

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Departamento de Física de Partículas Instituto Galego de Física de Altas Enerxías



Study of the very rare decay $B_s \rightarrow \mu^+ \mu^-$ in LHCb

Created: 29th Dec 2009
Last modified: 10th May 2010

Prepared By: Diego Martínez Santos

Jury

Dr. Jonathan R. Ellis (president)
Dr. Juan José Saborido Silva (secretary)
Prof. Martine Fernández-Bosman
Dr. Frederic Teubert
Dra. Raluca Anca Mureşan

Substitutes: Dr. Marcos Seco Miguélez, Dr. Giampiero Mancinelli

Informs for European Mention: Dr. Guy Wilkinson, Dr. Frederic Teubert

D. BERNARDO ADEVA ANDANY, Catedrático de Física Atómica, Molecular e Nuclear da Universidade de Santiago de Compostela e D. JOSÉ ÁNGEL HERNANDO MORATA, Profesor Titular de Física Atómica, Molecular e Nuclear da Universidade de Santiago de Compostela, informan:

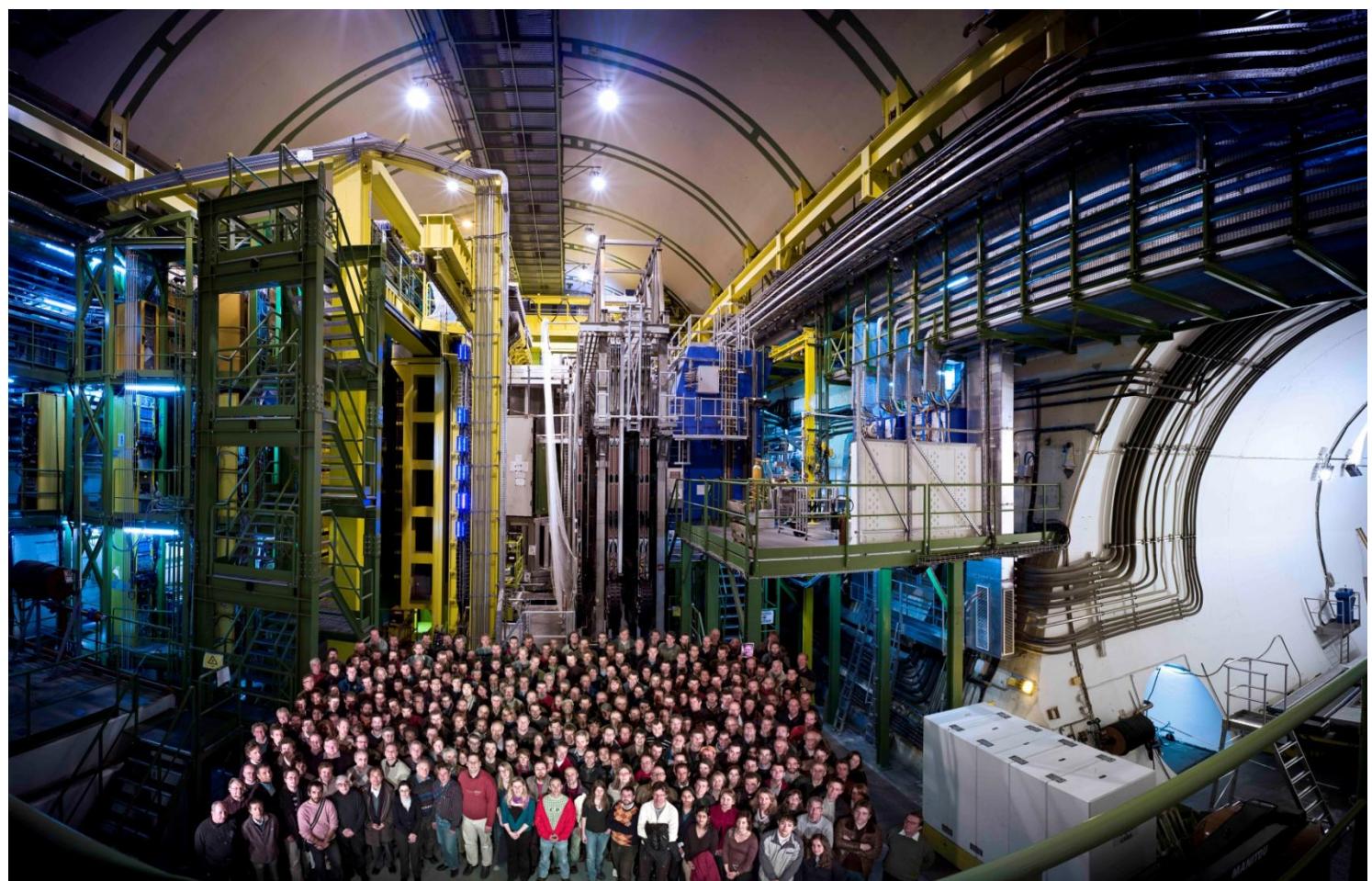
Que a memoria titulada "*Study of the very rare decay $B_s \rightarrow \mu^+ \mu^-$ in LHCb*" foi realizada por Diego Martínez Santos baixo a súa dirección e que constitúe a Tese que presenta para optar ao Grao de Doutor en Física.

Santiago de Compostela, a 2 de Febreiro de 2010.

Asdo. Bernardo Adeva Andany

Asdo. José Ángel Hernando Morata

Asdo. Diego Martínez Santos



Acknowledgements

I would like to thank my supervisors Bernardo Adeva and José Ángel Hernando Morata for dedicating their time to my learning and to the supervision of this work. Of course this thesis would have not been possible without them.

I thank Frederic Teubert for all the things I learned from him and the excellent guidelines and coordination in the development of the analysis presented in this thesis. The articles wrote in collaboration with him and José Ángel are the main pillars of this thesis.

I also thank all the other members of the $B_s \rightarrow \mu^+ \mu^-$ Working Group, especially those who I had frequent and useful conversations with: Marc-Olivier Bettler, Hugo Ruiz, Elías López, Gaia Lanfranchi, Leandro De Paula, Giampiero Mancinelli and Justine Serrano. The explanation of Marco about unexpected secrets of SPS will be difficult to forget. Xabier, don't be jealous you were going to appear in USC's paragraph...

I thank Raluca Anca Mureşan for the kind and nice help in the quick analysis of 2009 real data, the guidelines in the strange world of K_s and Λ and the review and corrections of chapter 6.

I thank as well the members of LHCb collaboration, all those who I exchanged conversations and thousands of mails with, for helping me in the understanding of the experiment and its sophisticated software. I thank Thomas Ruf for his help each time a new kind of data appears, Hans Dijkstra and Olivier Schneider for the supervision in the early times of my work at the HLT Generic and the inclusive b biased sample respectively.

I would like to thank all my colleagues from Universidade de Santiago de Compostela, especially Xabier Cid Vidal, José Luis Fungueiriño “LUCAS”, Pablo Estévez, Pablo Rodríguez, Celestino Rodríguez and Paula Álvarez for the daily interaction in the hopeless live of being PhD students. I also thank Marcos Seco Miguélez for his answers and bibliography in such different fields like Theoretical Physics and Distributed Analysis.

I also thank Andrei Golutvin and the organizers of the XXXVI ITEP Winter School in which I learned several things now applied on this thesis.

I thank the financial support provided by the Programa de Formación de Profesorado Universitario of Ministerio de Educación y Ciencia.

I also would like to thank Alfredo De Sena for converting such a bad student like me into a scientist. I really learned a lot from him. Without his help I would probably never even access the University.

I also thank the unconditional support of my parents, José Martínez Ben and Laudelina Santos Pereira, my brothers Borja and Tamara, and my brother in law Segundo Robledo Quispe.

I dedicate this thesis to the memory of José Santos Insúa, his sailor stories and his brave, happy and funny mood.

Abstract

This thesis shows the strategy to extract the $B_s \rightarrow \mu^+ \mu^-$ branching ratio from LHCb data, calibrating all the steps using control channels and not relying on the simulation. This branching ratio is very sensitive to New Physics effects, and can get large enhancements within SuperSymmetry or other Standard Model extensions.

The signal is separated from background according to three properties: the invariant mass, the muon identification, and the geometrical properties of the decay. The multivariate analysis designed to combine the geometrical properties is also shown here.

The ratio of offline reconstruction efficiencies between signal ($B_s \rightarrow \mu^+ \mu^-$) and normalization channels ($B^+ \rightarrow J/\psi (\mu^+ \mu^-) K^+$ and/or $B_d \rightarrow K^+ \pi^-$) can be extracted using the ratio of different control channels (for instance, $B_d \rightarrow J/\psi (\mu^+ \mu^-) K^{*0} (K^+ \pi^-)$) with a few percent precision. The ratio of trigger efficiencies can be extracted using events triggered independently of the signal, which with enough integrated luminosity will give a few percent precision. The invariant mass and the geometrical properties can be extracted using $B \rightarrow h^+ h^-$ events as signal candidates and the events in the sidebands of the mass distribution as background candidates, without relying on the simulation. There are several good control channels (for instance $J/\psi \rightarrow \mu^+ \mu^-$ and $\Lambda \rightarrow p \pi^-$) to be able to calibrate the muon identification efficiency and the muon misidentification probability.

This strategy will allow LHCb to perform a measurement of the $B_s \rightarrow \mu^+ \mu^-$ branching ratio that should not depend on how well our simulation reproduces real data. The strategy is tested with a sophisticated toy MC analysis for a hypothetical integrated luminosity of 150 pb^{-1} .

The potential of all LHC experiments in this measurement is also studied, showing that LHCb has the best performance for a given integrated luminosity. From this study, LHCb could overtake Tevatron's limit with the data of 2010 having already an important impact on New Physics searches. Within five nominal years, LHCb could observe values even smaller than SM prediction.

Finally, the first data produced by the LHC at the end of 2009 is used for validate, to first order the potential described in this thesis using MC simulations.

Contents

LIST OF TABLES	XVI
LIST OF FIGURES	XVIII
1. OVERVIEW.....	1
2. THEORY.....	3
2.1. EFFECTIVE HAMILTONIAN AND WILSON COEFFICIENTS	3
2.2. STANDARD MODEL.....	5
2.2.1. <i>Introduction</i>	5
2.2.2. <i>Mass generation and eigenstates</i>	6
2.2.3. <i>Fermions</i>	8
2.2.4. <i>Bosons</i>	9
2.2.5. $B_s \rightarrow \mu\mu$ <i>in the Standard Model</i>	9
2.3. NEW PHYSICS	12
2.3.1. <i>Motivation for New Physics</i>	12
2.3.2. <i>Two Higgs Doublet Models (2HDM)</i>	13
2.3.3. $B_s \rightarrow \mu\mu$ <i>in 2HDM-II</i>	14
2.3.4. <i>MFV and relation with other observables</i>	15
2.3.5. <i>SuperSymmetric Models</i>	15
2.3.6. <i>Minimal SuperSymmetric Standard Model (MSSM)</i>	16
2.3.7. $B_s \rightarrow \mu\mu$ <i>in MSSM and constrained versions</i>	18
2.3.8. <i>Tree level contributions in RPV MSSM</i>	23
2.3.9. <i>Extra dimensions</i>	24
2.3.10. $B_s \rightarrow \mu\mu$ <i>in Warped Extra Dimensions</i>	25
2.3.11. <i>Technicolor models</i>	26
2.3.12. <i>Topcolor assisted Technicolor (TC2)</i>	27
2.3.13. $B_s \rightarrow \mu\mu$ <i>in TC2</i>	28
2.3.14. <i>Little Higgs Models</i>	29
2.3.15. <i>Littlest Higgs with T – parity (LHT)</i>	29
2.3.16. $B_s \rightarrow \mu\mu$ <i>in LHT</i>	29
3. EXPERIMENTAL CONDITIONS	31
3.1. LHC ACCELERATOR.....	31
3.1.1. <i>LHC experiments</i>	32
3.2. LHCb EXPERIMENT.....	32
3.2.1. <i>LHCb detector</i>	33
3.2.2. <i>Tracking and Vertexing Performance</i>	40
3.2.3. <i>Particle Identification Performance</i>	44
3.2.4. <i>Trigger System</i>	47
3.2.5. <i>Stripping selections</i>	51
3.2.6. <i>Performance of selection and trigger algorithms</i>	56
4. ANALYSIS OF $B_s \rightarrow \mu\mu$ IN LHCb	59
4.1. SIMULATION	59
4.1.1. <i>Software Environment</i>	59
4.1.2. <i>Signal, Background and Control Channels Samples</i>	60
4.2. ANALYSIS OVERVIEW	61
4.2.1. <i>Mathematical Method to Combine Correlated Variables</i>	62
4.2.2. <i>Geometrical likelihood</i>	67
4.2.3. <i>Background studies</i>	67
4.3. EXPECTED SENSITIVITY	70
4.3.1. <i>Sensitivity in nominal conditions</i>	70

4.3.2.	<i>Sensitivity at LHC startup</i>	74
4.3.3.	<i>Sensitivity to the B_d mode</i>	75
4.4.	THE USE OF CONTROL CHANNELS	76
4.4.1.	<i>Normalization</i>	76
4.4.2.	<i>Trigger efficiencies and biases</i>	79
4.4.3.	<i>Invariant mass and Geometry likelihood calibration using $B \rightarrow hh$</i>	81
4.4.4.	<i>Background</i>	87
4.4.5.	<i>Muon ID calibration</i>	87
5.	FULL ANALYSIS EXAMPLE	91
5.1.	INTRODUCTION	91
5.2.	GENERATION OF LARGE STATISTICS BACKGROUND SAMPLES	91
5.2.1.	<i>Basic procedure</i>	91
5.2.2.	$B \rightarrow h^+ h^-$ sample	91
5.2.3.	$B^+ \rightarrow J/\psi(\mu\mu) K^+$ and $B_d \rightarrow J/\psi(\mu\mu) K^{*0}(K\pi)$ sample	93
5.2.4.	$B_s \rightarrow \mu^+ \mu^-$ sample	93
5.3.	CALIBRATION	94
5.3.1.	<i>Invariant mass and $B \rightarrow h^+ h^-$ exclusive fractions</i>	94
5.3.2.	<i>Geometrical likelihood</i>	103
5.3.3.	<i>Background level</i>	105
5.4.	NORMALIZATION	107
5.4.1.	<i>Offline efficiencies</i>	107
5.4.2.	<i>Trigger efficiencies</i>	110
5.4.3.	<i>Final normalization factor</i>	110
5.5.	RESULT.....	112
5.5.1.	<i>Summary of normalization and calibration parameters</i>	112
5.5.2.	<i>Binning</i>	112
5.5.3.	<i>Obtained measurement and limits</i>	113
5.5.4.	<i>Result neglecting normalization and calibration errors</i>	114
5.5.5.	<i>Comparison using MC truth efficiencies and pdf's</i>	115
5.6.	IMPLICATIONS.....	116
5.6.1.	<i>Test of the SM</i>	116
5.6.2.	<i>2HDM-II</i>	117
5.6.3.	<i>mSUGRA</i>	117
6.	LOOK AT 2009 DATA	119
6.1.	EVENT SELECTION, INVARIANT MASS AND STATISTICS	119
6.1.1.	<i>Event selection</i>	119
6.1.2.	<i>Invariant mass description and parameters</i>	121
6.1.3.	<i>Signal yield</i>	124
6.2.	SIGNAL DISTRIBUTIONS.....	125
6.2.1.	<i>Geometry Likelihood and related variables</i>	125
6.2.2.	<i>Effective proper time, selection efficiency for LL K_s</i>	128
6.2.3.	<i>Hits distribution</i>	129
6.2.4.	<i>Angular distributions</i>	130
6.2.5.	<i>Angular distributions in Centre of Mass</i>	132
6.3.	BACKGROUND STUDIES	134
6.3.1.	<i>Geometrical properties of LL background</i>	134
6.3.2.	<i>Muon misid</i>	136
7.	CONCLUSIONS	139
8.	APPENDIX	141
9.	RESUMO	143
9.1.	LHC	143

9.1.1.	<i>Experimentos do LHC</i>	143
9.2.	LHCb	144
9.2.1.	<i>Subdetectores do LHCb</i>	144
9.2.2.	<i>Reconstrucción de Vértices</i>	145
9.2.3.	<i>Medida do momento</i>	145
9.2.4.	<i>Identificación de Partículas</i>	145
9.2.5.	<i>Trigger</i>	146
9.3.	ANÁLISE DO DECAIMENTO RARO $B_s \rightarrow M^+ M^-$	146
9.3.1.	<i>Motivación para o estudo de $B_s \rightarrow \mu^+ \mu^-$</i>	146
9.3.2.	<i>Estratexia da Análise</i>	147
9.3.3.	<i>Sensibilidade</i>	148
9.3.4.	<i>Uso dos datos de Decembro de 2009</i>	148
10.	REFERENCES	149

List of Tables

Table 2-I: Standard Model fermion content.....	8
Table 2-II: Superpartners of SM particles.....	16
Table 3-I: IP, mass and momentum resolutions.....	43
Table 3-II: Fraction of particles with calorimeter signal below the threshold.	45
Table 3-III: L0 thresholds.	48
Table 3-IV: Stripping selection cuts for $B \rightarrow hh$	51
Table 3-V: Stripping selection cuts for $B^+ \rightarrow J/\psi(\mu\mu) K^+$	53
Table 3-VI: Stripping selection cuts for $B_d \rightarrow J/\psi(\mu\mu) K^{*0}(K^+\pi^-)$	54
Table 3-VII: Stripping selection efficiencies on offline reconstructed events.....	56
Table 3-VIII: Trigger efficiencies.....	56
Table 3-IX: Analysis output rate.....	57
Table 4-I: PYTHIA processes in DC06 minimum bias definition.....	59
Table 4-II: Simulated b-quark hadronization fractions	60
Table 4-III: Analyzed events.....	60
Table 4-IV: Expected number of bb selected events.....	67
Table 4-V: bb background events after $B_s \rightarrow \mu^+\mu^-$ selection.....	68
Table 4-VI: Signal and $\mathbf{bb} \rightarrow \mu + \mu - X$ annual yields in Mass and GL bins.	70
Table 4-VII: Expected yields per fb^{-1} in ATLAS and CMS.....	71
Table 4-VIII: CMS 3σ evidence sensitivity.....	73
Table 4-IX: CMS 90% CL exclusion sensitivity.	74
Table 4-X: Fraction of $B \rightarrow h^+h^-$ in four bins of GL and GL with normalized isolation.	85
Table 4-XI: Effect of multiplicity correction in GL calibration.....	87
Table 5-I: Invariant mass resolution in different GL bins.....	99
Table 5-II: Measured number of $B \rightarrow hh$ events in each GL bin.....	103
Table 5-III: Tight selection for the analysis example.	104
Table 5-IV: Measured number of $B \rightarrow hh$ events in each GL bin after tight selection.	105
Table 5-V: Measured number of $B \rightarrow hh$ events in each GL bin not surviving the tight selection.	105
Table 5-VI: Signal fraction in the sensitive bins.....	105
Table 5-VII: Parameters for background interpolation.	106
Table 5-VIII: Parameters for computing the normalization factor.	111

Table 5-IX: The nine independent parameters related to normalization and calibration.....	112
Table 6-I: Invariant mass parameters for observed strange hadrons.....	121
Table 6-II: Invariant mass parameters for observed MC	122
Table 6-III: Measured number of K_s , Λ and $\bar{\Lambda}$ in pp collisions, after background and beam-gas subtraction.....	124
Table 6-IV: Number of MC truth matched V^0 s	124
Table 6-V: Measured number of K_s , Λ and $\bar{\Lambda}$ in beam-gas events, after background subtraction.	125
Table 6-VI: Amount of combinatorial background for this test.....	134

List of Figures

Figure 2-1: Neutron decay.	3
Figure 2-2: Higgs-like potential $V(\phi) = -2 \phi ^2 + 4 \phi ^4$	6
Figure 2-3: $\mathbf{Bs} - \mathbf{Bs}$ oscillation diagrams.....	8
Figure 2-4: SM diagrams contributing to $B_s \rightarrow \mu\mu$	10
Figure 2-5: Example of QCD corrections to SM W box diagram.	10
Figure 2-6: Final state radiation photon energy distribution ($2E\gamma/M_{Bs}$).....	11
Figure 2-7: Gravitational lensing associated to dark matter.	12
Figure 2-8: Running of coupling constants in the SM.	13
Figure 2-9: Feynman diagrams in 2HDM-II.....	14
Figure 2-10: $BR(Bs \rightarrow \mu\mu)$ 2HDM-II as a function of M_{H+} , for $\tan\beta$ in the range 25 - 100.....	15
Figure 2-11: Dominant MSSM diagram at high $\tan\beta$	18
Figure 2-12: mSUGRA parameter space regions compatible with different BR ($B_s \rightarrow \mu^+\mu^-$).....	19
Figure 2-13: Masses of SUSY particles in CMSSM.....	20
Figure 2-14: $\Delta\chi^2$ of CMSSM fit [42] as a function of $BR(B_s \rightarrow \mu^+\mu^-)$	20
Figure 2-15: Masses of SUSY particles in NUHM.....	21
Figure 2-16: $\Delta\chi^2$ of NUHM fit [42] as a function of $BR(B_s \rightarrow \mu^+\mu^-)$ and different CL regions as a function of $\tan\beta$	22
Figure 2-17: $BR(Bs \rightarrow \mu\mu)$ dependence with $\tan\beta$ in MCPMFVMSSM.....	22
Figure 2-18: AMSB predictions for $BR(B_s \rightarrow \mu^+\mu^-)$	23
Figure 2-19: RPV Feynman diagrams.....	23
Figure 2-20: Dependence with RPV couplings in RPV mSUGRA.	24
Figure 2-21: $BR(B_s \rightarrow \mu\mu)$ vs $BR(K_L \rightarrow \mu\mu)_{SD}$ in the R-S model [58].	25
Figure 2-22: $BR(B_s \rightarrow \mu\mu)$ vs $BR(B_d \rightarrow \mu\mu)$ in the R-S model [59].	26
Figure 2-23: Fermion masses in ETC.	27
Figure 2-24: TC2 Feynman diagrams.	28
Figure 2-25: Example of mass spectrum in Littlest Higgs model.....	29
Figure 2-26: $BR(B_s \rightarrow \mu\mu)/BR(B_s \rightarrow \mu\mu)_{SM}$ versus $BR(K^+ \rightarrow \pi^+ \nu\nu)$ in LHT.	30
Figure 2-27: Variation of r as a function of 813d in LHT.....	30
Figure 3-1: LHC accelerator chain.....	31
Figure 3-2: LHCb detector.....	32
Figure 3-3: Unitarity Triangle.....	33

Figure 3-4: Probability of a given number of interactions as a function of luminosity.	34
Figure 3-5: bottom production angle.....	34
Figure 3-6: Beam pipe layout.....	35
Figure 3-7: LHCb magnet.	36
Figure 3-8: Magnetic field along z axis.	36
Figure 3-9: VELO layout.	37
Figure 3-10: IT (left) and TT (right) modules.....	37
Figure 3-11: OT layout and cross section of one OT module.	38
Figure 3-12: RICH detectors.	38
Figure 3-13: Lateral segmentation of calorimeters.	39
Figure 3-14: Muon stations.	40
Figure 3-15: Offline track classification.	41
Figure 3-16: Number of hits per subdetector for $B_s \rightarrow \mu\mu$ long tracks.	41
Figure 3-17: Tracking resolutions.	42
Figure 3-18: Primary vertex resolutions.	42
Figure 3-20: Invariant mass resolution.	43
Figure 3-19: $B_s \rightarrow \mu^+\mu^-$ decay vertex resolution.....	43
Figure 3-21: DLL obtained from RICH system.	44
Figure 3-22: Energy deposited in calorimeters.	45
Figure 3-23: IsMuon efficiency and missid rate as a function of momentum.	46
Figure 3-24: Muon identification in different subdetectors.	47
Figure 3-25: Stability of common selection cuts.	52
Figure 3-26: Common selection cuts for $B_d \rightarrow J/\psi(\mu\mu) K^{*0}(K^+\pi^-)$	54
Figure 3-27: $K\pi$ (from K^{*0}) mass.	55
Figure 3-28: Variation of K^{*0} mass cut efficiency with the resolution.	55
Figure 4-1: Momentum of the muons and its cumulative and Gaussian equivalents.....	63
Figure 4-2: Correlation in initial and Gaussian space.	63
Figure 4-3: Transformation to uncorrelated Gaussians.	64
Figure 4-4: Gaussian variables in signal and background space.....	65
Figure 4-5: Discriminating variable distributions.	65
Figure 4-6: Transformation of original Gaussian axes in the ideal case.	66
Figure 4-7: Fraction of events in the inclusive bb sample as a function of a cut in the GL used in [103].	68
Figure 4-8: Misidentified $B \rightarrow hh$ as dimuon background.	69
Figure 4-9: Geometry Likelihood distribution.	70

Figure 4-10: Sensitivity in the absence of signal.	72
Figure 4-11: Sensitivity in the presence of signal.	73
Figure 4-12: Sensitivity at LHC startup.	75
Figure 4-13: Ratio of reconstruction efficiencies as a function of B momentum.	78
Figure 4-14: Measurement of exclusive $B \rightarrow hh$ fractions.	79
Figure 4-15: TIS-TOS-TOB nomenclature.	79
Figure 4-17: K^+K^- invariant mass distribution of the $B \rightarrow hh$ inclusive sample.	82
Figure 4-16: Fitted invariant mass parameters in $B \rightarrow hh$ as a function of kaon purity.	82
Figure 4-18: $B_s \rightarrow \mu^+\mu^-$ invariant mass distribution (MeV).	83
Figure 4-19: GL for different triggers.	83
Figure 4-20: GL distribution in $B \rightarrow h^+h^-$	84
Figure 4-21: Comparison of isolation in offline with respect to TIS events.	85
Figure 4-22: GL distribution in $B \rightarrow h^+h^-$ TIS.	85
Figure 4-23: Correlation of the number of long tracks with the p_t of the B meson.	86
Figure 4-24: TIS observable efficiency as a function of B p_t and multiplicity.	86
Figure 4-25: Energy deposited by MIPs.	88
Figure 4-26: $J/\psi \rightarrow \mu^+\mu^-$ in minimum bias sample.	88
Figure 4-27: $\Lambda \rightarrow p\pi$ in minimum bias sample.	89
Figure 4-28: Calibration of Muon ID DLL.	89
Figure 5-1: Mass distributions of $B \rightarrow h^+h^-$ background.	92
Figure 5-2: Distributions of generated $B \rightarrow h^+h^-$ combinatorial background events.	92
Figure 5-3: Correlations of generated $B \rightarrow h^+h^-$ combinatorial background events.	93
Figure 5-4: K^+K^- invariant mass distribution for $DLL(K-\pi) > 25$ and $GL > 0.5$	94
Figure 5-5: $K^+\pi^-$ invariant mass distribution for $DLL(K-\pi) > 20$ in K^+ , $DLL(K-\pi) < -20$ in π^- and $GL > 0.5$	94
Figure 5-6: π^+K^- invariant mass distribution for $DLL(K-\pi) > 20$ in K^- , $DLL(K-\pi) < -20$ in π^+ and $GL > 0.5$	95
Figure 5-7: $\pi^+\pi^-$ invariant mass distribution for $DLL(K-\pi) < -30$ in and $GL > 0.5$	95
Figure 5-8: Measured positions of the B mass peaks.	96
Figure 5-9: Measured mass resolution for different PID cuts.	97
Figure 5-10: $\pi^+\pi^-$ invariant mass distribution for $GL > 0.85$ and no PID cuts.	98
Figure 5-11: Measured mass resolution for different PID cuts after correction of momentum bias.	99
Figure 5-12: Result of fitting a double Gaussian to $B_s \rightarrow K^+K^-$ invariant mass distribution.	100
Figure 5-13: Dependence of the transition point with the resolution.	101
Figure 5-14: Measured $B \rightarrow h^+h^-$ fractions as a function of PID cut.	102

Figure 5-15: Measured dependence of TIS efficiency with B p_t and multiplicity.....	103
Figure 5-16: Mass distribution of the $B \rightarrow hh$ candidates in each GL bin.	104
Figure 5-17: Invariant mass versus GL for the selected $B_s \rightarrow \mu^+ \mu^-$ candidates.....	106
Figure 5-18: Invariant mass of $B^+ \rightarrow J/\psi(\mu\mu) K^+$ after trigger and stripping selection.	108
Figure 5-19: B_d and B^+ mass peaks after the tight selection shown in Table 5-III.	109
Figure 5-20: Experiment binning.....	113
Figure 5-21: Observed invariant mass distribution of $B_s \rightarrow \mu^+ \mu^-$ candidates.....	113
Figure 5-22: CL_s and CL_b curves as a function of $BR(B_s \rightarrow \mu^+ \mu^-)$ including all calibration and normalization errors.....	114
Figure 5-23: CL_s and CL_b curves as a function of $BR(B_s \rightarrow \mu^+ \mu^-)$ without including calibration or normalization errors.....	115
Figure 5-24: CL_s and CL_b curves as a function of $BR(B_s \rightarrow \mu^+ \mu^-)$ using MC truth efficiencies and pdf's.....	116
Figure 5-25: Implications in the context of 2HDM-II for the result in the example.....	117
Figure 5-26: mSUGRA parameter space allowed by the measurement at one sigma.	118
Figure 6-1: Reconstructed invariant mass of V^0 candidates selected in 2009 preliminary LHCb data.....	120
Figure 6-2: Armenteros-Podolanski plot for LL candidates in 2009 data.....	121
Figure 6-3: Armenteros-Podolanski plot for DD candidates in 2009 data.....	121
Figure 6-4: LL $\pi^+ \pi^-$ invariant mass distribution.	122
Figure 6-5: MC and data superimposed AP plots for LL V^0 candidates.....	123
Figure 6-6: Analytical AP ellipses superimposed to LL V^0 candidates from data.....	123
Figure 6-7: GL distributions for Ks	125
Figure 6-8: DOCA distributions for Ks	126
Figure 6-9: IP distributions for Ks	126
Figure 6-10: Lifetime distributions for Ks , in meantime units.	126
Figure 6-11: minimum IPS for Ks daughters.....	127
Figure 6-12: IPS and vertex χ^2 of Ks	128
Figure 6-13: Effective mean time in LHCb.	129
Figure 6-14: Hits in OT from Ks daughters.....	129
Figure 6-15: Hits in VELO from Ks daughters.....	130
Figure 6-16: Hits in ST from Ks daughters.....	130
Figure 6-17: ϕ distribution of Ks	131
Figure 6-18: η distribution of Ks	131
Figure 6-19: Λ/Λ as a function of η in laboratory system.	132

Figure 6-20: Effect of boost to pp CM system.....	132
Figure 6-21: ϕ distribution of Ks and Λ in CM system.	133
Figure 6-22: $\Lambda/\bar{\Lambda}$ as a function of ϕ in CM system.	133
Figure 6-23: η distribution of Ks and Λ in CM system.	134
Figure 6-24: Ratios as a function of η	134
Figure 6-25: Armenteros-Podolanski selection for combinatorial background.	135
Figure 6-26: GL distribution for Ks combinatorial background.	136
Figure 6-27: Ks with positive pion flagged as IsMuon.....	136
Figure 6-28: Momentum spectra for V^0 daughters.....	137

1. Overview

Precision observables at low energy allow to access information at higher energy scales, constraining possible New Physics (NP) scenarios. The branching ratio $\text{BR}(B_s \rightarrow \mu^+ \mu^-)$ has been identified as a very interesting potential constraint on the parameter space of NP models. The Standard Model (SM) prediction is $\text{BR}(B_s \rightarrow \mu^+ \mu^-) = (3.35 \pm 0.32) \times 10^{-9}$ while the current upper limit given by Tevatron is $\text{BR}(B_s \rightarrow \mu^+ \mu^-) < 36 \times 10^{-9}$ @ 90% CL. Hence, NP can still contribute to increase the BR value up to one order of magnitude with respect to the SM expectation. In this thesis the analysis for the measurement of $\text{BR}(B_s \rightarrow \mu^+ \mu^-)$ in LHCb experiment is presented.

Section 2 reviews the theoretical prediction of the SM and the differences of the alternative expansions, such as supersymmetric models. Section 3 summarizes the experimental conditions offered by LHC accelerator and LHCb experiment, including the trigger and data stripping. Section 4 describes the analysis as it is planned to be done in LHCb, and reviews the potential of the experiment. Section 5 uses a toy MC to give a detailed example of such analysis, for a luminosity of 150 pb^{-1} and including some physical interpretations of the hypothetical measurement. In section 6, the first collisions at 900 GeV recorded by LHCb experiment are studied, using V^0 's to validate at first order several aspects of the analysis.

My contribution to this analysis includes the evaluation of reconstruction and trigger efficiency, design of the selection, the design a multivariate analysis to disentangle such rare decay from background, the search for a strategy to perform the analysis without relying in MC, the evaluation of the potential of the experiment, and the validation of LHCb performance at first order with the very first collisions of December 2009. This work is explained in section 4 (but 4.4.5, which corresponds entirely to the Muon ID group), 5 and 6, as well as the design of the stripping selection in section 3 and the evaluation of the reconstruction, trigger and PID performances shown there for $B_s \rightarrow \mu^+ \mu^-$ and its control channels. Several aspects of section 4 were done in collaboration with people from several institutions.

2. Theory

2.1. Effective Hamiltonian and Wilson Coefficients

Hadronic weak decays are often studied in terms of effective Hamiltonians. These Hamiltonians are expressed in terms of Operator Product Expansion (OPE):

$$H_{\text{eff}} = G_F \sum_i C_i \hat{Q}_i$$

Where G_F is the Fermi constant, \hat{Q}_i are local operators including only the initial and final state fields, multiplied by Wilson [1] coefficients C_i that contain the information about short distance physics, such that the masses of particles entering in internal loops. An example of such effective descriptions was Fermi theory of neutron decay [2] $n \rightarrow p^+ + e^- + \bar{\nu}_e$, where the degrees of freedom corresponding to W boson exchanged are integrated out giving a Hamiltonian with a four fermion local interaction:

$$H^F = G_F (\bar{p} \gamma^\mu n) (\bar{e} \gamma_\mu v) + h.c.$$

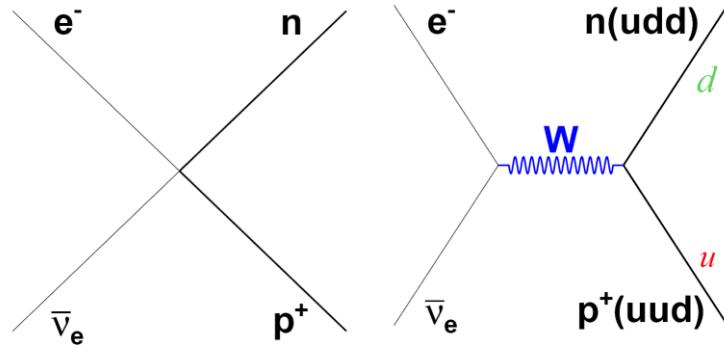


Figure 2-1: Neutron decay.

Left: Effective four-fermion theory. *Right:* Leading contribution in the fundamental electroweak theory.

The most general effective Hamiltonian for $B_q \rightarrow l^+ l^-$ transitions in the standard basis of Clifford algebra has the form:

$$H_{\text{eff}} = G_F \sum_i (\bar{q} \Gamma_i b) [\bar{l} (C_i \Gamma_i + \bar{l} C_i \Gamma_i \gamma_5) l]$$

Where Γ_i are the bilinears $\{I, \gamma_5, \gamma^\mu, \gamma^\mu \gamma_5, \sigma^{\mu\nu}\}$. In order to compute the amplitudes, the Hamiltonian is enclosed between final and initial states:

$$A(I \rightarrow F) = \langle F | H_{\text{eff}} | I \rangle = G_F \sum_i C_i \langle F | \hat{Q}_i | I \rangle$$

The matrix elements $\langle F/Q/I \rangle$ are usually factorized (strictly valid for pure leptonic decays) using vacuum insertion in order to cope with the fact that the quarks in the initial state are bounded into a hadron. The corresponding hadronic matrix elements for the $(\bar{q}\Gamma_b)$ terms are:

$$\begin{aligned} \langle 0|(\bar{q}b)|\bar{B} \rangle &= \langle 0|(\bar{q}\gamma^\mu b)|\bar{B} \rangle = \langle 0|(\bar{q}\sigma^{\mu\nu}b)|\bar{B} \rangle = 0 \\ \langle 0|(\bar{q}\gamma^\mu\gamma_5 b)|\bar{B} \rangle &= ip_{Bq}^\mu f_{Bq}, \quad \text{and} \quad \langle 0|(\bar{q}\gamma_5 b)|\bar{B} \rangle = -if_{Bq} \frac{M_{Bq}^2}{m_b + m_q} \end{aligned}$$

Where f_{Bq} is the decay constant of the B meson and is computed from experimental data. Then, from the ten possible terms in the effective Hamiltonian, six will vanish and, moreover, when doing the contractions of $\langle 0|(\bar{q}\gamma^\mu\gamma_5 b)|\bar{B} \rangle$:

$$\begin{aligned} p_{Bq}^\mu (\bar{l}\gamma_\mu l) &= (p_l^\mu + p_{\bar{l}}^\mu)(\bar{l}\gamma_\mu l) = m_l \bar{l}l - m_{\bar{l}} \bar{l}l = 0 \\ p_{Bq}^\mu (\bar{l}\gamma_\mu\gamma_5 l) &= (p_l^\mu + p_{\bar{l}}^\mu)(\bar{l}\gamma_\mu\gamma_5 l) = m_l \bar{l}\gamma_5 l - (-m_{\bar{l}} \bar{l}\gamma_5 l) = 2m_l \bar{l}\gamma_5 l \end{aligned}$$

another vanishing term will appear, so finally the effective Hamiltonian could be written in terms of only three operators. However, in order to classify different contributions not only in the Standard Model (SM) but also in its extensions, the effective Hamiltonian is often rewritten as in [3]:

$$H_{\text{eff}} = \frac{G_F \alpha}{\sqrt{2\pi} \sin^2 \theta_W} V_{tb}^* V_{tq} \left[C_S \hat{Q}_S + C_P \hat{Q}_P + C_A \hat{Q}_A + C'_S \hat{Q}'_S + C'_P \hat{Q}'_P + C'_A \hat{Q}'_A \right]$$

Being α the electromagnetic fine structure constant and θ_W the Weinberg angle, and the operators (note different quark mass in S', P' w.r.t S, P).

$$\begin{aligned} \hat{Q}_S &= m_b (\bar{q}P_R b)(\bar{l}l) & \hat{Q}_P &= m_b (\bar{q}P_R b)(\bar{l}\gamma_5 l) & \hat{Q}_A &= (\bar{q}\gamma^\mu P_L b)(\bar{l}\gamma_\mu\gamma_5 l) \\ \hat{Q}'_S &= m_q (\bar{q}P_L b)(\bar{l}l) & \hat{Q}'_P &= m_q (\bar{q}P_L b)(\bar{l}\gamma_5 l) & \hat{Q}'_A &= (\bar{q}\gamma^\mu P_R b)(\bar{l}\gamma_\mu\gamma_5 l) \end{aligned}$$

There m_b is the mass of the b quark, $(q = d, s)$ is the field of the companion quark of the b in the B_q meson, $P_{R,L} = (1 \pm \gamma_5)$ and l is the muon field. From that Hamiltonian, $\text{BR}(B_q \rightarrow \mu^+ \mu^-)$ can be expressed as:

$$\begin{aligned} \text{BR}(B_q \rightarrow \mu^+ \mu^-) &= \frac{G_F^2 \alpha^2}{64\pi^3 \sin^4 \theta_W} |V_{tb}^* V_{tq}|^2 \tau_{Bq} M_{Bq}^3 f_{Bq}^2 \sqrt{1 - \frac{4m_\mu^2}{M_{Bq}^2}} \times \\ &\times \left\{ M_{Bq}^2 \left(1 - \frac{4m_\mu^2}{M_{Bq}^2} \right) \left(\frac{C_S - \mu_q C'_S}{1 + \mu_q} \right)^2 + \left[M_{Bq} \left(\frac{C_P - \mu_q C'_P}{1 + \mu_q} \right) + \frac{2m_\mu}{M_{Bq}} (C_A - C'_A) \right]^2 \right\} \end{aligned}$$

Where τ_{Bq} , M_{Bq} and f_{Bq} are the mean lifetime, mass and decay constant of B_q meson, m_μ the muon mass and μ_q the ratio of masses m_q/m_b . As $C_{S,P}$ and $C'_{S,P}$ are in general of comparable size [3] and $\mu_q \ll 1$, the terms with μ_q can be neglected simplifying the expression:

$$BR(B_q \rightarrow \mu^+ \mu^-) = \frac{G_F^2 \alpha^2}{64\pi^3 \sin^4 \theta_W} |V_{tb}^* V_{tq}|^2 \tau_{Bq} M_{Bq}^3 f_{Bq}^2 \sqrt{1 - \frac{4m_\mu^2}{M_{Bq}^2}} \times \\ \times \left\{ M_{Bq}^2 \left(1 - \frac{4m_\mu^2}{M_{Bq}^2} \right) C_S^2 + \left[M_{Bq} C_P + \frac{2m_\mu}{M_{Bq}} (C_A - C'_A) \right]^2 \right\}$$

Note that, using same dimensions for the different Wilson coefficients, the axial contributions are suppressed with respect to scalar and pseudoscalar by a factor $\sim m_\mu/M_B$. In this sense, $BR(B_q \rightarrow \mu^+ \mu^-)$ is expected to be more sensitive to new physics in the scalar sector. Moreover, from measured rates of $B \rightarrow (K, K^*) l^+ l^-$, the NP effects are highly constrained in the vector/axial-vector is expected for $BR(B_q \rightarrow \mu^+ \mu^-)$ [4].

Even if in LHC the number of produced B_d mesons is four times larger than the number of B_s , the main $B_q \rightarrow \mu^+ \mu^-$ decay to search for is the B_s , as the B_d mode will be suppressed by $\sim |V_{tb}^* V_{td}|^2 / |V_{tb}^* V_{ts}|^2 \sim 1/20$. So we focus our interest mainly in the B_s case, although the B_d is also study not only for its individual sensitivity, but also because NP effects different from those of B_s can happen in case of non Minimal Flavor Violation (see for instance sections 2.3.10 and 2.3.16).

2.2. Standard Model

2.2.1. Introduction

The Standard Model (SM) is a quantum field theory (QFT) based on strong and electroweak (EW) interactions. The strong interactions are described by Quantum Chromodynamics (QCD) corresponding to the symmetry group $SU(3)_C$ of *color* (C), while the EW interaction is described by the group $SU(2)_T \otimes U(1)_Y$ of weak-isospin (T) and hypercharge(Y), being then $SU(3)_C \otimes SU(2)_T \otimes U(1)_Y$ the full group of gauge symmetry for the SM.

$$G_{SM} = SU(3)_C \times SU(2)_T \times U(1)_Y$$

This symmetry is spontaneously broken into $SU(3)_C \otimes U(1)_{EM}$ by the vacuum expectation value (VEV) of (the neutral component of) a scalar isospin doublet, with hypercharge 1/2, called Higgs:

$$G_{SM} \xrightarrow{\text{Higgs } (1,2)_{1/2}} SU(3)_C \times U(1)_{EM}$$

As a result of the interaction with the Higgs field, EW bosons combine into the massive particles W^\pm and Z^0 and the massless photon. The interaction with Higgs gives also masses to the fermions.

Each fermion generation, out of a total of three, has five representations of the SM gauge symmetry:

$$Q_{L,i}(3,2)_{+1/6} \quad U_{R,i}(3,1)_{+2/3} \quad D_{R,i}(3,1)_{-1/3} \quad L_{L,i}(1,2)_{-1/2} \quad E_{R,i}(1,1)_{-1}$$

The subscript number is the hypercharge, and the numbers in parenthesis indicate if it acts as a triplet or singlet in $SU(3)_C$ and as a doublet or singlet in $SU(2)_T$. The subscript $i = 1, 2, 3$ indicates fermion generation.

The EW symmetry breaking (EWSB) and the effects induced by Higgs field such like CP violation and flavor depending processes are explained in 2.2.2. The fermion and boson content of SM is explained in more detail in sections 2.2.3 and 2.2.4.

Thus, the SM lagrangian can be decomposed in three parts:

$$\mathcal{L} = \mathcal{L}_{Kin} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yuk}$$

The kinetic part includes the corresponding covariant derivative to preserve the gauge invariance.

$$D^\mu = \partial^\mu + ig_s G_a^\mu L_a + ig W_b^\mu T_b + ig' B^\mu Y$$

G_a are the gluon fields, W_b the three weak bosons and B the hypercharge boson. g_s , g and g' are the corresponding gauge couplings. The ratio $g'/g = \tan\theta_w$ defines the Weinberg angle.

The Higgs part includes Higgs self interactions and Yukawa part includes Higgs-fermion interactions. This last part contains all the CP violation sources in the SM, as well as the flavor depending terms.

2.2.2. Mass generation and eigenstates

2.2.2.1. Boson masses and EWSB

A Lagrangian containing only the terms of the gauge symmetry G_{SM} is not enough to build a model where the particles are massive. The gauge bosons are massless if the symmetry is unbroken, and masses for the fermions as self-interactions such like $\bar{\Psi}_L \Psi_R$ (Dirac mass) or $\Psi_L \Psi_L$ (Majorana mass) would explicitly break the SU(2) symmetry.

Non-abelian broken gauge theories are not renormalizable, thus in the SM the masses of the EW gauge bosons and the fermions are given by a spontaneously symmetry breakdown. This is achieved by the introduction of the Higgs, a scalar isospin doublet with hypercharge +1/2:

$$\varphi = \begin{pmatrix} \varphi^+ \\ \varphi^0 \end{pmatrix}$$

Which has a self interaction of the form:

$$\mathcal{L}_{Higgs} = \mu^2 \varphi^\dagger \varphi - \lambda (\varphi^\dagger \varphi)^2$$

The first term is like a mass term but with opposite sign. Such quadratic potential does not minimize at 0 (see example in Figure 2-2) thus φ acquires a VEV $v = \mu/\sqrt{\lambda}$.

$$\langle \varphi \rangle_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}$$

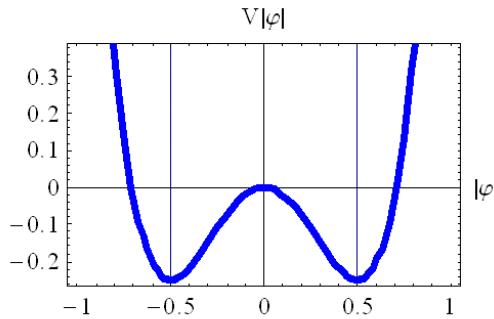


Figure 2-2: Higgs-like potential $V(|\varphi|) = -2|\varphi|^2 + 4|\varphi|^4$

The VEV gives masses, through the Higgs kinetic term plus the Higgs self-interaction Lagrangian, to the following boson combinations:

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (W_\mu^{(1)} \mp W_\mu^{(2)}) \rightarrow M_W = g \frac{v}{2}, \quad Z_\mu^0 = \frac{1}{\sqrt{g^2 + g'^2}} (g W_\mu^{(3)} - g' B_\mu) \rightarrow M_Z = \sqrt{g^2 + g'^2} \frac{v}{2}$$

From the degrees of freedom of the original Higgs field:

$$\varphi = \begin{pmatrix} \varphi^+ \\ \varphi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} G_1^+ + iG_2^+ \\ v + (H^0 + iG_3^0) \end{pmatrix}$$

H^0 will be a massive scalar particle with $M_H = \mu\sqrt{2}$ and the massless Goldstone bosons G_i are “eaten” by the gauge bosons W^\pm and Z^0 giving rise to their longitudinal polarizations and masses.

2.2.2.2. Fermion masses and CKM matrix

In order to give masses to the fermions, the corresponding couplings between them and the Higgs field are added, while keeping the Lagrangian SU(2) invariant. For example, for a single generation:

$$\Delta\mathcal{L} = -\lambda_e \bar{E}_L \varphi E_R - \lambda_d \bar{Q}_L \varphi D_R - \lambda_u \epsilon^{ab} \bar{Q}_{La} \varphi_b^\dagger U_R + h.c.$$

Substituting the VEV the fermion masses have the form:

$$m_e = \frac{v\lambda_e}{\sqrt{2}}, \quad m_u = \frac{v\lambda_u}{\sqrt{2}}, \quad m_d = \frac{v\lambda_d}{\sqrt{2}}$$

Those λ_i are inputs in the SM and thus they allow having very different masses for different fermions.

When the three fermion generations are added to the theory, additional terms mixing quarks of different generations are possible. Alternatively, it is possible to diagonalize the Higgs couplings by switching to a different basis for the quark fields. Writing the lagrangian in this alternative basis (hereafter “mass basis” or “physical basis”) will of course simplify \mathcal{L}_{Yuk} but with the cost of causing a complication in the gauge side.

Calling q the interaction eigenstates and q' the mass eigenstates, both bases are related through the unitary relations:

$$u_L^i = U_u^{ij} u'^j_L, \quad d_L^i = U_d^{ij} d'^j_L$$

And thus the weak current $\bar{u}_L^i \gamma^\mu d_L^i$ transforms to $\bar{u}'_L^i \gamma^\mu (U_u^\dagger U_d)_{ij} d'^j_L \equiv \bar{u}'_L^i \gamma^\mu V_{ij}^{CKM} d'^j_L$. Being V^{CKM} called the CKM [5][6] matrix (from Cabibbo-Kobayashi-Maskawa). Its coefficients use to be written as:

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

V^{CKM} , is not diagonal (the experimental value of the coefficients can be found in[7]) and such structure allows transitions between the different quark generations, giving rise to processes in which quarks change flavor without changing its electric charge. These processes are called Flavor Changing Neutral Currents (FCNC) and in particular include the decay $B_s \rightarrow \mu^+ \mu^-$. CP violation also arises from the non diagonal structure of V^{CKM} . Equivalently, if V^{CKM} were the identity matrix CP violation and FCNC’s would not exist within the SM.

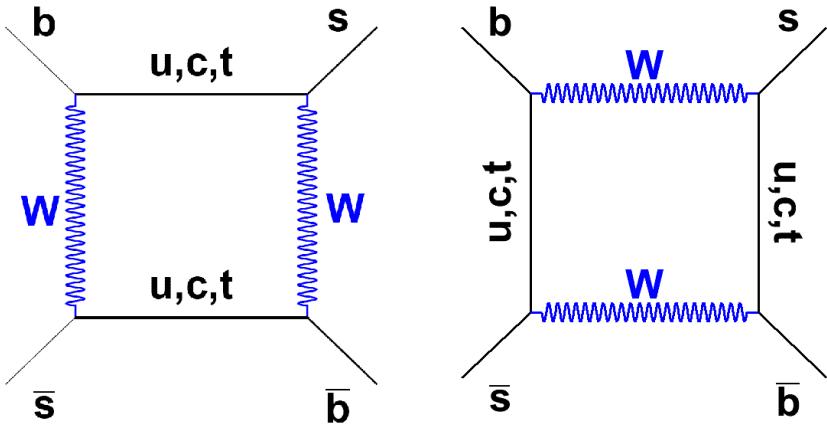


Figure 2-3: $B_s - \bar{B}_s$ oscillation diagrams.

A particular but very important example is the oscillation of neutral mesons composed by quarks of different generations. The off-diagonal terms of CKM matrix allows particles such like D^0 , K^0 , B_d or B_s to perform particle-antiparticle oscillations (see Figure 2-3).

The number of free parameters of the CKM matrix with the three generations is 4: 3 rotation angles and one phase that is the only source of CPV in the SM. It can be noticed that in the case of only two generations that phase can be removed, which implies that CPV processes must involve the three quark families.

2.2.3. Fermions

The Standard Model fermions can be divided in two groups depending if they are affected by strong interaction (quarks) or not (leptons). Each quark has three possible color states and (at low energy) only exist in bound states of color singlets, called hadrons. Hadrons are then composed by quarks (and gluons, the gauge bosons of QCD), being the most common states quark – antiquark (mesons), and three quarks (baryons). Due to spin addition, baryons are also fermions, while mesons are bosons.

Leptons are e , μ , τ and a neutrino (ν) for each one. In the SM neutrinos are massless particles thus their helicity becomes equivalent to chirality. It means that there are not right-handed neutrinos in the SM and, equivalently, there are not left-handed antineutrinos.

Table 2-I: Standard Model fermion content.

<i>Leptons</i>	T	T₃	Y	Q
ν_e, ν_μ, ν_τ	$\frac{1}{2}$	$\frac{1}{2}$	-1	0
e_L, μ_L, τ_L	$\frac{1}{2}$	$-\frac{1}{2}$	-1	-1
e_R, μ_R, τ_R	0	0	-2	-1
<i>Quarks</i>				
u_L, c_L, t_L	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{2}{3}$
u_R, c_R, t_R	0	0	$\frac{4}{3}$	$\frac{2}{3}$

d_L', s_L', b_L'	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{3}$	$-\frac{1}{3}$
d_R', s_R', b_R'	0	0	$-2/3$	$-1/3$

2.2.4. Bosons

Apart from mesons, the SM contains the gauge bosons corresponding to strong and EW interactions, and Higgs (H^0) boson, responsible of the masses of SM particles.

The gauge bosons of QCD are massless particles of spin 1, called gluons, and have eight possible color states. QCD couplings have the property of become small at high energies (or small distances); this effect is known as “asymptotic freedom”.

The gauge bosons corresponding to $SU(2)_T \otimes U(1)_Y$ are W_μ^i ($i = 1, 2, 3$) and B_μ , for $SU(2)$ and $U(1)$ respectively and the four should be massless in order to conserve the symmetry. However, the symmetry breaking induced by Higgs field, changes them into W_μ^+ , W_μ^- , Z_μ^0 and photon (A_μ), where only the photon is massless. All have spin 1.

2.2.5. $B_s \rightarrow \mu\mu$ in the Standard Model

A couple of muons only can be directly produced from a photon, Higgs, or Z^0 . But none of these bosons can be originated by a $b - s$ quark interaction directly, those neutral bosons can be produced only by particles of the same flavor. Thus, there are not tree diagrams for the process $B_s \rightarrow \mu^+ \mu^-$ in the SM (neither for any other FCNC); the main contributions are weak interactions of fourth order: Z^0 penguins and W^\pm box. The corresponding diagrams are shown in Figure 2-4 and were first calculated in [8] for the case of $K_L \rightarrow \mu^+ \mu^-$. Quarks entering in the loops can be u , c and t , but due to larger values of V_{tb} and m_t , u and c contributions can be neglected with respect to t contribution. As explained in 2.2.2.2, those diagrams would be 0 for $V^{CKM} = 1$.

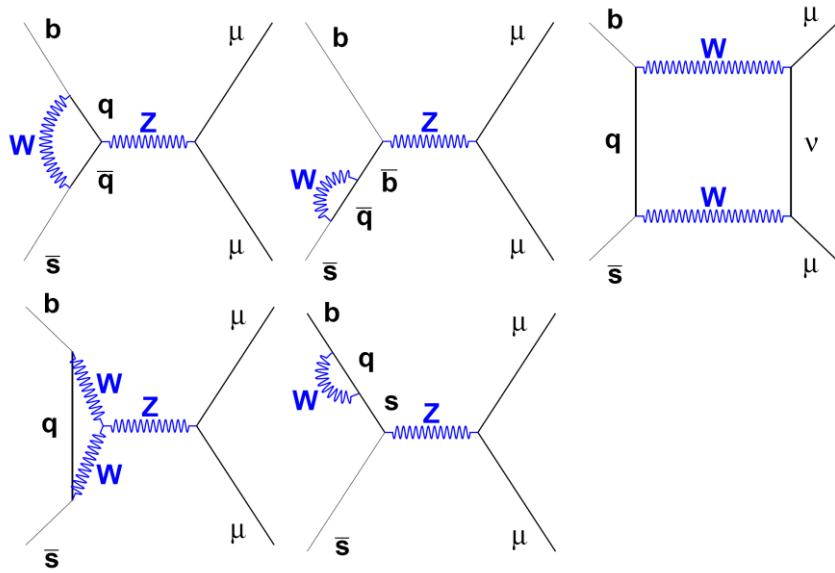


Figure 2-4: SM diagrams contributing to $B_s \rightarrow \mu\mu$.

Diagrams of Figure 2-4 contribute to C_A . Contributions from Higgs penguin[9] to C_S and Goldstone boson penguin[10] to C_P are suppressed by m_b^2/M_W^2 , being these two coefficients negligible in the SM.

2.2.5.1. QCD corrections

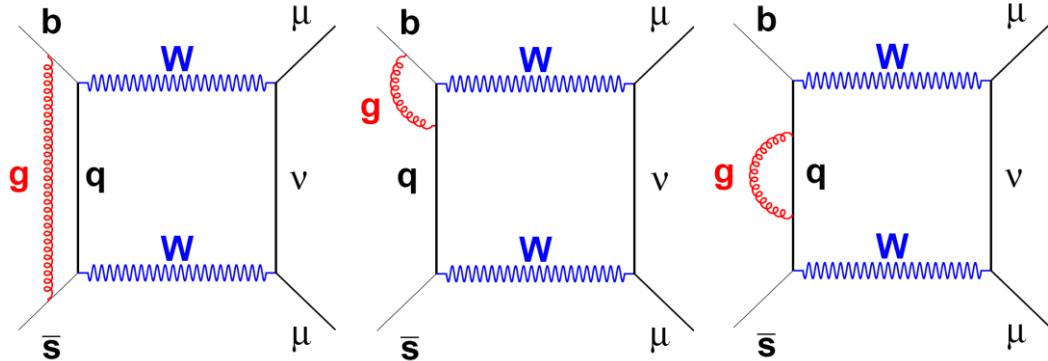


Figure 2-5: Example of QCD corrections to SM W box diagram.

Gluon exchange between the quarks b, s(d) and t entering in the diagrams needs to be included in order to get the right value of Wilson coefficients. C_A has been computed at the Next-to-Leading-Order (NLO)[11] [12] as a function of \overline{m}_t , the top quark mass computed in the Minimal Subtraction renormalization scheme (\overline{MS}) [13].

$$C_A^{(SM)} = -Y \left(\frac{\overline{m}_t^2}{M_W^2} \right); \quad \text{The expression of } Y, \text{ called the Inami-Lim function, can be found in the equation (16) of ref[12].}$$

The accuracy on the prediction can be improved by exploiting the SM correlation with the oscillation frequency of the B_s [14], giving finally [15] :

$$BR(B_s \rightarrow \mu\mu)^{SM} = (3.35 \pm 0.32) \cdot 10^{-9}$$

Which is still one order of magnitude away from current experimental 90% CL upper limit [16]:

$$BR(B_s \rightarrow \mu\mu) < 3.6 \cdot 10^{-8}$$

For the case of the Bd, the SM prediction is:

$$BR(B_d \rightarrow \mu\mu)^{SM} = (1.03 \pm 0.09) \cdot 10^{-10}$$

Also far from the current upper limit[16]:

$$BR(B_d \rightarrow \mu\mu) < 6.0 \cdot 10^{-9}$$

2.2.5.2. QED corrections (final state radiation)

When the B_s decays into the muons, the creation of the charged particles causes some energy to be radiated through soft photons. In a small fraction of those events, the radiated energy can be enough to make the $\mu^+\mu^-$ mass to be below the search window and hence they are not going to be selected for the experimental analysis.

The distribution of the photon energy follows the expression [17]:

$$\rho(x) \propto - \left[\left(2 + \frac{4r}{x} - \frac{2}{x} - x \right) \ln \frac{1 + \sqrt{1 - \frac{4r}{1-x}}}{1 - \sqrt{1 - \frac{4r}{1-x}}} + \frac{2}{x} (1-x) \sqrt{1 - \frac{4r}{1-x}} \right]$$

With x being $2E\gamma/M_{B_s}$ and r the ratio $(m_\mu/M_{B_s})^2$. That distribution follows qualitatively the usual $1/x$ behavior (see Figure 2-6).

QED corrections may also affect the actual value of the BR, however the effects are very small and negligible in comparison to the theoretical uncertainty[18].

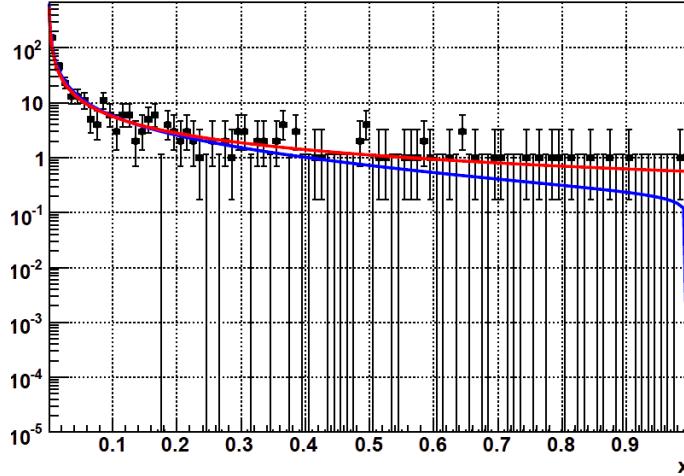


Figure 2-6: Final state radiation photon energy distribution ($2E\gamma/M_{B_s}$).

Black points: MC simulation. Blue line: Analytical expression. Red line: Approach $\rho(E\gamma) = 1/E\gamma$.

2.3. New Physics

2.3.1. Motivation for New Physics

2.3.1.1. Experimental indications

There are some experimental observations not explained by SM. Recent observation of neutrino flavor oscillation[19] implies that this particle is massive, in contradiction with SM postulates although it can be solved by adding some mechanism like seesaw[20], without changing too much the basic ideas of the SM.

In addition, astronomical observations indicate the Universe is populated by a kind of neutral particles known only by its gravitational effects, called “dark matter” (DM)[21], which contribution to the mass of the Universe is $\sim 5\text{-}6$ times larger than ordinary matter.

While some baryons, as well as neutrinos, could contribute to DM, the majority is non baryonic cold (in the sense of non-relativistic, i.e., opposite to neutrinos) DM and has no explanation within the SM. The astronomical data leads then to an inconsistency amongst SM and Gravitation Theory. Hence, strictly speaking, either the gravitation theory or the SM content need to be changed, but astronomical observations indicate that the second is more likely. In particular, the studies of Bullet cluster[22] or MACS J0025.4-1222 [23], where the gravitational lensing seems incompatible with the distribution of ordinary matter, are considered one of the most direct evidences of DM.

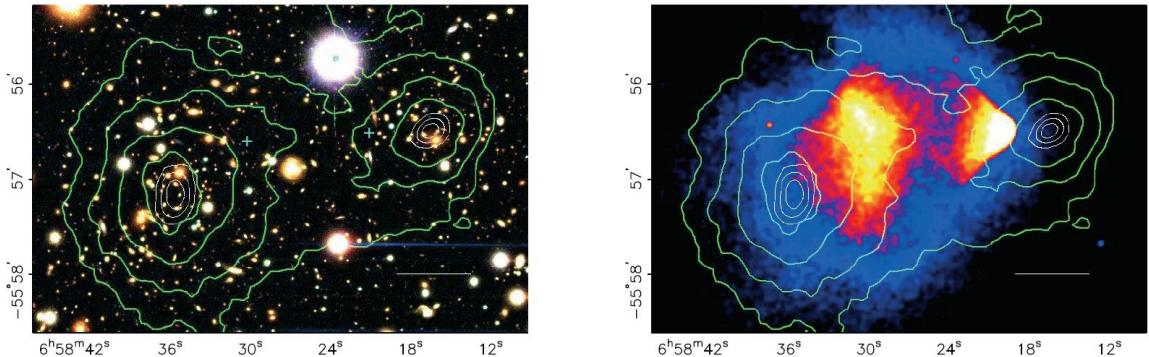


Figure 2-7: Gravitational lensing associated to dark matter.

Two galaxy clusters have collided. DM (if any) components did not interact and passed through each other. Galaxies also behave as non colliding particles. Colliding plasma keeps in the center of the collision. Left: Optical spectrum (galaxies). Right: X-ray spectrum (plasma). Green contours show the observed gravitational lensing. White contours show the errors (68.3%, 95.5%, and 99.7% confidence levels, respectively) on the positions of the two gravitational lensing peaks, measured to be at $\sim 8\sigma$ from the baryonic mass peaks[22].

Candidates for cold DM are primordial black holes [24][25], axions[26][27][28], and weak interacting massive particles (WIMPs), such as the lightest supersymmetric particle (LSP).

Also, the measurement of the anomalous magnetic dipole moment of the muon[29] [30] deviates more than 3σ from SM prediction, which could be interpreted as NP contributions.

2.3.1.2. Theoretical motivations

Apart from experimental requirements, some theoretical motivations exist.

- The number of free parameters (most of them in the Yukawa part) is large and makes the SM to look more like an effective low energy theory.
- Moreover, the number of fermion families is an input and so the SM does not offer an explanation for it.
- Gravity is not included in SM, so it cannot be valid at energies of the order of Planck scale $M_{Pl} \sim 10^{19}$ GeV.

- The mass of the Higgs boson posses quadratically divergences and naturally grows up to some huge scale unless a very delicate cancelation (i.e., fine tuning of input parameters) between the bare mass and the radiative corrections occur. If the mass of the Higgs is left to grow free then the EW scale would be large as well. Because of the relation of the Higgs mass with the EW scale, and thus with the question of why such scale and gravity scale are so different, this is known as hierarchy problem.

An extra theoretical indication stands on the idea of unification of gauge interactions, motivated by the direct product structure of G_{SM} , the reducibility of the fermion representations and the arbitrary assignment of hypercharge values. So a higher symmetry group (such as SU(5) or SO(10)) breaking down to G_{SM} is preferred. For instance, SO(10) would fit all the fermions (plus and extra particle with the quantum numbers of a right handed neutrino, which moreover is needed for explaining neutrino masses) into a single representation.

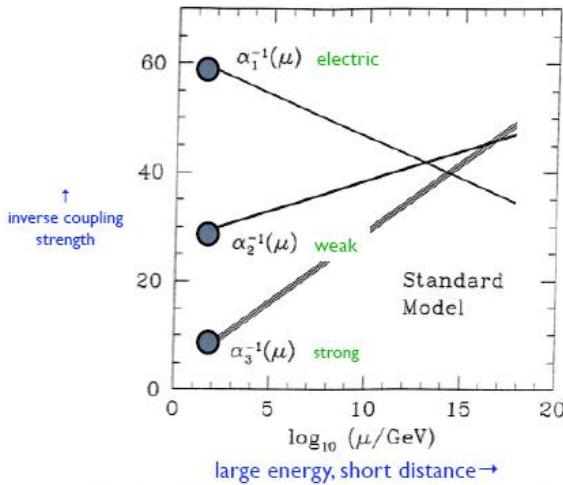


Figure 2-8: Running of coupling constants in the SM.

That unification of gauge interactions would require also a unification of coupling constants at the scales where the original symmetry remains unbroken. If the effect of the SM fields (through vacuum polarization) on the coupling constants is calculated, an approximated (but not successful within experimental uncertainty) unification is reached (see Figure 2-8). This is sometimes used as an argument in favor of new symmetries/fields (in particular supersymmetry) that would slightly modify the effects of vacuum polarization reaching more accurate unification.

2.3.2. Two Higgs Doublet Models (2HDM)

The Two Higgs Doublet Model (2HDM)[31] is an extension of the SM containing a larger Higgs sector, coming from two Higgs doublets both with a VEV $\neq 0$.

$$\varphi_a = \begin{pmatrix} \varphi_a^+ \\ \varphi_a^0 \end{pmatrix}, \quad \langle \varphi_a \rangle_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_a \end{pmatrix}; \quad \varphi_b = \begin{pmatrix} \varphi_b^+ \\ \varphi_b^0 \end{pmatrix}, \quad \langle \varphi_b \rangle_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_b \end{pmatrix}$$

The VEV's are not independent as they are constrained by $v_a^2 + v_b^2 = v_{SM}^2 \approx (174 \text{ GeV})^2$. The above fields can be rotated by an angle $\beta = \arcsin \left(\frac{v_b}{\sqrt{v_a^2 + v_b^2}} \right)$ giving:

$$\varphi_1 = \sin(\beta) \varphi_b + \cos(\beta) \varphi_a, \quad \varphi_2 = -\sin(\beta) \varphi_a + \cos(\beta) \varphi_b$$

$$\langle \varphi_1 \rangle_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_{SM} \end{pmatrix}; \quad \langle \varphi_2 \rangle_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The ratio $\tan\beta = v_b/v_a$ is an important free parameter of the theory. The combined field φ_1 will be the equivalent of the SM Higgs and the responsible of EWSB. The Goldstone bosons that will constitute the longitudinal component of the gauge fields Z^0 and W^\pm are then $Im(\varphi_1^0)$ and φ_1^\pm respectively. The physical particles will be two neutral scalars H^0 and h^0 being linear combinations of $Re(\varphi_1^0)$ and $Re(\varphi_2^0)$, a neutral pseudoscalar $A^0 \propto Im(\varphi_2^0)$ and charged scalars $H^\pm = \varphi_2^\pm$.

Then, depending on how the Yukawa couplings are implemented, 2HDM is divided into two types:

In Type-I both up-type and down-type quarks get masses from the same Higgs field, namely φ_a .

In Type-II the masses are obtained depending on the charge so $\varphi_{a,b}$ can be named as $\varphi_{u,d}$ being $\tan\beta = v_u/v_d$. Charged leptons also get their masses from φ_d .

2.3.3. $B_s \rightarrow \mu\mu$ in 2HDM-II

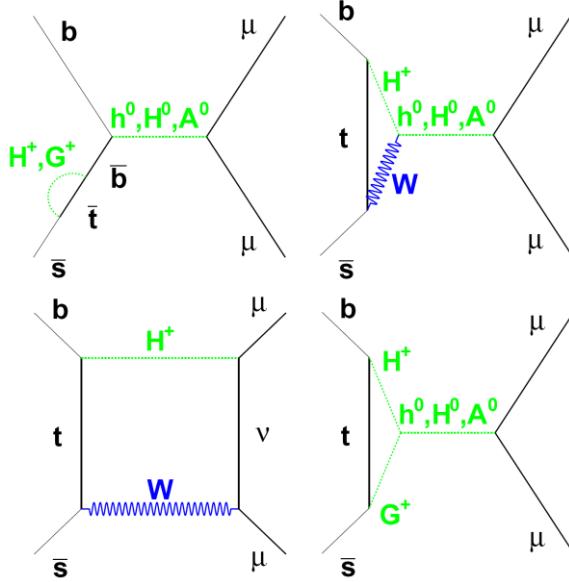


Figure 2-9: Feynman diagrams in 2HDM-II.

The leading order contributions to $BR(B_s \rightarrow \mu\mu)$ in this model have been computed in [32]. The main diagrams provided by 2HDM-II are shown in Figure 2-9, which give a contribution proportional to $\tan^2\beta$ in the scalar and pseudoscalar Wilson coefficients¹ (thus $\tan^4\beta$ once the amplitude is squared):

$$C_S^{2HDM-II} = C_P^{2HDM-II} = \frac{m_\mu}{4M_W^2} \tan^2\beta \frac{\text{Log}(\frac{M_{H+}^2}{m_t^2})}{\frac{M_{H+}^2}{m_t^2} - 1}$$

Where the mass of the top quark is evaluated in \overline{MS} at the scale of top mass. Although from direct search $M_{H+} > 80$ GeV, the agreement between the SM prediction and the measurement of the inclusive FCNC in $b \rightarrow s\gamma$ puts a higher limit of 295 GeV in the case of 2HDM-II[33]. This limit reduces the $\tan\beta$ enhancement and, because the negative relative sign between C_P and C_A , gives $BR(B_s \rightarrow \mu\mu) \lesssim$ SM prediction unless $\tan\beta \gtrsim 60$, as shown in Figure 2-10.

¹ The factor 2 difference from [32] is just because the global factor in BR expression used there is 1/256 instead of the 1/64 used here.

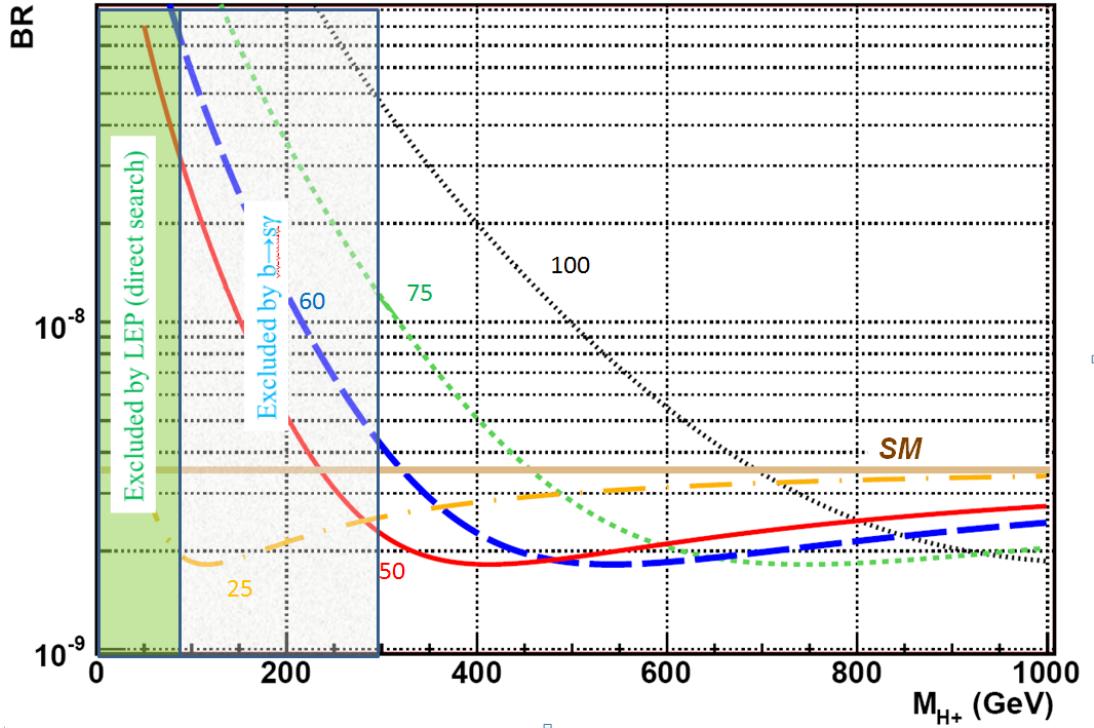


Figure 2-10: $\text{BR}(\text{Bs} \rightarrow \mu\mu)$ 2HDM-II as a function of M_{H^+} , for $\tan\beta$ in the range 25 - 100.

2.3.4. MFV and relation with other observables

Under the definition of MFV (*Minimal Flavor Violation*) in [34], which requires that the dynamics of flavor and CP violation are governed by the known structure of the Yukawa couplings, the $\text{BR}(\text{Bs} \rightarrow \mu\mu)$ can be related with other observables in a model independent way. This is because in MFV all FCNC's are governed by the same effective coupling constant $\propto V_{td_1} V_{td_2}^*$ where d_1 and d_2 are the external down-type quarks.

$\text{BR}(\text{B}_s \rightarrow \mu\mu)$ can be related [35] with the lepton universality in $\text{B}_d \rightarrow \text{K}^* \ell^+ \ell^-$ or the $\text{BR}(\text{B} \rightarrow \text{X}_s \tau\bar{\tau})$, but the most constraining is the B_d to B_s ratio:

$$\frac{\text{BR}(\text{B}_d \rightarrow \mu\mu)}{\text{BR}(\text{B}_s \rightarrow \mu\mu)} \approx \frac{\tau_{Bd} f_{Bd} m_{Bd}}{\tau_{Bs} f_{Bs} m_{Bs}} \left| \frac{V_{td}}{V_{ts}} \right|^2$$

The MFV condition can be constrained (CMFV) by also imposing that the only relevant operators in the effective Hamiltonian below the weak scale are those that are also relevant in the SM. In this case a relation with the oscillation frequency of the corresponding meson can be derived[36].

2.3.5. SuperSymmetric Models

The symmetry which requires the theory to be invariant under the transformation of fermions to bosons (and vice versa) is called SuperSymmetry, and can be considered as a generalization of space-time symmetries in QFT. It implies that the model should have the same number of fermions than bosons, establishing for each SM fermion the corresponding *sfermion* (selectron, squark...) and for each boson the corresponding fermion, named by adding the suffix “ino” (gluino, photino, higgsino...). The symmetry is extended of course to new fermions/bosons added to the theory (graviton \rightarrow gravitino).

But no supersymmetric particle has been yet observed in accelerators. One should have observed them since their masses were equal to the ones of their SM superpartners if the supersymmetry was unbroken. Therefore, if the supersymmetry exists, it is broken at low energy scales ($\equiv M_S$) and, in principle, the mechanism responsible of that SUSY-breaking has several ways to occur. Soft-SUSY breaking refers to those mechanisms which do not lead to ultraviolet divergences in scalar masses.

One quantity used to classify supersymmetric models is called R-parity, mathematically defined as:

$$R = (-1)^{3B+L+2S}$$

where B is the baryon number, L the lepton number and S the spin. Then, SM particles have $R = 1$, while their superpartners have $R = -1$. $B - L$ invariance implies that this quantity is conserved, predicting therefore at least one stable supersymmetric particle, the so called LSP (Lightest Supersymmetric Particle). The LSP is one of the most supported candidates to explain dark matter composition. Cosmological constraints[37] indicate that LSP should be electrically neutral. On the other hand, there are R-parity violating (RPV) SUSY models which add new couplings violating either baryon number or lepton number.

2.3.6. Minimal SuperSymmetric Standard Model (MSSM)

The MSSM is the minimal SUSY extension of the SM. It is constructed by adding the corresponding partners of SM particles and two hypercharged Higgs doublets, needed for avoiding anomalies. They generate separately the masses of “up” type (s)quarks (the Higgs fields H_u^0 and H_u^+) and “down”-type (s)quarks and charged (s)leptons (the Higgs fields H_d^0 and H_d^+). The mass eigenstates corresponding to those two Higgs doublets are the five physical higgses h^0, H^0, A^0 and H^\pm . This is essentially the same Higgs structure than 2HDM-II. Neutrinos are massless like in SM, in order to generate their masses, mechanisms like seesaw[20] can be used.

Table 2-II: Superpartners of SM particles.

Field Name	Symbol	Spin
sleptons	$\tilde{e}_L, \tilde{e}_R, \tilde{\mu}_L, \tilde{\mu}_R, \tilde{\tau}_L, \tilde{\tau}_R, \tilde{\nu}_e, \tilde{\nu}_\mu, \tilde{\nu}_\tau$	0
Squarks	\tilde{q}_L, \tilde{q}_R	0
neutralinos	$\tilde{B}^0, \tilde{W}^0, \tilde{H}_u^0, \tilde{H}_d^0$	$\frac{1}{2}$
charginos	$\tilde{W}^\pm, \tilde{H}_u^\pm, \tilde{H}_d^\pm$	$\frac{1}{2}$
Gluinos	\tilde{g}	$\frac{1}{2}$
Goldstino or gravitino	\tilde{G}	3/2

The MSSM is constructed in consistency with $B-L$ number conservation, thus the R – parity is conserved and the LSP of the model must be stable. It also includes all renormalizable soft-SUSY breaking terms consistent with SM gauge symmetry in order to consider all possible mechanism of supersymmetry breaking. This leads to a large number of free parameters on the MSSM.

The fraction between the vacuum expected value of H_u and H_d is a free parameter of the model, called $\tan\beta$ ($\tan\beta = v_u/v_d$). This parameter enters on $BR(B_s \rightarrow \mu^+\mu^-)$ computation in SUSY models and, through it, this channel can be a good test on supersymmetry.

The superpartners of SM particles are shown in Table 2-II. $\tilde{B}^0, \tilde{W}^{0,\pm}, \tilde{H}_{u,d}^0$ are also called bino, winos and higgsinos, respectively. Neutralinos and charginos defined above are not, in principle, mass eigenstates. Their corresponding mass eigenstates are indicated by $\tilde{\chi}_i^0$ ($i=1\dots 4$), and $\tilde{\chi}_j^\pm$ ($j=1,2$) where lower index means lighter particle. The set of superpartners of gauge bosons of the $SU(3)_C \otimes SU(2)_T \otimes U(1)_Y$ are termed *gauginos*. Goldstino field arises if SUSY is a spontaneously-broken global symmetry. But if SUSY is a local symmetry then it is replaced (absorbed) by the gravitino, the superpartner of the gravity boson (graviton). The inclusion of graviton (and hence, gravitino) is needed because local SUSY is only possible if space-time are curved and then, the gravity included[38].

One theoretically interesting prediction of MSSM is that it improves the unification of gauge coupling constants at some high energy scale, M_U , in the order of 2×10^{16} GeV. This unification is kept if SUSY is broken at a scale $M_S \leq O(1 \text{ TeV})$. Even if gravity is included, its coupling constant seems to roughly point to the same value at the same M_U .

For phenomenological predictions, MSSM with no other constraint becomes hard to treat, because the total number of free parameters of the model is 124:

- + 18 SM parameters
- + 1 Higgs sector parameter (analogue to SM Higgs mass)
- + 5 real and 3 CP violating phases in gaugino/higgsino sector
- + 21 squark and slepton masses
- + 36 real mixing angles for squark and slepton mass eigenstates
- + 40 CP-violating phases in squark and slepton interactions.

The soft-SUSY breaking part of MSSM lagrangian density is:

$$\begin{aligned}
L_{SOFT} = & -\{\tilde{l}_L^*(M_{\tilde{l}}^2)^{ij}\tilde{l}_{Lj} + \tilde{q}_L^*(M_{\tilde{q}}^2)^{ij}\tilde{q}_{Lj} + \tilde{u}_R^*(M_{\tilde{u}R}^2)^{ij}\tilde{u}_{Rj} + \tilde{d}_R^*(M_{\tilde{d}R}^2)^{ij}\tilde{d}_{Rj} + \tilde{e}_R^*(M_{\tilde{e}R}^2)^{ij}\tilde{e}_{Rj} \\
& + m_{Hu}^2|H_u|^2 + m_{Hd}^2|H_d|^2 + (B \cdot \mu H_u H_d + h.c.) + (H_d[\tilde{l}_L^*(h_{\tilde{e}} A_{\tilde{e}})^{ij}\tilde{e}_{Rj} + \tilde{q}_L^*(h_{\tilde{d}} A_{\tilde{d}})^{ij}\tilde{d}_{Rj}] + \\
& + H_u \tilde{q}_L^*(h_{\tilde{u}} A_{\tilde{u}})^{ij}\tilde{u}_{Rj} + h.c.) + \frac{1}{2}(m_{\tilde{B}} \bar{\tilde{B}}^o P_L \tilde{B}^o + m_{\tilde{B}}^* \bar{\tilde{B}}^o P_R \tilde{B}^o) + \frac{1}{2}(m_{\tilde{W}} \bar{\tilde{W}} P_L \tilde{W} + m_{\tilde{W}}^* \bar{\tilde{W}} P_R \tilde{W}) \\
& + \frac{1}{2}(m_{\tilde{g}} \bar{\tilde{g}}^a P_L \tilde{g}^a + m_{\tilde{g}}^* \bar{\tilde{g}}^a P_R \tilde{g}^a)\}
\end{aligned}$$

But it can be reduced by imposing mathematical constraints, specifying the way in which SUSY is broken. The most common are scenarios where SUSY breaking is mediated by gravity[39], gauge fields (Gauge Mediated Supersymmetry Breaking, GMSB)[40] or as a consequence of a dominating super-Weyl anomaly (Anomaly Mediated Supersymmetry Breaking, AMSB)[41]. The simplest (minimal) versions of those scenarios are mSUGRA, mGMSB and mAMSB.

2.3.7. $B_s \rightarrow \mu^+ \mu^-$ in MSSM and constrained versions

MSSM will contain the contributions from SM and 2HDM-II, which would be the limit for high masses of SUSY particles, but also we can take the diagrams of those models and exchange the particles in the loop by its superpartners while keeping R-parity conserved. In particular, the dominant diagram at high $\tan\beta$ is shown in Figure 2-11. Apart from those, MSSM contains diagrams including quartic squark coupling.

The $\text{BR}(B_s \rightarrow \mu^+ \mu^-)$ in MSSM, including QCD corrections, is computed in [3]. Z^0 penguins, neutral Higgs penguins and Box diagrams are affected by chargino, charged Higgs and quartic squark contributions. A typical feature of this channel is the enhancement with powers of $\tan\beta$, as it happens in the case of 2HDM-II, although again values even lower than SM are also possible. The SUSY effects in the different Wilson coefficients can be summarized as follows:

- C_A : Z^0 penguin and Box diagrams can receive contributions due to charginos and quartic squark couplings. Z^0 penguin can also be affected by charged Higgs.
- $C_{A'}$: chargino, quartic squark couplings and charged Higgs can make this coefficient to be different from 0. Charged Higgs produces $\tan^2\beta$ ($\tan^4\beta$) contributions to Z^0 penguin (Box diagram) but are strongly suppressed by $\sim m_s m_b / M_W^2$ ($m_s m_b m_\mu^2 / M_W^2 M_{H^+}^2$), and even for very large $\tan\beta$ are far from being dominant.
- $C_{S,P}$: charged Higgs produces $\tan^2\beta m_\mu / M_W^2$ in both neutral Higgs and Box diagrams. A $\tan^3\beta m_\mu / M_W M_{A0}$ term arises from chargino contributions (a $\tan^4\beta$ term also appears, but suppressed by s(d) quark mass) to neutral Higgs penguin. Finally, quartic squark couplings in neutral Higgs penguin also produce $\tan^3\beta m_\mu / M_W M_{A0}$ terms.
- $C_{S',P'}$: In MSSM have comparable size to $C_{S,P}$ and hence will be suppressed.

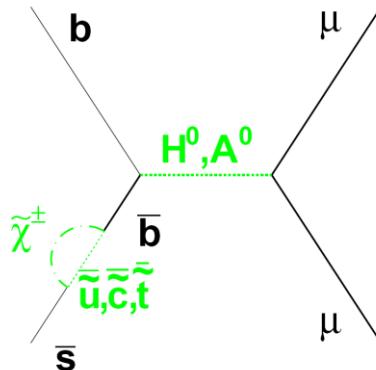


Figure 2-11: Dominant MSSM diagram at high $\tan\beta$.

A wide range of values for $\text{BR}(B_s \rightarrow \mu^+ \mu^-)$ are possible within MSSM, from lower than SM to current experimental upper limit. In order to make more explicit predictions and understand the impact of this measurement in the parameter space, constrained MSSM's become more suitable. Several constrained MSSM's, as well as the corresponding relations with $\text{BR}(B_s \rightarrow \mu^+ \mu^-)$, are analyzed in the following sections.

2.3.7.1. mSUGRA

In mSUGRA, SUSY breaking is mediated by gravity. The number of free parameters of MSSM is reduced by a set of universality conditions at the GUT scale (M_U):

Universal gaugino mass parameters:

$$m_{\tilde{B}}(M_U) = m_{\tilde{W}}(M_U) = m_{\tilde{g}}(M_U) \equiv m_{1/2}$$

Universal scalar mass parameters:

$$M_{\tilde{l}}^2(M_U) = M_{\tilde{q}}^2(M_U) \equiv m_0^2 I_3$$

$$M_{\tilde{u}}^2(M_U) = M_{\tilde{e}}^2(M_U) = M_{\tilde{d}}^2(M_U) = m_0^2 I_3$$

$$m_{Hu}^2 = m_{Hd}^2 = m_0^2$$

Universal trilinear couplings:

$$A_{\tilde{u}}(M_U) = A_{\tilde{e}}(M_U) = A_{\tilde{d}}(M_U) \equiv A_0 I_3$$

Here I_3 is the 3x3 identity matrix.

Those conditions let mSUGRA with only 5 free parameters apart from SM ones:

$$A_0, \tan \beta, m_0, m_{1/2}, \text{sign}(\mu)$$

The gravitino mass, $m_{3/2}$ is equal to the scalar mass m_0 within mSUGRA.

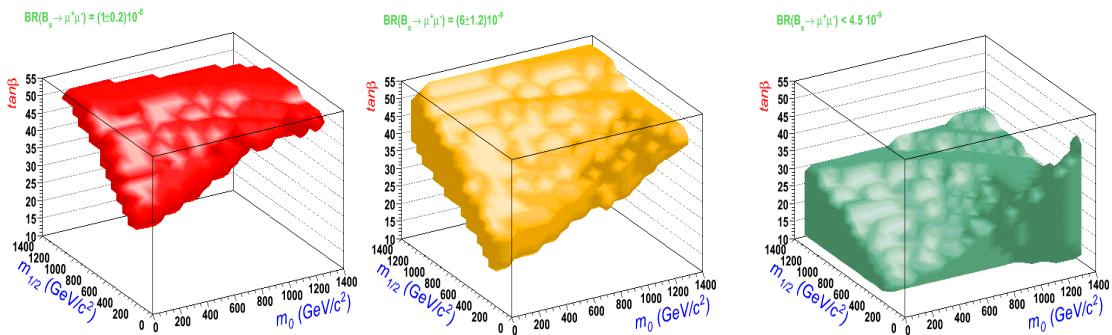


Figure 2-12: mSUGRA parameter space regions compatible with different BR ($B_s \rightarrow \mu^+ \mu^-$).

Left: $\text{BR} = (1 \pm 0.2) \cdot 10^{-8}$. Center: $4.8 \times 10^{-9} < \text{BR} < 7.2 \times 10^{-9}$. Right: $\text{BR} < 4.5 \times 10^{-9}$. The following constraints were applied: $\mu > 0$, $A_0 = 0$, m_h (lightest Higgs mass) > 114 GeV and $M_{W^\pm} = 80.398 \pm 0.025$ GeV.

Figure 2-12 shows three 3D plots, which are the mSUGRA regions compatible with $\text{BR}(B_s \rightarrow \mu^+ \mu^-) = (1 \pm 0.2) \cdot 10^{-8}$, $4.8 \times 10^{-9} < \text{BR}(B_s \rightarrow \mu^+ \mu^-) < 7.2 \times 10^{-9}$ and $\text{BR}(B_s \rightarrow \mu^+ \mu^-) < 4.5 \times 10^{-9}$. The plots have been produced using SoftSUSY² and SUITY v-1.0³.

2.3.7.2. CMSSM

CMSSM stands for Constrained MSSM. The boundaries applied are similar to those of mSUGRA, although mSUGRA is more restrictive. In particular CMSSM does not include the relation of the gravitino mass, $m_{3/2} = m_0$.

² Provided by Ben Allanach (Cambridge).

³ Provided by Athanasios Dedes (Durham).

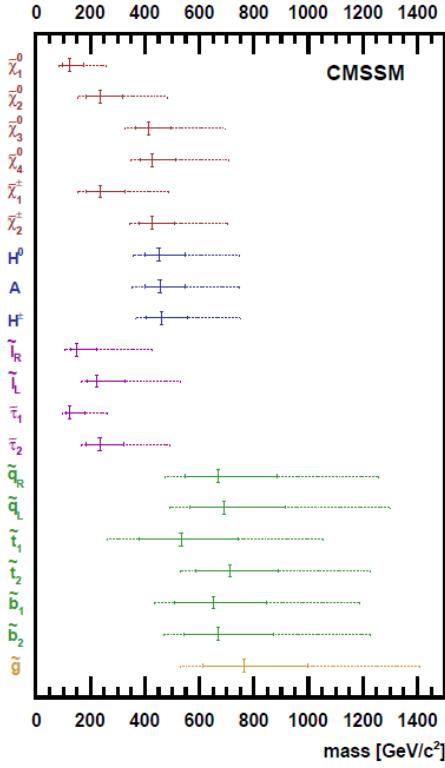


Figure 2-13: Masses of SUSY particles in CMSSM.

The vertical solid lines indicate the best-fit values in [42], the horizontal solid lines are the 68% C.L. ranges, and the horizontal dashed line are the 95% C.L. ranges for the indicated mass parameters.

Taking into account the EW precision data, $(g-2)_\mu$, B physics and cosmological measurements, [42] studies the parameter space of CMSSM and the implications to other observables including $\text{BR}(B_s \rightarrow \mu^+ \mu^-)$. The obtained masses for the SUSY particles are shown in Figure 2-13. Figure 2-14 shows the variation of the χ^2 of CMSSM fit with the value of $\text{BR}(B_s \rightarrow \mu^+ \mu^-)$. From that plot we can see that CMSSM prefers SM-like values of this BR, and enhancements are strongly disfavored.

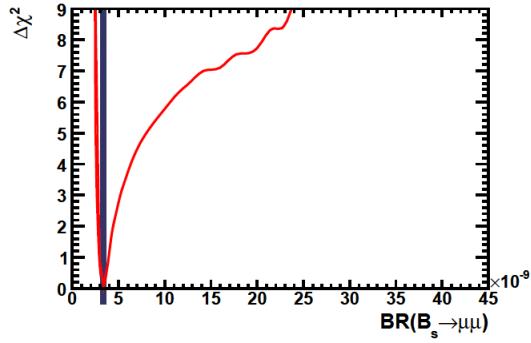


Figure 2-14: $\Delta\chi^2$ of CMSSM fit[42] as a function of $\text{BR}(B_s \rightarrow \mu^+ \mu^-)$.

The vertical line indicates the SM value with its theoretical error.

2.3.7.3. NUHM

NUHM (Non Universal Higgs Masses)[43] [44] reduces the MSSM parameter phase space in almost the same way as CMSSM, but avoiding the condition of universal Higgs masses, providing a framework for the study of the Higgs sector. Relaxing such condition lets the parameters M_A and μ to be free parameters for any specified value of m_0 , $m_{1/2}$, $\tan\beta$ and A_0 . The parameter M_A is the mass of the physical CP-odd neutral higgs A^0 , $M_A = \frac{\mu B}{v_u v_d} (v_u^2 + v_d^2) = \frac{2\mu B}{\sin 2\beta}$.

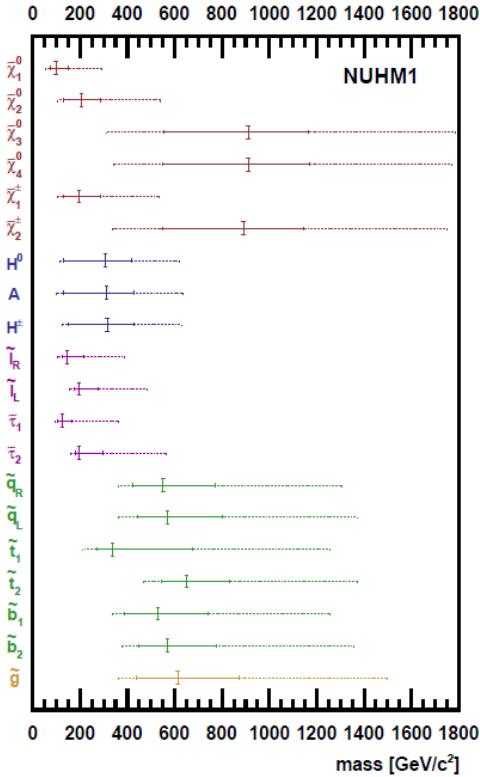


Figure 2-15: Masses of SUSY particles in NUHM.

The vertical solid lines indicate the best-fit values in [42], the horizontal solid lines are the 68% C.L. ranges, and the horizontal dashed line are the 95% C.L. ranges for the indicated mass parameters.

Reference [42] performs on NUHM the same study done for CMSSM. Figure 2-15 shows the obtained SUSY particle mass spectrum. We can see in this plot that NUHM prefers lower values for the Higgs masses than the CMSSM, which allows for larger departures of $\text{BR}(B_s \rightarrow \mu^+ \mu^-)$ with respect to the SM prediction. Figure 2-16 shows the variation of the χ^2 of NUHM fit with the value of $\text{BR}(B_s \rightarrow \mu^+ \mu^-)$. From that plot we can see that NUHM allows enhancements up to current experimental upper limit. The right plot on Figure 2-16 shows the relation of $\text{BR}(B_s \rightarrow \mu^+ \mu^-)$ with $\tan\beta$, where we see that the larger enhancements occur at high values of $\tan\beta$.

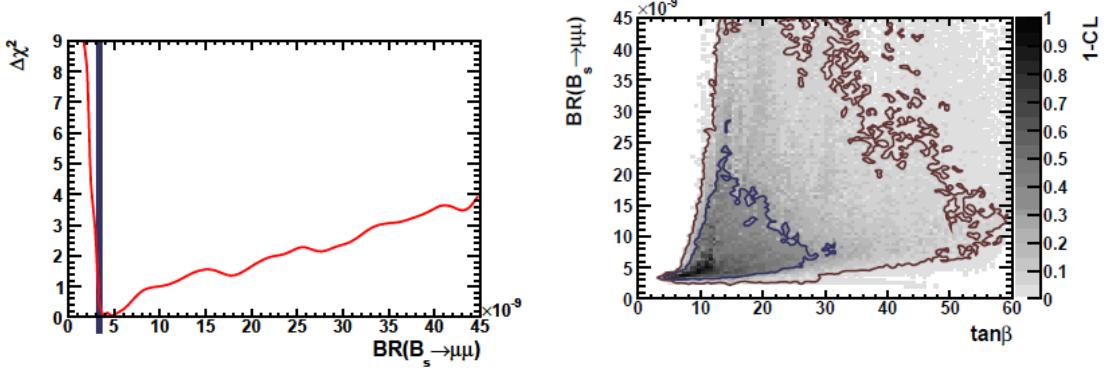


Figure 2-16: $\Delta\chi^2$ of NUHM fit[42] as a function of $\text{BR}(B_s \rightarrow \mu^+\mu^-)$ and different CL regions as a function of $\tan\beta$.

The vertical line on the left plot indicates the SM value with its theoretical error.

2.3.7.4. Maximal CP violation Minimal Flavor Violation (MCPMFV)-MSSM

MCPMFVMSSM[45] lets $m_{1/2}$ and A_0 to be complex with CP-odd phases. In this context, FCNC observables still vanish if the CKM matrix is equal to I_3 , and that is the reason to be named MFV. Then, the maximum number of CP-violating phases and extra flavor-singlet mass scales are added while keeping that MFV notion. The total number of parameters added to SM is then 19: 6 CP-violating phases and 13 real mass parameters. Thus, this model becomes useful to study effects of new CP violation sources introduced by SUSY.

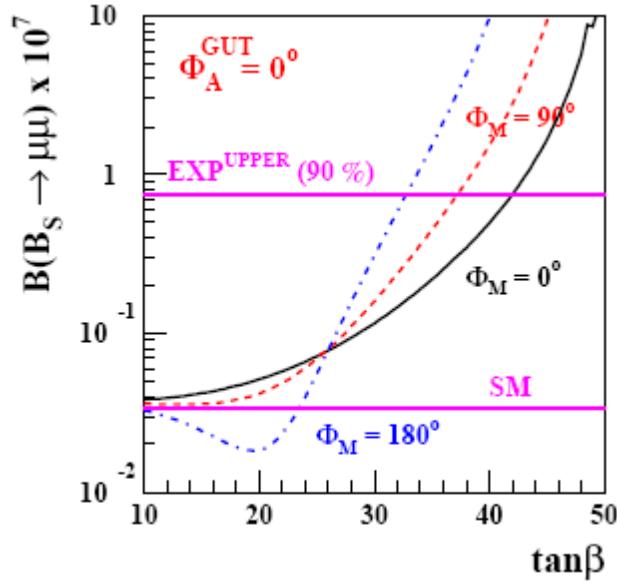


Figure 2-17: $\text{BR}(B_s \rightarrow \mu^+\mu^-)$ dependence with $\tan\beta$ in MCPMFVMSSM.

The dependence is shown for three different values of the phase Φ_M of gaugino mass parameters, and the phase of trilinear terms fixed to $\Phi_A^{\text{GUT}} = 0^\circ$. The values of the real mass parameters were chosen as: $m_{1/2} = 250$ GeV, $m_0 = 100$ GeV, $A_0 = 100$ GeV.

In Figure 2-17 shows an example taken from [45] is shown. In general, $\text{BR}(B_s \rightarrow \mu^+\mu^-)$ increases with $\tan\beta$ due to the enhancement of Wilson coefficients C_S and C_P with this parameter. However, for $\Phi_M \sim 180^\circ$ and $\tan\beta \sim 20$, a value below SM is predicted. This is due to a cancellation of C_P term with the SM term (C_A).

2.3.7.5. AMSB

The prediction of this model for $\text{BR}(\text{B}_s \rightarrow \mu^+ \mu^-)$ is published in [46]. The BR as a function of $\tan\beta$ and $m_{3/2}$ is shown in Figure 2-18 for $\mu < 0$ (left) and $\mu > 0$ (right). The blue dashed line corresponds to the region compatible with 5×10^{-9} and the green continuous line to 1×10^{-8} . Yellow dot-dashed line marks the limit excluded by direct searches of SM Higgs at LEP and magenta dotted line marks the limit from $b \rightarrow s\gamma$. An additional constraint can be imposed if one takes into account the results of $(g-2)_\mu$, which interpreted in the context of AMSB forces the μ parameter to be positive. In this case there is no possibility of large enhancement of $\text{BR}(\text{B}_s \rightarrow \mu^+ \mu^-)$ within AMSB MSSM.

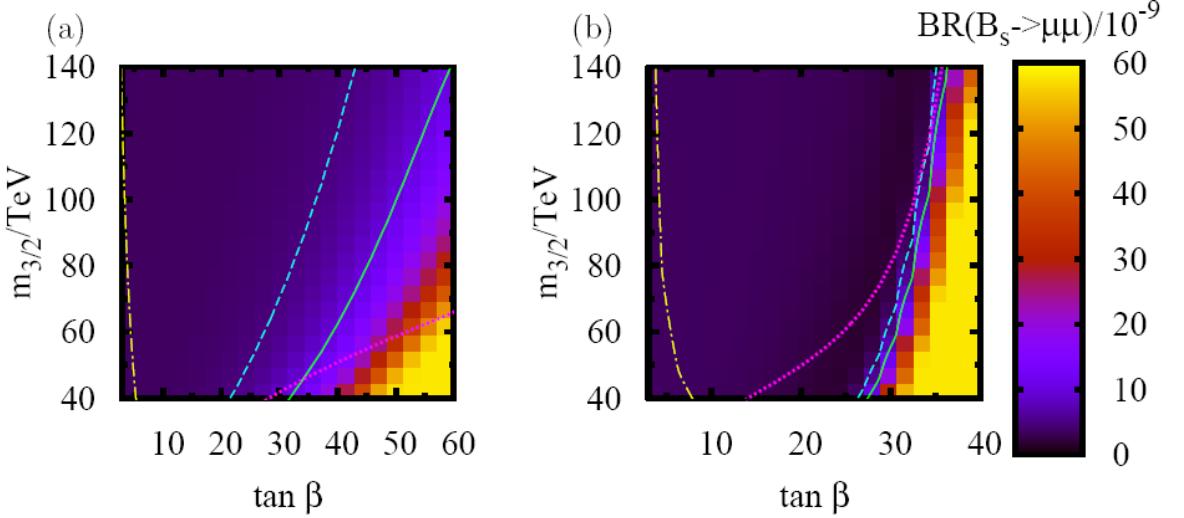


Figure 2-18: AMSB predictions for $\text{BR}(\text{B}_s \rightarrow \mu^+ \mu^-)$.

2.3.8. Tree level contributions in RPV MSSM

In order to allow R-parity violation in MSSM, the following terms appear in the superpotential[47]:

$$W_{RPV} = \mu_i H_u L_i + \frac{1}{2} \lambda_{ijk} L_i L_j E_k^c + \lambda'_{ijk} L_i Q_j D_k^c + \frac{1}{2} \lambda''_{ijk} U_i^c D_j^c D_k^c$$

i, j, k are generation indices, L_i and Q_i the lepton and quark SU(2)_L doublets and E^c , D^c and U^c are the charge conjugated of the right handed singlets. The terms with λ and λ' violate lepton number and the terms with λ'' violate baryon number. The constant λ_{ijk} (λ''_{ijk}) is antisymmetric in the first (last) two indices.

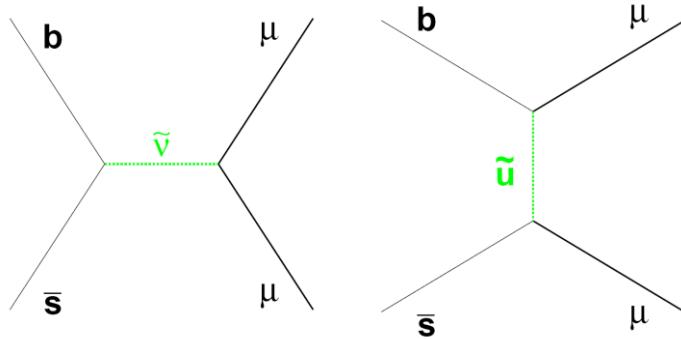


Figure 2-19: RPV Feynman diagrams.

A feature of RPV SUSY is that the decay $\text{B}_s \rightarrow \mu^+ \mu^-$ can happen at tree level, mediated by a sneutrino or up-type squark, as shown in Figure 2-19. The contributions to the different Wilson coefficients are [48][49].

$$\begin{aligned}
C_A^{RPV \text{ term}} &= \frac{\sqrt{2}\pi \sin^2 \theta_W}{4V_{tb}V_{tq}^*G_F\alpha} \sum_{m,n,i=1}^3 V_{ni}^\dagger V_{im} \frac{\lambda'_{\mu n 3} \lambda'^*_{\mu m q}}{m_u^2 c_i} \\
C_S^{RPV \text{ term}} &= \frac{\sqrt{2}\pi \sin^2 \theta_W}{(m_b + m_q)V_{tb}V_{tq}^*G_F\alpha} \sum_{i=1}^3 \frac{\lambda_{i\mu\mu}^* \lambda'_{iq3} - \lambda_{i\mu\mu} \lambda'^*_{i3q}}{m_{\tilde{\nu}_{L,i}}^2} \\
C_P^{RPV \text{ term}} &= (-) \frac{\sqrt{2}\pi \sin^2 \theta_W}{(m_b + m_q)V_{tb}V_{tq}^*G_F\alpha} \sum_{i=1}^3 \frac{\lambda_{i\mu\mu}^* \lambda'_{iq3} + \lambda_{i\mu\mu} \lambda'^*_{i3q}}{m_{\tilde{\nu}_{L,i}}^2}
\end{aligned}$$

Thus extra NP effects can appear apart from those of the MSSM if RPV is allowed. The negative sign in parenthesis indicates the relative sign when combined with the 2HDM-II contribution as it is written in 2.3.3. Note that the EW and CKM coefficients appearing there are just to cancel the global ones in the BR expression. It implies that the reduction factor $\sim |V_{tb}^* V_{ts}|^2 / |V_{tb}^* V_{ts}|^2 \sim 1/20$ that penalizes $B_d \rightarrow \mu^+ \mu^-$ with respect to $B_s \rightarrow \mu^+ \mu^-$ does not apply in case of R-parity violation. In [50] the case of RPV mSUGRA is analyzed, the dependence with the strength of RPV couplings is shown in Figure 2-20.

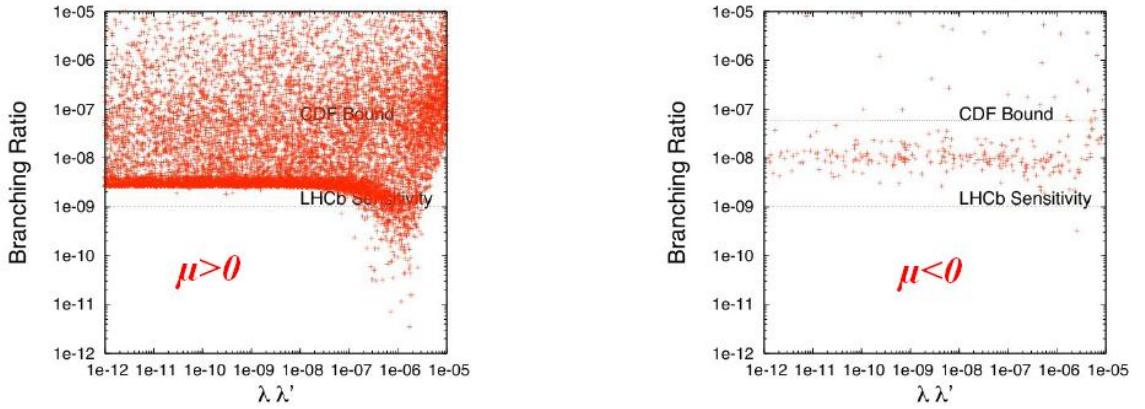


Figure 2-20: Dependence with RPV couplings in RPV mSUGRA.

2.3.9. Extra dimensions

The possibility of adding extra dimensions (hereafter, EDs) to the usual 4-D space-time was first proposed by Kaluza[51] and Klein [52] in order to unify electromagnetism and gravity into a common origin. The unification of gauge and gravity interactions is indeed one of the main motivations for the studies of extra dimensions, but they also provide a good framework for gravity quantization, addressing the Higgs mass hierarchy problem or the cosmological constant.

In the Kaluza-Klein picture, the extra dimensions are compact, with a microscopic size L ensuring that space-time is effectively four dimensional at distances $\gg L$. However, in the so called *braneworld* picture the mechanism for hiding extra dimensions is performed by trapping the ordinary matter (apart from graviton⁴) into a 3D submanifold (*brane*) embedded in the fundamental multidimensional space so that the extra dimensions can be large or even infinite.

Depending on the structure of EDs, the braneworlds are classified as[53]:

- Braneworlds with compact EDs: Also called ADD [54] (from Askani-Hamed, Dimopoulos and Dvali).
- Braneworlds with warped EDs or Randall-Sundrum (R-S) models [55].

⁴ Allowing the graviton to propagate through EDs is often used in order to explain the weakness of gravity w.r.t the other interactions.

- Infinite volume EDs [56].

2.3.10. $B_s \rightarrow \mu\mu$ in Warped Extra Dimensions

The $\text{BR}(B_s \rightarrow \mu\mu)$ has been computed in the context of Randall-Sundrum models. In reference [57] the computations corresponding to the specific model [58] with gauge symmetry in the bulk $SU(3)_C \otimes SU(2)_R \otimes SU(2)_L \otimes P_{L,R} \otimes U(1)_X$ are shown. In that case the deviations from SM predicted are very small unless the custodial symmetry imposed for consistency with $Z b_L \bar{b}_L$ is removed (see Figure 2-21).

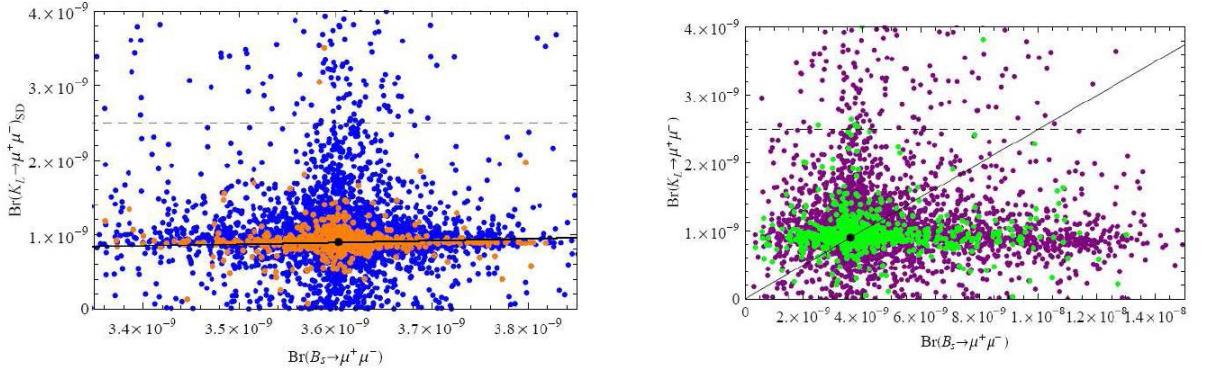


Figure 2-21: $\text{BR}(B_s \rightarrow \mu\mu)$ vs $\text{BR}(K_L \rightarrow \mu\mu)_{\text{SD}}$ in the R-S model[58].

Left: with custodial protection. Right: without custodial protection. The dashed line shows the experimental upper limit in $\text{BR}(K_L \rightarrow \mu\mu)_{\text{SD}}$ (the short distance contribution to $\text{BR}(K_L \rightarrow \mu\mu)$). Solid line represents the CMFV prediction.

However, in the studies of Randall-Sundrum models performed in [59], using $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ bulk gauge symmetry and brane localized Higgs sector, larger departures from the SM are found. In that context, $\text{BR}(B_s \rightarrow \mu\mu)$ can be as large as current experimental upper limit, and also different NP effects are seen in B_s and B_d so that the ratio of branching ratios can be very different from SM prediction (see Figure 2-22).

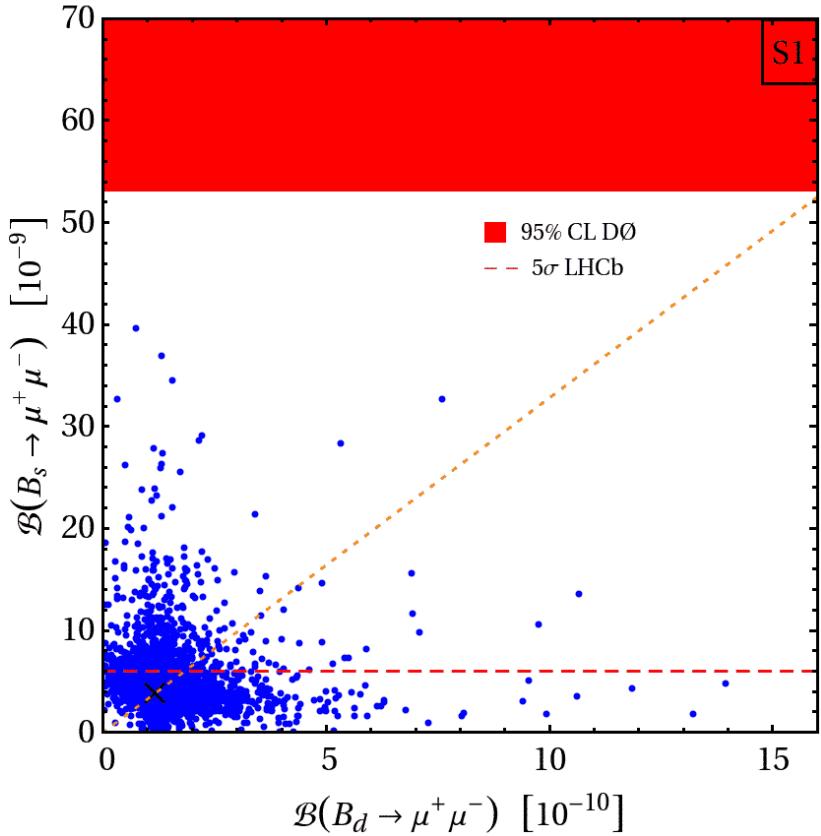


Figure 2-22: $\text{BR}(B_s \rightarrow \mu\mu)$ vs $\text{BR}(B_d \rightarrow \mu\mu)$ in the R-S model[59].

The black cross indicates the SM point. The 95% CL upper limit on $\text{BR}(B_s \rightarrow \mu\mu)$ from D0 is indicated by the red band. The orange dotted line represents the CMFV correlation between the two purely leptonic modes.

2.3.11. Technicolor models

Technicolor (TC) theories base on the idea of composite scalar fields as mechanism for EWSB instead of an elementary Higgs. This is motivated because, in fact, even in the case of no Higgs the W and Z bosons would acquire mass due to QCD effects: if $m_u = m_d = 0$ the pions would act as a (pseudo)Goldstone boson. However, the mass of the EW bosons obtained by this mechanism would be only few MeV. Technicolor uses then a new strong interaction which generates technihadrons that can account for the observed vector boson masses.

As there is no Higgs sector, in such theories in principle everything could be determined in terms of the gauge couplings of the theory. However, without the Higgs, the masses of the fermion field have to arise by radiative corrections from the exchange of a heavy boson. In order to explain the large differences between the fermion generations another gauge group, called extended Technicolor (ETC), is postulated. The (minimal) ETC that could give masses to the fermions would have the final group:

$$G = G_{ETC} \times G_{SM}$$

Where G_{ETC} would then break in the way:

$$G_{ETC} \longrightarrow G_{ETC} \longrightarrow G_{E''TC} \longrightarrow G_{TC}$$

And each of the symmetry breaking would be responsible of each of the generations. Essentially, this translates the problem of understanding the pattern of fermion masses to the understanding of that sequence of breakings.

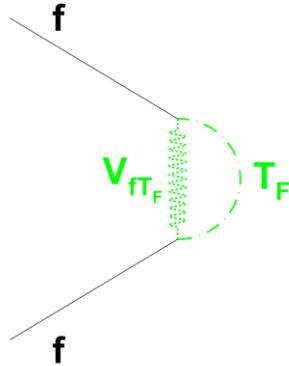


Figure 2-23: Fermion masses in ETC.

G_{ETC} includes then the bosons coupling fermions to technifermions that give masses to the later (see Figure 2-23) and, in order to close the algebra, it predicts also bosons coupling fermions of different generations.

Technicolor theories naturally provide a DM candidate, the lightest technibaryon[60].

However, the simplest TC models suffer from inconsistency with experimental observations, either for the prediction of unobserved particles or large FCNC because its prediction of flavor changing bosons.

2.3.12. Topcolor assisted Technicolor (TC2)

One way of allowing technicolor to be consistent with observed phenomenology is through the introduction of a $t - \bar{t}$ condensate due to a new strong interaction in the top system[61][62]. This condensate can produce several Goldstone bosons and an extra Higgs boson. The combination of this interaction with ETC is called Topcolor assisted Technicolor (TC2) [63][64]. The top condensation at the weak scale also gives rise to the large observed top quark mass, while ETC gives masses mostly to the light fermions and a small fraction of the top mass ϵm_t . The combination of topcolor with ETC makes one theory to help each other in solve its deficiencies.

The topcolor (or coloron) group $G_{CLN} = [SU(3)_1 \times U(1)_1] \times [SU(3)_2 \times U(1)_2]$ is broken by a condensate of techniquarks into $SU(3)_C \times U(1)_Y$, which is then broken by another techniquark condensate into $SU(3)_C \times U(1)_{EM}$:

$$[SU(3)_1 \times U(1)_1] \times [SU(3)_2 \times U(1)_2] \times SU(2)_T \xrightarrow{\langle \bar{Q}Q \rangle_0} G_{SM} \xrightarrow{\langle \bar{T}T \rangle_0} SU(3)_C \times U(1)_{EM}$$

From the first breaking, 9 new massive vector bosons appear: 8 colorons and a (non universal) color singlet Z' . All the nine should have masses in the order of $\sim 10^{3-4}$ GeV.

The $t - \bar{t}$ condensate also enters in the EWSB mechanism, but its contribution is rather small in TC2. Thus the main component of the EW Goldstone bosons is provided by technifermions, and then a triplet of pseudo Goldstone bosons from the $t - \bar{t}$ condensate remains “uneaten”. The mass fraction ϵm_t of the top quark induced by ETC gives mass (\sim hundreds of GeV) to the particles of this triplet that are called then top-pions.

The combination of topcolor masses and ETC masses produces a general fermion mass matrix which can be diagonalized leading to the CKM matrix.

The exchanges of top-pions, top-pions or Z' can lead to flavor changing processes, and in particular to large enhancement of FCNC's in the B sector.

2.3.13. $B_s \rightarrow \mu\mu$ in TC2

The main contributions to $B_s \rightarrow \mu\mu$ from TC2 can be found in [65]. Although this decay can happen at tree level due to the Z' boson (see first diagram on Figure 2-24), this contribution has been found small with respect to SM due to weak couplings of Z' to first and second generation fermions (it is different in the case of $B \rightarrow \tau\tau$, where the Z' can produce significant enhancement). This is therefore in consistency with the results in [4], from which no significant NP is expected in C_A . Charged top-pions (see last three diagrams on Figure 2-24) would also contribute to C_A so their masses should be large enough to make such contribution negligible.

On the other hand top-higgs and top-pions (2nd and 3rd diagrams on Figure 2-24) can contribute to the scalar and pseudoscalar Wilson coefficients.

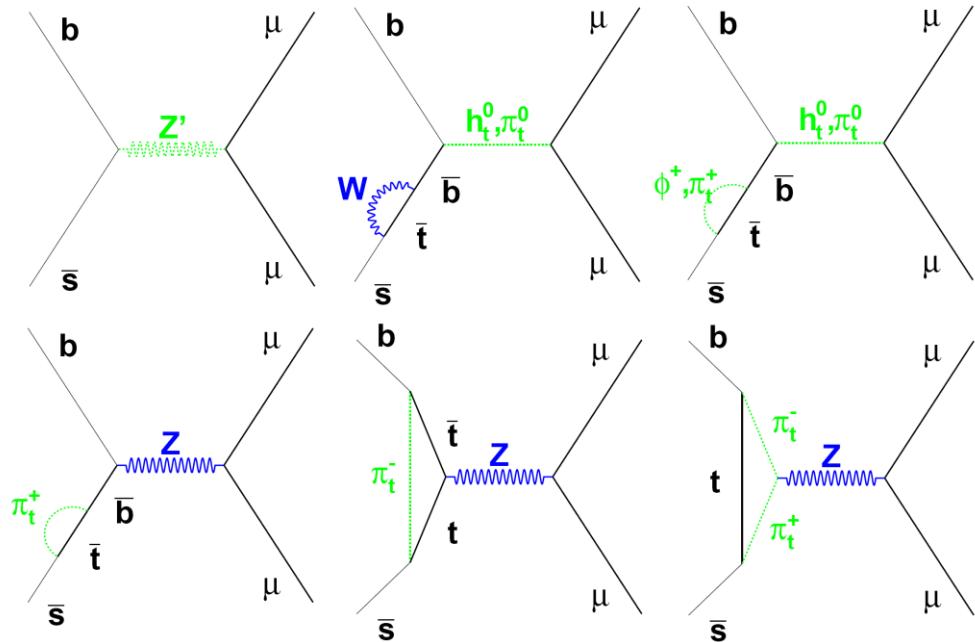


Figure 2-24: TC2 Feynman diagrams.

The contributions from TC2 to the scalar and pseudoscalar Wilson coefficients have been computed to be:

$$C_S^{TC2} = \frac{m_b^* m_\mu \sqrt{v^2/2 - F_\pi^2}}{2(m_b + m_q) F_\pi m_{h_0^t}^2} C\left(\frac{m_t^2}{M_W^2}\right) + \frac{m_b^{*2} m_\mu m_t M_W^2 V_{ts} \sin^2 \theta_W \sqrt{v^2/2 - F_\pi^2}}{4(m_b + m_q) F_\pi^3 v^2 m_{h_0^t}^2 g_2^4} C\left(\frac{m_t^{*2}}{M_S^2}\right)$$

$$C_P^{TC2} = \frac{m_b^* m_\mu \sqrt{v^2/2 - F_\pi^2}}{2(m_b + m_q) F_\pi M_S^2} C\left(\frac{m_t^2}{M_W^2}\right) + \frac{m_b^{*2} m_\mu m_t M_W^2 V_{ts} \sin^2 \theta_W \sqrt{v^2/2 - F_\pi^2}}{4(m_b + m_q) F_\pi^3 v^2 M_S^2 g_2^4} C\left(\frac{m_t^{*2}}{M_S^2}\right)$$

Being F_π the top-pion decay constant, $F_\pi \sim 50$ GeV, $m_t^* = m_t(1-\epsilon)$ is the topcolor contribution to the top mass ($\epsilon \sim 0.01 - 0.1$), $m_b^* = m_b - 0.1 \epsilon m_t$, $m_{h_0^t}$ the mass of the top-higgs and M_S the mass of the top-pions. The function C is the expression (55) in [65]. Although those coefficients are about 100 times larger than in the SM, the total contribution to $\text{BR}(B_s \rightarrow \mu\mu)$ is still very small, and thus no large deviations from the SM are expected in TC2.

2.3.14. Little Higgs Models

Little Higgs models (LHM)[66][67] attempt to stabilize the Higgs mass using the idea of light composite Higgs, which arises as a pseudo-Goldstone boson produced by the breakdown of some global approximate symmetry at the TeV scale. They are based on extended gauge interactions with a breaking pattern producing three energy scales of order $\sim 10\text{-}100 \text{ TeV}$, $\sim 1\text{-}5 \text{ TeV}$ and 250 GeV . The gauge symmetry should contain a subgroup breaking like:

$$G_{SM'} \times G_{New} \rightarrow G_{SM}$$

The minimal version has the form:

$$SU(3) \times [SU(2)_1 \times U(1)_1] \times [SU(2)_2 \times U(1)_2] \rightarrow G_{SM}$$

And thus at least four new heavy vector bosons (B' , Z' and W^{\pm}), are predicted. The particle content of LHM includes light Higgs doublets (with possibility of more extra light scalars), heavy vector bosons, heavy Higgs multiplets and heavy up-type-like quarks.

The Littlest Higgs model [68] contains $[SU(2)_1 \times U(1)_1] \times [SU(2)_2 \times U(1)_2] \rightarrow SU(2)_T \times U(1)_Y$ gauge symmetry breaking, with the group $[SU(2)_1 \times U(1)_1] \times [SU(2)_2 \times U(1)_2]$ embedded into $SU(5)$. Global $SU(5)$ breaks down to $SO(5)$ generating 14 Goldstone bosons. 4 of them are eaten by the B' , Z' , W^{\pm} . The remaining 10 scalars will constitute the SM Higgs doublet and six heavy scalars: A^0 , H^0 , H^{\pm} , $H^{\pm\pm}$.

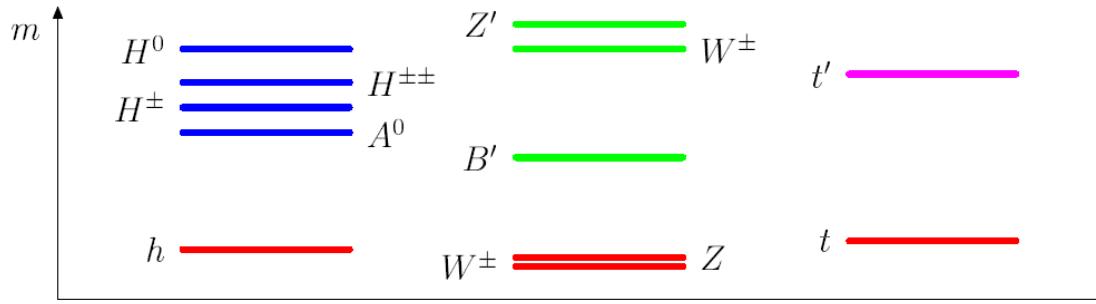


Figure 2-25: Example of mass spectrum in Littlest Higgs model.

2.3.15. Littlest Higgs with T – parity (LHT)

T-parity[69][70] is a discrete symmetry introduced in Littlest Higgs Model in order to keep NP at the TeV scale and then stabilize the EW scale. This also allows the model to be consistent with precision data without fine tuning and to provide a DM candidate (like the R-parity does for SUSY), the lightest T-odd particle.

The new bosons predicted by LHT are T-odd so do not contribute to tree processes with external SM particles. The heavy-top quark, t' , is T-even. In order to be phenomenologically viable, LHT requires the introduction of large mass “mirror” fermions: three doublets of T-odd “mirror leptons”, three doublets of T-odd “mirror quarks” and a T-odd heavy-top quark.

2.3.16. $B_s \rightarrow \mu\mu$ in LHT

The NP contributions from LHT to the process $B_s \rightarrow \mu\mu$ [71] come (like in TC2) from the axial coefficient C_A and thus they cannot exceed largely the SM prediction. Indeed, it can be seen in Figure 2-26 that this enhancement is 30-50 % at most.

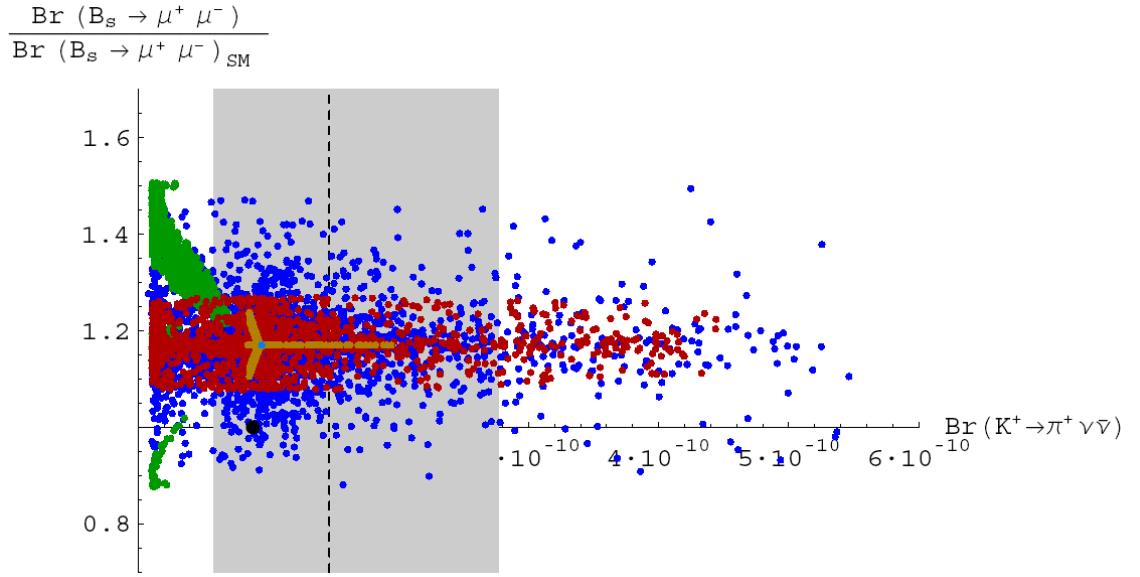


Figure 2-26: $\text{BR}(\text{B}_s \rightarrow \mu^+ \mu^-) / \text{BR}(\text{B}_s \rightarrow \mu^+ \mu^-)_{\text{SM}}$ versus $\text{BR}(K^+ \rightarrow \pi^+ \nu \bar{\nu})$ in LHT.

The black dot is the SM prediction, the dashed area is the 1σ region of $\text{BR}(K^+ \rightarrow \pi^+ \nu \bar{\nu})$ and the points in the different colors are different LHT scenarios (see[71]).

Larger effects have been predicted in the case of the golden CMFV relation (see 2.3.4), where the ratio $r = \left| \frac{c_A^{Bs}}{c_A^{Bd}} \right|^2 \frac{\Delta M_d / \Delta M_d^{SM}}{\Delta M_s / \Delta M_s^{SM}}$ which is 1 in SM and any other CMFV, can vary in the range $0.6 \leq r \leq 1.7$, with some dependency with δ_{13}^d , one of the phases of the CKM-like mixing matrix for the down-type mirror quarks (see Figure 2-27).

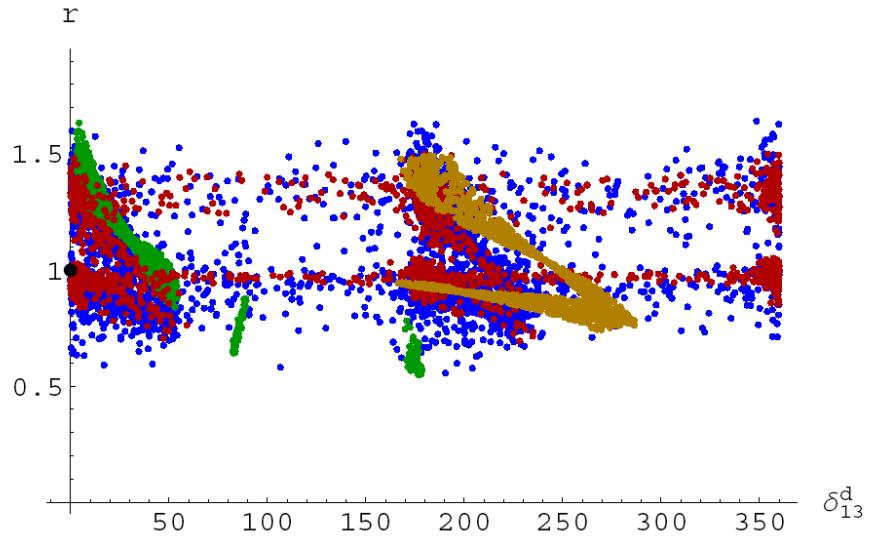


Figure 2-27: Variation of r as a function of δ_{13}^d in LHT.

The black dot is the SM prediction, the phase δ_{13}^d is in degrees and the points in the different colors are different LHT scenarios (see[71]).

3.Experimental conditions

3.1.LHC Accelerator

The Large Hadron Collider[72] (LHC) is a 27 km circular accelerator at CERN, designed to produce proton-proton collisions at energy in center of masses up to 14 TeV, allowing studying TeV energy scale, where New Physics is expected to appear to address the hierarchy problem.

Protons are extracted from ion sources like duoplasmatron, as described in[73] and [74]. The accelerator chain is shown in Figure 3-1, the energies reached by the protons at the end of each accelerator are:

- Proton LINear ACcelerator (LINAC): Up to 50 MeV
- Proton Synchrotron Booster (PSB) : 1.4 GeV
- Proton Synchrotron (PS) : 26 GeV
- Super Proton Synchrotron (SPS): 450 GeV
- LHC: 7 TeV

Also heavy ion collisions are planned for the ALICE experiment. ATLAS and CMS will also collect heavy ion collisions data.

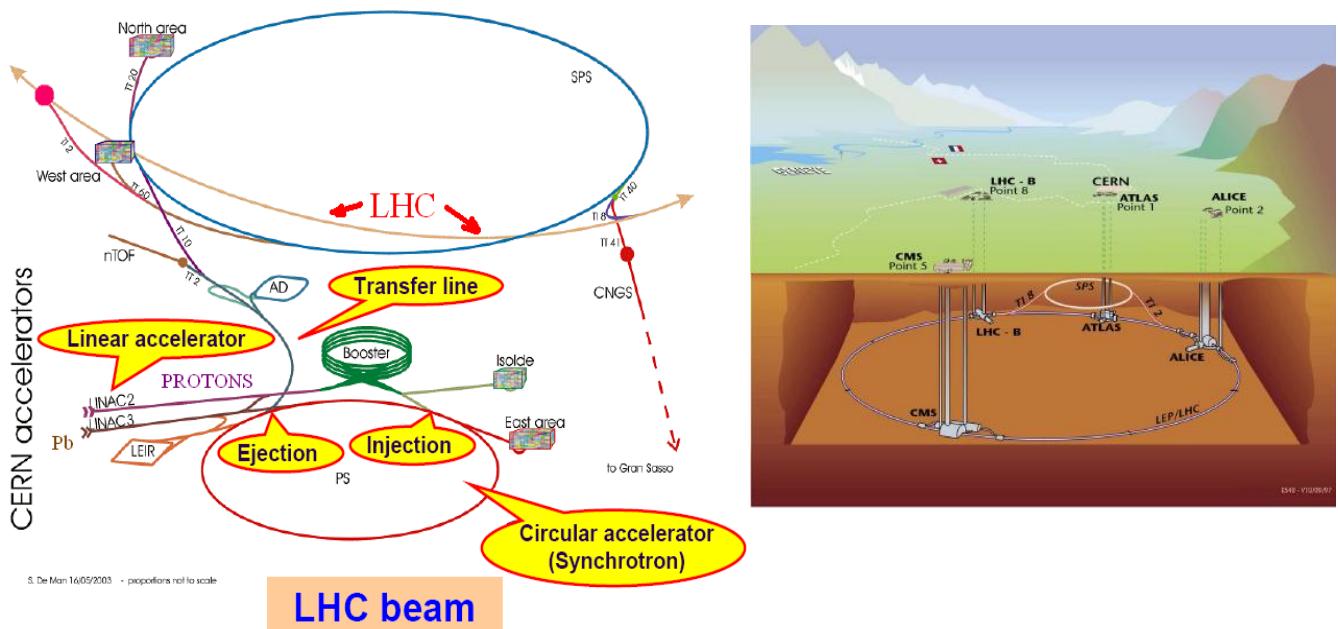


Figure 3-1: LHC accelerator chain

Protons are sent on bunches containing up to 1.15×10^{11} particles, and crossing with a rate of 40 MHz. The maximum luminosity of the accelerator is $10^{34} \text{ cm}^{-2} \text{s}^{-1}$.

3.1.1. LHC experiments

The experiments installed at LHC are:

- **ALICE** (*A Large Ion Collider Experiment*)[75]: ALICE is dedicated to the study of heavy nuclei collisions, looking for better understanding of nuclear matter phase transition and evidences of quark-gluon plasma.
- **ATLAS** (*A Toroidal LHC Apparatus*)[76]: Dedicated to search the Higgs boson, b and t quark physics, SUSY and, in general, signals of NP
- **CMS** (*Compact Muon Solenoid*)[77]: With the same objectives than ATLAS, CMS is in competition with it.
 - **TOTEM**[78]: Measures the elastic and inelastic p-p crossections at the LHC, allowing then the monitoring of the luminosity at the CMS interaction point.
- **LHCb** (*Large Hadron Collider beauty*)[79]: Study of CP violation in the b sector and rare decays of b hadrons.
- **LHCf** [80]: Understanding of high energy cosmic ray phenomena through the analysis of high energy photons and neutrons produced at low angle in ATLAS interaction point.

3.2. LHCb Experiment

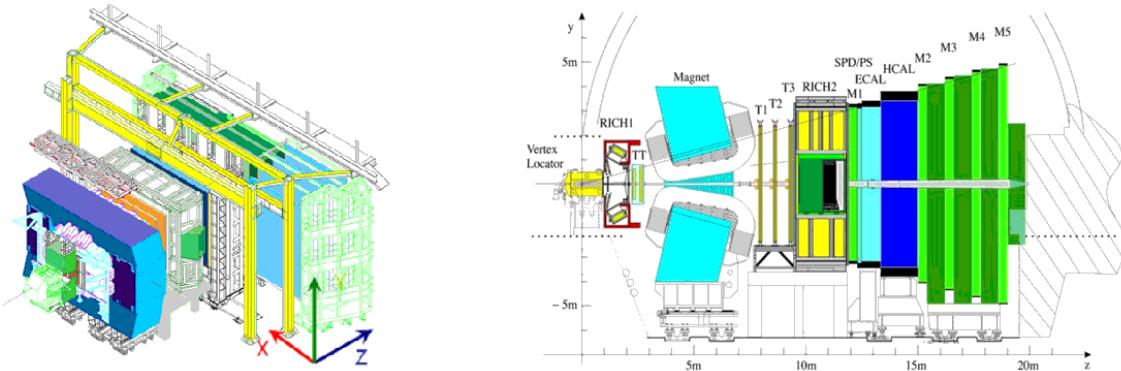


Figure 3-2: LHCb detector.

The LHCb experiment was designed for precise study of B decays. The key channels that are going to be studied and their analysis are explained in [81]. Apart from $B_s \rightarrow \mu^+ \mu^-$, the list includes the studies of Unitarity Triangle through the measurement of the angle γ (see Figure 3-3), the measurement of the electroweak phase ϕ_s , the study of exclusive $b \rightarrow s\gamma$ decays, and the rare decay $B_d \rightarrow K^{*0} \mu^+ \mu^-$.

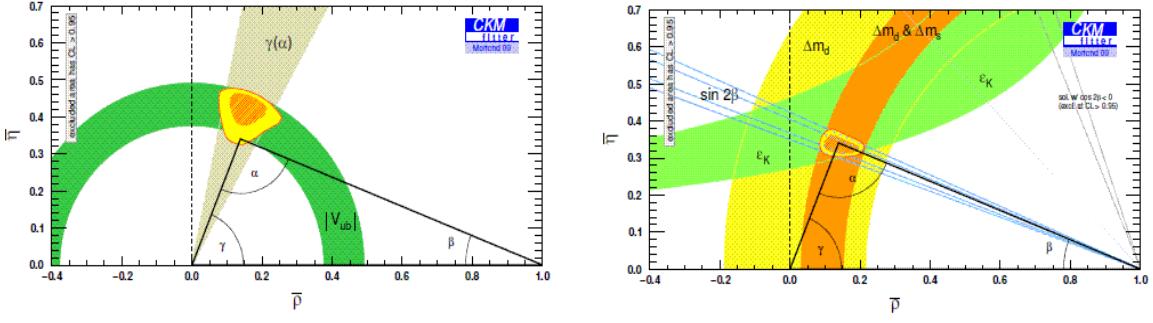


Figure 3-3: Unitarity Triangle.

Left: From tree-level processes (SM UT). Right: From loop processes. Inconsistency amongst them would reveal NP in the loops.

The dedication to B physics determines the forward geometry of the detector (see Figure 3-2), to take advantage of the large crosssection at low angle. Apart from B physics, the LHCb program also includes charm physics[82], inclusive production[83] -were the unique pseudorapidity range of LHCb will allow precise discrimination of the different models- and quarckonia studies[84].

3.2.1. LHCb detector

3.2.1.1. Layout

Due to the high b quark production crosssection at low angle, the LHCb detector[79] is a spectrometer covering a region of 300 mrad (in XZ plane) and 250 mrad (in YZ plane) from the beam line in the positive z direction. As p-p collisions are in CM and then symmetric, the same physics should happen in the negative z region. That geometry also allows the installation of Ring Imaging Cherenkov (RICH) detectors, used for K/ π separation which is needed in several decay channels studied in the experiment.

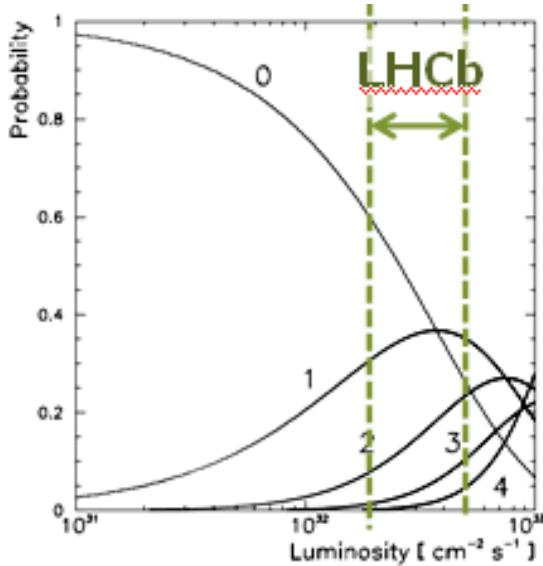


Figure 3-4: Probability of a given number of interactions as a function of luminosity.

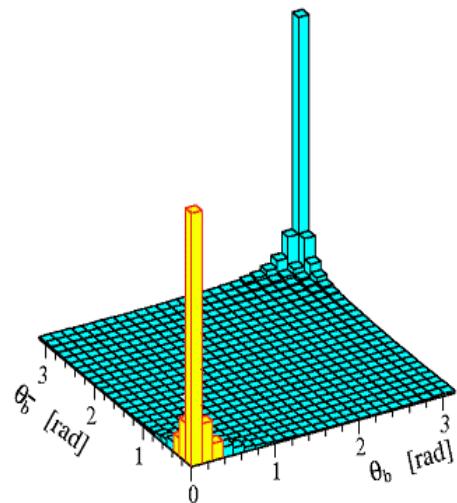


Figure 3-5: bottom production angle.

The separation between the Primary Vertex (PV), where the b 's were produced, and the Secondary Vertex (SV), where the b -hadron have decayed, is essential throughout all the LHCb analysis. With the nominal luminosity of the LHC, a mean of 23 PV's are produced for each bunch crossing; hence at that luminosity the PV-SV separation, as well as the tagging[85] of neutral B mesons, becomes very difficult. Thus, LHCb experiment is planned to run with lower luminosity than ATLAS and CMS, being tuned to get a mean value of ~ 1 PV per bunch crossing. It happens for instant luminosities in the order of $2 - 5 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$. This luminosity is tunned by defocusing proton bunches in the proximities of the interaction point. Also low luminosity helps to reduce the higher radiation damage at low angle.

The mechanisms related with b quark production in LHC p-p collisions are described in[86]. The high probability of proton crossing partons to have very different momenta causes the $b - \bar{b}$ pair to be produced with a large boost, having then high momentum and low angle. The production angle of both quark and antiquark are therefore highly correlated (see Figure 3-5).

LHCb[79] has the following subdetectors:

- **VELO (VErtex LOcator):** Silicon detector around the interaction point, used for the identification of primary and secondary vertices.
- **RICH (Ring Imaging Cherenkov Detector):** Allows the determination of the velocity of charged particles trough Cherenkov effect. Once the momentum is known, it can be used to compute the mass of the particle improving the particle identification performance, specially the separation K- π . LHCb detector has two RICH: RICH-1 before the magnet and RICH-2 after the magnet.
- **TT (Trigger Tracker) and Tracking Stations.** Detectors which together with the VELO, allow to reconstruct the trajectory of charged particles. Because of the magnetic field introduced by the magnet, measured tracks provide trimomentum information.
- **Calorimeter system:** SPD(*Scintillator Pad Detector*), PS(*PreShower*), ECAL (*Electromagnetic CALorimeter*) and HCAL (*Hadronic CALorimeter*). The calorimeter system stops all kind of long lived particles (apart from muons and neutrinos) and measures the energy deposited there.
- **Muon system:** The muon stations allow to identify muons that passed calorimeters, using MWPC's (and triple-GEMs for high occupancy regions).

3.2.1.2. Beam pipe, vacuum chamber and BCM

The beam pipe has a length of about 19 m, passing through all the subdetectors but the VELO. 12 out of the 19 m are made of beryllium (see Figure 3-6) because of its high transparency to particles created in the collision. However, beryllium has the disadvantages of fragility, high costs and toxicity, thus the region far away from the interaction point is made of stainless steel. VELO exit window, as well as bellows and

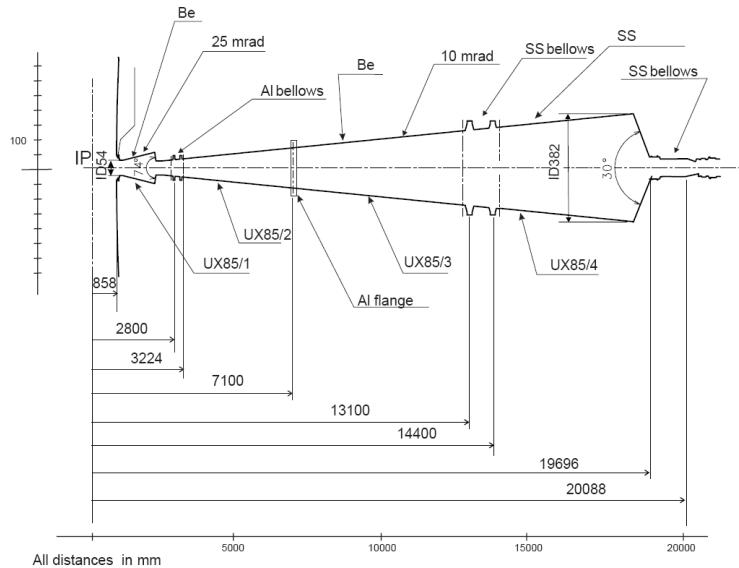


Figure 3-6: Beam pipe layout.

flanges in regions closer to interaction point are made of high strength aluminum.

The Beam Conditions Monitor (BCM) is a safety system consisting in chemical-vapor deposition diamond sensors monitoring the particle flux and located at 2.1 (upstream) and 2.8 (downstream) meters from the interaction point, close to the vacuum chamber. It allows detecting unstable beam conditions and requesting a beam dump if needed.

3.2.1.3. Magnet

In order to measure the momentum of the charged particles produced in LHCb, the experiment uses a dipole magnet (see Figure 3-7) with Al-99.7 coils and Fe yoke that provides an integrated magnetic field of 4Tm, which allows momentum resolutions better than 0.5 % (see section 3.2.2) for particles of about 200 GeV. The angular acceptance is ± 300 mrad in XZ (bending) plane and ± 250 mrad in YZ plane.

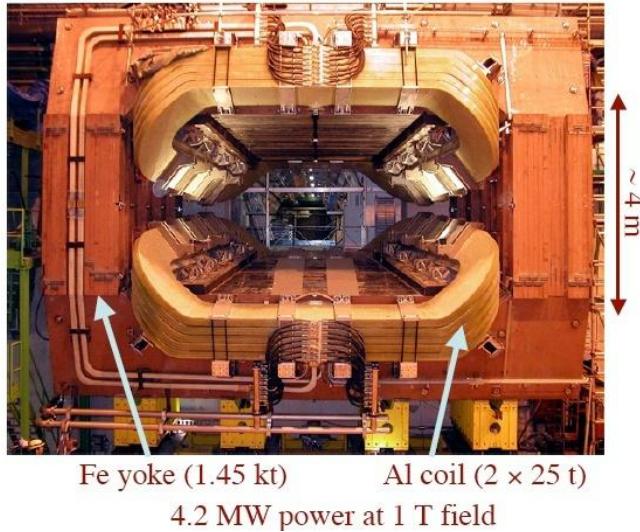


Figure 3-7: LHCb magnet.

The magnetic field along the Z axis is shown on Figure 3-8. In RICH-1 region, photon detectors are protected from magnetic field by iron shielding boxes.

The magnet can also be set to invert its polarity, which should reduce systematic errors that might affect to CP violation studies.

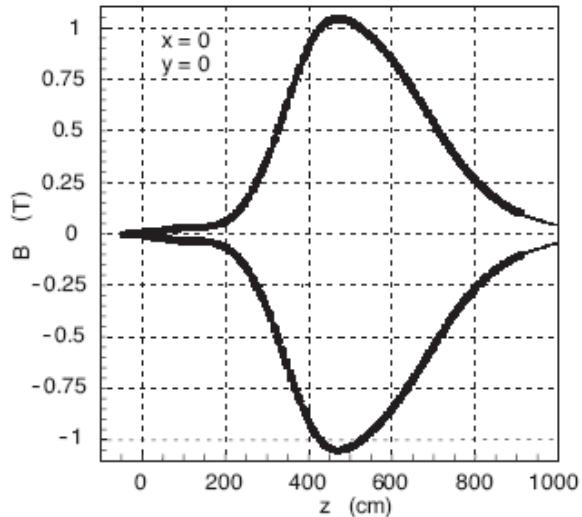


Figure 3-8: Magnetic field along z axis.

3.2.1.4. VELO

The VELO is used to get precise reconstruction of the tracks close to the interaction point. This is needed in order to separate the decay vertex of the b hadron from the primary one. It is made of semicircular silicon modules each one providing measurement of cylindrical coordinates r and φ using microstrips, and distributed as in Figure 3-9. The four VELO sensors situated before the interaction region form the pileup system, which can be used at the L0 trigger to eliminate high multiplicity events. The minimum pitch (at innermost radius) is $38\mu\text{m}$, and increases linearly up to $101.6\mu\text{m}$.

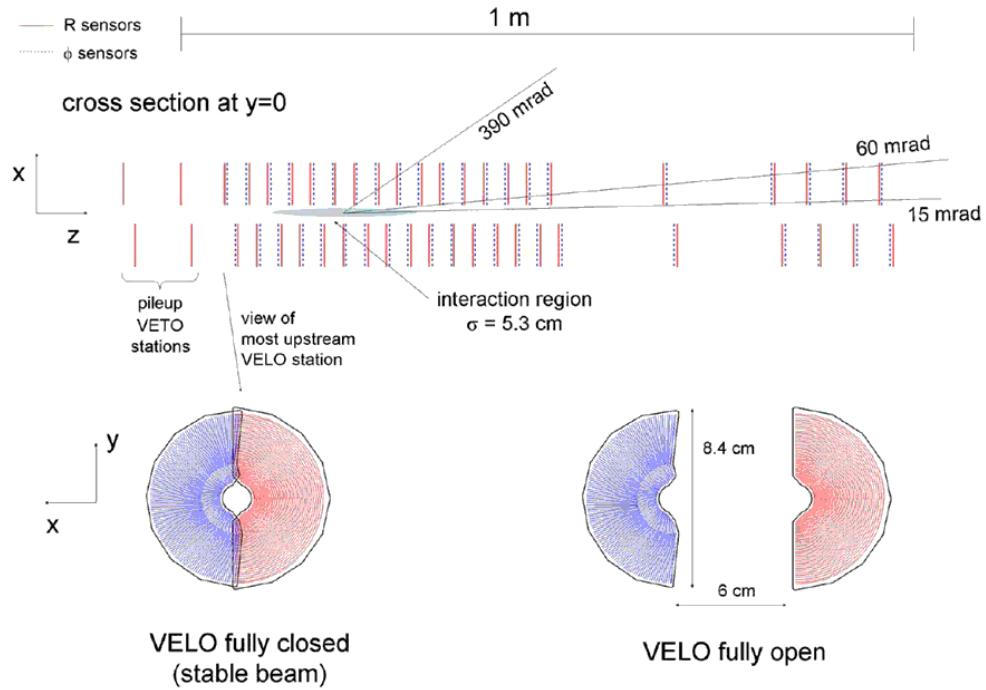


Figure 3-9: VELO layout.

The sensors are situated in a vessel that keeps them in a vacuum separated from the machine vacuum by a thin walled corrugated aluminum sheet, and are retractable thus that can be separated from the beam line during injection (or unstable beam conditions).

3.2.1.5. TT and Tracking Stations

The TT (Trigger Tracker or Tracker Turicensis) and Tracking Stations (hereafter Tx) constitute, together with the VELO, the Tracking System. The inner part –where the occupancy is higher- of Tx (Inner Tracker, IT), as well as the TT use silicon microstrip sensors and for that reason the system TT – IT is also called Silicon Tracker (ST). The outer part of Tx is a drift-time detector called Outer Tracker (OT).

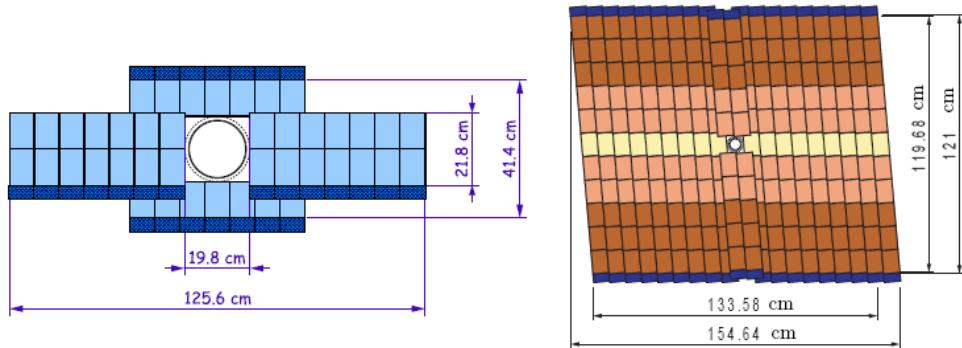


Figure 3-10: IT (left) and TT (right) modules.

The TT and each one of the three IT stations has four detection layers, with vertical strips for the first and last layer, and rotated at -5° ($+5^\circ$) in the second (third) layer, in order to get 3D reconstruction. The layout for IT and TT stations is shown in Figure 3-10. The pitch is about $200\mu\text{m}$ which gives a single hit resolution of $50\mu\text{m}$. Momentum resolution is then dominated by multiple scattering.

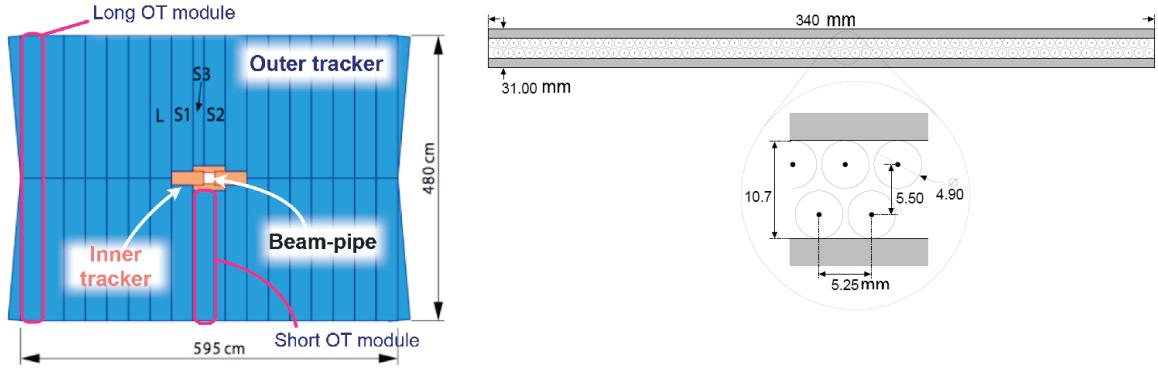


Figure 3-11: OT layout and cross section of one OT module.

The OT uses gas-tight straw-tube modules. As in the case of the ST, each station has four layers: first and last are vertical oriented and second and third have a $\pm 5^\circ$ rotation. The inner diameter of the drift tubes is 4.9 mm, the drift time less than 50 ns and the drift coordinate resolution about 200 μm .

3.2.1.6. RICH detectors

LHCb has two RICH detectors (see Figure 3-12) in order to cover full momentum range. While the RICH-1 (upstream) covers the momentum region from 1 to 60 GeV, using aerogel and C_4F_{10} as radiators, RICH-2 covers momenta larger than 15 GeV using CF_4 .

RICH-1 acceptance goes from 25 mrad to 250(300) mrad in the vertical (horizontal) direction. RICH-2 has a lower angle acceptance, from 15 mrad to 100(120) mrad in the vertical (horizontal) direction, as high momentum particles are produced at low angle.

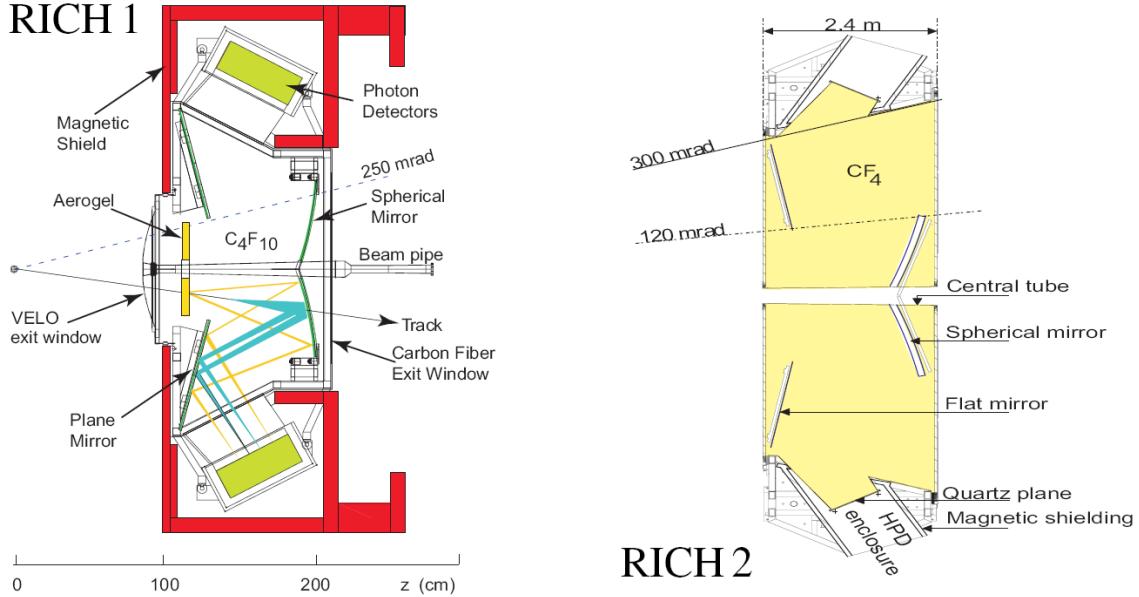


Figure 3-12: RICH detectors.

The rings produced by the particles passing through the radiators are guided to the Hybrid Photon Detectors (HPD, which are outside the acceptance) using a mirror system. HPDs are protected from the dipole field by magnetic shielding.

3.2.1.7. Calorimeter System

Calorimeter system is used in L0 trigger to identify high transverse energy hadrons, photons or electrons. It is also part of the particle identification system, allowing mainly the separation of hadrons, electrons and photons, as well as avoiding the pass of those particles to the muon system. LHCb uses an electromagnetic

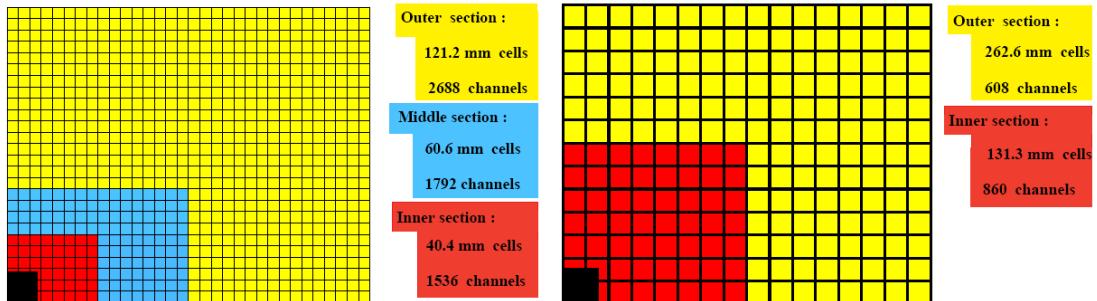


Figure 3-13: Lateral segmentation of calorimeters.

Left: ECAL, PS and SPD. Right: HCAL. Cell size is larger in HCAL because of the dimensions of hadronic showers.

calorimeter (ECAL) followed by a hadronic calorimeter (HCAL) to identify electromagnetic and hadronic showers, respectively.

In order to get a longitudinal segmentation of the electromagnetic shower, a preshower (PS) detector is installed before the ECAL. It allows to eliminate a large background of charged pions that otherwise might be identified as electrons. Rejecting neutral pions ($\pi^0 \rightarrow \gamma\gamma$) is achieved by inserting a scintillator pad detector (SPD) before the PS, such that charged and neutral particles can be separated. A thin lead converter is installed between PS and SPD.

3.2.1.8. Muon Stations

The Muon System allows identifying the muons that pass the calorimeter system. It provides also an estimation of the momentum of those muons, playing an important role in the L0 trigger. Muon System has 5 stations M1 – M5, from which M1 is located before the calorimeter in order to improve p_t measurement in the trigger (see Figure 3-14). M1-3 have good x resolution in order to provide track direction and a p_t measurement with ~20 % precision, while M4 and M5 are mainly designed to identify the most penetrating particles. Three iron shields are interleaved between the muon stations acting as muon filters. A fourth one is situated after M5 to remove machine related background.

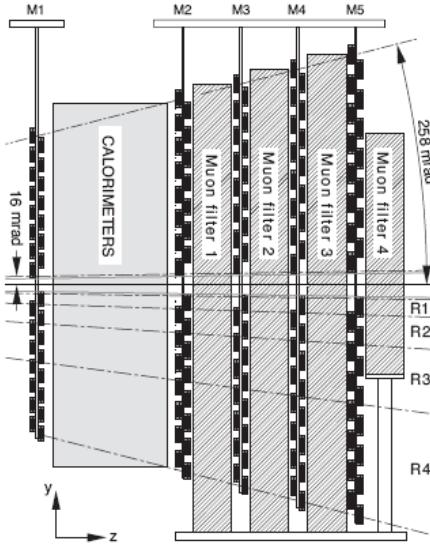


Figure 3-14: Muon stations.

The regions R1 – R4 in Figure 3-14 are defined such that the occupancy on each one is roughly the same within a given station. The spatial resolution worsens far from the beam axis, but anyway there it is limited by the higher multiple scattering at large angles.

3.2.2. Tracking and Vertexing Performance

Tracking software uses hits in VELO, TT and Tx to reconstruct the tracks of the charged particles. Depending on the subdetectors used for the reconstruction, offline tracks are classified in the following categories (see also Figure 3-15):

- Long tracks: The track trajectory goes from VELO to Tracking station, being then the largest possible and providing therefore the most precise momentum measurement.
- Upstream tracks: use VELO and TT. Upstream tracks are mainly due to low momentum particles that were bent out of the acceptance by the magnetic field.
- Downstream tracks: use TT and Tx. Decay products from K_s^0 and Λ are a typical source of downstream tracks.
- VELO tracks: only hits in the VELO. Because of the VELO has several modules upstream and downstream of the interaction point, it can reconstruct tracks from backward and large angle particles, used for primary vertex reconstruction.
- T tracks: only hits in the Tx, typically produced by secondary interactions.

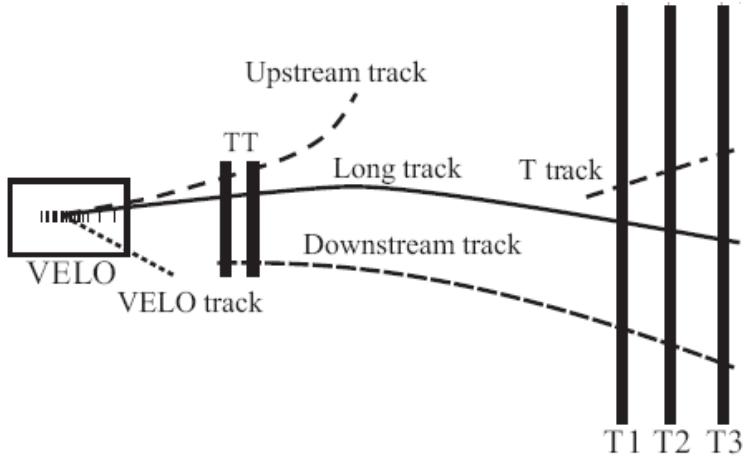


Figure 3-15: Offline track classification.

Track reconstruction starts by identification of track candidates, also called seeds, in VELO and Tx, as the magnetic field in those subdetectors is lower. The next step consists in trying to associate those seeds to hits in the other tracking subsystems to form tracks. First, taking a VELO seed and a hit in a T station, hits in the other T stations in a search window opened around the track candidate trajectory are searched for. If hits are found to confirm the track candidate, it becomes a long track. Hits in the TT corresponding to the track are added to it. Finally, the hits associated with the tracks are removed from the list of hits on which the algorithm has to run. This algorithm, called the forward tracking, reconstructs about 90% of the reconstructible (i.e., with enough hits in each subdetector to be reconstructed) long tracks. Another 4 % is recovered by the track matching algorithm, which starts by matching pairs of VELO and T seeds. There exist other algorithms to reconstruct the other types of track listed above. Once the track is found, it is refitted using Kalman filter, which accounts for multiple scattering. In the track reconstruction process, it is possible for a physical track to be reconstructed by more than one algorithm, resulting in two clone tracks. In that case, only the best out of the two tracks is kept.

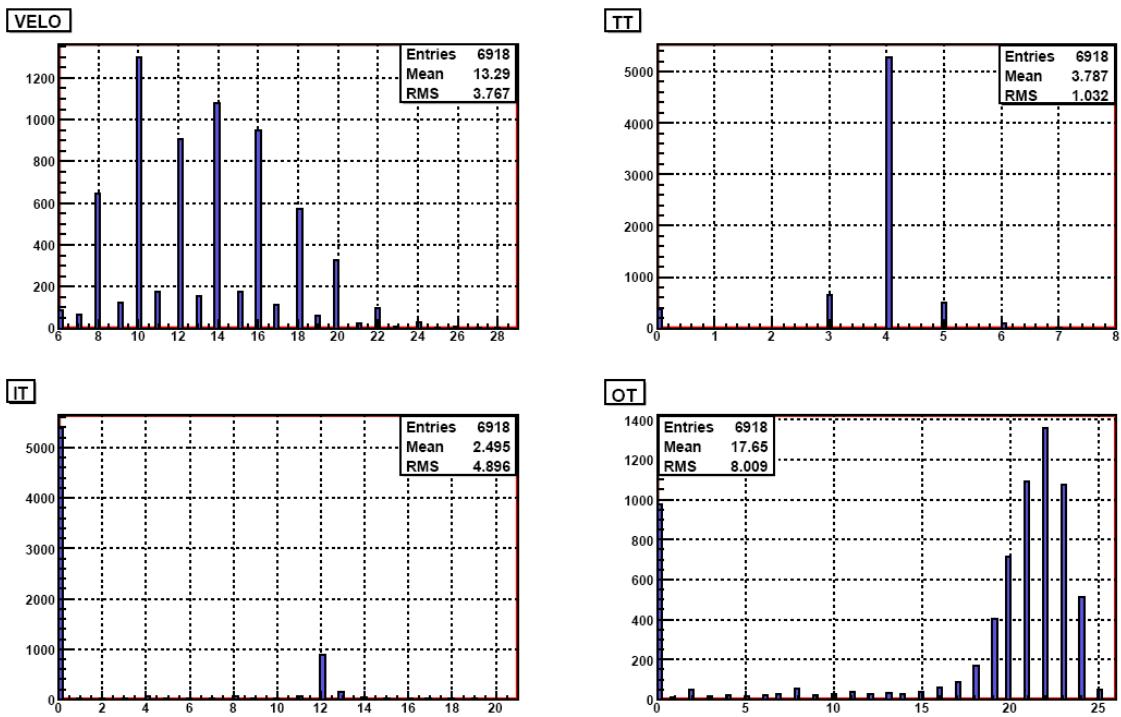


Figure 3-16: Number of hits per subdetector for $B_s \rightarrow \mu\mu$ long tracks.

The number of hits in each tracking subdetector for the long tracks produced by (positive charged) muons from $B_s \rightarrow \mu\mu$ is shown in Figure 3-16. Signal tracks tend to have an even number of hits in the VELO due to the proximity between consecutive r and φ sensors (see Figure 3-9). In average, the couple of tracks making the candidate have ~ 74 hits.

For tracks coming from $B_s \rightarrow \mu\mu$ the momentum and IP resolution are $\sim 0.5\%$ and $\sim 23\mu\text{m}$, having some dependence on the momentum of the associated particle. For low momenta the IP resolution is worse, but the momentum resolution itself becomes better. Figure 3-17 shows the dependence of p and IP resolution with the momentum, as well as the χ^2 of the track fit versus the number of degrees of freedom.

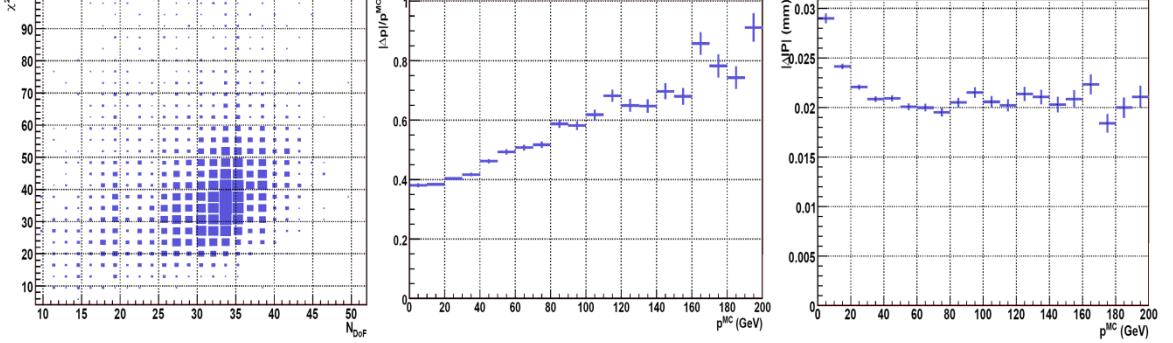


Figure 3-17: Tracking resolutions.

Left: χ^2 vs number of degrees of freedom. Center: momentum resolution, in percentage, as a function of momentum. Right: Impact parameter resolution as a function of momentum.

VELO tracks are used for reconstruction of the primary vertices. The fitting procedure for primary vertices is explained in [87]. Figure 3-18 shows the differences, in the three coordinates, between the MC Truth origin vertex of the B_s and the reconstructed and associated one, chosen as the PV that minimizes the impact parameter significance of the reconstructed B_s . The difference to zero in the mean of z resolution comes from two different contributions: tracks coming from SV's close to the PV, and bias in VELO RZ cluster position, present in DC06 simulation.

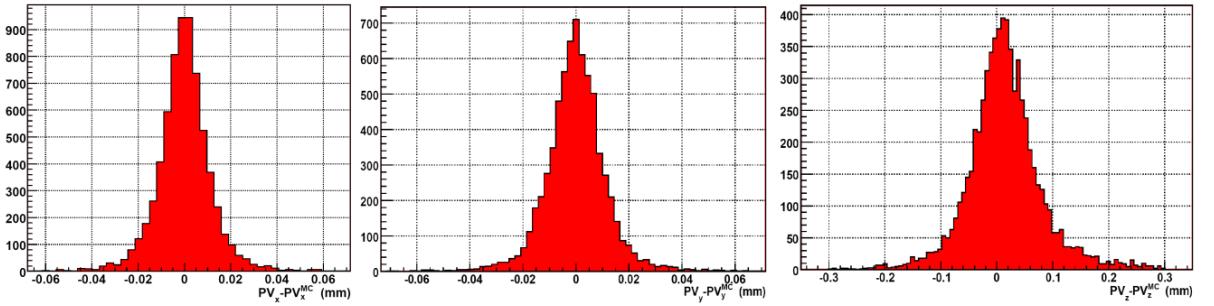


Figure 3-18: Primary vertex resolutions.

The $B_s \rightarrow \mu^+\mu^-$ decay has only two charged tracks so then the resolution of the secondary vertex is worse if compared with PV resolution, as can be seen in Figure 3-19. Note that, as the average flight is ~ 10 mm, the resolutions shown will provide a very good PV – SV separation.

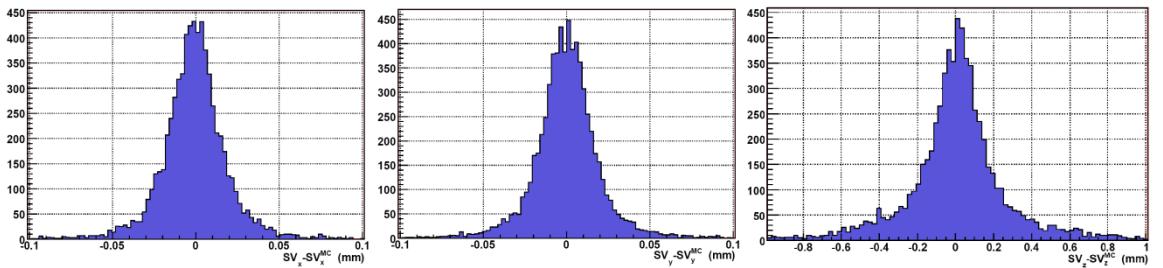


Figure 3-19: $B_s \rightarrow \mu^+ \mu^-$ decay vertex resolution.

A good invariant mass resolution is crucial in the analysis of $B_s \rightarrow \mu^+ \mu^-$, as it will be one of the clearest ways to distinguish signal from background. The resolution in LHCb for tracks coming from $B_s \rightarrow \mu^+ \mu^-$ is shown in Figure 3-20, where it has been fitted to two Gaussians: a 74 % component with 19 MeV resolution and 26 % with 41 MeV resolution, giving an averaged resolution of ~ 25 MeV. Also, the mean of the distribution is significantly biased to ~ -1 MeV, which implies that measured masses are slightly minor than true ones. This is because of an imperfect tuning of the energy loss[88].

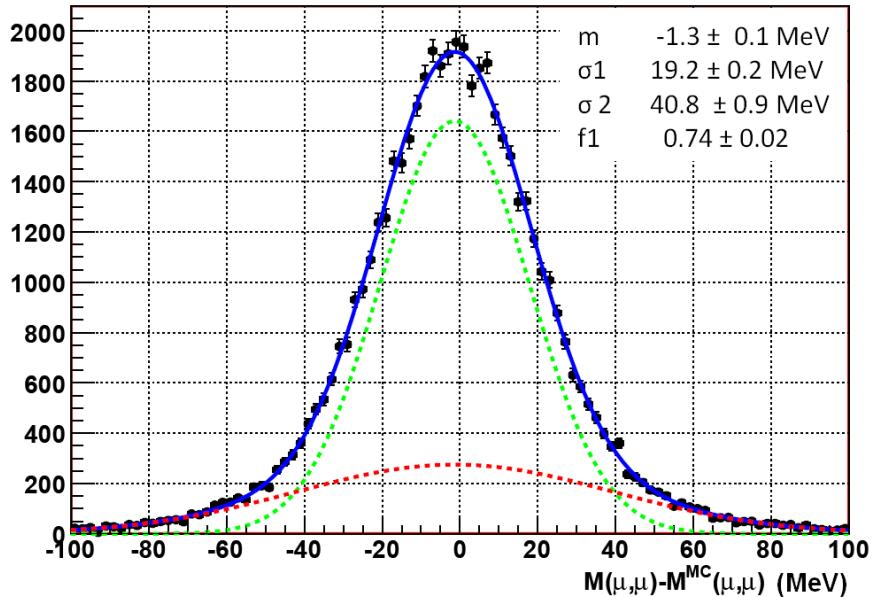


Figure 3-20: Invariant mass resolution.

Table 3-I: IP, mass and momentum resolutions.

Parameter	Resolution
Muon momentum	$\sim 0.5\%$
Muon IP	$\sim 23\mu\text{m}$
Dimuon invariant masss	74% : 19 MeV, 26 % : 41 MeV

3.2.3. Particle Identification Performance

3.2.3.1. Hadron ID

Separation of different hadron types is possible due to the RICH system. It uses the trajectory of a given particle to find where photons should be found on photodetector plane for a given mass hypothesis. The prediction is compared to the observed distribution of detected photoelectrons, giving the maximum likelihood and the difference on likelihood for the different mass hypothesis. It is expressed usually as a DLL (Delta Log Likelihood): $DLL(hyp.1 - hyp.2) = \text{Log}[\text{Prob}(hyp.1)/\text{Prob}(hyp.2)]$. Figure 3-21 shows the separation power of the RICH for $k-\pi$ (left) and $\pi-\mu$ (right). Of course, distinguishing between muons and pions using RICH system is difficult as the masses are very similar[7].

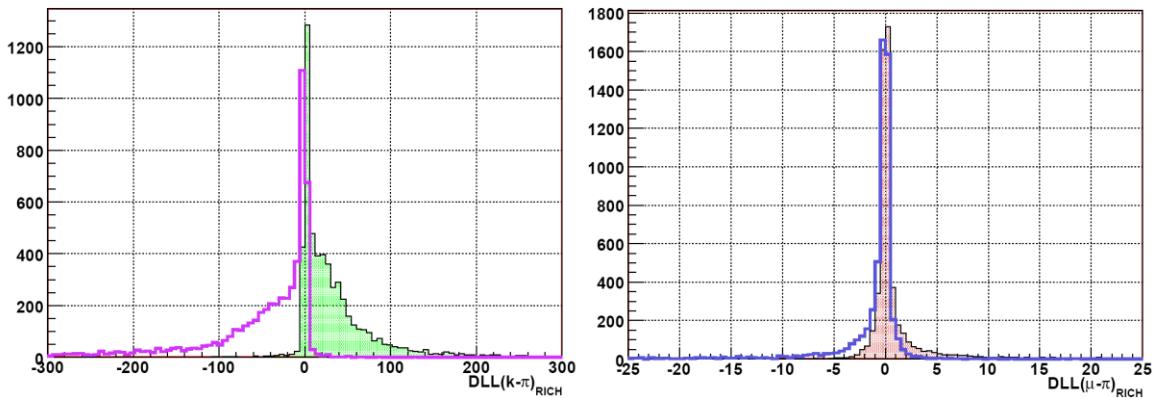


Figure 3-21: DLL obtained from RICH system.

Left: $DLL(k-\pi)$. Right: $DLL(\mu-\pi)$. Open histograms: pions $B_d \rightarrow \pi\pi$. Filled: kaons from $B_s \rightarrow kk$ (left) and muons from $B_s \rightarrow \mu\mu$ (right)

The calorimeter system provides $DLL(e-\pi)$ and $DLL(\mu-\pi)$, helping in the separation of charged particles between electrons, hadrons, and muons. The energy deposit in the calorimeters for kaons, muons and pions is shown in Figure 3-22. There are an important fraction of particles that do not contribute to those histograms as their produced signal falls below the required threshold. As the ADC counts in the calorimeter are calibrated by transverse energy, $E_t = E \sin(\theta)$, for a given E the sensitivity drops at low angle. The fraction of particles which do not pass the threshold is shown in Table 3-II.

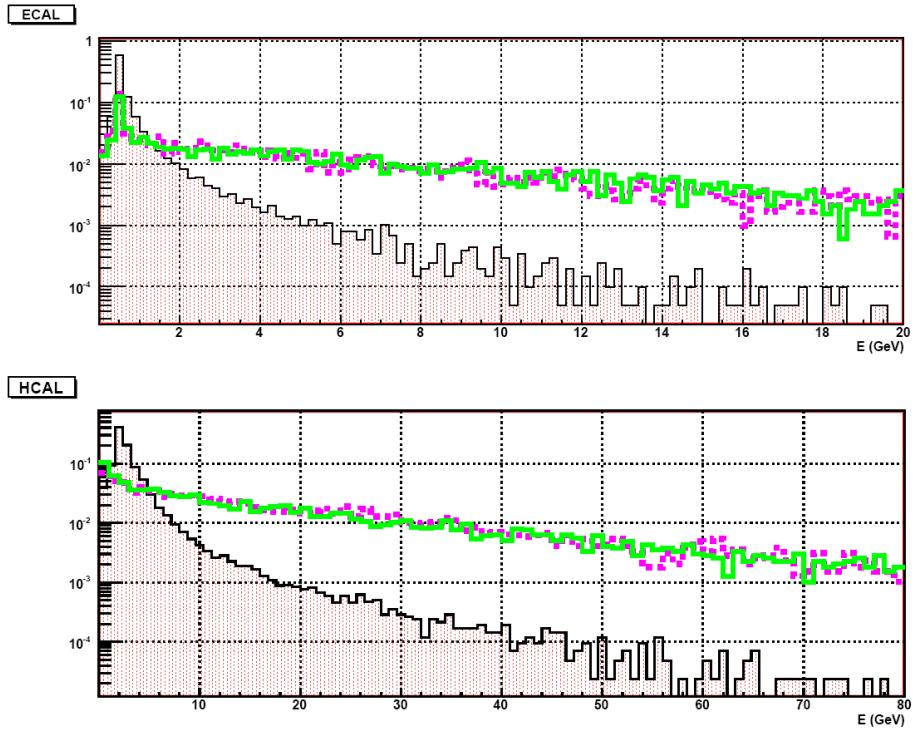


Figure 3-22: Energy deposited in calorimeters.

Filled: muons from $B_s \rightarrow \mu\mu$. Magenta dashed: kaons from $B_s \rightarrow KK$. Green Line: pions from $B_d \rightarrow \pi\pi$. The histograms are normalized to 1.

Figure 3-22 shows that the energy deposited in HCAL is in average ~ 4 times larger than in ECAL. The MIP signal peak in ECAL is about 400 MeV while in HCAL corresponds to approximately 2 GeV.

Table 3-II: Fraction of particles with calorimeter signal below the threshold.

Particles	Below thres. in ECAL	Below thres. in HCAL	Below thres. in both
K (from $B_s \rightarrow KK$)	(35±1) %	(23±1) %	(16±1) %
π (from $B_d \rightarrow \pi\pi$)	(31±1) %	(26±1) %	(16±1) %
μ (from $B_s \rightarrow \mu\mu$)	(62.1±0.2) %	(22.3±0.2) %	(19.4±0.2) %

3.2.3.2. Muon ID

Muons are identified by extrapolating long tracks into the muon stations. Depending on the momentum of the tracks, fields of interest (FOI) are opened in each muon station in order to find hits close to track extrapolation.

A track is flagged as *IsMuon* when satisfying:

- $3 \text{ GeV} < p < 6 \text{ GeV}$ and hits in M2 and M3
- $6 \text{ GeV} < p < 10 \text{ GeV}$ and hits in M2, M3 and M4 (or M5)
- $p > 10 \text{ GeV}$ and hits in M2-5.

The efficiency of *IsMuon* depends on the phase space of the particles and especially on the momentum (see Figure 3-23, left). It grows rapidly up to $p \sim 20$ GeV, but for higher momentum it starts to decrease as more muons go through the beam pipe (the momentum is higher for lower angles) so outside of the muon chambers acceptance. For tracks coming from $B_s \rightarrow \mu^+ \mu^-$ this efficiency is ~92 % (85 % to identify both), while for muons from $B^+ \rightarrow J/\psi(\mu^+ \mu^-) K^+$ it is a bit smaller (~90 %) as the momenta are lower.

The fraction of pions from $B_d \rightarrow \pi^+ \pi^-$ identified as muons is in the order of 1% (note that in order to missidentify a $B_d \rightarrow \pi^+ \pi^-$ as $B_s \rightarrow \mu^+ \mu^-$ both pions must be flagged as muons, which gives some few in 10000). The missid rate (fraction of non muons identified as muons) also depends on the momentum (see Figure 3-23, right) and is about a factor ~2 larger for kaons than for pions.

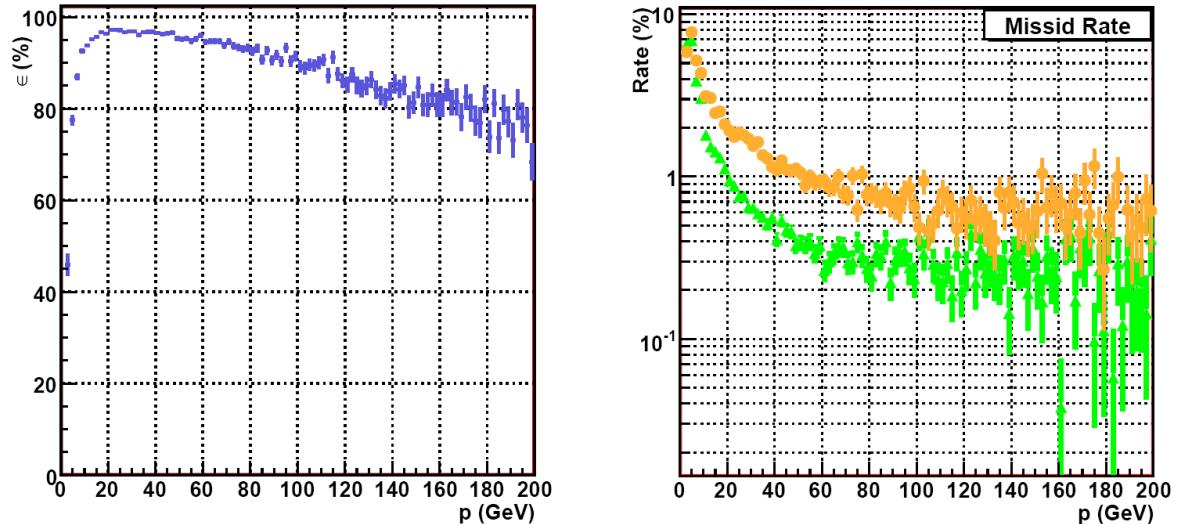


Figure 3-23: IsMuon efficiency and missid rate as a function of momentum.

Left: Efficiency for muons from $B_s \rightarrow \mu\mu$. Right: rate of hadrons misidentified as muons for kaons from $B_s \rightarrow kk$ (orange dots) and pions from $B_d \rightarrow \pi\pi$ (green triangles).

The distance of the hits to the track extrapolation is used to construct a likelihood ratio from the muon stations. Together with the energy associated in the calorimeters and the mass expected from RICH rings, an overall likelihood is then constructed as discriminating variable to separate muons from other particles.

Different DLLs are shown in Figure 3-24. The first two are the combined DLL($\mu-\pi$) and DLL($\mu-k$), while the other four are obtained from different subdetectors alone. Note that the particles used to fill those histograms were required to be identified as *IsMuon*. The best muon to hadron separation is provided by the muon system, although for the specific case of the kaons the RICH DLL has better rejection power.

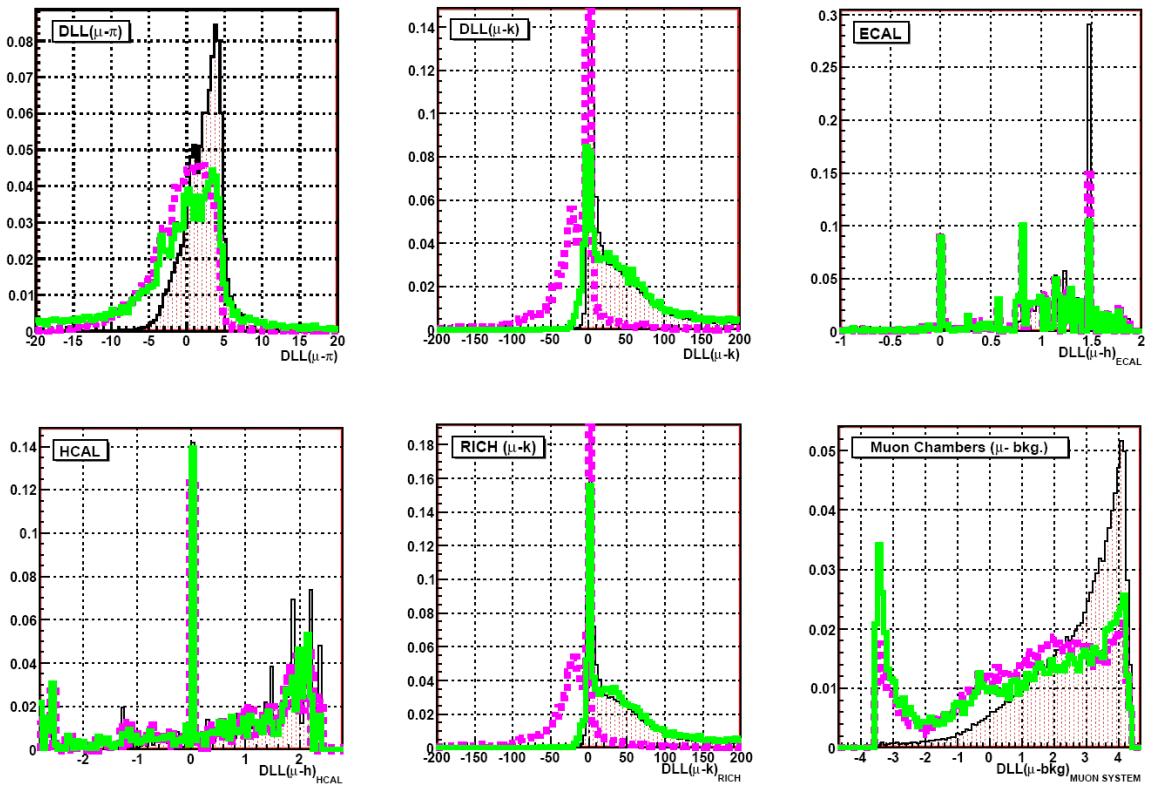


Figure 3-24: Muon identification in different subdetectors.

Top left: Combined DLL(μ - π). Top center: Combined DLL(μ -K). Top right: DLL(μ -h) from ECAL.

Bottom left: DLL(μ -h) from HCAL. Bottom center: DLL(μ -K) from RICH. Bottom right: DLL(μ -non μ) from Muon System.

Filled: muons from $B_s \rightarrow \mu\mu$. Magenta dashed: kaons from $B_s \rightarrow kk$. Green Line: pions from $B_d \rightarrow \pi\pi$. The histograms are normalized to 1. All particles were required to be identified as IsMuon.

3.2.4. Trigger System

At a luminosity of $2 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$, the rate of events with charged particles within the detector acceptance is expected to be ~ 10 MHz. LHCb trigger system must be capable to reduce the rate down to the 2 kHz allowed by the long-term data storage resources, while providing the best possible efficiency on interesting b decays. As the rate of b quark production will be as high as 100 kHz, the LHCb trigger aims for a subset of the b decay modes. The main signatures to allow identifying particles from b decays are relatively high p_T and vertex displacement (which leads to high IP of the b daughters).

3.2.4.1. Trigger architecture

The LHCb trigger is composed of two levels: Level-0 (L0) and High Level Trigger (HLT). L0 is implemented in custom electronics, and reduces the non-empty rate from 30 to 1 MHz at a fixed latency of 4 μ s. The HLT is a software trigger running in computing nodes forming the Event Filter Farm (EFF).

L0 uses information from the calorimeters and muon chambers to provide high E_t and p_t candidates, and VELO pile-up system provides a fast estimation of the number of proton-proton interactions that occurred in the bunch crossing. L0 positive decisions are sent back to the front-end electronics of all the sub-

detectors, which pick-up the pieces of the relevant events from buffers and send them through a read-out network to the EFF.

The HLT has hence access to information from the whole detector. However, a complete event reconstruction is not feasible at 1 MHz. The HLT reconstructs first regions of interest defined by the L0 candidates. Events where candidates do not satisfy the required conditions when tracking and vertexing information is added are promptly discarded.

3.2.4.2. L0 trigger

The L0 trigger decision is taken by the L0 Decision Unit, based on the following pieces of information provided by the trigger boards:

- Calorimeter clusters classified as electrons, photons, neutral pions or charged hadrons according to the energy deposition in the different calorimeter layers.
- Muon (μ^+) and dimuon ($\mu^-\mu^+$) candidates.
- The result of a fast proton-proton interaction vertex search based on two dedicated layers of the vertex detector (the Pileup System).
- The multiplicities in the Pileup System and in a scintillator layer in front of the calorimeter (SPD).

If the running conditions require it, vetoes on the number of interaction vertices and on multiplicities can be applied to any L0 line. This would allow removing events which are harder to process at the HLT. The Pileup system of the VELO will be used to reconstruct the longitudinal position of the interaction vertices and reject events with two or more such vertices. Together with the multiplicity measured in the Pileup detector and the Scintillator Pad Detector (SPD), and the total ET measured in the calorimeter are referred to as Global Event Cuts (GEC) in the following.

In the calorimeter, 2×2 -cell clusters are formed. For the particular case of hadronic clusters, the transverse energy deposited in the ECAL cells in front of the relevant HCAL cells is also added to compute the E_T of the candidate. The relative precision on E_T of the hadronic clusters is roughly 30%.

For building muon candidates, straight segments are searched for in the four muon stations downstream of the calorimeter, where the occupancy is relatively low. The search is performed within projective towers, under the assumption that the muon tracks roughly point to the interaction point. The segments found are confirmed by looking for a compatible hit in the muon chamber upstream of the calorimeter. The momentum is then estimated by using a look-up table built under the assumption that the muon tracks originate at the interaction point. The momentum resolution obtained is 20%. A positive decision is taken when a calorimeter or muon candidate above the p_t or E_t threshold is found. The threshold values for different types of candidates are shown in Table 3-III.

Table 3-III: L0 thresholds.

Electron	Photon	Hadron	Muon	DiMuon
2.8 GeV	2.8 GeV	3.84 GeV	1.36 GeV	$p_{t1} + p_{t2} > 1.48 \text{ GeV}$, $\min(p_{t1}, p_{t2}) > 80 \text{ MeV}$

3.2.4.3. HLT trigger

The HLT algorithm is divided in two sequential phases called HLT1 and HLT2. HLT1 applies a progressive, partial reconstruction seeded by the L0 candidates. Different reconstruction sequences (called alleys) with different algorithms and selection cuts are applied according to the L0 candidate type. In general, track segments pointing to candidates are first searched for at the main tracking stations. If they are

found, the corresponding L0 p_T cut is applied again, but with 3% resolution. Finally, matching segments are searched for at the VELO. In order to minimize the time consumption, a fast reconstruction is performed first by using the r -sensors that measure the distance to the beam axis. This allows reconstructing the so-called VELO two-dimensional (2D) tracks.

If any of such tracks is matched to a L0 candidate, primary vertices are then reconstructed using VELO 2D tracks. For the tracks matched to L0 candidates, the complete 3D reconstruction at the VELO is performed by adding information from the φ -sensors, which measure the azimuthal angle. If the 3D track is found, then an IP cut is applied (except in the case of some muon lines).

In some alleys, extra tracks forming a good vertex with the original candidate are required before accepting the event. The search of such tracks uses VELO 2D track segments as seeds. The main alleys for the analysis explained in this thesis are the muon alley[89], muon+track alley[90] and hadron alley [91].

3.2.4.4. HLT-1 Muon Alley

The muon reconstruction efficiency in L0 is low for low pt. In the HLT1 some of the muons not identified by L0 are recovered by reconstructing other muon tracks, called muon segments. Hence muon alley has four selecting lines:

- 2 L0 Single Muons
- DiMuon from 1 L0 Single muon + 1 muon segment
- DiMuon from the L0DiMuon
- Single Muon

The algorithm can be schematized as follows:

1. Single Muons:

a) Selection of L0 muons that fired L0.

b) T Confirmation and VELO Matching: The L0 candidate is confirmed with tracks in the T stations. The momentum is also confirmed. Then a similar procedure to the offline *IsMuon* is applied and a full track is reconstructed by matching with the VELO 2D and 3D tracks.

c) Single Muon Decision. The event is selected by HLT-1 if the single muon passes one of the following conditions:

- c.1) $IP > 0.07$ mm and $p_T > 1.8$ GeV
- c.2) $p_T > 6.2$ GeV

2. Dimuons From Two L0Single:

a) Proceed as in 1a and 1b for 2 L0 muons.

b) Combine two tracks with a maximum DOCA⁵ of 1.6 mm.

c) Dimuons from two L0Single Decision. The event is selected by HLT-1 if the Dimuon satisfies one of the following conditions:

- c.1) Invariant mass $(\mu, \mu) > 300$ MeV and $IP(\mu, \mu, PV) < 0.08$ mm.
- c.2) Invariant mass $(\mu, \mu) > 2.5$ GeV and $p_{T,\mu 1} + p_{T,\mu 2} > 1480$ MeV.

⁵ Distance Of Closest Approach. It is computed by taking the first states of the two tracks, and measuring the distance between the tangents.

3. Dimuons From Muon Segment:

- a) Proceed as in 1a and 1b for one L0 muons.
- b) Reconstruct muon segments (M5-M2) and proceed as in 1b using the muon segment instead of the L0 muon candidate.
- c) Prepare pairs with one confirmed L0 muon and one confirmed muon segment requiring a maximum distance of closest approach.
- d) Dimuons from Muon Segment Decision. The cuts applied are the same as in 2c.

4. Dimuons From L0DiMuons:

- a) L0Dimuon Trigger with two L0muons.
- b) Proceed as in 1b for each L0 muon.
- c) Recombine the tracks in pairs with a maximum distance of closest approach. Check if the pair is from an L0 DiMuon pair.
- d) Dimuons from L0DiMuons Decision. Same cuts as in 2c.

3.2.4.5. HLT-1 Muon + Track Alley

The muon+track alley is executed only on those events triggered by a single-muon candidate at L0. This candidate is then confirmed as in the Muon Alley, and it is requested to satisfy $IP > 0.07$ mm and $p_T > 950$ MeV.

Then VELO 3D tracks within a DOCA to it less than 0.2 mm are searched for. The z distance between the reconstructed SV and the PV is requested to be larger than 2 mm. If a VELO 3D track satisfying those conditions is found, the corresponding segments in the Tx are added in order to provide momentum measurement. Then the following cuts are applied:

- $p_T > 550$ MeV, $IP > 0.1$ mm, $\chi^2/\text{ndf} < 10$.
- $\frac{|\vec{p}_\mu + \vec{p}_{tr}| \sin \theta}{|\vec{p}_\mu + \vec{p}_{tr}| \sin \theta + p_{T,\mu} + p_{T,tr}}$ < 0.4, where \vec{p}_μ , \vec{p}_{tr} are the momenta of the muon and the track, $p_{T,\mu}$, $p_{T,tr}$ the corresponding transverse momenta and θ is the angle between $\vec{p}_\mu + \vec{p}_{tr}$, and the flight direction $\overrightarrow{SV} - \overrightarrow{PV}$. Note that in case of being the muon and the track the only daughters of the B and if the momentum reconstruction were perfect then the value of the variable would be 0.
- Invariant mass (μ , track) > 1 GeV.

3.2.4.6. HLT-1 Hadron Alley

In the first step (L0confirmation), a track is reconstructed starting from a hadron cluster that fired the L0 hadron trigger. The alley is divided into two independent sequences of algorithms (lines), called Single Hadron and DiHadron. The first one sets a trigger decision based only on the L0 confirmed track, while in the second one, the decision is taken on a secondary vertex made with the L0 confirmed track and an extra track (companion track). All the tracks used are required to have $\chi^2/\text{ndf} < 10$.

1. In the Single Hadron, the track is requested to have $p_T > 4.45$ GeV, $IP > 0.09$ mm.
2. In the DiHadron, the following cuts are applied:
 - L0 confirmed track: $p_T > 1.6$ GeV, $IP > 0.09$ mm.
 - Companion track: $p_T > 950$ MeV, $IP > 0.13$ mm.
 - $SV_z - PV_z > 0.05$ mm.
 - DOCA < 200 μm .

3.2.4.7. HLT-2

HLT-1 selects about 30 kHz of events, on which the HLT-2 algorithm is then executed. The whole event reconstruction – although faster than the offline and hence reaching slightly worse quality- is performed and the so-called HLT-2 selections are then applied, aiming for any of the following signatures:

- Displaced vertices formed by two, three or four tracks. These selections are called topological lines.
- Single muons with high p_T and IP.
- Pairs of leptons (muons or electrons) with moderate p_T and IP and forming a good vertex.
- Leptons (electrons or muons) forming a secondary vertex with an additional track.
- A photon cluster in the calorimeter and a pair of tracks forming a secondary vertex, to select radiative B decays ($B \rightarrow X\gamma$).
- Inclusive $B \rightarrow DX$ and $\phi \rightarrow KK$.
- Several selections looking for exclusive decays.

3.2.5. Stripping selections

The triggered events are then processed offline, with enough time to run the full reconstruction, and several selections are ran in order to reduce the amount of data to be analyzed. About 10 Hz are given for the analysis of $B_s \rightarrow \mu\mu$.

In the designing of the selections for such analysis, the main constraints are:

- The rate of the OR of all the selections used, including those for control channels must be bellow (or in the order of) 10 Hz.
- Common cuts of $B \rightarrow hh$ and $B_s \rightarrow \mu\mu$.
- Even if topologically different, the $B^+ \rightarrow J/\psi(\mu\mu) K^+$ needs to be selected in a way as similar as possible to the signal, in order to avoid introducing systematic errors on the normalization (see 4.4.1).
- In a similar way, there should be an analogous selection for $B^+ \rightarrow J/\psi(\mu\mu) K^+$ and $B_d \rightarrow J/\psi(\mu\mu) K^{*0}$. It does not necessarily imply that $B^+ \rightarrow J/\psi(\mu\mu) K^+$ is selected in the same way as in the previous point.

The reason for these requirements will become clear in section 4.4.

The cuts used for $B \rightarrow hh$ in the analysis shown in this thesis are shown in Table 3-IV.

Table 3-IV: Stripping selection cuts for $B \rightarrow hh$.

DOFS stands for Distance Of Flight Significance, and corresponds to the distance between PV and SV, in sigmas. It is considered negative if $SV_z < PV_z$. IPS stands for Impact Paramter Significance, the impact parameter of a particle trajectory to the PV, in sigmas.

$ M - M_{Bd,PDG} < 500 \text{ MeV}$	DOFS(B) > 15
B decay vertex $\chi^2 < 9$	Minimum IPS (h^+, h^-) > 5



The quantity $|M - M_{B_d, \text{PDG}}|$ refers to the absolute difference of reconstructed B mass with respect to the nominal value. The above selection, together with the effect of L0&HLT-1&HLT-2 provides a reasonable rate of 3.2 ± 0.9 Hz. The same cuts are applied to $B_s \rightarrow \mu\mu$ with wider sidebands of 600 MeV, apart from the muon ID, giving a negligible rate which anyway will have a large overlap with $B \rightarrow hh$ stripping selection.

In order to clone that selection into one suitable for 3-body in the final state, the J/ψ is used for the cuts applied in two-track vertices. This criterion stands on the assumption that the momentum distributions for the three meson types $B^+/B_d/B_s$ are similar, so taking into account also the known likeness in the lifetimes, the distance of the two-track J/ψ vertex to the PV should be very similar to that of $B \rightarrow hh$ and $B_s \rightarrow \mu\mu$.

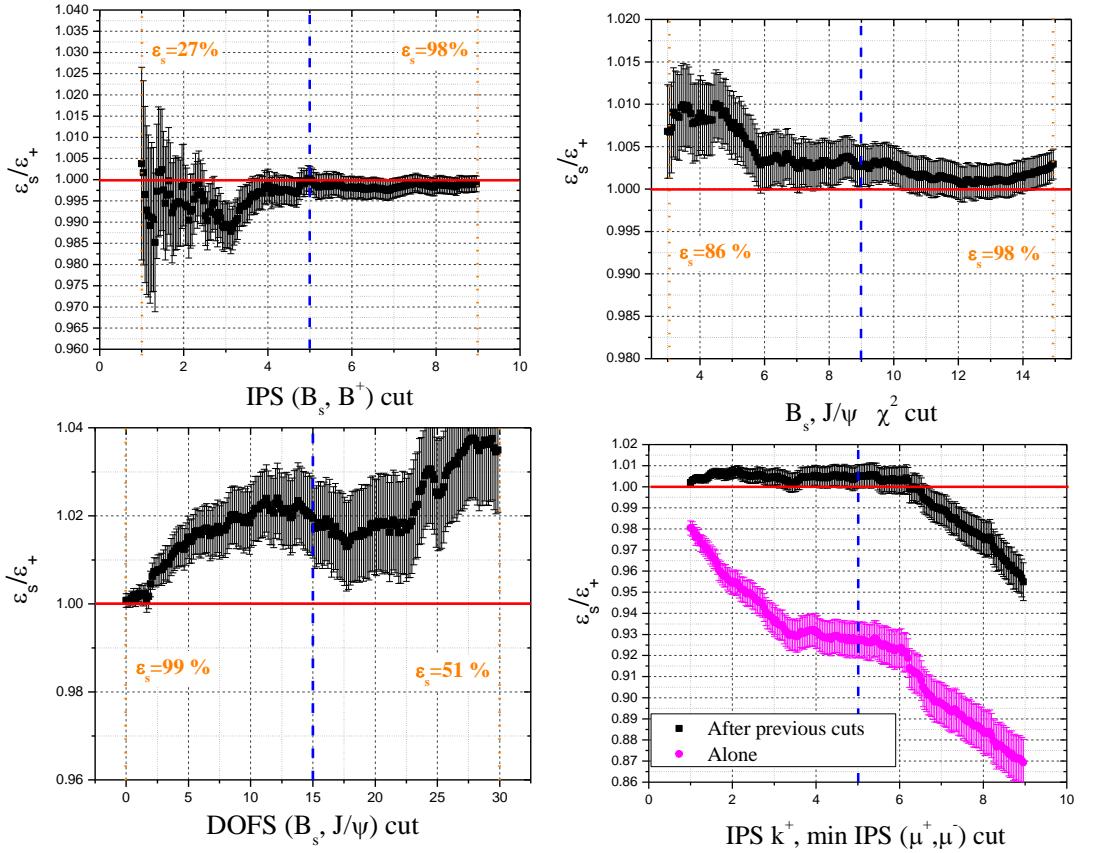


Figure 3-25: Stability of common selection cuts.

The y axis represents the ratio of efficiencies for $B_s \rightarrow \mu\mu$ and $B^+ \rightarrow J/\psi(\mu\mu) K^+$ of the cuts. Blue-dashed vertical lines show the actual value of the cut. Orange-dotted vertical lines show maximum and minimum efficiencies in the scanned range.

With those arguments, the cuts for $B \rightarrow hh$ are translated to $B^+ \rightarrow J/\psi(\mu\mu) K^+$ as the cuts shown in Table 3-V.

Table 3-V: Stripping selection cuts for $B^+ \rightarrow J/\psi(\mu\mu) K^+$.

DOFS stands for Distance Of Flight Significance, and corresponds to the distance between PV and SV, in sigmas. It is considered negative if $SV_z < PV_z$. IPS stands for Impact Paramter Significance, the impact parameter of a particle trajectory to the PV, in sigmas.

$ M - M_{B+,PDG} < 500 \text{ MeV}$	$\text{DOFS}(J/\psi) > 15$
$J/\psi(\text{ decay vertex } \chi^2) < 9$	$K^+ \text{IPS} > 5$
$B^+ \text{IPS} < 5$	$ M(\mu^+\mu^-) - M_{J/\Psi,PDG} < 60 \text{ MeV}$

Indeed, it can be seen in Figure 3-25 (plots 1-3) how the ratio of efficiencies for each cut keep close to one (within 2-4 %) for a wide range of values, even if each individual efficiency varies largely. However, a more delicate case is the IPS of the daughters (see Figure 3-25, plot 4). The IPS distribution of the kaon is not the same as the minimum IPS of h^+ , h^- (see ratio of efficiencies in magenta). Nevertheless, after the strong DOFS cut is applied (black), the efficiencies of both cuts keep being the same within ~2-4 % up to cut values of 9 sigmas.

Finally, in order to improve background rejection, only events with a reconstructed J/Ψ invariant mass in a window of 60 MeV around the nominal value are kept.

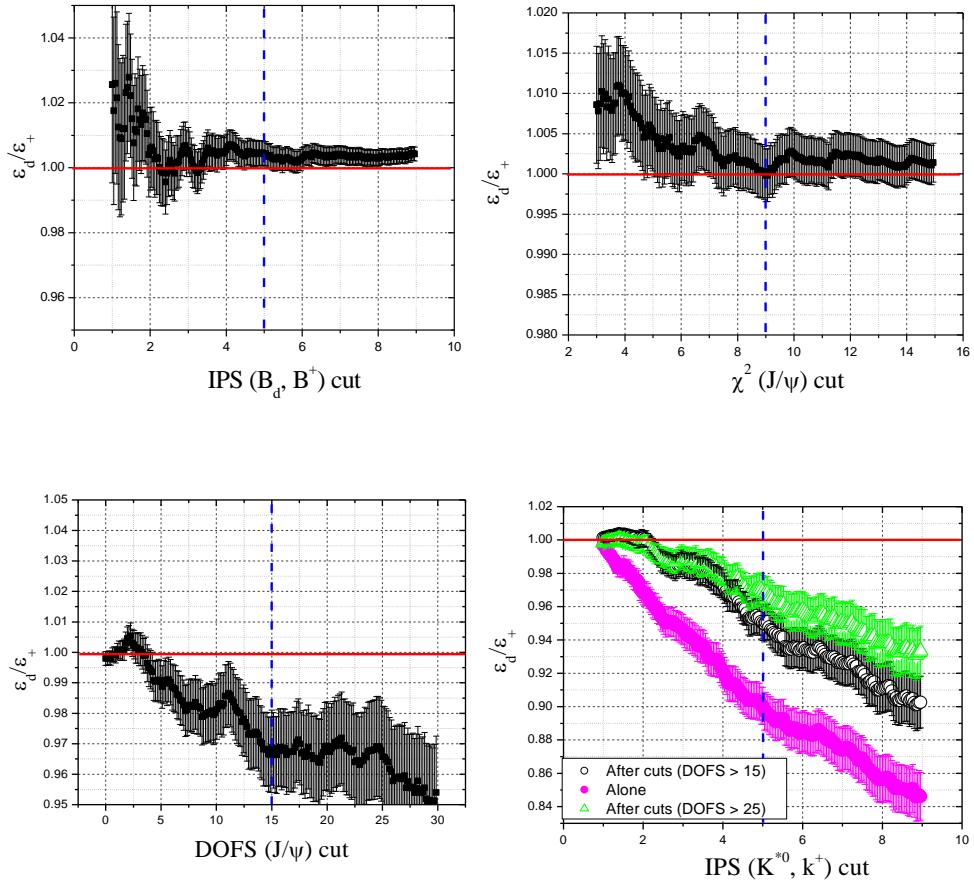


Figure 3-26: Common selection cuts for $B_d \rightarrow J/\psi(\mu\mu) K^{*0}(K^+\pi^-)$.

The y axis represents the ratio of efficiencies for $B_d \rightarrow J/\psi(\mu\mu) K^{*0}(K^+\pi^-)$ and $B^+ \rightarrow J/\psi(\mu\mu) K^+$ of the cuts. Blue-dashed vertical lines show the actual value of the cut.

Finally, the $B_d \rightarrow J/\psi(\mu\mu) K^{*0}(K^+\pi^-)$ is selected with the selection in Table 3-VI.

Table 3-VI: Stripping selection cuts for $B_d \rightarrow J/\psi(\mu\mu) K^{*0}(K^+\pi^-)$.

DOFS stands for Distance Of Flight Significance, and corresponds to the distance between PV and SV, in sigmas. It is considered negative if $SV_z < PV_z$. IPS stands for Impact Paramter Significance, the impact parameter of a particle trajectory to the PV, in sigmas.

$ M - M_{B_d, PDG} < 500 \text{ MeV}$	$\text{DOFS}(J/\psi) > 25$
$J/\psi \text{ (decay vertex } \chi^2) < 9$	$K, \pi \text{ IPS} > 2$
$B_d \text{ IPS} < 5$	$K^{*0} \text{ IPS} > 5$

$ M(K, \pi) - M_{K^*} < 40 \text{ MeV}$	$ M(\mu^+ \mu^-) - M_{J/\Psi, \text{PDG}} < 60 \text{ MeV}$
--	--

The cut on the distance of flight is tighter than in the previous case in order to force the daughters to have high IPS and therefore minimize systematic effects when cutting on those. The ratio of efficiencies with respect to $B^+ \rightarrow J/\psi(\mu\mu) K^+$ is shown in Figure 3-26.

The cut on the K^* mass window is motivated by the fact that it is a wide resonance with a Breit-Wigner width $\Gamma = 50.3 \pm 0.6 \text{ MeV}$, much larger than the expected resolution at that mass ($\sigma_{K^*} \sim 3-4 \text{ MeV}$, see Figure 3-27). Therefore, the line shape does not differ significantly from the physical distribution, and hence the efficiency of the cut $|M(K, \pi) - M_{K^*, \text{PDG}}| < 40 \text{ MeV}$ is insensitive to a good knowledge of the actual resolution. To illustrate this, Figure 3-28 shows an estimation of cut efficiency as a function of a hypothetical resolution. It can be seen that even for a factor 3-4 worse than nominal resolution, the efficiency only moves by a $\sim 3\%$. Thus, such cut (even if not optimal for background rejection) allows a rate reduction without the introduction of any MC dependence or significant systematic.

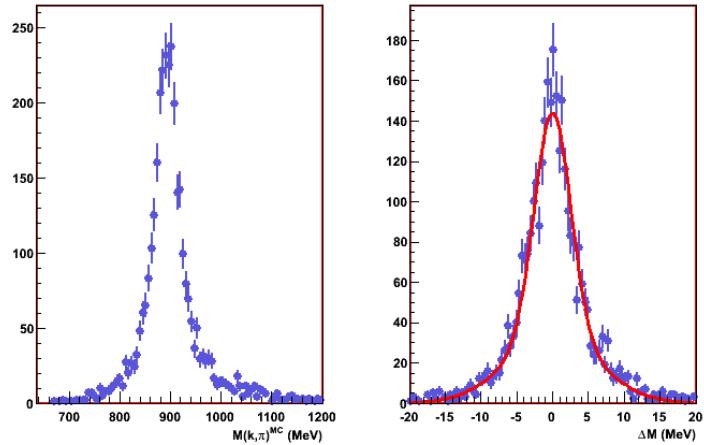


Figure 3-27: $K\pi$ (from K^{*0}) mass.

Left: MC Truth. Right: Measured mass minus MC Truth. The red line shows a fit of the resolution to double Gaussian, with a $(78 \pm 4)\%$ component of $\sigma_1 = 2.5 \pm 0.2 \text{ MeV}$ and a $(22 \pm 3)\%$ component of $\sigma_2 = 6.5 \pm 0.5 \text{ MeV}$.

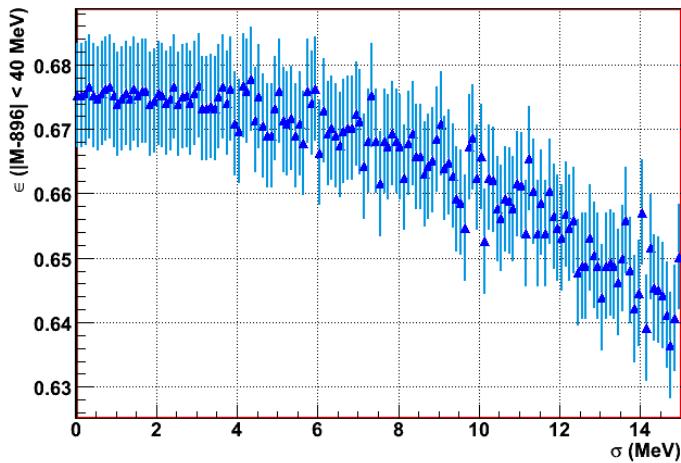


Figure 3-28: Variation of K^{*0} mass cut efficiency with the resolution.

It was estimated by single Gaussian degradation of MC truth spectrum.

3.2.6. Performance of selection and trigger algorithms

The efficiency of the stripping selections –and each of their cuts- on offline reconstructed signal candidates is shown in Table 3-VII. The highest efficiency cost comes from the DOFS cut, as this is the variable with higher rejection power.

Table 3-VII: Stripping selection efficiencies on offline reconstructed events.

	$B_s \rightarrow \mu\mu$	$B_d \rightarrow \pi\pi$	$B^+ \rightarrow J/\psi(\mu\mu) K^+$	$B_d \rightarrow J/\psi(\mu\mu) K^{*0}$
$ M - M_{B,\text{PDG}} < 600 (500) \text{ MeV}$	97.0 %	95.2 %	97.4 %	97.6 %
$B, J/\psi$ vertex $\chi^2 < 9$	98.2 %	97.6 %	97.1 %	97.0 %
B IPS < 5	97.8 %	97.3 %	97.4 %	97.4 %
$\text{DOFS}(B, J/\psi) > 15 (25)$	70.6 %	72.8 %	69.1 (54.7)%	(53.0 %)
<i>All the common</i>	65.8 %	65.8 %	63.7 (50.4) %	(48.9 %)
Min. IPS ($\mu/h^+, \mu/h^-$) > 5 , IPS $K^+, K^{*0} > 5$ (in parenthesis for DOFS > 25)	91.1 %	90.5 %	90.5 (93.2) %	(91.3 %)
$ M(\mu^+\mu^-) - M_{J/\psi,\text{PDG}} < 60 \text{ MeV}$	-	-	95.4 (95.6) %	(95.6 %)
$ M(K, \pi) - M_{K^*} < 40 \text{ MeV}$	-	-	-	(65.9 %)
K IPS > 2	-	-	-	(99.2 %)
π IPS > 1	-	-	-	(98.4%)
TOTAL	59.9 ± 0.6 %	59.5 ± 0.7 %	54.9 ± 0.6 % $(44.9 \pm 0.6$ %)	(27.5 ± 0.8) %

The trigger efficiencies on signal events selected by those stripping selections are shown in Table 3-VIII. As it can be seen, LHCb trigger system is highly (~90 %) efficient in the channels with muons in the final state, and reaches about 45 % efficiency for $B \rightarrow hh$.

Table 3-VIII: Trigger efficiencies.

The numbers in parenthesis correspond to efficiency on candidates selected by the DOFS > 25 cut.

	L0	HLT-1	HLT-2	TOTAL
$B_s \rightarrow \mu\mu$	98.7 %	98.1 %	98.6 %	95.4 ± 0.6 %

$B_d \rightarrow \pi\pi$	61.6 %	77.9 %	91.2 %	44 ± 1 %
$B^+ \rightarrow J/\psi(\mu\mu) K^+$	94.8 (94.9) %	94.9 (95.6) %	98.9 (99.2) %	89.1 ± 0.9 % $(90.0 \pm 0.9$ %)
$B_d \rightarrow J/\psi(\mu\mu) K^{*0}$	95.4 %	95.4 %	98.0 %	89 ± 1 %

This set of selections, together with the trigger system, gives the following output rates (measured on a 4 million events minimum bias sample, selected by L0 thus corresponding to 4 seconds of data taking):

Table 3-IX: Analysis output rate.

	L0&Stripping (Hz)	L0&HLT&Stripping (Hz)
$B_s \rightarrow \mu\mu$	< 0.6 @ 90% CL	< 0.6 @ 90% CL
$B \rightarrow hh$	~ 13	~ 2.8
$B^+ \rightarrow J/\psi(\mu\mu) K^+$	~ 5.3	~ 2.5
$B_d \rightarrow J/\psi(\mu\mu) K^{*0}$	~ 9.5	~ 3.0
All together	26 ± 3	7.8 ± 1.4

Hence the designed stripping cuts provide a rate lower than 10 Hz and thus fitting on the requirements.

4. Analysis of $B_s \rightarrow \mu\mu$ in LHCb

4.1. Simulation

4.1.1. Software Environment

All the Monte Carlo simulation samples used in this analysis were generated within the LHCb data challenge DC06⁶, using the event generator PYTHIA[92][93], the package GEANT [94] for simulation of particle interactions with material, and the detector geometry as described by Gauss (v25)[95]. The detector digitization is simulated using Boole (v12)[96] and the reconstruction is performed using Brunel (v30 & v31) [97].

The samples are obtained filtering a large data-set of minimum bias proton-proton interactions at $\sqrt{s} = 14$ TeV. The list of processes included in minimum bias definition is shown in Table 4-I.

Table 4-I: PYTHIA processes in DC06 minimum bias definition

Process number in PYTHIA	Description
11	$f + f' \rightarrow f + f'$ (QCD)
12	$f + f\bar{b} \rightarrow f + f\bar{b}'$
13	$f + f\bar{b} \rightarrow g + g$
28	$f + g \rightarrow f + g$
53	$g + g \rightarrow f + f\bar{b}$
68	$g + g \rightarrow g + g$
91	Elastic scattering
92	Single diffractive ($AB \rightarrow XB$)
93	Single diffractive ($AB \rightarrow AX$)
94	Double diffractive
95	Low-pT scattering
86	$g + g \rightarrow J/\Psi + g$
87	$g + g \rightarrow \chi_0 c + g$
88	$g + g \rightarrow \chi_1 c + g$
89	$g + g \rightarrow \chi_2 c + g$
106	$g + g \rightarrow J/\Psi + \gamma$

In order to optimize the production, when the event produces a b-meson it is required to have a forward direction within 400 mrad with respect to the beam axis. The efficiency of this cut for inclusive events with at least one of the two b-quarks satisfying the cut is $43.4 \pm 0.3\%$. In the case of exclusive B-decays specific cuts depending on the decay products are also used.

⁶ Apart from those used for the definition of the Geometrical Likelihood, which correspond to DC04.

Pileup events are simulated assuming an instantaneous luminosity of $2 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ and $5 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$, an inelastic cross-section of 80 mb and a non-empty crossing rate of 30 MHz.

4.1.2. Signal, Background and Control Channels Samples

A b quark production cross-section of $5 \times 10^{11} \text{ fb}$ is assumed throughout this document, as well as the hadronization fractions shown in Table 4-II, where Λ_b is considered to account for almost the totality of b-baryons.

Table 4-II: Simulated b-quark hadronization fractions

b hadron type	B_d	B^+	B_s	Λ_b^0	Other hadrons
$P(b \rightarrow H)$	0.4	0.4	0.1	0.1	small

The signal sample is obtained from minimum bias events, including pileup, where a b-quark is produced. The hadronization process is repeated until the correct hadron type B_s is produced. The B_s meson is then forced to decay into two muons using the EvtGen software[98][99]. The number of events analyzed is shown in Table 4-III. Given the extremely low branching ratio of the signal, a detailed understanding of the background is crucial in this analysis. There are several sources of background considered here: combinatorial background (where two real muons in the event combine to form a signal candidate), misidentified hadrons and exclusive decays that could simulate the signal.

Table 4-III: Analyzed events.

Process	BR	Crossection (fb)	Generator cut efficiency	#Events Analyzed	Integrated Luminosity (fb ⁻¹)
L0-Minimum Bias	-	5×10^{12}	-	4.01M	8.02×10^{-7}
Inclusive bb.	-	5×10^{11}	0.432	14.0 M	6.49×10^{-5}
$b\bar{b} \rightarrow \mu\bar{\mu}X$	-	-	-	23.7 M	4.90×10^{-3}
$B^+ \rightarrow J/\psi(\mu\mu) K^+$	5.97×10^{-5}	2.39×10^7	0.178	631 K	0.148
$B_d \rightarrow J/\psi(\mu\mu) K^*(K^+\pi^-)$	5.26×10^{-5}	2.10×10^7	0.173	2.48 M	0.683
$B_c^\pm \rightarrow J/\psi(\mu\mu) \mu^\pm \nu$	-	1.19×10^6	0.208	50.7 K	0.205
$B_s \rightarrow K^+ K^-$	1.94×10^{-5}	1.94×10^6	0.346	318 K	0.474
$B_s \rightarrow \pi^+ \pi^-$	5.13×10^{-6}	5.13×10^5	0.204	79.5 K	0.752
$B_d \rightarrow K^+ \pi^-$	1.94×10^{-5}	7.76×10^6	0.202	1.27 M	0.803
$B_d \rightarrow \pi^+ \pi^-$	5.13×10^{-6}	2.05×10^6	0.199	318 K	0.779
$\Lambda_b \rightarrow p K$	5.00×10^{-6}	5.00×10^5	0.210	86.4 K	0.823
$\Lambda_b \rightarrow p \pi$	3.10×10^{-6}	3.10×10^5	0.208	184 K	2.85
$B_s \rightarrow \mu^+ \mu^- \gamma$	1.20×10^{-8}	1.20×10^3	0.343	600	1.46
$B_s \rightarrow \mu^+ \mu^-$	3.35×10^{-9}	335	0.201	30 K	446

The combinatorial background and the misidentification are studied using inclusive samples of events containing a couple of b-quarks. The combinatorial background is studied in more detail using an inclusive

sample of events containing b-quarks and two muons of opposite charge within 400 mrad with respect to the beam axis; this is the so called, $b\bar{b} \rightarrow \mu^+\mu^-X$ sample in Table 4-III. This sample includes events where the two muons come from intermediate resonances and even a small fraction where the muon(s) are not related to the b-quark(s) in the event.

Several two body decays, very similar to the signal except for the Particle Identification likelihood (PID) of the candidates, have been studied and listed in Table 4-III. In addition, the process $B_c^\pm \rightarrow J/\psi(\mu\mu) \mu^\pm\nu$, identified as a possible source of background given the fact that the mass of the B_c^\pm (6286 MeV) is larger than the mass of the B_s (5369.6 MeV), has been studied in detail as well as a look at the possible contribution from the radiative decay $B_s \rightarrow \mu^+\mu^-\gamma$.

The decays $B^+ \rightarrow J/\psi(\mu\mu) K^+$ and $B_d \rightarrow J/\psi(\mu\mu) K^*(K^+\pi^-)$ are important control channels, together with the two-body decays $B \rightarrow h^+h^-$ as it will become clear along this note.

4.2. Analysis overview

The data analysis for extracting $\text{BR}(B_s \rightarrow \mu^+\mu^-)$ can be schematized as follows:

1. Reconstruction of all $\mu^+\mu^-$ combinations on triggered events.
2. Selection of $B_s \rightarrow \mu^+\mu^-$ candidates according to the stripping cuts (see 3.2.5).
3. Classification of each selected event in a binned 3D phase space, according to the following properties:
 - Geometry Likelihood (GL): Takes into account the geometrical properties of the candidate. It is explained more in detail in 4.2.2. The mathematical method to combine the different geometrical variables is explained in 4.2.1.
 - Particle Identification (PID) likelihood: It is the combined probability of the muon candidates to be real muons, over all the other particle hypotheses.
 - Invariant mass of the $\mu^+\mu^-$ couple. Only events in a window of ± 60 MeV (3σ) around B_s peak (5369.6 MeV) will be considered for the signal search.
4. The number of background events is computed interpolating from the mass sidebands (events outside the ± 60 MeV window), where the amount of signal is negligible. This is explained in more detail in 5.3.3.
5. The probability for a signal event to fall in each bin of the phase space is determined. Several control channels are used for that:
 - For the geometrical properties and the invariant mass, the two body decays of $B \rightarrow h^+h^-$ (where $h = K, \pi$ and the positive and negative hadrons are not necessarily of the same type), are used. Sections 5.3.1 (invariant mass) and 5.3.2 (GL) describe in more detail this subject.
 - For PID properties the control channels are: calibration muons (such as MIPs in the calorimeter, prompt $J/\psi \rightarrow \mu\mu$ decays or $B \rightarrow J/\psi(\mu\mu)X$) and hadrons coming from decays of prompt $K_s \rightarrow \pi^+\pi^-$, $\Lambda \rightarrow p\pi$ and $D^{*+} \rightarrow D^0(K\pi)\pi^+$. The momentum range of signal particles is known from $B \rightarrow hh$, thus the appropriate kinematical region from those PID control channels can be chosen.

6. Normalization: in order to translate the number of signal events into a BR, a measurement of the number of events from a decay with a known BR is used. As there is no accurate measurement of any B_s BR, normalization to B^+ or B_d decays is preferred. However, this implies a systematic error of 14% coming from the ratio of hadronization fractions of the b quark [7] (or, equivalently, the ratio of crosssections for B_s and the other B mesons). The normalization channels chosen are the decay $B^+ \rightarrow J/\psi (\mu\mu)K^+$, because of its large statistics and the J/ψ muons which correlate the reconstruction and trigger efficiencies for both signal and normalization channel, or $B_d \rightarrow K\pi$, which also has large statistics and has the same kinematics as the signal. A detailed explanation of the normalization procedure is given in 4.4.1.
7. Modified Frequentist Approach (MFA), the statistical method described in [100] [101], is used, in the case of exclusion, to get the maximum number of signal events compatible with the observed data configuration. In case of observation, it is used to get the significance.

4.2.1. Mathematical Method to Combine Correlated Variables

The following method is used to build a variable that contains most of the information related to the geometry of the event in an optimal way, taking correlations properly into account, as long as the transformation is linear. A similar method was described in reference [102]. For a given set of n variables, the procedure is the following:

1. Each of the variables, X_i , is transformed into a Gaussian (with $\sigma = 1$) variable, G_i . A convenient way to do this is first transform X_i into a uniform distributed variable U_i as,

$$U_i(X_i) = \frac{\int_{X_{\min}}^{X_i} \rho(x'_i) dx'_i}{\int_{X_{\min}}^{X_{\max}} \rho(x'_i) dx'_i}.$$

The Gaussian variable G_i is obtained just using the inverse error function,

$$G_i(X_i) = \sqrt{2} \operatorname{erf}^{-1}(2 \times U_i(X_i) - 1)$$

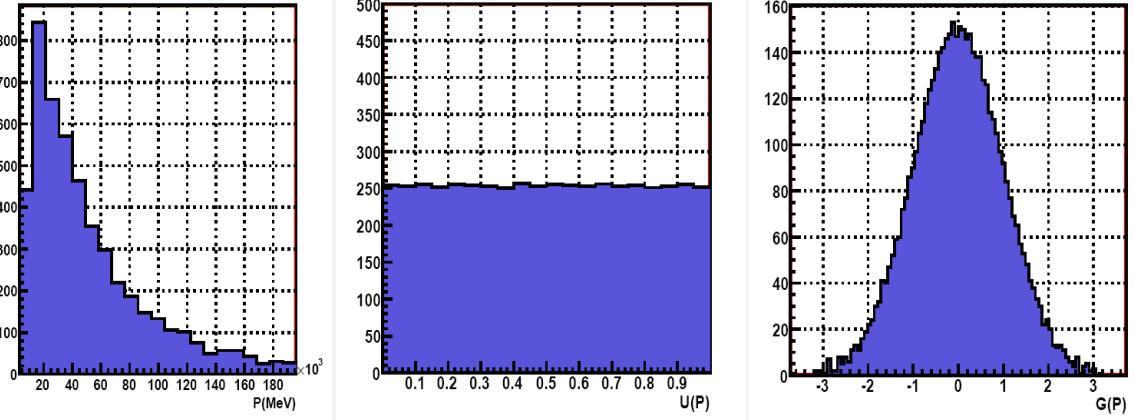


Figure 4-1: Momentum of the muons and its cumulative and Gaussian equivalents.

An example of that procedure is shown on Figure 4-1, where the momentum distribution of negative muons from $B_s \rightarrow \mu^+ \mu^-$ (left) is made uniform (center) and then Gaussian (right). The advantages of using those Gaussian distributions are:

- Easy computation of probabilities: In the Gaussian space the probability density for a given G_i is just:

$$\rho(G_i) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}G_i^2}$$

- Easy treatment of correlations: As all the variables will have the same distribution, ranges and correlations become easier to understand. In fact the correlations between the $\{G_i\}$ variables tend to get linear even if in the $\{X_i\}$ space they were not. This linearization can be illustrated with an analytical example. Suppose two original variables x, y 100 % correlated, that means $y = f(x)$. If $f(x)$ satisfies $\frac{df(x)}{dx} > 0 \forall x$ or $\frac{df(x)}{dx} < 0 \forall x$ then it is trivial to show that in the Gaussian space $G_x(x) = \pm G_y(y)$ (the sign is just the sign of the derivative in the original space), i.e., the correlation there is linear (with slope = 1) whatever it was in the original space. Figure 4-2 shows a graphical example being x the momentum of the muon and $y = \sqrt{p + p^4}$

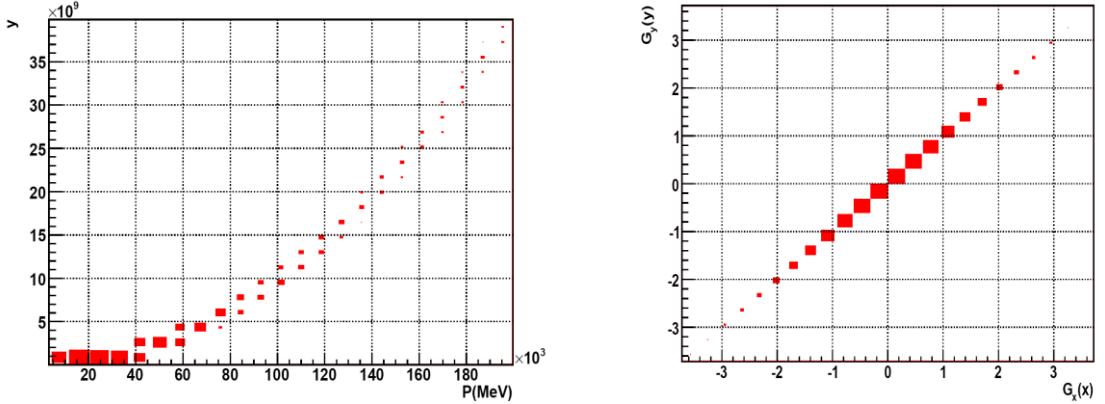


Figure 4-2: Correlation in initial and Gaussian space.

2. The linear correlation matrix $\bar{\bar{C}}$ and its corresponding rotation matrix $\bar{\bar{R}}$ are computed, so that the set of variables is rotated in order to be independent.

$$\bar{\bar{C}} \equiv C_{ij} = \sum_{k=1}^n (G_{i,k} - \langle G_i \rangle)(G_{j,k} - \langle G_j \rangle) = \sum_{k=1}^n G_{i,k}G_{j,k}$$

In the case of two variables having the same distribution, the rotation matrix is always (equivalent to) a $\pi/4$ rotation. This is due to the fact that the diagonal terms of the linear correlation matrix are equal as the distribution of x and y was forced to be the same. The off-diagonal terms are also equal each other by construction of the linear correlation matrix, so one always ends up with a correlation matrix $\bar{\bar{C}} = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$.

These new variables are transformed again into Gaussian variables repeating the procedure described above. A graphical example is shown on Figure 4-3, where the life time and the smaller of the IPS of the muons are transformed into a set of two uncorrelated Gaussian variables.

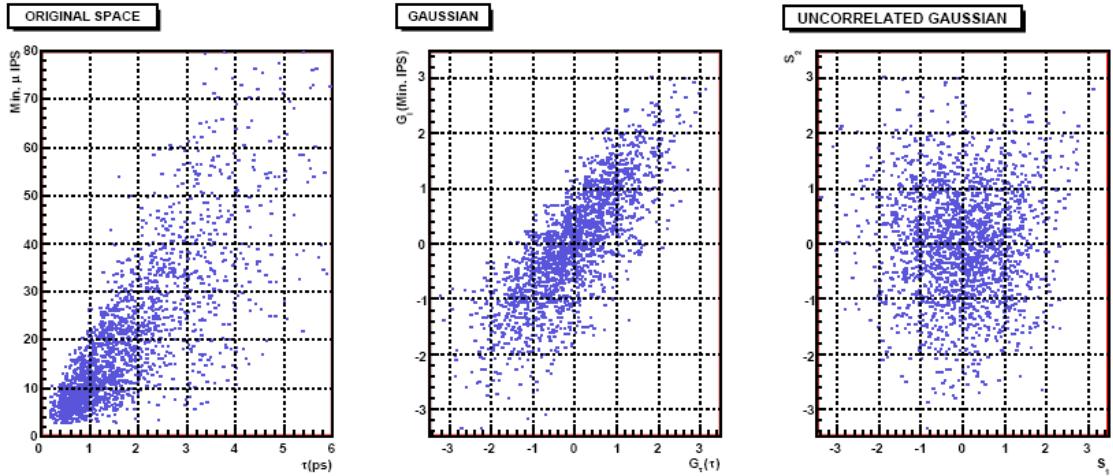


Figure 4-3: Transformation to uncorrelated Gaussians.

3. As the distributions and their correlations should be different in signal and background, the transformations above are computed twice: in a signal sample and in a background sample. Thus, from the original set of $N \{X_i\}$ variables, $2N$ output variables are created: $N \{S_i\}$, which would be Gaussian (and first order uncorrelated) in a signal sample and $N \{B_i\}$, which would have those properties in a background sample. An example is shown in Figure 4-4 where $S_{I,2}$ are the same as in Figure 4-3 and $B_{I,2}$ are those obtained using rotation matrix coefficients and cumulative functions from background sample.

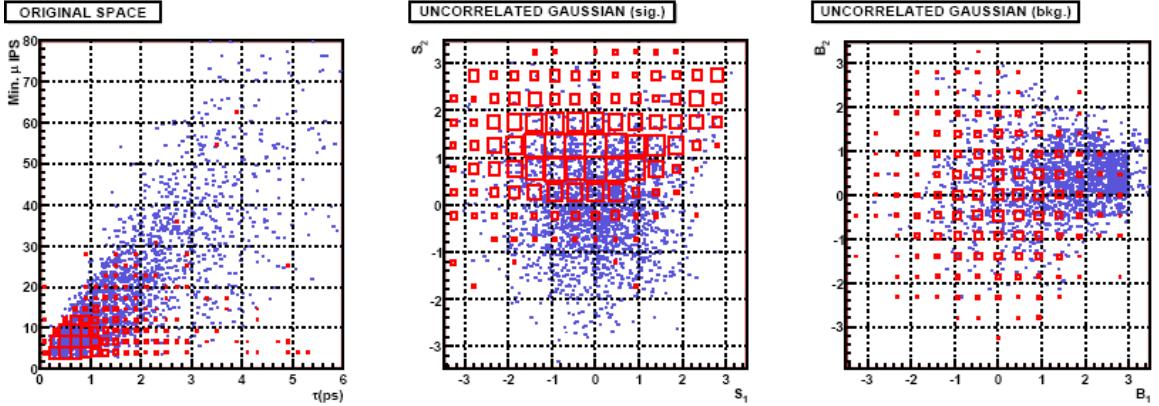


Figure 4-4: Gaussian variables in signal and background space.
Blue points are signal. Red boxes correspond to background.

4. For each event a χ^2 of signal hypothesis is then computed as $\chi_s^2 = \sum S_i^2$ where s_i are the Gaussian variables. The same procedure is applied to the background sample to compute $\chi_b^2 = \sum B_i^2$. For each event the quantity $\Delta\chi^2 = \chi_b^2 - \chi_s^2$ is used as discriminating variable. Again, transform this variable into a uniform variable (for the signal sample) distributed between 0 and 1 is convenient and the background will then peak at 0. Of course this last transformation does not increase the discriminating power, but is useful for visualization and straightforward computation of the efficiency for a given cut, as the efficiency will become just one minus the value of the cut. In Figure 4-5 the final variable from the combination in Figure 4-4 is shown.

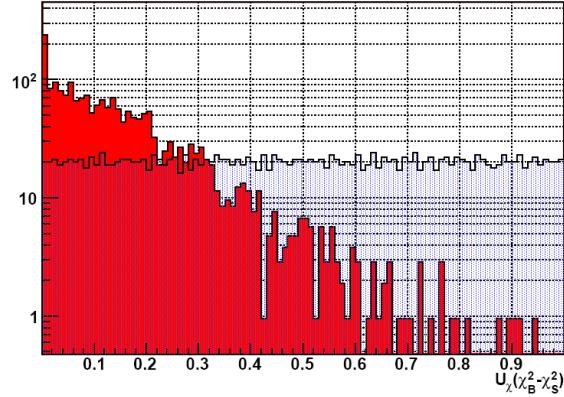


Figure 4-5: Discriminating variable distributions.

Blue dotted: signal (flat distribution). Red filled: background (peaked at 0).

4.2.1.1. Keeping one of the original variables

The steps 1-2 above allow getting a set of independent variables by mixing the original ones. If the uncorrelation was perfect, the distribution in the final space will be a Gaussian in N dimensions. That implies that there is an infinite set of combinations of final Gaussian uncorrelated variables. One might be interested in choosing a specific one.

For instance, suppose we want to calibrate some PID using a control sample. PID uses to depend on p and p_t , of the particles, in order to easily divide the phase space of those particles, p and p_t can be converted into two uncorrelated variables Q_1 and Q_2 . But for muon ID, the number of chambers used depends on the momentum p (equivalently, for hadron ID the radiator chosen at RICH system also depends on the momentum) so it could be interesting to keep it (or its Gaussian) as one of the variables and identify the

corresponding orthogonal. Therefore, a procedure to switch from $\{Q_1, Q_2\}$ to $\{G(p), G'\}$. Here the 2D case is analyzed.

Let $\{x, y\}$ the initial variables, $\{G_x, G_y\}$ its corresponding Gaussian, $\{q_1, q_2\}$ the rotated ($q_1 = \frac{1}{\sqrt{2}}(G_x + G_y)$, $q_2 = \frac{1}{\sqrt{2}}(G_x - G_y)$) and $\{Q_1, Q_2\}$ the final variables. We will consider (for detailed discussions see Appendix) that, if the uncorrelation was perfect, in the q_1, q_2 space (see Figure 4-6):

$$\rho_q(q_1, q_2) = \frac{1}{2\pi} e^{-\frac{1}{2}\left(\frac{q_1}{\sigma_1}\right)^2} e^{-\frac{1}{2}\left(\frac{q_2}{\sigma_2}\right)^2}, \quad \text{with } \sigma_1^2 + \sigma_2^2 = 2$$

Where the relation between $\sigma_{1,2}$ can be easily obtained by using that G_x and G_y are Gaussian distributed with $\sigma = 1$.

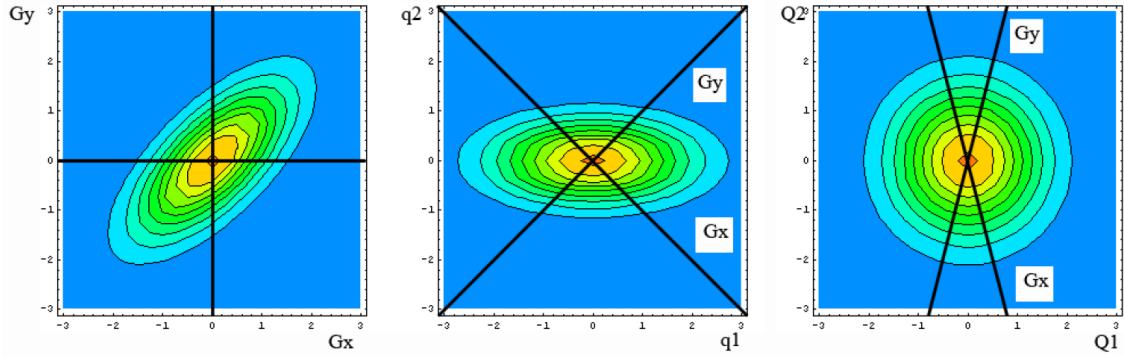


Figure 4-6: Transformation of original Gaussian axes in the ideal case.

So then $Q_1 = q_1/\sigma_1$ and $Q_2 = q_2/\sigma_2$. Thus the axis corresponding to G_x in $\{Q_1, Q_2\}$ plane can be identified, as well as its orthogonal one (and same for y). Calling σ_i the major of $\sigma_{1,2}$ (as in Figure 4-6), the expression for the orthogonal variables are⁷:

$$G_x = \frac{1}{\sqrt{2}} \left(\sigma_1 Q_1 - \sqrt{2 - \sigma_1^2} Q_2 \right), \quad G_x^\perp = \frac{1}{\sqrt{2}} \left(\sqrt{2 - \sigma_1^2} Q_1 + \sigma_1 Q_2 \right)$$

$$G_y = \frac{1}{\sqrt{2}} \left(\sigma_1 Q_1 + \sqrt{2 - \sigma_1^2} Q_2 \right), \quad G_y^\perp = \frac{1}{\sqrt{2}} \left(\sqrt{2 - \sigma_1^2} Q_1 - \sigma_1 Q_2 \right)$$

Note that the formulae above are only for positive correlation. In case of negative, calling again σ_i the major of $\sigma_{1,2}$:

$$G_x = \frac{1}{\sqrt{2}} \left(\sigma_1 Q_1 + \sqrt{2 - \sigma_1^2} Q_2 \right), \quad G_x^\perp = \frac{1}{\sqrt{2}} \left(\sqrt{2 - \sigma_1^2} Q_1 - \sigma_1 Q_2 \right)$$

$$G_y = \frac{1}{\sqrt{2}} \left(-\sigma_1 Q_1 + \sqrt{2 - \sigma_1^2} Q_2 \right), \quad G_y^\perp = \frac{1}{\sqrt{2}} \left(\sqrt{2 - \sigma_1^2} Q_1 + \sigma_1 Q_2 \right)$$

The σ_i parameter can be obtained from q_1 distribution, as well as from the major of the eigenvalues of correlation matrix as (see Appendix) $\sigma_{1,2}^2 = \lambda_{1,2}$.

⁷ Ambiguity in the sign of the orthogonal variables is of course possible.

4.2.2. Geometrical likelihood

The mathematical procedure described in 4.2.1 is used to combine the information about the geometrical properties of the candidate. The variables used are:

- DOCA of the $\mu^+ \mu^-$ pair (see ⁵ for definition of DOCA).
- Minimum IP of the B with respect of any reconstructed Primary Vertex.
- Proper time of the B, computed using the distance between the reconstructed Secondary and Primary vertices, and the reconstructed momentum of the B_s candidate. When more than one PV is reconstructed, the one that gives the minimum B_s impact parameter is chosen.
- The lowest impact parameter significance of the two muon candidates with respect to any of the primary vertices reconstructed in the event.
- Isolation of the muons: For each of the muon candidates, a search is performed for Long Tracks (excluding the companion muon candidate), that can make a “good” vertex with the muon candidate (i.e. $\text{DOCA} < 200 \mu\text{m}$ and the vertex coordinates along the beam axis should satisfy $0 < Z_{\mu+tr} < 3 \text{ cm}$). If we define $\alpha^{\mu+tr,PV}$ as the angle between the sum of the momentum of the muon and extra track and the direction defined by the PV and the vertex reconstructed using the muon and the extra track candidates, then the sum of the momenta is required to satisfy:

$$\frac{|\vec{p}_\mu + \vec{p}_{tr}| \cdot \alpha^{\mu+tr,PV}}{|\vec{p}_\mu + \vec{p}_{tr}| \cdot \alpha^{\mu+tr,PV} + p_{t,\mu} + p_{t,tr}} < 0.4,$$
where $p_{t,\mu}$ and $p_{t,tr}$ are the transverse momentum (with respect to the beam line) of the muon candidate and the extra track. This last pointing variable is similar to those used in the trigger, and was inherited from it as a good way to identify generic b decays. The number of tracks that satisfy these conditions is used as a discriminating variable for each of the muon candidates. Basically what the isolation does is to search for the number of $\mu+$ track candidates for each muon.

Taking into account that the isolation for each muon is used, the total number of variables entering in the GL definition is six.

4.2.3. Background studies

The background for this analysis is expected to be fully dominated by $b\bar{b}$ events, because of the $b\bar{b}$ purity after the trigger and the tight cuts on distance of flight and IPS applied in the stripping selection. The amount of candidates selected in $b\bar{b}$ for the signal and control channels is shown in Table 4-IV. As the control channels are present in the inclusive sample, the number of true signal candidates is also shown.

Comparing the corresponding rate with minimum bias rate, we can also estimate how representative is the $b\bar{b}$ background, at least for the control channels. This was done by scaling appropriately the number of events with the $b\bar{b}$ crosssection using in DC06 minimum bias generation, $\sigma_{b\bar{b}}^{\text{DC06}} = 698 \mu\text{b}$. The corresponding estimated fractions are shown in the 6th column.

Table 4-IV: Expected number of $b\bar{b}$ selected events.

The numbers in brackets correspond to one sigma interval. The fraction with respect to the total rate is also indicated.

	Selected Events	Selected Candidates	Signal Events	evts/fb ⁻¹	Fraction of $b\bar{b}$ in the rate
$B_s \rightarrow \mu\mu$	7	7	0	110k [70k,170k]	-
$B \rightarrow hh$	374	383	18	5.8 M	~60%

$B^+ \rightarrow J/\psi(\mu\mu) K^+$	552	745	45	8.5 M	~90%
$B_d \rightarrow J/\psi(\mu\mu) K^{*0}$	580	1709	13	8.9 M	~80%

This sample is enough to get an estimation of the amount of background for the control channels, but is not the case for $B_s \rightarrow \mu\mu$. In the previous studies we made, using looser selections [103][81], we found that the fraction of $B_s \rightarrow \mu\mu$ background candidates made with true muons was dominant and increasing rapidly with the GL (see plot Figure 4-7), being ~100% in the sensitive bins.

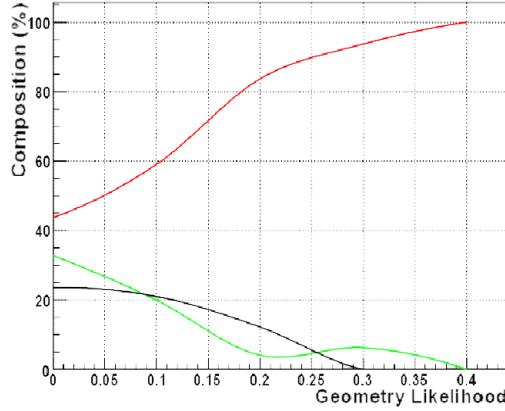


Figure 4-7: Fraction of events in the inclusive bb sample as a function of a cut in the GL used in [103].

Red (top line): two real muons from b-decays, black (bottom line at GL = 0): two real muons where at least one is not from b-decay, green (in the middle line at GL=0): at least one misidentified hadron. The sample corresponds to DC04 MC. The selection used in [103] is based on very loose cuts which selected in total about 90% of reconstructed signal.

We cannot do a similar study with only 7 events, but looking into them we see that 5 are $b \rightarrow \mu X, \bar{b} \rightarrow \mu Y$, one is $b \rightarrow \mu X, \bar{b} \rightarrow \bar{c} Y \rightarrow \mu Z$ and the remaining candidate has a missid pion $B^+ \rightarrow \mu^+ X, \bar{B}_d \rightarrow f(1270)Y \rightarrow \pi^- Z$. This confirms that the dominant background is $b\bar{b} \rightarrow \mu^+ \mu^- X$, as it includes 6 out of the 7 events, and moreover the missid (the only one that is not included in $b\bar{b} \rightarrow \mu^+ \mu^- X$ category) is the one with lowest GL (see Table 4-V) and hence the one with lowest impact on the sensitivity.

Table 4-V: $b\bar{b}$ background events after $B_s \rightarrow \mu^+ \mu^-$ selection.

Run number	Event number	Decay	GL
135401	15239756	$B_d \rightarrow \mu^+ X, \bar{B}_d \rightarrow \mu^- Y$	0.0066
135401	5978399	$B^+ \rightarrow \mu^+ X, B^- \rightarrow D^- Y \rightarrow \mu^- Z$	0.0031
135401	4885136	$B^+ \rightarrow \mu^+ X, B^- \rightarrow \mu^- Y$	0.0018
135601	5015725	$B^+ \rightarrow \mu^+ X, \bar{B}_d \rightarrow f(1270)Y \rightarrow \pi^- Z$	0.00079
136801	1348517	$B_s \rightarrow \mu^+ X, \bar{B}_d \rightarrow \mu^- Y$	0.0021
136801	4791219	$B^+ \rightarrow \mu^+ X, \bar{B}_s \rightarrow \mu^- Y$	0.011
135401	16910530	$B^+ \rightarrow \mu^+ X, \Lambda_b \rightarrow \mu^- Y$	0.15

Using a $b\bar{b} \rightarrow \mu^+\mu^-X$ inclusive sample, we estimate the amount of this background to be ~ 20.5 events per fb^{-1} in the signal search window region and in $\text{GL} > 0.5$, hereafter called “sensitive region” as it accumulates most of the sensitivity.

Several specific backgrounds were studied as well: $B_c^\pm \rightarrow J/\psi(\mu\mu) \mu^\pm v$, $B_s \rightarrow \mu^+\mu^- \gamma$ and misidentified $B \rightarrow hh$, finding that all of them are negligible in the signal window [81] [103]. The decays $B_c^\pm \rightarrow J/\psi(\mu\mu) \mu^\pm v$ and $B_s \rightarrow \mu^+\mu^- \gamma$ are not included in the $b\bar{b} \rightarrow \mu^+\mu^-X$ sample because of their low crosssections, so for this reason were studied separately. The misidentified $B \rightarrow hh$ might be not negligible in the left mass sideband; in this case special care is needed for the estimation of the background under the signal peak (see 4.4.4 and 5.3.3). The reasons which make these specific backgrounds to be negligible are:

- $B_s \rightarrow \mu^+\mu^- \gamma$. Even if the BR is larger than the one of the signal, it is still a rare decay. As the photon takes an important fraction of the momentum in the B_s rest frame, the invariant mass and the geometrical properties get strongly affected. In a sample corresponding to 1.46 fb^{-1} , no event was found in the signal region.
- $B_c^\pm \rightarrow J/\psi(\mu\mu) \mu^\pm v$. The $\mu\mu$ mass can reach the value of the B_s mass in a small fraction of cases. In a sample of 0.205 fb^{-1} , we find 3 events in the signal mass. However, the amount of momentum carried by the 3rd muon and the neutrino, affects geometrical variables like the IP of the reconstructed B_s . The isolation is different to that of the signal. At the end, among the 3 events in the tight mass region, the highest GL is smaller than 0.2. No events above $\text{GL} > 0.5$ were found even in the sidebands.
- $B \rightarrow hh$. The probability to identify both hadrons as muons is in the order of 10^{-4} , which strongly reduces the effective BR of those decays, up to the order of the one of SM signal and most of them come from the B_d , being outside of the B_s search window due to the good invariant mass resolution of LHCb. Furthermore, an important fraction of them are decays in flight, thus the track is a combination of the original hadron plus the originated muon. This effects both the geometrical properties and the invariant mass, moving the $B \rightarrow hh$ outside the sensitive region. Finally, the $B \rightarrow hh$ from the B_s meson is dominated by $B_s \rightarrow K\pi$ and $B_s \rightarrow KK$ and exchanging the mass of the kaon by the mass of the muon leads to an underestimation of the invariant mass, which pushes those backgrounds to the left sideband. About two events are expected in the sensitive region (see Figure 4-8) per fb^{-1} . Note that the PID likelihood would also help in the fighting of the remaining background.

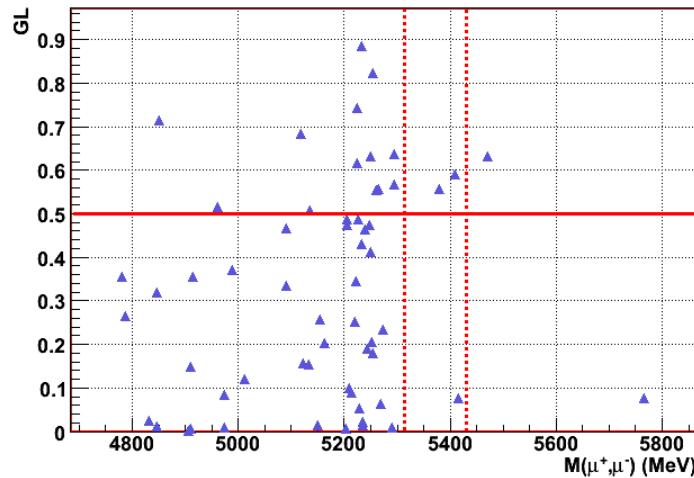


Figure 4-8: Misidentified $B \rightarrow hh$ as dimuon background.

The horizontal line corresponds to $\text{GL} = 0.5$. The dashed vertical lines show the limits of the signal search window.

4.3.Expected sensitivity

4.3.1. Sensitivity in nominal conditions

After the stripping selection, L0, HLT-1 and HLT-2, 41 SM signal and 25k $b\bar{b} \rightarrow \mu^+\mu^-X$ background events are expected in the tight mass window for each 2 fb^{-1} , the integrated luminosity corresponding to one LHCb nominal year. However, most of the background candidates fall in low values of the GL (see Figure 4-9). In particular above $\text{GL} > 0.5$, we expect 20 SM signal events and only 41 background events. In the central bin of invariant mass (see Table 4-VI) and $\text{GL} > 0.5$, a $S/B \sim 1$ is reached.

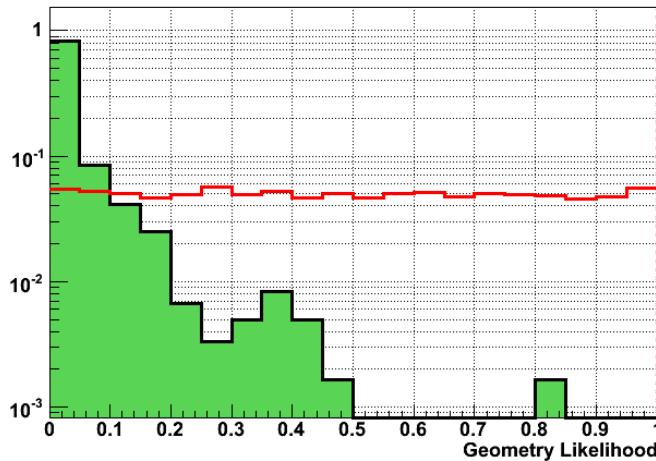


Figure 4-9: Geometry Likelihood distribution.

Green filled histogram: background. Red open histogram: signal. Both histograms are normalized to unity. Y axis is in logarithmic scale.

The signal and background annual yields, in the mass versus GL plane, are shown on Table 4-VI. The background yields were computed from the $b\bar{b} \rightarrow \mu^+\mu^-X$ sample, where one event falls in the last GL bin ($\text{GL} > 0.5$). Invariant mass and Geometrical Likelihood were assumed uncorrelated for the background. The numbers in brackets indicate 1σ interval of background expectation for the bins of the sensitive region ($\text{GL} > 0.5$).

Table 4-VI: Signal and $b\bar{b} \rightarrow \mu^+\mu^-X$ annual yields in Mass and GL bins.

	GL		
Mass (MeV)	0–0.25	0.25–0.5	0.5–1
5406.6 - 5429.6	S = 0.36	S = 0.36	S = 0.81
	B = 4700	B = 110	B = 7.8 [1.4,26]
5384.1 - 5406.6	S = 2.0	S = 1.8	S = 3.6
	B = 4500	B = 110	B = 7.6 [1.3,25]

5353.4 - 5384.1	S = 5.4	S = 5.5	S = 10
	B = 6000	B = 150	B = 10 [1.7,33]
5331.5 – 5353.4	S = 1.9	S = 2.1	S = 4.0
	B = 4400	B = 110	B = 7.6 [1.3,25]
5309.6 – 5331.5	S = 0.63	S = 0.67	S = 1.2
	B = 4400	B = 110	B = 7.6 [1.3,25]

From the MC predictions of signal and background properties and annual yields, the potential of the LHCb experiment on the exclusion or measurement of $\text{BR}(\text{B}_s \rightarrow \mu^+ \mu^-)$ can be extracted. MFA[100][101] is used in LHCb to get signal and background confidence levels for the analysis of $\text{BR}(\text{B}_s \rightarrow \mu^+ \mu^-)$. Figure 4-10 shows the BR excluded at 90% CL as a function of the integrated luminosity (left) and nominal time (right). Systematic errors were not added to this estimation, as are found to be small compared to the statistical errors. The expected limit at the end of Tevatron's Run II, 2×10^{-8} , is overtaken with less than 0.1 fb^{-1} . With 1 fb^{-1} limits around the SM prediction are set if no signal is present. The sensitivity of ATLAS and CMS is also shown, and has been computed using MFA from the quoted numbers of expected signal and background events[104] [105] (see Table 4-VII), without considering the effect of systematic errors.

Table 4-VII: Expected yields per fb^{-1} in ATLAS and CMS.

	Signal for $\text{BR}(\text{B}_s \rightarrow \mu^+ \mu^-) = 3.35 \times 10^{-9}$	Background
ATLAS	0.56	$1.4^{+1.3}_{-1.0}$
CMS	2.05	~ 6.53

For a given integrated luminosity, it is clear that LHCb performance is the best of the three experiments. This advantage is due to the capability to trigger at lower p_t 's, which gives to LHCb almost 3 times more effective $\text{B}_s \rightarrow \mu^+ \mu^-$ cross section than ATLAS and CMS [106]. The good detector performance and especially the invariant mass resolution are also advantages for LHCb. As a function of nominal time, considering 2 fb^{-1} per nominal year at LHCb and 10 fb^{-1} at ATLAS and CMS, the sensitivity of CMS becomes similar to that of LHCb.

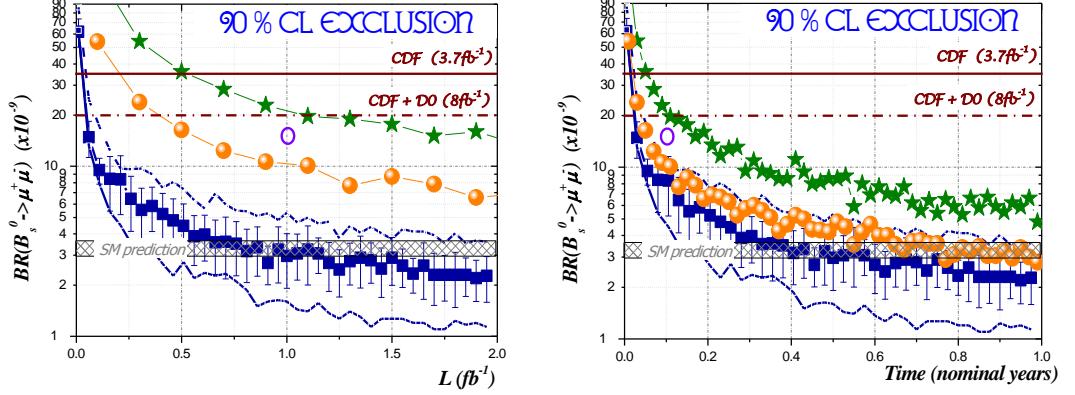


Figure 4-10: Sensitivity in the absence of signal.

The plot shows the BR excluded at 90% CL as a function of integrated luminosity (left) and nominal time (right). The blue squares correspond to LHCb sensitivity, where the error bars illustrate the rms due to statistical fluctuations of the background according to the expected value. The dashed lines show the uncertainty on LHCb sensitivity due to the limited statistics in the current simulation of the background, and correspond to 90% CL unified confidence interval of background estimation in $GL > 0.5$. The open violet circle corresponds to the expectation from CMS for $1 fb^{-1}$ including systematic errors[105] : $BR(B_s \rightarrow \mu^+ \mu^-) < 1.6 \times 10^{-8}$ at 90% CL. The horizontal full line shows the current upper limit from CDF, the dot-dashed one the expected limit at the end of Tevatron's Run-II assuming $8 fb^{-1}$ for each experiment. The filled horizontal bar shows the SM prediction. The sensitivity for ATLAS (green stars) and CMS (orange circles) is also shown and has been computed using MFA from the quoted signal and background annual yields.

In the case of presence of signal, the luminosity needed for the 3σ evidence of a given BR is shown in Figure 4-11. About $1-2 fb^{-1}$ are enough for a 3σ evidence if the BR is the SM prediction. This implies that a 5σ observation⁸ of a SM-like BR would require $\sim 3-5 fb^{-1}$. As in Figure 4-10, ATLAS and CMS sensitivities are also shown for comparison.

⁸ As the statistics on background for that luminosity is large enough to consider Gaussian errors, and then the luminosity needed for a 5σ observation is $25/9$ times the luminosity needed for a 3σ evidence.

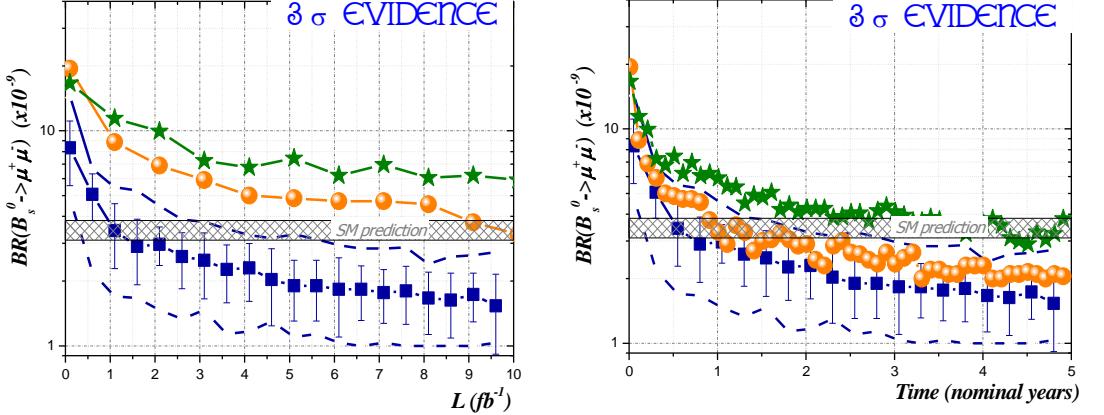


Figure 4-11: Sensitivity in the presence of signal.

The plot shows the minimum BR that would produce a 3σ evidence as a function of integrated luminosity (left) and nominal time (right). The blue squares correspond to LHCb sensitivity, where the error bars illustrate the rms due to statistical fluctuations of the background according to the expected value. The dashed lines show the uncertainty on LHCb sensitivity due to the limited statistics in the current simulation of the background, and correspond to 90% CL unified confidence interval of background estimation in $GL > 0.5$. The filled horizontal bar shows the SM prediction. The sensitivity for ATLAS (green stars) and CMS (orange circles) is also shown and has been computed using MFA from the quoted signal and background annual yields.

4.3.1.1. Analytical check

If we compare Figure 4-10 with Figure 4-11 it seems that, for a given BR, the amount of luminosity needed for a 3 sigma evidence is not much more than the needed for 90 % CL exclusion. This is seen for any of the three experiments. To get a better understanding of this, we can take the case of CMS which uses a cut analysis with 2.05 SM signal and 6.53 background expected events per fb^{-1} . For integrated luminosities larger than $4 fb^{-1}$, the background expectation is large enough to be considered a Gaussian number and then the statistical error is the squared root. A 3 sigma discrepancy with only-background hypothesis is equivalent to a up or down three sigma fluctuation of the amount background, which correspond to a background confidence level of 99.73 %. But as the signal contribution cannot be negative, the amount of signal that gives a $CL_b = 99.73\%$ is equivalent to a 2.81 sigma up fluctuation. Therefore, the 3 sigma evidence for CMS should follow:

$$S/\sqrt{B} = 2.81 \rightarrow 2.05 \cdot L \cdot \frac{BR}{3.35 \cdot 10^{-9}} \frac{1}{\sqrt{6.53 \cdot L}} = 2.81 \rightarrow BR = \frac{1.17 \cdot 10^{-8}}{\sqrt{L}}$$

Computing this for several luminosities as shown in Table 4-VIII, we find compatible numbers with what is seen in Figure 4-11.

Table 4-VIII: CMS 3σ evidence sensitivity.

$L(fb^{-1})$	$BR_{3\sigma}^{CMS}$
4	$5.9 \cdot 10^{-9}$
6	$4.8 \cdot 10^{-9}$
8	$4.1 \cdot 10^{-9}$

10	$3.7 \cdot 10^{-9}$
50 (five years)	$1.7 \cdot 10^{-9}$

The case of 90% CL exclusion is less simple. A signal hypothesis is excluded when $CL_s < 0.1$, being:

$$CL_s = \frac{CL_{s+b}}{CL_b}$$

When only the expected background is observed, $CL_b = 0.5$, so in this particular case $CL_s < 0.1$ means $CL_{s+b} < 0.05$, i.e. signal hypothesis is excluded at 90% CL if signal plus background hypothesis is excluded at 95 % CL.

Let's study the 1fb^{-1} case, i.e., CMS observes⁹ only the 6.53 expected background events. Then, using Poisson statistics, the number of expected events (S+B) is less than 11.84 at 95 % CL (or, equivalently, $CL_{s+b} < 0.05$ for $S+B > 11.84$) in order to be compatible with an observation of only 6.53 events. Then, the signal contribution that gives this $CL_{s+b} < 0.05$, and hence $CL_s < 0.1$, is $S < 11.84 - 6.53 = 5.31$ which corresponds to a BR of $8.7 \cdot 10^{-9}$. In summary:

$$(S + B)^{obs} = 6.53, \quad \langle B \rangle = 6.53; \quad \langle B \rangle = (S + B)^{obs} \rightarrow CL_b = 0.5 \rightarrow CL_s = 2 \cdot CL_{s+b}; \\ (S + B)^{obs} = 6.53 \rightarrow \langle S + B \rangle < 11.84 @ 95\% CL \rightarrow \langle S \rangle < 5.31 @ 90\% CL$$

Table 4-IX shows the result of this test for several luminosities, where we see results compatible with Figure 4-10.

Table 4-IX: CMS 90% CL exclusion sensitivity.

$L(\text{fb}^{-1})$	$\text{BR}_{90}^{\text{CMS}}$
0.5	$15 \cdot 10^{-9}$
1	$8.7 \cdot 10^{-9}$
1.5	$6.4 \cdot 10^{-9}$
2	$6.2 \cdot 10^{-9}$
10 (one year)	$2.4 \cdot 10^{-9}$

4.3.2. Sensitivity at LHC startup

As discussed in previous section, LHCb can provide interesting results in this channel with just few hundreds of pb^{-1} . It is hence interesting to study which sensitivity can be reached with the very first data, taking into account that the LHC will not be running at nominal performance since the beginning. The current expectation is that the accelerator will provide collisions at 7 TeV (instead of the nominal 14 TeV) and to increase the energy when possible. The instant luminosity will also be below the nominal value. The current schedule is to deliver 300 pb^{-1} per experiment in the first 10 months.

⁹ Of course only an integer number of events can be observed, the mean value is taken as an average of all possible situations.

Figure 4-12 shows, for the first 300 pb⁻¹, the BR excluded at 90% CL for absence of signal in the LHCb data sample. The $b\bar{b}$ cross section was assumed to be 45 % of the nominal (i.e 225 μb), corresponding to the ratio between the values at 7 TeV and 14 TeV in PYTHIA 6.4 [107][93]. The ratio is taken, instead of the absolute value, in order to be consistent with the assumption made in previous sections, 500 μb at 14 TeV. We can see from the plot that those first 10 months could be enough to allow LHCb overtaking any exclusion limit from Tevatron, if no signal is present.

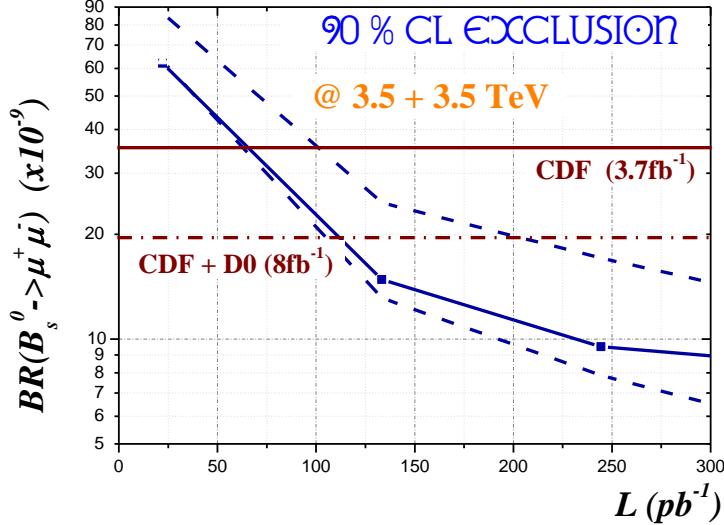


Figure 4-12: Sensitivity at LHC startup.

The plot shows the BR excluded at 90% CL as a function of integrated luminosity for a center of mass energy of 7 TeV. The dashed lines show the uncertainty on LHCb sensitivity due to the limited statistics in the current simulation of the background, and correspond to 90% CL unified confidence interval of background estimation in $GL > 0.5$. The horizontal full line shows the current upper limit from CDF, the dot-dashed one the expected limit at the end of Tevatron's Run-II assuming 8fb⁻¹ for each experiment.

4.3.3. Sensitivity to the B_d mode

The reconstruction, selection and trigger efficiency for $B_d \rightarrow \mu\mu$ must be the same as for the B_s mode. Due to the factor 4 larger crosssection for B_d than for B_s , the same yield in $B_d \rightarrow \mu\mu$ as the one of $B_s \rightarrow \mu\mu$ is reached for a 4 times smaller BR. Thus, one can in principle access smaller values of $BR(B_d \rightarrow \mu\mu)$ than for $B_s \rightarrow \mu\mu$. However, the SM prediction is about 30 times smaller for $BR(B_d \rightarrow \mu\mu)$ than for $B_s \rightarrow \mu\mu$ and of course the factor 4 larger B_d crosssection is not enough to compensate this. Hence, the amount of luminosity needed in order to reach values in the order of the Standard Model prediction is much larger for the case of $B_d \rightarrow \mu\mu$.

More quantitatively, considering that the background in the B_d mass is at most a factor ~ 2 larger than in the B_s mass (mainly due to the misid) one reaches the same significance to the case of B_s for a $BR(B_d \rightarrow \mu\mu)$:

$$\frac{BR(B_s \rightarrow \mu\mu)}{4} \lesssim BR(B_d \rightarrow \mu\mu) \lesssim \frac{\sqrt{2}BR(B_s \rightarrow \mu\mu)}{4}$$

Numerically:

$$0.25 BR(B_s \rightarrow \mu\mu) \lesssim BR(B_d \rightarrow \mu\mu) \lesssim 0.35 BR(B_s \rightarrow \mu\mu)$$

As the smallest values of $\text{BR}(B_s \rightarrow \mu\mu)$ that LHCb can access within five nominal years are, at most, in the order of $1-2 \times 10^{-9}$, we conclude that LHCb cannot observe $B_d \rightarrow \mu\mu$ if the BR is the SM prediction. However it can, of course, overtake the current limit from Tevatron.

4.4. The use of control channels

4.4.1. Normalization

The events that survive the selection described in section 3.2.5 are distributed in bins of a 3-dimensional space: Invariant Mass Likelihood, Geometry Likelihood and Muon ID likelihood. The procedure to calibrate these probabilities without relying on the MC simulation will be explained in 4.4.3 and 4.4.5. However in order to translate the observed number of candidate events in each 3D bin into a measurement of the Branching Ratio a global normalization factor is needed. If we decide to normalize to a calibration channel with a known Branching Ratio (BR_{cal}), then this factor can be written as:

$$BR = \frac{BR_{cal} \cdot \epsilon_{cal}^{REC} \epsilon_{cal}^{SEL/REC} \epsilon_{cal}^{TRIG/SEL}}{\epsilon_{sig}^{REC} \epsilon_{sig}^{SEL/REC} \epsilon_{sig}^{TRIG/SEL}} \cdot \frac{f_{cal}}{f_{Bs}} \cdot \frac{N_{sig}}{N_{cal}}$$

There “ f ” refers to the probability that a b-quark hadronizes into a B_s or B_x meson depending on the calibration channel chosen ($B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+$ or $B \rightarrow h^+h^-$). The ratio $f_{B_d(d)/f_{Bs}}$ is known with an uncertainty of $\sim 14\%$ [7]. The $BR(B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+) = (5.97 \pm 0.02) \times 10^{-5}$ and $BR(B \rightarrow K^+\pi^-) = (1.88 \pm 0.07) \times 10^{-5}$ are both known with a relative good precision of $\sim 3-4\%$, so they are good candidates as calibration channels. Moreover, they play a complementary role, because while for the $B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+$ the main difference w.r.t. the signal is the reconstruction of the extra kaon and the reconstruction of the lower momentum muons, for the $B_d \rightarrow K^+\pi^-$ the main difference is the trigger and the muon ID.

In the previous formula, the total efficiency has been separated for convenience in three components: ϵ^{REC} is the efficiency to reconstruct all the tracks needed for the signal and calibration channel, including the effect of the limited acceptance of the detector. $\epsilon^{SEL/REC}$ is the efficiency to select the events once they have been reconstructed. $\epsilon^{TRIG/SEL}$ is the efficiency of the trigger on reconstructed and selected events. Each of these components is discussed in the following sections, where it should become clear the convenience of these definitions.

4.4.1.1. Ratio of Reconstruction and Selection Efficiencies

The difference in the reconstruction efficiency, ϵ^{REC} , between $B_s \rightarrow \mu^+\mu^-$ and $B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+$ arises primarily from the efficiency to reconstruct an extra charged track in the control channel and in second order from the difference in space phase of the muon pair. Notice that if we decide to normalize to $B \rightarrow h^+h^-$ these effects should cancel out, except for the efficiency of the muon ID discussed in 3.2.3.2 and the different interaction of hadrons and muons with the material. For this reason it is an interesting alternative normalization. In the rest of this section we describe the case where the normalization channel is $B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+$.

The probability that all the tracks in the final state are within the LHCb acceptance depends on the phase space of these tracks. The charged particles may not traverse sufficient detector layers to be found by the track finding algorithms. In particular, the magnetic field introduces a cut-off at low momentum. The momentum spectrum of the B decays produced in LHCb has considerable uncertainty. Therefore, until the

spectrum has been properly calibrated, it would introduce a systematic uncertainty if the efficiencies are estimated using the simulation.

Track finding algorithms are also sensitive to occupancy. Therefore, the reconstruction efficiency depends on how busy are the events. It is unlikely that the existing Monte Carlo simulation estimates the occupancy reliably. All these arguments justify the need for an alternative method to evaluate these efficiencies (or rather the ratio between signal and control channel efficiencies) without relying on the simulation. So we plan to pursue two approaches:

- The efficiency ratio can be determined solely from the simulation. Using the control channel the simulation is either tuned or reweighted to properly represent the distributions to which the ratio is most sensitive, namely the momentum spectrum of the B meson and the occupancy of the event. Remaining differences between data and simulation are used to assign a systematic uncertainty.
- The efficiency ratio can be also estimated by considering another ratio of control channels, in order to probe explicitly the efficiency for reconstructing an extra track in the final state. A suitable control channel is $B_d \rightarrow J/\psi(\mu^+\mu^-)K^{*0}(K^+\pi^-)$. If the probability for a track to be reconstructed is independent of the momentum of the track, one may expect:

$$\frac{\varepsilon^{REC}(B^+ \rightarrow J/\psi K^+)}{\varepsilon^{REC}(B_s \rightarrow \mu^+\mu^-)} \approx \frac{\varepsilon^{REC}(B \rightarrow J/\psi K^*)}{\varepsilon^{REC}(B^+ \rightarrow J/\psi K^+)}$$

The first method is being worked out in the context of the tracking working group, and is not discussed further here.

The Branching Ratios for these control channels are known with enough precision: $BR(B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+)$ = $(5.97 \pm 0.02) \times 10^{-5}$ and $BR(B_d \rightarrow J/\psi(\mu^+\mu^-)K^*(K^+\pi^-))$ = $(5.26 \pm 0.24) \times 10^{-5}$. Hence, applying the stripping selections described in section 3.2.5 and neglecting the effect of the trigger (which will be discussed in section 4.4.2) one can obtain from the total number of selected $B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+$, N^{B^+} , and $B_d \rightarrow J/\psi(\mu^+\mu^-)K^*(K^+\pi^-)$, N^{B_d} , the ratio¹⁰:

$$\frac{\varepsilon_+^{REC} \times \varepsilon_+^{SEL/REC}}{\varepsilon_{sig}^{REC} \times \varepsilon_{sig}^{SEL/REC}} \approx \frac{BR(B^+ \rightarrow J/\psi K^+) \times N^{B^+}}{BR(B_d \rightarrow J/\psi K^*) \times N^{B_d}} \times \varepsilon_+^{M_{J/\psi}} \times \frac{1}{\varepsilon_d^{M_{K^*}} \varepsilon_d^{K,\pi_{IPS}}}$$

which is the first factor needed to compute the normalization. There most of the selection cuts cancelled to a good approximation, remaining only the J/ψ mass cut (which efficiency can be computed from data, thanks to the large statistics and purity on J/ψ samples), the K^* mass cut (which has been shown in 3.2.5 to be a systematic free quantity) and the IPS cuts applied on the daughters of the K^* , which have an efficiency close to 100% and can be checked by requiring even tighter cuts on the distance of flight. Hence, with an approximation valid at the percent level we can measure the ratio of offline efficiencies using the $B_d \rightarrow J/\psi(\mu^+\mu^-)K^*(K^+\pi^-)$ and $B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+$ control channels, without relying on any knowledge from the simulation.

In Figure 4-13 the dependence of the ratio calculated using the above method with the momentum of the B candidate is shown. The 3-body to 2-body ratio is calculated using $B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+$ and $B_d \rightarrow K^+\pi^-$ decays while the 4-body to 3-body ratio is calculated using $B_d \rightarrow J/\psi(\mu^+\mu^-)K^*(K^+\pi^-)$ and $B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+$ decays. It can be seen that both ratios are very similar in all the phase space. This sort of checks will be absolutely necessary when real data is in hand to understand if the approximations that seem to work fine in our simulation still hold with real data. Other ratios such as $B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+ / B_d \rightarrow K^+\pi^-$ without muon ID or $B \rightarrow 3h / B_d \rightarrow K^+\pi^-$ will also allow to study the 3 to 2 body ratio as well as the differences in the tracking reconstruction of hadrons with respect to the muons.

¹⁰ $P(b \rightarrow B_d) = P(b \rightarrow B^+)$ is assumed.

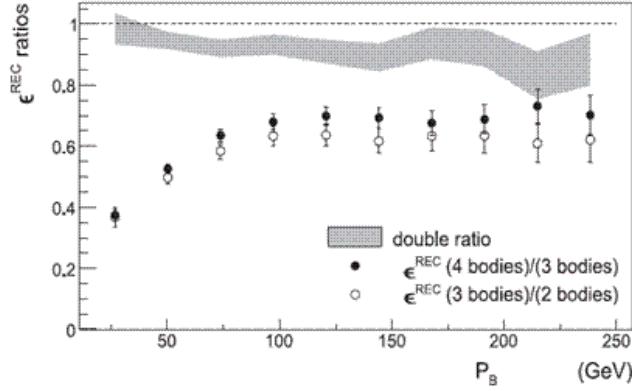


Figure 4-13: Ratio of reconstruction efficiencies as a function of B momentum.

Open circles: 3-body decays to 2-body decays ($B \rightarrow K^+\pi^-$). Black points: 4-body decays ($B \rightarrow J/\psi(\mu^+\mu^-)K^*(K^+\pi^-)$) to 3-body decays ($B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+$). The grey band shows the ratio between open and filled circles, showing that the 4 to 3 body ratio can be used as an estimator for the 3 to 2 body ratio in the full momentum range.

4.4.1.2. Normalization to an exclusive $B \rightarrow hh$ mode ($B_d \rightarrow K\pi$)

If we want to normalize to $B_d \rightarrow K\pi$, we need to separate the exclusive mode from the inclusive two-body decay. If we do not want to rely on the PID efficiencies obtained from the MC, one possible strategy is to measure the fraction (f') of the different two-body B decays in the limit of high purity, i.e. PID cuts high enough to ensure that only one of the modes (plus some possible combinatorial background) is present in the sample. Naming ε_π the efficiency (on $B \rightarrow hh$ pions) of the pion ID cut needed to reach high purity on pions, and ε_K the efficiency (on $B \rightarrow hh$ kaons) of the kaon ID cut needed to reach high purity on kaons, then:

- Require one pion and one kaon: The number of observed $B \rightarrow hh$ events will be $N'_{K\pi} = \varepsilon_K \varepsilon_\pi N_{K\pi}$. Where is the number of $K\pi$ events on the PID-unbiased sample. Note that B_s and B_d components can be separated through the invariant mass.
- Two pions: The number of observed $B \rightarrow hh$ events will be $N'_{\pi\pi} = \varepsilon_\pi^2 N_{\pi\pi}$
- Two kaons: The number of observed $B \rightarrow hh$ events will be $N'_{KK} = \varepsilon_K^2 N_{KK}$

The fraction f in the original sample can be then defined as a function of the fraction in the high purity limit:

$$f_{K\pi} = \frac{f'_{K\pi}}{f'_{K\pi} + f'_{KK} \left(\frac{\varepsilon_\pi}{\varepsilon_K} \right) + f'_{\pi\pi} \left(\frac{\varepsilon_K}{\varepsilon_\pi} \right)}$$

The ratio of PID efficiencies can be computed from the well known B_d Branching Ratios, as:

$$\frac{\varepsilon_\pi}{\varepsilon_K} = \frac{BR(B_d \rightarrow K\pi)}{BR(B_d \rightarrow \pi\pi)} \frac{N'_{\pi\pi}}{N'_{K\pi}} = (0.264 \pm 0.015) \frac{N'_{\pi\pi}}{N'_{K\pi}}$$

To ensure that the high purity limit is reached, the procedure can be repeated as a function of the applied cut, thus a plateau in the computed fractions should appear once the cuts are hard enough. This is illustrated in Figure 4-14, where a sample of $B \rightarrow hh$ being 68.1 % $B_d \rightarrow K\pi$, the computed fraction using this procedure was $f_{K\pi} = 0.68 \pm 0.03$, in very good agreement.

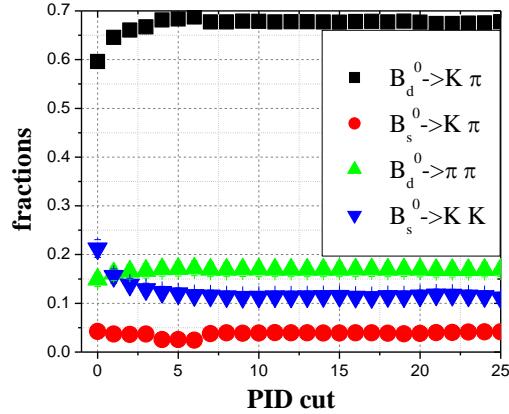


Figure 4-14: Measurement of exclusive $B \rightarrow hh$ fractions.

Also, the use of D^0 pions from the chain $D^{*+} \rightarrow D^0(k\pi)\pi^+$ to get the PID distribution in $B \rightarrow hh$ has been recently shown [81].

4.4.2. Trigger efficiencies and biases

4.4.2.1. Nomenclature

The nomenclature we use in this document is summarized in Figure 4-15. The square bubbles represent the event samples after successive levels of filtering. The first one contains all signal events, as *generated* by MC simulation or by nature. The second contains those events that would be accepted by the offline selection, while the third one contains the subset of the previous that pass the trigger.

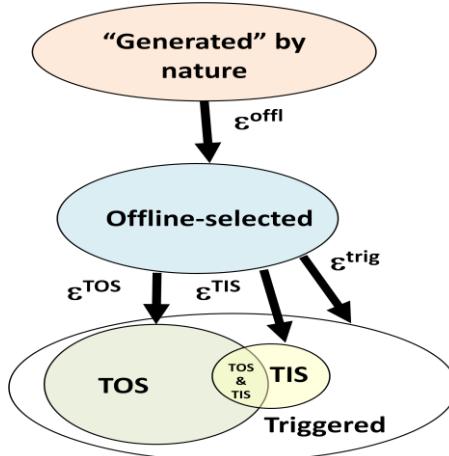


Figure 4-15: TIS-TOS-TOB nomenclature.

The method proposed in this note is based in the classification of observed events in the three categories proposed in [108]:

- Triggered on signal (TOS): events in which the detector information associated to the offline selected candidate is sufficient to trigger.

- Triggered independent of signal (TIS): events in which the detector information associated to the offline selected candidate is not necessary to trigger (for example, those events where the other B as fired the trigger). Note that TIS and TOS categories are not exclusive (for instance, both B's firing the trigger).
- Trigger on both (TOB): None of the previous, i.e., events in which the information associated to the offline candidate is needed but not enough to trigger.

The LHCb trigger system records all the information needed for such classification.

The efficiencies associated to each category are therefore defined as:

$$\epsilon^{TIS} \equiv \frac{N^{TIS \& SEL}}{N^{SEL}} \quad \epsilon^{TOS} \equiv \frac{N^{TOS \& SEL}}{N^{SEL}} \quad \epsilon^{TIS \& TOS} \equiv \frac{N^{TIS \& TOS}}{N^{SEL}}$$

4.4.2.2. Trigger efficiencies

From the definitions above, the total trigger efficiency can be written as:

$$\epsilon^{TRIG} \equiv \frac{N^{TRIG \& SEL}}{N^{SEL}} = \epsilon^{TIS} \frac{N^{TRIG \& SEL}}{N^{TIS \& SEL}}$$

Note that $N^{TRIG \& SEL}$ and $N^{TIS \& SEL}$ are observable quantities. If ϵ^{TIS} was the same for different channels, the above equation would allow computing the ratio of trigger efficiencies between them. That would be the case if particles from the signal decay were completely uncorrelated with the rest of the event (the *underlying event*).

As correlations between signal and the underlying event happen only through the three-momentum of the signal B (supposing all the B mesons are produced in the interaction point), ϵ^{TIS} will be the same for samples in which the B is selected in the same phase space. In that case:

$$\frac{\epsilon_1^{TRIG / SEL}}{\epsilon_2^{TRIG / SEL}} = \frac{\frac{N_1^{TRIG \& SEL}}{N_1^{TIS \& SEL}}}{\frac{N_2^{TRIG \& SEL}}{N_2^{TIS \& SEL}}}$$

However, using the fact that the underlying event and the B can only be correlated through tri-momentum of the B, even the absolute trigger efficiency can be extracted without use of MC simulation. In a given bin of tri-momentum:

$$\epsilon_i^{TIS} = \frac{N_i^{TIS \& TOS}}{N_i^{TOS}}$$

Hence, the number of events that would be offline selected can be computed as:

$$N^{SEL} = \sum_i N_i^{TIS} \Bigg/ \epsilon_i^{TIS}$$

Thus:

$$\epsilon^{TRIG / SEL} \equiv \frac{N^{TRIG}}{N^{SEL}} = \frac{N^{TRIG}}{\sum_i N_i^{TIS} \Bigg/ \epsilon_i^{TIS}} = \frac{N^{TRIG}}{\sum_i \frac{(N_i^{TIS} \cdot N_i^{TOS})}{N_i^{TIS \& TOS}}}$$

Where all the numbers are observable and, what is more, the ϵ^{TIS} in each bin can be computed in any channel or even combining several of them.

4.4.2.3. Trigger biases

The previous formalism also allows deconvoluting trigger effects from the observed event properties. In principle, already the offline selected candidates in TIS events are almost unbiased and hence have very similar properties than they would have in case of a 100% efficient trigger.

However, in order to do a rigorous recovering of the original properties each TIS event should be weighted accordingly to the phase space of the offline selected candidate, $w_i = 1/\varepsilon^{TIS,i}$.

This weight might make nonsense for background events, but anyway the background must be subtracted (using, for instance, sideband interpolation) in order to get the right distribution and this weighting procedure does not affect to subtraction techniques.

4.4.2.4. Trigger emulation

It has to be noticed that this strategy, as defined above, is not suitable if used to compute the efficiency for $B_s \rightarrow \mu\mu$ as we do not have enough $B_s \rightarrow \mu\mu$ (TIS) events in our data sample. One possible way out, without relying on the simulation is to use the sample of TIS¹¹ $B \rightarrow hh$ events and rerun the trigger with an appropriate parameterization of the online muon id obtained from an independent sample. In a similar way we can use the two muons from $B^+ \rightarrow J/\psi(\mu^+\mu^-)K^+$ decays and emulate the trigger response on them taking the momentum and IP distributions from the TIS $B \rightarrow hh$ events.

4.4.3. Invariant mass and Geometry likelihood calibration using $B \rightarrow hh$

In the previous section we have described how to translate the number of signal candidates into a measurement of the Branching Ratio. This normalization factor multiplies the number of candidates in each bin of the likelihood function. In this section we will describe the procedure to calibrate the likelihood assigned to each candidate. The likelihood is composed of three independent probabilities: invariant mass likelihood, geometrical likelihood and PID likelihood. The calibration of the Muon ID will be discussed in 4.4.5; here we focus in the first two components.

4.4.3.1. Invariant Mass Calibration

The invariant mass distribution of the signal is parameterized with a (simplified) crystal-ball function, which takes into account not only the experimental resolution, but also the $1/x$ behavior at low masses coming from the FSR (see 2.2.5.2):

$$\rho(m) \propto \begin{cases} e^{-\frac{1}{2}\left(\frac{m-M}{\sigma}\right)^2}, & \text{if } \left(\frac{m-M}{\sigma}\right) \geq -a \\ \frac{1}{\frac{1}{a} - \left(a + \frac{m-M}{\sigma}\right)}, & \text{if } \left(\frac{m-M}{\sigma}\right) < -a \end{cases}$$

The strategy to evaluate the invariant mass distribution for a signal candidate is to use the decay $B_s \rightarrow K^+K^-$. The mean and the resolution should be the same for both channels, and the a parameter accounting for the radiative tail can be corrected as the energy distribution for both channels is known. In order to select this decay from the inclusive $B \rightarrow hh$ with enough purity, relative strong cuts on the kaon ID are needed. The procedure is shown in Figure 4-16 where the kaon ID is tightened until a plateau (which ensures high purity on this mode) in the fitted parameters is reached. However, those strong kaon ID cuts bias the momentum of the kaons and hence changes the invariant mass resolution we are trying to measure.

¹¹ Note that, if a bias is expected even in the TIS sample, it can be removed as it was explained in 4.4.2.3.

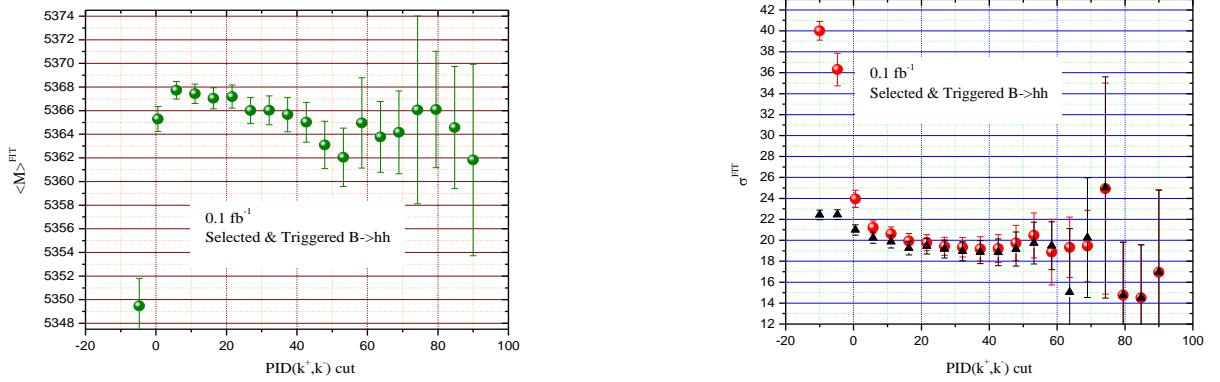


Figure 4-16: Fitted invariant mass parameters in $B \rightarrow hh$ as a function of kaon purity.

The black triangles in the right plot show the variation of the fitted sigma in an exclusive $B_s \rightarrow K^+K^-$ sample. A ~20% bias can be seen when strong kaon ID cuts are applied.

In order to evaluate this bias, which we will see is ~20%, the full $B \rightarrow hh$ sample can be used. There an effective sigma can be measured without PID cuts. Then the variation of that parameter when only well identified hadrons are taken into account should follow, at first order, the same evolution as the exclusive $B_s \rightarrow K^+K^-$ sigma, as the momenta are being biased in the same way. Note that first order knowledge of the correction is enough, as the correction itself is ~20%, a ~10 % error would lead only to a ~2% systematic.

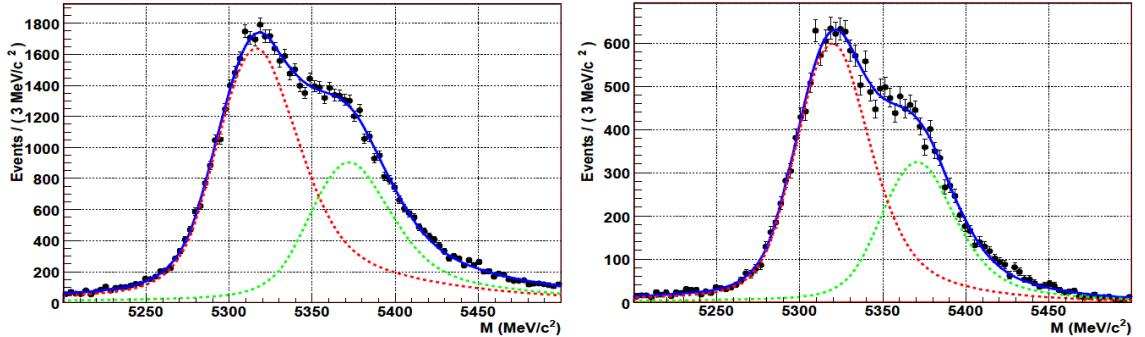


Figure 4-17: K^+K^- invariant mass distribution of the $B \rightarrow hh$ inclusive sample.

Left: No PID cut, giving $\sigma^{\text{eff}} = 24.5$ MeV. Right: after PID requirement $|DLL| > 20$, giving $\sigma^{\text{eff}} = 21.5$ MeV.

It is shown in Figure 4-17, where both B_d and B_s decays are forced to have the same resolution. Once the correction factor is known, it is applied to the distribution extracted from $B_s \rightarrow K^+K^-$. As can be seen in Figure 4-18, this correction is not negligible, ~20%, and once applied agrees nicely with the expected resolution from the MC.

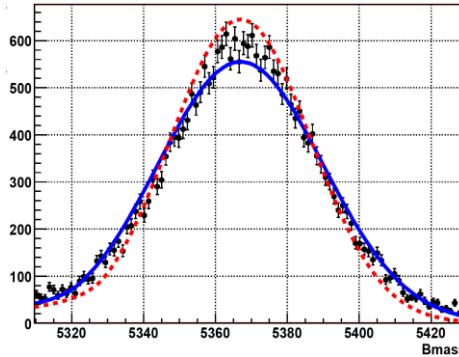


Figure 4-18: $B_s \rightarrow \mu^+ \mu^-$ invariant mass distribution (MeV).

Red dashed: Line shape without bias correction. Blue line: corrected line shape.

4.4.3.2. Geometry Likelihood Calibration

The transformation from the five variables into a single discriminating variable (GL) distributed between zero and one was computed from DC04 MC simulation. There is no need to change this transformation when doing the analysis with real data, at least for the first iteration. Even if the simulation of the variables entering the GL calculation does not agree with what we observe in the data, the only consequence will be a reduction in the discriminating power, but GL will be probably still a good discriminating variable. What needs to be calibrated is the distribution of GL itself, as this is what is used to compute the signal and background probability.

The GL distribution for signal-like events can be evaluated in principle using $B \rightarrow h^+ h^-$, however there are several issues to be considered and they are discussed in the following. The GL distribution for $B_s \rightarrow \mu^+ \mu^-$ events is designed to be flat between zero and one, and identical to the one obtained from $B \rightarrow h^+ h^-$ when no trigger is applied. Figure 4-19 shows the GL distribution computed from the simulation using $B_s \rightarrow \mu^+ \mu^-$ events separated according to the different triggers. The total distribution is very similar to the offline, as the trigger efficiency is very high. However, the shape would be very different if the efficiency was worse and the single muon trigger was dominating.

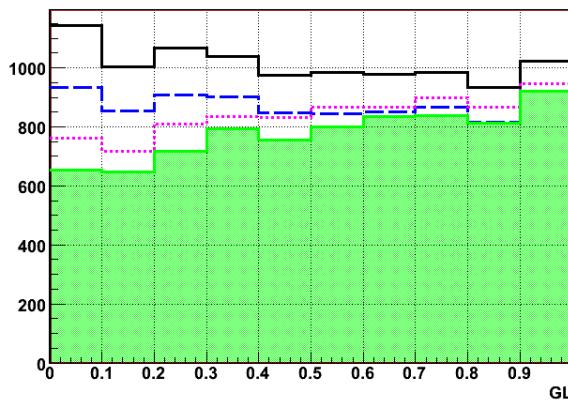


Figure 4-19: GL for different triggers.

Black solid line: all triggers. Blue dashed: dimuon triggers. Magenta dotted: muon plus track triggers. Green filled: single muon triggers.

The reason why the events triggered using the single muon trigger have such a different GL distribution compared with the events triggered with the Dimuon trigger are the hard cuts in p_t and IP. In order to

reproduce the muon trigger biases in the $B \rightarrow hh$ sample, we should use the trigger emulation described in 4.4.2.4.

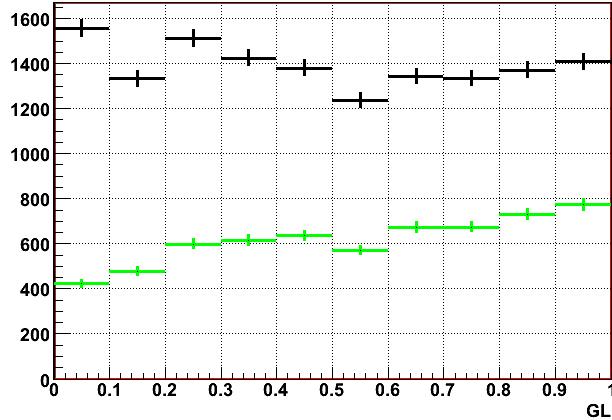


Figure 4-20: GL distribution in $B \rightarrow h^+h^-$.

Upper black points are offline selected events. Lower green points correspond to those in the sample that also pass the trigger.

In a similar case, the shape obtained from triggered $B \rightarrow h^+h^-$ candidates is going to be strongly affected as well due to the hard cuts in the hadron trigger. This is what is observed in Figure 4-20 (see, for instance, the first bin [0,0.1] has only 50% of population of the last one [0.9,1]) and thus we cannot use $B \rightarrow h^+h^-$ triggered events in a trivial way. The solution is to use again TIS events where the biases due to the $B p_t$, if visible, could be corrected as explained in 4.4.2.3. However, in the particular case of the GL, there is a second effect that is not accounted by that procedure. The isolation, which is one of the variables entering in the GL, has information not only about the signal, but about the underlying event as well. TIS events tend to have larger multiplicities than average signal events, and this is reflected in the isolation.

This effect was not visible in [81] probably due to the trigger version used for that note, which relied in GEC that removed high multiplicity events. GEC are not anymore present in the trigger version used for this thesis. Moreover, the tighter cuts in IP of the daughters and distance of flight make the flight related variables which enter in the GL to lose importance in favor of the others and hence distortions in the isolation have more impact in the final GL shape. Thus, to improve exactitude in the GL shape, we should also correct in bins of multiplicity. Alternatively, the GL can be redefined with a normalized isolation: $Iso' = Iso/N_{LongTracks}$. This redefinition makes the isolation to be more independent of the multiplicity improving slightly the agreement between TIS and would-be-offline selected events, as Figure 4-21 shows. Of course, the fact that there is more offline events with 0 associated tracks is not corrected by the denominator and the good agreement even at low values is due to a partial compensation in the vicinity of $Iso' = 0$.

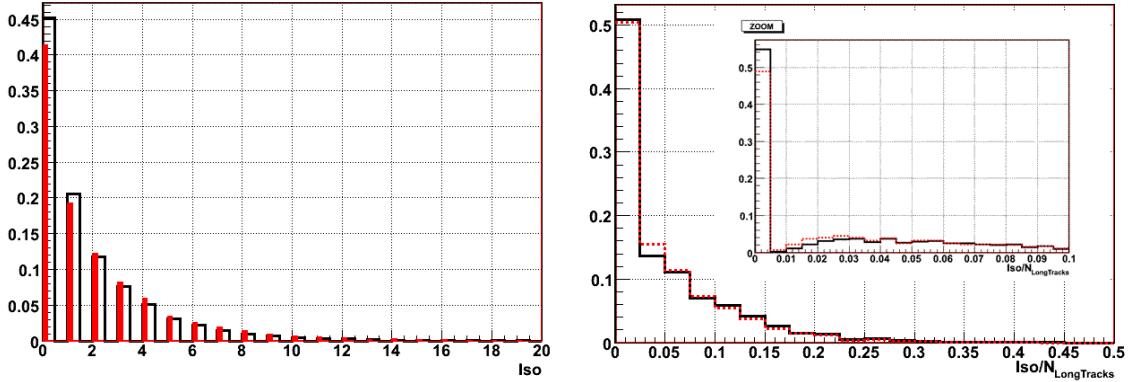


Figure 4-21: Comparison of isolation in offline with respect to TIS events.

Red filled/dotted histograms are TIS events. Black open histograms are offline selected events without any trigger effect. Left: Standard isolation. Right: Isolation divided by the number of long tracks. A zoom in the region close to 0 is included to show the remaining discrepancies.

The alternative of redefinition of the isolation is interesting as the resulting new GL shows the same rejection power as the original so it would not have an impact on the sensitivity.

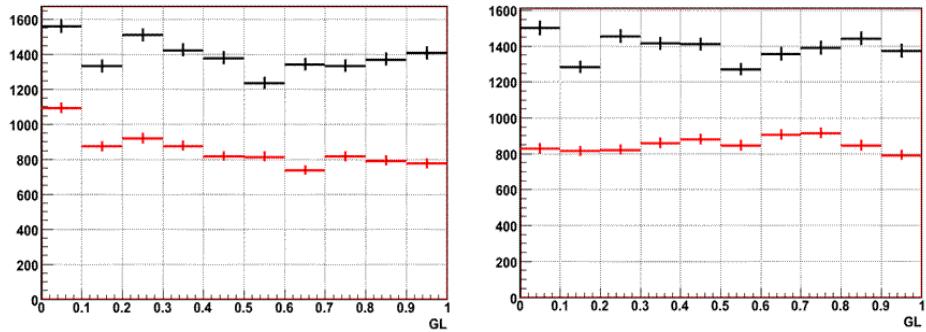


Figure 4-22: GL distribution in $B \rightarrow h^+h^-$ TIS.

The black upper points are for offline selected events. The red lower points are TIS events. Left: Standard GL. Right: GL defined with the isolation normalized to the number of long tracks in the event.

The comparison of the similarities between offline and TIS events are shown in In order include the effect of biases coming from the p_t dependence of ϵ^{TIS} , we weight the TIS sample according to the inverse efficiency. The weighted fractions are shown as TIS^{pt} on the table. We see that, even if the agreement improves slightly with the normalized isolation, some big (~8%) discrepancies remain.

Table 4-X: Fraction of $B \rightarrow h^+h^-$ in four bins of GL and GL with normalized isolation.

	GL			GL with normalized isolation		
	Offline	TIS	TIS ^{pt}	Offline	TIS	TIS ^{pt}
0-0.25	0.262 ± 0.004	0.284 ± 0.005	0.291	0.253 ± 0.004	0.238 ± 0.005	0.248
0.25-0.5	0.256 ± 0.004	0.254 ± 0.005	0.256	0.256 ± 0.004	0.256 ± 0.005	0.257
0.5-0.75	0.234 ± 0.004	0.232 ± 0.005	0.230	0.242 ± 0.004	0.261 ± 0.005	0.261

0.75-1.0	0.248±0.004	0.229±0.005	0.223	0.250±0.004	0.244±0.005	0.234
-----------------	-------------	-------------	-------	-------------	-------------	-------

Now let's compare with the results we can get by reweighting according to the multiplicity, in this case expressed as the number of long tracks in the events. There is some correlation between the number of tracks in the event and the p_t of the B, as more energetic collisions imply higher multiplicity. This correlation is shown in Figure 4-23, the scatter plot of the cumulative distributions of p_t and the number of long tracks. In case of no correlation, we should see a uniformly populated plane, but we can observe that for very high p_t , events with very low multiplicity are forbidden. Figure 4-24 shows the TIS probabilities measured (as $N^{TIS\&TOS}/N^{TOS}$) in a $B \rightarrow hh$ sample, as a function of p_t and the number of long tracks, n^{long} . In the case of n^{long} , the effect of weighting the numerator with $w(p_t) \propto 1/\varepsilon^{TIS}(p_t)$ to avoid overestimation of ε^{TIS} at high p_t is also shown.

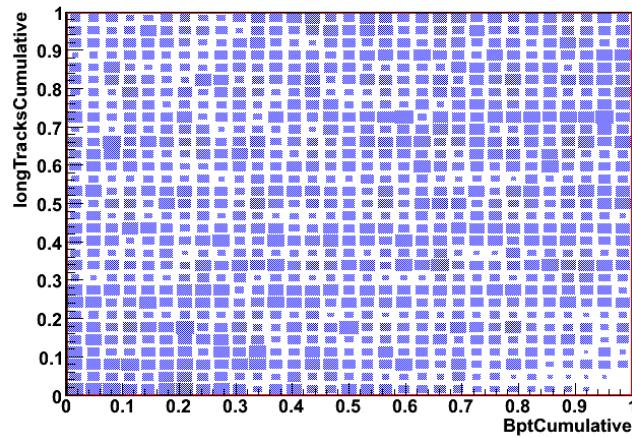


Figure 4-23: Correlation of the number of long tracks with the p_t of the B meson.

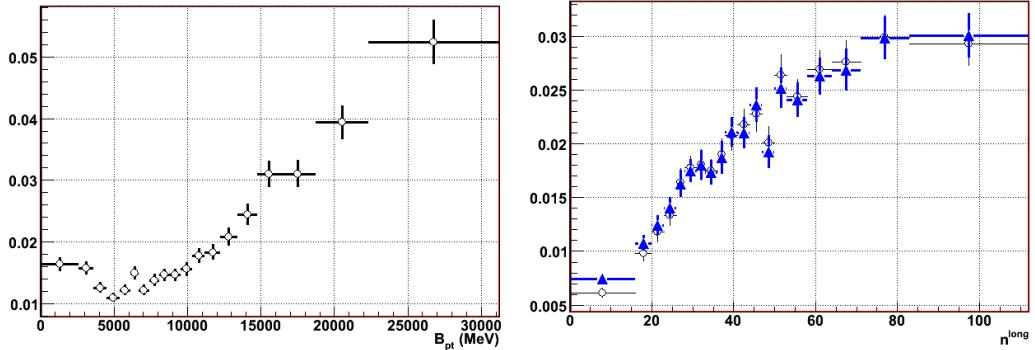


Figure 4-24: TIS observable efficiency as a function of B p_t and multiplicity.

The efficiency has been computed as $N^{TIS\&TOS}/N^{TOS}$. Left: Usual dependence with p_t . Right: Dependence with the number of long tracks. The open circles correspond to the measured quantity without any weighting. The blue triangles correspond to a weighting of $N^{TIS\&TOS}$ with $w(p_t) \propto 1/\varepsilon^{TIS}(p_t)$.

In Table 4-XI the effect of the multiplicity correction in the standard GL is shown. The columns **Offline** and **TIS^{pt}** are the same as in Table 4-X. The effect of correcting p_t and multiplicity is shown by **TIS^{pt,n}** and **TIS^{pt,n'}** where the events have been weighted according to

$$w(p_t, n^{long}) \propto 1/(\varepsilon^{TIS}(p_t) \cdot \varepsilon'^{TIS}(n^{long}))$$

The second case ($\mathbf{TIS}^{\text{pt},n'}$) also corrects the correlation $p_t \cdot n^{\text{long}}$, that means, the p_t spectrum of TIS events has been corrected to compute $\varepsilon^{\text{TIS}}(n^{\text{long}})$ as the p_t dependence is going to be accounted by $\varepsilon^{\text{TIS}}(p_t)$. Note that in both cases $\mathbf{TIS}^{\text{pt},n}$ and $\mathbf{TIS}^{\text{pt},n'}$ the TIS efficiencies are then considered uncorrelated and thus factorized.

Table 4-XI: Effect of multiplicity correction in GL calibration.

	Offline	TIS^{pt}	TIS^{pt,n}	TIS^{pt,n'}
0-0.25	0.262 ± 0.004	0.291	0.264	0.266
0.25-0.5	0.256 ± 0.004	0.256	0.247	0.248
0.5-0.75	0.234 ± 0.004	0.230	0.239	0.239
0.75-1.0	0.248 ± 0.004	0.223	0.250	0.248

We observe then that, after this multiplicity correction, the largest discrepancy is $\sim 3\%$ and not very different from the statistical fluctuations, achieving better accuracy than redefinition of the isolation studied above.

In the presence of background, the precision of the amount of signal in bins of GL is of course distorted, especially for low values of GL. Depending on the luminosity, the precision in the first bin of GL may be not very accurate. This is not a problem for normalization to $B \rightarrow h^+ h^-$, as a cut in GL can be performed and would have the same efficiency in signal and control channel. For normalization to $B^+ \rightarrow J/\psi(\mu^+ \mu^-) K^+$ we can use a tighter common selection to remove most of that background. An example of that procedure, for an integrated luminosity of 0.15 fb^{-1} , is shown in 5.3.2.2.

In principle, using exclusive $B \rightarrow h^+ h^-$ decays would help to reduce the background, but this implies applying relatively strong PID cuts, that again modify the shape of the GL distribution we are trying to obtain. The idea is then to use the inclusive $B \rightarrow h^+ h^-$ TIS events to determine the GL distribution for signal events. If the PID cuts have to be used, the bias could be corrected using $D^{*+} \rightarrow D^0(K\pi)\pi^+$.

4.4.4. Background

The strategy to know the amount of background in each bin relies in use the events in the sidebands of the invariant mass. The selected sidebands are 10 times larger than the window used for searching the signal. A linear or exponential interpolation is in principle enough to describe the background in the search window, as the background for $B_s \rightarrow \mu^+ \mu^-$ is dominated by $b\bar{b} \rightarrow \mu^+ \mu^- X$. However, in high values of the GL and higher luminosities, this interpolation can be distorted due to the presence of specific backgrounds in the left sideband, like $B \rightarrow h^+ h^-$. In those cases, an extrapolation from the right sideband or a more accurate description of the background will be needed. Some examples are given in section 5.

4.4.5. Muon ID calibration

The standard Muon ID algorithm[109] is a two steps procedure. In the first step, a track is required to have a given number of muon hits inside a Field of Interest (FoI) to be considered as a muon. For example a muon with a momentum above 10 GeV should have hits in all muons stations inside the FoI. In the second step, the DLL (see 3.2.3.2) is used. This procedure has some limitations as far as concern its applicability to real data: the *IsMuon* variable is very sensitive to small variations of the MWPC efficiency. Moreover the discriminating variable parameterization does not take properly into account the momentum dependence and it is therefore strongly dependent on the momentum spectrum of the analyzed sample.

In order to overcome these problems several solutions are under study [110]. Here we will describe a generic strategy to calibrate the discriminating variable, and we will use the standard Muon ID algorithm as example.

One possible strategy to calibrate the Muon ID likelihood is to measure the Muon ID efficiency and misidentification rate using a sample of muons from the decay $J/\psi \rightarrow \mu^+ \mu^-$ and a sample of hadrons from the decay $\Lambda \rightarrow p \pi$.

The $J/\psi \rightarrow \mu^+ \mu^-$ decay is an abundant process at LHC with a very clean signature and therefore it can be used to have a pure sample of muons. The inclusive J/ψ cross section is $\sim 290 \mu\text{b}$ with 93% from prompt production and only 7% from B decays. The expected yield in LHCb is 1.7×10^9 events per fb^{-1} .

The strategy is to use $J/\psi \rightarrow \mu^+ \mu^-$ events in which one of the muons is identified without the information from the muon stations. This second muon is required to be TIS (see 4.4.2.1) and selected by using the energy released on the electromagnetic (ECAL) and hadronic (HCAL) calorimeters by a Minimum Ionizing Particle (MIP). Figure 4-25 shows the energy distribution in ECAL and HCAL deposited by MIPs. The distributions obtained from the MC simulation show a clear peak which allows a clean selection.

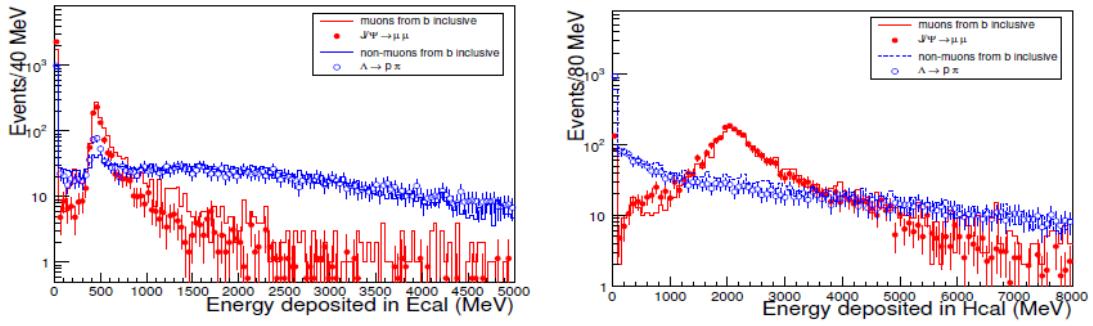


Figure 4-25: Energy deposited by MIPs.

Left: ECAL. Right: HCAL. The blue histogram corresponds to the energy distribution for non-muons while the red histogram is for muons. The blue open and red filled data points correspond to the energy distribution obtained from the calibration

Simple cuts on the vertex χ^2 and the energy distribution in the calorimeters allow reaching a very high purity. Figure 4-26 shows the J/ψ invariant mass distribution obtained from a sample of $\sim 4 \times 10^6$ Minimum Bias events after the L0 trigger (i.e. 4 seconds). These calibration muons allow knowing *IsMuon* efficiency as a function of phase space, as well as the DLL distribution.

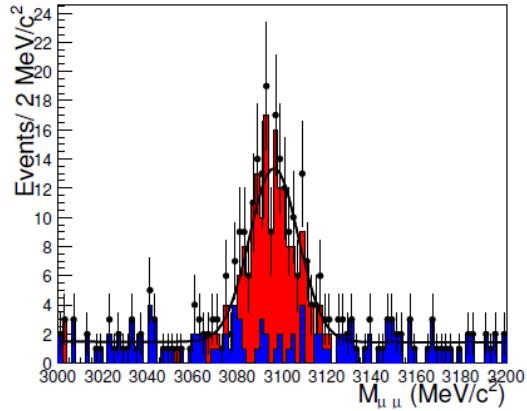


Figure 4-26: $J/\psi \rightarrow \mu^+ \mu^-$ in minimum bias sample.

In order to calibrate the misidentification probability the strategy is to use the decay $\Lambda \rightarrow p\pi^-$. In fact this is a quite abundant process at LHC (a $\Lambda \rightarrow p\pi^-$ decay every 6 Minimum Bias events is produced) and can be used even in the initial phase of the experiment, in the low luminosity regime. This channel is particularly interesting for studying separately the misidentification due to hadrons decaying and non-decaying in flight, due to the fact that contains both protons and π among the decay products.

Moreover the Λ is a very narrow resonance (~ 1 MeV) and has a very long lifetime (7.9 cm), therefore can be easily selected without any particle ID using only tight cuts in impact parameters, flight distance and invariant mass.

Figure 4-27 shows the Λ mass peak emerging from a sample of $\sim 4 \times 10^6$ Minimum Bias events, selected with a purity larger than 95%.

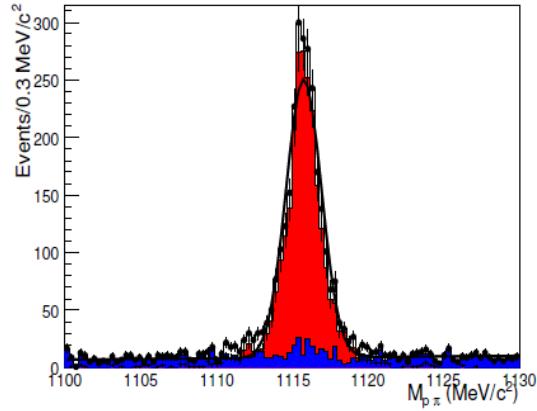


Figure 4-27: $\Lambda \rightarrow p\pi^-$ in minimum bias sample.

With the $J/\psi \rightarrow \mu^+ \mu^-$ and $\Lambda \rightarrow p\pi^-$ samples we can now extract the muon id likelihood from data. Figure 4-28 shows the DLL corresponding to muons and non-muons for an inclusive $b\bar{b}$ sample together with the results obtained from the calibration samples described above. Even if the agreement is not perfect, it indicates that this strategy is feasible. For more details on muon PID calibration see [111].

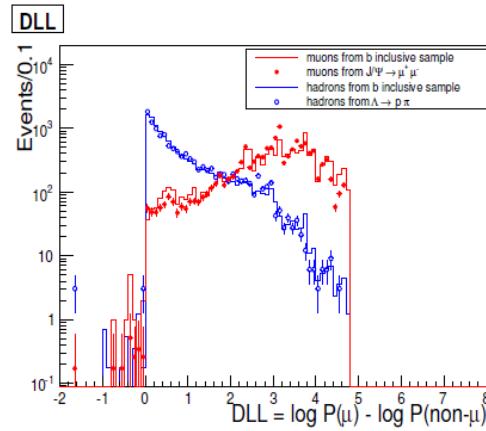


Figure 4-28: Calibration of Muon ID DLL.

The blue histogram corresponds to DLL for non-muons while the red histogram is the DLL for muons, in a sample of inclusive $b\bar{b}$ events. The blue open and red filled data points correspond to the DLL obtained from the calibration samples.

Then, once the *IsMuon* efficiency and the distribution of the DLL are known as a function of muons phase space, the distribution for the signal can be computed using the phase space of $B \rightarrow hh$ TIS events.

5. Full analysis example

5.1. Introduction

To illustrate the procedures described in 4, an example of a full analysis is given in this section. The integrated luminosity corresponding to this analysis is approximately 0.15 fb^{-1} . A $\text{BR}(B_s \rightarrow \mu^+ \mu^-)$ of 2×10^{-8} was used for the signal, in order to show both BR exclusion and signal significance. DC06 MC was used for signal, control channels and some specific backgrounds, while dominant background samples were generated by the procedure in 5.2.

5.2. Generation of large statistics background samples

5.2.1. Basic procedure

As seen in Table 4-III, the available statistics of full MC background samples corresponds at most to an effective integrated luminosity 50 pb^{-1} . To generate the background events for the analysis example, the following procedure was used:

1. Take the available full MC sample and transform the n variables that are going to be used into a set of n (almost) uncorrelated Gaussians according to the procedure described in 4.2.1.
2. Generate the desired number of events with n random Gaussian (with mean 0 and sigma 1) numbers.
3. Apply the inverse transformation on those Gaussian distributed events, such that they will have similar distributions and correlations like the original full MC sample.

With the above procedure we can generate large statistics of background reproducing its distributions and correlations between the different variables. Of course this procedure does not add new information to what is available from the full MC, so the events generated in this way cannot be used to improve our understanding of LHCb performance, but they are suitable for analysis examples like this one.

5.2.2. $B \rightarrow h^+ h^-$ sample

$B \rightarrow h^+ h^-$ is used for obtaining the shape of the GL and the mass. Apart from this two variables, we need to generate the p_t of the B, the PID(K- π) and the number of long tracks, as they are needed to compute the right shapes of those wanted variables. The IP and the IPS of the B, as well as the DOFS and the minor of the IPS of the muons, are going to be generated as well. This is motivated by the fact that a tighter selection will be needed in order to understand the lower GL bins with such a low integrated luminosity.

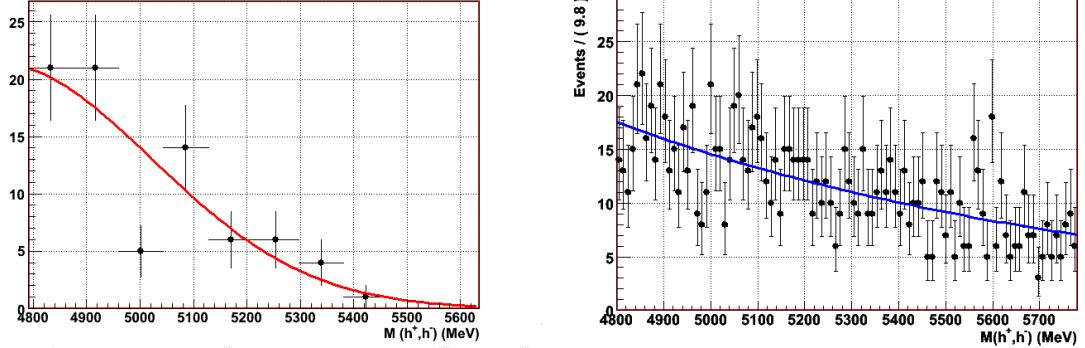


Figure 5-1: Mass distributions of $B \rightarrow h^+ h^-$ background.

Left: Physical background, fitted to a Gaussian of mean 4729 MeV and sigma 294 MeV Right: Combinatorial background, fitted to an exponential with parameter $-9.29 \times 10^{-4} \text{ MeV}^{-1}$.

The background is generated based on full MC $b\bar{b}$ events, separated into two categories:

- Physical background: This category corresponds to the case in which the two particles that define the candidate come from the same b hadron.
- Combinatorial background: This is the complementary category to the physical background.

As the $b\bar{b}$ sample accounts only for $\sim 60\%$ of the total background of $B \rightarrow h^+ h^-$, an extra component is added, accounting for the extra 40 % of minimum bias. This extra contribution is assumed to have the same properties as the combinatorial background. Both samples are generated separately using the procedure described in 5.2.1 for all the variables but the invariant mass, which is added afterwards using the p.d.f from the fits shown in Figure 5-1.

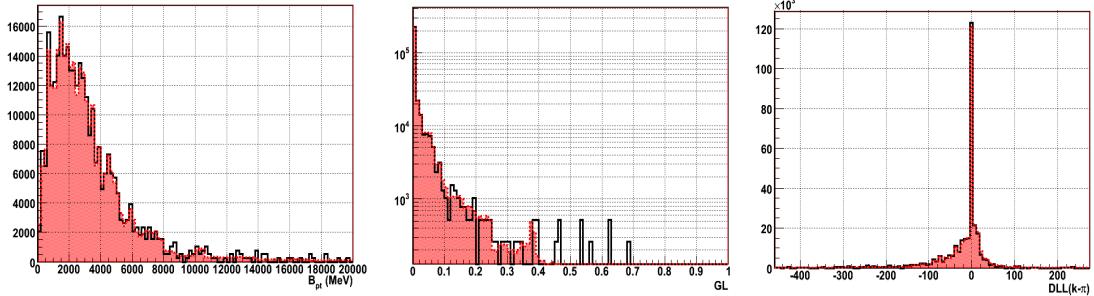


Figure 5-2: Distributions of generated $B \rightarrow h^+ h^-$ combinatorial background events.

Open histogram: original full MC sample. Red filled histogram: sample generated as described in 5.2.1. From left to right: p_T of the B candidate, GL and $DLL(k-\pi)$.

The TIS probability has been taken into account in order to provide the right fraction of physical and combinatorial background corresponding to a TIS sample. The TIS efficiency used was 2.0 % for physical background and 6.7 % for combinatorial. Those numbers were taken from full MC simulation. A total of 3652 physical and 311626 combinatorial events were generated in this way, to compose the background of the $B \rightarrow h^+ h^-$ TIS data. Figure 5-2 and Figure 5-3 show the good similarity between the properties of the generated sample and the ones observed in full MC simulation. The background from Λ_b is accounted by adding 55 $\Lambda_b \rightarrow pK$ and 36 $\Lambda_b \rightarrow p\pi$ from the full simulation sample as the statistics available in any of them is larger than 0.15 fb^{-1} . The signal is composed by 1596 $B \rightarrow h^+ h^-$ TIS from the full simulation sample.

For trigger efficiency studies, a sample of xTOS events was generated including 490755 combinatorial background, 60092 physical background, 1533 $\Lambda_b \rightarrow pK$, 949 $\Lambda_b \rightarrow p\pi$ and 40958 $B \rightarrow h^+h^-$ events mixed up using the following fractions:

$$\begin{aligned} f_{B_s \rightarrow KK} &= 0.100 \\ f_{B_d \rightarrow K\pi} &= 0.684 \\ f_{B_d \rightarrow \pi\pi} &= 0.174 \\ f_{B_s \rightarrow K\pi} &= 0.042 \end{aligned}$$

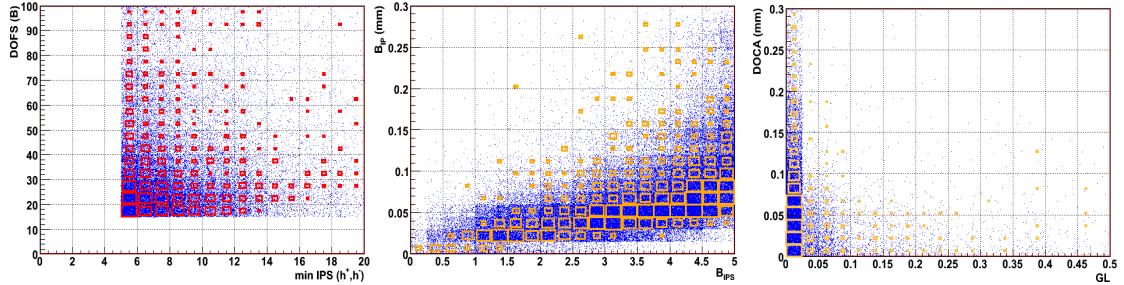


Figure 5-3: Correlations of generated $B \rightarrow h^+h^-$ combinatorial background events.

Open boxes: original full MC sample. Blue marks: generated sample. From left to right: IPS of the daughters versus DOFS, IPS of the B versus IP of the B in mm and GL versus DOCA.

5.2.3. $B^+ \rightarrow J/\psi(\mu\mu) K^+$ and $B_d \rightarrow J/\psi(\mu\mu) K^{*0}(K\pi)$ sample

For these control channels, the main distributions to be reproduced are the p_t of the B , the IP and IPS of the B , DOFS and DOCA of the J/ψ , the IPS of the kaon and p_t and total momentum of the muons as well as their DLL of muon to pion hypothesis. The binary labels of TOS and TIS are also generated in order to have the right proportion of background when computing the trigger efficiencies. As in the $B \rightarrow h^+h^-$ case, the invariant mass is added afterwards accordingly to the analytical shape.

A total of 1779661 background events for $B^+ \rightarrow J/\psi(\mu\mu) K^+$ and 2083302 background events for $B_d \rightarrow J/\psi(\mu\mu) K^{*0}(K\pi)$ were generated.

The amount of signal added to those samples was 126411 $B^+ \rightarrow J/\psi(\mu\mu) K^+$ events and 37084 $B_d \rightarrow J/\psi(\mu\mu) K^{*0}(K\pi)$ events.

5.2.4. $B_s \rightarrow \mu^+\mu^-$ sample

For $B_s \rightarrow \mu^+\mu^-$ background, a sample of 18489 $b\bar{b} \rightarrow \mu^+\mu^-X$ was generated. Then the specific backgrounds $B_s \rightarrow \mu^+\mu^-\gamma$ (2 events), $B_c^\pm \rightarrow J/\psi(\mu\mu) \mu^\pm v$ (49 events) and misidentified $B \rightarrow h^+h^-$ (12 events) are added from the full simulation samples. The GL of several $b\bar{b} \rightarrow \mu^+\mu^-X$ events had to be modified to properly match the expected number of events in $GL > 0.5$. After the modification, 34 $b\bar{b} \rightarrow \mu^+\mu^-X$ events fall in $GL > 0.5$ (31 expected).

Finally, 20 signal events were added, which correspond to a $BR(B_s \rightarrow \mu^+\mu^-) = 2 \cdot 10^{-8}$.

5.3. Calibration

5.3.1. Invariant mass and $B \rightarrow h^+ h^-$ exclusive fractions

The invariant mass is calibrated as explained in 4.4.3.1. The signal (cristalball), physical and combinatorial background (see 5.2.2) and Lambdas (described by a simple Gaussian) are fitted for different values of PID cut. In the KK mass hypothesis, a plateau is expected once the PID cut is tight enough to make the $B_s \rightarrow K^+ K^-$ to be the dominant mode. The $B_s \rightarrow K^+ K^-$ gives us the mean of the B_s peak as well as the resolution. But this last one is expected to be the same that in the B_d peak, so we will take advantage of the statistics in the B_d modes to improve the knowledge of the invariant mass resolution. Also, the centre of the B_d peak is needed in order to perform the search for $B_d \rightarrow \mu^+ \mu^-$. Hence, all the four particle combinations are studied; see for instance Figure 5-4, Figure 5-5, Figure 5-6 and Figure 5-7.

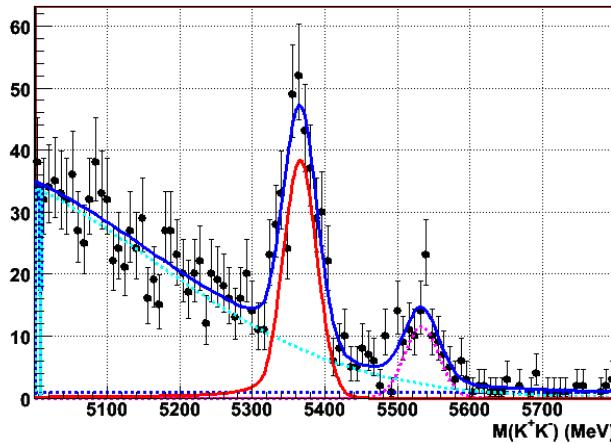


Figure 5-4: $K^+ K^-$ invariant mass distribution for $DLL(K-\pi) > 25$ and $GL > 0.5$.

The data corresponds to the xTOS sample in order to profit from the larger statistics. The central peak is dominated by true $B_s \rightarrow K^+ K^-$. The peak on the right is populated by the Lambdas ($\Lambda_b \rightarrow pK$ should have also large values of $DLL(K-\pi)$).

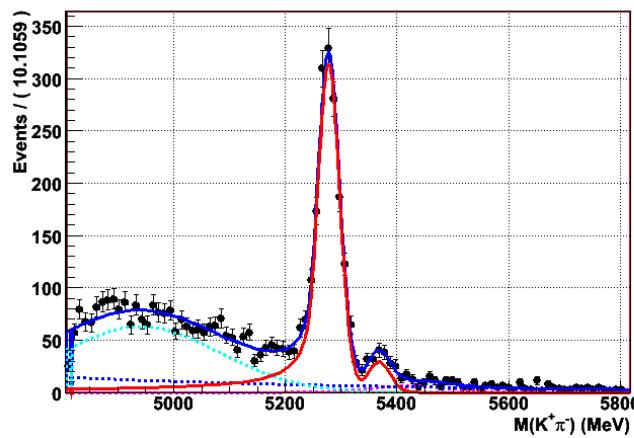


Figure 5-5: $K^+ \pi^-$ invariant mass distribution for $DLL(K-\pi) > 20$ in K^+ , $DLL(K-\pi) < -20$ in π^- and $GL > 0.5$.

The data corresponds to the xTOS sample in order to profit from the larger statistics. The central peak is dominated by true $B_d \rightarrow K^+ \pi^-$. The peak on the right is populated by $B_s \rightarrow K^+ \pi^-$. In this case Lambdas ($\Lambda_b \rightarrow p\pi$) contribution are smaller because of the smaller BR of $\Lambda_b \rightarrow p\pi$ with respect to $\Lambda_b \rightarrow pK$ and (mainly) because the antiparticle does not contribute as the pion ID cut will be applied on the proton.

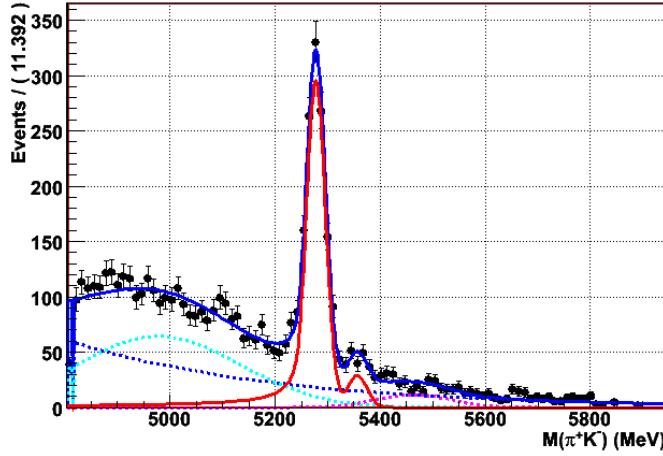


Figure 5-6: π^+K^- invariant mass distribution for $DLL(K-\pi) > 20$ in K^- , $DLL(K-\pi) < -20$ in π^+ and $GL > 0.5$.

The data corresponds to the xTOS sample in order to profit from the larger statistics. The central peak is dominated by true $B_d \rightarrow \pi^+K^-$. The peak on the right is populated by $B_s \rightarrow \pi^+K^-$. In this case Lambdas ($\Lambda_b \rightarrow p\pi$) contribution are smaller because of the smaller BR of $\Lambda_b \rightarrow p\pi$ with respect to $\Lambda_b \rightarrow pK$ and (mainly) because the antiparticle does not contribute as the pion ID cut will be applied on the proton.

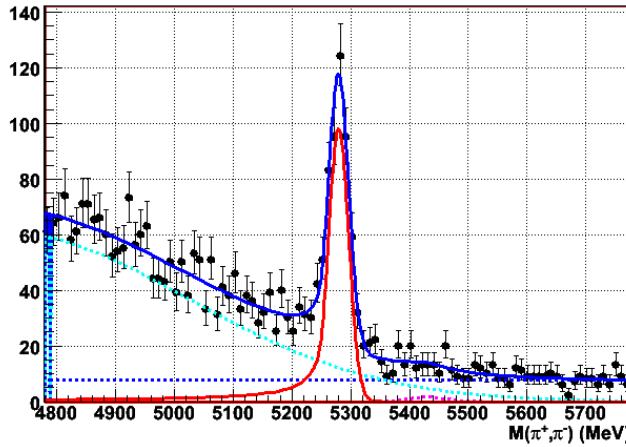


Figure 5-7: $\pi^+\pi^-$ invariant mass distribution for $DLL(K-\pi) < -30$ in and $GL > 0.5$.

The data corresponds to the xTOS sample in order to profit from the larger statistics. The central peak is dominated by true $B_d \rightarrow \pi^+\pi^-$.

5.3.1.1. Mean and sigma

The main parameters needed for the description of $B_s \rightarrow \mu^+\mu^-$ mass line shape are the mean and the sigma of the cristalball. The transition point has to be extracted indirectly as the radiative tail of $B_s \rightarrow \mu^+\mu^-$ is different than any of the $B \rightarrow h^+h^-$, and will be computed in 5.3.1.1. The measured values for the mean and sigma are shown in Figure 5-8 and Figure 5-9.

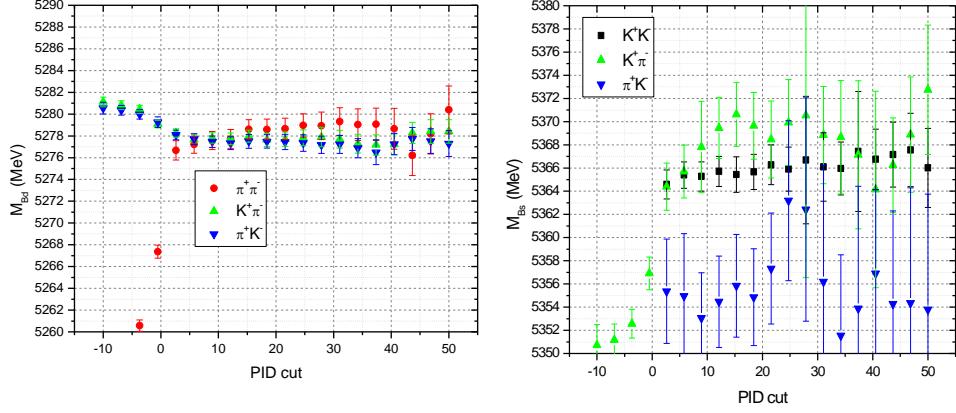


Figure 5-8: Measured positions of the B mass peaks.

Black squares: tightening the DLL for high purity on K^+K^- . Red dots: for high purity in $\pi^+\pi^-$. Green triangles: for high purity in $K^+\pi^-$. Blue inverse triangles: for high purity in π^+K^- .

The mean is calculated via weighted average over all the fits in $B_s \rightarrow K^+K^-$ for $PID > 20$ to get the final number. The error on the mean is averaged in the same way, i.e., it is not reduced by the number of measurements as all the points are highly correlated. We could, in principle, combine the measurement with those obtained from $K^+\pi^-$ and π^+K^- . But those one come from the decay $B_s \rightarrow K\pi$, with lower statistics and hence larger errors. Moreover, it is not clear, looking at Figure 5-8, whether high purity has been reached for $B_s \rightarrow \pi^+K^-$ and it might be introducing a systematic uncertainty. Therefore, the measured position of the B_s peak that will be used in the search of $B_s \rightarrow \mu^+\mu^-$ is:

$$M_{B_s \rightarrow \mu\mu} = M_{B_s \rightarrow KK} = 5366.4 \pm 2.5 \text{ MeV}$$

In the case of the B_d , the three final states ($B_d \rightarrow \pi^+K^-$, $B_d \rightarrow K^+\pi^-$, and $B_d \rightarrow \pi^+\pi^-$) can be combined, obtaining:

$$\left. \begin{array}{l} M_{B_d \rightarrow K^+\pi^-} = 5277.75 \pm 0.86 \text{ MeV} \\ M_{B_d \rightarrow \pi^+K^-} = 5277.25 \pm 0.89 \text{ MeV} \\ M_{B_d \rightarrow \pi^+\pi^-} = 5278.8 \pm 1.4 \text{ MeV} \end{array} \right\} M_{B_d \rightarrow \mu^+\mu^-} = 5277.73 \pm 0.31 \text{ MeV}$$

The nominal B_d mass is 5279 MeV which is incompatible by several sigmas with what we obtained. But as we see in 3.2.2 (see also Figure 3-20), DC06 reconstruction produce a bias of -1.3 MeV which is exactly what we have measured.

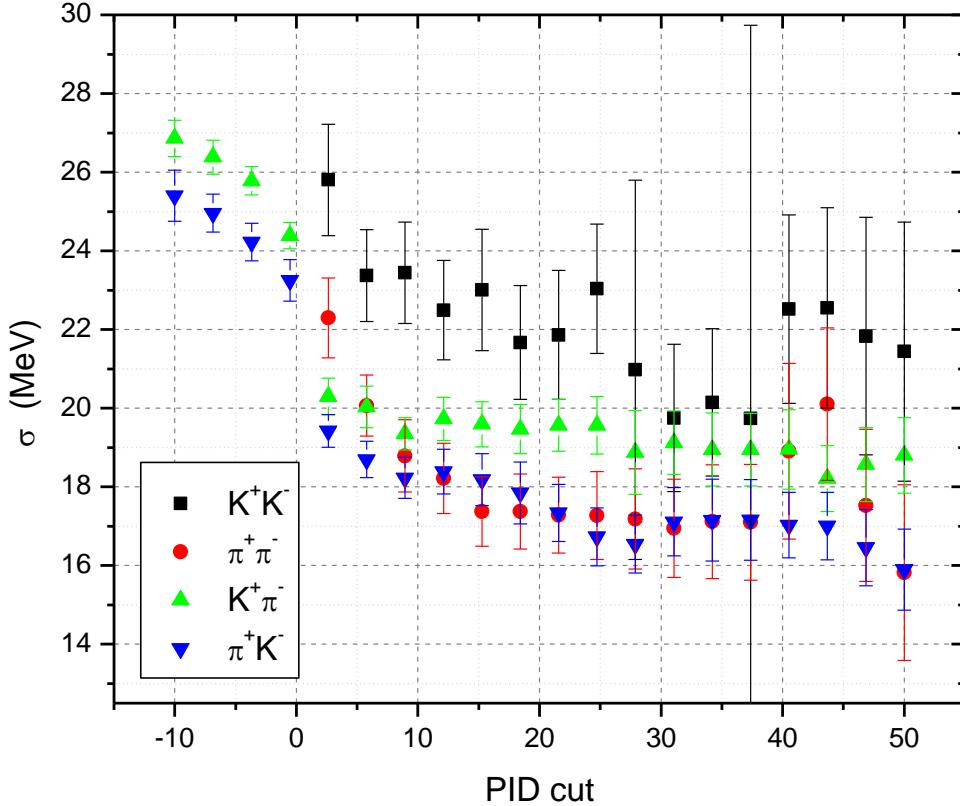


Figure 5-9: Measured mass resolution for different PID cuts.

Black squares: tightening the DLL for high purity on K^+K^- . Red dots: for high purity in $\pi^+\pi^-$. Green triangles: for high purity in $K^+\pi^-$. Blue inverse triangles: for high purity in π^+K^- .

The inclusive sample is fitted as well for several values of the absolute PID cuts up to the high purity limit (i.e., all well identified kaons and pions). The ratio of the widths provides us the correction factor for the invariant mass resolution; and the number of inclusive $B \rightarrow h^+h^-$ events, together with the number of exclusive events for different PID cuts gives us the relative fraction of each one of the decays in the $B \rightarrow h^+h^-$ sample (see 4.4.1.2). The $\pi\pi$ mass hypothesis is used in this case, and the parameterization of the signal line shape is composed of two double cristalballs:

$$\rho(m) \propto \begin{cases} e^{-\frac{1}{2}\left(\frac{m-M}{\sigma}\right)^2}, & \text{if } b \geq \left(\frac{m-M}{\sigma}\right) \geq -a \\ \frac{1}{\frac{1}{a} - \left(a + \frac{m-M}{\sigma}\right)}, & \text{if } \left(\frac{m-M}{\sigma}\right) < -a \\ \frac{1}{\left[\frac{1}{b} - \left(b + \frac{m-M}{\sigma}\right)\right]^n}, & \text{if } \left(\frac{m-M}{\sigma}\right) > b \end{cases}$$

The parameters a , b , σ and n are the same for both cristalballs, only the mean is different. The fitting of the tail parameters (a , b and n) is not easy in the presence of large amount of background. To solve this, we extract the related parameters from the xTOS sample with very high GL (> 0.85 in this case), and we will keep them fixed to those values in any other fit of the inclusive $B \rightarrow h^+h^-$ sample.

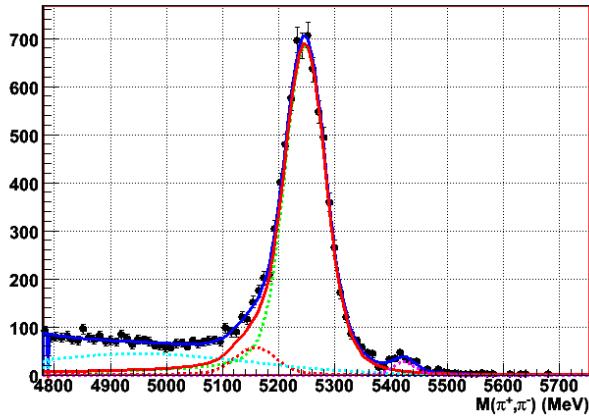


Figure 5-10: $\pi^+\pi^-$ invariant mass distribution for GL > 0.85 and no PID cuts.

The measured tail parameters are:

Parameter	Fitted Value
A	1.210 ± 0.073
B	1.51 ± 0.12
N	5.1 ± 1.3

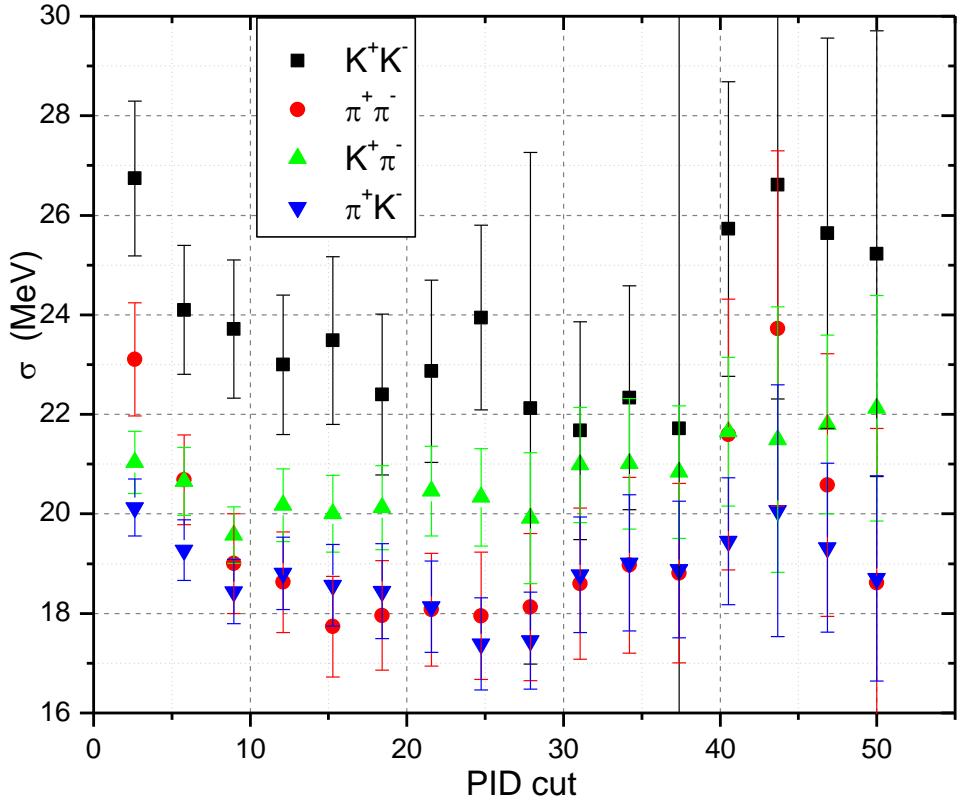


Figure 5-11: Measured mass resolution for different PID cuts after correction of momentum bias.

Black squares: tightening the DLL for high purity on K^+K^- . Red dots: for high purity in $\pi^+\pi^-$. Green triangles: for high purity in $K^+\pi^-$. Blue inverse triangles: for high purity in π^+K^- . The errors include the propagation of the statistical error from the correction factor.

We will assume that the variation of signal invariant mass distribution with the Geometrical Likelihood is negligible. To check this assumption, we compare the resolution measured in $K^+\pi^-$ hypothesis for xTOS in $0.5 < GL < 0.75$ and in $GL > 0.75$. The results of this test are shown in Table 5-I. In both cases a $DLL(K^-\pi) > 20$ for the positive particle and a $DLL(K^-\pi) < -20$ for the negative was applied to ensure purity of the $K^+\pi^-$ mode.

Table 5-I: Invariant mass resolution in different GL bins.

Sample	Mean(B_d) (MeV)	Resolution (MeV)
xTOS, $0.5 < GL < 0.75$	5279.31 ± 0.38	20.5 ± 1.0
xTOS , $GL > 0.75$	5277.24 ± 0.81	19.17 ± 0.71

In order to extract the sigma of the cristalball we are going to use in the final analysis, we proceed in a similar way to what was done for the mean, i.e, performing a weighted average of the σ and the error on the σ for the plateau of each decay mode (see Figure 5-11), and a weighted combination of what was obtained in each decay. The fact that error coming from the correction factor is the same for all the decays has been taken into account. The obtained resolution is:

$$\left. \begin{array}{l} \sigma_{B \rightarrow K^+ \pi^-} = 20.8 \pm 1.3 \text{ MeV} \\ \sigma_{B \rightarrow \pi^+ K^-} = 18.4 \pm 1.2 \text{ MeV} \\ \sigma_{B_d \rightarrow \pi^+ \pi^-} = 18.7 \pm 1.6 \text{ MeV} \\ \sigma_{B_s \rightarrow K^+ K^-} = 23.4 \pm 2.5 \text{ MeV} \end{array} \right\} \sigma_{B \rightarrow \mu^+ \mu^-} = 20 \pm 1 \text{ MeV}$$

In principle, the invariant mass description could be improved by adding a second resolution to the signal model. However, when trying to fit these two components to the $B_s \rightarrow K^+ K^-$ data we do not get two clearly different resolutions (see Figure 5-12), so for this example we will keep the simpler model of a single cristalball.

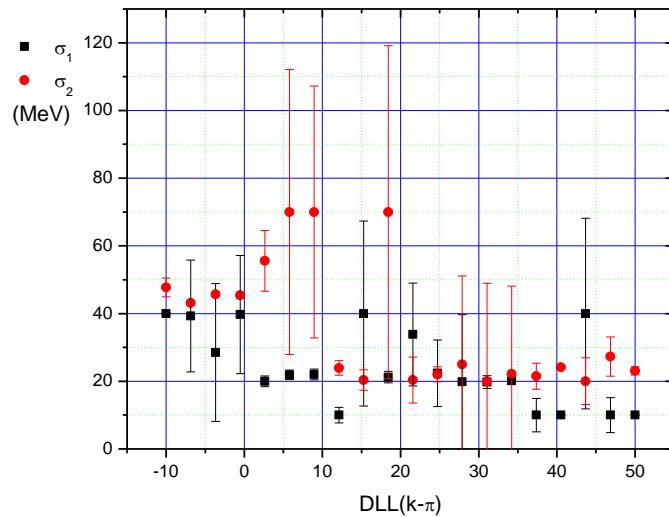


Figure 5-12: Result of fitting a double Gaussian to $B_s \rightarrow K^+ K^-$ invariant mass distribution.

The two resolutions, σ_1 and σ_2 , are shown as a function of the cut applied on the DLL of the daughters.

5.3.1.2. Transition point of the cristalball

The transition point of the cristalball sets the change from Gaussian distribution to exponential tail, in sigmas. In order to know which the actual value is that corresponds to $B_s \rightarrow \mu^+ \mu^-$ for a resolution of ~ 20 MeV, MC truth events (generated with PHOTOS) were smeared for different values of the resolution and fitted to a cristalball. The dependence of the transition point with the resolution is plotted in Figure 5-13:

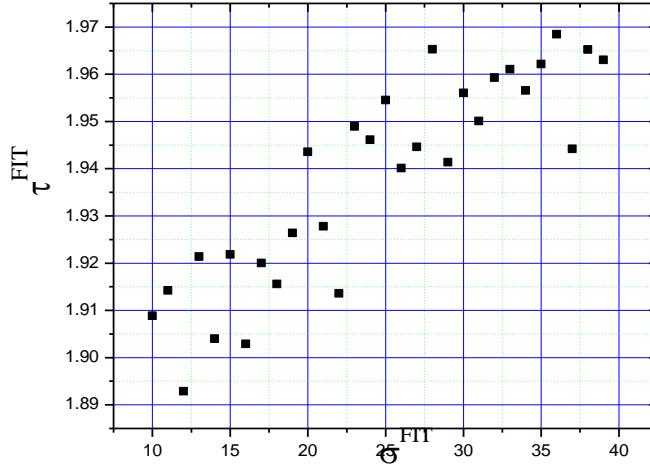


Figure 5-13: Dependence of the transition point with the resolution.

x axis: resolution, y axis: transition point.

We can see from Figure 5-13 a linear dependence between both parameters. However, such dependence is small so we will just assume a safe margin for the error:

$$a_{B \rightarrow \mu^+ \mu^-} = 1.9 \pm 0.1$$

5.3.1.3. Measurement of the different $B \rightarrow h^+ h^-$ fractions

As explained in 4.4.1.2, in order to measure the fraction of the different components we can take the advantage of the fact that the ratio of crosssections $B_d \rightarrow \pi^+ \pi^- / B_d \rightarrow K^+ \pi^-$ is very well known and that the acceptance, reconstruction, selection and trigger efficiencies of all the $B \rightarrow h^+ h^-$ channels should be the same to good approximation. The ratio $BR(B_d \rightarrow \pi^+ \pi^-)/BR(B_d \rightarrow K^+ \pi^-)$ is

$$\frac{BR(B_d \rightarrow K\pi)}{BR(B_d \rightarrow \pi\pi)} = 0.264 \pm 0.015$$

But taking into account the way in which the sample used here was generated; we must use $BR(B_d \rightarrow \pi^+ \pi^-)/BR(B_d \rightarrow K^+ \pi^-) = 0.25$ and assign the equivalent relative error, so:

$$\left. \frac{BR(B_d \rightarrow K\pi)}{BR(B_d \rightarrow \pi\pi)} \right|_{examp.} = 0.250 \pm 0.014$$

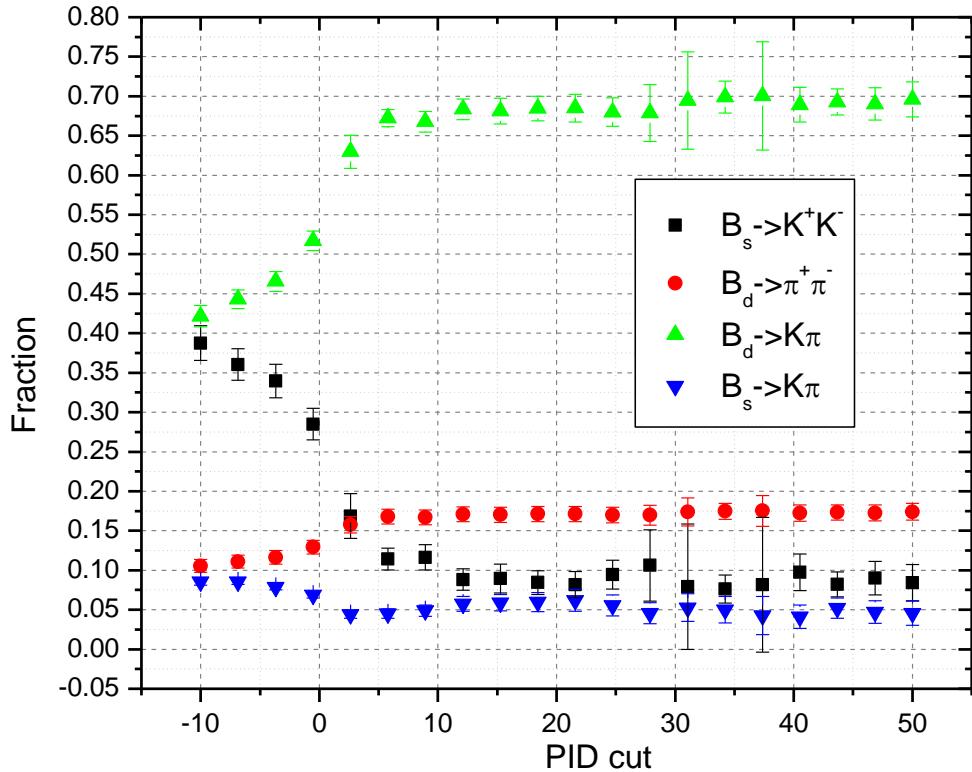


Figure 5-14: Measured $B \rightarrow h^+ h^-$ fractions as a function of PID cut.

Black squares: $B_s \rightarrow K^+ K^-$. Red dots: $B_d \rightarrow \pi^+ \pi^-$. Green triangles: $B_d \rightarrow K^+ \pi^- + B_d \rightarrow \pi^+ K^-$. Blue inverse triangles: $B_s \rightarrow K^+ \pi^- + B_s \rightarrow \pi^+ K^-$.

The computed fractions as a function of the PID cut is shown in Figure 5-14. The measured fractions are then:

$$\begin{aligned} f_{B_s \rightarrow KK} &= 0.086 \pm 0.020 \\ f_{B_d \rightarrow K\pi} &= 0.690 \pm 0.021 \\ f_{B_d \rightarrow \pi\pi} &= 0.172 \pm 0.011 \\ f_{B_s \rightarrow K\pi} &= 0.050 \pm 0.015 \end{aligned}$$

Which are in good agreement with the ones used in the generation of the sample (see 5.2.2):

$$\begin{aligned} f_{B_s \rightarrow KK} &= 0.100 \\ f_{B_d \rightarrow K\pi} &= 0.684 \\ f_{B_d \rightarrow \pi\pi} &= 0.174 \\ f_{B_s \rightarrow K\pi} &= 0.042 \end{aligned}$$

5.3.2. Geometrical likelihood

5.3.2.1. Computation of $B \rightarrow hh$ trigger unbiased GL distribution

Here we will use the standard Geometrical Likelihood, without modifying the definition of the isolation. To perform a good calibration, $B \rightarrow h^+h^-$ TIS events are used with an event-by-event weight accordingly to the dependence of the TIS efficiency with the p_t of the B and the number of long tracks in the event:

$$w(p_t, n^{long}) \propto 1/(\varepsilon^{TIS}(p_t) \cdot \varepsilon'^{TIS}(n^{long}))$$

The weights are normalized such as the measured amount of signal in the full GL range remains the same with and without the use of those weights. This normalization will imply some iterative process.

The TIS efficiencies can be measured from data as explained in 4.4.2 and 4.4.3.2. In principle they can be extracted from any B decay selected in any way, but here for simplicity we will use the $B \rightarrow hh$ itself, in particular those events in $GL > 0.5$. The measured dependence of the TIS probability with the p_t of the B meson and the multiplicity is illustrated in Figure 5-15. There we can see the agreement with the dependences we would get with larger statistics and in absence of background (compare with Figure 4-24).

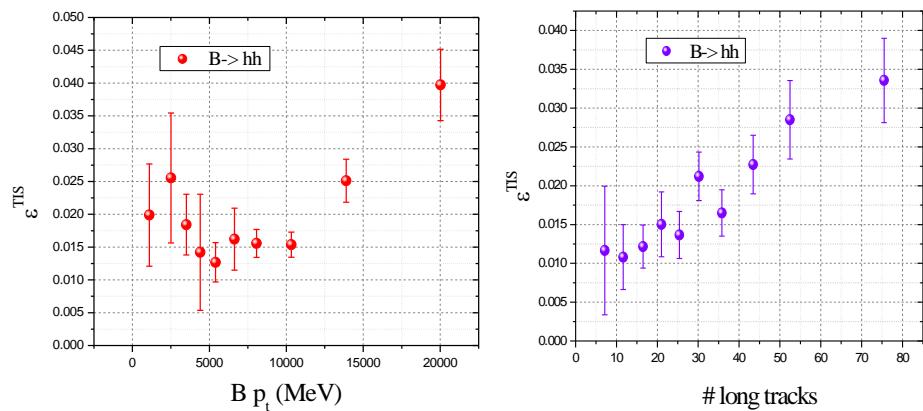


Figure 5-15: Measured dependence of TIS efficiency with $B p_t$ and multiplicity.

Then, the invariant mass spectrum is fitted, using those weights, in the different bins of the GL to compute its distribution. The measured numbers are

Table 5-II: Measured number of $B \rightarrow hh$ events in each GL bin.

	0.0-0.25	0.25-0.5	0.5-0.75	0.75-1
# No weights	820 ± 430	438 ± 52	347 ± 33	395 ± 25
# weighted	800 ± 320	411 ± 55	340 ± 35	434 ± 27

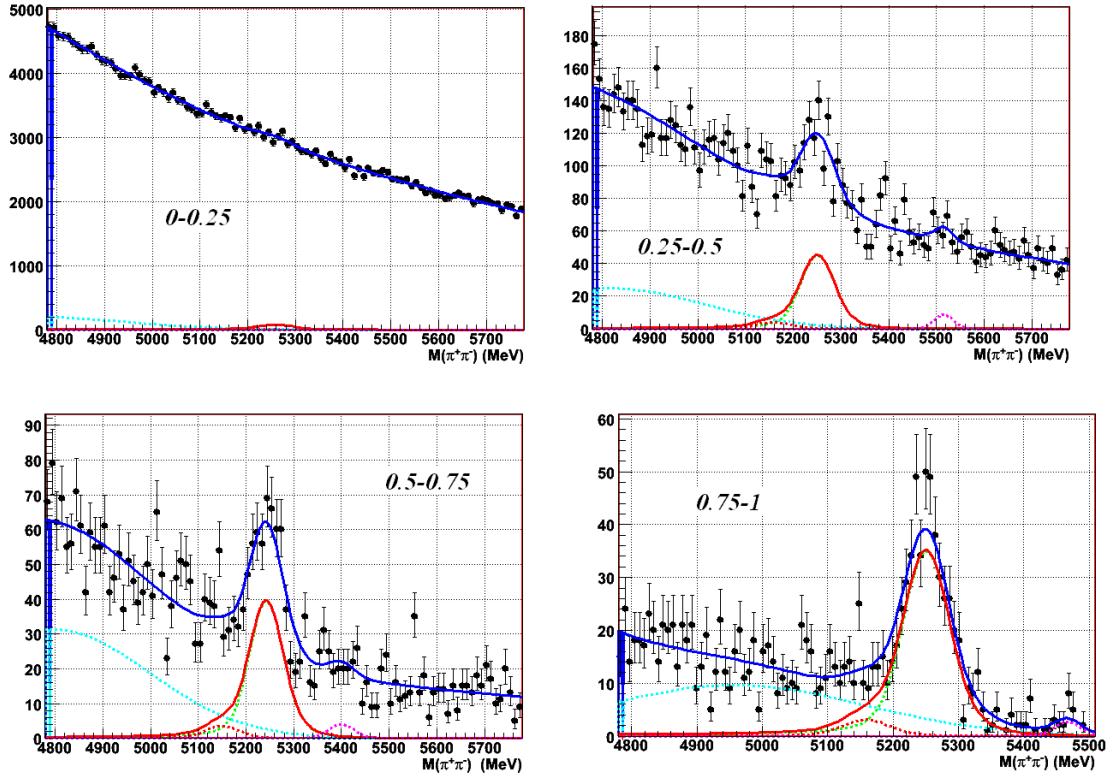


Figure 5-16: Mass distribution of the $B \rightarrow hh$ candidates in each GL bin.

If we look, for instance, at the first of the bins, we see the signal is overwhelmed by the background (see Figure 5-16). This implies that the low GL region will not be calibrated with accuracy. In order to deal with this, we will use a tighter common selection in the next section.

5.3.2.2. Use of a tighter selection

This is not a problem for normalization to $B_d \rightarrow K^+ \pi^-$ as we can just cut on $GL > 0.5$ for both signal and normalization channel that the corresponding efficiencies will cancel in the ratio. But for normalization to $B^+ \rightarrow J/\psi(\mu\mu) K^+$ it is not so straight forward. A similar approach is to apply tighter cuts on the common selection so that the $B \rightarrow h^+ h^-$ gets observable even for low values of the GL. For this analysis example we will use:

Table 5-III: Tight selection for the analysis example.

Min IPS (h^+, h^-) / $K^+ \text{IPS} / K^* \text{IPS} > 9$
DOFS ($B, J/\psi$) > 25
$B \text{IP} < 40 \mu\text{m}$, $B \text{IPS} < 3$
DOCA ($B, J/\psi$) < 60 μm

After including this set of cuts, we measure the following numbers per bin:

Table 5-IV: Measured number of $B \rightarrow hh$ events in each GL bin after tight selection.

	0.0-0.25	0.25-0.5	0.5-0.75	0.75-1
# No weights	152±52	276±28	292±24	384±24
# weighted	113±46	260±29	293±26	439±26

The tight selection will remove $B_s \rightarrow \mu^+ \mu^-$ events in the sensitive region and hence cause a lack of sensitivity. In order to avoid this, we can cut on the GL even for normalization $B^+ \rightarrow J/\psi(\mu\mu) K^+$, and just add the corresponding ratio of selection efficiencies, referred to the tight selection:

$$\frac{\varepsilon^{GL}}{\varepsilon^T} = \frac{N^{GL}}{N^T} = \frac{N^{GL\&T} + N^{GL\¬T}}{N^{GL\&T} + N^{T\¬GL}}$$

Where GL superscript refers to events that pass a given GL cut and the T superscript refers to those that pass the tight selection.

Table 5-V: Measured number of $B \rightarrow hh$ events in each GL bin not surviving the tight selection.

	0.0-0.25	0.25-0.5	0.5-0.75	0.75-1
# No weights	720±290	167±44	54±21	14.4±6.0
# weighted	637±468	183±47	72±25	23.4±8.3

Hence for a cut $GL > 0.5$:

$$N^{GL\&T} = (293 \pm 26) + (439 \pm 26) = 732 \pm 37$$

$$N^{GL\¬T} = (72 \pm 25) + (23.4 \pm 8.3) = 95 \pm 26$$

$$N^{T\¬GL} = (113 \pm 46) + (260 \pm 29) = 373 \pm 53$$

$$\frac{\varepsilon^{GL,05}}{\varepsilon^T} = 0.748 \pm 0.044 \Leftrightarrow \frac{\varepsilon^T}{\varepsilon^{GL,05}} = 1.336 \pm 0.078$$

Using $GL > 0.5$ as a selection, the expected signal fractions in the remaining two bins are:

Table 5-VI: Signal fraction in the sensitive bins.

	0.5-0.75	0.75-1
f	0.439±0.030	0.561±0.030

5.3.3. Background level

The background level is estimated using B_s sidebands, chosen to be the region outside 5200-5430 MeV to avoid contamination from possible two body decays with double misid (see Figure 5-17).

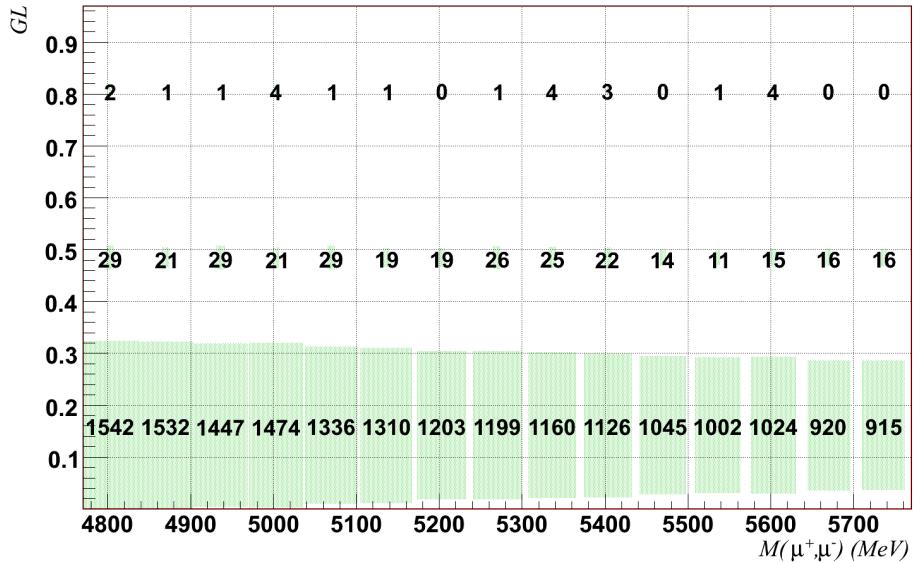


Figure 5-17: Invariant mass versus GL for the selected $B_s \rightarrow \mu^+ \mu^-$ candidates.

For this example, the background will be interpolated independently in each GL bin, without assuming that the mass shape is the same in all bins. As we see no strong evidence of missid in the B_d peak, this contribution is assumed to be negligible in the signal region, and the PID likelihood axis will not be used.

The background is fitted to an exponential e^{-KM} in the sideband region, giving:

Table 5-VII: Parameters for background interpolation.

GL	K (MeV ⁻¹)	N in sidebands
0-0.25	$(5.90 \pm 0.20) \cdot 10^{-4}$	13710
0.25-0.5	$(7.4 \pm 1.5) \cdot 10^{-4}$	404
0.5-0.75	$(0.0 \pm 6.6) \cdot 10^{-4}$	17
0.75-1.0	$(0.0 \pm 8.6) \cdot 10^{-4}$	8

Then, the number of events in a given bin of mass [x,y] is:

$$N = \frac{N_{sb} (e^{-Kx} - e^{-Ky})}{(e^{-4769.6 \cdot K} + e^{-5450 \cdot K} - e^{-5969.6 \cdot K} - e^{-5200 \cdot K})}$$

Where x,y are the limits of the bin and the numbers in the denominator are the limits of the sideband regions.

Hence, the systematic for the background estimation in the mass bins for a given GL will be determined by two numbers: the amount of background in the sidebands and the parameter that describes the line shape. Variations of those two numbers cause of course correlated variations for all the mass bins in the same GL.

The number of expected background events in the search mass window [5309.6-5429.6] above GL = 0.5 are, using above formula:

$$0.5-0.75 : 2.15^{+0.64}_{-0.52}; \quad 0.75-1.0 : 1.01^{+0.49}_{-0.38}$$

5.4.Normalization

This section shows the normalization to $B^+ \rightarrow J/\psi(\mu\mu) K^+$, discussing the relevant fractions which enter in the normalization expression:

$$BR = \frac{BR_{cal} \cdot \mathcal{E}_{cal}^{REC} \mathcal{E}_{cal}^{SEL/REC} \mathcal{E}_{cal}^{TRIG/SEL}}{\mathcal{E}_{sig}^{REC} \mathcal{E}_{sig}^{SEL/REC} \mathcal{E}_{sig}^{TRIG/SEL}} \cdot \frac{f_{cal}}{f_{Bs}} \cdot \frac{N_{sig}}{N_{cal}}$$

and decomposing it into independent quantities to estimate the overall uncertainty of the normalization factor. The ratio of hadronization fractions is, to the current knowledge [7]:

$$\frac{f_d}{f_s} = \frac{f_+}{f_s} = 3.52 \pm 0.48$$

But for consistency with the hadronization fractions used in simulation, we will use in this example:

$$\left. \frac{f_d}{f_s} \right|_{exa.} = \left. \frac{f_+}{f_s} \right|_{exa.} = 4.00 \pm 0.55$$

5.4.1. Offline efficiencies

The ratio of offline efficiencies is computed as:

$$\frac{\mathcal{E}_+^{REC} \times \mathcal{E}_+^{SEL/REC}}{\mathcal{E}_{sig}^{REC} \times \mathcal{E}_{sig}^{SEL/REC}} = \left(\frac{\mathcal{E}_+^{REC} \times \mathcal{E}_+^{SEL-CT/REC}}{\mathcal{E}_{sig}^{REC} \times \mathcal{E}_{sig}^{SEL-CT/REC}} \right) \times \mathcal{E}_+^{M_{J/\psi}} \times \frac{1}{\mathcal{E}_s^{M_{Bs}}} \approx \left(\frac{\mathcal{E}_d^{REC} \times \mathcal{E}_d^{SEL-CT'/REC}}{\mathcal{E}_+^{REC} \times \mathcal{E}_+^{SEL-CT'/REC}} \right) \times \mathcal{E}_+^{M_{J/\psi}} \times \frac{1}{\mathcal{E}_s^{M_{Bs}}} = \\ \left(\frac{BR(B^+ \rightarrow J/\psi K^+) \times N^{Bd}}{BR(B_d \rightarrow J/\psi K^*) \times N^{B+}} \times \frac{1}{\mathcal{E}_d^{M_{K^*}} \mathcal{E}_d^{K,\pi_{IPS}}} \right) \times \mathcal{E}_+^{M_{J/\psi}} \times \frac{1}{\mathcal{E}_s^{M_{Bs}}}$$

Where CT stands from “common cuts” and means the efficiency of all the cuts of the common $B_s \rightarrow \mu^+ \mu^- / B^+ \rightarrow J/\psi(\mu\mu) K^+$ selection including the IPS of the kaon/muon. CT' refers to the common selection for $B^+ \rightarrow J/\psi(\mu\mu) K^+ / B_d \rightarrow J/\psi(\mu\mu) K^0 (\bar{K}\pi)$ based on 25 sigma detached J/ψ . $\mathcal{E}_d^{M_{K^*}}$, the efficiency related to K^* mass window, is taken from a simple degradation of the Breit Wigner for a wide range of values of single Gaussian resolution, as it has been shown to be robust against ignorance of the actual resolutions. From Figure 3-28 we assign $67 \pm 2\%$ to this efficiency. The very tight cut on the distance of flight of the J/ψ was on purpose to make the IPS cut of K and π to be close to 100% efficient. In this example we will assume a safe margin of $98 \pm 2\%$ where anyway that ‘systematic’ is small compared with the ratio of hadronization fractions.

The factor $\mathcal{E}_s^{M_{Bs}}$ accounts for the fact that only events in the 60 MeV mass window around B_s peak are used in the 3D parameter space. As the window is large enough compared with the resolution, the events falling out of it are due to the radiated photon energy or mismatch between the center of the B_s mass peak and the center of the search window.

The final state radiation is a well known physical process and we can perform a first approximation using analytical formulae. In the B_s rest frame:

$$M_{Bs} = E_{Bs} = E_{\mu+} + E_{\mu-} + E_\gamma$$

Now, as the photon energy is, in general, small, we can the B_s rest frame and $\mu^+ \mu^-$ rest frame should be almost equivalent and hence $E_{\mu+} + E_{\mu-} \approx E_{\mu+} + E_{\mu-}|_{\mu\mu \text{ rest fr.}} = M_{\mu\mu}$. Therefore, at first order we can

suppose $M_{\mu\mu} \approx M_{Bs} - E_\gamma$ and thus the corresponding efficiency is just the fraction of events with photon energy (at B_s rest frame) in the range 60-600 MeV (note that a 600 MeV window was already applied in the stripping). Indeed, using [17] we get $\varepsilon_s^{M_{Bs}} = 93\%$. To improve this estimation, MC truth events (generated with PHOTOS) were smeared according to the observed resolution at $B \rightarrow hh$ (see 5.3.1.1). For a single resolution of 20-25 MeV we see no effect and the efficiency keeps being 93%.

To evaluate the effect of the mismatch of the peak center, we also displace the smeared events accordingly to the observed value of the mean (see 5.3.1.1) $M_{Bs \rightarrow \mu\mu} = 5366.4 \pm 2.5$ MeV. This effect was found to be negligible, providing again a 93 % efficiency. So, without taking into account the second Gaussian resolution we obtain $\varepsilon_s^{M_{Bs}} = 93\%$ with small variations < 1%. This is very similar to what we would obtain using MC information, 91% (although not very compatible).

The same exercise was done for the J/ψ (using $\sigma_{J/\psi} = 13 \pm 1$ MeV) as the ratio of both efficiencies is the actual number appearing in the fraction of offline efficiencies. The obtained efficiency was $\varepsilon_s^{M_{J/\psi}} = 96\%$. Residual effects such as biases in the radiative tail due to other cuts in the selection are not considered for this example. Hence:

$$\frac{\varepsilon_s^{M_{J/\psi}}}{\varepsilon_s^{M_{Bs}}} = 1.03 \pm \sim 0.02$$

In order to measure the number of $B^+ \rightarrow J/\psi(\mu\mu) K^+$ or $B_d \rightarrow J/\psi(\mu\mu) K^{*0}(K\pi)$ events, fits of the mass distribution are performed. The background is described with an exponential function and the signal is parameterized with a sum of a Gaussian and a cristalball. The mean of the Gaussian and the cristalball is forced to be the same, and thus the signal description has four free parameters: the mean, the two sigmas and the transition point of the cristalball.

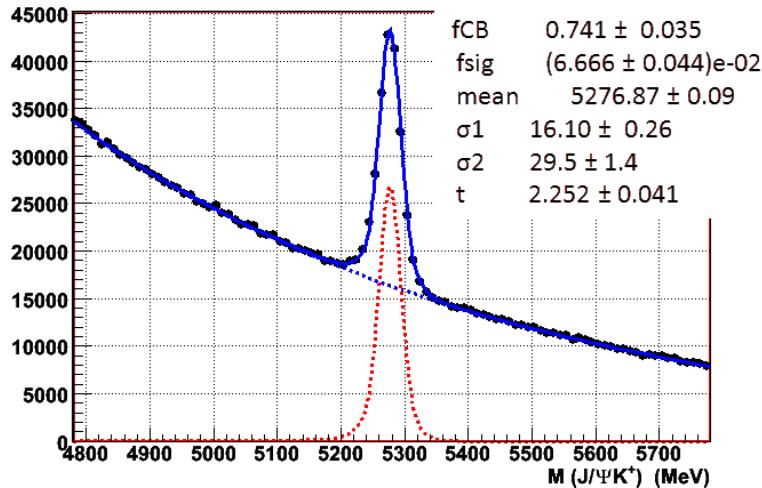


Figure 5-18: Invariant mass of $B^+ \rightarrow J/\psi(\mu\mu) K^+$ after trigger and stripping selection.

The result of the fit to the signal parameters is shown in the top right corner: f_{CB} is the fraction of the cristalball in the signal model, f_{sig} is the fraction of the signal in the total sample, $mean$ the mean of both Gaussian and cristalball contribution, σ_1 is the resolution for the cristalball, σ_2 is the resolution of the Gaussian contribution and t the transition point of the cristalball.

The number of $B^+ \rightarrow J/\psi(\mu\mu) K^+$ observed after trigger and stripping is 127059 ± 829 (see Figure 5-18).

As we need to compute the ratio of offline efficiencies, those events passing the selection that is common with the $B_d \rightarrow J/\psi(\mu\mu) K^{*0}(K\pi)$ have to be measured.

The number of $B^+ \rightarrow J/\psi(\mu\mu) K^+$ and $B_d \rightarrow J/\psi(\mu\mu) K^{*0}(K\pi)$ with the $DOFS(J/\psi) > 25$ cut are:

$$N_{Bd}^{CT'} = 37134 \pm 695, N_{B+}^{CT'} = 103348 \pm 674 \rightarrow \frac{N_{Bd}^{CT'}}{N_{B+}^{CT'}} = 0.3593 \pm 0.0071$$

(5:1)

But, as explained in 0 a tighter selection (Table 5-III) is needed in order to perform a proper calibration of the Geometrical likelihood. The ratio of N_{Bd}/N_{B+} accounts mainly for the reconstruction of the extra tracks, but it will also reflect the small differences in selection efficiencies, so here we compute it applying the tight cuts mentioned in 0.

$$N_{Bd}^T = 26996 \pm 308, N_{B+}^T = 78352 \pm 375 \rightarrow \frac{N_{Bd}^T}{N_{B+}^T} = 0.3442 \pm 0.0043$$

(5:2)

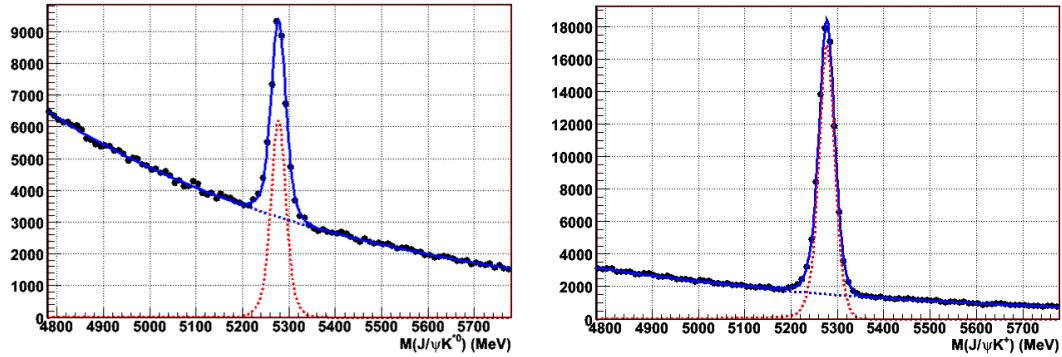


Figure 5-19: B_d and B^+ mass peaks after the tight selection shown in Table 5-III.

The difference of (5:2) with respect to (5:1) reflects the differences of the efficiencies of the common selections when applied to 3 bodies with respect to 4 bodies. Here we will assume that it is representative of the deviations from one of this ratio of efficiencies, and that similar deviations are expected in the ratio of selection efficiencies in 2 to 3 body ratio. This difference is hence added as a systematic error. This systematic ($\sim 5\%$) is already significantly larger than the statistical error on this ratio; therefore after adding both in quadrature the combined error is uncorrelated with the other sources in the normalization expression to a very good approximation.

$$\begin{aligned} \frac{\mathcal{E}_+^{REC} \times \mathcal{E}_+^{SEL/REC}}{\mathcal{E}_{sig}^{REC} \times \mathcal{E}_{sig}^{SEL/REC}} &= \left(\frac{BR(B^+ \rightarrow J/\psi K^+) \times N^{Bd}}{BR(B_d \rightarrow J/\psi K^*) \times N^{B+}} \times \frac{1}{\mathcal{E}_d^{M_{K^*}} \mathcal{E}_d^{K, \pi_{PS}}} \right) \times \frac{\mathcal{E}_+^{M_{J/\psi}}}{\mathcal{E}_s^{M_{Bs}}} = \\ &= \left(\frac{BR(B^+ \rightarrow J/\psi K^+)}{BR(B_d \rightarrow J/\psi K^*)} \times \frac{0.34 \pm 0.02}{(0.67 \pm 0.02)(0.98 \pm 0.02)} \right) \times \frac{\mathcal{E}_+^{M_{J/\psi}}}{\mathcal{E}_s^{M_{Bs}}} = \\ &= ((1.135 \pm 0.065) \times (0.518 \pm 0.036)) \times 1.03 \pm 0.02 = 0.606 \pm 0.055 \end{aligned}$$

Finally, as we preferred to replace, for $B_s \rightarrow \mu^+ \mu^-$, the tight selection by a cut in $GL > 0.5$, we have to multiply by the factor (see 0):

$$\frac{\mathcal{E}^T}{\mathcal{E}^{GL,05}} = 1.336 \pm 0.078$$

5.4.2. Trigger efficiencies

We will first compute the trigger efficiency for $B^+ \rightarrow J/\psi(\mu\mu) K^+$. For would be offline selected events:

$$\varepsilon^{TRIG} \equiv \frac{N^{TRIG}}{N} = \varepsilon^{TIS} \frac{N^{TRIG}}{N^{TIS}}$$

In the first approximation (no p_t binning)

$$\varepsilon^{TIS} \approx \frac{N^{TIS \& TOS}}{N^{TOS}} \rightarrow \varepsilon^{TRIG} \approx \frac{N^{TIS \& TOS}}{N^{TOS}} \frac{N^{TRIG}}{N^{TIS}}$$

Where all the four numbers are observables. Here we get, with respect to events that would be offline selected by the tight selection, an efficiency for $B^+ \rightarrow J/\psi(\mu\mu) K^+$ of about 93.4 % with high statistical accuracy. We could bin in p_t to improve the precision of the number, but as the efficiency is so high, the correction will be very small.

For the signal we could use the trigger emulation, but as $B_s \rightarrow \mu^+ \mu^-$ muons have larger p , p_t and IP, the trigger efficiency for $B_s \rightarrow \mu^+ \mu^-$, will be even higher than for $B^+ \rightarrow J/\psi(\mu\mu) K^+$. As this last one is already very high, the range for $\varepsilon_s^{TRIG/SEL}$ is already very small:

$$0.934 \leq \varepsilon_s^{TRIG/SEL} \leq 1$$

So, taking this safest margin:

$$\frac{\varepsilon_+^{TRIG/SEL}}{\varepsilon_s^{TRIG/SEL}} = 0.966 \pm 0.033$$

Which has a precision of 3.4% and hence good enough for this example.

5.4.3. Final normalization factor

Once that all the factors entering in the normalization and the calibration have been studied, we will write down the final normalization factor, as a function of independent sources of errors.

$$BR = BR_+ \frac{\varepsilon_+^{off}}{\varepsilon_s^{off}} \frac{\varepsilon_+^{tr}}{\varepsilon_s^{tr}} f_+ \frac{N_s}{N_+} = (6.07 \pm 0.93) \left(\frac{BR_+}{BR_d} \frac{N_d}{N_+^2} \frac{\varepsilon^T}{\varepsilon_{GL>0.5}^{GL>0.5}} \right) N_s = (1.81 \pm 0.31) \cdot 10^{-9} \frac{\varepsilon^T}{\varepsilon_{GL>0.5}^{GL>0.5}} N_s$$

Where BR_+ is $BR(B^+ \rightarrow J/\psi(\mu\mu) K^+)$, BR_d is $BR(B_d \rightarrow J/\psi(\mu\mu) K^{*0}(K\pi))$ N_d and N_+ are the number of $B_d \rightarrow J/\psi(\mu\mu) K^{*0}(K\pi)$ and events $B^+ \rightarrow J/\psi(\mu\mu) K^+$ after the tight selection. The number $(1.86 \pm 0.32) \cdot 10^{-9}$ is actually the normalization factor we would have to use if we were applying the tight selection to $B_s \rightarrow \mu^+ \mu^-$, and agrees very well with the MC truth number $(1.75 \cdot 10^{-9})$. $\frac{\varepsilon^T}{\varepsilon_{GL>0.5}^{GL>0.5}}$ is the ratio of efficiencies of the tight selection and the cut $GL > 0.5$. This ratio was evaluated to be 1.336 ± 0.078 in 5.3.2.2, and is also in very good agreement with the MC truth number, 1.31. As the numbers which enter in the calculation of this last ratio are closely related to those of the GL calibration, we could leave the expression as a function of the appropriate numbers. But as the error is so small compared to the first factor of the normalization, we just combine them and add the error in quadrature, providing the final normalization factor:

$$BR = BR_+ \frac{\varepsilon_+^{off}}{\varepsilon_s^{off}} \frac{\varepsilon_+^{tr}}{\varepsilon_s^{tr}} f_+ \frac{N_s}{N_+} = (2.42 \pm 0.43) \cdot 10^{-9} N_s \equiv \alpha N_s$$

The summary of the numbers which enter into it is shown in

Table 5-VIII: Parameters for computing the normalization factor.

Parameter	Value	Description
$\frac{f_+}{f_s}$	4.00 ± 0.55	Ratio of hadronization fractions used in this example. Not exactly equal to PDG, but with equivalent error
$\frac{\varepsilon_+^{tr}}{\varepsilon_s^{tr}}$	0.967 ± 0.033	Ratio of trigger efficiencies
$\frac{\varepsilon^{M_{J/\psi}}}{\varepsilon^{M_{Bs}}}$	1.03 ± 0.02	Ratio of efficiencies between J/ψ and B_s for a 60 MeV mass window cut
$\varepsilon^{M_{K^{*0}}}$	0.67 ± 0.02	Efficiency of a 40 MeV mass window cut in the K^{*0} . Goes in the denominator
$\varepsilon^{K,\pi_{IPS}}$	0.98 ± 0.02	Efficiency of IPS cut on K^{*0} daughters after all the cuts are applied. Goes in the denominator
-	~5% systematic	Differences in ratio 4 body/3 body for different common selections
All above	6.07 ± 0.93	
$\frac{BR_+^2}{BR_d}$	$(6.80 \pm 0.48) \cdot 10^{-5}$	BR_+ is $BR(B^+ \rightarrow J/\psi(\mu\mu) K^+)$, BR_d is $BR(B_d \rightarrow J/\psi(\mu\mu) K^{*0}(K\pi))$
$\frac{N_d}{N_+^2}$	$(4.397 \pm 0.065) \cdot 10^{-6}$	N_d and N_+ are the number of $B_d \rightarrow J/\psi(\mu\mu) K^{*0}(K\pi)$ and events $B^+ \rightarrow J/\psi(\mu\mu) K^+$ after the tight selection
$\frac{\varepsilon^T}{\varepsilon^{GL>0.5}}$	1.336 ± 0.078	Ratio of efficiencies of the tight selection and the cut $GL > 0.5$ (see 5.3.2.2)
All	$(2.42 \pm 0.43) \cdot 10^{-9}$	Normalization factor

5.5.Result

5.5.1. Summary of normalization and calibration parameters

The result of this analysis relies on 5 signal related parameters and 4 background related parameters that have to do with normalization and modelling of signal and background pdf's. These nine independent parameters are shown in Table 5-IX:

Table 5-IX: The nine independent parameters related to normalization and calibration.

Parameter	Value	Description
α	$(2.42 \pm 0.43) \cdot 10^{-9}$	Normalization factor (see 5.4)
$M_{B_s \rightarrow \mu\mu}$	5366.4 ± 2.5 MeV	B_s meson invariant mass (see 5.3.1.1)
$\sigma_{B \rightarrow \mu^+ \mu^-}$	20 ± 1 MeV	Invariant mass resolution for signal (see 5.3.1.1)
$a_{B \rightarrow \mu^+ \mu^-}$	1.9 ± 0.1	Transition point of the crystalball (see 5.3.1.2)
f_{GL}	0.439 ± 0.030	Signal fraction in $0.5 < GL < 0.75$ (see 5.3.2.2)
N_{sb1}	17^{+5}_{-4}	Expected value of background yield in sidebands for $0.5 < GL < 0.75$ (see 5.3.3)
N_{sb2}	8^{+4}_{-3}	Expected value of background yield in sidebands for $GL > 0.75$ (see 5.3.3)
k_{sb1}	$0.0 \pm 6.6 \cdot 10^{-4}$ MeV $^{-1}$	Exponential parameter of background mass line shape for $0.5 < GL < 0.75$ (see 5.3.3)
k_{sb2}	$0.0 \pm 8.8 \cdot 10^{-4}$ MeV $^{-1}$	Exponential parameter of background mass line shape for $GL > 0.75$ (see 5.3.3)

5.5.2. Binning

Here we will not perform any binning optimization, we will keep the GL binning 0.50 -0.75 - 1.00 used for illustration of the GL calibration. The invariant mass is binned in intervals of one sigma, and the PID likelihood is not being used. The number of observed events in each bin is shown in Figure 5-20.

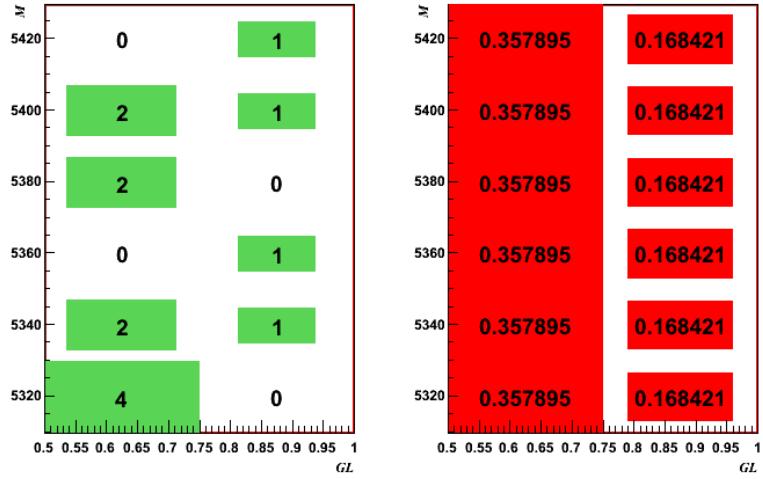


Figure 5-20: Experiment binning.

Left: Observed data. Right: Expected background. Invariant mass is in MeV.

5.5.3. Obtained measurement and limits

Using the binning defined in the previous section and the CLs method, we will provide the final measurement of the $\text{BR}(\text{B}_s \rightarrow \mu^+ \mu^-)$. If we look at the highest GL bins (see Figure 5-21) we can differentiate some excess in the B_s peak

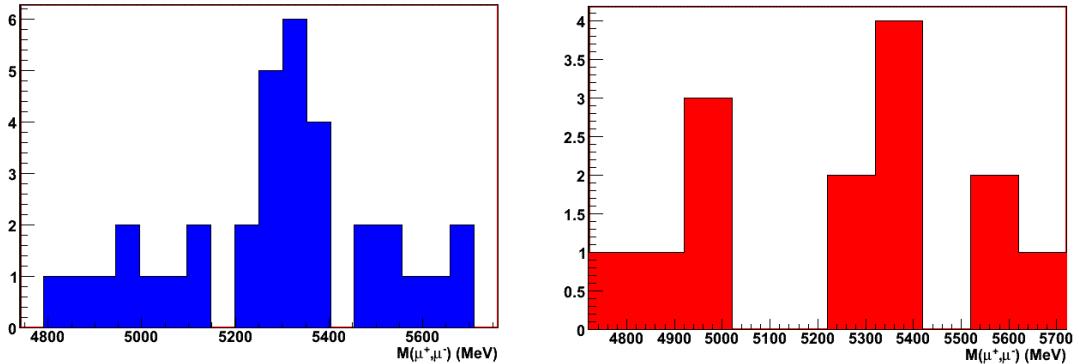


Figure 5-21: Observed invariant mass distribution of $\text{B}_s \rightarrow \mu^+ \mu^-$ candidates.

Left: in $0.5 < \text{GL} < 0.75$. Right: in $\text{GL} > 0.75$.

Using the code [112] for the CL_s method with inclusion of nuisance parameters, we get the values of CL_s and CL_b shown in Figure 5-22.

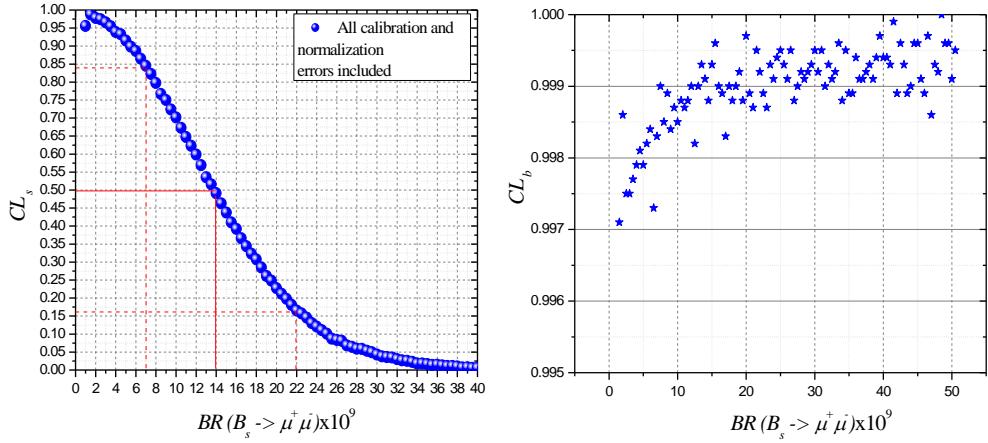


Figure 5-22: CL_s and CL_b curves as a function of $BR(B_s \rightarrow \mu^+ \mu^-)$ including all calibration and normalization errors.

From which we get the following results:

- $BR(B_s \rightarrow \mu^+ \mu^-) < 2.5 \cdot 10^{-8}$ @ 90 % CL
- $BR(B_s \rightarrow \mu^+ \mu^-) < 2.9 \cdot 10^{-8}$ @ 95 % CL
- $CL_s (BR < 5 \cdot 10^{-9}) > 0.9$, $CL_s (BR = 3.5 \cdot 10^{-9}) = 0.95$, which imply that SM-like values would be excluded at 90 – 95 % CL.
- CL_b in the range $0.9980 - 0.9995$. This corresponds to a bit more than three sigma evidence.
- Finally, the measured value is:

$$BR(B_s^0 \rightarrow \mu^+ \mu^+) = (1.40^{+0.85}) \cdot 10^{-8}$$

With a 3.1-3.5 sigma evidence over background hypothesis.

5.5.4. Result neglecting normalization and calibration errors

In this section we will perform the analysis only with the central values of the nine input parameters related to normalization and calibration, neglecting their errors. This quick study will give us a feeling of the impact of those errors on the sensitivity. In this case we obtain (see Figure 5-23):

- $BR(B_s \rightarrow \mu^+ \mu^-) < 2.7 \cdot 10^{-8}$ @ 90 % CL
- $BR(B_s \rightarrow \mu^+ \mu^-) < 3.1 \cdot 10^{-8}$ @ 95 % CL
- $CL_s (BR < 6 \cdot 10^{-9}) > 0.9$, $CL_s (BR = 4.0 \cdot 10^{-9}) = 0.95$, which imply that SM-like values would be excluded at 90 – 95 % CL.
- $CL_b > 0.9985$. This implies more than three sigma evidence. To disentangle whether it corresponds to a 5 sigma discovery we would need, in principle, to follow another approach or at least to use software different from [112].

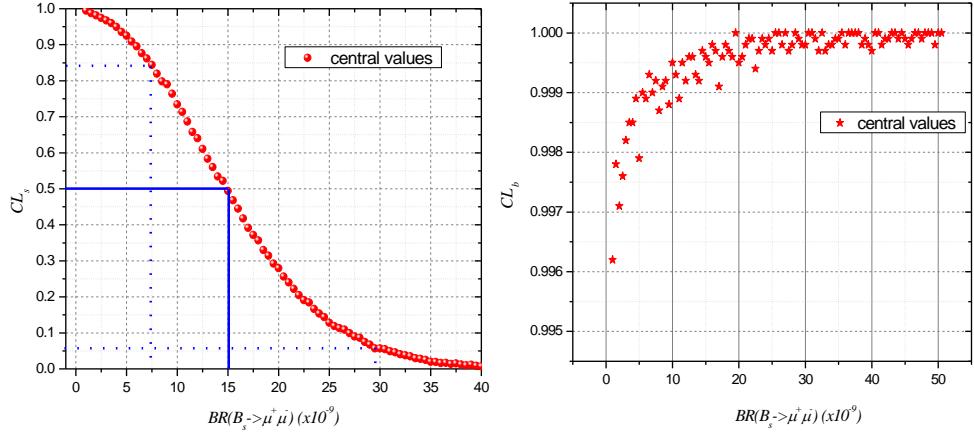


Figure 5-23: CL_s and CL_b curves as a function of $BR(B_s \rightarrow \mu^+ \mu^-)$ without including calibration or normalization errors.

Finally, the measured value –neglecting normalization and calibration errors– is:

$$BR(B_s^0 \rightarrow \mu^+ \mu^+)|_{cv.} = (1.50^{+0.85}) \cdot 10^{-8}$$

We see that the effect of including all those errors is small in the one-sigma interval, but the effects in the CL_b are clearly visible, distorting the significance over background hypothesis.

5.5.5. Comparison using MC truth efficiencies and pdf's

In this section we will redo the analysis using MC truth information about signal and background pdf's, efficiencies and amount of background in the search window. In this case we obtain (see Figure 5-24):

- $BR(B_s \rightarrow \mu^+ \mu^-) < 2.6 \cdot 10^{-8}$ @ 90 % CL.
- $BR(B_s \rightarrow \mu^+ \mu^-) < 2.9 \cdot 10^{-8}$ @ 95 % CL.
- $CL_s (BR < 5 \cdot 10^{-9}) > 0.9$, $CL_s (BR = 3.5 \cdot 10^{-9}) = 0.95$, which imply that SM-like values would be excluded at 90 – 95 % CL.
- CL_b in the range $0.998 - 1$. This implies more than three sigma evidence. To disentangle whether it corresponds to a 5 sigma discovery we would need, in principle, to follow another approach or at least to use software different from [112].

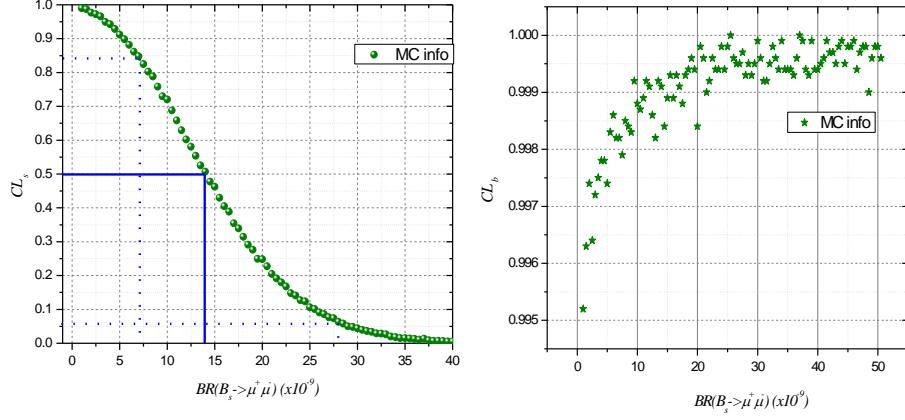


Figure 5-24: CL_s and CL_b curves as a function of $BR(B_s \rightarrow \mu^+ \mu^-)$ using MC truth efficiencies and pdf's.

Finally, the measured value would be:

$$BR(B_s^0 \rightarrow \mu^+ \mu^+)|_{MCinfo} = (1.40^{+0.85}_{-0.70}) \cdot 10^{-8}$$

Of course in agreement with the input used: $2.0 \cdot 10^{-8}$.

5.6.Implications

5.6.1. Test of the SM

From the results shown in 5.5.3, we concluded that SM should be excluded at $\sim 90\text{-}95\%$ CL. Here we will do a more accurate analysis by testing the signal hypothesis corresponding to SM, absorbing the theoretical error of the SM prediction into the normalization factor.

Assuming no NP in ΔM_s

$$BR(B_s \rightarrow \mu\mu)^{SM} = (3.35 \pm 0.32) \cdot 10^{-9}$$

Which gives $CL_{SM} = 0.932$, $CL_{SM+b} = 0.926$. So, if there is no NP in ΔM_s , our measurement is 1.3σ deviated from SM.

If we allow for possible NP in ΔM_s , then we should use instead [113][16]:

$$BR(B_s \rightarrow \mu\mu)^{SM} = (3.86 \pm 0.57) \cdot 10^{-9}$$

Which gives $CL_{SM} = 0.911$, $CL_{SM+b} = 0.905$.

So we conclude that the result is incompatible with SM at $\gtrsim 90\%$ CL. The deviation from SM is 1.2σ , and it becomes 1.3σ if we assume no NP in the oscillation of the B_s meson. Note also that if we consider that NP contribution can only be positive, the excluded areas are equivalent to 1.7σ and 1.8σ respectively.

5.6.2. 2HDM-II

In this section we will use the one sigma interval to map the allowed region of the parameter space for 2HDM-II. The result is shown in Figure 5-25.

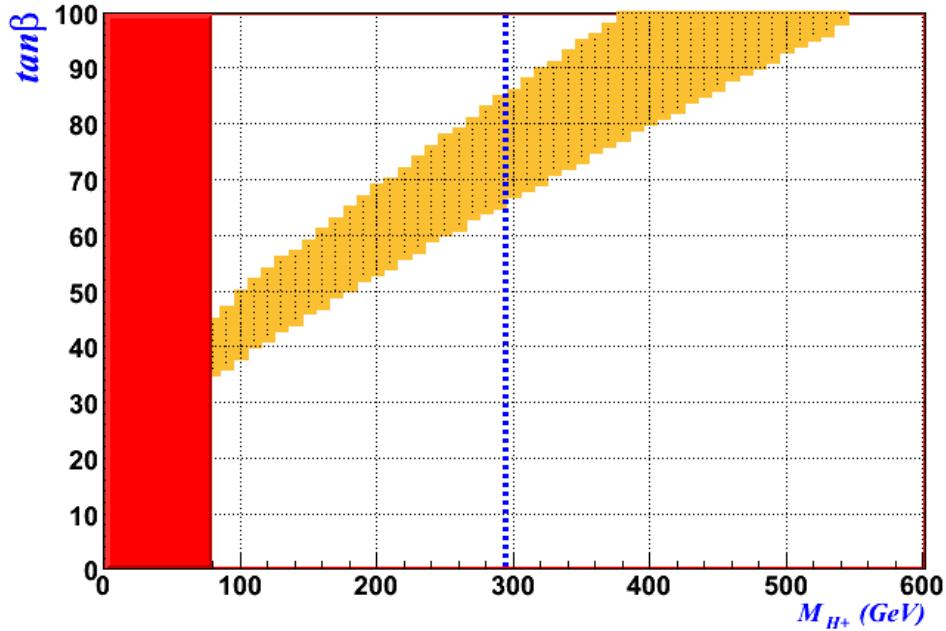


Figure 5-25: Implications in the context of 2HDM-II for the result in the example.

The orange diagonal band is the region compatible with the measurement at one sigma. The limit on the Higgs mass $M_{H^+} > 80$ GeV (red band) from direct search and the indirect limit from $b \rightarrow s\gamma M_{H^+} > 295$ GeV are also shown.

We see that $\tan\beta > 35$ in this context, and if we add the limit from $b \rightarrow s\gamma$ it becomes $\tan\beta > 65$.

5.6.3. mSUGRA

In the context of mSUGRA the values of $\tan\beta$ are lower than in the case of 2HDM-II as the BR grows with higher powers. We see in Figure 5-26 that values of $\tan\beta \sim 50$ are preferred, and it can be even lower (30-40) if the SUSY mass parameters are below 1 TeV. The figure was made using SoftSUSY¹² and SUITY v-1.0¹³, with the constraints $A_0 = 0$, $\mu > 0$, $M_{h0} > 114$ GeV (lightest Higgs mass) and $M_W = 80.398 \pm 0.025$ GeV.

¹² Provided by Ben Allanach (Cambridge)

¹³ Provided by Athanasios Dedes (Durham)

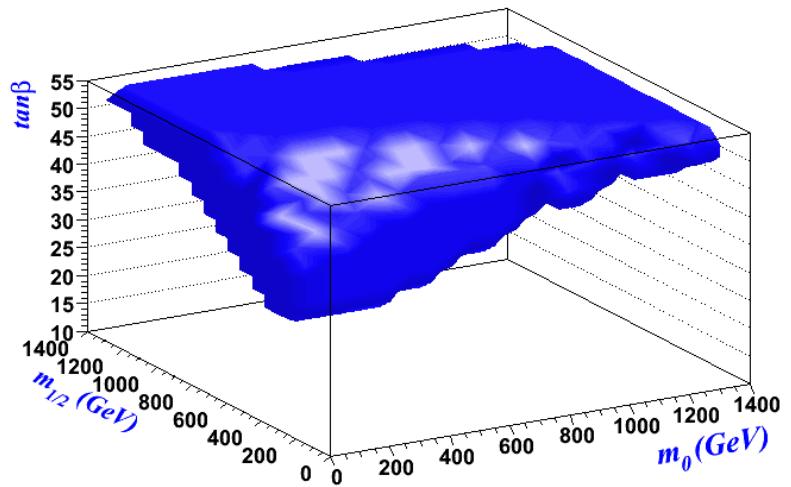


Figure 5-26: mSUGRA parameter space allowed by the measurement at one sigma.

6. Look at 2009 data

During December 2009, LHCb took its first collision data, at energy of 450 MeV per beam. A large amount of these data was taken with the VELO at 15 mm (all the rest at 29 mm or switched off), and a minimum bias trigger consisting on loose L0 calorimeter, muon and Pile-up cuts.

In this section we will study the distributions of V^0 's K_s and Λ on this first data and compare them with MC in order to check how well it is reproducing the current performance of the detector. We will focus on LL candidates (i.e., those formed by two long tracks) as this is the case of $B_s \rightarrow \mu\mu$, but DD candidates (formed by two downstream tracks) will be looked into as well.

About 224 K MC events were generated to start comparisons with data and validation of the detector performance. This MC has also the VELO at 15 mm, and the L0 was rerun to satisfy similar conditions. As not all the real data had the L0 Pile-up activated, here we decide to not run it on MC and just ignore those real events that rely only on this line. The amount of data rejected by this veto was less than per mil. There were some remaining differences in the trigger in MC as the simulation added ECAL and HCAL energies, while the trigger in real data was using only the HCAL. In any case, the efficiency of the calorimeter trigger is about 99 % on MC offline selected candidates, and 97-99 % on data as was obtained using the procedure described in 4.4.2 [114]. This means the bias in the distributions could only be very small, and hence the effect of the different treatment of calorimeter energy is negligible for the comparisons data-MC we are doing in this chapter.

All the runs/events with the VELO not at 15 mm are also ignored, as well as those in which the calorimeter and muon threshold of the L0 were higher than in the main data sample. In total, after veto we keep 91 % of the selected V^0 candidates.

6.1. Event selection, invariant mass and statistics

6.1.1. Event selection

The V^0 candidates were reconstructed by combining pions and protons made with long or downstream tracks. The IP of the V^0 with respect of the primary vertex was requested to be less than 6 mm and a cut on the lifetime larger than 10% of a meantime (i.e the numerical cut values are 8.953 ps for K_s and 26.31 ps for Λ) was applied. This lifetime is computed using the distance to the PV, and corresponds to the rest frame of the V^0 .

The invariant mass windows were $310 \text{ MeV} < M < 600 \text{ MeV}$ for K_s and $1100 \text{ MeV} < M < 1140 \text{ MeV}$ for Λ . These windows are large enough to allow an accurate background subtraction from sidebands. The fact that the lower sideband for K_s is larger than upper sideband is motivated to allow a large amount of Λ to be selected also by the K_s selection¹⁴ and hence could be used for crosschecks with those coming from Λ selection, as well as to allow looking at Armenteros-Podolanski [115](AP) plot in a single sample. Any PID requirement is applied neither in the K_s nor in the Λ selection, so all the tracks are considered twice: as pions and as protons. Kinematical cuts are enough to disentangle K_s , Λ and $\bar{\Lambda}$ as will become clear in the next paragraphs.

¹⁴ Λ daughters in Λ rest frame have lower momentum compared to K_s daughters in K_s rest frame due to the proximity to the p- π threshold. As a numerical example, computing Λ mass under $\pi\pi$ hypothesis in Λ rest frame would give a mass of ~ 340 MeV.

We will now have a look at the invariant mass resolution of K_s and Λ in this first data. As explained in the previous chapters, the invariant mass is one of the key points in the study of $B_s \rightarrow \mu\mu$ and one of the main advantages of LHCb with respect to ATLAS and CMS.

The selection above is enough to allow us to see the invariant mass peak of the two hadrons (see Figure 6-1).

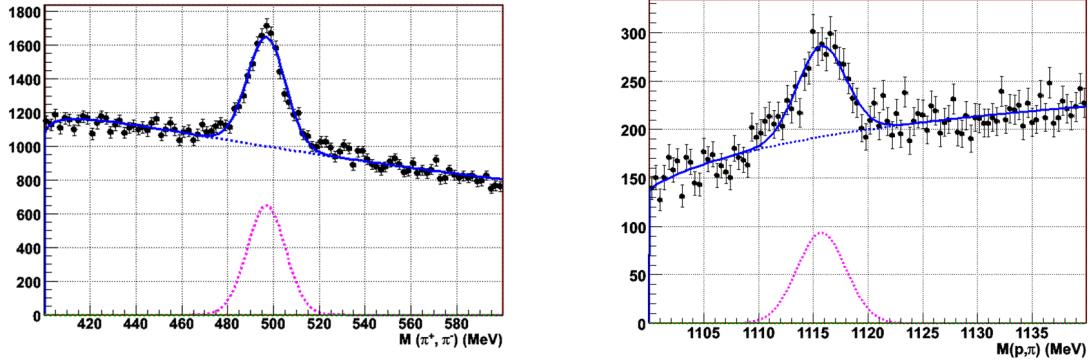


Figure 6-1: Reconstructed invariant mass of V^0 candidates selected in 2009 preliminary LHCb data.

Left: K_s , Right: Λ .

In principle, the sidebands of Λ selection are tight enough to not allow the Λ to be identified as $\bar{\Lambda}$ (and vice versa, of course) as exchanging the daughters mass hypothesis leads to a sizable overestimation of the invariant mass. But even though, to completely ensure the particle-antiparticle separation, we will use that fact that the proton takes most of the Λ momentum. This is illustrated in the Armenteros-Podolanski plot (see Figure 6-2 and Figure 6-3), which corresponds to the scatter plot of the variables $\alpha = \frac{p_L^+ - p_L^-}{p_L^+ + p_L^-}$ - being p_L^+ (p_L^-) the momentum of the positive (negative) daughter in the flight direction of the mother- and p'_T , the transverse momentum to the flight direction of the mother of any of the daughters (it must be the same for both). In this plot three ellipses are seen:

- One large ellipse that goes from $\alpha = -1$ to $\alpha = 1$ and reaches $p'_T \sim 206$ MeV. This corresponds to K_s . Analytically corresponds to (velocity $\beta = 1$):

$$\frac{4p'^2_T}{m_{K_s}^2} + \alpha^2 = 1 - \frac{4m_\pi^2}{m_{K_s}^2}$$

- A small ellipse that goes from $\alpha = -1$ to $\alpha = -0.5$. Those are the $\bar{\Lambda}$, where almost all the momentum is taken by the negative particle (the \bar{p}).
- A small ellipse that goes from $\alpha = 0.5$ to $\alpha = 1$. Those are the Λ , where almost all the momentum is taken by the positive particle (the p). Analytically corresponds to (velocity $\beta = 1$):

$$\frac{4p'^2_T}{m_\Lambda^2} + (\alpha - \alpha_0)^2 = \frac{4p'^2}{m_\Lambda^2}$$

Being p' the momentum of the daughters of Λ in Λ rest frame, approximately 100.7 MeV, and α_0 approximately 0.6916 (-0.6916) for Λ ($\bar{\Lambda}$).

Hence a simple cut $\alpha > 0$ gives us the baryon and $\alpha < 0$ the anti-baryon.

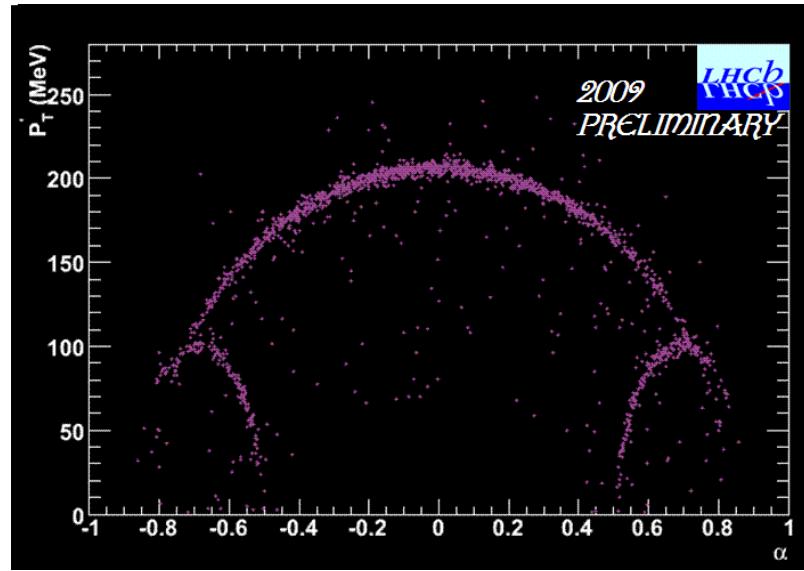


Figure 6-2: Armenteros-Podolanski plot for LL candidates in 2009 data.

The sample corresponds to $GL_{LL} > 0.05$ (see 6.2.1).

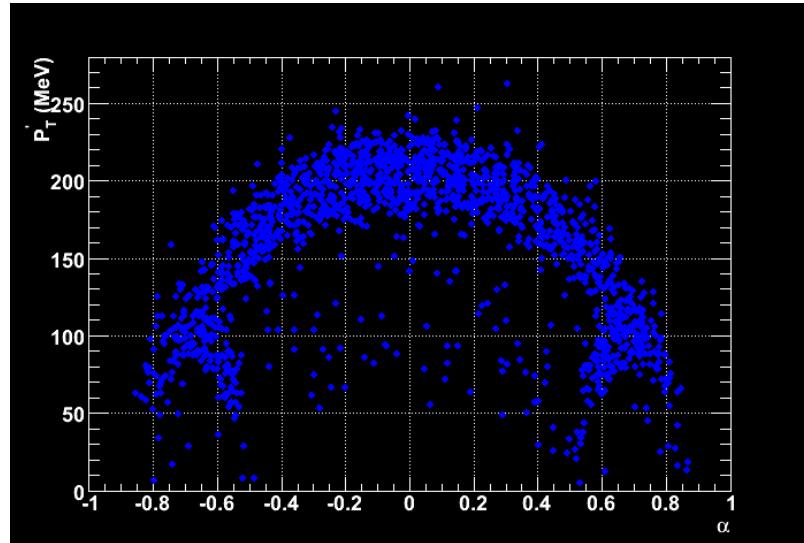


Figure 6-3: Armenteros-Podolanski plot for DD candidates in 2009 data.

The sample corresponds to $GL_{DD} > 0.4$ (see 6.2.1).

6.1.2. Invariant mass description and parameters

The signal peak is parameterized with a single or double Gaussian depending on the agreement with observed distribution. The background for K_s is well described by an exponential function. However, for the case of Λ , due to the proximity to the $p-\pi$ threshold, a logarithmic ($\rho(m) = \log(m - m_0)$) component is also added to the background PDF.

The measured parameters for single Gaussian signal are shown in Table 6-I.

Table 6-I: Invariant mass parameters for observed strange hadrons.

All numbers are in MeV. Label ‘x’ refers to events with beam crossing, while ‘bg’ refers to events with only one beam and therefore beam-gas events.

	m_Λ	σ_Λ	$m_{\bar{\Lambda}}$	$\sigma_{\bar{\Lambda}}$	$m_{K_S^0}$	$\sigma_{K_S^0}$
450 GeV LL (x)	1115.5 ± 0.2	1.4 ± 0.2	1115.8 ± 0.2	1.6 ± 0.2	497.4 ± 0.2	4.2 ± 0.2
450 GeV LL (bg)	1115.6 ± 0.6	1.5 ± 0.5	-	-	499.0 ± 0.5	4.3 ± 0.5
450 GeV DD (x)	1114.8 ± 0.3	2.8 ± 0.2	1117.1 ± 0.4	2.0 ± 0.5	496.2 ± 0.4	11.0 ± 0.5
450 GeV DD (bg)	1113.9 ± 0.7	1.7 ± 0.7	-	-	500 ± 3	11.0 ± 2

To be compared (Table 6-II) with the ones obtained performing the same fit on MC (the background was not filtered out in order to perform a fair comparison).

Table 6-II: Invariant mass parameters for observed MC.

All numbers are in MeV.

	m_Λ	σ_Λ	$m_{\bar{\Lambda}}$	$\sigma_{\bar{\Lambda}}$	$m_{K_S^0}$	$\sigma_{K_S^0}$
450 GeV LL	-	-	-	-	497.6 ± 0.2	3.6 ± 0.2
450 GeV DD	1115.7 ± 0.2	1.2 ± 0.1	1115.6 ± 0.2	1.1 ± 0.2	498.3 ± 0.2	6.0 ± 0.2

The good similarity for the LL resolution is a promising result for the $B_s \rightarrow \mu\mu$ analysis with future data.

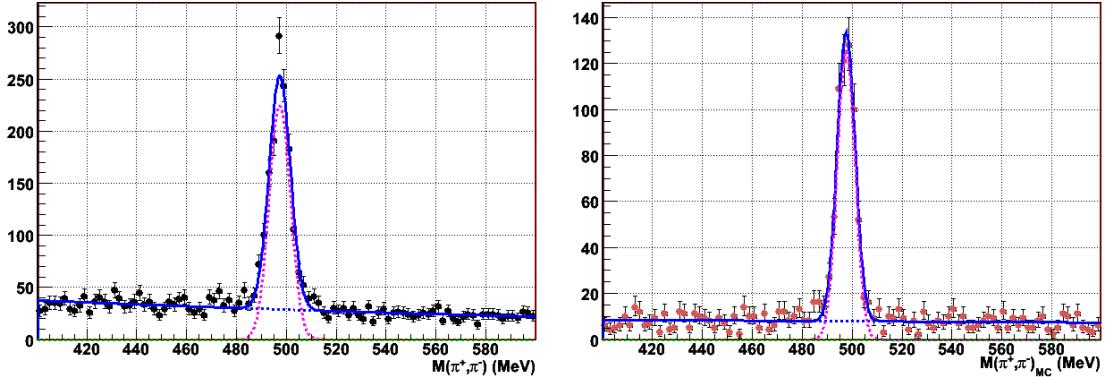


Figure 6-4: LL $\pi^+\pi^-$ invariant mass distribution.

Left: 2009 LHCb preliminary data. Right: MC.

Another good check for the quality of invariant mass reconstruction is superimposing the AP plot of data and MC. There is no information about the hypothesis about daughter masses in such plot, all the information there comes entirely from kinematic information reconstructed by the detector. We see in Figure 6-5 that the agreement among both samples is very good. Figure 6-6 superimposes the analytical formulae to the real data points.

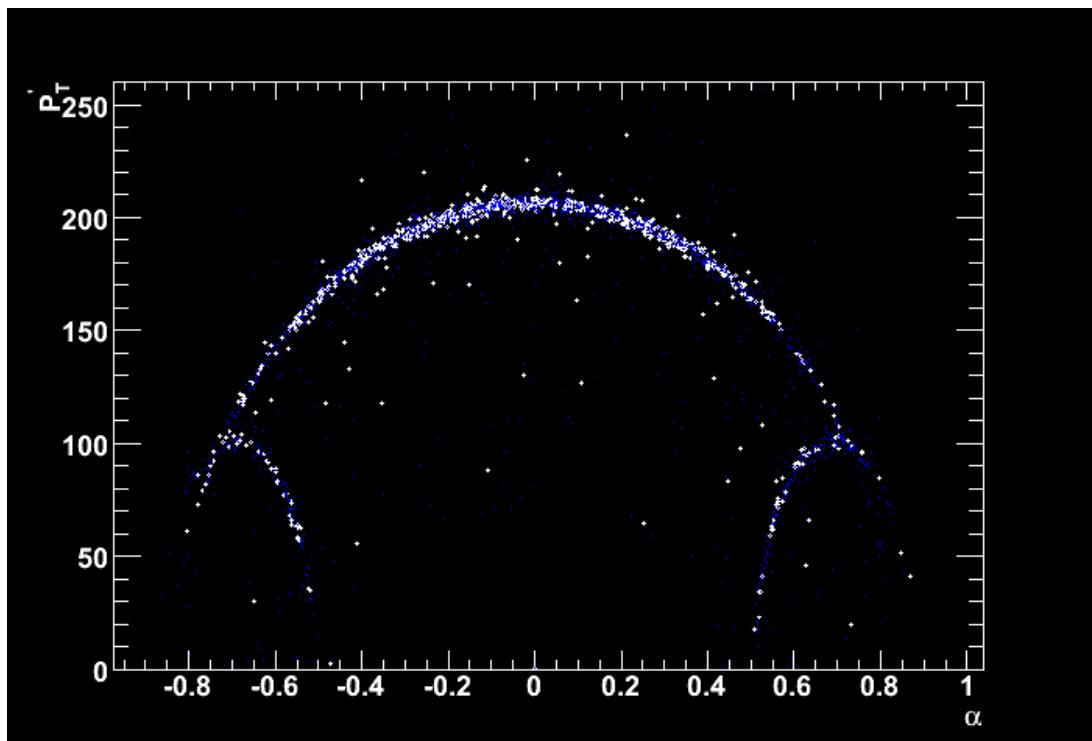


Figure 6-5: MC and data superimposed AP plots for LL V^0 candidates.

Blue: real data. White: MC. The sample corresponds to $GL_{LL} > 0.05$ (see 6.2.1).

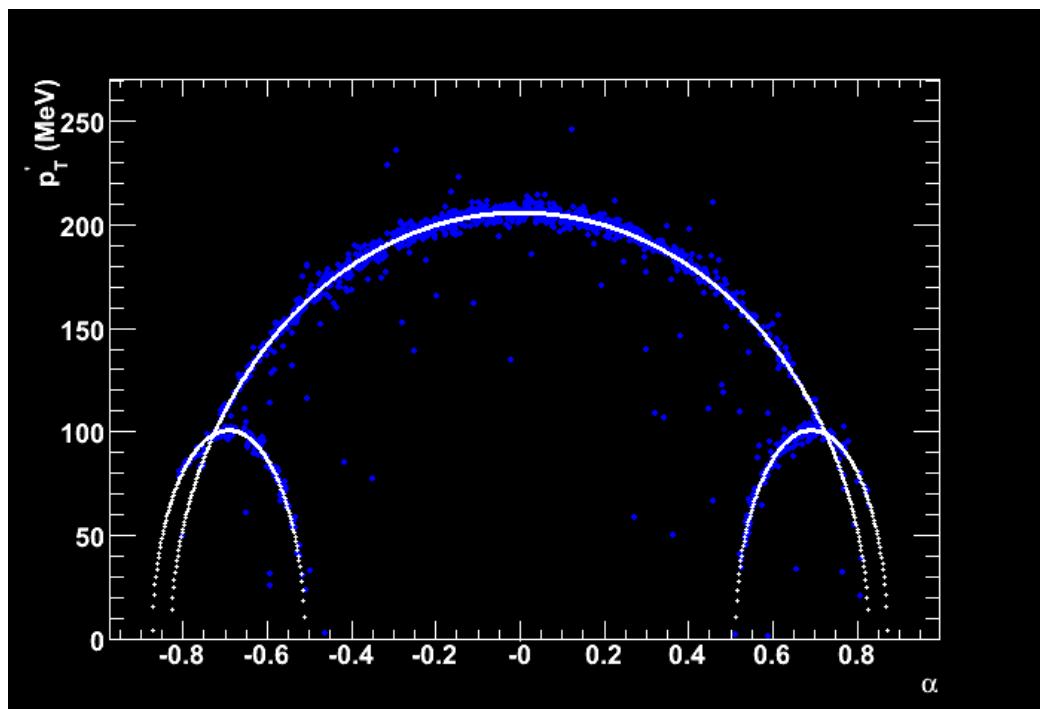


Figure 6-6: Analytical AP ellipses superimposed to LL V^0 candidates from data.

Blue: real data. White: points from analytical formulae. The sample corresponds to $GL_{LL} > 0.1$ (see 6.2.1).

6.1.3. Signal yield

6.1.3.1. Extraction of beam gas component

Even in those events where the two beams are crossing, we have a non negligible probability of getting a beam-gas interaction. Therefore, the beam-gas component has to be statistically subtracted from the beam-beam component to get the distributions that correspond to pp collisions. This is done under the assumption that the probability to get a beam-gas is the same in beam-beam as in beam-empty events. The ratio of beam-beam to empty beam crossings is extracted from the number of bunch ids that correspond to each case. Averaging for all the runs used in this chapter, we have computed the fraction beam-beam to beam-empty to be very close to one. Hence, the real number corresponding to 900 GeV centre of mass collisions, for a given bin of any distribution is:

$$N_i^{pp} = N_i^{bb} - nN_i^{eb} = N_i^{bb} - N_i^{eb}$$

Where *bb* correspond to beam-beam events, and *eb* to beam-empty events.

Because of the right description of the invariant mass in the Λ sample is more complicated and the statistics in *eb* very small, the parameters for signal description (mean and sigma) in *eb* are considered the same as in *bb* in each bin where the subtraction is performed.

6.1.3.2. Number of K_s , Λ and $\bar{\Lambda}$,

Here we will measure the number of K_s , Λ and $\bar{\Lambda}$, through an invariant mass fit. The signal is modelled as a single and as a double Gaussian, to check the dependence on the accuracy of the invariant mass line shape model. The width of the second Gaussian was restricted to be at most 4 MeV larger than the first component.

The measured numbers after background and beam-gas subtraction are shown in Table 6-III.

Table 6-III: Measured number of K_s , Λ and $\bar{\Lambda}$ in pp collisions, after background and beam-gas subtraction.

	Single Gaussian			Double Gaussian		
	N_Λ	$N_{\bar{\Lambda}}$	N_{K_s}	N_Λ	$N_{\bar{\Lambda}}$	N_{K_s}
LL	170 ± 19	119 ± 19	1081 ± 42	184 ± 21	130 ± 24	1142 ± 55
DD	593 ± 76	490 ± 82	3420 ± 140	810 ± 110	696 ± 80	3590 ± 140
LD or DL	-	-	1910 ± 160	-	-	1970 ± 160

The number of MC truth matched V^0 's used for comparisons with data is shown in Table 6-IV, and Table 6-V shows the measured number of signal candidates on the beam-gas events after background subtraction.

Table 6-IV: Number of MC truth matched V^0 's.

	N_Λ	$N_{\bar{\Lambda}}$	N_{K_s}
LL	42	42	597
DD	275	261	2400
LD or DL	33	33	1099

Table 6-V: Measured number of K_s , Λ and $\bar{\Lambda}$ in beam-gas events, after background subtraction.

	Single Gaussian			Double Gaussian		
	N_Λ	$N_{\bar{\Lambda}}$	N_{K_s}	N_Λ	$N_{\bar{\Lambda}}$	N_{K_s}
LL	21 ± 6	8 ± 5	113 ± 12	22 ± 6	10 ± 6	121 ± 14
DD	39 ± 14	2 ± 14	147 ± 33	48 ± 16	2 ± 20	146 ± 32
LD or DL	-	-	109 ± 31	-	-	~ 110

The number of pp collisions stored in 2009 data corresponds to about ~ 300 K, while the number of MC events ran on is 224 K. However, as the aim of this chapter is to compare distributions, we do not need the relative luminosities between the two samples.

6.2.Signal Distributions

6.2.1. Geometry Likelihood and related variables

One of the most important points of the $B_s \rightarrow \mu\mu$ analysis is the Geometry Likelihood. Due to the fact that $B_s \rightarrow \mu\mu$ is also a V0, the relevant variables that enter in the GL can be tested with 2009 data. We started by creating two GL's, one for K_s LL candidates and another for K_s DD candidates. The GL's were tuned with 597 MC signal events and 1019 MC background events for LL, and 2400 signal and 14151 background for DD. Only combinatorial background was used (i.e., the Λ were removed) for the tuning of the GL. The isolation was not used as most of the events, even the background, were populating the bin at 0.

The real data distributions shown in the following figures were obtained by signal + background fit in each bin, both in bb events and eb events to correctly perform the beam gas subtraction as described in 6.1.3.1. The signal was described by single Gaussian.

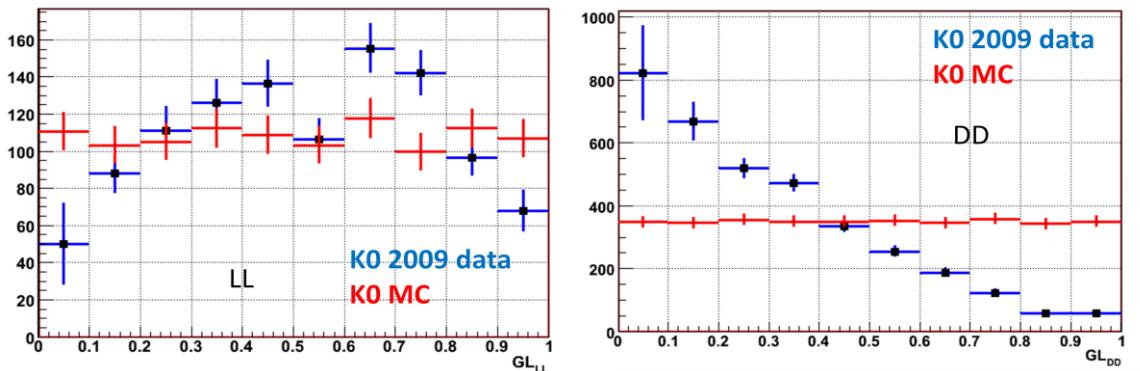


Figure 6-7: GL distributions for K_s .

Left: for LL. Right: for DD. The blue histogram with black squares is K_s signal in pp collisions data. Red histogram is MC truth K_s signal.

From Figure 6-7 we see that the agreement for LL is reasonable, in the sense that real data access also all the range 0-1 even if not fully flat. However, the DD combinations have clearly worse geometry than in

MC.

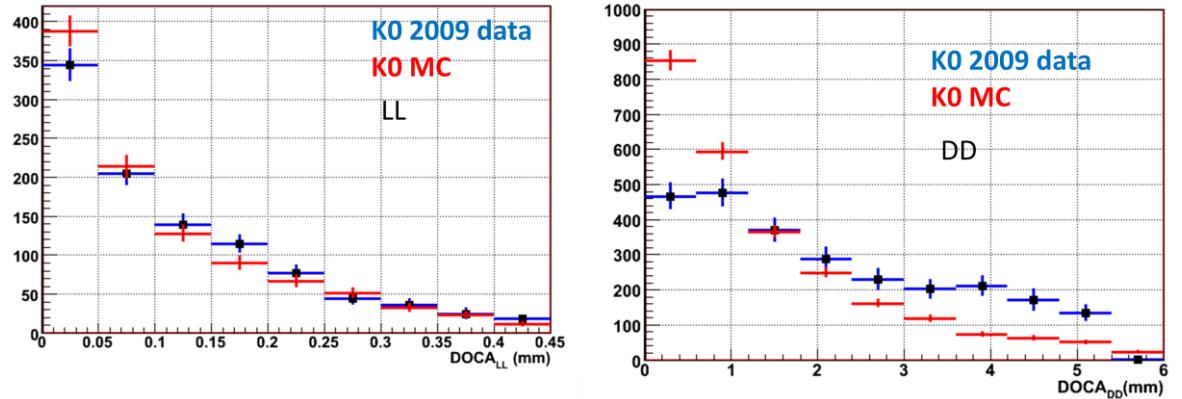


Figure 6-8: DOCA distributions for K_s .

Left: for LL. Right: for DD. The blue histogram with black squares is K_s signal in pp collisions data. Red histogram is MC truth K_s signal.

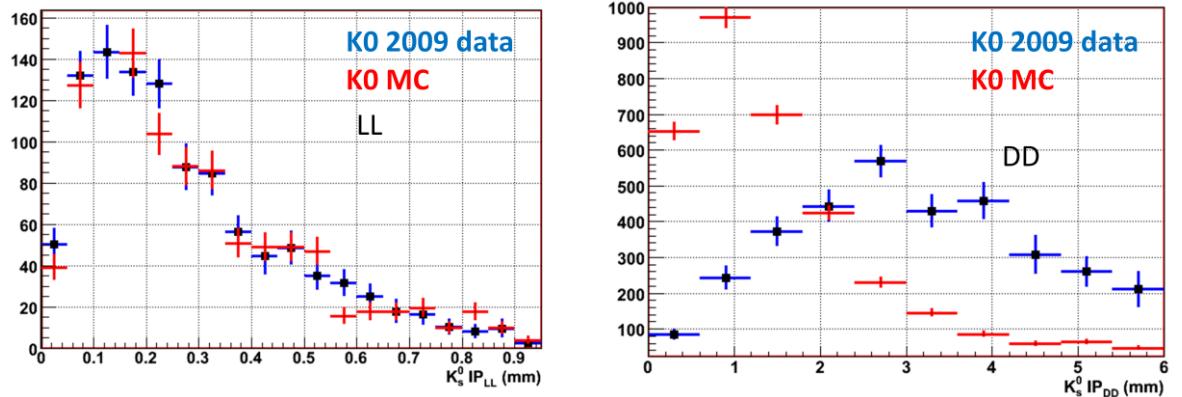


Figure 6-9: IP distributions for K_s .

Left: for LL. Right: for DD. The blue histogram with black squares is K_s signal in pp collisions data. Red histogram is MC truth K_s signal.

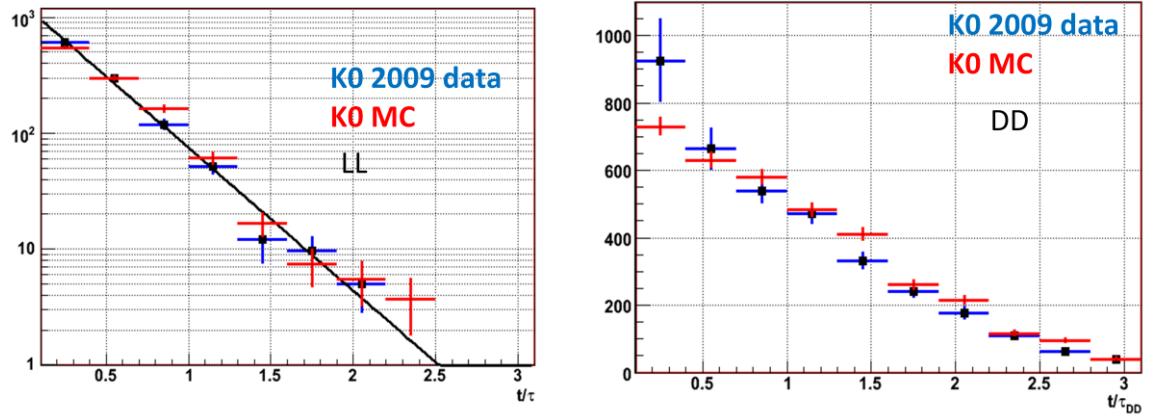


Figure 6-10: Lifetime distributions for K_s , in meantime units.

Left: for LL. Right: for DD. The blue histogram with black squares is K_s signal in pp collisions data. Red histogram is MC truth K_s signal.

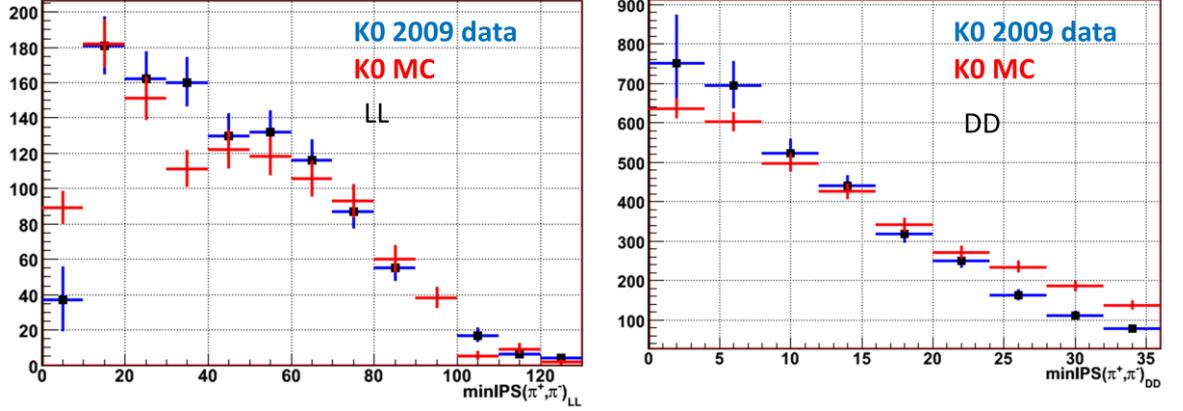


Figure 6-11: minimum IPS for K_s daughters.

Left: for LL. Right: for DD. The blue histogram with black squares is K_s signal in pp collisions data. Red histogram is MC truth K_s signal.

A look at the input variables (see Figure 6-8 to Figure 6-11) confirms that the agreement for LL is much better than for DD. The very good agreement for LL already with the very first data is good news for $B_s \rightarrow \mu\mu$ search as those are the track types used on this analysis. The reason for the differences in DD is the early status of the OT alignment and pattern recognition, while for LL the geometry is dominated by the VELO, which seems to be very well aligned (see hit distributions in 6.2.3). The worse distributions in DD candidates were also confirmed with the Λ .

Other geometrical variables not being part of the GL are shown in Figure 6-12.

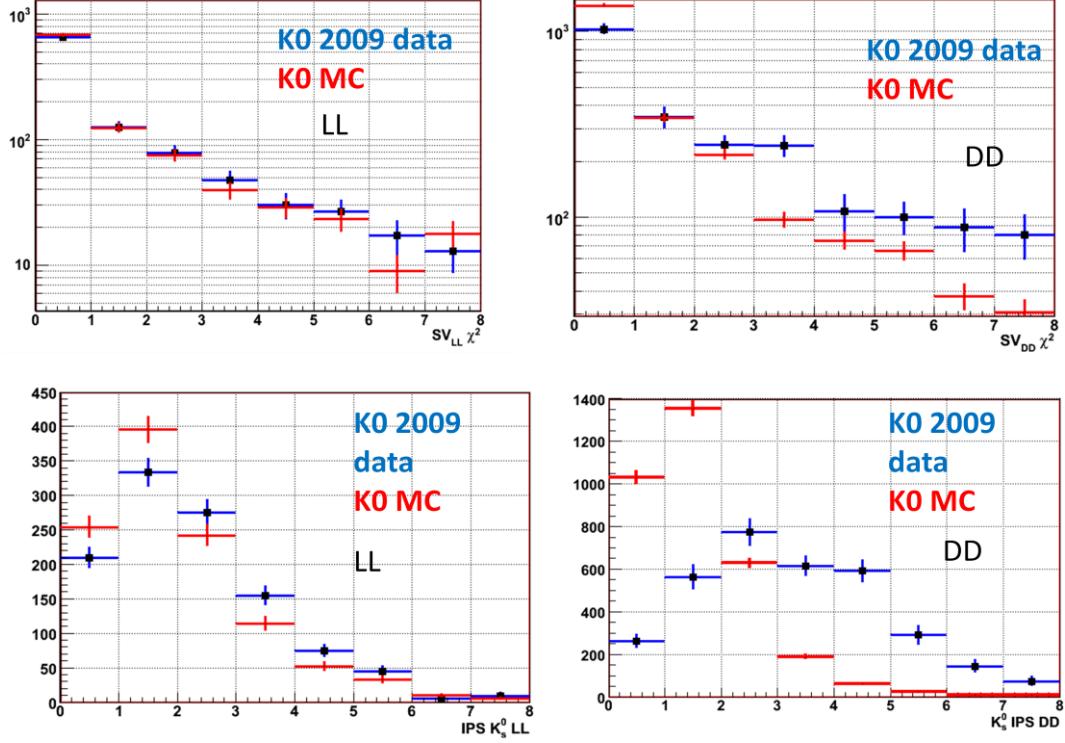


Figure 6-12: IPS and vertex χ^2 of K_s .

Left: for LL. Right: for DD. Top: χ^2 , Bottom: IPS. The blue histogram with black squares is K_s signal in pp collisions data. Red histogram is MC truth K_s signal.

6.2.2. Effective proper time, selection efficiency for LL K_s

The large mean time of the K_s and Λ makes them to flight large distances and decay most of times outside the VELO. This is the main reason of having more DD candidates than LL candidates. This feature of decay lengths being of comparable size as LHCb detector makes the efficiency to decrease at large values of the lifetime, as the decay products for longer lived particles will have less probability to cross enough stations to be reconstructed. Hence, the reconstructed mean time is effectively smaller than what one would get from PDG.

Looking at Figure 6-10, we see the distribution of lifetimes to be still exponential for LL (not the case for DD). The result of the fit to an exponential decay gives us, for the K_s :

$$\frac{\tau_{PDG}}{\tau_{LL}} = 2.84 \pm 0.09$$

In other words, the mean time of LL K_s is effectively three times shorter than the physical one. With this result we can also estimate which is the lack of LL statistics produced because of the selection cuts. From the IP distribution of LL (see Figure 6-9), we expect the cut $IP < 6$ mm to be $\sim 100\%$ efficient on LL prompt. The lifetime cut of 0.1 (PDG) mean times translates then to ~ 0.3 effective LL mean times, so:

$$\varepsilon_{K0_LL}^{SEL/REC} \approx e^{-0.3} = 74\%$$

Combining all tracks we also get exponential PDF's, which allow us to compute the effective mean time for the strange V^0 's in LHCb (see Figure 6-13):

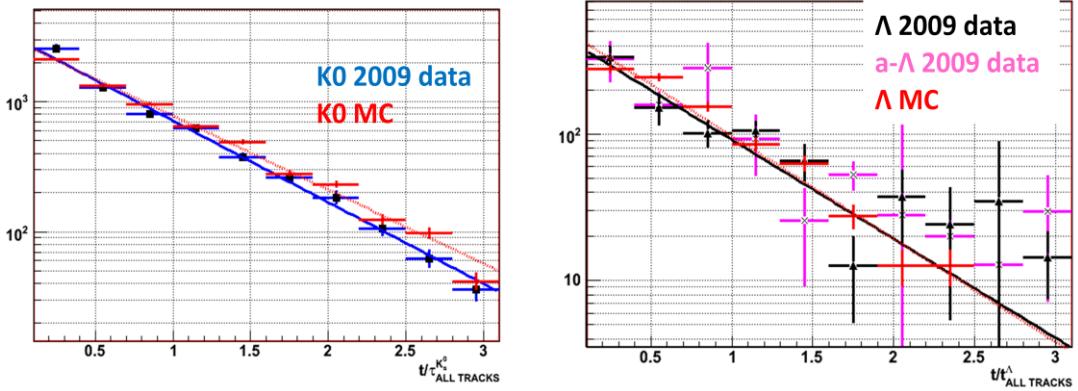


Figure 6-13: Effective mean time in LHCb.

Left: for K_s . Right: for Λ . The blue histogram with black squares is K_s signal in pp collisions data. Black with triangle are Λ . Magenta with small x in the centre is $\bar{\Lambda}$. Red histogram is MC truth signal.

The measured values are hence:

$$\left. \frac{\tau_{PDG}}{\tau_{LHCb}} \right|_{K_s} = 1.44 \pm 0.04, \quad \left. \frac{\tau_{PDG}}{\tau_{LHCb}} \right|_{\Lambda} = 1.6 \pm 0.2$$

Very similar (although not compatible for K_s) to what is seen in MC:

$$\left. \frac{\tau_{PDG}}{\tau_{LHCb}} \right|_{K_s, MC} = 1.31 \pm 0.02, \quad \left. \frac{\tau_{PDG}}{\tau_{LHCb}} \right|_{\Lambda, MC} = 1.6 \pm 0.1$$

i.e, the effective mean time for LHCb reconstructed V0s is approximately 3/2 smaller than the actual one.

6.2.3. Hits distribution

For a better understanding of the different agreement with MC in LL and DD, we have a look at the distribution of the number of hits in the relevant subdetectors, for signal events. The following figures show the sum of hits of both tracks forming the K_s in OT (Figure 6-14), VELO (Figure 6-15) and ST (Figure 6-16).

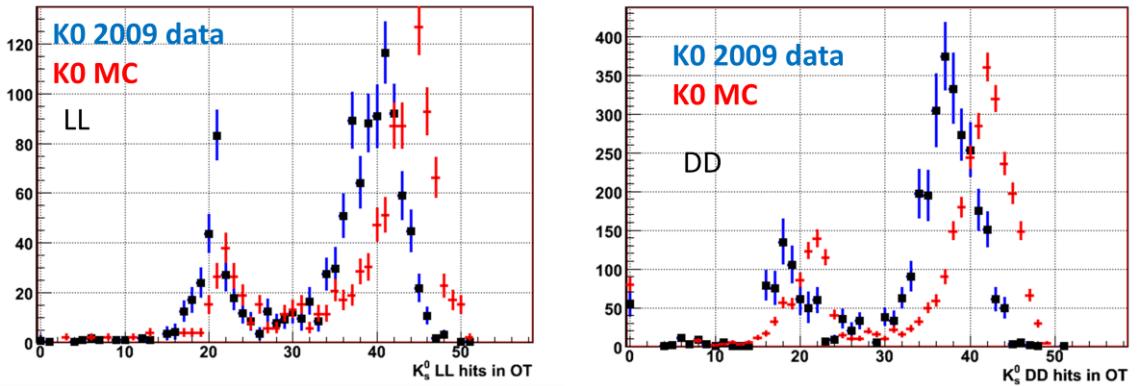


Figure 6-14: Hits in OT from K_s daughters.

Left: for LL. Right: for DD. The blue histogram with black squares is K_s signal in pp collisions data. Red histogram is MC truth K_s signal.

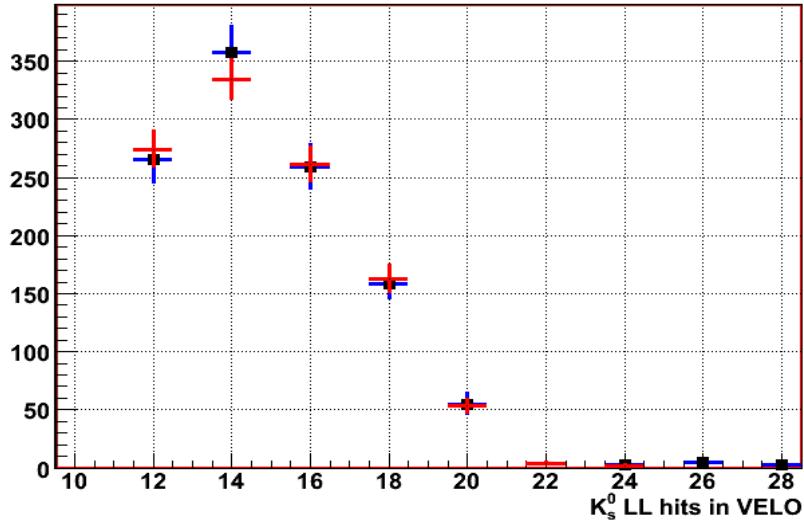


Figure 6-15: Hits in VELO from K_s daughters.

Only LL combinations. The blue histogram with black squares is K_s signal in pp collisions data. Red histogram is MC truth K_s signal.

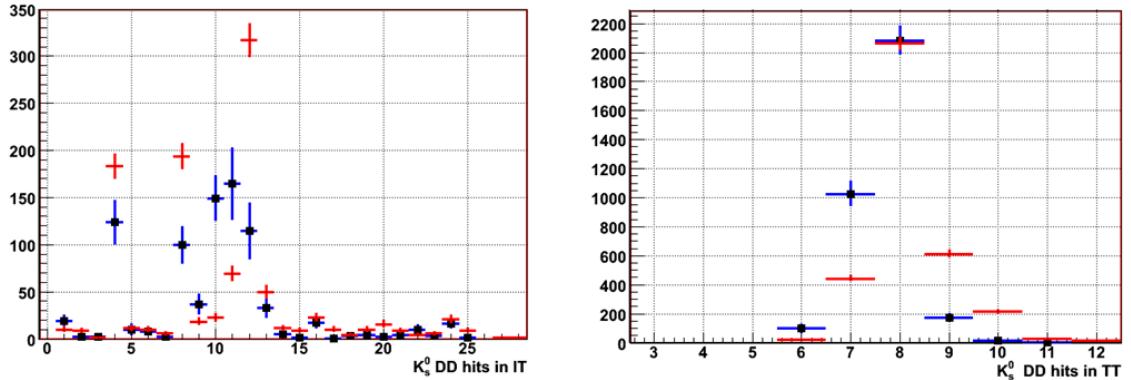


Figure 6-16: Hits in ST from K_s daughters.

Only DD combinations. The blue histogram with black squares is K_s signal in pp collisions data. Red histogram is MC truth K_s signal.
Left: IT, the bin at 0 is not included. Right: TT.

We see that the number of obtained hits in OT is in general smaller in data than in MC, but improvements on alignment and pattern recognition are expected. The distribution of hits in the VELO is very similar to MC and, as the VELO dominates the geometry for LL, this explains the good agreement in the geometrical properties shown in 6.2.1. We also see that here we only get even number of hits, while in Figure 3-16 we had seen a small fraction of long tracks with odd number of VELO hits. This is due to the fact that tracking reconstruction for open VELO requires space points, which need double hits (r and ϕ).

6.2.4. Angular distributions

In this section we will have a look at the angular distributions, to perform a quick comparison of acceptance effects.

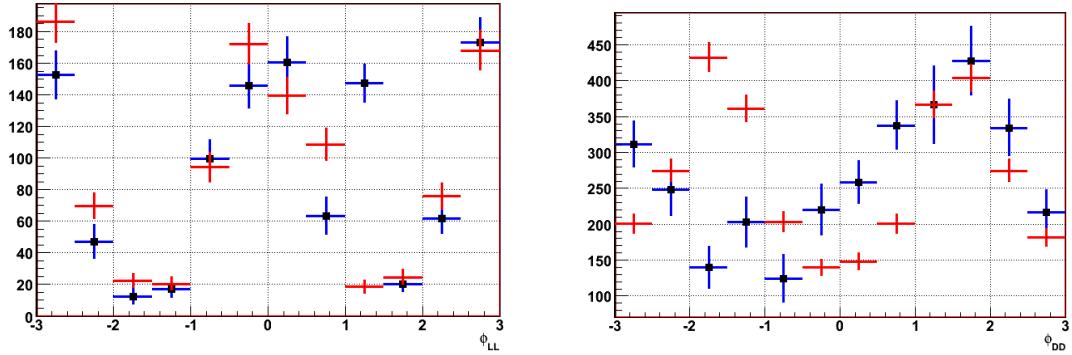


Figure 6-17: ϕ distribution of K_s .

Left: for LL. Right: for DD. The blue histogram with black squares is K_s signal in pp collisions data. Red histogram is MC truth K_s signal.

Figure 6-17 shows the distribution of the azimuthal angle, ϕ , for LL and DD combinations. In case of no acceptance effects, such distribution would be roughly flat, decreasing when going to zero because of the pp crossing angle (see 6.2.5). Hence, in LL we see an acceptance drop at $\pm 90^\circ$ which has to do with the RF foil and with the fact that the VELO is not closed. However, even in DD we see a drop at -90° which cannot be explained by the VELO and that moreover does not appear in MC. This acceptance drop was related to some TT module not yet properly aligned [116]. Figure 6-18 shows the distribution on pseudorapidity, η , where we see that LHCb covers approximately from 2 to 5. In the case of LL, η acceptance drops already at 4, both in data and MC.

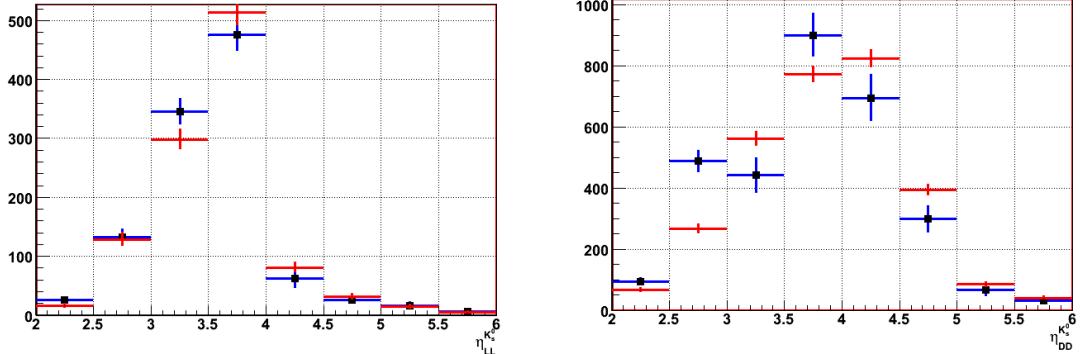


Figure 6-18: η distribution of K_s .

Left: for LL. Right: for DD. The blue histogram with black squares is K_s signal in pp collisions data. Red histogram is MC truth K_s signal.

The $\bar{\Lambda}/\Lambda$ ratio as a function of η is shown in Figure 6-19, where the particles were selected with a lifetime cut $t/\tau > 1/2$. The ratio seems to increase with the pseudorapidity, which could be explained in principle by acceptance effects[117].

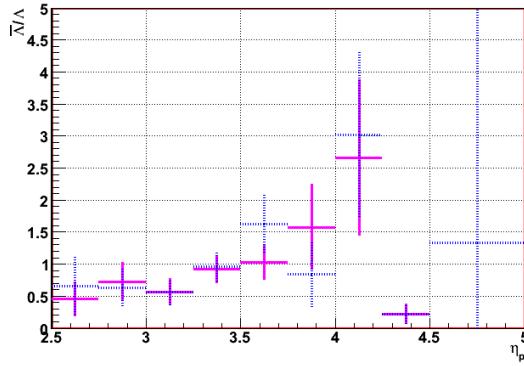


Figure 6-19: $\bar{\Lambda}/\Lambda$ as a function of η in laboratory system.

The magenta histogram was obtained with single Gaussian fit. The dashed blue histogram corresponds to double Gaussian.

6.2.5. Angular distributions in Centre of Mass

Due to the pp crossing angle being different from 0, the laboratory system in LHCb is not exactly the centre of mass system. In the conditions of 2009 data (injection energy without ramping), the protons are crossing with a -2.1 mrad angle in the XZ plane thus the sum of 3-momenta is not exactly 0. The total 4-momentum to which the particles have to be boosted in order to get centre of mass distributions is then:

$$\Sigma p^\mu = \left(2\sqrt{m_p^2 + (450 \text{ GeV})^2}, (2 \sin(-2.1 \cdot 10^{-3}) 450 \text{ GeV}, 0, 0) \right)$$

The distributions in centre of mass are the ones usually used for studies of production models, and in particular the ones of strange particles are of particular interest as probe of fragmentation (there is no valence strange quarks in the initial state). The effect of boosting to CM can be not negligible, as it is shown in Figure 6-20, where the increment in the angular variables is plotted as a function of the value in the laboratory frame.

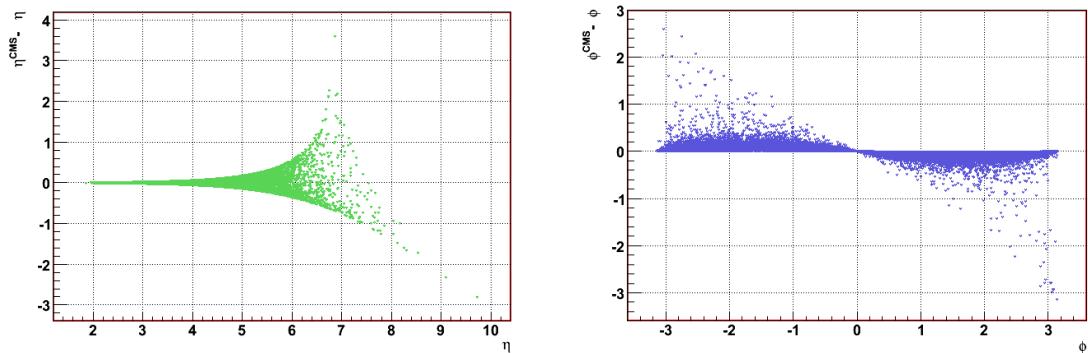


Figure 6-20: Effect of boost to pp CM system.

Left: Increment in pseudorapidity as a function of the value in the laboratory frame. Right: Increment in the azimuthal angle as a function of its value in the laboratory frame. The data corresponds to K_s selected candidates in 2009 data, beam-gas is not subtracted.

In this section we will compare the distributions of the different species and the ratios in bins of the angular variables in centre of masses. The acceptance is not deconvoluted as it is not yet known with precision.

To facilitate the fitting of Λ mass distribution in all the bins, a cut in the GL defined for DD K_s was used. Both particles have similar geometrical properties so the K_s GL also works for selecting Λ . The cut chosen is $GL > 0.2$, which has an efficiency of 61 % for both Λ and $\bar{\Lambda}$ in MC. Evaluating it in data, we obtain $54 \pm$

8 % for Λ and $38 \pm 8\%$ for $\bar{\Lambda}$, where the compatibility with being the same for baryon and anti-baryon is dubious.

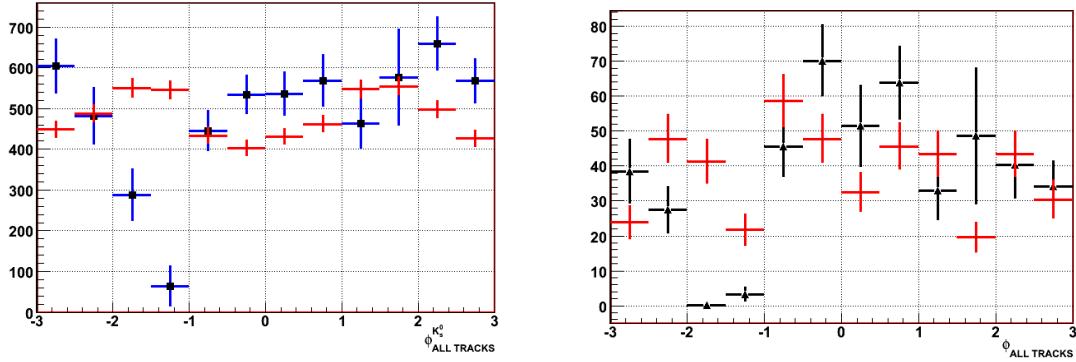


Figure 6-21: ϕ distribution of K_s and Λ in CM system.

Left: for K_s . Right: for Λ (anti baryon not included). The blue histogram with black squares is K_s signal in pp collisions data, Black histogram with triangles is Λ . Red histogram is MC truth signal.

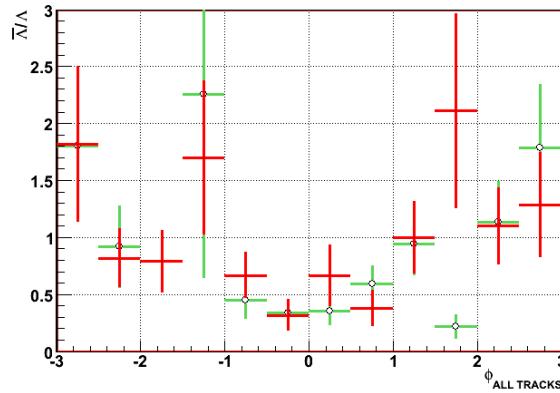


Figure 6-22: $\bar{\Lambda}/\Lambda$ as a function of ϕ in CM system.

Green with open circles: 2009 pp collision data. Red histogram: MC.

Figure 6-21 shows the distribution of ϕ in CM system. Those distributions must be flat if no acceptance effect, due to the azimuthal symmetry of the collision. However we see that we obtain not flat distributions, in particular we see the drop at -90° observed in the laboratory system (see 6.2.4).

The fact that the acceptance in ϕ is also different for positive and negative particles may cause the $\bar{\Lambda}/\Lambda$ ratio to effectively vary with ϕ , as seems to happen in Figure 6-22. The ratio seems in better agreement with MC than the Λ and $\bar{\Lambda}$ distributions, probably due to a partial cancelation of acceptance effects.

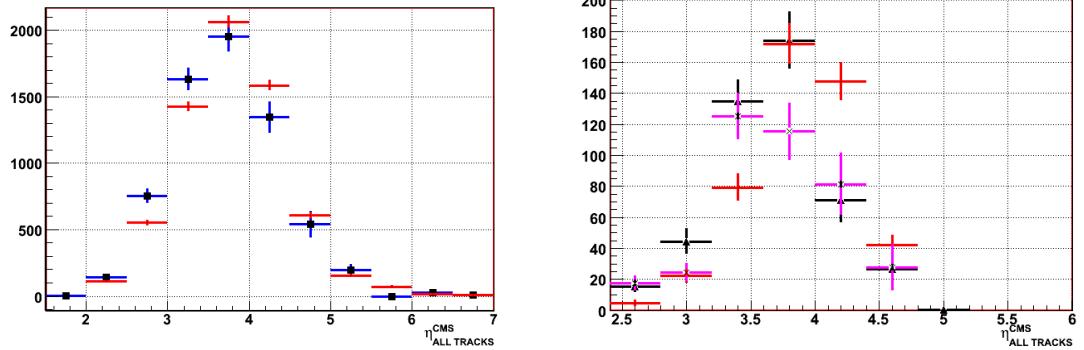


Figure 6-23: η distribution of K_s and Λ in CM system.

Left: for K_s . Right: for Λ . The blue histogram with black squares is K_s signal in pp collisions data. Black histogram with triangles is Λ . Magenta with crosses is $\bar{\Lambda}$. Red histogram is MC truth signal.

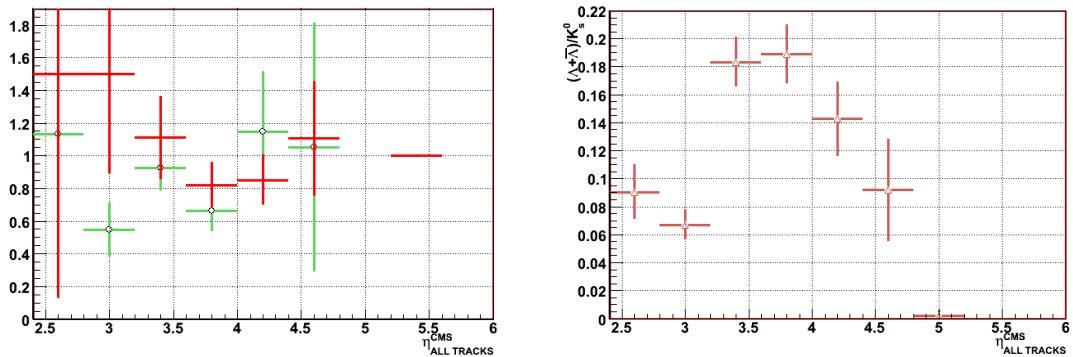


Figure 6-24: Ratios as a function of η .

Green with open circles: $\bar{\Lambda}/\Lambda$ in 2009 pp collision data. Brown with open triangles: $(\bar{\Lambda} + \Lambda)/K_s$ in 2009 pp collision data. Red histogram: MC.

Figure 6-24 shows the distribution of η in CM system. The distributions are similar to those obtained from MC, but inconsistent. Those discrepancies are most likely due to the early status of detector alignment, and hence to the description of the detector in MC. The $\bar{\Lambda}/\Lambda$ ratio seems compatible within the large errors.

6.3. Background studies

6.3.1. Geometrical properties of LL background

Here the properties of the combinatorial background for LL K_s are studied. In order to remove actual V^{0° 's from the sample, we take only the K_s sidebands $M < 470$ MeV and $M > 530$ MeV. The Λ are eliminated from the background sample by requiring p_t in the AP plot to be larger than 130 MeV (see Figure 6-25). First, we check the total amount of background with respect to signal in the LL K_s sample following the above definition:

Table 6-VI: Amount of combinatorial background for this test.

Sample	# Combinatorial

Bb	1585
Eb	153
MC	428

So, using the normalization factor in 6.1.3.1, the expected number of combinatorial LL background in pp collisions is 1432 ± 42 . Adding the information from Table 6-III and Table 6-IV we get, at selection level:

$$\frac{\text{Combinatorial}/\text{signal}|_{\text{data}}}{\text{Combinatorial}/\text{signal}|_{\text{MC}}} = 1.85 \pm 0.15$$

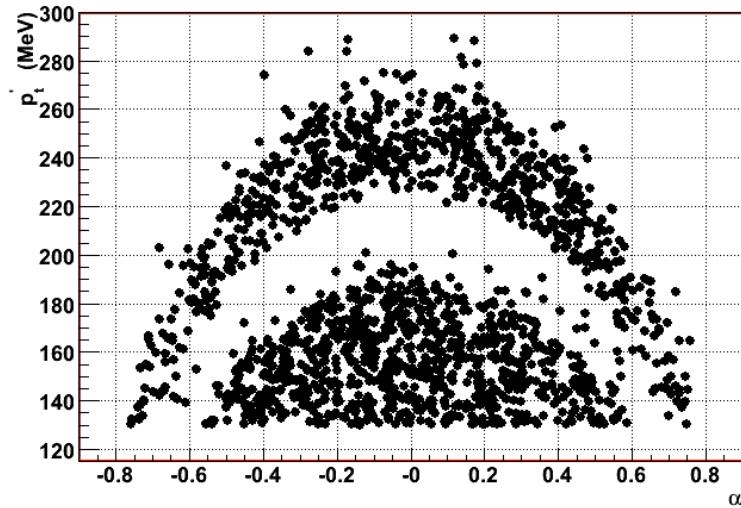


Figure 6-25: Armenteros-Podolanski selection for combinatorial background.

The selection of combinatorial background as detached vertices is mainly driven by the lifetime. The result of fitting the lifetime of the combinatorial background gives us

$$\left. \frac{\tau_{PDG}}{\tau_{LL}} \right|_{bkg} = 12.3 \pm 0.3$$

Doing the same in the MC sample, we obtain:

$$\left. \frac{\tau_{PDG}}{\tau_{LL}} \right|_{bkg, MC} = 11.2 \pm 0.6$$

So the lifetime looks very similar to MC, and it is about four times shorter than the mean time of LL K_s signal. It is worth to look at the full GL, to see the overall geometrical similarity. This is shown in Figure 6-26.

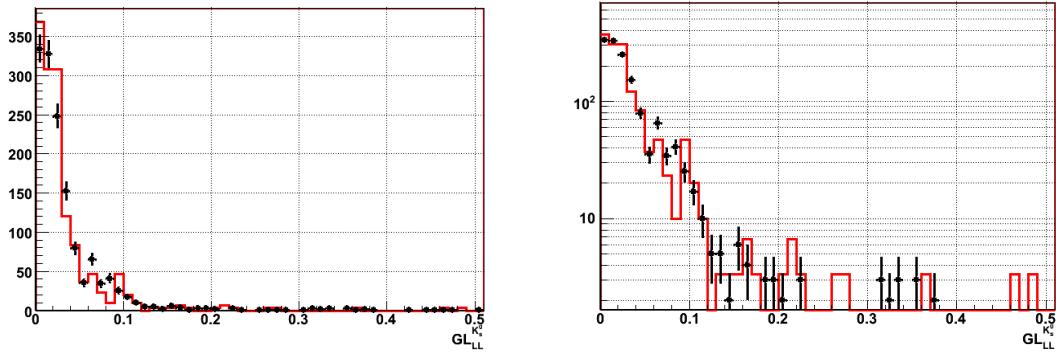


Figure 6-26: GL distribution for K_s combinatorial background.

Red histogram: MC. Black points: data. Left: Normal scale. Right: Logarithmic scale.

6.3.2. Muon misid

The probability to misidentify hadrons as muons is a very important issue for $B_s \rightarrow \mu\mu$ analysis. The V^0 's used here give us a clean sample of hadrons where the PID is known a priori. Looking for signal where the hadron has the ISMUON flag gives us the probability to misidentify them (including also the effect of acceptance of muon chambers). As the misidentified signal is rather small, to increase the significance over background we cut harder on the lifetime, up to $1/2$ PDG mean times.

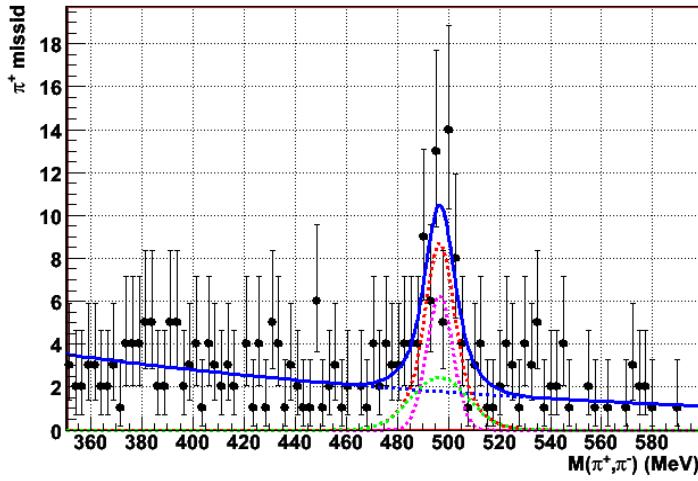


Figure 6-27: K_s with positive pion flagged as IsMuon.

Then, the measured misid probabilities are:

- $\pi^+ (K_s)$: $1.7 \pm 0.3 \%$ (MC: $1.0 \pm 0.2 \%$)
- $\pi^- (K_s)$: $1.6 \pm 0.4 \%$
- $\pi^+ (\bar{\Lambda})$: $1.0 \pm 1.0 \%$
- $\pi^- (\Lambda)$: $\lesssim 0.7 \%$
- $p (\Lambda)$: $< 0.5 \%$

- $\bar{p}(\bar{\Lambda}) = 0.5 \pm 0.4 \%$

So the pion misid evaluated at momenta of ~ 5 GeV (see spectrum in Figure 6-28) is, within errors, compatible with MC, and at most a factor 2-2.5 worse. Misid from protons remains unseen.

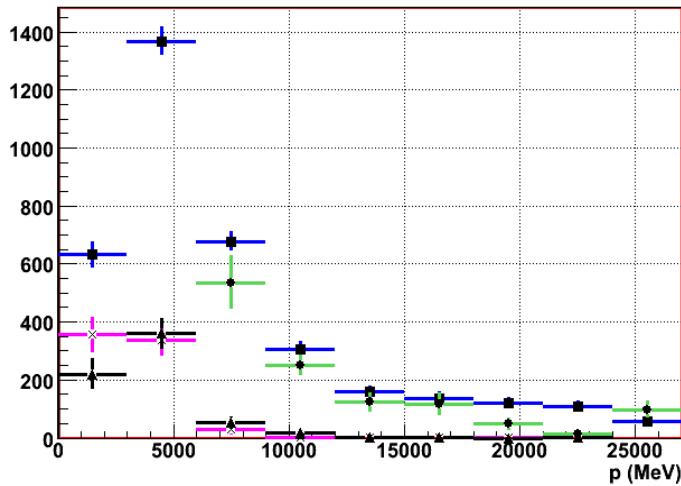


Figure 6-28: Momentum spectra for V^0 daughters.

The blue histogram with black squares is K_s signal in pp collisions data. Black with triangle are pions from Λ . Magenta with small x in the centre are pions from $\bar{\Lambda}$. Green with circles are protons from Λ .

7. Conclusions

The strategy to extract the $B_s \rightarrow \mu^+ \mu^-$ branching ratio from LHCb data has been shown, calibrating all the steps using control channels and not relying on the simulation. The analysis has also been illustrated with an example of a hypothetical evidence for 150 pb^{-1} . A multivariate analysis has been developed to extract this BR, and has been shown to be also useful for the fast generation of sophisticated toy MC samples.

This analysis allows LHCb to have the best performance on this search and with the first 300 pb^{-1} (the luminosity expected for 2010) it could overtake Tevatron's limit having already an important impact on New Physics searches. With one nominal year, LHCb can exclude any enhancement with respect to the SM prediction. Within five nominal years, LHCb could observe values even smaller than SM prediction.

The first data stored by LHCb experiment at the end of 2009 (about 300k pp collisions at 900 GeV) has been also studied showing good agreement geometrical with MC in invariant mass and geometrical properties for V^0 's made with two long tracks (LL). The muon misid probability evaluated with V^0 daughters (protons and pions with momenta around 5 GeV) is compatible with MC, and not much larger than a factor two. The combinatorial background in LL K_S has been found to be approximately a factor two larger than in simulation, but with very similar geometrical properties. These results are very promising to validate the potential of the analysis obtained in simulation.

8. Appendix

We will try to find some specific features for the case in which the decorrelation works perfectly.

If there is no remaining correlation after the rotation, in q_1, q_2 space the p.d.f can be written as:

$$\rho_q(q_1, q_2) dq_1 dq_2 = f_1(q_1) f_2(q_2) dq_1 dq_2$$

Being $f_{1,2}$ individual p.d.f's (i.e: $f_i(x) \geq 0, f_i(|x| \rightarrow \infty) \rightarrow 0$).

Thus, in $\{G_x, G_y\}$, as the rotation transformation produces a $J=1$ Jacobian:

$$\rho_G(G_x, G_y) dG_x dG_y = f_1\left(\frac{1}{\sqrt{2}}(G_x + G_y)\right) f_2\left(\frac{1}{\sqrt{2}}(G_x - G_y)\right) dG_x dG_y$$

Imposing the condition of Gaussian projections:

$$\int_{-\infty}^{\infty} \rho_G(G_x, G_y) dG_x = C e^{-\frac{1}{2}G_y^2}, \quad \int_{-\infty}^{\infty} \rho_G(G_x, G_y) dG_y = C e^{-\frac{1}{2}G_x^2}$$

One gets

$$\begin{aligned} \int_{-\infty}^{\infty} f_1\left(\frac{1}{\sqrt{2}}(G_x + G_y)\right) f_2\left(\frac{1}{\sqrt{2}}(G_x - G_y)\right) dG_x &= C e^{-\frac{1}{2}G_y^2}, \\ \int_{-\infty}^{\infty} f_1\left(\frac{1}{\sqrt{2}}(G_x + G_y)\right) f_2\left(\frac{1}{\sqrt{2}}(G_x - G_y)\right) dG_y &= C e^{-\frac{1}{2}G_x^2} \end{aligned}$$

Taking the first one, replacing $\frac{1}{\sqrt{2}}(G_x + G_y) \equiv t$, and absorbing any overall constant in C

$$\int_{-\infty}^{\infty} f_1(t) f_2(t - \sqrt{2}G_y) dt = C e^{-\frac{1}{2}G_y^2}$$

Applying similar replacement to the second equation we get that $f_{1,2}$ must be even functions. If we express both $f_{1,2}$ in the Fourier space, the equation above becomes:

$$F_1(\omega) F_2(\omega) = C e^{-\omega^2}$$

One solution is just $F_1(\omega) = A e^{-a\omega^2}$ and $F_2(\omega) = \frac{C}{A} e^{-(1-a)\omega^2}$, i.e, two Gaussian of different width that would drive us to the solution postulated in 4.2.1.1. But it does not mean it is the most general one, in principle it can be generalized to $F_1(\omega) = A e^{-a\omega^2 + \varphi(\omega)}$ and $F_2(\omega) = \frac{C}{A} e^{-(1-a)\omega^2 - \varphi(\omega)}$ although after several trials no $\varphi(\omega)$ was found such that $f_{1,2}$ were both valid as p.d.f's (in particular positive in all the range). Considering then the solution for $\varphi(\omega) = 0$ and undoing the Fourier transform, one gets:

$$f_1(q_1) = C_1 e^{-\frac{1}{2}(\frac{q_1}{\sigma_1})^2}, \quad f_2(q_2) = C_2 e^{-\frac{1}{2}(\frac{q_2}{\sigma_2})^2} \text{ with } \sigma_1^2 + \sigma_2^2 = 2$$

Finally, imposing normalization:

$$\rho_q(q_1, q_2) dq_1 dq_2 = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}(\frac{q_1}{\sigma_1})^2} e^{-\frac{1}{2}(\frac{q_2}{\sigma_2})^2} dq_1 dq_2, \quad \text{with } \sigma_1^2 + \sigma_2^2 = 2$$

$$\rho_G(G_x, G_y) dG_x dG_y = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{4}(G_x^2 + G_y^2)(1/\sigma_1^2 + 1/\sigma_2^2) + \frac{1}{2}(1/\sigma_2^2 - 1/\sigma_1^2)G_x G_y} dG_x dG_y$$

The value of σ can be taken from the eigenvalues of correlation matrix. The off-diagonal term is:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G_x G_y \rho_G(G_x, G_y) dG_x dG_y = 1 - \sigma_2^2$$

Thus:

$$\begin{vmatrix} 1 - \lambda & 1 - \sigma_2^2 \\ 1 - \sigma_2^2 & 1 - \lambda \end{vmatrix} = 0 \Rightarrow \begin{cases} \lambda_1 = \sigma_2^2 \\ \lambda_2 = 2 - \sigma_2^2 = \sigma_1^2 \end{cases}$$

9. Resumo

O *Standard Model* da física de partículas é unha teoría que describe as interaccións fundamentais entre as partículas e leva demostrado un gran poder predictivo dende a súa formulación nos anos 70. Mais o *Standard Model* non é una teoría completa, xa que non inclúe gravidade, non dá unha explicación á composición da materia escura e postula que os neutrinos non teñen masa o que contradí as medidas experimentais. Na actualidade hai formuladas varias extensións posibles do *Standard Model* e que proporcionan solucións aos seus defectos. Entre estas ideas atópase a supersimetría, as dimensións extra, a unificación das diferentes interaccións nunha soa, ou a formulación de novas interaccións. Hai, polo tanto, diferentes modelos alternativos ao *Standard Model* e son precisas mais medidas experimentais para descartar aqueles que resulten erróneos.

Os experimentos no LHC do CERN, e en particular LHCb, analizarán una gran cantidade de datos que porán a proba o *Standard Model* e as súas extensións. Esta tese presenta a estratexia de análise do decaimento raro $B_s \rightarrow \mu^+ \mu^-$ no experimento LHCb. A tasa de desintegración de $B_s \rightarrow \mu^+ \mu^-$ é moi sensible ás distintas extensións do *Standard Model*, e o seu estudo podería ter un impacto importante na búsqueda de Nova Física usando os datos de LHCb no 2010.

9.1. LHC

O colisionador LHC (*Large Hadron Collider*) é un acelerador circular de aproximadamente 27 km que está a producir datos a finais do ano 2009 no CERN (*Organisation Européenne pour la Recherche Nucléaire*, orixinariamente *Conseil Européen pour la Recherche Nucléaire*), laboratorio de física de partículas con sede en Suiza. O LHC está diseñado para producir millóns de colisións protón protón a unha enerxía en centro de masas de 14 TeV. A análise de ditas colisións permitirá poñer a proba varias das alternativas ao *Standard Model*, así como atopar a única partícula deste que ainda non foi observada, o bosón de Higgs.

Ademais de colisións protón – protón, o LHC tamén producirá choques de núcleos pesados (Pb). Sobre a producción das partículas iniciais poden consultarse¹⁵ e¹⁶.

Antes de alcanzar a enerxía final de 7 TeV, os protons son acelerados pola seguinte secuencia (fig I.I.2):

- LINAC (Linear ACcelerator): ata 50 MeV
- Booster: ata 1.4 GeV
- PS (Proton Synchrotron): ata 26 GeV
- SPS (Super Proton Synchrotron): ata 450 GeV

9.1.1. Experimentos do LHC

As colisións producidas no LHC serán analizadas polos seguintes experimentos (fig I.I.3):

- **ALICE** (*A Large Ion Collider Experiment*)¹⁷: Experimento adicado ao estudo de colisións nucleares, co obxectivo de analizar as transicións de fase entre a materia nuclear e o plasma de quarks e gluons.

¹⁵ <http://linac2.home.cern.ch/linac2/sources/source.pdf>

¹⁶ <http://linac2.home.cern.ch/linac2/seminar/seminar.pdf>

¹⁷ ALICE Collaboration, “A Large Ion Collider Experiment, Technical Proposal”, CERN/LHCC/95-71, LHCC/P3

- **ATLAS** (*A Toroidal LHC Appratus*)¹⁸: Experimento adicado á procura do bosón de Higgs do *Standard Model* (SM) ou outros modelos, física de quarks *b* e *t*, e, en xeral, sinais de nova física coma, por exemplo, supersimetría (SUSY).
- **CMS** (*Compact Muon Solenoid*)¹⁹: Experimento de propósito xeral, ao igual que ATLAS e en competición directa con el.
 - **TOTEM**²⁰: Situado xunto a CMS no punto de interacción IP5, pretende determinar a sección eficaz total de colisión no LHC, así como a sección inelástica e procesos difractivos, permitindo a monitorización da luminosidade no IP5.
- **LHCb** (*Large Hadron Collider beauty experiment*)²¹ : Experimento adicado á física dos quarks *b* : violación da simetría CP e decaimentos raros de mesóns *B*.

9.2.LHCb

O LHCb é un dos catro experimentos principais que están a tomar datos no colisionador LHC do CERN. LHCb é una colaboración de aproximadamente 600 persoas de diversos países e institucións. O detector utilizado é un espectrómetro de baixo ángulo, semellante aos detectores de branco fixo, ainda que as colisións estudiadas son en centro de masas. Isto débese ao facto de producírense os quarks *b* a baixo ángulo, nas colisións p-p que teñen lugar no acelerador. A aceptancia do detector é de 300 mrad no plano XZ e de 250 mrad no plano YZ, perdéndose polo tanto os *b*'s producidos na dirección $Z < 0$. A xeometría do detector (a diferenza do caso da xeometría de barril) tamén presenta a avantage de poder instalar detectores RICH, que permiten unha clara separación entre kaóns e pións, esencial para a análise de moitos dos canais de interés.

O experimento LHCb está pensado para traballar a unha luminosidade instantánea de $2 - 5 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$, menor que a nominal do LHC, e que se consegue defocalizando os feixes de protóns nas proximidades do punto de interacción. A razón disto é que a correcta identificación dos vértices primario (PV) e secundario (SV), é esencial para a medida dos parámetros que se pretenden determinar; dado que o número de vértices primarios medra coa luminosidade (sendo aproximadamente 23 en promedio para a luminosidade nominal do LHC), reduciuse de maneira que se maximizase a probabilidade de obter tan só un único PV. O feito de seres maior a radiación (e polo tanto o deterioro do dos detectores) a baixo ángulo tamén fai que unha luminosidade menor sexa mais adecuada.

9.2.1. Subdetectores do LHCb

O detector LHCb está formado por:

- **VELO** (*VErtex LOcator*)²² : Detector de silicio no entorno do punto de interacción, para a identificación dos vértices primarios e secundarios
- **RICH** (*Ring Imaging Cherenkov Detector*)²³ 1 e 2: Permite determinar a velocidade das partículas cargadas valéndose do efecto Cherenkov. Coñecido o momento das mesmas, pode estimarse o valor das masas e polo tanto mellorar notablemente a identificación de partículas.
- **TT** (*Trigger Tracker*) e **Tracking Stations**^{24, 25}: Detectores de silicio/gas que, xunto á información do VELO, permiten determinar as traxectorias das partículas. A presencia do imán (*magnet*)²⁶ permite determinar con precisión os momentos das mesmas.

¹⁸ ATLAS Collaboration, “ATLAS Technical Proposal”, CERN/LHCC/94-43, LHCC/P2

¹⁹ CMS Collaboration, “The Compact Muon Solenoid, Technical Proposal”, CERN/LHCC/94-38, LHCC/P1

²⁰ TOTEM Collaboration, “TOTEM, Total Cross Section, Elastic Scattering and Diffraction Disociation at LHC: Technical Proposal”, CERN/LHCC/99-07, LHCC/P5.

²¹ LHCb Collaboration, “LHCb Technical Proposal”, CERN/LHCC/98-1, LHCC/P4

²² http://lhcb-vd.web.cern.ch/lhcb-vd/TDR/velo_tdr.pdf

²³ <http://lhcb-rich.web.cern.ch/lhcb-rich/richtdr/tdr.pdf>

- SPD (*Scintillator Pad Detector*), PS(*PreShower*), ECAL (*Electromagnetic CALorimeter*) e HCAL (*Hadronic CALorimeter*)²⁷: A excepción dos muons, as partículas detectadas en LHCb son frenadas polos materiais destos subdetectores, medíndose a enerxía depositada neles.
- Muon Stations²⁸: Detectores de gas intercalados con filtros muónicos (que deteñen ás partículas que poidesen ter pasado os calorímetros). Permiten diferenciar muons doutras partículas cargadas.

9.2.2. Reconstrucción de Vértices

Mediante o VELO, situado a 8 mm do punto de interacción, os vértices primarios e secundarios son reconstruídos con gran precisión, obténdose una resolución excelente en parámetros de impacto (IP). A resolución típica na posición do PV (en promedio, unhas 70 trazas no VELO) é de 47 μm na dirección Z e 8 μm en X e Y.

Debido a que os mesóns B voan apreciablemente en LHC (~ 1 cm), as trazas procedentes do correspondente SV están, en xeral, separadas do PV (i.e., un rasgo distintivo son altos valores de IP).

9.2.3. Medida do momento

A medida do momento realiza-se a partir da curvatura (inducida por un campo integrado de 4 Tm, xerado polo imán do detector) das trazas reconstruídas mediante o VELO, TT e Track Stations. A resolución na rexión de momento propia das trazas procedentes de desintegracións de B's é de 0.35 - 0.55 %, levando a unha resolución en masa invariante de 18 MeV, esencial para a separación Bd - Bs, e de gran utilidade á hora de reducir a ventá de búsqueda dos decaimentos rares e polo tanto o nivel de ruído.

9.2.4. Identificación de Partículas

A identificación de partículas (PID) en LHCb codifícase en forma de DLL's, logaritmo neperiano do cociente de densidades de probabilidade de dúas hipóteses diferentes, por exemplo, no caso muon - pión:

$$DLL(\mu - \pi) = \ln\left(\frac{\rho_\mu}{\rho_\pi}\right)$$

combinando informacións dos RICH, calorímetros e cámaras de muons. Ademais, no caso particular dos muons, requírense sinais en diferentes rexións do sistema de muons, chamadas FOI's (*Fields Of Interest*). O número, posición e tamaño destas rexións dependen do momento da partícula candidata.

A calibración do PID será estudiada no decaimento $D^{*+} \rightarrow D^0\pi^+$, onde $D^0 \rightarrow K^-\pi^+$ debido á súa abundante producción e clara selección (a pequena diferencia de masas $D^{*+} - D^0$ apenas deixa espacio fásico disponible ó pión), ademais, o modo $D^0 \rightarrow K^+\pi^-$ está doblemente suprimido por Cabibbo sendo ~ 400 veces menor que $D^0 \rightarrow K^-\pi^+$. Así, neste decaimento cada traza ten unha pureza maior que o 99 % sen necesidade do RICH.

²⁴ <http://lhcb.web.cern.ch/lhcb/TDR/front%20cover/LHCb-IT-TDR.pdf>

²⁵ <http://www.nikhef.nl/pub/experiments/bfys/lhcb/outerTracker/tdr/final.pdf>

²⁶ <http://lhcb.web.cern.ch/lhcb/magnet/TDR/pdf/lhcb-magnet.pdf>

²⁷ http://lhcb.web.cern.ch/lhcb/calorimeters/html/TDR/calo_tdr.pdf

²⁸ <http://lhcb-muon.web.cern.ch/lhcb-muon/results/tdr.pdf>

9.2.5. Trigger

O número de veces no que se producen alomenos dous trazas na aceptancia de LHCb é aproximadamente de 10 millóns por segundo (10 MHz), mais só 2 KHz serán finalmente almacenados. O trigger (fig I.VII.1) é o sistema encargado de seleccionar os sucesos, baseándose en sinais características de decaimentos de B 's (trazas separadas do PV, partículas de alto momento transverso...). Consta de dousas etapas Level 0 (L0) e High Level Trigger (HLT). A primeira é unha etapa puramente electrónica, mentres que a segunda consiste nunha serie de algoritmos que correrán nunha granxa de aproximadamente 1800 CPU's.

O L0 deberá tomar 40 millóns de decisiones por segundo reducindo a cantidad de sucesos seleccionados a 1 MHz, que virá a ser a entrada do HLT. Á velocidade á que debe tomar as decisions o L0 non é posible ler ao completo o detector, así que a decision baséase só nos calorímetros – sinais de alta enerxía na dirección transversa ao feixe – e nas cámaras de muons – alto pt, baixo a hipótese de que o muon provén da rexión cercana ao punto de interacción).

Se un suceso é seleccionado polo L0 será procesado no HLT nunha das diferentes *alleys* (secuencias de algoritmos) dependendo da decisión do propio L0. Durante a execución do HLT, danse diferentes niveis de reconstrucción (aumentando a información lida do detector e polo tanto tamén a precisión de certas cantidades como o momento) segundo un suceso vai sendo aceptado polos diferentes algoritmos dunha *alley*. As eficiencias do HLT para os canais de estudio en LHCb son aproximadamente 80 – 90 % (respecto a sucesos seleccionados *offline*).

9.3. Análise do decaimento raro $B_s \rightarrow \mu^+ \mu^-$

9.3.1. Motivación para o estudo de $B_s \rightarrow \mu^+ \mu^-$

Un dos decaimentos raros de maior interese é $B_s \rightarrow \mu^+ \mu^-$ cuxa fracción de desintegración e tan só de $(3.35 \pm 0.32) \times 10^{-9}$ no *Standard Model*, poidendo ser moito maior no caso de Nova Física. A desintegración $B_d \rightarrow \mu^+ \mu^-$, que pode ser estudiada coa mesma estratexia de análise, tamén é sensible a Nova Física, pero ten a desvantaxe de ter una tasa de desintegración moito menor e polo tanto requírese un maior tempo de toma de datos para o seu estudo. Un feito importante é que, inda que moitos modelos de Nova Física predín os mesmos efectos en $B_s \rightarrow \mu^+ \mu^-$ e $B_d \rightarrow \mu^+ \mu^-$, tamén hai modelos que predín efectos diferentes.

Un breve esquema de cómo os diferentes escenarios de Nova Física afectarían á desintegración $B_s \rightarrow \mu^+ \mu^-$ sería (resumo da sección 2.3):

- Varios modelos supersimétricos tenden a predecir valores maiores desta tasa de desintegración, poidendo ser tan altos como o límite experimental obtido polo experimento CDF do Tevatron. A supersimetría tamén permite valores menores que o predito polo *Standard Model*, dependendo dos signos relativos entre as distintas contribucións.
- Varios modelos de dimensíons extras son explicados na sección 2.3.9. Estos escenarios tamén permiten desviacións do valor da tasa de desintegración $B_s \rightarrow \mu^+ \mu^-$ con respecto á predicción do *Standard Model*. Tamén é salientable o feito de que as contribucións de Nova Física nestes modelos afectan en xeral de forma distinta a $B_s \rightarrow \mu^+ \mu^-$ que a $B_d \rightarrow \mu^+ \mu^-$.
- Varios modelos de referencia nos que o Higgs non é elemental senón un composto que aparece como resultado dalgunha nova interacción, son explicados nas seccións 2.3.12 (*Topcolor assisted technicolor*, TC2) e 2.3.15 (*Littlest Higgs with T parity*, LHT). Nestes modelos, a tasa de desintegración $B_s \rightarrow \mu^+ \mu^-$ admite pequenas desviacións respecto á predicción do *Standard Model*, mais bastante restrinxidas debido a que estas contribucións deberían afectar tamén a outros procesos xa medidos e nos que non se observaron discrepancias significativas respecto ás prediccións do *Standard Model*.

9.3.2. Estratexia da Análise

No caso de que o valor desta tasa de desintegración sexa similar á prerita polo *Standard Model*, LHCb disporía duns 70 sucesos en cinta por ano nominal de toma de datos. Estos 70 sucesos deberán ser identificados entre os millóns de posibles candidatos almacenados.

O presente traballo descrebe a análise deste decaimento no experimento LHCb. A análise está baseada na clasificación de candidatos a $B_s \rightarrow \mu^+ \mu^-$ nun espazo de 3 parámetros: a masa invariante, a probabilidade de ser muóns reais frente a outras hipótesis de identidade das partículas fillas, e unha combinación de varias variables xeométricas chamada GL (de *Geometrical Likelihood*).

A GL é a principal ferramenta de separación entre a sinal verdadeira e o fondo. Baséase nun procedimento matemático que ten en conta non só os valores das variables en sí, mais tamén as súas correlacións. Este procedemento preséntase na sección 4.2.1, e foi desenrolado durante o verán de 2006 específicamente para ser usado nesta análise. A idea principal consiste en tomar as variables iniciais e facer dousas transformacións. As variables resultantes da primeira transformación terían, no caso de tratarse de sucesos de sinal, una distribución gaussiana sen correlacións entre elas. As variables resultantes da segunda transformación terían esas mesmas propiedades pero para una mostra de sucesos de fondo. Deste xeito, o resultado da primeira transformación danos o χ^2 respecto á hipótese de sinal, e a segunda danos o χ^2 respecto á hipótese de fondo. A diferenza entre estes dous χ^2 é a variable discriminatoria que utiliza a GL.

Tódolos sucesos almacenados en cinta e que superan una serie de criterios de semellanza a decaimentos $B_s \rightarrow \mu^+ \mu^-$ son entón clasificados neste espacio de 3 parámetros.

O uso de canais de control permite calibrar as propiedades do sinal que se pretende estudiar, e obter a distribución de $B_s \rightarrow \mu^+ \mu^-$ nese espacio de parámetros. Os decaimentos $B \rightarrow h^+ h^-$ ($h = K, \pi$), dos que LHCb disporá de gran cantidadade, permitirán obter as propiedades xeométricas e a masa invariante do sinal.

A principal dificultade no uso dos $B \rightarrow h^+ h^-$ ven do feito de que estos sucesos son escollidos polo *trigger* de hadróns, mentres que o sinal $B_s \rightarrow \mu^+ \mu^-$ é escollido polo *trigger* de muóns. Esto introduce diferencias notables entre as propiedades das dousas mostras. En particular, o trigger de muóns é moi eficiente e a mostra $B_s \rightarrow \mu^+ \mu^-$ non presenta sesgos respecto da *offline* (mostra ideal que se seleccionaría para a análise). Pero o trigger de hadróns precisa de cortes más duros debido á gran cantidadade de hadróns producidos no LHC, o que fai que a mostra $B \rightarrow h^+ h^-$ esté moi sesgada respecto da mostra ideal. Isto solucionase mediante a clasificación dos sucesos dacordo ao obxeto que disparou o *trigger*. Así, aqueles sucesos $B \rightarrow h^+ h^-$ no que o *trigger* foi disparado por un obxeto distinto (sucesos TIS) dos produtos da desintegración da B están prácticamente libres de sesgos, e os efectos residuais poden ser fácilmente corregidos. O estudio dos sucesos TIS tamén permite o cálculo de cocientes de eficiencias do *trigger* entre distintos canais, ou incluso a determinación da eficiencia absoluta. Os métodos para calcular ditas eficiencias e corregir sesgos producidos polo *trigger* están descritos na sección 4.4.2, e están a ser usados de forma xeral no experimento LHCb.

A identificación de muóns virá dada polas desintegracións $J/\psi \rightarrow \mu^+ \mu^-$, unha vez que o espacio fásico dos muóns é coñecido, de novo usando os $B \rightarrow h^+ h^-$.

A distribución do fondo (candidatos que non son realmente $B_s \rightarrow \mu^+ \mu^-$) será calculada a partir dos sucesos seleccionados fóra da ventá de masas (aproximadamente 3 desviacións típicas) do mesón B_s .

Este procedemento de calibración das propiedades do sinal e do fondo permite coñecer cantos dos candidatos seleccionados son desintegracións reais $B_s \rightarrow \mu^+ \mu^-$. Para convertir dito número de sucesos nunha medida da tasa de desintegración, divídese entre o número de sucesos observados noutro canal cunha tasa xa coñecida, chamado canal de normalización.

Un dos principais candidatos a canal de normalización é o decaimento $B^+ \rightarrow J/\psi (\mu^+ \mu^-) K^+$, do que LHCb disporá de gran cantidadade. Os dous muóns no estado final permiten que este daciamento sexa seleccionado polo mesmo trigger que o sinal, tendo en común ademais a eficiencia das cámaras de muóns. A principal desvantaxe é a presenza dunha traza mais no estado final, o K^+ . A probabilidade de reconstruir esta traza extra é polo tanto una cantidade que debe ser medida dalgún xeito. Para iso, fanse uso doutros canais de control. Nesta tese preséntase o exemplo de $B_d \rightarrow J/\psi (\mu^+ \mu^-) K^{*0} (K^+ \pi^-)$, donde o estudo do cociente do seu

número de sucesos respecto do número de $B^+ \rightarrow J/\psi (\mu^+\mu^-)K^+$ estaría afectado por un problema similar, neste caso catro trazas frente a tres. Como a tasa de desintegración de $B_d \rightarrow J/\psi (\mu^+\mu^-)K^{*0}(K^+\pi^-)$ é coñecida, a eficiencia de reconstruir una traza extra pode entón ser calculada.

9.3.3. Sensibilidade

A sensibilidade do experimento LHCb na búsquedas de $B_s \rightarrow \mu^+\mu^-$ estudiouse mediante a simuación MC de colisións protón protón e a resposta do detector incluindo o *trigger*. Obtívose que LHCb é o experimento con maior potencial neste análise, puidendo observar tasas incluso menores que a pedita polo *Standard Model* utilizando os datos correspondentes a cinco anos nominais. En caso de ser correcta a predicción do Standard Model, LHCb precisaría de aproximadamente un ano nominal para obter evidencia deste decaimento. Cos datos de 2010, que esperan ser tomados a menor enerxía e luminosidade, LHCb podería excluir/observar valores no rango de 10^{-8} , batendo o límite actual obtido no Tevatron e tendo un impacto importante na búsquedas de Nova Física.

9.3.4. Uso dos datos de Decembro de 2009

Durante o mes de Decembro de 2009, LHCb tomou os seus primeiros datos. As colisións foron producidas a enerxía de inxección, 450 GeV, moi por debaixo dos 7 TeV das condicións nominais. O número de colisións, aproximadamente 300 000, é tamén moito menor que o correspondente a un mes nominal. Os datos de 2009 non son, polo tanto, aptos para a búsquedas de mesóns B, mais son de grande utilidade para comprender a resposta do detector e comparara coa que forá obtida cos datos simulados. Esta tese inclúe no capítulo 6 varios estudos preliminares usando os V^0 's K_s e Λ . Ditas partículas teñen propiedades cualitativamente semellantes ás do $B_s \rightarrow \mu^+\mu^-$ xa que se detectan como vértices de dúas trazas separados do vértice da colisión, e cuxa suma vectorial de momentos apunta dende este. Comprobouse que a resolución de masa invariante e as propiedades xeométricas utilizando o mesmo tipo de trazas que se utilizarán na reconstrucción do $B_s \rightarrow \mu^+\mu^-$ son razoablemente semillantes ás obtidas na simulación. As distribucións do fondo son tamén semellantes, se ben a fracción total de fondo respecto ó sinal e aproximadamente un factor dous maior. Os K_s tamén se utilizaron para evaluar a probabilidade de identificar pions coma muóns, obténdose que esta está entre o valor obtido na simulación e un factor dous maior. Debe terse en conta que estos resultados ainda son preliminares, xa que o alinamento do detectorinda non foi completado. De todos os xeitos, indican que o coñecemento obtido da simulación parece correcto, o que é un gran avance na validación dos estudos respecto ao potencial de LHCb na análise do decaimento raro $B_s \rightarrow \mu^+\mu^-$.

10. References

1. K. G. Wilson, Phys. Rev. 179, 5 (1968).
2. E. Fermi, Ricerca Scientifica, 4(2), 491–495 (1933).
3. C. Bobeth et al., Nucl. Phys. B 630 (2002) 87.
4. A K. Alok and S. U. Sankar, Phys.Lett. B620 (2005) 61-68, arXiv:hep-ph/0502120v2 17 Mar 2005.
5. N. Cabibbo, Phys. Rev. Lett. 10 (1963) 531.
6. M. Kobayashi and T . Maskawa, Prog. Theor. Phys. 49 (1973) 652.
7. C. Amsler et al. (Particle Data Group), Physics Letters B667, 1 (2008).
8. T. Inami and C. S. Lim, Prog. Theor. Phys. 65, 297 (1981) [erratum 65, 1772 (1981)].
9. Grzadkowski and P. Krawczyk, Z. Phys. C 18, 43 (1983).
10. P. Krawczyk, Z. Phys. C 44, 509 (1989).
11. M. Misiak and J. Urban, Phys. Lett. B 451, 161 (1999) [arXiv:hep-ph/9901278].
12. G. Buchalla and A. J. Buras, Nucl. Phys. B 548, 309 (1999) [arXiv:hep-ph/9901288].
13. G. 't Hooft, Nucl. Phys. B 1 61 (1973) 455; B82 (1973) 444.
14. A. J. Buras, Phys. Lett. B 566 (2003) 115-119, hep-ph/0303060.
15. M. Blanke, A.J. Buras, D. Guadagnoli and C. Tarantino, arXiv:hep-ph/0604057 (2006).
16. CDF Collab., CDF Public Note 9892.
17. T.M. Aliev et al., Phys. Lett. B 424 (1998) 175–184.
18. Zbigniew Was, private communication.
19. M. Maltoni et. al [arXiv:hep-ph/0405172].
20. J.Ellis et. al, Phys. Rev. D 66 115013 (2002).
21. G. Bertone, D. Hooper and J. Silk, Phys. Rept. 405, 279 (2005) [arXiv:hep-ph/0404175].
22. D. Clowe et al, ApJ 648 No 2 (2006 September 10) L109-L113 .
23. M. Bradač et al., ApJ 687 No 2 (2008 November 10) 959-967.
24. Ya. B. Zel'dovich and I. D. Novikov, Astron. Zh. 43, 758 (1966) [Sov. Astron. 10, 602 (1966)].
25. P. Ivanov, P. Naselsky and I. Novikov, Phys. Rev. D 52 12 (1994).
26. R.D. Peccei and H. Quinn, Phys. Rev. Lett. 38, 1440 (1977), Phys. Rev. D16, 1791 (1977).
27. S. Weinberg, Phys. Rev. Lett. 40, 223 (1978).
28. F. Wilczek, Phys. Rev. Lett. 40, 279 (1978).
29. G.W.Bennett et al. [Muon g-2 Collaboration], Phys. Rev. Lett. 92 (2004) 161802.
30. F. Jegerlehner and A. Nyffeler, Phys. Rept. 477, 1 (2009) [hep-ph/0902.3360].
31. L. J. Hall and M.B. Wise, Nucl. Phys. B187 (1981) 397-408.
32. H. Logan and U.Nierste, preprint FERMILAB-PUB-00/84-T (hep-ph/0004139).
33. M. Misiak et al., Phys. Rev. Lett. 98, 022002 (2007).

34. G. D'Ambrosio et al., hep-ph/0207036 CERN–TH/2002–147 IFUP–TH/2002–17.
35. T. Hurth et al., Nucl. Phys. B808 (2009) 326–346, [arXiv:0807.5039].
36. A. J. Buras, arXiv:0904.4917v1 [hep-ph] 30 Apr 2009.
37. J. Ellis et. al. Nucl. Phys. B238, 453 (1984).
38. P. van Nieuwenhuizen Phys. Rep. 68, 189 (1981).
39. L. Hall, J. Lykken and S. Weinberg Phys. Rev. D 27 10 (1983).
40. G. Giudice and R. Rattazzi, Phys. Rept. 322 (1999) 419 [arXiv:hep-ph/9801271].
41. T. Gherghetta, G. Giudice and J. Wells, Nucl. Phys. B 559 (1999) 27 [arXiv:hep-ph/9904378].
42. O. Buchmueller et al, Eur.Phys.J.C64:391-415,2009 [arXiv:0907.5568v1 [hep-ph] 31 Jul 2009].
43. J. Ellis et. al. CERN-PH-TH/2007-138 [arXiv:0709.0098v1 [hep-ph]] (2007).
44. J. Ellis et. al. CERN-TH/2002-81 [arXiv:hep-ph/0204192] (2002).
45. J. Ellis et. al. CERN-PH-TH/2007-136 [arXiv:0708.2079v1 [hep-ph]] (2007).
46. B. C. Allanach et al., arXiv:0902.4880v2 [hep-ph] 20 Apr 2009.
47. R. Babier et al., Phys. Rep. 420 (2005) 1 – 195.
48. H. K. Dreiner, M. Kramer, Ben O' Leary arXiv:hep-ph/0612278v2 5 Feb 2007.
49. A. Mir, F. Tahir and K. Ahmed, Europhys. Lett. 83 41002 (2008).
50. A. K. Alok and S. K. Gupta, arXiv:0904.1878v1 [hep-ph] 12 Apr 2009.
51. Th. Kaluza, Sitzungsber. Preuss.Akad.Wiss. Phys. Math. Klasse 996 (1921).
52. O. Klein, Z.F. Physik, 37, 895 (1926).
53. G. Gabadadze, CERN –TH/2003-157.
54. N. Arkani-Hamed, S. Dimopoulos and G. R. Dvali, Phys. Lett. B 429, 263 (1998).
55. L. Randall and R. Sundrum,Phys. Rev. Lett. 83, 4690 (1999).
56. G. R. Dvali, G. Gabadadze and M. Poratti, Phys. Lett. B 485, 208 (2000).
57. M. Blanke et al., arXiv:0812.3803v3 [hep-ph] 24 Mar 2009.
58. M. E. Albrecht et al., arXiv:0903.2415v1 [hep-ph] 13 Mar 2009.
59. M. Bauer et al., arXiv:0912.1625v1 [hep-ph] 9 Dec 2009.
60. S. B. Gudnason, C. Kouvaris and F. Sannino, Phys. Rev.D 74, 095008 (2006) [arXiv:hep-ph/0608055].
61. T. Hill, FERMILAB-Conf-97/032-T, arXiv:hep-ph/9702320v1 12 Feb 1997.
62. G. Cvetič, Rev. Mod. Phys. 71,513 (1999).
63. T. Hill, Phys. Lett. B 345 (1995) 483-489.
64. K. Lane, Phys. Lett. B 433 (1998) 96-101.
65. W. Liu, C-X. Yue and H-D. Yuang, arXiv:0901.3463v3 [hep-ph] 30 May 2009.
66. N. Arkani-Hamed, A. G. Cohen and H. Georgi, Phys. Rev. Lett. 86 (2001) 4757 [arXiv:hep-th/0104005].
67. N. Arkani-Hamed, A. G. Cohen and H. Georgi, Phys. Lett.B 513 (2001) 232 [arXiv:hep-ph/0105239].
68. N. Arkani-Hamed, A. G. Cohen, E. Katz and A. E. Nelson, JHEP 0207 (2002) 034 [arXiv:hep-ph/0206021].

69. H. C. Cheng and I. Low, JHEP 0309 (2003) 051 [arXiv:hep-ph/0308199].
70. H. C. Cheng and I. Low, JHEP 0408 (2004) 061 [arXiv:hep-ph/0405243].
71. M.Blanke et al., JHEP 0701, 066 (2007).
72. Lyndon Evans and Philip Bryant (editors) 2008 JINST 3 S08001.
73. <http://linac2.home.cern.ch/linac2/sources/source.pdf>.
74. <http://linac2.home.cern.ch/linac2/seminar/seminar.pdf>.
75. The ALICE Collaboration, K Aamodt et al, 2008 JINST 3 S08002.
76. The ATLAS Collaboration, G Aad et al,2008 JINST 3 S08003.
77. The CMS Collaboration, S Chatrchyan et al, 2008 JINST 3 S08004.
78. The TOTEM Collaboration, G Anelli et al, 2008 JINST 3 S08007.
79. The LHCb Collaboration, A Augusto Alves Jr et al, 2008 JINST 3 S08005.
80. The LHCf Collaboration, O Adriani et al, 2008 JINST 3 S08006.
81. LHCb Collab., LHCb-PUB-2009-029.
82. J.Magnin, Proceedings on 12th international conference on B-physics at hadron machines.
83. R. Mureşan, Proceedings on 12th international conference on B-physics at hadron machines.
84. Y.Gao, Proceedings on 12th international conference on B-physics at hadron machines.
85. M. Calvi, O. Leroy and M. Musy, CERN-LHCb-2007-058.
86. J. Baines et al., CERN-2000-004.
87. M. Krasowski et al., LHCb-2007-011.
88. M. D. Needham, LHCb 2008-037.
89. S. Amato et al., CERN-LHCb-2008-058.
90. A. M. Perez-Calero Yzquierdo, H. Ruiz, CERN-LHCb-2008-075.
91. J.A. Hernando et al., CERN-LHCb-2009-034.
92. T. Sjostrand et al. , Comp. Phys. Comm. 135 (2001) 238.
93. <http://www.thep.lu.se/~torbjorn/Pythia.html>.
94. GEANT detector description and simulation tool, CERN Program Library long writeup W5013 (1994).
95. <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/gauss/>.
96. <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole/>.
97. <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/>.
98. D.J. Lange, NIM A 462 (2001) 152.
99. <http://www.slac.stanford.edu/~lange/EvtGen/>.
100. T. Junk, CARLETON/OPAL PHYS 99-01 (1999).
101. A. L. Read, CERN Yellow Report 2000-005.
102. D.Karlen, Computers in Physics Vol 12, N.4, Jul/Aug 1998 .
103. D. Martínez Santos, J. A. Hernando, F. Teubert. LHCb-PUB-2007-033.
104. ATLAS collab., CERN-OPEN-2008-020 [arXiv:0901.0512] (B-physics chapter).
105. CMS collab., CMS PAS BPH-07-001 (2009).

106. D. Martínez Santos, Proceedings on 12th international conference on B-physics at hadron machines.
107. Torbjörn Sjöstrand et al., JHEP 05 026 (2006).
108. B. Souza de Paula, LHCb-2006-046.
109. E. Polycarpo and M. Gandelman, CERN-LHCb-2007-145.
110. G. Lanfranchi et al., LHCb note 2009-013 (2009).
111. A. Sarti et al., LHCb-PUB-2010-002.
112. T. Junk, CDF/DOC/STATISTICS/PUBLIC/8128 (2007).
113. G. Buchalla et al., Eur. Phys. J. C (2008) 57: 309–492.
114. Y. Amhis. Discussion in LHCb V0 meeting. January 26th, 2010.
115. J. Podolanski and R. Armenteros, Phil. Mag. 45(1954)13.
116. J. Van Tilburg, private communication.
117. M. Britsch, DIS 2009 conference, Madrid.