired by muon and hadron candidates from the signal decay.
- HLT1: `Hlt1TrackMuonDecision_TOS||Hlt1TrackAllL0Decision_TOS`, lines that select muon of generic candidates with displaced secondary vertex.
- HLT2: `Hlt2CharmSemilepD02KPiMuMuDecision_TOS||` `Hlt2CharmSemilepD02PiPiMuMuDecision_TOS||` `Hlt2CharmSemilepD02KKMuMuDecision_TOS`, i.e. dedicated HLT2 trigger lines.
- Stripping selection, i.e. preselection (see Table 4.6 in Section 4.2).
- K and π: $PIDK(K^+) > 0$ and $PIDK(\pi^+) < 0$, i.e. identification selections.

- μ: $NShared = 0$, which means that the muon candidates cannot share hits in muon stations, and $isMuon = 1$, binary muon identification variable.
- Mass difference between D^{*+} and D: $\Delta m = m(D^{*+}) - m(D^0) \in (144.5; 146.5)$ MeV/c^2.

Cut optimization

To find a maximal signal significance we use 40% of the 2012 data sample of $D^0 \to K^-\pi^+\mu^+\mu^-$ decay. We perform a scan over BDT, the MVA output variable, vs. $ProbNNmu$, the muon identification variable. When requiring harder cut on BDT, more combinatorial background is removed. The $ProbNNmu$ cut removes peaking background from $D^0 \to K^-\pi^+\pi^+\pi^-$, where two pions can be misidentified as muons (see Section 5.7.4). For each cut on the pair $(BDT; ProbNNmu)$, as illustrated in Figure 6.6, we perform a fit, similar to that used in Section 5.7. This fit has three components: signal, peaking background and combinatorial background. Examples of the fit are also shown in Figure 6.6. We obtain the signal yield, N^{RS}_{signal}, as well as the fraction of combinatorial and peaking backgrounds under the D^0 mass peak, N^{RS}_{comb} and $N^{RS}_{peaking}$. The significance is determined as $N^{WS}_{signal}/\sqrt{N^{WS}_{signal} + N^{WS}_{comb} + N^{WS}_{peaking}}$, for the $D^0 \to K^+\pi^-\mu^+\mu^-$ mode, where each of the yields are scaled:

- $N^{WS}_{signal} = \frac{\mathcal{B}^{WS}_{signal}}{\mathcal{B}^{RS}_{signal}} \cdot N^{RS}_{signal} \cdot \kappa$;
- $N^{WS}_{peaking} = \frac{\mathcal{B}^{WS}_{peaking}}{\mathcal{B}^{RS}_{peaking}} \cdot N^{RS}_{peaking} \cdot \kappa$;
- $N^{WS}_{comb} = N^{RS}_{comb} \cdot \kappa$.

The coefficient κ corresponds to scale factor between 40% to 100% of the data sample.

This data-driven method is a priori more reliable than if the MC is used. We do not have to rely on the MC efficiencies matching these in data, nor on an imprecise prediction of the absolute $\mathcal{B}^{WS}$ since only the $\frac{\mathcal{B}^{WS}_{signal}}{\mathcal{B}^{RS}_{signal}}$ is necessary.

Comparison of the 11 MVA approaches

The training, testing and BDT vs. $ProbNNmu$ scans are performed for each of the 11 MVA approaches we have decided to study. The main chosen variables[3] are:

- for the daughter particles: p, p_t, IP χ^2, Track χ^2
- for the D^0 candidate: p, p_t, IP χ^2, Vertex χ^2, FD χ^2, DIRA, DOCA
- for the D^{+*} candidate: Vertex χ^2, DIRA, DOCA
- for the π^+_{slow} candidate: p_t, IP χ^2

[3]Definitions of the variables can be found in Section 2.11

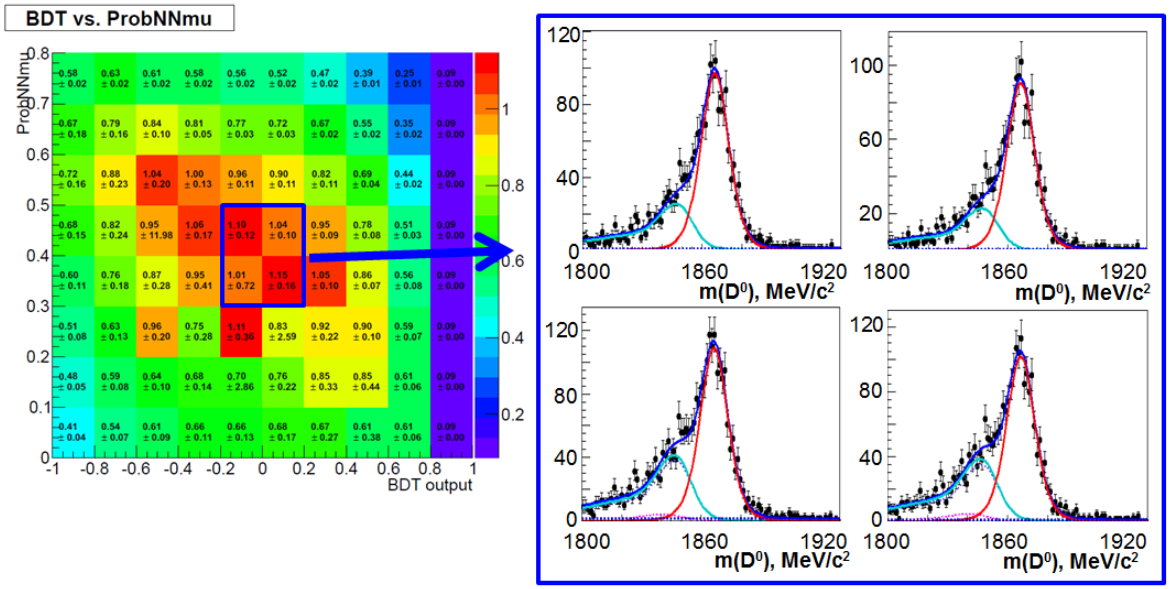


Figure 6.6: Illustration of the scan over BDT vs. $ProbNNmu$ parameters (on the left). For each level of cut a fit in real data is performed (on the right). The components of the fit: red - D^0 signal, cyan - peaking background, dashed blue - combinatorial background.

However, some of the methods use a different set of variables. These changes are pointed out for each method.

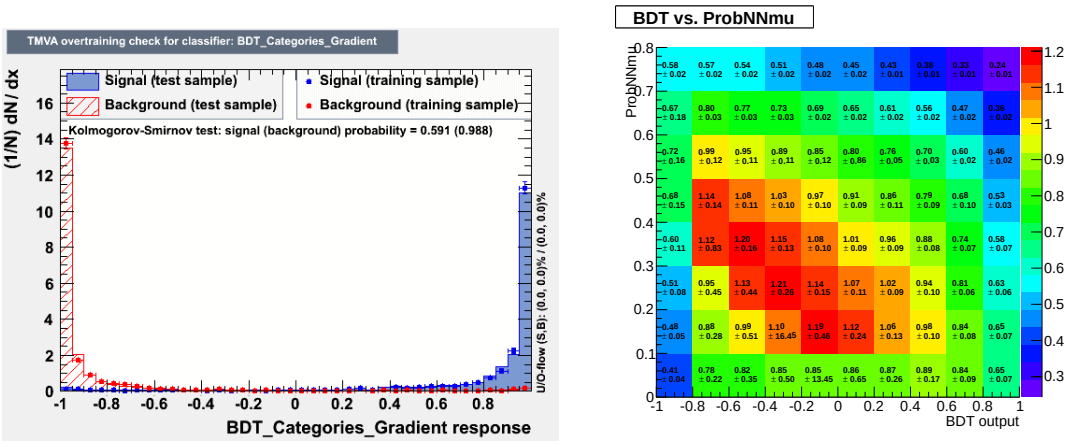
1. Rectangular cuts. The simplest and most common classifier for selecting signal events from a mixed sample of signal and background events is the application of an ensemble of rectangular cuts on discriminating variables. Unlike all other classifiers in TMVA, the cut classifier only returns a binary response (signal or background). The maximum significance obtained with this method is 0.72 ± 0.16.

2. Basic. This is a BDT with GradBoost boosting. It was adapted from the $D^+ \to \pi^+\mu^+\mu^-$ and $D^0 \to \pi^+\pi^-\mu^+\mu^-$ analyses. Training parameters are chosen to be $N_{Trees} = 800$, $Shrinkage$[4] $= 0.1$, $GradBaggingFraction$[5] $= 0.7$ and $N_{NodesMax} = 10$. This BDT has a category on $p_t^{asy} = 0$ or $p_t^{asy} \neq 0$, where p_t^{asy} is an isolation variable defined in Section 2.11. Due to this variable containing a delta peak at unity - when only the reconstructed mother is in the cone - the algorithms treat this circumstance differently from when there are other tracks from the event present in the cone.

An overtraining and significance scan are shown in Figure 6.7. One can see that maximum significance obtained is 1.21 ± 0.26.

3. Without ptasy. This is a BDT with GradBoost similar to the previous one but without categories on p_t^{asy} and some variables which contain less useful information, such as D^{+*} Vertex χ^2, DIRA and π^+_{slow} IP χ^2. The training parameters are chosen to

[4]Learning rate for GradBoost algorithm
[5]Defines the fraction of events to be used in each iteration

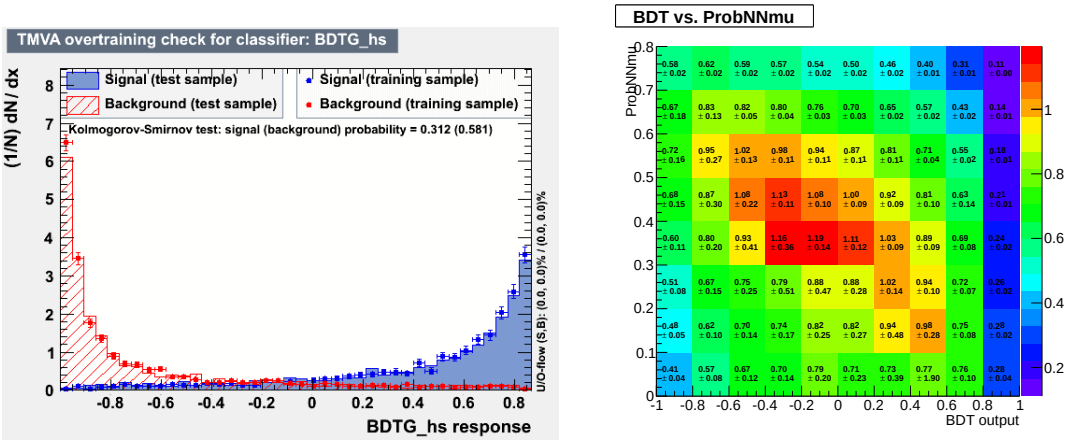


(a) Overtraining test (b) Scan over $(BDTCat; ProbNNmu)$

Figure 6.7: Tests of the "Basic" method.

be $N_{Trees} = 100$, $Shrinkage = 0.1$, $GradBaggingFraction = 0.7$ and $N_{NodesMax} = 5$.

An overtraining and significance scan are shown in Figure 6.8. One can see that maximum significance is 1.16 ± 0.12.



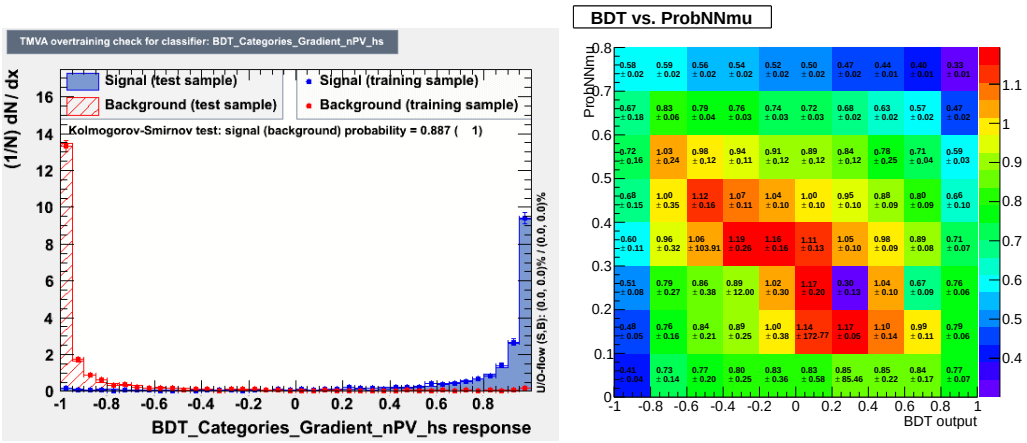
(a) Overtraining test (b) Scan over $(BDTCat; ProbNNmu)$

Figure 6.8: Tests of the "Without ptasy" method.

4. nPV. This is a BDT with GradBoost using the same variables as the previous one, "Without ptasy", but with categories on number of primary vertexes present in the event, nPV, $nPV = 1$ or $nPV > 1$. The events with only one primary vertex have lower multiplicity and as a consequence less combinatorial background. Thus, the treatment of such events independently could give higher performance. The training parameters are chosen to be $N_{Trees} = 800$, $Shrinkage = 0.1$, $GradBaggingFraction = 0.7$ are $N_{NodesMax} = 10$.

An overtraining and significance scan are shown in Figure 6.9. One can see that maximum significance is 1.19 ± 0.26.

5. newVarB. This is a BDT with GradBoost using the same variables as the "Without ptasy" training, but with two new variables added: the lifetime of the D^0, τ,

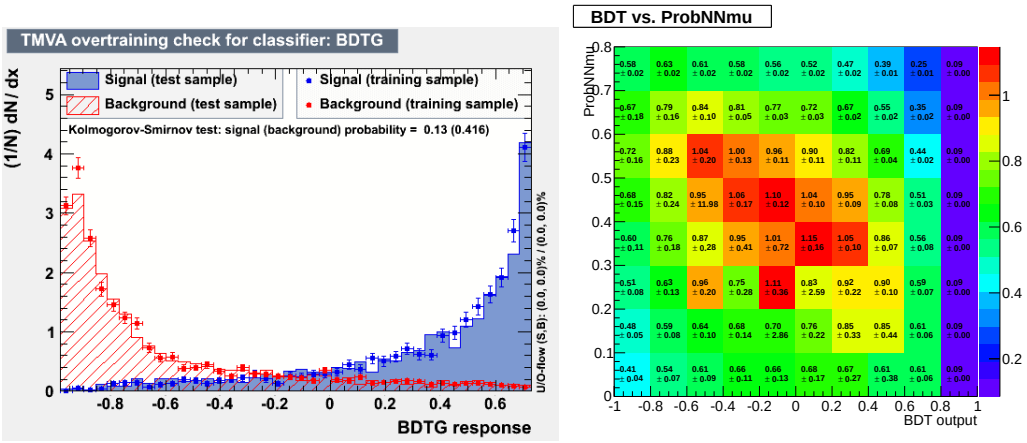


(a) Overtraining test (b) Scan over $(BDTCat; ProbNNmu)$

Figure 6.9: Tests of the "nPV" method.

and the decay tree fitted χ^2. The training parameters are chosen to be $N_{Trees} = 50$, $Shrinkage = 0.1$, $GradBaggingFraction = 0.7$ and $N_{NodesMax} = 4$.

An overtraining and significance scan are shown in Figure 6.10. One can see that maximum significance is 1.15 ± 0.16.

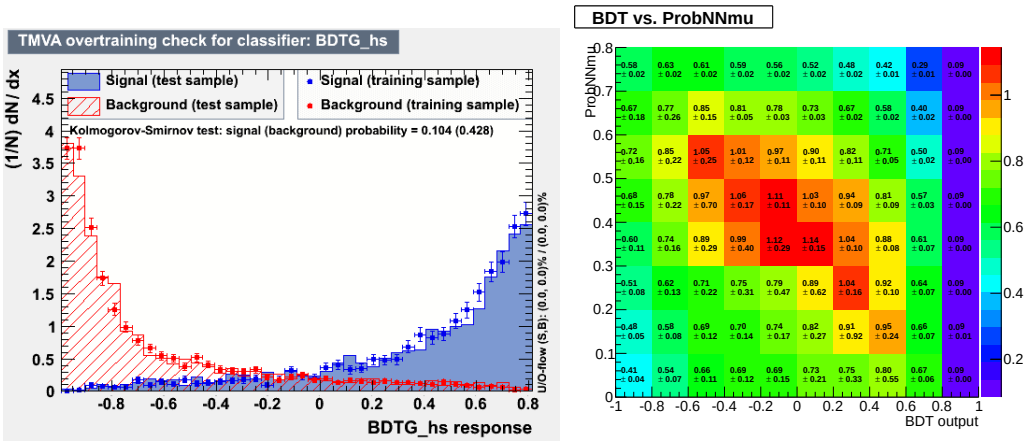


(a) Overtraining test (b) Scan over $(BDTCat; ProbNNmu)$

Figure 6.10: Tests of the "newVarB" method.

6. newVarW. This is a BDT with GradBoost using the same variables as the "Without ptasy" training, but with two new variables added. These are a goodness-of-fit, Match χ^2, quantifying the track matching between the VELO and TT station, for all daughter particles and a muon likelihood variable, $MuonMuLL$, for muon candidates. The inclusion of these variables was suggested by [97]. The training parameters are chosen to be $N_{Trees} = 70$, $Shrinkage = 0.1$, $GradBaggingFraction = 0.7$ and $N_{NodesMax} = 4$.

An overtraining and significance scan are shown in Figure 6.11. One can see that maximum significance is 1.14 ± 0.15.

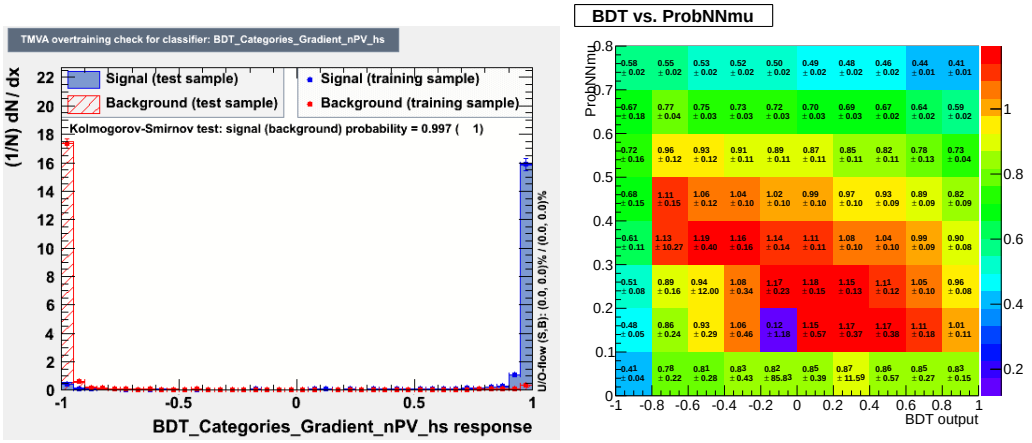


(a) Overtraining test (b) Scan over $(BDTCat; ProbNNmu)$

Figure 6.11: Tests of the "newVarW" method.

7. nPVnewVarB. This is a BDT with GradBoost using the same variables as the "newVarB" training, but with categories on the number of reconstructed primary vertices in the event, nPV, $nPV = 1$ or $nPV > 1$. The training parameters are chosen to be $N_{Trees} = 2500$, $Shrinkage = 0.1$, $GradBaggingFraction = 0.7$ and $N_{NodesMax} = 10$.

An overtraining and significance scan are shown in Figure 6.12. One can see that maximum significance is 1.19 ± 0.40.



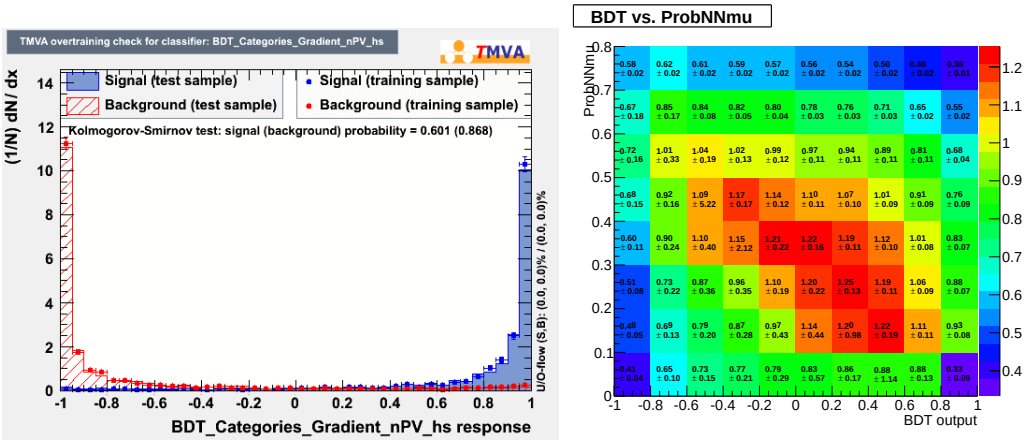
(a) Overtraining test (b) Scan over $(BDTCat; ProbNNmu)$

Figure 6.12: Tests of the "nPVnewVarB" method.

8. nPVnewVarW. This is a BDT with GradBoost using the same variables as the "newVarW" training, but with categories on nPV, $nPV = 1$ or $nPV > 1$. The training parameters are chosen to be $N_{Trees} = 1000$, $Shrinkage = 0.1$, $GradBaggingFraction = 0.7$ and $N_{NodesMax} = 10$.

An overtraining and significance scan are shown in Figure 6.13. One can see that maximum significance is 1.25 ± 0.13.

9. AdaB. This is a BDT with AdaBoost using the same variables as the "Without ptasy" training, but with the AdaBoost method instead of GradBoost. The training

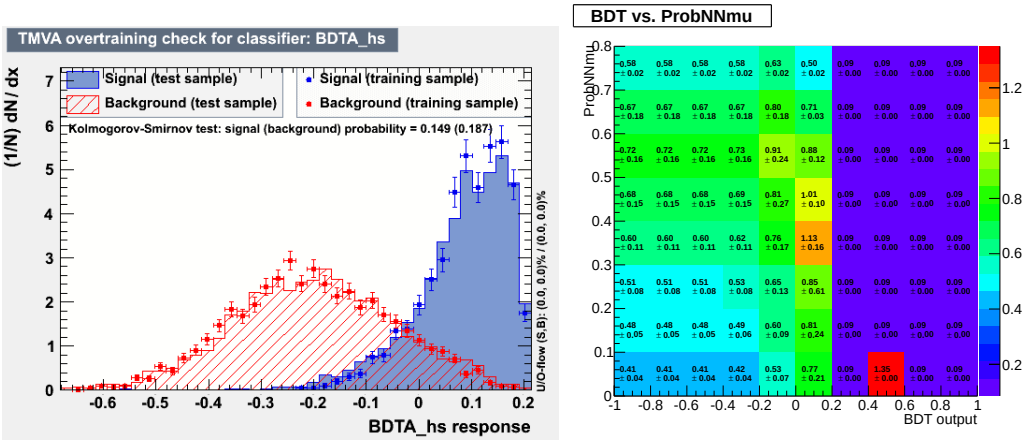


(a) Overtraining test (b) Scan over $(BDTCat; ProbNNmu)$

Figure 6.13: Tests of the "nPVnewVarW" method.

parameters are chosen to be $N_{Trees} = 500$, $AdaBoostBeta$[6] $= 0.5$ and $N_{NodesMax} = 6$.

An overtraining and significance scan are shown in Figure 6.13. One can see that maximum significance is 1.13 ± 0.16.



(a) Overtraining test (b) Scan over $(BDTCat; ProbNNmu)$

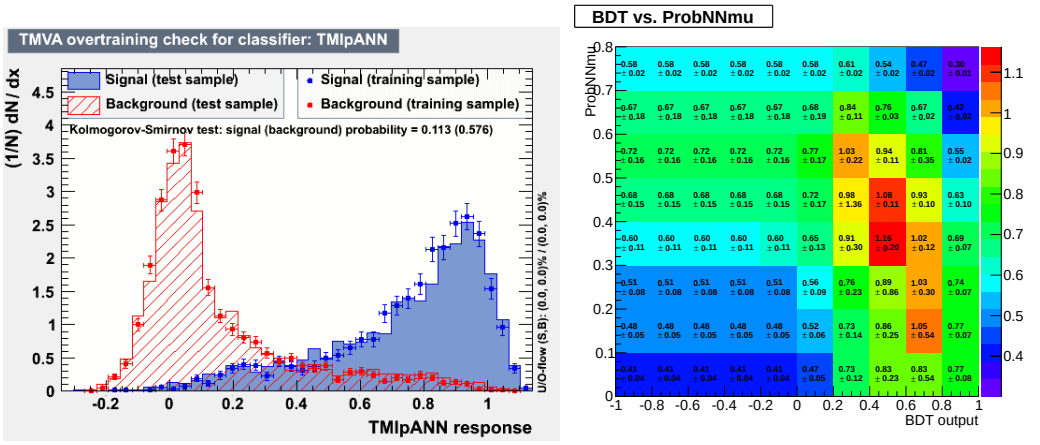
Figure 6.14: Tests of the "AdaB" method.

10. NN. This is an Artificial Neural Network with Multilayer perceptron using the same variables as the "without ptasy" training. The multi-layer perceptron is the type of neural network, where, in order to reduce complexity, connections between neurons are organized in layers and only direct connections a given layer to the following layer are allowed. The first layer is the input layer, the last one the output layer, and all others are *hidden layers*. For a classification problem with n input variables the input layer consists of n neurons that hold the input values, $x_1; ...; x_n$, and one neuron in the output layer that holds the output variable, the neural net estimator (Figure 6.2). The

[6]Learning rate for AdaBoost algorithm. This parameter is given as an exponent to the boost weight $\alpha \to \alpha^\beta$.

training parameters are chosen to be $N_{Cycles} = 50$ and $HiddenLayers = N$[7]

An overtraining and significance scan are shown in Figure 6.15. One can see that maximum significance is 1.16 ± 0.20.

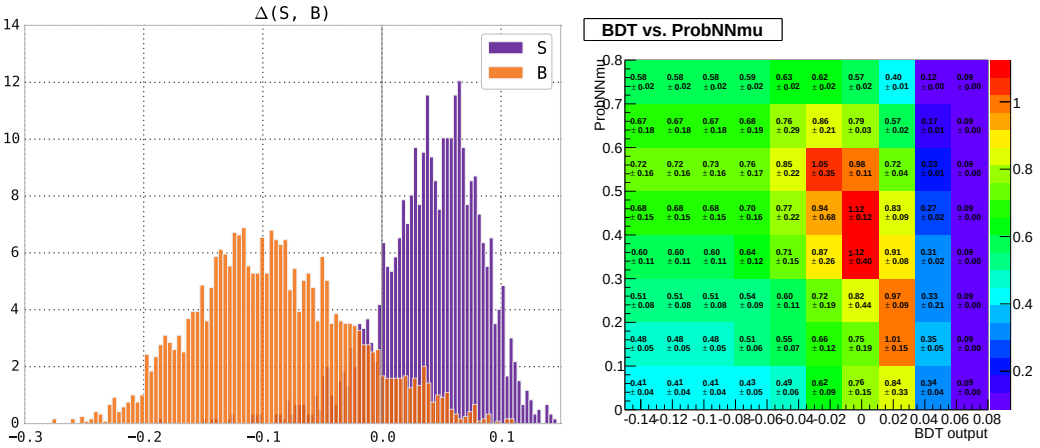


(a) Overtraining test (b) Scan over $(BDTCat; ProbNNmu)$

Figure 6.15: Tests of the "NN" method.

11. MultiBoost. This is a BDT based on MultiBoost package [96] developed by the AppStat group at LAL. The variables used are the same as in "nPVnewVarB" approach. It is a BDT with AdaBoost with the training parameters: $N_{Trees} = 500$, $AdaBoostBeta$[8] $= 0.5$ and $N_{NodesMax} = 6$.

An overtraining and significance scan are shown in Figure 6.16. One can see that maximum significance is 1.12 ± 0.12.



(a) Scores distribution (b) Scan over $(BDTCat; ProbNNmu)$

Figure 6.16: Tests of the "MultiBoost" training.

[7] A network with one hidden layer with $n = n_{variables}$ neurons

[8] Learning rate for AdaBoost algorithm

6.1.5 Results

The ROC curves, i.e. the signal efficiency vs. background rejection, for the eleven tested methods are presented in Figure 6.17 and 6.18. One can see from this figure that the rectangular cut method is less performant than BDT and ANN. The most performant methods are "nPVnewVarB", "nPVs" and "Basic". The various methods seem to yield very different signal significances, according to Figure 6.17. However, if we restrict our study to cuts for which the signal efficiency is below 85% (measurements with a lower background rejection would be very hard !), the difference is limited. This is confirmed by Table 6.1, which contains the maximal significances obtained by the scan for each method. It shows that the variation in significance is not decisive. Using nPV or p_t^{asy} categories might not compensate for the additional systematic uncertainty due to the fact these variables are not easy to describe by the MC. Only the result of the rectangular cuts approach is again lower and far from the best methods.

A similar conclusion is reaches from inspecting Table 6.1, which contains the maximal significances, obtained from the scans of the tested methods. The values in Table 6.1 do not differ much BDT and ANN methods while the significance of the rectangular cuts is again lower.

Note that we have checked that each method has been tuned in an optimal way by studying the distributions of the input variables before and after applying a cut on the MVA output variable (see for instance Figure 6.18).

Consequently, all MVA methods perform similarly on the current data. A possible explanation for this fact is that once the trigger and stripping selections (which are already quite tight) have been applied, not enough margin is left to fully exploit the potential of the various MVA's. It is supported by a study with the MultiBoost package (Figure 6.4). In this plot the error rate for the testing sample stops decreasing already after $\sim$ 220 trees. This indicates that it is rather straightforward for the MVA to perform the optimisation, and suggests that all methods should manage to converge to the same conclusion.

For further improvement we need more information and complexity, for example, the raw information from the muon system used for $ProbNNmu$ could be included directly in the MVA to optimize the muon PID for the specific case we are interested in. In order to gain in performance, MVA methods can also be used not only at the last step of selection, but also at the stripping and trigger level.

The maximal significance working point obtained with "newVarB" approach is used for the sensitivity studies for $D^0 \to K^+\pi^-\mu^+\mu^-$, $D^0 \to K^-\pi^+\mu^+\mu^-$ and $D^0 \to K^+K^-\mu^+\mu^-$ modes, that can be expected during future LHCb runs. These studies are discussed in Section 6.2. The choice of "newVarB" is dictated by the fact that this approach uses variables expected to be well reproduced by Monte Carlo.

	Name of training method	Significance
1	"Rectangular cuts"	0.72 ± 0.16
2	"Basic"	1.21 ± 0.26
3	"Without ptasy"	1.16 ± 0.12
4	"nPV"	1.19 ± 0.26
5	"newVarB"	1.15 ± 0.16
6	"newVarW"	1.14 ± 0.15
7	"nPVnewVarB"	1.19 ± 0.40
8	"nPVnewVarW"	1.25 ± 0.13
9	"AdaB"	1.13 ± 0.16
10	"NN"	1.16 ± 0.20
11	"MultiBoost BDT"	1.12 ± 0.12

Table 6.1: The maximal signal significance from different training methods.

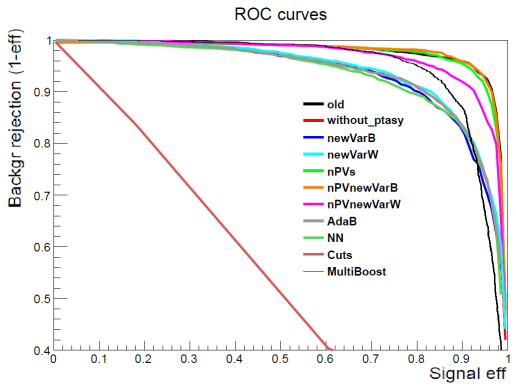


Figure 6.17: Superimposed ROC curves for all tested methods.

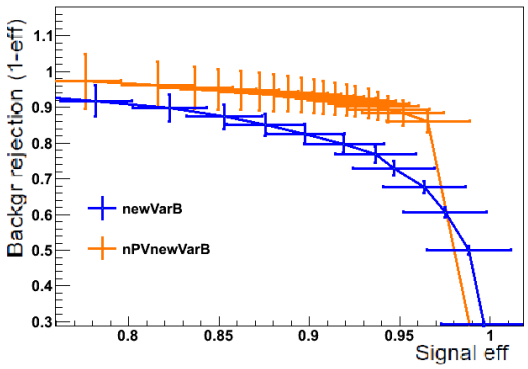
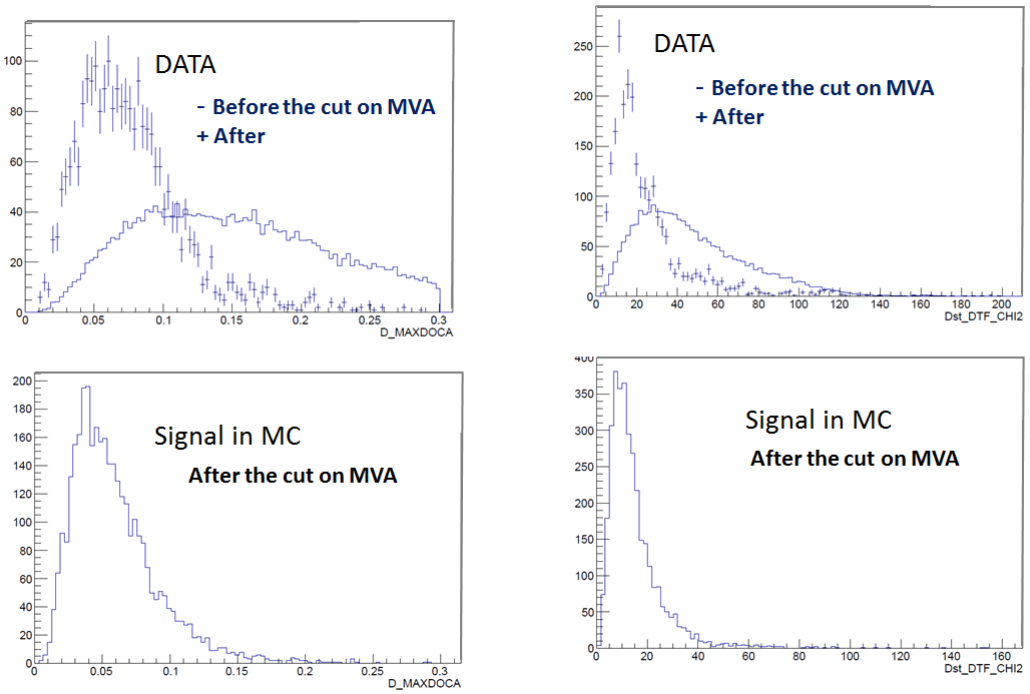


Figure 6.18: Two superimposed ROC curves plotted with error bars.



(a) DOCA(D^0) (b) DTF $\chi^2(D^0)$

Figure 6.19: Example of two distributions of discriminative variables, DOCA and DTFχ^2, in data before and after the MVA selection (on the top) and in signal Monte Carlo (on the bottom). The distributions in data change their forms drastically and become very signal-like. The BDT "nPVnewVarB" is used.

6.2 Sensitivity to branching ratios and asymmetries in $D^0 \to h^- h^{(\prime)+} \mu^+ \mu^-$ decays

In Section 6.1.4, we chose the muon identification criteria and the cut on the BDT discriminator that maximise the sensitivity to the total $D^0 \to K^+\pi^-\mu^+\mu^-$ branching ratio. This analysis is also focused on partial branching ratios in regions of the $m(\mu^+\mu^-)$ spectrum where the contribution of vector resonances, like $\rho \to \mu^+\mu^-$, is minimum. An upper limit will eventually be set in these regions. However, we consider that in the framework of this thesis, it is more interesting to evaluate LHCb's sensitivity to branching ratios that would trigger an interest in the theoretical community.

We study the evolution of the sensitivity to the total and partial branching ratio of $D^0 \to K^-\pi^+\mu^+\mu^-$, $D^0 \to K^+\pi^-\mu^+\mu^-$ and $D^0 \to K^+K^-\mu^+\mu^-$, as well as to asymmetries in these channels, for the three data taking conditions: LHCb's Run I, Run II and Upgrade. During Run I, 2 fb^{-1} of proton-proton data were collected in 2012 at energy $\sqrt{s} = 8$ TeV. The goal of Run II is to obtain 5 fb^{-1} at $\sqrt{s} = 13$ TeV. The Upgrade of LHCb [68] aims to collecting of 50 fb^{-1} at $\sqrt{s} = 14$ TeV.

The $D^0 \to \pi^+\pi^-\mu^+\mu^-$ mode is not studied here. It is left aside since upper limits have already been set by LHCb in 2013 [14].

6.2.1 Sensitivity to the total and partial branching ratios

In Section 6.1.4, eleven MVA approaches were compared, all yielding an offline selection of equivalent performance. We selected "newVarB" as the working hypothesis, as it uses variables which are expected to be reasonably well described in simulation. The scan shown in Figure 6.10 indicates the sensitivity to $\mathcal{B}(D^0 \to K^+\pi^-\mu^+\mu^-)$ that we can reach with this selection with LHCb's Run I data. Applying $BDT > -0.2$ and $ProbNNmu > 0.4$ criteria on top of the preselection described in Section 6.1.4, a significance of 1.15 ± 0.16 can be reached. Unless the branching ratio in [18] is underestimated, than more data will be necessary to measure it.

Total branching ratios

The sensitivities to the total branching ratios are extrapolated with obtained one in Section 6.1.4 for the Run I data, i.e. 2 fb^{-1} at $\sqrt{s} = 8$ TeV. The signal and background yields fitted to the $D^0 \to K^-\pi^+\mu^+\mu^-$ sample to perform the scans presented in Section 6.1.4 are re-used to evaluate the significance that can be reached with the Run II and Upgrade datasets. The signal and peaking background yields are rescaled according to

$$f_{\mathrm{R_I}\to\mathrm{R_{II}}} = \frac{5 \text{ fb}^{-1}}{2 \text{ fb}^{-1}} \cdot \frac{13\ TeV}{8\ TeV} \text{ and } f_{\mathrm{R_I}\to\mathrm{Upgrade}} = \frac{50 \text{ fb}^{-1}}{2 \text{ fb}^{-1}} \cdot \frac{14\ TeV}{8\ TeV}. \tag{6.3}$$

These factors account for the increment in integrated luminosity and for the fact that $\sigma(pp \to c\bar{c})$ is almost linear with the energy of the proton-proton collision. The same factors are applied as well to the combinatorial background. Indeed, we assume that is it

dominated by $D \to K + X$ decays (this is to some extent verified in Section 5.7.3, where building an artificial a background according to the assumptions seems to reproduce the background in data well).

In the case of $D^0 \to K^+K^-\mu^+\mu^-$, the signal and $D^0 \to K^+K^-\pi^+\pi^-$ peaking background yields expected in the future LHCb datasets are derived from the fit to the $D^0 \to K^-\pi^+\mu^+\mu^-$ sample, used in Section 6.1.4. This extrapolation is for each point of the BDT vs. $ProbNNmu$ scan, where the relative branching ratios and efficiencies between the $D^0 \to K^-\pi^+\mu^+\mu^-$ and $D^0 \to K^+K^-\mu^+\mu^-$ modes are taken in account. We assume $\mathcal{B}(D^0 \to K^+K^-\mu^+\mu^-) = 1.1 \times 10^{-7}$ (see Table I in [18]).

Partial branching ratios

As explained in Section 1.2.4, NP effects could be discovered by measuring Short Distance (SD) branching ratios higher than the SM expectations. The analysis of D^*-tagged $D^0 \to h^\pm h^{(\prime)\mp}\mu^+\mu^-$ decays with 2 fb^{-1} will set an upper limit on them. Interpreting these limits in terms of constraints on NP models will be a difficult theoretical challenge since it requires a precise determination of the SM contributions. We believe that more will be learned if a signal is observed instead of an upper limit. Based on Section 1.2.4, we expect that most NP models are unlikely to generate more than $\mathcal{B}^{SD}(D^0 \to h^\pm h^{(\prime)\mp}\mu^+\mu^-) \sim 10^{-8}$. We evaluated LHCb's sensitivity to this branching ratio.

$D^0 \to K^+\pi^-\mu^+\mu^-$ decays

Short Distance branching ratios can be accessed only in regions of the $m(\mu^+\mu^-)$ spectrum where the Long Distance (LD) contributions due to η, ρ, ω, η' or $\phi \to \mu^+\mu^-$ are minimal. A partial branching ratio can therefore be measured in the region below the η region: $m(\mu^+\mu^-) < 525$ MeV/c^2. Assuming a phase space decay model, 45% of $D^0 \to K^+\pi^-\mu^+\mu^-$ decays lie in this region. This means the accessible branching ratio would be $\sim 5 \times 10^{-9}$. A generator level simulation shows that this region contains only 6.5% of the $\rho \to \mu^+\mu^-$ decays. Assuming the value of $\mathcal{B}^{SM}(D^0 \to K^+\pi^-\mu^+\mu^-)$ is dominated by $D^0 \to K^+\pi^-\rho(\to \mu^+\mu^-)$, which is implied in [18], the corresponding contribution to the $m(\mu^+\mu^-) < 525$ MeV/c^2 region is $\sim 10^{-9}$. Thus, measuring $\mathcal{B}(D^0 \to K^+\pi^-\mu^+\mu^-)_{m(\mu^+\mu^-)<525\ MeV/c^2} \sim 5 \times 10^{-9}$ could not be attributed to a LD effect.

$D^0 \to K^+K^-\mu^+\mu^-$ decays

In the case of $D^0 \to K^+K^-\mu^+\mu^-$ decays, assuming the total branching ratio predicted in [18] is dominated by the $\rho \to \mu^+\mu^-$ contributions, the resonant branching ratio in this region is $\mathcal{B}(D^0 \to K^+K^-(\rho \to \mu^+\mu^-))_{m(\mu^+\mu^-)<525\ MeV/c^2} \sim 7 \times 10^{-9}$. Therefore, to attribute a signal to NP rather than this resonant contribution would require a branching ratio in this region of $\mathcal{B}(D^0 \to K^+K^-(\rho \to \mu^+\mu^-))_{m(\mu^+\mu^-)<525\ MeV/c^2} \sim 1.5 \times 10^{-8}$.

$D^0 \to K^-\pi^+\mu^+\mu^-$ decays

The signal and background yields fitted to the $D^0 \to K^-\pi^+\mu^+\mu^-$ sample to perform the scans presented in Section 6.1.4 are again used for a new scan. The signal yields are extrapolated to the $D^0 \to K^+\pi^-\mu^+\mu^-$ and $D^0 \to K^+K^-\mu^+\mu^-$ low-$m(\mu^+\mu^-)$ branching ratio discussed above: $\mathcal{B}^{SM}(D^0 \to K^+\pi^-\mu^+\mu^-)_{m(\mu^+\mu^-)<525\ MeV/c^2} \sim 5 \times 10^{-9}$ and $\mathcal{B}(D^0 \to K^+K^-(\rho \to \mu^+\mu^-))_{m(\mu^+\mu^-)<525\ MeV/c^2} \sim 1.5 \times 10^{-8}$. Also, the differences in efficiencies between full spectrum and low-$m(\mu^+\mu^-)$ regions are taken into account, as well as the fraction of the total signal and background yields that lie in this region.

Data	to total $\mathcal{B}$, σ	to partial $\mathcal{B}$, σ
2 fb^{-1}	~ 1.10	~ 0.5
5 fb^{-1}	~ 2.4	~ 1
50 fb^{-1}	~ 8	~ 3

Table 6.2: Sensitivity, expressed in σ of signal significance, to total and partial branching ratios for the $D^0 \to K^+\pi^-\mu^+\mu^-$ decay.

Data	to total $\mathcal{B}$, σ	to partial $\mathcal{B}$, σ
2 fb^{-1}	~ 4.6	~ 1.2
5 fb^{-1}	~ 9.2	~ 2.4
50 fb^{-1}	~ 30	~ 7.8

Table 6.3: Sensitivity, expressed in σ of signal significance, to total and partial branching ratios for the $D^0 \to K^+K^-\mu^+\mu^-$ decay.

The obtained sensitivities to total and partial branching ratios for $D^0 \to K^+\pi^-\mu^+\mu^-$ and $D^0 \to K^+K^-\mu^+\mu^-$ decays are presented in Tables 6.2 and 6.3 assuming $\mathcal{B}(D^0 \to K^+\pi^-\mu^+\mu^-) = 1.7 \sim 10^{-8}$ and $\mathcal{B}(D^0 \to K^+K^-\mu^+\mu^-) = 1.1 \times 10^{-7}$ [18]. From these Tables one can conclude that:

- with Run I data it is possible to observe the total $\mathcal{B}(D^0 \to K^+K^-\mu^+\mu^-)$ decays;
- with Run II data it is might be possible to observe the total $\mathcal{B}(D^0 \to K^+\pi^-\mu^+\mu^-)$ and the partial $\mathcal{B}(D^0 \to K^+K^-\mu^+\mu^-)$ while the total $\mathcal{B}(D^0 \to K^+K^-\mu^+\mu^-)$ branching fraction should be measured;
- with Upgrade data the measurements of total and partial branching fractions for $D^0 \to K^+K^-\mu^+\mu^-$ should be achieved. The total $\mathcal{B}(D^0 \to K^+\pi^-\mu^+\mu^-)$ should also be measured, while the partial $\mathcal{B}(D^0 \to K^+\pi^-\mu^+\mu^-)$ has a chance to be observed.

6.2.2 Sensitivity to asymmetries

In this section we evaluate the statistical uncertainty with which asymmetries can be measured in the two modes, $D^0 \to K^+\pi^-\mu^+\mu^-$ and $D^0 \to K^+K^-\mu^+\mu^-$. Generically, an asymmetry is expressed this way:

$$\mathcal{A} = \frac{N_1 - N_2}{N_1 + N_2}. \tag{6.4}$$

By simple error propagation, one obtains:

$$\sigma_{\mathcal{A}} = \frac{2}{(N_1 + N_2)^2}\sqrt{N_1^2\sigma_{N_2}^2 + N_2^2\sigma_{N_1}^2} \tag{6.5}$$

Assuming a small asymmetry, $N_1 \sim N_2$ is about half the signal yield found in a given sample. In practice, $N_{1,2}$ are taken as half the signal yields with which the best significances were obtained in the Section 6.2.1. The $\sigma_{N_{1,2}}$ uncertainties are evaluated using fits to toy samples. This samples comprise three contributions which shapes are the same as those found in the $D^0 \to K^-\pi^+\mu^+\mu^-$ sample for the signal, hadronic peaking background and combinatorial background. The generated background yields are also provided from the scans performed in Section 6.2.1.

Based on theoretical predictions [18, 46, 47, 98], we conclude that it is not necessary to restrict asymmetry measurements to a particular region of the $m(\mu^+\mu^-)$ spectrum. Therefore, evaluated sensitivities are obtained for the signal, including even the resonant contributions.

The summary of these studies is presented in Table 6.4. We include only the cases where the predicted uncertainties on asymmetries is at the $\sim 10\%$ or better. Based on these numbers, we conclude that:

- with 2 fb^{-1} only asymmetries in $D^0 \to K^-\pi^+\mu^+\mu^-$ decays could be observed. A priori, this mode is not the most promising to detect NP since it has the largest SM branching ratio among studied modes. However, the T-odd asymmetry predicted in [18], 7%, is the same for all the $D^0 \to h^\pm h^{(\prime)\mp}\mu^+\mu^-$ modes. Any NP effects involved here are expected to follow the CKM hierarchy of the SM.

- with 5 fb^{-1} asymmetries of the order of 5% should become measurable in $D^0 \to K^-\pi^+\mu^+\mu^-$.

- with 50 fb^{-1}, asymmetries of the order of 1-2% should be possible to measure in the $D^0 \to K^-\pi^+\mu^+\mu^-$ mode. For $D^0 \to K^+K^-\mu^+\mu^-$ it seems possible to measure asymmetries of the order of 15%. In case of $D^0 \to K^+\pi^-\mu^+\mu^-$ modes only asymmetries of the order of 50% seem detectable.

It is possible to improve the signal vs. background ratio if the asymmetries are studied not in a full dimuon mass spectrum, but only around resonances, like $D^0 \to \phi(\to K^+K^-)\mu^+\mu^-$, where according to [18] the dominate contribution to the total BF is. Cutting around the ϕ mass also reduces by a factor 3 the background yields.

Besides the higher luminosity and cross sections, LHCb's upgrade will bring additional improvements. The L0 trigger will be replaced by a software trigger. The online reconstruction should also become closer to offline one. If enough priority is given to these modes, it will be possible to use HLT lines as efficient as the rest of the offline selection. In MC, we observe that removing the L0 and HLT requirements improves the efficiency by a factor 2.3 in case of $D^0 \to K^+\pi^-\mu^+\mu^-$ and 3 in case of $D^0 \to K^+K^-\mu^+\mu^-$. Note that another way to improve the available statistics is to include D^0's produced in beauty hadron decays. However, the corresponding gain remains to be evaluated.

A possibility of improvement lies in the branching ratios predicted in [18]. They are a subject to an important theoretical uncertainty. The prediction for $D^0 \to K^-\pi^+\mu^+\mu^-$, $\mathcal{B}(D^0 \to K^-\pi^+\mu^+\mu^-) = 6.2 \times 10^{-6}$, assumed this mode is by far dominated by $D^0 \to K^-\pi^+(\rho \to \mu^+\mu^-)$. Due to the large width of the ρ, this correspond to $\mathcal{B}(D^0 \to$

$K^-\pi^+\mu^+\mu^-)_{m(\mu^+\mu^-)\in[675;875]MeV/c^2} \sim 4 \times 10^{-6}$ in the ρ/ω region. The measurement presented in Chapter 5 is consistent at 3 σ with 1.5 times this value. Assuming the $\mathcal{B}(D^0 \to K^+\pi^-\mu^+\mu^-)$ and $\mathcal{B}(D^0 \to K^+K^-\mu^+\mu^-)$ are actually underestimated by 50%, and combining this with the gains presented in the previous paragraph, we obtain the uncertainties in the bottom-most line of Table 6.4. In this case, $D^0 \to K^+\pi^-\mu^+\mu^-$ would be sensitive to asymmetries of the order of 25%, while $D^0 \to K^+K^-\mu^+\mu^-$ would be able to detect asymmetries of the order of 5%, which is in the regime of the prediction of several theoretical papers. Measuring asymmetries at the 1% level, however, will be difficult except in the Cabibbo favoured mode. In most of the multibody decay modes, it will also be difficult to measure NP-generated branching ratios at the level of a few 10^{-9}. In the future, searching for NP in rare charm modes, i.e. constraining C_7, C_9 and C_{10}, might rely more on the combination of all the available modes and observables rather than on individual measurements. It should also be noted that full angular analyses (as in the case of $B^0 \to K^{*0}\mu^+\mu^-$ [99]) could be more sensitive to NP than the mere asymmetries investigated in this thesis.

Data	$D^0 \to K^+\pi^-\mu^+\mu^-$	$D^0 \to K^-\pi^+\mu^+\mu^-$	$D^0 \to K^+K^-\mu^+\mu^-$
2 fb^{-1}		$\sim 2\%$	
5 fb^{-1}		$\sim 1\%$	$\sim 11\%$
50 fb^{-1}	$\sim 12\%$	$\sim 0.3\%$	$\sim 3.5\%$
50 fb^{-1} + improve	$\sim 6\%$	$\sim 0.15\%$	$\sim 1.5\%$

Table 6.4: Uncertainties on possible asymmetries in $D^+ \to h^+ h^{(\prime)-} \mu^+ \mu^-$ decays for three LHCb data taking conditions. Uncertainties higher than 15% are not shown.

Chapter 7

Test of the upgraded calorimeter front-end electronics

I participated in the beam tests of the prototype of the front-end board that will equip the HCAL and ECAL for the upgraded LHCb experiment. Both calorimeters use the same electronics, that has to be replaced.

During the Upgrade era the electromagnetic and hadron calorimeters will contribute for the particle identification and the Low Level Trigger (LLT), the new trigger which will filter the raw data events. For an efficient LLT selection the readout of the ECAL and HCAL must be performed at a rate of 40MHz. This constrains the requirements on the electronics of these two detector systems.

The new data taking conditions, discussed in Section 2.7, force the upgrade of the front-end electronics and a full revision of electronic architecture [100]. The consequences of the upgrade for the electronics are:

- the readout of the data is performed at 40 MHz;
- the gain of the PMT is reduced by a factor 5 in order to reduce ageing. It is compensated by an increase in the gain of the electronics.

This leads to the requirements for the calorimeter analogue Front-End boards (FEB) [100] listed in Table 7.1.

A total of 238 new FEB is necessary to perform the acquisition of the ECAL and HCAL modules. Each FEB is designed to receive the signals of 32 PMTs. The boards are distributed in 18 crates: 14 for the ECAL and 4 for the HCAL. A new control board located in the central slot of each crate must also be designed. It will receive the clock, slow and fast control signals from the Timing and Fast Control system (TFC) of the experiment and propagate them to the FEB via the backplane. This architecture is compatible with the parts of the existing electronics that will be kept for the upgrade (backplanes and crates, for example) in order to reduce the cost of the project.

The first FEB prototype was tested during a test beam carried out at CERN in November 2012, which is discussed in this section.

Parameter	Requirement
Energy range	$0 \leq E_t \leq 10$ GeV (ECAL)
Calibration/Resolution	4 fC/2.4 MeV per ADC count
Dynamic range	4096-356=3840: 12 bits
Noise	$\lesssim 1$ ADC cnt (ENC < 4 fC)
Termination	$50 \pm 5\ \Omega$
Baseline shift prevention	Dynamic pedestal subtraction
Max. peak current	4-5 mA over 50 Ω
Spill-over residue level	$\leq$ 1 %
Non-linearity	$<$ 1 %
Cross-talk	$<$ 0.5 %
Timing	Individual (per channel)

Table 7.1: Summary of the requirements for the calorimeter analogue Front-End.

7.1 Front-End Board

The FEB consists of two parts for analogue and digital processing. The analogue part discussed here performs the amplification of signal pulse, its integration and digitization. The digital part of FEB is composed of four major components with different functions:

- The Front-End-Programmable-Gate-Array (FE-PGA) processes 8 channels and receives data from 8 ADCs that provide a 12 bit digitalisation. It consists of six firmware blocks: data synchronisation, pedestal and trigger processing, event building, handling the test and TFC signals and sending the data to the GigaBit Transceiver (GBT);
- The Trigger-PGA (Trig-PGA) receives from the FE-PGA the data required to handle the processing of the Low Level Trigger on the board. The result is sent to the GBT;
- The Control-PGA (Ctrl-PGA) performs the tasks which are common to the channels of the FEB. It ensures that the four FE-PGAs run synchronously, provides the crate and FEB identification numbers to the links and distributes the slow control signals to the different components of the board;
- The power and slow control blocks.

The picture of the FEB prototype is presented in Figure 7.1. The analogue mezzanine is visible on the left. The two FPGAs are in the central part of the board, in black and under the copper socket.

The analogue signal processing on the FEB is mainly performed by a circuit that integrates the PMT signal pulse. The calorimeter measurements must be performed every 40 MHz, this means that the signal pulse must be integrated completely during 25 ns. The tail extending beyond 25 ns should be less than 1% of the signal.

Another requirement comes from the reduction of the PMT gain by a factor 5: it must be compensated by a similar enhancement of the electronics gain. This is

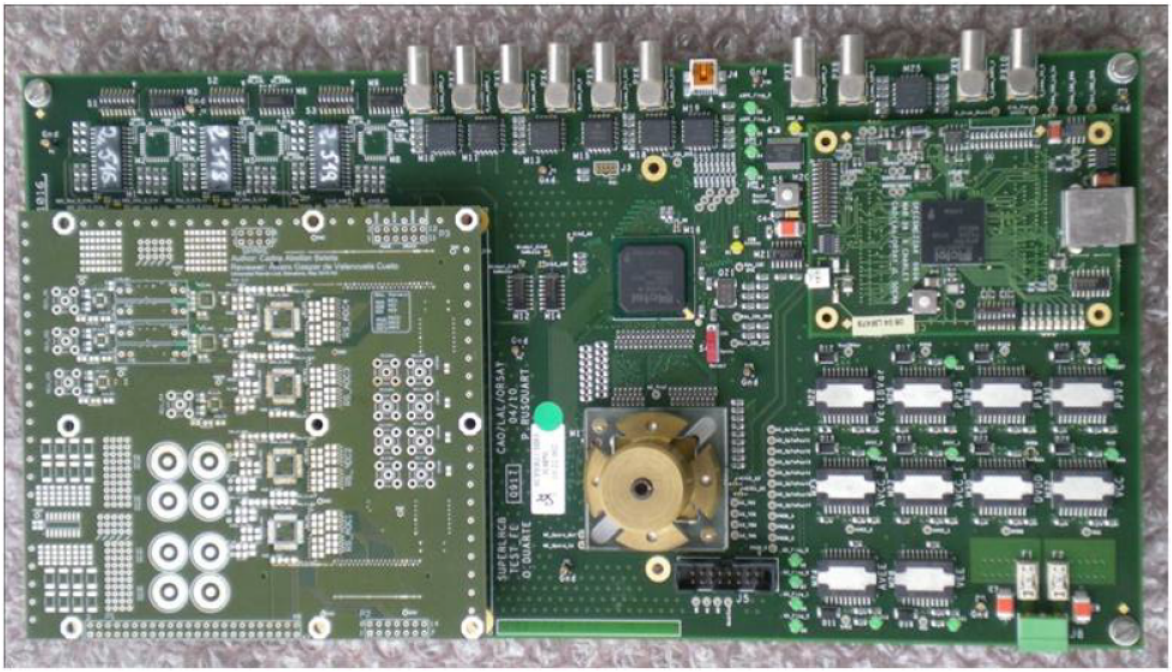


Figure 7.1: Picture of the prototype of the Front-End board. The analogue mezzanine is visible on the left. The two FPGAs are in the central part of the board, one is black and other is under the copper socket.

challenging as the electronic noise should be kept at the same level as for the current electronics although the amplification of the electronics is drastically increased.

Two analogue solutions are proposed: an Application-Specific Integrated Circuit (ASIC) [101] and a Commercial Off-The-Self (COTS) [102] solutions. The ASIC (Figure 7.2) performs the analogue processing. It amplifies the PMT pulse and injects the produced fully differential signal alternatively on two paths. Each path is connected to an integrator running at 20 MHz. While the active integrator is accumulating the charges from the amplified PMT pulse, the other signal is readout by a track and hold (TH) which presents the integrated value to the output and to the ADC. The readout integrator is reset before becoming active. The previously active integrator is then readout by the TH.

The COTS (Figure 7.3) also performs a fully differential processing (as for the ASIC) in order to match the requested noise performances. The signal is first amplified and injected to the integrator input with two polarities. The positive one is directly injected into the integrator. The negative one is delayed by 25 ns before being sent to the integrator input. The resulting signal (integrator output) is shown in Figure 7.4; it rises for 25 ns (positive polarity) reaches a plateau and decreases for 25 ns (negative polarity). This technique permits the integrator to be reset without switch which may induce noise in the system.

For both solutions, it is necessary to clip the signal which is originally wider than 25 ns. The *clipping* is an element that consist of a T[1], and a delay line or a cable. A

[1]Connection of three cables at one point

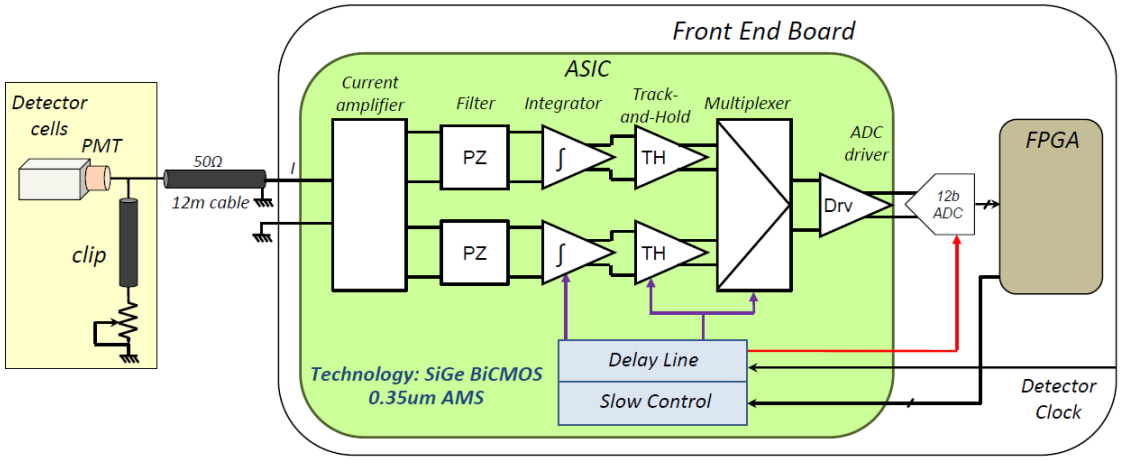


Figure 7.2: ASIC solution schematics.

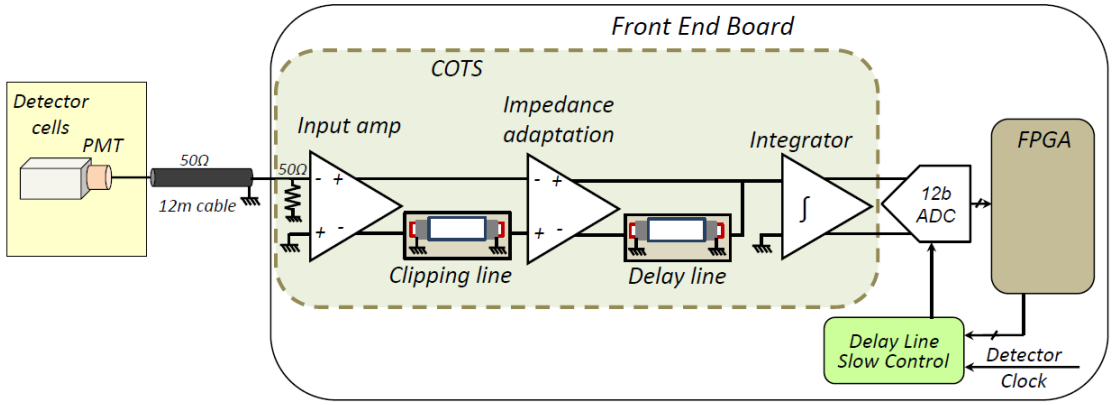


Figure 7.3: COTS solution schematics.

fraction of the signal goes to the clipping, is reflected and comes back with a delay and an additional phase, making the subtraction from the initial signal. During the clipping 2/3 of the charge is lost. Thus, the clipping element reduces approximatively the width of the signal to 25 ns. However, it induces some noise in the line. The ASIC solution has very good performances in terms of noise. Hence, it is not necessary to perform the amplification before the clipping (see results in Section 7.3.1). The noise performances of the COTS solution are not as good and the clipping needs to be moved after the amplification as seen in Figure 7.3. The PMT bases of the ECAL and HCAL include the clipping already. The ASIC solution does not request any modification of the PMT bases. Therefore, in the case that COTS solution is adopted an intervention will be needed on the ∼8000 PMT bases of the ECAL and the HCAL to remove the clipping. This intervention is simple (a line on the PMT base circuit to be cut), but will be time consuming. The clipping would be then integrated in the COTS design (see Figure 7.3).

The two prototypes were tested in the test beam, which took place in November 2012 at CERN.

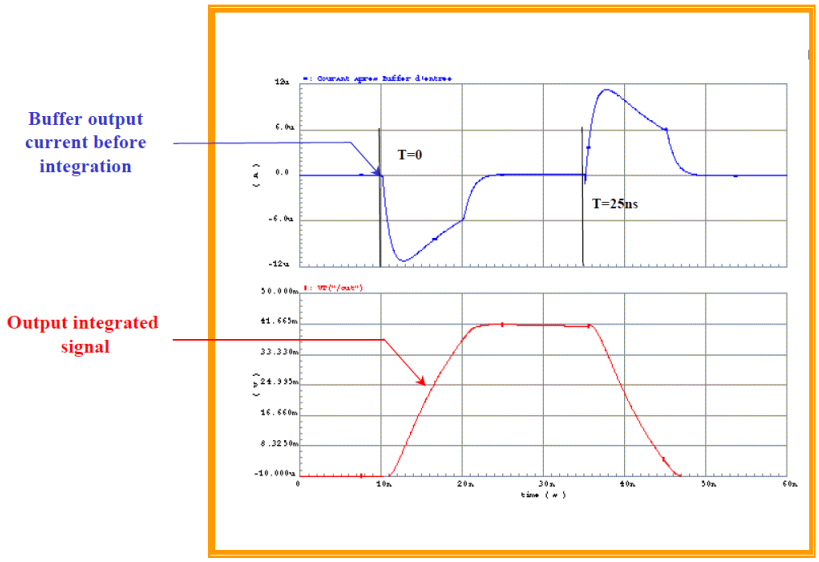


Figure 7.4: Simulated input and output pulses of the integrator at COTS solution.

7.2 Test beam setup

The aim of the beam test of the FEB is to check its performance characteristics, such as the noise level, linearity, plateau, spill over, etc. during realistic data taking conditions. A spare ECAL module was exposed to an electron beam. The modules contains also a Light-Emitting Diodes (LED).

The test beam was provided by the T4-H8 beam line at CERN, which delivers electrons with energies of 50 GeV, 100 GeV and 125 GeV. The scheme of the setup is presented in Figure 7.5. Electrons hit the module and the scintillation light is collected by the PMT. The obtained signal pulse is sent in parallel to two systems: a front-end prototype, integrating over 25 ns, and a Lecroy integrator, which has a long period of integration and cannot be used during LHCb data taking. The second integrator is required for the linearity tests. An adapted T is included so that the impedance seen from the output is correct.

The chronograph in Figure 7.6 shows the acquisition system. The trigger is produced by the coincidence of two scintillators placed on the beam axis in front of the module. It sends the positive decision to the Nuclear Instrumentation Module (NIM) crate, that transports it to the FEB and Lecroy ADC, and receives the signal "Busy" from these two systems during acquisition time; no other trigger can be fired during this period. To start the integration a 40MHz clock is used, nevertheless particles are unsynchronised with a clock phase. Therefore, a trigger is sent to the Time-to-Digital Converter (TDC) where the time/phase of electrons arriving with clock is stored. This provides the possibility to select the phase corresponding to the full integration of the signal, which is required for the plateau tests.

In the LHCb cavern the PMT and the FEB are connected by 15 m cables, which means that the signal can be reflected at connection points and come back as an input to the integrator in $\sim$ 75 ns after signal arrival. To estimate the level of reflected

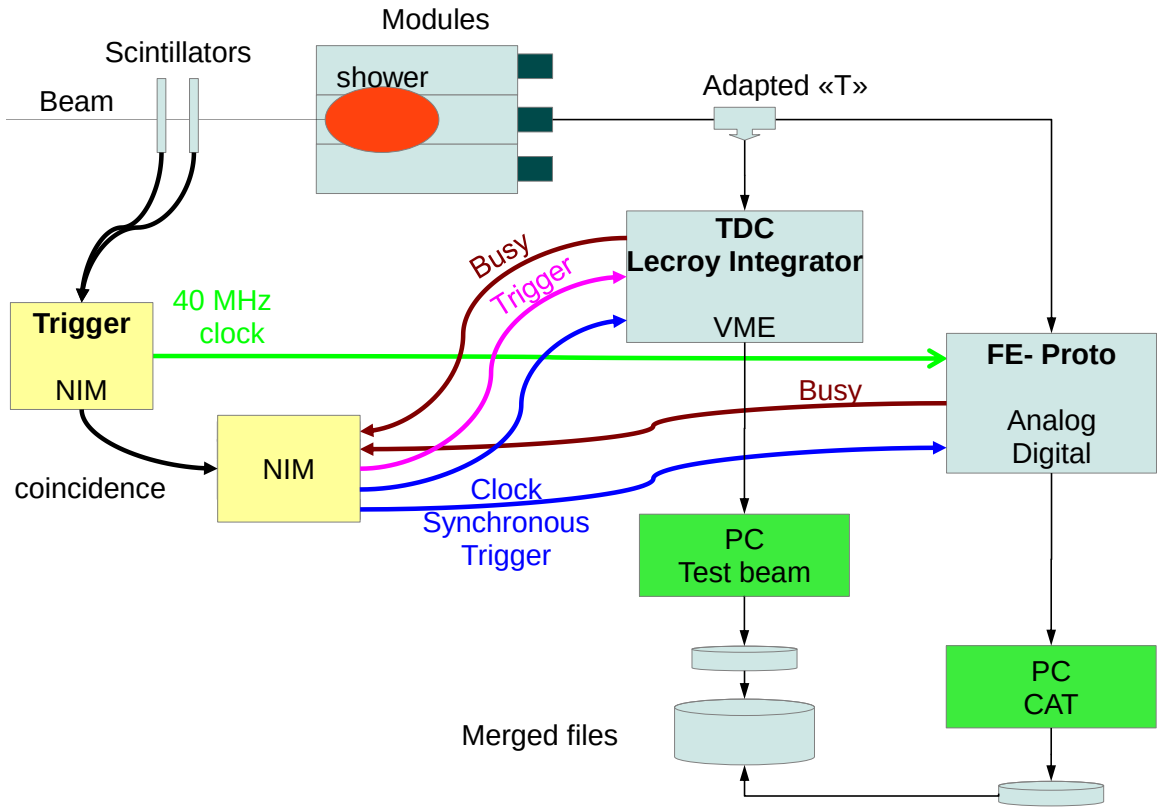


Figure 7.5: Test beam setup.

fraction the sequence of samples before and after the trigger was stored.

The tests of noise, linearity, integrator plateau, spill over and reflection are discussed below in this section. The results of the test beam exposed certain problems, discussed in Section 7.4 that could then be corrected in the Upgrade PID Technical Design Report (TDR) [100].

7.3 Test beam results

Firstly, low frequency noise has to be removed. For this purpose a pedestal subtraction is applied: the smaller of the two preceding samples before the signal is subtracted. For the test beam this was done offline. During data taking it will be done online in the digital processing.

7.3.1 Noise

The results of the noise study for the ASIC and for the COTS prototypes are plotted in Figure 7.7 and 7.8. After applying the pedestal subtraction the noise is reduced to 1.6 and 2.6 ADC counts for the ASIC and the COTS solutions, respectively. To check the fraction of the noise comes from the T connection and Lecroy ADC these components were removed from the setup and the noise level was measured again. From this study

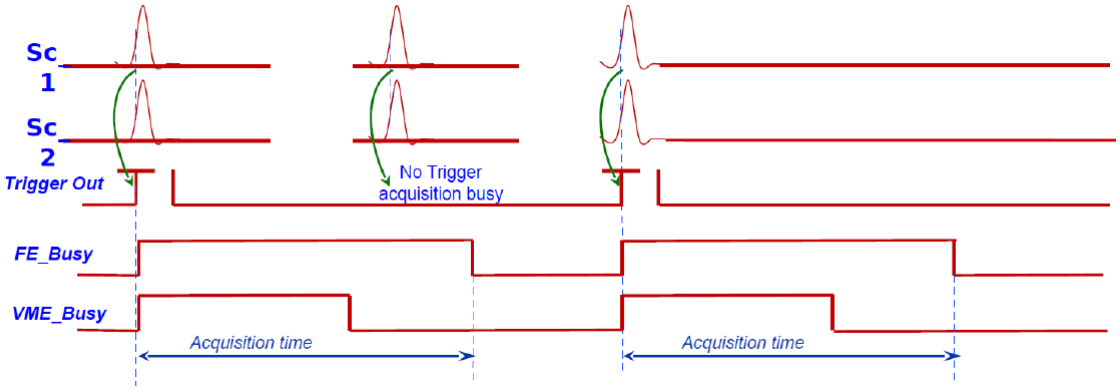


Figure 7.6: Chronograph showing the control signals (scintillator pulses, trigger and "busy" lines) before, during and after an acquisition in different cases, "busy" and no "busy" states.

they were found to contribute about 10-15% to the total. Thus, the expected noise is about 1.4 and 2.3 ADC counts for the ASIC and COTS solutions, respectively.

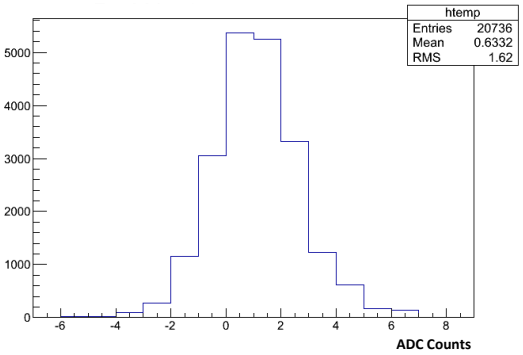


Figure 7.7: ASIC electronics noise after pedestal dynamic subtraction.

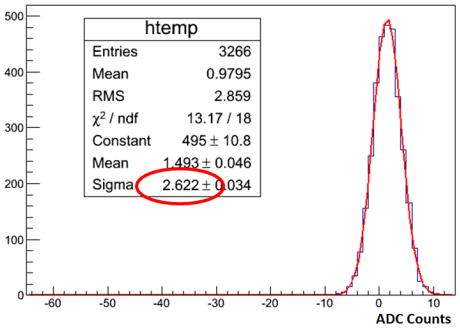


Figure 7.8: COTS electronics noise after pedestal dynamic subtraction.

7.3.2 Linearity

An important parameter of the electronic performance is the linearity of the integrator output with the amplitude of the input signal. For this test the two acquisitions, the FEB prototype and the Lecroy integrator, were compared. The ADC outputs *(Lecroy,Prototype)* for three beam energies 50 GeV, 100 GeV and 125 GeV, provide a measurement of the linearity of the fast prototype integrator.

The linearity was found to be better than 1% over the electron energy range. The results for the ASIC are shown in Figure 7.9.

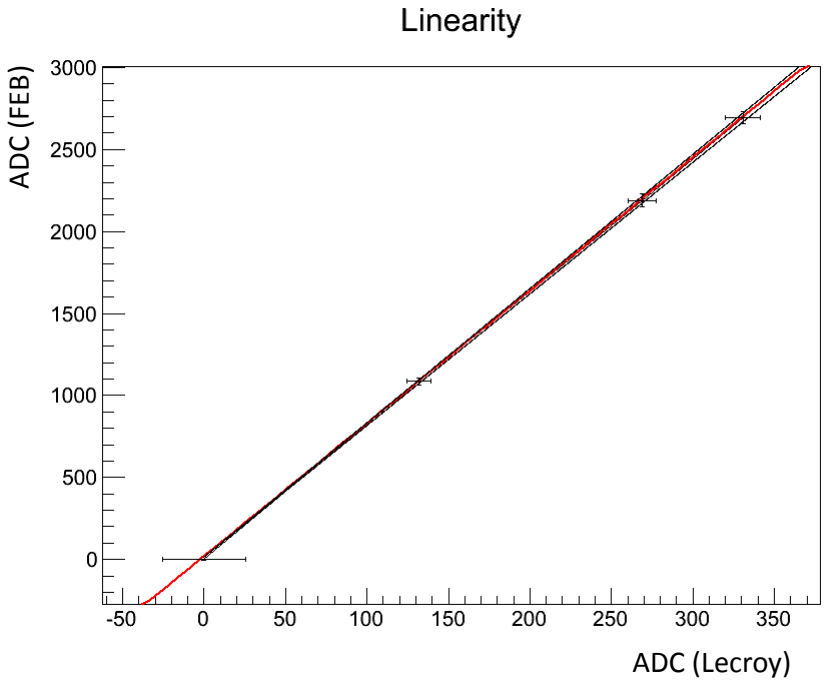


Figure 7.9: ASIC prototype linearity. The two black lines represent a ±1% deviation.

7.3.3 Integrator plateau

The integration starts at a front of a 40 MHz clock. However, the beam electrons are unsynchronised with this clock phase. This means that the pulse is not always fully integrated: for a fraction of the triggered events, only a part of the pulse is integrated. The time/phase of electrons arriving with respect to the clock is stored. Figure 7.11 illustrates the ASIC integrator output with respect to TDC counts, representing the phase of the integration start. The plateau around $TDC \sim 2000$ ns corresponds to the full integration of the signal. The output of the integrator has to be stable over 4 ns. Therefore, the fractions at ±1 and ±2 ns before and after the maximum were checked. Figure 7.11 represents of the relative variation around the maximum for the ASIC solution. The measurements for the ASIC and the COTS solutions are presented in Table 7.2. For the COTS the variation is about 1% at ±2 ns, while it is about $\sim 1.5\%$ for the ASIC. The last number shows that the plateau is not stable enough and should be improved. The solution to the problem is presented in Section 7.4.

Discrete component solution	Energy	±1 ns	±2 ns
ASIC	50 GeV	$(0.374 \pm 0.007)\%$	$(1.50 \pm 0.03)\%$
	100 GeV	$(0.374 \pm 0.008)\%$	$(1.50 \pm 0.03)\%$
	125 GeV	$(0.39 \pm 0.01)\%$	$(1.54 \pm 0.05)\%$
COTS	50 GeV	$(0.235 \pm 0.008)\%$	$(0.94 \pm 0.03)\%$

Table 7.2: Relative variation with respect to the maximum at ±1 and ±2 ns for 50 GeV, 100 GeV and 125 GeV energies for the ASIC solution and for 50 GeV for the COTS.

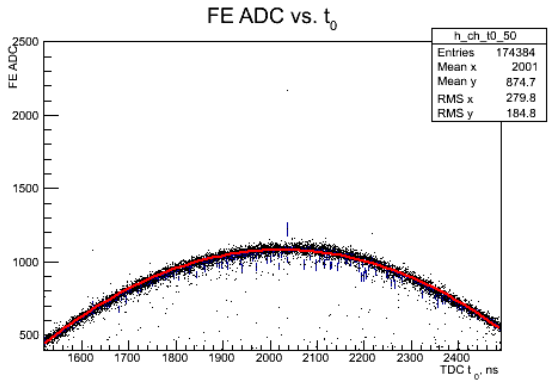


Figure 7.10: Data points around the maximum value of the integrator output plateau for the ASIC solution (ADC vs. time). The curve is fit to find the relative values at ±1 and ± 2 ns.

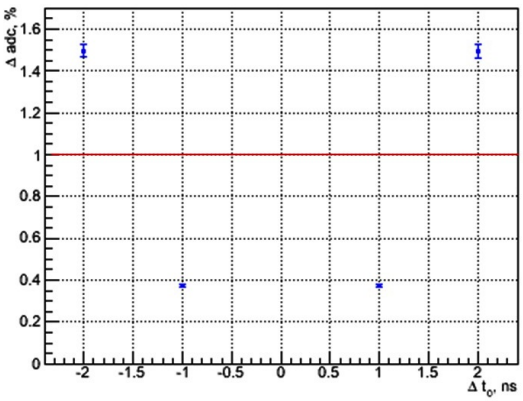


Figure 7.11: Example of the relative variation with respect to the maximum at ±1 and ±2 ns for 50 GeV energy for the ASIC solution.

7.3.4 Spill over

The signal pulse must not be wider than 25 ns. However, the signal shape displays a long tail. This is an effect of long cables, connecting the PMT and the FEB, that are also used in the LHCb cavern. This makes the pulse to be wider and the signal could leak in the 25 ns before or 25 ns after. This means that over the 25 ns of interest the signal is not fully integrated and it *spills over* to the sample before or after. It is desired that this spill over does not exceed 1 %.

To study the spill over, the shape of the signal pulses was recorded with an oscilloscope. The software integration of the signal and clipping were done offline. The integral shape is represented in Figure 7.12 in black without clipping and in red with clipping. At the point `t=0` the integral value is at maximum. The spill over affecting the samples before or after the current one is evaluated by the integral 25 ns, i.e. at the point `t=-25ns`/`t=25ns`. It is also useful to check the spill over 50 ns after, i.e. `t=50ns`.

The observed spill over is presented in Table 7.3 for the ASIC component. It is about 8% 25 ns after the maximum and 1.6% after 50 ns. These values indicate that the signal has a too wide shape to be fully integrated during 25 ns. One solution proposed to deal with this problem was to change the clipping parameters in order to improve the signal shape Figure 7.12. The results of using the variety of the clipping parameters are presented in Table 7.3. One can see that there is no significant improvement and additional corrections are needed. These are discussed in Section 7.4.

7.3.5 Reflection

Part of the signal pulse can be reflected at the connection points between the electronic board and the cables connecting it to the other elements of the apparatus. It can contribute to the next signal pulse. If the reflected fraction is underestimated the amplitude of the next signal pulse will be biased. The effect of the reflection is seen as a pulse with a delay depending on the length of the connecting cables. Thus, the data

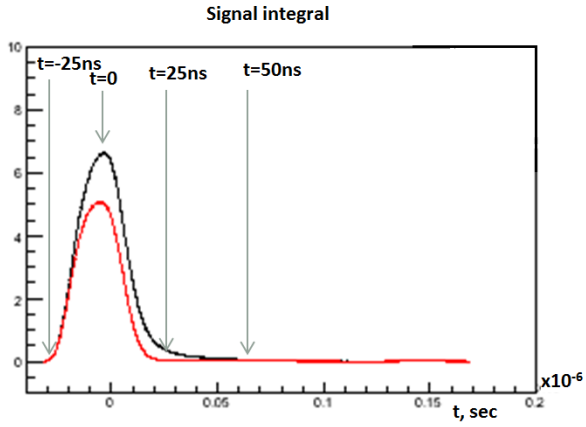


Figure 7.12: The integral shapes without (black) and with (red) correction as a function of time to check the spill over. The vertical arrows represent the samples: `t=0` - the main sample, where the integration of the signal is performed; `t=-25ns`/`t=25ns` - the sample before/after the main one; `t=50ns` - two samples after the main one.

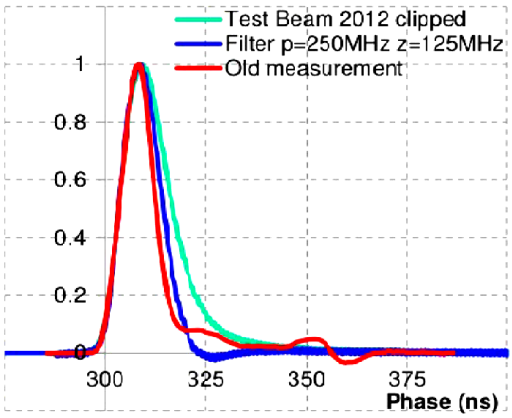


Figure 7.13: The signal shapes obtained during the test beam (cyan), measurements with the FEB currently used by LHCb (red) and with the filter pole-zero (blue) simulated after the test beam.

Time	no clip	clip with fraction 0.16	clip with fraction 0.24	clip with fraction 0.3
t=-25ns	0.14%	0.01%	0.004%	-0.11%
t=25ns	8.5%	5.6 %	3.6%	1.8%
t=50ns	1.5%	1.3%	1.2%	0.96%

Table 7.3: Spill over in percent of the signal to the samples before and after the main one. See Figure 7.12 for the illustration.

are stored during a certain period of time after the signal pulse.

The observed reflected fraction is $\sim 0.5\%$. This effect can be neglected.

7.4 Conclusions and consequences

The measurements of noise, resolution, linearity, spill over and integrator plateau show that the behaviour of the FEB lies within the specifications, except for the width of the pulse, as revealed with the integrator plateau and the spill over tests. To overcome this pulse width problem, it was proposed to implement a pole-zero filter for both the analogue ASIC and COTS the solutions which allows the tail to be reduced and the specifications to be satisfied. In Figure 7.2 the pole-zero filters are denoted as "PZ". They are installed after the amplifier and before the integrators. The pole-zero is a system of two resistance elements and a capacitor. It removes certain frequencies from the pulse in this way improving its shape. Simulations indicate that, indeed, the plateau should be significantly improved with the pole-zero filter.

The signal shape obtained from the simulations with the PZ is presented in Figure 7.13 as well as the shapes without the filter and corresponding to the currently used FEB, for comparison.

Chapter 8

Conclusion

This thesis presented my work on rare charm decays at LHCb. In particular, I focussed on $D^0 \to h^{\pm}h^{(\prime)\mp}\mu^+\mu^-$ decays. Measuring their branching ratio using LHCb's Run I data is a way to search for physics beyond the SM. It's also a way to evaluate the potential of future LHCb's datasets, including the Upgrade of LHCb that will start taking data in 2019, for the measurement of CP- and forward-backward asymmetries. The latter observables open another gate to Short Distance physics. Systematic uncertainties can be minimized by measuring these branching ratios relative to a normalization mode sharing similar features. The best mode for this is the $D^0 \to K^-\pi^+\mu^+\mu^-$ decay in the ρ/ω region. Measuring its branching ratio is the main work of this thesis. Using a 2 fb^{-1} sample of proton-proton collisions at a centre-of-mass energy of $\sqrt{s}$=8 TeV, we obtain the following preliminary result:

$$\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)) = (4.37 \pm 0.12_{stat.} \pm 0.53_{syst.}) \times 10^{-6}. \tag{8.1}$$

This measurement is also a laboratory to prepare the D^*-tagged measurement of the $D^0 \to K^-\pi^+\mu^+\mu^-$, $D^0 \to \pi^+\pi^-\mu^+\mu^-$, $D^0 \to K^+K^-\mu^+\mu^-$ and $D^0 \to K^+\pi^-\mu^+\mu^-$ modes. I prepared the selections used by these analyses. Based on this, I studied the potential of LHCb's future datasets for asymmetry measurements. If asymmetries of the order of 5 to 10% exist in the $D^0 \to K^+K^-\mu^+\mu^-$ mode, it should be with reach of the LHCb Upgrade. Asymmetries at the percent level could be discovered in the $D^0 \to K^-\pi^+\mu^+\mu^-$ mode.

I also took part in the study of the 3-body decay modes, $D^+_{(s)} \to \pi^+\mu^+\mu^-$ and $D^+_{(s)} \to \pi^-\mu^+\mu^+$, carried out by LHCb based on the data collected in 2011 (1 fb^{-1}). I evaluated the systematic uncertainties related to the selection efficiency determination. This helped obtaining upper limits about 50 times better than the ones available before the LHC era.

One chapter of the thesis is focused on tests of front-end electronic board for the Upgrade of LHCb. We studied the noise, linearity, integrator plateau, spill-over and reflection. The obtained results show a good performance of the card except for the width of the integrated pulse. This problem was solved at the time of Technical Design Report by using a pole-zero filter.

Appendix A

BDT. Input variable distributions, Correlation matrices and variable ranking for $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ mode.

```
--- TFHandler_Factory: Ranking input variables (method unspecific)...
--- IdTransformation: Ranking result (top variable is best ranked)
--- IdTransformation: ----------------------------------
--- IdTransformation: Rank: Variable: Separation
--- IdTransformation: ----------------------------------
--- IdTransformation: 1: EVCHI2: 2.757e-01
--- IdTransformation: 2: MAXDOCA: 2.219e-01
--- IdTransformation: 3: COSDIR: 2.036e-01
--- IdTransformation: 4: IPCHI2: 1.400e-01
--- IdTransformation: 5: MINHIPCHI2: 1.232e-01
--- IdTransformation: 6: MINHPT: 9.461e-02
--- IdTransformation: 7: MAXHIPCHI2: 9.392e-02
--- IdTransformation: 8: MAXHPT: 8.876e-02
--- IdTransformation: 9: MINMUIPCHI2: 8.659e-02
--- IdTransformation: 10: FDCHI2: 7.979e-02
--- IdTransformation: 11: DPT: 7.956e-02
--- IdTransformation: 12: MINMUPT: 3.581e-02
--- IdTransformation: 13: MAXMUIPCHI2: 2.472e-02
--- IdTransformation: 14: MAXMUPT: 1.792e-02
--- IdTransformation: ----------------------------------
```

Figure A.1: Ranking of BDT input variables.

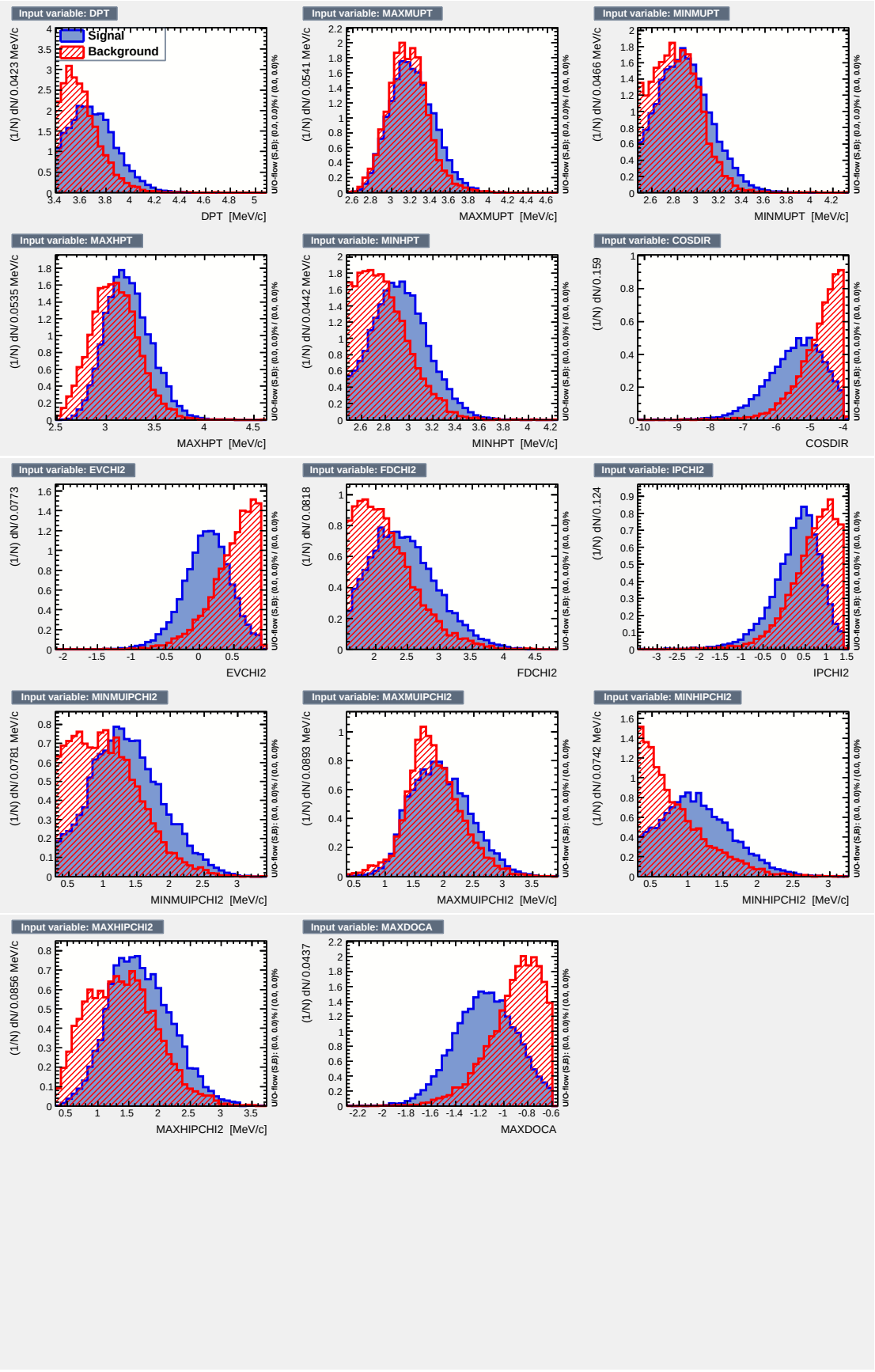


Figure A.2: Input variable distributions.

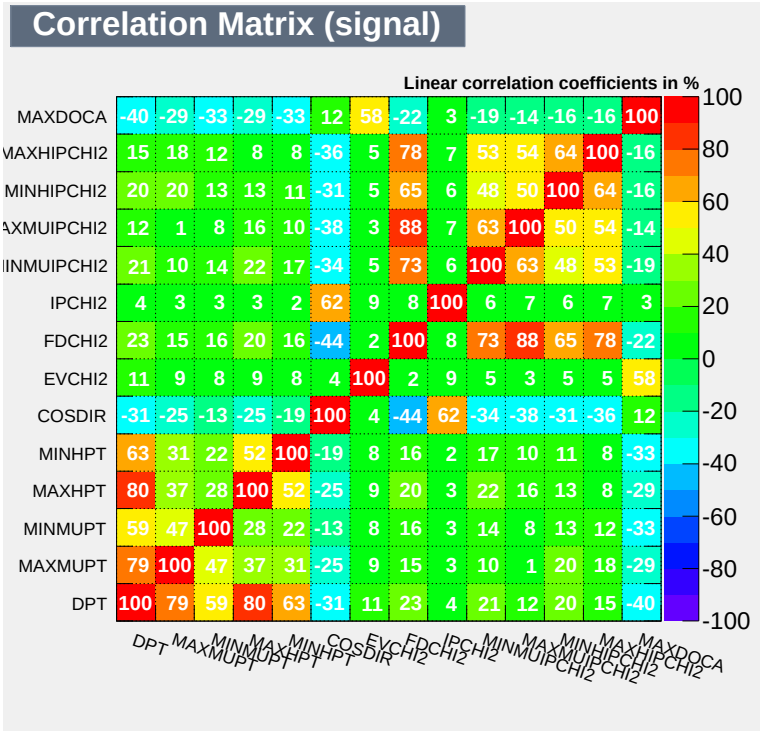


Figure A.3: BDT input variables linear correlation matrix for signal

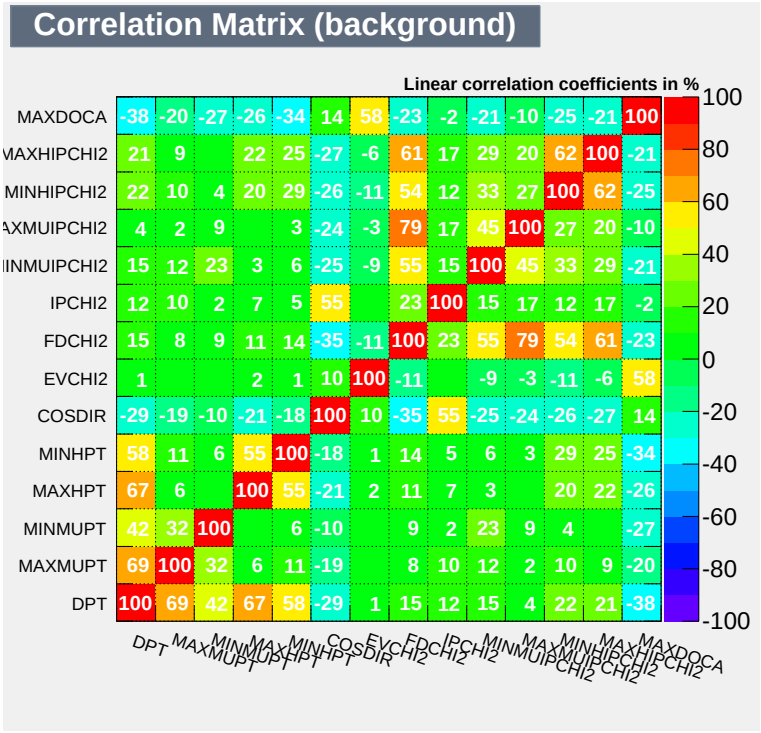


Figure A.4: BDT input variables linear correlation matrix for background

Appendix B

Additional plots related to Section 5.7

B.1 Tests with Double Crystal Ball function

An alternative to the Cruijff function is the Double Crystal Ball function:

$$f(m) = \begin{cases} exp\left(\frac{-(m-m_{D^0})^2}{2\sigma^2}\right) & \text{if } \frac{(m-m_{D^0})}{\sigma} < (\geq) 0 \text{ and } \frac{(m-m_{D^0})}{\sigma} < -|\alpha_L|(+|\alpha_R|) \\ \frac{\left(\frac{n_{L(R)}}{|\alpha_{L(R)}|}\right)^{n_{L(R)}} \times e^{-\frac{1}{2}\alpha^2_{L(R)}}}{\left(\frac{n_{L(R)}}{|\alpha_{L(R)}|} - |\alpha_{L(R)}| - (+)\left(\frac{m-m_{D^0}}{\sigma}\right)\right)} & \text{if } \frac{(m-m_{D^0})}{\sigma} < (\geq) 0 \text{ and } \frac{(m-m_{D^0})}{\sigma} > -|\alpha_L|(+|\alpha_R|) \end{cases} \tag{B.1}$$

The same samples are fitted again using this function. Table B.1 shows the values of the corresponding parameters, while Figures B.1 to B.5, show that this parametrisation also provides a satisfactory description of the peaking structures. However, it requires two more parameters to describe the tails. Toy studies showed that the $n_{L/R}$ parameters are difficult to determine. This can worsen the performance of the fit. For that reason, we chose the Cruijff function as our default. The measurement will be repeated using this PDF, and the difference with the default result will be used as a cross-check or a systematic uncertainty.

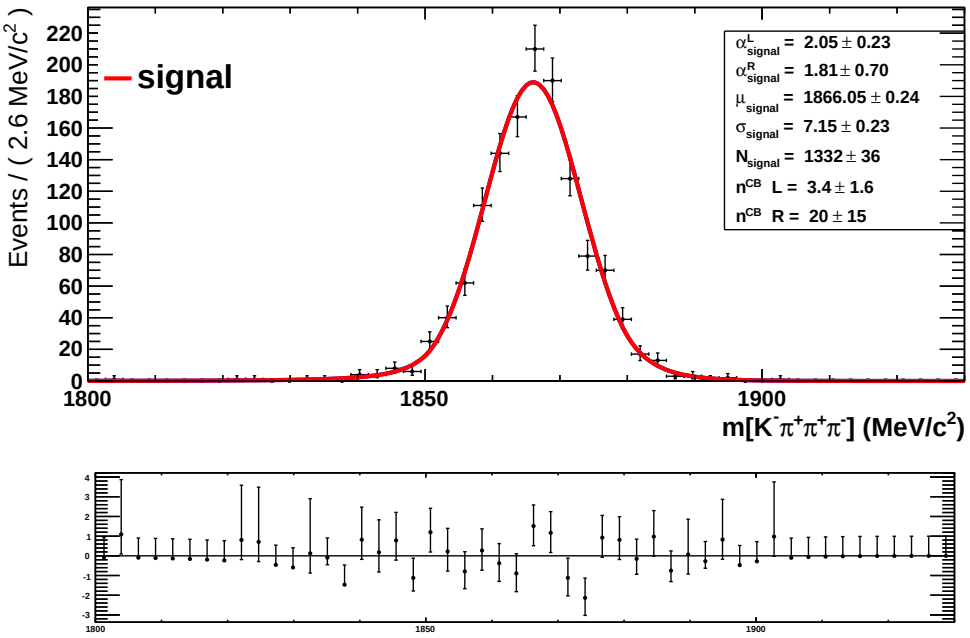


Figure B.1: Fit to the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ MC sample, selected with `L0HadronTIS` requirement, with the Double Crystal Ball function.

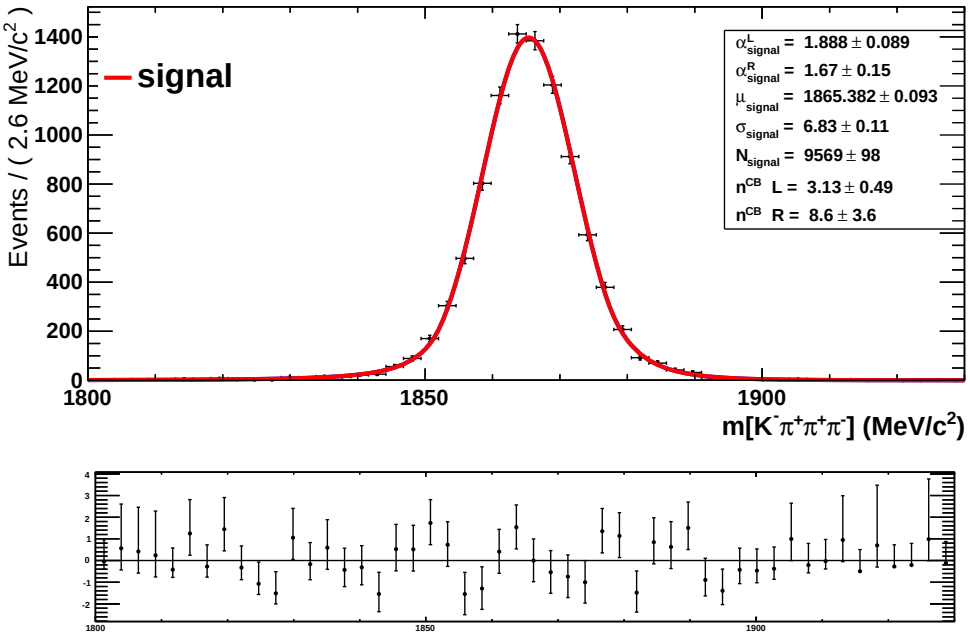


Figure B.2: Fit to the $D^0 \to K^-\pi^+\pi^+\pi^-$ MC sample with the Double Crystal Ball function.

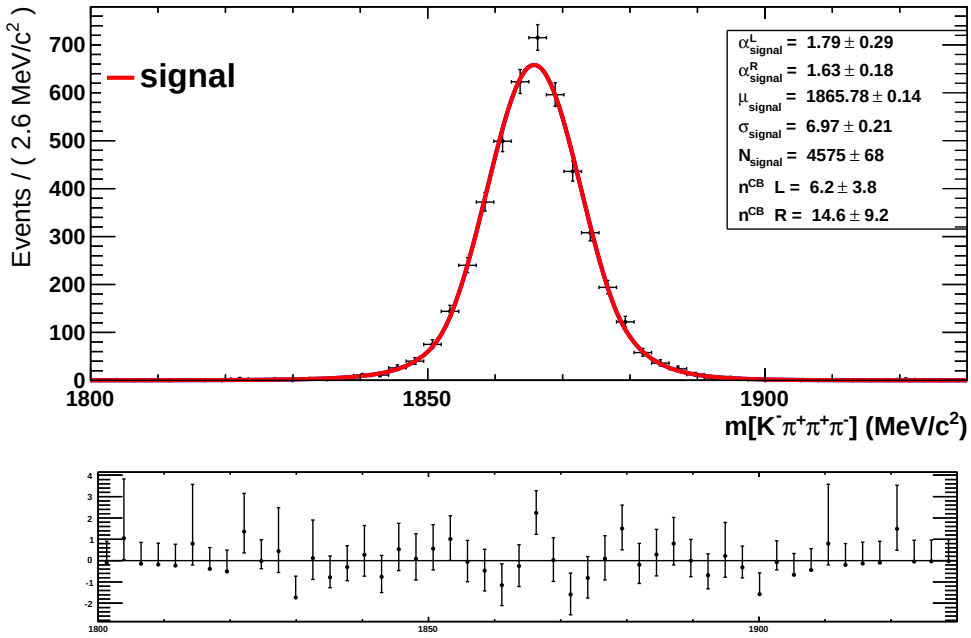


Figure B.3: Fit to the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ MC sample, selected with `L0HadronTOS||L0MuonTOS` requirement, with the Double Crystal Ball function.

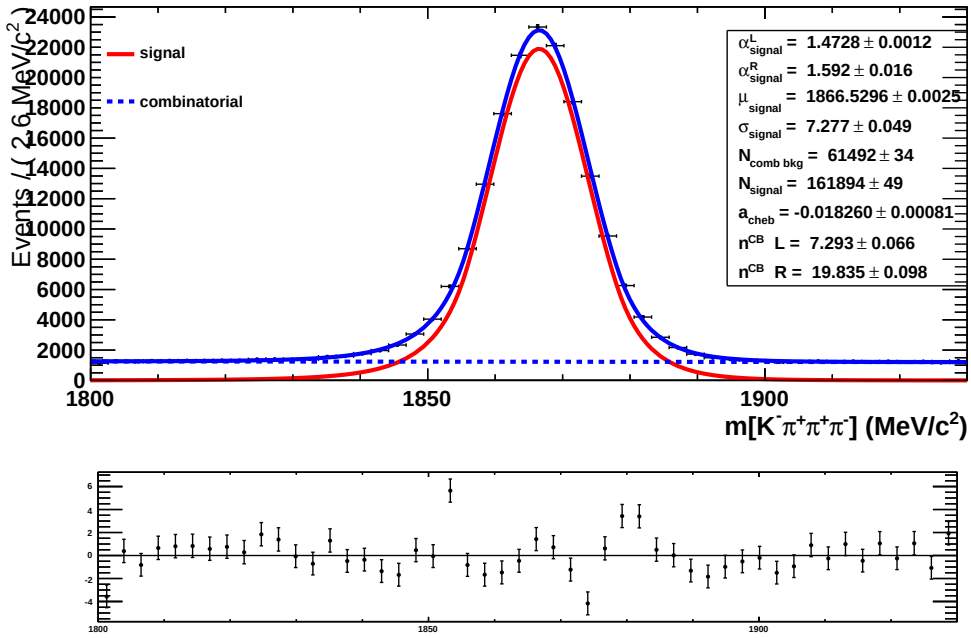


Figure B.4: Fit to the $D^0 \to K^-\pi^+\pi^+\pi^-$ data sample, using the Double Crystal Ball function to parametrize the peak. The treatment of the non-peaking component is described in Section 5.7.3.

Cruijff Parameters	Signal	Normalisation	Signal L0HadronTOS||L0MuonTOS
Monte Carlo			
m_{D^0}, MeV/c^2	1866.1 ± 0.2	1865.38 ± 0.09	1865.8 ± 0.1
σ, MeV/c^2	7.2 ± 0.2	6.8 ± 0.1	7.0 ± 0.2
α_L	2.1 ± 0.2	1.89 ± 0.09	1.8 ± 0.3
α_R	1.8 ± 0.7	1.7 ± 0.2	1.6 ± 0.2
n_L	3.4 ± 1.6	3.1 ± 0.5	6.2 ± 3.8
n_R	20.0 ± 15.1	8.6 ± 3.6	14.6 ± 9.2
Data			
m_{D^0}, MeV/c^2		1866.53 ± 0.003	1866.43 ± 0.02
σ, MeV/c^2		7.28 ± 0.05	7.40 ± 0.09
α_L		1.47 ± 0.001	1.59 ± 0.01
α_R		1.59 ± 0.02	1.54 ± 0.01
n_L		7.29 ± 0.07	11.7 ± 14.2
n_R		19.8 ± 0.1	19.9 ± 29.8

Table B.1: Values of the Double Crystal Ball function shape parameters, obtained by fits to several data and MC samples, chosen to validate the choice of this PDF as an alternative to the Cruijff function.

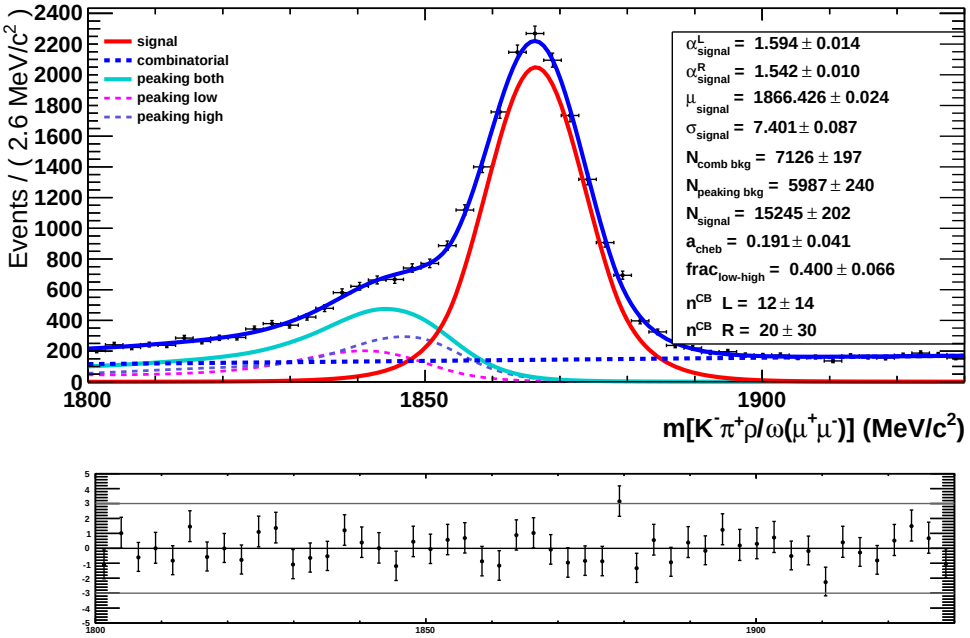


Figure B.5: Fit to the $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ data sample, selected with `L0HadronTOS||L0MuonTOS` requirement, using the Double Crystal Ball function to parametrize the signal peak. The other contributions to the total PDF are described in detail in sections 5.7.3 and 5.7.4.

B.2 Toy Study

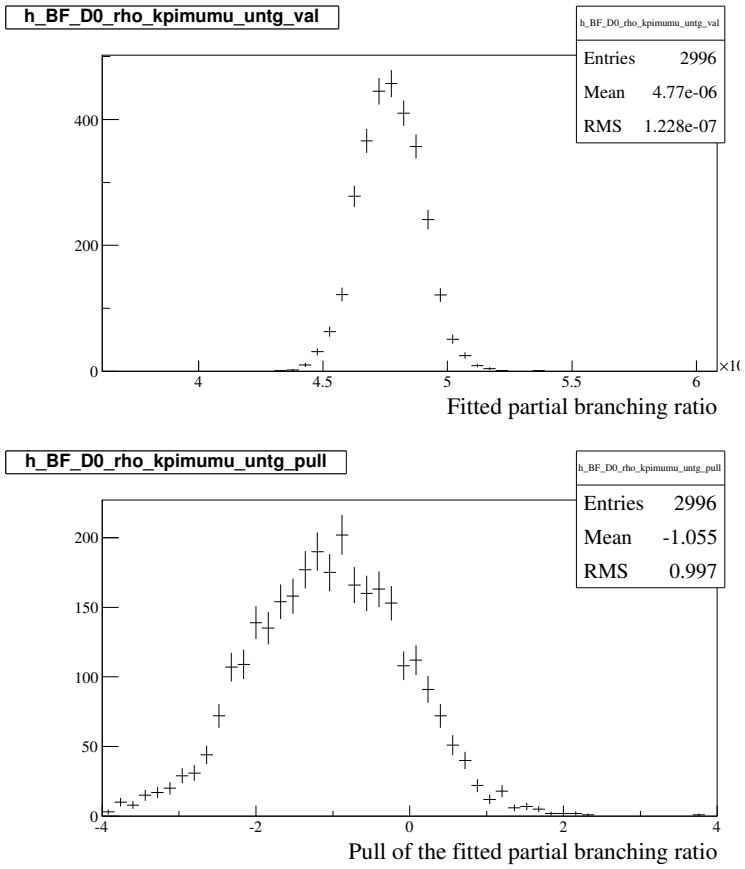


Figure B.6: Distribution of the values of $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$ obtained by fitting Toy experiments (top), and corresponding pull distribution (bottom). The toy samples are similar to the sample that will be used to perform the actual measurement. However, the component of the fit model for the $D^0 \to K^-\pi^+\pi^+\pi^-$ peaking background ignores the tails due to the pions decays in flight. The distributions show this causes a bias in the determination of $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$, equal to 3% of its generated value.

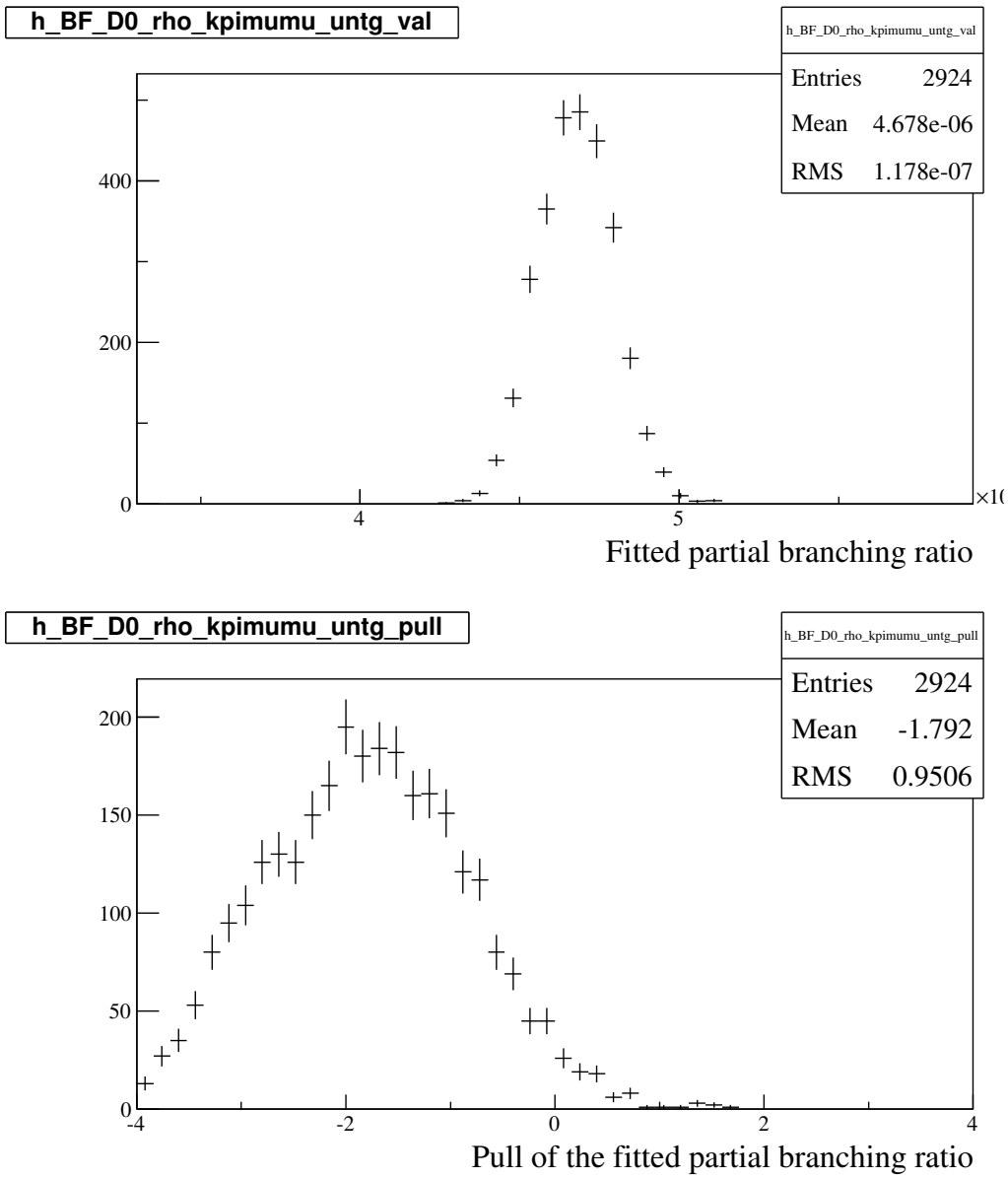


Figure B.7: Distribution of the values of $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$ obtained by fitting Toy experiments (top), and corresponding pull distribution (bottom). The toy samples are similar to the sample that will be used to perform the actual measurement. However, the component of the fit model describing the signal peak uses fixed values of α_L and α_R, equal to the values found by the fit to the signal sample obtained with the `L0HadronTOS||L0MuonTOS` requirement (Table 5.12, fourth column). The distributions show this causes a bias in the determination of $\mathcal{B}(D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-))$, equal to 4% of its generated value.

Appendix C

The comparison of $ProbNNmu$ distributions in data and MC.

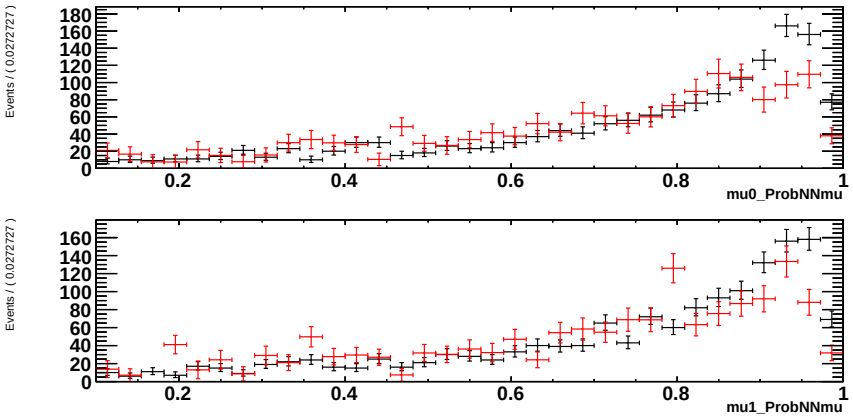


Figure C.1: $ProbNNmu$ distributions in MC (black) and sPlotted data (red) in 20% sample, selected with nominal L0 trigger cuts (`L0HadronTIS`).

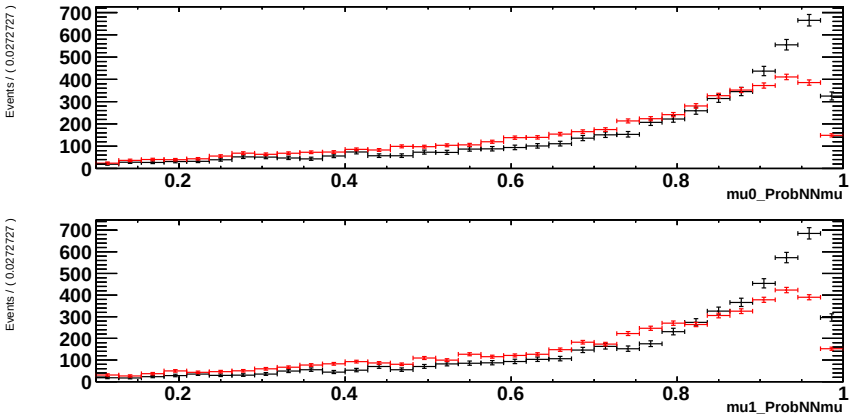


Figure C.2: $ProbNNmu$ distributions in MC (black) and sPlotted data (red) in the sample, selected with `L0HadronTOS||L0MuonTOS` L0 trigger cuts.

Appendix D

Checks of binnings for PIDCalib

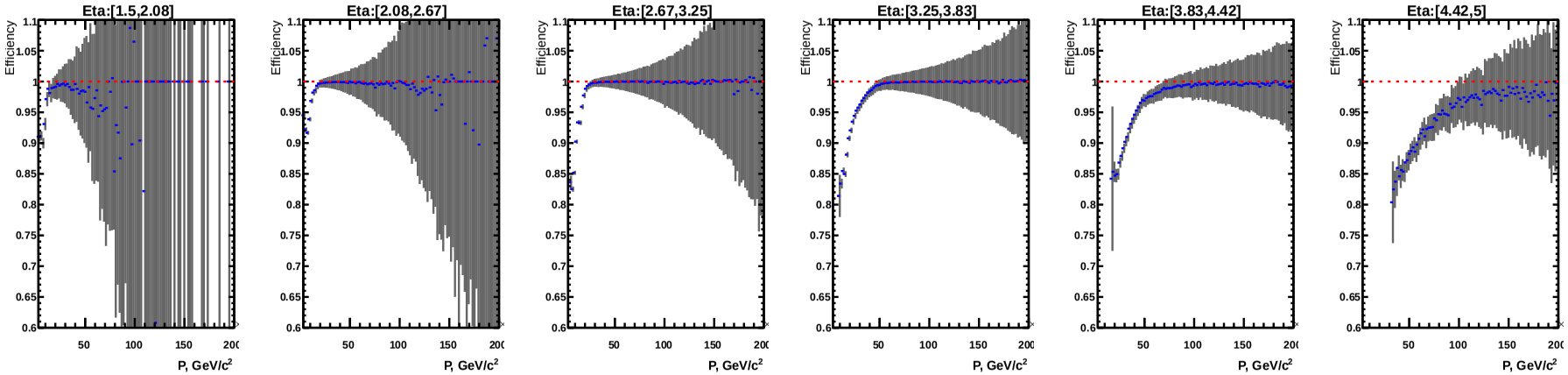


Figure D.1: *ProbNNmu* efficiency as a function of muon momentum in 6 equidistant pseudorapidity bins. Number of bins on P variable is equal 100.

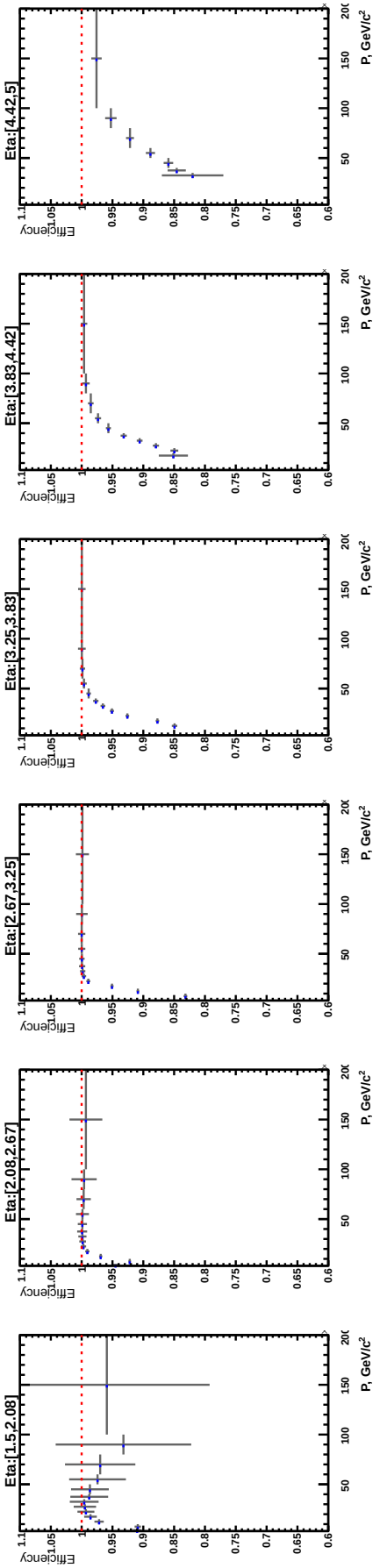
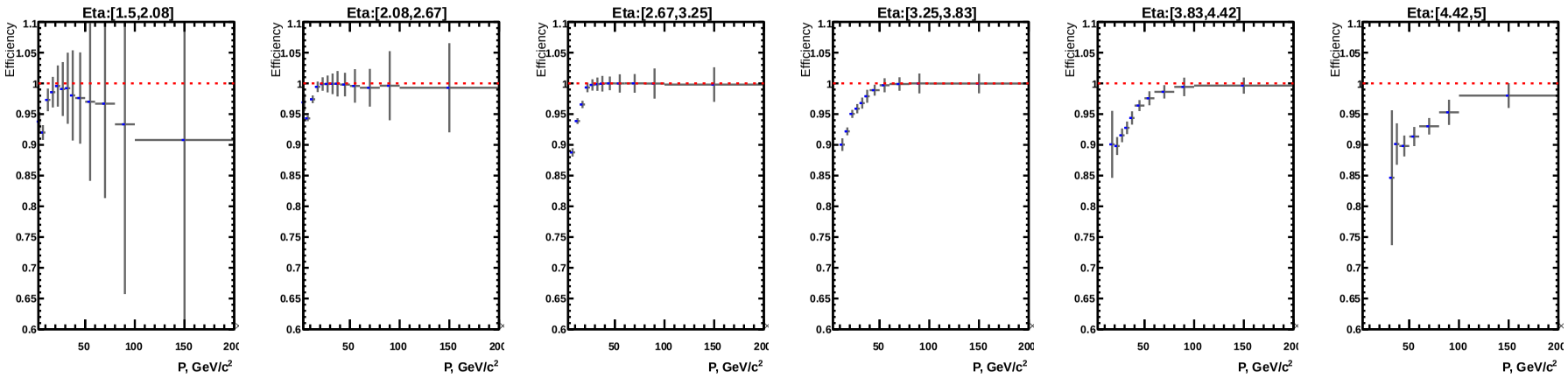
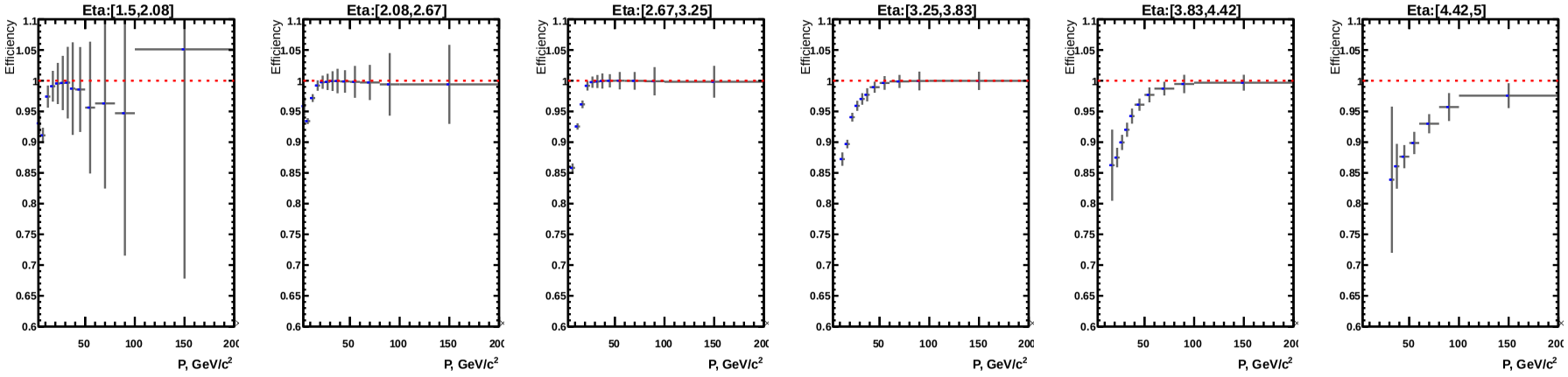


Figure D.2: *ProbNNmu* efficiency as a function of muon momentum in 6 equidistant pseudorapidity bins. Number of bins on P variable is equal 13.

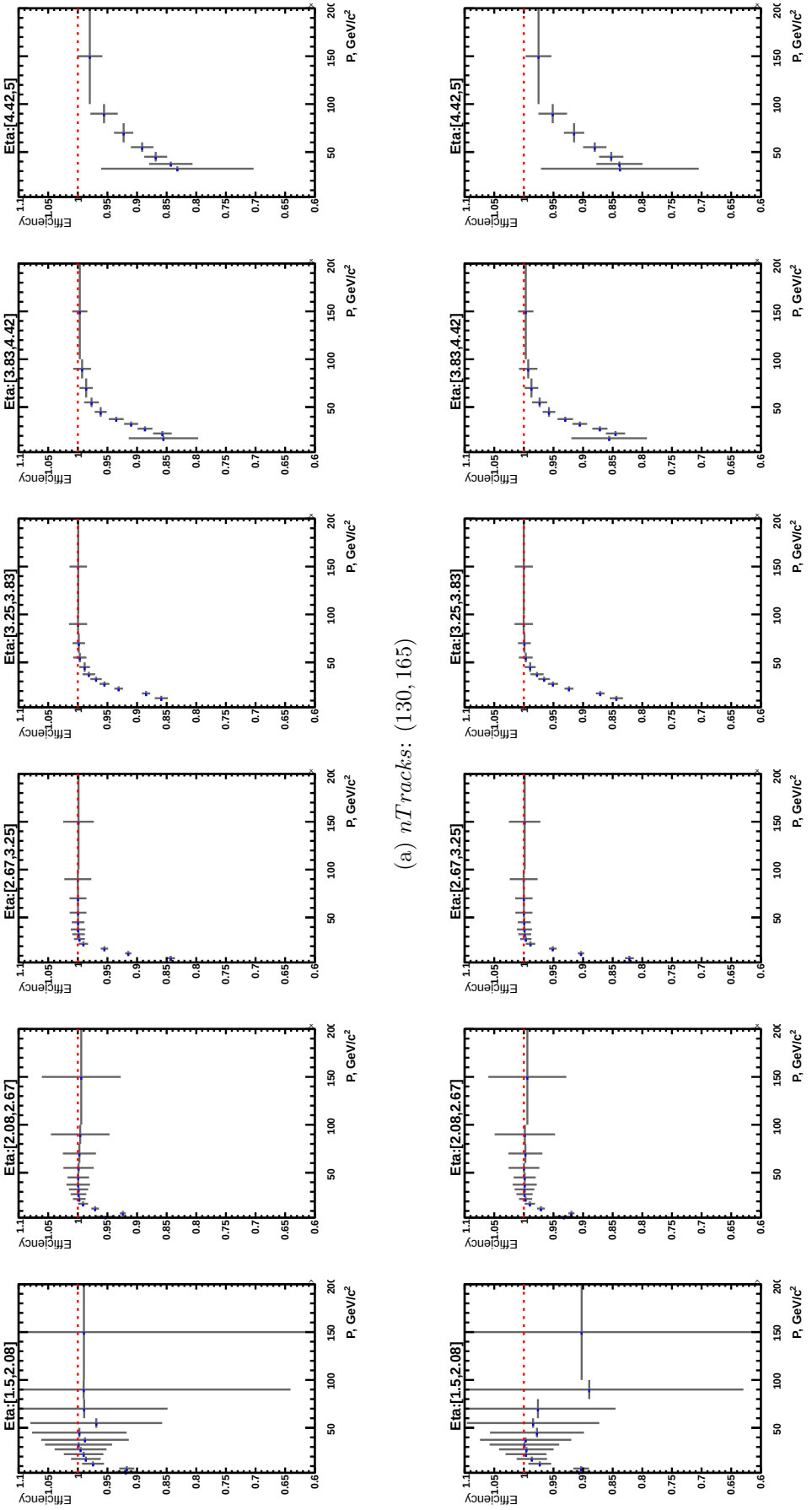


(a) $nTracks$: $(0, 95)$



(b) $nTracks$: $(95, 130)$

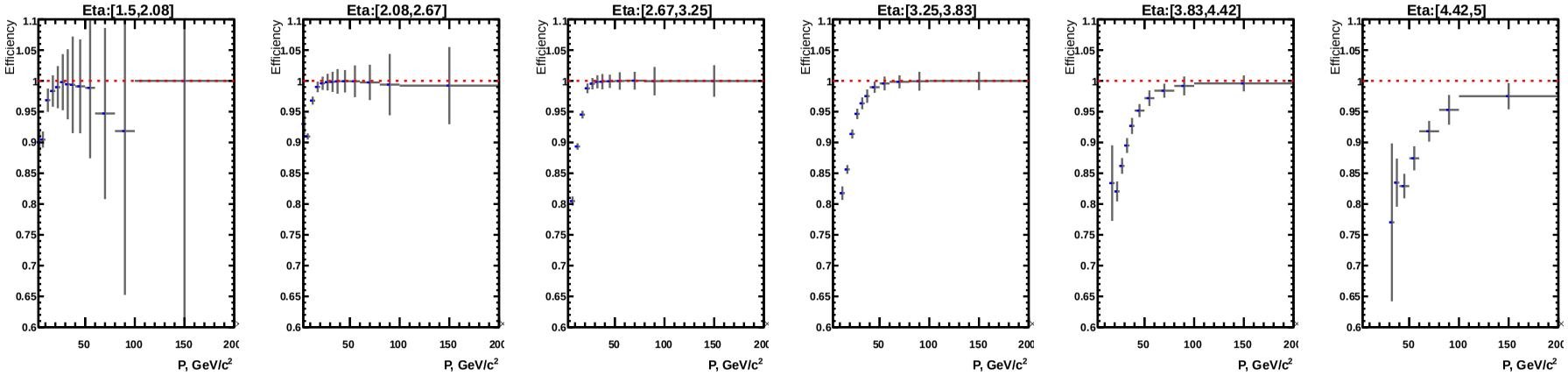
Figure D.3: $ProbNNmu$ efficiency as a function of muon momentum in 6 equidistant pseudorapidity bins and 2 bins of similar statistics on $nTracks$: $nTracks$: $(0, 95)$ and $nTracks$: $(95, 130)$. Number of bins on P variable is equal 13.



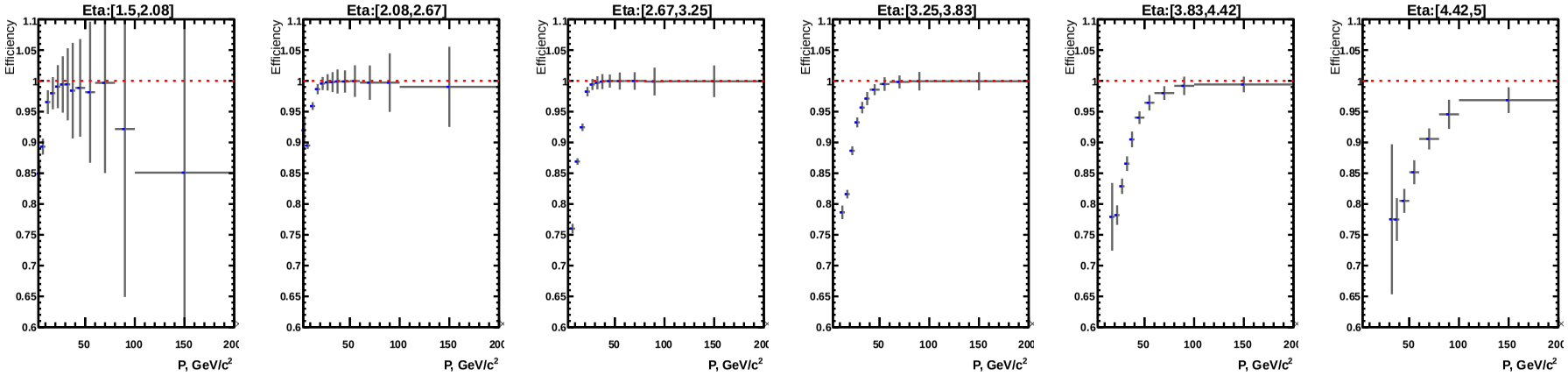
(a) $nTracks$: $(130, 165)$

(b) $nTracks$: $(165, 205)$

Figure D.4: $ProbNNmu$ efficiency as a function of muon momentum in 6 equidistant pseudorapidity bins and 2 bins of similar statistics on $nTracks$: $(130, 165)$ and $nTracks$: $(165, 205)$. Number of bins on P variable is equal 13.

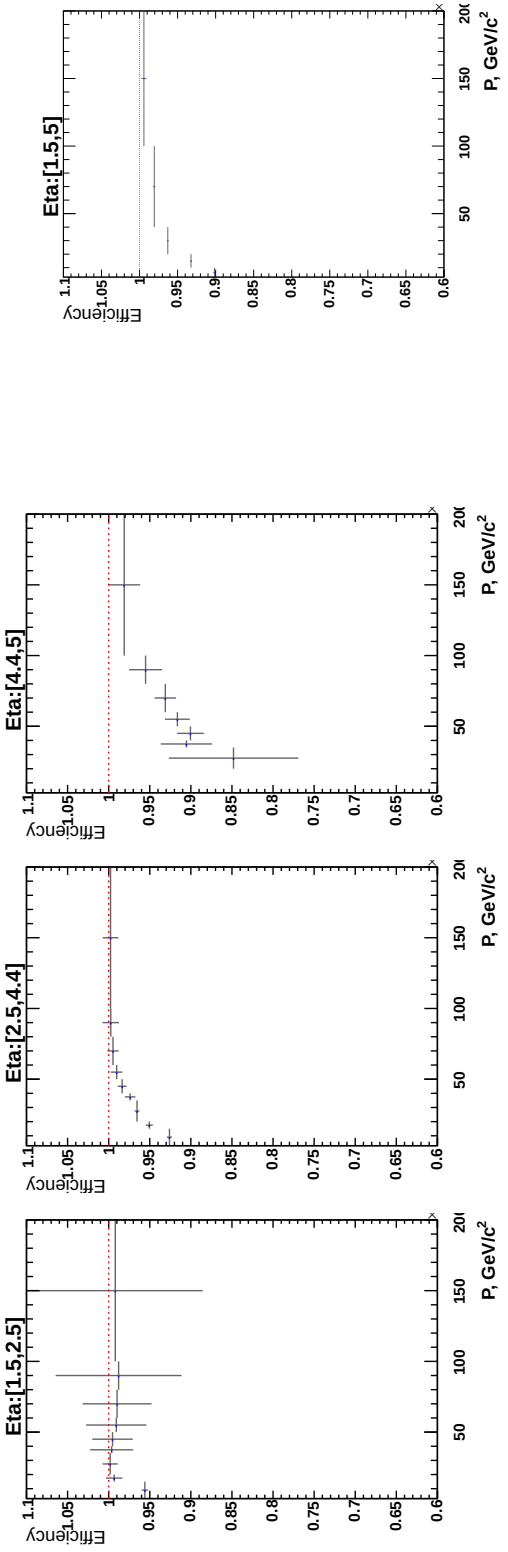


(a) $nTracks$: $(205, 265)$



(b) $nTracks$: $(265, 800)$

Figure D.5: $ProbNNmu$ efficiency as a function of muon momentum in 6 equidistant pseudorapidity bins and 2 bins of similar statistics on $nTracks$: $nTracks$: $(205, 265)$ and $nTracks$: $(265, 800)$. Number of bins on P variable is equal 13.

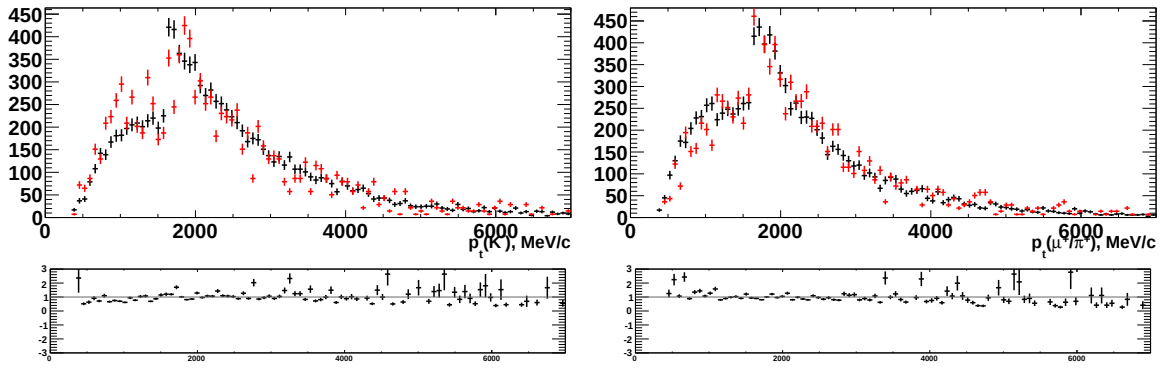


(a) Binning with 9 bins on P, 3 bins on η and 6 bins on $nTracks$. Here bin of $nTracks$: (0,95) is presented.

(b) Binning with 5 bins on P, 1 bin on η and 1 bin on $nTracks$.

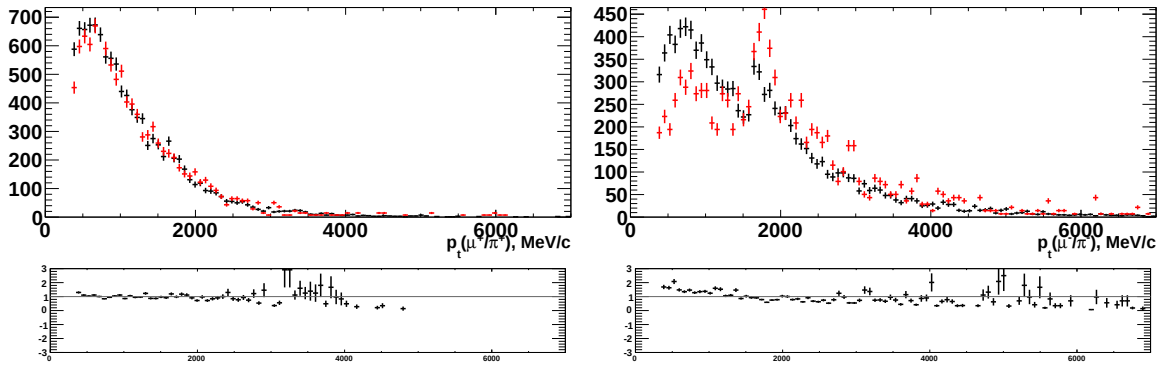
Figure D.6: Examples of $ProbNNmu$ efficiency as a function of muon momentum in coarser binning than on Figures D.3, D.4, D.5

D.1 $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ vs $D^0 \to K^-\pi^+\pi^+\pi^-$ in MC



(a) $p_t(K^-)$

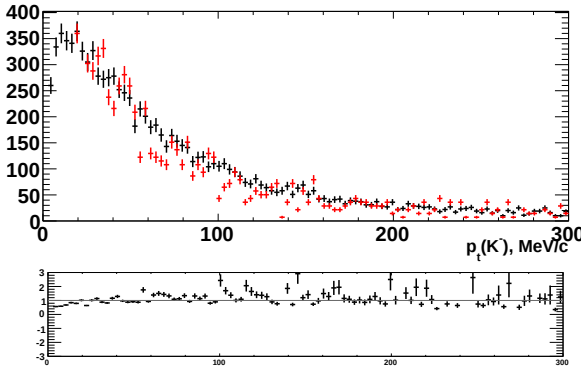
(b) max $p_t(\mu^+/\pi^+)$



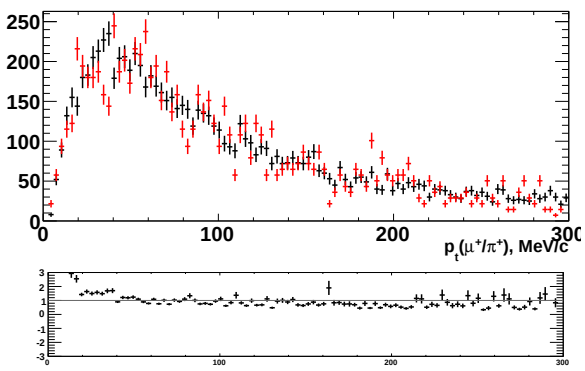
(c) min $p_t(\mu^+/\pi^+)$

(d) $p_t(\mu^-/\pi^-)$

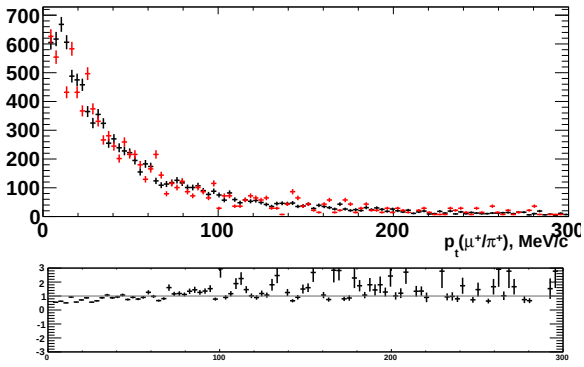
Figure D.7: Transverse momentum distributions of daughters tracks in MC for signal (red) and normalization (black).



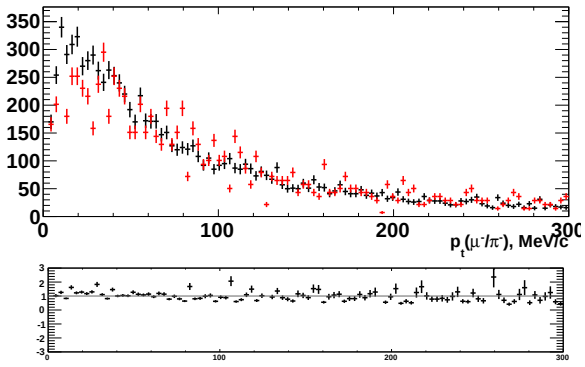
(a) $IP\chi^2(K^-)$



(b) max $IP\chi^2(\mu^+/\pi^+)$

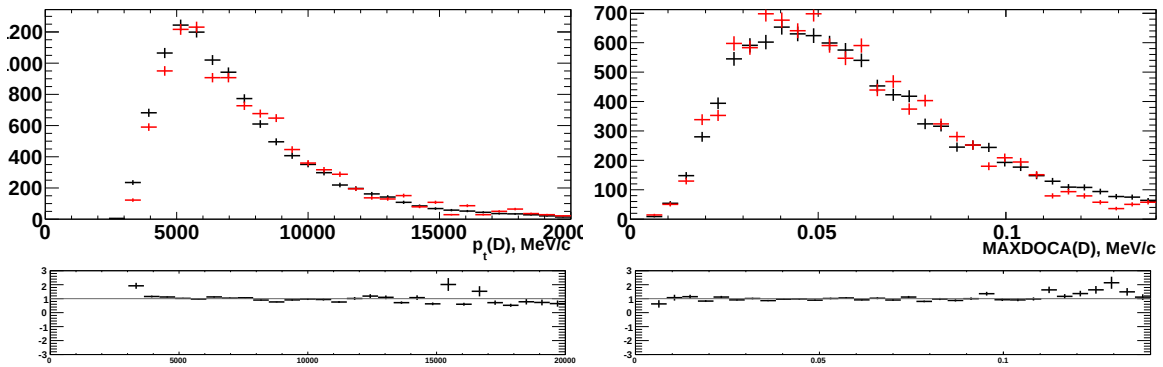


(c) min $IP\chi^2(\mu^+/\pi^+)$

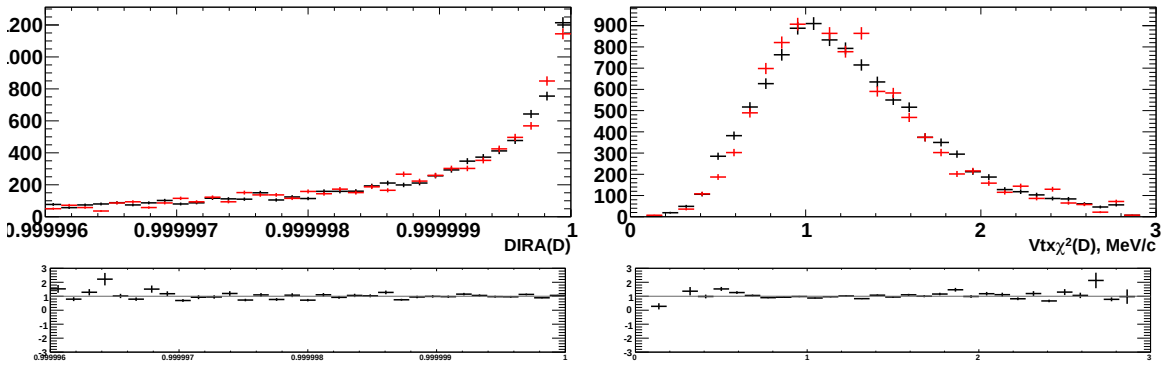


(d) $IP\chi^2(\mu^-/\pi^-)$

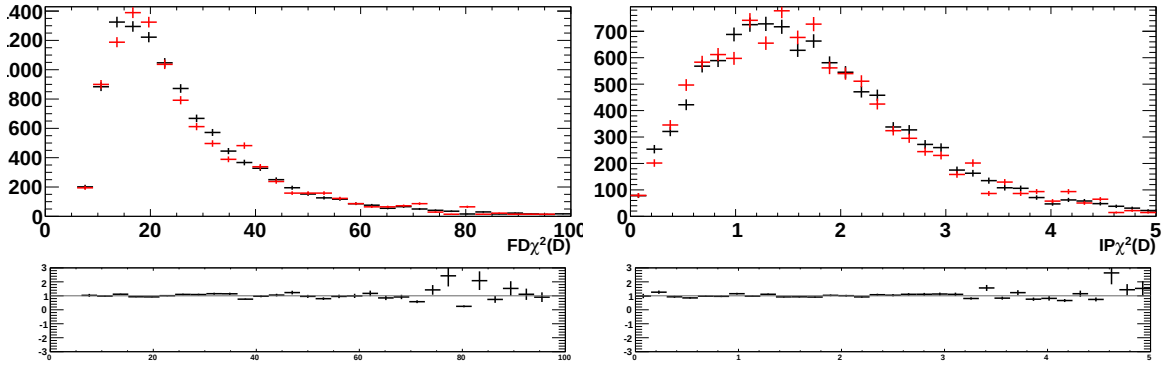
Figure D.8: IPχ^2 distributions of daughters tracks in MC for signal (red) and normalization (black).



(a) $p_t(D^0)$ (b) $DOCA(D^0)$



(c) $DIRA(D^0)$ (d) $Vtx\chi^2(D^0)$



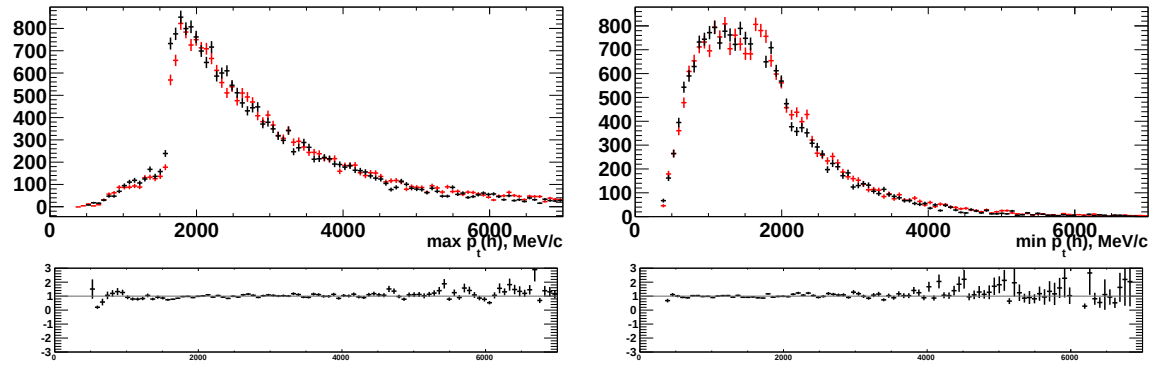
(e) $FD\chi^2(D^0)$ (f) $IP\chi^2$

Figure D.9: Momentum, DOCA, DIRA, Vertex χ^2, Flight distance χ^2 and IPχ^2 of D^0 in MC for signal (red) and normalization (black).

Appendix E

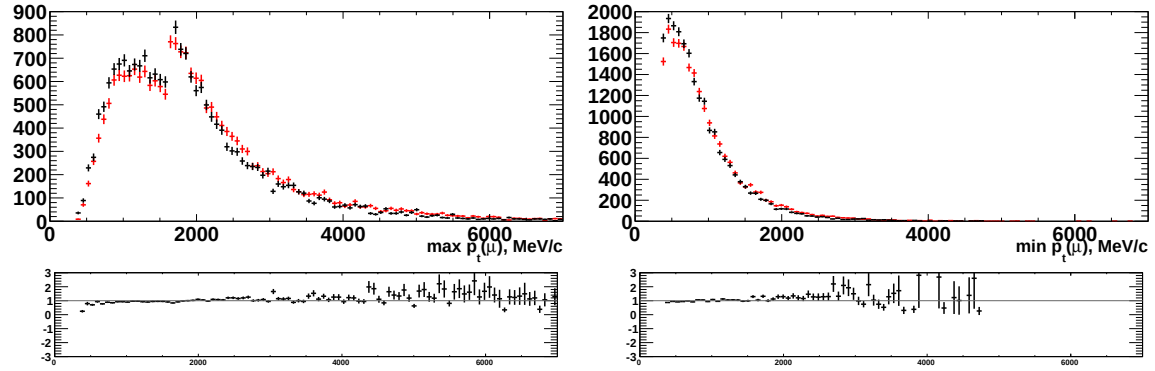
Distributions of variables used for BDT training in data and MC. `L0HadronTOS||L0MuonTOS` trigger selection.

E.1 $D^0 \to K^-\pi^+\pi^+\pi^-$



(a) max $p_t(h)$

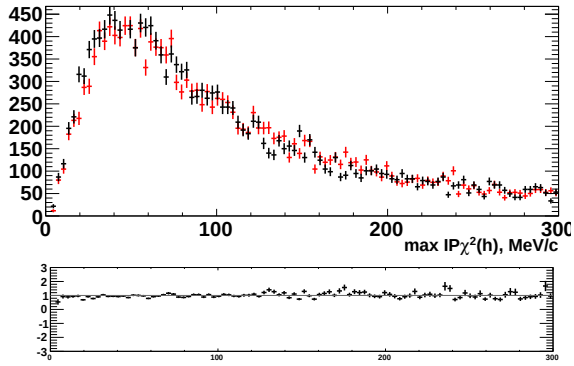
(b) min $p_t(h)$



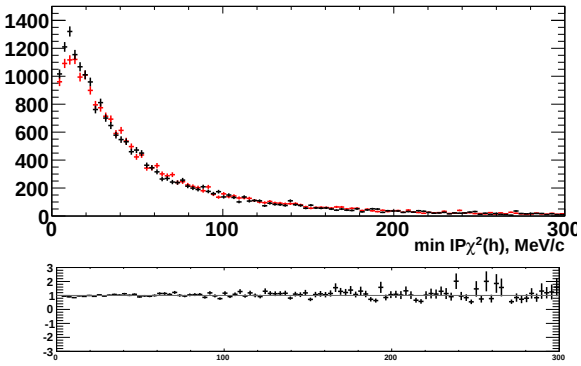
(c) max $p_t(\mu)$

(d) max $p_t(\mu)$

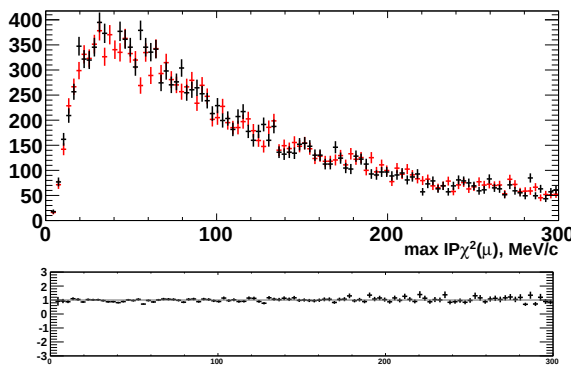
Figure E.1: Momentum distributions of daughters tracks in data (red) and MC (black) in $D^0 \to K^-\pi^+\pi^+\pi^-$ with nominal selection (`L0HadronTOS||L0MuonTOS`).



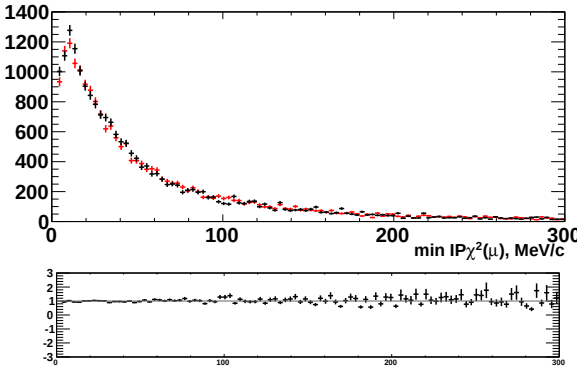
(a) max $IP\chi^2(h)$



(b) min $IP\chi^2(h)$

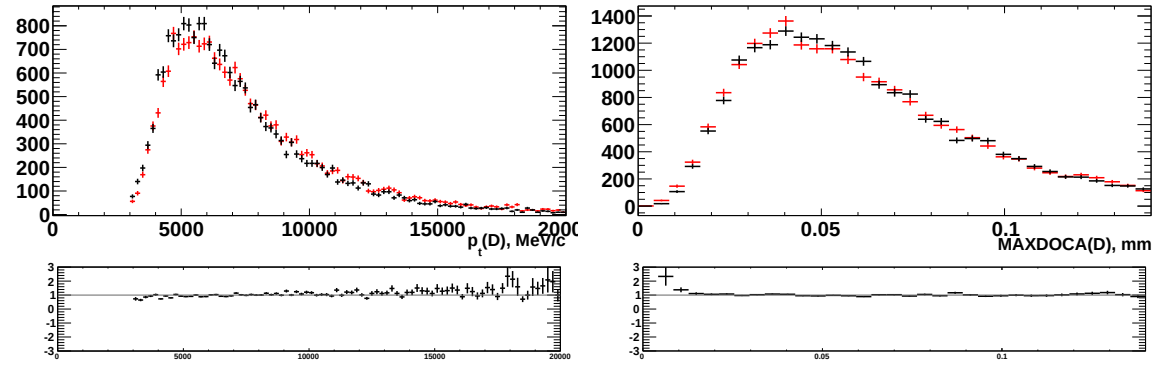


(c) max $IP\chi^2(\mu)$

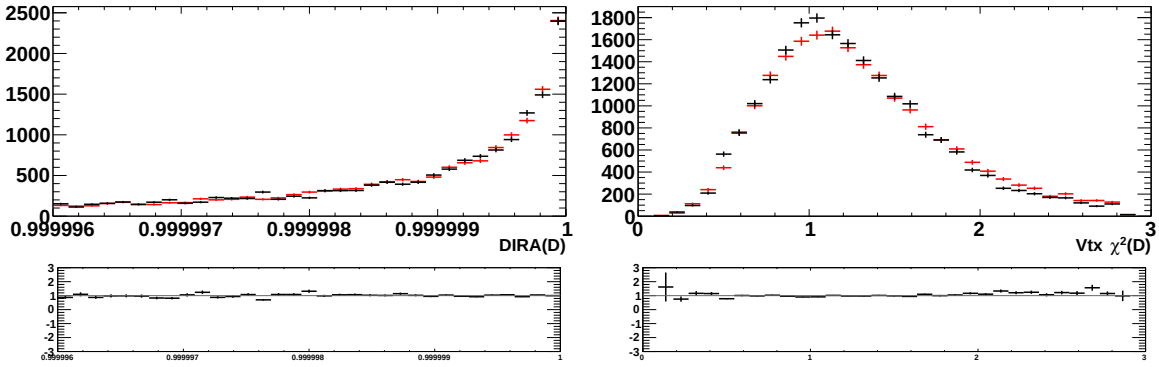


(d) max $IP\chi^2(\mu)$

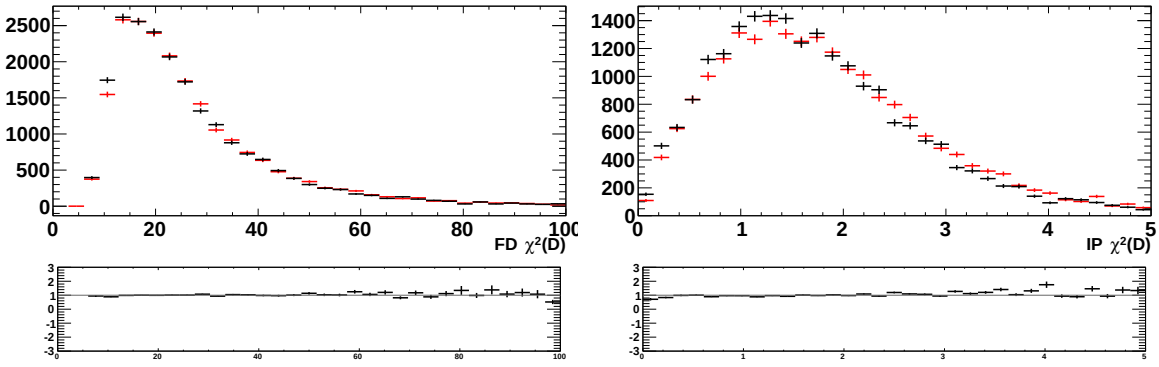
Figure E.2: Impact parameter χ^2 distributions of daughters tracks in data (red) and MC (black) in $D^0 \to K^-\pi^+\pi^+\pi^-$ with nominal selection (`L0HadronTOS||L0MuonTOS`).



(a) $p_t(D^0)$ (b) $DOCA(D^0)$



(c) $DIRA(D^0)$ (d) $Vtx\chi^2(D^0)$



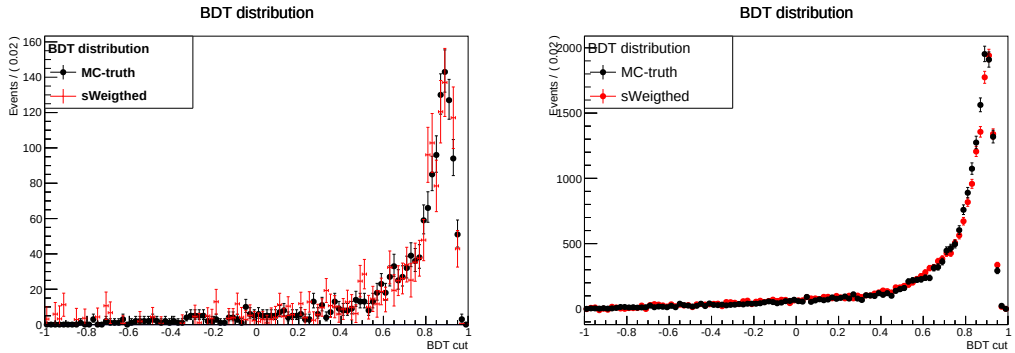
(e) $FD\chi^2(D^0)$ (f) $IP\chi^2$

Figure E.3: Momentum, Distance of Closest Approach (DOCA), Direction Angle (DIRA), Vertex χ^2, Flight distance χ^2 and impact parameter IPχ^2 of D^0 in data (red) and MC (black) in $D^0 \to K^-\pi^+\pi^+\pi^-$ with nominal selection (`L0HadronTOS||L0MuonTOS`).

E.2 Tests of BDT distributions for compatibility of data/MC on 20% $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ sample.

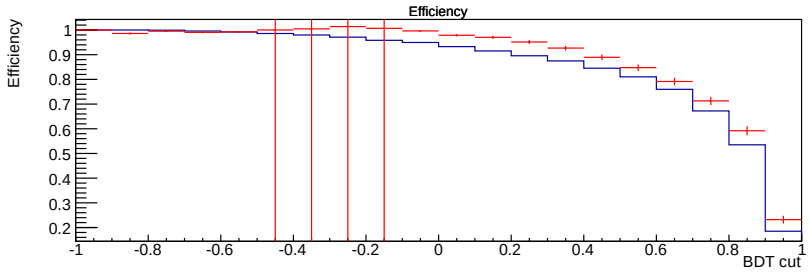
To study possible systematic discrepancies between data and MC we use *sPlot* method to subtract background in 20% signal $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ sample and 10% normalization sample. Individual BDT distributions in data and in MC are presented on Figure E.4, efficiencies and relative data-MC differences on Figure E.5 and E.6. The plot for efficiency ratio in data and MC and variation of their difference are presented on Figure E.7. One can see the large discrepancy up to 20% on Figure E.7(b) for the difference of efficiency ratios in MC and data. This effect can be explained by lack of statistics for making *sPlot* in 20% signal sample. In Figure E.4(a) there are regions, where sWeights are negative (empty regions on the plot). In Figure E.5(a) and Figure E.6 there are four badly calculated bins between $BDT = [-0.5, -0.1]$, where we have empty region on Figure E.4(a). It is clear that behaviour of the data trend on Figure E.5(a) changes after these BDT values. This signals that some statistical fluctuation or just errors during *sPlotting* due to lack of statistics in 20% signal sample takes place. On Figure E.8 there is one more prove of the statistical fluctuation. We present here several fits of 20% data sample for different cuts on BDT. One can notice that signal yields increase from $BDT > -0.6$ to $BDT > -0.2$. However it is impossible to cut harder on BDT and get more signal events. Such effect can be explained only by a statistical fluctuation.

In case of MC the efficiency ratio stays close to 1. It's a bit different in data but this difference should be similar in case of the signal and normalization. Looking at kinematical distributions in Section 5.5.5 one can see that data and MC distributions are consistent for $D^0 \to K^-\pi^+\pi^+\pi^-$ and $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$, up to a statistical effect in case of signal. This means that efficiency ratios in data and MC should also be similar.

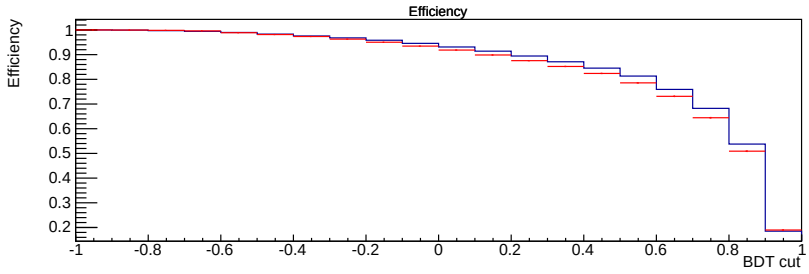


(a) 20% $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ sample. (b) $D^0 \to K^-\pi^+\pi^+\pi^-$ sample.

Figure E.4: BDT distributions in data (red) and MC (black) for the signal and normalization samples selected with nominal cuts.

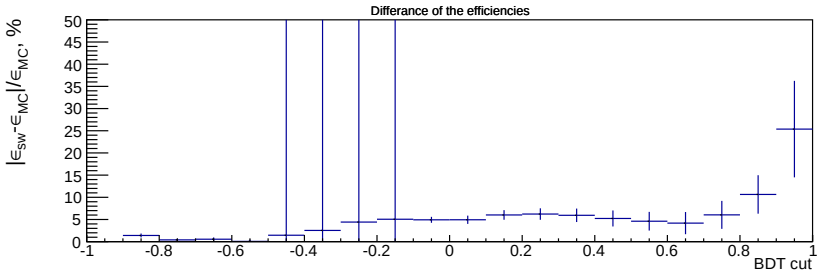


(a) 20% $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ sample.

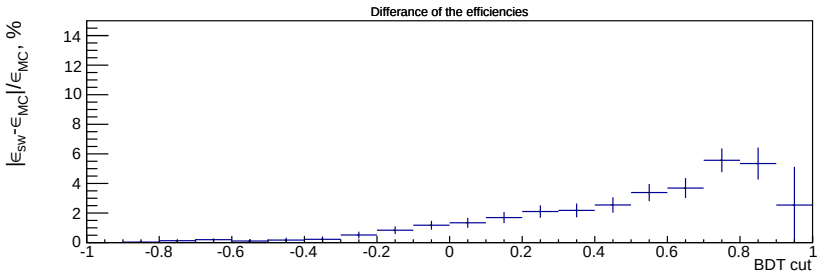


(b) $D^0 \to K^-\pi^+\pi^+\pi^-$ sample.

Figure E.5: Efficiency as a function of the cut on the discriminating variable produced by the BDT, in data (red) and MC (blue) for the signal and normalization samples selected with nominal cuts.

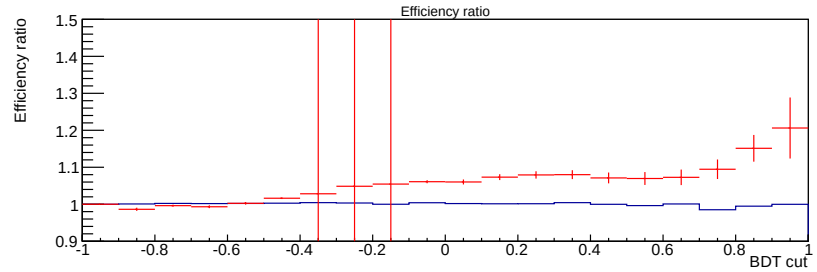


(a) 20% $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ sample.

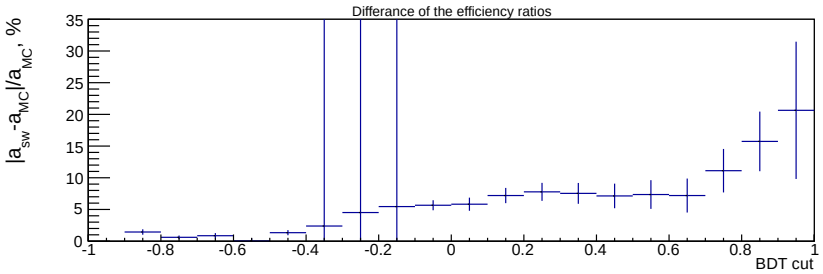


(b) $D^0 \to K^-\pi^+\pi^+\pi^-$ sample.

Figure E.6: Relative difference in efficiencies between data and MC for the signal and normalization samples selected with nominal cuts.

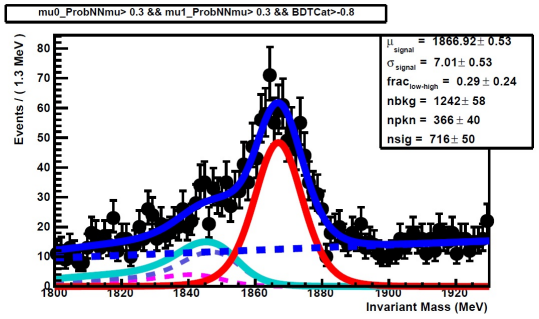


(a) Efficiency ratio in data (red) and MC (blue).

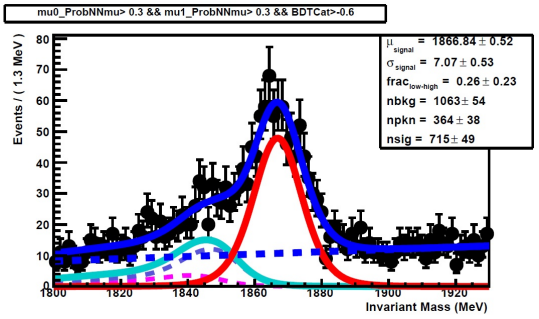


(b) Difference of efficiency ratio in MC and data

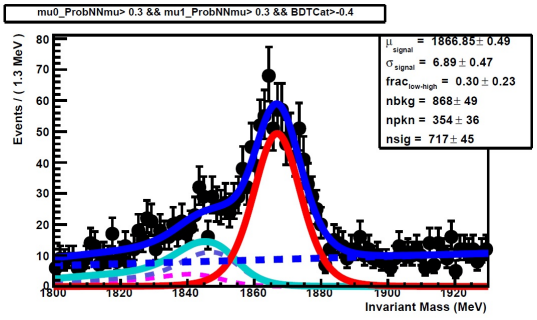
Figure E.7: Efficiency ratio in data and MC (top) and relative difference between these two ratios (bottom). The cut used in the analysis is $BDT > 0.2$



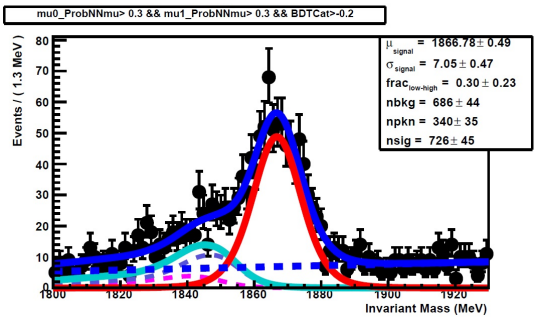
(a) $BDT > -0.8$, $N_{sig} = 716 \pm 50$



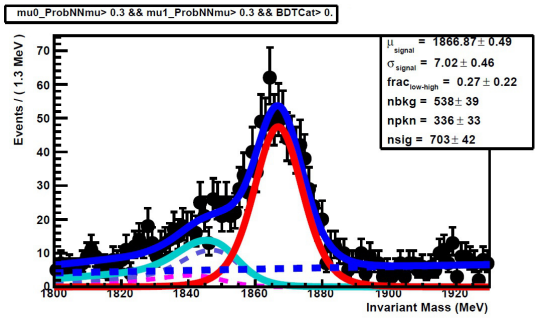
(b) $BDT > -0.6$, $N_{sig} = 715 \pm 49$



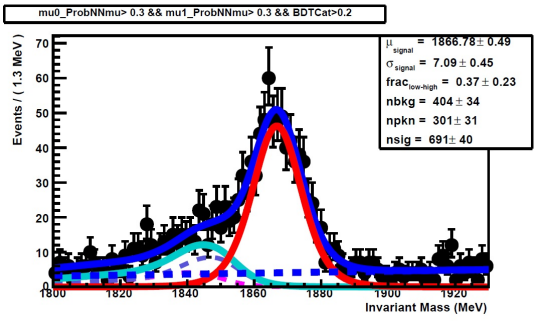
(c) $BDT > -0.4$, $N_{sig} = 717 \pm 45$



(d) $BDT > -0.2$, $N_{sig} = 726 \pm 45$



(e) $BDT > 0$, $N_{sig} = 703 \pm 42$

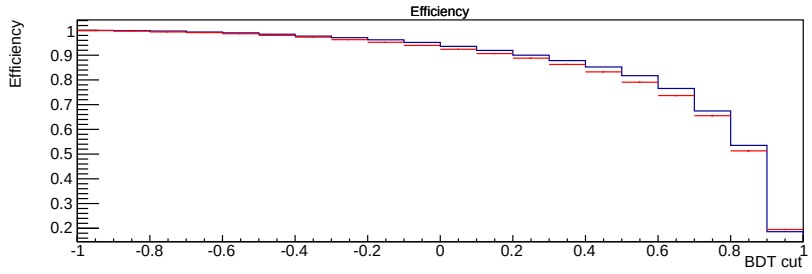


(f) $BDT > 0.2$, $N_{sig} = 691 \pm 40$

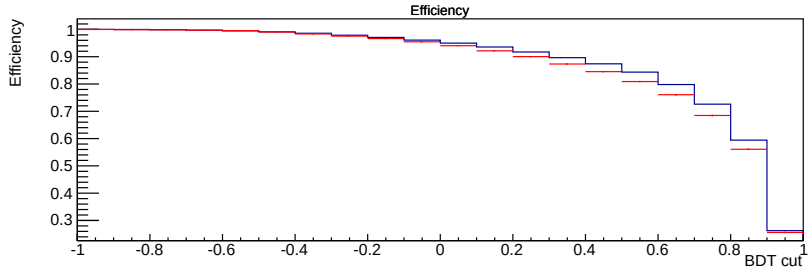
Figure E.8: Fits of 20% of $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ sample at several levels of BDT cut to show the statistical variation of the signal yield.

E.3 Tests of efficiency of BDT selection on `L0HadronTOS||L0MuonTOS` sample.

On Figures E.9 and E.10 distributions for efficiencies and differences between data and MC are given. The difference is higher at high BDT value but stays less than 6%. It means that there is some difference between data and MC, but it doesn't cause problems for the analysis as far as data/MC discrepancies are the same for both $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ and $D^0 \to K^-\pi^+\pi^+\pi^-$.

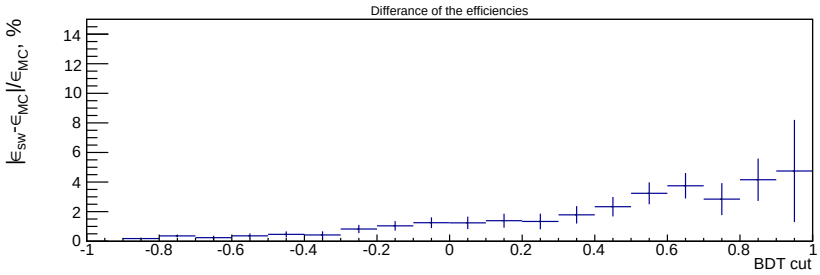


(a) $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ sample.

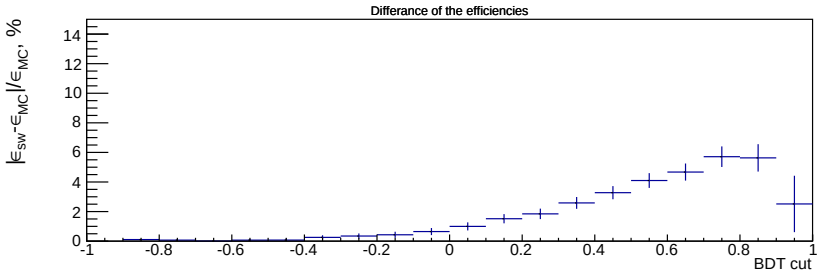


(b) $D^0 \to K^-\pi^+\pi^+\pi^-$ sample.

Figure E.9: Efficiency as a function of the cut on the discriminating variable produced by the BDT, in data (red) and MC (blue) for the signal and normalization samples selected with `L0HadronTOS||L0MuonTOS`.



(a) $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ sample.

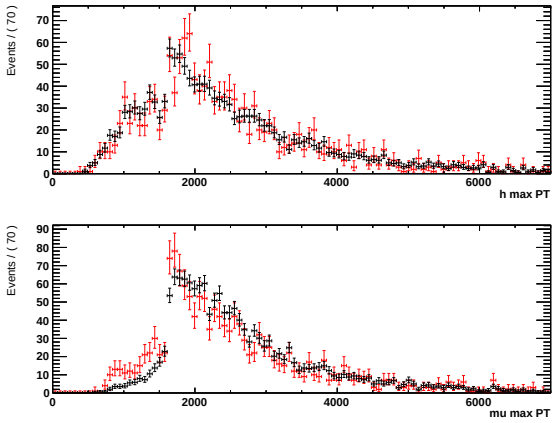


(b) $D^0 \to K^-\pi^+\pi^+\pi^-$ sample.

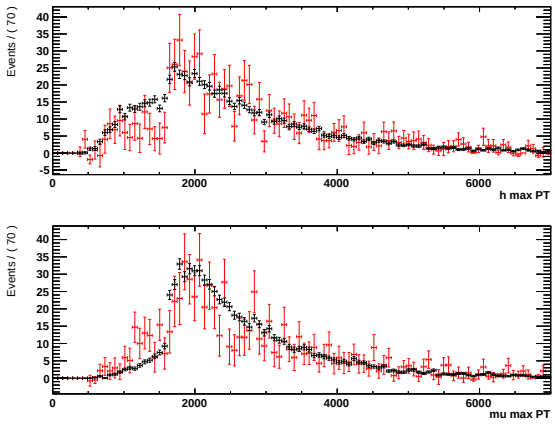
Figure E.10: Differences in efficiencies between data and MC for the signal (a) and normalization (b) samples selected with `L0HadronTOS||L0MuonTOS`.

Appendix F

Kinematical distributions in data and MC for $D^0 \to K^-\pi^+\rho/\omega(\to \mu^+\mu^-)$ selected with `L0HadronTOS||L0MuonTOS` and `L0HadronTIS` (nominal).

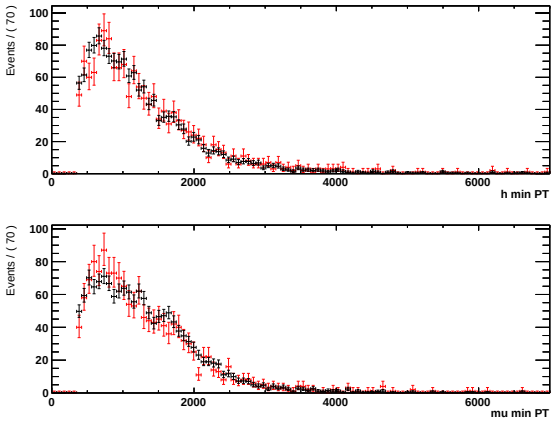


(a) MC

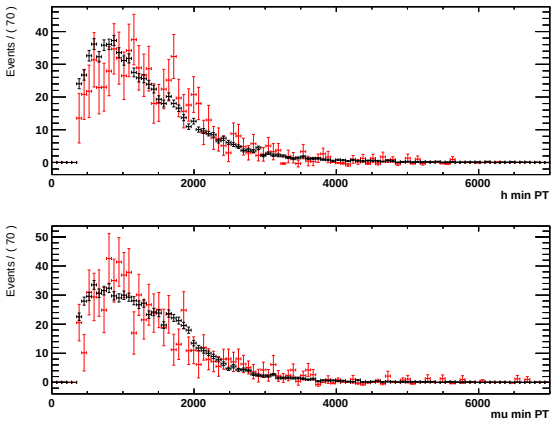


(b) Data

Figure F.1: Transverse momentum distributions in data and MC, selected with `L0HadronTOS||L0MuonTOS` (black) and nominal `L0HadronTIS` (red).

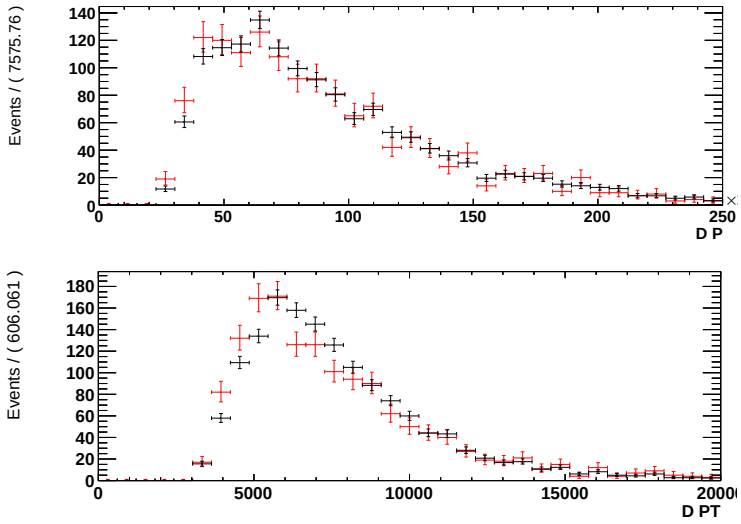


(a) MC

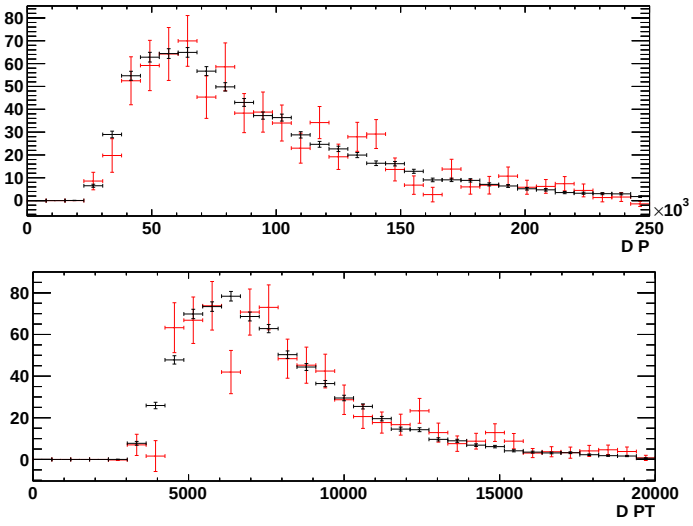


(b) Data

Figure F.2: Transverse momentum distributions in data and MC, selected with `L0HadronTOS||L0MuonTOS` (black) and nominal `L0HadronTIS` (red).



(a) MC

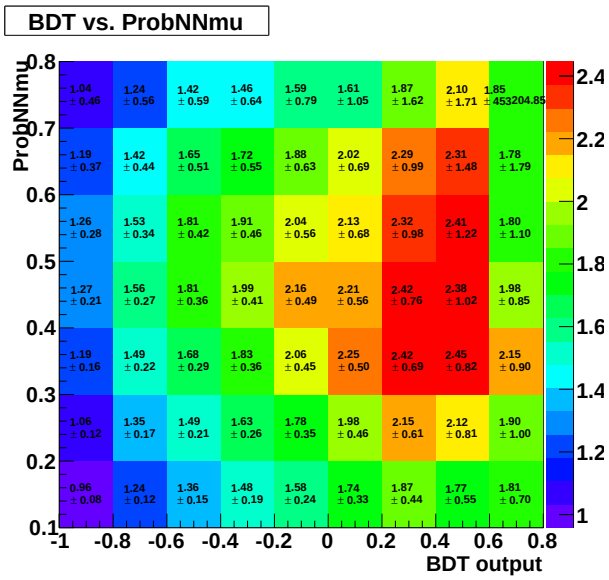


(b) Data

Figure F.3: Transverse momentum distributions in data and MC, selected with `L0HadronTOS||L0MuonTOS` (black) and nominal `L0HadronTIS` (red).

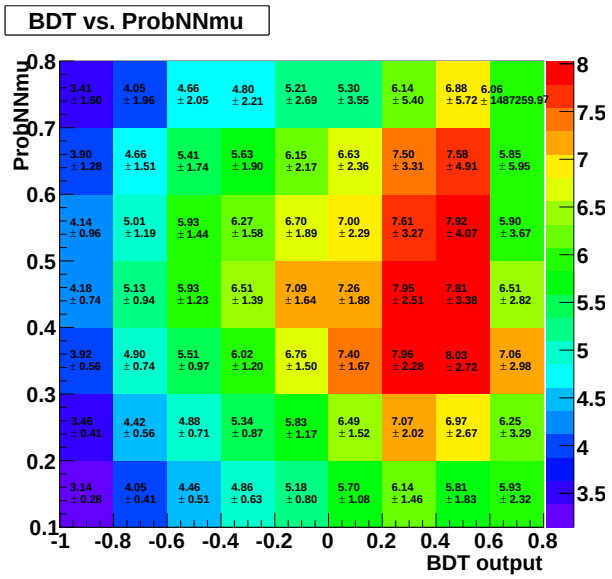
Appendix G

Scans for sensitivity studies



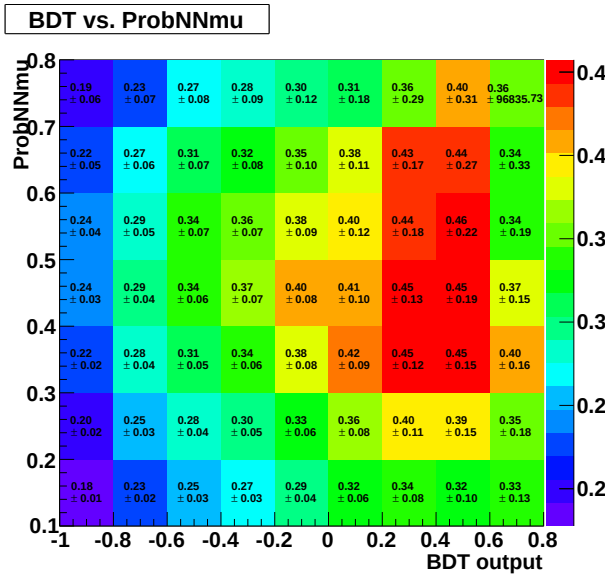
(a)

Figure G.1: Scan for the best significance on the total $D^0 \to K^+\pi^-\mu^+\mu^-$ branching ratio, with LHCb's Run II data.



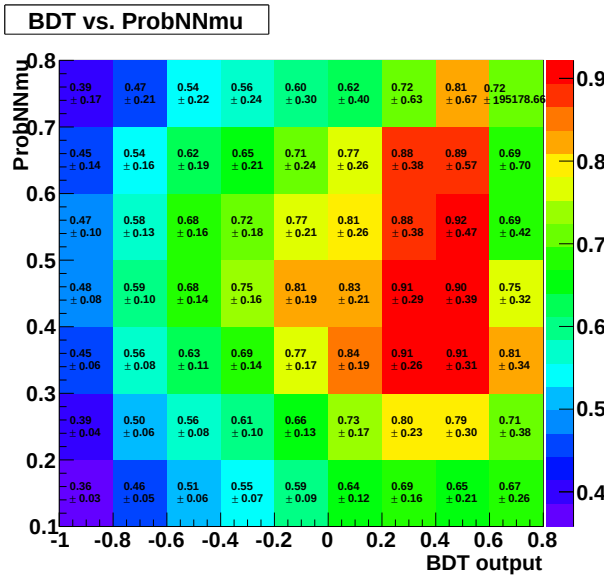
(a)

Figure G.2: Scan for the best significance that can be reached on the total $D^0 \to K^+\pi^-\mu^+\mu^-$ branching ratio, using LHCb's Upgrade dataset.



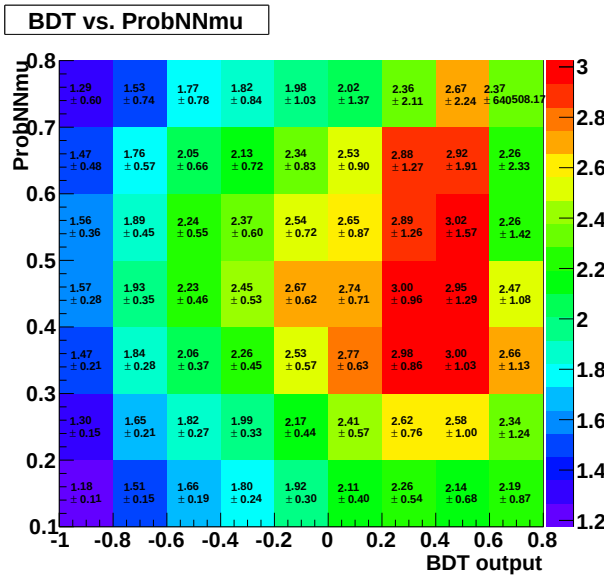
(a)

Figure G.3: Scan for the best significance that can be reached on the partial $D^0 \to K^+\pi^-\mu^+\mu^-$ branching ratio in the $m(\mu^+\mu^-) < 525$ MeV/c^2 region, using LHCb's Run I dataset and assuming $\Delta\mathcal{B}^{SM}(D^0 \to K^+\pi^-\mu^+\mu^-)_{m(\mu^+\mu^-)<525\ MeV/c^2} = 5 \times 10^{-9}$.



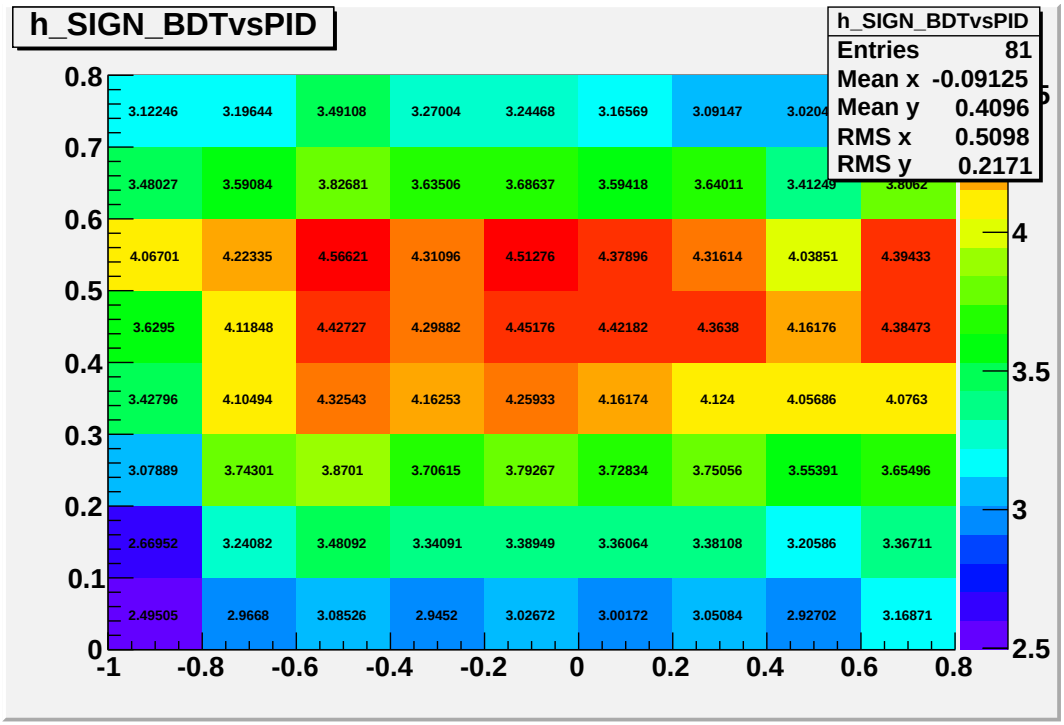
(a)

Figure G.4: Scan for the best significance that can be reached on the partial $D^0 \to K^+\pi^-\mu^+\mu^-$ branching ratio in the $m(\mu^+\mu^-) < 525$ MeV/c^2 region, using LHCb's Run II dataset and assuming $\Delta\mathcal{B}^{SM}(D^0 \to K^+\pi^-\mu^+\mu^-)_{m(\mu^+\mu^-)<525\ MeV/c^2} = 5 \times 10^{-9}$.

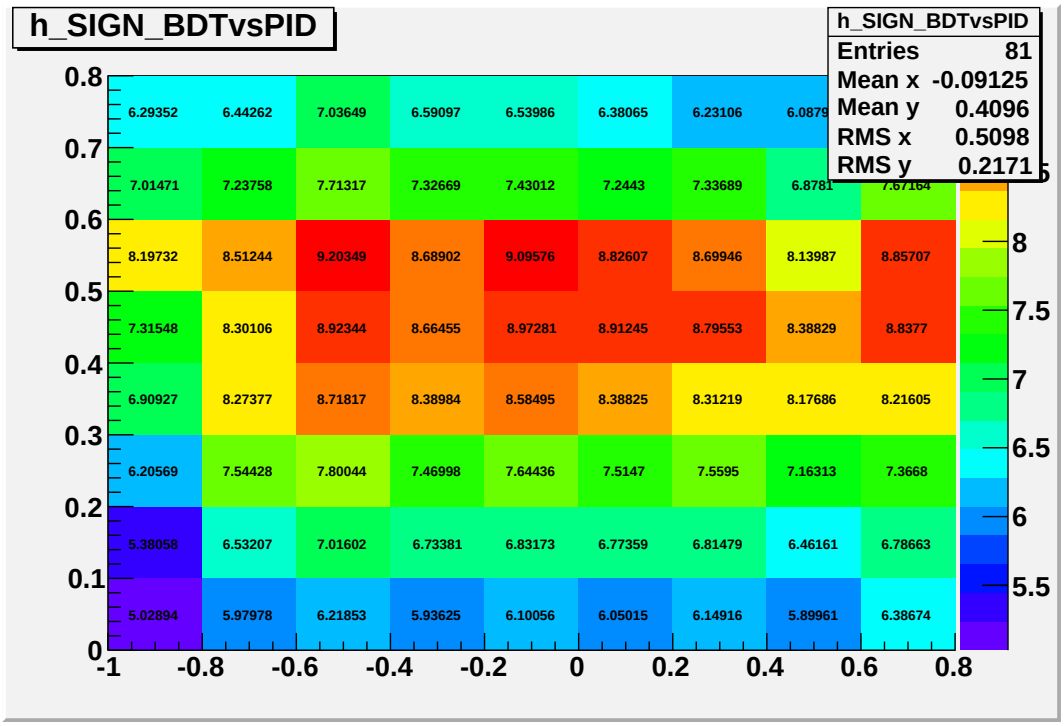


(a)

Figure G.5: Scan for the best significance that can be reached on the partial $D^0 \to K^+\pi^-\mu^+\mu^-$ branching ratio in the $m(\mu^+\mu^-) < 525$ MeV/c^2 region, using LHCb's Upgrade dataset and assuming $\Delta\mathcal{B}^{SM}(D^0 \to K^+\pi^-\mu^+\mu^-)_{m(\mu^+\mu^-)<525\ MeV/c^2} = 5 \times 10^{-9}$.

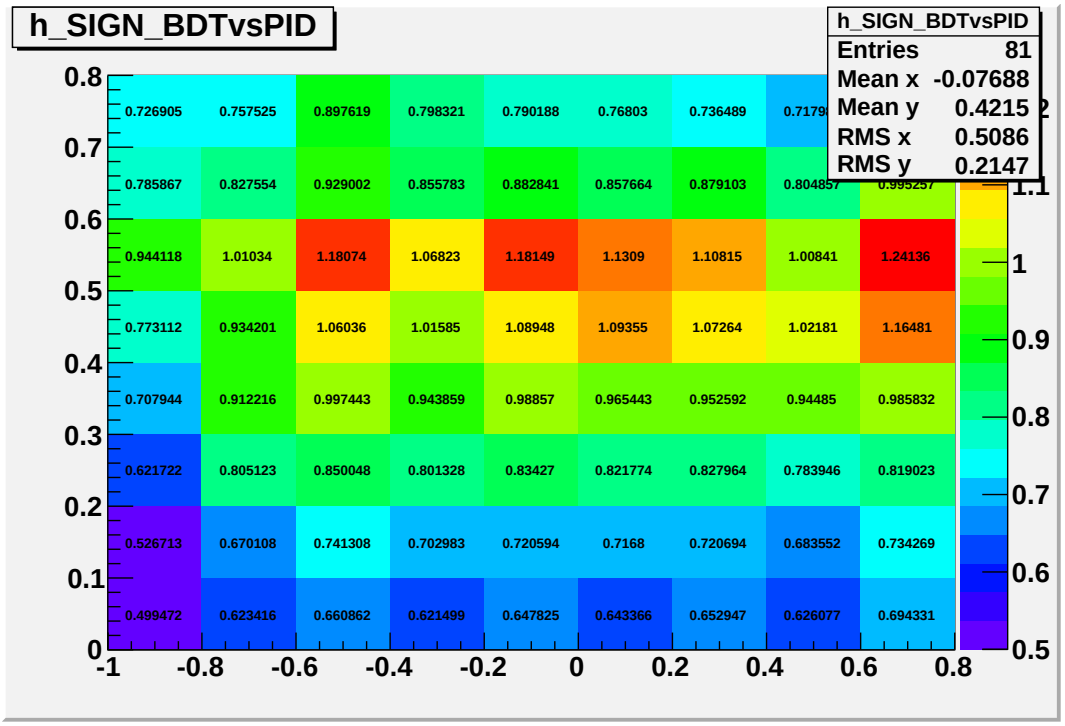


(a) LHCb Run I

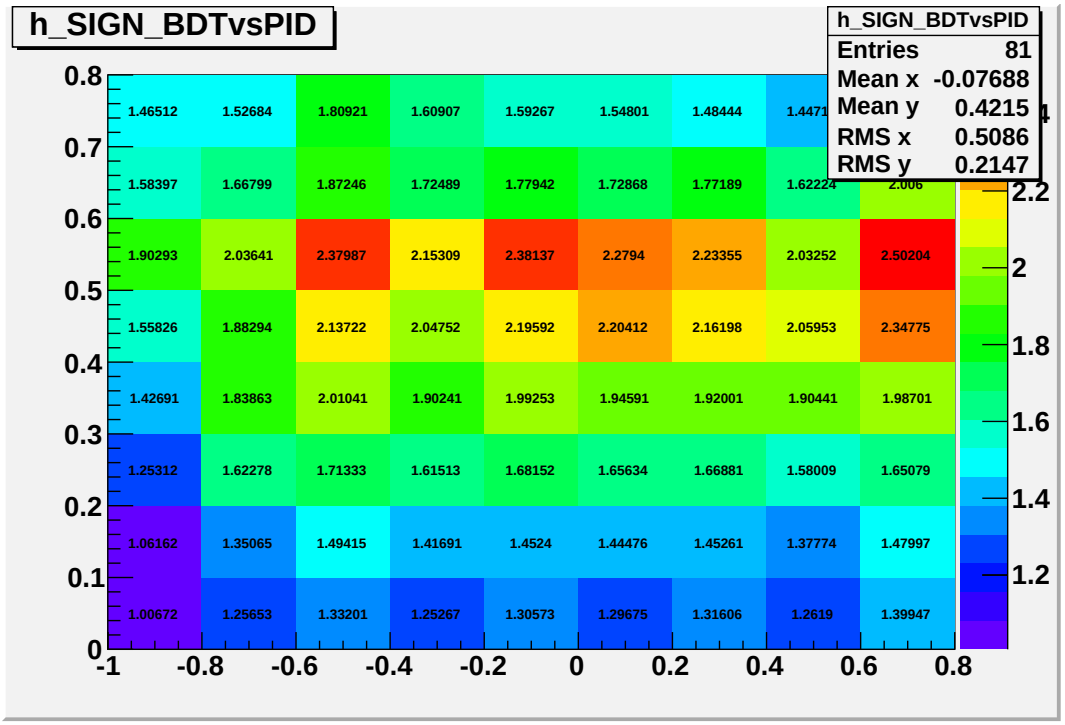


(b) LHCb Run II

Figure G.6: Scan for the best significance that can be reached on the total $D^0 \to K^+K^-\mu^+\mu^-$ branching ratio using LHCb's Run I (a) or Run II (b) datasets.

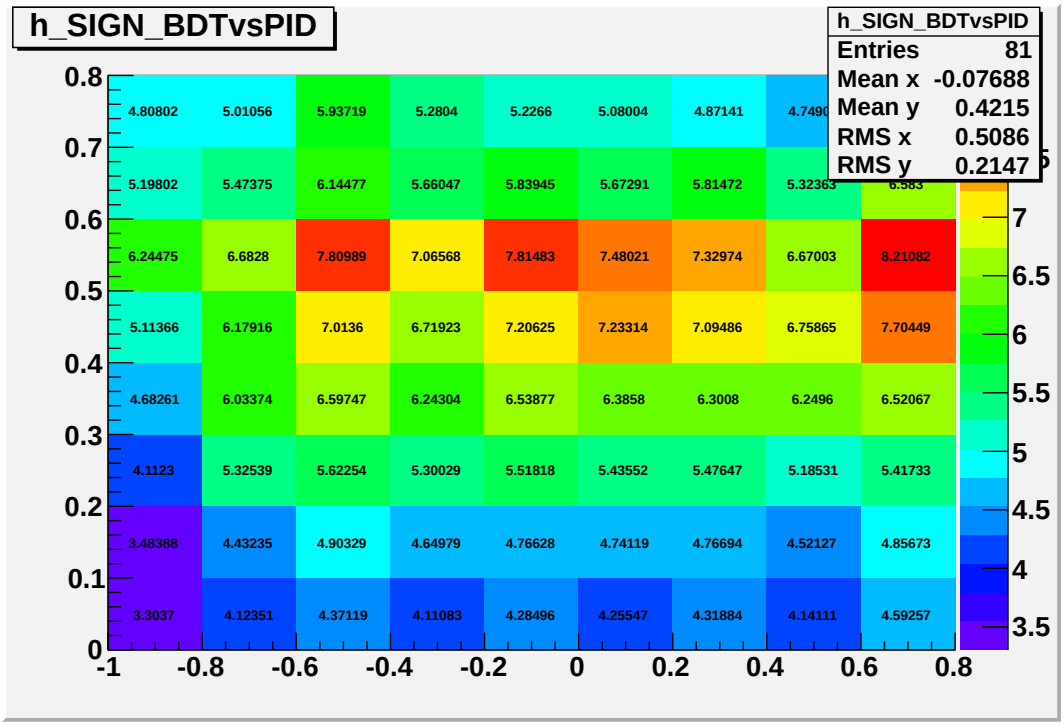


(a) LHCb Run I



(b) LHCb Run II

Figure G.7: Scan for the best significance that can be reached on the partial $D^0 \to K^+K^-\mu^+\mu^-$ branching ratio in the $m(\mu^+\mu^-) < 525$ MeV/c^2 region, using LHCb's Run I (a), Run II (b) datasets and assuming $\Delta\mathcal{B}^{SM}(D^0 \to K^+\pi^-\mu^+\mu^-)_{m(\mu^+\mu^-)<525\ MeV/c^2} = 1.5 \times 10^{-8}$.



(a) LHCb Upgrade

Figure G.8: Scan for the best significance that can be reached on the partial $D^0 \to K^+K^-\mu^+\mu^-$ branching ratio in the $m(\mu^+\mu^-) < 525$ MeV/c^2 region, using LHCb's Upgrade datasets and assuming $\Delta\mathcal{B}^{SM}(D^0 \to K^+\pi^-\mu^+\mu^-)_{m(\mu^+\mu^-)<525\ MeV/c^2} = 1.5 \times 10^{-8}$.

Bibliography

[1] LHCb Collaboration, R. Aaij et al., *Measurement of the CP-violating phase ϕ_s in the decay $B_s^0 \to J/\psi\phi$*, Phys. Rev. Lett. **108** (2012) 101803, arXiv:1112.3183.

[2] LHCb collaboration, R. Aaij et al., *Measurement of the CP asymmetry in $B^+ \to K^+\mu^+\mu^-$ decays*, Phys. Rev. Lett. **111** (2013) 151801, arXiv:1308.1340.

[3] LHCb collaboration, R. Aaij et al., *Differential branching fractions and isospin asymmetries of $B \to K^{(*)}\mu^+\mu^-$ decays*, arXiv:1403.8044.

[4] LHCb collaboration, R. Aaij et al., *Angular analysis of charged and neutral $B \to K\mu^+\mu^-$ decays*, JHEP **1405** (2014) 082, arXiv:1403.8045.

[5] LHCb collaboration, R. Aaij et al., *A study of CP violation in $B^\pm \to DK^\pm$ and $B^\pm \to D\pi^\pm$ decays with $D \to K^0_S K^\pm\pi^\mp$ final states*, Phys. Lett. **B733** (2014) 36, arXiv:1402.2982.

[6] LHCb collaboration, R. Aaij et al., *Measurement of CP observables in $B^0 \to DK^{*0}$ with $D \to K^+K^-$*, JHEP **1303** (2013) 067, arXiv:1212.5205.

[7] LHCb collaboration, R. Aaij et al., *Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and search for $B^0 \to \mu^+\mu^-$ decays at the LHCb experiment*, Phys. Rev. Lett. **111** (2013) 101805, arXiv:1307.5024.

[8] LHCb Collaboration, R. Aaij et al., *Differential branching fraction and angular analysis of the $B^+ \to K^+\mu^+\mu^-$ decay*, JHEP **1302** (2013) 105, arXiv:1209.4284.

[9] LHCb Collaboration, R. Aaij et al., *Measurement of the ratio of branching fractions $BR(B^0 \to K^{*0}\gamma)/BR(B_s^0 \to \phi\gamma)$ and the direct CP asymmetry in $B^0 \to K^{*0}\gamma$*, Nucl. Phys. **B867** (2013) 1, arXiv:1209.0313.

[10] LHCb collaboration, R. Aaij et al., *Measurement of CP asymmetry in $D^0 \to K^-K^+$ and $D^0 \to \pi^-\pi^+$ decays*, arXiv:1405.2797.

[11] LHCb collaboration, R. Aaij et al., *Measurement of $D^0 - \bar{D}^0$ mixing parameters and search for CP violation using $D^0 \to K^+\pi^-$ decays*, Phys. Rev. Lett. **111** (2013) 251801, arXiv:1309.6534.

[12] LHCb collaboration, R. Aaij et al., *Search for the rare decay $D^0 \to \mu^+\mu^-$*, Phys. Lett. **B725** (2013) 15, arXiv:1305.5059.

[13] LHCb Collaboration, R. Aaij et al., *Search for $D^+_{(s)} \to \pi^+\mu^+\mu^-$ and $D^+_{(s)} \to \pi^-\mu^+\mu^+$ decays*, Phys. Lett. B **724** (2013) 203, Comments: 21 pages, 5 figures.

[14] LHCb collaboration, R. Aaij et al., *Search for the decay $D^0 \to \pi^+\pi^-\mu^+\mu^-$*, Phys. Lett. B **728** (2013) 234.

[15] LHCb Collaboration, J. Alves, A. Augusto et al., *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

[16] E791 Collaboration, E. Aitala et al., *Search for rare and forbidden charm meson decays $D^0 \to Vl^+l^-$ and hhll*, Phys. Rev. Lett. **86** (2001) 3969, arXiv:hep-ex/0011077.

[17] M. Schmelling et al., *Integrated Charm Cross-Sections Measured by LHCb*, Internal note.

[18] L. Cappiello, O. Cata, and G. D'Ambrosio, *Standard Model prediction and new physics tests for $D^0 \to h^+h^-l^+l^-$ ($h = \pi, K$: $l = e, \mu$)*, JHEP **1304** (2013) 135, arXiv:1209.4235.

[19] CLEO Collaboration, G. Bonvicini et al., *Updated measurements of absolute D^+ and D^0 hadronic branching fractions and $\sigma(e^+e^- \to D\overline{D})$ at $E_{\rm cm} = 3774$ MeV*, Phys. Rev. **D89** (2014) 072002, arXiv:1312.6775.

[20] E. Noether, *Invariant Variation Problems*, Gott. Nachr. **1918** (1918) 235, arXiv:physics/0503066.

[21] T. Lee and C.-N. Yang, *Question of Parity Conservation in Weak Interactions*, Phys. Rev. **104** (1956) 254.

[22] C. Wu, E. Ambler, R. Hayward, D. Hoppes, and R. Hudson, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (1957) 1413.

[23] M. Goldhaber, L. Grodzins, and A. Sunyar, *Helicity of Neutrinos*, Phys. Rev. **109** (1958) 1015.

[24] J. Christenson, J. Cronin, V. Fitch, and R. Turlay, *Evidence for the 2 pi Decay of the k(2)0 Meson*, Phys. Rev. Lett. **13** (1964) 138.

[25] E. Fermi, *An attempt of a theory of beta radiation. 1.*, Z. Phys. **88** (1934) 161.

[26] S. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579.

[27] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[28] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) 367.

[29] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321.

[30] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132.

[31] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508.

[32] G. Guralnik, C. Hagen, and T. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585.

[33] ATLAS Collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, arXiv:1207.7214.

[34] CMS Collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, arXiv:1207.7235.

[35] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8** (1964) 214.

[36] G. Zweig, *An SU(3) model for strong interaction symmetry and its breaking. Version 2*, .

[37] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

[38] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak interactions with lepton-hadron symmetry*, Phys. Rev. D **2** (1970) 1285.

[39] M. Kobayashi and T. Maskawa, *CP violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973) 652.

[40] Particle Data Group, J. Beringer et al., *Review of Particle Physics (RPP)*, Phys. Rev. **D86** (2012) 010001.

[41] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[42] SLAC-SP-017 Collaboration, J. Augustin et al., *Discovery of a Narrow Resonance in e+ e- Annihilation*, Phys. Rev. Lett. **33** (1974) 1406.

[43] E598 Collaboration, J. Aubert et al., *Experimental Observation of a Heavy Particle J*, Phys. Rev. Lett. **33** (1974) 1404.

[44] G. Buchalla, *Heavy quark theory*, arXiv:hep-ph/0202092.

[45] J. Sakurai, *Theory of strong interactions*, Annals Phys. **11** (1960) 1.

[46] S. Fajfer and N. Kosnik, *Resonance catalyzed CP asymmetries in $D \to P l^+ l^-$*, Phys. Rev. **D87** (2013) 054026, arXiv:1208.0759.

[47] A. Paul, I. I. Bigi, and S. Recksiegel, *On $D \to X_u l^+ l^-$ within the Standard Model and Frameworks like the Littlest Higgs Model with T Parity*, Phys. Rev. **D83** (2011) 114006, arXiv:1101.6053.

[48] S. Fajfer, S. Prelovsek, and P. Singer, *Rare charm meson decays $D \to P\, l^+ l^-$ and $c \to u\, l^+ l^-$ in SM and MSSM*, Phys. Rev. **D64** (2001) 114009, arXiv:hep-ph/0106333.

[49] G. Burdman, E. Golowich, J. L. Hewett, and S. Pakvasa, *Rare charm decays in the standard model and beyond*, Phys. Rev. **D66** (2002) 014009, arXiv:hep-ph/0112235.

[50] G. Burdman and I. Shipsey, *D^0 - $\bar{D}^0$ mixing and rare charm decays*, Ann. Rev. Nucl. Part. Sci. **53** (2003) 431, arXiv:hep-ph/0310076.

[51] A. Paul, A. de La Puente, and I. I. Bigi, *Manifestations of Warped Extra Dimension in Rare Charm Decays and Asymmetries*, arXiv:1212.4849.

[52] S. Fajfer, N. Kosnik, and S. Prelovsek, *Updated constraints on new physics in rare charm decays*, Phys. Rev. **D76** (2007) 074010, arXiv:0706.1133.

[53] LHCb Collaboration, R. Aaij et al., *Evidence for CP violation in time-integrated $D^0 \to h^- h^+$ decay rates*, Phys. Rev. Lett. **108** (2012) 111602, arXiv:1112.0938.

[54] I. Dorsner, S. Fajfer, J. F. Kamenik, and N. Kosnik, *Can scalar leptoquarks explain the f(D(s)) puzzle?*, Phys. Lett. **B682** (2009) 67, arXiv:0906.5585.

[55] N. Kosnik, *Talk, Theory implication for Rare charm decays on the workshop on the Implications of LHCb measurements and future prospects, CERN, 14-16 October 2013* , https://indico. cern. ch/event/255380/session/2/contribution/14 (2013).

[56] S. Fajfer and S. Prelovsek, *Search for new physics in rare D decays*, Conf. Proc. **C060726** (2006) 811, arXiv:hep-ph/0610032.

[57] S. Fajfer and S. Prelovsek, *Effects of littlest Higgs model in rare D meson decays*, Phys. Rev. **D73** (2006) 054026, arXiv:hep-ph/0511048.

[58] S. Fajfer and N. Kosnik, *Leptoquarks in FCNC charm decays*, Phys. Rev. **D79** (2009) 017502, arXiv:0810.4858.

[59] G. Isidori and J. F. Kamenik, *Shedding light on CP violation in the charm system via D to V gamma decays*, Phys. Rev. Lett. **109** (2012) 171801, arXiv:1205.3164.

[60] J. Lyon and R. Zwicky, *Anomalously large $\mathcal{O}_8$ and long-distance chirality from $A_{\rm CP}[D^0 \to (\rho^0, \omega)\gamma](t)$*, arXiv:1210.6546.

[61] G. F. Giudice, G. Isidori, and P. Paradisi, *Direct CP violation in charm and flavor mixing beyond the SM*, JHEP **1204** (2012) 060, arXiv:1201.6204.

[62] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[63] CMS Collaboration, S. Chatrchyan et al., *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.

[64] ALICE Collaboration, K. Aamodt et al., *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[65] The Gaudi project, *http://proj-gaudi.web.cern.ch/proj-gaudi* .

[66] The Brunel project, *http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/*.

[67] ROOT, *http://root.cern.ch*.

[68] LHCb collaboration, I. Bediaga et al., *Framework TDR for the LHCb Upgrade: Technical Design Report*, Tech. Rep. CERN-LHCC-2012-007. LHCb-TDR-12, CERN, Geneva, Apr, 2012.

[69] C. LHCb, *LHCb Trigger and Online Upgrade Technical Design Report*, Tech. Rep. CERN-LHCC-2014-016. LHCB-TDR-016, CERN, Geneva, May, 2014.

[70] L. Collaboration, *LHCb VELO Upgrade Technical Design Report*, Tech. Rep. CERN-LHCC-2013-021. LHCB-TDR-013, CERN, Geneva, Nov, 2013.

[71] L. Collaboration, *LHCb Tracker Upgrade Technical Design Report*, Tech. Rep. CERN-LHCC-2014-001. LHCB-TDR-015, CERN, Geneva, Feb, 2014.

[72] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **0605** (2006) 026, `arXiv:hep-ph/0603175`.

[73] LHCb Collaboration, I. Belyaev et al., *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, IEEE Nucl. Sci. Symp. Conf. Rec. **2010** (2010) 1155.

[74] D. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[75] GEANT4, S. Agostinelli et al., *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[76] LHCb Collaboration, M. Clemencic et al., *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[77] R. Kalman, *A new approach to linear filtering and prediction problems*, Journal of Basic Engineering **D82** (1960) 35, `arXiv:physics/0402083`.

[78] R. Aaij et al., *The LHCb Trigger and its Performance in 2011*, J. Instrum. **8** (2012) P04022. 31 p.

[79] LHCb MuonID group, F. Archilli et al., *Performance of the Muon Identification at LHCb*, J. Instrum. **8** (2013) P10020. 17 p, Comments: 17 pages, 10 figures.

[80] M. De Cian et al., *Measurement of the track finding efficiency*, Tech. Rep. LHCb-PUB-2011-025. CERN-LHCb-PUB-2011-025, CERN, Geneva, Apr, 2012.

[81] D0 Collaboration, V. Abazov et al., *Search for flavor-changing-neutral-current D meson decays*, Phys. Rev. Lett. **100** (2008) 101801, arXiv:0708.2094.

[82] E. Majorana, *Teoria simmetrica dell'elettrone e del positrone*, Il Nuovo Cimento **14** (1937), no. 4 171.

[83] BaBar Collaboration, J. Lees et al., *Searches for Rare or Forbidden Semileptonic Charm Decays*, Phys. Rev. **D84** (2011) 072006, arXiv:1107.4465.

[84] A. L. Read, *Presentation of search results: the CLs technique*, Journal of Physics G: Nuclear and Particle Physics **28** (2002), no. 10 2693.

[85] S. Stahl, *Smearing of track parameters to account for resolution differences in data and monte carlo*, LHCb-INT-2011-027.

[86] CLEO Collaboration, M. Artuso et al., *Amplitude analysis of $D^0 \to K^+K^-\pi^+\pi^-$*, Phys. Rev. **D85** (2012) 122002, arXiv:1201.5716.

[87] P. Nason, *MINT: A Computer program for adaptive Monte Carlo integration and generation of unweighted distributions*, arXiv:0709.2085.

[88] The LHCb Particle ID Group, https://twiki.cern.ch/twiki/bin/view/LHCb/PIDCalibPackage.

[89] LHCb collaboration, R. Aaij et al., *Measurement of the track reconstruction efficiency at LHCb*, arXiv:1408.1251.

[90] LHCb collaboration, R. Aaij et al., *Measurement of the track reconstruction efficiency at LHCb*, Tech. Rep. arXiv:1408.1251. CERN-LHCB-DP-2013-002, Aug, 2014.

[91] LHCb collaboration, R. Aaij et al., *Searches for violation of lepton flavour and baryon number in tau lepton decays at LHCb*, Phys. Lett. **B724** (2013) 36, arXiv:1304.4518.

[92] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, arXiv:physics/0306116.

[93] A. Hocker et al., *TMVA - Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039.

[94] Y. Freund and R. E. Schapire, *A decision-theoretic generalization of on-line learning and an application to boosting*, Journal of Computer and System Sciences **55** (1997), no. 1 119 .

[95] J. H. Friedman, *Greedy function approximation: A gradient boosting machine.*, The Annals of Statistics **29** (2001) 1189.

[96] D. Benbouzid, R. Busa-Fekete, N. Casagrande, F.-D. Collin, and B. Kégl, *Multiboost: a multi-purpose boosting package*, The Journal of Machine Learning Research **13** (2012), no. 1 549.

[97] M. Williams, *Generic D from B Selections*, Tech. Rep. LHCb-INT-2012-002. CERN-LHCb-INT-2012-002, CERN, Geneva, Jan, 2012.

[98] I. I. Bigi and A. Paul, *On CP Asymmetries in Two-, Three- and Four-Body D Decays*, JHEP **1203** (2012) 021, arXiv:1110.2862.

[99] LHCb collaboration, R. Aaij et al., *Measurement of Form-Factor-Independent Observables in the Decay $B^0 \to K^{*0}\mu^+\mu^-$*, Phys. Rev. Lett. **111** (2013), no. 19 191801, arXiv:1308.1707.

[100] L. Collaboration, *LHCb PID Upgrade Technical Design Report*, Tech. Rep. CERN-LHCC-2013-022. LHCB-TDR-014, CERN, Geneva, Nov, 2013.

[101] D. Gascon et al., *Low noise front end ASIC with current mode active cooled termination for the Upgrade of the LHCb Calorimeter*, IEEE Nucl. Sci. Symp. Conf. Rec. **2011** (2011) 1445.

[102] C. Abellan Betetaa et al., *Study of a solution with COTS for the LHCb calorimeter upgrade*, Physics Procedia **37** (2012) 1744.

Acknowledgements

First of all I would like to thank the director of the LAL, Achille Stocchi, for giving me the opportunity to work in France, in this scientifically strong laboratory. I am also grateful to Fabien Cavalier, the assistant director, thanks to whom my participation to EPS-HEP conference and CERN-JINR school was possible.

I owe my deepest gratitude to the members of the jury, Svjetlana Fajfer and Urs Langenegger, that have accepted to read and evaluate my work. I have been really honoured. I am very thankful to Vera Luth and Guy Wikinson for being my referees. I appreciate very much your effort of reading my thesis so carefully and giving me such useful and interesting comments.

I would like to thank Marie-Hélène Schune for accepting me to do the internship and giving me the opportunity to work within such a wonderful group. I am thankful to her for millions of advice concerning physics, life and future. During these three years and a half she was a very important person for me, looking at whom I was always thinking: "This is how Women in science should be!". I am really sad to leave her noisy but very comfortable office.

I owe a very important debt to my supervisor, Benoit Viaud, without whom this thesis would not have been possible. I have learnt so much from you. All this work, all what I know and can is thanks to you. Thank you for this huge effort to share your knowledge with me. I remember how you were preparing me for the interview for the école doctoral, at that moment I realised that I am very lucky to have such supervisor. You were always available to answer any question, solve any problem. I am very thankful to you!

I would particularly like to thank Yasmine, who was supporting me since she arrived. It is impressive how many precious advice she gave me, how many new ideas she brought to the analysis. I will miss her so much, which makes me think that I will still need her advice and I will not stop asking them. I will miss her nice tea and delicious candies that always helped to cheer me up when needed.

I am truly and deeply indebted to many people who helped me a lot to prepare and correct this thesis, of course Benoit, Yasmine but also Emi Kou, Nejc Kosnic, Djalel Benbouzid for fruitful comments. I want to say "thank you" to Frederic Machefert that helped me a lot with the part of the thesis concerning the test beam of the electronic board.

I would like to show my greatest appreciation to Grigory Feofilov, my supervisor at Saint-Petersburg State University. I have discovered CERN thanks to this person. He is more than just a scientific supervisor for me, he is also my mentor and almost a member of my family.

I am deeply grateful to my school teacher of physics, Dobrohotov Sergey Borisovich.

He awakened in me love for physics, which determined my future profession. I still remember his original explanations of quite ordinary things (for example, why elephants have big ears), which helped to realize that the physics is everywhere around! Without his first impulse I would not have come to this thesis.

I am very thankful to all LHCb@LAL group: Patrick, Guy, Sergey, Jacques, Bernard. You were always here for any kind of help and advice. I also want to thank LAL students that became my dear friends: Alexandra, Alexis, Claire, Michelle, Maksim, Marta, Maria, Martino. I loved our pause café on the grass, our student breakfasts, our trips to Seillac and JRJC. These were great moments!

I would like to thank my close friends outside the LAL: Denis (thanks to him I arrived at LAL), Katya&Adrien, Antoine, Axel, Sasha, Igor', Liza; friends that are very far away from me now, in cold Russia: Anya, Sveta&Sasha, Denis&Nastya, Zhenya, Dashka, Ilya, Olkin. I am very fortunate and very lucky to have you in my life.

I would like to thank my family, my mother and my sister, who unfortunately could not come to share this moment with me. I am deeply grateful to my mom for making possible for me to get a good education. I also want to thank my cousin Lena, who is the only one from my family who was able to come for the defence and who also helped me with this acknowledgement.

Je tiens à remercier ma belle famille. Il est injuste de dire "belle famille", car ils sont devenus tout simplement "famille" pour moi. Je suis très reconnaissante qu'ils aient accepté cette étrange fille russe (c'est-à-dire moi) avec le coeur et l'âme ouvert. Je voudrais leur exprimer ma plus grande gratitude pour avoir trouvé le temps de venir pour ma soutenance et pour leur aide.

Enfin, je tiens à dire merci à une personne merveilleuse en plus, mon conjoint. Merci beaucoup de me faire découvrir la France, la culture francaise, la cuisine et de m'apprendre le francais. Je sais que parfois je n'étais pas facile, surtout depuis quelques semaines, mais je veux que tu saches que tu m'as parfaitement soutenu! Tu as préparé mes plats préférés, t'es occupé de l'appartement, m'as prêté ta voiture. Merci pour tes millions de marques d'affections dans les moment difficiles, j'étais vraiment inquiète et j'en avais besoin. Elle m'ont toujours aidées! Je suis tellement heureuse que nous nous soyons trouvésé. C'est un tel plaisir de partager des moments agréables et durs avec toi. Je t'aime tellement!

Спасибо всем!