HAL
open science

A phenomenological study of D+ → π+ I+ I− and the search for B(s) → τ+ τ− at LHCb

Cedric Meaux

To cite this version:

Cedric Meaux. A phenomenological study of D+ → π+ I+ I− and the search for B(s) → τ+ τ− at LHCb. High Energy Physics - Experiment [hep-ex]. Aix Marseille Université (AMU); CPPM; CPT, 2019. English. NNT : . tel-03597402

HAL Id: tel-03597402

https://hal.science/tel-03597402v1

Submitted on 4 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

AIX-MARSEILLE UNIVERSITÉ
ECOLE DOCTORALE PHYSIQUE ET SCIENCES DE LA MATIÈRE

Centre de Physique des Particules de Marseille (CPPM) UMR 7346

Thèse présentée pour obtenir le grade universitaire de docteur

Discipline: Physique et Sciences de la Matière
Spécialité: Physique des Particules et Astroparticules

Cédric MÉAUX

A phenomenological study of $D^+ \to \pi^+\ell^+\ell^-$ and the search for $B^0_{(s)} \to \tau^+\tau^-$ at LHCb

Soutenue le 28/10/2019 devant le jury composé de :

Viola SORDINI	IP2I/Université C. Bernard	Rapportrice
Svjetlana FAJFER	IJS/Ljubljana University	Rapportrice
Eli BEN-HAIM	LPNHE/Sorbonne Université	Examinateur
Cristinel DIACONU	CPPM/Aix-Marseille Université	Examinateur
Giampiero MANCINELLI	CPPM/Aix-Marseille Université	Directeur de thèse
Aoife BHARUCHA	CPT/Aix-Marseille Université	Co-directrice de thèse

This work has been carried out thanks to the support of the OCEVU Labex (ANR-11- LABX-0060) and the A*MIDEX project (ANR-11-IDEX-0001-02) funded by the "Investissements d'Avenir" French government program managed by the ANR.

Numéro national de thése/suffixe local : 2019AIXM0370/059ED352

Suffixe laboratoire : CPPM-T-2019-05

Contents

Remerciements **2**

Résumé **3**

Acronyms **23**

1 Probing physics beyond the Standard Model with FCNC meson decays **27**
1.1 The Standard Model of particle physics 27
1.1.1 On the flavor structure of the Standard Model 27
1.1.2 On the color structure of the Standard Model 30
1.2 FCNC decays as a powerful probe of SM 31
1.3 Effective Field Theory 34
1.4 On the purely leptonic decay $M \to \ell^+\ell^-$ 38
1.4.1 Theoretical background 38
1.4.2 The beauty case 39
1.4.3 The charmed case 41
1.5 On the semileptonic decay $D^+ \to \pi^+\ell^+\ell^-$ 41

I A phenomenological study of $D^+ \to \pi^+\ell^+\ell^-$ **47**

2 Theoretical framework for the non-resonant $D^+ \to \pi^+\ell^+\ell^-$ amplitude **49**
2.1 The naive amplitude 49
2.2 QCD factorization 51
2.2.1 Weak annihilation corrections 53
2.2.2 Hard spectator scattering corrections 54
2.2.3 Form factor corrections 55
2.3 Calculation of the Wilson coefficients 55
2.3.1 Solving the RGE 56
2.3.2 Calculation of the C_{1-8} Wilson coefficients 57
2.3.3 Calculation of the C_9 Wilson coefficient 58

3 Modeling the resonant structure and phenomenological analysis **61**
3.1 Modeling the quark vacuum polarizations 61
3.2 Numerical and phenomenological analysis 67
3.2.1 Parametrization and numerical inputs 68
3.2.2 Full differential decay width distribution 71
3.2.3 The BSM Wilson coefficients constraints 76

3.2.4 Observables . . . 80
3.3 Conclusions . . . 84

II Search for $B^0_{(s)} \to \tau^+\tau^-$ at LHCb **87**

4 The LHCb experiment **89**
4.1 The LHC Environment . . . 89
4.1.1 The CERN accelerator complex . . . 89
4.1.2 B physics at the LHC . . . 91
4.2 The LHCb detector . . . 93
4.2.1 LHCb geometry . . . 93
4.2.2 Tracking system . . . 95
4.2.3 Particle identification . . . 98
4.2.4 Trigger system and the stripping selection . . . 100
4.2.5 Simulation . . . 104
4.3 Experimental prospects . . . 104
4.3.1 e^+e^- colliders versus pp colliders . . . 104
4.3.2 The future of the LHCb detector . . . 105

5 Introduction to the search for the $B^0_{(s)} \to \tau^+\tau^-$ decays **107**
5.1 General context . . . 107
5.1.1 On the τ lepton properties . . . 107
5.1.2 Previous results . . . 109
5.2 Overview of the published LHCb analysis . . . 109
5.2.1 Analysis regions . . . 110
5.2.2 Analysis workflow . . . 111
5.3 Signal reconstruction and background rejection . . . 114
5.3.1 Signal reconstruction . . . 114
5.3.2 Main backgrounds for the $(3\pi, 3\pi)$ final state . . . 115
5.3.3 Main backgrounds for the $(3\pi, \mu)$ final state . . . 117
5.3.4 Discriminating variables . . . 117
5.4 Multivariate analyses . . . 122
5.4.1 Introduction to multivariate analysis techniques . . . 122
5.4.2 Boosted Decision Tree . . . 124
5.4.3 Artificial Neural Network . . . 125
5.4.4 k-folding . . . 125
5.4.5 Flattening . . . 126
5.4.6 Iterative procedure to select input variables entering the BDT 126
5.4.7 Optimization of the BDT parameters . . . 128

6 Search for $B^0_{(s)} \to \tau^+\tau^-$ via the $(3\pi, \mu)$ final state **129**
6.1 Analysis overview . . . 129
6.2 Selection . . . 131
6.2.1 Trigger and stripping selection . . . 132
6.2.2 Loose cut-based selection . . . 133
6.2.3 BDT-based selection . . . 134
6.2.4 Exclusive background suppression . . . 136
6.2.5 Qualitative study of the background composition . . . 142
6.2.6 Selection efficiencies . . . 142
6.3 Likelihood fit . . . 144

6.3.1 The fitted variable 144
6.3.2 Fit strategy 145
6.4 Normalization 148
6.4.1 Overview of the $B^0 \to D^- \pi^+$ analysis 149
6.4.2 Normalization factor 150
6.5 Results and conclusions 151
6.5.1 Limit estimate 152
6.5.2 A last test 152
6.5.3 Conclusions 156

7 Search for $B^0_{(s)} \to \tau^+\tau^-$ via the $(3\pi, 3\pi)$ final state 157
7.1 Analysis overview 157
7.2 Selection 158
7.3 Likelihood fit 164
7.3.1 The fitted variable 164
7.3.2 Fit strategy 168
7.4 Normalization 170
7.4.1 Overview of the $B^0 \to D^- D_s^+$ analysis 170
7.4.2 Normalization factors 173
7.5 Results and conclusions 174

8 Conclusions 177
8.1 Conclusions on the phenomenological study of the $D^+ \to \pi^+\ell^+\ell^-$ decay 177
8.2 Conclusions on the search for $B^0_{(s)} \to \tau^+\tau^-$ decays at LHCb 180

References 182

Remerciements

Je n'aurais jamais imaginé, avant de m'y avoir attelé, que la rédaction des remerciements d'une thèse était un exercice si délicat. Se remémorer trois années de sa vie, revenir sur les durs comme les bons moments, trouver les mots justes et finalement en faire une synthèse d'une page m'ont pris beaucoup plus de temps que ce que j'avais pensé. Tout les remerciements que j'ai lu mélangent vie privée et professionnelle et les miens n'échappent pas à cette règle.

Je tiens à remercier avant tout le LabEx OCEVU, dirigé par le Dr. Kajfasz, pour m'avoir fait confiance pour la réalisation de ce travail de recherche interdisciplinaire (un volet théorique et expérimental) ; les moyens financiers offerts par le LabEx m'ont permis d'arriver sereinement à l'achèvement de cette thèse. Ma reconnaissance va également à l'ensemble de mon jury de thèse : Le Dr. Sordini et le Pr. Fajfer pour leur lecture attentive de ce document et ce en un temps (très) limité, le Pr. Eli Ben-Haim et le Pr. Diaconu (Directeur du CPPM lors de ma soutenance), pour leurs rôles d'examinateurs, et enfin, mes deux directeurs de thèse, le Dr. Giampiero Mancinelli et le Dr. Aoife Bharucha.

La patience d'Aoife, sa capacité à me remotiver et ses efforts pour combler le fossé entre théoriciens et expérimentateurs (dont je suis issu) ont été indispensables à la prise en main, parfois difficile, de mon sujet théorique. J'ai été son premier thésard mais sa supervision n'en a rien laissé paraître. Giampiero, toujours plein d'énergie, a su me redonner de l'allant les jours où il m'en manquait. Sa rigueur et sa logique dans les techniques d'analyse ont été pour moi des rails auxquels je me suis parfois heurté mais qui ont finalement réussi à me faire garder le cap. Bien que n'étant pas officiellement ma directrice de thèse, Justine Serrano, s'est comportée en tant que telle tout du long. Elle m'a fortement aidé, avec Giampiero, à comprendre les subtilités des analyses et les aspects techniques des données enregistrées par LHCb. Leurs supervisions a su se faire discrètes quand nécessaire et appuyées quand il le fallait. Ils ont su apporter leur aide, tout en me laissant la liberté d'explorer mes propres idées, marge de liberté que je crois indispensable à la recherche.

Je tiens également à remercier toutes les personnes travaillant au CPPM pour leur accueil. Notamment, le personnel de l'équipe de direction pour son encadrement, le personnel du service informatique pour son aide à l'installation et le service comptabilité pour la gestion de mes nombreuses missions.

Aussi une thèse sans amitiés ne serait que bien peu de chose et je souhaite remercier tout ceux avec qui j'ai partagé des verres, des passions et des débats animés. Notamment, Vladimir, Clara et Christophe pour les bons moments passés autour d'une bière ou sur le rocher. Nghia et Grieg, camarades de promotion embarqués sur

le même bateau pendant plus de quatre ans. Guillaume, mes proches et ma famille pour leur soutien et leur patience indéfectible. Et enfin, Boris et ses acolytes pour leur soutien moral indispensable. La liste n'est bien sûr pas exhaustive mais que personne ne se sente oublié.

Résumé

La théorie physique utilisée pour décrire l'ensemble des particules élémentaires connues et leurs intéractions est le Modèle Standard de la physique des particules (MS) [1–8]. En dépit d'une remarquable capacité prédictive, le MS apparaît comme incomplet et la grande majorité des physiciens s'accordent pour le percevoir comme une version à basse énergie d'une théorie plus générale restant à découvrir et que l'on qualifiera par la suite de Nouvelle Physique (NP). Deux grandes stratégies sont utilisées pour la recherche de NP : la recherche directe et la recherche indirecte. Nous nous concentrons dans cette thèse sur la recherche indirecte et plus particulièrement sur les désintégrations de mésons lourds procédant par Changement de Saveur par Courant Neutre (CSCN). Ces désintégrations sont très rares dans le MS car elles procèdent via des diagrammes de Feynman en boucles et sont de plus chiralement défavorisées. Ces caractéristiques confèrent aux désintégrations CSCN une sensibilité accrue aux particules de NP. Expérimentalement, plusieurs anomalies ont été observées dans les transitions CSCN de type $b \to s\ell^+\ell^-$ (sondant le couplage de la NP avec les quarks de type down). Elles semblent suggérer des scénarios de nouvelle physique où les particules de NP auraient un couplage proportionnel à la masse des leptons dans l'état final, brisant ainsi une des hypothèses du MS appelée l'Universalité de Saveur Leptonique (USL), et mettant l'accent sur les recherches avec des leptons τ dans l'état final.
La recherche indirecte de NP repose sur deux ingrédients : des mesures expérimentales précises et des prédictions précises. Cette thèse porte sur ces deux ingrédients et est dédiée aux thématiques suivantes : Améliorer la recherche expérimentale de la désintégration $B^0_{(s)} \to \tau^+\tau^-$ avec le détecteur LHCb, sondant ainsi les transitions $b \to d(s)\tau^+\tau^-$, et améliorer les prédictions et la phénoménologie autour de la désintégration $D^+ \to \pi^+\ell^+\ell^-$, sondant ainsi le couplage de la NP avec les quarks de type up $c \to u\ell^+\ell^-$.

Le Modèle Standard et les transitions CNCS

Le contenu en matière du MS, visible dans la figure 1, se répartie en quarks et leptons. Les quarks sont soit de type *up* (saveurs u, c, t avec une charge électrique de $+2/3$) ou de type *down* (saveurs d, s, b avec une charge électrique de -1/3). Deux théories de jauge décrivent trois des quatre intéractions fondamentales : la chromodynamique quantique (QCD) et la théorie électrofaible. La QCD décrit l'interaction forte entre les quarks et les gluons. La théorie électrofaible repose sur une symétrie qui est est brisée via le mécanisme de Higgs [9]. Cette brisure génèrent deux sous-théories : l'électromagnétisme, dont le médiateur est le photon, et l'intéraction faible, dont les médiateurs sont les bosons massifs neutre Z^0 et chargés $W^\pm$. L'intéraction faible par courants chargés (via $W^\pm$) est la seule interaction permettant un changement de saveur des quarks.

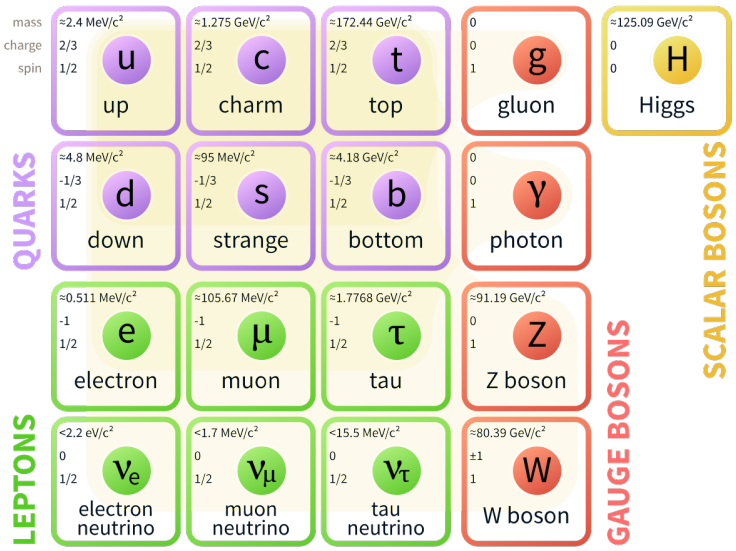


Figure 1: Contenu en particule élémentaire du Modèle Standard avec les quarks, de type *up* (u, c, t) avec une charge électrique $+2/3$ ou de type *down* avec charge -1/3 (d, s, b) et les leptons. Les bosons de jauge sont les médiateurs des intéractions entre particules.

Ce trait du MS a une conséquence importante : les transitions CSCN, par exemple de type $q_1 \to q_2\ell^+\ell^-$, où q_1 et q_2 sont tous deux des quarks de type up (ou down) et ℓ des leptons, procèdent dans le MS uniquement via des diagrammes de Feynman en boucle comme ceux montrés en figure 2. Les particules intervenant dans les boucles sont dites virtuelles, car elles existent en dehors de leurs couches de masse. C'est cette caractéristique qui confère aux transitions CSCN une grande sensibilité aux potentielles particules de NP, ces dernières pouvant remplacer ou générer de nouveaux diagrammes de Feynman, modifiant ainsi de manière importante des observables tel que le rapport d'embranchement[1].

Suivant une des propriétés de la QCD, les quarks ne peuvent être observés seuls et sont confinés au sein d'hadrons. Si les quarks q_1 et q_2 font partie du même état initial, on parlera de désintégrations purement leptonique : $M \to \ell^+\ell^-$ où M est un méson neutre composé de q_1 et q_2. Si q_1 est lié à un autre quark (que l'on qualifiera de quark spectateur) et que q_2 fait partie de l'état final, on parlera de désintégrations semileptoniques : $M \to P\ell^+\ell^-$ où P est le méson de l'état final constitué du quark q_2 et du quark spectateur.

Motivations pour la recherche de $B^0_{(s)} \to \tau^+\tau^-$

Il s'avère que plusieurs anomalies ont été mesurées dans les désintégrations semileptoniques invoquant des transitions $b \to s\ell^+\ell^-$. Certaines de ces anomalies ont trait directement à l'USL, i.e. : $R_K = [\mathcal{B}(B^+ \to K^+\mu^+\mu^-)] / [\mathcal{B}(B^+ \to K^+e^+e^-)]$ [10–14] et $R_{K^*} = \left[\mathcal{B}(B^0 \to K^{*0}\mu^+\mu^-)\right] / \left[\mathcal{B}(B^+ \to K^{*0}e^+e^-)\right]$ [10, 11, 15]. De plus, la mesure de $R(D^{(*)}) = \left[\mathcal{B}(B^0 \to D^{(*)}\tau^+\nu_\tau)\right] / \left[\mathcal{B}(B^0 \to D^{(*)}\ell^+\nu_\ell)\right]$, où ℓ^+ est soit un

[1] Le rapport d'embranchement ($\mathcal{B}$) d'une désintégration $A \to X$ peut être vu comme la probabilité que A se désintègre dans l'état final X.

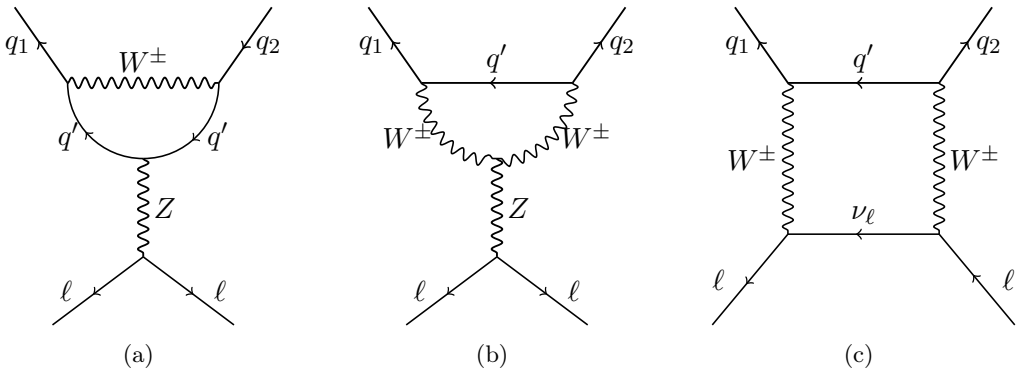


Figure 2: Diagrammes de Feynman dominant dans le MS pour les transitions CSCN $q_1 \to q_2\ell^+\ell^-$ de type pingouins (a,b) et boîte (c). Le type (down or up) de q_1 and q_2 détermine le type de q' (up or down) et la charge du boson $W^\pm$.

muon ou un électron a été trouvée plus que grande que la prédiction du MS [16–20], pointant vers une violation de l'USL dans la transition[2] $b \to c\ell\nu$. L'ensemble de ces anomalies suggèrent des scénarios de NP où le couplage des particules de NP est proportionnel à la masse des leptons. Il est donc particulièrement intéressant d'étudier des états finaux avec les leptons les plus massifs, à savoir les leptons τ.

De possible explications pour ces anomalies sont les leptoquarks (e.g. [21, 22]), des théories prédisant de nouveaux bosons de jauge W'/Z' (e.g. [23]) et les modèles avec doublet de Higgs (e.g. [24]). Dans ces modèles, le rapport d'embranchement des désintégrations purement leptoniques $B^0_{(s)} \to \tau^+\tau^-$ pourrait être augmenté par rapport à la valeur du MS[3], $\mathcal{B}(B^0 \to \tau^+\tau^-) = (2.22 \pm 0.19) \times 10^{-8}$ et $\mathcal{B}(B^0_s \to \tau^+\tau^-) = (7.73 \pm 0.49) \times 10^{-7}$ [28], de plusieurs ordres de grandeur [24, 29–31], il est donc fondamental de poursuivre la recherche de ces désintégrations.

Motivations pour la recherche de $D^+ \to \pi^+\ell^+\ell^-$

La plupart des efforts (expérimentaux et théoriques) se sont concentrés initialement sur les transitions invoquant des quarks de type *down*, car les prédictions théoriques liées à ces derniers sont souvent plus faciles que celles associées aux transitions $c \to u\ell^+\ell^-$. Ces transitions peuvent cependant être sensibles elles aussi à de la NP et il apparaît nécessaire aujourd'hui d'améliorer leurs prédictions théoriques. Un des meilleurs canaux au LHC invoquant cette transition est sans aucun doute la désintégration $D^+ \to \pi^+\ell^+\ell^-$. Plusieurs modèles de NP, tel que la supersymétrie avec violation de la parité R ou des modèles invoquant une quatrième génération de quarks, prévoient des valeurs de rapport d'embranchement beaucoup plus grande que celle du MS [32, 33]; il est donc fondamental d'améliorer les prédictions autour de cette désintégration.

[2] La transition $b \to c\ell\nu$ n'est pas une transition de type CSCN, il s'agit d'une transition de type Changement de Saveur par Courant Chargé (CSCC).

[3] Et ce bien que $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ ait été trouvé en accord avec le MS [25–27].

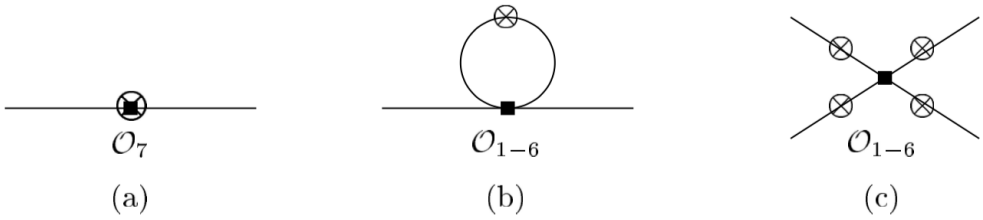


Figure 3: Trois des quatre contributions principales à $D^+ \to \pi^+\ell^+\ell^-$. Le carré noir symbolise l'opérateur effectif $\mathcal{O}_i$; la croix cerclée marque l'insertion possible d'un photon virtuel d'où les deux leptons de l'état final peuvent émerger.

Étude phénoménologie de $D^+ \to \pi^+\ell^+\ell^-$

La prédiction de désintégrations CSCN se fait dans le cadre des théories de champs effectives. Dans ce cadre, l'Hamiltonien $\mathcal{H}$ décrivant la transition est réécrit comme une somme de coefficients de Wilson C_i, encodant la physique à haute énergie, et d'opérateurs $\mathcal{O}_i$, encodant la physique à basse énergie (souvent non perturbative et donc plus difficile à calculer). Schématiquement, l'Hamiltonien peut être réécrit comme :

$$\mathcal{H} \to \mathcal{H}_{\text{effectif}} \propto \sum_i C_i \mathcal{O}_i \,. \tag{1}$$

Cette paramétrisation a deux grands avantages : elle simplifie grandement les calculs et elle permet de paramétriser facilement la NP. En effet cette dernière peut se manifester par la modification d'un des coefficients de Wilson ou par l'apparition de nouveaux opérateurs non présents dans le MS. Pour les désintégrations CSCN, seul quatre nouveaux opérateurs sont nécessaires pour paramétriser la NP : les opérateurs scalaires ($\mathcal{O}_S$ et $\mathcal{O}_P$) et les opérateurs tensoriels ($\mathcal{O}_T$ et $\mathcal{O}_{T5}$).

La prédiction de désintégrations CSCN semileptoniques inclut les contributions représentées par les diagrammes de la figure 3. Le carré noir symbolise l'opérateur effectif $\mathcal{O}_i$ et la croix cerclée marque l'insertion possible d'un photon virtuel d'où les deux leptons de l'état final peuvent émerger. Le quark spectateur n'est pas représenté dans les diagrammes (a) et (b). Les prédictions pour les mésons D sont plus difficiles que les prédictions pour les mésons B, et ce, pour deux raisons principales :

1. Ces dernières sont encore plus défavorisées que les désintégrations de mésons B dans le MS[4]. En conséquence, les corrections liées aux échanges de gluons et à l'annihilation ne peuvent plus être négligées et doivent être calculées. Le diagramme de l'annihilation est montré en figure 3 (c). Ces corrections sont calculées dans le cadre de la factorisation QCD (QCDf). Le calcul des corrections QCDf a été présenté pour la première fois pour les désintégrations de mésons B dans [35] et a été adapté aux désintégrations de mésons D dans [36].

2. Lorsque la masse invariante des deux leptons (son carré sera noté s dans la suite) correspond à la masse d'une résonance (i.e. : ρ^0, ω^0 et ϕ), la boucle de quark, présent dans le diagramme (b) de la figure 3, devient non perturbative

[4]Ceci est dû au mécanisme de GIM [34] et à l'absence d'un quark down aussi massif que le quark top.

et donc difficile à calculer. La stratégie employée pour les désintégrations de mésons B est d'ignorer l'espace des phases où les résonances jouent un rôle important. Cette stratégie ne peut être employée dans le cas des mésons D car les résonances dominent l'intégralité de l'espace des phases. Un soin particulier doit donc être apporté à leurs descriptions.

Le travail présenté dans cette thèse se démarque des références précédentes sur la désintégration $D^+ \to \pi^+\ell^+\ell^-$, les plus récentes étant [37,38], par la réalisation d'une étude phénoménologique incluant toutes les corrections QCDf et par l'utilisation d'une description alternative des résonances.

Sur la description des résonances

Sans résonances, la boucle de quark décrite dans le diagramme (b) de la figure 3 est décrite par la fonction perturbative (pt) $\tilde{h}^{(\mathrm{pt})}(s, m_q)$. Dans les premières références sur $D^+ \to \pi^+\ell^+\ell^-$, les résonances étaient modélisées par une fonction de Breit-Wigner, ajoutée "à la main" au dessus du résultat perturbatif $\tilde{h}^{(\mathrm{pt})}(s)$. Dans [36], une approche différente est utilisée : la partie imaginaire de $\tilde{h}$ est modélisée suivant une suggestion par M. Shifman [39]. Dans ce modèle, la résonance principale est modélisée par une fonction de Breit-Wigner "améliorée" plus une somme infinie de résonances, démarrant à partir de la première excitation et dont les masses suivent des trajectoires de Regge; cette somme étant calculée analytiquement. La fonction $\tilde{h}$ est ensuite reconstruite à partir de sa partie imaginaire via une relation de dispersion :

$$\tilde{h}_q(s) \to \tilde{h}^{(\mathrm{pt})}(-s_0, m_q) + \frac{1}{\pi}\int_0^\infty ds' \frac{s_0+s}{s_0+s'}\frac{\mathcal{I}m\tilde{h}_q(s')}{s'-s-i\epsilon}. \tag{2}$$

Dans cette approche, le résultat perturbatif est retrouvé asymptotiquement. Nous suivons leur approche pour la description de $\tilde{h}$ à trois modifications majeures près :

1. Dans [36], les contributions d'isospin $I = 0$ (ω^0) et $I = 1$ (ρ^0) sont décrites par une seule fonction. Nous les modélisons par deux fonctions différentes afin d'affiner leurs descriptions.

2. Nous avons introduit une phase relative entre les résonances afin d'observer l'influence de cette phase dans notre étude phénoménologique.

3. Les auteurs de [36] fixent les nombreux paramètres entrant dans les fonctions $\mathcal{I}m\tilde{h}$ sur des considérations théoriques, nous avons choisi d'adopter une stratégie inspirée de [40] et d'extraire la valeur de ces paramètres à partir de données expérimentales.

Ce dernier point repose sur le théorème optique; ce théorème permet de relier l'observable expérimental $R(s)$ à la partie imaginaire de $\tilde{h}$ pour chaque résonance :

$$R(s) \simeq \frac{\sigma_{e^+e^-\to hadrons(\gamma)}}{\sigma_{e^+e^-\to\mu^+\mu^-(\gamma)}} \propto \left(\frac{1}{2}\frac{\mathcal{I}m\tilde{h}_{I=1}(s)}{\pi} + \frac{1}{18}\frac{\mathcal{I}m\tilde{h}_{I=0}(s)}{\pi} + \frac{1}{9}\frac{\mathcal{I}m\tilde{h}_s(s)}{\pi}\right). \tag{3}$$

La figure 4 compare les données expérimentales de $R(s)$ (en violet), le modèle de [36] (en vert) et notre modèle (en rouge). Le bon accord entre notre modèle et les données expérimentales contraste avec le désaccord entre les données et le modèle de [36], ce qui illustre le bien-fondé de notre démarche. Armés de cette description des résonances, nous avons mené l'étude phénoménologique résumée dans la suite.

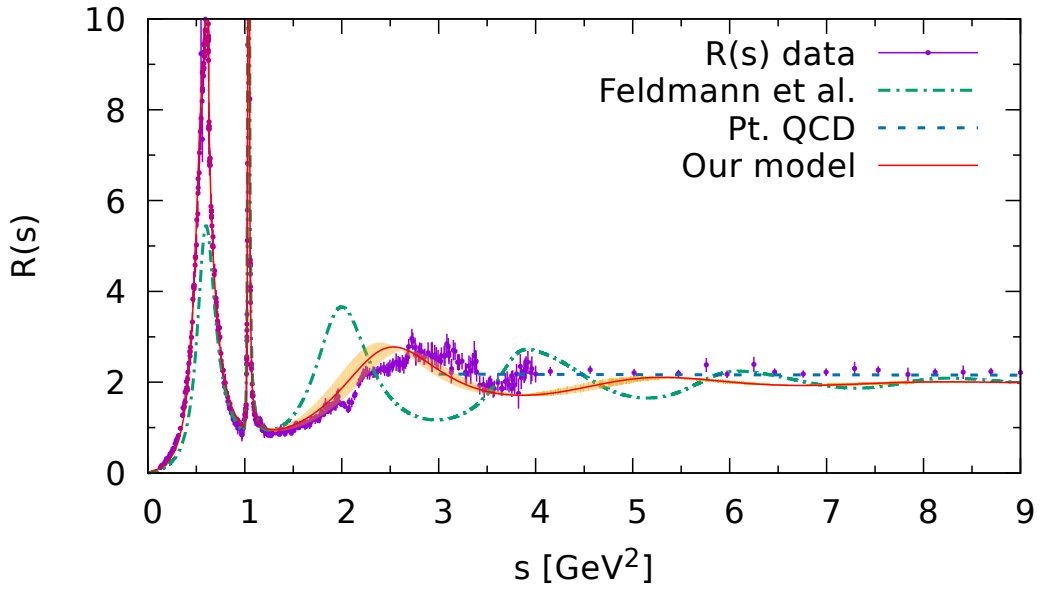


Figure 4: Les données expérimentales de $R(s)$ (en violet) comparées à notre modèle (en rouge) et au modèle de [36] (en vert). Le résultat perturbatif (sans résonances) est montré en bleu.

Une étude phénoménologique

L'ingrédient principal d'une étude phénoménologique est la distribution du taux de transition $\Gamma(D^+ \to \pi^+\ell^+\ell^-)$:

$$\frac{d^2\Gamma(D^+ \to \pi^+\ell^+\ell^-)}{ds\ d\cos\theta} \propto [a_\ell(s) + b_\ell(s)\cos\theta + c_\ell(s)\cos^2\theta], \tag{4}$$

où θ est l'angle entre l'impulsion du D^+ et celle de ℓ^- dans le référentiel où la paire de leptons est au repos. L'avantage de cette paramétrisation est que la distribution peut être écrite comme un simple polynôme en $\cos\theta$ de degré 2, les coefficients de ce polynôme a_ℓ, b_ℓ et c_ℓ dépendant uniquement de la masse invariante au carré des deux leptons (s). A partir de cette distribution, plusieurs observables peuvent être construites, une des plus évidentes est la distribution en s du rapport d'embranchement.

Il est intéressant d'analyser la sensibilité de cette observable à la NP, pour cela nous avons mené une analyse indépendant de tout modèle de NP. Ce type d'analyse repose sur la paramétrisation en terme de coefficients de Wilson C_i de notre taux de transition. Nous avons utilisé les limites expérimentales sur $D^0 \to \mu^+\mu^-$ [41] et $D^+ \to \pi^+\mu^+\mu^-$ [42] ainsi que notre prédiction pour extraire des valeurs maximales sur les coefficients de Wilson de NP. Ces valeurs maximales sont visibles dans le tableau 1. Il est pertinent de noter que certain modèle de NP comme le modèle de leptoquarks présenté en [38] peuvent produire des valeurs similaires à celles reportées dans le tableau 1.

La distribution du rapport d'embranchement est visible en figure 5 pour le MS et divers scénarios de NP. La prédiction MS est en noir et sa bande d'incertitude en jaune. Les zones en gris représentent les régions exclues par LHCb [42]. Nous reprenons la terminologie de [42] et appelons la région à bas s la région I (Reg. I)

NP C_i	Valeur maximale
$C_7^{\rm NP}$	1.03
$C_9^{\rm NP}$	1.3
C_{10}	0.087
C_S	0.86
C_P	0.86
C_T	0.84
C_{T5}	0.90

Table 1: Résumé des valeurs maximales autorisées par les limites expérimentales sur $D^0 \to \ell^+\ell^-$ [41] et $D^+ \to \pi^+\ell^+\ell^-$ [42] à 90 % C.L. si nous supposons les coefficients réels.

et la région à haut s la région II (Reg. II). Pour des raisons illustratives, nous montrons des scénarios de NP où les coefficients de Wilson varient individuellement et prennent les valeurs maximales reportées dans le tableau 1. Nous montrons l'effet sur le rapport d'embranchement de varier individuellement C_T (vert), C_P (rouge), C_{10} (violet) et C_9 (rose).

Le premier pic dans les distributions (autour de $s \sim 0.77$ GeV^2) est dû aux résonances ρ^0 et ω^0, le deuxième pic (autour de $s \sim 1$ GeV^2) est dû à la résonance ϕ. Comme nous pouvons le voir, la limite dans la Reg. I est très proche de notre prédiction pour le MS, elle est même contenue dans notre bande d'incertitude. De surcroît, tous les scénarios de NP sont compris dans cette bande. La Reg. II est plus intéressante car les bandes d'incertitude dans ces régions sont plus fines et la plupart des scénarios de NP pourraient être distingués du MS (dépendant de l'incertitude expérimentale).

La distribution du rapport d'embranchement est très intéressante mais d'autres observables peuvent être construites. Dans le MS, le coefficient angulaire b_ℓ est égal à zéro et $a_\ell \simeq -c_\ell$. Il est donc pertinent de construire un observable sensible à b_ℓ, il s'agit de l'asymétrie "Forward-Backward" $A_{\rm FB}$ ainsi qu'un observable sensible à la combinaison $(a_\ell + c_\ell)$, il s'agit du "flat term" F_H.

La distribution du flat term F_H est montrée en figure 6 pour le MS (en tiret noir) et pour différents scénarios de NP. Comme pour le rapport d'embranchement, nous montrons les effets de varier individuellement C_T (vert), C_{T5} (bleu), C_{10} (violet) et C_9 (rouge) et nous supposons qu'ils atteignent les valeurs maximales reportées dans le tableau 1. Dans le MS, F_H est juste une fonction cinématique et la bande d'incertitude est très petite, elle n'est donc pas montrée dans la figure 6. En revanche, l'incertitude pour le scénario de NP ou C_P est non nul est tracée en rouge. Toutes les prédictions de NP présentent deux pics inversés dû aux résonances. Il s'avère que cette observable est intéressante dans les régions à haut s, où la prédiction du MS est plus proche de zéro et où l'effet des résonances s'atténuent. Nous recommandons donc la mesure du flat term F_H dans la région au delà de la résonance ϕ. Une version intégrée de cette observable dans la région $s > 1.2$ GeV^2 serait une observable pertinente :

$$F_H^{\rm int} = \int_{s>1.2\ {\rm GeV}^2} ds\ F_H(s). \tag{5}$$

Contrairement à F_H, l'asymétrie "Forward-Backward" $A_{\rm FB}$ est sensible uniquement

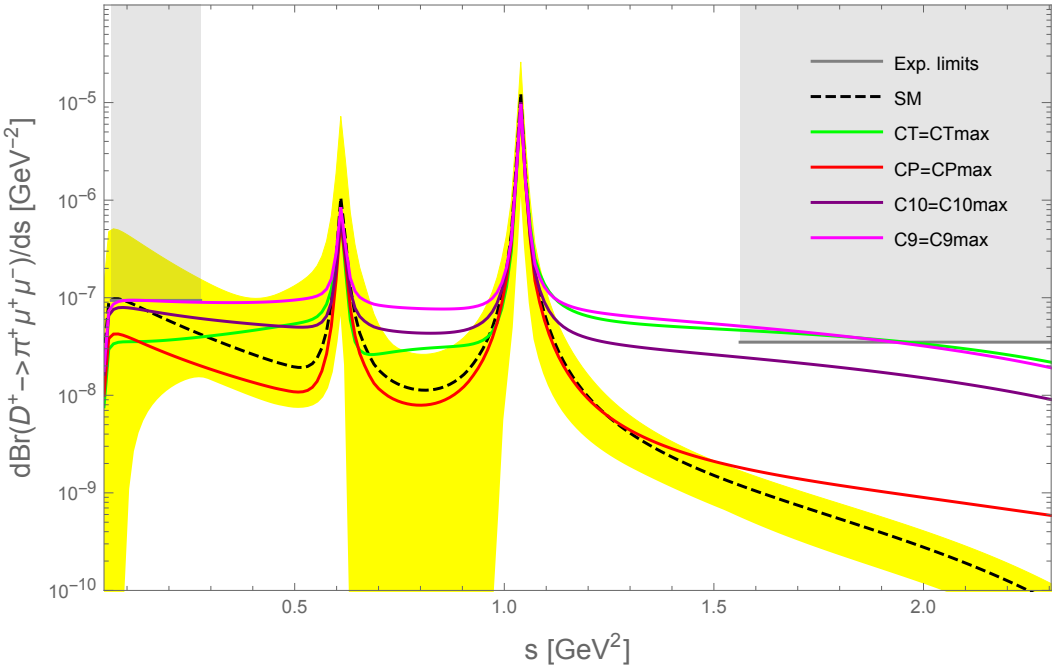


Figure 5: Distribution en s du rapport d'embranchement de $D^+ \to \pi^+\mu^+\mu^-$ pour le MS et différents scénarios de NP. La prédiction pour le MS est en tiré noir, la bande d'incertitude associée est en jaune et les régions exclues expérimentalement par [42] sont en grises. Pour des raisons illustratives, les coefficients de Wilson de NP sont variés individuellement et sont supposés prendre les valeurs maximales reportées dans le tableau 1. Nous montrons l'effet sur le rapport d'embranchement de varier individuellement C_T (vert), C_P (rouge), C_{10} (violet) et C_9 (rose).

à des combinaisons de coefficients de Wilson, soit C_P et C_{T5}, soit C_S et C_T. La distribution de $A_{\rm FB}$ est montrée en figure 7. La prédiction pour le MS est égale à zéro et n'est donc pas montrée. Les scénarios de NP où C_P et C_{T5} atteignent leurs valeurs maximales est montrée en noir (incertitude en rose). Il est notable que même pour des petites valeurs de C_{T5}, divisée par 10 (vert) et par 100 (rouge), $A_{\rm FB}$ est toujours significativement différent de zéro. Comme pour F_H, l'effet des résonances nous invitent à recommander des mesures expérimentales dans la région $s > 1.2$ GeV^2.

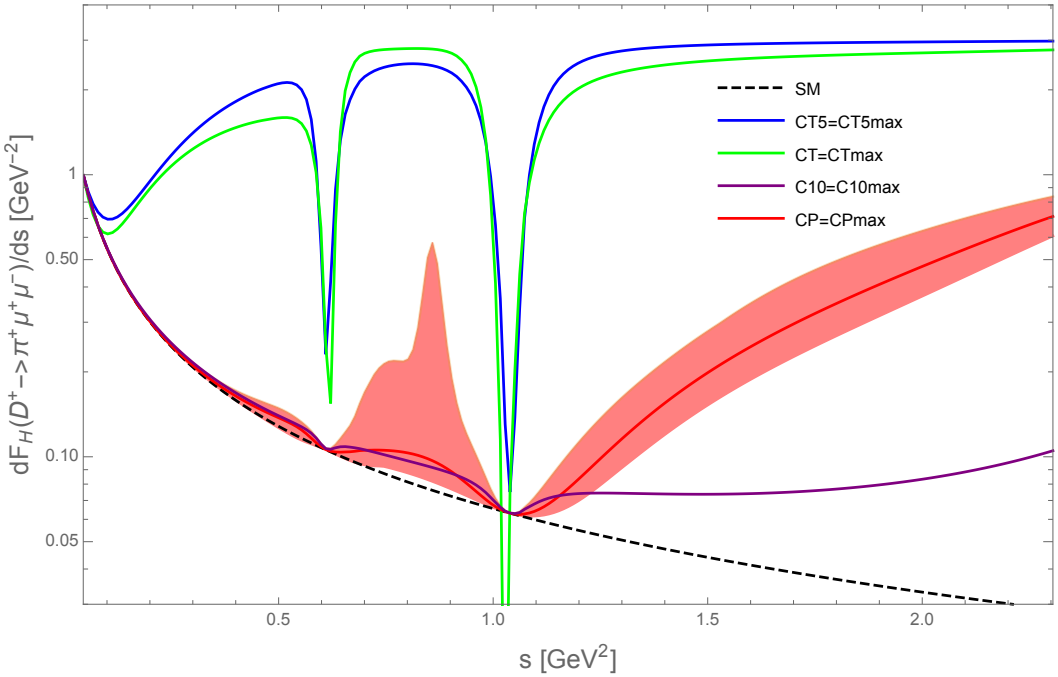


Figure 6: Distribution du flat term F_H pour le MS (en tiré noir) et pour différents scénarios de NP. Pour des raisons illustratives, les coefficients de Wilson de NP sont variés individuellement et sont supposés prendre les valeurs maximales reportées dans le tableau 1. Nous montrons l'effet sur le rapport d'embranchement de varier individuellement C_T (vert), C_{T5} (bleu), C_{10} (violet), et C_P (rouge). La bande d'incertitude pour ce dernier scénario est en rouge.

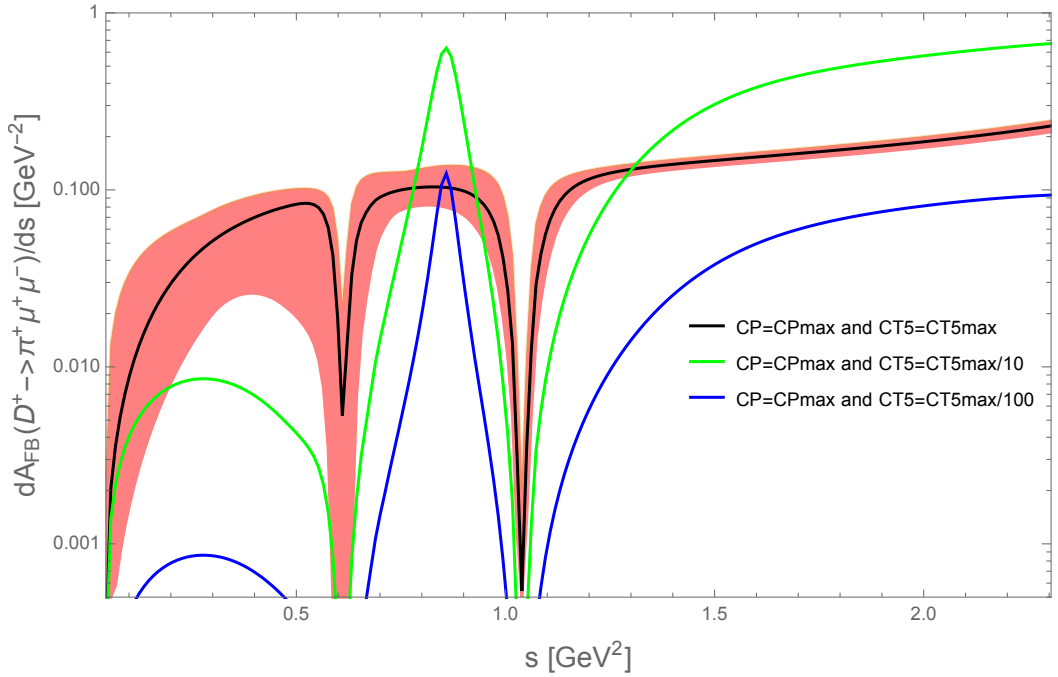


Figure 7: Distribution de l'asymétrie Forward-Backward $A_{\rm FB}$ pour différents scénarios de NP où C_P et C_{T5} sont simultanément non nuls. Sont montrés un scénario où C_P et C_{T5} atteignent les valeurs maximales reportées dans le tableau 1 (valeur centrale en noir, incertitude en rose) et deux scénarios où la valeur de C_{T5} est divisée par 10 (vert) et par 100 (bleu).

Conclusions sur l'étude phénoménologique de $D^+ \to \pi^+\ell^+\ell^-$

Nous avons trouvé que la contribution due à l'annihilation joue un rôle majeur dans les régions à bas s et ne doit donc pas être négligée. Cela a comme conséquence importante de "rehausser" par rapport aux prédictions précédentes [37,38], le rapport d'embranchement pour le MS dans la région à bas s. Notre prédiction pour le MS se trouve désormais très proche de la limite expérimentale actuelle. Nous avons également montré que la NP est difficilement distinguable du MS dans cette région et il parait plus judicieux de se concentrer vers la région à haut s où les scénarios de NP sont potentiellement distinguable du MS (dépendant de la sensibilité expérimentale). Un constat similaire a été dressé pour les observables F_H et $A_{\rm FB}$ où les résonances peuvent potentiellement masquer des effets de NP. En somme, les trois observables : rapport d'embranchement, F_H et $A_{\rm FB}$ sont de bon tests du MS quand on se place au delà de la résonance ϕ, soit à $s > 1.2$ GeV^2.

Recherche de $B^0_{(s)} \to \tau^+\tau^-$ avec le détecteur LHCb

Le détecteur LHCb [43] est une expérience dédiée à la physique des mésons D et B. Elle est située au CERN auprès du grand collisionneur de proton (LHC) [44]. Une vue schématique du détecteur est visible en figure 8. Au niveau du "vertex locator"[5], les protons collisionnent à une fréquence de 40 MHz, produisant à chaque collision une pléthore de particules, notamment des mésons B. Le détecteur se compose de sous-systèmes visant à remplir trois fonctions principales : la reconstruction des traces et des vertex (cerclé en rouge sur la figure 8), l'identification des particules (cerclé en bleu sur la figure 8) et enfin la sélection des événements intéressants par le trigger[6], permettant de réduire le flot d'événements produit par les collisions pp du LHC de 40 MHz à quelques dizaines de kHz. L'acquisition de données au LHC se décomposent en période appelée *Run*. Les données analysées dans cette thèse sont issues du Run1 et du Run2, plus d'informations à propos de ces Runs sont données dans le tableau 2. La luminosité intégrée ($\mathcal{L}$) est proportionnelle au nombre de mésons B produits.

Les mésons B, produits lors de la collision entre deux protons, volent sur environ 1cm avant de se désintégrer en deux leptons τ; eux-même volent sur environ 1mm avant de se désintégrer. Il faut donc choisir dans quel état final les leptons τ vont être reconstruits. En 2016, la collaboration LHCb a analysé les données du Run1 via l'état final $B^0_{(s)} \to \tau^+(\to 3\pi^\pm\bar{\nu_\tau})\tau^-(\to 3\pi^\mp\nu_\tau)$ [45], cette analyse a mené à la première limite expérimentale sur $B^0_s \to \tau^+\tau^-$ et la meilleure limite mondiale sur $B^0 \to \tau^+\tau^-$. La limite sur B^0_s est obtenue en supposant aucune contribution du B^0 et vice versa :

$$\begin{aligned} \mathcal{B}(B^0_s \to \tau^+\tau^-) &< 5.2\ (6.8) \times 10^{-3} \quad \text{à } 90\ (95)\%\ \text{C.L.}\ , \\ \mathcal{B}(B^0 \to \tau^+\tau^-) &< 1.6\ (2.1) \times 10^{-3} \quad \text{à } 90\ (95)\%\ \text{C.L.}\ . \end{aligned} \tag{6}$$

Deux axes ont été explorés dans cette thèse afin d'améliorer ce résultat :

1. Analyser les données du Run2 avec le même état final.

[5] Un vertex est un point reconstruit à partir de l'intersection supposée de traces.

[6] Le trigger est un terme anglais, il se réfère au système qui déclenche l'enregistrement d'un événement.

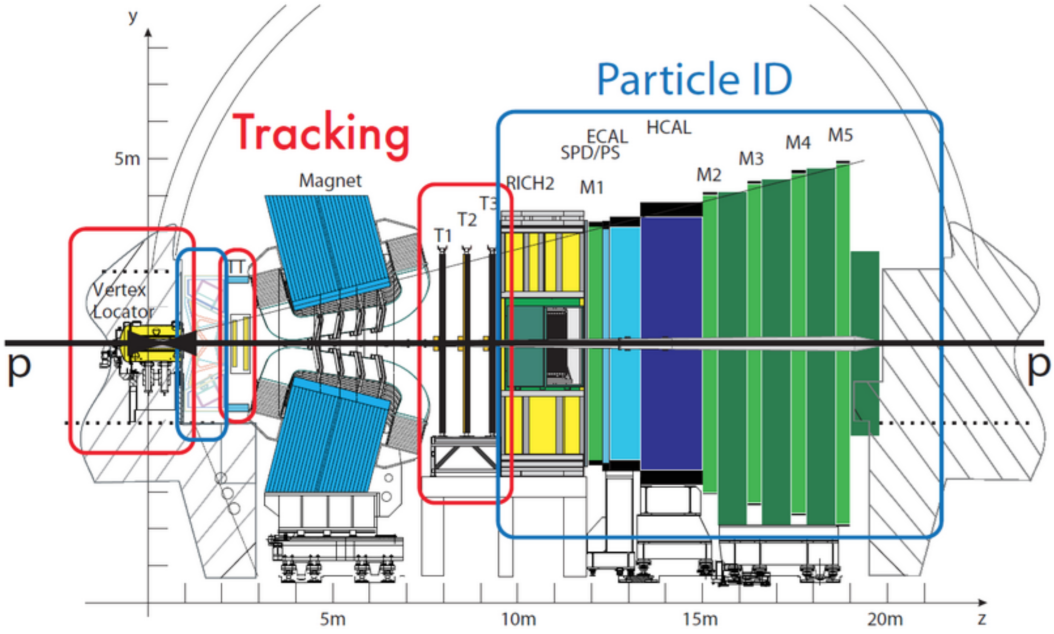


Figure 8: Vue schématique du détecteur LHCb ; le détecteur se compose de sous-systèmes visant à remplir trois fonctions principales : la reconstruction des traces et des vertex (cerclé en rouge), l'identification des particules (cerclé en bleu) et enfin la sélection des événements intéressants par le trigger.

	Run1	Run2
Période	2010-2012	2015-2018
Énergie du centre de masse (TeV)	7 (8 in 2012)	13
Section efficace au sein du détecteur (μb)	$\sim$ 72 [46]	$\sim$ 144 [46]
$\mathcal{L}$ enregistrée (fb^{-1})	3.1	5.9

Table 2: Quelques caractéristiques du Run1 et du Run2.

2. Analyser un autre état final; nous avons choisi d'analyser l'état final $B^0_{(s)} \to \tau^+(\to 3\pi^\pm\bar{\nu_\tau})\tau^-(\to \mu^-\bar{\mu_\nu}\nu_\tau)$.

Dans la suite, les états finaux $B^0_{(s)} \to \tau^+(\to 3\pi^\pm\bar{\nu_\tau})\tau^-(\to 3\pi^\mp\nu_\tau)$ et $B^0_{(s)} \to \tau^+(\to 3\pi^\pm\bar{\nu_\tau})\tau^-(\to \mu^-\bar{\mu_\nu}\nu_\tau)$ seront respectivement notés $(3\pi, 3\pi)$ et $(3\pi, \mu)$. L'état final $(3\pi, \mu)$ présente les avantages suivants par rapport à l'état final $(3\pi, 3\pi)$:

- Seul quatre traces sont requises au sein du détecteur (un muon et trois pions) alors que l'état final $(3\pi, 3\pi)$ requiert six traces de pions.
- Le détecteur LHCb est doté d'un excellent trigger à muon ce qui augmente l'efficacité de sélection de cet état final.
- La probabilité d'un tel événement est plus grande car $\mathcal{B}(\tau^- \to \mu^-\bar{\nu_\mu}\nu_\tau) \sim 2\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau)$.

Cependant certains désavantages sont également prévisibles :

- Nous disposons que d'un vertex de τ et donc moins de contraintes pour discriminer le bruit de fond.

- Nous aurons à discriminer le signal d'un bruit de fond abondant issu des désintégrations de mésons B en mésons D, e.g. : $B^0 \to D(3\pi\pi^0), \mu\nu$ ou $B^0 \to D(K^0\mu\nu), 3\pi$.

Une étude a été menée lors de cette thèse afin de déterminer le potentiel de cet état final; elles est résumée dans la suite.

L'état final $(3\pi, \mu)$

Nous avons choisi d'analyser les données du Run1 pour cet état final. La stratégie pour analyser ces données peut se décomposer en quatre étapes principales :

1. Une sélection visant à retirer le plus de bruit de fond possible tout en gardant une haute efficacité de signal. Cette dernière se décompose elle même en plusieurs étapes : une première sélection imposée par le trigger de l'expérience et la méthode de reconstruction, une sélection lâche basée sur des coupures, une coupure sur la sortie d'un algorithme d'analyse multivariée[7] (MVA1), et pour finir, la coupure sur la sortie d'un deuxième MVA (MVAH) afin d'affiner la réjection des désintégrations hadroniques de méson B.

2. La mesure du nombre d'événements de signal N_{sig} via l'ajustement d'un modèle sur les données; on utilisera l'anglicisme "fit" dans la suite. Dans beaucoup d'analyse en physique du B, la variable fittée est la masse invariante du méson B reconstruit. Dans le cas de $B^0_{(s)} \to \tau^+\tau^-$, cette masse invariante est très peu discriminante dû à la présence de neutrinos dans l'état final; il a donc été choisi de fitter la sortie d'un troisième MVA (MVA2).

3. La conversion du nombre d'événements fitté N_{sig} en une mesure de rapport d'embranchement se fait via un canal normalisation. Le canal de normalisation doit avoir de préférence un rapport d'embranchement large, mesuré précisément et un état final similaire avec le signal recherché, ce qui permet de réduire certaines incertitudes. Le canal choisi pour l'état final $(3\pi, \mu)$ est le canal : $B^0 \to D^-(\to \pi^-K^+\pi^-)\pi^+$. Au final, le rapport d'embranchement peut être réécrit comme :

$$\mathcal{B}(B^0_{(s)} \to \tau^+\tau^-) = \alpha_{(d/s)} N_{sig}, \tag{7}$$

   où $\alpha_{(d/s)}$ sont les facteurs de normalisation.

4. Si le nombre d'événements de signal obtenu est compatible avec zéro, le calcul d'une limite supérieure sur le rapport d'embranchement est effectué.

Des simulations Monte Carlo (MC) sont disponibles pour le signal et le canal de normalisation. Ces simulations sont utilisées afin d'optimiser les étapes de sélection, la variable fittée (MVA2) et modéliser le signal lors des étapes de fit mais aucune simulation de taille suffisamment grande n'est disponible pour modéliser le bruit de fond. L'idée pour contourner ce problème est de diviser les données en sous-régions. Une région dite de bruit de fond est utilisée afin de modéliser le bruit de fond lors de la sélection et l'optimisation du MVA2, une région de contrôle est utilisée comme "proxy" pour modéliser le bruit de fond lors du fit, et une région de signal sert comme

[7] Un algorithme Multivariée (MVA) exploitent les corrélations entre plusieurs variables $x_1, x_2, ..., x_D$, afin de produire une unique variable de sortie $\chi = \mathcal{F}(x_1, x_2, ..., x_D)$, la coupure sur χ étant plus discriminante que D coupures individuelles sur les variables x_i.

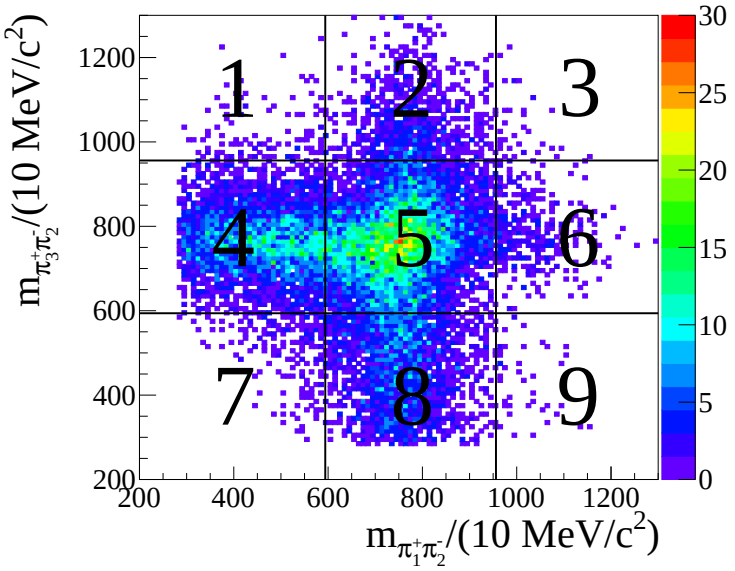


Figure 9: Illustration de la projection bi-dimensionnelle pour des événements de $B_s^0 \to \tau^+\tau^-$ simulés, utilisée dans cette analyse pour séparer les données en différente régions.

dernière étape de sélection. Pour créer ces régions, nous exploitons le fait que la désintégration $\tau^- \to 3\pi^\pm\nu_\tau$ procède majoritairement via les résonances $\rho^0(770)$ et $a_1(1260)$ [47] comme schématisé ci-dessous :

$$\begin{aligned}\tau^- \to & a_1(1260)^-\nu_\tau \\ & \hookrightarrow \pi_1^- \ \rho(770)^0 \\ & \qquad \hookrightarrow \pi_2^+\pi_3-\end{aligned}$$

Par conséquent, les événements de signal se répartissent suivant une forme de "plus" dans la projection bi-dimensionnelle ayant pour axes $m_{\pi_1^+\pi_2^-}$ et $m_{\pi_2^-\pi_3^+}$. Cette projection bi-dimensionnelle est divisée en neuf carrés comme illustré en figure 9. Dans l'analyse de l'état final $(3\pi, \mu)$, trois sous régions des données sont ainsi construites :

- région de signal : le candidat τ se trouve dans le carré 5,
- région de contrôle : le candidat τ se trouve dans le carré 4 ou 8,
- région de bruit de fond : le candidat τ se trouve dans le carré 1, 3, 7 ou 9.

Le choix des variables entrant dans les algorithmes MVA's est très important et peut être une tâche laborieuse, j'ai donc crée un algorithme de sélection basé sur une procédure itérative pour automatiser et optimiser ce choix. Cet algorithme a été utilisé pour sélectionner les variables entrant dans le MVA1, le MVAH et le MVA2. Les MVAs utilisés dans cette étude sont des arbres de décision boosté (BDT) et sont construits via le ToolKit TMVA [48]. Après la sélection, 13 715 événements de données sont présents dans la région de signal et l'efficacité totale de la sélection sur le signal est de $\epsilon_{B_s^0} \sim \epsilon_{B^0} = (1.42 \pm 0.06) \times 10^{-3}$ %. Comme expliqué précédemment, la région de contrôle est utilisée pour modéliser le bruit de fond dans la région de signal. Comme nous supposons que les régions de signal et de contrôle sont dominées par le bruit de fond, un bon accord de la distribution du MVA2 entre ces deux régions est attendu. La figure 10 montre la distribution de la sortie du MVA2 (BDT2) pour

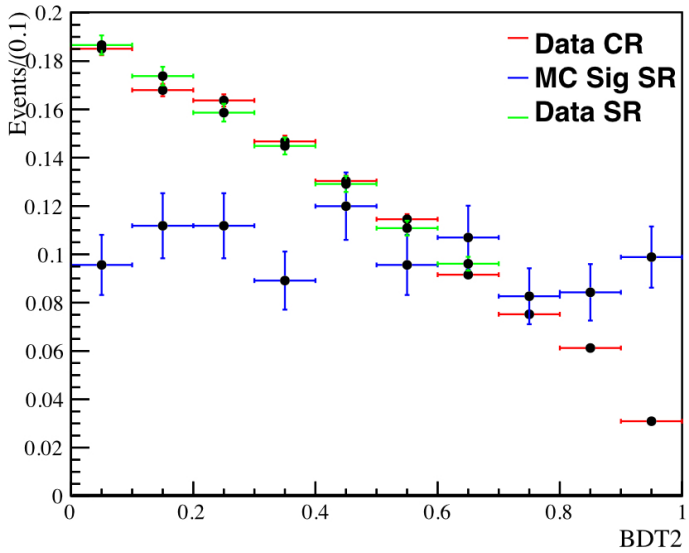


Figure 10: Distribution de la sortie du MVA2 (BDT2) pour les données dans la région de contrôle (en rouge), dans la région de signal (en vert) et pour le MC signal dans la région de signal (en bleu).

les données dans la région contrôle (en rouge), dans la région de signal (en vert) et pour le MC signal dans la région de signal (en bleu). Pour éviter d'éventuel biais, tous les choix de l'analyse sont faits sans regarder la région où se trouve le signal, on dit que l'analyse est faite à l'aveugle. Techniquement, la distribution du MVA2 dans la région de signal est masquée quand ces valeurs sont supérieures à 0.7, région la plus sensible à la présence de signal.

Le modèle du fit est donné ci-dessous :

$$\mathcal{N}^{\mathrm{SR}}_{\mathrm{data}} = N_{sig} \times \widehat{\mathcal{N}}^{\mathrm{SR}}_{\mathrm{sim}} + f_b \times \left(\mathcal{N}^{\mathrm{CR}}_{\mathrm{data}} - s \times \frac{\epsilon^{\mathrm{CR}}}{\epsilon^{\mathrm{SR}}} \times \widehat{\mathcal{N}}^{\mathrm{CR}}_{\mathrm{sim}} \right), \tag{8}$$

où $\mathcal{N}^{\mathrm{SR}}_{\mathrm{data}}$ ($\mathcal{N}^{\mathrm{CR}}_{\mathrm{sim/data}}$) sont les histogrammes de sortie du MVA2 dans la région de signal (contrôle) pour la simulation MC du signal/les données. Ces histogrammes sont visibles en figure 10. Le chapeau circonflexe sur le haut de la quantité $\mathcal{N}$ signifie que la distribution est normalisée à l'unité. Le facteur f_b factorisant le terme entre parenthèse est un facteur d'échelle. La soustraction dans le terme entre parenthèse permet de prendre en compte la présence d'éventuels événements de signal dans la région de contrôle, en effet la fraction de signal dans la région de contrôle (48.23 ± 1.16 %) est supérieure à la fraction de signal dans la région de signal (34.41 ± 1.10 %), ce potentiel signal doit donc être retiré du modèle pour le bruit de fond. Le fit utilise l'interface HistFactory de ROOT [49].

Le fit est validé via une méthode dite de simulation "Toys". Un "Toy" est une pseudo-expérience obtenue par ré-échantillonnage (BootStrapping [50]) des histogrammes présentés en figure 10. Cette méthode nous permet de simuler la sensibilité de notre fit à une variation statistique des données. La distribution de N_{sig} pour 500 toys est montrée en figure 11. Cette distribution est supposée être gaussienne, son écart-type

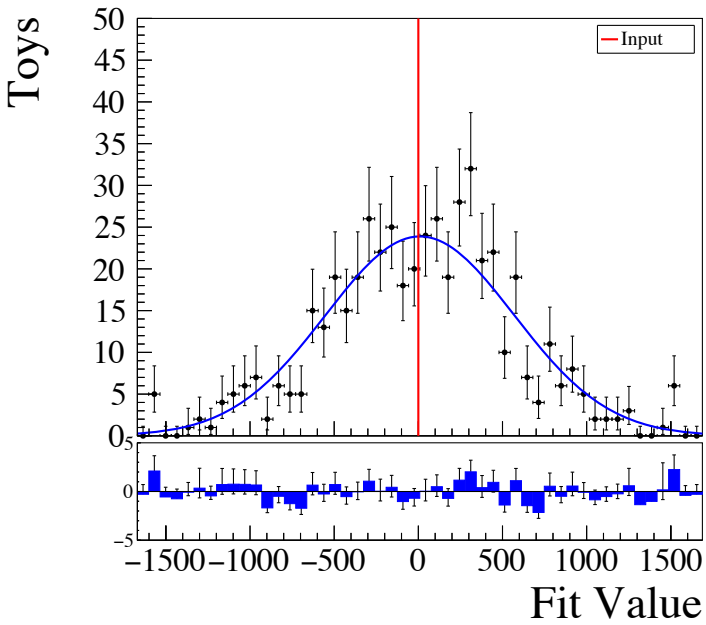


Figure 11: Distribution pour 500 toys de N_{sig}.

	B_s^0		B^0	
	$(3\pi, \mu)$	$(3\pi, 3\pi)$	$(3\pi, \mu)$	$(3\pi, 3\pi)$
$\epsilon^{\text{tot}} \times 10^5$	1.42	1.8	1.42	1.7
$\alpha \times 10^5$	3.5	4.7	0.9	1.3
σ^{toys}	444	58	490	61
$\mathcal{UL}$ at 95 % C.L.	2.4×10^{-2}	4.3×10^{-3}	7.0×10^{-3}	1.3×10^{-3}

Table 3: Efficacité totale sur le signal, facteur de normalisation, erreur statistique du fit, estimation $\mathcal{UL}$ sur la limite supérieur à 95 %C.L., pour B_s^0 et B^0, pour l'analyse du Run1 de l'état final $(3\pi, 3\pi)$ et de l'état final $(3\pi, \mu)$.

σ^{toys} nous indique la sensibilité du fit. Cette sensibilité peut être convertie en une estimation de la limite supérieure sur le rapport d'embranchement via :

$$\mathcal{UL} \sim 1.3(1.6) \times \alpha \times \sigma^{\text{toys}} \text{ at } 90(95) \text{ \% C.L.}, \tag{9}$$

où α est le facteur de normalisation. Le tableau 3 reporte l'efficacité totale de la sélection (non corrigée de potentielles erreurs dans la simulation MC), le facteur de normalisation, l'erreur sur le fit σ^{toys} ainsi que l'estimation de la limite supérieure $\mathcal{UL}$ à 95% C.L. pour B_s^0 et B^0. Les même grandeurs sont reportées pour l'analyse du Run1 dans l'état final $(3\pi, 3\pi)$ publiée dans [45].

L'estimation $\mathcal{UL}$ est environ cinq fois plus grand pour l'état final $(3\pi, \mu)$ que pour l'état final $(3\pi, 3\pi)$. Nous concluons que les avantages de l'état final $(3\pi, \mu)$ ne semblent pas compenser le désavantage principal, à savoir très peu de contraintes cinématiques et géométriques permettant de discriminer le signal d'un bruit de fond semileptonique très abondant. Cette étude a été reportée dans la note interne [CERN-LHCb-INT-2018-021] [51].

L'état final $(3\pi, 3\pi)$

Dans cette étude, seules les données 2016 et 2017 du Run2 ont été analysées, ce qui correspond environ à la même luminosité intégrée que pour le Run1. Comme la section efficace est environ multipliée par deux entre le Run1 et le Run2, on estime qu'environ deux fois plus de mésons B sont analysés dans cette étude; ce qui devrait abaisser la limite par un facteur $\sqrt{2}$. Notre analyse se veut la plus proche possible de celle réalisée pour le Run1 [45] afin d'obtenir une première idée du potentiel de l'analyse des données du Run2. La structure de l'analyse du Run1 [45] est la suivante :

1. Une sélection due au trigger et à la méthode de reconstruction, une sélection lâche basée sur des coupures puis la coupure sur la sortie d'un premier MVA (MVA1). Les MVA's utilisés dans cette analyse sont des réseaux artificiels de neurones (NN).

2. La mesure du nombre d'événements de signal N_{sig} via un fit sur la sortie d'un MVA (MVA2). Le même modèle de fit présenté dans l'équation (8) est utilisé.

3. La conversion du nombre d'événement fitté N_{sig} en une mesure sur le rapport d'embranchement. Le canal normalisation choisi pour l'état final $(3\pi, 3\pi)$ est : $B^0 \to D_s^+(K^+K^-\pi^+)D^-(\pi^+\pi^-K^-)$.

4. Le calcul de la limite supérieure (ou d'une estimation).

Cette fois, deux leptons τ se désintègrent en trois pions. Les régions sont donc définies différemment :

- région de signal : les deux candidats τ se trouve dans le carré 5,

- région de contrôle : un des candidats τ se trouve dans le carré 4 ou 8 tandis que l'autre se trouve dans le carré 4, 5 ou 8.

- région de bruit de fond : un des deux candidats τ se trouve dans le carré 1, 3, 7 or 9.

Nous appliquons la même sélection aux données du Run2 que celle appliqué aux données du Run1. Notamment le choix des variables entrant dans le MVA1 est le même mais l'algorithme a été réoptimisé sur les données du Run2. A la fin de la sélection, 53 876 événements de données se trouvent dans la région de signal et l'efficacité totale sur le signal est de $(31.32 \pm 0.85) \times 10^{-2}$ % pour 2016.

Comme pour le MVA1, le même ensemble de variables entrant dans le MVA2 (la variable fittée) est utilisé (tout du moins dans un premier temps) mais l'algorithme a été réoptimisé sur les données du Run2. La stratégie du fit est la même que celle présentée pour l'état final $(3\pi, \mu)$, il est donc important de vérifier l'accord entre la région de contrôle et la région de signal pour la distribution du MVA2 des données. Une figure tel que celle réalisée pour le canal $(3\pi, \mu)$ (montrée en figure 10) est montrée en figure 12. Il s'avère que les distributions sont très différentes dans la région où la sortie du NN2 est proche de zéro. Après analyse de chaque variables entrant dans le NN2, nous avons trouvé que ce désaccord est dû aux masses des τ. L'accord est presque parfait après les avoir écarter du NN2, comme visible en figure 13. Ces variables sont très corrélées avec les variables utilisées pour définir les

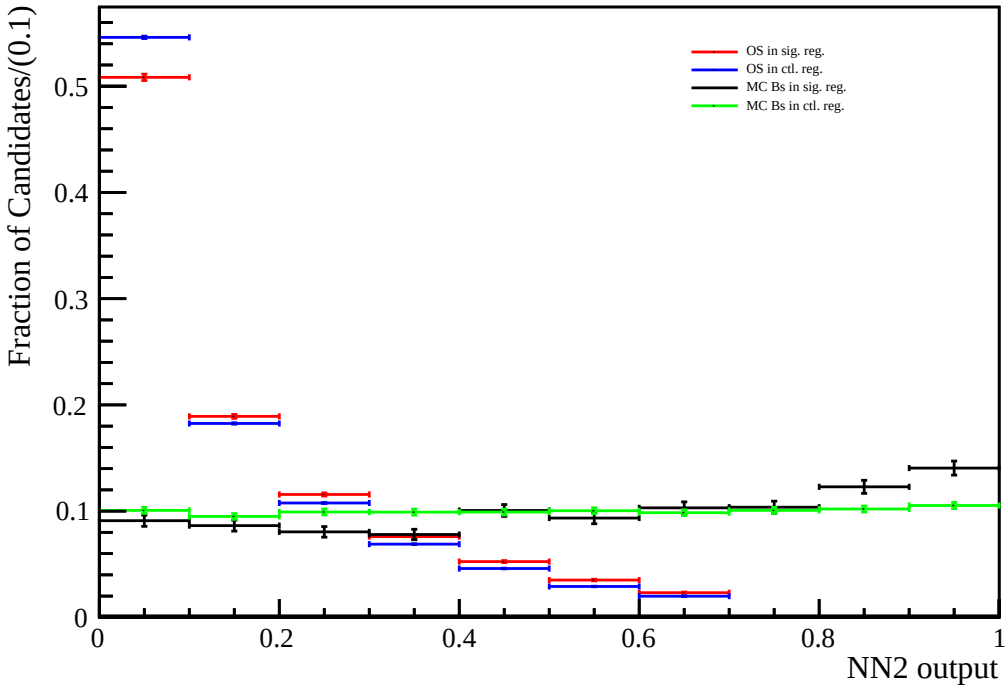


Figure 12: Distribution de la sortie du MVA2 (NN2) pour les données dans la région de contrôle (en rouge), dans la région de signal (en vert) et pour le MC signal dans la région de signal (en bleu). Le même ensemble de variables que dans l'analyse du Run1 est utilisé.

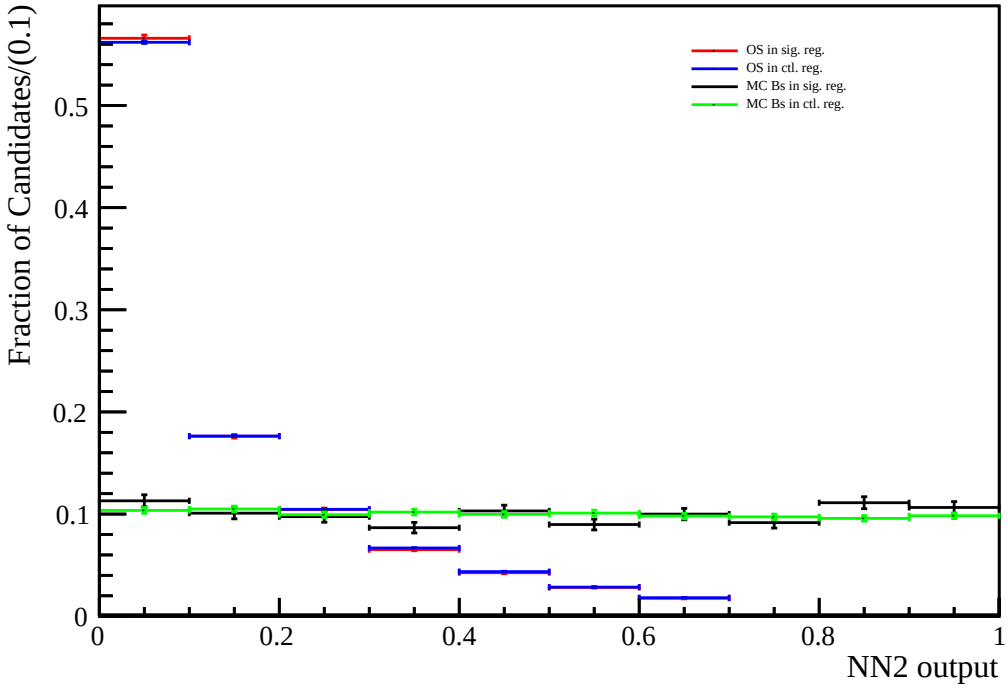


Figure 13: Distribution de la sortie du MVA2 (NN2) pour les données dans la région de contrôle (en rouge), dans la région de signal (en vert) et pour le MC signal dans la région de signal (en bleu). Le même ensemble de variables que dans l'analyse du Run1 est utilisé à l'exception des deux variables $m_{\tau}^{\pm}$.

	B_s^0		B^0	
	Run2	Run1	Run2	Run1
$\epsilon^{\text{tot}} \times 10^5$	3.7	1.8	3.3	1.7
$\alpha \times 10^5$	1.4	4.7	0.41	1.3
σ^{toys}	140	58	179	61
$\mathcal{UL}$ at 95 % C.L.	3.1×10^{-3}	4.4×10^{-3}	1.1×10^{-3}	1.3×10^{-3}

Table 4: Efficacité totale sur le signal, facteur de normalisation, erreur statistique du fit, estimation de $\mathcal{UL}$ à 95 %C.L., pour B^0 et B_s^0, pour l'analyse du Run1 de l'état final $(3\pi, 3\pi)$ et l'analyse du Run2 (2016+2017) du même état final.

régions ($m_{\pi_1^+\pi_2^-}$ et $m_{\pi_2^-\pi_3^+}$) et il n'est donc pas surprenant qu'elles soient responsables de cet effet. Leur suppression du MVA2 rend la variable fittée moins discriminante mais l'analyse y gagne en robustesse et en fiabilité.

Finalement, une validation du fit par simulation "Toys" a été réalisée aboutissant à une erreur statistique du fit σ^{toys}. Cette erreur, l'efficacité totale sur le signal, le facteur de normalisation et l'estimation $\mathcal{UL}$ sont visibles dans le tableau 4. Les estimations sur la limite supérieure sont plus petits d'un facteur d'environ 0.7 pour B_s^0 et 0.8 pour B^0 entre le Run2 (2016+2017) et le Run1, ce qui correspond environ au facteur attendu ($\sqrt{2}$).

Nous rappelons au lecteur que cette analyse n'a pas été optimisée pour les données du Run2, et un gain non négligeable sur la limite pourrait être obtenu en réoptimisant la procédure de sélection et la variable fittée (notamment le choix des variables entrant dans le MVA1 et le MVA2). De plus, les masses des τ ont été écartées du MVA2, et ce, en dépit de leur haut potentiel discriminant; il peut donc être intéressant d'utiliser ces variables plus intensivement en amont dans la sélection.

Conclusions sur la recherche de $B^0_{(s)} \to \tau^+\tau^-$ à LHCb

Une analyse de l'état final $(3\pi, \mu)$ a été réalisée pour la première fois. Nous avons montré que son potentiel était bien en dessous de celui de l'état final $(3\pi, 3\pi)$. L'analyse de l'ensemble des données du Run2 combinées aux données du Run1 via l'état final $(3\pi, 3\pi)$ devrait abaisser les limites actuelles par un facteur ~ 2.3 :

$$\mathcal{UL}(B_s^0 \to \tau^+\tau^-)^{\text{prospective}}_{(\text{Run1}-2)} \sim 3 \times 10^{-3} \text{ à } 95 \text{ \% C.L.}. \tag{10}$$

Cette amélioration sur la limite pourrait exclure certains modèles de NP dont la prédiction sur $B^0_{(s)} \to \tau^+\tau^-$ se trouve proche des limites expérimentales actuelles. Par exemple, dans [21], les auteurs présentent un modèle de vecteur leptoquark dont les paramètres sont fittés sur plusieurs anomalies expérimentales, $R_K^{(*)}$, $R_D^{(*)}$, ainsi que sur d'autre mesures. Leur modèle suggère que le rapport d'embranchement de $B_s^0 \to \tau^+\tau^-$ et de $B \to K\tau^+\tau^-$ pourraient être augmentés de plusieurs ordres de grandeur au dessus du MS. Ils illustrent leur propos avec la figure 14 (tirée de [21]) où est montrée la région de prédilection de leur fit pour les observables $\mathcal{B}(B_s^0 \to \tau^+\tau^-)$ et $\mathcal{B}(B \to K\tau^+\tau^-)$. La limite présentée dans l'équation (10) s'étend dans leur région $\Delta^2_\chi \leq 6.18(2\sigma)$ (en bleu dans la figure), réduisant ainsi l'espace des phases de leur modèle.

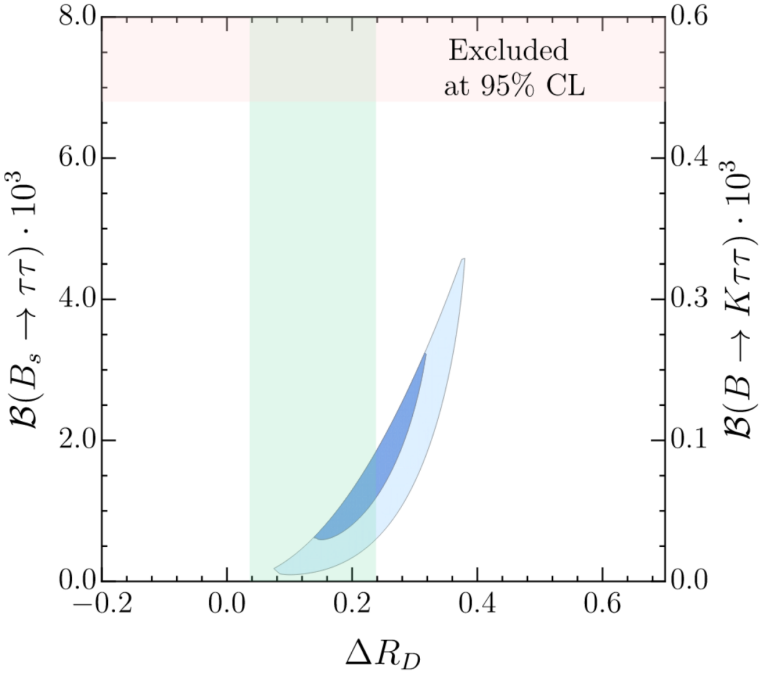


Figure 14: Régions du fit préférées d'un modèle de leptoquark vecteur [21] pour les observables $\mathcal{B}(B_s^0 \to \tau^+\tau^-)$ et $\mathcal{B}(B \to K\tau^+\tau^-)$. Les régions $\Delta_\chi^2 \leq 2.30(1\sigma)$ et $\Delta_\chi^2 \leq 6.18(2\sigma)$ sont montrées respectivement en bleu et en bleu clair. L'intervalle en vert montre la mesure actuelle de ΔR_D qui est l'un des observables entrant dans leur fit. La bande rouge montre la région exclue à 95% C.L. par la limite sur $B_s^0 \to \tau^+\tau^-$ [45]. La figure est tirée de [21].

Acronyms

ANN Artificial Neural Network. 102

BDT Boosted Decision Tree. 99, 102

BEH Brout-Englert-Higgs. 7

BSM Beyond the Standard Model. 5

BW Breit-Wigner. 41

CERN European Organization for Nuclear Research. 69

CKM Cabibbo-Kobayashi-Maskawa. 9

CP Charge-Parity. 9

DOCA Distance Of Closest Approach. 75

EFT Effective Field Theory. 7, 14

EWSB Electroweak Symmetry Breaking. 8

FCCC Flavor Changing Charged Current. 11

FCNC Flavor Changing Neutral Current. 5, 7, 11

FoM Figure of Merit. 104

GIM Glashow–Iliopoulos–Maiani. 12

Hlt High level trigger. 81

IP Impact Parameter. 75

LCDA Light Cone Distribution Amplitude. 23

LCSR Light Cone Sum Rules. 18

LFU Lepton Flavor Universality. 12

LHC Large Hadron Collider. 69

LL Leading Logarithmic. 11

LO Leading Order. 11

LQCD Lattice Quantum Chromodynamics. 18

LS Long Shutdown. 69

MC Monte Carlo. 83

MVA Multivariate Analysis. 87

NLL Next-to-Leading Logarithmic. 11

NLO Next-to-Leading Order. 11

OPE Operator Product Expansion. 14

PID Particle IDentification. 79

PV Primary Vertex. 72

QCD Quantum Chromodynamics. 10

QCDF Quantum Chromodynamics factorization. 23, 32

RGE Renormalization Group Equation. 11

ROC Receiver Operating Characteristic. 107

SM Standard Model. 5

TIS Triggered Independently of the Signal. 81

TOS Triggered on Signal. 82

VELO VErtex LOcator. 75

Introduction

The physics theory used to describe matter and its interactions at high energy is the so-called Standard Model (SM) of particle physics. Unfortunately, this theory is to some aspect incomplete and the particle physics community agrees nowadays to perceive the SM only as an effective low energy version of a more general theory to be discovered, referred to in the following as physics Beyond the Standard Model (BSM). One of the main portals towards these new dynamics are the heavy flavor meson decays happening through Flavor Changing Neutral Current (FCNC) as they proceed entirely through loop effects and are additionally suppressed by the chiral nature of the weak interactions. Thus their sensitivity to possible BSM virtual effects could be enhanced. $b \to s\ell^+\ell^-$ transitions provide some of the strongest constraints on BSM physics. Interestingly, a number of anomalies of various significances have shown up in the first LHCb data for some of these modes. These anomalies seem to suggest the possibility of BSM physics with a hierarchy of the generations in the coupling to BSM particles, where the τ leptons should couple most strongly. Less studied because more challenging from a theoretical point of view, the charmed transition $c \to u\ell^+\ell^-$ might also exhibit similar anomalies. This kind of searches, often qualified as indirect, relies on two ingredients: precision measurements and accurate predictions; this thesis is devoted to both aspects. At first, it focuses on how to improve the predictions for the less studied charmed decay $D^+ \to \pi^+\ell^+\ell^-$; then, it presents the experimental measurement of the $B^0_{(s)} \to \tau^+\tau^-$ decay.

Chap. 1 gives a brief review of key features of the SM and explains how the study of these decays can give access to BSM physics.
In contrast to the theory of B meson decay, the D meson decays are so far poorly understood, being dominated by resonances and long-distance hadronic contributions. The theoretical background needed to understand the long-distance hadronic contributions for the $D^+ \to \pi^+\ell^+\ell^-$ decay is introduced in Chap. 2. After presenting a possible route to improve the resonance's treatment, Chap. 3 presents the phenomenological results.
The LHCb detector and its environment are presented in Chap. 4. As explained in Chap. 5, that introduces the analyses, several final states are accessible for the $B^0_{(s)} \to \tau^+\tau^-$ decay. Chap. 6 gives an overview of the analysis performed with the $B^0_{(s)} \to \tau^+(\to 3\pi^\pm\bar{\nu_\tau})\tau^-(\to \mu^-\bar{\mu_\nu}\nu_\tau)$ final state. Chap. 7 focuses on the final state $B^0_{(s)} \to \tau^+(\to 3\pi^\pm\bar{\nu_\tau})\tau^-(\to 3\pi^\mp\nu_\tau)$.

Chapter 1

Probing physics beyond the Standard Model with FCNC meson decays

After introducing the key-features of the Standard Model relevant to this thesis in Sec. 1.1, this chapter explains in Sec. 1.2 why Flavor Changing Neutral Current (FCNC) decays offer an excellent opportunity to look for BSM physics. Then, the Effective Field Theory (EFT), one of the main tool used to set predictions is introduced in Sec. 1.3. Finally, the peculiarities of the $B^0_{(s)} \to \ell^+\ell^-$ and $D^+ \to \pi^+\ell^+\ell^-$ decays are introduced respectively in Sec. 1.4 and 1.5.

1.1 The Standard Model of particle physics

Two main features of the Standard Model are presented in the next two sections; i.e., its flavor structure related to the electroweak interaction and its color structure governed by the strong interaction responsible for the quark-gluon interaction[1].

1.1.1 On the flavor structure of the Standard Model

The Standard Model [1–8] is obtained by spontaneously breaking the $SU(3)_c \times SU(2)_L \times U(1)_Y$ non-abelian gauge theory into a $SU(3)_c \times U(1)_{EM}$ via the Brout-Englert-Higgs (BEH) mechanism [9]. It respects Lorentz symmetry and translation invariance. Its field content is given in Tab. 1.1. The subscript L/R refers to the Left/Right-chiral field and the superscript index $i = 1, 2$ or 3 to the fermion generations. When not needed, it will be suppressed for ease of reading. Q, u_R, d_R and L, e_R are the quark and lepton fields respectively. ϕ is the BEH field and g, A, B are the gauge fields.

A detailed version of the SM Lagrangian can be found in [52]. In the following we will mainly present its most relevant features for the thesis. It can be expressed as

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{Kin}} + \mathcal{L}_{\Phi} - \mathcal{L}^{q}_{\text{Yuk}} - \mathcal{L}^{\ell}_{\text{Yuk}}, \tag{1.1}$$

where $\mathcal{L}_{\text{Kin}}$ contains the massless kinetic and gauge terms for the electroweak and strong interactions, $\mathcal{L}_{\Phi}$ the BEH kinetic and potential terms, $\mathcal{L}^{q/\ell}_{\text{Yuk}}$ the Yukawa term

[1]The quark and gluon charges under the strong interaction are called color charges.

Field	Spin	Rep.
$Q^i = \begin{pmatrix} u_L^i \\ d_L^i \end{pmatrix}$	1/2	$(3,\ 2,\ \frac{1}{6})$
u_R^i	1/2	$(3,\ 1,\ \frac{2}{3})$
d_R^i	1/2	$(3,\ 1,\ -\frac{1}{3})$
$L^i = \begin{pmatrix} e_L^i \\ \nu_L^i \end{pmatrix}$	1/2	$(1,\ 2,\ -\frac{1}{2})$
e_R^i	1/2	(1, 1, -1)
$\phi = \begin{pmatrix} \varphi^+ \\ \varphi^0 \end{pmatrix}$	0	$(1,\ 2,\ \frac{1}{2})$
g	1	(8, 1, 0)
A	1	(1, 3, 0)
B	1	(1, 1, 0)

Table 1.1: Standard Model field content with their spins and their representations under the unbroken gauge group $SU(3)_c \times SU(2)_L \times U(1)_Y$.

for quark/lepton. This latter is the term relevant for flavor study and is given by

$$\mathcal{L}^q_{\text{Yuk}} = \lambda_D \bar{Q}\phi d_R + \lambda_U \bar{Q}\tilde{\phi} u_R, \tag{1.2}$$

where $\tilde{\phi} = i\sigma^2\phi$ is the $SU(2)_L$ conjugate of the BEH field and $\lambda_{D/U}$ are the flavor space matrices of Yukawa coupling for the down ($d_{L/R}$) and up-type ($u_{L/R}$) quarks respectively.

All terms of the SM Lagrangian respect a global flavor symmetry $SU(3)^5$ defined by rotating each field in flavor space, except the quark Yukawa terms. Indeed, the flavor space matrices are not diagonal resulting in a mixing of left- and right-chiral fields from different generations. The SM respects a chiral symmetry that prohibits the introduction of fermion mass terms in its Lagrangian. Masses are given to fermion via the Electroweak Symmetry Breaking (EWSB) mechanism. Upon this breaking, the BEH field settles to its minimum potential, called the vacuum expectation value v, and the Yukawa Lagrangian can be rewritten to make appear the following mass terms

$$\mathcal{L}^q_{\text{Yuk}} \to \mathcal{L}_{\text{mass}} = \bar{u_L} M_U u_R + \bar{d_L} M_D d_R + h.c., \tag{1.3}$$

where $M_{U/D} = (v/\sqrt{2})\lambda_{U/D}$ are the mass matrices. These mass matrices being not diagonal, quarks contained in doublets and singlets are not the physical states of theory. To find the physical states, one needs to diagonalize the mass matrices and hence the Yukawa matrices. This can be done via a bi-unitary transformation, $M \to \hat{M} = S^\dagger M T$ where the hat indicates diagonal matrices and S, T are two unitary matrices. Thus the mass Lagrangian becomes

$$\mathcal{L}_{\text{mass}} = \bar{u_L} S_U \hat{M}_U T_U^\dagger u_R + \bar{d_L} S_D \hat{M}_D T_D^\dagger d_R + h.c., \tag{1.4}$$

where $S_{U/D}$ and $T_{U/D}$ are the unitary matrices used to diagonalize the mass matrices. It is not possible to absorb both $S_{U/D}$ and $T_{U/D}$ into the definition of the chiral fields. By convention, we diagonalize the up-quark sector resulting in the following field

transformation

$$\begin{aligned} Q_L \to Q'_L &= S_U \; Q_L, \\ u_R \to u'_R &= T_U \; u_R, \\ d_R \to d'_R &= T_D \; d_R. \end{aligned} \tag{1.5}$$

Hence, the mass Lagrangian for the quark sector becomes

$$\mathcal{L}_{\text{mass}} = \bar{u_L}' \hat{M}_u u'_R + \bar{d_L}' \underbrace{S_U^\dagger S_D}_{V_{\text{CKM}}} \hat{M}_d d'_R + h.c., \tag{1.6}$$

where $V_{\text{CKM}} \equiv S_U^\dagger S_D$ is the so-called Cabibbo-Kobayashi-Maskawa (CKM) matrix [53, 54]. Note that the CKM matrix is unitary.

Upon electroweak symmetry breaking, two new gauge bosons are generated, the Z boson and the $W^\pm$ boson. This latter mix down- and up-type quarks from different generations via charged current interaction. The gauge vertex for the Z boson is sketched by the Feynman diagram in Fig. 1.1(a). Its vertex factor is given by

$$-i\frac{g_e}{2\cos\theta_W \sin\theta_W}\left[\left(\frac{1}{2} - \frac{4}{3}\sin^2\theta_W\right)\gamma^\mu - \frac{1}{2}\gamma^\mu\gamma^5\right], \tag{1.7}$$

in the case of the interaction $u_i \to Z u_i$, and

$$i\frac{g_e}{2\cos\theta_W \sin\theta_W}\left[\left(-\frac{1}{2} + \frac{2}{3}\sin^2\theta_W\right)\gamma^\mu - \frac{1}{2}\gamma^\mu\gamma^5\right], \tag{1.8}$$

in the case of the interaction $d_i \to Z d_i$. $\cos\theta_W$ is the Weinberg angle and g_e is the elementary electric charge. The gauge vertex for the $W^\pm$ boson is sketched by the Feynman diagram in Fig. 1.1(b). Its vertex factor is given by

$$-i\frac{g}{2\sqrt{2}\cos\theta_W} V_{ij}\gamma^\mu(1-\gamma^5), \tag{1.9}$$

in the case of the transition $u_i \to d_j W^+$ and

$$-i\frac{g}{2\sqrt{2}\cos\theta_W} V_{ji}^*\gamma^\mu(1-\gamma^5), \tag{1.10}$$

in the case of the transition $d_i \to u_j W^-$.

The quantity $|V_{ij}|^2$ involved in the W boson gauge vertices is the probability of transition of the quark i into the quark j via charged current. The CKM matrix can be parametrized in term of this V_{ij} and can be seen as the link between the down-type quark flavor states (d, s, b) and their physical eigenstates (d', s', b') via

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \underbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}_{V_{\text{CKM}}} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.11}$$

The unitarity of the CKM matrix results in six orthogonal relations which are often represented in the complex plane by the famous "CKM triangles". For example between the first and second row, we have

$$V_{cd}^* V_{ud}^* + V_{cs}^* V_{us}^* + V_{cb}^* V_{ub}^* = 0. \tag{1.12}$$

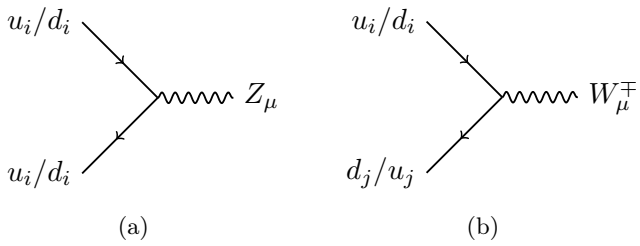


Figure 1.1: Feynman diagrams of the Z boson gauge vertices in (a) and the $W^\pm$ boson gauge vertices in (b). The index i, j refer to the fermion generations.

The CKM matrix can be described by 3 real angles and a Charge-Parity (CP) violating phase. Experimental evidence of this phase was first found in $K^0 - \bar{K}^0$ mixing [55] and later in $B^0 - \bar{B}^0$ mixing at BaBar [56] and Belle [57]. Following the experimental observation that b quarks decay predominantly to c, Wolfenstein suggested the following parametrization enhancing the relative importance of the different CKM elements [58],

$$V_{\text{CKM}} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \tag{1.13}$$

with $\lambda \sim 0.22$, $A \sim 0.81$, $\rho \sim 0.13$ and $\eta \sim 0.35$.

The theory of weak interaction governs the flavor structure of the SM from the free quark point of view. Nonetheless, this theory is not sufficient in itself as free quarks do not exist in nature since quarks are confined within hadronic states via the strong interaction. This interaction is described by the Quantum Chromodynamics (QCD) theory whose key features are detailed in the following.

1.1.2 On the color structure of the Standard Model

Quantum Chromodynamics is the theory of the strong interaction which confines quarks and gluons together within hadrons. This theory is invariant under the $SU(3)_c$ gauge group introduced above. QCD exhibits two main properties: color confinement and asymptotic freedom. Both are linked to the running of the strong coupling constant $\alpha_s(\mu)$ with the energy scale μ. Theoretically, this running is a direct consequence of a procedure called *renormalization*[2], introduced to remove divergences contained in the theory.

This divergence suppression is done in two steps, first the theory is regulated to allow manipulation of quantities, second the divergent terms are removed. We work in the following in naive dimension regularization [60], where the number of dimension is set to $4 - 2\epsilon$, ϵ being an arbitrarily small quantity. We use the bar minimal subtraction scheme $\overline{\text{MS}}$ [61], in which divergences along the associated $\ln 4\pi$ and the Euler-Mascheroni constant γ_E terms are simply subtracted. This implies multiplying bare quantities by a renormalization constant $Z(\mu)$, which leaves the renormalized quantities finite. This leads for the mass matrix m and the strong coupling constant

[2] A detailed overview on the renormalization technique can be seen in [59].

g_s to

$$m_0 = Z_m m \quad \text{and} \quad g_{s,0} = Z_g g_s \mu^\epsilon. \tag{1.14}$$

Knowing the bare quantities in Eq. (1.14) are scale independent [62], a simple derivation of Eq. (1.14) leads to the Renormalization Group Equation (RGE) which gives the scale dependence of the renormalized quantities:

$$\frac{dg_s}{d\ln\mu} = -\beta g_s \quad \text{and} \quad \frac{dm}{d\ln\mu} = -\gamma_m m, \tag{1.15}$$

where the β function and the anomalous dimension matrix γ_m are defined by

$$\beta = Z_g^{-1}\frac{dZ_g}{d\ln\mu} \quad \text{and} \quad \gamma_m = Z_m^{-1}\frac{dZ_m}{d\ln\mu}. \tag{1.16}$$

These latter can be expanded in perturbation theory in the strong coupling,

$$\beta = \beta_0 a_s + \beta_1 a_s^2 + ... \quad \text{and} \quad \gamma_m = \gamma_{m,0} a_s + \gamma_{m,1} a_s^2 + ... \tag{1.17}$$

with

$$\beta_0 = \frac{11N_c - 2n_f}{3}, \quad \beta_1 = \frac{34}{3}N_c^2 - \frac{10}{3}N_c n_f - 2C_F n_f \quad \text{and} \quad C_F = \frac{N_c^2 - 1}{2N_c} \tag{1.18}$$

where $a_s = \alpha_s/(4\pi) = g_s^2/(16\pi^2)$, n_f the number of active flavors in the theory and N_c the number of color charge. Note that we often use a_s instead of α_s or g_s in the following. Using Eq. (1.15), we obtain the running of a_s with respect to the scale μ:

$$a_s(\mu) = \frac{a_s(\mu_0)}{v(\mu)}\left(1 - a_s(\mu_0)\frac{\beta_1}{\beta_0}\frac{\ln v(\mu)}{v(\mu)}\right), \tag{1.19}$$

with

$$v(\mu) = 1 - \beta_0 a_s(\mu_0)\ln\frac{\mu_0^2}{\mu^2}. \tag{1.20}$$

Expanded in $a_s(\mu_0)$, the leading term of Eq. (1.19) becomes

$$a_s(\mu) = a_s(\mu_0)\sum_{m=0}^{\infty}\left(\beta_0 a_s(\mu_0)\ln\frac{\mu_0^2}{\mu^2}\right)^m. \tag{1.21}$$

This expansion exhibits the fact that Eq. (1.19) automatically sums the logarithms $\ln(\mu_0^2/\mu^2)$ which become large when $\mu \ll \mu_0$ and then spoil the convergence of the perturbation series. In general, solving the RGE to Leading Order (LO) in a_s, allows us in the Leading Logarithmic (LL) approximation to sum up the terms $(a_s(\ln(\mu_0/\mu))^n$ to all orders in n. In the Next-to-Leading Logarithmic (NLL) approximation, going to Next-to-Leading Order (NLO) in the RGE means that terms $a_s(a_s(\ln(\mu_0/\mu))^n$ are summed to all orders in n, etc... This leads to the RG improved perturbation theory. The reader is referred to Ref. [63] for more details.

1.2 FCNC decays as a powerful probe of SM

Due to the unitarity of the CKM matrix, Flavor Changing Charged Current (FCCC) transitions from a down- to a up-type quark or vice versa, are allowed at tree level in SM. On the other hand, FCNC transitions from a down(up)-type to a down(up)-type quark, are prohibited since the unitary transformations between two left-handed

down(up) quarks will cancel one another exactly. FCNC's are however allowed via loop processes, although suppressed by the Glashow–Iliopoulos–Maiani (GIM) mechanism [34]. For example, the $q_1 \to q_2 \ell^- \ell^+$ transition, where q_1 and q_2 are both either up or down-type quarks of different flavors, proceed via the so-called penguin and box process sketched by the Feynman diagrams shown in Fig. 1.2.

To summarize, FCNC transitions are interesting because they proceed entirely through loop effects and are additionally suppressed by the chiral nature of the weak interactions (GIM mechanism). Thus, the effects of BSM particles could show up clearly giving to these transitions a potentially large sensitivity to virtual effects in BSM theories.

Following the color confinement property of QCD, quarks are bound with other quarks to form hadronic states. If quarks q_1 and q_2 are part of the same initial hadronic state, the decay is of the type $M \to \ell^+ \ell^-$, where M is a neutral meson made of the anti-quark q_1 and the quark q_2. We will refer to this type of decay as a *purely leptonic* decay. Note that in this case, q_1 has to be a quark and q_2 an anti-quark or vice versa.

If q_1 is bound to a light quark q that do not participate to the weak FCNC transition, called spectator quark, the decay is of the type $M \to P \ell^+ \ell^-$ where M is the initial meson made of q_1 and q, and P is the final state meson made of q_2 (also called the recoiling quark) and q. As we are interested in the case where the initial meson is much heavier than the final one, we will refer to this type of decay as a *heavy-to-light semileptonic* decay.

Experimentally, semileptonic decays can be measured inclusively over all the final state containing q_2 or exclusively by tagging a particular final state. The inclusive measurement is theoretically easier to interpret since the decay rate can be related to the decay of a free q_1 quark into light quark q_2 but is more challenging experimentally, particularly in a busy hadronic environment like at the LHC. Hence the study of exclusive modes is often favored over inclusive.

On the $b \to s\ell^+\ell^-$ transition

If $q_{1/2}$ are down-type quarks, the possible transitions are $s \to d\ell^+\ell^-$ and $b \to d(s)\ell^+\ell^-$. They govern respectively the Kaon and B meson FCNC decays. The $b \to s\ell^+\ell^-$ transition is particularly interesting since it has revealed several hints of Lepton Flavor Universality (LFU) violation in the measurements of $R_K = [\mathcal{B}(B^+ \to K^+\mu^+\mu^-)] / [\mathcal{B}(B^+ \to K^+e^+e^-)]$[3] [10–14] and $R_{K^*} = \left[\mathcal{B}(B^0 \to K^{*0}\mu^+\mu^-)\right] / \left[\mathcal{B}(B^+ \to K^{*0}e^+e^-)\right]$ [10, 11, 15]. Although the fact that no single measurements has a high statistical significance, and that recent data slightly decrease the overall significance of these anomalies; the consistency of data is good and, once combined, the significance of LFU violating observables exceeds 3.7 standard deviations (σ) [21].

In addition to these anomalies, the measurement of $R(D^{(*)}) =$

[3]The branching ratio $\mathcal{B}$ of a decay can be seen as the rate of occurrence of this decay among all the other possible final states. Note that the inclusion of the charge-conjugate processes is implied through this thesis except if explicitly mentioned.

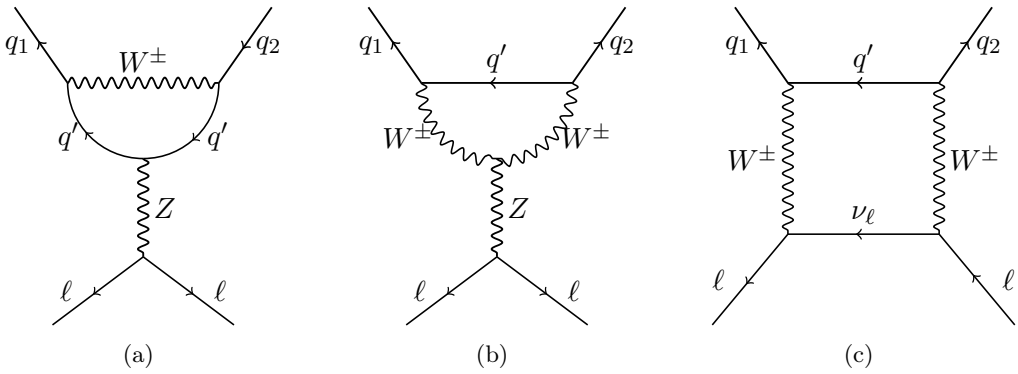


Figure 1.2: $q_1 \to q_2\ell^+\ell^-$ penguins (a,b) and box (c) dominant SM Feynman diagrams. Note that the type of q_1, q_2 (down or up) determines the type of the quark q' (up or down), the charge of the $W^\pm$ boson and the Charge-state of the neutrino in the box diagram (particle or antiparticle).

$\left[\mathcal{B}(B^0 \to D^{(*)}\tau^+\nu_\tau)\right] / \left[\mathcal{B}(B^0 \to D^{(*)}\ell^+\nu_\ell)\right]$, where ℓ^+ is either a muon, an electron or both, has been found to be larger than the SM prediction by 3.1 standard deviations (σ) [16–20]. This points towards LFU violation in the FCCC semileptonic $b \to c\ell\nu$ transition. Both sets of anomalies seem to suggest a hierarchy in the coupling of the generations to BSM particles, where the τ leptons could couple most strongly.

Possible explanations for these and other deviations from their SM expectations include leptoquarks (e.g. see Refs. [21, 22]), W'/Z' bosons (e.g. see Ref. [23]) and two-Higgs-doublet models (e.g. see Ref. [24]). In these models, the purely leptonic $B^0_{(s)} \to \tau^+\tau^-$ branching ratio could be enhanced, with respect to the SM predictions, $\mathcal{B}(B^0 \to \tau^+\tau^-) = (2.22 \pm 0.19) \times 10^{-8}$ and $\mathcal{B}(B^0_s \to \tau^+\tau^-) = (7.73 \pm 0.49) \times 10^{-7}$ [28], by several order or magnitude [24, 29–31].

On the $c \to u\ell^+\ell^-$ transition

As the top quark does not hadronize, the only possible transition, if q_1 and q_2 are up-type quarks, is the $c \to u\ell^+\ell^-$ transition. It governs the D meson FCNC decays and is even more suppressed by the GIM mechanism than the $b \to d(s)\ell^+\ell^-$ transition, due to the absence of a high mass down-type quark. As for B meson decays, the purely leptonic $D^0 \to \ell^+\ell^-$ decay is well predicted and its measurements already allowed to set several constraints on possible BSM scenarios [64,65].

In contrast to B meson physics, charmed semileptonic decay are dominated by large long-distance hadronic contributions. Such effects then screen the short distance contributions of interest. Nonetheless, semileptonic D meson decays could be subject to similar BSM loop effects and are an excellent portal to search for BSM physics in the up-type quark sector. Many BSM models, such as Supersymmetric models with R-parity violation or models involving a fourth quark generation, predict higher value for the branching ratio [32, 33].

Note that the decays $D^+ \to \pi^+\ell^+\ell^-$ and $D^0 \to \ell^+\ell^-$ are related to the decays $B^+ \to K^+\ell^+\ell^-$ and $B^0_s \to \ell^+\ell^-$ respectively via flavor symmetries.

This thesis summarizes the work done to:

- Improve the experimental search for the $B^0_{(s)} \to \tau^+\tau^-$ decay within the LHCb collaboration. An introduction to the $M \to \ell^+\ell^-$ decay phenomenology and experimental aspects is given in Sec. 1.4. The analyses performed to improve the experimental search for the $B^0_{(s)} \to \tau^+\tau^-$ decay are presented in Chap. [4-7].
- Improve the understanding of the long-distance hadronic effects dominating the $D^+ \to \pi^+\ell^+\ell^-$ decay, and thus improve the predictions of the related observables. After a brief overview in Sec. 1.5 of the theoretical and experimental aspect of the $D^+ \to \pi^+\ell^+\ell^-$ decay, Chap. 2 presents the framework used to compute the non-resonant amplitude. Then, Chap. 3 describes the procedure used to model the resonances, that screen the short distance contributions of interest, and the model-independent phenomenological analysis.

The SM predictions for $B^0_{(s)} \to \tau^+\tau^-$ and $D^+ \to \pi^+\ell^+\ell^-$ decays heavily rely on a tool called Effective Field Theory; next section is devoted to this tool.

1.3 Effective Field Theory

The calculation of semileptonic $M \to P\ell^-\ell^+$ and purely leptonic $M \to \ell^-\ell^+$ decays, where M is a D or B meson and P is a light meson requires a wide range of techniques. First, these decays involve a multitude of physical scales: the electroweak scale set by $M_W \sim 80$ GeV that governs the weak interaction responsible for the flavor change; the QCD scale $\Lambda_{QCD} \sim 0.2$ GeV where QCD describes the dynamics of hadrons, as well as the $b(c)$ quark mass in the case of B (D) meson decays. This multi-scale problem is efficiently tackled via Effective Field Theory (EFT) methods.

The key-principle of an effective theory is to take into account the fact that observables at one scale are not directly sensitive to the physics at other scales. As an example, to describe the propagation of a ray of light, we do not need to know anything about quantum electrodynamics. In other words, EFT includes all effects relevant at a given energy scale E, but not those that only play a role at significantly higher scales. The Fermi theory describing the charged current interaction between hadrons, leptons and neutrinos in terms of the Fermi constant G_F and a contact interaction between four fermions, is a textbook example of EFT. In the language of path-integral, the W boson has been "integrated out", meaning removed as dynamic degree of freedom [63].

The differential decay rate $d\Gamma$ of a decay process $M \to F$, where F represents a final state made of several particles, is given by the so-called Fermi's golden rule

$$d\Gamma = \frac{4\pi^4}{2E_M}|\langle F|\mathcal{H}|M\rangle|^2 d\Phi, \tag{1.22}$$

where E_M is the energy of the decaying particle and $d\Phi$ the differential phase space factor. The $M \to F$ decay amplitude $\langle F|\mathcal{H}|M\rangle$ contains the low- and high-energy contributions "mixed" together and can be very complicated or even impossible to compute. The first ingredient in the EFT is the so-called Operator Product Expansion (OPE). It allows one to write the Hamiltonian as a sum of product of

Wilson coefficients C_i that encode the relevant information about the high scale of the theory and local operators $\mathcal{O}_i$ that represent the low-energy effective theory. The effective Hamiltonian is then expressed as

$$\mathcal{H}_{eff} \sim \frac{G_F}{\sqrt{2}} \sum_i \lambda C_i(\mu) \mathcal{O}_i(\mu), \tag{1.23}$$

where λ represents some combinations of CKM elements and μ is the renormalization scale that separate the low and high scales of the EFT. Substituting the full Hamiltonian with the effective Hamiltonian in the decay amplitude, we obtain

$$\langle F|\mathcal{H}|M\rangle \to \langle F|\mathcal{H}_{\text{eff}}|M\rangle = \frac{G_F}{\sqrt{2}} \sum_i \lambda C_i(\mu) \langle F|\mathcal{O}_i(\mu)|M\rangle \tag{1.24}$$

where $\langle F|\mathcal{O}_i(\mu)|M\rangle$ are the matrix elements containing the non-perturbative hadronic interactions.

The operator basis

The operator basis governing rare charmed and beauty decays contains operators contributing to $q_1 \to q_2\gamma$ and $q_1 \to q_2\ell^-\ell^+$ transitions with $q_1 = b$, $q_2 = s$ or d in the case of beauty decay, and $q_1 = c$ and $q_2 = u$ in the case of charmed decay. The basis arises from the general Feynman diagrams given in Fig. 1.3.

The complete set of operators used in this thesis consists of the current-current operators, Fig. 1.3 (a),

$$\mathcal{O}_1^p = (\bar{q}_{2,L}\gamma_\mu T^a p_L)(\bar{p}_L\gamma_\mu T^a q_{1,L}), \tag{1.25}$$
$$\mathcal{O}_2^p = (\bar{q}_{2,L}\gamma_\mu p_L)(\bar{p}_L\gamma_\mu q_{1,L}), \tag{1.26}$$

where $p = u$ or c for beauty decays and $p = d, s$ or b for charmed decay. The strong penguin operators, Fig. 1.3 (b), read as

$$\mathcal{O}_3 = (\bar{q}_{2,L}\gamma_\mu q_{1,L}) \sum_{\{p:m_p\leq\mu\}} (\bar{p}\gamma^\mu p), \tag{1.27}$$
$$\mathcal{O}_4 = (\bar{q}_{2,L}\gamma_\mu T^a q_{1,L}) \sum_{\{p:m_p\leq\mu\}} (\bar{p}\gamma^\mu T^a p), \tag{1.28}$$
$$\mathcal{O}_5 = (\bar{q}_{2,L}\gamma_\mu\gamma_\nu\gamma_\rho q_{1,L}) \sum_{\{p:m_p\leq\mu\}} (\bar{p}\gamma^\mu\gamma_\nu\gamma_\rho p), \tag{1.29}$$
$$\mathcal{O}_6 = (\bar{q}_{2,L}\gamma_\mu\gamma_\nu\gamma_\rho T^a q_{1,L}) \sum_{\{p:m_p\leq\mu\}} (\bar{p}\gamma^\mu\gamma_\nu\gamma_\rho T^a p). \tag{1.30}$$

The electro- and chromo-magnetic penguin operators, Fig. 1.3 (c,d), are given by

$$\mathcal{O}_7 = -\frac{g_e}{16\pi^2} m_{q_2} (\bar{q}_{2,L}\sigma^{\mu\nu} q_{1,R}) F_{\mu\nu}, \tag{1.31}$$
$$\mathcal{O}_8 = -\frac{g_s}{16\pi^2} m_{q_2} (\bar{q}_{2,L}\sigma^{\mu\nu} T^a q_{1,R}) G^a_{\mu\nu}. \tag{1.32}$$

Finally, the semileptonic operators, Fig. 1.3 (e), are written as

$$\mathcal{O}_9 = -\frac{\alpha_e}{4\pi} (\bar{q}_{2,L}\gamma^\mu q_{1,L})(\bar{\ell}\gamma^\mu \ell), \tag{1.33}$$
$$\mathcal{O}_{10} = -\frac{\alpha_e}{4\pi} (\bar{q}_{2,L}\gamma^\mu q_{1,L})(\bar{\ell}\gamma^\mu\gamma_5 \ell). \tag{1.34}$$

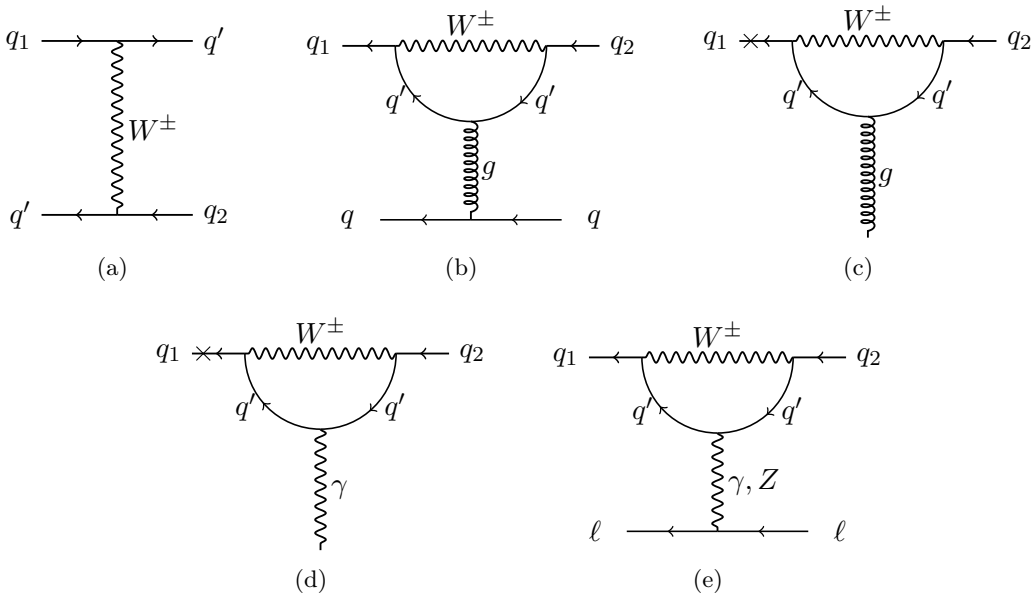


Figure 1.3: Tree and penguin diagrams involved in the $q_1 \to q_2\ell^+\ell^-$ and the $q_1 \to q_2\gamma$ transition in the SM. The $\times$ marks a chirality flip of the external quark line.

T^a are the $SU(3)_c$ generators, $q_{L/R} = (1 \mp \gamma_5)q/2$ denotes the left/right-handed quark fields, m_{q_2} is the q_2 quark mass in the $\overline{\text{MS}}$ scheme, $g_e = \sqrt{4\pi\alpha_e}$ is the electromagnetic coupling and α_e the fine structure constant, $F_{\mu\nu}$ and $G^a_{\mu\nu}$ are respectively the electromagnetic and chromomagnetic field strength tensor. In the semileptonic operators, ℓ and $\bar{\ell}$ are respectively the lepton ℓ^- and the anti-lepton ℓ^+.

To be exhaustive, the SM basis should also contain the chirality-flipped operators $\mathcal{O}'_i$ identical to the $\mathcal{O}_i$ up to the transformation $q_{,L/R} \to q_{,R/L}$ but the contributions of which are extremely small in the SM. A schematic representation of the semileptonic operators $\mathcal{O}_9$ and $\mathcal{O}_{10}$ is given in Fig. 1.4.

If new states are heavy compared to the quark scale, their effect on low-energy scale observables can be described by a modification of the Wilson coefficients associated to the SM operators or by the appearance of new operators not present in the SM. This is an example of the so-called bottom-up approach, when EFT are used to parametrize our ignorance of the full underlying theory, here the BSM physics. The basis can be extended to include the scalar and the pseudoscalar operators:

$$\mathcal{O}_S^{(\prime)} = \frac{g_e^2}{16\pi^2}(\bar{q}_1 q_{2,R(L)})(\bar{\ell}\ell), \tag{1.35}$$

$$\mathcal{O}_P^{(\prime)} = \frac{g_e^2}{16\pi^2}(\bar{q}_1 q_{2,R(L)})(\bar{\ell}\gamma_5\ell), \tag{1.36}$$

and the tensor operators:

$$\mathcal{O}_T = \frac{g_e^2}{16\pi^2}(\bar{q}_1\sigma_{\mu\nu}q_2)(\bar{\ell}\sigma^{\mu\nu}\ell), \tag{1.37}$$

$$\mathcal{O}_{T5} = \frac{g_e^2}{16\pi^2}(\bar{q}_1\sigma_{\mu\nu}q_2)(\bar{\ell}\sigma^{\mu\nu}\gamma_5\ell). \tag{1.38}$$

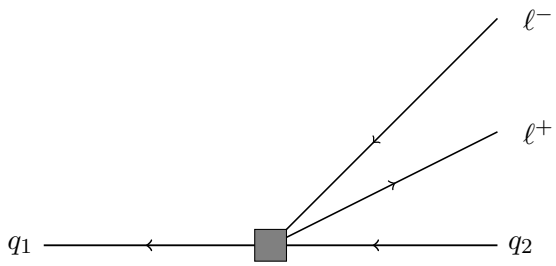


Figure 1.4: Schematic representation of the semileptonic operators $\mathcal{O}_9$ and $\mathcal{O}_{10}$ describing the transition . The short-distance contribution involving particles at mass scales above the scale of interest are integrated out and are represented by the shaded box.

Calculation of the Wilson coefficients

The calculation of the Wilson coefficient values in the SM at the physical mass scale relevant for B/D meson decays ($\mu \sim m_{b/c}$) proceeds in two steps. First, they are determined by calculating Feynman diagrams in the full SM as those shown in Fig. 1.3 and then matching the result to the effective theory. At the electroweak scale $\mu \sim M_W$, QCD is perturbative[4] and the Wilson coefficients can be expanded as

$$C_i = C_i^{(0)} + a_s C_i^{(1)} + a_s^2 C_i^{(2)} + \mathcal{O}(a_s^3), \tag{1.39}$$

where $C_i^{(0)}$ and $C_i^{(n)}$ denotes respectively the tree-level contribution and the n-loop contribution. Secondly, they are run down to the scale of interest via

$$\vec{C}(\mu_2) = U^{(n_f)}(\mu_2, \mu_1)\ \vec{C}(\mu_1), \tag{1.40}$$

where $\vec{C}$ is a vector of Wilson coefficients and $U^{n_f}(\mu_2, \mu_1)$ is the evolution matrix from the scale μ_1 to μ_2 and n_f refers to the number of active flavors in the EFT. Note that in the case of the D decays, we will cross a threshold ($\sim \mu_b$) where the number of active quark flavors in the EFT passes from 5 to 4. Hence a second matching procedure between the (n_f=5) EFT and the (n_f=4) EFT is necessary. The evolution matrix is obtained by solving the following Renormalization Group Equation for the Wilson coefficient,

$$\frac{d\vec{C}(\mu)}{d\ln\mu} = \gamma^T \vec{C}(\mu), \tag{1.41}$$

where γ is the *anomalous dimension matrix* which describes the mixing of the different operators and the evolution to the low-scale. The NNLL accuracy of our calculations, that will be described in Sec. 2.3, requires the Wilson coefficients at $\mu \sim \mu_W$ to two-loop accuracy and the anomalous dimension matrix at order 2.

We are now armed with one of the main tool, required to obtain predictions for the $D^+ \to \pi^+\ell^+\ell^-$ decay and the $B^0_{(s)} \to \ell^+\ell^-$ decay, namely EFT. But his tool is not enough in itself to do any predictions; the matrix elements $\langle F|\mathcal{O}_i(\mu)|M\rangle$, introduced in Eq. (1.24), are non-perturbative quantities and can be challenging to calculate. The methods to determine them depend strongly on the type of decay considered. The case of purely leptonic is rather simple and is covered in Sec. 1.4. The case of the $D^+ \to \pi^+\ell^+\ell^-$ decay is more complicated since it requires the introduction of form-factors, QCD factorization corrections and a realistic resonance description; it is introduced in Sec. 1.5.

[4] $\alpha_s(\mu)$ is a good perturbation expansion parameter when μ is above $\sim$1 GeV.

1.4 On the purely leptonic decay $M \to \ell^+\ell^-$

This section is devoted to the $M \to \ell^+\ell^-$ decay, M being a neutral pseudoscalar meson. A pseudoscalar meson is a meson with total spin $J = 0$ and an odd parity P ($J^P = 0^-$). After introducing the theoretical background in Sec. 1.4.1, we will focus on the beauty decays $B^0_{(s)} \to \ell^+\ell^-$ in Sec. 1.4.2 and then on the charmed decay $D^0 \to \ell^+\ell^-$ in Sec. 1.4.3.

1.4.1 Theoretical background

The case of the purely leptonic decay is particularly easy since the operators can be written as a product of a leptonic current j_ℓ and a quark current j_q. This is the so-called "naive factorization". Moreover all hadronic uncertainties are encoded by a single quantity, the meson decay constant f_M,

$$\langle \ell\ell | j_\ell \cdot j_q | M \rangle = \langle \ell\ell | j_\ell | 0 \rangle \cdot \langle 0 | j_q | M \rangle \sim \langle \ell\ell | j_\ell | 0 \rangle \cdot f_M. \tag{1.42}$$

In other words, the meson decay constant parametrizes the transition between the vacuum and the meson M, and is consequently a property of the meson and not of the decay mode. Being a non-perturbative quantity, the decay constant has to be computed by non-perturbative techniques. It is now computed to a precision of the percent level thanks to Lattice Quantum Chromodynamics (LQCD) methods [66] and Light Cone Sum Rules (LCSR) [67–69].

The $M \to \ell^+\ell^-$ decay can receive contributions from the semileptonic operator $\mathcal{O}_{10}$ and from the potential BSM scalar and pseudoscalar operators. Following the framework introduced in Sec. 1.3, the effective Hamiltonian governing the decay is then expressed by

$$\mathcal{H}_{eff}(M \to \ell^+\ell^-) = -\frac{G_F \alpha_e}{\sqrt{2}\pi} \sum_{i=10,S,P} (C_i^{M\ell} \mathcal{O}_i^{M\ell} + C_i'^{M\ell} \mathcal{O}_i'^{M\ell}) + h.c., \tag{1.43}$$

where the local operators $\mathcal{O}_i$ are those given in Sec. 1.3. Note that the operators, C_{10}, C_S and C_P depend on q_1 and q_2 and hence are not the same for the $c \to u\ell^+\ell^-$ and $b \to s\ell^+\ell^-$ transitions. This difference has been made explicit by introducing the superscript M on the Wilson coefficients and the operators. The superscript ℓ is also added to enlighten the fact that the contributions can be different between the three generations of leptons. The branching ratio is given by [70]:

$$\mathcal{B}(M \to \ell^+\ell^-) = \frac{1}{\Gamma_M} \frac{G_F^2 \alpha_e^2}{64\pi^3} f_M^2 m_M^5 \beta(m_M^2) \left(|P^{M\ell}|^2 + \beta^2(m_M^2) |S^{M\ell}|^2 \right) \tag{1.44}$$

where

$$P^{M\ell} = \frac{1}{m_{q_1}} (C_P^{M\ell} - C_P'^{M\ell}) + \frac{2m_\ell}{m_M^2} (C_{10}^{M\ell} - C_{10}'^{M\ell}) \tag{1.45}$$

$$S^{M\ell} = \frac{1}{m_{q_1}} (C_S^{M\ell} - C_S'^{M\ell}) \tag{1.46}$$

where $\beta(s) = \sqrt{1 - \frac{4m_\ell^2}{s}}$ is a kinematic function and S and P are the purely BSM scalar and pseudoscalar contributions respectively. The expression of the branching

ratio given in Eq. (1.44) is valid for the $D^0 \to \ell^+\ell^-$ and $B^0_{(s)} \to \ell^+\ell^-$ decay. Nonetheless, the different value of Wilson coefficients between the charmed and the beauty decay leads to different phenomenology. The beauty case is presented in Sec. 1.4.2 while the charmed case is explored in Sec. 1.4.3.

1.4.2 The beauty case

For the $B^0_{(s)} \to \ell^+\ell^-$ decay in the SM, $C_P = 0$, $C_S = 0$ and the branching ratio only depends on $C^\ell_{10,\mathrm{SM}}$. The value of this latter is the same for every generation of lepton in agreement with the SM lepton flavor universality assumption.

We note that the $C^\ell_{10,\mathrm{SM}}$ is suppressed by m_ℓ in Eq. (1.44). This particularity is due to the so-called helicity-suppression[5] which occurs since the two spin-1/2 leptons originate from a mother particle of spin 0. Indeed, in the M meson rest frame, the two daughter particles are emitted back-to-back along a given direction, say x. In order to conserve the spin projection along this x-axis, the spin projections of the two daughters must be opposite. This means that the two daughters have the same helicity. Now, a massless fermion and its antiparticle have opposite chirality. Since in the massless limit chirality eigenstates are also helicity eigenstates, it follows that the decay of a spin zero particle into a pair of massless spin-1 particle-antiparticle is strictly forbidden.

Contrary to $\mathcal{O}_{10}$, scalar and pseudoscalar contributions are not helicity-suppressed and a large enhancement of the branching ratio can be seen for models presenting such type of contributions.

In the case of the B^0_s decay, the difference of lifetime between the light and heavy mass eigenstates, $y_s = \frac{\Gamma_L - \Gamma_H}{\Gamma_L + \Gamma_H} = 0.063(5)$ [71], leads to a sizable difference between the prompt branching ratio, given in Eq. (1.44) and denoted $\mathcal{B}(B^0_s(t=0) \to \ell^+\ell^-)$ in the following, and the time-integrated branching ratio measured by experiments

$$\mathcal{B}(B^0_s \to \ell^+\ell^-) = \left[\frac{1 + \mathcal{A}_{\Delta\Gamma} \cdot y_s}{1 - y_s^2}\right] \mathcal{B}(B^0_s(t=0) \to \ell^+\ell^-). \tag{1.47}$$

$\mathcal{A}_{\Delta\Gamma}$ is the relative difference between the light and the heavy mass eigenstate decay width $\Gamma_{B^0_{s,L} \to \ell\ell}$ and $\Gamma_{B^0_{s,H} \to \ell\ell}$,

$$\mathcal{A}_{\Delta\Gamma} = \frac{\Gamma_{B^0_{s,H} \to \ell\ell} - \Gamma_{B^0_{s,L} \to \ell\ell}}{\Gamma_{B^0_{s,H} \to \ell\ell} + \Gamma_{B^0_{s,L} \to \ell\ell}}. \tag{1.48}$$

In the SM, $\mathcal{A}_{\Delta\Gamma} = 1$ which leads to the following correction to the $B^0_s \to \ell^+\ell^-$ decay time-integrated branching ratio

$$\mathcal{B}(B^0_s \to \ell^+\ell^-) \sim 1.06\, \mathcal{B}(B^0_s(t=0) \to \ell^+\ell^-). \tag{1.49}$$

Taking into account additional corrections, the values of the SM predictions [28] for the time-integrated branching ratio for the three generations of leptons ($\ell = e, \mu, \tau$) and the two flavor of meson (B^0 and B^0_s) are reported in Tab. 1.2. We note that the branching ratio values increase with the lepton mass in agreement with helicity-suppression and that the $B^0 \to \ell^+\ell^-$ decay is more CKM suppressed than

ℓ	$\mathcal{B}(B_s^0 \to \ell^+\ell^-)$	$\mathcal{B}(B^0 \to \ell^+\ell^-)$
e	$(8.54 \pm 0.55) \times 10^{-14}$	$(2.48 \pm 0.21) \times 10^{-15}$
μ	$(3.96 \pm 0.23) \times 10^{-9}$	$(1.06 \pm 0.09) \times 10^{-10}$
τ	$(7.73 \pm 0.49) \times 10^{-7}$	$(2.22 \pm 0.19) \times 10^{-8}$

Table 1.2: Values of the time-integrated branching ratio SM predictions for the three generations of leptons ($\ell = e, \mu, \tau$) and the two flavors of meson (B^0 and B_s^0) [28].

the $B_s^0 \to \ell^+\ell^-$ decay.

The phenomenology of the $B_s^0 \to \ell^+\ell^-$ decay is very rich since the measurement of $\mathcal{A}_{\Delta\Gamma}$ and $\mathcal{B}(B_s^0 \to \tau^+\tau^-)$ would provide different but complementary information on the quantities P and S, since both measurements are sensitive to different combinations of P and S. Technically, $\mathcal{A}_{\Delta\Gamma}$ could be extracted from a measurement of the $B_s^0 \to \ell^+\ell^-$ effective lifetime since

$$\mathcal{A}_{\Delta\Gamma} = \frac{\mathrm{Re}(P^2 - S^2)}{|P|^2 + |S|^2}, \tag{1.50}$$

$$\mathcal{A}_{\Delta\Gamma} = \frac{1}{y_s} \frac{(1 - y_s^2)\tau_{B_s^0 \to \ell^+\ell^-} - (1 + y_s^1)\tau_{B_s^0}}{2\tau_{B_s^0} - (1 - y_s^2)\tau_{B_s^0 \to \ell^+\ell^-}}, \tag{1.51}$$

with

$$\tau_{B_s^0 \to \ell^+\ell^-} = \frac{\int_0^\infty t \langle \Gamma(B_s^0(t) \to \ell^+\ell^-) \rangle}{\int_0^\infty \langle \Gamma(B_s^0(t) \to \ell^+\ell^-) \rangle}. \tag{1.52}$$

Experimental aspects

Before the start of data taking at the LHC, no experimental evidence for any of the $B_{(s)}^0 \to \ell^+\ell^-$ decays had been found and only upper limits were set for all six decays except $B_s^0 \to \tau^+\tau^-$. LHCb has been the first single experiment to observe the $B_s^0 \to \mu^+\mu^-$ decay (with a statistical significance of 7.8 σ) and to set an upper limit on the $B_s^0 \to \tau^+\tau^-$ decay. The collaboration also sets the best world limit on the $B^0 \to \tau^+\tau^-$ and the $B_s^0 \to \tau^+\tau^-$ decays and performed the first measurement of the $B_s^0 \to \mu^+\mu^-$ effective lifetime, $\tau(B_s^0 \to \mu^+\mu^-) = 2.04 \pm 0.44 \pm 0.05$ps [25], value in agreement with the SM prediction.

An overview of the current best measurements of the branching ratio for the six decays is shown in Tab. 1.3. All results are in agreement with the SM. The $B^0 \to \mu^+\mu^-$ upper limit is only twice above the SM predictions and should be soon discovered if its value is in agreement with the SM prediction. For the electronic channel, limits are still seven order of magnitude above the SM predictions due to a strong helicity suppression and to the fact that electrons are experimentally more challenging to reconstruct than muon due to Bremsstrahlung effect. On the tauonic side, SM predictions for the branching fractions are larger than the one for the muonic channel, due to helicity suppression. However, as we will see in Sec. 5.1.1, the final state contains between 2 and 4 neutrinos, making the $B_{(s)}^0 \to \tau^+\tau^-$ decay considerably more difficult to reconstruct than its electronic and muonic counterparts.

[5]The helicity is the spin projection along the momentum direction.

Decay	Measurement	Experiments
$B^0 \to e^+e^-$	$< 8.3 \times 10^{-8}$ at 90% C.L.	CDF [72]
$B_s^0 \to e^+e^-$	$< 2.8 \times 10^{-7}$ at 90% C.L.	CDF [72]
$B^0 \to \mu^+\mu^-$	$< 2.1 \times 10^{-10}$ at 95% C.L.	ATLAS [26]
$B_s^0 \to \mu^+\mu^-$	$(3.0 \pm 0.6^{+0.3}_{-0.2}) \times 10^{-9}$	LHCb [25]
$B_s^0 \to \mu^+\mu^-$	$(2.8^{+0.8}_{-0.7}) \times 10^{-9}$	ATLAS [26]
$B_s^0 \to \mu^+\mu^-$	$(3.9^{+1.0}_{-0.9}) \times 10^{-9}$	CMS [27]
$B^0 \to \tau^+\tau^-$	$< 1.6 \times 10^{-3}$ at 90% C.L.	LHCb [45]
$B_s^0 \to \tau^+\tau^-$	$< 5.2 \times 10^{-3}$ at 90% C.L.	LHCb [45]

Table 1.3: Overview of the current best branching ratio measurements for the six decays $B^0_{(s)} \to \ell^+\ell^-$. Limits are not independent for the τ channel: when setting limit on $\mathcal{B}(B_s^0 \to \ell^+\ell^-)$ it is assumed that $\mathcal{B}(B_s^0 \to \ell^+\ell^-) = 0$ and vice versa.

Decay	Limit at 90% C.L.	Experiments
$D^0 \to e^+e^-$	$< 7.9 \times 10^{-8}$	Belle [73]
$D^0 \to \mu^+\mu^-$	$< 6.2 \times 10^{-9}$	LHCb [41]

Table 1.4: Overview of the current best branching ratio measurements for the two decays $D^0 \to \ell^+\ell^-$.

1.4.3 The charmed case

In the SM, the $D^0 \to \ell^+\ell^-$ decay is even more suppressed by the GIM mechanism than the $B^0_{(s)} \to \ell^+\ell^-$ decays due to the absence of a high mass down-type quark. The short distance contribution to the branching ratio is of the order of 10^{-18} (C_{10}) and the long distance contribution, dominated by the two photons intermediate state, is about $2.7 \times 10^{-5} \times \mathcal{B}(D^0 \to \gamma\gamma)$ [74]. This translates, for the $D^0 \to \mu^+\mu^-$ decay, into the following bound:

$$\mathcal{B}(D^0 \to \mu^+\mu^-)^{\text{SM}} < 6 \times 10^{-11}. \tag{1.53}$$

The branching ratio measurement of the $D^0 \to \mu^+\mu^-$ decay provides precious information on the bounds on the Wilson coefficients $C_{10/S/P}$ and we will use these bounds in the phenomenological study of the $D^+ \to \pi^+\ell^+\ell^-$ decay in Sec. 3.2.

The current best experimental upper limits for the $D^0 \to \ell^+\ell^-$ decay are reported in Tab. 1.4. Note that the tauonic final state is not possible as the di-tau system invariant mass is larger than the D^0 mass. The limit on the muonic final state has been obtained using a data sample of 0.9 fb^{-1} collected at a center-of-mass energy of 7 TeV by the LHCb experiment.

1.5 On the semileptonic decay $D^+ \to \pi^+\ell^+\ell^-$

For D decays, the EFT expansion is divided into two steps. First, the W boson is integrated out at a scale $\mu_W \sim M_W$ (as for B decays). At this step, no penguin operators

are generated and only operators $\mathcal{O}_1$ and $\mathcal{O}_2$ are present in the SM Hamiltonian:

$$\mathcal{H}_{\text{eff}}^{\text{SM}}(\mu_W > \mu > \mu_b) = \frac{4\,G_F}{\sqrt{2}} \sum_{q=d,s,b} \lambda_q \left(C_1 \mathcal{O}_1^q + C_2 \mathcal{O}_2^q\right), \tag{1.54}$$

where λ_i is the combination of CKM elements $\lambda_i = V_{ci}^* V_{ui}$. Next, we integrate out the b-quark at a scale $\mu_b \sim m_b$. This generates penguin operators C_{3-9} with Wilson coefficients depending on μ_W only through $C_{1/2}$. Thus, the effective Hamiltonian is given by

$$\mathcal{H}_{\text{eff}}^{\text{SM}}(\mu_W > \mu > \mu_b) = \frac{4\,G_F}{\sqrt{2}} \sum_{q=d,s} \lambda_q \left(C_1 \mathcal{O}_1^q + C_2 \mathcal{O}_2^q + \sum_{i=3}^{9} C_i \mathcal{O}_i\right). \tag{1.55}$$

This SM Hamiltonian for $D^+ \to \pi^+ \ell^+ \ell^-$ can be extended to receive BSM contributions:

$$\mathcal{H}_{\text{eff}} = \mathcal{H}_{\text{eff}}^{\text{SM}} + \frac{4G_F}{\sqrt{2}} \left(\sum_{i=10,S,P,T,T5} C_i \mathcal{O}_i + \sum_{i=7,9,10,S,P} C_i' \mathcal{O}_i' \right). \tag{1.56}$$

The CKM factor combinations have been integrated in the BSM Wilson coefficients. Note that compared to the purely leptonic decay $D^0 \to \ell^+ \ell^-$ decays, which are only sensitive to effects in $\mathcal{O}_{10}^{(\prime)}$, $\mathcal{O}_S^{(\prime)}$ and $\mathcal{O}_P^{(\prime)}$, the semileptonic decay $D^+ \to \pi^+ \ell^+ \ell^-$ is sensitive to a more diverse range of BSM effects.

Contrary to purely leptonic decays, the hadronic part is not just a single constant but a set of functions of the dilepton invariant mass squared noted $q^2 = (p-k)^2$ where p^μ, k^μ and q^μ are respectively the four-momenta of the D, the π and the dilepton system. Note that the dilepton invariant mass squared will sometimes be dubbed $s \equiv q^2$ in the following. These functions are the so-called *form factors* functions $f(s)$; there are only three independent functions in the D to pseudoscalar transition, the scalar f_0, the vector f_+ and the tensor f_T form-factors. Their factorization formula relating matrix elements to form factors are given by:

$$\langle \pi(k) | \bar{u} \gamma^\mu (1 \pm \gamma_5) c | D(p) \rangle = f_+(q^2) \left[(p+k)^\mu - q^\mu \frac{m_D^2 - m_\pi^2}{q^2} \right] + f_0(q^2) \frac{m_D^2 - m_\pi^2}{q^2} q^\mu, \tag{1.57}$$

and

$$\langle \pi(k) | \bar{u} \sigma^{\mu\nu} (1 \pm \gamma_5) c | D(p) \rangle = i \frac{f_T(q^2)}{m_D + m_\pi} \Big[(p+k)^\mu q^\nu - (p+k)^\nu q^\mu \pm i \epsilon^{\mu\nu\alpha\beta} (p+k)_\alpha q_\beta \Big]. \tag{1.58}$$

More details about the computation and the expressions of these quantities are given in Sec. 2.1. The leading contributions in an expansion in the strong coupling α_s for the $D^+ \to \pi^+ \ell^+ \ell^-$ decay arise from:

- diagrams in Fig. 1.4 described by the semileptonic operator $\mathcal{O}_9$,
- diagrams in Fig. 1.5(a) and Fig. 1.5(b) where the charged lepton pair originates from a virtual photon γ^* coming from a purely flavor-conserving interaction,

- diagrams in Fig. 1.5(c) that represent annihilation topologies.

In naive factorization, only the two first types of contributions are taken into account. Annihilation topologies are part of the so-called *non-factorizable* contributions. By non-factorizable, we mean all contributions that do not enter in the definition of the hadronic form factors. Hence, in contrast to the corresponding B decays, the naive factorization hypothesis does not even serve as a first approximation here. As we will see, the decay amplitude turns out to be dominated by non-factorizable dynamics, through annihilation topologies. These non-factorizable contributions are computed in the framework of Quantum Chromodynamics factorization (QCDF). In this framework, the decay amplitude is schematically expressed as

$$\langle \pi \ell^+ \ell^- | \mathcal{H}_{\text{eff}}^{(q)} | D \rangle \sim C^{(q)} f(s) + \phi_D^{\pm} \otimes T^{(q)} \otimes \phi_\pi + \mathcal{O}(\frac{1}{m_c}). \tag{1.59}$$

The non-factorizable corrections enter the coefficients $C^{(q)}$ as well as the second term made of a convolution product between the so-called hard kernel $T^{(q)}$ where $\phi_D^{\pm}$ is the Light Cone Distribution Amplitude (LCDA) of the D meson and is given by

$$\langle 0 | \bar{d}_\beta(z) P(z,0) c_\alpha(0) | \bar{D}(p) \rangle = -\frac{i f_D m_D}{4} \left[\frac{1+\not{v}}{2} \left(2\phi_D^+(t) + \frac{\phi_D^-(t) - \phi_D^+(t)}{t} \not{z} \right) \right]_{\alpha\beta}. \tag{1.60}$$

We take the notation from [75], f_D and v are the decay constant and the velocity of the D meson respectively. ϕ_π is the LCDA of the π meson and is defined at the scale μ by [76]:

$$\langle \pi(k) | \bar{u}(0) \gamma_\mu \gamma_5 d(x) | 0 \rangle = -i f_\pi k_\mu \int_0^1 du \; e^{i\bar{u} k \cdot x} \phi_\pi(u, \mu^2) \tag{1.61}$$

More details on these non-factorizable corrections are given in Sec. 2.2.

The closed fermion loop in diagram (b) of Fig. 1.5 (or by closing two quark lines in the annihilation topology diagram) is calculable perturbatively as long as the $q\bar{q}$ pairs remains off-shell; the perturbative result at the scale μ is given by

$$\tilde{h}^{(\text{pt})}(s, m_q) = -ln\frac{m_q^2}{\mu^2} - \frac{2}{3} - \zeta - (2+\zeta)\sqrt{1-\zeta} \ln \left[\frac{1+\sqrt{1-\zeta}}{\sqrt{-\zeta}} \right], \tag{1.62}$$

where m_q is the mass of the quark entering in the loop and $\zeta = 4m_q^2/(s+i\epsilon)$.

In the region where the invariant dilepton mass squared s corresponds to that of hadronic resonances this factorization breaks down and the quark loop becomes non-perturbative.
In the case of B meson observables, the resonances due to the light resonances (ρ^0, ω^0 and ϕ) are neglected as the resulting effects in binned observables are negligible since the bin size is large compared to the width of the states[6]. Nonetheless, polluting resonant effects due to the c-quark loop, are much larger and the kinematic regime where charmonium resonances are produced is ignored and often vetoed in the experimental analysis. However, in the case of D decays such an approach cannot be followed as this would result in most of the available phase space being thrown away. Hence predictions are made in the resonant region, relying on modeling the structure

[6]Note that this is for a similar reason that we ignore the contributions from pseudoscalar meson, e.g. η, η' in our $D^+ \to \pi^+ \ell^+ \ell^-$ prediction.

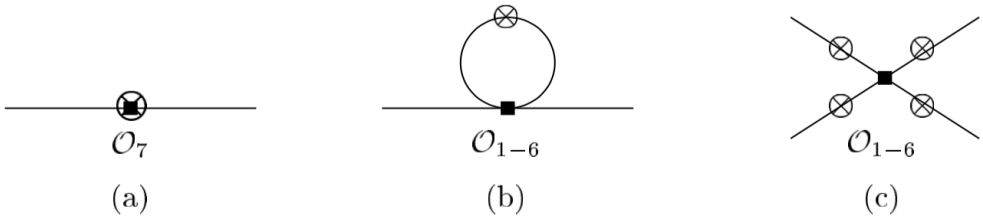


Figure 1.5: Leading contribution to $\langle \pi \ell^+ \ell^- | \mathcal{H}_{\rm eff}^{(q)} | D \rangle$ in an expansion in the strong coupling. The circled cross marks the possible insertion of a virtual photon.

of the resonances.

Initially, for D decays, the resonances were added by hand, by means of Breit-Wigner functions, on top of a non-resonant background described by the partonic result for the quark vacuum polarization. In [36] an alternative approach is advocated, where they use a once-subtracted dispersion relation with the subtraction constant fixed from the perturbative description:

$$\tilde{h}_q(s) \to \tilde{h}^{(\rm pt)}(-s_0, m_q) + \frac{1}{\pi} \int_0^\infty ds' \frac{s_0 + s}{s_0 + s'} \frac{\mathcal{I}m \tilde{h}_q(s')}{s' - s - i\epsilon}, \tag{1.63}$$

where $\mathcal{I}m \tilde{h}_q(s')$ are the hadronic spectral functions. In this approach, the partonic result is recovered asymptotically and the resonances are modeled following Regge trajectories with a number of simplifications, the main one being that the isospin 1 and isospin 0 channels are not treated separately, but in terms of a single tower of resonances with "effective" parameters.

A possible route to improve upon the model of [36] is to follow the strategy described in [40]. In this reference, the authors investigate the interference pattern of the charm resonance with the $\mathcal{O}_9$ operators in the branching ratio of $B^+ \to K^+ \mu^+ \mu^-$. For this purpose, they extract the charm vacuum polarization from $e^+ e^- \to$ (hadrons) data by means of a dispersion relation and the optical theorem. The framework of [40] is employed in this thesis too to extract the u, d and s vacuum polarization, although the fact that in the low-energy region the three active quark flavors makes the interplay between the different resonances much more intricate. More details about the framework used to model the resonance structure are given in Sec. 3.1. As the predictions of $D^+ \to \pi^+ \ell^+ \ell^-$ is one of the projects of this thesis, the results are not reported here but in Sec. 3.2.

Experimental aspects

On the experimental side, the current best-world limits on $\mathcal{B}(D^+ \to \pi^+ \mu^+ \mu^-)$ have been obtained by the LHCb experiment [42]. The search has been performed for the non-resonant decay using proton-proton collision data, corresponding to an integrated luminosity of $1.0\,\text{fb}^{-1}$ collected at the center-of-mass energy of 7 TeV. Limits are set on branching fractions in two bins of the dilepton invariant mass squared s and on the total branching fraction excluding the resonant contributions assuming a phase space model. Results are summarized in Tab. 1.5.
The current most stringent limits on $\mathcal{B}(D^+ \to \pi^+ e^+ e^-)$ has been obtained by the BaBar experiment [77]. The measurement is done excluding the $e^+ e^-$ mass region

s-bin	90% C.L. Limit	95% C.L. Limit
full s-region	7.3×10^{-8}	8.3×10^{-8}
Reg. I: $s \in [0.250^2, 0.525^2]$ GeV^2	2.0×10^{-8}	2.5×10^{-8}
Reg. II: $s > 1.25^2$ GeV^2	2.6×10^{-8}	2.9×10^{-8}

Table 1.5: Current best-world limits on $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$ in two bins of the dilepton invariant mass squared s [42]. The total branching fraction excluding the resonant contributions is extrapolated assuming a phase space model.

around the ϕ resonance, $0.95 < m(e^+e^-) < 1.05$ GeV. The non-resonant limit is

$$\mathcal{B}(D^+ \to \pi^+e^+e^-) < 1.1 \times 10^{-6} \text{ at 90 \% C.L.} \tag{1.64}$$

Conclusions

We have now enough elements on $B^0_{(s)} \to \ell^+\ell^-$ decays to introduce the search for $B^0_{(s)} \to \tau^+\tau^-$ decay at LHCb. In addition, we are armed with one of the main tool, required to obtain predictions for the $D^+ \to \pi^+\ell^+\ell^-$ decay, namely effective field theory. The other tools needed: form-factors, QCD factorization and an appropriate resonance description, are described in the next two chapters.

Part I

A phenomenological study of $D^+ \to \pi^+ \ell^+ \ell^-$

Chapter 2

Theoretical framework for the non-resonant $D^+ \to \pi^+\ell^+\ell^-$ amplitude

This chapter presents the framework needed to compute the non-resonant $D^+ \to \pi^+\ell^+\ell^-$ amplitude. Sec. 2.1 is devoted to the the naive amplitude; by "naive" we mean that the result does not include non-factorizable contributions. The QCDF corrections and then the full non-resonant amplitude for the SM are given in Sec. 2.2. Details on how to obtain the Wilson coefficients at the appropriate scale are given in Sec. 2.3.

2.1 The naive amplitude

The SM effective Hamiltonian for the $c \to u\ell^+\ell^-$ transition has already been given in Eq. (1.55). For convenience, we rewrite it with the help of Eq. (1.12) into

$$\mathcal{H}_{\text{eff}}^{\text{SM}}(m_b > \mu > m_c) = -\frac{4\,G_F}{\sqrt{2}} \sum_{q=d,b} \lambda_q \mathcal{H}_{\text{eff}}^{(q)}, \tag{2.1}$$

where

$$\mathcal{H}_{\text{eff}}^{(b)} = C_1\mathcal{O}_1^s + C_2\mathcal{O}_2^s + \sum_{i=3}^{9} C_i\mathcal{O}_i, \tag{2.2}$$

$$\mathcal{H}_{\text{eff}}^{(d)} = C_1(\mathcal{O}_1^s - \mathcal{O}_1^d) + C_2(\mathcal{O}_2^s - \mathcal{O}_2^d). \tag{2.3}$$

This rewriting makes the CKM suppression explicit, i.e. as $\lambda_b \ll \lambda_d$[1], all contributions entering $\mathcal{H}_{\text{eff}}^{(b)}$ are heavily CKM suppressed.

Based on this Hamiltonian, we can write the amplitude $\mathcal{M}(D^+(p) \to \pi^+\ell^-\ell^+)$ as [38]:

$$\begin{aligned}\mathcal{M}(D^+(p) \to \pi^+\ell^-\ell^+) =& i\frac{G_F\alpha_e}{\sqrt{2}\pi}\big[F_V p^\mu(\bar{\ell}\gamma_\mu\ell) + F_A p^\mu(\bar{\ell}\gamma_\mu\gamma_5\ell) \\ &+ (F_S + \cos\theta\; F_T)(\bar{\ell}\ell) + (F_P + \cos\theta\; F_{T5})(\bar{\ell}\gamma_5\ell)\big],\end{aligned} \tag{2.4}$$

[1] $\lambda_b \sim \lambda^4$ and $\lambda_b \sim \lambda^2$ where $\lambda \sim 0.22$ following the Wolfenstein parametrization.

where the functions F_A, F_S, F_P, F_T, and F_{T5} depend on the di-lepton invariant mass squared s, and are expressed in terms of the form factors and Wilson coefficients as

$$F_V(s) = \left[\lambda_b C_9^{(b)}(s) + \lambda_d C_9^{(d)}(s)\right] f_+(s) + \frac{8m_l}{m_D + m_\pi} C_T \; f_T(s), \tag{2.5}$$

$$F_A(s) = (C_{10} + C'_{10}) \; f_+(s), \tag{2.6}$$

$$F_S(s) = \frac{m_D^2 - m_\pi^2}{2m_c}(C_S + C'_S) \; f_0(s), \tag{2.7}$$

$$\begin{aligned} F_P(s) =& \frac{m_D^2 - m_\pi^2}{2m_c}(C_P + C'_P) \; f_0(s) \\ & - m_\ell (C_{10} + C'_{10}) \left(f_+ - \frac{m_D^2 - m_\pi^2}{s} \left[f_0(s) - f_+(s) \right] \right), \end{aligned} \tag{2.8}$$

$$F_T(s) = \frac{2\beta(s)\lambda(s)^{1/2}}{m_D + m_\pi} C_T \; f_T(s), \tag{2.9}$$

$$F_{T5}(s) = \frac{2\beta(s)\lambda(s)^{1/2}}{m_D + m_\pi} C_{T5} \; f_T(s). \tag{2.10}$$

The function $\beta(s)$ and the Kallen function $\lambda(s)$ entering the expression of the tensor functions $F_T(s)$ and $F_{T5}(s)$ are two kinematic functions given by

$$\begin{aligned} \lambda(s) &= (m_D^2 + m_\pi^2 + s)^2 - 4(m_D^2 m_\pi^2 + m_D^2 s + m_\pi^2 s), \\ \beta(s) &= \sqrt{1 - \frac{4m_\ell^2}{s}}. \end{aligned} \tag{2.11}$$

In the Standard Model, all functions F are equal to zero except F_V, i.e.:

$$F^{\rm SM}_{P,A,S,P,T,T5}(s) = 0, \tag{2.12}$$

$F_V^{\rm SM}$ is given by:

$$F_V^{\rm SM}(s) = \left[\lambda_b C_9^{(b)}(s) + \lambda_d C_9^{(d)}(s)\right] f_+(s). \tag{2.13}$$

The functions $C_9^{(b)}(s)$ and $C_9^{(d)}(s)$ are taken from [36] and can be expressed by

$$C_9^{(q)}(s) = \delta^{qb} C_9 + \frac{2m_c}{m_D} \frac{\mathcal{T}^{(q)}(s)}{f_+(s)}. \tag{2.14}$$

We use the terminology of the Ref. [36] and call the following quantities $\mathcal{T}^{(q)}$ the generalized form factors. In naive factorization, they are simply expressed as

$$\mathcal{T}^{(q)}(s)\big|_{\rm Naive} = -f_+(s) C^{(0,q)}(s), \tag{2.15}$$

where the coefficient $C^{(0,q)}$ contains only factorizable and leading contributions and is given by

$$C^{(0,q)}(s) = -\delta^{qb} C_7 - \frac{m_D}{2m_c} Y^{(q)}(s). \tag{2.16}$$

Eq. (2.14) and (2.15) could have been written in a single equation without the need of the quantity $\mathcal{T}^{(q)}$. The role of $\mathcal{T}^{(q)}$ seems a bit convoluted now but will become more clear in the context of QCD factorization discussed in Sec. 2.2. The 1-loop functions $Y^{(q)}$ combine the contribution from the four-quark operators $\mathcal{O}_{1-6}$,

$$Y^{(d)}(s) = -\frac{4}{9}\left(\frac{2}{3}C_1 + \frac{1}{2}C_2\right)\left[\tilde{h}(s,m_s) - \tilde{h}(s,m_d)\right], \tag{2.17}$$

$$\begin{aligned} Y^{(b)}(s) =& \frac{4}{9}\Bigg[\left(7C_3 + \frac{4}{3}C_4 + 76C_5 + \frac{64}{3}C_6\right)\left[\tilde{h}(s,m_c) + \tilde{h}(s,m_u)\right] \\ &- \left(\frac{2}{3}C_1 + \frac{1}{2}C_2 + 3C_3 + 30C_5\right)\tilde{h}(s,m_s) \\ &- (3C_3 + 30C_5)\,\tilde{h}(s,m_d) + 2\left(3C_3 + 16C_5 + \frac{16}{3}C_6\right)\Bigg], \end{aligned} \tag{2.18}$$

where $\tilde{h}(s,m_q)$ is the closed quark loop function described in Eq. (1.62). Note that our closed fermion loop $\tilde{h}_q(s)$ is related to the one of [36], noted $h(s,m_q)$, by $\tilde{h}_q(s) = \frac{9}{4}h(s,m_q)$.

As explained in Sec. 1.5, the perturbative result does not include any long-distance hadronic effect and will be modified following the dispersion relation described in Eq. (1.63). Combining Eq. (2.14), (2.15) and (2.16), the functions $C_9^{(q)}$ in the naive factorization framework, denoted $C_9^{(q)}\big|_{\text{Naive}}$, can be summarized by

$$C_9^{(d)}(s)\big|_{\text{Naive}} = Y^{(d)}(s), \tag{2.19}$$

$$C_9^{(b)}(s)\big|_{\text{Naive}} = Y^{(b)}(s) + C_9 + \frac{2m_c}{m_D + m_\pi}C_7\frac{f_T(s)}{f_+(s)}. \tag{2.20}$$

As explained previously, the naive result is not enough to make reliable predictions as we have to take into account non-factorizable corrections. They can be computed in the framework called QCD factorization, presented in the next section.

2.2 QCD factorization

We have seen in Sec. 1.5 that several contributions are not factorizable, i.e. they are not included in the definition of the form factors and are hence beyond the naive factorization. At leading order in an expansion in α_s and $1/m_c$[2], only the annihilation topologies seen in Fig. 1.5(c) are not factorizable. At NLO, several contributions with a gluon exchange occur and are not factorizable. Following [35], the non-factorizable contributions can be classified into two categories:

- A first category where the spectator quark participates in the hard scattering; annihilation topologies enter in this category. At NLO, the relevant contributions are sketched in Fig. 2.1(a) and (b). In these diagrams, the spectator quark participates in the FCNC process via hard gluon exchange. The calculation of these diagrams leads to the so-called *hard spectator scattering* corrections.

[2]Note that the QCD factorization works better for B than for D meson as the expansion parameter $1/m_b$ is smaller than $1/m_c$.

- A second category where the spectator quark is connected to the hard process only through soft interactions and the hadronic transition can be described by the form factors. These contributions are sketched by diagrams in the second row of Fig. 2.1. Note that the spectator quark line is not drawn for these diagrams. The calculation of which leads to the so-called *form-factor* corrections.

As we can see, these NLO diagrams arise from QCD corrections to the matrix element of purely hadronic operators $\mathcal{O}_{1-6}$ and $\mathcal{O}_8$, of the weak effective Hamiltonian defined in Eq. (2.1). These corrections can be computed in QCDF the combined heavy quark and large energy (recoil) limit [78]. Here, large energy refers to that of the final state meson, where E is related to the dilepton invariant mass s via:

$$E = \frac{m_D^2 + m_\pi^2 - s}{2m_D} \to \sim \frac{m_D}{2}. \tag{2.21}$$

In other words, QCDF approach removes some of the theoretical uncertainties in the kinematic region of small invariant mass of the photon, also called region of large hadronic recoil.

Still following [35], QCDF proposes an expression consisting of two terms, separating the first and the second category of contributions. The first category of corrections (annihilation and hard spectator scattering corrections) are calculated by the convolution of a perturbative hard-scattering kernel $T^{(q)}$ with the light cone distribution amplitudes of the mesons $\phi_{D,\pm}$ and ϕ_π. The second category, form-factor corrections, enters into the coefficient $C^{(q)}$. In analogy to Eq. (1.59), the naive generalized form factors described in Eq.(2.15) can be extended to receive non-factorizable corrections as

$$\mathcal{T}^{(q)}(s) = -C^{(q)} f_+(s) + \frac{\pi^2}{N_c}\frac{f_D f_\pi}{m_D}\sum_{\pm}\int \frac{d\omega}{\omega}\,\phi_{D,\pm}(\omega)\int_0^1 du\,\phi_\pi(u)\,T_\pm^{(q)}(u,\omega). \tag{2.22}$$

where the $\pm$ subscript refers to the projection of the amplitude on the D meson LCDA. The perturbative quantities $C^{(q)}$ and $T_\pm^{(q)}$ are given by

$$\begin{aligned} C^{(q)} &= C^{(0,q)} + a_s C_F C_\pm^{(1,q)}, \\ T_\pm^{(q)} &= T_\pm^{(0,q)} + a_s C_F T_\pm^{(1,q)}, \end{aligned} \tag{2.23}$$

where we remind the reader that $a_s \equiv \alpha_s/(4\pi)$ and $C^{(0,q)}$ is defined in Eq. (2.16). The form-factor corrections are contained in $C^{(1,q)}$, the annihilation corrections in $T^{(0,q)}$ and the hard spectator scattering corrections in $T^{(1,q)}$.

Note that if one takes only the naive contributions into account, Eq. (2.15) and (2.22) are equivalent. Finally, in the QCDF framework the function $C_9^{(q)}$ defined in Eq. (2.14) can be written by:

$$C_9^{(q)} = C_9^{(q)}\big|_{\text{Naive}} + C_9^{(q)}\big|_{\text{Ann}} + C_9^{(q)}\big|_{\text{SS}} + C_9^{(q)}\big|_{\text{FF}}, \tag{2.24}$$

where $C_9^{(q)}\big|_{\text{Naive}}$ has been defined in Eq. (2.20) and $C_9^{(q)}\big|_{\text{Ann}}$, $C_9^{(q)}\big|_{\text{SS}}$, $C_9^{(q)}\big|_{\text{FF}}$ are the annihilation, the spectator scattering and the form factors corrections to the $C_9^{(q)}$ functions respectively. The expressions for these corrections are given in the next three subsections; they are all taken from [36] in which the authors adapted expressions for the $B \to K^*\ell^+\ell^-$ from Ref. [35] to the case of $D \to \rho(\pi)\ell^+\ell^-$.

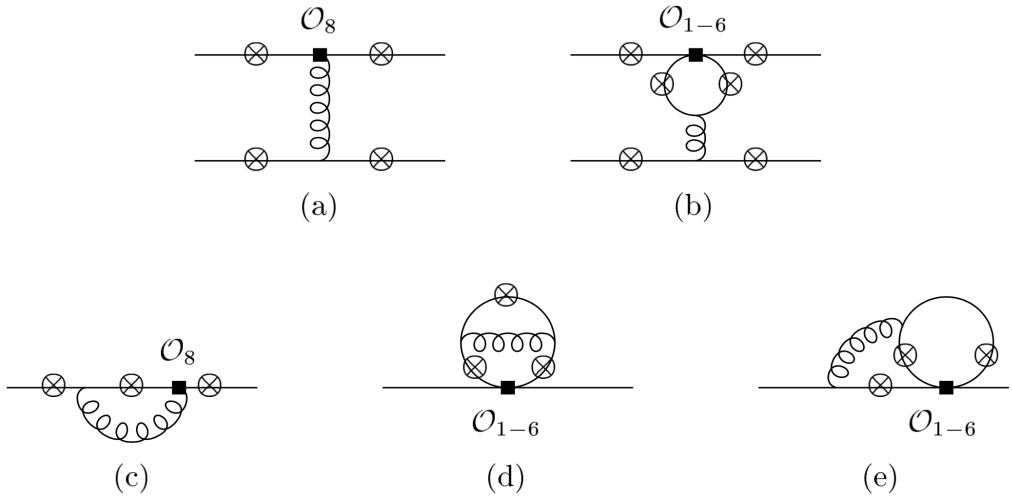


Figure 2.1: Next-to-leading contribution to $\langle \ell^+\ell^-\pi|\mathcal{H}_{\rm eff}^{(q)}|D\rangle$ in an expansion in the strong coupling α_s and $1/m_c$.

2.2.1 Weak annihilation corrections

To compute the annihilation corrections, the amplitude must be projected on the D meson LCDA as explained in [79]. The four diagrams in diagram (c) of Fig. 1.5 contribute at different powers in the $1/m_c$ expansion. With the convention that the π^+ meson momentum is nearly light-like in the minus light-cone direction, the amplitude for the surviving contributions depends only on the minus component and $T_+^{(0)} = 0$. The result depends on the charge factor of the spectator quark, in our case $e_d = -1/3$,

$$T_-^{(0,b)}(\omega) = \frac{e_d}{m_c}\frac{4m_D\omega}{\omega - s/m_D - i\epsilon}\left[-C_3 - \frac{4}{3}(C_4 + 12C_5 + 16C_6)\right], \tag{2.25}$$

$$T_-^{(0,d)}(\omega) = \frac{e_d}{m_c}\frac{4m_D\omega}{\omega - s/m_D - i\epsilon}\, 3C_2. \tag{2.26}$$

$T_-^{(0,d)}$ being u-independent, corrections to $C_9^{(q)}$ are then given by

$$C_9^{(d)}(s)\big|_{\rm Ann} = 8e_d\frac{\pi^2}{N_c}\frac{f_D f_\pi}{m_D}\frac{1}{f_+(s)}\frac{1}{\lambda_D^-(s)}\, 3C_2 \tag{2.27}$$

$$C_9^{(b)}(s)\big|_{\rm Ann} = 8e_d\frac{\pi^2}{N_c}\frac{f_D f_\pi}{m_D}\frac{1}{f_+(s)}\frac{1}{\lambda_D^-(s)}\left[-C_3 - \frac{4}{3}(C_4 + 12C_5 + 16C_6)\right] \tag{2.28}$$

with the s-dependent moment $\lambda_D^-(s)$ given by

$$\frac{1}{\lambda_D^-(s)} = \int_0^\infty d\omega \frac{\phi_D^-(\omega)}{\omega - s/m_D - i\epsilon}. \tag{2.29}$$

We note that $C_9^{(q)}\big|_{\rm Ann}$ leads to a sizable contribution as C_2 appears without any cancellation from other Wilson coefficients. It turns out that annihilation gives very large contributions to the decay rate, therefore it will be essential in the following to take into account the associated hadronic uncertainties. More details about the annihilation uncertainties will be given in Sec. 3.2.2.

2.2.2 Hard spectator scattering corrections

The hard spectator scattering corrections arise at NLO and have been obtained in [79] by computing matrix elements of four-quark operators $\mathcal{O}_{1-6}$ and chromomagnetic $\mathcal{O}_8$ operators represented by diagrams (a) and (b) in Fig. 2.1. The perturbative quantities $T_{\pm}^{(1,q)}$ in Eq. (2.23) can be expressed as

$$T_{+}^{(1,b)} = \frac{m_D}{m_c}\Bigg[(t(u,m_c)+t(u,m_u))\left(C_3 - \frac{1}{6}C_4 + 16C_5 + \frac{10}{3}C_6\right) + e_d t(u,m_s)\left(C_2 - \frac{1}{6}C_1 + 6C_6\right) + e_d t(u,m_d)6C_6\Bigg], \tag{2.30}$$

$$T_{+}^{(1,d)} = e_d \frac{m_D}{m_c}\left(C_2 - \frac{1}{6}C_1\right)(t(u,m_s) - t(u,m_d)), \tag{2.31}$$

and

$$T_{-}^{(1,q)} = e_d \frac{\omega}{\omega - s/m_D - i\epsilon}\left[\delta^{ib}\frac{8C_8^{\text{eff}}}{\bar{u} + us/m_D^2} + \frac{6m_D}{m_c}F_V^{(q)}(\bar{u}m_D^2 + us)\right]. \tag{2.32}$$

The $t(u,m_q)$ functions, that arise from the diagrams of Fig. 2.1(b) in which the photon is attached to the internal quark loop, are given by

$$t(u,m_q) = \frac{2m_D}{\bar{u}E}I_1(m_q) + \frac{\bar{u}m_D^2 + us}{\bar{u}^2E^2}\left[B_0(\bar{u}m_D^2 + us, m_q) - B_0(s,m_q)\right], \tag{2.33}$$

where B_0 and I_1 are defined by

$$B_0(s,m_q) = -\sqrt{4m_q^2/s - 1}\arctan\frac{1}{\sqrt{4m_q^2/s-1}}, \tag{2.34}$$

$$I_1(m_q) = 1 + \frac{2m_q^2}{\bar{u}(m_D^2 - s)}\left[L_1(x_+) + L_1(x_-) + L_1(y_+) + L_1(y_-)\right], \tag{2.35}$$

and

$$L_1(x) = \ln\frac{x-1}{x}\ln(1-x) - \frac{\pi^2}{6} + \text{Li}_2\left(\frac{x}{x-1}\right), \tag{2.36}$$

$$x_{\pm} = \frac{1}{2} \pm \left(\frac{1}{4} - \frac{m_q^2}{\bar{u}m_D^2 + us}\right)^{1/2}, \qquad \frac{1}{2} \pm \left(\frac{1}{4} - \frac{m_q^2}{s}\right)^{1/2}. \tag{2.37}$$

The function $\text{Li}_2(z)$ is the dilogarithm function defined by

$$\text{Li}_2(z) = -\int_0^z du \frac{\ln(1-u)}{u}, \quad z \in \mathbb{C}. \tag{2.38}$$

The corrections $C_9^{(q)}\big|_{\text{SS}}$ are finally given by

$$C_9^{(q)}\big|_{\text{SS}}(s) = \frac{2m_c}{m_D}\frac{1}{f_+(s)}\frac{\pi^2}{N_c}\frac{f_D f_\pi}{m_D}a_s C_F \sum_{\pm}\int\frac{d\omega}{\omega}\phi_{D,\pm}(\omega)\int_0^1 du\ \phi_\pi(u)\ T_{\pm}^{(1,q)}(u,\omega). \tag{2.39}$$

Note that the functions $T_{+}^{(1,q)}$ are independent of ω which simplifies the integration in Eq. (2.39).

2.2.3 Form factor corrections

The corrections have been obtained in Ref. [79] by computing matrix elements of four-quark operators and chromomagnetic operators represented by diagrams (c), (d) and (e) in Fig. 2.1. The perturbative quantities $C_{\pm}^{(1,q)}$ in Eq. (2.23) can be expressed as

$$C_F C^{(1,q)} = C_1 \left(\delta_{qd} F_{1,d}^{(7)} - F_1^{(7)}\right) + C_2 \left(\delta_{qd} F_{2,d}^{(7)} - F_2^{(7)}\right) + \delta_{qb} C_8^{\text{eff}} F_8^{(7)} \tag{2.40}$$

$$+ \frac{m_D}{2m_c} \left[C_1 \left(\delta_{qd} F_{1,d}^{(9)} - F_1^{(9)}\right) + C_2 \left(\delta_{qd} F_{2,d}^{(9)} - F_2^{(9)}\right) + \delta_{qb} C_8^{\text{eff}} F_8^{(9)}\right]. \tag{2.41}$$

The functions $F_8^{(7)}$ and $F_8^{(9)}$ are taken from [80]

$$F_8^{(7)} = \frac{64}{9} \ln \frac{\mu}{m_c} \frac{16}{9} \frac{\hat{s}}{1-\hat{s}} \ln \hat{s} \frac{16}{9} i\pi + \frac{8}{9} \frac{11 - 16\hat{s} + 8\hat{s}^2}{(1-\hat{s})^2}$$
$$- \frac{8}{9} \frac{1}{(1-\hat{s})^3} \left[(9\hat{s} - 5\hat{s}^2 + 2\hat{s}^3) B_0(\hat{s}) - (4 + 2\hat{s}) C_0(\hat{s})\right], \tag{2.42}$$

$$F_8^{(9)} = -\frac{32}{9} \frac{1}{1-\hat{s}} \ln \hat{s} - \frac{16}{9} \frac{5 - 2\hat{s}}{(1-\hat{s})^2} + \frac{16}{9} \frac{4 - \hat{s}}{(1-\hat{s})^3} \left[(1+\hat{s}) B_0(\hat{s}) - 2C_0(\hat{s})\right], \tag{2.43}$$

where $\hat{s} = s/m_c^2$, $B_0(\hat{s}) = B_0(s, m_c^2)$ is given in Eq. (2.34) and C_0:

$$C_0(\hat{s}) = \int_0^1 dx \frac{1}{x(1-\hat{s}) + 1} \ln \frac{x^2}{1 - x(1-x)\hat{s}}. \tag{2.44}$$

As explained in [36], the functions $F_{1/2(,d)}^{(7/9)}$ are taken from the unrenormalized (bare) function from b-decays [81] replacing the charge factors and renormalizing the functions. The renormalization constant with general charge factor Z can be found in [36]. The results for the functions $F_{1/2(,d)}^{(7/9)}$ have been reconstructed from a Mathematica notebook kindly provided by Christoph Greub (related to the work published in [82]).

Finally, the corrections to $C_9^{(q)}$ are given by

$$C_9^{(q)}\big|_{\text{FF}} = -\frac{2m_c}{m_D} a_s C_F C^{(1,q)}. \tag{2.45}$$

2.3 Calculation of the Wilson coefficients

In this section we present the two-step running of the Wilson coefficients for the $c \to u\ell^+\ell^-$ transition. These coefficients are computed at the electroweak scale $\mu_W \sim M_W$ and then run down to the typical mass scale of the decay under consideration, here $\mu_c \sim m_c$ via the use of the evolution matrices. This running involves the intermediate scale (here $\mu_b \sim m_b$), where the bottom quark is "integrated out". The matching coefficients and anomalous dimensions matrices are taken to the required order by generalizing and extending results from $b \to d/s$ transitions. The Wilson coefficients for the $c \to u\ell^+\ell^-$ transition at NNLL accuracy were calculated for the first time in [83]; we have implemented the authors prescription in order to get control on the renormalization scales (μ_W, μ_b, μ_c).

As seen in Sec. 1.5, the running of Wilson coefficients involves solving the RGE; an overview of the standard procedure was presented in Sec. 2.3.1. Sec. 2.3.2 provides a global picture of how the Wilson coefficient C_{1-8} at the scale $\mu < m_b$ can be obtained. The calculation procedure for the last coefficient C_9 is reported in Sec. 2.3.3. The numerical values of the Wilson coefficients are given in the next chapter (Sec. 3.2.1) where all numerical results are contained.

2.3.1 Solving the RGE

As explained previously, the running of the Wilson coefficient is performed via the following equation

$$\vec{C}(\mu_2) = U^{(n_f)}(\mu_2, \mu_1)\ \vec{C}(\mu_1), \tag{2.46}$$

where $\vec{C}(\mu)$ is the vector of Wilson coefficients C_{1-8} at the scale μ and $U^{(n_f)}(\mu_2, \mu_1)$ the so-called evolution matrix. This latter is obtained by solving the RGE

$$\frac{dU^{(n_f)}(\mu)}{d\ln\mu} = \gamma^T U^{(n_f)}(\mu). \tag{2.47}$$

Note that this equation is similar to Eq. (1.41) and that $U^{(n_f)}$ only depends on the anomalous dimension matrix γ. The relations needed to solve this equation are taken from [35] for the $b \to d(s)\ell^+\ell^-$ transition and are summarized below. First, the anomalous matrix is expanded as

$$\gamma = \gamma^{(0)} a_s + \gamma^{(1)} a_s^2 + \gamma^{(2)} a_s^3 + ... \tag{2.48}$$

Let V be the matrix that diagonalizes ${\gamma^{(0)}}^T$, so that

$$V^{-1} {\gamma^{(0)}}^T V = \left[\gamma_i^{(0)}\right]_{\text{diag}}, \tag{2.49}$$

and define

$$U^{(0)}(\mu, M_W) = V \left[\left(\frac{a_s(\mu)}{a_s(M_W)} \right)^{-\gamma_i^{(0)}/(2\beta_0)} \right] V^{-1}, \tag{2.50}$$

which solves Eq. (2.47) to leading order in $a_s(\mu)$. Then the NNLL expression for the evolution matrix

$$\begin{aligned} U(\mu, M_W) = &(1 + a_s(\mu) J^{(1)} + a_s(\mu)^2 J^{(2)}) \\ &\times U^{(0)}(\mu, M_W) \left(1 - a_s(M_W) J^{(1)} - a_s(M_W)^2 \left[J^{(2)} - {J^{(1)}}^2\right]\right) \end{aligned} \tag{2.51}$$

where

$$J^{(n)} = V H^{(n)} V^{-1} \tag{2.52}$$

The matrices $H^{(1)}$ and $H^{(2)}$ are given by

$$H_{ij}^{(1)} = \delta_{ij}\gamma_i^{(0)}\frac{\beta_1}{2\beta_0^2} - \frac{G_{ij}^{(1)}}{2\beta_0 + \gamma_i^{(0)} - \gamma_j^{(0)}} \tag{2.53}$$

$$H_{ij}^{(1)} = \delta_{ij}\gamma_i^{(0)}\frac{\beta_2}{4\beta_0^2} - \sum_k \frac{2\beta_0 + \gamma_i^{(0)} - \gamma_k^{(0)}}{4\beta_0 + \gamma_i^{(0)} - \gamma_j^{(0)}}\left(H_{ik}^{(1)}H_{kj}^{(1)} - \frac{\beta_1}{\beta_0}H_{ij}^{(1)}\delta_{jk}\right) - \frac{G_{ij}^{(2)}}{4\beta_0 + \gamma_i^{(0)} - \gamma_j^{(0)}} \tag{2.54}$$

where $G_{ij}^{(n)}$ is defined by

$$G^{(n)} = V^{-1}\gamma^{(n)^T}V \tag{2.55}$$

2.3.2 Calculation of the C_{1-8} Wilson coefficients

The full procedure to calculate the C_{1-8} Wilson coefficients can be broken down in the following four steps:

1. As noted in Sec. 1.5, only $C_{1/2}$ receive non-zero contributions from the matching procedure at $\mu \sim \mu_W$, the values of which are computed at the scale μ_W, in a perturbative expansion in power of $a_s(\mu_W)$. The perturbative result at NLO results is given by [83]:

$$C_1(\mu_W) = 15a_s + a_s^2\left[(16x+8)\sqrt{4x-1}\mathrm{Cl}_2\left(2\arcsin\frac{1}{2\sqrt{x}}\right) - \left(16x + \frac{20}{3}\right)\ln x - 32x + \frac{7091}{72} + \frac{17}{3}\pi^2\right], \tag{2.56}$$

$$C_2(\mu_W) = 1 + a_s^2\left(\frac{127}{18} + \frac{4}{3}\pi^2\right), \tag{2.57}$$

   where $x = [m_t(M_W)/M_W]^2$ with m_t the top quark mass given in the $\overline{\mathrm{MS}}$ scheme. The Wilson coefficient vector $\vec{C}(\mu_W)$ is then made of $C_1(\mu_W)$ and $C_2(\mu_W)$ for the two first entries and zero for the left-over six entries.

2. Secondly, C_1 and C_2 are run down to the scale μ_b following Eq. (2.46). The evolution matrix $U^{(n_f=5)}(\mu_b, \mu_W)$ needed for this running is calculated following the procedure described in Sec. 2.3.1. The full 8×8 matrices $\gamma^{(0/1/2)}$ in the RGE can be broken down in the following way

$$\gamma^{(i)} \equiv \begin{pmatrix} Q_1^{(i)} & Q_2^{(i)} \\ Q_3^{(i)} & Q_4^{(i)} \end{pmatrix} \tag{2.58}$$

   where $Q_1^{(i)}$ are the 6×6 three-loop anomalous dimension matrices describing the mixing of the four quark operators C_{1-6}, they are taken from [84]. $Q_2^{(i)}$ are the 6×2 matrices describing the mixing of the four quark operators and the dipole operators $C_{7/8}$, they are taken from [83]. $Q_3^{(i)} = 0$ at any order. Finally, the 2×2 matrix from self-mixing in the dipole operator sector is extracted from [85].

3. Then, we perform the matching from the five quarks ($n_f = 5$) to the four quarks ($n_f = 4$) effective field theory. The matching matrix R is different from the unit matrix because the operators $\mathcal{O}^b_{1/2}$ are absent below the b-quark threshold. $C_{3/9}$ receive non-zero contributions only from the matching of the five-flavor effective theory above the scale μ_b to the four-flavor EFT below that scale and from the mixing of $\mathcal{O}_{1/2}$ into $\mathcal{O}_{3/9}$. The R matrix is given by:

$$R = \delta_{ij} + a_s(m_b) R^{(1)}_{ij} + ..., \tag{2.59}$$

where the non-zero elements of $R^{(1)}_{ij}$ are given by

$$\begin{aligned} R^{(1)}_{41} &= -R^{(1)}_{42}/6 = 1/9, \\ R^{(1)}_{71} &= -R^{(1)}_{72}/6 = 8/81, \\ R^{(1)}_{81} &= -R^{(1)}_{82}/6 = -1/54. \end{aligned} \tag{2.60}$$

4. Finally, C_{1-8} are run down to the scale μ_c via the 8×8 evolution matrix $U_2(\mu_c, \mu_b)$. This latter is also computed following the procedure described in Sec. 2.3.1 using this time the full 8×8 anomalous matrices $\gamma^{(0/1/2)}$.

The full procedure can be summarized by the following equation:

$$C(\mu) = U^{(n_f=4)}(\mu, \mu_b)\ R\ U^{(n_f=5)}(\mu_b, \mu_W)\ C(\mu_W) \tag{2.61}$$

Note that C_{10} does not mix under renormalization and thus is zero at all scales to leading order in the $1/M_W$ expansion. Moreover, the framework introduced above does not compute the coefficients $C_{7/8}$ but the renormalization-scheme independent effective ones, defined as:

$$C^{\text{eff}}_{7/8} = C_{7/8} + \sum_{i=1}^{6} y^{(7/8)}_i C_i, \tag{2.62}$$

with $y^{(7)} = Q_d \left(0, 0, 1, \frac{4}{3}, 20, \frac{80}{3}\right)$, $Q_d = -1/3$ and $y^{(8)} = \left(0, 0, 1, -\frac{1}{6}, 20, -\frac{10}{3}\right)$.

2.3.3 Calculation of the C_9 Wilson coefficient

For C_9 we need to solve the following RGE

$$\frac{dC_9(\mu)}{d\ln\mu} = \kappa(\mu) C(\mu). \tag{2.63}$$

with κ the 1×6 matrix that describes the mixing into $\mathcal{O}_9$, κ can be expanded as

$$\kappa = \kappa^{(-1)} + \kappa^{(0)} a_s + ... \tag{2.64}$$

The full procedure for C_9 can be summarized by the following equation:

$$C_9(\mu_c) = C_9(\mu_b) + W^{n_f=4}(\mu_c, \mu_b)\ R\ U^{n_f=5}(\mu_b, \mu_W)\ C(\mu_W) \tag{2.65}$$

where $C(\mu_W)$ is a the vector made of $C_1(\mu_W)$ and $C_2(\mu_W)$ for the two first entries and zero for the left-over six entries and $C_9(\mu_b)$ given by

$$C_9(\mu_b) = -\frac{8}{27}\left[C_1(\mu_b) + \frac{3}{4}C_2(\mu_b)\right] \tag{2.66}$$

The 1×6 matrix $W^{n_f=4}(\mu_c, \mu_b)$ is expressed as:

$$W^{n_f=4}(\mu_c, \mu_b) = -\frac{1}{2}\int_{a_s(\mu_b)}^{a_s(\mu_c)} da_s \frac{\kappa(a_s)}{\beta(a_s)} U^{n_f=4}(\mu_c, \mu_b). \tag{2.67}$$

Here $U^{n_f=4}(\mu_c, \mu_b)$ and R are the 6×6 sub-matrices from the corresponding quantities defined above. The solution to NNLL accuracy is obtained by inserting $U^{n_f=4}(\mu_c, \mu_b)$ to this accuracy. The solution of Eq. (2.67) can be found in appendix C of [35].

Chapter 3

Modeling the resonant structure and phenomenological analysis

As mentioned previously the strategy adopted for B meson decays (i.e. cutting out the resonant region) is not followed for D decays as this would result in most of the available kinematic region being thrown away. Instead, predictions are made in the resonant region, relying on a description of the resonant structure.

Sec. 3.1 is devoted to the resonance description. It is followed by Sec. 3.2 that presents the numerical results and the phenomenological analysis.

3.1 Modeling the quark vacuum polarizations

In the calculation of $D^+ \to \pi^+\ell^+\ell^-$ one needs the quark vacuum polarization functions $\tilde{h}_q(s)$ which intervene in the 1-loop functions $Y^{(q)}$ described in Eq. (2.17) and Eq. (2.18). The kinematical region for $D^+ \to \pi^+\ell^+\ell^-$ decay lies between $4m_l^2$ and $(m_D - m_\pi - 2m_\ell)^2$, numerically $0 \lesssim s \lesssim 2.3$ GeV for the muonic final state. This region is free of charmonia (bound state of a charm quark and a charm anti-quark) as the first charmonium resonance is the J/Ψ whose mass is around 3 GeV, hence $h_c(s)$ can be described by the perturbative QCD result $\tilde{h}^{(\mathrm{pt})}(s, m_c)$ given in Eq. (1.62).

For the light quarks (u,d,s), it is too simplistic to assume that the perturbative QCD result can model the vacuum polarizations as the region of interest for $D^+ \to \pi^+\ell^+\ell^-$ decay contains several light resonances. In the region of these resonances, the partonic description of the closed quark loop, from which both leptons are emitted, breaks down and the decay is dominated by $D^+ \to \pi^+ V(\to \ell^-\ell^+)$ where V is a vector meson $V = \rho^0, \omega^0$ or ϕ. The mass, the quark content and quantum numbers of these resonances are given in Tab. 3.1.
Initially, the resonances were added by hand, by means of Breit-Wigner (BW) functions, on top of a non-resonant background described by the partonic result for the quark vacuum polarization $\tilde{h}^{(\mathrm{pt})}(s, m_q)$ described in Eq. (1.62), e.g. see [37].

In [36] a different approach is advocated, in which the imaginary part of the vacuum polarizations $\tilde{h}_q(s)$ (also called the hadronic spectrum) are modeled following a proposal by M. Shifman [39]. In this model, the imaginary part of each channel is modeled with a dominant vector resonance plus the sum of an infinite tower of resonances, starting from the first excitation, with masses following Regge trajectories.

Resonances	Mass	Quark content	$I(J^{PC})$
ρ^0	775.4 ± 0.34	$\frac{\lvert u\bar{u}\rangle - \lvert d\bar{d}\rangle}{\sqrt{2}}$	$1(1^{--})$
ω^0	782.65 ± 0.12	$\frac{\lvert u\bar{u}\rangle + \lvert d\bar{d}\rangle}{\sqrt{2}}$	$0(1^{--})$
Φ	1019.461 ± 0.020	$\lvert ss\rangle$	$0(1^{--})$

Table 3.1: Properties of the vector meson ρ^0, ω^0 and Φ.

The dominant resonance in each channel is modeled by a Breit-Wigner type function and the sum over the infinite tower of resonances is performed analytically. In this approach, the partonic result is recovered asymptotically.

Then the full function $\tilde{h}(s, m_q)$ is reconstructed from its imaginary part using a dispersion relation as discussed in App. B of the Ref. [36]. We follow their suggestion and use a once-subtracted dispersion relation with the subtraction constant fixed from the perturbative description. Accordingly, the subtraction point is chosen in the deep Euclidean regime at $-s_0$ and the functions $h_q(s)$ are given by

$$\tilde{h}_q(s) \to \tilde{h}^{(\mathrm{pt})}(-s_0, m_q) + \frac{1}{\pi}\int_0^\infty ds' \frac{s_0 + s}{s_0 + s'} \frac{\mathcal{I}m\tilde{h}_q(s')}{s' - s - i\epsilon}. \tag{3.1}$$

The subtraction constant is calculated from the perturbative description:

$$\tilde{h}_q^{(\mathrm{pt})}(s) = -\ln\frac{m_q}{\mu^2} + \frac{2}{3} + \zeta - (2+\zeta)\sqrt{1-\zeta}\ln\left[\frac{1+\sqrt{1-\zeta}}{\sqrt{-\zeta}}\right], \tag{3.2}$$

where m_q is the mass of the quark q and $\zeta = 4m_s/(s + i\epsilon)$. When reconstructing the functions $\tilde{h}_q(s)$ we have checked that the using a dispersion relation with more subtractions leads to very similar results. We have also checked that the results are stable upon variation of the subtraction point in Eq. (3.1). For our final results we use $s_0 = 10$ GeV^2.

In Ref. [36], the isospin 1 (ρ^0) and isospin 0 (ω^0) light resonances are not treated separately, but in terms a single tower of resonances with "effective" parameters. We tried to improve upon the model of [36] by:

- First, treating separately the isospin contribution entering the $\tilde{h}_{d/u}(s)$ functions.
- Second, implementing the strategy of Ref. [40], where the authors relate the imaginary part of the charm hadronic spectrum to the observable

$$R(s) \equiv \frac{3s}{4\pi\alpha_e^2}\sigma_{e^+e^-\to hadrons(\gamma)}(s) = \frac{\sigma_{e^+e^-\to hadrons(\gamma)}(s)}{\sigma_{e^+e^-\to\mu^+\mu^-(\gamma)}(s)}, \tag{3.3}$$

  where α_e is the fine-structure constant. The second equality holds for values of s for which we can neglect the muon mass. The γ in parentheses indicates that hadronic states with final-state radiation are included in addition to purely hadronic states.

This latter point can be performed via the optical theorem. The framework described in Ref. [40] can be employed here too, although the fact that in the low-energy region we have three active quark flavors makes the interplay between the different resonances much more intricate.

From the experimental observable R to the imaginary part of the quark vacuum polarizations

Below charm threshold, experimental data for $R(s)$ contains the entangled contributions from u, d and s quarks. In the case of the isospin 1 combination, additional information exists, since they govern hadronic τ decays and experimental information about the imaginary part are also available [86, 87], which can be exploited in order to disentangle this part from the $I = 0$ and the strange quark contribution.

On the theory side, $R(s)$ can be related via the optical theorem[1] in terms of the imaginary part of the $\bar{q}q$ contributions to the photon vacuum polarization as

$$R(s) = 12\pi\, \mathcal{I}m\Pi(s), \tag{3.4}$$

where the vacuum polarization is defined as

$$\Pi_{\mu\nu}(s) = i\int d^4x\, e^{ix\cdot q}\langle 0|T\{J_\mu^{\rm EM}(x)J_\nu^{\rm EM}(0)\}|0\rangle = (q_\mu q_\nu - g_{\mu\nu}s)\Pi(s), \tag{3.5}$$

with $J_\mu^{\rm EM}$ is the electromagnetic current given by

$$J_\mu^{\rm EM} = Q_u(\bar{u}\gamma_\mu u) + Q_d(\bar{d}\gamma_\mu d) + Q_s(\bar{s}\gamma_\mu s) \tag{3.6}$$

(where we omitted the sum over color indices). Below charm threshold, we then write the R ratio as the sum of the three quark-flavor contributions $R = \sum_q R_q = R_u + R_d + R_s$. We then define the correlators $\Pi^{(q)}(s)$ which are formed from the flavor-singlet currents as

$$\Pi_{\mu\nu}^{(q)}(s) = i\int d^4x\, e^{ix\cdot q}\langle 0|T\{(\bar{q}\gamma_\mu q)(x)(\bar{q}\gamma_\nu q)(0)\}|0\rangle = (q_\mu q_\nu - g_{\mu\nu}s)\Pi^{(q)}(s). \tag{3.7}$$

We define:

$$\tilde{h}_q(s) = \frac{12\pi^2}{N_c}\Pi^{(q)}(s). \tag{3.8}$$

At sufficiently high momenta in the deep Euclidean region, $s \to -\infty$, we have, after renormalization in the $\overline{\rm MS}$ scheme,

$$\tilde{h}_q(s) \to -\log\left(-\frac{s}{\mu^2}\right) + \frac{2}{3}. \tag{3.9}$$

Upon analytic continuation one obtains, for $s > 0$, in the chiral limit, $\mathcal{I}m\,\tilde{h}_q(s) = \pi$. Within our assumptions we can cast the result for the R ratio as

$$R_{uds} = N_c \sum_{q=u,d,s} Q_s\, \frac{\mathcal{I}m\tilde{h}_q(s)}{\pi} = 2\left(1 + \frac{\alpha_s}{\pi} + \cdots\right)\left(1 + \mathcal{O}\left(\frac{m^4}{s^2}\right)\right), \tag{3.10}$$

where we show the leading perturbative correction on the right-hand side.

For the purpose of modeling the resonances it is appropriate to consider the different isospin channels separately. The electromagnetic current of Eq. (3.6) contains an $I = 1$ and $I = 0$ light-quark part, as well as the strange quark contributions, the respective dominant vector resonances being the $\rho(770)$, the $\omega(782)$, and the $\phi(1020)$.

[1] A good introduction to the optical theorem can be seen in [59].

Taking into account the quark-charge factors the electromagnetic current can be written as

$$J_\mu^{\rm EM} = \frac{1}{2}\left(\bar{u}\gamma^\mu u - \bar{d}\gamma^\mu d\right) + \frac{1}{6}\left(\bar{u}\gamma^\mu u + \bar{d}\gamma^\mu d\right) - \frac{1}{3}\bar{s}\gamma_\mu s, \tag{3.11}$$

where we made explicit the $I = 1$ and $I = 0$ light-quark contributions in the first and second terms in between parenthesis on the right-hand side, respectively, where $J^\mu_{1/0} = 1/\sqrt{2}(\bar{u}\gamma^\mu u \mp \bar{d}\gamma^\mu d)$. For the R ratio one has then[2]

$$R_{uds} = N_c\left(\frac{1}{2}\frac{\mathcal{I}m\tilde{h}_{I=1}(s)}{\pi} + \frac{1}{18}\frac{\mathcal{I}m\tilde{h}_{I=0}(s)}{\pi} + \frac{1}{9}\frac{\mathcal{I}m\tilde{h}_s(s)}{\pi}\right), \tag{3.12}$$

where $\tilde{h}_{1/0}(s)$ represent the light-quark $I = 1$ and $I = 0$ contributions and $\tilde{h}_s(s)$ is the contribution from the strange quark. The imaginary part of these functions are proportional to respective hadronic spectral functions. In the particular case of $I = 1$, experimental data for the spectral function exists from hadronic τ decays [86, 87, 90]. Ultimately, in the application to $D^+ \to \pi^+\mu^+\mu^-$ we need the functions $\tilde{h}_u(s)$, $\tilde{h}_d(s)$ and $\tilde{h}_s(s)$. The two first ones can be obtained from $\tilde{h}_{I=1/0}(s)$, which contribute equally to the u- and d-quark vacuum polarization:

$$\mathcal{I}m\tilde{h}_{d/u}(s) \simeq \frac{\mathcal{I}m\tilde{h}_{I=0}(s) + \mathcal{I}m\tilde{h}_{I=1}(s)}{2} \tag{3.13}$$

Hadronic spectrum model

We need a concrete model for the imaginary part of the different $\tilde{h}_I(s)$ functions and $\tilde{h}_s(s)$. Our description is based on the model suggested in App. B of Ref. [36] which, in turn, is based on a proposal by Shifman [39] to which we refer for more details about the model. The model can be summarized as follows. The imaginary part of each channel is modeled with a dominant vector resonance plus the sum of an infinite tower of resonances, starting from the first excitation, with masses following Regge trajectories. The dominant resonance in each channel is modeled by a Breit-Wigner type function, $f^{BW}(s)$, and the sum over the infinite tower of resonances is performed analytically within Shifman's model [39]. Concretely, we have for the $I = 0$ and $I = 1$ contributions:

$$\mathcal{I}m\tilde{h}_I(s) = \mathcal{I}m f_I^{BW}(s) - \mathcal{I}m\left[\frac{\Psi(z_I + a_I)}{1 - b_I/\pi}\right]. \tag{3.14}$$

The first term corresponds to the dominant vector resonance and is described by a Breit-Wigner-type function. We adopt the form as suggested in [36], where we add a phase φ_I:

$$f_I^{BW}(s) = n_I e^{i\varphi_I}\frac{1}{1 + \hat{\sigma}_I^2 z_I}, \tag{3.15}$$

with

$$\hat{\sigma}_I = \frac{\sigma_I}{m_I}, \quad z_I = \left(\frac{-s - i\epsilon}{\sigma_I^2}\right)^{1-b_I/\pi} \quad \text{and} \quad b_I = \frac{\Gamma_I}{m_I}, \tag{3.16}$$

where m_I and Γ_I are the resonance mass and width respectively and n_I is a normalization factor. In our case, the resonances are $\rho(770)$ for $I = 1$ and $\omega(782)$ for $I = 0$.

[2] For related discussions see Ref. [88, 89]

The imaginary part of this modified BW ansatz can be written as

$$\begin{aligned}\mathcal{I}m f_I^{\rm BW} =& \frac{\sin\varphi_I + |z_I|\hat{\sigma}_I^2 \sin(b_I - \varphi_I)}{1 - 2|z_I|\hat{\sigma}_I^2 \cos b_I + |z_I|^2 \hat{\sigma}_I^4} \\ \simeq& \frac{\sin\varphi_I m_I^2(m_I^2 - s) + \theta(s) s\Gamma_I m_I \cos\varphi_I}{(s - m_I^2)^2 + s\Gamma_I^2}\end{aligned} \tag{3.17}$$

where $\theta(s)$ is the Heaviside function. In the last line we have taken $\hat{\sigma}^2 \simeq \sigma_I^2/m_I^2 \simeq 1$ for simplicity and used $b_I \ll 1$. Note that the previous equation is equivalent to Eq. (B.3) of the Ref. [36] if $\varphi_I = 0$.

The second term represents the sum over the infinite tower of equally spaced resonances with masses $m_I^2(n) = (n + a_I)\sigma_I^2$ and widths $\Gamma_I(n) = b_I\, m_I(n)$ and $\Psi(z)$ is the digamma function.

For the strange quark, we use an expression similar to Eq. (3.14):

$$\mathcal{I}m\tilde{h}_s(s) = \mathcal{I}m f_s^{BW}(s) - \mathcal{I}m\left[\frac{\Psi(z_s + a_s)}{1 - b_s/\pi}\right]. \tag{3.18}$$

where $\mathcal{I}m f_s^{BW}(s)$ and z_s are slightly different from $\mathcal{I}m f_I^{BW}(s)$ and z_I since they take into account the finite Kaon mass m_K. For z_s we have:

$$z_s = \left(\frac{4\, m_K^2 - s - i\epsilon}{\sigma_s^2}\right)^{1-b_s/\pi}. \tag{3.19}$$

For $\mathcal{I}m f_s^{\rm BW}(s)$ we use the expression given in Eq. (B.6) from [36]:

$$\mathcal{I}m f_{\rm BW}^{(\phi)}(s) = n_\phi \frac{(s - 4m_K^2)\theta(s - 4m_K^2)}{M_\phi^2 - 4m_K^2} \frac{m_\phi^3 \Gamma_\phi}{(s - M_\phi^2)^2 + \frac{s - 4m_K^2}{M_\phi^2 - 4m_K^2} M_\phi^2 \Gamma_\phi^2}. \tag{3.20}$$

Note that we have treated the light quarks and the pions as massless. The use of these Breit-Wigner-like functions, which are slightly different from the more widely employed forms [91], means that the resonance parameters will be shifted when compared with some of the values in the literature. One should, however, keep in mind that Breit-Wigner parameters are not physical, as opposed to the pole positions, and these differences are no cause for alarm.

Adjustment of parameters on data

In Ref. [36] the parameters of the model were fixed based only on the resonance masses and widths. Furthermore, no separation between the $I = 0$ and $I = 1$ components was introduced. Here, our purpose is to improve upon the description of Ref. [36] by fixing the parameters of our model from a comparison to $e^+e^- \to$ (hadrons) and $\tau \to$ (hadrons) $+ \nu_\tau$ data.

It is known that models related to the one we are employing here agree well with the data in an asymptotic region [88, 90, 92], where s is large enough (in practice this means $s \gtrsim 1.5$ GeV^2). We are being more ambitious here, because we are attempting to the describe the spectral functions in the whole kinematic range, including the

leading resonances and the intermediate and low-energy regions. The simplicity and elegance of the model we use for the spectral functions, Eq. (3.14), is very appealing but limits our ability to describe all the features of the data. When fixing the parameters we will inevitably have discrepancies since the model is not equally suitable for the whole kinematic range. We can expect, therefore, that our model, although improved with respect to the one from Ref. [36], will still not be enough for a precise description of the experimental data available.

Here we use the publicly available Particle Data Group compilation of $R(s)$ data [91] supplemented with $R(s)$ measurements from the BES and KEDR collaborations published recently in Refs. [93–95]. We also use the ALEPH [87] and OPAL [86,90] data for the vector isovector spectral function from $\tau \to (\text{hadrons}) + \nu_\tau$.

The imaginary parts of the functions $\tilde{h}_I(s)$ are then fixed as to provide a reasonable description of the data, while keeping the parameters of the model within physically acceptable values. For simplicity, the masses of the Regge tower of resonances in $I = 0$ and $I = 1$ channels will be described by the same parameters and the phases φ_I were set to zero. The final parameters are shown in Tab. 3.2. The result for $R(s)$ within our model is shown in Fig. 3.1 where we also give the results of the model of [36] for comparison.

Our description agrees relatively well with the data in the resonant region. At higher energies, it is systematically below data due to the lack of perturbative corrections, which are of the order of 15%. We also show the comparison of our description and τ decay data in Fig. 3.2. In both cases, our spectral functions provide reasonable descriptions of the data while maintaining the model rather simple. For the purposes of the present work, this should be sufficient.

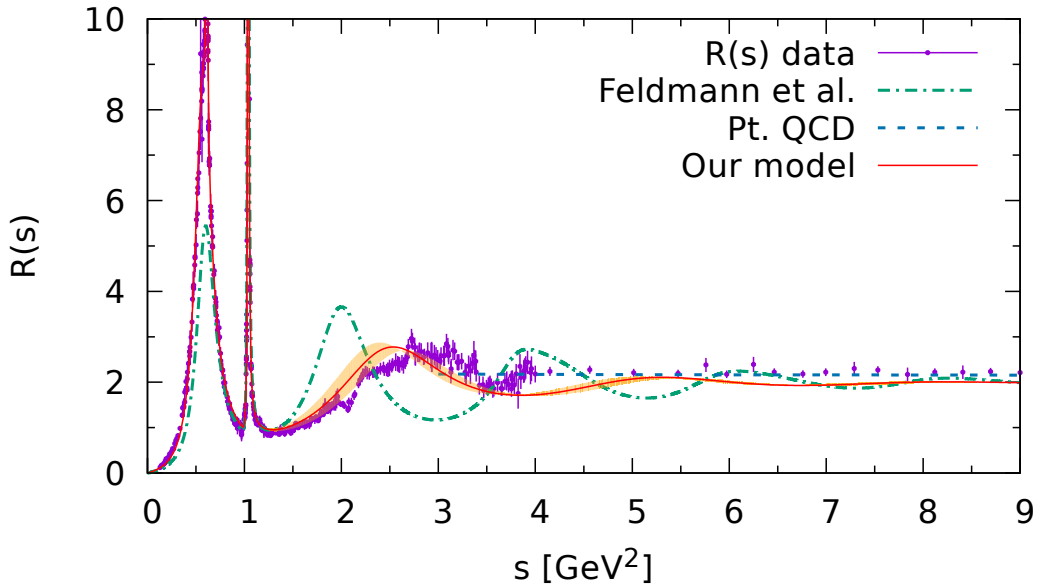


Figure 3.1: $R(s)$ described by Eq. (3.12) with parameters given in Tab. 3.2 (solid line). Results from the model of Ref. [36] are also shown for comparison (dashed line). The perturbative QCD result at four loops with $N_f = 3$ in the chiral limit is also shown.

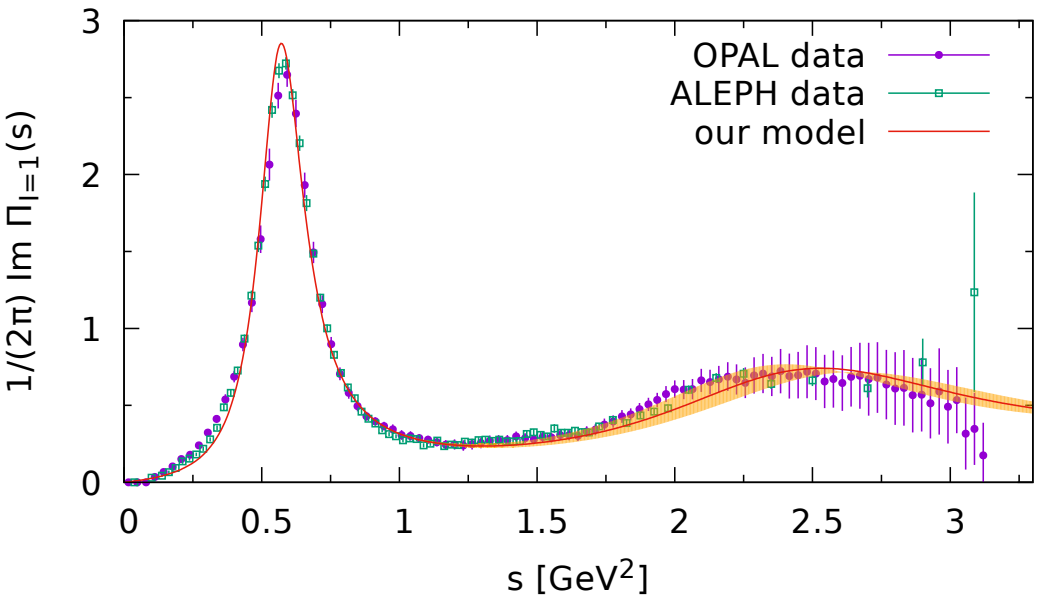


Figure 3.2: Vector isovector spectral function from hadronic tau decay data [86, 87, 90] compared with our model with parameters given in Tab. 3.2.

3.2 Numerical and phenomenological analysis

We first present the parametrization of the form factors and LCDA's used in the thesis as well as the numerical inputs in Sec. 3.2.1. It is followed by the prediction for the full decay width in Sec. 3.2.2. Experimental constraints on the BSM Wilson coefficients are developed in Sec. 3.2.3. Finally, a phenomenological study of the $D^+ \to \pi^+\mu^+\mu^-$ decay is presented in Sec. 3.2.4.

Parameter	Central value	Uncertainty
n_ρ	3.12	–
m_ρ	0.756 GeV	–
Γ_ρ	0.132 GeV	–
$b_{I=1}$	0.26	2.3%
$\sigma^2_{I=1}$	2.39 GeV^2	3.5%
$a_{I=1}$	1.05	3.5%
n_ω	2.58	–
m_ω	0.782 GeV	–
Γ_ω	0.0091 GeV	–
$b_{I=0}$	0.26	2.3%
$\sigma^2_{I=0} = \sigma^2_{I=1}$	2.39 GeV^2	3.5%
$a_{I=0} = a_{I=1}$	1.08	3.5%
n_ϕ	1.86	–
m_ϕ	1.019 GeV	–
Γ_ϕ	0.0042 GeV	–
σ^2_s	2.0 GeV^2	3%
a_s	0.71	5%
b_s	0.93	6%

Table 3.2: Parameters for the spectral functions

3.2.1 Parametrization and numerical inputs

Parametrization of the form factors

As stated previously, there are only three independent functions in the D to pseudoscalar transition, the scalar f_0, the vector f_+ and the tensor f_T form-factors. They are defined in Eq. (1.57) and Eq. (1.58) and can be expressed following a z-expansion with z given by

$$z = \frac{\sqrt{t_+ - s} - \sqrt{t_+ - t_0}}{\sqrt{t_+ - s} + \sqrt{t_+ - t_0}}, \tag{3.21}$$

where t_+ and t_0 are:

$$\begin{aligned} t_+ &= (m_D = m_\pi)^2, \\ t_0 &= (m_D + m_\pi)(\sqrt{m_D} - \sqrt{m_\pi})^2. \end{aligned} \tag{3.22}$$

The vector and scalar $f_+^{D\to\pi}$ and $f_0^{D\to\pi}$ form factors are taken from [96]:

$$f_+(s) = \frac{f^{D\to\pi}(0) + c_+(z - z_0)(1 + \frac{z+z_0}{2})}{1 - P_V s}, \tag{3.23}$$

$$f_0(s) = \frac{f^{D\to\pi}(0) + c_0(z - z_0)(1 + \frac{z+z_0}{2})}{1 - P_S s}, \tag{3.24}$$

where $z_0 \equiv z(s = 0)$. The tensor form factor $f_T^{D\to\pi}$ is taken from [97] and is given by a similar expression

$$f_T(s) = \frac{f_T^{D\to\pi}(0) + c_T(z - z_0)(1 + \frac{z+z_0}{2})}{1 - P_T s}. \tag{3.25}$$

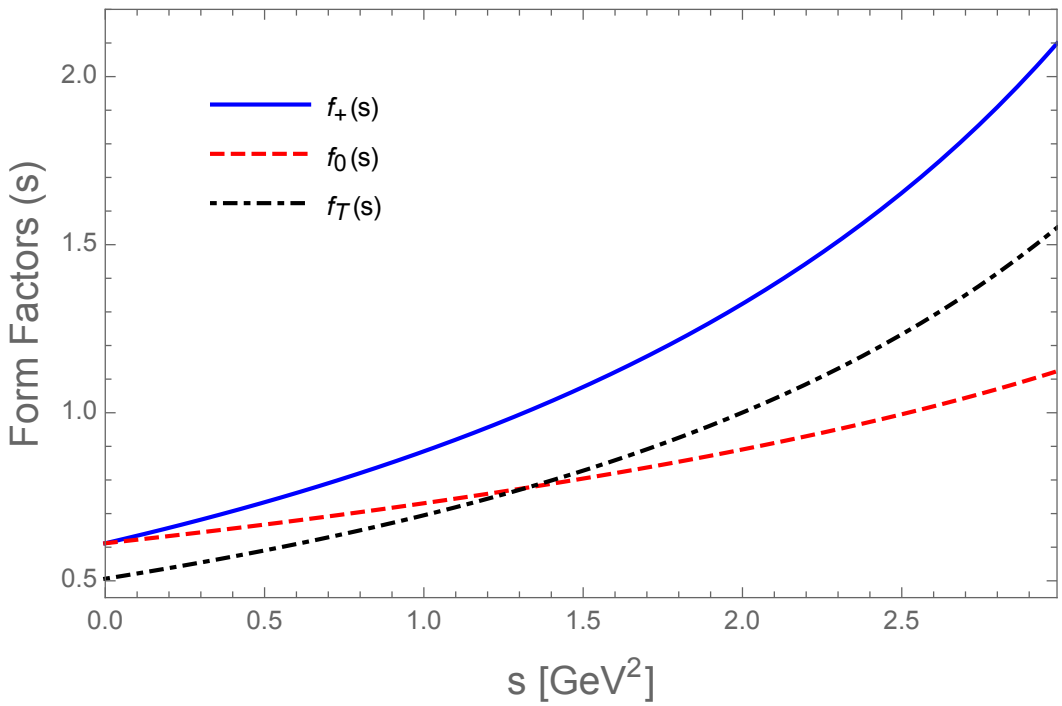


Figure 3.3: Distribution of the three $D \to \pi$ form factors used in this analysis.

Note that, at zero momentum transfer, the additional relation $f_+(0) = f_0(0)$ holds. The value of the parameters $f^{D\to\pi}(0), f_T^{D\to\pi}(0), c_+, c_0, c_T, P_V, P_S, P_T$ are given in Tab. 3.4. The distributions for the three form-factors are shown in Fig. 3.3. They are found to be in good agreement with the ones obtained by LCSR calculation in Ref. [98].

Parametrization of the π and D meson LCDA

The parametrization for the π meson light cone distribution amplitude ϕ_π, defined in Eq. (1.61), is taken from [99]. In this reference, it is expanded in a series of Gegenbauer polynomials $C_n^\alpha(z)$:

$$\phi_\pi(u, \mu^2) = 6u(1-u)\sum_{n=0}^{\infty} a_n(\mu^2) C_n^{3/2}(2u-1) \tag{3.26}$$

The $a_n(\mu^2)$ are the Gegenbauer moments, a_0 is 1 from the normalization condition and the odd moments are equal to zero in case of the pions. The values of these moments at the scale $\mu = 1$ GeV are taken from [99]:

$$a_0 = 1, \quad a_2(1\text{ GeV}) = 0.17 \pm 0.08, \quad a_4(1\text{ GeV}) = 0.06 \pm 0.10 \tag{3.27}$$

We need to obtain these moments at the scale value μ_c. This can be achieved by solving the following equation [100]

$$c_i(\mu) = L^{\gamma_{c_i}/\beta_0} c_i(1\text{ GeV}), \tag{3.28}$$

where $L = \alpha_s(\mu)/\alpha_s(1\text{ GeV})$ and γ_{c_i} is given by

$$\gamma_{c_i} = C_F\left(1 - \frac{2}{(n+1)(n+2)} - \sum_{m=2}^{n+1}\frac{1}{m}\right) \tag{3.29}$$

	C_1	C_2	C_3	C_4	C_5	C_6	$C_7^{\rm eff}$	$C_8^{\rm eff}$	C_9
LL	-0.890	1.072	-0.002	-0.041	0.000	0.000	0.057	-0.042	-0.095
NLL	-0.603	1.029	-0.003	-0.065	0.000	0.000	0.035	-0.045	-0.270
NNLL	-0.529	1.026	-0.004	-0.063	0.000	0.000	0.036	-0.048	-0.413

Table 3.3: Value of the Wilson coefficients at the NNLL approximation given at scale $\mu_c = 1.67$ GeV, $\mu_b = 4.18$ GeV, $\mu_W = 80.4$ GeV.

Using the one-loop running of α_s to be consistent with the fact that we only have the leading anomalous dimension (and using $n_f = 3$) we find for $\mu_c = m_c^{\rm pole}$:

$$a_0 = 1, \quad a_2(\mu_c) = 0.8450\, a_2(1\ \text{GeV}), \quad a_4(\mu_c) = 0.7825\, a_4(1\ \text{GeV}). \tag{3.30}$$

The D meson LCDA, defined in Eq. (1.60) is given by a simple exponential model as in [101]:

$$\phi_D^-(\omega) = \frac{1}{\omega_0} e^{-\omega/\omega_0}, \qquad \phi_D^+(\omega) = \frac{1}{\omega_0^2} e^{-\omega/\omega_0} \tag{3.31}$$

where $\omega_0 = 450 \pm 300$ MeV.

Numerical inputs

The numerical inputs used in our work are given in Tab. 3.4. The value used for the scale μ_c is the pole mass of the quark c. The scales μ_c, μ_b, μ_W vary between $\mu/\sqrt{2}$ and $\mu\sqrt{2}$.

The values of the Wilson coefficients C_{1-9} depend on the renormalization scale μ_c, μ_b and μ_W used in the procedure described in Sec. 2.3. The numerical results for the Wilson coefficients, for the central value of these scale (as given in Tab. 3.4), are summarized in Tab. 3.3. We note that only C_1, C_2 and C_9 have sizable values and the coefficients related to the strong penguin and the dipole operators are numerically small.

Parameters	Value
μ_c [GeV]	$1.67^{+0.69}_{-0.49}$
μ_b [GeV]	$4.18^{+1.73}_{-1.23}$
μ_W [GeV]	$80.4^{+33.3}_{-23.6}$
m_s [MeV]	95 ± 3
m_c [MeV]	1.67 ± 0.07
$\omega_0(\phi_D)$[MeV]	450 ± 300
f_{π^+} [MeV]	130.5 ± 16
f_{D^+} [MeV]	212.15 ± 1.45
$a_2(\phi_\pi)$	0.143 ± 0.07
$a_4(\phi_\pi)$	0.046 ± 0.08
$f_{(0)}$ (f_+ and f_0)	0.6117 ± 0.0354
c_+ (f_+)	-1.985 ± 0.347
c_0 (f_0)	-1.188 ± 0.256
P_V (f_+)	0.1314 ± 0.0127
P_S (f_0)	0.0342 ± 0.0122
$f_T(0)$ (f_T)	0.5063 ± 0.0786
C_T (f_T)	-1.10 ± 1.03
P_T (f_T)	0.1461 ± 0.0681
τ_{D^+} [ps]	1040 ± 7
λ_d	0.208 ± 0.009
λ_b $\times 10^5$	$(4.8 \pm 2) - (15.9 \pm 1.7)i$
b_ω	0.26 ± 0.023
σ_ω	2.39 ± 0.035
a_ω	1.08 ± 0.035
a_Φ	0.71 ± 0.05
b_Φ	0.93 ± 0.06

Table 3.4: Summary of the numerical inputs used in the study of $D^+ \to \pi^+\ell^+\ell^-$.

3.2.2 Full differential decay width distribution

Following [38, 102], the double differential $D^+ \to \pi^+\ell^+\ell^-$ decay width with respect to s the dilepton invariant mass and θ, the angle between the three-momenta of D^+ and ℓ^- in the rest frame of lepton pair ($\cos\theta$ vary between -1 and +1), is given by

$$\frac{d^2\Gamma(D^+ \to \pi^+\ell^+\ell^-)}{ds\ d\cos\theta} = N\lambda^{1/2}(s)\beta(s)[a_\ell(s) + b_\ell(s)\cos\theta + c_\ell(s)\cos^2\theta] \tag{3.32}$$

where the kinematic function $\beta(s)$ and $\lambda(s)$ were already given in Eq. (2.11). The normalization factor N is given by:

$$N = \frac{G_F^2\alpha_e^2}{(4\pi)^5 m_D^3} \tag{3.33}$$

and the three angular coefficients $a_\ell(s)$, $b_\ell(s)$ and $c_\ell(s)$ by:

$$a_\ell(s) = \frac{\lambda}{2}(|F_V|^2 + |F_A|^2) + 8m_\ell^2 m_D^2 |F_A|^2 + 2s(\beta^2|F_S|^2 + |F_P|^2), \tag{3.34}$$

$$b_\ell(s) = 4Re\left[s(\beta^2 F_S F_T^* + F_P F_{T5}^*) + m_\ell(\lambda^{1/2}\beta F_V F_S^* + (m_D^2 - m_\pi^2 + s)F_A F_{T5}^*)\right], \tag{3.35}$$

$$c_\ell(s) = -\frac{\lambda\beta^2}{2}(|F_V|^2 + |F_A|^2) + 2s(\beta^2|F_T|^2 + |F_{T5}|^2) + 4m_\ell\beta\lambda^{1/2}Re[F_V F_T^*]. \tag{3.36}$$

If we integrate over $\cos\theta$, the decay width distribution becomes:

$$\frac{d\Gamma(D^+ \to \pi^+\ell^+\ell^-)}{ds} = 2N\lambda^{1/2}\beta\left[a_\ell(s) + \frac{c_\ell(s)}{3}\right] \equiv N\lambda^{1/2}(s)\beta(s)d\hat{\Gamma}(s) \tag{3.37}$$

It is convenient to define $d\hat{\Gamma}(s)$ since this enters in the definition of several observables. In terms of Wilson coefficients and form factors, $d\hat{\Gamma}(s)$ can be written

$$\begin{aligned} d\hat{\Gamma}(s) =& \frac{2}{3}\lambda(|F_V^{\rm SM}|^2 + |C_{10}|^2 f_+^2) + (m_D^2 - m_\pi^2)\frac{s}{m_c^2} f_0^2(|C_P|^2 + |C_S|^2) \\ &+ \frac{16}{3}\lambda s \frac{f_T^2(|C_T|^2 + |C_{T5}|^2)}{(m_D + m_\pi)^2}. \end{aligned} \tag{3.38}$$

Breakdown of the different SM contributions

In the SM, the following relations hold approximately

$$b_\ell(s) = 0, \tag{3.39}$$

$$a_\ell(s) = -\frac{c_\ell(s)}{\beta^2(s)} = \frac{\lambda}{2}|F_V^{\rm SM}(s)|^2, \tag{3.40}$$

thus the full differential decay rate can be written

$$\frac{d^2\Gamma(D^+ \to \pi^+\ell^+\ell^-)}{ds\ d\cos\theta} = N\lambda^{1/2}\beta a_\ell\left(1 - \beta^2(s)\cos^2\theta\right). \tag{3.41}$$

Integrating over $\cos\theta$ we have

$$\frac{d\Gamma(D^+ \to \pi^+\ell^+\ell^-)}{ds} = N\lambda^{3/2}\beta\left(1 - \frac{\beta^2(s)}{3}\right)|F_V^{\rm SM}(s)|^2. \tag{3.42}$$

This last expression is a multiplication of simple purely kinematic functions (λ and β) and a more complicated function $F_V^{\rm SM}(s)$. In the following, we examine the different parts $F_V^{\rm SM}(s)$ in more in details. We remind the reader that $F_V^{\rm SM}(s)$ is given by

$$F_V^{\rm SM}(s) = \left[\lambda_b C_9^{(b)}(s) + \lambda_d C_9^{(d)}(s)\right] f_+(s), \tag{3.43}$$

where

$$C_9^{(q)} = C_9^{(q)}\big|_{\rm Naive} + C_9^{(q)}\big|_{\rm Ann} + C_9^{(q)}\big|_{\rm SS} + C_9^{(q)}\big|_{\rm FF}. \tag{3.44}$$

The naive results for $C_9^{(q)}$ are given by

$$C_9^{(d)}(s)\big|_{\rm Naive} = Y^{(d)}(s), \tag{3.45}$$

$$C_9^{(b)}(s)\big|_{\rm Naive} = Y^{(b)}(s) + C_9 + \frac{2m_c}{m_D + m_\pi}C_7\frac{f_T(s)}{f_+(s)}, \tag{3.46}$$

Contribution	$\propto \lambda_b$	$\propto \lambda_d$
C_9	-0.413	0
$Y^{(q)}$	$-1.303 + 0.034i$	$1.345 + 0.981i$
$C_9^{(q)}\big\vert_{\rm FF}$	$-0.287 - 0.457i$	$-0.028 - 0.002i$
$C_9^{(q)}\big\vert_{\rm Ann}$	$0.013 - 0.054i$	$0.503 - 2.100i$
$C_9^{(q)}\big\vert_{\rm SS}$	$0.028 - 0.033i$	$0.005 + 0.002i$

Table 3.5: Breakdown of individual contribution at NLO for the $D^+ \to \pi^+\ell^+\ell^-$ decay at $s = 0.5$ GeV^2.

and the expressions for the QCDF corrections $C_9^{(q)}\big\vert_{\rm Ann}$, $C_9^{(q)}\big\vert_{\rm SS}$ and $C_9^{(q)}\big\vert_{\rm FF}$ are given in Sec. 2.2. The numerical results for these different contributions at $s = 0.5$ GeV^2 are presented in Tab. 3.5. We note that $Y^{(q)}$ (that contains the resonances) and $C_9^{(q)}\big\vert_{\rm Ann}$ are the only sizable contributions. Moreover, we remind the reader that the $(q = b)$ contributions are CKM suppressed since $\lambda_b \ll \lambda_d$. Hence, $F_V^{\rm SM}(s)$ can be approximated by

$$F_V^{\rm SM}(s) \sim \lambda_d \left(Y^{(d)} + C_9^{(d)}\big\vert_{\rm Ann} \right) f_+(s). \tag{3.47}$$

As it was first noticed in Ref. [37], $Y^{(d)}$ is suppressed by cancellations between C_1 and C_2 occurring at the scale m_c in the factor $(2/3C_1 + 1/2C_2) \sim 0.34$.

Annihilation uncertainties

As the contribution due to the annihilation topology $C_9^{(d)}\big\vert_{\rm Ann}$ are large (see 3.5), it is important to have a correct estimate of the related uncertainty. The QCDF correction due to the annihilation topology is given in Eq. (2.27) which we provide here as a reminder,

$$C_9^{(d)}(s)\big\vert_{\rm Ann} = 8e_d \frac{\pi^2}{N_c} \frac{f_D f_\pi}{m_D} \frac{1}{f_+(s)} \frac{1}{\lambda_D^-(s)}\ 3C_2 \tag{3.48}$$

It depends, among other parameters, on the moment $\lambda_D^-(s)$ of the D meson LCDA defined by:

$$\frac{1}{\lambda_D^-(s)} = \int_0^\infty \frac{ds'}{s' - s - i\epsilon} \phi_D^-(s'/m_D). \tag{3.49}$$

The authors of Ref. [36] wished to include the possible non-perturbative hadronic effects near $s = m_{\rho/\omega}^2$ which would modify the annihilation result. In order to model this, they assumed that the s spectrum factorizes into the D meson LCDA and a hadronic model for the spectrum associated to the light vector current, resulting in a modified moment:

$$\frac{1}{\lambda_D^-(s)} \to \int_0^\infty \frac{ds'}{s' - s - i\epsilon} \phi_D^-(s'/m_D)\ \mathcal{I}m\tilde{h}_q(s). \tag{3.50}$$

Note that the hadronic spectrum $\mathcal{I}m\tilde{h}_q(s)$ is denoted $j_q(s)$ in [36]. The parameters in the function $\mathcal{I}m\tilde{h}_q(s)$ are adjusted to reproduce the perturbative result in the limit $s \ll -m_D\Lambda$. More concretely, the value of the parameter n_V (this parameter factorizes the BW component of their function $j_{u/d}$) is tuned by imposing:

$$\int_0^\infty ds\ \phi_D^-(s/m_D)\ (j_q(s) - 1) \stackrel{!}{=} 0. \tag{3.51}$$

s-bin	90% C.L. Limit	95% C.L. Limit
Reg. I: $s \in [0.250^2, 0.525^2]$ GeV^2	2.0×10^{-8}	2.5×10^{-8}
Reg. II: $s > 1.25^2$ GeV^2	2.6×10^{-8}	2.9×10^{-8}

Table 3.6: Current best-world limits on $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$ in two bins of the dilepton invariant mass s [42]. The total branching fraction excluding the resonant contributions is extrapolated assuming a phase space model.

Using the exponential model for the D meson LCDA given in Eq. (3.31), $n_V = 2.40$ is obtained.

We don't follow this strategy (that consist in taking the hadronically modified moment to calculate the central value of their prediction). Nevertheless we agree that the physical s spectrum might be influenced by the light-vector resonances and we decided to include this possible hadronic modification of the moment in our uncertainties.

We then need to solve Eq. (3.51) for our hadronic spectrum $\mathcal{I}m\tilde{h}_d(s)$, which depends on the ρ and ω contributions, as explained in Sec. 3.1. These two BW's are factorized by n_ρ and n_ω. Imposing $n_\rho = n_\omega$, solving Eq. (3.51) leads to:

$$n_\rho = n_\omega = 2.48. \tag{3.52}$$

Note that the value is similar to that obtained for the function $j_{u/d}$ in Appendix B of Ref. [36] ($n_V = 2.40$). Finally, we use this hadronic modified moment to find the resulting error on the annihilation uncertainty. In addition, we extend the uncertainty on ω_0 from 150 MeV (used in [36]) to 300 MeV.

Phases of the resonances

The branching ratio distribution is related to the decay width distribution defined in Eq. (3.37) via

$$\frac{d\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)}{ds} = \frac{1}{\Gamma_D}\frac{d\Gamma(D^+ \to \pi^+\ell^+\ell^-)}{ds} \tag{3.53}$$

where Γ_D is the total decay width of the D^0 meson. The distribution of the $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$ for the SM is given in Fig. 3.4. The green line is the distribution where both phases φ_I are set to zero. The blue (red) lines show various distributions where $\varphi_\rho > 0$ ($\varphi_\rho < 0$). The error band is in yellow, it includes all parameter uncertainties reported in Tab. 3.4 as well as the error on the annihilation contributions as explained in Sec. 3.2.2. The experimentally excluded regions at 90% C.L. are in gray. The definition of the regions and the limits [42] are summarized in Tab. 3.6. The region I (Reg. I) is on the left (low s value) and the region II (Reg. II) is on the right (high s value) of the plot. A similar plot for various scenarios where $\varphi_\omega \neq 0$ is shown in Fig. 3.5.

We note that only scenarios where $\varphi_\omega > 0$ lie outside of the error band; the maximal discrepancy with respect to the central value (in green) is given for $\varphi_\omega \sim 7\pi/10$.

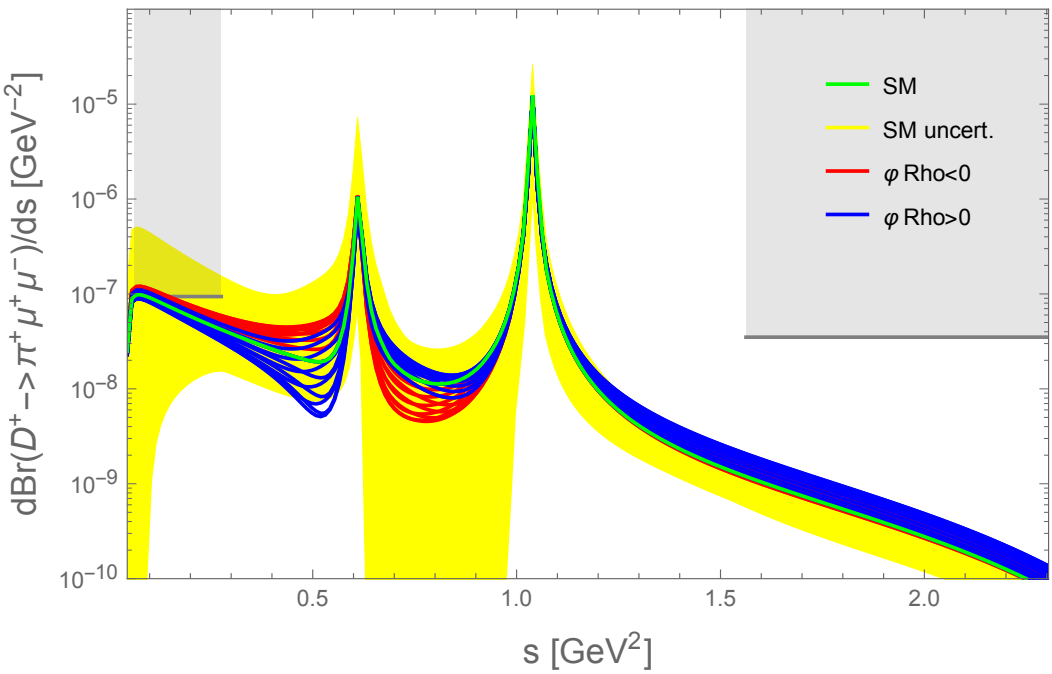


Figure 3.4: Distribution of the branching ratio in SM with both phases φ_I set to zero in green and the related uncertainty band in yellow. The same distribution is plotted on top for several negative (in red) and positive (in blue) value of the ρ^0 phase. The ω^0 phase being set to zero. The gray regions represent the experimental exclusions at 90% C.L..

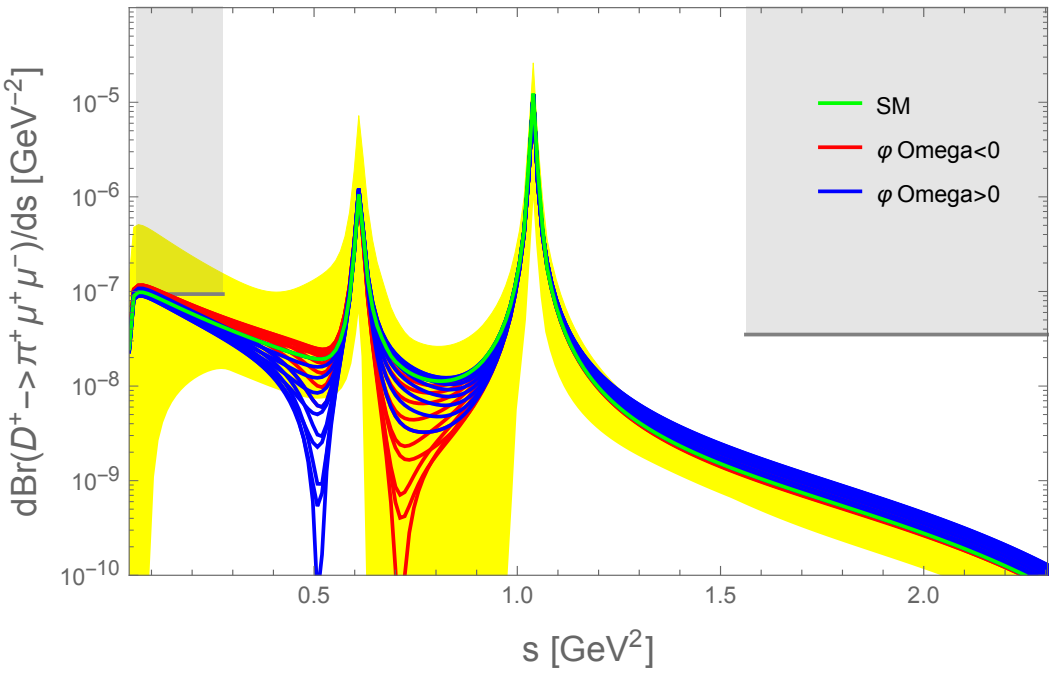


Figure 3.5: Distribution of the branching ratio in SM with both phases φ_I set to zero in green and the related uncertainty band in yellow. The same distribution is plotted on top for several negative (in red) and positive (in blue) value of the ω^0 phase. The ρ^0 phase being set to zero. The gray regions represent the experimental exclusions at 90% C.L..

3.2.3 The BSM Wilson coefficients constraints

In this section, we interpret the experimental limits on the branching ratio of the $D^0 \to \ell^+\ell^-$ and $D^+ \to \pi^+\ell^+\ell^-$ decay as constraints on BSM Wilson coefficients. We then compare these constraints with the possible size of the Wilson coefficients in some BSM models. Finally, some distributions of the branching ratio are shown for different model-independent scenarios.

Constraints from the $D^0 \to \ell^+\ell^-$ decay measurement

The $D^0 \to \ell^+\ell^-$ decay can helps us to set experimental constraints on the BSM Wilson coefficients C_{10}, C_S and C_P. A generic expression for the branching ratio of the $M \to \ell^+\ell^-$ decay (M being a $B^0_{(s)}$ or a D^0 meson) has been given in Eq. (1.44). For convenience, it is provided here for the $D^0 \to \ell^+\ell^-$ decay:

$$\mathcal{B}(D^0 \to \ell^+\ell^-) = \frac{1}{\Gamma_D}\frac{G_F^2\alpha_e^2}{64\pi^3} f_D^2 m_D^5 \beta(m_D^2)\left(|P|^2 + \beta^2(m_D^2)|S|^2\right), \tag{3.54}$$

where

$$P = \frac{1}{m_c}(C_P - C'_P) + \frac{2m_\ell}{m_D^2}(C_{10} - C'_{10}), \tag{3.55}$$

$$S = \frac{1}{m_c}(C_S - C'_S). \tag{3.56}$$

The constraints obtained from the $D^0 \to e^+e^-$ decay are much weaker than those obtained from $D^0 \to \mu^+\mu^-$. For $D^0 \to \mu^+\mu^-$, using the current best world limit at 90%C.L. given in Tab. 1.4, we find:

$$|C_S - C'_S|^2 + |C_P - C'_P + 0.1(C_{10} - C'_{10})|^2 \lesssim 0.008. \tag{3.57}$$

Barring cancellations between coefficients, this leads to

$$|C_{10} - C'_{10}| \leq 0.86, \tag{3.58}$$

$$|C_{P/S} - C'_{P/S}| \leq 0.087. \tag{3.59}$$

These constraints are similar to that obtained in Ref. [37]. We note that the constraint on C_{10} is ten times weaker than the constraints on C_P and C_S. This is due to the fact that this contribution is helicity suppressed, hence C_{10} is multiplied by m_ℓ.

Constraints from the $D^+ \to \pi^+\mu^+\mu^-$ decay measurement

As for $D^0 \to \ell^+\ell^-$, constraints obtained from the $c \to ue^+e^-$ transition are much weaker than the one obtained from the $c \to u\mu^+\mu^-$ transition. Compared to the leptonic $D^0 \to \mu^+\mu^-$ decay, which is only sensitive to effects in $C_{10}^{(\prime)}$, $C_S^{(\prime)}$ and $C_P^{(\prime)}$, the $D^+ \to \pi^+\mu^+\mu^-$ decay is sensitive to a more diverse range of new physics effects, namely $C_{10}^{(\prime)}$, $C_S^{(\prime)}$, $C_P^{(\prime)}$, C_T, C_{T5} but also to possible BSM contributions to C_7 and C_9. In the SM, these Wilson coefficients are multiplied by λ_b as visible in Eq. (3.46) and (3.47). We adopt the definition of [38] and define possible BSM contributions to C_7 and C_9 like

$$C_7 = C_7^{\rm SM} + \frac{C_7^{\rm BSM}}{\lambda_b}, \tag{3.60}$$

$$C_9 = C_9^{\rm SM} + \frac{C_9^{\rm BSM}}{\lambda_b}, \tag{3.61}$$

	Reg. I	Reg. II
$\lvert C_7^{\rm BSM}\rvert$	1.58	1.03
$\lvert C_9^{\rm BSM}\rvert$	2.11	1.3
$\lvert C_{10}+C'_{10}\rvert$	1.07	1.26
$\lvert C_S+C'_S\rvert$	4.29	0.82
$\lvert C_P+C'_P\rvert$	3.70	0.81
$\lvert C_T\rvert$	2.20	0.84
$\lvert C_{T5}\rvert$	2.79	0.90

Table 3.7: Constraints obtained from the limit at 90% C.L. on the branching ratio of $D^+ \to \pi^+\mu^+\mu^-$ decay.

where we assume that $C_7^{\rm BSM}$ and $C_9^{\rm BSM}$ are not sensitives to the same CKM dynamics than the SM contributions.

The constraints on the BSM Wilson coefficients from the $D^+ \to \pi^+\mu^+\mu^-$ decay are obtained from the limits at 90% C.L. given in Tab. 1.5. We use the central value of the parameters given in Tab. 3.4 and our ansatz for the branching ratio distribution given in Eq. (3.53) to compute the constraints on the Wilson coefficients $C_7^{\rm BSM}$, $C_9^{\rm BSM}$, C_{10}, $C_{S/P/T/T5}$. They are summarized in Tab. 3.7.

It is relevant to note that the upper limit lie really close to the SM prediction in Reg. I; however bounds on BSM coefficients are not stronger in this region. Indeed, this region is dominated by the weak annihilation and we don't consider new physics for the coefficients entering it, namely C_1 and C_2 essentially.

Discussion on the constraints

Several models generating $c \to u\ell^+\ell^-$ transition have been studied in the past, e.g. Minimal Supersymmetric Standard Models [32, 74, 103, 104], two Higgs doublet models [104], Little Higgs model [105, 106], vector-like quark singlet [32], scalar leptoquarks [38] and it is relevant to note that the size of Wilson coefficients that we obtain from the experimental constraints on $\mathcal{B}(D^0 \to \mu^+\mu^-)$ and $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$ can also be obtained from some of these concrete BSM models. For example, in Ref. [38], the following bounds on Wilson coefficients are obtained for a scalar leptoquark:

$$-C'_{10} = C'_9 = 0.63, \quad 4C_T = 4C_{T5} = C_P = C_S = -0.049, \tag{3.62}$$

where C_P and C_S reach the maximal value allowed from the experimental constraints on $\mathcal{B}(D^0 \to \mu^+\mu^-)$ as computed in [38]. Bounds are different from those we obtained as the authors use the limit at 95% C.L. and a different $\mathcal{B}(D^0 \to \mu^+\mu^-)$ expression. The following values are also obtained for a vector leptoquark scenario:

$$C'_9 = C'_{10} < 0.24. \tag{3.63}$$

Here we do not aim to perform a full analysis of all BSM models but to demonstrate that leptoquark models provide an example of a concrete model where it is possible to obtain observable effects. The upper limits on C_P, C_S and C'_9 from the scalar leptoquark model are of the same order as the model independent bounds from $\mathcal{B}(D^0 \to \mu^+\mu^-)$ and $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$. In the case of the tensor

BSM Wilson coeff.	Maximal values
$C_7^{\rm BSM}$	1.03
$C_9^{\rm BSM}$	1.3
C_{10}	0.087
C_S	0.86
C_P	0.86
C_T	0.84
C_{T5}	0.90

Table 3.8: Summary of the maximal value allowed by the current experimental limits on $D^0 \to \ell^+\ell^-$ and $D^+ \to \pi^+\ell^+\ell^-$ at 90% C.L. if we assume the BSM Wilson coefficients are real and that the chirality-flipped Wilson coefficients are all equal to zero.

contributions, C_T and C_{T5}, the leptoquark prediction are smaller by a factor $\sim$80 but as we will see are still large enough to produce effects distinguishable from the SM.

As in Ref. [38], we note that the upper limit on the $\mathcal{B}(D^0 \to \mu^+\mu^-)$ decay is more restrictive for C_S, C_P and C_{10} than any of the bounds from the $D^+ \to \pi^+\ell^+\ell^-$ decay. We also remark that C_7, C_9, C_T and C_{T5} are better constrained by the limit set in the experimental region Reg. II than the one set in Reg. I.

If we assume that the BSM Wilson coefficients are real and that the chirality-flipped Wilson coefficients are all equal to zero, the maximal values allowed by the current experimental limits on $\mathcal{B}(D^0 \to \ell^+\ell^-)$ and $\mathcal{B}(D^+ \to \pi^+\ell^+\ell^-)$ are summarized in Tab. 3.8.

Branching ratio

The distribution of $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$ is given in Fig. 3.6 for the SM (dashed black) and various BSM scenarios. We show the effect of varying only C_T (green), C_P (red), C_{10} (purple) and C_9 (pink). For illustrative purposes, the BSM Wilson coefficients are varied individually and are assumed to take the maximal values allowed by the experimental measurements as summarized in Tab. 3.8.

As we can see, the upper limit in Reg. I lies close to the SM prediction and is below the upper line of our error band (yellow). In addition all the BSM scenarios lie within this error band. Reg. II is more interesting as the bounds on BSM physics are such that the BSM predictions could be distinguished from the SM in a future measurement (depending on the experimental errors). In addition, the error band obtained in this region on the SM prediction is smaller. The C_T and C_9 lines lie close to the experimental limits as expected since the values used are coming from the experimental limit on $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$ in Reg. II. The C_{10} distribution is slightly lower in this region and the C_P line lie close the SM prediction. The C_{10} and C_P lines lie far from the experimental limits on $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$ as their bounds come from limits on $\mathcal{B}(D^0 \to \mu^+\mu^-)$.

As illustrated in Fig. 3.7, the error band (pink) obtained for the BSM scenario where C_{T5} is maximal (black line) is really small in this region. To illustrate the impact on increased bounds on these coefficients, scenarios where $C_{T5} = C_{T5}^{\rm max}/2$ (purple) and $C_{T5} = C_{T5}^{\rm max}/10$ (red) are shown. We note that the error band on C_{T5}

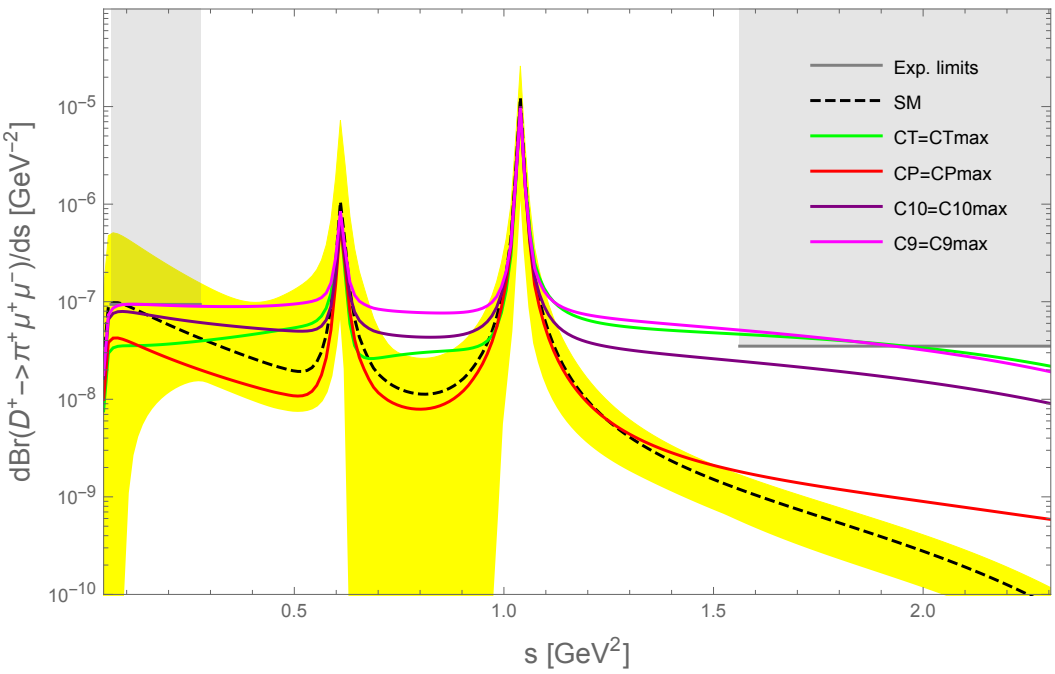


Figure 3.6: Distribution of the $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$ for the SM and various BSM scenarios. The error band on the SM prediction is shown in yellow and the experimentally excluded regions are in gray. For illustrative purposes, the BSM Wilson coefficients are varied individually and are assumed to take the maximal values allowed by the experimental measurements as summarized in Tab. 3.8.

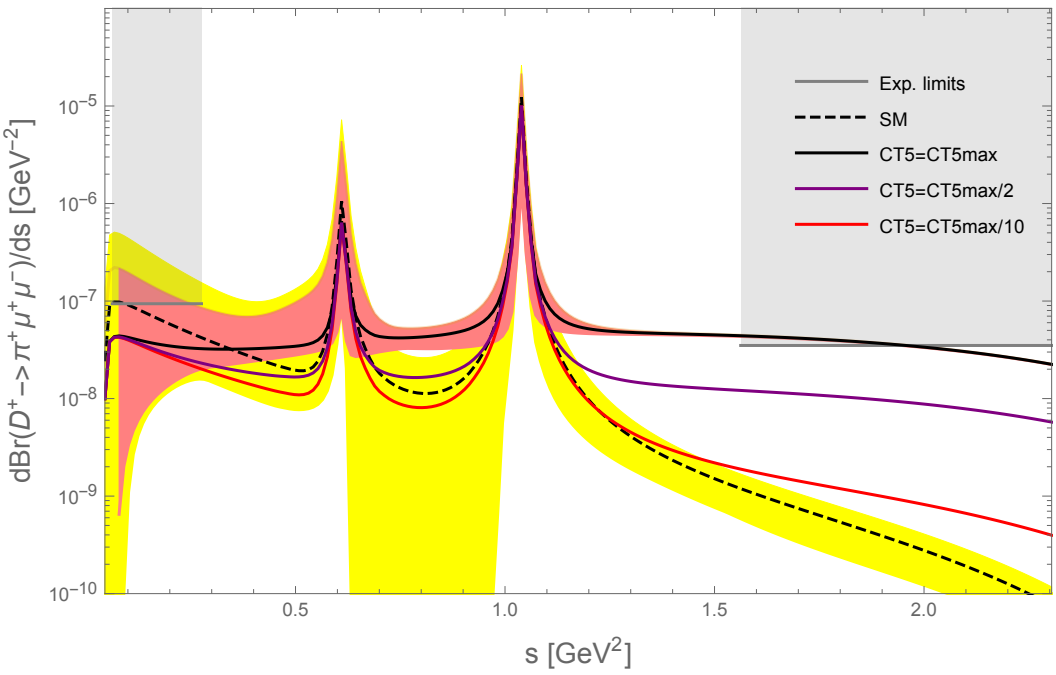


Figure 3.7: Distribution of the $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$ for the SM and various BSM scenarios where C_{T5} is non zero. The error band on the SM prediction is shown in yellow and the experimentally excluded regions are in gray. The error band for the scenario where $C_{T5} = C_{T5}^{\max}$ is shown in pink.

is extremely thin in high s region which give even more interest to this region. The scenario where $C_{T5} = C_{T5}^{\max}/2 \simeq 0.45$ and the SM can be distinguished in this region when the experimental error is smaller than the theory error. The scenario where $C_{T5} = C_{T5}^{\max}/10 \simeq 0.09$ shows the approximate limit where a scenario only including a contribution to C_{T5} could be distinguished from the SM via the branching ratio measurement.

While not shown here, we found that the branching ratio distribution for the maximum allowed Wilson coefficients in the scalar leptoquark model of Ref. [38] (bounds are provided in Eq. (3.62)) is similar to that for C_T (green line) and hence distinguishable from the SM in the high s region.

3.2.4 Observables

The flat-term F_H

As in the SM, $c_\ell(s) = -\beta^2(s)a_\ell(s) \simeq a_\ell(s)$ in high s region and this equality is broken in many BSM models, hence an observable sensitive $a_\ell + c_\ell$ would be a clean null experimental test of the SM. Following [64], the flat term F_H can be defined as

$$\begin{aligned} F_H(s) &\equiv \frac{a_\ell(s) + c_\ell(s)}{a_\ell(s) + c_\ell(s)/3} \\ &= \frac{s}{d\hat{\Gamma}(s)} \left(\frac{m_D^2 - m_\pi^2}{m_c^2} f_0^2(|C_P|^2 + |C_S|^2) + 16\lambda \frac{f_T^2\left(|C_T|^2 + |C_{T5}|^2\right)}{(m_D + m_\pi)^2} \right). \end{aligned} \tag{3.64}$$

The last equality in Eq. (3.64) is valid only in the high s region where $\beta(s) \sim 1$. In the SM, the flat term is only a function of $\beta(s)$ up to $\mathcal{O}(m_\ell)$ corrections:

$$F_H^{\rm SM}(s) = \frac{1 - \beta(s)^2}{1 - \beta(s)^2/3} + \mathcal{O}(m_\ell), \tag{3.65}$$

resulting in a small uncertainty on the SM distribution. Distributions of the flat term are shown in Fig. 3.8 for the SM (black dashed) and different BSM scenarios. We show the effect of varying only C_T (green), C_{T5} (blue), C_{10} (purple) and C_9 (red). For illustrative purposes, the BSM Wilson coefficients are varied individually and are assumed to take the maximal values allowed by the experimental measurements as summarized in Tab. 3.8. The error band on scenario where C_P is varied is also drawn in pink. This latter is includes the same set of uncertainties as the SM error band for the branching ratio described earlier. The error band for the SM is not drawn as it is expected to be very small.

The blue and green lines (C_{T5} and C_T) present an inverse double peak structure. This structure is also present in the red and purple lines (C_{10} and C_P) but the effect is less visible. This structure is due to the quantity $d\hat{\Gamma}(s)$ present in the denominator of F_H, the definition of which is given in Eq. (3.38). This quantity is proportional to the branching ratio distribution, it is then expected to see the effect of the resonances on F_H albeit inverse.

The C_T and C_{T5} lines (blue and green) lie much above the SM prediction in the ($s > 1.2$ GeV^2) region. The C_P line (red) is a bit closer to the SM prediction and the C_{10} line even more. Nevertheless, even the C_{10} line can be distinguished from the SM. The smallness of the SM uncertainty and the effects of BSM scenarios on this

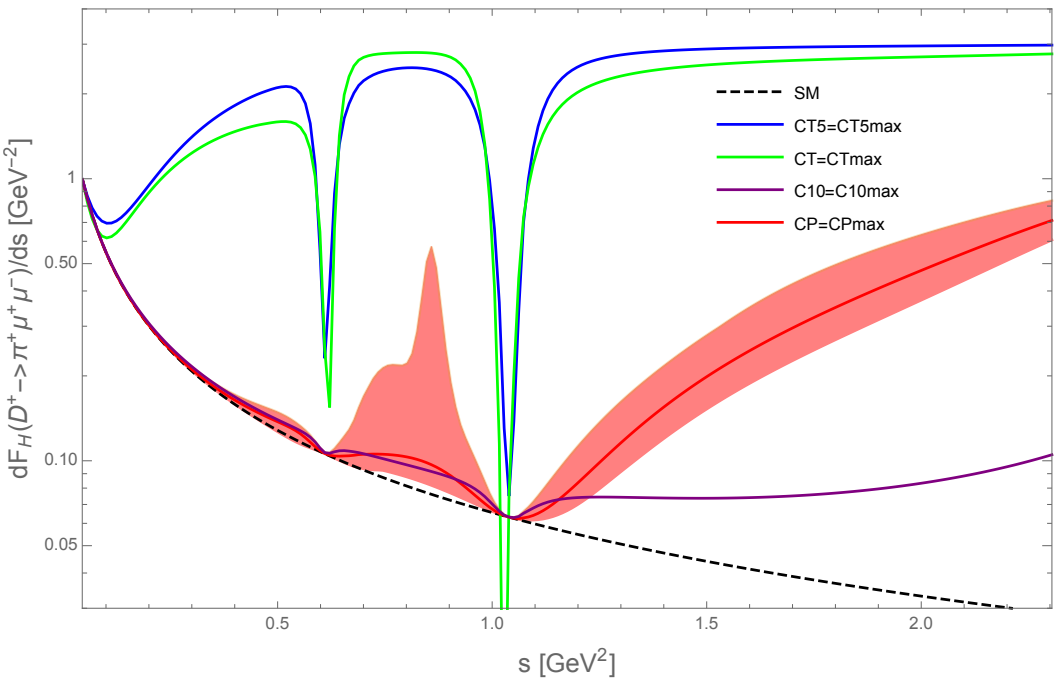


Figure 3.8: Distribution of the flat term for the $D^+ \to \pi^+\mu^+\mu^-$ decay for the SM and different BSM scenarios. For illustrative purposes, the BSM Wilson coefficients are varied individually and are assumed to take the maximal values allowed by the experimental measurements as summarized in Tab. 3.8. The error band on the the scenario where $C_P = C_P^{\rm max}$ is shown in pink.

observable make F_H a very clean observable to probe BSM physics in $D^+ \to \pi^+\mu^+\mu^-$.

While not shown here, we found that the F_H distribution for the maximum allowed Wilson coefficients in the scalar leptoquark model of Ref. [38] (bounds are provided in Eq. (3.62)), is found to be similar to that for C_{10} (purple line) and hence distinguishable from the SM in the high s region.

As the values for F_H are quite large in the SM for low value of s and the downward fluctuations due to the resonances can make BSM physics more difficult to detect, we advocate the measurement of F_H in region above the ϕ resonance: ($s > 1.2$ GeV^2). Thus, it would be interesting to measure the following observable:

$$F_H^{\rm int} = \int_{s>1.2\ {\rm GeV}^2} ds\ F_H(s) \tag{3.66}$$

The Forward-Backward asymmetry $A_{\rm FB}$

Noticing that Eq. (3.39) might not hold anymore in BSM scenarios, it is interesting to build observables sensitive to the angular coefficient $b_\ell(s)$, one example of which is

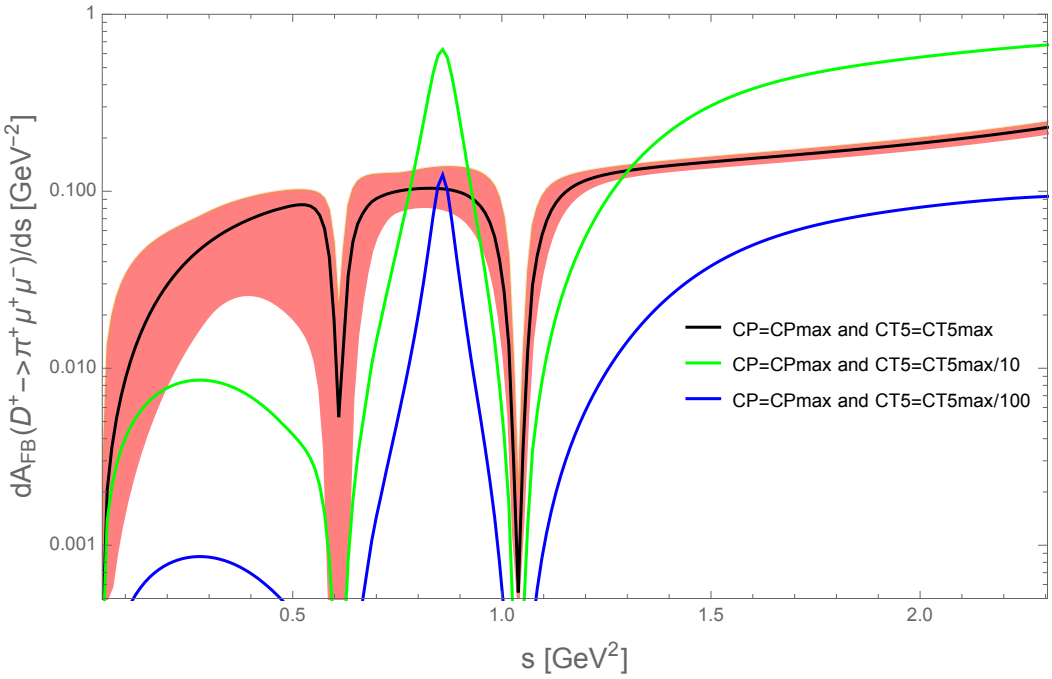


Figure 3.9: Distribution of $A_{\rm FB}$ for various BSM scenarios where C_P and C_{T5} are not zeros. A scenario where where both C_P and C_{T5} reach the maximal allowed value given in Tab. 3.8 is shown (black) and the associated uncertainty (pink band). A similar scenario but where C_{T5} is divided by 10 (100) is shown in green (blue).

the Forward-Backward asymmetry $A_{\rm FB}$. Following [38], $A_{\rm FB}$ can be defined as

$$\begin{aligned} A_{\rm FB}(s) &\equiv \frac{\left(\int_0^1 - \int_{-1}^0\right) d\cos\theta \frac{d^2\Gamma}{ds\ d\cos\theta}}{\int_{-1}^1 d\cos\theta \frac{d^2\Gamma}{ds\ d\cos\theta}} \\ &= \frac{b_\ell(s)}{2\left[a_\ell(s) + c_\ell(s)/3\right]} \\ &= \frac{2s\sqrt{\lambda}}{m_c\, d\hat{\Gamma}(s)} f_0 f_T \left[Re(C_{T5}C_P^*) + Re(C_T C_S^*)\right] + \mathcal{O}(m_\ell) \end{aligned} \tag{3.67}$$

As visible in Eq. (3.67), $A_{\rm FB}$ depends, up to $\mathcal{O}(m_\ell)$ corrections, only on BSM Wilson coefficients and $A_{\rm FB} \simeq 0$ in the SM. Moreover, none of the BSM Wilson coefficients are individually enough to have $A_{\rm FB} > 0$ and we need a combination of Wilson coefficients, either both C_{T5} and C_P different from zero or both C_T and C_S different from zero.

The distribution of $A_{\rm FB}$ is shown in Fig. 3.9 for various BSM scenarios where C_P and C_{T5} are not zeros. The SM distribution is not shown as $A_{\rm FB} \simeq 0$ in the SM. A scenario where both C_P and C_{T5} reach the maximal allowed value given in Tab. 3.8 is shown (black) and the associated uncertainty (pink band). A similar scenario but where C_{T5} is divided by 10 (100) is shown in green (blue). This shows that even for values of tensor Wilson coefficients ($C_{T5} \simeq C_{T5}/100 \simeq 0.09$) much below their limit the BSM effects should be observable, depending on the future experimental sensitivity. This very interesting effect is due to the presence of C_{T5} in the quantity $d\hat{\Gamma}(s)$ entering the denominator.

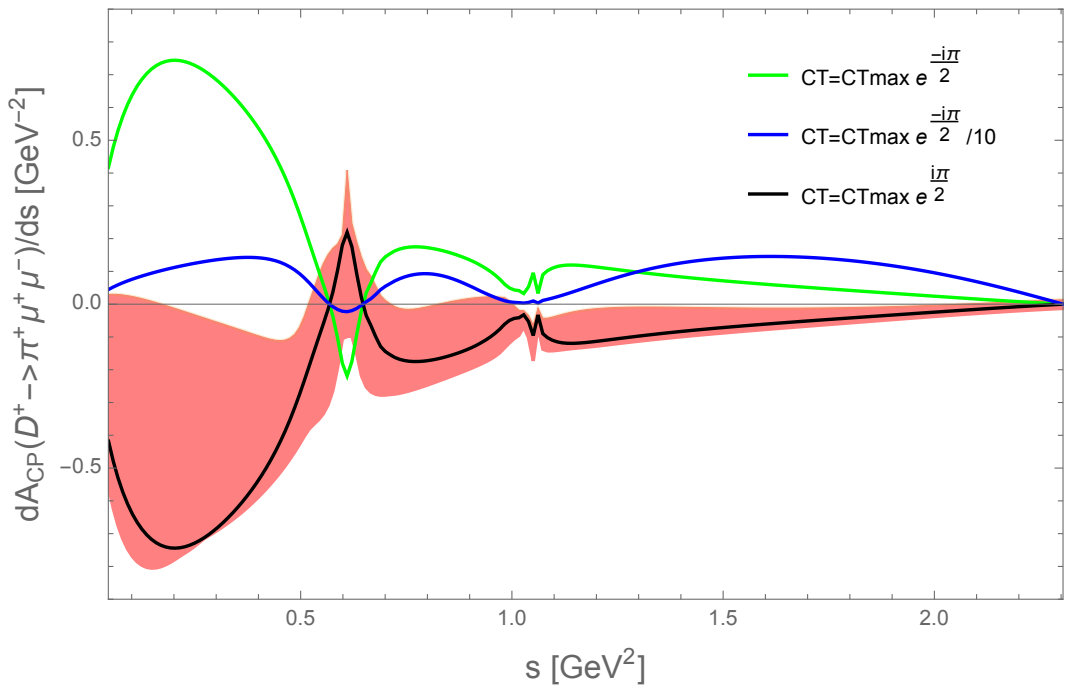


Figure 3.10: Distribution of A_{CP} for different BSM scenarios.

We observe, as for F_H, the inverse resonant structure due to the presence of $d\hat{\Gamma}(s)$ in the denominator, see Eq. (3.67). As for the F_H observable, $A_{\rm FB}$ turns out to be cleaner in the high than in the low region of s where effects of BSM physics can be hidden by the resonances or simply be invisible if the Wilson coefficients are too smalls.

The CP-asymmetry A_{CP}

The CP-asymmetry A_{CP} is sensitive to the phases of the BSM Wilson coefficients (the phases are supposed strong), it is defined by [37]:

$$A_{CP}(s) = \frac{d\Gamma/ds - d\bar{\Gamma}/ds}{\int_{s_{\rm min}}^{s_{\rm max}} ds(d\Gamma/ds + d\bar{\Gamma}/ds)} \tag{3.68}$$

where $d\bar{\Gamma}/ds$ is the differential decay rate of the CP-conjugated mode, $D^- \to \pi^-\ell^+\ell^-$.

A_{CP} is zero in the SM up to $\mathcal{O}(m_\ell)$ terms and for BSM scenarios where Wilson coefficients are real. Distributions of A_{CP} for different BSM scenarios are shown in Fig. 3.10. A scenario where C_T reaches the maximal allowed value given in Tab. 3.8 and where a phase $\pi/2$ is added (this is the phase where A_{CP} reaches its maximal absolute values) is shown (black) along with the associated uncertainty (pink band). A similar scenario but where the phase is opposite ($-\pi/2$) is shown in green. Finally, a scenario where the C_T phase is still $-\pi/2$ but C_T only reaches a tenth of its maximal allowed value is shown in blue. The SM prediction is not shown as very small.

We note that the resonant structure is also visible for A_{CP} and that larger values are found in the low region of s. This implies that a measurement of A_{CP} in the low s region is more appropriate than in the high s region. Nevertheless we observe that the related uncertainty is large and even the maximal value of C_T (black line

and pink uncertainty) is compatible with zero through the majority of the spectrum. Therefore, for the moment, A_{CP} is not a good observable candidate to probe BSM physics in the $D^+ \to \pi^+\mu^+\mu^-$ decay.

3.3 Conclusions

We presented in this chapter an alternative to describe the resonance contributions. Following the idea of Ref. [36], the vacuum polarizations are reconstructed from their imaginary part using a dispersion relation. The imaginary parts are modeled following a proposal by M. Shifman [39], i.e. a Breit-Wigner type function modeling the dominant vector resonance plus the sum of an infinite tower of resonances, starting from the first excitation. Our work is based on this resonance description, with the following improvement:

- In Ref. [36], the isospin 1 (ρ^0) and isospin 0 (ω^0) light resonances are not treated separately, but in terms of a single tower of resonances with "effective" parameters. We improve upon this model by treating separately the isospin contributions.

- Instead of using "effective" parameters, we follow the strategy of [40], where the authors relate the imaginary part of the charm hadronic spectrum to the experimental ($e^+e^- \to$ hadrons) data.

Armed with this resonance description and the framework presented in Chap. 2, we gave predictions for the $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$ distribution. As in [36] for the $D^0 \to \rho^0\ell^+\ell^-$, we found that the non-factorizable annihilation contribution, neglected in [37, 38], is one of the largest SM contributions in the low s region. Its impact on the SM distribution of the branching ratio is visible by the large increase in low s region with respect to the ones shown in Ref. [37, 38].

We then computed, with the help of the branching ratio distribution, new bounds on BSM Wilson coefficients imposed by the current experimental upper limits on the $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$ [42] and $\mathcal{B}(D^0 \to \mu^+\mu^-)$ [41]. Our results are summarized in Tab. 3.8. We compared our results to the upper limits from the scalar leptoquark model presented in Ref. [38], whose Wilson coefficient values were provided in Eq. (3.62), and found that C_P, C_S and C_9' are of the same order as our model independent bounds. In the case of the tensor contributions, C_T and C_{T5}, the leptoquark predictions are smaller by a factor $\sim$80 but as we saw, they are still large enough to produce effects distinguishable from the SM.

The high s region is more interesting as the bounds on BSM physics are such that the BSM predictions could be distinguished from the SM in future measurements. In addition, the error band obtained in this region on the SM prediction is thiner than that in low s region. For illustrative purposes, we showed several distributions where the BSM Wilson coefficients were varied individually. We found that scenarios where C_T, C_P, C_{10} and C_9 take the maximal values allowed by the experimental measurements as summarized in Tab. 3.8, were fully distinguishable from the SM. We also showed that BSM effects can be seen for value of C_{T5} up to $C_{T5}/10 \simeq 0.08$. In addition, we found that distribution for the maximal Wilson coefficients allowed by the scalar leptoquark model of Ref. [38] lies just below the upper limit in Reg. II

which reinforce the interest for future measurements in this region.

The effect of the phases of the resonances on the branching ratio distribution has also been explored and we found that they were not measurable except perhaps for some value of φ_ω around $7\pi/10$.

We then pursued our phenomenological analysis with the observables F_H, $A_{\rm FB}$ and A_{CP}. We observed that the SM prediction for F_H is very clean and is not zero in the low s region. For illustrative purposes, we showed distributions where C_T, C_{T5}, C_P and C_{10} take the maximal values allowed by the experimental measurements as summarized in Tab. 3.8. We found that all these BSM distributions are sensitive to the resonances but can be distinguished from SM in the high s region, more particularly the tensors contributions. In addition, the F_H distribution for the maximum allowed Wilson coefficients in the scalar leptoquark model of Ref. [38] appears to be also distinguishable from the SM in the high s region.

The study of $A_{\rm FB}$ shows that this observable is only sensitive to combinations of BSM coefficients, i.e. C_{T5} and C_P or C_T and C_S. We plotted the distribution of $A_{\rm FB}$ for BSM scenarios where both C_P and C_{T5} reaches the maximal allowed value given in Tab. 3.8, as well as similar scenarios but where C_{T5} is divided by 10 and 100. We found that, as for F_H, the observable is sensitive to the resonances and that even for small value of C_{T5} ($\sim$0.08), the distribution can still be distinguished from the SM which is approximately zero for the full spectrum. Therefore, even for values of tensor Wilson coefficients much below their limit the BSM effects should be observable, depending on the future experimental sensitivity.

The observable A_{CP} was found to be inefficient to probe BSM physics as the related uncertainty is large and even for the maximal value of C_T, the observable is compatible with zero through the majority of the spectrum.

To summarize, the branching ratio, F_H and $A_{\rm FB}$ were found to be very good candidates to probe BSM models. $A_{\rm FB}$ and F_H are very clean and null observables in the SM in the high s region. The branching ratio is not zero in this region but we found that BSM physics, like the one allowed by the scalar leptoquark model of Ref. [38], can be distinguished from the SM. As mentioned, the three observables are sensitive to the resonance. In the case of $A_{\rm FB}$ and F_H, they decrease the observables rendering the BSM effects harder to detect. The observables also show a larger discrepancy from the SM in the high s region. Therefore, we advocate the measurement in the s region above the ϕ resonance: $s > 1.2$ GeV^2 for these three observables.

An overall conclusion of the work performed on the $D^+ \to \pi^+\ell^+\ell^-$ decay is given in Chap. 8.

Part II

Search for $B^0_{(s)} \to \tau^+\tau^-$ at LHCb

Chapter 4

The LHCb experiment

The data analyzed in our search for $B^0_{(s)} \to \tau^+\tau^-$ decays are collected by a detector called LHCb, a dedicated B (and D) physics experiment installed at the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN). The LHCb experiment is mainly devoted to the study of rare decays of b and c-hadrons and CP-violation.

An overview of the LHC environment from the point of view of B-physics is given in Sec. 4.1. The LHCb detector and the simulation chain are described in Sec. 4.2. Finally, some prospects regarding B-physics experiments are discussed in Sec. 4.3.

4.1 The LHC Environment

A brief overview of how protons are generated and accelerated at the LHC is given in Sec. 4.1.1. Then, Sec. 4.1.2 explains how B mesons are produced in proton collisions and introduces two important quantities, the cross-section and the luminosity.

4.1.1 The CERN accelerator complex

The LHC is a circular collider installed deep underground ($\sim$ -100 m) in the 27 km-long tunnel previously used for the Large Electron-Positron collider (LEP). It can accelerate protons but also heavy ions in two opposite directions to make them collide at different points of the ring, where several detectors are located.

The LHC data taking is split in periods called "Runs". A Run typically lasts several years and is separated from the next one by a long phase of maintenance and possible upgrading called Long Shutdown (LS). Run 1 started in 2010 and finished in 2012. The collision energy for protons was of 7 TeV in 2011 and 8 TeV in 2012. Run2 started in 2015, after the LS1, and finished in 2018. Although the LHC was designed to run at a nominal energy of 14 TeV, the energy during the full Run 2 was of 13 TeV.

Protons generation and acceleration

Hydrogen atoms are submitted to an intense electric field via a duoplasmatron in order to be stripped of their electrons. The resulting protons are accelerated by different machines present on the CERN site before their injection into the LHC. They are first accelerated up to 50 MeV by a first linear accelerator (LINAC) before being boosted by the Proton Synchrotron booster up to 1.8 GeV. This booster allows

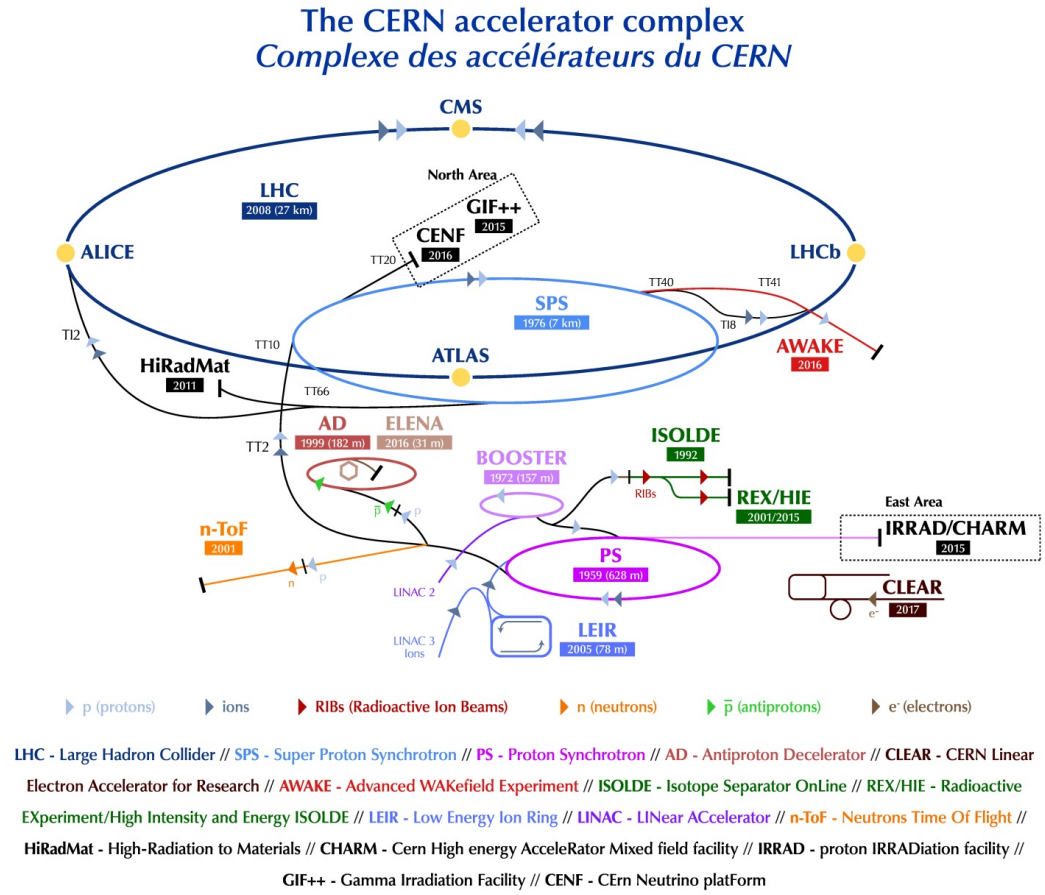


Figure 4.1: The CERN accelerator complex.

to inject protons into the Super Proton Synchroton (SPS) which speeds them up to an energy of 450 GeV. At this stage, protons are injected into the LHC. This acceleration chain is shown schematically in Fig. 4.1.
Within the LHC, the protons are accelerated and driven by superconducting electromagnets: 1232 15 m long dipole magnets that guide the proton beams along the ring and 392 5/7 m long quadrupole magnets that keep the beam focused. For Run1, the beam is structured in 1800 bunches of $\sim 1.2 \times 10^{11}$ protons, each bunch being separated by a bunch spacing of 50 ns. For Run2, the number of bunches has been increased to 2835 and the bunch spacing reduced to 25 ns, i.e. a bunch crossing frequency f_{rev} of 40 MHz. More information about the LHC machine can be found in [44].

The two beams cross each other at four points of the LHC machine where four main detectors are located. Two of them are polyvalent and general-purpose experiments: "A Toroidal LHC ApparatuS" (ATLAS) and the "Compact Muon Solenoid" (CMS). They both discovered the Higgs boson in 2012 and aim at observing new physics particles. Another experiment, "A large Ion Collider Experiment" (ALICE) is dedicated to the study of quark-gluon plasma in extreme conditions of pressure and temperature. Finally, the LHCb experiment is dedicated to the study of b and c-hadrons.

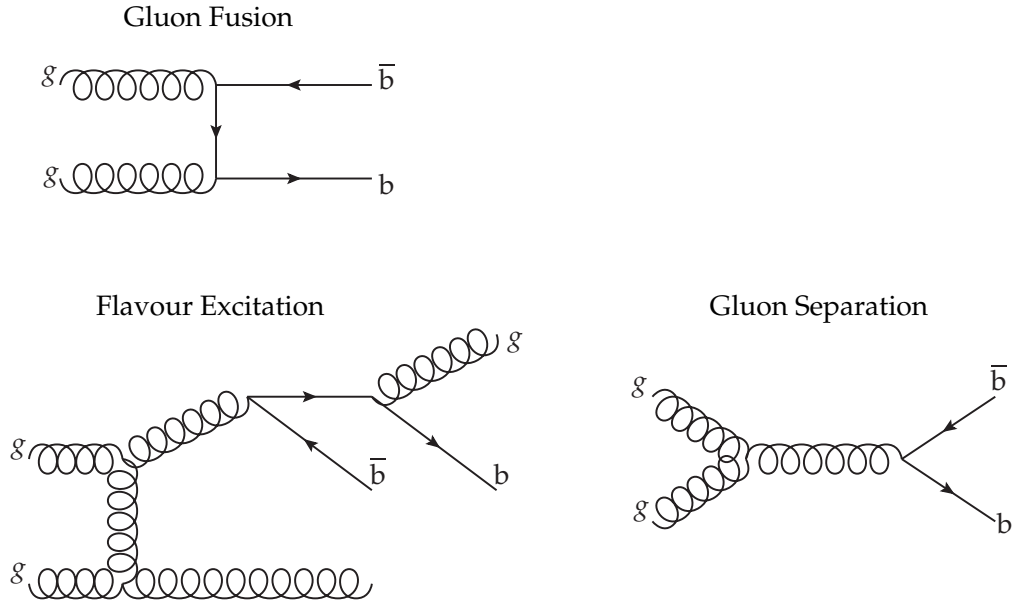


Figure 4.2: The main processes at work for $b\bar{b}$ production at the LHC.

4.1.2 *B* physics at the LHC

An important quantity in B physics is the number of b-hadrons N_b that a collider can deliver. This number can be expressed as:

$$N_b = 2 \cdot \sigma_{b\bar{b}} \cdot L, \tag{4.1}$$

where $\sigma_{b\bar{b}}$ is the cross-section for the creation of pairs of quark and anti-quark $b\bar{b}$, often given in femto-barn (fb) and L the time-integrated luminosity given in fb^{-1} [107]. These two quantities are described in the following.

The $b\bar{b}$ production at the LHC

The production of $b\bar{b}$ pairs is the result of interactions among the partons during the proton collisions. The main processes [108], illustrated in Fig. 4.2, are the following:

- parton fusion (16% of the production): the pair arises from gluon fusion or quark annihilation: $q\bar{q} \rightarrow b\bar{b}$ or $g\bar{g} \rightarrow b\bar{b}$.
- flavor excitation (54% of the production): it takes place when one of the two $b\bar{b}$ virtual quarks in the proton quark pair undergoes an interaction with a parton of the other proton.
- gluon separation (27% of the production): hard scattering produces a gluon splitting into a $b\bar{b}$ pair.

The production cross-section $\sigma_{b\bar{b}}$ also increases with the center-of-mass energy and basically doubled between Run1 and Run2.

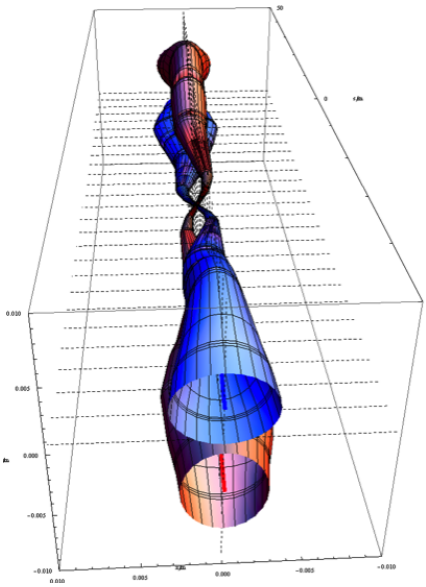


Figure 4.3: Schematic view of a beam collision at LHCb. The two beams (one represented in red, the other in blue) are squeezed and cross each other at the interaction point. The angle between the two beams is taken into account by the geometrical factor F. The overlap between the two beams in the transverse plane to beam axis is called the transverse overlap.

The luminosity

The time-integrated luminosity L is given by:

$$L = \int \mathcal{L} dt, \tag{4.2}$$

where $\mathcal{L}$ is the instantaneous luminosity often given in [cm^{-2}s^{-1}]. It can be expressed as:

$$\mathcal{L} = \frac{N_p^2 n_b f_{rev} F}{4\pi\sigma_{xy}}, \tag{4.3}$$

where N_p and n_b are respectively the number of protons per bunch and the number of bunches, f_{rev} is the revolution frequency, σ_{xy} the size of the beam in the transverse plane and F a geometric factor accounting for the luminosity reduction due to the crossing-angle of the beams. A schematic view of the beam collision at the LHCb detector, illustrating the importance of these different parameters, can be seen in Fig. 4.3.

While the ATLAS, CMS and ALICE experiments are able to handle a large pile-up (number of collisions per bunch crossing) $\mu \sim 50$, LHCb tries to keep $\mu \sim 2$ in order to minimize background in the detector and to optimize a secondary vertex reconstruction. A secondary vertex is a vertex displaced from the reconstructed Primary Vertex (PV), which is simply one of the interaction points of the two proton beams. Indeed, the B and D meson has a lifetime long enough to "fly" within the detector before decaying.

The instantaneous luminosity is not constant as, due to collisions, the density of protons decreases with time. Every few hours, the beams are dumped and new beams are injected into the LHC. This period of few hours from when the beams are put in collision until the beams are dumped, is called a *fill*.

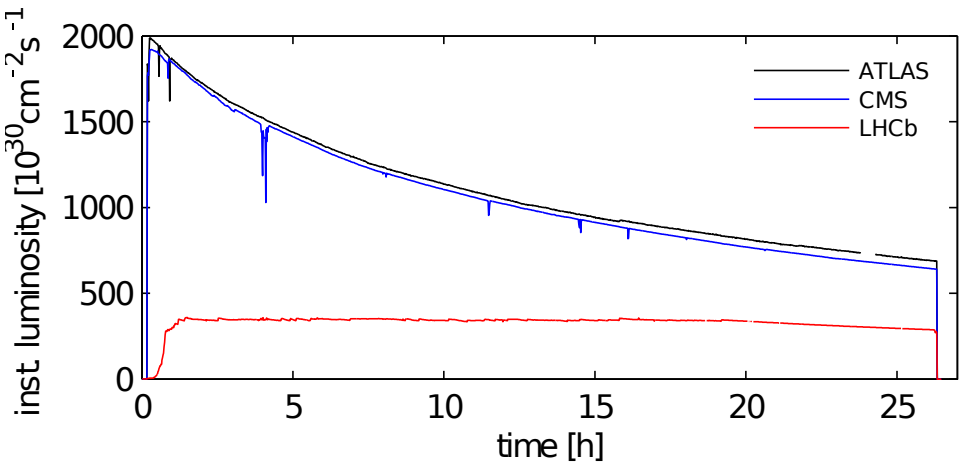


Figure 4.4: Leveling of the instantaneous luminosity in LHCb compared to the exponential decrease of the CMS and ATLAS luminosity during a fill.

	Run1	Run2
Period of data taking	2010-2012	2015-2018
Center-of-mass energy (TeV)	7 (8 in 2012)	13
Cross-section in the detector acceptance (μb)	$\sim$ 72 [46]	$\sim$ 144 [46]
Bunch spacing (ns)	50	25
$\mathcal{L}$ recorded (fb^{-1})	3.1	5.9

Table 4.1: Overview of the LHCb data taking conditions for Run1 and Run2.

In LHCb, the transverse overlap of the beams is adjusted as the fill evolves in time in order to keep constant luminosity and hence $\mu \sim$2. This technique is called *luminosity leveling* [109] and is illustrated in Fig. 4.4.
A summary of the Run1 and Run2 LHCb data taking conditions can be seen in Tab. 4.1. More details on theses conditions can be found respectively in [110] and [111]. The integrated luminosity recorded in LHCb for pp collisions year by year can be seen in Fig. 4.5.

4.2 The LHCb detector

The b and c-hadron properties, as well as their production at the LHC, drive fully the design of the detector from its *geometry*, to its main characteristics: a precise *tracking*, a powerful *particle identification* and an efficient *trigger system*. These four main features of the LHCb detector are detailed in the following sections. An overview of the LHCb detector layout is shown in Fig. 4.6. More details on the LHCb detector can be found in [43].

4.2.1 LHCb geometry

The $b\bar{b}$ production cross section $\sigma_{b\bar{b}}$ presented in Sec. 4.1.2 has an angular distribution mainly peaking in the backward/forward regions with respect to the beam axis, implying that most b-hadrons at LHC, and consequently their decay products, are produced in a cone of small aperture. Hence, a detector with a conic shape covering a region from approximately $\theta = 0.9°$ to $17°(14°)$ in the horizontal (vertical) plane, is

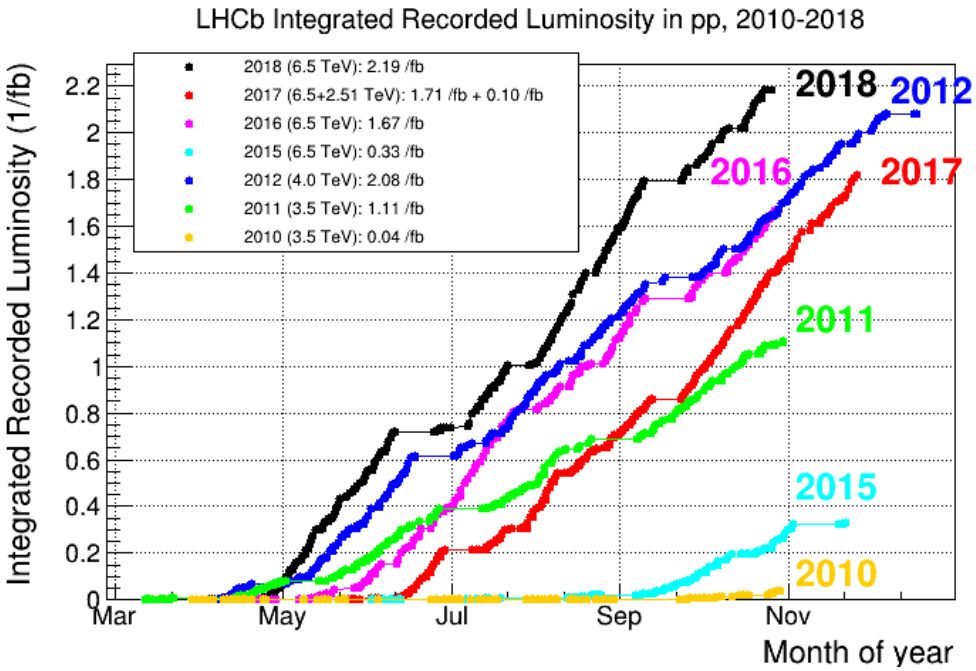


Figure 4.5: Integrated luminosity recorded in LHCb for pp collisions for Run1 and Run2.

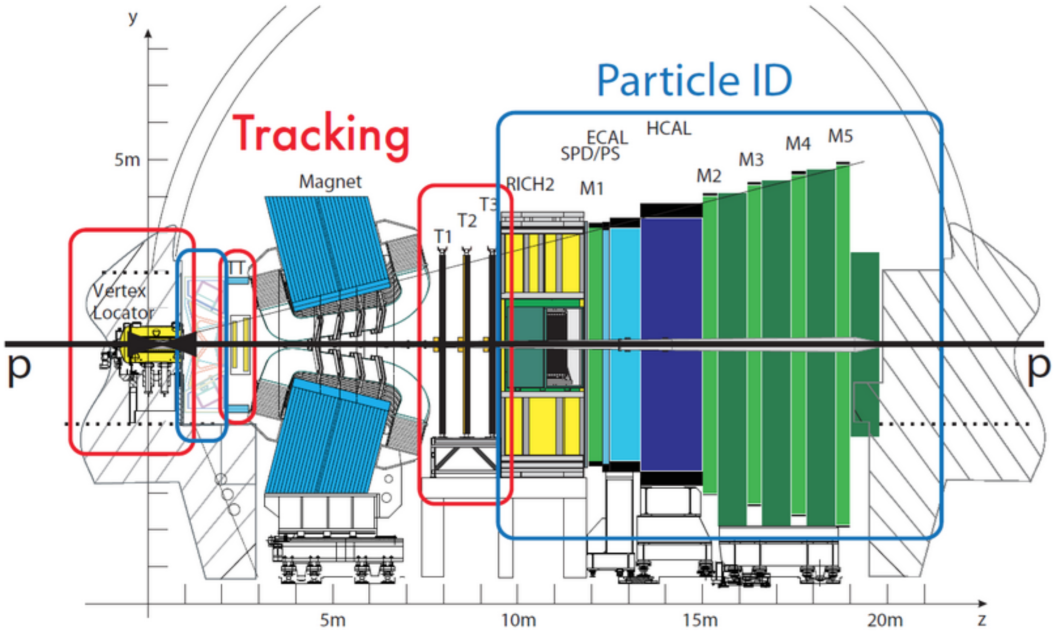


Figure 4.6: Schematic view of the LHCb detector, with circled in red (blue) the main components of its tracking (particle identification) system. The collisions take place within the Vertex Locator, which is the most upstream sub-detector.

enough to detect one fourth of all b-hadrons produced. In terms of pseudo-rapidity η,

$$\eta = -\ln \tan \frac{\theta}{2} \tag{4.4}$$

the LHCb detector coverage range is approximately $2 < \eta < 5$, which is extremely forward with respect to the general-purpose detectors ATLAS and CMS. A comparison

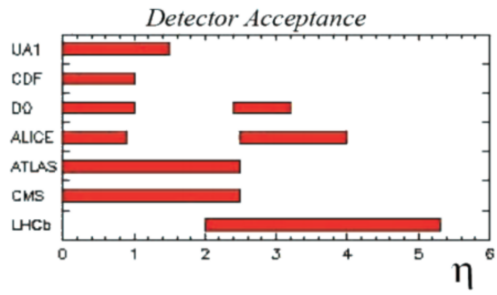


Figure 4.7: Different detector coverages in pseudo-rapidity. LHCb is the most forward detector.

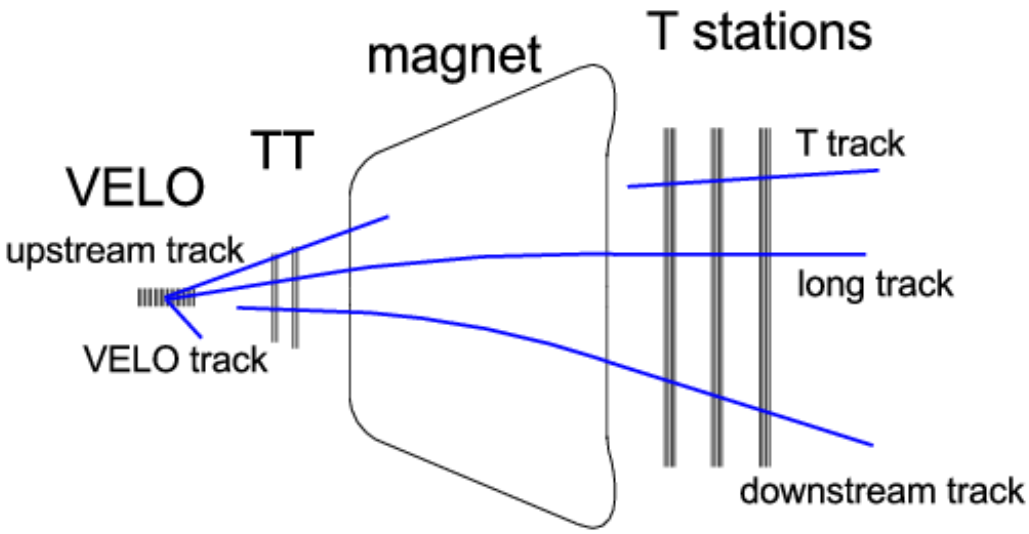


Figure 4.8: LHCb track categories

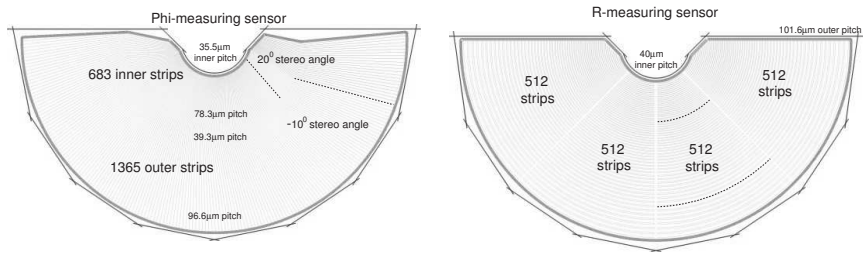
of different detector coverages can be seen in Fig. 4.7.

4.2.2 Tracking system

The tracking system's goal is to measure the trajectories and the momenta of charged particles. It is composed of the magnet [112] (that bends the charged particle trajectories), the VErtex LOcator (VELO) [113] and the Tracking stations (TT, T1, T2 and T3) [114]. A *tracking algorithm* [115] is used to translate the set of hits collected in the sub-detectors into tracks. Depending on which sub-detectors are used, tracks can be classified in different categories as shown in Fig. 4.8. Long tracks are those with best momenta and Impact Parameter (IP) resolution. The IP with respect to the PV is defined as the Distance Of Closest Approach (DOCA) of the reconstructed PV and the track. Hence, they are the main tracks used in physics analyses. The tracking system allows to reconstruct long tracks with a 96% efficiency.

The Magnet

The magnet allows to calculate the momenta of the charged particles by deflecting their trajectory in the horizontal plane thanks to a magnetic field of $\sim$4 T. The momentum resolution is $\delta p/p = 0.4\%$ for p=5 GeV. As such a strong magnetic field



(a) (b)

Figure 4.9: Phi (a) and R-measuring (b) VELO sensors.

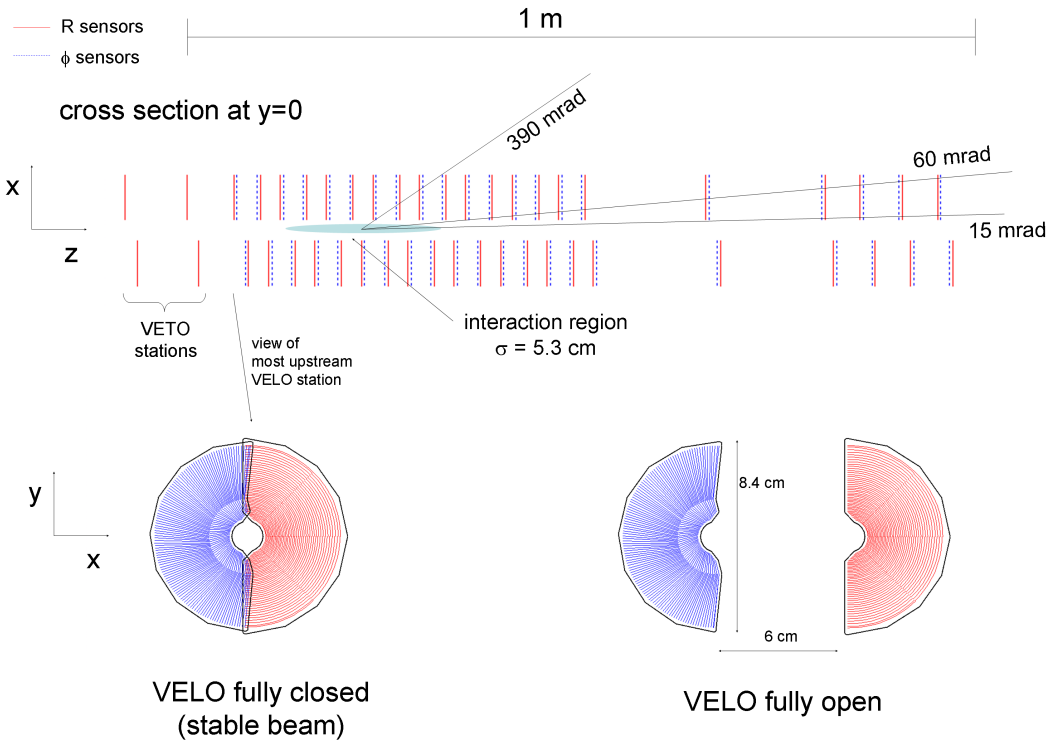


Figure 4.10: Overview of the VErtex LOcator

has also an impact on the beam trajectory, the beam deflection is corrected by three dipole magnets placed downstream the detector.

The Vertex Locator

B mesons have a long lifetime (on average of 1.5 ps) and the boost acquired during their production is enough to make them "fly" about 1 cm within the detector before decaying. This imposes to the detector to be able to distinguish the secondary decay vertex from the primary(ies) one(s). For this purpose, a special tracking device, the VELO, is placed as close as possible to the interaction point. This detector is made out of half circular sensors with two kinds of geometry: A Phi and R-measuring sensors segmented respectively along their polar ϕ and r coordinates as shown in Fig. 4.9. 42 R-Phi pairs of these semi-circled sensors are arranged in two 1 m long "semi-cylinder" making two halves of the VELO. These sensors are placed at 8 mm from the collision point. This distance being smaller than the beam diameter during

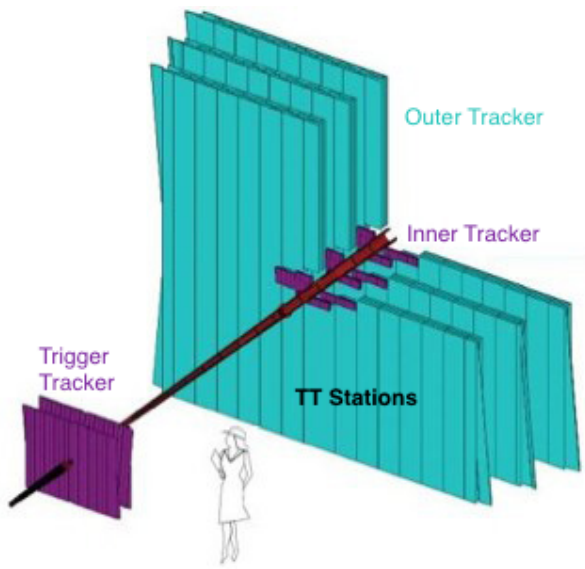


Figure 4.11: Overview of TT and T1-3 tracking stations. The purple (cyan) components are made of silicon tracker (straw tube drift chamber).

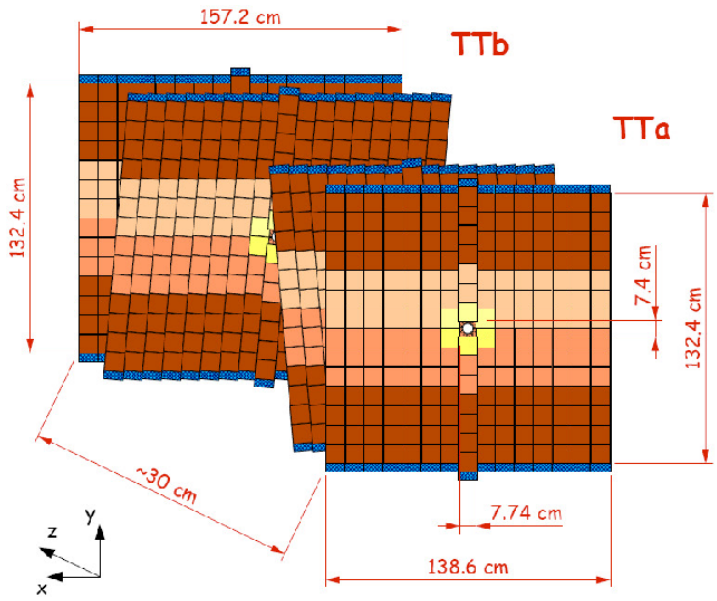


Figure 4.12: Layout of the 4 layers making up the TT tracker.

injection, the VELO has the ability to open while beams are injected and to close the halves when beams are declared stable. An overview of the VELO can be seen in Fig. 4.10.

Tracking stations TT, T1-3

An overview of the Trigger Tracker (TT) and the T1-3 stations can be seen in Fig. 4.11. The TT and inner part of the T1-3 stations (cross-shaped area around the beam pipe in stations T1-T3) use silicon microstrip detectors with a strip pitch of approximately 200 μm. Silicon pads allow a very fine spatial resolution but are

expensive. Therefore, they are only used in regions where the density of particles is high, those close to the beam pipe. Each station is made of four detection layers, the second and third layers are rotated respectively by an angle of $-5°$ and $5°$ as can be seen in Fig. 4.12. This specific layout allows to reach a spatial resolution of about 50μm in the inner region.

In the outer regions, the particle density is less and the detector features a straw tube drift chamber technology, which is cheaper but has a coarser spatial resolution of about 200μm. The outer trackers are also made of 4 layers arranged with the same layout as for the inner tracker.

Tracking algorithm

At the level of raw data, tracks don't exist yet and only a collection of hits in the different detectors is seen. The track reconstruction is performed in two stages: first, an algorithm searches for the detector hits which are most likely coming from the same particle making a track candidate. Then the track candidates are fitted into tracks. The same particle can be associated to several track candidates. These multiples candidates are called clones and are eliminated by running a clone killer algorithm selecting only the best clone. Additionally, some tracks, called ghost, are also created from random combinations of hits in the detector.

4.2.3 Particle identification

The Particle IDentification (PID) system is composed of the RICH detectors, the calorimeters and the muon stations.

The RICH system

The LHCb detector includes two Ring Imaging Cherenkov detectors (RICH): one located between the VELO and the TT (RICH1) and a second one located right after the T stations (RICH2). They are mainly devoted to tell apart pions and kaons but can also provide important information about protons, muons and electrons. They take benefit of the Cherenkov radiation to access the particle velocity v, which, combined with the momentum, given by the tracking system, allows to measure the mass of the particle (and consequently to identify the particle).
When a charged particle travels in a dielectric medium at a speed greater than the speed of light in such medium, it emits Cherenkov light following a cone of aperture θ_c along the particle trajectory.

$$\cos\theta_c = \frac{c}{nv}, \tag{4.5}$$

where c is the speed of light and n the medium refraction index. The measured angle versus the measured particle's momentum is shown in Fig. 4.13. The RICH1 covers the full angular acceptance of LHCb and is used to identify particles whose momenta are in the range [1, 60] GeV/c. It uses an aerogel and fluorobutane gas (C_4F_{10}). The RICH2 is used for particles in the range [15, 1000] GeV/c and has a more limited coverage as high momentum particles tend to propagate forward. It uses a CF_4 gas medium.

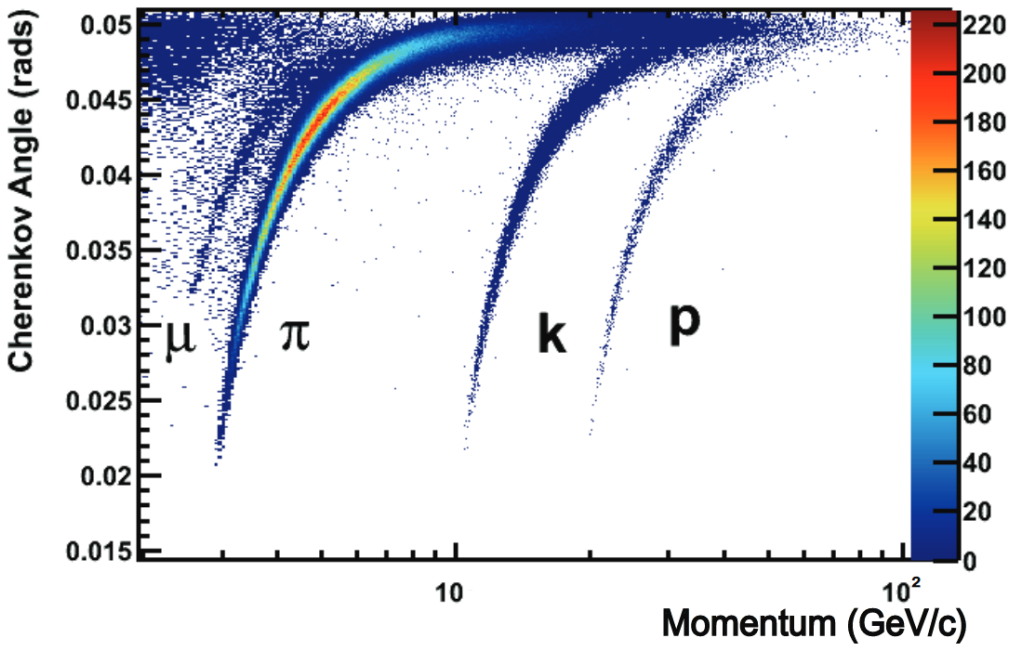


Figure 4.13: Cherenkov angle with respect to the measured particle momentum for different particles in the RICH1.

The Calorimeter system

The calorimeters purpose is to collect and measure the energy of a particle. They are the only sub-detectors allowing the observation of neutral particles like photons or neutral pions (which are invisible to the RICH, trackers and are not deflected by the magnet). As they stop all particles but muons and neutrinos, they are placed downstream of the tracking system (but upstream of the muon stations). The calorimeter system is composed of different sub-detectors that can be seen in Fig. 4.14. We find from up to downstream position:

- The Scintillating Pad shower Detector (SPD), in which only charged particles depose energy, hence its role is mainly to tell apart electrons from photons.
- A thin layer of lead of 12 mm, in order to initiate the electromagnetic shower.
- The Pre-Shower (PS), inside which electrons and photons partly deposit their energy.
- The Electromagnetic Calorimeter (ECAL), where electrons and photons finally deposit all their energy. The ECAL is made of interleaved layers of dense material (lead) and scintillator tiles. The particle interacts with the dense material producing showers, then the scintillator tiles emit UV light when interacting with the showers. The amount and the pattern of collected UV light provide information on the particle energy (but also on its direction and identity). The ECAL energy resolution is about $(1 \pm 10/\sqrt{E})\%$ for E from 15 to 1000 GeV.
- The Hadronic Calorimeter HCAL has a similar structure as the ECAL but the absorber is made of iron instead of lead. The resolution varies from 23 to 12% for momenta between 15 and 100 GeV.

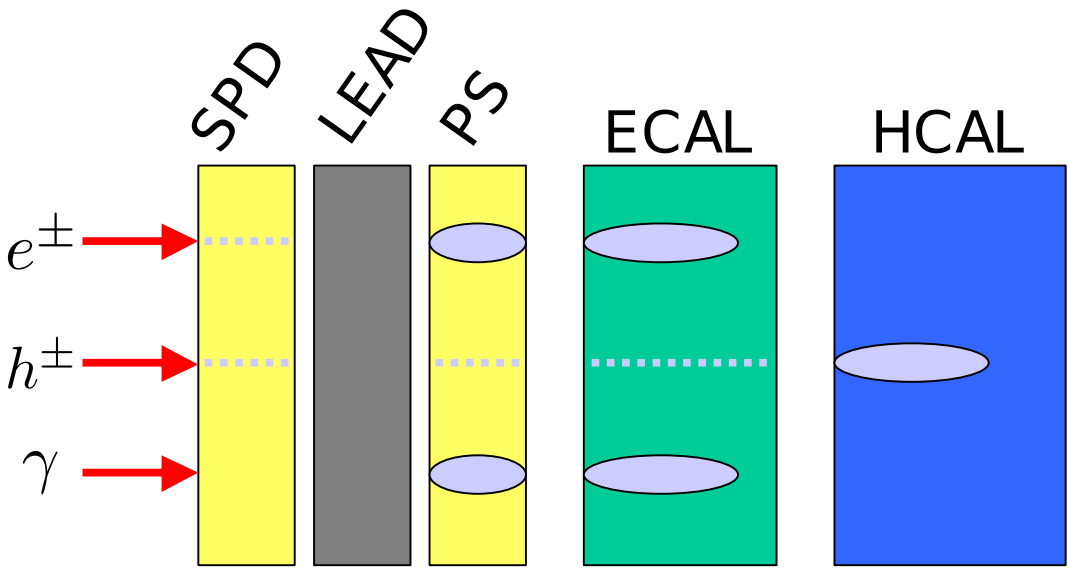


Figure 4.14: The main layers of the calorimeter devices which allow to measure the energy and tell apart photons, electrons and hadrons.

The Muon stations

Muons cross the full detector without much interaction. LHCb features an excellent muon detection system. This latter is composed of five muons stations, one station before the calorimeters (M1) and four (M2-5) after them. Muon first station's purpose is to improve the transverse momentum resolution. Each station is divided into four regions, R1 to R4, with increasing distance from the beam axis as can be seen in Fig. 4.15. Their granularity decreases from R1 to R4 in order to keep occupancy approximately constant over the station. Nevertheless, the granularity is higher in the horizontal plane, in order to give an accurate measurement of the muon momentum. The stations are interleaved with iron absorber so that only muons with momenta larger than 5 GeVcan penetrate through all chambers.
Information must be delivered to the trigger within 20 nanoseconds, so the detectors are optimized for speed. The system is therefore equipped with Multi Wire Proportional Chambers except the region R1 of the station M1 which is made of Triple-GEM detectors. This type of detector has indeed better aging properties and can support the high particle rate in this region. The Muon stations allow a muon identification efficiency of 97% for a misidentification rate ($\pi \to \mu$) of (1-3)%.

4.2.4 Trigger system and the stripping selection

The trigger

Most of the events produced at the LHC are not relevant for physics analyses (e.g.: many of them don't contain heavy flavor hadrons in the acceptance of the LHCb detector). A trigger is then implemented to select and record online only interesting events. LHCb features a two-level trigger. The first one is implemented at the hardware level (L0) and it reduces the amount of data from 40 MHz (the bunch crossing frequency) to 1 MHz. Of this 1 MHz, 0.45 (0.4) MHz are events triggered by hadrons (muons). The hardware trigger uses basic fast available calorimeter and muon system sub-detector information.
At this lower rate, all sub-detectors can be read and a partial reconstruction is

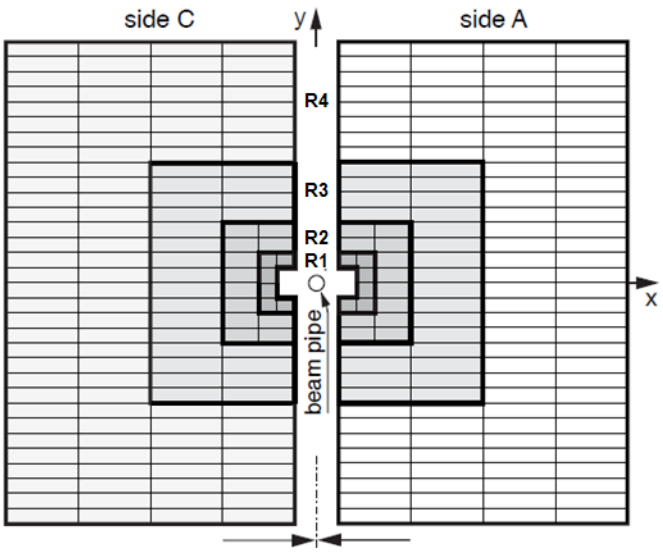


Figure 4.15: The different regions used in the muon stations.

performed. A software trigger called High level trigger (Hlt) runs on a computer farm called Event Filter Farm (EFF) and performs a final selection by reducing the final rate to 12.5 (5) kHz for Run2 (Run1). A scheme of the trigger system for Run1 and Run2 is shown in Fig. 4.16.

Two triggers conditions are used:

- *Triggered Independently of the Signal (TIS)*: the event fires the specified trigger line independently of the features of the signal considered.
- *Triggered on Signal (TOS)*: the signal properties are enough to fire the specified trigger line.

More informations about the LHCb trigger system in Run1 and Run2 can be found respectively in [116] and [111].

Hardware trigger (L0)

The hardware trigger stage (L0) requires events to have a hadron, photon or electron with high transverse energy in the calorimeters or a muon with high transverse momentum with respect to the beam line (p_T). For hadrons, the transverse energy threshold slightly depends on the data-taking conditions but is around 3.5 GeV. The L0 level is composed of two independent triggers called L0-Calo and L0-Muon. The L0-Calo triggers events if the calorimeter detects at least one hadron, photon or electron with a large transverse energy (E_T). The L0-Muon triggers events if the muon stations detect one or two muon candidates with a large transverse momentum (p_T).

High level trigger

The high level trigger refines the selection performed by the L0 trigger by selecting events which have a secondary vertex (which is the case for B and D mesons). For such a search, the event need to be reconstructed. Given the L0 output rate and the

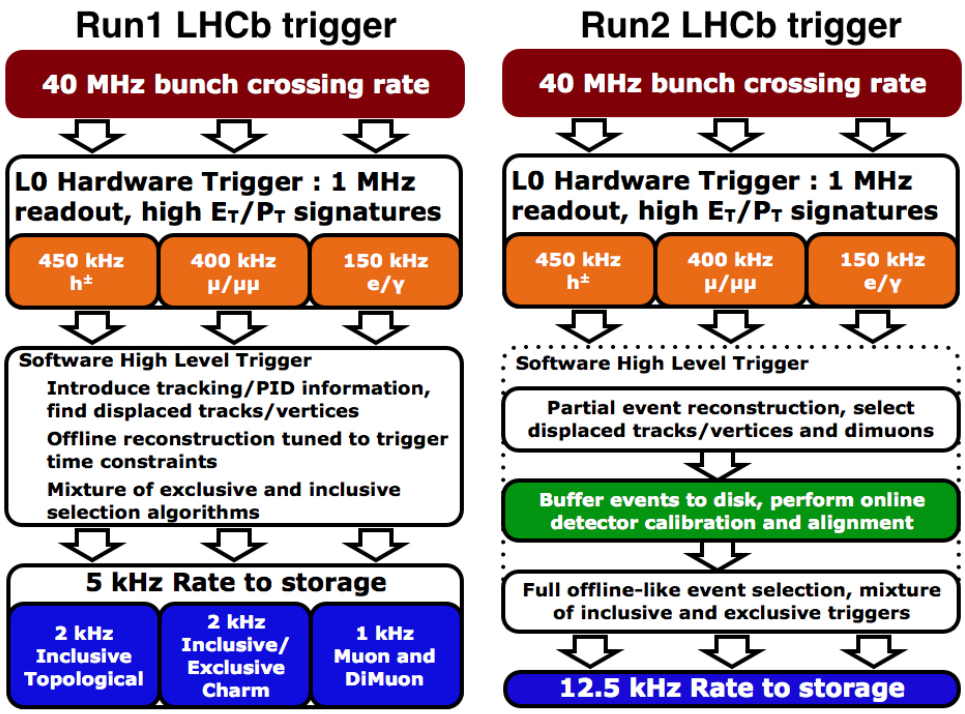


Figure 4.16: Diagrams of the trigger system for Run1 (left) and Run2 (right).

EFF resources, the event cannot be fully reconstructed in real time. Hence, the Hlt triggers is split into two: the Hlt1 level, that runs in real time and the Hlt2 level that works off-line.
The Hlt1 refines the selection by combining VELO and tracking station's information and reduces the event rate down to 40kHz. Selected events are then stored on the EFF machines.
The Hlt2 completes the reconstruction when more EFF resources are available (typically when there is no beam). Our analysis uses the *topological trigger*, which requires a two-, three- or four-track secondary vertex with a significant displacement from any PV. A multivariate classifier [117] is used for the identification of secondary vertices that are significantly displaced from the PVs, and are consistent with the decay of a b-hadron. At least one charged particle must have $p_T > 1.7$ GeV/c and be inconsistent with originating from any PV.
Finally, selected events are written on tape.

The stripping

A preselection of potentially interesting signals, called *stripping*, is performed as part of the candidate reconstruction process. It is applied after the complete off-line reconstruction. A stripping line typically corresponds to the reconstruction and selection of a D or B hadron decay. It allows to save space and provides a fast and easy access to data for the analysts.

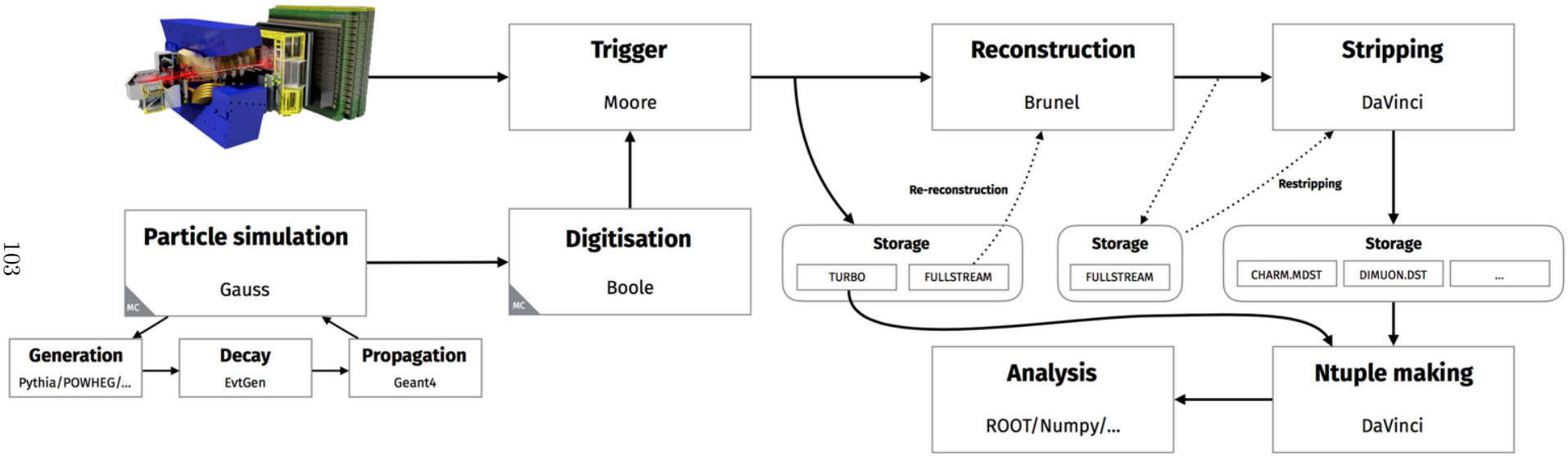


Figure 4.17: The LHCb data flow (real and simulated) with the software implemented at each step.

4.2.5 Simulation

This chapter on the LHCb experiment would not be complete without a paragraph on the simulation chain, intensively used in all analyses. Indeed Monte Carlo (MC) simulations are used to understand the data, reproduce the sought signal but also understand the detector's response. A schematic view of the data flow can be seen in Fig. 4.17. First, the pp collision and the hadronization process are simulated using generators. The main one used in LHCb is Pythia [118]. Then, other modules (eg: EvtGen [119], Tauola [120]) simulate the decay and propagation of the particles. In order to save space and computing power, a set of cuts are applied on the simulated event, the so-called *generator level cuts*. The detector's responses (hit digitization, showers, Cherenkov light,..) are simulated using Geant4 [121].

4.3 Experimental prospects

In 2018, the Run2 data taking has ended and LHCb has entered in a double stage upgrade phase in order to increase the instantaneous luminosity. In parallel, the Belle II experiment based on an e^+e^--collider started to collect data (without the vertex detector for the moment).

Sec. 4.3.1 gives a comparison of the two types of colliders and associated detectors: e^+e^- colliders and pp colliders. Sec. 4.3.2 describes the future of the LHCb detector.

4.3.1 e^+e^- colliders versus pp colliders

Two main strategies exist to produce b-hadrons at large scale: e^+e^- colliders and hadron colliders.

e^+e^- colliders

e^+e^- colliders, also called B factories, run mainly at the center-of-mass energy matching the $\Upsilon(4S)$ resonance mass, which is just above twice the B^0 mass. At this energy, they produce only pairs of $B^0\bar{B}^0$ and B^+B^- mesons. They can also run at the $\Upsilon(5S)$ resonance allowing to produce B_s^0 mesons, but at the cost of lower yields. Contrary to hadronic machines where parton momenta remain unknown, the total energy of the event is known in e^+e^- colliders. This justifies the construction of a 4π detector and simplifies the analyses with missing particles in the final state, since the missing energy can be measured. On the other side, the cross-section $\sigma_{b\bar{b}}$ is extremely low ($\sim$1 nb), resulting in low statistics.

The first two factories were KEKB located in Japan with the Belle experiment [122, 123], and PEP-II at the SLAC National Accelerator Laboratory with the BaBar experiment [124]. BaBar and Belle have finished collecting data in 2008 and 2010 respectively. An upgrade of KEKB and Belle, namely SuperKEKB and Belle II, started collecting data in 2018 and should finish the data acquisition period by the end of 2027. At the end of its operation, Belle II would have collected 50 ab^{-1}.

To summarize, e^+e^- colliders allow a relatively easy reconstruction of the B decay in an extremely clean environment, and with a trigger efficiency of almost 100% but with lower statistics. They are also optimized to the study of B^0 and B^+ decays although they can also produce B_s^0 mesons.

	LHC Run	Period of data taking	$\mathcal{L}$ [cm^{-2}s^{-1}]	L [fb^{-1}]
Initial detector	1 & 2	2011-2012, 2015-2018	4×10^{32}	9
Phase-I upgrade	3 & 4	2021-2013, 2026-2029	2×10^{33}	50
Phase-II upgrade	5 & ...	2031-2033, 2035-...	2×10^{34}	300

Table 4.2: Summary of LHCb running and data taking conditions for the initial experiment, the Phase-I and the Phase-II upgrades.

Hadron colliders

Another completely different strategy is to produce B mesons in a hadronic environment. This allows to produce all type of B hadrons in very large amounts but their reconstruction is challenging due to busy underlying events ($\sim$200 tracks per events in LHCb).

The Tevatron ($p\bar{p}$) at Fermilab in the USA with the D0 [125] and CDF experiment [126] has demonstrated that B-physics was possible at hadron colliders. The LHCb, CMS and ATLAS experiments are the only experiments based at hadron colliders to study B-physics in service nowadays.

4.3.2 The future of the LHCb detector

In addition to the Run1 and Run2, the LHC plans to deliver data during at least 3 additional runs, Run3-5. During the LS3 [2024-2026], the High-Luminosity LHC (HL-LHC) will be installed increasing the LHC nominal luminosity by a factor $\sim$10. While the ATLAS and CMS collaboration will enter into the HL-LHC era at the beginning of Run4, the LHCb collaboration will potentially enter the high luminosity phase at the beginning of Run5.

A first phase, called the Phase-I upgrade, is currently under installation and will allow to take data at a luminosity of 2×10^{33} cm^{-2}s^{-1}. Then a second phase, the Phase-II upgrade that should be installed in 2030 will allow the experiment to run at the spectacular luminosity of 2×10^{34} cm^{-2}s^{-1} [127]. An overview of the LHCb running and data taking conditions is given in Tab. 4.2.

The phase-I upgrade

The hardware trigger (L0) will be replaced by a full software trigger, which will allow the experiment to reconstruct every event at the rate of 40MHz. This new software trigger and higher luminosity will provide higher efficiency for hadronic final states. Most of the sub-detectors will be replaced during this upgrade: a new pixel VELO detector, a new silicon tracking station before the magnet (UT) and a large scale downstream Scintillating Fiber tracker will be installed.

The phase-II upgrade

If approved, a phase-II upgrade will be achieved in order to record data at the luminosity of 2×10^{34} cm^{-2}s^{-1}. The key point, at such high luminosity, is to handle a μ of $\sim$50. To disentangle the combinatorial background, a promising solution would be a precise track timing information from the VELO but also downstream of the magnet. A resolution a few tens of ps per particle will allow charged tracks

to be associated to the correct interaction vertex. Additionally, the installation of a tungsten re-sampling electromagnetic calorimeter is foreseen, which would increase the detector's capabilities for decays involving photons, electrons, π^0 and η mesons. At the end of Run5, the experiment will have accumulated 300 fb^{-1}.

Chapter 5

Introduction to the search for the $B^0_{(s)} \to \tau^+\tau^-$ decays

In this thesis, two studies based on two different final states have been performed to improve the experimental limit on the $B^0_{(s)} \to \tau^+\tau^-$ branching ratio. One study is based on the $B^0_{(s)} \to \tau^+(\to 3\pi^\pm\bar{\nu_\tau})\tau^-(\to \mu^-\bar{\mu_\nu}\nu_\tau)$ final state and the other one on the $B^0_{(s)} \to \tau^+(\to 3\pi^\pm\bar{\nu_\tau})\tau^-(\to 3\pi^\mp\nu_\tau)$ final state. For ease of reading, the final states will be referred in the following as the $(3\pi, \mu)$ and the $(3\pi, 3\pi)$ final state respectively. Both studies are similar on many aspects and share several technical tools. It is hence natural to describe their common features in this introductory chapter.

Sec. 5.1 presents some basic properties of the τ leptons and the previous results on the search for $B^0_{(s)} \to \tau^+\tau^-$ decays. Sec. 5.2 gives an overview of the first LHCb analysis. Background and variables used in the current study are described in Sec. 5.3. Finally, Sec. 5.4 presents the Multivariate Analysis (MVA) techniques.

5.1 General context

According to the SM, the $B^0_{(s)} \to \tau^+\tau^-$ decay branching ratio is higher than the $B^0_{(s)} \to e^+e^-$ and $B^0_{(s)} \to \mu^+\mu^-$ decay ones, as visible in Tab. 1.2. Nevertheless, the experimental search for the tauonic mode is more recent than the searches for the two lighter lepton modes. The reason why this search started later is due to the challenge that represents the final state reconstruction and its separation from the background.

Sec. 5.1.1 presents the τ lepton properties. The different searches performed so far for the $B^0_{(s)} \to \tau^+\tau^-$ decays are presented in Sec. 5.1.2.

5.1.1 On the τ lepton properties

The τ lepton lifetime is relatively brief as visible in Tab. 5.1. Even with a high boost, the τ lepton propagates only on a very short distance before decaying. In LHCb, the typical distance is $\sim$3 mm. As a consequence, the detector only detects the τ daughter particles. This would not be a problem in itself if each daughter could be directly detected. The problem arises from the presence of at least one ν_τ per τ decay (and possibly other neutral particles), which escapes detection.

Quantity	Value
τ lifetime (fs)	290.3 ± 0.5
τ mass (MeV/c^2)	1776.86 ± 0.12
Decay	Branching ratio measurement [%]
$\tau^- \to \pi^-\pi^0\nu_\tau$	25.94 ± 0.09
$\tau^- \to \pi^-\nu_\tau$	10.82 ± 0.05
$\tau^- \to \pi^-\pi^0\pi^0\nu_\tau$	9.26 ± 0.10
$\tau^- \to \pi^+\pi^-\pi^-\nu_\tau$	9.31 ± 0.05
$\tau^- \to \pi^+\pi^-\pi^-\pi^0\nu_\tau$	4.62 ± 0.05
$\tau^- \to \mu^-\bar{\nu_\mu}\nu_\tau$	17.39 ± 0.04
$\tau^- \to e^-\bar{\nu_e}\nu_\tau$	17.82 ± 0.04

Table 5.1: Overview of some τ properties: mass, lifetime, its main decay modes and their measured branching ratios [128].

This partially reconstructed final state implies a deterioration of the discriminating power of all variables related to the kinematic of the decay, and thus a loss of sensitivity. In our particular case, where two τ leptons are presents in the final state, this loss of information is particularly large and variables such as the τ lifetime or the di-τ invariant mass $m_{\tau\tau}$ are deteriorated.

Since the τ leptons decay before being detected, we have to choose which of their decay modes is the most suitable for an analysis at LHCb. The τ main decay modes and the value of their measured branching ratios are reported in Tab. 5.1. Three decays are very abundant: the hadronic decay $\tau^- \to \pi^-\pi^0\nu_\tau$ and the two purely leptonic decays $\tau^- \to e^-\bar{\nu_e}\nu_\tau$ and $\tau^- \to \mu^-\bar{\nu_\mu}\nu_\tau$. At first sight, these last two modes do not seem ideal to reconstruct the $B^0_{(s)} \to \tau^+\tau^-$ decay in LHCb since the neutrinos in the final state (and the neutral pion in the case of $\tau^- \to \pi^-\pi^0\nu_\tau$) will result in missing energy and LHCb is not able to measure the missing energy. Nonetheless, LHCb is very efficient at identifying and triggering muons. Hence, one of the study presented in this thesis is focused on the potential of reconstructing one of the τ leptons decaying into a muon and two neutrinos.

The $\tau^- \to \pi^+\pi^-\pi^-\nu_\tau$ decay is less abundant but gives more handles to reconstruct the τ lepton since the final state consists of 3 charged tracks and only one invisible particle. Additionally, this decay presents a feature which is intensively used in the analyses presented in this thesis: the decay proceeds predominantly through the intermediate resonances $\rho^0(770)$ and $a_1(1260)$ [47]:

$$\begin{aligned}\tau^- \to & a_1^-(1260)\nu_\tau \\ & \hookrightarrow \pi_1^-\ \rho^0(770) \\ & \qquad \hookrightarrow \pi_2^+\pi_3^- .\end{aligned}$$

Hence, the 3π and the $\pi_2^+\pi_3^-$ invariant mass should respectively match the mass of the $a_1^-(1260)$ and $\rho^0(770)$ resonances. Some properties of these resonances are given in Tab. 5.2.

Resonances	Mass (MeV/c^2)	Γ (MeV/c^2)
$a_1^-(1260)$	1230 ± 40	[250-600]
$\rho^0(770)$	775.26 ± 0.25	149.1 ± 0.8

Table 5.2: Values of the masses and decay widths (Γ) of the intermediate resonances for the $\tau^- \to \pi^+\pi^-\pi^-\nu_\tau$ decay [128].

5.1.2 Previous results

So far, only the BaBar and the LHCb experiments have searched for $B^0_{(s)} \to \tau^+\tau^-$ decays.

BaBar results

The BaBar experiment has been the first one to publish a result on the $B^0 \to \tau^+\tau^-$ decay in 2006. The data sample consists of 210 fb^{-1} collected at the energy of the $\Upsilon(4S)$ resonance. The $\Upsilon(4S)$ decays in equal proportion to $B^0\bar{B}^0$ and B^+B^- pairs. To avoid the mis-assignment of decay products to the parent B, the BaBar analysis completely reconstruct one B candidate in each event and searches for the signal decay among the remaining detected particles. Signal events that are consistent with each τ decaying to a single charged particle (and one or two ν) are selected. The resulting upper limit, published in [129], is

$$\mathcal{B}(B^0 \to \tau^+\tau^-) < 4.10 \times 10^{-3} \quad \text{at 90\% C.L.} . \tag{5.1}$$

LHCb results

In 2016, no experimental results for the $B^0_s \to \tau^+\tau^-$ mode existed. Nevertheless its branching ratio was indirectly constrained to be less than 3% at 90% C.L. [130–132]. Using Run1 data, LHCb has obtained a first experimental limit on the B^0_s mode and the best world limit on the B^0. For this analysis, both τ leptons were reconstructed through their 3-prong decay into three charged pions and one neutrino. Due to the missing particles in the final state, the invariant mass resolution does not allow the separation of the B^0 and B^0_s mesons. The limit on the B^0_s mode is thus obtained assuming no contribution from the B^0 and vice versa. The upper limits, published in [45], are:

$$\mathcal{B}(B^0_s \to \tau^+\tau^-) < 5.2\ (6.8) \times 10^{-3} \quad \text{at 90 (95)\% C.L.} , \tag{5.2}$$

$$\mathcal{B}(B^0 \to \tau^+\tau^-) < 1.6\ (2.1) \times 10^{-3} \quad \text{at 90 (95)\% C.L.} . \tag{5.3}$$

Both studies performed for this thesis are the continuation of this analysis, referred in the following as the *published LHCb analysis*. It is then necessary to give a more detailed overview of it, as done in Sec. 5.2.

5.2 Overview of the published LHCb analysis

The published analysis workflow can be divided into three main steps: the event selection, the fit of the signal yield and the transformation of the fit result into an upper limit on the $B^0_{(s)} \to \tau^+\tau^-$ branching ratio.
In many B-physics analyses, the number of signal events is obtained from a fit to the

reconstructed B mass. In our case, the B mass reconstruction is complicated by the presence of two neutrinos, originating from the τ decays. As a consequence, the B mass provides only a weak discrimination between signal and background and cannot be used as a fit variable. Instead, the number of signal candidates is obtained from a fit to the output of a multivariate classifier.
Simulated data are used to optimize the selection, obtain the signal model for the fit and determine the selection efficiencies. Data-driven methods are used to determine background models used for the fit and for the selection optimization. They rely on a subdivision of the data in different regions described in the following.

5.2.1 Analysis regions

As said in section 5.1.1, the τ leptons are reconstructed in the 3π final state, which proceeds predominantly through the intermediate resonances $\rho^0(770)$ and $a_1(1260)$ [47]. This implies that τ events form a plus-like shape in the $(m_{\pi_1^+\pi_2^-}, m_{\pi_3^+\pi_2^-})$ scatter plot, as illustrated in Fig. 5.1. In the following we will refer to this plot as a "Dalitz plot"[1]. It is divided into nine rectangles, using two cuts on each of the two invariant mass variables. These nine regions are referred to by the position on a classic dial pad, as illustrated in Fig. 5.1. As an example, region "5" is defined as

$$\text{box 5} = \left(m_{\text{low}} \leq m_{\pi_1^+\pi_2^-} \leq m_{\text{high}}\right) \,\&\&\, \left(m_{\text{low}} \leq m_{\pi_3^+\pi_2^-} \leq m_{\text{high}}\right) . \tag{5.4}$$

The values of m_{low} and m_{high} are identical for $m_{\pi_1^+\pi_2^-}$ and $m_{\pi_3^+\pi_2^-}$ but differ between the event selection and the fit due to a re-optimization of the signal search window. Their values for the event selection and the fit are given respectively in Eq. (5.5) and (5.6).

$$m_{\text{low/high}} = 775 \pm 181 \text{ MeV/c}^2, \tag{5.5}$$

$$m_{\text{low/high}} = 775 \pm 160 \text{ MeV/c}^2. \tag{5.6}$$

Based on the above described division the τ^+ and τ^- Dalitz plots, three data subregions are built:

Signal region: (τ^+ in box 5) && (τ^- in box 5).
These events are considered to be the most signal-like, and will therefore define the to-be-fitted data.

Background region: (τ^+ in boxes 1, 3, 7 or 9) || (τ^- in boxes 1, 3, 7 or 9).
At least one of the two τ candidates must end up in one of the four corners of its Dalitz plot. These data events are considered to be the least signal-like, and are considered background when optimizing the signal selection.

Control region: (τ^+ in boxes 4, 5 or 8) && (τ^- in boxes 4 or 8), or vice versa.
These data events provide the background model for the fit.

[1] A Dalitz plot is classically defined in terms of the squares of these masses.

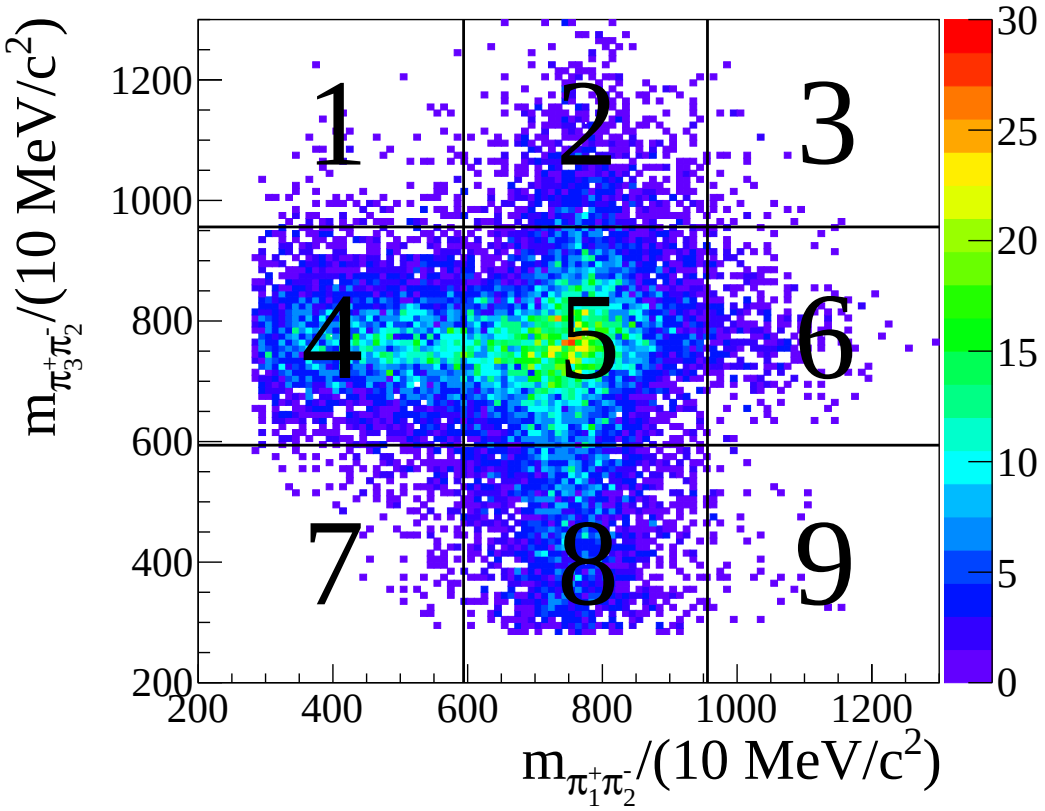


Figure 5.1: Illustration of the "Dalitz"-like plot [45] for simulated $B_s^0 \to \tau^+\tau^-$ events, used in this analysis to divide the data into subregions. The values $m_{\mathrm{low/high}} = 775 \pm 181$ MeV/c^2 are used.

5.2.2 Analysis workflow

Selection

The event selection is made of several steps:

1. The first step of the selection is an online selection performed by the trigger described in Sec. 4.2.4.

2. The second step is imposed by the stripping described in Sec. 4.2.4.

3. The third step consists of a loose, cut-based selection. Eleven variables are identified from a long list of (highly) discriminating variables that are individually able to remove more than 50% of the candidates in data while still maintaining a signal efficiency of more than 95%. The list of variables and their cut values are optimized using only the $B_s^0 \to \tau^+\tau^-$ and $B^0 \to \tau^+\tau^-$ MC samples, and aim to retain as much signal as possible by only rejecting the most obvious backgrounds.

4. The fourth step of the selection is based on a MVA output set to keep 80% of the signal according to the simulation. This selection is optimized on MC for signal and data in the background region for background.

At the end of the full selection process, only 10187 data events are present in the signal region. The full selection efficiency averaged over 2011 and 2012 is $(14.21\pm0.38)\times10^{-6}$ for the $B_s^0\to\tau^+\tau^-$ MC signal.

Likelihood fit

To determine the number of signal events in the data signal region, μ, a binned maximum likelihood fit is performed to the output of a second MVA. This second MVA is a Neural Network built using the NeuroBayes setup [133]. The signal model is taken from the MC simulation while the background model is taken from the data in a control region, keeping under consideration that a sizable fraction of signal can be present in the control region. The fit model is given below by

$$\mathcal{N}_{\text{data}}^{\text{SR}} = s\times\widehat{\mathcal{N}}_{\text{sim}}^{\text{SR}} + f_b\times\left(\mathcal{N}_{\text{data}}^{\text{CR}} - s\times\frac{\epsilon^{\text{CR}}}{\epsilon^{\text{SR}}}\times\widehat{\mathcal{N}}_{\text{sim}}^{\text{CR}}\right), \tag{5.7}$$

where $\mathcal{N}_{\text{data}}^{\text{SR}}$ ($\mathcal{N}_{\text{sim/data}}^{\text{CR}}$) is the MVA output distribution in the signal (control) region from simulation/data, s is the signal yield, f_b is a scaling factor for the the background template, and ϵ^{SR} (ϵ^{CR}) is the signal efficiency in the signal (control) region. The hat on top of the MVA output distributions means that the distributions are normalized to unity. The fit is set up using the HistFactory framework [49] and linked to the RooStats framework [134] for the limit calculation. Bin-by-bin systematic uncertainties are determined. The plot of the fit for $B_s^0\to\tau^+\tau^-$ can be seen in Fig. 5.2. The signal yields obtained for Run1 from the fit on data are:

$$s_{B_s^0} = -23^{+63}_{-53}(\text{stat})^{+41}_{-40}(\text{syst}), \tag{5.8}$$

$$s_{B^0} = -15^{+67}_{-56}(\text{stat})^{+44}_{-42}(\text{syst}). \tag{5.9}$$

Normalization and branching ratio limit

The observed number of $B^0\to\tau^+\tau^-$ events is related to the $B^0\to\tau^+\tau^-$ branching ratio through the normalization factor:

$$\begin{aligned}\mathcal{B}(B^0\to\tau^+\tau^-) =& \frac{\mathcal{B}(B^0\to\tau^+(\to 3\pi^\pm\bar{\nu_\tau})\tau^-(\to 3\pi^\mp\nu_\tau))}{\mathcal{B}(\tau^-\to\pi^-\pi^+\pi^-\nu_\tau)^2}\\ =& \frac{N^{\text{obs}}_{B^0\to\tau^+\tau^-}/\epsilon^{\text{tot}}_{B^0\to\tau^+\tau^-}}{\mathcal{B}(\tau^-\to\pi^-\pi^+\pi^-\nu_\tau)^2 N_{B^0}} \equiv \alpha_d\times N^{\text{obs}}_{\tau^+\tau^-}\,,\end{aligned} \tag{5.10}$$

where $\epsilon^{\text{tot}}_{B^0\to\tau^+\tau^-}$ is the total $B^0\to\tau^+(\to 3\pi^\pm\bar{\nu_\tau})\tau^-(\to 3\pi^\mp\nu_\tau)$ reconstruction and selection efficiency and N_{B^0} is the number of produced B^0 mesons in the data.

At hadron colliders, it is difficult to accurately determine the number N_{B^0} directly from the measured luminosity. The normalization factor α_d is therefore determined using a second B decay, the normalization mode, which is chosen for its (relatively) high yield and similarity to the signal. For this analysis, the normalization channel is the $B^0\to D_s^+(K^+K^-\pi^+)D^-(\pi^+\pi^-K^-)$ decay, which has the same number of charged tracks in the final state and a similar decay topology with $B^0\to\tau^+(\to 3\pi^\pm\bar{\nu_\tau})\tau^-(\to 3\pi^\mp\nu_\tau)$. It has a relatively high branching ratio of $(3.9\pm0.6)\times10^{-5}$ and has previously been studied by the LHCb collaboration [135, 136]. The normalization factor is then given by

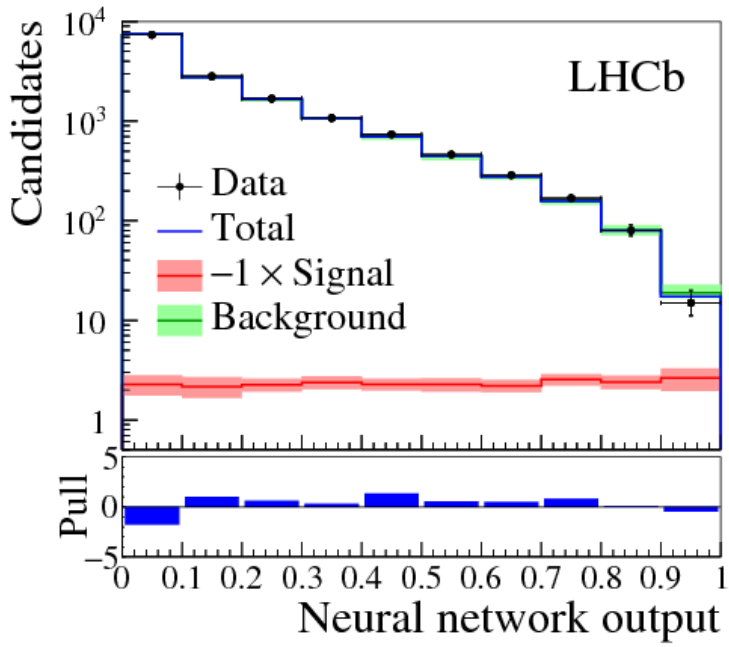


Figure 5.2: Distribution of the NN output in the signal region (black points) with the total fit result (blue line) and the background component (green line). The fitted $B_s^0 \to \tau^+\tau^-$ signal component is negative and is therefore shown multiplied by -1 (red line). For each bin of the signal and background component the combined statistical and systematic uncertainty on the template is shown as a light-colored band. The difference between data and fit divided by its uncertainty (pull) is shown underneath.

$$\alpha_d = \frac{\epsilon^{D^-D_s^+} \cdot \mathcal{B}(B^0 \to D^-D_s^+) \cdot \mathcal{B}(D^+ \to \pi^+K^-\pi^+) \cdot \mathcal{B}(D_s^+ \to K^+K^-\pi^+)}{N^{\text{obs}}_{D^-D_s^+} \cdot \epsilon^{B^0\to\tau^+\tau^-} \cdot [\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau)]^2}, \quad (5.11)$$

where $\epsilon^{D^-D_s^+}$ is the total $B^0 \to D_s^+(K^+K^-\pi^+)D^-(\pi^+\pi^-K^-)$ reconstruction and selection efficiency In the case of the B_s^0 meson, the normalization factor α_s must be divided by the ratio of B_s^0 to B^0 production factor $f_s/f_d = 0.259 \pm 0.015$ [137] and the efficiency $\epsilon_{B^0\to 6\pi^\pm}$ must be replaced by $\epsilon_{B_s^0\to 6\pi^\pm}$.

The normalization factor α and the expected SM yield expressed by:

$$N^{\text{SM}}_{B^0_{(s)}} = \frac{\mathcal{B}(B^0_{(s)} \to \tau^+\tau^-)^{\text{SM}}}{\alpha_{d/s}}, \quad (5.12)$$

are given in Tab. 5.3.

Upper limits

Yields for $B_s^0 \to \tau^+\tau^-$ and $B^0 \to \tau^+\tau^-$ are compatible with zero, uppers limit are set on their branching ratios. This is done using the CLs method [138], using the RooStats framework [134]. Due to the broad mass resolution, it is not possible to distinguish $B^0 \to \tau^+\tau^-$ from $B_s^0 \to \tau^+\tau^-$ decays. The limit is therefore set on the $B_s^0 \to \tau^+\tau^-$ branching ratio under the assumption of zero $B^0 \to \tau^+\tau^-$ events (and vice versa). The published upper limits are given in Eq. (5.3).

Mode	α	SM Yield $N^{\mathrm{SM}}_{B^0_{(s)}}$
$B^0_s \to \tau^+(3\pi)\tau^-(3\pi)$	$(3.16 \pm 0.43) \times 10^{-5}$	0.0245 ± 0.0037
$B^0 \to \tau^+(3\pi)\tau^-(3\pi)$	$(0.94 \pm 0.16) \times 10^{-5}$	0.0024 ± 0.0004

Table 5.3: Overview of the normalization factor α and expected SM yield for the two signal channels.

5.3 Signal reconstruction and background rejection

Two final states are used in this thesis to reconstruct the $B^0_{(s)} \to \tau^+\tau^-$ signal, the $(3\pi, \mu)$ and $(3\pi, 3\pi)$ final states. They share a large set of variables that help in discriminating signal from background.

Signal reconstruction for both final states is described in Sec. 5.3.1. The main backgrounds for the $(3\pi, 3\pi)$ and the $(3\pi, \mu)$ final state are respectively described in Sec. 5.3.2 and 5.3.3. The description of the discriminating variables used in this thesis is reported in Sec. 5.3.4.

5.3.1 Signal reconstruction

Analysis with the $(3\pi, \mu)$ final state

A schematic view of the $(3\pi, \mu)$ final state is given in Fig. 5.3. Only the B production vertex and one of the two three-prongs τ decay vertex can be directly reconstructed. Candidates are reconstructed in events having at least three charged π tracks and one μ track in the detector acceptance. The total charge of the three π has to be -1 or +1. The total net charge of the 4 tracks has to be 0. We can breakdown the $B^0_{(s)} \to \tau^+\tau^-$ candidates reconstruction into two steps:

1. The τ candidate is built with three charged tracks previously identified as π and coming from the same well defined (with a low χ^2 value) vertex. This vertex has to be a secondary vertex, meaning it has to be well displaced with respect to any PV in the event. The invariant mass of this 3π system ($m_{3\pi}$) has to be compatible with the mass of the a_1 resonance, visible in Tab. 5.2. An approximate momentum, equal to the vectorial sum of the momenta of the three π tracks is associated to the τ.

2. The $B^0_{(s)} \to \tau^+\tau^-$ candidate is built using a reconstructed τ and a charged track identified as a μ. An approximate momentum, equal to the vectorial sum of the momenta of the μ and the τ candidates is associated to the B candidate.

This final state is expected to have a larger trigger efficiency and a larger reconstruction efficiency than the $(3\pi, 3\pi)$ final state due to the excellent performance of the muon trigger and the requirement of only 4 charged tracks in the detector acceptance.

On the other side, one of the two τ decay vertices cannot be reconstructed. Additionally, large backgrounds from semileptonic B decays can pollute the signal region. A more detailed study is then needed to figure out the potential of the $(3\pi, \mu)$ final state, which is described in Chap. 6.

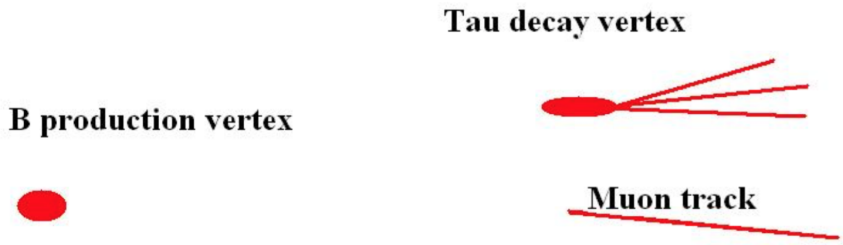


Figure 5.3: Schematic view of a $B^0_{(s)} \to \tau^+(\to 3\pi^\pm \bar{\nu_\tau})\tau^-(\to \mu^- \bar{\nu_\mu}\nu_\tau)$ decay in the LHCb detector. Courtesy of A. Morda [139].

Analysis of the Run2 data with the $(3\pi, 3\pi)$ final state

A schematic view of the $(3\pi, 3\pi)$ final state is given in Fig. 5.4. The B production vertex and, this time, both τ decay vertices can be directly reconstructed. Candidates are reconstructed in events having at least six charged π tracks in the detector acceptance with a total net charge of zero. We can breakdown the $B^0_{(s)} \to \tau^+\tau^-$ candidates reconstruction into two steps:

1. The τ candidate is built with three charged tracks previously identified as π and coming from the same well defined secondary vertex. The invariant mass of both 3π systems have to be compatible with the mass of the a_1 resonance. An approximate momentum, equal to the vectorial sum of the momenta of the three π tracks is associated to each τ candidate.

2. The $B^0_{(s)} \to \tau^+\tau^-$ candidate is built using two reconstructed τ candidates. The momenta of both τ leptons (sum of the momenta of the 3 pions) are used to reconstruct the position of the $B^0_{(s)}$ meson decay vertex. A visible momentum, equal to the vectorial sum of the 6 pions is calculated for the B candidate.

The possibility to reconstruct both τ decay vertices and the B production vertex gives access to the plane of the decay and make possible the reconstruction the two τ four-momenta. The branching ratio of the three-prong τ decay suppresses the process by a factor:

$$\frac{\mathcal{B}(\tau^- \to \mu^- \bar{\nu_\mu}\nu_\tau)}{\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau)} \sim 1.9 \tag{5.13}$$

with respect to the $(3\pi, \mu)$ final state. As visible in Tab. 4.1, the cross-section $\sigma_{b\bar{b}}$ in Run2 as well as the integrated luminosity are doubled with respect to those of Run1. Hence, we expect in Run2 roughly 4 times more B mesons produced than in Run1. This analysis is described in Chap. 7.

5.3.2 Main backgrounds for the $(3\pi, 3\pi)$ final state

Misidentified $\tau \to 3\pi\nu_\tau$ candidates

Both final states have at least one τ reconstructed from the decay $\tau^- \to \pi^+\pi^-\pi^-\nu_\tau$. Two main sources of background can mimic a $\tau \to 3\pi\nu_\tau$ decay:

- *purely combinatoric* τ made with at least one random π track forming a common vertex.

Figure 5.4: Schematic view of a $B^0_{(s)} \to \tau^+(\to 3\pi^\pm\bar{\nu_\tau})\tau^-(\to 3\pi^\pm\nu_\tau)$ decay in the LHCb detector. Courtesy of A. Morda [139].

	Mass (MeV)	τ (fs)
τ	1776.86 ± 0.12	290.3 ± 0.5
D^0	1864.84 ± 0.05	410 ± 1.5
$D^\pm$	1869.5 ± 0.05	1040 ± 7
$D_s^\pm$	1969.0 ± 1.4	504 ± 4

Table 5.4: Some properties of the τ lepton compared to various D mesons [128].

- *partially reconstructed D meson*: $D^{(0,\pm)}_{(s)} \to 3\pi^\pm X$ with X being a generic hadron. The D meson represents the most abundant decay product of b-hadrons due to the high $b \to cW^-$ transition rate. Additionally, although the masses of D mesons are slightly higher than the mass of the τ lepton, the presence of additional non reconstructed hadron in the final state can shift the reconstructed invariant mass of the three π towards the mass of the a_1 resonance. Moreover we can also have real a_1 resonances from the $D_{(s)}$ decay.

The majority of the purely combinatoric τ is removed by requirements on the τ decay vertex χ^2 (defining the quality of the vertex) and on the 3π or 2π invariant masses, $m_{3\pi}$ and $m_{2\pi}$. $m_{3\pi}$ and $m_{2\pi}$ are indeed more likely to match respectively the masses of the a_1 and the ρ^0 resonances if the pions are coming from a real $\tau \to 3\pi\nu_\tau$ decay.

Removing the *partially reconstructed D meson* background is more difficult. The criteria on the three π common vertex χ^2 become ineffective since the 3 pions are really originating from the same vertex. The requirements on the 3π and 2π invariant masses remove most of the $D^\pm \to 3\pi X$ decays but not the one that proceed via the a_1 and ρ^0 resonances. The key variable to discriminate this background is the τ lifetime. Indeed $D^\pm$ mesons have a longer lifetime than the τ lepton as shown in Tab. 5.4.
These fake τ candidates are just part of the background composition. Combined to other type of backgrounds, they enter the full $B^0_{(s)} \to \tau^+(3\pi)\tau^-(3\pi)$ background composition described below.

$B^0_{(s)} \to \tau^+(3\pi)\tau^-(3\pi)$ fake candidates

The high number of tracks in the final state allows a lot of processes to mimic the signal:

- $B^0_{(s)}$ candidate made of two misidentified-τ leptons,
- random combinations of τ and partially reconstructed D mesons coming from decays of two different b-hadrons in the events,
- physical backgrounds generated by the semileptonic decay of the B meson into a true τ and a partially reconstructed D or into two misidentified D mesons, also referred to as *exclusive background*.

Most background candidates coming from combinatorics are removed by requirements on the τ lifetime, the τ and B decay vertex χ^2 and other variables related to the B reconstructed candidate. The most difficult background to suppress comes from exclusive decays of the B meson. In this case, the τ and D mesons are coming from the same parent particle and previous requirements becomes less efficient.

5.3.3 Main backgrounds for the $(3\pi, \mu)$ final state

Additionally to the $\tau \to 3\pi\nu_\tau$ fake candidates described in Sec. 5.3.2, we have to deal with another type of fake τ decay:

$\tau \to \mu\nu_\mu\nu_\tau$ fake candidates

The main source of background that can mimic a $\tau \to \mu\nu_\tau\bar{\nu_\mu}$ decay is decays of the type $D^{(*,\pm)}_{(s)} \to X\mu\nu_\mu$, which will be referred to as $D \to X\mu$ *background*.

$B^0_{(s)} \to \tau^+(3\pi)\tau^-(\mu)$ fake candidates

- $B^0_{(s)}$ candidate made of a purely combinatoric τ and μ coming from a b-hadron decay,
- random combinations of τ and D mesons coming from decays of two different b-hadrons in the events, where either τ or D decays to μ.
- physical backgrounds generated by the semileptonic decay of the B meson in a true τ and a D or into two D mesons (e.g. $B \to D\tau\nu_\tau$ or $B \to DD$),
- physical background coming from semileptonic decay of type $B \to D^{(*,**)}(\to 3\pi(X))\mu\nu_\mu$ with X being a generic hadron,
- physical background coming from purely hadronic decay of type $B \to D(\to (Y)\mu\nu_\mu)3\pi(X)$. Here the three π can also proceed through a_1 and ρ^0 resonances.

Most background coming from combinatorics is removed by requirements on the τ lifetime, τ and B decay vertex χ^2 and on variables related to the B reconstructed candidate. The most difficult backgrounds to suppress comes from the two last physical backgrounds, the purely hadronic and the semileptonic exclusive decays of the B meson.

5.3.4 Discriminating variables

Besides the standard kinematic and geometric variables introduced in previous sections, both analyses make extensive use of vertex and tracks related variables, variables related to the candidate multiplicity, the B meson corrected mass and custom-made

isolation variables. As mentioned previously, a partial reconstruction of the two τ four-momenta is possible for the $(3\pi, 3\pi)$ final state and gives access to a large set of reconstruction variables. Also, some variables cannot be used due to a bad data-MC agreement as explained below.

Data-MC agreement

For various reasons, some variables are not well simulated. They can be identified by comparing the distribution for data and MC simulations. We need for this comparison a decay mode whose branching ratio is large and whose signal can be extracted easily from background. In our analyses, the normalization channel is used for this comparison. All variables whose data-MC disagreement is too large are discarded from the analyses.

Vertex and tracks related geometrical variables

Vertex positions are reconstructed by fitting together the reconstructed tracks that are compatible with originating from the same point. The quality of the reconstructed vertices is quantified by the fit χ^2. To quantify the distance between two reconstructed vertices, for instance between a PV and a candidate B meson decay vertex, we define the two following variables:

- "flight distance (FD)" which is the distance between two vertices,
- "flight distance χ^2 (FD χ^2)" which can be seen as the ratio of the flight distance over its related uncertainty.

In addition to the flight distance, we also define:

- the "impact parameter" with respect to a reconstructed PV. This is defined as the Distance of Closest Approach (DOCA) of the reconstructed PV and the track. Decay vertices of b or c-hadron are well displaced with respect to the PV, hence tracks originating from such decays are expected to have a large value of the IP.
- the "impact parameter χ^2" (IP χ^2) is defined for each reconstructed track as the difference in the χ^2 of the related PV when including or not the considered track to the set of those used to fit the PV. For b or c-hadron decay, the difference in the χ^2 is expected to be large.

We also use:

- the "θ_{DIRA}" variable, defined as the cosine of the angle between the momentum of the particle and the flight direction from the best PV to the decay vertex.
- the "θ_{DIRA} ORIVX" variable, defined as the cosine of the angle between the momentum of the particle and the flight direction from its vertex of origin.

Track isolation variables

The isolation variables measure the presence of other particles in the vicinity of the selected tracks or their displaced vertices. With more activity near the signal candidate, it becomes less likely that the candidate is a true $B \to \tau^+\tau^-$ decay.

Three custom-made track isolation variables are used in the selection. The construction of these variables was initially developed for the $B_s^0 \to \mu^+\mu^-$ analysis. The isolation level is derived either on a cut-based strategy or on the outcome of a Boosted Decision Tree (BDT), which has been optimized for the $B_s^0 \to \tau^+(3\pi)\tau^-(3\pi)$ search. For the training of this BDT, the non-signal tracks in the event are divided into two categories: *non-isolating* tracks, i.e. they make the signal track non-isolated, versus *isolating* tracks, i.e. they leave the signal track isolated. The non-isolating tracks are all tracks coming from displaced B and D decay vertices that are part of the same true decay chain as the signal track. The isolating tracks are all remaining tracks, primarily coming directly from the PV or other B/D meson decays, which are essentially unrelated to the signal track.
For each non-signal track the common vertex $\mathcal{V}$ with the signal track is defined as the midpoint between the two tracks at their point of closest approach [2]. The BDT combines the following seven discriminating variables:

- The minimum IP χ^2 of the track with respect to any PV.
- The transverse momentum of the track
- The angle between the track and the signal track
- The quantity:
$$f_c = \frac{|P_S + P_{tr}|\alpha}{|P_S + P_{tr}|\alpha + P_{TS} + P_{Ttr}} \,, \tag{5.14}$$
where P and P_T are the momentum and transverse momentum of the tracks, S identifies the signal track, tr identifies the considered track, α is the angle between the vectorial sum of the combined track and the signal track momenta and the vector between the PV and the vertex $\mathcal{V}$.
- The distance of closest approach of the track and signal track.
- The distance between the vertex $\mathcal{V}$ and the B decay vertex.
- The distance between the vertex $\mathcal{V}$ and the PV.

The BDT is trained using the non-isolating tracks as background target, and the isolating tracks as signal target. Isolating tracks (high BDT values) are preferred, whereas non-isolating tracks (low BDT values) are harmful.
For each long track in the event which is not part of the signal candidate, the BDT response is calculated. Let a, b and c be the number of long tracks with a BDT value smaller than -0.09, -0.05 and 0, respectively, the track isolation variables are defined as:

- *BDT Iso 1:* The sum $a + 100 \times b + 10000 \times c$.
The values of a, b and c are recovered for the selection as the 1st digit (a), 3rd digit (b) and 5th digit (c) of BDT Iso 1.
- *BDT Iso 2:* The sum $\sum_{\mathrm{BDT}(x)<-0.05} \mathrm{BDT}(x)$ of BDT values for all tracks with a BDT output smaller than -0.05.
- *BDT Iso 3:* The sum of BDT Iso 2 and the minimal BDT value of all tracks with a BDT output in the range $[-0.05, 0]$.

[2] An actual refit of this vertex for every track in the event is too CPU intense, and not practical.

Vertex isolation variables

Three vertex isolation variables are used in the selection. For their construction, the tracks making up the τ or B candidate's decay vertex $\mathcal{V}$ are combined with a single other track from the event, and fitted together into a new vertex $\mathcal{V}^*$. The vertex isolation variables are then defined as

- *NumVtxWithinChi2WindowOneTrack:* The number of other tracks in the event for which the χ^2 of the new vertex fit is less than nine, $\chi^2_{\mathcal{V}^*} < 9$.
- *SmallestDeltaChi2MassOneTrack:* The invariant mass of the tracks making up the τ or B candidate and the track leading to the smallest difference in χ^2

$$\Delta\chi^2 \equiv \chi^2_{\mathcal{V}} - \chi^2_{\mathcal{V}^*}, \tag{5.15}$$

  between the $\mathcal{V}$ and $\mathcal{V}^*$ vertex fits.
- *SmallestDeltaChi2MassTwoTracks:* The invariant mass of the tracks making up the τ or B candidate and the two tracks leading to the smallest difference in χ^2 between the $\mathcal{V}$ and $\mathcal{V}^{**}$ vertex fits, where $\mathcal{V}^{**}$ is constructed iteratively from the $\mathcal{V}^*$ vertex that itself has the smallest $\Delta\chi^2$.

Isolation from neutral objects

Four isolation variables that target neutral objects are used in the selection. For their construction, a cone of size 0.5 in pseudo-rapidity η and polar angle ϕ around the B candidate is defined. The size of this cone was optimized by looking at the signal efficiency versus background rejection, and by comparing the ranking of different cone sizes in the preprocessing stage of the NeuroBayes algorithm [133]. The B neutral cone (nc) isolation variables are then defined as

- *nc-Mult:* The number of neutral objects found inside the cone.
- *nc-vPT:* The vector-summed transverse momentum of the neutral objects inside the cone.
- *nc-sPT:* The scalar-summed transverse momentum of the neutral objects inside the cone.
- *nc-PZ:* The z component of the total momentum of the neutral objects inside the cone.

Candidate multiplicity

The $\tau \to 3\pi$ reconstruction algorithm looks for all combinations of three charged tracks identified as a π and forming a common vertex with a net charge $e = \pm 1$. Due to the high boost of the B meson, the angle between the momenta of the two τ leptons is very small. Hence the two τ decay vertices are almost aligned with the B primary vertex and a π track coming from one vertex can be reconstructed as coming from the other. The six final state tracks can be arranged in several ways in group of three and this creates more than one B candidate per event.
Additionally, extra charged tracks close enough to the τ candidate's decay vertices can also enter the combinatorics increasing again the candidate multiplicity. Although the $(3\pi, \mu)$ final state has less tracks in its final state, it is also sensitive to combinatorics

and most events have more than one candidate. The variable "totCandidates", also referred as candidate multiplicity, which is the number of candidates per event is used in the selection for both final states.

B meson corrected mass

The selection uses standard LHCb variables, except for the corrected mass variable which was used for the first time in the SLD experiment [140] to correct for the missing energy from neutral and/or undetected particles in the decay chain. It includes the transverse momentum relative to the flight direction as the minimum missing momentum to the invariant mass of the charged tracks, such that the momentum vector and flight direction align. This variable is calculated as[3]

$$M_{corr} = \sqrt{M_B^2 + \left(P_B \times \sqrt{1-\theta_{DIRA}^2}\right)^2} + P_B \times \sqrt{1-\theta_{DIRA}^2}\,, \tag{5.16}$$

where M_B is the invariant mass calculated from the charged tracks, and θ_{DIRA} is the cosine of the angle between the momentum of the B and its flight direction.

Pion variables

The three pion candidates are labeled 1, 2 and 3. By definition, the pion 2 has the same sign than its mother (one of the τ leptons). There is no physical reason to use a variable associated to a particular pion rather than its neighbors. Hence, we use in this study some combinations over the three pions like the sum, the minimum or the maximum of a given observable.

$B_s^0 \to \tau^+(3\pi)\tau^-(3\pi)$ reconstruction variables

A method for the full reconstruction of the $B_s^0 \to \tau^+(3\pi)\tau^-(3\pi)$ decay has been developed in Ref. [139] based on geometrical information from measurable quantities of the decay, combined with mass constraints for the particles in the decay chain. With this method, the reconstruction of the two τ momenta is equivalent to finding the roots of a 4th degree polynomial:

$$\mathcal{P}^{(4)}(\xi) = \sum_{i=0}^{4} a^{(i)}(\theta)\xi^i = 0\,, \tag{5.17}$$

where the explicit expressions for the coefficients $a^{(i)}(\theta)$ as well as the definition of the unknown parameters ξ are shown in Appendix D of Ref. [139], and θ is a rotation angle physically related to the asymmetry in the triangle formed by the PV and the two τ decay vertices. The solutions of the fourth degree equation are in the form

$$\xi_1 = a - ib\,, \quad \xi_2 = a + ib\,, \quad \xi_3 = c - id\,, \quad \xi_4 = c + id\,, \tag{5.18}$$

with $a, b, c, d \in \mathcal{C}$.

The angle θ is not known and has to be approximated. The approximated value of the θ parameter is used to compute the coefficients of the polynomial in Eq. (5.17)

[3]The expression $M = \sqrt{m^2 + p^2} + p$ represents the invariant mass of a particle decaying into a particle with mass m and a massless particle.

and to solve for the unknown ξ. Once the correct solution is chosen among the ξ_i, the two (approximate) τ momenta can be reconstructed, as well as the topology of the decay. The knowledge of the τ momenta allows to reconstruct their common origin vertex and the B candidate momentum.

All the kinematic variables such as the masses of the τ mesons or the B meson, as well the ones of the two ν have already been imposed as external constraint in the derivation of the Eq. ((5.17)) and thus cannot be reconstructed. Nevertheless, in the various steps of the reconstruction algorithm, several variables which are functions of observable quantities appear and turn out to be useful in discriminating signal against background. These variables do not have a trivial immediate meaning, being much more related to the mathematics needed to derive the relevant equations. Among these variables, the most discriminating ones, which are further considered in this analysis, are [139]:

- $\mathcal{R}e[x_3]$: $Re(|\vec{\tilde{p}}_+(\theta = \bar{\theta}, \xi_3)|)$
- $\mathcal{I}m[\tilde{s}_1^{\pm}]$: $Im(|\vec{\tilde{p}}_+(\theta = \theta^\star, \xi_1)|)$
- $\mathcal{I}m[x_1]$: $Im(\tilde{p}_+(\theta = \theta^\star, \xi_1)\tilde{p}_-(\theta = \theta^\star, \xi_1))$
- $\mathcal{I}m[\tilde{s}_3^{\pm}]$: $Im(\tilde{p}_+(\theta = \theta^\star, \xi_3)\tilde{p}_-(\theta = \theta^\star, \xi_3))$
- $\mathcal{R}e[\tilde{s}_1^{\pm}]$: $Re(\tilde{p}_+(\theta = \pi/4, \xi_1)\tilde{p}_-(\theta = \pi/4, \xi_1))$
- $\mathcal{R}e[\xi_1]$
- $\bar{\theta}$
- $p_+ \cdot p_-$

where $\bar{\theta}$ and $\theta^\star$ are different approximations of the angle θ, $s \equiv p_+ \cdot p_-$ (for signal events $s = \frac{M_B^2 - 2m_\tau^2}{2}$) and $\tilde{p}_\pm^2 \equiv M_\pm^2 = m_\tau^2 \pm \sin(2\theta)s$, $p_\pm$ being the four momenta of the $\tau^\pm$.

5.4 Multivariate analyses

This section is devoted to the MVA techniques which are intensively used in this thesis. An introduction is given in Sec. 5.4.1. Two types of MVA, the Boosted Decision Tree (BDT) and the Artificial Neural Network (ANN), are presented respectively in Sec. 5.4.2 and 5.4.3. The k-folding and the flattening are detailed respectively in Sec. 5.4.4 and 5.4.5. Finally, a method for selecting variables entering a MVA is detailed in Sec. 5.4.6.

5.4.1 Introduction to multivariate analysis techniques

Multivariate analysis is a generic term covering a wide range of techniques. It involves analysis of more than one variable at a time in order to take into account the effects of all variables on the response of interest. In physics analyses, it is often used for regression or for classification purposes. In our case, MVA's are used to design a binary classification (signal versus background). As mentioned in Sec. 5.2, MVA's are used in two steps of both analyses presented in this thesis, where a cut is

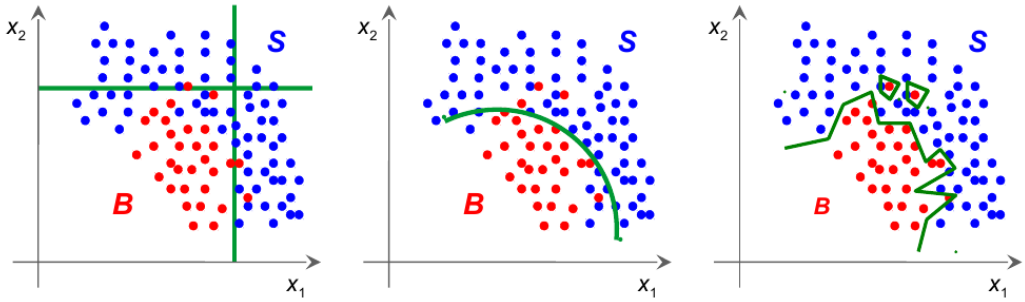


Figure 5.5: Scatter plot of two variables x_1 and x_2 with from left to right: a rectangular cut, a non linear cut, a non linear cut with a non analytic function.

applied to an MVA output, and in the likelihood fit.

A convenient way to introduce MVA technique is to provide a visual illustration. For clarity reason, we limit to two the number of variables in the following. Nevertheless, in real cases, a MVA classification has interest only when performed on more than two variables. We call x_1 and x_2 these two variables. Fig. 5.5 shows the scatter plot of the variables x_1 and x_2 for two kinds of population, in red the signal and in blue the background.

The most basic selection we can imagine is to apply two independent linear cuts on the variables. This is illustrated with the green straight lines in the left plot of Fig. 5.5. As we can see, the classification is not optimal and can be refined if we take into account the correlation of the two variables. In this simple case, a cut on the combination $\sqrt{x_1^2 + x_2^2}$ has a stronger discriminating power than two independent linear cuts applied on the two variables. This is illustrated with the green curve in the middle plot of the same figure.

This is the idea behind MVA classification, although it is performed with $D > 2$ variables. In this case, the $(x_1,\ x_2)$ plan becomes a space of dimension D and the green curves an hyper-surface of dimension $D - 1$. In other terms, a MVA classification consists in building a new variable χ, function of D variables, $\chi = \mathcal{F}(x_1, x_2, ..., x_D; w_i)$ which separates efficiently two populations. This function F is not necessary analytic. It depends on a set of weights w_i. The optimization of these weights is referred to as *training* or *machine learning*. Two families of training methods exist, the supervised and the non-supervised one. In this thesis, only supervised training are used. The term "supervised" means that we previously need to tell the machine what is signal and what is background. Technically, we provide as input to the MVA two training samples, one of pure signal and another one of pure background. The weights w_i are then optimized based on these samples.

Both, ANN and BDT, have no limitation in learning how to discriminate the training samples and each event can be perfectly classified in the correct category as illustrated in the most right plot of Fig. 5.5. Unfortunately, training samples are statistically limited and the MVA can end up learning statistical fluctuation of these training samples. Once the MVA is applied to unknown samples, also called test samples, the classification error rate will not be zero anymore. This is not a problem

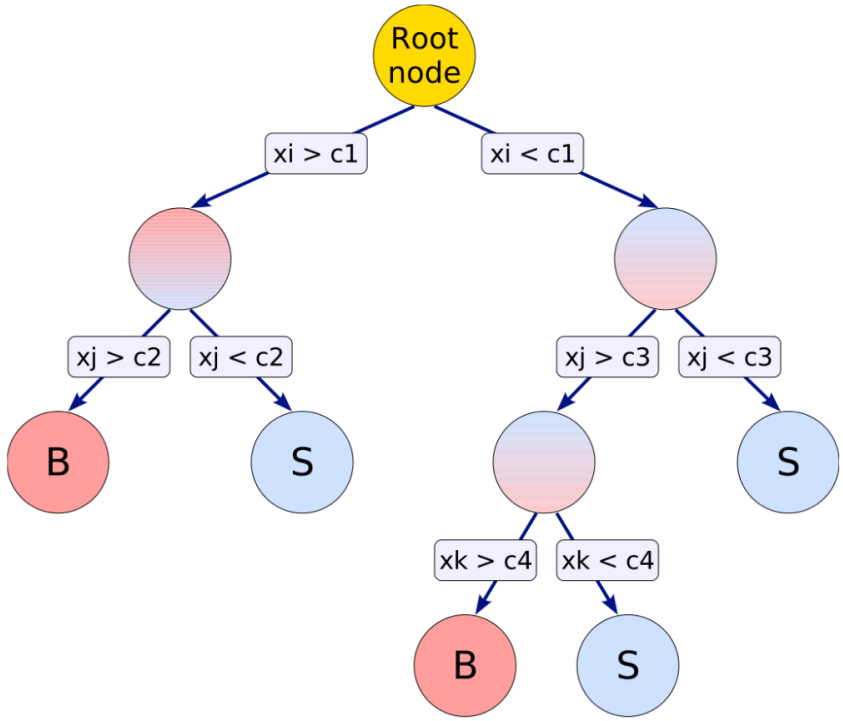


Figure 5.6: Decision Tree of depth 3. The leaves are tagged "S" or "B" for Signal or Background.

in itself as it is natural that the MVA knows better how to classify the samples on which it was trained than unknown samples.

The performance of the algorithm has to be evaluated on the test sample following a certain Figure of Merit (FoM) (the choice of the FoM depends on the algorithm's purpose). This FoM increases with the amount of training until a certain point where it reaches a plateau. In certain cases, the FoM can decrease after reaching its maximum, this effect, prejudicial for analyses, is referred as *overtraining* and has to be avoided.

5.4.2 Boosted Decision Tree

As said previously, we need to provide to the MVA two training samples, one of pure signal and another of pure background. From these samples, the BDT algorithm looks for the single variable cut maximizing the separation between signal and background. The samples are then split into two sub-samples and the process is repeated in each sub-sample. After d iterations, a decision tree (DT) of depth d is built and the last produced sub-samples are called leaves. If a leave contain mainly signal (background), it is tagged "Signal" ("Background"). An example of decision tree is shown in Fig. 5.6. An event passing through a DT is assigned a number of -1 (+1) if it falls in a background (signal) leave. After building the first DT, the training events are re-weighted in a way that the weights for previously misclassified events will be more important. This event weighting is called *boosting*. A collection of DT (called a forest) is then built and their outputs are combined by taking the weighted average of single output DT. At the end of the full process, a number between -1 and 1 is assigned to each event. An event whose output is close to -1 (+1) is most likely a

background (signal) event.

The main parameters controlling the BDT optimization are the maximal allowed depth of a single tree $d_{\max}$, the number of trees in the forest N_t and the minimal percentage of training events required in a leaf $MinNodeSize$.

The CPU consumption of a BDT scales as $n \times N \times \log N$ with N the number of events in the training sample and n the number of variables used in the BDT. The software used to build the BDT in this thesis is TMVA [48]. TMVA is a ROOT-integrated toolkit for multivariate classification and regression analysis. It performs the training, testing and performance evaluation of a large variety of multivariate methods such as BDT, but also ANN.

5.4.3 Artificial Neural Network

Artificial Neural Networks are vaguely inspired from biological neural networks. They consist of a collection of nodes (like the neurons in a biological brain) connected in a network structure. Each connection (like the synapses in a biological brain) can transmit information from one node to another one. A node, labeled j, receives from other nodes labeled i inputs o_i and processes it into a single output o_j. In turn, this output o_j can be used as input by other nodes directly connected to the node j.

The treatment of the input information by a node is done in two steps:

1. First, input information o_i are weighted by a weight factor w_i and combined together in the output $o_j = \sum_j w_i o_i$.
2. Second, the output o_j is passed through an *activation function* which transforms the range of the output $[-\infty, \infty]$ into [-1,1]. The Sigmoid function is used in our case:
$$S(o_j) = \frac{2}{1 + e^{-o_j}} - 1. \tag{5.19}$$

The analyses presented in this thesis make use of *feed-forward network* with a single output. In this type of network, neurons are arranged in layers: one input layer, one or several hidden layers and an output layer made of a unique node. Information is only passed unidirectionally between consecutive layers; there are hence no cycles inside the network. A schematic view of this neural network structure is shown in Fig. 5.7. The input layer has $n + 1$ neurons; one for each input distribution plus an additional one for a so-called bias term that is added to improve the learning performance.
The hidden layer can take any number of nodes, but the best performance is achieved with $\mathcal{O}(n)$ nodes. The purpose of training is to optimize the value of the weight factors in order to maximize the signal to background separation. The software used to build the ANN's in this thesis is NeuroBayes [133].

5.4.4 *k*-folding

If events on which the MVA is trained are used later on in the analysis, a bias can be introduced. It is typically the case of the MC signal sample which is used and re-used multiple times in the analysis to train the MVA's. To bypass this problem, the k-folding technique is used. It consists of randomly splitting the MC signal sample

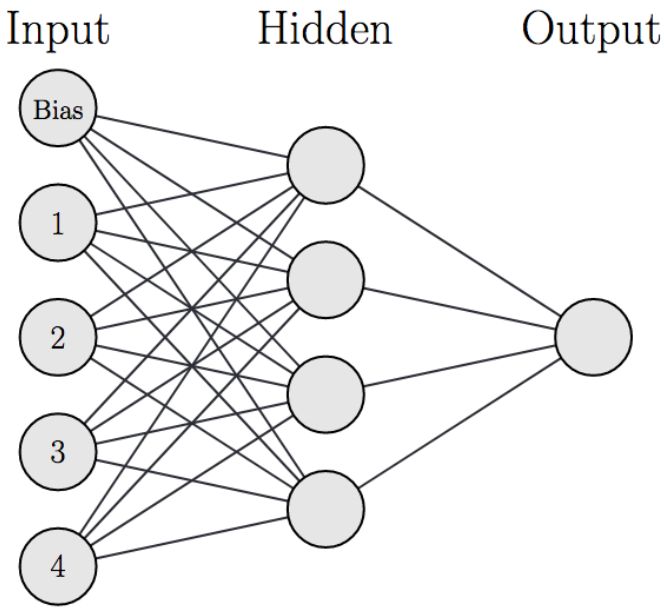


Figure 5.7: Layout of a feed-forward ANN.

into k equally-sized subsamples (folds). (k-1) folds are used as training data and then applied to the remaining one. No k-folding is applied on data but the k MVA outputs are averaged to produce a single output. Strictly speaking, the data events are also used in further steps of the analysis but the very small statistics of the data used in the BDT training compared to large statistics of the full data sample allows us to consider this bias as negligible.

5.4.5 Flattening

The output of a BDT or a ANN is a non-physical value in the range [-1, 1]. This means that the shapes for the signal and background distributions have no physical meaning and can vary between different folds. Before combining the information of the 10 trained MVA outputs, the distributions $F_{sig(bkg)}$ of the output variable χ of the MVAs are individually transformed through the change of variable:

$$\chi \rightarrow \chi_{flat} = \frac{\int_{-1}^{\chi} F_{sig}(y)dy}{\int_{-1}^{+1} F_{sig}(y)dy} \tag{5.20}$$

As a consequence, the variable χ_{flat} is uniformly distributed between [0,1] for signal, while it is peaking at zero for background events.

5.4.6 Iterative procedure to select input variables entering the BDT

The choice of the parameters controlling a MVA training is fundamental, but the MVA performance also depends in a large way on the choice of the input variables. We designed an iterative procedure to select these input variables for any MVA implemented in the TMVA framework. The method described in the following is also valid for NeuroBayes's ANN but has not been implemented in this framework.

The procedure consists of "picking up" variables one after the other among a list of "potentially usable variables", by testing different combinations. This list is called *list_all* in the following. We also define the list *list_picked* containing the variables that definitively enter the BDT. The *list_picked* is built iteratively by:

1. Training BDT for all variable combinations made of *list_picked* plus one variable from *list_all*.
2. Moving the variable which gives the best FoM from *list_all* to *list_picked*.
3. Repeating steps 1. and 2. to select one more variable.

Although this procedure is quite automatic, it allows some degrees of freedom: the choice of the initial list of variables, the choice of the FoM, the choice of the final number of variables, the choice of the BDT parameters, the choice of optimizing or not the BDT parameters during the selection procedure.

As an example, we show the result of this procedure for the BDT1 presented in Sec. 6.2.3. For reasons explained later, the chosen FoM is the area under the Receiver Operating Characteristic (ROC) curve[4] for the test sample. The initial list of variable is made of only one variable, the most discriminating one selected thanks to the TMVA criteria: "Separation" [5].
The FoM strongly increases at the beginning of the variable selection process and then keep increasing but at a slower pace as visible in Fig. 5.8. At this stage, adding more variables does not increase significantly the FoM. Finally, the FoM is expected to reach a plateau where adding more variables do not increase the BDT performance. Note that we stop the algorithm before the plateau for matter of CPU time. Furthermore each variable entering a MVA has a cost since it has to be corrected from possible data-MC disagreement. Additionally, possible uncertainties can also be associated to theses variables which can be prejudicial for the analysis. Hence, a compromise between the number of variables entering the MVA and the MVA performance has to be found. Finally, the BDT parameters are optimized only once all variables are selected also for matter of CPU time.

Validation study

Another method was previously used to select input variables for the BDT. This section aims at comparing this previous method to the iterative procedure used in this study.

The previous method was based on the following procedure:
• Rank all variables with respect to the criteria of TMVA "Separation".
• Remove variables with too high linear correlation. The correlation values produced by TMVA are used in this step. If two variables have a linear correlation coefficient above 70% and this coefficient is similar for both signal and background, then the variable with the lowest "Separation" power is removed from the list.

[4]The background rejection is projected on the y axis, the signal efficiency on the x axis.

[5]The "separation" $\langle S^2 \rangle$ of a variable y is defined by the integral $\langle S^2 \rangle = \frac{1}{2} \int \frac{(\hat{y}_S(y) - \hat{y}_B(y))^2}{\hat{y}_S(y) + \hat{y}_B(y)} dy$, where $\hat{y}_S(y)$ and $\hat{y}_B(y)$ are the signal and background PDFs of y, respectively. The separation is zero for identical signal and background shapes, and it is one for shapes with no overlap.

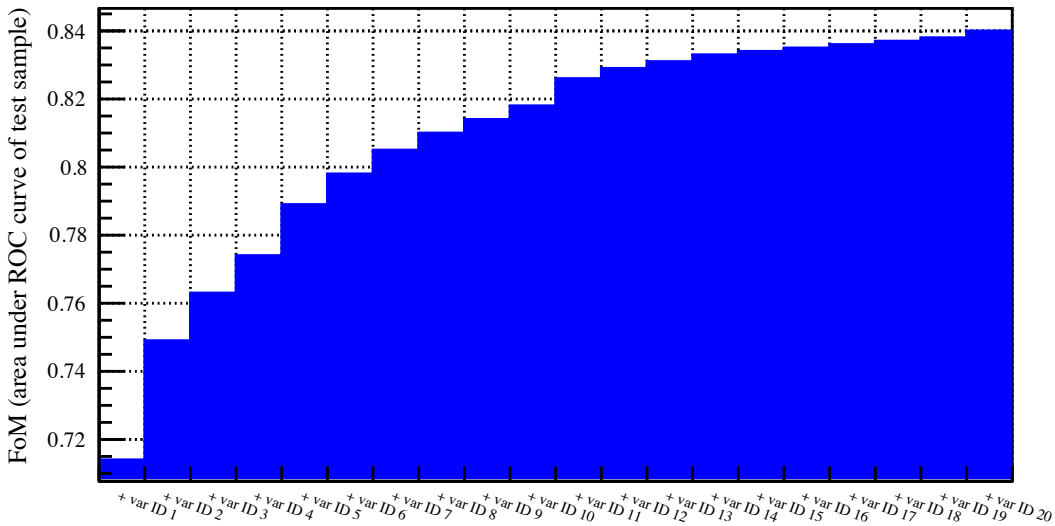


Figure 5.8: Evolution of the FoM (area under the ROC curve for the test sample) at each step of the variable selection procedure for the BDT1 described in Sec. 6.2.3.

Parameters	Value	Number of points
N_t	from 1 to 951, step of 50	20
$MinNodeSize$ [%]	from 0.1 to 2.9, step of 0.2 %	15
Total number of points		300 (15×20)

Table 5.5: Grid of parameters used to optimize the BDTs

The comparison has been realized for some BDTs used in this study, the iterative procedure described in this Section reaches systematically better results than the previous one. As an example, for BDT1, the iterative procedure returns a FoM of 0.8 with 8 variables versus 20 variables for the previous method. Moreover the previous method does not allow to reach the final FoM of 0.84 accessible with the iterative procedure.

5.4.7 Optimization of the BDT parameters

Once all variables are selected, we perform the BDT optimization. This optimization consists in finding the value of the BDT parameters giving the largest FoM (In this thesis the area under the ROC curve for the test sample is used as a FoM).
The main BDT parameters are the following: number of trees in the forest (N_t), maximum allowed depth of the trees ($d_{\max}$), minimum amount of data contained in a leaf in % ($MinNodeSize$). All BDT's used in this study use adaptive boosting, controlled by the parameter AdaBoostFactor (β).
As N_t and $d_{\max}$ are, in principle, extremely correlated we decided to set $d_{\max}$ to 3 and to optimize only the value of N_t. Moreover, as the influence of β is small, we decided to set it to 0.2. The FoM is computed for different values of N_t and $d_{\max}$ that can be seen in Tab. 5.5. We then interpolate the function of the FoM (N_t, $MinNodeSize$) and select parameters giving the maximum of this function for the test sample. In each BDT, the signal and background sample are split in the following way: 7/10 for training, 3/10 for testing.

Chapter 6

Search for $B^0_{(s)} \to \tau^+\tau^-$ via the $(3\pi, \mu)$ final state

This chapter describes the feasibility study performed to improve the experimental upper limit on the $B^0_{(s)} \to \tau^+\tau^-$ decay via the final state $B^0_{(s)} \to \tau^+(\to 3\pi^\pm\bar{\nu_\tau})\tau^-(\to \mu^-\bar{\mu_\nu}\nu_\tau)$, also referred to as the $(3\pi, \mu)$ final state. It is the first time this final state is studied in LHCb. The strategy of this study is similar to the one used in the published LHCb analysis [45] developed for the $B^0_{(s)} \to \tau^+(\to 3\pi^\pm\bar{\nu_\tau})\tau^-(\to 3\pi^\mp\nu_\tau)$ final state. As in the published analysis, the workflow can be split into three main steps: the selection, the fit and the normalization to translate the signal yield into a branching ratio measurement (or limit).

Sec. 6.1 gives the analysis overview and the main changes with respect to the published one. The selection, the fit and the normalization are respectively described in Sec. 6.2, 6.3 and 6.4. Finally, conclusions are drawn in Sec. 6.5.

6.1 Analysis overview

A comparison between the $(3\pi, 3\pi)$ and $(3\pi, \mu)$ final states has already been developed in Sec. 5.3. The two main differences are reminded below:

1. We can reconstruct only one τ decay vertex.
2. We have to suppress an abundant source of background coming from semileptonic decays.

Analysis regions

Having only one Dalitz plane available imposes a redefinition of the regions used in the analysis. We use a similar splitting of the τ-"Dalitz"-like plane as for the $(3\pi, 3\pi)$ final state, see Sec. 5.2.1 for a reminder. The τ-"Dalitz"-like plane is illustrated in Fig. 6.1. The values of m_{low} and m_{up} are identical for $m_{\pi_1^+\pi_2^-}$ and $m_{\pi_3^+\pi_2^-}$, and are the same for the event selection and the fit. They are set to

$$m_{\text{low}} = 640 \text{ MeV/c}^2 \text{ and } m_{\text{up}} = 900 \text{ MeV/c}^2. \tag{6.1}$$

Four data subregions are built from the subdivisions of the τ-"Dalitz"-like plane :
Signal region: τ in box 5.

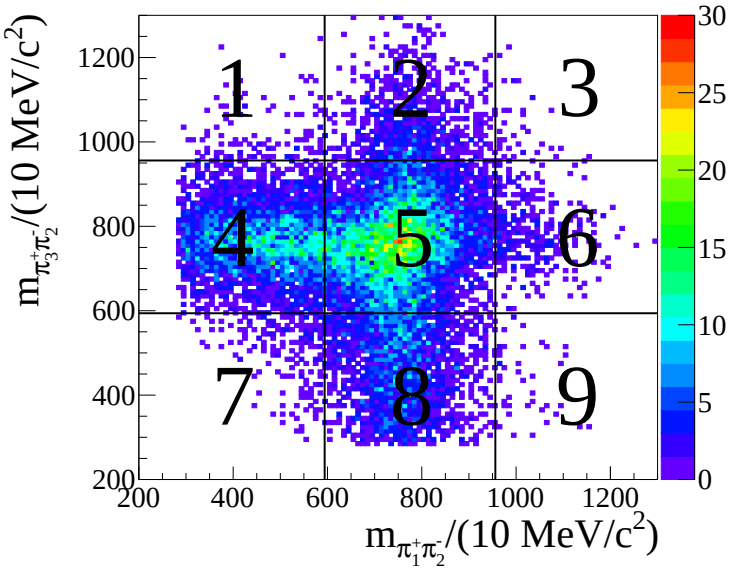


Figure 6.1: Reminder of the "Dalitz"-like plot for simulated $B_s^0 \to \tau^+\tau^-$ events used in the Run1 analysis of the $(3\pi, 3\pi)$ final state. This reminder allows us to define the subregions of data used in the $(3\pi, \mu)$ analysis .

Control region: τ in box 4 or 8.
Background region: τ in box 1, 3, 7 or 9.
Left-over region: τ in box 2 or 6.

Data samples

The analysis is performed using data corresponding to an integrated luminosity of 3 fb^{-1} of proton–proton collisions recorded with the LHCb experiment at center-of-mass energies of $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV in 2011 and 2012, respectively. A sample of Same Sign (SS) data was also available (unphysical candidates reconstructed with two τ of same sign, $\tau^+\tau^+$ or $\tau^-\tau^-$). The distributions of the SS data and the data in the background region are different for most of the variables, therefore the SS data are not used in this analysis.

Simulated data samples

Several MC samples are used to study the properties of the signal and background components, as well as those of the normalization sample. A brief description of the MC simulation chain is given in Sec. 4.2.5. An overview of the available set of MC simulated samples is given in Tab. 6.1, where it should be noted that, except for the signal $B^0_{(s)} \to \tau^+\tau^-$ and the normalization channel $B^0 \to D(\to K\pi\pi)\pi$, all samples are generated for the 2012 running conditions. In order to save CPU time, the samples are generated requiring that all particles in the final state be in the LHCb acceptance. In addition, all samples generated for the signal and exclusive backgrounds have been produced with the following generator level cuts on the daughter particles: $p_{\mathrm{T}}(\pi) > 250\,\mathrm{MeV}/c$, $p(\pi) > 2\,\mathrm{GeV}/c$ and $p_{\mathrm{T}}(\mu) > 250\,\mathrm{MeV}/c$. In order to study the background properties, several samples of exclusive B decays with at least 3 charged pions and a muon in the final state were produced. The exclusive samples list is not meant to be exhaustive but it is built to provide qualitative

Channel	Year	Yield
Signal		
$B_s^0 \to \tau^+(3\pi)\tau^-(\mu)$	2011	0.7 M
$B_s^0 \to \tau^+(3\pi)\tau^-(\mu)$	2012	1.4 M
$B^0 \to \tau^+(3\pi)\tau^-(\mu)$	2011	0.7 M
$B^0 \to \tau^+(3\pi)\tau^-(\mu)$	2012	1.4 M
Control/normalization channels		
$B^0 \to D^-\pi^+$	2011	0.5 M
$B^0 \to D^-\pi^+$	2012	1 M
Exclusive backgrounds		
$B^0 \to D^{*-}\pi^+\pi^-\pi^+$, $D \to \mu\nu$	2012	400k
$B^0 \to D^{*-}\pi^+\pi^-\pi^+\pi^0$, $D \to \mu\nu$	2012	400k
$B^0 \to D^-\pi^+\pi^-\pi^+$, $D \to \mu\nu$	2012	400k
$B^0 \to D^-\pi^+\pi^-\pi^+$, $D \to K\mu\nu$	2012	800k
$B^0 \to D^-\tau\nu$, $D \to \mu\nu$	2012	400k
$B^0 \to D^*\mu\nu$, $D \to 3\pi\pi^0\nu$	2012	800k
$B^0 \to D\mu\nu$, $D \to 3\pi\pi^0\nu$	2012	800k
cocktail $B^+ \to D^{**}\mu\nu$, $D \to 3\pi\pi^0$	2012	800k
cocktail $B^+ \to D^{**}\tau\nu$, $D \to 3\pi\pi^0$,$\tau \to \mu$	2012	800k
$B^0 \to D^*\tau\nu$, $D \to 3\pi\pi^0$,$\tau \to \mu$	2012	400k
$B^0 \to D\tau\nu$, $D \to 3\pi\pi^0$,$\tau \to \mu$	2012	800k
$B_s^0 \to D_s\mu\nu$, $D \to \tau\nu$	2012	400k
$B_s^0 \to D_s\pi^+\pi^-\pi^+$, $D_s \to \mu\nu$	2012	400k

Table 6.1: Monte Carlo samples used in this analysis and their yields.

information about the background behavior and categorization.

In addition, running over large existing samples of 360M events simulated with one of the b hadrons decaying into specific modes, a custom generic sample of 19k events, where the other b hadron is giving a candidate passing the stripping selection, is built to study the background from $b\bar{b}$ events. This sample is referred to as inclusive sample.

Regarding the τ decays, the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ channel is generated using the resonance chiral Lagrangian model [141] with a tuning based on the BaBar results for the $\tau^- \to \pi^-\pi^+\pi^-\nu_\tau$ decays [142] implemented in the Tauola generator [143]. The EVTGEN generator is used for the normalization channel sample.

The selection and the final neural networks are optimized on *truth matched* signal MC sample, i.e. the signal candidates are matched to the MC generated particles in order to select only well reconstructed candidates.

6.2 Selection

The first two steps of the selection are the so-called online trigger selection and the stripping. Their requirements are described in Sec. 6.2.1. Candidates are then submitted to a tighter selection made of a loose cut-based selection, and a BDT-based

selection, respectively described in Sec. 6.2.2 and 6.2.3. At this level of the selection, a study of the background composition, described in Sec. 6.2.4, is performed and a final BDT-based selection is applied. The full selection efficiency tables as well as a study of the remaining background composition in data are given respectively in Sec. 6.2.5 and 6.2.6.

6.2.1 Trigger and stripping selection

Trigger selection

We remind the reader that an overview of the trigger is provided in Sec. 4.2.4. Although the recorded events have passed the trigger, a refined trigger selection is applied offline. The B meson candidates of the selected events are required to pass the following TOS lines:

L0: L0HadronDecision or L0MuonDecision,

Hlt1: TrackAllL0Decision or TrackMuonDecision,

Hlt2: TopoMu2BodyBBDTDecision or TopoMu3BodyBBDTDecision or TopoMu4BodyBBDTDecision.

Stripping selection

An overview of the requirements applied on the stripping selection for the $B^0_{(s)} \to \tau^+(\to 3\pi^\pm\bar{\nu_\tau})\tau^-(\to \mu^-\bar{\mu_\nu}\nu_\tau)$ signal and the normalization channel $B^0 \to D^\pm(\to K^\pm\pi^\mp\pi^\pm)\pi^\mp$ is given in Tab. 6.2. We note that the requirements on the TOS Hlt2 trigger lines are actually applied in the stripping selection.

		$B^0_{(s)} \to \tau^+\tau^-$		$B^0 \to D^\pm\pi^\mp$
cut	on	value	on	value
p	π	$> 2000\,\text{MeV}$	π/K	$> 2000\,\text{MeV}$
p_T		$> 250\,\text{MeV}$		$> 250\,\text{MeV}$
min $\mathcal{V}$ dist. χ^2		> 16		> 16
Track χ^2		< 3		< 3
Ghost probability		< 0.3		< 0.3
ProbNN π		> 0.55	π	> 0.55
PID K		-	K	> -5
p	μ	$> 6000\,\text{MeV}$	π (from B)	$> 2000\,\text{MeV}$
p_T		$> 1000\,\text{MeV}$		$> 250\,\text{MeV}$
min $\mathcal{V}$ dist. χ^2		> 16		> 16
Track χ^2		< 3		< 3
Ghost prob.		< 0.3		< 0.3
PID μ		> 0		-
p_T	B	$> 5000\,\text{MeV}$	B	$> 5000\,\text{MeV}$
M		$[2000, 7000]\,\text{MeV}$		$[2000, 7000]\,\text{MeV}$
M_{corr}		$< 10000\,\text{MeV}$		$< 10000\,\text{MeV}$
IP χ^2		< 200		< 200
$\mathcal{V}$ dist.		$< 35\,\text{mm}$		$< 35\,\text{mm}$
$p_{T(\mu+3\pi/\pi+3\pi)}$		$> 2500\,\text{MeV}$		$> 2500\,\text{MeV}$
M	τ	$[400, 2100]\,\text{MeV}$	D	$[1750, 2080]\,\text{MeV}$
childs min $\mathcal{V}$ dist. χ^2		> 50		> 50
At least 1 daughter with P_T		$> 800\,\text{MeV}$		$> 800\,\text{MeV}$
DOCA max		$< 0.2\,\text{mm}$		$< 0.2\,\text{mm}$
p_T		$> 1000\,\text{MeV}$		$> 1000\,\text{MeV}$
θ_{DIRA}		> 0.99		> 0.99
$\mathcal{V}$ χ^2		< 12		< 12
FD χ^2		$[16,\ 4000]$		$[16,\ 4000]$
$\mathcal{V}$ ρ-dist		$[0.1, 7]\,\text{mm}$		$[0.1, 7]\,\text{mm}$
$\mathcal{V}$ z-dist		$5\,\text{mm}$		$5\,\text{mm}$
TOS Hlt2 (Topo*BodyBBDT or TopoMu or SingleMuon)	B	yes		no
hasMuon		yes		-

Table 6.2: Stripping requirements for the $B^0_{(s)} \to \tau^+(\to 3\pi^\pm\bar{\nu_\tau})\tau^-(\to \mu^-\bar{\mu_\nu}\nu_\tau)$ signal and the normalization channel $B^0 \to D^\pm(\to K^\pm\pi^\mp\pi^\pm)\pi^\mp$.

6.2.2 Loose cut-based selection

The first cut-based selection aims at removing the most obvious backgrounds while keeping a high signal efficiency. This selection is optimized on a unique signal MC sample merging the B^0, B^0_s, 2011 and 2012 samples, this latter is referred to as *merged signal MC sample* in the following. The background is modeled by the background region of data. The procedure to select the variables used in this selection is the following: First, we rank all variables with respect to their discrimination power; the FoM used for this ranking is the area under the ROC curve. A cut on the

ID	Step	Yield	Efficiency [%]
1	Total candidates = 1	15 350	97.44 ± 0.13
2	$\sum_\pi$ BDTiso2 $\geq$ -0.4	14 568	94.91 ± 0.18
3	$950 \leq m_\tau \leq 1450$ MeV	12 510	85.87 ± 0.29
4	τ and B^0 TOF signif.	12 454	99.55 ± 0.06
Total			79.06 ± 0.32

Table 6.3: Cut, yield and efficiency (on top of the previous cut) at each steps of the based-cut selection, for the merged signal MC sample. The cut ID 4 is detailed below.

selected variable is applied in order to remove the most obvious backgrounds. Then the ranking is reperformed and a new variable is selected.

Following this procedure three variables have been selected in order to obtain a signal efficiency around $\sim 80\%$: the total number of candidates in the event "Total candidates", the mass of the reconstructed τ lepton and the sum of the BDTiso2 isolation variable for the three pions. The set of applied cuts is summarized in Tab. 6.3. The cut on "Total candidates" leads to a multiplicity of 1 (only one candidate per event).
The cut ID 4 "τ and B^0 TOF signif." is a set of cuts on the time of flight significance of the B meson and the τ lepton. The time of flight significance is given by the ratio of the time of flight of the particle over its error; it aims at removing some "pathologic" events whose time of flight significance is too far away from the main distribution. This set of cuts is detailed below:

$$\text{Cut ID 4} = (-10 < \tau \text{ TOF signif.} < 50 \text{ and } 0 < B^0 \text{ TOF signif.} < 250). \quad (6.2)$$

The global efficiency on top of trigger and stripping for this selection is $(79.06 \pm 0.32)\%$ for merged MC signal and $(12.94 \pm 0.01)\%$ for data in the background region.

6.2.3 BDT-based selection

The signal training sample for the BDT1 consists of all signal MC events passing the cut-based selection. The background training sample consists of 50k data events from the background region, passing the cut-based selection. The BDT1 uses 20 variables, listed in Tab. 6.4. These variables have been selected following the procedure described in Sec. 5.4.6. The plot showing the evolution of the FoM at every step of the procedure is shown in Fig. 5.8. As the signal events are used later on in the analysis, the MC signal sample has been k-folded following the procedure described in Sec. 5.4.4. Strictly speaking, the data events are also used in further steps of the analysis but the very small statistics of the data used in the BDT training (50k) compared to large statistics of the full data sample (4.2M) allows us to consider this bias as negligible. The BDT parameters are optimized following the procedure described in 5.4.7. The chosen FoM is the area under the test ROC curve. The BDT parameters are given in Tab. 6.5.
The distribution of the BDT1 output for data in the background region and merged MC signal as well as the associated ROC curve are shown in Fig. 6.2. The output of the BDT1 is used as a discriminating variable. A selection cut is set in order to keep about 18k data events in the signal region (like in the Run1 analysis of the $(3\pi, 3\pi)$

ID	Variable
0	τ nc-PZ
1	$\min_\pi p_T$
2	$\sum_\pi$ BDT iso 2
3	B M_{corr}
4	B DOCA
5	τ lifetime significance
6	τ p_T
7	$\min_\pi$ (min IP χ^2)
8	B M
9	τ $\mathcal{V}$ dist.χ^2
10	τ IP
11	$\max_\pi$ p_T
12	$\min_\pi$ IP
13	μ p_T
14	τ IP χ^2
15	B nc-sPT
16	τ BDT iso 1 (1st digit)
17	τ NumVtxWithinChi2WindowOneTrack
18	τ ENDVERTEX χ^2
19	$\max_\pi$ (min IP χ^2)

Table 6.4: Variables used in BDT1 training ranked by selection order.

Parameters	Value
N_t	900
$d_{\max}$	3
$MinNodeSize$ [%]	1
β	0.2

Table 6.5: Parameters used in BDT1.

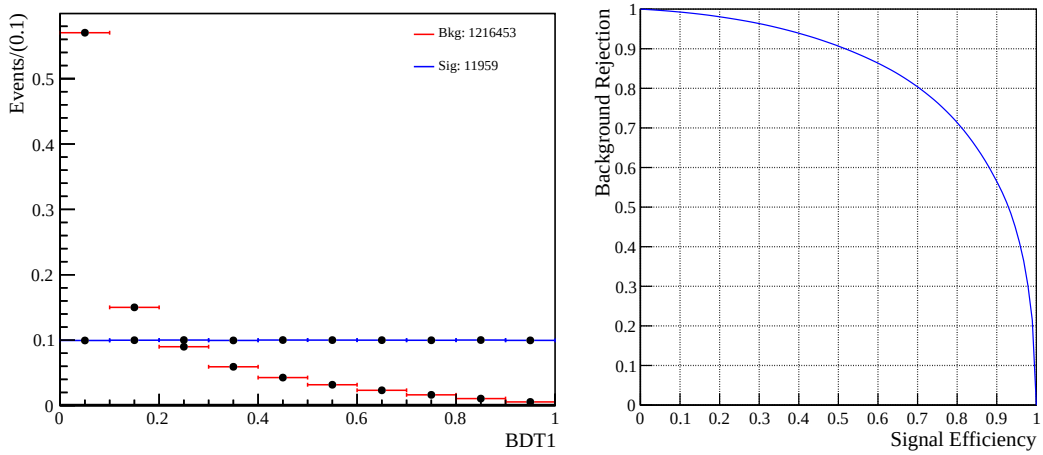


Figure 6.2: Distribution of the BDT1 output for data in background region [Red] and MC signal [Blue] (left plot) and associated ROC curve (right plot).

final state). This leads to the cut: BDT1 output$>$ 0.835. The yield and the efficiency on top of the cut-based selection for MC signal, data and data in the background region are given in Tab. 6.6.

BDT1	Yield	Efficiency[%]
Signal	2056	16.51 ± 0.33
Data in full Dalitz	75175	1.78 ± 0.01
Data in bkg. region	14478	1.19 ± 0.01

Table 6.6: Yield after the cut on the BDT1 and efficiencies of the BDT1 cut for merged MC signal, data and data in the background region.

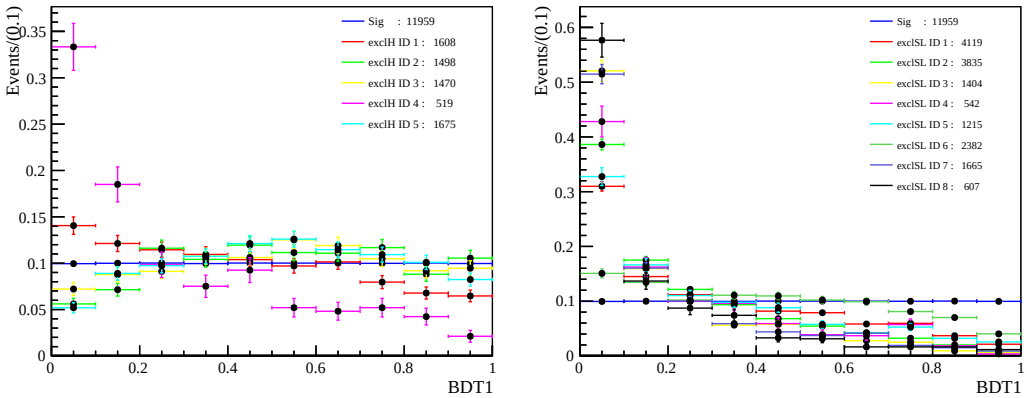


Figure 6.3: Distribution of the BDT1 output for the exclusive MC backgrounds making up the hadronic (semileptonic) cocktail in the left (right) plot. The IDs for each exclusive sample match the ones given in Tab. 6.7 and 6.8. The merged MC signal is in blue.

6.2.4 Exclusive background suppression

As mentioned in Sec. 5.3.3, two of the most difficult backgrounds to suppress have the following topologies:

- semileptonic decay as $B \to D^{(*,**)}(\to 3\pi(X))\mu\nu_\mu$,
- purely hadronic decay of type $B \to D(\to (X)\mu\nu_\mu)3\pi(Y)$,

where X and Y are generic hadrons, e.g.: π^0. In order to better understand this background, a set of exclusive decays have been simulated; they are given in Tab. 6.1. We selected these decays for their relatively high branching ratio and because they have the same final state than the signal, i.e. three charged pions, one muon and missing energy. The list is not exhaustive but it gives insights on the nature of these backgrounds and how to suppress them. This list of exclusive backgrounds has been split into two according to the two different topologies identified above:

- One category, called the semileptonic cocktail, where the B meson decay is semileptonic, giving most of the time a D meson and a τ or μ.
- A second category, called the hadronic cocktail, where the B meson decay is purely hadronic. Most of the time the B meson decays into a D meson and at least 3 pions. The D meson then decays either leptonically or semileptonically.

The exact composition of these cocktails, as well as the expected yield of each mode (taking into account the branching ratio and the selection efficiency) at the end of the BDT1 selection are given in Tab. 6.7 and 6.8 for the semileptonic cocktail and hadronic cocktail respectively. As the uncertainty on the branching ratio of some of

ID	Decay	Absolute yield	Expected yield
1	$B^0 \to D(3\pi\pi^0), \mu\nu$	187	2397
2	$B^0 \to D^*(3\pi\pi^0), \mu\nu$	66	493
3	$B^0 \to D^{**}(3\pi\pi^0), \mu\nu$	15	49
4	$B^0 \to D^*(3\pi\pi^0), \tau\nu(\mu\nu)$	7	6
5	$B^0 \to D(3\pi\pi^0), \tau\nu(\mu\nu)$	61	55
6	$B^0_s \to D_s(\tau\nu), \mu\nu$	190	1024
7	$B^+ \to D^{**}(3\pi\pi^0), \mu\nu$	27	27
8	$B^+ \to D^{**}(3\pi\pi^0), \tau\nu(\mu\nu)$	12	1
Total		565	4052

Table 6.7: Semileptonic cocktail composition, the yields in the MC sample and the expected yields in Run1 after the BDT1 selection.

ID	Decay	Absolute yield	Expected yield
1	$B^0 \to D(K^0\mu\nu), 3\pi$	173	4207
2	$B^0 \to D(\mu\nu), 3\pi$	244	78
3	$B^0 \to D^*(\mu\nu), 3\pi$	228	22
4	$B^0 \to D^*(\mu\nu), 3\pi\pi^0$	22	4
5	$B^0_s \to D^+_s(\mu\nu), 3\pi$	244	549
Total		911	4860

Table 6.8: Hadronic cocktail composition, the yields in the MC sample and the expected yields in Run1 after the BDT1 selection.

these modes is large, the expected yield is just an estimate. The distribution of the BDT1 output for each exclusive background of the hadronic/semileptonic cocktail can be seen in Fig. 6.3. We note that the BDT1 seems to be more discriminating for the semileptonic cocktail than for the hadronic one. Motivated by this difference in shape, we build two BDTs trained each on a specific topology, BDTH and BDTSL, respectively trained on the hadronic and semileptonic background topology.

BDT against semileptonic background

The signal training sample for BDTSL consists of all truth matched signal MC passing the cut on BDT1, i.e. 1981 events. The background training sample consists of all events from the semileptonic cocktail, described in Tab. 6.7, passing the cut on BDT1, i.e. 565 events. BDTSL uses the same setup as BDT1. The BDT uses 12 variables, listed in Tab. 6.9. The BDT parameters are given in Tab. 6.10.

The distribution of the BDTSL output for semileptonic cocktail and merged MC signal as well as the associated ROC curve are shown in Fig. 6.4. The distribution for data in the background region and MC signal are shown in Fig. 6.5. We note that BDTSL seems to perform badly on data; as a consequence we don't use it in the selection. This poor performance will be explained later.

ID	Variable
0	τ ρ-flightdistance
1	B p_Z
2	μ p_Z
3	B p_T
4	$\max_\pi$ BDTiso2
5	$\min_\pi$ BDTiso1, 3rd digit
6	μ IP χ^2
7	$\max_\pi$ p_Z
8	τ IP
9	μ IP
10	$\sum_\pi$ IP
11	$\max_\pi$ BDTiso1, 5th digit

Table 6.9: Variables used in BDTSL training ranked by selection order.

Parameters	Value
N_t	150
$d_{\max}$	3
$MinNodeSize$ [%]	1.1
β	0.2

Table 6.10: Parameters used in BDTSL.

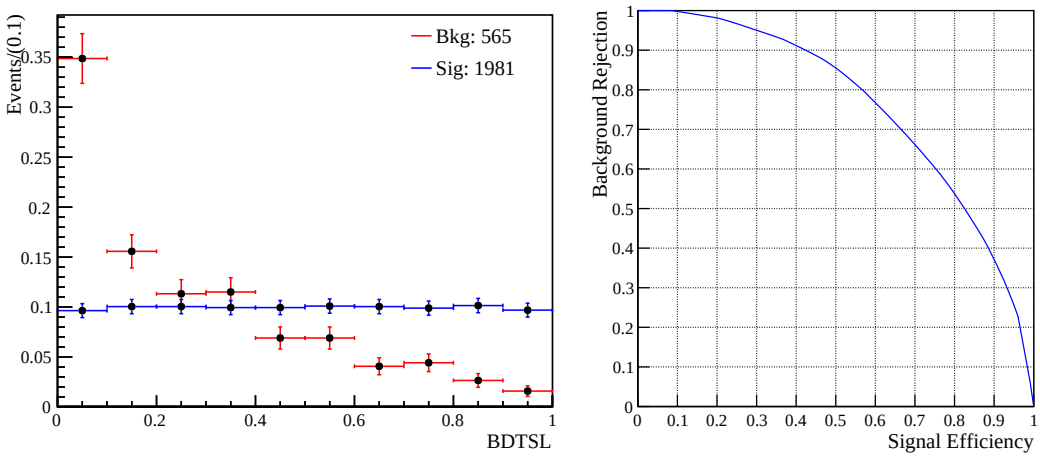


Figure 6.4: Distribution of the BDTSL output for data in the background region [Red] and MC signal [Blue] (left plot) and the associated ROC curve (right plot).

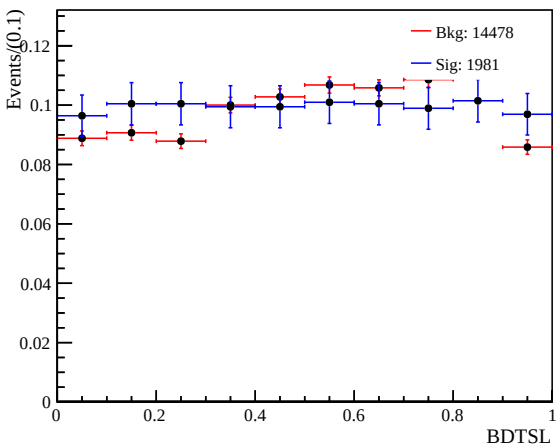


Figure 6.5: Distribution of the BDTSL output for data in the background region [Red] and MC signal [Blue].

ID	Variable
0	τ ρ-flightdistance
1	BM_{corr}
2	τ θ_{DIRA} ORIVX
3	B DOCA
4	B mass
5	$\max_{\pi} p_Z$
6	τ θ_{DIRA} OWNPV
7	$\sum_{\pi} p_Z$

Table 6.11: Variables used in BDTH ranked by selection order.

Parameter	Value
N_t	660
$d_{\max}$	3
$MinNodeSize$ [%]	0.1
β	2

Table 6.12: Parameters used in BDTH

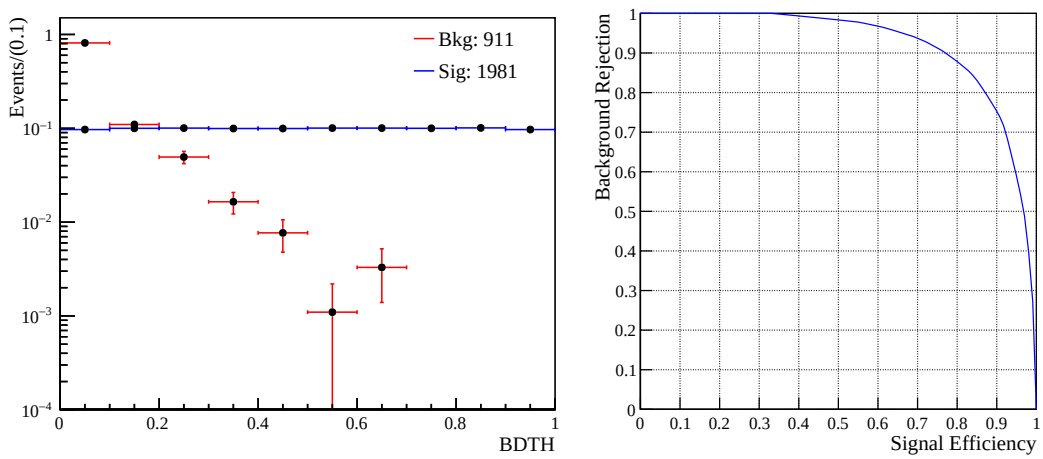


Figure 6.6: Distribution of the BDTH output for hadronic cocktail [Red] and MC signal [Blue] in logarithmic scale (left plot) and the associated ROC curve (right plot).

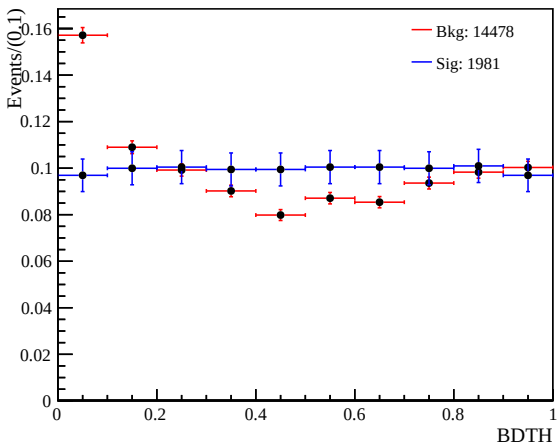


Figure 6.7: Distribution of the BDTH output for data in the background region [Red] and MC signal [Blue].

BDT against purely hadronic background

The signal training sample for BDTH consists of all truth matched signal MC passing the cut on BDT1, i.e. 1981 events. The background training sample

consists of all events from the hadronic cocktail, described in Tab. 6.8, passing the cut on BDT1, i.e. 911 events. BDTH uses the same setup than BDT1. The BDT uses 8 variables, listed in Tab. 6.11. The BDT parameters are given in Tab. 6.12.

The distribution of the BDTH output for hadronic cocktail and MC signal and the associated ROC cure are shown in Fig. 6.6. The distribution of the BDTH output for data in the background region and MC signal is shown in Fig. 6.7. The BDTH output is used as a selection variable. We apply the cut: BDTH> 0.1. This cut corresponds to an efficiency of $(90.47 \pm 0.65)\%$ for signal and $(84.29 \pm 0.30)\%$ for background. 60 690 data events remain after this cut.

Concluding discussion on BDTSL and BDTH

The distribution of the BDTH (BDTSL) output for each exclusive background of the hadronic cocktail and the semileptonic cocktail can be seen on Fig. 6.8 (6.9). As we can see, BDTH (BDTSL) badly performs on the semileptonic (hadronic) cocktail. This means that the topology of the two categories are really different. Moreover in both cases, the most discriminating variable entering the BDT is the τ "ρ-flight-distance". This variable is the projection of the τ flight-distance on the plan orthogonal to the beam.
We also note that these BDTs are much less effective on data than MC. It is not clear why but several explanations can be given:

- The exclusive backgrounds used to train BDTH and BDTSL are not exactly the same modes we find in data, as we will see in Sec. 6.2.5.
- The performance of the BDT is "diluted" once it is applied on data, since data are made not only of these specific backgrounds.

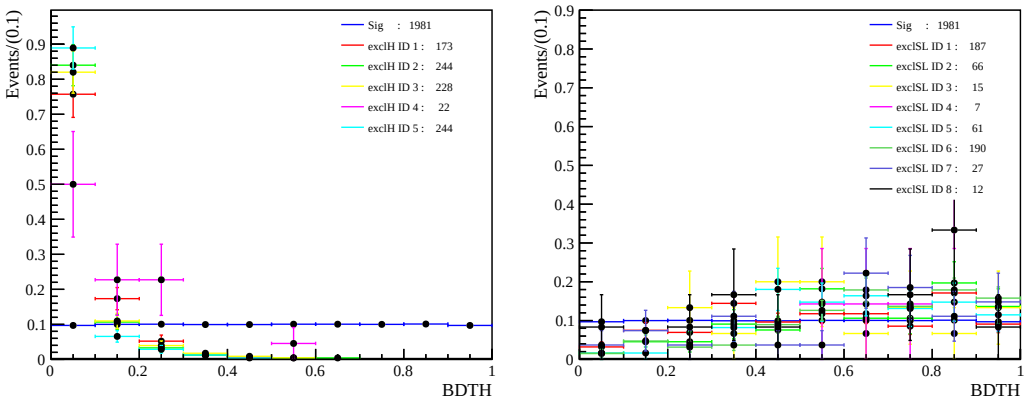


Figure 6.8: Distribution of the BDTH output for the exclusive MC backgrounds making up the hadronic (semileptonic) cocktail in the left (right) plot. The IDs for each exclusive sample match the ones given in Tab. 6.7 and 6.8. The MC signal is in blue.

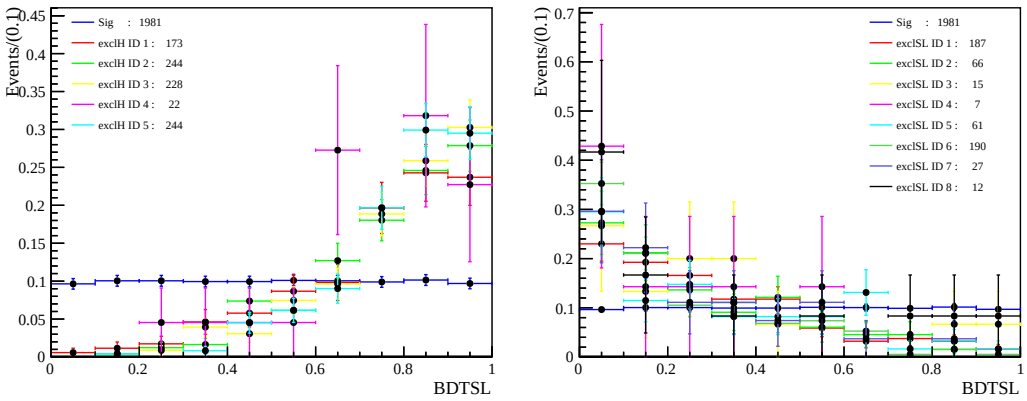


Figure 6.9: Distribution of the BDTSL output for the exclusive MC backgrounds making up the hadronic (semileptonic) cocktail in the left (right) plot. The IDs for each exclusive sample match the ones given in Tab. 6.7 and 6.8. The MC signal is in blue.

Category	Sig. region	Ctl. region
Hadronic B^0	2	7
Hadronic B_s^0	0	0
Hadronic B^+	2	2
Semileptonic B^0	3	13
Semileptonic B_s^0	4	4
Semileptonic B^+	6	4
Others (Λ_b meson, ...)	0	2
Combinatorial	0	2
Total Hadronic	4	9
Total Semileptonic	13	21
Total	17	34

Table 6.13: Yields of the inclusive MC sample after a cut on the BDT1, classified according to the type of B hadron decay.

6.2.5 Qualitative study of the background composition

This section presents a qualitative study of the background composition using the custom MC inclusive sample described in Sec. 6.1. A handful of events from this sample pass the selection: 22093 candidates pass the trigger and stripping selection, 3607 pass the loose cut-based selection and 68 survive the BDT1 selection. These 68 events are classified in Tab. 6.13 according to the type of B-hadron decay. The statistics used to produce this table is really low and the yield in the signal and the control region given in the table should be taken with caution. Nevertheless, we can make the following observations:

- Very few (2) combinatorial background events remain after the BDT1 selection, most of background is composed of exclusive decays, where the four tracks come from the same B meson.
- The background composition is approximately the same for the signal and the control region: the semileptonic B decay appears to be dominant in the control region (21 over 34 events) and in the signal region (13 over 17 events).
- Most of the semileptonic decays contain a K meson, e.g. the topology $B^+ \to (D^{0(*)} \to a_1^- K^+)\mu\nu$ seems to dominate in the signal region (4 over 13 events). Several modes also contain other neutral particle, e.g.: photon, ρ^0, ...

To summarize, most events contain a kaon and sometimes another neutral meson. There is no a single mode that appear more than once. The complexity of the semileptonic background remaining in data after the BDT1 selection and its difference with the semileptonic cocktail on which the BDTSL is trained could explain the bad performance of the BDTSL on data.

6.2.6 Selection efficiencies

The selection efficiency ϵ_f, where f labels the final state, can be decomposed into its individual sources as

$$\epsilon_f = \epsilon_{\text{Geo.}} \times \epsilon_{\text{Reco.}} \times \epsilon_{\text{Strip.}} \times \epsilon_{\text{Trig.}} \times \epsilon_{\text{Sel.}} \,. \tag{6.3}$$

$\epsilon_{\mathbf{Geo.}}$ is the geometrical acceptance of the LHCb detector, i.e. the requirement that all four final-state particles fly through the LHCb detector. This includes as well the generator cuts applied at the MC production level.

$\epsilon_{\mathbf{Reco.}}$ is the candidate reconstruction efficiency.

$\epsilon_{\mathbf{Strip.}}$ is the stripping selection efficiency

$\epsilon_{\mathbf{Trig.}}$ is the efficiency for the events to be triggered. It can be further decomposed into a trigger efficiency for the `L0Hadron` and the Hlt1 trigger levels. We remind that Hlt2 is contained in the stripping requirements.

$\epsilon_{\mathbf{Sel.}}$ is the efficiency of the further offline selection, i.e. the loose cut-based selection and the BDT-based selection (BDT1 and BDTH).

No systematic uncertainties or correction of possible data/MC differences have been taken into account in the computation of the efficiencies. An overview of the yields at various stages of the selection is given in Tab. 6.14 for MC simulation and Tab. 6.15 for data and data in the background region. From these yields, the selection efficiencies are determined. The efficiencies and the uncertainties on the efficiencies are calculated following the frequentist binomial method:

$$\epsilon = \langle\epsilon\rangle \pm \sigma_\epsilon \,, \qquad \langle\epsilon\rangle = \frac{S}{N} \,, \qquad \sigma_\epsilon = \langle\epsilon\rangle\sqrt{\frac{1}{S} - \frac{1}{N}} \,, \tag{6.4}$$

where S is the number of events passing the selection requirement and N is the total number of events before the cut. The yields after the full selection in each region are given in Tab. 6.16.

Step	Yield	Efficiency [%]
+ Geo. Acc. + Gen. Cut	2621360	5.83 ± 0.01
+ Strip + Reco	18616	0.71 ± 0.01
+ Trigger L0	17203	92.41 ± 0.19
+ Trigger Hlt	15753	91.57 ± 0.21
+ Loose Sel	12454	79.06 ± 0.32
+ BDT1	2056	16.51 ± 0.33
+ BDTH	1860	90.47 ± 0.65
+ Signal region	640	34.41 ± 1.10
Total		$(1.43 \pm 0.06) \times 10^{-3}$

Table 6.14: Selection step, yield and efficiency (after the previous step) for the merged signal sample.

	Data		Data in bkg. region	
Step	Yield	Efficiency [%]	Yield	Efficiency [%]
+ Reco + Strip	24184604		11142707	
+ Trigger L0	21871852	90.44 ± 0.01	10080651	90.47 ± 0.01
+ Trigger Hlt	20446991	93.49 ± 0.01	9399259	93.24 ± 0.01
+ Loose Sel	4219857	20.64 ± 0.01	1216453	12.94 ± 0.01
+ BDT1	75175	1.78 ± 0.01	14478	1.19 ± 0.01
+ BDTH	60690	80.73 ± 0.14	12203	84.29 ± 0.30
+ Signal region	13715	22.60 ± 0.17		

Table 6.15: Selection step, yield and efficiency (after the previous step) used in the full selection for data and data in the background region.

	$B_s^0 \to \tau^+\tau^-$ MC		Data	
Sample	Yield	Fraction [%]	Yield	Fraction [%]
Signal	640	34.41 ± 1.10	13715	22.60 ± 0.17
Control	897	48.23 ± 1.16	30007	49.44 ± 0.20
Background	160	8.60 ± 0.65	12203	20.11 ± 0.16
Left-over	163	8.76 ± 0.66	4765	7.85 ± 0.11
Total	1860		60690	

Table 6.16: Overview of the event yields for the different subregions after applying the full selection.

6.3 Likelihood fit

As mentioned previously we fit the distribution of a new BDT output, denoted BDT2. Sec. 6.3.1 is devoted to this BDT. The fit procedure is given in Sec. 6.3.2.

6.3.1 The fitted variable

The training signal sample for the BDT2 consists of all the signal MC candidates passing the BDTH selection, i.e. all the Dalitz regions are used (1 780 events).

ID	Variable
0	BDT1
1	$\sum_\pi$ BDT iso 3
2	B SmallestDeltaChi2MassTwoTracks
3	B NumVtxWithinChi2WindowOneTrack
4	τ nc sPT
5	B decay vertex χ^2
6	B mass
7	B FDχ^2
8	τ z-flightdistance
9	τ BDTiso3
10	τ IPχ^2
11	τ θ_{DIRA}

Table 6.17: Variables used in BDT2 training ranked by selection order.

Parameter	Value
N_t	400
$d_{\max}$	3
$MinNodeSize$ [%]	1.9
β	0.2

Table 6.18: Parameters used in BDT2.

The background training sample consists of all data candidates remaining in the background region after the BDTH selection (12 203 candidates). The BDT2 uses the same setup as BDT1. The BDT2 uses 12 variables, listed in Tab. 6.17. The BDT parameters are given in Tab. 6.18. Note that the BDT2 has been k-folded for signal following the procedure described in 5.4.4.
The distribution of the BDT2 output for data in the background region and merged MC signal as well as the associated ROC curve are shown in Fig. 6.10. We aim at using the control region of data as a background model for the likelihood fit, hence the data BDT2 shapes in the control and in the signal region have to agree at least in the part of the distribution free of potential signal (low values of BDT2). The output distribution for data in the control and signal region (the signal region is blinded for BDT2 > 0.7) is shown in Fig. 6.11. The agreement is good, hence we use this variable for the likelihood fit.

Note that the input variables used for this BDT2 have been selected via the iterative procedure described in Sec. 5.4.6. This procedure selects the τ-invariant masses, due to their good discrimination power. Nevertheless, the agreement of the BDT2 distribution between the control region and the signal region is spoiled once these variables are added to the BDT2. This is explained by their strong correlations with the variables used to define the signal and the control region ($m_{\pi_1^+\pi_2^-}$ and $m_{\pi_2^+\pi_3^-}$). Hence these variables have been removed from the iterative procedure that selects input variables.

6.3.2 Fit strategy

Our fit strategy is to use the control region of data as a "proxy" to construct the background template. It relies on the fact that the BDT2 shape in the control and signal regions are similar as we can see on Fig. 6.11. However, as Tab. 6.16 indicates, the control region contains an important fraction of $B_s^0 \to \tau^+\tau^-$ events. The possible signal component in the control region should therefore be taken into account when constructing the background template. The resulting fit model takes the form:

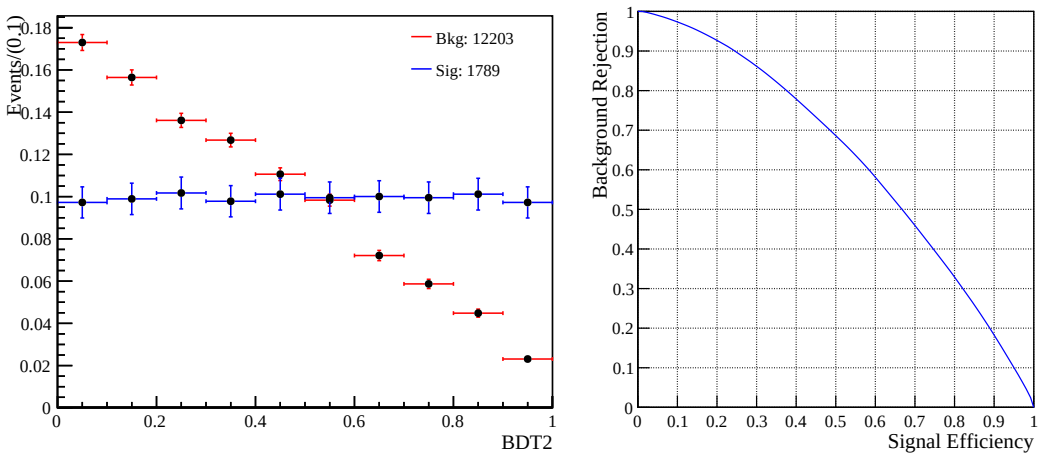


Figure 6.10: Distribution of the BDT2 output for data in background region [Red] and merged MC signal [Blue] (left plot) and the associated ROC curve (right plot).

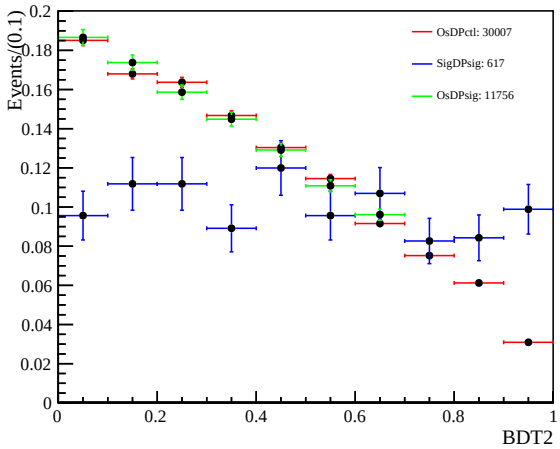


Figure 6.11: Distribution of the BDT2 output for data in control region [Red], in the signal region [Green] and MC signal in the signal region [Blue].

$$\mathcal{N}^{\mathrm{SR}}_{\mathrm{data}} = s \times \widehat{\mathcal{N}}^{\mathrm{SR}}_{\mathrm{sim}} + f_b \times \left(\mathcal{N}^{\mathrm{CR}}_{\mathrm{data}} - s \times \frac{\epsilon^{\mathrm{CR}}}{\epsilon^{\mathrm{SR}}} \times \widehat{\mathcal{N}}^{\mathrm{CR}}_{\mathrm{sim}} \right), \tag{6.5}$$

where $\mathcal{N}^{\mathrm{SR}}_{\mathrm{data}}$ ($\mathcal{N}^{\mathrm{CR}}_{\mathrm{sim/data}}$) is the BDT2 output distribution in the signal (control) region from simulation/data, s is the signal yield, f_b is a scaling factor for the background template, and ϵ^{SR} (ϵ^{CR}) is the signal efficiency in the signal (control) region. The hat on top of the BDT2 output distribution means that the distribution is normalized to unity. The fit is set up using the HistFactory framework [49].

Fit validation

The fit model is validated using toy simulations, also called sometimes pseudo experiments. Lacking analytic probability density functions from which toy data could be generated, the templates and pseudo data samples for the toys are obtained by re-sampling the original distributions following the bootstrapping method [50]. The following templates are re-sampled: signal region of the $B^0_s \to \tau^+\tau^-$ MC (for signal template), control region of the data and signal (for the background templates).

Parameter	Value
s	9.2 ± 25.0
σ^{toys}	444.7 ± 4.4
Pull – mean	0.038 ± 0.056
Pull – σ	1.24 ± 0.04

Table 6.19: Results for the fit validation study based on 500 pseudo experiments and obtained by fitting a Gaussian function to the distribution of fit values, fit uncertainties and pulls on the signal yield s, shown in Fig. 6.12.

The validation of the fit model is based on 500 toys. The settings include the statistical uncertainties following the Beeston–Barlow technique [144]. For the situation representing the SM expectation, i.e. no signal, the distribution of fit values, fit uncertainties and pulls for the $B_s^0 \to \tau^+\tau^-$ signal yield s and the scale factor s are given in Fig. 6.12. These distributions are fitted with a Gaussian function, also shown in Fig. 6.12, whose parameters are given in Tab. 6.19. As expected the s-value is compatible with zero. The pull is given by the ratio of s-value over the s-error. If the fit performs correctly, the distribution of the pull is expected to match the one of a standard Gaussian. As we see the mean of the pull is compatible with zero but the deviation σ is slightly above 1.

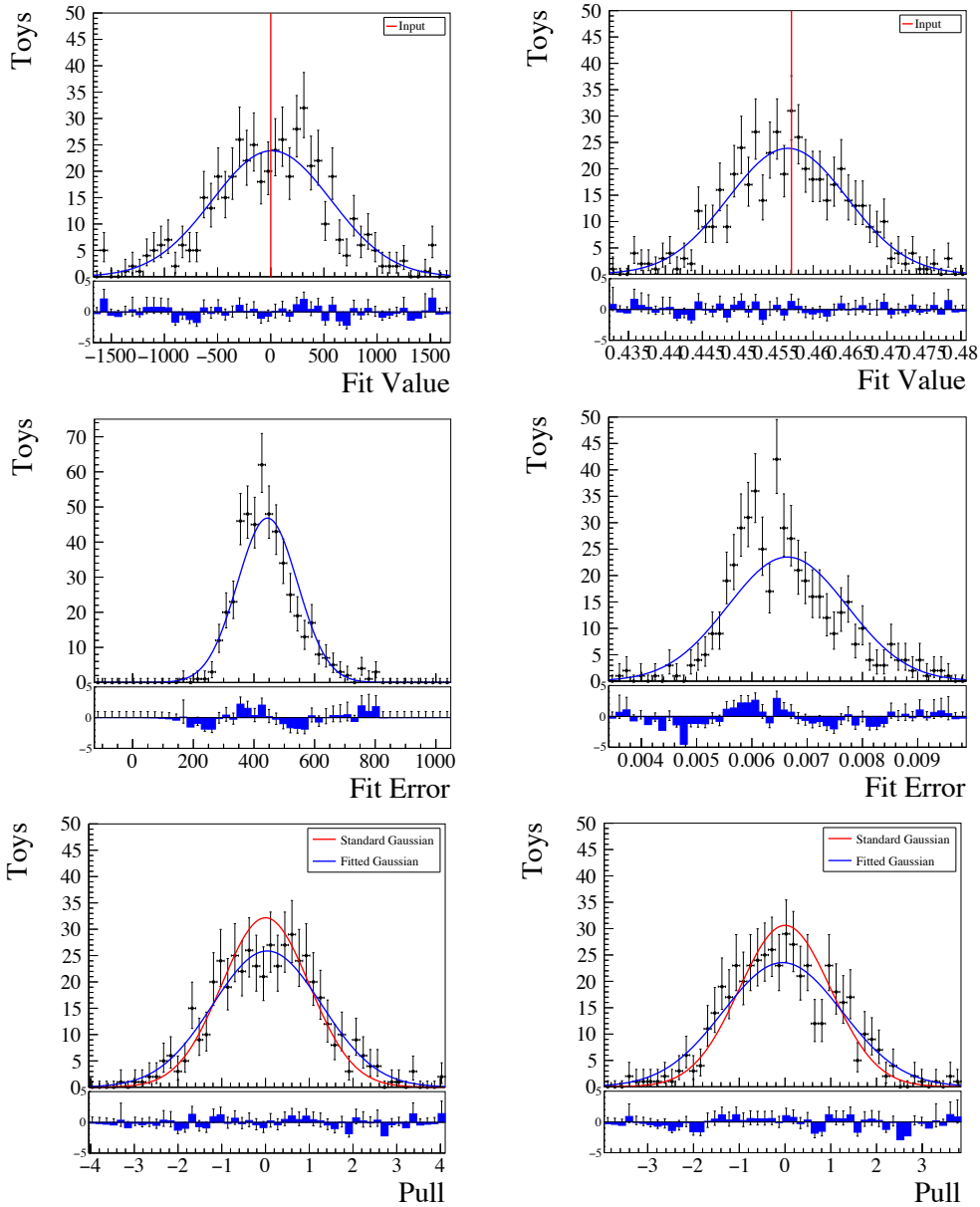


Figure 6.12: Results for the fit validation study based on 500 pseudo experiments. Shown are the distribution of fit values [Top], fit uncertainties [Middle] and pulls [Bottom] on the signal yield s [Left] and background scale f_b [Right] for the nominal setting including only statistical uncertainties. The distributions are fitted with a Gaussian function, of which the fitted parameters are given in Tab. 6.19.

6.4 Normalization

This section is devoted to the normalization channel $B^0 \to D^-\pi^+$, which has the same number of charged tracks in the final state and a similar decay topology. Sec. 6.4.1 gives an overview of the analysis of the normalization channel. Sec. 6.4.2 presents the normalization factor. We remind the reader that the normalization procedure is explained in detail in Sec. 5.2.2 (for the $(3\pi, 3\pi)$ final state). The $B^0 \to D^\pm(\to K^\pm\pi^\mp\pi^\pm)\pi^\mp$ channel is also used as a control channel to check the agreement of the variables used in this analysis between data and simulation.

Variable	Associated to	Value
PID requirements		
ProbNNpi	π from B	> 0.55
PID K	K from D	> 15
ProbNNpi	π's from D	> 0.55
Kinematic requirements		
min IP	B	< 0.05
θ_{DIRA}		< 0.0085
decay vertex χ^2		< 5
FD		> 2
PT		$> 5050\,\mathrm{MeV}/c$
TAU χ^2		< 10
DOCA		< 0.05
M	D	$[1850, 1900]\,\mathrm{MeV}/c$
decay vertex χ^2		< 6
FD χ^2		> 200
PT	K from D	$> 750\,\mathrm{MeV}/c$

Table 6.20: Offline selection requirements for the normalization channel $B^0 \to D^{\pm}(\to K^{\pm}\pi^{\mp}\pi^{\pm})\pi^{\mp}$.

6.4.1 Overview of the $B^0 \to D^-\pi^+$ analysis

This normalization channel has already been used in the $B^0_{(s)} \to \tau^-\mu^+$ analysis [145]. We give here a brief overview of the analysis.

Selection

After the stripping and trigger selection, see Sec. 6.2.1, a set of requirements on the PID and kinematic variables is applied. These requirements are listed in Tab. 6.20. The total $B^0 \to D^{\pm}(\to K^{\pm}\pi^{\mp}\pi^{\pm})\pi^{\mp}$ efficiency corrected for data/MC differences is

$$\epsilon^{\mathrm{tot}}_{D^+\pi^-} = (2.00 \pm 0.08\ (\mathrm{stat}) \pm 0.17\ (\mathrm{syst})) \times 10^{-4}. \tag{6.6}$$

Likelihood fit

The B invariant mass distribution after the selection process is fitted using an unbinned maximum likelihood fit in the B invariant mass range $[5200, 5400]\,\mathrm{MeV}/c^2$. The $D^{\pm}\pi^{\mp}$ invariant mass distribution in data is modeled with a double tailed Crystal Ball (CB) function [146] parameterizing the signal contribution, and an exponential function, accounting for the combinatorial background.
The parameters of the tails of the signal PDF ($n1$ and $n2$) are fixed from a fit to the MC signal. The fit on data is shown in Fig. 6.13 and the fit results are given in Tab. 6.21. The yield of the $B^0 \to D^{\pm}(\to K^{\pm}\pi^{\mp}\pi^{\pm})\pi^{\mp}$ mode is:

$$N^{\mathrm{obs}}_{D^+\pi^-} = 22588 \pm 176. \tag{6.7}$$

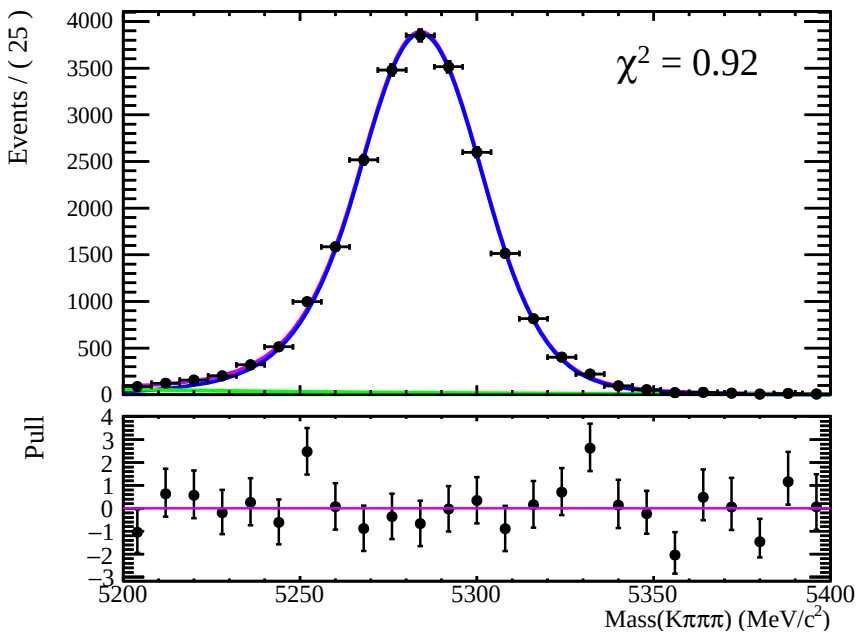


Figure 6.13: Fit of B^0 invariant mass distribution for the $B^0 \to D^-\pi^+$ channel. The total PDF is represented in purple, the signal in blue, and the background in green.

Parameter	Value
Crystal ball parameters	
Yield	22588 ± 176
Mean	5284.3 ± 1.2
Width	17.4 ± 0.2
$\alpha 1$	1.1 ± 0.3
n1	40
$\alpha 2$	-1.2 ± 0.4
n2	39
Exponential parameters	
Yield	559 ± 95
Exponent	$(-1.2 \pm 0.2) \cdot 10^{-2}$

Table 6.21: Results of the fit to the B^0 invariant mass distribution for the $B^0 \to D^-\pi^+$ channel.

6.4.2 Normalization factor

The observed number of $B^0 \to \tau^+\tau^-$ events is related to the $B^0 \to \tau^+\tau^-$ branching ratio through the normalization factor:

$$\begin{aligned}\mathcal{B}(B^0 \to \tau^+\tau^-) &= \frac{\mathcal{B}(B^0 \to (3\pi, \mu))}{\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau)\mathcal{B}(\tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau)} \\ &= \frac{N^{\text{obs}}_{B^0\to(3\pi,\mu)}/\epsilon^{\text{tot}}_{B^0\to(3\pi,\mu)}}{\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau)\mathcal{B}(\tau^- \to \mu^-\bar{\nu}_\mu\nu_\tau)N_{B^0}} \equiv \alpha_d \times N^{\text{obs}}_{B^0\to(3\pi,\mu)} \,,\end{aligned} \tag{6.8}$$

Quantity	Value
$\mathcal{B}(B^0 \to D^+\pi^-)$	$(2.52 \pm 0.13) \times 10^{-3}$
$\mathcal{B}(D^+ \to K^+\pi^-\pi^+)$	$(8.98 \pm 0.28) \times 10^{-2}$
$\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^+\nu_\tau)$	$(9.31 \pm 0.06) \times 10^{-2}$
$\mathcal{B}(\tau^- \to \mu^-\bar{\nu_\mu}\nu_\tau)$	$(17.39 \pm 0.04) \times 10^{-2}$

Table 6.22: Overview of the input branching ratios needed to calculate the normalization factors $\alpha_{d/s}$. Values are taken from [47].

	B_s^0	B^0
$\alpha_{d/s}$	$(3.5 \pm 0.4) \times 10^{-5}$	$(8.8 \pm 1.1) \times 10^{-6}$
$N^{\mathrm{SM}}_{B^0_{(s)}}$	$\sim 2.2 \times 10^{-2}$	$\sim 2.5 \times 10^{-3}$

Table 6.23: Normalization factors and expected yields in SM for B_s^0 and $B^0 \to \tau^+\tau^-$ $(3\pi, \mu)$ decays.

where $\epsilon^{\mathrm{tot}}_{B^0\to\tau^+\tau^-}$ is the total $B^0 \to \tau^+(\to 3\pi^\pm\bar{\nu_\tau})\tau^-(\to \mu^-\bar{\nu_\mu}\nu_\tau)$ reconstruction and selection efficiency and N_{B^0} is the number of produced B^0 mesons in the data. The normalization factor for B^0 is then given by

$$\alpha_d = \frac{\epsilon^{\mathrm{tot}}_{D^+\pi^-} \cdot \mathcal{B}(B^0 \to D^+\pi^-) \cdot \mathcal{B}(D^+ \to K^+\pi^-\pi^+)}{N^{\mathrm{obs}}_{D^+\pi^-} \cdot \epsilon^{\mathrm{tot}}_{B^0\to\tau^+\tau^-} \cdot \mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau) \cdot \mathcal{B}(\tau^- \to \mu^-\bar{\nu_\mu}\nu_\tau)}, \tag{6.9}$$

where $\epsilon^{\mathrm{tot}}_{D^+\pi^-}$ is the total $B^0 \to D^-(\to \pi^-K^+\pi^-)\pi^+$ reconstruction and selection efficiency In the case of the B_s^0 meson, the normalization factor α_s must be divided by the ratio of B_s^0 to B^0 production factor $f_s/f_d = 0.259 \pm 0.015$ [137] and the efficiency $\epsilon^{\mathrm{tot}}_{B^0\to\tau^+\tau^-}$ must be replaced by $\epsilon^{\mathrm{tot}}_{B_s^0\to\tau^+\tau^-}$.
For the $(3\pi, 3\pi)$ final state, the total MC signal efficiencies for B_s^0 and B^0 are similar (2.4×10^{-5} versus 2.2×10^{-5} [45]), hence we make the approximation $\epsilon^{\mathrm{tot}}_{B^0\to\tau^+\tau^-} \simeq \epsilon^{\mathrm{tot}}_{B_s^0\to\tau^+\tau^-} \simeq \epsilon^{\mathrm{tot}}_{B^0_{(s)}\to\tau^+\tau^-}$ where $\epsilon^{\mathrm{tot}}_{B^0_{(s)}\to\tau^+\tau^-}$ is the efficiency for the merged sample (B^0 and B_s^0) given in Tab. 6.14. The normalization factor α_s and α_d are then related by

$$\alpha_s \simeq \alpha_d \cdot \frac{f_d}{f_s}\,. \tag{6.10}$$

The branching ratios in Eq. (6.9) are taken from [47] and listed in Tab. 6.22. The value of $\epsilon^{\mathrm{tot}}_{D^+\pi^-}$ and the signal yield $N^{\mathrm{obs}}_{D^+\pi^-}$ are respectively given in Eq. (6.6) and (6.7).

The value of the normalization factor and the expected yields for B^0 and B_s^0 are given in Tab. 6.23.

6.5 Results and conclusions

In this section, estimates of the reachable limits in case no signal events are observed are computed. A comparison between the $(3\pi, 3\pi)$ and $(3\pi, \mu)$ final state of the main quantities entering this estimate for Run1 is given in Sec. 6.5.1. A cross-check of the limit estimate is described in Sec. 6.5.2. Finally, conclusions are drawn in Sec. 6.5.3.

	B_s^0		B^0	
	$(3\pi,\mu)$	$(3\pi,3\pi)$	$(3\pi,\mu)$	$(3\pi,3\pi)$
$\epsilon^{\text{tot}} \times 10^5$	1.42	1.8	1.42	1.7
$\alpha \times 10^5$	3.5	4.7	0.9	1.3
σ^{toys}	444	58	490	61
$\mathcal{UL}$ at 95 % C.L.	2.4×10^{-2}	4.4×10^{-3}	7.0×10^{-3}	1.3×10^{-3}
$\frac{\mathcal{UL}_{(3\pi,\mu)}}{\mathcal{UL}_{(3\pi,3\pi)}}$	5.7		5.4	

Table 6.24: Signal efficiency, normalization factor, fit error, limit estimate for the $(3\pi,\mu)$ final state and limit at 95% C.L. for the $(3\pi,3\pi)$ final state.

6.5.1 Limit estimate

We have now all the ingredients to compute a first estimate of the reachable upper limit, noted $\mathcal{UL}$, in the absence of signal events. An estimate of the upper limit is given by:

$$\mathcal{UL} \sim 1.3(1.6) \times \alpha \times \sigma^{\text{toys}} \text{ at } 90(95) \text{ \% C.L.} \tag{6.11}$$

The normalization factors for B_s^0 and B^0 can be found in Tab. 6.23, the fit errors on the signal yield, σ^{toys}, are given in Tab. 6.19. The value of the limit estimate can be found in Tab. 6.24. We have to keep in mind that our estimate is based on a fit framework which does not include any kind of systematics and that the σ^{toys} entering the estimate is purely statistic. Moreover, the signal efficiency entering the normalization factor and the signal BDT2 distribution are not corrected from possible data-MC disagreements. Our estimates are compared to the expected limit estimates at 95% C.L. obtained via the $(3\pi,3\pi)$ final state in the same condition: no data/MC corrections applied to the efficiencies and a σ^{toys} extracted from a similar toys study. The worse limit estimates we obtain with respect to the analysis of the $(3\pi,3\pi)$ final state seem to be mainly due to the fit error which is one order of magnitude above the fit error for the $(3\pi,3\pi)$ final state.

6.5.2 A last test

We need to understand if the high fit error is only due to the bad discriminating power of the BDT2 or to a difference of distributions for the input variables entering the BDT between the background and the control region for data, which would result in a non optimal BDT.
To answer this question, a new BDT is trained on the data control region (instead of the background region as BDT2) and a new limit is computed. This BDT represents the best BDT we could obtain as it is not possible to find a better region to approximate the background in the signal region. This BDT is called BDT2ctl and is described in the following. Note that, in principle, it is not possible to fit this variable as using the same region of data to train the BDT and as a model for background in the signal region could lead to biases.

BDT training

The signal training sample for the BDT2ctl consists of all truth matched signal MC events passing the full selection (1789 events). The background training sample consists of all remaining events in the control region of data (30007 events).

ID	Variable
0	BDT1
1	μ BDTiso2
2	τ ρ-FD
3	$\max_\pi(p_T)$
4	τ FD χ^2
5	B nc-PZ
6	BDTSL
7	B-NumVtxWithinChi2WindowOneTrack
8	τ θ_{DIRA}
9	B DOCA
10	$\max_\pi(\mathrm{iso})$

Table 6.25: Variables used in BDT2ctl training ranked by selection order.

Parameter	Value
N_t	400
$d_{\max}$	3
$MinNodeSize$ [%]	1.5
β	0.2

Table 6.26: Parameters used in BDT2ctl.

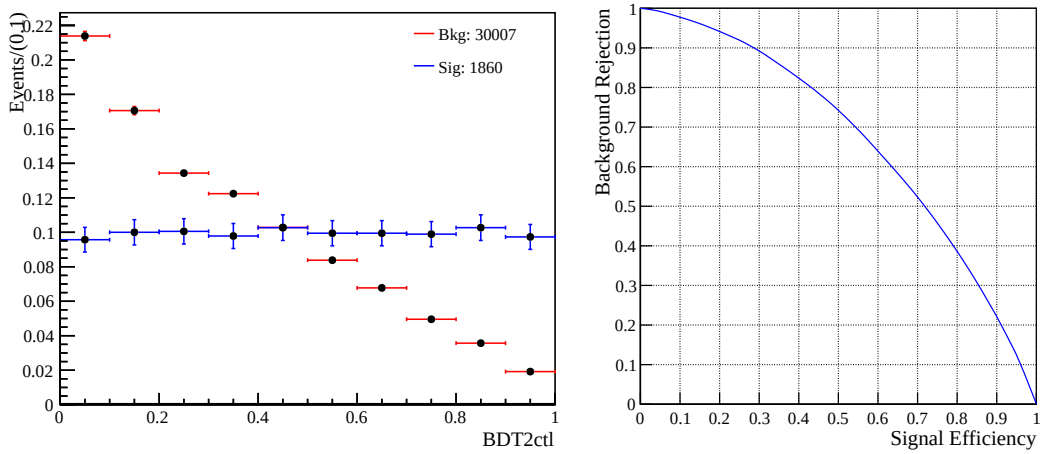


Figure 6.14: Distribution of the BDT2ctl output for data in control region [Red] and MC signal [Blue] (left plot) and the associated ROC curve (right plot).

The BDT setup is described in Sec. 5.4.6. The BDT uses 11 variables, listed in Tab. 6.25. The BDT parameters are given in Tab. 6.26. It is relevant to note that only two variables are in common with BDT2, the BDT1 output and B-NumVtxWithinChi2WindowOneTrack. Moreover, BDTSL that performs badly on data in the background region seems to be useful to BDT2ctl.

BDT results

The distribution of the BDT2ctl output for data in the control region and MC signal as well as the associated ROC curve are shown in Fig. 6.14. By comparing the ROC curves, we note that BDT2ctl is slightly more discriminating than BDT2.

Fit results and limit estimate

We apply the fit strategy described in Sec. 6.3.2. For the situation representing the SM expectation, i.e. no signal, the distribution of fit values, fit uncertainties and pulls for the $B_s^0 \to \tau^+\tau^-$ signal yield s and the scale factor f_b are given in Fig. 6.15. These distributions are fitted with a Gaussian function, also shown in Fig. 6.15, for

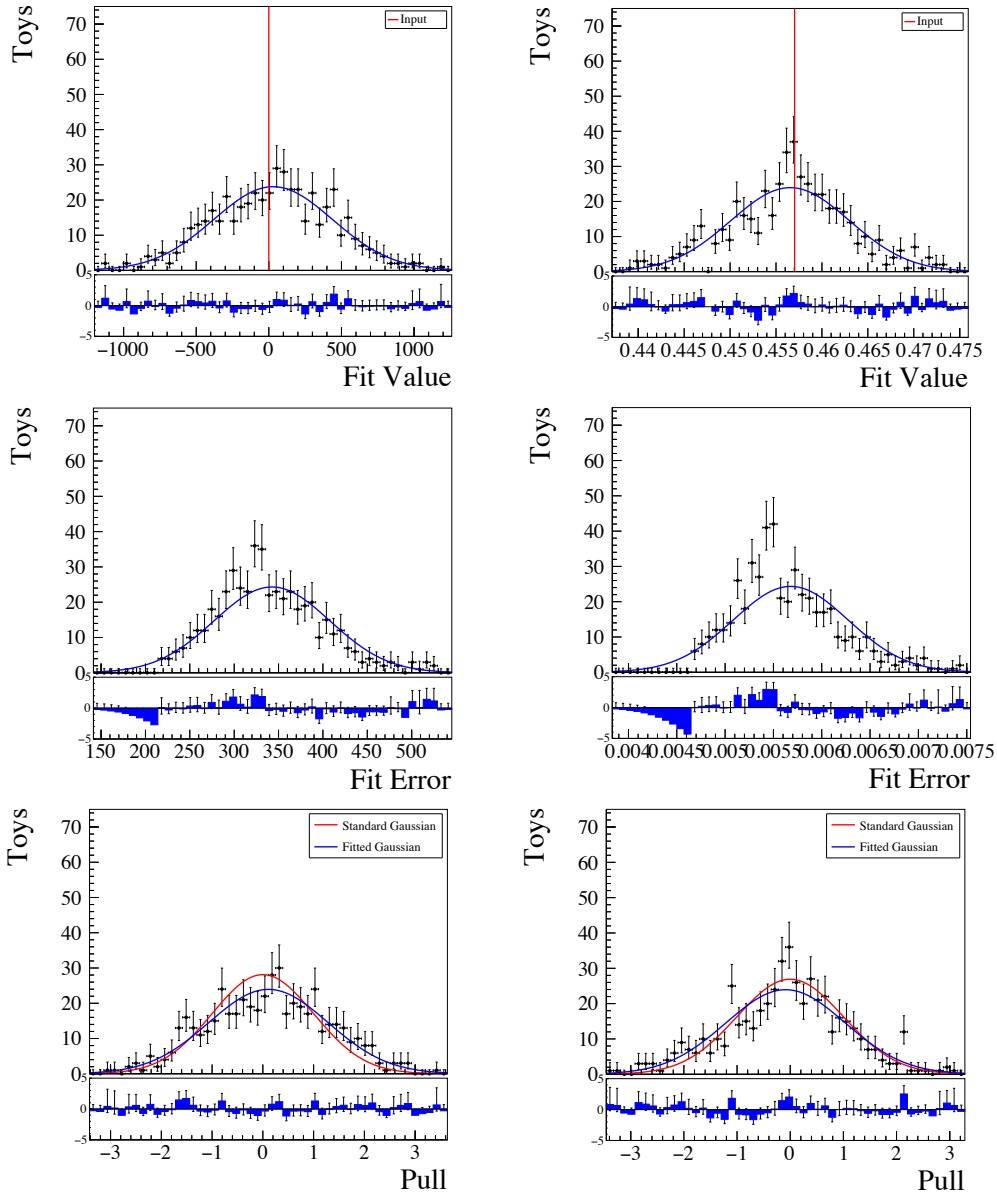


Figure 6.15: Results for the fit validation study for BDT2ctl, based on 500 pseudo experiments. Shown are the distribution of fit values [Top], fit uncertainties [Middle] and pulls [Bottom] on the signal yield s [Left] and scale background f_b [Right] for the nominal setting including only statistical uncertainties. The distributions are fitted with a Gaussian function, of which the fitted parameters are given in Tab. 6.27.

which the parameters are given in Tab. 6.27.

From the fit error in Tab. 6.27 and the normalization factor in Tab. 6.23, we can compute the limit estimate via Eq. (7.9) for BDT2ctl. This estimate for B_s^0 and B^0, as well as the ratio with limit for $(3\pi, 3\pi)$ final state, are given in Tab. 6.28.

Concluding discussion on BDT2ctl

The s-error obtained from the toys study is smaller (342 versus 444), which leads to better limit estimate for B_s^0 and B^0. Nevertheless, the limits are still much above

Parameter	Value
s	29.4 ± 18.0
σ^{toys}	342.6 ± 2.9
Pull – mean	0.11 ± 0.052
Pull – σ	1.17 ± 0.037

Table 6.27: Results for the fit validation study for BDT2ctl based on 500 pseudo experiments and obtained by fitting a Gaussian function to the distribution of fit values, fit uncertainties and pulls on the signal yield s, shown in Fig. 6.15.

	B_s^0	B^0
$\mathcal{UL}_{(3\pi,\mu)}$	1.9×10^{-2}	4.9×10^{-3}
$\frac{\mathcal{UL}_{(3\pi,\mu)}}{\mathcal{UL}_{(3\pi,3\pi)}}$	4.4	3.7

Table 6.28: Limit estimate for the $(3\pi, \mu)$ final state for BDT2ctl as well as its ratio over the limit at 95% C.L. for the $(3\pi, 3\pi)$ final state.

the ones obtained with the $(3\pi, 3\pi)$ final state.

This test shows that, even using the control region (which cannot be done in a real world analysis) for the BDT training, the final BDT is not discriminating enough to make the $(3\pi, \mu)$ final state competitive with the $(3\pi, 3\pi)$ final state.

6.5.3 Conclusions

This chapter has presented a feasibility study for the search for $B^0_{(s)} \to \tau^+\tau^-$ via the $(3\pi, \mu)$ final state. A toy-based estimate of the upper limit at 95% C.L. on the branching ratio in absence of signal events has been computed in Sec. 6.5.1 and gives:

$$\mathcal{B}(B^0_s \to \tau^+\tau^-) < 2.4 \times 10^{-2},$$
$$\mathcal{B}(B^0 \to \tau^+\tau^-) < 7.0 \times 10^{-3}.$$

These estimates do not take into account corrections from possible data-MC disagreements or systematics uncertainties. They are around five times larger than the expected limit obtained via the $(3\pi, 3\pi)$ final state in the same conditions (no corrections for possible data-MC disagreements and no systematics uncertainties.)

This feasibility study has shown that the advantage of the $(3\pi, \mu)$ final state (higher acceptance and trigger efficiency, larger branching ratio entering the normalization factor) do not compensate its drawbacks, namely only one reconstructed τ vertex and an abundant source of semileptonic background.

We show that a refined selection reduces a bit the hadronic background but fails in removing most semileptonic background. This latter could be further studied by generating new MC exclusive backgrounds matching the main modes found in the analysis of the inclusive sample. In particular, as a lot of backgrounds seems to come from $D^+ \to D^+(\to K^{0,+}3\pi^\pm)\mu^-\nu_\mu$, we could try to improve the π isolation variables. Nevertheless, the last test (BDT2ctl) shows that it is very difficult to improve the discriminating power of the final BDT, even using a data region closer to the signal one.

Finally, this feasibility study indicates that the $(3\pi, \mu)$ final state can contribute marginally to the search on the $B^0_{(s)} \to \tau^+\tau^-$ decay. The study has been reported into the LHCb internal note CERN-LHCb-INT-2018-021 [51].

The next chapter is devoted to the $(3\pi, 3\pi)$ final state for the Run2 data.

Chapter 7

Search for $B^0_{(s)} \to \tau^+\tau^-$ via the $(3\pi, 3\pi)$ final state

This chapter describes the analysis performed to improve the experimental upper limit on the $B^0_{(s)} \to \tau^+\tau^-$ decay via the final state $B^0_{(s)} \to \tau^+(\to 3\pi^\pm\bar{\nu_\tau})\tau^-(\to 3\pi^\mp\nu_\tau)$, also referred to as the $(3\pi, 3\pi)$ final state, with the Run2 data. We present a study whose strategy is almost identical to the one used in the published LHCb analysis [45] developed for the $(3\pi, 3\pi)$ final state and optimized for Run1 data. We remind the reader that an overview of the published analysis is given in Sec. 5.2.

Sec. 7.1 gives the analysis overview and the main changes with respect to the published LHCb analysis. The selection, the fit and the normalization are described respectively in Sec. 7.2, 7.3 and 7.4. Finally, conclusions are drawn in Sec. 7.5.

7.1 Analysis overview

The idea behind this study is to obtain an estimate of the reachable upper limit without reoptimizing or improving the strategy of the published analysis. Thus this estimate can be used as a *baseline* for future reoptimizations/improvements of the analysis. This analysis is as close as possible to the published analysis. Nonetheless, some differences exist between Run1 and Run2 and cannot be ignored. The differences are made explicits below.

First, the Hlt trigger lines have changed between Run1 and Run2. We use lines as close as possible to the lines used in the published analysis. The stripping process is the same up to the Hlt2 lines (we remind the reader that the Hlt2 trigger is applied at the stripping level). The loose cut-based selection has not been reoptimized and the set of cuts is the same. The same set of variables and parameters tuning is imposed to NN1, nonetheless the NN1 algorithm has been retrained with Run2 data and MC simulations. The NN2 algorithm has also the same set of variables than the published analysis, except for two variables, the two τ masses, that have been removed for reasons explained later. The same k-folding technique and flattening to the NN2 output are applied. Finally, the same fit strategy is used. The available samples for Run2 are presented below.

Channel	Year	Yield
$B_s^0 \to \tau^+\tau^-$	2015	400k
$B^0 \to \tau^+\tau^-$	2015	400k
$B_s^0 \to \tau^+\tau^-$	2016	1.6M
$B^0 \to \tau^+\tau^-$	2016	1.6M
$B_s^0 \to \tau^+\tau^-$	2017	1.6M
$B^0 \to \tau^+\tau^-$	2017	1.6M
$B^0 \to D^- D_s^+$	2016	1M
$B^0 \to D^- D_s^+$	2017	1M

Table 7.1: Monte Carlo samples used in the baseline analysis and their yields.

Data samples

The analysis is performed using data corresponding to an integrated luminosity of $\sim 3.4\,\mathrm{fb}^{-1}$ of proton–proton collisions recorded with the LHCb experiment at center-of-mass energies of $\sqrt{s} = 13$ TeV in 2016 and 2017, respectively. We will refer to these samples as the Run2 samples although the Run2 contained also data recorded in 2015 and 2018.

Simulated samples

Several MC samples are used to study the properties of the signal, as well as those of the normalization sample. A brief description of the MC simulation chain is given in Sec. 4.2.5. An overview of the available set of MC simulated samples is given in Tab. 7.1. In order to save CPU time, the samples are generated requiring that all particles in the final state be in the LHCb acceptance. In addition, all samples generated for the signal and exclusive backgrounds have been produced with the following generator level cuts on the daughter particles: $p_{\mathrm{T}}(\pi) > 250\,\mathrm{MeV}/c$, $p(\pi) > 2\,\mathrm{GeV}/c$ and $p_{\mathrm{T}}(\mu) > 250\,\mathrm{MeV}/c$.

We discovered a problem with the version of the variable "ProbNN π" entering the stripping for the 2015 samples (data and MC). A procedure called re-stripping would have been necessary to correct this version. As this procedure can be long and the statistics of the 2015 samples are low, we decided to not analyze the 2015 data in this feasibility analysis.
The 2018 MC simulation were not available at the time of this work. They are now available and samples have been produced.

7.2 Selection

Stripping selection

An overview of the requirements applied on the stripping selection for the signal $B_{(s)}^0 \to \tau^+\tau^-$ and the normalization channel $B^0 \to D^- D_s^+$ for Run2 is given in Tab. 7.2. The statistics for the data samples after the stripping selection are given in Tab. 7.3.

		$B^0_{(s)} \to \tau^+\tau^-$		$B^0 \to D^- D_s^+$
Cut	on	value	on	value
Track χ^2/ndf	π	<3	π/K	<3
Ghost probability		< 0.3		< 0.3
IPχ^2		>16		>16
p_T		> 0.25 GeV/c		> 0.25 GeV/c
p		> 2 GeV/c		> 2 GeV/c
ProbNN π		>0.55	π	>0.55
PID K		N/A	K	>-5
at least 1 daughter with p_T	τ	>0.8 GeV/c	D	>0.8 GeV/c
$\mathcal{V}$ χ^2		<16		<16
DOCA max		<0.2 mm		0.2 mm
M		[500-2000] MeV/c		[1800-2030] MeV/c
$\mathcal{V}$ dist. χ^2		>16		>16
$\mathcal{V}$ ρ-dist		[0.1 − 7] mm		[0.1 − 7] mm
$\mathcal{V}$ z-dist		>5.0 mm		>5.0 mm
θ_{DIRA}		>0.99		>0.99
p_T		>1 GeV/c		>1 GeV/c
p_T	B	> 2 GeV/c	B	> 2 GeV/c
M		[2-7] GeV/c		[5-7] GeV/c
M_{corr}		<10 GeV/c		<10 GeV/c
$\mathcal{V}$ χ^2		<90		<90
θ_{DIRA}		>0.99		>0.99
$\mathcal{V}$ dist. χ^2		>225		>225
$\mathcal{V}$ dist.		<90 mm		<90 mm
max p_T of D/τ		>4 GeV/c		>4 GeV/c
max(min IPχ^2 D/τ_1,min IPχ^2 D/τ_2)		>150		>150
min(min IPχ^2 D/τ_1,min IPχ^2 D/τ_2)		>16		>16
max(min (IPχ^2 $K/\pi_{1,2,3}$),min (IPχ^2 $K/\pi_{4,5,6}$))		>20		>20
max p_T of K/π		>2 GeV/c		>2 GeV/c
sum p_T of K/π		>7 GeV/c		>7 GeV/c
TOS on Hlt2Topo[2,3,4]Body		yes		N/A

Table 7.2: Stripping requirements for the $B^0_{(s)} \to \tau^+(\to 3\pi^\pm\bar{\nu_\tau})\tau^-(\to 3\pi^\mp\nu_\tau)$ signal and the normalization channel $B^0 \to D^- D_s^+$ for Run2.

Year	Polarity	Candidates	Events	Multiplicity
2016	MagDw	113 956 890	17 812 922	6.397±0.076
2016	MagUp	108 596 297	17 001 648	6.387±0.077
2017	MagDw	123 419 266	18 540 776	6.657±0.079
2017	MagUp	117 652 793	17 718 263	6.640±0.080
Sum		463 625 246	71 073 609	6.523

Table 7.3: Statistics of the data samples after the stripping selection.

Trigger selection

The Hlt trigger lines have changed between Run1 and Run2. All trigger lines are given in Tab. 7.4. The L0Global[TIS] requires that at least one L0 line fired under TIS condition. The L0 trigger efficiency for B_s^0 signal MC on top of the stripping is $(49.95 \pm 0.29)\%$ for 2012, $(54.56 \pm 0.29)\%$ for 2016 and $(70.27 \pm 0.26)\%$ for 2017. We showed that this large difference between 2016 and 2017 efficiencies is explained by a

Step	Run1	Run2
L0	L0Hadron[TOS] || L0Global[TIS]	idem
Hlt1	TrackAllL0[TOS]	TrackMVA[TOS]
Hlt2	Topo[2/3/4]BodyBBDT[TOS]	Topo[2/3/4]Body[TOS]

Table 7.4: Trigger lines used in Run1 and Run2.

ID	Description
0	$802 \leq m_\tau \leq 1598\,\mathrm{MeV}/c^2$
1	$m_B - m_{\tau^+} - m_{\tau^-} \geq 756\,\mathrm{MeV}/c^2$
2	$\tau_\tau \geq -0.12\,\mathrm{ps}$
3	B SmallestDeltaChi2MassOneTrack ≥ 3499
4	τ SmallestDeltaChi2MassTwoTracks ≥ 1791
5	Total Candidates ≤ 6
6	$\sum \mathrm{Iso}(\pi) \leq 3$
7	$\sum_\pi$ BDTiso1, 1st digit ≤ 16
8	$\sum_\pi$ BDTiso1, 3rd digit ≤ 5
9	$\sum_\pi$ BDTiso3 ≥ -0.96
10	B nc-vPT ≤ 3850

Table 7.5: Overview of the cuts used in the loose cut-based selection.

difference of threshold of the transverse energy E_T in L0Hadron:

$$\begin{aligned} E_T &\geq 3744\text{MeV for 2016,} \\ E_T &\geq 3456\text{MeV for 2017.} \end{aligned} \tag{7.1}$$

The combined Hlt1 and Hlt2 trigger efficiency for B_s^0 signal MC on top of the stripping is $(77.83 \pm 0.34)\%$ for 2012. Efficiency for the 2016 and 2017 signal MC are slightly larger, $(85.34 \pm 0.27)\%$ and $(84.89 \pm 0.24)\%$ respectively.

Loose cut-based selection

In Run1, eleven variables have been identified from a long list of (highly) discriminating ones and are individually able to remove more than 50% of the candidates in data while still maintaining a signal efficiency of more than 95%. The cuts have been optimized to retain as close to, but more than, 98% of the signal candidates in simulation. The same set of cuts, given in Tab. 7.5, is applied to Run2 data. The MC signal efficiencies for B_s^0 on top of the stripping and trigger selection are in good agreement between 2012 $(82.45 \pm 0.35\%)$, 2016 $(82.20 \pm 0.32\%)$ and 2017 $(81.87 \pm 0.28\%)$.

Neural network selection

The neural network (NN) selection is the last step of the offline selection before the signal region selection. The implementation is the same as for Run1. The signal training sample for the NN1 consists of all signal MC events passing the cut-based selection. The background training sample consists of 50k events from the data background region. These 50k events are split equally across the four subsamples (2011/2012, Mag Up/Down) of the data, and only one candidate per event. The NN1 uses the 7 variables, listed in Tab. 7.6, that most discriminate between signal and

Rank		Variable	Sig.
1	$\tau^{\pm}$	reconstruction variable $\mathcal{R}e[x_3]$	109.3 σ
2	τ^{-}	Mass	66.1 σ
3	τ^{+}	Mass	62.7 σ
4	B_s^0	nc-vPT	50.7 σ
5	$\pi^{\pm}$	$\sum_{6\pi}$ BDTiso3	45.2 σ
6	τ^{+}	Lifetime	44.0 σ
7	τ^{-}	Lifetime	41.9 σ

Table 7.6: List of input variables entering the NN1 together with their significance. See the NeuroBayes manual [133] for the definition of the significance.

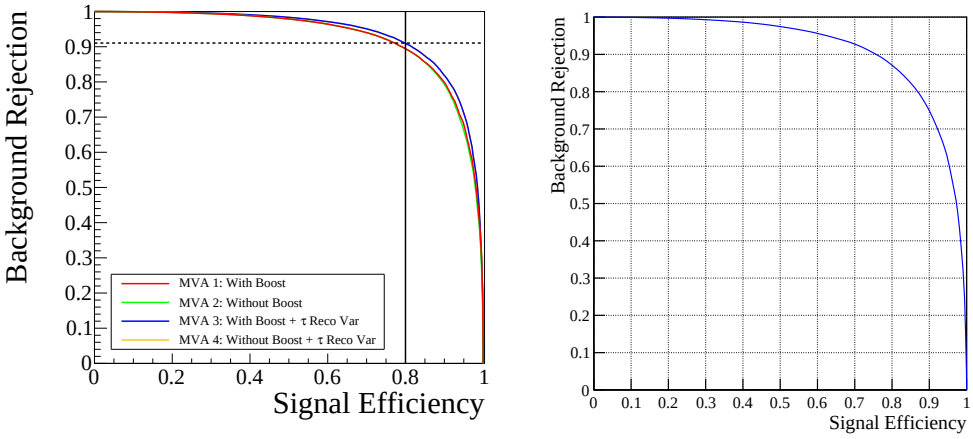


Figure 7.1: ROC curves for the NN1 of Run1 (Run2) on the left (right) plot. The MVA 3 (blue curve) has been used for Run1.

background. They are identified from the ranking provided by NeuroBayes at the preprocessing stage of the NN training, where a significance of at least 40σ is required.

The same set of variables as for Run1 has been used to built the NN1 for Run2. The ROC curves for Run1 and Run2 are shown in Fig. 7.1. The performance of the NN1 are slightly better for Run1, e.g. for a signal efficiency of 80%, the background rejection is around 91% for Run1 and around 87% for Run2. We have chosen the same figure of merit as for Run1, i.e. the selection cut is optimized in order to retain 80% of MC signal, which leads to the cut: NN1 output ≥ -0.43.

Selection efficiencies

As for the $(3\pi, \mu)$ final state, the total selection efficiency can be broken down in the following way

$$\epsilon_f = \epsilon_{\text{Geo.}} \times \epsilon_{\text{Reco.}} \times \epsilon_{\text{Strip.}} \times \epsilon_{\text{Trig.}} \times \epsilon_{\text{Sel.}} \,. \tag{7.2}$$

An overview of the yields at various stages of the selection is given in Tab. 7.7 for B_s^0 and 7.8 for B^0 MC simulations. From these yields, the default selection efficiencies are determined. No systematic uncertainties or corrections for possible data/MC differences are taken into account in the computation of the efficiencies. The efficiencies and the associated uncertainties are calculated following the frequentist binomial method described in Sec. 6.2.6.

For completeness, the signal efficiencies for B_s^0 are also compared in Tab. 7.9 to those obtained in the published analysis.

2016	Cut-and-Count	
Step	Yield	Efficiency[%]
+Geom. Acc.	1 607 659	3.654 ± 0.008
+Reco + Strip	55115	3.43 ± 0.01
+Truth Match	49147	89.17 ± 0.13
+PID	30510	62.08 ± 0.22
+L0	16645	54.56 ± 0.29
+Hlt	14205	85.34 ± 0.27
+Loose Sel	11753	82.20 ± 0.32
+NN1	9354	79.70 ± 0.37
+Dalitz Sig	1385	14.81 ± 0.37
Total		$(31.32 \pm 0.85) \times 10^{-4}$

2017	Cut-and-Count	
Step	Yield	Efficiency[%]
+Geom. Acc.	1 584 829	3.667 ± 0.012
+Reco + Strip	55683	3.51 ± 0.01
+Truth Match	49601	89.08 ± 0.13
+PID	31528	63.56 ± 0.22
+L0	22155	70.27 ± 0.26
+Hlt	18808	84.89 ± 0.24
+Loose Sel	15502	81.87 ± 0.28
+NN1	12600	81.47 ± 0.31
+Dalitz Sig	1856	14.71 ± 0.32
Total		$(42.94 \pm 1.00) \times 10^{-4}$

Average		$(37.17 \pm 0.66) \times 10^{-4}$

Table 7.7: Overview of the selection efficiencies with respect to the previous step for the $B_s^0 \to \tau^+\tau^-$ samples, split by step in the selection process.

2016	Cut-and-Count	
Step	Yield	Efficiency[%]
+Geom. Acc.	1 597 951	3.591 ± 0.007
+Reco + Strip	53265	3.33 ± 0.01
+Truth Match	47224	88.66 ± 0.14
+PID	28760	60.90 ± 0.22
+L0	15673	54.50 ± 0.29
+Hlt	13257	84.58 ± 0.29
+Loose Sel	10656	80.38 ± 0.34
+NN1	8331	77.63 ± 0.40
+Dalitz Sig	1252	15.03 ± 0.34
Total		$(28.14 \pm 0.80) \times 10^{-4}$

2017	Cut-and-Count	
Step	Yield	Efficiency[%]
+Geom. Acc.	1 580 386	3.593 ± 0.007
+Reco + Strip	53714	3.40 ± 0.01
+Truth Match	47633	88.68 ± 0.14
+PID	29515	61.96 ± 0.22
+L0	20781	70.41 ± 0.27
+Hlt	17544	84.42 ± 0.25
+Loose Sel	14206	80.97 ± 0.30
+NN1	11242	79.14 ± 0.34
+Dalitz Sig	1639	14.58 ± 0.33
Total		$(37.24 \pm 0.92) \times 10^{-4}$

Average		$(32.64 \pm 0.61) \times 10^{-4}$

Table 7.8: Overview of the selection efficiencies for the $B^0 \to \tau^+\tau^-$ samples, split by step in the selection process.

Step	2012 MC B_s^0 [%]	2016 MC B_s^0 [%]	2017 MC B_s^0 [%]
+Geom. Acc.	3.124 ± 0.008	3.654 ± 0.011	3.667 ± 0.012
+Reco + Strip	2.708 ± 0.011	3.43 ± 0.01	3.51 ± 0.01
+Truth Match	88.54 ± 0.14	89.17 ± 0.13	89.08 ± 0.13
+PID	61.08 ± 0.22	62.08 ± 0.22	63.56 ± 0.22
+L0	49.95 ± 0.29	54.56 ± 0.29	70.27 ± 0.26
+Hlt	77.83 ± 0.34	85.34 ± 0.27	84.89 ± 0.24
+Loose Sel	82.45 ± 0.35	82.20 ± 0.32	81.87 ± 0.28
+NN1	77.16 ± 0.43	79.70 ± 0.37	81.47 ± 0.31
+Dalitz Sig	15.82 ± 0.42	14.81 ± 0.37	14.71 ± 0.32
Total	$(17.91 \pm 0.53) \times 10^{-4}$	$(31.32 \pm 0.85) \times 10^{-4}$	$(42.94 \pm 1.00) \times 10^{-4}$

Table 7.9: Overview of the selection efficiencies for the $B_s^0 \to \tau^+\tau^-$ samples, split by step in the selection process, for 2012, 2016 and 2017.

Data statistics

The multiplicity of the data after the full selection can be seen in Tab. 7.10. It is about 1 for all samples as expected. A comparison between the Run1 and Run2 data efficiencies, split by the main steps of the selection process, is given in Tab. 7.11. The efficiencies are in agreement between Run1 and Run2.

Year	Polarity	Candidates	Events	Multiplicity
2016	MagDw	14 467	13 728	1.054 ± 0.016
2016	MagUp	13 702	13 014	1.053 ± 0.016
2017	MagDw	14 519	13 807	1.052 ± 0.015
2017	MagUp	13 977	13 327	1.049 ± 0.015
Sum		56 665	53 876	1.052 ± 0.008

Table 7.10: Statistics for the Run2 data samples after the NN1 stage in the signal region.

	Run1		Run2	
Step	Yield	Efficiency [%]	Yield	Efficiency [%]
+Reco Strip	18 101 739	-	71 073 609	-
+Loose Sel	1 424 442	7.87 ± 0.01	5 221 579	7.35 ± 0.00
+NN1	213 756	15.01 ± 0.00	883 177	16.91 ± 0.02
+Dalitz	14 690	6.87 ± 0.05	53 876	6.10 ± 0.02
Total		$(8.12 \pm 0.07) \times 10^{-2}$		$(7.58 \pm 0.03) \times 10^{-2}$

Table 7.11: Overview of some data efficiency and events yield for Run1 and Run2.

7.3 Likelihood fit

As for Run1, the fitted variable is the output of a second neural network, called NN2. Sec. 7.3.1 is devoted to this second neural network. The fit of the NN2 output is given in Sec. 7.3.2.

7.3.1 The fitted variable

The signal training sample for the NN2 consists of all signal MC events passing the full selection. The background training sample consists of all events from the data background region passing the full selection and having one candidate per event. The NN2 is based on 36 possible variables, listed in Tab. 7.12, including general event properties, τ properties, B properties, isolation variables and τ reconstruction variables.

Both the signal and background events on which the NN2 is trained are used later on to model the signal and background templates of the likelihood fit. To avoid overtraining, the k-folding technique (with $k = 10$) is therefore applied. For candidates in the signal and control region of the data the outputs of the 10 NNs are first flattened to uniformize their distribution and then averaged.

As for Run1, the data in the signal region whose NN2 output ≥ 0.7 are blinded. The fit strategy relies on the fact that the background composition of data in the control and the non blinded signal regions are similar. It is then important to have a good agreement of the shape of the NN2 output distribution for data between the control and the non-blinded signal regions. The distributions of the NN2 output in the control and the non blinded signal region for data and MC signal are shown in the upper plot of Fig. 7.2 for Run2. For reference, a similar plot for the NN2 output built for Run1 is shown in the central plot of Fig. 7.2. We note that the agreement between the control and the signal region for data is not good for Run2.

To know if one variable in particular is responsible for this disagreement, we compared the distribution in the control and the signal region of all the variables entering the NN2. We found that the agreement is not good for the τ masses. It is not surprising as these variables are extremely correlated with the variables used to build the regions, $m_{\pi_1^+\pi_2^-}$ and $m_{\pi_3^+\pi_2^-}$. Note that we observed a similar disagreement for the BDT2 built for the $(3\pi, \mu)$ final state analysis, we finally discarded the τ masses from the list of variables entering BDT2.

To check if the disagreement in the NN2 output is due to the τ masses, we build a new NN2 based on the same list of variables described in Tab. 7.12 but after τ masses have been removed. The distributions of this new version of the NN2 output in the control and the signal region for data and MC signal are shown in the lower plot of Fig. 7.2. The agreement between the control and the signal region is much better than for the initial NN2 (with the τ masses). The agreement is even better than for Run1. Hence we retain this version of the NN2 output to be fitted.

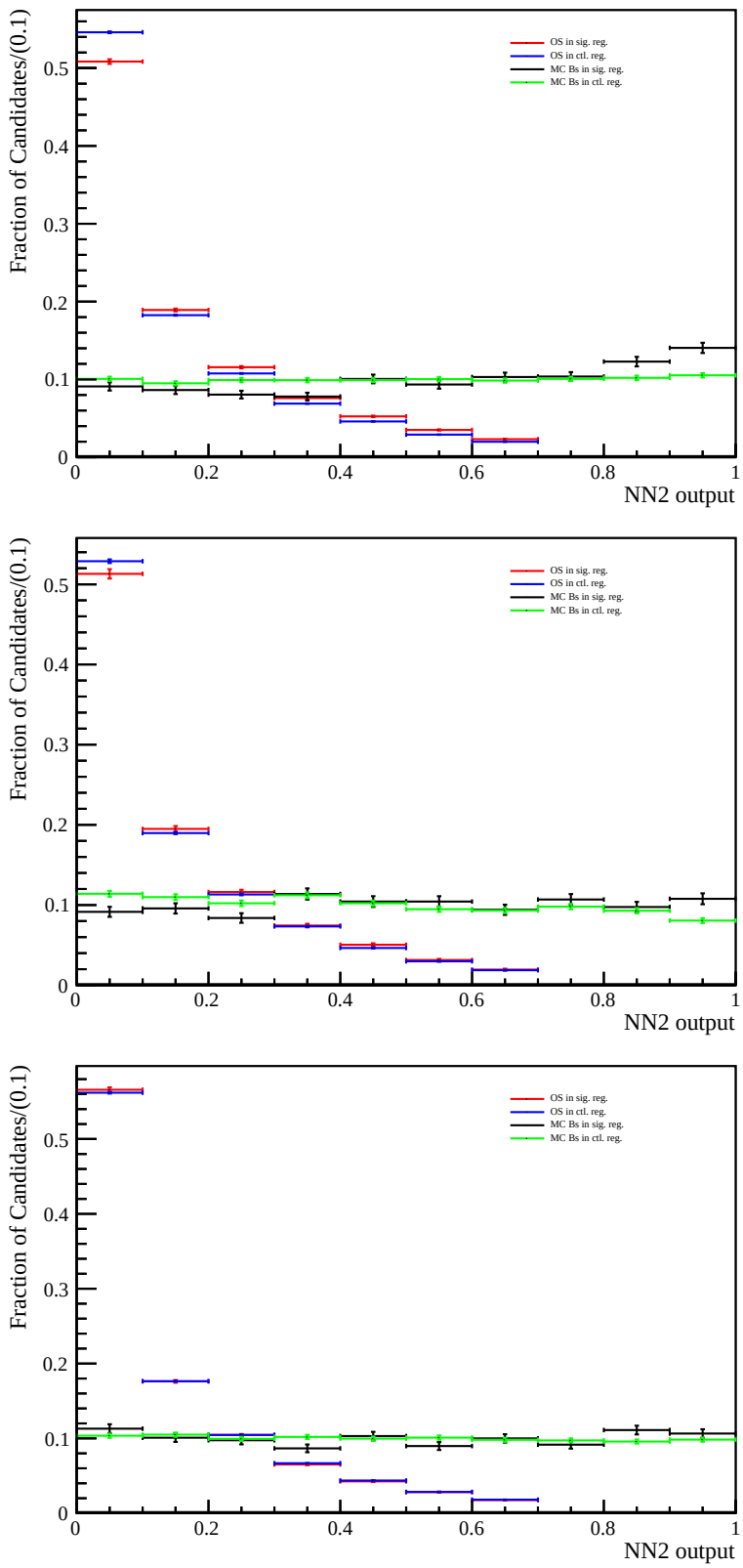


Figure 7.2: Distributions of the NN2 output in the control and the signal region for data and MC signal. Data in the signal region whose NN2 output ≥ 0.7 are blinded. The upper (central) plot shows the distribution for the NN2 trained on Run2 (Run1) samples. The lower plot shows the distribution for the NN2 trained on Run2 samples but where the τ masses have been discarded from the initial list of 36 variables.

Rank		Variable	Sig.
1	τ^+	Lifetime	32.3 σ
2	τ^-	Lifetime	33.2 σ
3	B_s^0	nc-vPT	32.3 σ
4	τ^-	Mass	32.3 σ
5	τ^+	Mass	35.9 σ
6		τ reconstruction variable $\mathcal{R}e[x_3]$	40.5 σ
7		$\sum_\pi$ BDTiso3	44.6 σ
8	B_s^0	p_T	29.2 σ
9	B_s^0	DIRA	26.2 σ
10	τ^+	Isolation smallest $\Delta\chi^2$	21.1 σ
11	τ^+	Decay vertex χ^2	18.8 σ
12	τ^-	Decay vertex χ^2	16.9 σ
13		τ reconstruction variable $\mathcal{I}m[\tilde{s}_3^\pm]$	14.8 σ
14		Missing mass	15.5 σ
15	B_s^0	Decay vertex χ^2	11.1 σ
16	B_s^0	Corrected mass	8.5 σ
17		τ reconstruction variable $\bar{\theta}_W$	7.8 σ
18	τ^-	Distance of closest approach	7.1 σ
19	τ^+	Distance of closest approach	7.1 σ
20		τ reconstruction variable $p^+ \cdot p^-$	6.8 σ
21		Number of real solutions for $p^+ \cdot p^-$	6.4 σ
22	B_s^0	Mass	5.8 σ
23		$\sum_\pi$ first isolation BDT output, 5th digit	5.2 σ
24		Total number of candidates	4.9 σ
25	τ^+	z displacement	4.0 σ
26		τ reconstruction variable $\mathcal{I}m[x_1]$	4.1 σ
27	τ^-	z displacement	4.3 σ
28		$\sum_\pi$ first isolation BDT output, 1st digit	4.1 σ
29		τ reconstruction variable $\mathcal{R}e[\xi_1]$	3.3 σ
30	B_s^0	Isolation smallest $\Delta\chi^2$	2.4 σ
31	τ^-	Transverse momentum	1.6 σ
32	τ^+	Transverse momentum	2.1 σ
33		τ reconstruction variable $\mathcal{R}e[\tilde{s}_1^\pm]$	2.1 σ
34	B_s^0	Neutral isolation cone variable PZ	2.0 σ
35		$\sum_\pi$ BDTiso1, 3rd digit	1.2 σ
36	B_s^0	Neutral isolation cone variable mult	0.8 σ

Table 7.12: List of input variables for the NN2, together with their significance. See the NeuroBayes manual [133] for the definition of the significance.

7.3.2 Fit strategy

The same fit strategy as for Run1 is used. We remind the reader that the fit model is given by

$$\mathcal{N}^{\mathrm{SR}}_{\mathrm{data}} = s \times \widehat{\mathcal{N}}^{\mathrm{SR}}_{\mathrm{sim}} + f_b \times \left(\mathcal{N}^{\mathrm{CR}}_{\mathrm{data}} - s \times \frac{\epsilon^{\mathrm{CR}}}{\epsilon^{\mathrm{SR}}} \times \widehat{\mathcal{N}}^{\mathrm{CR}}_{\mathrm{sim}} \right), \tag{7.3}$$

where $\mathcal{N}^{\mathrm{SR}}_{\mathrm{data}}$ ($\mathcal{N}^{\mathrm{CR}}_{\mathrm{sim/data}}$) is the NN2 output distribution in the signal (control) region from simulation/data, s is the signal yield, f_b is a scaling factor for the the background template, and ϵ^{SR} (ϵ^{CR}) is the signal efficiency in the signal (control) region. The hat on top of the NN2 output distributions means that the distributions are normalized to unity.

A toy validation has been performed with 500 toys. Only the setting including the statistical uncertainties, i.e. the Beeston Barlow technique has been tested. For the situation representing the SM (no signal), the distribution of fit values, fit uncertainties and pulls for the $B^0_s \to \tau^+\tau^-$ signal yield are shown on Fig. 7.3. Based on these toys, we expect to find on average a fit error of

$$\begin{aligned} \sigma^{toys}_{B^0_s} &= 140.2 \text{ (stat)}, \\ \sigma^{toys}_{B^0} &= 179.0 \text{ (stat)}. \end{aligned} \tag{7.4}$$

For Run1, the same toy analysis led to a fit error of

$$\begin{aligned} \sigma^{toys}_{B^0_s} &= 58.3 \text{ (stat)}, \\ \sigma^{toys}_{B^0} &= 60.7 \text{ (stat)}. \end{aligned} \tag{7.5}$$

We note that the fit errors for Run2 are between 2 and 3 times larger than for Run1. On one hand, the discriminating power of the NN2 is approximately the same between Run1 and Run2. This can be seen by comparing distributions between the central (Run1) and the lower plot (Run2) in the Fig. 7.2. On the other hand, the data yield in the signal region is about 3.6 times larger for Run2 than for Run1 (53 876 vs 14 690), which explain the difference of fit errors.

The fit error for Run2 is larger on B^0 than on B^0_s. This can be explained by a difference of signal templates. The templates $\mathcal{N}^{\mathrm{SR}}_{\mathrm{sim}}$ for B^0_s and B^0 are shown in Fig. 7.4. The larger fit error seems to be due to the difference of trends between B^0_s and B^0. The B^0 template is slightly more peaked to lower value of NN2 (like background). This difference of trends has not been observed in the published analysis or the analysis of the $(3\pi, \mu)$ final state; it could be due to statistical fluctuation.

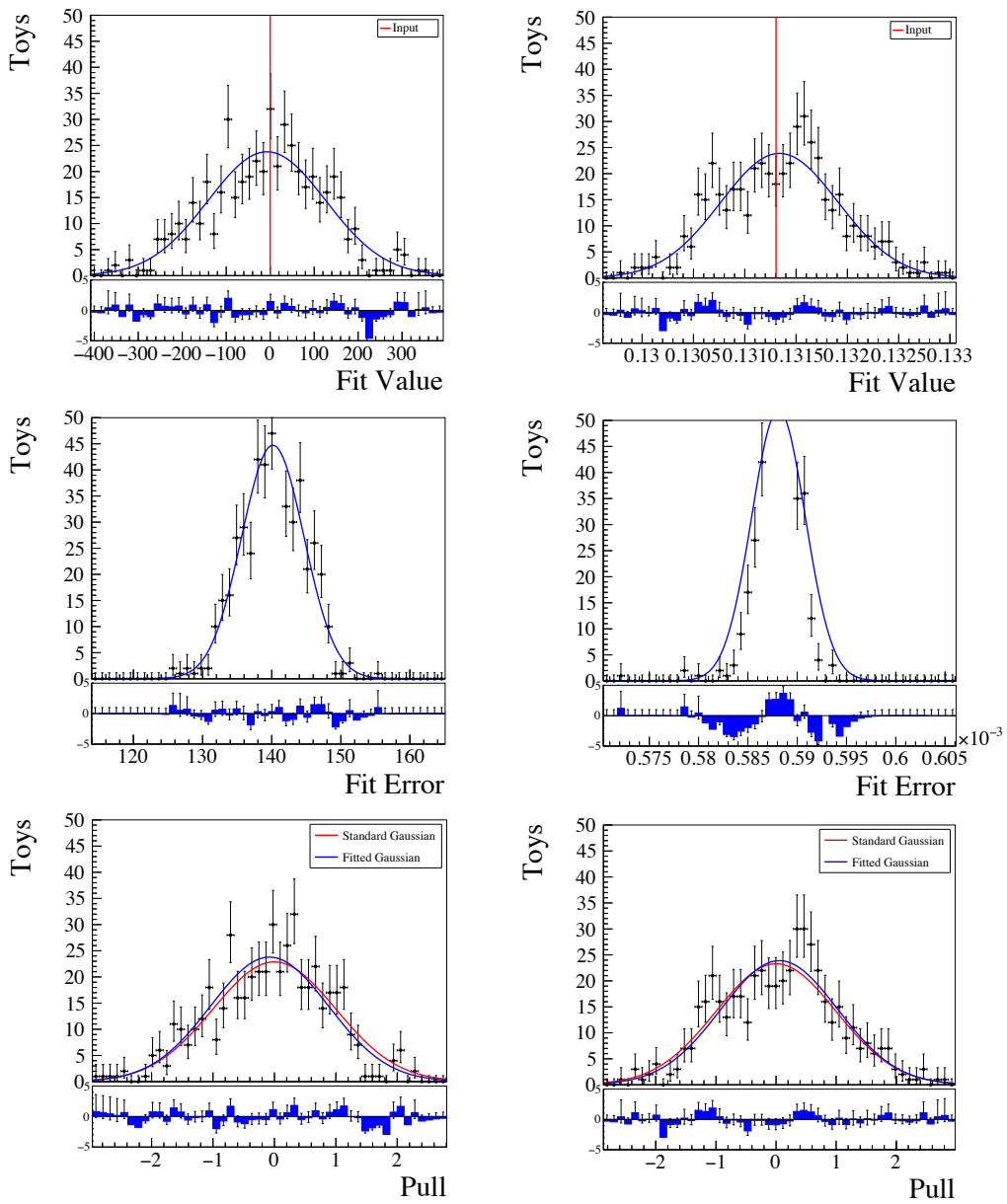


Figure 7.3: Results for the fit validation study for the $B_s^0 \to \tau^+\tau^-$ mode, based on 500 pseudo experiments. Shown are the distribution of fit values [Top], fit uncertainties [Middle] and pulls [Bottom] on the signal yield s [Left] and on the scale factor f_b [Right] for the nominal setting including only statistical uncertainties.

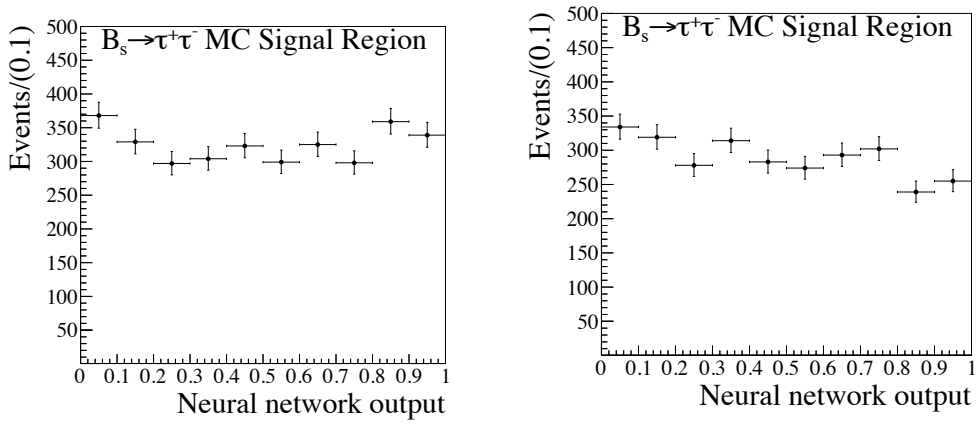


Figure 7.4: $\mathcal{N}_{\text{sim}}^{\text{SR}}$ templates for B_s^0 (B^0) on the left (right) side used in the likelihood fit.

7.4 Normalization

This section is devoted to the normalization channel $B^0 \to D_s^+(K^+K^-\pi^+)D^-(\pi^+\pi^-K^-)$, which has the same number of charged tracks in the final state as $B_{(s)}^0 \to \tau^+\tau^-$ and a similar decay topology.

Sec. 7.4.1 gives an overview of the analysis of the normalization channel. Sec. 7.4.2 presents the calculation of the normalization factors.

7.4.1 Overview of the $B^0 \to D^-D_s^+$ analysis

The same selection as for Run1 is applied to the Run2 samples. An overview of the selection cuts applied in the stripping is given in Tab. 7.2. Successfully reconstructed $B^0 \to D^-D_s^+$ candidates are required to pass the following additional selection cuts.

- *Trigger*: The same trigger conditions as for $B_s^0 \to \tau^+\tau^-$ are required (see Tab. 7.4).
- *D meson mass cuts*: To select the DD_s final state, one D is required to decay into the $\pi^+K^-\pi^+$ final state and have an invariant mass $m \in [1855, 1885]$ MeV/c^2, while the other D is required to decay into the $K^+K^-\pi^+$ final state and have an invariant mass $m \in [1955, 1985]$ MeV/c^2.

In case multiple candidates still pass these selection requirements, one candidate per event is chosen at random and all others are discarded. The efficiencies split by steps of the selection are given in Tab. 7.13.

Likelihood Fit

We fit the mass distribution of the selected $B^0 \to D^-D_s^+$ candidates following the strategy used in the published analysis, described in the following.

The reconstructed $B^0 \to D^-D_s^+$ candidates receive contributions from four main sources: the $B^0 \to D^-D_s^+$ signal, a combinatorial background component and two partially reconstructed peaking backgrounds, $B^0 \to D^{*-}D_s^+$ and $B^0 \to D^-D_s^{*+}$.

2016	Cut-and-Count	
Step	Yield	Efficiency [%]
+Geom. Acc.	1011701	12.545±0.022
+Reco + Strip	19585	1.94±0.01
+Truth Match	18378	93.84±0.17
+PID	14373	78.21±0.30
+L0	7644	53.18±0.42
+Hlt	7054	92.28±0.31
+Loose Sel	5982	84.80±0.43
Total		$(7.42 \pm 0.10) \times 10^{-2}$

2017	Cut-and-Count	
Step	Yield	Efficiency [%]
+Geom. Acc.	1064533	12.560±0.034
+Reco + Strip	21016	1.97±0.01
+Truth Match	19742	93.94±0.16
+PID	15652	79.28±0.29
+L0	10814	69.09±0.37
+Hlt	9899	91.54±0.27
+Loose Sel	8494	85.81±0.35
Total		$(10.02 \pm 0.11) \times 10^{-2}$

Average		$(8.753 \pm 0.074) \times 10^{-2}$

Table 7.13: Overview of the selection efficiencies for the $B^0 \to D^- D_s^+$ samples for Run2.

These latter are characterized by a double-horned structure. To model these contributions, the strategy set out in Ref. [135, 136], where the shape of the $B^0 \to D^{*-} D_s^+$ and $B^0 \to D^- D_s^{*+}$ mass are obtained from simulation, is followed. The fitted PDF contains the followed components:

- $B^0 \to D^- D_s^+$ Signal: The B^0 signal distribution is modeled using an double-sided Hypatia function, described in Ref. [147], with the parameters β and ζ fixed to zero. The parameters describing the tails, a_i and n_i, as well as λ are obtained from simulation. The parameters μ and σ are fitted directly on the data.

- Combinatorial Background: The mass line shape of the combinatorial background component is modeled by an exponential function, and fitted directly on the data.

- $B^0 \to D^{*-} D_s^+$: The peaking background from $B^0 \to D^{*-} D_s^+$ has two main contributions: one from the subsequent $D^{*-} \to D^- \pi^0$ decay, and one from the $D^{*-} \to D^- \gamma$ decays. The mass PDF of the $B^0 \to D^{*-} D_s^+$ peaking background is obtained from simulation. Truth matched events are taken from the $B^0 \to D^{*-} D_s^+$ MC sample, and required to pass the same stripping and offline selection as the $B^0 \to D^- D_s^+$ mode. The line shape is modeled using

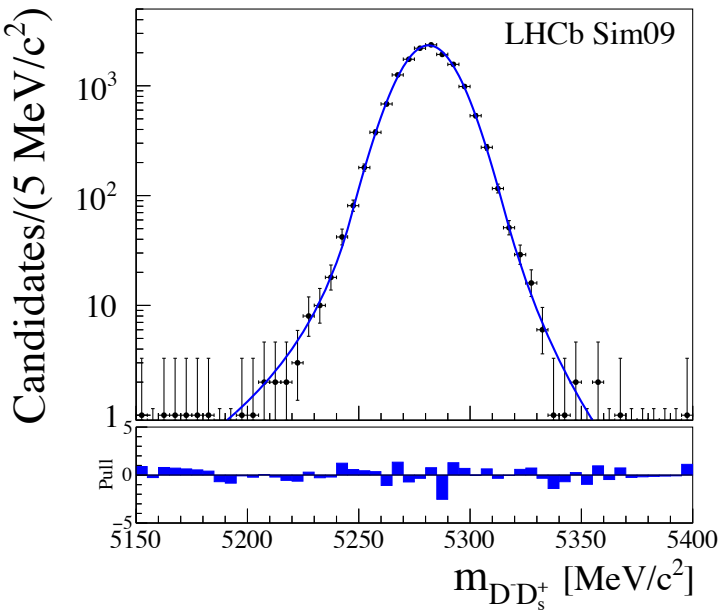


Figure 7.5: The invariant mass distribution [Black points] of the $B^0 \to D^- D_s^+$ signal, fitted by an Hypatia function in blue.

Par.	Simulation	Data
μ	5281.1 ± 0.11	5283.6 ± 0.15
σ	12.843 ± 0.144	13.61 ± 0.14
λ	-9.938 ± 3.06	N/A
a_1	2.718 ± 0.25	N/A
n_1	2.878 ± 0.81	N/A
a_2	2.675 ± 0.39	N/A
n_2	5.159 ± 2.45	N/A

Table 7.14: Results on the parameters of the Hypatia function obtained from a fit to the mass distribution of the $B^0 \to D^- D_s^+$ signal in Run2 simulation and data.

three Gaussian functions.

- $B^0 \to D^- D_s^{*+}$: The mass PDF of the $B^0 \to D^- D_s^{*+}$ peaking background is obtained from simulation, using its partner decay $B_s^0 \to D_s^- D_s^{*+}$. Truth matched events are taken from the $B_s^0 \to D_s^- D_s^{*+}$ MC sample, and required to pass the same stripping and trigger selection as the $B^0 \to D^- D_s^+$ mode. In addition, both reconstructed D mesons are required to decay into $K^+K^-\pi^+$ with $m \in [1955, 1985]$ MeV/c^2. The line shape is modeled using three Gaussian functions.

Note that we have used Run1 samples to model these partially reconstructed peaking backgrounds as the Run2 samples were not produced at the time of this work. The 2016 and 2017 B^0 signal distribution is modeled using an double-sided Hypatia function, described in Ref. [147], with the parameters β and ζ fixed to zero. The fit of the $B^0 \to D^- D_s^+$ signal component is shown in Fig. 7.5, and listed in Tab. 7.14. The nominal model used to fit the $D^- D_s^+$ invariant mass distribution in data combines the four above described PDF components. The resulting fit is shown in Fig. 7.6, and

Mode	Yield
$B^0 \to D^- D_s^+$	$26\,340 \pm 176$
Comb. Bkg	$21\,355 \pm 416$
$B^0 \to D^{*-} D_s^+$	$6\,188 \pm 280$
$B^0 \to D^- D_s^{*+}$	$22\,633 \pm 544$

Table 7.15: Yields of the four different components used in the nominal fit to describe the $D^- D_s^+$ invariant mass distribution in data.

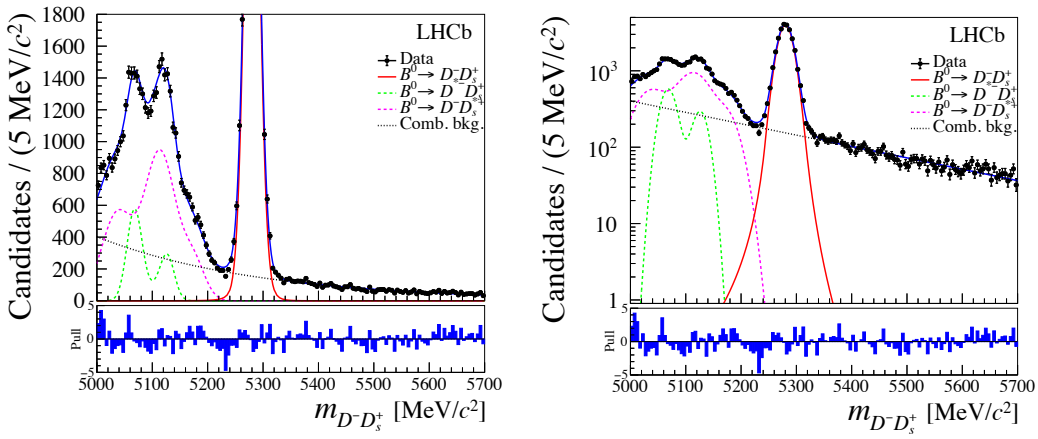


Figure 7.6: The invariant mass distribution of the reconstructed $B^0 \to D^- D_s^+$ events in data [Black points], together with the nominal PDF [Blue] used to extract the $B^0 \to D^- D_s^+$ yield, in linear scale [Left] and in log scale [Right] for Run2. The individual components are described in the text.

the resulting yields for each of the components is listed in Tab. 7.15. It leads to a $B^0 \to D^- D_s^+$ yield of

$$N^{\text{obs}}_{D^- D_s^+} = 26\,340 \pm 176\,, \tag{7.6}$$

where the uncertainty is statistical only. Note that the yield for Run1 was of $10\,629 \pm 114$.

7.4.2 Normalization factors

The observed number of $B^0_{(s)} \to \tau^+\tau^-$ events is related to the $B^0_{(s)} \to \tau^+\tau^-$ branching ratio through the normalization factors $\alpha_{d/s}$ via:

$$\begin{aligned}\mathcal{B}(B^0_{(s)} \to \tau^+\tau^-) &= \frac{\mathcal{B}(B^0_{(s)} \to \tau^+\tau^-(6\pi))}{\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau)^2} \\ &= \frac{N^{\text{obs}}_{B^0_{(s)}\to\tau^+\tau^-(6\pi)} / \epsilon^{\text{tot}}_{B^0_{(s)}\to\tau^+\tau^-(6\pi)}}{\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau)^2 N_{B^0_{(s)}}} \equiv \alpha_{d/s} \times N^{\text{obs}}_{\tau^+\tau^-(6\pi)}\,,\end{aligned} \tag{7.7}$$

where $\epsilon^{\text{tot}}_{B^0_{(s)}\to\tau^+\tau^-}$ is the total $B^0_{(s)} \to \tau^+\tau^-$ reconstruction and selection efficiency and $N_{B^0_{(s)}}$ is the number of produced $B^0_{(s)}$ mesons in the data using the normalization channel. The normalization factor for B^0 is given by

$$\alpha_d = \frac{\epsilon^{D^- D_s^+} \cdot \mathcal{B}(B^0 \to D^- D_s^+) \cdot \mathcal{B}(D^+ \to \pi^+ K^- \pi^+) \cdot \mathcal{B}(D_s^+ \to K^+ K^- \pi^+)}{N^{\text{obs}}_{D^- D_s^+} \cdot \epsilon_{B^0 \to 6\pi^\pm} \cdot [\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau)]^2}\,. \tag{7.8}$$

Quantity	Value	Ref.
$\mathcal{B}(D^+ \to \pi^+K^-\pi^+)$	$(8.98 \pm 0.28) \times 10^{-2}$	[128]
$\mathcal{B}(D_s^+ \to K^+K^-\pi^+)$	$(5.45 \pm 0.17) \times 10^{-2}$	[128]
$\mathcal{B}(B^0 \to D^-D_s^+)$	$(7.2 \pm 0.8) \times 10^{-3}$	[148]
$\mathcal{B}(\tau^- \to \pi^-\pi^+\pi^-\nu_\tau)$	$(9.31 \pm 0.05) \times 10^{-2}$	[128]

Table 7.16: Overview of the input branching ratios needed to calculate the normalization factor $\alpha_{d/s}$.

	Run2	Run1	Run2/Run1
$N^{\text{obs}}_{D^-D_s^+}$	$(2.63 \pm 0.02) \times 10^4$	$(1.06 \pm 0.01) \times 10^4$	2.48 ± 0.03
$\epsilon^{D^-D_s^+}$	$(8.75 \pm 0.07) \times 10^{-4}$	$(5.79 \pm 0.03) \times 10^{-4}$	1.51 ± 0.01
$\epsilon^{B^0 \to 6\pi^\pm}$	$(3.27 \pm 0.06) \times 10^{-5}$	$(1.68 \pm 0.05) \times 10^{-5}$	1.95 ± 0.07
$\epsilon^{B_s^0 \to 6\pi^\pm}$	$(3.72 \pm 0.07) \times 10^{-5}$	$(1.84 \pm 0.04) \times 10^{-5}$	2.02 ± 0.06
α_d	$(4.10 \pm 0.50) \times 10^{-6}$	$(1.32 \pm 0.23) \times 10^{-5}$	0.31 ± 0.01
α_s	$(1.40 \pm 0.19) \times 10^{-5}$	$(4.70 \pm 0.60) \times 10^{-5}$	0.30 ± 0.01

Table 7.17: Overview of the normalization factor $\alpha_{d/s}$ as well as different quantities entering it.

Mode	SM Yield $N^{\text{exp}}_{\tau\tau}$
$B_s^0 \to \tau^+(3\pi)\tau^-(3\pi)$	0.0551 ± 0.0082
$B^0 \to \tau^+(3\pi)\tau^-(3\pi)$	0.0054 ± 0.0008

Table 7.18: Overview of the expected SM yield for $B_s^0 \to \tau^+\tau^-$ and $B^0 \to \tau^+\tau^-$ for Run2.

In the case of the B_s^0 meson, the normalization factor α_s must be divided by the ratio of B_s^0 to B^0 production factor $f_s/f_d = 0.259 \pm 0.015$ [137] and the efficiency $\epsilon_{B^0 \to 6\pi^\pm}$ must be replaced by $\epsilon_{B_s^0 \to 6\pi^\pm}$.
The branching ratios are listed in Tab. 7.16. The total efficiencies are determined from MC simulation. The $\alpha_{d/s}$ factors for Run1 and Run2 are compiled in Tab. 7.17. Assuming the SM branching ratio for $B_s^0 \to \tau^+\tau^-$ and $B^0 \to \tau^+\tau^-$, the expected number of signal events is given in Tab. 7.18.

7.5 Results and conclusions

As for the $(3\pi, \mu)$ final state, the statistical fit error and the normalization factor can be combined into an estimate on the limit on the $B^0 \to \tau^+\tau^-$ and $B_s^0 \to \tau^+\tau^-$ branching ratios via:

$$\mathcal{UL} = 1.3(1.6) \times \alpha \times \sigma^{\text{toys}} \text{ at } 90(95) \text{ \% C.L.}, \tag{7.9}$$

where σ^{toys} is the statistical fit error. The efficiencies, the normalization factors, the fit errors on the signal yield, σ^{toys}, and the limit estimate for B_s^0 and B^0 are compiled in Tab. 7.19.
We have to keep in mind that our estimate is based on a fit framework which does not include any kind of systematics and that the σ^{toys} entering the estimate is purely statistic. Moreover, the signal efficiency entering the normalization factor and the

	B_s^0		B^0	
	Run2	Run1	Run2	Run1
$\epsilon^{\text{tot}} \times 10^5$	3.7	1.8	3.3	1.7
$\alpha \times 10^5$	1.4	4.7	0.41	1.3
σ^{toys}	140	58	179	61
$\mathcal{UL}$ at 95 % C.L.	3.1×10^{-3}	4.4×10^{-3}	1.1×10^{-3}	1.3×10^{-3}
$\frac{\mathcal{UL}_{\text{Run2}}}{\mathcal{UL}_{\text{Run1}}}$	0.7		0.8	

Table 7.19: Signal efficiencies, normalization factors, fit errors and limit estimates at 95% C.L. for the $(3\pi, 3\pi)$ final state and for Run1 and Run2.

	B_s^0	B^0
$\epsilon^{\text{tot}}(10^{-5})$ corrected	2.4	2.2
$\alpha(10^{-5})$ corrected	4.1	1.2
σ^{toys} included syst. uncertainties	67	70
$\mathcal{UL}$ at 95 % C.L. (10^{-3}) expected	7.4	2.1
$\mathcal{UL}$ at 95 % C.L. (10^{-3}) real	6.8	2.1

Table 7.20: Efficiencies and normalization factors corrected from data-MC disagreements, fit results including systematics uncertainties, and the published limits at 95% C.L. for the published analysis [45].

signal NN2 distribution are not corrected from possible data-MC disagreements. Our estimates are compared to the expected limit estimates at 95% C.L. obtained using the Run1 data in the same condition: no data-MC corrections applied to the efficiencies and a σ^{toys} extracted from a similar toys study.
On one hand, the fit errors σ^{toys} obtained for $B_s^0 \to \tau^+\tau^-$ and $B^0 \to \tau^+\tau^-$ are larger for Run2 than for Run1. On the other hand, the signal efficiencies are also larger resulting in smaller normalization factors for Run2 than for Run1. Once these two trends combined, it turns out that the expected upper-limit obtained with Run2 (2016+2017) are smaller than the one obtained for Run1 by a factor 0.7 for B_s^0 and 0.8 for B^0.

The worse ratio $\frac{\mathcal{UL}_{\text{Run2}}}{\mathcal{UL}_{\text{Run1}}}$ for B^0 is due to larger fit error σ^{toys} on B^0. This larger fit error is explained in Sec. 7.3.2.

For completeness, the efficiencies and normalization factors corrected from data-MC disagreements, the fit results including systematics uncertainties, and the published limits of [45] are given in Tab. 7.20.
We remind the reader that this analysis has not been optimized for Run2, and some gain could be obtained by optimizing the selection process (loose cut-based selection, choice of variables entering the NN1, etc...) as well as the fitted variables (mainly the choice of variables entering NN2).

Moreover the τ masses have been discarded from the NN2 in spite of their good discriminating power; using their discriminating power more intensively during the selection process could be a good handle to increase the background rejection.

The large fit errors obtained for Run2 are due to the large data yields in the signal

region. A more drastic selection would decreases these fit errors but also decreases the signal efficiencies, hence a compromise has to be found.

In addition, only $\sim$3.4 fb^{-1} of data have been analyzed (2016 and 2017), which is approximately the same amount of data than for Run1 (3 fb^{-1}). The amount of data for the full Run2 (2015-2018) is $\sim$6 fb^{-1}, which is roughly twice that of Run1. As mentioned previously, the cross-section $\sigma_{b\bar{b}}$ in Run2 is doubled with respect to those of Run1. Hence, we expect in Run2 roughly four times more B mesons produced than in Run1 resulting approximatively in a factor 2 improvement in the limit.

Chapter 8

Conclusions

In this thesis, two FCNC decays have been tackled from two different angles. We performed a phenomenological study of the $D^+ \to \pi^+\ell^+\ell^-$ decay and an experimental search for the $B^0_{(s)} \to \tau^+\tau^-$ decays.

8.1 Conclusions on the phenomenological study of the $D^+ \to \pi^+\ell^+\ell^-$ decay

We presented in this thesis the work performed to improve the theory description of the $D^+ \to \pi^+\ell^+\ell^-$ decay. This decay is dominated by resonances and long-distance hadronic contributions and is so far poorly understood. Following [36], our result includes the most recent calculation of the Wilson coefficients for the $c \to u\ell^+\ell^-$ transition at the next-to-leading logarithm approximation [83] and the main next-to-leading order contributions in an expansion in the strong coupling including the non-factorizable corrections, calculated first for B decays in [35]. We also use the latest lattice results for the form factors that were not available for previous studies. Compared to the past phenomenological analysis, we use different resonance description.

Initially, the resonances were added by hand, by means of Breit-Wigner functions, on top of a non-resonant background described by the partonic result for the quark vacuum polarization, e.g. see [37,38]. In Ref. [36] an alternative is presented, in which vacuum polarizations are reconstructed from their imaginary part using a dispersion relation. The imaginary part of the vacuum polarizations are modeled following a proposal by M. Shifman [39], i.e. a Breit-Wigner type function modeling the dominant vector resonance plus the sum of an infinite tower of resonances, starting from the first excitation. In this approach, the partonic result is recovered asymptotically.
Our work improves upon this resonance description in the following ways:

- In Ref. [36], the isospin 1 (ρ^0) and isospin 0 (ω^0) light resonances are not treated separately, but in terms a single tower of resonances with "effective" parameters. We improve upon this model by treating separately the isospin contributions.

- Instead of using "effective" parameters, we followed the strategy of [40], where the authors relate the imaginary part of the charm hadronic spectrum to the experimental observable:

$$R(s) \equiv \frac{3s}{4\pi\alpha_e^2}\sigma_{e^+e^-\to hadrons(\gamma)}(s), \tag{8.1}$$

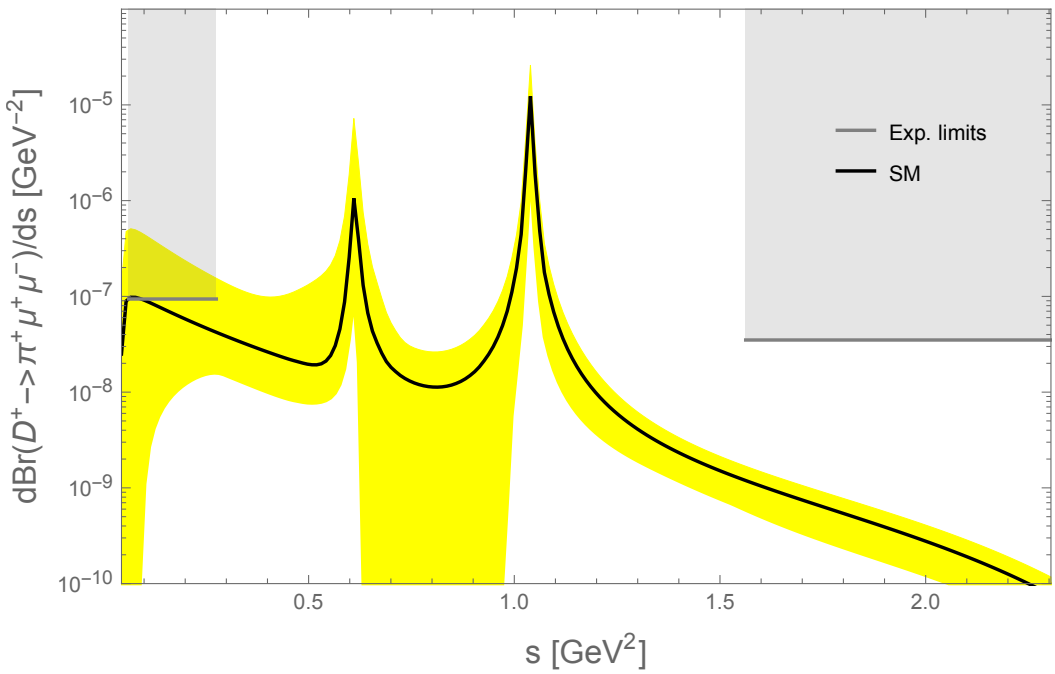


Figure 8.1: Distribution over s (the dilepton invariant mass squared) of the $D^+ \to \pi^+\mu^+\mu^-$ branching ratio. The related uncertainty is in yellow and the experimental excluded regions are in gray [42].

Armed with this resonance description and the framework presented in Chap. 2, we gave predictions for the $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$ distribution as shown in Fig. 8.1. As in [36] for the $D^0 \to \rho^0\ell^+\ell^-$, we found that the non-factorizable annihilation contribution, neglected in [37, 38], is one of the largest SM contributions in the low s region. Its impact on the SM distribution of the branching ratio is visible by the large increase in low s region with respect to the ones shown in Ref. [37,38].

We then computed, with the help of the branching ratio distribution, new bounds on BSM Wilson coefficients imposed by the current experimental upper limits on the $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$ [42] and $\mathcal{B}(D^0 \to \mu^+\mu^-)$ [41]. Our results are summarized in Tab. 3.8. We compared our results to the upper limits from the scalar leptoquark model presented in Ref. [38], whose Wilson coefficient values were provided in Eq. (3.62), and found that C_P, C_S and C_9' are of the same order as our model independent bounds. In the case of the tensor contributions, C_T and C_{T5}, the leptoquark bounds are smaller by a factor $\sim$80 but as we saw, they are still large enough to produce effects distinguishable from the SM.

The high s region is more interesting as the bounds on BSM physics are such that the BSM predictions could be distinguished from the SM in a future measurement. In addition, the error band obtained in this region on the SM prediction is thiner than that in low s region. For illustrative purposes, we showed several distributions where the BSM Wilson coefficients were varied individually. We found that scenarios where C_T, C_P, C_{10} and C_9 take the maximal values allowed by the experimental measurements as summarized in Tab. 3.8, were fully distinguishable from the SM. We also showed that BSM effects can be seen for value of C_{T5} up to $C_{T5}/10 \simeq 0.08$. In addition, we found that distribution for the maximal Wilson coefficients allowed by the scalar leptoquark model of Ref. [38] lies just below the upper limit in Reg. II

which reinforce the interest for future measurements in this region.

The effect of the phases of the resonances on the branching ratio distribution has also been explored and we found that they were not measurables except perhaps for some value of φ_ω around $7\pi/10$.

We then pursued our phenomenological analysis with the observables F_H, $A_{\rm FB}$ and A_{CP}. We observed that the SM prediction for F_H is very clean and is not zero in the low s region. For illustrative purposes, we showed distributions where C_T, C_{T5}, C_P and C_{10} take the maximal values allowed by the experimental measurements as summarized in Tab. 3.8. We found that all these BSM distributions are sensitive to the resonances but can be distinguished from SM in the high s region, particularly the contributions of the tensors. In addition, the F_H distribution for the maximum allowed Wilson coefficients in the scalar leptoquark model of Ref. [38] appears to be also distinguishable from the SM in the high s region.

The study of $A_{\rm FB}$ shows that this observable is only sensitive to combinations of BSM coefficients, i.e. C_{T5} and C_P or C_T and C_S. We plotted the distribution of $A_{\rm FB}$ for BSM scenarios where both C_P and C_{T5} reaches the maximal allowed value given in Tab. 3.8, as well as similar scenarios but where C_{T5} is divided by 10 and 100. We found that even for small value of C_{T5} ($\sim$0.08), the distribution can still be distinguished from the SM which is approximately zero for the full spectrum. Therefore, even for values of tensor coefficients much below their limit the BSM effects should be observable, depending on the future experimental sensitivity. We also found that, as for F_H, the observable is sensitive to the resonances (which decrease the observables rendering the BSM effects harder to detect).

The observable A_{CP} was found to be inefficient to probe BSM physics as the related uncertainty is large and even for the maximal value of C_T, the observable is compatible with zero through the majority of the spectrum.

To summarize, the branching ratio, F_H and $A_{\rm FB}$ were found to be very good candidates to probe BSM models. $A_{\rm FB}$ and F_H are very clean and null observables in the SM in the high s region. The branching ratio is not zero in this region but we found that BSM physics, like the one allowed by the scalar leptoquark model of Ref. [38], can be distinguished from the SM. As mentioned, the three observables are sensitive to the resonance. In the case of $A_{\rm FB}$ and F_H, they decrease the observables rendering the BSM effects harder to detect. They also show a larger discrepancy from the SM in the high s region. Therefore, we advocate the measurement in the s region above the ϕ resonance: $s > 1.2$ GeV^2 for these three observables.

The measurement presented in Ref. [42] is based on proton-proton collision data, corresponding to an integrated luminosity of $1.0\,\mathrm{fb}^{-1}$ collected by LHCb at the center-of-mass energy of 7 TeV. A measurement of the full Run1 and Run2 samples ($\sim 9\,\mathrm{fb}^{-1}$) is on going at LHCb. According to our prediction, it is not impossible that LHCb observes some signal in the low s region. In Ref. [149], the LHCb collaboration claims that the Upgrade II (if approved) is expected to improve the limit on $\mathcal{B}(D^+ \to \pi^+\mu^+\mu^-)$, currently set at 7.3×10^{-8} at 90% CL, by more than one order of magnitude. The expected upper limits are about 1.3×10^{-8} with 23 fb^{-1} and 0.37×10^{-8} with $300\,\mathrm{fb}^{-1}$. In addition, LHCb will have the ability to measure

angular observables such as the forward-background asymmetry A_{FB}, which would provide additional handles to separate the long distance from the short distance contributions.

8.2 Conclusions on the search for $B^0_{(s)} \to \tau^+\tau^-$ decays at LHCb

Two analyses on the search for $B^0_{(s)} \to \tau^+\tau^-$ decays at LHCb have been presented in this thesis. A first one where we analyze Run1 data via the $(3\pi, \mu)$ final state, and a second one where we use Run2 data and the $(3\pi, 3\pi)$ final state. The upper limits at 95% C.L. for the Run1 analysis of the $(3\pi, 3\pi)$ final state are reminded in Tab. 8.1. The expected upper limits for the two analyses presented in this thesis, as well as the ratio of which over the expected upper limits on $\mathcal{B}(B^0_s \to \tau^+\tau^-)$ for the Run1 analysis of the $(3\pi, 3\pi)$ final state, are given in Tab. 8.2. The upper limits do not take into account corrections from possible data-MC disagreements and systematic uncertainties.

The worse limits obtained with the $(3\pi, \mu)$ final state is due to the difficulty to discriminate the signal from the abundant semileptonic background. The analysis of Run2 data with the $(3\pi, 3\pi)$ final state shows promising results and I advocate the use this final state for the full Run2 analysis.

The prospects on the evolution of the upper limit on the $B^0_s \to \tau^+\tau^-$ decay at 95% C.L. by the LHCb collaboration are shown in Fig. 8.2. The red curve takes only into account the luminosity gain and the analysis based only on the $(3\pi, 3\pi)$ final state without improvement. The red curve extrapolates a situation where the limit is improved by a factor 2 from the trigger improvement and the inclusion of a tracking system inside the magnet for low momentum pions and another factor 2 from the analysis improvement. The SM prediction and its uncertainty are in green, we remind the reader that $\mathcal{B}(B^0_s \to \tau^+\tau^-)^{\mathrm{SM}} = (7.73 \pm 0.49) \times 10^{-7}$ [28].

As we can see on Fig. 8.2, the analysis of the Run1 and Run2 data combined leads to the following limit estimate:

$$\mathcal{UL}(B^0_s \to \tau^+\tau^-)^{\mathrm{prospect}}_{(\mathrm{Run1-2})} \sim 3 \times 10^{-3} \text{ at 95 \% C.L..} \tag{8.2}$$

This estimate is ~ 2.3 smaller than the published limit set with the full Run1 data. This improvement on the limit could already exclude some of the phase space of models that lie close to the current experimental limit. For example, in Ref. [21], the authors present a U_1 vector leptoquark model whose parameters are fitted to several experimental anomalies, $R_K^{(*)}$, $R_D^{(*)}$, but also others measurements. Their model suggest that the branching ratio of the $B^0_s \to \tau^+\tau^-$ and $B \to K\tau^+\tau^-$ decays could be enhanced by several orders of magnitude above the SM predictions. They show on Fig. 8.3 (taken from [21]) the preferred 2D fit regions of their model for the $\mathcal{B}(B^0_s \to \tau^+\tau^-)$ and $\mathcal{B}(B \to K\tau^+\tau^-)$ observables. The limit presented in Eq. (8.3) would corner their $\Delta^2_\chi \leq 6.18(2\sigma)$ region (in blue on the plot).

At the end of the LHC data-taking period, the LHCb collaboration expects to set

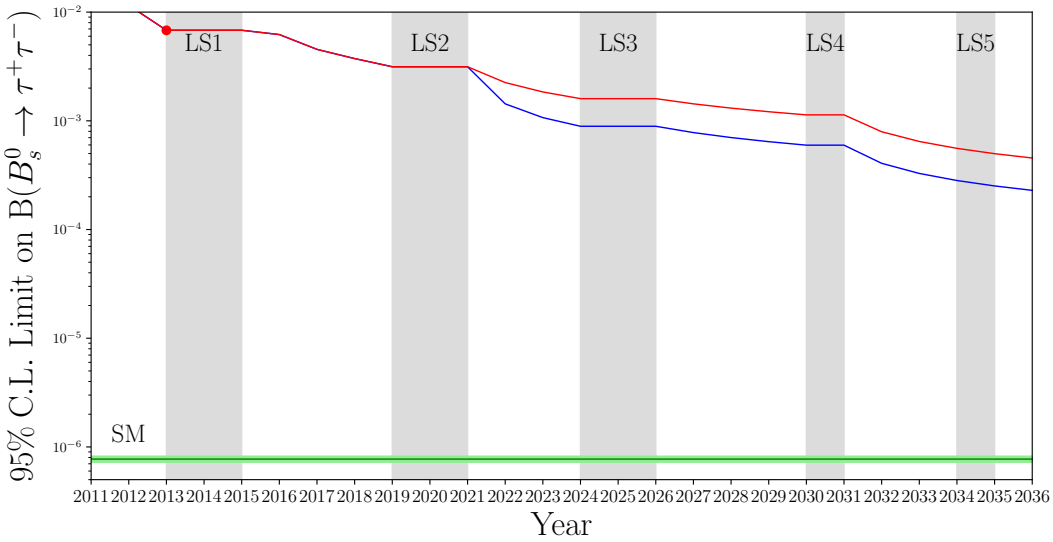


Figure 8.2: Prospects on the evolution of the upper limit on the $B_s^0 \to \tau^+\tau^-$ decay at 95 % C.L. by the LHCb collaboration. The red curve assumes only luminosity gain and analysis based only on the $(3\pi, 3\pi)$ final state without improvement. The red curve extrapolates a situation where the limit is improved by a factor 2 from the trigger improvement and the inclusion of a tracking system inside the magnet for low momentum pions and another factor 2 from the analysis improvement. The SM prediction and its uncertainty are shown in green.

	$(3\pi, 3\pi)$ Run1 B_s^0	$(3\pi, 3\pi)$ Run1 B^0
$\mathcal{UL}$ at 95 % C.L.	4.4×10^{-3}	1.3×10^{-3}

Table 8.1: Expected upper limits on $\mathcal{B}(B_s^0 \to \tau^+\tau^-)$ at 95% C.L. for the Run1 analysis of the $(3\pi, 3\pi)$ final state.

	B_s^0		B^0	
	$(3\pi, \mu)$ Run1	$(3\pi, 3\pi)$ Run2	$(3\pi, \mu)$ Run1	$(3\pi, 3\pi)$ Run2
$\mathcal{UL}$ at 95 % C.L.	2.4×10^{-2}	3.1×10^{-3}	7.0×10^{-3}	1.1×10^{-3}
$\frac{\mathcal{UL}}{\mathcal{UL}^{\text{Run1}}_{(3\pi,3\pi)}}$	5.7	0.7	5.4	0.8

Table 8.2: Expected upper limits on $\mathcal{B}(B^0 \to \tau^+\tau^-)$ at 95% C.L. for the two analysis presented in this thesis as well as the ratio of which over the expected upper limits on $\mathcal{B}(B_s^0 \to \tau^+\tau^-)$ for the Run1 analysis of the $(3\pi, 3\pi)$ final state.

Observables	LHCb after Run3	LHCb after Run5	Belle II 5 ab^{-1}	Belle II 50 ab^{-1}
$\mathcal{B}(B_s^0 \to \tau^+\tau^-) \times 10^4$	6.4	3.6	8.1	-
$\mathcal{B}(B^0 \to \tau^+\tau^-) \times 10^4$	1.8	0.46	3.0	0.96

Table 8.3: Estimate of the limit at 90% C.L. at different stages of the data acquisition period for Belle II [150] and LHCb.

the following limit at 95% C.L. on the $B_s^0 \to \tau^+\tau^-$ decay:

$$\mathcal{UL}(B_s^0 \to \tau^+\tau^-)^{\text{prospect}}_{(\text{Run1}-5)} \sim 4.5 \times 10^{-4} \text{ at 95 \% C.L..} \tag{8.3}$$

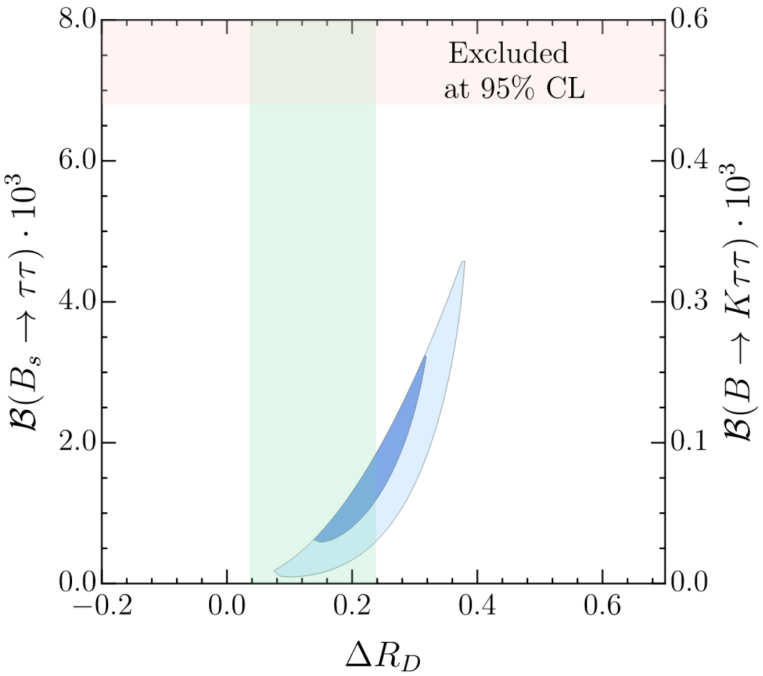


Figure 8.3: Preferred 2D fit regions of an U_1 vector leptoquark model [21] for the $\mathcal{B}(B_s^0 \to \tau^+\tau^-)$ and $\mathcal{B}(B \to K\tau^+\tau^-)$ observables. The $\Delta_\chi^2 \leq 2.30(1\sigma)$ and $\Delta_\chi^2 \leq 6.18(2\sigma)$ regions are shown in blue and light blue, respectively. The interval in green shows the current experimental measurements of ΔR_D with 1σ errors which is one of the observables entering the fit. The red bands show the 95% CL experimental exclusion limits from $B_s^0 \to \tau^+\tau^-$. The plot is taken from Ref. [21].

This would improve the current limit by a factor $\sim$35. Nonetheless, we note that the limit will still be more than 2 orders of magnitude above the SM prediction. In parallel, Belle II started to collect data in 2019 and should finish its data acquisition period by the end of 2027. We remind the reader that a brief overview of the Belle II experiment has been given in Sec. 4.3. Tab. 8.3 gives some estimates of the upper limits at 90% C.L on $\mathcal{B}(B_{(s)}^0 \to \tau^+\tau^-)$ that Belle II and LHCb could reach in the future.

Bibliography

[1] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579.

[2] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[3] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8** (1964) 214.

[4] S. Weinberg, *Nonabelian Gauge Theories of the Strong Interactions*, Phys. Rev. Lett. **31** (1973) 494.

[5] M. Y. Han and Y. Nambu, *Three Triplet Model with Double SU(3) Symmetry*, Phys. Rev. **139** (1965) B1006, [,187(1965)].

[6] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, Phys. Rev. Lett. **30** (1973) 1343, [,271(1973)].

[7] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett. **30** (1973) 1346, [,274(1973)].

[8] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, Phys. Lett. **47B** (1973) 365.

[9] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508, [,160(1964)].

[10] BaBar, J. P. Lees *et al.*, *Measurement of Branching Fractions and Rate Asymmetries in the Rare Decays $B \to K^{(*)} l^+ l^-$*, Phys. Rev. **D86** (2012) 032012, `arXiv:1204.3933`.

[11] Belle, J.-T. Wei *et al.*, *Measurement of the Differential Branching Fraction and Forward-Backword Asymmetry for $B \to K^{(*)} \ell^+ \ell^-$*, Phys. Rev. Lett. **103** (2009) 171801, `arXiv:0904.0770`.

[12] LHCb, R. Aaij *et al.*, *Test of lepton universality using $B^+ \to K^+ \ell^+ \ell^-$ decays*, Phys. Rev. Lett. **113** (2014) 151601, `arXiv:1406.6482`.

[13] LHCb, R. Aaij *et al.*, *Search for lepton-universality violation in $B^+ \to K^+ \ell^+ \ell^-$ decays*, Phys. Rev. Lett. **122** (2019), no. 19 191801, `arXiv:1903.09252`.

[14] Belle II, M. T. Prim, *Semileptonic and leptonic B decay results from early Belle II data*, in *17th Conference on Flavor Physics and CP Violation (FPCP 2019) Victoria, BC, Canada, May 6-10, 2019*, 2019. `arXiv:1906.09337`.

[15] LHCb, R. Aaij *et al.*, *Test of lepton universality with $B^0 \to K^{*0} \ell^+ \ell^-$ decays*, JHEP **08** (2017) 055, `arXiv:1705.05802`.

[16] BaBar, J. P. Lees *et al.*, *Measurement of an Excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ Decays and Implications for Charged Higgs Bosons*, Phys. Rev. **D88** (2013), no. 7 072012, `arXiv:1303.0571`.

[17] LHCb, R. Aaij *et al.*, *Measurement of the ratio of branching fractions $\mathcal{B}(\bar{B}^0 \to D^{*+}\tau^-\bar{\nu}_\tau)/\mathcal{B}(\bar{B}^0 \to D^{*+}\mu^-\bar{\nu}_\mu)$*, Phys. Rev. Lett. **115** (2015), no. 11 111803, `arXiv:1506.08614`, [Erratum: Phys. Rev. Lett.115,no.15,159901(2015)].

[18] Belle, S. Hirose *et al.*, *Measurement of the τ lepton polarization and $R(D^*)$ in the decay $\bar{B} \to D^*\tau^-\bar{\nu}_\tau$*, Phys. Rev. Lett. **118** (2017), no. 21 211801, `arXiv:1612.00529`.

[19] LHCb, R. Aaij *et al.*, *Test of Lepton Flavor Universality by the measurement of the $B^0 \to D^{*-}\tau^+\nu_\tau$ branching fraction using three-prong τ decays*, Phys. Rev. **D97** (2018), no. 7 072013, `arXiv:1711.02505`.

[20] Belle, A. Abdesselam *et al.*, *Measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with a semileptonic tagging method*, `arXiv:1904.08794`.

[21] C. Cornella, J. Fuentes-Martin, and G. Isidori, *Revisiting the vector leptoquark explanation of the B-physics anomalies*, `arXiv:1903.11517arXiv:1903.11517`.

[22] D. Bečirević, S. Fajfer, N. Košnik, and O. Sumensari, *Leptoquark model to explain the B-physics anomalies, R_K and R_D*, Phys. Rev. **D94** (2016), no. 11 115021, `arXiv:1608.08501`.

[23] A. Crivellin, G. D'Ambrosio, and J. Heeck, *Addressing the LHC flavor anomalies with horizontal gauge symmetries*, Phys. Rev. **D91** (2015), no. 7 075006, `arXiv:1503.03477`.

[24] J. M. Cline, *Scalar doublet models confront τ and b anomalies*, Phys. Rev. **D93** (2016), no. 7 075017, `arXiv:1512.02210`.

[25] LHCb, R. Aaij *et al.*, *Measurement of the $B_s^0 \to \mu^+\mu^-$ branching fraction and effective lifetime and search for $B^0 \to \mu^+\mu^-$ decays*, Phys. Rev. Lett. **118** (2017), no. 19 191801, `arXiv:1703.05747`.

[26] ATLAS, M. Aaboud *et al.*, *Study of the rare decays of B_s^0 and B^0 mesons into muon pairs using data collected during 2015 and 2016 with the ATLAS detector*, JHEP **04** (2019) 098, `arXiv:1812.03017`.

[27] CMS, S. Chatrchyan *et al.*, *Measurement of the $B_s^0 \to \mu^+\mu^-$ Branching Fraction and Search for $B^0 \to \mu^+\mu^-$ with the CMS Experiment*, Phys. Rev. Lett. **111** (2013) 101804, `arXiv:1307.5025`.

[28] C. Bobeth *et al.*, *$B_{s,d} \to l^+l^-$ in the Standard Model with Reduced Theoretical Uncertainty*, Phys. Rev. Lett. **112** (2014) 101801, `arXiv:1311.0903`.

[29] R. Alonso, B. Grinstein, and J. Martin Camalich, *Lepton universality violation and lepton flavor conservation in B-meson decays*, JHEP **10** (2015) 184, `arXiv:1505.05164`.

[30] D. Bečirević, N. Košnik, O. Sumensari, and R. Zukanovich Funchal, *Palatable Leptoquark Scenarios for Lepton Flavor Violation in Exclusive $b \to s\ell_1\ell_2$ modes*, JHEP **11** (2016) 035, `arXiv:1608.07583`.

[31] B. Capdevila *et al.*, *Searching for New Physics with $b \to s\tau^+\tau^-$ processes*, Phys. Rev. Lett. **120** (2018), no. 18 181802, `arXiv:1712.01919`.

[32] S. Fajfer, N. Kosnik, and S. Prelovsek, *Updated constraints on new physics in rare charm decays*, Phys. Rev. **D76** (2007) 074010, `arXiv:0706.1133`.

[33] G. Buchalla *et al.*, *B, D and K decays*, Eur. Phys. J. **C57** (2008) 309, `arXiv:0801.1833`.

[34] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. **D2** (1970) 1285.

[35] M. Beneke, T. Feldmann, and D. Seidel, *Systematic approach to exclusive $B \to Vl^+l^-$, $V\gamma$ decays*, Nucl. Phys. **B612** (2001) 25, `arXiv:hep-ph/0106067`.

[36] T. Feldmann, B. Müller, and D. Seidel, *$D \to \rho\,\ell^+\ell^-$ decays in the QCD factorization approach*, JHEP **08** (2017) 105, `arXiv:1705.05891`.

[37] S. de Boer and G. Hiller, *Flavor and new physics opportunities with rare charm decays into leptons*, Phys. Rev. **D93** (2016), no. 7 074001, `arXiv:1510.00311`.

[38] S. Fajfer and N. Košnik, *Prospects of discovering new physics in rare charm decays*, Eur. Phys. J. **C75** (2015), no. 12 567, `arXiv:1510.00965`.

[39] M. A. Shifman, *Quark hadron duality*, in *At the frontier of particle physics. Handbook of QCD. Vol. 1-3*, (Singapore), pp. 1447–1494, World Scientific, World Scientific, 2001. `arXiv:hep-ph/0009131`. [3,1447(2000)].

[40] J. Lyon and R. Zwicky, *Resonances gone topsy turvy - the charm of QCD or new physics in $b \to s\ell^+\ell^-$?*, `arXiv:1406.0566`.

[41] LHCb, R. Aaij *et al.*, *Search for the rare decay $D^0 \to \mu^+\mu^-$*, Phys. Lett. **B725** (2013) 15, `arXiv:1305.5059`.

[42] LHCb, R. Aaij *et al.*, *Search for $D^+_{(s)} \to \pi^+\mu^+\mu^-$ and $D^+_{(s)} \to \pi^-\mu^+\mu^+$ decays*, Phys. Lett. **B724** (2013) 203, `arXiv:1304.6365`.

[43] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[44] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[45] LHCb Collaboration, R. Aaij *et al.*, *Search for the decays $B^0_s \to \tau^+\tau^-$ and $B^0 \to \tau^+\tau^-$*, Phys. Rev. Lett. **118** (2017) 251802. 10 p.

[46] LHCb collaboration, R. Aaij *et al.*, *Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions*, Phys. Rev. Lett. **118** (2017) 052002, Erratum ibid. **119** (2017) 169901, `arXiv:1612.05140`.

[47] Particle Data Group, K. A. Olive *et al.*, *Review of particle physics*, Chin. Phys. **C38** (2014) 090001, and 2015 update.

[48] A. Hoecker *et al.*, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, `arXiv:physics/0703039`.

[49] ROOT Collaboration, K. Cranmer *et al.*, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, Tech. Rep. CERN-OPEN-2012-016, New York U., New York, Jan, 2012.

[50] Y. Xie, *sFit: concept, implementation and possible pitfalls*, Talk at the 50th Analysis and Software Week [Indico 201535] January 30th, 2013.

[51] C. Meaux *et al.*, *Search for $B^0_{(s)} \to \tau^+\tau^-$ in the $(3\pi, \mu)$ final state*, Tech. Rep. LHCb-INT-2018-021. CERN-LHCb-INT-2018-021, CERN, Geneva, Jun, 2018.

[52] G. Isidori, Y. Nir, and G. Perez, *Flavor Physics Constraints for Physics Beyond the Standard Model*, Ann. Rev. Nucl. Part. Sci. **60** (2010) 355, `arXiv:1002.0900`.

[53] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

[54] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[55] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π Decay of the K^0_2 Meson*, Phys. Rev. Lett. **13** (1964) 138.

[56] BaBar, B. Aubert *et al.*, *Observation of CP violation in the B^0 meson system*, Phys. Rev. Lett. **87** (2001) 091801, `arXiv:hep-ex/0107013`.

[57] Belle, K. Abe *et al.*, *Observation of large CP violation in the neutral B meson system*, Phys. Rev. Lett. **87** (2001) 091802, `arXiv:hep-ex/0107061`.

[58] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[59] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*, Addison-Wesley, Reading, USA, 1995.

[60] C. G. Bollini and J. J. Giambiagi, *Dimensional Renormalization: The Number of Dimensions as a Regularizing Parameter*, Nuovo Cim. **B12** (1972) 20.

[61] M. S. Chanowitz, M. Furman, and I. Hinchliffe, *The Axial Current in Dimensional Regularization*, Nucl. Phys. **B159** (1979) 225.

[62] G. 't Hooft, *Dimensional regularization and the renormalization group*, Nucl. Phys. **B61** (1973) 455.

[63] A. J. Buras, *Weak Hamiltonian, CP violation and rare decays*, in *Probing the standard model of particle interactions. Proceedings, Summer School in Theoretical Physics, NATO Advanced Study Institute, 68th session, Les Houches, France, July 28-September 5, 1997. Pt. 1, 2*, pp. 281–539, 1998. `arXiv:hep-ph/9806471`.

[64] S. Fajfer and N. Košnik, *Prospects of discovering new physics in rare charm decays*, Eur. Phys. J. **C75** (2015), no. 12 567, `arXiv:1510.00965`.

[65] S. de Boer and G. Hiller, *Flavor and new physics opportunities with rare charm decays into leptons*, Phys. Rev. **D93** (2016), no. 7 074001, `arXiv:1510.00311`.

[66] S. Aoki *et al.*, *Review of lattice results concerning low-energy particle physics*, Eur. Phys. J. **C74** (2014) 2890, `arXiv:1310.8555`.

[67] A. Khodjamirian and R. Ruckl, *QCD sum rules for exclusive decays of heavy mesons*, Adv. Ser. Direct. High Energy Phys. **15** (1998) 345, `arXiv:hep-ph/9801443`, [,345(1998)].

[68] V. M. Braun, *QCD sum rules for heavy flavors*, `arXiv:hep-ph/9911206arXiv:hep-ph/9911206`, [PoShf8,006(1999)].

[69] P. Colangelo and A. Khodjamirian, *QCD sum rules, a modern perspective*, `arXiv:hep-ph/0010175arXiv:hep-ph/0010175`.

[70] W. Altmannshofer *et al.*, *Symmetries and Asymmetries of $B \to K^* \mu^+ \mu^-$ Decays in the Standard Model and Beyond*, JHEP **01** (2009) 019, `arXiv:0811.1214`.

[71] Heavy Flavor Averaging Group (HFAG), Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2014*, `arXiv:1412.7515`.

[72] CDF, T. Aaltonen *et al.*, *Search for the Decays $B_s^0 \to e^+ \mu^-$ and $B_s^0 \to e^+ e^-$ in CDF Run II*, Phys. Rev. Lett. **102** (2009) 201801, `arXiv:0901.3803`.

[73] Belle, M. Petric *et al.*, *Search for leptonic decays of D^0 mesons*, Phys. Rev. **D81** (2010) 091102, `arXiv:1003.2345`.

[74] 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`.

[75] M. Beneke, T. Feldmann, and D. Seidel, *Exclusive radiative and electroweak $b \to d$ and $b \to s$ penguin decays at NLO*, Eur. Phys. J. **C41** (2005) 173, `arXiv:hep-ph/0412400`.

[76] V. M. Braun and I. E. Filyanov, *QCD Sum Rules in Exclusive Kinematics and Pion Wave Function*, Z. Phys. **C44** (1989) 157, [Yad. Fiz.50,818(1989)].

[77] BaBar, J. P. Lees *et al.*, *Searches for Rare or Forbidden Semileptonic Charm Decays*, Phys. Rev. **D84** (2011) 072006, `arXiv:1107.4465`.

[78] M. Beneke, T. Feldmann, and D. Seidel, *Systematic approach to exclusive $B \to V l^+ l^-$, $V\gamma$ decays*, Nucl. Phys. **B612** (2001) 25, `arXiv:hep-ph/0106067`.

[79] M. Beneke and T. Feldmann, *Symmetry breaking corrections to heavy to light B meson form-factors at large recoil*, Nucl. Phys. **B592** (2001) 3, `arXiv:hep-ph/0008255`.

[80] H. H. Asatrian, H. M. Asatrian, C. Greub, and M. Walker, *Two loop virtual corrections to $B \to X_s l^+ l^-$ in the standard model*, Phys. Lett. **B507** (2001) 162, `arXiv:hep-ph/0103087`.

[81] D. Seidel, *Analytic two loop virtual corrections to b —> d l+ l-*, Phys. Rev. **D70** (2004) 094038, `arXiv:hep-ph/0403185`.

[82] C. Greub, V. Pilipp, and C. Schupbach, *Analytic calculation of two-loop QCD corrections to $b \to s l^+ l^-$ in the high q^2 region*, JHEP **12** (2008) 040, `arXiv:0810.4077`.

[83] S. de Boer, B. Müller, and D. Seidel, *Higher-order Wilson coefficients for $c \to u$ transitions in the standard model*, JHEP **08** (2016) 091, `arXiv:1606.05521`.

[84] M. Gorbahn and U. Haisch, *Effective Hamiltonian for non-leptonic $|\Delta F| = 1$ decays at NNLO in QCD*, Nucl. Phys. **B713** (2005) 291, `arXiv:hep-ph/0411071`.

[85] M. Gorbahn, U. Haisch, and M. Misiak, *Three-loop mixing of dipole operators*, Phys. Rev. Lett. **95** (2005) 102004, `arXiv:hep-ph/0504194`.

[86] OPAL, K. Ackerstaff *et al.*, *Measurement of the strong coupling constant alpha(s) and the vector and axial vector spectral functions in hadronic tau decays*, Eur. Phys. J. **C7** (1999) 571, `arXiv:hep-ex/9808019`.

[87] M. Davier *et al.*, *Update of the ALEPH non-strange spectral functions from hadronic τ decays*, Eur. Phys. J. **C74** (2014), no. 3 2803, `arXiv:1312.1501`.

[88] D. Boito *et al.*, *Strong coupling from $e^+e^- \to$ hadrons below charm*, Phys. Rev. **D98** (2018), no. 7 074030, `arXiv:1805.08176`.

[89] J. T. Daub, C. Hanhart, and B. Kubis, *A model-independent analysis of final-state interactions in $\overline{B}^0_{d/s} \to J/\psi\pi\pi$*, JHEP **02** (2016) 009, `arXiv:1508.06841`.

[90] A. Pich and A. Rodríguez-Sánchez, *Updated determination of $\alpha_s(m_\tau^2)$ from tau decays*, Mod. Phys. Lett. **A31** (2016), no. 30 1630032, `arXiv:1606.07764`.

[91] Particle Data Group, M. Tanabashi *et al.*, *Review of Particle Physics*, Phys. Rev. **D98** (2018), no. 3 030001.

[92] D. Boito *et al.*, *Strong coupling from the revised ALEPH data for hadronic τ decays*, Phys. Rev. **D91** (2015), no. 3 034003, `arXiv:1410.3528`.

[93] BES, M. Ablikim *et al.*, *R value measurements for e+ e- annihilation at 2.60-GeV, 3.07-GeV and 3.65-GeV*, Phys. Lett. **B677** (2009) 239, `arXiv:0903.0900`.

[94] V. V. Anashin *et al.*, *Measurement of R_{uds} and R between 3.12 and 3.72 GeV at the KEDR detector*, Phys. Lett. **B753** (2016) 533, `arXiv:1510.02667`.

[95] V. V. Anashin *et al.*, *Measurement of R between 1.84 and 3.05 GeV at the KEDR detector*, Phys. Lett. **B770** (2017) 174, `arXiv:1610.02827`.

[96] ETM, V. Lubicz *et al.*, *Scalar and vector form factors of $D \to \pi(K)\ell\nu$ decays with $N_f = 2+1+1$ twisted fermions*, Phys. Rev. **D96** (2017), no. 5 054514, `arXiv:1706.03017`.

[97] ETM, V. Lubicz *et al.*, *Tensor form factor of $D \to \pi(K)\ell\nu$ and $D \to \pi(K)\ell\ell$ decays with $N_f = 2+1+1$ twisted-mass fermions*, Phys. Rev. **D98** (2018), no. 1 014516, `arXiv:1803.04807`.

[98] A. Khodjamirian, C. Klein, T. Mannel, and N. Offen, *Semileptonic charm decays D —> pi l nu(l) and D —> K l nu(l) from QCD Light-Cone Sum Rules*, Phys. Rev. **D80** (2009) 114005, `arXiv:0907.2842`.

[99] A. Bharucha, *Two-loop Corrections to the Btoπ Form Factor from QCD Sum Rules on the Light-Cone and $|V_{ub}|$*, JHEP **05** (2012) 092, `arXiv:1203.1359`.

[100] P. Ball and R. Zwicky, *New results on $B \to \pi, K, \eta$ decay formfactors from light-cone sum rules*, Phys. Rev. **D71** (2005) 014015, `arXiv:hep-ph/0406232`.

[101] A. G. Grozin and M. Neubert, *Asymptotics of heavy meson form-factors*, Phys. Rev. **D55** (1997) 272, arXiv:hep-ph/9607366.

[102] C. Bobeth, G. Hiller, and G. Piranishvili, *Angular distributions of $\bar{B} \to \bar{K}\ell^+\ell^-$ decays*, JHEP **12** (2007) 040, arXiv:0709.4174.

[103] R.-M. Wang *et al.*, *Decays $D^+_{(s)} \to \pi(K)^+\ell^+\ell^-$ and $D^0 \to \ell^+\ell^-$ in the MSSM with and without R-parity*, Int. J. Mod. Phys. **A30** (2015), no. 12 1550063, arXiv:1409.0181.

[104] S. Fajfer, S. Prelovsek, and P. Singer, *Rare charm meson decays D —> P lepton+ lepton- and c —> u lepton+ lepton- in SM and MSSM*, Phys. Rev. **D64** (2001) 114009, arXiv:hep-ph/0106333.

[105] 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.

[106] S. Fajfer and S. Prelovsek, *Effects of littlest Higgs model in rare D meson decays*, Phys. Rev. **D73** (2006) 054026, arXiv:hep-ph/0511048.

[107] W. Herr and B. Muratori, *Concept of luminosity*, in *Intermediate accelerator physics. Proceedings, CERN Accelerator School, Zeuthen, Germany, September 15-26, 2003*, pp. 361–377, 2003.

[108] E. Norrbin and T. Sjostrand, *Production and hadronization of heavy quarks*, Eur. Phys. J. **C17** (2000) 137, arXiv:hep-ph/0005110.

[109] F. Follin and D. Jacquet, *Implementation and experience with luminosity levelling with offset beam*, in *Proceedings, ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders (BB2013): CERN, Geneva, Switzerland, March 18-22 2013*, pp. 183–187, 2014. arXiv:1410.3667. [183(2014)], doi: 10.5170/CERN-2014-004.183.

[110] LHCb, R. Aaij *et al.*, *LHCb Detector Performance*, Int. J. Mod. Phys. **A30** (2015), no. 07 1530022, arXiv:1412.6352.

[111] D. Francesca, *LHCb detector and trigger performance in Run2*, EPJ Web Conf. **164** (2017) 01016.

[112] LHCb Collaboration, S. Amato *et al.*, *LHCb magnet: Technical Design Report*, Technical Design Report LHCb, CERN, Geneva, 2000.

[113] L. Collaboration, *LHCb VELO Upgrade Technical Design Report*, Tech. Rep. CERN-LHCC-2013-021. LHCB-TDR-013, Nov, 2013.

[114] L. Collaboration, *LHCb Tracker Upgrade Technical Design Report*, Tech. Rep. CERN-LHCC-2014-001. LHCB-TDR-015, Feb, 2014.

[115] M. Needham, *Performance of the LHCb Track Reconstruction Software*, Tech. Rep. LHCb-2007-144. CERN-LHCb-2007-144, CERN, Geneva, Jan, 2008.

[116] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055.

[117] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, `arXiv:1210.6861`.

[118] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006) 026, `arXiv:hep-ph/0603175`.

[119] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[120] Z. Was, *TAUOLA for simulation of tau decay and production: perspectives for precision low energy and LHC applications*, Nucl. Phys. Proc. Suppl. **218** (2011) 249, `arXiv:1101.1652`.

[121] M. G. Pia, T. Basaglia, Z. W. Bell, and P. V. Dressendorfer, *Geant4 in Scientific Literature*, in *Proceedings, 2009 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC 2009): Orlando, Florida, October 25-31, 2009*, pp. 189–194, 2009. `arXiv:0912.0360`. doi: 10.1109/NSSMIC.2009.5401810.

[122] Belle, D. Liventsev, *Recent results from the Belle experiment*, EPJ Web Conf. **125** (2016) 01008, `arXiv:1608.04500`.

[123] Belle-II, B. Wang, *The Belle II Experiment and SuperKEKB Upgrade*, J. Univ. Sci. Tech. China **46** (2016), no. 7 617, `arXiv:1511.09434`.

[124] BaBar, B. Aubert *et al.*, *The BABAR Detector: Upgrades, Operation and Performance*, Nucl. Instrum. Meth. **A729** (2013) 615, `arXiv:1305.3560`.

[125] "Abazov and others", *The upgraded D0 detector*, Nuclear Instruments and Methods in Physics Research A **565** (2006) 463, `arXiv:physics/0507191`.

[126] CDF, A. Meyer, *The CDF experiment at the Tevatron: The First two year of Run II*, Mod. Phys. Lett. **A18** (2003) 1643, `arXiv:hep-ex/0306046`.

[127] LHCb Collaboration, *Expression of Interest for a Phase-II LHCb Upgrade: Opportunities in flavour physics, and beyond, in the HL-LHC era*, Tech. Rep. CERN-LHCC-2017-003, CERN, Geneva, Feb, 2017.

[128] Particle Data Group, C. Patrignani *et al.*, *Review of Particle Physics*, Chin. Phys. **C40** (2016), no. 10 100001.

[129] BaBar collaboration, B. Aubert *et al.*, *A search for the rare decay $B^0 \to \tau^+\tau^-$ at BaBar*, Phys. Rev. Lett. **96** (2006) 241802, `arXiv:hep-ex/0511015`.

[130] Y. Grossman, Z. Ligeti, and E. Nardi, *$b \to \tau^+\tau^-(x)$ decays: First constraints and phenomenological implications*, Phys. Rev. D **55** (1997) 2768.

[131] A. Dighe, A. Kundu, and S. Nandi, *Enhanced $B_s - -\bar{b}_s$ lifetime difference and anomalous like-sign dimuon charge asymmetry from new physics in $B_s \to \tau^+\tau^-$*, Phys. Rev. D **82** (2010) 031502.

[132] C. Bobeth and U. Haisch, *New Physics in Γ_{12}^s: $(\bar{s}b)(\bar{\tau}\tau)$ Operators*, Acta Phys. Polon. **B44** (2013) 127, `arXiv:1109.1826`.

[133] M. Feindt and U. Kerzel, *The NeuroBayes neural network package*, Nucl. Instrum. Meth. **A559** (2006) 190.

[134] L. Moneta *et al.*, *The RooStats project*, PoS **ACAT2010** (2010) 057, arXiv:1009.1003.

[135] LHCb collaboration, R. Aaij *et al.*, *First observations of $\overline{B}^0_s \to D^+D^-$, $D^+_sD^-$ and $D^0\overline{D}^0$ decays*, Phys. Rev. **D87** (2013) 092007, arXiv:1302.5854.

[136] LHCb collaboration, S. Blusk, V. Gligorov, and M. Williams, *First observation and branching fraction measurement of $\overline{B}^0_s$ to double-charm final states*, LHCb-ANA-2012-100, 2012.

[137] LHCb collaboration, R. Aaij *et al.*, *Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on $\mathcal{B}$ meson kinematics*, JHEP **04** (2013) 001, arXiv:1301.5286; LHCb collaboration, *Updated average f_s/f_d b-hadron production fraction ratio for 7 TeV pp collisions*, LHCb-CONF-2013-011.

[138] A. L. Read, *Presentation of search results: The CL(s) technique*, J. Phys. **G28** (2002) 2693, [11(2002)].

[139] A. Mordá, *Rare dileptonic $B^0_{(s)}$ meson decays at LHCb, PhD Thesis*, Aix-Marseille Université (2015), CERN-THESIS-2015-264.

[140] SLD collaboration, E. Etzion, *Measurement of R_b at SLD*, arXiv:hep-ex/9606008.

[141] I. M. Nugent *et al.*, *Resonance chiral Lagrangian currents and experimental data for $\tau^- \to \pi^-\pi^-\pi^+\nu_\tau$*, Phys. Rev. **D88** (2013) 093012, arXiv:1310.1053.

[142] BaBar, I. M. Nugent, *Invariant mass spectra of $\tau^- \to h^-h^-h^+\nu_\tau$ decays*, Nucl. Phys. Proc. Suppl. **253-255** (2014) 38, arXiv:1301.7105.

[143] N. Davidson *et al.*, *Universal Interface of TAUOLA Technical and Physics Documentation*, Comput. Phys. Commun. **183** (2012) 821, arXiv:1002.0543.

[144] R. J. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Comput. Phys. Commun. **77** (1993) 219.

[145] LHCb, R. Aaij *et al.*, *Search for the lepton-flavour-violating decays $B^0_s \to \tau^\pm\mu^\mp$ and $B^0 \to \tau^\pm\mu^\mp$*, arXiv:1905.06614.

[146] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986, DESY-F31-86-02.

[147] D. M. Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nucl. Instrum. Meth. **A764** (2014) 150, arXiv:1312.5000.

[148] Belle collaboration, A. Zupanc *et al.*, *Improved measurement of $\overline{B}^0 \to D^-_sD^+$ and search for $\overline{B}^0 \to D^+_sD^-_s$ at Belle*, Phys. Rev. **D75** (2007) 091102, arXiv:hep-ex/0703040.

[149] LHCb, R. Aaij *et al.*, *Physics case for an LHCb Upgrade II - Opportunities in flavour physics, and beyond, in the HL-LHC era*, arXiv:1808.08865.

[150] Belle-II, W. Altmannshofer *et al.*, *The Belle II Physics Book*, arXiv:1808.10567.