

# LHCb实验上双粲味重子的寻找

(申请清华大学理学博士学位论文)

培 养 单 位 : 工程物理系

学 科 : 物理学

研 究 生 : 钟 亮

指 导 教 师 : 高 原 宁 教 授



二〇一五年一月



# **Search for the Doubly Charmed Baryon at LHCb**

Dissertation Submitted to  
**Tsinghua University**  
in partial fulfillment of the requirement  
for the degree of  
**Doctor of Philosophy**  
in  
**Physics**

by

**ZHONG Liang**

Dissertation Supervisor : Professor Gao Yuanning

**January, 2015**





## 关于学位论文使用授权的说明

本人完全了解清华大学有关保留、使用学位论文的规定，即：

清华大学拥有在著作权法规定范围内学位论文的使用权，其中包括：（1）已获学位的研究生必须按学校规定提交学位论文，学校可以采用影印、缩印或其他复制手段保存研究生上交的学位论文；

（2）为教学和科研目的，学校可以将公开的学位论文作为资料在图书馆、资料室等场所供校内师生阅读，或在校园网上供校内师生浏览部分内容；（3）根据《中华人民共和国学位条例暂行实施办法》，向国家图书馆报送可以公开的学位论文。

本人保证遵守上述规定。

**（保密的论文在解密后应遵守此规定）**

作者签名：\_\_\_\_\_

导师签名：\_\_\_\_\_

日 期：\_\_\_\_\_

日 期：\_\_\_\_\_



## 摘要

双粲味重子 $\Xi_{cc}^+$ 是SU(4)夸克模型预言的包含两个粲夸克的重子, 目前实验上尚未有确切证据表明其存在。许多基于量子色动力学的理论模型都给出了对这个粒子性质的预言, 其质量在3500-3700 MeV/ $c^2$ 之间, 寿命在110-250 fs之间。实验上寻找 $\Xi_{cc}^+$ 并测量其性质可以直接检验这些模型, 对理解量子色动力学的非微扰性质有着重要意义。

2003年SELEX实验组声称在 $\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$ 衰变道中观察到了双粲味重子信号, 但是所测量的信号寿命远小于理论预言。FOCUS, Belle和BaBar等实验组在同样的衰变道中寻找 $\Xi_{cc}^+$ , 都没有观测到信号。但是这些实验并没有直接排除SELEX的结果, 因为它们的产生环境都与SELEX不相同。大型强子对撞机(LHC)是一个高能质子-质子对撞机, 理论预期LHC上 $\Xi_{cc}^+$ 的产生截面可达0.1  $\mu\text{b}$ 或更高。LHCb探测器是LHC上的一个专为重味物理设计的大型探测器, 具有良好的寻迹和粒子鉴别系统。利用2011年LHCb积累的0.65  $\text{fb}^{-1}$ 数据, 本论文用 $\Xi_{cc}^+ \rightarrow (\Lambda_c^+ \rightarrow p K^- \pi^+) K^- \pi^+$ 这一衰变道对 $\Xi_{cc}^+$ 进行了寻找, 没有观测到信号。由于 $\Xi_{cc}^+$ 的质量和寿命均未知, 本论文给出了在5个不同的寿命假设下,  $\Xi_{cc}^+$ 相对 $\Lambda_c^+$ 的截面上限随质量假设的变化曲线。

论文中给出的上限与理论预言没有直接矛盾, 但远低于SELEX的结果, 可能的原因是LHCb和SELEX的产生环境不同或 $\Xi_{cc}^+$ 的寿命确实很小。今后LHCb将用更多的数据和更多的衰变道来寻找 $\Xi_{cc}^+$ , 对信号的触发也有很大改进, 可以预期LHCb对 $\Xi_{cc}^+$ 信号灵敏度会有很大提高。

**关键词:** 大型强子对撞机; 双粲味重子; 量子色动力学; 粲强子谱

## Abstract

The doubly charmed baryon  $\Xi_{cc}^+$ , containing two charm quarks, is a baryon predicted by the SU(4) quark model. Experimentally its existence has not been established yet. Many Quantum Chromodynamics (QCD) based theoretical models have predicted its properties with a mass in the range 3500-3700 MeV/ $c^2$  and a lifetime in the range 110-250 fs. The experimental searches for the  $\Xi_{cc}^+$  baryon and the measurements of its properties can test these models directly, providing an important input for the understanding of the non-perturbative aspect of QCD.

The SELEX collaboration claimed the observation of the  $\Xi_{cc}^+$  baryon in the  $\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$  decay in 2003. However, the measured lifetime was much shorter than theoretical predictions. Searches for the  $\Xi_{cc}^+$  baryon in the same decay mode by FOCUS, Belle and BaBar experiments failed to reproduce the results. This does not mean that the SELEX result is excluded, however, since production environments at these experiments were different from that of SELEX. The Large Hadron Collider (LHC) at CERN is a high energy proton-proton collider with an expected  $\Xi_{cc}^+$  production cross-section of the order of 0.1  $\mu\text{b}$ . LHCb is the dedicated heavy flavor physics experiment at the LHC and has good tracking, flexible trigger and particle identification systems. Using the 0.65  $\text{fb}^{-1}$  data collected by LHCb in 2011, a search for the  $\Xi_{cc}^+$  baryon in the decay  $\Xi_{cc}^+ \rightarrow (\Lambda_c^+ \rightarrow p K^- \pi^+) K^- \pi^+$  is reported in this thesis. No signal is observed. Since the mass and lifetime of the  $\Xi_{cc}^+$  baryon are a priori unknown, the upper limit of production cross-section of the  $\Xi_{cc}^+$  baryon relative to that of the  $\Lambda_c^+$  baryon is given as a function of the invariant mass for 5 different lifetime hypotheses.

The measured upper limits have no direct contradiction with theoretical predictions, but much smaller than the value reported by SELEX. This could be explained by the fact that the production environments of LHCb and SELEX are different, or that the lifetime of  $\Xi_{cc}^+$  is indeed smaller than theoretical predictions. In future the  $\Xi_{cc}^+$  baryon will be searched with 2011 and 2012 data using more decay modes, and the trigger for the signal will be much improved. It is expected that the sensitivity of the  $\Xi_{cc}^+$  baryon search will be much enhanced.

**Key words:** LHC; Doubly Charmed Baryon; QCD; Charmed Baryon Spectroscopy

## Contents

<b>Chapter 1 Introduction</b> .....	1
1.1 Theoretical background .....	1
1.1.1 The Standard Model .....	1
1.1.2 Quantum chromodynamics .....	3
1.1.3 The quark model .....	6
1.2 Doubly charmed baryon .....	6
1.2.1 Theoretical predictions .....	6
1.2.2 Experimental status .....	15
<b>Chapter 2 The LHCb experiment</b> .....	26
2.1 The Large Hadron Collider .....	26
2.2 The LHCb experiment .....	29
2.2.1 Tracking system .....	30
2.2.2 PID system .....	41
2.3 LHCb trigger .....	52
2.3.1 Level-0 trigger .....	53
2.3.2 High level trigger .....	54
2.3.3 TISTOS method .....	55
2.3.4 Performance of the trigger system .....	56
2.4 LHCb data processing .....	56
2.5 LHCb software framework .....	58
2.5.1 Simulation .....	58
2.5.2 Trigger .....	59
2.5.3 Track reconstruction .....	59
2.5.4 Data analysis .....	59
<b>Chapter 3 Search for the <math>\Xi_{cc}^+</math> baryon using 2011 data</b> .....	60
3.1 Analysis strategy .....	60
3.2 Data set .....	61
3.3 Selection .....	62
3.3.1 $\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$ decay topology .....	62
3.3.2 Stripping and offline selection .....	63
3.3.3 Multivariate analysis .....	67
3.3.4 Trigger requirements .....	71

---

## Contents

---

3.4	Efficiency ratio determination .....	73
3.4.1	Ratio of acceptance efficiencies.....	75
3.4.2	Ratio of stripping and offline selection efficiencies .....	75
3.4.3	Ratio of PID cut efficiencies .....	79
3.4.4	Ratio of the MLP selection efficiency .....	86
3.4.5	Ratio of trigger efficiency .....	86
3.4.6	Summary of efficiency ratios.....	89
3.5	Yield determination .....	89
3.5.1	Fit of the invariant mass distribution for $\Lambda_c^+$ candidates from the control sample in data.....	89
3.5.2	Fit of the invariant mass distribution for $\Xi_{cc}^+$ candidates from the control sample in MC .....	90
3.5.3	Procedure to measure the $\Xi_{cc}^+$ yield on data .....	91
3.6	Decays through an intermediate $\Sigma_c$ baryon .....	96
3.7	Systematic uncertainties .....	98
3.7.1	Tracking efficiency .....	98
3.7.2	Stripping, cut-based offline selection and MLP efficiencies .....	99
3.7.3	PID calibration.....	99
3.7.4	Trigger efficiency .....	104
3.7.5	Yield determination.....	104
3.7.6	Decay models.....	104
3.7.7	Summary .....	106
3.8	Variation of the efficiency ratio .....	106
3.8.1	Variation of the efficiency ratio depending on the $\Xi_{cc}^+$ mass .....	107
3.8.2	Variation of the efficiency ratio depending on the $\Xi_{cc}^+$ lifetime .....	110
3.9	Upper limit calculation.....	115
3.10	Results .....	118
<b>Chapter 4</b>	<b>Search for the <math>\Xi_{cc}^+</math> baryon using 2011 and 2012 data .....</b>	<b>124</b>
4.1	Analysis strategy .....	124
4.2	Data set .....	124
4.3	Selection .....	125
4.3.1	Stripping and offline selection .....	125
4.3.2	Multivariate analysis .....	128
<b>Chapter 5</b>	<b>Summary and outlook .....</b>	<b>132</b>
5.1	Summary.....	132
5.2	Outlook .....	133

## Contents

---

<b>Reference</b> .....	137
<b>Acknowledgement</b> .....	144
声 明 .....	145
<b>Appendix A Tables for PID calibration</b> .....	146
A.1 Calibration of the $\Xi_{cc}^+$ MC using the multiplicity in the $\Lambda_c^+$ sample of data...	146
A.2 Calibration of the $\Xi_{cc}^+$ MC without the multiplicity weighting .....	150
<b>Appendix B Calculation of the uncertainty of L0 trigger efficiency</b> .....	154
<b>Appendix C Uncertainty of the efficiency ratio relative to that of 3500 MeV/c<sup>2</sup></b> ..	156
<b>Appendix D Background Subtraction using 25-Tiles method</b> .....	157
D.1 Introduction.....	157
D.2 Lemma: combinatorial background .....	157
D.2.1 Signal box.....	158
D.2.2 Corner box .....	159
D.2.3 Vertical sidebands .....	159
D.2.4 Horizontal sidebands .....	159
D.3 Lemma: $\Lambda_c^+$ background .....	160
D.4 Combining the above results.....	160
Resume and publications .....	163

## List of main acronyms

SM	Standard Model
QCD	Quantum ChromoDynamics
LO	Leading-Order
NLO	Next-to-Leading Order
PDF	Probability Density Function
OPE	Operator Product Expansion
pQCD	Perturbative QCD
NRQCD	Non-relativistic QCD
DCB	Doubly Charmed Baryon
PID	Particle IDentification
PV	Primary Vertex
IP	Impact Parameter
HLT	High Level Trigger
TCK	Trigger Configuration Key
DST	Data Summary Tape
GEC	Global Event Cut
MC	Monte Carlo
$E$	Energy
$\eta$	Pseudo-rapidity
$p$	Momentum
$p_T$	Transverse momentum
BF	Branching fraction
TeV	Tera electron volt
GeV	Giga electron volt
MeV	Mega electron volt
LHC	Large Hadron Collider
LHCb	Large Hadron Collider beauty experiment
$pp$	proton-proton
SELEX	Segmented Large $x$ baryon spectrometer
LEE	Look-Elsewhere Effect
CL	Confidence Level



## Chapter 1 Introduction

The thesis describes the work of searching for doubly charmed baryon  $\Xi_{cc}^+$  using proton-proton ( $pp$ ) collision data collected by the LHCb experiment <sup>①</sup>. The thesis is organized as below: Chapter 1 introduces the theoretical framework of DCBs, followed by an overview of the LHC collider and the LHCb spectrometer where the search is performed in Chapter 2. Chapter 3 presents the selection criteria, the associated systematic uncertainties and the results of the search for the  $\Xi_{cc}^+$  baryon using data collected in 2011. Chapter 4 shows the improvements in the selection of the  $\Xi_{cc}^+$  search using data collected in 2011 and 2012. Chapter 5 gives a summary of the results and future prospects.

### 1.1 Theoretical background

#### 1.1.1 The Standard Model

Particle physics is the branch of physics which studies the basic constituents of matter and interactions between them. The best understanding of laws governing the fundamental particles and interactions are called the Standard Model (SM). It includes 12 fermions and 4 gauge bosons as the fundamental particles and describes interactions between them. It is a remarkably successful theory in that it is able to explain almost all experimental observations and makes precise predictions for a wide variety of phenomena. Figure 1.1 shows the elementary particles in the SM.

In the SM the fermions are grouped into two families, quarks and leptons, depending on whether they participate in strong interaction. We distinguish six quark flavors corresponding to three generations of matter, *i.e.* the up quark ( $u$ ) and the down quark ( $d$ ) form the first generation, the charm quark ( $c$ ) and the strange quark ( $s$ ) form the second generation, and the top quark ( $t$ ) and the bottom quark ( $b$ ) form the third generation. They carry fractional electric charge and an additional quantum number called color charge, introduced to describe strong interaction. A single free quark cannot be observed in nature due to a phenomenon called color confinement; all quarks are compelled to form

---

① The charge conjugation will be implied implicitly unless specified otherwise.

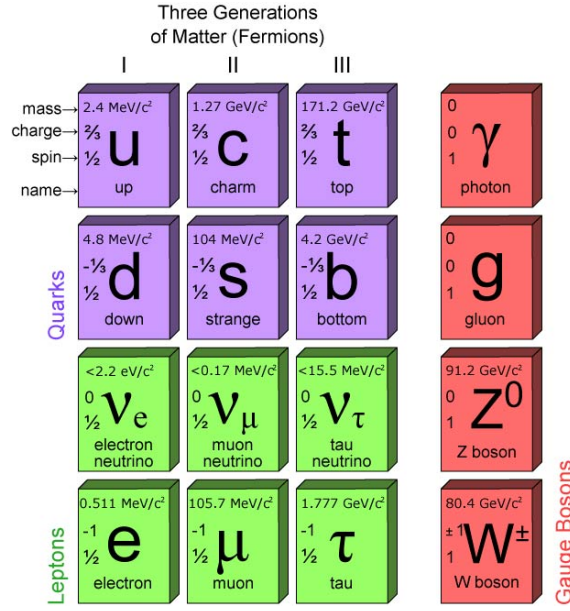


Figure 1.1 Elementary particles in the SM. The left three columns shows fermions, which are divided into three generations, and the right column represents gauge bosons.

Table 1.1 Properties of quarks and leptons

Quarks (spin = $\frac{1}{2}$ )			Leptons (spin = $\frac{1}{2}$ )		
Flavor	Electric Charge	Mass (GeV/ $c^2$ )	Flavor	Electric Charge	Mass (MeV/ $c^2$ )
$u$	$+\frac{2}{3}$	0.002	$e$	-1	0.511
$b$	$-\frac{1}{3}$	0.005	$\nu_e$	0	< 0.002
$c$	$+\frac{2}{3}$	1.28	$\mu$	-1	105.6
$s$	$-\frac{1}{3}$	0.095	$\nu_\mu$	0	< 0.17
$t$	$+\frac{2}{3}$	173.2	$\tau$	-1	1777.8
$b$	$-\frac{1}{3}$	4.2	$\nu_\tau$	0	< 18.2

colorless structures called hadrons, which comprise a quark and an antiquark (mesons), or three quarks or three antiquarks (baryons). The leptons also have six flavors and are arranged into three generations, *i.e.* electron and electron neutrino, muon and muon neutrino, and tau lepton and tau neutrino. The electron, the muon and the tau lepton are massive particles and carry a negative elementary electric charge, while the neutrinos have small but non-zero masses and carry zero electric charge. For each fermion, there is a corresponding antiparticle which has the same mass and opposite quantum numbers. A summary of fermion properties is given in Table 1.1.

Table 1.2 Fundamental forces in the SM

Force	Theory	Strength	Range (m)	Mediators
Weak	Electroweak	$10^{-6}$	$10^{-18}$	$W^+$ , $W^-$ and $Z$
EM	QED	1/137	$\infty$	photon
Strong	QCD	1	$10^{-15}$	gluon

All known phenomena from the small scale of quarks and leptons to the large scale of galaxies can be explained in terms of four fundamental interactions: gravity, weak, electromagnetic (EM) and strong interactions. The gravity is described by Einstein's general relativity, and the latter three are described by the SM, in which they occur through the exchange of gauge bosons.

The weak interaction acts upon all the fermions and is responsible for some of the spontaneous decays of unstable particles, *e.g.* radioactive  $\beta$  decay of neutrons. It has three massive mediating gauge bosons,  $W^+$ ,  $W^-$ , and  $Z$ . The two charged bosons,  $W^+$  and  $W^-$ , both have a mass of  $80.385 \text{ GeV}/c^2$ , and the neutral boson,  $Z$ , has a mass of  $91.187 \text{ GeV}/c^2$ . The large mass of these gauge bosons makes the weak interaction a short-range force and a feeble interaction at low energy scale.

The EM interaction acts at particles possessing electric charge. It is mediated by a massless and chargeless photon and has an infinite interaction range. Though having very different behavior at low energy scale, the weak and the EM interactions are described uniformly in the electroweak theory.

The strong interaction acts between particles with color charge. Its interaction carriers are eight mediating vector particles called gluons, which are also massless and chargeless vector bosons. Unlike electrically neutral photons, gluons carry non-zero color charge and therefore couple to other gluons. The theory of the strong force is called quantum chromodynamics (QCD), which is described later in Section 1.1.2. Properties of the three fundamental forces are summarized in Table 1.2.

### 1.1.2 Quantum chromodynamics

QCD is a non-Abelian gauge field theory based on the gauge group  $SU(3)$  and describes the strong interaction sector of the SM<sup>[1]</sup>. The Lagrangian of QCD is given

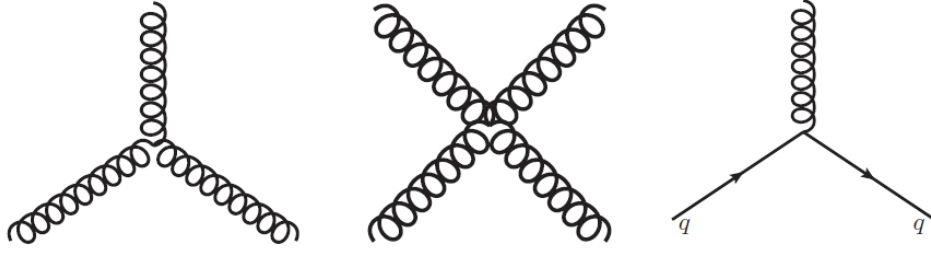


Figure 1.2 Feynman rules of QCD. Curly lines represent gluons, and solid lines indicate quarks. The first and the second diagrams show the triple and quartic gluon self-interactions, respectively, and the third diagram shows the gluon bremsstrahlung of the quark.

by

$$\mathcal{L} = \bar{\psi}_q^i (i\gamma^\mu) (D_\mu)_{ij} \psi_q^j - m_q \bar{\psi}_q^i \psi_{qi} - \frac{1}{4} G_{\mu\nu}^a G^{a\mu\nu}, \quad (1-1)$$

where  $\psi_q^i$  is the Dirac field of the quark with flavor  $q$  and color index  $i$ ,  $\psi_q = (\psi_{qR}, \psi_{qG}, \psi_{qB})$ ,  $\gamma^\mu$  is the Dirac matrix,  $m_q$  is the mass of the quark,  $G_{\mu\nu}^a$  is the gluon field strength tensor for a gluon with color index  $a$ , and  $D_\mu$  is the covariant derivative in QCD,

$$(D_\mu)_{ij} = \delta_{ij} \partial_\mu - ig_s \frac{\lambda_{ij}^a}{2} A_\mu^a, \quad (1-2)$$

with  $g_s$  the strong coupling constant,  $A_\mu^a$  the gluon field with color index  $a$ , and  $\lambda_{ij}^a$  the Gell-Mann matrices of  $SU(3)^{[1]}$ ,

$$\begin{aligned} \lambda^1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \\ \lambda^5 &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \lambda^8 = \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 \\ 0 & 0 & \frac{-2}{\sqrt{3}} \end{pmatrix}. \end{aligned} \quad (1-3)$$

Feynman rules of QCD can be derived from the Lagrangian, as shown in Figure 1.2. Due to the non-Abelian nature of QCD, gluons interact with other gluons through triple and quartic vertices.

The energy scale dependence of the strong coupling constant is determined by a

differential equation

$$\frac{\partial \alpha_s}{\partial \ln \mu} = \beta(\alpha_s). \quad (1-4)$$

Here,  $\mu$  is the energy scale and the  $\beta$  function is

$$\beta(\alpha_s) = -\alpha_s^2(b_0 + b_1\alpha_s + \dots) \quad (1-5)$$

with leading order (LO) and next-to-leading-order (NLO) coefficients

$$\begin{aligned} b_0 &= \frac{33 - 2n_f}{6\pi} \\ b_1 &= \frac{153 - 19n_f}{24\pi^2}, \end{aligned} \quad (1-6)$$

where  $n_f$  is the number of active quark flavors. Since in the SM the maximum possible value of  $n_f$  is 6,  $b_0$  is always a positive number.

The dependence is more clear if we ignore high orders and solve these equations explicitly,

$$\alpha_s(\mu) = \frac{1}{b_0 \ln(\mu/\Lambda_{\text{QCD}})}, \quad (1-7)$$

where  $\Lambda_{\text{QCD}} \sim 250 \text{ MeV}$  is the strong interaction scale, where the coupling constant becomes large and the perturbative formalism starts to break down<sup>[1]</sup>. If the energy scale  $\mu$  goes to infinity (or the distance approaches zero), the strong coupling decreases to zero. This remarkable feature is called asymptotic freedom, which implies that the perturbative formalism is legitimate at energy scale much higher than  $\Lambda_{\text{QCD}}$ . Here the conclusion is derived by removing high order terms, but consideration of the complete expression of Eq. 1-5 have shown that asymptotic freedom occurs for the full QCD<sup>[2,3]</sup>. Towards lower energies, the perturbative coupling becomes divergent at the energy scale close to  $\Lambda_{\text{QCD}}$ , which indicates that physics at this energy scale is non-perturbative. Universal functions, effective theories or phenomenological models must be used to extract the results. The running of the coupling constant is demonstrated in Figure 1.3.

Another prominent feature of QCD is color confinement, which states that at zero temperature only color singlet states can exist at distances larger than  $1/\Lambda_{\text{QCD}}$ <sup>[5]</sup>. In other words, isolated free quarks or gluons cannot exist in nature. The color confinement is still a theoretical conjecture that is consistent with experimental observations, however, it has not been proved yet.

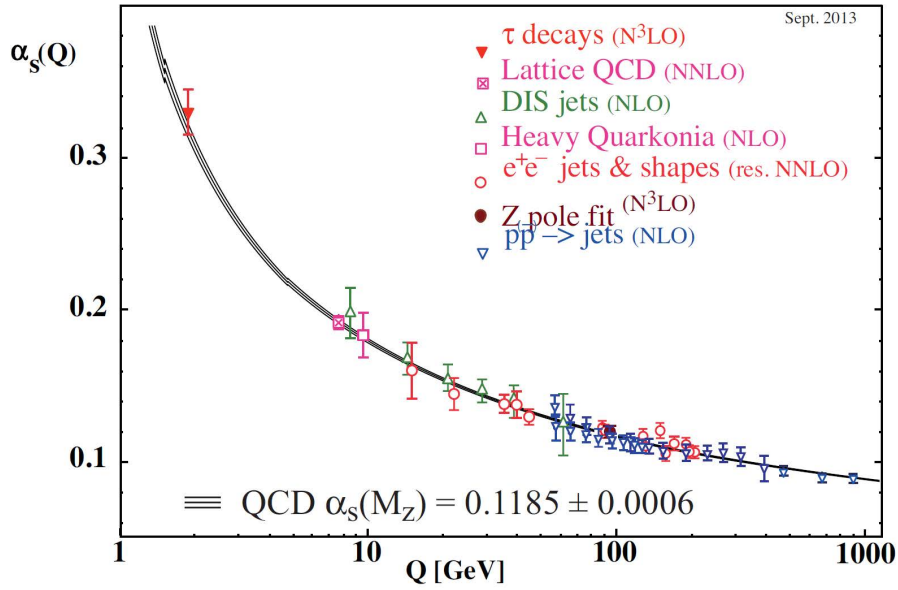


Figure 1.3 Illustration of the running of the coupling constant<sup>[4]</sup>.

### 1.1.3 The quark model

The quark model is an effective tool to classify and predict hadron structures<sup>[6,7]</sup>. Considering the first four quarks, *i.e.*  $u$ ,  $d$ ,  $s$  and  $c$  quarks, the quark model predicts two  $SU(4)$  multiplets of baryons which consist of three of these four quarks, as shown in Figure 1.4. Ground states of the baryons with zero or one charm quark are all discovered, while the baryons with two or three charm quarks are still not reliably observed. These particles are difficult to produce due to the presence of two or three heavy quarks, but their existence is highly probable because of the great success of the quark model.

The lightest doubly charmed baryons (DCBs) are  $\Xi_{cc}^+ = ccd$  and  $\Xi_{cc}^{++} = ccu$ , which form an isospin doublet with  $J_P = \frac{1}{2}^+$ . This thesis focuses on the search of the  $\Xi_{cc}^+$  baryon.

## 1.2 Doubly charmed baryon

### 1.2.1 Theoretical predictions

The DCBs are particularly interesting from a theoretical perspective since they provide valuable information about the non-perturbative QCD and serve as a ground to test various models<sup>[8]</sup>. Though consisting of three quarks, theorists consider that the structure of a DCB is very similar to that of a Hydrogen atom. Since the mass of the charm quark is substantially larger than the QCD scale  $\Lambda_{\text{QCD}}$ , the two charm quarks are

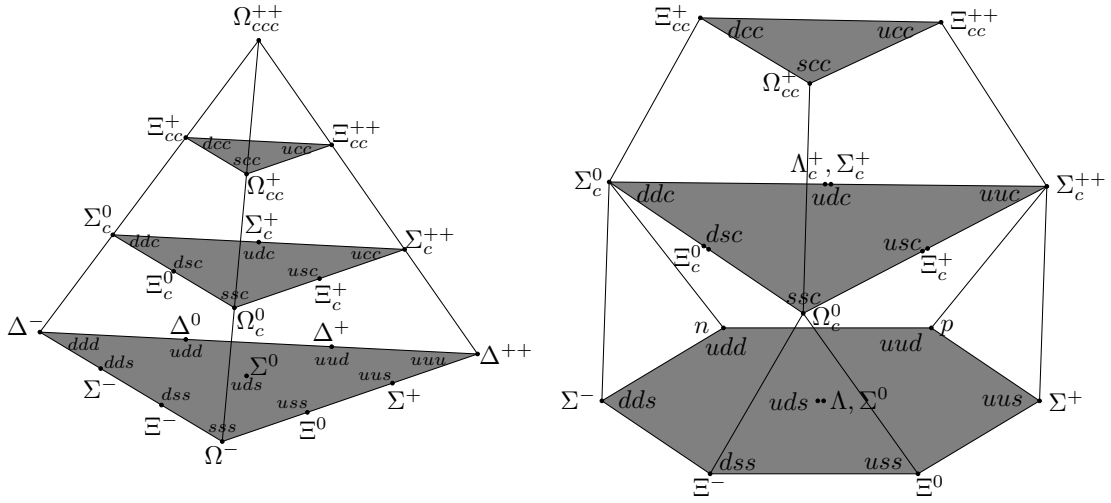


Figure 1.4 SU(4) multiplets of baryons composed of  $u$ ,  $d$ ,  $s$  and  $c$  quarks. The left (right) plot shows the 20-plet with an SU(3) octet (decuplet). Baryons in the same multiplet have the same spin and parity.

bound into a compact diquark system with a dimension much smaller than the scale of the light quark confinement<sup>[8,9]</sup>. The diquark acts as a static color source and is surrounded by the light quark, just like the proton is surrounded by the electron in the Hydrogen atom. With this assumption, the properties of DCBs can be modeled easier. However, although this approximation is generally considered valid by theorists, it has never been verified by experiments.

### 1.2.1.1 Mass

The mass of the  $\Xi_{cc}^+$  baryon is considered by various methods, *e.g.* relativistic quark model<sup>[10,11]</sup>, non-relativistic quark model<sup>[12–15]</sup>, QCD sum rules<sup>[16–18]</sup> and the MIT bag model<sup>[8]</sup>. This section does not attempt to provide a comprehensive review of the literature, but only gives a brief introduction to some of these methods and summarizes the results.

Potential models provide an effective approach to predict mass spectra of hadrons. Extracting spectrum information directly from the QCD Lagrangian is a challenging task since non-perturbative contributions dominate in the infrared region. However, one may recall that Schrödinger equation precisely predicts the spectrum of the Hydrogen atom. If a potential between two quarks is known, reasonable predictions for hadron spectra can be made by solving the corresponding Schrödinger equation. This method provides

a clear physics picture and requires relatively easy calculations, although its predictions have discrepancies with experimental results since non-perturbative processes and high order corrections are simply replaced by the potential. The accuracy of the potential model method depends on the potential used. For predictions of DCBs, there are various choices for the potential, *e.g.* the relativistic quark model<sup>[10,11]</sup> and the non-relativistic quark model<sup>[12–15]</sup>.

QCD sum rules provide model-independent predictions of hadronic parameters<sup>[19]</sup>. Though analytical calculation of the bound state system is quite difficult due to the failure of perturbative formalism in the infrared regime, QCD sum rules extract information on the bound states by connecting the bound state problem with short-distance calculable amplitudes. The idea is to start from the ultraviolet side, where perturbative methods apply, and move towards the infrared region. As the distance increases, confinement effects become significant, perturbative methods start to fail, and resonances emerge as a reflection of the fact that gluons and quarks are confined within hadrons<sup>[20]</sup>. On one hand, the correlation function can be expanded by inserting a full set of intermediate hadronic states, and the spectral function can be related to the expansion by the dispersion relation. On the other hand, the correlation function can also be decomposed by the operator product expansion (OPE)<sup>[21]</sup>, where the coefficients contain short distance contributions, which can be expressed in terms of the Lagrangian parameters of the theory in a perturbative way. Long distance effects present in higher dimensional operators of the expansion as vacuum expectation values of gluons and quarks<sup>[18]</sup>, which are determined by experiment. This way information on the hadronic states can be extracted.

Theory predictions are listed in Table 1.3. Most calculations yield mass predictions in the range 3500 – 3700 MeV/c<sup>2</sup><sup>Ⓣ</sup>. Uncertainty of the predictions depends on the methods, and it ranges from 10 MeV/c<sup>2</sup> to 200 MeV/c<sup>2</sup>.

### 1.2.1.2 Lifetime

The lifetime of the  $\Xi_{cc}^+$  baryon can be calculated on the basis of optical theorem, which relates the total width of the heavy hadron with the imaginary part of the forward scattering amplitude, and OPE, which allows to separate the short-distance contributions and the long-distance ones<sup>[28–30]</sup>.

Ⓣ There are calculations that predict a  $\Xi_{cc}^+$  mass larger than 4000 MeV/c<sup>2</sup>, but they are not considered in this thesis.



Table 1.3 Theoretical predictions for the  $\Xi_{cc}^+$  mass

Method	$\Xi_{cc}^+$ mass[ MeV/ $c^2$ ]	Reference
MIT bag model	3520	D. He <i>et al.</i> [8]
QCD sum rule	3570	Z. Wang [22]
QCD sum rule	3560	V. Kiselev <i>et al.</i> [23]
Potential model	3510	A. Martynenko [11]
Potential model	3732	S. Tong <i>et al.</i> [24]
Potential model	3676	D. Lichtenberg <i>et al.</i> [25]
Lattice QCD	3608	R. Lewis <i>et al.</i> [26]
Lattice QCD	3549	J. Flynn <i>et al.</i> [27]

The total decay width of a charmed hadron  $H_c$  is then written as

$$\begin{aligned}\Gamma(H_c \rightarrow X) &= \frac{1}{m_{H_c}} \text{Im} \int d^4x \langle H_c | \hat{T} | H_c \rangle \\ &= \frac{1}{2m_{H_c}} \langle H_c | \hat{\Gamma} | H_c \rangle,\end{aligned}\tag{1-8}$$

Here,  $\hat{T} = T\{i\mathcal{L}_{\text{eff}}(x), \mathcal{L}_{\text{eff}}(0)\}$  is the non-local operator in which  $T$  denotes the time ordered product and  $\mathcal{L}_{\text{eff}}$  represents the effective Lagrangian, and  $m_{H_c}$  is the mass of hadron.

If the decay releases a sufficiently large energy, the non-local operator can be expressed as an infinite sum of local operators  $O_i$  with  $1/m_c$  as the expansion parameter, where  $m_c$  the short distance mass of the charm quark, *i.e.* the mass used in the modified minimal subtraction scheme [31]. The decay width of a singly charmed hadron takes the following form

$$\begin{aligned}\Gamma(H_c \rightarrow f) &= \frac{G_F^2 m_c^5}{192\pi^3} |V|^2 \frac{1}{M_{H_c}} \left\{ c_3^f \langle H_c | \bar{c}c | H_c \rangle + c_5^f \frac{\langle H_c | \bar{c} g_s \sigma^{\mu\nu} G_{\mu\nu} c | H_c \rangle}{m_c^2} \right. \\ &\quad \left. + \sum_i c_6^f \frac{\langle H_c | (\bar{c} \Gamma_i q)(\bar{q} \Gamma_i c) | H_c \rangle}{m_c^3} + (1/m_c^4) \right\}.\end{aligned}\tag{1-9}$$

where  $G_F$  is the Fermi constant,  $V$  is the corresponding element in the CKM matrix and  $c_i^f$  is the Wilson coefficient functions.

The first term on the right-hand side of Eq. 1-9 is a dimension-three operator, which represents the kinematic energy of the heavy quark and some non-perturbative contributions [32]. This contribution is doubled for DCBs since they contain two charm

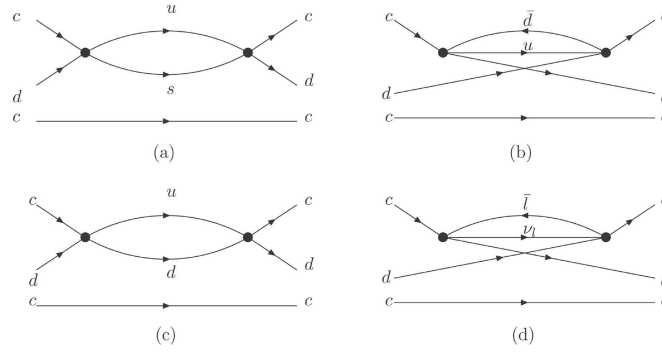


Figure 1.5 Non-spectator effects contributions to the lifetime of the  $\Xi_{cc}^+$  baryon<sup>[30]</sup>. Figures (a) and (c) are the weak scattering diagrams, and (b) and (d) are the Pauli interference diagrams.

quarks<sup>[28]</sup>. The second term is the chromomagnetic operator, which describes the spin interaction<sup>[32]</sup>. This contribution is also doubled for DCBs. The last term represents the non-spectator effects, which are the interactions of the light quark with the heavy quark and consist of two processes, namely the Pauli interference and the weak scattering, as are illustrated in Figure 1.5. The non-spectator effects are dominated by non-perturbative processes, and therefore cannot be reliably calculated from QCD yet. Phenomenological models are used to carry out the calculation. The weak scattering contribution could be as large as 60%-120% of the total width depending on the model used<sup>[30,33]</sup>, while the Pauli interference contribution is suppressed.

Using the above considerations and appropriate input parameters, the lifetime of the  $\Xi_{cc}^+$  baryon can be obtained. Results from different articles are listed in Table 1.5. The predictions for the  $\Xi_{cc}^+$  lifetime mainly range between 110 and 250 fs, which means the  $\Xi_{cc}^+$  baryon will fly a small but non-zero distance before it decays. The uncertainty of the predictions is about 20%-30%. The smallest prediction is  $0.11 \pm 0.03$  ps, and the largest prediction is  $0.25 \pm 0.05$  ps. The difference between the smallest and largest value is about 2.4 standard deviation.

### 1.2.1.3 Production cross-section

In general, the formation of DCBs is divided into three steps according to the time scale of the process:

1. Production of two  $c$  quarks in the  $pp$  collision.
2. Binding of the two  $c$  quarks into a diquark structure with a configuration of

Table 1.4 Theoretical predictions for the  $\Xi_{cc}^+$  lifetime.

$\Xi_{cc}^+$ lifetime[fs]	Reference
110	V. Kiselev <i>et al.</i> <sup>[9]</sup>
120	V. Kiselev <i>et al.</i> <sup>[23]</sup>
220	B. Guberina <i>et al.</i> <sup>[28]</sup>
200	A. Likhoded <i>et al.</i> <sup>[29]</sup>
250	C. Chang <i>et al.</i> <sup>[30]</sup>

 Table 1.5 Theoretical predictions for the  $\Xi_{cc}^+$  lifetime.

$\Xi_{cc}^+$ lifetime[fs]	Reference
110	PRD 60 (1999) 014007
120	Phys.Usp. 45 (2002) 455
220	EPJC 9 (1999) 213
200	arXiv:hep-ph/9912425
250	Commun.Theor.Phys 49 (2008) 993

$(cc)_3[{}^3S_1]$  or  $(cc)_6[{}^1S_0]$ .<sup>①</sup>

### 3. Hadronization of the diquark structure into a DCB.

The first step mainly proceeds via the following four processes<sup>[35]</sup>,

$$g + g \rightarrow cc\bar{c}\bar{c},$$

$$g + c \rightarrow cc\bar{c},$$

$$c + c \rightarrow ccg,$$

$$q + \bar{q} \rightarrow cc\bar{c}\bar{c}.$$

At high energy colliders, *e.g.* the Large Hadron Collider (LHC), the contribution from the quark-antiquark annihilation is much smaller than those from gluon-gluon fusion and gluon-charm scattering<sup>[35,36]</sup>. Therefore this process is ignored in the following calculations.

① In the ground state of a DCB, the  $cc$ -diquark can only be produced in the configuration  $(cc)_3[{}^3S_1]$  since the whole baryon should be colorless. However, the configuration  $(cc)_6[{}^1S_0]$  is also possible if an additional gluon is emitted which makes the baryon system a color singlet. In this case the corresponding DCB is an excited state, which can decay to the ground state and generate substantial contribution to the  $\Xi_{cc}^+$  sample. Therefore, in this thesis both configurations are considered<sup>[34,35]</sup>.

Using the perturbative QCD (pQCD) factorization theorem and the general-mass variable-flavor-number (GM-VFN) scheme, the production cross-section of two  $c$  quarks is written as follows:

$$\begin{aligned}
 \sigma &= F_{H_1}^g(x_1, \mu, m_c) F_{H_2}^g(x_2, \mu, m_c) \otimes \hat{\sigma}_{gg \rightarrow \Xi_{cc}}(x_1, x_2, \mu, m_c) \\
 &+ \sum_{i,j=1,2; i \neq j} F_{H_i}^g(x_1, \mu, m_c) \left[ F_{H_j}^c(x_2, \mu, m_c) - F_{H_j}^c(x_2, \mu, m_c)_{\text{SUB}} \right] \otimes \hat{\sigma}_{gc \rightarrow \Xi_{cc}}(x_1, x_2, \mu, m_c) \\
 &+ \sum_{i,j=1,2; i \neq j} \left[ \left( F_{H_i}^c(x_1, \mu, m_c) - F_{H_i}^c(x_1, \mu, m_c)_{\text{SUB}} \right) \left( F_{H_j}^c(x_2, \mu, m_c) - F_{H_j}^c(x_2, \mu, m_c)_{\text{SUB}} \right) \right] \\
 &\otimes \hat{\sigma}_{cc \rightarrow \Xi_{cc}}(x_1, x_2, \mu, m_c) + \dots, \tag{1-10}
 \end{aligned}$$

where high orders and the quark-antiquark annihilation processes are not included,  $F_H^i(x, \mu, m_c)$  (with  $H = H_1$  or  $H_2$  and  $x = x_1$  or  $x_2$ ) is the distribution function of parton  $i$  in the hadron  $H$ ,  $\hat{\sigma}$  is the cross-section of the corresponding process,  $\mu$  is the renormalization or factorization scale. <sup>①</sup> The subtraction for  $F_H^c(x, \mu, m_c)$ , defined to avoid ‘‘double counting’’, takes the following form

$$F_H^c(x, \mu, m_c)_{\text{SUB}} = F_H^g(x, \mu, m_c) \otimes F_g^c(x, \mu, m_c) = \int_x^1 \frac{dy}{y} F_g^c(y, \mu) F_H^g\left(\frac{x}{y}, \mu, m_c\right). \tag{1-11}$$

The second step is the fusion of the two  $c$  quarks into a  $(cc)$ -diquark with a configuration of  $(cc)_{\bar{3}}[{}^3S_1]$  or  $(cc)_6[{}^1S_0]$ . As a summary of the first two steps, the hadronic production of DCBs has contributions from

- LO:  $g + g \rightarrow (cc)_{\bar{3}}[{}^3S_1]c\bar{c}$  and  $g + g \rightarrow (cc)_6[{}^1S_0]c\bar{c}$ ,
- LO:  $g + c \rightarrow (cc)_{\bar{3}}[{}^3S_1]c\bar{c}$  and  $g + c \rightarrow (cc)_6[{}^1S_0]c\bar{c}$ ,
- NLO:  $c + c \rightarrow (cc)_{\bar{3}}[{}^3S_1]c\bar{c}$  and  $c + c \rightarrow (cc)_6[{}^1S_0]c\bar{c}$ , <sup>②</sup>

as shown in Figures 1.6, 1.7 and 1.8. The formation of the  $(cc)$ -diquark is a non-perturbative process but can be described by a matrix element in the non-relativistic QCD (NRQCD) framework, which is formulated as

$$\begin{aligned}
 (cc)_6[{}^1S_0] : h_1 &= \frac{1}{48} \langle 0 | [\psi^{a_1} \epsilon \psi^{a_2} + \psi^{a_2} \epsilon \psi^{a_1}] (a^\dagger a) \psi^{a_2 \dagger} \epsilon \psi^{a_1 \dagger} | 0 \rangle, \\
 (cc)_{\bar{3}}[{}^3S_1] : h_3 &= \frac{1}{72} \langle 0 | [\psi^{a_1} \epsilon \sigma^i \psi^{a_2} - \psi^{a_2} \epsilon \sigma^i \psi^{a_1}] (a^\dagger a) \psi^{a_2 \dagger} \sigma^i \epsilon \psi^{a_1 \dagger} | 0 \rangle, \tag{1-12}
 \end{aligned}$$

① The renormalization scale of the process,  $\mu_R$ , and the factorization scale for PDFs and hard processes factorization,  $\mu_F$ , are taken to be the same for convenience, *i.e.*  $\mu_R = \mu_F = \mu$ .

② The LO extrinsic charm fusion mechanisms only contribute to the purely longitudinal production, *i.e.*  $p_T = 0$ <sup>[35]</sup>, hence only the NLO mechanisms are considered.

where  $a_j(j = 1, 2)$  is the color of the valence quark fields,  $\sigma^i(i = 1, 2, 3)$  are Pauli matrices,  $\epsilon = i\sigma^2$ , and  $h_1$  and  $h_3$  represent the probability that the two  $c$  quarks merge into a diquark with  $(cc)_6[{}^1S_0]$  and  $(cc)_3[{}^3S_1]$ , respectively<sup>[35,37]</sup>. Note that the calculation of the matrix element involves the diquark wave function at the origin, but the diquark wave function is poorly known. Therefore, there could be sizable uncertainty in the calculation of the diquark fusion probability.

The third step is the hadronization of the diquark. It is assumed that the diquark is bound tightly enough so that the probability that a diquark fragments to a DCB is very close to 1<sup>[38,39]</sup>. Therefore, the production of a DCB is equivalent to that of a  $(cc)$ -diquark. Combining Eq. 1-10, 1-11 and 1-12, the numerical value of the production probability can be evaluated, and the results considering only the  $\Xi_{cc}^+$  baryon but not the  $\Xi_{cc}^-$  baryon are listed in Table 1.6. To estimate the lower bound, the value of the DCB cross-section at 14 TeV used for calculation is the smallest one of the predictions, *i.e.* the sum of the third and fourth rows of Table 1.6 since their fiducial ranges are almost complementary of each other. That is to say,  $\sigma_{\Xi_{cc}} \approx (63 + 59) = 122$  nb.

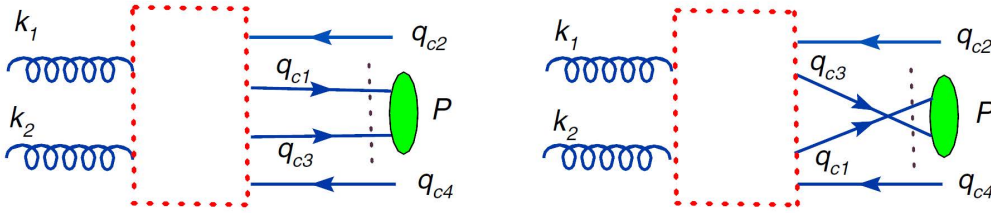


Figure 1.6 Schematic Feynman diagrams of the gluon-gluon fusion subprocess for the hadronic production of  $\Xi_{cc}$  baryons<sup>[35]</sup>. The dashed box denotes the interaction kernel,  $k_1$  and  $k_2$  are the momenta of the two gluons,  $q_{c2}$  and  $q_{c4}$  are the momenta of the two outgoing  $\bar{c}$  and  $P$  is the momentum of the diquark.

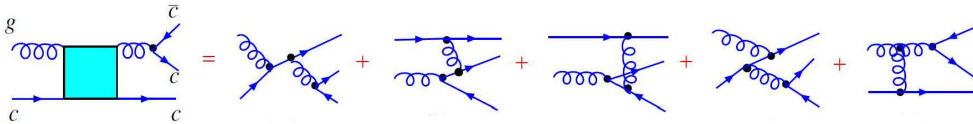


Figure 1.7 Typical Feynman diagrams of the gluon-charm scattering subprocess for the hadronic production of  $\Xi_{cc}$  baryons<sup>[35]</sup>.

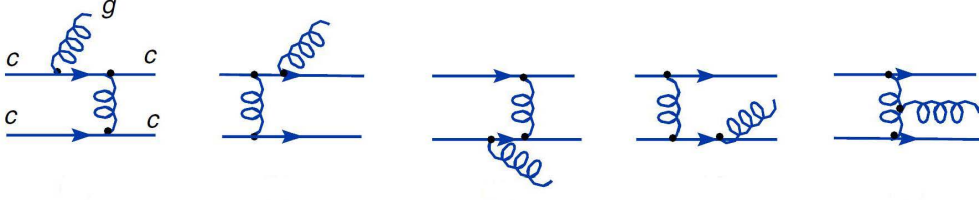


Figure 1.8 Typical Feynman diagrams of the charm-charm scattering subprocess for the hadronic production of  $\Xi_{cc}$  baryons<sup>[35]</sup>.

Table 1.6  $\Xi_{cc}$  production cross-section at the LHC

$\Xi_{cc}$ production cross-section (nb)	Fiducial Cut	Comment
1800 <sup>[34]</sup>	Not mentioned	
122 <sup>[36]</sup>	$ \eta  < 1$	$(cc)_6[{}^1S_0]$ not considered
63 <sup>[35]</sup>	$ \eta  < 1.5, p_T > 4 \text{ GeV}/c$	
59 <sup>[40]</sup>	$1.8 \leq  \eta  \leq 5.0, p_T > 4 \text{ GeV}/c$	LHCb acceptance $1.8 \leq \eta \leq 5.0$

The above calculations predict the production cross-section for all kinds of DCBs inclusively, not solely for the  $\Xi_{cc}^+$  baryon. Since there is no information about the decay of excited DCBs, for simplicity the fraction that an excited DCB decays to its corresponding ground state is assumed to be 100%. Then as long as a  $(cc)$ -diquark snatches a  $d$  quark, a  $\Xi_{cc}^+$  baryon will be produced, either directly from the collision or from the decay of the excited state. The probability that a  $(cc)$ -diquark captures a  $d$  quark is assumed to be  $(45 \pm 10)\%$ , according to the measured cross-section of charmed mesons<sup>[41]</sup><sup>①</sup>. Besides, the theoretical references only calculated the production of the  $\Xi_{cc}^+$  baryon, but not for the charge conjugation state  $\Xi_{cc}^-$ . For the three steps in which the DCB generation proceeds, the first two steps produce the same number of  $(cc)$ -diquark and  $(\bar{c}\bar{c})$ -diquark. At the third step, a  $cc$ -diquark needs a  $d$  quark to form a  $\Xi_{cc}^+$  baryon, while a  $(\bar{c}\bar{c})$ -diquark needs a  $\bar{d}$  antiquark to form a  $\Xi_{cc}^-$  baryon. The probabilities of these two processes are assumed to be the same, then it can be concluded that the production cross-sections of the  $\Xi_{cc}^+$  and  $\Xi_{cc}^-$  baryons are the same at the LHC. With these assumptions, the cross-section of the

① Note Table 2 of Ref.<sup>[41]</sup> contains contributions from higher excited charm states. These contributions should be subtracted before we compare the relative production ratio of  $D^0$ ,  $D^+$ , and  $D_s^+$ .

$\Xi_{cc}^+$  baryon (including  $\Xi_{cc}^+$  and  $\Xi_{cc}^-$ ) is obtained as

$$\sigma(pp \rightarrow \Xi_{cc}^\pm X) \approx 2 \times 122 \times 0.45 \approx 110 \text{ nb.} \quad (1-13)$$

Although calculations have been made, theoretical predictions of the  $\Xi_{cc}^+$  production cross-section have several uncertainties. First, the binding probability of the two  $c$  quarks might have considerable uncertainty since the diquark wave function is not known very well. Second, the assumption that a diquark has a probability close to 1 to hadronize to a  $\Xi_{cc}$  baryon still needs to be tested. Third, all the calculation is based on the picture that the two  $c$  quarks form a diquark in the  $\Xi_{cc}^+$  baryon, which is not verified by experiments.

#### 1.2.1.4 $\Xi_{cc}^+$ decay

The  $\Xi_{cc}^+$  baryon is a ground state and is expected to decay weakly. Although its decay modes have never been reliably observed yet, reasonable speculations can be made using the quark model. Since associated combinatorial background near the  $pp$  collision point is large and LHCb has a low efficiency to trigger and reconstruct neutral particles (see Chapter 2), final states used to reconstruct the  $\Xi_{cc}^+$  baryon should preferably contain charged particles and, if possible, have a long-lived daughter. Promising decays are listed below.

- $\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$
- $\Xi_{cc}^+ \rightarrow D^0 p K^- \pi^+$
- $\Xi_{cc}^+ \rightarrow D^+ p K^-$
- $\Xi_{cc}^+ \rightarrow \Xi_c^+ \pi^+ \pi^-$
- $\Xi_{cc}^+ \rightarrow \Xi_c^0 \pi^+$

This thesis focuses on the  $\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$  decay. The branching fraction (BF) of this decay is a priori unknown, therefore the BF of  $\Lambda_c^+ \rightarrow p K^- \pi^+$ , 5%, is used as an appropriate approximation since Feynman diagrams of these two modes are very similar.

#### 1.2.2 Experimental status

DCBs have been searched by many experiments, but the experimental knowledge is still limited. The only reported signal came from the SELEX experiment, but the measured properties did not agree with theoretical predictions very well. Other experiments all failed to observe this state.

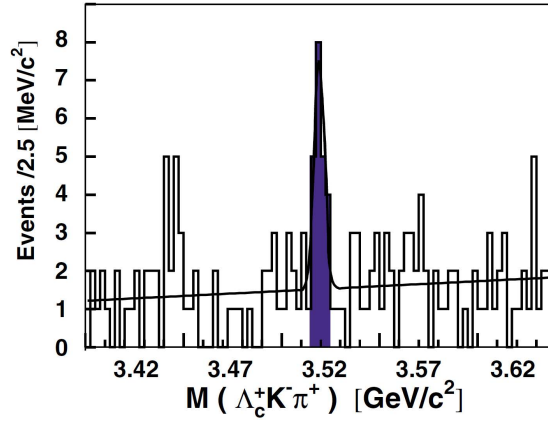


Figure 1.9 Invariant mass distribution of  $\Lambda_c^+ K^- \pi^+$  at SELEX<sup>[43]</sup>. The histogram is fitted using a Gaussian signal and a linear background.

### 1.2.2.1 SELEX

SELEX was a fixed target experiment employing  $\Sigma^-$ ,  $\pi^-$ , and proton beams of momenta around 600 GeV/c to study properties of charmed hadrons<sup>[42]</sup>. In 2002 the SELEX collaboration observed a state with a mass of 3520 MeV/c<sup>2</sup> at a significance of 6.3  $\sigma$  in the final state  $\Lambda_c^+ K^- \pi^+$ <sup>[43]</sup>. The observed invariant mass distribution is shown in Figure 1.9. This state was interpreted as the  $\Xi_{cc}^+$  baryon. However, one flaw was found in the computation of the signal significance: since SELEX searched the signal over a large mass range, a statistically significant observation may arise due to the large parameter space searched. This phenomenon, called the Look-Elsewhere Effect (LEE)<sup>[44]</sup>, was not taken into account. After the LEE correction, the probability that the peak is a fluctuation increases from  $1.0 \times 10^{-6}$  (6.3  $\sigma$ ) to  $1.1 \times 10^{-4}$  (3.9  $\sigma$ )<sup>[4]</sup>, which is below “the threshold” to announce a discovery. Trying to confirm the results, SELEX searched this state in other decay channels. In 2003 they declared the confirmation of the  $\Xi_{cc}^+(3520)$  state in the  $pD^+K^-$  final state with a significance of 4.8  $\sigma$ <sup>[45]</sup>, and the mass distribution is shown in Figure 1.10. In 2006 the same state was reported using the  $\Xi_c^+ \pi^+ \pi^-$  final state<sup>[46]</sup>, and the mass distribution is shown in Figure 1.11.

Beside the  $\Xi_{cc}^+(3520)$  state, SELEX also reported observation of other five states, which were all interpreted as DCBs. One excited DCB,  $\Xi_{cc}^{++}(3780)$ , and three DCBs,  $\Xi_{cc}^{++}(3460)$ ,  $\Xi_{cc}^{++}(3452)$  and  $\Xi_{cc}^{++}(3541)$ , were claimed in the  $\Lambda_c^+ K^- \pi^+ \pi^+$  final state, and one  $\Xi_{cc}^+(3443)$  baryon was declared in the  $\Lambda_c^+ K^- \pi^+$  final state<sup>[47,48]</sup>. Mass distributions of these final states are illustrated in Figure 1.12.

Although these states were all reported with high statistical significances, there were



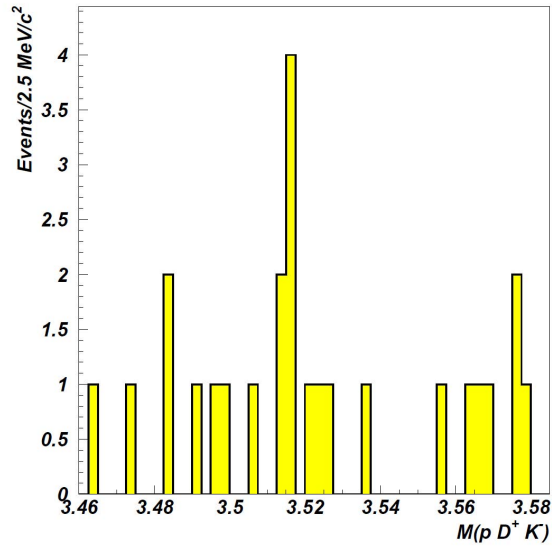


Figure 1.10 Invariant mass distribution of  $pD^+K^-$  at SELEX<sup>[45]</sup>.

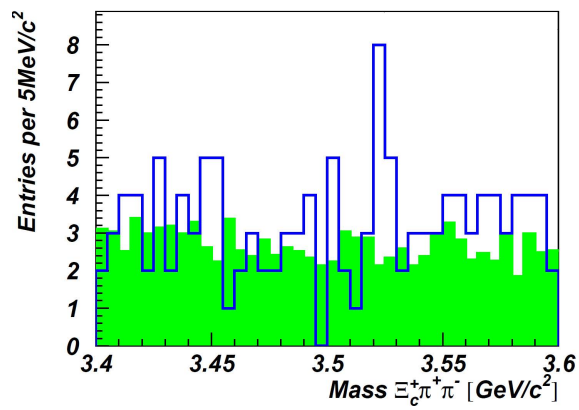


Figure 1.11 Invariant mass distribution of  $\Xi_c^+ \pi^+ \pi^-$  at SELEX<sup>[46]</sup>. The green histogram is the estimate of the combinatoric background.

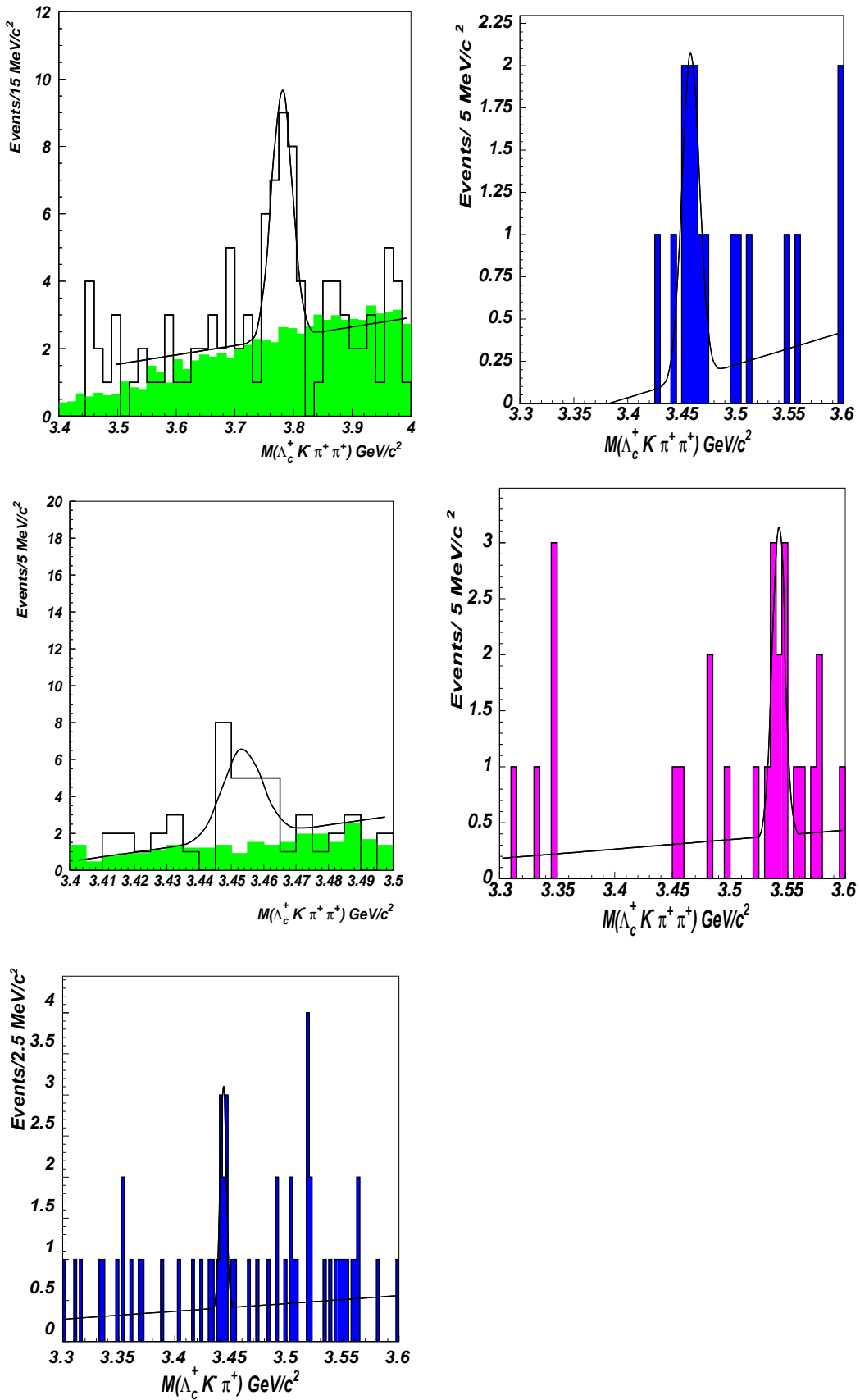


Figure 1.12 Other DCBs reported by SELEX. Top left for the observation of  $\Xi_{cc}^{++}(3780)$ <sup>[47]</sup>, top right for  $\Xi_{cc}^{++}(3460)$ <sup>[47]</sup>, middle left for  $\Xi_{cc}^{++}(3452)$ <sup>[47]</sup>, middle right for  $\Xi_{cc}^{++}(3541)$ <sup>[48]</sup>, and bottom left for  $\Xi_{cc}^+(3443)$ <sup>[47]</sup>.

several anomalies in SELEX observations. Let us focus on the  $\Xi_{cc}^+(3520)$  state. Besides the vulnerability of the computation of the signal significance, the disagreement between the predicted and measured lifetime and the disagreement between the measured  $R$  value by SELEX and other experiments, where  $R$  is the relative production rate of  $\Xi_{cc}^+$  to  $\Lambda_c^+$ , also make the results less convincing. An additional criticism is that the production rate of this state is much higher than theoretical predictions, *i.e.* the production ratio of the  $\Xi_{cc}^+$  baryon to all charm hadrons at SELEX is estimated to be  $\sigma_{\Xi_{cc}^+}/\sigma_{\text{charm}} \sim 10^{-6} - 10^{-5}$ , while SELEX measured the ratio to be approximately  $2 \times 10^{-2}$ , which is several order of magnitudes higher than predictions<sup>[23]</sup>. (The total sample of SELEX contains 1630 reconstructed  $\Lambda_c^+$ , 10210 reconstructed  $D^0$ <sup>[49]</sup>, and 3900 reconstructed  $D^\pm$ <sup>[45]</sup>.) But note Ref.<sup>[23]</sup> calculates the ratio based on  $\pi^-p$  and  $p-p$  scattering, while SELEX claimed that the  $\Xi_{cc}^+$  baryon are mostly generated in the  $\Sigma^-$  beam. Therefore no obvious contradiction can be derived from these comparisons. SELEX measured a small but non-zero lifetime for this state (smaller than 33 fs @ 90% C.L.), which suggests it to be the ground state  $\Xi_{cc}^+$  rather than an excited state. However, theoretical predictions of the  $\Xi_{cc}^+$  lifetime mainly lie between 100 – 250 fs, which is incompatible with the experimental result. The  $R$  measured at SELEX is two orders of magnitude higher than other experiments, *e.g.* FOCUS, as described in the next section.

### 1.2.2.2 FOCUS

The FOCUS experiment was a heavy flavor photo-production program with a center-of-mass energy of 200 GeV<sup>[50]</sup>. They searched for the  $\Xi_{cc}^+$  and  $\Xi_{cc}^{++}$  baryons in various decay modes with the assumption that the DCBs have a mass of 3.6 GeV/ $c^2$  and a lifetime of 1.0 ps ( $\Xi_{cc}^{++}$ ) and 0.2 ps ( $\Xi_{cc}^+$ )<sup>[51]</sup>, which do not overlap with the corresponding values of SELEX, especially for the lifetime. Figure 1.13 shows the mass distributions of the  $\Xi_{cc}^+$  and  $\Xi_{cc}^{++}$  baryons in each decay mode, and Figure 1.14 shows the sum of mass distributions of all decay modes. No significant evidence for either of the DCBs was found.

In particular, if only the two decay modes searched for by SELEX,  $\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$  and  $\Xi_{cc}^+ \rightarrow D^+ p K^-$ , are studied, no clear evidence for these states is observed in the SELEX signal region, as shown in Figure 1.15. FOCUS set a Rolke-Lopez upper limit of 2.21 events @ 90% C.L. for both  $\Xi_{cc}^+$  and  $\Xi_{cc}^{++}$  baryons<sup>[52]</sup>. A comparison of the results of FOCUS and SELEX is shown in Table 1.7. The  $R$  value measured at SELEX is two

Table 1.7 Comparison between the results of FOCUS and SELEX

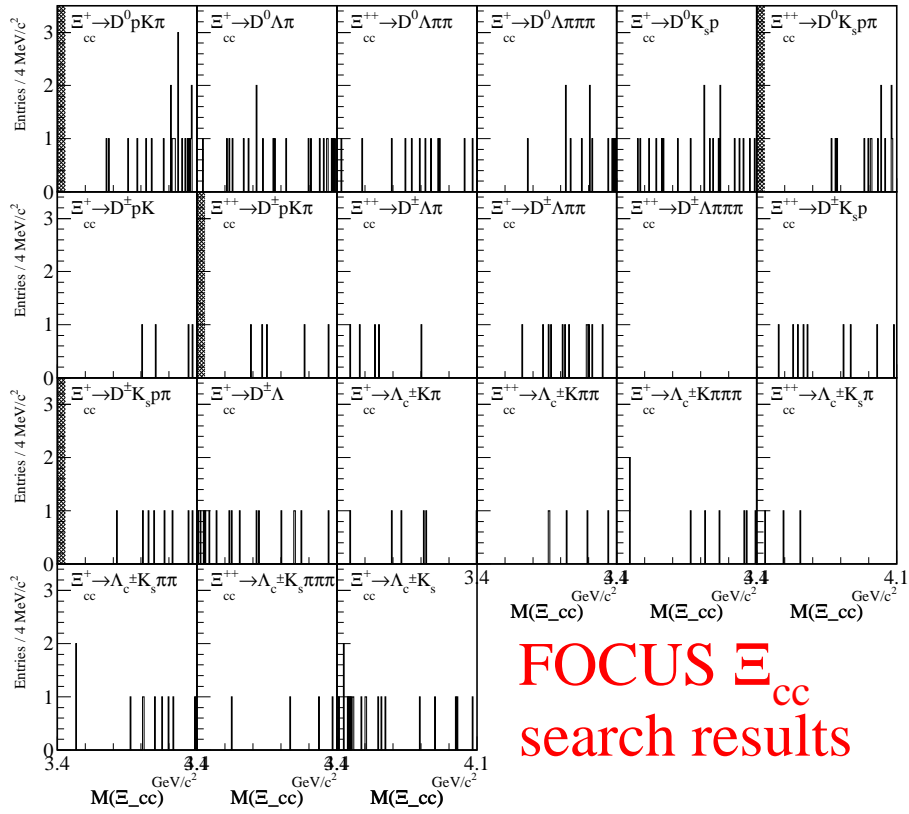
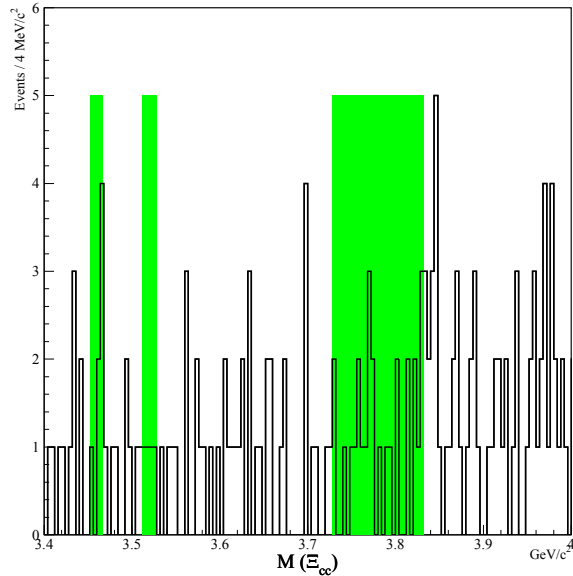
Decay mode	$\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$		$\Xi_{cc}^{++} \rightarrow \Lambda_c^+ K^- 2\pi^+$	
	FOCUS	SELEX	FOCUS	SELEX
Experiment				
$\Xi_{cc}$ candidates	< 2.21 @ 90% CL	15.8	< 2.21 @ 90% CL	8
$\Lambda_c^+$ candidates	$19444 \pm 262$	1650	$19444 \pm 262$	1650
$\epsilon(\Xi_{cc})$ relative to $\Lambda_c^+$	5%	10%	13%	5%
$R$	0.23% @ 90% CL	9.6%	0.09% @ 90% CL	9.7%
Ratio of $R$ at SELEX and FOCUS	> 42 @ 90% CL		> 111 @ 90% CL	

orders of magnitude higher than that of FOCUS, although it should be noted that the production environments are different and the efficiency at FOCUS is estimated based on  $\tau(\Xi_{cc}^+) = 0.2$  ps and  $\tau(\Xi_{cc}^{++}) = 0.1$  ps, while SELEX measured  $\tau(\Xi_{cc}^+) < 0.033$  ps @ 90% C.L. Nevertheless, if we assume that the efficiency of the selection criteria of FOCUS is insensitive to the lifetimes of DCBs, then the results of FOCUS and SELEX are difficult to reconcile<sup>[53]</sup>.

### 1.2.2.3 Belle

The Belle detector is a magnetic spectrometer dedicated to the study of  $b$  physics<sup>[54]</sup>. It is located at the KEK asymmetric-energy  $e^+e^-$  collider. In 2006 they searched for the DCB  $\Xi_{cc}^+$  in the  $\Lambda_c^+ K^- \pi^+$  final state using a sample corresponding to an integrated luminosity of  $461.5 \text{ fb}^{-1}$ <sup>[55]</sup>. No significant evidence for the  $\Xi_{cc}^+$  baryon was found, as shown in Figure 1.16.

In 2013 Belle updated the search using more final states with the full data set, which corresponds to an integrated luminosity of  $980 \text{ fb}^{-1}$ <sup>[56]</sup>. Note that search at Belle is insensitive to the lifetime of  $\Xi_{cc}$  baryons. The  $\Xi_{cc}^+$  ( $\Xi_{cc}^{++}$ ) baryon is searched in the decay modes  $\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+(\pi^+)$  and  $\Xi_{cc}^+ \rightarrow \Xi_c^0 \pi^+(\pi^+)$ . For the final states containing the  $\Lambda_c^+$  baryon, the decay modes  $pK^- \pi^+$  and  $pK_S^0$  are used to reconstruct the  $\Lambda_c^+$  candidates. The invariant mass distributions of  $\Xi_{cc}^+$  and  $\Xi_{cc}^{++}$  candidates are shown in Figure 1.17. For the final states containing the  $\Xi_c^0$ , three final states,  $\Xi^- \pi^+$ ,  $\Lambda K^- \pi^+$  and  $pK^- K^- \pi^+$ , are used to reconstruct the  $\Xi_c^0$  candidates. The mass distributions of  $\Xi_{cc}^+$  and  $\Xi_{cc}^{++}$  candidates in the final state  $\Xi_c^0 \pi^+(\pi^+)$  are shown in Figure 1.18 and Figure 1.19, respectively. No significant  $\Xi_{cc}$  signal is observed in these decay modes.


 Figure 1.13 Mass distribution of DCBs candidates in various decay modes at FOCUS<sup>[51]</sup>.

 Figure 1.14 Sum of mass distributions of all decay modes at FOCUS<sup>[51]</sup>. The three shaded regions correspond to the three states reported by SELEX,  $\Xi_{cc}^{++}$  (3460),  $\Xi_{cc}^+$  (3520), and  $\Xi_{cc}^{++}$  (3780).

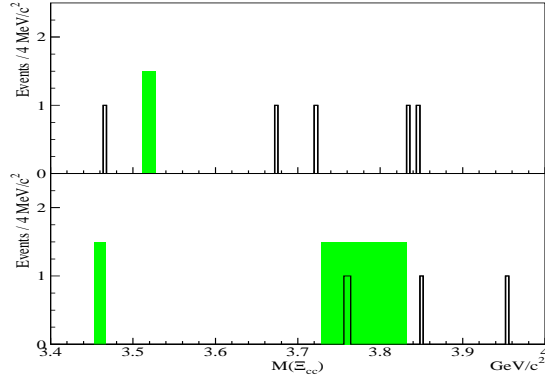


Figure 1.15 Invariant mass distribution of the  $\Xi_{cc}^+$  candidates (top) and the  $\Xi_{cc}^{++}$  candidates (bottom) for FOCUS<sup>[51]</sup>. The shaded region on the top plot indicates the state  $\Xi_{cc}^+(3520)$ , and the regions in the bottom plot correspond to the states  $\Xi_{cc}^{++}(3460)$  and  $\Xi_{cc}^{++}(3780)$ . No events is observed in the signal region of  $\Xi_{cc}^+(3520)$  and  $\Xi_{cc}^{++}(3460)$ , and only one event is observed in the signal region of  $\Xi_{cc}^{++}(3780)$ .

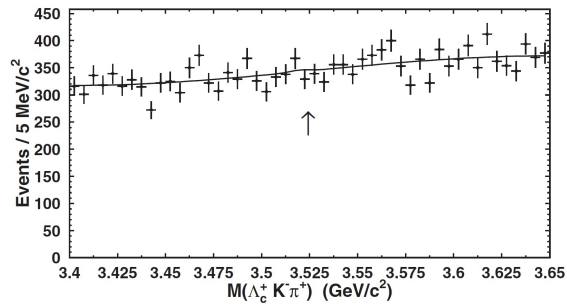


Figure 1.16 Invariant mass distribution of  $\Lambda_c^+ K^- \pi^+$  at Belle<sup>[55]</sup>. The mass of the SELEX candidate is indicated by an arrow.

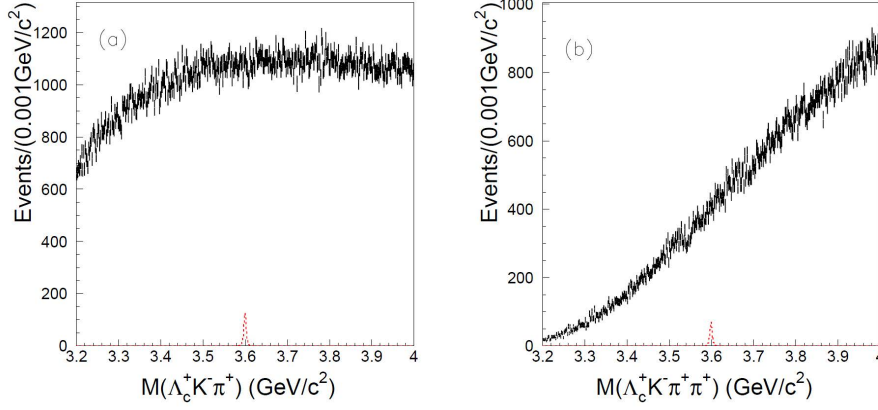


Figure 1.17 Invariant mass distribution of the  $\Xi_{cc}$  candidates for  $\Lambda_c^+ K^- \pi^+$  (left) and  $\Lambda_c^+ K^- \pi^+ \pi^+$  (right) final states at Belle<sup>[56]</sup>. The dashed histograms are signal MC assuming  $\sigma(e^+e^- \rightarrow \Xi_{cc}^{+(+)}) = 500 \text{ fb}^{-1}$  and  $\mathcal{B}(\Xi_{cc}^{+(+)} \rightarrow \Lambda_c^+ K^- \pi^+ (\pi^+)) = 5\%$ .

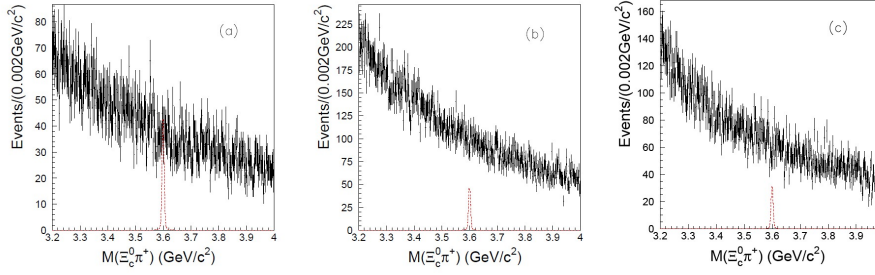


Figure 1.18 Invariant mass distribution of the  $\Xi_{cc}^+$  candidates for  $\Xi_{cc}^0 \pi^+$  final state with  $\Xi_{cc}^0$  reconstructed using  $\Xi^- \pi^+$ ,  $\Lambda K^- \pi^+$ , and  $p K^- K^- \pi^+$  final states at Belle<sup>[56]</sup>. The dashed histograms are signal MC assuming  $\sigma(e^+e^- \rightarrow \Xi_{cc}^+ X) = 500 \text{ fb}^{-1}$ ,  $\mathcal{B}(\Xi_{cc}^+ \rightarrow \Xi_{cc}^0 \pi^+) = 5\%$  and  $\mathcal{B}(\Xi_{cc}^0 \rightarrow \Xi^- \pi^+) = 5\%$ .

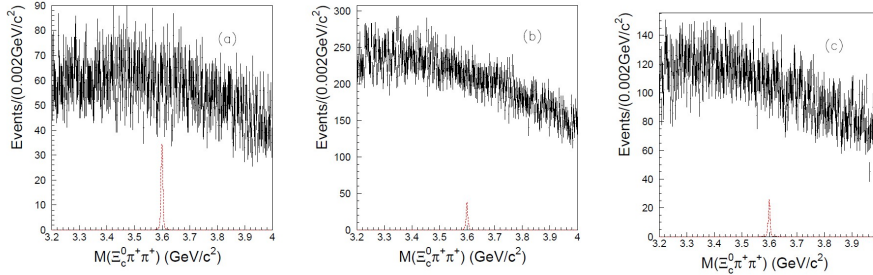


Figure 1.19 Invariant mass distribution of the  $\Xi_{cc}^{++}$  candidates for  $\Xi_{cc}^0 \pi^+$  final state with  $\Xi_{cc}^0$  reconstructed using  $\Xi^- \pi^+$ ,  $\Lambda K^- \pi^+$ , and  $p K^- K^- \pi^+$  final states at Belle<sup>[56]</sup>. The dashed histograms are signal MC assuming  $\sigma(e^+e^- \rightarrow \Xi_{cc}^{++} X) = 500 \text{ fb}^{-1}$ ,  $\mathcal{B}(\Xi_{cc}^{++} \rightarrow \Xi_{cc}^0 \pi^+) = 5\%$  and  $\mathcal{B}(\Xi_{cc}^0 \rightarrow \Xi^- \pi^+) = 5\%$ .

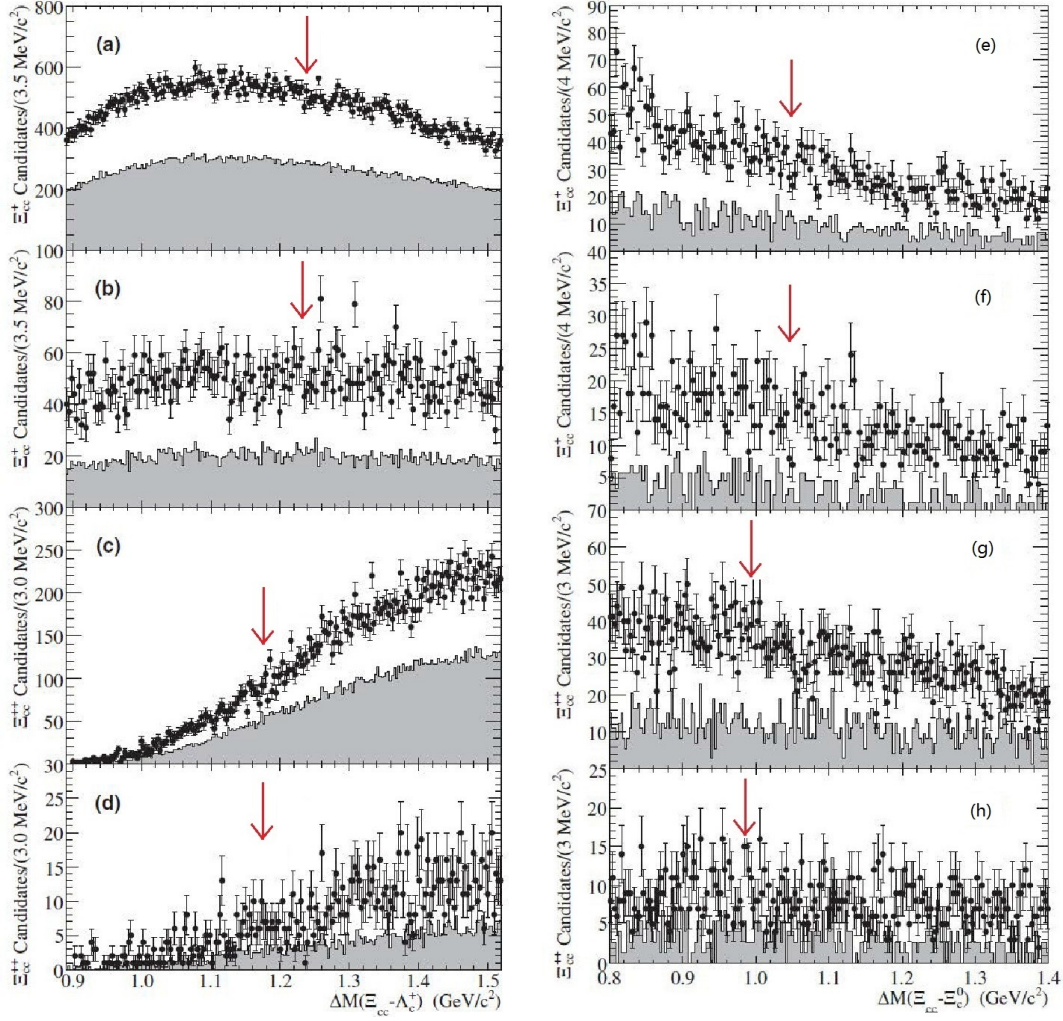


Figure 1.20 Distributions of  $\Delta M(\Xi_{cc} - \Lambda_c^+)$  (left) and  $\Delta M(\Xi_{cc} - \Xi_c^0)$  (right) for (a,b,e,f)  $\Xi_{cc}^+$  and (c,d,g,h)  $\Xi_{cc}^{++}$  candidates with (a,c,e,g) no  $p^*$  requirements and (b,d,g,h)  $p^* > 2.3$  GeV/c (2.0 GeV/c for right plot), where  $p^*$  is the momentum of  $\Xi_{cc}$  candidates in the center-of-mass system. Shaded histograms correspond to candidates in the sideband region of  $\Lambda_c^+$ .

#### 1.2.2.4 BaBar

The BaBar detector, a spectrometer at the SLAC PEP-II asymmetric  $e^+e^-$  collider, was designed to perform comprehensive studies of CP-violation in B-meson decays<sup>[57]</sup>. They also searched for the  $\Xi_{cc}^+$  ( $\Xi_{cc}^{++}$ ) baryon in the decays  $\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+ (\pi^+)$  and  $\Xi_{cc}^+ \rightarrow \Xi_c^0 \pi^+ (\pi^+)$ <sup>[58]</sup>. To improve resolution, the difference between the reconstructed mass of the  $\Xi_{cc}$  candidates and the singly charmed baryon,  $\Delta M$ , is used as the observable. The distributions of  $\Delta M(\Xi_{cc} - \Lambda_c^+)$  and  $\Delta M(\Xi_{cc} - \Xi_c^0)$  are illustrated in Figure 1.20. No clear evidence for the  $\Xi_{cc}$  signal is observed in all these distributions.



### 1.2.2.5 Summary

Many experiments have searched for the  $\Xi_{cc}^+$  baryon, but different results have been reported. SELEX claimed the observation of the  $\Xi_{cc}^+$  baryon with a lifetime much shorter than the predictions using the OPE formalism. However, the OPE formalism allowed to precisely predict lifetimes of many hadrons, including many charm hadrons. It would be surprising to find that the OPE formalism only fails for the case of the  $\Xi_{cc}^+$  baryon. Besides, SELEX collaboration computed the signal significance in a vulnerable way; the probability that the peak is a fluctuation is larger than they suggested. Other experiments failed to find the signal of the  $\Xi_{cc}^+$  baryon, but given different production environments and different lifetime assumptions, it is imprudent to conclude that SELEX observed a false signal. For the time being no solid conclusion about the experimental situation of the  $\Xi_{cc}^+$  baryon can be drawn. The only way to solve this puzzle is to keep searching at other experiments until we find the genuine  $\Xi_{cc}^+$  baryon or a state with the same properties as the one found by SELEX.

## Chapter 2 The LHCb experiment

### 2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a two-ring super-conducting proton-proton collider located in a 27 km long tunnel at CERN<sup>[59]</sup>. It is designed to be a multi-TeV machine, since the no-lose theorem guarantees that the LHC is certain to discover something at this energy scale<sup>[60]</sup>. First, one of the most important pieces of the SM, the Higgs boson, was still missing at the time the LHC was proposed. The mass of the Higgs boson cannot be calculated from the SM directly, but it can be constrained by precision measurements of the weak interaction<sup>[61]</sup>. The global fit shows the Higgs boson must have a mass smaller than  $250 \text{ GeV}/c^2$ , otherwise the SM will be strongly inconsistent with data<sup>[62]</sup>, as shown in Figure. Second, there are cosmological evidences as well as aesthetic arguments which indicate that there should be physics beyond the SM. Supersymmetry (SUSY) is the most well studied theory among those possible extensions of the SM. The naturalness of SUSY implies that new physics must appear at a scale of  $\mathcal{O}(1 \text{ TeV})$  or below<sup>[63]</sup>. With the ability to directly access physics at terascale, the LHC discovered the Higgs boson and searched for new physics beyond the SM.

In the planned running conditions, the LHC has two 7 TeV counter-rotating proton beams, each of which contains 2,808 bunches. These bunches, each carrying  $1.15 \times 10^{11}$  protons, are separated by 25 ns intervals<sup>[64]</sup>. The layout of the LHC acceleration chain is shown in Figure 2.2. The acceleration begins at the Linear accelerator 2 (Linac 2). As the pre-injector of the LHC inject chain, the machine produces protons by stripping the electrons from hydrogen atoms, accelerates them to 50 MeV using three Alvarez tanks and feeds them to the Proton Synchrotron Booster(PSB)<sup>[65]</sup>. Protons are further accelerated to 1.4 GeV in one of the PSB's four rings and then delivered to the Proton-Synchrotron (PS), which accelerates protons to 25 GeV and divides them to form nominal LHC 25 ns bunch trains. The beams are then accepted by the Super-Proton-Synchrotron (SPS), which accelerates them to 450 GeV and injects them to the LHC. Finally, the LHC accelerates protons to the desired energy. To manipulate proton beams with energies of multiple TeV, an extremely strong bending magnetic field

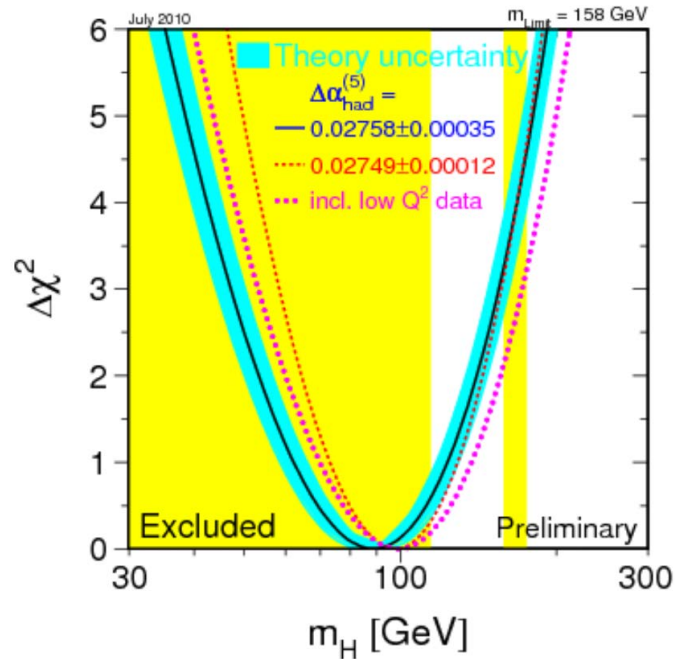


Figure 2.1 Global fit of the Higgs mass in 2010. The yellow shaded area indicates the Higgs mass region that had been excluded.

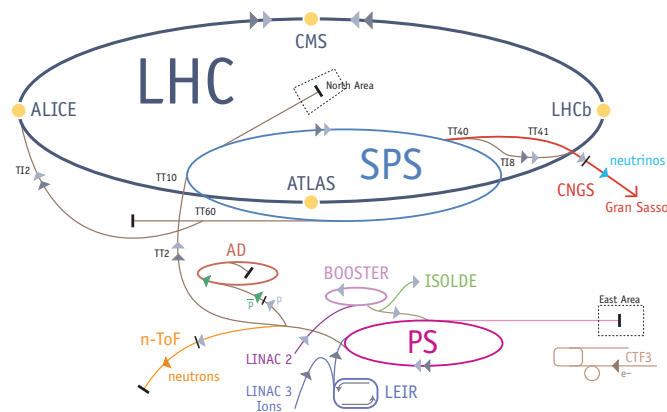


Figure 2.2 Layout of the LHC accelerator complex.

is required. It is achieved by cooling 1,232 magnets using super-fluid Helium to an operating temperature of 1.9 K. The field has a strength of 8.33 T, which is near the limit of current technologies.

Four main experiments are installed along the LHC ring for different physics purposes. Two general purpose experiment, ATLAS<sup>[66]</sup> and CMS<sup>[67]</sup>, have been designed to discover the Higgs boson and search for new physics directly. The LHCb<sup>[68]</sup> is a dedicated heavy flavor physics experiment, whose primary goals are to perform precise

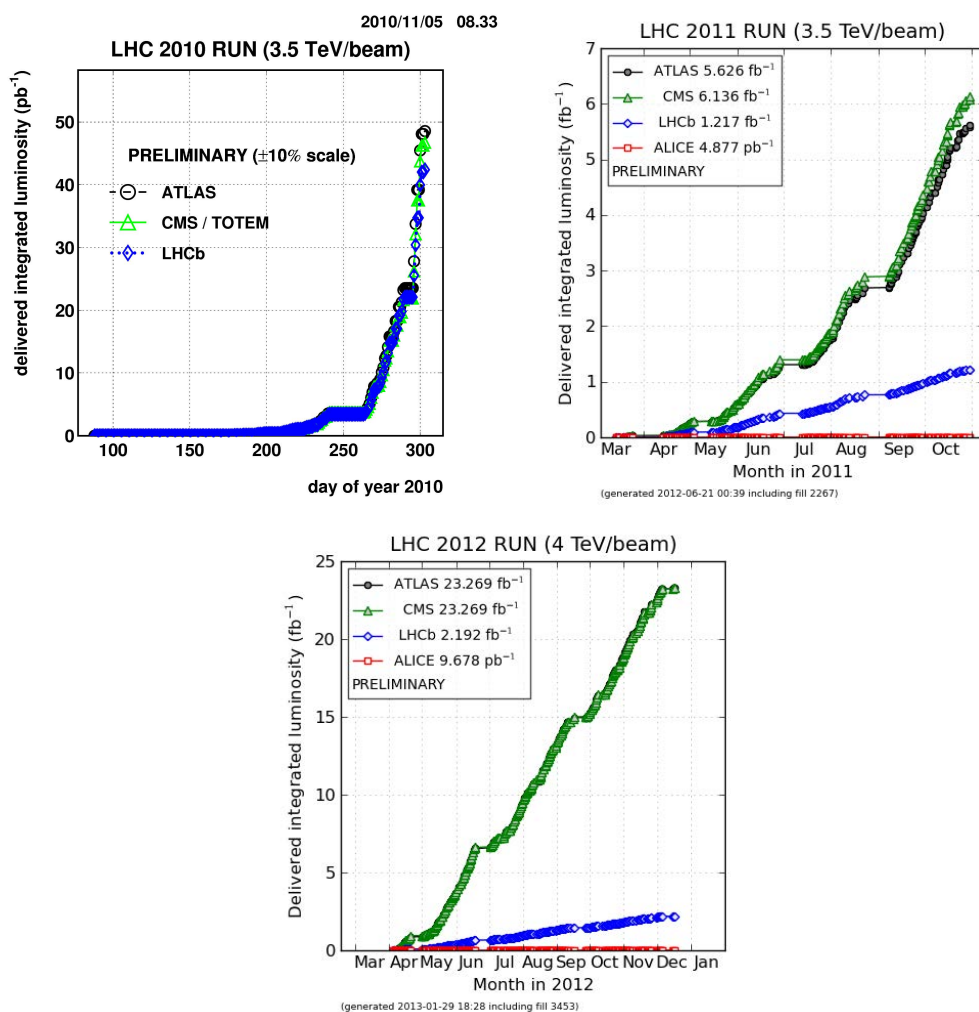


Figure 2.3 Delivered integrated luminosity for each detector in 2010 (top left), 2011 (top right) and 2012 (bottom), respectively. The LHC delivered integrated luminosity of 50  $\text{pb}^{-1}$ , 6  $\text{fb}^{-1}$  and 25  $\text{fb}^{-1}$  for ATLAS and CMS, and 40  $\text{pb}^{-1}$ , 1.2  $\text{fb}^{-1}$  and 2.2  $\text{fb}^{-1}$  for LHCb in 2010, 2011 and 2012, respectively.

CKM measurements and to search for new physics in CP symmetry violation and rare decays via processes involving loops. The ALICE is optimized for heavy-ion physics to study the physics of strongly interacting matter at extreme energy densities<sup>[69]</sup>.

The LHC circulated two proton beams simultaneously for the first time in November 2009. In March 2010 it successfully collided two proton beams at a center-of-mass energy of 7 TeV and was providing data to experiments at this energy until end of 2011. The center-of-mass energy was raised to 8 TeV in 2012. Figure 2.3 shows the integrated luminosity delivered to each detector. LHCb recorded integrated luminosity of 37.0  $\text{pb}^{-1}$  in 2010, 1.0  $\text{fb}^{-1}$  in 2011 and 2  $\text{fb}^{-1}$  in 2012.

## 2.2 The LHCb experiment

LHCb is an experiment designed to precisely test the SM and to search for indirect evidence of new physics in CP violation and rare decays of beauty and charm hadrons<sup>[68]</sup>. The SM is a very successful theory, which explains almost all experimental observations. The only source of CP violation in the SM is the CKM mechanism<sup>[70]</sup>. However, it is believed to be insufficient to explain the observed baryon asymmetry in the universe. New sources of CP violation are needed. Numerous new physics models produce additional contributions to CP violation and permit decay modes, which are forbidden or highly suppressed in the SM. Precise measurements of the CP violation parameters and the search for rare decays can potentially unveil new physics. Physical processes involving  $b$  quarks are theoretically cleaner in the SM, which makes the separation of the SM and new physics easier. Therefore it is an ideal place for the new physics search at LHC energies. LHCb profits from large cross-sections for  $b\bar{b}$  ( $288 \pm 4(\text{stat.}) \pm 48(\text{syst.}) \text{ mb}^{\text{[71]}}$ ). With the large samples of beauty hadrons and an optimized detector, LHCb is able to determine the SM parameters to an unprecedented precision.

The  $b\bar{b}$  pair production process at the LHC is dominated by the gluon-gluon fusion process<sup>[72]</sup>. Since the two gluons usually have a large boost along the beam axis, the produced  $b$  quark and  $\bar{b}$  antiquark tend to have similar rapidity, as shown in Figure 2.4. With this important observation, the LHCb detector is designed to be a single arm spectrometer, with a forward angular acceptance of 10 mrad to 250 mrad vertically and 10 mrad to 300 mrad horizontally. Covering only 4% of the full solid angle, LHCb collects 27% of  $b$  or  $\bar{b}$  hadrons produced.

The nominal LHC luminosity is of the order of  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , which is not suitable for LHCb due to multiple reasons. The high luminosity causes  $\mathcal{O}(10)$  proton-proton interactions per bunch crossing. The consequent detector occupancy will be too high to be accepted by LHCb. Difficulties also emerge in  $b$ -tagging and lifetime measurements as a result of ambiguities in the primary vertex determination or . The radiation damage is also harsher at high luminosity. Therefore, the luminosity at LHCb is decreased to  $2 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ . This is achieved by a local de-focusing of the LHC beams at the LHCb interaction point.

A precise tracking system, an efficient particle identification (PID) system and a flexible trigger system are essential for precise measurements of heavy flavor physics. At LHCb the tracking system is composed of a vertex locator (VELO), a trigger tracker

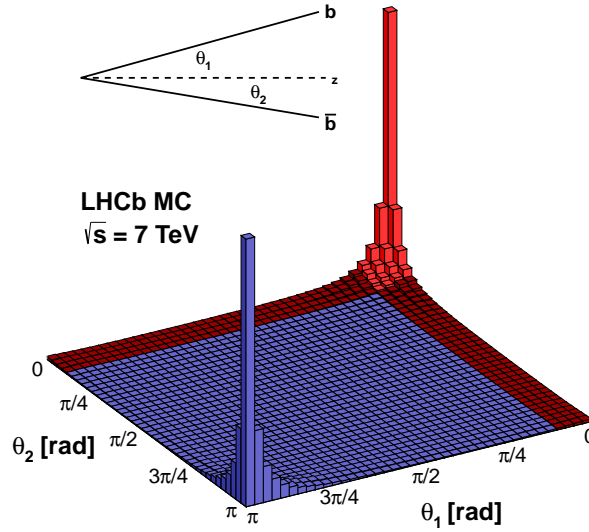


Figure 2.4 Simulated distribution of polar angles of  $b$  or  $\bar{b}$  hadrons with respect to the beam axis, which shows that the kinematics of the two  $b$  hadrons are highly correlated.

(TT), an inner tracker (IT), an outer tracker (OT) and a dipole magnet. The PID system is composed of two ring imaging Cherenkov detectors (RICH1 and RICH2), a calorimeter system (comprised of a pre-shower(PS), a scintillating pad detector (SPD), an electromagnetic calorimeter (ECAL) and a hadron calorimeter (HCAL)) and a muon detector. The layout of the LHCb detector is shown in Figure 2.5. Detailed description of these sub-detectors are given in the following subsections.

## 2.2.1 Tracking system

The tracking system serves to reconstruct trajectories of charged particles and to determine their momentum precisely. Tracks are reconstructed across the spectrometer by combining hits in the VELO with those reconstructed in the main tracker. Reconstructed tracks are used to perform the vertex fit<sup>[73]</sup>.

### 2.2.1.1 Vertex locator

The VELO provides information to reconstruct primary vertices and displaced secondary vertices, which provide a signature of heavy flavor decays<sup>[68]</sup>, and is therefore an important part of the LHCb tracking system. It consists of 21 silicon tracking stations,

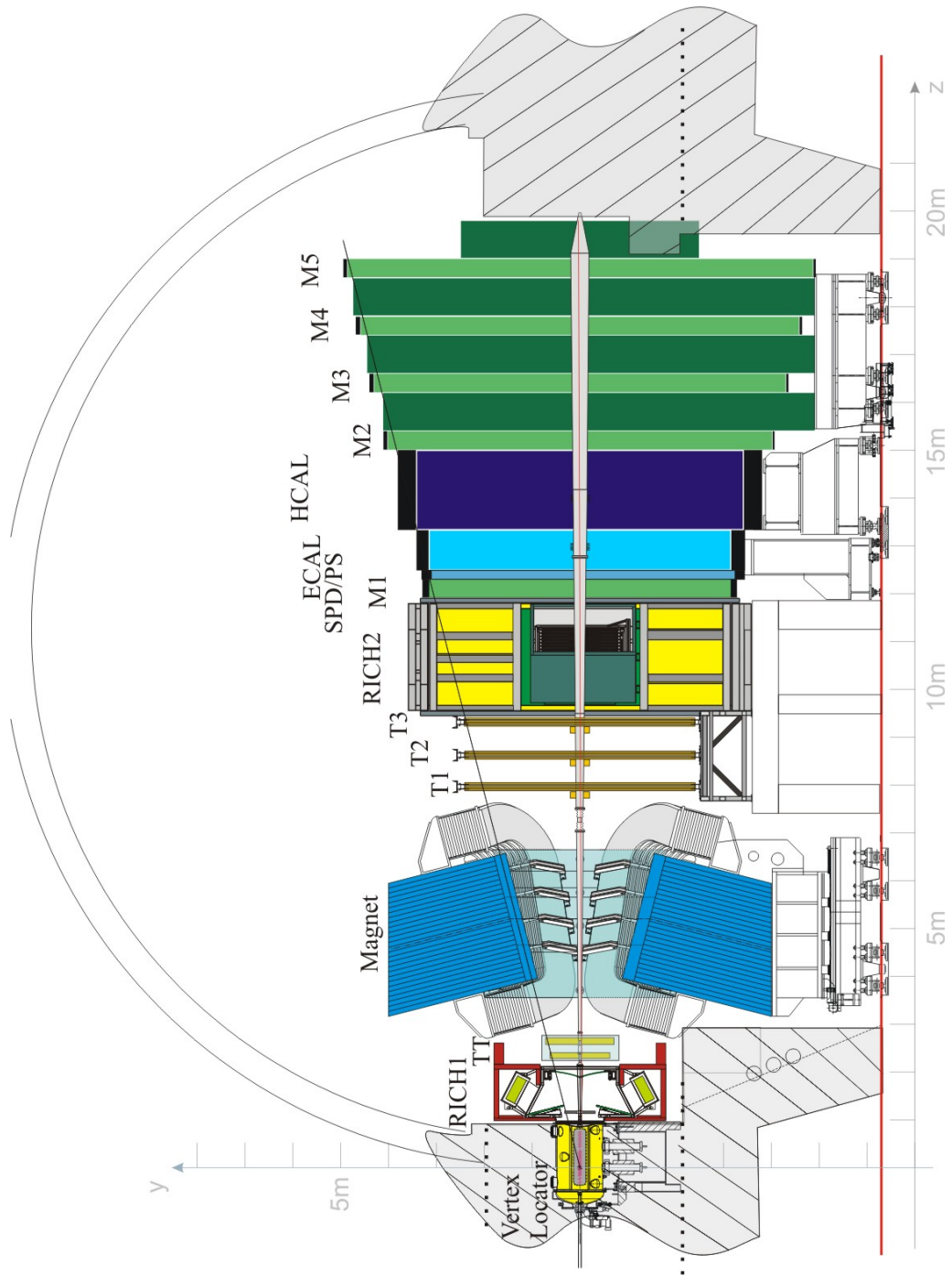


Figure 2.5 Layout of the LHCb detector. The tracking system includes the vertex locator (VELO) situated around the proton-proton interaction point, the trigger tracker, three tracking stations and the dipole magnet. The particle identification system includes two RICH detectors, the calorimeter system (SPD, PS, ECAL and HCAL) and the muon detector (M1-M5) [68].

as shown in Figure 2.6. To improve the resolution of the primary vertex (PV), some stations are installed upstream of the nominal interaction point. Each station is divided into two retractable modules, which are designed as a semi-circular plate rather than a simpler rectangular scheme to adopt the cylindrical geometry ( $R\phi$  coordinates), which allows faster track and vertex reconstruction in the LHCb trigger. To achieve better resolution of impact parameters, the track extrapolation length should be reduced, which requires the inner radius of the module to be as close to the beam axis as possible. However, due to the yet-unknown closed-orbit variation of the LHC, the allowed closest approach to the beam axis is 5 mm, with the sensitive area begins at a distance of  $\sim 8$  mm<sup>[68]</sup>. During beam injection and acceleration, the increased size of the beam spot requires two halves of the VELO to be retracted by a distance of 3 cm. When closed, the two modules of each station are required to overlap to cover the full azimuthal acceptance and to improve the detector alignment. The module has an outer radius of 42 mm. Each module contains two different kinds of silicon sensor on two sides, one to measure the radial distance (the  $R$ -sensor) and another to measure the azimuthal angle (the  $\phi$ -sensor), as shown in Figure 2.7. Each sensor,  $R$ -type or  $\phi$ -type, has 2048 readout channels. To minimize the occupancy and reduce strip capacitance, the  $R$ -sensor is divided into four regions, each with 512 concentric strips. The strip pitch increases linearly from the inner edge to the outer edge to stabilize the strip occupancy across the region. The  $\phi$ -sensor is divided into two subsections, inner and outer, to reduce strip occupancies and to keep the strip pitch not too large at the outer edge. The inner subsection has 683 strips and the pitch increases linearly from the inner radius to the border, while the outer subsection has 1365 strips and the pitch increases linearly from the border to the outer radius. The  $\phi$ -sensors in adjacent modules have opposite skews with respect to each other so that ghost hits can be distinguished from true hits through a stereo view<sup>[68]</sup>.

The VELO shows an excellent performance during the data-taking period. Figure 2.8 shows the reconstructed primary vertex resolution in 2010 data as a function of the number of tracks used to reconstruct the vertex. For a vertex reconstructed by 25 tracks, a resolution of 14  $\mu\text{m}$  in  $x$  direction and 75  $\mu\text{m}$  in  $z$  direction can be achieved. The performance is similar for early 2011 data, as shown in Figure 2.9. Figure 2.10 illustrates the impact parameter resolution for the early 2011 data and the simulated sample as a function of the inverse of the transverse momentum ( $p_T$ ). Due to imperfect Monte Carlo modeling, *e.g.* inaccurate description of multiple scattering and the detector material,



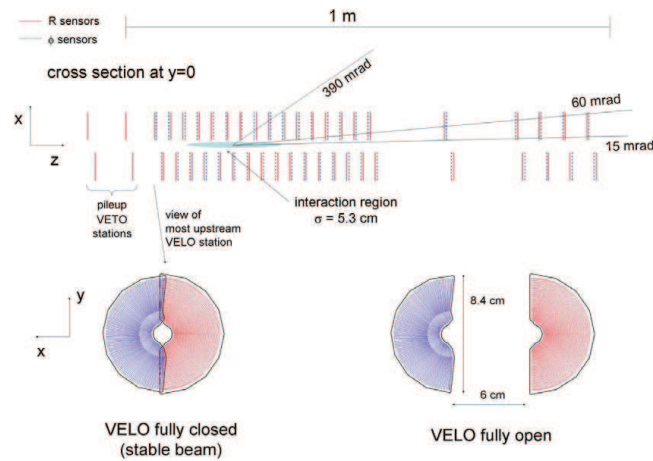


Figure 2.6 Top plot shows the cross-section of the VELO in the  $x-z$  plane, with the sub-detector in the fully closed position. The VELO consists of 21 tracking stations, each divided into two retractable modules. The two pile-up stations located upstream of the VELO are also shown. The bottom plots show the front-face of the module in the fully closed (left) and fully open (right) position<sup>[68]</sup>.

discrepancies are observed between the data and the simulated samples. The VELO meets the requirements of tracking and vertexing at LHCb, and has key contributions to LHCb physics results.

### 2.2.1.2 Trigger tracker

The TT is a silicon detector located between the RICH1 detector and the dipole magnet and performs fast momentum measurements for trigger decision using residual magnetic field. It consists of four rectangular detection layers, which are approximately 150 cm in width and 130 cm in height to cover the full LHCb acceptance. The layout of the third TT detection layer is shown in Figure 2.11. The four layers are settled in the sequence of  $x-u-v-x$  with strips in the  $x$  layers arranged vertically, while strips in the  $u$  and  $v$  layers rotated  $+5^\circ$  and  $-5^\circ$ , respectively. To facilitate tracking finding, the four layers are divided into two groups,  $(x, u)$  and  $(v, x)$ . The detector layer consists of two half modules, an upper half and a lower half, each comprised of a row of seven silicon sensors organized into two or three readout sectors depending on their position relative to the beam pipe. The organization of a half module close to the beam pipe is shown in Figure 2.12.

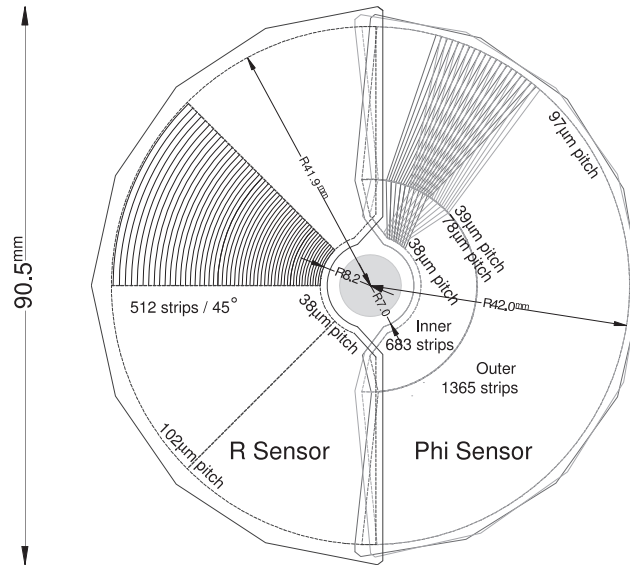


Figure 2.7  $R\phi$  geometry of the VELO sensors with only a portion of the strips shown. For the  $\phi$ -sensors strips on two adjacent modules are indicated [68].

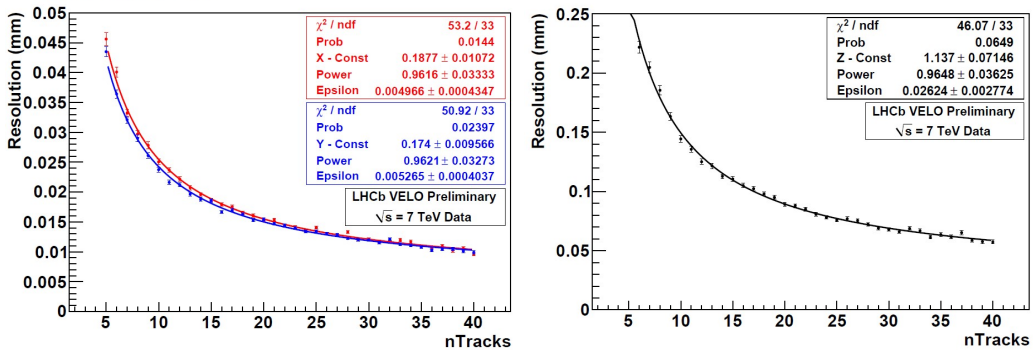


Figure 2.8 Resolution of VELO reconstructed PV in (left)  $x$  and  $y$  direction and (right)  $z$  direction.

### 2.2.1.3 Inner tracker

The tracking stations (T-stations) T1-T3 are divided into two regions: the Inner Tracker (IT) and the Outer Tracker (OT). Each station of the IT is a silicon micro-strip detector consisting of four layers, which are 120 cm in width and 40 cm in height. The same arrangement as the four layers in the TT is repeated in each station of the IT. The layout of an  $x$  layer in the second IT station is shown in Figure 2.13. The IT only covers about 2% of the LHCb detector area, but approximately 20% of the tracks pass through it.

The OT is a drift-time detector which covers the rest of the detector acceptance at

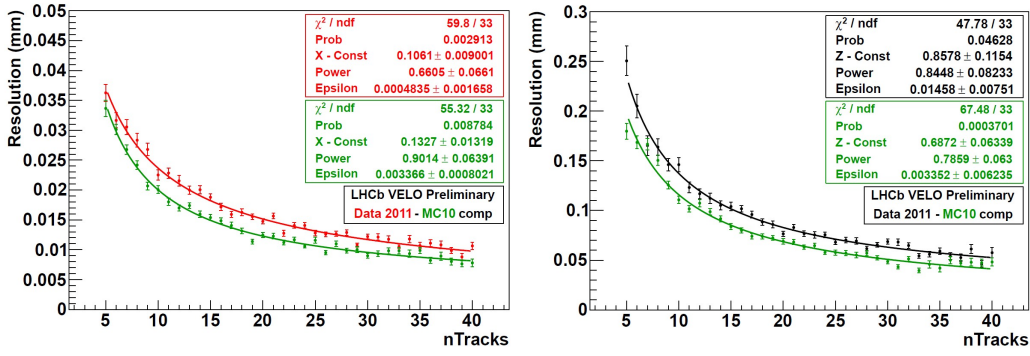


Figure 2.9 Resolution of VELO reconstructed PV in (left)  $x$  and (right)  $z$  direction for early 2011 data.

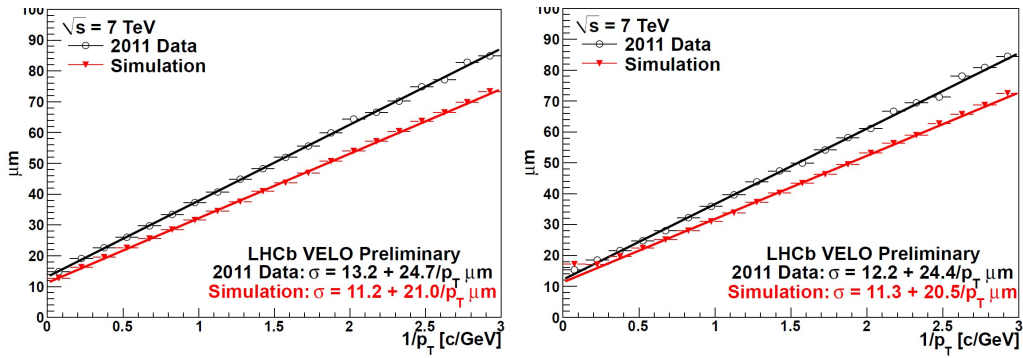


Figure 2.10 Impact parameter resolution in (left)  $x$  and (right)  $y$  direction.

the tracking stations T1-T3, as shown in Figure 2.14. Each OT station consists of four layers with the same pattern as for the TT and the IT. Each layer has several gas-tight straw-tube modules, each of which contains two staggered layers of drift-tubes filled with a gas mixture of Argon (70%) and  $\text{CO}_2$  (30%) for a fast drift time (below 50 ns) and sufficient drift-coordinate resolution ( $200 \mu\text{m}$ )<sup>[68]</sup>. The cross-section of a straw-tube module is depicted in Figure 2.15.

### 2.2.1.4 Dipole magnet

A strong magnetic field is essential for high precision momentum measurements. Originally the LHCb magnet was proposed to be super-conducting. Later it was found to require an unacceptable investigation cost and take too long time to construct, which would cause serious issue for LHCb since the polarity of the magnetic field needs to be regularly reversed to minimize systematic uncertainties. The magnetic field at LHCb is therefore provided by a dipole magnet, located between the TT and the tracking station

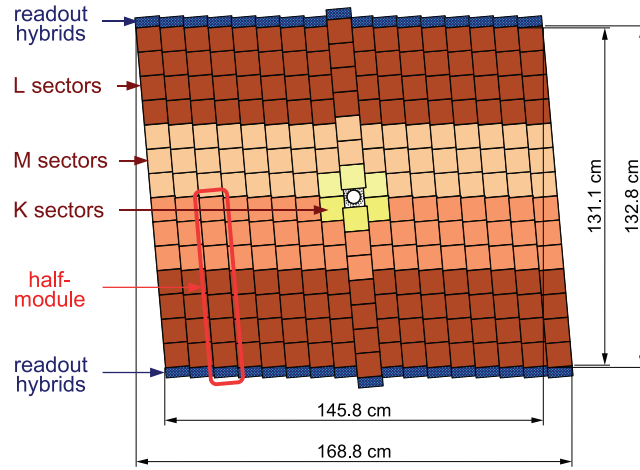


Figure 2.11 Layout of the third TT detection layer with different readout sectors labeled with different colors<sup>[68]</sup>.

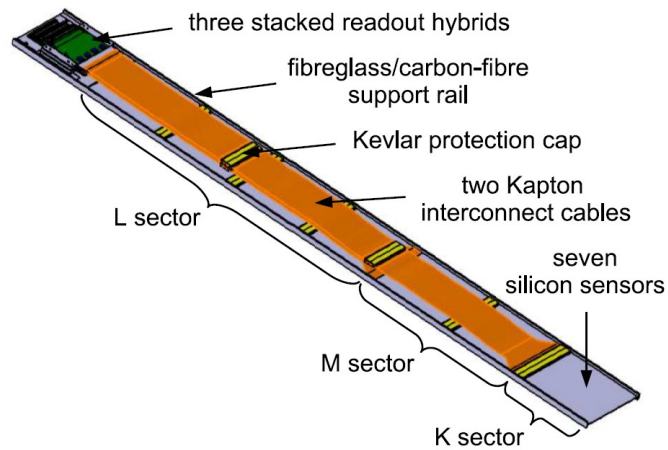


Figure 2.12 Schematic view of a TT detector module close to the beam pipe<sup>[68]</sup>.

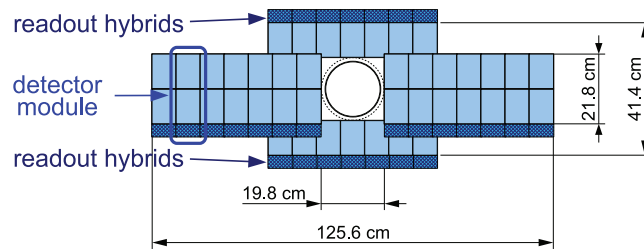


Figure 2.13 Layout of an  $x$  detection layer in the second IT station.

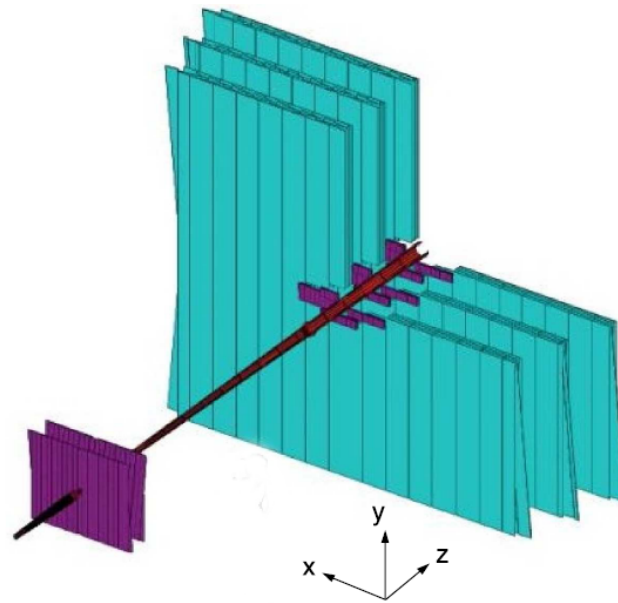


Figure 2.14 Arrangement of straw tube modules of the OT (cyan) in layers and stations. The TT and IT (purple) are also shown.

T1. The magnet covers the full LHCb acceptance of  $\pm 250$  mrad vertically and of  $\pm 300$  mrad horizontally. The designed momentum resolution for charged particles requires the magnetic field integral  $\int B dl$  determined with a relative precision of the order of  $10^{-4}$ , and the position of the B-field peak determined with a precision of a few millimeters. The magnetic field was mapped with an array of Hall probes. The measured vertical field on  $z$  axis is shown in Figure 2.16 for both magnet polarities.

### 2.2.1.5 Track finding algorithm

Charged particles produce signal in one or more of the tracking sub-detectors. Depending on the path through which the charged particle traverses the detector, the following classes of tracks are defined<sup>[68]</sup>, as illustrated in Figure 2.17.

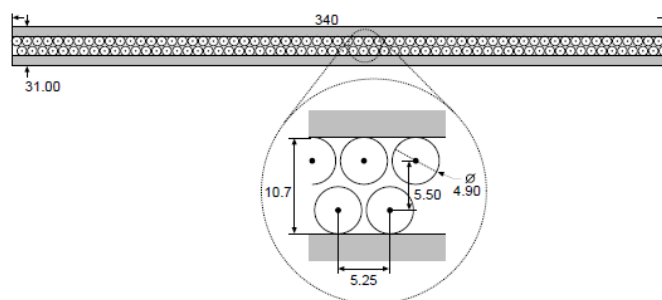


Figure 2.15 Cross-section of a straw-tube module.

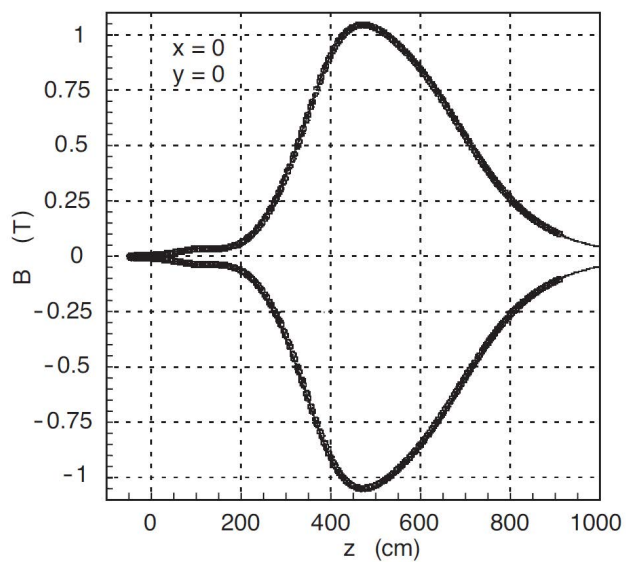


Figure 2.16 Vertical magnetic field along the  $z$  axis for both polarities.

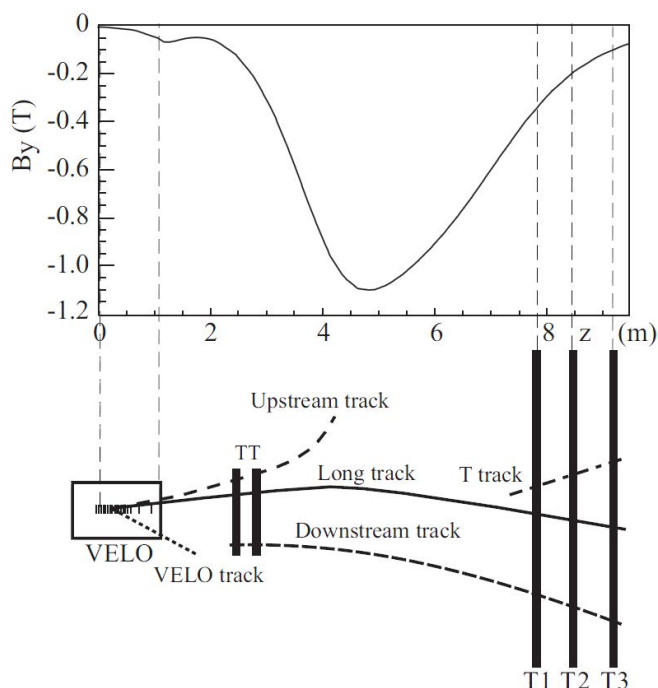


Figure 2.17 Illustration of the LHCb track types. For reference the  $y$  component of the magnetic field is also plotted<sup>[68]</sup>.

- **Long tracks** pass through the complete tracking system, from the VELO to T stations. They have the highest reconstruction quality and are the most important tracks for physics analysis.
- **Upstream tracks** traverse only the VELO and the TT. They are usually low momentum tracks swept out of the detector acceptance by the magnetic field, and have relatively poor momentum resolution. However, they can be used to understand the background in the RICH PID algorithm since they pass through RICH1 and may generate Cherenkov photons if having large enough momentum. They can also be used for flavor tagging and  $b$ -hadron decay reconstruction.
- **Downstream tracks** pass through the TT and the T stations. Typically they are daughters of long-lived particles, such as  $K_s^0$  or  $\Lambda$ .
- **VELO tracks** leave hits solely in the VELO, and have large angles with respect to the beam axis or move backward. They can be used to reconstruct primary vertices and veto non-exclusive events.
- **T tracks** only leave hits in the T stations. They help to perform RICH2 global pattern recognition.

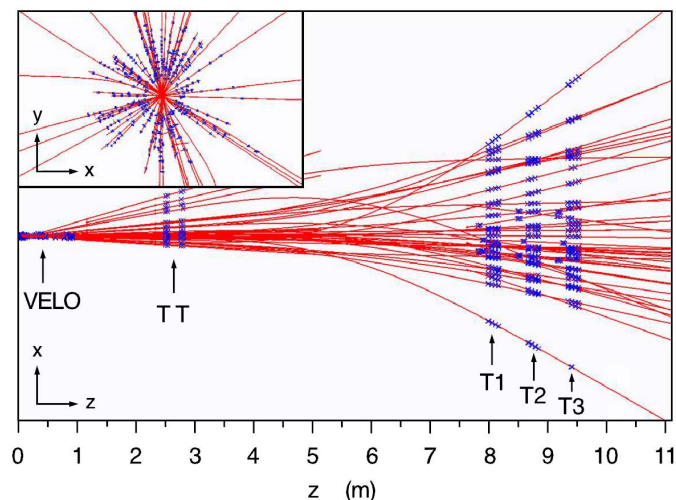


Figure 2.18 Reconstructed tracks and assigned hits in an event. The insert plot shows a projection in the  $x - y$  plane at  $z$  corresponding to VELO.

The reconstruction procedure begins with a search of track seeds in the VELO, where the magnetic field is sufficiently low and tracks can be considered as straight lines. The seeds are also searched in the T stations. The VELO seeds are extrapolated to the TT and T stations to form long tracks. This algorithm, called forward tracking, already find a large fraction of long tracks. The hits used by forward tracking are removed from further track finding to save computing time. To increase track finding efficiency for long tracks, a second algorithm called track matching is performed. T seeds are matched to the VELO seeds left from the forward tracking using the Kalman filter<sup>[74]</sup>. VELO seeds not used by the forward tracking or track matching are extrapolated to TT stations to form upstream tracks. Downstream tracks are reconstructed by adding the TT hits to the T seeds. The VELO and T seeds that have not been used for a long, upstream or downstream track are classified as a VELO or T track. An example of a reconstructed event is displayed in Figure 2.18.

#### 2.2.1.6 Track finding efficiency and momentum resolution

The track finding efficiency is measured using a data-driven method called *tag-and-probe*<sup>[75]</sup>. This method employs two-body decays with one daughter particle, the “tag” leg, reconstructed using the full detector information, while the other daughter particle, the “probe” leg, reconstructed only using information from part of sub-detectors. The information carried by the tag leg and the probe leg is sufficient to reconstruct the



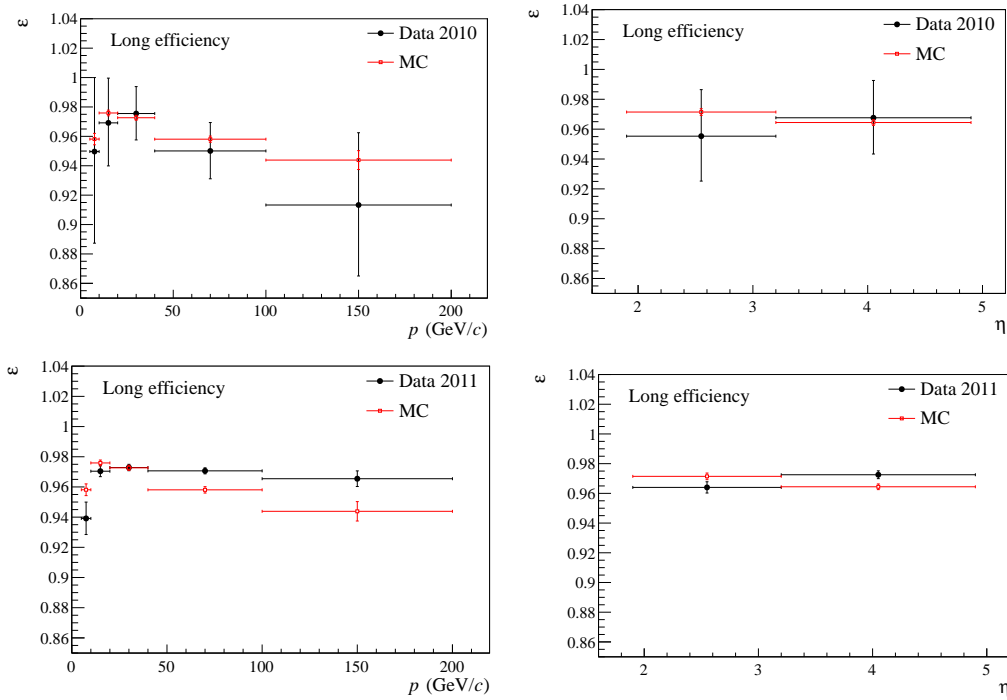


Figure 2.19 Track finding efficiency for data and MC in 2010 (top) and 2011 (bottom) as a function of momentum (left) and pseudo-rapidity (right). Efficiency from MC simulation is also shown.

mother particle, thus the number of mother particles can be obtained from fit. The probe leg is then matched to fully reconstructed tracks, and the number of the mother particle after matching can again be obtained. The track finding efficiency is defined as the ratio of the number of mother particles after and before matching. Using  $J/\psi \rightarrow \mu^+\mu^-$  decays, the track finding efficiencies for data and MC in 2010 and 2011 are measured as a function of momentum and pseudorapidity, as shown in Figure 2.19. The discrepancy between the absolute value of efficiency in data and MC is small and corrections are applied. The application of this correction will be described in Section 3.4.2.1.

Using  $J/\psi \rightarrow \mu^+\mu^-$  decays, momentum resolution for long tracks is determined to be  $\delta p/p = (0.4 - 0.8)\%$  depending on track momentum<sup>[76]</sup>, as shown in Figure 2.20.

## 2.2.2 PID system

A good PID system is essential for a precision measurement experiment. In many instances the signal mode has several partner modes with similar topologies but only one or more daughters changed. These partner modes can only be suppressed using particle identification information. The PID system in LHCb uses information from the

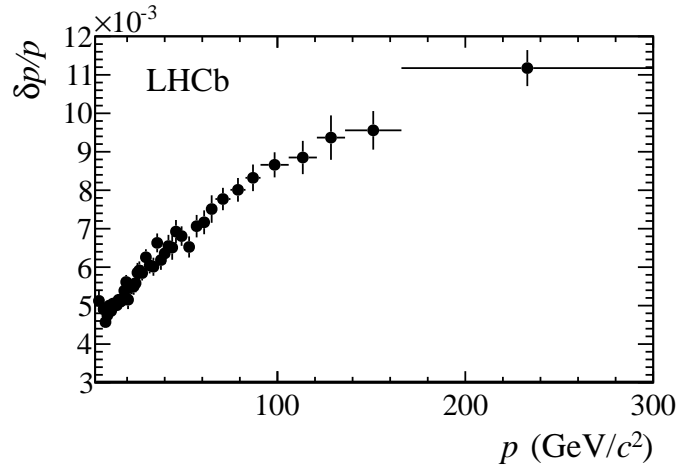


Figure 2.20 Relative momentum resolution as a function of momentum of long tracks.

following sub-detectors: two RICH detectors, the calorimeter system (SPD, PS, ECAL and HCAL) and the muon detector. In general, RICH detectors are dedicated to charged hadron identification; calorimeter system provides information to discriminate electrons and photons, photons and  $\pi^0$  and electrons and charged hadrons; the muon system is used for muon identification.

### 2.2.2.1 RICH detectors

Separation of hadrons, in particular pions and kaons, is one of the fundamental requirements for LHCb. The discrimination is accomplished by the RICH systems, which use the Cherenkov effect to identify different types of hadrons<sup>[77]</sup>. The Cherenkov effect is the phenomenon that a charged particle emits photons in the medium when it travels at a speed greater than the speed of light in that medium. The angle between the emitted photon and the momentum of the particle is

$$\cos \theta = \frac{c}{nv}, \quad (2-1)$$

where  $c$  is the speed of light in vacuum,  $n$  is the refractive index of the medium and  $v$  is the speed of the particle. If the emission angle  $\theta$  and the refractive index  $n$  is known, the velocity of the particle can be determined. Together with the momentum measurement obtained from the tracking system, the identity of the track can be determined. The Cherenkov angle dependence on track momenta of different particles in three LHCb radiators is presented in Figure 2.21.

To perform hadron identification over a momentum range of 2 GeV/c to beyond 100

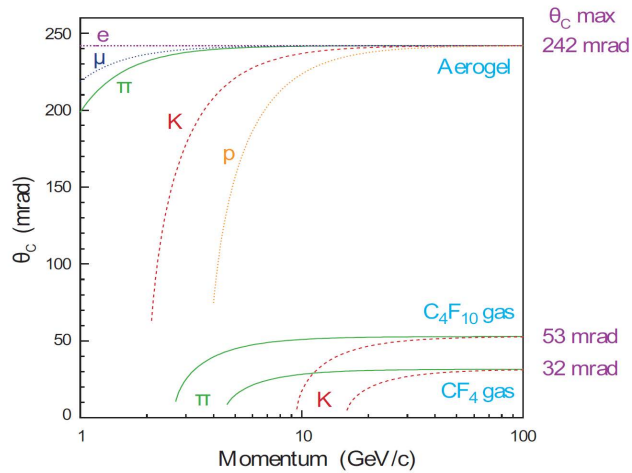


Figure 2.21 Cherenkov angle versus track momentum for different particles in three different radiators. These radiators are used in the LHCb RICH detectors<sup>[68]</sup>.

GeV/ $c$ , two RICH detectors are installed. The RICH1 detector, covering the whole LHCb acceptance, is located upstream of the dipole magnet to identify low momentum particles that may be swept out the acceptance by the magnet. It identifies charged particles with momenta of  $\sim 1$  GeV/ $c$  to 60 GeV/ $c$  by using aerogel ( $n = 1.03$ , suitable for tracks with momentum of a few GeV/ $c$ ) and C<sub>4</sub>F<sub>10</sub> gas ( $n = 1.0014$ ) as the radiators. To minimize the material budget in front of ECAL and to install hybrid photon detectors (HPD) in radiation tolerance area, spherical and flat mirrors are introduced to guide Cherenkov photons out of LHCb acceptance and focus them onto HPDs. A schematic view of the RICH1 is shown in Figure 2.22. The RICH2 detector, downstream of the T stations, has a reduced acceptance of 120 mrad horizontally and 100 mrad vertically. It uses CF<sub>4</sub> ( $n = 1.0005$ ) as the radiator and covers the momentum range of  $\sim 15$  GeV/ $c$  to beyond 100 GeV/ $c$ . Similar to RICH1, system of mirrors guides Cherenkov photons to the array of HPDs. Figure 2.23 shows a schematic view of the RICH2.

The particle identities are determined by a global likelihood algorithm, which considers information from all tracks, all radiators, and all HPD photons of a given event simultaneously<sup>[78–80]</sup>. The algorithm calculates a likelihood for each particle hypothesis by comparing the predicted and observed distribution of the photoelectrons detected, and chooses the hypothesis with the maximum likelihood. At the beginning of the algorithm, all the charged tracks are assumed to be pions, the most common type of particles produced at LHCb, and the likelihood for this initial set of hypothesis is computed. The

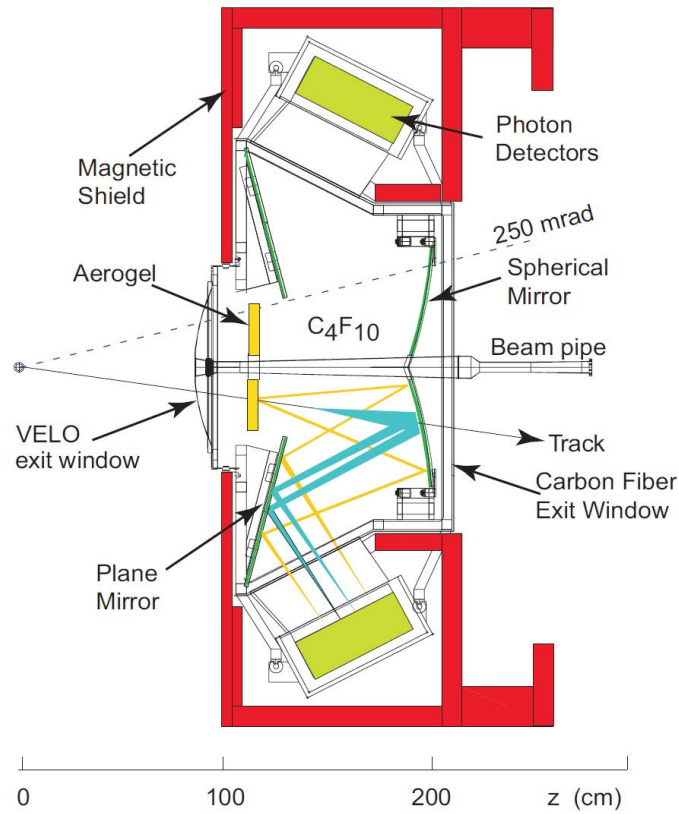


Figure 2.22 Schematic view of the RICH1 detector.

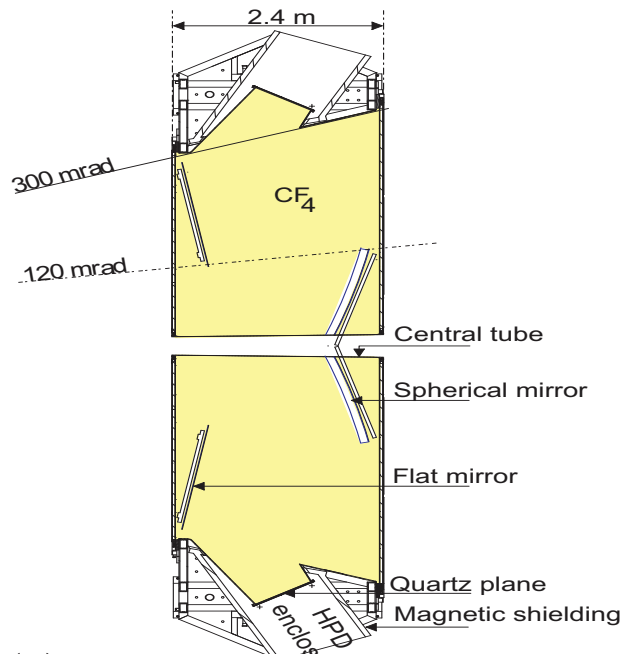


Figure 2.23 Schematic view of the RICH2 detector.

mass hypothesis of the first track is changed to electron, while hypotheses of other tracks are fixed, and a new likelihood for this set of hypotheses is calculated. The first track is also assumed to be muon, kaon and proton and the likelihoods for these set of hypotheses are evaluated. The identity of the first track is determined from the mass hypothesis which gives the largest increase in the likelihood. This process is repeated for every track; the iteration stops when all tracks have been set to their optimal hypotheses, and no further improvement in the event likelihood can be found<sup>[81]</sup>. The discrimination variable used for particle identification is the delta-log-likelihood, or DLL, which is defined as the log-likelihood difference between two different particle hypotheses of a track while keeping the other tracks unchanged. As an example, the variable to discriminate the kaon and pion hypotheses for track  $i$  is

$$\text{DLL}_{K\pi} = \ln \left( \frac{L(t_1, t_2, \dots, t_i = K, \dots, t_n)}{L(t_1, t_2, \dots, t_i = \pi, \dots, t_n)} \right), \quad (2-2)$$

where  $t_i$  is the  $i$ -th track of the event, and  $L(t_1, t_2, \dots, t_i, \dots, t_n)$  is the likelihood function.

To investigate the performance of the RICH system, large samples of genuine  $K$ ,  $\pi$ , and  $p$  are needed<sup>[81]</sup>. Such control samples should be selected without any PID-related selection applied, otherwise the efficiency will be biased. The following decays,  $\Lambda^0 \rightarrow p\pi^-$ , and  $D^* \rightarrow (D^0 \rightarrow K^-\pi^+)\pi^+$  are selected only using kinematic information. The residual background is subtracted using the *sPlot* technique<sup>[82]</sup>, where the invariant mass of  $\Lambda^0$  and the mass difference between  $D^*$  and  $D^0$  are chosen as discriminating variables. chosen as the discriminating variable. Requiring kaons to be more consistent with the kaon hypothesis than the pion hypothesis, *i.e.*  $\text{DLL}(K - \pi) > 0$ , the average kaon efficiency over the momentum range of 2 – 100 GeV/ $c$  is found to be  $\sim 95\%$ , while the misidentification rate of pion is 10%. With tighter cuts the misidentification rate of pions could be reduced to  $\sim 3\%$  while keeping kaon identification efficiency at  $\sim 85\%$ , as shown in Figure 2.24. The discrimination between protons and pions are demonstrated in Figure 2.25. For  $\text{DLL}(p - \pi) > 0$  and  $\text{DLL}(p - \pi) > 5$ , the average selection efficiency for protons is about 95% and 85%, while on average about 5% and 3% of pions are misidentified, respectively. Figure 2.26 shows the separation for protons and kaons. The performance of  $p$ - $K$  separation is similar as for  $p$ - $\pi$  at high momentum, but at low momentum the discrimination is much worse.

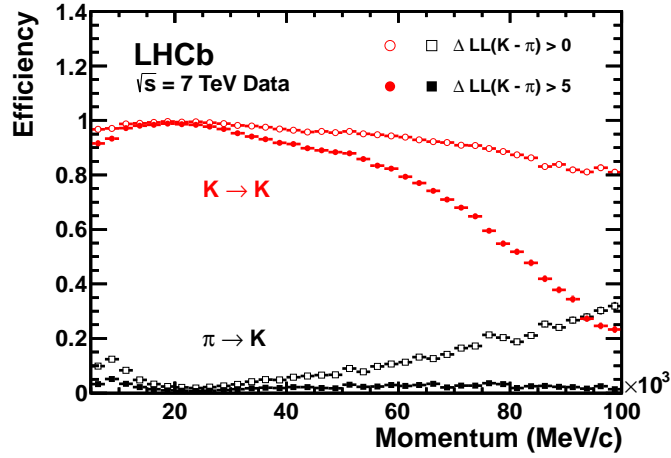


Figure 2.24 Kaon identification efficiency and pion misidentification rate as a function of track momentum. The open and filled marker distributions indicate two different  $\Delta \log \mathcal{L}(K - \pi)$  cuts<sup>[81]</sup>.

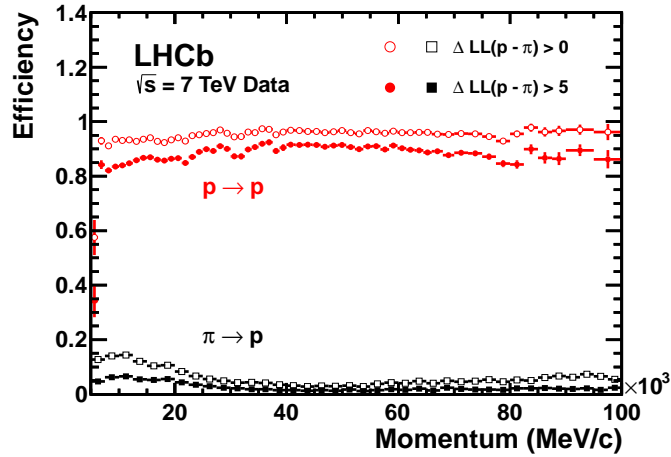


Figure 2.25 Proton identification efficiency and pion misidentification rate as a function of track momentum. The open and filled marker distributions indicate two different  $\Delta \log \mathcal{L}(p - \pi)$  cuts<sup>[81]</sup>.

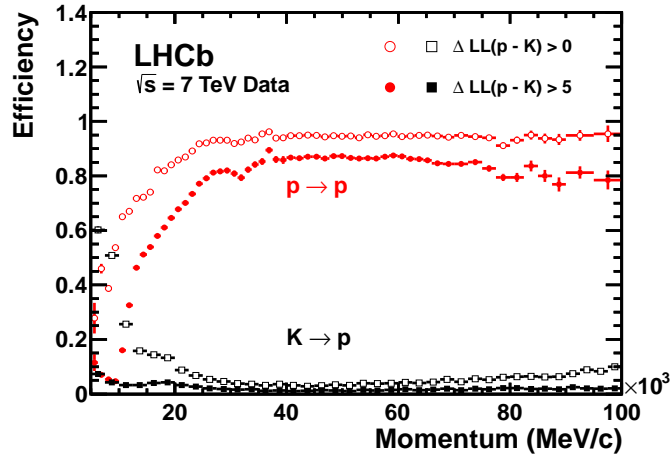


Figure 2.26 Proton identification efficiency and kaon misidentification rate as a function of track momentum. The open and filled marker distributions indicate two different  $\Delta \log \mathcal{L}(p-K)$  cuts<sup>[81]</sup>.

### 2.2.2.2 Calorimeters

The LHCb calorimeter system is used to identify both charged (electrons, charged hadrons) and neutral ( $\pi^0$ , photons and neutrons) particles. It also provides information for the hardware level trigger. It consists of four sub-detectors: the scintillator pad detector (SPD), the preshower (PS), the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL)<sup>[68]</sup>. The system is located downstream of the first muon station. ECAL (HCAL) covers the acceptance of 25 (20) mrad to 300 mrad in the horizontal plane and 20 (20) mrad to 250 mrad in the vertical plane, where the inner acceptance is determined by the acceptable radiation dose level.

The SPD/PS are composed of two almost identical planes of rectangular scintillator pads with high granularity sandwiching a 15-mm thin lead converter<sup>[68]</sup>. The SPD makes distinction between high  $E_T$  charged particles and high  $E_T$  photons and  $\pi^0$ s, since only charged particles can interact with the scintillator of the SPD, where  $E_T$  is the transverse energy. The PS rejects the background of high  $E_T$  charged pions by detecting the longitudinal segmentation of the electromagnetic shower initiated by the lead converter<sup>[83]</sup>. The scintillation photons of the SPD/PS cell are transmitted to a multianode photomultiplier tubes (MAPMT) by a single wavelength-shifting (WLS) fibre, as shown in Figure 2.27.

The ECAL, downstream of the SPD/PS, is required to distinguish electrons and hadrons and to give a modest energy resolution while having an acceptable radiation

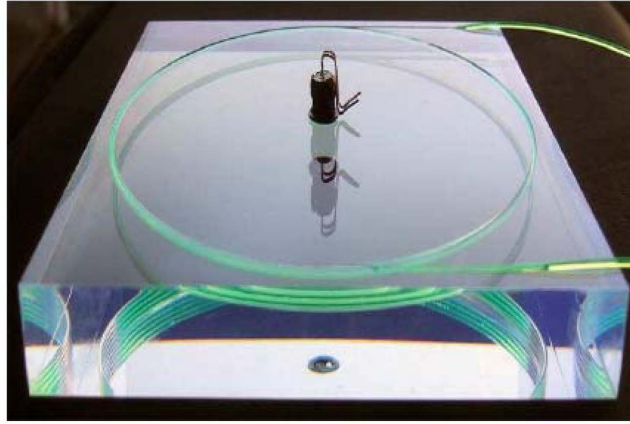


Figure 2.27 SPD/PS detector cell with the WLS fibre layout and the LED installed in the center<sup>[68]</sup>.

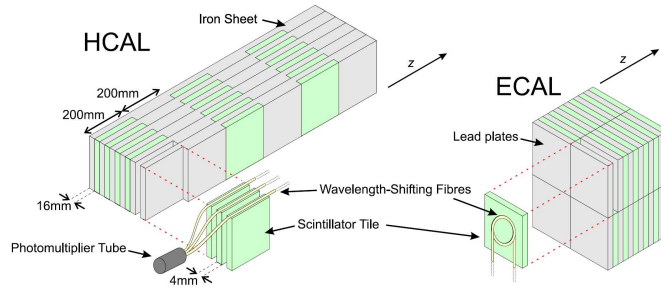


Figure 2.28 Schematic view of the LHCb HCAL (left) and ECAL (right), showing scintillator tiles, absorber plates and WLS readout fibres.

resistance. Therefore it adopts the reliable shashlik calorimeter technology, *i.e.* individual modules made of 4- mm thick scintillator tiles interleaved by 2- mm thick lead absorber plates (see Figure 2.28). The scintillation light is absorbed, re-emitted and transported by WLS fibres traversing the entire module. The energy resolution of the ECAL is

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 1\%, \quad (2-3)$$

where  $E$  is measured in GeV.

The particle density varies drastically from the inner section to the outer section. The SPD, PS and ECAL are divided to three regions with different segmentation to balance the detector occupancy, as shown in Figure 2.29. The segmentation in the SPD and PS is scaled to match that of the ECAL projectively to allow for a simpler energy reconstruction and a faster hardware-level trigger decision.



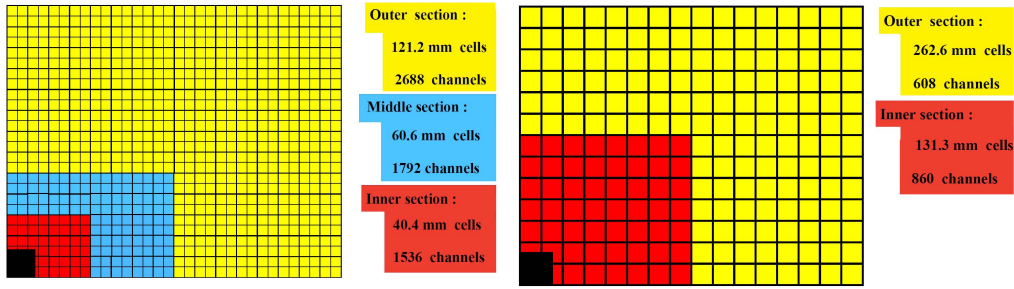


Figure 2.29 Lateral segmentation of the SPD/PS and ECAL (left) and HCAL (right) for one quarter of the front face. The cell dimensions in the left figure are for the ECAL.

The HCAL, downstream of the ECAL, is mainly used for the hardware trigger. Therefore the primary goal is a fast response time. To make hadrons deposit all the energy using limited volume, the HCAL adopts a sampling structure with iron as the absorber and scintillating tiles as the active material (see Figure 2.28). The scintillating tiles have their orientation running parallel to the beam axis. The scintillation light is transported to PMTs situated at the back side by WLS fibres, which run along the tile edges. The energy resolution of the HCAL is

$$\frac{\sigma_E}{E} = \frac{(69 \pm 5)\%}{\sqrt{E}} \oplus (9 \pm 2)\%, \quad (2-4)$$

where  $E$  is measured in GeV. The HCAL also matches pseudo-projectile geometry, as shown in Figure 2.29.

### 2.2.2.3 MUON detector

The muon system provides excellent identification for muons. Due to its lepton nature, a relatively large mass and a long lifetime, muons penetrate much further than other particles. Consequently M2-M5 stations of the muon system are installed the most downstream of the LHCb detector, while M1 station is located before the calorimeters to improve the  $p_T$  measurement for trigger and full reconstruction. Transverse dimensions of muon stations are adjusted according to their distance from the interaction point so that they cover the same angular acceptance. The particle flux decreases rapidly in the lateral direction, therefore each station is divided into four regions (R1-R4) to make channel occupancies roughly the same for each region of a given station. A side view of the muon system is shown in Figure 2.30.

The inner region (R1) of M1 station has the highest particle flux in all the regions,

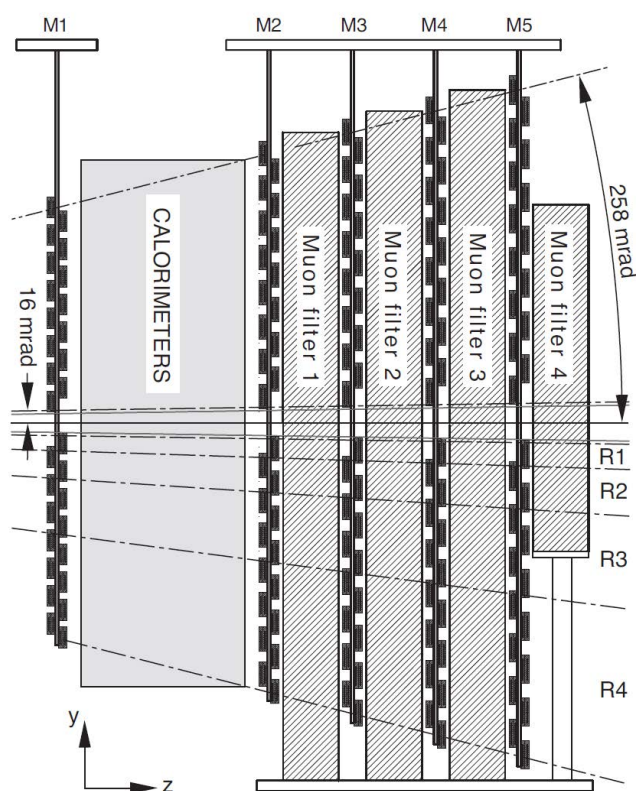


Figure 2.30 Side view of the R1–R4 regions of the muon system<sup>[68]</sup>.

therefore it consists of two superimposed triple Gas-Electron-Multiplier (triple-GEM) detectors in order to enhance the radiation resistance<sup>[84]</sup>. The rest of M1 and M2-M5 consist of multi-wire proportional chambers (MWPC)<sup>[85]</sup>. The triple-GEM detector is composed of three perforated gas electron multiplier foils which are sandwiched between anode and cathode planes, and the drift gap between foils is filled with a mixture of  $\text{Ar}/\text{CO}_2/\text{CF}_4$ <sup>[84]</sup>. A cross-section of the triple-GEM detector is shown in Figure 2.31. The decision of M1R1 is the logical OR of the two triple-GEM detectors. In M2-M5 chambers comprise four equal gas gaps with the electronics of two adjacent gaps logically OR-ed to create a double gap layer. In R2-R4 of M1 station the MWPCs have only two gas gaps, which are readout independently, to reduce the material in front of the ECAL. The gas used for MWPCs is also a mixture of  $\text{Ar}/\text{CO}_2/\text{CF}_4$  but with different fractions.

#### 2.2.2.4 Muon identification

The muon identification proceeds in three steps<sup>[86,87]</sup>:

- Select muon candidates (called IsMuon) based on the penetration length of the

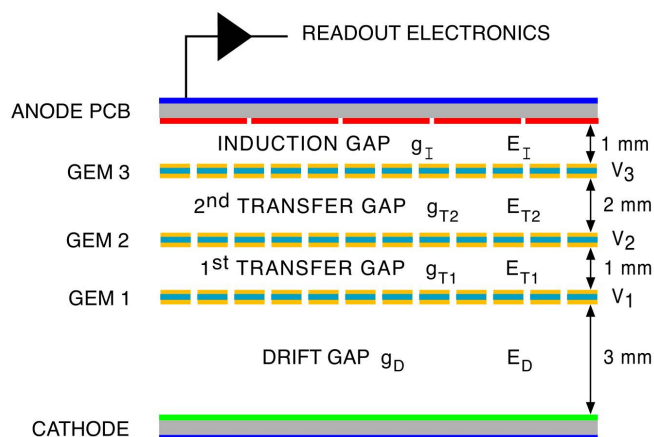


Figure 2.31 Schematic view of the cross-section of a triple-GEM detector<sup>[68]</sup>.

muons through the calorimeters and iron filters. This simple selection provides a high efficiency for muons while suppresses the hadron misidentification rate to a percent level.

- Extrapolate charged particle trajectories to different muon stations, and compute likelihoods for the muon and non-muon hypotheses based on the pattern of hits around the extrapolation. The logarithm of the likelihood ratio between the muon and non-muon hypotheses, called muDLL, is used as the discriminating variable.
- Combine information from the calorimeter and the RICH systems to calculate a combined likelihood for different particle hypotheses.

The performance of the muon system is extracted using large samples of muons, pions, kaons, and protons, which are selected using kinematical information only. The pion, kaon, and proton samples are selected using the same decays as those used to study the RICH performance. A high purity muon sample is provided by the  $J/\psi \rightarrow \mu^+\mu^-$  decays, and muons can further be enriched if the inclusive  $B \rightarrow J/\psi X$  decays are required. The identification efficiency for muons and misidentification rate for pions, kaons, and protons are shown in Figure 2.32. The performance of muon identification depends on the track momentum. With the IsMuon and muDLL cuts, the average muon efficiency could be kept at the level of 93%, while the misidentification rate is suppressed below 0.6%<sup>[87]</sup>.

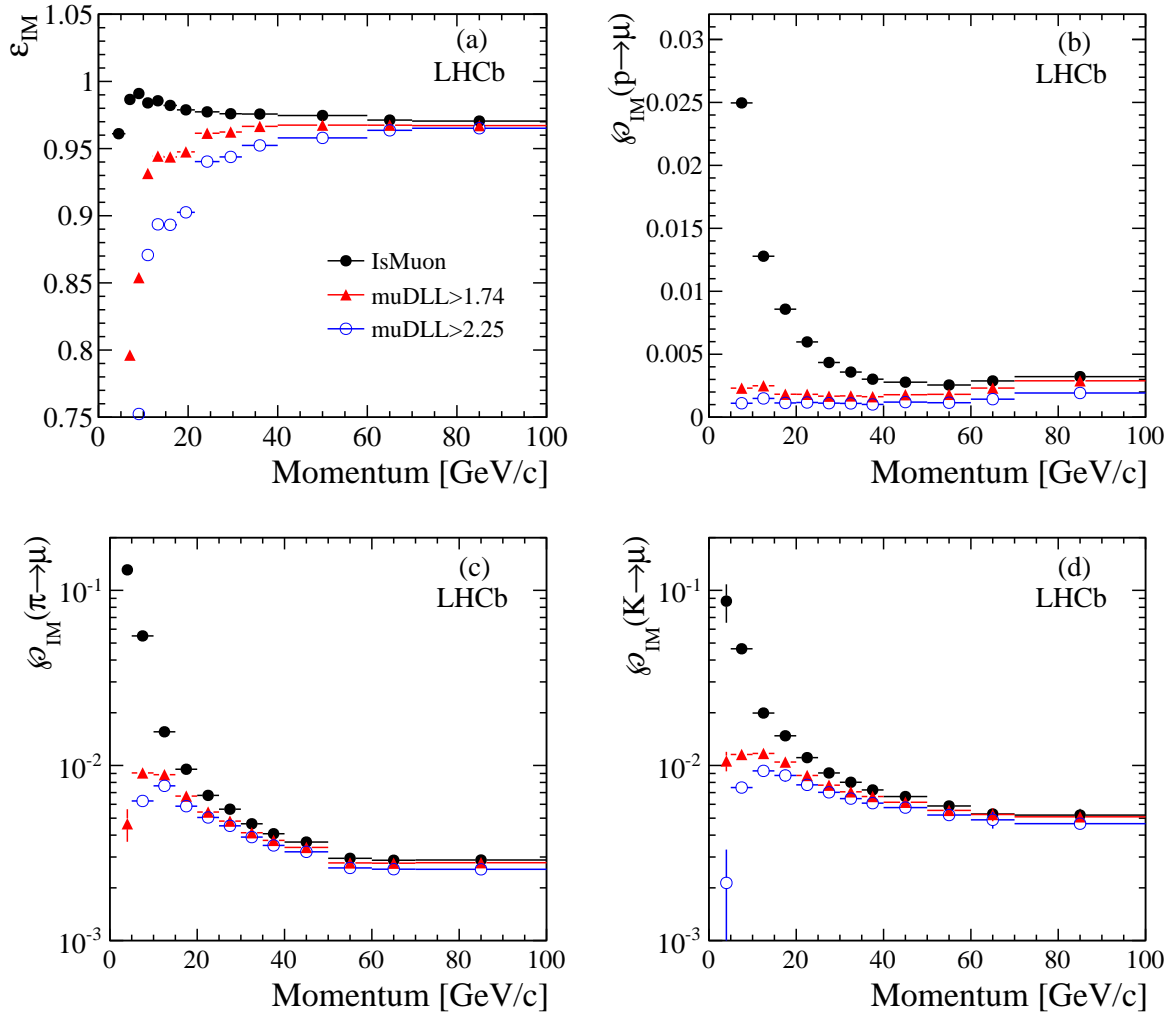


Figure 2.32 Identification efficiency for muons (a) and misidentification rate for protons (b), pions (c) and kaons (d) as functions of track momentum. The black solid circles show the efficiency for the IsMuon requirement alone, red triangles and blue open circles show the efficiencies with the additional cuts  $\text{muDLL} \geq 1.74$  and  $\text{muDLL} \geq 2.25$ , respectively<sup>[87]</sup>.

### 2.3 LHCb trigger

Most collision events at the LHC do not involve interesting decays. An efficient trigger system is essential to reject boring events and only keep the events that probably contain heavy flavor hadrons. At LHCb this filtration task is accomplished by a two-level system. The Level-0 trigger (L0) is implemented at the hardware level, and the High Level Trigger (HLT) is a software trigger running on online computing farm called the Event Filter Farm (EFF)<sup>[68,88,89]</sup>. Both L0 and HLT are configured via a unique hexadecimal key, the Trigger Configuration Key (TCK), which defines the sequence of lines included in the trigger and selection criteria of these lines.

### 2.3.1 Level-0 trigger

The L0 trigger reduces the inelastic event rate from 16 MHz to about 1 MHz, the maximum rate at which the full detector response can be readout<sup>[90]</sup>. L0 is divided into three independent parts, L0-Calorimeter trigger, L0-Muon trigger and L0-PileUp trigger. The trigger decisions from these triggers are combined in L0 Decision Unit (L0DU), and further transferred to the Readout Supervisor board (RS). The RS decides to accept or throttle a L0 trigger based on the information of L0DU and the status of other sub-detectors and the EFF.

The L0-Calorimeter trigger selects events based on the information from SPD, PS, ECAL, and HCAL<sup>[89]</sup>. The calorimeter segment was optimized to follow the detector occupancy. The L0-Calorimeter system calculates the transverse energy ( $E_T$ ) deposited in the clusters, formed by the arrays of  $2 \times 2$  cells of the same size. The transverse energy is defined as

$$E_T = \sum_{i=1}^4 E_i \sin \theta_i \quad (2-5)$$

where  $E_i$  is the energy deposited in cell  $i$  and  $\theta_i$  is the angle between the beam axis and the hypothetical line connecting the interaction point and the center of the cell. The L0-Calorimeter system defines three types of candidates based on the hits and  $E_T$  information:

- the hadron candidate (L0Hadron) is the HCAL cluster with the highest  $E_T$ ;
- the photon candidate (L0Photon) is the ECAL cluster with the highest  $E_T$ , which has 1 or 2 hits in the corresponding PS cell and no hits in the corresponding SPD cell;
- the electron candidate (L0Electron) has the same requirements as for a photon candidate except that it requires at least one SPD cell hit in front of the PS cells.

If at least one of the candidates have  $E_T$  above the corresponding threshold, the event will trigger L0. The total number of hits in the SPD is also recorded to veto events which contain too many tracks to be processed in a reasonable time in the HLT.

The muon system is divided into four quadrants in the  $x-y$  plane, and each quadrant is connected to an L0 muon processor. The processors pick up two muon track candidates with the largest and the second largest  $p_T$  in their corresponding quadrant. The event will trigger L0 if the largest  $p_T$  of these eight candidates is larger than a threshold (L0Muon), or if the product  $p_T^{\text{largest}} \times p_T^{\text{2nd largest}}$  is greater than the given threshold (L0DiMuon).

The L0-PileUp system consists of two  $R$ -sensor planes placed upstream of the VELO. It was designed to reject events with multiple primary vertices, but as LHCb is running at a higher  $\mu$  than foreseen in the design phase, events with multiple proton-proton interactions are not rejected. Now the system provides the trigger for beam-gas events, which are used in the luminosity determination.

### 2.3.2 High level trigger

The High Level Trigger (HLT) is a series of algorithms written in C++, which run on the EFF that contains 26,110 processors to process events accepted by L0. It is split into two stages, HLT1 and HLT2, with each stage consisting of several selection algorithms called trigger lines. Each trigger line is dedicated to a certain class of events of interest. Trigger lines can be configured through python scripts, which define reconstruction algorithms and selection criteria.

The HLT1 begins with a fast VELO track pattern recognition, then all primary vertices are constructed using these VELO tracks. HLT1 line not requiring muons select VELO tracks with good track qualities whose impact parameters to the closest primary vertex are larger than a threshold. A fast muon reconstruction algorithm is performed on VELO tracks by extrapolating them to M3 and searching for hits in M2, M4, and M5. The selected VELO tracks are matched to OT and IT stations, and their momenta are determined using a Kalman filter based track fitter with a simplified material geometry description. HLT1 lines requiring muons select muon candidates based on the impact parameter and momentum.

HLT2 is only executed for the events which are accepted by at least one of the HLT1 lines. The output rate of the HLT1 is about 50 kHz, which is low enough so the forward tracking of all VELO seeds can be performed in the HLT2, although the track matching between VELO seeds and T seeds is not performed due to time constraint. The HLT2 consists of several inclusive beauty hadron and many exclusive charm and beauty hadron trigger lines. Inclusive beauty hadron trigger lines select  $b$ -hadron signals using a multivariate method based on two signatures of  $b$ -decays: a displaced and high transverse momentum track, and a displaced vertex containing this track and other one to three tracks<sup>[91]</sup>. Exclusive trigger lines select prompt charm decays, and other interesting decays which cannot be triggered by the inclusive lines. Due to the limited bandwidth, trigger lines with plethoric production rate are pre-scaled: events which meet all the

selection criteria of these trigger lines are discarded randomly at a given rate.

### 2.3.2.1 HLT deferral

HLT deferral is a novel feature of the LHCb trigger system which postpone the processing of some events accepted by L0 until the LHC is preparing new beams and no collision happens<sup>[92]</sup>. The LHC beams intensities decrease with time. When the intensity goes below some threshold, the beams are dumped and a new fill is needed. It could take a few hours to start another fill. Before the implementation of the HLT deferral system, the EFF was largely idle during this period and the computing power was wasted. With the HLT deferral system, a fraction of the events which are accepted by L0 are not processed immediately but saved in the EFF storage. The EFF process these events when the LHC is filling new beams. Due to deferral HLT, LHCb recorded approximately 20% more data in 2012.

### 2.3.2.2 TCK

Both L0 and HLT are configured via a unique hexadecimal key, named Trigger Configuration Key (TCK), which defines the sequence of lines included in the trigger and selection criteria of these lines.

### 2.3.3 TISTOS method

Events after off-line selection are required to pass specific trigger lines so that data samples can be better understood. The trigger efficiency is therefore defined relative to off-line selections. However, the naive data-based trigger efficiency computation is biased since only events firing some trigger lines are recorded. The TISTOS method is used to extract the trigger efficiency from data correctly. Two categories of the data sample, TOS and TIS, are defined:

- Trigger On Signal(TOS): the signal candidate or its component parts fire the trigger line;
- Trigger Independent of Signal(TIS): the event can be triggered without the presence of the signal candidate.

By definition, the efficiency of TOS events can be determined as

$$\epsilon^{\text{TOS}} = \frac{N_{\text{TOS} \cap \text{Sel}}}{N_{\text{Sel}}}, \quad (2-6)$$

where  $N_{\text{Sel}}$  is the number of events which pass the selection criteria, regardless whether they can be triggered or not, and  $N_{\text{TOS} \cap \text{Sel}}$  is the number of events which pass the selection and TOS on the trigger line.  $N_{\text{Sel}}$  is not an observable since all the events must fire some trigger line and samples without trigger requirements applied do not exist. Therefore, Eq.2-6 cannot be used to compute the trigger efficiency directly. However, if the correlation between TIS and selections is small enough, TIS-ed sample is a good approximation of the sample that are only filtered by the selection criteria. In this case, the TOS efficiency can be expressed as

$$\varepsilon^{\text{TOS}} = \frac{N_{\text{TOS} \cap \text{Sel}}}{N_{\text{Sel}}} = \frac{N_{\text{TIS} \cap \text{TOS} \cap \text{Sel}}}{N_{\text{TIS} \cap \text{Sel}}}. \quad (2-7)$$

where  $N$  is the number of events, and subscripts indicate requirements the events should pass.

### 2.3.4 Performance of the trigger system

The performance of the trigger system is extracted using the TISTOS method. In this section only performance of the trigger lines that are relevant to the  $\Xi_{cc}^+$  analysis, *i.e.* L0Hadron and Hlt1TrackAllL0, is presented.

L0Hadron selects decays with hadrons in the final state. The performance of L0Hadron is shown in Figure 2.33 for  $B^0 \rightarrow D^+\pi^-$ ,  $B^+ \rightarrow D^0\pi^+$ ,  $D^0 \rightarrow K^-\pi^+$  and  $D^+ \rightarrow K^-\pi^+\pi^+$  as a function of  $p_T$  of signal mesons. The L0Hadron TOS efficiency is very sensitive to the  $p_T$  of signal mesons due to the presence of  $E_T$  requirements in L0.

The Hlt1TrackAllL0 is an HLT1 line executed for all events which pass L0. It selects hadron decays with a non-zero lifetime, *e.g.*  $\Lambda_c^+ \rightarrow pK^-\pi^+$  decay. The performance of the Hlt1TrackAllL0 is shown in Figure 2.34 for  $B^0 \rightarrow D^+\pi^-$ ,  $B^+ \rightarrow D^0\pi^+$ ,  $D^0 \rightarrow K^-\pi^+$  and  $D^+ \rightarrow K^-\pi^+\pi^+$  as a function of  $p_T$  of signal mesons. The Hlt1TrackAllL0 provides a very efficient trigger for heavy flavor decays with a significant distance from their primary vertices.

## 2.4 LHCb data processing

The data processing at LHCb proceeds in several steps. First, events selected by the LHCb trigger system are recorded as streamed files, called RAW files, by the LHCb online system. Second, reconstruction jobs are run for the RAW files since they do not explicitly contain information of physical objects. The output containing



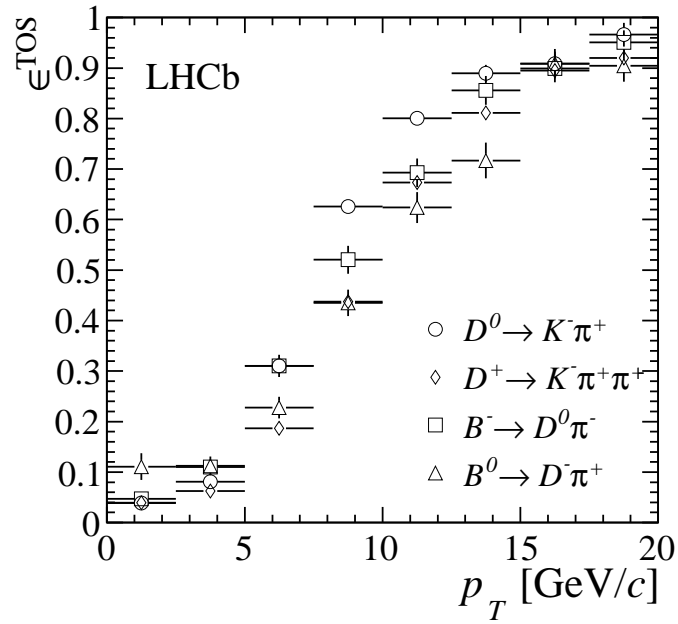


Figure 2.33 TOS efficiency of the LOHadron as a function of  $p_T$  of signal mesons<sup>[89]</sup>.

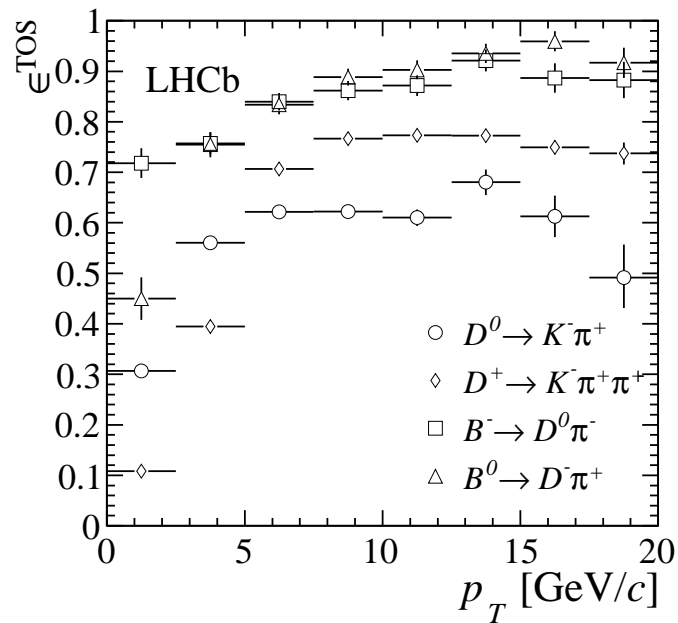


Figure 2.34 TOS efficiency of the Hlt1TrackAllL0 as a function of  $p_T$  of signal mesons<sup>[89]</sup>.

all reconstructed objects of the events are produced in a format called selected data summary tape. Third, to perform physics analysis, the sDST files are centrally processed by a procedure called stripping, which consists of many sets of selection algorithms called stripping lines. Most stripping lines process all events, but some lines may be prescaled, *i.e.* only a fraction of the total events, selected randomly, will be processed by that line; other events will be discarded no matter they satisfy the selection criteria or not. The output of the stripping procedure can be in two formats: data summary tape or micro data summary tape, with the former containing all the information of the event while the latter only containing information relevant for a particular analysis. The same RAW files may be processed several times using different reconstruction algorithms and different stripping selection criteria. To distinguish these results, the output files are tagged with the reconstruction version and stripping version, *e.g.* the 2011 data can be tagged with “Reco12-Stripping17” or “Reco14-Stripping20”. Finally, LHCb physicists read information of the selected events from the DST or MDST files and perform the offline analysis.

## 2.5 LHCb software framework

The LHCb software framework consists of several projects built on Gaudi, which is an object-oriented C++ framework that provides common interfaces and services for building HEP experiment data processing frameworks<sup>[93]</sup>. Many projects are developed for different purposes, including GAUSS (event generation and detector simulation), BOOLE (detector digitization), MOORE (trigger), BRUNEL (track reconstruction) and DAVINCI (data analysis).

### 2.5.1 Simulation

Monte Carlo (MC) events are essential for physics analysis since the detector is an extremely complex system and it is very difficult to foresee the behavior of the signal without simulation. MC event generation at LHCb is handled by the GAUSS<sup>[94-97]</sup> and BOOLE<sup>[98]</sup> projects. The whole process is split into three steps:

- Event generation. The GAUSS project simulates proton-proton collisions, invokes PYTHIA<sup>[99]</sup> to generate desired signals, and then forces them to decay through the desired channel using EVTGEN<sup>[100]</sup>. It is also possible to use other toolkits

for generation. This could be useful if the signal has a very small production cross-section in PYTHIA, *e.g.* the  $\Xi_{cc}^+$  baryon, which is generated using the GENXICC generator<sup>[101,102]</sup>.

- Detector simulation. GAUSS implements the interaction of generated particles with the detector by calling the GEANT4 toolkit<sup>[103,104]</sup>.
- Detector digitization. The output of GAUSS is transferred to BOOLE, which emulates the detector response, including the readout electronics and the hardware trigger as well as the noise, cross-talk and spill-over from previous bunch crossings.

The output of BOOLE has the same format as for real data, and after this step MC samples and real data are processed in the same way.

## 2.5.2 Trigger

The Moore project is the application for high level trigger<sup>[105]</sup>. It can run online to filter data from the LHCb data acquisition system, or offline to process real data or MC samples digitalized by BOOLE. Moore can be configured in two modes: the rejection mode, in which only events accepted by the HLT are saved, and the flagging mode, in which all the events are saved and pass/fail information for each trigger line are added for each event.

## 2.5.3 Track reconstruction

The BRUNEL project performs event reconstruction at LHCb<sup>[106]</sup>. It reconstructs charged tracks from detector responses using pattern recognition algorithms, and passes these tracks to PID routines to perform particle identification. The tracking algorithm and its performance can be found in Section 2.2.1.6.

## 2.5.4 Data analysis

DAVINCI is the physics analysis framework for LHCb<sup>[107]</sup>. It provides interfaces to manipulate reconstructed tracks and vertices. DAVINCI allows to plug in user-defined algorithm to perform specific tasks. Many dedicated packages, *e.g.* flavor tagging, lifetime fit, MC match, are developed based on DAVINCI. The stripping procedure is also accomplished by DAVINCI.

## Chapter 3 Search for the $\Xi_{cc}^+$ baryon using 2011 data

This chapter presents the search for the  $\Xi_{cc}^+$  baryon using  $pp$  collision data collected by LHCb in 2011. Section 3.1 introduces the overall analysis strategy. A description of the data sample is given in Section 3.2. Selection criteria are detailed in Section 3.3. Numerical results of the efficiency ratio are given in Section 3.4. Section 3.5 demonstrates two methods to determine the signal yield. Systematic uncertainties are evaluated in Section 3.7. Section 3.8 discusses the efficiency variation with the mass and lifetime of the  $\Xi_{cc}^+$  baryon, and Section 3.9 explains the upper limit setting procedure. Finally, Section 3.10 presents the results.

### 3.1 Analysis strategy

The  $\Xi_{cc}^+$  baryon is searched through its decay  $\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$  with the  $\Lambda_c^+$  baryon reconstructed via  $\Lambda_c^+ \rightarrow p K^- \pi^+$ . To facilitate the comparison between our measurements and theoretical predictions or results of other experiments, the cross-section, or rather, the ratio of cross-sections of the  $\Xi_{cc}^+$  baryon and an appropriate control mode should be given. However, the kinematics and topology of the signal mode are not so common: the moderate lifetime and small  $Q$ -value make it very different from  $b$ -hadron decays, and the non-zero lifetime separates it from strong decays of  $\Xi_c$  resonances. Instead, the signal mode is normalized relative to inclusive  $\Lambda_c^+$  production, again reconstructed through  $\Lambda_c^+ \rightarrow p K^- \pi^+$ . Although this mode do not share common topology with the signal mode, it has several benefits: the lifetimes of  $\Lambda_c^+$  and  $\Xi_{cc}^+$  baryons and the  $Q$ -value of the two decays are very close; many systematic uncertainties (including the uncertainty of the  $\Lambda_c^+$  branching fraction) cancel in the ratio; it also allows a direct comparison with previous experiments that quoted a similar ratio. Note that uncertainty associated to additional kaon and pion is not canceled.

The production cross-section of the  $\Xi_{cc}^+$  baryon is formulated as

$$\sigma_{\Xi_{cc}^+} = \frac{N_{\Xi_{cc}^+}}{\mathcal{L} \times \varepsilon_{\Xi_{cc}^+} \times \mathcal{B}(\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+) \times \mathcal{B}(\Lambda_c^+ \rightarrow p K^- \pi^+)}, \quad (3-1)$$

where  $\sigma_{\Xi_{cc}^+}$  is the production cross-section of the  $\Xi_{cc}^+$  baryon,  $N_{\Xi_{cc}^+}$  is the yield of the  $\Xi_{cc}^+$  baryon after all the selections,  $\mathcal{L}$  is the integrated luminosity,  $\varepsilon_{\Xi_{cc}^+}$  represents the selection

efficiency for this decay, and  $\mathcal{B}(\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+) \times \mathcal{B}(\Lambda_c^+ \rightarrow p K^- \pi^+)$  indicates the adjoint branching fraction (BF).

Similarly, the inclusive production cross-section of the  $\Lambda_c^+$  baryon is defined as

$$\sigma_{\Lambda_c^+} = \frac{N_{\Lambda_c^+}}{\mathcal{L} \times \varepsilon_{\Lambda_c^+} \times \mathcal{B}(\Lambda_c^+ \rightarrow p K^- \pi^+)}. \quad (3-2)$$

Therefore, the relative production ratio,  $R$ , is defined as

$$R \equiv \frac{\sigma(\Xi_{cc}^+) \mathcal{B}(\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+)}{\sigma(\Lambda_c^+)} = \frac{N_{\text{sig}} \varepsilon_{\text{con}}}{N_{\text{con}} \varepsilon_{\text{sig}}} \quad (3-3)$$

where ‘‘sig’’ and ‘‘con’’ refer to the signal ( $\Xi_{cc}^+$ ) and control ( $\Lambda_c^+$ ) modes,  $N$  is the event yield, and  $\varepsilon$  is the efficiency. Our results will be presented using  $R$ , *i.e.* we will measure  $R$  instead of the cross-section.

To avoid preconceptions, which may bias the result in the direction of previous measurements or theoretical exceptions, the analysis is performed in a blind approach, *i.e.* the signal region is blinded during the design of the selection criteria and the fit procedure, and can only be unblinded when all the procedures are frozen<sup>[108]</sup>. The mass difference,  $\delta m$ , is used as the observable, where  $\delta m$  is defined as

$$\delta m \equiv m(p K^- \pi^+ K^- \pi^+) - m(p K^- \pi^+) - m(K^-) - m(\pi^+). \quad (3-4)$$

It has a better resolution than  $m(p K^- \pi^+ K^- \pi^+)$  since the invariant mass of the  $\Lambda_c^+$  and  $\Xi_{cc}^+$  candidates are positively correlated and the subtraction cancels part of the error. The blind region is defined as  $380 < \delta m < 880$  MeV/ $c^2$  to cover most theoretical predictions, and approximately correspond to a  $\Xi_{cc}^+$  mass range of  $3300 < m(\Xi_{cc}^+) < 3800$  MeV/ $c^2$ .

## 3.2 Data set

The search for the  $\Xi_{cc}^+$  baryon is performed using a 7 TeV  $pp$  collision data sample corresponding an integrated luminosity of  $0.65 \text{ fb}^{-1}$ . The data was collected in the late period of the 2011 data taking since the dedicated trigger line for  $\Lambda_c^+$  baryons was not presented in the HLT2 until July, 2011. Several TCKs were used during the data-taking period, but selection cuts of the trigger lines which are used to trigger the signal and control modes did not change in that period. The processing pass for data is *Reco12-Stripping17* and ntuples are made with *DAVINCI v29r1*.

The MC sample of the control mode, with a size of 2.02 M, is generated in the conventional way with *PYTHIA*. For the signal mode, however, the default *PYTHIA*

configuration is very inefficient due to the presence of two  $c$  quarks in the  $\Xi_{cc}^+$  baryon. The generator GENXICC is therefore employed to generate  $\Xi_{cc}^+$  baryons at LHCb. The mass and lifetime of the  $\Xi_{cc}^+$  baryon are both unknown a priori, and these parameters are chosen as  $m(\Xi_{cc}^+) = 3500 \text{ MeV}/c^2$  and  $\tau(\Xi_{cc}^+) = 333 \text{ fs}$  for the generator based on theoretical predictions (see Section 1.2.1). Note that the efficiency varies with the mass and the lifetime, and this needs to be taken into account when we calculate upper limits for different mass and lifetime hypotheses. The MC sample for the signal mode has a size of 2.26 M. Both modes have their MC samples produced using the following software versions: GAUSS v41r3, BOOLE v23r1, MOORE v12r8g1, BRUNEL v41r1p1 and DAVINCI v29r1p1. The samples have been flagged with TCK 0x40760037 and Reco12-Stripping17, the same criteria as data, except for the prescale factors that were set to 1.

The  $\Xi_{cc}^+$  kinematic distribution of GENXICC is assumed, and in the absence of a signal an upper limit on the cross-section will be quoted under this assumption. No model uncertainty will be additionally assigned <sup>①</sup>.

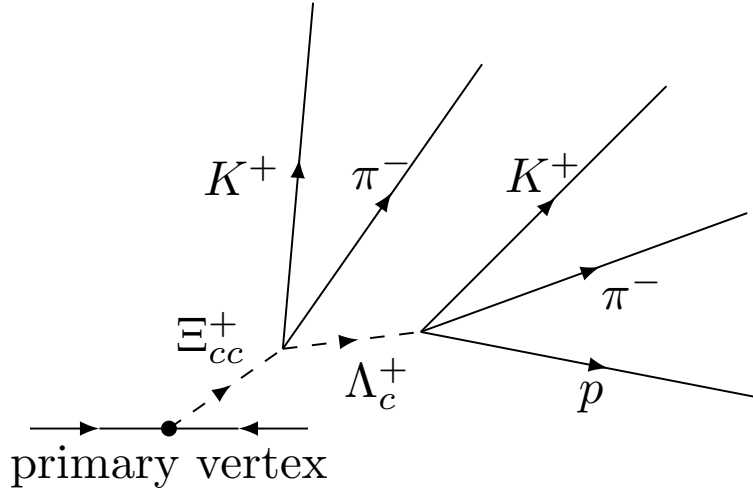
### 3.3 Selection

Both signal and control modes are required to pass stripping and offline selections which are chosen to be as close to each other as possible. In addition, we apply a multivariate selection to the signal mode only, *i.e.* not to the  $\Lambda_c^+$  control mode. This is needed because the expected  $\Xi_{cc}^+$  yield is many orders of magnitude smaller than that of  $\Lambda_c^+$ . At the last step specific trigger lines are required for the two modes.

#### 3.3.1 $\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$ decay topology

There could be two sources of  $\Xi_{cc}^+$  baryons, the  $\Xi_{cc}^+$  directly from the PV and the  $\Xi_{cc}^+$  from the decay of  $b$ -hadrons. In this thesis only the first case is considered, and the decay topology of the corresponding  $\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$  is illustrated in Figure 3.1. Theoretical calculations predict a moderate lifetime for the  $\Xi_{cc}^+$  baryon, therefore the  $\Xi_{cc}^+$  baryon would travel a short but sizable distance before it decays and forms a secondary vertex.

<sup>①</sup> Given the lack of knowledge of the production mechanism, we do not feel that a model uncertainty can meaningfully be assigned. The only other approach would be to quote upper limits separately in various kinematic bins—but keeping in mind that we already plan to quote upper limits as a function of mass and lifetime, that would become hideously complex.

Figure 3.1 Topology of the decay  $\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$ .

The  $\Lambda_c^+$  baryon has a lifetime of 200 fs, therefore, it also flies a significant distance before decaying to  $p$ ,  $K^-$  and  $\pi^+$ . In the reconstruction process, three tracks that are consistent to the proton, kaon and pion hypotheses are searched for and combined to form  $\Lambda_c^+$  candidates; two more tracks which are consistent to the kaon and pion hypotheses and have not been used in the  $\Lambda_c^+$  reconstruction will be combined with the  $\Lambda_c^+$  candidate to form a  $\Xi_{cc}^+$  candidate.

### 3.3.2 Stripping and offline selection

Stripping selection aims to use loose selection criteria to keep signal events and to reject the most obvious background under the constraint of the available bandwidth. After that, tighter offline selection criteria are applied to purify the sample. Variables used in the stripping and offline selection include:

- the track impact parameter (IP)  $\chi^2$  of a candidate with respect to the closest PV. The IP is the shortest distance between the PV and the track of the candidate. The IP  $\chi^2$  is defined as the difference between the PV fit  $\chi^2$  with and without the candidate included in the track set used for the PV reconstruction. The larger the IP  $\chi^2$  is, the smaller the probability that the candidate originates from the PV is.
- the decay vertex  $\chi^2$  per degree of freedom, or  $\chi^2/\text{ndof}$ , of a mother particle, which measures the goodness of the fit to the decay vertex. A small decay vertex  $\chi^2/\text{ndof}$  indicates the mother to most probably be a genuine signal rather than a randomly combined background;

- the constrained DecayTreeFitter (DTF)<sup>[109]</sup> vertex  $\chi^2/\text{ndof}$  of a mother particle with the requirement that the mother particle originates from the PV. The DTF is a technique to fit the decay chain from upstream to downstream, which propagates the information from the mother particle to the daughters and hence has better resolution for kinematic variables;
- the track  $\chi^2/\text{ndof}$ , which measures the quality of the track fit and is useful to remove poorly-reconstructed tracks;
- the flight distance (FD)  $\chi^2$ , which is calculated for an unstable particle candidate as the difference between the PV fit  $\chi^2$  with and without the tracks from this candidate under consideration. The FD  $\chi^2$  is large for long-lived particles;
- the cosine of the decay angle with respect to the direction from PV, or DIRA, where the decay angle of a candidate is the angle between its momentum vector and the vector between the PV and its decay vertex. For well-reconstructed signal decays, the DIRA should peak at 1;
- the maximum distance of the closest approach between all possible pairs of daughters of a candidate, or MAXDOCA. A small MAXDOCA implies that the tracks of the daughters intersect at a common point, and the candidate is likely to be a true signal;
- the distance in  $z$  direction between the decay vertices of  $\Lambda_c^+$  and  $\Xi_{cc}^+$  baryons. The  $\Xi_{cc}^+$  baryon inside LHCb acceptance is highly boosted in the forward direction, hence the  $\Lambda_c^+$  baryon from the  $\Xi_{cc}^+$  decay also travels forward and flies a significant distance before it decays. Therefore, the  $z$  component difference for a signal decay is greater than zero;
- the Delta-Log-Likelihood, or DLL, constructed from the PID log-likelihood, which discriminates between different particle hypotheses. As an example, a cut of  $\text{DLL}(K - \pi) > 0$  requires the track to be more consistent with a kaon hypothesis than with a pion one;
- the transverse momentum,  $p_T$ , which is the component of the momentum in the  $xy$  plane. The  $p_T$  spectra of charm hadrons and their daughters are harder than those of prompt background.



Table 3.1 Selection criteria for the signal mode in stripping.

$\Xi_{cc}^+$	DIRA	$> 0.999$
	Vertex fit $\chi^2/\text{ndof}$	$< 10$
	Flight distance $\chi^2/\text{ndof}$	$> 16$
	Sum of daughters $p_T$	$> 2000 \text{ MeV}/c$
	Invariant mass	$< 4500 \text{ MeV}/c^2$
	$z$ separation between $\Xi_{cc}^+$ and $\Lambda_c^+$ vertices	$> 0.01 \text{ mm}$
Tracks from $\Xi_{cc}^+$	$p$	$> 2000 \text{ MeV}/c$
	$p_T$	$> 250 \text{ MeV}/c$
	Track $\chi^2/\text{ndof}$	$< 4$
	Kaon PID DLL( $K - \pi$ )	$> 0$
	Pion PID DLL( $\pi - K$ )	$> 0$
	IP $\chi^2$	$> 4$
$\Lambda_c^+$	DIRA	$> 0.98$
	Vertex fit $\chi^2/\text{ndof}$	$< 10$
	Flight distance $\chi^2/\text{ndof}$	$> 36$
	MAXDOCA	$< 0.5 \text{ mm}$
	Invariant mass	$2185 < M < 2385 \text{ MeV}/c^2$
	At least one daughter with IP $\chi^2$	$> 30$
Tracks from $\Lambda_c^+$	$p$	$> 2000 \text{ MeV}/c$
	$p_T$	$> 250 \text{ MeV}/c$
	Track $\chi^2/\text{ndof}$	$< 4$
	Proton PID DLL( $p - \pi$ )	$> 5$
	Kaon PID DLL( $K - \pi$ )	$> 5$
	Pion PID DLL( $\pi - K$ )	$> 0$
	IP $\chi^2$	$> 4$

### 3.3.2.1 Stripping

Stripping criteria for the signal and control modes are listed in Table 3.1 and 3.2. Same criteria are used for the  $\Lambda_c^+$  baryon from the  $\Xi_{cc}^+$  decays and the control  $\Lambda_c^+$  baryons. The exception is that for the control mode the DIRA of  $\Lambda_c^+$  is required to be  $> 0.999$  and a prescale factor of 0.05 is applied, which means 5% of the total data are randomly chosen and processed by the control  $\Lambda_c^+$  selection criteria and the other 95% are not.

Table 3.2 Selection criteria for the control mode in stripping.

$\Lambda_c^+$	DIRA	$> 0.999$
	Vertex fit $\chi^2/\text{ndof}$	$< 10$
	Flight distance $\chi^2/\text{ndof}$	$> 36$
	$p_T$	$> 1000 \text{ MeV}/c$
	MAXDOCA	$< 0.5 \text{ mm}$
	Invariant mass	$2185 < M < 2385 \text{ MeV}/c^2$
	At least one daughter with IP $\chi^2$	$> 30$
	$z$ separation between $\Xi_{cc}^+$ and $\Lambda_c^+$ vertices	$> 0.01 \text{ mm}$
Tracks from $\Lambda_c^+$	$p$	$> 2000 \text{ MeV}/c$
	$p_T$	$> 250 \text{ MeV}/c$
	Track $\chi^2/\text{ndof}$	$< 4$
	Proton PID DLL( $p - \pi$ )	$> 10$
	Kaon PID DLL( $K - \pi$ )	$> 10$
	Pion PID DLL( $\pi - K$ )	$> 5$
	IP $\chi^2$	$> 4$

Table 3.3 Offline selection criteria for the signal mode. These criteria are applied in addition to stripping selection criteria.

Particle	Variable	Cut value
$\Xi_{cc}^+$	PV Fit $\chi^2$	$< 50$
	IP $\chi^2$	$< 20$
$K$ from $\Xi_{cc}^+$	DLL( $K - \pi$ )	$> 10$
$K$ from $\Lambda_c^+$	DLL( $K - \pi$ )	$> 10$
$\pi$ from $\Xi_{cc}^+$	DLL( $\pi - K$ )	$> 5$
$\pi$ from $\Lambda_c^+$	DLL( $\pi - K$ )	$> 5$
$p$ from $\Lambda_c^+$	DLL( $p - \pi$ )	$> 10$

### 3.3.2.2 Offline selection

The offline selection criteria for the signal and control modes are listed in Tables 3.3 and 3.4. Same criteria are used for the  $\Lambda_c^+$  baryon from the  $\Xi_{cc}^+$  decay and the control  $\Lambda_c^+$  baryon.

Table 3.4 Offline selection criteria for the control mode. These criteria are applied in addition to stripping selection criteria.

Particle	Variable	Cut value
$K$ from $\Lambda_c^+$	$DLL(K - \pi)$	$> 10$
$\pi$ from $\Lambda_c^+$	$DLL(\pi - K)$	$> 5$
$p$ from $\Lambda_c^+$	$DLL(p - \pi)$	$> 10$

### 3.3.3 Multivariate analysis

After stripping and offline selections, the background level is still overwhelming, as shown in Figure 3.2. To enhance the signal efficiency as much as possible, an artificial neural network (ANN) is developed to extract every bit of information out of the available data.

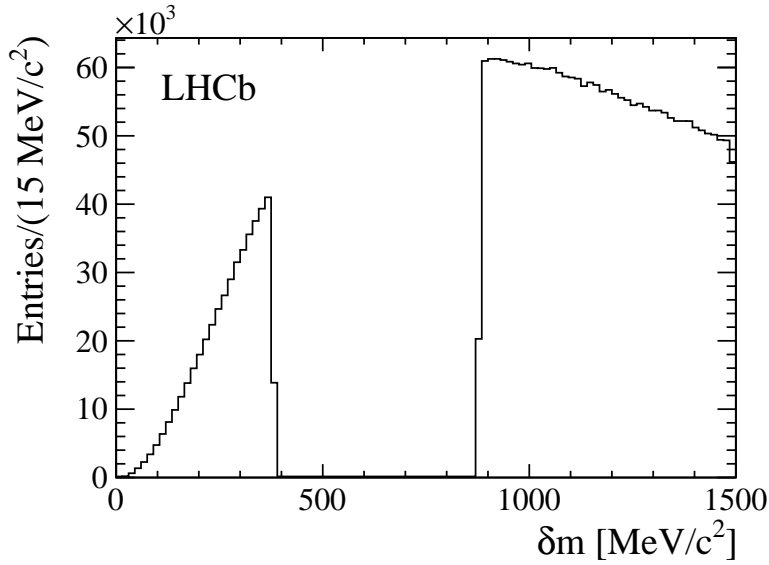


Figure 3.2 Mass distribution for  $\Xi_{cc}^+$  candidates after all the stripping and offline selection criteria. The signal region is blinded.

#### 3.3.3.1 Neural Network

The neural network is a sophisticated and powerful tool for statistical classification and prediction<sup>[110]</sup>. It is a mathematical model which simulates the learning process of the human brain, in which the learning task is accomplished by a series of interconnected

neural neurons. The neurons receive signals from synapses. If the signal transmitted to a neural neuron has a strength larger than a certain threshold, the neuron is activated and sends signals to synapses through axones. The new signals might activate other neurons depending on their strength. The modeling of ANN follows basically the same principle but with a much simpler implementation. Neural neurons in the ANN are called nodes, which receives inputs from other nodes and sends outputs to other nodes if it is activated, as shown in Figure 3.3. Whether a node is activated is determined by its total input, *i.e.* the sum of all inputs from other nodes, through the activation function. For the sake of simplicity, the nodes are usually organized in a structure called multilayer perceptron (MLP), in which they are divided into several layers, and only nodes in adjacent layers are connected, as illustrated in Figure 3.4. If a neuron is activated, the output of the neuron is weighted and transferred to another, and the weights are tuned so that the final node is activated if the input event is a signal, and is not if the input event is a background. The determination of the weights is called training. The ANNs used in the HEP community are all supervised ANNs, which evaluate the weights from categorized data, *i.e.* the category of the input event, either signal or background, is labeled in the input sample<sup>[111]</sup>. If the training sample does not have enough statistics to determine the weights precisely, the neural network may suffer from statistical fluctuation and describes random errors of the training sample instead of underlying relationship, then the response of the network will be biased, and the efficiency will be incorrect. This phenomenon is called overtraining. To prevent the false efficiency calculation, the trained ANN is tested using a statistically independent sample. An over-trained ANN produces very different response distributions for training and testing samples. The multivariate selection used in our selection is implemented with the MLP neural network from the TMVA package<sup>[112]</sup>.

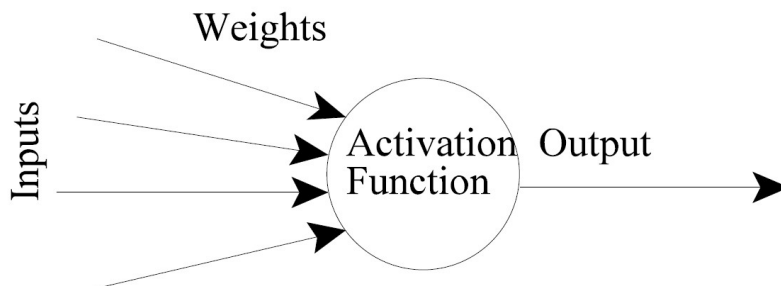


Figure 3.3 Model of the artificial neural network.

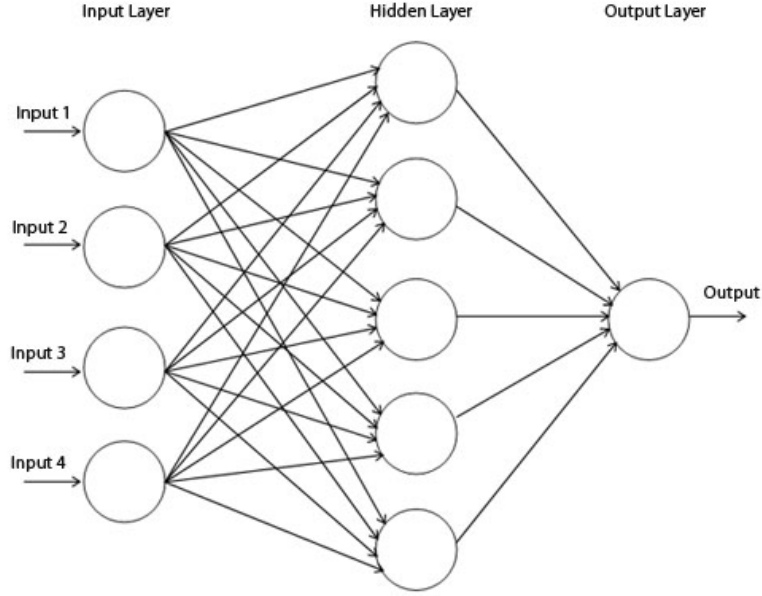


Figure 3.4 Example of the multilayer perceptron.

The input nodes, or variables, are chosen so as to only weakly depend on the  $\Xi_{cc}^+$  lifetime. The signal sample is the truth-matched  $\Xi_{cc}^+$  MC events with the same selection criteria as applied on data. The truth-matching is done with the background category tool<sup>[113]</sup>. Since the trigger efficiency is low and the MC statistics is limited, the ANN training and the selection optimization are done before applying trigger requirements, otherwise we cannot have enough events for training. The signal sample contains 3690 events, half used for training and another half for testing. The background sample is obtained from  $\Xi_{cc}^+$  mass sidebands. Since the mass range of [3300, 3800] MeV/ $c^2$  is blind, the lower sideband is chosen to be [3100, 3300] MeV/ $c^2$ , and the upper sideband is [3800, 4000] MeV/ $c^2$ . The background sample has 5000 events, with all the cuts applied, also half used for training and another half for testing. For the purposes of the MLP training and testing, a  $\Lambda_c^+$  mass window of  $\pm 40$  MeV/ $c^2$  is chosen. This background sample is not excluded from the main analysis, because there is no indication of overtraining, and the training sample contains only 3.5% of the total background candidates.

Input variables used for the neural network training are listed in the following:

- $\Xi_{cc}^+$  MAXDOCA
- $\Xi_{cc}^+$  IP  $\chi^2$
- $\Xi_{cc}^+$  decay vertex  $\chi^2$
- The smallest  $p_T$  of the three  $\Xi_{cc}^+$  daughters

- $\Xi_{cc}^+$  DecayTreeFitter  $\chi^2$  with a PV constraint
- $\Lambda_c^+$  decay vertex  $\chi^2$
- $\Lambda_c^+$  IP  $\chi^2$
- $\Lambda_c^+$  flight distance  $\chi^2$
- $\Lambda_c^+$  MAXDOCA

Distributions of the input variables are shown in Figure 3.5. Blue lines indicate the signal MC, and red shaded regions show the sideband background.

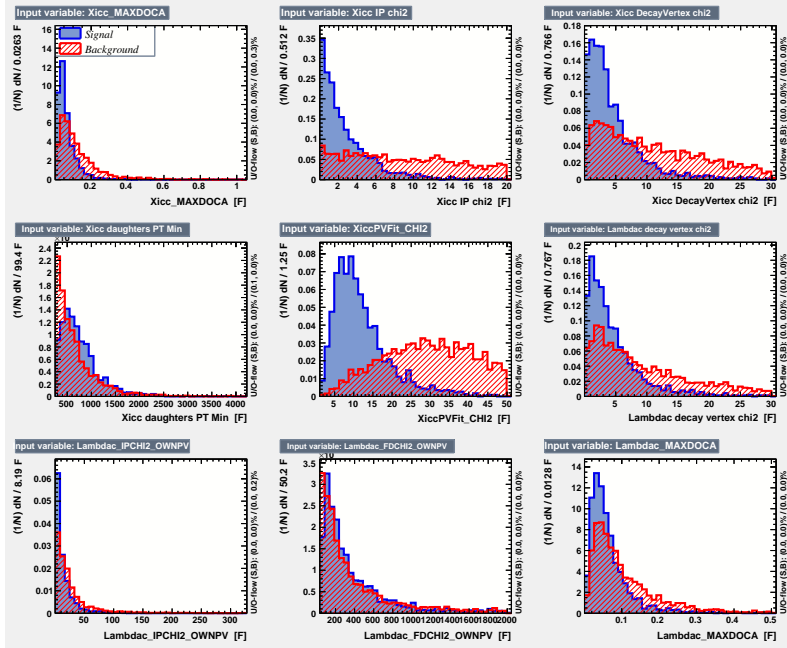


Figure 3.5 Distribution of input variables for the MLP. Blue lines indicate the signal MC, and red shaded regions show the sideband background.

MLP responses for the signal (blue) and the background (red) are shown in Figure 3.6. Testing and training samples agree well for both signal and background, which indicates that the overtraining is small. The corresponding receiver operating characteristic (ROC) curve is displayed in Figure 3.7. The optimal cut is determined by maximizing the expected significance  $s/\sqrt{s+b}$ , where  $s$  is the expected signal yield estimated based on the theoretical cross-section, the luminosity and the efficiency from MC, and  $b$  is the background counts in the signal region estimated with the extrapolation from  $m(\Xi_{cc}^+)$  sidebands. The  $\text{MLP} > 0.8$  is chosen. The efficiency of this cut relative to the MLP input sample is  $\epsilon_{\text{sig}}^{\text{MLP}} = 55.7\%$ , while the estimated retention on combinatoric background in the  $\Lambda_c^+$  signal region is 4.2%.

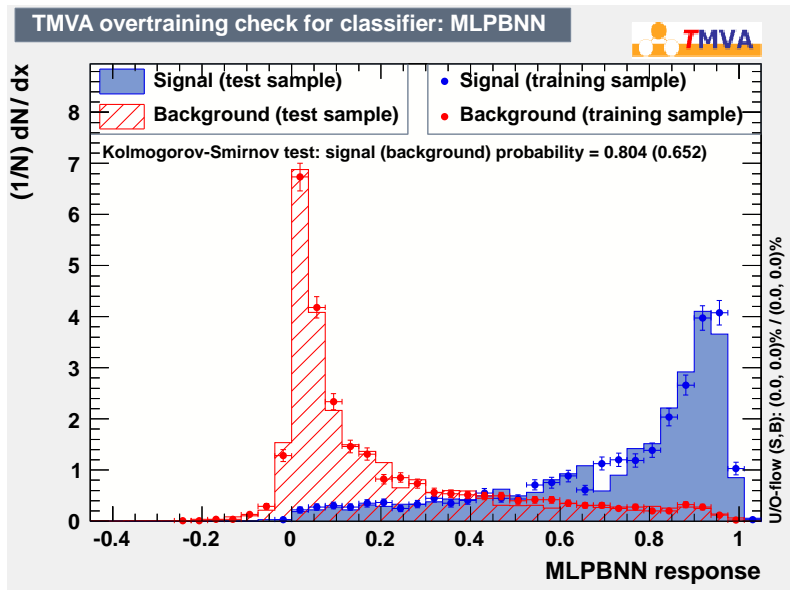


Figure 3.6 MLP response for signal (blue) and background (red).

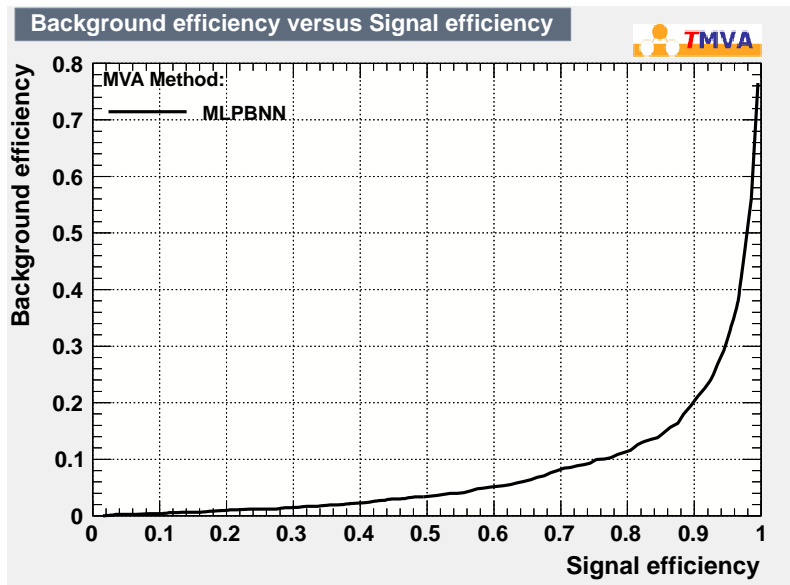


Figure 3.7 ROC curve for the MLP method. The  $x$ -axis is the signal efficiency, and the  $y$ -axis is the background retention rate.

### 3.3.4 Trigger requirements

To better understand signal and background behavior and avoid biases, the data is required to TOS on specific trigger lines. To use similar selection criteria, the trigger lines are chosen to be a TOS chain on the  $\Lambda_c^+$  so that the same trigger lines could be used for both the signal and the control modes; TIS triggers cannot be used since we do not know

Table 3.5 L0 trigger selection criteria of the L0Hadron line.

Variable	Cut value
SPD hits	< 600
$E_T$ in HCAL	> 3500 MeV

the efficiency for  $\Xi_{cc}^+$  events and cannot assume that they will be the same as for the  $\Lambda_c^+$  events since the associated production may be very different. The trigger lines for each trigger phase are

- L0: L0Hadron TOS on a  $\Lambda_c^+$  daughter track
- HLT1: Hlt1TrackAllL0 TOS on a  $\Lambda_c^+$  daughter track
- HLT2: Hlt2CharmHadLambdaC2KPPi TOS on the  $\Lambda_c^+$  baryon

Selection criteria applied in the trigger lines are briefly introduced below. Note that throughout this section all selection criteria are appeared at the trigger (online) level, and related quantities have typically worse resolutions compared to those in the offline analysis.

#### 3.3.4.1 L0 trigger

L0Hadron selection criteria comprise a global event cut (GEC) and an  $E_T$  cut, as listed in Table 3.5. The event should satisfy the GEC and at least one of the tracks from the  $\Lambda_c^+$  candidate should satisfy the  $E_T$  requirement, otherwise the event will not be triggered.

#### 3.3.4.2 HLT1 trigger lines

The HLT1TrackAllL0 trigger line contains four steps to select signal and to efficiently reject background. First it applies GECs to remove high multiplicity events which would consume too much time to process. Then VELO tracks and primary vertices are reconstructed, and selection criteria are applied on track quality and IPs of VELO tracks. At the third step, selected VELO tracks are fully reconstructed and further requirements on track quality,  $p$  and  $p_T$  of these tracks are applied. Finally, tracks passing previous selection criteria are refitted using a BiDirectional Kalman filter and required to satisfy final track quality and IP  $\chi^2$  selection criteria. All requirements of HLT1 are listed in Table 3.6. The event should satisfy the GECs and at least one of the tracks from the



Table 3.6 HLT1 trigger selection criteria of the HLT1TrackAllL0 line.

	Variable	Cut value
Step 1	OT hits	< 15000
	IT hits	< 3000
	VELO hits	< 10000
Step 2	$ n_{\text{exp}} - n_{\text{obs}} $ hits in VELO	< 3
	VELO hits	> 9
	IP to any PV	< 100 mm
Step 3	hits in tracking system	> 16
	$p_T$	> 1700 MeV/c
	$p$	> 10000 MeV/c
Step 4	track $\chi^2/\text{ndof}$	< 2
	min IP $\chi^2$ to all PV	> 16

$\Lambda_c^+$  candidate should satisfy all the HLT1 trigger requirements, otherwise the event will be rejected.

### 3.3.4.3 HLT2 trigger lines

Both L0 and HLT1 lines are generic selections for all hadronic decays, while the HLT2 line is dedicated to charm baryon studies and is described here. The selection criteria applied in the trigger line are reported in Table 3.7. An unusual feature of this trigger is that it uses RICH PID information. Note that this line was only introduced into the online trigger system in July 2011. So the TOS requirement means implicitly that only data taken starting from this time are used in the analysis.

## 3.4 Efficiency ratio determination

With the selection procedure described above, the efficiency of the signal mode can be decomposed as follows:

$$\mathcal{E}_{\text{sig}} = \mathcal{E}_{\text{sig}}^{\text{acc}} \mathcal{E}_{\text{sig}}^{\text{sel|acc}} \mathcal{E}_{\text{sig}}^{\text{PID|sel}} \mathcal{E}_{\text{sig}}^{\text{MLP|PID}} \mathcal{E}_{\text{sig}}^{\text{trigger|MLP}}, \quad (3-5)$$

where the efficiency components are: acceptance (acc), stripping and offline selections (sel), further offline PID cuts (PID), MLP multivariate selection (MLP), and

Table 3.7 HLT2 trigger selection criteria of the Hlt2CharmHadLambdaC2KPPi line.

Selection type	Variable	Value
Combination criteria	Invariant mass calculated from four-momenta of daughters	$2150 \text{ MeV}/c^2 < M < 2430 \text{ MeV}/c^2$
Mother criteria	$p_T$	$> 2500 \text{ MeV}/c$
	cosine of decay angle w.r.t PV	$> 0.99985$
	Flight distance $\chi^2/\text{ndof}$	$> 16$
	Vertex Fit $\chi^2/\text{ndof}$	$< 15$
All daughters	Track $\chi^2/\text{ndof}$	$< 3$
	$p_T$	$> 500 \text{ MeV}/c$
	IP $\chi^2$ to own PV	$> 9$
Extra criteria for Proton	$ p $	$> 10 \text{ GeV}/c$
	PID DLL( $p - \pi$ )	$> 0$
	PID DLL( $p - K$ )	$> 0$

trigger requirements (trigger). Because of the inaccurate description of RICH information in the MC samples, PID cut efficiency should be determined by a data-driven method, instead of being evaluated from MC samples, like for other selection criteria in stripping and cut-based offline selection. PID cuts are therefore separated from main selection efficiency. The | symbol means “given” (*e.g.*  $\varepsilon_{\text{sig}}^{\text{sel|acc}}$  is the efficiency to pass the selection given that a candidate is in the LHCb acceptance) and all selections are cumulative. The corresponding decomposition for the control mode is

$$\varepsilon_{\text{con}} = \varepsilon_{\text{con}}^{\text{acc}} \varepsilon_{\text{con}}^{\text{sel|acc}} \varepsilon_{\text{con}}^{\text{PID|sel}} \varepsilon_{\text{con}}^{\text{trigger|sel}}, \quad (3-6)$$

with no multivariate selection. We can therefore express the production ratio  $R$  defined in Eq. 3-3 as:

$$R = \frac{\varepsilon_{\text{con}}^{\text{acc}}}{\varepsilon_{\text{sig}}^{\text{acc}}} \frac{\varepsilon_{\text{con}}^{\text{sel|acc}}}{\varepsilon_{\text{sig}}^{\text{sel|acc}}} \frac{\varepsilon_{\text{con}}^{\text{PID|sel}}}{\varepsilon_{\text{sig}}^{\text{PID|sel}}} \frac{1}{\varepsilon_{\text{sig}}^{\text{MLP|PID}}} \frac{\varepsilon_{\text{con}}^{\text{trigger|PID}}}{\varepsilon_{\text{sig}}^{\text{trigger|MLP}}} \frac{N_{\text{sig}}}{N_{\text{con}}} = \alpha N_{\text{sig}}, \quad (3-7)$$

where  $\alpha$  is called the single event sensitivity (SES).

The whole efficiency is decomposed into five components: the acceptance efficiency, the stripping and offline efficiency, the PID efficiency, the MLP efficiency and the trigger efficiency. In the following sections, the ratio of each efficiency type for the  $\Lambda_c^+$  control

Table 3.8 Acceptance efficiencies.

Decay mode	MagUp	MagDown	Average
Signal mode	$0.176 \pm 0.001$	$0.178 \pm 0.001$	$0.177 \pm 0.001$
Control mode	$0.213 \pm 0.002$	$0.215 \pm 0.002$	$0.214 \pm 0.001$

Table 3.9 Acceptance efficiencies.

Decay mode	MagUp	MagDown	Average
$\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$ , $\Lambda_c^+ \rightarrow p K^- \pi^+$	$0.176 \pm 0.001$	$0.178 \pm 0.001$	$0.177 \pm 0.001$
$\Lambda_c^+ \rightarrow p K^- \pi^+$	$0.213 \pm 0.002$	$0.215 \pm 0.002$	$0.214 \pm 0.001$

mode and the  $\Xi_{cc}^+$  signal mode will be determined in order to calculate  $\alpha$  as defined in Eq. 3-7.

### 3.4.1 Ratio of acceptance efficiencies

The PYTHIA and GENXICC generate particles in the full solid angle, but LHCb only covers a limited solid cone (10 to 300 mrad in the horizontal plane and 10 to 250 mrad in the vertical plane). It would be a waste of computing resources if we perform full simulation for all these generated particles. LHCb then discards the events in which at least one signal track is outside the solid angle of  $10 < \theta < 400$  mrad, where  $\theta$  is the angle between the momentum of the track and the  $z$  axis. This angular cut does not exactly correspond to the LHCb acceptance. But the overall efficiency is not overestimated, as the real effect of the acceptance will be accounted for by the full simulation. The acceptance efficiency is defined as the fraction of the events that are accepted for later full simulation and is computed using a generator-level Gauss simulation. The results are listed in Table 3.9, and the acceptance efficiencies for the two magnet polarities are consistent. The ratio of acceptance efficiencies is  $1.21 \pm 0.01$ .

### 3.4.2 Ratio of stripping and offline selection efficiencies

Stripping and offline selection efficiencies are calculated using MC samples. Note that the PID related cuts are considered separately in the next subsection due to the

Table 3.10 Selection efficiency (stripping and non-PID offline selection criteria).

Decay mode	# of events generated	# of events selected	selection efficiency
Signal mode	$2.26 \times 10^6$	7187	$(3.18 \pm 0.04) \times 10^{-3}$
Control mode	$2.0 \times 10^6$	$26 \times 10^3$	$(1.22 \pm 0.01) \times 10^{-2}$

imperfect modeling of RICH by MC. For the signal mode, 7,187 out of  $2.26 \times 10^6$   $\Xi_{cc}^+$  MC candidates are retained after the selection cuts; for the control mode,  $2.6 \times 10^4$  out of  $2.02 \times 10^6$   $\Lambda_c^+$  MC candidates are retained. The MC stripping is run in the flagging mode, and the prescale factor of the control mode is not taken into account. The ratio of stripping and cut-based offline selection efficiencies between the control and the signal modes is  $3.84 \pm 0.05$ .

### 3.4.2.1 Tracking efficiency corrections

The tracking efficiency is included implicitly in the stripping efficiency, since only signal candidates which have all the final tracks reconstructed could pass the stripping selection. Since the detector response might not be described perfectly and some second-order effects might be neglected in MC, the tracking efficiency is not properly reproduced by MC. These data/MC discrepancies have to be corrected. There is a standard procedure described below to correct for these differences. An additional systematic uncertainty arises due to limitations in the correction procedure, and is also mentioned below.

The LHCb tracking group have compared the tracking efficiency in data and MC, and calculated the data/MC efficiency ratio for individual tracks as a function of momentum and pseudorapidity<sup>[75]</sup>. However, before considering that correction there is another effect that needs to be considered: besides track kinematics, the efficiency depends on track multiplicity in the event. This latter effect is neither described by MC, so the track multiplicity in MC needs to be reweighted to match that expected in data. For the  $\Lambda_c^+$  control mode we extract the track multiplicity from data. For the signal mode, however, there is no  $\Xi_{cc}^+$  data sample that can be used for reweighting. Instead, a suitable proxy should be chosen. There are a few possibilities but none provides an exact match. The  $B_s^0$  decays are chosen since they also involve two non-light quark-antiquark pairs ( $b\bar{b}$  and  $s\bar{s}$  instead of two  $c\bar{c}$  pairs). The specific control sample used is a sample of

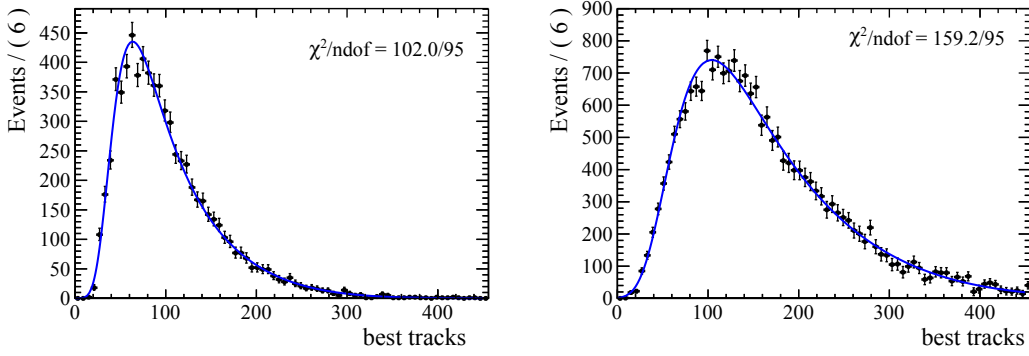


Figure 3.8 Best track multiplicity distribution for  $\Xi_{cc}^+$  MC (left) and  $B_s^0 \rightarrow J/\psi \phi$  (right).

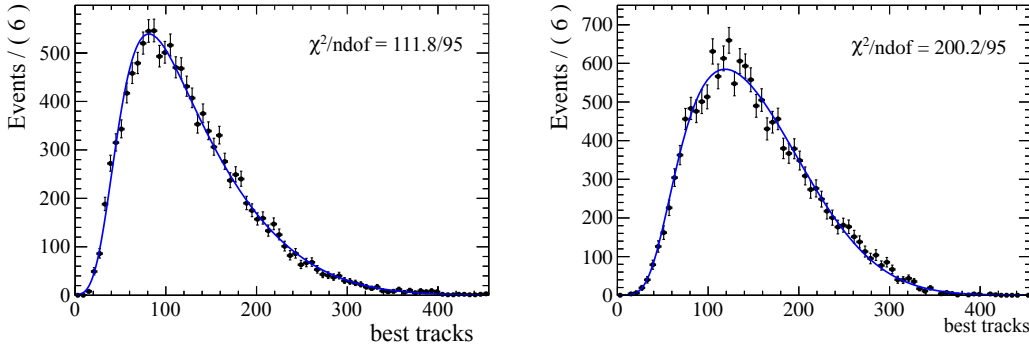


Figure 3.9 Best track multiplicity distribution for  $\Lambda_c^+$  MC (left) and  $\Lambda_c^+$  data (right).

$B_s^0 \rightarrow J/\psi \phi$  decays. The *sPlot* method is used to extract the multiplicity distribution for events with  $B_s^0$  meson decays<sup>[82]</sup>. For both the  $\Lambda_c^+$  and  $B_s$  samples, the multiplicity distribution of best tracks is adequately described by an empirical function consisting of a Landau function multiplied by a Gaussian function, where best tracks are the set of reconstructed tracks used for physics analysis. The fitted distributions are shown in Figure 3.8 and Figure 3.9.

Using the tag-and-probe method on the  $J/\psi \rightarrow \mu^+ \mu^-$  sample, the tracking efficiency ratio between data/MC is evaluated as a function of momentum and pseudorapidity of the track, as shown in Figure 3.10.

A toy MC method is used to apply this correction<sup>[114,115]</sup>. For each toy experiment, the central values in the table are smeared according to their errors. The smeared values are saved as the correction table for that toy. Then for each daughter track of each candidate in the signal MC samples the correction can be calculated from the correction table. Combining the per-daughter corrections gives us a per-candidate correction, and

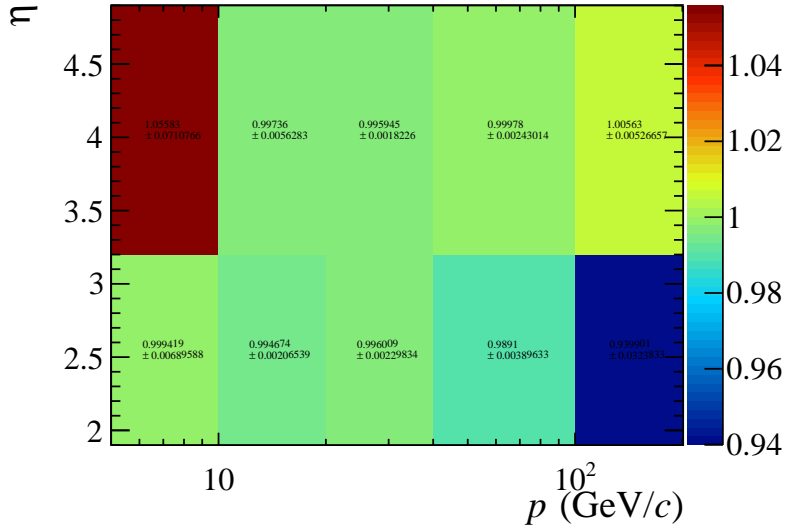


Figure 3.10 Tracking efficiency ratio between data and MC in bins of momentum and pseudorapidity of the track.

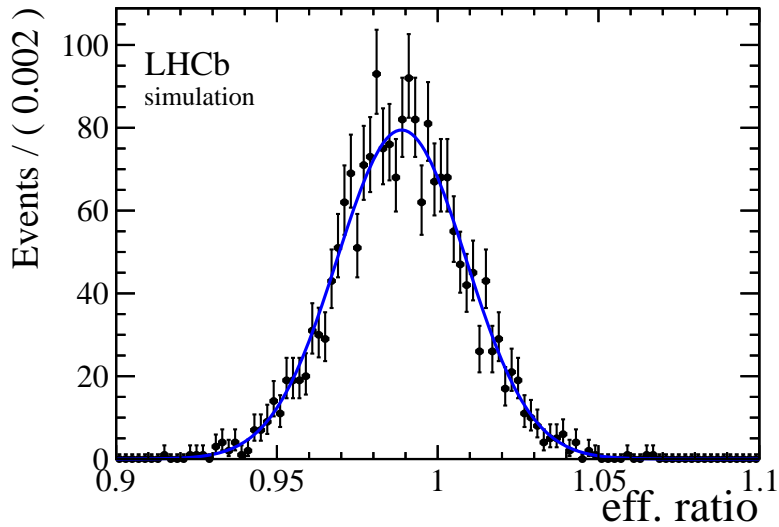


Figure 3.11 Fit of the efficiency ratio distribution from toy MC.

we then average across the candidates, whose multiplicity is weighted as described above, to give a per-toy correction for the  $\Xi_{cc}^+$  candidate and for the  $\Lambda_c^+$  candidate. From this procedure a per-toy correction to  $\varepsilon_{\text{con}}^{\text{sel|acc}} / \varepsilon_{\text{sig}}^{\text{sel|acc}}$  is obtained. By performing many toy experiments we obtain a distribution of these correction ratios, shown in Figure 3.11. Fitting this distribution we obtain a mean correction of  $0.989 \pm 0.020$ .

In this method kinematic correlations among daughter tracks are neglected—a

systematic uncertainty corresponding to this effect will be assigned in Section 3.7.1.

Finally, note that the efficiency tables described above is obtained with a muon sample, while the signal and control modes involve hadronic final states which will suffer additional material interactions. A systematic uncertainty corresponding to this effect will be assigned in Section 3.7.1.

### 3.4.3 Ratio of PID cut efficiencies

It is known that the distributions of PID variables are different between data and MC, because several 2nd-order effects are not considered in simulation and the performance of RICH varies with running conditions. The LHCb PID working group provides the PIDCalib package to solve this problem. It is assumed that the PID selection efficiency depends on track momentum and pseudorapidity as well as the track multiplicity in the event. Momentum and pseudorapidity distributions of tracks, which are from standard calibration samples selected without PID cuts, can then be reweighted to match the distributions expected for tracks produced in  $\Xi_{cc}^+$  decays according to the signal MC. The efficiency of our PID cuts can then be obtained by applying them to the weighted, background-subtracted calibration samples. The calibration sample used for kaons and pions is  $D^{*+} \rightarrow D^0\pi^+$ ,  $D^0 \rightarrow K^-\pi^+$ , and that used for protons is  $\Lambda^0 \rightarrow p\pi^-$ .

#### 3.4.3.1 Calibration procedure

To reweight the calibration samples the kinematic distribution for the data before PID cuts is needed. However, the standard stripping applied to signal MC includes PID cuts (and so can have a biased distribution). Therefore, the stripping with the PID cuts disabled is rerun for the signal and the control modes, and then the cut-based offline selections (again without PID cuts) are applied.

To account for the variation in the RICH performance over time, calibration samples are split into several subsamples according to changes in the data-taking conditions. The PID cut efficiencies are different between these samples. To minimize systematic effects, the efficiency ratios between control and signal modes are calculated for each run range separately, then combined in a lumi-weighted average. Assuming the luminosity of sample  $i$  is  $L_i$ , and the efficiency ratio is  $r_i$ , then the average efficiency ratio  $r$  is

$$r = \frac{\sum_i L_i r_i}{\sum_i L_i} \quad (3-8)$$

Table 3.11 Binning scheme for  $K$  and  $\pi$  momentum and pseudorapidity and event multiplicity used for PID calibration.

Variable	Binning scheme
$p$ [ MeV/ $c$ ]	[2000, 9300, 15600, 17675, 20000, 26000, 29650, 40000, 55000, 100000]
$\eta$	[1.5, 2.8, 3.6, 5.0]
multiplicity	[0, 120, 200, 400]

Table 3.12 Binning scheme for  $p$  momentum and pseudorapidity and event multiplicity used for PID calibration.

Variable	Binning scheme
$p$ [ MeV/ $c$ ]	[2000, 9300, 15600, 17675, 20000, 25000, 29650, 50000, 125000]
$\eta$	[1.5, 2.8, 3.6, 5.0]
multiplicity	[0, 120, 200, 400]

The multiplicity of the events with  $\Xi_{cc}^+$  baryon is possibly considerably under-estimated in MC. Since the PID efficiency is known to depend on the track multiplicity, this difference must be taken into account. We use a similar approach as that discussed in Section 3.4.2.1. The multiplicity distribution of the  $\Xi_{cc}^+$  MC is weighted to match that of the  $B_s \rightarrow J/\psi\phi$  data, and the  $\Lambda_c^+$  MC is weighted to match the  $\Lambda_c^+$  data.

Since kinematic distributions of the signal mode are very different from those of the calibration mode, kinematical reweighting is needed to match these distributions. Due to the limited statistics of the MC sample, a binning scheme better suited to our analysis is used (see Table 3.11 for  $K$  and  $\pi$  and Table 3.12 for  $p$ ). The efficiencies obtained for the  $\Xi_{cc}^+$  final state tracks can be found in Tables 3.13, 3.14, 3.15, 3.16, and 3.17. Those for the  $\Lambda_c^+$  final state tracks can be found in Tables 3.18, 3.19, and 3.19.

### 3.4.3.2 Luminosity of each run range

The luminosity of each sample is listed in Table 3.21 and the efficiency ratio in each run range is given in Table 3.22.



Table 3.13 Calibration of  $K$  identification for kaons from  $\Lambda_c^+$  in the  $\Xi_{cc}^+$  MC.

SubID	MagDown		MagUp	
	Run range	PIDK > 10 Efficiency	Run Runge	PIDK > 10 Efficiency
2	92821-93224	–	94261-96641	$(86.43 \pm 0.25)\%$
3	93225-97884	$(87.09 \pm 0.21)\%$	96642-100243	$(86.48 \pm 0.26)\%$
4	97885-98198	$(87.65 \pm 0.20)\%$	100244-102505	$(86.17 \pm 0.23)\%$
5	98199 -101905	$(87.28 \pm 0.21)\%$	102506-102893	$(86.32 \pm 0.23)\%$
6	101906-102378	$(87.09 \pm 0.21)\%$	102894-104263	$(85.90 \pm 0.36)\%$
7	102379-103361	$(86.90 \pm 0.19)\%$		
8	103362-103686	$(87.42 \pm 0.20)\%$		

Table 3.14 Calibration of  $p$  identification for protons from  $\Lambda_c^+$  in the  $\Xi_{cc}^+$  MC.

SubID	MagDown		MagUp	
	Run range	PIDp > 10 Efficiency	Run Runge	PIDp > 10 Efficiency
2	92821-93224	–	94261-96641	$(88.90 \pm 0.38)\%$
3	93225-97884	$(89.73 \pm 0.40)\%$	96642-100243	$(89.49 \pm 0.74)\%$
4	97885-98198	$(89.42 \pm 0.28)\%$	100244-102505	$(88.65 \pm 0.50)\%$
5	98199 -101905	$(89.87 \pm 0.93)\%$	102506-102893	$(88.55 \pm 0.49)\%$
6	101906-102378	$(89.42 \pm 0.44)\%$	102894-104263	$(89.05 \pm 0.94)\%$
7	102379-103361	$(90.17 \pm 0.66)\%$		
8	103362-103686	$(88.63 \pm 0.68)\%$		

Table 3.15 Calibration of  $\pi$  identification for pions from  $\Lambda_c^+$  in the  $\Xi_{cc}^+$  MC.

SubID	MagDown		MagUp	
	Run range	PIDK < -5 Efficiency	Run Runge	PIDK < -5 Efficiency
2	92821-93224	–	94261-96641	(70.67 ± 0.21)%
3	93225-97884	(69.57 ± 0.18)%	96642-100243	(67.19 ± 0.22)%
4	97885-98198	(66.62 ± 0.19)%	100244-102505	(65.62 ± 0.22)%
5	98199 -101905	(68.09 ± 0.19)%	102506-102893	(67.27 ± 0.22)%
6	101906-102378	(67.85 ± 0.18)%	102894-104263	(65.81 ± 0.36)%
7	102379-103361	(67.84 ± 0.18)%		
8	103362-103686	(65.50 ± 0.20)%		

Table 3.16 Calibration of  $K$  identification for kaons from  $\Xi_{cc}^+$  in the  $\Xi_{cc}^+$  MC.

SubID	MagDown		MagUp	
	Run range	PIDK > 10 Efficiency	Run Runge	PIDK > 10 Efficiency
2	92821-93224	–	94261-96641	(87.02 ± 0.26)%
3	93225-97884	(86.90 ± 0.23)%	96642-100243	(87.06 ± 0.26)%
4	97885-98198	(87.63 ± 0.22)%	100244-102505	(86.89 ± 0.24)%
5	98199 -101905	(87.21 ± 0.23)%	102506-102893	(86.94 ± 0.23)%
6	101906-102378	(86.99 ± 0.23)%	102894-104263	(86.56 ± 0.37)%
7	102379-103361	(86.79 ± 0.21)%		
8	103362-103686	(87.42 ± 0.22)%		

Table 3.17 Calibration of  $\pi$  identification for pions from  $\Xi_{cc}^+$  in the  $\Xi_{cc}^+$  MC.

SubID	MagDown		MagUp	
	Run range	PIDK < -5 Efficiency	Run Runge	PIDK < -5 Efficiency
2	92821-93224	–	94261-96641	(69.50 ± 0.20)%
3	93225-97884	(67.96 ± 0.21)%	96642-100243	(65.76 ± 0.21)%
4	97885-98198	(64.86 ± 0.22)%	100244-102505	(64.21 ± 0.21)%
5	98199 -101905	(66.40 ± 0.22)%	102506-102893	(65.98 ± 0.21)%
6	101906-102378	(66.22 ± 0.22)%	102894-104263	(64.34 ± 0.34)%
7	102379-103361	(66.18 ± 0.22)%		
8	103362-103686	(63.71 ± 0.23)%		

 Table 3.18 Calibration of  $K$  identification for kaons from  $\Lambda_c^+$  in the  $\Lambda_c^+$  MC.

SubID	MagDown		MagUp	
	Run range	PIDK > 10 Efficiency	Run Runge	PIDK > 10 Efficiency
2	92821-93224	–	94261-96641	(86.10 ± 0.18)%
3	93225-97884	(85.86 ± 0.17)%	96642-100243	(85.97 ± 0.19)%
4	97885-98198	(86.38 ± 0.17)%	100244-102505	(85.84 ± 0.17)%
5	98199 -101905	(86.14 ± 0.18)%	102506-102893	(85.79 ± 0.17)%
6	101906-102378	(85.77 ± 0.18)%	102894-104263	(85.62 ± 0.27)%
7	102379-103361	(85.64 ± 0.16)%		
8	103362-103686	(86.12 ± 0.17)%		

Table 3.19 Calibration of  $p$  identification for protons from  $\Lambda_c^+$  in the  $\Lambda_c^+$  MC.

SubID	MagDown		MagUp	
	Run range	PIDp > 10 Efficiency	Run Runge	PIDp > 10 Efficiency
2	92821-93224	–	94261-96641	$(88.92 \pm 0.54)\%$
3	93225-97884	$(89.56 \pm 0.60)\%$	96642-100243	$(89.42 \pm 0.89)\%$
4	97885-98198	$(89.03 \pm 0.49)\%$	100244-102505	$(88.65 \pm 0.67)\%$
5	98199 -101905	$(94.03 \pm 1.22)\%$	102506-102893	$(88.39 \pm 0.59)\%$
6	101906-102378	$(89.38 \pm 0.70)\%$	102894-104263	$(91.50 \pm 1.21)\%$
7	102379-103361	$(90.37 \pm 1.17)\%$		
8	103362-103686	$(88.76 \pm 0.71)\%$		

Table 3.20 Calibration of  $\pi$  identification for pions from  $\Lambda_c^+$  in the  $\Lambda_c^+$  MC.

SubID	MagDown		MagUp	
	Run range	PIDK < -5 Efficiency	Run Runge	PIDK < -5 Efficiency
2	92821-93224	–	94261-96641	$(71.61 \pm 0.18)\%$
3	93225-97884	$(69.55 \pm 0.18)\%$	96642-100243	$(68.26 \pm 0.19)\%$
4	97885-98198	$(66.67 \pm 0.18)\%$	100244-102505	$(66.74 \pm 0.18)\%$
5	98199 -101905	$(68.09 \pm 0.18)\%$	102506-102893	$(68.27 \pm 0.19)\%$
6	101906-102378	$(67.82 \pm 0.18)\%$	102894-104263	$(66.91 \pm 0.31)\%$
7	102379-103361	$(67.82 \pm 0.18)\%$		
8	103362-103686	$(65.55 \pm 0.20)\%$		

Table 3.21 Luminosity of each subsample.

SubID	MagDown		MagUp	
	Run range	Lumi ( $\text{pb}^{-1}$ )	Run Runge	Lumi ( $\text{pb}^{-1}$ )
2	92821-93224	–	94261 -96641	$33.02 \pm 0.14$
3	93225-97884	$57.19 \pm 0.19$	96642 -100243	$67.39 \pm 0.21$
4	97885-98198	$68.23 \pm 0.21$	100244 -102505	$69.76 \pm 0.22$
5	98199 -101905	$68.94 \pm 0.21$	102506 -102893	$67.94 \pm 0.21$
6	101906-102378	$69.59 \pm 0.21$	102894-104263	$25.26 \pm 0.12$
7	102379-103361	$68.97 \pm 0.21$		
8	103362-103686	$57.59 \pm 0.18$		

Table 3.22 Efficiency ratio for different subsamples.

SubID	MagDown		MagUp	
	Run range	Efficiency ratio	Run Runge	Efficiency ratio
2	92821-93224	–	94261 -96641	$1.67 \pm 0.02$
3	93225-97884	$1.67 \pm 0.02$	96642 -100243	$1.76 \pm 0.03$
4	97885-98198	$1.73 \pm 0.02$	100244 -102505	$1.82 \pm 0.02$
5	98199 -101905	$1.78 \pm 0.03$	102506 -102893	$1.76 \pm 0.02$
6	101906-102378	$1.71 \pm 0.02$	102894-104263	$1.87 \pm 0.04$
7	102379-103361	$1.72 \pm 0.03$		
8	103362-103686	$1.77 \pm 0.02$		

### 3.4.3.3 Ratio of PID selection efficiency

With all the above factors taken into account, the ratio of PID selection efficiency is calculated to be  $1.75 \pm 0.03$ .

### 3.4.4 Ratio of the MLP selection efficiency

To ensure the signal efficiency is not overestimated, the test sample, which is independent of the training sample, is used to evaluate the MLP selection efficiency. For the chosen MLP cut of  $> 0.8$  (see Section 3.3.3), the efficiency is  $\varepsilon_{\text{sig}}^{\text{MLP}} = 0.557 \pm 0.012$ .

Note the MLP cut is only applied to the signal mode, not to the control mode. So taking the MLP efficiency for the control mode to be 1 by construction, the ratio of efficiencies is obtained as  $(1/\varepsilon_{\text{sig}}^{\text{MLP}}) = 1.80 \pm 0.04$ .

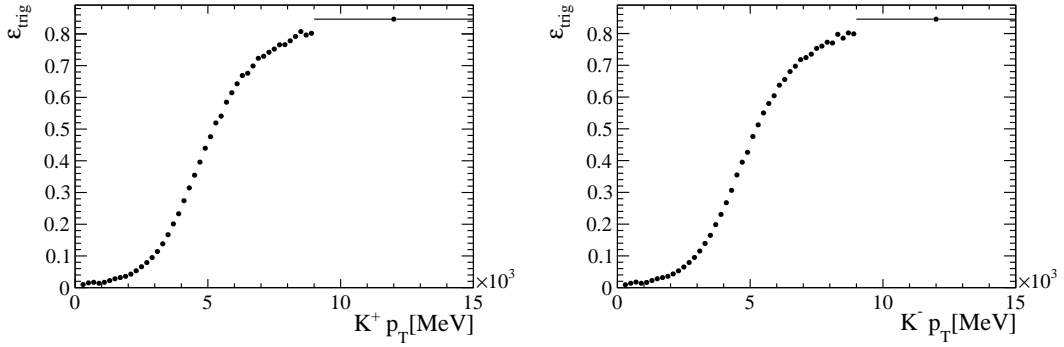
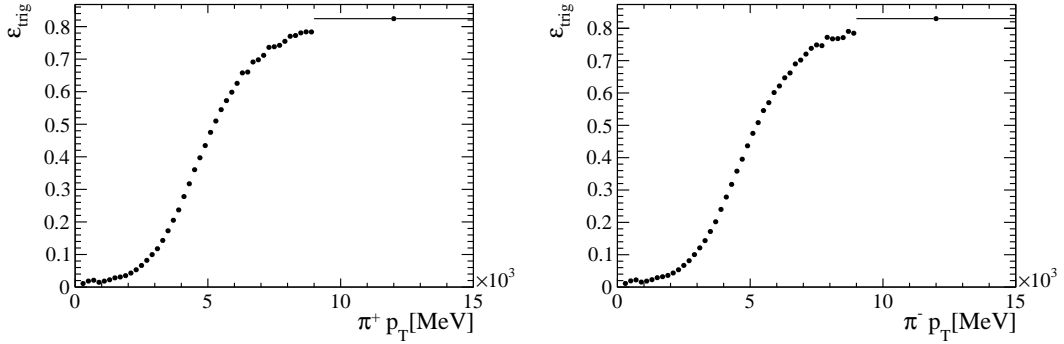
### 3.4.5 Ratio of trigger efficiency

The calculation of trigger efficiency is split into two parts according to whether MC describes the detector response well.

#### 3.4.5.1 L0 efficiency

The MC does not reproduce the L0 efficiency well for a number of reasons, such as that the L0Calo efficiency is not constant over time due to HCAL and ECAL ageing and recalibration<sup>[116]</sup>. It is very difficult to take this into account in MC since the variations for different zones are not the same. The Calo working group provides several trigger efficiency tables in bins of  $p_T$  for different types of track. These efficiencies are calculated using tracks from the PIDCalib sample with the TISTOS method<sup>[117]</sup>. We have the same problem as for the PID calibration, namely that the kinematic distribution of protons in our  $\Xi_{cc}^+$  signal sample does not properly match that in the calibration sample: protons from inclusive  $\Lambda$  decays to long tracks have low momentum, and hence events with  $p_T > 4600 \text{ MeV}/c$  are very rare, but the signal sample has many events in this region and the trigger efficiency is a strong function of  $p_T$ . To get a more accurate result, the control sample ( $\Lambda_c^+ \rightarrow pK^-\pi^+$ ) is used to evaluate the L0 trigger efficiency in the high  $p_T$  region, using the method as described below (following the documentation<sup>[118,119]</sup> linked from the L0Calo twiki page<sup>[117]</sup>).

The trigger efficiency is calculated using the efficiency tables with a toy MC method. For each toy MC experiment, the efficiency for each track is smeared according to the


 Figure 3.12 Trigger efficiency in bins of  $p_T$  for (a)  $K^+$  and (b)  $K^-$ .

 Figure 3.13 Trigger efficiency in bins of  $p_T$  for (a)  $\pi^+$  and (b)  $\pi^-$ .

uncertainty in the corresponding  $p_T$  bin, then the trigger efficiency for this event is obtained as

$$\varepsilon = 1 - (1 - \varepsilon_p)(1 - \varepsilon_K)(1 - \varepsilon_\pi) \quad (3-9)$$

The individual track L0 efficiencies are plotted in Figures 3.12, 3.13, and 3.14. After many toy experiments, the distribution of the efficiency could be drawn (Figure 3.15) and the efficiency and the uncertainty can be extracted from a fit. Using a Gaussian function to fit the distribution, the ratio of L0 efficiencies between the control and signal modes is found to be  $1.451 \pm 0.002$ .

### 3.4.5.2 HLT1 and HLT2 efficiency

The HLT1 and HLT2 efficiencies are calculated from the MC directly. The results are shown in Table 3.23. The ratio of HLT efficiencies between the control and signal modes is measured to be  $0.93 \pm 0.17$ .

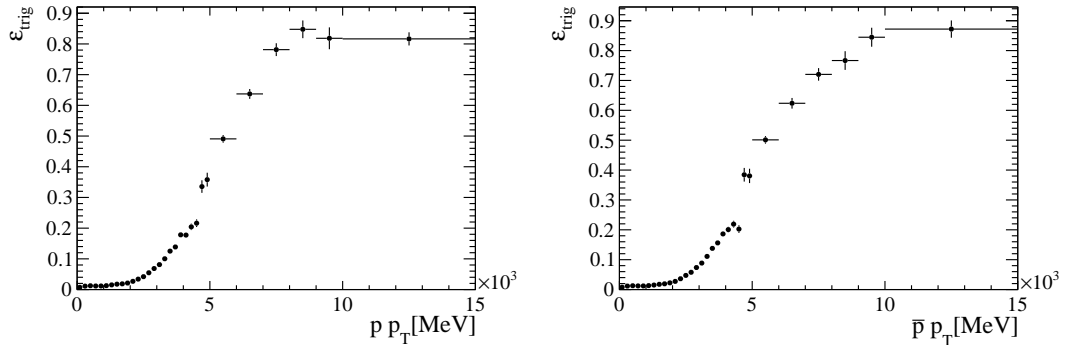
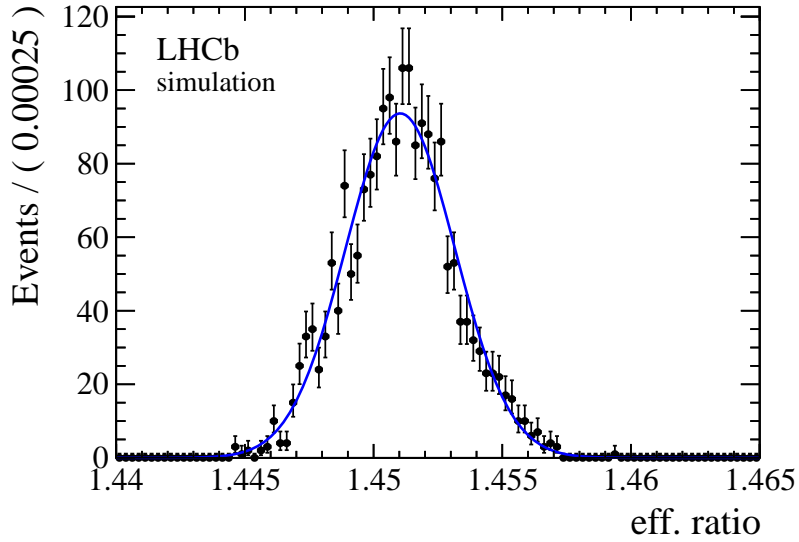

 Figure 3.14 Trigger efficiency in bins of  $p_T$  for (a)  $p$  and (b)  $\bar{p}$ .


Figure 3.15 Fit of the trigger efficiency ratio of control mode and signal mode.

Table 3.23 Trigger efficiency (with respect to the previous trigger level).

Trigger line	Signal mode	Control mode
-	1028	15880
L0	76	2207
L0+Hlt1	46((60.53 $\pm$ 5.61)%)	1314((59.54 $\pm$ 1.04)%)
L0+Hlt1+Hlt2	23((50.00 $\pm$ 7.37)%)	622((47.34 $\pm$ 1.38)%)



### 3.4.5.3 Ratio of trigger efficiencies

Combining the L0, HLT1 and HLT2 efficiencies together, the ratio of trigger efficiencies between the control and signal modes is obtained as  $1.35 \pm 0.25$ .

### 3.4.6 Summary of efficiency ratios

Each of the efficiency ratios in Eq. 3-7 has been considered and evaluated, and numerical results of the ratios are summarized in Table 3.24. The systematic uncertainties on the efficiency ratios are considered in Section 3.7, and the variations of the combined efficiency ratio with the mass and lifetime of the  $\Xi_{cc}^+$  baryon are discussed in Section 3.8.1 and Section 3.8.2.

Table 3.24 Summary of efficiency ratios

Efficiency ratio	Result
$\mathcal{E}_{\text{con}}^{\text{acc}} / \mathcal{E}_{\text{sig}}^{\text{acc}}$	$1.21 \pm 0.01$
$\mathcal{E}_{\text{con}}^{\text{sel acc}} / \mathcal{E}_{\text{sig}}^{\text{sel acc}}$	$3.84 \pm 0.05$
$\mathcal{E}_{\text{con}}^{\text{PID sel}} / \mathcal{E}_{\text{sig}}^{\text{PID sel}}$	$1.75 \pm 0.03$
$1 / \mathcal{E}_{\text{sig}}^{\text{MLP PID}}$	$1.80 \pm 0.04$
$\mathcal{E}_{\text{con}}^{\text{trigger PID}} / \mathcal{E}_{\text{sig}}^{\text{trigger MLP}}$	$1.35 \pm 0.25$

## 3.5 Yield determination

In this section the yield extraction procedures for the signal and control modes are explained.

### 3.5.1 Fit of the invariant mass distribution for $\Lambda_c^+$ candidates from the control sample in data

After all selection and trigger requirements, an unbinned maximum likelihood is performed to extract the  $\Lambda_c^+$  candidates in the control mode. The Probability Density Function (PDF) for the signal is a double Gaussian function, *i.e.* the sum of two Gaussian functions with a common mean and different resolution, and the PDF for the background is a first order Chebychev polynomial. The fit curve is shown in Figure 3.16. The signal yield obtained from the fit is  $N_{\text{con}} = 40,911 \pm 364$ . Note that there is a prescale factor of

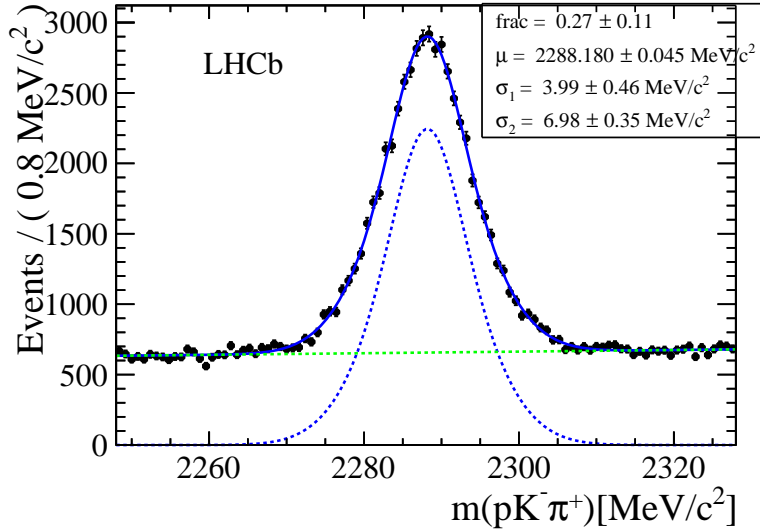


Figure 3.16 Yield of the control  $\Lambda_c^+$  in data. A prescale factor of 0.05 is present in the stripping.

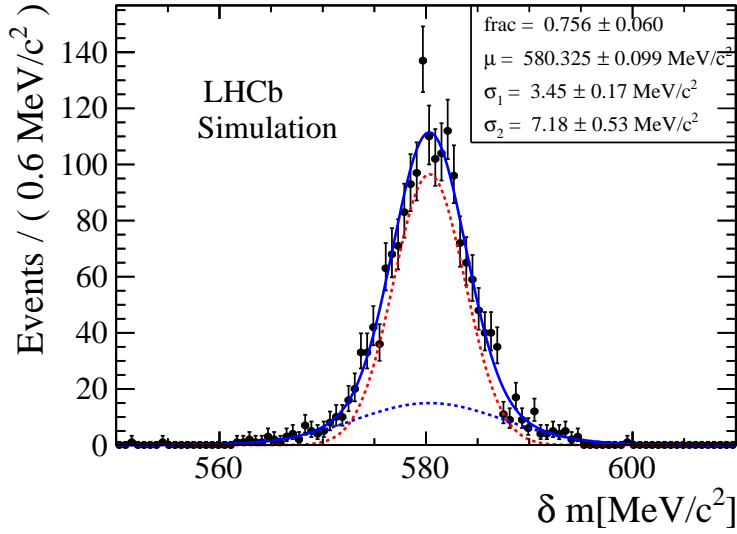
0.05 in the stripping line for the  $\Lambda_c^+$  control mode, so the actual signal in the data sample is 20 times larger:  $(818 \pm 7) \times 10^3$ . This will be included in the calculation of  $\alpha$ .

The resolution measured from this high-statistics fit will be used as a guide to determine an appropriate signal window size in Section 3.5.3. However, we do not assume that the lineshape is the same, since the kinematics of the  $\Lambda_c^+$  baryon may differ between the control and signal samples.

### 3.5.2 Fit of the invariant mass distribution for $\Xi_{cc}^+$ candidates from the control sample in MC

Although the signal yield is not obtained from the fit directly (see next Section), it is still useful to have an idea of the signal lineshape and resolution before unblinding, therefore the signal MC is used as a reference.

Figure 3.17 shows the distribution of  $\delta m$  in truth-matched signal events. The lineshape for fitting is a double Gaussian function. The fraction-weighted average of the resolution is  $4.4 \text{ MeV}/c^2$ , and we take  $5 \text{ MeV}/c^2$  as the resolution of the  $\Xi_{cc}^+$  candidates, and the signal region is defined as  $\pm 2\sigma$  around the mass.

Figure 3.17 The  $\delta m$  distribution in the truth-matched signal MC.

### 3.5.3 Procedure to measure the $\Xi_{cc}^+$ yield on data

#### 3.5.3.1 Overview

After all selection criteria applied, the background level becomes very low and the expected signal yield is zero or close to zero. This makes a  $\chi^2$  fit impossible, since  $\chi^2$  fits are biased when the yield per bin is small. It also makes likelihood fits difficult to interpret, since negative PDFs are not allowed so the signal yield is forced to be positive, which also causes biases. Therefore, the signal yield is not obtained from the fit directly but determined by subtracting the expected background from the total event number in the signal region, which is not statistically optimal but more robust against fluctuations.

Two complementary procedures are used to extract the yield in data. The general approach for both methods are defined below:

1. Define a signal mass hypothesis (e.g.  $\delta m_0 = 602 \text{ MeV}/c^2$ ).
2. Define a symmetric signal window around that mass value (e.g.  $\delta m_0 - 2\sigma < \delta m < \delta m_0 + 2\sigma$  with  $\sigma$  taken to be  $5 \text{ MeV}/c^2$ ).
3. Use the data outside that window to determine the background level, and estimate the expected background  $N_b$  inside the signal window.
4. Count the number of candidates in the signal window  $N_{s+b}$ .
5. The signal yield is  $N_s = N_{s+b} - N_b$ .

In both methods, we assume that the background can be described as the sum of

two components: a) combinatorial background related to  $\Xi_{cc}^+$  decay and b) combinatorial background related to  $\Lambda_c^+$  decay. Both have a smooth distribution in  $\delta m$  (though not necessarily the same distribution). The combinatorial component a) also has smooth distribution in  $m(\Lambda_c^+)$ , whereas the combinatorial component b) is constructed from a real  $\Lambda_c^+$  and therefore peaks at the known mass of  $\Lambda_c^+$ .

The first method, referred to as “25-Tiles”, is a 2D local sideband subtraction in  $(m(\Lambda_c^+), \delta m)$  that does not use any fitting at all. It is basically a generalization of 1D sideband subtraction but assuming a quadratic background. We define a big rectangle with size of  $80 \text{ MeV}/c^2$  in  $m(\Lambda_c^+)$  and size of  $200 \text{ MeV}/c^2$  in  $\delta m$ , and a small signal window of size of  $30 \text{ MeV}/c^2$  in  $m(\Lambda_c^+)$  and  $20 \text{ MeV}/c^2$  in  $\delta m$  at the center of the rectangle. It is assumed that the combinatorial background can be described by a smooth 2D quadratic function and the  $\Lambda_c^+$  background can be described as the product of a signal peak in  $m(\Lambda_c^+)$  and a smooth 1D quadratic function in  $\delta m$  within the rectangle<sup>①</sup>. Then the 2D sidebands are used to make an analytic estimate of the expected background in the small signal window along with its statistical uncertainty. A detailed description of the procedure is available in Appendix D along with a diagram showing the 2D windows.

The second method, referred to as “1D Fit & Count”, involves a 1D fit to the  $\delta m$  spectrum across a wide range, from threshold at  $\delta m = 0$  up to an upper bound of  $\delta m = 1500 \text{ MeV}/c^2$  but excluding the signal window of width  $20 \text{ MeV}/c^2$ . The fit function is then integrated across the signal window to give the expected background. The function used is a two-sided Landau distribution, which is defined as

$$f(\delta m) = \begin{cases} L(\delta m; \mu, \sigma_L) & \delta m \leq \mu \\ aL(\delta m; \mu, \sigma_R) & \delta m \geq \mu \end{cases} \quad (3-10)$$

where  $L(\delta m; \mu, \sigma)$  is a Landau distribution and  $a$  is chosen such that  $L(\delta m; \mu, \sigma_L) = aL(\delta m; \mu, \sigma_R)$ . A fit to data in the  $m(\Lambda_c^+)$  sidebands is shown in Figure 3.18.

The final results will be quoted with the 25 Tiles method, which is considered to be more robust because it does not require us to guess the shape of the background in the blinded region before unblinding.

① In effect, we are making a Taylor expansion of the background. If you make your extrapolation window narrow enough, the background is linear; the wider the range you have to describe, the higher the order of polynomial required. Toy MC tests with background models shows that a first-order polynomial is not sufficient and lead to a bias but second-order polynomial is applicable.

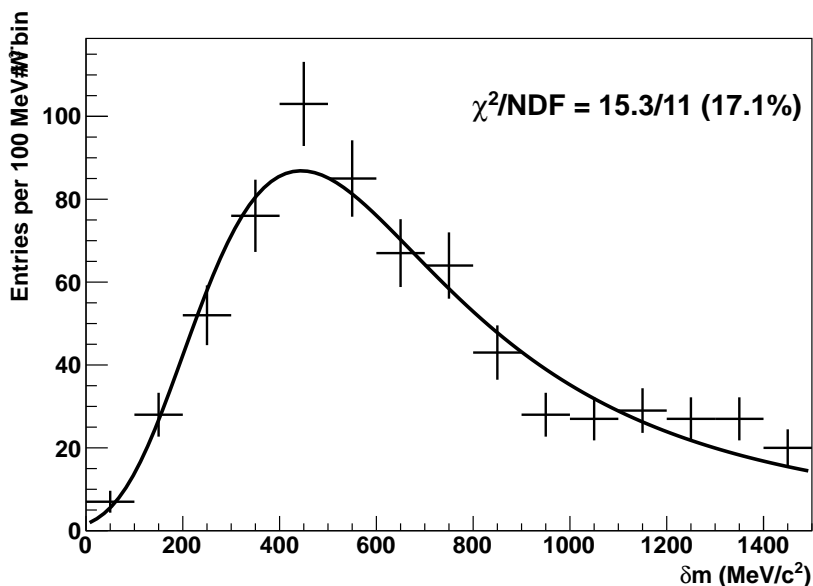


Figure 3.18 Fit of a two-sided Landau distribution to the  $\delta m$  spectrum on data in the  $m(\Lambda_c^+)$  sidebands. All candidates in the region 2248–2328  $\text{MeV}/c^2$  are excluded. In this fit:  $\mu = 484 \pm 33 \text{ MeV}/c^2$ ,  $\sigma_L = 177 \pm 18 \text{ MeV}/c^2$ ,  $\sigma_R = 170 \pm 19 \text{ MeV}/c^2$ .

### 3.5.3.2 Validation using toy MC

Before applying these two methods to data, toy MC studies are performed to ensure that the yield measurement procedure is unbiased and reports correct statistical uncertainties. In the MC studies the data is assumed to consist of three components:

- Signal: Gaussian in both  $m(\Lambda_c^+)$  and  $\delta m$ ,
- $\Lambda_c^+$  combinatorial: Gaussian in  $m(\Lambda_c^+)$  and smooth in  $\delta m$ ,
- $\Xi_{cc}^+$  combinatorial: Smooth in both  $m(\Lambda_c^+)$  and  $\delta m$ .

The  $m(\Lambda_c^+)$  shapes are taken from the fit to the inclusive  $m(\Lambda_c^+)$  spectrum in the  $\delta m$  sidebands in data, with the peak described by a single Gaussian and the background described by a first-order polynomial. The  $\delta m$  shape for signal is taken to be a single Gaussian, and that for background, both the  $\Lambda_c^+$  combinatorial and the  $\Xi_{cc}^+$  combinatorial, is a two-sided Landau distribution fitted to the  $\delta m$  spectrum in the  $m(\Lambda_c^+)$  sidebands on data, as shown in Figure 3.18. Note here the  $\delta m$  distributions for the  $\Lambda_c^+$  combinatorial and the  $\Xi_{cc}^+$  combinatorics components are assumed to be the same, since we cannot check the  $\Lambda_c^+$  combinatorial background shape in  $\delta m$  without unblinding.

A toy experiment consists of the following steps:

1. Choose expected background yields to be comparable to those expected in data

using an interpolation from sidebands.

2. Choose expected signal yield, which is set to zero in the toy studies.
3. Generate background data according to the distributions described above, with the number of background events of each category Poisson-fluctuated around the expected value.
4. Generate signal data according to the distribution described above, with the number of signal events Poisson-fluctuated around the expected value.
5. Scan over the allowed range in  $\delta m$ . For each  $\delta m$  value, run the yield extraction procedure and record the p-value for consistency with the null hypothesis. (See below for how this is defined.)
6. Record the results at each step, as well as the average p-value and the most significant, *i.e.* smallest, p-value seen in the toy.

The whole blinded range 380–880 MeV/ $c^2$  is scanned through using a  $\delta m$  step of 25 MeV/ $c^2$ , *i.e.* 21 points are tested for each toy experiment, and for each point 1000 toys are carried out. An average p-value across all steps and toys of 0.504 is found for the 25-Tiles method and 0.511 for the 1D Fit & Count method.

Due to the difficulty in the error estimate, the local p-values are not obtained directly. This however is not important in the end since the local p-values are not used for upper limit settings and is only used indirectly for the look-elsewhere-effect-corrected (LEE-corrected) significance (see Section 3.5.3.3: the final LEE-corrected p-values will be set using ensembles of toys). Nonetheless, it is worth explaining how it is calculated. The local p-value  $p_{\text{local}}$  is defined via the significance  $s$  as follows:

$$s = \frac{y_{\text{tot}} - y_{\text{bkg}}}{\sqrt{\sigma_{\text{tot}}^2 + \sigma_{\text{bkg}}^2}} \quad (3-11)$$

$$p_{\text{local}} = \frac{1}{2} \left( 1 - \text{Erf}(s/\sqrt{2}) \right) \quad (3-12)$$

where  $y_{\text{tot}}$  is the observed yield in the signal box,  $y_{\text{bkg}}$  is the expected background yield estimated from sidebands,  $\sigma_{\text{tot}}$  is the uncertainty on the observed yield, and  $\sigma_{\text{bkg}}$  is the uncertainty on the estimate of the expected background yield. The problem is caused by  $\sigma_{\text{tot}}$ , and more specifically by how it is correlated with the signal yield. It cannot be simply taken as  $\sqrt{y_{\text{tot}}}$ , otherwise the p-values become biased towards 1: upward fluctuations in the signal box yield get bigger uncertainties (so the significance is reduced) and downward fluctuations in the signal box yield get smaller uncertainties (so the

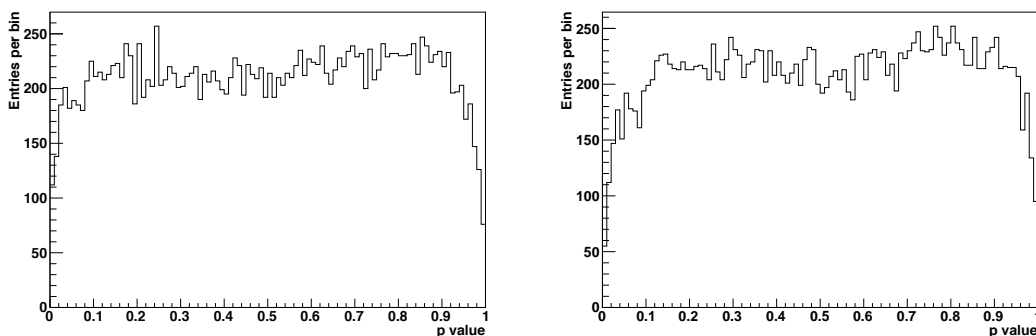


Figure 3.19 Distribution of local p-values obtained as described in Section 3.5.3.2 for the 25-Tiles method (left) and the 1D Fit & Count method (right) using a sample of 1000 toy experiments, taking 21 points per toy. Note that these are not the final corrected p-values.

significance is again reduced). Conversely, even under the null hypothesis it cannot be taken as  $\sqrt{y_{\text{bkg}}}$ , otherwise the p-values become biased towards 0: downward fluctuations in the sidebands reduce the expected background, giving a bigger signal excess with a smaller uncertainty (vs. a smaller signal with a bigger uncertainty for upward sideband fluctuations). A compromise solution is

$$\sigma_{\text{tot}} = \max(\sqrt{y_{\text{tot}}}, \sqrt{y_{\text{bkg}}}) \quad (3-13)$$

which tends to over-estimate the errors for large fluctuations in either direction but in a reasonably symmetric way. The resulting local p-value distributions are shown in Figure 3.19.

### 3.5.3.3 Look-Elsewhere Effect

Although the expected signal yield is close to zero after the whole selection criteria, the significance calculation procedure should still be established to ensure that every situation is handled. Since the signal is searched in a wide mass range, which increases the probability of observing small p-value fluctuations, the significance of the structure observed in data, if any, needs to take the LEE into account. The procedure is described below:

1. Run  $N$  toy experiments as defined in Section 3.5.3.2 with zero signal yield. In each toy experiment we do 501 steps of  $1 \text{ MeV}/c^2$  as in data. The smallest p-value observed in each toy experiment is recorded.
2. For any per-test p-value  $p_{\text{test}}$ , the number of toys that has a per-test p-value equal to

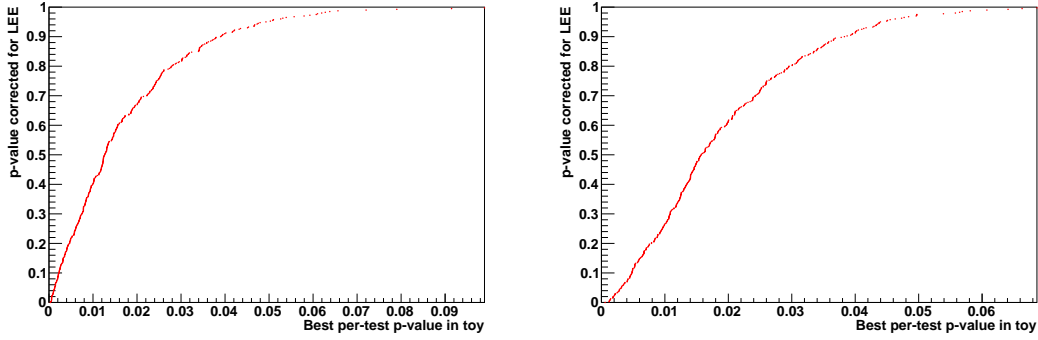


Figure 3.20 Example of LEE corrections derived from toy studies, for the 25-Tiles method (left) and the 1D Fit & Count method (right). Each is made with a sample of 400 toys.

or smaller than  $p_{\text{test}}$  is computed, which is called  $n(p_{\text{test}})$ .

3. If the lowest per-test p-value observed in the real data experiment is  $p_{\text{data}}$ , then the corresponding LEE-corrected p-value is  $n(p_{\text{data}})/N$ .

If every step is completely independent, the LEE correction would roughly equal to a factor of 501. But since the measurements are highly correlated—especially for the 1D Fit & Count method, where essentially the same background sample is used for every step, the correction does not equal to a simple factor. The easiest way to estimate it is toy MC. LEE correction curves are shown in Figure 3.20.

Note that the efficiency correction and its associated systematic uncertainty are not needed to check for the existence of a peak.

### 3.6 Decays through an intermediate $\Sigma_c$ baryon

So far only the case, in which the  $\Xi_{cc}^+$  baryon decays as a pure 3-body model, is considered. However, the  $\Xi_{cc}^+$  baryon may also decay through an intermediate resonance such as  $\Xi_{cc}^+ \rightarrow \Sigma_c^{++} K^-$ ,  $\Sigma_c^{++} \rightarrow \Lambda_c^+ \pi^+$ . Requiring an intermediate  $\Sigma_c(2455)$  or  $\Sigma_c(2520)$  baryon reduces the BF, but it also strongly suppresses the background.

The mass and natural width parameters of the  $\Sigma_c(2455)$  and  $\Sigma_c(2520)$  baryons in PDG<sup>[4]</sup> are listed below:

$$m(\Sigma_c^{++}(2455)) - m(\Lambda_c^+) = 167.52 \pm 0.08 \text{ MeV}/c^2$$

$$m(\Sigma_c^{++}(2520)) - m(\Lambda_c^+) = 231.4 \pm 0.6 \text{ MeV}/c^2$$

$$\Gamma(\Sigma_c^{++}(2455)) = 2.26 \pm 0.25 \text{ MeV}$$

$$\Gamma(\Sigma_c^{++}(2520)) = 14.9 \pm 1.5 \text{ MeV}$$



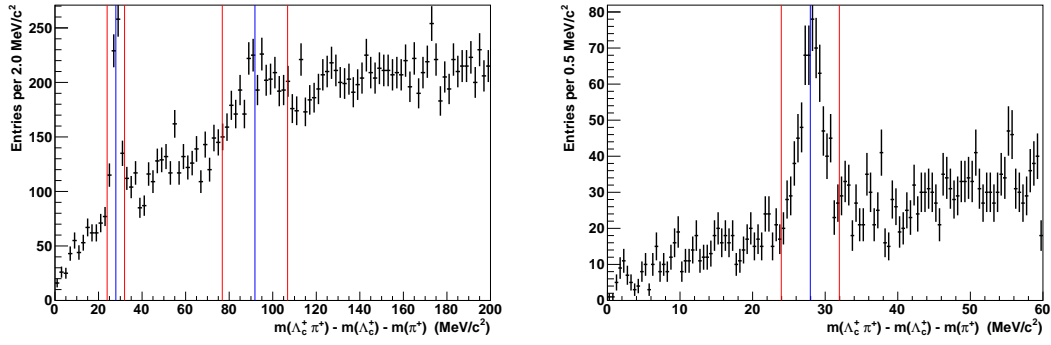


Figure 3.21 Inclusive  $m(\Lambda_c^+\pi^+) - m(\Lambda_c^+) - m(\pi^+)$  distribution in data close to threshold. A cut of  $2755 < m(\Lambda_c^+) < 2300$  MeV/ $c^2$  has been applied. The blue vertical lines indicate the nominal mass difference for the  $\Sigma_c^{++}(2455)$  and the  $\Sigma_c^{++}(2520)$ , and the red lines show the windows that will be used to select these states. The two plots show the same data, but over different  $x$ -axis ranges.

Taking  $m(\pi^+) = 139.57$  MeV/ $c^2$ , this corresponds to:

$$\begin{aligned} m(\Sigma_c^{++}(2455)) - m(\Lambda_c^+) - m(\pi^+) &= 27.95 \pm 0.08 \text{ MeV}/c^2, \\ m(\Sigma_c^{++}(2520)) - m(\Lambda_c^+) - m(\pi^+) &= 91.83 \pm 0.6 \text{ MeV}/c^2. \end{aligned}$$

Both the natural width and the expected experimental resolution are narrower for the  $\Sigma_c^{++}(2455)$  baryon than for the  $\Sigma_c^{++}(2520)$  baryon. For the former, a mass window of  $\pm 4.0$  MeV/ $c^2$  around the nominal mass difference defined above is used to capture nearly all the signal peak. For the latter, the mass window is taken to be  $\pm 15$  MeV/ $c^2$  (*i.e.*  $\pm \Gamma$ ). This is illustrated in Figure 3.21. After applying the appropriate mass window, the yield extraction study discussed above (and LEE correction <sup>①</sup>, if required) will be repeated. But note that the check of the contribution from the  $\Sigma_c$  resonances only serves as a cross-check. Separate upper limits for  $R$  on the  $\Sigma_c$  resonances will not be quoted.

① Note that for the  $\Sigma_c$  tests we'll make a LEE correction for testing different  $\Xi_{cc}^+$   $\delta m$  hypotheses as usual if a peak is seen, but we will not add a further correction for the fact that we check both 3-body and quasi-two-body decays.

### 3.7 Systematic uncertainties

In this section the main systematic uncertainties affecting the results are described. They mainly arise from the efficiency ratio calculation, in which many assumptions have been made and data-MC differences have been ignored. To be more specific, MC does not well reproduce the tracking efficiency and IP  $\chi^2$ ; signals are assumed to decay through the phase space model; many assumptions are made in the PID calibration; etc.

#### 3.7.1 Tracking efficiency

As mentioned in Section 3.4.2.1, the tracking efficiency is not accurately described by MC, which causes inaccuracy in the efficiency ratio calculation. According to the method provided by the LHCb tracking group, the systematic uncertainty is split into three parts.

The first part is the data-MC tracking efficiency ratio. The uncertainty of the ratio in each  $p - \eta$  bin leads to an uncertainty of the overall efficiency ratio. We fit the distribution of the data-MC tracking efficiency ratio with a Gaussian function. This overall uncertainty can then be estimated by the mean of the fit divided by the Gaussian width. This source contributes an uncertainty of 2.1%.

The second part is due to the neglecting of kinematic correlations between daughter tracks in the calculation of the ratio in the previous part. We assign 0.7% correlated systematic uncertainty per track following the procedure developed in Ref.<sup>[115]</sup>. The  $\Lambda_c^+$  part of the uncertainty cancels in the ratio and only the two daughter tracks from  $\Xi_{cc}^+$  contribute since both the signal and control channel contain the  $\Lambda_c^+$  decay. The systematic uncertainty due to kinematics correlations is 1.4%.

The third part is caused by hadronic interactions of daughter tracks. The tracks used for signal reconstruction are long tracks, which leave hits in the downstream-most tracking stations. Daughter tracks having hadronic interactions with detector material are not reconstructed, which causes the loss of the signal decay. This source of inefficiency is not corrected by the data-MC ratio table, as the data-MC ratio table is obtained by a tag-and-probe method using  $J/\psi \rightarrow \mu^+\mu^-$ , which do not have hadronic interactions. Following the procedure of Ref.<sup>[115]</sup>, we assign uncertainty of 2% per hadron for the uncertainty on material interactions. The uncertainty is fully correlated between daughter tracks since as long as one of the tracks has the hadronic interaction, the event will not

be reconstructed. Again, the  $\Lambda_c^+$  part of the uncertainty is expected to cancel in the ratio. The systematic uncertainty due to hadronic interactions is 4%.

Combining the previous numbers in quadrature yields the systematic uncertainty due to tracking efficiency of 4.7%.

### 3.7.2 Stripping, cut-based offline selection and MLP efficiencies

The efficiencies of these three selections are evaluated using MC directly. Most variables are well reproduced by MC (the PID related variables are not considered here), with one notable exception: the IP  $\chi^2$ . Hence an estimation of the uncertainty needs to be made to account for this data-MC difference. LHCb tracking group developed a tool which smears the track parameters of MC to match the IP  $\chi^2$  distribution observed in data<sup>[120]</sup>. But unfortunately this tool will bias the vertex  $\chi^2$  distribution and therefore cannot be used as a part of the standard MC processing procedure. We will use different strategy to study this effect for (a) stripping + offline selection, and (b) the MVA efficiency.

For stripping and offline selection, we calculate the ratio of efficiencies between control and signal modes both with and without track smearing. The difference seen when using smearing is 6.6% and is taken as the systematic uncertainty estimate (Table 3.25). For the MLP, we cannot just smear the MC sample and recheck the efficiency because of the effect on the vertex  $\chi^2$  cut discussed above. Instead, we weight unsmearred test MC sample so that its IP  $\chi^2$  distribution matches that of the smeared sample, then apply the standard MLP with the standard requirement of  $\text{MLP} > 0.8$ . The difference in the efficiency before and after reweighting is found to be 6.7% and is taken as the systematic uncertainty (Figure 3.22).

Finally, the two systematic uncertainties described above are combined assuming full correlation since they stem from the same inaccuracy. The systematic uncertainty is estimated to be 13.3%.

### 3.7.3 PID calibration

The procedure of PID calibration is described in Section 3.4.3. Several assumptions have been made to extract the PID requirement efficiency, and they cause systematic uncertainty.

Table 3.25 Efficiency of the stripping and offline cuts (excluding PID) for MC samples with and without track smearing for IP  $\chi^2$ .

Sample	signal mode	control mode
standard	$(3.18 \pm 0.04) \times 10^{-3}$	$(1.22 \pm 0.01) \times 10^{-2}$
smearred	$(3.16 \pm 0.04) \times 10^{-3}$	$(1.29 \pm 0.01) \times 10^{-2}$

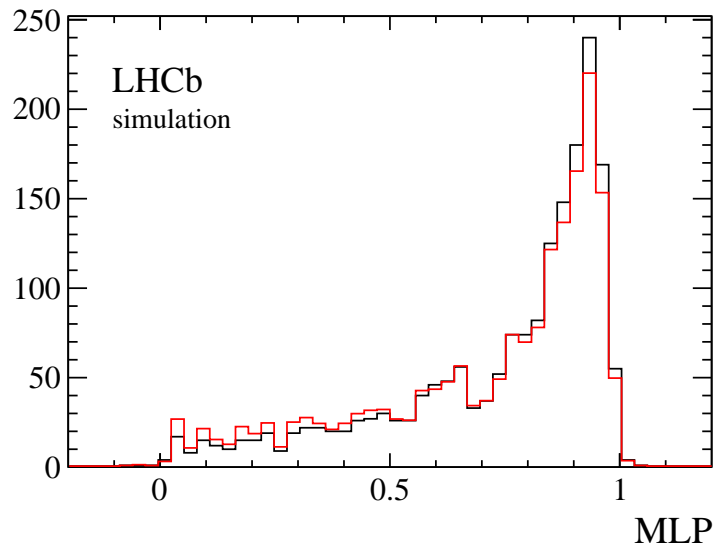


Figure 3.22 MLP distributions for the  $\Xi_{cc}^+$  MC sample before (black) and after (red) the IP  $\chi^2$  is weighted to match that of the smeared sample.

Table 3.26  $\Xi_{cc}^+$  MC: daughter tracks correlation check. The product of the individual efficiencies is 0.483 and the efficiency when applying all requirements simultaneously is 0.509.

Requirement	$\Lambda_c^+ p$ PIDp	$\Lambda_c^+ K$ PIDK	$\Lambda_c^+ \pi$ PIDK	$\Xi_{cc}^+ K$ PIDK	$\Xi_{cc}^+ \pi$ PIDK	events	efficiency
Sample0	-	-	-	-	-	7187	-
Sample1	> 10	-	-	-	-	6893	0.959
Sample2	-	> 10	-	-	-	6547	0.911
Sample3	-	-	< -5	-	-	5725	0.797
Sample4	-	-	-	> 10	-	6544	0.911
Sample5	-	-	-	-	< -5	5478	0.762
Sample6	> 10	> 10	< -5	> 10	< -5	3655	0.509

### 3.7.3.1 The calibration result

The PID efficiency ratio estimated with the calibration samples is  $1.75 \pm 0.03$ , so the quasi-statistical contribution to the systematic uncertainty is 1.95%.

### 3.7.3.2 Correlation between kinematics of daughter tracks

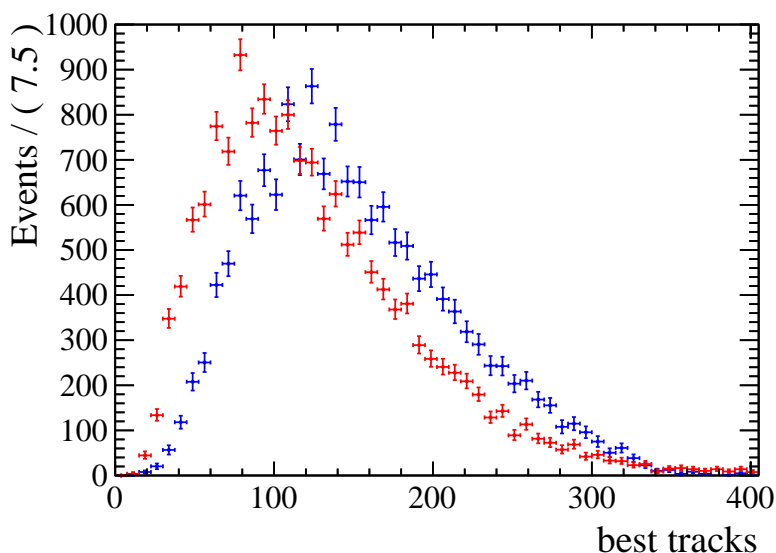
Since the daughter tracks share the energy of their mother, their kinematics is correlated, which causes correlation between the PID variables. However, the overall efficiency is calculated from individual efficiencies for daughter tracks, which assumes that the PID efficiencies factorize, and these correlations are ignored. To investigate the uncertainty from this assumption we compared the ratio calculated in the baseline way (assuming that the PID efficiencies factorize) to the ratio calculated by applying all PID requirements simultaneously using MC samples. The individual and simultaneous efficiencies are shown in Table 3.26 for  $\Xi_{cc}^+$  MC and Table 3.27 for  $\Lambda_c^+$  MC. The difference between the two approaches (ratio of 0.634/0.483 vs. ratio of 0.648/0.509) is found to be 3%, and is taken as the corresponding systematic uncertainty estimate.

### 3.7.3.3 MC multiplicity

As explained in Section 3.4.3, the PID efficiency depends on the RICH occupancy and hence on the event multiplicity. But the true  $\Xi_{cc}^+$  multiplicity distribution is unknown, so there is an intrinsic uncertainty in the PID calibration. In the baseline calculation of Section 3.4.3, we reweighted the  $\Xi_{cc}^+$  MC to match the multiplicity distribution of  $B_s^0$  data,

Table 3.27  $\Lambda_c^+$  MC: daughter tracks correlation check The product of the individual efficiencies is 0.634 and the efficiency when applying all requirements simultaneously is 0.648.

Requirement	$\Lambda_c^+ p$ PIDp	$\Lambda_c^+ K$ PIDK	$\Lambda_c^+ \pi$ PIDK	events	efficiency
Sample0	-	-	-	24689	-
Sample1	> 10	-	-	23234	0.941
Sample2	-	> 10	-	21786	0.882
Sample3	-	-	< -5	18870	0.764
Sample4	> 10	> 10	< -5	15999	0.648


 Figure 3.23 Comparison of best track distributions on MC (red) and on data (blue) for inclusive  $\Lambda_c^+$  baryons.

and the  $\Lambda_c^+$  MC to match that of  $\Lambda_c^+$  data. We now consider two other cases: first as above except that the  $\Xi_{cc}^+$  MC is reweighted to match  $\Lambda_c^+$  data, and second using unweighted MC, *i.e.* taking the multiplicity distributions from MC. For the first case the efficiency ratio becomes 1.740 (a shift of 0.5% from the baseline), and for the second case the ratio becomes 1.479 (a shift of 15% relative to the baseline). The results for the calibration are listed in Appendix A. Taking the mean of these two cases, we assign a systematic uncertainty of 7.8%.

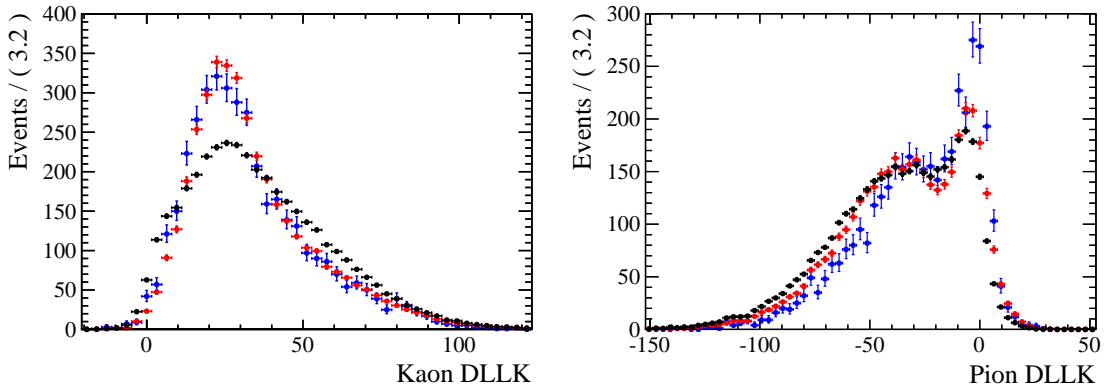


Figure 3.24  $DLL_{K\pi}$  distribution of kaons (left) and pions (right) for  $\Xi_{cc}^+$  MC (blue), weighted MC calibration sample (red), and unweighted MC calibration sample (black).

### 3.7.3.4 Calibration method

The PID calibration method assumes that the PID distribution only depends on momentum, pseudo-rapidity and the event multiplicity. However, this may not be the case, *i.e.* there may be other variables which also affect the PID distribution. Besides, the variable binnings used in calibration causes information loss and is another possible source of systematic uncertainty. Therefore, there are inherent uncertainties in the PID calibration procedure itself. Corresponding MC samples of standard calibration decays are used to calibrate the PID efficiency of  $\Xi_{cc}^+$  MC. Weighted distributions of the calibration sample are compared with true distributions from MC, and the efficiency difference can be quoted as the systematic uncertainty. Since the same calibration procedure is used for the signal and control modes, systematic uncertainties associated with the  $\Lambda_c^+$  part of the decay cancel and therefore only the effect associated with the bachelor  $K$  and  $\pi$  mesons contribute.

The the weighting procedure behaves quite well for the kaon, but does not work so nicely for the pion, as shown in Figure 3.24. This leads to a modest difference in the ratio of 2.0% for the kaon and a larger difference of 7.7% for the pion. Combining these in quadrature a corresponding systematic uncertainty of 8.0% is obtained.

### 3.7.3.5 Summary for PID efficiency

Combining uncertainties related to the PID efficiency in quadrature, the total uncertainty from PID calibration is 11.8%, which is dominated by the best track multiplicity uncertainty.

### 3.7.4 Trigger efficiency

The HLT reconstruction is similar to that used offline, so we treat the HLT as well modeled in the MC given the existing systematic uncertainties on tracking etc. However, L0 is rather different and is known not to be well reproduced. We corrected for this in Section 3.4.5.1. However, there are uncertainties associated with this correction, namely errors on the efficiencies in the look-up tables and the limited size of our signal MC sample. The efficiencies in the table have small uncertainties ( $\sim 0.1\%$ ) so the associated systematic uncertainty is negligible. The size of the MC sample does play a role. Since the size of our  $\Lambda_c^+$  MC sample is much larger than that of the  $\Xi_{cc}^+$  MC, the latter will dominate the corresponding uncertainty. The method used to calculate the systematic uncertainty can be found in Appendix B. Applying this method yields an uncertainty of 3.3%.

### 3.7.5 Yield determination

The number of signal decays is obtained from the background subtraction (see Section 3.5.3.1), therefore the method used for background estimate and the efficiency of the signal window definition can cause systematic uncertainty.

Two different methods are used for the background estimate and they have very similar outputs, so we do not assign any further systematic uncertainty due to the background estimate method. Inaccurate estimate of the  $\delta m$  resolution could cause incorrect estimate of the signal window efficiency. To assess the difference between the mass resolution in data and MC, the  $\Lambda_c^+$  data and MC samples are both fitted with a double-Gaussian function, and the weighted resolution is found to be very close (5.11 MeV/ $c^2$  in MC and 5.75 MeV/ $c^2$  in data). Hence the scale constant for the mass resolution is 1.12. We apply the constant to the  $\delta m$  resolution and find the efficiency difference of about 2%.

The fitting model of the control mode could also cause an uncertainty, which is estimated by fitting the signal with sum of two Crystal Ball functions. The variation in the signal yield is 2.7%, and this is taken as the systematic uncertainty estimate.

### 3.7.6 Decay models

The signal MC is generated with both the  $\Xi_{cc}^+$  baryon and the  $\Lambda_c^+$  baryon decaying according to a phase-space distribution. This is unlikely to be realistic since intermediate



Table 3.28 Resonance structures in the  $\Lambda_c^+ \rightarrow pK^-\pi^+$  decay<sup>[4]</sup>.

Decay modes	Branching fraction
$p \bar{K}^{*0}$	$(1.6 \pm 0.5)\%$
$\Delta^{++} K^-$	$(8.6 \pm 3.0) \times 10^{-3}$
$\Lambda(1520)\pi^+$	$(1.8 \pm 0.6)\%$
$p K^- \pi^+$ nonresonant	$(2.8 \pm 0.8)\%$
Total	$(5.0 \pm 1.3)\%$

resonance structures can be present in the  $\Xi_{cc}^+$  and the  $\Lambda_c^+$  decays. For the  $\Lambda_c^+$  baryon we do have some information on what resonances are present but no proper amplitude model, as shown in Table 3.28. Note that about half of the BF is considered to be non-resonant.

For the  $\Xi_{cc}^+$  baryon neither experimental data nor theoretical predictions are available. However, we can make reasonable speculations about what known resonances might be present. The list is quite short:  $\Sigma_c(2455)^{++} \rightarrow \Lambda_c^+\pi^+$ ,  $\Sigma_c(2520)^{++} \rightarrow \Lambda_c^+\pi^+$ ,  $\Sigma_c(2800)^{++} \rightarrow \Lambda_c^+\pi^+$ ,  $K^*(892)^0 \rightarrow K^-\pi^+$ . There is also one resonance that has been seen to decay to  $\Lambda_c^+K^-$ , namely the  $\Xi_c(2930)$  baryon. However, this is not considered to be confirmed in PDG and has not been seen in inclusive  $\Lambda_c^+K^-$ , so it is ignored here.

To check how the efficiency varies across the  $\Xi_{cc}^+$  Dalitz plot<sup>①</sup>, we define regions where resonances might be present, and compute the expected yield fractions in these regions using the proportion of the region to the whole Dalitz plot. The three regions defined are

- $\Sigma_c(2455)/\Sigma_c(2520)$  region:  $m(\Lambda_c^+\pi^+) < 2540 \text{ MeV}/c^2$
- $\Sigma_c(2800)$  region:  $2750 < m(\Lambda_c^+\pi^+) < 2850 \text{ MeV}/c^2$
- $K^*(892)^0$  region:  $846 < m(K^-\pi^+) < 946 \text{ MeV}/c^2$

However, in the end we are limited by MC statistics: only 23 signal MC events left after applying all the requirements. The efficiency variation within the Dalitz plot can not be analysed. Results of the expected and measured fractions are shown in Table 3.29. The point here is that within the big MC statistical uncertainty the measured fractions are compatible with the expected fractions, and also none of these resonant regions shows serious depletion.

① Strictly speaking it's not a real Dalitz plot, since the  $\Xi_{cc}^+$  and  $\Lambda_c^+$  are not pseudoscalars, but it still can be a useful approach.

Table 3.29 Signal MC event counts and fractions in various regions of the  $\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$  Dalitz plot after the complete selection, and after applying only the stripping and trigger requirements. The expected fractions are also shown assuming a pure phase-space distribution. (MC PID values are used without calibration/correction.)

Region	Complete selection		Stripping + trigger		Expected
Whole D.P.	23	100%	151	100%	100%
$\Sigma_c(2455)^{++}/\Sigma_c(2520)^{++}$	4	$(17 \pm 8)\%$	17	$(11 \pm 3)\%$	11%
$\Sigma_c(2800)^{++}$	10	$(43 \pm 10)\%$	43	$(28 \pm 4)\%$	22%
$K^*(892)^0$	9	$(39 \pm 10)\%$	41	$(27 \pm 4)\%$	21%

No additional uncertainty due to the decay models is assigned: we are fully limited by the signal MC sample size for which there is already a large systematic uncertainty assigned.

### 3.7.7 Summary

The uncertainties described in the previous sections are summarized in Table 3.30. The total uncertainty is found to be 26%. The dominate source of uncertainty is MC statistics, which can be reduced in future by using larger MC samples. Other large systematic uncertainties are caused by IP smearing and PID calibration. The former can be pushed down by a better description of detector response. For the latter, the main discrepancy is caused by the under-estimate of the event multiplicity in  $\Xi_{cc}^+$  MC. This can be reduced by a better fragmentation description of GENXICC, or by using PYTHIA to simulate the production of the  $\Xi_{cc}^+$  baryon. <sup>①</sup>

## 3.8 Variation of the efficiency ratio

The signal efficiency depends on the mass and lifetime of the  $\Xi_{cc}^+$  baryon. To reduce the model dependence as much as possible, the efficiency ratio is given in a large range

<sup>①</sup> It may sounds strange that the biggest source of the uncertainty is MC statistics. However, there is one reason why we decide not to use a much larger MC sample. The expected yield of the signal is consistent with zero, while there are about 10 background events in the signal window, which indicates that the statistical uncertainty is of order hundreds percent, and there will be very little gain to push the systematic uncertainty down. We have taken the systematic uncertainty halved within a simple toy, but the upper limit only goes down by 10%. Therefore, we decide not to use a larger MC sample.

Table 3.30 Summary of systematic uncertainties.

Source of uncertainty	$\sigma_R/R$
Tracking efficiency	4.7%
IP Smearing	13.3%
PID calibration	11.8%
Trigger efficiency	3.3%
Mass resolution difference in data and MC	2%
Mass fit of $\Lambda_c^+$ control sample	2.7%
MC statistics	18.3%
Total uncertainty	26.4%

of invariant mass for five lifetime hypotheses. As the mass and the lifetime are two independent variables, the efficiency ratio variations can be considered separately.

### 3.8.1 Variation of the efficiency ratio depending on the $\Xi_{cc}^+$ mass

The mass of the  $\Xi_{cc}^+$  baryon affects the efficiency in two ways. First, the efficiencies, especially the trigger and reconstruction efficiencies, rely on the kinematics of daughter tracks, and therefore rely on the  $\Xi_{cc}^+$  mass. Second, in the yield measurement the signal window is defined as the  $\pm 10$  MeV/ $c^2$  region around the  $\delta m$  value being tested. Since the  $\delta m$  resolution is a function of  $\delta m$ , this signal window cut efficiency also depends on the mass hypothesis. But as the mass of the  $\Xi_{cc}^+$  baryon is a priori unknown<sup>①</sup>, the efficiency ratio should be determined for each  $\delta m$  point in the the range 380 – 880 MeV/ $c^2$ . The treatment of the efficiency variation is described below.

#### 3.8.1.1 Effects from variation of kinematic distributions

The efficiency variation with kinematic distributions is investigated using generator-level MC samples with the weighting technique. The overall efficiency is expected to depend on the mass value linearly, therefore only two mass hypotheses are studied:  $m(\Xi_{cc}^+) = 3300$  MeV/ $c^2$ , 3700 MeV/ $c^2$ ; efficiencies of other mass hypotheses are obtained by a linear interpolation.

① The SELEX collaboration did measure the mass of the  $\Xi_{cc}^+$  baryon, but in the case of lack of confirmation, we still consider the  $\Xi_{cc}^+$  mass to be unknown.

For the mass hypothesis  $m$  being tested, two generator-level MC samples with  $m(\Xi_{cc}^+) = m$  and  $m(\Xi_{cc}^+) = 3500 \text{ MeV}/c^2$  are generated. A three-dimensional reweighting with the binnings shown in Table 3.31 is carried out on the  $p_T$  distributions of  $\Xi_{cc}^+$  daughters, *i.e.* the  $\Lambda_c^+$  baryons, the bachelor  $K$  meson, and the bachelor  $\pi$  meson. The weights obtained are used to reweight the  $p_T$  distributions of  $\Xi_{cc}^+$  daughters in the full MC sample generated with  $m(\Xi_{cc}^+) = 3500 \text{ MeV}/c^2$  to estimate the efficiency for the mass hypothesis  $m$ . The single event sensitivity (SES) for the mass hypothesis  $m$ ,  $\alpha_m$ , is calculated as

$$\alpha_m = \frac{\varepsilon_m}{\varepsilon_{3500}} \alpha_{3500}, \quad (3-14)$$

where  $\varepsilon_m$  and  $\varepsilon_{3500}$  are the efficiency for mass equal to  $m$  and  $3500 \text{ MeV}/c^2$ , respectively, and  $\alpha_{3500}$  is the SES for the normal sample with corrections from data. Here we assume that data/MC differences are independent of  $\Xi_{cc}^+$  mass hypotheses and therefore cancel in the ratio. Hence the PID and L0 calibration is not performed for other mass hypotheses.

 Table 3.31 Binning scheme for  $\Xi_{cc}^+$  daughters

Variable	Binning
$\Lambda_c^+ p_T$ (MeV/c)	[0, 700, 1000, 1450, 1900, 2600, 3300, 4300, 15000]
$K p_T$ (MeV/c)	[0, 150, 250, 350, 450, 600, 850, 1150, 5000]
$\pi p_T$ (MeV/c)	[0, 150, 250, 350, 450, 600, 850, 1150, 5000]

The weighted efficiencies are shown in Table 3.32 and the variation is shown in Figure 3.25. Each component of the efficiency is found to depend on  $m(\Xi_{cc}^+)$ , but these effects largely cancel out and the overall efficiency ratio only shows a weak dependence. Nonetheless, this variation is taken into account when setting upper limits.

### 3.8.1.2 Corrections corresponding to the $\Xi_{cc}^+$ mass window definition

As described in Section 3.5.3, the signal region is defined as a mass window of  $\pm 10 \text{ MeV}/c^2$  around the  $\delta m$  value being tested. This is equivalent to a mass window cut. The  $\delta m$  lineshape is measured in the signal MC for  $m(\Xi_{cc}^+) = 3500 \text{ MeV}/c^2$  (corresponding to  $\delta m = 580 \text{ MeV}/c^2$ ) and found to be described by the sum of two Gaussian functions, whose weighted average  $\sigma$  is  $4.36 \text{ MeV}/c^2$ . Integrating the fitted lineshape across the signal window we find that this corresponds to an efficiency of 96%.

Table 3.32 Acceptance efficiency for different mass hypotheses, and corresponding values of  $\alpha$ . Note that the errors are mainly driven by the limited full MC sample size and are highly correlated between different mass hypotheses. The detail of the calculation of the error can be found in Appendix C.

Mass (MeV/ $c^2$ )	3300	3500	3700
$\varepsilon_{\text{acc}}(\times 10^{-2})$	$18.25 \pm 0.08$	$17.67 \pm 0.08$	$17.09 \pm 0.08$
$\varepsilon_{\text{strip}}(\times 10^{-3})$	$2.914 \pm 0.038$	$3.168 \pm 0.037$	$3.316 \pm 0.041$
$\varepsilon_{\text{PID}}(\times 10^{-2})$	$50.18 \pm 0.65$	$51.47 \pm 0.59$	$51.68 \pm 0.64$
$\varepsilon_{\text{MLP}}(\times 10^{-2})$	$56.45 \pm 1.26$	$55.84 \pm 1.16$	$55.53 \pm 1.22$
$\varepsilon_{\text{L0}}(\times 10^{-2})$	$7.89 \pm 0.91$	$7.39 \pm 0.82$	$7.07 \pm 0.81$
$\varepsilon_{\text{HLT}}(\times 10^{-2})$	$31.52 \pm 5.62$	$30.26 \pm 5.27$	$31.55 \pm 5.58$
Ratio to 3500	$1.041 \pm 0.067$	1.000	$1.008 \pm 0.061$
$\alpha[\times 10^{-5}]$	$2.285 \pm 0.612$	$2.379 \pm 0.618$	$2.360 \pm 0.630$

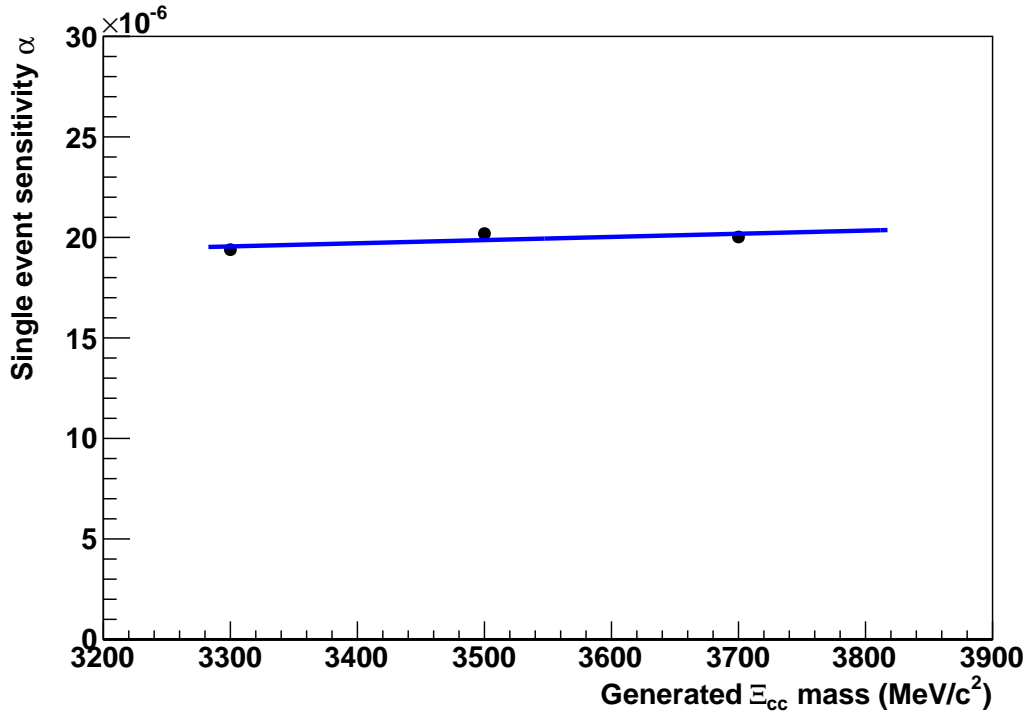


Figure 3.25 Variation of the single event sensitivity  $\alpha$  with  $m(\Xi_{cc}^+)$ . The points are obtained from a sample of full MC with  $m(\Xi_{cc}^+) = 3500$  MeV/ $c^2$  plus generator-level MC with  $m(\Xi_{cc}^+) = 3300$  and  $3700$  MeV/ $c^2$ . To obtain efficiencies at  $m(\Xi_{cc}^+) = 3300$  and  $3700$  MeV/ $c^2$ , the full MC with  $m(\Xi_{cc}^+) = 3500$  MeV/ $c^2$  is reweighted so that kinematic distributions of the final state daughters match those at a different  $\Xi_{cc}^+$  mass hypothesis as described in the text.

The mass resolution depends linearly on  $\delta m$  over the energy range we test. Assuming that the resolution of the lineshape does scale linearly with  $\delta m$ , the efficiency of this signal window cut at any point considered can be calculated. The efficiency evolves from 99% for  $\delta m = 380 \text{ MeV}/c^2$  to 87% for  $\delta m = 880 \text{ MeV}/c^2$ . Together with the kinematic dependence, the overall  $\delta m$ -dependent correction to  $\alpha$  is obtained.

### 3.8.2 Variation of the efficiency ratio depending on the $\Xi_{cc}^+$ lifetime

As a result of lifetime-related cuts in the trigger and offline selection, the efficiency as well as the production ratio  $R$  strongly depend on the lifetime hypothesis. However, the  $\Xi_{cc}^+$  lifetime is not accurately predicted but has a considerable theoretical uncertainty, as described in Chapter 1. To cover the whole prediction range, a discrete set of lifetimes (100 fs, 150 fs, 250 fs, 333 fs, 400 fs) is considered. We recompute the efficiency ratio and  $\alpha$ , and then quote upper limits on  $R$  for each value of the  $\Xi_{cc}^+$  lifetime <sup>①</sup>.

As the only full simulation MC sample is generated with a lifetime of 333 fs, a weighting technique is employed to estimate the efficiency for other lifetimes. For the lifetime hypothesis of  $\tau$ , the weight  $w$  is defined as:

$$w = \frac{\frac{1}{\tau} \exp^{-\frac{t}{\tau}}}{\frac{1}{\tau_0} \exp^{-\frac{t}{\tau_0}}},$$

where  $t$  is the decay time of the  $\Xi_{cc}^+$  in the event, and  $\tau_0$  is the lifetime in the current MC sample, *i.e.*  $\tau_0 = 333 \text{ fs}$ . The efficiency for a requirement is

$$\varepsilon = \frac{\sum_i^{\text{after requirement}} w}{\sum_i^{\text{before requirement}} w} = \frac{\sum_i^{\text{pass}} w_i}{\sum_i^{\text{pass}} w_i + \sum_j^{\text{fail}} w_j},$$

where the sum  $i$  runs over the events which pass the requirement and  $j$  runs over the events that fail to pass.

The statistical uncertainty of the efficiency is not trivial as the binomial distribution, but needs some subtle calculations:

$$\Delta\varepsilon = \frac{\sqrt{(\sum_j^{\text{fail}} w_j)^2 (\sum_i^{\text{pass}} w_i^2) + (\sum_i^{\text{pass}} w_i)^2 (\sum_j^{\text{fail}} w_j^2)}}{(\sum_i^{\text{pass}} w_i + \sum_j^{\text{fail}} w_j)^2}. \quad (3-15)$$

<sup>①</sup> The SELEX lifetime value, 33 fs, is not put in consideration. The reason is that, as we can find in the following text, the systematic uncertainty increases very sharply as we move towards smaller lifetime hypotheses. At the lifetime of 33 fs, the systematic uncertainty is so large that the upper limit on  $R$  does not contain meaningful information.

Table 3.33 Stripping and offline selection efficiency for different lifetime hypotheses

$\tau$ (fs)	$\sum_i^{\text{after cut}} w_i$	$\epsilon_{\text{sig}}^{\text{sel}}$	$\epsilon_{\text{con}}^{\text{sel/acc}} / \epsilon_{\text{sig}}^{\text{sel/acc}}$
100	543.44	$(2.40 \pm 0.07) \times 10^{-4}$	$50.70 \pm 1.41$
150	1546.43	$(6.84 \pm 0.12) \times 10^{-4}$	$17.81 \pm 0.33$
250	4461.78	$(1.97 \pm 0.02) \times 10^{-3}$	$6.17 \pm 0.09$
333	7187	$(3.18 \pm 0.04) \times 10^{-3}$	$3.84 \pm 0.05$
400	9363.05	$(4.14 \pm 0.05) \times 10^{-3}$	$2.94 \pm 0.04$

The selection criteria are optimized for the default lifetime hypothesis 333 fs, which may lead to a non-optimal selection for other lifetime hypotheses, but the important thing here is that the efficiency is evaluated correctly. It should be noted that we only have one single  $\delta m$  spectrum, therefore the measured yield for all lifetime hypotheses are the same. Each part of the efficiency ratio is considered in turn in the following.

### 3.8.2.1 Ratio of the acceptance efficiency

The generator-level cut requires that the initial four-momenta of all stable charged daughters are in the polar angle range  $10 \text{ mrad} < \theta < 400 \text{ mrad}$ . This is fully independent of the lifetime, hence the value of  $\epsilon_{\text{con}}^{\text{acc}} / \epsilon_{\text{sig}}^{\text{acc}}$  obtained in Section 3.4.1 applies to all lifetime hypotheses.

### 3.8.2.2 Ratio of the non-PID stripping and offline requirement efficiency

The acceptance requirement does not bias the lifetime distribution (at the first order), so (in the limit of large sample) the sum of weights before the stripping and offline selection are the same for all lifetime hypotheses, namely 2.26 M (the size of the MC sample generated for the default lifetime hypothesis). The sums of weights after the stripping and offline selection are given in Table 3.33, along with the resulting efficiency ratios. The ratio drops rapidly as the lifetime increases due to IP  $\chi^2$  related requirements in the selection.

### 3.8.2.3 Ratio of the PID efficiency

The PID calibration procedure is repeated for each lifetime hypothesis, using the lifetime-weighted MC sample. The ratio only weakly depends on the lifetime (see

Table 3.34).

Table 3.34 PID efficiency for different lifetime hypotheses

$\tau$ (fs)	$\mathcal{E}_{\text{con}}^{\text{PID sel}} / \mathcal{E}_{\text{sig}}^{\text{PID sel}}$
100	$1.680 \pm 0.034$
150	$1.714 \pm 0.034$
250	$1.734 \pm 0.034$
333	$1.748 \pm 0.034$
400	$1.757 \pm 0.034$

### 3.8.2.4 Ratio of MLP efficiency

To have the same  $\delta m$  spectrum, the same ANN (trained with the MC sample of  $\tau = 333$  fs) and the same requirement (MLP > 0.8) are applied for all lifetime hypotheses. The efficiency ratio is basically flat with lifetime, possibly with a gradual increase at very short lifetimes, as shown in Table 3.35.

### 3.8.2.5 Ratio of the L0 efficiency

The L0 efficiency is obtained from the trigger efficiency tables. Since the lifetime is now weighted, the efficiency calculation should be modified as

$$\bar{\varepsilon} = \frac{\sum_i w_i n_i \varepsilon_i}{\sum_i w_i n_i}, \quad (3-16)$$

where  $w_i$  is the lifetime weight,  $n_i$  is the number of events in the  $p_T$  bin, and  $\varepsilon_i$  is the probability that the event in that  $p_T$  bin is triggered. The statistical uncertainty on

Table 3.35 MLP efficiency for different lifetime hypotheses.

$\tau$ (fs)	$\mathcal{E}_{\text{sig}}^{\text{MLP}}$	$1/\mathcal{E}_{\text{sig}}^{\text{MLP}}$
100	$0.512 \pm 0.027$	$1.953 \pm 0.102$
150	$0.554 \pm 0.017$	$1.805 \pm 0.056$
250	$0.565 \pm 0.012$	$1.770 \pm 0.038$
333	$0.557 \pm 0.012$	$1.795 \pm 0.037$
400	$0.549 \pm 0.012$	$1.822 \pm 0.040$



Table 3.36 L0 efficiency for different lifetime hypotheses.

$\tau$ (fs)	$\epsilon_{\text{con}}^{\text{trig}} / \epsilon_{\text{sig}}^{\text{trig}}$
100	$0.937 \pm 0.003$
150	$1.161 \pm 0.002$
250	$1.358 \pm 0.002$
333	$1.451 \pm 0.002$
400	$1.507 \pm 0.002$

Table 3.37 HLT efficiencies for different lifetime hypotheses.

$\tau$ (fs)	HLT efficiency	$\epsilon_{\text{con}}^{\text{HLT}} / \epsilon_{\text{sig}}^{\text{HLT}}$
100	$0.107 \pm 0.046$	$2.636 \pm 1.134$
150	$0.187 \pm 0.050$	$1.503 \pm 0.403$
250	$0.268 \pm 0.051$	$1.053 \pm 0.203$
333	$0.303 \pm 0.053$	$0.931 \pm 0.165$
400	$0.323 \pm 0.055$	$0.872 \pm 0.153$

the L0 efficiency is quoted as the corresponding systematic uncertainty, which can be deduced from error propagation formula, as detailed in Appendix B. The ratio is shown in Table 3.36 and shows variation with lifetime. Note that this is not a feature of L0Hadron itself—the trigger is basically lifetime-unbiased—but rather due to the correlation with the applied selection criteria. At short lifetimes, only high-momentum signal survives the impact parameter and flight distance selection in the stripping; at longer lifetimes, lower-momentum signal which survive the stripping is probably rejected by L0Hadron. This is why the ratio is lower, *i.e.* the signal efficiency is higher at short lifetimes.

### 3.8.2.6 Ratio of the HLT efficiency

The HLT (HLT1 and HLT2) efficiency is calculated directly from MC. The statistical error are very large due to very limited MC sample size and the weighting procedure, as shown in Table 3.37. The trigger is clearly lifetime-biased in favor of long-lived  $\Xi_{cc}^+$  baryons.

Table 3.38 Systematic uncertainties for different lifetime hypotheses.

source of uncertainty	100 fs	150 fs	250 fs	333 fs	400 fs
Tracking efficiency	4.72%	4.72%	4.72%	4.72%	4.72%
IP smearing	13.32%	13.32%	13.32%	13.32%	13.32%
PID calibration	11.76%	11.76%	11.76%	11.76%	11.76%
L0 efficiency	12.70%	6.73%	3.89%	3.27%	3.03%
MC statistics	43.46%	27.10%	19.57%	18.02%	17.87%
Total uncertainty	48.86%	33.43%	27.12%	25.95%	25.81%

Table 3.39 SES  $\alpha$  for different lifetime hypotheses.

$\tau$ (fs)	Systematic uncertainty	$\alpha[\times 10^{-5}]$
100	48.86%	(60.04 $\pm$ 29.34)
150	33.43%	(14.05 $\pm$ 4.70)
250	27.13%	(3.96 $\pm$ 1.07)
333	25.95%	(2.38 $\pm$ 0.62)
400	25.81%	(1.81 $\pm$ 0.47)

### 3.8.2.7 Systematic uncertainties for different lifetime hypotheses

MC statistical uncertainty (including L0) is the only source of systematic uncertainty which depends on the  $\Xi_{cc}^+$  lifetime; others do not vary with lifetime at the first order. As a result, only the MC sample statistical uncertainty of  $\alpha$  is evaluated; all other systematic uncertainties are left unchanged. The overall systematic uncertainty is the quadratic sum of them. The systematic uncertainty for different lifetime hypotheses is summarized in Table 3.38.

### 3.8.2.8 SES for different lifetime hypotheses

Finally, putting all the above pieces together, the modified values of  $\alpha$  and their uncertainties for the five lifetime hypotheses are obtained, as listed in Table 3.40.

Table 3.40 SES  $\alpha$  for different lifetime hypotheses.

$\tau$ (fs)	$\alpha[\times 10^{-5}]$
100	$60.04 \pm 29.34$
150	$14.05 \pm 4.70$
250	$3.96 \pm 1.07$
333	$2.38 \pm 0.62$
400	$1.81 \pm 0.47$

### 3.9 Upper limit calculation

Since the expected signal yield is consistent with zero, upper limits (ULs) for the cross-section ratio  $R$  will need to be set, which is done using the  $CL_s$  method. If the expected signal and background events, and the SES  $\alpha$  are known very precisely, then the UL is easy to calculate. Define  $CL_{s+b}(n_{\text{obs}}, n_{\text{exp}})$  as the probability of observing a count smaller than or equal to  $n_{\text{obs}}$  under a signal-plus-background hypothesis with expected signal yield  $n_{\text{exp}}$ , and define  $CL_b(n_{\text{obs}})$  as the probability of seeing a count smaller than or equal to  $n_{\text{obs}}$  under the background-only hypothesis. Then  $CL_s$  is defined as:

$$CL_s(n_{\text{obs}}, n_{\text{exp}}) = \frac{CL_{s+b}(n_{\text{obs}}, n_{\text{exp}})}{CL_b(n_{\text{obs}})}.$$

The 95% confidence level UL on the yield is the value of  $n_{\text{exp}}$  for which  $1 - CL_s(n_{\text{obs}}, n_{\text{exp}}) = 0.95$  given the observed result  $n_{\text{obs}}$ . This calculation is simple to do: we adjust the expected yield until the sum over Poisson probabilities gives the desired result. Then the UL on  $R$  can be obtained from the UL on the yield multiplied by the SES.

However, since the expected background events and the SES are only known within certain uncertainties (including statistical and systematic uncertainties), the UL on  $R$  (from Eq. 3-3) cannot be derived from the simple calculation. Therefore, an MC-based evaluation is employed for the ULs setting. Assuming all the uncertainties are correctly evaluated, the procedure for UL settings is described below.

- Vary  $R$  in many small steps.
- For each value of  $R$ , we generate many random configurations to determine the corresponding  $CL_s$ , as described below.
- For each configuration, the value of  $R$  is the same but the values of  $\alpha$  and  $b$  are

fluctuated within their uncertainties <sup>①</sup>. This produces the expected signal yield  $s$ , which equals to  $R/\alpha$ , and the expected number of background events  $b$  for that configuration. Then random yields for signal and background are generated from the Poisson fluctuation of the expected yields. We then test whether the background-only and signal-plus-background hypotheses are less than or equal to  $n_{\text{obs}}$ .

- From the ensemble of configurations,  $CL_b$  is the fraction of background-only tests with yield below  $n_{\text{obs}}$ , and  $CL_{s+b}$  is the fraction of signal-plus-background tests with yield below  $n_{\text{obs}}$ . Those fractions have (binomial) statistical uncertainties due to the finite number of configurations generated.
- Then  $CL_s$  for this value of  $R$  is the ratio  $CL_{s+b}/CL_b$ . As before, there is a statistical uncertainty on that value due to the finite number of configurations generated.

In this way a curve of  $CL_s$  vs  $R$  is produced. The UL on  $R$  at the 95% CL is the value of  $R$  which gives  $CL_s = 0.05$  on that curve. An example of the scan running over toy data is shown in Figure 3.26. Note that for different mass and lifetime hypotheses, the SES  $\alpha$  is changed accordingly, while the signal yield is the same.

The above procedure will be repeated for 501 mass hypotheses, but these numbers would contain too much information to be interpreted. Therefore, the main results are quoted as follows (assuming a null result):

- Quote the largest, *i.e.* the worst, UL across the entire mass range, *i.e.* all 501 points in  $380 < \delta m < 880 \text{ MeV}/c^2$  in steps of  $1 \text{ MeV}/c^2$ .
- Divide the data into disjoint  $50 \text{ MeV}/c^2$  intervals and quote the largest UL in each interval. This is probably more useful for a theorist who has a model with a given mass for the state.

<sup>①</sup> There is a subtlety here: the uncertainties on  $\alpha$  can be quite large and therefore non-Gaussian. As we will see in the next Section, at  $\tau = 100 \text{ fs}$  the relative error on  $\alpha$  is a little above 50% due to the limited MC sample size at short decay times (see Section 3.8.2). So we cannot simply treat the errors as Gaussian otherwise we will have a negative efficiency in a few percent of cases. Instead the following algorithm is used to fluctuate  $\alpha$ : we generate a random number  $r$  according to a Gaussian distribution with  $\mu = 0$  and  $\sigma = 1$ . If  $r > 0$  we fluctuate  $\alpha$  upward by  $r\sigma_\alpha$  in the usual way. If  $r < 0$  we fluctuate  $\alpha$  downward, but we do this by fluctuating  $1/\alpha$  upward by  $-r\sigma_{1/\alpha}$  where  $\sigma_{1/\alpha} = \sigma_\alpha/\alpha^2$  from standard error propagation.

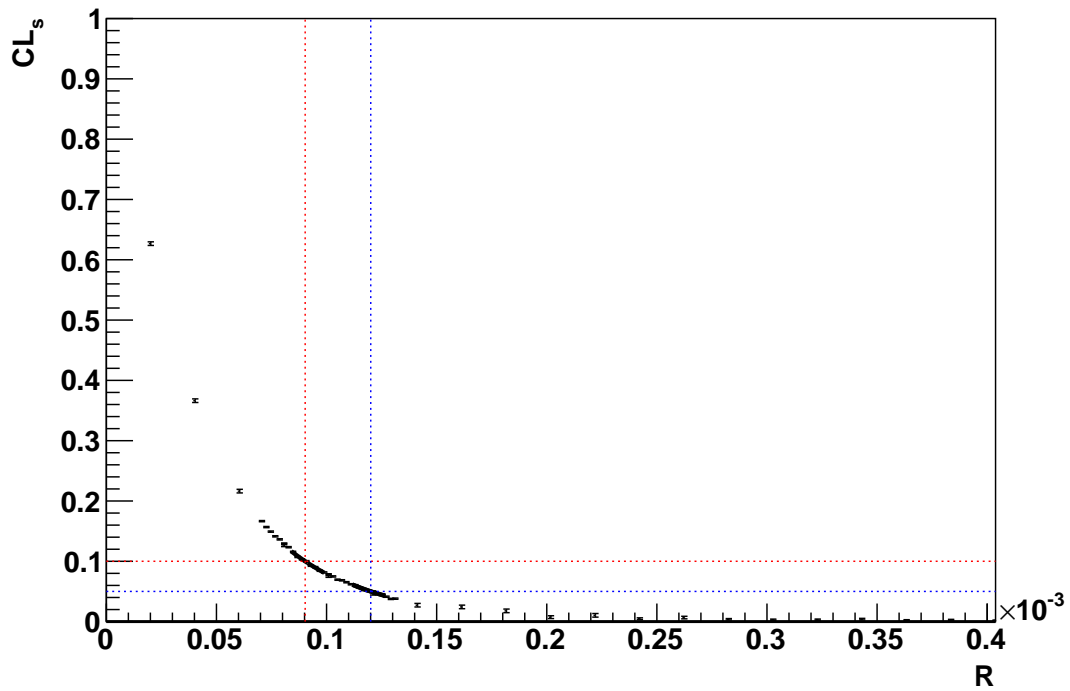


Figure 3.26 Example of the  $CL_s$  scan for the following input values:  $n_{\text{obs}} = 2$ ,  $b = 2.0 \pm 0.1$ ,  $\alpha = (2.019 \pm 0.622) \times 10^{-5}$ . The estimated  $CL_s$  ULs for these example parameters are  $R < 9.0 \times 10^{-5}$  at 90% CL,  $R < 12.0 \times 10^{-5}$  at 95% CL. Toy MC is used to obtain the points, and local exponential fits are used to interpolate between points in the vicinity of  $CL_s = 0.1$  and  $CL_s = 0.05$ . The error bars are due to finite MC sample size. Note that the density of points and their statistical uncertainties are not uniform across the plot.

### 3.10 Results

With the selection criteria and the procedures for the upper limit calculations defined, the signal region can be unblinded. The unblinded  $\delta m$  spectra is shown in Figure 3.27. No significant peak is found across the mass range  $380 < \delta m < 880 \text{ MeV}/c^2$ . Applying the yield measurement procedure described in Section 3.5.3, we obtain the signal yield spectra shown in Figure 3.28. The yields fluctuate around zero but do not show a large excess. The local significance and local p-values (before LEE correction) are shown in Figure 3.29 and 3.30. The largest local significance seen with the baseline 25-Tiles method is  $1.46 \sigma$ , and with the crosscheck 1D Fit & Count method it is  $2.16 \sigma$ . To account for the LEE, we use an ensemble of toy experiments as described in Section 3.5.3.3. In the ensemble we find that 99.1% of toys contain a smaller local 25-Tiles p-value than that seen in data, and 52.5% contain a smaller local 1D Fit & Count p-value than that seen in data (displayed in Figure 3.31). Thus, no significant excess above background is observed. The upper limits on the quantity  $R$ , defined in Eq. 3-3, are set following the procedure described in Section 3.9. The limits obtained are shown in Figure 3.32 and tabulated in Tables 3.41 and 3.42.

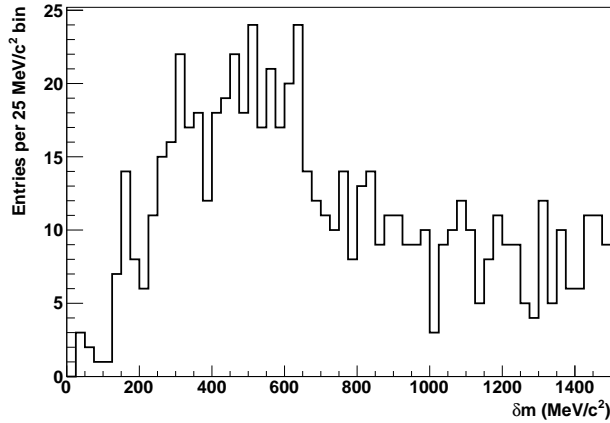


Figure 3.27 The  $\delta m$  spectrum of selected  $\Xi_{cc}^+$  candidates in the unblinded data set.

The mass spectrum after requiring an intermediate  $\Sigma_c$  resonance in the decay, as described in Section 3.6, is also checked. The resulting raw mass spectrum is shown in Figure 3.33. There is an interesting fluctuation around  $\delta m = 550 \text{ MeV}/c^2$  when requiring the presence of intermediate  $\Sigma_c(2455)^{++}$  baryon, but it is not statistically significant.

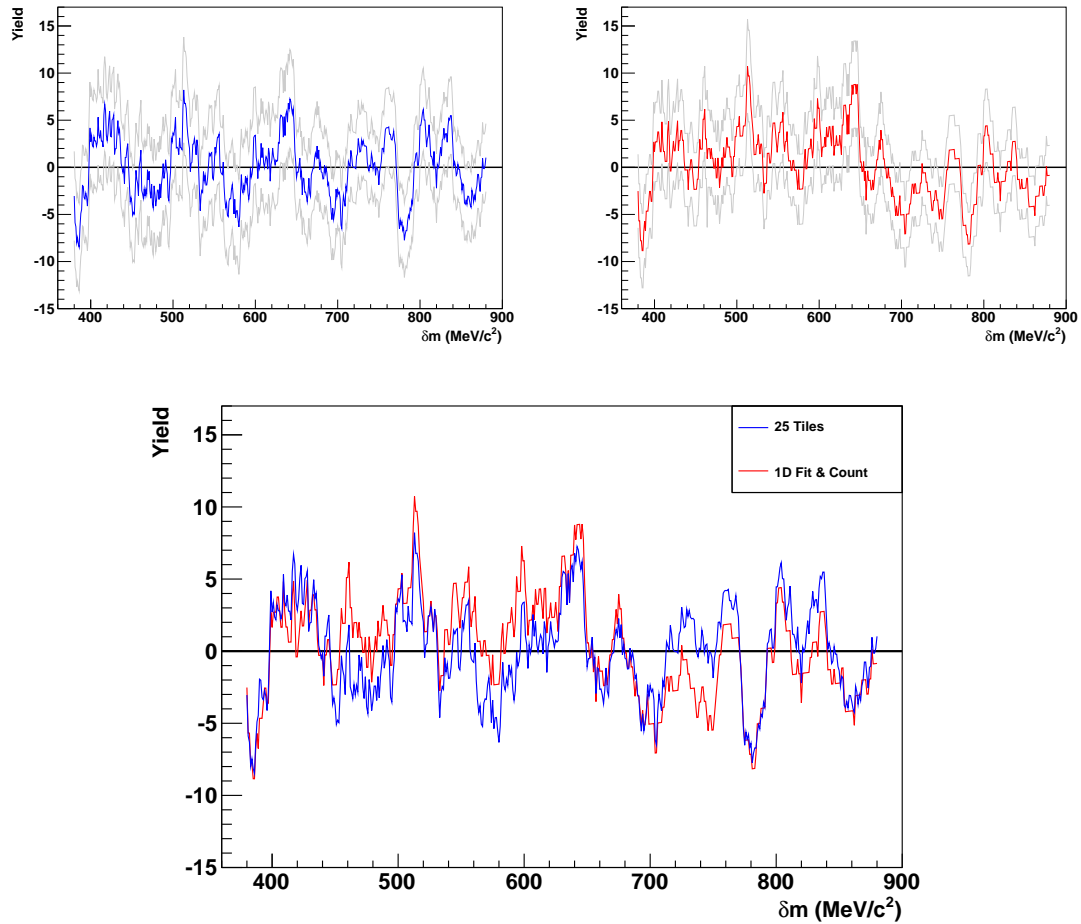


Figure 3.28 Measured signal yields as a function of  $\delta m$  in the unblinded data set. The upper row shows the estimated signal yield as a colored line and the  $\pm 1\sigma$  statistical uncertainty bands as grey lines for the baseline 25-Tiles method (upper left) and the crosscheck 1D Fit & Count method (upper right). The central values of the two methods are compared in the lower plot and found to agree well.

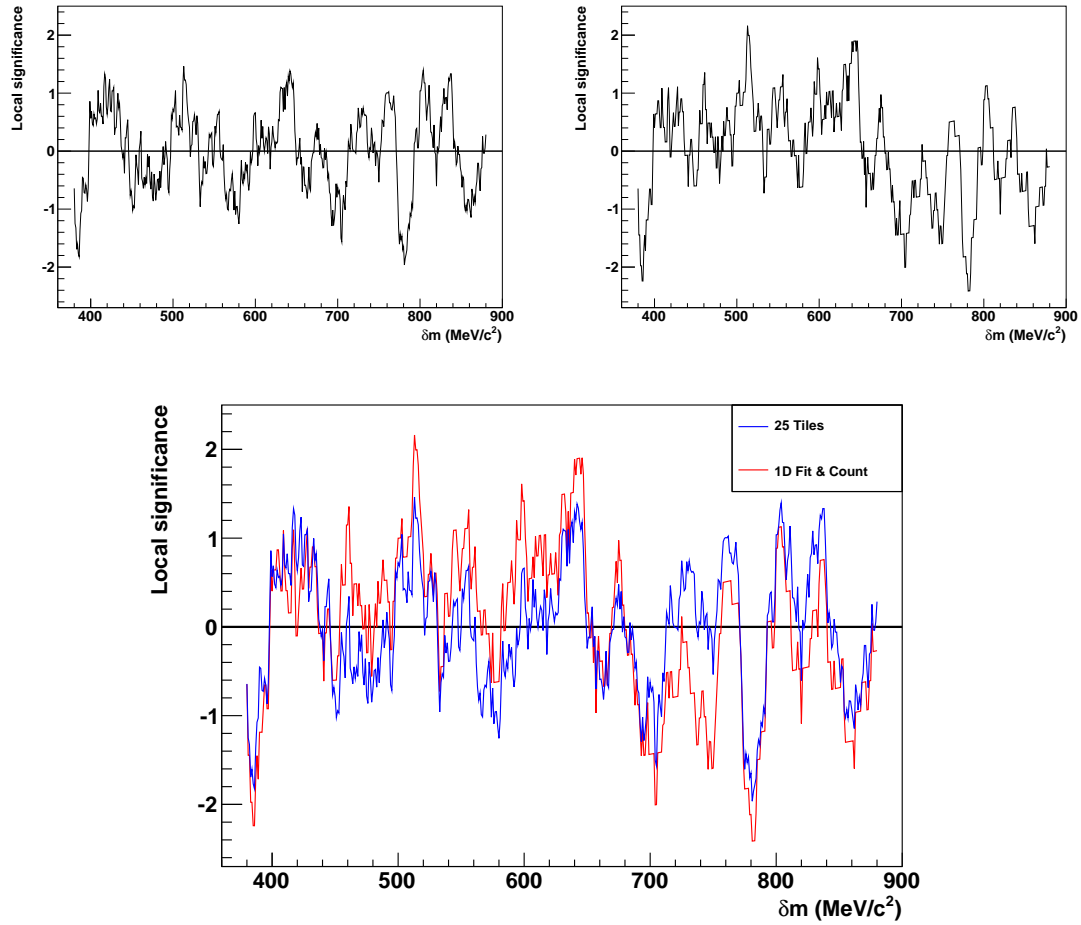


Figure 3.29 Local signal significance as a function of  $\delta m$  in the unblinded data set, for the baseline 25-Tiles method (upper left), and the crosscheck 1D Fit & Count method (upper right). The results from the two methods are compared in the lower plot.



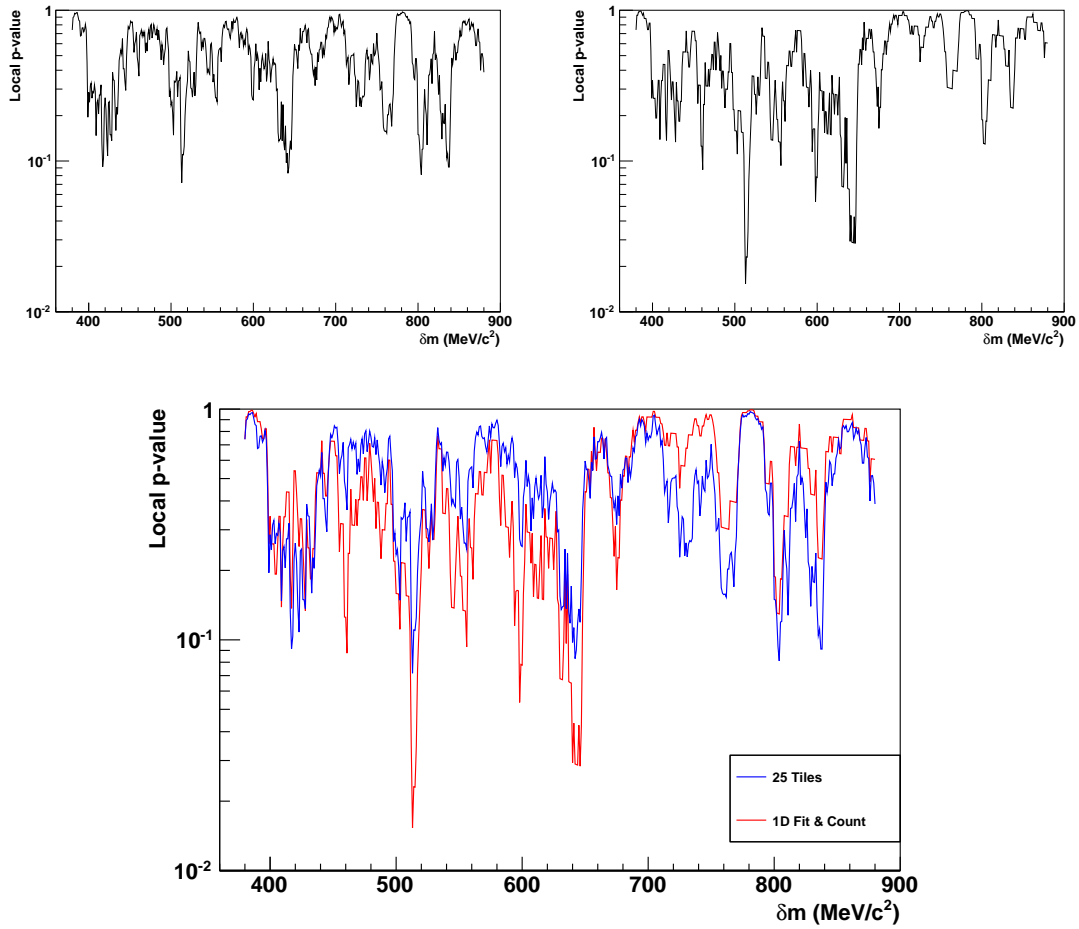


Figure 3.30 Local p-values as a function of  $\delta m$  in the unblinded data set, for the baseline 25-Tiles method (upper left), and the crosscheck 1D Fit & Count method (upper right). The results from the two methods are compared in the lower plot.

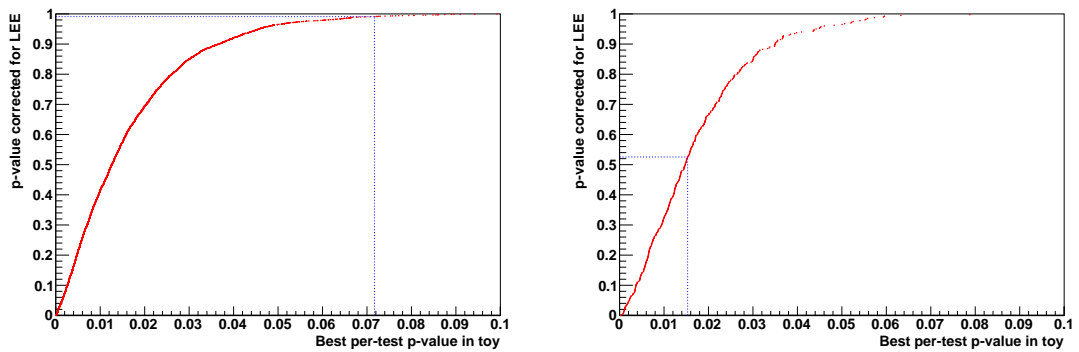


Figure 3.31 Look Elsewhere Effect correction for the baseline 25-Tiles method (left), and the crosscheck 1D Fit & Count method (right). The dotted lines indicate the smallest local p-value seen in the unblinded data.

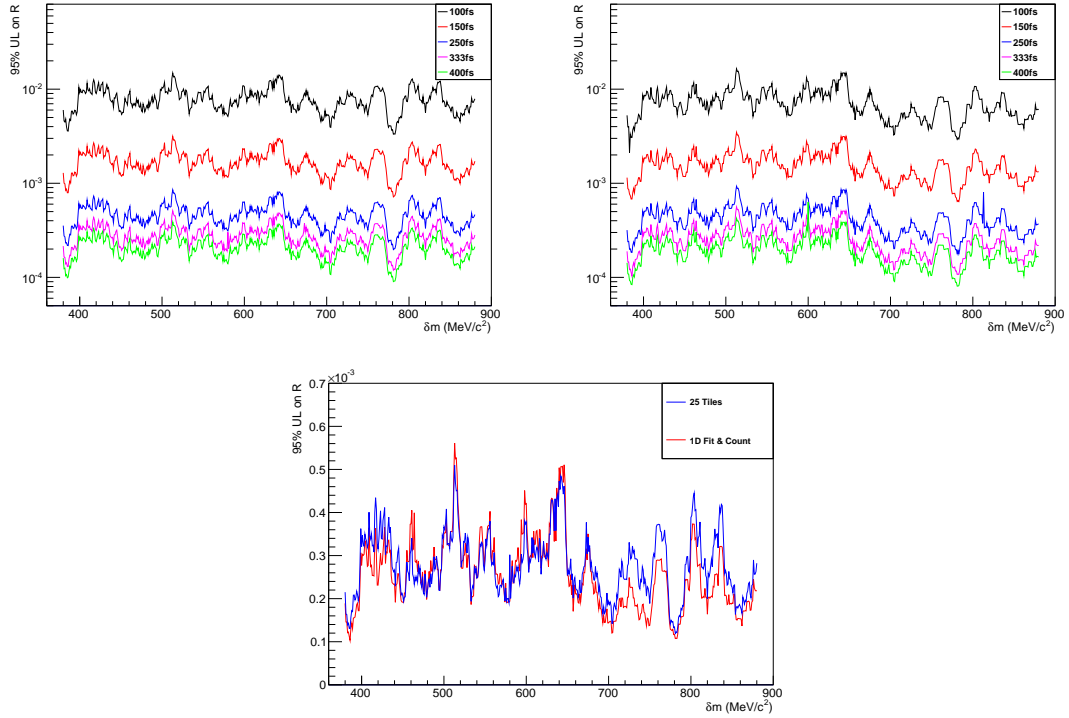


Figure 3.32 Upper limits on  $R$  at 95% CL, as a function of  $\delta m$ , for the baseline 25-Tiles method (upper left), and the crosscheck 1D Fit & Count method (upper right). For comparison, the limits obtained with the two methods are plotted together for the 333 fs lifetime hypothesis in the lower plot.

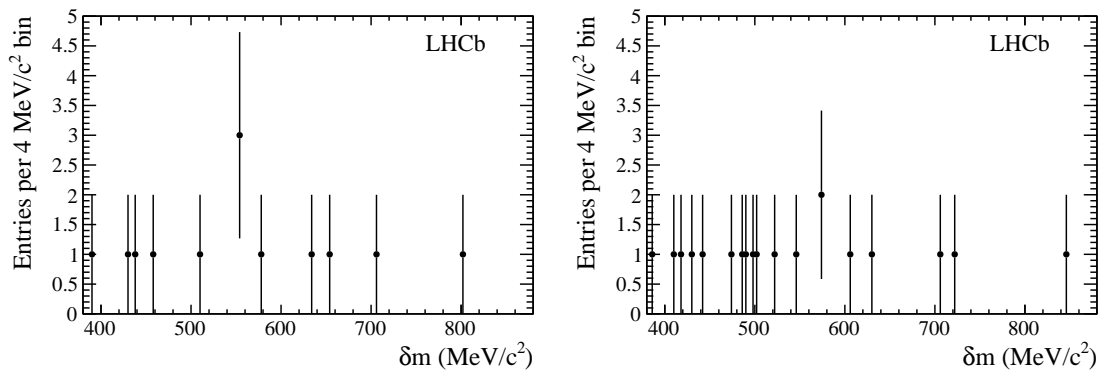


Figure 3.33 Raw  $\delta m$  spectra of selected  $\Xi_{cc}^+$  candidates in the unblinded data set, requiring an intermediate  $\Sigma_c(2455)^{++}$  (left) or  $\Sigma_c(2520)^{++}$  (right) resonance in the decay. A requirement of  $2273 < m(\Lambda_c^+) < 2303$  MeV/ $c^2$  has been applied.

Table 3.41 Upper limits on  $R$  at 95% CL in ranges of  $\delta m$  with the 25-Tiles method, for a set of lifetime hypotheses. The upper limits across the entire 500 MeV/ $c^2$  range are also shown.

$\delta m$ (MeV/ $c^2$ )	$R$ , largest 95% CL UL in range $\times 10^3$				
	100 fs	150 fs	250 fs	333 fs	400 fs
380–429	13	2.7	0.73	0.43	0.33
430–479	11	2.4	0.65	0.39	0.29
480–529	15	3.2	0.85	0.51	0.39
530–579	11	2.3	0.63	0.38	0.29
580–629	11	2.3	0.63	0.38	0.29
630–679	14	3.0	0.81	0.48	0.37
680–729	9.5	2.0	0.56	0.33	0.25
730–779	11	2.3	0.63	0.37	0.28
780–829	13	2.7	0.74	0.45	0.33
830–880	12	2.6	0.70	0.42	0.32
380–880	15	3.2	0.85	0.51	0.39

 Table 3.42 Upper limits on  $R$  at 95% CL in ranges of  $\delta m$  with the crosscheck 1D Fit & Count method, for a set of lifetime hypotheses. The upper limits across the entire 500 MeV/ $c^2$  range are also shown.

$\delta m$ (MeV/ $c^2$ )	$R$ , largest 95% CL UL in range $\times 10^3$				
	100 fs	150 fs	250 fs	333 fs	400 fs
380–429	11	2.3	0.61	0.37	0.28
430–479	12	2.5	0.67	0.41	0.32
480–529	16	3.5	0.93	0.56	0.42
530–579	12	2.5	0.67	0.40	0.31
580–629	13	2.8	0.76	0.45	0.63
630–679	15	3.2	0.86	0.51	0.39
680–729	7.0	1.5	0.41	0.25	0.19
730–779	8.3	1.8	0.49	0.29	0.22
780–829	11	2.3	0.81	0.37	0.28
830–880	9.2	2.0	0.54	0.32	0.24
380–880	16	3.5	0.93	0.56	0.63

## Chapter 4 Search for the $\Xi_{cc}^+$ baryon using 2011 and 2012 data

This chapter presents the search for the  $\Xi_{cc}^+$  baryon using  $pp$  collision data collected by LHCb in 2011 and 2012. Section 4.1 introduces the overall analysis strategy. A description of the data sample is given in Section 4.2. Selection criteria and expected sensitivity are discussed in Section 4.3.

### 4.1 Analysis strategy

The  $\Xi_{cc}^+$  baryon is searched through the decay  $\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$  with the  $\Lambda_c^+$  baryon reconstructed via  $\Lambda_c^+ \rightarrow p K^- \pi^+$ . With more data and a better design of the selection, there might be a possibility to observe the  $\Xi_{cc}^+$  baryon. Therefore, instead of measuring the relative production ratio, we suggest to focus on the existence of the  $\Xi_{cc}^+$  baryon—after all, this is the first important physical information from the  $\Xi_{cc}^+$  search. To avoid preconceptions that may bias the results, the analysis is performed using a blind approach. To have a better understanding of the background, wrong charge (WC) candidates, in which a  $\Lambda_c^+$  candidate is combined with a  $K^+$  track and a  $\pi^+$  track to form a  $\Xi_{cc}^+$  candidate, and the doubly Cabibbo-suppressed (DCS) decay, in which a  $\Lambda_c^+$  candidate combined with a  $K^+$  track and a  $\pi^-$  track to form a  $\Xi_{cc}^+$  candidate, are also selected to train the MVA and to estimate the lineshape of the background.

### 4.2 Data set

The search for the  $\Xi_{cc}^+$  baryon is performed using the  $pp$  collision data sample collected by LHCb in 2011 and 2012, at a center-of-mass energy of 7 TeV and 8 TeV and correspond to an integrated luminosity of  $0.65 \text{ fb}^{-1}$  and  $2.0 \text{ fb}^{-1}$ , respectively. The processing pass for data is *Reco14-Stripping20* and ntuples are made with DAVINCI v32r2p1.

To produce the MC sample, the GENXICC generator is used to generate  $\Xi_{cc}^+$  baryons. Unfortunately, due to a change in the LHCb simulation software, GAUSS, GENXICC was not compatible well with GAUSS. A quick fix has been provided, but it takes some time to incorporate the fix into the official GAUSS. Therefore, we do not have MC samples

which have the same processing procedures as the data at the time of writing this thesis. The old MC samples, which were produced for the analysis using 2011 data, have a reconstruction version *Reco12*, and are used to optimize the selection. These samples provide an appropriate guide for the efficiency calculations, although the statistics is fairly insufficient.

### 4.3 Selection

The signal decay is required to pass trigger requirements, stripping, cut-based offline selection and a multivariate selection. To be more sensitive to the low lifetime region, selection criteria that are harmful for a short lifetime  $\Xi_{cc}^+$  baryon, *e.g.* requirements on IP  $\chi^2$  for daughter tracks of the  $\Xi_{cc}^+$  baryon, should not be applied.

#### 4.3.1 Stripping and offline selection

Stripping selection aims at retaining signal events using loose criteria, while rejecting the most obvious background. To increase sensitivity to a low lifetime  $\Xi_{cc}^+$  baryon, the selection criteria that are only efficient for a long lifetime  $\Xi_{cc}^+$  baryon are removed or significantly relaxed at the stripping level. This drastically increases the output retention rate. To keep the rate at an acceptable level, candidates are required to TOS for inclusive charm trigger lines in stripping, *i.e.* the event should at least TOS on one of the HLT2 trigger lines: the  $D_{(s)}^+$  trigger line, the  $D^0$  trigger line and the  $\Lambda_c^+$  trigger line. The WC and DCS modes have the same selection criteria, taking into account that daughter tracks of  $\Xi_{cc}^+$  baryons are replaced by the corresponding antiparticle tracks, *e.g.* in the WC mode a  $K^+$  track are combined with a  $\Lambda_c^+$  candidate and a  $\pi^+$  track to form a  $\Xi_{cc}^+$  candidate.

In addition to the variables used in the analysis using 2011 data (Section 3.3.2), the following variables are also employed to further differentiate the signal and the background:

- the neural net particle probability, for a specific particle hypothesis, or ProbNN, which is the probability that the track has the identity of a given type. It is computed by an ANN, which is trained using the DLL, various track quality variables, and sub-detector acceptance flags<sup>[121]</sup>. It contains more information than the DLL and has a better performance;

- the ghost track probability, or ProbNNghost, which is the probability that the track is not a true particle but a ghost track, *i.e.* a mis-combination of detector hits that is recognized as a track by the reconstruction algorithm. The larger the ProbNNghost is, the greater is the probability that the track is a ghost track;
- the cosine of the helicity angle, or  $\cos\theta_H$ , where the helicity angle is the angle between the momentum vector of  $\Lambda_c^+$  in the  $\Xi_{cc}^+$  rest frame and the momentum vector of the  $\Xi_{cc}^+$  baryon in the laboratory frame. Figure 4.1 shows an example of  $\theta_H$ . The  $\cos\theta_H$  distribution is flat for the signal, while it is peaking close to 1 for the background;
- the cosine of the maximum projection angle, or  $\cos\alpha$ , where the projection angle is the angle between the projected momentum of two particles in the  $x$ - $y$  plane. Figure 4.1 shows an example of the projection angle. There are six projection angles since we have three final state tracks from  $\Lambda_c^+$  and the two final state tracks from  $\Xi_{cc}^+$  baryons, and  $\alpha$  is the smallest one of them. The signal tends to have a small  $\alpha$  since all the tracks are boosted in the same direction, while the background could have a large  $\alpha$  since in background events tracks in different directions are randomly combined.

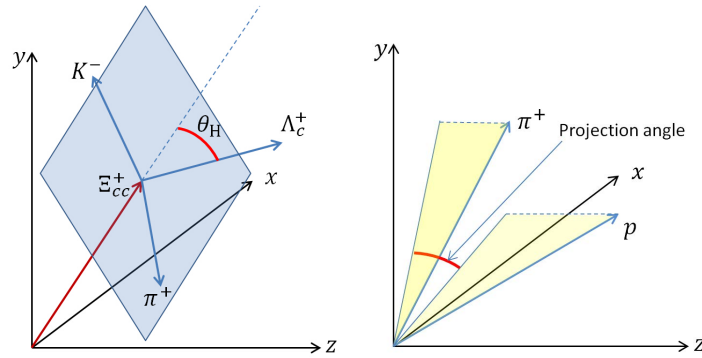


Figure 4.1 Example of the helicity angle (left) and the projection angle (right).

#### 4.3.1.1 Stripping

Stripping criteria for the signal mode, except for the inclusive charm hadron trigger requirements, are listed in Table 4.1.

Table 4.1 Selection criteria of the  $\Xi_{cc}^+$  decay at the stripping level.

Particle	Variable	Cut value
$\Xi_{cc}^+$	DIRA	$> 0$
	Vertex fit $\chi^2/\text{ndof}$	$< 10$
	$z$ separation between $\Xi_{cc}^+$ and $\Lambda_c^+$ vertices	$> 0.01$ mm
	Sum of daughters $p_T$	$> 2000$ MeV/ $c$
	Invariant mass	$< 4000$ MeV/ $c^2$
	MAXDOCA	$< 0.5$ mm
Tracks from $\Xi_{cc}^+$	$p$	$> 2000$ MeV/ $c$
	$p_T$	$> 250$ MeV/ $c$
	Track $\chi^2/\text{ndof}$	$< 5$
	Has RICH information	True
$\Lambda_c^+$	DIRA	$> 0.99$
	Vertex fit $\chi^2/\text{ndof}$	$< 10$
	Flight distance $\chi^2/\text{ndof}$	$> -1$
	MAXDOCA	$< 0.5$ mm
	Invariant mass	$2211 < M < 2361$ MeV/ $c^2$
	Sum of daughters $p_T$	$> 1000$ MeV/ $c$
Tracks from $\Lambda_c^+$	$p$	$> 2000$ MeV/ $c$
	$p_T$	$> 250$ MeV/ $c$
	HASRICH	True
	Track $\chi^2/\text{ndof}$	$< 5$
	Proton PID DLL( $p - \pi$ )	$> 5$
	Proton PID DLL( $p - K$ )	$> 0$
	Kaon PID DLL( $K - \pi$ )	$> 5$
	Pion PID DLL( $\pi - K$ )	$> 0$
IP $\chi^2$	$> -1$	

Table 4.2 Offline selection criteria for the  $\Xi_{cc}^+$  decay.

Particle	Variable	Cut value
$\Xi_{cc}^+$	$p_T$	$> 4000 \text{ MeV}/c$
	$\ln(\text{IP } \chi^2)$	$< 4$
	MAXDOCA of daughters	$< 0.2 \text{ mm}$
All final-state tracks	ProbNNghost	$< 0.9$
$K$ from $\Xi_{cc}^+$ and $K$ from $\Lambda_c^+$	ProbNNk	$> 0.05$
$K$ from $\Xi_{cc}^+$ and $K$ from $\Lambda_c^+$	ProbNNk - ProbNNpi	$> 0$

#### 4.3.1.2 Offline selection

Offline selection criteria are applied after stripping, as listed in Table 4.2.

#### 4.3.1.3 Trigger requirements

After the selection criteria, the following trigger lines are applied:

- L0: L0Hadron TIS or TOS for  $\Xi_{cc}^+$  candidates
- Hlt1: Hlt1TrackAllL0 TOS for  $\Xi_{cc}^+$  candidates
- Hlt2: Hlt2CharmHadLambdaC2KPPi TOS for  $\Xi_{cc}^+$  candidates

#### 4.3.1.4 Results

After the stripping and offline selection, the MC sample has 449 candidates, which corresponds to an overall efficiency, including the acceptance efficiency, of  $4 \times 10^{-5}$ . Compared to the analysis using 2011 data, which has 108 events before MVA training, the efficiency has increased by a factor of 4. The increase mainly benefits from the relax of lifetime-related selection criteria. If we assume the following parameters, *i.e.*  $\sigma(pp \rightarrow \Xi_{cc}^+ + X) = 50 \text{ nb}$ ,  $\mathcal{B}(\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+) = 5\%$  and  $\mathcal{B}(\Lambda_c^+ \rightarrow p K^- \pi^+) = 5\%$ , then the expected signal yield in 2011 data before the MVA selection is 3.6.

#### 4.3.2 Multivariate analysis

The MVA method used to optimize the selection is Gradient Boosted Decision Tree (BDTG). The signal sample is the truth-matched MC events and contains 449 events. The background sample is provided by the WC sample and contains 3329 events. To avoid



the reduction of the sensitivity at low lifetime region, only variables that are independent or weakly correlate with the lifetime are used for the MVA, as listed below:

- $\Xi_{cc}^+$  decay vertex  $\chi^2$
- $\Xi_{cc}^+$  MAXDOCA
- $\Xi_{cc}^+$   $p_T$
- $\cos\theta_H$
- $\cos\alpha$
- $\Lambda_c^+$  decay vertex  $\chi^2$
- $\Lambda_c^+$  MAXDOCA

Distributions of these variables are shown in Figure 4.2. Blue lines indicate the signal MC, and red shaded are the sideband background.

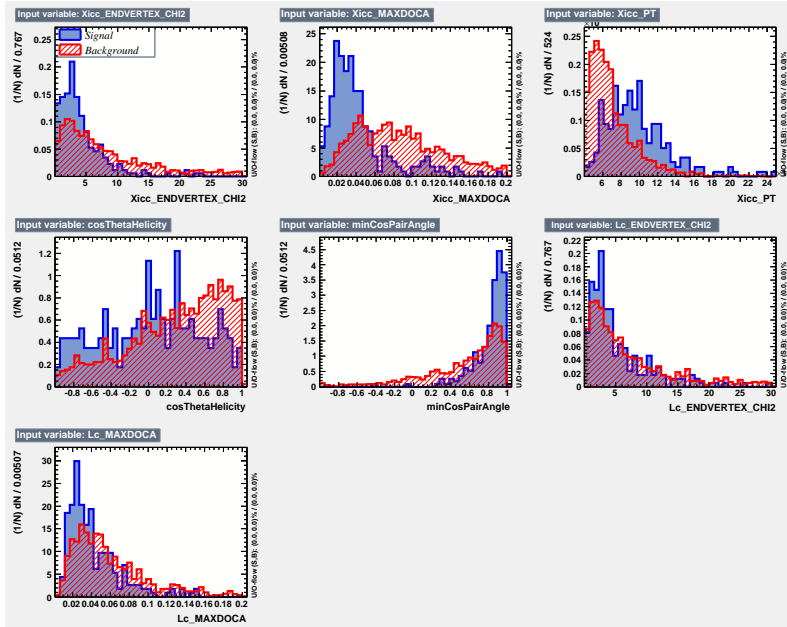


Figure 4.2 Distributions of input variables for the MVA. Blue lines indicate the signal MC, and red shaded are the sideband background.

The performance of the BDTG method is shown in Figure 4.3. Due to the small statistics of the input sample, the BDTG is seriously overtrained, *i.e.* the method recognizes the statistical fluctuation of the input training sample as the characteristic of the signal. Although this do not cause an incorrect calculation of the efficiency since only the efficiency evaluated using a statistical independent sample will be used, this may cause an underestimate of the efficiency since the method is optimized based on fluctuation. It is expected that the performance will improve with larger MC samples.

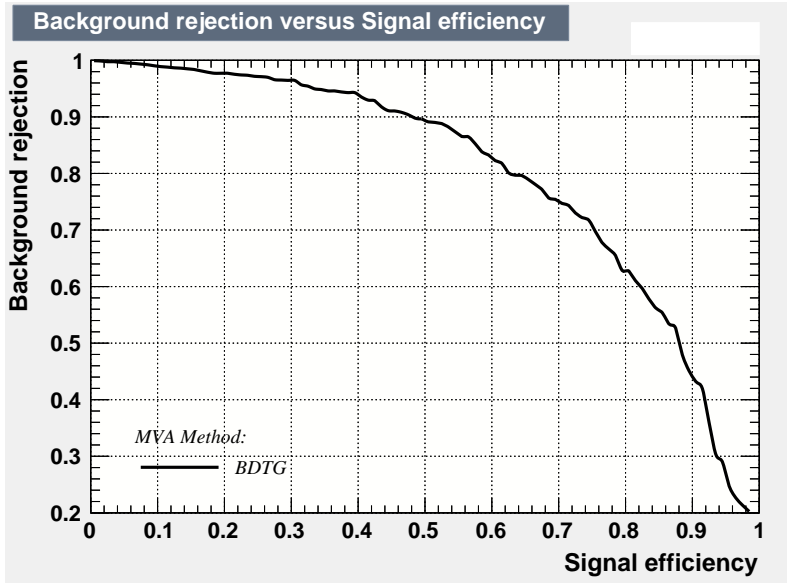


Figure 4.3 ROC curve for the BDTG method. The  $x$ -axis is the signal efficiency, and the  $y$ -axis is the background retention rate.

The optimal cut is determined by maximizing the expected significance  $s/\sqrt{s+b}$ , where  $s$  is the expected signal yield estimated based on the theoretical cross-section, the luminosity and the efficiency from the MC, and  $b$  is the background counts in the signal region estimated from the WC sample. For the signal, we know  $s = 3.6$  from the previous section. Since no known process could significantly contribute physical background to this decay, it is expected that the dominate background is combinatorial. Therefore, the WC sample, a  $\Lambda_c^+$  candidate plus a  $K^+$  track and a  $\pi^+$  track, provides a reasonable reference for the background shape of the signal mode. This is justified by the overlap of the invariant mass distributions for WC and right charge (RC) samples, as shown in Figure 4.4. The estimated background entries in the signal region is  $1.63 \times 10^4$  according to the calculation in Table 4.3, where the signal region is defined as  $|m(\Lambda_c^+) - m_{\text{PDG}}(\Lambda_c^+)| < 15 \text{ MeV}/c^2$  and  $|m(\Xi_{cc}^+) - 3500| < 15 \text{ MeV}/c^2$ . With these pieces, if only the 2011 data is used, the optimal cut is determined to be  $\text{BDTG} > -0.34$ , and the expected significance is 0.04. If the full data set is used and the efficiency of the 2012 data is assumed to be the same as for 2011, then the optimal cut is  $\text{BDTG} > -0.34$  and the expected significance is 0.1. Although the expected significance seems small, it should be noted that the results presented here are very preliminary: the selection can be much improved using large MC samples; more decay modes, *e.g.*  $\Xi_{cc}^+ \rightarrow D^0 p K^- \pi^+$ ,  $\Xi_{cc}^+ \rightarrow D^+ p K^-$ ,  $\Xi_{cc}^+ \rightarrow \Xi_c^+ \pi^+ \pi^-$  and  $\Xi_{cc}^+ \rightarrow \Xi_c^0 \pi^+$ , will complementarily be used for the  $\Xi_{cc}^+$

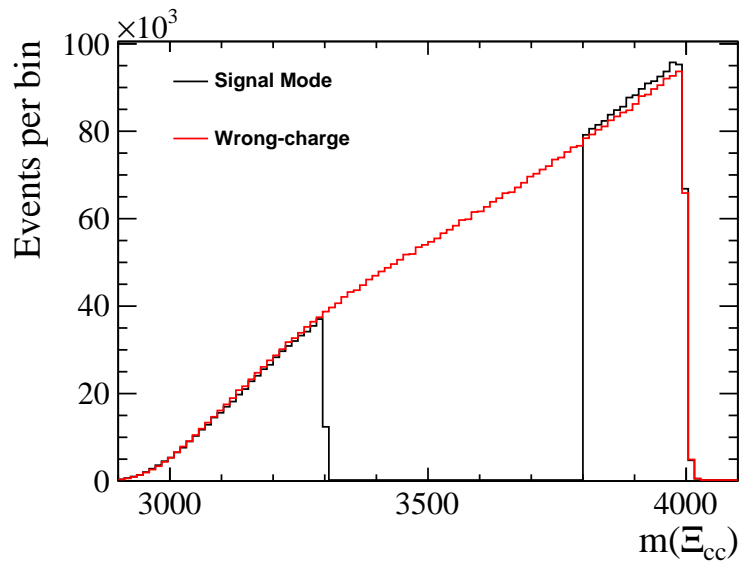


Figure 4.4 Invariant mass distribution of  $\Xi_{cc}^+$  candidates for the signal mode (black line with the signal region blinded) and the WC mode. The scale factor for the WC mode is 0.96.

Table 4.3 Estimated background in the signal region.

WC sample	Scale factor	Signal mode
$1.73 \times 10^4$	0.96	$1.63 \times 10^4$

search.

## Chapter 5 Summary and outlook

### 5.1 Summary

This thesis describes the work of searching for the  $\Xi_{cc}^+$  baryon at LHCb. The data is recorded in  $pp$  collisions at a center-of-mass energy of 7 TeV and corresponds to an integrated luminosity of  $0.65 \text{ fb}^{-1}$ .

The  $\Xi_{cc}^+$  baryon is one of the lightest baryons which contain two charm quarks. It is of great interest, both theoretical and experimental. From the theoretical perspective, the  $\Xi_{cc}^+$  baryon has a structure very similar to those of heavy mesons, since the two charm quarks are tightly combined into a diquark structure and act as a static color source. Therefore, similar to heavy mesons, it provides a tool to explore non-perturbative QCD. Besides, the heavy-meson-like structure also makes calculations of properties of the  $\Xi_{cc}^+$  much easier. Thus the  $\Xi_{cc}^+$  also serves as an excellent testing ground to examine various models. On the other side, conclusive experimental study of the  $\Xi_{cc}^+$  is still missing. The only reported signal came from the SELEX experiment, but the measured properties do not agree with theoretical predictions in a satisfactory way. All other experiments failed to reproduce the SELEX results. New measurements are definitely needed.

The  $\Xi_{cc}^+$  baryon is searched through its decay  $\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$ , and the production ratio of the  $\Xi_{cc}^+$  baryon relative to that of inclusive  $\Lambda_c^+$  baryon,  $R$ , is measured. To avoid biases from preconceptions, the search is performed in a blind approach, *i.e.* the signal window is not looked into until all the selection and fit procedures are fixed. We use loose cuts to reject the most obvious background first, and optimize the signal selection using an artificial neural network. No significant signal is observed, and upper limits for  $R$  are set for a range of mass and lifetime hypotheses, since the mass and lifetime of  $\Xi_{cc}^+$  baryons are a priori unknown. The results are tabulated in Table 5.1 for blocks of the invariant mass and five different lifetime hypotheses. The largest upper limit over the entire mass range is also given. Due to the presence of lifetime-related requirements at the stripping level, the overall selection efficiency is lower for smaller lifetime hypotheses, and the statistical uncertainty on the selection efficiency is larger. Recall that the dominate source of the systematic uncertainty is the MC statistics, therefore, the systematic uncertainty is

larger for smaller lifetime hypotheses. This implies a higher upper limit on  $R$  for smaller lifetime hypotheses.

Table 5.1 Largest values of the upper limits (UL) on  $R$  at the 95% CL in ranges of  $\delta m$  for a set of lifetime hypotheses, given in units of  $10^{-3}$ . The largest values across the entire mass range are also shown.

$\delta m$ (MeV/ $c^2$ )	$R$ , largest 95% CL UL in range, $\times 10^3$				
	100 fs	150 fs	250 fs	333 fs	400 fs
380–429	13	2.7	0.73	0.43	0.33
430–479	11	2.4	0.65	0.39	0.29
480–529	15	3.2	0.85	0.51	0.39
530–579	11	2.3	0.63	0.38	0.29
580–629	11	2.3	0.63	0.38	0.29
630–679	14	3.0	0.81	0.48	0.37
680–729	9.5	2.0	0.56	0.33	0.25
730–779	11	2.3	0.63	0.37	0.28
780–829	13	2.7	0.74	0.45	0.33
830–880	12	2.6	0.70	0.42	0.32
380–880	15	3.2	0.85	0.51	0.39

The upper limit of the relative production ratio for a  $\Xi_{cc}^+$  baryon with a mass of 3520 MeV/ $c^2$  and a lifetime of 100 fs is  $R < 9.5 \times 10^{-3}$  @ 95% CL, which is an order of magnitude lower than the  $R$  value obtained by SELEX, which measured  $R = 9.6 \times 10^{-2}$ . This discrepancy can be explained by different production environment, or that the  $\Xi_{cc}^+$  indeed has a much shorter lifetime than expected. Based on the measured cross-section of prompt  $\Lambda_c^+$  baryons and the assumption that  $\mathcal{B}(\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+) = 5\%$ , the upper limit of the absolute production cross-section of the  $\Xi_{cc}^+$  is  $\sigma(\Xi_{cc}^+) < 40 \mu\text{b}$  @ 95% CL, which does not directly contradict with theoretical predictions, which are of the order of 110 nb.

## 5.2 Outlook

The  $\Xi_{cc}^+$  baryon is not observed in the decay  $\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$  using the data collected by LHCb in 2011, but we are optimistic about the sensitivity of the future search, which will be improved in the following three ways:

1. The 2012 data will be included, which increases the sample size by a factor of 3.

2. More decay modes, particularly final states with a charmed meson instead of a charmed baryon, for which the HLT and offline efficiency is much larger due to the longer lifetime, will be added.

$$\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$$

$$\Xi_{cc}^+ \rightarrow D^0 p K^- \pi^+$$

$$\Xi_{cc}^+ \rightarrow D^+ p K^-$$

$$\Xi_{cc}^+ \rightarrow \Xi_c^+ \pi^+ \pi^-$$

$$\Xi_{cc}^+ \rightarrow \Xi_c^0 \pi^+$$

3. The stripping cuts that reduce our sensitivity in small lifetime region will be removed.

To have a more quantitative estimate of the future search, the following assumptions are made:

- Assume that the production cross-section of the  $\Xi_{cc}^+$  baryon at a center-of-mass energy  $\sqrt{s}$ ,  $\sigma(\sqrt{s})$ , varies linearly with  $\sqrt{s}$ . Since  $\sigma(14 \text{ TeV})$  is known to be 110 nb based on theoretical predictions, we expect that  $\sigma(7 \text{ TeV}) = 55 \text{ nb}$ ,  $\sigma(8 \text{ TeV}) = 63 \text{ nb}$ .
- Assume that the branching fractions of  $\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$ ,  $\Xi_{cc}^+ \rightarrow D^0 p K^- \pi^+$  and  $\Xi_{cc}^+ \rightarrow D^+ p K^-$  are all 5% to simplify calculations.
- For each decay mode, the overall efficiency is assumed to be the same for 2011 and 2012 data.
- MVA efficiency for all the modes and lifetime hypotheses is all assumed to be 40%, the efficiency of the BDTG cut in the previous chapter, which is a conservative estimate.
- Available luminosity for each decay mode are listed in Table 5.2.

Table 5.2 Expected yields of the  $\Xi_{cc}^+$  baryon, including all decay channels and using 2011 and 2012 data, for different lifetime hypotheses.

Decay mode	Luminosity
$\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$	2.65 fb <sup>-1</sup>
$\Xi_{cc}^+ \rightarrow D^0 p K^- \pi^+$	3 fb <sup>-1</sup>
$\Xi_{cc}^+ \rightarrow D^+ p K^-$	3 fb <sup>-1</sup>

The decays  $\Xi_{cc}^+ \rightarrow \Xi_c^+ \pi^+ \pi^-$  and  $\Xi_{cc}^+ \rightarrow \Xi_c^0 \pi^+$  are not included since the overall efficiency of these modes are expected to be low due to the long decay chains.

Table 5.3 Expected yields of the  $\Xi_{cc}^+$  baryon, including three decay channels,  $\Xi_{cc}^+ \rightarrow \Lambda_c^+ K^- \pi^+$ ,  $\Xi_{cc}^+ \rightarrow D^0 p K^- \pi^+$  and  $\Xi_{cc}^+ \rightarrow D^+ p K^-$ , and using 2011 and 2012 data, for different lifetime hypotheses.

Lifetime hypothesis(fs)	Expected yield
100	21
150	26
250	30
333	38
400	41

With these assumptions, the expected yields of the  $\Xi_{cc}^+$  baryon for different lifetime hypotheses, including the three decay modes and using 2011 and 2012 data, are listed in Table 5.3. Depending on the lifetime, 21-38 signal events are expected after all the selections. Keep in mind this is still a conservative estimate. LHCb has a very good opportunity to discover the  $\Xi_{cc}^+$  baryon.

During the Run II phase (2015–2018), two major changes will be made for LHCb. First, the LHC center-of-mass energy will be raised to 13 or 14 TeV. The cross-sections of charm and beauty hadron production, including that of the  $\Xi_{cc}^+$  baryon, will increase significantly. Second, more powerful machines will be used for the HLT farm, which allow to apply looser cuts on the impact parameter of the tracks during the HLT1 stage. This will largely increase the HLT1 efficiencies for charm hadrons. According to MC simulation, the HLT1 efficiency of  $D^0$  decays will increase by 70%<sup>[122]</sup>. And we expect a larger increase for  $\Lambda_c^+$  decays since the  $\Lambda_c^+$  baryon has a shorter lifetime and hence will benefit more from the looser cuts. Assuming that the efficiency increase by a factor of 1.7, the cross-section of the  $\Xi_{cc}^+$  baryon is 102 nb at 13 TeV and with the lifetime of the  $\Xi_{cc}^+$  baryon of 100 fs, at least 120 events are expected in the  $8 \text{ fb}^{-1}$  sample during the whole Run I + Run II phase.

The Belle II experiment is the upgrade of the Belle experiment<sup>[123]</sup>. The production cross-section of the  $\Xi_{cc}^+$  baryon at Belle II is several orders of magnitude smaller (70 – 230 fb) than that at LHCb<sup>[34,39]</sup>, but Belle II is expected to have a much higher efficiency

( $0.15^{[56]}$ ) and much larger integrated luminosity (at order of  $\text{ab}^{-1}$ ), Belle II is expected to collect an integrated luminosity of  $5 \text{ ab}^{-1}$  by around 2019<sup>[124]</sup>. Combining all the decay modes, the expected  $\Xi_{cc}^+$  yield will be of the order of 500. This means that Belle II will be a very competitive experiment for the  $\Xi_{cc}^+$  baryon search. Nevertheless, LHCb could be the first to discover the  $\Xi_{cc}^+$  baryon if the selection is will optimized.



## Reference

- [1] P. Skands. *Introduction to QCD*. CERN-PH-TH-2012-196, arXiv:1207.2389.
- [2] D. J. Gross and F. Wilczek. *Ultraviolet Behavior of Nonabelian Gauge Theories*. Phys.Rev.Lett., 1973, 30:1343–1346.
- [3] H. D. Politzer. *Reliable Perturbative Results for Strong Interactions?* Phys.Rev.Lett., 1973, 30:1346–1349.
- [4] J. Beringer *et al.* *Review of Particle Physics (RPP)*. Phys.Rev., 2012, D86:010001.
- [5] M. Baker, J. Ball, and F. Zachariasen. *The Deconfinement Transition in Dual QCD*. Phys.Rev.Lett., 1988, 61:521.
- [6] G. Zweig. *An SU(3) model for strong interaction symmetry and its breaking. Version 1*. CERN-TH-401.
- [7] M. Gell-Mann. *A Schematic Model of Baryons and Mesons*. Phys.Lett., 1964, 8:214–215.
- [8] D.-H. He, K. Qian, Y.-B. Ding, X.-Q. Li, and P.-N. Shen. *Evaluation of spectra of baryons containing two heavy quarks in bag model*. Phys.Rev., 2004, D70:094004, arXiv:hep-ph/0403301.
- [9] V. Kiselev, A. Likhoded, and A. Onishchenko. *Lifetimes of doubly charmed baryons:  $\Xi_{cc}^+$  and  $\Xi_{cc}^{++}$* . Phys.Rev., 1999, D60:014007, arXiv:hep-ph/9807354.
- [10] D. Ebert, R. Faustov, V. Galkin, and A. Martynenko. *Mass spectra of doubly heavy baryons in the relativistic quark model*. Phys.Rev., 2002, D66:014008, arXiv:hep-ph/0201217.
- [11] A. Martynenko. *Ground-state triply and doubly heavy baryons in a relativistic three-quark model*. Phys.Lett., 2008, B663:317–321, arXiv:0708.2033.
- [12] W. Roberts and M. Pervin. *Heavy baryons in a quark model*. Int.J.Mod.Phys., 2008, A23:2817–2860, arXiv:0711.2492.
- [13] C. Albertus, E. Hernandez, J. Nieves, and J. Verde-Velasco. *Static properties and semileptonic decays of doubly heavy baryons in a nonrelativistic quark model*. Eur.Phys.J., 2007, A32:183–199, arXiv:hep-ph/0610030.
- [14] J. Vijande, H. Garcilazo, A. Valcarce, and F. Fernandez. *Spectroscopy of doubly charmed baryons*. Phys.Rev., 2004, D70:054022, arXiv:hep-ph/0408274.
- [15] S. S. Gershtein, V. V. Kiselev, A. K. Likhoded, and A. I. a. Onishchenko. *Spectroscopy of doubly heavy baryons*. Phys.Rev., 2000, D62:054021.
- [16] L. Tang, X.-H. Yuan, C.-F. Qiao, and X.-Q. Li. *Study of Doubly Heavy Baryon Spectrum via QCD Sum Rules*. Commun.Theor.Phys., 2012, 57:435–444, arXiv:1104.4934.
- [17] E. Bagan, M. Chabab, and S. Narison. *Baryons with two heavy quarks from QCD spectral sum rules*. Phys.Lett., 1993, B306:350–356.
- [18] J.-R. Zhang and M.-Q. Huang. *Doubly heavy baryons in QCD sum rules*. Phys.Rev., 2008, D78:094007, arXiv:0810.5396.

## Reference

---

- [19] M. A. Shifman, A. Vainshtein, and V. I. Zakharov. *QCD and Resonance Physics. Sum Rules.* Nucl.Phys., 1979, B147:385–447.
- [20] L. Reinders, H. Rubinstein, and S. Yazaki. *Hadron Properties from QCD Sum Rules.* Phys.Rept., 1985, 127:1.
- [21] M. A. Shifman and M. Voloshin. *Preasymptotic Effects in Inclusive Weak Decays of Charmed Particles.* Sov.J.Nucl.Phys., 1985, 41:120.
- [22] Z.-G. Wang. *Analysis of the  $\frac{1}{2}^+$  doubly heavy baryon states with QCD sum rules.* Eur.Phys.J., 2010, A45:267–274, arXiv:1001.4693.
- [23] V. Kiselev and A. Likhoded. *Baryons with two heavy quarks.* Phys.Usp., 2002, 45:455–506, arXiv:hep-ph/0103169.
- [24] S.-P. Tong, Y.-B. Ding, X.-H. Guo, H.-Y. Jin, X.-Q. Li, et al. *Spectra of baryons containing two heavy quarks in potential model.* Phys.Rev., 2000, D62:054024, arXiv:hep-ph/9910259.
- [25] D. Lichtenberg, R. Roncaglia, and E. Predazzi. *Mass sum rules for singly and doubly heavy flavored hadrons.* Phys.Rev., 1996, D53:6678–6681, arXiv:hep-ph/9511461.
- [26] R. Lewis, N. Mathur, and R. Woloshyn. *Charmed baryons in lattice QCD.* Phys.Rev., 2001, D64:094509, arXiv:hep-ph/0107037.
- [27] J. Flynn, F. Mescia, and A. S. B. Tariq. *Spectroscopy of doubly charmed baryons in lattice QCD.* JHEP, 2003, 0307:066, arXiv:hep-lat/0307025.
- [28] B. Guberina, B. Melic, and H. Stefancic. *Inclusive decays and lifetimes of doubly charmed baryons.* Eur.Phys.J., 1999, C9:213–219, arXiv:hep-ph/9901323.
- [29] A. Likhoded and A. Onishchenko. *Lifetimes of doubly heavy baryons.* arXiv:hep-ph/9912425.
- [30] C.-H. Chang, T. Li, X.-Q. Li, and Y.-M. Wang. *Lifetime of doubly charmed baryons.* Commun.Theor.Phys., 2008, 49:993–1000, arXiv:0704.0016.
- [31] K. Chetyrkin and M. Steinhauser. *Short distance mass of a heavy quark at order  $\alpha_s^3$ .* Phys.Rev.Lett., 1999, 83:4001–4004, arXiv:hep-ph/9907509.
- [32] G. Bellini, I. I. Bigi, and P. Dornan. *Lifetimes of charm and beauty hadrons.* Phys.Rept., 1997, 289:1–155.
- [33] V. Kiselev and A. Likhoded. *Comment on ‘First observation of doubly charmed baryon  $\Xi_{cc}^+$ ’.* arXiv:hep-ph/0208231.
- [34] J. Ma and Z. Si. *Factorization approach for inclusive production of doubly heavy baryon.* Phys.Lett., 2003, B568:135–145, arXiv:hep-ph/0305079.
- [35] C.-H. Chang, C.-F. Qiao, J.-X. Wang, and X.-G. Wu. *Estimate of the hadronic production of the doubly charmed baryon  $\Xi_{cc}^+$  under GM-VFN scheme.* Phys.Rev., 2006, D73:094022, arXiv:hep-ph/0601032.
- [36] A. Berezhnoy, V. Kiselev, A. Likhoded, and A. Onishchenko. *Doubly charmed baryon production in hadronic experiments.* Phys.Rev., 1998, D57:4385–4392, arXiv:hep-ph/9710339.

## Reference

---

- [37] G. T. Bodwin, E. Braaten, and G. P. Lepage. *Rigorous QCD analysis of inclusive annihilation and production of heavy quarkonium*. Phys.Rev., 1995, D51:1125–1171, arXiv:hep-ph/9407339.
- [38] A. F. Falk, M. E. Luke, M. J. Savage, and M. B. Wise. *Heavy quark fragmentation to baryons containing two heavy quarks*. Phys.Rev., 1994, D49:555–558, arXiv:hep-ph/9305315.
- [39] V. Kiselev, A. Likhoded, and M. Shevlyagin. *Double charmed baryon production at B factory*. Phys.Lett., 1994, B332:411–414, arXiv:hep-ph/9408407.
- [40] C.-H. Chang, J.-P. Ma, C.-F. Qiao, and X.-G. Wu. *Hadronic production of the doubly charmed baryon  $\Xi_{cc}$  with intrinsic charm*. J.Phys., 2007, G34:845, arXiv:hep-ph/0610205.
- [41] R. Aaij *et al.* *Prompt charm production in pp collisions at  $\sqrt{s} = 7$  TeV*. Nucl.Phys., 2013, B871:1–20, arXiv:1302.2864.
- [42] J. Russ *et al.* *First charm hadroproduction results from SELEX*. CMU-HEP-98-07, DOE-ER-40682-103, arXiv:hep-ex/9812031.
- [43] M. Mattson *et al.* *First observation of the doubly charmed baryon  $\Xi_{cc}^+$* . Phys.Rev.Lett., 2002, 89:112001, arXiv:hep-ex/0208014.
- [44] E. Gross and O. Vitells. *Trial factors or the look elsewhere effect in high energy physics*. Eur.Phys.J., 2010, C70:525–530, arXiv:1005.1891.
- [45] A. Ocherashvili *et al.* *Confirmation of the double charm baryon  $\Xi_{cc}^+(3520)$  via its decay to  $pD^+K^-$* . Phys.Lett., 2005, B628:18–24, arXiv:hep-ex/0406033.
- [46] J. Engelfried. *SELEX: Recent Progress in the Analysis of Charm-Strange and Double-Charm Baryons*. eConf, 2006, C0610161:003, arXiv:hep-ex/0702001.
- [47] J. Engelfried. *The experimental discovery of double-charm baryons*. Nucl.Phys., 2005, A752:121–128.
- [48] J. Engelfried. *Doubly charmed baryon results from SELEX*. Talk at ECTStar 2013.
- [49] A. Kushnirenko *et al.* *Precision measurements of the  $\Lambda_{c^+}$  and  $D^0$  lifetimes*. Phys.Rev.Lett., 2001, 86:5243–5246, arXiv:hep-ex/0010014.
- [50] P. Frabetti, V. Giordano, G. Molinari, C. Bogart, H. Cheung, P. Coteus, S. Culy, J. Cumalat, C. Dallapiccola, E. Erdos, *et al.* *Description and performance of the Fermilab E687 spectrometer*. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 1992, 320(3):519–547.
- [51] *FOCUS double charm baryon search*. [http://www-focus.fnal.gov/xicc/xicc\\_focus.html](http://www-focus.fnal.gov/xicc/xicc_focus.html).
- [52] W. A. Rolke, A. M. Lopez, and J. Conrad. *Limits and confidence intervals in the presence of nuisance parameters*. Nucl.Instrum.Meth., 2005, A551:493–503, arXiv:physics/0403059.
- [53] S. Ratti, J. Link, M. Reyes, P. Yager, J. Anjos, I. Bediaga, C. Gobel, J. Magnin, A. Massafferri, J. de Miranda, *et al.* *New results on c-baryons and a search for cc-baryons in FOCUS*. Nuclear Physics-Section B-PS-Proceedings Supplements, 2003, 115:33–36.
- [54] A. Bondar. *The BELLE detector*. Nucl.Instrum.Meth., 1998, A408:64–76.

## Reference

---

- [55] R. Chistov *et al.* *Observation of new states decaying into  $\Lambda_c^+ K^- \pi^+$  and  $\Lambda_c^+ K_S^0 \pi^-$ .* Phys.Rev.Lett., 2006, 97:162001, [arXiv:hep-ex/0606051](#).
- [56] Y. Kato *et al.* *Search for doubly charmed baryons and study of charmed strange baryons at Belle.* Phys.Rev., 2014, D89:052003, [arXiv:1312.1026](#).
- [57] B. Aubert *et al.* *The BaBar detector.* Nucl.Instrum.Meth., 2002, A479:1–116, [arXiv:hep-ex/0105044](#).
- [58] B. Aubert *et al.* *Search for doubly charmed baryons  $\Xi_{cc}^+$  and  $\Xi_{cc}^{++}$  in BaBar.* Phys.Rev., 2006, D74:011103, [arXiv:hep-ex/0605075](#).
- [59] L. Evans and P. Bryant. *LHC Machine.* JINST, 2008, 3:S08001.
- [60] J. R. Espinosa and J. F. Gunion. *No-Lose Theorem for Higgs Boson Searches at a Future Linear Collider.* Phys. Rev. Lett., 1999, 82:1084–1087.
- [61] D. Abbaneo *et al.* *A Combination of preliminary electroweak measurements and constraints on the standard model.* SLAC-REPRINT-2000-098, CERN-EP-2000-016.
- [62] M. E. Peskin and J. D. Wells. *How can a heavy Higgs boson be consistent with the precision electroweak measurements?* Phys.Rev., 2001, D64:093003, [arXiv:hep-ph/0101342](#).
- [63] H. E. Haber and G. L. Kane. *The Search for Supersymmetry: Probing Physics Beyond the Standard Model.* Phys.Rept., 1985, 117:75–263.
- [64] M. Benedikt, P. Collier, V. Mertens, J. Poole, and K. Schindl. *LHC Design Report.*
- [65] A. Feschenko, C. Hill, A. Lombardi, A. Liou, R. Scrivens, *et al.* *Tests of the CERN proton linac performance for LHC-type beams.* CERN-PS-2000-053-PP.
- [66] G. Aad *et al.* *The ATLAS Experiment at the CERN Large Hadron Collider.* JINST, 2008, 3:S08003.
- [67] S. Chatrchyan *et al.* *The CMS experiment at the CERN LHC.* JINST, 2008, 3:S08004.
- [68] J. Alves, A. Augusto *et al.* *The LHCb Detector at the LHC.* JINST, 2008, 3:S08005.
- [69] K. Aamodt *et al.* *The ALICE experiment at the CERN LHC.* JINST, 2008, 3:S08002.
- [70] M. Kobayashi and T. Maskawa. *CP-violation in the renormalizable theory of weak interaction.* Progress of Theoretical Physics, 1973, 49(2):652–657.
- [71] R. Aaij *et al.* *Measurement of  $J/\psi$  production in  $pp$  collisions at  $\sqrt{s} = 7$  TeV.* Eur.Phys.J., 2011, C71:1645, [arXiv:1103.0423](#).
- [72] M. L. Mangano. *Two lectures on heavy quark production in hadronic collisions.* CERN-TH-97-328, [arXiv:hep-ph/9711337](#).
- [73] M. Krasowski, M. Kucharczyk, W. Manner, W. Maenner, G. Polok, *et al.* *Primary vertex reconstruction.* LHCB-2007-011.
- [74] R. E. Kalman. *A new approach to linear filtering and prediction problems.* Journal of Fluids Engineering, 1960, 82(1):35–45.
- [75] A. J. *et al.* *Measurement of the track finding efficiency.* LHCB-PUB-2011-025.
- [76] T. Ruf, J. v.Tilburg, C. Jones, *et al.* *LHCb Performance Paper.* <https://twiki.cern.ch/twiki/bin/view/LHCB/LHCBPerformancePaper>.

## Reference

---

- [77] P. A. Cherenkov. *Visible emission of clean liquids by action of  $\gamma$  radiation*. Doklady Akademii Nauk SSSR, 1934, 2:451.
- [78] R. Forty and O. Schneider. *RICH pattern recognition*. LHCb-98-040.
- [79] R. Forty. *RICH pattern recognition for LHCb*. Nucl.Instrum.Meth., 1999, A433:257–261.
- [80] A. Powell. *Reconstruction and PID performance of the LHCb RICH detectors*. Nucl.Instrum.Meth., 2011, A639:260–263.
- [81] M. Adinolfi *et al.* *Performance of the LHCb RICH detector at the LHC*. Eur.Phys.J., 2013, C73:2431, [arXiv:1211.6759](https://arxiv.org/abs/1211.6759).
- [82] M. Pivk and F. R. Le Diberder. *sPlot: a statistical tool to unfold data distributions*. Nucl.Instrum.Meth., 2005, A555:356–369, [arXiv:physics/0402083](https://arxiv.org/abs/physics/0402083).
- [83] *LHCb calorimeters: Technical design report*. CERN-LHCC-2000-036.
- [84] G. Bencivenni, G. Felici, F. Murtas, P. Valente, W. Bonivento, *et al.* *A triple-GEM detector with pad readout for the inner region of the first LHCb muon station*. LHCb-2001-051.
- [85] B. Botchine, V. Lazarev, N. Saguidova, A. Vorobyov, A. Kachtchouk, *et al.* *Wire pad chambers and cathode pad chambers for the LHCb muon system*. LHCb-2000-114.
- [86] *LHCb muon system technical design report*. CERN-LHCC-2001-010.
- [87] F. Archilli, W. Baldini, G. Bencivenni, N. Bondar, W. Bonivento, *et al.* *Performance of the Muon Identification at LHCb*. JINST, 2013, 8:P10020, [arXiv:1306.0249](https://arxiv.org/abs/1306.0249).
- [88] E. Rodrigues. *The LHCb trigger system*. Nucl.Phys.Proc.Suppl., 2007, 170:298–302.
- [89] R. Aaij, J. Albrecht, F. Alessio, S. Amato, E. Aslanides, *et al.* *The LHCb Trigger and its Performance in 2011*. JINST, 2013, 8:P04022, [arXiv:1211.3055](https://arxiv.org/abs/1211.3055).
- [90] A. Barczyk, G. Haefeli, R. Jacobsson, B. Jost, and N. Neufeld. *1 MHz Readout*. LHCb-2005-062.
- [91] V. Gligorov, C. Thomas, and M. Williams. *The HLT inclusive B triggers*. LHCb-PUB-2011-016, LHCb-INT-2011-030.
- [92] P. Spradlin. *The LHCb prompt charm triggers*. LHCb-PROC-2013-076, [arXiv:1311.7585](https://arxiv.org/abs/1311.7585).
- [93] G. Barrand, I. Belyaev, P. Binko, M. Cattaneo, R. Chytracsek, *et al.* *GAUDI - A software architecture and framework for building HEP data processing applications*. Comput.Phys.Commun., 2001, 140:45–55.
- [94] *The GAUSS Project*. <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/gauss>.
- [95] I. Belyaev, P. Charpentier, S. Easo, P. Mato, J. Palacios, *et al.* *Simulation application for the LHCb experiment*. eConf, 2003, C0303241:TUMT003, [arXiv:physics/0306035](https://arxiv.org/abs/physics/0306035).
- [96] M. Clemencic *et al.* *The LHCb simulation application, Gauss: Design, evolution and experience*. J.Phys.Conf.Ser., 2011, 331:032023.
- [97] I. Belyaev *et al.* *Handling of the generation of primary events in Gauss, the LHCb simulation framework*. IEEE Nucl.Sci.Symp.Conf.Rec., 2010, 2010:1155–1161.
- [98] *The BOOLE Project*. <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/boole>.
- [99] T. Sjöstrand, S. Mrenna, and P. Skands. *PYTHIA 6.4 physics and manual*. JHEP, 2006, 05:026, [arXiv:hep-ph/0603175](https://arxiv.org/abs/hep-ph/0603175).

## Reference

---

- [100] D. J. Lange. *The EvtGen particle decay simulation package*. Nucl. Instrum. Meth., 2001, A462:152–155.
- [101] C.-H. Chang, J.-X. Wang, and X.-G. Wu. *GENXICC: A Generator for hadronic production of the double heavy baryons  $\Xi_{cc}$ ,  $\Xi_{bc}$  and  $\Xi_{bb}$* . Comput.Phys.Commun., 2007, 177:467–478, arXiv:hep-ph/0702054.
- [102] C.-H. Chang, J.-X. Wang, and X.-G. Wu. *GENXICC2.0: An upgraded version of the generator for hadronic production of double heavy baryons  $\Xi_{cc}$ ,  $\Xi_{bc}$  and  $\Xi_{bb}$* . Comput.Phys.Commun., 2010, 181:1144–1149, arXiv:0910.4462.
- [103] S. Agostinelli *et al.* *Geant4: a simulation toolkit*. Nucl. Instrum. Meth., 2003, A506:250.
- [104] J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Dubois, *et al.* *Geant4 developments and applications*. IEEE Trans.Nucl.Sci., 2006, 53:270.
- [105] *The MOORE Project*. <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/moore>.
- [106] *The BRUNEL Project*. <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel>.
- [107] *The DAVINCI Project*. <http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci>.
- [108] J. R. Klein and A. Roodman. *Blind analysis in nuclear and particle physics*. Annu. Rev. Nucl. Part. Sci., 2005, 55:141–163.
- [109] W. D. Hulsbergen. *Decay chain fitting with a Kalman filter*. Nucl.Instrum.Meth., 2005, A552:566–575, arXiv:physics/0503191.
- [110] D. O. Hebb. *The organization of behavior: A neuropsychological theory*. Psychology Press, 2005.
- [111] M. Mohri, A. Rostamizadeh, and A. Talwalkar. *Foundations of machine learning*. MIT Press, 2012.
- [112] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss. *TMVA: the Toolkit for Multivariate Data Analysis with ROOT*. PoS, 2007, ACAT:040, arXiv:physics/0703039.
- [113] V. Gligorov. *Background Category*. <https://indico.cern.ch/event/262630>.
- [114] B. Liu. *Tracking efficiency*. <https://indico.cern.ch/event/189746>.
- [115] M. Needham. *Tracking Systematic for X(3872)*. <https://indico.cern.ch/event/135014>.
- [116] LHCb CALO group. *L0Calo efficiency*. <https://indico.cern.ch/event/198143>.
- [117] LHCb CALO group. *L0Calo efficiency twiki*. <https://twiki.cern.ch/twiki/bin/viewauth/LHCbPhysics/CalorimeterObjectsToolsGroupDOC>.
- [118] A. M. Sanchez *et al.* *Performances of the LHCb L0 Calorimeter Trigger*. LHCb-PUB-2011-026.
- [119] A. M. Sanchez. *Talk on TupleToolL0Calo*. <https://indico.cern.ch/event/217143>.
- [120] LHCb tracking group. *TrackSmearingTool*. <https://twiki.cern.ch/twiki/bin/viewauth/LHCbPhysics/TrackSmearingTool>. LHCb internal material.
- [121] C. Jones. *Charged PID*. <https://indico.cern.ch/event/159519>.

## Reference

---

- [122] J. Albrecht, M. Cattaneo, T. Gershon, V. Gligorov, P. Koppenburg, and G. Raven. *Implications of post-LS1 running conditions on LHCb's data processing*. Technical Report LHCb-PUB-2013-008, CERN, Geneva, Jun, 2013.
- [123] T. Abe *et al.* *Belle II Technical Design Report*. KEK-REPORT-2010-1, arXiv:1011.0352.
- [124] D. M. Asner, E. Dart, and T. Hara. *Belle II Experiment Network and Computing*. PNNL-SA-97204, arXiv:1308.0672.

## Acknowledgement

衷心感谢我的导师高原宁教授对我的悉心指导，他的言传身教将让我受益终身。高老师让我认识了高能物理这样一个激动人心的领域，在科研上给予了大量细致的指导，没有高老师的帮助本论文不可能如期完成。

感谢张锋博士、Matthew Charles博士和Patrick Spradlin博士，在合作期间他们给了我巨大的帮助，即使是非常细小的技术问题，他们也不厌其烦地指导我。Matthew Charles博士在合作组内部评审和英文文档写作方面都给了我很大的帮助，没有他的帮忙我的论文工作将会大大延迟。感谢杨振伟副教授对我的关心和在论文上的帮助。

感谢陈少敏教授，朱相雷副教授和王喆副教授对我的指导和帮助。感谢在高能所期间莫小虎老师的指导。

感谢实验室全体同仁，你们让我的博士生活更加丰富多彩。

最后感谢我的家人和女朋友，你们的支持是我前进的动力。没有你们，本论文无法顺利完成。



## 声 明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本学位论文的研究成果不包含任何他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。

签 名：\_\_\_\_\_ 日 期：\_\_\_\_\_

## Appendix A Tables for PID calibration

### A.1 Calibration of the $\Xi_{cc}^+$ MC using the multiplicity in the $\Lambda_c^+$ sample of data

Table A.1 Calibration of  $K$  identification from  $\Lambda_c^+$  decays in the  $\Xi_{cc}^+$  MC, with the multiplicity of  $\Xi_{cc}^+$  MC weighted to that in  $\Lambda_c^+$  data.

SubID	MagDown		MagUp	
	Run range	PIDK > 10 Efficiency	Run Runge	PIDK > 10 Efficiency
2	92821-93224	–	94261-96641	(86.99 ± 0.23)%
3	93225-97884	(87.01 ± 0.21)%	96642-100243	(87.02 ± 0.24)%
4	97885-98198	(87.64 ± 0.20)%	100244-102505	(86.76 ± 0.22)%
5	98199 -101905	(87.24 ± 0.22)%	102506-102893	(86.79 ± 0.21)%
6	101906-102378	(86.97 ± 0.22)%	102894-104263	(86.55 ± 0.33)%
7	102379-103361	(86.86 ± 0.20)%		
8	103362-103686	(87.37 ± 0.20)%		

Table A.2 Calibration of  $p$  identification from  $\Lambda_c^+$  decay in the  $\Xi_{cc}^+$  MC, with the multiplicity of  $\Xi_{cc}^+$  MC weighted to that in  $\Lambda_c^+$  data.

SubID	MagDown		MagUp	
	Run range	PIDp > 10 Efficiency	Run Runge	PIDp > 10 Efficiency
2	92821-93224	–	94261-96641	(89.12 ± 0.38)%
3	93225-97884	(89.81 ± 0.40)%	96642-100243	(89.61 ± 0.70)%
4	97885-98198	(89.58 ± 0.30)%	100244-102505	(88.91 ± 0.50)%
5	98199 -101905	(90.29 ± 0.87)%	102506-102893	(88.74 ± 0.47)%
6	101906-102378	(89.69 ± 0.49)%	102894-104263	(89.29 ± 0.91)%
7	102379-103361	(90.41 ± 0.70)%		
8	103362-103686	(88.79 ± 0.67)%		

Table A.3 Calibration of  $\pi$  identification from  $\Lambda_c^+$  decay in the  $\Xi_{cc}^+$  MC, with the multiplicity of  $\Xi_{cc}^+$  MC weighted to that in  $\Lambda_c^+$  data.

SubID	MagDown		MagUp	
	Run range	PIDK < -5 Efficiency	Run Runge	PIDK < -5 Efficiency
2	92821-93224	–	94261-96641	(70.49 ± 0.21)%
3	93225-97884	(69.63 ± 0.18)%	96642-100243	(67.06 ± 0.22)%
4	97885-98198	(66.68 ± 0.19)%	100244-102505	(65.36 ± 0.21)%
5	98199 -101905	(68.19 ± 0.19)%	102506-102893	(67.01 ± 0.22)%
6	101906-102378	(67.92 ± 0.19)%	102894-104263	(65.57 ± 0.37)%
7	102379-103361	(67.91 ± 0.19)%		
8	103362-103686	(65.54 ± 0.20)%		

 Table A.4 Calibration of  $K$  identification from  $\Xi_{cc}^+$  decay in the  $\Xi_{cc}^+$  MC, with the multiplicity of  $\Xi_{cc}^+$  MC weighted to that in  $\Lambda_c^+$  data.

SubID	MagDown		MagUp	
	Run range	PIDK > 10 Efficiency	Run Runge	PIDK > 10 Efficiency
2	92821-93224	–	94261-96641	(87.24 ± 0.23)%
3	93225-97884	(87.16 ± 0.22)%	96642-100243	(87.27 ± 0.24)%
4	97885-98198	(87.87 ± 0.21)%	100244-102505	(87.15 ± 0.23)%
5	98199 -101905	(87.41 ± 0.22)%	102506-102893	(87.12 ± 0.22)%
6	101906-102378	(87.12 ± 0.23)%	102894-104263	(86.88 ± 0.35)%
7	102379-103361	(87.08 ± 0.21)%		
8	103362-103686	(87.59 ± 0.21)%		

Table A.5 Calibration of  $\pi$  identification from  $\Xi_{cc}^+$  decay in the  $\Xi_{cc}^+$  MC, with the multiplicity of  $\Xi_{cc}^+$  MC weighted to that in  $\Lambda_c^+$  data.

SubID	MagDown		MagUp	
	Run range	PIDK < -5 Efficiency	Run Runge	PIDK < -5 Efficiency
2	92821-93224	–	94261-96641	(69.42 ± 0.20)%
3	93225-97884	(67.80 ± 0.22)%	96642-100243	(65.72 ± 0.21)%
4	97885-98198	(64.67 ± 0.22)%	100244-102505	(64.09 ± 0.21)%
5	98199 -101905	(66.26 ± 0.22)%	102506-102893	(65.84 ± 0.21)%
6	101906-102378	(66.03 ± 0.22)%	102894-104263	(64.23 ± 0.35)%
7	102379-103361	(66.00 ± 0.23)%		
8	103362-103686	(63.52 ± 0.24)%		

 Table A.6 Calibration of  $K$  identification from  $\Lambda_c^+$  decay in the  $\Lambda_c^+$  MC, with the multiplicity of  $\Lambda_c^+$  MC weighted to that in  $\Lambda_c^+$  data.

SubID	MagDown		MagUp	
	Run range	PIDK > 10 Efficiency	Run Runge	PIDK > 10 Efficiency
2	92821-93224	–	94261-96641	(86.10 ± 0.18)%
3	93225-97884	(85.86 ± 0.17)%	96642-100243	(85.97 ± 0.19)%
4	97885-98198	(86.38 ± 0.17)%	100244-102505	(85.84 ± 0.17)%
5	98199 -101905	(86.14 ± 0.18)%	102506-102893	(85.79 ± 0.17)%
6	101906-102378	(85.77 ± 0.18)%	102894-104263	(85.62 ± 0.27)%
7	102379-103361	(85.64 ± 0.16)%		
8	103362-103686	(86.12 ± 0.17)%		

Table A.7 Calibration of  $p$  identification from  $\Lambda_c^+$  decay in the  $\Lambda_c^+$  MC, with the multiplicity of  $\Lambda_c^+$  MC weighted to that in  $\Lambda_c^+$  data.

SubID	MagDown		MagUp	
	Run range	PIDp > 10 Efficiency	Run Runge	PIDp > 10 Efficiency
2	92821-93224	–	94261-96641	(88.92 ± 0.54)%
3	93225-97884	(89.56 ± 0.60)%	96642-100243	(89.42 ± 0.89)%
4	97885-98198	(89.03 ± 0.49)%	100244-102505	(88.65 ± 0.67)%
5	98199 -101905	(94.03 ± 1.22)%	102506-102893	(88.39 ± 0.59)%
6	101906-102378	(89.38 ± 0.70)%	102894-104263	(91.50 ± 1.21)%
7	102379-103361	(90.37 ± 1.17)%		
8	103362-103686	(88.76 ± 0.71)%		

 Table A.8 Calibration of  $\pi$  identification from  $\Lambda_c^+$  decay in the  $\Lambda_c^+$  MC, with the  $\Lambda_c^+$  MC multiplicity weighted to the  $\Lambda_c^+$  data.

SubID	MagDown		MagUp	
	Run range	PIDK < -5 Efficiency	Run Runge	PIDK < -5 Efficiency
2	92821-93224	–	94261-96641	(71.61 ± 0.18)%
3	93225-97884	(69.55 ± 0.18)%	96642-100243	(68.26 ± 0.19)%
4	97885-98198	(66.67 ± 0.18)%	100244-102505	(66.74 ± 0.18)%
5	98199 -101905	(68.09 ± 0.18)%	102506-102893	(68.27 ± 0.19)%
6	101906-102378	(67.82 ± 0.18)%	102894-104263	(66.91 ± 0.31)%
7	102379-103361	(67.82 ± 0.18)%		
8	103362-103686	(65.55 ± 0.20)%		

A.2 Calibration of the  $\Xi_{cc}^+$  MC without the multiplicity weighting

 Table A.9 Calibration of  $K$  identification from  $\Lambda_c^+$  decay in the  $\Xi_{cc}^+$  MC.

SubID	MagDown		MagUp	
	Run range	PIDK > 10 Efficiency	Run Runge	PIDK > 10 Efficiency
2	92821-93224	–	94261-96641	(89.81 ± 0.29)%
3	93225-97884	(90.16 ± 0.31)%	96642-100243	(89.65 ± 0.30)%
4	97885-98198	(90.26 ± 0.29)%	100244-102505	(89.40 ± 0.26)%
5	98199 -101905	(90.16 ± 0.27)%	102506-102893	(89.05 ± 0.28)%
6	101906-102378	(90.36 ± 0.33)%	102894-104263	(88.83 ± 0.50)%
7	102379-103361	(89.75 ± 0.29)%		
8	103362 -103686	(90.47 ± 0.27)%		

 Table A.10 Calibration of  $p$  identification from  $\Lambda_c^+$  decay in the  $\Xi_{cc}^+$  MC.

SubID	MagDown		MagUp	
	Run range	PIDp > 10 Efficiency	Run Runge	PIDp > 10 Efficiency
2	92821-93224	–	94261-96641	(90.85 ± 0.30)%
3	93225-97884	(91.93 ± 0.43)%	96642-100243	(91.64 ± 0.55)%
4	97885-98198	(91.37 ± 0.26)%	100244-102505	(90.85 ± 0.41)%
5	98199 -101905	(91.02 ± 0.82)%	102506-102893	(90.44 ± 0.31)%
6	101906-102378	(91.21 ± 0.41)%	102894-104263	(89.88 ± 1.15)%
7	102379-103361	(91.91 ± 0.53)%		
8	103362 -103686	(91.18 ± 0.77)%		

Table A.11 Calibration of  $\pi$  identification from  $\Lambda_c^+$  decay in the  $\Xi_{cc}^+$  MC.

SubID	MagDown		MagUp	
	Run range	PIDK < -5 Efficiency	Run Runge	PIDK < -5 Efficiency
2	92821-93224	–	94261-96641	(75.53 ± 0.27)%
3	93225-97884	(73.85 ± 0.25)%	96642-100243	(71.84 ± 0.30)%
4	97885-98198	(71.08 ± 0.27)%	100244-102505	(70.27 ± 0.33)%
5	98199 -101905	(72.34 ± 0.28)%	102506-102893	(72.05 ± 0.29)%
6	101906-102378	(72.31 ± 0.27)%	102894-104263	(70.43 ± 0.51)%
7	102379-103361	(72.39 ± 0.25)%		
8	103362 -103686	(69.77 ± 0.30)%		

 Table A.12 Calibration of  $K$  identification from  $\Xi_{cc}^+$  decay in the  $\Xi_{cc}^+$  MC.

SubID	MagDown		MagUp	
	Run range	PIDK > 10 Efficiency	Run Runge	PIDK > 10 Efficiency
2	92821-93224	–	94261-96641	(90.18 ± 0.31)%
3	93225-97884	(90.46 ± 0.31)%	96642-100243	(90.05 ± 0.32)%
4	97885-98198	(90.60 ± 0.29)%	100244-102505	(89.88 ± 0.29)%
5	98199 -101905	(90.45 ± 0.26)%	102506-102893	(89.46 ± 0.31)%
6	101906-102378	(90.59 ± 0.31)%	102894-104263	(89.21 ± 0.58)%
7	102379-103361	(90.11 ± 0.28)%		
8	103362 -103686	(90.82 ± 0.27)%		

Table A.13 Calibration of  $\pi$  identification from  $\Xi_{cc}^+$  decay in the  $\Xi_{cc}^+$  MC.

SubID	MagDown		MagUp	
	Run range	PIDK < -5 Efficiency	Run Runge	PIDK < -5 Efficiency
2	92821-93224	–	94261-96641	(74.38 ± 0.27)%
3	93225-97884	(72.57 ± 0.27)%	96642-100243	(70.45 ± 0.29)%
4	97885-98198	(69.70 ± 0.29)%	100244-102505	(68.89 ± 0.32)%
5	98199 -101905	(70.99 ± 0.31)%	102506-102893	(70.75 ± 0.29)%
6	101906-102378	(71.01 ± 0.29)%	102894-104263	(69.07 ± 0.50)%
7	102379-103361	(71.05 ± 0.28)%		
8	103362 -103686	(68.33 ± 0.33)%		

 Table A.14 Calibration of  $K$  identification from  $\Lambda_c^+$  decay in the  $\Lambda_c^+$  MC.

SubID	MagDown		MagUp	
	Run range	PIDK > 10 Efficiency	Run Runge	PIDK > 10 Efficiency
2	92821-93224	–	94261-96641	(87.77 ± 0.20)%
3	93225-97884	(87.59 ± 0.18)%	96642-100243	(87.57 ± 0.21)%
4	97885-98198	(87.91 ± 0.17)%	100244-102505	(87.43 ± 0.18)%
5	98199 -101905	(87.77 ± 0.18)%	102506-102893	(87.22 ± 0.19)%
6	101906-102378	(87.55 ± 0.19)%	102894-104263	(87.04 ± 0.33)%
7	102379-103361	(87.24 ± 0.17)%		
8	103362 -103686	(87.85 ± 0.18)%		



Table A.15 Calibration of  $p$  identification from  $\Lambda_c^+$  decay in the  $\Lambda_c^+$  MC.

SubID	MagDown		MagUp	
	Run range	PID $p$ > 10 Efficiency	Run Runge	PID $p$ > 10 Efficiency
2	92821-93224	–	94261-96641	(90.08 $\pm$ 0.47)%
3	93225-97884	(90.92 $\pm$ 0.57)%	96642-100243	(90.74 $\pm$ 0.82)%
4	97885-98198	(90.18 $\pm$ 0.42)%	100244-102505	(89.95 $\pm$ 0.61)%
5	98199 -101905	(90.92 $\pm$ 0.92)%	102506-102893	(89.49 $\pm$ 0.48)%
6	101906-102378	(90.38 $\pm$ 0.63)%	102894-104263	(88.79 $\pm$ 0.88)%
7	102379-103361	(91.37 $\pm$ 0.87)%		
8	103362 -103686	(88.91 $\pm$ 0.72)%		

 Table A.16 Calibration of  $\pi$  identification from  $\Lambda_c^+$  decay in the  $\Lambda_c^+$  MC.

SubID	MagDown		MagUp	
	Run range	PID $K$ < -5 Efficiency	Run Runge	PID $K$ < -5 Efficiency
2	92821-93224	–	94261-96641	(74.00 $\pm$ 0.19)%
3	93225-97884	(71.84 $\pm$ 0.19)%	96642-100243	(70.55 $\pm$ 0.20)%
4	97885-98198	(69.07 $\pm$ 0.20)%	100244-102505	(69.04 $\pm$ 0.21)%
5	98199 -101905	(70.39 $\pm$ 0.21)%	102506-102893	(70.63 $\pm$ 0.20)%
6	101906-102378	(70.20 $\pm$ 0.20)%	102894-104263	(69.25 $\pm$ 0.35)%
7	102379-103361	(70.22 $\pm$ 0.20)%		
8	103362 -103686	(67.87 $\pm$ 0.22)%		

## Appendix B Calculation of the uncertainty of L0 trigger efficiency

The L0 efficiency is calculated using the trigger efficiency table. For each MC event, the probability of the events being triggered can be estimated from the  $p_T$  distribution of the daughters.

$$\varepsilon_i = 1 - (1 - \varepsilon_p)(1 - \varepsilon_K)(1 - \varepsilon_\pi),$$

where  $\varepsilon_i$  is the probability of the event being triggered,  $\varepsilon_p$ ,  $\varepsilon_K$  and  $\varepsilon_\pi$  is the L0 trigger efficiency for the proton, kaon and pion, respectively. Then the average efficiency is

$$\bar{\varepsilon} = \frac{\sum_i w_i n_i \varepsilon_i}{\sum_i w_i},$$

where  $w_i$  is the lifetime weight,  $n_i$  is the number of events in the  $p_T$  bin, and  $\varepsilon_i$  is the probability that the event in the given  $p_T$  bin is triggered.

The uncertainty of  $\bar{\varepsilon}$  consists of two parts: the systematic part and the statistical part. The systematic part is the uncertainty due to the method itself, mainly due to the trigger efficiency uncertainty of the final state hadron. Since the uncertainties in the trigger efficiency table are small, so is the systematic uncertainty. The statistical part is the uncertainty due to limited MC sample used to estimate the efficiency. The uncertainty can be deduced from the error propagation formula:

$$\frac{\partial \bar{\varepsilon}}{\partial n_i} = \frac{w_i \varepsilon_i \cdot (\sum_j w_j n_j) - (\sum_j w_j n_j \varepsilon_j) \cdot w_i}{(\sum_i w_i n_i)^2}.$$

Since  $n_i$  follows Poisson distribution, we then have

$$\Delta n_i = \sqrt{n_i},$$

and

$$\frac{\partial \bar{\varepsilon}}{\partial \varepsilon_i} = \frac{w_i n_i}{\sum_j w_j n_j}.$$

And the uncertainty of  $\varepsilon_i$  can be known from

$$(\Delta \varepsilon_i)^2 = (1 - \varepsilon_K)^2 (1 - \varepsilon_\pi)^2 (\Delta \varepsilon_p)^2 + (1 - \varepsilon_p)^2 (1 - \varepsilon_\pi)^2 (\Delta \varepsilon_K)^2 + (1 - \varepsilon_p)^2 (1 - \varepsilon_K)^2 (\Delta \varepsilon_\pi)^2$$

Then  $\Delta\bar{\varepsilon}$  can be expressed as

$$\Delta\bar{\varepsilon} = \sqrt{\sum_i \left[ \left( \frac{\partial\bar{\varepsilon}}{\partial n_i} \right)^2 (\Delta n_i)^2 + \left( \frac{\partial\bar{\varepsilon}}{\partial \varepsilon_i} \right)^2 (\Delta \varepsilon_i)^2 \right]}.$$

## Appendix C Uncertainty of the efficiency ratio relative to that of 3500 MeV/c<sup>2</sup>

We weight the  $p_T$  distribution of the  $\Lambda_c^+$  baryon, the bachelor  $K$  and  $\pi$  mesons to estimate the efficiencies at 3300 MeV/c<sup>2</sup> and 3700 MeV/c<sup>2</sup>. The efficiency ratio is

$$r = \frac{\frac{\sum_i w_i n_i}{\sum_i w_i (n_i + n'_i)}}{\frac{\sum_i n_i}{\sum_i (n_i + n'_i)}},$$

where  $i$  is the index of the histogram (note we have used  $i$  to denote the three-dimensional indexes to simplify the expression),  $w_i$ ,  $n_{ijk}$  and  $n'_{ijk}$  are the weight, the number of events that passes the selection and the number of events that fails to pass the selection of the  $i$ -th bin, respectively.  $n_i$  and  $n'_i$  are considered to follow Poisson fluctuation, the uncertainty due to this fluctuation can be deduced from the uncertainty propagation formula:

$$\begin{aligned} \frac{\partial r}{\partial n_i} &= \frac{\frac{\sum_i w_i n'_i \cdot w_i}{[\sum_i w_i (n_i + n'_i)]^2} \cdot \frac{\sum_i n_i}{\sum_i (n_i + n'_i)} - \frac{\sum_i w_i n_i}{\sum_i w_i (n_i + n'_i)} \cdot \frac{\sum_i n'_i}{[\sum_i (n_i + n'_i)]^2}}{\left[ \frac{\sum_i n_i}{\sum_i (n_i + n'_i)} \right]^2} \\ &= \frac{sw' \cdot w_i \cdot n(n+n') - sw \cdot n' \cdot (sw + sw')}{n^2 \cdot (sw + sw')^2}. \end{aligned}$$

Similarly,

$$\frac{\partial r}{\partial n'_i} = \frac{-sw \cdot w_i \cdot n(n+n') + sw \cdot n \cdot (sw + sw')}{n^2 \cdot (sw + sw')^2}.$$

where  $sw$  is the weight sum for the events that pass the selection,  $sw'$  is the weight sum for the events that fail to pass the selection,  $n$  is the total number of events that pass the selection, and  $n'$  is the total number of events that fail to pass the selection.

Then the total uncertainty is

$$\Delta r = \sqrt{\sum_i \left[ \left( \frac{\partial r}{\partial n_i} \right)^2 n_i + \left( \frac{\partial r}{\partial n'_i} \right)^2 n'_i \right]}.$$

## Appendix D Background Subtraction using 25-Tiles method

### D.1 Introduction

In this appendix, the expressions needed for the 25-Tiles background subtraction are derived. First, we introduce the problem. Suppose that there are two components of the background: a combinatoric component described by a function  $f$ , which is smooth in  $\delta m$  and  $m(\Lambda_c^+)$ , and a  $\Lambda_c^+$  component described by a function  $g$ , which is smooth in  $\delta m$  but sharply peaked in  $m(\Lambda_c^+)$ . We also assume that  $f$  factorizes into functions of  $\delta m$  and  $m(\Lambda_c^+)$ , i.e. that its  $\delta m$  projection is independent of  $m(\Lambda_c^+)$ . With these assumptions, we will count background events in the 2D sidebands and use those to estimate the background projection in the signal box. The sidebands and signal box are illustrated in Figure D.1, along with the parameters  $(a, b, c, d, \alpha, \beta, \gamma, \delta)$ , which define the grid. By construction the corner boxes are all of equal size, i.e.  $d - c = c - b$  and  $\delta - \gamma = \gamma - \beta$ .

In the following derivation,  $x$  will be the mass of the  $\Lambda_c^+$  candidate (but offset so that 0 is at the expected  $\Lambda_c^+$  mass), and  $y$  will be  $\delta m$  (but offset so that 0 is at the expected value).

### D.2 Lemma: combinatorial background

Consider a function  $f(x, y)$  which is smooth in  $x$  and  $y$  and whose  $x$  and  $y$  projections are uncorrelated. Then the second-order Taylor expansion of  $f$  is given by:

$$f(x, y) \simeq f_0 + \left(\frac{df}{dx}\right)_0 x + \left(\frac{d^2f}{dx^2}\right)_0 \frac{x^2}{2} + \left(\frac{df}{dy}\right)_0 y + \left(\frac{d^2f}{dy^2}\right)_0 \frac{y^2}{2}$$

where  $f_0 = f(0, 0)$  and the derivatives are evaluated at  $(0, 0)$ . Then the integral over a box in the  $xy$  plane ( $x_0 < x < x_1, y_0 < y < y_1$ ) is given by:

$$\begin{aligned} I &= \int_{y=y_0}^{y=y_1} \int_{x=x_0}^{x=x_1} f_0 + \left(\frac{df}{dx}\right)_0 x + \left(\frac{d^2f}{dx^2}\right)_0 \frac{x^2}{2} + \left(\frac{df}{dy}\right)_0 y + \left(\frac{d^2f}{dy^2}\right)_0 \frac{y^2}{2} dx dy \\ &= \int_{y=y_0}^{y=y_1} f_0 [x_1 - x_0] + \left(\frac{df}{dx}\right)_0 \frac{[x_1^2 - x_0^2]}{2} + \left(\frac{d^2f}{dx^2}\right)_0 \frac{[x_1^3 - x_0^3]}{6} + \left(\frac{df}{dy}\right)_0 y [x_1 - x_0] \end{aligned}$$

$$\begin{aligned}
 & + \left( \frac{d^2 f}{dy^2} \right)_0 \frac{y^2}{2} [x_1 - x_0] dy \\
 = & f_0 [x_1 - x_0] [y_1 - y_0] + \left( \frac{df}{dx} \right)_0 \frac{[x_1^2 - x_0^2]}{2} [y_1 - y_0] + \left( \frac{d^2 f}{dx^2} \right)_0 \frac{[x_1^3 - x_0^3]}{6} [y_1 - y_0] \\
 & + \left( \frac{df}{dy} \right)_0 [x_1 - x_0] \frac{[y_1^2 - y_0^2]}{2} + \left( \frac{d^2 f}{dy^2} \right)_0 [x_1 - x_0] \frac{[y_1^3 - y_0^3]}{6}
 \end{aligned}$$

We now consider four particular cases: signal box, corner box, vertical sidebands and horizontal sidebands.

### D.2.1 Signal box

For a signal box centered at the origin,  $(-a < x < a, -\alpha < y < \alpha)$ , the integral is:

$$\begin{aligned}
 I_{\text{sig}} & = f_0(2a)(2\alpha) + \left( \frac{d^2 f}{dx^2} \right)_0 \frac{(2a^2)(2\alpha)}{6} + \left( \frac{d^2 f}{dy^2} \right)_0 \frac{(2a)(2\alpha^3)}{6} \\
 & = 4a\alpha f_0 + \frac{4a^3\alpha}{3} \left( \frac{d^2 f}{dx^2} \right)_0 + \frac{4a\alpha^3}{3} \left( \frac{d^2 f}{dy^2} \right)_0.
 \end{aligned}$$

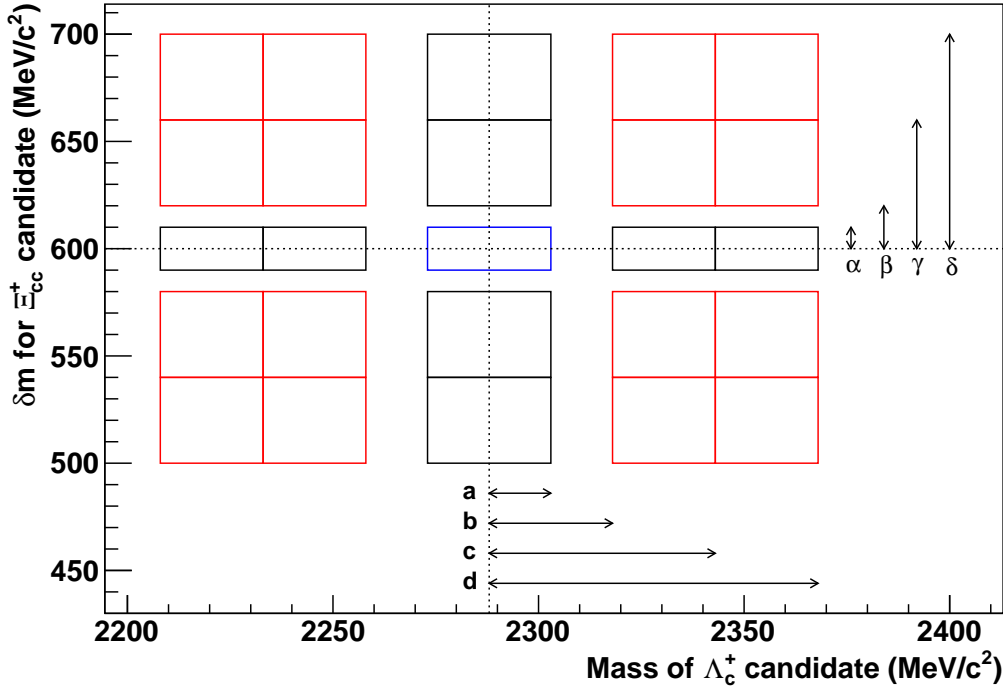


Figure D.1 Definitions of signal and sideband boxes for the 25-Tiles method. The numerical values can vary—in particular, the central point in  $\delta m$  will be moved in steps.

### D.2.2 Corner box

For a corner box spanning  $(b < |x| < c, \beta < |y| < \gamma)$ , the integral is:

$$I_{\text{corner}} = f_0(c-b)(\alpha-\beta) + (\pm) \left( \frac{df}{dx} \right)_0 \frac{(c^2-b^2)(\gamma-\beta)}{2} + \left( \frac{d^2f}{dx^2} \right)_0 \frac{(c^3-b^3)(\gamma-\beta)}{6} \\ + (\pm) \left( \frac{df}{dy} \right)_0 \frac{(c-b)(\gamma^2-\beta^2)}{2} + \left( \frac{d^2f}{dy^2} \right)_0 \frac{(c-b)(\gamma^3-\beta^3)}{6},$$

where the first  $\pm$  is positive for  $x > 0$  and negative for  $x < 0$ , and the second  $\pm$  is positive for  $y > 0$  and negative for  $y < 0$ . The sum of the four corners is thus:

$$\sum I_{\text{corner}} = 4f_0(c-b)(\gamma-\beta) + \left( \frac{d^2f}{dx^2} \right)_0 \frac{2(c^3-b^3)(\gamma-\beta)}{3} + \left( \frac{d^2f}{dy^2} \right)_0 \frac{2(c-b)(\gamma^3-\beta^3)}{3}.$$

### D.2.3 Vertical sidebands

For a vertical sideband spanning  $(-a < x < a, \beta < |y| < \gamma)$ , the integral is:

$$I_{\text{vert}} = f_0(2a)(\gamma-\beta) + \left( \frac{d^2f}{dx^2} \right)_0 \frac{(2a^3)(\gamma-\beta)}{6} \pm \left( \frac{df}{dy} \right)_0 \frac{(2a)(\gamma^2-\beta^2)}{2} + \left( \frac{d^2f}{dy^2} \right)_0 \frac{(2a)(\gamma^3-\beta^3)}{6} \\ = 2f_0a(\gamma-\beta) + \left( \frac{d^2f}{dx^2} \right)_0 \frac{a^3(\gamma-\beta)}{3} \pm \left( \frac{df}{dy} \right)_0 a(\gamma^2-\beta^2) + \left( \frac{d^2f}{dy^2} \right)_0 \frac{a(\gamma^3-\beta^3)}{3},$$

where the  $\pm$  is positive for  $y > 0$  and negative for  $y < 0$ . Thus, the sum of the top and bottom sidebands is:

$$\sum I_{\text{vert}} = 4f_0a(\gamma-\beta) + \left( \frac{d^2f}{dx^2} \right)_0 \frac{2a^3(\gamma-\beta)}{3} + \left( \frac{d^2f}{dy^2} \right)_0 \frac{2a(\gamma^3-\beta^3)}{3}.$$

### D.2.4 Horizontal sidebands

Similarly, for a horizontal sideband spanning  $(b < |x| < c, -\alpha < y < \alpha)$ , the integral is:

$$I_{\text{horiz}} = 2f_0\alpha(c-b) + \left( \frac{d^2f}{dy^2} \right)_0 \frac{\alpha^3(c-b)}{3} \pm \left( \frac{df}{dx} \right)_0 \alpha(c^2-b^2) + \left( \frac{d^2f}{dx^2} \right)_0 \frac{\alpha(c^3-b^3)}{3},$$

where the  $\pm$  is positive for  $x > 0$  and negative for  $x < 0$ , and the sum of the left and right sidebands is:

$$\sum I_{\text{horiz}} = 4f_0\alpha(c-b) + \left( \frac{d^2f}{dy^2} \right)_0 \frac{2\alpha^3(c-b)}{3} + \left( \frac{d^2f}{dx^2} \right)_0 \frac{2\alpha(c^3-b^3)}{3}.$$

### D.3 Lemma: $\Lambda_c^+$ background

We now consider the  $\Lambda_c^+$  background, which is a simplified version of the above, in which the distribution is smooth in  $y$  but sharply peaked in  $x$  so that everything falls within the interval  $-a < x < a$ . Then we only need to know the  $y$  projection of the distribution,  $g(y)$ . We find the integral in the signal box to be:

$$(I_{\text{sig}})_{\Lambda_c} = 4\alpha g_0 + \frac{4\alpha^3}{3} \left( \frac{d^2 g}{dy^2} \right)_0.$$

Note that similar to the term in  $d^2 g/dx^2$ , the factors of  $a$  have disappeared compared to the combinatoric version: this is because the width of the signal box in  $x$  is irrelevant as long as it is wider than the  $\Lambda_c^+$  lineshape.

The integral in the vertical sidebands is:

$$(I_{\text{vert}})_{\Lambda_c} = 2g_0(\gamma - \beta) \pm \left( \frac{dg}{dy} \right)_0 (\gamma^2 - \beta^2) + \left( \frac{d^2 g}{dy^2} \right)_0 \frac{(\gamma^3 - \beta^3)}{3},$$

where the  $\pm$  is positive for  $y > 0$  and negative for  $y < 0$ , and hence the sum of the top and bottom sidebands is:

$$\left( \sum I_{\text{vert}} \right)_{\Lambda_c} = 4g_0(\gamma - \beta) + \left( \frac{d^2 g}{dy^2} \right)_0 \frac{2(\gamma^3 - \beta^3)}{3}.$$

The integral of the  $\Lambda_c^+$  background in the horizontal sideband is zero by definition, as is the integral in the corner boxes.

### D.4 Combining the above results

The goal is to obtain the background estimate in the signal box, *i.e.* the sum of the integrals of the combinatoric and  $\Lambda_c^+$  backgrounds components in the central box. These are given by:

$$(I_{\text{sig}})_{\text{comb}} = 4a\alpha f_0 + \frac{4a^3\alpha}{3} \left( \frac{d^2 f}{dx^2} \right)_0 + \frac{4a\alpha^3}{3} \left( \frac{d^2 f}{dy^2} \right)_0$$

and:

$$(I_{\text{sig}})_{\Lambda_c^+} = 4\alpha g_0 + \frac{4\alpha^3}{3} \left( \frac{d^2 g}{dy^2} \right)_0,$$

and their sum is:

$$(I_{\text{sig}})_{\text{comb}+\Lambda_c^+} = 4\alpha(af_0 + g_0) + \frac{4a^3\alpha}{3} \left( \frac{d^2 f}{dx^2} \right)_0 + \frac{4\alpha^3}{3} \left[ a \left( \frac{d^2 f}{dy^2} \right)_0 + \left( \frac{d^2 g}{dy^2} \right)_0 \right]. \quad (\text{D-1})$$



Now, consider the sum of the integrals in the near vertical sidebands (i.e.  $\beta < |y| < \gamma$ ) and in the far vertical sidebands (i.e.  $\gamma < |y| < \delta$ ), summing  $x > 0$  and  $x < 0$ :

$$\begin{aligned}\sum (I_{\text{vert,near}})_{\text{comb}+\Lambda_c^+} &= 4f_0a(\gamma - \beta) + \left(\frac{d^2f}{dx^2}\right)_0 \frac{2a^3(\gamma - \beta)}{3} \\ &\quad + \left(\frac{d^2f}{dy^2}\right)_0 \frac{2a(\gamma^3 - \beta^3)}{3} + 4g_0(\gamma - \beta) + \left(\frac{d^2g}{dy^2}\right)_0 \frac{2(\gamma^3 - \beta^3)}{3}, \\ \sum (I_{\text{vert,far}})_{\text{comb}+\Lambda_c^+} &= 4(f_0a + g_0)(\delta - \gamma) + \left(\frac{d^2f}{dx^2}\right)_0 \frac{2a^3(\delta - \gamma)}{3} \\ &\quad + \left(\frac{d^2f}{dy^2}\right)_0 \frac{2a(\delta^3 - \gamma^3)}{3} + \left(\frac{d^2g}{dy^2}\right)_0 \frac{2(\delta^3 - \gamma^3)}{3}.\end{aligned}$$

Using the fact that the sidebands are of equal width so that  $\delta - \gamma = \gamma - \beta$ , the difference is:

$$\begin{aligned}\sum (I_{\text{vert,far}})_{\text{comb}+\Lambda_c^+} - \sum (I_{\text{vert,near}})_{\text{comb}+\Lambda_c^+} &= \left[ \left(\frac{d^2f}{dy^2}\right)_0 \frac{2a(\delta^3 - \gamma^3)}{3} + \left(\frac{d^2g}{dy^2}\right)_0 \frac{2(\delta^3 - \gamma^3)}{3} \right] \\ &\quad - \left[ \left(\frac{d^2f}{dy^2}\right)_0 \frac{2a(\gamma^3 - \beta^3)}{3} + \left(\frac{d^2g}{dy^2}\right)_0 \frac{2(\gamma^3 - \beta^3)}{3} \right] \\ &= \frac{2}{3}(\delta^3 - 2\gamma^3 + \beta^3) \left[ a \left(\frac{d^2f}{dy^2}\right)_0 + \left(\frac{d^2g}{dy^2}\right)_0 \right].\end{aligned}$$

This allows us to evaluate

$$a \left(\frac{d^2f}{dy^2}\right)_0 + \left(\frac{d^2g}{dy^2}\right)_0$$

in terms of known or measured quantities. Note that this term appears in Eq. D-1.

We can form a similar expression for the horizontal sidebands, and it is simpler because the  $\Lambda_c^+$  background doesn't contribute:

$$\begin{aligned}\sum (I_{\text{horiz,far}})_{\text{comb}} - \sum (I_{\text{horiz,near}})_{\text{comb}} &= \left(\frac{d^2f}{dx^2}\right)_0 \frac{2\alpha(d^3 - c^3)}{3} - \left(\frac{d^2f}{dx^2}\right)_0 \frac{2\alpha(c^3 - b^3)}{3} \\ &= \frac{2}{3}(d^3 - 2c^3 + b^3) \alpha \left(\frac{d^2f}{dy^2}\right)_0.\end{aligned}$$

This allows us to evaluate  $\left(\frac{d^2f}{dy^2}\right)_0$  in terms of known or measured quantities. Note that this term also appears in Eq. D-1.

We now go back to the vertical sidebands, but this time take the sum of the near and far components instead of their difference:

$$\sum (I_{\text{vert,far}})_{\text{comb}+\Lambda_c^+} + \sum (I_{\text{vert,near}})_{\text{comb}+\Lambda_c^+} = 4(f_0a + g_0)(\delta - \beta) + \left(\frac{d^2f}{dx^2}\right)_0 \frac{2a^3(\delta - \beta)}{3}$$

$$\begin{aligned}
 & + \left( \frac{d^2 f}{dy^2} \right)_0 \frac{2a(\delta^3 - \beta^3)}{3} + \left( \frac{d^2 g}{dy^2} \right)_0 \frac{2(\delta^3 - \beta^3)}{3} \\
 = & 4(f_0 a + g_0)(\delta - \beta) + \left( \frac{d^2 f}{dx^2} \right)_0 \frac{2a^3(\delta - \beta)}{3} \\
 & + \frac{2(\delta^3 - \beta^3)}{3} \left[ a \left( \frac{d^2 f}{dy^2} \right)_0 + \left( \frac{d^2 g}{dy^2} \right)_0 \right].
 \end{aligned}$$

Note that we have already evaluated the derivatives, and therefore can use this expression to evaluate  $(af_0 + g_0)$ .

At this point we have enough information to evaluate all terms in Eq. D-1. However, we can obtain some more information by using the corner tiles. The inner half of the corner tiles defined by  $(b < |x| < d, \beta < |y| < \gamma)$  sums to:

$$\begin{aligned}
 \left( \sum I_{\text{corner, near}} \right)_{\text{comb}} & = 4f_0(d - b)(\gamma - \beta) + \left( \frac{d^2 f}{dx^2} \right)_0 \frac{2(d^3 - b^3)(\gamma - \beta)}{3} \\
 & + \left( \frac{d^2 f}{dy^2} \right)_0 \frac{2(d - b)(\gamma^3 - \beta^3)}{3}
 \end{aligned}$$

and likewise the outer half defined by  $(b < |x| < d, \gamma < |y| < \delta)$  sums to:

$$\begin{aligned}
 \left( \sum I_{\text{corner, far}} \right)_{\text{comb}} & = 4f_0(d - b)(\delta - \gamma) + \left( \frac{d^2 f}{dx^2} \right)_0 \frac{2(d^3 - b^3)(\delta - \gamma)}{3} \\
 & + \left( \frac{d^2 f}{dy^2} \right)_0 \frac{2(d - b)(\delta^3 - \gamma^3)}{3}.
 \end{aligned}$$

The sum of these two terms evaluates to:

$$\begin{aligned}
 \left( \sum I_{\text{corner, far}} \right)_{\text{comb}} + \left( \sum I_{\text{corner, near}} \right)_{\text{comb}} & = 4f_0(d - b)(\delta - \beta) + \left( \frac{d^2 f}{dx^2} \right)_0 \frac{2(d^3 - b^3)(\delta - \beta)}{3} \\
 & + \left( \frac{d^2 f}{dy^2} \right)_0 \frac{2(d - b)(\delta^3 - \beta^3)}{3}
 \end{aligned}$$

and the difference to:

$$\left( \sum I_{\text{corner, far}} \right)_{\text{comb}} - \left( \sum I_{\text{corner, near}} \right)_{\text{comb}} = \left( \frac{d^2 f}{dy^2} \right)_0 \frac{2(d - b)}{3} (\delta^3 - 2\gamma^3 + \beta^3).$$

Combining these expressions with those previous, we can deduce  $f_0$  and  $g_0$  separately.

## Resume and publications

### 个人简历

1987年8月出生於福建省連江縣。

2005年9月考入清華大學工程物理系工程物理專業，2009年7月本科畢業並獲得工學學士學位。

2009年9月免試進入清華大學工程物理系攻讀物理學博士學位至今。

### 發表的學術論文

- [1] The LHCb Collaboration, R. Aaij *et al.*, Search for the doubly charmed baryon  $\Xi_{cc}^+$ , JHEP 12 (2013) 090, arXiv:1310.2538[hep-ex]
- [2] Liang Zhong (on behalf of the LHCb Collaboration), *b* and *c* hadron production and spectroscopy at LHCb, PoS (Hadron2013) 069
- [3] Liang Zhong (on behalf of the LHCb Collaboration), Latest LHCb results from the *pA* data, J.Phys.:Conf.Ser.509 012020