Use of $B \to h^+h'^-$ as control channels for the measurement of $\mathbf{BR}(B_s^0 \to \mu^+\mu^-)$ at LHCb

Memòria de la tesi presentada per

Elías López Asamar

per optar al títol de Doctor en Física

Directors de tesi:

Lluís Garrido Beltran, Hugo Ruiz Pérez

Programa de doctorat en Física

Departament d'Estructura i Constituents de la Matèria

Barcelona, octubre de 2011

Agraïments

Primer de tot vull agraïr als meus directors, Lluís Garrido i Hugo Ruiz, per tot el temps i tota l'atenció que m'han dedicat durant el doctorat. I tot seguit vull donar les gràcies a diverses persones vinculades al grup d'LHCb a Barcelona, no tan sols per haver compartit molts moments durant aquests últims anys, sinó també per l'ajut ofert quan ha estat necessari: a Eugeni Graugés i Ricardo Graciani; a Jordi Garra, Antonio Pérez-Calero i Cédric Potterat; a Alessandro Camboni, Marc Grabalosa, Albert Puig, Ricardo Vázquez i Vicente Rives; a Adrià Cassajús, Albert Comerma, David Gascón i Eduard Picatoste; a Xavier Vilassís i Míriam Calvo; i a Carlos Conzález.

També vull incloure en aquests agraïments a diversos membres i ex-membres del Departament d'Estructura i Constituents de la Matèria i el Departament de Física Fonamental de la Universitat de Barcelona: a Miguel Ángel Escobedo, Clemens Bauer i Aldo Déctor; a Domènec Espriu, Joan Soto i Federico Mescia; i a Jorge Mondéjar, María José Rodríguez i Raúl Sánchez.

En el context del grup de treball de $B_s^0 \to \mu^+\mu^-$ vull agraïr a varis membres d'LHCb per tot el suport que m'han donat: a Frederic Teubert; a Jose Ángel Hernando i Diego Martínez; i a Gaia Lanfranchi.

Entrant a un nivell més personal vull esmentar al meu amic Adrià Parejo, i a diversos ex-companys del Colegio Mayor Loyola de Madrid: Alfredo Álvarez, Rodrigo Bravo, Carlos García, Daniel García, Gregorio Munoz, Jesús Ortega, Miguel Ángel Quintero, Ignacio de los Reyes, José María Rodriguez, Lucas Sánchez i Francisco Sánchez.

Finalment vull donar les gràcies a la meva família per tot el que m'han donat: al meu germà Abraham, i als meus pares Pedro i Mercè; i als meus dos estimats estels Susana i Andrés.

Resum

Prediccions per al ritme de desintegració de $B_s^0 \to \mu^+\mu^-$

Actualment el Model Estàndar (SM) ofereix la descripció més fonamental de la matèria i de les seves interaccions. La seva validesa es recolza en un gran nombre de mesures experimentals que han permès verificar la seva consistència amb una extraordinària precisió. No obstant, alguns fets com el problema de la jerarquia de masses, l'evidència que la massa dels neutrins no és nul·la o certes mesures en l'àmbit de la cosmologia posen en dubte que el SM sigui la teoria última i més fonamental. Per aquest motiu s'han desenvolupat diverses teories que superen el SM per a intentar oferir una explicació natural a aquests fets, prenent com a base hipòtesis com supersimetria, technicolor o dimensions addicionals.

La fenomenologia dels models que amplien el SM ha estat estudiada en detall per tal de trobar mesures que permetin validar-los. En particular, s'ha trobat que millorant la precisió actual en la mesura de certs observables referents a desintegracions de mesons B es poden detectar discrepàncies sensibles respecte a les prediccions del SM. Alguns casos remarcables són la presència de violació de la simetria CP en $B_s^0 \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$, les distribucions angulars dels productes de $B^0 \to K^{*0}\mu^+\mu^-$ i el ritme de desintegració (BR) de $B_s^0 \to \mu^+\mu^-$. Abans de la posada en funcionament de l'LHC, aquest últim observable estava restringit per Tevatron a valors per sota de 2.4×10^{-8}, amb un nivell de confiança del 95%.

En el SM, el procés $B_s^0 \to \mu^+\mu^-$ està fortament suprimit degut a que l'acoblament feble restringeix l'espai de fase de l'estat final. En aquests context, el seu BR es pot predir amb una baixa incertesa amb l'ajut de la mesura actual de la diferència de massa dels dos autoestats d'evolució del mesó B_s^0, obtenint un valor igual a $(3.2\pm0.2)\times10^{-9}$. D'altra banda, diferents models més enllà del SM basats en supersimetria o en dimensions adicionals prediuen importants increments d'aquest valor que poden arribar fins a l'actual fita experimental en algunes regions permeses del seu espai de paràmetres.

L'experiment LHCb

El detector LHCb va ser concebut per a estudiar la física de les desintegracions de mesons B produïts en les interaccions protó-protó a 14 TeV a l'LHC. L'acceptància geomètrica d'LHCb, motivada per la distribució angular dels mesons B produïts sota

aquestes condicions, comprèn únicament direccions properes a l'eix d'incidència, i es restringeix a un sol costat de la regió d'interacció. Es preveu que dins d'aquesta acceptància es produeixin aproximadament 10^{12} parells $b\bar{b}$ per any nominal (2 fb^{-1}).

Els subdetectors que constitueixen LHCb , ordenats segons la distància a la regió d'interacció, són els següents: un detector de traces d'alta precisió anomenat *Vertex Locator* (VeLo), destinat a la reconstrucció de vèrtexs de desintegració de mesons B; dos detectors d'anells de radiació Cerenkov (RICH), dissenyats per a discriminar entre diferents tipus de partícules, principalment pions i kaons; un sistema de detectors de traces format pel *Tracker Turicensis* (TT), l'*Inner Tracker* (IT) i l'*Outer Tracker* (OT), per a mesurar la trajectòria de partícules carregades sota l'efecte d'un camp magnètic; un sistema de calorímetres format per l'*Scintillator Pad Detector* (SPD), el *Pre-Shower* (PS), un calorímetre electromagnètic (ECAL) i un calorímetre hadrònic (HCAL), usats en els primers nivells del trigger i en la detecció de partícules neutres; i un sistema de detectors de traces situat més enllà dels calorímetres anomenat *Muon System*, necessari per a la identificació de muons.

El trigger d'LHCb redueix el ritme d'esdeveniments dels 10 MHz produïts en les col·lisions de l'LHC fins als 2 kHz permesos per la capacitat d'emmagatzematge, i es separa en dos nivells degut a les limitacions de còmput. El primer nivell, anomenat *Level-0* (L0), rebaixa el ritme d'esdeveniments fins a 1 MHz mitjançant dispositius electrònics adaptats operant a la sortida de dades del VeLo, dels calorímetres i del sistema de muons. El segon nivell, anomenat *High Level Trigger* (HLT) aconsegueix la reducció restant fent ús d'algoritmes operant en una granja de computació.

La presa de dades i el control d'LHCb s'efectua a través de tres sistemes: el sistema d'adquisició de dades (DAQ), constituït pels dispositius de lectura i transport de dades, i pels mitjans de computació usats pels algoritmes de l'HLT (*Event Filter Farm*, EFF); el sistema de *timing* i control ràpid (TFC), que controla el flux de dades entre el detector i la EFF; i el sistema de control de l'experiment (ECS), que permet fer el seguiment i manipular els paràmetres del detector, del DAQ i del TFC.

En el procés de reconstrucció dels esdeveniments emmagatzemats es busquen els vèrtexs primaris amb la informació del VeLo, i es formen les traces de les partícules carregades amb la informació del VeLo, el TT, l'IT i l'OT. A més, a cada traça s'associa informació per a la seva identificació, proveïda pel RICH, els calorímetres i el sistema de muons. Finalment, aquests esdeveniments són classificats segons les seves propietats (*stripping*), per tal de facilitar la seva anàlisi posterior.

La mesura de BR($B_s^0 \to \mu^+\mu^-$) a LHCb

El procediment per a mesurar BR($B_s^0 \to \mu^+\mu^-$) a LHCb es basa en l'ús de variables discriminants per a assignar a cada esdeveniment una mesura de la versemblança de ser senyal. Per aquest motiu la selecció d'esdeveniments mitjançant talls queda reduïda a uns pocs talls poc restrictius aplicats al primer nivell de l'anàlisi.

La selecció inicial d'esdeveniments requereix la presència de dues traces reconstruïdes i identificades com a muons. L'estat final de $B_s^0 \to \mu^+\mu^-$ es reconstrueix

amb una eficiència igual a 0.682±0.003, mentre que aquest estat final reconstruït s'identifica com a una parella de muons amb una eficiència igual a 0.838±0.003. Els talls de la selecció es basen en la qualitat del vèrtex format pel dos muons, la seva massa invariant i la compatibilitat entre la trajectòria del mesó B reconstruït i alguna interacció primària. No s'inclouen talls que puguin suprimir parts de l'espai de fases dels productes de desintegració, com els que involucren el moment transvers o el paràmetre d'impacte d'aquestes partícules. La eficiència d'aquesta selecció sobre el senyal reconstruït és de 0.922±0.002. El trigger és considerat com un efecte residual sobre la mostra d'esdeveniments seleccionats. En aquest sentit, l'impacte del trigger sobre el senyal és baix degut a l'alta eficiència dels talls orientats a canals amb muons, que en aquest cas porten a una eficiència total igual a 0.925±0.001 sobre el senyal seleccionat. Considerant els efectes de reconstrucció, identificació de muons, selecció i trigger, la quantitat total de senyal esperat per a 2 fb^{-1} és igual a 53 esdeveniments.

S'ha comprovat que el fons prové principalment de la combinació de dos muons procedents de la desintegració de diferents mesons B, esperant uns $2{\times}10^5$ esdeveniments d'aquest tipus en 2 fb^{-1}. Altres fonts de fons considerades són la combinació de dos muons procedents de $B_c^+ \to J/\psi(\mu^+\mu^-)\mu^+\nu_\mu$ i la identificació errònia dels dos hadrons procedents de $B \to h^+h'^-$ com muons, que suposarien 70 i 17 esdeveniments respectivament en 2 fb^{-1}.

Les variables discriminants que es fan servir per a extreure BR($B_s^0 \to \mu^+\mu^-$) a partir dels esdeveniments seleccionats són la massa invariant (IM) del parell de muons, i una variable anomenada *versemblança geomètrica* (GL). Aquesta última variable es construeix a partir d'altres variables relacionades amb la geometria de l'esdeveniment, que es combina mitjançant tècniques d'anàlisi multivariant. La GL no inclou explícitament informació relativa al mòdul del moment dels productes de desintegració, per tal d'evitar correlacions entre aquesta variable i la IM. Aquest fet permet la calibració independent de cada una d'aquestes variables discriminants. Per raons pràctiques, els valors permesos de la GL estan compresos entre 0 i 1. A més, la GL es defineix de manera que la seva distribució sigui plana per al senyal i molt propera a l'origen per al fons. Per aquest motiu, l'intèrval comprès entre 0.5 i 1 es coneix com la *regió sensible*.

Per a tal d'extreure BR($B_s^0 \to \mu^+\mu^-$) els esdeveniments seleccionats es classifiquen en diferents categories segons els seus valors de la IM i la GL. Els continguts observats es comparen amb els respectius continguts esperats de senyal i fons, mitjançant un mètode de base freqüentista anomenat CL_s.

El rendiment del mètode proposat per a mesurar BR($B_s^0 \to \mu^+\mu^-$) a LHCb es mostra a la Fig. 1 en funció de la lluminositat integrada. Segons aquests resultats, LHCb té el potencial d'excloure valors de BR($B_s^0 \to \mu^+\mu^-$) fins al valor predit pel SM amb menys de 2 fb^{-1}. A més, l'observació d'aquest valor a 3σ i 5σ es preveu amb 2 i 10 fb^{-1} respectivament.

Els continguts esperats de senyal usats pel mètode CL_s es determinen amb canals de control per tal d'evitar els efectes sistemàtics derivats de l'ús de simulacions. Dos canals han estat proposats per a dur a terme aquesta tasca: $B^+ \to J/\psi(\mu^+\mu^-)K^+$, que presenta un parell de muons en l'estat final, que permet cancel·lar parcialment

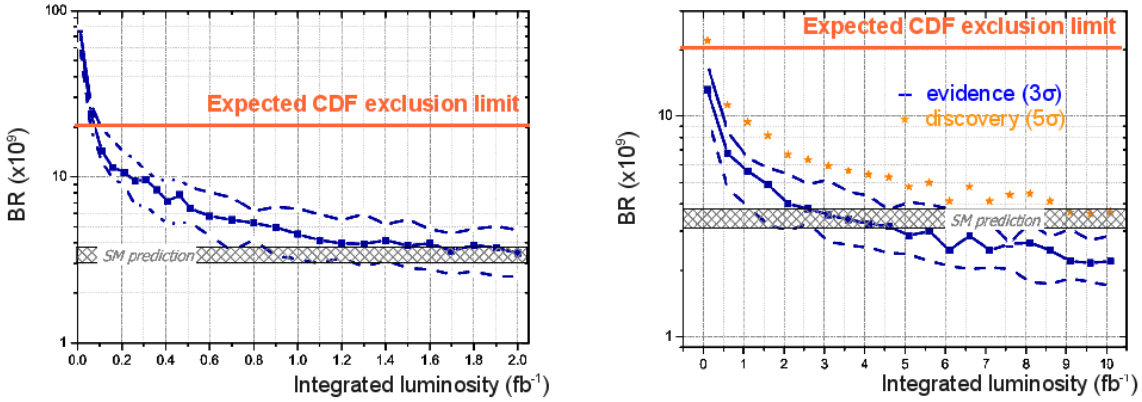

Figure 1: Límit d'exclusió a 90% CL (esquerra), i observació a 3σ i 5σ (dreta), en funció de la lluminositat integrada $\mathcal{L}$.

efectes del trigger i de la identificació de muons; i $B \to h^+h'^-$, on h representa un pió o kaó carregat, que té les mateixes propietats geomètriques i cinemàtiques que el senyal, i això permet calibrar les dues variables discriminants i cancel·lar efectes de selecció.

L'ús de $B \to h^+h'^-$ com a canals de control per a la mesura de BR($B_s^0 \to \mu^+\mu^-$)

Les desintegracions de tipus $B \to h^+h'^-$ tenen un BR de l'ordre de 10^{-6}, com es pot veure a la Taula 1. El procés més abundant és $B^0 \to K^+\pi^-$, seguit de $B_s^0 \to K^+K^-$ i de $B^0 \to \pi^+\pi^-$. En total, la producció esperada de $B \to h^+h'^-$ dins l'acceptància d'LHCb és d'uns 14 milions de desintegracions per a 2 fb^{-1}. D'altra banda, mentre el BR dels canals $B^0 \to K^+\pi^-$ i $B^0 \to \pi^+\pi^-$ s'ha determinat amb alta precisió a les B *factories*, la mesura del BR dels canals $B_s^0 \to K^+K^-$ i $B_s^0 \to \pi^+K^-$ es veu afectada per la incertesa del 13% en la mesura de la fracció de producció del mesó B_s^0.

Procés	BR ($\times 10^5$)	$2\sigma_{b\bar{b}} f_{\rm B}$BR (nb)
$B^0 \to \pi^+\pi^-$	0.513 ± 0.024	1.05 ± 0.05
$B^0 \to K^+\pi^-$	1.94 ± 0.06	4.0 ± 0.1
$B_d \to KK$	<0.12, 90% CL	<0.06, 90% CL
$B_s \to \pi\pi$	<0.041, 90% CL	<0.08, 90% CL
$B_s^0 \to \pi^+K^-$	0.49 ± 0.10	0.24 ± 0.05
$B_s^0 \to K^+K^-$	3.3 ± 0.9	1.6 ± 0.4

Table 1: Mesures actuals del BR i ritmes de producció esperats per a $B \to h^+h'^-$.

Les diferències en massa entre els mesons B^0 i B_s^0 i entre els muons, pions i kaons són despreciables respecte al moment dels productes de desintegració de $B_s^0 \to \mu^+\mu^-$ i $B \to h^+h'^-$. Aquest fet justifica que, per a molts propòsits, les propietats cinemàtiques de $B \to h^+h'^-$ a nivell de producció es puguin considerar iguals que les de $B_s^0 \to \mu^+\mu^-$.

El fet de que una part important dels hadrons és incapaç de travessar tot el sistema de detecció de traces degut a la interacció amb el detector introdueix diferències entre $B_s^0 \to \mu^+\mu^-$ i $B \to h^+h'^-$ a nivell de reconstrucció. L'efecte d'aquestes diferències és pràcticament independent del moment dels productes de desintegració, de manera que es pot considerar que les propietats cinemàtiques de $B_s^0 \to \mu^+\mu^-$ i $B \to h^+h'^-$ romanen similars després de la reconstrucció. Sota aquestes circumstàncies, la selecció de $B_s^0 \to \mu^+\mu^-$ i $B \to h^+h'^-$ mitjançant talls en comú garanteix que les propietats cinemàtiques d'aquests dos canals es mantenen similars després de la selecció, i permet la cancel·lació de l'efecte d'aquests talls. Això justifica l'ús de $B \to h^+h'^-$ per a calibrar la IM i la GL de $B_s^0 \to \mu^+\mu^-$, i per a normalitzar la mesura del seu BR. En cas de considerar els modes $B^0 \to K^+\pi^-$ o $B^0 \to \pi^+\pi^-$ en aquesta normalització, la principal font d'incertesa provindria del ràtio de fraccions de producció dels mesons B^0 i B_s^0, que actualment es coneix amb una incertesa del 13%. Finalment, el trigger introdueix importants diferències entre $B_s^0 \to \mu^+\mu^-$ i $B \to h^+h'^-$, que han de ser corregides per a la normalització de la mesura i la calibració de la GL.

La selecció original de $B_s^0 \to \mu^+\mu^-$ no ofereix prou rebuig pel fons combinatori de $B \to h^+h'^-$, degut a que el nombre mitjà d'hadrons per esdeveniment es situa dos ordres de magnitud per sobre del de muons. Per aquest motiu s'estudia la viabilitat de modificar la selecció original de $B_s^0 \to \mu^+\mu^-$ afegint talls en moment transvers i parametre d'impacte dels productes de desintegració. En particular, els talls sobre aquestes variables continguts en l'*stripping* de $B \to h^+h'^-$ destinat a la mesura de l'angle γ es poden afegir a la selecció de $B_s^0 \to \mu^+\mu^-$ sense afectar substancialment el senyal en la regió sensible. Aquests canvis en la selecció original de $B_s^0 \to \mu^+\mu^-$, amb algunes lleus odificacions adicionals, porten a una selecció comuna per a $B_s^0 \to \mu^+\mu^-$ i $B \to h^+h'^-$, tant per a controlar la mesura de $\mathrm{BR}(B_s^0 \to \mu^+\mu^-)$ com per a mesurar l'angle γ.

El ràtio entre senyal i fons (S/B) per a $B \to h^+h'^-$ seleccionat amb la nova selecció de $B_s^0 \to \mu^+\mu^-$ és igual a 0.19 ± 0.07. La major part d'aquest fons prové de la combinació de dues partícules completament independents, només el $(2.2\pm0.6)\%$ és degut a la combinació de dues partícules associades al mateix mesó B (*fons físic*). Les principals fonts de fons físic que s'han identificat són les desintegracions de mesons B amb tres hadrons a l'estat final, i les desintegracions d'hadrons Λ_b. La distribució de la IM d'aquests fons podria dur a una substracció incorrecta del fons total, i a dificultar la calibració de la IM.

La selecció de $B \to h^+h'^-$ es pot complementar amb l'ús de talls en la informació del tipus de partícula (PID), per tal d'obtenir la massa invariant correcta i reduir el fons. L'ús de definicions complementàries per a pions i kaons permet la selecció de tots els modes de $B \to h^+h'^-$ rebutjant la major part del fons físic degut a desintegracions de mesons B amb tres hadrons a l'estat final. D'altra banda,

per a seleccionar un canal en particular són necessaris talls més forts. També és possible aplicar talls per a identificar protons, que permeten reduir el fons degut a desintegracions d'hadrons Λ_b gairebé a la meitat.

Els talls en PID permeten calibrar la IM de $B_s^0 \to \mu^+\mu^-$ amb un únic canal de $B \to h^+h'^-$. En aquest cas, és necessari considerar els biaixos deguts tant a la ineficiència dels talls en PID com a la contaminació per part dels altres canals de $B \to h^+h'^-$.

La GL és independent de la hipotesi de massa, de manera que la seva calibració no requereix la selecció d'un únic mode de $B \to h^+h'^-$ mitjançant talls en PID. En aquest cas es troba que l'ús de definicions complementàries per a pions i kaons permet seleccionar tots els modes de $B \to h^+h'^-$ introduint un biaix negligible. Aquest procediment permet la calibració de la GL amb una incertesa de l'ordre del tant per cent per a 2 fb^{-1}, amb la major part de la incertesa deguda a l'alt nivell de fons per a la categoria d'esdeveniments amb GL<0.25. Per aquest motiu es proposa realitzar la mesura de BR($B_s^0 \to \mu^+\mu^-$) considerant només esdeveniments amb GL>0.25, donat que la sensibilitat oferta pel senyal rebutjat és negligible i es millora notablement la precisió de la calibració. Finalment, es demostra que és possible usar talls basats en la identificació de protons per a reduir el fons de desinteracions d'hadrons Λ_b sense introduir un biaix apreciable en la calibració de la GL.

Correcció dels efectes de reconstrucció, PID i trigger

Els efectes associats a la diferent reconstrucció de traces de muons i hadrons necessiten ser considerats en la normalització de la mesura de BR($B_s^0 \to \mu^+\mu^-$) degut a que el ràtio entre les eficiències de reconstrucció de $B_s^0 \to \mu^+\mu^-$ i $B \to h^+h'^-$ no es cancel·la. Aquest ràtio es pot estimar assumint que l'eficiència de reconstrucció d'un determinat canal de desintegració es pot factoritzar segons les eficiències de reconstrucció de les partícules del seu estat final. Sota aquest supòsit, a partir del nombre observat d'esdeveniments n i de la mesura del BR de les desintegracions $B^+ \to J/\psi(\mu^+\mu^-)K^+$ i $B^+ \to \bar{D}^0(K^+\pi^-)\pi^+$ es troba que

$$\frac{\epsilon_{rec,S}}{\epsilon_{rec,C}} = \frac{\epsilon_{trig,C'} BR_{C'}}{\epsilon_{trig,S'} BR_{S'}} \frac{n_{S'}}{n_{C'}},$$

on els subíndexs S i C es refereixen als canals $B_s^0 \to \mu^+\mu^-$ i $B \to h^+h'^-$ respectivament, i els subíndexs S' i C' es refereixen als canals $B^+ \to J/\psi(\mu^+\mu^-)K^+$ i $B^+ \to \bar{D}^0(K^+\pi^-)\pi^+$ respectivament.

La calibració de la IM fent servir un únic mode de $B \to h^+h'^-$ requereix un mètode per a controlar els efectes associats als talls en PID. L'efecte d'aplicar talls forts en PID no es pot corregir adequadament mitjançant una ponderació dels esdeveniments degut a que part de l'espai de fases dels productes de desintegració està suprimit. Per aquest motiu, es relaxen els talls en PID per tal de conservar tot l'espai de fases de l'estat final, amb el consegüent increment de la contaminació deguda als altres modes de $B \to h^+h'^-$. En primer lloc, l'efecte de la ineficiència dels talls en PID es

corregeix ponderant la contribució a la IM mesurada en *bins* de l'espai de fases de l'estat final. Aquesta ponderació aprofita el fet que l'espai de fases de l'estat final d'un únic mode de $B \to h^+h'^-$ és pràcticament igual que el de la suma de tots els modes de $B \to h^+h'^-$. I en segon lloc, l'efecte de la contaminació dels altres modes de $B \to h^+h'^-$ s'elimina considerant intérvals de la IM de $B^0 \to K^+\pi^-$ i $B_s^0 \to K^+K^-$ amb baix nivell de fons d'aquest tipus.

La correcció dels efectes del trigger és necessària per a la normalització de la mesura de BR($B_s^0 \to \mu^+\mu^-$), degut a que el ràtio entre les eficiències del trigger per a $B_s^0 \to \mu^+\mu^-$ i $B \to h^+h'^-$ no es cancel·la. A més, també és necessària per a la calibració de la GL. Aquests efectes poden ser tractats adequadament mitjançant la classificació d'esdeveniments segons si passen el trigger per propietats del senyal (TOS) o de la resta de l'esdeveniment (TIS). Els mètodes basats en aquesta classificació necessiten considerar les correlacions entre el senyal i la resta de l'esdeveniment. Aquestes correlacions tenen lloc només a través del moment del mesó B, de manera que el seu efecte es redueix prenent *bins* en l'espai de fases del mesó B. Això permet l'ús de l'equació

$$\frac{n_{TOS\&TIS}}{n_{TOS}} = \frac{n_{TIS}}{n_{sel}}$$

en cada un d'aquests *bins*, de manera que és possible extreure el nombre d'esdeveniments seleccionats en absència de trigger a partir només de quantitats observables i, per tant, conèixer l'eficiència del trigger. Fent servir tan sols tres *bins* en el moment transvers del mesó B, aquest mètode permet determinar l'eficiència del trigger per a $B \to h^+h'^-$ amb una desviació sistemática per sota de l'1%. Incloent l'efecte del fons, la incertesa estadística en aquest cas és del 2% per a 2 fb^{-1}, tot i que es pot reduir a 0.2% si només es consideren esdeveniments amb GL>0.25. D'altra banda, aquest mètode es pot aplicar per a corregir els biaixos en propietats del senyal i, amb lleus modificacions, també per a corregir biaixos en variables que involucrin propietats de la resta de l'esdeveniment.

Resultats de BR($B_s^0 \to \mu^+\mu^-$) amb les dades de 2010 i 2011

Les dades analitzades a LHCb fins a l'octubre de 2011 corresponen a un total de 300 pb^{-1}, i han servit per a donar una mesura de BR($B_s^0 \to \mu^+\mu^-$) competent amb la donada per Tevatron. L'anàlisi d'aquestes dades inclou diversos plantejaments desenvolupats en aquest document. Les desintegracions de tipus $B \to h^+h'^-$ han estat usades per a calibrar les variables discriminants i per a normalitzar la mesura de BR($B_s^0 \to \mu^+\mu^-$). Aquests canals han estat seleccionats en comú amb $B_s^0 \to \mu^+\mu^-$, mitjançant talls similars als proposats en aquesta tesi. Tot i que la calibració de la IM s'ha dut a terme interpolant les propietats de les desintegracions dimuòniques de resonàncies $c\bar{c}$ i $b\bar{b}$ a la massa de la B_s^0, els canals de tipus $B \to h^+h'^-$ s'han usat per a comprovar els resultats. La GL ha estat substituïda per una nova variable basada en arbres de decisió (BDT). Aquesta nova variable s'ha calibrat amb $B \to h^+h'^-$, i els biaixos introduïts pel trigger s'han corregit mitjançant l'ús d'esdeveniments TIS.

La normalització de la mesura s'ha dut a terme amb tres canals diferents, un dels quals ha estat $B^0 \to K^+\pi^-$. En aquest cas, el ràtio d'eficiències de reconstrucció s'ha estimat a partir de simulacions, mentre que el ràtio d'eficiències del trigger s'ha estimat amb l'ús d'esdeveniments TIS.

El resultat d'aquest anàlisi amb els 300 pb^{-1} de dades preses a LHCb estableix $\text{BR}(B_s^0 \to \mu^+\mu^-) < 1.6{\times}10^{-8}$ a 95% CL, que combinat amb el respectiu resultat per part de CMS dóna $\text{BR}(B_s^0 \to \mu^+\mu^-) < 1.08{\times}10^{-8}$ a 95% CL.

Summary

Theoretical predictions for the branching ratio of $B_s^0 \to \mu^+\mu^-$

Currently the Standard Model (SM) gives the most fundamental description of matter and its interactions. It has been validated by a large number of measurements that verify its consistency with an extraordinary precision. However, facts as the mass hierarchy problem, the evidences for non-zero neutrino masses or some cosmological measurements lead to consider the SM as an effective model of a more fundamental theory. For this reason, some theories beyond the SM have been developed in order to give a natural explanation to these facts, based on hypothesis as supersymmetry, technicolor or extra dimensions.

The phenomenology of models beyond the SM has been studied in detail in order to find measurements allowing their validation. In particular, by improving the current precision of certain observables in the B meson sector it would be possible to detect significant discrepancies with respect to the SM predictions. Some important examples are the violation of the CP symmetry in $B_s^0 \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$, the angular distributions of the $B^0 \to K^{*0}\mu^+\mu^-$ decay products and the branching ratio (BR) of $B_s^0 \to \mu^+\mu^-$. Before the LHC start-up, an upper limit for this last observable was set by Tevatron to 2.4×10^{-8} at 95% CL.

The $B_s^0 \to \mu^+\mu^-$ decay is strongly suppressed in the SM due to the fact that the weak coupling restricts the phase space of the final state. In this context, its BR can be predicted with low uncertainty given the current measurement of the mass difference between the two hamiltonian eigenstates of the B_s^0 system, finding a value equal to $(3.2\pm0.2)\times10^{-9}$. Several models beyond the SM based on supersymmetry or extra dimensions predict important departures from this value that could reach the current upper limit for some allowed regions of their parameter space.

The LHCb experiment

The LHCb detector has been specially designed for B physics studies in the environment of the 14 TeV proton-proton collisions provided by the LHC. The geometrical acceptance of LHCb, motivated by the angular distribution of B mesons produced in such conditions, comprise only directions close to the beam axis, and is restricted to one side of the interaction point. Approximately 10^{12} $b\bar{b}$ pairs are expected to be produced within this acceptance per nominal year (2 fb^{-1}).

The LHCb subdetectors are the following, ordered according to their distance to the interaction point: a high resolution tracker called *Vertex Locator* (VeLo), aiming to identify the decay vertices of B mesons; two detectors of Cerenkov radiation rings (RICH), designed in order to discriminate among different types of particles, in particular pions and kaons; a tracker system composed by the *Tracker Turicensis* (TT), the *Inner Tracker* (IT) and the *Outer Tracker* (OT), measuring the track of charged particles in a magnetic field; a calometry system built by the *Scintillator Pad Detector* (SPD), the *Pre-Shower* (PS), an electromagnetic calorimeter (ECAL) and an hadronic calorimeter (HCAL), used at the first stages of the trigger and for the detection of neutral particles; and a system of track chamber located beyond the calorimeters called *Muon System*, needed for the muon identification.

The LHCb trigger reduces the event rate from the 10 MHz produced by the LHC collisions down to the 2 kHz allowed by the storage resources. It is divided in two levels due to the computing limitations. The first level, called *Level-0* (L0), reduces the event rate down to 1 MHz using custom front-end electronics processing the information given by the VeLo, the calorimeters and the muon system. The second leve, called *High Level Trigger* (HLT) achieves the remaining rate reduction by means of algorithms operating in a computing farm.

Data taking and the control of LHCb is performed by three systems: the Data Acquisition system (DAQ), built by the readout and data transfer devices, and the computing resources used by the HLT algorithms (*Event Filter Farm*, EFF); the Timing and Fast Control system (TFC), which controls the data flux between the detector and the EFF; and the Experiment Control System (ECS), which allows to monitor and control the detector parameters, the DAQ and the TFC.

In the event reconstruction, primary vertices are searched using the VeLo information, and tracks are built from the VeLo, TT, IT and OT information. In addition, information for particle identification is given to each track, provided by the RICH, the calorimeters and the muon system. Finally, events are classified according to their properties (*stripping*).

The measurement of the branching ratio of $B_s^0 \to \mu^+\mu^-$ at LHCb

The procedure for measuring BR($B_s^0 \to \mu^+\mu^-$) at LHCb is based on the use of the statistical information contained in the distribution of discriminant variables. For this reason, the event selection is reduced to some loose cuts applied at the first level of the analysis.

The event selection starts from two reconstructed tracks, both being identified as muons. The final state of $B_s^0 \to \mu^+\mu^-$ is reconstructed with an efficiency equal to 0.682 ± 0.003, while this reconstructed final state is properly identified as a pair of muons with an efficiency equal to 0.838 ± 0.003. The cuts of the event selection are based on the quality of the dimuon vertex, its invariant mass and the compatibility between the flight direction of the reconstructed B meson and a primary interaction.

Cuts that could suppress parts of the phase space of the decay products, as those involving the transverse momentum or the impact parameter of these particles, are not included. The efficiency of this selection on reconstructed signal is equal to 0.922±0.002. Trigger is considered as a residual effect acting on the sample of selected events. The impact of the trigger on signal is low due to the high efficiency of the trigger cuts dedicated to muon channels, which enable a total trigger efficiency equal to 0.925±0.001 on selected signal. Including the effects of reconstruction, muon identification, selection and trigger, the total ammount of expected signal for 2 fb^{-1} is equal to 53 events.

The main background source is the combination of two muons coming from the decay of different B mesons, expecting around 2×10^5 of such events for 2 fb^{-1}. Other background sources are the combination of two muons from $B_c^+ \to J/\psi(\mu^+\mu^-)\mu^+\nu_\mu$ and the misidentification of the two hadrons from $B \to h^+h'^-$ as muons, leading to 70 and 17 events respectively for 2 fb^{-1}.

The discriminant variables used for extracting BR($B_s^0 \to \mu^+\mu^-$) from the sample of selected events are the invariant mass (IM) of the muon pair, and a variable called *geometrical likelihood* (GL). The latter variable contains information about the event geometry only, which is combined by means of multivariant analysis techniques. The GL does not include explicit information related to the momentum modulus of the decay products, in order to avoid correlations between this variable and the IM. This fact allows the independent calibration of each discriminant variable. For practical purposes, the allowed values of the GL range between 0 and 1. In addition, the GL is defined in order to be uniformly distributed for signal, and close to the origin for background. For this reason, the GL range between 0.5 and 1 is known as the *sensitive region*.

In order to extract BR($B_s^0 \to \mu^+\mu^-$), selected events are classified in different categories according to their values for the IM and the GL. The observed contents are compared with the respective contents expected for signal and background, by means of a modified frequentist approach called CL_s.

The performance of the method proposed for measuring BR($B_s^0 \to \mu^+\mu^-$) at LHCb is shown in Fig. 2 as a function of the integrated luminosity. LHCb has the potential to exclude values of BR($B_s^0 \to \mu^+\mu^-$) down to the SM prediction with less than 2 fb^{-1}. In addition, the observation of the SM prediction for this value at 3σ and 5σ is expected with 2 and 10 fb^{-1} respectively.

The expected signal contents used by the CL_s method are determined with control channels in order to avoid systematic effects derived from the use of simulations. Two control channels have been proposed: $B^+ \to J/\psi(\mu^+\mu^-)K^+$, which contains two muons in the final state that allow the partial cancellation of trigger and muon identification effects; and $B \to h^+h'^-$, where h represents either a pion or a kaon, which has the same geometrical and kinematical properties as signal, hence allowing the calibration of the two discriminant variables and the cancellation of event selection effects.

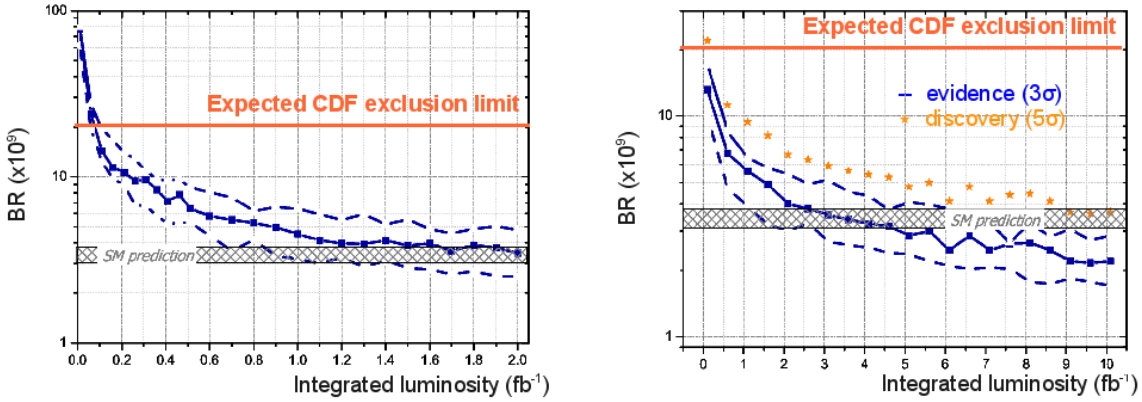

Figure 2: Exclusion limit at 90% CL (left), and observation at 3σ and 5σ (right), as a function of the integrated luminosity $\mathcal{L}$.

The use of $B \to h^+h'^-$ as control channels for BR($B_s^0 \to \mu^+\mu^-$)

The BR of $B \to h^+h'^-$ decays are of the order of 10^{-6}, as shown in Table 2. The most abundant decay is $B^0 \to K^+\pi^-$, followed by $B_s^0 \to K^+K^-$ and $B^0 \to \pi^+\pi^-$. The total expected ammount of $B \to h^+h'^-$ within the LHCb acceptance is around 14 million events per 2 fb^{-1}. While the BR of $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$ has been determined with high precision at the B *factories*, the measure of the BR of $B_s^0 \to K^+K^-$ and $B_s^0 \to \pi^+K^-$ is affected by the 13% uncertainty in the measure of the production fraction of the B_s^0 meson.

Decay	BR $(\times 10^5)$	$2\sigma_{b\bar{b}} f_{\rm B}$BR (nb)
$B^0 \to \pi^+\pi^-$	0.513 ± 0.024	1.05 ± 0.05
$B^0 \to K^+\pi^-$	1.94 ± 0.06	4.0 ± 0.1
$B_d \to KK$	<0.12, 90% CL	<0.06, 90% CL
$B_s \to \pi\pi$	<0.041, 90% CL	<0.08, 90% CL
$B_s^0 \to \pi^+K^-$	0.49 ± 0.10	0.24 ± 0.05
$B_s^0 \to K^+K^-$	3.3 ± 0.9	1.6 ± 0.4

Table 2: Current averages of the BR and expected production rate of $B \to h^+h'^-$ decays.

The mass difference between the B^0 and B_s^0 mesons and between muons, pions and kaons are negligible with respect to the momentum of the decay products of $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$. This fact justifies to assume that the differences between the kinematical properties of $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ can be neglected at production level.

The fact that a significant fraction of hadrons is unable to cross the entire tracking system due to their interaction with the detector material introduces differences between $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ at reconstruction level. The effect of such

differences is practically independent of the momentum of the decay products, hence the kinematical properties of $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ can still be considered similar after the reconstruction. In addition, the selection of $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ by means of common cuts ensures that the kinematical properties of these channels remain similar after the event selection, allowing the cancellation of the effect of such cuts. This justifies the use of $B \to h^+h'^-$ for calibrating the IM and the GL of $B_s^0 \to \mu^+\mu^-$, and for normalizing the measurement of its BR. In case of considering $B^0 \to K^+\pi^-$ or $B^0 \to \pi^+\pi^-$ for such normalization, the main source of uncertainty would come from the ratio of production fractions of the B^0 and B_s^0 mesons, which is currently measured with 13% accuracy. Finally, the trigger introduces important differences between $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$, that need to be corrected for both the normalization of the measurement and the calibration of the GL.

The original $B_s^0 \to \mu^+\mu^-$ selection does not provide enough background rejection for $B \to h^+h'^-$ due to the fact that the average number of hadrons per event is two orders of magnitude above the average number of muons per event. For this reason it is necessary to consider the feasibility of modifying the original $B_s^0 \to \mu^+\mu^-$ selection by adding cuts on the transverse momentum and impact parameter of the decay products. In particular, the cuts on these variables provided by the stripping of $B \to h^+h'^-$ used for the measurement of the γ angle can be added to the selection of $B_s^0 \to \mu^+\mu^-$ without affecting substantially the sensitive signal. These changes in the original selection of $B_s^0 \to \mu^+\mu^-$, plus slight additional modifications, lead to a common selection for $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$, either for control the measurement of $\mathrm{BR}(B_s^0 \to \mu^+\mu^-)$ or for measuring the γ angle.

The signal-to-background ratio (S/B) for $B \to h^+h'^-$ selected using the new $B_s^0 \to \mu^+\mu^-$ selection is equal to 0.19 ± 0.07. The bulk of this background is due to the combination of two independent particles, while only $(2.2\pm0.6)\%$ is due to the combination of two particles coming from the same B decay (*physical background*). The main sources of physical background that have been identified are B decays to three hadrons in the final state, and Λ_b decays. The IM distributions of these backgrounds could difficult the proper background subtraction using mass sidebands.

The selection of $B \to h^+h'^-$ can be complemented with the use of the information for particle identification (PID), in order to obtain the correct mass hypothesis and reduce the background levels. The use of complementary definitions for pions and kaons allows the selection of all $B \to h^+h'^-$ modes while rejecting the bulk of the physical background due to B decays with three hadrons in the final state. The selection of a single $B \to h^+h'^-$ mode requires tighter PID cuts. In addition, the PID information can be used for proton identification, which allows to reject half of the physical background due to Λ_b.

PID cuts allow the calibration of the IM of $B_s^0 \to \mu^+\mu^-$ with a single $B \to h^+h'^-$ mode. In this case, it is necessary to consider the biases due to both the inefficiency of the PID cuts and the pollution due to the remaining $B \to h^+h'^-$ modes.

The GL is independent of the mass hypothesis, hence its calibration does not require the selection of a single $B \to h^+h'^-$ mode by means of PID cuts. In this case,

the use of complementary definitions for pions and kaons allows the selection of all $B \to h^+h'^-$ modes introducing a negligible bias. This leads to an uncertainty at the percent level in the calibration of the GL for 2 fb^{-1}, with most of this uncertainty due to the large background levels for the category of events having GL<0.25. For this reason, it is proposed to measure BR($B_s^0 \to \mu^+\mu^-$) by considering only events satisfying GL>0.25, as the lost in sensitivity is negligible and the precision of the calibration is significantly improved. The use of cuts based on the proton identification in order to reject the physical background due to Λ_b decays do not introduce significant biases in the calibration of the GL.

Correction of reconstruction, PID and trigger effects

Effects arising from the different track reconstruction of muons and kaons need to be considered in the normalization of the measurement of BR($B_s^0 \to \mu^+\mu^-$) as the ratio between the reconstruction efficiencies of $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ does not cancel out. This ratio can be estimated by assuming that the reconstruction efficiency of a given decay channel can be factored as the product of the reconstruction efficiencies of the particles composing its final state. Under such assumption, from the observed number of events n and the measure of the BR for $B^+ \to J/\psi(\mu^+\mu^-)K^+$ and $B^+ \to \bar{D}^0(K^+\pi^-)\pi^+$ decays it is possible to write

$$\frac{\epsilon_{rec,S}}{\epsilon_{rec,C}} = \frac{\epsilon_{trig,C'} BR_{C'}}{\epsilon_{trig,S'} BR_{S'}} \frac{n_{S'}}{n_{C'}},$$

where the subindices S and C denote the $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ channels respectively, and the subindices S' and C' denote the $B^+ \to J/\psi(\mu^+\mu^-)K^+$ and $B^+ \to \bar{D}^0(K^+\pi^-)\pi^+$ channels respectively.

The calibration of the IM using a single $B \to h^+h'^-$ mode needs to control the effects of the PID cuts. The effect of tight PID cuts can not be corrected properly by weighting the event sample as part of the phase space of the decay products is suppressed. For this reason, PID cuts are relaxed in order to preserve the whole of the phase space of the final state, even though this leads to an increased pollution due to the remaining $B \to h^+h'^-$ modes. The effects due to the inefficiency of the PID cuts are corrected by weighting the contribution to the IM measured in bins of the phase space of the final state. This weighting uses the fact that the phase space of the final state for a single $B \to h^+h'^-$ mode is practically the same as that for the sum of all $B \to h^+h'^-$ modes. The effects due to the pollution of the remaining $B \to h^+h'^-$ modes are avoided by considering only IM ranges for $B^0 \to K^+\pi^-$ and $B_s^0 \to K^+K^-$ with low levels of such background.

The correction of trigger effects is necessary in the normalization of the measurement of BR($B_s^0 \to \mu^+\mu^-$), due to the fact that the ratio between the trigger efficiencies of $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ does not cancel out. It is also necessary for the calibration of the GL. These effects can be handled properly by classifying events as passing the trigger by means of signal properties (TOS) or properties in-

dependent of signal (TIS). The methods based on this classification need to consider the correlations between the properties of signal and those for the rest of the event Such correlations occur through the momentum of the signal B meson only, hence their effects can be reduced by binning the phase space of this B meson. This allows to use the equation

$$\frac{n_{TOS\&TIS}}{n_{TOS}} = \frac{n_{TIS}}{n_{sel}}$$

in each such bin, enabling the extraction of the number of selected events with no trigger requirements from observable quantities only, and hence leading to the estimation of the trigger efficiency. Using three bins in the transverse momentum of the signal B meson, this method allows to estimate the trigger efficiency for $B \to h^+h'^-$ with a systematic deviation below 1%. By including the effect of background, the statistical uncertainty is equal to 2% for 2 fb^{-1}, although it can be reduced to 0.2% if only events having GL>0.25 are considered. In addition, this method can be applied for correcting the trigger biases from signal properties and, with slight modifications, for correcting the trigger biases from variables involving non-signal properties.

Results on BR($B_s^0 \to \mu^+\mu^-$) with the data from 2010 and 2011

Currently the LHCb detector has collected 300 pb^{-1} of data, which has been used for giving a measure of BR($B_s^0 \to \mu^+\mu^-$) competitive with that provided by Tevatron. The analysis of this data includes several issues developed in this document. $B \to h^+h'^-$ decays have been used for calibrating the discriminant variables and for normalizing the measure of BR($B_s^0 \to \mu^+\mu^-$). These channels have been selected in common with $B_s^0 \to \mu^+\mu^-$, by means of cuts similar to those proposed in this thesis. Although the calibration of the IM has been performed by interpolating the properties of dimuonic decays of $c\bar{c}$ and $b\bar{b}$ resonances to the B_s^0 mass, methods using $B \to h^+h'^-$ channels have been used for checking these results. The GL has been replaced by a new variable based on boosted decision trees (BDT). It has been calibrated with $B \to h^+h'^-$, and the biases introduced by the trigger have been corrected using TIS events. The normalization of the measurement has been performed with three different channels, including $B^0 \to K^+\pi^-$. In this case, the ratio of reconstruction efficiencies has been estimated from simulations, while the ratio of trigger efficiencies has been estimated using TIS events.

The result of this analysis with the first 300 pb^{-1} collected by LHCb sets BR($B_s^0 \to \mu^+\mu^-$) $< 1.6\times10^{-8}$ at 95 CL%, while its combination with the respective result from CMS gives BR($B_s^0 \to \mu^+\mu^-$) $< 1.08\times10^{-8}$ a 95 CL%.

Contents

Introduction

In spite of the impressive success of the Standard Model, some theoretical and experimental facts suggest that it might not be the *final* theory, and hence that *new physics* exists. Decays of B meson are an exceptional source of observables where physics beyond the SM could lead to unambiguous effects. Indeed, measurements in this sector provide some of the first possibilities of observing effects from new physics at the LHC. This is the case of the branching ratio of the yet unobserved decay $B_s^0 \to \mu^+\mu^-$. This observable is precisely predicted by the Standard Model and hence it is also very sensitive to new physics. The conception of LHCb as a dedicated experiment for B physics studies puts it in an outstanding position for measuring this branching ratio with respect to other experiments operating at the LHC.

The measurement of the branching ratio of $B_s^0 \to \mu^+\mu^-$ at LHCb widely uses methods for both calibrating and normalizing independently of the Monte Carlo simulations, based on control channels. The family of two-body decays of B mesons into charged pions and kaons, referred to as $B \to h^+h'^-$, are particularly interesting due to the very similar kinematical properties with respect to those of $B_s^0 \to \mu^+\mu^-$. The possible uses of such channels within the measurement of the branching ratio of $B_s^0 \to \mu^+\mu^-$, as well as the methods that have been developed in order to correct the differences between $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$, constitute the subject of this thesis.

Chapter 1 justifies the interest of the measurement of the branching ratio of $B_s^0 \to \mu^+\mu^-$, as a test of the Standard Model and its extensions. Chapter 2 discusses the requirements for a B physics experiment at the LHC and describes the LHCb detector. Chapter 3 explains the analysis that has been developed at LHCb for the measurement of the branching ratio of $B_s^0 \to \mu^+\mu^-$. The event selection and the definition of discriminant variables are discussed in detail. In addition, the use of $B \to h^+h'^-$ as control channels for this analysis is introduced. Chapter 4 explains the uses of $B \to h^+h'^-$ for calibrating and normalizing the measurement, and shows the similarities and the differences between this family of channels and $B_s^0 \to \mu^+\mu^-$. The common selection of $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ and the use of information for particle identification are discussed in detail. Chapter 5 describes methods for correcting the differences between $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ caused by track reconstruction, particle identification and trigger. Finally, Chapter 6 presents the current measurement of the upper limit for $\mathrm{BR}(B_s^0 \to \mu^+\mu^-)$ set by LHCb, and

discusses the impact of the studies developed in this document.

Chapter 1

Theoretical predictions for the branching ratio of $B_s^0 \to \mu^+\mu^-$

This chapter justifies the interest of the measurement of the branching ratio (BR) of $B_s^0 \to \mu^+\mu^-$, called BR($B_s^0 \to \mu^+\mu^-$) hereafter, as a test of the Standard Model (SM) and its extensions. Section 1.1 reviews the SM, as well as the motivation for searching for a more complete theory. Some important examples of such theories are presented. Section 1.2 describes the rich phenomenology of B meson decays, and its importance for the precise determination of parameters of the SM, as well as for the search for physics beyond the SM. Section 1.3 introduces the use of effective theories in phenomenological analyses, focusing on the case of B decays. Finally, Section 1.4 discusses the prediction of BR($B_s^0 \to \mu^+\mu^-$) in the SM, as well as the departures that are expected in some of its extensions.

1.1 Models of fundamental interactions

So far, practically all the observations in high energy physics have been in excellent agreement with the SM. However, in spite of this experimental success, the SM does not provide a satisfactory answer to some fundamental questions. This fact has motivated the rise of theories extending the SM, such as supersymmetry, technicolor or extra dimensions.

1.1.1 The Standard Model

The SM is constructed within the framework of Quantum Field Theory. As such, it is described by a lagrangian from which measurable quantities can be derived, mainly cross sections and BRs. This lagrangian has been developed according to the experimental facts and some invariance principles.

In the SM, fermion fields represent the concept of fundamental constituents of matter. Interactions between fermions arise in a natural way by simply requiring invariance of the lagrangian under the gauge transformation of certain multiplets of

these fields. By this mechanism, a definite set of boson fields is introduced in the theory, representing the interaction carriers.

The matter constituents and the interactions of the SM are discussed with more detail in the following.

Matter constituents

The content of the fermionic sector of the SM is shown in Table 1.1 [1]. There are three families of fermion fields, with a common structure, differing only in the mass of their components. Each family is composed of one charged lepton, one neutral lepton, and two sets of three quarks each. The charged leptons and the quarks have a non-zero mass, while the neutrinos are massless. The label used for quarks of a given mass is known as *flavor*. In turn, each flavor implies three degenerate mass states, which are conventionally labelled as *blue*, *green* and *red*.

	leptons		quarks	
Family 1	electron, e^-	electronic neutrino, ν_e	down, d	up, u
Family 2	muon, μ^-	muonic neutrino, ν_μ	strange, s	charm, c
Family 3	tauon, τ^-	tauonic neutrino, ν_τ	bottom, b	top, t
Electric charge	-1	0	-1/3	2/3

Table 1.1: Fermionic content of the SM. Each quark flavor has a three-fold mass degeneracy.

The electromagnetic and weak interactions

Both the electromagnetic and weak interactions are described in the SM lagrangian by terms consisting in a vector current coupled to a fermionic current. This similarity pointed out to a unified approach to these interactions [2][3][4], based on the invariance of the lagrangian under two gauge transformations:

- The structure of the fermionic currents that couple to the weak currents, which involve left-handed components only, suggests invariance under the $SU(2)$ transformation of doublets of left-handed fermions of the form $(\nu_L l_L)$ and $(q_L q'_L)$, where each member of the doublet belong to the same family.
- The fact that electromagnetic interaction affects on equal foot fermions of both helicities suggests invariance under the $U(1)$ transformation of any fermion field of definite helicity, either left-handed or right-handed.

Apart from these symmetries, the SM description of electroweak interactions has two more fundamental features: the spontaneous symmetry breaking of the $SU(2) \times U(1)$ invariance (EWSB); and the mixing of the quark flavors through the Cabibbo-Kobayashi-Maskawa (CKM) matrix.

Spontaneous symmetry breaking of $SU(2) \times U(1)$. Experiments have strongly stablished that the SM fermions and the weak gauge bosons are massive. If these masses

are explained by just introducing mass terms in the SM lagrangian, the gauge invariance is lost. This is overcome by requiring spontaneous symmetry breaking [5] of the $SU(2) \times U(1)$ invariance through the so-called *Higgs mechanism* [6][7][8]. This mechanism yields a gauge invariant lagrangian, enabling the fermions and gauge bosons to have mass by introducing a degenerate ground state that is not symmetric under $SU(2) \times U(1)$. The specific realization of the Higgs mechanism in the SM operates using two complex scalar fields ϕ^+ and ϕ^0 with an interaction term $\lambda(\Phi^\dagger\Phi)^2$, where $\Phi \equiv (\phi^+\phi^0)$. Three physical states arising from electroweak gauge fields, identified with the W and Z bosons respectively, acquire a non-zero mass by absorbing three of the four degrees of freedom introduced by this doublet of complex scalar fields. The remaining degree of freedom results in a physical scalar boson with non-zero mass, called the *Higgs boson*, which generates the mass of the SM fermions through its coupling to their fields. So far, the Higgs boson still remains unobserved. Its mass is a free parameter of the SM, and from direct searches the range below 114.4 GeV/c^2 has been excluded at 95% CL [9].

Quark flavor mixing matrix. Since the doublets that are invariant under the $SU(2)$ gauge transformation contain only quarks belonging to the same family, weak processes arising only from the invariance under $SU(2) \times U(1)$ are unable to mix quarks belonging to different families. Such mixing is observed in Nature, and it is introduced in the SM by postulating that the quark fields entering in the weak interaction term are not expressed in the basis that diagonalizes the time evolution operator of the theory. The mixing between different families of quarks in weak interactions arise when these fields are expressed in such basis. The change of basis is expressed by means of a matrix representing an unitary transformation, referred as *CKM mixing matrix* [10][11]. This matrix depends on three independent real parameters and one irreducible complex phase. A key property of the CKM matrix is that a non-zero value of this complex phase leads to violation of the charge-parity (CP) symmetry in weak processes involving mixing of different families of quarks. Within the SM, this is the unique source of violation of the CP symmetry.

The strong interaction

Strong interactions are independent of the color of the quarks. This feature suggests the invariance of the lagrangian under a gauge $SU(3)$ transformation of quark triplets of the form (q_B, q_G, q_R), where q and its subscript indicate the flavor and the color of the quark respectively [12][13]. This requirement is sufficient to explain the observed features of the strong interaction [14][15][16], mediated by eight massless bosons called *gluons*.

1.1.2 Motivations for extending the SM

The predictions of the SM have been confirmed in practically all experiments, reaching in several cases a high degree of accuracy. However, this model suffers from some

theoretical limitations, and is unable to explain some observational facts.

Theoretical limitations

The main concerns of the theoretical framework of the SM are the mass hierarchy problem, and the lack of a description of the gravitational interaction.

The mass hierarchy problem. In the SM, the quantum corrections to the mass of the Higgs boson (m_H) are heavily affected by any physics existing at some scale Λ above the typical scale of the electroweak interactions ($\Lambda_{EW} \sim m_{W,Z}$), setting inevitably $m_H \sim \Lambda$ [17][18][19]. In particular, the existence of the gravitational interaction, related to the Plack scale, implies that m_H must be at least 17 orders of magnitude above Λ_{EW}. However, an indirect upper bound for m_H is already set from precision measurements of electroweak parameters, quoting $m_H < 192$ GeV/c^2 at 99% CL [20]. This result indicates that within the SM the mass of the Higgs boson must be affected by very delicate cancellations. These cancellations are rather artificial, and spoil the naturalness of the SM.

Number of free parameters. The SM lagrangian contains several parameters whose value is not predicted by the theory, as the mass of the Higgs boson, or the strength of the couplings between the Higgs and each SM fermion. Assuming that Nature only contains few fundamental constants, there must exist a more fundamental theory determining the value of most of the parameters of the SM.

Impossibility of describing the gravitational interaction. The attempts of including this interaction in the framework of the SM fail due to the lack of renormalizability of the resulting theory.

Observational inconsistencies

Few results from particle physics experiments, as well as measurements from cosmological observations, are inconsistent with the SM.

Massive neutrinos. The neutrino fields entering the SM are massless by definition. This assumption has been questioned after the observation of neutrino oscillations [21][22][23]. Massive neutrinos offer a natural explanation of this phenomenon, and can be accomodated in the SM without introducing substantial changes in the theory [25].

Anomalous magnetic moment of the muon. The latest measurement of this quantity shows a discrepancy of 3σ with respect to the SM prediction [26]. This represents an interesting result since this observable is relatively sensitive to contributions of physics beyond the SM.

Matter content in the Universe. Astronomical observations of the rotation of galaxies have made evident that the gravitational effect of the visible matter alone can not account for the observed rotational velocities [27][28][29]. In fact, observations of the cosmic microwave background indicate that visible matter would represent roughly 5% of the total ammount of non-relativistic matter of the Universe [30]. The remaining ammount of matter would remain undetected, at least to explorations in the electromagnetic spectrum, and hence it is called *dark matter (DM)* [31]. This fact leads to assume that its fundamental constituents are electrically neutral. The SM does not contain stable massive neutral particles that can contribute to the cosmical content of non-relativistic matter.

Matter/antimatter asymmmetry. Cosmological models assume that both particles and antiparticles were equally numerous in the initial Universe. However, a large asymmetry between the ammounts of matter and antimatter is observed in astronomical observations. A necessary condition in order for the early Universe to evolve into such asymmetric state is the violation of the CP symmetry [32]. The CKM mechanism does introduce such violation in the SM, but with a magnitude that is not able to account for the total difference between the actual contents of matter and antimatter [33].

The previous considerations show that the SM cannot be considered as a final theory. Several extensions of the SM have been developed in order to overcome at least some of these problematic issues, usually with special attention on the hierarchy problem.

1.1.3 Extensions of the SM

The main theories extending the SM are based on new mechanisms enabling the EWSB [34]. Such models can be classified as either *linear* or *non-linear*, according to whether the fields producing the EWSB transform linearly or non-linearly under $SU(2) \times U(1)$. Linear models imply elementary Higgs bosons, as in the SM extension using two Higgs doublets, or in supersymmetric theories. Non-linear models imply effective Higgs bosons, as composite Higgs bosons in technicolor models, or four-dimensional projections of elementary Higgs boson fields in theories with extra dimensions. The fundamental differences among these models lead to radical variations among their respective phenomenological predictions. In the following, some of the most currently accepted extensions of the SM are briefly described.

SM with two Higgs doublets (2HDM)

In the SM, the EWSB is achieved by means of one doublet of complex scalar Higgs fields Φ [35]. This represents the simplest implementation of the Higgs mechanism. However, the existence of two doublets of complex scalar Higgs fields Φ_1 and Φ_2, each one developing a non-zero vacuum expectation value v_1 and v_2, also leads to

the required EWSB. This is often considered as the minimal extension of the SM. The two doublets of complex scalar fields introduce eight degrees of freedom. Three of them are absorbed by electroweak gauge fields in order to produce the physical states identified with the W and Z bosons, just as in the SM description of the EWSB. The remaining degrees of freedom result in five physical Higgs bosons:

- Two charged scalar states with equal masses ($H^{\pm}$).
- Two neutral scalar states that mix together by an angle α to form two mass eigenstates (h^0 and H^0).
- One neutral pseudoscalar state (A^0).

This model contains six free parameters: the four Higgs masses; the mixing angle α; and the ratio of vacuum expectation values $\tan\beta = v_1/v_2$. It ramifies into two submodels, according to the assumptions about the interactions between the Higgs and fermionic sectors, named 2HDM-I and 2HDM-II. In the first case, fermions only couple to one of the two doublets. In the second case, down-type quarks and charged leptons couple to one doublet, and up-type quarks and neutrinos couple to the other doublet.

Supersymmetry

The concept of supersymmetry (SUSY) refers to the invariance of a theory under transformations mixing fermion and boson fields [36][37][38]. Any set of such fermion and boson fields that is invariant under these transformations is called a *supermultiplet*, and contains the same number of fermionic and bosonic degrees of freedom.

In the context of SUSY, the mass hierarchy problem is naturally handled by assuming that any particle in Nature is a member of a certain supermultiplet. Then, the quantum corrections introduced in the propagator of the Higgs boson by the fermionic and bosonic sectors of a given supermultiplet cancel out, and hence its mass is kept within the electroweak scale, regardless of the physics above this scale. The minimal extension of the SM satisfying this assumption is named *Minimal Supersymmetric Standard Model* (MSSM) [39]. Its structure is that of the 2HDM-II, together with the corresponding supersymmetric sector. There are two additional arguments supporting this extension of the SM. First, it provides a natural candidate for DM by assuming conservation of the R-parity, which is defined as $(-1)^{3B+L+2S}$, with B and L being the baryon and lepton numbers respectively, and S being the spin. Under this symmetry, introduced in order to protect the proton decay, the lightest supersymmetric particle turns out to be stable within the MSSM. And second, unlike the SM, the three coupling constants associated to the $U(1)$, $SU(2)$ and $SU(3)$ gauge symmetries respectively unify at a common scale, of the order of 10^{16} GeV, called *unification scale* (Λ_U). Notoriously, this last statement still remains valid in SUSY theories including the gravitational interaction.

By now, experiments rule out the existence of supersymmetric particles with masses equal to those of their SM partners. This empirical fact implies that, if

the actual lagrangian is invariant under SUSY transformations, then SUSY is spontaneously broken [40][41]. The main possibilities that have been proposed comprise symmetry breakings mediated by either gravity (SUGRA) [42], gauge fields (GMSB) [43] or anomalies (AMSB) [44]. If no explicit SUSY breaking mechanism is required, the theory is left with more than one hundred parameters. This number is usually lowered by assuming that some parameters are *universal*, *i.e.* they have the same value, at Λ_{U}:

- The *constrained MSSM* (CMSSM) [45][46][47]. It assumes universality for the mass of scalar fields, for the mass of gaugino fields, and for the coupling of triple scalar vertices. These three universal quantities, together with the ratio of the vacuum expectation values generated by the Higgs fields, $\tan\beta$, and the sign of the Higgsino mixing parameter, $\mathrm{sign}(\mu)$, are conventionally chosen as the free parameters of this model.
- The *MSSM with non-universal Higgs masses* (MSSM-NUHM) [48]. In the CMSSM, both the Higgs and the scalar superpartners of the fermions have universal masses at the unification scale. However, the assumption of universality in the Higgs masses lacks of a strict justification. This leads to the MSSM-NUHM, that inherits all the universalities assumptions of the CMSSM except those in the Higgs masses. Two new free parameters are then necessary, that are chosen to be the modulus of μ and the mass of the pseudoscalar Higgs M_A.

Technicolor

In the SM, the EWSB is assumed to be originated by elementary scalar fields. This symmetry breaking can also be achieved by means of fermion condensates [49][50]. To that purpose, it is necessary to introduce new fermions (*technifermions*), and a new interaction among them arising from some non-abelian gauge symmetry (*technicolor*[1]). Then, technifermion condensates act as the Goldstone bosons generating the mass of the W and Z bosons.

Although technifermion condensates enable massive weak gauge bosons, the SM fermions still remain massless. This problem can be overcome by embedding the color and technicolor gauge symmetries into a larger gauge symmetry group, called *extended techicolor* [51], which is broken down to color and technicolor. In addition, a further interaction is introduced in order to provide a natural explanation for the large value of the top quark mass, called *topcolor*, featuring a strong coupling to the top quark field [52][53]. The combination of topcolor and technicolor is named *topcolor assisted technicolor* (TC2) [54], and represents a rather satisfactory extension of the SM.

[1]This name is due to the fact that, in analogy with QCD, this interaction is related to a non-abelian gauge symmetry.

Extra dimensions

A different approach for extending the SM is to consider variations of the standard four-dimensional spacetime structure. Based on this idea, two different frameworks offer models in which the large difference between the energy scales of gravity and the SM interactions is not of intrinsic nature, but due to the spacetime structure. This fact enables the intrinsic energy scales of gravity and the SM interactions to be of the same order, hence solving automatically the hierarchy problem. Both frameworks modify the standard four-dimensional spacetime by adding extra dimensions. However, the underlying assumptions in each case are fundamentally different.

Flat extra dimensions (ADD models). The large energy scale associated to gravity is due to the small values of its strength that have been observed experimentally. By assuming that such measured strength is just an effective value arising from a much strong gravitational coupling, it is possible to build a model in which the energy scale of gravity is situated at the TeV scale. Under such conditions, it is necessary to introduce a mechanism in order to explain the large difference between the *observed* and the *intrinsic* strengths of this interaction. ADD models postulate that the effect of gravity is distributed among extra dimensions of finite size, while that of the SM interactions is confined in the subspace constituting the conventional four-dimensional spacetime [55][56]. Restricting to this subspace, the ADD setup leads to an effective dillution of the strength of gravity with respect to those of the SM interactions, scaling with the number of extra dimensions and their respective sizes.

Warped extra dimensions (RS models). In curved spacetimes, the energy scale of a given interaction depends on the position at which it is considered. This is in analogy with the dependence of the frequency of an electromagnetic emitter on its position in a spacetime curved by the effect of gravity. RS models introduce extra dimensions having an extreme instrinsic curvature [57]. In this case, large differences between the energy scales of gravity and the SM interactions can develop at certain regions of the spacetime due only to the intense intrinsic curvature [58][59]. A satisfactory model has been proposed by enlarging the standard four-dimensional spacetime with a single finite extra dimension. In addition, SM matter is assumed to be confined to the four-dimensional subspace located at one end of the extra coordinate. For this setup, Einstein equations predict an exponential warping along the extra dimension, which naturally leads to a large difference between the energy scales of gravity and the SM interactions in the subspace where SM matter is confined.

1.2 The B meson sector

B mesons represent a very attractive area to investigate not only flavor physics in the SM but also signatures of physics beyond the SM. Although all the current measurements related to this sector are consistent with the SM, for some cases the

experimental accuracy has not yet matched the precision of the theoretical prediction. Besides, some rare B decays very sensitive to new physics (NP) have not been observed or studied in depth yet, as in the case of $B_s^0 \to \mu^+\mu^-$ or $B^0 \to K^{*0}\mu^+\mu^-$.

1.2.1 Weak processes in the B meson sector

Weak interactions describe all the decays of the B^0, $B^\pm$ and B_s mesons, as well as the mixing of neutral B mesons. Measurements concerning these processes allow the determination of the parameters of the CKM matrix, and enable SM tests by checking the consistency of such results with those coming from the K and D sectors.

Of particular importance are the measurements concerning the violation of the CP symmetry (CPV). For convenience, the information about the CPV in a given sector is usually condensed in a relation derived from the unitarity of the CKM matrix, called *unitary triangle*. In the case of the sector of the B^0 mesons, the corresponding unitary triangle is $V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$, where each quantity of the form $V_{qq'}$ refers to the element of the CKM matrix mixing the q and q' quark flavors. Viewing the complex numbers $V_{ud}V_{ub}^*$, $V_{cd}V_{cb}^*$, and $V_{td}V_{tb}^*$ as vectors in the complex plane, this relation is equivalent to state that these vectors form a triangle in this plane, see Fig. 1.1.

The current measurements related to the parameters of the CKM matrix involving B^0 mesons come from both the B factories and Tevatron. Particularly, the three angles of the unitarity triangle of the B^0 meson are found to be $\alpha = (88^{+6}_{-5})^\circ$ [65], $\sin 2\beta = 0.681 \pm 0.025$ [66] and $\gamma = (77^{+30}_{-32})^\circ$ [65]. LHCb is expected to become an important step forward in reducing the large uncertainties that still affect some measurements, as the γ angle.

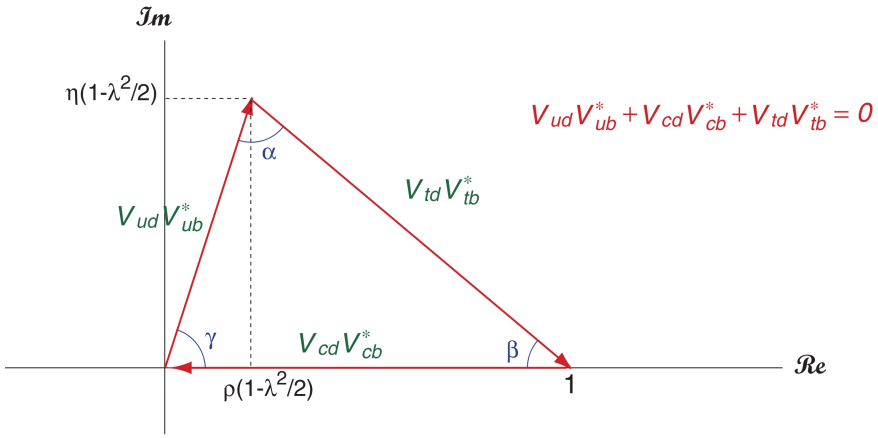

Figure 1.1: The unitary triangle of the B^0 sector.

1.2.2 The B meson sector beyond the SM

The approaches used in the search of physics beyond the SM can be classified as either *direct* or *indirect*. In the first case, the new particles are produced on-shell, and they can be identified through the reconstruction of the final state of their decay. In the second case, the existence of NP is made evident by its effect on some measurable quantity such as angular distributions or BRs. The searches carried out in the B meson sector are instances of the indirect approach, based on the fact that new heavy particles produced virtually in loops could affect the value of certain observables. Such effects provide an unambiguous evidence of NP if their strength is at least of the size of the SM uncertainty.

The main searches for physics beyond the SM in the B sector are related to measurements involving either CPV or flavor changing neutral currents (FCNC). In the latter case, no tree-level contribution is allowed within the SM, leading to the prediction of small BRs. For this reason, such decays are often referred to as *rare decays*. At the moment, some of the most important measurements that are considered as more promising at LHCb are the following:

- CPV in $B_s^0 \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ [67]. The effect of CPV in the interference between the decay of B^0 and $\bar{B}^0$ mesons to $J/\psi\phi$ either directly or via the $B^0 - \bar{B}^0$ oscillation is parametrized by a mixing phase called ϕ_s. Within the SM, this quantity is double-Cabibbo suppressed, being its predicted value equal to $\phi_s = (0.0360^{+0.0020}_{-0.0016})^\circ$. Current analyses from Tevatron agree with the SM [68][69], but further data is necessary in order to match with the accuracy of the theoretical prediction.

- Angle γ from $B \to h^+h'^-$ decays [70][71]. Unlike other methods aimed to extract γ, the measurement of this angle using $B \to h^+h'^-$ decays use processes having CPV effects that receive contributions from penguin diagrams. The value of γ obtained from such processes could be affected by NP arising in the loop of the penguin diagram, hence implying potential discrepancies between this result and that given by other procedures. $B \to h^+h'^-$ channels have BRs of the order of 10^{-6}, having been observed at the B factories [72][73], Tevatron [74] and recently at LHCb, where $\sim 10^3$ $B \to h^+h'^-$ events have been collected by the end of 2010. However, an analysis addressed to the measurement of the γ angle from these processes is still lacking.

- Angular distributions of the decay products of $B^0 \to K^{*0}\mu^+\mu^-$ [75]. The most interesting observable is the angular distribution of the two muons of the final state, as a function of their invariant mass. Currently, there are some analyses coming from both the B factories [76][77][78] and Tevatron [79], but lacking accuracy due to the low statistics.

- Polarization of the photon in $B_s^0 \to \phi\gamma$. Within the SM, the production of a left-handed photon in the decay $B_s^0 \to \phi\gamma$ is suppressed by a factor m_s/m_b with respect to that of a right-handed photon, while in some theories beyond the SM

it is considerably enhanced. The $B_s^0 \to \phi\gamma$ decay has been observed at Belle [80], measuring a BR in agreement with the SM. However, no analysis addressing the polarization of the photon produced in this decay has been performed yet.

- Branching ratio of $B_s^0 \to \mu^+\mu^-$. This BR is predicted to be $(3.2\pm0.2)\times10^{-9}$ in the SM [81], while an upper limit was set to 2.4×10^{-8} at 95% CL by Tevatron [82] before the start-up of LHC. This quantity is specially sensitive to NP, expecting important departures from the SM prediction in several models. Further phenomenological considerations about this process will be given in detail in Section 1.4.

Although there are results of some of these measurements, the large experimental uncertainties do not allow to provide a final conclusion about the existence on NP in such processes. In this situation, the role of the LHCb experiment is crucial in order to achieve the necessary accuracy [83].

1.3 Effective theories

Given a process involving the fields $\phi_1(x_1)\dots\phi_n(x_n)$ in either the initial or final states, the associated correlation function is defined as

$$G^{(n)}(x_1,\dots,x_n) \equiv \langle 0|\phi_1(x_1)\dots\phi_n(x_n)|0\rangle,$$

where $|0\rangle$ is the vacuum of the interacting theory. In B decays, this function enables the perturbative description of the underlying weak interaction. However, such processes involve hadrons at least in the initial state, implying that the quark fields forming this state are not assymptotically free[2]. Under these conditions, the Lehmann-Symanzik-Zimmerman (LSZ) formalism [60] does not allow the derivation of the final quantum amplitudes using the correlation function. For this reason, the computation of such amplitudes in phenomenological analyses of B requires requires alternative methods in order to render an effective separation between the perturbative and non-perturbative effects.

1.3.1 Effective hamiltonians for B decays

B decays occur by means of weak interactions between four fermion fields. With the exception of some particular cases such as radiative B decays, these four fields belong to either the initial or final state of the process, and hence they are called *external* fields. In the case of B decays, the momentum exchange k among these states is of the order of the mass of the decaying meson, namely $k \sim$5 GeV. The masses of the intermediate weak bosons M_W and M_Z are of the order of 100 GeV, largely above the scale k. This makes possible the use of an effective interaction hamiltonian H_{eff} in which the effect of any field of mass $M \gg k$ is absorbed into some numerical

[2]This is due to the fact that the low-energy QCD interaction among the quarks forming the B hadron does not vanish as $t \to \pm\infty$.

coefficients, called *Wilson coefficients* [61][62][63]. Such fields are referred as *heavy* fields, and include not only the intermediate weak bosons, but also the top quark and any other heavy particle not yet discovered.

The required effective interaction hamiltonian H_{eff} is obtained by developing the exact expression using operator product expansions:

$$H = \sum_{n=1}^{\infty} \sum_{i_n} \frac{w_{i_n}}{M_W^{2n}} O_{i_n}^{(2n+4)}, \tag{1.1}$$

where the w_{i_n} are adimensional coefficients having a momentum-independent complex value, and $O_{i_n}^{(2n+4)}$ are Lorentz-invariant operators of dimension $(2n+4)$. Each operator has a definite Lorentz structure, which is labelled by the subindex i_n. Note that, by definition, the value of the coefficients w_{i_n} can depend only on fields entering virtually in the process, since their momentum dependence is integrated out. Any possible relation with the momentum of the external fields is thus excluded.

In general, operators entering in Eq. 1.1 depend on the fields involved in the process, either real or virtual, and their derivatives. Specifically, operators related to $n = 1$, namely $O_{i_1}^{(6)}$, depend only on the external fields. The contributions to H resulting from these operators are dominant, provided that $M_W, M_Z \gg k$, since contributions from the remaining operators, corresponding to $n > 2$, are suppressed by $\mathcal{O}\left(k^{2n}/M_W^{2n}\right)$. This condition allows to truncate the expansion in Eq. 1.1 above $n = 1$, thus obtaining the required effective interaction hamiltonian H_{eff}:

$$H_{eff} = \sum_{i_1} \frac{w_{i_1}}{M_W^2} O_{i_1}^{(6)} \equiv \sum_i c_i O_i. \tag{1.2}$$

As the operators $O_i \equiv O_{i_1}^{(6)}$ are composed only of external fields, the effect of heavy fields is exclusively contained in the coefficients c_i, which are thus identified with the Wilson coefficients defined above. Since H_{eff} represents a small contribution to the total hamiltonian, the corresponding amplitudes can be computed using the Born approximation:

$$\mathcal{M} \approx \langle X|H_{eff}|B\rangle = \sum_i c_i \langle X|O_i|B\rangle, \tag{1.3}$$

where the states $|B\rangle$ and $|X\rangle$ refer to the initial B meson and the resulting decay products respectively.

According to the discussion in the previous paragraph, heavy fields only affect the value of the Wilson coefficients c_i, while external fields are only involved in the operators O_i. This effective separation allows the use of different techniques for treating each object. On one side, the weak interaction and the respective QCD corrections determining the value of the Wilson coefficients c_i can be addressed using perturbation theory. On the other side, the low-energy QCD evolution providing the operator elements $\langle X|O_i|B\rangle$ needs to be handled by means of either non-perturbative techniques, such as lattice QCD, or experimentally, using specific processes sharing initial or final bound states.

1.3.2 The $bs\ell\ell$ transitions

Weak processes whose four external fields consist of two quarks of b and s flavor, plus two charged-conjugated leptons, can be treated using the effective theory described in Subsection 1.3.1. They represent the basis of the phenomenological analysis of some important B decays, including $B^0 \to K^{*0}\mu^+\mu^-$ and $B_s^0 \to \mu^+\mu^-$. The general expression of the effective interaction hamiltonian for these transitions is [64]:

$$H_{eff}^{(bs\ell\ell)} = \sum_{i=7}^{10} c_i O_i + c_S O_S + c_P O_P + \sum_{i=7}^{10} c_i' O_i' + c_S' O_S' + c_P' O_P'. \tag{1.4}$$

Defining $\xi \equiv \{7, \ldots, 10, S, P\}$, the symbols O_ξ, O_ξ' represent the operators, while c_ξ, c_ξ' correspond to their corresponding Wilson coefficients. Subindices refer to the different Lorentz structure of each operator. Specifically, i =10, S and P correspond to axial, scalar and pseudoscalar contributions respectively. The explicit expressions of the unprimed operators are

$$\begin{aligned}
O_7 &= m_b(\bar{s}_u \sigma_{\mu\nu} P_R b_u) F^{\mu\nu}, \\
O_8 &= m_b(\bar{s}_u T^a_{uv} \sigma_{\mu\nu} P_R b_v) G^{a\mu\nu}, \\
O_9 &= (\bar{s}_u \gamma^\mu P_L b_u)(\bar{\ell}\gamma_\mu \ell), \\
O_{10} \equiv O_A &= (\bar{s}_u \gamma^\mu P_L b_u)(\bar{\ell}\gamma_\mu \gamma_5 \ell), \\
O_S &= m_b(\bar{s}_u P_R b_u)(\bar{\ell}\ell), \\
O_P &= m_b(\bar{s}_u P_R b_u)(\bar{\ell}\gamma_5 \ell),
\end{aligned}$$

where the subindices u and v denote the quark color, and T^a_{uv} refers to the SU(3) generators. The primed partners of these operators are obtained by just changing the operator P_L by P_R, and the mass m_b by m_s. Operators O_1, O_1' to O_6, O_6' do not contribute at tree level, since their external fields only include quarks. In addition, at higher order in the coupling constant α, their effects are mainly absorbed by redefinition of the c_i, c_S and c_P and their primed partners.

1.4 The branching ratio of $B_s^0 \to \mu^+\mu^-$

The SM and a wide variety of NP theories provide predictions for BR($B_s^0 \to \mu^+\mu^-$). In all the cases, the physics governing the transition from the b and s quarks to the final dimuon pair operates at scales of at least $M_W \sim 100$ GeV. This fact enables the use of the effective theory introduced in Subsection 1.3.2 for $bs\ell\ell$ transitions. The effect of each model is then studied by considering the different contributions of the corresponding Wilson coefficients. To that purpose, it is convenient to express BR($B_s^0 \to \mu^+\mu^-$) in terms of these coefficients only. Eqs. 1.3 and 1.4 provide the amplitude for the general process $B_s^0 \to \ell\bar{\ell}$:

$$\mathcal{M} \approx \langle \ell\bar{\ell} | H_{eff}^{(bs\ell\ell)} | B_s \rangle = F_S \bar{\ell}\ell + F_P \bar{\ell}\gamma_5 \ell + F_A p^\mu_{B_s} \bar{\ell}\gamma_\mu\gamma_5 \ell, \tag{1.5}$$

$$F_{S,P} \equiv i\mathcal{F}_{B_s} M_{B_s}^2 \frac{c_{S,P} m_b - c'_{S,P} m_s}{m_b + m_s}, \tag{1.6}$$

$$F_A \equiv i\mathcal{F}_{B_s}(c_{10} - c'_{10}), \tag{1.7}$$

being p_{B_s} and $\mathcal{F}_{B_s}$ the four-momentum and the decay factor of the B_s meson respectively. The latter fully contains the effects of low-energy non-perturbative QCD on the temporal evolution of this particle. Its value has been calculated using lattice QCD, finding $\mathcal{F}_{B_s} = 238.8 \pm 9.5$ MeV [84]. Only the operators O_S, O_P, $O_A \equiv O_{10}$ and their primed partners contribute to Eq. 1.5. Terms for $i = 7$ and $i = 8$ vanish as they involve the matrix element $\langle 0|s\sigma_{\mu\nu}\bar{b}|B_s\rangle$ and its charge conjugate, which are equal to zero due to the impossibility of having antisymmetric quantities with respect to the index interchange $\mu \leftrightarrow \nu$, built only from the four-momentum of the B_s. Besides, terms with O_9 and O'_9 do not contribute since they involve hadronic matrix elements that are proportional to p_{B_s}. This four-momentum is equal to the sum of those of the two decaying leptons, leading to the contraction between such matrix elements and the leptonic vector current $\bar{\ell}\gamma_\mu\ell$ to vanish. The decay width for $B_s^0 \to \mu^+\mu^-$ is obtained by adding the squared amplitudes $\mathcal{M}^2$ of all the possible momentum and spin configurations of the final state, and considering muons as the final leptons. The average over the possible configurations of the initial state is not necessary since it consist in a single particle of spin zero. The resulting expression is

$$\Gamma(B_s^0 \to \mu^+\mu^-) = \frac{M_{B_s}}{2\pi}\sqrt{\beta}\left(\beta|F_S|^2 + |F_P + 2m_\mu F_A|^2\right), \quad \beta \equiv \left(1 - \frac{4m_\mu^2}{M_{B_s}^2}\right). \tag{1.8}$$

The factor m_μ multiplying F_A reflects the suppression of the phase space of the final state due to the conservation of the angular momentum (*helicity suppression*). Its final effect is the suppression of the axial contribution by a factor m_μ/m_{B_s}, with respect to the scalar and pseudoscalar contributions. From the decay width $\Gamma(B_s^0 \to \mu^+\mu^-)$, the BR is automatically calculated as

$$\mathrm{BR}(B_s^0 \to \mu^+\mu^-) = \frac{\Gamma(B_s^0 \to \mu^+\mu^-)}{\Gamma_{B_s}} = \tau_{B_s}\frac{M_{B_s}}{2\pi}\sqrt{\beta}\left(\beta|F_S|^2 + |F_P + 2m_\mu F_A|^2\right). \tag{1.9}$$

1.4.1 The branching ratio of $B_s^0 \to \mu^+\mu^-$ within the SM

In order to obtain $\mathrm{BR}(B_s^0 \to \mu^+\mu^-)$ in the SM, it is necessary to calculate the value of the corresponding Wilson coefficients entering in Eq. 1.9 through Eqs. 1.61.7. These coefficients are related to the high-energy regime of the process, and hence it is possible to computate them by using perturbative methods. The leading order (LO) diagrams governing this decay in the SM, shown in Fig. 1.2, involve only the weak interaction. Due to the $V - A$ nature of the weak interaction, only the

coefficient $c_A \equiv c_{10}$ is different from zero. The fermions that enter virtually in loops are either of u, c or t flavor. However, since the contribution from the u and c flavors is negligible compared to that from t, the c_{10} coefficient can be expressed in terms of $V_{ts}^* V_{tb}$ only:

$$c_{10} = \frac{G_F \alpha}{4\sqrt{2}\pi \sin^2\theta_W} V_{ts}^* V_{tb} Y(m_t/M_W). \tag{1.10}$$

The contribution from each Feynmann diagram enters this expression through the function $Y(x)$ (*Inami-Lim function*) [85][86]. The effect of the next-to-leading order (NLO) QCD corrections to the purely weak LO contribution is obtained from the ratio

$$\eta_Y \equiv \frac{Y_{NLO-QCD}(x_t)}{Y_{LO}(x_t)} = (1.026 \pm 0.006), \quad x_t = \frac{m_t}{m_W},$$

where the functions $Y_{NLO-QCD}(x)$ and $Y_{LO}(x)$ refer to the Inami-Lim functions with and without such corrections respectively. This ratio translates into an effect of 5% on the BR. Taking into account these corrections, the resulting BR is

$$\mathrm{BR}(B_s^0 \to \mu^+\mu^-) = \tau_{B_s} \frac{G_F^2 \alpha^2}{16\pi^3 \sin^4\theta_W} |V_{ts}^* V_{tb}|^2 \mathcal{F}_{B_s}^2 M_{B_s} m_\mu^2 \sqrt{\beta} Y_{NLO-QCD}^2(m_t/M_W). \tag{1.11}$$

For the sake of precision, the factor $|V_{ts}^* V_{tb}|^2 \mathcal{F}_{B_s}^2$ in this equation is extracted from the expression of the SM prediction for the mass difference between the two mass eigenstates of the B_s meson, ΔM_{B_s}:

$$\Delta M_{B_s} = \frac{G_F^2}{6\pi^2} |V_{ts}^* V_{tb}|^2 \mathcal{B}_{B_s} \mathcal{F}_{B_s}^2 M_{B_s} M_W S(m_t/M_W), \tag{1.12}$$

where $\mathcal{B}_{B_s}$ is the bag parameter of the B_s meson, which, analogously to $\mathcal{F}_{B_s}$, is related to the effects of low-energy QCD on the temporal evolution of this particle. In this case, calculations using lattice QCD result in $\mathcal{B}_{B_s} = 1.33 \pm 0.06$ [84]. The value of ΔM_{B_s} is available from experimental observations, finding $\Delta M_{B_s} = 17.77 \pm 0.12$ ps [66]. Eq. 1.12 allows to write Eq. 1.11 as

$$\mathrm{BR}(B_s^0 \to \mu^+\mu^-) = \frac{3}{8\pi} \frac{\alpha^2}{\sin^4\theta_W} \left(\frac{m_\mu}{M_W}\right)^2 \frac{\tau_{B_s}}{\mathcal{B}_{B_s}} \sqrt{\beta} \frac{Y^2(m_t/M_W)}{S(m_t/M_W)} \Delta M_{B_s}. \tag{1.13}$$

Note that the factor $|V_{ts}^* V_{tb}|^2 \mathcal{F}_{B_s}^2$ has been replaced by the more accurate quantities ΔM_{B_s} and $\mathcal{B}_{B_s}$. Using Eq. 1.13, the SM prediction of $\mathrm{BR}(B_s^0 \to \mu^+\mu^-)$ becomes [81]

$$\mathrm{BR}(B_s^0 \to \mu^+\mu^-) = (3.2 \pm 0.2) \times 10^{-9}, \tag{1.14}$$

which has an uncertainty of 6%.

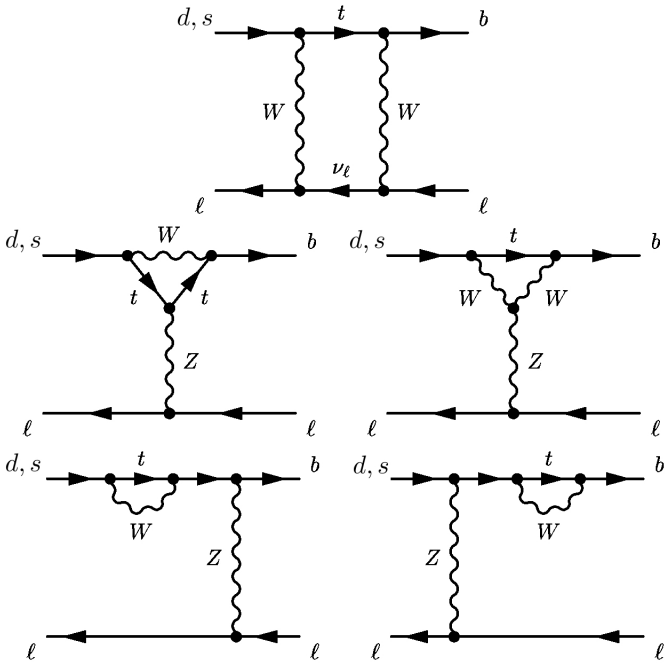

Figure 1.2: Main Feynmann diagrams contributing to the decays $B^0_{(s)} \to \ell\bar{\ell}$ in the SM.

1.4.2 Contributions from extensions of the SM

The fact that only loop diagrams govern the $B^0_s \to \mu^+\mu^-$ decay in the SM, together with the helicity suppression of these contributions, make this channel specially sensitive to indirect effects of NP. These potential effects are already constrained by current precision measurements of electroweak parameters, such as the anomalous magnetic moment of the muon $g_\mu - 2$, and direct searches of heavy particles, including that of the Higgs boson. Specifically, these observations are only compatible with values of c_A close to that of the SM. This automatically rejects any significant effect in $B^0_s \to \mu^+\mu^-$ coming from TC2, since the main contributions in this case enter through c_A [92]. Besides, the sensitivity of $B^0_s \to \mu^+\mu^-$ to SUSY with AMSB has been drastically reduced by the constraints imposed by the current measurements of $g_\mu - 2$ and $b \to s\gamma$ to the parameter space of this model. In spite of these constrains, several models predict important departures from the SM while still being compatible with the current observations:

- 2HDM-II. In this case, c_S and c_P receive contributions scaling with $\tan^2\beta$. The prediction for the BR as a function of the parameter M_{H^+}, and for different values of $\tan\beta$, is shown in Fig. 1.3 [94]. Values of M_{H^+} below 295 GeV are excluded by the measurement of $b \to s\gamma$ [95]. In the remaining parameter space, the value of $\mathrm{BR}(B^0_s \to \mu^+\mu^-)$ has significant enhancements for values of $\tan\beta >$

60.

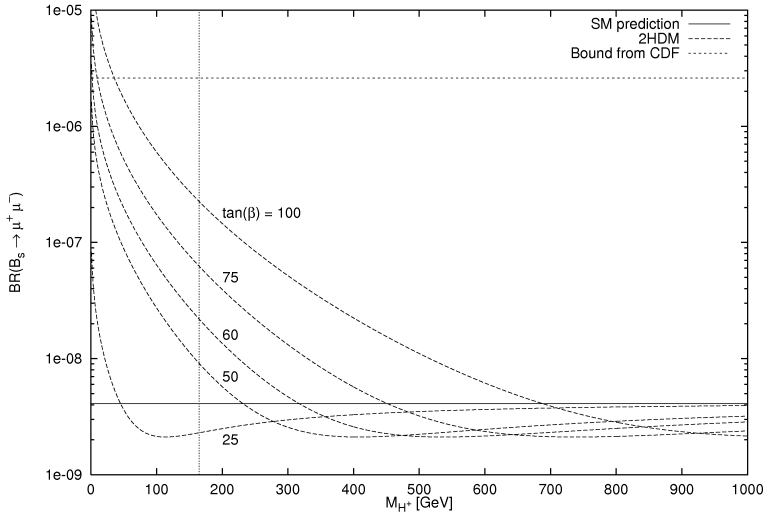

Figure 1.3: The value of BR($B_s^0 \to \mu^+\mu^-$) in 2HDMII, as a function of the parameter M_{H^+}.

- MSSM. As for the 2HDM-II, the additional diagrams due to MSSM affecting BR($B_s^0 \to \mu^+\mu^-$) contribute with terms proportional to powers of tan β. Considering the constraints on the contribution of NP to c_A, significant departures from the SM prediction are only possible for large values of tan β. Specific versions of the MSSM have been studied in detail, showing important effect on BR($B_s^0 \to \mu^+\mu^-$).

  The value of BR($B_s^0 \to \mu^+\mu^-$) has been studied within the CMSSM and the MSSM-NUHM by means of a likelihood built according to the constraints set by electroweak precision measurements, B physics data and cosmological observations [96]. The result is shown in Fig. 1.4. No substantial departures from the SM prediction are expected for the CMSSM. On the contrary, for the MSSM-NUHM, enhancements up to 40% are clearly compatible with the empirical input of the likelihood. Furthermore, there are non-negligible possibilities of having larger departures within this framework. This difference with respect to the CMSSM is due to the freedom in the choice of the Higgs masses in the MSSM-NUHM.

  The effect of R-parity violation (RPV) in the value of BR($B_s^0 \to \mu^+\mu^-$) has been considered in the context of the minimal version of SUGRA (MSUGRA). The assumption of RPV implies new interaction terms in the lagrangian of the model [97]. The new couplings can be grouped in four classes, with their respective constants referred as ϵ_i, λ_{ijk}, λ'_{ijk} and λ''_{ijk}, where the subindices denote the fermion family. The first three constants are related with the violation of the lepton number, while the last constant is involved in the violation of the baryon number. The predictions for BR($B_s^0 \to \mu^+\mu^-$) under these conditions are shown

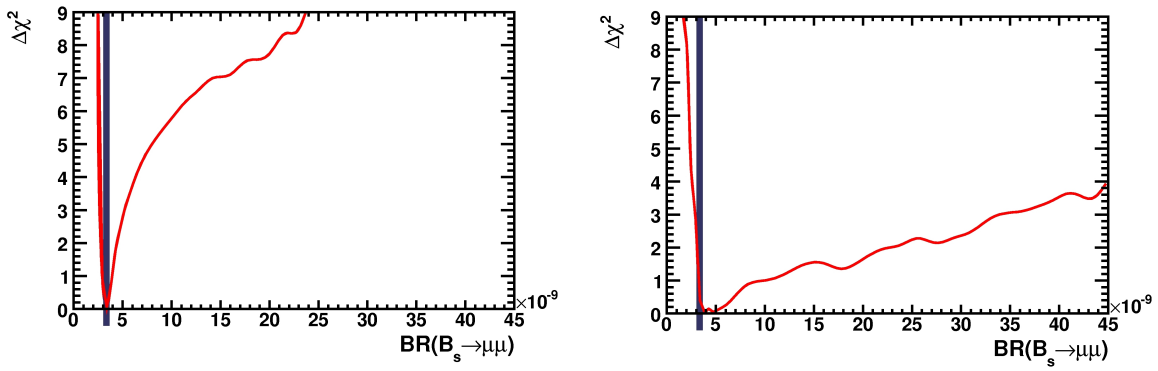

Figure 1.4: Likelihood for the value of BR($B_s^0 \to \mu^+\mu^-$) in the CMSSM (left) and in the MSSM-NUHM (right), considering the constraints set by electroweak precision measurements, B physics data and cosmological observations. The vertical line shows the SM prediction.

in Fig. 1.5 [98]. They include the constraints coming from electroweak precision measurements, B physics data, direct searches of the Higgs boson, and tests of the R-parity conservation. In addition, it assumes universal values for both λ_{ijk} and λ'_{ijk} constants at the electroweak scale, denoted by λ and λ' respectively. The predictions are classified according to the sign of the Higgsino mixing parameter μ. Negative values of this parameter lead to a moderate enhancement with respect the SM prediction, independent of the product $\lambda\lambda'$. For positive values, considerable departures from the SM prediction are possible, specially for $\lambda\lambda' > 10^{-7}$.

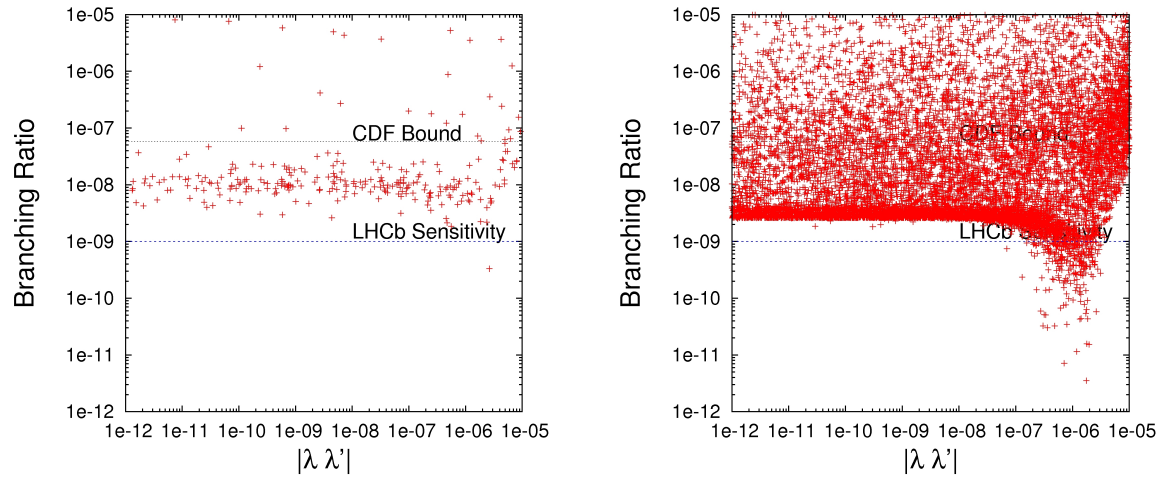

Figure 1.5: The value of BR($B_s^0 \to \mu^+\mu^-$) in MSUGRA, assuming RPV, for negative (left) and positive (right) values of the Higgsino mixing parameter μ.

- Warped extra dimensions (RS models). The simplest realization of the RS theory considers a single finite extra dimension having an exponential warping along. Within this setup, the value of BR($B_s^0 \to \mu^+\mu^-$) has been studied together with that of $B^0 \to \mu^+\mu^-$ and $B^0 \to X_s\nu\nu$ [99]. The results are summarized in Fig. 1.6. Important departures from the SM expectation are predicted for BR($B_s^0 \to \mu^+\mu^-$), that can reach a factor two for several configurations of the parameter space. Besides, these departures are expected to be related only to those arising in the BR of $B^0 \to X_s\nu\nu$.

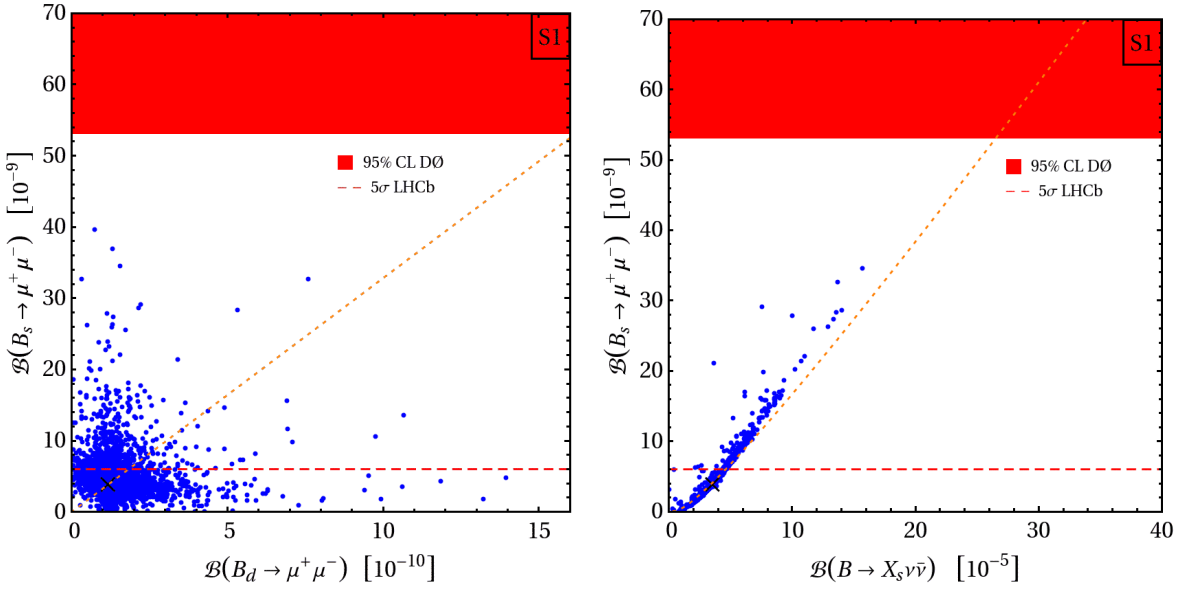

Figure 1.6: Relation between BRs in RS models: $B_s^0 \to \mu^+\mu^-$ against $B^0 \to \mu^+\mu^-$ (left), and $B_s^0 \to \mu^+\mu^-$ against $B^0 \to X_s\nu\nu$ (right).

Chapter 2

The LHCb experiment

This chapter describes the LHCb detector [100]. Section 2.1 discusses the requirements for a B physics experiment at the LHC. Section 2.2 contains the detailed description of each subdetector. Sections 2.3 and 2.4 explain the LHCb trigger and the detector readout respectively. Finally, Section 2.5 discusses the offline data processing.

Throughout this chapter, coordinates are expressed with respect to the following axes: the positive x axis is directed towards the center of the LHC ring; the positive y axis is directed upwards; and the positive z axis is parallel to the incoming beams, and arranged in order to form a left-handed system of coordinates together with the previous axis. The latter axis will be also referred as *beam axis*. Directions defined by vectors having either negative or positive values of their z coordinate are labelled as either *backward* or *forward* respectively. Besides, the origin of this reference frame is situated in the interaction point.

2.1 Requirements for a B physics experiment at the LHC

At the LHC, the unprecedented ammount of B mesons will allow to perform crucial measurements in this sector. The LHCb detector has been specially conceived for the analysis of B decays occurring in the interaction of two proton bunches at the nominal energy of 14 TeV.

The production of B mesons at the LHC is dominated by the production of $b\bar{b}$ quark pairs in hard QCD interactions. The expected angular distribution of these quarks is basically restricted to the forward and backward regions adjacent to the beam axis, as shown in Fig. 2.1. Note that, in addition, the two quarks tend to follow close directions in space. Due to the large average momentum of the resulting B mesons with respect to their mass, the angular distribution of the B decay products is similar to that of the original quarks. This feature, together with budget constraints, motivated the acceptance of LHCb to comprise only forward directions satisfying $0.01 < \theta_{xz} < 0.25$ rad and $0.01 < \theta_{yz} < 0.30$ rad, where θ_{xz} and θ_{yz} refer to the angle with respect to the xz and yz planes respectively. It is expected

that approximately 20% of the B meson decays will be fully contained within these acceptance limits.

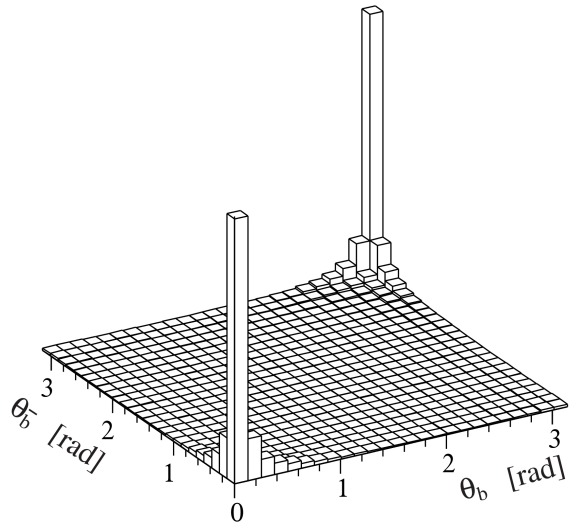

Figure 2.1: Angular distribution of the $b\bar{b}$ quarks produced in LHC collisions under nominal conditions, with respect to the beam axis.

The mean lifetimes of the B mesons are of the order of 1 ps. This relatively long lifetime, together with the large ratio between the average momentum and the mass of the B mesons produced at LHCb, leads to decay vertices with a mean separation of the order of 1 cm from the primary vertex. The presence of displaced decay vertices is a clear signature of the B decays, and its identification is an essential component of B physics experiments. For this reason, events having multiple primary interactions introduce severe complications to the analysis of B mesons. This fact has motivated the decision of having only 0.7 collisions per bunch crossing on average, resulting in an instantaneous luminosity equal to 2×10^{32} cm^{-2}s^{-1}. Regarding these conditions, the integrated luminosity is expected to be 2 fb^{-1} after one nominal year of data taking, translating approximately into 10^{12} $b\bar{b}$ pairs produced within the LHCb acceptance during such time period.

The general structure of the LHCb detector is displayed in Fig. 2.2. The subdetectors and their characteristics are determined by the properties of the final states of the B meson decays, as discussed in the following subsections.

2.1.1 Vertexing

As discussed above, the detection of displaced decay vertices is crucial for identifying B decays. In addition, a precise knowledge of the separation between the primary and secondary vertices is necessary in studies involving a time-dependent analysis.

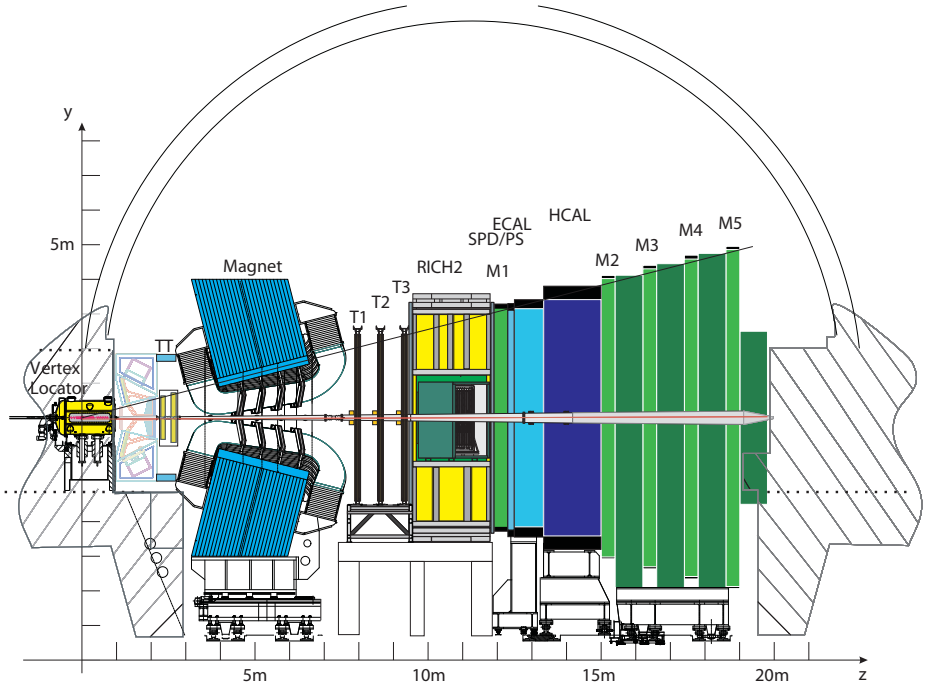

Figure 2.2: Layout of the LHCb detector, showing the different subdetectors.

Some important examples are the measurement of the angle γ using $B^0 \to D^-\pi^+$ or $B_s^0 \to D_s^\pm K^\mp$, and the determination of the CP violating phase ϕ_s of $B_s^0 \to J/\psi\phi$. In order to have accurate measurements of the position of primary and secondary vertices, a tracking detector is located very close to the interaction point, the Vertex Locator (*VeLo*).

2.1.2 Momentum and energy measurement

Two important quantities related to the kinematics of the products of the B decays allow for a considerable background rejection, namely the transverse momentum (p_T) of individual tracks, or equivalently their transverse energy (E_T), and the invariant mass of combinations of tracks. The former rejects background from low-energy interactions, while the latter discriminates combinatorial background. For these reasons, systems measuring the momentum or the energy of the decay products are crucial. Depending on the purpose, this information is provided by either mass spectrometry and calorimetry. The former involves the VeLo and the tracking system: the Tracker Turicensis (TT), Inner Tracker (IT) and Outer Tracker (OT). They are devoted to the precise measurement of the momentum of charged particles. The complexity of the information provided by such subdetectors restricts their use to the offline analysis and the last stages of the trigger. The latter uses three calorimeters: the Pre-Shower (PS), the Electromagnetic Calorimeter ($ECAL$) and the Hadronic

Calorimeter (*HCAL*). Although their accuracy is lower than that of the tracking detectors, they enable the fast measurement of the transverse energy in the stages of the trigger where the use of mass spectrometry is unfeasible. Besides, calorimeters provide the measurement of the energy of neutral particles for both the trigger and some offline analyses as those involving radiative B decays.

2.1.3 Particle identification (PID)

As discussed in Section 1.2, the most decisive measurements in the B sector contain either photons, muons, pions or kaons in the final state. For example, the angular analysis of $B^0 \to K^{*0}\mu^+\mu^-$ and the determination of $\mathrm{BR}(B_s^0 \to \mu^+\mu^-)$ rely on the muon identification, while the use of $B \to h^+h'^-$ for either measure the angle γ or control the analysis of $B_s^0 \to \mu^+\mu^-$ require the separation between pions and kaons. In addition, there are some important semileptonic decays involving electrons, such as $B^+ \to K^{*+}e^+e^-$. The proper identification of these particles is achieved by combining the information coming from the following subdetectors:

- Ring Imaging Cerenkov detectors (RICH). These subdetectors use the dependence of the radius of a Cerenkov ring with the velocity of the particle that produces it. In combination with the momentum information provided by the tracking system, this information allows to extract the mass of the particle. The main application of this component is the separation among pions, kaons and protons.
- Calorimeters. In addition to the three calorimeters mentioned in the previous subsection, another device is used for PID, called Scintillator Pad Detector (*SPD*). The SPD allows for discriminate between photons and electrons, while the PS separates between electromagnetic and hadronic particles.
- Muon system. The bulk of particles remaining after the calorimeters are muons, and hence they are recognized by means of a system of trackers located after these subdetectors.

2.2 LHCb subdetectors

The LHC*b* detector is composed of a vertex tracker (VeLo), Cerenkov ring detectors (RICH), a tracking system (magnet plus TT, IT and OT), calorimeters (SPD, PS, ECAL and HCAL), and muon detectors (muon system). These subdetectors are described in the following, ordered according to their distance to the interaction point.

2.2.1 Vertex Locator (VeLo)

The VeLo measures the trajectories of charged particles in the region adjacent to the interaction point. It is used for detect the primary and secondary vertices, as well as

to construct seeds for the track reconstruction. It consist of two arrays of 21 stations each, covering respectively the two sides of the interaction region, as shown in Fig. 2.3. Furthermore, each array contains two additional stations situated backwards the interaction point, forming the so-called *Pile-Up* system. It is used in the first stages of the trigger for a fast estimation of the number of primary interactions. Each VeLo station has two silicon sensors, called r and ϕ sensors respectively. For a given particle crossing a VeLo station, r sensors measure the radial coordinate, while ϕ sensors measure an analogue of the azimuthal coordinate. In addition, the z coordinate is known from the position of the sensors. The layout of the r and ϕ sensors is displayed in Fig. 2.4. The strip pitch of these sensors increases linearly with the radial coordinate. For r sensors it ranges from 40 μm at r =8.2 mm, to 101.6 μm at $r = 41.9$ mm. For ϕ sensors it depends on the region, varying from 37.7 μm at $r = 8.2$ mm, to 79.5 μm at $r = 17.2$ mm, and from 39.8 μm at $r = 17.2$ mm, to 96.9 μm at $r = 41.9$ mm.

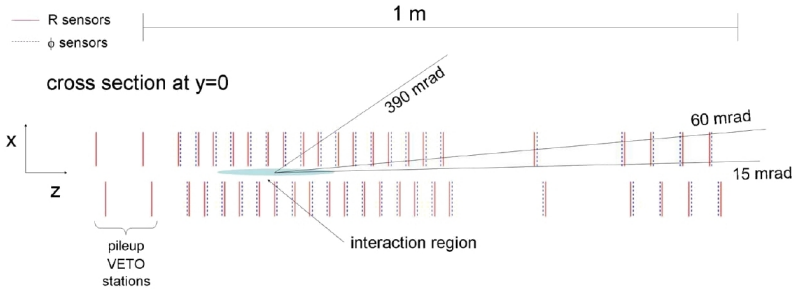

Figure 2.3: Arrangement of the silicon stations in the VeLo.

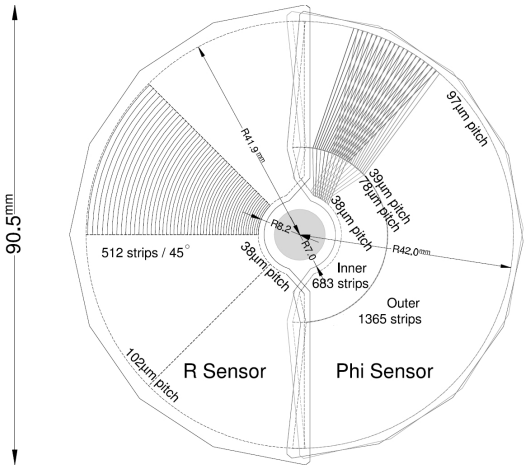

Figure 2.4: Characteristics of the VeLo stations. The disposition of the r and ϕ angles is shown on the left and right halves respectively.

2.2.2 RICH detectors

The RICH system provides information for the identification of charged particles by means of the Cerenkov rings produced by their passage through certain radiators. It is composed of two similar stations called RICH1 and RICH2, located before and after the tracking system respectively. Both RICH1 and RICH2 are shown in Fig. 2.6. Each station is optimized for complementary momentum ranges, depending on the radiatior material, see Fig. 2.5. RICH1 is designed for momenta below 60 GeV by using silicon aerogel and gaseous C_4F_{10}, while RICH2 covers the range between 15 and 100 GeV by using CF_4. The Cerenkov light is directed outside the LHCb acceptance by means of a set of mirrors, and then it is collected by an hexagonal array of pixel hybrid photodetectors (HPD). The packing factor of this array is 0.64. In turn, each HPD has a granularity of 2.5×2.5 cm^2.

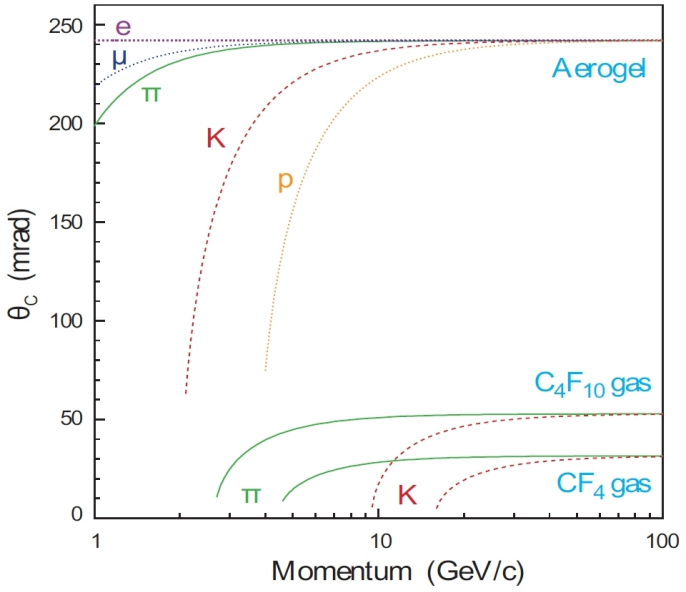

Figure 2.5: Cerenkov angle as a function of the momentum, for different particles and materials.

2.2.3 Tracking system

The tracking system measures the trajectory of charged particles under the influence of a vertical magnetic field, enabling the determination of their momentum by mass spectrometry. This magnetic field is supplied by the Magnet, and has an integrated strength of 4 Tm. It can be directed either downwards or upwards, in order to control the systematic effects of the detector in the measurement of CP assymmetries.

The tracking system includes the Tracker Turicensis (TT), Inner Tracker (IT) and Outer Tracker (OT), which are displayed in Fig. 2.7. These subdetectors are made of layers providing either the x, u or v coordinate of the position of particles,

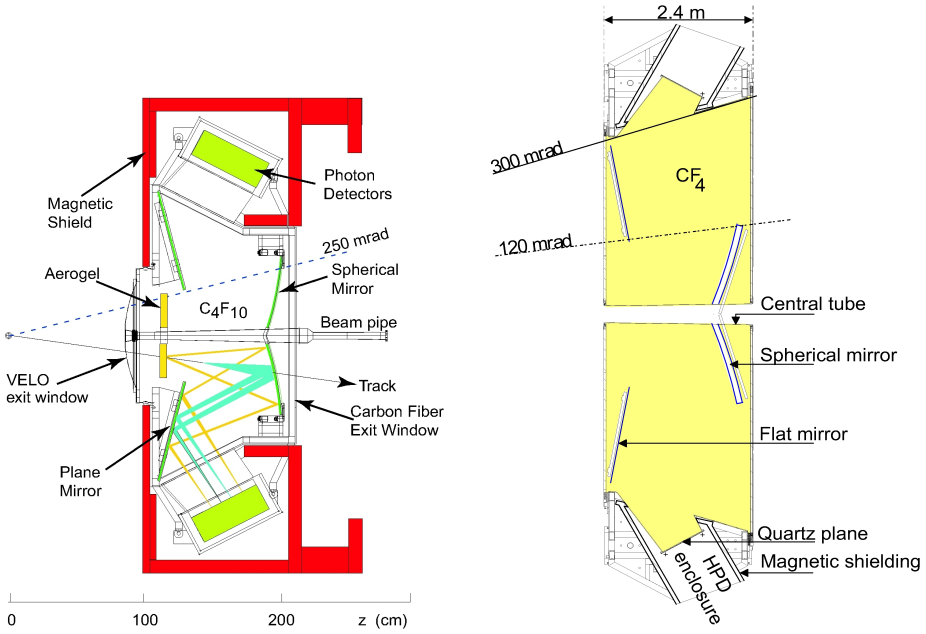

Figure 2.6: Side view of RICH1 (left) and top view of RICH2 (right).

where the u and v axes are defined as the result of rotating the x axis by 5 and -5 degrees respectively around the beam line. This is illustrated in Fig. 2.8 by showing the arrangement of the TT layers. The choice of the x, u and v axes is motivated by the fact that the bending of the trajectory of charged particles by the magnetic field occurs in the horizontal plane. As the u and v coordinates depend slightly on the y coordinate, they allow for the tridimensional reconstruction of the tracks. More details about the subdetectors composing the tracking system are given in the following.

- Tracker Turicensis (TT). It is situated in front of the Magnet. It consist of four layers of silicon sensors, grouped in two stations, one with two layers measuring the x and u coordinate, plus another one with two layers measuring the v and x coordinate. The strip pitch of the silicon sensors is 183 μm.

- Inner Tracker (IT) and Outer Tracker (OT). These devices are also known as *T-stations*. They are situated behind the Magnet, covering complementary areas of the LHCb acceptance. The IT consist of silicon sensors covering a small cross-shaped area around the beam line, while the OT uses drift tubes extended over the remaining area of the LHCb acceptance. This setup arise from the necessity of ensuring a low occupancy around the beam axis, where the flux of particles coming from proton-proton interactions is large, while fitting to a moderate cost. Both trackers are composed of three stations situated along the beam axis, and each station in turn is made up of four layers providing consecutive measures of the x, u, v and x coordinates. The strip pitch of the IT silicon sensors is 198

μm. In the case of OT, each layer consist of two staggered sublayers of drift tubes, with a pitch of 5.25 mm. In turn, tubes have a drift-coordinate resolution of $\sim$200 μm.

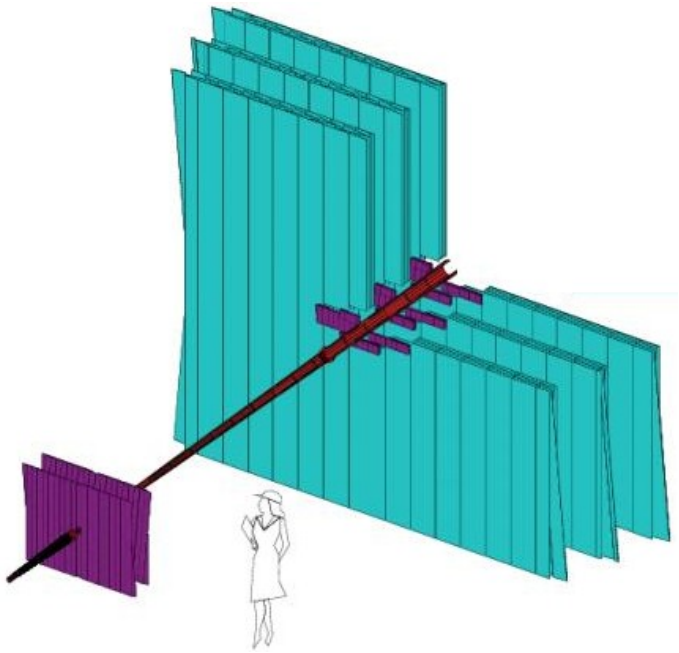

Figure 2.7: Arrangement of the tracking stations, with one quadrant of the IT and OT removed for viewing purposes. TT and IT are displayed in purple, while OT is represented in blue. The beam pipe is shown for reference.

2.2.4 Calorimeters

The measurement of the energy deposited in the calorimeters by both charged and neutral particles gives an estimation of their energy and provides information for their identification. Four calorimeters are used in the LHCb detector, which are the following, according to their order in the z coordinate:

- Scintillator Pad Detector (SPD). It consist of a single layer of scintillator cells situated in front of the remaining calorimeters. Each cell provides a binary outcome according to the deposited energy, which is used to distinguish between photons and electrons.
- Pre-shower (PS) detector. It uses a layer of scintillator cells identical to that of the SPD, which measure the energy of the particle shower originated by a 15 mm thick lead barrier placed just in front of it, corresponding to 2.5 radiation lengths. This information is used to distinguish photons and electrons from hadrons.

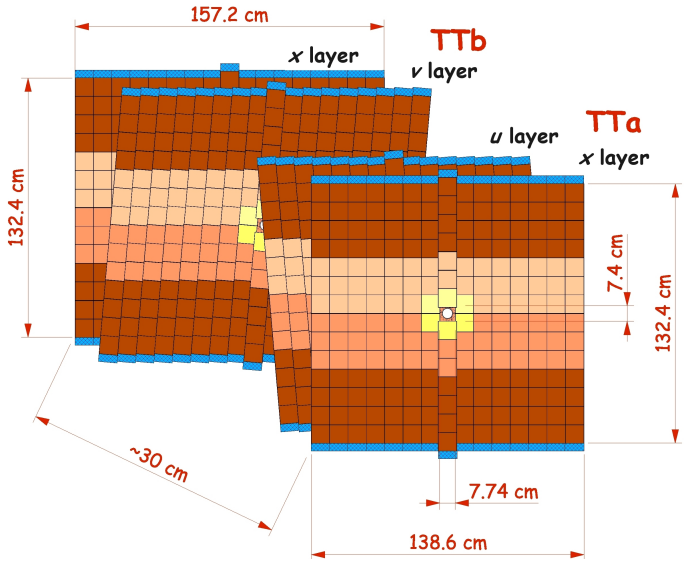

Figure 2.8: Arrangement of the four layers of the TT, illustrating the definition of the x, u and v coordinates. The IT and OT stations have analogous structures.

- Electromagnetic calorimeter (ECAL). ECAL cells consist in stacks of alternating lead-scintillator plates, representing 25 radiation lenghts in total. The energy of the particle shower is measured in combination with the PS, providing a resolution of $\sigma_E/E = 10\%/\sqrt{E} \oplus 1\%$, with E expressed in GeV.
- Hadronic calorimeter (HCAL). The HCAL is composed of iron and scintillator plates arranged parallel to the yz plane, adding up to 5.6 interaction lengths. The resolution in the measure of the shower energy is $\sigma_E/E = 70\%/\sqrt{E} \oplus 9\%$, with E expressed in GeV.

The cell patterns of the different calorimeters are shown in Fig. 2.9. For the SPD, PS and ECAL they match projectively with respect to the interaction point. Their occupancies are kept low by defining three areas of different granularity according to the distance to the beam axis, called *inner*, *middle* and *outer* regions. For ECAL, such granularities are 40.4 mm, 60.6 mm and 121.2 mm respectively, scaling accordingly for the SPD and the PS. In HCAL, the width of hadronic showers enable the use of wider cells than those in ECAL. Only two different regions are defined in this case, namely the inner and outer regions, having granularities equal to 131.3 mm and 262.6 mm respectively.

2.2.5 Muon system

The identification of muons coming from the interaction point is handled by the muon system. This subdetector is made up of five stations situated along the z axis, identified respectively by the name succession from M1 to M5. M1 is located in

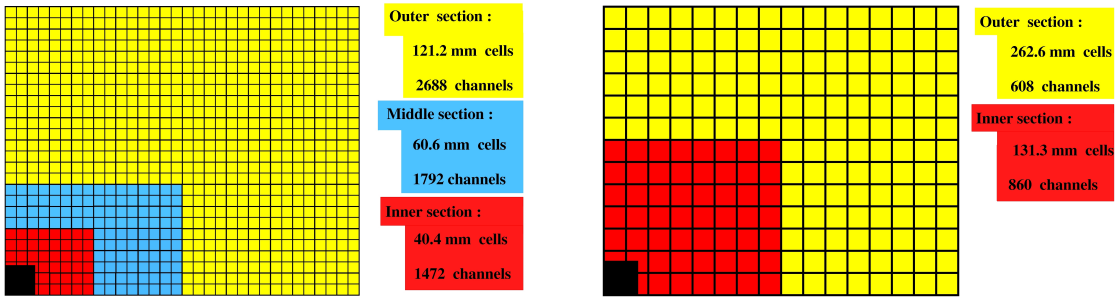

Figure 2.9: Cell pattern of SPD, PS and ECAL (left), and of HCAL (right).

front of the calorimeters, while the rest are situated behind them, separated each other by 80 cm thick iron filters. An additional iron filter is placed behind the M5 in order to prevent contamination from low energy backward muons from beam backgrounds. Stations are segmented into rectangular pads, as shown in Fig. 2.10. The pad patterns of the different muon stations match projectively with respect the interaction point. In order to ensure low occupancies, four different granularities are used depending on the distance to the beam axis, as summarized in Table 2.1. The muon system use multiwire proportional chambers (MWPC) except in the innermost region R1 of station M1, where gas electron multiplier (GEM) detectors are necessary in order to handle the high flux of particles.

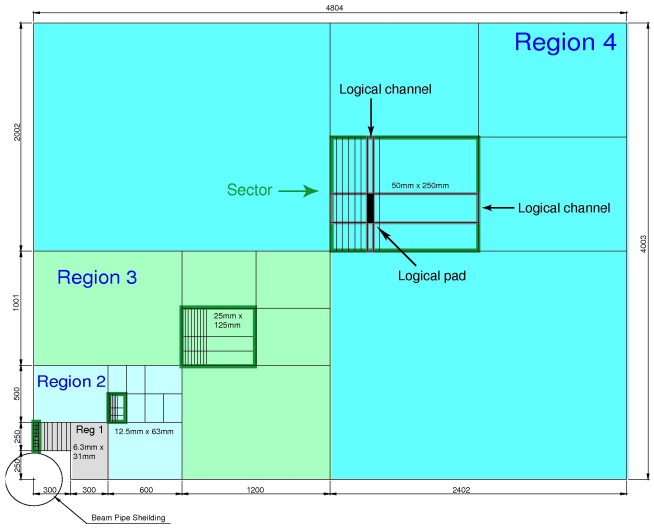

Figure 2.10: Layout of one quadrant of M2, showing the four different regions, and the definition of logical pad.

	M1	M2	M3	M4	M5
R1	1×2.5	0.63×3.1	0.67×3.4	2.9×3.6	3.1×3.9
R2	2×5	1.25×6.3	1.35×6.8	5.8×7.3	6.2×7.7
R3	4×10	2.5×12.5	2.7×13.5	11.6×14.5	12.4×15.5
R4	8×20	5×25	5.4×27	23.1×29	24.8×30.9

Table 2.1: Granularity in the different regions of the muon stations, expressed in cm.

2.3 Trigger

The LHCb trigger reduces the event rate from the 10 MHz of bunch crossings with visible interactions at nominal LHC conditions, down to the 2 kHz limit set by the available storage resources. Besides, the available computing resources dedicated to the LHCb trigger limit the reading of the full detector to a rate of at most 1 MHz. This fact has motivated the design of the trigger as two consecutive levels, called Level-0 (L0) trigger and High Level Trigger (HLT) respectively. The L0 trigger is implemented on custom hardware devices, and reduces the initial event rate down to 1 MHz by using only the Pile-Up system, the calorimeters and the muon system. The HLT is implemented on software running in a cumputer farm, and performs the remaining reduction of the event rate by accessing to the information of the entire detector.

2.3.1 Level-0 trigger

The L0 trigger operates in coincidence with the bunch crossing rate of LHC of 40 MHz. It is implemented in custom electronic devices, using only the following information:

- Transverse energy of candidates in the calorimeters. These candidates are built from clusters of 2×2 cells. This size is optimal to contain most of the energy of the shower of a given particle, while preventing the overlap with showers of other particles. Each candidate is classified as either a photon, electron or hadron according to the information provided by the SPD and the PS.
- Transverse momentum of candidates in the muon system. Using the projectivity of the muon system, muon candidates require hits in the five muon stations, contained within the same projective region. The p_T of the candidate is measured with a resolution of about 20%, by using only the information of the hit position in the first and second layer of the muon system.
- Number of primary interactions estimated by the Pile-Up system.
- Charged track multiplicity estimated by the SPD.

The component of the L0 trigger combining this information in order to either accept or reject events is called L0 Decision Unit (L0DU). Its decision is based on

the presence of certain signatures of B decays. Each signature is addressed by means of a set of specific requirements called *line*. These requirements imply the presence of high p_T or E_T candidates, and in some cases include requirements on global event properties as the estimations of the number of primary interactions or the charged track multiplicity. The configuration of the L0DU is defined by the active lines. For nominal running conditions, the standard configuration include three lines based on single particles with high p_T, being identified in each case as either photons, electrons or hadrons respectively, and two lines based on muons of high p_T, either single or forming a pair. Events are accepted if the requirements of at least one active line are satisfied.

2.3.2 High Level Trigger

The HLT receives events from the L0 trigger at a rate of 1 MHz. It is implemented in software running in a farm of commercial computers, thus supplying the HLT with a high flexibility. The computing resources of the HLT consist in a farm of 2000 CPUs placed in the LHCb pit, called *Event Filter Farm* (EFF).

HLT algorithms have access to the information of the full detector, which is added gradually in successive steps. At each step, events are required to fulfill certain requirements, otherwise they are automatically rejected. On this basis, the HLT is subdivided in two main stages: the HLT1, where only the information from the VeLo and the tracking stations is used to refine L0 candidates and some possible companions; and the HLT2, which performs a full reconstruction of the event. Although the general structure of both stages is well stablished, their specific implementation will evolve according to the data taking conditions. The intermediate rate between HLT1 and HLT2 is not completely fixed, but it is regarded to be of the order of 10 kHz.

HLT1

The HLT1 is composed of independent algorithms called *alleys*. Each alley starts with the candidates of a definite L0 line, and operates by supplying it with additional information in subsequent steps. At each step, the information that is added to the candidate must satisfy certain requirements, otherwise the candidate is rejected by the alley. The use of the VeLo and the tracking system enables an improved measurement of the momentum of candidates, and a fast estimation of their impact parameter with respect to the primary vertices. Besides, some alleys search for additional objects accompanying the candidate, involving requirements as the distance of closest approach between two tracks, or the pointing of the sum of two momenta to the primary vertex. Events are accepted if at least one alley is providing candidates. The current structure of the HLT1 involves few tens of alleys. Among them, there are groups of alleys dedicated to muonic and hadronic decays.

HLT2

The HLT2 is able to perform a full track reconstruction similar to that in the offline processing, and search for common intermediate resonances as $K^{*0} \to K^{\pm}\pi^{\mp}$, $\phi \to K^+K^-$ or $J/\psi \to \mu^+\mu^-$. This information is used to either accept or reject events according to the requirements of several selections. Such selections are oriented to either exclusive or inclusive processes, and hence they are named *exclusive* and *inclusive* selections respectively. Their design is substantially influenced by the offline analyses, specially in the case of exclusive selections. As a consequence, the effect of the HLT2 on signal will be practically negligible for most of the physics analysis at LHC*b*.

2.4 Online System

The data acquisition and the control of the detector is performed by the Online System. It consist of three parts:

- Data Acquisition (DAQ) system. It is the infrastructure allowing for the data flow from the front-end electronics of each subdetector to permanent storage. In case of positive L0 trigger decision, each subdetector is read by means of standarized electronic boards (TELL1), with the exception of the RICH detectors that use a specific board having the same functionality. Data are sent to the EFF through the readout network. The EFF provides the computing resources for running the HLT, and has a buffer space of 40 TB. Finally, events passing the HLT are sent to permanent storage.
- Timing and Fast Control (TFC) system. It handles the automatic signals that control the data flow between the detector and the EFF. The main element is the Readout Supervisor. It synchronizes the L0 trigger decision with the LHC clock, and provides the final decision of either accepting or rejecting an event, based on the result of the L0DU, the calibration requirements of the subdetectors and the load of the readout network.
- Experiment Control System (ECS). It enables the control and monitoring of the operation of the LHC*b*, handling not only the detector parameters but also the trigger and the DAQ and TFC systems.

2.5 Storage and processing of data

The LHC*b* experiment uses the computing resources provided by the Grid infrastucture. It is organized as a hyerarchy of computing sites called *Tiers*. The central site is Tier-0, placed at CERN. Next, there are seven Tier-1, each one located in a different european country. The last level consist in many Tier-2 spread across all the countries adscribed to CERN.

Raw data sent from the LHCb pit is received and saved by the CERN Tier-0. In addition, a copy is transferred to each Tier-1 for permanent archiving. Then, the processing of the data is distributed by CERN Tier-0 among itself and all the Tier-1. This leads to the generation of new data, called *Data Summary Tape* (DST), which is in turn copied to Tier-0 and each Tier-1 for permanent storage. Besides, data may be subjected to further reprocessings due to various reasons, such as improvements in the software versions or in the calibration and alignment of the detector. In these cases, the procedure is analogous to that of the first processing.

The processing of data consist in two steps: the event reconstruction and the stripping, which are described in the following.

2.5.1 Event reconstruction

In the event reconstruction, the physical information of the event is extracted from the raw data. Energy depositions in the calorimeters are grouped in clusters associated to particle showers, and hits in the tracking stations are combined to form tracks with definite position and momentum. In addition, the PID information provided by the respective subdetectors is processed and supplied to these tracks, turning them into new objects called *protoparticles*. Note that, although protoparticles contain PID information, its type, and hence its mass, remains undefined. The issues related to the event reconstruction that are necessary for the analysis of $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ are discussed in the following.

Track reconstruction

Tracks are built from segments constructed in either the VeLo or the T-stations, called *seeds*, where particles follow approximately straight lines due to the low magnetic field. The reconstruction algorithms operate by matching seeds with hits in the remaining tracking subdetectors, or by associating VeLo seeds with those in the T-stations. The resulting trajectories are refitted using a Kalman filter, which accounts for effects due to the interaction with the material, such as the multiple scattering and the energy loss.

Tracks traversing the full tracking system, namely the VeLo, the TT and the T-stations, are called *long tracks*, and have the most precise momentum determination. They represent the most important sample of tracks in the analysis of B decays.

For tracks with momentum above 10 GeV, the algorithms used for long track reconstruction at LHCb have an efficiency of 94% for a ghost rate of 9%. However, most of the ghost tracks have low reconstructed p_T. The resolution on the momentum is shown in Fig. 2.11. It ranges from 0.35% to 0.55%, for low and high momenta respectively. The resolution on the impact parameter (IP) with respect to primary vertices is shown in Fig. 2.11 as a function of $1/p_T$. It is dominated by multiple scattering in the VeLo.

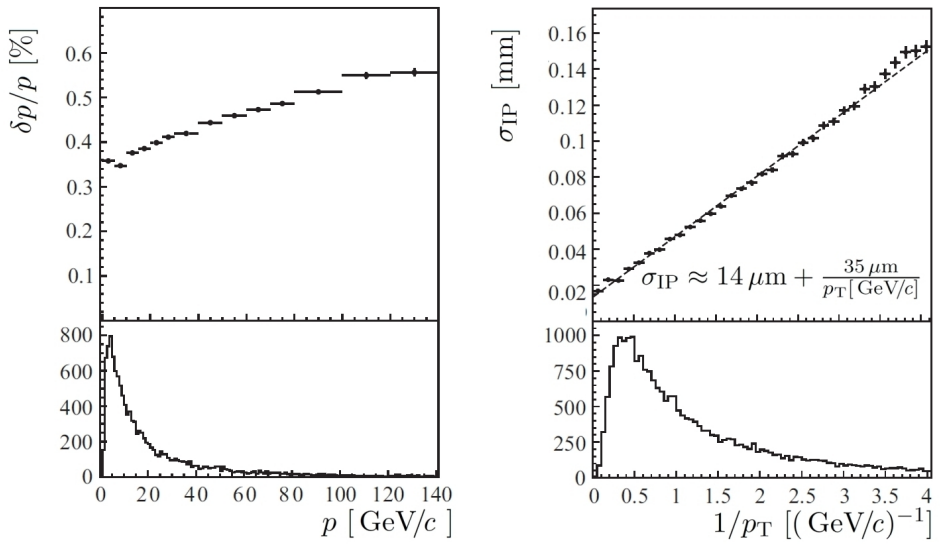

Figure 2.11: Momentum resolution (left) and impact parameter resolution (right) at the production vertex of the track. In each case, the corresponding spectrum is shown below.

Muon identification

Muons are identified by extrapolating tracks into muon stations, and then requiring hits in these stations within a certain *field of interest* (FOI) around the extrapolated trajectory. Both the requirements on the number of hits and the FOI depend on the momentum of the track.

Most muons with momentum below 3 GeV are unable to reach the muon chambers situated after the calorimeters. Above this threshold, the efficiency of the muon identification is 93% for a misidentification rate of 3% due to pions. However, this misidentification rate can be lowered to 1% by using the information from the RICH, while maintaining the efficiency at 93%.

Hadron separation

Tracks not identified as either electrons or muons are considered hadrons. The separation between pions, kaons and protons is accomplished by using the information provided by the RICH subdetectors. Tracks are given different PID hypotheses, and in each case the compatibility with the actual measurements provided from the RICH subdetectors is quantified by means of a likelihood. The best PID assignments are thus obtained by maximizing this likelihood. Note that this procedure considers all the tracks in the event simultaneously, and hence it is called *global pattern recognition*. Based in the best PID choices, the information that is passed to the protoparticles is the variation in the likelihood resulting from changing its PID hypothesis without modifying that of the remaining tracks.

For kaons between 2 and 100 GeV, the efficiency of this algorithm is 97% for a

misidentification rate of 5% due to pions.

2.5.2 Stripping

The stripping represents the last stage before the offline selection for physics analysis. Protoparticles are provided with a definite type, and hence with a definite mass, according to its PID information. The resulting objects are named *particles*. In addition, primary vertices are reconstructed from the VeLo tracks. Events are classified in streams by means of pre-selection algorithms provided by the groups of physics analysis. These algorithms commmonly operate by searching decay candidates, using a loose version of the offline selection. All the information produced at this level, which comprise particles, vertices and decay candidates, is added to the reconstructed data, and kept in DST.

Chapter 3

The measurement of the branching ratio of $B_s^0 \to \mu^+\mu^-$ at LHCb

This chapter explains the details of the analysis of the $B_s^0 \to \mu^+\mu^-$ decay at LHCb. Section 3.1 introduces the simulated samples that are used for this work. Section 3.2 presents the analysis strategy, focussing on the event selection, the definition of discriminant variables, and the use of the so-called CL_s method for extracting $\mathrm{BR}(B_s^0 \to \mu^+\mu^-)$. Section 3.3 provides an estimation of the sensitivity that could be achieved using this procedure as a function of the integrated luminosity. Finally, Section 3.4 discusses the calibration of the discriminant variables and the normalization of the measurement, showing the convenience of using control channels. The family of $B \to h^+h'^-$ decays is introduced in this context, discussing their advantages for control purposes in the analysis of $B_s^0 \to \mu^+\mu^-$.

3.1 Simulated data samples

The study presented in this document is based on Monte Carlo (MC) simulations generated in the LHCb Data Challenge 06 (DC06) [101]. All the samples under consideration used Pythia [102] for the simulation of the process $pp \to b\bar{b}X$ and the subsequent hadronization of the resulting b quarks. In order to optimize the MC production, events were required to have at least one b quark within a cone of 0.4 radians around the forward z-axis, which roughly covers the LHCb acceptance. The simulation included pile-up, assuming a luminosity of $2{\times}10^{32}$ cm^{-2}s^{-1}. The decay of the B mesons was simulated by EvtGen [103], while the interaction of the final state with the detector was performed with Geant4 [104]. Finally, the digitization of the detector response and the reconstruction was achieved by specific LHCb software packages called Boole [105] and Brunel [106] respectively.

In this document, the study of the signal, the control channels and some specific backgrounds is carried out using exclusive samples. The combinatoric background is estimated by means of an inclusive $b\bar{b}$ sample and the so-called *dimuon* sample, which is explained below.

The exclusive samples are built by forcing one b quark to hold the required process, while the other b quark decays according to an extensive list of predefined channels based on the most frequent b hadron decays. The b hadron holding the decay of interest must satisfy the cut on the production angle defined at the beginning of this section. In addition, most of the exclusive samples require the final state of the decay of interest to be contained within the geometrical acceptance of LHCb. Several samples of this type are considered: the $B_s^0 \to \mu^+\mu^-$ sample is used to describe the signal; the $B \to h^+h'^-$ samples are used either as control channels, or for studying the background due to misidentification; finally, the $B_c^+ \to J/\psi(\mu^+\mu^-)\mu^+\nu_\mu$ and $\Lambda_c \to \pi p$ samples are considered as specific backgrounds. Throughout this document the contribution of each $B \to h^+h'^-$ sample to the mixture of all $B \to h^+h'^-$ modes is weighted in order to match that expected from the current measurements.

Inclusive $b\bar{b}$ samples are generated by allowing both b hadrons to decay according to the list of channels introduced above. The *dimuon* sample was obtained by generating inclusive $b\bar{b}$ events, and keeping only those having two oppositely charged muons coming directly from a b hadron. The symbol $b\bar{b} \to \mu^+\mu^-$ is often used to refer to this type of events.

Table 3.1 shows the total number of generated events n_{gen}, the efficiency of the generation requirements ϵ_{gen}, and the equivalent integrated luminosity $\mathcal{L}_{gen}$ for all the samples discussed above. These three variables are related through the equation

$$n_{gen} = \epsilon_{gen}\sigma\mathcal{L}_{gen}, \tag{3.1}$$

where σ is the cross section of the process under consideration. For samples dedicated to exclusive B decays, this cross section can be expressed as

$$\sigma = 2\sigma_{b\bar{b}} f_B \mathrm{BR}, \tag{3.2}$$

where $\sigma_{b\bar{b}}$ refer to the cross section for the $b\bar{b}$ production, and f_B represents the probability of a b quark to hadronize into the corresponding B meson. The factor 2 is due to the presence of two B mesons in the event, associated with each b quark. DC06 simulations assume $\sigma_{b\bar{b}}$ equal to 0.5 mb, and f_B equal to 0.405, 0.405 and 0.1 for B_u, B_d and B_s mesons respectively.

3.2 Description of the analysis

The procedure devised for measuring $\mathrm{BR}(B_s^0 \to \mu^+\mu^-)$ at LHCb attempts to improve the sensitivity of the classical cut-and-count methods by extracting the maximum information contained in the distribution of discriminant variables [107]. For this reason, the use of an event selection based on cuts is restricted to few loose cuts applied at the first level of the analysis. This procedure allows to retain the bulk of the signal, at the expenses of low purity. Events are then classified in bins according to two discriminant variables. The corresponding bin contents are combined in order to extract the BR, or to set an exclusion limit.

Process	n_{gen}	ϵ_{gen}	$\mathcal{L}_{gen}$ (fb^{-1})
$B_s^0 \to \mu^+\mu^-$	107k	0.200	1.68k
$B^0 \to K^+\pi^-$	1.27M	0.202	0.802
$B_s^0 \to K^+K^-$	1.28M	0.346	1.12
$B^0 \to \pi^+\pi^-$	1.32M	0.199	3.19
$B_s^0 \to \pi^+K^-$	128k	0.204	1.28
$B_c^+ \to J/\psi(\mu^+\mu^-)\mu^+\nu_\mu$	50.7k	0.231	0.142
$\Lambda_b \to \pi p$	97.9k	0.208	1.37
Inclusive $b\bar{b}$	852k	0.437	3.90×10^{-6}
$b\bar{b} \to \mu^+\mu^-$	21.8M	0.438	1.11×10^{-2}

Table 3.1: List of the MC samples used in this document, showing in each case the total number of events N_{gen}, the efficiency of the generation requirements ϵ_{gen} and their equivalent integrated luminosity $\mathcal{L}_{gen}$. The different value of ϵ_{gen} for $B_s^0 \to K^+K^-$ with respect to those for the other $B \to h^+h'^-$ modes is due to the lack of requirements on the final state applied in this specific exclusive sample.

3.2.1 Event reconstruction

The reconstruction of the final state of the $B_s^0 \to \mu^+\mu^-$ use long tracks only. As explained in Section 2.5.1, these tracks have the best momentum resolution, and have the necessary information for searching decay vertices. In addition, they can be safely extrapolated to other regions of the detector, as the muon system. The reconstruction efficiency of the tracks from the final state of $B_s^0 \to \mu^+\mu^-$ is defined as

$$\epsilon_{rec} \equiv \frac{n_{rec}}{n_{gen}}, \tag{3.3}$$

where n_{gen} is the number of generated events, and n_{rec} is the number of events with both muons from $B_s^0 \to \mu^+\mu^-$ reconstructed as long tracks. Note that, due to the requirements on the final state applied in the $B_s^0 \to \mu^+\mu^-$ sample, n_{gen} only count generated events with both muons from $B_s^0 \to \mu^+\mu^-$ contained within the geometrical acceptance of LHCb. This efficiency is found to be $\epsilon_{rec} = 0.682\pm0.003$.

A reconstructed track is identified as a muon by requiring the presence of hits within a definite region around its extrapolation into the muon system, called *field of interest*. The efficiency of the muon identification on reconstructed events is given by the ratio

$$\epsilon_{\mu ID/rec} \equiv \frac{n_{rec+\mu ID}}{n_{rec}}, \tag{3.4}$$

where $n_{rec+\mu ID}$ is the number of events with both decay products from $B_s^0 \to \mu^+\mu^-$ reconstructed as long tracks, and identified as muons. This efficiency is found to be $\epsilon_{\mu ID/rec} = 0.838\pm0.003$.

3.2.2 Event selection

The event selection is carried out by searching pairs of muons with opposite electric charge satisfying the following cuts:

- Both muons must form a vertex satisfying $\chi^2 < 14$.
- The invariant mass of the dimuon pair (IM) is required to be contained within a mass window of 60 MeV around the mass of the B_s (M_{B_s}), which should be enough for including the mass sidebands.
- The significance of the IP (IPS) of the reconstructed B meson with respect to the primary vertex has to be less than six.
- The angle between the momentum of the reconstructed B meson and the line joining the primary and secondary vertex, called *pointing angle*, must be below 0.1 radians.
- The secondary vertex is required to be downstream from the primary vertex for which the IPS of the reconstructed B meson is minimum.

These cuts are summarized in Table 3.2, and they are referred to as the *original* $B_s^0 \to \mu^+\mu^-$ selection. Note that requirements aiming to cut on the phase space of the final state of $B_s^0 \to \mu^+\mu^-$, as those involving the p_T or the IPS of the decay products, are not included in this selection. The selection efficiency is defined on reconstructed events as

$$\epsilon_{sel} \equiv \frac{n_{sel}}{n_{rec}}, \tag{3.5}$$

where n_{sel} is the number of reconstructed events satisfying the cuts of the original $B_s^0 \to \mu^+\mu^-$ selection except that on the IM of the dimuon pair. This cut is not included in order to be able to compare this efficiency for $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$, as the fact of having several $B \to h^+h'^-$ modes leads to the total IM distribution for these decays to be substantially different to that for $B_s^0 \to \mu^+\mu^-$. Note that events considered by n_{sel} are not required to pass the trigger. This efficiency is found to be $\epsilon_{sel} = 0.922 \pm 0.002$. The effect of the muon identification on events passing the offline selection, given by

$$\epsilon_{\mu ID/sel} \equiv \frac{n_{sel+\mu ID}}{n_{sel}}, \tag{3.6}$$

where $n_{sel+\mu ID}$ is the number of selected events with both decay products from $B_s^0 \to \mu^+\mu^-$ identified as muons, is found to be $\epsilon_{\mu ID/sel} = 0.838 \pm 0.003$. The expected ammount of signal for 2 fb^{-1} is equal to 57 events, if only the effects of the track reconstruction, the muon identification and the offline selection are considered.

Value	Cut
$\chi^2(B_s)$	< 14
$\lvert\Delta\mathrm{M}(B_s)\rvert$	< 60 MeV/c^2
$\mathrm{IPS}(B_s)$	< 6
$\alpha_{PV}(B_s)$	> 0.1 rad
$\Delta z_{PV}(B_s)$	> 0 mm

Table 3.2: Cut values of the original $B_s^0 \to \mu^+\mu^-$ selection.

3.2.3 Trigger

The trigger is considered as a residual effect on the sample of offline selected events, hence all the efficiencies quoting the trigger performance are evaluated on events passing the offline selection. The most important L0 lines used for triggering $B_s^0 \to \mu^+\mu^-$ are the *muon* line and the *dimuon* line, requiring a high p_T muon and a pair of high p_T muons respectively. The efficiency of these lines on signal events is 0.969±0.002 and 0.753±0.006 respectively. At HLT1, $B_s^0 \to \mu^+\mu^-$ is selected mainly by a *single muon* alley, requiring a muon with high p_T and high IPS with respect to the primary vertex, and the *muon+track* alley, requiring a muon forming a good vertex with a long track. Both alleys require the muon to be associated with a L0 muon candidates. The efficiency of these alleys on signal events passing the L0 is 0.936±0.004 and 0.823±0.006 respectively. In addition, six alleys based on muon pairs give efficiencies between 0.5 and 0.65 for signal events passing the L0. Two inclusive HLT2 selections are relevant for $B_s^0 \to \mu^+\mu^-$, based respectively on the topology of generic two-body B decays and on the presence of a muon and a track forming a displaced vertex, with efficiences of 0.861±0.005 and 0.828±0.006. The HLT2 also includes two exclusive selections based on the presence of a pair of muons, having efficiences equal to 0.810±0.006 and 0.797±0.006 respectively on $B_s^0 \to \mu^+\mu^-$.

The total trigger efficiency is defined as

$$\epsilon_{trig} \equiv \frac{n_{trig}}{n_{sel}}, \tag{3.7}$$

where n_{trig} is the number of events passing both the trigger and the offline selection. This efficiency is found to be $\epsilon_{trig} = 0.925\pm0.001$.

3.2.4 Sources of background

The study of the sample of inclusive $b\bar{b}$ events allowed to conclude that the only relevant source of combinatoric background are $b\bar{b} \to \mu^+\mu^-$ events. In addition, other relevant contributions from exclusive processes have been identified:

- Combination of two muons from $B_c^+ \to J/\psi(\mu^+\mu^-)\mu^+\nu_\mu$ decays. The fact that the mass of the B_c is above M_{B_s} implies that combinations of one muon from the J/ψ decay with the muon coming directly from the B_c may satisfy the IM

requirement. Assuming that the BR for $B_c^+ \to J/\psi\mu^+\nu_\mu$ is equal to that for $B_c^+ \to J/\psi e^+\nu_e$, Tevatron quotes $f_{B_c} \times \mathrm{BR}(B_c^+ \to J/\psi\mu^+\nu_\mu) = (2.6^{+1.2}_{-1.1})\times10^{-5}$ [108]. Considering the specific decay of the J/ψ into two muons, the BR of $B_c^+ \to J/\psi(\mu^+\mu^-)\mu^+\nu_\mu$ is found to be $(1.4^{+0.7}_{-0.6})\times10^{-6}$. For comparison, the value of $f_{B_s} \times \mathrm{BR}(B_s^0 \to \mu^+\mu^-)$ is expected to be equal to 3.2×10^{-10} within the SM. Muons of $B_s^0 \to \mu^+\mu^-$ candidates built from $B_c^+ \to J/\psi(\mu^+\mu^-)\mu^+\nu_\mu$ tend to form a good vertex, hence this variable does not allow the discrimination of this background. However, considering the B meson built from these muons, the compatibility of its IPS and pointing angle with the values allowed by the selection requirements is expected to be low due to the missing decay products.

- Misidentification of the two hadrons in $B \to h^+h'^-$ decays. The main source of misidentification is the decay of the produced pions and kaons into a muon and a muonic neutrino. The BR of this process is practically 100% for pions, and (63.55±0.11)% for kaons. The probability of misidentifying a single hadron as a muon due to this process is $\mathcal{O}(10^{-2})$. Besides, the sum of the product $f_B\times$BR for $B \to h^+h'^-$ decays is $\mathcal{O}(10^{-5})$. Hence the probability of having a $b\bar{b}$ event with a $B \to h^+h'^-$ decay whose both hadrons are identified as muons is of the order of 10^{-9}. Note that, within the SM, the probability of having a $b\bar{b}$ event with a true $B_s^0 \to \mu^+\mu^-$ decay is $f_{B_s}\times\mathrm{BR}(B_s^0 \to \mu^+\mu^-) = 3.2\times10^{-10}$. An important part of candidates coming from misidentified $B \to h^+h'^-$ decays is rejected by the IM requirement due to the wrong mass hypothesis and the missing four-momentum due to the neutrinos.

The expected ammount of events for 2 fb^{-1}, for signal and each background category, is shown in Table 3.3. The resulting signal-to-background ratio (S/B) is equal to $(2.3\pm0.1)\times10^{-4}$.

Process	$n_{yr,sel}$	$n_{yr,sel+trig}$
$B_s^0 \to \mu^+\mu^-$	57±1	53±1
$bb \to \mu^+\mu^-$	$(276\pm8)\times10^3$	$(227\pm7)\times10^3$
$B^0 \to K^+\pi^-$	<11, 90% CL	<11, 90% CL
$B_s^0 \to K^+K^-$	2±3	2±3
$B^0 \to \pi^+\pi^-$	1±2	1±2
$B_s^0 \to \pi^+K^-$	2±3	2±3
$B_c^+ \to J/\psi(\mu^+\mu^-)\mu^+\nu_\mu$	100±40	70±30

Table 3.3: Expected yields for 2 fb^{-1}, for signal and the different sources of background, using the offline selection shown in Table 3.2. Values in the first column do not take into account the effect of the trigger.

3.2.5 Event categorization

The analysis of the sample accepted by the offline selection is based on two independent discriminant event variables. They are identified respectively as the IM

and a likelihood containing the information about the geometry of the event called *geometrical likelihood* (GL). These variables are explained in the following.

Invariant mass

Due to the relatively small decay width of the B_s meson, the distribution of the IM for $B_s^0 \to \mu^+\mu^-$ is dominated by the experimental momentum resolution, leading to a gaussian dispersion of 25 MeV approximately. Final state radiation leads to a non-negligible radiative tail at low values of the IM.

Combinatorial background from inclusive $b\bar{b}$ follows a typical exponential decreasing distribution, while backgrounds due to $B_c^+ \to J/\psi(\mu^+\mu^-)\mu^+\nu_\mu$ and $B \to h^+h'^-$ show different structures, see Fig. 3.1. Specifically, the $B \to h^+h'^-$ component has a distribution similar to that for signal, but slightly shifted and with a degraded width, due to the incorrect mass assignment, the missing neutrinos produced in the semileptonic decays and the kick in the tracks caused by the decays.

Geometrical likelihood

The GL combines the minimal IPS and distance of closest approach (DOCA) of the two muons, the IP and lifetime of the B meson, and two variables parametrizing the isolation of each muon, called *isolations*, which are described below. It does not contain explicitly the modulus of the momentum of the decay products. This choice avoids correlations between the GL and the IM, enabling the independent calibration of each variable.

The two isolations aim to reject the background due to decays to more than two bodies, by examining the presence of long tracks accompanying each muon candidate. Excluding the two long tracks used to recontruct the $B_s^0 \to \mu^+\mu^-$ candidate, the isolation of each muon is defined as the number of long tracks satisfying the following requirements:

- Each track must be compatible with forming a vertex with the muon, by requiring their DOCA to be less than 200 μm. The position of this vertex is defined as $\mathbf{r}_{\mu+trk} = (\mathbf{r}_\mu + \mathbf{r}_{trk})/2$, being $\mathbf{r}_\mu$ and $\mathbf{r}_{trk}$ the positions on the muon trajectory and the long track respectively having the closest mutual approach.

- For each track, the z coordinate of the position $\mathbf{r}_{\mu+trk}$, denoted by $z_{\mu+trk}$, is required to satisfy $0 < z_{\mu+trk} - z_{PV} < 30$ mm, where z_{PV} is the z coordinate of the primary vertex for which the IPS of the B_s candidate is minimum.

- For each track, the sum of its momentum with the momentum of the muon, represented by $\mathbf{p}_{\mu+trk}$, must be contained within a cone around the line joining the primary vertex and the vertex of the muon+track combination. The angle α of this cone scales with the sum of the p_T of the muon and the long track, following the relation

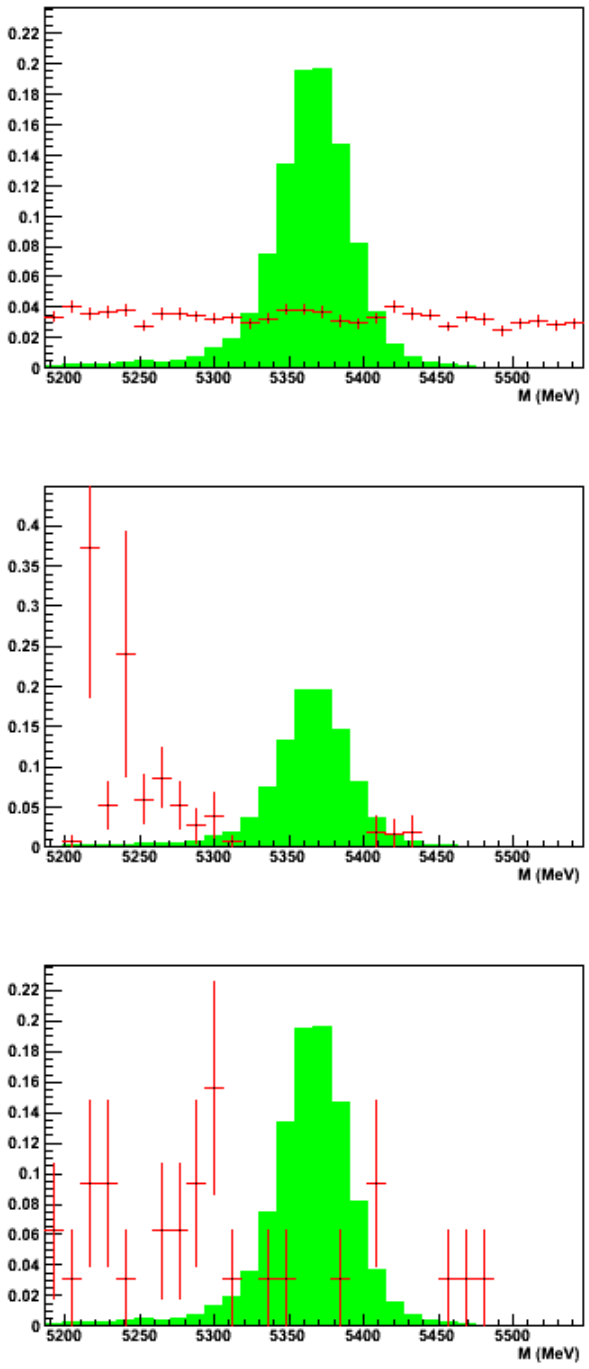

Figure 3.1: IM for events satisfying the offline selection, not including the trigger. Each plot represents signal (green filled histrogram) against a specific background contribution (red histogram). From top to bottom: combinatorial dimuon background, misidentification from $B^0 \to \pi^+\pi^-$, and fake combination from $B_c^+ \to J/\psi(\mu^+\mu^-)\mu^+\nu_\mu$. All histograms are normalized to unity.

$$\alpha = \frac{2}{3}\,\frac{p_{T,mu} + p_{T,trk}}{|\mathbf{p}_{\mu+trk}|}, \tag{3.8}$$

where $p_{T,\mu}$ and $p_{T,trk}$ represent the transverse momenta of the muon and the long track respectively.

The specific method used for combining the six variables of the GL in order to give a single value is discussed in the following. For convenience, the initial variables are represented by a vector $\mathbf{x}$. Two linear transformations $\mathbf{T_s}$ and $\mathbf{T_b}$ act separately on $\mathbf{x}$, aiming to convert the expected distribution of this vector for signal and background respectively into uncorrelated standard normal distributions [109]. Considering this property of the transformed variables, the two following quantities

$$\chi_s^2 \equiv \sum_{i=1}^{6} (\mathbf{y_s})_i^2, \quad \chi_b^2 \equiv \sum_{i=1}^{6} (\mathbf{y_b})_i^2 \tag{3.9}$$

should distribute as a χ^2 of six degrees of freedom for signal and background respectively, and hence quantify the compatibility of the observed input variables with the expected distribution in each case. The discriminant information is thus carried by the difference $\chi_s^2 - \chi_b^2$. For representation purposes, the GL is obtained by transforming this difference into a value defined between 0 and 1, and following a flat distribution for the expected signal.

The analytical expression of $\mathbf{T_s}$ and $\mathbf{T_b}$ is not obtainable as the analytical expression of the distribution of the vector $\mathbf{x}$ is not available neither for signal nor for background. Two effective operators have been introduced in order to evaluate the action of $\mathbf{T_s}$ and $\mathbf{T_b}$. They have been built according to the following facts:

- In many cases, non-linear dependences between two random variables are due mainly to the fact that their respective probability density functions are related through non-linear transformations. In such situation, if these variables are transformed into two new variables respectively distributing according to a common probability density function, the dependences between them become practically linear.
- If two variables are independent, their distribution can be transformed into a two-dimensional gaussian by converting separately each one-dimensional projection into a gaussian.

These considerations, stated for two random variables, are generalizable to an arbitrary number of random variables. Based on them, the effective transformations associated to $\mathbf{T_s}$ and $\mathbf{T_b}$ proceed in three steps. First, each original variable is transformed into a new quantity following a definite probability density function. According to the considerations in the first point above, this probability density function is common to each variable, choosing a standard normal distribution in the current studies. Then, the covariance matrix is used to determine the rotation

of the reference axes that is necessary to remove the remaining correlations. The resulting values are independent as long as only linear correlations were present before the rotation. Finally, each new variable is transformed into a quantity following a standard normal distribution, motivated by the statement in the second point above. This method provides an exact n-dimensional uncorrelated gaussian if the distribution yield by the first transformation contains only linear correlations. Note that the implementation of this procedure requires a numerical description of the distributions, which in the current study is achieved by means of histograms.

By construction, the distribution of the GL is flat for signal, see Fig. 3.2. Backgrounds are restricted to low values of the GL, very sharply in the case of combinatoric due to muons coming directly from B decays. For this reason, the region with GL greater than 0.5 is known as *sensitive region*.

3.2.6 Extraction of the branching ratio

The analysis of $B_s^0 \to \mu^+\mu^-$ at LHCb attempts to improve the sensitivity by using the distribution of the two discriminant variables defined in the previous subsection. To that purpose, the space spanned by such variables is divided in a certain number n of bins. The relevant information about the i-th bin is the expected ammount of signal s_i and background b_i, and the yield d_i measured experimentally. Note that, assuming a specific luminosity, s_i depends only the BR, while b_i is completelly determined. In order to perform the hypothesis test, the expected and measured bin contents are combined into a test-statistic Q, which is processed on the basis of a modified frequentist approach.

Construction of the test-statistic

The frequentist approach is applied typically using test-statistics that depend monotonically on the unknown parameters. A widely used test-statistic satisfying this property is the ratio of the two probability densities for the possible experimental results associated to the signal-plus-background and background-only hypotheses respectively. Representing the ammounts observed in n independent bins by $\{d_1, \ldots, d_n\}$, and the corresponding expected ammounts of signal and background by $\{s_1, \ldots, s_n\}$ and $\{b_1, \ldots, b_n\}$ respectively, this ratio reads

$$Q = \frac{\prod_{i=1}^{n} \text{Poisson}(d_i, s_i + b_i)}{\prod_{i=1}^{n} \text{Poisson}(d_i, b_i)}, \tag{3.10}$$

where Poisson(k,λ) refers to the Poisson distribution of the variable k with parameter λ. Note that this test-statistic depends monotonically on the BR, due to the s_i entering in the Poisson distribution of the numerator of Q_i.

Motivation for a modified frequentist approach

The conventional frequentist approach is based on the definition of the confidence level (CL) as

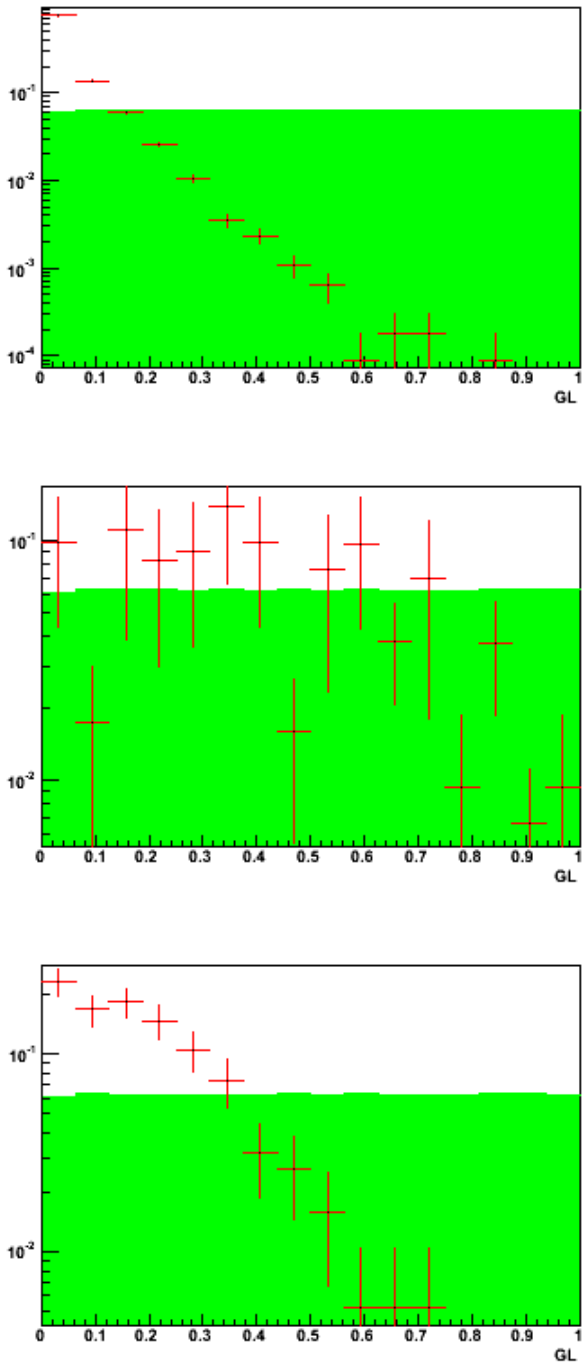

Figure 3.2: GL for events satisfying the offline selection, not including the trigger. Each plot represents signal (green filled histrogram) against a specific background contribution (red histogram). From top to bottom: combinatorial dimuon background, misidentification from $B^0 \to \pi^+\pi^-$, and fake combination from $B_c^+ \to J/\psi(\mu^+\mu^-)\mu^+\nu_\mu$. All histograms are normalized to unity.

$$CL_{s+b} \equiv P_{s+b}(Q \leq Q_{obs}), \tag{3.11}$$

being $P_{s+b}(Q \leq Q_{obs})$ the probability of having a value of Q less than or equal to that obtained experimentally, under the signal-plus-background hypothesis. Representing by $\mathcal{D}$ the outcome $\{d_1, \ldots, d_n\}$ of a possible experiment, $P_{s+b}(Q \leq Q_{obs})$ is expressed as

$$P_{s+b}(Q \leq Q_{obs}) = \sum_{\mathcal{D} \,/\, Q(\mathcal{D}) \leq Q_{obs}} \left[\prod_i \text{Poisson}(d_i, s_i + b_i) \right], \tag{3.12}$$

enabling an univocal relation among the confidence level CL_{s+b}, the BR, and the outcome $\mathcal{D}_{obs} \equiv \{d_{1,obs}, \ldots, d_{n,obs}\}$ given by the experiment. Fixing the value of CL_{s+b}, this relation provides a BR associated to a given experimental observation. The BR obtained using this procedure is interpreted as the upper bound defining a range that, if computed in many independent experiments, would cover the true value of the BR in a fraction $\alpha = CL_{s+b}$ of the cases.

Focussing on a single experiment, the BR provided by the conventional frequentist approach can not be interpreted according to the intuitive idea of an interval of values of the BR for which no signal is observed, which has more physical meaning. For instance, in cases of small signals, a downward background fluctuation can lead to BR $= 0$ at high CL_{s+b}. Although a perfectly valid result from a probabilistic point of view, it does not account for small values of the BR that might show signal using an increased statistics. This feature is due to the fact that CL_{s+b} checks only the signal-plus-background hypothesis, leaving aside the compatibility with the background-only hypothesis. For this reason, in order to have a method providing confidence intervals involving a more physical idea of exclusion, the basis of the conventional frequentist approach must be modified in order to include the confidence in the background-only hypothesis.

The CL_s method

The CL_{s+b} method [110][111] is aimed to offer an alternative interpretation to that of the conventional frequentist approach, more related with the physical meaning of exclusion. To that purpose, it checks the signal+background hypothesis against the background-only hypothesis, by using the ratio of their corresponding conventional confidence levels

$$CL_s \equiv \frac{CL_{s+b}}{CL_b}, \tag{3.13}$$

where, in analogy with the CL_{s+b}, the CL_b is defined as the probability of having a value of Q less than or equal to that obtained experimentally, but under the background-only hypothesis instead. The value of BR excluded at a CL equal to α is found within this framework by solving the equation $1 - CL_s \leq \alpha$.

The effect of the normalization of the CL_{s+b} using CL_b can be illustrated using the previous example describing small signals in presence of a background fluctuating downwards. In such situation, even a zero value for BR could be excluded at high CL using the conventional approach, due to the very low CL_{s+b} that would result from the observed outcome $\mathcal{D}_{obs}$. However, in the framework of the CL_s method, this downwards fluctuation of the background would also imply a low value of the CL_b as well, thus compensating the low confidence in the signal-plus-background hypothesis.

So far the method has been described for exclusion. It can be used for quoting an observation in a straightforward way. The central value is such that it is excluded half of the times, thus it corresponds to the value satisfying $CL_s = 0.5$. The significance is extracted using the confidence in the background-only hypothesis,

$$1 - CL_b \leq \frac{1}{\sqrt{2\pi}} \int_{-n}^{n} dx\, e^{\frac{-x^2}{2}}, \tag{3.14}$$

where n represent the number of standard deviations quantifying how much the ammounts provided by the experiment depart from the background-only hypothesis.

3.3 Expected performance

The potential to either exclude or measure the $B_s^0 \to \mu^+\mu^-$ decay, applying the techniques described in the previous sections, has been studied as a function of the luminosity, by taking into account the statistical uncertainties only. To that purpose, the distributions of signal and background were obtained from MC samples, in order to perform different toy experiments.

The description of signal and background was provided by the MC samples of $B_s^0 \to \mu^+\mu^-$ and $b\bar{b} \to \mu^+\mu^-$ respectively. These samples were filtered by the offline selection presented in Subsection 3.2.2. The effect of the trigger was not considered as it is expected to have a low impact on signal. The remaining signal and background events were classified separately under five bins in the IM and four bins in the GL.

Due to the limited statistics of the background sample, some bins have low or no background content. Representing by b_i the background content in the i-th bin, the statistical uncertainty introduced by bins with $b_i = 0$ was handled by considering new background contents $\tilde{b}_i$ obtained by shifting each b_i upwards by a common ammount ϵ,

$$b_i \quad \longrightarrow \quad \tilde{b}_i \equiv b_i + \epsilon,$$

such that the total number of background events $\sum_i b_i$ has a 90% probability of being below the sum of the shifted values $\sum_i \tilde{b}_i$.

For each luminosity, the distributions of signal and background were scaled accordingly, assuming different values of the BR. Then, the CL_s method was applied under two different assumptions:

- Absence of signal. The observed bin contents assumed in this case only included the ammount of expected background. The value of the BR excluded at 90% CL was computed for each luminosity. The statistical uncertainty due to bins having low or no background content was estimated by comparing the results obtained with the nominal and shifted background distributions respectively.
- Presence of signal. The observed bin contents were estimated as the sum of the corresponding ammounts of expected signal and background. In particular, signal was scaled according to the assumed BR. For each luminosity, the lower value of the BR enabling a significance of either 3σ or 5σ was computed. Since bins with no background imply an infinite contribution to the significance, only the shifted background was considered.

The value of the BR that would be either excluded or observed using this method are displayed in Fig. 3.3, for different values of the luminosity. According to these results, LHCb has the potential to exclude BR($B_s^0 \to \mu^+\mu^-$) down to the SM value with less than 2 fb^{-1}. Besides, the observation of the SM value at 3σ and 5σ is expected after 2 and 10 fb^{-1} respectively. These prospects imply that enhancements of BR($B_s^0 \to \mu^+\mu^-$) departing from the SM prediction can be explored within the first years of LHCb, thus providing one of the first tests of New Physics at the LHC.

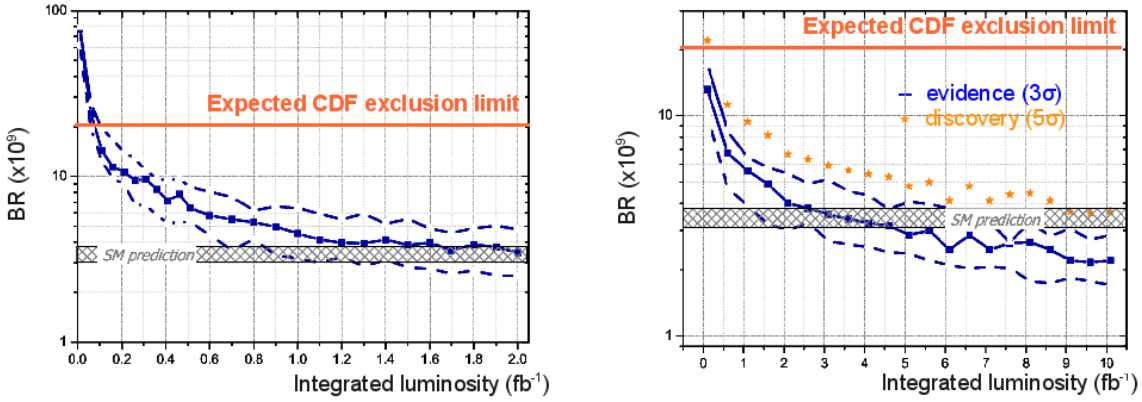

Figure 3.3: Exclusion limit at 90% CL (left), and observation at 3σ and 5σ (right), as a function of the integrated luminosity $\mathcal{L}$.

3.4 Calibration and normalization using control channels

The performance quoted in the previous section relies on the expected bin contens of signal and background, which in this specific case were obtained from simulations. However, poorly known effects as hadronization, or oversimplifications of the detector structure, may imply MC distributions to depart with respect to that occurring in the real experiment. In addition, further systematic uncertainties and biases enter

in the analysis if the integrated luminosity and the cross section of the $b\bar{b}$ production are used explicitly to estimate the overall normalization factor of these distributions.

These facts reflect the necessity of a proper calibration and normalization of the distributions used in the analysis, independent of the MC. Two control channels have been proposed in order to perform such tasks, taking advantage of their similarities with $B_s^0 \to \mu^+\mu^-$:

- $B^+ \to J/\psi(\mu^+\mu^-)K^+$. The BR of this process is $(1.014\pm0.034)\times10^{-3}$, leading to expect about 810M events of this type for 2 fb^{-1}. Although the phase space of the two muons from the J/ψ differ from the muons from $B_s^0 \to \mu^+\mu^-$, they are still characterized by a relatively high p_T and IP, due to the fact of coming from a B meson.

  For this reason, it is expected that the effect the trigger will be similar in both channels. In addition, no the PID and the trigger will be similar in both channels. This fact suggests the use of $B^+ \to J/\psi(\mu^+\mu^-)K+$ for normalization, thus enabling a partial cancellation of the trigger and PID effects. The kinematical properties of this channel depart from those of $B_s^0 \to \mu^+\mu^-$ due to the extra kaon, thus difficulting the treatment of the effects of the offline selection.

- $B \to h^+h'^-$ channels. The BRs of these channels are of the order of 10^5 and 10^6, expecting about 25M events for 2 fb^{-1}. Their use is motivated by the fact of sharing common kinematical features with signal. For this reason, any set of cuts based on kinematical variables would have similar effects on both channels. In particular, signal and control channels would still share similar kinematical properties after applying these requirements. Then, using a common offline selection with no PID requirements, such channels could be used for normalization, leading to a cancellation of the effects of the offline requirements, and for calibration of the IM and the GL, since they only involve kinematical information. Besides, the differences between these channels and signal are concerned basically with the effect of the trigger.

The previous discussion shows the central role represented by $B \to h^+h'^-$ channels in the analysis of $B_s^0 \to \mu^+\mu^-$. The use of these processes for control purposes imply to consider their offline selection, and the development of methods for correcting the effects due to the differences between $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$. These topics constitute the subject of this document, and are discussed in detail in the next three chapters.

Chapter 4

The use of $B \to h^+h'^-$ as control channels for BR($B_s^0 \to \mu^+\mu^-$)

This chapter describes how $B \to h^+h'^-$ is used to control the measurement of BR($B_s^0 \to \mu^+\mu^-$), and hence to make this measurement as little dependent as possible on MC simulation. Section 4.1 introduces the properties of $B \to h^+h'^-$ and discusses the use of these channels for normalization and for calibration of the variables used for assign a signal-likelihood to the events. Section 4.2 argues that the use of $B \to h^+h'^-$ requires the modification of the offline selection for $B_s^0 \to \mu^+\mu^-$ described in Chapter 3. The expected performance in terms of efficiency and purity of the modified selection for $B \to h^+h'^-$ decays is given, including a detailed discussion about the background sources. Finally, section 4.3 discusses the convenience of using PID cuts for selecting $B \to h^+h'^-$.

4.1 Comparison between $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ decays

The two-body decays of B mesons into charged pions and kaons are usually referred to as $B \to h^+h'^-$ decays. The experimental values of their BRs are displayed in Table 4.1. The table shows in addition the respective production rates, computed by assuming a $b\bar{b}$ cross section of 0.5 mb, and f_{B_d} and f_{B_s} equal to 0.405 and 0.1 respectively. The most frequent $B \to h^+h'^-$ decay is $B^0 \to K^+\pi^-$, followed by $B^0 \to \pi^+\pi^-$ and $B_s^0 \to K^+K^-$. Both $B^0 \to K^+\pi^-$ and $B^0 \to \pi^+\pi^-$ have been widely studied at B factories, reaching a precision of few per cent in the measurement of their BRs. Channels involving the B_s meson, such as $B_s^0 \to K^+K^-$ or $B_s^0 \to \pi^+K^-$, have been studied at Tevatron only. The measurement of their BRs is affected by the 13% uncertainty in f_{B_s}.

The kinematical properties of $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ depend on the mass of the B meson and the masses of decay products. The difference between the mass of the B_d and B_s is 86.8 MeV, which is 1.6% of M_{B_s}, while the difference between the masses of the possible decay products can reach 388 MeV at most. These differences are small compared to the momentum scale of the decay products

Decay	BR ($\times 10^5$)	$2\sigma_{b\bar{b}} f_{\rm B}$BR (nb)
$B^0 \to \pi^+\pi^-$	0.513 ± 0.024	1.05 ± 0.05
$B^0 \to K^+\pi^-$	1.94 ± 0.06	4.0 ± 0.1
$B^0 \to KK$	<0.12, 90% CL	<0.06, 90% CL
$B_s^0 \to \pi\pi$	<0.041, 90% CL	<0.08, 90% CL
$B_s^0 \to \pi^+K^-$	0.49 ± 0.10	0.24 ± 0.05
$B_s^0 \to K^+K^-$	3.3 ± 0.9	1.6 ± 0.4

Table 4.1: Current averages of the BR and expected production rate of $B \to h^+h'^-$ decays.

in the center-of-mass frame of the B meson (around 2 GeV), and to the typical boost of these particles at LHCb (comprised between 10 and 100 GeV). This fact leads to consider that most differences between the kinematical properties of $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ can be neglected. The use of these similarities for calibrating the kinematical properties of $B_s^0 \to \mu^+\mu^-$ and normalizing the measurement of its BR is simplified if reconstruction, trigger and offline selection are also similar for both signal and control channels. As explained in Section 4.2, it is possible to have an offline selection satisfying this requirement. However, track reconstruction and trigger introduce sizable systematic effects that have to be taken into account, as explained in Subsecs. 4.1.1 and 4.1.3 respectively. The methods aiming to correct such effects are discussed in Chapter 5.

4.1.1 Differences in track reconstruction between $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$

The different physical properties of muons and hadrons may affect the signature left by these particles in the tracking system. The reconstruction efficiency, as defined by Eq. 3.3, is 0.682±0.003 and 0.529±0.002 for $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ respectively. The difference between these values is due to the different penetration depth into the LHCb detector for muons and hadrons. Fig. 4.1 shows that, contrary to muons from $B_s^0 \to \mu^+\mu^-$, a substantial fraction of hadrons from $B \to h^+h'^-$ undergo an inelastic hard interaction within the tracking system of LHCb. This fraction is found to be (20.07±0.02)%.

The dependence of the track reconstruction efficiency for muons and hadrons with their kinematical properties is displayed in Fig. 4.2. Aside from the global difference discussed above, the dependences are similar for $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$. This allows to conclude that the selection efficiency on reconstructed events, as defined by Eq. 3.5, and the distributions of kinematical variables for reconstructed events should remain similar for $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$.

4.1.2 Calibration and normalization of BR($B_s^0 \to \mu^+\mu^-$) using $B \to h^+h'^-$

The effect of the event selection on reconstructed events is expected to be similar for $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$, as the differences from track reconstruction do

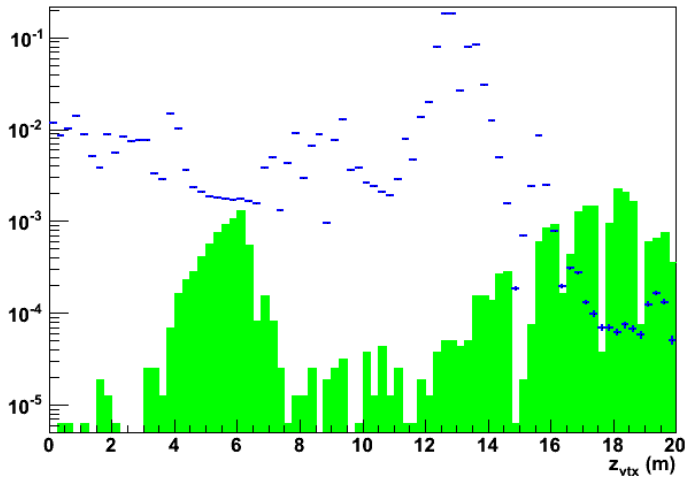

Figure 4.1: z-coordinate of the end vertex of the decay products of $B_s^0 \to \mu^+\mu^-$ (green filled histogram) and $B \to h^+h'^-$ (blue histogram), along the z axis. The tracking system of LHCb covers the range in z between the origin and 9.5 m, and the muon system starts at $z \sim$15 m. All histograms are normalized to the respective fraction of decay products with z-coordinate of the end vertex below 20 m.

not affect substantially the distribution of the kinematical variables. Fig. 4.3 shows the distribution of the variables that we would like to calibrate using $B \to h^+h'^-$ instead of MC (IM and GL) for $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$, using for each decay product the correct mass from the generator. The IM distribution of each $B \to h^+h'^-$ mode has been shifted so that the true mass of the decaying B meson [1] is situated at M_{B_s}. The agreement between the distributions of these two variables for both channels supports the use of $B \to h^+h'^-$ for their calibration.

This similarity between the kinematical properties of $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ implies that selection cuts involving only kinematical quantities might have practically the same effect on both channels. Aside from the additional advantage introduced by this feature in the calibration of the IM and the GL, it suggests the use of $B \to h^+h'^-$ for normalizing the measurement of $\mathrm{BR}(B_s^0 \to \mu^+\mu^-)$,

$$\mathrm{BR}_S = \frac{\epsilon_C}{\epsilon_S}\frac{n_S}{n_C}\frac{f_{B_C}}{f_{B_s}}\mathrm{BR}_C, \tag{4.1}$$

where the subindices S and C represent the signal and control channels respectively, and ϵ and n refer to the total efficiency and the observed ammount of events respectively.

The normalization of $\mathrm{BR}(B_s^0 \to \mu^+\mu^-)$ with $B \to h^+h'^-$ would allow the cancellation of the effects due to selection requirement based on kinematical properties, if applied in common to both channels. The efficiency of the original $B_s^0 \to \mu^+\mu^-$ selection on reconstructed events, as defined by Eq. 3.5, is 0.922±0.002 and 0.893±0.002

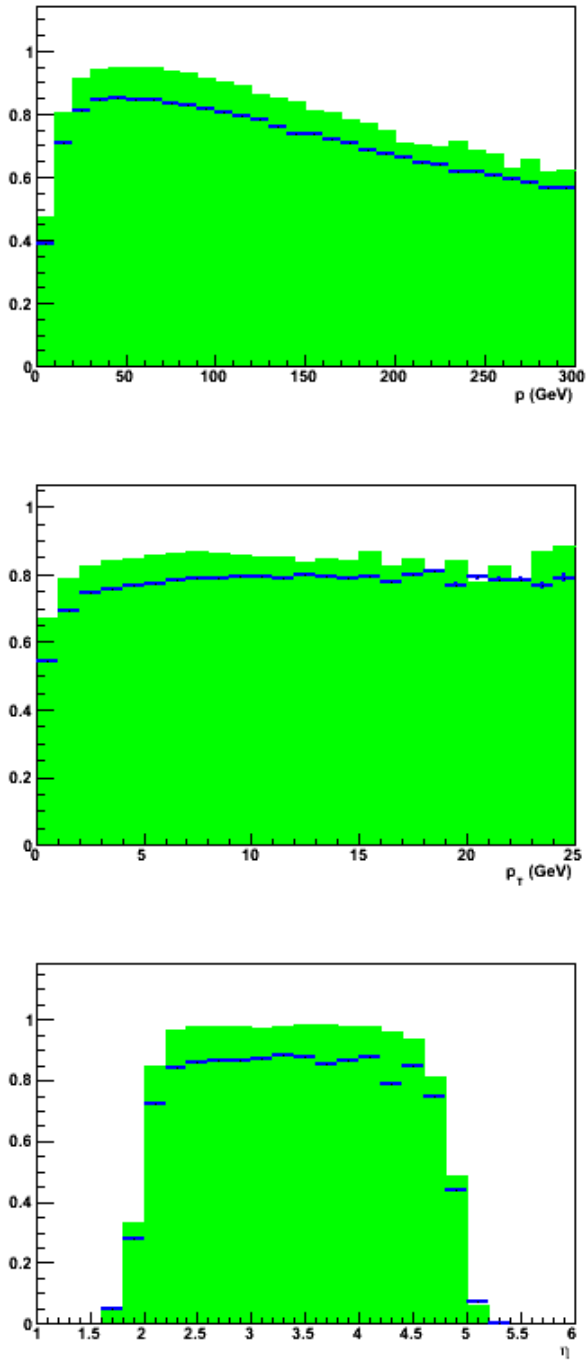

Figure 4.2: From top to bottom, track reconstruction efficiency of the decay products of $B_s^0 \to \mu^+\mu^-$ (green filled histogram) and $B \to h^+h'^-$ (blue histogram), as a function of the modulus of the momentum, the p_T and the pseudo-rapidity.

for $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ respectively. The small difference is a residual effect of the differences introduced by the track reconstruction.

The normalization using $B \to h^+h'^-$ channels involving the B_d meson require the knowledge of the fraction f_{B_C}/f_{B_s} in Eq. 4.1. This ratio depends on the conditions at which b quarks are produced, hence the current measurements provided by Tevatron and Belle can not be directly used by LHCb. Using $B_d \to D^-K^+$ and $B_s^0 \to D_s^-\pi^+$ decays [112] with 35 pb^{-1} of data at LHCb, this ratio has been measured to be $0.250\pm0.024^{\text{stat}}\pm0.017^{\text{syst}}\pm0.017^{\text{theor}}$ [113].

4.1.3 Triggering $B_s^0 \to \mu^+\mu^-$ vs. triggering $B \to h^+h'^-$

The relative abundance of hadrons in LHC interactions is two orders of magnitude above the abundance of muons, as shown in Table 4.2. For this reason, the trigger requirements oriented to channels containing only hadrons in the final state are tighter than those oriented to channels including muons. This fact introduces important differences between $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$, which need to be taken into account.

Trigger efficiency. The trigger efficiency, as defined by Eq. 3.7, is found to be 0.925 ± 0.001 and 0.4839 ± 0.0006 for $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ respectively. Note that these efficiencies are quoted over events that already satisfy the offline selection. The difference between these two trigger efficiencies affects the normalization of BR$(B_s^0 \to \mu^+\mu^-)$ using $B \to h^+h'^-$.

Biases on distributions due to the trigger. The biases introduced by the trigger in the distribution of the IM and the GL for $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ respectively are shown in Fig. 4.4. These biases are given for events that already satisfy the offline selection. In particular, substantial differences are found between signal and control channels on the GL distribution after the trigger.

μ	π	K	p
0.2789 ± 0.0006	20.96 ± 0.02	2.681 ± 0.003	1.494 ± 0.002

Table 4.2: From left to right, average number of reconstructed muons, pions, kaons and protons per event, for a simulated sample of inclusive $b\bar{b}$ events.

4.2 Towards a common selection of $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$

In order to avoid systematic effects due to the selection requirements in the calibration and normalization of the analysis of $B_s^0 \to \mu^+\mu^-$ using $B \to h^+h'^-$, the cuts applied to $B \to h^+h'^-$ should be as similar as possible to those applied to $B_s^0 \to \mu^+\mu^-$ However, applying the original $B_s^0 \to \mu^+\mu^-$ selection to $B \to h^+h'^-$ would lead to

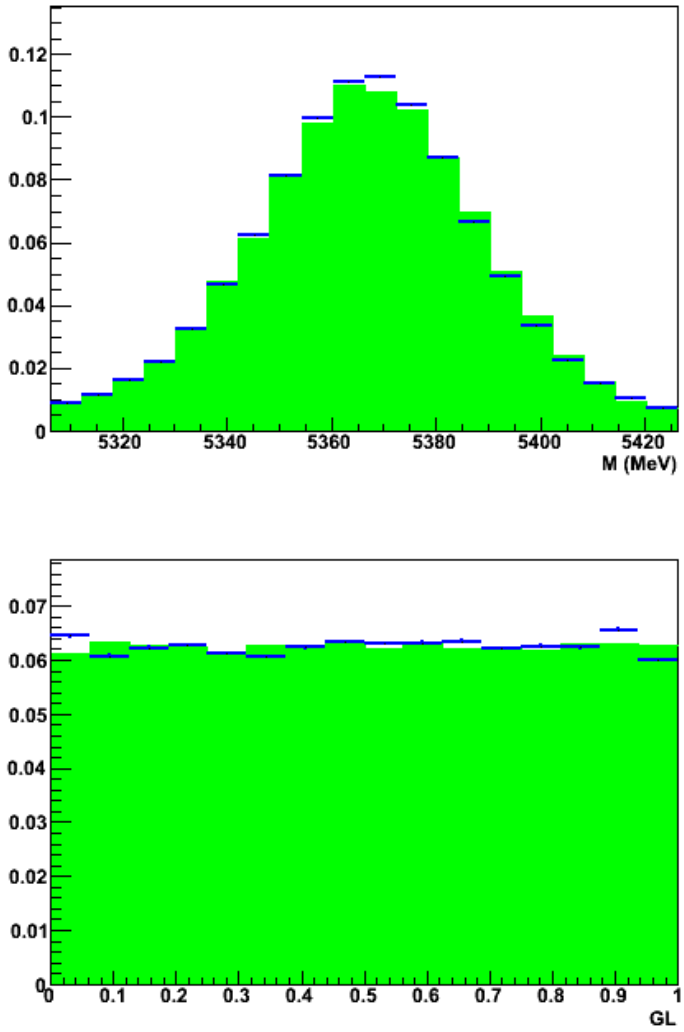

Figure 4.3: IM (top) and GL (bottom), for $B_s^0 \to \mu^+\mu^-$ (green filled histogram) and $B \to h^+h'^-$ (blue histogram), after the original $B_s^0 \to \mu^+\mu^-$ selection. The IM is computed using for each decay product the correct mass from the generator. In addition, the IM distribution of each $B \to h^+h'^-$ mode has been shifted so that the true mass of the decaying B meson is situated at M_{B_s}. All histograms are normalized to unity.

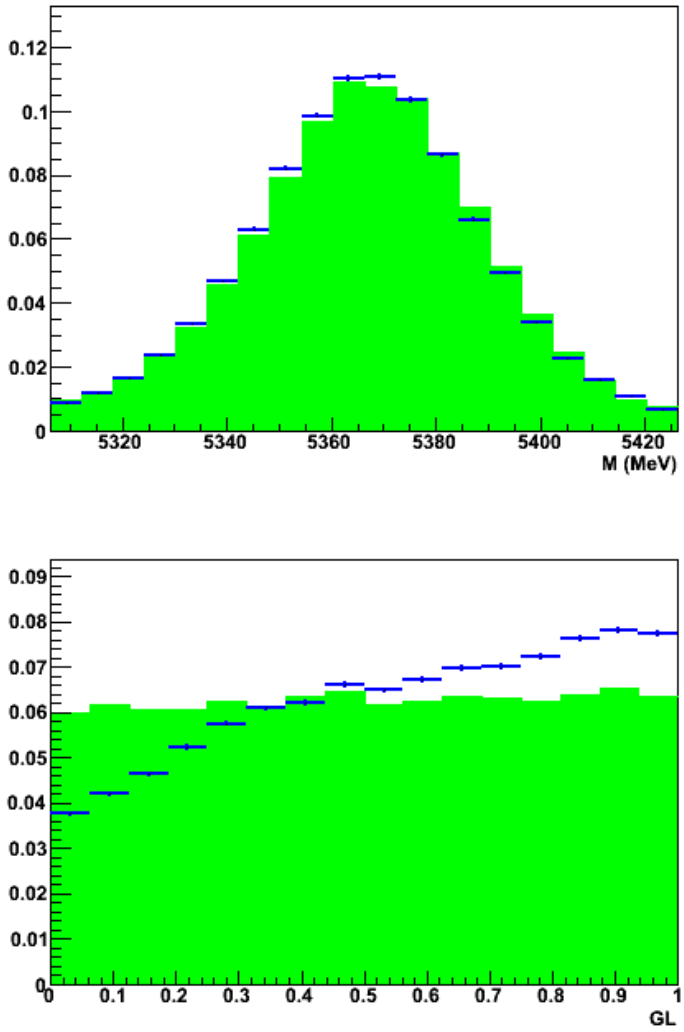

Figure 4.4: IM (top) and GL (bottom), for $B_s^0 \to \mu^+\mu^-$ (green filled histogram) and $B \to h^+h'^-$ (blue histogram), after the trigger and the original $B_s^0 \to \mu^+\mu^-$ selection. The IM for each $B \to h^+h'^-$ mode is computed assuming the mass hypotheses used at the generator level. In addition, the IM distribution of each $B \to h^+h'^-$ mode has been shifted so that the true mass of the decaying B meson is situated at the mass of the B_s. All histograms are normalized to unity.

an unnaffordable ammount of combinatoric background, as the average number of hadrons per event is two orders of magnitude above that for muons, see Table 4.2. This reason motivates a study aimed to a common selection for $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$, which provides high efficiency on the former and enough rejection for the latter.

4.2.1 The common selection

The design of a new selection for both $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ was guided by existing selection algorithms for $B \to h^+h'^-$ within the LHCb experiment. In particular, the framework for measuring the γ angle at LHCb from $B \to h^+h'^-$ decays [114] provides two algorithms used for stripping and offline selection. In this document, these algorithms are referred respectively to as the $B \to h^+h'^-$ stripping and selection *for measuring* γ. Both algorithms involve requirements on the same set of variables, but using different cut values in each case, as shown in Table 4.3.

Value	Cut (stripping)	Cut (selection)
$\mathrm{p}_{T,min}(h_1,h_2)$	> 0.6 GeV/c	> 1 GeV/c
$\mathrm{p}_{T,max}(h_1,h_2)$	> 2 GeV/c	> 3 GeV/c
$\mathrm{IPS}_{min}(h_1,h_2)$	> 3.5	> 6
$\mathrm{IPS}_{max}(h_1,h_2)$	> 5.5	> 12
$\chi^2(B_{d,s})$	< 6	< 5
$\lvert\Delta\mathrm{M}(B_{d,s})\rvert$	< 2 GeV/c^2	< 0.05 GeV/c^2
$\mathrm{p}_T(B_{d,s})$	> 0.7 GeV/c	> 1 GeV/c
$\mathrm{IPS}(B_{d,s})$	< 3.15	< 2.5
$\mathrm{DOFS}(B_{d,s})$	> 12	> 18
$\Delta z_{PV}(B_{d,s})$	> 0 mm	> 0 mm

Table 4.3: Cut values for the stripping and the selection of $B \to h^+h'^-$ for measuring γ.

Some cuts in the stripping and selection of $B \to h^+h'^-$ for measuring γ involve quantities already included in the original $B_s^0 \to \mu^+\mu^-$ selection, as the χ^2 of the vertex, the IM and IPS of the B meson, and the z coordinate of the secondary vertex. The IM used by these algorithms is computed by assuming for both hadrons the pion mass hypothesis by default. The cut values for the χ^2 of the vertex and and IPS of the B meson for these algorithms are tighter than those used for the original selection of $B_s^0 \to \mu^+\mu^-$.

The cuts on the IPS and the distance-of-flight significance (DOFS) of the B meson are related to the cut on the pointing angle of the B meson $\alpha_{PV}(B)$ in the original $B_s^0 \to \mu^+\mu^-$ selection, according to the relation

$$\alpha_{PV}(B) \approx \frac{\mathrm{IP}(B)}{\mathrm{DOF}(B)}, \tag{4.2}$$

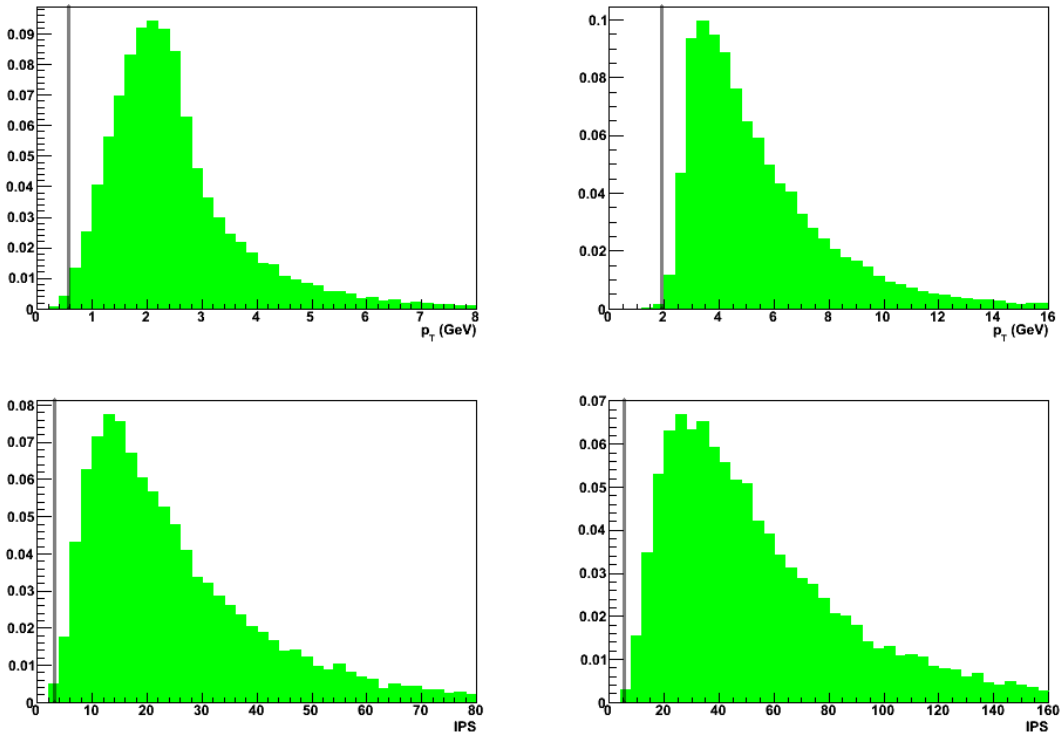

Figure 4.5: Minimal and maximal values of the p_T (top left and top right respectively) and minimal and maximal values of the IPS (bottom left and bottom right respectively) for $B_s^0 \to \mu^+\mu^-$ events satisfying GL > 0.5. Vertical lines show the respective cut values used in the original $B \to h^+h'^-$ stripping for measuring γ.

where IP(B) and DOF(B) refer to the IP and the distance of flight (DOF) of the B meson.

The remaining requirements of the stripping and selection of $B \to h^+h'^-$ for measuring γ, namely p_T and IPS of the decay products and p_T of the B meson, have no counterpart in the original selection of $B_s^0 \to \mu^+\mu^-$. These new cuts would allow to reduce drastically the combinatorics in the selection of $B \to h^+h'^-$. In order not to affect the framework developed in Chapter 3 the addition of cuts on p_T and IPS of the decay products to the original $B_s^0 \to \mu^+\mu^-$ selection should not to affect substantially the sensitive signal. The distribution of these cut variables for signal passing the original $B_s^0 \to \mu^+\mu^-$ selection and having GL > 0.5 in Fig. 4.5 shows that the cut values of the $B \to h^+h'^-$ stripping for measuring γ are highly efficient. Indeed, the efficiency on such $B_s^0 \to \mu^+\mu^-$ events of the additional cuts on p_T and IPS of the decay products is found to be 0.9903 ± 0.0006.

For practical reasons, it is convenient that the new common selection for $B_s^0 \to \mu^+\mu^-$ and its control channels is also used as stripping selection for $B \to h^+h'^-$ for measuring γ. The stripping of $B \to h^+h'^-$ using the looser cut values for the χ^2 of

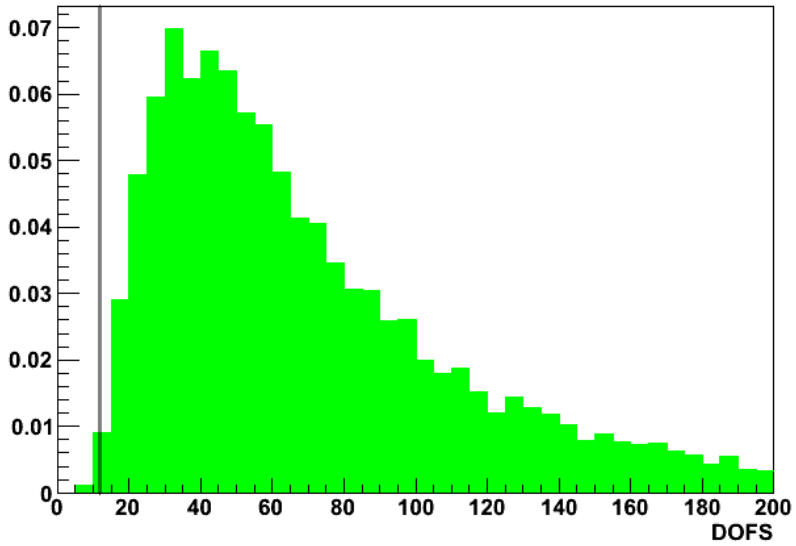

Figure 4.6: DOFS of the B meson for $B_s^0 \to \mu^+\mu^-$ passing the original $B_s^0 \to \mu^+\mu^-$ selection. The vertical line shows the cut value used in the original $B \to h^+h'^-$ stripping for measuring γ.

the vertex and the IPS of the B meson from the original $B_s^0 \to \mu^+\mu^-$ selection allows a reasonable background rejection if the mass window is set to 0.6 GeV [115].

The lack of the cut on the p_T of the B meson in the new selection with respect to the original $B \to h^+h'^-$ stripping for measuring γ do not affect severely the background rejection. This is due to the fact that the effect of this requirement on combinatorics is practically achieved by the cuts on the p_T of the decay products [116].

The cut on the DOFS of the B meson is necessary for keeping the background rejection within reasonable limits. For this reason, and based on the relation given by Eq. 4.2, the cut on the pointing angle of the B meson in the original $B_s^0 \to \mu^+\mu^-$ selection is replaced by that on the DOFS of the B meson. This change has a low impact on sensitive signal, as shown in Fig. 4.6. Finally, the effect of using the pion mass hypothesis by default on the total IM distribution of $B \to h^+h'^-$ is handled by enlarging the mass window to the range comprised between 5140 and 5340 MeV, as shown in Fig. 4.7. The use of the 60 MeV mass window around the mass of the B meson by means of the PID information is discussed in Section 4.3.

The selection built from the considerations above is summarized in Table 4.4. This new selection of $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ is referred to as the *common* selection. For signal events in the sensitive region, the ratio between the number of events passing these requirements and the number of events passing the original $B_s^0 \to \mu^+\mu^-$ selection is found to be 1.007 ± 0.009. This confirms the low impact of the new selection on sensitive signal. The total efficiency on reconstructed events, as

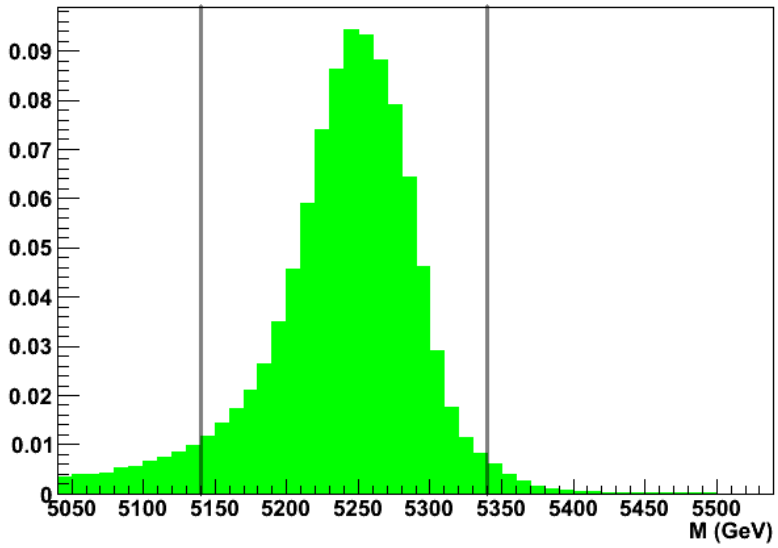

Figure 4.7: IM for $B \to h^+h'^-$ passing the common selection, assuming the pion mass hypothesis by default. Vertical lines show the mass window proposed in the text.

defined by Eq. 3.5, is equal to 0.650 ± 0.004.

The expected yields for $B_s^0 \to \mu^+\mu^-$ and the backgrounds explained in Chapter 3, given by the common selection and applying muon identification, are shown in Fig. 4.5. Note that this new selection improves the S/B for $B_s^0 \to \mu^+\mu^-$ by one order of magnitude with respect to that obtained with the original $B_s^0 \to \mu^+\mu^-$ selection, namely from $(2.3\pm0.1)\times10^{-4}$ to $(1.5\pm0.1)\times10^{-3}$.

Value	Cut
$p_{T,min}(h_1,h_2)$	> 0.6 GeV/c
$p_{T,max}(h_1,h_2)$	> 2 GeV/c
$IPS_{min}(h_1,h_2)$	> 3.5
$IPS_{max}(h_1,h_2)$	> 5.5
$\chi^2(B_{d,s})$	< 14
$\lvert\Delta M(B_{d,s})\rvert$	< 60 MeV/c^2
$IPS(B_{d,s})$	< 6
$DOFS(B_{d,s})$	> 12
$\Delta z_{PV}(B_{d,s})$	> 0 mm

Table 4.4: Cut values used in the common $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ selection.

Process	$n_{yr,sel}$	$n_{yr,sel+trig}$
$B_s^0 \to \mu^+\mu^-$	43±1	41±1
$b\bar{b} \to \mu^+\mu^-$	(33±3)×10^3	(28±3)×10^3
$B^0 \to K^+\pi^-$	<11, 90% CL	<11, 90% CL
$B_s^0 \to K^+K^-$	<3, 90% CL	<3, 90% CL
$B^0 \to \pi^+\pi^-$	<2, 90% CL	<2, 90% CL
$B_s^0 \to \pi^+K^-$	<3, 90% CL	<3, 90% CL
$B_c^+ \to J/\psi(\mu^+\mu^-)\mu^+\nu_\mu$	30±22	15±16

Table 4.5: Expected yields for 2 fb^{-1}, for signal and the different sources of background considered in Chapter 3, using the common selection and applying muon identification. Values in the first column do not take into account the effect of the trigger.

4.2.2 Performance of the new common selection on $B \to h^+h'^-$

The efficiency on reconstructed events for $B \to h^+h'^-$, as defined by Eq. 3.5, is found to be 0.630±0.003, which is similar to that found for $B_s^0 \to \mu^+\mu^-$, namely 0.650±0.004. The small difference is a residual effect of the differences introduced by the track reconstruction. The expected yields of each $B \to h^+h'^-$ mode obtained by using such selection are shown in Table 4.6, for 2 fb^{-1}. The mass window considered is [5140, 5340] MeV.

Process	$n_{yr,sel}$ (×10^{-3})	$n_{yr,sel+trig}$ (×10^{-3})
$B^0 \to K^+\pi^-$	947±3	455±2
$B_s^0 \to K^+K^-$	397±1	195±1
$B^0 \to \pi^+\pi^-$	253±1	121±1
$B_s^0 \to \pi^+K^-$	60±1	29±1
Total	1657±3	800±2

Table 4.6: Expected yields for 2 fb^{-1}, for different $B \to h^+h'^-$ decay modes, using the common selection on combinations of two reconstructed tracks. Values in the first column do not take into account the effect of the trigger.

The distribution of the IM and GL for $B \to h^+h'^-$ events satifying the common selection is shown in Fig. 4.8. Note that there are no significant differences with respect to the respective distributions obtained for $B_s^0 \to \mu^+\mu^-$, hence confirming the validity of assuming the same kinematical properties for both $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$.

By using inclusive $b\bar{b}$ events for estimating the background, the S/B obtained for $B \to h^+h'^-$ is found to be 0.19 ± 0.07 in the mass window [5140, 5340] MeV. The background is handled here by simply subtracting it using mass sidebands, assuming a linear background distribution. This is motivated by the fact that most of the background is combinatorial. Only (2.2 ± 0.6)% of the background is due to two particles coming from the same B meson (*physical background*). The distribution of

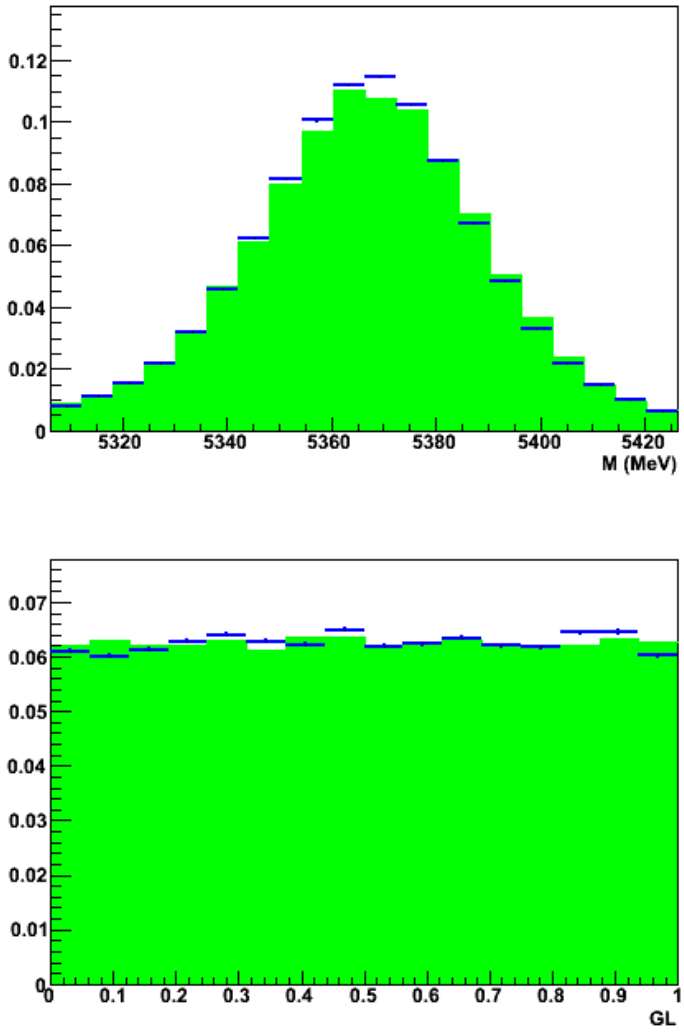

Figure 4.8: IM (top) and GL (bottom), for $B_s^0 \to \mu^+\mu^-$ (green filled histogram) and $B \to h^+h'^-$ (blue histogram), after the common selection. The IM is computed using for each decay product the correct mass from the generator. In addition, the IM distribution of each $B \to h^+h'^-$ mode has been shifted so that the true mass of the decaying B meson is situated at M_{B_s}. All histograms are normalized to unity.

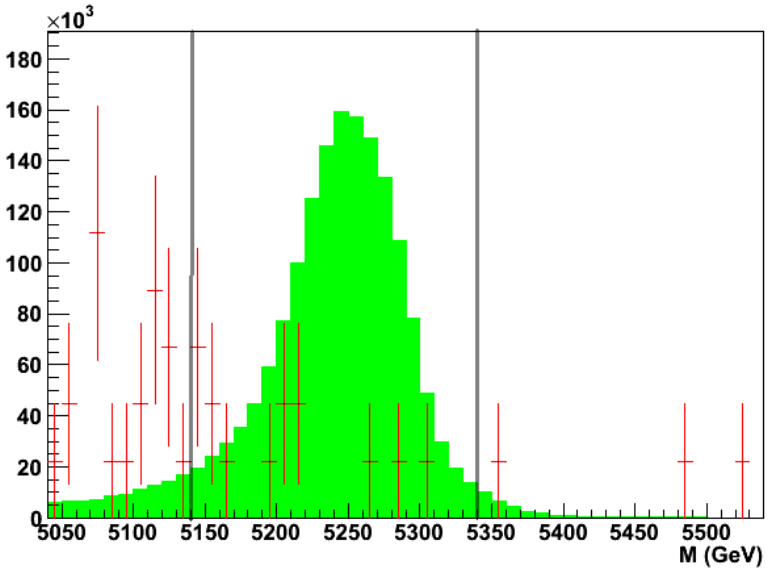

Figure 4.9: IM for $B \to h^+h'^-$ (green filled histogram) and for physical background (red histogram), passing the common selection, and using the pion mass hypothesis by default. Vertical lines show the mass window proposed in the text. All histograms are normalized to 2 fb^{-1}.

the IM of the physical background, displayed in Fig. 4.9, may not allow its proper subtraction using mass sidebands, as its IM distribution is different from that of the combinatoric background. The ammount of background b_{est} estimated for the mass window using the mass sidebands may depart systematically from the true ammount of background b falling in this window. The impact of this bias $b_{est} - b$ depends on the ammount of signal s, thus it can be quantified by means of the ratio

$$\Delta\mathrm{B/S} \equiv \frac{b - b_{est}}{s}. \tag{4.3}$$

Note that a value of ΔB/S compatible with zero would imply that the physical background can be safely subtracted by means of the estimation provided by the mass sidebands. By estimating the physical background from inclusive $b\bar{b}$ events the value of ΔB/S is found to be -0.10$\pm$0.08 by using the mass window [5140, 5340] MeV and two mass sidebands of 100 MeV on each side.

Two processes have been identified as the main sources of physical background:

- Three-body B decays, with at least two decay products being charged pions or kaons. The combination built from these two particles may provide fake $B \to h^+h'^-$ candidates. The BR of these decays are typically around 10^{-5} or below, and using the inclusive $b\bar{b}$ sample the yield for 2 fb^{-1} is estimated to be (90$\pm$40)$\times 10^3$ events. The IM for this background would be distributed below the mass of the B meson, as shown in Fig. 4.10. These exclusive decays

represent the main contribution to the value of ΔB/S quoted above. This is checked by computing ΔB/S considering only such processes for estimating the physical background, which gives -0.10±0.06 by using the mass window [5140, 5340] MeV and one mass sideband of 100 MeV on each side. However, due to the IM distribution of this background, this value can be lowered by using two sidebands on the right side of the mass window.

- Decays of beauty baryons. These processes frequently contain a proton in their final states. In particular, the two-body decay $\Lambda_b^0 \to ph^-$, where the proton is mainly identified as either a pion or a kaon, is considered in detail. The BR of this process is $(9.8 \pm 2.3)\times10^{-6}$ [1], and using the inclusive $b\bar{b}$ sample its expected yield for 2 fb^{-1} is equal to $(220\pm70)\times10^3$ events. Considering only the contribution from Λ_b^0 decays to the physical background, the ΔB/S is found to be 0.028±0.011 by using the mass window between 5140 and 5340 MeV and one mass sideband of 100 MeV on each side. In addition, using inclusive $b\bar{b}$ events the IM distribution for Λ_b^0 decays involving more than two bodies in the final state is found to be practically linear in the mass range used for $B \to h^+h'^-$, as shown in Fig. 4.10. This motivates to consider that such processes have a low impact on the value of ΔB/S.

4.3 Use of the PID information for selecting $B \to h^+h'^-$

The computation of the IM for $B \to h^+h'^-$ using the pion mass hypothesis by default leads to important differences with respect to the mass of the B meson, as shown in Fig. 4.11. The proper identification of the decay products as either pions or kaons allows to obtain the right mass, which also can be used for resolving the contribution from single $B \to h^+h'^-$ modes. These issues are important for calibrating the IM distribution of $B_s^0 \to \mu^+\mu^-$, as well as for improving the S/B. However, the use of PID requirements may involve some inefficiencies and biases that need to be taken into account.

As explained in Subsection 2.5.1, the separation between pions and kaons at LHCb is carried out using information from the RICH subdetectors. The relative likelihood for several mass hypotheses is computed for each track, and expressed by means of the difference between the logarithms of two such likelihoods, called *delta log-likelihood* (ΔLL) [117]. In particular, one ΔLL relates the likelihoods for the pion and kaon hypotheses, denoted by $\Delta\text{LL}_{K-\pi}$. Using this variable only, pions and kaons can be defined as those tracks having the value of $\Delta\text{LL}_{K-\pi}$ below and above some thresholds c_π and c_K respectively. The generic *loose* and *tight* values of c_π and c_K used in LHCb are summarized in Table 4.7. Note that these definitions are not exclusive, as it is possible to have particles satisfying both the pion and the kaon requirements.

The performance of the pion identification is given in terms of its efficiency, defined as the fraction of true pions correctly identified, and its misidentification rate, defined as the fraction of true kaons identified as pions. Analogous ratios are used for quoting

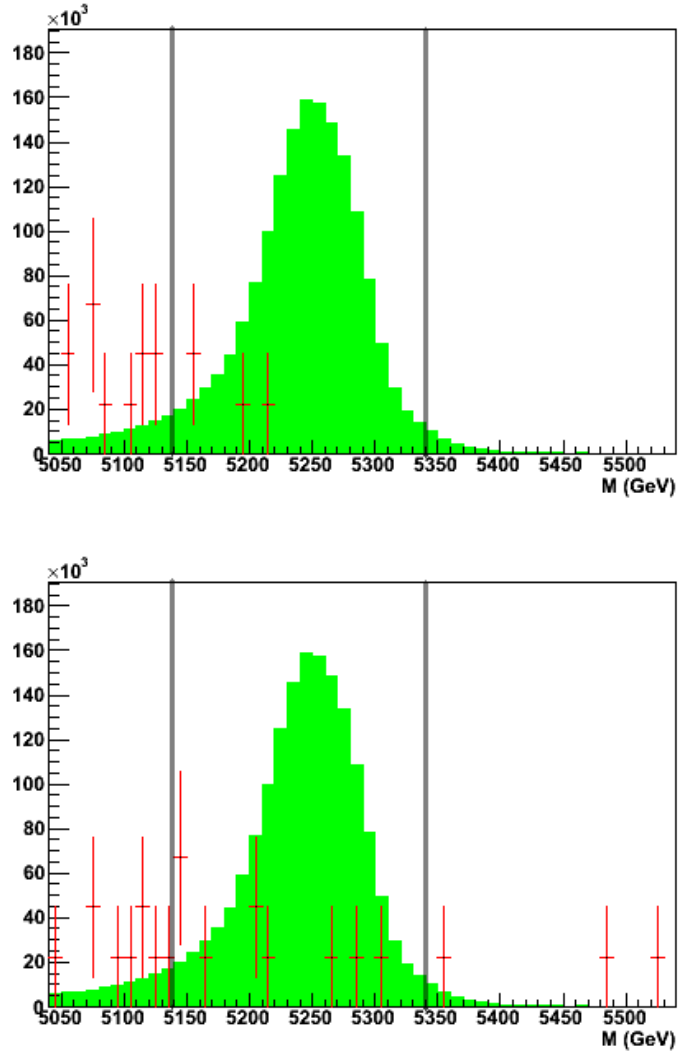

Figure 4.10: IM for $B \to h^+h'^-$ (green filled histograms) and for charmless three-body B decays (top red histogram) and Λ_b^0 decays (bottom red histogram), passing the common selection, and using the pion mass hypothesis. Vertical lines show the mass window proposed in the text. All histograms are normalized to 2 fb^{-1}.

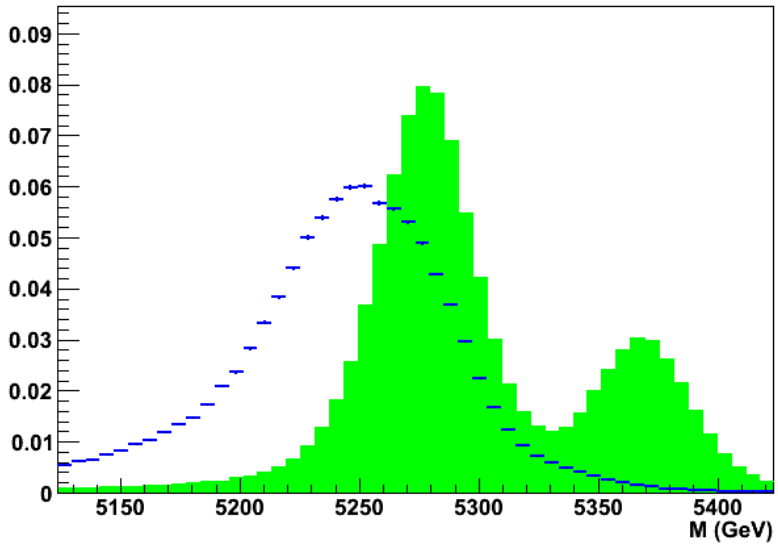

Figure 4.11: IM for $B \to h^+h'^-$ passing the common selection, assuming the mass hypotheses used at the generator level (green filled histogram) and the default mass hypothesis (blue histogram). All histograms are normalized to unity.

Category	Requirements
Standard Loose Pions	–
Standard Loose Kaons	$\Delta\text{LL}_{K-\pi} > -5$
Standard Tight Pions	$\Delta\text{LL}_{K-\pi} < 5$
Standard Tight Kaons	$\Delta\text{LL}_{K-\pi} > 0$

Table 4.7: Standard definitions used for the identification of pions and kaons at LHCb.

the performance of the kaon identification. Fig. 4.12 shows the performance of the pion and kaon identification as a function of the threshold for $\Delta\text{LL}_{K-\pi}$, for hadrons produced in $B \to h^+h'^-$ decays.

The efficiency of the pion and kaon identification depends on the momentum of these particles, as shown in Fig. 4.13, for the cases in which both thresholds c_π and c_K are set to -0.5 and -2.5. The motivation for considering these particular thresholds is given below. This efficiency decreases for increasing momentum, due to the loss of resolution between the Cerenkov rings of pions and kaons respectively, as shown in Fig. 2.5.

By setting both thresholds c_π and c_K to the same value, a particle is exclusively classified as either a pion or a kaon. The particular choice for the current study is to set c_π and c_K equal to zero, as the performance of the PID cuts with such thresholds is practically the same for pions and kaons, as shown in Fig. 4.12. If

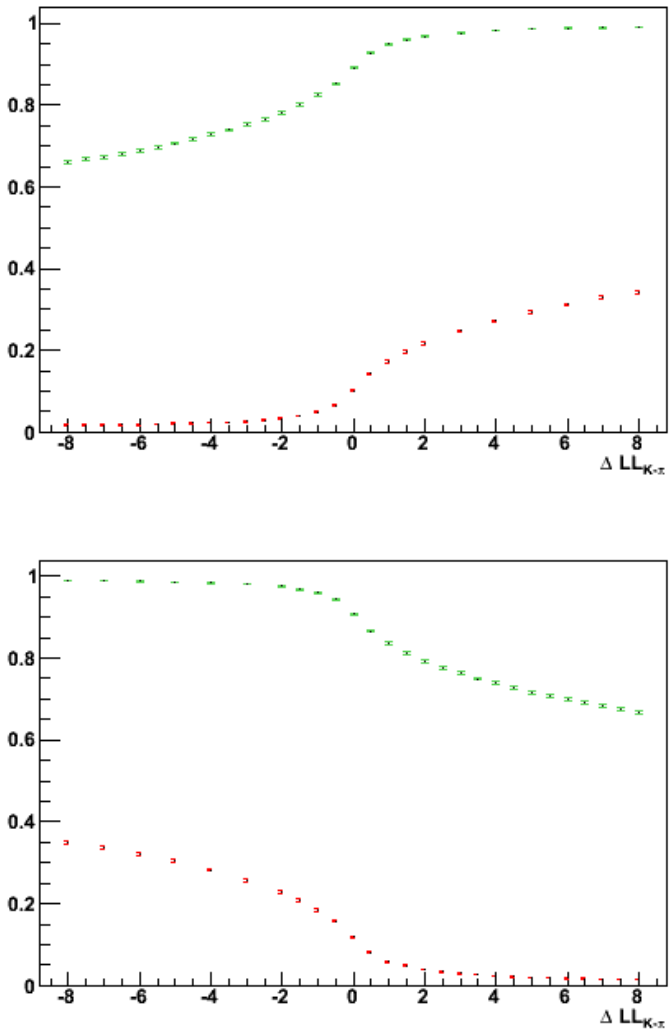

Figure 4.12: Efficiency (green points) and misidentification rate (red points) for hadrons produced in $B \to h^+h'^-$ decays identified as pions (top) and kaons (bottom), as a function of the threshold for $\Delta\mathrm{LL}_{K-\pi}$. Events are required to satisfy the common selection.

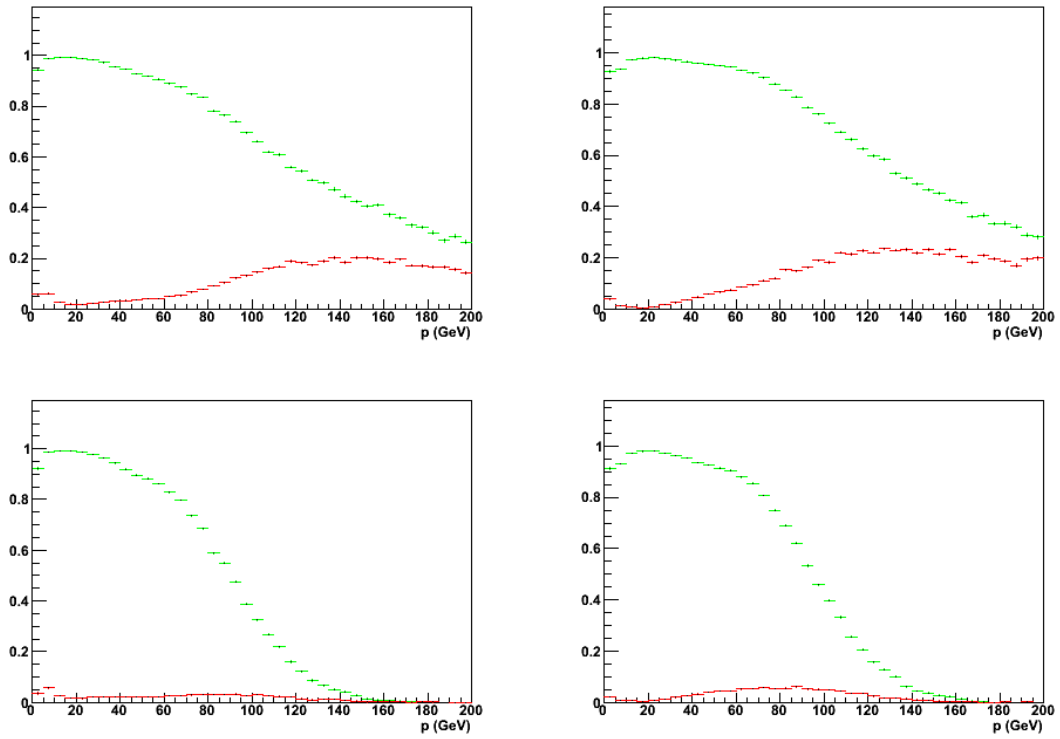

Figure 4.13: Efficiency (green points) and misidentification rate (red points) of the pion (left) and kaon (right) identification as a function of the momentum of the hadron, setting both thresholds c_π and c_K equal to -0.5 (top) and -2.5 (bottom). Events are required to satisfy the common selection.

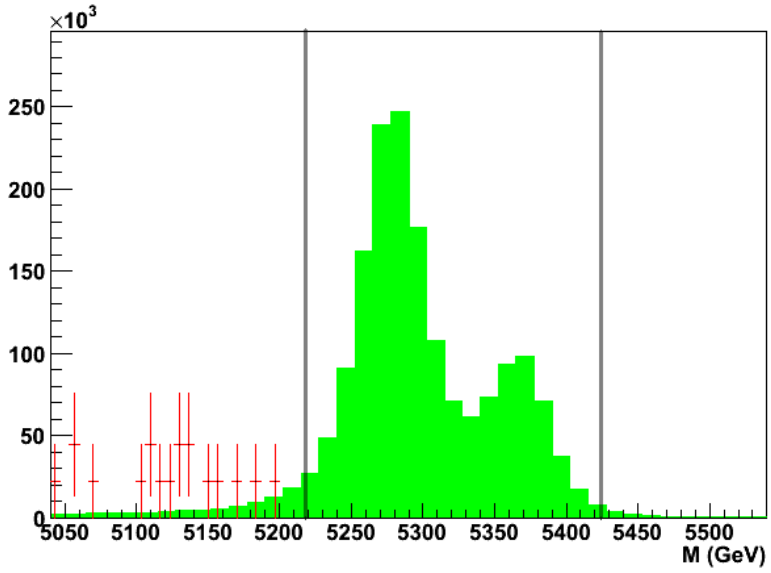

Figure 4.14: IM for $B \to h^+h'^-$ (green filled histogram) and for charmless three-body B decays (red histogram), passing the common selection, and using the most likely mass hypotheses. Vertical lines show the mass window proposed in the text. All histograms are normalized to 2 fb^{-1}.

this particle identification hypothesis is used, the resulting IM distribution shows two peaks centered at the B^0 and B_s^0 masses respectively, with widths similar to those of single modes. This is displayed in Fig. 4.14. $B \to h^+h'^-$ events from all modes can then be selected by requiring the IM to lie on the union of the two mass windows of 60 MeV around the mass of the B^0 and B_s^0 mesons. The requirement on the IM for the selection of generic $B \to h^+h'^-$ modes thus becomes more similar to that for $B_s^0 \to \mu^+\mu^-$. Note that this procedure for selecting $B \to h^+h'^-$ improves substantially the discrimination between these decays and the background due to charmless three-body B decays, as shown in Fig. 4.14. The S/B found by using this method is equal to 0.15±0.05, which is similar to that obtained by using the pion mass hypothesis by default, namely 0.19±0.07.

In order to resolve single $B \to h^+h'^-$ contributions both thresholds c_π and c_K are set to 2.5 and 0.5 for hadrons below and above 50 GeV respectively. These cut values are chosen in order to have little contamination at low momenta, and to keep signal at high momenta. More details about this choice are given in Subsection 5.2. Note that these cuts provide exclusive pion and kaon definitions. Table 4.8 shows the performance of the identification of single $B \to h^+h'^-$ modes using these requirements.

In addition to $\Delta\text{LL}_{K-\pi}$, the ΔLL relating the proton and pion hypothesis, called $\Delta\text{LL}_{p-\pi}$, can be used for improving the particle identification. In particular, it allows to discriminate background coming from Λ_b^0 decays. Fig. 4.15 shows the maximum

Mode	Efficiency	Purity	Yield (2 fb^{-1}, $\times 10^{-3}$)	S/B
$B^0 \to K^+\pi^-$	0.749±0.001	0.9316±0.0003	347±2	0.22±0.13
$B_s^0 \to K^+K^-$	0.7330± 0.0009	0.893±0.002	156±1	0.3±0.3
$B^0 \to \pi^+\pi^-$	0.7139± 0.0007	0.829±0.003	101±1	0.06± 0.04

Table 4.8: Performance of the identification of single $B \to h^+h'^-$ modes, by setting both thresholds c_π and c_K to 0.5 and 2.5 for particles below and above 50 GeV respectively. The efficiency and the purity refer to events passing the common selection. They are defined respectively as the fraction of $B \to h^+h'^-$ candidates of the given mode that satisfy the PID requirements, and the fraction of total $B \to h^+h'^-$ candidates passing these requirements that are correctly identified. Recall that the contribution of each $B \to h^+h'^-$ mode is weighted in order to match the measured relative abundances.

value of $\Delta\text{LL}_{p-\pi}$ of the reconstructed decay products, for $B \to h^+h'^-$ and $\Lambda_b^0 \to p\pi^-$.

Based on the distributions shown in Fig. 4.15, it was proposed to require both decay products to satisfy $\Delta\text{LL}_{p-\pi} < 30$. The efficiency of this cut, for $B \to h^+h'^-$ and Λ_b^0 decays from inclusive $b\bar{b}$ events passing the common selection, is found to be 0.9810 ± 0.0003 and 0.6 ± 0.2 respectively. This cut allows to reduce substantially the effect of possible non-linearities in the IM distribution of Λ_b^0 decays on the background subtraction for $B \to h^+h'^-$.

4.3.1 Impact of PID requirements on the calibration of the invariant mass

The use of PID requirements for selecting a single $B \to h^+h'^-$ mode biases the distribution of the IM to values closer to the mass of the decaying B meson. This is due to the dependence of the efficiency of such requirements with the momentum of the decay products, and to the fact that the momentum resolution of these particles decreases with their momentum. In addition, the effect of the background from the misidentification of the remaining $B \to h^+h'^-$ modes, called *cross-feed* background, may introduce additional biases on the IM distribution. The result of such effects, shown in Fig. 4.16, strongly depends on the performance of the RICH. This motivates the development of methods for calibrating the IM less dependent on such performance, which are discussed in Section 5.2.

4.3.2 Impact of PID requirements on the calibration of the GL

Unlike the IM, the GL is independent of the assumed mass hypothesis, hence the distributions of this variable for $B_s^0 \to \mu^+\mu^-$ and the mixture of all $B \to h^+h'^-$ contributions should be similar. However, in presence of background, the small S/B for low values of the GL could introduce large statistical uncertainties in the calibration of this distribution. The use of PID requirements for improving this S/B biases slightly the distribution of the GL, as shown in Fig. 4.17, due to residual correlations between the GL and the modulus of the momentum of the decay products. Such biases depend on the performance of the PID requirements. For

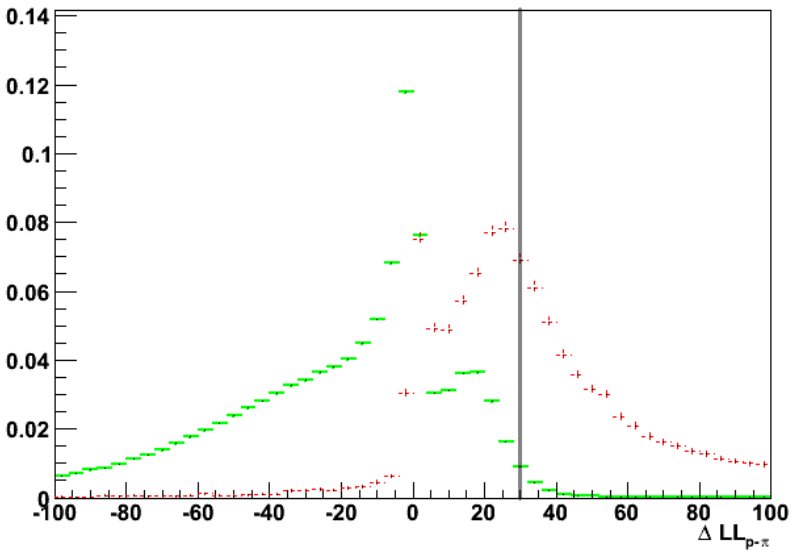

Figure 4.15: Maximum value of $\Delta\mathrm{LL}_{p-\pi}$ of the reconstructed decay products, for $B \to h^+h'^-$ (green solid line) and $\Lambda_b^0 \to p\pi^-$ (red dashed line) satisfying the common selection. The vertical line shows the cut value proposed in the text. All histograms are normalized to unity.

this reason, it is worth considering whether the use of the common selection alone is sufficient to ensure the proper calibration of the GL distribution of $B_s^0 \to \mu^+\mu^-$ using $B \to h^+h'^-$.

The relevant quantities necessary for the calibration of the GL are the relative bin contents f^i, where the subindex i is the bin label, defined as

$$f^i \equiv \frac{n^i}{\sum_j n^j}, \tag{4.4}$$

where n^i are the number of events in the i-th bin left after the background subtraction. Note that the uncertainties in n^i for bins at low GL, which are expected to give the most important contribution to the total uncertainty, propagate to the remaining bins via the sum of the denominator. For the current study, the GL range is divided in four equal bins, hence the value of f^i for $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ should be close to 1/4 in each bin. Although no single $B \to h^+h'^-$ mode is selected, the PID information is used for measuring the IM as described above. The effect of the cut on the IM does not introduce important biases in the GL distribution, as shown in Fig. 4.18.

For each bin in the GL, the signal and background contents in the mass window, denoted as s^i and b^i respectively, are estimated using MC samples, for a luminosity of 2 fb^{-1}. Only combinatoric background is considered, motivated by the fact that no physical background is found in the mass window using the inclusive $b\bar{b}$ MC sample.

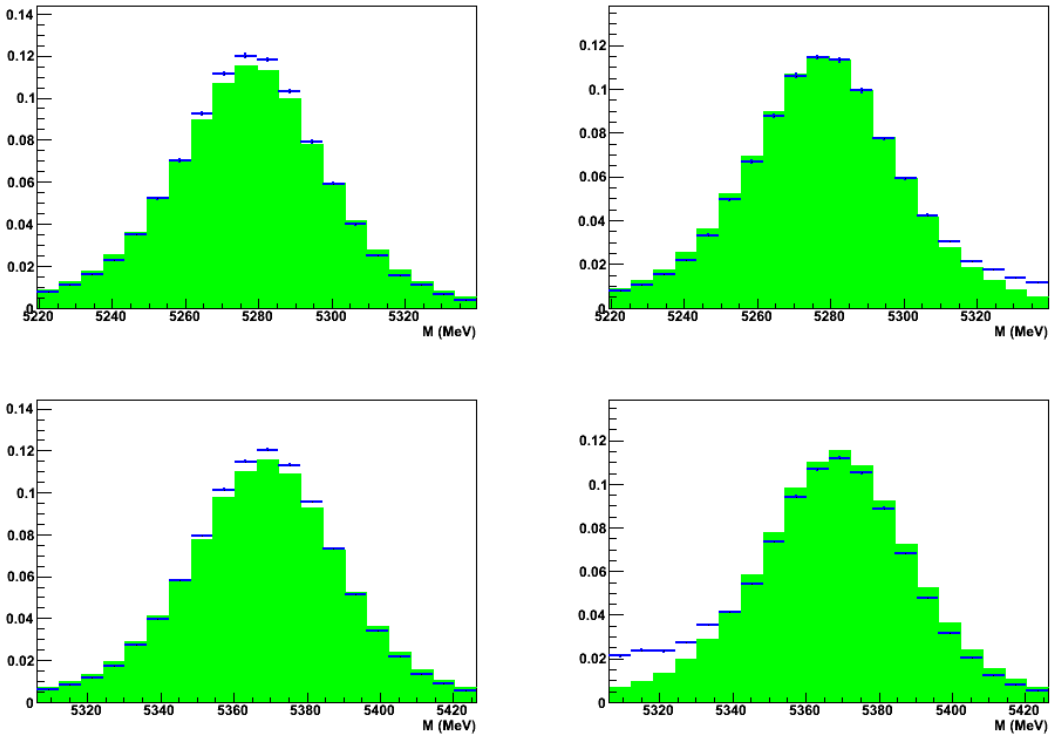

Figure 4.16: IM for $B^0 \to K^+\pi^-$ and $B_s^0 \to K^+K^-$ (top and bottom green filled histograms respectively), and $B \to h^+h'^-$ selected as $B^0 \to K^+\pi^-$ and $B_s^0 \to K^+K^-$ using the PID requirements explained in the text (top and bottom blue histograms respectively). Histograms on the left do not include the cross-feed background from $B \to h^+h'^-$, that is, only the effect from the PID inefficiency is shown. Events are required to satisfy the common selection. All histograms are normalized to unity.

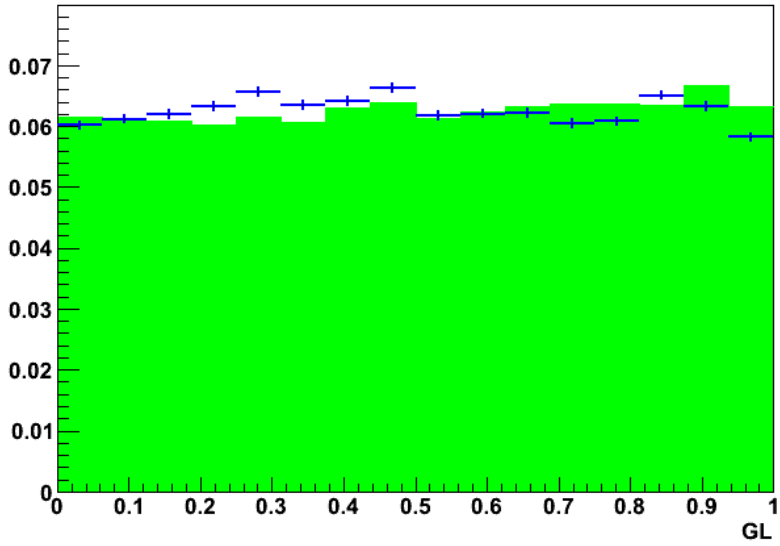

Figure 4.17: GL for $B^0 \to K^+\pi^-$ (green filled histogram) and $B \to h^+h'^-$ selected as $B^0 \to K^+\pi^-$ using the PID requirements explained in the text (blue histogram). Events are required to satisfy the common selection. All histograms are normalized to unity.

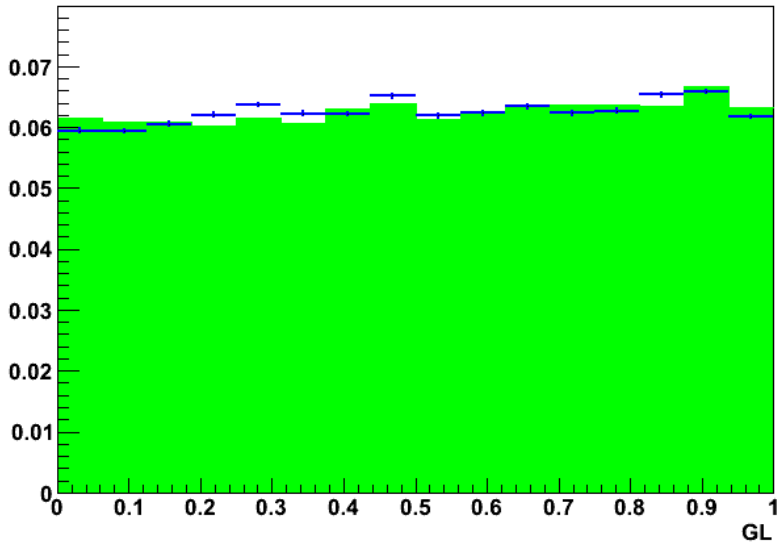

Figure 4.18: GL for $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ events passing the common selection. For $B \to h^+h'^-$ the cut on the IM is applied to the value computed using the most likely mass hypotheses. All histograms are normalized to unity.

These bin contents are used in 2000 toy experiments in order to estimate at each step the numer of events in the mass window and its sidebands, called n^i_{MW} and n^i_{SB}, as

$$\begin{array}{rcl} n^i_{MW} & = & \text{Poisson}(s^i) + \text{Poisson}(b^i), \\ n^i_{SB} & = & \text{Poisson}(b^i), \end{array} \tag{4.5}$$

where the function Poisson(x) gives random values distributed according to a Poisson distribution of parameter x. Note that the ammount of background in the mass sidebands is estimated by means of the background found in the mass window, keeping the same normalization factor. This assumes that the sum of the sizes of the mass sidebands is equal to the size of the mass window. The result n^i of the background subtraction is thus given as

$$n^i = n^i_{MW} - n^i_{SB}. \tag{4.6}$$

The relative dispersion of the values of f_i computed by the toy experiments is shown in Table 4.9. The statistical uncertainties are of the order of few per cent or less, even though the S/B in the first bin is of the order of 10^{-1}. This effect is understood by considering the ammount of signal and background in that bin, which is of the order of 10^6 and 10^7 respectively. In this case, the fluctuations in s'_i due to the background are around $\sqrt{10^7}$, which are two orders of magnitude below the ammount of signal.

Bin (i)	$\sigma(f_i)/f_i$ (%)
1	1.38
2	0.47
3	0.49
4	0.46

Table 4.9: Relative dispersion for the values of f_i for $B \to h^+h'^-$ passing the common selection and with no additional PID requirements, for 2 fb^{-1}. The cut on the IM is applied to the value computed using the most likely mass hypotheses.

The contribution to the total uncertainty due to the first bin in GL could be suppressed by using only the remaining three bins for the measurement of BR($B^0_s \to \mu^+\mu^-$). This implies to consider only events with GL above 0.25. In this case, the impact of discarding the first bin in the final result would be negligible as its sensitivity for the $B^0_s \to \mu^+\mu^-$ analysis is very low. In addition, the effect of this cut would be the same for $B^0_s \to \mu^+\mu^-$ and $B \to h^+h'^-$. This leads to a S/B equal to 2.4±0.6, which is one order of magnitude above from that obtained in Subsection 4.2.2.

The use of PID requirements aiming to discriminate protons from pions and kaons is important for the calibration of the GL as they allow to reject physical backgrounds as those due to Λ^0_b decays. Recall that such backgrounds could introduce biases in this variable as they might not be properly subtracted using mass sidebands. Fig. 4.19 shows that the impact of requiring $\Delta\text{LL}_{p-\pi} < 30$ on the GL distribution of $B \to h^+h'^-$ is low.

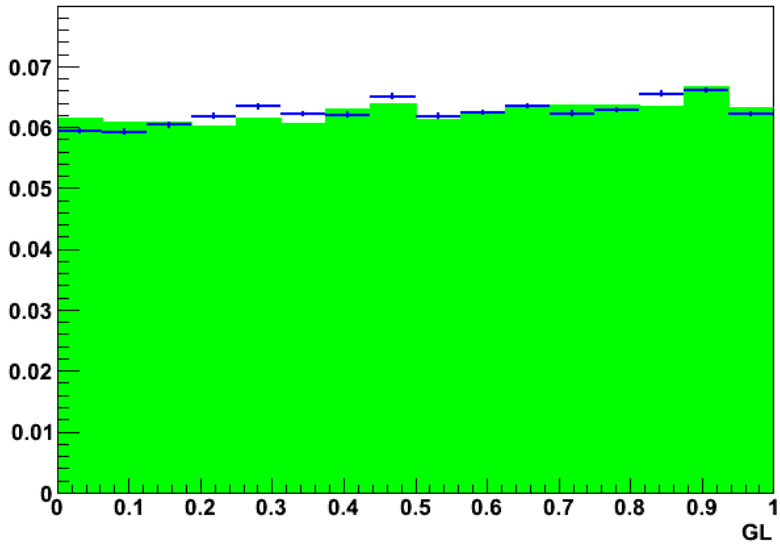

Figure 4.19: GL for $B \to h^+h'^-$ passing the common selection, without including further cuts (green filled histogram), and requiring in addition both decay products to satisfy $\Delta\mathrm{LL}_{p-\pi} < 30$ (blue histogram). All histograms are normalized to unity.

Chapter 5

Correction of reconstruction, PID and trigger effects

This chapter presents the methods developed for correcting the effects introduced by track reconstruction, PID requirements and trigger in the calibration and normalization of the measurement of $B_s^0 \to \mu^+\mu^-$ using $B \to h^+h'^-$. Section 5.1 describes a method for estimating the ratio between the reconstruction efficiencies of signal and control channels. Section 5.2 shows the procedure used for avoiding the biases introduced by the PID requirements in the calibration of the IM of $B_s^0 \to \mu^+\mu^-$. Finally, Section 5.3 explains a generic method for correcting the inefficiencies and biases due to the trigger.

5.1 Estimation of the ratio of reconstruction efficiencies

The fact that differences between $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ due to track reconstruction introduce basically a global normalization factor implies that such effects are not relevant for the calibration of the IM and the GL. However this difference needs to be considered for the normalization of $\mathrm{BR}(B_s^0 \to \mu^+\mu^-)$, see Eq. 4.1.

The ratio between the reconstruction efficiency of $B \to h^+h'^-$ and $B_s^0 \to \mu^+\mu^-$ can be estimated by using the decays $B^+ \to J/\psi(\mu^+\mu^-)K^+$ and $B^+ \to \bar{D}^0(K^+\pi^-)\pi^+$, and making the following assumptions:

- The track reconstruction efficiency for pions and kaons is the same.
- The reconstruction efficiency of each channel can be factored as the product of the track reconstruction efficiency of each decay product,

  $$\epsilon_{rec,S} = \epsilon_{rec,\mu}^2, \quad \epsilon_{rec,C} = \epsilon_{rec,h}^2,$$

  hence

  $$\frac{\epsilon_{rec,C}}{\epsilon_{rec,S}} = \left(\frac{\epsilon_{rec,h}}{\epsilon_{rec,\mu}}\right)^2.$$

This is justified by considering that the track reconstruction efficiency is practically independent of the kinematical properties of the particles, as discussed in Subsection 4.1.1.

Under these assumptions, the ratio $\epsilon_{rec,h}/\epsilon_{rec,\mu}$ can be estimated by means of the two channels introduced above by means of the ratio

$$\frac{n_{S'}}{n_{C'}} = \frac{\epsilon_{rec,S'}\epsilon_{sel,S'}\epsilon_{trig,S'}}{\epsilon_{rec,C'}\epsilon_{sel,C'}\epsilon_{trig,C'}} \frac{f_{S'}}{f_{C'}} \frac{BR_{S'}}{BR_{C'}} \tag{5.1}$$

where the subindices S' and C' refer to the channels $B^+ \to J/\psi(\mu^+\mu^-)K^+$ and $B^+ \to \bar{D}^0(K^+\pi^-)\pi^+$ respectively. The ratio $f_{S'}/f_{C'}$ is 1 as both channels involve the same B meson, while the offline selection can be built in order to have similar effects on both channels. In this case, by rearranging Eq. 5.1 the following expression is obtained:

$$\frac{\epsilon_{rec,S'}}{\epsilon_{rec,C'}} = \frac{\epsilon_{trig,C'}BR_{C'}}{\epsilon_{trig,S'}BR_{S'}} \frac{n_{S'}}{n_{C'}}. \tag{5.2}$$

By factoring the left hand side using the reconstruction efficiency of each decay product, this expression reads

$$\left(\frac{\epsilon_{rec,\mu}}{\epsilon_{rec,h}}\right)^2 = \frac{\epsilon_{trig,C'}BR_{C'}}{\epsilon_{trig,S'}BR_{S'}} \frac{n_{S'}}{n_{C'}}, \tag{5.3}$$

which is the required ratio. The ratio $BR_{C'}/BR_{S'}$ is known from the current measurements with an accuracy of 4.8% [1], while $n_{S'}$ and $n_{S'}$ are observable quantities. Finally, the ratio of trigger efficiencies $\epsilon_{trig,C'}/\epsilon_{trig,S'}$ can be estimated by means of the methods explained in Section 5.3.

5.2 Correction of the bias due to the PID requirements in the calibration of the invariant mass

As discussed in Subsection 4.3.1, the calibration of the IM of $B_s^0 \to \mu^+\mu^-$ using $B \to h^+h'^-$ requires the use of PID requirements. The biases in the IM distribution due to such requirements strongly depend on the performance of the RICH, and this motivates the development of a robust method for calibrating this distribution.

The biases due to the inefficiency of PID cuts on a given $B \to h^+h'^-$ mode could be corrected by weighting each event by the inverse of the efficiency of such cuts provided by the generic calibration of the RICH. However, this requires to select the specific $B \to h^+h'^-$ channel with high purity in order to avoid weighting events from other $B \to h^+h'^-$ modes. The thresholds that are necessary for that purpose highly suppress the part of the phase space of the decay products of high momenta, as seen in Fig. 4.13. In this case the contribution from $B \to h^+h'^-$ with high momentum products can not be recovered by this method as there are no events left for weighting, hence the bias due to the PID requirements can not be fully corrected.

This motivates the use PID cuts that preserve the whole phase space of the decay products. Following the results displayed in Fig. 4.13, the choice is to set both thresholds c_π and c_K equal to 2.5 and 0.5 for hadrons below and above 50 GeV respectively. Under these conditions, the biases introduced in the IM calibration are not due only to the inefficiency of the PID cuts but also to the contamination from other $B \to h^+h'^-$ modes. The correction of these two effects is addressed in the following.

The bias in the IM distribution for a particular $B \to h^+h'^-$ channel due to the momentum dependence of the PID efficiency can be handled by binning in the momentum of the decay products. In each such bin, the PID efficiency is the same for all the events. Then the full IM distribution should be recovered by adding consecutively the respective contribution from each bin, scaled according to the relative contribution of such bin to the total number of events of the chosen $B \to h^+h'^-$ mode in absence of PID cuts. This scaling factor can be extracted by taking into account that the momentum distributions of all $B \to h^+h'^-$ modes are similar, hence it can be estimated as the fraction of all $B \to h^+h'^-$ events falling in that bin without applying PID cuts. As the inefficiency of the PID biases affects mostly high momentum hadrons, and the probability of having both hadrons from $B \to h^+h'^-$ of high momentum is low, it is sufficient to bin only in the maximum momentum of the hadron pair.

The method aiming to extract the IM distribution in each bin of the maximal momentum of the decay products needs to handle the background due to the cross-feed from $B \to h^+h'^-$. The IM distribution of this background for the selection of the $B^0 \to K^+\pi^-$ and $B_s^0 \to K^+K^-$ modes is displayed in Fig. 5.1. Note that, if both distributions are centered at the same mass value, the respective intervals of the IM range that contain the bulk of this background do not overlap. In particular, if the mass of the decaying B meson is used as reference, the selection of $B^0 \to K^+\pi^-$ and $B_s^0 \to K^+K^-$ is very pure in the mass intervals [-60,0] and [-20,60] respectively. By assuming that in each bin of the maximal momentum of the decay products the IM distribution is described by Crystal-Ball (CB) functions [118], their parameters can be determined from such intervals alone as follows: first the width of the gaussian core can be extracted from $B_s^0 \to K^+K^-$; then the two parameters of the radiative tail can be easily extracted from $B^0 \to K^+\pi^-$, by assuming the width already obtained with $B_s^0 \to K^+K^-$.

The result of calibrating the IM distribution as explained above is shown in Fig. 5.2, for four bins in the maximum of the decay products defined by the intervals [0,80], [80,120], [120,160] and [160,∞] GeV respectively. In each bin the width of the gaussian core and the two parameters of the radiative tail are extracted from the fit, while the center of the gaussian core is kept fixed. For each decay mode, the value of this parameter is estimated from events having both decay products with momenta below 50 GeV and belonging to the interval $[M_{B^0} - 30, M_{B^0} + 30]$ for $B^0 \to K^+\pi^-$ and $[M_{B_s} - 20, M_{B_s} + 40]$ for $B_s^0 \to K^+K^-$, by fitting their IM to a gaussian distribution. Note that in this case the requirements on the momentum of the decay products ensure that the biases due to both the PID inefficiency and the

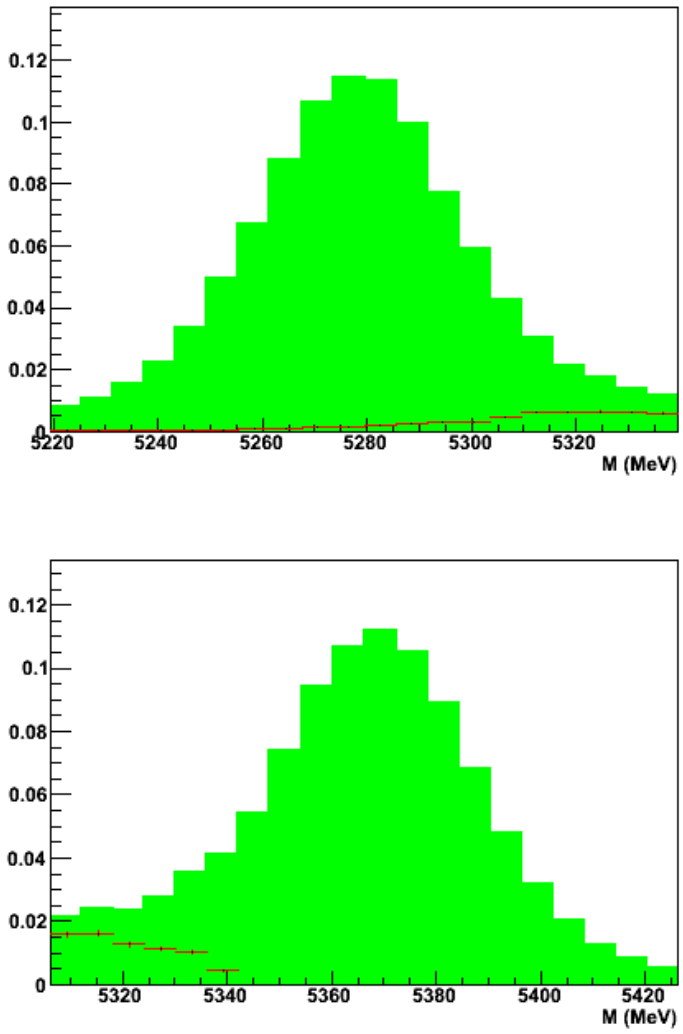

Figure 5.1: IM for $B \to h^+h'^-$ selected as $B^0 \to K^+\pi^-$ and $B_s^0 \to K^+K^-$ (top and bottom green filled histograms respectively), and for the respective background due to the cross-feed from $B \to h^+h'^-$ (red histograms). Events are required to satisfy the common selection. Signal histograms are normalized to unity, while background histograms are normalized respecting the ratio of yields.

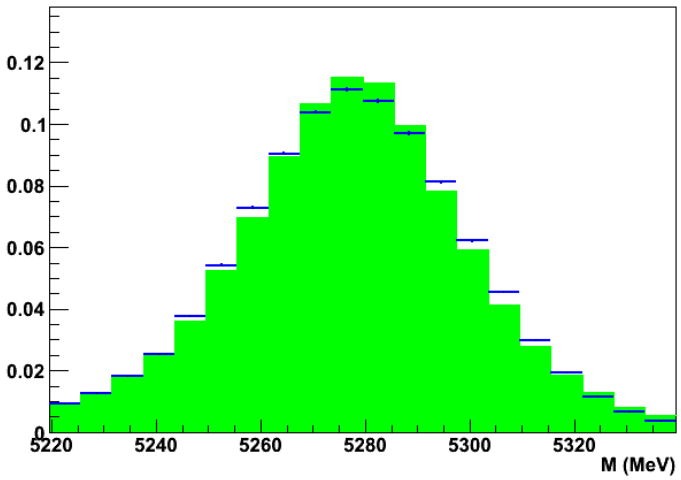

Figure 5.2: IM for the $B^0 \to K^+\pi^-$ mode (green filled histogram), using for each decay product the correct mass from the generator. The cross-feed contribution from $B \to h^+h'^-$ is not included. The IM obtained by fitting the IM distribution of $B^0 \to K^+\pi^-$ and $B_s^0 \to K^+K^-$ in ranges of high purity is also shown (blue histogram). In each bin of the maximum momentum of the decay products the fit assumes a simple CB distribution. All histograms are normalized to unity.

background from the other $B \to h^+h'^-$ modes are low.

The residual biases arising in Fig. 5.2 are due to the fact that the IM resolution depends on the momentum of the decay products, hence even in absence of PID cuts the total IM distribution is the convolution of CBs with different parameters. By increasing the binning in the momentum of the decay products this bias would decrease. However, this effect is handled effectively by assuming that in a given bin the IM distribution consist of the sum of two CBs differing only in the width of the gaussian core, and leaving free the relative contribution of each one. The result of this improvement is displayed in Figs. 5.3 and 5.4, showing that the residual biases are now corrected.

5.3 Correction of inefficiencies and biases due to the trigger

The fact that a fraction of signal events pass the trigger selection because of particles not related to signal allows to correct the effect of the trigger on signal properties. To that purpose, it is convenient to use the following classification of events, according to the relation between the reconstructed signal candidate and the information used to trigger:

- Triggered on signal (TOS): events in which signal particles are sufficient to trigger.

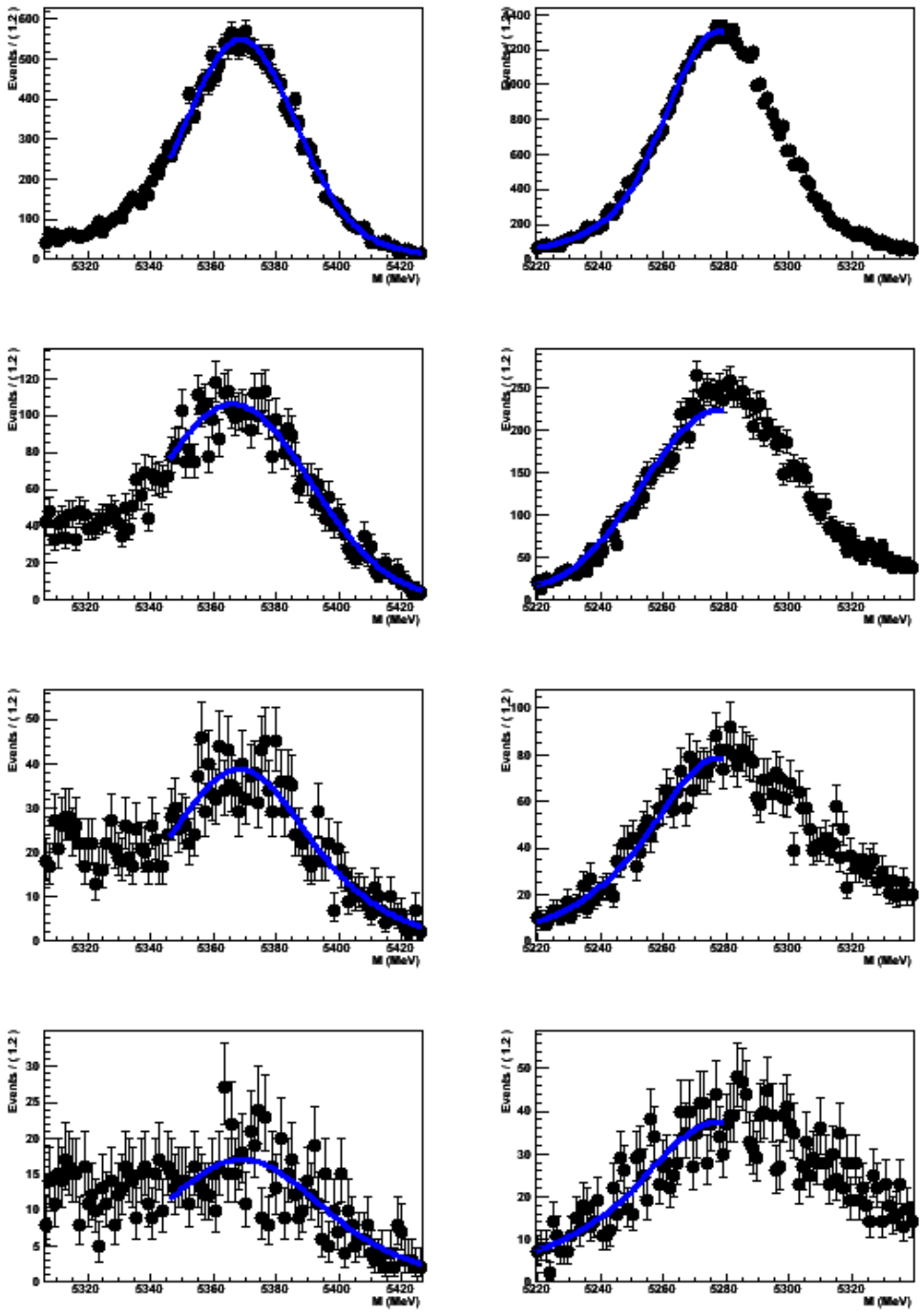

Figure 5.3: Result of the fit of the IM to a sum of two CB distributions in the range of high purity only, for $B_s^0 \to K^+K^-$ (left) and $B^0 \to K^+\pi^-$ (right), and in four successive bins of the maximum momentum of the decay products (from top to bottom). These bins are defined by the intervals [0,80], [80,120], [120,160] and [160,∞] GeV.

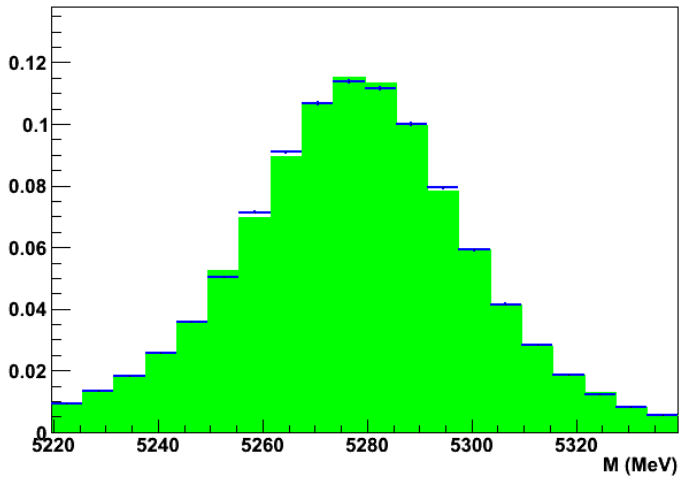

Figure 5.4: IM for the $B^0 \to K^+\pi^-$ mode (green filled histogram), using for each decay product the correct mass from the generator. The cross-feed contribution from $B \to h^+h'^-$ is not included. The IM obtained by fitting the IM distribution of $B^0 \to K^+\pi^-$ and $B_s^0 \to K^+K^-$ in ranges of high purity is also shown (blue histogram). In each bin of the maximum momentum of the decay products the fit assumes a the addition of two CB distributions. All histograms are normalized to unity.

- Triggered independent of signal (TIS): events in which non-signal particles are sufficient to trigger.
- Trigger on both (TOB): events in which both signal and non-signal particles are necessary to trigger.

The LHCb trigger system records all the information needed for such classification. Note that TOS and TIS categories are not exclusive and that, by definition, TOB events can never be TIS or TOS, and viceversa. TOS and TIS efficiencies are defined with respect to the number of reconstructed events passing the offline selection n_{sel}, as

$$\epsilon_{TOS} \equiv \frac{n_{TOS}}{n_{sel}}, \quad \epsilon_{TIS} \equiv \frac{n_{TIS}}{n_{sel}}, \tag{5.4}$$

where n_{TOS} and n_{TIS} are the number of events that are classified as TOS and TIS respectively, in addition to passing the offline selection. Note that the definition of both TOS and TIS efficiencies is analogous to the definition of trigger efficiency given by Eq. 3.7.

Signal properties are related to those of non-signal particles only through the three-momentum of the signal B meson. On average, a hard signal is associated with hard signatures in the rest of the event. This allows to develop methods for computing relative and absolute trigger efficiencies, with potential applications in a wide variety of analyses.

5.3.1 Estimation of the ratio of trigger efficiencies

From Eqs. 3.7 and 5.4, the trigger efficiency can be written as

$$\epsilon_{trig} = \epsilon_{TIS} \frac{n_{trig}}{n_{TIS}}. \tag{5.5}$$

Recall that both n^{trig} and n^{TIS} are observable quantities, while ϵ_{TIS} is not. As discussed above, the probability of being TIS is correlated with the signal properties only through the three-momentum of the signal B meson. Hence two samples having the same three-momentum distribution of the signal B meson after the offline selection should have the same value of ϵ_{TIS}. In this case, Eq. 5.5 leads to

$$\frac{\epsilon_{trig,S1}}{\epsilon_{trig,S2}} = \frac{n_{trig,S1}}{n_{TIS,S1}} \frac{n_{TIS,S2}}{n_{trig,S2}}, \tag{5.6}$$

where the subindices $S1$ and $S2$ represent each sample. This result allows to normalize the measurement of $\mathrm{BR}(B_s^0 \to \mu^+\mu^-)$ using $B \to h^+h'^-$ decays involving the B_s^0 meson, as $B_s^0 \to K^+K^-$. In this case the fact that both $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ are required to satisfy the same offline selection would ensure that the three-momentum distribution of the signal B meson is the same for both channels. By normalizing $\mathrm{BR}(B_s^0 \to \mu^+\mu^-)$ with $B_s^0 \to K^+K^-$, the ratio ϵ_C/ϵ_S is found to be 0.53 ± 0.03 using Eq. 5.6, while its true value is 0.535 ± 0.001. Besides, only slight differences are expected between the phase space of B^0 and B_s^0 mesons, for $B \to h^+h'^-$ satisfying the common selection. This allows to estimate the ratio ϵ_C/ϵ_S from Eq. 5.6 by using all the $B \to h^+h'^-$ modes. In this case the result is found to be 0.51 ± 0.03 , while the true value is 0.523 ± 0.001.

5.3.2 Absolute estimation of the trigger efficiency

The fact that the probability of being TIS is correlated with the signal properties only through the three-momentum of the signal B meson allows to estimate the absolute value of the trigger efficiency. If the phase space of the signal B meson is divided in small enough bins, then in each such bin the correlation between the probability of being TIS and the signal properties can be neglected. In other words, in any bin of the phase space of the signal B meson, any selection of events based on signal properties would not affect the probability of being TIS. A selection of events according to their classification as TOS would not be an exception. Therefore, in each bin, the value of the ratio defining the TIS efficiency in Eq. 5.4 would remain unchanged if calculated on events that, in addition of satisfying the offline selection, are classified as TOS. That is, in the i-th bin of the phase space of the signal B meson,

$$\frac{n^i_{TOS\&TIS}}{n^i_{TOS}} = \frac{n^i_{TIS}}{n^i_{sel}}, \tag{5.7}$$

where n_{TOS} is the number of events passing the offline selection and classified as TOS only, and $n_{TOS\&TIS}$ is the number of events passing the offline selection and

classified as both TOS and TIS. Eq. 5.7 is checked in Fig. 5.5, by showing the values of ϵ_{TIS} and their respective estimations using Eq. 5.7 in bins of p_T of the signal B meson (p_T^B), for $B \to h^+h'^-$. For comparison, the values of the total ϵ_{TIS} and its estimation by means of the ratio $n_{TOS\&TIS}/n_{TOS}$ for $B \to h^+h'^-$ are equal to 0.0116 ± 0.0001 and 0.0133 ± 0.0002 respectively. The difference between these two values is significant, due to the fact that the full phase space of the signal B meson is considered.

By rearranging Eq. 5.7, the value of n^i_{sel} is given as

$$n^i_{sel} = \frac{n^i_{TOS} n^i_{TIS}}{n^i_{TOS\&TIS}}. \tag{5.8}$$

Note that Eq. 5.8 could also be derived in an analogous manner by interchanging the role of the TOS and TIS categories. The value of n_{sel} can thus be obtained by summing the contribution of each bin,

$$n_{sel} = \sum_i n^i_{sel} = \sum_i \frac{n^i_{TOS} n^i_{TIS}}{n^i_{TOS\&TIS}}. \tag{5.9}$$

The last member of this expression contains only observable quantities, and should give an unbiased result in the limit of high statistics and high number of bins in the signal B phase space. Note that substituting this expression in Eq. 3.7 allows measuring the absolute value of ϵ_{trig} for a given channel from data.

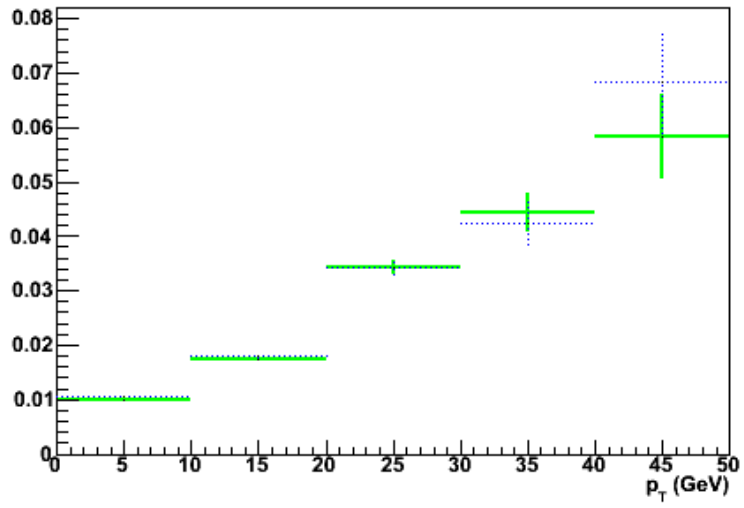

Figure 5.5: TIS efficiency in bins of p_T^B for $B \to h^+h'^-$ events passing the common selection, showing the true value (blue dotted line) and the respective estimations using Eq. 5.7 (green solid line).

The performance of the method has been evaluated on $B \to h^+h'^-$ MC data. For Eq. 5.9 to hold exactly, the binning of the phase space of the signal B meson

would have to be performed in the three components of the three-momentum of this particle. Considering the rough azimuthal symmetry of LHCb, the number of bins can be kept reasonably low by binning in the transversal (p_T^B) and longitudinal (p_z^B) components only. Other combinations of variables using the modulus (p^B) and the angle with the beam axis (θ^B) have been indeed tried, yielding slightly worse results.

Bins for which $n^i_{TOS\&TIS} = 0$ lead to singular contributions to n_{sel}, as this quantity enters in the denominators of the last member of Eq. 5.9. In order to allow a large number of bins while avoiding these singularities, we perform a variable-size binning. To that purpose, the p_z^B axis is first bined keeping the number of TOS&TIS events per bin constant, and then each of such bins is bined along the p_T^B axis by also keeping the number of TOS&TIS events per bin constant. Note that this leaves the number of TOS&TIS events per bin constant in each two-dimensional bin.

The performance of the method depends on the number of bins in both axis of the phase space of the signal B meson. A measure of the bias introduced by the method is given by:

$$\delta = \frac{\epsilon^*_{trig} - \epsilon_{trig}}{\epsilon_{trig}}, \tag{5.10}$$

where ϵ^*_{trig} is the value of the trigger efficiency computed using the method. This systematic deviation is shown in Fig. 5.6 for $B \to h^+h'^-$, as a function of the number of bins in p_T^B, and for different number of bins in p_z^B. Note that the bias decreases as the number of bins in the phase space of the signal B meson increases. In particular δ reaches the percent level by using few bins only in p_T^B. The expected statistical uncertainty is computed for 2 fb^{-1}, by means of 2000 MC toy-experiments using values of n^i_{TOS}, n^i_{TIS} and $n^i_{TOS\&TIS}$ built from the Poisson-fluctuated contents of the following exclusive categories:

- Exclusive-TOS (xTOS): events that are TOS, but not TIS.
- Exclusive-TIS (xTIS): events that are TIS, but not TOS.
- TOS and TIS (TOS&TIS): events that are both TIS and TOS.

The use of these exclusive categories avoids inconsistencies due to the fact of fluctuating the contents of overlapping categories. Using only three bins in p_T^B the bias is found to be $\delta = 0.85\%$, and the relative statistical uncertainty in ϵ_{trig} is equal to 0.03% with 2 fb^{-1}.

When large numbers of bins are used both in p_T^B and in p_z^B, new biases arise due to the fact that $n_{TOS\&TIS}$, the smallest number in each term of Eq. 5.9, appears in the denominator. In this situation, downward fluctuations imply larger dispersions than those due to upwards fluctuations, and hence the most probable value is shifted away from the central value. However, this does not represent a problem in the current situation, as a small δ is obtained if the number of bins is kept moderately low.

The validity of the method presented here relies on events in which the trigger decision is either based on the reconstructed signal or on the rest of the event. This

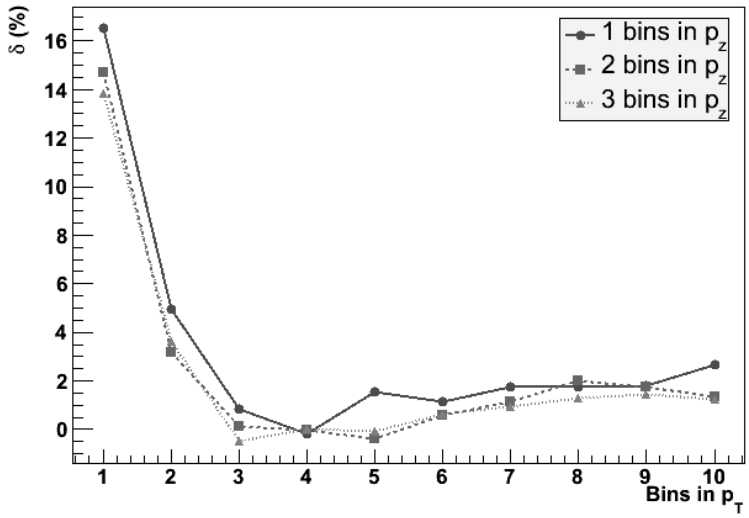

Figure 5.6: Relative bias in the efficiency estimation δ, as defined in the text, as a function of the number of bins in p_T^B and for several numbers of bins in p_z^B, for $B \to h^+h'^-$. Vertical bars showing the dispersion of the bias computed by the different MC toy-experiments assuming 2 fb^{-1} are not displayed as they are smaller than the size of the points.

method formally breaks down if the trigger includes requirements involving both the signal and the rest of the event, called *global event cuts* (GEC). As an example, the classification of an event as TOS is based on whether signal particles are sufficient for triggering. If GEC are used, the answer depends on the rest of the event. At LHC*b*, GEC are present in some configurations of the L0, and are based on the number of primary interactions estimated by the Pile-Up system and the charged track multiplicity estimated by the SPD. The settings of the trigger used in the current study do not contain GEC. If GEC are finally applied, the simplest turn-around to overcome the problem would be to consider as TIS only those events passing L0 lines free from GEC, as the muon lines. This will have the effect of reducing the number of TIS events available by roughly a 40%.

Effect of the background subtraction

The background subtraction is done using IM sidebands. The full size of the IM sidebands is chosen to be equal to the mass window used in the selection, namely 120 MeV in total. In order to compute ϵ_{trig} using the method explained above, the subtraction needs to be applied in each bin of the signal B-meson phase space, and for each of the three trigger categories $c =$ xTOS, xTIS, TOS&TIS:

$$n_c^i = n_{c,MW}^i - n_{c,SB}^i, \tag{5.11}$$

where the subindices MW and SB refer to the mass window and its respective side-

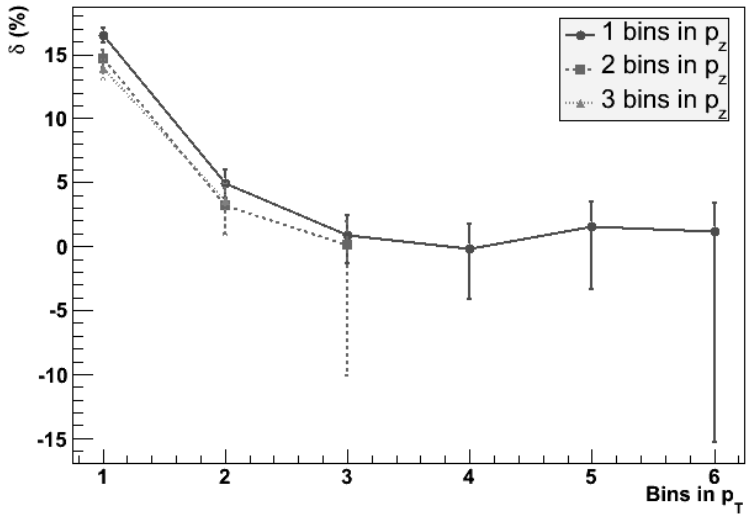

Figure 5.7: Relative bias in the efficiency estimation δ as a function of the number of bins in p_T^B and for several numbers of bins in p_z^B, for $B \to h^+h'^-$, and having added and subtracted background from inclusive $b\bar{b}$. Vertical bars show the dispersion of the bias computed by the different MC toy-experiments, assuming 2 fb^{-1}. The uncertainty on the bias is hence $\sim \sqrt{2000}$ times smaller. Points having vertical bars of size above 20% are not displayed for visual reasons.

bands respectively. The systematic deviation δ including the effect of the background subtraction is shown in Fig. 5.7 for $B \to h^+h'^-$, as a function of the number of bins in p_T^B, and for different number of bins in p_z^B. The expected statistical uncertainty is computed for 2 fb^{-1}, by means of 2000 MC toy-experiments using at each step the following values for $n^i_{c,MW}$ and $n^i_{c,SB}$:

$$\begin{array}{rcl} n^i_{c,MW} & = & \text{Poisson}(s^i_c) + \text{Poisson}(b^i_c), \\ n^i_{c,SB} & = & \text{Poisson}(b^i_c), \end{array} \tag{5.12}$$

where s^i_c and b^i_c are the ammounts of signal and background respectively obtained from MC. Note that this assumes that the background is independent of the IM, as the ammount of background used for $n^i_{c,SB}$ is the same as those used for $n^i_{c,MW}$. For this reason, the results in Fig. 5.7 are independent on the position of each mass sideband with respect to the mass window used in the offline selection. In order to increase the background statistics, the values of b^i_c are obtained by considering events in a mass window of 600 MeV. Using only three bins in p_T^B it is possible to determine ϵ_{trig} for 2 fb^{-1} with a statistical relative precision of 2% and a systematic bias δ equal to 0.85%.

As discussed in Section 4.3.2, a more stringent signal selection can be achieved by requiring GL > 0.25. In this case, using four bins in p_T^B and two bins in p_z^B the relative precision becomes 0.2% and the systematic bias δ is equal to 1.2%, as shown in Fig. 5.8.

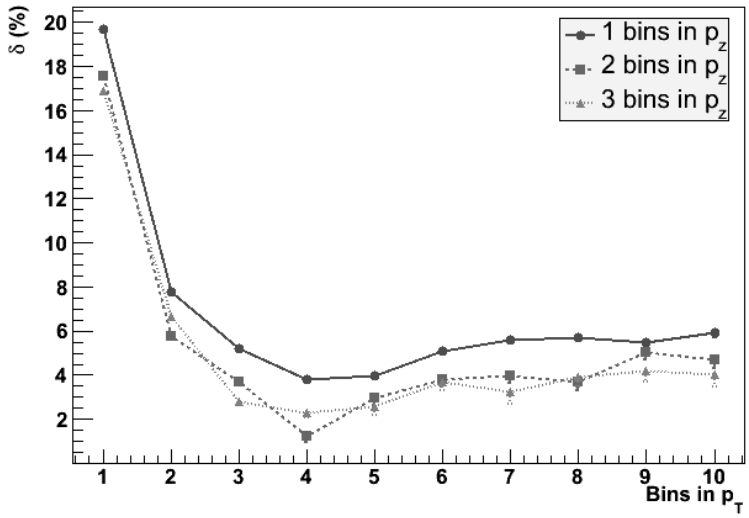

Figure 5.8: Relative bias in the efficiency estimation δ as a function of the number of bins in p_T^B and for several numbers of bins in p_z^B, for $B \to h^+h'^-$ satisfying GL > 0.25, and having added and subtracted background from inclusive $b\bar{b}$. Vertical bars show the dispersion of the bias computed by the different MC toy-experiments, assuming 2 fb^{-1}. The uncertainty on the bias is hence $\sim \sqrt{2000}$ times smaller.

5.3.3 Correction of trigger biases from the distribution of signal properties

The formalism developed in Subsection 5.3.2 can be applied for correcting the effect of trigger on the observed signal properties for events passing the offline selection. That is, it allows to obtain the distributions of such properties after the offline selection as if signal was passed a fully efficient trigger. In order to illustrate the performance of the method, the following six variables are considered: the minimal p_T; the maximal p_T; the minimal IPS and the maximal isolation of the $B \to h^+h'^-$ products; the DOCA between the two $B \to h^+h'^-$ products; and the GL of the $B \to h^+h'^-$ candidate.

If signal was completely independent of the rest of the event, the distribution of signal properties for TIS events would not be biased by the trigger, as these properties would not be related with the properties used for trigger the event. As explained above, signal and the rest of the event are in fact related through the three-momentum of the signal B meson. For this reason the trigger still introduces significant biases in the distribution of signal properties for TIS events, as shown in Fig. 5.9.

As discussed in Subsection 5.3.2, the correlation between signal and the rest of the event is negligible in small enough bins of the phase space of the signal B meson. Hence in each of such bins the distribution of signal properties for TIS events should

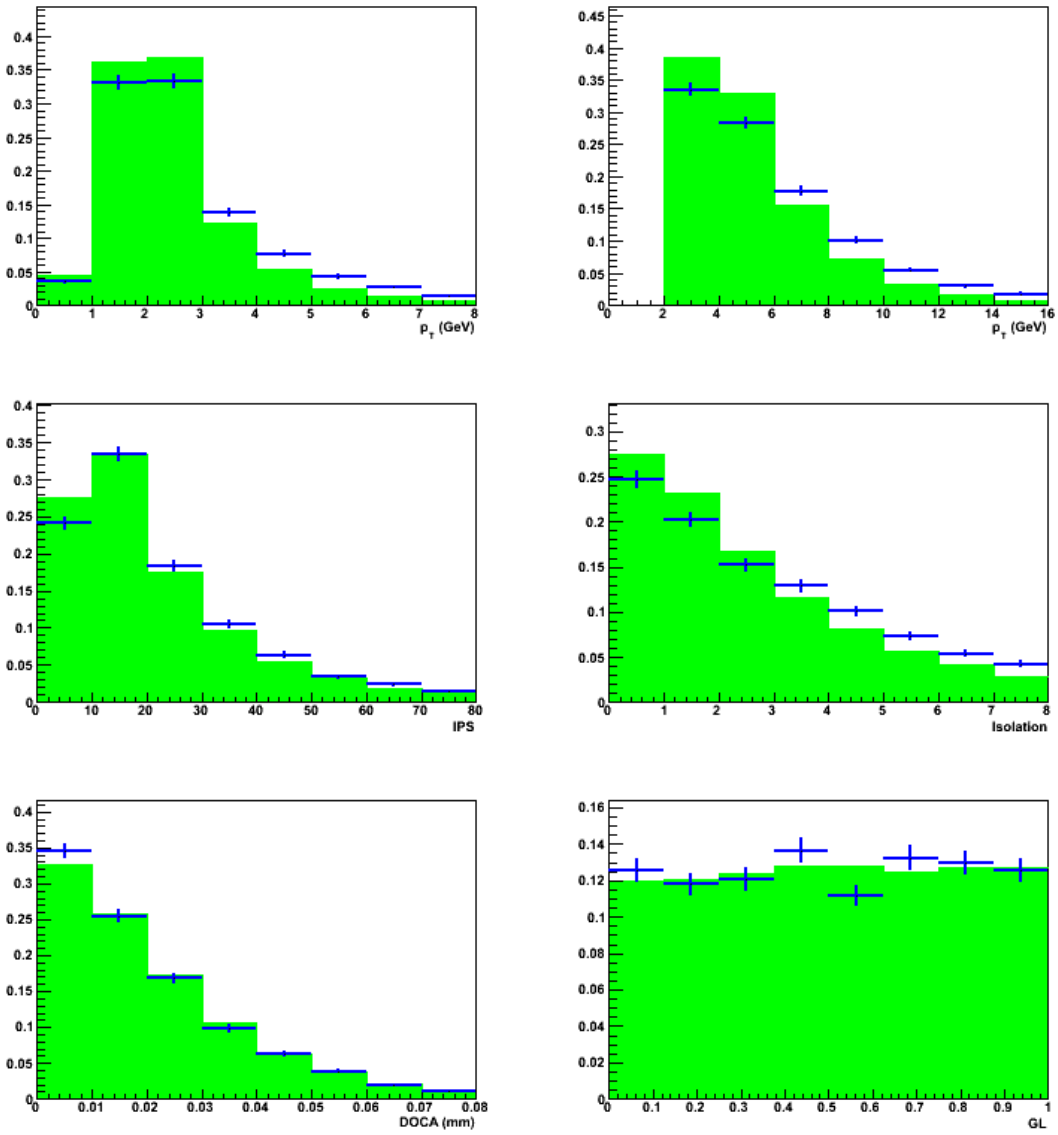

Figure 5.9: From left to right, and from top to bottom: minimal p_T of the $B \to h^+h'^-$ products; maximal p_T of the $B \to h^+h'^-$ products; minimal IPS of the $B \to h^+h'^-$ products; maximal isolation of the $B \to h^+h'^-$ products; DOCA between the two $B \to h^+h'^-$ products; and GL of the $B \to h^+h'^-$ candidate. Their distribution is shown for events passing the common selection only (green filled histogram) and for events being also classified as TIS (blue histogram).

not be affected by any bias due to the trigger. In addition, in each bins the value of ϵ_{TIS} can be estimated by means of Eq. 5.7, that is,

$$\epsilon_{TIS} \equiv \frac{n^i_{TIS}}{n^i_{sel}} = \frac{n^i_{TOS\&TIS}}{n^i_{TOS}}. \tag{5.13}$$

Note that the effect of the trigger in each bin of the phase space of the signal B meson is only to scale the distribution of any signal property in that bin by the global factor ϵ^i_{TIS}, without introducing any bias. Hence, the probability density function (PDF) of a given signal property x as if no trigger requirements were applied, denoted by $f_{sel}(x)$, can be obtained as

$$\begin{aligned} f_{sel}(x) &= C_{sel} \sum_i \frac{1}{\epsilon^i_{TIS}} f^i_{TIS}(x) \\ &= C^*_{sel} \sum_i \frac{n^i_{TOS}}{n^i_{TOS\&TIS}} f^i_{TIS}(x), \end{aligned} \tag{5.14}$$

where $f^i_{TIS}(x)$ is the PDF of the signal property x for TIS events in the i-th bin of the phase space of the signal B meson, and C_{sel} and C^*_{sel} are normalization coefficients defined as

$$C_{sel} = \left(\sum_i \frac{1}{\epsilon^i_{TIS}} \right)^{-1},$$

$$C^*_{sel} = \left(\sum_i \frac{n^i_{TOS}}{n^i_{TOS\&TIS}} \right)^{-1}.$$

The last member of Eq. 5.14 contains only observable quantities, and should give an unbiased result in the limit of high statistics and high number of bins in the signal B phase space.

For testing the method, the same procedure for dividing the phase space of the signal B meson in bins as in Subsection 5.3.2 is used. The validity of the assumptions of the method is illustrated by Fig. 5.10 by using three bins in p^B_T and a single bin in p^B_z. The statistical uncertainty is estimated for 2 fb^{-1} by fluctuating the content of the xTOS, xTIS and TOS&TIS categories in each bin of the phase space of the signal B meson and in each bin of the variable, over 2000 MC toy-experiments. The effect of the trigger is properly corrected in most of the variables.

There is a clear discrepancy for the case of the isolation. The reason is that, as discussed in Section 3.2.5, the definition of this variable correlates particles from the signal B decay with the underlying event, and this is in conflict with the assumptions of the method. That affects also the result for the GL, as isolations are included in the definition of this variable. Alternative isolation criteria are currently being studied in order to minimize this effect. Not that if only signal properties are considered, the method is able to remove the effect of the trigger from all the remaining distributions.

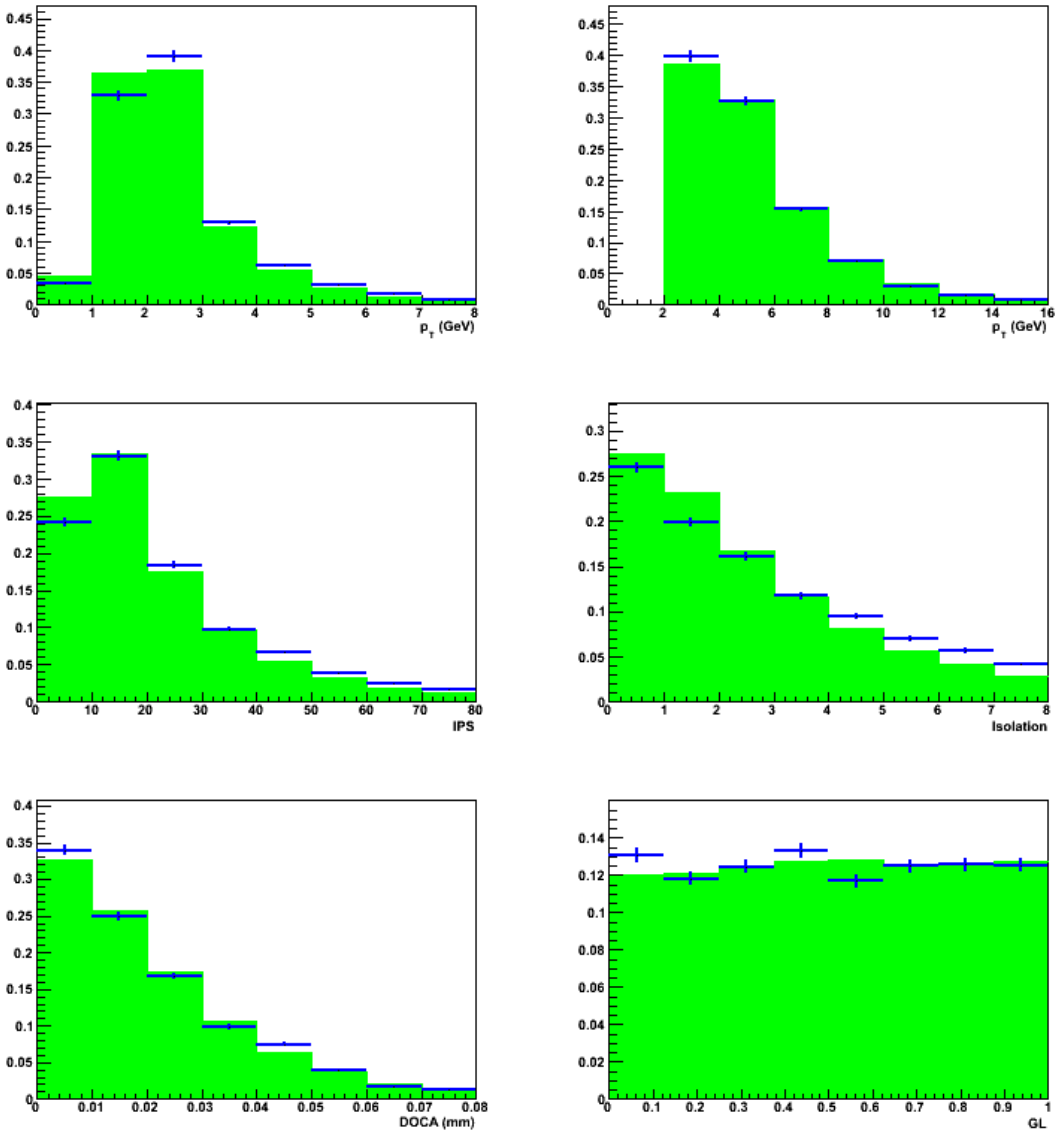

Figure 5.10: From left to right, and from top to bottom: minimal p_T of the $B \to h^+h'^-$ products; maximal p_T of the $B \to h^+h'^-$ products; minimal IPS of the $B \to h^+h'^-$ products; maximal isolation of the $B \to h^+h'^-$ products; DOCA between the two $B \to h^+h'^-$ products; and GL of the $B \to h^+h'^-$ candidate. Their distribution is shown for events passing the common selection only (green filled histograms) and for events being also classified as TIS and weighted by the factor $n^i_{TOS}/n^i_{TOS\&TIS}$ (blue histograms). The error bars show the statistical uncertainty expected for 2 fb^{-1} in absence of background.

The method discussed above can still be used for variables involving non-signal properties if few slight modifications are introduced. As the failure of this method is due to the fact that the variable that is being calibrated depends on non-signal properties, the most straightforward solution would be to consider bins on such variables. However, in addition to the difficulties that could imply the identification of the proper non-signal variables necessary for such purpose, this could lead to an unfeasible number of bins.

An alternative solution is to consider each bin of the variable that is being calibrated as an independent sample where the method described in Subsection 5.3.2 for estimating n_{sel} can be applied. In this case only the global value of ϵ_{TIS} has to be estimated in each bin of such variable. Note that the fact of requiring events lying in a given bin of this variable can be considered as an additional selection requirement. This does not invalidate the basic condition imposed by Eq. 5.7, hence the use of the method described in Subsection 5.3.2 for computing the value of ϵ_{TIS} in such bin is legitimate. The result of correcting the biases due to the trigger from distributions by means of this procedure is shown in Fig. 5.11.

Effect of the background subtraction

The effect of background subtraction is studied by applying the same procedure as in Subsection 5.3.2 in each bin of the variable whose distribution needs to be corrected. Due to limited background statistics, some of the bins of the phase space of the signal B meson are empty for the inclusive $b\bar{b}$ sample. For this reason each trigger category c = xTOS, xTIS, TOS&TIS is assumed to have background distributions for the phase space variables and the variable that is being calibrated equal to those for the total background sample. In particular, the background content in each bin is estimated as follows: first, the global fraction of background classified as c, called f_c, is computed; second, the total ammount of background falling in each bin i of the phase space of the signal B meson and in each bin j of the variable that is being calibrated, denoted as n^{ij}, is found; with these values, for each bin i of the phase space of the signal B meson and for each bin j of the variable that is being calibrated, the background content is set to $f_c n^{ij}$. For signal and background, the content of the xTOS, xTIS and TOS&TIS categories, in each bin of the phase space of the signal B meson and in each bin of the variable, are fluctuated over 2000 MC toy-experiments assuming 2 fb^{-1}. The performance of the method including the effect of the background subtraction is shown in Fig. 5.12.

The effect of the background subtraction has also been checked for the alternative method discussed above oriented to both signal and non-signal properties, using 2000 MC toy-experiments. The contents used at each step are obtained using the procedure explained above. The respective results are shown in Fig. 5.13.

The effects of the background shown in Figs. 5.12 and 5.13 can be largely reduced by using only events with GL>0.25 for measuring BR($B_s^0 \to \mu^+\mu^-$). The statistical uncertainties expected in this case for the two methods studied above are shown in Figs. 5.14 and 5.15, by using four bins in p_T^B and two bins in p_z^B.

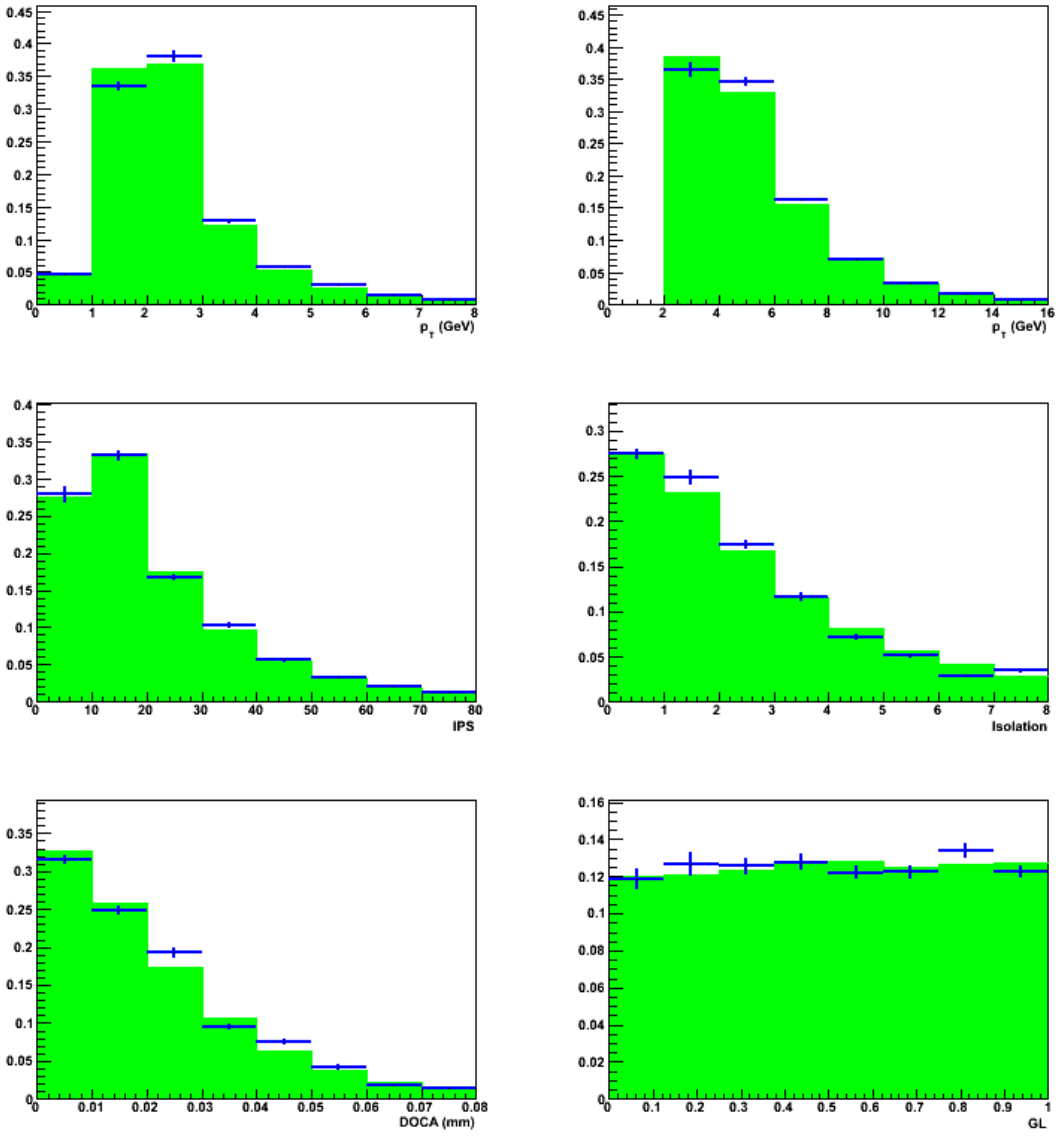

Figure 5.11: From left to right, and from top to bottom: minimal p_T of the $B \to h^+h'^-$ products; maximal p_T of the $B \to h^+h'^-$ products; minimal IPS of the $B \to h^+h'^-$ products; maximal isolation of the $B \to h^+h'^-$ products; DOCA between the two $B \to h^+h'^-$ products; and GL of the $B \to h^+h'^-$ candidate. Their distribution is shown for events passing the common selection only (green filled histograms) and for events for which also Eq. 5.9 has been applied in each bin of the variable (blue histograms). The error bars show the statistical uncertainty expected for 2 fb^{-1} in absence of background.

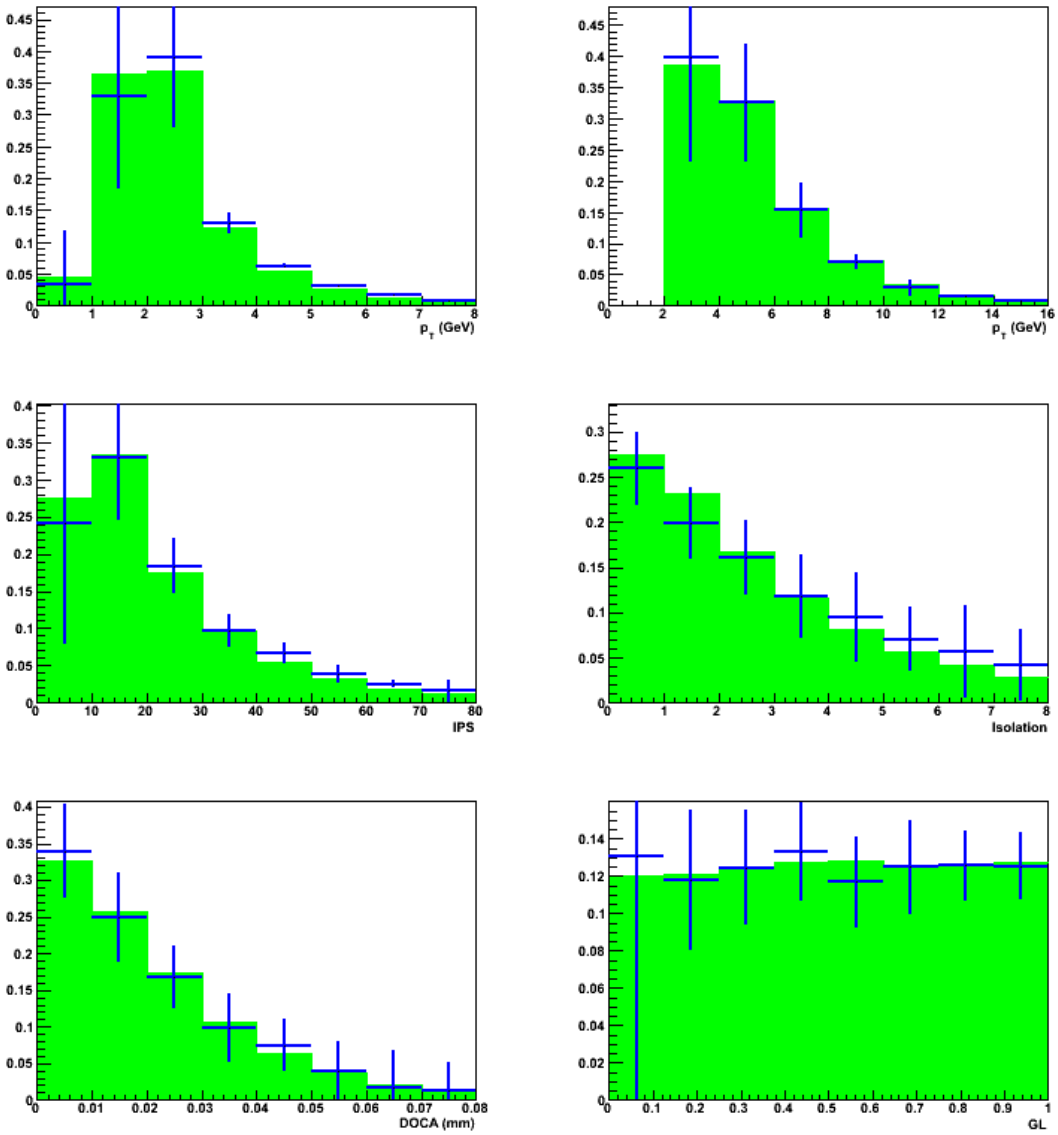

Figure 5.12: From left to right, and from top to bottom: minimal p_T of the $B \to h^+h'^-$ products; maximal p_T of the $B \to h^+h'^-$ products; minimal IPS of the $B \to h^+h'^-$ products; maximal isolation of the $B \to h^+h'^-$ products; DOCA between the two $B \to h^+h'^-$ products; and GL of the $B \to h^+h'^-$ candidate. Their distribution is shown for events passing the common selection only (green filled histograms) and for events being also classified as TIS and weighted by the factor $n^i_{TOS}/n^i_{TOS\&TIS}$ (blue histograms). The error bars show the statistical uncertainty expected for 2 fb^{-1} after the background subtraction, while the central values are computed assuming no background.

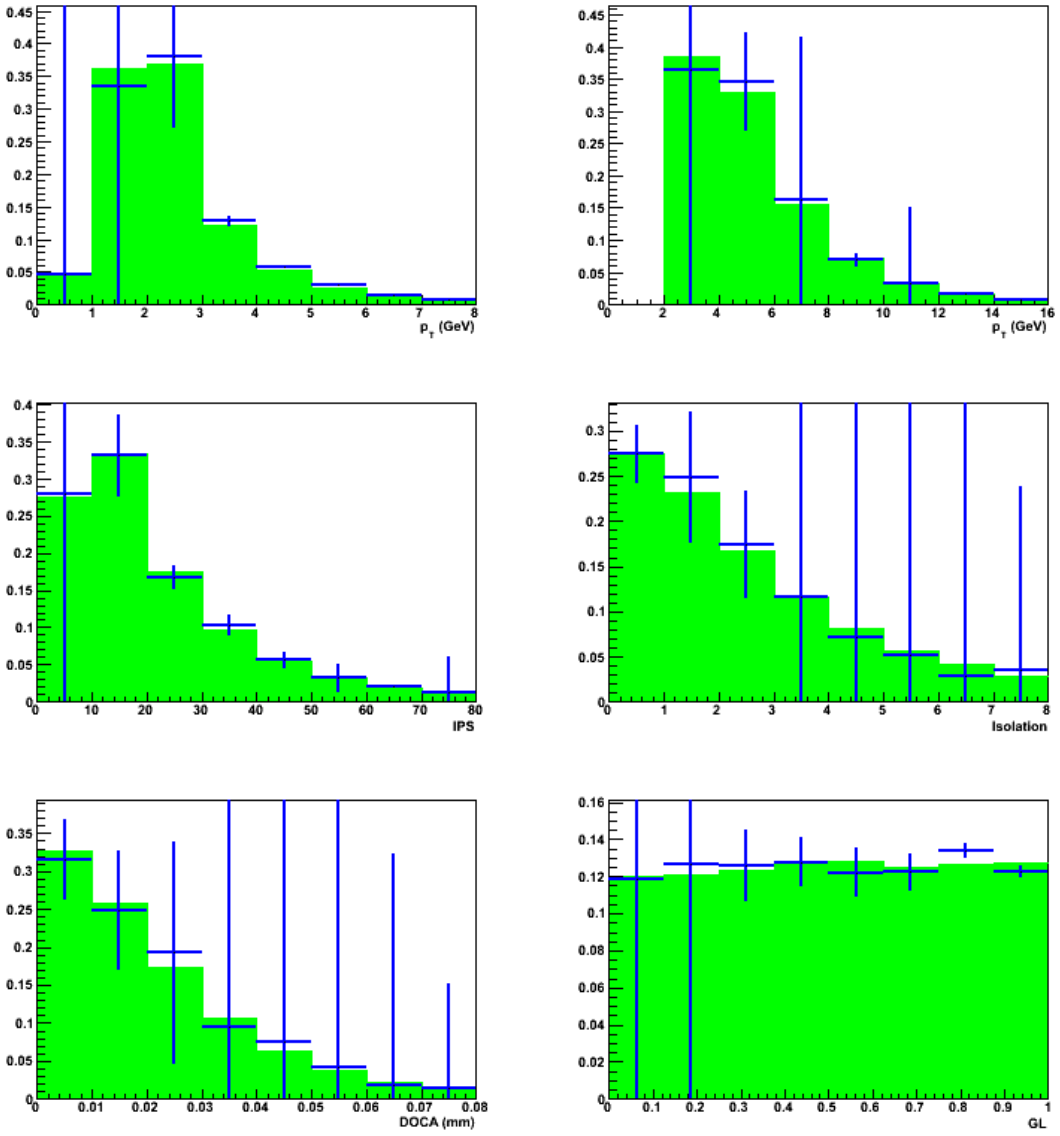

Figure 5.13: From left to right, and from top to bottom: minimal p_T of the $B \to h^+h'^-$ products; maximal p_T of the $B \to h^+h'^-$ products; minimal IPS of the $B \to h^+h'^-$ products; maximal isolation of the $B \to h^+h'^-$ products; DOCA between the two $B \to h^+h'^-$ products; and GL of the $B \to h^+h'^-$ candidate. Their distribution is shown for events passing the common selection only (green filled histograms) and for events for which also Eq. 5.9 has been applied in each bin of the variable (blue histograms). The error bars show the statistical uncertainty expected for 2 fb^{-1} after the background subtraction, while the central values are computed assuming no background.

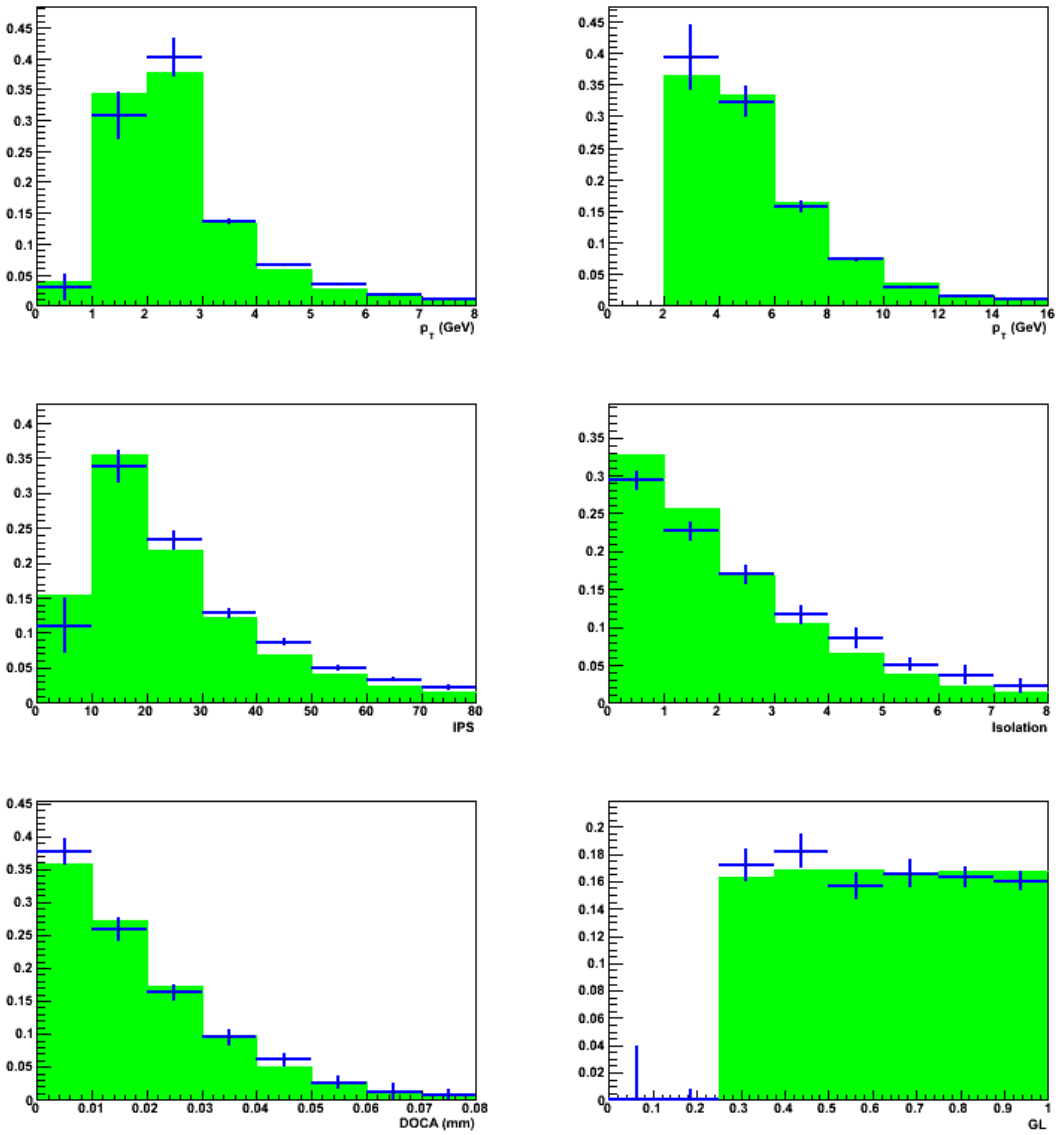

Figure 5.14: From left to right, and from top to bottom: minimal p_T of the $B \to h^+h'^-$ products; maximal p_T of the $B \to h^+h'^-$ products; minimal IPS of the $B \to h^+h'^-$ products; maximal isolation of the $B \to h^+h'^-$ products; DOCA between the two $B \to h^+h'^-$ products; and GL of the $B \to h^+h'^-$ candidate. Their distribution is shown for events passing the common selection and satisfying GL>0.25 only (green filled histograms) and for events being also classified as TIS and weighted by the factor $n^i_{TOS}/n^i_{TOS\&TIS}$ (blue histograms). The error bars show the statistical uncertainty expected for 2 fb^{-1} after the background subtraction, while the central values are computed assuming no background.

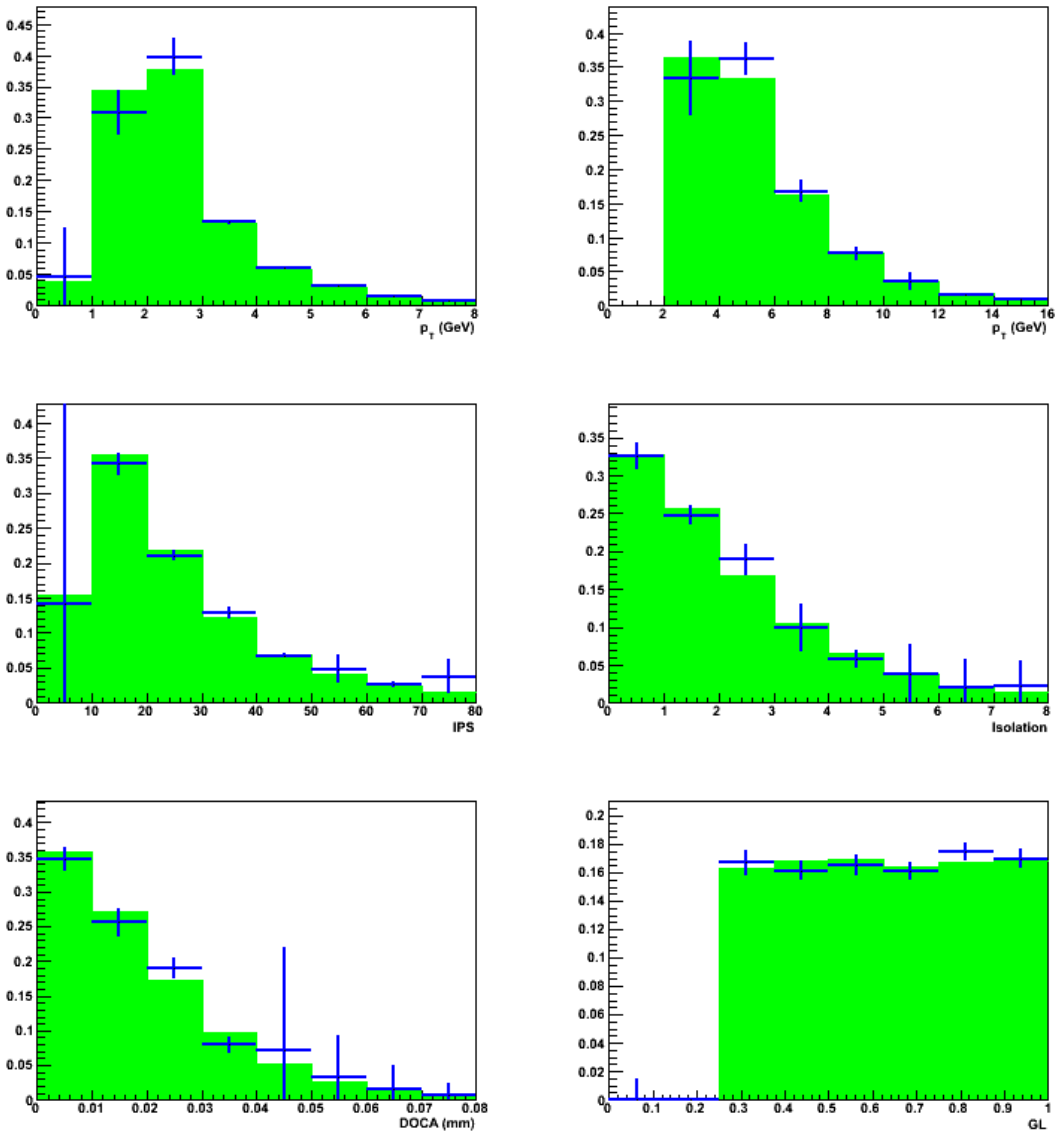

Figure 5.15: From left to right, and from top to bottom: minimal p_T of the $B \to h^+h'^-$ products; maximal p_T of the $B \to h^+h'^-$ products; minimal IPS of the $B \to h^+h'^-$ products; maximal isolation of the $B \to h^+h'^-$ products; DOCA between the two $B \to h^+h'^-$ products; and GL of the $B \to h^+h'^-$ candidate. Their distribution is shown for events passing the common selection and satisfying GL>0.25 only (green filled histograms) and for events for which also Eq. 5.9 has been applied in each bin of the variable (blue histograms). The error bars show the statistical uncertainty expected for 2 fb^{-1} after the background subtraction, while the central values are computed assuming no background.

Chapter 6

Results on BR($B_s^0 \to \mu^+\mu^-$) with the data from 2010 and 2011

The data analyzed by the LHCb collaboration by October 2011 has allowed to set an upper limit for BR($B_s^0 \to \mu^+\mu^-$) [119][120] competitive with that already provided by Tevatron [119][120]. The latest result is based on 300 pb^{-1} of proton-proton interactions at 7 TeV, and includes several concepts developed in the current document.

The measurement of an upper limit for BR($B_s^0 \to \mu^+\mu^-$) at LHCb with 300 pb^{-1} relies on the use of control channels for most calibration and normalization purposes. The particular choices have been $B^+ \to J/\psi(\mu^+\mu^-)K^+$ and $B \to h^+h'^-$, already discussed in Section 3.4, and $B_s^0 \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$. Both $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ are selected by means of a common selection, very similar to that proposed in Subsection 4.2.1, as shown in Table 6.1. Note the presence of some new cuts requiring the momentum of the two muons to be below 1 TeV and the time of flight (TOF) of the B meson candidate to be below five times its mean lifetime respectively. In addition, the p_T of the B meson candidate is required to be above 500 MeV in order to reject the background from elastic diphoton production. The ammount of $B \to h^+h'^-$ collected with 300 pb^{-1} is around 6×10^4 events.

The extraction of an upper limit for BR($B_s^0 \to \mu^+\mu^-$) at LHCb with 300 pb^{-1} uses the CL_s method on the two-dimensional space defined by the IM and the outcome of a boosted decision tree (BDT) combining geometrical and kinematical information. The variables chosen for building the BDT are such that its outcome is practically independent of the IM. They include the information used by the GL plus the minimal p_T of the two muons and two variables related to the isolation and the polarization of the B meson candidate. Two methods are applied for calibrating the IM of $B_s^0 \to \mu^+\mu^-$: the first interpolates linearly the parameters of the IM distribution for the dimuon decays of $c\bar{c}$ and $b\bar{b}$ resonances to the mass of the B_s^0; the second fits the IM distribution of $B \to h^+h'^-$, either considering all its decay modes globally or selecting a single contribution by using the PID information. The mass resolution measured by these methods is displayed in Table 6.2. The outcome of the BDT is calibrated using $B \to h^+h'^-$, selected with no PID requirements. In order to minimize the bias due to hadronic trigger, only events classified as TIS at the L0

Value	Cut
χ^2_{trk}	< 5
$\chi^2_{PV'} - \chi^2_{PV}$	> 12.5
p_{trk}	< 1 TeV/c
DOCA	< 0.3 mm
$\chi^2(B_{d,s})$	< 9
$\|\Delta\mathrm{M}(B_{d,s})\|$	< 600 MeV/c^2
$p_T(B_{d,s})$	> 500 MeV/c
$\chi^2_{PV'} - \chi^2_{PV}$	< 12.5
$\mathrm{DOFS}(B_{d,s})$	> 15
$\mathrm{TOF}(B_{d,s})$	$< 5\tau_{B_s}$
$\Delta z_{PV}(B_{d,s})$	> 0 mm

Table 6.1: Cut values used in the common $B^0_s \to \mu^+\mu^-$ and $B \to h^+h'^-$ selection. For a given track, the quantities $\chi^2_{PV'}$ and χ^2_{PV} refer to the χ^2 of the primary vertex built with and without the considered track respectively. The differerence $\chi^2_{PV'} - \chi^2_{PV}$ quantifies the significance of the track displacement with respect to the primary vertex.

and HLT1 are considered. Events are not required to be classified as TIS at the HLT2 due to the low statistics. The biases due to this last step of the trigger are estimated from MC to be below 5%. The distribution of the outcome of the BDT for $B^0_s \to \mu^+\mu^-$ obtained by this method is shown in Fig. 6.1.

Method	Resolution (MeV/c^2)
Interpolation	$24.6\pm0.2_{stat}\pm1.0_{sys}$
$B \to h^+h'^-$, globally	$23.7\pm0.4_{stat}\pm1.5_{sys}$
$B \to h^+h'^-$, single mode	$23.5\pm0.2_{stat}\pm1.3_{sys}$

Table 6.2: Result of the methods used for calibrating the IM.

The normalization of $\mathrm{BR}(B^0_s \to \mu^+\mu^-)$ using $B \to h^+h'^-$ is based on the selection of a single decay mode, namely $B^0 \to K^+\pi^-$, by using PID requirements. The impact of the trigger is minimized by considering only events classified as TIS at the L0 and HLT1, and applying the following relation obtained from Eqs. 4.1 and 5.4:

$$\mathrm{BR}_S = \frac{\epsilon_{rec+sel,C}}{\epsilon_{rec+sel,S}} \frac{\epsilon_{TIS,C}}{\epsilon_{trig,S}} \frac{n_S}{n_{TIS,C}} \frac{f_{B_C}}{f_{B_s}} \mathrm{BR}_C \equiv \alpha n_S, \tag{6.1}$$

The ratio $\epsilon_{rec+sel,C}/\epsilon_{rec+sel,S}$ is taken from MC as 0.82 ± 0.06. In order to maximize statistics, the value of $\epsilon_{TIS,C}$ is not directly estimated from $B \to h^+h'^-$ but from $B^+ \to J/\psi(\mu^+\mu^-)K^+$ by using Eq. 5.8 with a single bin, obtaining $\epsilon_{TIS,C} = (6.9\pm0.6)\%$. The trigger efficiency for signal is determined by parametrizing the trigger efficiency of dimuons from J/ψ as a function of the higher p_T and IP of the pair, and convoluting the result by the phase space of muons for $B^0_s \to \mu^+\mu^-$. This method leads to $\epsilon_{trig,S} = (89.9\pm0.8^{\mathrm{stat}}\pm4.0^{\mathrm{syst}})\%$. The normalization factor α in

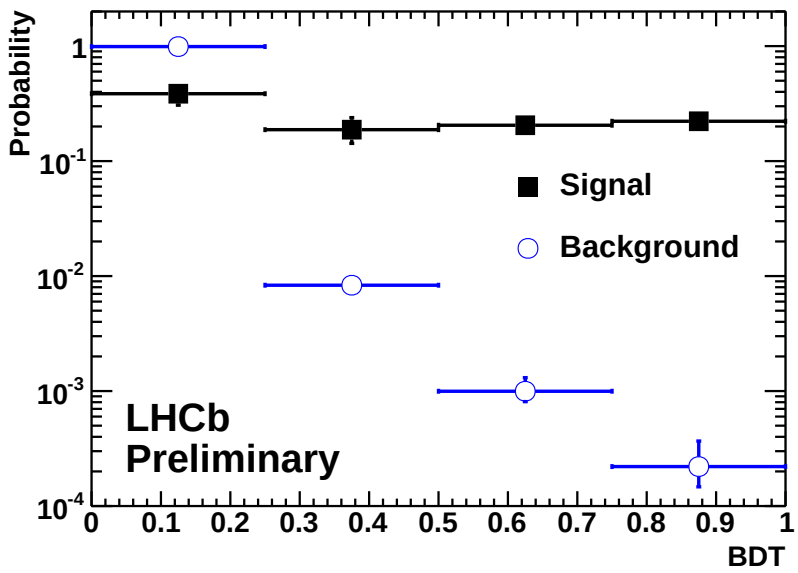

Figure 6.1: Outcome of the BDT for $B \to h^+h'^-$ events classified as TIS at the L0 and HLT1 (solid squares), and for the dimuon background obtained by interpolating the content of the mass sidebands to the mass window of 60 MeV around the mass of the B_s^0 meson (open circles).

Eq. 6.1 obtained by using $B^0 \to K^+\pi^-$ events and considering the above values for $\epsilon_{rec+sel,C}/\epsilon_{rec+sel,S}$, $\epsilon_{TIS,C}$ and $\epsilon_{trig,S}$ is found to be 0.92±0.22. This result is compatible with those obtained using the other two control channels, equal to 0.966±0.096 and 1.27±0.35 for $B^+ \to J/\psi(\mu^+\mu^-)K^+$ and $B_s^0 \to J/\psi(\mu^+\mu^-)\phi(K^+K^-)$ respectively.

The final result of the analysis performed with the first 300 pb^{-1} of data quotes BR($B_s^0 \to \mu^+\mu^-$) $< 1.6\times10^{-8}$ at 95% CL. The combination of this result with that currently provided by CMS sets the upper limit for this BR to 1.08×10^{-8} at 95% CL [121]. These results are already competitive with the latest measurement of BR($B_s^0 \to \mu^+\mu^-$) given by Tevatron, which quotes BR($B_s^0 \to \mu^+\mu^-$) = $1.8\pm^{+1.1}_{-0.9}\times10^{-8}$ [122].

The finer calibration methods presented in this thesis for the calibration of the invariant mass (Section 5.2) and the correction of the trigger effects (Section 5.3) will be used to improve the accuracy of the results when enough statistics is collected. Given the extraordinary performance of the LHC, this is expected to happen at the end of 2011.

Conclusions

The branching ratio of $B_s^0 \to \mu^+\mu^-$ can potentially give access to the first effect from physics beyond the Standard Model at the LHC. The LHCb experiment is particularly well suited to measure this observable because of its excellent vertexing and tracking capabilities, muon system, and dedicated trigger for B physics studies.

In order to rely as little as possible on Monte Carlo simulations, the measurement of the branching ratio of $B_s^0 \to \mu^+\mu^-$ at LHCb makes an intensive use of control channels. In this context, the main source of uncertainty is the fraction f_{B_C}/f_{B_s} entering into the normalization factor, whose current measurement has a systematic uncertainty of 7%, and a theoretical uncertainty of 7%. The family of $B \to h^+h'^-$ decays have wide applications as control channels for this analysis because of the very similar kinematical properties with respect to those of $B_s^0 \to \mu^+\mu^-$. These decays are used for calibrating the invariant mass and the geometrical likelihood, and to normalize the measurement.

The main differences between $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ enter at two levels, namely track reconstruction and trigger. In addition, particle identification introduces some effects that need to be considered in the calibration of the invariant mass. Track reconstruction effects have been found to be practically independent of the kinematical properties, introducing only a global reduction factor around 20% for $B \to h^+h'^-$ with respect to $B_s^0 \to \mu^+\mu^-$ that needs to be considered for normalization. The global trigger efficiency for $B \to h^+h'^-$ is approximately half of that for $B_s^0 \to \mu^+\mu^-$, and the dependence with the kinematical properties is different in each case. This leads to important effects in both the normalization and the calibration of the geometrical likelihood. The offline selection effects can be cancelled by applying common requirements to both signal and control channels. In addition, the particle identification can be used to separate $B \to h^+h'^-$ from its physical background and to select a single $B \to h^+h'^-$ mode for calibrating the invariant mass. The statistical uncertainties introduced by the use of $B \to h^+h'^-$ for controlling the measurement of the branching ratio of $B_s^0 \to \mu^+\mu^-$ can be reduced significantly by excluding from the analysis events with negligible sensitivity.

Several methods have been developed in order to correct the differences between $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$. The effects of the track reconstruction in the normalization can be corrected by estimating the ratio between the reconstruction efficiencies of $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ using $B^+ \to J/\psi(\mu^+\mu^-)K^+$ and $B^+ \to \bar{D}^0(K^+\pi^-)\pi^+$. For 2 fb^{-1}, the uncertainty in the estimation of this ratio is dominated by the 4.8% un-

certainty propagated from the accuracy in the current measurement of the branching ratio of these channels. The calibration of the invariant mass depends on the particle identification performance. This dependence can be minimized by calibrating this variable using relatively tight particle identification requirements, and combining the information from $B^0 \to K^+\pi^-$ and $B_s^0 \to K^+K^-$ in mass intervals where the cross-feed contamination from $B \to h^+h'^-$ is negligible. The systematic effects introduced by calibrating the invariant mass using this procedure are below the percent level. The use of events triggered independently of signal properties allows to correct the effect of the trigger in the normalization and the calibration of the geometrical likelihood. By handling the correlation between signal and the rest of the event properly, such events allow to estimate the ratio of trigger efficiencies with systematic effects around 1% and a statistical uncertainty of 2% for 2 fb^{-1}. Note that these methods for calibrating the invariant mass and for correcting the trigger effects lead to uncertainties that are negligible with respect to those arising from the knowledge of the fraction f_{B_C}/f_{B_s}.

The data collected by the LHCb detector so far allowed to set an upper limit for BR($B_s^0 \to \mu^+\mu^-$) competitive with that already provided by Tevatron [120], based on 300 pb^{-1} of proton-proton interactions at 7 TeV. This measurement used $B \to h^+h'^-$, selected in common with $B_s^0 \to \mu^+\mu^-$, for the normalization of the branching ratio and the calibration of both the invariant mass and geometrical likelihood. Some methods for correcting the differences between $B_s^0 \to \mu^+\mu^-$ and $B \to h^+h'^-$ have not yet been applied due to the low statistics, but it could be possible to use them within the next year.

Bibliography

[1] K. Nakamura *et al.* [Particle Data Group], *Review of Particle Physics*, J. Phys. **G37**, 075021 (2010).

[2] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl.Phys. **22**, 579 (1961).

[3] A. Salam and J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13**, 168 (1964).

[4] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19**, 1264 (1967).

[5] Y. Nambu, *Axial Vector Current Conservation in Weak Interactions*, Phys. Rev. Lett. **4**, 380 (1960).

[6] P. W. Higgs, *Broken Symmetries and Masses of Gauge Bosons*, Phys. Rev. Lett. **13**, 508 (1964).

[7] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13**, 321 (1964).

[8] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13**, 585 (1964).

[9] R. Barate *et al.* [LEP Working Group for Higgs Boson Searches], *Search for the Standard Model Higgs Boson at LEP*, Phys Lett. **B565**, 61 (2003) [hep-ex/0306033].

[10] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10**, 531 (1963).

[11] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interactions*, Prog. Theor. Phys. **49**, 652 (1973).

[12] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8**, 214 (1964).

[13] M. Gell-Mann, *Quarks: Developments in the Quark Theory of Hadrons*, Acta Phys. Austriaca Suppl. **9**, 733 (1972).

[14] H. Fritzsch, M. Gell-Mann and H. Leutwyler, *Advantages of the Colour Octet Gluon Picture*, Phys. Lett. **B47**, 365 (1973).

[15] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*, Phys. Rev. Lett. **30**, 1343 (1973).

[16] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett. **30**, 1346 (1973).

[17] S. Weinberg, *Implications of Dynamical Symmetry Breaking*, Phys. Rev. **D13**, 974 (1976).

[18] E. Gildener, *Gauge Symmetry Hierarchies*, Phys. Rev. **D14**, 1667 (1976).

[19] S. Weinberg, *Implications of Dynamical Symmetry Breaking: An Addendum*, Phys. Rev. **D19**, 1277 (1979).

[20] H. Flacher *et al.* [Gfitter Group], *Revisiting the Global Electroweak Fit of the Standard Model and Beyond with Gfitter*, Eur. Phys. J. **C60**, 543 (2009) [hep-ph/0811.0009].

[21] R. Davis, D. S. Harmer and K. C. Hoffman, *Search for Neutrinos from the Sun*, Phys. Rev. Lett. **20** 1205 (1968).

[22] S. Fukuda *et al.* [Super-Kamiokande collaboration], *Constraints on Neutrino Oscillations Using 1258 Days of Super-Kamiokande Solar Neutrino Data*, Phys. Rev. Lett. **86** 5656 (2001).

[23] Q. R. Ahmad *et al.* [Sudbury Neutrino Observatory collaboration], *Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^8B Solar Neutrinos at the Sudbury Neutrino Observatory*, Phys. Rev. Lett. **87**, 071301 (2001).

[24] R. Mohapatra and G. Senjanovic, *Neutrino Mass and Spontaneous Parity Nonconservation*, Phys. Rev. Lett. **44**, 912 (1980).

[25] J. R. Ellis, J. Hisano, M. Raidal and Y. Shimizu, *New Parametrization of the Seesaw Mechanism and Applications in Supersymmetric Models*, Phys. Rev. **D66**, 115013 (2002).

[26] G. W. Bennet *et al.* [Muon g-2 collaboration], *Final Report of the Muon E821 Anomalous Magnetic Moment Measurement at BNL*, Phys. Rev. **D73**, 072003 (2006) [hep-ex/0602035].

[27] F. Zwicky, *Spectral Displacement of Extra Galactic Nebulae*, Helv. Phys. Acta **6**, 110 (1933).

[28] F. Zwicky, *On the Masses of Nebulae and of Clusters of Nebulae*, Astrophys. J. **86**, 217 (1937).

[29] M. S. Roberts and A. H. Rots, *Comparison of Rotation Curves of Different Galaxy Types*, Astronomy and Astrophysics **26**, 483 (1973).

[30] J. Dunkley *et al.* [WMAP collaboration], *Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Likelihoods and Parameters from the WMAP Data*, Astrophys. J. Suppl. **180**, 306 (2009) [hep-/0803.0586]

[31] G. Bertone, D. Hooper and J. Silk, *Particle Dark Matter: Evidence, Candidates and Constraints*, Phys. Rept. **405**, 279 (2005) [hep-ph/0404175].

[32] A. D. Sakharov, *Violation of CP Symmetry, C Asymmetry, and Baryon Asymmetry of the Universe*, JETP Lett. **5**, 24 (1967).

[33] M. B. Gavela, P. Hernandez, J. Orloff and O. Pene, *Standard Model CP Violation and Baryon Asymmetry*, Mod. Phys. Lett. **A9**, 795 (1994).

[34] R. Rattazzi, *Physics Beyond the Standard Model*, PoS **HEP2005**, 399 (2006) [hep-ph/0607058].

[35] L. J. Hall and M. B. Wise, *Flavor Changing Higgs Boson Couplings*, Nucl. Phys. **B187**, 397 (1981).

[36] P. Fayet, S. Ferrara, *Supersymmetry*, Phys. Rept. **32**, 249 (1977).

[37] N. Sakai, *Naturalness in Supersymmetric GUTs*,

[38] H. P. Nilles, *Supersymmetry, Supergravity and Particle Physics*, Phys. Rept. **110**, 1 (1984).

[39] E. Witten, *Dynamical Breaking of Supersymmetry*, Nucl. Phys. **B188**, 513 (1981).

[40] S. Dimopoulos and H. Georgi, *Softly Broken Supersymmetry and $SU(5)$*, Nucl. Phys. **B193**, 150 (1981).

[41] L. Girardello and M. Grisaru, *Soft Breaking of Supersymmetry*, Nucl. Phys. **B194**, 65 (1982).

[42] P. van Nieuwenhuizen, *Supergravity*, Phys. Rept. **68**, 189 (1981).

[43] G. F. Giudice and R. Ratazzi, *Theories with Gauge-Mediated Supersymmetry Breaking*, Phys. Rept. **322**, 419 (1999) [hep-ph/9801271].

[44] T. Gherghetta, G. F. Giudice and J. D. Wells, *Phenomenological Consequences of Supersymmetrywith Anomaly-Induced Masses*, Nucl. Phys. **B559**, 27 (1999) [hep-ph/9904378].

[45] J. R. Ellis, K. A. Olive and Y. Santoso, *Constraining Supersymmetry*, New J. Phys. **4**, 32 (2002) [hep-ph/0202110].

[46] J. R. Ellis, K. A. Olive, Y. Santoso and V. C. Spanos, *Gravitino Dark Matter in the CMSSM*, Phys. Lett. **B588**, 7 (2004) [hep-ph/0312262].

[47] J. R. Ellis *et al.*, *The Supersymmetric Parameter Space in Light of B-Physics Observables and Electroweak Precision Data*, JHEP **08**, 083 (2007) [hep-ph/0706.0652].

[48] J. R. Ellis, K. A. Olive and Y. Santoso, *The MSSM Parameter Space with Non-Universal Higgs Masses*, Phys. Lett. **B539**, 107 (2002) [hep-ph/0204192].

[49] L. Susskind, *Dynamics of Spontaneous Symmetry Breaking in the Weinberg-Salam Theory*, Phys. Rev. **D20**, 2619 (1979).

[50] E. Eichten and K. D. Kane, *Dynamical Breaking of Weak Interaction Symmetries*, Phys. Lett. **B90**, 125 (1980).

[51] S. Dimopoulos and L. Susskind, *Mass without Scalars*, Nucl. Phys. **B155**, 237 (1979).

[52] V. A. Miransky, M. Tanabashi and K. Yamawaki, *Dynamical Electroweak Symmetry Breaking with Large Anomalous Dimension and t Quark Condensate*, Phys. Lett. **B221**, 177 (1989).

[53] V. A. Miransky, M. Tanabashi and K. Yamawaki, *Is the t Quark Responsible for the Mass of the W and Z Bosons?*, Mod. Phys. Lett. **4**, 1043 (1989).

[54] C. T. Hill, *Topcolor Assisted Technicolor*, Phys. Lett. **B345**, 483 (1995).

[55] N. Arkani-Hamed, S. Dimopoulos and G. Dvali, *The Hierarchy Problem and New Dimensions at a Millimeter*, Phys. Lett. **B436**, 263 (1998).

[56] I. Antoniadis, N. Arkani-Hamed, S. Dimopoulos and G. Dvali, *New Dimensions at a Millimeter to a Fermi and Superstrings at a TeV*, Phys. Lett. **B436**, 257 (1998).

[57] V. A. Rubakov and M. E. Shaposhnikov, *Do We Live Inside a Domain Wall?*, Phys. Lett. **B125**, 136 (1983).

[58] L. Randall and R. Sundrum, *A Large Mass Hierarchy from a Small Extra Dimension*, Phys. Rev. Lett. **83**, 3370 (1999).

[59] L. Randall and R. Sundrum, *An Alternative to Compactification*, Phys. Rev. Lett. **83**, 4690 (1999).

[60] H. Lehmann, K. Symanzik and W. Zimmermann, *On the formulation of Quantized Field Theories*, Nuovo Cimento **1**, 205 (1955).

[61] K. G. Wilson, *Non-Lagrangian Models of Current Algebra*, Phys. Rev. **179**, 1499 (1969).

[62] K. G. Wilson, *Operator Product Expansion and Composite Field Operators in the General Framework of Quantum Field Theory*, Commu. Math. Phys. **24**, 87 (1972).

[63] E. Witten, *Short Distance Analysis of Weak Interactions*, Nucl. Phys. **B122**, 109 (1977).

[64] G. Buchalla, A. Buras and M. Lautenbacher, *Weak Decays Beyond Leading Logarithms*, Rev. Mod. Phys. **68**, 1125 (1996) [hep-ph/9512380].

[65] `http://ckmfitter.in2p3.fr/`.

[66] `http://www.slac.stanford.edu/xorg/hfag/`.

[67] I. Dunietz, R. Fleischer and U. Nierste, *In Pursuit of New Physics with B_s Decays*, Phys. Rev. **D63** 114015 (2001) [hep-ph/0012219].

[68] T. Aaltonen *et al.* [CDF collaboration], *Combination of D0 and CDF Results on $\Delta\Gamma_s$ and the CP-Violating Phase $\beta_s^{J/\psi\phi}$*, CDF/PHYS/BOTTOM/CDFR/9787 (2009).

[69] V. M. Abazov *et al.* [D0 collaboration], *Combination of D0 and CDF Results on $\Delta\Gamma_s$ and the CP-Violating Phase $\beta_s^{J/\psi\phi}$*, D0 Note 5928-CONF (2009).

[70] R. Fleischer, *New Strategies to Extract β and γ from $B_d \to \pi^+\pi^-$ and $B_s \to K^+K^-$*, Phys. Lett. **B459**, 306 (1999) [hep-ph/9903456].

[71] R. Fleischer, *$B_{s,d} \to \pi\pi, \pi K, KK$: Status and Prospects*, Eur. Phys. J. **C52**, 267 (2007) [hep-ph/0705.1121].

[72] S.-W. Lin *et al.* [Belle collaboration], *Measurements of Branching Fractions for $B \to K\pi$ and $B \to \pi\pi$ Decays*, Phys. Rev. Lett. **99**, 121601 (2007).

[73] B. Aubert *et al.* [BaBar collaboration], *Improved Measurements of the Branching Fractions for $B^0 \to \pi^+\pi^-$ and $B^0 \to K^+\pi^-$, and Search for $B^0 \to K^+K^-$*, Phys. Rev. **D75**, 012008 (2007).

[74] A. Abulencia *et al.* [CDF collaboration], *Observation of $B_s^0 \to K^+K^-$ and Measurements of Branching Fractions of Charmless Two Body Decays of B^0 and B_s^0 Mesons in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. Lett. **97**, 211802 (2006).

[75] A. Ali, T. Mannel and T. Morozumi, *Forward-Backward Assymetry of Diletpon Angular Distribution in the Decay $b \to s\ell^+\ell^-$*, Phys. Lett. **B273**, 505 (1991).

[76] B. Aubert *et al.* [BaBar collaboration], *Measurement of Branching Fractions, Rate Assymetries, and Angular Distributions in the Rare Decays $B \to K\ell^+\ell^-$ and $B_d \to K^{*0}\ell^+\ell^-$*, Phys. Rev. **D73**, 092001 (2006).

[77] I. Adachi *et al.* [Belle collaboration], *Measurement of the Differential Branching Fraction and Forward-Backward Assymetry for $B \to K(^*)\ell^+\ell^-$ and $B_d \to K^{*0}\ell^+\ell^-$*, BELLE-CONF-822 (2008) [hep-ex/0810.0335].

[78] J. T. Wei *et al.* [Belle collaboration], *Measurement of the Differential Branching Fraction and Forward-Backward Assymetry for $B \to K(^*)\ell^+\ell^-$ and $B_d \to K^{*0}\ell^+\ell^-$*, Phys. Rev. Lett. **103**, 171801 (2009) [hep-ex/0904.0770].

[79] T. Aaltonen *et al.* [CDF collaboration], *Search for the Rare Decays* $B^+ \to \mu^+\mu^- K^+$, $B^0 \to \mu^+\mu^- K^{*0}(892)$ *and* $B_s \to \mu^+\mu^-\phi$ *at CDF*, Phys. Rev. **D79** 0111104 (2009) [hep-ex/0804.3908].

[80] J. Wicht *et al.* [Belle collaboration], *Observation of* $B_s^0 \to \phi\gamma$ *and Search for* $B_s^0 \to \gamma\gamma$ *Decays at Belle*, Phys. Rev. Lett. **100**, 121801 (2008) [hep-ex/0712.2659].

[81] J. Shigemitsu *et al.*, *Recent Results on B Mixing and Decay Constants from HPQCD* (2009) [hep-lat/0910.4131].

[82] T. Aaltonen *et al.* [CDF collaboration], *Search for* $B_s^0 \to \mu^+\mu^-$ *and* $B_s^0 \to \mu^+\mu^-$ *Decays in 3.7* fb^{-1} *of* $p\bar{p}$ *Collisions with CDF II*, CDF Public note 9892 (2009).

[83] B. Adeva *et al.* [LHC*b* collaboration], *Roadmap for Selected Key Measurements of LHCb* , LHCb-PUB-2009-029 (2009) [hep-ex/0912.4179].

[84] J. Laiho, R. S. Van de Water and E.Lunghi, *Lattice QCD Inputs to the CKM Unitarity Triangle Analysis* (2009) [hep-ph/0910.2928]

[85] T. Inami and C. S. Lim, *Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes* $K_L^0 \to \mu\bar{m}u$, $K^+ \to \pi^+\nu\bar{n}u$ *and* $K^0 leftrightarrow \bar{K}^0$, Prog. Theor. Phys. **65**, 297 (1981).

[86] T. Inami and C. S. Lim, *Errata: Effects of Superheavy Quarks and Leptons in Low-Energy Weak Processes* $K_L^0 \to \mu\bar{m}u$, $K^+ \to \pi^+\nu\bar{n}u$ *and* $K^0 leftrightarrow \bar{K}^0$, Prog. Theor. Phys. **65**, 1772 (1981).

[87] A. J. Buras, *Relations between* $\Delta M_{s,d}$ *and* $B_{s,d} \to \mu\bar{\mu}$ *in Models with Minimal Flavor Violation*, Phys. Lett. **B566**, 115 (2003) [hep-ph/0303060].

[88] M. Blanke, A. Buras, D. Guadagnoli and C. Tarantino, *Minimal Flavour Violation Waiting for Precise Measurements of* ΔM_s, $S_{\psi\phi}$, A_{SL}^s, $|V_{ub}|$, γ *and* $B_{s,d}^0 \to \mu^+\mu^-$, JHEP **10**, 003 (2006) [hep-ph/0604057].

[89] E. Gamiz, C. T. H. Davies, G. P. Lepage, J. Shigemitsu and M. Wingate [HPQCD collaboration], *Neutral B Meson Mixing in Unquenched Lattice QCD*, Pys. Rev. **D80**, 014503 (2009) [hep-lat/0902.1815].

[90] A. J. Buras *et al.*, *Patterns of Flavour Violation in the Presence of a Fourth Generation of Quarks and Leptons* (2010) [hep-ph/1002.2126].

[91] A. K. Alok and S. U. Sankar, *New Physics Upper Bound on the Branching Ratio of* $B_s \to \ell^+\ell^-$, Phys. Lett. **B620**, 61 (2005) [hep-ph/0502120].

[92] W. Liu, C.-X. Yue and H. D. Yuang, *Rare Decays* $B_s \to \ell^+\ell^-$ *and* $B \to K\ell^+\ell^-$ *in the topcolor-assisted technicolor* (2009) [hep-ph/0901.3463].

[93] B. C. Allanach, G. Hiller, D. Jones and P. Slavich, *Flavour Violation in Anomaly Mediated Supersymmetry Breaking*, JHEP **04**, 088 (2009) [hep-ph/0902.4880].

[94] H. Logan and U. Nierste, $B_{s,d} \to \ell^+\ell^-$ *in a Two Higgs Doublet Model*, Nucl. Phys. **B586**, 39 (2000) [hep-ph/0004139].

[95] M.Misiak *et al.*, *The First Estimate of* $\mathcal{B}(\bar{B} \to X_s\gamma)$ *at* $O(\alpha_s^2)$, Phys. Rev. Lett, **98**, 022002 (2007) [hep-ph/0609232].

[96] O. Buchmueller *et al.*, *Likelihood Functions for Supersymmetric Observables in Frequentist Analyses of the CMSSM and NUHM1*, Eur. Phys. J.**C64**, 391 (2009) [hep-ph/0907.5568].

[97] R. Barbier *et al.*, *R-Parity Violating Supersymmetry*, Phys. Rept. **420**, 1 (2005) [hep-ph/0406039].

[98] A. K. Alok and S. K. Gupta, $B_s \to \mu^+\mu^-$ *Decay in the R-Parity Violating Minimal Supergravity* (2009) [hep-ph/0904.1878].

[99] M. Bauer, S. Casagrande, U. Haisch and M. Neubert, *Flavor Physics in the Randall-Sundrum Model: II. Tree-Level Weak-Interaction Processes* (2009) [hep-ph/0912.1625].

[100] A. Alves *et al.* [LHC*b* collaboration], *The LHCb Detector at the LHC*, J. Inst. **3**, S08005 (2008).

[101] R. Nandakumar *et. al.*, *The LHCb Computing Data Challenge DC06*, J. Phys. **119**, 072023 (2008).

[102] T. Sjostrand, S. Mrenna, P. Skands, *Pythia 6.4 - Physics and Manual*, JHEP **05**, 026 (2006) [hep-ph/0603175].

[103] D. J. Lange, *The EvtGen particle decay simulation package*, NIM **A462**, 152 (2001).

[104] S. Agostinelli *et al.* [GEANT4 collaboration], *Geant4 - A simulation toolkit*, NIM **A506**, 250 (2003).

[105] `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/`.

[106] `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/`.

[107] D. Martinez, J. A. Hernando and F. Teubert, *LHCb Potential to Measure/Exclude the Branching Ratio of the Decay* $B_s \to \mu^+\mu^-$, CERN-LHCb-2007-033 (2007).

[108] F. Abe *et al.* [CDF collaboration], *Observation of the B(c) meson in p anti-p collisions at S**(1/2)=1.8TeV*, Phys. Rev. Lett. **81**, 2432 (1998) [hep-ex/9805034].

[109] D. Karlen *Using Projections and Correlations to Approximate Probability Distributions*, Comput. Phys. **12**, 380 (1998).

[110] A. L. Read, *Modified Frequentist Analysis of Search Results (the CL_s Method)*, CERN Yellow Report 2000-005, 81 (2000).

[111] A. L. Read, *Presentation of Search Results: the CL_s Method*, J. Phys. **G28**, 2693 (2002).

[112] R. Fleischer, N. Serra and N. Tuning, *New Strategy for B_s Branching Ratio Measurements and the Search for New Physics in $B_s^0 \to \mu^+\mu^-$*, Phys. Rev. **D82**, 034038 (2010).

[113] R. Aaij *et al.* [LHC*b* collaboration], *Determination of f_s/f_d for 7 TeV pp collisions and a measurement of the branching fraction of the decay $B_d \to D^-K^+$*, CERN-PH-EP-2011-075 (2011) [hep-ex/1106.4435].

[114] A. Carbone *et al.*, *Charmless charged two-body decays at LHCb* , CERN-LHCb-2007-059 (2007).

[115] A. Carbone, private communication.

[116] A. Carbone, private communication.

[117] R. W. Forty and O. Schneider, *RICH Pattern Recognition*, CERN-LHCb-1998-040 (1998).

[118] J. E. Gaiser, *Charmonium Spectroscopy from Radiative Decays of the J/Ψ and Ψ'*, SLAC-R-255 (1982).

[119] R. Aaij *et al.* [LHC*b* collaboration], *Search for the Rare Decays $B_s^0 \to \mu^+\mu^-$ and $B_s^0 \to \mu^+\mu^-$*, Phys. Lett. **B699**, 330 (2011) [hep-ex/1103.2465].

[120] R. Aaij *et al.* [LHC*b* collaboration], *Search for the Rare Decays $B_{(s)}^0 \to \mu^+\mu^-$ with 300 pb^{-1} at LHCb* , LHCb-CONF-2011-037, (2011).

[121] LHC*b* and CMS collaborations, *Search for the Rare Decay $B_s^0 \to \mu^+\mu^-$ at the LHC with the CMS and LHCb Experiments*, LHCb-CONF-2011-047 and CMS-PAS-PPH-11-019 (2011).

[122] T. Aaltonen *et al.* [CDF collaboration], *Search for $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ Decays with CDF II*, Fermilab-Pub-11-315-E (2011) [hep-ex/1107.2304].