Radiation damage studies in the LHCb VELO detector and searches for lepton flavour and baryon number violating tau decays

A thesis

submitted to The University of Manchester

for the degree of

doctor of philosophy (PhD)

in the faculty of Engineering and Physical Sciences

Jonathan Robert Harrison

Particle Physics Group

School of Physics and Astronomy

2014

The University of Manchester

CONTENTS

Final word count: 31420

ABSTRACT

This thesis presents work carried out using data from the LHCb experiment during the first three years of data taking, 2010 – 2012. A study of the effects of radiation damage on the silicon sensors of the LHCb Vertex Locator is performed, with an emphasis on the implications for the long term performance of the detector. Following three years of operation the sensors have received a maximum delivered neutron equivalent fluence of $\sim 1.6 \times 10^{12}$ per square centimeter, leading to a number of radiation induced effects. In particular the change in charge collection efficiency and signal/noise with fluence is compared to theoretical expectations, and the current trends are extrapolated to the fluences expected at the end of the LHCb detector lifetime. The development of an unexpected effect due to the structure of the routing lines in the sensors is described in detail.

Searches for lepton flavour and baryon number violating decays of the τ lepton using the 2011 LHCb dataset are described. Observation of any lepton flavour or baryon number violation would be an unambiguous sign of new physics, whilst setting improved limits helps to constrain a number of Beyond the Standard Model theories. First LHCb limits are set on the branching fractions of the decays $\tau^- \to \mu^-\mu^+\mu^-$, $\tau^- \to \overline{p}\mu^+\mu^-$ and $\tau^- \to p\mu^-\mu^-$, with these results also representing the first limits on lepton flavour violating tau decays at a hadron collider. The limit on $\tau^- \to \mu^-\mu^+\mu^-$ is expected to approach the world's best result from Belle in the coming years whilst the $\tau^- \to \overline{p}\mu^+\mu^-$ and $\tau^- \to p\mu^-\mu^-$ results constitute the first limits on the branching fractions of these decays. The future prospects for these measurements with further data are briefly described.

DECLARATION

The University of Manchester

PhD by published work Candidate Declaration

Candidate Name: Jonathan Robert Harrison

Faculty: Engineering and Physical Sciences

Thesis Title: Radiation damage studies in the LHCb VELO detector and searches for lepton flavour and baryon number violating tau decays

Declaration to be completed by the candidate:

This work represents the combined efforts of the LHCb collaboration. Some of the content has been published elsewhere and/or presented to several audiences. No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Signed: Date: June 25, 2014

COPYRIGHT

ACKNOWLEDGEMENTS

The completion of this thesis, and my PhD in general, has been a thoroughly enjoyable experience from start to finish. It would not have been possible without a number of people who I would like to thank:

For the tireless support, advice, proofreading and all-round fantastic supervision: George.

For the guidance, patience and help with various aspects of my PhD studies: Chris, Johannes, Paula and Lars.

For the discussions, hard work and camaraderie required to meet the many deadlines and produce successful analyses: Paul and Marcin.

For showing me the ropes and being on hand to answer my endless questions about LHCb, physics and computing: Grant, Marius, Marco, Dermot, Tom and Alex.

For "taking me under his wing": Adam.

For making my time based at CERN a truly memorable experience: Kara, Espen, Ed, Sarah, Ewen, Clara and Liz.

For inspiring me to push myself that extra mile, whether on grass, tarmac or snow: Nick, Mika and Joel.

For squeezing every last penny out of the travel budget and providing excellent coffee time entertainment: Fred.

For all things Mancunian and administrative: Anne and Matthew.

For being brilliant housemates: Simon B, Mike and Simon O.

And last, but by no means least, for the encouragement and the support that has enabled me to get this far: my parents.

"Difficulties are just things to overcome, after all."

Ernest Shackleton

1 INTRODUCTION

Particle physics is the study of the fundamental constituents of matter and the forces acting between them. Particle physicists attempts to answer a number of questions about the nature of the universe by discovering the smallest building blocks and describing how they behave.

Whilst a great deal of work is underway in both the theoretical and experimental areas of the field, this thesis will largely focus on experimental data from the LHCb experiment at the Large Hadron Collider (LHC) at CERN. At the LHC, beams of particles are collided together at high energies with the aim of learning something about the underlying physical processes. The LHCb detector is one of several located at the LHC that examine the debris resulting from these collisions. Through its many unique features it enables the study of a wide range of both well established and predicted, but so far unobserved, phenomena.

Particle physics also provides a rich environment for research and development at the frontiers of a number of related fields such as computing, engineering and instrumentation. In particular, the detectors themselves include a large number of novel technologies that are subject to unique conditions. Therefore a portion of this thesis will be dedicated to the study of silicon strip technology in the extreme radiation environment of the LHC.

Chapter 2 provides a theoretical motivation for the experimental studies that follow. A brief overview of our current understanding of particle physics is

given, followed by expectations for physics beyond the Standard Model and how this may manifest itself in the decays of tau leptons, which can be searched for at LHCb.

Chapter 3 describes both hardware and software aspects of the LHCb detector, with an emphasis on their contribution to the overall physics performance.

Chapter 4 introduces the concept of particle tracking in semiconductor detectors and describes the LHCb Vertex Locator in more detail. It then describes studies to understand the current and future performance as a result of radiation damage from several years of operation.

Chapter 5 details experimental searches for both lepton flavour violation and baryon number violation in the decays of tau leptons at LHCb with the dataset collected in 2011. Future prospects at both LHCb and upcoming experiments are discussed.

Appendix A describes updates to the searches for lepton flavour violating tau decays at LHCb using the full dataset collected before the first long shutdown of the LHC.

Throughout this thesis the convention $c = 1$ is used and charge conjugation is implied.

2 THEORETICAL REVIEW

In this chapter the theoretical motivation is described for the studies of lepton flavour violating and baryon number violating τ decays presented in Chapter 5. Section 2.1 gives a brief overview of the Standard Model of particle physics. Section 2.2 describes theoretical predictions for the branching fractions of lepton flavour violating decays within the Standard Model and in various Beyond the Standard Model scenarios. Section 2.3 motivates the need for baryon number violation, and suggests various ways in which this could manifest itself in the decays of τ leptons.

2.1 THE STANDARD MODEL OF PARTICLE PHYSICS

Our current understanding of particle physics is based on a quantum field theory called the Standard Model (SM). It describes three of the four fundamental forces of nature — the electromagnetic, weak and strong interactions — but does not include gravity.

The SM is based on a number of symmetries and their associated conservation laws which result as a consequence of the Noether theorem [1]. For example, the electromagnetic interaction is unchanged, or 'invariant', under a global phase change, and as a result the net quantity of electric charge is always conserved. A quantum mechanical system can be defined in terms of the values of these conserved quantities, or 'quantum numbers', such as the electric charge, Q, colour charge and weak isospin, I_3. A number of approximate conservation laws also exist which do not have an underlying symmetry. The values of the

associated conserved quantities are described by 'internal quantum numbers', which are described in more detail in the following.

The particle content of the SM consists of 'fermions' (particles with half-integer spin) and 'bosons' (particles with integer spin), with the bosons, mediating the various interactions. The fermions can be classified into either 'quarks' or 'leptons' of three generations, which, along with the gauge bosons, are briefly described in the following.

2.1.1 Quarks

The SM quarks are elementary particles which experience all four fundamental forces. Quarks come in six types, or 'flavours': *up, down, charm, strange, top* and *bottom*, with a corresponding 'anti-quark' for each flavour. In the following they are referred to as u, d, c, s, t and b. The anti-quarks have the same mass as their respective quark, but have opposite charge and internal quantum numbers, and are denoted by an overline, such as $\overline{u}$ for the anti-up quark.

The quarks form three generations of increasing mass, with one 'up-type' and one 'down-type' quark in each generation or 'doublet',

$$\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix}.$$

Several of their properties are listed in Table 1, where C, S, T and B are the charm, strangeness, topness and bottomness internal quantum numbers respectively. These 'flavour quantum numbers' are additive, such that their sum is conserved in interactions. They are conserved under the electromagnetic and strong interactions, but not under the weak interaction. As a result of the strong interaction, quarks combine to form composite particles or 'hadrons', in either two-quark ('meson') or three-quark ('baryon') configurations.

Particle	Q	Mass [MeV]	I_3	C	S	T	B
u	$+2/3$	$2.3^{+0.7}_{-0.5}$	$+1/2$	0	0	0	0
d	$-2/3$	$4.8^{+0.7}_{-0.3}$	$-1/2$	0	0	0	0
c	$+2/3$	1275 ± 25	$+1/2$	$+1$	0	0	0
s	$-2/3$	95 ± 5	$-1/2$	0	-1	0	0
t	$+2/3$	$(160^{+5}_{-4}) \times 10^3$	$+1/2$	0	0	$+1$	0
b	$-2/3$	4180 ± 30	$-1/2$	0	0	0	-1

Table 1: Quark content of the SM. The c, b and t quark masses are according to the so-called $\overline{MS}$ scheme [2]. The particle masses are taken from Ref. [3].

2.1.2 Leptons

Similarly to the quarks, the SM leptons come in six flavours: *electron, electron neutrino, muon, muon neutrino, tau,* and *tau neutrino*, with a corresponding 'anti-lepton' for each flavour. In the following they are referred to as $e, \nu_e, \mu, \nu_\mu, \tau$ and ν_τ. The charged leptons (e, μ, τ) are defined as having negative charge, whilst the anti-leptons have positive charge.

The leptons form three generations, with one 'charged lepton' and one neutrino in each doublet,

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix}, \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}, \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}.$$

Several of their properties are listed in Table 2, where L_e, L_μ and L_τ are the electron number, the muon number and the tau number respectively. Like the flavour quantum numbers of the quarks, lepton flavours are internal quantum numbers and are additive. An additional 'lepton number' internal quantum number, L, can be defined as the total number of leptons of any generation minus the total number of anti-leptons of any generation. Lepton number is

Particle	Q	Mass [MeV]	L_e	L_μ	L_τ
e^-	-1	0.510998928(11)	$+1$	0	0
ν_e	0	$< 2 \times 10^{-6}$	$+1$	0	0
μ^-	-1	105.6583715(35)	0	$+1$	0
ν_μ	0	< 0.19	0	$+1$	0
τ^-	-1	1776.82(16)	0	0	$+1$
ν_τ	0	< 18.2	0	0	$+1$

Table 2: Lepton content of the SM. The particle masses are taken from Ref. [3].

always conserved in SM interactions. Unlike the quarks, the leptons do not form composite particles as they do not participate in the strong interaction. The neutrinos participate in neither the strong interaction nor the electromagnetic interaction, and thus only interact weakly in the SM.

2.1.3 Bosons

The gauge bosons of the SM are listed in Table 3. The exchange of photons, γ, mediates the electromagnetic interaction, whilst the exchange of the Z and $W^\pm$ bosons mediates the weak interaction. The exchange of gluons, g, of which there are eight types, mediates the strong interaction. The Higgs boson, H, which was finally confirmed to exist in 2012 [4, 5], was the last particle of the SM to be discovered. It is responsible for giving mass to the other SM particles.

Particle	Q	Mass [GeV]	Spin
γ	0	0	1
Z	0	91.1876 ± 0.0021	1
$W^{\pm}$	± 1	80.385 ± 0.015	1
g	0	0	1
H	0	125.9 ± 0.4	0

Table 3: Boson content of the SM. The particle masses are taken from Ref. [3].

2.1.4 The Weak Interaction and Charge-Parity

The 'helicity' of a particle is defined as the normalised component of its spin, along its direction of flight,

$$h \equiv \frac{\mathbf{S} \cdot \mathbf{p}}{p},$$

where $\mathbf{S}$ and $\mathbf{p}$ are the spin and momentum vectors respectively. For a spin-half particle, if the spin and momentum are in the same direction, the particle state is referred to as 'right-handed' ($h = +1/2$). If instead, the spin and momentum are in the opposite direction the particle state is referred to as 'left-handed' ($h = -1/2$). 'Chirality' is the Lorentz-invariant equivalent of helicity, and is the same in the limit $E \gg m$, or for massless particles.

The SM weak interaction is a 'chiral' interaction, as it only couples to left-handed chiral particle states and right-handed anti-particle states. If a change of sign of all three spatial coordinates, or 'parity inversion', is applied the momentum of a particle changes sign, but its spin does not. As a result a right-handed particle changes to a left-handed particle and vice-versa. Under 'charge-conjugation', or the exchange or particle and anti-particle, the handedness is invariant. Therefore under the combined charge-parity operation, or CP, a right-

handed particle changes to a left-handed anti-particle and vice-versa[1]. As the SM neutrinos only interact via the weak interaction, the right-handed neutrinos and the left-handed anti-neutrinos do not interact via any of the fundamental interactions of the SM. This allows for a number of theoretical interpretations, several of which are discussed in Section 2.2.2.

2.2 LEPTON FLAVOUR VIOLATION

2.2.1 Flavour Violation in the Standard Model

In the weak charged-current interaction [6], the up quark is coupled to a superposition of down type quarks, d', via a term proportional to

$$d' = \cos\theta_c d + \sin\theta_c s,$$

where θ_c is the Cabibbo angle [7]. Similarly the charm quark is coupled to a superposition of down type quarks, s', via a term proportional to

$$s' = -\sin\theta_c d + \cos\theta_c s.$$

These two equations can be expressed in matrix notation as

$$\begin{bmatrix} d' \\ s' \end{bmatrix} = \begin{bmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{bmatrix} \begin{bmatrix} d \\ s \end{bmatrix},$$

yielding a 2×2 unitary rotation matrix called the 'Cabibbo matrix'. Expanding this matrix to include the top and bottom quarks yields the 'Cabibbo-Kobayashi-Maskawa matrix' [8], or CKM matrix,

$$\begin{bmatrix} d' \\ s' \\ b' \end{bmatrix} = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} \begin{bmatrix} d \\ s \\ b \end{bmatrix},$$

1 Whilst the electromagnetic and the strong interactions are CP symmetric, CP symmetry is violated in the quark sector of the weak interaction.

where $|V_{ij}|^2$ is proportional to the probability for a quark of flavour j to couple to a quark of flavour i via the weak interaction. The full weak charged-current is given by the commutator of a term containing the CKM matrix and a term containing its Hermitian conjugate, such that the resulting matrix is diagonal in flavour space. The neutral-current, which is a linear superposition of the weak charged-current and the electromagnetic current is also diagonal. Therefore to first-order, or at 'tree level', there are no flavour changing neutral-currents (FCNC) in quark decays in the SM.

FCNC are allowed at the 'one-loop' level, as shown in Fig. 1 (left) for the $K^0 \to \mu^+\mu^-$ decay, which involves an s to d transition. However, FCNC are highly suppressed, as according to the 'Glashow-Iliopoulos-Maini mechanism', or GIM mechanism [9], there is a second diagram involving the c quark, as shown in Fig. 1 (right). If the flavour symmetry of the SM was exact these two diagrams would exactly cancel, and there would be no FCNC at one-loop level either. As this is not the case, the sum of the two diagrams is approximately $\alpha^2(m_c^2 - m_u^2)/m_W^2 \sim \alpha^2 m_c^2/m_W^2$, where α is the fine structure constant, and m_u, m_c and m_W are the masses of the u quark, the c quark, and the W bosons respectively.

Whilst quark flavour violating decays, such as the decay $K_L \to \mu^+\mu^-$, have been experimentally confirmed [3], FCNC in the leptons, and hence lepton

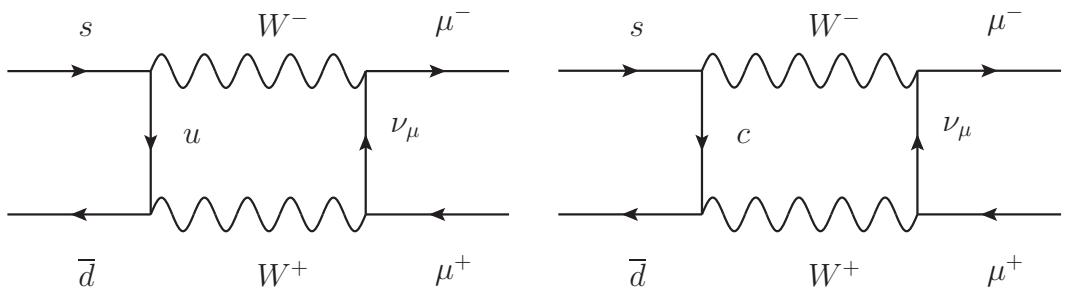


Figure 1: Left: One-loop contribution to $K^0 \to \mu^+\mu^-$ in the SM from the u quark. Right: One-loop contribution to $K^0 \to \mu^+\mu^-$ in the SM from the c quark.

flavour violation (LFV), was unexpected due to the apparent zero mass of the neutrinos. The observation of the effect of neutrino oscillations at the Super-Kamiokande experiment [10] was the first evidence of LFV, and in order to incorporate these results (and their subsequent confirmation by other experiments such as K2K [11] and SNO [12]) the SM can be extended to include neutrino masses [13]. Within this framework neutrinos can 'oscillate' from one flavour to another as the 'flavour eigenstates', $\nu_{e,\mu,\tau}$, which participate in the weak interaction are a mixture of the 'mass eigenstates', $\nu_{1,2,3}$, which correspond to the fundamental particles. The magnitude of the mixing is defined by the Pontecorvo-Maki-Nakagawa-Sakata matrix [14], or PMNS matrix,

$$\begin{bmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{bmatrix} = \begin{bmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{bmatrix},$$

where $|U_{li}|^2$ is proportional to the probability of a neutrino of flavour l to be found in mass eigenstate i. The PMNS matrix for lepton mixing can be viewed as analogous to the CKM matrix for quark mixing [15], and as it is also unitary, FCNC in the lepton sector are forbidden at tree level. At the one-loop level FCNC are again suppressed via the GIM mechanism, but the suppression is significantly larger due to the small size of the neutrino mass. The charged LFV decay $\tau^- \to \mu^-\mu^+\mu^-$, which is the subject of Chapter 5, is then allowed via the one-loop level process shown in Fig. 2. However, the branching fraction [16] is approximately

$$\mathcal{B}\left(\tau^- \to \mu^-\mu^+\mu^-\right) \sim \left| \sum_{i=2,3} U^*_{\tau i} U_{\mu i} \frac{\Delta m^2_{i1}}{m^2_W} \right|^2 < 10^{-54},$$

where Δm^2_{ij} are the measured neutrino mass differences. The values of Δm^2_{ij} and the PMNS matrix elements are taken from Ref. [17].

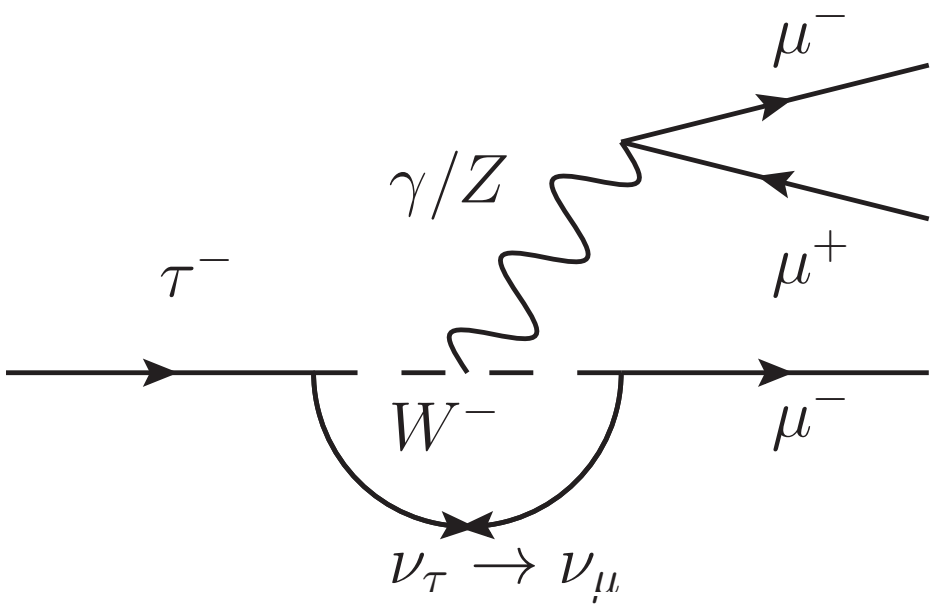


Figure 2: Example Feynman diagram for the $\tau^- \rightarrow \mu^-\mu^+\mu^-$ decay in the SM.

2.2.2 Lepton Flavour Violation Beyond the Standard Model

As the branching fractions for charged LFV decays in the SM are so small, they are beyond the sensitivity of any current or likely future experiment. Any observation of charged LFV at rates above the SM expectation would be an unambiguous sign of New Physics (NP).

Charged LFV could manifest itself in the decays of either the μ or the τ leptons. If the coupling constants in a NP model are mass-dependent, then LFV τ decays could be favoured over LFV μ decays. As the τ lepton is the heaviest charged lepton, it can decay to both quarks and leptons of the first and second generations, giving a larger number of decay modes for which the branching fractions can be modified by NP, and which can be studied to search for charged LFV. However, for any NP theory to be tested the expected branching fractions must be within experimental sensitivities. Therefore in the following, the charged LFV decay $\tau^- \rightarrow \mu^-\mu^+\mu^-$ is studied, as the possible NP modifications to the branching fraction are large, and the presence of three muons in the final state give

the decay a clean experimental signature. Further details of the experimental motivation for studying this decay are given in Chapter 5.

The enhancement of LFV branching fractions is often not the main motivation for the development of NP models. Instead, the possibility for observable charged LFV arises as a consequence of new symmetries, particles or interactions, which are introduced to address limitations of the SM. Setting limits on the branching fractions of charged LFV decays can therefore constrain the parameter spaces of these models, or in some cases exclude them completely. Similarly an observation of charged LFV and a precision measurement of the branching fraction may enable the origin of NP to be identified.

The NP models which modify the branching fraction for $\tau^- \to \mu^-\mu^+\mu^-$ can be broadly classified into two types. In the first, LFV decays are still forbidden at the tree level, but contributions at the loop level are enhanced. In the second, LFV decays are allowed directly at the tree level. Whilst many models of both of these types exist, brief descriptions of a selection of models are given in the following. The expected branching fractions for $\tau^- \to \mu^-\mu^+\mu^-$ decays from these models, along with those from various others, are summarised in Table 4.

2.2.2.1 *One-loop level modifications to the Standard Model*

The 'seesaw mechanism' offers a possible explanation for the small size of the neutrino masses in the SM, via the introduction of a number of additional right-handed neutrino fields and their associated heavy gauge bosons, W' and Z'. The SM neutrino masses can then be calculated via the relation

$$m_{\nu_l} \simeq m_e^2/g m_{W'},$$

where g is the 'coupling constant' of the weak interaction [18]. 'Supersymmetry' (SUSY) on the other hand, attempts to address the difference between the electroweak scale and the Planck scale (the 'hierarchy problem' [19]), by introducing a 'superpartner' for every SM particle [20]. Amongst other new particles, for every SM charged lepton there is a slepton, $\tilde{l}$, and for every SM gauge boson

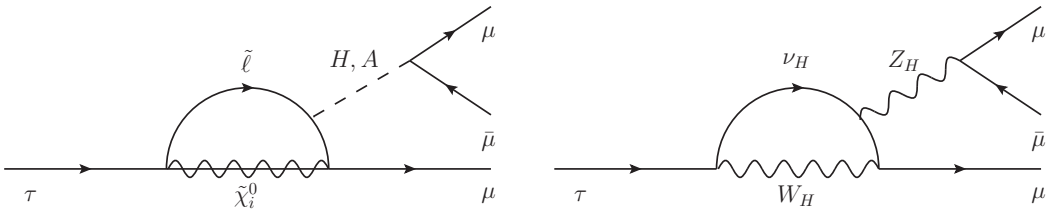


Figure 3: Example Feynman diagrams for the $\tau^- \to \mu^-\mu^+\mu^-$ decay at the one-loop level in (left) a Minimal Supersymmetric Model with massive neutrinos and (right) a Little Higgs Model with $\mathcal{T}$-parity.

there is an associated gaugino. The latter can mix to form four eigenstates of the mass operator, the neutralinos, $\tilde{\chi}_i^0$. There are also a number of additional Higgs particles, such as the neutral scalars, H and A. The Minimal Supersymmetric Standard Model (MSSM) with heavy neutrinos is a combination of elements of both of these models, and within it $\tau^- \to \mu^-\mu^+\mu^-$ decays can proceed via the diagram shown in Fig. 3 (left). The predicted branching fractions vary depending on the parameters of the model, but can be as large as approximately 10^{-9} [21].

As described previously as the motivation for SUSY, the GeV scale mass of the Higgs boson in the SM cannot be 'naturally' explained without the addition of NP [22]. Whilst many models attempt to solve this problem with the addition of new particles at the TeV scale, precision electroweak measurements suggest no evidence for NP below $5 - 7\,\text{TeV}$ [23], creating a so-called 'little hierarchy problem'. The 'Little Higgs' class of models address this problem by requiring that the SM Higgs boson is in fact a 'Nambu-Goldstone boson' [24]. This framework can be extended further by the introduction of a new symmetry at the TeV scale, $\mathcal{T}$-parity, which acts only on the new particles, such as the heavy bosons, W_H and Z_H, and the heavy neutrino, ν_H. As a result of this, tree level contributions from the model are forbidden, but additional loop diagrams are allowed. Figure 3 (right) shows an example of how this model could modify the branching fraction for the decay $\tau^- \to \mu^-\mu^+\mu^-$. Depending on the NP scale, branching fractions in the range $1 \times 10^{-10} - 2 \times 10^{-8}$ are predicted [25].

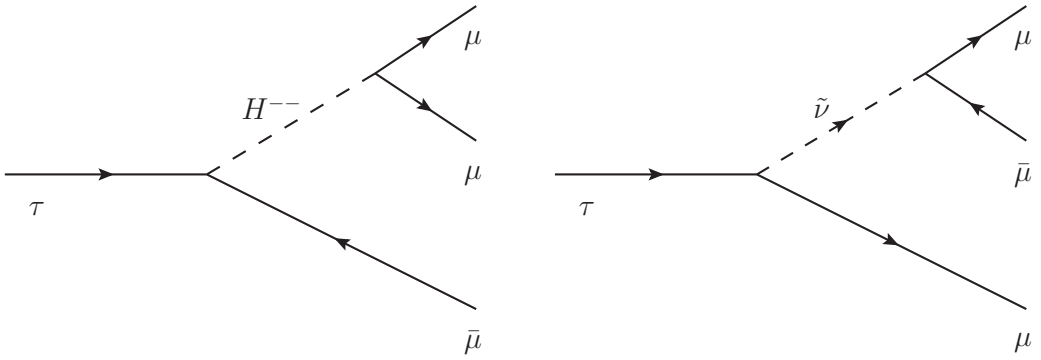


Figure 4: Example Feynman diagrams for the $\tau^- \to \mu^-\mu^+\mu^-$ decay at tree level in (left) a model with doubly-charged Higgs bosons and (right) a supersymmetric model with $\mathcal{R}$-parity violation.

2.2.2.2 *Tree level modifications to the Standard Model*

In 'left-right symmetric' models such as the Pati-Salam model [26], the SM is extended to include a right-handed weak interaction, in a similar manner to the seesaw mechanism described in the previous section. The symmetry also requires the introduction of two doubly-charged Higgs bosons, $H_L^{\pm\pm}$ and $H_R^{\pm\pm}$, with masses of the order of several hundred GeV or greater. The mass of the right-handed neutrino, ν_R, is then found to be proportional to the masses of the doubly-charged Higgs bosons, and is typically at the TeV scale. The doubly-charged Higgs bosons allow for the decay $\tau^- \to \mu^-\mu^+\mu^-$ to occur at the tree level, via diagrams such as that shown in Fig. 4 (left) [27]. The predicted branching fractions depend on a number of parameters of the model, such as the values of the PMNS matrix elements, but branching fractions as large as 10^{-7} are predicted for $M_{H_L^{\pm\pm}} = 3\,\text{TeV}$.

In supersymmetric extensions of the SM, the introduction of additional particles and interactions allows for the possibility of baryon number and lepton number violation. As the conservation of these quantities has been tested to a high level of precision experimentally, this is often an unwanted feature of these models. To account for this, an additional symmetry can be introduced with

Model	$\mathcal{B}(\tau^- \to \mu^-\mu^+\mu^-)$	Reference
SM + neutrino oscillations	$\sim 10^{-54}$	[16]
MSSM + seesaw mechanism	$< 10^{-9}$	[21]
Littlest Higgs with $\mathcal{T}$-parity	$< 10^{-8}$	[25, 29]
Left-right symmetric model	$< 10^{-7}$	[27]
$\mathcal{R}$-parity violating SUSY	$< 10^{-8}$	[28]
Non-universal gauge interaction	$< 10^{-8}$	[30, 31]
Unparticles	$10^{-3} - 10^{-11}$	[32]
Seesaw mechanism + Higgs mediated decays	$< 10^{-10}$	[33]
SUSY SO(10) + seesaw mechanism	$< 10^{-10}$	[34]
SM + heavy Majorana neutrino	$< 10^{-10}$	[35]

Table 4: Branching fraction predictions for the decay $\tau^- \to \mu^-\mu^+\mu^-$ in various models.

an associated quantum number, known as $\mathcal{R}$-parity [20]. $\mathcal{R}$-parity is defined as $+1$ for SM particles and -1 for their superpartners, such that in any SUSY decay the lightest supersymmetric particle cannot decay. Whilst the conservation of $\mathcal{R}$-parity is often favoured as it provides a dark matter candidate, a number of models exist where $\mathcal{R}$-parity is violated [28]. In these models the decay $\tau^- \to \mu^-\mu^+\mu^-$ can then be mediated by the exchange of a sneutrino, $\tilde{\nu}$, as shown in Fig. 4 (right), with branching fractions of 10^{-8} or greater, depending on the mass of the sneutrino.

2.2.2.3 *Other lepton flavour violating τ and μ decays*

In addition to the $\tau^- \to \mu^-\mu^+\mu^-$ decay described previously, a number of other LFV decays of either the τ or the μ could occur as a result of NP. Historically (as summarised in Fig. 5), the experimental upper limits on μ decays have been several orders of magnitude smaller than those for τ decays, as muons are produced in abundance at high intensity proton accelerators. However, as in certain NP

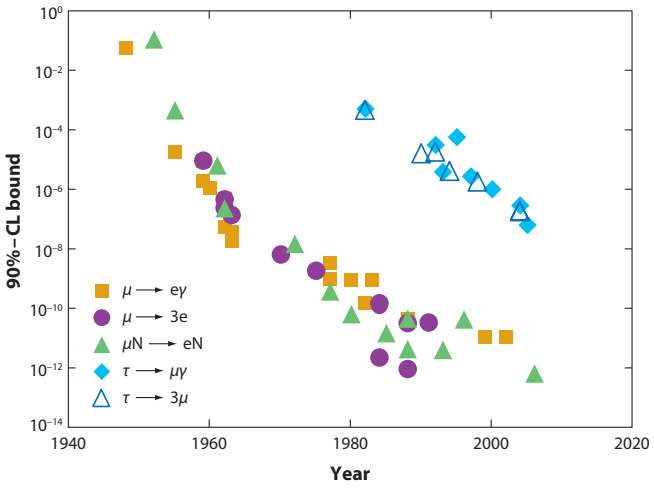


Figure 5: Improvements on the experimental upper limits on LFV τ and μ decays up to 2008. Figure from Ref. [36].

models the predicted branching fractions for τ LFV decays can be much larger than those for μ decays, there is a strong motivation to increase the sensitivity of searches for both types of decays.

Table 5 summarises the current experimental limits of several LFV τ and μ decays, along with theoretical predictions for their branching fractions within the Littlest Higgs Model described in Ref. [25]. Whilst an observation of any of these decays would be evidence for NP, it would not be sufficient to identify the NP model responsible. However, the ratio of the branching fraction predictions for different LFV decays is sensitive to the choice of model, and for example, given an experimental measurement of the ratio of the $\tau^- \to \mu^-\mu^+\mu^-$ and $\tau^- \to \mu^-\gamma$ branching fractions, the Littlest Higgs Model could be distinguished from the MSSM [29]. Furthermore, as demonstrated in Fig. 6, the branching fraction predictions within a given model are correlated, and therefore a measurement of the branching fractions of several LFV decays could enable the values of the parameters of the NP model to be identified, or the parameter space of the model to be restricted in the case of improved experimental constraints, but no observation of LFV.

Channel	Exp. limit	Theoretical expectation	Exp. ref.
$\tau^- \to \mu^-\gamma$	4.4×10^{-8}	$< 2 \times 10^{-8}$	[37]
$\tau^- \to e^-e^+e^-$	2.7×10^{-8}	$< 2 \times 10^{-8}$	[38]
$\mu^- \to e^-\gamma$	5.7×10^{-13}	$< 1.2 \times 10^{-11}$	[39]
$\mu^- \to e^-e^+e^-$	1.0×10^{-12}	$< 1.0 \times 10^{-12}$	[40]
$\mu^- N \to e^- N$	7×10^{-13}	$< 10^{-10}$	[41]

Table 5: Experimental limits on the branching fractions of LFV decays and their theoretical expectations in the Littlest Higgs Model described in Ref. [25].

2.3 BARYON NUMBER VIOLATION

The apparent excess of baryons in the Universe is one of the biggest unexplained mysteries in physics. According to the 'Sakharov conditions' [43] the creation of this baryon asymmetry, or 'baryogenesis', could only have occurred in the early universe if there was:

1. Baryon number violation (BNV), such that the number of baryons minus the number of anti-baryons is not constant;

2. Charge-parity violation, such that the baryon number violating processes were not cancelled out by their CP conjugate processes;

3. Interactions out of thermal equilibrium, such that the baryon number violating processes were not cancelled out by their time-reversed processes.

Whilst CP violating processes are allowed within the SM through the weak interaction, the observed CP violation in the quark sector is insufficient to account for the baryon asymmetry [44]. Possibilities exist for another source of CP violation, such as in the lepton sector, but this has yet to be experimentally determined.

No BNV processes have yet been observed [3], but they are predicted in a number of NP models, such as $\mathcal{R}$-parity violating SUSY [28]. Due to angular

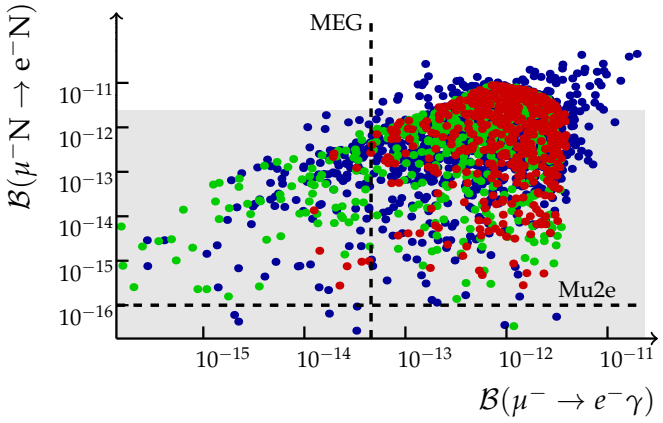


Figure 6: Correlation between the μ-e conversion rate and $\mathcal{B}(\mu^- \to e^-\gamma)$ in the Littlest Higgs Model described in Ref. [29]. The three colours represent different benchmark scenarios of the model. The current experimental upper limit on $\mathcal{B}(\mu^- \to e^-\gamma)$ from the MEG experiment [39] and the expected experimental sensitivity from the Mu2e experiment [42] are shown for reference. Figure adapted from Ref. [29].

momentum conservation, the change in the number of baryons, ΔB, must be equal to the change in the number of leptons, ΔL, via the relation, $\Delta B = \pm\Delta L$, such that $|\Delta(B-L)| = 0$ or 2. The SM and most of its extensions [45] require $|\Delta(B-L)| = 0$. The two BNV decays presented in Chapter 5, $\tau^- \to \bar{p}\mu^+\mu^-$ and $\tau^- \to p\mu^-\mu^-$, have $|\Delta(B-L)| = 0$.

Experimental constraints on the lifetime of the proton ($\tau_p > 10^{31}$ years [3]) place indirect bounds [46] on the branching fractions for $\tau \to \bar{p} + X$,

$$\mathcal{B}(\tau \to \bar{p} + X) < 10^{-40}.$$

However, these constraints may not be applicable to models with dimension-9 operators or other mechanisms [47], and therefore it is of interest to study a range of BNV decays experimentally. Further details of BNV models can be found in Refs. [44, 48].

3 THE LARGE HADRON COLLIDER AND THE LHCB EXPERIMENT

3.1 THE LARGE HADRON COLLIDER

The Large Hadron Collider (LHC) is a particle accelerator located at the CERN complex in Geneva, Switzerland. It accelerates beams of protons or heavy ions in opposite directions around a 27 km ring, bringing them to collide at four points where experiments are located (Fig. 7). Of the four major LHC experiments, two of them, ATLAS and CMS, are referred to as 'general purpose detectors' (GPDs), whilst another, ALICE, specialises in the study of heavy ion collisions and the fourth, LHCb, specialises in the study of rare decays and CP violation in heavy flavours and is the focus of this thesis. There are also three smaller experiments of which TOTEM and LHCf study particles on either side of the GPD collision points and MoEDAL searches for magnetic monopoles. During nominal running conditions the LHC beams consist of protons, but it is also possible to perform collisions of heavy ions or a heavy ion and a proton in either direction. In the following, only the collisions of protons are considered.

The performance of the accelerator is typically defined by its centre-of-mass energy, $\sqrt{s}$, which indicates the energy available for producing particles in collisions, and the luminosity, $\mathcal{L}$, which gives the rate for an interaction when multiplied by its cross-section, σ. Since the startup and commissioning of the LHC in 2008, both of these parameters have increased towards their design values of 14 TeV and $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ respectively. Whilst the peak luminosity approached this value in 2012, the accelerator is not planned to operate at full energy until it is fully upgraded, following the first long shutdown (LS1) from 2013 to 2015. As

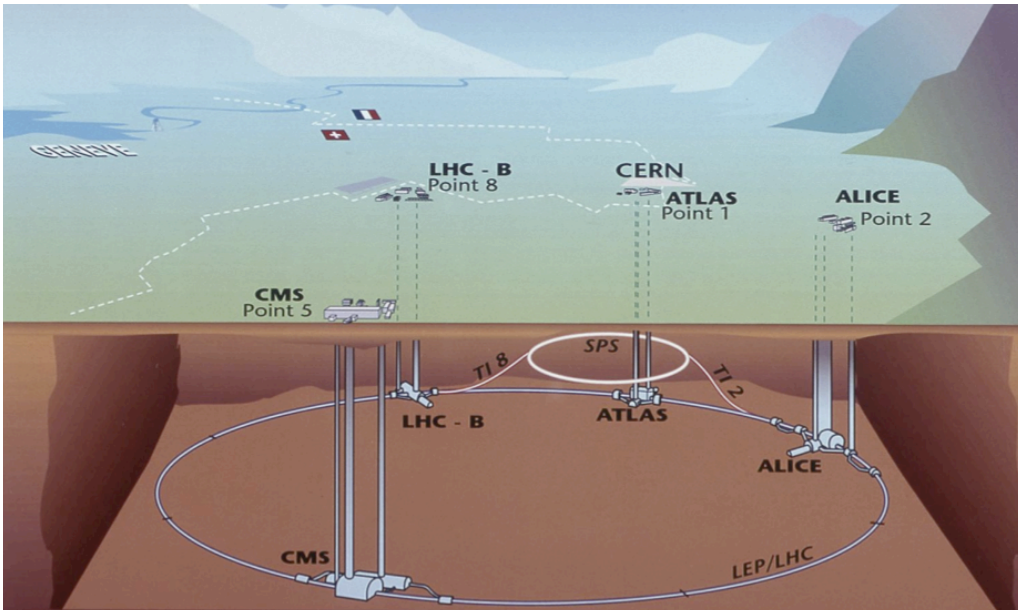


Figure 7: Overall view of the LHC, showing the position of the four major experiments around the accelerator ring. Figure from Ref. [49].

a result the data collected by all LHC experiments up until early 2013 includes a number of different beam conditions, which can be broadly distinguished by the collision energy. In the following, only the largest of these datasets, those collected at 7 and 8 TeV, are considered. Each dataset can be further divided into a number of 'fills', with each representing one complete filling and subsequent circulation of protons within the LHC. Each fill itself contains a number of 'runs' which represent a single non-stop period of data taking by a particular LHC experiment.

3.2 THE LHCB EXPERIMENT

The LHCb detector [50], shown schematically in Fig. 8, is a forward spectrometer, which provides instrumental coverage of the angular region close to the beam direction (in this case defined as the z-axis) in one direction only. This

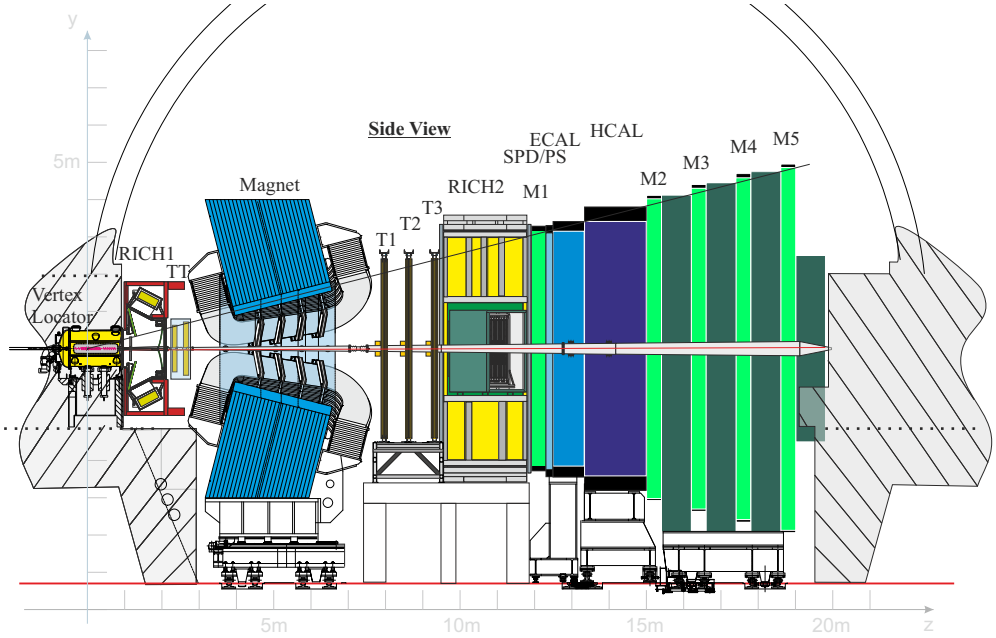


Figure 8: Schematic diagram of the LHCb detector. Figure from Ref. [51].

coverage is defined by the pseudorapidity range $2 < \eta < 5$, which extends to much larger values than that of the GPDs. As the production of heavy quark pairs (defined here as $b\bar{b}$ and $c\bar{c}$) is predominantly in the forward region, LHCb is able to detect $\sim 40\%$ of the heavy quark pairs despite only covering $\sim 4\%$ of the total solid angle.

Its design is further optimised for the study of particles containing b or c quarks via the features described in the following sections.

3.2.1 The Dipole Magnet

The LHCb dipole magnet is used to measure the momentum (p) of charged particles via the curvature of their trajectories in a known magnetic field. It consists of a warm magnet design with sloping poles to match the detector acceptance, and provides an integrated magnetic field of 4 Tm for tracks of 10 m length. The region of momentum measurement covers ± 250 mrad vertically and ± 300 mrad horizontally. Its design is constrained by the need to maximise the field in the

tracking subdetectors whilst keeping the field inside the RICH detectors (described later) below 2 mT. It must also fit within the dimensions of the experimental hall, which was previously occupied by the DELPHI experiment [52] at the Large Electron Positron (LEP) collider [53].

For the measurement of physics asymmetries the systematic effects of the detector have a large contribution and these can be controlled by periodically changing the direction of the magnetic field. The LHCb dataset contains roughly an equal amount of data collected at each polarity for this purpose.

For physics measurements the magnetic field integral is required to be known with a relative precision of $\sim 10^{-4}$, whilst the position of the B-field peak must be known to within a few millimetres. Therefore a detailed knowledge of the three components of the magnetic field inside the tracking volume is required for both magnet polarities. This is obtained via the remote scanning of an array of Hall probes throughout the detector over a fine grid of $8 \times 8 \times 10\,\mathrm{cm}^3$, resulting in a field map with a relative precision of $\sim 4 \times 10^{-4}$.

3.2.2 Tracking

The LHCb tracking system, shown in Fig. 8 consists of three parts: the Vertex Locator (VELO), the Tracker Turicensis (TT) and the tracking stations, T1–T3. Whilst the VELO and TT are upstream (closer to the interaction point) of the dipole magnet, T1–T3 are placed downstream (further from the interaction point). The TT and T1–T3 are split into two projects based on their technology choices, with the TT and the Inner Tracker (IT) forming the Silicon Tracker (ST), where the IT constitutes the innermost regions of T1–T3. The Outer Tracker (OT) then makes up the remaining outer parts of T1–T3.

3.2.2.1 *Vertex Locator*

In a typical LHCb event, if a b or c hadron is produced directly in the collision of the proton beams at the interaction point, the position in 3D-space where

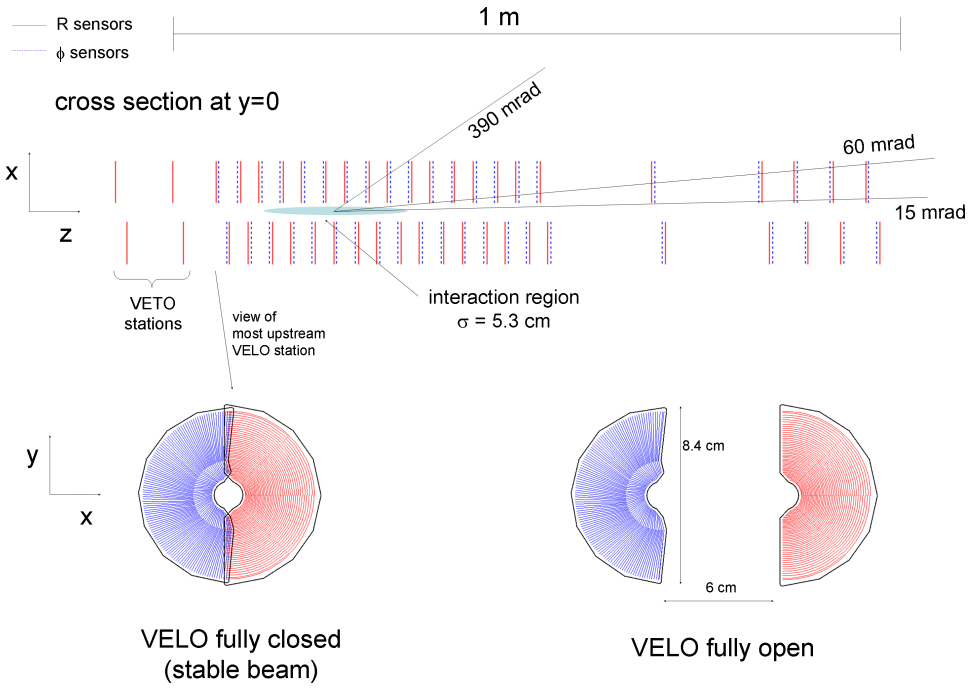


Figure 9: Top: Schematic view of the placement of the VELO modules along the LHCb beam axis. Bottom: Position of the VELO sensors before and after closing. Figures from Ref. [50].

this occurs is known as the primary vertex (PV). These heavy particles are long lived, with mean lifetimes of 10^{-13} – 10^{-12} seconds, and travel at velocities close to the speed of light, such that in the detector frame of reference the typical distance between the PV and the secondary vertex (SV: the point where the heavy particle decays) can be several mm.

The VELO is a silicon strip detector located around the LHCb interaction point, designed to provide identification and the precise measurement of SVs which are used to infer the presence of b and c hadrons and τ leptons within events. A schematic layout of the VELO is shown in Fig. 9. It consists of 42 standard modules, where each module contains an R and a ϕ sensor which provide the R (radial) and ϕ (azimuthal) coordinates of a particular hit, whilst the z-coordinate is determined from the module position along the beam axis.

The choice of cylindrical geometry is motivated by the faster reconstruction of tracks in the trigger with respect to rectangular geometry. The 42 modules are complimented by two sets of R sensors upstream of the interaction point, called the pile-up veto system which is used to improve the PV reconstruction and to veto events with more than one pp interaction.

The measurement of the impact parameter (IP) of a track is an essential feature of the VELO. The IP is defined as the distance of closest approach of a track to a particular point in 3D-space, e.g. the PV. The IP resolution improves as the radial distance between the inner-most region of the sensors and the LHC beam decreases. Therefore the sensors are positioned at the minimum possible radial distance of 8.2 mm, defined by the LHC safety margin to account for the RMS spread of the beam (5 mm) plus the distance to the active area of the sensors (3 mm). The active area of the sensors extends radially outwards to an outer radius of 42 mm. As the inner radial distance is smaller than the aperture required by the LHC beam during injection, the modules must therefore be retractable. For optimum performance the sensors must operate in vacuum, which itself must be independent of the machine vacuum. In order to separate the two at the innermost parts of the VELO a thin corrugated aluminium sheet (RF foil) encloses the sensors. Whilst the thickness of the RF foil should be limited to minimise the detector material budget, it is also required to be at least 0.3 mm thick to shield against RF pickup, suppress beam wakefields, prevent outgassing and avoid inelastic deformations due to the vacuum.

To cover the full angular acceptance of the downstream detectors, the VELO must detect particles in the pseudorapidity range $1.6 < \eta < 4.9$ that emerge from PVs within 10.6 cm of the nominal interaction point ($z = 0$) in both the positive and negative z-directions. Figure 10 shows the VELO tracking efficiency as a function of pseudorapidity (left) and azimuthal angle (right). The overall efficiency is approximately 99%, with a dip in the efficiency at $|\phi| \approx \pi/2$ due to multiple scattering from the RF foil. Any track should traverse a minimum of four VELO stations. While only three VELO stations are needed for an ac-

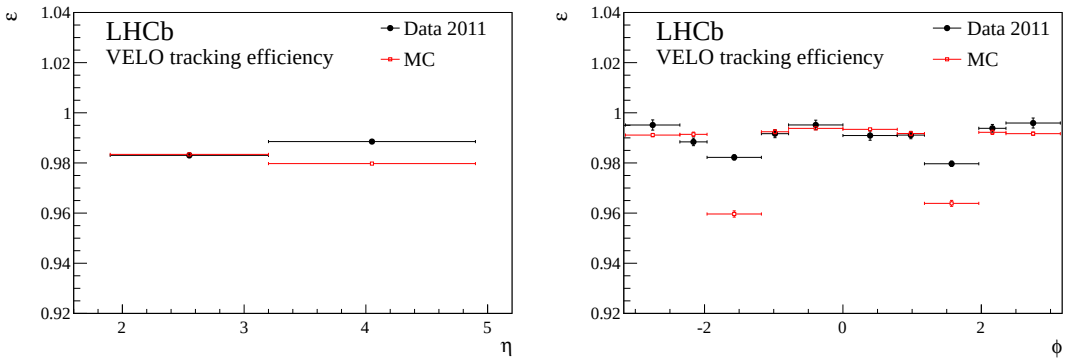


Figure 10: Tracking efficiency as a function of η (left) and ϕ (right) for VELO tracks in 2011 data. Figures from Ref. [54].

curate measurement, requiring four allows for the possibility of a missing hit. These requirements act to define the geometrical layout of the sensors within the VELO. As a result the first of the three most downstream sensors is positioned at $z = 65\,\text{cm}$, and given that the sensors themselves are limited to a maximum active outer radius of 42 mm, the distance between sensors in the central region is less than 3.5 cm. To fully cover the inner angular acceptance the sensors overlap slightly, which is achieved by shifting one half of the detector by 1.5 cm in the z-direction, such that the R and ϕ sensors in opposite halves can be brought closer together.

The performance requirements of the detector demand that the signal-to-noise ratio (S/N) should initially be greater than 14, and remain above 10 for the detector lifetime. When a particle traverses a silicon sensor it undergoes energy loss and deposits electrons in the silicon. These electrons are read out as analogue-to-digital converter (ADC) counts and form a distribution that can be fitted with the convolution of a gaussian and a landau distribution to obtain the most probable value (MPV), which defines the signal, S. The noise, N, is subsequently defined as the RMS value of the intrinsic ADC counts for any channel. For a cut of $S/N > 5$ the overall channel efficiency is required to be better than 99%. The S/N of typical R and ϕ sensors before irradiation are shown in Fig. 11. The average S/N is around 20 with higher values for ϕ sensors than R sensors,

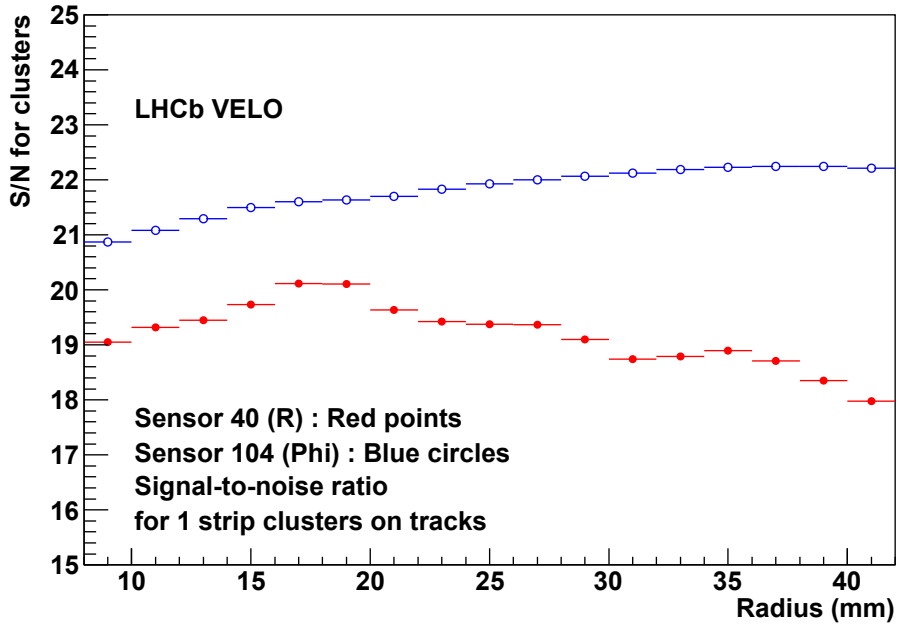


Figure 11: S/N as a function of sensor radius for R (red) and ϕ (blue) sensors before irradiation. Figure from Ref. [54].

due to differences in the strip and routing line layout between the two sensor types.

To obtain the desired IP resolution a charge cluster (several adjacent strips with charge above a predefined threshold) must be resolvable to within 4 µm for tracks within the smallest strip pitch region. Figure 12 (left) shows the single hit resolution in data and the IP resolution in the x-direction (right). For particles with transverse momentum (p_T) greater than 1 GeV, IP resolutions of less than 35 µm are obtained. The tracking also imposes strict requirements on the alignment of the detector. Within each half the R sensors should be positioned to an accuracy of 20 µm in x and y relative to one another, whilst the two halves should be aligned to within 100 µm. Once the VELO is closed around the LHC beam the two halves of the detector are adjustable in x and y to account for variations in beam position.

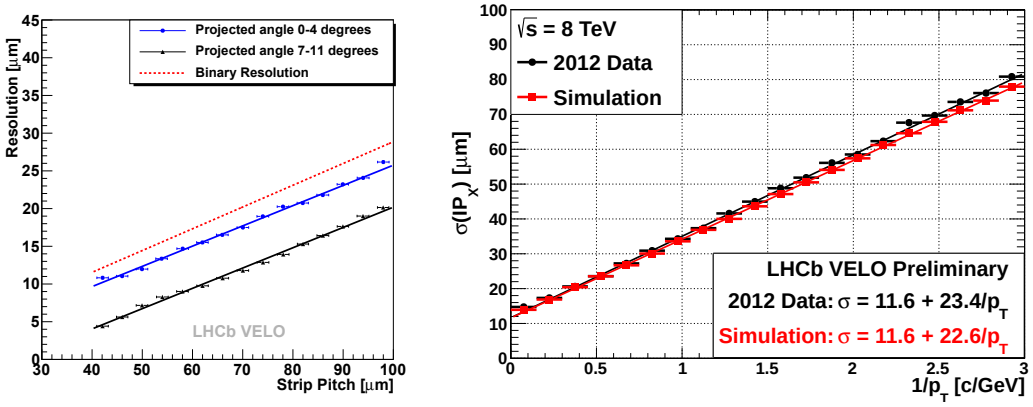


Figure 12: Left: Single hit resolution as a function of strip pitch for tracks of different projected angle in data. Right: IP resolution in the x-direction as a function of $1/p_T$ in data and MC. Figures from Ref. [54].

Further performance parameters of the VELO are summarised in Fig. 13. Precise vertex reconstruction is essential for the separation of production and decay vertices. The vertex resolution is highly dependent on the track multiplicity as shown in Fig. 13 (left). In the x and y-directions it is found to be better than 10 μm for vertices with 40 tracks. The accurate determination of decay times is required to resolve the fast oscillations which occur in heavy meson mixing. The decay time resolution is dominated by the track parameter resolution for short decay times, and is determined separately for each final state of interest. A resolution of 50 fs is obtained for $B_s^0 \to J/\psi\phi \to \mu^+\mu^-K^+K^-$ decays and is measured using prompt combinations which fake signal candidates. The shape of the prompt peak, shown in Fig. 13 (right), is then determined only by the resolution.

The VELO sensors are described in more detail in Section 4.3.

3.2.2.2 *Silicon Tracker*

The two detectors in the ST project are grouped together due to their similar requirements and design considerations. Both use silicon microstrip sensors with

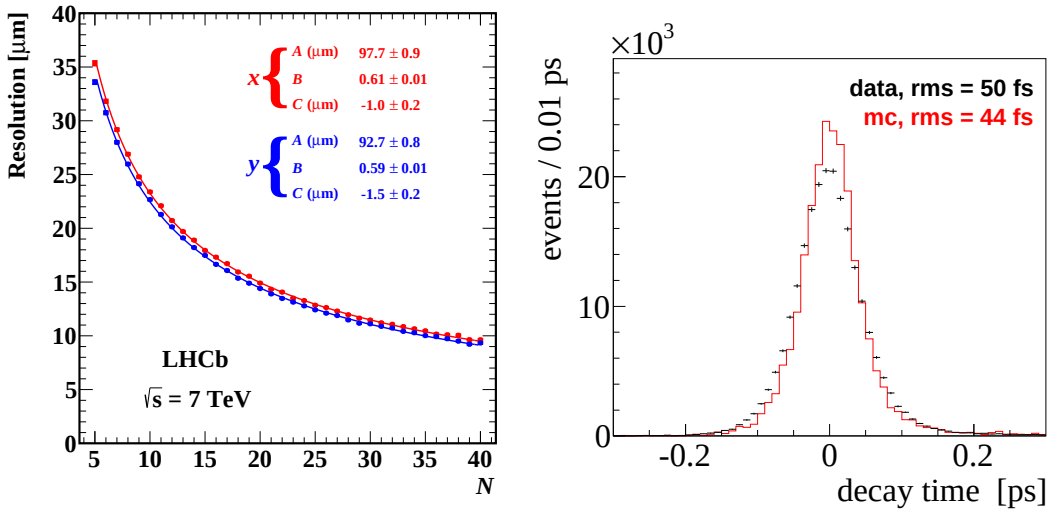


Figure 13: Left: Primary vertex resolution in the x (red) and y (blue) directions in 2011 data as a function of track multiplicity. Exactly one PV is required. Right: Decay time distribution for fake, prompt $B_s^0 \to J/\psi\phi \to \mu^+\mu^-K^+K^-$ decays. Figures from Ref. [54].

a strip pitch of 200 μm, to achieve a 50 μm single hit resolution such that the momentum resolution is dominated by multiple scattering.

The TT, shown in Fig. 14, is located upstream of the dipole magnet, and is 150 cm wide and 130 cm high. It covers the full LHCb acceptance, whilst the IT covers the 120 cm by 40 cm region in the centre of the downstream tracking stations T1–T3. The TT has a much larger active area at 8.4 m^2, compared to 4.0 m^2 for the IT. Despite this the total number of readout strips is similar, at 143 360 and 129 024 respectively, as the number of channels is a significant contribution to the overall cost of the detectors. As the charged particle densities vary between the two detectors, different strip lengths are employed such that the maximum strip occupancies are kept at a few percent whilst minimising the number of channels. With the short LHC bunch crossing rate, pile-up effects from consecutive interactions can have a considerable effect on the detector performance. Therefore fast front-end amplifiers with a signal shaping time of the order of 25 ns (the nominal bunch crossing rate) are used.

As with the VELO, the key parameter of detector performance is the S/N ratio and it has been shown that provided this remains above 10 : 1 a full particle detection efficiency of 99.8% is achieved. Both the TT and the IT are designed such that the S/N ratio should remain in excess of 12 : 1 after 10 years of operation at the nominal luminosity. In the following the '1 MeV neutron equivalent fluence per cm^2', $n_{eq}\,cm^{-2}$, is used as a measure of the integrated particle flux, with expected delivered fluences of $5 \times 10^{14}\,n_{eq}\,cm^{-2}$ and $9 \times 10^{12}\,n_{eq}\,cm^{-2}$ for the TT and IT respectively after 10 years. These large fluences also cause the required bias voltage and leakage currents to increase and as such the detectors have been designed and tested to operate at up to 500 V whilst staying below 5 °C. All of the above requirements are achieved with a material budget of on average 0.04 radiation lengths (X_0) for the TT and $0.035X_0$ for the IT, with some variations depending on η and ϕ, such that the effects of multiple scattering are minimised.

3.2.2.3 *Outer Tracker*

The OT utilises a drift-time detector to track charged particles and measure their momentum with high precision, such that the invariant mass of reconstructed hadrons can be accurately determined. The OT detector itself consists of a number of straw-tube modules, containing two staggered layers of drift-tubes of 4.9 mm inner diameter. They are filled with Argon and CO_2 in a 70 : 30 mixture, such that a drift time of less than 50 ns and a drift-coordinate resolution of 200 μm are achieved.

Each of T1–T3 contains four layers of modules, arranged in an x-u-v-x layout, such that the u and v layers are tilted by $\pm 5^\circ$ with respect to the vertical x layer to give a total active area of $5971 \times 4850\,mm^2$. The stations provide coverage out to 300 and 250 mrad in the horizontal and vertical planes respectively, whilst the transition to the IT occurs once the occupancies approach 10% at $2 \times 10^{32}\,cm^{-2}\,s^{-1}$.

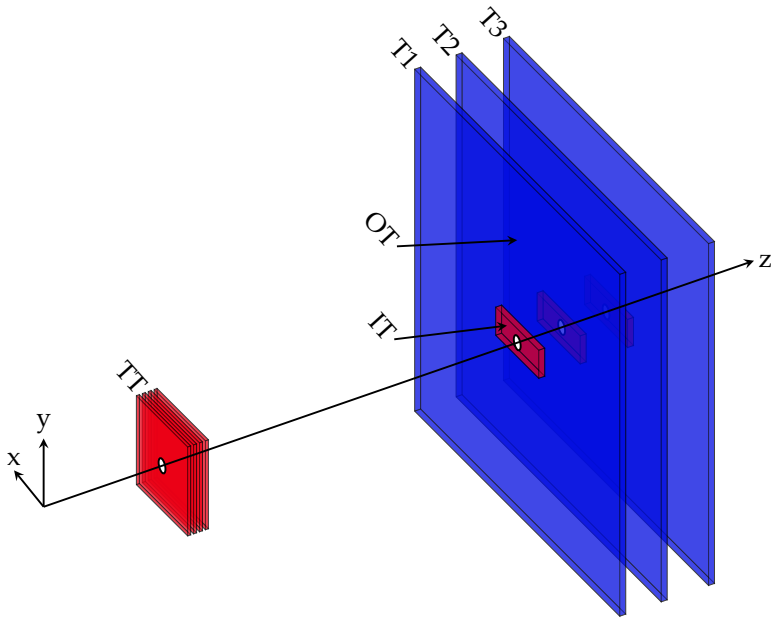


Figure 14: Layout of the LHCb tracking system, highlighting the IT and OT portions of T1–T3. The TT and IT collectively form the ST.

The straw-tube modules must be mechanically stable such that their position can be known to within 100 and 500 µm in the x and z-directions respectively. Furthermore the anode wire must be centred to within 50 µm over the entire length of the tube and the module boxes must be able to withstand an overpressure of 10 mbar, whilst maintaining a leak rate below $8 \times 10^{-4}\,\mathrm{l\,s^{-1}}$. To reduce the impact of multiple scattering the material budget must be kept below a few percent per station, whilst the tubes must be shielded to avoid noise. The detector should also be able to operate for 10 years at the nominal luminosity, requiring all detector materials to be radiation-hard.

As many of the final states of interest in heavy flavour decays involve a large number of particles it is required to maintain a tracking efficiency of close to 100% such that the overall reconstruction efficiency, which is the product of the individual efficiencies, remains high. In practice efficiencies of $\sim$ 99% are achieved. The OT must also minimise the number of 'ghost tracks', defined as tracks formed from a combination of pseudo-random hits in the detector that

do not actually originate from a particle. The treatment of ghost tracks is discussed in Appendix A. The second possible form of misidentified tracks, 'clone tracks', results from track segments or whole tracks being found by different algorithms and therefore being duplicated. These are removed by clone killing algorithms in the LHCb software framework, but these are not 100% effective and the treatment of clones is discussed in more detail in Chapter 5.

3.2.3 Particle Identification

Excellent separation of charged particle species across a range of momenta is an essential requirement of the LHCb experiment and is achieved by the use of three subdetectors: the two ring imaging Cherenkov (RICH) detectors, the calorimeters and the muon stations.

3.2.3.1 *RICH*

The LHCb detector includes two RICH detectors with different radiators to cover the full momentum range. By exploiting the relation between momentum and Cherenkov angle, charged particle identification (PID) can be achieved between hadrons, and to some extent muons.

RICH1, which is upstream of the dipole magnet, uses aerogel and C_4F_{10} radiators to cover low momentum (up to 60 GeV) over the full angular acceptance (see Fig. 15). RICH2, which is downstream of the dipole magnet, uses a CF_4 radiator to cover high momentum (up to 100 GeV) between ± 15 and ± 120 (100) mrad in the vertical (horizontal) plane. The large difference in momentum range between the two RICHs is due to the low momentum particles being swept out of the detector by the magnetic field, and thus the remaining particles occupying a narrower angular acceptance. The resulting Cherenkov light is focused by a combination of spherical and flat mirrors onto Hybrid Photon Detectors (HPDs) outside the detector acceptance, which detect wavelengths in the range 200 – 600 nm. The HPDs operate at a bias voltage of 20 kV, where photons are con-

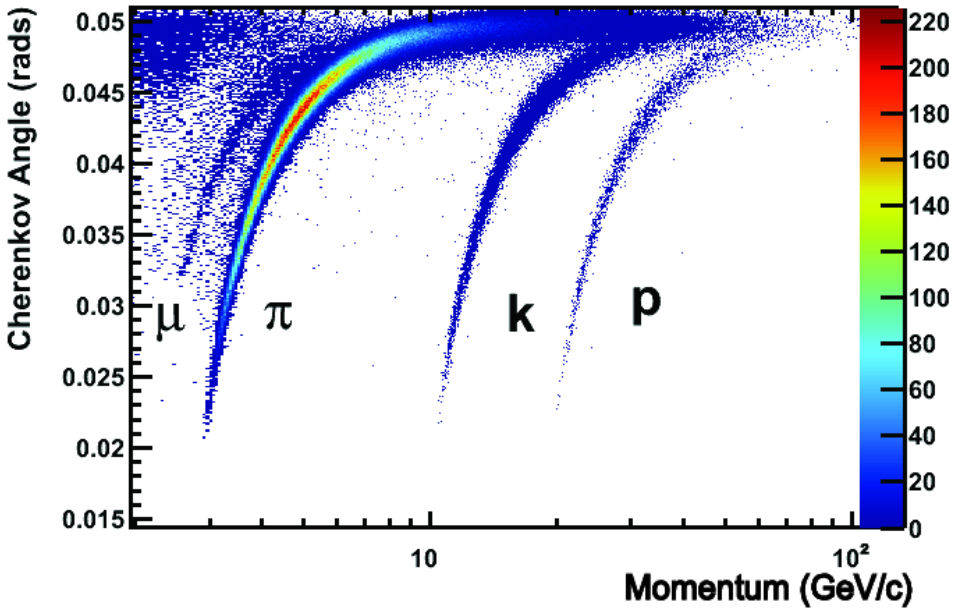


Figure 15: Cherenkov angle versus momentum for different particles in the C_4F_{10} radiator in RICH1. Figure from Ref. [55].

verted into photoelectrons and captured on a segmented silicon detector of 1024 pixels of 500 μm^2. They cover an area of 3.5 m^2 with an active area of 64%, a high granularity of 2.5 mm^2 and 25 ns timing resolution. The Cherenkov angle resolution is determined to be 1.618 ± 0.002 mrad for RICH1 and 0.68 ± 0.02 mrad for RICH2.

In order to operate in the high magnetic field environment at either side of the dipole magnet, the HPDs are surrounded by iron shields and placed in MuMetal (an alloy with a high magnetic permeability) cylinders. Residual stray fields still have a significant effect on the images recorded on the HPDs and therefore individual correction factors are applied to each. Changes in the effect of the field are periodically tested for during dedicated calibration runs and then corrected for offline.

RICH1 is required to minimise the material budget within the detector acceptance such that the mirrors must be lightweight and all other components, such as the HPDs with their iron shielding, are kept out of the acceptance (see Fig. 16

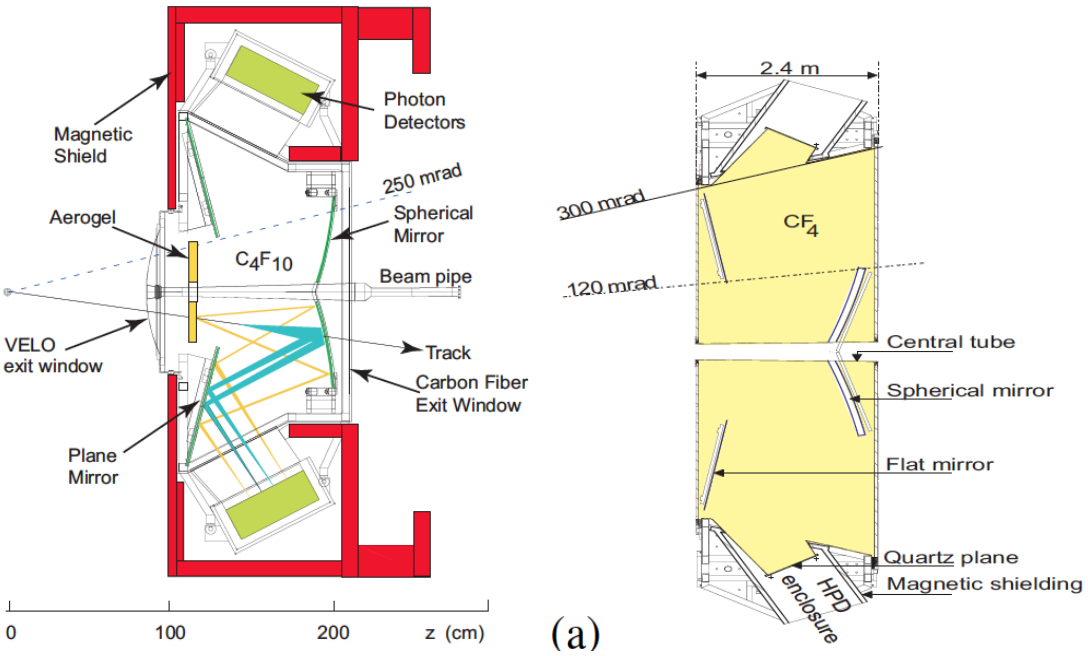


Figure 16: Schematic side views of RICH1 (left) and RICH2 (right). The relative dimensions are incorrect. Figure from Ref. [50].

left). As a result the total radiation length is $\sim 0.08X_0$. The lower limit on the angular acceptance (± 25 mrad) is defined by the position of the LHCb beampipe, which passes through the centre of the detector.

RICH2, which is the much larger of the two RICHs, is also required to minimise the material budget within the detector acceptance. This is achieved by keeping the supporting structures and HPDs outside the detector acceptance, whilst shortening the overall length of the detector by including a set of flat secondary mirrors (see Fig. 16 right). This results in a total radiation length of $\sim 0.15X_0$.

The radiators are required to have a large photoelectron yield per track that is stable against fluences up to $5.5 \times 10^{13}\,n_{eq}\,cm^{-2}$. The small volume of aerogel used ($\sim 30\,l$) allows for the option of a full replacement if required. To ensure good photon transparency the gaseous radiators must maintain a high purity at all times and this is periodically tested by measuring the speed of sound in the

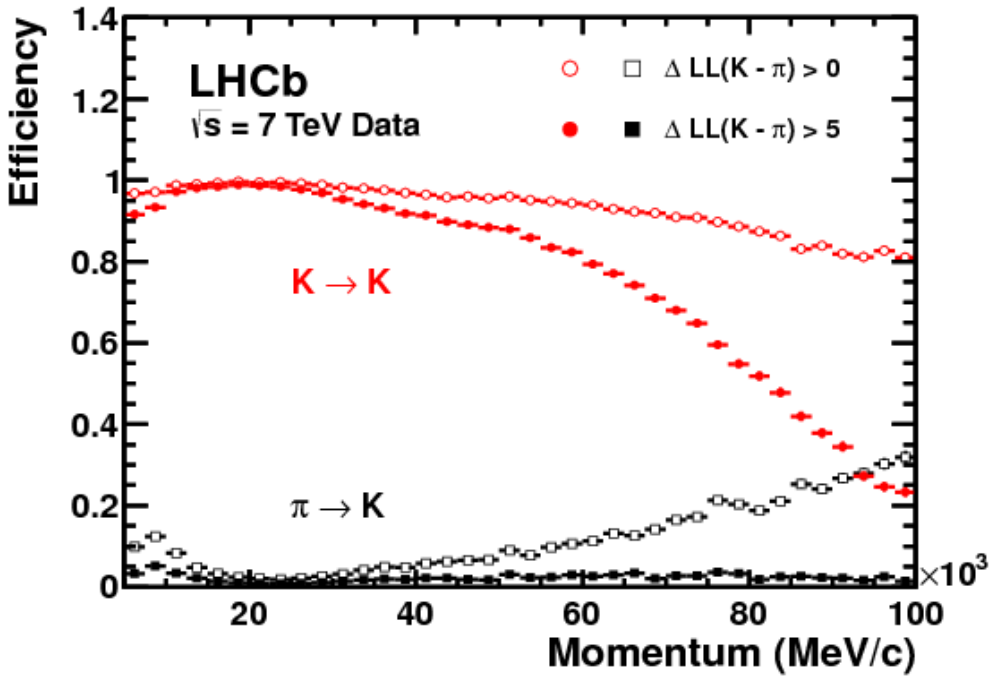


Figure 17: Data determined kaon identification efficiency (red) and pion misidentification rate (black) as a function of momentum, for different cuts on the particle DLLs. Figure from Ref. [55].

gas, via the time taken for a sound pulse to travel to the other side of the vessel and back.

A reflective coating of Al + SiO_2 + HfO_2 is chosen for all RICH mirrors such that the reflectivity is largest in the photon energy range where the HPD quantum efficiency is maximum. This results in a 5% increase in detected photon yield with respect to other coatings. The detecting layer is also hard and chemically inert. The mirrors themselves are required to be carefully aligned to maintain the required Cherenkov angular resolution throughout the data taking. Their movements are monitored by a CCD camera which compares the position of a laser light spot with that reflected off the mirror. Changes in relative position of the two spots are then corrected for offline.

By combining the Cherenkov angle information with the track momentum from the tracking system the likelihood for a track to be of a given charged

particle type can be determined. As the majority of charged particles produced in pp collisions are pions, the difference in likelihood with respect to the pion hypothesis is used, giving a delta log-likelihood or DLL. Figure 17 shows that for a cut of $DLL(K-\pi) > 0$ the efficiency for identifying kaons and the rate for misidentifying pions in data are 95% and 10%, respectively up to momentum of 60 GeV. By increasing this cut to $DLL(K-\pi) > 5$ the misidentification rate can be reduced to 3% for a kaon identification efficiency of 85%. The performance of the RICH is described in detail in Ref. [55].

3.2.3.2 *Calorimeters*

The calorimeters, shown schematically in Fig. 18, select hadron, photon and electron candidates for the trigger via measurements of their transverse energy, E_T. Full performance is only obtained at the offline analysis level, where they also provide PID and position measurements. The constraints on the trigger performance and PID define the overall structure, which is an electromagnetic calorimeter (ECAL) followed by a hadronic calorimeter (HCAL).

The electron trigger looks for high E_T particles to reject 99% of the background. As this background is predominantly charged pions that can be mistaken for electrons, the ECAL is longitudinally segmented with a preshower detector (PS) before the main ECAL, where the different species are distinguished via their energy deposition. Electrons are separated from neutral pions and photons by the placement of a scintillator pad detector (SPD) in front of the PS, in which only charged particles deposit energy. A 15 mm lead converter (corresponding to $2.5X_0$) is placed between the SPD and the PS such that they are spaced 56 mm apart. The SPD/PS alone has been shown to achieve pion rejection factors of $\sim 99.6\%$ with electron efficiencies from 91 – 97% in a test beam environment [50].

For the best possible energy resolution the thickness of the ECAL is chosen to be $25X_0$ such that the full electromagnetic shower of a high energy photon is contained within the calorimeter. As the trigger performance of the HCAL is not strongly dependent on hadronic energy resolution its thickness is set to 5.6

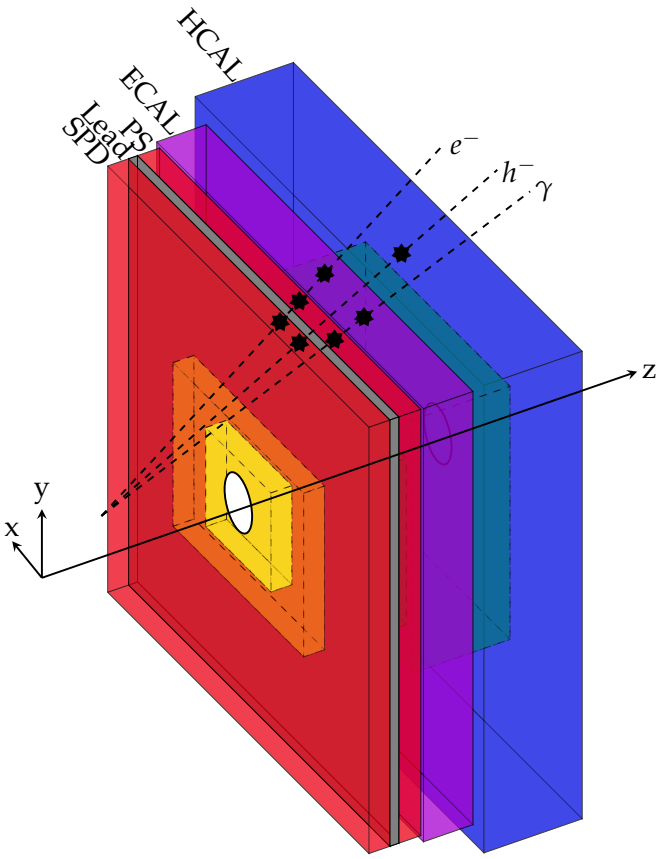


Figure 18: Layout of the SPD/PS, ECAL and HCAL showing the segmentation and the interactions of different particle species. The relative dimensions of the ECAL and HCAL are correct, but the z-scale of the SPD/PS is exaggerated.

hadronic interaction lengths to save space. As the hit density varies by several orders of magnitude with radius in all sections of the calorimeter the detectors are laterally segmented into zones. Whilst the SPD/PS and ECAL are segmented into three zones, the HCAL is only segmented into two due to the larger size of hadronic showers. All sections of the calorimeter are also divided vertically into two halves to allow for access.

The scintillators consist of polystyrene doped with paraterphenyl and POPOP. The scintillation light is collected by wavelength-shifting fibres and detected by PMTs. The gain on the phototubes is scaled according to their radial distance

such that a constant E_T scale is achieved. The HCAL is operated at 30 times the gain of the ECAL to compensate for the lower light yield.

The SPD/PS detector consists of two planes of scintillator pads with a total of 12 032 channels coupled to 64 channel multi-anode PMTs. The active area is 7.6 m wide and 6.2 m high, with the dimensions of the SPD smaller by 0.45% due to the angular projection. The segmentation of the three radial zones projects exactly onto the corresponding ECAL zones, with dimensions of 4×4, 6×6 and $12 \times 12\,\text{cm}^2$ for the individual cells. The multi-anode PMTs are shielded with iron and MuMetal to reduce the effects of the residual magnetic field.

The ECAL uses a structure of alternating layers of scintillator and lead read-out by scintillating fibres due to its energy resolution, response rate, radiation resistance and reliability. Placed 12.5 m from the interaction point, it matches the outer angular acceptance of the tracking system, whilst being limited to an inner acceptance of > 25 mrad in both the x and y-planes due to radiation levels. A single module consists of 2 mm of lead, 120 µm of white, reflecting paper and 4 mm of scintillator tile. A total of 66 modules are then combined to create a 42 cm stack of $25X_0$ and 3.5 cm Moliere radius. In total there are 3312 modules comprising 6016 channels. The ECAL is designed to achieve an energy resolution of $\sigma_E/E = 10\%/\sqrt{E} \oplus 1\%$, and test beam studies, the results of which are shown in Fig. 19 (left), have confirmed that a performance close to this is obtained.

The HCAL combines iron and scintillating tiles in a structure that runs parallel to the beam axis. Tiles are interspersed with 1 cm of iron laterally for a total of 5.6 hadronic interaction lengths, whilst in the longitudinal direction the distance corresponds to a single interaction length. The detector is transversely segmented into cells of 131.3 mm and 262.6 mm square in the inner and outer radial zones respectively, where a cell is defined by the grouping together of multiple fibres onto a single PMT. The detector is built as a wall at 13.33 m from the interaction point with a height of 8.4 m, width of 6.8 m and depth of 1.65 m. A ^{137}Cs gamma source is embedded within the detector allowing for

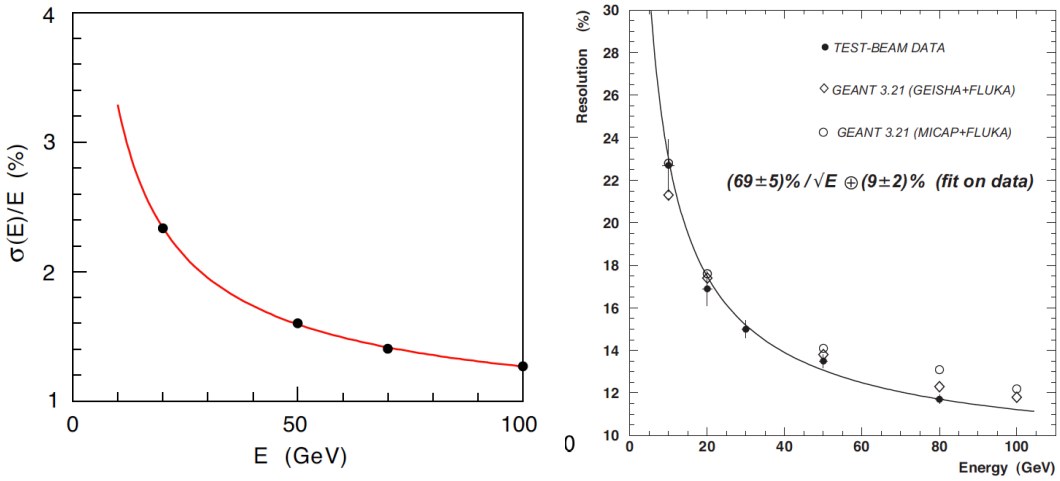


Figure 19: Energy resolution as a function of energy for the ECAL (left) and HCAL (right). Figure from Ref. [50].

self-calibration via the circulation of the source with a hydraulic pump. The absolute calibration provided in this way also allows for cross checks of upstream detector calibrations. From a fit to test beam data at several energies an HCAL resolution of $\sigma_E/E = (69 \pm 5)\%/\sqrt{E} \oplus (9 \pm 2)\%$ is determined, as shown in Fig 19 (right).

3.2.3.3 *Muon system*

The identification of muons by the muon system forms a crucial part of all stages of the trigger and is essential for offline selections, as the presence of a high p_T muon can often be used to distinguish signal from background.

The muon system consists of five stations, M1–M5, covering a total area of 435 m^2 across 1380 chambers. One quarter of a muon chamber is shown schematically in Fig. 20. The stations themselves are divided into 276 chambers (see Fig. 20 left), which are subsequently partitioned into logical pads in the x, y-direction and hence define the resolution. Each station is also divided into four regions (see Fig. 20 right), R1–R4, where the linear dimensions and number of pads in each scale in the ratio $1 : 2 : 4 : 8$, such that the total flux and occupancy

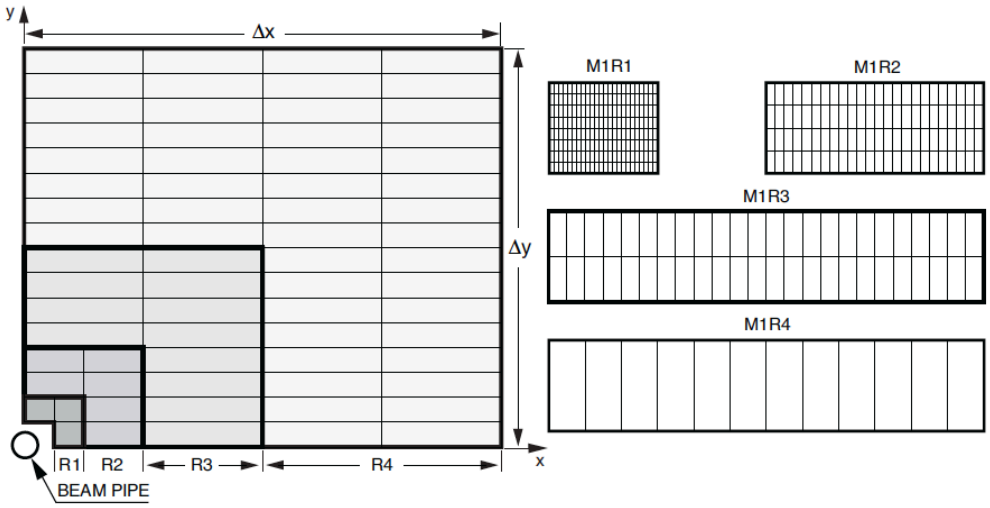


Figure 20: Left: Layout of one quarter of a muon chamber where each segment represents one chamber. Right: Layout of one chamber in each of the different radial regions of M1, where each segment represents one logical pad. Figure from Ref. [50].

are roughly the same in each region. The use of fewer pads at large angles is also justified by increased multiple scattering limiting the spatial resolution in this area of the system. The number of pads in each chamber also depends on the station number, with M2 and M3 containing twice as many pads as M1, whilst M4 and M5 contain half as many, such that the largest number of pads coincides with where the best p_T resolution is required. M1 is located upstream of the calorimeters, whilst M2–M5 are downstream and are interleaved with 80 cm thick iron absorbers, such that the total absorption length is $\sim 20X_0$. The transverse dimensions of the system scale with distance from the interaction point.

With the exception of the inner region (R1) of M1, multi-wire proportional chambers (MWPC) are used throughout the system with a mixture of Ar, CO_2 and CF_4 to ensure an adequate time resolution to cope with the bunch crossings every 25 ns. In M1 the chambers contain two gas gaps to minimise the material in front of the calorimeter, whilst in M2–M5 four gas gaps in two independent

Momentum range	Muon stations
$3\,\text{GeV} < p < 6\,\text{GeV}$	M2 and M3
$6\,\text{GeV} < p < 10\,\text{GeV}$	M2 and M3 and (M4 or M5)
$p > 10\,\text{GeV}$	M2 and M3 and M4 and M5

Table 6: The muon stations in which hits are required to give an IsMuon decision.

layers are used to give extra redundancy. The angular acceptance ranges from 20 mrad to 306 mrad in the horizontal plane and 16 mrad to 258 mrad in the vertical plane.

The MWPCs are insufficiently radiation hard to cope with the charged particle rate of $500\,\text{kHz}\,\text{cm}^{-2}$ in the R1 region of M1 for 10 years of LHCb operation. Therefore triple-GEM chambers with an active area of $20 \times 24\,\text{cm}^2$ are used instead, with each chamber consisting of two triple-GEM detectors superimposed.

For the detector to trigger on a muon hits are required in all five stations, corresponding to a minimum muon momentum of $\sim 6\,\text{GeV}$. The efficiency of each station is sufficient to give a total trigger efficiency of 95% or greater. M1–M3 are used to calculate the track direction and p_T with a resolution of 20%, whilst M4–M5 identify penetrating particles. The system has a built in redundancy so that it can operate with one of M1, M4 or M5 missing, albeit with reduced p_T resolution.

Using the position of hits in M1–M5 as defined in Table 6, a loose binary decision on whether a track is a muon or not can be defined. This decision (known as IsMuon) alone is sufficient to reduce the hadron misidentification rate to $\sim 1\%$ for muon efficiencies of up to 98% as shown in Fig. 21 (left). A muon DLL based on the pattern of hits around a track can be calculated to improve the performance further as shown in Fig 21 (right), and by adding information from the RICH and the calorimeters hadron misidentification probabilities of $< 0.6\%$ can be achieved for muon efficiencies of $\sim 93\%$.

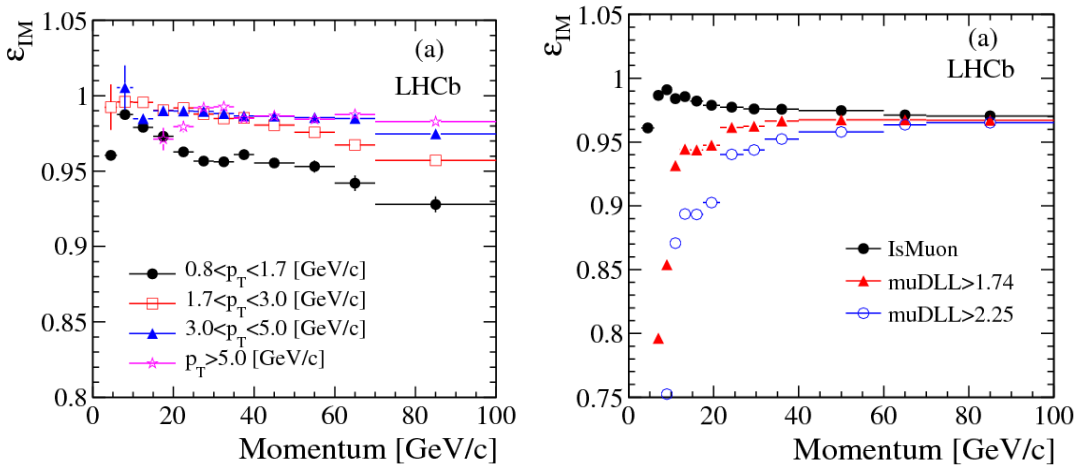


Figure 21: Efficiency of the IsMuon requirement as a function of momentum for different transverse momenta (left) and cuts on the muon DLL (right). Figure from Ref. [56].

The performance of the muon system is described in detail in Refs. [57, 56].

3.2.4 LHCb Performance

The LHCb detector has been online and collecting data since early 2010. The centre-of-mass energy and recorded integrated luminosity are summarised by year in Fig. 22, along with the current expectations for future operations, including the LHCb Upgrade, which is briefly discussed in Section 5.5.2. Lead-lead and lead-proton data taking is not included but the contribution to the total integrated luminosity is negligible. The uncertainty on the integrated luminosity is 3.5%, determined from a combination of 'van der Meer scan' and direct beam imaging methods [58]. It can be seen that the detector has operated at above 90% efficiency throughout this period, with inefficiency arising from detector deadtime or certain subdetectors not being in a fully functional state.

The average luminosity has risen slowly throughout 2010 – 2012, but remains significantly below the LHC luminosity which is delivered to the GPDs, being almost a factor of 20 smaller in 2012. Although a large as possible average lu-

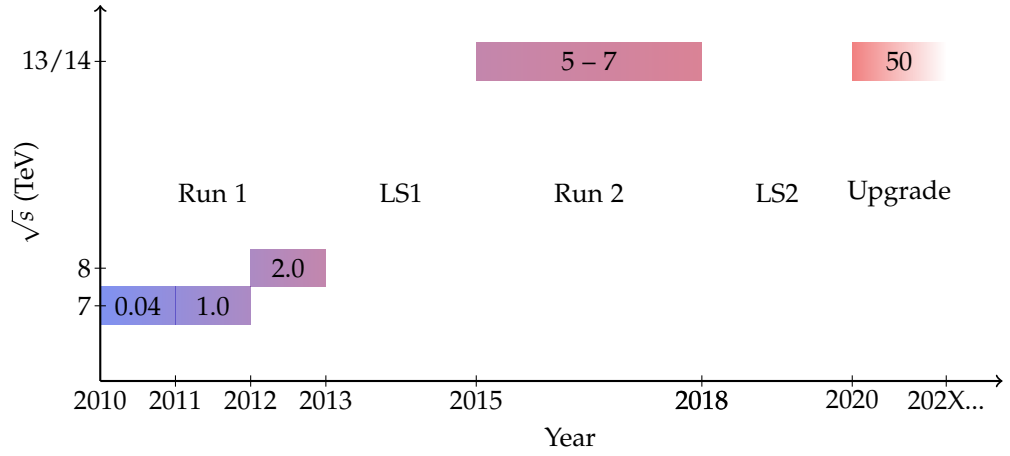


Figure 22: Summary of the three main LHCb data taking periods in Run 1 and the current expectations for future operations. The numbers in the boxes represent the recorded integrated luminosity in fb^{-1} during each period.

minosity is desired by LHCb, there is a hardware imposed limit to the number of events that can be handled per bunch crossing. Therefore whilst the LHC beam circulates at the luminosity required by the GPDs, a technique known as 'luminosity levelling' is employed at the LHCb collision point, where the transverse distance between the beams is varied during a fill to achieve a stable, yet reduced average luminosity. Whilst the maximum luminosity has been as large as $1 \times 10^{33}\,\text{cm}^{-1}\,\text{s}^{-1}$, the luminosity after LS1 is expected to remain the same as the average luminosity of $4 \times 10^{32}\,\text{cm}^{-1}\,\text{s}^{-1}$ in 2012.

3.3 EVENT TRIGGERING

As discussed in the previous section, hardware limitations restrict the maximum luminosity that the detector can receive. Nonetheless the crossing frequency for interactions which are reconstructible within the detector is still around 15 MHz [59], and bandwidth, CPU and offline storage limits require that this be reduced further before offline analysis can take place. A two stage trigger

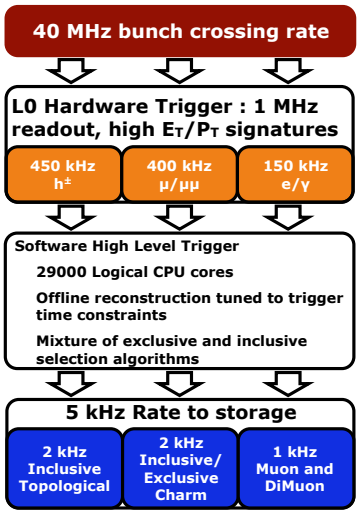


Figure 23: The workflow of the LHCb trigger. The 'Software High Level Trigger' includes both HLT1 and HLT2. Figure from Ref. [60].

system [59] is imposed to achieve this, where the Level-0 trigger (L0) uses a hardware trigger to reduce the rate to around 1 MHz, and then the High Level Trigger (HLT) uses a combination of C++ algorithms to reduce it further to 5 kHz, after which the events are written to storage, as shown in Fig. 23.

The overall performance of the trigger is optimised to achieve the highest efficiency for events likely to be selected in offline analyses of heavy flavour decays, whilst rejecting background events as effectively as possible.

As the data taking environment, computational resources and physics goals of the experiment change over time, the specific details of the trigger lines described in the following sections are also subject to change. For example, over the course of 2010 – 2012 many of the cuts in certain trigger lines have been tightened to reduce the demand for bandwidth, whilst other cuts have been loosened or supplemented by additional trigger lines to allow for the study of a new process. It is crucial to be able to reproduce these trigger conditions offline and thus the combination of algorithms and their cuts is assigned a 32 bit hex-

adecimal Trigger Configuration Key (TCK). A total of 7 TCKs were used for the majority of the 2011 physics data taking, with a further 13 in 2012.

3.3.1 Hardware: L0

The L0 trigger uses information from the calorimeters and muon chambers to attempt to reconstruct the highest E_T hadron, electron and photon clusters, and the two highest p_T muons. It is supplemented by a pile-up system in the VELO which uses the four most upstream R-sensors to estimate the number of primary interactions in a bunch crossing. Furthermore the number of hits in the SPD is used to estimate the number of tracks in the event, and along with the total energy in the calorimeters this is used to reject events with a large number of combinations of tracks which would use a disproportionately large amount of trigger resources. The L0 Decision Unit uses all of the available information to arrive at the final L0 trigger decision, which consists of a decision by one or more of the following channels for the majority of physics studies. The exact cut values are dependent on TCK but examples of the typical values in the 2012 data taking period are shown in brackets. The p_T reconstruction in the muon chambers has an uncertainty of 20 – 30%.

- L0Muon: Searches for one high p_T track in the muon chamber ($> 1.76\,\text{GeV}$).
- L0DiMuon: Searches for pairs of tracks in the muon chambers with the product of their p_T above some threshold ($> 1.6\,\text{GeV}^2$).
- L0Hadron: Searches for particles with a total cluster E_T in the HCAL above some threshold ($> 3.5\,\text{GeV}$).
- L0Photon: Searches for a photon with a total cluster E_T in the ECAL above some threshold ($> 2.72\,\text{GeV}$), hits in several PS cells in front of the ECAL cluster and no hits in the corresponding SPD cells.

- L0Electron: Searches for an electron with a total cluster E_T in the ECAL above some threshold ($> 2.72\,\text{GeV}$), hits in several PS cells in front of the ECAL cluster and at least one hit in the corresponding SPD cells.

Based on the division of bandwidth between the L0 channels, an SPD multiplicity cut of < 900 is applied to L0DiMuon, whilst a cut of < 600 is applied to all other L0 triggers.

3.3.2 Software: HLT1 and HLT2

The HLT refines candidates which passed L0 by adding information in two stages. At HLT1 the 1 MHz of L0 output is reduced to around 30 kHz by requiring tracks to be reconstructed in the VELO and the T-stations and to have a combination of high p_T and large IP. It starts with 'alleys' corresponding to the channels detailed in the previous section, with the aim of confirming the L0 decisions whilst reducing the rate. The position of the PV is also computed in 2D using VELO tracks independently of the alley.

The 30 kHz output rate of HLT1 is then sufficient to allow a full offline track reconstruction using all available data. In HLT2, tracks are combined to form composite particles, e.g. $J/\psi \to \mu^+\mu^-$ and $\phi \to K^+K^-$ and used in all selections to avoid wasting computing resources by recreating the same final state particle in different trigger lines.

Cuts on invariant mass, pointing angle towards the PV and other quantities can then be applied in either inclusive or exclusive selections. An event only needs to pass a single trigger line to be selected by HLT2.

There are a substantially higher number of trigger lines at both HLT1 and HLT2 compared to L0; therefore details of specific lines are omitted here. However those relevant to the analyses presented in Chapter 5 and Appendix A are described in detail there, where applicable.

After the first full year of LHCb data taking in 2011, it became apparent that the HLT CPU farm was largely unused in the periods between fills. Therefore 20% of L0 accepted events were stored to disk and processed interfill, effectively providing a 20% increase in processing power, allowing for reduced p_T thresholds in the HLT and a subsequent increase in trigger efficiency for a number of channels. This method is known as deferred triggering.

More information on the trigger and its performance can be found in Ref. [59].

3.4 LHCB SOFTWARE

The LHCb Software is broadly grouped into one of several projects, depending on its purpose. Of relevance to the following are Gauss, for simulation, Brunel, for reconstruction, DaVinci for physics analysis and Vetra for various VELO tasks.

As the use of simulation is an important aspect of the physics analyses detailed in Chapter 5 and Appendix A a brief description of the Gauss project is given in the following.

3.4.1 Gauss: The LHCb Simulation Program

In LHCb, Gauss is used to produce various types of simulated data, referred to in the following as Monte Carlo (MC). Typically, this can be separated into two steps, generation and simulation. In the first phase pp collisions are generated using an event generator with a specific LHCb configuration. In almost all cases this is Pythia, with Pythia 6 (written in Fortran) being used in 2010 and 2011 and Pythia 8 (written in C++) being phased in as of 2012. The generated particles are subsequently decayed according to the EvtGen [61] program in which final-state radiation is generated using Photos [62]. Various models exist

for describing the distribution of particles in the final state, but typically a phase-space model is used, unless otherwise specified.

The second phase corresponds to the interaction of the generated particles with the detector, and the response to these interactions. This is implemented using the GEANT4 toolkit [63, 64] as described in Ref. [65]. It is possible to perform the generation step alone, and this is often done to perform basic calculations, as the second step is the most CPU intensive part of the process by several orders of magnitude.

In order to produce MC corresponding to a particular inclusive or exclusive decay within the LHCb framework a decay file must first be created. This contains information on the branching fractions (BF) of the decays of interest and is able to be varied from SM values, such that large numbers of extremely rare or forbidden events can be generated. It is also possible to include cuts on certain kinematic variables, such that all events created pass a particular requirement e.g. $p_T > 300\,\text{MeV}$. This prevents the generation and simulation of large numbers of events which will subsequently be rejected by the trigger or offline selections. These are referred to in the following as generator level cuts.

3.4.2 Stripping

As described previously, the selection of interesting events by the trigger is an essential step in reducing the amount of data collected by the experiment to a manageable level. Following this initial selection the associated data (referred to in the following as 'raw data') are stored on disk and are effectively structurally identical to the raw simulated data produced by GAUSS. Both types of data then undergo processing, where physical objects such as track hits and calorimeter clusters are reconstructed and PID hypotheses are calculated for the associated proto-particles. The term proto-particle is used to describe a particle-like object that is yet to be associated with a particular mass hypothesis, charge, momentum etc.

Stripping campaign	Year data collected	$\sqrt{s}$ (TeV)	$\mathcal{L}$ (fb^{-1})
S17	2011	7	1.02
S20r1	2011	7	1.02
S20	2012	8	2.06

Table 7: Summary of the three stripping campaigns considered in this thesis.

This processing results in the creation of a Data Summary Tape (DST) which contains all the necessary information for further pre-selection algorithms to be run. However, as these DSTs are extremely large ($\sim$tens of TB) it would be prohibitively resource-intensive to allow LHCb users direct access to them. Therefore in advance of the data taking a given analysis group can define a collection of cuts known as a 'stripping selection', with the aim of loosely selecting their events of interest. These stripping selections are then collectively run over the DSTs in 'streams', which contain selections of a similar type. For example, in the analyses described in Chapter 5 and Appendix A the DiMuon stream, which contains a number of lines with $\mu\mu X$ final states (where X can be 0 or more particles), is used. The output of the stripping is a new DST (referred to as simply the DST in the following) of a more manageable size which can be processed by individual users. As the number of raw events in simulated data is typically much smaller than in real data, in most cases the stripping is run over them in 'flagged' mode, where the decision of a particular stripping stream is stored, but events which fail to pass it are still retained.

Another benefit of the stripping is the ability to be able to accommodate improvements in the various reconstruction algorithms, detector calibrations and detector alignments which contribute to the quality of the data processing. As such raw data are first stripped promptly after they are collected and are then periodically 'restripped' as required, but typically once a year as part of a full

scale stripping campaign. The different stripping campaigns relevant to the analyses presented in this thesis are shown in Table 7.

Finally a number of other types of raw data are collected over the course of data taking which are used for detector calibration, minimum bias physics, etc., and are not subject to the trigger and stripping requirements described above. This is usually achieved by recording the data at a reduced bandwidth (e.g. 10 Hz) so as to limit the storage space required. Data of this type are relevant to the VELO studies described in Chapter 4.

4 RADIATION DAMAGE IN THE LHCB VERTEX LOCATOR

In this chapter results on several studies of radiation damage in the LHCb VELO detector are described. Section 4.1 gives an overview of the theory of particle tracking in silicon detectors and the basic properties of these detectors, whilst Section 4.2 describes how these properties can undergo radiation induced changes. Section 4.3 gives details of the silicon sensors used in the LHCb VELO, and Section 4.4 explains how some of their properties can be determined using specific datasets. Sections 4.5 – 4.8 describe the results of studies of these datasets, and a summary is given in Section 4.9.

4.1 PARTICLE TRACKING IN SILICON DETECTORS

As described in Chapter 3, the tracking of charged particles in particle physics experiments is an essential part of the full event reconstruction. Whilst a variety of technologies exist for achieving this, such as solid state (RICH, scintillator counters) and gaseous (time projection chambers, MWPC) detectors, semiconductor detectors are currently the best choice for achieving high precision. As cost considerations often limit their use, they are placed where the highest resolution is required, such as around the interaction point in the LHCb detector.

Of the several possible material choices for a semiconductor detector, silicon is the most common, and has been used in many of the major collider experiments in recent decades. Requirements on the cost and quality of the silicon used for

particle physics experiments mean that the silicon is typically produced via the float zone method [66] or the Czochralski process [67].

Silicon is an abundant chemical element, distributed throughout nature in a variety of compounds, such as silicon dioxide (SiO_2). With an atomic number and group number of 14, it has 14 electrons in three shells, in a 2, 8, 4 arrangement. As a result of this, a silicon atom can covalently bond with four others such that its outer shell becomes completely full, forming a crystalline lattice. The periodic potential of the crystalline structure results in the formation of electron energy bands. At a temperature of absolute zero the highest energy electrons lie in the 'valence band'. At higher temperatures electrons also occupy higher energy states known as the 'conduction band'.

At absolute zero the filled valence band is separated from the empty conduction band by the band gap energy, E_g, and the silicon acts as an insulator. Figure 24 shows the classification of materials according to their band gap energy. As the temperature of the silicon is increased, electrons are thermally excited from the valence band into the conduction band and the silicon becomes semiconducting. The removal of an electron from the valence band gives rise to an absence of negative charge, or 'hole', which also carries charge. The band gap energy of silicon is 1.12 eV at room temperature [68], but since an increase in temperature causes atomic vibrations to increase and the interatomic spacing to get larger, the band gap energy is temperature dependent. For the temperature range of interest in the VELO an 'effective band gap energy' of 1.21 eV is expected [69]. However, an average value of 1.16 eV is determined from sensor current measurements [70], with a number of reasons proposed for this discrepancy [69]. The band gap in silicon is 'indirect', such that the maximal-energy state in the valence band and the minimal-energy state in the conduction band occur at different momentum values.

The conductivity of a semiconductor depends on the value of E_g/k_BT, where k_B is the Boltzmann constant and T is the temperature. It is typically many orders of magnitude smaller than that of a conductor, where there is zero band

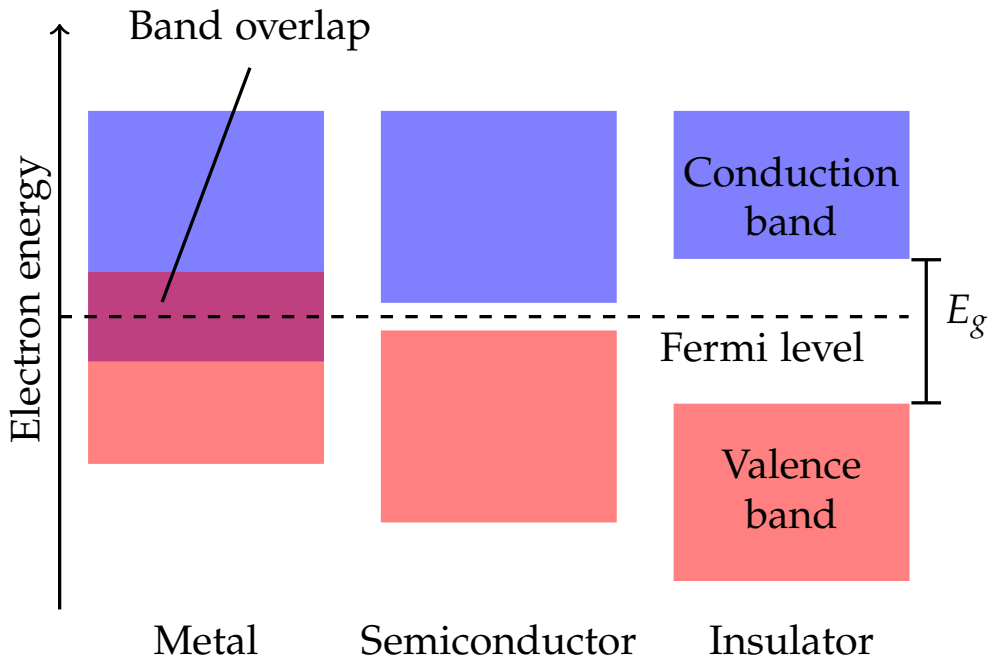


Figure 24: Classification of metals, semiconductors and insulators according to their band gap energy.

gap as the energy bands overlap. The lack of free electrons in the crystalline structure gives pure silicon a low conductivity at room temperature, such that its semiconducting properties are limited. However, via the addition of small quantities of impurities, or 'doping', the conductivity of silicon can be greatly increased. These impurities are classified as either 'n-type' or 'p-type', depending on the charge of the resulting conducting particles. The process of n-type doping adds a group 15 element (phosphorus or arsenic) such that after bonding, the additional electron in the dopant is free to move through the lattice. Similarly, p-type doping adds a group 13 element (boron or gallium) such that after bonding, the absence of an electron in the lattice creates the effect of a positive charge, or 'hole', which can carry charge. The electrons and holes are referred to as negative and positive charge carriers respectively.

Figure 25 shows a 'p-n junction', formed at the boundary between an n-type and p-type semiconductor. The electrons diffuse into the p-type side and the

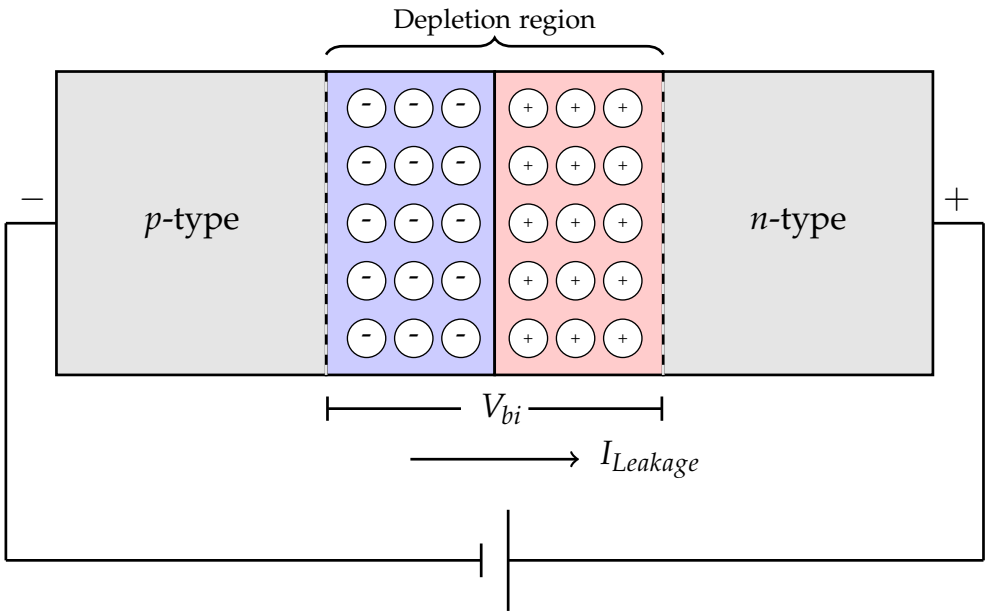


Figure 25: Schematic diagram of a reverse biased p-n junction.

holes diffuse into the n-type side, and the two charged species recombine, creating a 'space charge region' where there are a net number of positive and negative ions in the n-type side and p-type side respectively. An electric field is established across the junction with 'built-in' potential difference V_{bi}, which opposes further charge transfer across the junction. A small reverse current, or 'leakage current', still flows across the junction due to the thermal excitation of electrons from the valence band to the conduction band. The magnitude of this current can be controlled by reducing the temperature of the diode.

Due to the absence of free charge carriers in the space charge region it is referred to as the 'depletion region'. In an 'abrupt p-n junction', where one side is more heavily-doped than the other, the depletion region extends further into the less heavily-doped side. It can be assumed that the density of doping atoms changes abruptly from one side of the junction to the other and the space charge is constant in the depleted region. This arrangement is typically used in silicon detectors [71].

If a potential difference is applied in the same direction as V_{bi}, the size of the depleted region grows. This 'reverse bias' prevents the passage of large currents and can be increased until the entire diode is depleted. The corresponding voltage required to achieve this is called the 'depletion voltage', V_{dep}, and is related to the effective doping concentration, N_{eff}, by

$$V_{dep} + V_{bi} = \frac{q_0}{2\epsilon_r\epsilon_0}|N_{eff}|d^2.$$

Here q_0 is the electron charge, d is the depth of the depleted region and $\epsilon_r\epsilon_0$ is the permittivity of silicon. The effective doping concentration is defined as

$$N_{eff} = |N_D - N_A|,$$

where N_D and N_A are the concentrations of donors and acceptors respectively. Typically the depletion voltage is more than an order of magnitude higher than V_{bi} and the contribution of V_{bi} can be ignored. The reverse bias cannot be increased indefinitely, as above a certain 'breakdown voltage' the current flowing across the junction rapidly increases and can potentially destroy the device. Therefore if V_{dep} increases (as a result of radiation damage for example) and approaches the breakdown voltage, it is possible that the diode will be unable to be operated fully depleted.

A charged particle traversing silicon will create electron-hole pairs as it interacts with the atomic lattice and deposits energy. Due to the indirect band gap in silicon, on average, an electron-hole pair is created for every 3.6 eV of deposited energy, as a change in electron momentum is required in addition to the band gap energy [72]. If this energy deposition occurs within the depleted region, the presence of the electric field pulls the charges in opposite directions, such that they can be collected by contacts at the edge of the silicon. If instead, this process occurs outside of the depleted region, the electron-hole pairs typically recombine or diffuse randomly through the lattice due to the absence of a potential, and will either be collected over a much longer period, or not be collected at all.

By measuring the position and magnitude of the induced charges, the position and energy of the incident particle can be determined. This forms the basis of particle detection using silicon.

In order to achieve maximum particle tracking efficiency, a silicon detector must collect all of the deposited charge, such that the 'charge collection efficiency' (CCE) is 100%. As this is most likely to occur if the depleted region extends across the whole silicon sensor, it is desirable to operate the sensor at a voltage equal to, or above the depletion voltage.

4.2 RADIATION DAMAGE

Silicon detectors subjected to irradiation undergo radiation damage effects which can be divided into two types: surface and bulk damage. A brief overview of bulk damage effects is given in the following, whilst further information on surface damage effects can be found in Ref. [73].

Particles incident on a silicon detector can undergo non-ionising energy loss (NIEL) interactions with a lattice silicon atom [74]. A schematic view of some of the types of defects resulting from this process can be seen in Fig. 26. If the energy transferred in the interaction is sufficient to remove the silicon atom from its position in the lattice, then the atom leaves behind a 'vacancy' and forms an 'interstitial', which are collectively known as a 'Frenkel pair'. If the displaced atom moves to the surface a 'Schottky defect' is created. The silicon atoms displaced in this way can also remove further atoms from the silicon lattice in a knock-on effect, resulting in clusters of high concentrations of Frenkel pairs. Whilst the majority of the vacancies formed in this way recombine with interstitials, a number combine with other vacancies to form 'divacancy' structures, or react with impurities in the silicon crystal. Under the NIEL hypothesis, the total displacement damage cross section is proportional to the energy deposited in the displacement interactions. At the LHC, this is predominantly from the inter-

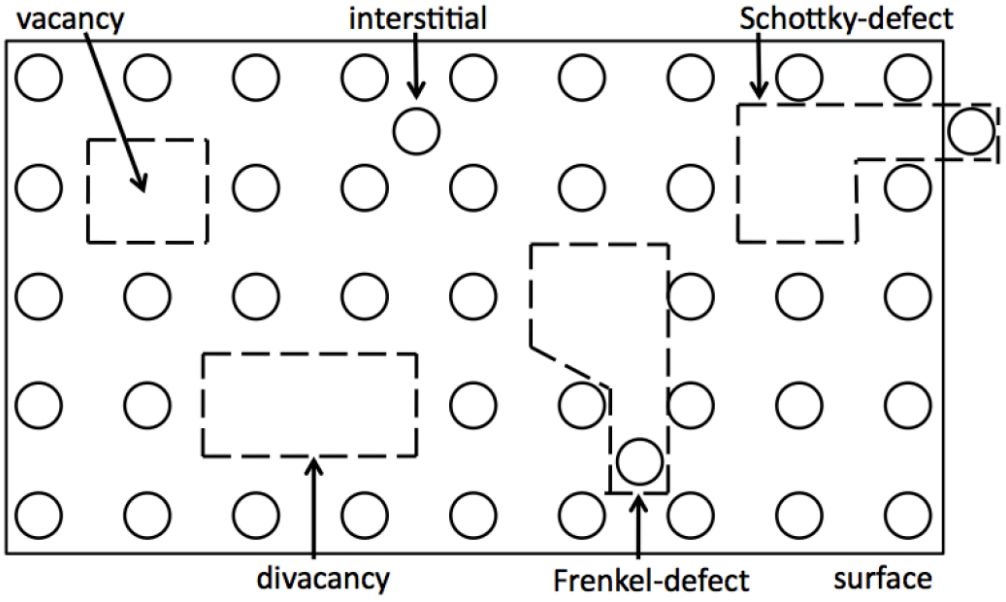


Figure 26: Examples of radiation induced defects in a silicon lattice. Figure from Ref. [75].

actions of protons, neutrons or pions with the silicon, with a varying degree of damage arising from each.

Thermally stable defects can lead to changes in the structure of the energy levels in the band gap and have a direct influence on the electrical properties of the silicon. If a defect level is present in the band gap, then an electron can be excited from the valence band to the conduction band via the defect and increase the leakage current. This process has been well measured in the VELO, where the sensor leakage currents have been observed to increase at a rate directly proportional to the delivered integrated luminosity [76], as shown in Fig. 27.

In a similar process, the presence of an intermediate defect level can also lead to 'charge trapping', where a carrier from the conduction band is trapped before being emitted back to the conduction band. As a result the total time required to collect a fixed amount of charge is increased and the carriers themselves become less mobile. In a detector such as the VELO where the collection of the deposited charge occurs over a fixed period of time, this will have the effect of reducing

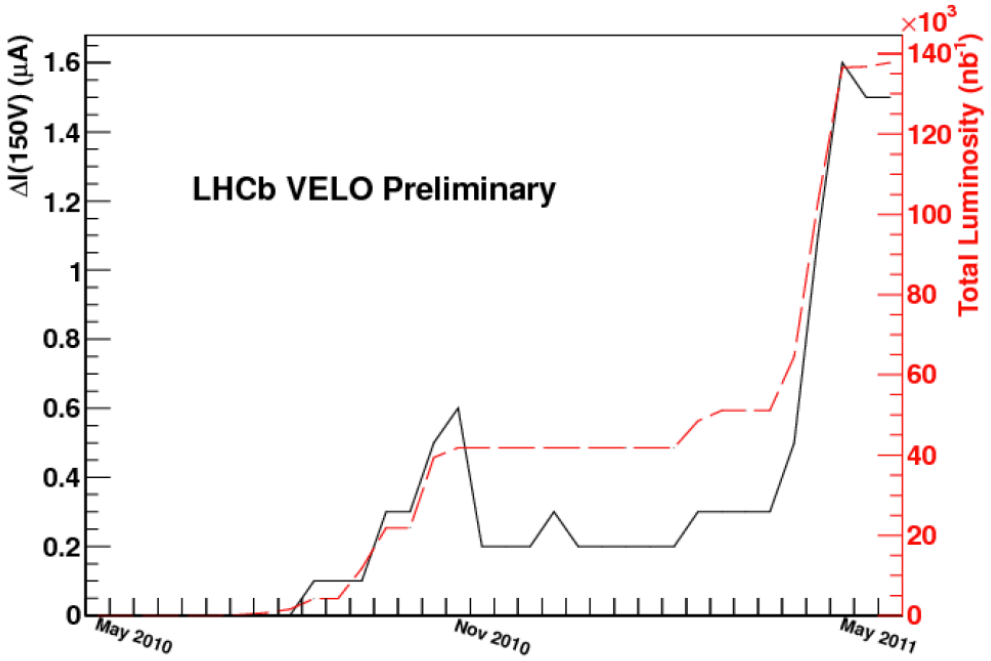


Figure 27: Change in current for a VELO sensor as a function of time and delivered luminosity. The sharp decrease in the current around November 2010 corresponds to the start of a period of annealing between 2010 and 2011. Figure from Ref. [77].

the total integrated charge and is expected to have a strong influence on the operational lifetime of the detector.

Radiation damage also leads to a change in the effective doping concentration of the silicon (defined in Section 4.1). As n-type silicon is irradiated, acceptor defects are introduced whilst donor defects are removed (for example, via their combination with vacancies or carbon interstitials), and N_{eff} is found to decrease until the acceptor concentration outweighs the donor concentration. At this point the diode effectively inverts to p-type and is said to have 'type-inverted'. Subsequently, the process continues, and the p-type N_{eff} increases with the further addition of acceptor defects and the removal of donor defects. This mechanism is illustrated in Fig. 28, with type-inversion occurring at approximately $2 \times 10^{12}\, n_{eq}\, cm^{-2}$.

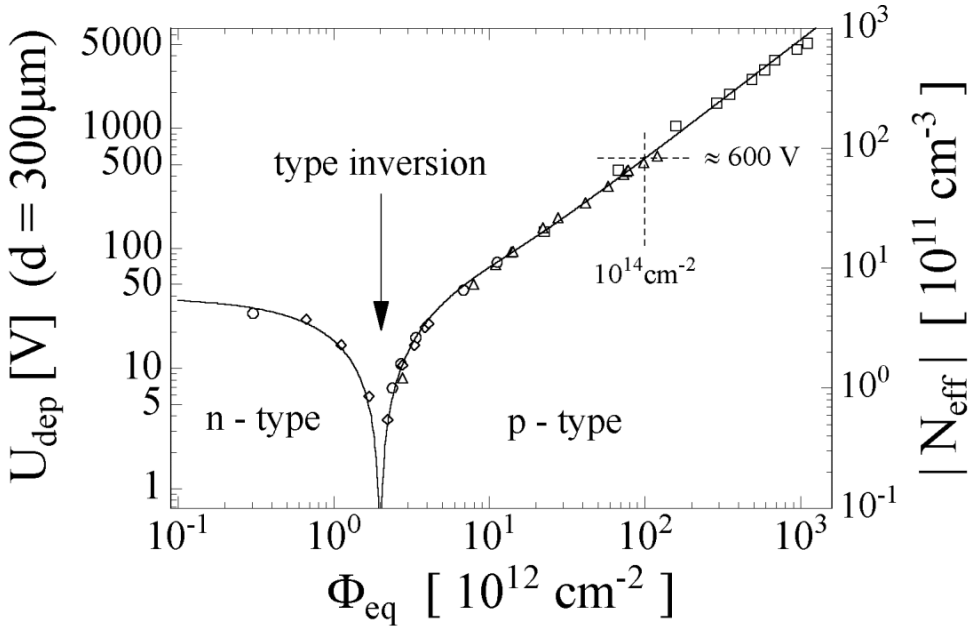


Figure 28: Change in the depletion voltage and the effective doping concentration as a function of fluence. Figure from Ref. [78].

For a number of the changes due to radiation damage, the long term effects are partially dependent on the time after irradiation. In the case of inverted n-type silicon, if the detector is stored at room temperature initially the p-type N_{eff} is found to decrease for a number of days [79], as defects move through the silicon and interact with other defects. This process is known as 'annealing' and allows the possibility of reversing some of the radiation damage by careful control of the detector temperature for a fixed period after irradiation. However, following the beneficial annealing, N_{eff} is then observed ultimately to increase once again, as defects interact further to form more complex defects, such that the magnitude of the recovery afforded by this method is limited. Nonetheless, it is an important option to consider in the long term operation of a silicon detector.

Parameter	Value	
	R	ϕ
Oxygen enhancement	$> 1 \times 10^{17}\,\mathrm{cm}^{-3}$	
Silicon thickness	300 µm	
Number of strips	2048	
Strip width	11 – 38 µm	
Routing line width	~ 11 µm	
Stereo angle	-	10 – 20°
Strip pitch	40 – 102 µm	38 – 97 µm
Strip length	3.8 – 33.8 mm	5.9 – 24.9 mm

Table 8: Parameters of the VELO sensors.

4.3 THE VERTEX LOCATOR

The properties of the silicon sensors used in the VELO are summarised in Table 8. The silicon is enriched with oxygen to improve its radiation hardness, as oxygen can interact with vacancies and interstitials created as a result of radiation damage. Further information on the mechanisms causing this improvement in performance can be found in Refs. [80, 81].

The rate of energy loss in the VELO detector is close to a minimum at the momentum values typical of the particles produced at the LHC. According to the distribution of deposited energy versus momentum from the Bethe-Bloch equation [3], the MPV of deposited charge is 0.7 of the mean, as shown in Fig. 29. Therefore as the VELO sensors are 300 µm thick, and the mean energy loss of a minimum ionising particle in silicon is 388 eV/µm [3], a particle interaction with a VELO sensor generates approximately 22 400 electron-hole pairs on average.

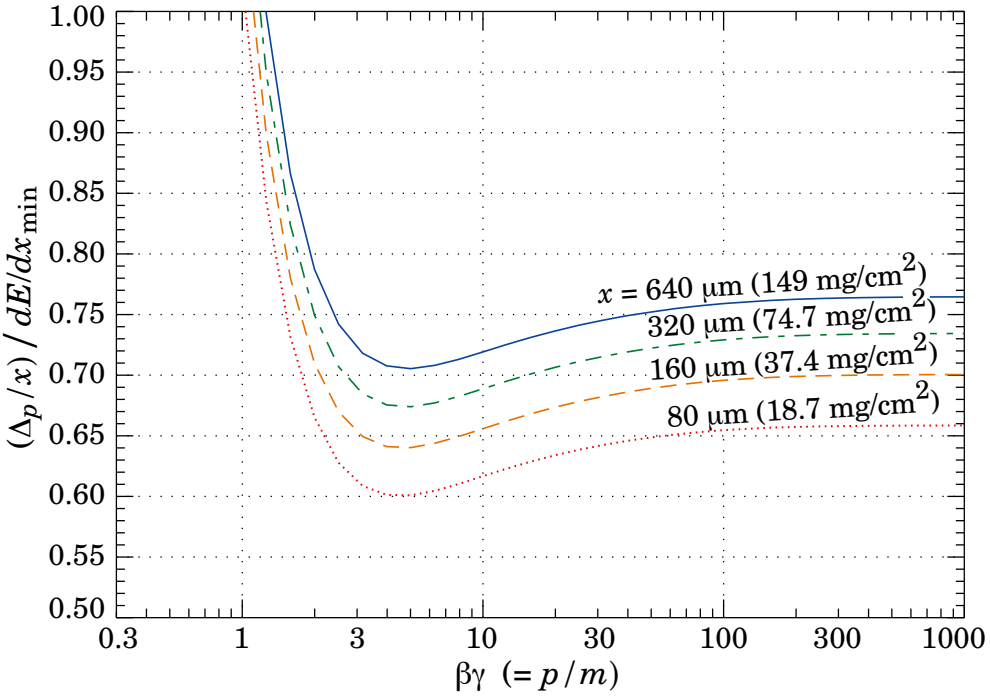


Figure 29: Most probable energy loss in silicon for different sensor thickness, scaled to the mean loss of a minimum ionising particle. Figure from Ref. [3].

For a typical VELO sensor before irradiation, the MPV of the ADC distribution is around 35 counts, such that one ADC count corresponds to 640 electrons.

Of the 84 standard VELO sensors (excluding the pile-up sensors), 82 have an n^+-on-n structure, with a highly-doped n^+ implant in an n-type bulk. In the following these are referred to as the 'n-type' sensors. The remaining two sensors are n^+-on-p, and are the only two of their type in an LHC detector. They are located at the most upstream end of the VELO, where their potential as a sensor technology for use in LHC upgrade detectors can be studied without risk of greatly affecting the detector performance. In the following they are referred to as the 'p-type' sensors. Unless otherwise specified the default sensor type shown is n-type. For further information on the p-type sensors see Refs. [82, 83].

Figure 30 shows a schematic diagram of the cross-section of an n-type R sensor. The highly-doped p^+ backplane and lightly-doped bulk cause the depleted region to extend further into the n-type bulk. The detector is said to be fully

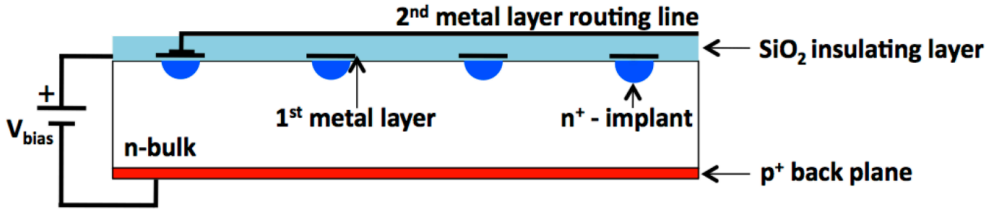


Figure 30: Schematic cross-section of an n-type R sensor. The strips (first metal layer) and n^+ implants run into and out of the page, whilst the routing line runs horizontally. Figure from Ref. [83].

depleted when the depleted region extends through the full substrate to the highly-doped n^+ implants on the front surface. The charge deposited in the silicon travels along the electric field lines within the sensor, which terminate at the n^+ implants. The charge is then collected by the first metal layer (or 'strips' in the following) which is coupled to the implants. The segmentation of the sensor surface via these implants allows for the determination of the position of the deposited charge and ensures that the majority of the signal reaches the first metal layer. Each sensor has 2048 strips, the approximate layout of which is shown in Fig. 31. The strip pitch varies as a function of radius, with a pitch of 40 µm at the innermost radial region of the sensors and a pitch of 120 µm at the outermost radial region of the sensors. The change in strip pitch improves the vertex resolution, and allows for each measurement along a track to contribute to the IP resolution with an approximately equal weight. The width of the strips also varies, from a minimum of 11 µm in the innermost region, to a maximum of 38 µm in the outermost region.

In an R sensor the strips are arranged into four 45° segments containing 512 strips each to reduce the occupancy. The strips are placed perpendicular to the radial direction, with pitch linearly increasing with radius. In a ϕ sensor the strips are split into inner and outer radial regions, containing 683 and 1365 strips respectively. The inner region has a strip pitch varying from 38 – 78 µm whilst the outer region has a strip pitch varying from 39 – 97 µm. The transition

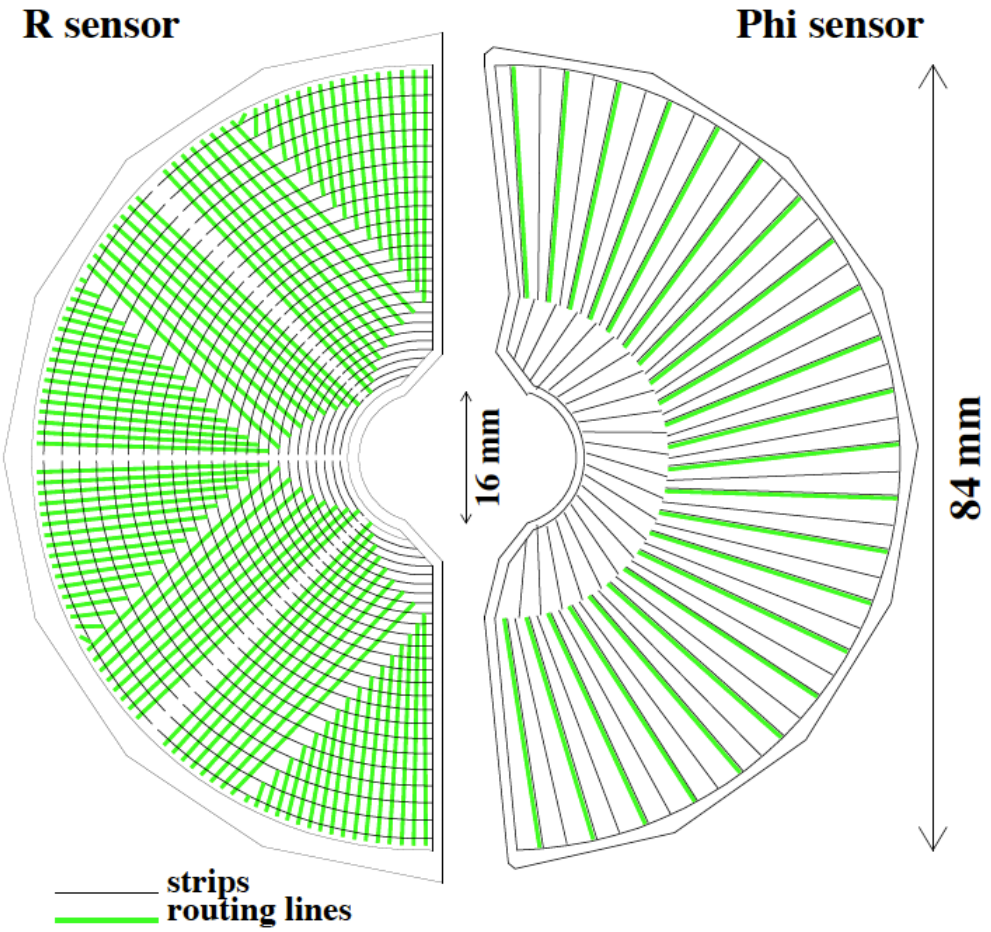


Figure 31: Schematic distribution of the strips and routing lines in R and ϕ sensors. Figure from Ref. [83].

from the inner to the outer region occurs at approximately 17 mm. The strips are skewed by 20° and 10° with respect to the radial direction in the inner and outer regions respectively, to improve pattern recognition and allow ghost hits to be identified.

The readout of the sensors occurs on their outer radial edge. Therefore for all of the strips in an R sensor, and the strips in the inner radial region of a ϕ sensor, the charge must be transported to the outer edge of the sensor via 'routing lines' (the second metal layer), as indicated by the green lines in Fig. 31. The routing lines are approximately 11 µm wide and are insulated from the rest of the sensor by 3.8 ± 0.3 µm of SiO_2. The arrangement of the routing lines is

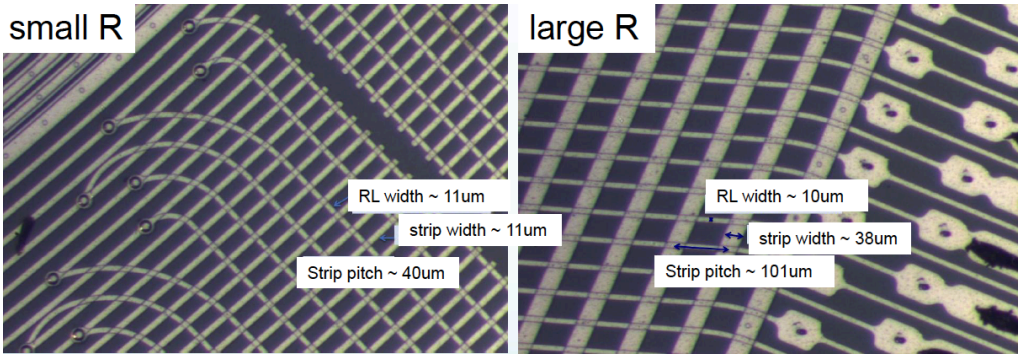


Figure 32: Structure of the routing lines (RL) and strips of an R sensor at small radius (left) and large radius (right).

such that they run perpendicular to the strips in an R sensor, and parallel to the strips in a ϕ sensor. The structure of the routing lines and the strips in an R sensor can be seen in more detail in Fig. 32. In the following, the terms 'routing lines' and 'second metal layer' are used interchangeably.

4.4 CHARGE COLLECTION EFFICIENCY SCANS

Before the installation of the VELO sensors into the detector, the capacitance of each sensor at a number of bias voltages was measured. By plotting the inverse of the capacitance squared against the bias voltage, the depletion voltage can be determined as the point at which the distribution plateaus. In the following this is referred to as the 'CV method' [84]. A range of depletion voltages were observed for the R and ϕ sensors, with a maximum of approximately 65 V.

After installation this direct method of measurement is no longer possible. Therefore an analogous quantity known as the 'effective depletion voltage' (EDV) is used to quantify the change in CCE due to radiation damage. It is also possible to obtain information on the depletion voltage by measuring the noise present in the sensors prior to type-inversion [83], but this method is not considered here.

4.4.1 Determination of the Effective Depletion Voltage

During typical LHCb data taking in 2010 to 2013, all VELO sensors were operated at a nominal bias voltage of 150 V to ensure consistent performance across the detector. In order to measure the EDV, data must be collected at a range of sensor voltages, in a process known as a 'CCE scan'. Therefore the beam time available for such measurements is restricted as they are performed at the expense of physics data. This imposes a limitation on both the number of CCE scans and their length, such that only a small number of bias voltages can be sampled, with limited statistics at each voltage.

Table 9 gives details of the nine CCE scans collected up until the end of Run 1, with the net delivered integrated luminosity up to the time of each scan. The points at which the scans occurred were dictated both by the availability of the detector for such a scan, and the requirement that the increase in delivered integrated luminosity between scans should not be too large. A short period of annealing occurred between the fifth scan at the end of 2011 and the sixth scan at the start of the 2012 data taking, due to the replacement of the cooling line insulation.

Initially only seven voltages were studied with a maximum of 150 V, but as further CCE scans were taken extra voltages were added to improve the accuracy of the method. Voltages beyond the nominal bias voltage were added from July 2012, as it became apparent that the nominal bias voltage may have to be increased. This will be described in detail later in this chapter. The final configuration of thirteen voltages is likely to be the maximum possible number due to timing constraints. In the future when even higher voltages are required, lower voltages may have to be omitted, or the difference between subsequent voltages steps may have to be increased.

The EDV method makes use of the fact that the amount of charge collected by an under-depleted silicon strip increases with the bias voltage, until the silicon becomes fully depleted. To determine with a CCE scan the bias voltage at which

Date	$\sqrt{s}$ (TeV)	Del. $\mathcal{L}$ (fb^{-1})	Bias voltages sampled
2010	7	0	0, 10, 30, 50, 70, 100, 150
April 2011	7	0.04	0, 10, 20
July 2011	7	0.4	30, 40, 50
September 2011	7	0.8	60, 70, 80
October 2011	7	1.2	90, 100, 150
April 2012	8	1.2	0, 10, 20, 30 40, 50, 60, 70 80, 100, 120, 150
July 2012	8	1.9	10, 20, 30, 40
September 2012	8	2.7	50, 60, 70, 80
January 2013	8	3.4	100, 120, 150, 170, 200

Table 9: Details of the nine CCE scans collected up until the end of Run 1. Between the fifth and sixth scan there was a short period of annealing.

this occurs for a given sensor, the sensor of interest, known as the 'test sensor', is operated at one of the voltages listed in Table 9. The four sensors either side of it (VELO geometry permitting) remain at the nominal bias voltage. The test sensor is then removed from the tracking algorithms, and the position of a particle track in the test sensor is determined by extrapolating the hit positions from the surrounding sensors, as shown in Fig. 33.

This one-in-five pattern allows for the measurement of multiple sensors simultaneously. The total charge deposited in the five strips nearest to the extrapolated position is recorded, resulting in a distribution of analogue-to-digital converter (ADC) counts for each sensor at each voltage. In practice each ADC value has a fixed number of ADC counts (or 'pedestal') removed such that only

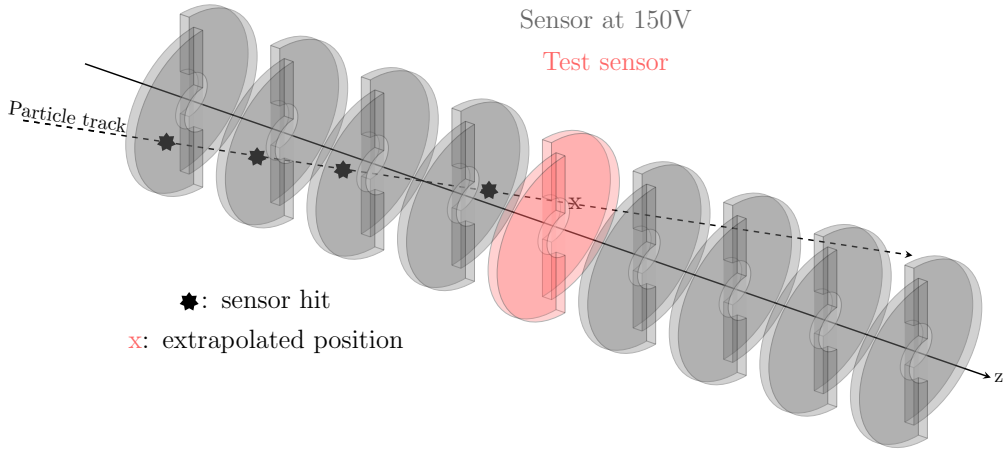


Figure 33: Schematic illustration of the method for determining the EDV and the CFE.

the deposited charge relevant to the particle track is recorded. This pedestal-subtracted distribution is then fitted with a gaussian function convoluted with a landau function, as shown in Fig. 34 (left) for several voltages. The MPV at each voltage is determined from the fit and plotted as a function of bias voltage, as shown in Fig. 34 (right). It can be seen that once a certain bias voltage is exceeded, the MPV of the ADC distribution reaches a plateau, indicating the sensor is fully depleted. The EDV is then defined as the voltage at which a sensor would have an MPV of its ADC distribution that is 80% of the plateau value. This 80% value was chosen to give the best agreement between the predetermined depletion voltage and the EDV method from the first CCE scan, which occurred prior to any irradiation. As shown in Fig. 35 the agreement between the CV and EDV methods is generally very good, with a difference between the EDV and depletion voltage of less than 10 V for the majority of sensors. In the following a bias voltage of equal to, or above the EDV is assumed to be sufficient to fully deplete a sensor.

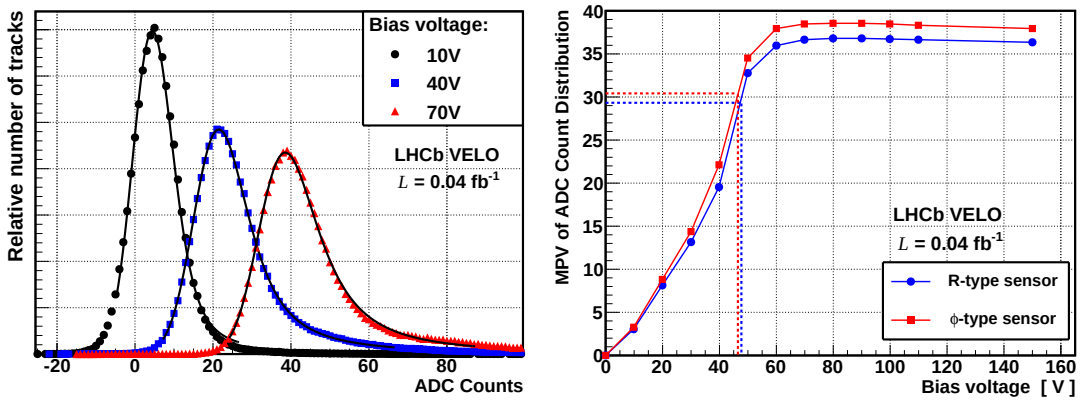
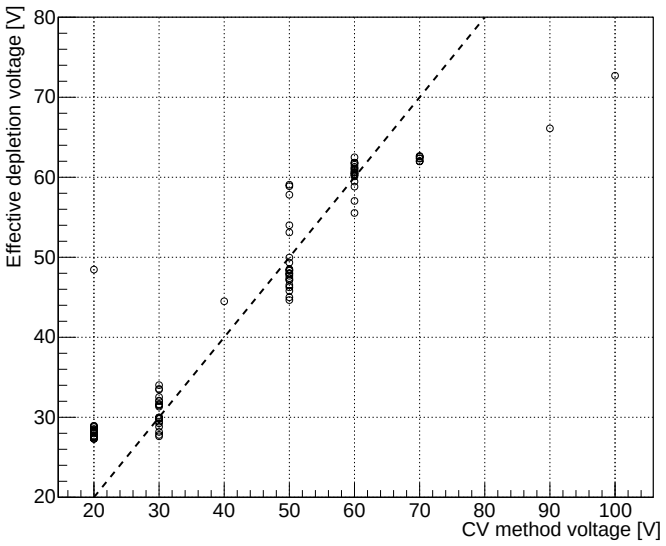


Figure 34: Left: Fits to the pedestal-subtracted ADC distributions at different bias voltages for an R sensor. Right: MPV of the ADC distribution as a function of bias voltage for an R sensor. The EDV is indicated by the vertical dashed lines. Figures from Ref. [83].

4.5 CHARGE COLLECTION EFFICIENCY

Once the EDV for each sensor is determined, the behaviour as a function of delivered integrated luminosity can be studied. In practice, rather than considering the EDV of an entire VELO sensor, each sensor is split into five radial regions as the amount of radiation received has a strong radial dependence. The fluence varies with sensor radius and z-position of the sensor, as shown in Fig. 36, where the fluences are determined from simulation with an uncertainty of 8%. In the figure the term 'station' refers to an R or ϕ sensor. The fluence as a function of radius is fitted with the function Ar^{-k} for each sensor, and the variation of the fitted exponent, k, with z-position of the sensor is shown in the top right. The fluences follow an approximately $1/r^{1.75}$ dependence [83].

The sizes of the five radial regions are defined such that the fluence changes by approximately a factor of two across a region. This definition and the information in Fig. 36 can be combined to give the expected fluence per inverse femtobarn of delivered luminosity for each radial region of each sensor. This is shown in Fig. 37 and is used as an input to many of the calculations in the



CV voltage [V]	N
20	18
30	19
40	1
50	21
60	17
70	6
90	1
100	1

Figure 35: Left: Comparison of the EDV and the depletion voltage measured from the CV method. The dashed line indicates equal EDV and CV method voltage. In the CV method the measured voltages are rounded to the nearest 10 volts. Right: Number of VELO sensors at each depletion voltage as determined by the CV method.

remainder of the chapter. For ϕ sensors the expected fluence of the R sensor at the same z-position is used. For a given radial region, the fluence per inverse femtobarn can vary by as much as close to a factor of two, depending on the z-position of the sensor.

Figure 38 shows the EDV in bins of sensor radius for an n-type R sensor for each of the nine CCE scans taken up until the end of Run 1. Here, and in many of the following plots, the uncertainties on the variables derived from the EDV method are not included. However, these are considered in Section 4.8. If only the innermost radial region is considered it can be seen that at first the EDV decreases to a minimum at $0.4\,\text{fb}^{-1}$ due to bulk radiation damage. This is subsequently followed by an increase in EDV, indicating that the silicon has type-inverted in this region between 0.4 and $0.8\,\text{fb}^{-1}$. With the exception of the decrease in EDV between the fifth and sixth CCE scans due to annealing, the

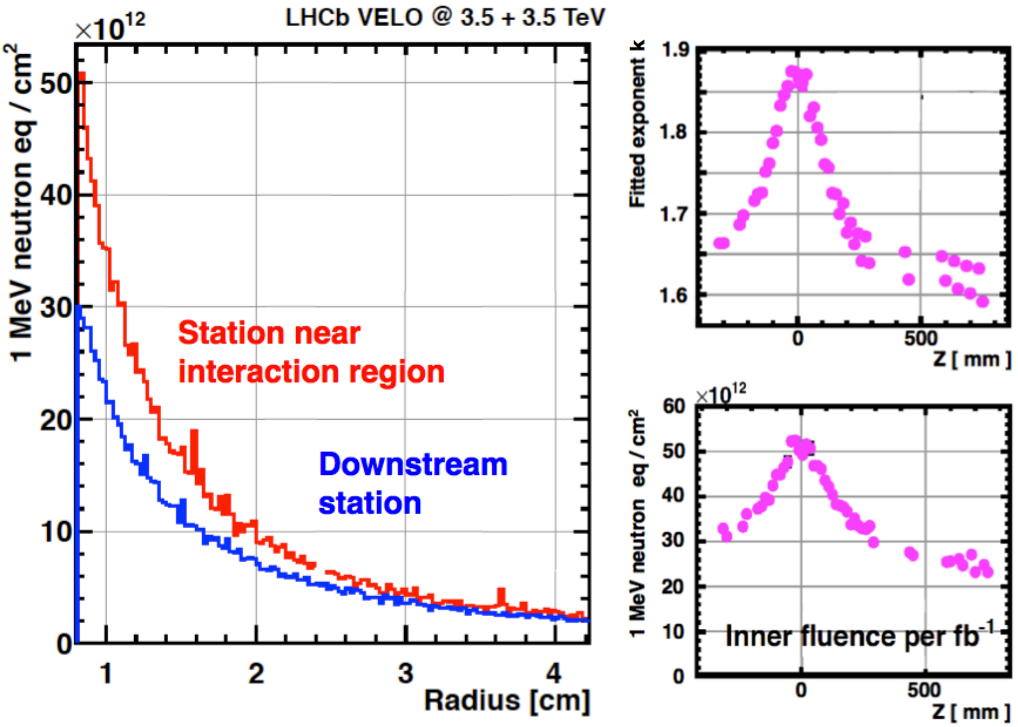


Figure 36: Variation of fluence with radius (left) and z-position of the module (bottom right) as determined from MC at 7 TeV. The exact radial dependence of the fluence varies with z-position of the module (top right). Figures from Ref. [83].

EDV then continues to rise, such that after 3.4 fb^{-1} it is larger than its value prior to irradiation. The other radial regions also experience an initial decrease due to bulk radiation damage, but as a result of the strong radial dependence on fluence only the next two innermost radial regions are subject to enough fluence to undergo type-inversion during Run 1.

As the VELO sensors are subject to large sensor-by-sensor variations in initial depletion voltage, defects, bad strips and other effects, it is useful to study the behaviour of all sensors. By plotting the EDV as a function of fluence rather than integrated luminosity, all radial regions and z-positions can be considered at the same time. This can be achieved by combining the information in Figure 37 with the total delivered integrated luminosity of a particular CCE scan. This global behaviour is shown in Fig. 39 for all sensors using data from the CCE scans

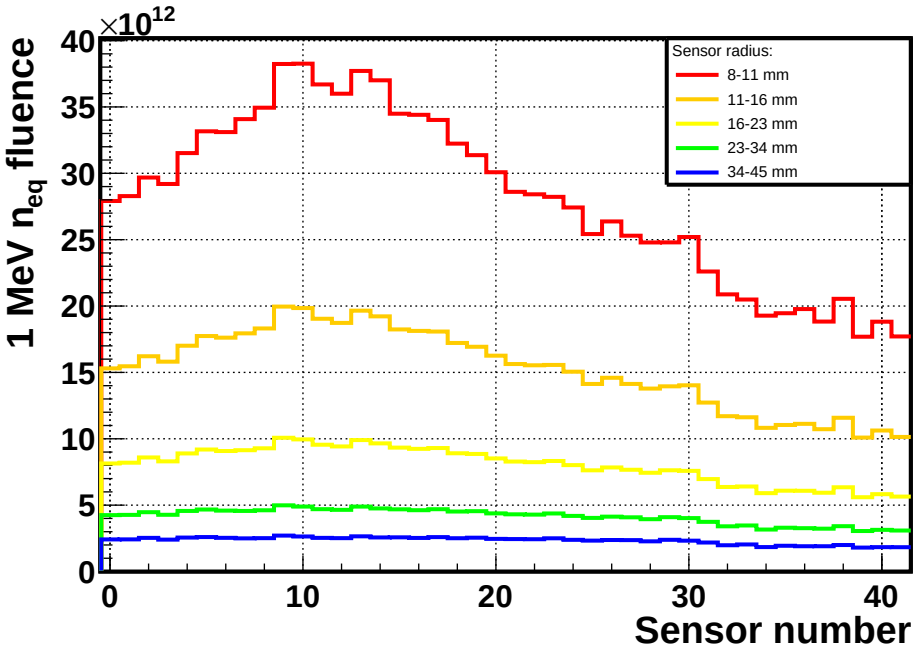


Figure 37: Predicted fluence per fb^{-1} at 7 TeV as a function of sensor number and the radial regions used in this chapter.

taken in 2012 and 2013. Due to the strong variation of fluence with radial and z-position, a range of fluences spanning two orders of magnitude is obtained. From Fig. 39 the minimum EDV for any sensor is found to be about 18 V. This differs substantially from the behaviour in Fig. 28 where the depletion voltage decreases to 0 V, due to the sensor electronics failing to collect the deposited charge within the finite sampling window at low bias voltages. A common type-inversion point is observed at $10 - 20 \times 10^{12}\,\text{n}_{\text{eq}}\,\text{cm}^{-2}$, after which the EDV is found to be independent of its initial value, and to exhibit a linear increase with further fluence. This behaviour is in agreement with the Hamburg model (see Ref. [83] for more information) and test beam studies at fluences beyond those currently experienced by the VELO sensors [85], and is useful for predicting the future operating conditions of the VELO, as discussed in Section 4.8.

The p-type sensors (highlighted in the dashed box in Fig. 39) exhibit a similar linear increase in EDV with fluence to that of the n-type sensors, but require much larger voltages to be fully depleted at a given fluence. This suggests that if

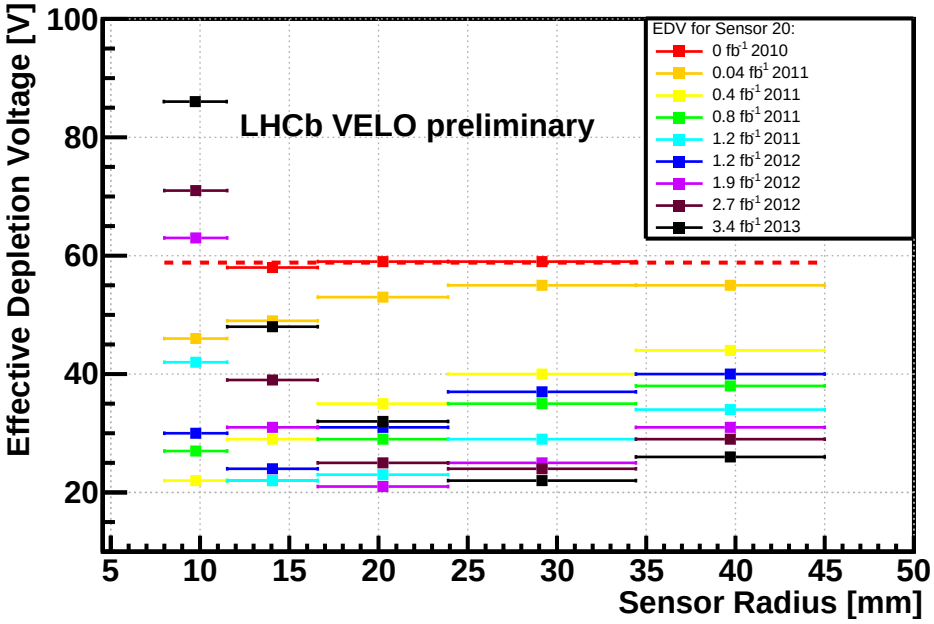


Figure 38: EDV in bins of sensor radius for an n-type R sensor for the nine CCE scans taken before LS1. The horizontal line indicates the initial depletion voltage determined from the CV method.

the currently observed trends continue, the overall lifetime of the p-type sensors will be reduced compared to that of the n-type sensors.

4.6 CLUSTER FINDING EFFICIENCY

In a similar manner to the method used to determine the EDV, the cluster finding efficiency (CFE) can be determined by measuring the fraction of times a cluster is obtained at the track intercept on the test sensor. As mentioned in Section 3.2.2.1, a cluster is defined as one or several adjacent strips with charge above a particular clustering threshold. In principle it is possible to change this threshold, but it has been kept fixed throughout the 2010 – 2012 data taking. As the LHCb track reconstruction relies upon clusters in the VELO to determine the

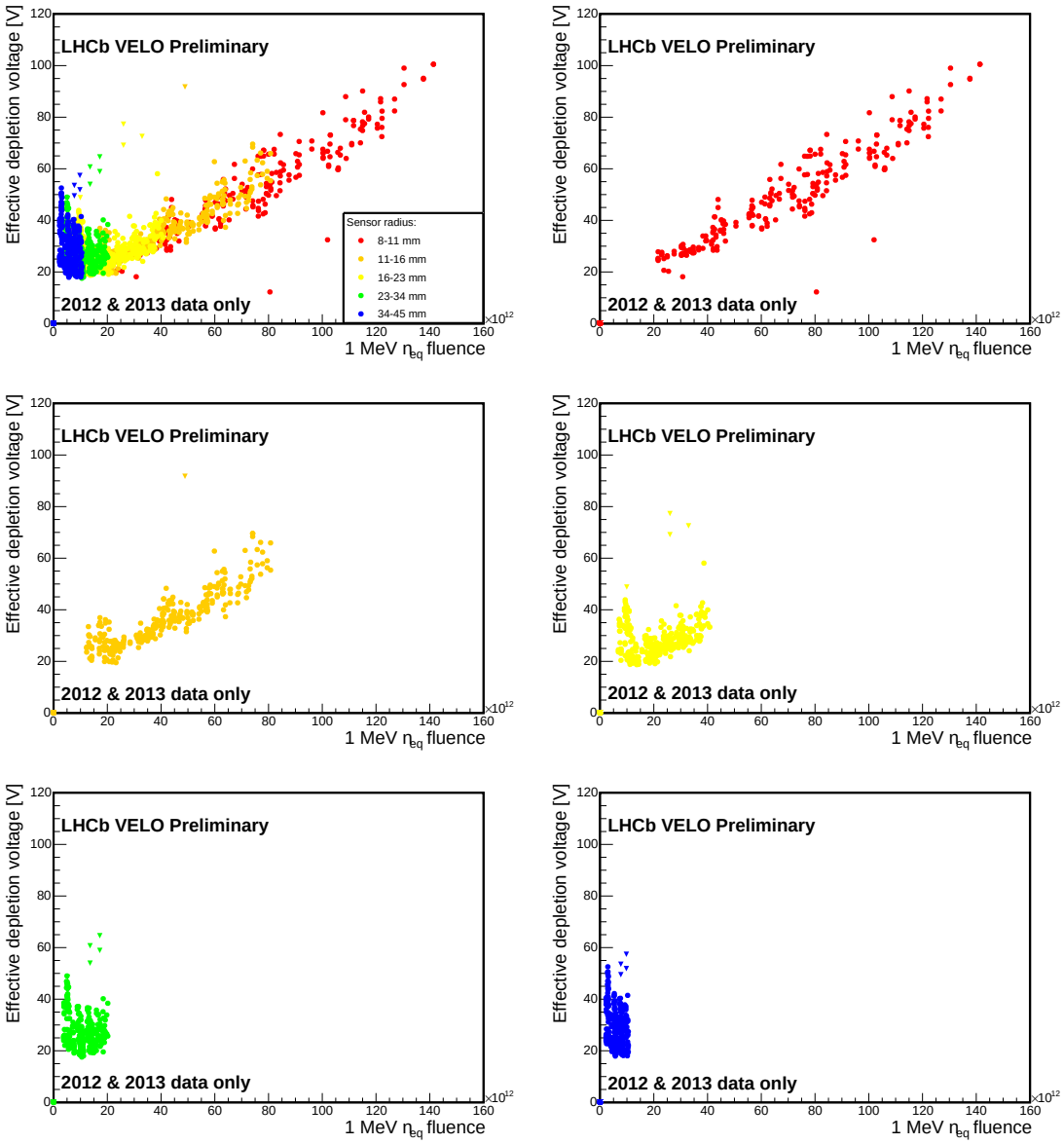


Figure 39: Change in EDV with fluence for all sensors before LS1 in all radial regions (top left) and in order of increasing radius (top right to bottom right). For clarity data from the CCE scans prior to the annealing period between 2011 and 2012 are omitted.

trajectory of tracks, it is essential that the CFE remain as high as possible. Before irradiation the mean CFE of the VELO sensors was measured to be 99.97% [54].

The change in CFE with fluence for Run 1 is shown in Figs. 40 and 41. A general decrease in CFE with fluence is observed in all ϕ sensors as a result of bulk radiation damage. As the inner radial regions receive the most fluence the decrease is greatest in these regions. For the R sensors a large and unexpected decrease is found in the outer radial regions, with some evidence of a recovery with further fluence in the 16 – 23 mm radial region. The potential cause of this behaviour is described in the following. Currently no degradation in the VELO tracking efficiencies as a result of these changes have been observed within uncertainties.

4.6.1 The Second Metal Layer Effect

The CFE can also be determined using regular physics data if measurements are restricted to the nominal bias voltage of 150 V. This allows for large statistics samples to be obtained, such that the CFE as a function of the x and y-position within a sensor can be determined. Figure 42 compares the change in CFE across a downstream R sensor (top) at the end of Run 1 to the distribution of the routing lines (bottom) described in Section 4.3. The absence of routing lines, indicated by the lighter regions in Fig. 42 (bottom), is found to correspond to areas of CFE close to 100% in Fig. 42 (top). Figure 43 shows the variation of the CFE with distance to the nearest routing line (RL distance) and distance to the nearest strip. The minimum CFE is found in the regions closest to routing lines and furthest away from strips, a characteristic which is also observed in the silicon detectors of other LHC experiments [86, 87].

The loss in CFE can be attributed to the induction of charge on the routing lines, originating from radiation induced modification of the field line structure within the sensor. In the following, this mechanism is referred to as the 'second metal layer effect'. The strong radial dependence observed is a result of the

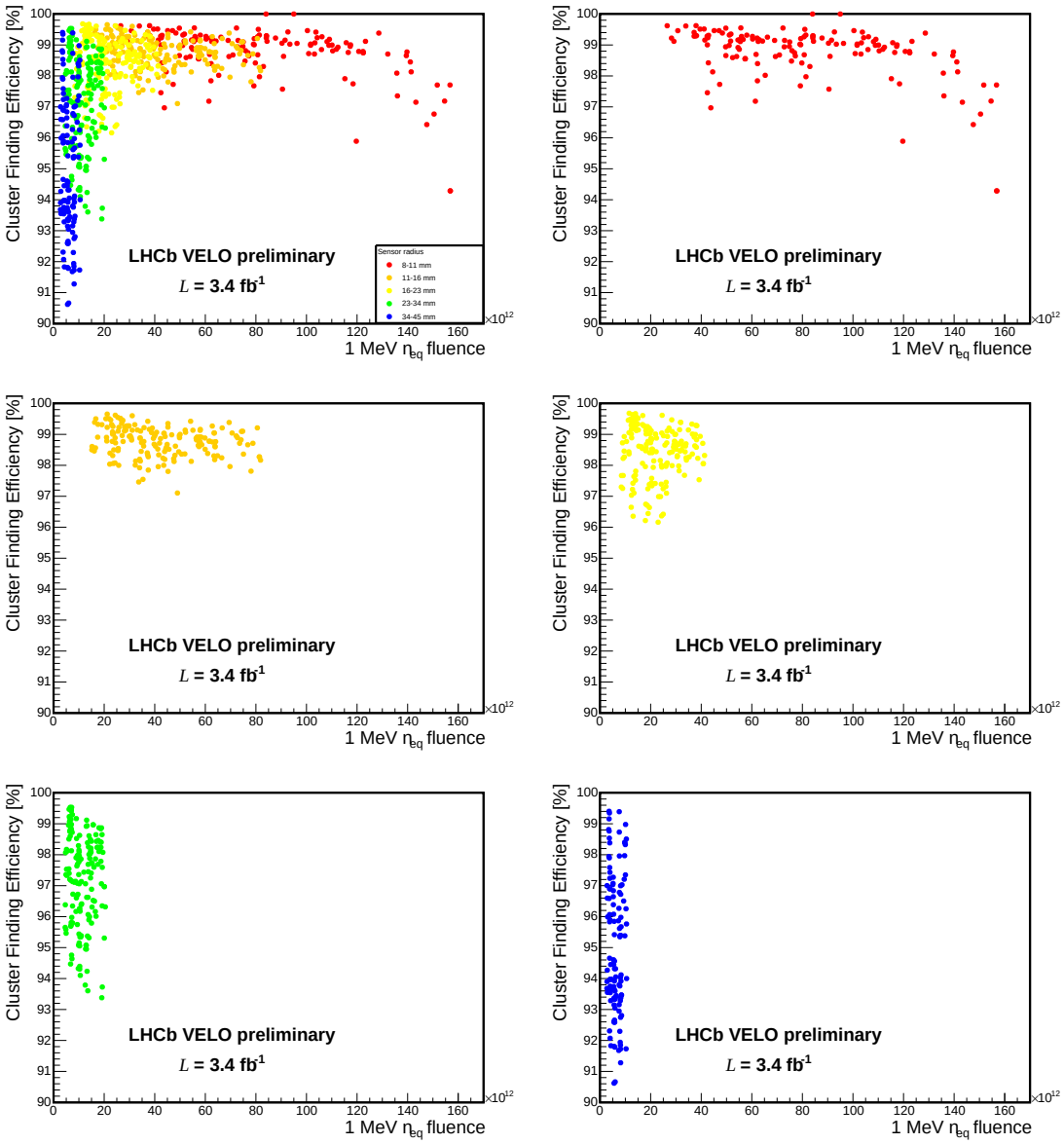


Figure 40: Change in CFE with fluence for all R sensors before LS1 in all radial regions (top left) and in order of increasing radius (top right to bottom right). All data were taken at a bias voltage of 150 V. For clarity data from the CCE scans prior to the annealing period between 2011 and 2012 are omitted.

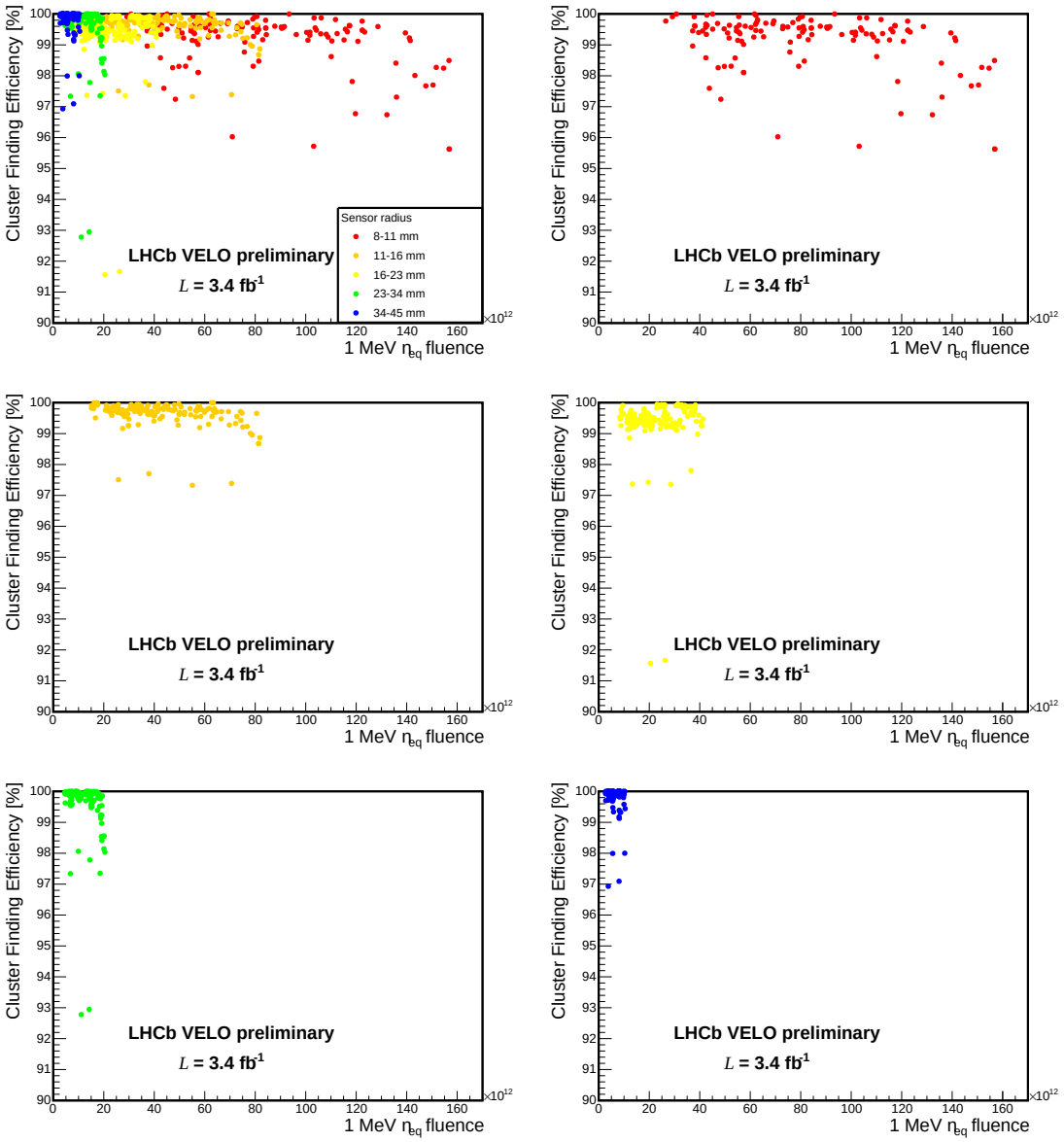


Figure 41: Change in CFE with fluence for all ϕ sensors before LS1 in all radial regions (top left) and in order of increasing radius (top right to bottom right). All data were taken at a bias voltage of 150 V. For clarity data from the CCE scans prior to the annealing period between 2011 and 2012 are omitted.

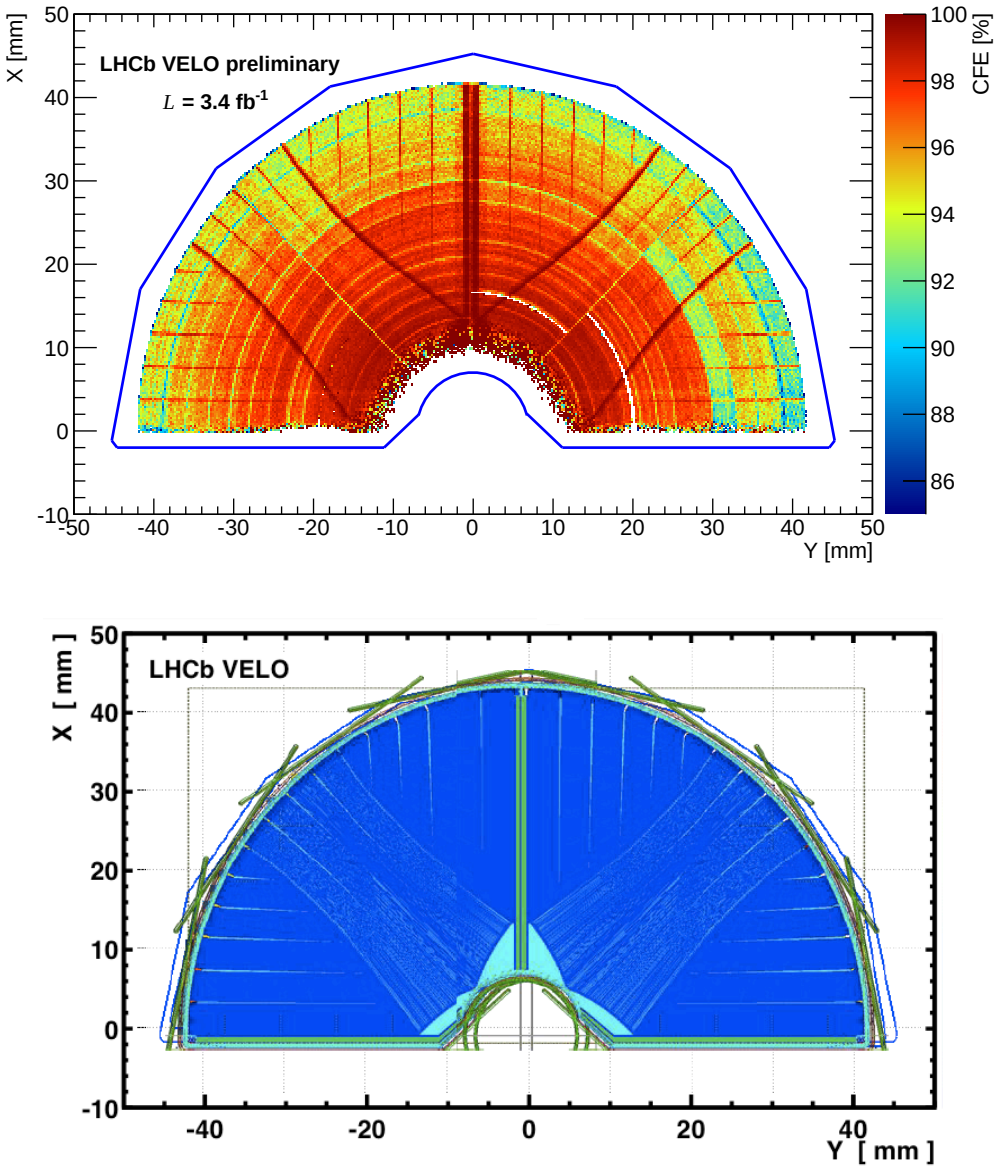


Figure 42: Top: The CFE in different regions of a downstream R sensor at the end of Run 1. The data were taken at a bias voltage of 150 V. Bottom: Distribution of routing lines across an R sensor. Lighter regions represent the absence of routing lines.

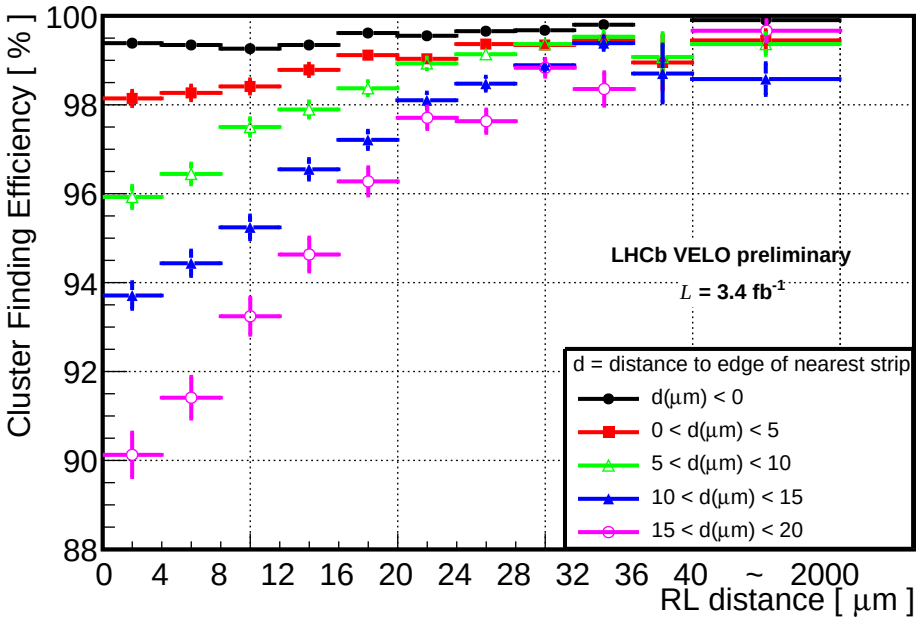


Figure 43: Change in CFE with distance from the nearest routing line (RL distance) and the nearest strip for a downstream R sensor at the end of Run 1. Strip distances of less than zero correspond to track intercepts underneath a strip. All data were taken at a bias voltage of 150 V.

structure of the routing lines and the strips shown in Fig. 32. At small radius (left) the width of the strips and the fractional area covered by the routing lines are small. At large radius (right) both of these quantities increase, and the loss in charge resulting from charge induction to the routing lines becomes significant. The reduction in CFE due to the second metal layer effect is not observed in ϕ sensors, as shown in Fig. 44, as the geometry of the sensors allows for the routing lines to be placed directly over the strips, minimising charge induction. However, there is a small drop in CFE corresponding to the transition between the inner and outer radial regions.

The decrease in CFE resulting from the second metal layer effect is also found to be dependent on the bias voltage of the sensors, as shown in Fig. 45 for a downstream R sensor. In the outer parts of the sensor the reduction in CFE is largest at high bias voltages, as a greater amount of charge is lost to the routing

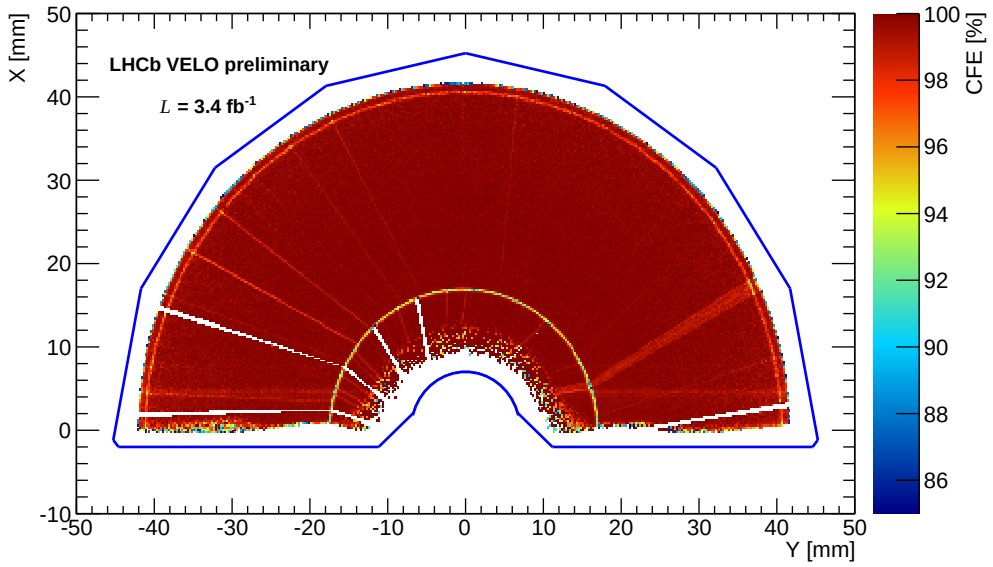


Figure 44: The CFE in different regions of a downstream ϕ sensor at the end of Run 1. The data were taken at a bias voltage of 150 V.

lines at large radius. However, in the inner parts of the sensor the reduction in CFE is largest at low bias voltages. This can be attributed to a combination of the reduced effect of the routing lines at small radius, and the fact that these parts of the sensor have undergone type-inversion. As the depletion region of the silicon grows in the opposite direction after type-inversion, the magnitude of the second metal layer effect may be reduced. This idea is supported by the increase of the CFE in the 16 – 23 mm radial region shown in Fig. 40 after approximately $20 \times 10^{12}\,n_{eq}\,cm^{-2}$, which corresponds to the average type-inversion fluence of VELO sensors, as shown in Fig. 39. Therefore after type-inversion the CFE decrease is instead dominated by charge trapping and can be recovered by increasing the bias voltage.

As a result of these two competing effects, the optimum bias voltage for operating a sensor on the basis of CFE is a compromise between the performance of the inner and outer radial regions. As the magnitude of the CFE decrease is currently largest in the outer regions, a reduction of the bias voltages is favoured

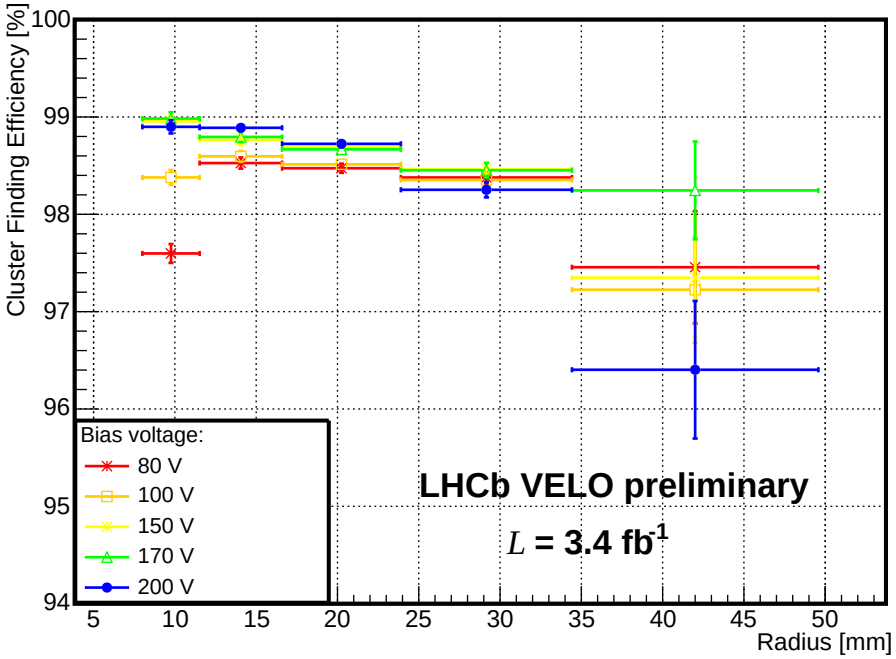


Figure 45: Dependence of the CFE on the sensor bias voltage for a downstream R sensor at the end of Run 1.

as these regions have yet to type-invert in any of the sensors (see Fig. 39). However, in the limit of full type-inversion of all of the sensors, the reduction in CFE from the second metal layer effect should be negligible, and increasing the bias voltage is the best solution for reducing the effect of bulk radiation damage to the CFE.

More details on the CFE and the effect of the second metal layer can be found in Refs. [83, 82].

4.7 THE SIGNAL TO NOISE RATIO

As described in Chapter 3, the signal to noise ratio, S/N, is an important operational parameter of many detectors. In the case of the VELO, the S/N can be

calculated by comparing the MPV of the cluster ADC distribution (the signal) to the RMS of the pedestal subtracted ADC distribution (the noise).

Figures 46 and 47 show the MPV as a function of fluence for all sensors. The MPV decreases approximately linearly with fluence, for the ϕ sensors, as a result of bulk radiation damage. For the R sensors, the MPV decreases with fluence, but with a radius dependent rate, such that the lowest MPVs in the outer radial regions of the sensors are comparable to those in the inner radial regions, despite almost an order of magnitude difference in the fluence. The difference in behaviour between the R and ϕ sensors can be attributed to the second metal layer effect described in the previous section. In the case of the S/N, the charge induction on the routing lines causes a decrease in the amount of collected charge, and hence the signal is reduced. Furthermore, regardless of the radial region, at a given fluence the R sensors typically have lower MPV values than the ϕ sensors.

Figures 48 and 49 show the noise as a function of fluence for all sensors. The level of noise in the sensors appears to be approximately constant with fluence. The R sensors have slightly higher noise on average, due to the difference in strip and routing line configurations, and this, combined with the reduced signal, means that the average S/N is lower in R sensors.

By combining the information in Figs. 46–49 the change in S/N with fluence can be determined, as shown in Figs. 50 and 51. The S/N is observed to start at 4 – 5 counts higher for ϕ sensors than for R sensors. As the sensor noise is approximately independent of fluence and radius, the S/N changes with fluence similarly to the MPV. The S/N decreases approximately linearly with fluence for the ϕ sensors, whilst for the R sensors the S/N decreases with fluence with a radius dependent rate, due to the second metal layer effect. Therefore although the inner radial regions experience larger fluences at a given integrated luminosity, the S/N remains comparable across the whole sensor.

In a similar manner to the CFE, the S/N is observed to be dependent on the bias voltage. Figure 52 shows the S/N versus voltage for a downstream R sensor.

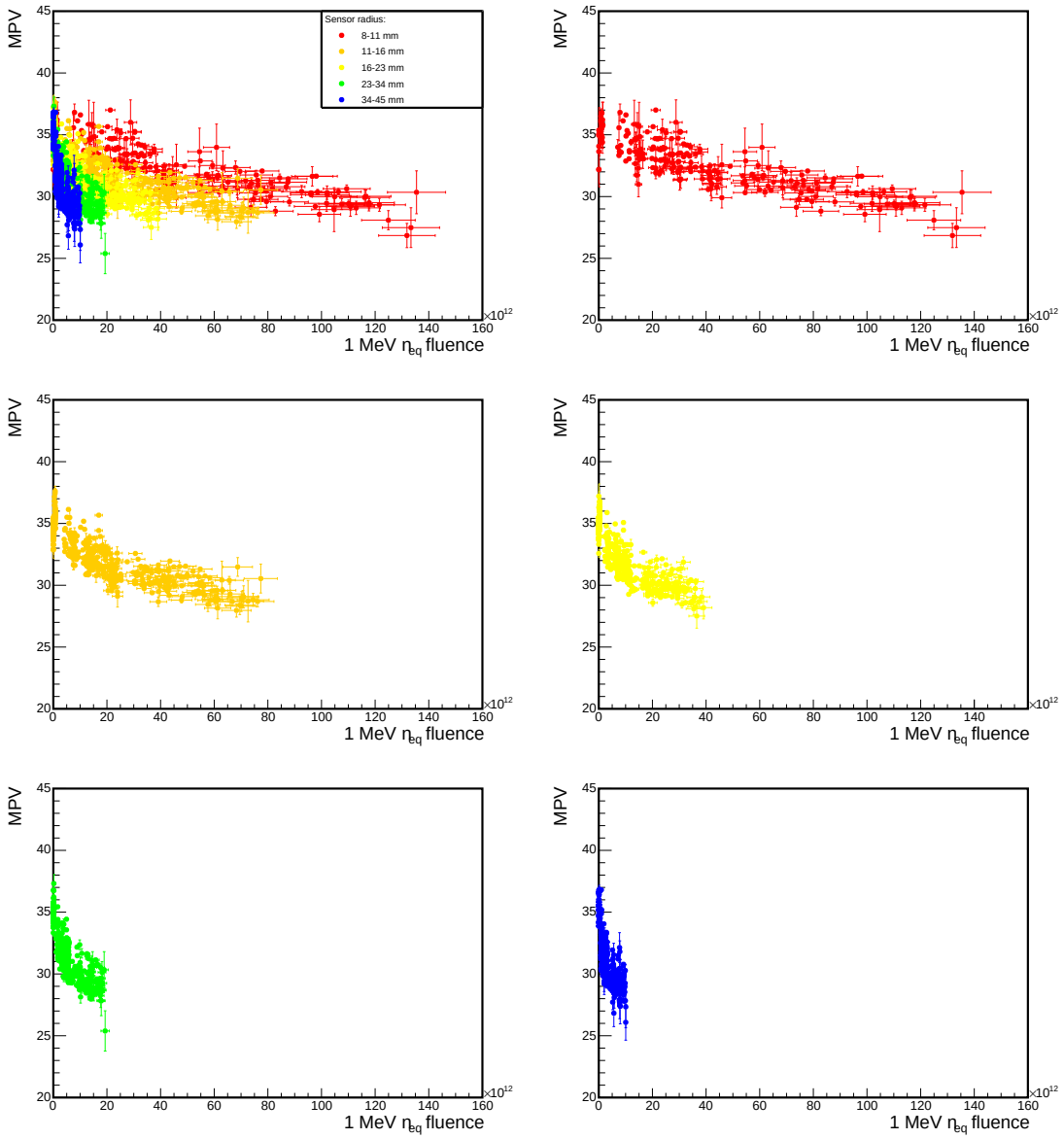


Figure 46: Change in MPV with fluence for all R sensors before LS1 in all radial regions (top left) and in order of increasing radius (top right to bottom right). All data were taken at a bias voltage of 150 V.

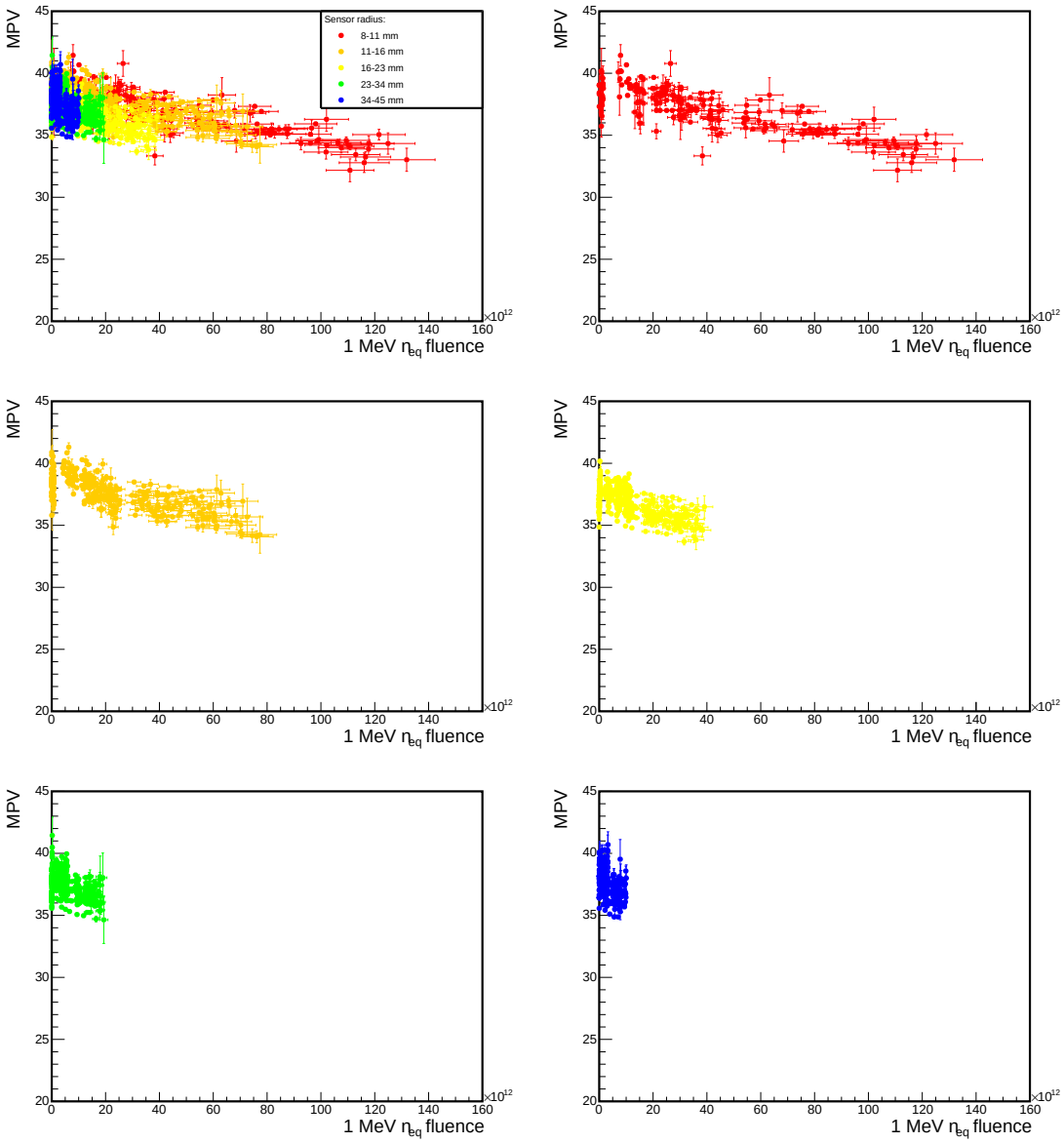


Figure 47: Change in MPV with fluence for all ϕ sensors before LS1 in all radial regions (top left) and in order of increasing radius (top right to bottom right). All data were taken at a bias voltage of 150 V.

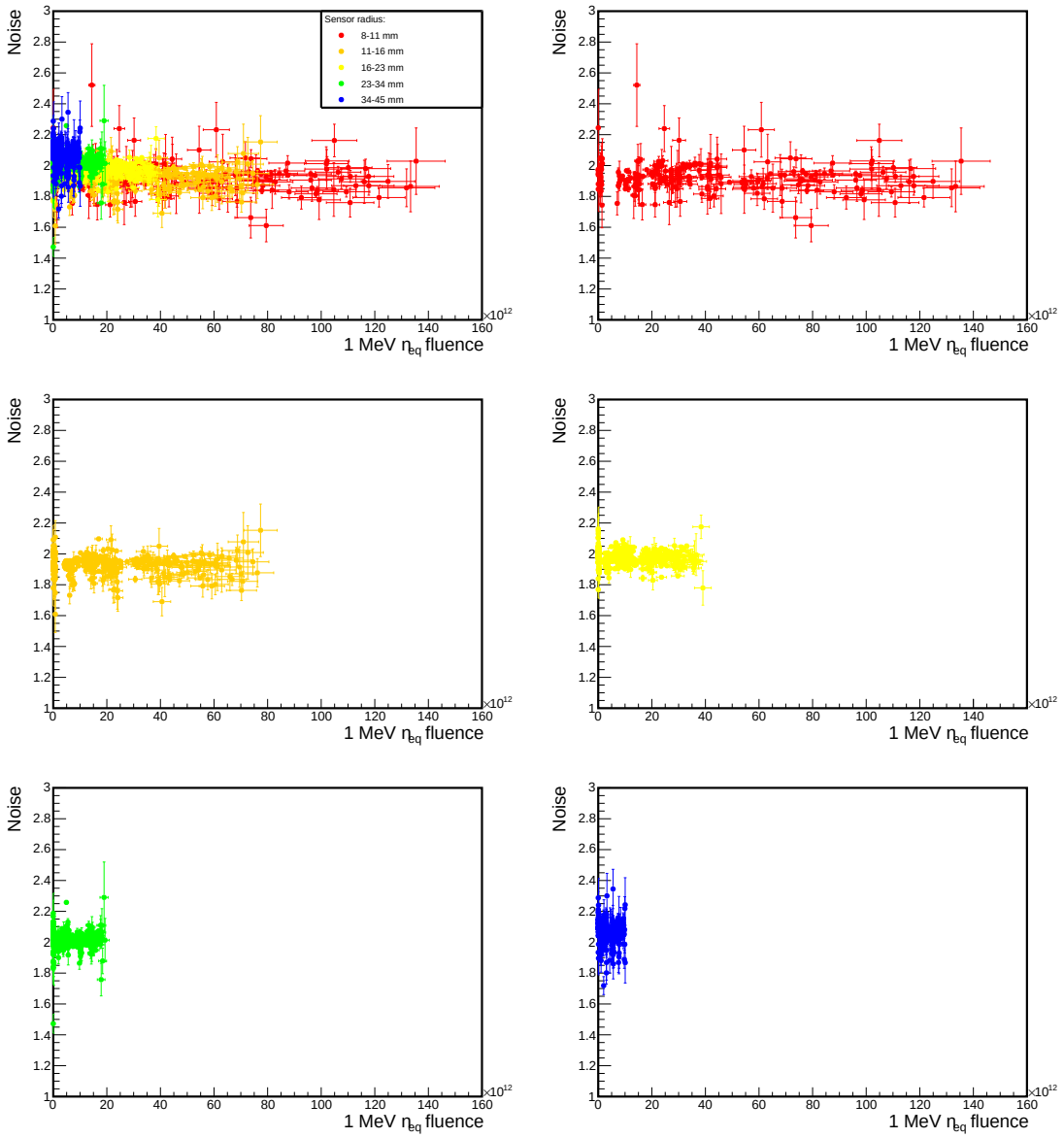


Figure 48: Change in noise with fluence for all R sensors before LS1 in all radial regions (top left) and in order of increasing radius (top right to bottom right). All data were taken at a bias voltage of 150 V.

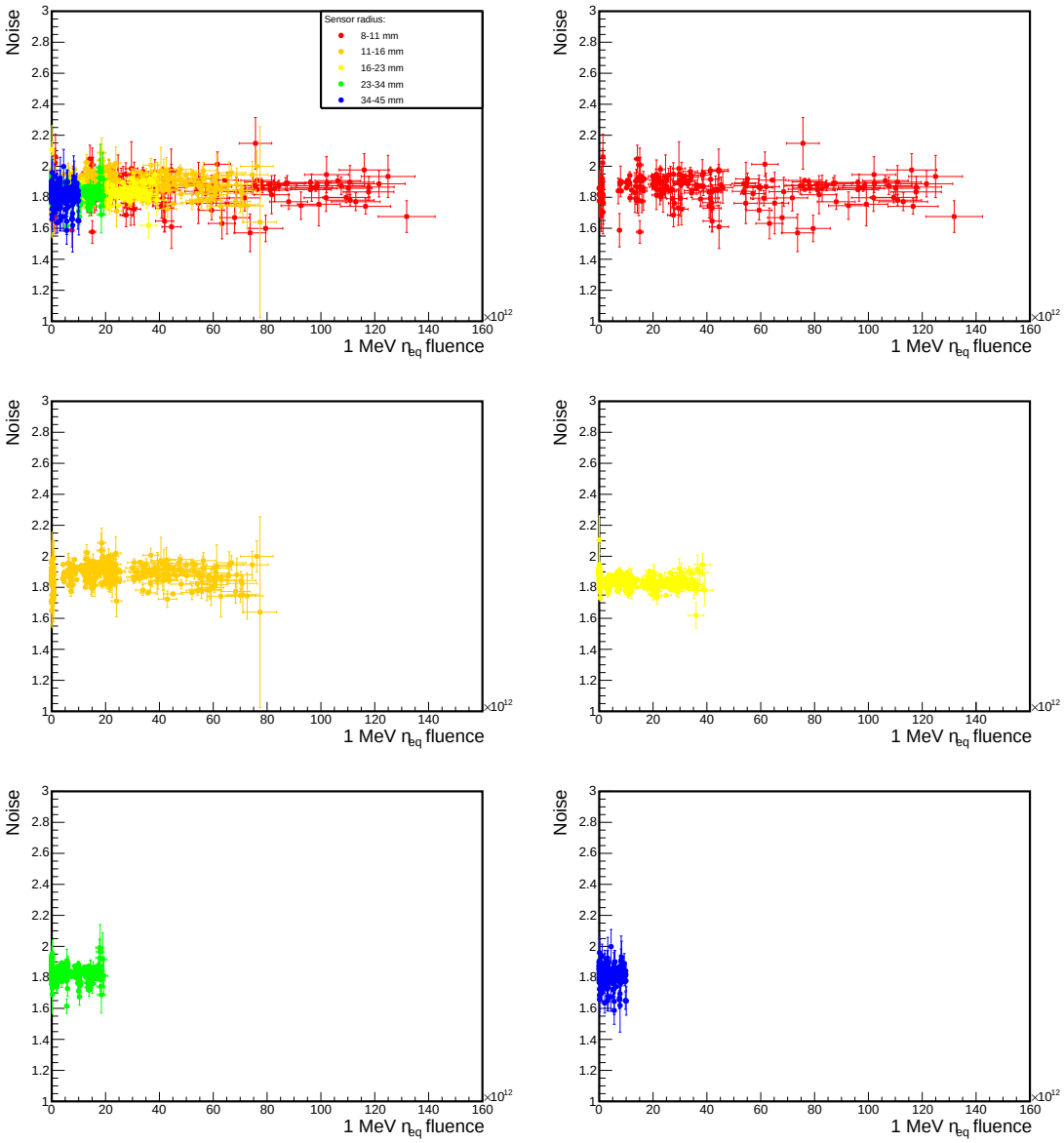


Figure 49: Change in noise with fluence for all ϕ sensors before LS1 in all radial regions (top left) and in order of increasing radius (top right to bottom right). All data were taken at a bias voltage of 150 V.

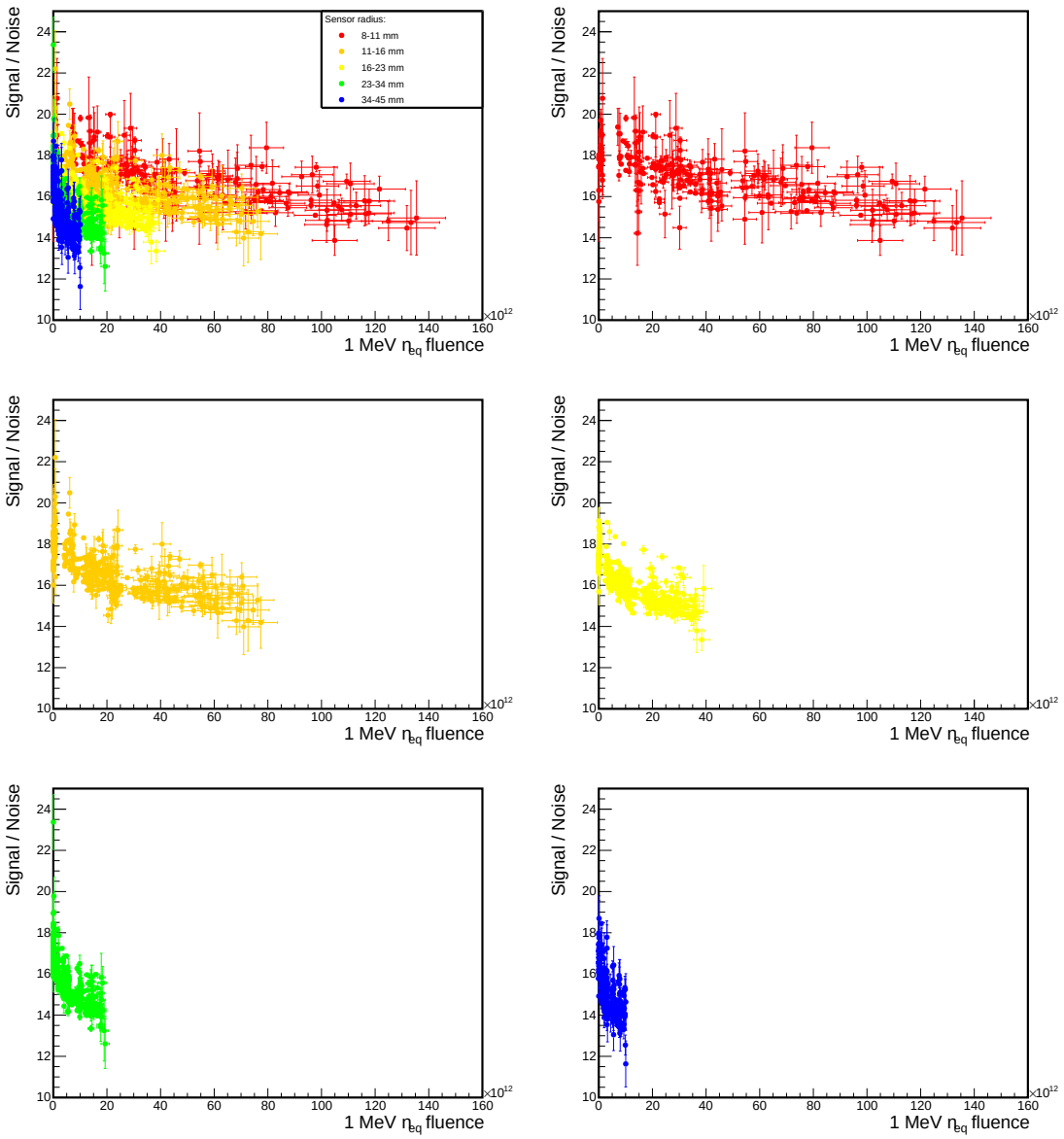


Figure 50: Change in S/N with fluence for all R sensors before LS1 in all radial regions (top left) and in order of increasing radius (top right to bottom right). All data were taken at a bias voltage of 150 V.

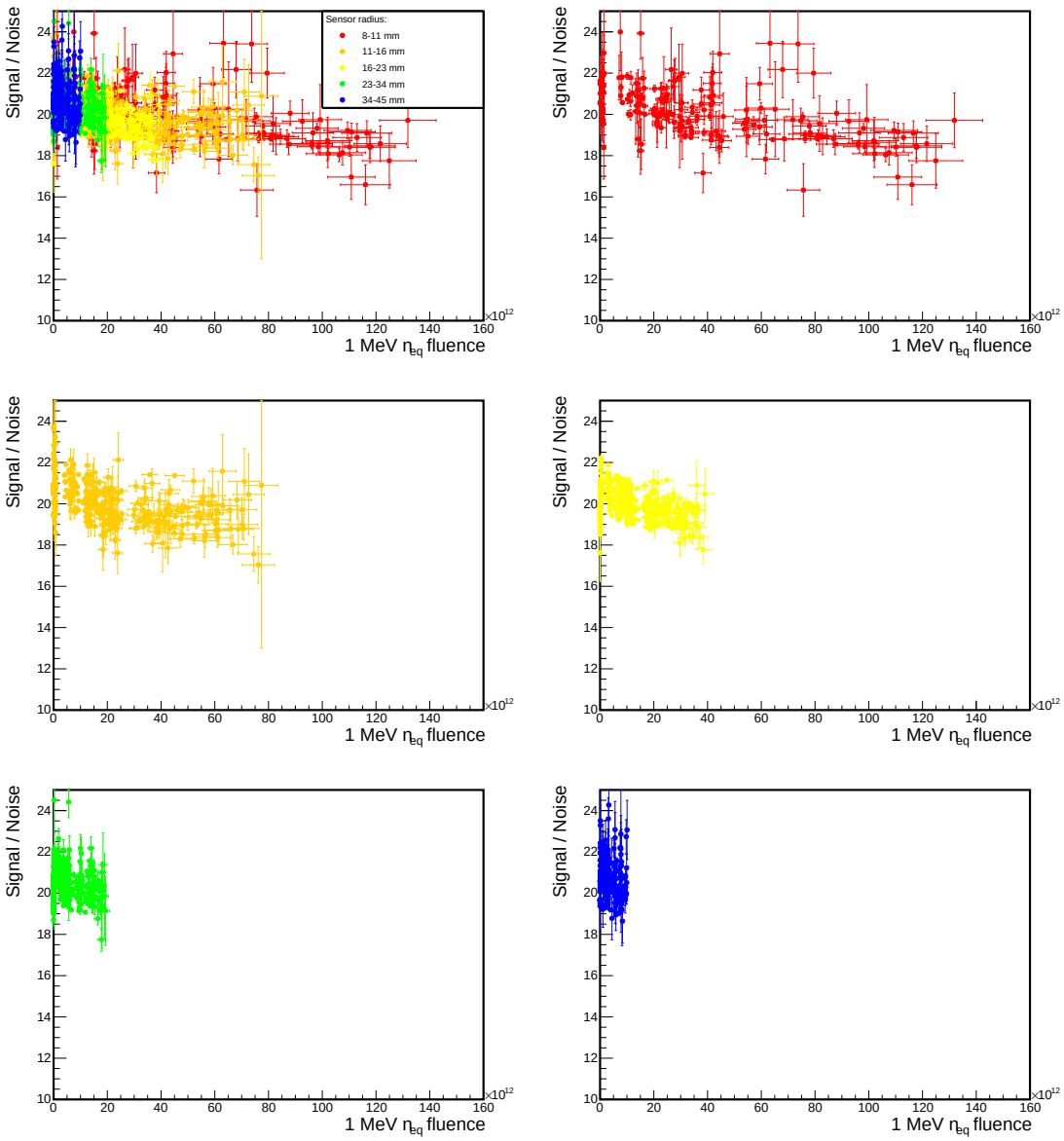


Figure 51: Change in S/N with fluence for all ϕ sensors before LS1 in all radial regions (top left) and in order of increasing radius (top right to bottom right). All data were taken at a bias voltage of 150 V.

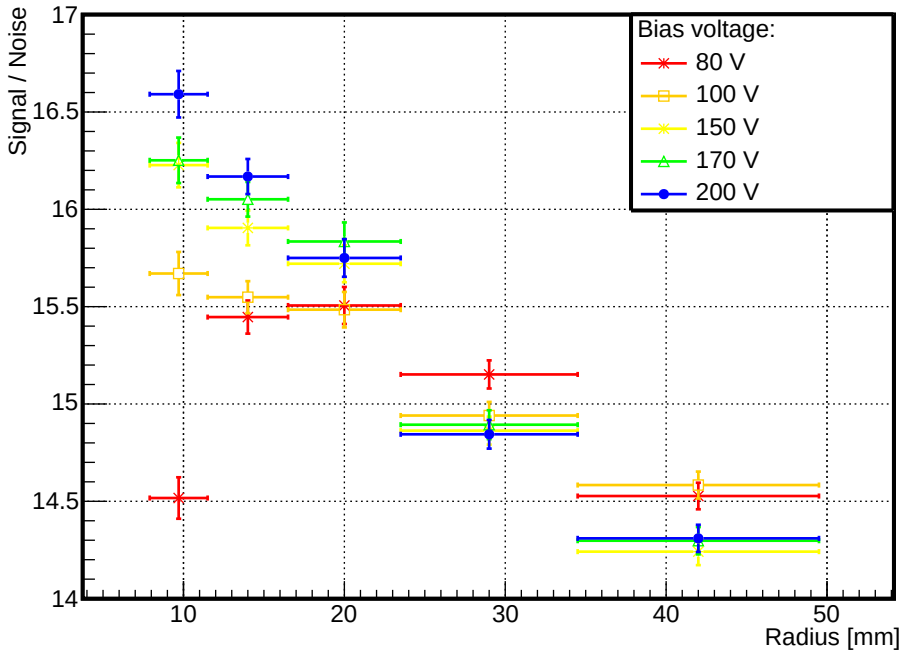


Figure 52: Dependence of the S/N on the sensor bias voltage for a downstream R sensor at the end of Run 1.

In the inner radial regions the amount of collected charge increases at higher bias voltages. This is consistent with the reduction in the S/N being caused by charge trapping. In the outer radial regions the amount of collected charge decreases with bias voltage due to the second metal layer effect. However, in the limit of full type-inversion of all of the sensors, increasing the bias voltage is also the best option for reducing the decrease in the S/N, if the magnitude of the second metal layer effect is reduced as expected.

The implications of these observations for the long term operation of the VELO are discussed in the following section.

4.8 FUTURE PERFORMANCE

Until the start of LS1 the VELO sensors were operated at a common bias voltage of 150 V, with no observed decrease in tracking efficiency as a result of radiation damage. However, by the end of Run 2 a total of 10 fb^{-1} of integrated luminosity may have been delivered to the detector. Therefore with almost twice the current fluence and significant additional radiation damage yet to occur, if the observed trends in the EDV, CFE and S/N continue, a number of problems may become apparent.

Firstly, if the bias voltage required to fully deplete the sensors continues to increase, it will shortly exceed the nominal bias voltage of 150 V. This will result in reduced CCE and directly impact the tracking capabilities of the VELO. The VELO has a hardware-imposed maximum operating voltage of 500 V, and if the EDV exceeds this value the additional loss in CCE cannot be recovered.

Secondly, as the CFE decreases, the fraction of missing hits per track will increase and the number of tracks which are accurately reconstructed by the VELO will be reduced. This effect can be compensated by reducing the clustering thresholds (currently set at six times the average noise) but this may also increase the number of fake clusters and is not considered further here.

Finally, as the S/N decreases it will approach a value of 10, which is conventionally regarded within the silicon detector community as the minimum acceptable S/N for the operation of a detector. As a result, the number of fake clusters resulting from noise may start to become significant.

Whilst the magnitude of these effects could be substantially reduced by immediately increasing the bias voltages of the sensors to the hardware limit, this should be avoided as a large increase in bias voltages may introduce a number of unexpected changes to the detector performance. In the case of the CFE and the S/N, a minimal increase in the bias voltages is favoured to reduce the magnitude of the second metal layer effect in the outer radial regions of the sensors. It has also been shown that operating the sensors in an overdepleted state leads

to a reduction in the tracking resolution, as the number of multi-strip clusters is reduced [82]. Therefore it is beneficial to be able to predict the minimum required bias voltage at a given integrated luminosity. In addition, the expected values of the EDV, CFE and S/N at a total of $10\,\mathrm{fb}^{-1}$ are of interest for determining the expected lifetime of the VELO and whether or not a replacement is required. A replacement is available and can be used if essential, but this would require the extensive work involved in the installation and commissioning of a new detector.

4.8.1 Assumptions

The behaviour of the EDV, CFE and S/N are extrapolated to higher fluences via a number of fits to the current CCE scan data. In order to do this a number of additional assumptions are made:

- The expected values of all variables are calculated based on the fluences at the inner radial edge of the sensors. As the change in fluence is approximately a factor of two across each radial region, the expected fluences for the inner edge of the sensors can be calculated by multiplying the fluences of the 8 – 11 mm radial regions by a factor of $\sqrt{2}$.
- Large sensor-by-sensor variations make the uncertainties on the individual measurements unreliable. As the average expected value at a given fluence is the quantity of interest, the data are first binned in fluence and the uncertainties are calculated as the standard errors of the means for the different sensors within each bin.
- Only data from type-inverted sensors are considered. Therefore a minimum fluence cut of $50 \times 10^{12}\,\mathrm{n_{eq}\,cm^{-2}}$ is applied.

- A maximum of 1 fb^{-1} of luminosity is delivered between subsequent CCE scans, such that further radiation damage studies with additional data can be performed. The next CCE scan is therefore assumed to be at 4.4 fb^{-1}.
- The centre-of-mass energy is 14 TeV for all subsequent data taking. As the fluence is dependent on the centre-of-mass energy, the expected fluence per inverse femtobarn is higher than that at the previous centre-of-mass energies of 7 and 8 TeV. Figure 53 shows the MC predicted fluences in the inner regions of the sensors at 14 TeV. The maximum increase in the inner fluence with respect to the same sensor at 7 TeV (Fig. 36) is taken as a conservative estimate of the scaling and is found to be approximately 40% for the most downstream sensors.

The expected bias voltages of all sensors at the fluences corresponding to the next CCE scan are calculated. If these voltages are found to be above the current sensor bias voltages then the bias voltages are increased in increments of 50 V until they are above the expected values, to ensure all sensors are fully depleted. The expected bias voltages at 10 fb^{-1} are also calculated.

The CFE and the S/N of all sensors are then calculated and extrapolated to 10 fb^{-1}, using data collected at a bias voltage corresponding to that determined in the previous step. As only the behaviour of the inner radial regions is considered, this is a measure of the effects of bulk radiation damage and does not account for the reduction in the CFE and S/N in the outer radial regions of the sensors due to the second metal layer effect. Whilst the magnitude of the second metal layer effect is expected to decrease with further type-inversion, the changes in the CFE and S/N from this mechanism do not currently have a clear fluence dependence and should be monitored separately with further CCE scan data.

If the extrapolations suggest that for any sensors at 10 fb^{-1} the EDV exceeds 500 V, then the performance of the VELO could be substantially impacted and a replacement should be considered. If this is not the case, but instead the CFE

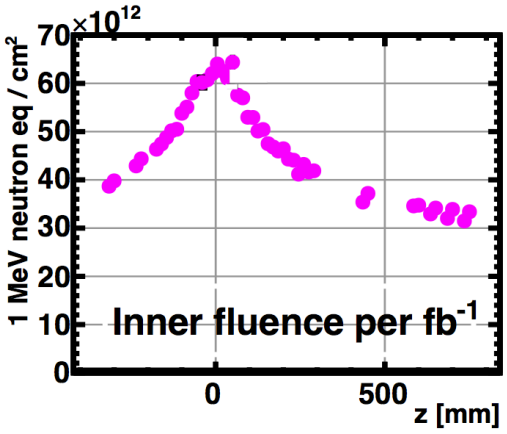


Figure 53: Variation of fluence with z-position of the sensor as determined from MC at 14 TeV. Figure from Ref. [77].

drops below 95%, or the S/N drops below 10, then raising the sensor bias voltages further above the EDV could partially compensate for these losses.

4.8.2 Results

The results of linear fits to the EDV of the inner radius regions of all sensors can be seen in Fig. 54. As the change in EDV is not directly affected by the second metal layer effect, both R and ϕ sensors exhibit a similar linear increase with fluence.

Table 10 shows the results of extrapolating the fits to the estimated fluence at the next CCE scan ($+1\,\text{fb}^{-1}$) and at $10\,\text{fb}^{-1}$ total delivered luminosity. It can be seen that in order to keep the VELO sensors fully depleted until the next CCE scan, the sensor bias voltages should be increased before the start of Run 2. Furthermore, even with the conservative assumptions used in this extrapolation, the required sensor bias voltages should not exceed the hardware limit of 500 V at $10\,\text{fb}^{-1}$.

To avoid having to account for many run periods with different bias voltages it is practical to keep the number of changes of the bias voltages to a minimum.

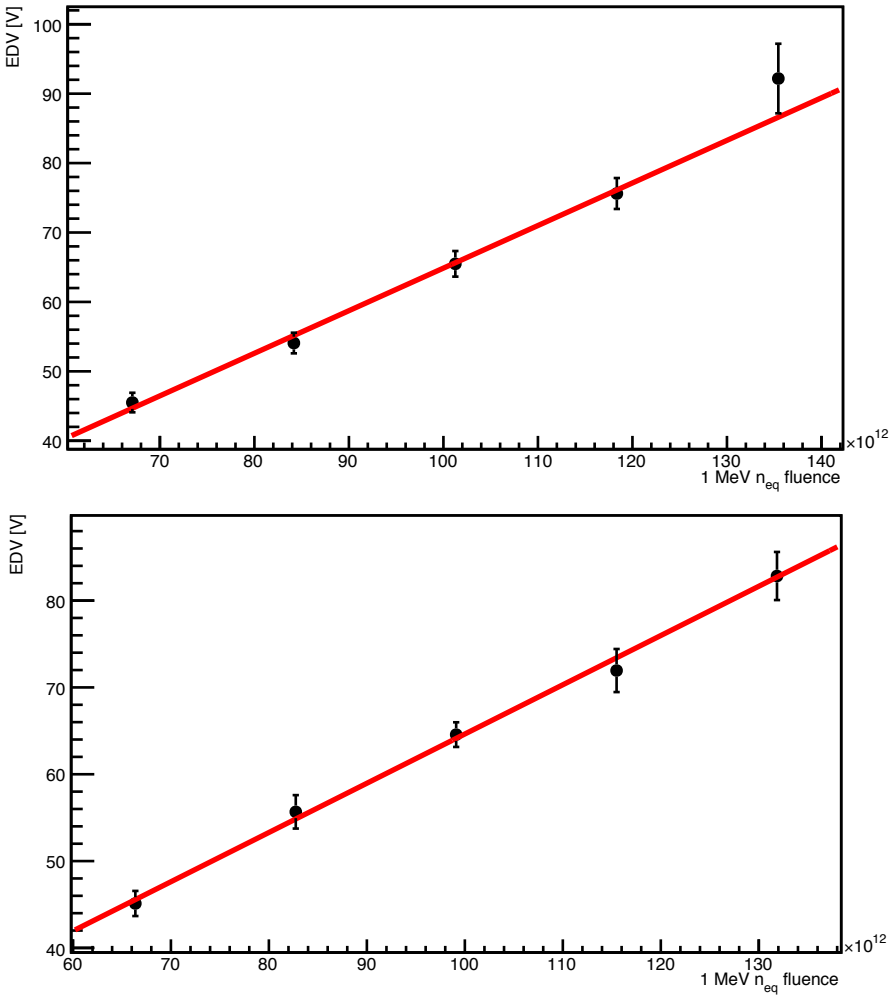


Figure 54: Linear fit to the EDV of the inner radius region of R (top) and ϕ (bottom) sensors.

Sensor bias voltages of 200 V at the start of Run 2 would ensure that the next inverse femtobarn could be collected without changing the voltages again. As the increase in bias voltages will cause a reduction in CFE and the S/N in the outer parts of the sensors, it would be possible to minimise this effect by only changing the bias voltages of certain sensors. One possible method of grouping the sensors is indicated in Fig. 55 (top), where three groups of sensors are defined according to their expected fluence per inverse femtobarn. The group of sensors which receive the highest fluence, defined here as those with 30×10^{12}

Sensor	Expected EDV at 4.4 fb^{-1} (V)	Expected EDV at 10 fb^{-1} (V)
R	172 ± 12	432 ± 30
ϕ	164 ± 11	404 ± 26

Table 10: Extrapolated fit results for the EDV of sensors at different fluences using a linear function.

$n_{eq}\,cm^{-2}$ per fb^{-1} and above, would have their bias voltages increased to 200 V. The remaining two groups, whose sensors receive $20 - 30 \times 10^{12}$ $n_{eq}\,cm^{-2}$ and $< 20 \times 10^{12}$ $n_{eq}\,cm^{-2}$ per fb^{-1}, would be kept at 150 V, as after 4.4 fb^{-1} the maximum expected EDVs are only (136 ± 10) V and (87 ± 6.8) V respectively. Various other sensor groupings are possible and can be adopted as further fluence is received.

Figure 55 (bottom) shows the predicted EDV of the R sensors after 10 fb^{-1}. The p-type sensor is excluded. The predicted EDV varies by as much as a factor of two, depending on the z-position of the sensor. For all sensors the EDV is approximately 200 V or greater.

The results of the fit to the CFE of the inner radius region of all sensors can be seen in Fig. 56. Data at 200 V are used since the results of the EDV fits strongly suggest that the sensor operating voltages will be increased before further luminosity is collected. A similar behaviour is observed for both R and ϕ sensors, as the difference in CFE due to the second metal layer effect is largely confined to the outer radial regions of the sensors.

Table 11 shows the results of extrapolating the fits to higher fluences. For both the R and ϕ sensors the CFE should remain above 95% in the inner radial region until 10 fb^{-1}.

The results of fits to the S/N of the inner radius region of all sensors can be seen in Fig. 57. Again data at 200 V are used as an increase in sensor bias voltages is expected before further luminosity is collected. Both a linear and

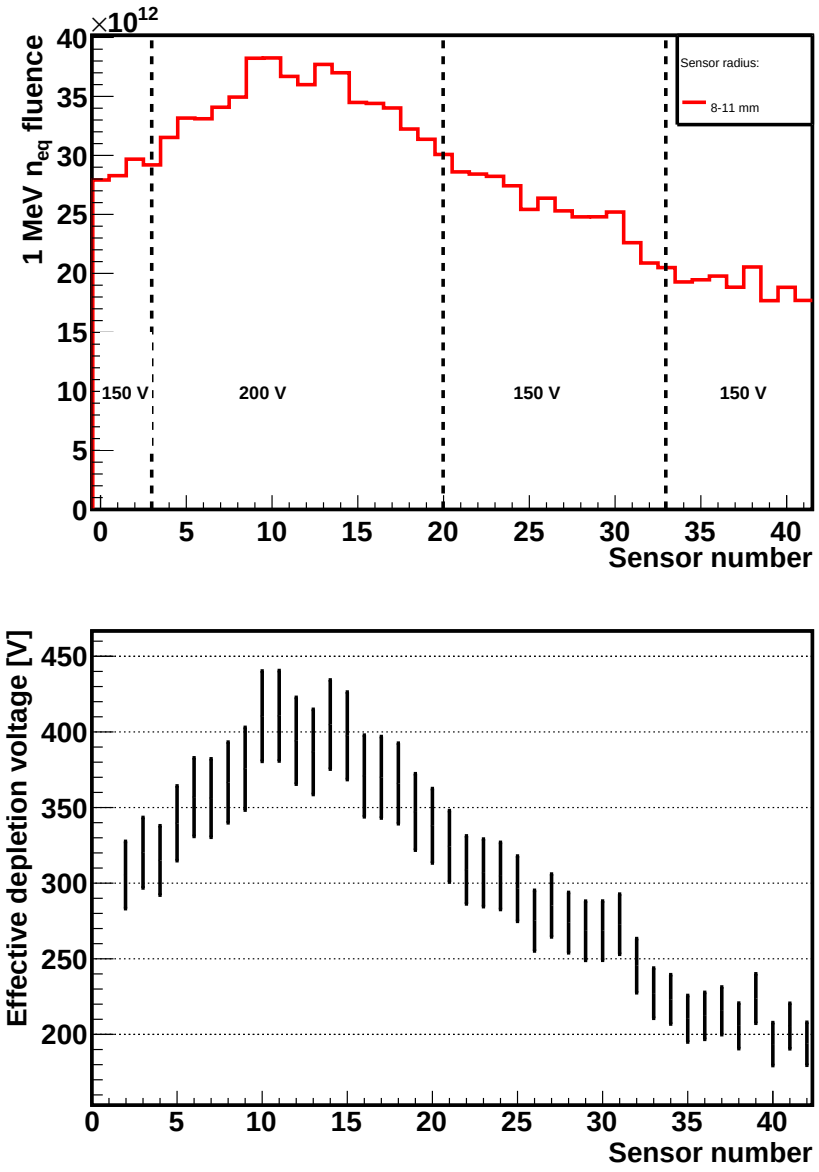


Figure 55: Top: The fluence per fb^{-1} as a function of sensor number at 7 TeV, showing four possible groupings of sensors operated at the same voltage for the start of Run 2. Bottom: Expected EDV for all n-type R sensors at $10\,\mathrm{fb}^{-1}$.

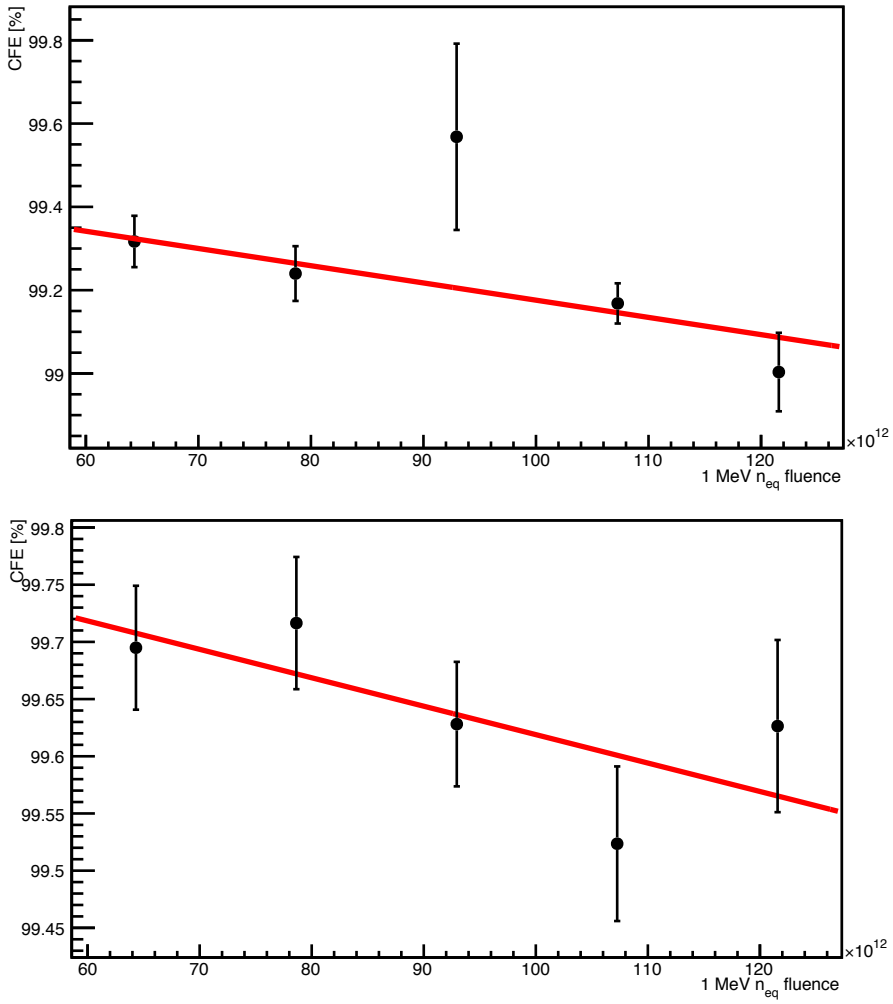


Figure 56: Linear fit to the CFE of the inner radius region of R (top) and ϕ (bottom) sensors.

third-order polynomial fit are performed. The linear fit (red) is motivated by the apparent trend in the data and the linear increase in EDV observed in Fig. 54. However, fits to laboratory data from proton, neutron and pion irradiations suggest that a third-order polynomial may give the best description of the data at higher fluences, as shown in Fig. 58. As the fluences delivered to the VELO may approach these values by the end of Run 2 this function is also fitted (blue in Fig. 57) and it is found that the fitted parameters vary by less than 10% from the values determined in the laboratory data fit. Other parameterisations of the

Sensor	Expected CFE at 4.4 fb^{-1} (%)	Expected CFE at 10 fb^{-1} (%)
R	98.5 ± 0.4	96.7 ± 1.1
ϕ	99.2 ± 0.4	98.1 ± 1.0

Table 11: Extrapolated fit results for the CFE at 200 V of sensors at different fluences using a linear function.

Sensor	Expected S/N at 4.4 fb^{-1}		Expected S/N at 10 fb^{-1}	
	1st-order	3rd-order	1st-order	3rd-order
R	13.2 ± 1.5	13.6 ± 0.8	6.3 ± 3.5	9.5 ± 2.2
ϕ	15.8 ± 1.5	16.2 ± 1.3	7.2 ± 3.5	11.1 ± 3.3

Table 12: Extrapolated fit results for the S/N at 200 V of sensors at different fluences, using both a first-order and third-order polynomial function.

change in collected charge with fluence exist, such as that in Ref. [88], but are typically only applicable at fluences greater than those relevant to the VELO.

From Fig.57 it can be seen that the agreement between the two functions is excellent at the fluences currently received by the VELO, and the functions describe the data for both R and ϕ sensors. Table 12 shows the results of extrapolating the fits to higher fluences for both functions and agreement is maintained at 4.4 fb^{-1}. The ϕ sensors are approximately 2.5 ADC counts higher than the R sensors for both fits. At 10 fb^{-1} there is an approximately 3 – 4 ADC count discrepancy between the central values of the two fits for both R and ϕ sensors, although the results are consistent within their uncertainties.

Establishing which of these functions gives the best description of the data is important, as the severity of the S/N decrease may inform the choice of sensor bias voltages for future data taking. Based on data taken at 200 V, if the S/N

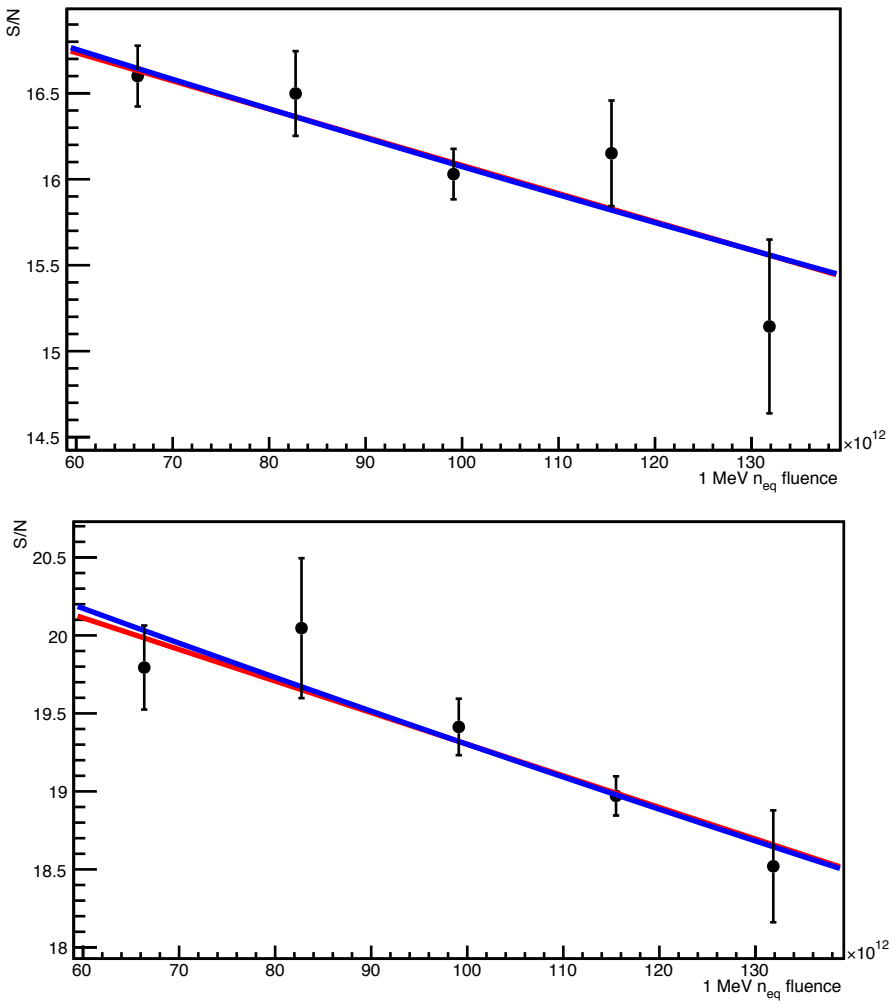


Figure 57: First-order (red) and third-order polynomial (blue) fits to the S/N of the inner radius region of R (top) and ϕ (bottom) sensors at 200 V.

continues to decrease with a linear dependence it will drop below the minimum S/N of 10 before $10\,\mathrm{fb}^{-1}$ for both R and ϕ sensors. The exact behaviour of the S/N with fluence should become apparent with further data. Therefore careful monitoring of this quantity should continue to the next CCE scan and beyond.

In practice, for both the CFE and the S/N, the performance at $10\,\mathrm{fb}^{-1}$ may be substantially better than the results in Tables 11 and 12 indicate, as the sensors will be operated at even higher bias voltages. Due to the timing restrictions on the CCE scan data taking, data at bias voltages above 200 V are not currently

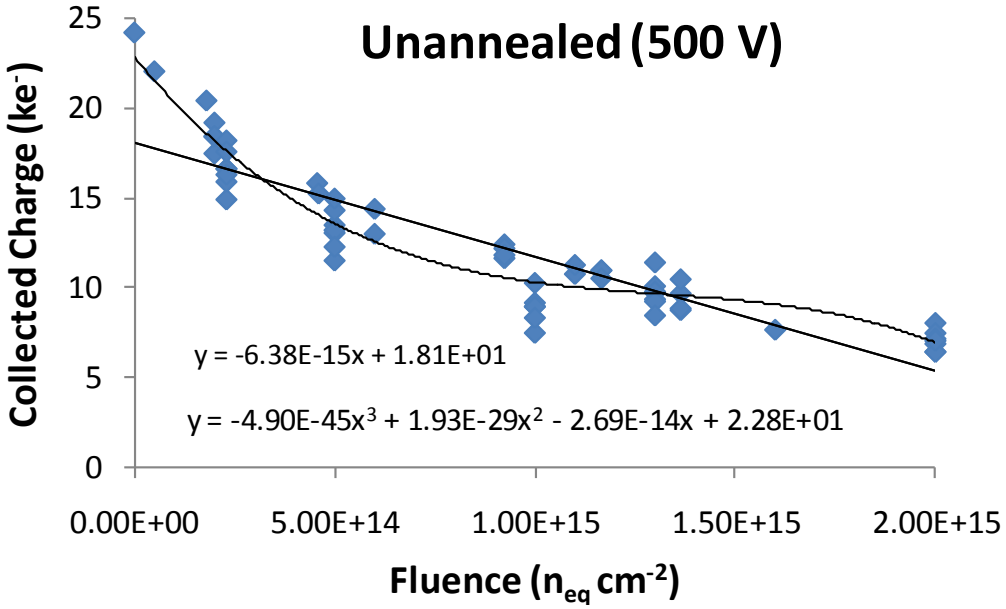


Figure 58: First-order and third-order polynomial fits to the collected charge versus fluence for laboratory data. Figure from A. Affolder, private communication.

available. However, data will be taken at these voltages in future CCE scans, allowing for further extrapolations. Based on the current data, it can be assumed that the CFE should remain above 95% and the S/N should remain above 10 for all sensors for the VELO lifetime.

4.9 SUMMARY

Changes in the EDV, CFE and the S/N of the n-type VELO sensors have been studied using CCE scan data collected up until the end of Run 1.

The decrease in EDV with fluence, followed by type-inversion and an increase in EDV, is consistent with expectations. A method has been developed to determine the operational voltage of the VELO sensors at a given delivered integrated luminosity, with extrapolations to $4.4\,\text{fb}^{-1}$ suggesting that the bias voltage of the sensors (or at least the sensors which receive the most fluence) should be in-

creased to 200 V before the start of Run 2. Extrapolations to 10 fb^{-1} show that the maximum EDV required to ensure the sensors remain fully depleted is below the hardware limit.

The CFE shows a steady decrease with fluence in the inner radial regions of all sensors. The CFE of the outer radial regions of the R sensors is found to be strongly affected by a charge coupling effect from the layout of the second metal layer. Data suggest that the magnitude of the second metal layer effect may be reduced after type-inversion. As the type-inverted region extends to higher radii in the VELO sensors with additional fluence, the effect may become restricted to the outermost radial regions only. Extrapolations of the CFE of the inner radial regions of the sensors to 10 fb^{-1} suggest the CFE will remain above 95%, even if the sensor bias voltages are not increased above 200 V.

The S/N shows a decrease with fluence for both R and ϕ sensors, with the R sensors having a lower S/N for a given fluence, due to the second metal layer effect. Extrapolations to 10 fb^{-1} with both a linear and third-order polynomial fit have been shown. Further data are required before a value for the expected S/N can be determined, but the combination of a reduction in the magnitude of the second metal layer effect and the increased S/N at larger bias voltages should ensure that the S/N remains above 10.

In conclusion, the development of radiation damage induced effects in the VELO should be monitored with further CCE scan data, but are not expected to limit the lifetime of the detector before 10 fb^{-1} of luminosity is delivered.

5 SEARCHES FOR LEPTON FLAVOUR AND BARYON NUMBER VIOLATING TAU DECAYS

In this chapter results on the searches for lepton flavour and baryon number violating τ decays in 2011 LHCb data are presented, based on the studies published in Ref. [89]. The theoretical motivation for these studies is described in Chapter 2.

A review of the current experimental results is given in Section 5.1. Details of the analyses of the decays $\tau^- \to \mu^-\mu^+\mu^-$, $\tau^- \to \overline{p}\mu^+\mu^-$ and $\tau^- \to p\mu^-\mu^-$ are given in Sections 5.2 and 5.3 and the results are summarised in Section 5.4.

An update to the $\tau^- \to \mu^-\mu^+\mu^-$ analysis (referred to as the 2012 analysis, compared to the 2011 analysis described here) is being performed using the full $3.0\,\text{fb}^{-1}$ collected before LS1 and the current status is described in Appendix A. The $\tau^- \to \overline{p}\mu^+\mu^-$ and $\tau^- \to p\mu^-\mu^-$ analyses are also being updated to $3.0\,\text{fb}^{-1}$ and some discussion of possible improvements for these measurements will be given here.

A summary of the LHCb results on lepton flavour and baryon number violating τ decays is given in Section 5.5, along with an overview of future prospects for these measurements.

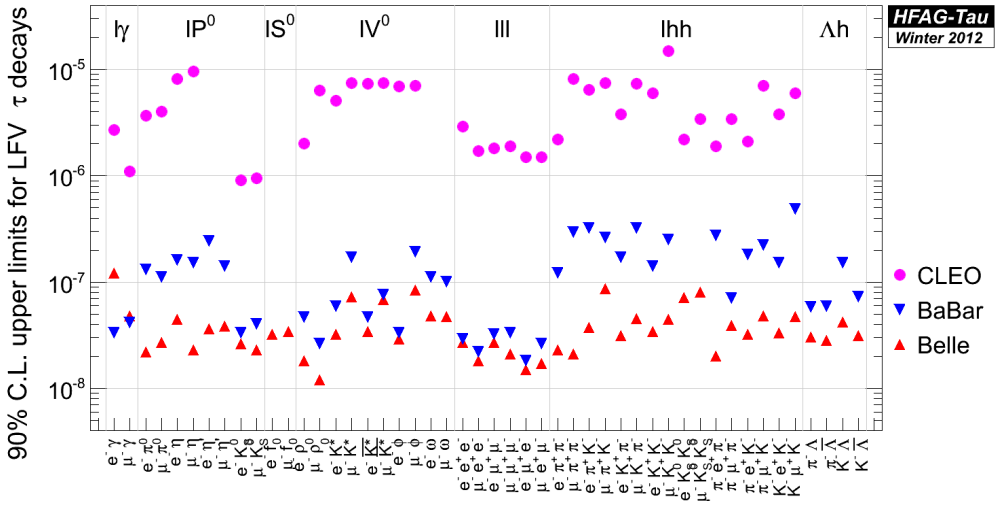


Figure 59: HFAG upper limits on τ LFV branching fractions. Figure from Ref. [90].

5.1 EXPERIMENTAL STATUS

A number of searches for LFV and BNV τ decays have been performed at various experiments. Up to the present day, no LFV or BNV τ-decay has yet been observed, so upper limits have been set, as summarised in Fig. 59.

The current best experimental upper limit on $\tau^- \to \mu^-\mu^+\mu^-$ is

$$\mathcal{B}\left(\tau^- \to \mu^-\mu^+\mu^-\right) < 2.1 \times 10^{-8},$$

at 90% confidence level (CL) from the Belle experiment [38] using $782\,\text{fb}^{-1}$ of electron-positron collision data. The BaBar experiment obtained a limit of $\mathcal{B}\left(\tau^- \to \mu^-\mu^+\mu^-\right) < 3.3 \times 10^{-8}$ [3] at 90% CL with an integrated luminosity of $480\,\text{fb}^{-1}$. Both results were obtained with almost the full data samples of the two experiments and no significant improvement can be expected. BaBar and Belle have searched for BNV τ decays with $|\Delta(B-L)| = 0$ and $|\Delta(B-L)| = 2$ using the modes $\tau^- \to \Lambda h^-$ and $\overline{\Lambda} h^-$ (with $h^- = \pi^-, K^-$), and upper limits on branching fractions of order 10^{-7} were obtained [90]. BaBar has also searched for the B meson decays $B^0 \to \Lambda_c^+ l^-$, $B^- \to \Lambda l^-$ (both having $|\Delta(B-L)| = 0$) and $B^- \to \overline{\Lambda} l^-$ ($|\Delta(B-L)| = 2$), obtaining upper limits at 90% CL on branching

fractions in the range $(3.2 - 520)\times10^{-8}$ [91]. The two BNV τ decays presented here, $\tau^- \to \bar{p}\mu^+\mu^-$ and $\tau^- \to p\mu^-\mu^-$, have $|\Delta(B-L)| = 0$ and have not been studied by any previous experiment.

5.2 TAU PHYSICS AT LHCB

All previous results in τ LFV and BNV have been obtained at electron-positron colliders, due to the clean experimental environment, the low background level and the ability to tag τ decays via SM decays in the other 'hemisphere' of the event. Whilst the hadronic environment of LHCb negates these benefits it also yields a number of advantages of its own, particularly the large cross-section for inclusive τ production, clean detector signatures from muons and excellent particle identification from the RICH detectors. Therefore initially, final states containing muons are of most interest, but it is expected that in the future, searches will be extended to final states containing electrons and hadrons. To reduce systematic uncertainties and avoid requiring a precise knowledge of the integrated luminosity, LHCb analyses typically involve the normalisation to a known decay channel which is as similar as possible to the signal channel of interest. For the τ decays described here the chosen normalisation channel is $D_s^- \to \phi\pi^-$ with $\phi \to \mu^+\mu^-$, (referred to in the following as $D_s^- \to \phi\,(\mu^+\mu^-)\,\pi^-$), due to the similar decay kinematics and the presence of two muons in the final state. However, for this to be possible, an understanding of the cross-section for τ production at LHCb, $\sigma(pp \to \tau X)$, and the fraction of τ from D_s decays, $f_\tau^{D_s}$, is required.

Channel	LHCb cross-section	p_T (GeV)	η	Reference
$\sigma\left(b\bar{b}X\right)$	$284 \pm 53\,\mu b$	$p_T < \infty$	$-\infty < \eta < \infty$	[92]
$\sigma\left(D_s^{\pm}\right)$	$194 \pm 23 \pm 3\,\mu b$	$p_T < 8$	$2 < \eta < 4.5$	[93]
$\sigma\left(D^{\pm}\right)$	$717 \pm 39 \pm 15\,\mu b$	$p_T < 8$	$2 < \eta < 4.5$	[93]
$\sigma\left(\psi(2S)\right)$	$1.69 \pm 0.01 \pm 0.12^{+0.20}_{-0.40}\,\mu b$	$p_T < 16$	$2 < \eta < 4.5$	[94]
$\sigma\left(\Upsilon(1S)\right)$	$92.3 \pm 0.4 \pm 4.4^{+7.7}_{-14.9}$ nb	$p_T < 15$	$2 < \eta < 4.5$	[95]
$\sigma\left(Z\right)$	$76.7 \pm 1.7 \pm 4.3$ pb	$p_T > 20$	$2 < \eta < 4.5$	[96]
$\sigma\left(W^+\right)$	$831 \pm 9 \pm 40$ pb	$p_T > 20$	$2 < \eta < 4.5$	[96]
$\sigma\left(W^-\right)$	$656 \pm 40 \pm 30$ pb	$p_T > 20$	$2 < \eta < 4.5$	[96]

Table 13: Production cross-sections measured at $\sqrt{s} = 7$ TeV at LHCb and the p_T and η ranges over which they were performed.

5.2.1 Tau Production

Due to the presence of one or more neutrinos in the final state of any SM τ decay, a direct measurement of $\sigma\left(pp \to \tau X\right)$ at LHCb is not possible. However, by studying the possible production mechanisms of the τ, the cross-section can be calculated.

Table 13 lists the cross-sections for all particles which are known to decay into τ leptons. It immediately becomes clear that even with the differences in p_T and η intervals of the measurements, the contributions to $\sigma\left(pp \to \tau X\right)$ from anything other than than b and c production are negligible. Drell-Yan production is also considered to be negligible, based on the LHCb measurements of Drell-Yan produced dimuon final states [97] and assuming a small mass-dependent correction to the cross-section for $\tau^+\tau^-$ final states. Therefore the inclusive τ cross-section can be calculated by considering the branching fractions for b and c hadrons to τ leptons.

5.2.1.1 *Contribution from B decays*

The LHCb inclusive $b\bar{b}$ cross-section measurement [92] was performed by reconstructing $D^0 \to K^-\pi^+$ decays, where the D^0 mesons come from the semileptonic decays of b-hadrons. No requirement was placed on the flavour of the b-hadrons, and so the individual flavour contributions to the inclusive $b\bar{b}$ cross-section are unknown. Rather than trying to predict the separate flavour cross-sections, it is more accurate to consider the BF of the measured b-admixture to τ leptons. From the PDG [3] this is

$$\mathcal{B}\left(\bar{b} \to \tau^+\nu_\tau X\right) = 2.41 \pm 0.23\%,$$

where the contribution calculated in this way is referred to as 'primary' (not including $b \to c \to \tau$). There are also contributions to $\sigma(pp \to \tau X)$ from the decays of D-mesons produced via decay of the b-admixture, i.e. $b \to c \to \tau$ transitions, which are denoted as 'secondary'. The BF for $D_s^+ \to \tau^+\nu_\tau$ is taken from Ref. [98] as

$$\mathcal{B}\left(D_s^+ \to \tau^+\nu_\tau\right) = 5.61 \pm 0.24\%,$$

whilst here and in Section 5.2.1.2 the contribution of D^+ is included with a BF of $0.1 \pm 0.1\%$ on the basis that theoretically it should exist [99] but its BF has yet to be measured. Therefore, to account for all of the secondary contributions the BFs of the b-admixture to $D^\pm$ and $D_s^\pm$ mesons are also required. From the PDG these are known to be:

$$\begin{aligned}
\mathcal{B}\left(\bar{b} \to D^- X\right) &= 23.3 \pm 1.7\%, \\
\mathcal{B}\left(\bar{b} \to D_s^+ X\right) &= 10.1 \pm 3.1\%, \\
\mathcal{B}\left(\bar{b} \to D_s^- X\right) &= 14.7 \pm 2.1\%.
\end{aligned}$$

It is expected that there should also be a contribution from $\bar{b} \to D^+ X$ production. This has yet to be measured, so it is included here with the BF of $\bar{b} \to D^- X$ scaled by the ratio of $\bar{b} \to c \to l^+\nu_l X$ / $\bar{b} \to \bar{c} \to l^-\bar{\nu}_l X$ $\left(1.6^{+0.4}_{-0.5} \;/\; 8.02 \pm 0.19\right)$ as determined by LEP measurements [3], to account for the difference between

the c and $\bar{c}$ decay channels. It is given a 100% uncertainty. However, the effect of this on the overall τ cross-section is negligible.

The contribution to $\sigma(pp \to \tau X)$ from prompt B decays is calculated via

$$\sigma\left(b_{\text{primary}} \to \tau X\right) = 2 \times \sigma\left(b\bar{b}X\right) \times \mathcal{B}\left(\bar{b} \to \tau^{+}\nu_{\tau}X\right),$$

where the factor of two accounts for the production of b quarks in pairs. This is found to be

$$\sigma\left(b_{\text{primary}} \to \tau X\right) = 13.9 \pm 2.7\,\mu b,$$

in the full η region and over the full p_T range, where here, and in the following, the uncertainty contains both systematic and statistical contributions added in quadrature.

The contribution to $\sigma(pp \to \tau X)$ from secondary B decays is calculated via

$$\sigma\left(b_{\text{secondary}} \to \tau X\right) = 2 \times \sigma\left(b\bar{b}X\right) \times \sum_{i} \mathcal{B}\left(\bar{b} \to D_i X\right) \times \mathcal{B}\left(D_i \to \tau\nu_{\tau}\right),$$

where i indicates a sum over $D^{\pm}$ and $D_s^{\pm}$ mesons. This is found to be

$$\sigma\left(b_{\text{secondary}} \to \tau X\right) = 8.2 \pm 1.6\,\mu b,$$

in the full η region and over the full p_T range.

5.2.1.2 *Contribution from D decays*

The LHCb inclusive $c\bar{c}$ cross-section measurement [93] was performed by fully reconstructing decays of D^0, $D^{*\pm}$, $D^{\pm}$ and $D_s^{\pm}$ mesons to calculate the individual D-meson cross-sections, using the production fractions to determine the total inclusive $c\bar{c}$ cross-section from each, and then averaging over the results. To avoid the uncertainty introduced by the production fractions and the averaging process, the decays of the individual mesons are considered directly. Therefore the contribution to $\sigma(pp \to \tau X)$ from D-meson decays is

$$\sigma\left(D \to \tau X\right) = \sum_{i} \sigma\left(D_i\right) \times \mathcal{B}\left(D_i \to \tau\nu_{\tau}\right),$$

where i indicates a sum over $D^{\pm}$ and $D_s^{\pm}$ mesons. There is no factor of two here as the measured cross-sections are for mesons, not quarks. There is also a difference in the pseudorapidity and p_T intervals covered by the $b\bar{b}$ and $c\bar{c}$ cross-section measurements, with the $b\bar{b}$ measurement covering the full η region and the full p_T range. To make these two measurements comparable, a phase space extrapolation factor of 4.943 ± 0.014 is calculated with PYTHIA (from Ref. [93]), and applied to the D-meson contribution to give

$$\sigma(D \to \tau X) = 57.9 \pm 9.5\,\mu b,$$

in the full η region and over the full p_T range.

The total inclusive τ cross-section is therefore the sum of the three contributions listed above and is found to be

$$\sigma(pp \to \tau X) = 80.0 \pm 9.9\,\mu b,$$

at $\sqrt{s} = 7\,\text{TeV}$ in the full η region and over the full p_T range. The fraction $f_\tau^{D_s}$ can then be calculated as the sum of the primary and secondary D_s contributions divided by the total inclusive τ cross-section. The current systematic error on the cross-section is about 10%, largely due to uncertainties on the $b\bar{b}$ and $c\bar{c}$ cross-section measurements.

5.2.2 MC and Data Samples

The results described in the following are obtained using the data collected at a centre-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$ in 2011, corresponding to $1.0\,\text{fb}^{-1}$ of integrated luminosity. The data are taken from the DiMuon stream of the Reco12-Stripping17b reconstruction campaign and are reconstructed with BRUNEL and stripped with DAVINCI.

The MC samples used can be broadly separated into four categories: signal, normalisation, physics backgrounds and inclusive backgrounds, and are summarised in Table 14. They are all from the MC11a production campaign and in-

Channel	Number of generated events
Signal:	
$\tau^- \to \mu^-\mu^+\mu^-$	1 104 490
$\tau^- \to \bar{p}\mu^+\mu^-$	268 300
$\tau^- \to p\mu^-\mu^-$	265 100
Control/normalisation:	
$D_s^- \to \phi\,(\mu^+\mu^-)\,\pi^-$	1 111 994
$D_s^- \to \eta\,(\mu^+\mu^-\gamma)\,\mu^-\bar{\nu}_\mu$	4 731 512
Background:	
Inclusive $b\bar{b} \to \mu^+\mu^-$	22 948 903
Inclusive $c\bar{c} \to \mu^+\mu^-$	26 714 154

Table 14: MC simulation samples used in the 2011 analysis of the 7 TeV dataset. All samples are produced with an approximately equal fraction of each magnet polarity.

clude the requirement that the particles of interest are in the LHCb acceptance, which is referred to in the following as 'DecProdCut' or DPC.

The trigger has been emulated during MC production using TCK 0x40760037, which was chosen to describe the most common trigger conditions throughout the year.

For the signal channels the events were simulated using a simple phase-space model for the decay of the τ.

5.2.3 MC Correction

Ideally the default LHCb MC would reproduce the fractional contributions of the different τ production subchannels described in Section 5.2.1. These are given in Table 15 as '$\text{Calc}_{4\pi}$'. However, this is not the case as the ratio of the inclusive $c\bar{c}$ and $b\bar{b}$ cross-sections differs from the latest LHCb measurements, and the MC first must be corrected before it can be used. This was achieved via the reweighting of the individual subchannels. As the τ cross-section determined in Section 5.2.1 is calculated over 4π solid angle and with no cuts, it cannot be compared directly to the proportions found in the MC sample. Therefore approximately 0.25 M signal MC events were generated with GAUSS using the same settings as the official production, but without the DPC or the detector simulation. This sample, referred to in Table 15 as 'No DPC', can then be compared directly to the calculation to determine the ratios of the contributions based on the cross-section calculation to those in the LHCb MC. The resultant weights used to correct the MC are given in Table 15 as 'w_{calc}'. The weights are independent of the final state particles from the τ decay and are therefore identical for $\tau^- \to \mu^-\mu^+\mu^-$ and $\tau \to p\mu\mu$.

For $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ reweighting, the only applicable channels are $D_s \to \tau$ and $B_x \to D_s \to \tau$, to account for prompt and secondary production respectively. Therefore when the reweighting procedure is performed these weights are normalised such that the total sample size is unchanged. In the following when referring to MC samples, it is assumed that they are reweighted.

Differences between IP resolution in MC and data are apparent due to, among other reasons, the poor description of the shape and mass of the VELO RF foil in the MC simulation. Therefore a smearing procedure is performed where the width of the IP distribution of tracks in a given bin of p_T and ϕ in MC is corrected according to the equivalent width in minimum bias data. To allow for evaluation of the systematic uncertainty of this method, samples with 50% over-

Decay chain	$\text{Calc}_{4\pi}$ (%)	No DPC (%)	w_{calc}
$D_s \to \tau$	68.5 ± 5.2	49.1 ± 0.1	1.4 ± 0.1
$B_x \to D_s \to \tau$	10.0 ± 2.1	7.45 ± 0.04	1.3 ± 0.3
Sum $D_s \to \tau$	78.5 ± 4.8	56.5 ± 0.1	
$D^- \to \tau$	4.0 ± 3.9	2.31 ± 0.02	1.7 ± 1.7
$B_x \to D^- \to \tau$	0.2 ± 0.2	0.22 ± 0.01	0.9 ± 0.8
Sum $D^- \to \tau$	4.2 ± 3.9	2.53 ± 0.02	
$B_x \to \tau$	17.4 ± 3.5	40.7 ± 0.1	0.43 ± 0.08

Table 15: Details of the MC reweighting in the 2011 analysis. The sum of the D_s and D^- subchannels are given in the third and sixth rows for reference, but are not used in the reweighting procedure.

smearing and without smearing are also prepared. In the following, the MC is assumed to be smeared unless otherwise specified.

5.3 ANALYSIS METHOD

As discussed previously, the measurement of the signal branching fractions are normalised to the $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ decay channel. Selection criteria are implemented for the three signal modes, $\tau^- \to \mu^-\mu^+\mu^-$, $\tau^- \to \bar{p}\mu^+\mu^-$ and $\tau^- \to p\mu^-\mu^-$, and for the normalisation channel $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ and are designed to be as loose as possible, whilst reducing the dataset to a manageable level. To avoid potential bias, $\mu^-\mu^+\mu^-$ and $p\mu\mu$ candidates with invariant mass within $\pm 30\,\text{MeV}$ ($\approx 3\sigma_m$) of the τ mass are initially blinded from the analysis, where σ_m denotes the expected mass resolution, as determined from MC. For the $\tau^- \to \mu^-\mu^+\mu^-$ channel, further discrimination between potential signal and background is achieved via a two-dimensional binned distribution in two

likelihood variables. One likelihood variable is based on the three-body decay topology and the other on muon identification. For the $\tau \to p\mu\mu$ channels, the use of the second likelihood function is replaced by cuts on the proton and muon PID variables. This was initially motivated by complications resulting from including the different PID behaviour of the final state proton into the likelihood. However, it is expected that the PID cuts will be replaced by a dedicated multivariate PID classifier in the analysis of the full $3.0\,\text{fb}^{-1}$, as discussed in the following.

5.3.1 Signal and Background Discrimination

5.3.1.1 *Selection*

The signal and normalisation channels have a similar topology, shown in Fig. 60 for $\tau^- \to \mu^-\mu^+\mu^-$ in the case where the τ comes from the decay of a promptly produced b or c hadron. The selection criteria are a combination of predefined cuts in the stripping and additional cuts, and are shown in Table 16. They are designed to take advantage of this 'prompt topology' and therefore do not specifically consider the τ leptons that come from the decays of secondary charm mesons (approximately 10%, as shown in Table 15). Despite having an additional vertex in their decay topology, the majority of the event properties of secondary decays are similar to those of prompt decays, and are therefore still efficiently selected by the cuts in Table 16. The following description assumes the prompt topology only.

As all τ leptons come from the decay of b or c hadrons, the measurable flight distances of the heavy hadrons give a secondary vertex displaced from the primary vertex. The decay vertex (DV) of the τ is also displaced from the primary vertex, and has three tracks that are reconstructed to give a mass close to that of the τ lepton (or D_s meson for the normalisation channel). Therefore well-reconstructed and well-identified muon, pion and proton tracks with large IP

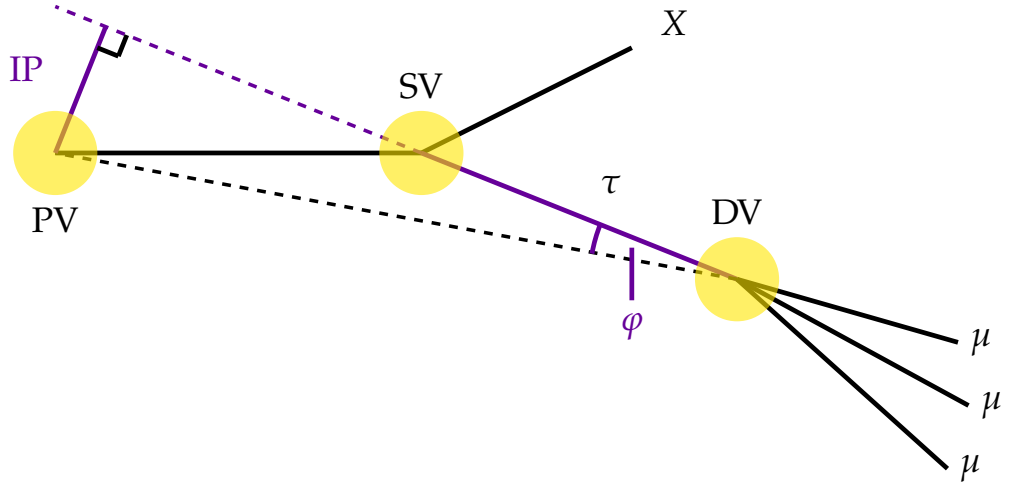


Figure 60: Topology of a $\tau^- \to \mu^-\mu^+\mu^-$ event where the τ comes from the decay of a prompt hadron.

are required. All muons are also required to register hits in the muon detectors and to pass the associated IsMuon requirement. The fitted three-track vertex has to be of good quality, and the measured decay time of the candidate forming the vertex has to be compatible with that of a heavy meson or τ lepton. Since the Q-values in decays of charm mesons to τ are relatively small, the pointing angle, φ, between the momentum vector of the three-track system and the line joining the primary and decay vertices is required to be small. For the $\tau \to p\mu\mu$ signal channels the muon and proton candidates must pass loose PID requirements. In the $\tau^- \to \mu^-\mu^+\mu^-$ channel, signal candidates with a $\mu^+\mu^-$ mass close to that of the ϕ meson mass are removed, and to eliminate irreducible background near the signal region arising from the decay $D_s^- \to \eta\,(\mu^+\mu^-\gamma)\,\mu^-\overline{\nu}_\mu$, candidates with low $\mu^+\mu^-$ mass are also rejected (see Section 5.3.2). Finally, to remove potential contamination from clone tracks, low mass same-sign muon pairs are removed in both the $\tau^- \to \mu^-\mu^+\mu^-$ and $\tau^- \to p\mu^-\mu^-$ channels. The $\tau \to p\mu\mu$ selection, shown in Table 17 includes tight p_T and cos φ cuts and is known to give a suboptimal selection efficiency, having been implemented in the stripping before the analysis was started. This has been updated for the 3.0 fb^{-1} measurement and is

	$\tau^- \to \mu^-\mu^+\mu^-$	$D_s^- \to \phi\,(\mu^+\mu^-)\,\pi^-$
$\mu^\pm$, $\pi^\pm$		
p_T	$> 300\,\text{MeV}$	
Track χ^2/ndf	< 4	
IP χ^2/ndf	> 9	
μ pairs		
$\lvert m_{\mu^+\mu^-} - m_\phi \rvert$	$> 20\,\text{MeV}$	$< 20\,\text{MeV}$
$m_{\mu^+\mu^-}$	$> 450\,\text{MeV}$	-
$m_{\mu^+\mu^+}$	$> 250\,\text{MeV}$	-
$\tau^\pm$ and $D_s^\pm$		
Δm	$< 400\,\text{MeV}$	$< 50\,\text{MeV}$
Vertex χ^2	< 15	
IP χ^2	< 225	
$\cos\varphi$	> 0.99	
$c\tau$	$> 100\,\mu\text{m}$	
Decay time	$> -0.01\,\text{ns}$ & $< 0.025\,\text{ns}$	

Table 16: Selections for the $\tau^- \to \mu^-\mu^+\mu^-$ and $D_s^- \to \phi\,(\mu^+\mu^-)\,\pi^-$ channels for the $\tau^- \to \mu^-\mu^+\mu^-$ analysis. All muons must also pass the IsMuon requirement. The symbols have their usual meanings.

	$\tau \to p\mu\mu$	$D_s^- \to \phi\,(\mu^+\mu^-)\,\pi^-$
$\mu^\pm$, p and $\pi^\pm$		
p_T	$> 300\,\text{MeV}$	
p	$> 5\,\text{GeV}$ & $< 100\,\text{GeV}$	
Track χ^2/ndf	< 3	
IP χ^2/ndf	> 9	
DLL$(\mu - \pi)$	> -5	
DLL$(\mu - K)$	> 0	
DLL$(p - \pi)$	> -5	-
μ pairs		
$\lvert m_{\mu^+\mu^-} - m_\phi \rvert$	-	$< 20\,\text{MeV}$
$m_{\mu^+\mu^+}$	$> 250\,\text{MeV}$	-
$\tau^\pm$ and $D_s^\pm$		
Δm	$< 250\,\text{MeV}$	$< 50\,\text{MeV}$
p_T	$> 4\,\text{GeV}$	
Vertex χ^2	< 15	
IP χ^2	< 225	
$\cos\varphi$	> 0.999	
$c\tau$	$> 100\,\mu\text{m}$	
Decay time	$> -0.01\,\text{ns}$ & $< 0.025\,\text{ns}$	

Table 17: Selections for the $\tau \to p\mu\mu$ and $D_s^- \to \phi\,(\mu^+\mu^-)\,\pi^-$ channels for the $\tau \to p\mu\mu$ analysis. All muons must also pass the IsMuon requirement. The DLL variables are described in Section 3.2.3.1. The symbols have their usual meanings.

expected to yield a significant improvement in signal efficiency. To maintain the correspondence between the cuts applied to the signal and normalisation channels, the $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ selection is different when used in the $\tau \to p\mu\mu$ measurement.

The signal regions are defined by $\pm 20\,\text{MeV}$ (approximately twice the signal mass resolution) windows around the nominal τ mass. The values of the τ, D_s and $\phi(1020)$ masses measured in MC (using the fit model described in Section 5.3.1.2) are used as the central values for any mass window cuts. Candidates within mass windows of $\pm 400\,\text{MeV}$ for $\tau^- \to \mu^-\mu^+\mu^-$ decays and $\pm 250\,\text{MeV}$ for $\tau \to p\mu\mu$ decays are kept to allow evaluation of the background contributions in the signal regions. These are referred to in the following as the data sidebands. A mass window of $\pm 20\,\text{MeV}$ is also used to define the signal region for the $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ channel, with the $\mu^+\mu^-$ mass required to be within $\pm 20\,\text{MeV}$ of the $\phi(1020)$ meson mass.

The selected events in MC must also pass the LHCb trigger, with the requirement that some combination of any of the physics triggers at L0, HLT1 and HLT2 has fired.

5.3.1.2 *Multivariate classification*

After the selection, each event is given a probability to be signal or background according to the values of several likelihoods. For $\tau^- \to \mu^-\mu^+\mu^-$ two likelihoods are used:

- **Geometry:** The geometric likelihood is a multivariate classifier based on the geometry and kinematics of the final state tracks and the reconstructed τ candidate. It aims to reject backgrounds from combinations of tracks which do not share a common vertex and those from N-body decays, where $N > 3$. In the following this classifier is referred to as $\mathcal{M}_{3\text{body}}$.

- **Particle identification:** The PID likelihood is a multivariate classifier provided by the LHCb PID group which gives the probability that each of the

three final state particles is compatible with the muon hypothesis. It aims to reject backgrounds from three-body decays involving misidentified particles. In the following this classifier is referred to as $\mathcal{M}_{\text{PID}}$.

In addition, the limit calculation is performed in bins of the mass of the reconstructed τ candidate. For $\tau \to p\mu\mu$ only the $\mathcal{M}_{\text{3body}}$ likelihood is used and the $\mathcal{M}_{\text{PID}}$ likelihood is replaced by PID cuts.

The geometric and PID likelihoods are developed using MC and are calibrated on data. The response of the $\mathcal{M}_{\text{3body}}$ for $\tau \to p\mu\mu$ is determined using the training from the $\tau^- \to \mu^-\mu^+\mu^-$ samples. The extent to which this is suboptimal depends on the kinematical and topological differences between the $\tau^- \to \mu^-\mu^+\mu^-$ and $\tau \to p\mu\mu$ channels. Whilst these are found to be small, in the $3.0\,\text{fb}^{-1}$ update of the $\tau \to p\mu\mu$ analysis $\mathcal{M}_{\text{3body}}$ will be specifically trained on $\tau \to p\mu\mu$ samples.

The choice of variables for the $\mathcal{M}_{\text{3body}}$ classifier is based on their power to discriminate signal from background. After highly correlated variables are excluded, the following minimal set is used:

- **Track quality:** The quality of the fit to the hits which form a track in the detector in terms of their χ^2 value. As each event consists of three tracks the smallest of the three χ^2 values is used.

- μ **impact parameter:** The IP of the muon candidate tracks with respect to the best primary vertex, in terms of their χ^2 per degree of freedom divided by the uncertainty (the χ^2 significance). As each event consists of three tracks the smallest value is used.

- **Track isolation:** The number of other tracks in the event that can make a vertex with each muon candidate according to the definition used in the LHCb $B_s \to \mu^+\mu^-$ analysis [100]. Details of the definition of the isolation are given in Appendix B. As each event consists of three muon candidates the sum of the track isolations is used.

- **Distance of closest approach:** The distance of closest approach (DOCA) of two of the three muon candidates forming the τ vertex. There are three possible two-muon pairings and all three are used.
- τ **transverse momentum:** The transverse momentum of the reconstructed τ candidate.
- τ **impact parameter:** The IP of the τ candidate with respect to the best primary vertex, in terms of the χ^2 significance.
- **Vertex fit quality:** The fit quality of the reconstructed τ vertex in terms of χ^2.
- **Decay time:** The measured decay time of the τ candidate assuming production at the primary vertex.
- **Pointing angle:** The angle, φ, between the τ candidate momentum vector and a straight line from the τ candidate decay vertex to the best primary vertex.
- τ **isolation:** The transverse momentum of the τ candidate divided by the sum of transverse momenta in a cone around it, as used in the CDF $B_s \to \mu^+\mu^-$ analysis [101]. Details of the definition of the isolation are given in Appendix B.

The $\mathcal{M}_{3body}$ likelihood uses a boosted decision tree (BDT [102]), with the AdaBoost algorithm [103], and is implemented via the TMVA Toolkit [104]. The response varies from -1 to $+1$, with -1 being most background-like and $+1$ most signal-like. Whilst several possible classifiers were considered, the BDT was chosen as it was found to give the best background rejection for 99% signal efficiency. The BDT is trained using $\tau^- \to \mu^-\mu^+\mu^-$ signal MC and the inclusive background samples listed in Table 14. The contribution of the background samples is weighted according to their relative abundance as measured in data.

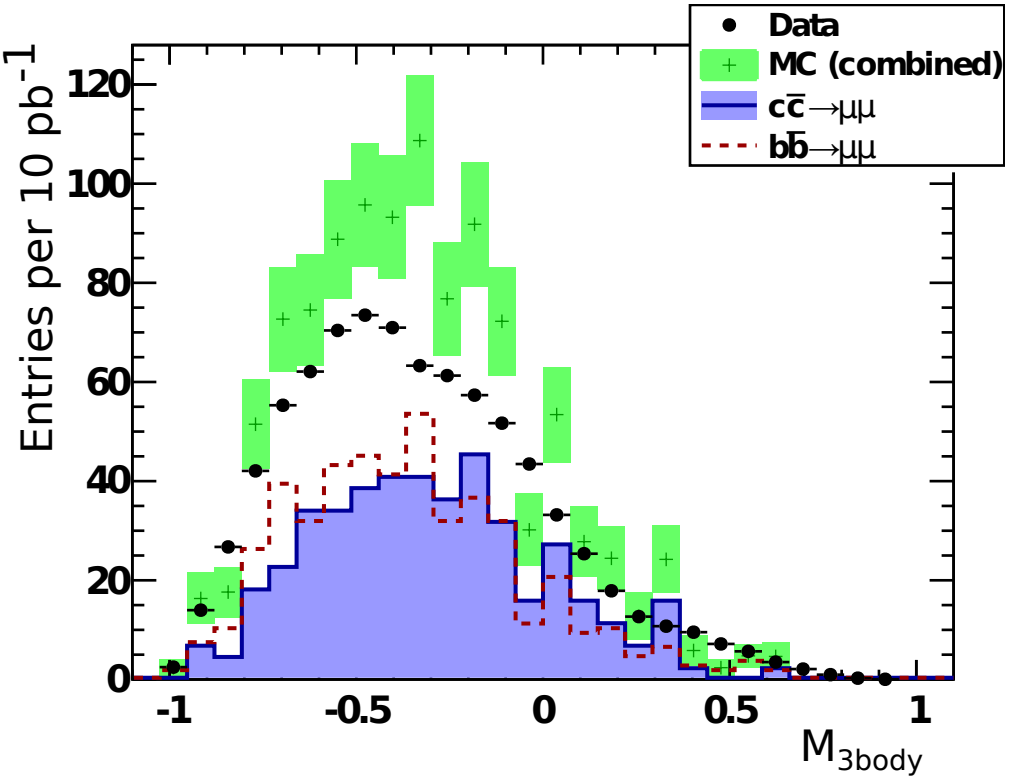


Figure 61: Comparison of the $\mathcal{M}_{3body}$ response for the $\tau^- \to \mu^-\mu^+\mu^-$ data sidebands and the sum of the inclusive MC samples.

Figure 61 compares the $\mathcal{M}_{3body}$ response for data and for the sum of the inclusive MC samples, and shows that the MC is sufficient to describe the major sources of background. However, as the MC samples are forced to include two real muons in each event at the generator level, the level of misidentified muons is incorrectly modelled in the MC, and the MC prediction is slightly above the data.

The $\mathcal{M}_{3body}$ response is calibrated on $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ events in data to account for potential differences between data and MC. This is justified by the excellent agreement between $\tau^- \to \mu^-\mu^+\mu^-$ and $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ MC for a number of the variables used in $\mathcal{M}_{3body}$, as shown in Fig. 62. Small differences between the two channels exist in some cases, such as the IP significance, where the addition of an extra vertex (e.g. promptly produced $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ compared to promptly produced $D_s \to \tau^- \to \mu^-\mu^+\mu^-$) causes a discrepancy.

Therefore correction factors are determined using MC and applied to $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ to give the expected $\tau^- \to \mu^-\mu^+\mu^-$ response for data. Correction factors are also calculated comparing unsmeared $\tau^- \to \mu^-\mu^+\mu^-$ MC to smeared $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ MC and the differences from the nominal values are assigned as systematic uncertainties.

Comparisons of $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ in data and in MC for several variables used in the $\mathcal{M}_{3body}$ classifier and the $\mathcal{M}_{3body}$ response itself are shown in Fig. 63. In general agreement is good and any discrepancies are accounted for in the calibration.

The $\mathcal{M}_{PID}$ likelihood uses a neural network, combining information from the PID subdetectors, tracking information and kinematical variables to give the probability of a track for a particular PID hypothesis. The resulting responses, known as 'ProbNNs', vary from 0 (most background-like) to +1 (most signal-like). In the case of $\tau^- \to \mu^-\mu^+\mu^-$, the muon hypothesis, ProbNNmu, is used. As there is one response for each final state particle the smallest of the three is used as the classifier. The $\mathcal{M}_{PID}$ response is calibrated on a pure sample of $J/\psi \to \mu^+\mu^-$ events in data, where the $\mathcal{M}_{PID}$ values for a single track in MC are corrected to the data response of a muon with similar kinematics. The statistical uncertainty on the number of muons in the calibration sample is combined with a systematic uncertainty of 1% per track to give the uncertainty on the calibrated $\mathcal{M}_{PID}$ response.

For $\tau \to p\mu\mu$ the same classifier cannot be used, due to the inclusion of the proton in the final state. Cuts on the DLLs (described in Section 3.2.3.1) are instead applied and are optimised on signal MC and outer data sidebands. In the $3.0\,fb^{-1}$ update the DLL variables will be replaced by ProbNNmu and the probability for the proton hypothesis, ProbNNp.

The PID cuts are tuned on the outer data sidebands, defined as the mass regions 1528.5 – 1643.5 MeV and 1913.5 – 2028.5 MeV, using a cut optimiser. The optimisation is performed over four PID variables: proton DLL$(p-\pi)$, proton

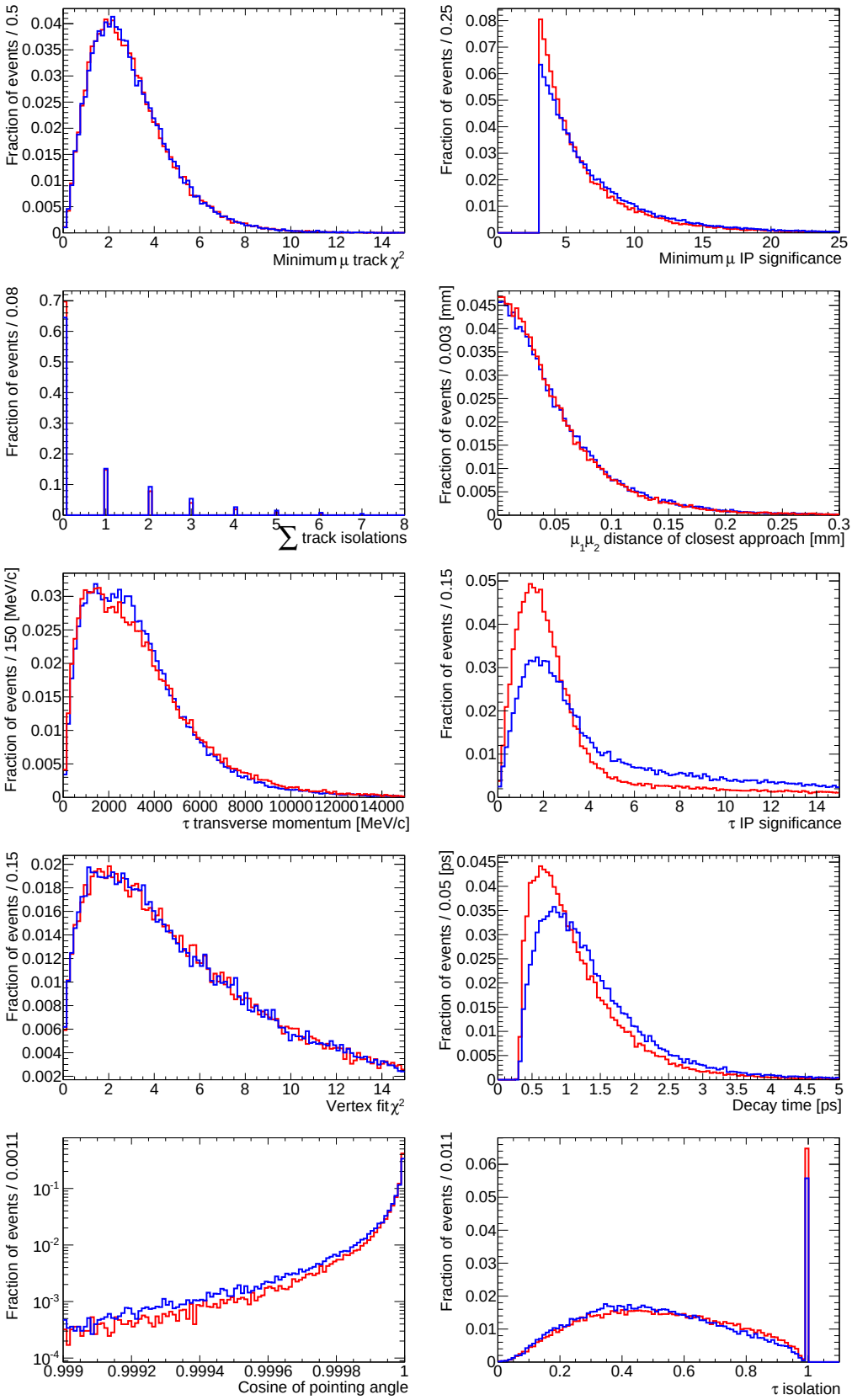


Figure 62: Comparison of $\tau^- \to \mu^-\mu^+\mu^-$ (blue) and $D_s^- \to \phi\,(\mu^+\mu^-)\,\pi^-$ (red) MC for the variables used in $\mathcal{M}_{3\text{body}}$.

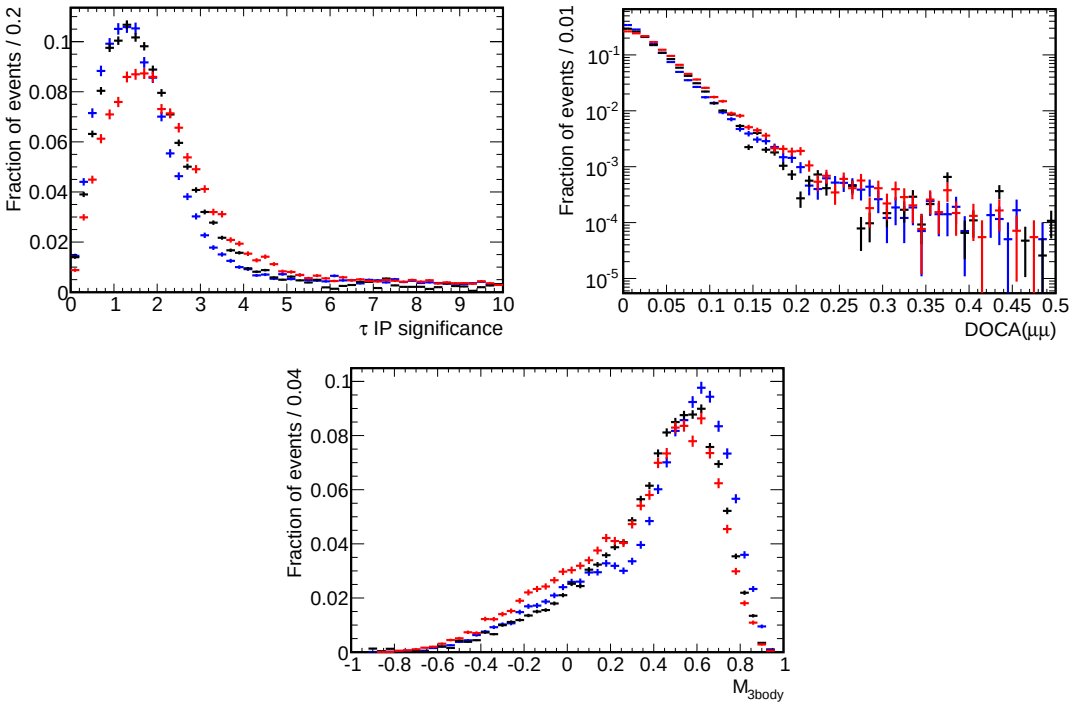


Figure 63: Comparison of τ IP significance (top left), DOCA (top right) and $\mathcal{M}_{3\text{body}}$ response (bottom) for $D_s^- \to \phi\,(\mu^+\mu^-)\,\pi^-$ in data (black), in MC (blue) and in smeared MC (red).

$\text{DLL}(p-K)$, muon $\text{DLL}(\mu-\pi)$ and muon $\text{DLL}(\mu-K)$, using the figure of merit (FoM)

$$\frac{S}{1+\sqrt{B}},$$

described by Punzi in Ref. [105], optimised for a $2\,\sigma$ significance. As the PID cuts are optimised to maximise the FoM, it is important that this optimisation is performed using PID distributions that correctly describe data, such that the signal efficiency is properly estimated. From Fig. 64 (prepared using the PIDCalib package, which is the default tool in LHCb for assessing the difference between MC and data PID response) it can be seen that there are large discrepancies between the MC and data response, as indicated by the difference between the raw MC (blue) and data corrected MC (red). The PID variables for protons are found to contain the largest discrepancy between MC and data, whilst the PID variables for muons are in reasonable agreement. Therefore for the cut optimi-

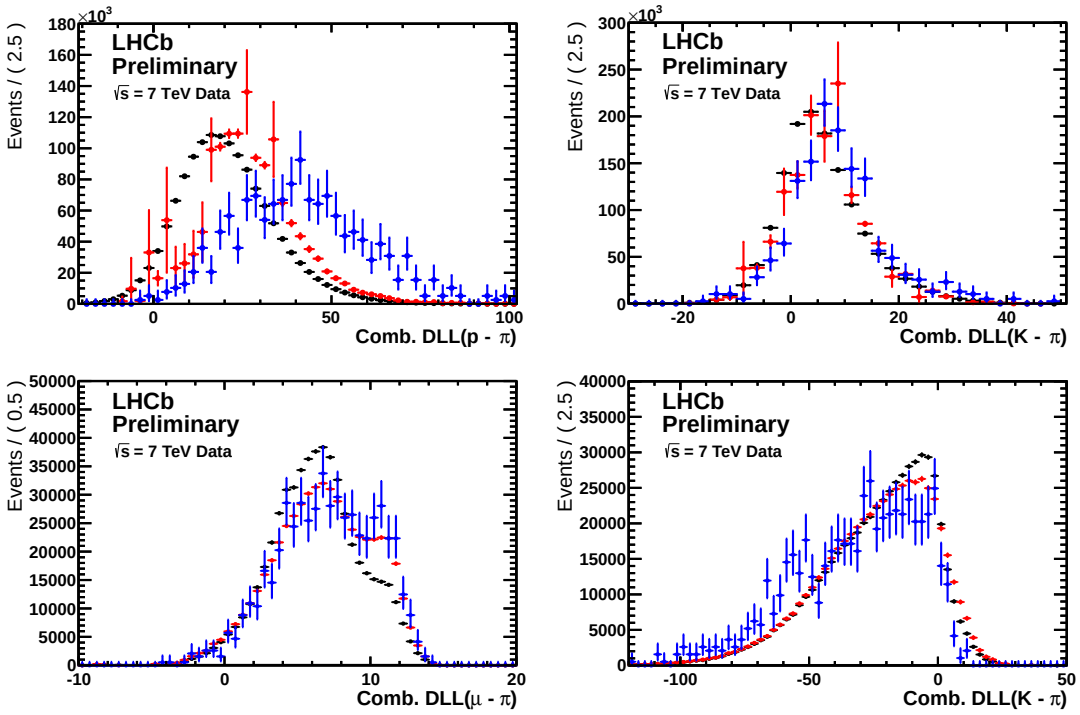


Figure 64: Output of the PIDCalib package, showing the MC PID distributions before (blue) and after (red) reweighting to the data response (black). Top left is DLL($p-\pi$), top right is DLL($K-\pi$), bottom left is DLL($\mu-\pi$) and bottom right is DLL($K-\pi$).

sation the proton PID in MC is corrected to give the corresponding response in data, whilst the muon PID is kept the same as the discrepancy between data and MC is minimal. Any remaining discrepancies between data and MC result only in a slightly suboptimal choice of cuts and in the calculation of the normalisation factor all PID cut efficiencies are corrected to data, as described in Section 5.3.3.6.

Three different $\mathcal{M}_{3\text{body}}$ binning schemes are considered during the optimisation, with the scenario selected that yielded the largest value of the FoM. The optimisation is found to perform most effectively if the lowest bin in $\mathcal{M}_{3\text{body}}$ (described later) is omitted, as the number of poorly reconstructed events in the data sample is reduced, such that the cuts are optimised on a larger proportion of events with genuine misidentified particles.

Variable	Cut
$\mathrm{DLL}(\mu - \pi)$	> 1
$\mathrm{DLL}(\mu - K)$	> 20
$\mathrm{DLL}(p - \pi)$	> 15
$\mathrm{DLL}(p - K)$	> 9

Table 18: PID cuts used in the selection of the $\tau \to p\mu\mu$ signal channels.

The final cut values are shown in Table 18 and are applied to both the $\tau^- \to \overline{p}\mu^+\mu^-$ and $\tau^- \to p\mu^-\mu^-$ data and MC samples.

The shape of the invariant mass spectrum for the $\tau^- \to \mu^-\mu^+\mu^-$ signal is taken from a fit to the $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ control channel, using the selection outlined in 5.3.1.1. The signal distribution is modelled with the sum of two gaussian functions with common central value, where the first gaussian function is restricted to be narrower than the second and contribute 70% of the total signal yield. The restrictions are required to achieve a common fit model across all samples. The combinatorial background (data only) is modelled with a linear function.

The result of the fits to $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ events in data after the $\tau^- \to \mu^-\mu^+\mu^-$ and $\tau \to p\mu\mu$ normalisation selections are shown in Fig. 65, and are summarised in Table 19. No PID requirements are placed on the pion. The $\phi \to \mu^+\mu^-$ peak in the $\mu^+\mu^-$ mass spectrum for the $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ channel is shown in Fig. 66 for reference. The substantially smaller yield of $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ events, N_{cal}, in the $\tau \to p\mu\mu$ channel is due to the additional cuts imposed in the selection and the application of PID cuts instead of the multivariate classification. The values of N_{cal} include the systematic error due to the fit model, determined by allowing the fractions of the gaussian functions in the fit to vary. The fit model is used to determine the central values for the masses in the $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ channel in both MC and data. A common central value

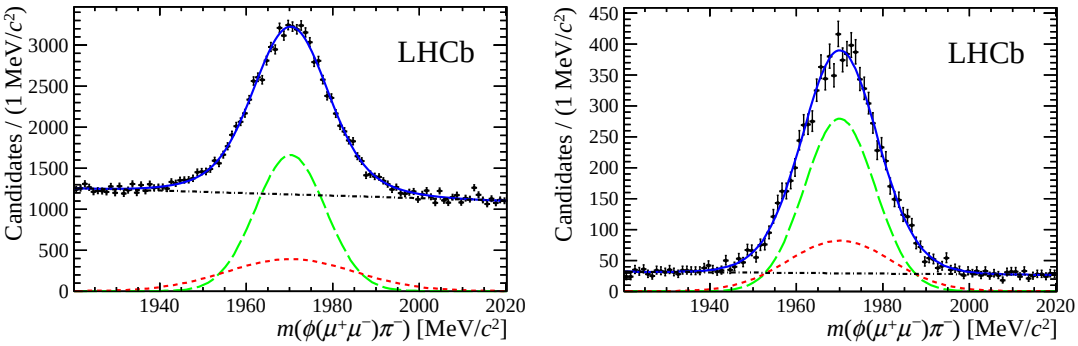


Figure 65: Fit to the invariant mass distribution of $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ events in 2011 data after the $\tau^- \to \mu^-\mu^+\mu^-$ selection and trigger (left) and the $\tau \to p\mu\mu$ selection and trigger (right). No PID requirements are placed on the pion. The blue lines show the overall fits, the green and red lines show the two gaussian components of the signal and the black lines show the combinatorial backgrounds.

of 1970.2 MeV is determined for all samples, including both $\tau^- \to \mu^-\mu^+\mu^-$ and $\tau \to p\mu\mu$ selections, and therefore the central values from MC are used for the signal channels. The $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ mass resolution is found to be slightly larger in data than in MC. The correction for the difference in mass resolution between MC and data is described further in Section 5.3.3.5.

To achieve optimum sensitivity to the $\tau^- \to \mu^-\mu^+\mu^-$ branching fraction, the limit calculation is performed in bins of the $\mathcal{M}_{3body}$ and $\mathcal{M}_{PID}$ likelihoods, rather than simply cutting on particular response values. As the limit is sensitive to the choice of binning scheme, the binning is chosen such that the separation power between the background-only and signal-plus-background hypotheses is maximised according to the FoM

$$\Delta LQ = 2\ln(Q_{SB}) - 2\ln(Q_B),$$

where,

$$Q_{SB} = \prod \frac{P(s_i + b_i, s_i + b_i)}{P(s_i + b_i, b_i)},$$
$$Q_B = \prod \frac{P(b_i, s_i + b_i)}{P(b_i, b_i)}.$$

Fitted parameter	$\tau^- \to \mu^-\mu^+\mu^-$	$\tau \to p\mu\mu$
N_{cal}	$48\,076 \pm 760 \pm 357$	$8145 \pm 127 \pm 128$
Mean (MeV)	1970.29 ± 0.08	1970.1 ± 0.1
σ_1 (MeV)	8.10 ± 0.09	8.15 ± 0.20
σ_2 (MeV)	14.8 ± 0.9	11.92 ± 0.84

Table 19: Results of the fits to $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ data events for the $\tau^- \to \mu^-\mu^+\mu^-$ and $\tau \to p\mu\mu$ channels. The widths of the gaussian functions are denoted by σ_1 and σ_2.

In the above, $P(a, e)$ is the probability to observe a when the quantity has a Poisson distribution with expectation value e, and the index i indicates the bin number for which the probabilities are computed; s_i and b_i are the number of signal and background events in each bin. The optimisation is performed after all selection cuts have been applied. The FoM tends to increase as further bins are added, with the improvement becoming smaller with each additional bin. As a large number of bins complicates the analysis, once the fractional improvement becomes below one percent no further increase in binning is considered.

The optimum number of bins is found to be six for $\mathcal{M}_{3body}$ and $\mathcal{M}_{PID}$ in the $\tau^- \to \mu^-\mu^+\mu^-$ channel, and five for $\mathcal{M}_{3body}$ in the $\tau \to p\mu\mu$ channel with the same binning used for both $\tau^- \to \bar{p}\mu^+\mu^-$ and $\tau^- \to p\mu^-\mu^-$. In the case of $\tau^- \to \mu^-\mu^+\mu^-$ the lowest bin in $\mathcal{M}_{PID}$ contains a large contribution from misidentified particles, such as $D^+_{(s)} \to K^-K^+\pi^+$ and $D^+_{(s)} \to \pi^-\pi^+\pi^+$ which peak close to the signal region. To avoid complications from attempting to model these contributions this bin is instead removed from the analysis as it does not contribute to the sensitivity. The same procedure is applied to the lowest bin in $\mathcal{M}_{3body}$ such that the number of sensitive bins is five in both cases, with the excluded bins referred to as the "trash bins" in the following. The effect of removing these bins on the signal efficiency is considered in Section 5.3.3.5. For the mass binning in the limit calculation, the choice of number of bins and their

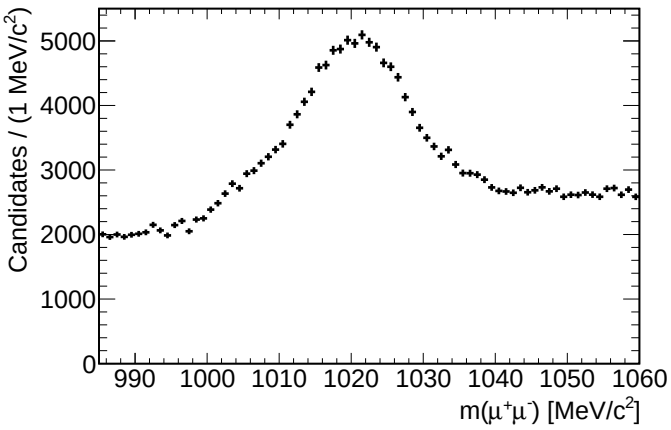


Figure 66: Invariant mass distribution of $\mu^+\mu^-$ in the $D_s^- \to \phi\,(\mu^+\mu^-)\,\pi^-$ channel after the $\tau^- \to \mu^-\mu^+\mu^-$ selection and trigger. For clarity, the $\phi(1020)$ mass cut has not been applied.

widths is less significant, and the data are divided into eight equally spaced bins in the $\pm 20\,\text{MeV}$ mass window around the nominal τ mass. The binned distributions of $\mathcal{M}_{3\text{body}}$, $\mathcal{M}_{\text{PID}}$ (where applicable) and invariant mass (before the correction for the difference in mass resolution between data and MC described in Section 5.3.3.5 is applied) are shown in Fig. 67 and Fig. 68 for $\tau^- \to \mu^-\mu^+\mu^-$ and $\tau \to p\mu\mu$ respectively. Figure 69 shows the distribution of $\tau^- \to \mu^-\mu^+\mu^-$ MC events in the $\mathcal{M}_{3\text{body}}$ and $\mathcal{M}_{\text{PID}}$ bins, and is used as the probability density function (PDF) in the limit calculation.

5.3.2 Background Characterisation

The background processes for the $\tau^- \to \mu^-\mu^+\mu^-$ decay consist mainly of a combination of heavy meson decays yielding three real muons in the final state, or one or two real muons in combination with two or one misidentified particles. There are also a large number of events with one or two muons from heavy meson decays combined with a random muon, or muons from elsewhere in the event. Decays with final state particles such as neutrinos, photons or neutral pions that are not detected can give large backgrounds which vary

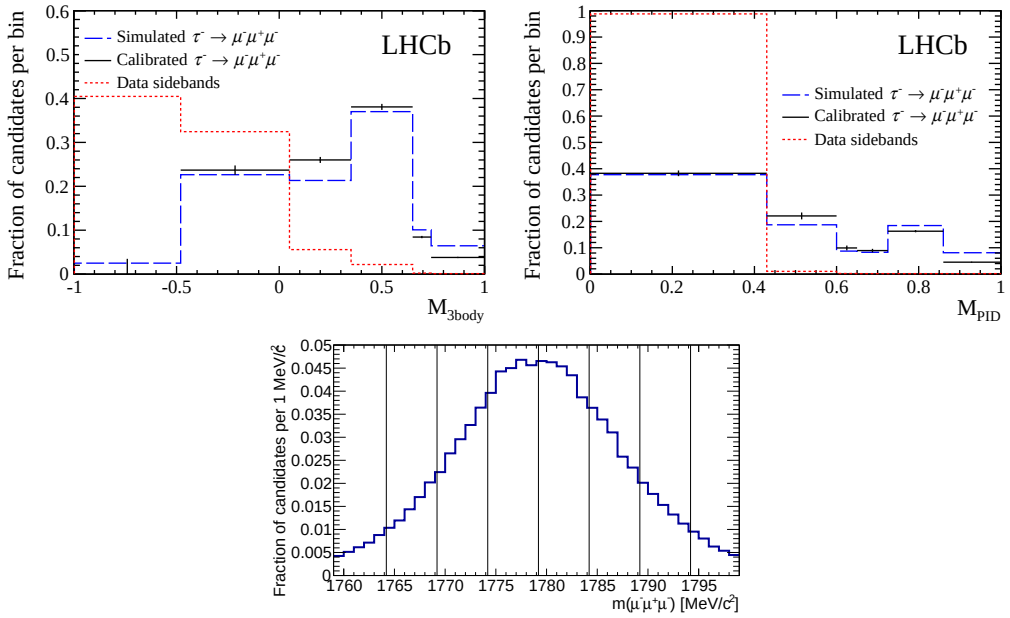


Figure 67: Distribution of $\mathcal{M}_{3body}$ (top left), $\mathcal{M}_{PID}$ (top right) and invariant mass (bottom) of the τ candidate for signal MC and data sidebands for $\tau^- \to \mu^-\mu^+\mu^-$. The binning corresponds to that used in the limit calculation.

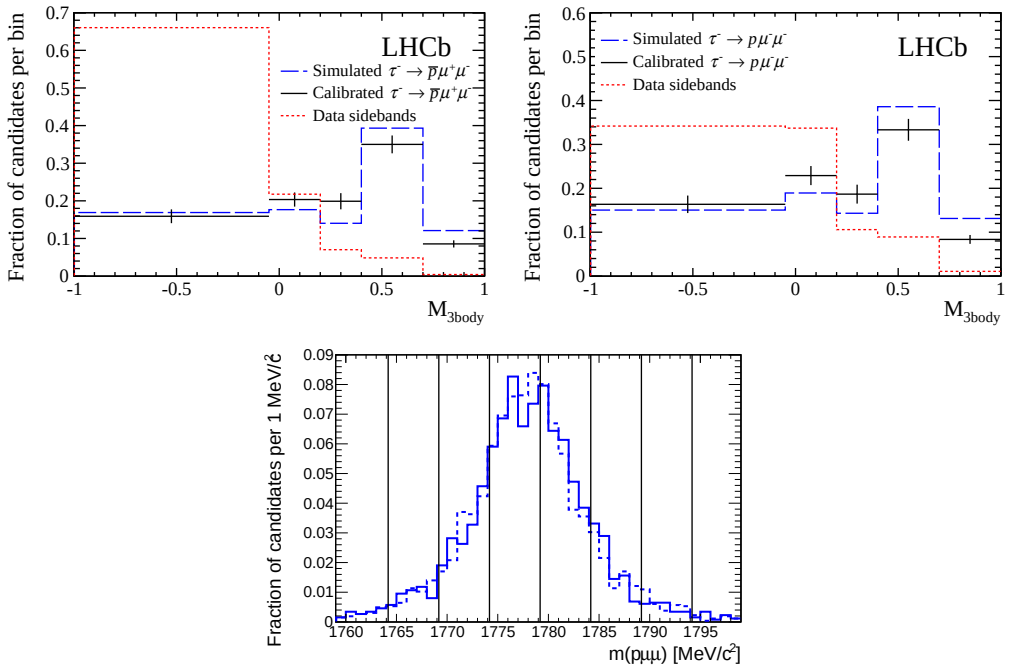


Figure 68: Distribution of $\mathcal{M}_{3body}$ for $\tau^- \to \bar{p}\mu^+\mu^-$ (top left), $\mathcal{M}_{3body}$ for $\tau^- \to p\mu^-\mu^-$ (top right) and invariant mass (bottom) of the τ candidate for $\tau \to p\mu\mu$, for signal MC and data sidebands. For invariant mass both $\tau^- \to \bar{p}\mu^+\mu^-$ (solid blue) and $\tau^- \to p\mu^-\mu^-$ (dashed blue) are shown. The binning corresponds to that used in the limit calculation.

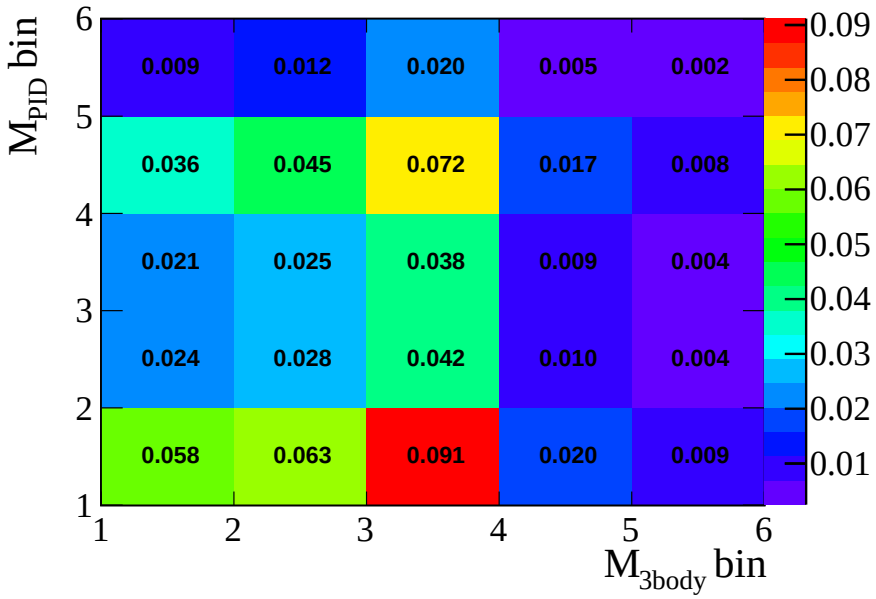


Figure 69: Distribution of $\tau^- \to \mu^-\mu^+\mu^-$ MC events in the $\mathcal{M}_{3body}$ and $\mathcal{M}_{PID}$ bins used in the 2011 analysis.

smoothly across the mass region close to the signal mass window. The most important background channel of this type is found to be $D_s^- \to \eta\,(\mu^+\mu^-\gamma)\,\mu^-\overline{\nu}_\mu$, about 80% of which is removed (see Section 5.3.1.1) by a cut on the dimuon mass. The small remaining background from this process is consistent with the smooth variation in the mass spectra of the other backgrounds in the mass range considered in the fit. As previously mentioned, the dominant contributions to the background from misidentified particles are from $D^+_{(s)} \to K^-K^+\pi^+$ and $D^+_{(s)} \to \pi^-\pi^+\pi^+$ decays. However, these events are strongly peaked in the low likelihood region of $\mathcal{M}_{PID}$ and can be reduced to a negligible level by the removal of the trash bins.

Based on MC simulation, no peaking backgrounds are expected in the $\tau \to p\mu\mu$ analyses. However, as the size of the data sample increases, specific backgrounds may become apparent and as such a detailed study will be performed in the update of the analysis with $3.0\,fb^{-1}$.

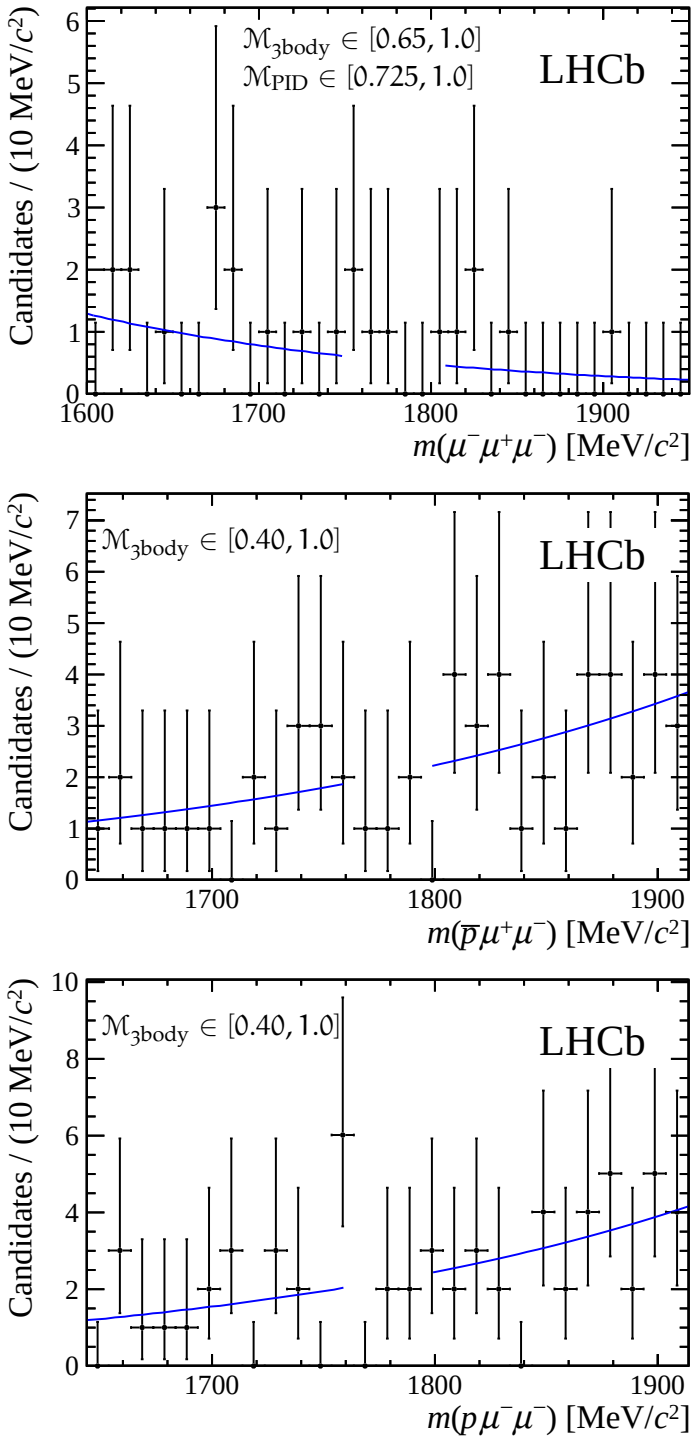


Figure 70: Invariant mass distributions and fits to the mass sidebands in data for $\mu^+\mu^-\mu^-$ candidates in the four merged bins that contain the highest signal probabilities (top), $\overline{p}\mu^+\mu^-$ candidates in the two merged bins with the highest signal probabilities (middle), and $p\mu^-\mu^-$ candidates in the two merged bins with the highest signal probabilities (bottom). The definitions of the merged bins are given in the figure legends.

The expected numbers of background events within the signal region, for each bin in $\mathcal{M}_{3body}$, $\mathcal{M}_{PID}$ (for $\tau^- \to \mu^-\mu^+\mu^-$) and mass, are evaluated by fitting the candidate mass spectra outside of the signal windows to an exponential function using an extended, unbinned maximum likelihood fit. The small differences obtained if the exponential curves are replaced by straight lines are included as systematic uncertainties. For $\tau^- \to \mu^-\mu^+\mu^-$ the data are fitted over the mass range 1600 – 1950 MeV, while for $\tau \to p\mu\mu$ the fitted mass range is 1650 – 1900 MeV, excluding windows around the expected signal mass of ± 30 MeV for $\mu^-\mu^+\mu^-$ and ± 20 MeV for $p\mu\mu$. The resulting fits to the data sidebands for a selection of bins for the three channels are shown in Fig. 70. Detailed fit results are given in Section 5.4.

5.3.3 Normalisation

To measure the signal branching fractions for $\tau^- \to \mu^-\mu^+\mu^-$ and $\tau \to p\mu\mu$ the numbers of observed signal events are normalised to the numbers of events in the $D_s^- \to \phi\,(\mu^+\mu^-)\,\pi^-$ channels using

$$\begin{aligned}
&\mathcal{B}\left(\tau^- \to \mu^-\mu^+\mu^-\right) \\
&\quad = \mathcal{B}\left(D_s^- \to \phi\left(\mu^+\mu^-\right)\pi^-\right) \times \frac{f_{D_s}^{\tau}}{\mathcal{B}\left(D_s^- \to \tau^-\overline{\nu}_\tau\right)} \times \frac{\epsilon_{norm}}{\epsilon_{sig}} \times \frac{N_{sig}}{N_{norm}} \\
&\quad = \alpha \times N_{sig}\,, \qquad (1)
\end{aligned}$$

where α is the overall normalisation factor and N_{sig} and N_{norm} are the number of observed events in the signal and calibration channels respectively. The efficiencies, ϵ, and the branching fractions, $\mathcal{B}(X)$, are discussed in detail in the following section. The value of $f_{D_s}^{\tau}$ is discussed in Section 5.2.1 and is required as $D_s^- \to \tau^-\overline{\nu}_\tau$ does not fully account for the production of τ leptons. The nor-

Channel	Branching fraction	Reference
$D_s^- \to \phi\,(K^+K^-)\,\pi^-$	$(2.27 \pm 0.12) \times 10^{-2}$	[106]
$\phi \to K^+K^-$	$(48.9 \pm 0.5) \times 10^{-2}$	[3]
$\phi \to \mu^+\mu^-$	$(2.87 \pm 0.19) \times 10^{-4}$	[3]
$D_s^- \to \tau^-\overline{\nu}_\tau$	$(5.61 \pm 0.24) \times 10^{-2}$	[98]

Table 20: Branching fractions used in the normalisation factor calculation.

malisation factor for $\tau \to p\mu\mu$ is identical except for the inclusion of a further term

$$\frac{\epsilon_{norm}^{PID}}{\epsilon_{sig}^{PID}},$$

to account for the effect of the PID cuts.

The values of the individual terms in the normalisation factors are given in Table 22.

5.3.3.1 *Input branching fractions*

The branching fractions used in the analysis are listed in Table 20. The branching fraction for $D_s^- \to \phi\,(\mu^+\mu^-)\,\pi^-$ is calculated via

$$\mathcal{B}\left(D_s^- \to \phi\left(\mu^+\mu^-\right)\pi^-\right) = \frac{\mathcal{B}\left(D_s^+ \to \phi\pi^+, \phi \to K^+K^-\right)}{\mathcal{B}\left(\phi \to K^+K^-\right)} \times \mathcal{B}\left(\phi \to \mu^+\mu^-\right),$$

where only the $\phi \to K^+K^-$ component of the $D_s^- \to K^+K^-\pi^-$ decay is considered, as the contribution of non-resonant $D_s^- \to \mu^+\mu^-\pi^-$ events in data is found to be negligible.

5.3.3.2 *Total efficiencies*

The total efficiencies, ϵ_{sig} and ϵ_{norm}, are the product of the efficiencies for generating the signal or normalisation channel of interest, reconstructing all of the

tracks in the final state, selecting the event with the selection cuts listed in Section 5.3.1.1, passing the LHCb trigger, and correcting for differences between data and MC:

$$\frac{\epsilon_{norm}}{\epsilon_{sig}} = \frac{\epsilon^{GEN}_{norm} \times \epsilon^{REC\&SEL}_{norm} \times \epsilon^{TRIG}_{norm}}{\epsilon^{GEN}_{sig} \times \epsilon^{REC\&SEL}_{sig} \times \epsilon^{TRIG}_{sig}} \times c^{track} \times c^{\mu ID} \times c^{\phi} \times c^{\tau} \times c^{trash}.$$

The individual terms are defined in the following sections.

5.3.3.3 *Generation, reconstruction and selection efficiencies*

The generator level efficiencies, ϵ^{GEN}_{sig} and ϵ^{GEN}_{norm}, are the efficiencies for all final state particles in the signal and normalisation channels to pass the LHCb detector acceptance cuts at the generator level. They are determined from the MC production statistics provided by the LHCb simulation group. The uncertainties are purely statistical.

The reconstruction and selection efficiencies, $\epsilon^{REC\&SEL}_{sig}$ and $\epsilon^{REC\&SEL}_{norm}$, are determined using the smeared MC samples, reweighted to give the correct τ production fractions, after the selection criteria detailed in Section 5.3.1.1 (including the $D_s^- \to \eta\,(\mu^+\mu^-\gamma)\,\mu^-\overline{\nu}_\mu$ veto described in Section 5.3.2) have been applied. The systematic uncertainties on $\epsilon^{REC\&SEL}$ come from the reweighting and smearing procedures, where the systematic uncertainty due to the reweighting is determined from the maximum difference between the efficiencies with the weight increased or decreased by the uncertainty. This is found to be 2.3% and 6.0% for $\tau^- \to \mu^-\mu^+\mu^-$ and $\tau \to p\mu\mu$ respectively. The large difference in the effect of the reweighting procedure between $\tau^- \to \mu^-\mu^+\mu^-$ and $\tau \to p\mu\mu$ comes from the additional cuts in the $\tau \to p\mu\mu$ selection changing the relative proportions of the different sources of τ in the MC samples. The difference in the ratio of $\epsilon^{REC\&SEL}$ for the signal and normalisation channels between the samples with different smearings is assigned as the systematic uncertainty due to the smearing procedure, and is found to be 1% for both $\tau^- \to \mu^-\mu^+\mu^-$ and $\tau \to p\mu\mu$.

5.3.3.4 *Trigger efficiencies*

The trigger efficiencies, ϵ_{sig}^{TRIG} and ϵ_{norm}^{TRIG}, are determined using the smeared MC samples. For $\tau \to p\mu\mu$ decays a 7% difference is found between the trigger efficiency of $\tau^- \to \bar{p}\mu^+\mu^-$ and $\tau^- \to p\mu^-\mu^-$ decays due to different HLT2 lines contributing to the total efficiency for each decay, as they require specific combinations of particle charges in the final state.

The systematic error is determined using $B^- \to J/\psi K^-$ events by comparing the difference between the MC trigger efficiency and that determined from data using the 'TISTOS' method. In this method the trigger efficiency is calculated via

$$\epsilon^{TRIG} = \frac{N^{TIS\&TOS}}{N^{TOS}} \times \frac{N^{TRIG}}{N^{TIS}},$$

where N^{TRIG} is the number of events passing the chosen triggers (or trigger) in the data sample, N^{TOS} is the number of these events triggered due to the channel of interest, N^{TIS} is the number of these events that would have triggered without the channel of interest being present, and $N^{TIS\&TOS}$ is the number of these events that would have been triggered both with and without the channel of interest.

The $B^- \to J/\psi K^-$ channel is chosen as, unlike in the $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ channel, there are sufficient statistics to allow for a full TISTOS analysis in kinematic bins. Whilst the behaviour of the trigger in MC is well tested at high track momentum, the kinematical similarities between the muons from B^- and τ^- allow this to also be checked at the low values of momenta found in $\tau^- \to \mu^-\mu^+\mu^-$ and $\tau \to p\mu\mu$ decays.

Comparing the trigger efficiencies in data and MC in this way yields a difference of 4.3% which is used as a conservative error estimate for the ratio $\epsilon_{norm}^{TRIG}/\epsilon_{sig}^{TRIG}$ for both $\tau^- \to \mu^-\mu^+\mu^-$ and $\tau \to p\mu\mu$, shown in Table 23.

An additional check was performed on $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ events, where L0 and HLT1 were ignored and a TISTOS analysis was performed on HLT2 trigger lines only. The data/MC agreement was found to be better than 1%. In addition

a TISTOS analysis was also performed on L0 and HLT1 only, ignoring HLT2, and the results were found to be consistent.

5.3.3.5 *Correction for data/MC differences*

The ratio of the product of all efficiencies is corrected by four factors to account for the differences between data and MC of track reconstruction efficiency, muon ID efficiency, $\phi(1020)$ mass cut efficiency and τ mass cut efficiency. A further correction factor is applied to account for the cuts removing the trash bins in the signal channels only.

The track reconstruction efficiency is determined from MC and corrected using the tracking efficiency map as provided by the LHCb tracking group [107]. The correction factor, c^{track}, depends on the pseudorapidity and the momentum of the tracks and is shown in Table 22, where the first error is the statistical error from the tracking efficiency map [107]. The second error accounts for the systematic uncertainty on the track finding efficiency (0.6%) plus the uncertainty due to the limited knowledge of the hadronic interaction length for the pion from the D_s decay (2.0%.)

The correction factor, $c^{\mu ID}$, to account for the precision with which the MC describes the IsMuon requirement [57], is calculated from the tables determined for the $B_s \to \mu^+\mu^-$ analysis [100]. Maps of efficiency times acceptance in bins of p and p_T for data and MC are combined with the p and p_T spectrum of the muons from reconstructed and selected $\tau^- \to \mu^-\mu^+\mu^-$ and $D_s^- \to \phi\,(\mu^+\mu^-)\,\pi^-$ MC events, and the average values $< \epsilon_\mu \times acc_\mu >_{data}$ and $< \epsilon_\mu \times acc_\mu >_{MC}$ are obtained. The ratio $< \epsilon_\mu \times acc_\mu >_{data} / < \epsilon_\mu \times acc_\mu >_{MC}$ is then used to correct the IsMuon efficiency evaluated using MC for a given channel. The differences between data and MC largely cancel in the ratio, and the error comes from the statistical error in the IsMuon map and takes into account correlations. Whilst using the ratio of data and MC is preferred as it cancels possible biases originating from the method, to determine the systematic error the correction is also calculated using the 2D data map and the measured IsMuon efficiency

for each of the $\tau^- \to \mu^-\mu^+\mu^-$ and $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ MC samples instead of the MC map. This tests the reliability of the method when applied to the $\tau^- \to \mu^-\mu^+\mu^-$ and $\tau \to p\mu\mu$ MC samples. The difference in the correction factor is then assigned as a systematic uncertainty of 2.6% for $\tau^- \to \mu^-\mu^+\mu^-$ and 3.3% for $\tau \to p\mu\mu$.

A correction must also be made to account for the truncation of the $\phi(1020)$ mass distribution in MC at a lower limit of 988 MeV (the KK threshold) and an upper limit of 1085 MeV. To calculate the correction factor, c^ϕ, a simple non-relativistic Breit-Wigner (BW) lineshape is generated according to the phase space available in the $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ decay. Then, if it is assumed that no ϕ is produced below the KK threshold, 1.0% of the distribution is found to be removed by the truncation at 1085 MeV. If the BW lineshape is instead allowed to go down to the $\mu\mu$ threshold, then a total of 2.8% of the distribution is removed by the cuts at 988 and 1085 MeV. In this full range BW, the area outside of the dimuon mass cut corresponds to 6.6% of the total area, whilst in the truncated MC it correspond to 3.8%. Therefore the fits to data and MC remove 6.6% and 3.8% of the total events respectively and the $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ selection efficiency should be corrected by a factor of $^{1.066}/_{1.038}$, or 2.7%. Assuming that the BW does not go below the KK threshold, the region outside of the cut in data corresponds to 4.8% of the total area, and the correction factor reduces to $^{1.048}/_{1.038}$, or 1.0%. Therefore we assume the correction factor is somewhere in this range and assign a value of $(2 \pm 1)\%$ to c^ϕ.

To account for differences in the efficiency of τ mass cuts between MC and data the correction factor, c^τ, is applied. This is calculated as the ratio of the integral of the MC-resolution gaussian functions at ± 30 MeV to the integral of the data-resolution gaussian functions at ± 20 MeV. This takes account of both the effect of reducing the mass window from ± 30 MeV in the selection to ± 20 MeV in data and the measured difference in mass resolutions between data and MC. The widths of the data-resolution gaussian functions, σ_1 and σ_2, are determined by correcting the raw signal MC widths by the ratio of data/MC widths in the

Parameter	$\tau^- \to \mu^-\mu^+\mu^-$	$\tau^- \to \bar{p}\mu^+\mu^-$	$\tau^- \to p\mu^-\mu^-$
Mean (MeV)	1779.18 ± 0.04	1778.0 ± 0.1	1778.0 ± 0.1
σ_1 (MeV)	8.1 ± 0.1	4.5 ± 0.2	4.6 ± 0.2
σ_2 (MeV)	16 ± 1	7.7 ± 0.7	8.0 ± 0.7

Table 21: Corrected gaussian parameters in data for the $\tau^- \to \mu^-\mu^+\mu^-$ and $\tau \to p\mu\mu$ signal channels. The widths of the gaussian functions are denoted by $sigma_1$ and σ_2.

$D_s^- \to \phi(\mu^+\mu^-)\pi^-$ channel. This results in the widths shown in Table 21, which are used as input to the integration. The uncertainty on c^τ represents the statistical uncertainty propagated through the calculation. The central value of the gaussian functions remain unchanged from the MC value of 1779.18 MeV and 1778.0 MeV for $\tau^- \to \mu^-\mu^+\mu^-$ and $\tau \to p\mu\mu$ respectively. The efficiency of the ± 20 MeV mass cut on the smeared and reweighted MC samples is found to be 93.9%, 94.0% and 94.2% for $\tau^- \to \mu^-\mu^+\mu^-$, $\tau^- \to \bar{p}\mu^+\mu^-$ and $\tau^- \to p\mu^-\mu^-$ respectively, after the offline selection and trigger have been applied, and assuming an uncorrected MC mass resolution.

As described in Section 5.3.2 the trash bins are excluded from the limit calculation. For $\tau^- \to \mu^-\mu^+\mu^-$ this corresponds to a cut of $\mathcal{M}_{3body} > -0.48$ and $\mathcal{M}_{PID} > 0.43$, whilst for $\tau \to p\mu\mu$ this corresponds to a cut of $\mathcal{M}_{3body} > -0.05$. The factor c^{trash} is calculated after all selection and trigger requirements have been applied. The uncertainties on the $\mathcal{M}_{3body}$ cuts are derived from the statistical uncertainty on the $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ MC sample. This is added in quadrature to the systematic error (1.3%) of the calibration strategy to give an efficiency of $100.0 \pm 3.5\%$ for a cut of $\mathcal{M}_{3body} > -0.48$ in $\tau^- \to \mu^-\mu^+\mu^-$ and efficiencies of $84.0 \pm 1.4\%$ and $83.3 \pm 1.4\%$ for cuts of $\mathcal{M}_{3body} > -0.05$ on $\tau^- \to \bar{p}\mu^+\mu^-$ and $\tau^- \to p\mu^-\mu^-$ respectively. The uncertainty on the cut at $\mathcal{M}_{PID} > 0.43$ is calculated using tools provided by the LHCb PID group described in Section 5.3.3.6. The statistical and systematic errors are added in quadrature to give an $\mathcal{M}_{PID}$

cut efficiency of $67.6 \pm 1.6\%$. The efficiencies for the removal of the trash bins in $\mathcal{M}_{3body}$ and $\mathcal{M}_{PID}$ are then multiplied together to give a total cut efficiency of $67.6 \pm 3.8\%$. By taking the inverse of this efficiency a value of $c^{trash} = 1.48 \pm 0.08$ is determined. A similar procedure is followed for $\tau \to p\mu\mu$.

5.3.3.6 *PID efficiencies for* $\tau \to p\mu\mu$

The PID efficiencies, ϵ^{PID}_{sig} and ϵ^{PID}_{norm}, for the PID cuts in the $\tau \to p\mu\mu$ analysis are calculated on the smeared MC samples. The same set of tools (the PIDCalib package) used to create the PID plots in Section 5.3.1.2 enables these to be corrected according to the response in data, which in the case of the proton PID cuts is substantially different from that determined in MC. The correction is achieved via the use of calibration samples which can be efficiently reconstructed without the need for PID cuts. For protons a sample of $\Lambda^0 \to p\pi^-$ decays is provided and for muons a sample of $J/\psi \to \mu^+\mu^-$ decays is provided. The MC and data calibration samples are then kinematically binned according to transverse momentum, pseudorapidity and track multiplicity, and the true efficiency for a particular PID cut can be determined. In particular, the MultiTrack tools [108] within the PIDCalib package enable the efficiency of multiple PID cuts on multiple final state tracks to be corrected simultaneously. As the RICH detectors only provide reliable PID in the momentum range 5 – 100 GeV, this method effectively assigns events outside this kinematic region with an efficiency of 0%. This effect is included in the efficiencies in Table 22, and is also applied as a cut on the $\tau \to p\mu\mu$ data samples.

The systematic uncertainty on ϵ^{PID} is a combination of the uncertainties on the muon and proton PID efficiency corrections due to the PIDCalib method, resulting from a loss of information in the binning of the MC and data calibration samples. In several other LHCb analyses this effect is evaluated by considering the resulting difference in PID distributions after the method is performed on a corresponding MC calibration sample instead of the data calibration samples. However, in the case of final states involving protons and muons, MC calibra-

	$\tau^- \to \mu^-\mu^+\mu^-$	$\tau^- \to \bar{p}\mu^+\mu^-$	$\tau^- \to p\mu^-\mu^-$
ϵ_{sig}^{GEN}	17.67 ± 0.04	20.11 ± 0.05	20.13 ± 0.05
ϵ_{norm}^{GEN}		17.81 ± 0.04	
$\epsilon_{sig}^{REC\&SEL}$	$4.76 \pm 0.03 \pm 0.11$	$1.01 \pm 0.02 \pm 0.06$	$1.01 \pm 0.02 \pm 0.06$
$\epsilon_{norm}^{REC\&SEL}$	$4.46 \pm 0.02 \pm 0.12$	$1.33 \pm 0.01 \pm 0.02$	
ϵ_{sig}^{TRIG}	$46.5 \pm 0.4 \pm 0.2$	$42.1 \pm 1.5 \pm 1.8$	$34.8 \pm 1.4 \pm 1.5$
ϵ_{norm}^{TRIG}	$35.0 \pm 0.3 \pm 0.2$	$70.8 \pm 0.9 \pm 3.0$	
c^{track}	$1.000 \pm 0.026 \pm 0.026$	$1.000 \pm 0.020 \pm 0.033$	
$c^{\mu ID}$	$0.9998 \pm 0.0001 \pm 0.0200$	$0.9927 \pm 0.0016 \pm 0.0377$	
c^{ϕ}		0.98 ± 0.01	
c^{τ}	1.07 ± 0.01	1.003 ± 0.002	1.004 ± 0.002
c^{trash}	1.48 ± 0.08	1.19 ± 0.02	1.20 ± 0.02
ϵ_{sig}^{PID}	-	$25.1 \pm 0.8 \pm 1.0$	$25.2 \pm 0.9 \pm 1.0$
ϵ_{norm}^{PID}	-	$35.8 \pm 0.4 \pm 1.0$	

Table 22: Normalisation terms for $\tau^- \to \mu^-\mu^+\mu^-$ (left), $\tau^- \to \bar{p}\mu^+\mu^-$ (centre) and $\tau^- \to p\mu^-\mu^-$ (right). All efficiencies, ϵ, are given in percent. In the cases where two uncertainties are quoted they are statistical and systematic respectively.

tion samples are not yet available. Therefore a systematic uncertainty of 1% per track is assigned as determined in the charm cross-section analysis [93], plus a further 1% to account for differences in binning with respect to this analysis.

5.3.3.7 *Normalisation summary*

The final normalisation factors are given in Table 23, including the values of N_{norm} from Section 5.3.1.2.

	$\tau^- \to \mu^-\mu^+\mu^-$	$\tau^- \to \overline{p}\mu^+\mu^-$	$\tau^- \to p\mu^-\mu^-$
$\mathcal{B}(D_s^- \to \phi(\mu^+\mu^-)\pi^-)$		$(1.33 \pm 0.12) \times 10^{-5}$	
$f_{D_s}^{\tau}$		0.78 ± 0.05	
$\mathcal{B}(D_s^- \to \tau^-\overline{\nu}_\tau)$		0.0561 ± 0.0024	
$\epsilon_{norm}/\epsilon_{sig}$	1.12 ± 0.11	2.27 ± 0.24	2.76 ± 0.30
$\epsilon_{norm}^{PID}/\epsilon_{sig}^{PID}$	-	1.43 ± 0.07	1.42 ± 0.08
N_{norm}	$48\,076 \pm 840$	8145 ± 180	
α	$(4.31 \pm 0.65) \times 10^{-9}$	$(7.4 \pm 1.2) \times 10^{-8}$	$(8.9 \pm 1.5) \times 10^{-8}$

Table 23: Summary of the factors and their combined statistical and systematic uncertainties entering in the normalisation factor for $\tau^- \to \mu^-\mu^+\mu^-$ (left), $\tau^- \to \overline{p}\mu^+\mu^-$ (centre) and $\tau^- \to p\mu^-\mu^-$ (right).

The value of the normalisation factor obtained using an absolute normalisation, α_{abs}, can be calculated as a cross-check of the relative normalisation above. Using the inclusive τ cross-section calculated in Section 5.2.1 the total number of τ produced per fb^{-1} at 7 TeV is found to be $N_\tau = (8.00 \pm 0.99) \times 10^{10}$. The number of expected reconstructed, selected and triggered $\tau^- \to \mu^-\mu^+\mu^-$ candidates per fb^{-1}, $N_{\tau^-\to\mu^-\mu^+\mu^-}$, can then be calculated by multiplying N_τ by all of the efficiencies and corrections factors in Table 22 which are applicable to the signal channel:

$$N_{\tau^-\to\mu^-\mu^+\mu^-} = N_\tau \times \epsilon_{sig}^{GEN} \times \epsilon_{sig}^{REC\&SEL} \times \epsilon_{sig}^{TRIG} \times \frac{1}{c^\tau} \times \frac{1}{c^{trash}}.$$

The tracking and muon ID correction factors are ignored as they are approximately equal to one, and the inverse of the remaining correction factors are required, as the signal channel is in the denominator of Equation 1. This results in a value of $N_{\tau^-\to\mu^-\mu^+\mu^-} = (2.0 \pm 0.3) \times 10^8$ from which the absolute branching ratio for $\tau^- \to \mu^-\mu^+\mu^-$ can be calculated via

$$\mathcal{B}(\tau^- \to \mu^-\mu^+\mu^-) = \frac{N_{sig}}{N_{\tau^-\to\mu^-\mu^+\mu^-}} = \alpha_{abs} \times N_{sig},$$

where N_{sig} is the observed number of $\tau^- \to \mu^-\mu^+\mu^-$ candidates. The absolute normalisation factor is then determined to be

$$\alpha_{abs} = (5.0 \pm 0.8) \times 10^{-9},$$

which is in agreement with the normalisation factor in Table 23.

5.4 RESULTS

Tables 24 and 25 give the expected and observed numbers of candidates for all three channels investigated, in each bin of the likelihood variables. No significant evidence for an excess of events is observed.

5.4.1 The CL_s Method

To calculate the limits on the branching fractions, the CL_s method [109, 110], which was developed for the Higgs searches at LEP, is used. It is also the statistical framework used in the $B_s \to \mu^+\mu^-$ analysis [100]. This modified frequentist approach provides overcoverage with respect to a purely frequentist method but avoids the problems normally present when setting a confidence limit for small signals in the presence of backgrounds.

Using the test statistic of choice (typically the likelihood ratio of the signal plus background to background only PDFs) two estimators, CL_{s+b} and CL_b are calculated, where CL_{s+b} and CL_b give the compatibility of the observed value of the test statistic with the signal-plus-background and background-only hypotheses respectively. CL_s is then calculated as the ratio of the two and gives the level of confidence in the signal hypothesis. The distributions of observed and expected CL_s values are calculated as functions of the assumed branching fractions, as shown in Fig. 71. The branching fraction values that have $CL_s < 1 - CL$ are then excluded at CL confidence. In the case of $\tau^- \to \bar{p}\mu^+\mu^-$

$\mathcal{M}_{\text{PID}}$	$\mathcal{M}_{\text{3body}}$	Expected	Observed
	−0.48 – 0.05	345.0 ± 6.7	409
	0.05 – 0.35	83.8 ± 3.3	68
0.43 – 0.6	0.35 – 0.65	30.2 ± 2.0	35
	0.65 – 0.74	4.3 ± 0.8	2
	0.74 – 1.0	1.4 ± 0.4	1
	−0.48 – 0.05	73.1 ± 3.1	64
	0.05 – 0.35	18.3 ± 1.5	15
0.6 – 0.65	0.35 – 0.65	8.6 ± 1.1	7
	0.65 – 0.74	0.4 ± 0.1	0
	0.74 – 1.0	0.6 ± 0.2	2
	−0.48 – 0.05	45.4 ± 2.4	51
	0.05 – 0.35	11.7 ± 1.2	6
0.65 – 0.725	0.35 – 0.65	5.3 ± 0.8	3
	0.65 – 0.74	0.8 ± 0.2	1
	0.74 – 1.0	0.4 ± 0.1	0
	−0.48 – 0.05	44.5 ± 2.4	62
	0.05 – 0.35	10.6 ± 1.2	13
0.725 – 0.86	0.35 – 0.65	7.3 ± 1.0	7
	0.65 – 0.74	1.0 ± 0.2	2
	0.74 – 1.0	0.4 ± 0.1	0
	−0.48 – 0.05	5.9 ± 0.9	7
	0.05 – 0.35	0.7 ± 0.2	1
0.86 – 1.0	0.35 – 0.65	1.0 ± 0.2	1
	0.65 – 0.74	0.5 ± 0.0	0
	0.74 – 1.0	0.4 ± 0.1	0

Table 24: Expected background candidate yields, with their systematic uncertainties, and observed candidate yields within the τ signal window in the different likelihood bins for the $\tau^- \to \mu^-\mu^+\mu^-$ analysis. The lowest likelihood bins have been excluded from the analysis.

	$\tau^- \to \overline{p}\mu^+\mu^-$		$\tau^- \to p\mu^-\mu^-$	
$\mathcal{M}_{3body}$	Expected	Observed	Expected	Observed
$-0.05 - 0.20$	37.9 ± 0.8	43	41.0 ± 0.9	41
$0.20 - 0.40$	12.6 ± 0.5	8	11.0 ± 0.5	13
$0.40 - 0.70$	6.76 ± 0.37	6	7.64 ± 0.39	10
$0.70 - 1.00$	0.96 ± 0.14	0	0.49 ± 0.12	0

Table 25: Expected background candidate yields, with their systematic uncertainties, and observed candidate yields within the τ mass window in the different likelihood bins for the $\tau \to p\mu\mu$ analysis. The lowest likelihood bin has been excluded from the analysis.

and $\tau^- \to p\mu^-\mu^-$ a small downward fluctuation occurs between the expected and the observed CL_s curves, but these are consistent within 1 σ of the expected results.

The expected limits at 90% (95%) CL for the branching fractions are found to be

$$\begin{aligned}
\mathcal{B}\left(\tau^- \to \mu^-\mu^+\mu^-\right) &< 8.3\ (10.2) \times 10^{-8}, \\
\mathcal{B}\left(\tau^- \to \overline{p}\mu^+\mu^-\right) &< 4.6\ (5.9) \times 10^{-7}, \\
\mathcal{B}\left(\tau^- \to p\mu^-\mu^-\right) &< 5.4\ (6.9) \times 10^{-7},
\end{aligned}$$

while the observed limits at 90% (95%) CL are found to be

$$\begin{aligned}
\mathcal{B}\left(\tau^- \to \mu^-\mu^+\mu^-\right) &< 8.0\ (9.8) \times 10^{-8}, \\
\mathcal{B}\left(\tau^- \to \overline{p}\mu^+\mu^-\right) &< 3.3\ (4.3) \times 10^{-7}, \\
\mathcal{B}\left(\tau^- \to p\mu^-\mu^-\right) &< 4.4\ (5.7) \times 10^{-7}.
\end{aligned}$$

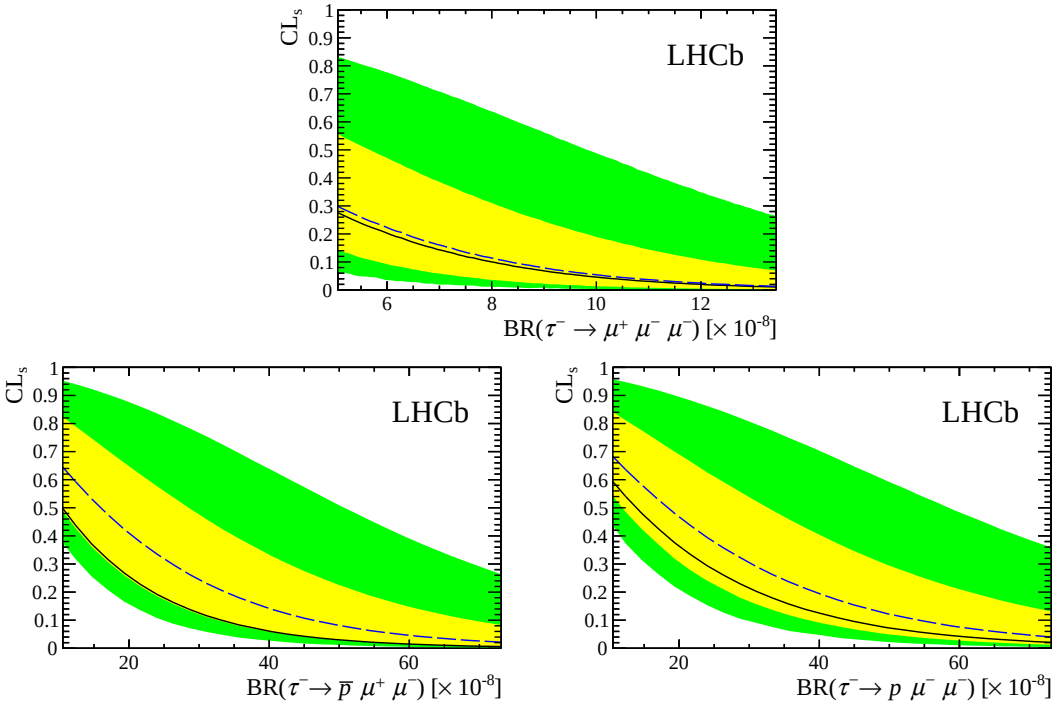


Figure 71: Distribution of CL_s values as functions of the assumed branching fractions, under the hypothesis to observe background events only, for $\tau^- \to \mu^-\mu^+\mu^-$ (top), $\tau^- \to \bar{p}\mu^+\mu^-$ (bottom left) and $\tau^- \to p\mu^-\mu^-$ (bottom right). The dashed lines indicate the expected curves and the solid lines the observed ones. The light (yellow) and dark (green) bands cover the regions of 68% and 95% confidence for the expected limits.

5.4.2 Model Dependence of the Limits

All limits have been calculated using the phase-space model of τ decays. The model dependences of the limits are investigated by studying the variation in the total signal efficiency with dimuon and $p\mu$ mass (in the case of $\tau \to p\mu\mu$). For $\tau^- \to \mu^-\mu^+\mu^-$ it is found to vary by less than 20% over the $\mu^-\mu^-$ mass range and by 10% over the $\mu^+\mu^-$ mass range, whilst for $\tau \to p\mu\mu$ it is found to vary by less than 20% over the dimuon mass range and less than 10% with $p\mu$ mass, suggesting that only a small model dependence exists.

5.4.3 Conclusions

The results from the $1.0\,\text{fb}^{-1}$ dataset represent the first limits on the branching fractions of τ lepton flavour violating and baryon number violating decays obtained at a hadron collider. Whilst the limit on $\tau^- \to \mu^-\mu^+\mu^-$ is compatible with previous results, it is approximately a factor of four above the world's best limit from Belle. However, studies detailed in Appendix A emphasise the improvements on this which are already possible with the addition of 2012 data and improved analysis techniques. The limits on $\tau^- \to \overline{p}\mu^+\mu^-$ and $\tau^- \to p\mu^-\mu^-$ represent the first direct upper limits on the branching fractions for these channels and will continue to be improved upon with the addition of further data.

5.5 SUMMARY AND OUTLOOK

LHCb results on the branching fractions of $\tau^- \to \mu^-\mu^+\mu^-$, $\tau^- \to \overline{p}\mu^+\mu^-$ and $\tau^- \to p\mu^-\mu^-$ have been presented in the previous section. In the following a brief discussion of the possibilities for further improvements on these limits at LHCb and future experiments is given.

5.5.1 Prospects for Beyond LS1

Following the first long shutdown of the LHC, which is currently scheduled to end in 2015, the LHCb experiment will resume taking data. The centre-of-mass energy will increase to 13 TeV, with the aim of reaching the design energy of 14 TeV shortly afterwards, whilst the luminosity will remain relatively stable with respect to 2012 operations. In the subsequent years as much as an additional $7\,\text{fb}^{-1}$ will be recorded, bringing the total dataset to $9 - 10\,\text{fb}^{-1}$ by approximately 2018.

For τ LFV and BNV analyses this will yield a number of benefits. The increase in the size of the dataset should allow the currently existing limits to be improved upon by a factor of about $\sqrt{10}$, assuming that the background increases linearly with the size of the dataset. The cross-sections for charm and beauty production will also increase, and assuming that they scale as determined by PYTHIA (approximately linearly with centre-of-mass energy), the number of τ per inverse femtobarn of data will be a factor of $14/7$ larger.

Assuming that the LHCb trigger efficiency for τ decays remains constant in these conditions, that the background per inverse femtobarn can be kept at a similar level without reducing the signal efficiency, and neglecting the difference in the cross-sections between 13 and 14 TeV, the expected limits can be calculated as

$$\begin{aligned}
\mathcal{B}\left(\tau^- \to \mu^-\mu^+\mu^-\right)_{10\,\mathrm{fb}^{-1}} &= \frac{8\times10^{-8}}{\sqrt{2}\times 8/7+\sqrt{7}\times 14/7} = \; < \; 1.2\times10^{-8},\\
\mathcal{B}\left(\tau^- \to \overline{p}\mu^+\mu^-\right)_{10\,\mathrm{fb}^{-1}} &= \frac{3.3\times10^{-7}}{\sqrt{2}\times 8/7+\sqrt{7}\times 14/7} = \; < \; 4.8\times10^{-8},\\
\mathcal{B}\left(\tau^- \to p\mu^-\mu^-\right)_{10\,\mathrm{fb}^{-1}} &= \frac{4.4\times10^{-7}}{\sqrt{2}\times 8/7+\sqrt{7}\times 14/7} = \; < \; 6.4\times10^{-8}.
\end{aligned}$$

For $\tau^- \to \mu^-\mu^+\mu^-$ this would represent the world's best limit (or evidence for NP) and would help to further constrain some of the models described in Chapter 2. The scaling also assumes no further improvements in the analysis methodology, and based on changes introduced from the $1.0\,\mathrm{fb}^{-1}$ to the $3.0\,\mathrm{fb}^{-1}$ measurements these may well be of the order of 10 – 20%, so even more restrictive limits are a possibility.

For the $\tau \to p\mu\mu$ channels this would be almost an order of magnitude improvement on the existing limits and again this is a rather pessimistic outlook. In particular, as a result of changing the stripping selection, large improvements in the selection efficiency are expected and potential exists for reducing the background level further by studying the contributions of individual processes.

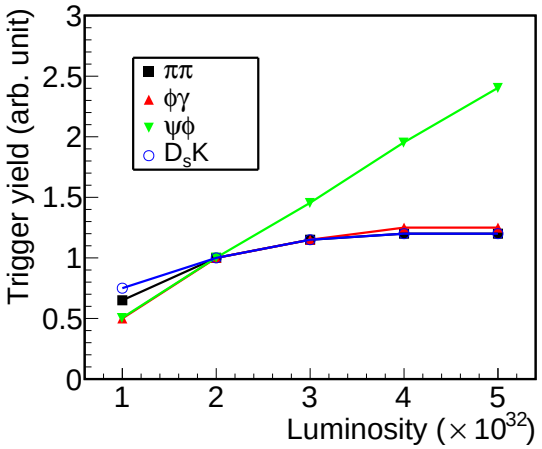


Figure 72: The trigger yield for B-meson decays as a function of luminosity at LHCb. Figure adapted from Ref. [111].

The prospects for other τ channels at LHCb are also good, with a competitive measurement of $\mathcal{B}(\tau^- \to \phi(K^+K^-)\mu^-)$ possible in the coming years.

5.5.2 Future Experiments

The requirement that the LHCb L0 trigger reduces the bandwidth to 1 MHz before the software level effectively limits the potential of LHCb at higher luminosities. As shown in Fig. 72, the trigger yield is already close to saturation for hadronic channels, as the increase in thresholds required to keep the rate of triggered events at an acceptable level causes the trigger efficiency to decrease.

A proposed solution, the LHCb Upgrade [111], would read out the detector at 40 MHz and run at an increased luminosity of $2 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. In order for this to be achieved, the majority of the detector components must be upgraded or replaced, and in doing so a number of improvements are expected, such as in the IP resolution, PID performance and trigger efficiency. Starting in 2019 – 2020 the experiment is expected to collect $50\,\mathrm{fb}^{-1}$ over 10 years, and with the increase in cross-sections at 14 TeV and the improved trigger efficiency, this will correspond to an effective integrated luminosity of approximately 150 times that collected in 2012. For $\tau^- \to \mu^-\mu^+\mu^-$ a sensitivity to branching fractions of the

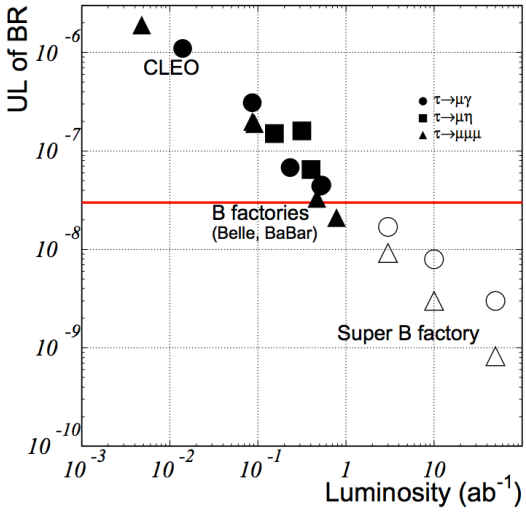


Figure 73: Expected upper limits on the BF of $\tau^- \to \mu^-\mu^+\mu^-$ and $\tau^- \to \mu^-\gamma$ at a Super B factory. Figure from Ref. [113].

order of 10^{-9} is expected and it should be possible to study a wide range of additional τ LFV decays.

Elsewhere, the Belle experiment at KEK, Japan will be replaced by Belle 2 [112] which will perform asymmetric e^+e^- collisions at a luminosity of $8 \times 10^{35}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. Starting from 2016, it is expected to collect $50\,\mathrm{ab}^{-1}$ by 2022, yielding order of magnitude improvements in τ LFV channels, as shown in Fig. 73.

A UPDATES TO SEARCHES FOR LEPTON FLAVOUR VIOLATING TAU DECAYS

In this appendix preliminary results on the search for the $\tau^- \to \mu^-\mu^+\mu^-$ decay in the combined 2011 and 2012 LHCb dataset are presented.

Experience gained during the study of the first $1.0\,\text{fb}^{-1}$ meant that a number of the analysis techniques described in Chapter 5 have been improved upon. These changes are detailed in Section A.1. Where parts of the analysis method are omitted it is assumed they are the same as in the analysis of 2011 data. Preliminary results of the updated search for the $\tau^- \to \mu^-\mu^+\mu^-$ decay are given in Section A.2.

A.1 CHANGES TO ANALYSIS PROCEDURES

A.1.1 Tau Production

As the cross-section for a given process is a function of the centre-of-mass energy, the calculation of the inclusive τ cross-section must be repeated using 8 TeV cross-section measurements. For the $b\bar{b}$ cross-section the value of $298 \pm 36\,\mu\text{b}$ is taken from Ref. [114]. As a measurement of the $c\bar{c}$ cross-section at 8 TeV has yet to be performed, the 7 TeV measurement is scaled by a factor of $8/7$ to take into account the difference in centre-of-mass energy, as validated by calculations in Pythia.

Decay chain	Fraction at 7 TeV (%)	Fraction at 8 TeV (%)
$D_s \to \tau$	68.4 ± 4.4	70.2 ± 4.0
$B_x \to D_s \to \tau$	10.0 ± 2.4	9.3 ± 2.0
Sum $D_s \to \tau$	78.4 ± 3.6	79.6 ± 3.1
$D^+ \to \tau$	4.00 ± 0.72	4.10 ± 0.75
$B_x \to D^+ \to \tau$	0.204 ± 0.048	0.190 ± 0.039
Sum $D^+ \to \tau$	4.20 ± 0.74	4.29 ± 0.76
$B_x \to \tau$	17.3 ± 3.4	16.2 ± 2.8

Table 26: Contributions to $\sigma(pp \to \tau X)$ at 7 and 8 TeV.

The contribution to the τ cross-section from D^+ decays is corrected to account for the theoretical expectation of the BF of $D^+ \to \tau^+\nu_\tau$, by taking the BF of $D^+ \to \mu^+\nu_\mu$, assuming lepton universality, and scaling the BF for differences in helicity suppression and phase space. Using $(3.82 \pm 0.33) \times 10^{-4}$ [99] as the value of the BF for $D^+ \to \mu^+\nu_\mu$, and 2.65 [99] as the helicity suppression and phase space factor, a BF of $(1.012 \pm 0.087) \times 10^{-3}$ is obtained.

These changes lead to an updated 7 TeV cross-section of 80.0 ± 9.4 μb and a larger cross-section of 89 ± 11 μb at 8 TeV due to the corresponding increases in the $b\bar{b}$ and $c\bar{c}$ production cross-sections. The fraction $f_\tau^{D_s}$ then becomes 0.78 ± 0.04 and 0.80 ± 0.03 at 7 and 8 TeV respectively, as shown in Table 26.

A.1.2 MC and Data Samples

The results described in the following are obtained using data collected at a centre-of-mass energy of $\sqrt{s} = 7$ TeV in 2011, corresponding to $1.0\,\text{fb}^{-1}$ of integrated luminosity, and at $\sqrt{s} = 8$ TeV in 2012, corresponding to $2.0\,\text{fb}^{-1}$ of integrated luminosity.

Variable	Signal, norm. and $D_s^- \to \eta\,(\mu^+\mu^-\gamma)\,\mu^-\overline{\nu}_\mu$	Incl. background
p_T	$> 250\,\text{MeV}$	$> 280\,\text{MeV}$
p	$> 2500\,\text{MeV}$	$> 2900\,\text{MeV}$
DOCA	-	$< 0.35\,\text{mm}$
Mass ($\mu\mu$)	-	$< 4500\,\text{MeV}$

Table 27: MC generator level cuts for the 2012 analysis. Cuts for signal, normalisation and $D_s^- \to \eta\,(\mu^+\mu^-\gamma)\,\mu^-\overline{\nu}_\mu$ MC are applied to all final state particles, whilst those for inclusive background MC are applied to muons only.

The MC samples used are updated to take into account the new stripping version and the associated improvements in reconstruction, particle identification and other effects. As this analysis involves the use of both 2011 and 2012 data, in some cases separate MC samples are used for each to take into account the differences in beam energy, trigger, and other effects.

To save CPU time and reduce the number of events which are generated and then subsequently removed by the selection, cuts were applied at the generation level. For the MC samples in the 2011 analysis this was just the requirement that the particles of interest were in the LHCb acceptance. In the 2012 analysis presented here further cuts were added as listed in Table 27.

For the 2012 MC samples the trigger has been emulated during MC production using TCK 0x409f0045. However, as some significant changes are expected between TCKs in several of the trigger lines of interest in the 2012 data sample, a number of additional TCKs are studied, and are taken into account in the calculation of the normalisation factor.

A.1.3 MC Correction

To avoid the need for reweighting, separate MC samples for each subchannel were produced and then added together to create a mixed MC sample with the proportions shown in Table 26, for both $\tau^- \to \mu^-\mu^+\mu^-$ and $D_s^- \to \phi\,(\mu^+\mu^-)\,\pi^-$. As the proportions vary between 7 and 8 TeV this procedure was performed separately for 2011 and 2012 MC.

Due to a more accurate description of the VELO RF foil in 2012 MC, correction of the IP resolution is no longer required for the MC samples used in the 2012 analysis.

A.1.4 Signal and Background Discrimination

A.1.4.1 *Selection*

Due to the high number of ghost tracks observed in the 2011 data, a cut on the probability for a track to be a ghost was introduced into the LHCb stripping by default from 2012 onwards. For most channels of interest this is found to have an extremely small effect on the efficiency and studies of $\tau^- \to \mu^-\mu^+\mu^-$ MC show that the effect of this cut is negligible. The selection for $\tau^- \to \mu^-\mu^+\mu^-$ and $D_s^- \to \phi\,(\mu^+\mu^-)\,\pi^-$ events remains almost identical to the 2011 analysis, with the exception of the track χ^2/ndf cut, which is tightened to be less than three. The PDG masses of the τ, D_s and $\phi(1020)$ are used as the central values for any mass window cuts.

To minimise the effect of differences between TCKs and to further improve the background rejection, a minimal subset of trigger decisions is required. These are determined by evaluating the Punzi FoM for a given trigger using signal MC and the data sidebands. The FoM is normalised to unity for all triggers such that a lower FoM indicates an improved performance over using all of the triggers. This process is repeated for each trigger level separately to give the

Trigger level	2011	2012
L0	L0Muon	
HLT1	Hlt1TrackMuon	
HLT2	Hlt2TriMuonTau or Hlt2CharmSemilepD2HMuMu	Hlt2TriMuonTau

Table 28: Trigger lines used in the 2012 analysis. All lines are required to be TOS.

trigger requirements shown in Table 28. The requirement that all events are TOS is also imposed as this is known to give a more accurate simulation of the true trigger efficiency in MC.

A.1.4.2 *Multivariate classification*

An additional variable is added to the definition of the $\mathcal{M}_{3\text{body}}$ classifier used in the 2011 analysis. In a similar manner to the track isolation, the BDT isolation is used to identify muon candidates which could form a vertex with an additional track in the event. A BDT response, the BDT isolation, is determined for every track-muon pair in the event such that there are three values per track. These three track-muon values are summed together and the maximum sum per event is used as the discriminating variable. The inclusion of this variable in the $\mathcal{M}_{3\text{body}}$ classifier is found to give an increased performance with respect to the definition of the classifier used in the 2011 analysis. The BDT isolation is described in more detail in Ref. [115].

The training of the $\mathcal{M}_{3\text{body}}$ classifier is updated to take advantage of the different production channels of the τ via the 'blending' technique. In this method, classifiers are trained separately on each of the five signal subchannels and then blended together according to their relative contributions in data (shown in Table 26) to give a single response. Various possible classifiers are considered, and the MatrixNet classifier developed by Yandex [116] is found to give the best performance.

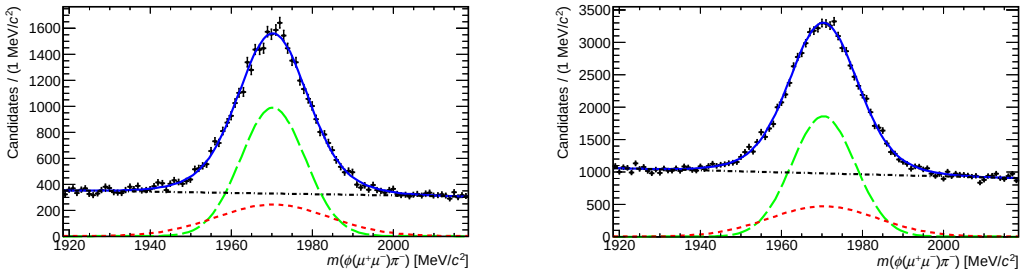


Figure 74: Fit to the invariant mass distribution of $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ events in data after the $\tau^- \to \mu^-\mu^+\mu^-$ selection and trigger for 2011 (left) and 2012 (right). The blue lines show the overall fits, the green and red lines show the two gaussian components of the signal and the black lines show the combinatorial backgrounds.

The binning optimisation procedure is updated to optimise the number of bins and the bin boundaries in both $\mathcal{M}_{3\text{body}}$ and $\mathcal{M}_{\text{PID}}$ simultaneously. The effect of this is to move the region of highest signal efficiency from the central bins into the high likelihood bins of both $\mathcal{M}_{3\text{body}}$ and $\mathcal{M}_{\text{PID}}$. As a result the position of the bin boundaries for both classifiers change, and the number of bins in the $\mathcal{M}_{3\text{body}}$ response increases. As in the 2011 analysis there are six $\mathcal{M}_{\text{PID}}$ bins with the lowest likelihood bin excluded. The binning of the $\mathcal{M}_{3\text{body}}$ response depends on the data sample, such that there are seven bins for the analysis of the 7 TeV sample, and eight bins for the analysis of the 8 TeV sample, with the lowest likelihood bin excluded in both cases.

Fits to $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ candidates in 2011 and 2012 data are shown in Fig. 74 and are summarised in Table 29. Whilst the absolute yield of $D_s^- \to \phi(\mu^+\mu^-)\pi^-$ events in 2011 data is observed to decrease with respect to the 2011 analysis, this is accounted for by the reduced trigger efficiency in the normalisation factor, and the corresponding decrease in background in the $\tau^- \to \mu^-\mu^+\mu^-$ data sample.

Fitted parameter	2011	2012
N_{cal}	$28\,024 \pm 398 \pm 600$	$51\,447 \pm 603 \pm 245$
Mean (MeV)	1970.33 ± 0.09	1970.42 ± 0.07
σ_1 (MeV)	8.0 ± 0.1	7.82 ± 0.08
σ_2 (MeV)	13.9 ± 0.9	13.3 ± 0.6

Table 29: Results of the fits to $D_s^- \to \phi\,(\mu^+\mu^-)\,\pi^-$ data in the 2012 analysis.

A.1.5 Normalisation

The normalisation factors for the 7 TeV and 8 TeV datasets are calculated separately, as several of the terms are dependent on the centre-of-mass energy and the integrated luminosity. The individual terms entering the normalisation factors are summarised in Table 30 and the final values of the normalisation factors are given in 31. The value of α for the 7 TeV data sample is not directly comparable to that determined in Chapter 5 as the change in binning scheme results in a different value for c^{trash}. In addition, the difference in the choice of trigger lines reduces the signal efficiency with respect to the 2011 analysis. However, this effect is compensated by an overall reduction in the level of background in the data sidebands, such that the expected limit for the 7 TeV sample alone is comparable to that obtained in the 2011 analysis.

	7 TeV	8 TeV
ϵ_{sig}^{GEN}	8.99 ± 0.40	9.21 ± 0.35
ϵ_{norm}^{GEN}	11.19 ± 0.34	11.53 ± 0.32
$\epsilon_{sig}^{REC\&SEL}$	9.927 ± 0.028	9.261 ± 0.023
$\epsilon_{norm}^{REC\&SEL}$	7.187 ± 0.022	6.690 ± 0.022
c^{track}	$0.997 \pm 0.009 \pm 0.026$	$0.996 \pm 0.009 \pm 0.026$
$c^{\mu ID}$	$0.9731 \pm 0.0031 \pm 0.0264$	$1.0071 \pm 0.0022 \pm 0.0204$
c^{ϕ}	0.98 ± 0.01	
c^{τ}	1.032 ± 0.006	1.026 ± 0.006
c^{trash}	1.89 ± 0.12	1.96 ± 0.12
ϵ_{sig}^{TRIG}	$35.52 \pm 0.14 \pm 0.14$	$39.3 \pm 1.7 \pm 2.0$
ϵ_{norm}^{TRIG}	$23.42 \pm 0.14 \pm 0.09$	$20.62 \pm 0.76 \pm 1.07$

Table 30: Normalisation terms for $\tau^- \to \mu^-\mu^+\mu^-$ at 7 TeV (left) and 8 TeV (right). All efficiencies, ϵ, are given in percent. In the cases where two uncertainties are quoted they are statistical and systematic respectively.

	7 TeV	8 TeV
$\mathcal{B}(D_s^- \to \phi(\mu^+\mu^-)\pi^-)$	$(1.317 \pm 0.099) \times 10^{-5}$	
$f_{D_s}^{\tau}$	0.78 ± 0.04	0.80 ± 0.03
$\mathcal{B}(D_s^- \to \tau^-\overline{\nu}_{\tau})$	0.0561 ± 0.0024	
$\epsilon_{norm}/\epsilon_{sig}$	0.898 ± 0.060	0.912 ± 0.054
N_{cal}	$28\,162 \pm 434$	$51\,998 \pm 684$
α	$(7.20 \pm 0.98) \times 10^{-9}$	$(3.37 \pm 0.50) \times 10^{-9}$

Table 31: Summary of the factors and their combined statistical and systematic uncertainties entering in the normalisation factor for 7 TeV (left) and 8 TeV (right).

A.2 RESULTS

As of June 2014, the expected limit for the combined 2011 and 2012 data samples is

$$\mathcal{B}\left(\tau^- \to \mu^-\mu^+\mu^-\right) \;<\; 5.0 \times 10^{-8},$$

at 90% CL, and

$$\mathcal{B}\left(\tau^- \to \mu^-\mu^+\mu^-\right) \;<\; 6.1 \times 10^{-8},$$

at 95% CL. The distribution of expected CL_s values as a function of the assumed branching fraction is shown in Fig. 75.

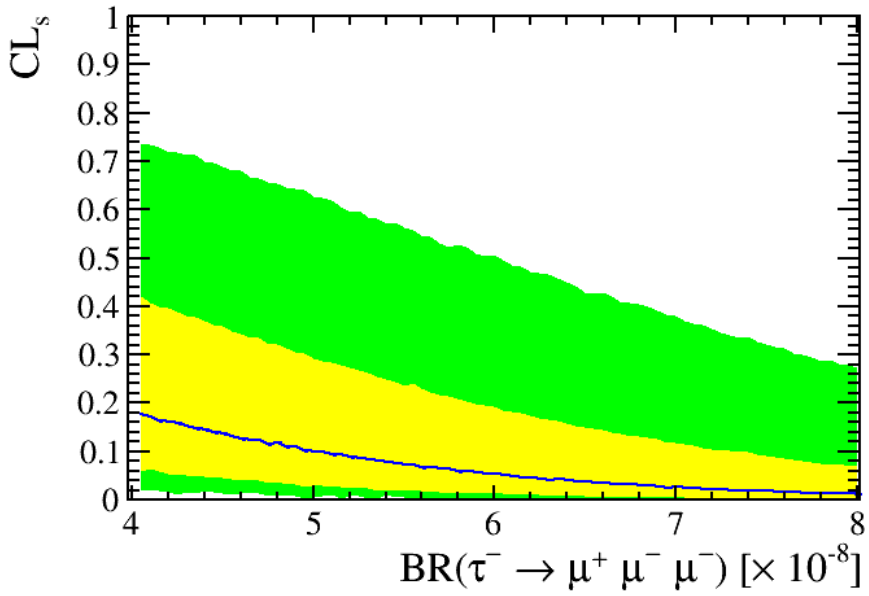


Figure 75: Distribution of CL_s values as a function of the assumed branching fraction for $\tau^- \to \mu^-\mu^+\mu^-$, under the hypothesis to observe background events only. The blue line indicates the expected limit. The light (yellow) and dark (green) bands cover the regions of 68% and 95% confidence for the expected limits.

B ISOLATION DEFINITIONS

B.1 TRACK ISOLATION

The calculation of the track isolation variable begins by searching for tracks that traverse all of the tracking detectors and have an IP significance with respect to any PV of greater than three. If the combination of any of these tracks and one of the muon candidates meet the criteria in Table 32, then the track-muon vertex is designated as a 'good' vertex. The track isolation variable is defined as the number of tracks which form a good vertex that satisfies the equation

$$\frac{|\vec{p_\mu} + \vec{p_{tr}}| \cdot \sin\left(\alpha^{\mu+tr,PV}\right)}{|\vec{p_\mu} + \vec{p_{tr}}| \cdot \sin\left(\alpha^{\mu+tr,PV}\right) + p_{T\,-} + p_{T\,tr}} < 0.6,$$

for each of the three muon candidates in the final state. Here $\vec{p_\mu}$ and $\vec{p_{tr}}$ are the momenta of the muon candidate and the track of interest, $p_{T\,-}$ and $p_{T\,tr}$ are the corresponding transverse momenta, and $\alpha^{\mu+tr,PV}$ is the angle between the

Track-muon variable	Criteria
Separation angle	< 0.27 rad
DOCA	$< 130\,\mu$m
$z_{vertex} - z_{PV}$	> 0.5 cm & < 4 cm
$z_{vertex} - z_{SV}$	> -0.15 cm & < 30 cm

Table 32: Criteria for a good vertex in the track isolation definition.

momentum of the track-muon combination and the line connecting the PV and the track-muon vertex. This definition is taken from Ref. [100].

B.2 τ ISOLATION

The τ isolation is defined as

$$I = \frac{p_T(\tau)}{p_T(\tau) + \Sigma_{\text{tracks}}\, p_T(\text{track})},$$

where $p_T(\tau)$ and $p_T(\text{track})$ are the transverse momenta of the τ candidate and the track respectively. The tracks are defined as all tracks in the event (excluding the muon candidates) with $\sqrt{\delta\eta^2 + \delta\phi^2} < 1$, where the first and second terms are the difference between the τ candidate and the track in pseudorapidity and azimuthal angle, respectively. This definition is taken from Ref. [101].

REFERENCES

[1] E. Noether, *Invariant variation problems*, *Transport Theory and Statistical Physics* **1** (1971) 186–207.

[2] W. A. Bardeen and others., *Deep-inelastic scattering beyond the leading order in asymptotically free gauge theories*, *Phys. Rev. D* **18** (1978) 3998–4017.

[3] Particle Data Group, J. Beringer et al., *Review of Particle Physics*, *Phys. Rev.* **D86** (2012) 010001.

[4] ATLAS Collaboration, G. Aad et al., *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Physics Letters B* **716** (2012), no. 1 1–29.

[5] CMS Collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Physics Letters B* **716** (2012), no. 1 30–61.

[6] S. Weinberg, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) 1264–1266.

[7] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531–533.

[8] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, *Progress of Theoretical Physics* **49** (1973), no. 2 652–657.

[9] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak interactions with lepton-hadron symmetry*, *Phys. Rev. D* **2** (1970) 1285–1292.

[10] Super-Kamiokande Collaboration, Y. Fukuda et al., *Evidence for Oscillation of Atmospheric Neutrinos*, *Phys. Rev. Lett.* **81** (1998) 1562–1567.

[11] M. H. Ahn et al., *Indications of Neutrino Oscillation in a 250 km Long-Baseline Experiment*, *Phys. Rev. Lett.* **90** (2003) 041801.

[12] Q. R. Ahmad et al., *Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory*, *Phys. Rev. Lett.* **89** (2002) 011301.

[13] R. Mohapatra and P. Pal, *Massive Neutrinos in Physics and Astrophysics*. Lecture Notes in Physics Series. World Scientific Publishing Company, Incorporated, 2004.

[14] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the Unified Model of Elementary Particles*, *Progress of Theoretical Physics* **28** (1962), no. 5 870–880.

[15] M. Thompson, *Modern Particle Physics*. Cambridge University Press, 2013.

[16] R. Bernstein and P. S. Cooper, *Charged Lepton Flavor Violation: An Experimenter's Guide*, *Physics Reports* **532** (2013), no. 2 27–64.

[17] Particle Data Group, J. Beringer et al., *Review of Particle Physics*, *Phys. Rev. D* **86** (2012 and 2013 partial update for the 2014 edition) 010001.

[18] R. N. Mohapatra and G. Senjanović, *Neutrino Mass and Spontaneous Parity Nonconservation*, *Phys. Rev. Lett.* **44** (1980) 912–915.

[19] S. Weinberg, *Implications of dynamical symmetry breaking*, *Phys. Rev. D* **13** (1976) 974–996.

[20] S. P. Martin, *A Supersymmetry primer*, `arXiv:hep-ph/9709356`.

[21] J. Ellis et al., *A new parametrization of the seesaw mechanism and applications in supersymmetric models*, *Phys. Rev.* **D66** (2002) 115013.

[22] H.-C. Cheng and I. Low, *TeV symmetry and the little hierarchy problem*, *JHEP* **309** (2003) 051.

[23] R. Barbieri and A. Strumia, *What is the limit on the Higgs mass?*, *Phys.Lett.* **B462** (1999) 144–149.

[24] M. Schmaltz and D. Tucker-Smith, *Little Higgs review*, *Ann.Rev.Nucl.Part.Sci.* **55** (2005) 229–270.

[25] M. Blanke et al., *FCNC processes in the Littlest Higgs Model with T-parity: a 2009 Look*, *Acta Phys. Polon.* **B41** (2010) 657.

[26] J. C. Pati and A. Salam, *Lepton number as the fourth 'color'*, *Phys. Rev. D* **10** (1974) 275–289.

[27] A. Akeroyd et al., *Lepton flavour violating τ decays in the left-right symmetric model*, *Phys. Rev.* **D76** (2007) 013004.

[28] R. Barbier et al., *R-parity-violating supersymmetry*, *Phys. Rept.* **420** (2005) 1–202.

[29] M. Blanke et al., *Charged lepton flavour violation and $(g-2)_\mu$ in the Littlest Higgs Model with T-parity: a clear distinction from supersymmetry*, *JHEP* **5** (2007) 013.

[30] K. Y. Lee, *Lepton flavour violation in a non-universal gauge interaction model*, arXiv:1009.0104.

[31] C.-X. Yue, Y.-M. Zhang, and L.-J. Liu, *Non-universal gauge bosons Z' and lepton flavor-violating tau decays*, *Phys.Lett.* **B547** (2002) 252–256.

[32] A. Hektor, Y. Kajiyama, and K. Kannike, *Muon anomalous magnetic moment and lepton flavor violating tau decay in unparticle physics*, *Phys. Rev.* **D78** (2008) 053008.

[33] A. Dedes et al., *Higgs mediated $B^0_{s,d} \to \mu\tau$, $e\tau$ and $\tau \to 3\mu$, $e\mu\mu$ decays in supersymmetric seesaw models*, *Phys. Lett.* **B549** (2002) 159–169.

[34] A. Masiero, S. K. Vempati, and O. Vives, *SUSY seesaw and FCNC*, *Nucl. Phys. Proc. Suppl.* **137** (2004) 156–168.

[35] G. Cvetic, C. Dib, C. Kim, and J. Kim, *Lepton flavor violation in tau decays*, *Phys.Rev.* **D66** (2002) 034008.

[36] W. J. Marciano, T. Mori, and J. M. Roney, *Charged lepton flavour violation experiments*, *Ann. Rev. Nucl. Part. Sci* **58** (2008) 315–341.

[37] BABAR Collaboration, B. Aubert et al., *Searches for Lepton Flavor Violation in the Decays* $\tau^{\pm} \to e^{\pm}\gamma$ *and* $\tau^{\pm} \to \mu^{\pm}\gamma$, *Phys. Rev. Lett.* **104** (2010) 021802.

[38] Belle Collaboration, K. Hayasaka et al., *Search for lepton-flavor-violating* τ *decays into three leptons with 719 million produced* $\tau^{+}\tau^{-}$ *pairs*, *Phys.Lett.* **B687** (2010) 139–143.

[39] MEG Collaboration, J. Adam et al., *New Constraint on the Existence of the* $\mu^{+} \to e^{+}\gamma$ *Decay*, *Phys. Rev. Lett.* **110** (2013) 201801.

[40] SINDRUM Collaboration, U. Bellgardt et al., *Search for the Decay* $\mu^{+} \to e^{+}e^{+}e^{-}$, *Nucl.Phys.* **B299** (1988) 1.

[41] SINDRUM II Collaboration, W. H. Bertl et al., *A Search for muon to electron conversion in muonic gold*, *Eur.Phys.J.* **C47** (2006) 337–346.

[42] The Mu2e Collaboration, F. Cervelli et al., *The Mu2e Experiment at Fermilab*, *Journal of Physics: Conference Series* **335** (2011), no. 1 012073.

[43] A. D. Sakharov, *Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe*, *Soviet Physics Uspekhi* **34** (1991), no. 5 392.

[44] A. Riotto and M. Trodden, *Recent progress in baryogenesis*, *Ann.Rev.Nucl.Part.Sci.* **49** (1999) 35–75.

[45] M. Raidal et al., *Flavour physics of leptons and dipole moments*, *Eur.Phys.J.* **C57** (2008) 13–182.

[46] W. J. Marciano, *Tau physics 1994 — a theoretical perspective*, *Nuclear Physics B - Proceedings Supplements* **40** (1995), no. 1-3 3–15.

[47] BaBar Collaboration, G. Lafferty, *Lepton Flavour Violation and Baryon Number Non-conservation in $\tau \to \Lambda + h$*, *Nucl.Phys.Proc.Suppl.* **169** (2007) 186–191.

[48] W.-S. Hou, M. Nagashima, and A. Soddu, *Baryon number violation involving higher generations*, *Phys.Rev.* **D72** (2005) 095001.

[49] LHC at Home, *A view of the LHC*, http://lhcathome.web.cern.ch/sixtrack/lhc-unique-tool.

[50] LHCb Collaboration, A. Alves et al., *The LHCb detector at the LHC*, *JINST* **3** (2008) S08005.

[51] LHCb Collaboration, *LHCb layout*, http://cds.cern.ch/record/1087860.

[52] DELPHI Collaboration, P. Abreu et al., *Performance of the DELPHI detector*, *Nucl. Instr. Meth. A* **378** (1996) 57–100.

[53] K. Österberg, *Performance of the vertex detectors at LEP2*, *Nucl. Instr. Meth. A* **435** (1999) 1–8.

[54] LHCb VELO Group, A. Affolder et al., *Performance of the LHCb Vertex Locator*, *To be submitted to JINST* (2014).

[55] LHCb RICH Group, M. Adinolfi et al., *Performance of the LHCb RICH detector at the LHC*, *Eur. Phys. J. C* **73** (2012) 2431.

[56] LHCb MuonID Group, F. Archilli et al., *Performance of the Muon Identification at LHCb*, *J. Instrum.* **8** (2013) P10020.

[57] G. Lanfranchi et al., *The muon identification procedure of the LHCb experiment for the first data*, Tech. Rep. CERN-LHCb-PUB-2009-013, 2009.

[58] LHCb Collaboration, R. Aaij et al., *Absolute luminosity measurements with the LHCb detector at the LHC*, *Journal of Instrumentation* **7** (2012), no. 1 P01010.

[59] R. Aaij et al., *The LHCb Trigger and its Performance in 2011*, *J. Instrum.* **8** (2012) P04022.

[60] LHCb Collaboration, *Trigger schemes*, `http://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html`.

[61] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth.* **A462** (2001) 152–155.

[62] P. Golonka and Z. Was, *PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays*, *Eur.Phys.J.* **C45** (2006) 97–107.

[63] GEANT4 Collaboration, J. Allison et al., *Geant4 developments and applications*, *IEEE Trans.Nucl.Sci.* **53** (2006) 270.

[64] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: a simulation toolkit*, *Nucl. Instrum. Meth.* **A506** (2003) 250.

[65] M. Clemencic et al., *The LHCb simulation application, Gauss: design, evolution and experience*, *J. of Phys: Conf. Ser.* **331** (2011) 032023.

[66] W. G. Pfann, *Zone Melting*. Wiley, J. Wiley & Sons, New York, 1966.

[67] K. Nakamura et al., *Formation process of grown-in defects in Czochralski grown silicon crystals*, *Journal of Crystal Growth* **180** (1997), no. 1 61–72.

[68] C. Kittel, *Introduction to Solid State Physics*. Wiley, J. Wiley & Sons, New York, 2004.

[69] A. Chilingarov, *Generation current temperature scaling*, Tech. Rep. EP-Tech-Note-2013-001, 2013.

[70] A. Hickling et al., *Use of IT (current vs temperature) scans to study radiation damage in the LHCb VELO*, Tech. Rep. CERN-LHCb-PUB-2011-021, 2011.

[71] G. Casse, *The effect of hadron irradiation on the electrical properties of particle detectors made from various silicon materials*. PhD thesis, Universite Joseph Fourier-Grenoble 1, 1998.

[72] R. C. Alig and S. Bloom, *Electron-Hole-Pair Creation Energies in Semiconductors*, *Phys. Rev. Lett.* **35** (1975) 1522–1525.

[73] ROSE Collaboration, C. Goessling et al., *Influence of Surface Damage on Highly Segmeneted Silicon Detectors*, Tech. Rep. ROSE/TN/2000-05, 2000.

[74] F. Hönniger, *Radiation Damage in Silicon*. PhD thesis, Hamburg University, 2007.

[75] G. Barbottin and A. Vapaille, *Instabilities in Silicon Devices*. Elsevier Science B. V., 1996.

[76] A. Gureja et al., *Use of IV (current vs voltage) scans to track radiation damage in the LHCb VELO*, Tech. Rep. CERN-LHCb-PUB-2011-020, 2011.

[77] LHCb VELO group, *Approved conference plots*, https://lbtwiki.cern.ch/bin/view/VELO/VELOConferencePlots.

[78] R. Wunstorf, *Systematische Untersuchungen zur Strahlenresistenz von Silizium-Detektoren fuer die Verwendung in Hochenergiephysik-Experimenten*. PhD thesis, Hamburg University, 1992.

[79] K. V. Ravi, *Imperfections and Impurities in Semiconductor Silicon*. Wiley-Interscience, J. Wiley & Sons, New York, 1981.

[80] M. Moll, E. Fretwurst, and G. Lindström, *Investigation on the improved radiation hardness of silicon detectors with high oxygen concentration*, *Nucl. Instr. Meth. A* **439** (2000), no. 2-3 282–292.

[81] G. Lindström et al., *Developments for radiation hard silicon detectors by defect engineering-results by the CERN RD48 (ROSE) collaboration*, *Nucl. Instr. Meth. A* **465** (2001), no. 1 60–69.

[82] A. D. Webber, *Radiation damage studies in the LHCb VELO detector and measurement of the flavour-specific asymmetry in semileptonic B-decays*. PhD thesis, The University of Manchester, 2013. CERN-THESIS-2013-117.

[83] LHCb VELO Group, A. Affolder et al., *Radiation damage in the LHCb vertex locator*, *Journal of Instrumentation* **8** (2013), no. 8 P08002.

[84] P. R. Turner, *Velo module production - sensor testing*, Tech. Rep. LHCb-2007-072, 2007.

[85] M. Moll, *Radiation Damage in Silicon Particle Detectors*. PhD thesis, Hamburg University, 1999.

[86] CMS Tracker Collaboration, C. Barth, *Evolution of silicon sensor characteristics of the CMS silicon strip tracker*, *Nucl. Instr. Meth. A* **699** (2013), no. 0 178–183.

[87] ATLAS Pixel Collaboration, T. Lari, *Test beam results of ATLAS Pixel sensors*, arXiv:hep-ex/0210045.

[88] G. Kramberger et al., *Charge collection studies on custom silicon detectors irradiated up to* $1.6 \times 10^{17} n_{eq}/cm^2$, *Journal of Instrumentation* **8** (2013), no. 8 P08004.

[89] LHCb Collaboration, R. Aaij et al., *Searches for violation of lepton flavour and baryon number in tau lepton decays at LHCb*, *Physics Letters B* **724** (2013), no. 1-3 36–45.

[90] Heavy Flavor Averaging Group, Y. Amhis et al., *Averages of b-hadron, c-hadron, and tau-lepton properties as of early 2012*, arXiv:1207.1158.

[91] BaBar Collaboration, P. del Amo Sanchez et al., *Searches for the baryon- and lepton-number violating decays* $B^0 \to \Lambda_c^+ \ell^-$, $B^- \to \Lambda \ell^-$, *and* $B^- \to \overline{\Lambda} \ell^-$, *Phys.Rev.* **D83** (2011) 091101.

[92] LHCb Collaboration, R. Aaij et al., *Measurement of* $\sigma(pp \to b\bar{b}X)$ *at* $\sqrt{s} = 7$ *TeV in the forward region*, *Phys.Lett.* **B694** (2010) 209–216.

[93] LHCb Collaboration, R. Aaij et al., *Prompt charm production in* pp *collisions at* $\sqrt{s} = 7$ *TeV*, *Nucl. Phys.* **B871** (2013) 1.

[94] LHCb Collaboration, R. Aaij et al., *Measurement of* $\psi(2S)$ *meson production in* pp *collisions at* $\sqrt{s} = 7$ *TeV*, *Eur. Phys. J. C* **72** (2012) 2100.

[95] LHCb Collaboration, R. Aaij et al., *Measurement of* Υ *production in* pp *collisions at* $\sqrt{s} = 7$ *TeV*, *Eur. Phys. J. C* **72** (2012) 2025.

[96] LHCb Collaboration, R. Aaij et al., *Inclusive* W *and* Z *production in the forward region at* $\sqrt{s} = 7$ *TeV*, *J. High Energy Phys.* **6** (2012) 058.

[97] LHCb Collaboration, LHCb Collaboration, *Inclusive low mass Drell-Yan production in the forward region at* $\sqrt{s} = 7$ *TeV*, Tech. Rep. LHCb-CONF-2012-013, 2012.

[98] J. L. Rosner and S. Stone, *Leptonic decays of charged pseudoscalar mesons - 2012*, arXiv:1201.2401.

[99] CLEO Collaboration, P. Rubin et al., *An Investigation of* $D^+ \to \tau^+\nu_\tau$, *Phys.Rev.* **D73** (2006) 112005.

[100] LHCb Collaboration, R. Aaij et al., *Measurement of the* $B_s^0 \to \mu^+\mu^-$ *Branching Fraction and Search for* $B^0 \to \mu^+\mu^-$ *Decays at the LHCb Experiment*, *Phys. Rev. Lett.* **111** (2013) 101805.

[101] CDF Collaboration, A. Abulencia et al., *Search for* $B_s \to \mu^+\mu^-$ *and* $B_d \to \mu^+\mu^-$ *Decays in* $p\bar{p}$ *Collisions with CDF II*, *Phys.Rev.Lett.* **95** (2005) 221805.

[102] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*. Wadsworth International Group, Belmont, California, USA, 1984.

[103] R. E. Schapire and Y. Freund, *A Decision-Theoretic Generalization of On-line Learning and an Application to Boosting*, *Jour. Comp. and Syst. Sc.* **55** (1997) 119.

[104] A. Hoecker et al., *TMVA: Toolkit for multivariate data analysis*, *PoS* **ACAT** (2007) 040.

[105] G. Punzi, *Sensitivity of searches for new signals and its optimization*, in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology* (L. Lyons, R. Mount, and R. Reitmeyer, eds.), p. 79, 2003. arXiv:physics/0308063.

[106] BaBar Collaboration, P. del Amo Sanchez et al., *Dalitz plot analysis of* $D_s^+ \to K^+K^-\pi^+$, *Phys.Rev.* **D83** (2011) 052001.

[107] A. Jaeger et al., *Measurement of the track finding efficiency*, Tech. Rep. CERN-LHCb-PUB-2011-025, 2012.

[108] LHCb Collaboration, R. Aaij et al., *Measurement of* CP *observables in* $B^0 \to DK^{*0}$ *with* $D \to K^+K^-$, *J. High Energy Phys.* **3** (2012) 067.

[109] A. L. Read, *Presentation of search results: the CL(s) technique*, *J. Phys.* **G28** (2002) 2693–2704.

[110] T. Junk, *Confidence level computation for combining searches with small statistics*, *Nucl. Instrum. Meth.* **A434** (1999) 435–443.

[111] LHCb Collaboration, *Letter of intent for the LHCb Upgrade*, Tech. Rep. LHCC-2011-001, 2011.

[112] T. Abe et al., *Belle II Technical Design Report*, arXiv:1011.0352.

[113] T. Aushev et al., *Physics at Super B Factory*, arXiv:1002.5012.

[114] LHCb Collaboration, R. Aaij et al., *Production of* J/ψ *and* Υ *mesons in* pp *collisions at* $\sqrt{s} = 8$ *TeV*, *Journal of High Energy Physics* **2013** (2013), no. 6 .

[115] L. Gavardi, *Search for lepton flavour violation in τ decays at the LHCb experiment*. PhD thesis, Milan Bicocca University, 2013. CERN-THESIS-2013-259.

[116] Yandex, *MatrixNet: New Level of Search Quality*, http://company.yandex.com/technologies/matrixnet.xml, 2014.