

Charged Lepton Flavour Violation and Optimised 4D Tracking at LHCb

By

Daniel J. D. Thompson

A thesis submitted to the University of Birmingham
for the degree of *Doctor Of Philosophy*



UNIVERSITY OF
BIRMINGHAM

Particle Physics Research Group
School of Physics and Astronomy
College of Engineering and Physical Sciences
University of Birmingham
July 2024

Abstract

The first estimated upper limit for the branching fraction of the forbidden decay $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ is presented. This is calculated using data collected by the LHCb experiment in 2011–12 and 2016–18, corresponding to an integrated luminosity of 8.7 fb^{-1} . This decay is forbidden in the Standard Model of particle physics and hence this constitutes a search for Charged Lepton Flavour Violation (cLFV) and an observation of signal would be unequivocal evidence for New Physics.

While any hints or evidence of signal would transform the analysis outlook, the search for $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ remains blinded and the presented results are based on a background-only dataset assuming no signal. An estimated upper limit of

$$\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp) \lesssim 2.8 \times 10^{-8} \text{ @ 95\% confidence level}$$

is evaluated, from an overall low-background analysis strategy. This limit is competitive with other cLFV searches in the decays of b -hadrons.

During Long-Shutdown 4 of the LHC, a major upgrade to the LHCb experiment is planned, to enable the detector to derive maximum benefit from the High-Luminosity LHC. An investigation of a potential sensor technology for the VELO sub-detector is presented, concluding that the LGADs tested are unable to withstand the high-radiation environment of the HL-LHC. A flexible approach to VELO Upgrade II simulation is also introduced to develop and evaluate performance of candidate designs. This lead to the conclusion that a detector with fast timing, improved spatial resolution and a larger inner radius is essential to meet the requirements of LHCb Upgrade II. Finally, a tool to allow performance optimisation of the entire LHCb Upgrade II detector has been developed, and used for the first investigations of inter-subdetector performance with LHCb core software.

Acknowledgements

Everyone tells you that a thesis is a marathon not a sprint, but unless my ignorance for long-distance running exceeds itself, marathons are a steady, isolated undertaking, keeping a constant pace until the end nears and generally with only your own thoughts driving you forwards. This thesis has been anything but steady and if the writing of these 200 pages had been driven by my thoughts alone I would warn against proceeding. Instead I am lucky enough to have had countless people around me supporting and cheering me on throughout my PhD and I truly could not have become the scientist (I guess I can say that now?) I am today without you all. Too many names to mention but many a name to highlight...

First, a tremendous thank you to Nigel, truly the backbone of my PhD. Your unwavering support over the last four years has kept me on the right path, assuring me that I was progressing and succeeding when the many troughs of a PhD hit. Not to mention the vast reservoir of knowledge you have shared with me, never failing to help, no matter how little context I may have provided with a question. I also extend my supervision gratitude to Niladri and Dan. Niladri, thanks for your advice, guidance and instruction on how to grow as a particle physicist and for always bringing a smile and a laugh to our analysis work and meetings. To Dan, I don't think I was aware of it at the time but your joining of the group came at a crucial time for me and brought a new motivation that has shown me a whole new side of particle physics and has inspired me to continue along its career path, thank you.

I want to thank my fellow students, now friends, at Birmingham, your companionship throughout the thesis and entire PhD process have definitely kept me (relatively) sane. To Adrien, Mihaela and Jon, your friendship alongside me over the four years has been essential, you will all make brilliant PostDocs and I can't wait to visit your exciting new institutes. To Paul, your advice, guidance and paving of the way for my analysis has been key (+ special thanks for the clarity of your documentation!). To Felicia, thank you for your recent help and continuous support while I've been writing, I am excited to work alongside you as a PostDoc very soon! To Marcus and Ellie, I never could have asked for better new students to start during my time. I hope I have imparted some knowledge somewhere in my frequent chatter and I am certain you will continue to excel and succeed throughout the remainder of your PhDs.

A quick mention must go to those that have helped directly with the work described in this thesis! In order of appearance: thank you to Chandler and Ryan for being great MPhys students and designing a highly performant MVA; thank you to the $A_b^0 \rightarrow A\ell^+\ell^-$ analysis team for all the advice and guidance in the early stages of my analysis; a huge

thanks again to Jon for designing and essentially teaching a hardware-novice how to hardware; and thank you to Mark for the technical work on our Performance Optimisation tool, proving anything is possible with software! A thank you must be extended to the LHCb VELO Upgrade II simulation group, in particular Kazu, Paula, Tim, Tom and Victor, your advice and support helped open up a new avenue of my PhD that I have very much enjoyed.

I want to most of all thank my family and friends that have been the most excellent teammates in this test-match of a PhD (I've given up on the marathon analogy). To Aidan, Alex and Niall, our shared LTA + PhD experiences should guarantee our paths cross at least once more, right? To Mel, Simon and Ollie, thanks for the joy your family have brought me. To Matt, Jess, Meg and Alex thank you for enduring and engaging with my occasionally mindless ramblings, your distractions from the deepest moments of writing have pulled me through and you can finally stop telling me to "just submit it". To my siblings, thank you to Meg for showing me that undertaking a PhD is possible, to Harry for always appreciating my musings on the railways and to Alfie for reminding us that there is much more to life than science, keep lighting up the stage! Finally, I am eternally grateful to my parents that have always and continue to inspire and encourage me, nothing I do would be possible without your love and support.

And the final whistle blows! (I'm pushing it now) My ultimate thanks and appreciation undoubtedly must go to my wonderful ~~partner~~ (edit: fiancée) Ella. There is not a second of this journey I can imagine without you, every day you have been there for me through the lows and the extra-lows, always reminding me that the highs make it all worth it. Thank you for everything.

Declaration of Author's contribution

The large majority of the analysis detailed in Chapters 3–5 was performed by myself apart from the the L0 trigger efficiency correction maps in Section 4.2.5 that were generated by Niladri Sahoo and the initial prototype of the MVA that was implemented by MPhys Project students at the University of Birmingham, Chandler Kenworthy and Ryan Hull.

The work detailed in Section 6.3 was carried out in the BILPA lab at the University of Birmingham by myself, Naomi Cooke, Niladri Sahoo and Marcus Madurai. The experimental setups were initially designed and built by Jonathan Mulvey. The initial simulation reconstruction framework detailed in Section 6.4.2.1 was implemented by Tim Evans and Laurent Dufour to produce results for the LHCb UII FTDR, but was adapted and extended by myself to the form detailed throughout Section 6.4. The work detailed in Section 6.5 was implemented with Mark Slater, who designed the backend for the LHCbPRGrid implementation and the Custom Stack builder, where as I prepared the dashboard, job config and submission.

Contents

1	Introduction and Theoretical Foundations	1
1.1	The Standard Model	3
1.1.1	Leptons, Quarks and Hadrons	4
1.2	Flavour Physics	5
1.2.1	Lepton Flavour Universality	8
1.3	Searching for New Physics with Lepton Flavour Violation	10
2	The LHCb Detector - Run 1 & 2	15
2.1	The Large Hadron Collider	15
2.1.1	The LHC Accelerator	15
2.1.2	The LHCb Experiment at the LHC	17
2.2	LHCb Tracking	18
2.2.1	VERtex LOcator in Run 1 and 2	18
2.2.2	Tracking Stations and Magnet	19
2.3	LHCb Particle Identification	20
2.4	Data Acquisition and Offline Reconstruction	24
2.4.1	Trigger in Run 1 and 2	24
2.4.2	Offline Processing	25
2.4.3	Simulation at LHCb	26
3	A Search for Charged Lepton Flavour Violation in $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ decays at LHCb	29
3.1	Motivation for $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ Lepton Flavour Violation (LFV) Search	29
3.2	Analysis Strategy	32

3.2.1	Data and Simulation Samples	34
3.3	Summary of Analysis Components	36
4	Event Selection and Corrections to Simulation	39
4.1	Event Selection	39
4.1.1	Stripping Selection	39
4.1.2	Truth Matching of the Simulation	40
4.1.3	Fiducial Selection	42
4.1.4	Trigger Selection	42
4.1.5	Loose Pre-Selection	46
4.1.6	Semileptonic Removal	49
4.2	Corrections to Simulation	56
4.2.1	Λ_b^0 Lifetime Correction	59
4.2.2	Modelling Correction for $\Lambda_b^0 \rightarrow pK^- J/\psi$	59
4.2.3	PID Efficiency Correction	60
4.2.4	Track Efficiency Correction	63
4.2.5	L0 Trigger Efficiency Correction	64
4.2.6	Λ_b^0 Production Kinematics Correction	65
4.2.7	Post-Correction Data-MC Alignment	67
4.3	Multi-Variate Algorithm Selection	68
4.3.1	Keras Sequential Neural Network	70
4.3.2	Variable Study	72
4.3.3	Performance, Validation and Optimisation	78
4.4	Particle Identification Selection and MVA-PID Optimisation	84
4.4.1	PID Selection Weights	86
4.4.2	MVA-PID Optimisation	87
4.5	Single Candidate Selection	90
4.6	Determination of Selection Efficiency	91
4.6.1	Efficiency Combinations across an Experimental Run	92
4.6.2	Total Efficiency of Signal and Control Modes	93

4.6.3	Total Efficiency of Background Modes	96
4.7	Determination of Systematic Uncertainty on Efficiency Ratio	98
5	Evaluation of Backgrounds and Yield Determination	101
5.1	Potential Backgrounds for $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$	102
5.1.1	Estimated Yield for $\Lambda_b^0 \rightarrow D^0 p \ell^- \bar{\nu}$ Backgrounds	104
5.2	Exclusive fits to backgrounds for $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$	105
5.2.1	$\Lambda_b^0 \rightarrow p K^- K^+ K^-$	105
5.2.2	$\Lambda_b^0 \rightarrow D^0 (\rightarrow K^- \ell^+ \nu) p \ell^- \bar{\nu}$	106
5.3	Mass Fits of the Control Mode	107
5.3.1	MC Signal Component	109
5.3.2	Fits to Data without Exclusive Background Components	109
5.3.3	Fits to the Data with Exclusive Background Components	112
5.3.4	Fit Validation - M_{pK} Spectrum	119
5.4	Mass Fits of the Signal Mode	122
5.4.1	MC Signal Component	123
5.4.2	Constraining the Combinatorial Shape with μe SS Data	124
5.4.3	Blinded Background fits in $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$	128
5.5	Determination of an Estimated Upper Limit for $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp)$	133
5.5.1	Single-Event Sensitivity	133
5.5.2	Blinded Upper Limit	135
5.5.3	Current Status and Future of Analysis	136
6	The LHCb Upgrade II	139
6.1	LHCb Upgrade II: A Flavour Physics Experiment in a High Luminosity Environment	139
6.1.1	High Luminosity LHC and LHCb UII Physics Potential	140
6.1.2	Upgrades to Tracking Detectors	144
6.1.3	Upgrades to Particle Identification Detectors and Data Processing	147
6.2	Future Vertex Detector - Sensor Technology	148
6.2.1	Characteristics for an Upgraded VELO Pixel Sensor	149

6.2.2	Sensor Technology Options	151
6.3	Feasibility Investigation of Micron Semiconductor LGADs	154
6.3.1	LGAD Design, Experimental Set-up and Irradiation	155
6.3.2	Current-Voltage (IV) Characterisation	156
6.3.3	Capacitance-Voltage (CV) Characterisation	159
6.3.4	Gain Characterisation	161
6.3.5	Feasibility for VELO Upgrade II	163
6.4	Future Vertex Detector - Development and Simulation	164
6.4.1	VELO UII Scenarios	165
6.4.2	VELO UII Simulation	167
6.5	Performance Optimisation Tool for LHCb Upgrade II	186
6.5.1	LHCb Performance and Regression System	186
6.5.2	LHCbPR Development for an Upgrade II Use Case	187
6.5.3	Results - VELO UII Scenario X	191
6.5.4	Results - UII Tracking System	192
7	Conclusions	195
	References	199
	Appendices	215
	Appendix A Additional Detail for Pre-Selection	217
A.1	Trigger Selection Options Efficiency Summary	217
	Appendix B Further Material for the Multivariate Analysis	219
	Appendix C Complete Tables for the Background Study	223
	Appendix D Additional for N_{Signal} and $N_{Control}$ Extraction	229
	Appendix E Additional Content for Performance Optimisation Tool	237

List of Figures

1.1	Feynman diagram representation of the Flavour-Changing-Neutral-Current (FCNC) $s \rightarrow d\mu^+\mu^-$ transition.	6
1.2	FCNC $b \rightarrow s\ell^{(\prime)+}\ell^-$ transition diagram.	8
1.3	LHCb R_H measurement summary.	10
1.4	Feynman diagrams of Charged Lepton Flavour Violation (cLFV) decay $\Lambda_b^0 \rightarrow \Lambda(1520)\ell^+\ell'^-$	12
1.5	Summary of $b \rightarrow s\ell^-\ell'^+$ searches at LHCb.	13
2.1	CERN accelerator complex during Run 2 of the LHC	16
2.2	Schematic of the LHCb detector during Run 1 and 2 of the LHC.	17
2.3	Track types and reconstruction effects in the LHCb detector.	20
2.4	Particle Identification (PID) system at LHCb.	21
2.5	Leptonic PID performance at LHCb.	23
2.6	The LHCb data processing flow-diagrams.	24
3.1	Differential $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-)$ analysis results.	30
3.2	M_{pK} distribution in data from the Run 1 LHCb dataset.	32
3.3	Simulation types for the $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ signal mode.	35
4.1	Reconstructed $M(\Lambda_b^0)$ in $\Lambda_b^0 \rightarrow \Lambda(1520)(\rightarrow pK^-)\mu^\pm e^\mp$ simulation.	41
4.2	Trigger Configuration Key (TCK) alignment of $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ Monte Carlo (MC)/Data	44
4.3	Effect of pre-selection on signal datasets.	46
4.4	2D scans of M_{pK} used to determine the optimum $\Lambda(1520)$ region cut in the signal mode.	47
4.5	Justification for mis-identification vetoes applied in the signal mode.	48

4.6	Effect of applying semileptonic selection on signal mode.	50
4.7	Derivation of the optimum inclusive semileptonic cut for the signal mode. . .	51
4.8	Comparison of inclusive semileptonic cuts for the signal mode $M(\Lambda_b^0)$	52
4.9	M_{pKe^+} distribution after partial semileptonic selection	53
4.10	Effect of semileptonic selection on signal datasets.	54
4.11	Angular distribution θ_ℓ for signal mode simulation.	55
4.12	Feasibility study for a D^0 Semileptonic Veto.	56
4.13	Reconstructed momentum distributions before reweighting is applied	57
4.14	Mass fit to $M(\Lambda_b^0)$ with control mode data after cut-based selection.	58
4.15	Correction applied by the $\Lambda_b^0 \rightarrow pK^- J/\psi$ reweighter.	60
4.16	Effect on PID variables from applying PID efficiency corrections.	61
4.17	PID efficiency maps for the stripping selection from PIDCalib.	62
4.18	Track reconstruction efficiency ratio map from TrackCalib.	64
4.19	A map of Trigger efficiency ratio for correcting L0 trigger selection in simulation.	65
4.20	Justification for including nTRACKS in the Gradient Boosted Re-weighter (GBR).	67
4.21	Distributions of kinematic and derived variables showing correction applied by the reweighting scheme.	69
4.22	Performance comparison for MVA with different prior cuts.	73
4.23	Effective modelling of variables used to train the MVA.	74
4.24	Consistency of variables used to train the MVA as a function of $M(\Lambda_b^0)$ ranges.	75
4.25	Correlation matrices of MVA training variables.	76
4.26	Relative feature importance for the discriminating MVA input variables. . .	78
4.27	MVA probability distributions for training and testing.	79
4.28	Receiver Operator Characteristic (ROC) curves from the trained MVA. . . .	80
4.29	MVA performance comparison between the individual trained folds.	81
4.30	ROC Curves evaluating MVA performance for the different analysis categories.	81
4.31	Study of correlation of p_{MVA} output with $M(\Lambda_b^0)$	82

4.32	Performance comparison for the MVA response between μeOS and μeSS datasets.	82
4.33	Performance comparison of ε_{MVA} for signal and control mode samples.	84
4.34	Effect of PID selection on the signal mode data.	86
4.35	Effect of PID selection on the control mode.	87
4.36	PID mis-identification efficiency maps for PID selection from PIDCalib.	88
4.37	figure-of-merit (FoM) scans of MVA output for the Run 2 analysis categories.	89
5.1	Mass fit to corrected $\Lambda_b^0 \rightarrow pK^-K^+K^-$ simulation.	106
5.2	Invariant mass fit to corrected $\Lambda_b^0 \rightarrow D^0(\rightarrow K^-\ell^+\nu) p\ell^-\bar{\nu}$ simulation	107
5.3	Decay Tree Fitter (DTF) comparison for $M(pK^-\mu^+\mu^-)$ fits to corrected simulation.	108
5.4	Signal shape template fits to control mode simulation.	109
5.5	Control mode data fits for Run 1/2 without exclusive background components.	110
5.6	Exclusive background template fits for the control mode.	112
5.7	Extraction of exclusive background ratio from control mode data.	114
5.8	Control mode data fits for Run 1/2 with exclusive background components.	116
5.9	Stability of control mode fits using pseudo-experiments.	117
5.10	Stability of yield extraction for control mode data fit options.	118
5.11	Independence test between reconstructed $M(\Lambda_b^0)$ and M_{pK} for Run 2 control mode data.	120
5.12	Invariant mass fit to the M_{pK} spectrum for Run 2 control mode data.	121
5.13	Signal shape template fits to signal mode simulation.	124
5.14	Cascade fits to μeSS data at different p_{MVA} WPs.	127
5.15	Consistency test of μeSS distribution with floating parameters.	128
5.16	Blinded fits to the $M(\Lambda_b^0)$ distribution of signal mode μeOS data.	130
5.17	Blinded fits for the Run 1 combined category.	131
5.18	Stability tests for blinded μeOS fits using pseudo-experiments.	131
5.19	Cross-compatibility of background fit schemes for the signal mode.	132
5.20	Fits to unblinded pseudo-experiments for the analysis categories.	134

6.1	The LHCb Upgrade I VELO.	141
6.2	LHCb luminosity summary during Run 1–6 of the Large Hadron Collider (LHC)	142
6.3	Demonstration of track density in the instrumented VELO region for High Luminosity LHC (HL-LHC) conditions.	142
6.4	Schematic side-view of the LHCb Upgrade II detector.	144
6.5	PV reconstruction performance of VELO Upgrade II for various 4D tracking scenarios.	145
6.6	Schematic diagrams of the UT and MT upgrades for LHCb UII.	147
6.7	Resolutions as a function of incident particle angle for silicon pixel detectors.	150
6.8	Diagrams of charge collection in three fast-timing sensor technologies.	152
6.9	Current-Voltage (IV) measurements for LGAD sensor B performed at a range of temperatures.	157
6.10	IV measurement for LGAD sensor D after irradiation to 9×10^{14} 1 MeV $n_{\text{eq}}/\text{cm}^2$	158
6.11	Capacitance-Voltage (CV) measurements for sensor D before and after irradiation to 9×10^{14} 1 MeV $n_{\text{eq}}/\text{cm}^2$	160
6.12	Gain measurements for sensor D and F comparing performance at room temperature and cold.	162
6.13	Gain measurements for sensors D and H before and after irradiation to 9×10^{14} 1 MeV $n_{\text{eq}}/\text{cm}^2$	163
6.14	Impact parameter resolution performance for VELO UII as a function of module design parameters.	165
6.15	Visualisation of the four VELO scenarios considered.	168
6.16	Distribution of Primary Vertex (PV)s before and after the implementation of the 4D beam-spot method.	169
6.17	Track Impact Parameter (IP) resolution (in x) for different VELO Upgrade II scenarios.	172
6.18	Reconstruction performance of D_s^\pm candidates for different VELO Upgrade II scenarios.	173
6.19	Separation power of track p_T selection as a function of VELO UII scenarios.	174
6.20	PV reconstruction efficiency as a function of VELO UII scenarios.	175
6.21	Reconstruction and selection performance of D_s^\pm candidates for VELO UII at different σ_t working points.	177

6.22	PV matching efficiency as a function of η for muons originating from b -hadrons.	178
6.23	Tracking efficiency for VELO UII as a function of track η	179
6.24	Requirement for improved spatial resolution in VELO UII.	181
6.25	Particle and pixel hit rates for VELO UII.	183
6.26	Particle rate per pixel per event as a function of radius for VELO UII.	184
6.27	Workflow for performance optimisation tool for LHCb UII.	189
6.28	Flow diagram formulating a Global Optimisation approach for LHCb Upgrade II.	190
6.29	Example VELO UII studies displayed on the prototype LHCbPR dashboard.	192
6.30	Tracking efficiency for LHCb UII tracking detectors, displayed on the prototype LHCbPR dashboard.	193
6.31	Mighty Tracker (MT) and Upstream Tracker (UT) occupancy as a function of detector material from the UII LHCbPR dashboard.	194
B.1	Consistency of all variables used to train the MVA as a function of $M(A_b^0)$ ranges.	220
B.2	Effective modelling of all variables used to train the MVA.	221
D.1	Stability tests for control mode data fits without exclusive background components.	230
D.2	Stability tests for control mode data fits with exclusive background components.	231
D.3	Stability of yield extraction for control mode data fit options, Run 1.	231
D.4	Invariant mass fit to the M_{pK} spectrum for Run 1 control mode data.	232
D.5	Blinded fits to the $M_{A_b^0}$ distribution of signal mode μe OS data without an exclusive background component.	232
D.6	Cascade fits to μe SS data at different p_{MVA} WPs.	233
D.7	Stability tests from blind μe OS fits.	234
D.8	Cross-compatibility of background fit schemes for the signal mode.	235
E.1	Example of configuration files for stack builder and LHCbPR jobs.	237

List of Tables

1.1	The 19 free Standard Model (SM) parameters.	4
3.1	Particle Data Group (PDG) branching fraction values used during analysis.	33
4.1	Summary of selection applied by the analysis stripping line.	40
4.2	Summary of Fiducial Selection Cuts for the analysis.	42
4.3	Summary of analysis trigger selection.	43
4.4	Summary of pre-selection cuts applied for the analysis.	45
4.5	Summary of semileptonic selection for the analysis.	49
4.6	Calibration dataset summary for deriving PID weights.	61
4.7	Number of events in analysis datasets for training the MVA.	71
4.8	Summary of PID selection applied to the analysis.	85
4.9	Summary of MVA working points for the four analysis categories.	90
4.10	Multiple-candidate summary for analysis datasets.	91
4.11	Generator efficiency summary for simulation samples used.	93
4.12	Weighted efficiency summary for $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ simulation at different selection stages.	94
4.13	Weighted efficiency summary for $\Lambda_b^0 \rightarrow pK^- J/\psi(\rightarrow \mu^+\mu^-)$ simulation at different selection stages.	95
4.14	Weighted efficiency summary for background $\Lambda_b^0 \rightarrow D^0(\rightarrow K^-\mu^+\nu_\mu)pe^-\bar{\nu}_e$ simulation at different selection stages.	96
4.15	Weighted efficiency summary for background $\Lambda_b^0 \rightarrow D^0(\rightarrow K^-e^+\nu_e)p\mu^-\bar{\nu}_\mu$ simulation at different selection stages.	96
4.16	Weighted efficiency summary for background $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ simulation at different selection stages.	97

5.1	Background yield estimates for the key background samples in the signal mode $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$	103
5.2	Background yield estimates for the key background samples in the control mode $\Lambda_b^0 \rightarrow pK^- J/\psi(\rightarrow \mu^+ \mu^-)$	111
5.3	Final parameter values from control mode mass fits.	112
5.4	Results from invariant mass fits to the M_{pK} spectrum in the control mode.	121
5.5	Parameter values from signal shape template fits for signal.control mode.	122
5.6	Final shape parameters from the μe SS cascade fits.	126
5.7	Blinded fit results for the μe OS fit to $M(\Lambda_b^0)$ in each analysis category.	129
5.8	Single-event sensitivities and preliminary CL_s upper limits for the analysis.	133
6.1	LHCb subdetector changes across Run 1–6 of the LHC.	140
6.2	Summary of LGAD sensors studied for VELO UII feasibility.	155
6.3	Summary of four VELO UII scenarios implemented for testing in simulation.	167
A.1	Performance of alternative trigger strategies.	217
C.1	Signal mode background yield estimates for the background samples where hadronic mis-ID ($h \rightarrow \ell$) is required.	223
C.2	Signal mode background yield estimates for the background samples where leptonic mis-ID ($\ell \rightarrow \ell'$) is required.	224
C.3	Signal mode background yield estimates for the background samples where a single-semileptonic ($\ell\nu_\ell$) decay is required.	225
C.4	Signal mode background yield estimates for the background samples where a double-semileptonic ($\ell\nu_\ell, \ell'\nu_{\ell'}$) decay is required.	225
C.4	Background yield estimates for the all background samples in the control mode $\Lambda_b^0 \rightarrow pK^- J/\psi(\rightarrow \mu^+ \mu^-)$	227

List of Acronyms

LHCb Large Hadron Collider beauty.

PS Proton Synchrotron.

SPS Super Proton Synchrotron.

ATLAS A Toroidal Large Hadron Collider Apparatus.

AUC Area Under Curve.

BDT Boosted Decision Tree.

BSM Beyond the Standard Model.

BX Bunch Crossing.

CKM Cabibbo-Kobayashi-Maskawa.

cLFV Charged Lepton Flavour Violation.

CLIC Compact Linear Collider.

CMS Compact Muon Solenoid.

CoM Centre-of-Mass.

CP Charge-Parity.

CPV Charge-Parity Violation.

CV Capacitance-Voltage.

DAQ Data Acquisition.

DCB Double-Sided Crystal Ball.

DIRA DIRection Angle.

DLL delta-log-likelihood.

DOCA distance-of-closest-approach.

DTF Decay Tree Fitter.

ECAL Electromagnetic CALorimeter.

EFT Effective Field Theory.

EM Electromagnetic.

EW ElectroWeak.

FCC Future Circular Collider.

FCNC Flavour-Changing-Neutral-Current.

FD Flight Distance.

FoM figure-of-merit.

FTDR Framework Technical Design Report.

GBR Gradient Boosted Re-weighter.

GEM Gas Electron Multiplier.

GIM Glashow-Iliopoulos-Maiani.

GKDE Gaussian Kernel Density Estimator.

HCAL Hadronic CALorimeter.

HL High Luminosity.

HL-LHC High Luminosity LHC.

HLT High Level Trigger.

HPD Hybrid Photodetector.

ILC International Linear Collider.

IP Impact Parameter.

IT Inner Tracker.

IV Current-Voltage.

JTE Junction Terminating Extension.

L0 Level-0.

LEP Large Electron-Positron collider.

LFU Lepton Flavour Universality.

LFV Lepton Flavour Violation.

LGAD Low-Gain Avalanche Detectors.

LHC Large Hadron Collider.

LHCbPR LHCb Performance and Regression.

LQ Lepto-Quark.

LS2 Long-Shutdown 2.

LS3 Long-Shutdown 3.

LS4 Long-Shutdown 4.

MAPS Monolithic-Active-Pixel-Sensors.

MC Monte Carlo.

MD Mag Down.

MS Magnet Stations.

MT Mighty Tracker.

MU Mag Up.

MVA Multivariate Analysis.

MWPC Multi-wire Proportional Chamber.

NLL Negative-Log-Likelihood.

NN Neural Network.

NP New Physics.

OS Opposite Sign.

OT Outer Tracker.

PDF Probability Density Function.

PDG Particle Data Group.

PID Particle Identification.

PiN Positive-intrinsic-Negative.

PMNS Pontecorvo–Maki–Nakagawa–Sakata.

PV Primary Vertex.

QCD Quantum ChromoDynamics.
QED Quantum ElectroDynamics.
QFT Quantum Field Theory.
R&D Research and Development.
RF Radio-frequency.
RICH Ring Imaging CHerenkov.
ROC Receiver Operator Characteristic.
SM Standard Model.
SPD Scintillating Pad Detector.
SS Same Sign.
SV Secondary Vertex.
TCK Trigger Configuration Key.
TCT Transient Current Technique.
TDR Technical Design Report.
TIS Trigger-Independent-of-Signal.
TM Truth Matching.
TORCH Time-Of-internally-Reflected-CHerenkov-light.
TOS Trigger-On-Signal.
TT Tracker Turicensis.
TV “Timing Velo”.
UT Upstream Tracker.
VELO VERTex LOcator.
WLCG Worldwide LHC Computing Grid.

Chapter 1

Introduction and Theoretical Foundations

Science, at its core, is the pursuit of knowledge. Particle Physics is no exception. Since J. J. Thomson’s discovery of the electron in 1897 [1], with a “cathode ray” in a glass tube, experimental particle physics, as well as our understanding of nature, has grown rapidly. It has since evolved to the study of cosmic rays with cloud chambers [2, 3], laying the foundations for the Standard Model (SM) of particle physics, before the dawn of accelerator-based “High-Energy” particle physics with giant machines such as the Synchro-Cyclotron built at CERN in 1957 [4]. This has culminated in the state-of-the-art Large Hadron Collider (LHC) facility at CERN [5], world-leading in energy and discovery potential. The rapid development has been driven by the core pursuit of knowledge, with each discovery or improved measurement refining the SM into the extremely precise and predictive framework that currently provides the best theoretical understanding of the universe. Yet it remains incomplete, so the pursuit continues powered by pioneering technology enabling more prolific accelerators and ultra-sensitive detectors as the objective of particle physics focusses onto the search for Beyond the Standard Model (BSM) physics.

Through its initial derivation in the 1950s, 60s and 70s, the SM has taken on many roles, initially developing in response to new observations before becoming a more “predictive” model, forecasting the presence of a fourth quark and a third-generation of fermions with experiments tasked with confirming these predictions. This prompted the development of machines to push the energy and intensity frontiers of particle physics, with the Tevatron discovering the heaviest “top” quark [6, 7] and experiments at the Large Electron-Positron collider (LEP) rigorously constraining SM parameters [8, 9], culminating in the ultimate SM discovery facility, the LHC, observing the Higgs boson in 2012 [10, 11], the “final piece” of the SM puzzle rendering the fundamental structure complete.

Since the discovery of the Higgs boson, particle physics has entered its “exploration era” with a focus on investigating any inconsistencies in the expected SM description of particle interactions, searching for “New Physics (NP)” that could point towards explanations for observed phenomena not described in the SM. These include the lack of a description for gravitational interaction between particles, the fundamental content and interaction of Dark Matter [12] as well as the origin of so-called “Dark Energy” that determines the accelerating expansion of the universe [13, 14]. All are without description in the current form of the SM, prompting BSM searches with the construction of gargantuan neutrino and dark matter detectors [15, 16] as well as studies of ever-increasing precision at the LHC to identify any deviation that could hint at a “crack” in the SM. The LHC facility and corresponding experiments will continue to collect data and scrutinise the SM until 2042, with a future collider expected to supersede it, either the Future Circular Collider (FCC) [17], or a high-intensity linear-collider, Compact Linear Collider (CLIC) [18] or International Linear Collider (ILC) [19].

Another omission of the SM is the prediction that matter and anti-matter are produced at a near-identical rate [20], with the presence of a matter-dominated universe thoroughly contradicting this, introducing a branch of NP searches focused on the flavour of decaying particles. The Large Hadron Collider beauty (LHCb) experiment is the primary flavour facility at the LHC, a single-armed forward spectrometer designed to be sensitive to matter-antimatter differences (Charge-Parity Violation (CPV)) [21, 22]. Since 2011 it has collected data and made both SM precision measurements and NP searches, proving to be a general-purpose detector in the forward region as well as the world-leading CPV experiment. It also has a rich programme of direct and indirect BSM physics searches, investigating the decays of particles forbidden in the SM as well as scrutinising SM predictions for any signature of inconsistency. Chapters 3–5 of this thesis will describe the motivation, methodology and results of an analysis searching for SM-breaking Charged Lepton Flavour Violation (cLFV) with the LHCb detector. This is a direct probe for NP and the first measurement in the decay mode $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$.

As a precision detector, the primary limitation of LHCb is the size of the data sample collected. Since a successful first upgrade to the detector between 2018–2022, referred to as LHCb Upgrade I (UI) [23], Research and Development (R&D) and design work have begun for the proposed next phase of LHCb, known as LHCb Upgrade II (UII) [24], intended to utilise the High Luminosity (HL) potential of the scheduled upgrade to the LHC, to improve dramatically the sensitivity of LHCb measurements, enabling ultimate scrutiny of the SM and opening new pathways for BSM searches. Chapter 6 will set out the motivations and plans for LHCb UII, focussing on the upgrade to the VErtex LOcator (VELO) subdetector [25], evaluating a potential new sensor technology and describing simulation that is aiding to determine the design of the upgraded subdetector. A new

framework will also be introduced intended to facilitate the overall “Global Optimisation” of LHCb UII as it strives to “fully realise the flavour-physics potential of the High Luminosity LHC (HL-LHC)” [24].

1.1 The Standard Model

The SM provides a unified classification for all known fundamental particles, describing their interactions with three of the four forces of nature in a self-consistent, non-abelian, gauge Quantum Field Theory (QFT). Defined by the symmetry group $U(1)_Y \times SU(2)_L \times SU(3)_C$ it governs the Electromagnetic (EM), Weak and Strong force (with gravity omitted), leading to 21 SM fields in total along with 18 distinct particles [26]. In its current form, particles in the SMs fall into two categories: 12 spin- $\frac{1}{2}$ matter particles, the fermions, and six bosons, with five spin-1 gauge bosons that mediate the three quantum forces, and the spin-0 Higgs boson. Each fermion has a corresponding anti-particle, identical but for inversion of all internal quantum numbers such as charge and parity. Fermions are also further split into two types: quarks, charged particles subject to the strong force, only observed in bound states and leptons, not interacting via the strong force but as a result able to exist as free charged and neutral fermions.

The three forces stated are reduced to two unified theories and an underlying field. The theory of Quantum ElectroDynamics (QED) mediates the interaction between light and matter, providing a quantum description of EM [27], related to the $U(1)_Y$ group. At a large enough energy scale, EM unifies with the weak interaction ($SU(2)_L$) [28–30] to form the ElectroWeak (EW) interaction [31], described by the gauge symmetry $U(1)_Y \times SU(2)_L$. The EW force is associated with one massless boson, the photon (γ), governing solely EM interactions [27, 32], as well as three massive weak bosons (W^\pm, Z^0) formed via the spontaneous symmetry breaking of the Higgs field [33–35]. The $SU(3)_C$ group corresponds directly to the strong force described by the theory of Quantum ChromoDynamics (QCD) [36–38], dictating the interaction of quarks via the exchange of gluons, mediating a transfer of an intrinsic “colour charge” binding quarks into multi-quark states, hadrons. The sixth boson, the Higgs boson (H^0), corresponding to the aforementioned field, is responsible for giving mass to the weak bosons (via spontaneous EW symmetry breaking that creates the Higgs itself [33–35]) and fermions via the Yukawa interaction [39, 40], manifesting in a H^0 - f coupling proportional to the fermion mass (m_f).

The unification of the SM reduces it to 26 free parameters, with the 19 primary defining observables stated in Table 1.1. The remaining seven correspond to the properties of the neutral leptons, the neutrinos. Of these, 11 are associated with the Higgs field (nine

Table 1.1: The 19 free SM parameters in the form of physical observables, with the parameter values from [41]. Adapted from [42].

Parameter	Value
Quark masses: m_u, m_c, m_t (MeV/ c^2)	$2.16_{-0.26}^{+0.49}, (1.27 \pm 0.020) \times 10^3, (172.69 \pm 0.30) \times 10^3$
Quark masses: m_d, m_s, m_b (MeV/ c^2)	$4.67_{-0.17}^{+0.48}, 93.4_{-3.4}^{+8.6}, 4.18_{-0.02}^{+0.03} \times 10^3$
Lepton masses: m_e, m_μ, m_τ (MeV/ c^2)	$0.51099895000(15), 105.6583755(23), 1776.86(12)$
Fine Structure constant, $\alpha_{EM} \equiv \frac{e^2}{4\pi}$	$(137.035999084(21))^{-1}$
Strong coupling, $g_s = \sqrt{4\pi\alpha_s}$	$0.1179(9)$
Fermi Constant, G_F (/GeV 2)	$1.1663788(6) \times 10^{-5}$
Higgs Mass, m_H (GeV/ c^2)	$125.25(17)$
Vacuum Expectation value, v (GeV/ c^2)	~ 246.22
CKM mixing angles: $\sin \theta_{12}, \sin \theta_{23}, \sin \theta_{13}$	$0.22500 \pm 0.00067, 0.04182_{-0.00074}^{+0.00085}, 0.00369 \pm 0.00011$
CKM matrix CPV phase, δ	$1.144(27)$
QCD phase angle, θ_{QCD} (rad)	$\lesssim 2 \times 10^{-10}$

m_f, m_{H^0} and the vacuum expectation energy v), three to the fundamental forces α_{EM} (EM), G_F (weak) and g_s (strong), four mixing parameters of the Cabibbo-Kobayashi-Maskawa (CKM) matrix (discussed further in Section 1.2) and a QCD phase. These parameters can not be predicted with the SM and must be determined experimentally with the current values presented in Table 1.1. This highlights the dichotomy of the SM, with some parts displaying a clear mathematical rigidity, e.g. the weak boson masses, while other aspects are far less elegant, with emerging patterns and structure currently unexplained by the SM, propelling searches for BSM physics.

1.1.1 Leptons, Quarks and Hadrons

Fermions constitute matter, combining to form nucleons, atoms and the entire universe. Right-handed fermions form right-handed singlets in the SM, however the left-handed fermions form $SU(2)_L$ doublets in three generations, separated by mass [43]. Beginning with the leptons, each left-handed doublet contains one massive charged lepton (ℓ^-) with an intrinsic charge¹ of -1 and a corresponding neutral, ultra-low-mass neutrino (ν_ℓ). Each doublet is associated with a “flavour”, electron, muon and tau (e, μ, τ) of increasing mass. As leptons do not interact with the strong force, they hold no colour charge, however they are assigned a “Lepton Number” L , with leptons holding $L = +1$ and anti-leptons -1 , which is conserved in SM interactions [44]. This extends to “Lepton Flavour Conservation”, where each generation of leptons holds an independent L_ℓ that

¹In units of $|e|$, the charge of the electron, which is implicit for subsequent references to charge.

has so far been observed to be conserved. However this “accidental” symmetry remains as an experimental observation, Section 1.3 will expand on how Lepton Flavour Violation (LFV) could naturally enter and extend the SM.

Quarks are arranged into left-handed doublets of one up-type (u, c, t) and one down-type quark (d, s, b), with the three generations following the trend of increasing mass seen in leptons. Up (down) type quarks hold an EM charge of $+\frac{2}{3}$ ($-\frac{1}{3}$) as well as a colour charge of either red, blue or green, with each corresponding anti-particle holding the opposite charges, $-\frac{2}{3}$ and \bar{r} for the example of \bar{u} . Unbound quarks have not been observed [45], instead forming bound hadron states. These are divided into $q\bar{q}$ mesons and qqq baryons, with corresponding anti-particles. Each quark in a hadron is assigned a $\pm\frac{1}{3}$ “baryon number”, with the bound states forming integer values of B and the conservation of this observed in SM interactions, forming another accidental symmetry [44]. The formation of hadrons is also dictated by the colour charge of the constituent quarks, forming colour neutral states of either $r\bar{r}, b\bar{b}, g\bar{g}$ or rbg .

The flavour of a hadron is defined by its valence quarks and the particle name typically follows the heaviest quark, e.g. the K^0 , D^0 and B^0 are strange, charm and “beauty” mesons with corresponding Λ , Λ_c^+ and Λ_b^0 baryons with increasing mass by generation as seen in quarks. There are no corresponding top-quark hadrons due to their extremely short life time (5×10^{-10} fs) resulting in decay before hadronisation, the process by which hadrons are formed [26]. There are also quarkonia meson states formed of $c\bar{c}$ and $b\bar{b}$ pairs as well as “exotic” higher-order hadrons, tetraquarks and pentaquarks, recently discovered as four and five quark bound states permitted and predicted within the SM but not observed until 2013–15 [46–49]. The study of the heavy flavour charm and beauty hadrons is referred to as “Flavour Physics” and is a focus of current experiments including LHCb, Belle II [50] and BESIII [51], with LHCb able to study both beauty mesons and baryons with potentially interesting differing properties such as the integer/non-integer spin from the even/odd number of fermions.

1.2 Flavour Physics

Charged-current interactions mediated by the W^\pm boson are the only process that can change quark flavour in the SM. In the early 1960s, observed differences in the decay rates of u/d and strange mesons through charged-current interactions prompted Cabibbo to formulate a new treatment of quark states to retain universality of the weak coupling constant g_W [52]. Cabibbo presented a separation between the mass eigenstates d, s and

flavour eigenstates d', s' of down-type quarks,

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos \theta_c & \sin \theta_c \\ -\sin \theta_c & \cos \theta_c \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}, \quad (1.2.1)$$

introducing eigenstate mixing via the ‘‘Cabibbo Angle’’ $\theta_c \approx 12.7^\circ$ [52], with preferential coupling within generations. This produces a splitting in the charged-current interaction, $W^+ \rightarrow u\bar{d}' = u\bar{d}, u\bar{s}$ with the probability of each proportional to the square of the corresponding matrix element V_{qq} , explaining the discrepancies in decay rates observed in π^\pm and K^\pm decays [53].

The initial derivation of Cabibbo mixing did not include the second matrix row in Equation 1.2.1, providing a single generation $\begin{pmatrix} u \\ d' \end{pmatrix}$ doublet, implying the existence of a $Z^0 \rightarrow s\bar{d}$ interaction [30, 54]. The branching fraction this predicted for the decay $K_L^0 \rightarrow \mu^+\mu^-$ disagreed with measurement, prompting the inclusion of a second doublet $\begin{pmatrix} c \\ s' \end{pmatrix}$, providing cancellation of different-flavour mass eigenstate pairs when expanding $Z^0 \rightarrow d'\bar{d}' + Z^0 \rightarrow s'\bar{s}'$, due to the negative $\sin \theta_c$ term in Equation 1.2.1. This constructs a rule in the SM that there are no Flavour-Changing-Neutral-Current (FCNC) at tree level, and any interaction must proceed through a suppressed ‘‘box’’ or ‘‘loop’’ diagram such as those shown in Figure 1.1, mediated by ‘‘virtual’’ u, c or t quarks. This is the Glashow-Iliopoulos-Maiani (GIM) mechanism [30] and its proposal in 1970 provided a concrete prediction for the presence of a fourth quark, with the charm quark discovered four years later via the observation of the charmonium meson J/ψ [55, 56].

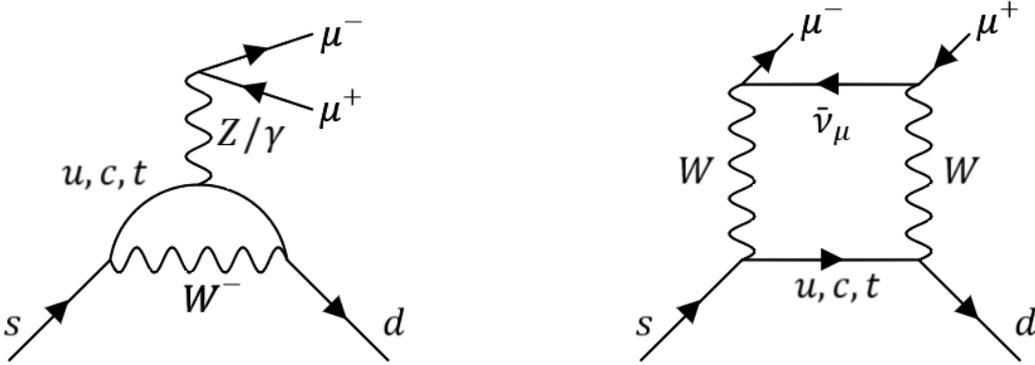


Figure 1.1: Feynman diagram representation of a $s \rightarrow d\mu^+\mu^-$ transition proceeding through virtual loops, representing a $K^0 \rightarrow \mu^+\mu^-$ decay in the SM, displaying the process of GIM suppression. (left) The EW ‘‘penguin’’ form. (right) The box form.

The Cabibbo matrix was subsequently generalised to include the third generations of quarks to allow CPV, forming the unitary CKM matrix [57],

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}, \quad (1.2.2)$$

where each s_{ij} or c_{ij} element correspond to the CKM mixing angles stated in Table 1.1 and δ the CPV phase [57]. V_{CKM} extends the trend in Equation 1.2.1, with the diagonal elements larger in magnitude corresponding to preferred interactions within generations and reduced-rate of quark mixing across generations, so-called CKM suppression. The relative suppression at higher generations is best illustrated by the Wolfenstein parameterisation of the CKM matrix [58]

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \quad (1.2.3)$$

with all parameters $\mathcal{O}(1)$ and $\lambda = 0.22500 \pm 0.00067$ [41], displaying the vanishing probability of a third-to-first generation transition with respect to the already suppressed $s \rightarrow u$ and $c \rightarrow d$ decays. As discussed, the mixing angles and CPV phase in Equation 1.2.2 are free parameters of the SM and determined from measurements of $|V_{qq}|$ via a plethora of decays of heavy flavour hadrons proceeding through the corresponding charged-current interactions, with the current experimental status [41]

$$|V_{CKM}| = \begin{pmatrix} 0.97435 \pm 0.00016 & 0.22500 \pm 0.00067 & 0.00369 \pm 0.00011 \\ 0.22486 \pm 0.00067 & 0.97349 \pm 0.00016 & 0.04182^{+0.00085}_{-0.00074} \\ 0.00857^{+0.00020}_{-0.00018} & 0.04110^{+0.00083}_{-0.00072} & 0.999118^{+0.000031}_{-0.000036} \end{pmatrix}. \quad (1.2.4)$$

Overall this results in a modification of the weak coupling, introducing flavour dependent decay rates dictating the experimentally observed branching fractions of light and heavy flavour hadrons [41]. At each interaction vertex in a Feynman diagram (such as Figure 1.1) the transformation of initial-state to final state particles is governed by the decay rate matrix element \mathcal{M}_{fi} modifying the differential decay rate,

$$d\Gamma(A \rightarrow f_1 + f_2 + \dots + f_n) = \frac{1}{2M_A} \|\mathcal{M}_{fi}\|^2 \left(\prod_f^n \frac{d^3p_f}{(2\pi)^3 2E_f} \right) (2\pi)^4 \delta^4 \left(p_A - \sum_f p_f \right), \quad (1.2.5)$$

for an n -body decay in the centre of mass frame [59], where the product term dictates the phase-space available in the decay and the four-dimensional Dirac-delta function ensures

four-momenta conservation [60]. For weak charged-current interactions $\mathcal{M}_{fi} \propto g_W V_{q_i q_j}$, mathematically enforcing CKM favoured and suppressed decays.

Equation 1.2.5 can also effectively describe GIM suppression, which combines with CKM suppression to produce FCNC decays with very small branching fractions, known as ‘‘Rare Decays’’. In the example of neutral $b \rightarrow s$ mixing, the total matrix element $\mathcal{M}_{fi} \propto \sum_{i,j=u,c,t} V_{ib}^* V_{is} V_{jb} V_{js}^*$ expands and factorises to form three components with $\sum_{q=u,c,t} V_{qb}^* V_{qs} = \delta_{bs} = 0$ by definition of the unitary CKM matrix, dramatically reducing the decay rate with respect to tree-level $V_{qb}^* V_{qs}$ processes. Higher order correction terms, from the mass of the u, c, t virtual mediating fermions, apply factors of $\mathcal{O}\left(\frac{m_q^2}{m_W^2}\right)$ to the matrix element [32] which do not cancel in the sum over generations. The quark loop is therefore dominated by virtual top quarks due to the gap in m_q (see Table 1.1), despite the CKM suppression the V_{ts} term applies. The experimental realisation of this can be seen, for example, by comparing the most fundamental $b \rightarrow s \ell^- \ell^+$ decay $B_s^0 \rightarrow \mu^+ \mu^-$ to the corresponding $b \rightarrow d \ell^- \ell^+$ decay $B^0 \rightarrow \mu^+ \mu^-$, with $\mathcal{B}(B_s^0 \rightarrow \mu^+ \mu^-) = (3.09 \pm 0.46) \times 10^{-9}$ and the most recent limit of $\mathcal{B}(B^0 \rightarrow \mu^+ \mu^-) < 2.6 \times 10^{-10}$ at 95% confidence [61]. Both are rare decays due to GIM suppression and their ratio provides a direct probe of $|V_{td}/V_{ts}|$.

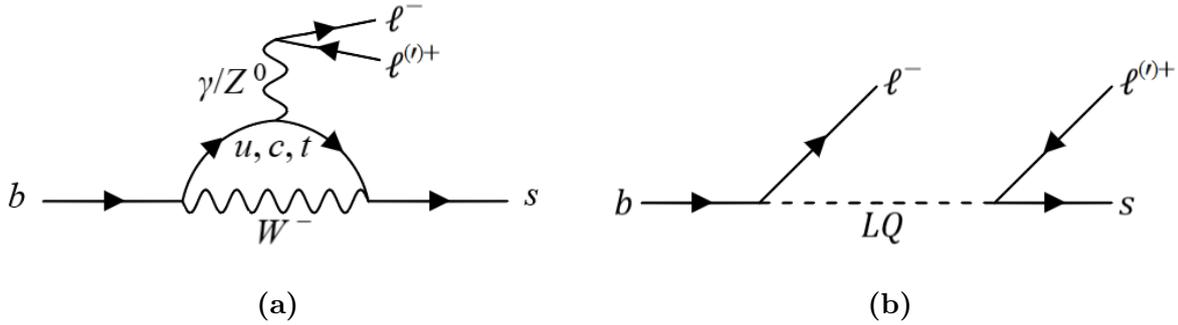


Figure 1.2: Feynman diagram representation of a $b \rightarrow s \ell^{(\prime)+} \ell^-$ transition proceeding through virtual loops. (a) Shows the SM box form EW penguin form and (b) shows the same transition mediated by a proposed BSM Lepto-Quark (LQ) mediator [62, 63].

1.2.1 Lepton Flavour Universality

Lepton Flavour Universality (LFU) is an experimentally observed accidental symmetry in the SM that originates from the leptonic $U(1)_Y$ group ($U(1)_L$) imposing equal first-order coupling to all generations of charged and neutral leptons for a given gauge-boson [26]. This results in near-identical decay rates for $W \rightarrow \ell \nu_\ell$ or $Z/\gamma \rightarrow \ell^+ \ell^-$ interactions of any flavour [41], with minor differences due to m_ℓ .

While there are tight experimental constraints on LFU for tree-level decays [41],

studying LFU with FCNCs is one of many methods to scrutinise the SM and search for NP. Using processes mediated by virtual loops can provide experiments with indirect sensitivity beyond the collision Centre-of-Mass (CoM) energy, with contribution to the decay rate from massive virtual BSM particles resulting in a deviation from the predicted SM value. A hypothetical example of a LQ mediating a $b \rightarrow s\ell^- \ell^+$ interaction is described in Figure 1.2b. Any NP contribution is expected to be small (hence the lack of observation thus far), but many BSM models [64–66] suggest NP in FCNC interactions to be within reach of current or next-generation experiments. This is referred to as an “indirect” search for NP.

Often the factor limiting NP-sensitivity for indirect searches is the precision of the SM prediction. As experimental precision improves, any change in theoretical precision or shift in μ_{SM} could alter the significance; this is the case in the recent $(g-2)$ measurement [67, 68] with one theoretical prediction producing a $> 5\sigma$ result [69] and another $< 2\sigma$ [70]. In an attempt to mitigate changes in the theoretical predictions of $b \rightarrow s\ell^- \ell^+$ branching fractions, recent studies of LFU at LHCb have utilised the ratio

$$R_H = \frac{\mathcal{B}(B \rightarrow H\mu^+\mu^-)}{\mathcal{B}(B \rightarrow HJ/\psi(\rightarrow \mu^+\mu^-))} \bigg/ \frac{\mathcal{B}(B \rightarrow He^+e^-)}{\mathcal{B}(B \rightarrow HJ/\psi(\rightarrow e^+e^-))}, \quad (1.2.6)$$

where H is the hadronic component of a B decay [71], with a “double-ratio” between same-flavour rare ($b \rightarrow s\ell^- \ell^+$) and resonant ($b \rightarrow sJ/\psi$) decay modes, as well as between the flavour generations, ensuring that any systematic effects in the theory derivation and experimental procedure propagate with maximal cancellation. This provides a robust test of LFU, yielding a predicted value of $R_H \simeq 1$ for all hadronic varieties with $< 1\%$ theoretical uncertainty [72, 73].

The consistency of LFU has been tested by many experiments, producing precise results in agreement with the SM for $W \rightarrow \ell\nu_\ell$, $Z \rightarrow \ell^+\ell^-$ and $J/\psi \rightarrow \ell^+\ell^-$ decays [9, 41]. The $\mu^+\mu^-/e^+e^-$ ratio of the latter has been measured to $r_{J/\psi} = 0.998 \pm 0.008$, ensuring that Equation 1.2.6 is testing the consistency of the $b \rightarrow s\ell^- \ell^+$ rare modes only. A summary of R_H measurements at LHCb is presented in Figure 1.3, showing consistency of $\mathcal{O}(1-2)\sigma$ in all cases [74–76]. However previous measurements of these ratios had hinted at NP [78], prompting a tide of focus on probing LFU in $b \rightarrow s\ell^- \ell^+$ decays, with deviations of $\mathcal{O}(2-3)\sigma$ remaining in other observables including angular [79–81] and individual branching fractions [82–84]. Further measurements with larger experimental datasets are required to draw definitive conclusions regarding the LFU symmetry, with current measurements at LHCb statistically limited. Nevertheless, the potential sensitivity to new physics in the rare decays of heavy flavour hadrons prompts further investigation of new and alternative observables.

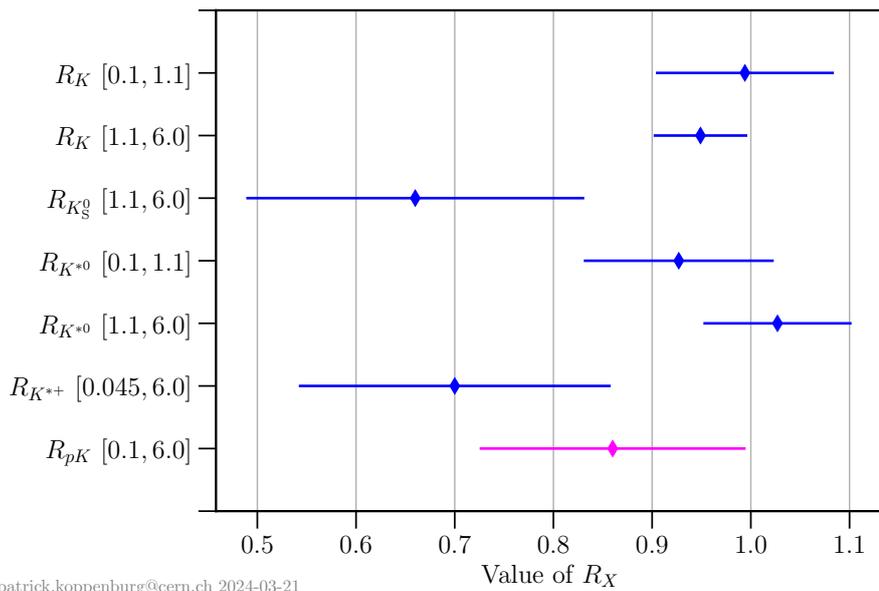


Figure 1.3: Summary of published LHCb R_H measurements [74–76]. Blue and pink points distinguish between a parent B -meson or B -baryon respectively. Taken from [77].

1.3 Searching for New Physics with Lepton Flavour Violation

While no unequivocal evidence of NP in the decays of heavy flavour hadrons have been observed up to now, the search continues for processes breaking LFU with the lack of a formal SM symmetry suggesting that the consistency observed could be local to the energy-scales probed [71]. Furthermore, the SM continues to not explain many clear, observed phenomena. Many of these limitations involve flavour physics:

CP Violation Quark CPV originates naturally in flavour-changing currents, but the observed rates are insufficient to justify the stark matter-antimatter asymmetry of the universe [20]. Possible origins of the missing violation include leptonic CPV within neutrino interactions [85] as well as from the yet undetermined QCD phase θ_{QCD} which governs CPV in the strong force, the so-called “Strong CP Problem” [26, 86];

Mass Hierarchy As discussed, both quarks and leptons exist in three distinct generations with a hierarchy of vastly different fermion masses that the SM does not predict or explain.

Most striking, is the observed phenomena of neutrino oscillations [87–89], where neutrino flavour (ν_e , ν_μ or ν_τ) can change when propagating over large distances. The original form

of the SMs assumed that neutrinos were massless, but an oscillation mechanism requires non-zero $\Delta m_{\nu\nu'}$ implying that at least one neutrino is massive. To facilitate this neutrinos are hypothesised to exist as mass eigenstates ($\nu_{i=1,2,3}$), a superposition of multiple flavour eigenstates, similar to down-type quarks in the CKM matrix and resulting in a required extension, the Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix [87, 90]. While the oscillation mechanism and elements of this matrix are not finalised in measurement [41], neutrino-mixing provides indisputable evidence that lepton flavour is not conserved absolutely in the SM when neutrinos propagate between production and observation. This justifies consideration of whether the same process that allows neutrino oscillation can result in a non-zero coupling between leptons of different generations, LFV; if neutrinos oscillate, why do the charged leptons not?

As yet, there is no evidence for LFV, surviving as another accidental symmetry of the SM with the EW bosons observed to only couple within lepton generations and processes such as $W \rightarrow \ell\nu_{\ell'}$ or $Z/\gamma \rightarrow \ell\ell'$ not seen. However a NP process could, in-theory, simultaneously mediate the observed neutrino-mixing as well as $\ell\nu_{\ell'}$ or $\ell\ell'$ final states and is therefore an avenue for BSM searches. Furthermore, evidence for coupling between lepton generations could aid with understanding of the flavour problem and could reveal a relationship between the charged lepton masses.

One method to search for LFV are bespoke experiments like “Mu2e” [91] and “Mu3e” [92], both attempting to observe neutrinoless muon-electron conversion, with an additional e^+e^- emission for Mu3e, focussing experimental effort on Charged Lepton Flavour Violation (cLFV) specifically with visible final states of charged leptons, without the difficulty of neutrino reconstruction. However, these experiments are impaired by high-rate SM processes, e.g. μ -decay with or without hard photon emission, and require further R&D to be fully-realised [93]. Other methods include searching for cLFV within large samples of light hadron decays, such as at NA62 [94].

Searches for cLFV have also been performed at collider experiments, e.g. by studying the production and decay of Higgs bosons or top quarks within pp collisions [95–97]. An alternative is to search for cLFV in the decays of b -hadrons, specifically $b \rightarrow s\ell^-\ell'^+$ processes, that involve at least one spectator quark. This produces a corresponding hadronic system that can be utilised experimentally for selection and background rejection and could potentially enhance NP [98, 99]. A $b \rightarrow s\ell^-\ell'^+$ search also complements the discussed studies of FCNC $b \rightarrow s\ell^-\ell^+$ processes, both by the sharing of experimental techniques for studying the rare decays of b -hadrons but also that many theoretical models that provide a mechanism towards lepton non-universality naturally produce LFV as a by-product [100, 101], with constraints from both methods able to be collated into global fits for potentially greater NP sensitivity [102–104].

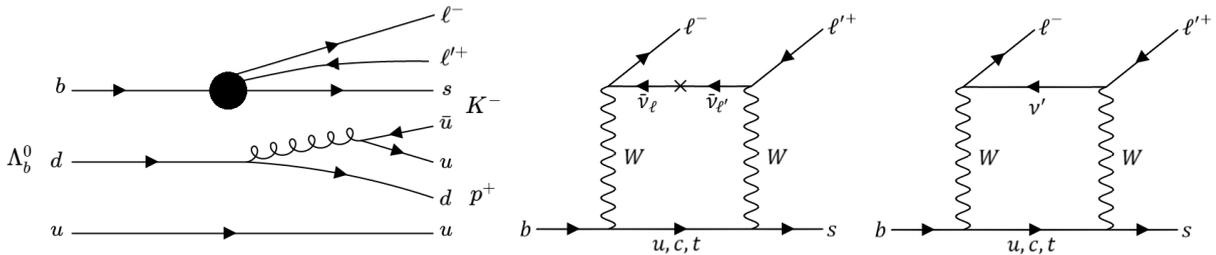


Figure 1.4: Feynman diagram representation of a $\Lambda_b^0 \rightarrow \Lambda(1520)\ell^+\ell^-$ transition with (left) the quark interaction (centre, right) different potential mediators replacing the four-particle interaction vertex including a SM oscillating neutrino as well as proposed NP interactions through BSM ν' with the LQ equivalent in Figure 1.2b.

Unlike the $b \rightarrow s\ell^-\ell^+$ interaction, $b \rightarrow s\ell^-\ell'^+$ decays are forbidden in the SM. The only viable SM method yielding a process such as $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ is via a virtual neutrino oscillation within a FCNC as shown in Figure 1.4, requiring a higher-order matrix-element correction term of $\mathcal{O}\left(\frac{m_\nu}{m_W}\right)^2$ which, for a SM like neutrino ($m_\nu \leq 1$ eV), inflicts a $\approx 10^{-44}$ scale-factor on the decay rate compared to the $b \rightarrow s\ell^-\ell^+$ equivalent. This projects $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp) \approx 10^{-51}$ from the measured $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-)$ [105], rendering this process unobservable for any current or future experiments if only the SM is considered. The null result for a cLFV search is therefore essentially $\mathcal{B} = 0$, removing dependence on the precision of a SM prediction from which to measure deviations, thus providing a “direct” search for NP. Therefore, any significant evidence of $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ would serve as unequivocal evidence for NP, a clear signature of a new interaction mediated by a BSM process either at the tree level or embedded within a virtual loop diagram.

Two potential NP models for mediating the $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ process have been presented in Figures 1.2 and 1.4, a sterile neutrino model [106], which can interact with both flavour ℓ^\pm at $W \rightarrow \ell\nu'$ vertices, as well as a scalar LQ model [63, 107] with direct coupling to both quarks and leptons at tree level. While both models produce a feasible interaction, a single observation of LFV would be unable to confirm one NP model. Therefore it is appropriate for a model-independent approach to be adopted for LFV searches, where the phase-space component of Equation 1.2.5 determines the expected kinematics of the target decay mode. This method maintains the ability for subsequent reinterpretation of results to either investigate sensitivity to a preferred NP model or to combine with complementary studies. One common method is to utilise an Effective Field Theory (EFT) approach [108], factorising SM symmetries and conservation laws into individual local operators \mathcal{O}_i , with corresponding Wilson coefficients \mathcal{C}_i that separate BSM contributions from any corresponding SM feature, $\mathcal{C}_i = \mathcal{C}_i^{\text{SM}} + \mathcal{C}_i^{\text{BSM}}$. For a $\Lambda_b^0 \rightarrow \Lambda(1520)\ell_1\ell_2$ decay, the EFT Hamiltonian can be described by

$$\mathcal{H}_{\text{eff}} = \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \sum_i (C_i^{\ell_1 \ell_2} O_i^{\ell_1 \ell_2} + C_i'^{\ell_1 \ell_2} O_i'^{\ell_1 \ell_2}), \quad (1.3.1)$$

where only the top-quark elements of the CKM matrix are retained due to the discussed dominance, and G_F is the Fermi constant [42, 101].

Another application of Equation 1.3.1 is to produce predictions for BSM decay modes, and while no formal calculation has been made for $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$, a projection from the prediction for the analogous channel $\Lambda_b^0 \rightarrow \Lambda \ell \ell'$ [101] is possible. Applying a factor of $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-)/\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-)$,

$$\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp) \lesssim 2.2 \times 10^{-8}, \quad (1.3.2)$$

potentially within sensitivity of a Run 1 and 2 LHCb measurement and undoubtedly in the near-future with the LHCb dataset projected to increase over five-fold during Run 3 and 4 of the LHC, with a further factor of seven expected for LHCb Upgrade II. This motivates cLFV searches at LHCb with the potential to either observe NP or significantly constrain NP models, narrowing and improving understanding of the BSM theoretical landscape.

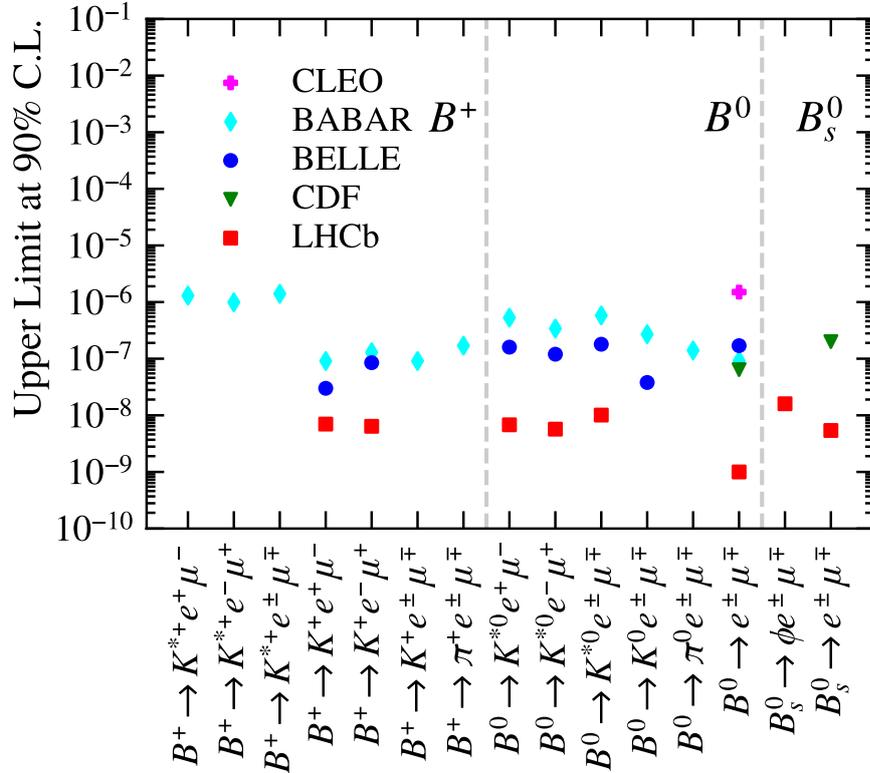


Figure 1.5: A summary of cLFV $b \rightarrow s \ell^- \ell'^+$ searches at LHCb and other experiments in the $\mu^\pm e^\mp$ final state [109–113]. All limits are at the 90% confidence level. Provided by [114].

A wide range of $B \rightarrow H\ell\ell'$ searches have been published by LHCb, with $H = K$ [113, 115], K^{*0} [112, 116], ϕ [112, 117] including searches for $H\mu^\pm e^\mp$ and $H\tau^\pm\mu^\mp$ as well as the purely leptonic $B \rightarrow \mu^\pm e^\mp$ [111] and $B \rightarrow \tau^\pm e^\mp$ [118], summarised in Figure 1.5. All have produced world-leading limits for LFV, providing constraints and inputs to NP models [109]. As yet, there are no published results for LFV in the decays of b -baryons, with one ongoing study of $\Lambda_b^0 \rightarrow \Lambda e^\mp \mu^\pm$ described in Ref. [42, 119]. This motivates further searches in the decays of Λ_b^0 baryons with distinct spectator quarks to mesonic searches, resulting in a non-integer initial spin and sensitivity to different hadronic form-factors that could potentially enhance NP contributions. Chapters 3–5 of this thesis describe the methodology and results for an analysis of $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$, designed to yield either the first measurement of cLFV or set a world-first, competitive limit on the branching fraction. Chapter 3 will introduce the methodology behind a cLFV search at LHCb, also justifying the hadronic choice of the excited strange-baryon $\Lambda(1520)$, before Chapters 4 and 5 describe the process from candidate selection through to result interpretation, with the ultimate result of this work a blinded limit on $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp)$.

Chapter 2

The LHCb Detector - Run 1 & 2

2.1 The Large Hadron Collider

2.1.1 The LHC Accelerator

The LHC is a 27 km circular hadron collider at the CERN laboratory near Geneva. In its standard operational mode, the LHC accelerates protons in opposite directions, colliding the beams with an extreme Centre-of-Mass (CoM) energy (\sqrt{s}) at four interaction points instrumented with the four primary LHC experiments: ATLAS, CMS, ALICE and LHCb [120–123]. The LHC can also operate with heavy-ion beams [124]. Chapters 3–5 of this thesis analyses pp collision data taken during Run 1 (2011–2012) and Run 2 (2015–2018) of the LHC, hence the accelerator during this period will be described. Minor improvements have since been implemented for Run 3 and a major upgrade, the HL-LHC, is scheduled for Run 4, detailed in Chapter 6 of this thesis.

As described in Figure 2.1, the LHC is the final accelerator of a series of smaller machines that step towards the CoM energy of $\sqrt{s} = 13$ TeV for Run 2 (7 and 8 TeV for 2011 and 2012). The protons are extracted from hydrogen gas and initially accelerated to 450 MeV within LINAC2 and injected into the Proton Synchrotron (PS) system, consisting of booster rings and a 0.63 km primary accelerator [126]. The PS achieves a beam energy of 25 GeV and separates the protons into bunches of $\approx 1.5 \times 10^{11}$ per bunch, with neighbouring bunches spaced by 25 ns intervals, tuned for the LHC clock frequency of 40 MHz. The Super Proton Synchrotron (SPS) provides the final acceleration to 450 GeV before proton bunches are injected into the LHC, with a design maximum of 2808 colliding bunches per beam, referred to as an LHC “fill” [127].

The LHC accelerates the proton bunches to their nominal beam energy with an

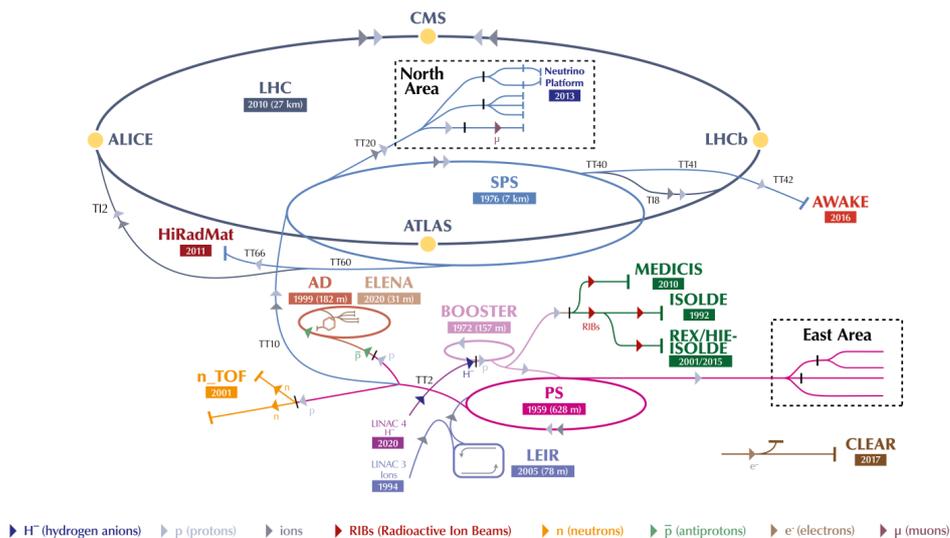


Figure 2.1: The status of the CERN accelerator complex during Run 2 of the LHC. Taken from [125].

Radio-frequency (RF) cavity installation operating at 40 MHz inline with the bunch frequency. The beam-trajectory is maintained by superconducting dipole magnets and quadrupole focussing magnets [126]. The LHC beam pipe contains an Ultra-High Vacuum of $10^{-10} - 10^{-11}$ mbar [126] minimising the likelihood of beam-gas collisions in the accelerator that could trigger a beam dump, a safety mechanism designed to automatically redirect the LHC beams into steel-concrete blocks if, for example, a problem with the beam condition or an issue at one of the experiments is detected.

At the four experimental interaction points dipole magnets direct the beams to cross, colliding bunches at the LHC clock frequency of 40 MHz. Not every Bunch Crossing (BX) contains two proton bunches, with some “beam-empty” or “empty-empty”, rendering an effective maximum average collision frequency of 30 MHz [126]. Collision BXs separated by 25 ns are still present however so experiments must tolerate this peak rate.

During a fill, the number of BXs producing pp collisions decays exponentially as the beams deplete and are repeatedly cleaned via collimation [126]. For ATLAS and CMS, attempting to yield the maximal data samples, this results in a decaying instantaneous luminosity (\mathcal{L}_{inst}) over a fill of $\lesssim 20$ h. LHCb, requiring a stable \mathcal{L}_{inst} , performs “luminosity levelling” to maintain $\mathcal{L}_{inst} = 4 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ by initially directing the colliding beams away from each other and gradually returning them to head-on as the fill progresses [128]. This enables control of “ μ ”, the average number of visible pp interactions per BX. For Run 1 and 2 of LHCb, $\mu \approx 1$ [128], in comparison to ATLAS and CMS which operate at $\langle \mu \rangle = 32$ for Run 2. Throughout this thesis each recorded BX is referred to as an “event”, with the measured products referred to as reconstructed “candidates”.

2.1.2 The LHCb Experiment at the LHC

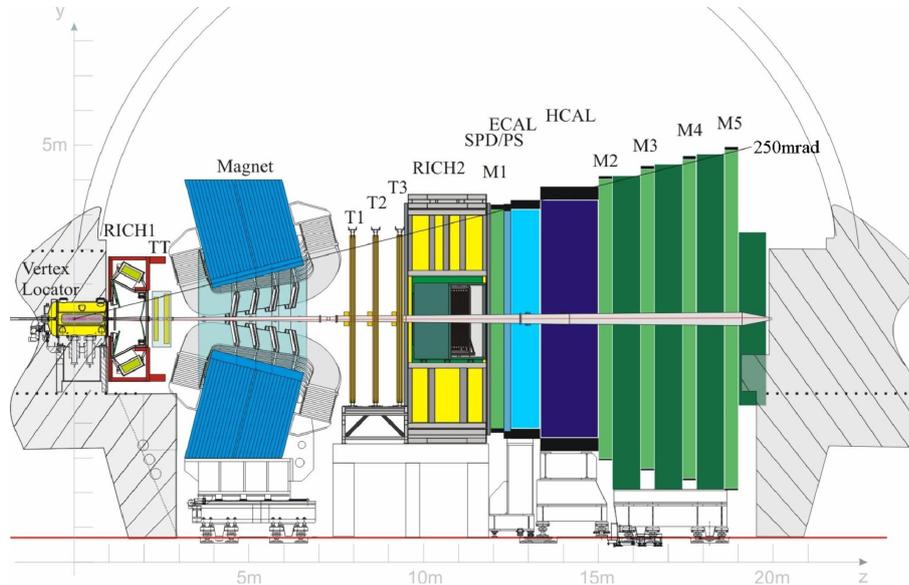


Figure 2.2: Schematic of the LHCb detector during Run 1 and 2 of the LHC. The (x, y, z) coordinate system is defined. Taken from [129].

The LHCb detector is a single-arm forward spectrometer instrumenting the pseudo-rapidity range $2 < \eta < 5$, in contrast to a hermetic design like ATLAS or CMS (primarily covering $\sim 0 < \eta \lesssim 3$) [120, 121]. This design choice is to optimise the geometrical acceptance to the production of b quarks in pp collisions with the study of b -hadron decays the primary design aim of LHCb. $b\bar{b}$ production proceeds primarily through gluon-gluon fusion and quark annihilation in pp collisions [130], where there is a high likelihood of a large momentum asymmetry between colliding protons producing a significant boost in the forward (or backward) direction [128]. The LHCb acceptance therefore contains $\sim 27\%$ of b quarks produced and $\sim 24\%$ of $b\bar{b}$ pairs [128], only a factor ~ 2 reduction on the hermetic experiments which cover a vastly higher total solid-angle [120, 121].

The LHCb detector is described in Figure 2.2, consisting of a high-precision tracking system, a sophisticated multi-part Particle Identification (PID) system and a flexible layered trigger designed to reconstruct large datasets of high-quality B^0 , B^+ , B_s^0 and Λ_b^0 hadrons to study both SM and new physics. LHCb is unique in its access to Λ_b^0 baryons, which are beyond the reach of other beauty experiments [50, 51]. LHCb has out-performed its specification with, for example, a rich charm, ElectroWeak and heavy-ion physics programme producing novel results beyond B -decays. This is enabled by the flexible trigger system used with lower thresholds than other experiments, possible with the reduced and constant \mathcal{L}_{inst} from luminosity levelling.

2.2 LHCb Tracking

2.2.1 VERtEX LOcator in Run 1 and 2

The VELO surrounds the pp -interaction region, performing the first measurement of charged particles produced in collisions at LHCb. It consists of 42 silicon-strip half-circle modules, arranged along a 1 m section of the beamline, with a greater concentration around the central $\sigma = 5.3$ cm interaction region. Each module consists of an R and Φ sensor, providing an (r, ϕ) coordinate measurement of a traversing particle “hit” with a pixel pitch (granularity of the sensor) ranging from $38 \mu\text{m}$ at the innermost region to $102 \mu\text{m}$ [128]. This corresponds to radial scaling of the sensitive area of each half-circle module with $r_{min} = 8.2$ mm to $r_{outer} = 42$ mm. Four pile-up modules are also installed in the backwards direction (negative z with respect to the interaction region) to remove high multiplicity events that would result in poor reconstruction efficiency.

The entire VELO detector is placed within a secondary vacuum separated by a thin 0.5 mm aluminium RF foil from the primary LHC vacuum [128]. This enables independence in operation, protecting the LHC vacuum from VELO components, as well as providing a shield from RF interference from the beam. However, during LHC beam injection the shield is insufficient to protect the modules from hadronic emission due to the unstable beam-profile, therefore each half of the VELO is retracted by 6 cm, only returning to the nominal r_{min} when stable beams are established. This complex closing system reduces radiation-damage while enabling instrumentation closer to the interaction region than any other LHC experiment [128].

Module pairs form 21 stations arranged in z to ensure that particles with $1.6 < \eta < 4.9$ traverse at least six stations, compensating for the minor hit-inefficiency of $\approx 0.6\%$ to measure a complement of hits that will reconstruct a corresponding track in the LHCb acceptance [128]. These initial “VELO tracks” are analysed for large coincidences to reconstruct Primary Vertices (PVs), the pp collision points. Smaller coincidences, displaced from the PV, are reconstructed as Secondary Vertices (SVs), the likely decay-vertex position of long-lived b and c hadrons. The VELO therefore provides a measurement of the Flight Distance (FD) and decay-time (τ) for candidate heavy-flavour hadrons, vital for isolating a pure sample of decays-of-interest with respect to the hadronic background present in pp collisions [128].

The VELO also measures the Impact Parameter (IP) of a reconstructed track, the distance between its trajectory and a given PV in the event, calculated at the z position of the PV. This is used to associate a particle to a specific PV and provides a measurement of likelihood of it originating from that PV, χ_{IP}^2 , vital for selecting displaced tracks to

form heavy-flavour candidates. χ_{IP}^2 is defined as the change in χ^2 of the PV fit if the given track is included. Proportional to $(IP/\sigma_{IP})^2$ it renders IP resolution as a critical VELO metric [24, 128]. It can approximately be parameterised as

$$\sigma_{IP} \approx \sigma_{\text{extrap}} \oplus \frac{\sigma_{\text{MSC}}}{p_T}, \quad (2.2.1)$$

where the track extrapolation term σ_{extrap} and multiple scattering term σ_{MSC} are determined by r_{min} , the amount of scattering material and the sensor hit-resolution [24]. The requirement to minimise this dictated the discussed Run 1 and 2 VELO design producing a resolution of $(15 + 29/p_T) \mu\text{m}$ for candidate “physics tracks” likely to be used in reconstruction of heavy-flavour candidates [128].

2.2.2 Tracking Stations and Magnet

In LHCb, the relative position between subdetectors is referred to as downstream or upstream, at a greater or smaller z by the direction of particle flow. The 4 Tm dipole magnet [128] is often used as the reference-point with the VELO and Tracker Turicensis (TT) comprising the upstream portion of the tracking system and tracking stations T1–3 downstream (see Figure 2.2). The aluminium room-temperature magnet deflects charged particles in the horizontal (x) detector plane, with charged particles requiring $p > 1.5 \text{ GeV}$ to remain in downstream acceptance of the T stations [128]. The deflection of tracks before and after the magnet enable a momentum measurement. As a dipole magnet, the polarity is swapped during each Run period, set to Mag Up (MU) and Mag Down (MD) for \sim equal periods of time, such that when the two datasets are averaged potential charge-dependent systematic biases can be cancelled, especially important for Charge-Parity (CP) violation measurements [123].

The upstream tracking station, TT, consists of silicon microstrip sensors with a pitch of $183 \mu\text{m}$ with an active area of 8 m^2 instrumenting the forward acceptance of the VELO at $z \sim 2.5 \text{ m}$ [128]. The two stations provide a second measurement of particles originating in the VELO, significantly reducing the native rate of reconstructed fake “ghost” tracks¹ to $\sim 6.5\%$ [128]. For long-lived particles, such as Λ and K_S^0 , $\sim \frac{1}{2}$ decay beyond the VELO [128], and the TT provides a vital first measurement enabling the determination of momentum.

Downstream of the magnet three tracking stations of 29.5 m^2 each reconstruct “T-Tracks” to be matched to objects formed in the VELO and/or TT to form long and down-

¹A track that is unlikely to be genuine, defined by having fewer than 70% of its hits from a single generated particle[131].

stream tracks as defined in Figure 2.3a via dedicated “track-seeding” algorithms [128]. Each T1–3 station is formed of a silicon-strip Inner Tracker (IT) 4 m² area to provide smaller granularity for the high particle multiplicity high- η region, with a larger gaseous straw drift-tube Outer Tracker (OT) in the lower occupancy region [128]. For reconstructed long tracks a relative momentum resolution of $\delta p/p = 0.5 - 1\%$ is achieved for particles with $p = 1.5 - 200$ GeV. This consistent performance across a wide momentum range propagates to versatile physics performance with a mass resolution of 14.3 MeV and 1727 MeV measured for J/ψ and Z^0 resonances respectively with Run 1 $\mu^+\mu^-$ calibration data [128].

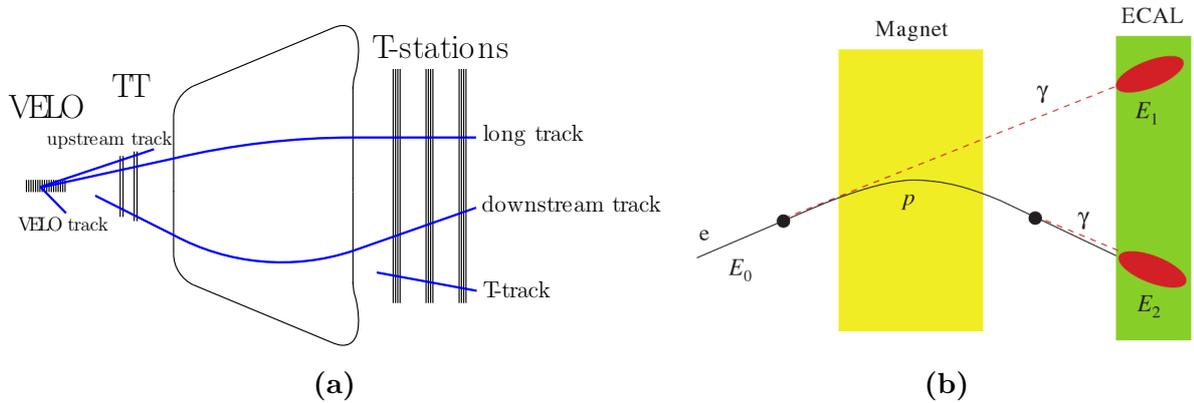


Figure 2.3: Track types and reconstruction effects in the LHCb detector. (a) Schematic of different LHCb track types assigned during Run 1 and 2 based on the information used to reconstruct a given track. Taken from [132]. (b) Effect of Bremsstrahlung radiation at different points in the detector and the “0 γ ” and “1 γ ” track categories this produces [123].

2.3 LHCb Particle Identification

The identity of particles is determined by a combination of specialised subdetectors that provide a complement of likelihood measurements for different particle hypotheses matched to each reconstructed track. This is crucial for flavour physics measurements, allowing separation between decay modes with similar topology of a given unstable hadron, such as $\Lambda_b^0 \rightarrow pK^- J/\psi$ and $\Lambda_b^0 \rightarrow p\pi^- J/\psi$.

As shown in Figure 2.2, the LHCb PID system consists of two Ring Imaging CHerenkov (RICH) detectors (up and downstream of the magnet), an Electromagnetic CALorimeter (ECAL) system, a Hadronic CALorimeter (HCAL) and five Muon chambers. Figure 2.4a describes how the presence or lack of a detector response in each sub-system can inform on the identity of the incoming particle, e.g. a muon with a track in the T-stations followed by a response in the RICH and Muon chambers. The PID system provides the only

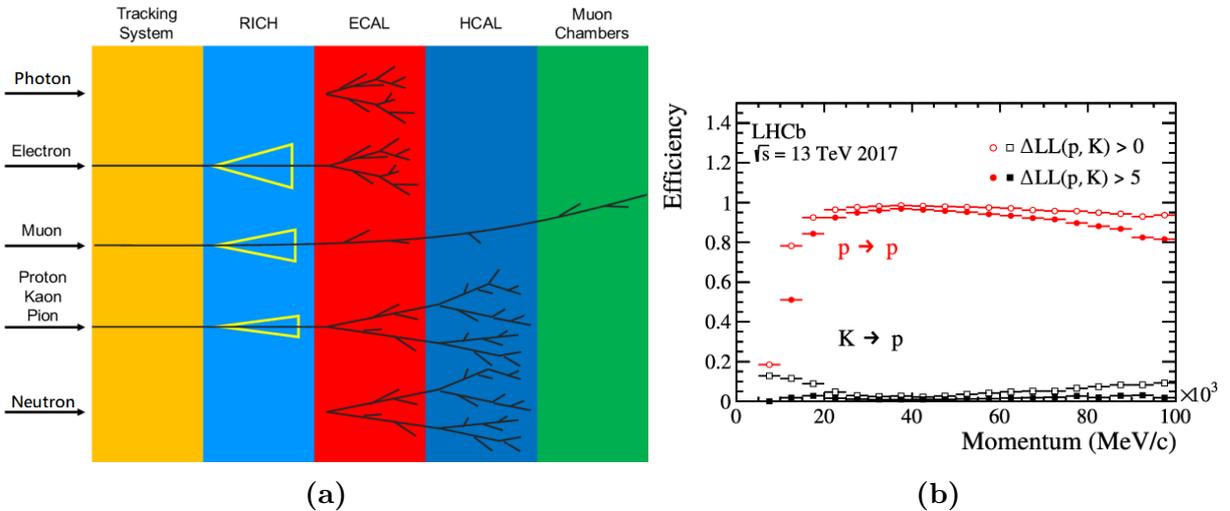


Figure 2.4: PID at LHCb. (a) Schematic of the LHCb PID system, displaying how different particles interact with the subdetectors. Adapted from [133]. (b) Overview of efficiency for the LHCb Run 2 RICH system to identify protons correctly and to mis-identify kaons as protons, as a function of track p at different PIDp working points. $\Delta LL(p, K) > 0,5$ are example cut values. Taken from [129].

measurement of neutral particles, γ and neutrons, with the missing track followed by a clear response in the ECAL and HCAL respectively implying their identity.

To distinguish between charged hadrons, RICH1 and RICH2 utilise Cherenkov radiation rings emitted at a specific angle dependent on the particle velocity to calculate an estimate for the particle mass, using the momentum from the matched track segment [128]. RICH1 covers the range $2 < p < 60$ GeV using higher refractive index radiators (C_4F_{10} and Aerogel, with the latter removed for Run 2 [129]) than RICH2, covering approximately $15 < p < 100$ GeV with its CF_4 radiator [128]. RICH2 has a reduced acceptance of $3 < \eta < 5$, prompting the use of binary acceptance variables such as “hasRich” in analyses [128]. HPDs read out the Cherenkov photons, reconstructing the corresponding rings and forming a Negative-Log-Likelihood (NLL) value for each charged particle-type hypothesis. A delta-log-likelihood (DLL) value is then computed between each particle-type and the corresponding pion hypothesis producing the discriminating variable “PIDx” (where $x = p, K, \dots$) which is subsequently used in analyses to tune efficiency and purity for selecting specific particle species. This allows calculation of specialised variables such as DLL_{pK} (PIDp-PIDk), vital for distinguishing, for example, $\Lambda_b^0 \rightarrow pK^- J/\psi$ and $B_s^0 \rightarrow K^+ K^- J/\psi$ candidates with the Run 2 performance described in Figure 2.4b. The drop in performance at low- p is due to protons and kaons only producing reconstructible Cherenkov rings above $p \approx 10$ GeV [129].

The ECAL system is composed of a Scintillating Pad Detector (SPD), a Preshower detector and the shashlik-type ECAL [128]. The Preshower initiates an EM shower in

electrons and photons but has a small hadronic radiation length, reducing the likelihood of a hadronic shower which could result in the mis-identification (mis-ID) of incoming hadrons. Hits in the SPD imply an electron interaction, providing a further distinguishing feature and simplifying the identification of electrons/photons from their showers in the ECAL. The ECAL, with alternating scintillator and lead-absorber plates provides a relative energy resolution of $(1 + 10/\sqrt{E})\%$ producing the only measurement of photon energy and an additional handle on electron energy as well as the initial momentum measurement. The precise energy resolution is crucial for electron reconstruction due to the frequent Bremsstrahlung emission described in Figure 2.3b, where emission before the magnet can result in a non-negligible fraction of the initial electron energy missing or within a distinct ECAL cell [128]. This will result in a under-reconstruction of the true electron p . For emission beyond the magnet region, the momentum measurement is unaffected and the forward boost on the Bremsstrahlung photon results in a high-likelihood of the $e + \gamma$ pair showering in the same ECAL cell [128].

To correct for Bremsstrahlung emission before the magnet, a recovery algorithm extrapolates the electron VELO/upstream track and searches for a photon shower with feasible energy in that area, adding P_γ to the initial electron reconstruction [128]. Multiple emissions are possible however which the algorithm can not reconstruct and random photons may be attached resulting in over-reconstruction of the momentum. This produces two distinct categories of electrons at LHCb, “1 γ ” and “0 γ ”, with and without Bremsstrahlung recovery, with both displaying consistently worse momentum resolution than other charged particles due to the missing energy or use of the ECALs poor energy-resolution, even if the recovery algorithm works as expected [128].

Hadrons begin showering in the ECAL and continue into the HCAL where they are entirely absorbed, providing the only measurement of neutral hadrons or an additional confirmation of h^\pm identity. The sampling device is made from iron and scintillating tiles, providing a fast response time for a hadronic shower [128], with its most common application as the hadronic component of the Level-0 (L0) hardware trigger (discussed in Section 2.4.1). The calorimetry system contributes to combined DLL variables, including information from the RICH, with binary acceptance variables “`InAcc(Ecal,Hcal,Prs,SPD)`” and “`HasCalo`” used to determine if the calorimeter response is valid for individual events in analysis.

The five Muon stations provide robust identification of muons that pass through the other PID systems with only minor, if any, interaction. M1 consists of fast Gas Electron Multiplier (GEM) detectors to manage the high particle-rate before the calorimeters, with M2–5 using Multi-wire Proportional Chamber (MWPC) covering an area of 435 m², the full LHCb acceptance, to select and identify penetrating muons [128]. Due to its fast

response the Muon chambers are used in the L0 trigger, with M1 providing an initial transverse momentum estimate $p_T^{L0}(\mu)$, with a tuneable threshold applied to control the event-rate in the experiment [128]. M2–3, with finer granularity, provide better position resolution, improving the $p_T^{L0}(\mu)$, while M4–5 are primarily for muon identification, with $p_T > 6$ GeV required to traverse both stations [128]. This produces two muon-only variables for analysis `HasMuon` and `IsMuon`, binary metrics to state whether a particle produced a response in the muon chambers and if the energy deposited was consistent with a high- p_T μ .

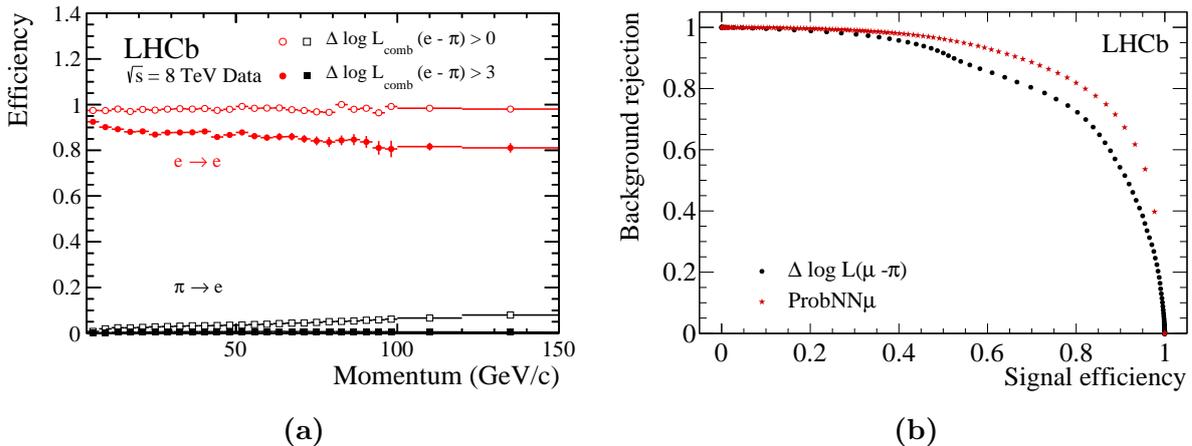


Figure 2.5: Leptonic PID performance at LHCb. (a) Analogous (mis)identification distribution to Figure 2.4b for electrons (with pions as the mis-ID), combining information from the RICH and ECAL systems during Run 1 (where 0 and 3 are example cut values). Taken from [128]. (b) Receiver Operator Characteristic (ROC) curve comparing μ - π separation power for the likelihood difference method and the Neural Network (NN) powered variable. Taken from [128].

Combined DLL variables are constructed with information from all PID sub-systems, implemented with the same strategy as the RICH DLL variables. This produces variables such as `PIDe`, providing tuneable separation between electrons and pions as shown in Figure 2.5a. As displayed, $h \rightarrow \ell$ mis-ID at LHCb is small, due to the distinct hadron response in the RICH compared to the expected behaviour in the ECAL/Muon chambers, and can be tuned to be negligible for a small reduction in selection efficiency. Further PID variables are constructed, called `ProbNNx`, based on a specialised, central NN taking input from all LHCb subdetectors. It is trained on simulation from each Run period, producing an output probability score for the particle hypothesis queried. The `ProbNN` variables are found to be more powerful at suppressing mis-ID background contributions [129] and are used in the analysis discussed in Chapters 3–5 of this thesis to efficiently select events. The `ProbNN μ` performance compared to `PID μ` is displayed in Figure 2.5b, showing consistently better background rejection for the high signal efficiencies desired in heavy-flavour analyses.

2.4 Data Acquisition and Offline Reconstruction

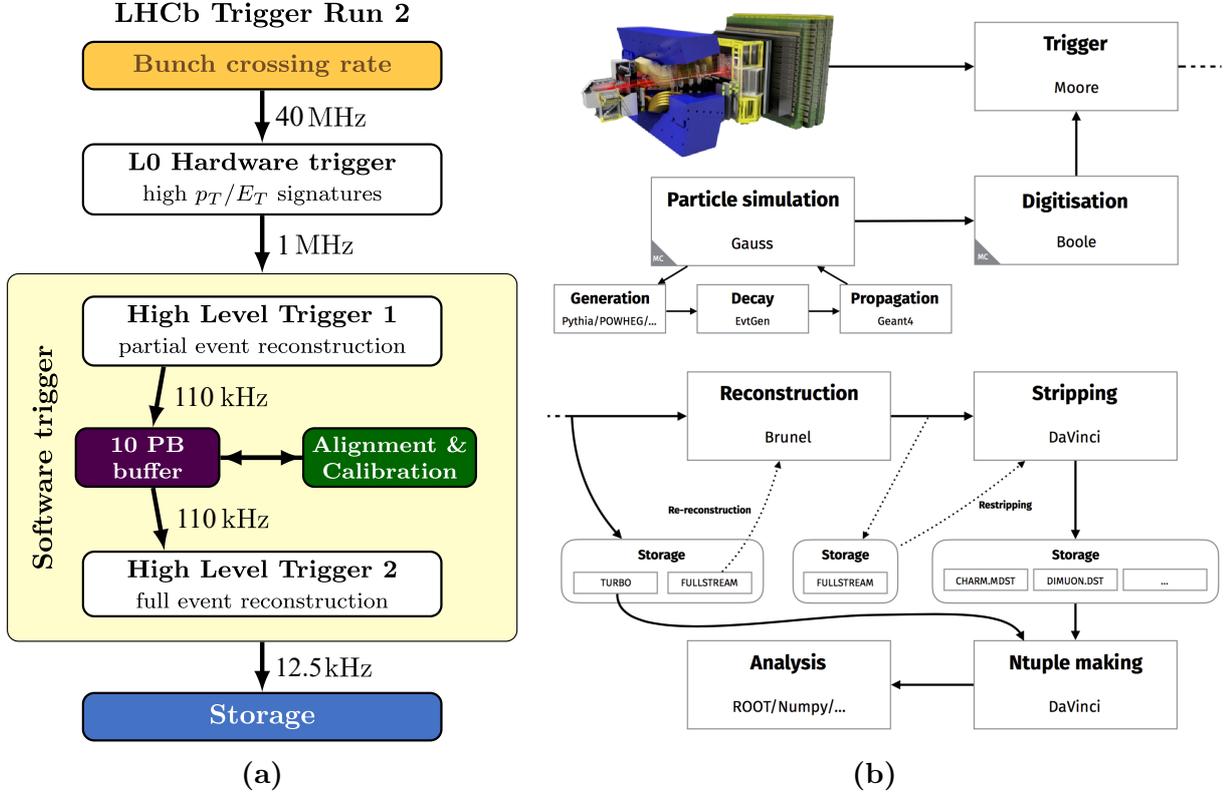


Figure 2.6: The LHCb data processing flow-diagrams for Run 2. (a) The trigger architecture. (b) The full data and simulation flowchart, from pp collisions to analysis. Taken from [132] and [134].

The LHCb Data Acquisition (DAQ) system is designed to allow data taking with minimal dead-time at the full LHC BX rate of 40 MHz [132]. Figure 2.6a describes the two-level trigger system employed to reduce the data-rate to $\mathcal{O}(10)$ kHz, optimised to select events containing signatures of the decays of b and c -hadrons [132]. The central data processing after the trigger stage in Figure 2.6b is referred to as “Offline” and are the stages required to prepare datasets for flexible use by LHCb analysts.

2.4.1 Trigger in Run 1 and 2

The first level of the trigger, L0, is implemented in hardware, making decisions within $4\ \mu\text{s}$ by retaining events with high- p_T or E_T particles using instant information from the calorimeters and muon stations [132]. While each event must pass at least one L0 threshold, separate L0 trigger-line response information is retained to allow offline curation of events with different muonic, EM and hadronic signatures.

Events passing L0 are passed to the software High Level Trigger (HLT), split into two stages, HLT1 and HLT2. The first stage uses reconstructed tracks to perform inclusive selection on signatures of individual tracks or track pairs, using the initial PV reconstruction to preference displaced tracks that are more likely to originate from b or c -hadron decays [132]. Single-track signatures are subject to a higher threshold than two-track composites. During Run 1, HLT1 was limited to tracks with $p_T > 1.2$ GeV [128], limiting the reach of charm and strange physics analyses. In Run 2 the CPU-bandwidth was increased by a factor ~ 2 and the buffer system improved allowing further processing of events during inter-fill periods and alignment performed in real-time [132], reducing the limit to $p_T > 0.5$ GeV [132].

During Run 2, HLT1 events are passed to HLT2 for full event reconstruction, with a wide range of inclusive and exclusive trigger lines allowing near-complete online selection. This replaced the Run 1 HLT2, limited by its requirement for offline alignment and lack of online PID information, providing only general trigger decisions [128]. This distinction between the trigger system in Run 1 and 2 results in different trigger lines available and corresponding performance, with analyses often using partitioned datasets.

All trigger lines in LHCb have two primary implementations describing the given response: “Trigger-On-Signal (TOS)” and “Trigger-Independent-of-Signal (TIS)”. The first is when the given trigger decision was fired by a particle or composite within the decay chain of the “signal candidate” (the feature of the event being reconstructed), and the latter when the trigger is caused by an independent object or track. This separation enables determination of the live data taking trigger efficiency,

$$\varepsilon_{\text{trig}} = \frac{N_{\text{TOS}}}{N_{\text{TOS}} + N_{\text{TIS \& TOS}}}, \quad (2.4.1)$$

where N_X refers to the yield of events classified as each trigger-response type [135].

2.4.2 Offline Processing

Events passing the trigger stage are recorded to tape for further reconstruction in both Run 1 and 2. Reconstructed tracks, often from inclusive trigger lines, can be combined with other features to form complex multi-body objects with further selection allowing categorisation into “Stripping Lines” that curate events into those likely to contain events of interest for a given analysis. Stripping lines can be trigger-line agnostic, maximising the number of events available for a rare FCNC search, for example.

For flexible analysis, the output of stripping lines are processed into offline Ntuples that are curated to contain a plethora of subdetector response information, individual

particle variables and reconstructed metrics such as the invariant mass and χ_{IP}^2 of candidates of interest on an event-by-event basis. This enables analysts to manipulate, further select and extract physics results in novel ways, diversifying the output of the LHCb experiment. For example, the work discussed in Chapters 3–5 of this thesis primarily used data analysis tools from the `ScikitHEP` ecosystem [136] as an alternative to the commonly used `ROOT` framework [137].

2.4.3 Simulation at LHCb

Simulation is essential to model the acceptance and resolution of LHCb as well as for determining the efficiency of selection for a specific physics channel under study. To deliver this LHCb uses the `GAUSS` framework [138], generating underlying pp collisions using `PYTHIA` [139, 140], controlling the decay of candidates-of-interest with `EVTGEN` [141] and propagating particles through a model of the detector in `GEANT4` [142].

While `PYTHIA` produces unbiased pp -events, for most analysis use cases `EVTGEN` is tuned to only accept events containing a viable “head” b or c -hadron with decay products produced within the LHCb acceptance. In addition, candidates that would be valid in the inverse LHCb acceptance ($-5 < \eta < -2$) are mirrored and the “Decay File” that controls the candidate decay can be further tuned to apply specific physics models and require certain kinematics, ensuring efficient production of simulation [138]. Properties of generated candidates before particle propagation and without simulation of material interaction and reconstruction are referred to as “generator-level” or “truth” information, retained alongside the reconstructed `Ntuples` for later analysis.

All particles produced in the pp event, including the candidate decay, are propagated through a detailed simulation of the LHCb detector in `GEANT4`, adjusted and aligned to represent the near-exact detector in a given year [142, 143]. This records particle traversal of sensitive detectors, emulates the curvature of charged particles in the B -field and simulates material interaction effects including multiple-scattering and Bremsstrahlung. To mimic the response of subdetectors, a digitisation step inflicts hit inefficiencies, dead channels and radiation damage effects [144] while converting the simulation to the input format ready for the LHCb trigger system. This allows the subsequent trigger, reconstruction and `Ntuple` creation for simulation to be aligned with real data as described in Figure 2.6b.

The discussed simulation generation and processing chain ensures that simulated “Monte Carlo (MC)” samples are the best available proxy for real data. But the emulation is not perfect. Often simplified candidate decay models are used for generation when specific physics models are not well-defined or too complex. In addition, certain subde-

tectors such as the RICH and ECAL produce large number of Cherenkov photons and complex shower objects with simulation limited by modelling of these processes and the computing power available [138]. This limits the accuracy of the simulated PID response at LHCb. Data-driven reweighting schemes are hence employed to tune the simulation to the required alignment with real data, an example of which is presented in Section 4.2. Simulation with unbiased modelling is often preferred to give analysts control of MC using reweighting suites, especially in searches for NP where the behaviour of the target decay mode is unknown. Furthermore, as the LHCb detector is upgraded, detailed in Chapter 6 of this thesis, the \mathcal{L}_{inst} is projected to increase by up to a factor ~ 50 [24], amplifying ν (the number of simulated pp collisions per BX) and dramatically increasing the particle rate to simulate, suggesting that faster, simpler simulation with offline correction will be desirable.

Chapter 3

A Search for Charged Lepton

Flavour Violation in

$\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ decays at LHCb

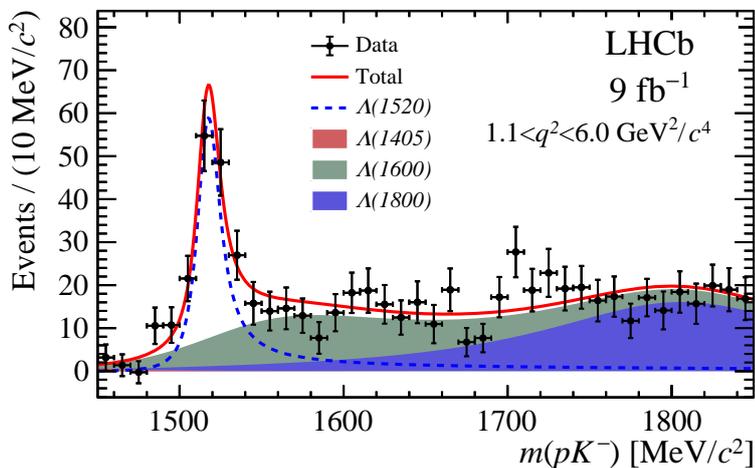
The following chapters will introduce the first search for the Charged Lepton Flavour Violation (cLFV) decay $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$. As discussed in Section 1.3, searches for cLFV are a clean test-bench to search for NP, with no observable mechanism in the SM that can create the signature. Therefore any significant signal measured in this decay mode would be unequivocal evidence for BSM physics.

3.1 Motivation for $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ LFV Search

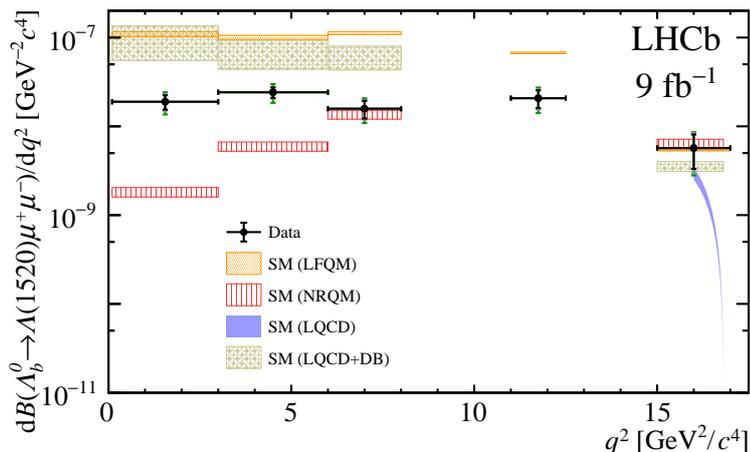
This analysis builds on recent efforts and follows a similar methodology to other searches for LFV at LHCb in inclusive $\mu^\pm e^\mp$ final states. Recent examples of these are the search for $B^+ \rightarrow K^+ e^\mp \mu^\pm$ [113] and $B_{(s)}^0 \rightarrow X \mu^\pm e^\mp$, ($X = K^{*0}, \phi$) [112], as well as the ongoing search for $\Lambda_b^0 \rightarrow \Lambda e^\mp \mu^\pm$ [42, 119]. The inclusive $\mu^\pm e^\mp$ dilepton final state is an attractive mode at LHCb: muon identification and reconstruction are reliably-high quality and electrons have distinctive signatures but due to detector effects such as Bremsstrahlung, they have lower reconstruction efficiency than muons. Decays of τ leptons are not trivial to reconstruct due to the unobserved energy carried away by neutrinos.

In parallel with the motivation for an LFV search, there has also been recent theoretical and experimental interest in the rare decay $\Lambda_b^0 \rightarrow p K^- \ell^+ \ell^-$, in particular on the exclusive mode $\Lambda_b^0 \rightarrow \Lambda(1520) \ell^+ \ell^-$ containing the dominant and narrow excited-strange baryon $\Lambda(1520)$. This $b \rightarrow s \ell^+ \ell^-$ FCNC process with b -baryons is attractive experi-

mentally due to the four “long” tracks in the final state that arise from the very-short lifetime of the strong $\Lambda^* \rightarrow pK$ resonances. This compares favourably with the analogous $\Lambda_b^0 \rightarrow \Lambda \ell^+ \ell^-$ mode which require reconstruction of downstream tracks to efficiently measure long-lived Λ baryons.



(a)



(b)

Figure 3.1: Measurement of the differential $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-)$ at LHCb with the Run 1 and 2 dataset [105]. (a) A Λ^* component fit to the M_{pK} invariant mass spectrum in $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ decays. (b) The measured differential $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-)$ as a function of the squared dilepton invariant mass q^2 , compared to different SM predictions from theory. Both taken from [105]. The predictions shown correspond to those derived using form factors from: the non-relativistic (NR) [145] and light-front (LF) [146] quark model (QM) as well as the lattice QCD [147] and dispersive bound (DB) [148] formulation. The blue LQCD prediction is only available for $q^2 > 16 \text{ GeV}^2$ and is given as a rate average due to the rapid changing from the kinematic limit on dilepton mass in this region.

Studies of the $\Lambda_b^0 \rightarrow pK^- \ell^+ \ell^-$ decay include the first measurement of the LFU ratio R_{pK} (as discussed in Section 1.2) which measured the branching fraction ratio of $\Lambda_b^0 \rightarrow$

$pK^- \mu^+ \mu^-$ to $\Lambda_b^0 \rightarrow pK^- e^+ e^-$ and found it to be consistent with the SM prediction [76]. While this was an inclusive study, covering the proton-kaon invariant mass range $1432 < M_{pK} < 2600$ MeV, there has been theoretical interest in an exclusive $R_{\Lambda(1520)}$ measurement as well as efforts to calculate the hadronic form-factors and predict the SM branching fraction for the exclusive $\Lambda_b^0 \rightarrow \Lambda(1520)\ell^+\ell^-$ modes [64, 148]. The predictions have since been tested by the LHCb measurement of the $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ differential branching fraction, as shown in Figure 3.1, which observes deviation from select theory predictions in many of the q^2 (dilepton invariant mass squared) bins studied [105]. Differences between the data and predictions are however comparable in size to the discrepancies between the theoretical models, therefore it is difficult to draw robust conclusions without improved understanding of the models themselves.

There is also theoretical and experimental focus on studying the angular distribution of $\Lambda_b^0 \rightarrow \Lambda(1520)\ell^+\ell^-$ with sensitivity to NP models predicted [145, 149]. One theoretical study has predicted the $\Lambda_b^0 \rightarrow pK^- \ell^+\ell^-$ angular distribution including all Λ^* resonances up to spin- $\frac{5}{2}$ [150], determining that the $\Lambda(1520)$ resonance is modified the most by NP, confirming it as the most suitable candidate for a BSM search. This has prompted the ongoing analysis at LHCb to experimentally measure the $\Lambda_b^0 \rightarrow \Lambda(1520)\ell^+\ell^-$ angular distribution [151].

The discussed collection of intriguing results and ongoing analyses motivates an LFV search in $\Lambda_b^0 \rightarrow pK^- \ell\ell'$ and specifically in the exclusive decay mode $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$, rather than the inclusive mode $\Lambda_b^0 \rightarrow pK^- \mu^\pm e^\mp$ which would include the full M_{pK} spectrum of resonances of varying spin as seen in Figure 3.2. While the inclusive mode would potentially provide a larger sample of signal candidates for the search, the ability to quote an exclusive measured branching fraction or limit with the $\Lambda(1520)$ resonance is preferred as it provides a clean and well-defined input for theory. This follows precedence of other LFV searches at LHCb such as $B_{(s)}^0 \rightarrow X\mu^\pm e^\mp$, ($X = K^{*0}, \phi$) setting limits on the exclusive K^{*0} and ϕ branching fraction rather than with the inclusive $M_{K^+\pi^-}$ and $M_{K^+K^-}$ spectrum [112].

At the LHC, the production of Λ_b^0 baryons is abundant, and LHCb provides a unique opportunity to perform studies into baryonic LFV, which have had only minor exploration in the past including the ongoing $\Lambda_b^0 \rightarrow \Lambda e^\mp \mu^\pm$ analysis at LHCb and a previous study of forbidden charmed baryon decays at BaBar which set limits on decay modes such as $\Lambda_c^+ \rightarrow p\mu^\pm e^\mp$ [152]. These factors together motivate this search, with the ultimate goal of the analysis to observe the decay $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ or set the first experimental upper limit on the branching fraction to aid in the constraining of BSM models such as those discussed in Section 1.3.

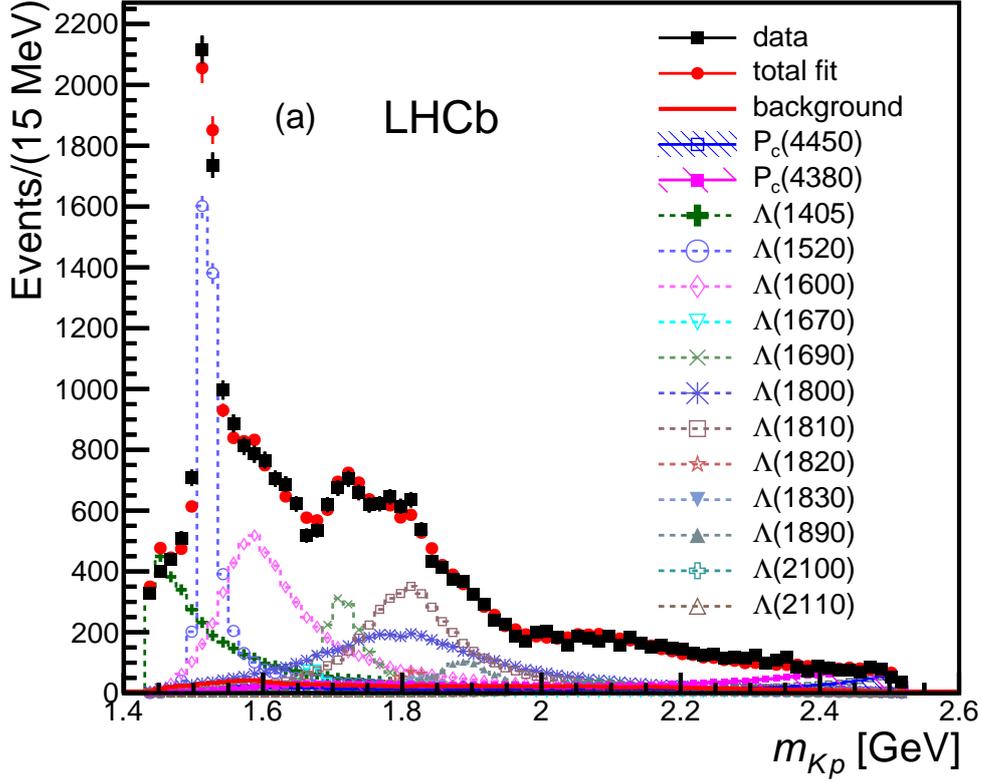


Figure 3.2: Measured M_{pK} distribution in $\Lambda_b^0 \rightarrow pK^- J/\psi$ data from the Run 1 LHCb dataset [49], overlaid with the fitted Λ^* components, including the dominant $\Lambda(1520)$ resonance.

3.2 Analysis Strategy

The LFV search will be achieved by measuring the “signal” mode $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ with respect to a well-known “control” mode, $\Lambda_b^0 \rightarrow pK^- J/\psi (\rightarrow \mu^+ \mu^-)$. The control mode has been chosen due to its large, well-measured branching fraction, same parent Λ_b^0 and similar topology/kinematics due to the four final state long tracks, with the primary difference the electron in the signal mode. The equation that underpins the analysis is given as

$$\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp) = \frac{N_{Signal}}{\varepsilon_{Signal}} \frac{\varepsilon_{Control}}{N_{Control}} \frac{\mathcal{B}(\Lambda_b^0 \rightarrow pK^- J/\psi) \mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)}{\mathcal{B}(\Lambda(1520) \rightarrow pK^-)}, \quad (3.2.1)$$

where N_{Signal} and $N_{Control}$ are the signal and control mode yields in data and ε_{Signal} and $\varepsilon_{Control}$ are the total selection and reconstruction efficiencies. The stated branching fractions provide the reference from which the signal branching fraction is measured, ensuring alignment of the definitions for N_{Signal} and $N_{Control}$. This equation can hence be simplified to

$$\mathcal{B}_{Signal} = \frac{N_{Signal}}{\varepsilon_{Signal}} \frac{\varepsilon_{Control}}{N_{Control}} \mathcal{B}_{norm} = \alpha N_{Signal}, \quad (3.2.2)$$

Table 3.1: Summary of the known PDG values for the branching fractions used in Equation 3.2.1 where the first uncertainty is statistical and the second systematic [41]. The latest LHCb result is also quoted where applicable. The last result is a fit fraction from an amplitude fit to $\Lambda_b^0 \rightarrow pK^- J/\psi$ decays [49] and can be used to estimate $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)J/\psi)$ for which there is no formal branching fraction measurement published.

Decay	Branching Fraction ($\mu \pm \sigma_{stat.} \pm \sigma_{sys.}$)	
	PDG Average	LHCb Latest
$\mathcal{B}(\Lambda_b^0 \rightarrow pK^- J/\psi)$	$(3.2^{+0.6}_{-0.5}) \times 10^{-4}$	$(3.17 \pm 0.04^{+0.57}_{-0.45}) \times 10^{-4}$ [153]
$\mathcal{B}(J/\psi \rightarrow \mu^+ \mu^-)$	$(5.961 \pm 0.033)\%$	-
$\mathcal{B}(\Lambda(1520) \rightarrow pK)$	$(22.5 \pm 0.5)\%$	-
$F(\Lambda(1520) \rightarrow pK^- / \Lambda^* \rightarrow pK^-)$	-	$(18.93 \pm 0.52 \pm 3.89)\%$ [49]

where α is a normalisation constant that scales the measured signal yield to the signal branching fraction. For the normalisation branching fractions the world-average values from the latest Particle Data Group (PDG) version are used, stated in Table 3.1.

The control mode, $\Lambda_b^0 \rightarrow pK^- J/\psi (\rightarrow \mu^+ \mu^-)$, was chosen over other options including $\Lambda_b^0 \rightarrow \Lambda(1520) (\rightarrow pK^-) J/\psi (\rightarrow \mu^+ \mu^-)$ and $\Lambda_b^0 \rightarrow pK^- J/\psi (\rightarrow e^+ e^-)$. The first alternative is a closer analogue to the signal mode, including the $\Lambda(1520)$ resonance, but extracting an $N_{Control}$ for this decay is not trivial and requires a tight cut or complex fit in the M_{pK} spectrum. This would either significantly reduce the control mode statistics or add uncertainty to the measurement of $N_{Control}$ when the analysis focus should be on the signal mode, with the control mode providing a precise and well-measured reference point. In addition, the branching fraction of $\Lambda_b^0 \rightarrow \Lambda(1520)J/\psi$ has not been exclusively measured and would rely on a fit fraction with relative large uncertainty (see Table 3.1). The second alternative, $\Lambda_b^0 \rightarrow pK^- J/\psi (\rightarrow e^+ e^-)$, would allow some insight and possible cancellation of systematic effects from the reconstruction of electrons but the difficulty of reconstructing two electrons and the subsequent degradation of the Λ_b^0 mass resolution outweighs this. Furthermore, with the chosen control mode, the same trigger strategy can be used as the signal mode.

To optimise sensitivity to the branching fraction in Equation 3.2.1, the analysis aims to maximise the signal efficiency while minimising the contamination from backgrounds, ensuring a pure and precise measurement of N_{Signal} . The efficiencies, ε_{Signal} and $\varepsilon_{Control}$ are determined using Monte-Carlo (MC) simulation of the $\Lambda_b^0 \rightarrow \Lambda(1520) (\rightarrow pK^-) \mu^\pm e^\mp$ and $\Lambda_b^0 \rightarrow pK^- J/\psi (\rightarrow \mu^+ \mu^-)$ processes. Simulated samples are corrected using a series of data-driven techniques to ensure that the efficiencies measured are as representative of real data as possible.

The control mode yield, $N_{Control}$, has a sufficiently high branching fraction that it

can be accurately determined through a fit to the invariant mass of the A_b^0 candidate in data. The signal yield, N_{Signal} is the key unknown of this analysis and therefore is not measured until all other aspects of the analysis are complete and approved. This is ensured by blinding a “signal region” in data by completely removing the candidates in the $M(A_b^0)$ invariant mass range of

$$(5200 < M(pK^- \mu^\pm e^\mp) < 5800) \text{ MeV}, \quad (3.2.3)$$

preventing any potential bias when developing the analysis selection chain or during the construction of the invariant mass fits. The inverted regions produced as a result of the blinding are denoted as the lower ($M(A_b^0) < 5200 \text{ MeV}$) and upper ($M(A_b^0) > 5800 \text{ MeV}$) mass sidebands. These background-dominated regions are used to test the effectiveness of selections at removing background, by assuming that the reduction in the upper and lower sidebands can be interpolated into the signal region. The two sidebands are hence often combined to perform “blinded fits” that can estimate the level of background in the signal region.

3.2.1 Data and Simulation Samples

The LHCb dataset used in the analysis corresponds to over 95% of the 9 fb^{-1} recorded in Run 1 (3 fb^{-1}) and Run 2 (6 fb^{-1}) respectively, where data collected during 2015 (0.3 fb^{-1}) is excluded due to the low luminosity and problems related to the trigger in this year. This results in a total luminosity of 8.7 fb^{-1} .

The analysis is split into four categories according to differences in kinematics and reconstruction of candidates across the dataset, observed to affect the efficiency. The partition is made by data taking period, Run 1 or Run 2, and on whether the electron in the signal decay has recovered a Bremsstrahlung photon: “without Brem.” (0γ) and “with Brem.” (1γ). The split by run is motivated by the large change in CoM energy, with b -hadrons in Run 2 produced with a greater boost, as well as the change in trigger strategy (discussed in Section 2.4.1). The split by Bremsstrahlung category is due to the distinct quality of the electron reconstruction and identification when a Bremsstrahlung photon is recovered (as outlined in Section 2.3), significantly altering the signal shape in $M(A_b^0)$. This difference also leads to distinct backgrounds in the two categories, therefore considering them separately improves the description of the mass distributions. Although the discussed components of Equation 3.2.1 are determined individually for each category in the analysis, they are treated collectively in a combined fit for the final analysis result.

3.2.1.1 Dataset to Model Combinatorial Background

While the sidebands provide information about the blinded signal region, the behaviour of the background in this region remains an assumption and the split mass range can lead to unstable invariant mass fits. To aid with this, a leptonic Same Sign (SS) dataset, $\Lambda_b^0 \rightarrow pK^- \mu^\pm e^\pm$, is constructed, analogous to the Opposite Sign (OS) signal mode, to act as a background-only proxy dataset. These can be studied without any blinding since the total charge creates an unphysical Λ_b^0 candidate.

The SS dataset is treated, after the same selection process as the OS datasets, as originating from random mixtures of tracks only and not from the decay of interest or any exclusive background, referred to as “combinatorial background”. The $\Lambda_b^0 \rightarrow pK^- \mu^\pm e^\pm$ dataset (subsequently referred to as μe SS) is, therefore, able to provide information about the excess physical backgrounds above the combinatorial level in the μe OS datasets which, by comparing the OS and SS invariant mass variables, is used in Chapter 4 to monitor the event selection. Creating a background-only dataset in this way is feasible at LHCb due to the negligible rate of charge mis-assignment and the excellent momentum resolution, which is maintained to 1% for 200 GeV particles.

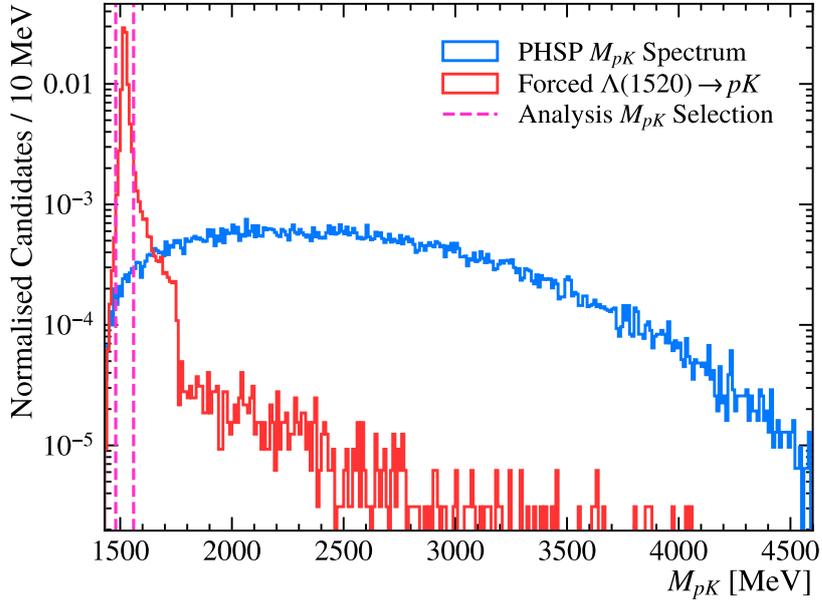


Figure 3.3: Different options for simulation of the $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ signal mode: where the pK pair must originate from a $\Lambda(1520)$ resonance (in red) and where the decay is according to a flat phase-space model (in blue). The selection applied in the analysis is overlaid.

3.2.1.2 Simulation

While the sidebands and the μe SS proxy datasets are used to test selection and evaluate the level of background in the analysis, MC simulation is used to evaluate the corresponding effect on the signal. Simulated samples are used to measure ε_{Signal} and $\varepsilon_{Control}$, optimise the selection chain, train the MVA, perform mass fits to extract shape parameters and generally act as a representation of the desired signal when designing components of the analysis.

Dedicated simulation samples were prepared for this analysis, corresponding to $\Lambda_b^0 \rightarrow \Lambda(1520)(\rightarrow pK^-)\mu^\pm e^\mp$ and $\Lambda_b^0 \rightarrow pK^- \mu^\pm e^\mp$ with the first requiring the hadrons to originate from the $\Lambda(1520)$ resonance, while the second forms an M_{pK} spectrum following the available phase-space according to the general decay rate equation (see Equation 1.2.5) with a constant matrix element. A comparison of the two simulated samples can be seen in Figure 3.3, produced during the initial prototyping of the analysis.

The advantage of $\Lambda_b^0 \rightarrow pK^- \mu^\pm e^\mp$, without the forced $\Lambda(1520) \rightarrow pK$ resonance, is that it can be reweighted using data-driven approaches to better emulate the expected M_{pK} spectrum for a SM $\Lambda_b^0 \rightarrow pK^- \ell^+ \ell^-$ decay seen in Figure 3.2 and 3.1a. However, the decay mechanism for the hypothetical $\Lambda_b^0 \rightarrow pK^- \mu^\pm e^\mp$ decay is unknown, and there is no reason to expect it would be identical to that seen in the $\Lambda_b^0 \rightarrow pK^- J/\psi$ or $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ data [105, 154]. Furthermore, as the analysis is searching for $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$, a simulated sample that directly produces the $\Lambda(1520)$ resonance in the M_{pK} spectrum is justified. As seen in Figure 3.3 it is also far more efficient to generate $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ directly considering the tight M_{pK} selection applied in Section 4.1.5.1.

The primary simulation Ntuples for the analysis are produced with ~ 20 million events each for the signal and control mode. Generator-level tuples are also retained for access to the complete truth information of the initial generated sample.

3.3 Summary of Analysis Components

The analysis selection chain can be summarised as:

Stripping Selection A loose selection applied when building candidates and derivation of the Ntuples.

Fiducial Selection Kinematic cuts to align the analysis data to external calibration datasets.

Pre-Selection Invariant mass selection and resonance vetoes to remove contamination from possible backgrounds.

Semileptonic Vetoes Selection to remove specific background semileptonic decays.

Multi-Variate Analysis (MVA) Selection A neural network trained to reduce combinatorial background.

Particle Identification (PID) Tightening the identity requirements on each particle species, further removing mis-identification backgrounds.

Multi-Candidate Removal Ensuring that only a single Λ_b^0 candidate is measured per event, inline with the ε calculations.

The given order is adhered to throughout the analysis, described in detail in Chapter 4. Mis-modelling and mis-reconstruction of the MC samples are corrected using data-driven techniques after the “Fiducial” selection stage (described in Section 4.2), amending known data-MC differences in kinematic, topological and multiplicity distributions. This ensures the simulation is as close a representation of real data as possible when used for optimisation of the subsequent selection stages and invariant mass fits.

After selection, a background study is constructed to estimate the possible remaining contributions from individual contaminants in the signal mode, using a combination of MC-derived efficiencies and measured branching fractions (presented in Section 5.1). Any of these contaminants with the potential to have a significant contribution in the final signal dataset are included in the signal mode invariant mass fit along with the remaining “combinatorial” background component. The combinatorial component is fixed from a fit to the μe SS distribution with the MVA initially loosened before Gaussian-constraining shape parameters and tightening the MVA to ensure fit stability (see Section 5.4.2). To determine the high-statistics control mode yield a simpler approach can be adopted with an invariant mass fit to the $\Lambda_b^0 \rightarrow pK^- J/\psi$ dataset with a floating combinatorial background and fixed exclusive background components, detailed in Section 5.3.2.

Upon full selection of the signal mode in data, with potential backgrounds determined and all other components of Equation 3.2.1 measured, the presence of the (unblinded) $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ signal in data would be determined using a simultaneous fit of all categories to extract a value for $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp)$. If a significant signal yield is measured, the analysis would present the first evidence for charged LFV and be a signature of NP. If no significant signal is observed, a confidence-interval scan will be performed using the CL_s method to obtain the first upper-limit for this decay mode [155]. At the time of writing, this analysis is still blinded, therefore a blind CL_s scan is performed to establish a blinded upper-limit instead, and presented in Section 5.5.

Chapter 4

Event Selection and Corrections to Simulation

4.1 Event Selection

The aim of the cut-based event selection is to build a pure dataset of high quality signal candidates, maintaining a high ε_{Signal} while reducing background contamination and aligning the selection with external samples used for simulation correction. In an attempt to maximise the cancellation of systematic effects in the efficiency ratio in Equation 3.2.1, the selection is aligned between the signal and control mode when possible, with the few necessary differences highlighted.

4.1.1 Stripping Selection

The `Bu2LLK.meLine` is used for the signal mode, and its selection criteria are summarised in Table 4.1. This stripping line applies relatively loose selection to candidates and hence yields a large initial sample with only minor bias towards the topology and identity of the signal mode, optimum for an analysis searching for a forbidden decay. Furthermore, by using the analogous $\mu^+\mu^-$ version of the stripping line nearly identical stripping selection can be applied to the control mode $\Lambda_b^0 \rightarrow pK^- J/\psi (\rightarrow \mu^+\mu^-)$, providing the best initial selection alignment possible, with the differences marked in Table 4.1.

In most simulation samples (including the primary signal and control mode samples), the PID requirements from the stripping line are removed by reprocessing, to be corrected and replaced by a data-driven approach (see Section 4.4), due to the imperfect simulation of the particle identification detectors in LHCb simulation.

Table 4.1: Stripping Line cuts for Bu211K_meLine, used for the signal and control mode. PID selection components removed by re-stripping are marked with †. “Det.=” refers to the PID detector response required by each particle and other acronyms defined in the glossary.

	Selection Type			
	Mass	Kinematics	Quality	PID
p		$p > 2000$ MeV $p_T > 250$ MeV	$\chi_{\text{IP}}^2 > 9$ Ghost Prob. < 0.35	$\text{DLL}_{p\pi} > -5^\dagger$ $\text{ProbNNp} > 0.05^\dagger$ Det. = 'RICH'
K		$p > 2000$ MeV $p_T > 250$ MeV	$\chi_{\text{IP}}^2 > 9$ Ghost Prob. < 0.35	$\text{DLL}_{K\pi} > -5^\dagger$ $\text{ProbNNk} > 0.05^\dagger$ Det. = 'RICH'
μ		$p_T > 350$ MeV	$\chi_{\text{IP}}^2 > 9$	HasMuon & IsMuon † Det. = 'MUON'
e		$p_T > 350$ MeV	$\chi_{\text{IP}}^2 > 9$	$\text{PIDe} > 0^\dagger$ Det. = 'CALO'
pK	$M_{pK} < 5620$ MeV	$p_T > 400$ MeV DOCA < 30 mm	$\chi_{\text{vtx}}^2 < 25$ $\chi_{\text{vtx}}^2 < 25$	
$\ell\ell^{(\prime)}$	$m_{\ell+\ell^-} > 100$ MeV $m_{\ell+\ell^-} < 5500$	$p_T > 0$ MeV DOCA < 30 mm	$\chi_{\text{vtx}}^2/\text{ndf} < 9$ $\chi_{\text{FD}}^2 > 16$	
Λ_b^0	$ m_{B^+} - M(\Lambda_b^0) < 1500$ MeV	DIRA > 0.9995	$\chi_{\text{vtx}}^2/\text{ndf} < 9$ $\chi_{\text{IP}}^2 < 25$ $\chi_{\text{FD}}^2 > 100$	
Global Event Cuts: nSPDHits < 600 (450) for Run 1 (2)				

4.1.2 Truth Matching of the Simulation

Truth Matching (TM) is required to select true $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ decays from within the simulation samples. There are two possible methods: ID matching all particles including the parent and grandparent of final state particles or using built-in variables that split MC into “background categories” that compare the truth-level decay hierarchy with the reconstruction-level hierarchy in DAVINCI and break down the simulation into sub-categories relating to their truth-level contents [156]. The second method is preferred

as it automates TM with a standard LHCb method that is generally used across the collaboration.

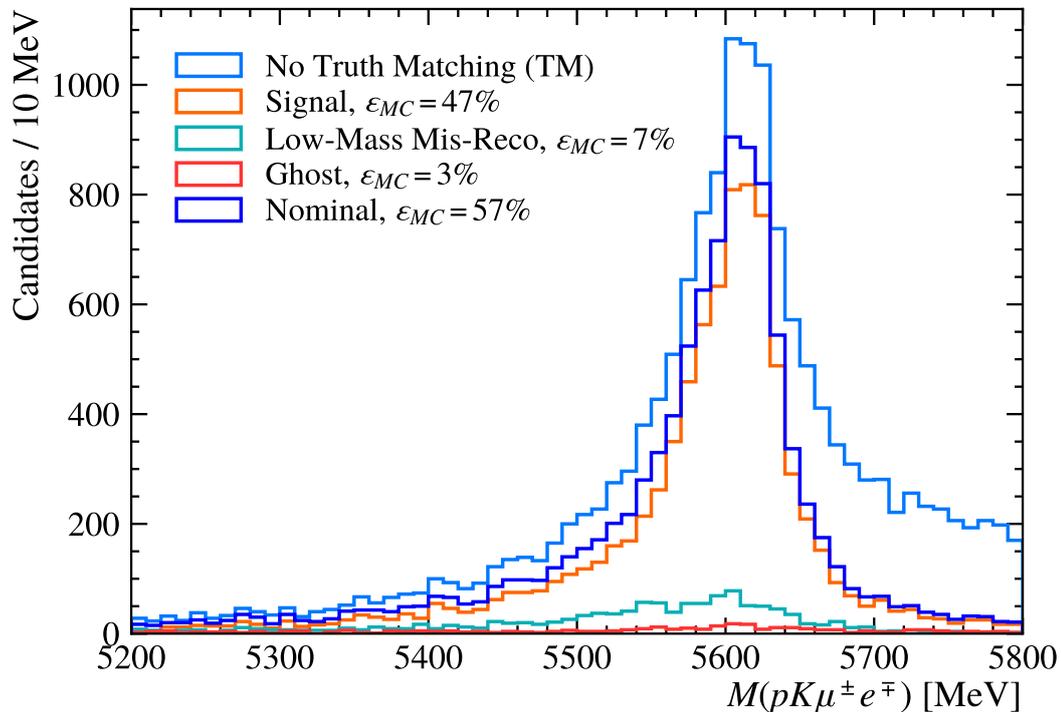


Figure 4.1: Reconstructed $M(\Lambda_b^0)$ in a representative sample of $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ simulation, comparing signal shape for different truth matching categories, defined in Section 4.1.2, as well as the initial shape without truth matching.

The choice was made to align with the topologically similar analysis, R_{pK} [157]. This required one of the following conditions to be met by the signal event:

Signal All pseudo-final state particles are correctly assigned their true-identity and originate from the same parent(s) also with the correct true-identity (Λ_b^0).

Low-Mass Mis-Reconstruction The same requirements as category 10 with the exception that the reconstructed head particle (Λ_b^0) could be of the incorrect identity and has other products that were not reconstructed.

Ghost At least one of the pseudo-final state particles is a ghost.

with the $M(\Lambda_b^0)$ shape of each of these choices shown in Figure 4.1 as well as the combined distribution with nominal truth matching. To reduce contamination from multiple ghosts, it is also required that only 1 ghost is retained in category 60. Allowing this limited mis-reconstruction into the simulation makes a more realistic representation of signal, as these effects will also be present in data and would be indistinguishable from “true” signal perfectly reconstructed signal.

Table 4.2: Summary of Fiducial Selection Cuts for the analysis, split by particle. The binary cuts `hasRich`, `hasCalo`, `InAccMuon`, `InAccEcal` and `InAccPrs` (Pre-Shower) queries if the particle meets the corresponding subdetectors acceptance requirements and hence if using that detectors response for further selection is valid. Other acronyms defined in the glossary

Particle	Run 1 Cut	Run 2 Cut	Particle	Run 1 Cut	Run 2 Cut
All Tracks	$\text{Trk_GhostProb} < 0.3$ $\chi_{trk}^2 < 5$		μ	$1.7 < \eta < 5.3$ $(3 < p < 150) \text{ GeV}$ $(0.8 < p_T < 30) \text{ GeV}$ <code>hasRich</code> & <code>InAccMuon</code>	
A_b^0	$2 < \eta < 5$			e	$1.7 < \eta < 5.3$ $(3 < p < 150) \text{ GeV}$ $(0.5 < p_T < 30) \text{ GeV}$ <code>hasRich</code> & <code>InAccPrs</code> <code>InAccEcal</code> & <code>hasCalo</code> In Active Ecal Region
p	$1.7 < \eta < 5.3$ $(9.3 < p < 150) \text{ GeV}$ $p_T > 250 \text{ MeV}$ $p_T > 1000 \text{ MeV}$ <code>hasRich</code>		$\ell_1 \ell_2$		$p_T(\ell_1) \cdot p_T(\ell_2) < 150 \text{ GeV}^2$
K^-	$1.7 < \eta < 5.3$ $(2 < p < 100) \text{ GeV}$ $p_T > 250 \text{ MeV}$ $p_T > 400 \text{ MeV}$ <code>hasRich</code>				

4.1.3 Fiducial Selection

Outlined in Section 4.2, a series of data-driven corrections are applied to the simulation using external calibration datasets. To ensure the techniques are valid, the parameter space from which the corrections are calculated must mirror the analysis samples, to prevent any loss of accuracy from extrapolation. Therefore “fiducial” kinematic cuts are made to the datasets according to the different correction techniques, summarised in Table 4.2 (see Section 4.2 for more detail on the origin of these cuts).

4.1.4 Trigger Selection

As discussed in Section 2.4.1, the LHCb Run 1 and 2 trigger had three stages: a hardware trigger (L0) and two software triggers (HLT1 and HLT2). The hardware trigger uses responses from the fast subdetectors, the muon chambers and calorimeters, while the software trigger creates tracks and forms loose candidates to make decisions.

To allow alignment between the signal and control modes, no electron specific triggers are used in the analysis, with the focus instead on the well-reconstructed muon that is shared between the signal and control mode. The final trigger selection used are defined in Table 4.3.

Table 4.3: Trigger selection applied to all analysis samples, only one of each HLT1 and HLT2 lines need to be triggered for the event to be retained. The different lines are defined in the text.

Trigger Level	Trigger Lines (Run 1 Only)	
	Run 1	Run 2
L0 (ℓ)	Muon_TOS	
HLT1 (A_b^0)	TrackAllL0_TOS TrackMuon_TOS	TrackMVA_TOS TrackMuon_TOS
HLT2 (A_b^0)	Topo2,3,4BodyBBDT_TOS TopoMu2,3,4BodyBBDT_TOS	Topo2,3,4Body_TOS TopoMu2,3,4Body_TOS

4.1.4.1 L0 Selection

Several trigger strategies are considered for L0 due to the two hadrons, muon and electron in the final state of the signal mode plus the options of triggering with TIS, TOS or both. To simplify the trigger efficiency estimation in later stages of the analysis, by enabling the use of the TISTOS method (see Equation 2.4.1), only TOS versions of the trigger lines are used. This is justified by the only small ε_{Signal} increase of 3.8% by including the TIS triggers.

The L0 muon trigger, which is very reliable due to the low occupancy and high efficiency of the muon chambers, is the most efficient single trigger available for the signal mode with a signal efficiency of 57%, referred to as L0Muon [158]. While a $\approx 16\%$ improvement could be realised by including an electron specific trigger for the signal mode, the L0Muon trigger alone can be shared between the signal and control mode maximising cancellation of systematic effects when calculating the trigger efficiencies, via the ε ratio in Equation 3.2.1. The efficiency of alternative trigger options are summarised in Appendix A.1.

The performance and efficiency of the L0Muon trigger depends on the threshold set for the trigger on $p_T^{L0}(\mu)$, an integer representation of the approximate p_T of the triggering muon, based on the fast-response in the muon chambers. During data taking, these thresholds are tuned dependent on the live data-rates and performance of the trigger system, with each data taking run assigned a specific TCK to document this. This results in the data containing a mixture of TCKs and therefore trigger thresholds, where as the simulation is generated with a single TCK, introducing mis-alignment between data and simulation. To correct for this with the L0Muon trigger, any data with a $p_T^{L0}(\mu)$ value equal or below the single threshold in simulation is removed, and by studying the remaining mixture of thresholds in data, the thresholds in simulation are systematically

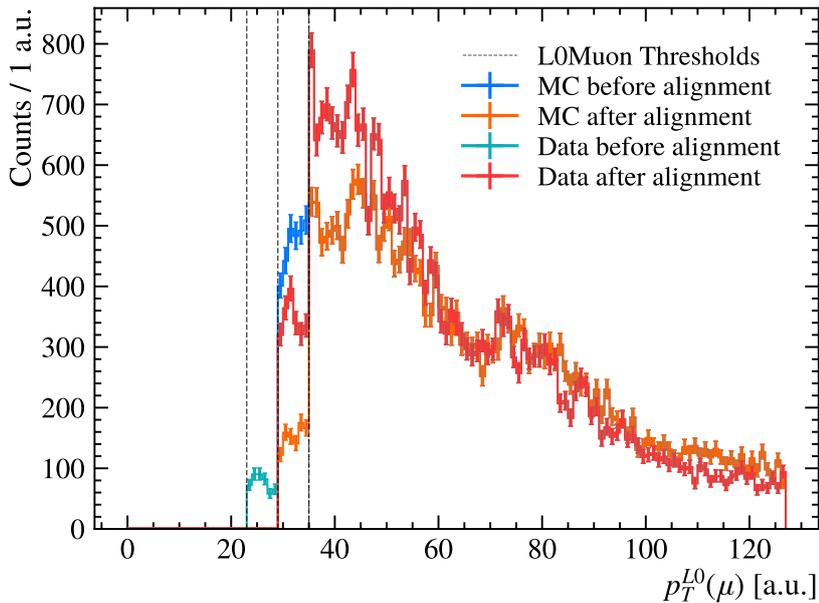


Figure 4.2: Distribution of $p_T^{L0}(\mu)$ in $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ MC/Data before and after Trigger Configuration Key (TCK) alignment of the L0Muon is applied for a subset of 2017 MD samples. The three thresholds used during 2017 MD data taking are marked. The features of the distributions correspond to the activation of different muon chambers at higher $p_T^{L0}(\mu)$ [158].

(but randomly) tightened such that the fractional luminosity of each TCK in data is reflected in the mixture of artificial $p_T^{L0}(\mu)$ thresholds in simulation. This process can be seen in Figure 4.2 for the 2017 MD sample, with the steps in counts in both data and MC coinciding with the marked thresholds. The complete set of TCKs for Run 1 and Run 2 of LHCb are centrally documented and retrieved for this analysis via an internal LHCb tool [42, 159].

4.1.4.2 HLT Selection

The content and topology of the signal and control modes are sufficiently similar that the same HLT trigger lines can be used to select candidates. The HLT1 triggers are summarised in Table 4.3 and correspond to cuts on the p_T and χ_{IP}^2 of reconstructed tracks that pass the L0 requirement. During Run 1 the `Track(A11L0/Muon)_TOS` trigger [160] required simple cuts on these variables where as in Run 2 `Track(MVA/Muon)_TOS` applied a more complex 2D selection cut of

$$\log \chi_{IP}^2 > \frac{b(25 - p_T)}{25(p_T - 1)^2} \cdot \log 7.4 \quad (4.1.1)$$

Table 4.4: Pre-selection cuts applied to the analysis datasets. Mis-ID vetoes are applied by making mass-hypothesis substitutions such as $M(pK^-)_{p\leftarrow K}$ where the invariant mass is calculated with the reconstructed p re-hypothesised as a K .

Description	Signal Mode Selection	Control Mode Selection
Mass Selection		
$pK^- \mu^\pm e^\mp$	$(4500 < M(pK^- \mu^\pm e^\mp) < 6750) \text{ MeV}$	
pK^-	$(1480 < M(pK^-) < 1560) \text{ MeV}$	$(1450 < M(pK^-) < 1850) \text{ MeV}$
J/ψ	$M(\mu^\pm e^\mp) \notin (\sqrt{9.0}, \sqrt{10.1}) \text{ GeV}$ $(2900 < M(\mu^+ \mu^-) < 3200) \text{ MeV}$	
Background Vetoes		
ϕ	$ M(pK^-)_{p\leftarrow K} - 1019.46 > 12 \text{ MeV}$	
J/ψ	$ M(K^- \ell^+)_{K\leftarrow \ell} - 3097 > 35 \text{ MeV}$	$ M(K^- \mu^+)_{K\leftarrow \mu} - 3097 > 35 \text{ MeV}$
$\psi(2S)$	$M(\mu^\pm e^\mp) \notin (\sqrt{13}, \sqrt{14}) \text{ GeV}$	
γ Pole	$M(\mu^\pm e^\mp) > \sqrt{0.1} \text{ GeV}$ $M(h^\pm e^\mp) > 10 \text{ MeV}$	
$D^0 \rightarrow K^- \pi^+$	$ M(K^- \ell^+)_{\ell\leftarrow \pi} - 1865 > 20 \text{ MeV}$	$ M(K^- \mu^+)_{\mu\leftarrow \pi} - 1865 > 20 \text{ MeV}$
$B^+ \rightarrow K^+ \ell^+ \ell^-$	$ M(p\ell^-)_{p\ell\leftarrow K\pi} - 1865 > 20 \text{ MeV}$ $M(h^+ \ell^+ \ell^-) < 5200 \text{ MeV}$	
Topological Cuts		
Clone Tracks	$\theta_{hh,hl,\ell\ell} > 1 \text{ mrad}$	

where b is the single tuneable parameter used to alter the rate of the trigger rather than directly changing the p_T and χ_{IP}^2 threshold as for Run 1 [132]. This selection ensured that high- p_T responses from L0 corresponded to high-quality tracks upon reconstruction.

For HLT2 the triggers stated in Table 4.3 correspond to selecting two, three or four track candidates creating the topology of a heavy-flavour hadron decay. This is determined by a lightweight Multivariate Analysis (MVA) built into the trigger line that determines and selects the optimum candidates within an event [132]. The `TopoMu` lines also require a loose PID requirement on the muon(s), requiring it has been flagged as `isMuon & HasMuon`, the same loose selection that is applied in the stripping selection discussed in Section 4.1.1.

The discussed choice of lines were one set of many options available, with the priority given to aligning selection between the signal and control mode by focusing on the shared muon of the decays. The efficiency of the different options is detailed in Appendix A.1, with the corrected efficiency of the nominal lines detailed in Tables 4.12 and 4.13, where the HLT lines are shown to select $> 95\%$ of signal candidates in Run 2.

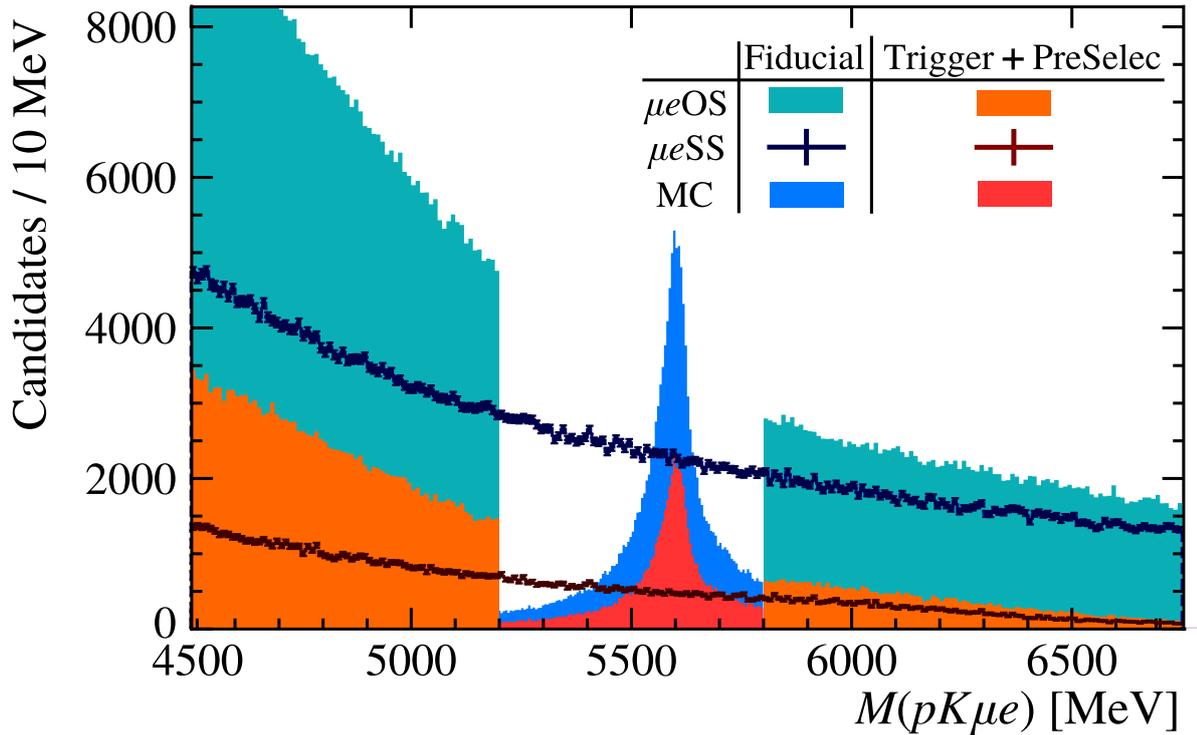


Figure 4.3: Overlaid distributions of $M(\Lambda_b^0)$ for $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ data and MC, showing how the datasets change in magnitude and shape from after the fiducial selection to after the trigger and loose pre-selection is added to the selection chain. The MC is scaled by $\frac{1}{4}$ to share axes with the data. The μeSS proxy dataset is shown with the dashed lines over the blinded μeOS signal.

4.1.5 Loose Pre-Selection

After stripping, fiducial and trigger selection, loose cut-based pre-selection is applied to remove physical and combinatorial backgrounds that otherwise dominate the contamination in the signal region. Table 4.4 summarises these cuts. The pre-selection largely follows the R_{pK} and $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ branching fraction analysis with the shared final state of $pK\ell\ell$ [76, 105].

The efficiency of these requirements are evaluated using simulation in Section 4.6, while their effect on the signal datasets can be seen in Figure 4.3, where the selection reduces the level of background in the sideband regions, a major step towards the near-zero background desired for a forbidden LFV search. Data taken during 2016 is displayed for illustrative purposes for Figures 4.3, 4.5, 4.6, 4.7 and 4.10, with consistent distributions in other data-taking years.

4.1.5.1 Mass Selection

The search for Λ_b^0 candidates in the $M(pK\mu^\pm e^\mp)$ invariant mass spectrum is performed within the range $4500 < M(\Lambda_b^0) < 6750$ MeV. This is derived to be within the limits set by the Bu2LLK stripping line ($3779 < M(\Lambda_b^0) < 6779$ MeV), without reaching into lower mass regions where non- b -hadron decays could contaminate. The same range is used for the control mode. This selection also defines the sideband regions, the areas of the signal data outside of the blinded region in the dataset. The lower sideband is set as $4500 < M(\Lambda_b^0) < 5200$ MeV and the upper $5800 < M(\Lambda_b^0) < 6750$ MeV.

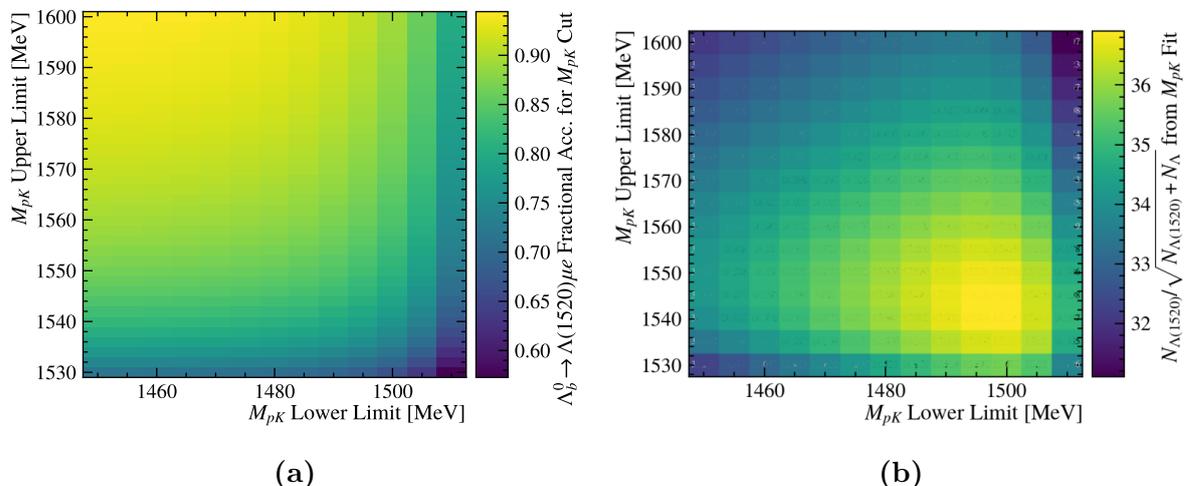


Figure 4.4: 2D scans of M_{pK} used to determine the optimum $\Lambda(1520)$ region cut for the analysis signal mode. Figure (a) shows ε_{MC} from corrected $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ simulation and (b) shows a constructed figure-of-merit for selecting $\Lambda(1520)$ in data with respect to other Λ^* resonances using control mode data.

The analysis is a search for $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$, hence $\Lambda(1520) \rightarrow pK^-$ candidates must be selected. While this is possible using the *sPlot* method (discussed in Section 4.2) and extracting the $\Lambda(1520)$ component from a wider spectrum, this would be very difficult for a search where very few, if any, events are expected to be observed. It is more robust to select a tight window around $m_{\Lambda(1520)} = (1519 \pm 6)$ MeV [41] to ensure that the analysis is measuring the branching fraction of the proposed $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ decay. The region chosen, $1480 < M_{pK} < 1560$ MeV, produced a high $\varepsilon_{MC} = 88.6\%$ on signal mode simulation while maximising the fraction of pK candidates measured to originate from a $\Lambda(1520)$ in this region from a fit to the M_{pK} distribution in control mode data discussed in Section 5.3.4. The 2D scans for this search are shown in Figure 4.4.

As discussed, the control mode for this analysis is inclusive in M_{pK} . Not selecting the $\Lambda(1520)$ resonance increases the control mode sample size improving the quality and reliability of measurements and shifting the focus to be on the important signal mode. The

control mode however is intended to emulate the signal mode in kinematics as much as possible to aid in the cancelling of systematic uncertainties, so a selection of $1450 < M_{pK} < 1850$ MeV is applied to limit the phase-space of the pK pair in comparison to the signal mode. This range was chosen to match the M_{pK} signal region in the $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ branching fraction analysis [105], covering the lower-mass Λ^* resonances in Figure 3.2, without allowing the leakage of complex pentaquark resonances.

The selected region of dilepton invariant mass, q , (more often used as q^2) dictates the kinematics of the $\ell^+\ell^-$ pair that enter into the analysis dataset. There is a rich mixture of charmonium ($c\bar{c}$) and other neutral resonances in this spectrum for an opposite-sign, same flavour lepton pair. For the control mode $\Lambda_b^0 \rightarrow pK^- J/\psi(\rightarrow \mu^+\mu^-)$ decays are used, hence a tight selection of $2900 < q < 3200$ MeV is applied to remove contributions from other $\mu^+\mu^-$ resonances. For the signal mode, no $\mu^\pm e^\mp$ spectrum is expected, with strong limits on $J/\psi \rightarrow \mu^\pm e^\mp$ and $\phi \rightarrow \mu^\pm e^\mp$ [161, 162]. Hence, to maximise sensitivity, q^2 selection is chosen to be inclusive as possible. However due to the high-rate J/ψ and $\psi(2S)$ resonances as well as the photon pole towards $q^2 = 0$, a series of vetoes are made to the signal mode to remove contamination from $\ell \rightarrow \ell'$ mis-identification, which only contribute significantly at the resonances due to lepton mis-ID being $< 1\%$ at LHCb [128]. These are summarised in the ‘‘Background Vetoes’’ section of Table 4.4. The overall changes to the mass spectrum of q^2 can be seen in Figure 4.6a.

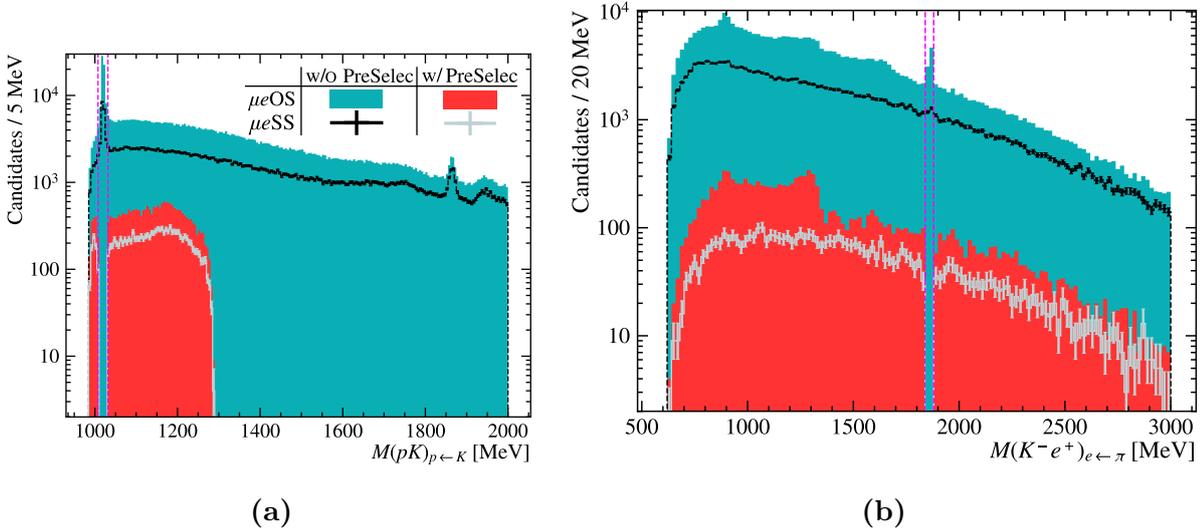


Figure 4.5: Invariant mass distributions of μe OS data showing mis-ID background resonances and corresponding vetoes applied (in pink). Figure (a) shows the dominant KK resonance ϕ in $M(pK^-)_{p \leftarrow K}$ (invariant mass calculated with the reconstructed p re-hypothesised as a K) and (b) a D^0 contribution in $M(K^- e^+)_{e \leftarrow \pi}$. The legend is shared.

Table 4.5: Selection targeting the removal of semileptonic backgrounds applied to the analysis datasets. Electron specific cuts applied only to the signal $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ mode are marked with †.

Description	Cut
Inclusive 2D Cut†	$M_{HOP} > 75 \times \log \chi_{\text{FD}}^2(\Lambda_b^0) + 3750 \text{ MeV}$
Λ_c^+ Veto	$M(pK^-\ell^+) > 2320 \text{ MeV}$
Λ_c^+ Veto $p \leftrightarrow K$ Swap	$M(p^-K^+\ell^-)_{pK \leftrightarrow Kp} > 2320 \text{ MeV}$
D_s^\pm Veto	$M(pK^-\ell^\pm)_{K \leftarrow p} > 2000 \text{ MeV}$
Λ_c^{*+} Veto	$M(pK^-\ell^+) > 3000 \text{ MeV}$
Λ_c^{*+} Veto $p \leftrightarrow K$ Swap	$M(p^-K^+\ell^-)_{pK \leftrightarrow Kp} > 3000 \text{ MeV}$

4.1.5.2 Background Vetoes

In addition to the $\ell^+\ell^-$ vetoes in q^2 mentioned, a series of vetoes removing other physics backgrounds are summarised in Table 4.4. These range from removing possible B^+ candidates formed from three tracks as well as removing $\ell^+\ell^-$ resonances from combinations of one lepton and one mis-identified hadron. The important ϕ and D^0 vetoes robustly remove these high rate narrow resonances with tight cuts around the measured masses of these particles, which were immediately visible in the sidebands of the signal region during the early stages of the analysis, as seen in Figure 4.5. This followed the general method when choosing the background vetoes to include; by searching in the sideband data for specific mis-ID resonances, before identifying and removing them.

Clone tracks are removed using the angular vetoes in Table 4.4. Clone tracks occur when the same, or a very similar complement of hits have been used to construct two near-identical versions of a track. This can be caused by hits in the detector being in very close proximity or an error in the reconstruction. While a specialised central reconstruction algorithm removes the majority of these clones, offline candidates where two particles have an opening angle < 1 mrad are removed to remove any residual clone pairs. This is especially important for the signal mode, where the $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ decay has a very low rate and any contamination from clones could be significant.

4.1.6 Semileptonic Removal

The q^2 , M_{pK} , $M(pKe^+)$ and $M(pKe^-)$ invariant mass distributions after applying the loose pre-selection are shown in Figure 4.6. Comparing the opposite-sign (signal data sidebands) and same-sign (background proxy) shapes after pre-selection (blue blocks and

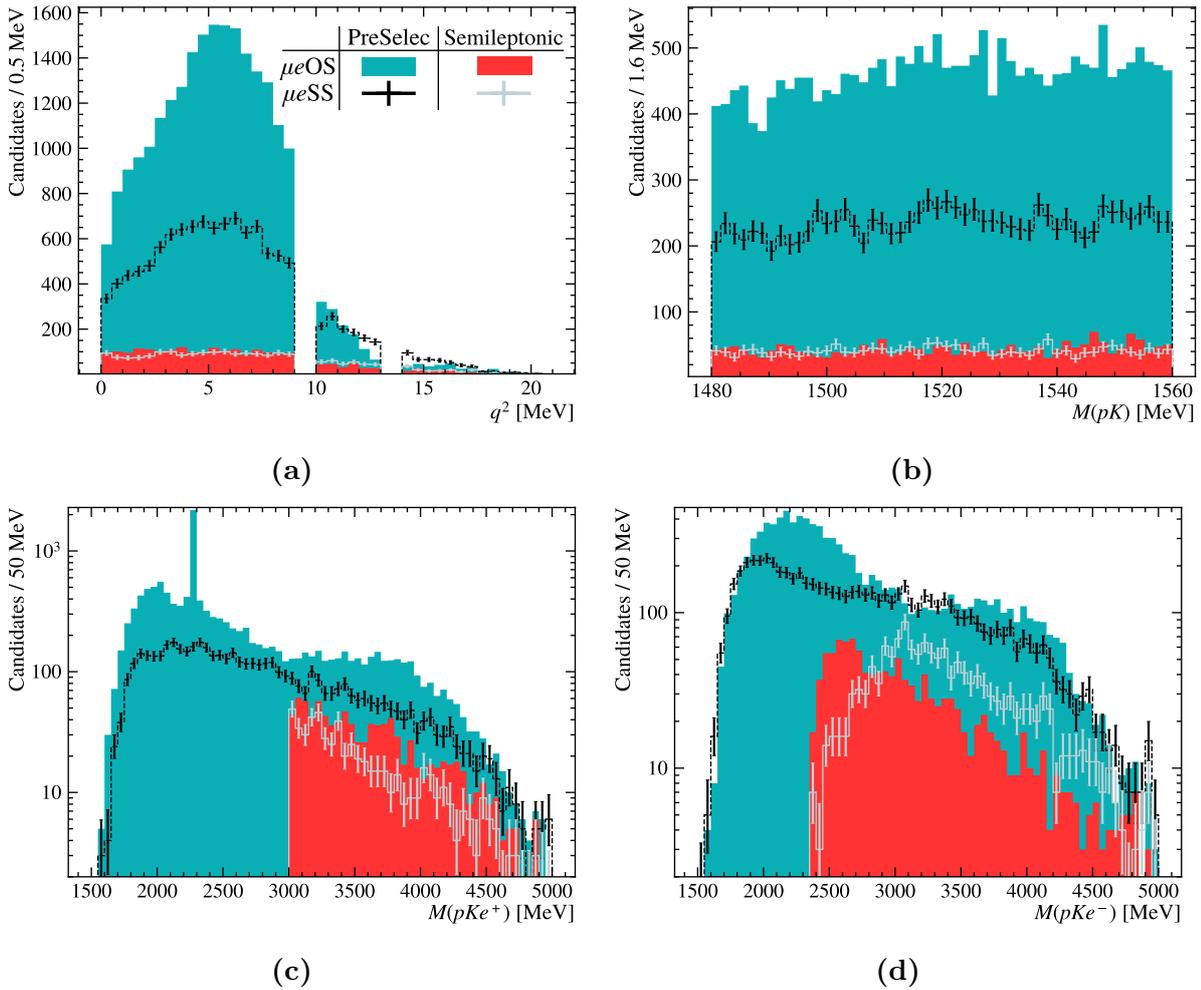


Figure 4.6: Invariant mass distributions of different final state particle combinations in μe OS and μe SS data before and after semileptonic selection. (a) shows the q^2 distribution, (b) $M(pK)$, (c) $M(pKe^+)$ and (d) $M(pKe^-)$. The legend is shared.

black points respectively) a general excess can be seen in the μe OS above the μe SS, biased to lower mass regions such as $M(pKe^\pm) < 3000$ MeV. This is also displayed in the $M(\Lambda_b^0)$ plot in Figure 4.3, far greater in the lower mass sideband than the upper sideband. This shows there is a large contamination remaining from partially reconstructed physics backgrounds with missing energy, either from semileptonic particle decays $X \rightarrow Hl\nu$ where ν is not detected, or from the mis-reconstruction of a particle with more than four children, such as an additional π^0 , that shifts $M(pK\mu^\pm e^\mp)$ down into the lower sideband. The first of these are expected to be dominant from Cabibbo-favoured $b \rightarrow cl^- \nu$ decays with the neutrino removing a large range of energy from the decaying system with no clear and accurate way to calculate the missing energy. Therefore two methods are employed to remove semileptonic decays, an inclusive ‘‘HOP FD’’ 2D cut, as well as a set of exclusive vetoes targeting specific high rate semileptonic processes, summarised in Table 4.5.

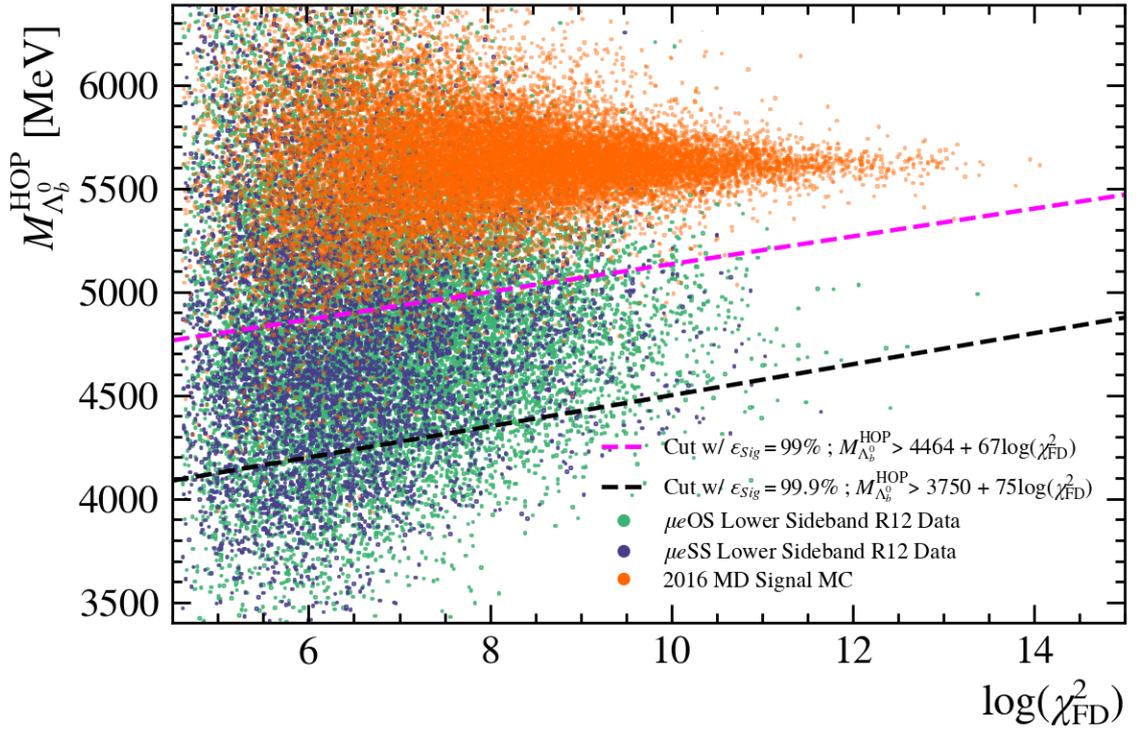


Figure 4.7: A distribution of M_{HOP} against $\log \chi_{FD}^2(\Lambda_b^0)$ for signal MC and the lower sidebands of the μeOS and μeSS datasets with the corresponding Run 2 samples. The two considered 2D inclusive semileptonic cuts are overlaid.

4.1.6.1 Implementation of the Inclusive Semileptonic Cut

The “HOP” variable is widely used in LHCb analyses and derives a scale factor that attempts to correct for the energy lost from electron Bremsstrahlung emission using the visible p_T asymmetry in the decays of b -hadrons [163]. The Figure 4.7 shows the distribution of M_{HOP} against $\log \chi_{FD}^2(\Lambda_b^0)$ for the signal MC and the lower sidebands of the μeOS and SS datasets, where M_{HOP} is the invariant mass after the HOP correction and $\chi_{FD}^2(\Lambda_b^0)$ is a measure of how far the reconstructed Λ_b^0 candidate has travelled between production and decay. The signal MC and sidebands appear initially to be distinct and able to be easily separated in the plot, however the signal MC leaks into lower values of M_{HOP} , especially at low $\chi_{FD}^2(\Lambda_b^0)$, therefore an inclusive 2D cut of $M_{HOP} > X_1 \log \chi_{FD}^2(\Lambda_b^0) + X_2$ is used to retain signal while removing partially reconstructed backgrounds present in the sidebands.

For this analysis, two inclusive semileptonic cuts were considered with $\varepsilon_{Signal} = 99\%$ and $\varepsilon_{Signal} = 99.9\%$, each with X_1 and X_2 optimised to maximise background rejection of the μeOS lower sideband. However, as Figure 4.8 displays, using the tighter $\varepsilon_{Signal} = 99\%$ option further sculpts the lower sideband combinatorial shape, shifting the core of the distribution into the signal region with the peak closer to $m_{\Lambda_b^0}$. While removing

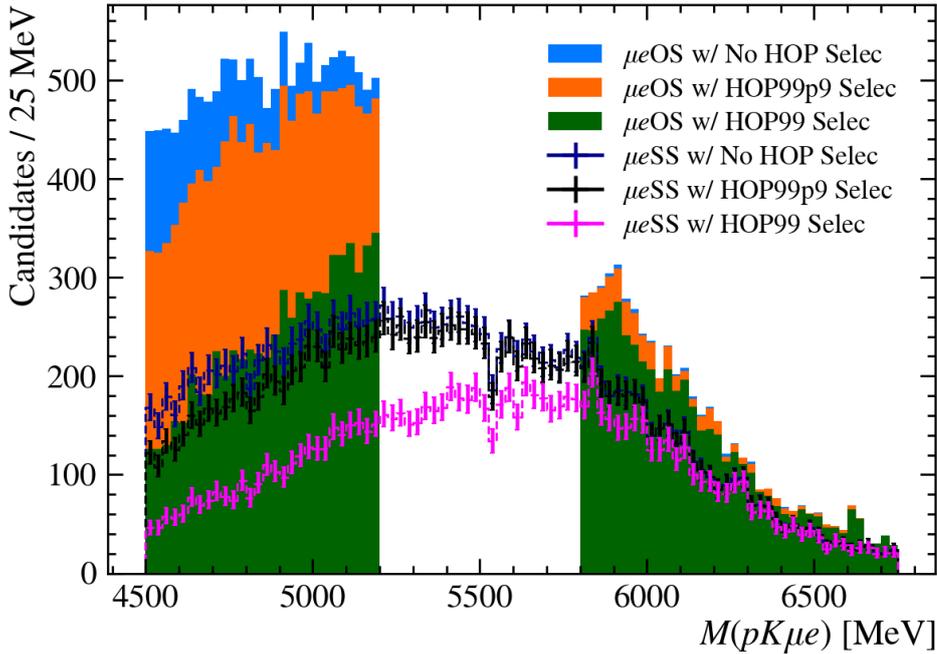


Figure 4.8: Distribution of $M(\Lambda_b^0)$ in the full Run 1 and 2 signal datasets as a function of the considered options for a inclusive semileptonic cut, where $\varepsilon_{Signal} = 99.9\%$ is ultimately chosen. The deficit observed at 5525 MeV was studied and deemed a fluctuation.

background is the aim of the event selection, the ultimate background contents will be measured by a NLL fit and a shape that forms a peak in the signal region is harder to parameterise in a blinded dataset, especially with the low background level derived for this analysis. Due to this as well as the increased signal efficiency, the $\varepsilon_{Signal} = 99.9\%$ option is chosen as described in Table 4.5.

4.1.6.2 $\Lambda_c^{(*)+}$ Veto

The only SM physics processes that can create the signal mode $pK\mu^\pm e^\mp$ final state are $\Lambda_b^0 \rightarrow \Lambda_c^+(\rightarrow pK^-\ell^+\nu)\ell^-\bar{\nu}$, $\Lambda_b^0 \rightarrow \Lambda_c^{*+}(\rightarrow D^0(\rightarrow K^-\ell^+\nu)p)\ell^-\bar{\nu}$ and $\Lambda_b^0 \rightarrow D^0(\rightarrow K^-\ell^+\nu)p\ell^-\bar{\nu}$, all double-semileptonic decays with large energy losses upon reconstruction due to the two neutrinos. While the energy loss will shift the background distributions into the lower sideband, any leakage into the signal would reduce the signal sensitivity and needs to be accounted for in the final $M(\Lambda_b^0)$ fits. A clear signature of the presence of these backgrounds can be seen in Figure 4.6c, with a large excess above the μeSS distribution below values of 3000 MeV, increasing below 2300 MeV, showing the contributions from Λ_c^{*+} and Λ_c^+ semileptonic decays respectively.

The most effective way of removing exclusive semileptonic backgrounds is to make a semileptonic veto. By reconstructing the visible invariant mass of a candidate particle

(such as Λ_c^+) and removing events below the known mass ($m_{\Lambda_c^+} = 2286.5 \text{ MeV}$ [41]), the remaining distribution could not have events originating from a semileptonic decay of the candidate particle due to the energy removed by the neutrino,

$$X \rightarrow H\ell\nu, P_\nu > 0 \therefore M_{H\ell} < m_X. \quad (4.1.2)$$

Hence a veto can be applied with a $M_{H\ell} > (m_X + \delta)$ selection, where δ prevents leak-

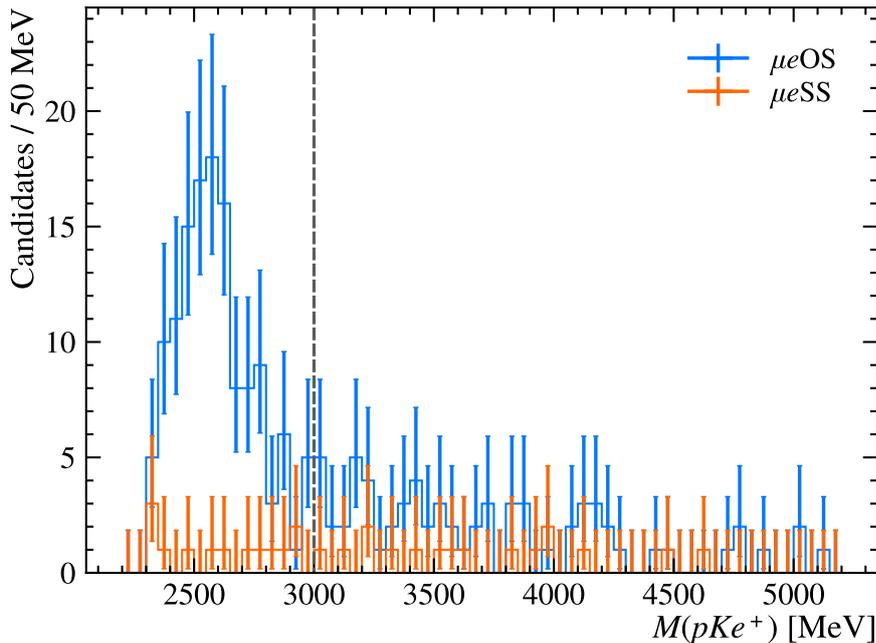


Figure 4.9: Invariant mass M_{pKe^+} distribution for the Run 1 and 2 μe OS and μe SS datasets, after applying all selection (including the Λ_c^+ veto) and the MVA discussed in Section 4.3 (at the medium working point) but without the Λ_c^+ veto, marked at 3000 MeV.

age from over-reconstructed contaminants. This relies on high quality reconstruction of the visible invariant mass, hence the performance of semileptonic vetoes are checked on simulated samples of the exclusive backgrounds, see Section 5.1.

Initially due to the relative branching fractions of the $\Lambda_b^0 \rightarrow \Lambda_c^{(*)+} \ell^- \bar{\nu}$ only a Λ_c^+ veto was considered for this analysis at $M(pK^-\ell^+) > 2320 \text{ MeV}$, but Figure 4.9, after applying this veto, shows that a significant excess was remaining below 3000 MeV due to contributions from $D^0 p$ candidates in $\Lambda_b^0 \rightarrow \Lambda_c^{*+} (\rightarrow D^0 (\rightarrow K^-\ell^+\nu) p) \ell^- \bar{\nu}$ decays, with three Λ_c^{*+} resonances, $\Lambda_c^{*+}(2860)$, $\Lambda_c^{*+}(2880)$ and $\Lambda_c^{*+}(2940)$ expected to dominate from studies of the analogous hadronic decay mode $\Lambda_b^0 \rightarrow D^0 p \pi^-$ [164]. Therefore, to remove these overlapping resonances from contaminating the signal region, the veto was tightened to $M(pK\ell) > 3000 \text{ MeV}$ which is $\approx M(\Lambda_c^{*+}(2940)) + \Gamma(\Lambda_c^{*+}(2940))$. The effectiveness of these vetoes are verified in the background study in Section 5.1.

In addition to the quoted veto, which are only applied to the correct charge combination, a pK swap version is implemented. This prevents a large asymmetry between the $pK\mu^+e^-$ and $pK\mu^-e^+$ datasets. In addition a D_s^\pm veto is applied, for the equivalent process $B_s^0 \rightarrow D_s^\pm(\rightarrow K^+K^-\ell^+\nu)\ell^-\bar{\nu}$, due to the common $p \leftrightarrow K$ mis-ID. This almost entirely overlaps with the $\Lambda_c^{(*)+}$ veto, but is retained for full clarity and bookkeeping, as in the summary Table 4.5.

While semileptonic vetoes are very effective at removing specific backgrounds and widely utilised in LHCb analyses, they also remove other candidates with $M(pK\ell)$ below the cut value despite not originating from a semileptonic decaying background. The effect of this on the signal efficiency and background shape is important to evaluate when deciding the implementation of these selections. The final signal efficiency on the ensemble of semileptonic vetoes is $\approx 50\%$ for 0γ events and $\approx 55\%$ for 1γ events, which is the largest reduction in efficiency by any single step of the analysis selection chain. Despite this, the power these vetoes offer in removing the only final state mimicking backgrounds justify their inclusion.

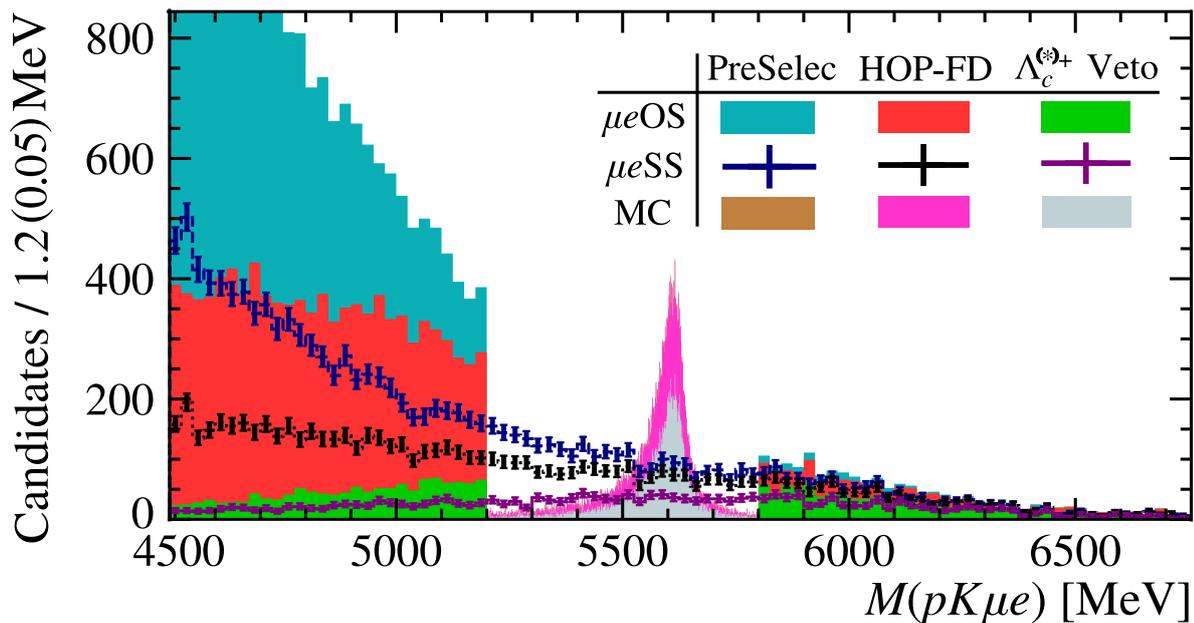


Figure 4.10: Overlaid distributions of data and MC for $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$, showing the changing dataset as the different semileptonic removal techniques are added to the selection chain. The MC is scaled by $\frac{1}{8}$ to share axes with the data, and is rebinned as marked in brackets on the y -axis. The μ SS proxy dataset is shown with the dashed lines over the blinded μ OS signal.

The semileptonic vetoes also considerably sculpt the combinatorial background distribution. Figure 4.10 shows the effect on the μ OS and μ SS distributions, with a clear reduction in the μ OS excess above the purely combinatorial μ SS in the lower sideband

but also the distribution tending towards zero at $M = 4500$ MeV. This is due to competing kinematic limits. The $pK\ell$ combination after the semileptonic vetoes is required to contribute a larger proportion of the Λ_b^0 candidate energy hence the other lepton must have a sufficiently low momentum to produce an $M(\Lambda_b^0) \rightarrow 4500$ MeV, yet simultaneously low momentum leptons are being removed by the fiducial cuts in Section 4.1.3.

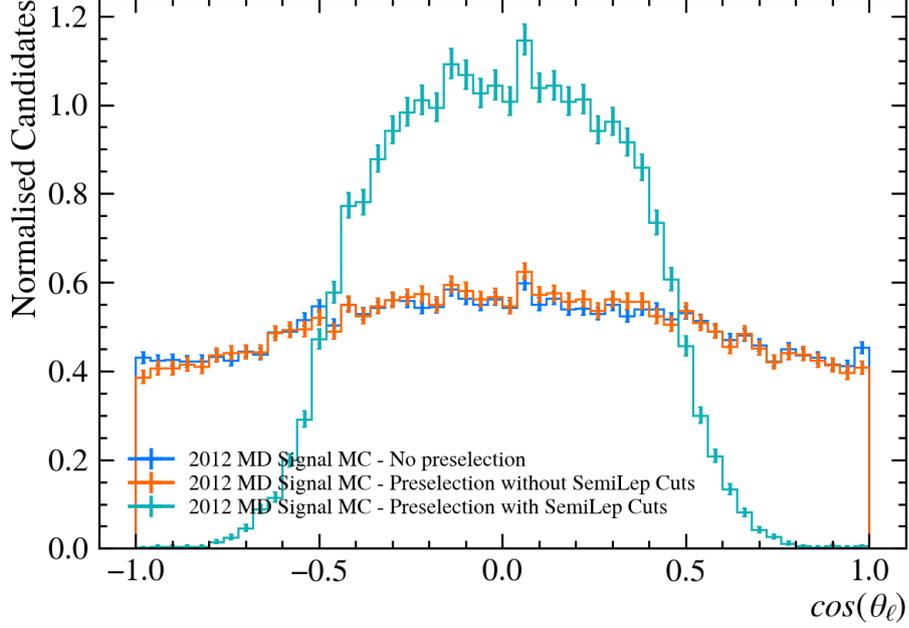


Figure 4.11: The angular distribution θ_ℓ in 2012 MD signal mode MC showing the change in shape as the semileptonic removal selection is added.

The kinematic sculpting can be seen explicitly in a distribution of θ_ℓ , the angle between the momentum vector of ℓ^+ in the $\ell^+\ell^-$ candidate rest frame and the momentum vector of the $\ell^+\ell^-$ pair in the Λ_b^0 rest frame. In Figure 4.11, adding the semileptonic vetoes dramatically changes the shape of the distribution, restricting θ_ℓ significantly. These vetoes would hence be unfeasible for an angular analysis of this decay mode, but as this is a search for a proposed forbidden decay, and removing any possible background is priority, the heavily sculpted mass and angular distributions are acceptable to reach the desired near-zero background state.

4.1.6.3 Feasibility of D^0 Semileptonic Veto

The only double-semileptonic background that can avoid the discussed vetoes is the non-resonant background $\Lambda_b^0 \rightarrow D^0(\rightarrow K^-\ell^+\nu)p\ell^-\bar{\nu}$ and only in the case where $M(D^0p) > 3000$ MeV. Figure 4.12a shows that while the large majority of these pairs will have been removed by the cut, a significant quantity are expected to remain in the analogue $\Lambda_b^0 \rightarrow D^0p\pi^-$ case.

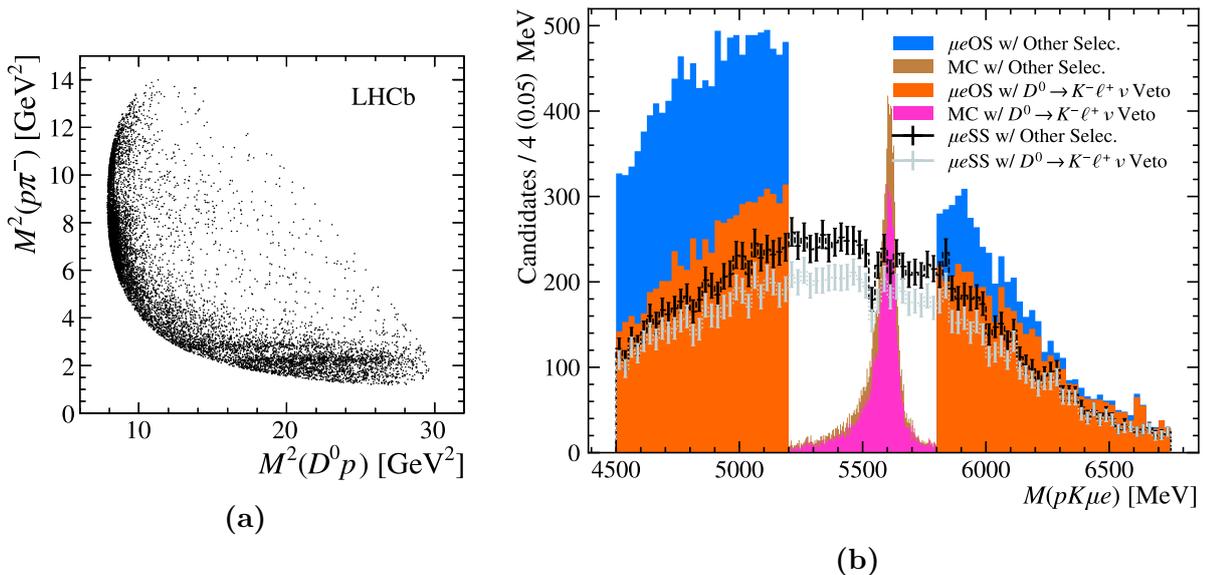


Figure 4.12: Feasibility study for a D^0 Semileptonic Veto. (a) A 2D distribution of $M_{p\pi^-}^2$ against $M_{D^0 p}^2$ in the decay mode $\Lambda_b^0 \rightarrow D^0 p \pi^-$ measured with LHCb Run 1 data [164]. (b) The overlaid $M(\Lambda_b^0)$ distributions for signal data and MC for Run 2 samples of this analysis, to study any potential sculpting. The MC is scaled by $\frac{1}{4}$ to share axes with the data, and is rebinned as marked in brackets on the y-axis.

Two options were considered to attempt removal of this contribution. The first is a robust two-body semileptonic D^0 veto of $M(K^- \ell^+) > m_{D^0}$ to target $D^0 \rightarrow K^- \ell^+ \nu$ decays. Figure 4.12b shows the small sculpting this would add to the combinatorial background, but the greater concern is the dramatic effect on signal efficiency, further reducing it by 28%, producing an overall $\varepsilon_{Signal} \approx 40\%$ for the full suite of semileptonic selection. This is considered to be unfeasible for the small amount of background these modes introduce (discussed in detail in Section 5.1). The alternative option is a targeted cut on this background, using the HOP variables as before, but as discussed, tightening this leads to major sculpting of the combinatorial background, and a stable background fit is of greater importance. The planned solution is to add an exclusive fit component for the remaining $\Lambda_b^0 \rightarrow D^0(\rightarrow K^- \ell^+ \nu) p \ell^- \bar{\nu}$ background contribution remaining after selection.

4.2 Corrections to Simulation

The MC samples are extensively used in the analysis for key components including the signal and control mode mass shapes, the target training sample for the MVA and importantly the samples used to calculate ε_{Signal} and $\varepsilon_{Control}$ for the $\Lambda_b^0 \rightarrow \Lambda(1520) \mu^\pm e^\mp$ branching fraction extraction. However the simulation samples produced are not a complete and accurate representation of real data. These differences arise from both recon-

struction effects such as simulation of complicated subdetectors like the RICH detectors, leading to incorrect evaluation of a particle’s PID response but also from incorrect modelling at the generation stage of simulation, such as particle decay kinematics that are not representative of real decays.

While these mis-alignments are minimised by using the latest LHCb simulation available, improvements to the agreement are made by deriving and applying a series of event-by-event weights to all MC samples, with each reconstructed event weight defined as

$$w_{reco} = w_{PID} \times w_{trk} \times w_{trig} \times w_{gen} \quad (4.2.1)$$

where

$$w_{gen} = w_{\tau} \times w_{kin} (\times w_{penta}). \quad (4.2.2)$$

Five weights are applied to the signal MC $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ sample to correct PID efficiency (w_{PID}), tracking efficiency (w_{trk}), L0 trigger efficiency (w_{trig}), Λ_b^0 lifetime (w_{τ}) and Λ_b^0 production kinematics (w_{kin}) with each weight derived with respect to the previous. For the control mode $\Lambda_b^0 \rightarrow pK^- J/\psi$ MC, an additional weight to correct the modelling and distribution of the M_{pK} spectrum is applied (w_{penta}) which is not required for the signal mode with $\Lambda(1520) \rightarrow pK^-$. For MC used in the background study (see Section 5.1), the reconstruction weights w_{PID} , w_{trk} and w_{trig} are always applied and the generation weights, w_{τ} and w_{kin} are applied to samples with a Λ_b^0 head particle and w_{penta} is applied if the sample is of type $\Lambda_b^0 \rightarrow pK^- \ell^+ \ell^-$ (including $c\bar{c} \rightarrow \ell^+ \ell^-$).

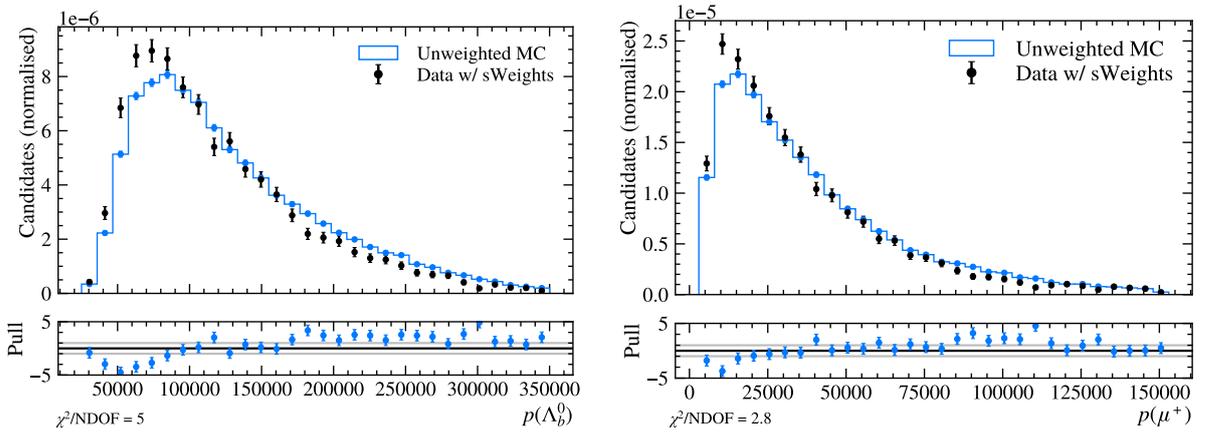


Figure 4.13: Reconstructed momentum distributions before reweighting is applied. The Λ_b^0 and μ measure is shown with 2016 MD control mode MC and corresponding sWeighted data (after cut-based selection).

An example of the need for this reweighting can be seen in Figure 4.13, where the Λ_b^0 and muon momentum distributions in the control mode MC versus signal-weighted (sWeighted) control mode data are not compatible before any reweighting is applied. The

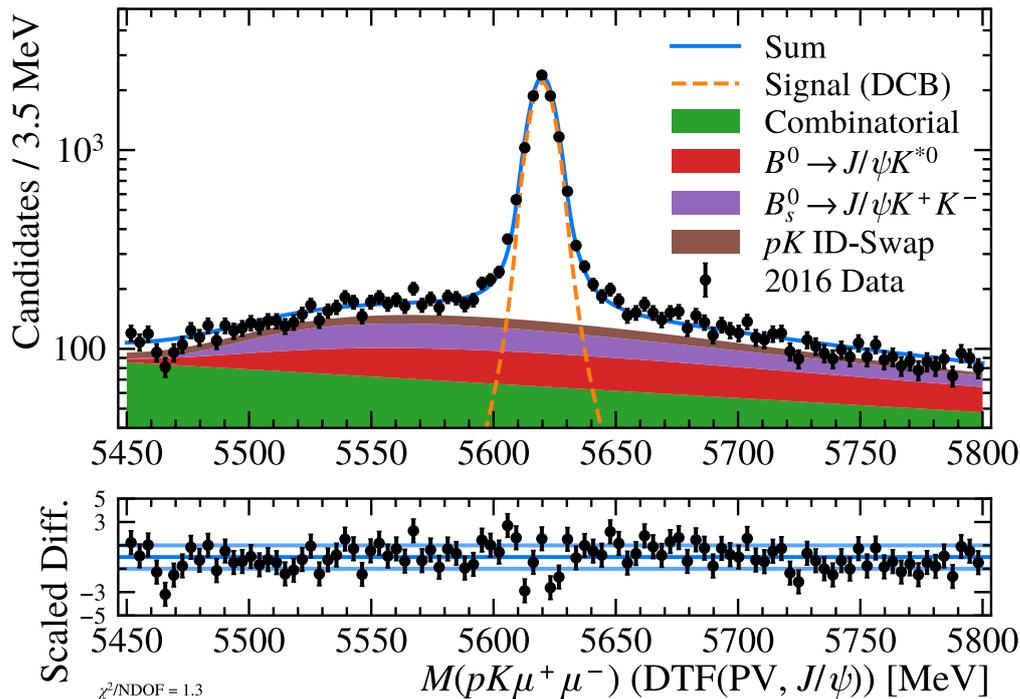


Figure 4.14: The fitted distribution of $M(\Lambda_b^0)$ in the 2016 control mode data of $\Lambda_b^0 \rightarrow pK^- J/\psi$ after the cut-based selection discussed in Section 4.1 and before the MVA (Section 4.3) is applied. A significant combinatorial background remains along with exclusive background contributions. The fitting technique follows the method discussed in Section 5.3.3.

sWeighted data is derived with the sPlot method [165], where the result of an extended-likelihood fit to data is used to separate the signal and background components by producing a scheme of signal and background “sWeights” that when applied, will render the distribution to be effectively signal-only or background-only by providing the most signal-like events with a higher weight and vice-versa. The signal-only sWeights are used to produce “target” distributions in other variables with minimal contamination from background and no mis-modelling that is present in the MC. The reweighting suite discussed in this section uses sWeighted data to correct towards, both by examining how adding weights improves the consistency of variables of interest (such as those in Figure 4.13) as well as providing these sWeighted distributions to algorithms to use directly in training.

For this analysis, the control mode $\Lambda_b^0 \rightarrow pK^- J/\psi (\rightarrow \mu^+ \mu^-)$ is used as the target sWeighted dataset on which to evaluate the performance of the MVA as it is most kinematically and topologically similar to the signal mode while having a large branching fraction that produces a statistically-sufficient dataset to derive sWeights in individual years and magnet polarities (discussed in Section 5.3), allowing discrete and detailed training and

analysis of the reweighting. The MC weights are derived after the full fiducial, trigger and pre-selection has been applied, but before the MVA, as corrected signal MC is a required input for the training of the MVA. The result of a mass fit for deriving sWeights can be seen in Figure 4.14, for the example of 2016 data. This year of data-taking is used as the illustrative example throughout this section.

4.2.1 Λ_b^0 Lifetime Correction

The simulation used in the analysis was generated with the Λ_b^0 lifetime ($\tau(\Lambda_b^0)$) set to an out-of-date value of $\tau(\Lambda_b^0)_{gen} = 1.451$ ps [166], and was therefore analytically reweighted to the current world average $\tau(\Lambda_b^0)_{wa} = 1.470$ ps [41] via the equation

$$w_\tau = \frac{\tau(\Lambda_b^0)_{gen}}{\tau(\Lambda_b^0)_{wa}} \cdot \exp \left[t(\Lambda_b^0) \cdot \frac{\tau(\Lambda_b^0)_{wa} - \tau(\Lambda_b^0)_{gen}}{\tau(\Lambda_b^0)_{wa} \tau(\Lambda_b^0)_{gen}} \right], \quad (4.2.3)$$

where $t(\Lambda_b^0)$ is the measured decay-time of the Λ_b^0 candidate. Figure 4.21f shows the distribution of $t(\Lambda_b^0)$ as weights are applied, showing that there is a relatively small difference in the sWeighted and simulated distribution.

4.2.2 Modelling Correction for $\Lambda_b^0 \rightarrow pK^- J/\psi$

As discussed in Section 3, there is a rich and complicated structure of Λ^* resonances in the M_{pK} spectrum, including pentaquarks (see Figure 3.2). In the generation of the simulation for $\Lambda_b^0 \rightarrow pK^- J/\psi (\rightarrow \mu^+ \mu^-)$ there is no attempt to model this and the spectrum is dictated only by the phase-space available from the decaying Λ_b^0 . Calculation of $\varepsilon_{Control}$ with the uncorrected simulation could lead to a systematic mis-representation of the efficiency, especially the cuts made in M_{pK} , having a potential knock on effect to the limit setting power of the analysis. Therefore a tool developed for the amplitude analysis of the $\Lambda_b^0 \rightarrow pK^- J/\psi$ decay [49], which uses the current knowledge of the Λ^* resonances and the discovered pJ/ψ pentaquark states, is used to derive a weight as a function of the truth-level four-momenta of the p , K^- and J/ψ . The tool was also used in the R_{pK} analysis as well as the $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ branching fraction/angular analyses [76, 105, 151].

The effect of the resulting weight, w_{penta} , can be seen clearly in Figure 4.15, where adding the weight re-introduces the resonance shapes as in the control mode data and Figure 3.2. The phase-space model was clearly underestimating the contribution in the lower-mass region, biasing the kinematics of the decay and as this analysis is to measure the exclusive $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ decay, correctly modelling the $\Lambda(1520)$ contribution in the control mode is essential.

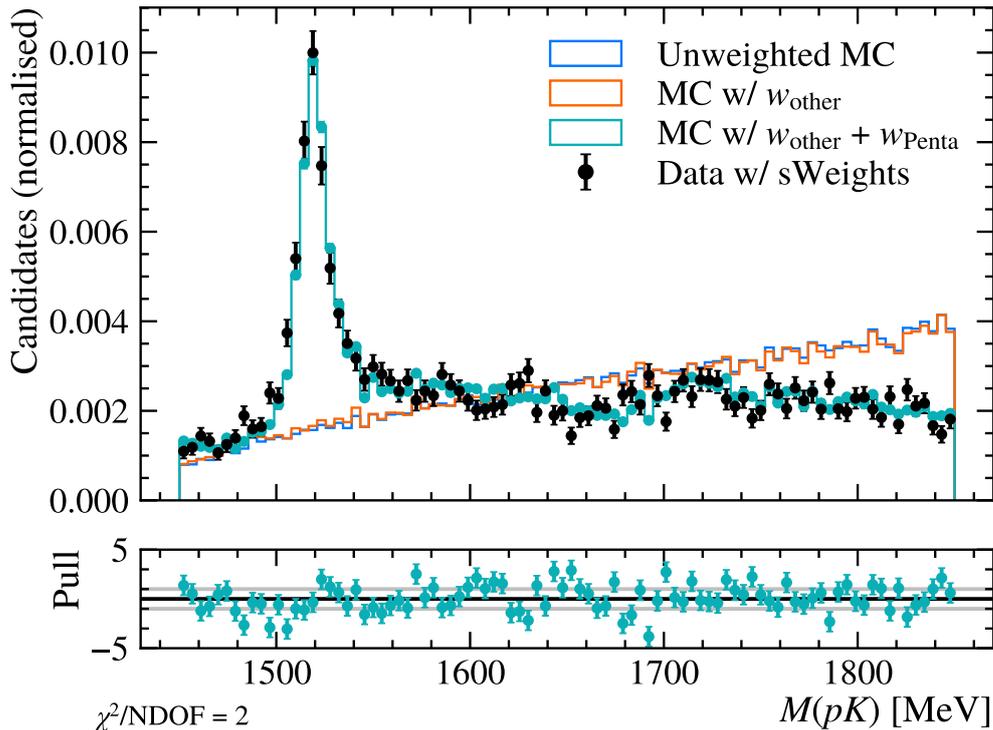


Figure 4.15: The result of the correction provided by the $\Lambda_b^0 \rightarrow pK^- J/\psi$ reweighter, with 2016 control mode simulation shown before and after applying w_{penta} (and with the remaining correction chain) with the corresponding sWeighted data overlaid to show the real “target” distribution.

The very good agreement shown in Figure 4.15 is testament to the performance of the powerful data-driven tool used across many analyses within LHCb. The correction is applied to all MC samples involving $\Lambda_b^0 \rightarrow pK^- \ell \ell$ decays, but not to the signal mode MC which, as discussed in Section 3, is due to efficiency savings on MC generation and the unknown decay structure of the LFV decay $\Lambda_b^0 \rightarrow pK^- \mu^\pm e^\mp$.

4.2.3 PID Efficiency Correction

The LHCb software package PIDCALIB [167] is used to obtain weights that describe the efficiency for the PID selection used in this analysis, both at the stripping and PID stage of event selection. These weights entirely replace the cut-based selection in simulation that is applied to the collision datasets due to mis-modelling of the PID detectors response and therefore PID variables in the MC.

PIDCALIB2 is the new python implementation of the commonly used tool at LHCb. The tool calculates a PID weight on a particle-by-particle basis using a central, high-statistics collision data sample of a given particle, evaluating the efficiency of applying

Table 4.6: The calibration datasets used to derive the PID weights for the analysis, with each a large, pure sample of the given particle species. The decay from which each particle is selected is stated.

Particle	Run 1	Run 2
p	P_IncLc (p from inclusive Λ_c^+ decays)	
K^-	Kaon (from $D^0 \rightarrow K^- \pi^+$)	
μ	Mu (from $J/\psi \rightarrow \mu^+ \mu^-$)	
e	e (from $J/\psi \rightarrow e^+ e^-$)	e_B_Jpsi (from $B \rightarrow J/\psi X$)

the users PID cuts using the detectors actual response. The PID response varies with the kinematics of the particle, so maps of efficiency are built by PIDCALIB in variables and binning of choice, which act as lookup-tables in histograms. To maximise the alignment between the PIDCALIB datasets and the analysis MC, any over or under-coverage is solved by applying cuts in both directions, with the fiducial cuts in Table 4.2 dictated by the reference PIDCALIB datasets used. An example of the need for this alignment is shown in Figure 4.16.

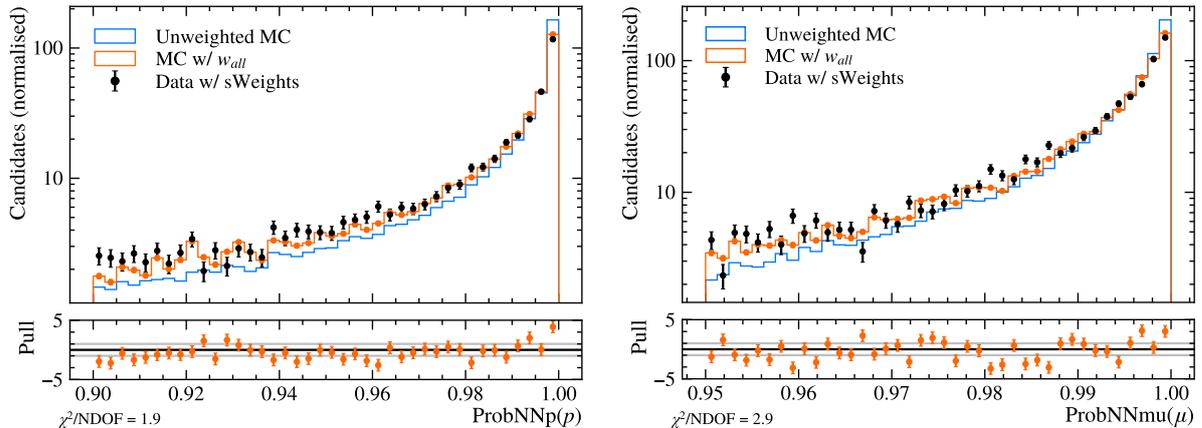


Figure 4.16: The effect on the distributions of PID variables as PID corrections are applied, with 2016 control mode simulation shown with corresponding sWeighted data overlaid. (left) Shows the proton ProbNN_p , used to select protons/reject kaons in the analysis and (right) the muon ProbNN_μ , important for ensuring muon purity.

PID efficiency histograms for p , K^- , μ and e are derived from the PIDCALIB datasets specified in Table 4.6. The maps use a binning scheme in particle p and η of isopopulated bins in the reference dataset. These were ported from the analogue analysis of $\Lambda_b^0 \rightarrow \Lambda e^\mp \mu^\pm$ [42, 119] with minor adjustments and merging of low statistics bins to ensure an error on the efficiency of $< 5\%$. Separate maps were produced for data taking years and magnet polarities due to the changing trigger lines by year and documented differences in PID performance for Mag Up versus Mag Down due to the

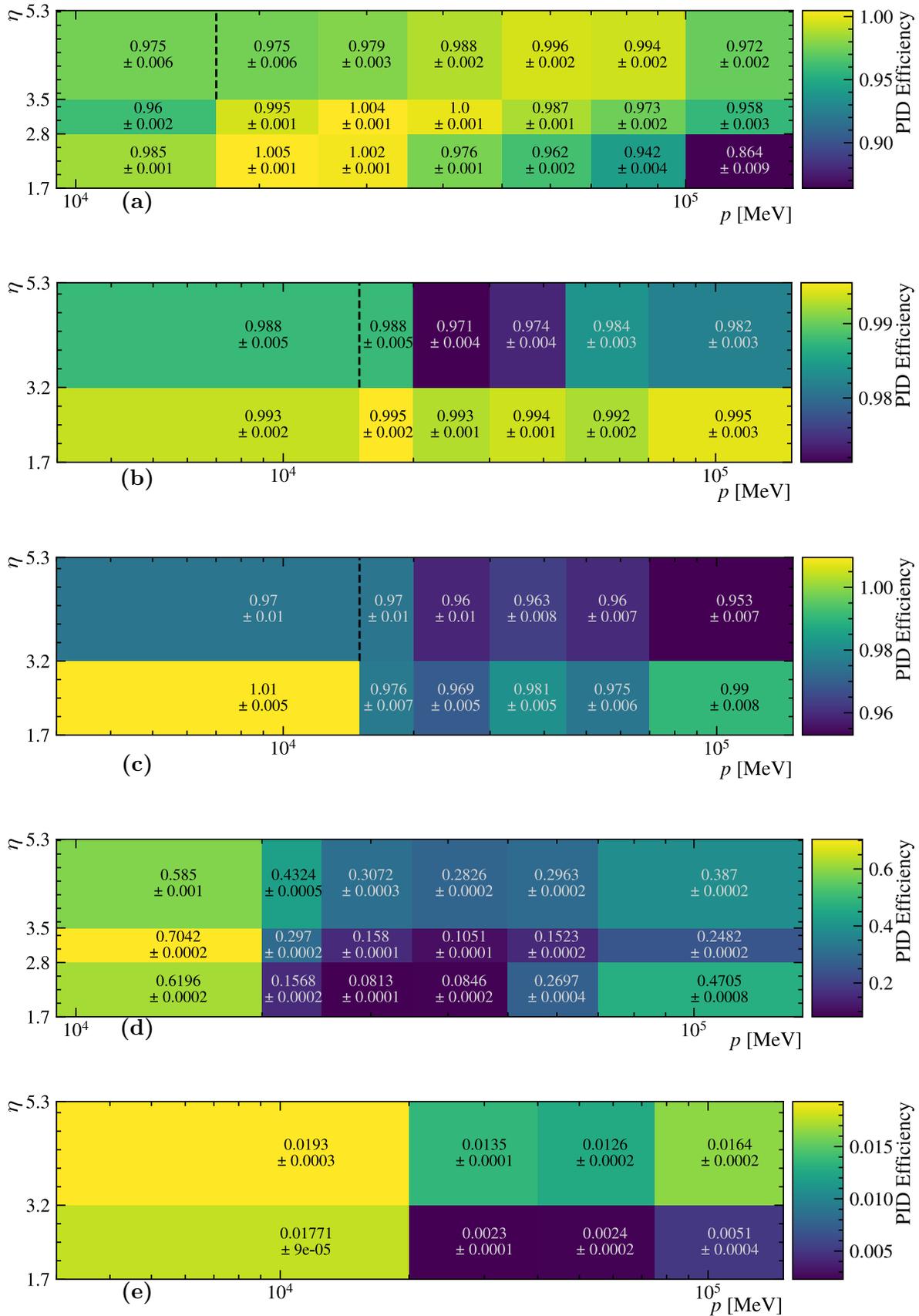


Figure 4.17: PID efficiency maps for 2016 MD calibration samples, displaying the stripping PID cut efficiency for (a) protons, (b)/(c) electrons with/without Bremsstrahlung and the efficiency for mis-identifying (d) a kaon as a proton and (e) a muon as an electron.

different luminosities and efficiency asymmetries within the detector [168]. The overall event PID weight is derived as

$$w_{PID}^{Event} = w_{PID}^p \times w_{PID}^K \times w_{PID}^{\ell_1} \times w_{PID}^{\ell_2} \quad (4.2.4)$$

where w_{PID} is derived for the stripping PID cuts and the selection PID cuts (with the stripping PID cuts as a prior), with more detail on the selection maps in Section 4.4. Example histograms of the stripping PID weights are shown in Figure 4.17 where the importance of binning in these kinematic variables is clear with a large variation in efficiency. As shown in the figure, the electron histograms are further split by Bremsstrahlung category due to the significant difference in the PID efficiency from the response in the ECAL for an electron with missing energy, with the 1γ category more efficiently selected (see Figure 4.17b and c).

In the signal and control mode, the weights applied to each particle align with the reconstructed identity, but to evaluate the PID efficiency on the background MC samples, mis-ID must be accounted for. “Mis-ID weights” are therefore produced by deriving efficiency maps for every combination of PIDCALIB sample and reconstructed particle species for the corresponding particle PID cut, with these maps displaying significantly lower efficiencies (as seen in Figure 4.17d and e) corresponding to the expected mis-ID rates discussed. For each particle in an MC sample, its true-ID is queried which decides the reference sample to use and retrieve a map for the particle’s reconstructed identity.

4.2.4 Track Efficiency Correction

Minor discrepancies are observed in the tracking efficiency between reconstructed data and simulation in LHCb [169]. This originates from differences in detector response between simulation and data affecting reconstruction. A central software package, TRACKCALIB2 was used to obtain weights to correct for this by deriving

$$w_{trk} = \frac{\varepsilon_{Long\ Trk}^{Data}}{\varepsilon_{Long\ Trk}^{MC}} \quad (4.2.5)$$

where $\varepsilon_{Long\ Trk}^{Data(MC)}$ corresponds to the tracking efficiency measured for long tracks in a large calibration sample of reference $J/\psi \rightarrow \mu^+\mu^-$ data and MC [169]. The “Tag-and-Probe” method is employed to calculate this for the different track types defined in Section 2.4.2. For the long track case the efficiency is measured by reconstructing one muon as a long “Tag” track with tight selection requirements and the other “Probe” track with unbiased selection using only hits in the TT and Muon chambers (`MuonTTTrack`), with the final

$$\varepsilon_{longtrk} = \frac{N_{\text{Tag, Long}}}{N_{\text{Probe, MuonTT}}}. \quad (4.2.6)$$

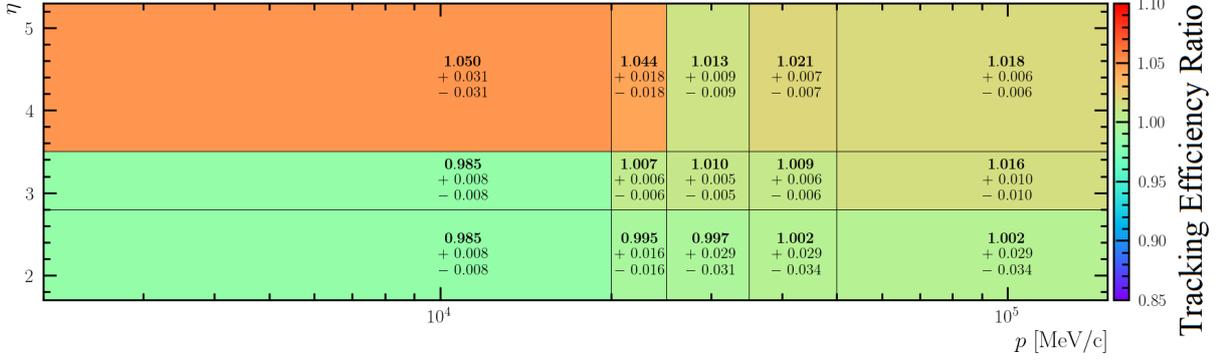


Figure 4.18: Track reconstruction efficiency ratio map for the 2016 calibration samples, displaying the measured efficiency of long tracks in data, divided by the efficiency in MC. Long tracks are the only track-type utilised in this analysis.

As with PIDCALIB, to parameterise the changing track efficiency as a function of particle kinematics, efficiency histograms in bins of track p and η are produced for each data taking year (example in Figure 4.18), using the same binning scheme as for the muon PID efficiency histograms. Alignment of the reference and analysis samples is made by adding the analysis track cuts to the reference “Tag” track sample. The final track efficiency weight is

$$w_{trk}^{Event} = w_{PID}^p \times w_{PID}^K \times w_{PID}^{\ell_1} \times w_{PID}^{\ell_2} \quad (4.2.7)$$

where each particles contribution is derived individually.

4.2.5 L0 Trigger Efficiency Correction

Trigger efficiency corrections are determined using the TISTOS method (described in Section 2.4) using an independent sample of $B^+ \rightarrow K^+ J/\psi (\rightarrow \mu^- \mu^+)$. This decay mode was chosen due to it being the highest rate b -hadron decay with similar leptonic topology to the signal mode, yielding a large data sample which will trigger with the same L0 trigger selection as the signal mode. The invariant mass fits to determine N_{TOS} and $N_{TIS\&TOS}$ for Equation 2.4.1 are described by Gaussian functions with power-law tails, and are performed for both MC and data with the final L0 trigger weight

$$w_{trig} = \varepsilon_{TISTOS}^{Data} / \varepsilon_{TISTOS}^{MC}. \quad (4.2.8)$$

To accurately represent the trigger efficiency across the phase-space of the B -decay the efficiencies are calculated in a 2D binning scheme and stored in maps as histograms. As

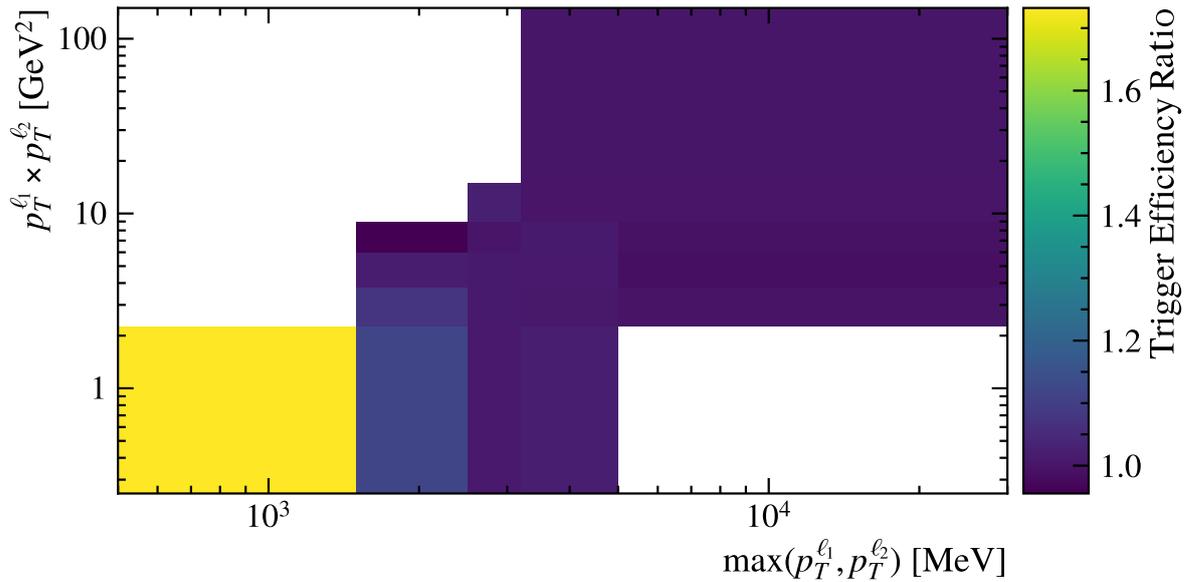


Figure 4.19: Trigger efficiency ratio map applied to 2016 signal MC samples, from a calibration sample of $B^+ \rightarrow K^+ J/\psi(\rightarrow \mu^- \mu^+)$ displaying the efficiency ratio of applying the L0_MuonTOS trigger, to correct simulation.

the L0 muon trigger is used, the important kinematics are those of the leptons, hence the efficiencies are split into bins of $\max(p_T^{\ell_1}, p_T^{\ell_2})$ and $p_T^{\ell_1} \times p_T^{\ell_2}$ with the corresponding alignment corrections in Table 4.2 from the bounds of the efficiency maps produced. An example is shown in Figure 4.19.

When applying this correction to the signal mode MC, where only one muon is present, $p_T(\mu)$ is used to decide the x -axis bin in Figure 4.19, with the y -axis retained to provide granularity in efficiency of the total momentum of the di-lepton system.

4.2.6 Λ_b^0 Production Kinematics Correction

The production kinematics of b -hadrons in the LHCb simulation framework is not modelled correctly. Hence a reweighting procedure was derived to correct this using a Gradient Boosted Re-weighter (GBR) from the `hep_ml` package, designed for use in high-energy physics experiments [170]. The GBR is provided with kinematic variables of a “target” dataset and an “original” dataset and trains an ensemble of regression trees that produce a weight that, when applied to the original distributions, will attempt to align them with the target distributions. This technique works simultaneously in multiple dimensions, its major advantage over binned reweighting techniques.

For this analysis, the target dataset is the sWeighted control mode data after pre-selection with the original dataset as the corresponding control mode MC at the same

selection-stage with all preceding weights applied. This is to ensure that the GBR is correcting the residual differences due to the Λ_b^0 production kinematics and not another systematic process. The GBR is trained with three variables, $p_T(\Lambda_b^0)$, $\eta(\Lambda_b^0)$ and NTRACKS (the total number of tracks in a given event), all showing significant differences between data and MC before kinematic reweighting (see Figures 4.21a, b and c). To prevent the GBR from “over-training”, the complexity of the model is limited to 50 estimators with a low learning rate, with each estimator subsampling only 60% of the training sample, so no local feature of the dataset will define the training scheme. In addition a k -folding technique is employed (discussed in detail in Section 4.3) with the datasets split into two distinct folds and two GBR models individually trained with the performance evaluated on the inverse dataset. No over-training is observed and the GBR is shown to be highly performant, with a simple classifier unable to separate the original and target distributions once the new weights are applied.

Separate GBR models are trained for each data taking year, with studies showing that also splitting by magnet polarity provides no improvement and more unstable models due to the reduction in training statistics. After training on the control mode, the models are “frozen” and applied to MC datasets by providing the candidate $p_T(\Lambda_b^0)$, $\eta(\Lambda_b^0)$ and NTRACKS and a corresponding weight being predicted. For the control mode MC only, if the event in question was used to train the model in one of the folds, the opposite fold is utilised for weight prediction, to further prevent any bias from over-training. Other MC samples predict a weight with both folds and take an average for better stability. As discussed, this reweighting technique is not applied to background MC samples where the head particle is not a Λ_b^0 .

4.2.6.1 GBR with nTracks as a Discriminant Variable

As discussed NTRACKS is included in the final model as a powerful discriminant variable, with the shown dramatic change in Figure 4.21c after applying the GBR. This variable is only available in reconstructed datasets however, and there is no equivalent in the generator-level tuples. Ideally, to derive weights for the generator-level tuples, the prediction of weights would temporarily ignore and de-correlate the NTRACKS, but with no in-built way to do this a work-around was developed.

By creating a random number sampler based on the distribution of NTRACKS in the reconstructed MC training dataset, a random and decorrelated yet realistic value for NTRACKS can be generated for each event in the generator-level tuple. This focuses the kinematic weight prediction onto $p_T(\Lambda_b^0)$ and $\eta(\Lambda_b^0)$ which are available at the generator-level. To validate this method, weights are predicted using both techniques using the

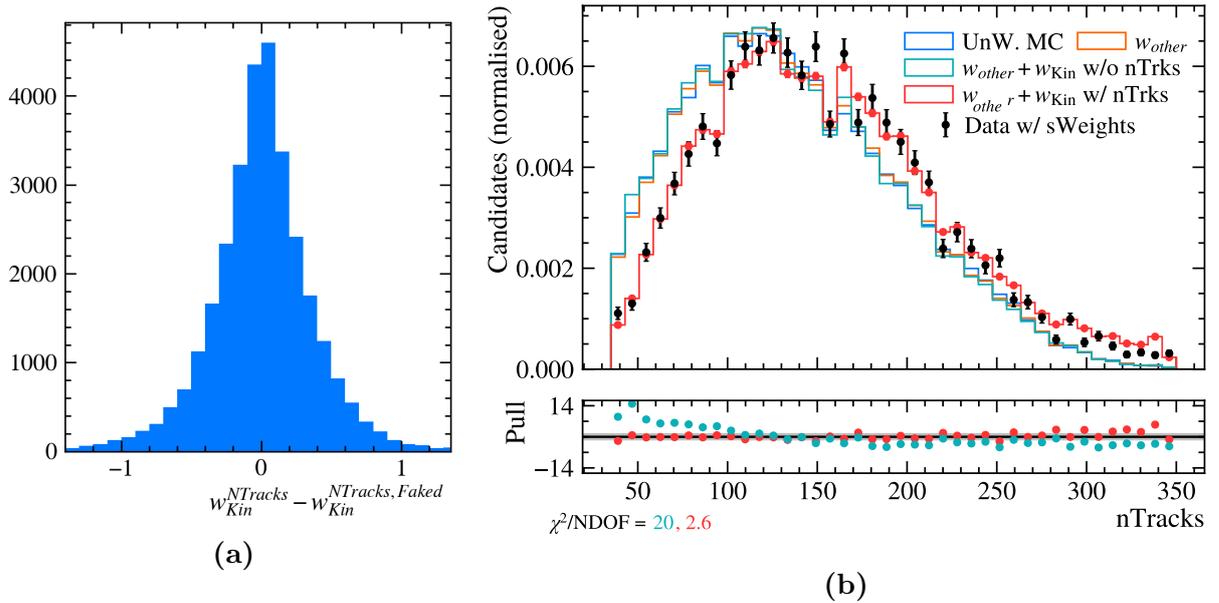


Figure 4.20: Justification for including nTRACKS in the GBR. (a) The difference in the kinematic weight derived using the real nTRACKS variable versus when generated using the sampler. (b) The corrected distribution of nTRACKS using the GBR with (in red) and without (in cyan) including nTRACKS in the training. The initial MC distribution without a kinematic (blue) is included for reference. The black points is the sWeighted data which the GBR targets, with the pull distributions of both model types calculated with respect to this.

same GBR model, where in one case the actual nTRACKS variable is used and the other it is generated from the sampler. The difference in the weight predicted is shown in Figure 4.20a, with the symmetric distribution centred on 0 implying that both methods would produce the same $\sum w_{PID}$, the important metric for generator-level tuples which are not considered event-by-event but as an ensemble (see Section 4.6).

The alternative to using the nTRACKS sampler would be to either build GBR models using only shared generator/reconstruction variables or to have two separate models, one for reconstructed and one for generator-level MC events. The latter could lead to systematic effects on the efficiencies derived, with no cancellation possible in the equations discussed in Section 4.6. Training a GBR without nTRACKS was implemented in the analysis as an option, but the worse alignment shown in Figure 4.20b compared to the nominal model used in Figure 4.21c is enough to reject this option.

4.2.7 Post-Correction Data-MC Alignment

As discussed, if the distributions in sWeighted data align with the MC after reweighting for the control mode $A_b^0 \rightarrow pK^- J/\psi (\rightarrow \mu^+ \mu^-)$, the reweighting has been successful and the

suite of corrections will provide a more accurate calculation of $\varepsilon_{Control}$. The assumption remains that the performance of the corrections will be reflected directly in the signal mode to calculate ε_{Signal} , but the choice of a kinematically and topologically similar control mode maximises the likelihood of this.

Figures 4.15, 4.16 and 4.21 across this section show the correction of a series of kinematic, invariant mass and derived variables, each showing significant improvement after reweighting. The improvement in agreement displayed in $p(\Lambda_b^0)$ and $\chi_{IP}^2(\Lambda_b^0)$ in Figures 4.21d and e are robust proof of the reweighting suite working as these variables were not directly used in any of the discussed methods. In addition the improvement in particle specific variables $p_T(\mu)$ and $p(p)$ (Figures 4.21g and h) imply the corrections applied to Λ_b^0 candidates are being propagated to the final state particles.

Residual differences do remain, including in the low momentum region of Figures 4.21e, g and h as well as in Figure 4.16. This shows that the methods discussed are not a solution for all data and MC differences, and generally the agreement of variables should be considered before use in measurements. The remaining residual difference will be accounted for by the propagation of a systematic uncertainty from each method to the efficiency calculation, which these weights directly affect (see Equation 4.6.1). This is clarified in more detail in Section 4.7. The overall agreement shown on the key variables targeted by the reweighting is considered sufficient to conclude that all reasonable corrections have been applied, and that the corrected simulation will better represent the datasets they correspond to when deriving efficiencies and performing invariant mass fits.

4.3 Multi-Variate Algorithm Selection

Upon application of the cut-based selection the datasets have been cleaned, with the relative background level reduced significantly. Figure 4.14 shows the resulting $pK\mu\mu$ invariant mass spectrum and Figure 4.3 and 4.10 describe the changing μeOS and μeSS distributions as the selection is applied in stages. Despite the reduction, the significant remaining combinatorial background component must be further suppressed to achieve the near-zero background aim of the analysis. This is achieved via the application of a Multivariate Analysis (MVA).

This analysis follows the common scheme for the application of an MVA in LHCb, using corrected $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ signal MC as the training “target” sample and the upper data-sideband in μeOS as the training background “rejection” sample. The upper sideband only is used as this is expected to be dominated by combinatorial background, which the MVA aims to reject, rather than the lower sideband, which is a mixture of

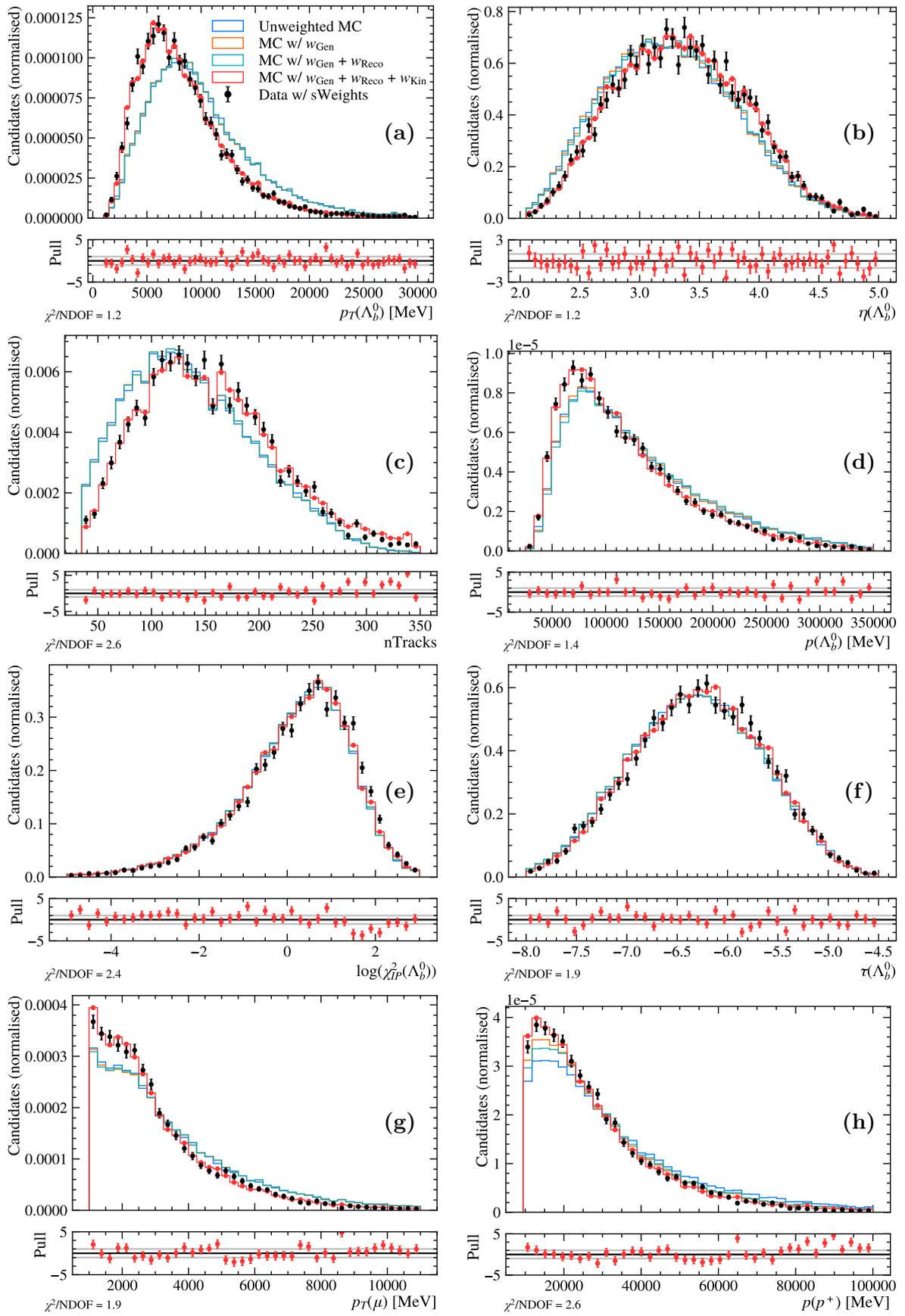


Figure 4.21: Important analysis variables in control mode simulation as different stages of the correction chain are applied (coloured lines) with the corresponding sWeighted data (black points) overlaid to show the expected distribution. The pull distribution between the fully corrected MC (red line) and sWeighted data is shown. The legend is shared.

combinatorial and exclusive semileptonic backgrounds as seen from the excess in the μe OS lower sideband above the μe SS compared to the small difference in the upper sideband in Figure 4.10.

To maximise the statistics of the training samples, all analysis years are combined when training the MVA. An equal amount of target and rejection events are retained as well as ensuring the ratio of Run 1 to Run 2 events in the MC target sample is the same as in the data sideband training sample. This is to avoid the introduction of potential bias between the samples and years and is performed by randomly removing MC events when building the training datasets. In addition, any events with multiple A_b^0 candidates are cleaned to leave a single-candidate per event, using the same random, reproducible method discussed in Section 4.5. Subsequently, the working point of the MVA is determined independently for each analysis category, allowing for differences in performance to be normalised (see Section 4.4).

Rather than the often-used technique of a Boosted Decision Tree (BDT) (such as TMVA, XGBoost [171, 172]), the nominal technique for this analysis is a Keras Sequential model [173]. This was initially developed as a masters thesis project at the University of Birmingham [174, 175] where students were provided with datasets and designed a range of models, with the most performant deployed into the analysis chain and used as described in this section. The advantage of NN approach over a BDT is the suite of in-built tools available for training, with the focus of NN and machine learning development generally to produce a maximally complex and intricate model with reliable performance, often when trained on small datasets. This includes tools to automatically prevent overtraining as well as training over many “epochs” to ensure the best performance is found. This suits the problem of rejecting combinatorial background for this analysis, as a simple binary classification effort, where the best performance is desired for relatively small input datasets, but robust performance is paramount.

In contrary to many analyses, this MVA is derived and applied after pre-selection but before any PID selection. This is due to the extensive pre-selection discussed, with very limited statistics remaining in the data sidebands, as summarised in Table 4.7.

4.3.1 Keras Sequential Neural Network

The Keras sequential model is employed for its simplicity as a stack of layers, with each receiving input from the previous only and passing output to the subsequent layer only [173]. This reduces the likelihood of over-training significantly and allows the intricacies of the model to be scrutinised as visible weights between nodes in neighbouring layers. The output parameter of interest is the “signal probability” p_{MVA} , a likelihood measure that

Table 4.7: A summary of the statistics of the data and simulation signal samples with all selection applied up to the MVA stage. The μe SS yields are represented as sidebands. The μe OS upper sideband statistics used for training the MVA are also shown, where the M_{pK} cut is loosened to $1450 < M_{pK} < 1850$ MeV to improve the model performance.

Sample	Total Number of Events for MVA Training				
	MC	μe OS Sidebands		μe SS Sidebands	
		Low	Upper (Loose M_{pK} Cut)	Low	Upper
Run 1 0γ	37917	1505	490 (2756)	470	317
Run 1 1γ	41283	902	373 (1762)	289	211
Run 2 0γ	65308	2194	632 (2705)	760	451
Run 2 1γ	70224	1814	583 (2414)	850	440

a given candidate is signal.

A series of model designs were tested of varying complexity, but the model converged on a two layer approach with 8 and 12 nodes, complex enough to produce the desired signal-background separation but with only 241 weights to train $N_w \ll N_{\text{events}}$ further reducing the likelihood of over-training. Between layers, batch normalisation [176] is employed to regulate and normalise the weights and over-training is prevented with a kernel regulariser [177] and a low learning rate that reduces as the model converges towards its plateau in performance.

The Keras model is trained in epochs where the state of the model after one epoch is used as the start point for the next. If over-training is detected the model will revert to the last stable point, a useful feature for automatically preventing over-training. At the performance plateau the model will stop and revert to the best performing epoch, hence while the maximum number of epochs is set to 2000, models very rarely reach this limit during training.

During training the datasets are split into three partitions: training, validation and testing. The training data is used directly by the model to draw correlations and derive weights between nodes of the NN. At each epoch, the validation dataset is used to automatically check whether the NN has over-trained, by checking that the performance that is seen in the training dataset is reflected in the separate validation dataset. The testing dataset is kept completely independent and is not used for training in any way, such that, post completion of the MVA training, the performance can be evaluated independently and over-training can be checked for. If the model performs significantly better on the training and/or validation datasets, the model is over-trained.

For this analysis, the datasets are split by 60% training, 20% validation and 20%

testing. By default, this would result in 20% of the dataset unable to be utilised to train the MVA at all, potentially reducing performance. Also, 20% of the dataset is a relatively small sample to evaluate the important independent performance of the MVA considering the statistics in Table 4.7. To remedy this, the k -folding method is employed [178], training five different independent neural networks, each with a distinct 20% testing dataset. Hence the complete analysis samples can be utilised to independently test the MVA performance by applying each fold’s 20% test data to the corresponding model. The remaining 80% for each model is used for training and validation as discussed.

To evaluate the signal probability measure for an incoming external dataset the default behaviour of the MVA is to measure the response on an event with all five folds, and average the probability. This gives the most stable response, averaging out any fluctuations in performance between folds. Performance in individual folds is studied as a cross check, as discussed in Section 4.3.3.

4.3.2 Variable Study

The selection of variables is paramount as their separation power defines the performance of the NN while any erroneous correlations could incorrectly train the model to select features inconsistent with signal events. Potential differences in the variables between Run 1 and 2 were monitored when selecting variables as the model will not know to separate the two when training so an intrinsic difference will damage the overall performance of the model.

Another important consideration was that the variables used were well-modelled in simulation, preventing the algorithm separating based on correlations that originated from simulation mis-modelling which would not appear in real data. This is checked by comparing control mode data and corrected simulation, discussed in Section 4.2. The MVA is derived on $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ signal samples and applied to both the signal and control mode, therefore variables that are electron exclusive variables are not considered.

4.3.2.1 Prior χ_{DTF}^2 Cut

On initial testing it became clear that the Decay Tree Fitter (DTF) goodness of fit measure “ χ_{DTF}^2 ” outperforms all other variables considered. This resulted in models trained with this variable focusing on this one attribute, suppressing discrimination power from other variables. Hence, tests were carried out where a loose χ_{DTF}^2 cut was applied to the samples prior to training, with initial signal-background separation possible with $\chi_{\text{DTF}}^2 < 30$ as shown in Figure 4.22a. This allows the more complicated discrimination to be

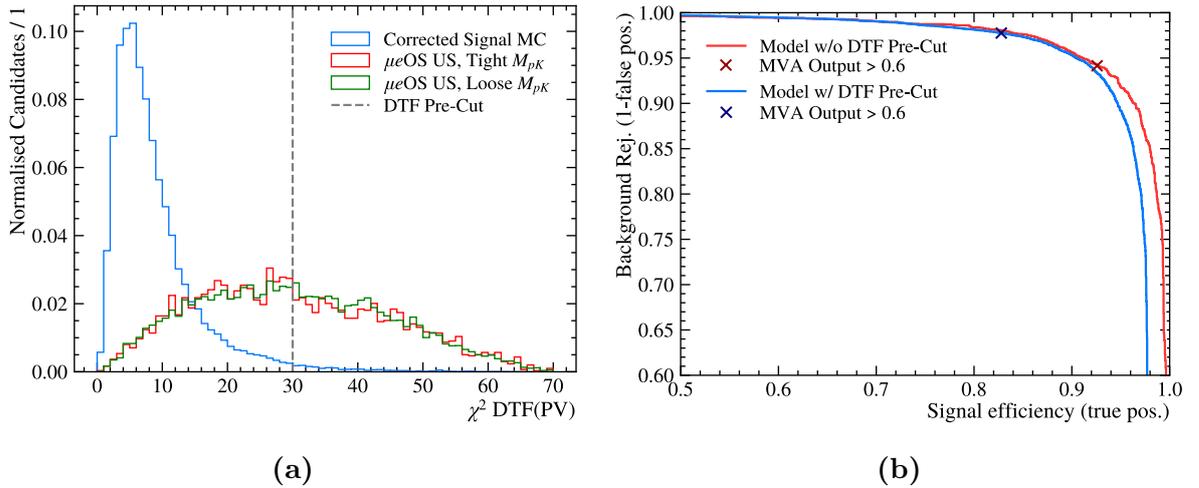


Figure 4.22: Justification for use of χ^2_{DTF} prior cut. (a) Distribution of χ^2_{DTF} in signal MC and the upper sideband μe OS data (the signal and background samples for the MVA). The μe OS distribution with the tight M_{pK} cut re-applied is also displayed. (b) Performance comparison for the MVA trained with and without applying the prior cut. The efficiencies stated for the nominal MVA include the pre-cut applied.

performed by the algorithm, significantly increasing relative performance of other variables as shown in Figure 4.26. In terms of selection performance, Figure 4.22b shows that, when combining the signal selection and background rejection into a ROC curve, the MVA with the prior χ^2_{DTF} cut is able to provide comparable performance overall, but crucially better background rejection at a fixed working point such as the $p_{\text{MVA}} = 0.6$ shown. Additionally, the working point is in a more stable location for the MVA with the prior cut, where slight variations in response will result in only a minor change in signal-background efficiency in comparison to the MVA without the χ^2_{DTF} cut. Since the χ^2_{DTF} selection is associated with the MVA, all quoted ε_{MVA} in subsequent sections refers to the compound efficiency of the χ^2_{DTF} cut and the MVA selection.

This “pre-MVA” selection significantly reduces the statistics in the signal data sidebands, and would reduce the number of events in the MVA training background sample by 60%, to 2166 events. Training the MVA with this ensemble was attempted, but the final fitted state was unstable and not as performant as had been observed before applying the χ^2_{DTF} selection. Therefore, the tight M_{pK} cut was loosened to $(1450 < M_{pK} < 1850)$ MeV for the MVA training only, allowing more combinatorial background from within the upper sideband. This reverts to the control mode M_{pK} selection, and improved the MVA performance as expected due to the increased sample size as shown in Table 4.7. Figure 4.22a describes how the prior-cut performance is consistent with the looser M_{pK} cut. This approach was also taken in the $B_{(s)}^0 \rightarrow X\mu^\pm e^\mp$ analysis, loosening the ϕ and K^{*0} selection accordingly [112].

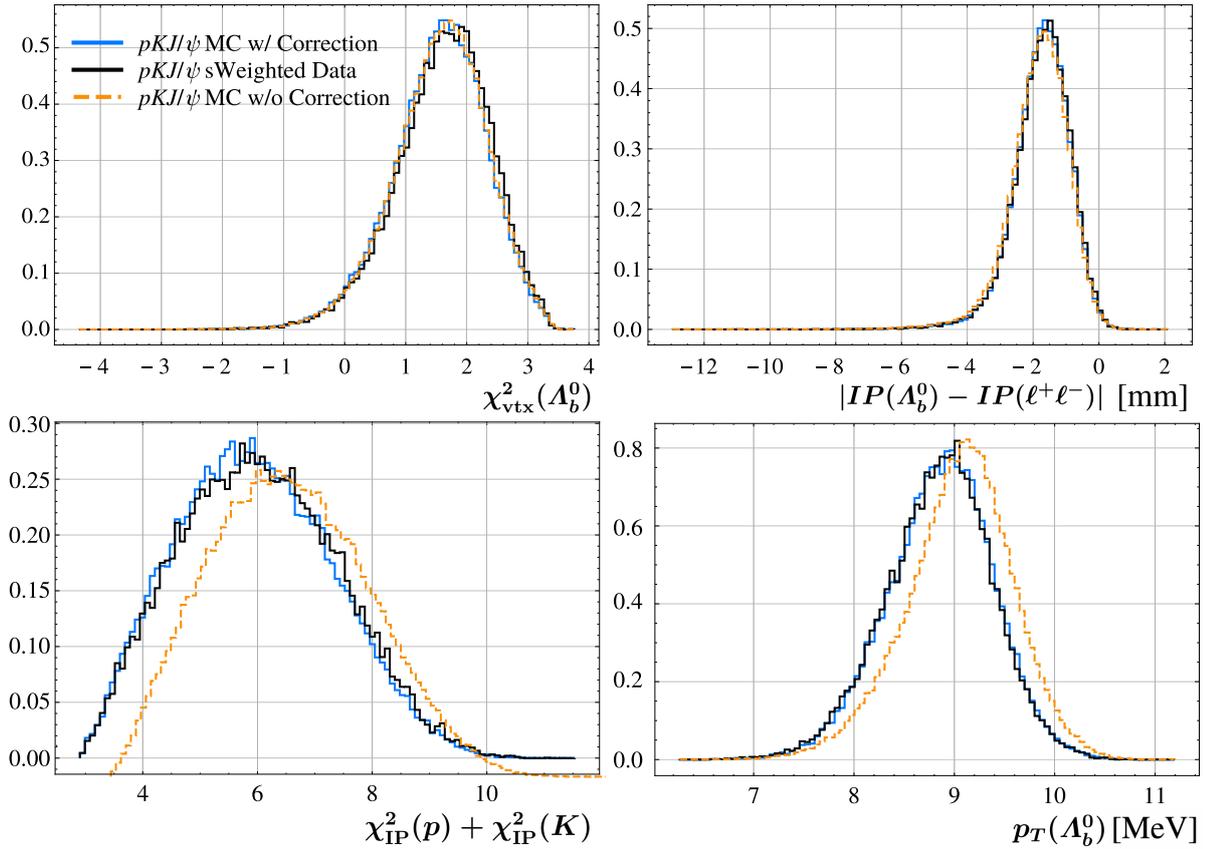


Figure 4.23: A selection of the most discriminant variables used to train the MVA, with the corrected $\Lambda_b^0 \rightarrow pK^- J/\psi (\rightarrow \mu^+ \mu^-)$ simulation and sWeighted data shown to have good agreement, implying that in the signal mode these variables are also well-modelled (where the same sWeight comparison is not possible). The simulation without weights is also included to display the agreement that the weights provide. All variables have the natural logarithm applied. The remaining distributions are shown in Figure B.2.

4.3.2.2 Variable Selection Process

Three methods are used to select variables. First of all any variables that are not well-modelled in simulation are discarded. Distributions of a selection of the final variable set are shown in Figure 4.23, for the control mode samples, allowing a comparison between simulation and background-subtracted data, showing good agreement. The fact that many distributions have good agreement before correction also implies that the MVA training is robust to any potential systematic issues in the weighting scheme.

An ideal MVA training variable will show separation between the μe OS upper-sideband rejection sample and the signal MC target sample, providing discriminating power for the MVA, while also showing consistency with the μe SS distribution, ensuring that the variable will provide the same discrimination across the $M(\Lambda_b^0)$ spectrum when the MVA is applied to the blinded signal region. This is monitored by overlaying the target

and rejection variable distributions along with the μe SS distribution for $M(\Lambda_b^0) < 5800$ as in Figure 4.24, with sufficient separation and agreement shown in each case. The μe OS lower sideband data is also checked to understand if any variables behave very differently between the upper and lower sideband, but differences here could be caused by physics backgrounds rather than the variable having significant mass correlation.

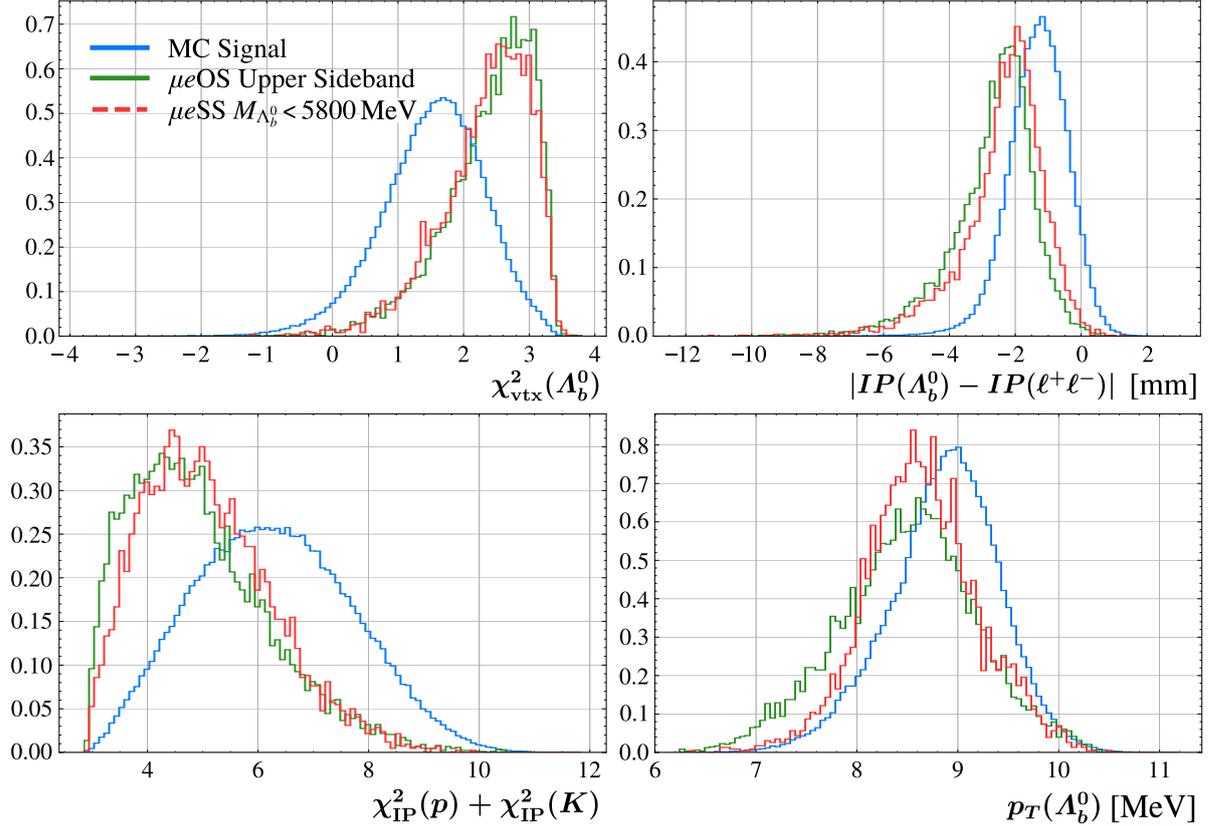


Figure 4.24: Target/Rejection sample distributions of a selection of the most discriminant variables used to train the MVA. The μe SS distribution for $M(\Lambda_b^0) < 5800$ is included to ensure the variable will remove background across the analysis $M(\Lambda_b^0)$ window. All variables have the natural logarithm applied. The remaining distributions are shown in Figure B.1.

Finally, correlation between the training variables is studied including important external variables, by plotting a Pearson correlation [179] coefficient grid as shown in Figure 4.25. Variables which are exploiting similar discrimination power can be identified by an abnormally high correlation, with the less performant variable removed or combined into a single measure. All training variables are also checked with the mass and analysis category variables, removing any with significant coefficients that cannot be justified, to prevent the possible introduction of a mass/category bias in the MVA response.

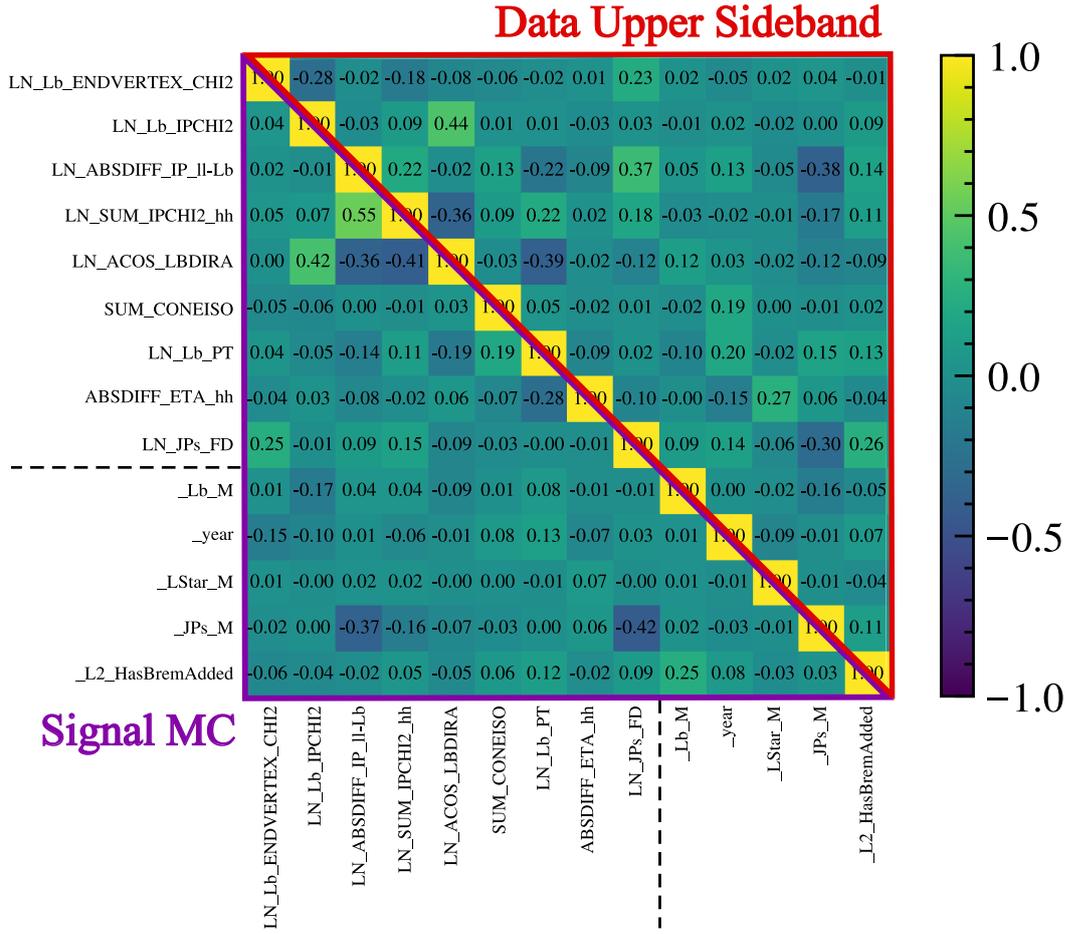


Figure 4.25: Correlation between MVA discriminating variables, shown as a 2D matrix. The lower left half of the matrix (purple surrounding) shows the correlations within the corrected signal MC dataset and the upper right half (red surrounding) shows the equivalent upper sideband measures. Some variables not used in training are included in the matrix (marked by starting with “_”, and divided by the dashed line) to allow checks for undesired large mass and analysis category correlations. The Pearson correlation technique is used [179].

4.3.2.3 Final Variables

Due to the complexity of the final model, nine variables were selected. These are as follows and their final distributions are shown in Figure 4.24 and Figure B.1 (their text-based names used in the figures are shown in brackets for reference):

$\chi^2_{\text{vtx}}(\Lambda_b^0)$ (LN_Lb_ENDVERTEX_CHI2) Quality of the vertex fit for the Λ_b^0 decay vertex.

$\chi^2_{\text{IP}}(\Lambda_b^0)$ (LN_Lb_IPCHI2) Quality measure of association between the momentum vector of the reconstructed Λ_b^0 candidate and the PV it is associated to, providing a measure of likelihood of the candidate originating from that pp collision, as defined in

Section 2.2.1.

$p_T(\Lambda_b^0)$ (LN_Lb_PT) Transverse momentum of the reconstructed Λ_b^0 candidate.

Λ_b^0 **DIRection Angle (DIRA)** (LN_ACOS_LBDIRA) The cosine of the angle between the Λ_b^0 reconstructed momentum vector and the position vector between its primary and decay vertex, providing a measure of consistency in momentum and coordinate space.

$\chi_{\text{IP}}^2(\mathbf{p}) + \chi_{\text{IP}}^2(\mathbf{K})$ (LN_SUM_IPCHI2_hh) Sum of the χ_{IP}^2 hadronic measures, ensuring separation from the PV, but not the large displacement consistent with a secondary vertex from a long lived hh candidate. Variable combined due to correlated performance shown individually in a previous iteration of the MVA.

$|\boldsymbol{\eta}(\mathbf{p}) - \boldsymbol{\eta}(\mathbf{K})|$ (ABSDIFF_ETA_hh) Provides a handle on the angular distribution of the hadronic system.

J/ψ **FD** (LN_JPs_FD) Flight distance of the $\ell^+\ell^-$ object (J/ψ) from its origin vertex (the Λ_b^0 decay vertex).

$|\text{IP}(\Lambda_b^0) - \text{IP}(\ell^+\ell^-)|$ (LN_ABSDIFF_IP_11-Lb) Difference between the IP values for the reconstructed Λ_b^0 and $\ell^+\ell^-$ candidates, ensuring the leptons are displaced with respect to the head particle.

Sum Cone Isolation (SUM_CONEISO) For each final state particle track, the p_T values of other tracks in a surrounding cone is summed, providing a measure of isolation of each particle. These are summed to provide a global isolation measure.

Certain variables have the natural logarithm applied if it was visible that the separation power increased by translating the distribution to a more regular shape (those with text-based names beginning with “LN_”). Additionally, neural networks have been shown to perform better with variable distributions centred around 0 with a standard deviation of unity [180], hence a column transformer [181] is used to transform the target and rejection samples in each variable before beginning the MVA training. The transformers use the training dataset only to prevent leakage from the other MVA datasets introducing any possible bias. The transformers are then frozen and saved so external datasets will always be transformed in the same way, providing an exact and reliable map between the analysis variable and the NN variable. The final variable importance is shown in Figure 4.26. It is clear that there is not one variable that absolutely dominates, a desirable feature of the final model.

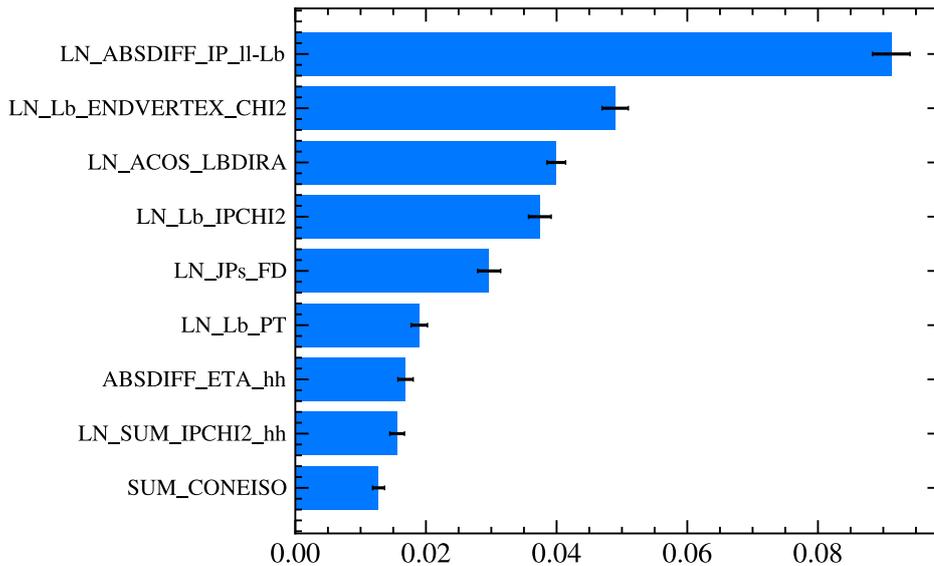


Figure 4.26: Relative feature importance for the input variables used to train the MVA to discern between signal and background. Importance is calculated by randomising each variable in turn and measuring the reduction in performance of the MVA (hence determining how much the variable “adds”). “LN” refers to the natural logarithm of the variable being used.

4.3.3 Performance, Validation and Optimisation

The performance of the MVA is evaluated using similar techniques to most LHCb analyses which employ BDTs as a base model. Figure 4.27 explicitly displays the performance, stability and robustness of the model trained as a function of the p_{MVA} output variable. The clear separation between the signal and background datasets in signal probability shows that the MVA is performing correctly and that a cut anywhere > 0.25 would remove $> 85\%$ of background and retain up to 95% of signal. In addition, the flatness observed in the central $0.25 < p_{\text{MVA}} < 0.75$ region (where the final selection will be made) demonstrates the stability of the MVA. If the response shown for signal simulation was slightly different to that on real data, the chosen p_{MVA} cut-value would result in only a minor change in efficiency. The figure also displays the lack of over-training in the model, with the signal probability of the test datasets inline with the training dataset across the entire p_{MVA} range in both signal and background. This shows that the MVA has trained on global correlations of the dataset rather than individual event characteristics. This is quantitatively represented by the KS p-values which state that there is a $> 85\%$ probability that the train/test dataset responses were drawn from the same distribution.

Figure 4.28 also displays the consistent high performance of the MVA, with a large AUC across all datasets, and no distinct or significant over-training with the three curves

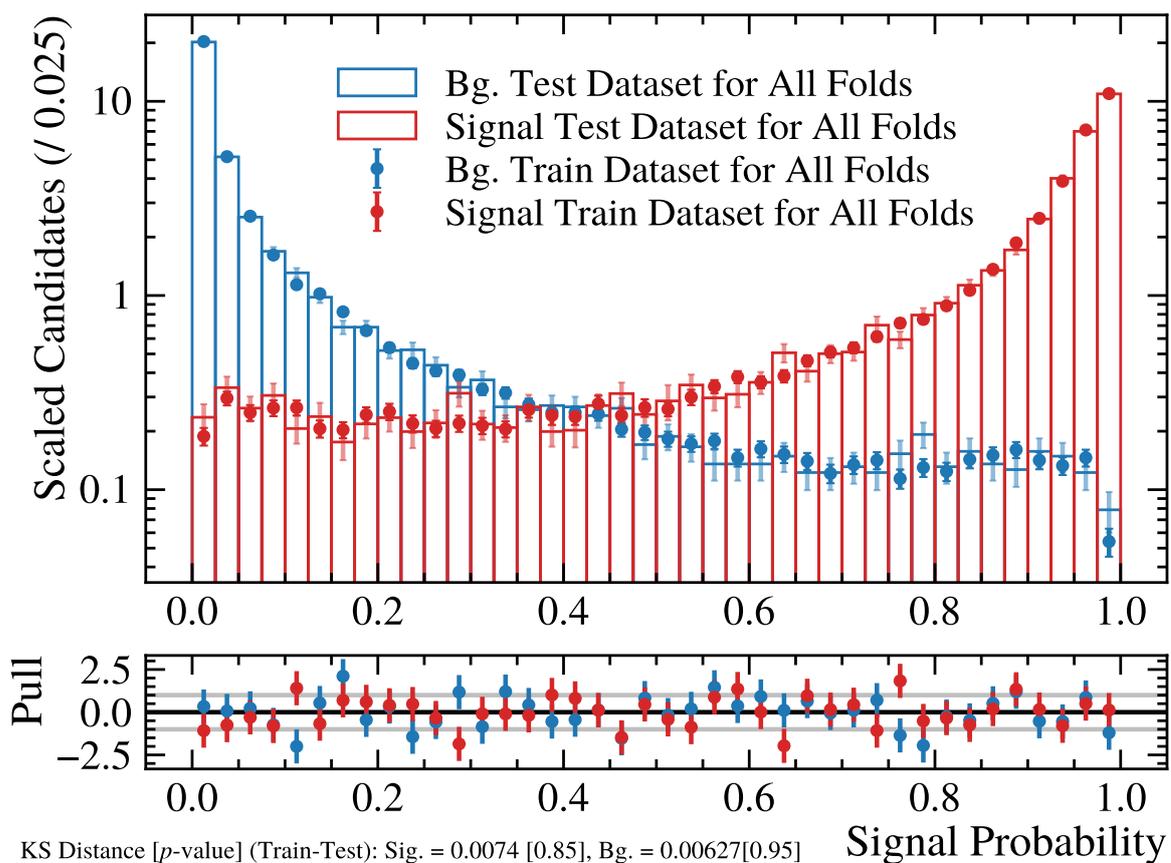


Figure 4.27: Overlaid MVA response probability distributions for combined signal and background datasets used in the MVA training / testing. Kolmogorov-Smirnov distance (and p-values) are evaluated between the respective train/test datasets’ signal probability distributions and are close to 0 (1) for datasets that are statistically similar [182, 183].

not separable at considered p_{MVA} cut-values, with Area Under Curve (AUC) values ranging $\leq 0.2\%$. The MVA delivers high background rejection ($> 90\%$) while retaining very good signal efficiency ($> 80\%$) in the stable region discussed.

The performance of each fold is also individually evaluated to check consistency. Figure 4.29 displays that while across the folds there is natural variation in MVA response, each shows high separation. Overall the performance is consistent both between folds and within each fold between individual test and train datasets. This supports the decision to average across all folds as the default behaviour when extracting a p_{MVA} value, giving a more precise and robust prediction overall.

The performance of the MVA is expected to be similar across the different analysis categories, however Figure 4.30 displays the Run 2 categories out-performing Run 1 consistently and that the events where the electron has a Bremsstrahlung photon reconstructed perform better. There are a few possible reasons for this, as discussed (see

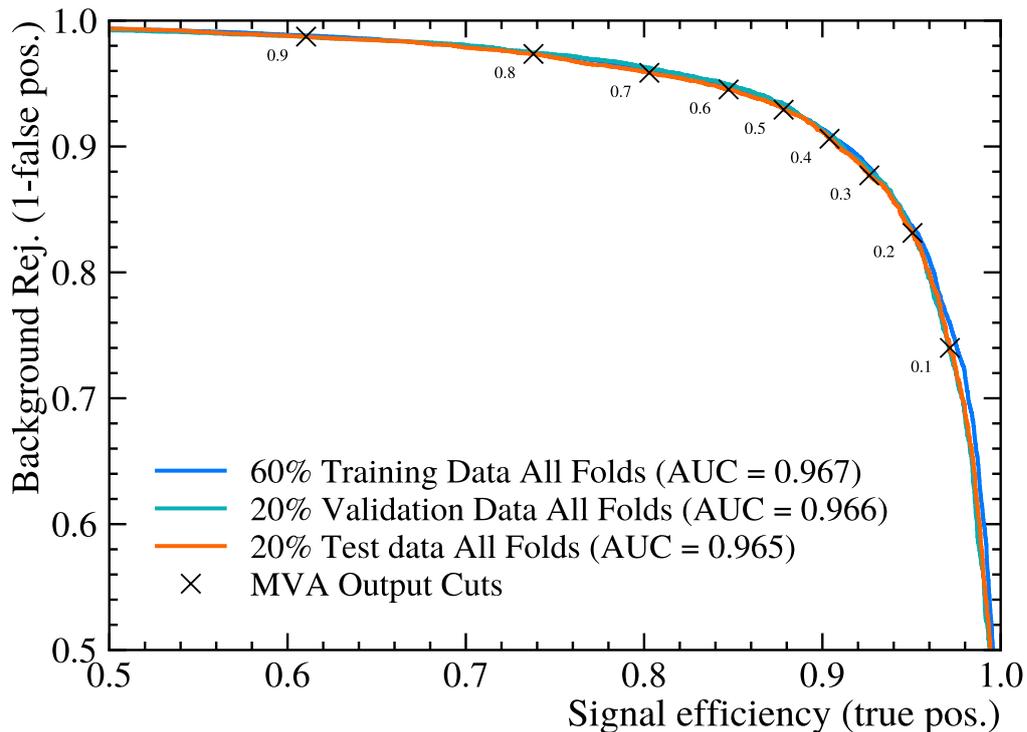


Figure 4.28: ROC curves for the training, validation and test datasets with the “Area-Under-Curve” (AUC) included and points marked where the given p_{MVA} cut produces the corresponding efficiencies.

Table 4.7) there are more events in Run 2 than in Run 1 in the overall analysis dataset, and the MVA trains with an ensemble of both, potentially introducing a bias towards performance in Run 2. In addition, there is an imbalance in the number of events in the different Bremsstrahlung categories. However forcing all four categories to contribute equally to the MVA datasets did not alter the performance in the figure.

This implies that the differing performance is down to the reconstruction quality of the variables used for training. If a variable is better reconstructed in one category and has a more distinct distribution between the target and rejection sample, it will better separate events. The LHCb Run 2 detector had improved reconstruction over Run 1 (see Section 2.4.2) and recovering a Bremsstrahlung photon will, in most cases, improve the quality of the electron reconstructed, explaining the trend seen.

With greater statistics available, an individual MVA for each category could be trained which potentially would give better global performance of the MVA. But as discussed, the MVA is already limited by the low number of background events. The differing performance shown prompted the approach discussed in Section 4.4.2 where the chosen p_{MVA} cuts are optimised on a per-category basis.

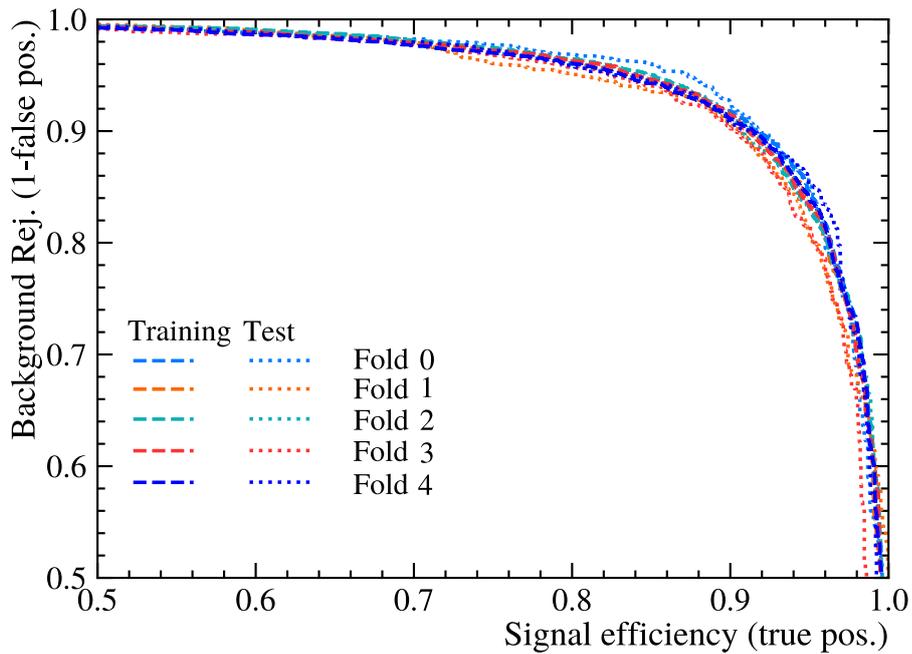


Figure 4.29: MVA performance for the individual trained folds, showing the overlaid ROC curves. The AUC values are evaluated to be consistent across folds (within 0.5% of each other), with all train and test datasets within each fold consistent to $< 0.5\%$ level in AUC.

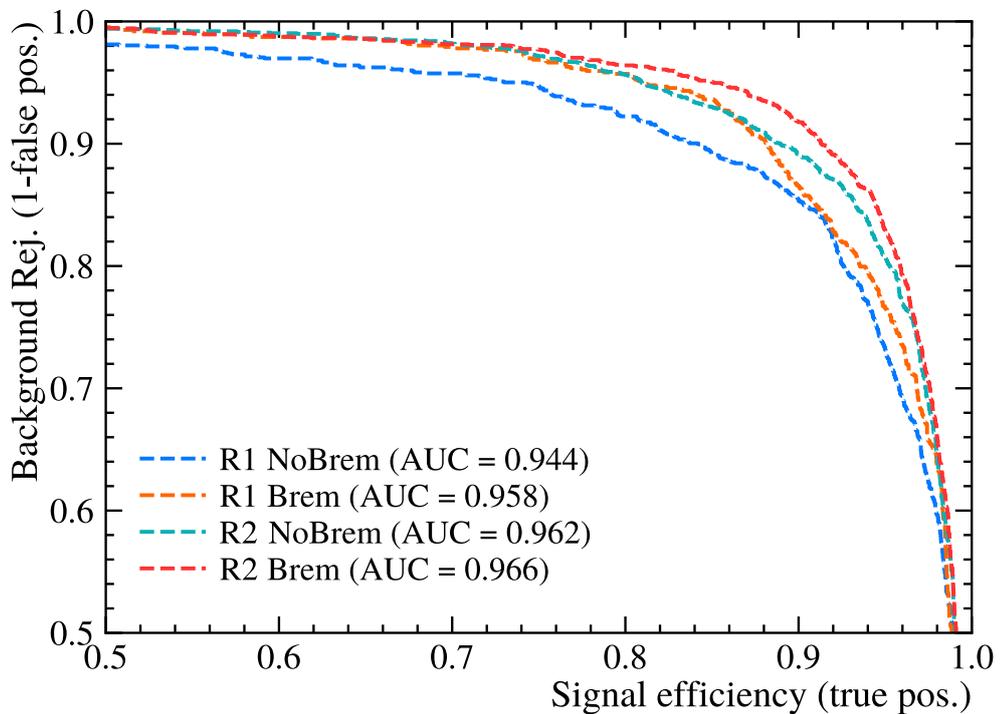


Figure 4.30: ROC Curves evaluating MVA performance for the different analysis categories.

Figure 4.31a visualises the response of the MVA as a function of mass in the signal data sidebands. This acts as a check for any artificial peaking structures being introduced or specific biases in the MVA performance in different mass ranges. This is an especially important check since the MVA is trained on upper-sideband μe OS data only, and is then applied to the full mass range to remove combinatorial background. No biases or erroneous structures are observed.

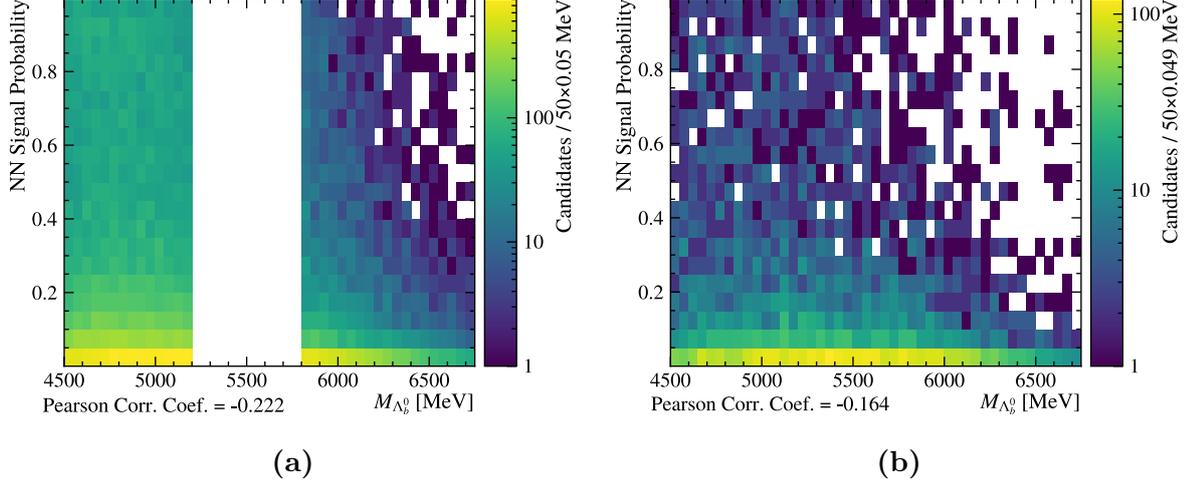


Figure 4.31: p_{MVA} output against the Λ_b^0 invariant mass distribution for: (a) the lower and upper sidebands in the μe OS dataset and (b) the combinatorial proxy μe SS dataset. The Pearson correlation coefficients are quoted.

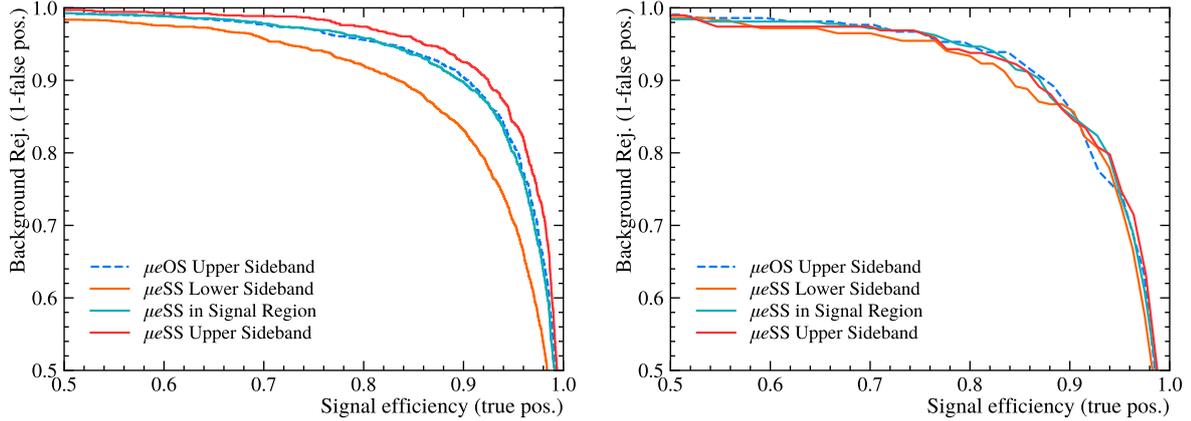


Figure 4.32: Performance comparison for the MVA response on μe SS background data split into three regions of $M(\Lambda_b^0)$ corresponding to the lower sideband, upper sideband and signal region with the curve for the μe OS upper sideband (with the tight M_{pK} selection) overlaid. The distributions before (left) and after (right) PID selection are shown (see Section 4.4).

4.3.3.1 Validation with Combinatorial Background Proxy

To evaluate the MVA performance at removing combinatorial background across the full $4500 < M(\Lambda_b^0) < 6750$ MeV analysis range, the μe SS dataset is analysed split by mass region in Figure 4.32 and the response as a function of mass is tested in Figure 4.31b. The ROC curve displays that before any PID selection the μe SS upper sideband out-performs the μe OS equivalent, with an anti-correlation with $M(\Lambda_b^0)$ implied from the signal and lower sideband regions. However, after applying the PID selection (discussed further in Section 4.4), the μe SS and μe OS performance becomes consistent across the mass regions, implying that the difference seen before PID selection was due to two or three body mis-ID backgrounds that are removed in the final selection. Crucially, the performance in the μe SS signal region is consistent with the MVA background training sample in both cases shown, confirming that the MVA will efficiently remove combinatorial background in the blinded region at the suggested rate shown from the μe OS sidebands and μe SS proxy.

Further justifying the robust response, Figure 4.31b shows no unexpected structure and the majority of combinatorial background is in the lowest p_{MVA} bin, with minor leakage into the higher bins for the higher statistics central region. Comparing this to the distribution in Figure 4.31a, the upper-sideband region shows a similar response, where as the μe OS lower-sideband has a large number of events in the higher p_{MVA} bins highlighting the physical, non-combinatorial backgrounds remaining in the dataset which produce a more “signal-like” response from the MVA. This strengthens the argument for training the MVA with the upper-sideband only.

4.3.3.2 Validation with Control Mode Data

While the MVA has been shown to efficiently select the signal mode simulation, upon unblinding, the MVA must be able to also select any possible signal in data with the same efficiency. Therefore cross-checks with the control mode simulation and data is performed to identify any bias in the MVA response on data versus corrected simulation. This also acts as an additional check of the MC correction suite discussed in Section 4.2.

A comparison of MVA selection efficiency is shown in Figure 4.33 for the control mode simulation and data, with corrections and background-subtraction applied respectively. The background-subtraction applied to the data allows it to be treated as “signal-only” and hence suitable for comparison to the simulation. The response is shown to be consistent between the two modes implying that there is only very minor, if any, bias in the response of the MVA. The figure displays that the $\Lambda_b^0 \rightarrow pK^- J/\psi$ control mode also under-performs with respect to the signal mode, which is expected due to the MVA being

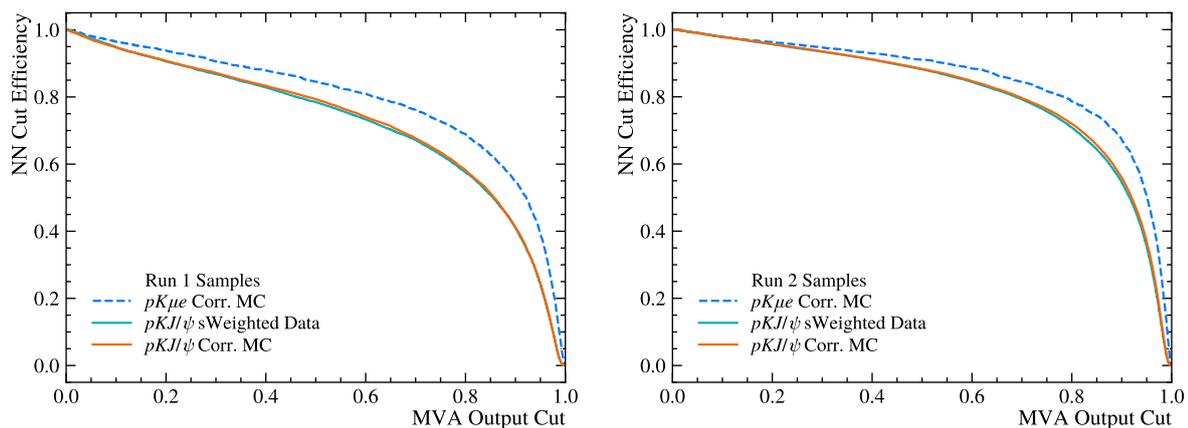


Figure 4.33: Performance comparison of ε_{MVA} at different cut values of p_{MVA} , shown for the corrected control mode simulation and the background-subtracted control mode data (following the sWeight process discussed in Section 4.2). Results are split by analysis category for the control mode, into Run 1 (left) and Run 2 (right). The corrected signal mode simulation response is included for reference.

trained on the signal mode samples. Despite this the performance is still sufficiently high, implying that the MVA has been trained in a general enough way and that the selected control mode data sample is suitable for fitting and extracting a precise N_{Control} for the signal branching fraction in Equation 3.2.1.

4.4 Particle Identification Selection and MVA-PID Optimisation

As discussed in Section 4.1.1, very loose PID cuts are applied at the stripping selection stage of the analysis summarised in Table 4.1. While these remove the clearest cases of particle mis-identification (mis-ID), there still remains a large likelihood of building $M(\Lambda_b^0)$ candidates from the wrong particle species, e.g. about 40% of selected protons after the stripping selection are expected to have a true identity of a kaon. Therefore this PID selection must be tightened, with the aim to reduce the background remaining in the analysis to be almost entirely combinatorial, in-line with the upper sideband.

The PID selection (summarised in Table 4.8) is designed around removing background in the signal mode, and then translated to the control mode and its performance verified. The only mis-ID types that can reproduce the signal mode are hadronic ($h \rightarrow \ell$) and leptonic ($\ell \rightarrow \ell'$) backgrounds. These are prevented and reduced by tightening the leptonic $\text{ProbNN}\ell$ cut values from 0.05 to 0.1. This reduces contamination from the high-rate hadronic backgrounds $\Lambda_b^0 \rightarrow pK^-K^+K^-$ and $\Lambda_b^0 \rightarrow pK^-K^+\pi^-$, where a high-momentum

Table 4.8: The PID selection applied to the analysis. Specific tunes are chosen for the ProbNN variables, following the R_{pK} analysis; for the Run 2 categories the cut type “MC15TuneV1” is used and for Run 1 the “MC12TuneV3” is used in most cases [76].

Particle	PID Selection
p	$\text{ProbNNp} > 0.3 \ \& \ \text{ProbNNk} < 0.8 \ \& \ \text{ProbNNpi} < 0.7$
K	$\text{ProbNNk} > 0.2 \ \& \ \text{ProbNNp} < 0.8$
μ	$\text{ProbNNmu} > 0.1$
e	$\text{ProbNNe} > 0.1 \ \& \ \text{PIDe} > 5$

h “punches through” to the muon chambers as well as a low-momentum h showering in the ECAL emulating a muon or electron respectively. It also prevents the kinematically similar analogous FCNC modes $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ and $\Lambda_b^0 \rightarrow \Lambda(1520)e^+e^-$ from contaminating where a low/high momentum lepton emulates the opposite type. The $\text{PIDe} > 5$ additional cut is made after $B^+ \rightarrow K^+K^+K^-$ events were discovered to be contaminating the rare decay $B^+ \rightarrow K^+e^+e^-$ in the R_X analysis at a higher rate than expected [71, 74].

While hadronic mis-ID ($h \rightarrow h'$) is possible and present in the analysis, it must accompany one of the previous mis-ID types in order to reproduce the $pK\mu e$ final state particles. Nevertheless, tighter hadronic PID selection is applied to prevent contamination from the decays of the higher production rate b -mesons into similar final states to those discussed above. These are applied in two ways, by requiring a minimum cut on $\text{ProbNN}h$ as well as a maximum limit on other hadronic $\text{ProbNN}h'$, as seen in the proton and kaon cuts in Table 4.8. These cuts especially target the high rate $p \leftrightarrow K$ mis-ID, which could lead to true pK^- candidates being misidentified as $K^+\bar{p}$ or light resonances decaying to K^+K^- contaminating the M_{pK} spectrum.

The suite of selections in Table 4.8 were mostly taken from the R_{pK} analysis which investigated small adjustments around these values and found these to be the most stable. The new additions for this analysis are the PIDe selections as well as bringing the Run 1 cut in line with the Run 2 [76]. This selection scheme was also used for the $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ branching fraction analysis and keeping these $\Lambda_b^0 \rightarrow pK^-\ell^+\ell^-$ analyses aligned enables simpler cross-checks.

The power of the PID selection is displayed in Figure 4.34, where the large excess in the $\mu e\text{OS}$ distribution above the $\mu e\text{SS}$ distribution is effectively removed by the tightening of the selection. The ability to purify and retain Λ_b^0 candidates in data is also displayed in Figure 4.35a and b, rendering a narrow signal peak in both analysis categories of the control mode, with the relative level of the surrounding background reduced. Ensuring that any pK^- candidates selected in the analysis originate from the $\Lambda(1520)$ resonance is

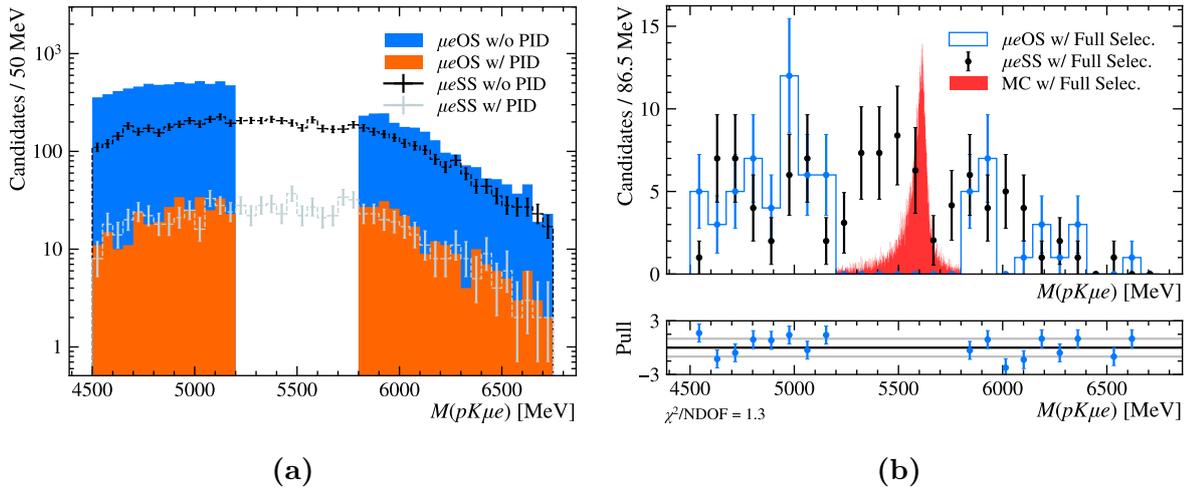


Figure 4.34: The effect of applying the nominal PID selection on the signal mode blinded μeOS data and the μeSS proxy on the combined Run 1 and 2 dataset. Figure (a) displays the reduction in the μeOS excess over the μeSS when PID cuts are applied, with no p_{MVA} cut applied for clearer demonstration. The trend is consistent across analysis categories, with the amount of background above the μeSS reducing by a compatible amount. Figure (b) displays the consistency between μeOS and μeSS after all selection is applied including MVA and PID selection, with a scaled MC distribution overlaid to show the signal shape.

vital to discovering or setting a first limit of the $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ branching fraction. This is cross-checked using the control mode, where the only mis-ID present should be hadronic due to the high rate J/ψ selection. Figure 4.35d shows the relative level of the $B^0 \rightarrow K^{*0} J/\psi$ background before and after applying the PID cuts, reducing $h \rightarrow h'$ mis-ID by over 90%. The equivalent hadron-only mis-ID source is not possible in the signal mode, with the tight limits set on the $B_{(s)}^0 \rightarrow X\mu^\pm e^\mp$, ($X = K^{*0}, \phi$) decays [112]. The hadronic PID selection chosen is also shown to be effective enough to purify the $\Lambda(1520)$ selection, as seen in the M_{pK} plot in Figure 4.35c.

4.4.1 PID Selection Weights

As discussed in Section 4.2.3, PID selection for simulation samples in the analysis is replaced by data-driven techniques using PID weights. New weights are hence derived using the same technique as before with the new selection in Table 4.8 for each particle species, as well as the different mis-identification combinations to be used for the background samples (to be used in Section 5.1). This is performed for each year and magnet polarity individually. The new weights are derived with the loose stripping PID cuts as the denominator, resulting in the final $w_{PID}^{Strip+Selec} = w_{PID}^{Strip} w_{PID}^{Selec}$. Figure 4.36 displays important examples of 2D maps of $w_{PID}^{Strip+Selec}$, confirming the low mis-identification rate of $h \rightarrow \ell$ and $\ell \rightarrow \ell'$ with the PID selection applied, which this analysis relies on to return

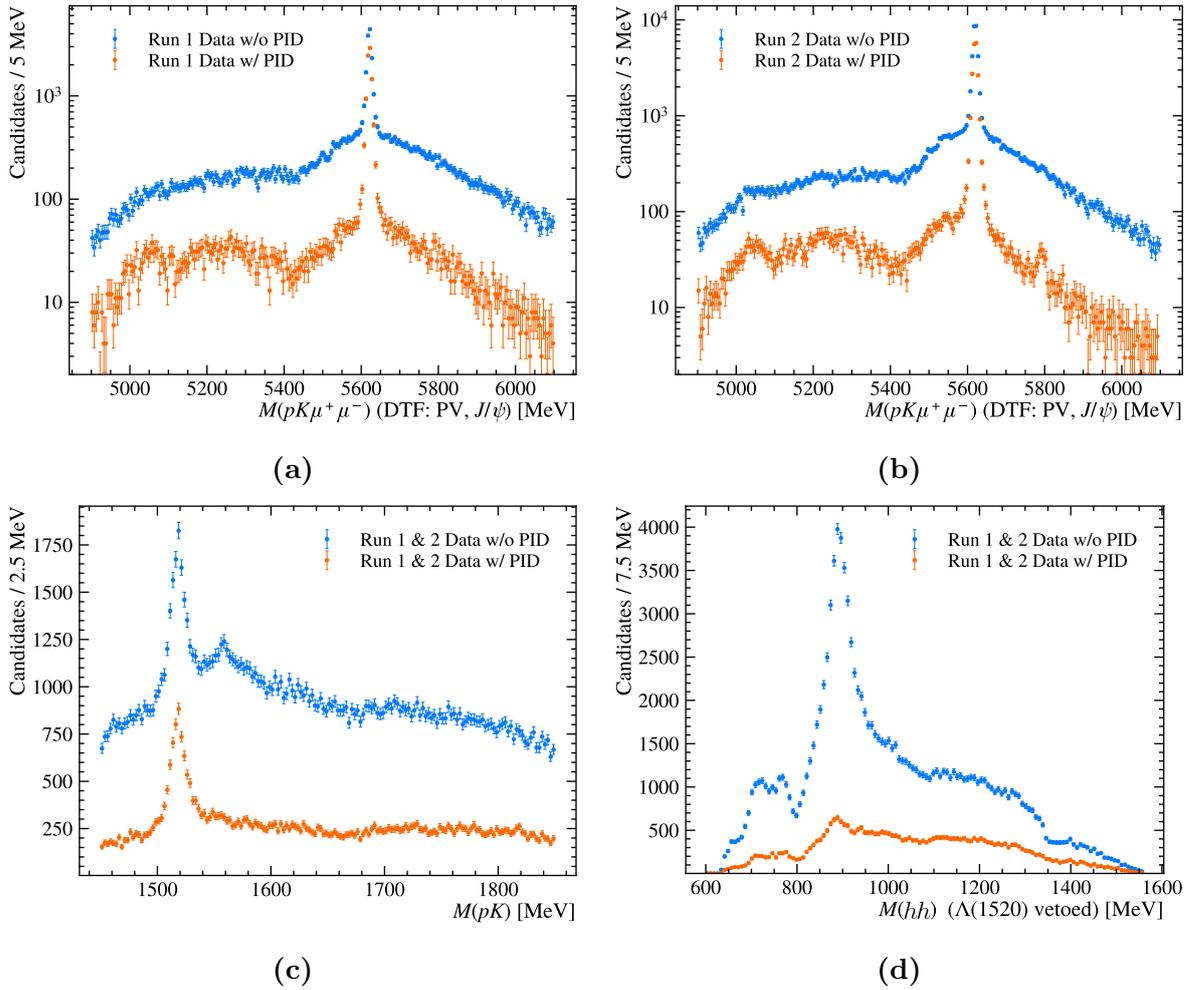


Figure 4.35: The effect of applying PID selection to the control mode displaying the removal of mis-identification backgrounds in the $M(\Lambda_b^0)$ distribution for (a) Run 1 and (b) Run 2. The purification of the M_{pK} distribution is shown in Figure (c) along with the removal of K^{*0} ($m_{K^{*0}} \approx 892$ MeV) backgrounds in Figure (d) where the mass hypothesis of each particle has been swapped to form a $K^{*0} \rightarrow K^-\pi^+$ decay and any signal $\Lambda(1520)$ candidates have been vetoed by $M_{pK} \notin (1500, 1540)$ MeV.

the near background-free datasets required to maximise sensitivity.

4.4.2 MVA-PID Optimisation

The MVA discussed in Section 4.3 was built and trained using datasets before introducing the PID selection, but the choice of a p_{MVA} working point needs to be optimised with the analysis dataset in its fully-selected final state. Therefore the p_{MVA} cut values are optimised by first applying the PID selection and evaluating the combined performance of the “MVA + PID” selection. The MVA was also trained with a combination of all of the analysis categories, so the optimisation is performed on a per category basis to get

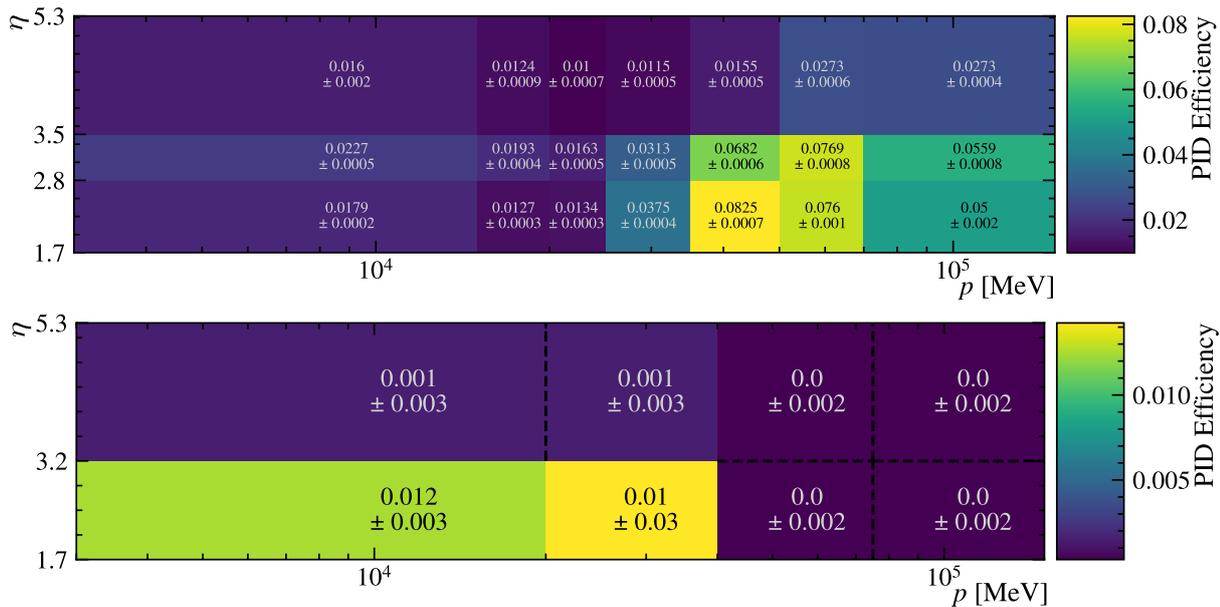


Figure 4.36: PID mis-identification efficiency maps for 2016 MD calibration samples, displaying the full selection PID cut efficiency for mis-identifying a kaon as an electron (upper) and a muon as an electron (lower).

the best p_{MVA} cut value for each sub-dataset of the analysis.

The two factors that dictate the MVA and PID performance are the signal efficiency and background rejection, evaluated using the corrected signal MC and the background μeOS and μeSS datasets. These are competing figure-of-merits (FoMs), the p_{MVA} cut that maximises background rejection would be expected to have a poor signal efficiency and vice-versa. Therefore a FoM that combines these two metrics is used [184],

$$FoM(p_{MVA}) = \frac{\varepsilon_{MVA}\varepsilon_{PID}}{\frac{a}{2} + \sqrt{N_{BG}}}, \quad (4.4.1)$$

where $\varepsilon_{MVA}\varepsilon_{PID}$ refers to the combined signal efficiency of the MVA and PID selection at the given working point and N_{BG} is the number of combinatorial background events in the signal region ($5200 < M(\Lambda_b^0) < 5800$ MeV) [184]. This FoM is chosen over $S/\sqrt{S+B}$ due to the lack of dependence on an estimated branching fraction. For the denominator offset $a/2$ in Equation 4.4.1, $a = 3$ is chosen since this analysis does not expect to observe a signal ($a = 5$), and targets evidence as the first step towards discovering LFV.

The N_{BG} measure in Equation 4.4.1 refers to the number of events in the signal region of the μeOS signal mode, which is blinded. Converging a fit with a large blinded window can lead to fit instabilities (as discussed in Section 5.4.2), hence the μeSS mode is utilised as a combinatorial background proxy, where a fit can be made to the full mass range and a yield in the signal region can be extracted. The method for the background

fit follows that discussed in Section 5.4.2, a third-order Chebyshev polynomial.

Up to now, the μe SS mode has been used to compare shapes and relative levels with the μe OS mode, but to translate a yield from one to the other the absolute statistics of each sample must be considered. Before any PID cuts, this would require a scale factor but, as shown in Figure 4.34, after the PID cuts are applied the upper sidebands become consistent with 21 and 24 events in μe OS and μe SS respectively. This enables the assumption and interpolation of the μe SS yield into the signal region for the FoM fits.

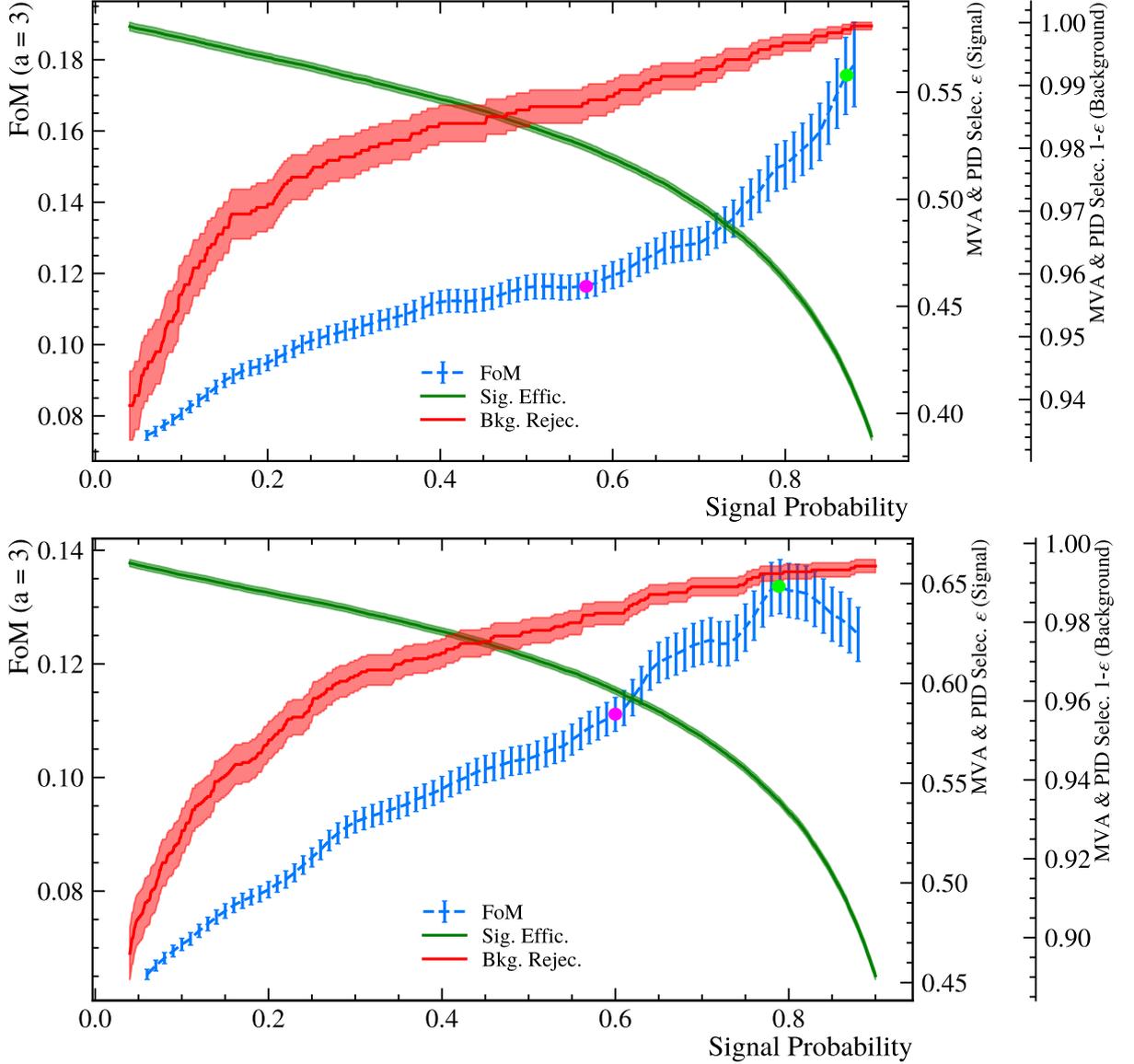


Figure 4.37: FoM scans of MVA output for the Run 2 analysis categories of (upper) 0γ and (lower) 1γ . The potential working points discussed are marked, with ϵ_{MVA}^{90} in pink and Max-FoM in green. The FoM score is calculated by Equation 4.4.1, with a rolling average of five neighbouring values at each p_{MVA} displayed to smooth the curve.

The FoM scans for Run 2 0γ and 1γ categories are shown in Figure 4.37 with the

Table 4.9: Summary of MVA working points for the four analysis categories. For the control mode dataset, the corresponding 1γ values are used as this closer corresponds to a fully-reconstructed candidate.

Working Point	p_{MVA} Optimal Cut Value			
	Run 1		Run 2	
	0γ	1γ	0γ	1γ
Max-FoM	0.76	0.6	0.87	0.79
$\varepsilon_{\text{MVA}} = 90\%$	0.35	0.39	0.57	0.6
$\varepsilon_{\text{MVA}} = 97\%$	0.1	0.12	0.2	0.22
$\varepsilon_{\text{MVA}} = 99.5\%$	0.02	0.03	0.04	0.04

$\varepsilon_{\text{MVA}}\varepsilon_{\text{PID}}$ and background rejection curves overlaid. The working points marked in the plots correspond to:

- “Max-FoM”, the p_{MVA} working point that maximises the FoM.
- $\varepsilon_{\text{MVA}}^{90}$, the point where the efficiency of the MVA crosses 90%, analytically independent of the N_{BG} .

with the $\varepsilon_{\text{MVA}} = 90\%$ working point the primary option for the final selection of the analysis. This is due to the unstable FoM values around the Max-FoM working points where a potential slight change in MVA response on signal data compared to MC would result in a large shift in FoM. Another advantage of the $\varepsilon_{\text{MVA}} = 90\%$ is that it can be analytically and consistently derived across the categories and is not subject to fluctuations in the low statistics $\mu e\text{SS}$ data sample. Calculating the FoM curve allows a cross check that the $\varepsilon_{\text{MVA}} = 90\%$ working points correspond to high FoM values that are not falling or rising rapidly.

The final values of the working points per category are summarised in Table 4.9 where the “medium” $\varepsilon_{\text{MVA}} = 97\%$ and loose $\varepsilon_{\text{MVA}} = 99.5\%$ values are stated for use in Section 5.4.2 as relaxed working points for invariant mass fitting validation.

4.5 Single Candidate Selection

After the entire selection chain is applied, a search for candidates which share the same event and run numbers are performed. Known as “multiple candidates”, these could be formed due to the same event producing two distinct real A_b^0 candidates or from two candidates with shared tracks producing two different head particles. These require removal

Table 4.10: Fraction of events that contain multiple-candidates after all other analysis selection applied. Split by analysis category as well as sample type. The uncertainties are statistical from a binomial efficiency calculation. In Run 2 1γ μe OS sideband data, one event had multiple candidates.

Dataset	Frac. of Events w/ Multi-Cand. (%)			
	Run 1		Run 2	
	0γ	1γ	0γ	1γ
Signal Simulation	0.16 ± 0.03	0.49 ± 0.05	0.18 ± 0.03	0.59 ± 0.05
μe OS Sideband Data	0	0	0	6 ± 6
Control Mode Sim.	0.12 ± 0.02		0.21 ± 0.01	
Control Mode Data	1.5 ± 0.1		0.34 ± 0.04	

before continuing the analysis chain as the efficiency and yield values are measured with respect to the number of events, not candidates. Single candidate selection is performed after all other selection is applied, to maximise the number of events in the analysis by preventing earlier removal of potential high quality candidates.

The removal of multiple candidates is performed randomly for both MC and data, ensuring that only one candidate per event is retained. The random approach is considered best, as any choice based on quality of the event could introduce a selection bias which relies on the variable itself, with extensive studies performed within the LHCb collaboration determining this [185]. The fraction of events with multiple candidates is summarised in Table 4.10 for the different signal/control mode data and MC modes. The results imply that multiple candidates in simulation (where all candidates are truth-matched) are from several high-quality A_b^0 candidates per event, hence the fraction increases as reconstruction improves with Run 1 versus Run 2 and 0γ versus 1γ . Conversely in control mode data, the results imply that the multiple candidates being removed are more likely to be background with a larger fraction in Run 1 versus Run 2.

4.6 Determination of Selection Efficiency

To determine a corrected yield (the overall number of signal/control mode candidates produced in the pp collisions studied) for Equation 3.2.1, the efficiency of reconstruction and selection requirements on each mode must be determined. These are calculated by applying the full selection chain to the MC samples with the correction weights applied. The overall efficiency is the ratio of the sum of weights across all events at generation and

after full selection, given as

$$\begin{aligned}\varepsilon_{\text{tot}} &= \varepsilon_{\text{gen}} \cdot \frac{\sum_{\text{rec+sel}} w_{\tau} \cdot (w_{\text{penta}}) \cdot w_{\text{kin}} \cdot w_{\text{PID}} \cdot w_{\text{trk}} \cdot w_{\text{trig}}}{\sum_{\text{gen}} w_{\tau} \cdot (w_{\text{penta}}) \cdot w_{\text{kin}}} \\ &= \varepsilon_{\text{gen}} \cdot \varepsilon_{\text{rec+sel}}\end{aligned}\tag{4.6.1}$$

where ε_{gen} is the efficiency of any generator-level cuts applied to the simulation and all other components are defined in Section 4.2.

The generator-level cut efficiency refers to loose acceptance and quality cuts that are applied directly to generated candidates, before any simulation of the LHCb detector. The main component of this is an acceptance cut of $10 < \theta < 400$ mrad on the pseudo-stable children of the head particle, but some simulation samples in the analysis have further kinematic or visible-mass (invariant mass ignoring any ν in the decay) cuts applied to increase the practical efficiency of generating MC, and this must also be accounted for. This can also be taken a step further by filtering MC to only include candidates that pass the chosen stripping line, producing a very concise MC sample that in reality originated from a huge number of simulated pp collisions which were discarded. This is all accounted for by breaking ε_{gen} into three components

$$\varepsilon_{\text{gen}} = \varepsilon_{\text{vis}} \cdot \varepsilon_{\text{acc|vis}} \cdot \varepsilon_{\text{filt|acc}}\tag{4.6.2}$$

where ε_{vis} is the generator efficiency of a visible-mass cut, ε_{acc} of the mentioned acceptance cut and $\varepsilon_{\text{filt}}$ the stripping line filtering efficiency all calculated with respect to the previous. For samples which do not have filtering or visible-mass cuts, these components are simply set to 1. The efficiencies are extracted by different methods, with examples shown in Table 4.11. The acceptance efficiency is centrally calculated when the MC is produced, while the visible-mass and filtering efficiencies are calculated using smaller samples of private generator-level simulation due to the manipulation of cuts required. The sample sizes are chosen to ensure a relative error of $\lesssim 1\%$ on the efficiency components.

The total efficiency to reconstruct and select candidates, $\varepsilon_{\text{rec+sel}}$, is determined from the simulated samples by the sum over the weights of the reconstructed and selected candidates divided by the sum of the weights of the generator-level candidates. This efficiency is evaluated in each analysis category.

4.6.1 Efficiency Combinations across an Experimental Run

Across the data taking years of LHCb during Run 1 and 2, the LHC operated at different CoM energies and for different lengths of time, with LHCb also retaining an indepen-

Table 4.11: Breakdown of the generator efficiencies (as %) for a selection of simulation samples for 2016 MD. Full and partially reconstructed samples are shown to display the visible mass cut used as well as a filtered simulation sample and the additional efficiency applied to that. The last mode in the table uses filtering performed for the $B_{(s)}^0 \rightarrow X\mu^\pm e^\mp$ analysis [112] with the excited charm resonance decaying $D^{*-}(2460) \rightarrow \bar{D}^0(\rightarrow K^+e^-\bar{\nu}_e)\pi^-$.

Decay Mode	ε_{vis} (%)	ε_{acc} (%)	$\varepsilon_{\text{filt}}$ (%)	ε_{gen} (%)
$\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$	100	17.80 ± 0.05	100	17.80 ± 0.05
$\Lambda_b^0 \rightarrow pK^- J/\psi$	100	17.43 ± 0.05	100	17.43 ± 0.05
$\Lambda_b^0 \rightarrow D^0(\rightarrow K^-e^+\nu_e)p\mu^-\bar{\nu}_\mu$	18.9 ± 0.2	17.55 ± 0.05	100	3.32 ± 0.04
$\Lambda_b^0 \rightarrow D^0(\rightarrow K^-\mu^+\nu_\mu)pe^-\bar{\nu}_e$	19.1 ± 0.2	17.36 ± 0.04	100	3.31 ± 0.04
$B^0 \rightarrow D^{*-}(2460)\mu^+\nu_\mu$	1.97 ± 0.03	17.0 ± 0.2	35.46 ± 0.02	0.119 ± 0.002

dent operational efficiency. This results in different total luminosities recorded in each year as well as differing values of $f_{\Lambda_b^0}$, the fragmentation fraction of hadronising $b \rightarrow \Lambda_b^0$ which is correlated to collision energy, both requiring consideration when combining efficiency measurements; years with a greater Λ_b^0 yield must be given greater weight in an efficiency combination. Therefore a scheme of ‘‘luminosity weights’’ are derived using the latest LHCb determined $f_{\Lambda_b^0}$ values [186] as well as the standard LHCb luminosity for each year [187], yielding weights of 0.07, 0.18, 0.24, 0.22, 0.29 for 2011, 12, 16, 17, 18 respectively.

4.6.2 Total Efficiency of Signal and Control Modes

The ensemble of efficiencies for the signal and control mode MC is summarised in Tables 4.12 and 4.13. The efficiencies of important exclusive background modes corresponding to $\Lambda_b^0 \rightarrow \Lambda_c^{*+}(\rightarrow D^0(\rightarrow K^-\ell^+\nu)p)\ell^-\bar{\nu}$ decays are shown in Tables 4.14 and 4.15, while the leptonic mis-ID $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ background is given in Table 4.16. The reconstruction and selection efficiencies in the tables are broken down into components by the following equation

$$\varepsilon_{\text{rec+sel}} = \varepsilon_{\text{rec}} \cdot \varepsilon_{\text{trig|rec}} \cdot \varepsilon_{\text{preselec|trig}} \cdot \varepsilon_{\text{MVA|preselec}} \cdot \varepsilon_{\text{PID|MVA}} \cdot \varepsilon_{\text{multi|PID}}, \quad (4.6.3)$$

allowing study and comparison of individual analysis components.

Some slight differences in efficiency can be seen for Run 1 and Run 2. For the signal and control mode the reconstruction efficiency is greater for Run 1 than Run 2 due to the different proton p_T alignment cuts from PIDCALIB while the trigger selection, especially

Table 4.12: Weighted efficiencies for $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ at different selection stages for the different analysis categories. For the efficiency steps, each is conditional based on the stage above. The incremental stages (right-aligned labels) give finer description of the main stages in bold. The incremental steps before “Fiducial” are not weighted. The generator-level and reconstruction efficiency cannot be split by Bremsstrahlung category.

Efficiency (%)	Analysis Categories			
Categories	Run 1 1γ	Run 1 0γ	Run 2 1γ	Run 2 0γ
Generator	16.78 ± 0.02	16.78 ± 0.02	17.83 ± 0.02	17.83 ± 0.02
Reconstruction	4.891 ± 0.008	4.891 ± 0.008	3.822 ± 0.005	3.822 ± 0.005
Stripping + Rec.	11.32 ± 0.01	11.6 ± 0.01	10.653 ± 0.009	11.502 ± 0.009
Truth Matching	49.62 ± 0.06	46.54 ± 0.06	49.66 ± 0.04	46.5 ± 0.04
Fiducial	44.47 ± 0.08	44.33 ± 0.08	35.74 ± 0.06	36.11 ± 0.06
Trigger	40.4 ± 0.1	42.8 ± 0.1	46.03 ± 0.08	46.43 ± 0.09
L0	53.6 ± 0.1	54.4 ± 0.1	51.5 ± 0.1	51.9 ± 0.1
HLT1	91.17 ± 0.09	90.4 ± 0.1	95.81 ± 0.06	96.06 ± 0.05
HLT2	82.9 ± 0.1	87.1 ± 0.1	95.47 ± 0.06	95.23 ± 0.06
Pre-Selection	37.8 ± 0.1	34.7 ± 0.1	38.2 ± 0.1	35.6 ± 0.1
Mass Selec.	85.0 ± 0.1	85.1 ± 0.1	86.1 ± 0.1	86.2 ± 0.1
Resonances Vetoes	78.8 ± 0.2	80.5 ± 0.2	79.2 ± 0.1	81.2 ± 0.1
Semilep. Vetoes	55.7 ± 0.2	49.8 ± 0.2	55.4 ± 0.2	50.3 ± 0.2
MVA Selection	86.3 ± 0.2	85.8 ± 0.2	87.1 ± 0.1	86.8 ± 0.1
PID Selection	71.3 ± 0.2	61.0 ± 0.3	67.2 ± 0.2	59.2 ± 0.2
Single Candidate	99.49 ± 0.04	99.83 ± 0.02	99.4 ± 0.03	99.83 ± 0.02
Reco.&Selec.	0.457 ± 0.003	0.38 ± 0.003	0.389 ± 0.002	0.322 ± 0.002
Total ($\times 10^{-2}$)	7.68 ± 0.05	6.38 ± 0.05	6.93 ± 0.04	5.74 ± 0.03

the HLT1 and HLT2 lines, perform better in Run 2 due to the improved trigger system and more complex lines discussed in Section 4.1.4. The difference in performance for the MVA between Run 1 and 2 can also be seen in these tables (see Section 4.3.3).

If the Bremsstrahlung categories are compared, a significant difference in efficiency is seen for the semileptonic and PID selection with the 1γ category more efficient in each case. This justifies the splitting of the analysis by Bremsstrahlung emission, as the recovered electron energy is preventing some signal candidates from being erroneously removed by the semileptonic vetoes as well as PID system returning a more electron-like response for candidates with recovered Bremsstrahlung, with the presence of this lost energy inferring the identity of the electron.

Table 4.13: Weighted efficiencies for $\Lambda_b^0 \rightarrow pK^- J/\psi(\rightarrow \mu^+\mu^-)$ at different selection stages for the different categories in the analysis. For the efficiency steps, each is conditional based on the stage above. The incremental stages (right-aligned labels) give finer description of the stages in bold and are not weighted for simplicity.

Efficiency (%)	Analysis Categories	
Categories	Run 1	Run 2
Generator	16.45 ± 0.02	17.47 ± 0.02
Reconstruction	6.06 ± 0.01	4.401 ± 0.005
Stripping + Rec.	30.34 ± 0.02	28.78 ± 0.01
Truth Matching	47.52 ± 0.05	47.7 ± 0.02
Fiducial	41.91 ± 0.07	32.0 ± 0.03
Trigger	72.08 ± 0.08	77.08 ± 0.04
L0	84.03 ± 0.08	82.88 ± 0.05
HLT1	95.39 ± 0.05	97.63 ± 0.02
HLT2	90.0 ± 0.07	96.26 ± 0.03
Pre-Selection	32.5 ± 0.1	31.71 ± 0.05
Mass Selec.	71.6 ± 0.1	72.05 ± 0.06
Resonances Vetoes	93.37 ± 0.08	93.52 ± 0.04
Semilep. Vetoes	48.6 ± 0.2	47.05 ± 0.09
MVA Selection	81.1 ± 0.2	83.6 ± 0.1
PID Selection	72.1 ± 0.3	73.4 ± 0.1
Single Candidate	99.86 ± 0.03	99.8 ± 0.01
Reco.&Selec.	0.829 ± 0.005	0.658 ± 0.002
Total	0.1363 ± 0.0008	0.1149 ± 0.0004

Comparison between the signal efficiency and control mode efficiency shows the expected trend. The trigger selection is significantly more efficient on the control mode due to two muons in the decay able to pass the L0 requirement of a single high quality muon. The stripping and reconstruction efficiency is also dramatically higher in the control mode than the signal, due to the looser requirements applied to muons due to the high quality track they produce by leaving a signature in the muon chambers.

The overall efficiencies to be used in Equation 3.2.1 are displayed on the last row of Tables 4.12 and 4.13, each showing their corresponding statistical uncertainty. All have a relative uncertainty of $< 1\%$, implying that other systematic uncertainties will dominate the error on these efficiency values. These are discussed in detail in Section 4.7.

Table 4.14: Weighted efficiencies for the double-semileptonic $\mu^\pm e^\mp$ background sample $\Lambda_b^0 \rightarrow D^0(\rightarrow K^- \mu^+ \nu_\mu) p e^- \bar{\nu}_e$ different selection stages for the different data taking years of the analysis. For the efficiency steps, each is conditional based on the stage above. The single candidate selection is ignored due to the low statistics.

Efficiency (%) Categories	Analysis Categories			
	Run 1 1 γ	Run 1 0 γ	Run 2 1 γ	Run 2 0 γ
Generator	3.13 ± 0.03	3.13 ± 0.03	3.29 ± 0.02	3.29 ± 0.02
Reconstruction	1.556 ± 0.009	1.556 ± 0.009	1.37 ± 0.006	1.37 ± 0.006
Trigger	19.1 ± 0.1	19.67 ± 0.07	26.0 ± 0.1	26.96 ± 0.06
Pre-Selection	0.39 ± 0.07	0.26 ± 0.03	0.35 ± 0.04	0.27 ± 0.02
MVA Selection	40.0 ± 7.0	56.0 ± 5.0	46.0 ± 5.0	55.0 ± 3.0
PID Selection	70.0 ± 10.0	47.0 ± 7.0	61.0 ± 7.0	58.0 ± 4.0
Reco.&Selec. ($\times 10^{-4}$)	3.1 ± 0.8	2.2 ± 0.3	3.3 ± 0.6	3.0 ± 0.3
Total ($\times 10^{-4}$)	0.10 ± 0.02	0.07 ± 0.01	0.11 ± 0.02	0.100 ± 0.008

Table 4.15: Weighted efficiencies for the double-semileptonic $\mu^\pm e^\mp$ background sample $\Lambda_b^0 \rightarrow D^0(\rightarrow K^- e^+ \nu_e) p \mu^- \bar{\nu}_\mu$ different selection stages for the different data taking years of the analysis. For the efficiency steps, each is conditional based on the stage above. The single candidate selection is ignored due to the low statistics.

Efficiency (%) Categories	Analysis Categories			
	Run 1 1 γ	Run 1 0 γ	Run 2 1 γ	Run 2 0 γ
Generator	3.14 ± 0.03	3.14 ± 0.03	3.3 ± 0.02	3.3 ± 0.02
Reconstruction	2.15 ± 0.01	2.15 ± 0.01	1.769 ± 0.006	1.769 ± 0.006
Trigger	30.0 ± 0.2	30.04 ± 0.08	40.8 ± 0.1	40.15 ± 0.06
Pre-Selection	0.35 ± 0.05	0.19 ± 0.02	0.28 ± 0.03	0.19 ± 0.02
MVA Selection	40.0 ± 6.0	36.0 ± 3.0	41.0 ± 4.0	48.0 ± 2.0
PID Selection	70.0 ± 8.0	52.0 ± 6.0	60.0 ± 6.0	56.0 ± 4.0
Reco.&Selec. ($\times 10^{-4}$)	6 ± 1	2.2 ± 0.4	5.0 ± 0.8	3.5 ± 0.3
Total ($\times 10^{-4}$)	0.20 ± 0.04	0.07 ± 0.01	0.16 ± 0.03	0.12 ± 0.01

4.6.3 Total Efficiency of Background Modes

Tables 4.14 to 4.16 show efficiency breakdowns for two double-semileptonic and one leptonic mis-ID background mode for the analysis. The overall selection efficiencies of these can be seen to be orders of magnitude lower than the signal mode, highlighting how performant the selection chain is at removing background and retaining signal. Specific

Table 4.16: Weighted efficiencies for the $\mu^\pm e^\mp$ leptonic mis-identification background sample $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ different selection stages for the different data taking years of the analysis. For the efficiency steps, each is conditional based on the stage above. The single candidate selection is ignored due to the low statistics.

Efficiency (%)	Analysis Categories			
Categories	Run 1 1 γ	Run 1 0 γ	Run 2 1 γ	Run 2 0 γ
Generator	16.85 ± 0.02	16.85 ± 0.02	17.92 ± 0.03	17.92 ± 0.03
Reconstruction	0.268 ± 0.004	0.268 ± 0.004	0.14 ± 0.002	0.14 ± 0.002
Trigger	51.5 ± 0.2	56.7 ± 0.04	62.5 ± 0.1	62.16 ± 0.02
Pre-Selection	13.3 ± 0.3	16.78 ± 0.05	25.8 ± 0.2	28.81 ± 0.04
MVA Selection	77.0 ± 1.0	85.4 ± 0.2	79.3 ± 0.6	89.45 ± 0.07
PID Selection	0.2 ± 0.1	0 ± 0.1	0.6 ± 0.1	0.61 ± 0.02
Reco.&Selec. ($\times 10^{-4}$)	0.2 ± 0.1	0 ± 0.01	0.80 ± 0.09	1.05 ± 0.02
Total ($\times 10^{-4}$)	0.03 ± 0.02	0 ± 0.002	0.14 ± 0.02	0.188 ± 0.004

examples of how different selection components are targeting different background types can be seen by comparing the pre-selection efficiency, which contains the semileptonic vetoes, between the signal and $\Lambda_b^0 \rightarrow D^0 p \ell^- \bar{\nu}$ mode or the PID selection efficiency for the signal and $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ modes. The semileptonic vetoes are clearly removing the semileptonic backgrounds more than the signal, as designed, while the PID selection removes the majority of the remaining FCNC background due to the low-rate $\ell \rightarrow \ell'$ mis-ID.

The total efficiencies of these modes however highlight the potential issues from evaluating exclusive background rejection in this way, with the small number of events remaining in the MC samples after full selection leading to efficiencies with $> 10\%$ relative uncertainty. A brute-force solution to this is to generate more MC samples as well as being more efficient with their generation using filtering and generator-level cuts as discussed, but this has already been done for the $\Lambda_b^0 \rightarrow D^0 p \ell^- \bar{\nu}$ modes (see Table 4.11). With increased sample size reducing statistical uncertainty by only \sqrt{N} , a balance must be found as done here, where the uncertainty is low enough that fluctuations would not dramatically alter the prediction while MC sample sizes do not lead to unfeasible computing requirements.

4.7 Determination of Systematic Uncertainty on Efficiency Ratio

The calculation of the signal and control mode efficiency has attempted to emulate the effect of analysis selection on real-data using MC with extensive data-driven corrections. While the MC has been shown to describe data effectively, residual differences are expected that could bias the calculation. Common differences affecting the efficiency should cancel in the ratio $\varepsilon_{Control}/\varepsilon_{Signal}$ in Equation 3.2.1, but this must be verified as a series of systematic choices have been made that could have mis-aligned the signal and control mode, further compounding the intrinsic differences between $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ and $\Lambda_b^0 \rightarrow pK^- J/\psi$ simulation.

The basis for estimating these systematic uncertainties is to re-perform the calculation of the efficiency ratio with a change implemented to a given aspect of the selection or correction chain. Such changes must be reasonable, i.e. one that represents a plausible alternative choice, which does not degrade the performance of the analysis.

The uncertainty introduced from the limited size of the MC samples has been calculated in line with the efficiency derivation (see Tables 4.12 and 4.13), with the hypothetical “alternative choice” of simulating a big enough sample for this to be neglected, this combines to a relative uncertainty on the efficiency ratios of 1%(0.9%) and 0.7%(0.6%) for Run 1 and 2, 0γ (1γ) respectively.

For the efficiency-map-based reweighters, the statistical uncertainty from the sample used to derive each bin (examples in Figures 4.17, 4.18 and 4.19), will be propagated to determine the effect on the efficiency ratio of the limited statistics in each bin of phase-space with the chosen binning scheme. This will not however evaluate the effect of any variation in efficiency across each bin, with the potential for large changes to be averaged out to the single value used. To evaluate the magnitude of this effect, an alternative finer binning scheme will be derived in each case and the average difference in the efficiency ratio determined due to the change in the event-by-event weights (folding in the statistical uncertainty on each bin). For the kinematic reweighter, the model will be retrained with alternative hyper-parameters and the difference in efficiency ratio determined similarly. From similar analyses each component is expected to yield 0.1 – 1% relative uncertainty.

The largest systematic effect is expected to originate from the M_{pK} selection and modelling in the MC samples. The signal and control mode use distinct methods, with a tight-cut and a simple approach requiring a resonant $\Lambda(1520) \rightarrow pK^-$ for the signal mode. The alternative choice, discussed in Section 3.2.1.2, would have been to simulate and reweight the complete pK^- -spectrum as in the control mode, aiding cancellation of potential

bias from the $pK J/\psi$ reweighter (see Section 4.2.2). This approach was adopted for the R_{pK} analysis, using one-dimensional w_{penta} [76]. To determine any residual systematic from the nominal choice with respect to the alternative, 2016 $\Lambda_b^0 \rightarrow pK^- \mu^\pm e^\mp$ simulation is generated without any $\Lambda(1520)$ bias and the absolute systematic uncertainty is derived as

$$\sigma_r = \left(\frac{\varepsilon_{Control}^{RW}}{\varepsilon_{Gen}^{Loose \rightarrow Tight}} \bigg/ \frac{\varepsilon_{Signal}^{RW}}{\mathcal{F}(\Lambda(1520))_{\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-}} \right) - \frac{\varepsilon_{Control}}{\varepsilon_{Signal}}, \quad (4.7.1)$$

where RW refers to efficiencies where the full pK -spectrum and the modelling reweighter were used, and ε' denotes that the control mode M_{pK} selection is tightened to the signal mode interval for this calculation. The generator efficiency of tightening the M_{pK} cut and $\mathcal{F}(\Lambda(1520))$, the measured $\Lambda(1520)$ fraction in $\Lambda_b^0 \rightarrow pK^- \mu^+ \mu^-$ decays [105], normalises the newly constructed ratio such that a comparison to the nominal ratio is valid. The calculation isolates the difference in efficiency from the modelling only, with the expected reduction from tight M_{pK} selection normalised in the first denominator. Due to availability of simulation samples, only a Run 2 value can be derived, producing a relative uncertainty of 8.1% that is applied to both Run 1 and 2. This is expected to be the dominant uncertainty from the efficiency calculation, and hence has been calculated first to verify that it is not restricting the overall analysis sensitivity. Further systematics are derived for the control mode yield in Section 5.3.3.3.

Chapter 5

Evaluation of Backgrounds and Yield Determination

To measure the signal and control mode yields in Equation 3.2.1, a fit model for the invariant mass distribution in data must be built in each analysis category. Each fit requires a signal and background component. The signal component is determined with the corresponding MC samples by fitting to a single Probability Density Function (PDF) and fixing the value of most if not all parameters, allowing some to shift or scale to take into account differences in resolution between the actual detector and the model in simulation. The background component can contain multiple sources. With the extensive tight selection described in Section 4.1 the dominant background is expected to be the combinatorial contribution from coincidences of random tracks, but the possibility of partially or misidentified exclusive backgrounds from other physics decays remain and their contribution must be evaluated.

As discussed in Chapter 3, the combinatorial component in the $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ signal mode will be controlled using the μe SS combinatorial proxy and checked on the blinded μe OS sidebands (Section 5.4.2). A background study is performed to determine which, if any, exclusive backgrounds have a significant predicted yield remaining, with invariant mass fits to the remaining distributions appended to the overall background fit and yields constrained from the study (see Section 5.1). The background components are then combined to perform final blinded fits in Section 5.4.3 before including the MC signal component to determine a blinded upper limit for N_{Signal} in Section 5.5 and hence determine the expected analysis sensitivity to $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp)$.

For the $\Lambda_b^0 \rightarrow pK^- J/\psi$ control mode mass fits, in Section 5.3, any exclusive background components are determined in a similar fashion but, without the blinded region, the combinatorial background yield and shape can be evaluated with the $M(pK^- \mu^+ \mu^-)$

distribution without introducing external control. The final control mode fit, including the signal-component partially-fixed from MC, is then performed in two categories of Run 1 and Run 2 to determine values of $N_{Control}$ for the different analysis categories, as the final input for Equation 3.2.1.

5.1 Potential Backgrounds for $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$

A range of exclusive backgrounds were considered for the $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ decay. Table 5.1 presents a summary of the dominant and most important modes with a full list included in Tables C.1 to C.4 split by the different types of background. To estimate the contribution from each individual background, MC samples were collected and processed through the full analysis reweighting and selection chain, recording the weighted selection efficiency as presented in Tables 4.14 to 4.16. By combining the full reconstruction and selection efficiency of each background mode, ε_{BG} , with an externally measured exclusive branching fraction, \mathcal{B}_{BG} , an estimate of the expected yield for each background mode in $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ data can be calculated as

$$N_{BG} = \varepsilon_{BG} \cdot \mathcal{B}_{BG} \cdot N_{\Lambda_b^0} = \varepsilon_{BG} \cdot \mathcal{B}_{BG} \cdot \frac{N_{Control}}{\mathcal{B}_{Control} \cdot \varepsilon_{Control}}, \quad (5.1.1)$$

where $N_{Control}$, $\mathcal{B}_{Control}$ and $\varepsilon_{Control}$ are the control mode yield, branching fraction and selection efficiency for the given analysis category, providing a calculation of the overall number of Λ_b^0 particles produced in pp collisions ($N_{\Lambda_b^0}$) for that category. This provides a baseline from which a calculation of the relative background contribution for the $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ mode can be made. The control mode is used as it is sufficiently high in statistics and already utilised in the analysis framework, with the potential for correlated systematic uncertainties cancelling in the ratio $\varepsilon_{BG}/\varepsilon_{Control}$. For backgrounds considered where the head particle is not a Λ_b^0 , the equation must be adjusted to

$$N_{BG} = \varepsilon_{BG} \cdot \mathcal{B}_{BG}^{corr} \cdot N_{\Lambda_b^0}, \text{ where } \mathcal{B}_{BG}^{corr} = B_{BG}^{PDG} \cdot \frac{f_{BG}}{f_{\Lambda_b^0}}, \quad (5.1.2)$$

to correct for the different fragmentation fractions of Λ_b^0 and the B mesons, where f_{BG} is the latest LHCb measurement for the head particle of the mode considered [186, 187]. The estimated yield is determined in each analysis category by including the splitting fraction of 0γ and 1γ in the value of ε_{BG} .

The estimated yield of each background is evaluated at three stages of selection: immediately before applying the MVA, after applying the MVA and after the full selection (including the PID and single candidate selection) by using different values of ε_{BG}

Table 5.1: Background yield estimates for the key background samples in the signal mode $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$. The predicted yields after pre-selection, MVA selection and PID selection are given separately. If no events remained after selection, $0 \pm \sigma$ is displayed with a single candidate propagation of the uncertainty. † highlights modes that use an estimated branching fraction to determine the yields, discussed in Section 5.1.1.

Background Sample Estimated Yield ($\mu^\pm e^\mp$)	Stage	Analysis Categories			
		Run 1 1 γ	Run 1 0 γ	Run 2 1 γ	Run 2 0 γ
$\Lambda_b^0 \rightarrow D^0(\rightarrow K^- e^+ \nu_e) p \mu^- \bar{\nu}_\mu$ †	PS:	29.0 ± 6.0	25.0 ± 5.0	61.0 ± 9.0	63.0 ± 9.0
	MVA:	12.0 ± 3.0	9.0 ± 2.0	25.0 ± 4.0	30.0 ± 5.0
	PID:	8.0 ± 2.0	4.5 ± 0.9	14.0 ± 3.0	17.0 ± 2.0
$\Lambda_b^0 \rightarrow D^0(\rightarrow K^- \mu^+ \nu_\mu) p e^- \bar{\nu}_e$ †	PS:	18.0 ± 4.0	13.0 ± 2.0	45.0 ± 7.0	38.0 ± 5.0
	MVA:	7.0 ± 2.0	8.0 ± 2.0	19.0 ± 4.0	22.0 ± 3.0
	PID:	5.0 ± 1.0	3.5 ± 0.7	12.0 ± 2.0	12.0 ± 2.0
$\Lambda_b^0 \rightarrow \Lambda_c(2860)^+ \mu^- \bar{\nu}_\mu$ † $\Lambda_c(2860)^+ \rightarrow D^0(\rightarrow K^- e^+ \nu_e) p$	PS:	2.5 ± 0.6	1.6 ± 0.4	5.0 ± 1.0	6.0 ± 1.0
	MVA:	1.2 ± 0.4	0.8 ± 0.2	1.8 ± 0.5	3.0 ± 0.6
	PID:	0.6 ± 0.2	0.6 ± 0.1	1.1 ± 0.3	2.1 ± 0.4
$\Lambda_b^0 \rightarrow \Lambda_c(2860)^+ e^- \bar{\nu}_e$ † $\Lambda_c(2860)^+ \rightarrow D^0(\rightarrow K^- \mu^+ \nu_\mu) p$	PS:	1.1 ± 0.3	1.2 ± 0.3	1.5 ± 0.5	5.1 ± 0.9
	MVA:	0.5 ± 0.2	0.1 ± 0.1	0.4 ± 0.3	1.5 ± 0.5
	PID:	0.2 ± 0.1	0.08 ± 0.04	0.2 ± 0.1	1.1 ± 0.2
$\Lambda_b^0 \rightarrow p K^- K^+ K^-$	PS:			0.0 ± 0.2	1.8 ± 0.5
	MVA:				1.7 ± 0.5
	PID:			0.0 ± 0.3	0.0 ± 0.2
$\Lambda_b^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) p K^-$	PS:	1.2 ± 0.2	7.0 ± 1.0	2.9 ± 0.4	6.5 ± 0.8
	MVA:	0.8 ± 0.1	5.8 ± 0.9	1.3 ± 0.2	2.9 ± 0.3
	PID:	0.001 ± 0.003	0.16 ± 0.04	0.013 ± 0.004	0.1 ± 0.02
$\Lambda_b^0 \rightarrow \Lambda_c(2940)^+ \mu^- \bar{\nu}_\mu$ † $\Lambda_c(2940)^+ \rightarrow D^0(\rightarrow K^- e^+ \nu_e) p$	PS:	0.17 ± 0.07	0.1 ± 0.04	1.2 ± 0.3	0.23 ± 0.07
	MVA:	0.08 ± 0.04	0.004 ± 0.003	0.27 ± 0.09	0.003 ± 0.003
	PID:	0.03 ± 0.02	0.0 ± 0.002	0.18 ± 0.06	0.0 ± 0.003
$\Lambda_b^0 \rightarrow D^0(\rightarrow K^- \mu^+ \nu_\mu) p \pi^-$	PS:	0.07 ± 0.03	2.6 ± 0.3		
	MVA:	0.001 ± 0.006	1.4 ± 0.2		
	PID:	0.0 ± 0.002	0.07 ± 0.02		
$\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ (PID × 10 ⁻⁴)	PS:	0.011 ± 0.002	0.3 ± 0.05	0.026 ± 0.003	0.8 ± 0.1
	MVA:	0.009 ± 0.001	0.26 ± 0.04	0.021 ± 0.003	0.74 ± 0.09
	PID:	0.12 ± 0.08	4 ± 1	1.0 ± 0.2	35 ± 4
$\Lambda_b^0 \rightarrow p K^- K^+ \pi^-$	PS:			0.0 ± 0.04	0.16 ± 0.09

measured at each selection stage. This can be seen in Table 5.1 with the stated error from the statistical uncertainty on measurements of $N_{Control}$ and ε as well as the combined systematic and statistical error for the external \mathcal{B} and f factors. The decay modes shown were identified as potential contaminants from either their large branching fraction or ability to mimic the signal kinematics. Appendix C includes the full results from the

~ 40 modes studied, with those not shown in Table 5.1 found to be consistent with zero.

Analysing Table 5.1, the PID selection, applied through mis-ID weights, are shown to undoubtedly remove the backgrounds requiring $\ell \rightarrow \ell'$ mis-ID as well as $h \rightarrow \ell$ mis-ID, with only one fully-hadronic background, $\Lambda_b^0 \rightarrow pK^-K^+K^-$, within 5σ of contributing a single event per category with predicted yields of $\approx 0 \pm 0.3$. The background is not included however in the final background fit as the yield is already small before PID selection, so any remaining background will be undoubtedly suppressed. The $\Lambda_b^0 \rightarrow pK^-K^+K^-$ background was also not seen in other $\Lambda_b^0 \rightarrow pK^-\ell^+\ell^-$ analyses, which used looser PID selection than this analysis [76, 105].

Considering the partially reconstructed backgrounds in Table 5.1, the single-semileptonic backgrounds have also been entirely removed by the selection chain, due to the requirement of an additional $h \rightarrow \ell'$ mis-ID to emulate the $pK\mu e$ final state. The only significant backgrounds remaining are from the double-semileptonic decays $\Lambda_b^0 \rightarrow D^0(\rightarrow K^-\ell^+\nu)p\ell^-\bar{\nu}$, with the resonant versions, $\Lambda_b^0 \rightarrow \Lambda_c^{*+}(\rightarrow D^0(\rightarrow K^-\ell^+\nu)p)\ell^-\bar{\nu}$ shown to be heavily suppressed in comparison by the $\Lambda_c^{(*)+}$ veto discussed in Section 4.1.6. The $\Lambda_b^0 \rightarrow D^0(\rightarrow K^-\ell^+\nu)p\ell^-\bar{\nu}$ background modes (subsequently referred to as $D^0p\ell$) require special consideration however, with an estimated branching fraction currently in use, see Section 5.1.1.

The background estimates at different selection stages further validate that the analysis selection suite is removing the backgrounds they are designed to. For example, the relatively small change in the estimated yield before and after PID selection for the $D^0p\ell$ modes (which produce a $pK\mu e$ visible final state) compared to the large change in $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ show the effectiveness of the PID cuts against $\ell \rightarrow \ell'$ mis-ID. In addition, the highest rate background processes that were specifically vetoed against have been removed including J/ψ or $\psi(2S)$ resonances and $\Lambda_b^0 \rightarrow \Lambda_c^+(\rightarrow pK^-\ell^+\nu)\ell^-\bar{\nu}$ decays. This overall performance supports that the analysis selection chain is working effectively, and suggests that the predicted yields are a sufficient estimate (with the exception of $D^0p\ell$ discussed in the following section) and hence that it can be assumed that the only background components that require fitting to are the combinatorial and $D^0p\ell$ shapes.

5.1.1 Estimated Yield for $\Lambda_b^0 \rightarrow D^0p\ell^-\bar{\nu}$ Backgrounds

The stated estimated yields in Table 5.1 for $D^0p\ell$ come with the large caveat that $\mathcal{B}(\Lambda_b^0 \rightarrow D^0p\ell^-\bar{\nu})$ has not been experimentally measured. While it has been seen as a contaminant background in other LHCb analyses [76, 186, 188] the energy lost by the two final state neutrinos result in it being absorbed into the smooth combinatorial shape for the relatively looser selection imposed on precision analyses of $b \rightarrow s\ell^-\ell^+$ decays. For this

analysis, it was expected that $D^0 p \ell$ would require factorisation considering the ultra-low combinatorial level remaining after full-selection (as shown in Figure 4.34), therefore it was incorporated into the background study. Simulation samples were produced, including approximate-modelling of the $D^0 p$ distribution in Figure 4.12a, using an estimation for

$$\mathcal{B}(\Lambda_b^0 \rightarrow D^0 p \ell^- \bar{\nu}) = \mathcal{B}(\Lambda_b^0 \rightarrow \Lambda_c^+ \ell^- \bar{\nu}) \cdot \frac{\mathcal{B}(\Lambda_b^0 \rightarrow p D^0 \pi^-)}{\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda_c^+ \pi^-)} = (0.8 \pm 0.2)\%, \quad (5.1.3)$$

naively assuming that the fractional difference in decay rate would propagate to the $\Lambda_b^0 \rightarrow D^0 p X$ system. This however is not necessarily the case, with different quark-structure and corresponding form-factors from the $\Lambda_b^0 \rightarrow \Lambda_c^+ X$ system. The estimated branching fraction returns a total $N_{BG} = 78 \pm 18$ integrated over analysis categories and the two flavour combinations (quoted per category in Table 5.1), which visually does not agree with the conclusion from Figure 4.34b where only a minor excess of 12 events is observed in the μeOS lower sideband over the μeSS proxy, validated by the robust consistency in the upper sideband. Therefore, it is reasonable to assume that the branching fraction estimate is incorrect by an unknown margin, with the errors underestimated and likely unable to cover the difference. Therefore the decision was made to allow this branching fraction to float or be fixed to zero in the subsequent signal background fits, freely quantifying the background in each analysis category or evaluating if the strategy of previous analyses of folding the contribution into the combinatorial is feasible.

5.2 Exclusive fits to backgrounds for $\Lambda_b^0 \rightarrow \Lambda(1520) \mu^\pm e^\mp$

The remaining backgrounds are added to overall the background fit by producing a template shape by an invariant mass fit in $M(pK^- \mu^\pm e^\mp)$ of the fully selected MC samples that were used to determine ε_{BG} previously.

5.2.1 $\Lambda_b^0 \rightarrow pK^- K^+ K^-$

The fully hadronic background $\Lambda_b^0 \rightarrow pK^- K^+ K^-$ was given special consideration in this analysis despite the predicted yield consistent with zero. This was due to a similar background, $B^+ \rightarrow K^+ K^+ K^-$, found to contaminate in the signal region in the R_X analysis [71], requiring significant re-evaluation of results with respect to the previous R_K measurement [71]. With no missing energy in this decay and simply a double $h \rightarrow \ell$ mis-ID, any background remaining from the $\Lambda_b^0 \rightarrow pK^- K^+ K^-$ would be contained within the signal region and produce an erroneous peak around the Λ_b^0 mass, as can be seen in

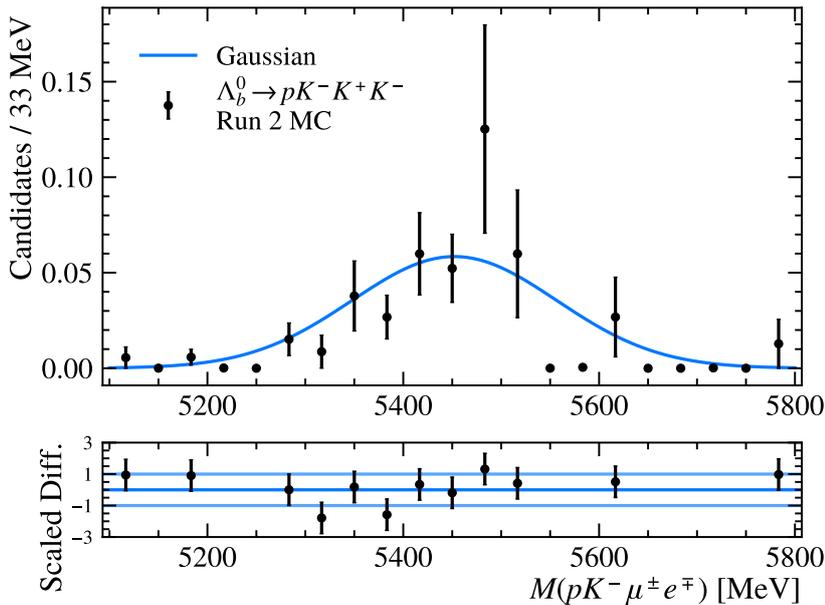


Figure 5.1: Invariant mass fit to corrected $\Lambda_b^0 \rightarrow pK^-K^+K^-$ simulation for Run 1 and 2 of LHCb after full application of the full selection chain. A Gaussian fit attempts to parameterise the shape.

the reweighted MC distribution in Figure 5.1. The figure highlights the limited statistics currently available in the simulation sample after full selection, insufficient for a reliable template fit, however due to the ultra-low predicted yield in Table 5.1 it is not included in the background ensemble. The $\sim 0.01^2$ $h \rightarrow \ell$ mis-ID suppression factor that is applied to any events passing the MVA supports this conclusion given the low statistics sample available.

5.2.2 $\Lambda_b^0 \rightarrow D^0(\rightarrow K^- \ell^+ \nu) p \ell^- \bar{\nu}$

The double-semileptonic $D^0 p \ell$ backgrounds are kinematically distinct to $\Lambda_b^0 \rightarrow pK^-K^+K^-$, with large energy losses from the two neutrinos biasing the reconstructed mass below the signal region and into the lower mass sideband as seen in Figure 5.2. While the two lepton combinations ($e^+\mu^-$ and μ^+e^-) differ, with each lepton originating from the D^0 or Λ_b^0 decay, observed consistency of the distributions motivates the combination of them when fitting, especially as the predicted contribution from each mode are similar and the combined yield is more precise. Furthermore, only minor differences were observed between simulation samples of different run periods, with the only clear distinction when splitting by Bremsstrahlung category, motivating the Run 1+2 $0\gamma/1\gamma$ category fits in Figure 5.2.

The smeared shape in Figure 5.2 is difficult to analytically fit to, and while a number of options were attempted including a Gaussian Kernel Density Estimator (GKDE) [189,

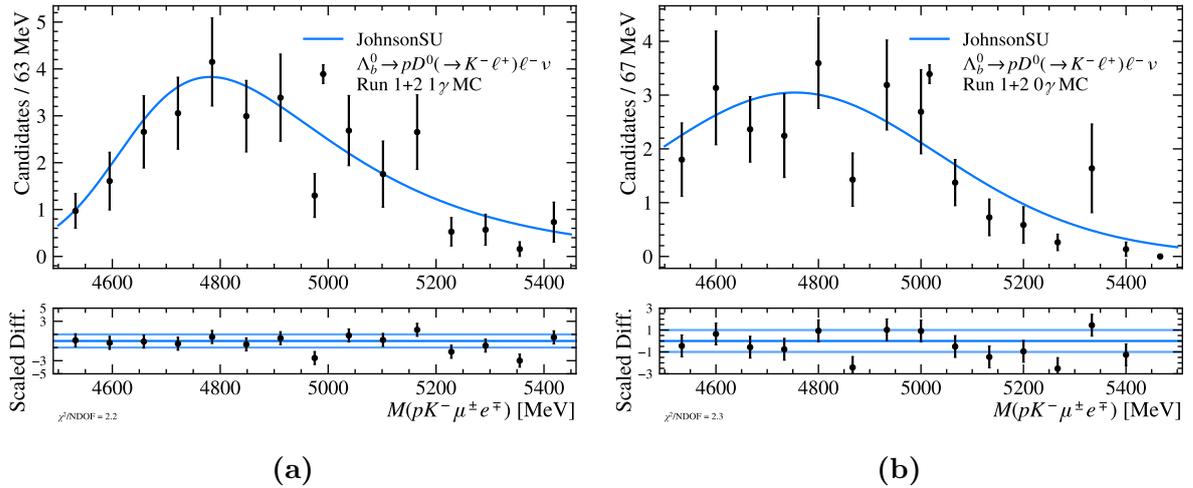


Figure 5.2: Invariant mass fit to corrected $\Lambda_b^0 \rightarrow D^0(\rightarrow K^- \ell^+ \nu) p \ell^- \bar{\nu}$ simulation for Run 1 and 2 of LHCb in categories of (left) with Bremsstrahlung reconstructed and (right) without Bremsstrahlung reconstructed. The difference in shape from the Bremsstrahlung categories was observed to be more significant than the difference between LHC runs therefore Run 1 and 2 samples were combined to improve fit stability from the very low statistics fits.

[190] and an Argus distribution [191], a JohnsonSU [192] was chosen as it best described the weighted MC samples. Two JohnsonSU shapes were determined independently for the $0\gamma/1\gamma$ Bremsstrahlung categories, prepared for the blinded data fits in Section 5.4.3, both supporting the feasibility of floating the $D^0 p \ell$ yield, with the core of the shape within the lower sideband.

5.3 Mass Fits of the Control Mode

To extract $N_{Control}$, an invariant mass fit to the control mode in $M(pK^- \mu^+ \mu^-)$ is performed. The $\Lambda_b^0 \rightarrow pK^- J/\psi(\rightarrow \mu^+ \mu^-)$ mode has the same full selection chain as the signal mode, with the muon-only trigger selection and choice of discriminating variables for the MVA chosen to not include any lepton-flavour specific options. The only differences in selection correspond to the removal of the J/ψ veto (and instead an inverted J/ψ selection), the wider M_{pK} range and the lack of HOP-FD cut. The p_{MVA} cut value for the control mode is taken from the corresponding 1γ signal mode analysis category due to the greater similarity in reconstruction of an electron with Bremsstrahlung recovered to a muon (shown in Sections 5.3.1 and 5.4.1).

The major difference between invariant mass fits in the control mode compared to the signal mode is the use of the DecayTreeFitter (DTF) tool [193] to create a constrained

$M(pK^-\mu^+\mu^-)$ variable in the control mode. The DTF algorithm is widely used in LHCb in addition to the default invariant mass calculation with the additional ability to provide one or many initial hypotheses for the topology of the decay-of-interest. The tool, developed initially for the BaBar experiment [193], takes the measured kinematics and topology of a given candidate and performs a global fit producing new kinematic and invariant mass variables and a χ_{DTF}^2 , a measure of how consistent the given candidate is to the decay hypothesis provided. This χ_{DTF}^2 measure, without any further constraints, is used as prior cut before the MVA discussed in Section 4.3.2.1.

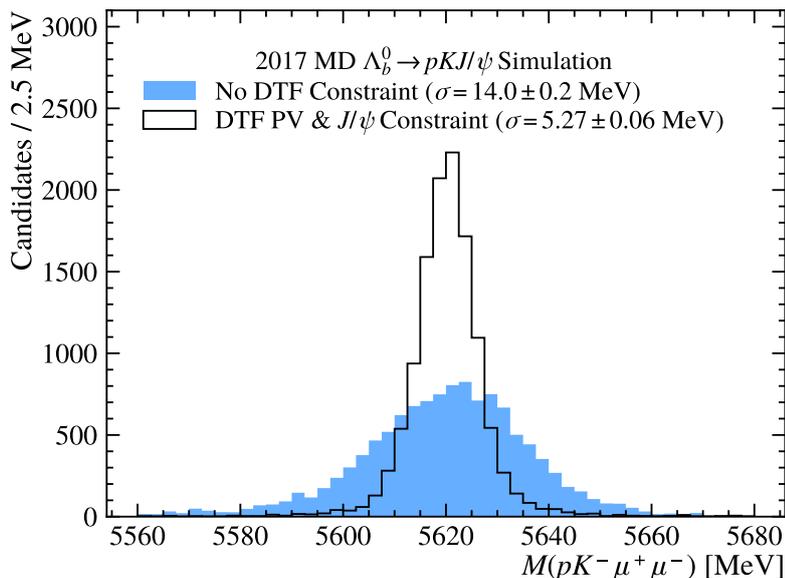


Figure 5.3: $M(pK^-\mu^+\mu^-)$ for 2017 corrected $\Lambda_b^0 \rightarrow pK^-J/\psi$ simulation, showing the improvement in resolution from applying DTF constraints. The σ shown are from a Double-Sided Crystal Ball (DCB) fit to each distribution, detailed in Section 5.3.1.

For the control mode fits, DTF was implemented adding constraints requiring that the di-muon pair originated from a J/ψ resonance and that the Λ_b^0 originates from the PV. Figure 5.3 shows the drastic improvement these two constraints have on the mass-resolution, therefore $M_{\Lambda_b^0}^{DTFJ/\psi,PV}$ becomes the nominal variable for control mode invariant mass fits. In a negligible fraction ($\approx 0\%$) of cases the DTF global fit can fail, returning a non-zero status, these candidates are therefore also removed under the “Single Candidate Selection”.

A series of options and variations for the control mode fit were considered and are discussed below, with the nominal method detailed in Section 5.3.3 where signal-shape parameters are taken from MC and the background is a combination of a floating Chebyshev 2nd order polynomial as the combinatorial background with two exclusive background components from b -meson decays.

5.3.1 MC Signal Component

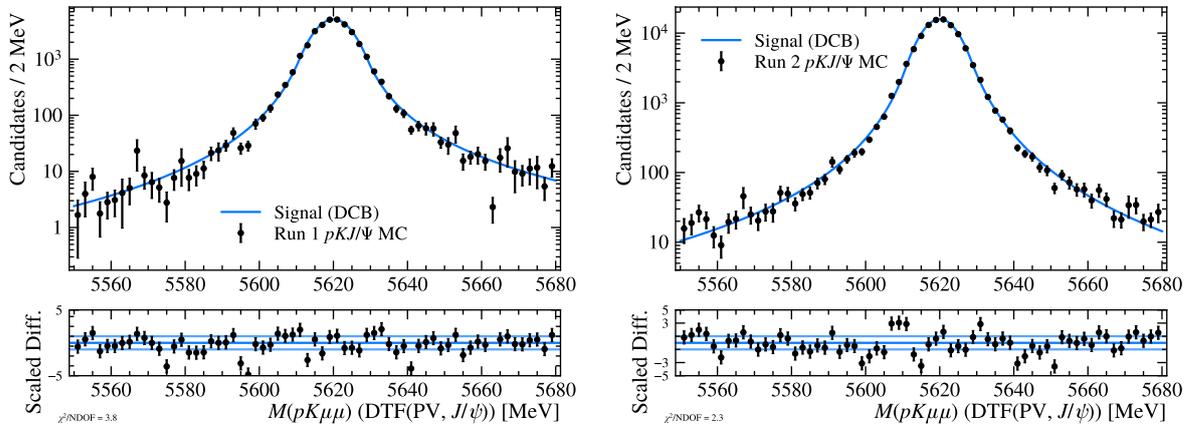


Figure 5.4: Invariant mass fits to the $M_{\Lambda_b^0}^{DTF J/\psi, PV}$ distribution of corrected $\Lambda_b^0 \rightarrow pK^- J/\psi$ control mode simulation for Run 1 (left) and Run 2 (right) of the LHC. The fit is performed with a DCB shape with all parameters floated.

The signal shape for the control mode is described by a DCB shape, which takes into account slight under or over reconstruction of the candidates momentum via the exponential tails on either side of the peak. This is performed in the two control mode categories of Run 1 and Run 2 with the $M_{\Lambda_b^0}^{DTF J/\psi, PV}$ variable, within a mass range tuned to $> 10\sigma$ from $m_{\Lambda_b^0}$ while removing the very low-statistics tail region observed to bias the core shape.

Figure 5.4 shows the result of these fits, with the pull between the sample points and fitted curve included and the final fit parameters summarised in Table 5.5. The fit describes the simulation sample effectively, with all pulls $< 5\sigma$ within the central region of the fit, with a few outliers from the statistically limited region. These define the signal shape for the subsequent data fits by fixing all tail parameters and allowing the σ (μ) parameter to scale (shift) via a new parameter, f_σ ($\Delta\mu$). This accounts for the expected minor change in resolution between simulation and data, which the $\Lambda_b^0 \rightarrow pK^- J/\psi$ is expected to be sensitive to due to the large decay rate.

5.3.2 Fits to Data without Exclusive Background Components

Figures 4.35a and 4.35b shows the status of the control mode data after full selection, plotted on a log scale to emphasise the surrounding distribution from the clear Λ_b^0 peak. There is clearly the presence of a peaking background at ≈ 5800 MeV in the Run 2 distribution, corresponding to the decay $\Xi_b^0 \rightarrow pK^- J/\psi$ as well as some non-combinatorial

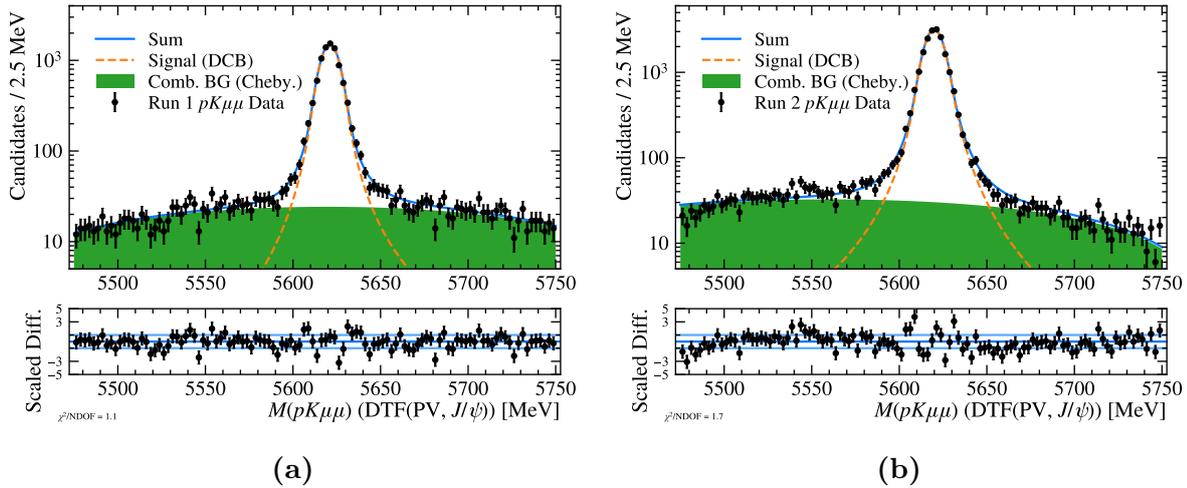


Figure 5.5: Invariant mass fits to the $M_{\Lambda_b^0}^{DTFJ/\psi, PV}$ distribution of $\Lambda_b^0 \rightarrow pK^- J/\psi$ control mode data for Run 1 (a) and Run 2 (b). The signal DCB shape is partially-fixed from the fits in Figure 5.4 and the combinatorial a second-order Chebyshev polynomial.

contributions below 5550 MeV in both categories, from other $B \rightarrow h^+ h^- J/\psi$ decays. However, by fitting in the mass range $5475 < M(\Lambda_b^0) < 5750$ MeV these contributions can be neglected in the invariant mass fit. This approach was taken in the $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ analysis (which shares control mode with this analysis) by fitting to a DCB as the signal and a second order Chebyshev polynomial as the combinatorial background, which is shown to describe the data effectively [105]. This implementation can be verified by conducting a background study similar to in Section 5.1, using a series of $B \rightarrow h(h)J/\psi$ modes, with results displayed in Table 5.2. The study shows that while there are highly-significant contributions from background when integrating across the whole analysis mass regime ($4500 < M(\Lambda_b^0) < 6750$ MeV), if the yields are recalculated in the fit range, the combinatorial-component dominates. This supports describing the background with a single combinatorial shape, absorbing any remaining exclusive backgrounds. In Section 5.3.3 a scheme with individual separate background components is detailed.

From the combinatorial shape either side of the peak in Figures 4.35a and 4.35b, the background distribution clearly cannot be described by an exponential, with the discussed sculpting in the lower mass regions due to the semileptonic selection in Section 4.1.6. The distribution that best describes this was found to be a Chebyshev polynomial, with two coefficients X_1 and X_2 floating in the fit. Higher-order polynomials were tested, producing near-identical results and higher coefficients consistent with zero.

The combinatorial component, with a yield $N_{Comb.}$, is combined with the partially-fixed signal component and fit to the $M_{\Lambda_b^0}^{DTFJ/\psi, PV}$ distributions in the two categories via a maximum-likelihood fit to extract the control mode yields $N_{Control}$, with results shown

Table 5.2: Background yield estimates for the key background samples in the control mode $\Lambda_b^0 \rightarrow pK^- J/\psi(\rightarrow \mu^+ \mu^-)$, split by LHC Run. The bottom-line numbers after full selection are shown, labelled “PID”, with the cascading estimates at different selection stages included in Table C.4. If there were no statistics remaining after selection $0 \pm \sigma$ is displayed from a single candidate propagation of the uncertainty. A second value labelled “Cut” describes the predicted yield of backgrounds in the range $5475 < M(\Lambda_b^0) < 5750$ MeV.

Background Datasets	Stage	Analysis Categories	
		Run 1	Run 2
$B^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) K^{*0}$	PID:	710.0 ± 80.0	1060.0 ± 90.0
	Cut:	350.0 ± 40.0	760.0 ± 60.0
$B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) K^+ K^-$	PID:	800.0 ± 100.0	760.0 ± 70.0
	Cut:	590.0 ± 70.0	670.0 ± 60.0
$\Lambda_b^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) pK^-$ (<i>pK</i> Swap)	PID:	300.0 ± 40.0	43.0 ± 5.0
	Cut:	210.0 ± 30.0	33.0 ± 4.0
$B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) K^{*0}$	PID:		18.0 ± 2.0
	Cut:		10.0 ± 1.0
$B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) \pi^+ \pi^-$	PID:	1.4 ± 0.5	3.0 ± 1.0
$B^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) \pi^+ \pi^-$	PID:	1.3 ± 0.4	2.1 ± 0.6
$B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) K^+$	PID:	0.0 ± 1.0	0.0 ± 2.0
$\Lambda_b^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) pK^- \pi^0$	PID:	0.31 ± 0.06	

in Table 5.3 and fit shapes overlaid in Figure 5.5. The fit describes the data effectively, with no large pulls or significant pattern in the distribution that could suggest a missing background.

5.3.2.1 Fit Validation - Stability

To validate the stability of the fit scheme chosen, 10000 pseudo-experiment datasets are produced from the fit result, each allowing the yield to fluctuate via a Poisson distribution. Each pseudo-experiment is then re-fit with the same model and floating parameters as in the initial fit. The distributions of how each parameter deviated from its generation parameters, $(x_{gen} - x_{fit})/\sigma(x)$, is subsequently fit to a Gaussian function. All parameters

Table 5.3: Final parameter values for maximum-likelihood fits to control mode data presented in Figures 5.5 and 5.8. All parameters were floated, except for the $f_{K^{*0}}$ which is externally Gaussian constrained at 0.039 ± 0.008 (Run 1) and 0.033 ± 0.007 (Run 2) from the background study in Table 5.2. The overall background yield is $N_{K^{*0}J/\psi} = f_{K^{*0}}N_{Control}$. The second exclusive background component refers to the $B_s^0 \rightarrow K^+K^-J/\psi$ mode, fixed relative to $N_{K^{*0}J/\psi}$ in the fit. $\Delta\mu$ has units MeV.

Fit Scheme/ Category	Control Mode Fit Parameters						
	$N_{Control}$	$N_{Comb.}$	$f_{K^{*0}}$	$\Delta\mu$	f_σ	X_1	X_2
w/ Excl. Bg.							
Run 1	8840 ± 100	1710 ± 140	0.039(8)	-1.04(7)	1.10(1)	0.10(5)	-0.24(6)
Run 2	19640 ± 150	1650 ± 180	0.035(4)	-0.10(5)	1.141(9)	-0.31(5)	-0.22(7)
Comb. Only							
Run 1	8820 ± 100	2231 ± 59	N/A	1.04(7)	1.10(1)	0.08(4)	-0.28(5)
Run 2	19560 ± 150	2800 ± 70	N/A	-0.11(5)	1.136(9)	-0.40(3)	-0.25(4)

for the control mode fit showed symmetrical distributions centred on zero with $\sigma = 1$, implying the overall fit is stable and robust to any fluctuations in the analysis datasets and that the fit uncertainties are correctly determined.

5.3.3 Fits to the Data with Exclusive Background Components

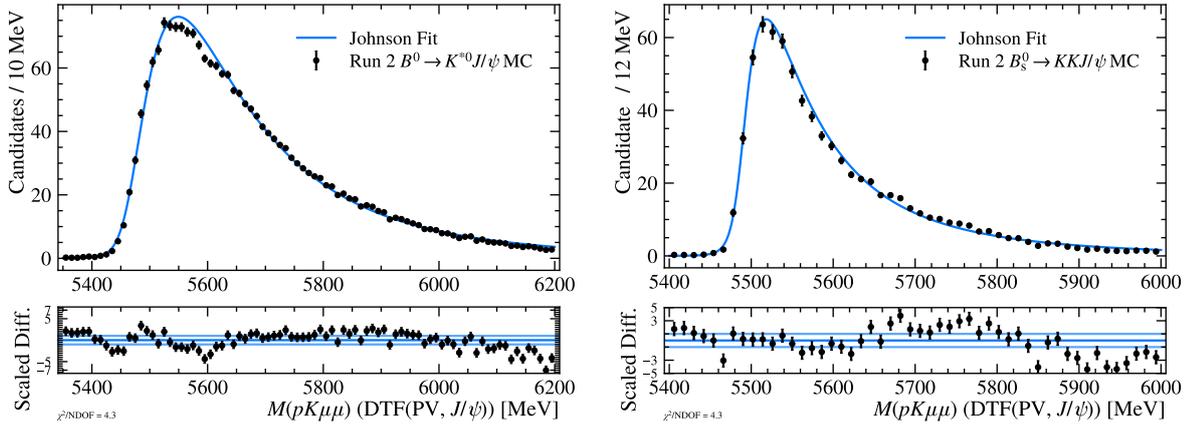


Figure 5.6: Exclusive background template fits in $M_{A_b^0}^{DTF, J/\psi, PV}$ for the two background components included in the nominal control mode fits. (left) The Run 2 template fit for $B^0 \rightarrow K^{*0}J/\psi$ corrected MC including mis-ID weights. (right) The corresponding Run 2 fit for $B_s^0 \rightarrow K^+K^-J/\psi$. Both backgrounds are described by a JohnsonSU [192].

Despite the good description of the control mode $M(A_b^0)$ distribution with the method discussed in Section 5.3.2, the fact the clearly significant exclusive background

yields were ignored in the fit and absorbed into the combinatorial component is not satisfactory. The method required a very specific $M(\Lambda_b^0)$ range and is also unlikely to be robust against changes to PID selection, with the assumption of the shapes being absorbed quickly failing for any looser selection. There is also potential for the signal component to unintentionally include some background contributions possibly over-estimating $N_{Control}$ as a result. Therefore an additional fitting method was implemented, with exclusive backgrounds factorised from the combinatorial component.

5.3.3.1 Modelling and Constraining Control Mode Exclusive Backgrounds

Table 5.2 implies that the primary backgrounds are $B_{(s)}^0 \rightarrow K^{*0} J/\psi$, $B_s^0 \rightarrow K^+ K^- J/\psi$ and a “ pK -swap” component of $\Lambda_b^0 \rightarrow p K^- J/\psi$, with a double mis-ID reflection. The $B_s^0 \rightarrow K^{*0} J/\psi$ can immediately be neglected, with the estimated yield in Run 2 significantly smaller than the uncertainty of $B^0 \rightarrow K^{*0} J/\psi$. The pK -swap component is also negligible for the same reason in Run 2, with the smeared shape from the incorrect mass hypothesis also able to be absorbed by the larger background components (see Figure 4.14 for a looser selection fit including this component). For Run 1, this conclusion is less palatable, but currently retained for consistency between the different category fits.

To parameterise the shape of the remaining $B^0 \rightarrow K^{*0} J/\psi$ and $B_s^0 \rightarrow K^+ K^- J/\psi$ backgrounds a template fit similar to that described in Section 5.2.2 is implemented. The resulting JohnsonSU fits for Run 2 are shown in Figure 5.6. The PDF describes the overall shape of the background effectively but fails to detail the local features of the corrected MC with sharp edges and steps in both distributions. Alternatively a GKDE could have been utilised to meticulously emulate the shape, but there is a possibility that these features are in-fact simulation artefacts, and with the predicted yields discussed significantly smaller than the MC sample sizes, it is unlikely they would be resolved in data. Therefore the smooth JohnsonSU PDF is utilised for the fit, with all parameters fixed to enhance fit convergence. A considered addition to the PDF would be to apply a Gaussian resolution function based on the f_σ from the signal component, but the already present smearing from the hadronic mis-ID is expected to dominate.

The expected method for handling the two exclusive background yields would be to apply a Gaussian constraint corresponding to the predicted N_{BG} and uncertainty in the fit range. However the R_{pK} [76] and $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ angular analyses [151] observed that the $B_s^0 \rightarrow K^+ K^- J/\psi$ background yield was over-estimated by the method in Equation 5.1.1, with the $K^+ K^-$ spectrum incorrectly described in the default available simulation and a full reweighting of this background beyond the scope of these analyses for a simple control mode background. Therefore a data-driven approach was utilised in

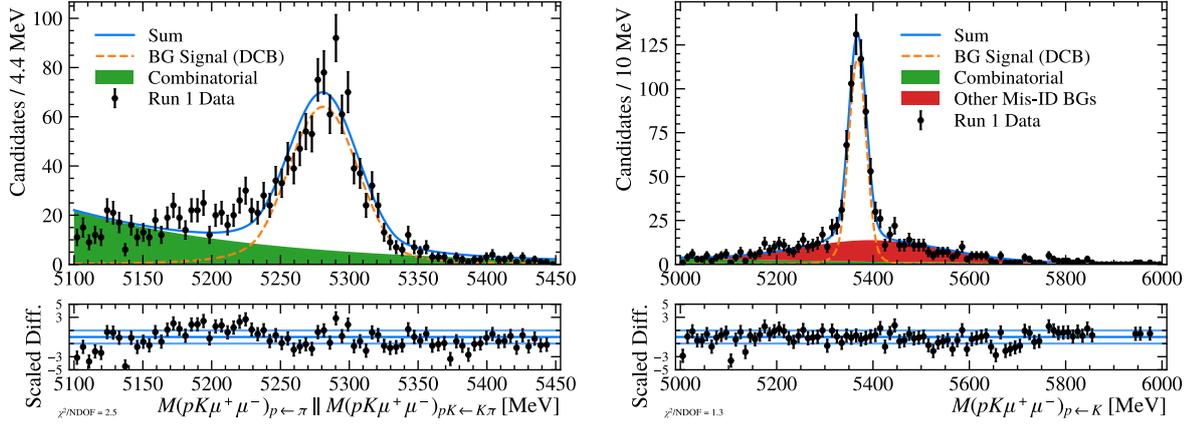


Figure 5.7: Investigation of $\Lambda_b^0 \rightarrow pK^- J/\psi$ control mode data to extract a data-driven relative yield for the backgrounds $B^0 \rightarrow K^{*0} J/\psi$ (left) and $B_s^0 \rightarrow K^+ K^- J/\psi$ (right) to reduce the number of floating parameters in the total $N_{Control}$ fit. Each shows Run 1 data in the control mode “sidebands” with $5580 < M_{\Lambda_b^0} < 5680$ MeV excluded. Fits were performed with a mass hypothesis consistent with the exclusive background. The PDFs are intentionally simplified, using double or single Crystal ball functions for the core and exponential and Gaussian components to describe the other background content.

each analysis to determine $F_{KK/K^{*0}}$, the expected fraction of $B_s^0 \rightarrow K^+ K^- J/\psi$ background events with respect to the $B^0 \rightarrow K^{*0} J/\psi$ mode that is well-modelled and reliable in default simulation. This also binds the $B_s^0 \rightarrow K^+ K^- J/\psi$ component to the other, dominant background reducing the number of free parameters in the fit. Figure 5.7 displays the simple invariant mass fits implemented to extract

$$F_{KK/K^{*0}} = \frac{N'_{KK}}{\varepsilon_{KK}} \bigg/ \frac{N'_{K^{*0}}}{\varepsilon_{K^{*0}}} \quad (5.3.1)$$

and hence,

$$N_{K^+ K^- J/\psi} = F_{KK/K^{*0}} N_{K^{*0} J/\psi}, \quad (5.3.2)$$

where N'_H is the yield from a simple invariant mass fit to an $M(\Lambda_b^0)$ distribution in data where the hadronic mass-hypotheses are altered to that of the background: $p \leftarrow K$ and $p \leftarrow \pi \mid pK \leftarrow K\pi$ for KK and K^{*0} respectively, following the convention introduced in Section 4.1.5.2. The efficiency ε_H encodes the selection applied to manipulate the distributions to be dominated by the target backgrounds, calculated with the corresponding simulation ensuring the ratio correctly describes the background ratio in the final control mode fit range. This selection excluded the region $5580 < M_{\Lambda_b^0}^{DTF J/\psi, PV} < 5680$ to remove the majority of signal Λ_b^0 events as well as a similar veto for Ξ_b . The region ± 75 MeV about $m_{K^{*0}}$ was also vetoed (selected) for the KK (K^{*0}), overall producing the well-defined background peaks in Figure 5.7.

The NLL fits in Figure 5.7 describe the data reasonably well in the core peaking regions of each fit, with otherwise poorly characterised regions in both background modes from combinatorial and other background content. This is not considered to be an issue however, since the yields N'_H propagated through to Equation 5.3.2 are dominated by the uncertainty on $N_{K^*0J/\psi}$ ($\sigma_N/N \approx 10\%$). This validates the fixing of $F_{KK/K^*0} = 0.835 \pm 0.060$ (0.830 ± 0.060) for Run 1 (2) in the total data-fit, coupling the yield of $B_s^0 \rightarrow K^+K^-J/\psi$ directly to $B^0 \rightarrow K^*0J/\psi$ in the current iteration of the analysis. The other analyses discussed, derived fractions of 1.25 ± 0.12 [76] and 0.72 ± 0.04 [151], with the differences likely from the tighter semileptonic selection and unique MVA used in this analysis. An alternative approach would have been to fix $N_{K^*0J/\psi}$ and instead constrain F_{KK/K^*0} to its value and uncertainty. Performing the nominal fit (described in Equation 5.3.3) with this change resulted in $F_{KK/K^*0}^{float} = 0.836 \pm 0.059$ (0.836 ± 0.057) for Run 1 (2) and crucially produced an $N_{Control}$ consistent at $\ll 1\sigma$ with the fixed- F_{KK/K^*0} version of the fit.

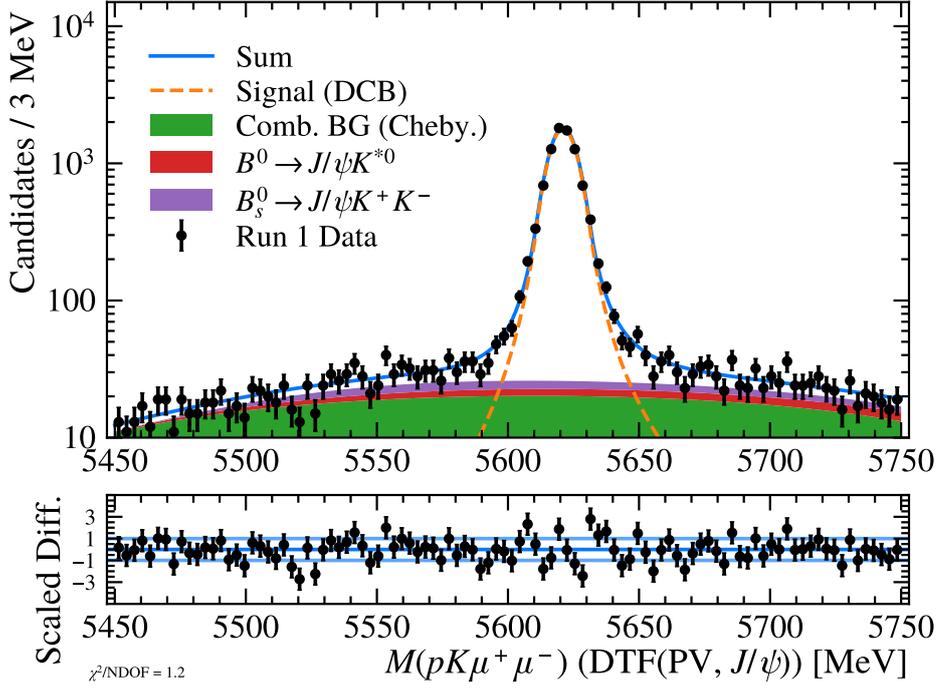
5.3.3.2 Extraction and Validation of Final $N_{Control}$ Yield

The total control mode fit therefore combines to

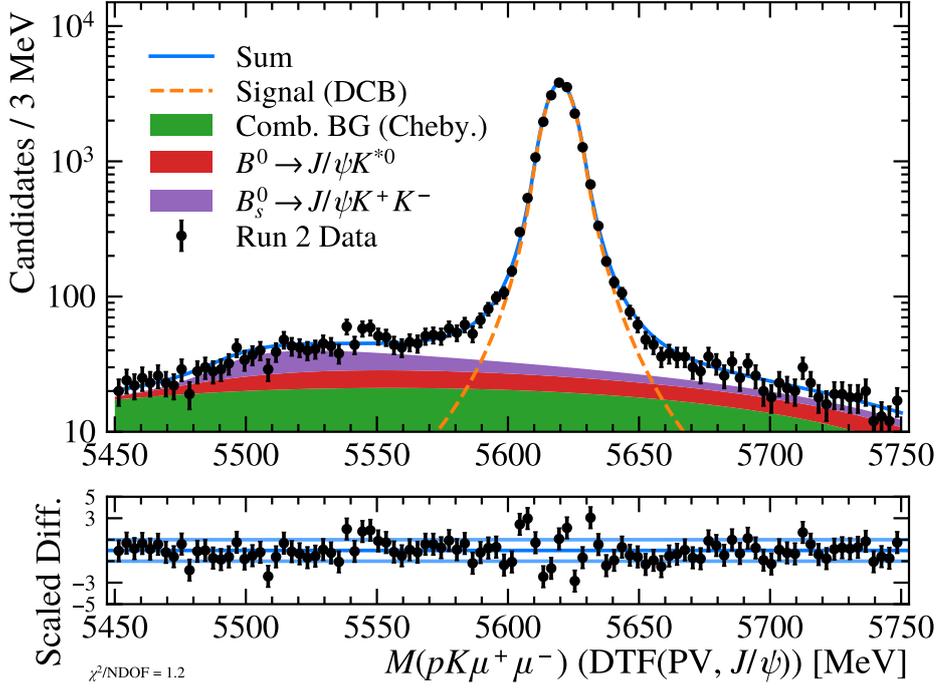
$$P(pK\mu^+\mu^-) = N_{Control}(P_{MC}^{DCB} + f_{K^*0}(P_{K^*0MC}^{John.} + F_{KK/K^*0}P_{KKMC}^{John.})) + N_{Comb}P^{Cheby}, \quad (5.3.3)$$

where P describes the PDFs of the signal and exclusive background shapes as described in Sections 5.3.1 and 5.3.3 as well as the second-order Chebyshev polynomial P^{Cheby} , which continues to describe the combinatorial background effectively, controlled by the mass region above the Λ_b^0 peak. As the estimated background yields in Equation 5.1.1 are measured relative to $N_{Control}$, the actual floating parameter is factorised to f_{K^*0} , which the data-driven background ratio F_{KK/K^*0} inherits as in Equation 5.3.2. The fit range is also loosened to $(5450 < M(\Lambda_b^0) < 5750)$ MeV. An analogous total fit model with the Chebyshev combinatorial swapped for an exponential was used in Figure 4.14 to extract sWeights for developing the correction suite for the analysis.

The final parameter results after a NLL fit are presented in Table 5.3 with the corresponding PDF shapes described in Figure 5.8. The addition of exclusive backgrounds has clearly improved the description of the data for Run 2 with respect to Figure 5.5b, describing the distribution more consistently over a wider fit range, including the $M(\Lambda_b^0) < 5500$ MeV region. A small feature continues to appear in the pulls at ~ 5550 MeV but it is consistent with a fluctuation. The Run 1 fit, compared to Figure 5.5a, is more similar, due to the relatively smaller exclusive background components in Run 1, with N_{Comb} more compatible but a significant f_{K^*0} component remaining, consistent with the prediction



(a)



(b)

Figure 5.8: Invariant mass fits to the $M_{\Lambda_b^0}^{DTF J/\psi, PV}$ distribution with exclusive background components included in the overall model. The figures show $\Lambda_b^0 \rightarrow pK^- J/\psi$ control mode data for Run 1 (a) and Run 2 (b). The signal DCB shape is partially-fixed from the fits in Figure 5.4, exclusive background components are JohnsonSU shapes and the combinatorial a second-order Chebyshev polynomial.

from Table 5.3, confirming that the with exclusive background approach is best for both categories.

Both distributions show hints of a pattern in the pulls around the signal peak. Although there is some ambiguity in the potential cause of this, it is a small effect and shifting/scaling in the μ and σ of the signal-component are expected to accommodate data-MC differences. Future refinements of the analysis could investigate an alternative signal shape or add further exclusive backgrounds to the fit, but the stability and cross-checks discussed in the following sections support the fit method in its current form for the dataset available.

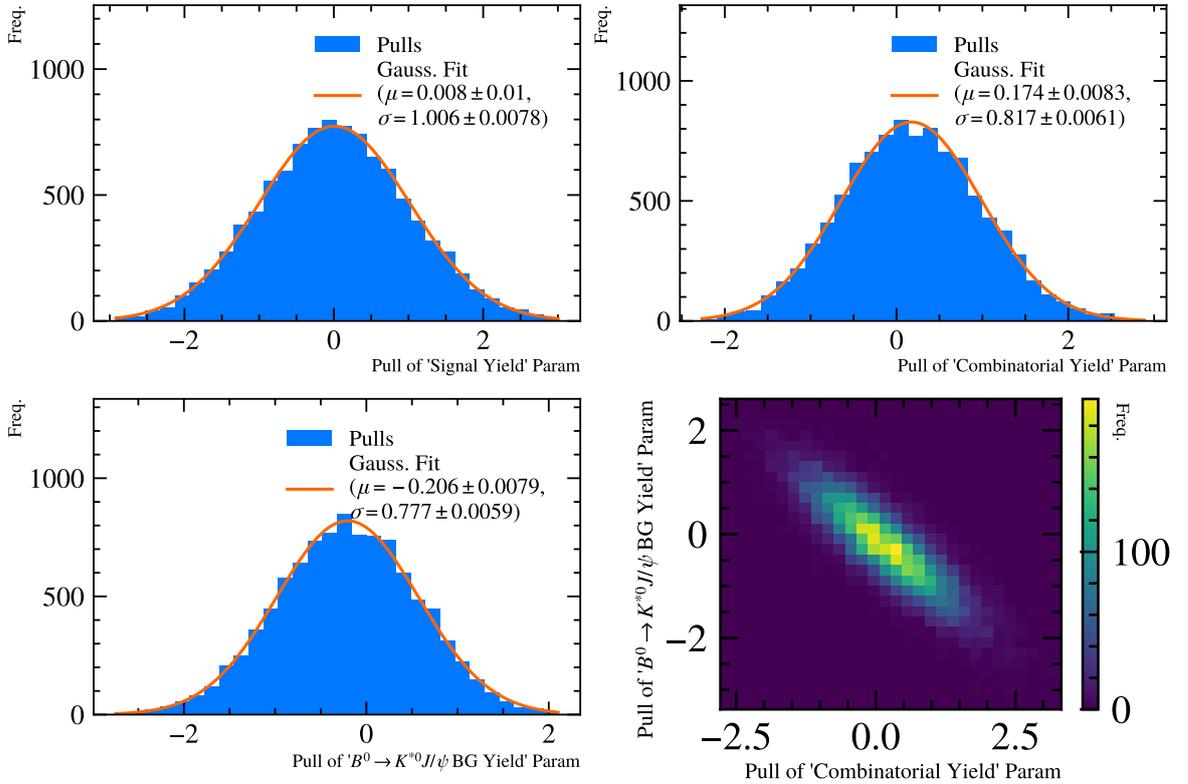


Figure 5.9: Pull distributions for the floating yield-like parameters for 10000 pseudo-experiments generated for the Run 2 control mode data fit displayed in Figure 5.8b. The $f_{K^{*0}}$ parameter is labelled as “B⁰ → K^{*0}J/ψ yield”. A 2D pull distribution showing an extreme correlation between the combinatorial and background yield. Distributions not included are displayed in Figure D.2.

The fit stability was validated as before with 10000 pseudo-experiments with the distribution of pulls of floating parameters presented in Figure 5.9. Importantly, $N_{Control}$ shows absolute stability, confirming that the method is robust against fluctuations in the dataset and the errors are correctly calculated. The exclusive background fraction $f_{K^{*0}}$ is constrained in the fit hence the non-standard pull distribution. This propagates to the floating combinatorial yield due to the high correlation between N_{Comb} and $f_{K^{*0}}$ shown

in the figure, implying that the overall background yield is stable. The signal-shape parameters $\Delta\mu$ and f_σ also show normally distributed pulls from pseudo-experiments further supporting the validity of the control mode yield extraction by stating that the shape of the A_b^0 peak will be consistently well-described under statistical fluctuations.

5.3.3.3 Systematic Comparison between Control Mode Yield Fit Methods

While the fit-method without exclusive background components is less complex and with similar fit-stability, there is a potential for the simpler shape to misrepresent the contribution of the combinatorial background to the total yield, which could bias $N_{Control}$. Table 5.3 describes good consistency between the two methods for the final yield, as well as for $\Delta\mu$ and f_σ controlling the signal shape. The fit with the exclusive background is preferred, considering the definite presence of backgrounds, as well as the expected stability to the fit range, unlike the simpler model that is observed to breakdown at any wider range.

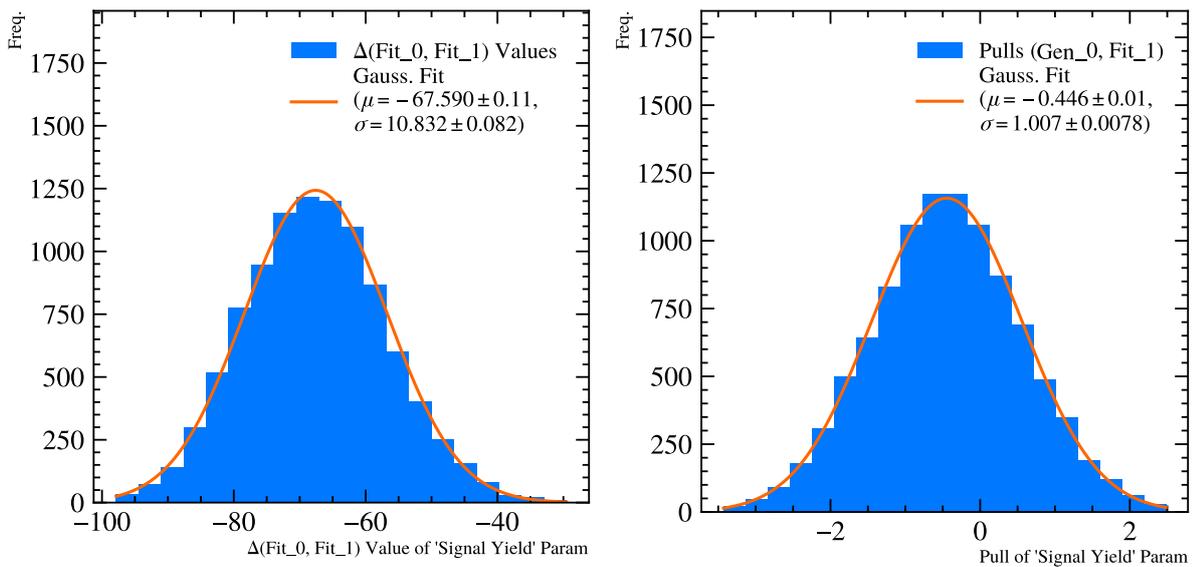


Figure 5.10: Distributions from cross-fit pseudo-experiments for the Run 2 control mode, where distributions are generated from the nominal model and fitted with the nominal and alternative. (left, a) The raw difference in yield between the nominal (with exclusive background) and alternative (without) control mode fitter. (right, b) The pull distribution of the alternative model fit, using the fitted yield uncertainty only. Figure D.3 shows the Run 1 result.

To understand the actual compatibility between the two methods, the pseudo-experiments discussed were cross-examined by also fitting each dataset generated by one model with the alternative model. This enables robust conclusions to be drawn on the

effect of the choice of model for the control mode fit, rather than relying on the singular statistic from comparing the real data results in Table 5.3. Figure 5.10 shows two metrics for measuring the systematic difference in $N_{Control}$, with 10000 pseudo-experiments generated by the nominal model. Figure 5.10a, showing the difference in $N_{Control}$ when both models are refitted to each pseudo-experiment, provides an absolute comparison between the fit-models, significant at -67.6 ± 0.1 , but convolves any shortcomings or mis-measurement within the nominal method as well as the alternative too. Figure 5.10b considers only the fit value using the alternative method, calculating the pull from the generated value using the nominal method, with the $\sigma \approx 1$ implying that the $N_{Control}^{Alt}$ is well-estimated but is consistently pulling -0.45σ from the nominal. The latter method is used to evaluate the systematic effect of choosing the nominal method for the analysis over the alternative. The mean of the pull is calculated in each analysis category and scaled by the nominal fit uncertainty for $N_{Control}$ resulting in $\sigma_{sys}(N_{Control}) = 86$ and 67 for Run 1 and 2 respectively, a 0.97% and 0.34% systematic effect. This ensures that the $N_{Control}$ uncertainty as an input to Equation 3.2.1 accounts for both the statistical fluctuation in the analysis dataset as well as any bias introduced when specific choices were made when constructing the fit.

5.3.4 Fit Validation - M_{pK} Spectrum

As discussed, the control mode is inclusive in the range $1450 < M_{pK} < 1850$ MeV selecting a mixture of Λ^* resonances, rather than exclusively selecting $\Lambda(1520)$ as in the signal mode. However it is still important to understand and cross-check that our M_{pK} distribution is consistent with other $\Lambda_b^0 \rightarrow pK^-\ell^+\ell^-$ analyses and that if the process of extracting a $\Lambda_b^0 \rightarrow \Lambda(1520)(\rightarrow pK^-)J/\psi(\rightarrow \mu^+\mu^-)$ yield is followed it produces the expected fit fraction.

In order to produce an M_{pK} distribution which represents contributions only from the signal component of the fit in Figure 5.5, the *sPlot* method is used as discussed in Section 4.2 to produce a set of signal sWeights that amplify the candidates more likely to be from the signal component and suppress those likely to be originating from the combinatorial background [165]. In contrast to Section 4.2, which used the *sPlot* method to create signal-like distributions in kinematic and event variables that are orthogonal from the sWeight fitting variable $M(\Lambda_b^0)$, extracting a signal-like distribution in M_{pK} has the potential for correlations between the two mass variables causing problems with the sWeight extraction. For the *sPlot* method to succeed, the two variables must be independent and factorisable, hence a *u*-statistic permutation test [194] is performed to determine the consistency of the “independent” hypothesis for $M(\Lambda_b^0)$ and M_{pK} , with the statistic distribution and result shown in Figure 5.11b. The *p*-value of 0.124 agrees

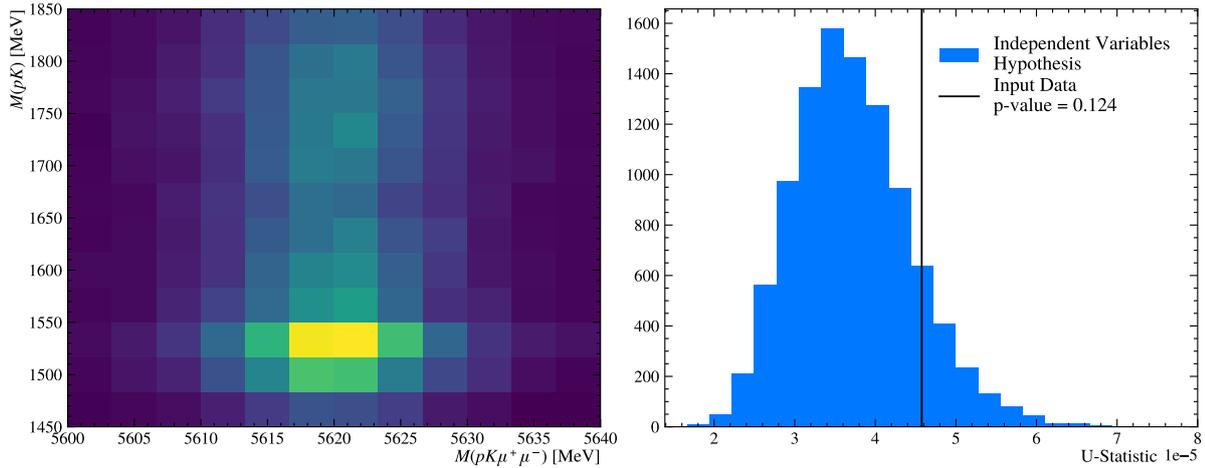


Figure 5.11: Independence test between reconstructed $M(\Lambda_b^0)$ and M_{pK} for Run 2 control mode data. (left) A two-dimensional histogram of masses, showing no unexpected correlation, besides the excess at the high-rate $\Lambda(1520)$ mass. (right) The u -statistic permutation distribution and corresponding test value for this analysis [194].

with the independent hypothesis at $\sim 1.15\sigma$ implying that using sWeights to create the signal-like M_{pK} distribution is valid.

To extract the yield from the dominant peaking $\Lambda(1520)$ component of this distribution, the fitting technique derived for and used by the $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ branching fraction analysis is directly translated to this analysis [105]. This uses the PDG values for the mass and widths of the five dominant Λ^* resonances $\Lambda(1520)$, $\Lambda(1405)$, $\Lambda(1600)$, $\Lambda(1690)$, and $\Lambda(1800)$ [41] to construct Λ^* line shape functions combining Relativistic Breit-Wigners with measured momentum and Blatt-Weisskopf barrier factors [195]. The resonance shapes are corrected for the varying acceptance efficiency across the M_{pK} spectrum with the $\Lambda(1520)$ shape also convolved by a Gaussian resolution function (G) to emulate the detector resolution of the dominant component. This produces an overall fit function of

$$P_{tot}(M_{pK}) = N_{\Lambda(1520)} \cdot P_{\Lambda(1520)} \otimes G(\Lambda(1520)) + N_{\Lambda^*} \cdot \sum_{\Lambda^*} f_{\Lambda^*} \cdot P_{\Lambda^*} \quad (5.3.4)$$

where P refers to the corrected Λ^* line shape, with the floating parameters $N_{\Lambda(1520)}$ (the yield of $\Lambda_b^0 \rightarrow \Lambda(1520)(\rightarrow pK^-)J/\psi(\rightarrow \mu^+\mu^-)$) and N_{Λ^*} the sum of yields of the other Λ^* components combined by their relative fractions f_{Λ^*} also floating in the fit [76].

The results for the maximum likelihood fits of Equation 5.3.4 are shown in Table 5.4 with the corresponding fit for Run 2 shown in Figure 5.12, with the different Λ^* components' contributions overlaid. The results produce a $\Lambda(1520)$ fit fraction in the range $(1450 < M_{pK} < 1850)$ MeV of $(15.7 \pm 0.7)\%$ for Run 1 and $(19.9 \pm 0.5)\%$ for Run 2, inconsistent with the corresponding Run 1 result from the $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ branch-

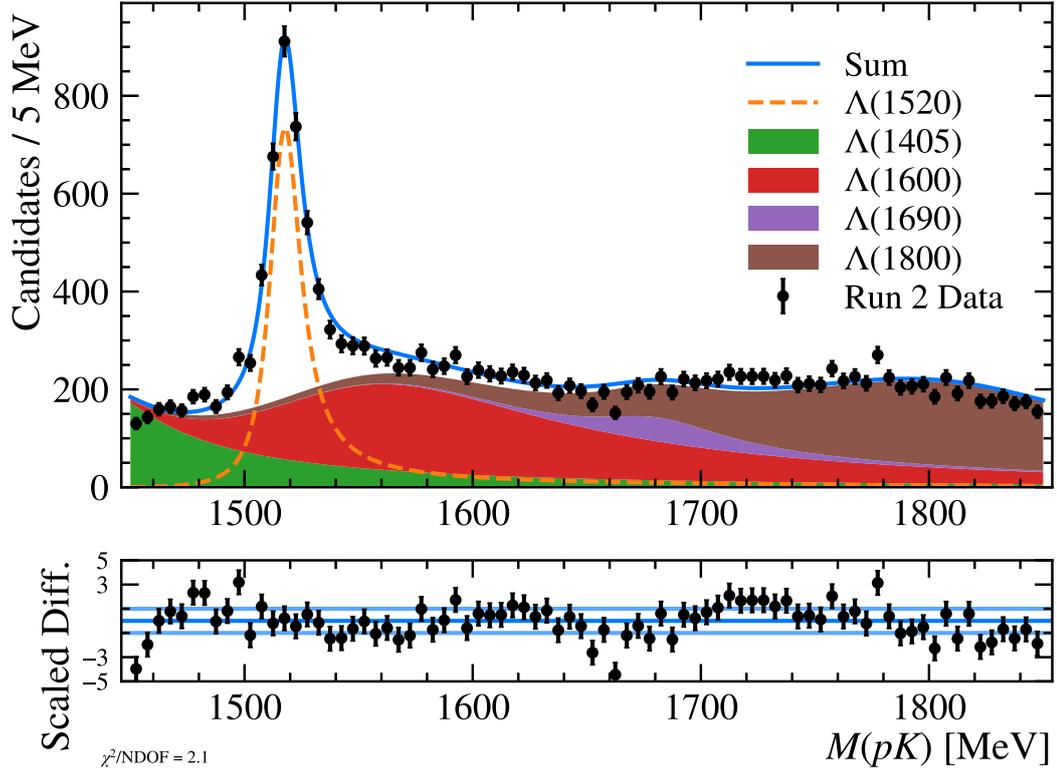


Figure 5.12: Invariant mass fit to the M_{pK} spectrum for Run 2 control mode data using the model defined in Equation 5.3.4. The corresponding Run 1 fit is included in Figure D.4.

Table 5.4: Parameter results for invariant mass fits to the M_{pK} spectrum with $\Lambda_b^0 \rightarrow pK^- J/\psi$ sWeighted data. Run 1 and 2 for this analysis are presented alongside the Run 1 results from the $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ analysis for a fit to $\Lambda_b^0 \rightarrow pK^- J/\psi$.

Analysis Category	M_{pK} Fit Parameters				
	$N_{\Lambda(1520)}$	N_{Λ^*}	$f_{\Lambda(1405)}$	$f_{\Lambda(1600)}$	$f_{\Lambda(1690)}$
This Analysis					
Run 1	1385 ± 62	7451 ± 99	0.174(9)	0.42(2)	0.05(1)
Run 2	3905 ± 98	15730 ± 150	0.163(6)	0.42(1)	0.056(7)
$\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$					
Run 1	6468 ± 127	17984 ± 222	0.1958(66)	0.5704(98)	N/A

ing fraction analysis producing an equivalent $\Lambda(1520)$ fit fraction $(26.5 \pm 0.6)\%$ despite near-identical implementation of the fitting technique.

There are a number of potential causes for this. $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ and this analysis use distinct stripping and pre-selection, with overall selection efficiencies a factor of 3-5

Table 5.5: Parameter results for DCB maximum-likelihood fits to the $M(pK^-\ell^+\ell'^{-})$ distribution in each analysis category for the corrected signal and control mode simulation.

Analysis Category	DCB Parameters					
	μ (MeV)	σ (MeV)	α_l	α_r	n_l	n_r
Signal Mode						
Run 1 0γ	5610.9(7)	16.1(4)	0.36(2)	2.5(2)	2.5(3)	1.7(6)
Run 1 1γ	5611.3(7)	35(1)	0.87(6)	1.11(7)	2.4(3)	1.8(3)
Run 2 0γ	5610.8(5)	17.1(4)	0.40(2)	2.4(1)	2.3(1)	2.3(6)
Run 2 1γ	5613.0(6)	34(1)	0.82(5)	1.04(6)	2.6(3)	2.0(3)
Control Mode						
Run 1	5619.97(3)	5.08(3)	1.97(3)	1.76(3)	2.03(7)	2.5(1)
Run 2	5620.02(2)	5.19(2)	1.92(2)	1.77(2)	2.06(4)	2.47(5)

different. There is also expectation of considerable systematic uncertainty from the M_{pK} fit technique, not estimated in this analysis, but assuming a $\sim 10\%$ uncertainty on both values of $N_{\Lambda(1520)}$ from studies in the $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ analysis, the consistency reduces to $< 3\sigma$, but this assumes no cancellation of shared systematic effects, highly unlikely for near-identical fitting schemes. The result suggests that for any future iteration of the analysis that relies on the extracted $N_{\Lambda(1520)}$ value, further investigation is required, but for this result $N_{Control}$ is independent of an M_{pK} fit.

As well as a cross check, the fitted PDF of the $\Lambda(1520)$ component is used in Section 4.1.5.1 as a data-driven approach for determining the optimal cut in M_{pK} for extracting $\Lambda(1520)$ contributions in the signal mode. The fit also validates and supports the choice to use the inclusive $1450 < M_{pK} < 1850$ MeV range for the control mode, with the additional process of *sPlot* and extracting $N_{\Lambda(1520)}$ increasing the relative yield uncertainty to 2.5% compared to $\sigma_N/N_{Control} = 0.73\%$ for the nominal method.

5.4 Mass Fits of the Signal Mode

The invariant mass fit for the signal mode requires a distinct approach and suite of cross-checks from the control mode. This is due to the blinding of the central ($5200 < M(pK^-\mu^\pm e^\mp) < 5800$) MeV region until the analysis is near-complete. Therefore invariant mass fits focus on using external inputs as well as the upper and lower sideband to constrain and predict the background content of the blinded signal region.

The nominal fit method takes the signal-shape from fits to $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$

simulation, the combinatorial background shape from template fits to the μe SS proxy distribution and fixes exclusive background components following the method discussed in Section 5.2. The overall signal fit equation is defined as

$$P(\Lambda(1520)\mu e) = N_{Signal}P_{MC}^{DCB} + N_{Comb}P_{\mu eSS}^{Cheby} + (N_{D^0pe} + N_{D^0p\mu})P_{MC}^{John}. \quad (5.4.1)$$

where P_{MC}^{DCB} is a DCB function describing the signal-shape, $P_{\mu eSS}^{Cheby}$ a third-order Chebyshev polynomial describing the combinatorial and P_{BG}^{John} the pre-fitted JohnsonSU function from Section 5.2.2 to describe the combined remaining exclusive backgrounds $\Lambda_b^0 \rightarrow D^0(\rightarrow K^-\ell^+\nu)p\ell^-\bar{\nu}$. The combined yield of $N_{D^0p\ell}$ is allowed to float in the fit for the presented iteration of the analysis, due to the unknown $\mathcal{B}(\Lambda_b^0 \rightarrow D^0p\ell^-\bar{\nu})$ discussed in Section 5.1.1, preventing the preferred method of applying a Gaussian constraint of predicted $N_{BG} \pm \sigma_{BG}$ on each background yield.

By default, the parameters of each PDF are fixed from their external sources leaving the only unconstrained floating parameters as N_{Signal} , N_{Comb} and $N_{D^0p\ell}$ allowing the maximum likelihood fit to converge even when performed with the limited statistics seen in Figure 4.34 after full-selection. The subsequent sections describe how the shape of the signal and combinatorial background PDFs are determined.

5.4.1 MC Signal Component

The result of a maximum-likelihood fit to the $M(pK^-\mu^\pm e^\mp)$ spectrum of the signal simulation in each analysis category is included in Figure 5.13, with the results summarised in Table 5.5. The distinct difference in the shape of the 0 γ and 1 γ categories, both visually and by the parameter values, further justify the decision to split the analysis by this metric. The category without Bremsstrahlung reconstructed, as expected, has a long low-mass tail from missing energy while the 1 γ category distributions avoid the lower-tail due to the recovered energy, but show a larger tail above the mean due to over-reconstruction. This difference manifests also in a shift in the μ of the two fits below the expected $M_{\Lambda_b^0} = (5619.6 \pm 0.2)$ MeV [41]. This is understood to be due to an overall under-reconstruction due to energy lost in electron reconstruction, significantly below the μ value from the $\Lambda_b^0 \rightarrow pK^- J/\psi$ fit from Section 5.3.1. The recovery of Bremsstrahlung improves this, with the 1 γ categories consistently above 0 γ across both LHC Runs, but the deficit remains. This effect has been observed in other LHCb analyses, the fits to $\Lambda_b^0 \rightarrow pK^- e^+ e^-$ simulation from the R_{pK} analysis show the same trend for the equivalent categories 0 γ , 1 γ and 2 γ (where both electrons are reconstructed with Bremsstrahlung), with $\mu_{0\gamma} < \mu_{1\gamma} < \mu_{2\gamma} < m_{\Lambda_b^0}$ [76].

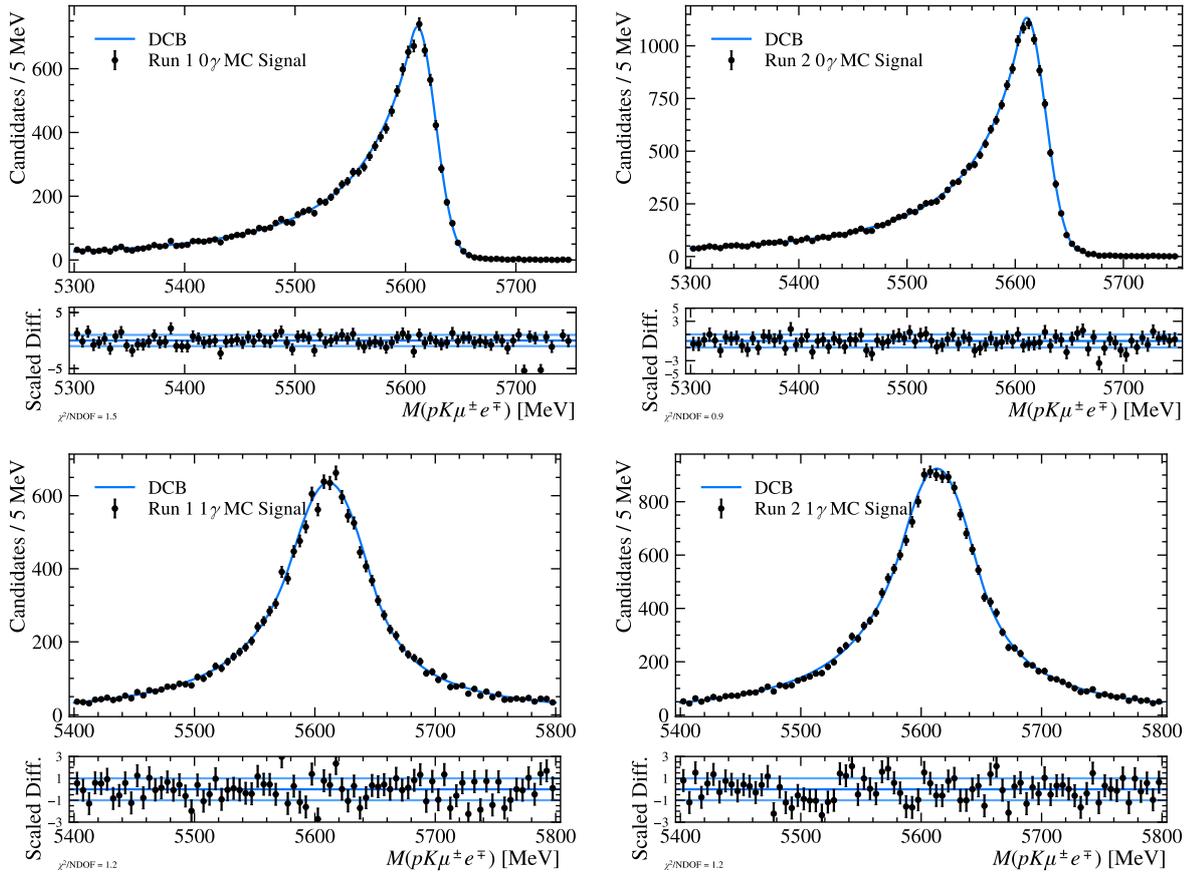


Figure 5.13: Invariant mass fits to the $M(\Lambda_b^0)$ distribution of corrected $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ signal mode simulation split by Run 1 (left), Run 2 (right) and categories 0γ (upper), 1γ (lower). The fit is performed with a DCB shape with all parameters floated.

The DTF algorithm is not used to constrain the $M(\Lambda_b^0)$ distribution in the signal mode as it is in the control mode. Without the narrow dominant $\ell^+\ell^-$ resonance, the only constraint that could be applied is a $\Lambda(1520)$ mass constraint or the requirement that the reconstructed Λ_b^0 originates from the PV. For similar analyses, such as the $\Lambda_b^0 \rightarrow \Lambda e^\mp \mu^\pm$ analysis, the DTF constraint is put on the hadronic resonance in the LFV mode, which improves the mass resolution [42], however DTF mass constraints perform poorly for resonances with larger widths [193], like the $\Lambda(1520)$, and attempting it in this analysis worsened the resolution. The PV constraint on Λ_b^0 was also checked, but made negligible difference to the mass resolution and was also rejected.

5.4.2 Constraining the Combinatorial Shape with μe SS Data

The limited statistics in the blinded signal mode of each analysis category render a floating-fit to the combinatorial background shape impossible, with 10 (23) and 20 (16)

events in Run 1 0γ (1γ) and Run 2 0γ (1γ) categories respectively. The situation eases slightly by utilising the μe SS dataset which is not blinded in the signal region increasing the statistics available for a possible fit, providing an analogous dataset that can be treated as “combinatorial only” (the status after full selection is shown in Figure 4.34). The ideal method to determine the combinatorial shape would be to simply perform a maximum-likelihood fit of $M(A_b^0)$ with the μe SS distribution in each category after full selection, fixing the parameters and transfer the shape to the overall blinded fit. The μe SS dataset is however also limited by sample size and floating fits were found to be unstable and rarely-converge. Hence, a “cascade” fitting method was derived where the p_{MVA} cut was initially loosened to increase the sample size before sequentially tightening and refitting via the following steps:

Loose Working Point

Relax to the loose p_{MVA} cut value where $\varepsilon_{MVA} = 99.5\%$ and perform a floating simultaneous fit to the Run 1 and 2 μe SS distributions separately for 0γ and 1γ , improving the sample-size while resolving the expected shape difference between Bremsstrahlung categories.

Medium Working Point

Tighten to the medium p_{MVA} with $\varepsilon_{MVA} = 97\%$, Gaussian constrain the parameter values from the previous step and refit in each category individually, allowing any slight change in shape from the tightening of p_{MVA} to be accounted for as well as any difference between the data taking periods.

Nominal Working Point

Fix the shape parameters and refit with the nominal p_{MVA} selection cut ($\varepsilon_{MVA} = 90\%$), providing a check that the looser combinatorial shape continues to describe the μe SS distribution at the ultimate analysis working point.

The three working points (WP) correspond to p_{MVA} cuts at the MVA values stated in Table 4.9. The final shape in each category is then transferred to the overall blinded fit where the N_{comb} yield is allowed to float to account for differences between the amount of μe SS and μe OS events.

A third-order Chebyshev polynomial was found to best describe the signal shape, with a second-order unable to describe the upper-sideband tailing off. The results from the μe SS fits are summarised in Table 5.6 with the data and overlaid-fit shape at each stage shown in Figure 5.14 for the Run 1 0γ and Run 2 1γ categories (with the remaining distributions in Appendix D). The fits describe the corresponding data well at all three working points, considering the sample size, and result in a smooth decaying shape across

Table 5.6: Results from the μe SS fits at each working point in the analysis categories. The sharing of parameters between LHC Runs is shown by the shared cells in the Loose column. The Nominal working point shows a yield from the μe SS fit only, across the full analysis range. The uncertainties are quoted to two significant figures for X_i to show detail of the small values in these simultaneous and constrained fits. Yields are quoted with a two-sided error.

Working Point	Loose			Medium			Nominal
Parameter	X_0	X_1	X_2	X_0	X_1	X_2	$N_{\mu eSS}$
Run 1 0γ	-0.47(13)	-0.58(16)	0.14(13)	-0.51(12)	-0.60(15)	0.10(1.2)	$16.0^{+4.3}_{-3.7}$
Run 2 0γ				-0.52(11)	-0.66(14)	0.05(12)	$23.0^{+5.1}_{-4.5}$
Run 1 1γ	-0.594(81)	-0.63(11)	0.372(89)	-0.589(74)	-0.62(10)	0.355(84)	$19.0^{+4.7}_{-4}$
Run 2 1γ				-0.623(69)	-0.635(89)	0.433(75)	$37.0^{+6.4}_{-5.8}$
Run 1 $(0+1)\gamma$	-0.552(67)	-0.618(88)	0.295(75)	-0.563(63)	-0.622(83)	0.263(70)	$35.0^{+6.3}_{-5.6}$

the blinded region aiding with the stability of the subsequent limit setting. Alternative fits including a JohnsonSU [192] and Argus distribution [191] were considered with the first forming a peak in the signal region, an undesirable feature for a combinatorial background, while the latter was unable to describe the tail of the upper-sideband in some categories. Future iterations of the analysis could reconsider the Argus distribution with the advantage of fewer floating parameters, as well as other polynomials.

To validate the method of determining the combinatorial shape, two checks were adopted. Pseudo-experiments were generated for the medium working point to study the stability of the combinatorial-shape before it is fixed for the μe OS blinded-fits and further pseudo-experiments are generated at the nominal working point to ensure that the combinatorial-proxy dataset at full selection is actually well-described, and not that the fits shown in Figure 5.14 are converging by fortune. Each category produced standard normal-distributions for the medium and nominal working point yields, suggesting the fits provide a good description of the fully-selected $M(\Lambda_b^0)$ distributions. Slight bias in the Chebyshev shape parameters were observed, producing non-standard distributions due to the inter-parameter correlations and Gaussian constraints in the fit. Over 95% of the pseudo-experiments at the medium-WP were successful however, sufficient considering the small yield and hence relatively large fluctuations expected between experiments.

In addition, fine-scans of p_{MVA} were made to study the consistency of the $M(\Lambda_b^0)$ distribution as the cut is tightened, evaluating the validity of the assumption that the shape at the loose WP is representative of the tight. Figure 5.15 describes the trend of the Chebyshev fit parameters if allowed to float in a maximum-likelihood fit at each p_{MVA} value for the Run 1 and 2 combined 1γ categories. Parameters X_1 and X_2 show robust self-consistency across the range of p_{MVA} values tested, implying the core shape

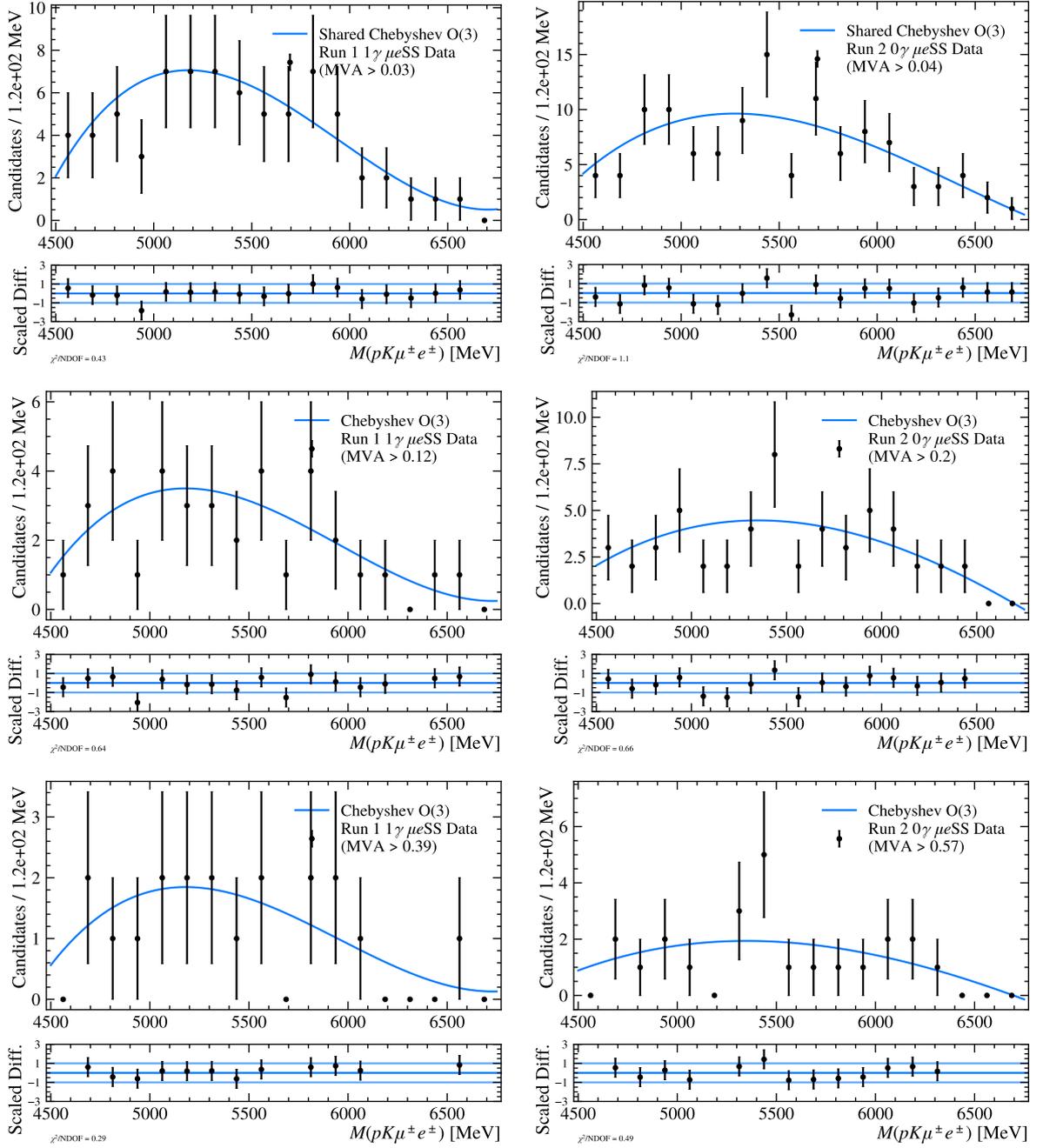


Figure 5.14: Invariant mass fits to the $M(\Lambda_b^0)$ distribution in the μe SS combinatorial proxy dataset for Run 1 1γ (left) and Run 2 0γ (right) with the plots vertically ordered by the loose, medium and nominal working points. Note, the loose fits are simultaneously derived with the other Run category, not shown but included in Figure D.6 with the remaining fits.

of the combinatorial background doesn't change. The third-order parameter X_3 appears to show a slight trend, but crucially it remains compatible with the actual results from the constrained medium working point for both categories. This compatibility is also displayed in X_1 and X_2 . This implies that the “cascade” method discussed is valid to

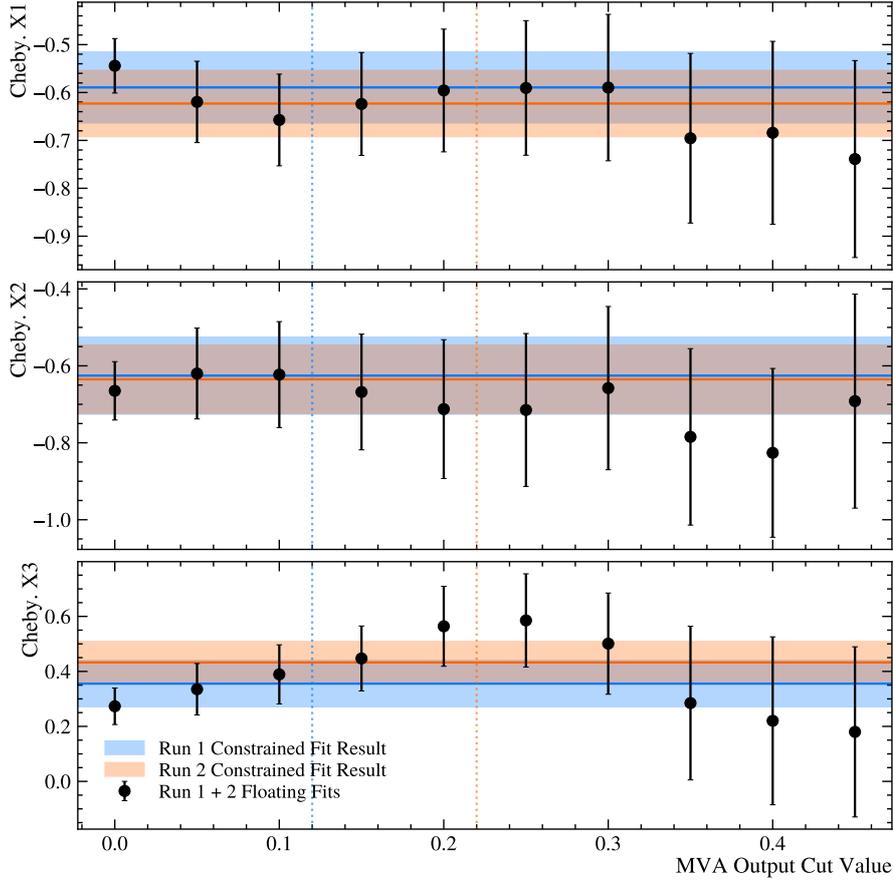


Figure 5.15: Consistency test of the μeSS distribution when parameters are left to float for Run 1 and 2 1γ categories combined. Black points describe the parameter value and uncertainties for each floating fit at the corresponding p_{MVA} cut values, with each dataset a tighter-selected subset of the previous. Above $p_{MVA} = 0.45$ the fit would not converge. The coloured bars show the Gaussian constrained results at the corresponding 1γ medium working point, with dashed lines to indicate which p_{MVA} cut value this corresponds to.

describe the shape of the combinatorial background and applying the Gaussian constraint from the low to medium WP essentially aids in reducing the parameter uncertainty, also shown in Figure 5.15, while not systematically altering the central values. This method breaks down due to lack of sample size before the Run 2 1γ nominal WP value of 0.6 is reached, but the compatibility at the Run 1 equivalent of 0.39 suggests that the cascade method is valid for determining a combinatorial background shape for the final blinded fits.

5.4.3 Blinded Background fits in $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$

With all components of Equation 5.4.1 prepared for each category, blinded fits to the signal mode can be made. Initially fits are performed individually in each category but a simul-

Table 5.7: Blinded fit results for the μe OS fit to $M(\Lambda_b^0)$ in each analysis category, including the fit without the exclusive $\Lambda_b^0 \rightarrow D^0(\rightarrow K^-\ell^+\nu)p\ell^-\bar{\nu}$ component. The blinded yields are displayed alongside yields integrated across the signal region ($5200 < M(\Lambda_b^0) < 5800$) in brackets. The μe SS combinatorial proxy equivalent result is included for reference.

Dataset Category Parameter	μe SS	μe OS Blinded Fit		
		Comb. + Excl. Bg.		Comb. Only
	N_{SS}^{blind}	$N_{Comb.}$	$N_{D^0 p\ell}$	$N_{Comb.}$
Run 1 0 γ	10.2 (5.8)	$8.3^{+4.8}_{-3.9}$ (4.7)	$1.7^{+3.9}_{-1.7}$ (0.18)	$10.0^{+3.5}_{-2.8}$ (4.7)
Run 1 1 γ	11.7 (7.3)	$23.1^{+5.1}_{-6.6}$ (14.5)	$0^{+5.6}_{-0}$ (0)	$23.0^{+5.1}_{-4.5}$ (14.5)
Run 2 0 γ	17.2 (8.5)	$14.8^{+5.9}_{-4.7}$ (8.6)	$5.2^{+4.4}_{-4.2}$ (0.52)	$20.0^{+4.8}_{-4.1}$ (8.6)
Run 2 1 γ	22.5 (14.5)	$11.1^{+8.0}_{-6.1}$ (7.2)	$4.8^{+6.6}_{-4.8}$ (0.76)	$16.0^{+4.3}_{-3.7}$ (7.2)
Run 1 (0+1) γ	21.8 (13.2)	$30.2^{+8.2}_{-8.4}$ (18.5)	$2.8^{+7.3}_{-2.8}$ (0.36)	$33.0^{+6.1}_{-5.4}$ (20.2)

taneous fit is also set-up such that, upon unblinding, a combined $\mathcal{B}(\Lambda_b^0 \rightarrow A(1520)\mu^\pm e^\mp)$ can be constructed from the the four yields. For the blinded fits, N_{Signal} is fixed to 0, hence the only floating parameters are N_{Comb} and the yield $N_{D^0 p\ell}$ (if included), enabling the fits to converge in each category despite the low statistics.

The results of the blinded fits are summarised in Table 5.7, with the corresponding distributions shown in Figure 5.16 for the floating $N_{D^0 p\ell}$ fit method in the four analysis categories. With such small statistics, it is difficult to draw conclusions about the consistency of the blinded μe SS yields and N_{Comb} values, suggesting the combinatorial proxy dataset should continue to be used for shape parameters only. The low statistics of the Run 1 0 γ category results in poor stability from pseudo-experiments. The Run 1 1 γ fit also converges on $N_{D^0 p\ell} = 0$, conflicting with the clear contribution across the Run 2 categories. While the analysis has thusfar focused on retaining separate 0/1 γ categories, with distinct signal shapes (see Figure 5.13) and a significant difference in ε_{Signal} observed, the limited statistics in Run 1 favours a combined Bremsstrahlung approach.

To combine the Run 1 categories a similar approach to other LHCb LFV analyses is followed using two MC signal shapes, fixed by the ratio of MC events, with the efficiencies calculated in a combined chain. This is derived for Run 1 only resulting in three final categories. The μe SS combinatorial shape parameters are recalculated, but as seen in Table 5.6, the dominance of the 1 γ category in Run 1 results in a compatible shape for the “(0+1) γ ”, Run 1 category. This fit is considerably more stable as a result of the increased statistics. The separation between the Run 2 categories is retained to understand the analysis sensitivity as a function of Bremsstrahlung recovery. The Run 2 recovery algorithm also performs better (see Section 2.3), so any difference will be better resolved than with the intended Run 1 categories.

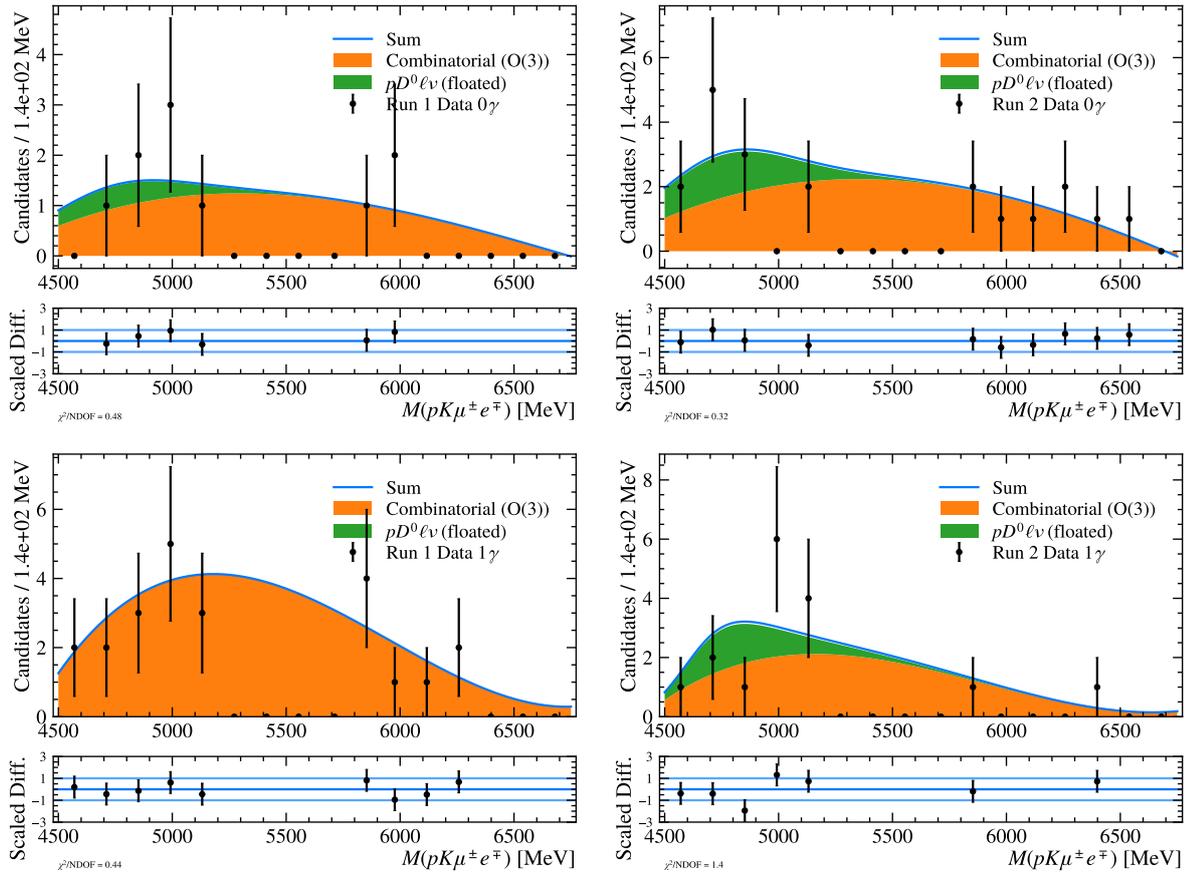


Figure 5.16: Blind fits to the $M(\Lambda_b^0)$ distribution of signal mode μe OS data split by Run 1 (left), Run 2 (right) and categories 0γ (upper), 1γ (lower). The fit is performed with the description in Equation 5.4.1 with the N_{Signal} fixed to zero. For the Run 1 0γ fit $N_{D^0 p \ell}$ converges at zero.

The final blinded fit for the Run 1 category is shown in Figure 5.17a, with the fit visually describing the data better than was observed in the split category case. The three final categories describe a stable trend across the signal region, with the contribution from the exclusive component almost entirely contained in the lower sideband ensuring that, upon unblinding the combinatorial shape will not significantly change. The yield for Run 1 is also included in Table 5.7, allowing comparison in each category of the $N_{D^0 p \ell}$ to the predicted N_{BG} prediction from Equation 5.1.1. They are, as discussed, an order of magnitude smaller than predicted, a factor of 7.5 (5.5) for Run 1 (2), however the ratio of predicted and measured $N_{D^0 p \ell}$ for the Run 2 categories are compatible (≈ 0.9), with only a 30% discrepancy for Run 1. This implies that the background estimation method is valid and that the discrepancy is caused by the unmeasured $\mathcal{B}(\Lambda_b^0 \rightarrow D^0 p \ell^- \bar{\nu})$ which is projected to 0.1-0.15% from the blinded fits.

Blinded pseudo-experiments are generated for each individual fit, 5000 in each category as before, showing that the yield determination is approximately stable upon small

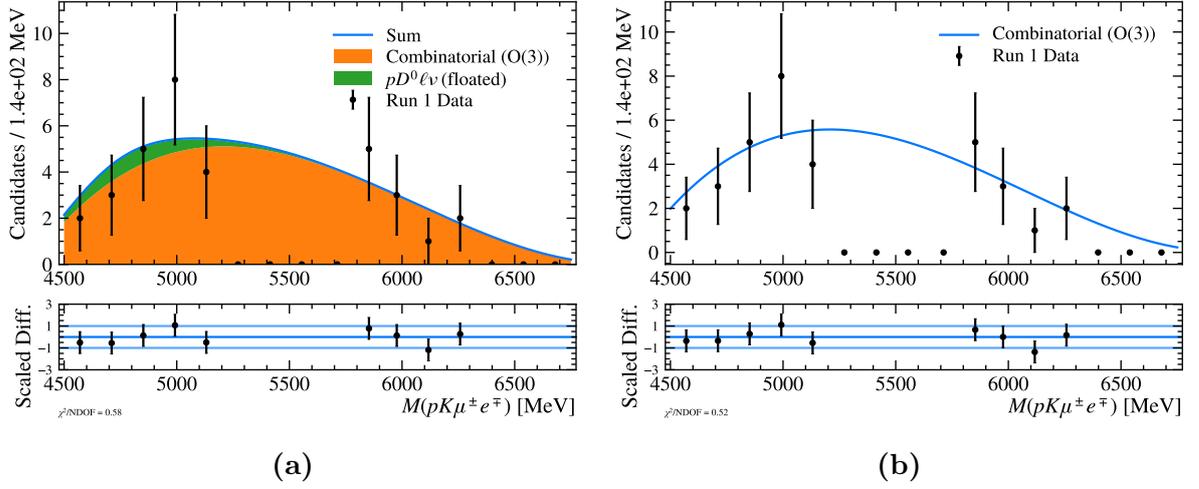


Figure 5.17: Comparison of blinded fits for the Run 1 combined Bremsstrahlung category with (a) and without (b) the exclusive $\Lambda_b^0 \rightarrow D^0(\rightarrow K^-\ell^+\nu)pl^-\bar{\nu}$ component. The remaining distributions for Run 2 0γ and 1γ are shown in Figure D.5.

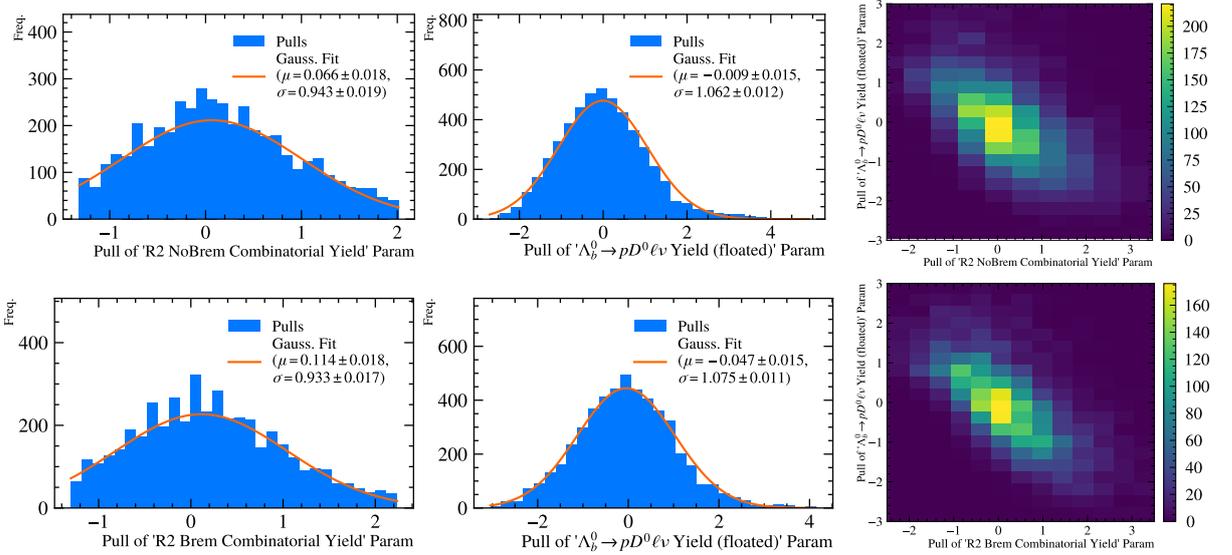


Figure 5.18: Example pull distributions of the floating parameters in blinded μe OS fits for Run 2 0γ (upper) and Run 2 1γ (lower), from 5000 generated blind pseudo-experiments for the nominal method with a floating N_{Comb} (left) and N_{D^0pl} (middle) with a 2D pull distribution (right) included to display correlation. Figure D.7 shows the Run 1 category results.

changes to the generated distribution (Run 2 examples in Figure 5.18), with unity $\sigma \approx 1$ stating the error of each parameter is determined correctly. The two yields show correlation across the categories, implying that the combinatorial is able to partially absorb the exclusive background in some pseudo-experiments and vice versa. However the N_{D^0pl} pull distributions all have a mean consistent with zero, supporting the inclusion of the exclusive background component in the final μe OS fit method.

In the Run 1 case in Figure 5.17b, the argument could be made that the small contribution of $N_{D^0 p \ell}$ produces a sufficient fit without it, but for the Run 2 categories the combinatorial shape would over-parameterise the upper sideband if it had to compensate for the removed exclusive component (Figure D.5 included for reference). Therefore for a consistent approach, the exclusive background is retained for all three final categories, allowing each to float in the individual fits. This is supported by the knowledge that the background is present, just at an unknown rate. If the analysis selection chain was looser, the exclusive component would be required and the distinct shape shown, especially in the Run 2 1γ category, confirms that this is the conservative and more correct approach. While the $\mathcal{B}(A_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp)$ sensitivity will be determined in each category, a simultaneous fit will also be constructed and retaining all $N_{D^0 p \ell}$ components will enable the three yields to be bound by their relative N_{BG} prediction, factoring out the unknown $\mathcal{B}(A_b^0 \rightarrow D^0 p \ell^- \bar{\nu})$.

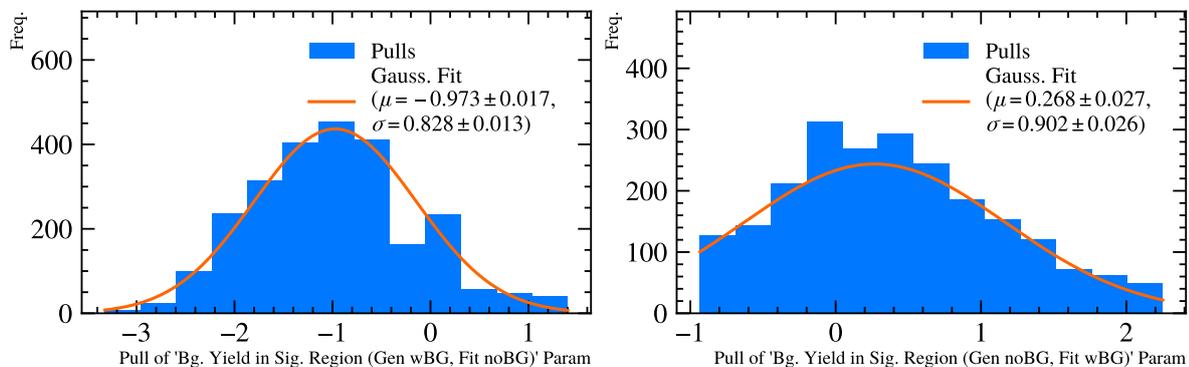


Figure 5.19: Cross-compatibility of the two background fit schemes considered for the signal mode, derived by generating 5000 pseudo-experiments for both models where the $D^0 p \ell$ is included or not, and fitting with the opposite hypothesis. (left) The pull distribution of the integrated signal region background yield with the $D^0 p \ell$ generated but not included in the subsequent fit and (right) the inverse. The pulls for Run 2 0γ are shown with other categories included in Figure D.8.

To understand the impact of the choice to include the $D^0 p \ell$ component in the final fit, the effect on the signal region needs to be determined. While the shape of the background definition in the sidebands is important, the sensitivity to discover or set a limit on the signal mode is determined by the level and stability of the background below any potential $A_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ NP. Therefore cross-fit pseudo-experiments, with generation and fitting performed with opposite model types, are constructed with and without $N_{D^0 p \ell}$ fixed to zero, with the parameter of interest the integral of the background model between $(5200 < M(A_b^0) < 5800)$ MeV. The absolute difference in integrated fit-yields as well as a “pull” is derived in Figure 5.19 for the Run 2 0γ example, with corresponding results included in Figure D.8. The mean of the pull distributions imply that fitting without the exclusive background when a sample has been generated with

Table 5.8: Single-event sensitivities and estimated upper limits for the signal mode. The upper limits are quoted at the 95% confidence level.

	Single-Event Sensitivity	CL _s Method
Run 1	$(18.4 \pm 3.0) \times 10^{-9}$	11×10^{-8}
Run 2 0 γ	$(8.6 \pm 1.4) \times 10^{-9}$	3.3×10^{-8}
Run 2 1 γ	$(7.1 \pm 1.1) \times 10^{-9}$	3.3×10^{-8}
Total	$(3.22 \pm 0.32) \times 10^{-9}$	2.8×10^{-8}

consistently over-estimates the background yield in the signal region due to the attempt to accommodate any excess in the lower sideband with a single component. The inverse pseudo-experiments show that the extra background component may under-estimate the yield but to a lesser extent ($\mu = 0.27(2)$ versus $-0.97(2)$ in Run 2 0 γ). This suggests that even if no background is truly present, the model with an exclusive background will better describe the signal region, vital for the best branching fraction limit. This is most clear in the Run 2 categories, where the upper sideband is best described by the fits with a $D^0 p \ell$ component, but overall, further supports the conclusion to include the exclusive component in all categories.

5.5 Determination of an Estimated Upper Limit for $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp)$

At the time of writing the analysis remains blinded, therefore a final measurement of the branching fraction or upper limit (if a signal was not observed) is not possible. To determine an estimated $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp)$ sensitivity for the analysis, a blinded upper limit can instead be calculated where the fitted state of the sidebands from Section 5.4.3 is extended into the signal region and the potential significance of the MC signal shape from Section 5.4.1 at different $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp)$ working points (scaled from N_{Signal} with Equation 3.2.2) is scanned. This estimates the minimum required rate for the LFV decay to be measured, which excludes the process at smaller values of $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp)$.

5.5.1 Single-Event Sensitivity

The single-event sensitivity is defined as the value of $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp)$ if one signal mode event was observed. It provides an initial order-of-magnitude estimate of the

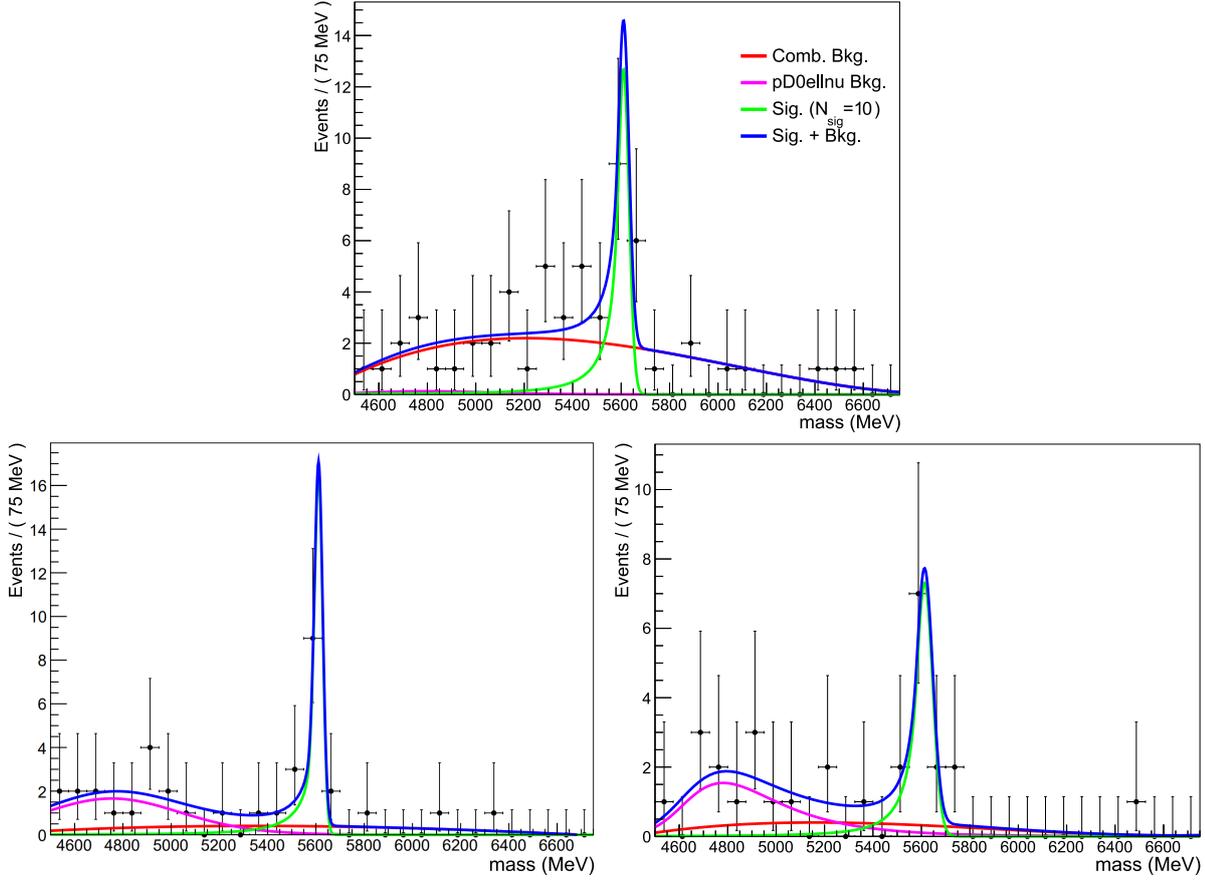


Figure 5.20: Signal fits to unblinded pseudo-experiments for each analysis category using the blinded fits in Section 5.4.3, integrated into the signal region. The signal yield is set to 10 for each pseudo-experiment to reveal how a potential measurement yielding evidence for LFV could look. (upper) Run 1, (lower left) Run 2 0γ and (lower right) Run 2 1γ with fitted $N_{Signal} = 14.1 \pm 5.6$, 13.2 ± 4.7 and 10.7 ± 4.0 respectively. “mass” refers to $M(pK^-\mu^\pm e^\mp)$ and the legend is shared.

expected upper limit using α defined in Equation 3.2.2, combining the other key results from the analysis ($N_{Control}$, $\varepsilon_{Control}$, ε_{Signal}). This is calculated for each analysis category, α_{cat} , and is subsequently combined by

$$\alpha_{tot} = \left(\sum_{cat} \frac{1}{\alpha_{cat}} \right)^{-1}. \quad (5.5.1)$$

The results are displayed in Table 5.8, where Run 2 is found to be more sensitive than Run 1, as expected with the larger initial data sample and the improvements to the LHCb trigger and reconstruction, overall improving the resolution. The recovery of Bremsstrahlung appears to slightly improve the sensitivity for Run 2, expected from the slightly greater ε_{Signal} in Table 4.12, but the two category sensitivities are compatible. The significant difference to Run 1 however justifies the analysis being performed in independent categories

and while the difference by Bremsstrahlung is small, the distinct signal and background shapes from the recovered energy support splitting the dataset where possible, as discussed. The total combined single-event sensitivity reaches 3.22×10^{-9} , a factor of 6.8 smaller than the projected prediction for $\mathcal{B}(A_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp)$ in Equation 1.3.2 [101]. This suggests that, after unblinding, if no signal is observed, the analysis could exclude certain NP models, and aid in the development and constraining of future models.

While a single event above the background model in the signal region would be insufficient to claim evidence or discovery of $A_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$, the very low background level in the analysis datasets mean as few as 10 signal events can produce a visible contribution to the $M(A_b^0)$ spectrum. Figure 5.20 displays a pseudo-experiment for each analysis category using the blinded fits in Section 5.4.3, where the dataset was generated with $\mathcal{B}(A_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp) = 10\alpha_{cat}$ and subsequently fit with the signal model defined in Equation 5.4.1. In all three categories the fitted pseudo-experiments produced N_{Signal} and hence $\mathcal{B}(A_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp)$ values of $2 - 3\sigma$ significance, meeting the threshold for evidence when combined. This justifies the ultra-low background approach followed in the analysis, as any looser selection would have required a larger $\mathcal{B}(A_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp)$ hypothesis to render the clear signal peaks shown.

5.5.2 Blinded Upper Limit

The pseudo-experiments generated in Figure 5.20 suggest that in an absence of signal, a $\mathcal{B}(A_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp) > 10\alpha_{cat}$ could definitely be excluded by the analysis, assuming the background shape extrapolated into the signal region behaves as is suggested from the sidebands. Any smaller branching fraction than this and it begins to become impossible to separate the signal peak from the background with sufficient significance. To systematically calculate this upper limit of $\mathcal{B}(A_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp)$ the branching fraction parameter space is scanned using the CL_s method, outlined in Ref. [155]. Specifically for this analysis, the process and framework created for the ongoing $A_b^0 \rightarrow \Lambda e^\mp \mu^\pm$ analysis is used, outlined in Ref. [42], with a similar final state and process of splitting the dataset into distinct analysis categories followed. As this analysis is currently blinded, estimated upper limits are calculated with the assumption that no signal is measured, using the sideband data as the input with the model defined in Equation 5.4.1. Initial upper limit results are stated in Table 5.8, using an asymptotic calculator implemented in HEPSTATS [136, 196]. A confidence level of 95% is reported.

The upper limit estimates differ from the single-event sensitivities as expected due to $N_{Signal} = 1$ unable to be discerned from the background. The CL_s method is a more realistic limit as it takes into account the background shape in the signal region which

has been shown to differ in each category. This is shown by the compatible limits for the two Run 2 categories, with fewer than 10 background events in the signal region and the relatively large $D^0 p \ell$ components restricting the background to remain mostly in the lower sideband. The Run 1 upper limit is a factor of ~ 3 greater than Run 2, a greater factor than was observed for the single-event sensitivities suggesting that the combinatorial dominated background shape, with compatible contributions in the lower sideband and signal region, has affected the sensitivity. Run 2 0γ provides a compatible limit to Run 2 1γ , as seen in the single-event sensitivities. Upon simultaneous determination of the upper limit in all three categories, a final $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp) \lesssim 2.8 \times 10^{-8}$ is derived. This is comparable to the projected prediction in Equation 1.3.2 [101].

This result will be the world-first limit on the $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ process and will provide a new constraint for future models in the b -baryon sector, which to date has no published LFV searches. The estimated limit is also competitive with other leading LFV results from searches in the decays of b -mesons in Figure 1.5, an indication of the significant Λ_b^0 sample only accessible by LHCb.

The results in Table 5.8 are estimated upper limits, using blinded analysis datasets and assuming no $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ signal is measured. They are potentially underestimated as the systematic uncertainties derived for the components of α_{cat} are yet to be propagated and no attempt has been made to calculate the uncertainty from the limit determination itself. This will take place after the signal region has been unblinded. However, considering the systematic effects already evaluated in Sections 4.7 and 5.3.3.3, the relative uncertainty is expected to remain $< 15\%$ and hence the analysis overall is statistically limited. The possibility remains that upon unblinding a signal is measured, providing hints, or the first evidence, for LFV. This would shift the analysis focus to quoting a measured branching fraction rather than an upper limit.

5.5.3 Current Status and Future of Analysis

The results presented are for a blinded version of the analysis. The general policy in LHCb experiment is that new analyses and searches, such as this search for $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$, remain blinded until approval by the corresponding collaboration working group and from the assigned review committee.

At time of writing the analysis has not entered working group review. A technical internal note is required before entering working group review and it is expected that all components of the analysis are near-finalised. For this analysis, the internal note is progressing and not expected to delay review. Since submission of this thesis, the analysis selection has been improved by replacing the tight semileptonic Λ_c^{*+} veto (described in

Section 4.1.6.2) with the D^0 veto (described in Section 4.1.6.3), in an attempt to remove the exclusive background in the fits described in Section 5.4 and 5.5. This has been a success, although it has required the reprocessing of many analysis components, and hence slightly delayed the analysis timeline.

Chapter 6

The LHCb Upgrade II

6.1 LHCb Upgrade II: A Flavour Physics Experiment in a High Luminosity Environment

Upon the completion of Run 2 of the LHC, Long-Shutdown 2 (LS2) began, during which the LHCb detector was significantly upgraded, referred to as LHCb Upgrade I (UI), with all subdetectors undergoing either a complete replacement or significant alterations to the readout and electronics. The changes to each subdetector are summarised in Table 6.1. Two major upgrades were replacing the VELO silicon strips with pixels and removing the L0 hardware trigger for a GPU-based HLT1. This is designed to process up to 5 TB/s for an increase to $\mathcal{L}_{inst} = 2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ with an expected Run 1–4 integrated luminosity of 50 fb^{-1} , a large increase in the LHCb data sample [23, 197, 198]. Run 3 began in 2022, and is currently expected to continue until the end of 2025.

The Run 3 VELO detector comprises of 26 two-module stations, with each module providing coverage for half a square either side of the LHC beamline, as seen in Figures 6.1a and 6.1c. Each module consists of four sensors, two on either side of the cooling substrate, with each sensor comprising of 768×256 square pixels of $55 \mu\text{m}$ side-length (referred to as pixel pitch from here on). Each sensor is read out by three VeloPix ASICs at 40 MHz, increasing the data rate of VELO UI to 1.2 Tbit/s, with the closest ASIC to the beamline reaching particle rates of 9 particles / event (see Figure 6.1b) [197]. The new Hamamatsu n-in-p sensors are reduced to $200 \mu\text{m}$ thickness with an increased radiation tolerance of $8 \times 10^{15} \text{ 1 MeV } n_{eq}/\text{cm}^2$, decreasing the material budget of the upgraded VELO [198]. In addition, the change in design means the closest sensor has an r_{min} of 5.1 mm which, with the enhanced spatial resolution from the small pitch, improves the IP resolution of long tracks to $(11 + 13.1/p_T) \mu\text{m}$, as seen in Figure 6.1d [197, 199].

Table 6.1: Summary of changes to each subdetector and other subsystems from the original LHCb detector operated during Run 1 and 2 of the LHC, through to the proposed LHCb Upgrade II (Run 5–6).

Sub-det.	Run 1–2	Run 3–4	Run 4–5
VELO	$r - \phi$ Silicon Strips	Silicon Pixel Detector	Pixel Detector with Timing
TT/UT	Microstrip Silicon Sensors	High Granularity Strips	Silicon Pixel Sensors
Tracker	Silicon-strip Inner Straw-tube Outer	Scint. Fibres (Sci-Fi)	Silicon Pixel Inner Scint. Fibre Outer
RICH	Ring-Imaging Cherenkov	Improved Optical Performance	Rad. Hard + Timing
ECAL	Shashlik	Upgraded Electronics	Rad. Hard Inner + Timing
HCAL	Iron-scint. sampling	Upgraded Electronics	Removed
Muon	5 MWPC	4 MWPC	Improved Inner region
Trigger	L0 hardware + CPU HLT1,2	GPU HLT1, CPU HLT2	GPU/FPGA HLT1,2

By the end of Run 4, many components of the UI detector will need to be replaced, and the long “data-doubling time” will render continued operation of this detector unattractive. Therefore LHCb Upgrade II (UII) has been proposed to take full advantage of the flavour-physics potential and opportunities at the HL-LHC, which will increase the peak luminosity by a factor of up to 7.5 [201]. LHCb UII is planned to be installed during Long-Shutdown 4 (LS4) and will feature major upgrades to all subdetectors enabling operation at up to $\mathcal{L} = 1.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. This corresponds to 50 fb^{-1} of data recorded per year bringing the projected full Run 1–6 estimated dataset to 350 fb^{-1} , as displayed in Figure 6.2 [201].

In this section, the HL-LHC motivation and corresponding LHCb physics potential will be summarised along with brief descriptions of the design changes to the tracking, PID and DAQ systems required to deliver the desired performance during Run 5 and 6.

6.1.1 High Luminosity LHC and LHCb UII Physics Potential

The HL-LHC upgrade will be completed during Long-Shutdown 3 (LS3) and will extend the LHC operation by a further 10 years beyond the end of LHC Run 3. The increased luminosity capability is delivered by:

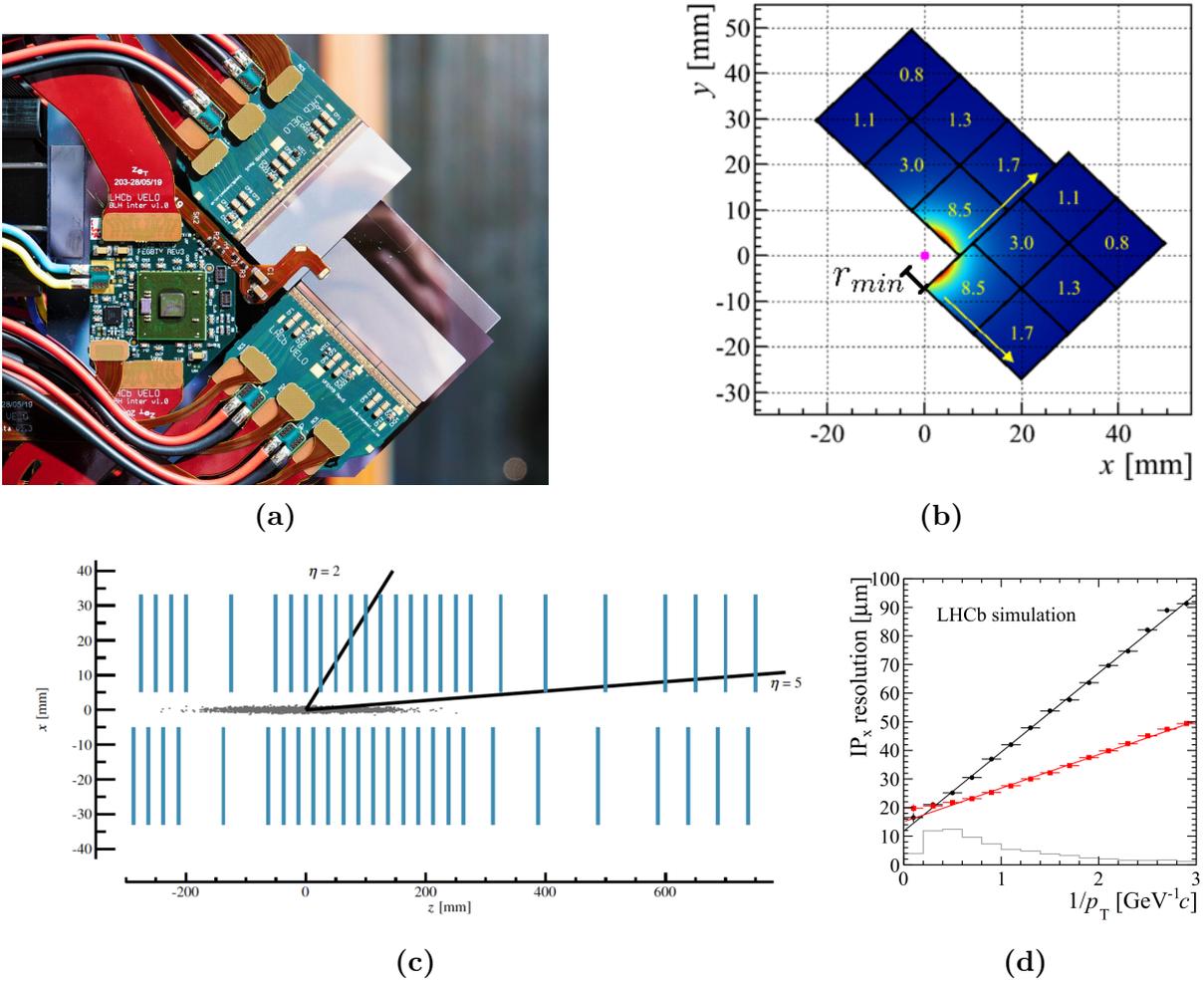


Figure 6.1: The LHCb Upgrade I VELO: (a) photo of a Run 3 VELO module, (b) the mean number of particles per ASIC per event for UI, (c) placement of 26 VELO UI stations in z and (d) the IP resolution performance of the original VELO (black) and UI VELO (red). (a) from [200], others taken from [197].

- the installation of more powerful superconducting quadrupole (focussing) magnets, with a peak field of 12 T compared to 8 T in the current LHC [201];
- use of new “Crab Cavities”, which provide particle bunches with transverse momentum immediately before the collision, reducing the area of overlap between two bunches [201];
- using high-temperature superconducting links that provide currents up to 100 kA to the new accelerator magnets without the need for ultra-low temperature cooling, operating at ≈ 50 K [201].

The installation of the improved accelerator coincides with the upgrades to the ATLAS and CMS detectors which will be able to take advantage of the expected peak

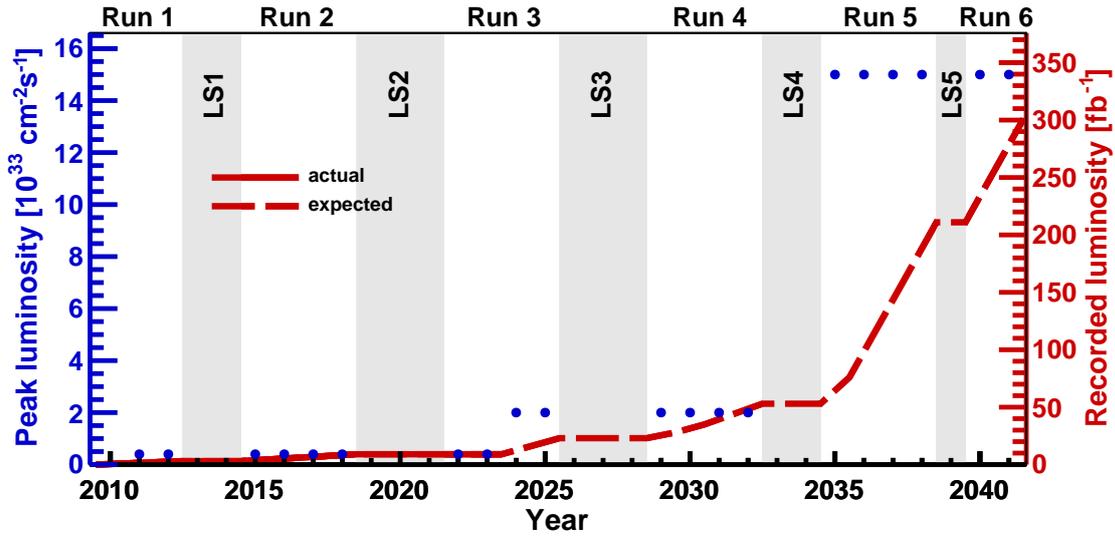


Figure 6.2: Measured and projected Run 1–6 instantaneous luminosities as well as the projected total dataset recorded by LHCb between 2011 and 2042 [24].

luminosity of $\mathcal{L} = 5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, corresponding to 140 pp collisions per bunch crossing, and a luminosity of up to 250 fb^{-1} per year, with the bunch-crossing rate remaining at 40 MHz [201].

Due to the fact that the shielding at IP8 is not as extensive as at IP1/5, LHCb UII will operate at a reduced luminosity with respect to A Toroidal Large Hadron Collider ApparatuS (ATLAS) and Compact Muon Solenoid (CMS), despite the removal of the levelling system employed for Run 1–4 of the LHC at LHCb [202]. The mean number of collisions per bunch crossing at $\mathcal{L} = 1.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ will reach $\mu > 40$, which poses problems for correctly reconstructing primary and secondary vertices. The higher particle rates also dramatically increase the operational fluence for detector components close to the collision point and beamline [24].

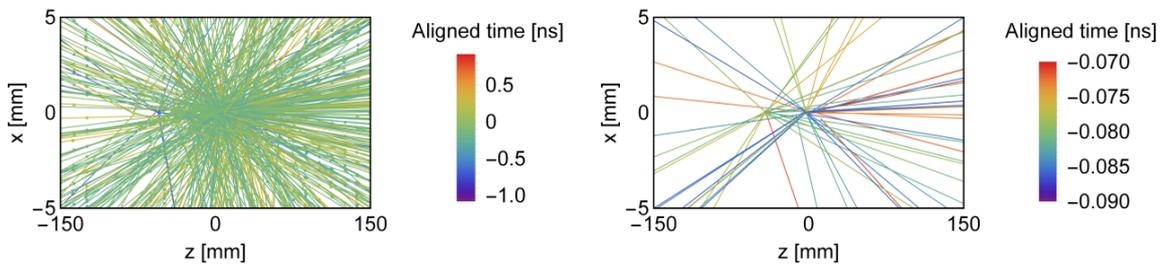


Figure 6.3: Demonstration of track density in the instrumented VELO region for different σ_t sensors in HL-LHC conditions. (left) Mock-up of all tracks within a bunch crossing of 2 ns. (right) Tracks within a 20 ps window, the target track-time resolution for VELO Upgrade II. Figure from [24].

The key tool to mitigate the challenges arising from operation at such high luminosities is the proposed addition of time resolution across LHCb, providing a fourth dimension to further split individual bunch crossings within their 25 ns windows. This improves the PV and track reconstruction as well as enhancing association with PID detector responses in the high track multiplicity environment. Temporal resolution will also enable continued building of multi-body candidates into high-quality secondary vertices, suppressing the large background from overlapping tracks with accurate PV association possible despite $\mu > 40$. The requirement for timing, especially for PV and SV reconstruction is visualised in Figure 6.3. If the discussed issues can be overcome, the increase in size of the LHCb dataset by a factor of seven will transform physics results across the entire LHCb physics programme, as illustrated by selected examples below.

Standard Model Benchmarks Searches for NP require precise experimental determination of SM benchmarks, to increase the significance of any deviation from the SM predicted value. LHCb UII’s unparalleled level of precision will allow the CKM triangle to be probed to small enough uncertainties that any contributing NP could be revealed, with the projected $\sigma_\gamma = 0.35^\circ$ in comparison to the LHCb Run 1 and 2 value of 4° [24, 203].

The increased dataset will also improve measurements of the lower-rate b -hadrons B_s^0 and A_b^0 as well as enabling precise measurements of the currently inaccessible hadrons B_c^+ , Σ_b and Ξ_b . The more performant subdetectors will also both directly and indirectly improve measurements, by increasing signal and background separation through, for example, reduction of material within the VELO, as well as lowering experimental systematic uncertainties by improving the PID accuracy with a new subdetector, Time-Of-internally-Reflected-CHerenkov-light (TORCH) [24].

New Physics in Rare Decays As discussed in Sections 1.2 and 1.3, rare decays that proceed through FCNC are sensitive to NP. The larger dataset size in UII will improve the precision of the measurements discussed, with the sensitivity of R_K and R_{K^*} projected to reach 0.7% and 0.9% respectively in comparison to the Run 1 and 2 measurements of 4.4% and 7.7% [24, 71, 204]. The improved ECAL will also significantly reduce the issues from electron reconstruction, while the improved tracking of the UII detector could render $R_{\mu\tau}$ measurements (see Equation 1.2.6) viable using $b \rightarrow s\tau^+\tau^-$ decays, which LHCb is not expected to be sensitive to throughout Runs 1–4. In addition $b \rightarrow d\ell\ell$ will be studied with greater precision after LHCb Upgrade II, providing new insight into the dynamics of FCNC decays, complementing analogous $b \rightarrow s\ell\ell$ measurements. LFV searches will be possible for a wider range of b -hadrons as well as more complicated final states, with the increased dataset enabling control modes to be closer in topology to signal modes

without affecting the precision of the search. The improved background rejection from the addition of timing will also aid in the “no-background” aim of LFV searches such as that introduced in Chapters 3–5 of this thesis.

A schematic view of the proposed LHCb Upgrade II is shown in Figure 6.4, with a similar overall envelope to the Run 1–4 detector, with the alterations summarised in Table 6.1 and changes to the Tracking and PID detectors discussed in the following sections.

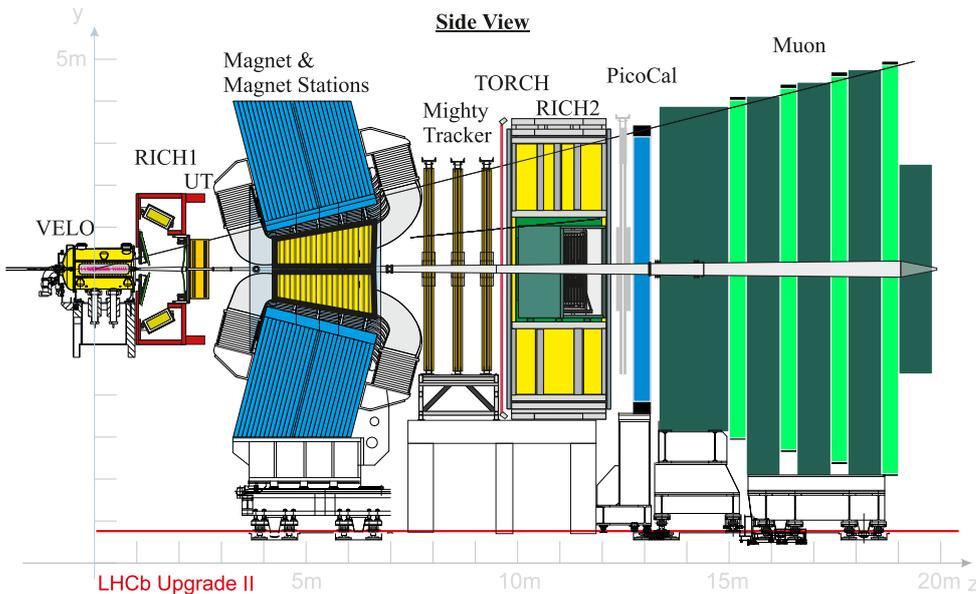


Figure 6.4: Schematic side-view of the proposed LHCb Upgrade II detector. Note the replacement of the HCAL with passive neutron shielding. Taken from [24]

6.1.2 Upgrades to Tracking Detectors

To deliver the physics aims for LHCb UII, the detector performance achieved with the UI detector must be (at least) maintained. To endure and benefit from the maximum luminosity of $\mathcal{L} = 1.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, with up to 2000 particles per BX within LHCb acceptance [24], the tracking system must be overhauled to increase granularity, improve radiation hardness and provide precise time resolution. This will enable LHCb to continue to achieve near-perfect tracking efficiency and the formation of high-quality candidates to deliver its physics goals. An overview of the changes with respect to Run 1–4 are laid out below.

The upgrade to the VELO will be based on similar high-granularity pixel technology to LHCb UI, with the sensors positioned a few millimetres from the LHC beams, providing

a precise first measurement for all tracks originating from PVs and most SVs. The changes are focused on providing a precise track timestamp, with a target of 20 ps, while maintaining radiation tolerance in the extreme environments, with fluence at the VELO UI inner radius (5.1 mm) projected to 8×10^{15} 1 MeV n_{eq}/cm^2 per year (equal to the UI tolerance for the entire eight year lifetime).

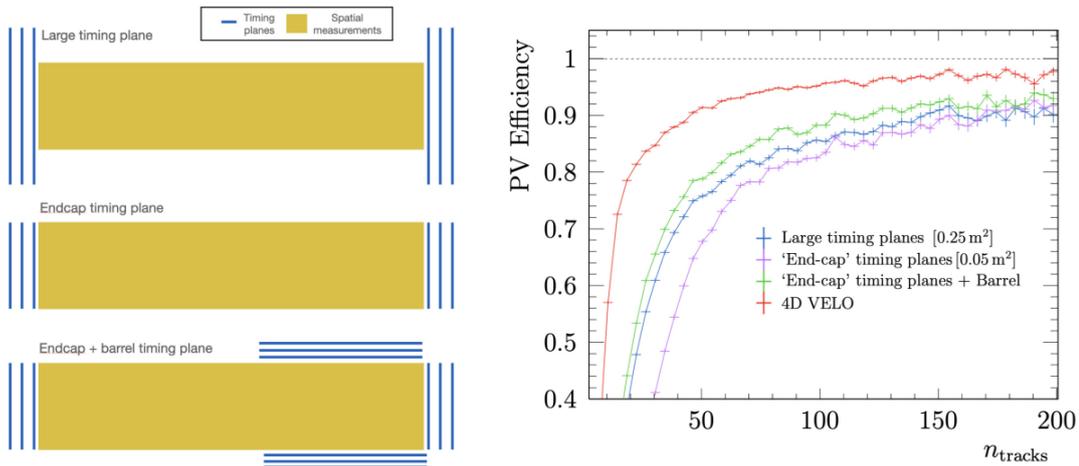


Figure 6.5: PV reconstruction performance of VELO Upgrade II for various 4D tracking scenarios. (left) Schematic of the three alternatives implemented. (right) PV reconstruction efficiency for each scenario. Taken from [24].

To provide a 20 ps track-time resolution, a 4D tracking approach is required, with all VELO stations equipped with timing capabilities and each sensor and readout ASIC providing a maximum resolution 50 ps per hit, assuming on average six VELO hits per track. Figure 6.5 displays the expected PV reconstruction efficiency for several alternative approaches using timing planes; the performance is degraded in all cases and are ruled out [24]. This highlights the significance of timing to precisely measure PVs position in UII, where on average 40 vertices must be reconstructed per bunch crossing. Without precise timestamps, overlapping tracks would be impossible to disentangle as illustrated in Figure 6.3. The timing performance of the VELO affects all other subdetectors, both for better separating individual tracks and the extremely precise t_0^{PV} timestamp as a result. Further justification for the timing requirement is discussed in Section 6.4 with the design of sensors further investigated in Section 6.2.1.

The UII VELO will be designed to withstand the extreme radiation environment in Run 5 and 6 by using improved radiation-hard sensor/ASIC technology as well as optimising r_{min} of the innermost sensors. This ensures the sensors can survive the entire Run 5 and 6 of the LHC, with a projected total fluence of $1-6 \times 10^{16}$ 1 MeV n_{eq}/cm^2 dependent on the final r_{min} [24]. The challenge this poses is developing technology that simultaneously provides the temporal and spatial resolution required while maintaining radiation

hardness, with the two naturally anticorrelated, as discussed further in Section 6.2.2.

The reduction in material for the UII VELO will reduce both the multiple scattering of incident particles and secondary particle production from material interactions within the VELO components. The primary contributor to material budget in the UI VELO is the 150 μm thick aluminium RF-shield, similar in design to the Run 1 and 2 shield, protecting the sensors from the LHC beam wake-fields and reducing the impedance experienced by the beams [24, 197]. Reducing the shield material will improve primarily the IP resolution, and will be delivered in UII by reducing the thickness of the RF-shield as well as a potential change in design; alternatives considered include a move from corrugations around each sensor to a uniform cylindrical shape with an ultra-low material budget. These design choices and their effect on performance will be explored in detail in Section 6.4.

The design of the main tracking system downstream of the VELO will follow that implemented for the UI detector, with performance improvements to cope with the increased luminosity. This consists of a redesigned Upstream Tracker of four pixel detector layers in the same location as the Run 1 and 2 TT and the Mighty Tracker, with three layers of a hybrid mix of silicon pixel detectors and scintillating-fibre mats. These will be installed in the location that the T1–3 stations occupied in Runs 1 and 2.

The Upstream Tracker (UT) will be upgraded from the new detector in UI (summarised in Table 6.1) and entirely instrumented with high granularity pixel sensors, with pitches as small as 50 μm considered feasible for the target Monolithic-Active-Pixel-Sensors (MAPS) technology [24]. This will improve the momentum resolution of long tracks and provide the first measurement of long-lived particles like Λ and K_S^0 that decay beyond the VELO acceptance [24].

The MT is designed with scintillating-fibre mat stations as were installed for LHCb UI, with the addition of high granularity silicon pixel sensors that cover the regions around the beam pipe with the highest particle densities. The MT stations will cover an area of 30 m^2 , with the hybrid design shown in Figure 6.6 optimised to instrument the area over which the LHCb magnet will sweep charged particles, retaining a close-to-uniform tracking efficiency across the MT area, rather than the estimated 50% efficiency in the central region without pixel instrumentation in UII conditions [24]. Mechanical constraints imposed by the HL-LHC beam pipe on the MT are expected to limit its acceptance to $\eta < 4.8$ [24], which has consequences for other subdetectors, expanded on in Section 6.4.

In combination with the VELO, the downstream detectors will provide a precise momentum measurement by the increased detector granularity and better track-segment

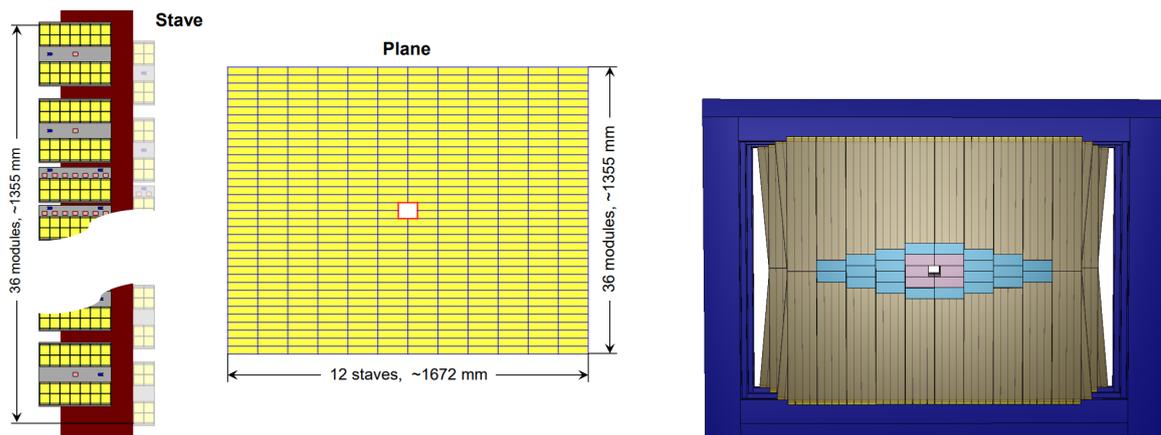


Figure 6.6: Schematic of (left) the upgrade to the UT detector and (right) one Mighty Tracker (MT) station from LHCb UII simulation geometry. For the MT station the central blue and pink regions are instrumented with high-granularity pixel detectors, and the outer regions using scintillating-fibre (Sci-Fi) mats. Taken from [24].

matching, which will reduce the proportion of fake tracks produced. The track matching mentioned will benefit from the precise timestamp of VELO tracks, with any segments overlapping in space becoming separable in time. A further option under consideration is the instrumentation of the internal surfaces of the magnet side-walls using scintillating bar detectors, the Magnet Stations, which will extend the LHCb acceptance to very-low momentum particles that would usually bend out of reach within the magnetic field [24].

6.1.3 Upgrades to Particle Identification Detectors and Data Processing

High quality particle identification is essential for flavour measurements, with the proposal for the UII PID subdetectors to maintain at-least UI performance [24]. This will be achieved through improvements in granularity as well as the ability to provide high precision timestamps to PID decisions, enhancing the matching of a track to its response in the PID detectors.

The upgraded RICH system for UII will continue to have two subdetectors in the same positions as Runs 1–4 of LHCb, with improved photo-detectors allowing Cherenkov photons to be separated into $\mathcal{O}(\text{ps})$ time-intervals [24]. This will enable identification of photons from individual PVs, aiding with the assignment of a PID response to tracks originating from a certain PV, with the discussed t_0^{PV} from the VELO enabling this.

The RICH will be complemented by a new Time-Of-internally-Reflected-Cherenkov-light (TORCH) detector that will allow low-momentum particle separation by utilising

totally-internally-reflected Cherenkov light to register and measure photons at high angles. This will enhance proton-kaon discrimination, especially below 10 GeV where current analyses often enforce a cut-off to reduce mis-identification [24, 76].

The calorimetry system will be overhauled by the removal of the HCAL and a significant redesign of the ECAL to continue identification in the high intensity central region around the LHC beam pipe. With the current ECAL this region would produce overlapping showers in UII conditions leading to degradation in energy resolution. This is prevented by the addition of fast-timing, aiding the separation of individual showers and improving the association of ECAL responses to individual PVs, in a similar way to discussed in the RICH.

The MUON detectors will be upgraded with a new technology to provide better granularity in the central region to match improvements in other subdetectors [24]. This is vital for many analyses, including the LFV analysis discussed in Chapters 3–5, where muons serve as an anchor in reconstruction and selection due to the precise momentum measurement and very low mis-identification rates [24].

At a peak luminosity of $\mathcal{L} = 1.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ LHCb is expected to produce 200 Tbit/s of data, which will need to be processed in real time while reducing the fraction of data written to storage. The trigger strategy will therefore have to switch from focusing on discarding bunch-crossings, to discarding individual PVs (pile-up suppression), by identifying and characterising each pp interaction in real-time [24]. The inclusion of fast-timing in many subdetectors will aid this, by allowing reconstructed objects to be associated with individual pp collisions and a high-quality track in the VELO or shower in the ECAL triggering an interval in time within which a PV should be reconstructed and higher level selections initiated. As with Run 3, the UII trigger system will have no hardware component implemented entirely on GPU or FPGA accelerators to utilise the parallel and expandable processing advantages vital for the high particle and data rate scenario [24].

6.2 Future Vertex Detector - Sensor Technology

With the VELO performing the first measurement for the majority of charged particles in LHCb, its design and characteristics are paramount to achieving the target performance of LHCb Upgrade II, with the pixel-based sensors the fundamental technology at the core of this. If a particle traversing a VELO station cannot be precisely measured in space and time, the corresponding track will be degraded in resolution, subsequently impacting reconstruction of physics candidates, damaging the overall physics performance of the

experiment.

In this section, the requirements of a 4D pixel sensor for VELO UII are presented, along with an overview of the feasible sensor technology options with an experimental investigation of one option, Low-Gain Avalanche Detectors (LGAD), presented. The results include the sensor performance before and after irradiation, to emulate the ageing effects expected over the course of Run 5 and 6 of the HL-LHC.

6.2.1 Characteristics for an Upgraded VELO Pixel Sensor

As discussed in Section 2.2.2, the performance of the VELO is critical because it propagates to many quantities used in physics analyses including primary and secondary vertex positioning, flight distances between vertices and measurements of association between vertices and tracks.

6.2.1.1 Spatial and IP Resolution Requirements

The track Impact Parameter resolution is often used as a single figure of merit because it encompasses the complete performance of the subdetector. Defined in Equation 2.2.1, σ_{extrap} is approximately proportional to σ_{xy} , the per hit spatial resolution in the x or y direction (equal for a sensor with square pixels) and σ_{MSC} is dictated by the material before the second hit on a track. σ_{xy} is defined as the uncertainty on the measurement of the x or y position of a particle traversing the sensor with respect to the true position. For a pixel detector, taking the case where a particle only traverses one pixel, the resolution is related to the pixel pitch p_{xy} , e.g. [205], by

$$\sigma_{xy} = \frac{p_{xy}}{\sqrt{12}}, \quad (6.2.1)$$

In reality, particles often produce multi-pixel clusters in silicon sensors, with the diffusion of charge deposited and the possible traversal of multiple pixels. This is known as charge sharing with improvements to σ_{xy} possible if exploited. Considering diffusion only, if a particle traverses a sensor close to the pixel-edge a sufficient amount of free electrons/holes (e/h) may enter the neighbouring pixel to register a second hit. If the clustering algorithm is able to associate these pixel hits, the average of the two pixels centres can already reduce the residual between the true and measured positions. If this is taken a step further, with the amount of charge in each pixel available for clustering, charge-weighting can further reduce the residual. The more pixels that are triggered by a particle traversing the sensor, the more information available and hence the better the σ_{xy} that can be obtained.

This also applies to particles entering the silicon at a non-zero incident angle (relative to the normal) depositing charge and triggering multiple pixels, with simulation studies in Figure 6.7a for the Run 3 VELO displaying the potential improvement in σ_x [197]. At large θ_x the charge is spread over many pixels, with some not meeting the charge threshold, degrading the resolution.

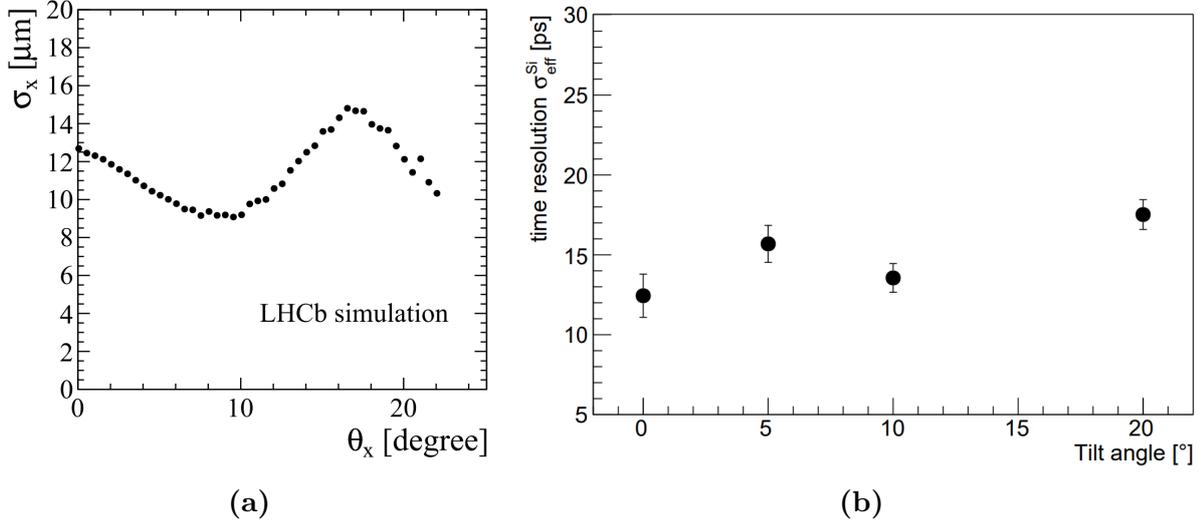


Figure 6.7: Performance of silicon sensors as a function of angle. (a) Spatial Resolution in x as a function of one-dimensional incident particle angle θ_x (with tracks restricted to $\theta_y < 2$) for the UI VELO, taken from [197]. The average σ_{xy} was determined to be $\approx 12 \mu\text{m}$ across all incident tracks. (b) Time resolution as a function of tilt angle (analogous to incident angle) for a 3D trench single pixel, taken from [206].

For UII however, it is an open question of how many pixels can be read out per particle traversal, with the increased luminosity leading to the potential for data rates to become unmanageable [24]. One considered solution would be performing clustering on the ASIC itself, reducing the readout size. A study of particle and pixel hit rates for VELO UII will be discussed in Section 6.4.2.3, as a precursor to a future clustering investigation.

The target IP resolution for VELO Upgrade II is $(12 + 14/p_T) \mu\text{m}$, similar to the UI target, but with the anticipated increase in r_{min} for the UII VELO design, there is a requirement for σ_{xy} to be reduced to compensate [24]. An improved σ_{xy} would also aid with the increase in luminosity with smaller pixels resulting in an overall lower occupancy per pixel, reducing hit inefficiencies from “Front-End Pileup”, when a particle passes through a pixel within the discharge time of the previous pixel activation (about 42 ns for a pixel silicon sensor [205]). Therefore investigation of new sensor technologies that can improve the VELO hit resolution, through either reducing p_{xy} of the sensors or by charge sharing/weighting methods, is required to ensure the design requirements are feasible at the extreme particle rates of the HL-LHC.

6.2.1.2 Temporal Resolution Requirements

In addition to improvements in spatial resolution, precise temporal resolution is required for the UII VELO sensor, with a 4D tracking approach discussed in Section 6.1.2 as the only feasible option, with a per hit $\sigma_t^{Hit} \leq 50$ ps. σ_t^{Hit} is a combination of both the sensor and ASIC response hence, assuming equal contributions, the VELO pixel sensor technology target is $\sigma_t^{Pix} \approx 35$ ps. This σ_t^{Pix} requirement must be achieved while retaining the σ_{xy} , hit efficiency and radiation tolerance discussed. For example, Figure 6.7 suggests that the preferred particle incident angle for spatial resolution may lead to a worsened time resolution and vice-versa, due to an angled track traversing multiple pixels and damaging σ_t by sharing the charge deposited between pixels. Scenarios for VELO Upgrade II where σ_t is degraded are presented in Sections 6.4.2.2 and 6.4.2.3, displaying the cost in performance from reduced timing ability.

6.2.1.3 Radiation Tolerance, Efficiency and Material Requirements

The VELO UII sensors must withstand unprecedented radiation doses from the increased luminosity while maintaining an acceptable efficiency across the lifetime of Runs 5 and 6. If the r_{min} was to be maintained at 5.1 mm (as is for the Run 3 VELO), the UII expected fluence reaches $\approx 6 \times 10^{16} \text{ 1 MeV } n_{eq}/\text{cm}^2$ [24], which is unfeasible for any existing technology. But increasing this radius will lead to a degradation in performance. Hence it is expected that R&D efforts will dictate the maximum fluence for a sensor technology, with the aim to minimise the corresponding r_{min} through optimisation in simulation, presented in Section 6.4 where the value of r_{min} is shown to be crucial for LHCb UII global performance.

To reconstruct a VELO track, at least three hits are required, needing a hit efficiency of $\geq 99\%$ [197]. Any value lower than this, or any degradation in hit efficiency over the lifetime of a sensor due to radiation damage, would have to be mitigated with more stations, which would increase the overall cost and material budget of the VELO. The sensor technologies discussed in Section 6.2.2 balance this requirement with the spatial, temporal and fluence restrictions to identify an optimum solution for a new 4D VELO module design.

6.2.2 Sensor Technology Options

Three sensor technologies under consideration for VELO Upgrade II are briefly outlined in the following section, with examples of their designs shown in Figure 6.8. The expectation

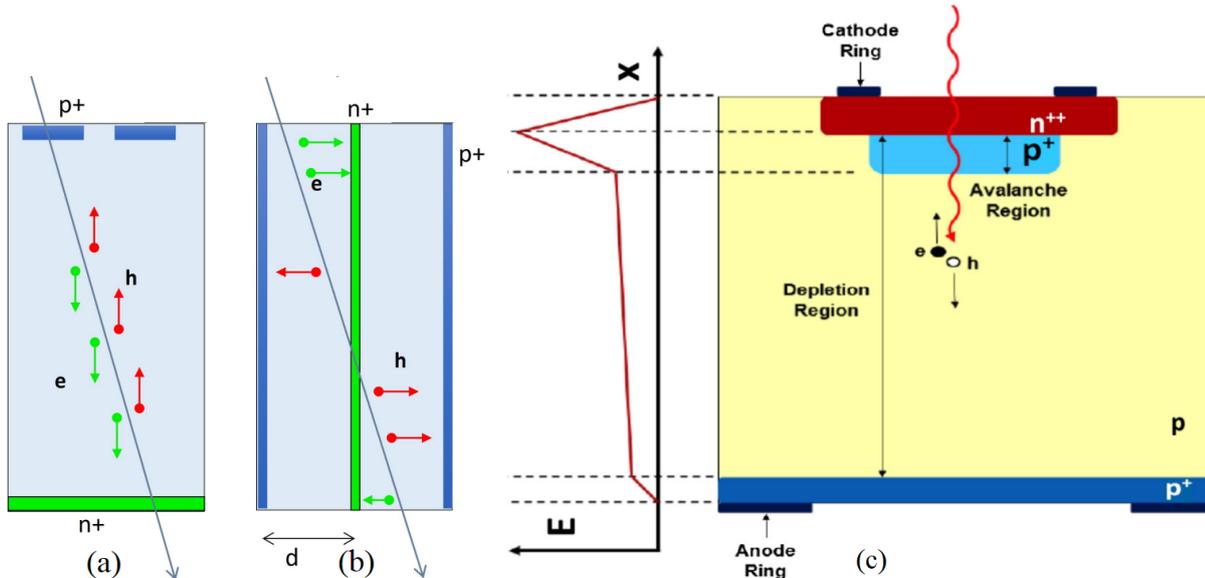


Figure 6.8: Illustrative diagrams of (a) a planar sensor, (b) a 3D column sensor and (c) an LGAD, comparing the charge collection process in each, with electron-hole pairs labelled along with position of the different electrodes. For the LGAD diagram, the electric field in different regions is described. (a) and (b) are taken from [206] and (c) from [207].

is that silicon sensors will be used, as in the Run 3 VELO. An alternative with significantly better radiation hardness would be diamond sensors, but the cost and relative immaturity of the technology render these unfeasible [208].

6.2.2.1 Planar Sensors

One option is to continue to use planar sensors with a focus on achieving ultra-fast timing by minimising sensor thickness, yielding shorter charge collection times, improving σ_t^{Pix} . This however reduces charge in the sensor, requiring very low noise in the readout to achieve an acceptable signal-to-noise ratio, expected to be difficult to maintain with the radiation damage the sensors would incur. Some planar silicon sensors have been proven to operate up to $1.6 \times 10^{17} \text{ 1 MeV } n_{eq}/\text{cm}^2$ [209], but the observed significant drop-off in charge collection efficiency above $1 \times 10^{16} \text{ 1 MeV } n_{eq}/\text{cm}^2$ [210] becomes problematic when thinning the sensors. To mitigate this, operation at very low thresholds and high bias voltages would be required, requiring ultra-low operating temperatures to reduce the leakage current and prevent any thermal runaway that leads to breakdown, a potentially catastrophic process discussed further in Section 6.3. Operating below the UI temperature of -20°C would pose challenges in itself, increasing costs [24].

6.2.2.2 3D Sensors

A new technology considered are 3D sensors, with the main advantage their radiation resistance, with consistent performance up to 3×10^{16} 1 MeV $n_{\text{eq}}/\text{cm}^2$ already proven [211], and the potential for further R&D to push this beyond the maximum fluence considered for VELO Upgrade II [212]. This is delivered while retaining the fast timing performance, with as low as $\sigma_t^{\text{Hit}} = 20$ ps proven, with an estimated $\sigma_t^{\text{Pix}} < 15$ ps contribution [213]. These characteristics are achieved by altering the position and plane of the electrodes within the silicon, moving one electrode to the side of the sensors and implementing a trench or column through the middle for the other, compared to planar detectors with electrodes parallel to the sensor plane (see Figures 6.8a and b). This reduces the inter-electrode distance for $e - h$ pairs to traverse leading to faster signal collection and hence better time resolution, with sensor thicknesses remaining at 150–200 μm .

At high fluences the major limitation of silicon sensors is the trapping of charge carriers, inhibiting particle detection by radiation-induced defects along the drift-path to the collecting electrode [211]. Therefore short inter-electrode distances are also the source of the improved radiation hardness in 3D sensors. The main drawback, apart from the increased manufacturing costs, is the insensitive volume introduced by the trench/columns, with geometrical efficiency reduced to 90% in some designs [211, 213]. This will reduce the hit efficiency of the sensor, but studies have shown that increasing the average incident angle for particles by tilting the sensors can mitigate this [211]. Nevertheless, the 99% target is likely to be unachievable, with additional VELO stations and optimisation of their z positions required to retain the target tracking efficiency and angular acceptance of the VELO. As discussed, this would introduce more material and increase particle scattering but the fast σ_t^{Pix} and proven radiation hardness make 3D sensors a strong candidate for the VELO Upgrade II sensor technology.

6.2.2.3 Low-Gain Avalanche Detectors

LGADs are designed to provide good timing performance by inserting an additional charge multiplication layer into a planar n-in-p sensor. Figure 6.8c shows the thin “gain” layer (“p⁺ Avalanche Region”) that creates a high E -field region initiating charge multiplication from the initial charge deposited by a traversing particle [207, 214–216]. The high E -field amplifies this effect to form a fast “avalanche”, which is collected quickly and then dissipates, resulting in a precise time resolution while the intrinsic low gain retaining a high signal-to-noise ratio.

The small initial charge required to produce an avalanche enables the operation of

very thin LGADs, with sensors fabricated to $35\ \mu\text{m}$ [217]. This improves the σ_t by reducing the drift time for the avalanche, with the latest R&D showing $\sigma_t^{Pix} \approx 30\ \text{ps}$ and $25\ \text{ps}$ for $50\ \mu\text{m}$ and $35\ \mu\text{m}$ sensors respectively [217]. The primary drawback of LGADs is the degradation of performance at moderate fluences compared to other sensor technologies. Current studies show that above $\approx 2 \times 10^{15}\ 1\ \text{MeV}\ n_{\text{eq}}/\text{cm}^2$ acceptor removal can completely degrade the gain layer to the point of failure [207]. However, the relatively low-cost to manufacture LGADs (compared to 3D sensors), makes them an interesting subject for R&D with novel designs being investigated to improve the radiation hardness. Results from irradiation tests with LGADs from a new manufacturer (Micron Semiconductor Ltd.) will be presented in Section 6.3; further R&D investigating the use of more radiation-hard dopants such as carbon is also in progress at other foundries [218].

6.3 Feasibility Investigation of Micron Semiconductor LGADs

In 2017, Micron Semiconductor Ltd. began fabricating LGADs in collaboration with the University of Glasgow as a candidate sensor for the in-development MediPix chip [215, 219]. With the clear preference for a 4D VELO for Upgrade II of LHCb, and the existing collaboration, an LHCb-UK focused R&D effort began between the University of Birmingham and Glasgow to characterise and test Micron LGADs to determine plausibility as a sensor technology for VELO Upgrade II. These efforts were mirrored across the LHCb collaboration with similar R&D work ongoing with other foundries including FBK, CNM, Teledyne-e2v and Hamamatsu [207, 214, 217, 220].

In the following sections, characterisation of a collection of thin ($50\ \mu\text{m}$) Micron LGADs of different designs will be presented, with results shown before and after irradiation followed by a discussion on the suitability of these sensors for VELO Upgrade II and what changes would be required for a production sensor. The testing was performed in collaboration with colleagues from the BILPA laboratory, at Birmingham, who had designed experimental apparatus for testing Teledyne-e2v LGADs [214]. Results for Current-Voltage (IV), Capacitance-Voltage (CV) and gain measurements performed at Birmingham will be presented, as well as a review of the timing ability of these sensors (performed at Glasgow) [216].

Table 6.2: Characteristics of the different sensors with results presented in this section, a subset of the overall Birmingham LGAD inventory.

Sensor	Normalised Gain Doping	Pixel Pitch (μm)	JTE Width (μm)
B	1.0	500	10
D	1.08	1000	20
F	1.08	500	10
H	1.08	1000	10
PiN	N/A	1000	10

6.3.1 LGAD Design, Experimental Set-up and Irradiation

Figure 6.8c shows the schematic design for a standard LGAD, with each pixel containing an n and p-type electrode and highly-doped gain layer. The main design choices for an LGAD are thickness, pixel size, the concentration of dopants in the gain layer and the width of the Junction Terminating Extension (JTE). The JTE is a region at the edge of the gain layer which modifies the electric field to prevent extremely high fields forming that could result in the early breakdown of sensors [221]; this must be prevented for a robust sensor technology, expanded on in Section 6.3.2. The designs of the sensors tested during the R&D work are summarised in Table 6.2. The Positive-intrinsic-Negative (PiN) diode is fabricated with the same process as the LGADs but without the gain layer so provides a “no gain” reference point for measurements.

An LGAD is made operational by applying a “reverse bias” voltage (V_R) to the backside of the sensor (labelled p^+ anode in Figure 6.8c), reducing the amount of free charge carriers until the active region is entirely depleted. Once depleted, a particle traversing the LGAD will deposit energy in the silicon and produce $e - h$ pairs, which separate and drift to be read out as a hit in the sensor. Potential sensor technologies should be investigated in environments emulating the experimental conditions for a silicon pixel detector, with cold operation $\leq -20^\circ\text{C}$ and appropriate reverse bias voltages. Hence the experimental setups used for characterisation are a Climate Chamber and a Transient Current Technique (TCT) laser station with climate control adaptations [214].

The temperature of the Climate Chamber can be varied between -40°C and $> 40^\circ\text{C}$ with the humidity kept stable, allowing measurements to be made in controlled conditions. Individual “floating” LGADs are tested in the climate chamber after being fixed to a PCB and a wire-bond connecting the top electrode to readout pads. The backside and top of the sensor are then connected in a series circuit to a high-precision, high-voltage power-supply/multi-meter that provides a reverse bias to the backside of the sensor and can read out current from the topside. IV and CV measurements are performed in this way,

by increasing V_R and recording the corresponding measurement change.

The TCT set-up is used for Gain measurements, by directing an infra-red laser through the sensor and recording the LGAD response to the $e - h$ pairs produced by the high-intensity photon beam [214]. Similarly this is recorded as function of V_R , with water-cooled Peltiers, nitrogen flushing and a sealed chamber allowing temperature and humidity control of the measurements, with stable temperatures of $\approx -20^\circ\text{C}$ possible [214].

To emulate radiation damage and annealing over multiple years of operating the VELO in HL-LHC conditions, with large fluences from pp collisions followed by long periods of time in cold and warm conditions, the LGADs were irradiated after initial characterisation in the University of Birmingham’s in-house MC40 Cyclotron facility [222, 223]. This produces protons up to 40 MeV and delivers doses of $\approx 1 \times 10^{16}$ 1 MeV $n_{\text{eq}}/\text{cm}^2$ within 120 minutes.

While the expected radiation dose of the VELO Upgrade II sensors closest to PV production could reach a design-dependent $\approx 6 \times 10^{16}$ 1 MeV $n_{\text{eq}}/\text{cm}^2$, previous R&D observed major damage to LGAD performance above 2×10^{15} 1 MeV $n_{\text{eq}}/\text{cm}^2$ [207]. Therefore, it was decided to lower the initial dose for the Micron LGADs to $\approx 9 \times 10^{14}$ 1 MeV $n_{\text{eq}}/\text{cm}^2$ to study and understand how performance of the sensor would change during the course of a single data taking year during Run 5, especially as this work presents the first post-irradiation tests of Micron LGADs.

The irradiation is followed by an intentional scheme of annealing by placing the sensors in a precisely temperature-controlled oven at 60°C for 40 minutes. This attempts to emulate the gradual annealing of a sensor over the lifetime of a detector [205]. The sensors are subsequently stored and tested at $\leq -20^\circ\text{C}$ to reduce the rate of further annealing. For a longer schedule of testing, these devices would be further irradiated and re-tested, but as discussed in the following sections a single dose of radiation was sufficient to draw conclusions with respect to radiation hardness.

6.3.2 Current-Voltage (IV) Characterisation

Within a silicon sensor that has been reverse biased, free charge carriers can be thermally generated at the surface or within the volume of the depleted sensor. This results in a current being able to flow, even after full depletion. This is usually negligible at low V_R and can be further reduced by lowering the operating temperature. As the bias voltage increases, the E -field within the sensor will increase, breaking covalent bonds and increasing the rate of $e - h$ pairs “leaking” out of the device, increasing this leakage current I_l . For a sufficient V_R , the E -field will accelerate thermally-produced free charges

to an energy at which further $e - h$ pairs are produced, leading to uncontrolled charge-multiplication and an avalanche where the current begins to increase according to a power law as a function of bias voltage, following the equation

$$\frac{dI_l}{dV_R} = \kappa \frac{I_l}{V_R}, \Rightarrow I_l(V) \propto V_R^\kappa, \quad (6.3.1)$$

where κ is a factor which numerically categorises the breakdown. This uncontrolled avalanche, unlike that induced by the passage of a charged particle, can result in breakdown of the sensor. If thermal runaway is not arrested, with the internal sensor temperature increasing and further amplifying charge, a catastrophic breakdown can occur damaging the sensor. Therefore, IV measurements, where the leakage current is measured as a function of V_R are required to understand the safe operating range for each sensor.

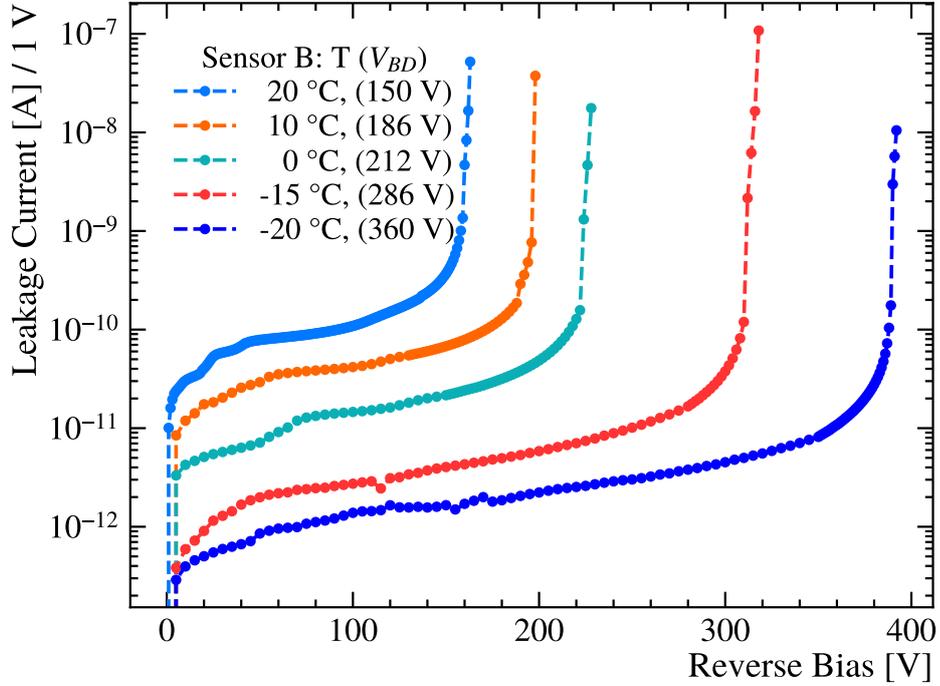


Figure 6.9: IV measurements for sensor B performed at a range of temperatures, before irradiation. The measurement at each V_R was repeated five times and averaged, with the error on the mean negligible. A relative humidity of $< 6\%$ was maintained for each measurement using a nitrogen flush. The breakdown voltages were derived at $\kappa = 10$ with the measurements halted when I_l enters the μA range.

Figure 6.9 presents a series of IV measurements for sensor B at a range of temperatures before irradiation. All measurements show sensor breakdown, where the current begins to rapidly increase for small increments in voltage.

The results show that sensor B breaks down at ≈ 150 V at 20°C , with the breakdown

voltage increasing as the operation temperature reduces, to ≈ 360 V at -20°C . While this trend is expected for a simple silicon detector, with the lower temperature reducing leakage current, other studies of LGADs at low temperatures have displayed the inverse trend [220]. These have concluded that, as the intrinsic gain of an LGAD increases at lower temperatures, there is more charge multiplication within the sensor leading to breakdown at lower V_R , despite the reduced leakage current [220]. For the Micron LGADs, the trend observed implies that the impact of temperature on absolute leakage current out-weighs the increase in gain for these very low gain devices. This is potentially a desirable feature implying that reducing the detector temperature could allow stable operation at higher V_R without inducing breakdown in the sensor [224].

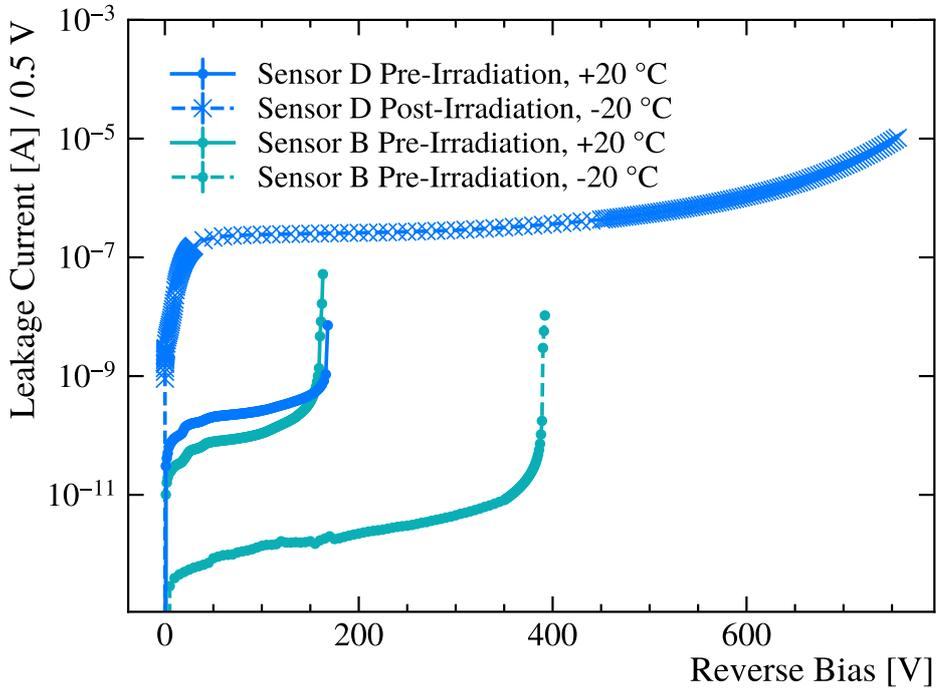


Figure 6.10: IV measurements for sensors B and D, showing sensor D after irradiation to 9×10^{14} 1 MeV $n_{\text{eq}}/\text{cm}^2$. An IV measurement of sensor D at $T = -20^\circ\text{C}$ was not performed prior to irradiation, hence two results for sensor B are included to infer the expected before irradiation curve for sensor D at $T = -20^\circ\text{C}$. The result for post-irradiation includes a scheme of annealing.

Operating at high V_R is preferred due to the increased gain and hence time resolution of LGADs. However, long-term stable operation is the priority in the case of the VELO detector, therefore characterisation of potential VELO sensors at different points in their lifetime is required to understand expected performance. IV measurements were hence repeated after irradiation in the cyclotron, with the result for sensor D shown in Figure 6.10. Radiation damage can be seen to dramatically change the sensor behaviour, with bulk damage providing new generation-recombination centres for electron-hole pairs,

increasing the overall leakage current [205]. Despite the increased leakage current, the sensor does not show clear signs of breakdown as in the pre-irradiation sensors, with tests up to $V_R = 750$ V only showing a small increase in dI/dV . This implies that the radiation damage has severely impaired the LGAD's ability to generate an avalanche, pointing towards damage to the gain layer, assumed to be due to acceptor removal.

IV measurements at different fluences would inform the values of V_R at which sensors should be operated to achieve stable performance throughout the sensor lifetime. The measurement discussed implies operation up to $V_R = 600$ V is feasible after a fluence of 9×10^{14} 1 MeV $n_{\text{eq}}/\text{cm}^2$. While this bias voltage is possible, any higher could lead to operational difficulties including increased power dissipation requiring better cooling solutions. In addition, Single-Event-Burnout has been observed in LGADs at high V_R [225], where the extreme potential difference results in rapid charge generation from a single particle traversing the sensor, causing irreversible damage. This has been measured to be possible for electric fields above $12 \text{ V } \mu\text{m}^{-1}$ [225], 600 V for these 50 μm Micron sensors.

6.3.3 Capacitance-Voltage (CV) Characterisation

In normal operation, only the charge produced in the depleted region of a silicon sensor can be detected. Sensors are therefore operated at a bias above the ‘‘Full-Depletion voltage’’ (V_{FD}) where the depleted region extends over the whole sensor [205]. Further increasing V_R beyond this increases the field by $(V_R - V_{FD})/d$ where d is the thickness of the sensor [205]. Silicon sensors hence must be characterised to record V_{FD} as a suitable minimum operating voltage and to study the doping concentration distribution through a sensor. This is performed by Capacitance-Voltage (CV) measurements with the expected behaviour of

$$C(V) \approx \begin{cases} \sqrt{\frac{\varepsilon_0 \varepsilon_{Si} N_D}{2V}} & \text{for } V < V_{FD}, \\ \frac{\varepsilon_0 \varepsilon_{Si} A}{d} & \text{for } V > V_{FD}. \end{cases} \quad (6.3.2)$$

derived from treating silicon sensors as parallel plate capacitors separated by the depletion width, d . $\varepsilon_0 \varepsilon_{Si}$ is the dielectric-constant in silicon and N_D is the dopant concentration in the sensor, producing a linear trend below V_{FD} and a uniform plateau above this threshold for $C^{-2}(V)$.

For LGADs there is an additional ‘‘Gain Depletion Voltage’’ (V_{GD}), the reverse bias at which the highly-doped P-type multiplication layer becomes depleted. This introduces an additional vertex in the CV curve as shown in the studies of sensor D in Figure 6.11, with, for the pre-irradiation curve, the gradual depletion of the gain layer as V_R increases up to $V_{GD} \approx 25$ V before the rapid depletion of the remaining sensor to $V_{FD} \approx 35$ V. This

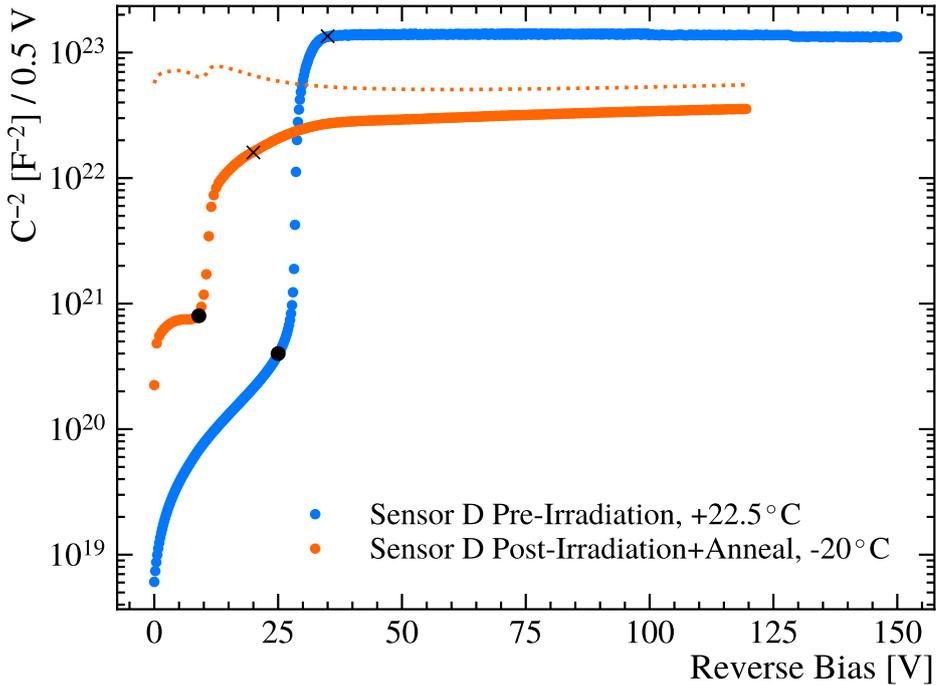


Figure 6.11: CV measurements for sensor D before and after irradiation to $9 \times 10^{14} \text{ 1 MeV } n_{\text{eq}}/\text{cm}^2$. The measurements were recorded with a low frequency of 1 kHz using an LCR meter [226]. The V_{GD} and V_{FD} on each curve is marked with a dot and cross respectively. The orange dashed curve shows a CV measurement post-irradiation with an LCR frequency of 10 kHz.

implies that this LGAD should be operated at $V_R > 35 \text{ V}$. The straight line segments between the depletion voltages imply that the concentration of charge is uniform within individual layers of the sensor, but that N_D is orders of magnitude higher in the gain layer, as shown by the reduced line gradient.

Post-irradiation, the concentration of free-charge carriers is reduced to the extent that the contribution to the overall capacitance of the sensor is minimal, requiring measurements at low temperature and frequency to reduce noise. The post-irradiation measurement for sensor D is shown in Figure 6.11, with significantly different absolute capacitance to pre-irradiation. This is interpreted as being due to a frequency dependence of the capacitance that has been observed in irradiated silicon detectors [226], where radiation-induced change to resistivity alters the simple relationships in Equation 6.3.2. The post-irradiation CV shown in Figure 6.11 was performed at 1 kHz, with other frequencies tested returning capacitance orders of magnitude different (as shown). However, the post-irradiation extracted values of $V_{GD} \approx 9 \text{ V}$ and $V_{FD} \approx 20 \text{ V}$ should remain correct [226] and are consistent across measurement frequencies. This implies that irradiation to $9 \times 10^{14} \text{ 1 MeV } n_{\text{eq}}/\text{cm}^2$ has significantly reduced the dopant concentration in the sensor,

especially in the gain layer where the gradient of the line for $V < V_{GD}$ is significantly larger after irradiation. The difference ($V_{FD} - V_{GD}$) remains roughly constant implying limited damage in the bulk of the sensor, but the reduction to V_{GD} indicates major impact on the ability to read out and provide precise time measurements.

To build from this, further CV measurements would be performed at a range of fluences, to understand the trend of V_{GD} and V_{FD} with increasing fluence as well as extracting reliable absolute capacitance measurements by applying a frequency-based correction to the CV curves. This is important for determining noise in the measurement amplifiers and to calculate the contribution to capacitance from the silicon pixel sensor itself, in comparison to the readout electronics and neighbouring pixels [205].

6.3.4 Gain Characterisation

LGADs are devices with an intrinsic gain, G_L . This dictates how charge deposited in the device is amplified by multiplication of charge carriers and can hence be summarised as

$$Q_L = G_L q, \quad (6.3.3)$$

where q is the charge deposited by the traversing particle and Q_L the total charge in the sensor. For an LGAD, the larger the intrinsic gain, the greater the electric field magnitude, the larger the avalanche produced by a hit and therefore an expected improved timing performance from the greater signal-to-noise ratio.

LGADs are designed to have a lower gain ($\lesssim 10$) than other silicon devices (such as Avalanche Photo-Diodes), to reduce noise and acute temperature dependence [221]. Understanding the performance as a function of gain is however important to ensure that the device is tuned correctly during operation, with an intrinsic dependence on gain with bias voltage due to the E -field dependence on V_R .

To understand this trend for the Micron LGADs, measurements of individual devices gain were performed as a function of V_R using the TCT technique. The infra-red laser deposits a column of charge in the sensor, producing a characteristic waveform that is read out by an oscilloscope with an integral of S_L , where

$$S_L = A Q_L = q A G_L, \quad (6.3.4)$$

with A a constant amplification factor from the equipment. A is difficult to measure for an unknown gain, but by repeating each measurement using a PiN diode (where $G_P = 1$)

and measuring the integral of the waveform S_P , Equation 6.3.4 can be normalised to

$$G_L = \frac{S_L}{S_P}, \quad (6.3.5)$$

with the advantage of cancellation of potential systematic effects from the measurement of the integrals.

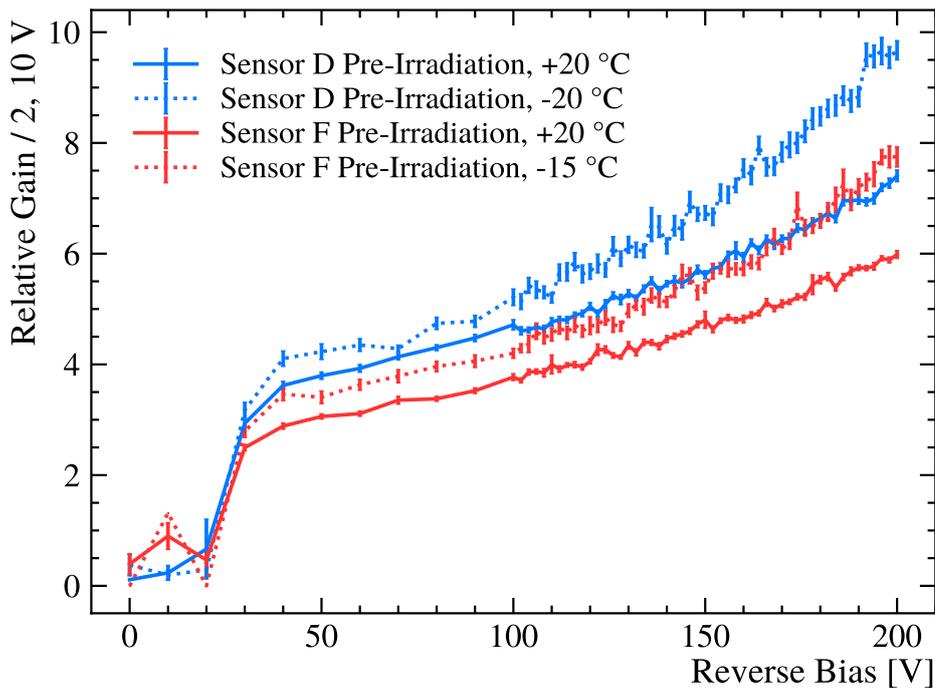


Figure 6.12: Gain measurements for sensor D and F comparing performance at room temperature and cold. The measurements are normalised to PiN measurements at the corresponding temperatures.

Measurements of gain are performed for each LGAD, up to a V_R determined from corresponding IV measurements, with different temperatures also tested and PiN measurements performed at each corresponding environment to ensure correct normalisation. Gain-voltage curves for sensors D and F are shown in Figure 6.12. The overall greater gain of sensor D is due to the four times larger active area, despite the larger JTE, with neither sensor exceeding a gain of 10 up to the V_R tested. The clear trend of increased gain for lower temperature agrees with the general trend for silicon devices with gain, where reduced temperature reduces the mean-free path of charge carriers, increasing charge multiplication within the sensor [220].

The device gain also correlates to the time resolution performance, hence measurements like those shown allow calibration of devices while operating a detector. If many measurements are performed in different conditions, slight changes in operation temper-

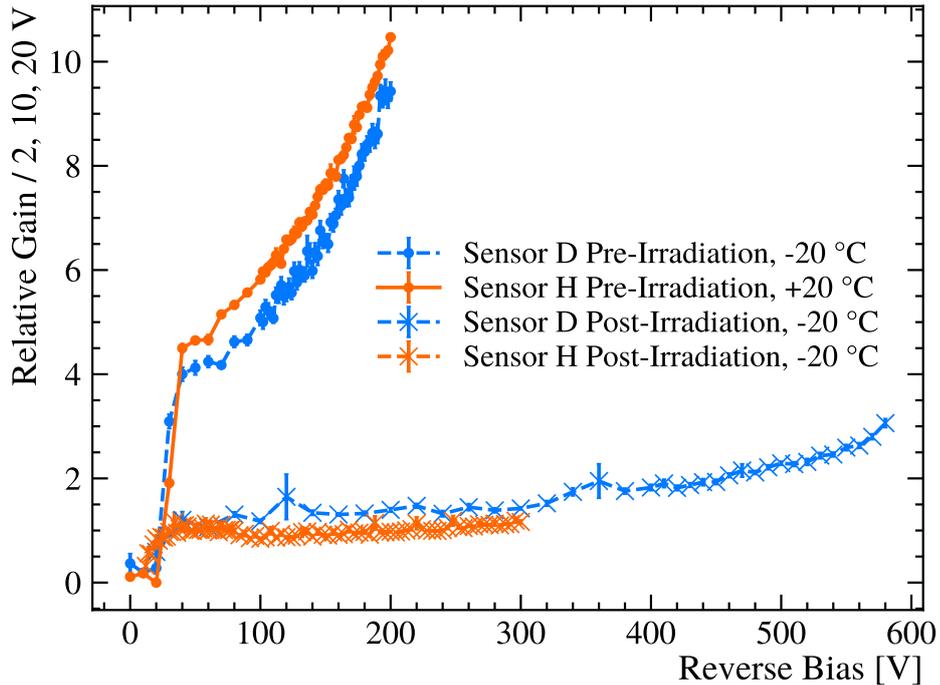


Figure 6.13: Gain measurements for sensors D and H displaying sensor performance before and after irradiation to 9×10^{14} $1 \text{ MeV } n_{\text{eq}}/\text{cm}^2$. The post-irradiation measurement of sensor H was stopped early due to temperature instabilities, but the trend is clear.

ature can be compensated by altering the bias voltage. During the sensor lifetime, where compounding radiation damage to the silicon results in acceptor removal and reduced gain at a fixed V_R , the same method can be applied to achieve a constant gain by progressively increasing the reverse bias. However, this fails if the device gain is degraded to an extent where the V_R required to compensate would be too large for operation. This was the case with the Micron sensors, with the gain-voltage measurements for sensors D and F shown in Figure 6.13. It is clear here that even up to the operational limit of 600 V the gain of the devices is almost entirely diminished in comparison to the pre-irradiation curves, implying the majority of dopants in the gain layer have been removed and that the corresponding time-resolution is expected to be similarly damaged.

6.3.5 Feasibility for VELO Upgrade II

After a fluence of $\Phi \approx 9 \times 10^{14}$ $1 \text{ MeV } n_{\text{eq}}/\text{cm}^2$, Figure 6.13 clearly demonstrates that the Micron LGAD performance has significantly worsened. For the minimum expected fluence for a VELO UII inner sensor across Run 5 and 6 of $\approx 1 \times 10^{16}$ $1 \text{ MeV } n_{\text{eq}}/\text{cm}^2$, it can be assumed that these LGADs would be damaged beyond operation in HL-LHC conditions.

Time resolution studies of other Micron LGADs from the same production batch were carried out in parallel at the University of Glasgow, yielding time resolutions as low as $\sigma_t = 26.5 \pm 0.7$ ps at -30°C [216]. While no time resolution measurements were carried out on devices post-irradiation, if a naive inverse-scaling approach is taken from other LGAD investigations ($\Delta\sigma_t \propto 1/G_L$) [217, 218], the resolution would be expected to degrade to $\sigma_t \sim 50\text{--}80$ ps for a fluence of 9×10^{14} 1 MeV $n_{\text{eq}}/\text{cm}^2$, exceeding the σ_t^{Pix} requirement for a VELO Upgrade II sensor, with simulated performance in Section 6.4.2.2 of different σ_t working points clarifying the importance of this metric.

These results, along with similar studies showing significant reduction in performance of LGADs above fluences of 2×10^{15} 1 MeV $n_{\text{eq}}/\text{cm}^2$ [207], mean that LGADs of the discussed design are unsuitable for the VELO Upgrade II sensor technology. Ongoing R&D at other foundries have investigated using carbon-infused dopants [218], designed to be more resilient to acceptor removal from irradiation. These LGADs have shown promising signs with radiation hardness increasing towards 6×10^{15} 1 MeV $n_{\text{eq}}/\text{cm}^2$ [227], but significant further R&D is required to understand if these relatively-cheap future pixel detectors are viable options for large scale projects such as the HL-LHC and beyond.

Despite the progress towards radiation hard LGADs, the option of 3D sensors remains more attractive, with current R&D showing radiation hardness meeting the minimum requirement for VELO Upgrade II and ultra-fast timing capabilities exceeding those of LGADs [211, 213]. 3D column sensors are therefore fast becoming the primary candidate for the upgrade of the VELO detector for Run 5 of LHCb, with extensive R&D by the IGNITE collaboration developing TimeSpot demonstrators [213] and work ongoing by the PicoPix and TimePix [228, 229] collaborations to optimise ASIC designs for this new sensor design [211, 212]. While these sensors are not without drawbacks, including the inefficient column regions, extensive simulation will be utilised to understand and optimise performance for the VELO Upgrade II detector, ensuring whichever technology is established delivers the physics performance required. The following Section 6.4 will explain techniques and present results for the current progress of VELO UII simulation with a natural future evolution of this to incorporate 3D sensor simulation.

6.4 Future Vertex Detector - Development and Simulation

The LHCb Upgrade II is scheduled to be installed during LS4 in 2033 and 2034, with Run 5 beginning the following year and physics scheduled at the LHC through to the end of Run 6 in 2041. To transform the subdetector design and performance requirements

discussed in Section 6.1 into optimised and realistic technical designs, detailed simulation studies are required alongside the comprehensive R&D projects. One such key area is constructing and utilising simulation of the VELO in Upgrade II to determine how cost and ongoing technical limitations could affect vertexing performance, and therefore the entire experiment in the HL-LHC environment.

In the following section the development of and results from simulation for VELO Upgrade II are discussed, including the process of moving from the hypothetical sub-detector scenarios presented in the LHCb Upgrade II Framework Technical Design Report (FTDR) [24] towards realistic optimised designs, with focus evolving from the detector metrics discussed in Section 6.2 to simulated physics performance.

6.4.1 VELO UII Scenarios

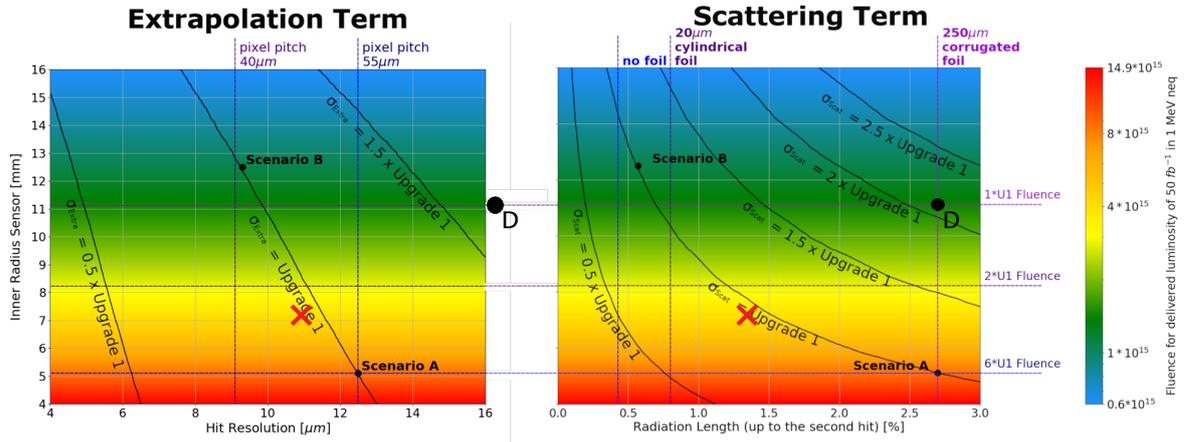


Figure 6.14: Impact parameter resolution represented by iso-lines of the extrapolation and scattering terms σ_{extrap} and σ_{MSC} in Equation 2.2.1 with respect to the UI values for a $\eta = 3.5$ particle. Described as a function of r_{min} against σ_{xy} (x/X_0) in the left (right) figure. On each, the heat map describes the fluence delivered per 50 fb⁻¹ of data taking (\sim a year for HL-LHC). Marked are the different VELO scenarios with pixel pitch and RF shield characteristics labelled.

The minimum requirement for VELO Upgrade II is to realise the same performance as the UI detector while managing the harsh environment with increased pixel occupancy and radiation damage. The key driving parameter for this is r_{min} , the radius of the innermost sensor, therefore to set the boundaries of design two “Scenarios” were established with the same performance as UI but exploring different technological limits. Scenario A (S_A) corresponds to retaining a near identical design to VELO UI design, with $r_{min} = 5.1$ mm, resulting in an innermost sensor fluence of $\geq 5 \times 10^{16}$ 1 MeV n_{eq}/cm^2 over the course of Run 5+6. Scenario B (S_B) instead considers the opposite limit where r_{min}

is relaxed to 12.5 mm, reducing the innermost sensor fluence to the UI level but requiring a significant improvement in σ_{xy} ($\approx 8 \mu\text{m}$) to recover the *IP* resolution performance in S_A , corresponding to a pitch of $\sim 40 \mu\text{m}$. Both options are considered challenging for R&D and are hence treated as theoretical boundaries, but are still important to evaluate and show that both routes for R&D of radiation hardness and improved intrinsic sensor performance are vital for a VELO to thrive in HL-LHC conditions.

Figure 6.14 displays how S_A and S_B vary significantly in design while located on the same iso-line for the σ_{extrap} and σ_{MSC} inputs to the IP resolution calculation (see Equation 2.2.1). To retain this performance S_B must reduce the detector material significantly, partially through a technically challenging ultra-thin RF shield marked on the Figure. The full description of the scenarios are stated in Table 6.3, including the specification for fast timing in both scenarios.

To complement the two hypothetical scenarios, two practical designs have also been derived based on limitations including detector cost and technology. Scenario D (S_D) refers to a descoped scenario with respect to S_A and S_B , with an intentional compromise on performance and physics reach to reduce the cost of the design by up to 50% [24]. Based on 65 nm ASIC technology, this imposes a stringent limit of $\Phi < 1 \times 10^{16} \text{ 1 MeV } n_{\text{eq}}/\text{cm}^2$ on fluence [230], limiting r_{min} to 11.5 mm. The design would be based on current-specification planar sensors, with an increased σ_{xy} and $\sigma_t = 200 \text{ ps}$, expected to damage the track time resolution and hence PV timestamp, with a knock-on effect to other subdetectors as discussed in Sections 6.1 and 6.2.

Scenario X (S_X) on the other hand, provides an initial realistic working point for VELO Upgrade II, derived to benefit from advantages of both S_A and S_B as well as projected technology developments to achieve, if not improve on, VELO UI performance. Marked in Figure 6.14, S_X is designed around improvements to ASIC technology using a 28 nm chip, expected to be radiation hard to fluences of $\approx 2.5 \times 10^{16}$ [231], enabling $r_{\text{min}} = 7.2 \text{ mm}$. The primary sensor option at this working point are 3D sensors, which are also expected to meet the $\sigma_t^{\text{Hit}} = 50 \text{ ps}$ requirement [213]. Despite the decrease in r_{min} compared to S_B or S_D , to retain UI performance the hit resolution for S_X would require improvement from the $12 \mu\text{m}$ in S_A . But while R&D is ongoing to reduce the pixel pitch in 3D sensors it is not yet known how much improvement must be derived from the method of clustering and whether the binary readout utilised for the UI detector can be re-implemented or improved upon.

Table 6.3: Summary of the four VELO UII scenarios implemented in simulation and discussed in this Section. For the RF shield design the “radius” refers to the innermost material of the shield, with the thickness stated referring to an aluminium design. $X_0(2)$ states the radiation length in the design up to the second hit. Fluence stated in units of $1 \text{ MeV } n_{\text{eq}}/\text{cm}^2$.

Parameter	VELO Upgrade II Scenarios			
	<i>A</i>	<i>B</i>	<i>D</i>	<i>X</i>
r_{min}	5.1 mm	12.5 mm	11.5 mm	7.2 mm
Pixel Pitch (σ_{xy})	55 (12) μm	40 (8) μm	60 (17) μm	55 (10) μm
σ_t^{Hit}	50 ps	50 ps	200 ps	50 ps
RF Shield Design	Corrugated	Cylindrical	Cylindrical	Cylindrical
RF Shield Thickness	250 μm	20 μm	200 μm	75 μm
RF Shield Radius	3.5 mm	11 mm	10 mm	3.5 mm
$X_0(2)$	2.7%	0.6%	2.7%	1.3%
Run 5+6 Fluence @ r_{min}	5×10^{16}	8×10^{15}	$\approx 8 \times 10^{15}$	2.5×10^{16}

6.4.2 VELO UII Simulation

The spectrum of scenarios discussed poses two questions. One of R&D focus, of reducing pixel pitch or improving radiation hardness, and the other of cost-benefit balance, can compromises be made while retaining a design in scope for building an experiment “to exploit the full forward physics potential of the HL-LHC” [24]. To answer these questions, an extensive simulation framework was designed to scrutinise each scenario quickly, testing the detector metrics discussed in Sections 6.1 and 6.2 but also investigating the effect on physics performance by processing signal simulation and performing analysis-like selection. While the exact specifications of each Scenario are not fixed, the designs used in the simulation are stated in Table 6.3, producing the VELO module cross-sections in Figure 6.15. This analysis will be followed by an in-depth study of S_X , which is currently the candidate design working point for VELO Upgrade II.

6.4.2.1 Simulation Framework

While a complete, detailed chain of simulation and reconstruction exists for the LHCb detector for Run 1–3 (as discussed in Section 2.4.3), the proposed UII detector poses significant differences in generation, detector modelling and reconstruction. Initially ignoring the re-design of many subdetectors, these include the requirement to simulate factor of ~ 10 more pp collisions per bunch crossing as well as the addition of timing to the simula-

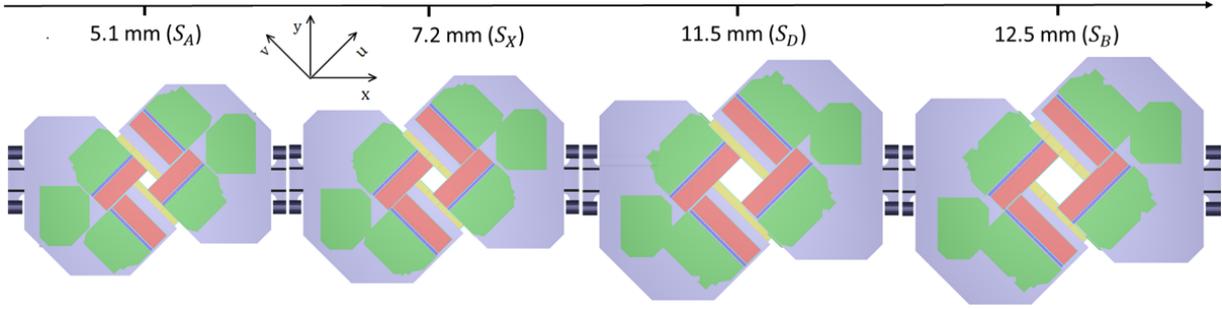


Figure 6.15: Visualisation of the four VELO scenarios in Table 6.3, displaying how altering r_{min} in the simulation automatically adjusts other parameters. The sensor layout follows the VELO UI layout in Figure 6.1a.

tion and reconstruction chain. The following will describe the developments implemented to aid with simulation of VELO Upgrade II.

Generation and Detector Geometry

The requirement to simulate more luminous pp collisions is solved by using the new “Gauss-on-Gaussino” LHCb simulation suite [232], with multi-threaded implementation allowing parallel generation and propagation of particles, enabling simulation of large samples in relatively short time ($\mathcal{O}(1)$ min per event at $\nu = 60$ for VELO-only simulation).

When simulating pp collisions for LHCb Run 1–4, there is no requirement to implement the distribution of PVs in time for each bunch crossing, with PVs produced during the intersection of bunches (≈ 1 ns) reduced to the same moment in time as no subdetector had the resolution to discern between them temporally. Therefore, for HL simulations, with the VELO targeting $\sigma_t^{\text{PV}} < 5$ ps, the natural spread of collisions within a bunch crossing must be emulated. As shown in Figure 6.16, the implementation of this essentially adds a Gaussian smear to each PV with a resolution derived from the beam crossing, spreading the PVs across the 1 ns window, enabling more-realistic reconstruction of different PVs at different times.

Gauss-on-Gaussino also introduced the use of the software “DD4HEP” to replace the detector geometry implementation from Run 1 and 2 [233]. DD4HEP provides a more flexible workbench to alter detector designs quickly, with a new C++ implementation and native visualisation and validation. This enabled the Run 3 VELO model to be modified into the different VELO UII scenarios discussed, producing a flexible geometry that could be rapidly tuned by altering a few important design parameters. The alterations can be broken down into the following parameters:

Inner Radius (r_{min}) By factorising the geometric parameters for the UI VELO to be as a function of r_{min} , changing a single parameter allows scaling of the VELO

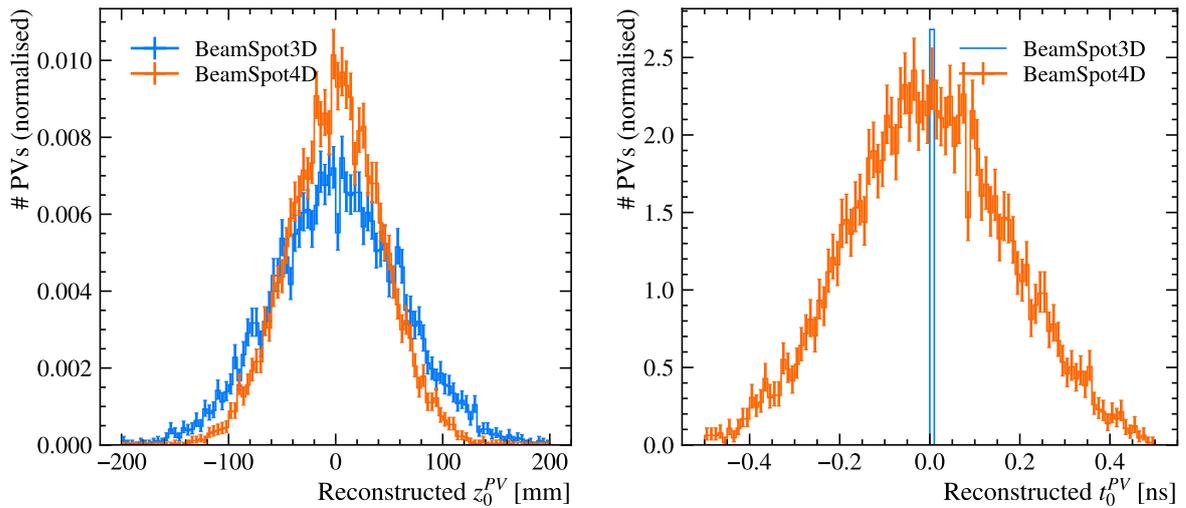


Figure 6.16: Distribution of PVs before and after the implementation of the 4D beam-spot for z (left) and time (right), showing the introduction of PV distribution in time during each bunch crossing rather than simply space.

module cross-section, as shown in Figure 6.15. This included translating the pairs of rectangular sensors in each corner radially out in local- y (v) as well as automatically extending them in local- x (u) such that each working point produced a square sensitive area with uniform acceptance. This extension in u was controlled by the addition of pixels, with the design shifting between three or four ASICs dependent on the r_{min} set. The substrate shape and hybrid position was designed to automatically adjust to approximately the correct position, with the inner sensors retaining an ≈ 5 mm overhang from the substrate edge.

RF Shield To test the change in design of the RF shield, a cylindrical tube of tuneable thickness and radius was implemented. This simplified quick alterations to the design and enabled simulation of an ultra-low material option. The UI corrugated shield can also be restored via a parameter, with scaling of its size and density implemented to approximately emulate changes in radius and thickness.

Pixel Layout As the size of each pixel changes, so does the VELO envelope. Any changes to pixel pitch in the simulation were therefore coupled to the substrate and hybrid shape and position. For Scenario X, specific pixel layouts were implemented, rather than the automatic adjustment. The parameterised model allows manual assignment of this, including requirements on the number and regular shape of ASICs, producing a 256×256 layout for S_X inline with the 2^N optimum channel size.

The parameterised model makes many approximations and the simulated designs are

not intended to be final. The geometry description is likely to be overhauled as Run 5 approaches. This implementation however enables design changes to be quickly realised in simulation and tested without major reworking of the model, producing results valid for comparison as the discussed parameters are altered.

Digitisation and Reconstruction

As discussed in Section 2.4.3, emulation of the detector response to a particle interaction is required to digitise simulation such that the reconstruction is aligned. Digitisation for the VELO detector in Run 1–3 included modelling spatial resolution, hit inefficiencies, noise and radiation damage [197]. For this to be implemented, detailed knowledge of the sensor technology is required.

For UII, with sensor R&D ongoing and no technology decision made, it is too premature to implement a new detailed digitisation and if the algorithms from Run 3 were reused this would neglect the emulation of temporal resolution, vital for UII reconstruction. Therefore, digitisation was replaced by a smearing approach, where the true spatial and time location of a particle traversing a VELO sensor were blurred in x, y and t by Gaussian resolution functions with tuneable σ_{xy} and σ_t . This method, referred to as “Fake Clustering”, enabled the intrinsic resolutions in each detector scenario to be decoupled from the design, allowing the pixel pitch to remain as an approximate value that primarily dictates the VELO envelope and enabling emulation of different sensor and ASIC technology, by altering σ_{xy} and σ_t . In reality, as shown in Figure 6.7, spatial and temporal resolution is not constant for a pixel detector with different particle incident angles leading to different amount of charge deposited as well as changing charge-collection times. But by using the VELO UII simulation suite to only draw comparisons between design working points, the conclusions using the constant σ_{xy} and σ_t are valid.

The “fake” clusters, complete with a timestamp, are combined and fitted into VELO tracks via tracking algorithms similar to those of Run 3, with the addition of a window in time enforced to ensure the complement of clusters are likely from the same particle. The window is tuned to $2.5\sigma_t$ for these studies. For a σ_t of 50 ps, this returns a track timestamp of ≤ 20 ps. VELO tracks are used to determine the position of PVs in the event, using the same method as for Run 1–3. The track-timestamp is vital to disentangle overlapping vertices, with the set of tracks from each PV constrained by windows in z and t , with the more precisely measured tracks given more weight in PV position determination. For an event with average track timestamp of 20 ps, PVs obtain an uncertainty of ≈ 3 ps (see Figure 6.20b), vital for separating responses in subsequent subdetectors by PV and for use in the potential pile-up suppression-based trigger.

Without simulating the downstream UT and MT tracking detectors, the full com-

plement of hits to form a long track is not available and hence a track momentum measurement is not possible. Therefore, for VELO UII simulation, a “downstream momentum parameterisation” is implemented by selecting tracks corresponding to particles that propagate into the approximate downstream tracking detectors acceptance ($2 < \eta < 5$) and smearing the true momentum by an estimated UII momentum resolution of $\Delta p/p = 0.5\%$ (consistent with LHCb Run 1–3). While there are shortcomings with this method, such as the constant momentum resolution, this enables selection of tracks likely to form physics candidates allowing study of quantities including the IP resolution and tracking efficiency of “long” tracks, important to understand the experiment’s physics-reach.

The simulation framework also goes a step further, moving beyond these discussed “detector metrics”, with a flexible physics selection framework allowing multi-body candidates to be built from long tracks and associated to PVs, with access to the MC decay-hierarchy information to determine the identity of particles and therefore study efficiency of reconstructing and selecting a given decay mode. The reconstruction of $B_s^0 \rightarrow D_s^\pm (\rightarrow K^+ K^- \pi^\pm) \pi^\mp$ decays and study of PV matching are presented in the following sections using this method.

6.4.2.2 Relative Performance of VELO Upgrade II “Scenarios”

Figures 6.17–6.20 display performance comparisons for the four VELO Upgrade II scenarios laid out in Table 6.3. As discussed, these results provide information about the relative performance and not the absolute values expected in data if the corresponding designs were implemented, due to the approximations in detector simulation and reconstruction.

Figure 6.17 shows the condition that Scenario A and Scenario B were defined on, with consistent IP resolution performance for tracks with $1/p_T < 1.5/\text{GeV}$, with departure only at low p_T where the reduced acceptance results in fewer hits-per-track and a lower resolution measurement. This is also shown in Figure 6.17b, with the S_B result increasing rapidly at high eta (out of frame). The results describe the reliance of S_B on an ultra-thin RF shield as, despite the significantly lower σ_{xy} , the increase to r_{min} with respect to S_A has damaged the resolution, which would be unfeasible for a thicker shield as in S_A .

Scenario X, with its reduced σ_{xy} and material is able to outperform both S_A and S_B , displaying that the degradation in performance from expanding to $r_{min} = 7.2\text{mm}$ can be recovered. The effect of changing shield design is highlighted by the difference between S_X and S_A in Figure 6.17b, where at high η particles traverse a greater amount of material in a cylindrical shield than one with corrugations around each module, where particles enter closer to the normal. For an ideal RF shield, the corrugated shape would be retained but with the low density that R&D has shown possible for cylindrical shields [24].

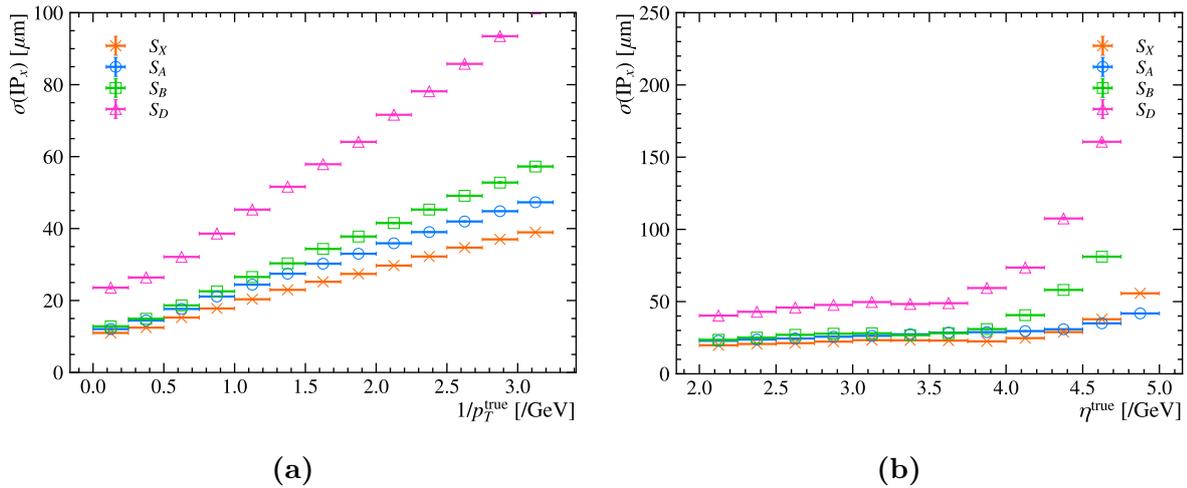


Figure 6.17: Track IP resolution (in x) for different VELO Upgrade II scenarios from the discussed simulation framework. (a) As a function of $1/p_T^{\text{true}}$. (b) As a function of η^{true} . Tracks are required to originate from a PV and be within long track acceptance with loose selection requirements. A linear fit to the S_X result yields $\sigma(IP_x) = 9.52 + 9.5/p_T$ with uncertainties of 0.08 and 0.2 respectively.

The results for Scenario D in Figure 6.17 display the dramatic deterioration in performance by the descoping scheme implemented, where a combination of the increased r_{min} , σ_{xy} and heavy cylindrical shield produce a consistently worse IP resolution than other scenarios. At low p_T and high η especially the resolution tends to impractical values, with association to individual PVs near-impossible at the average PV-PV distance $\approx 2.5 \mu\text{m}$ for a 500 ps time window in HL-LHC conditions.

To understand the direct effect on physics performance from altering the VELO UII design working points, simulation of 50000 $B_s^0 \rightarrow D_s^\pm (\rightarrow K^+ K^- \pi^\pm) \pi^\mp$ decays were generated with each scenario, propagating each sample through the reconstruction framework as well as the physics selection suite. Loose kinematic cuts are applied to select displaced tracks that could originate from the D_s^\pm , with “perfect” PID requirements imposed to separate K and π tracks by their true identity, with no simulation of the PID detectors available. Combining these tracks into D_s^\pm candidates enables study of the selection efficiency, mass resolution and background level from the decays of displaced vertices, a cornerstone of the LHCb physics programme.

Figure 6.18a displays the reconstructed D_s^\pm mass for the VELO UII layouts. The same selection is imposed on each scenarios selection chain, with the only difference of note that the S_D sample was generated at a reduced $\mathcal{L} = 1 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ corresponding to $\nu = 38$, inline with descoped operating conditions, rather than the nominal $\nu = 60$ with $\mathcal{L} = 1.58 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. The performance of S_D is strikingly worse than the other scenarios, unable to effectively separate signal from background and leading to a poorer

mass resolution. The other scenarios show compatible performance, with S_X the best performing as expected from Figure 6.17. The results show “signal” simulation, without realistic emulation of combinatorial background. Therefore the background level shown originates entirely from the other pile-up vertices in the event, further amplifying the poor performance of S_D , simulated with a lower luminosity. “Real” combinatorial background in data is likely to be of greater magnitude.

Figure 6.18b records the background fraction in each scenario as a function of $\max\eta$ of the D_s^\pm children. The performance degrades for all scenarios as η increases. This explicitly highlights the need for fast timing to suppress background from random combinations of tracks, especially at high-eta where the density increases and correct candidate building becomes more difficult with overlapping tracks associated to multiple PVs. The S_D result, with $\sigma_t = 200$ ps, displays this with an unfeasible background fraction at high- η . The compatible performance of S_A and S_B confirms the conclusion that performance can be recovered for a large r_{min} VELO. Demonstration of this using signal simulation, rather than one detector metric, includes factors such as differing mass resolution, material budget and acceptance from the different peak widths, background level and performance as a function of η . The S_X result implies that an optimum combination of the two approaches, reducing σ_{xy} and minimising r_{min} , is preferred.

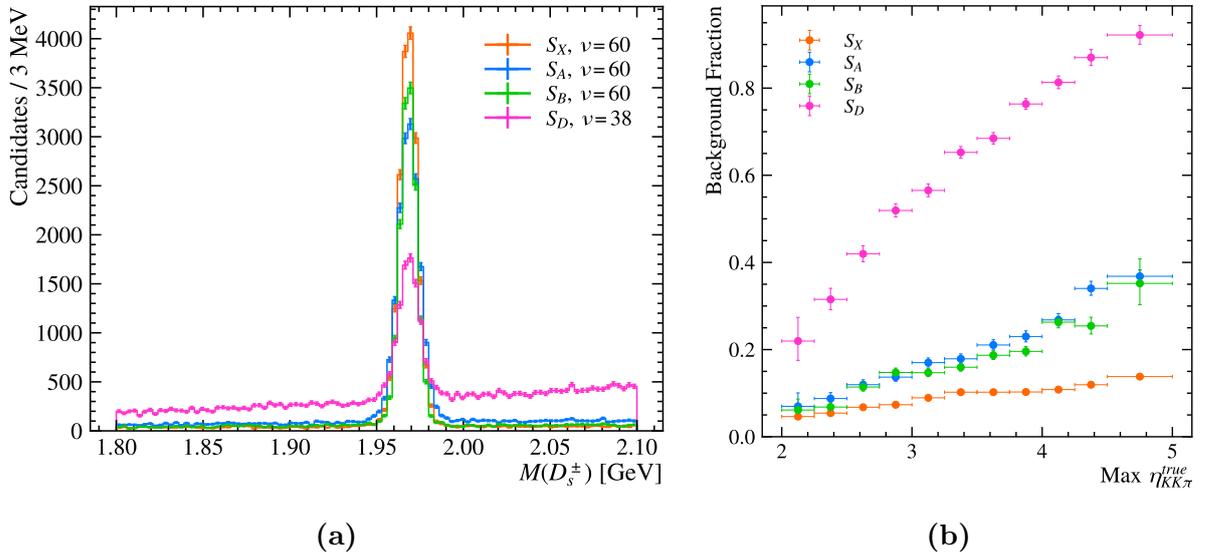


Figure 6.18: Reconstruction of D_s^\pm candidates for different VELO Upgrade II scenarios, using the discussed UII selections framework. (a) The reconstructed $M(D_s^\pm)$ after loose selection requirements. (b) The fraction of events that are background as a function of the maximum η^{true} of the D_s^\pm children from fits to (a), illustrating the worsening resolution from the inclusion of a high η track.

The fixed scheme of selections across scenarios allows direct comparison, but in reality selection would be optimised for the final VELO UII detector. Therefore a signal

purity study is performed on the $B_s^0 \rightarrow D_s^\pm(\rightarrow K^+K^-\pi^\pm)\pi^\mp$ samples by varying the minimum p_T selection on the children of reconstructed D_s^\pm candidates. The signal purity against signal efficiency (with respect to no minimum p_T selection) displayed as a ROC curve for each scenario in Figure 6.19a. To achieve a target 95% purity from this selection, S_D would result in a signal efficiency of 28% requiring a cut of 1176 MeV, where as S_X retains 68% of candidates for a significantly looser selection at 570 MeV. The difference in resulting reconstructed D_s^\pm mass for a purity of 95% is displayed in Figure 6.19b, another example of the unsuitability of S_D , with the ability to reconstruct extremely large, pure samples of displaced vertex decays (including $D_s^+ \rightarrow K^+K^-\pi$) an essential aspect of CP violation studies at LHCb [234].

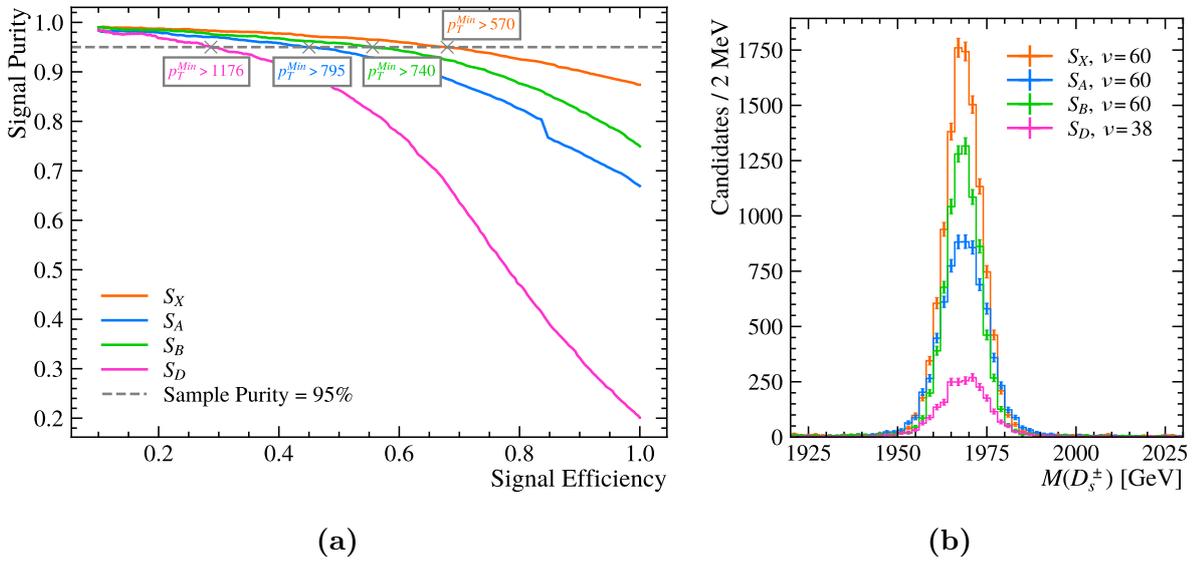


Figure 6.19: Effectiveness of track p_T selection as a function of different VELO Upgrade II scenarios. (a) The achievable signal efficiency and purity by tuning a selection on the minimum p_T of the children of reconstructed D_s^\pm candidates. (b) The $M(D_s^\pm)$ distributions with the cut tuned to a purity of 95% for each scenario.

To determine vital physics quantities such as the lifetime of b -hadrons, PVs must be reconstructed and the 4D position determined. In HL-LHC conditions, as discussed, this poses significant difficulty from overlapping PVs with a precise timestamp of both tracks and vertices essential. The efficiency to reconstruct PVs (and the corresponding time resolution) is presented for each scenario in Figure 6.20 as a function of the number of tracks produced at each PV. Despite the reduced luminosity of S_D , the performance is consistently worse than the other scenarios, unable to consistently reach the 94% efficiency target from UI performance [197]. The Scenario A, B and X results, all with 50 ps time resolution per hit, display the importance of spatial resolution as well as low- η acceptance, with $\eta < 2$ and backwards VELO tracks critical for PV reconstruction, further clarifying the need for an optimisation between σ_{xy} and r_{min} for the final VELO Upgrade II design.

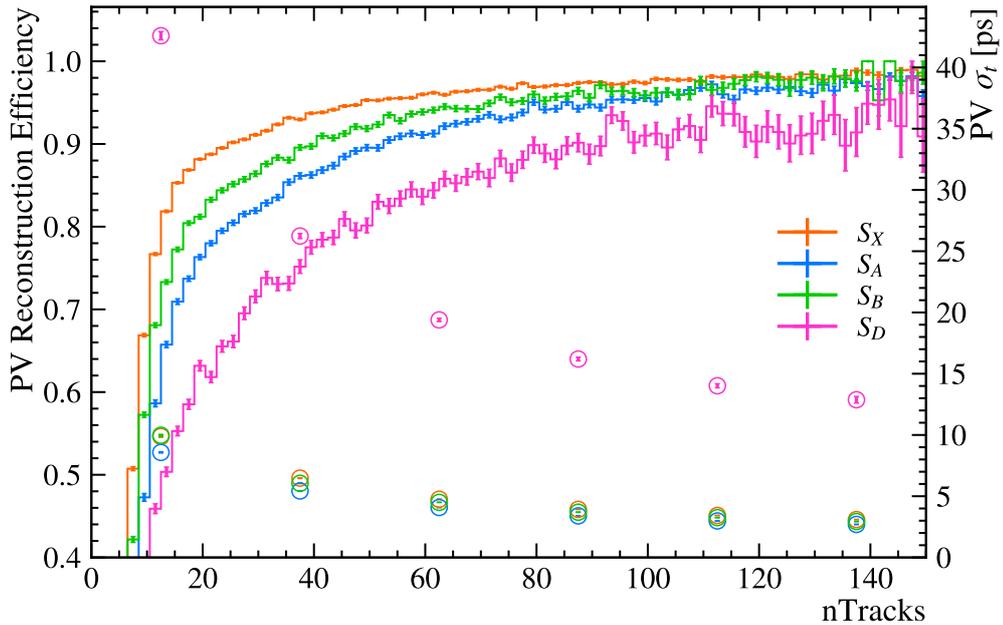


Figure 6.20: PV reconstruction efficiency (coloured lines) and PV time resolution (coloured markers) as a function of track multiplicity for different VELO Upgrade II scenarios.

These results for Scenario A, B and D were presented to the LHCb collaboration in 2023, displaying that the requested “Descoped” scenario is unfeasible for a performant VELO in HL-LHC conditions due to the dramatic impact on performance, despite being tested already at a reduced luminosity. This led to the collaboration and LHCb UII Resources-Review Board [235] concluding that S_D and any significant VELO descoping is not worth pursuing, with the quality of VELO reconstruction fundamental to downstream performance including the formation of long tracks and the initial complement of PV timestamps underpinning the fast-timing capabilities of subdetectors across LHCb Upgrade II.

6.4.2.3 VELO Upgrade II Performance by Characteristic

“Scenario X” has evolved into the current nominal layout for VELO Upgrade II. There is no doubt that the design will change as work for the VELO UII Technical Design Report (TDR) approaches, as constraints from sensor, cooling and RF shield R&D develop and more precise cost-estimates are possible. It is therefore important to understand which aspects of the current S_X design (here onwards referred to simply as “VELO UII”) have the most significant impact on performance, and where possible cost-savings or improvements could benefit the overall design. The following section will discuss the importance of timing resolution, spatial resolution and angular acceptance by varying characteristics

individually before reviewing the status of projected data-rates for the readout of the VELO, a potential development barrier not discussed up to now. Results as a function of RF shield design are also presented in Section 6.5.3.

Importance of $\sigma_t = 50$ ps

The reliance on fast-timing for VELO UII is displayed in Figure 6.21, by utilising the same $B_s^0 \rightarrow D_s^\pm (\rightarrow K^+ K^- \pi^\pm) \pi^\mp$ reconstruction and selection chain as before with four different σ_t^{Hit} working points. This displays that, as with S_D in Figure 6.18, increasing σ_t damages the mass resolution and introduces combinatorial background. While the D_s^\pm mass plot appears to imply a minor difference between $\sigma_t = 50$ ps and 100 ps, the ROC curve for min- p_T selection shows a 14% difference in selection efficiency for obtaining 95% purity. This could be potentially mitigated by more complex selection, suggesting a slight relaxation of the 50 ps timing constraint for VELO UII could be feasible for displaced candidate reconstruction. However the clear damage to PV reconstruction and time resolution disputes this, where a doubling of the resolution significantly degrades the low-Tracks PVs, resulting in $\sigma_t^{PV} > 5$ ps, above the UII target. This would not only damage performance in the VELO; the precision of this timestamp is relied on for associating EM showers and Cherenkov rings to individual PVs within each bunch crossing for the ECAL and RICH upgrades respectively. This timestamp also underpins the trigger strategy for UII, with real-time separation and selection of individual PVs within a bunch crossing only possible if the online system can rely on the timing information provided.

One of the essential and unique assets of LHCb is the ability to accurately “flavour tag” the decays of b -mesons, to determine the initial particle/anti-particle state of $B_{(s)}^0$ decays for CP violation studies of neutral flavour-oscillations [236]. The technique is built on reconstructing a B candidate and analysing other particles in the event originating from the same PV and hence inferring the flavour at production of the candidate. For Run 1–4, the low pile-up in events allows particles from a single PV to be efficiently isolated [236], but at Run 5 luminosity there is a greater likelihood for incorrect PV matching and mis-tagging introduced from either the B -candidate or surrounding tracks associating to the wrong PV. This reduces $\varepsilon_{\text{match}}$, the PV association efficiency. The addition of fast-timing across LHCb UII is a proposed fix for this, hence an initial study of the Muon Opposite-Side (μ OS) tagger was derived using VELO Upgrade II simulation of displaced muons originating from the decays of b -hadrons in UII conditions at $\mathcal{L} = 1 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, processed through the UII selection framework discussed.

The μ OS tagger utilises $b \rightarrow c\mu^-\bar{\nu}_\mu$ decays on the opposite-side of the detector to a candidate B decay of interest, implying that if both the B and μ share a PV they are from a fragmenting $b\bar{b}$ pair and hence the sign of the μ indicates the initial B flavour [236].

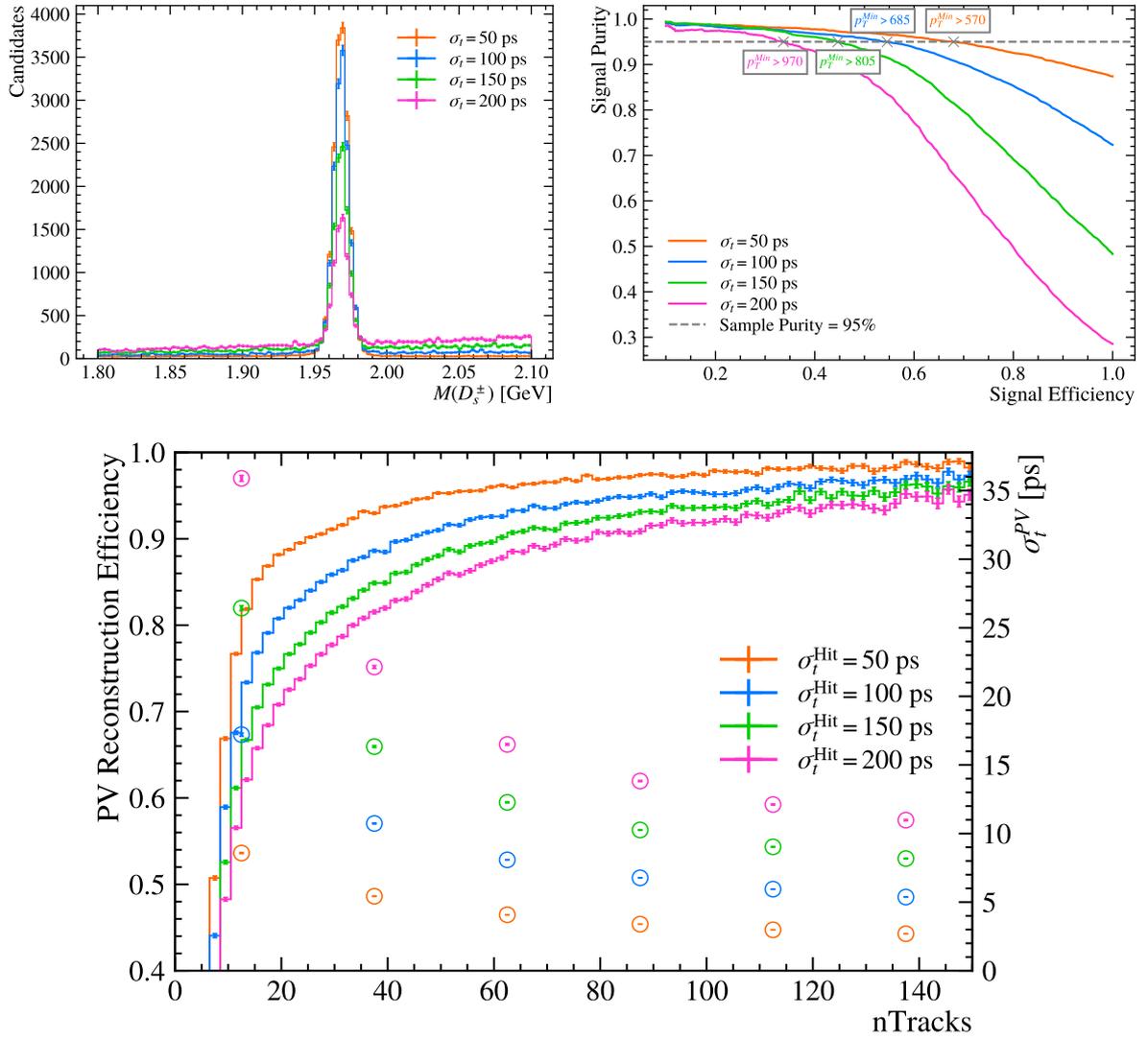


Figure 6.21: Reconstruction and selection of D_s^\pm candidates for VELO Upgrade II, with different σ_t imposed to display the importance of precise time resolution. (upper left) The reconstructed $M(D_s^\pm)$ after a series of loose selection requirements. (upper right) Selection purity curves for varying the minimum p_T of the children of reconstructed D_s^\pm candidates. (lower) PV reconstruction efficiency (coloured lines) and PV time resolution (coloured markers) as a function of track multiplicity.

Therefore, studying the PV matching efficiency of a displaced muon sample with different VELO UII designs and methods can ascertain the expected relative performance for a μ OS tagger in UII, without simulation of the candidate B or Muon chambers.

Figure 6.22 presents the muon $\varepsilon_{\text{match}}$ as a function of $\eta(\mu)$ for two VELO sensor σ_t working points as well as two techniques of PV association. It is clear that using spatial measures alone ($\chi_{IP,3D}^2$) returns worse performance, notably at high η where tracks will point to many of the ~ 27 pile-up vertices. By folding in track and PV resolution, a $\chi_{IP,4D}^2$

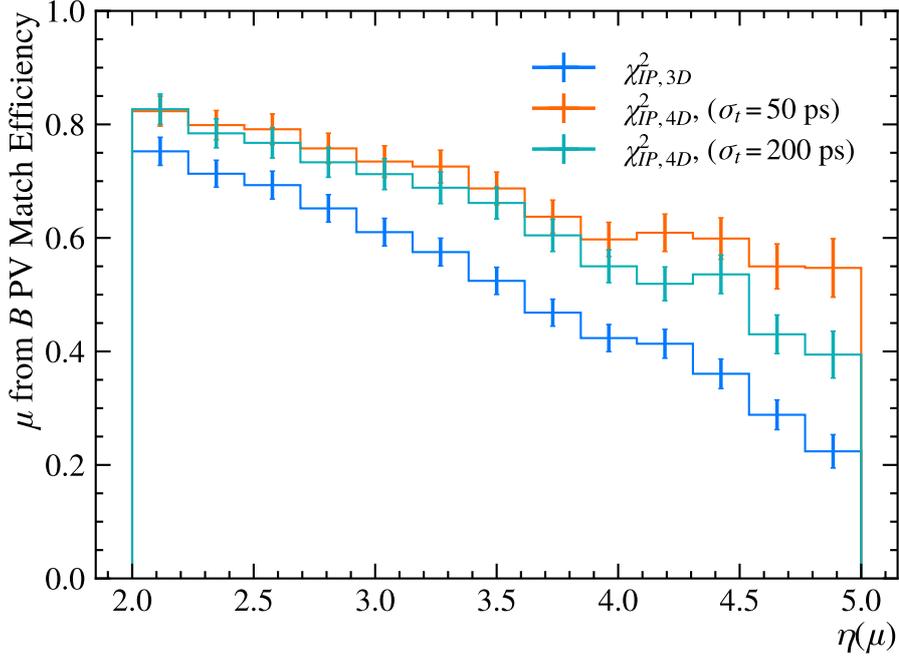


Figure 6.22: PV matching efficiency as a function of track η for muons from the decays of b -hadrons, an important metric for flavour tagging at LHCb. Distributions show different PV matching techniques using $\chi^2_{IP,4D}$ and $\chi^2_{IP,3D}$ as well as the degradation in performance for a reduced σ_t sensor. All results shown at $\mathcal{L} = 1 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ working point and timing was used to construct the PVs initially with the method discussed in Section 6.4.

is constructed, improving performance across the acceptance and providing significant recovery at high- η . However for a degraded σ_t the high- η matching ability is quickly harmed, corresponding to an increase in the rate of mis-tagging, displaying the importance of timing to retain LHCb physics reach in the quark CP violation sector.

The natural evolution of this study would be to build a complete μ OS tagger in simulation with reconstruction of candidate B decays and information from the Muon chambers, but this would require a more complete simulation and reconstruction framework, with the aim of the discussed VELO UII studies to promptly produce results to gain relative performance understanding.

Angular Acceptance with $r_{min} = 7.2 \text{ mm}$

Despite the overall improved performance shown in Section 6.4.2.2 for the VELO UII (S_X) design with respect to VELO UI (taken as S_A), some degradation can be seen in the most forward region in Figures 6.17, 6.18 and 6.22. This is generally expected for any designs, with a greater density of tracks in the most forward region, but the difference to UI attributed to the reduced acceptance from increasing r_{min} to 7.2 mm.

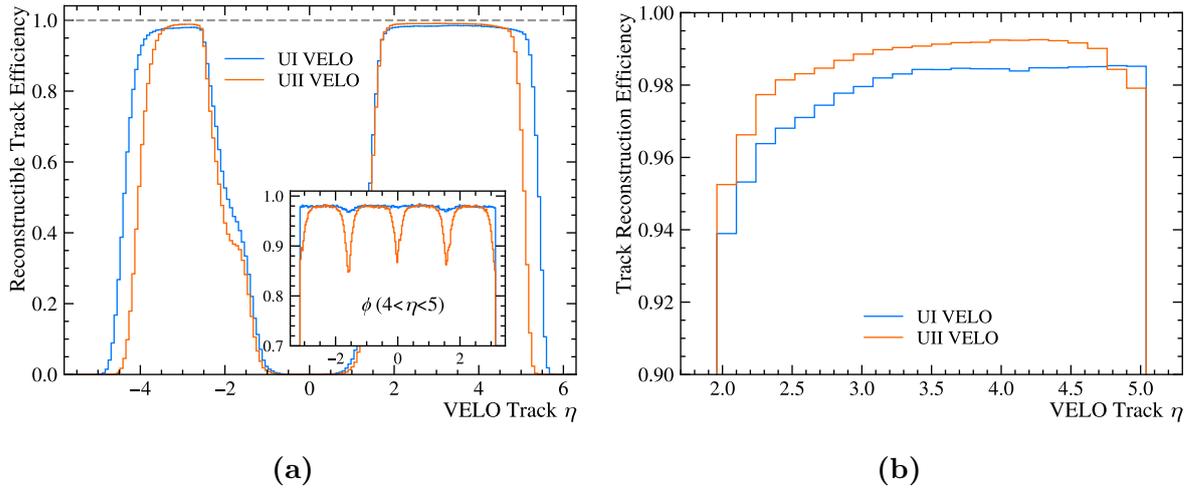


Figure 6.23: Tracking efficiency for VELO UII as a function of track η , comparing acceptance to the UI VELO. (a) the reconstructible track efficiency, with the $4 < \eta < 4.8$ region overlaid as a function of track ϕ to display acceptance gaps. (b) The track reconstruction efficiency, for all VELO tracks.

Figure 6.23 displays the difference in track acceptance and reconstruction performance between the UI and UII geometry (with S_A simulation shown for UI). Figure 6.23a describes the “reconstructible” track efficiency, what fraction of generated tracks originating from the interaction region traverses at least three VELO sensors, providing an efficiency metric for maximum performance if mis-reconstruction and detector effects like hit-efficiency and noise are ignored. The UII result is compatible with UI up to $\eta \approx 4.7$, showing flat efficiency as a function of ϕ preventing the need for angular-based corrections in physics analyses. Above $\eta = 4.7$, the increased r_{min} results in tracks that were in acceptance for UI no longer able to be reconstructed, with the overlaid distribution in ϕ (for $4 < \eta < 5$) showing a drop in efficiency from the corners of the square cut-out in the VELO design. Compounding with this, Figure 6.23b shows that the track reconstruction efficiency, the fraction of VELO tracks reconstructed with respect to the number in acceptance, begins to also reduce above $\eta = 4.7$ compared to UI, displaying the effects of overlapping tracks and increased multiple scattering in the cylindrical RF shield, without the radial coverage to compensate this.

One solution to regain this acceptance would be to extend the VELO in z , but strict mechanical constraints from RICH1 would likely prevent this. Re-optimisation of the z -layout will be performed to improve the acceptance of backwards tracks to aid with PV reconstruction, but with effects such as the expected reduced hit efficiency from new sensor technologies still to be implemented (see 3D sensors in Section 6.2.2), the required flat reconstructible and reconstruction efficiency up to $\eta = 5$ seems unfeasible without dramatically increasing the number of forward stations in the current footprint, which

introduces further structural constraints.

However, recent investigation of mechanical constraints from the LHC beam pipe on the Mighty Tracker has determined that an acceptance of $\eta > 4.8$ may be unobtainable for the central MightyPix sensors. This could relax long track acceptance requirements for the VELO, rendering the performance shown in Figure 6.23 more plausible, with an $\eta \approx 0.1$ recovery from 4.7 likely to be possible by re-optimisation. Considering Figures 6.17, 6.18 and 6.22, a relaxed acceptance would result in a better performance on average for VELO Upgrade II, especially for IP resolution and background suppression. The reduced acceptance would come at a cost to physics yield in some areas however, with dedicated studies required to determine this.

Ultra-Low Spatial Resolution Feasibility

For the current VELO UII design the most ambitious characteristic is $\sigma_{xy} \leq 10 \mu\text{m}$, which aids in the improved performance over S_A and S_B in Section 6.4.2.2. For this to be achieved for a $55 \mu\text{m}$ pixel pitch, significant improvements in clustering would be required compared to the UI performance, which achieved an average $\sigma_{xy} = 12 \mu\text{m}$, only reaching $10 \mu\text{m}$ at preferential increment angles as shown in Figure 6.7 [197]. This could potentially be delivered by moving to a charge-weighted clustering scheme, where readout of each fired pixels deposited charge improves the calculation of a particle's hit position. Investigation of this for VELO UI yielded an average $\sigma_{xy} = 10 \mu\text{m}$ [237] if all pixel hits were read out, associated and combined into the correct cluster. However this is an unfeasible scenario for HL-LHC conditions, with up to 20000 pixel hits projected per bunch crossing. Therefore investigation of alternative clustering methods and a review of performance with an increased σ_{xy} are required.

Performance comparison of relaxing the $\sigma_{xy} = 10 \mu\text{m}$ requirement is shown in Figure 6.24. The UII $\sigma_{xy} = 12 \mu\text{m}$ result approximately corresponds to utilising the UI pixel pitch and clustering technique (if the data rate issue is ignored) and the $\sigma_{xy} = 16 \mu\text{m}$ result is the minimal scenario of $p/\sqrt{12}$ as defined in Equation 6.2.1, where a single high-charge pixel hit is used to form a cluster. The UI VELO result is also emulated with the same simulation framework to enable direct comparison. A theoretical result for the UI VELO with a thin cylindrical shield is included to decorrelate the difference in design of the RF shield between UI and UII. Figure 6.24 shows that with the increase in r_{min} , $\sigma_{xy} = 10 \mu\text{m}$ is necessary to achieve the same intrinsic sensor-only performance as UI, but relaxing to $\sigma_{xy} = 12 \mu\text{m}$ appears feasible thanks to other improvements such as material budget. The $p/\sqrt{12}$ resolution however degrades the IP resolution to $15.0 + 12.2/p_T$, with σ_{extrap} substantially above the UII target ($12 \mu\text{m}$).

To achieve $\sigma_{xy} \leq 12 \mu\text{m}$, either developments in clustering at high particle-rates

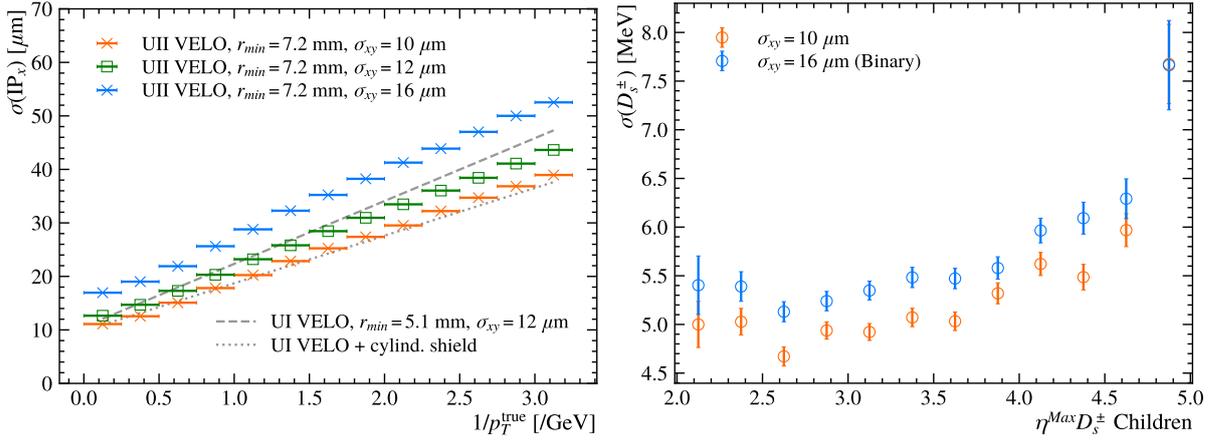


Figure 6.24: Requirement for spatial resolution beyond the “binary” resolution defined in Equation 6.2.1. (left) Track IP resolution comparison of Scenario X to an edited UI VELO design as it is moved to $r_{\text{min}} = 7.2 \text{ mm}$ and to S_X with binary resolution for a $55 \mu\text{m}$ pixel sensor. (right) Describes D_s^\pm mass resolution as a function of the maximum η^{true} of the D_s^\pm children for the nominal and binary resolutions for S_X .

are required to improve on the VELO UI method, or the pixel pitch must be reduced. Recent R&D shows that achieving $\sigma_{xy} = 10 \mu\text{m}$ by reducing pitch alone to $p \sim 35 \mu\text{m}$ is unattainable for the likely 3D sensor technology choice [238, 239] as a smaller pitch would increase the fractional inefficient area from the readout columns. Reducing to an intrinsic $\sigma_{xy} = 12 \mu\text{m}$ from $p \sim 42 \mu\text{m}$ pixels is also considered technically challenging for the timeline of VELO Upgrade II [238]. To achieve these resolutions an achievable combination of reduced pitch and improved clustering techniques at high-luminosity will be required, with proposed 3D sensor sizes of $45 - 49 \mu\text{m}$ to be fabricated and tested with VELO UII in mind, providing $\sigma_{xy} = 13 - 14 \mu\text{m}$ with single-pixel readout [238, 240]. If this effort is paired with dedicated test-beam and simulation studies of clustering techniques that are feasible for HL-LHC conditions, it is plausible that the data readout rate could be reduced while improving σ_{xy} by $2 - 3 \mu\text{m}$. This would produce the target σ_{xy} discussed producing a VELO UII that delivers on the LHCb Upgrade II physics potential.

6.4.2.4 Projected VELO UII Readout Data Rates

Despite the increase in r_{min} , operating in HL-LHC conditions poses a $5 - 7.5$ factor increase in luminosity for VELO Upgrade II, introducing a potential limitation from the bandwidth for reading out the detector response. As displayed in Figure 6.1b for VELO UI, the average number of particles per bunch crossing for the closest ASIC is ~ 10 , corresponding to an average data rate per ASIC of $\approx 10 \text{ Gbit/s}$ (assuming 44 bits per particle), with the peak data rate integrated across all stations and ASICs reaching 2.85 Tbit/s [197]. While a simple extrapolation to UII conditions is possible, the correlated changes to VELO

UII design, sensor technology and varying luminosity justified a simulation-based study into the expected data-rates, to understand the magnitude of the task for UII readout-electronics and identify optimal savings that could be made through design changes and new clustering techniques. Studies were carried out at two luminosity working points $\mathcal{L} = 1 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ and $\mathcal{L} = 1.58 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, to understand LHCb UII relative performance for a descoped upgrade.

For the “fake clustering” technique a simulated particle’s traversal of a VELO sensor is converted into a cluster through position alone, without requiring knowledge of which sensor, ASIC or individual pixel is hit. To study data-rates the simulation was extended to track the path of a particle through the sensor and determine which individual pixels were likely to be fired by emulating the deposition and diffusion of charge in silicon. This produces multiple pixel hits per particle that for full simulation require reconstruction into clusters, as discussed in Section 6.2.1.1, to determine an estimate of the particle’s position as it crosses the centre of the sensor, producing an intrinsic spatial resolution.

For this study of particle and pixel hit rates, the information was simply used to produce Figures 6.25 and 6.26, allowing projection to expected particle and hit-rates for the discussed VELO Upgrade II design with flexible study of different metrics. Figures 6.25a–b describe the total number of particles and hits per station per bunch crossing, for both the low and high luminosity working points. The distribution of particle rate shows a maximum for Station 15 at $z \approx 20 \text{ cm}$, instrumenting the luminous forward region close to the interaction region. The most central stations however are more likely to measure a track traversing at a high incident angle, with the number of pixel-hits dramatically increasing despite the reduced particle rate from the smaller solid-angle these stations cover. The differences between these distributions are important; while overall the UII VELO produces an average hit (particle) rate of 461 (225) s^{-1} for the low luminosity working point and 727 (355) s^{-1} for high, the final ASIC design must facilitate the peak data-rate. For the VELO UI technique, with effectively all pixel hits read out, this would be dictated by the central stations. If however clustering can be performed with onboard ASIC electronics, the peak would be significantly reduced, and dictated by Station 15 in the current model.

Figures 6.25c–d describe the distribution of particle and pixel-hit rates across the ASICs on the discussed peak stations, displaying the extreme environment the closest ASICs experience. Figure 6.25c can be directly compared to Figure 6.1b, showing that despite the increase in r_{min} with respect to UI, the particle rate has increased by a factor 5, with an average 52 particles per bunch crossing the “hottest” ASIC in UII. Assuming an average LHC frequency of 30 MHz and a packet size of 44 bits [24], a clustering-on-the-ASIC scheme is projected to require an average bandwidth of 69 Gbit/s, but with

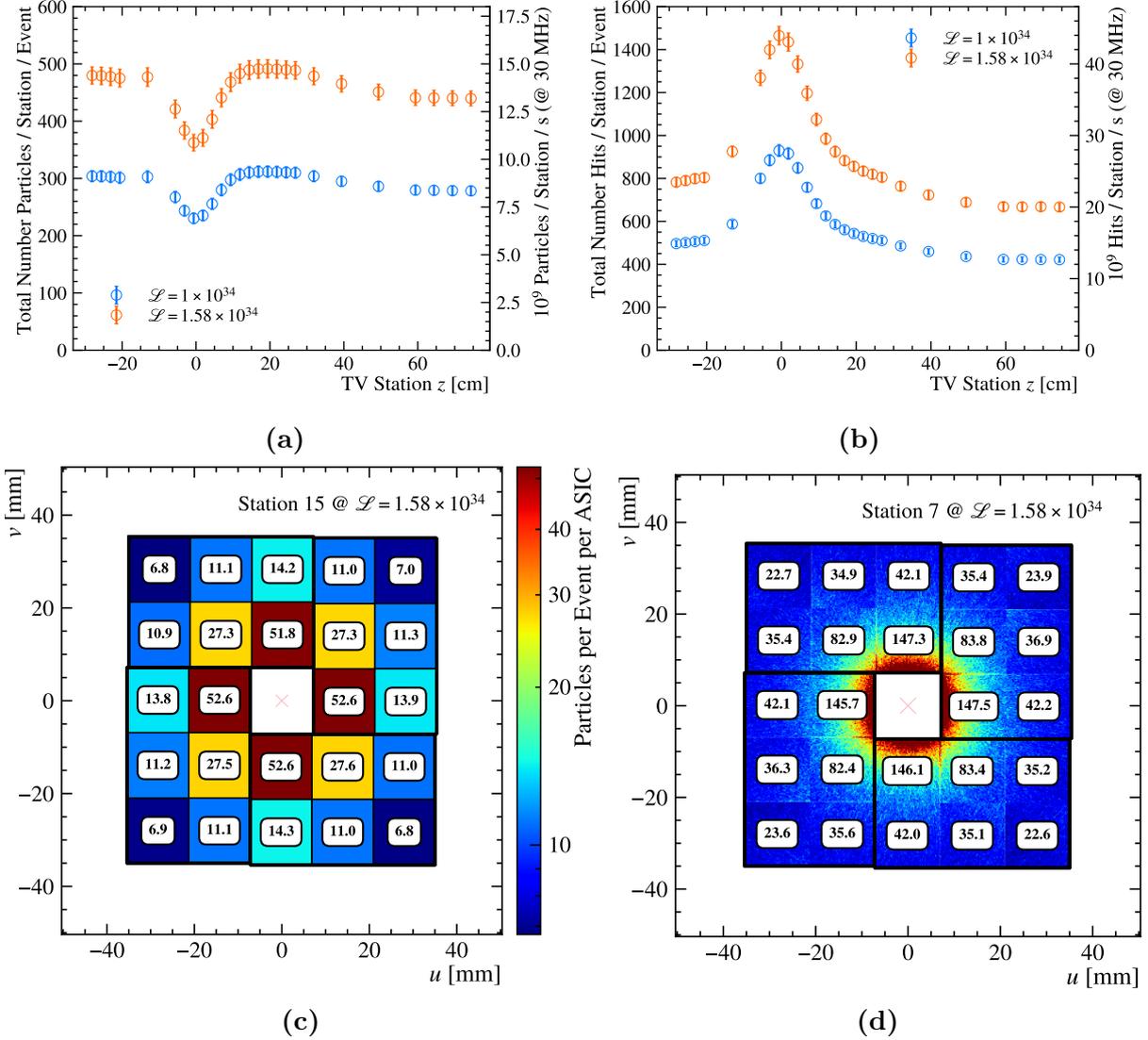


Figure 6.25: A study of particle and pixel hit rates for VELO Upgrade II. (a) The mean number of particles per station per event comparing $\mathcal{L} = 1 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ and $\mathcal{L} = 1.58 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. (b) The mean number of hits per station. (c) The mean number of particles per ASIC per event, for the peak particle-rate Station 15 ($z = 19.4 \text{ cm}$). (d) The mean number of pixel hits per ASIC per event, for the peak hit-rate Station 7 ($z = -0.6 \text{ cm}$), with a heat-map of pixel hits overlaid (heat-map scale is arbitrary). For (a) and (b) a second axis of rate / s is included, calculated for the 30 MHz average collision rate. Rates in (c) and (d) are at the higher luminosity.

a maximum clock of 40 MHz and expected fluctuations in particle-rate, a safety factor of ~ 1.5 is applied, increasing the peak bandwidth to ~ 100 Gbit/s. With the VELO UI clustering scheme, with all pixel hits read out, this explodes to > 300 Gbit/s with a greater safety factor required. While both are beyond the limitations of single copper links at 20 Gbit/s [241], the latter is considered unfeasible for even new readout techniques including state-of-the-art Silicon Photonics modules, which are currently limited to ~ 100 Gbit/s [242] per link, with one link per ASIC preferred for LHCb UII.

For the lower luminosity working point, the projected particle rate reduces to an average bandwidth requirement of 44 Gbit/s, with the peak at ~ 65 Gbit/s still requiring a novel readout method. The difference from the clustering technique highlights the need for R&D and simulation in this area. Simulation of the potential techniques is required to study their affect on performance, with the reach of clustering-on-the-ASIC unknown but likely to be unable to provide the spatial resolution that a complete pixel hit read out would deliver [228].

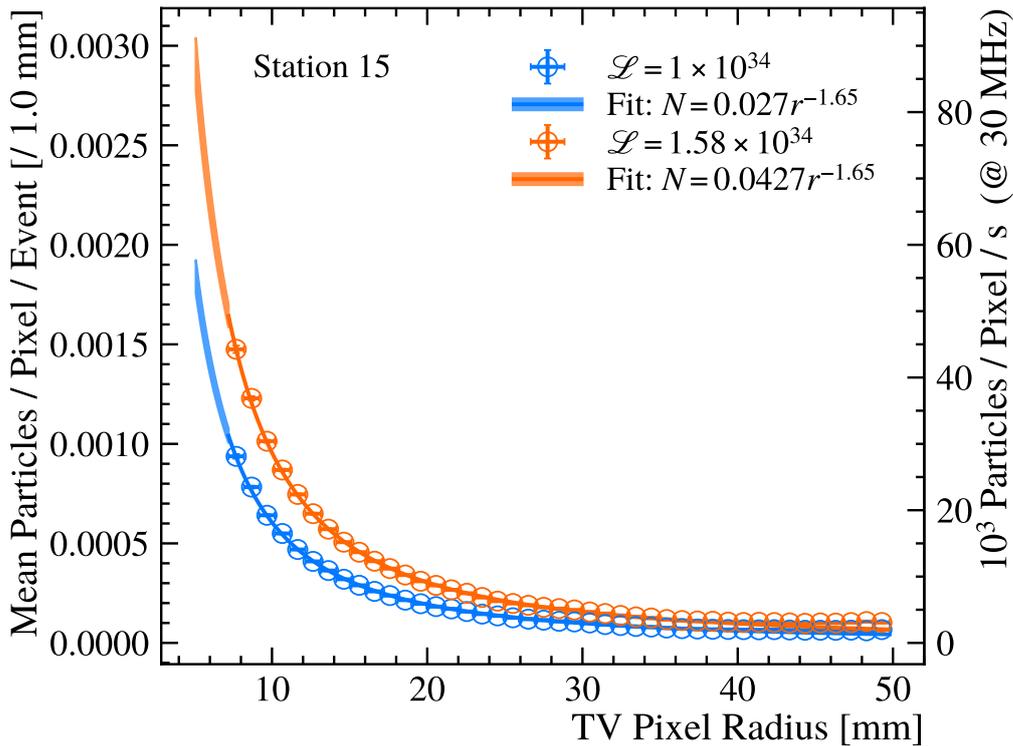


Figure 6.26: Mean particles traversing each pixel per event as a function of radius for VELO Upgrade II Station 15 at different luminosity working points with the corresponding average rate included. The radius is calculated at the midpoint of a particle’s traversal of the silicon sensor. The distribution is fit to a linear shape between $r = 7.2$ mm and 50 mm and then extrapolated to $r = 5.1$ mm to estimate the rate for a VELO UI design in HL-LHC conditions, with the fit result uncertainty shown by the coloured bands.

An alternative path to reducing the peak ASIC data-rate is to simply reduce the size of each ASIC. Without departing from the 2^N square pixel constraint, this would require reducing the pixel pitch considerably, with its own technological challenges. However, the fact remains that the pixels closest to the beamline will contribute the most to data-rate, implying that increasing r_{min} is a more suitable method to decrease the bandwidth requirements. To study this, Figure 6.26 describes the average number of particles traversing each pixel as a function of pixel radius for VELO Upgrade II. In the current design, the ASIC closest to the beamline covers the region 7.2–21.3 mm, with a potential smaller ASIC reducing coverage linearly with pixel pitch. However the Figure shows the majority of the overall rate is from the lower radius pixels, and hence a smaller ASIC would not significantly reduce the data rate, and actually increase the average pixel hit-rate in that ASIC. With the same ASIC design and specification required across the sensors for DAQ homogeneity and cost-saving on design, the extreme variations a smaller ASIC would bring is undesirable. Also, a greater number would be required per sensor to retain acceptance, likely increasing costs further.

Figure 6.26 also explicitly displays that increasing r_{min} from 5.1 mm to 7.2 mm for VELO Upgrade II reduces the innermost pixel particle rate by a factor of two, reflecting the working points in Figure 6.14 used to derive the initial “Scenario X”. To halve the particle rate again, r_{min} would have to be increased to ≈ 11.5 mm, approaching the S_B or S_D working point discussed, with a significant degradation in performance shown for VELO UII sensors operating at this high radius (see Scenario D). This would also only yield a small data-rate saving relative to the initial reduction from 5.1 mm, implying that the current VELO UII design is close to or at the performance-data rate optimum as a function of r_{min} .

Figure 6.26 can also be utilised to project the expected fluence as a function of radius using simulation. If the particle rate per pixel is scaled to area (49.6 particles/cm²/event at $r = 7.2$ mm), and assuming, in the high luminosity scenario, an overall $\mathcal{L}_{int} = 300 \text{ fb}^{-1}$, the projected particle rate for the innermost pixel reaches 2.84×10^{16} particles/cm² across the VELO Upgrade II lifetime. Using a fluence scaling of 1 particle = 0.9 1 MeV n_{eq} /cm² for Station 15 [197], the peak fluence for Run 5+6 for the innermost pixels in the current VELO Upgrade II design corresponds to 2.52×10^{16} 1 MeV n_{eq} /cm², inline with the previous analytical calculations [24], reducing to 1.6×10^{16} 1 MeV n_{eq} /cm² for the lower luminosity working point. This renders 3D sensors as the primary technology choice for VELO UII sensors, with the LGAD best case radiation-hardness of 6×10^{15} 1 MeV n_{eq} /cm², requiring either $r_{min} > 12$ mm or complete sensor replacement every two years of data taking even at the lower luminosity working point, both considered unfeasible for VELO performance and cost.

Upon the implementation of a more sophisticated simulation of the sensor technology choice, emulation of clustering and radiation damage can be integrated into the VELO UII simulation allowing a more detailed study of the expected performance for operating a fast-timing, radiation-hard detector in HL-LHC conditions. The studies shown in this section have however been instructive and allowed rapid evolution and understanding of a constantly changing design, providing approximate metrics and relative performance without the need for a detailed, time-consuming simulation chain.

6.5 Performance Optimisation Tool for LHCb Upgrade II

While individual subdetector design and optimisation is vital to a performant LHCb Upgrade II detector, of ultimate importance is the global optimisation of the detector, ensuring that complementary design choices are made such that the detector as a whole meets the requirements of an experiment to exploit the full forward physics potential of the HL-LHC. In the short term this poses issues due to subdetector projects working separately with no central, shared simulation.

The following section will present a new framework designed to enable the execution of prototype simulation on a large-scale using the Worldwide LHC Computing Grid (WLCG) [243, 244] with a focus on transparent record of applications, configurations and the results produced. The primary use case is LHCb Upgrade II, to aid with the production of results during the design phase of the detector upgrade, providing a tool to ease collaboration by hosting the building of applications, running of tests and presentation of results. Example use cases for the VELO Upgrade II will be presented along with an initial attempt to combine and produce results from the UII tracking system.

6.5.1 LHCb Performance and Regression System

The implementation of the new framework is modelled around the existing LHCb Performance and Regression (LHCbPR) system [245], which facilitates regular testing of LHCb Software components, providing systematic profiling and comparison of past results to understand changes made during continuous development of the LHCb architecture [246]. In summary, the system collects “nightly stacks”, curated chains of LHCb software that are rebuilt every day with the latest mixture of versions, and initiates a set of test jobs at regular intervals that study the current performance of simulation, digitisation, reconstruction and data acquisition before processing the results and uploading them to a

central website, enabling analysis and comparison of different variations on each job as well as retaining a historic record. The tool is currently used to identify problems with and track the development of LHCb Run 1, 2 and 3 software stacks as well as monitoring the trigger software, vital for Run 3 performance.

However, the bandwidth of LHCbPR is limited. It currently operates on four dedicated machines at CERN and as the number of stacks/jobs increases and complexity of each develops, limits on sample sizes being generated/tested have been imposed. While this remains feasible for LHCb Run 1–4 (if undesirable), for UII simulation the increased luminosity will have a knock-on effect on the runtime of jobs, and it is likely that simultaneous testing of Run 1+2, Run 3/4 and Run 5 tests will be required in the near future. The following Section 6.5.2 will describe the implementation of an extension to LHCbPR to enable use of the WLCG to process the more complex jobs that could benefit from increased sample-sizes and/or running time, taking strain from the central system.

In addition to bandwidth limits, the LHCbPR tests in their current form utilise central authorised software stacks only, maintained by the LHCb Core Software team. For a user to implement a change to one of these stacks and realise it in a LHCbPR test, a series of tests and rigorous merge request approvals required. This process is unsuitable for UII studies currently, with a focus on rapid iteration of designs, producing relative performance estimates without the requirements of full detailed simulation. In addition, no complete simulation stack for UII exists up to now, with each subdetector working independently, and only early progress realised to combine studies. The following sections will expand on developments to implement a solution to this using “Custom Stack Builds”, where users can specify software mixtures and make quick changes, independently to the central stacks, while still launching tests using the WLCG enabling large-scale jobs, stretching the reach of UII simulation studies that up to now are fulfilled privately. This development also retains the ethos of LHCbPR, with a full history of the stacks built and jobs initiated as well as using the same website to present the final results for full transparency to the collaboration.

6.5.2 LHCbPR Development for an Upgrade II Use Case

From an operational point of view, two new systems were developed to manage the building, submission and running of standard and custom stack based WLCG jobs. These are combined into one service with two distinct components, operating standalone from the CERN-based LHCbPR system on a machine hosted at the University of Birmingham, able to plug-in the final results to the central website when development is complete.

6.5.2.1 LHCbPR Grid Implementation

An extension to the LHCbPR system enabling jobs to be run on the WLCG was already in active development before the UII use case was derived. The initial idea was a coherent system drawing from the central LHCbPR schedule, isolating suitable jobs, converting the existing configuration details to a submission file for Ganga (a standard job submission tool used at LHCb [247]), which in turn automatically submits and monitors job running on the WLCG. The primary advantage was the in-built ability to split jobs into many sub-jobs, with this feature native to Ganga, before merging the outputs retaining compatibility with the subsequent LHCbPR system. This enables jobs to seamlessly run on more input files or generate bigger samples, improving the benchmarking of simulation and the precision of monitors for trigger development, which were limited by the sample size previously.

The WLCG implementation with Ganga also lent itself naturally to an extension with “chained” jobs for LHCbPR. Up to now a single LHCbPR job operated with a single application on a specific stack, but in normal experimental data-flow many applications are often used in a chain. Ganga natively manages the organisation and submission of this, which was implemented into LHCbPRGrid by specifying the unique title of the required job of the previous step. In testing, this enabled a chain of Simulation → Digitisation → Reconstruction to be performed with LHCbPR, with only the output from Reconstruction needing to be retained. With separate configuration files for each step, utilising different stacks is trivial (for example, simulating with a released version and reconstructing with an in-development “nightly” stack). Example configuration files including chained jobs are displayed in Figure E.1.

6.5.2.2 Custom Stack Builder

In order to facilitate rapid, large-scale testing for LHCb Upgrade II a “Custom” LHCb stack builder was implemented into the LHCbPRGrid system. This utilised the same system users have access to, allowing curation of a specific LHCb software stack to be built and run privately [248]. Each new custom stack is specified by a configuration file of Git commits identifying the exact mixture of repositories required along with other build information (see example in Figure 6.27), before building on dedicated machines at the University of Birmingham. This is followed by a novel process of cutting away the majority of unnecessary source code and artefacts, allowing packaging of the stack into a relatively small executable (from 5–10 GB to < 500 MB. These are subsequently uploaded to WLCG storage, where the discussed LHCbPRGrid jobs can download the stack, requested via the configuration file (e.g. `'stackName': 'U2_Trk'` in Figure E.1),

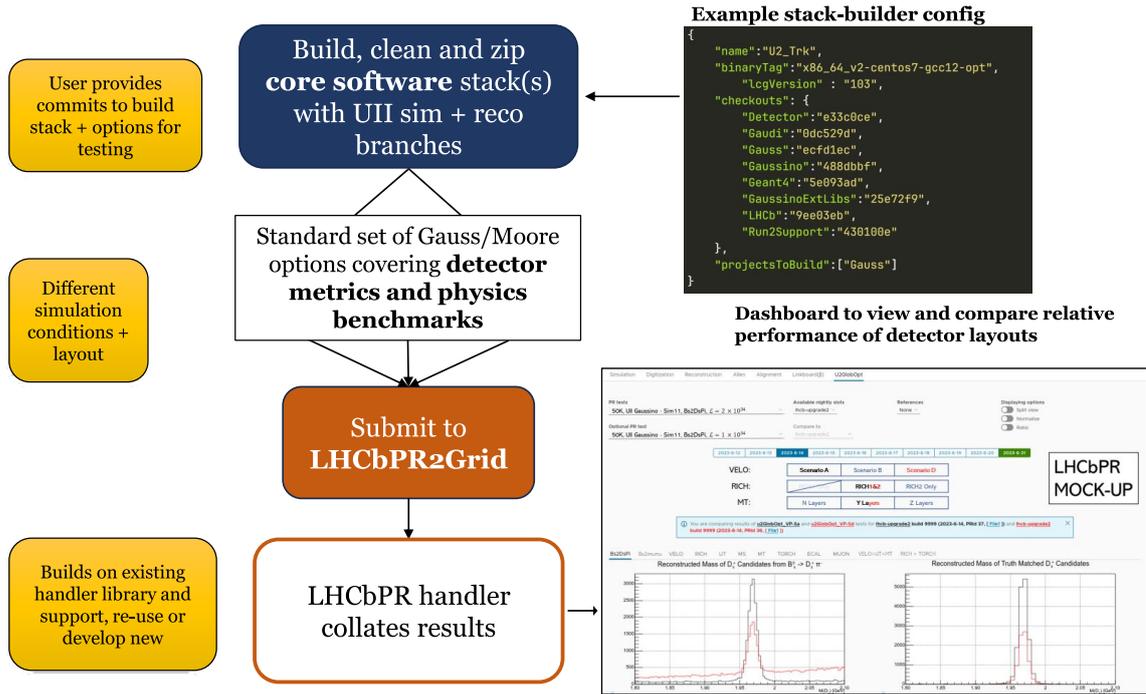


Figure 6.27: Summary of framework workflow from constructing a custom stack, through to running jobs on the WLCG before collating and uploading results to an online dashboard. The dashboard shown is a mock-up of the intended final design, with the current state of the working prototype displayed in Figure 6.29.

and run the specified job on the WLCG as if it were released or a central software stack.

With this method, adjustments to a stack can be implemented and reflected in the packaged applications quickly before being tested using LHCbPRGrid jobs, producing updated results with large statistics within 1–2 days, rather than the extensive review and approval procedure for a central software stack. The method also reduces the need for computationally intensive stack builds by individual users, usually on shared clusters, and further saves computing resources by naturally enabling the sharing of stacks, with a clear record via the configuration files of the exact contents of the application. The custom stacks are also compatible with the chained LHCbPRGrid jobs, either by building with multiple applications (example in Section 6.5.3), or chaining one stage using central software with another that uses a custom stack (see Section 6.5.4). This allows smooth transfer to released and central software stacks if/when these become available.

6.5.2.3 User Interface and UII Approach

The implementation of the LHCbPR2Grid system and Custom Stack Builder enables the workflow summarised in Figure 6.27. From the user side, a stack configuration file is

curated, producing a custom software stack. The user then provides the configuration for a job or set of jobs that use this stack along with an “LHCbPR Handler” [245] to process any output, with the results automatically uploaded to the discussed website where an “LHCbPR Dashboard” is configured to display these in a given format, with further example of the working prototype dashboard in Figure 6.29.

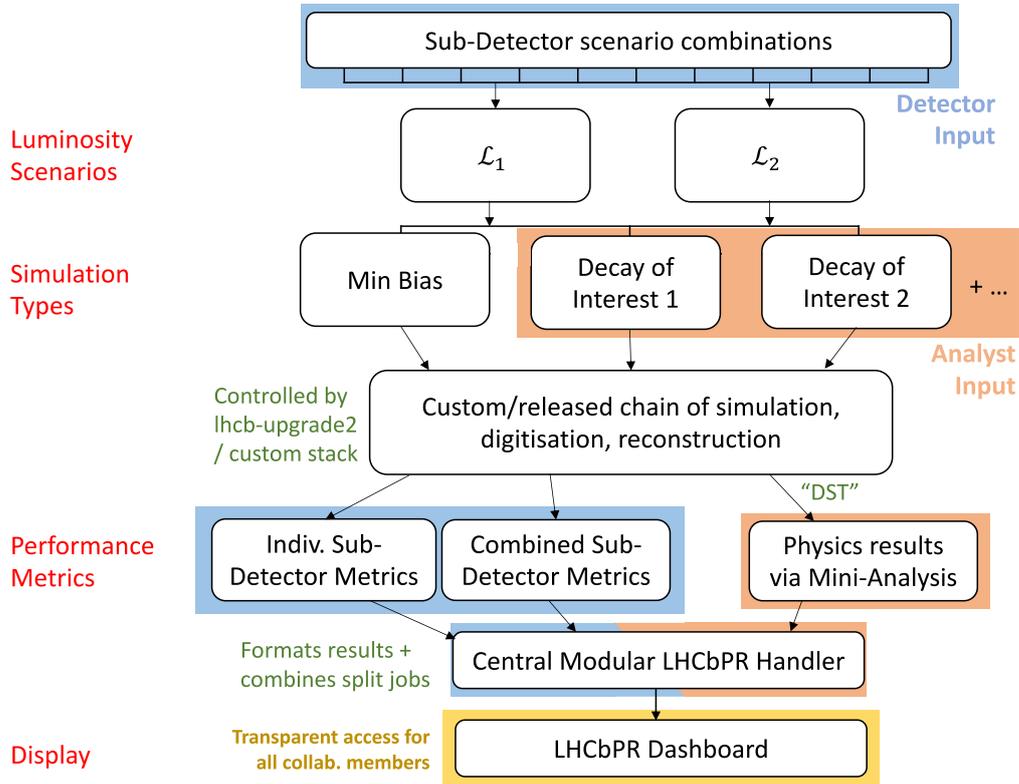


Figure 6.28: Flow diagram formulating a Global Optimisation approach for LHCb Upgrade II using the LHCbPRGrid and custom stack builder. The orange shading denotes inputs from physics analysts, blue from detector experts and yellow the interface with the wider collaboration.

Figure 6.28, describes how a “Global” optimisation approach for LHCb Upgrade II could be attained with the discussed system. Upon realisation of a distinct set of realistic “Subdetector scenario combinations” and their geometry/reconstruction differences implemented, a custom stack could be designed and distributed as described in Section 6.5.2.2. A series of jobs could then be set-up to test each combination with different simulation types shown, all conducting relevant analysis with shared scripts simplifying the implementation. Splitting into sub-jobs is vital for this scale of analysis to be feasible on a semi-regular basis, especially for the simulation of specific physics decay-modes which require large samples. The subsequent “Central Modular LHCbPR Handler” would uniformly format the results and upload them to a fully-equipped LHCbPR dashboard, where study and comparison of results for different simulation types, luminosity scenarios

and subdetector combinations could be made by all interested, a mock-up of which is shown in Figure 6.27.

This tool, realised as stated, would be able to identify design combinations that benefit certain metrics over others, potentially revealing issues or advantages not foreseen by studying each subdetector and simulation type individually. It is therefore ideal for optimisation of the LHCb Upgrade II detector design as the collaboration moves towards the Technical Design Report, requiring a conclusion and projection of global performance in terms of physics reach, rather than individual subdetector metrics and performance extrapolations.

Sections 6.5.3 and 6.5.4 describe the initial steps towards this combined detector tool by initially studying VELO performance as a function of RF shield design before performing a first performance study of the combined Upgrade II Tracking system.

6.5.3 Results - VELO UII Scenario X

Figure 6.29 displays a selection of results from using the custom stack tester and LHCbPRGrid tool to generate a scheme of VELO Upgrade II designs where the RF shield is altered from the nominal described in Table 6.3. A custom stack is curated with each shield design implemented in the detector geometry, with a customised simulation and reconstruction chain, similar to that discussed in Section 6.4.2, formed of two LHCbPRGrid jobs chained together. By splitting the jobs into 100 sub-jobs, 10000 events are generated for each design, with a runtime of < 2 hours.

The selected results in Figure 6.29 compare the performance for VELO Upgrade II if a corrugated RF shield design is re-implemented, similar to that of the Run 3 VELO. The increase in material from a corrugated design degrades the IP resolution overall, with improvements only shown at high track η due to a particle on average traversing the shield closer to the normal for a corrugated shield. The distributions of average number of hits per track look as expected in η but identified a potential gap in the VELO acceptance at $\frac{\pi}{6}$ (and $-\frac{5\pi}{6}$), which prompted cross-checks and determined that this is due to an intentional 1 mm gap between the outer perpendicular sensors in each VELO module [197]. This displays one use case of the dashboard, automatically displaying many metrics which may not have been considered, to identify problems or features of interest.

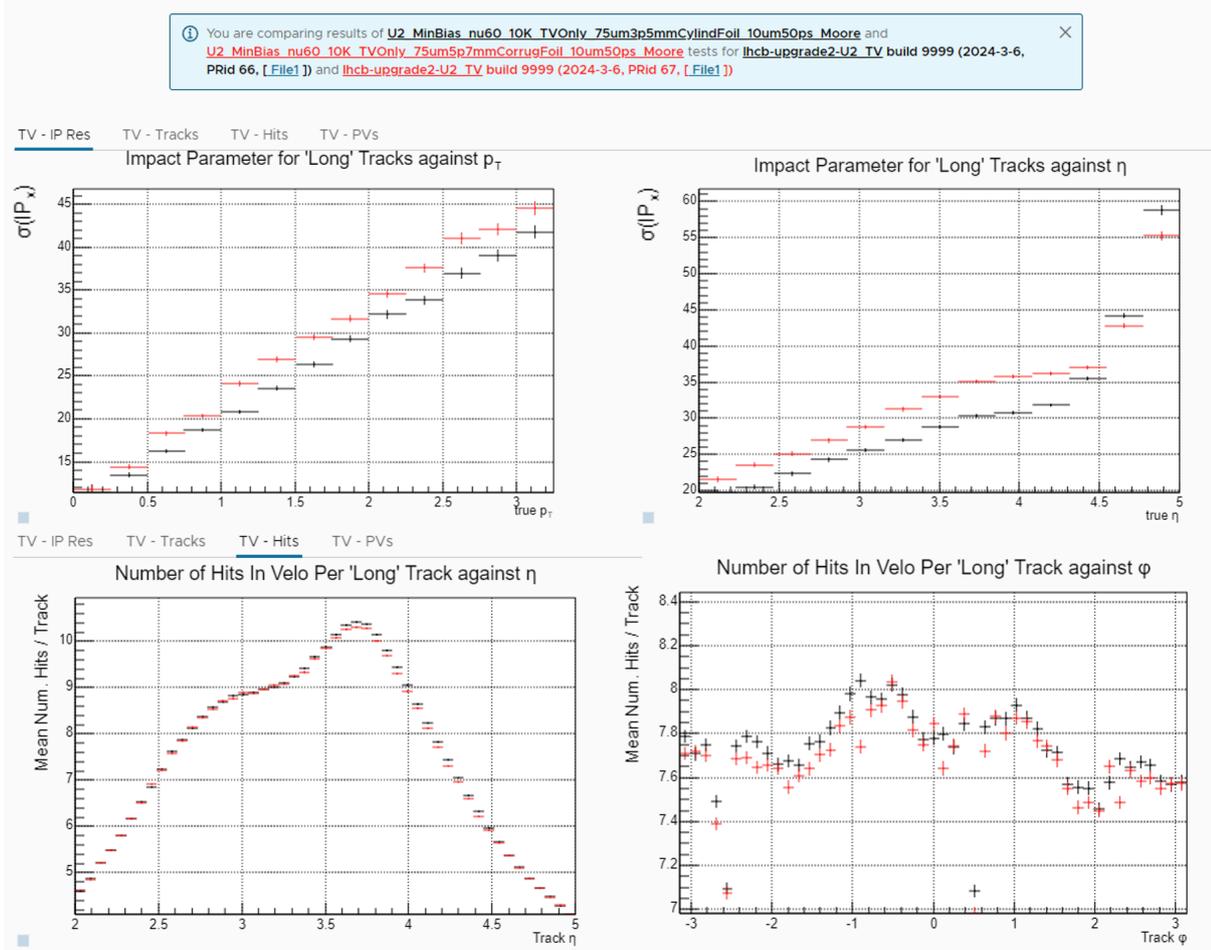


Figure 6.29: Screenshots from prototype LHCbPR dashboard for Upgrade II, displaying VELO (referred to as “Timing Velo” (TV) in the figure) studies of IP resolution as a function of both $1/p_T$ (upper left) and track η (upper left) as well as a measure of the average number of hits per track as function of track η (lower left) and ϕ (lower right). The results compare the VELO model discussed in Section 6.4.2.3 with different RF shield designs: 75 μm thick cylindrical shield with radius 3.5 mm (black) and an alternative design with a corrugated shape similar to UI, with 75 μm thickness at the closest point to the modules, 5.7 mm from the z -axis (red).

6.5.4 Results - UII Tracking System

To move towards a global optimisation system, an investigation of changing the UII tracking components is attempted as a first look at inter-subdetector performance with core simulation software for Upgrade II at LHCb. A custom stack is curated to simulate the combined UII VELO, UT and MT designs, along with the beam pipe and LHCb magnet (corresponding to the configuration stated in Figure 6.27). The simulation stage was chained with the recently implemented “lhcb-run5” nightly stack for reconstruction, using a central UII reconstruction script, automatically inheriting development from other collaborators as updates are made. For each combination 2500 events were generated, with

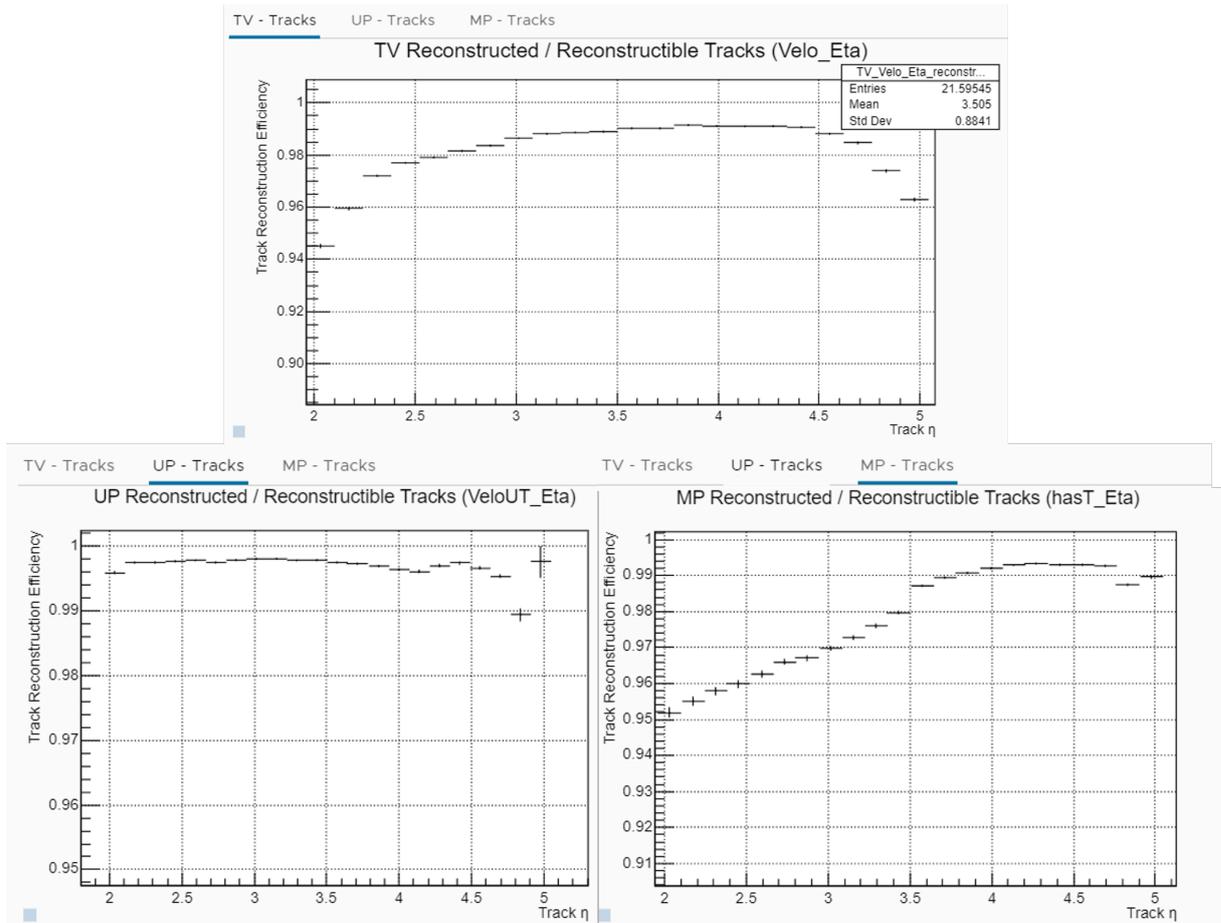


Figure 6.30: Screenshots from prototype LHCbPR dashboard for Upgrade II, displaying track reconstruction efficiencies as a function of η for the three primary tracking detectors VELO (upper), Upstream Tracker (lower left) and the pixel segment of the Mighty Tracker (lower right). Each displaying tracks reconstructed within the corresponding subdetector.

an example of one job configuration displayed in [E.1](#).

Figure 6.30 describes individual tracking efficiencies for the three subdetectors. Retaining individual metrics like these is important to monitor subdetector performance while the global metrics are investigated. The results shown are compatible with offline tests by subdetector experts.

Figure 6.31 studies a more global metric by monitoring the number of hits in the MT as a function of material in the UT. The immediate expectation would be that if the material in the UT is reduced, the number of secondary particles produced in material interactions would also, resulting in reduced occupancy in the MT layers and stations. By simulating a layout of the UT reduced from four planes to three (by scaling all material densities by a factor of 0.75), and comparing it to the nominal design, only minor reduction in occupancy was realised. This supported an external investigation that had concluded

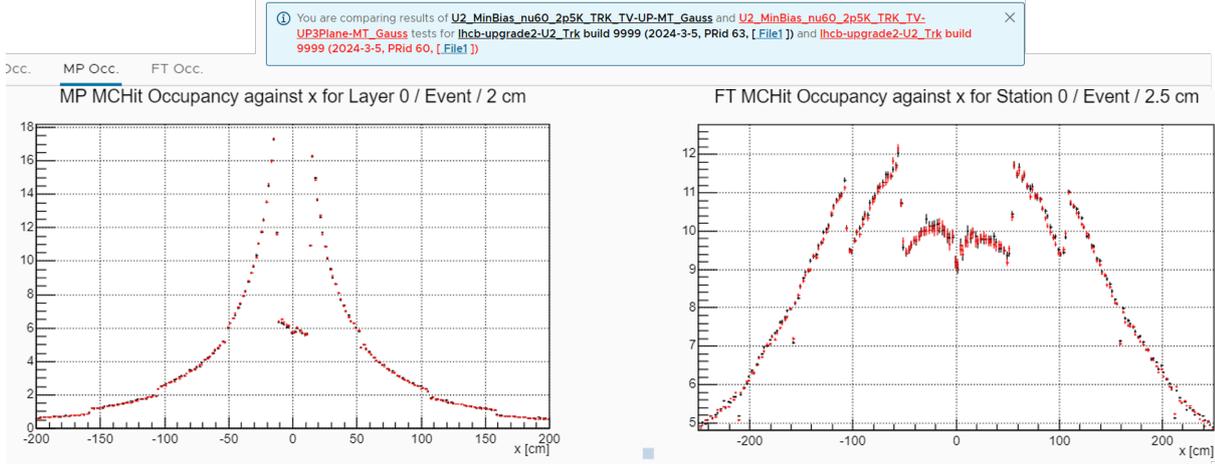


Figure 6.31: Screenshots from prototype LHCbPR dashboard for Upgrade II, displaying “Occupancy” in the MT central pixel component (left) and the outer fibre region (right), measuring the affect of the change in material of the UT with four planes (black) and three planes (red). Occupancy here refers to the total number of hits per layer or station of the MT per event.

that, for UII conditions, showering from material interaction with the beam pipe is the dominant contributor to occupancy in the MT.

While these are only initial studies, there is clear potential for expansion with the displayed set-up. A primary use case would be monitoring long track reconstruction as a function of tracking detector design, vital for LHCb physics performance [24]. The results displayed in Figures 6.30 and 6.31 also show a route to broaden the analysis to a complete detector description, with optimisation based on a distinct set of realistic subdetector designs. For this to be realised significant collaboration will be required, covering all subdetectors as well as drawing expertise from physics analysis teams to implement results describing LHCb physics benchmarks. Only then can actual global performance be determined, paramount for presenting LHCb Upgrade II as the only way to exploit the full forward physics potential of the HL-LHC.

Chapter 7

Conclusions

The LHC at CERN is the world-leading particle physics facility offering unprecedented sensitivity and energy reach, with LHCb its primary flavour physics experiment that has successfully become a general purpose detector in the forward region. While the dataset collected by LHCb during Run 1 and 2 of the LHC has produced significant contributions to both SM measurements and NP searches, the experiment must evolve as the LHC upgrades, with the HL-LHC designed to produce more luminous collisions and increase the discovery potential of the facility. This thesis explored the current and future status of LHCb, presenting the first search for cLFV in the process $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ using the LHCb Run 1 and 2 dataset before investigating the future of the VELO and LHCb detector as a whole as R&D and design work for LHCb Upgrade II progresses.

Chapters 3–5 of this thesis described the world-first search for the forbidden decay $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$. This process would require cLFV to produce a final state with two different flavour leptons and hence any significant measurement would be unequivocal, direct evidence for NP. Chapter 4 presented how the LHCb Run 1 and 2 detector was used to refine the available dataset into an ultra-low background environment maximising sensitivity to cLFV. The development of data-driven corrections to the analysis simulation samples is documented in Section 4.2, allowing confident use of simulation to validate the selection chain and calculate accurate inputs to Equation 3.2.1. With this Equation the well-understood and high-rate process $\Lambda_b^0 \rightarrow pK^- J/\psi(\rightarrow \mu^+\mu^-)$ was utilised as a control mode (see Section 5.3), scaling any potential signal contribution in the $M(pK^-\mu^\pm e^\mp)$ mass spectrum to measure $\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp)$. Currently the signal region for $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ is blinded, but evaluation of the data sidebands followed by extrapolation allows an estimated upper limit to be derived using the CL_s method. This calculates the branching fraction that must be exceeded in order to measure a significant contribution from potential $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ events, excluding larger values if no signal is observed

upon unblinding. Initial upper limit estimates were calculated in Section 5.5, yielding

$$\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp) \lesssim 2.8 \times 10^{-8} \text{ @ 95\% confidence level,} \quad (7.0.1)$$

competitive with other cLFV searches at LHCb. This provides a further constraint on the forbidden decays of b -hadrons that may be able to inform future BSM models, providing a complementary result to b -meson searches shown in Figure 1.5. A study of b -baryons as presented would only have been possible at LHCb with its unique ability to collect a substantial Λ_b^0 sample and make exclusive selections in a low background environment.

While the search for $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ remains blinded, and any hints or evidence of signal would transform the analysis outlook, previous cLFV measurements suggest that a limit similar to Equation 7.0.1 will become the final result, the world-first upper limit on this decay mode.

The analysis measurement is ultimately statistically limited due to the low number of events in each analysis category for the final background fits, with $< 15\%$ systematic uncertainty expected from the priority elements already calculated and insight from similar analyses. Therefore as the LHCb dataset grows there is scope for enhancement of a search for $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ to make this limit more stringent, potentially providing strict constraints on BSM models if not observing cLFV. After installation of LHCb Upgrade II during LS4 and data-taking through Run 5 and 6, the LHCb dataset will reach its ultimate form of $\mathcal{L}_{int} = 350 \text{ fb}^{-1}$, a factor 40 increase from Run 1 and 2. This would significantly improve the precision of α in Equation 3.2.2 and allow more intricate and tighter selections such as more stringent vetoes or an increased p_{MVA} cut value. The increased statistics would also allow a more detailed final background fit that will improve the CL_s method stability. This will undoubtedly improve the sensitivity of the analysis overall, and a simple projection yields

$$\mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp)_{\text{UII pred.}} = \mathcal{B}(\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp)_{\text{est.}} / \sqrt{40} \lesssim 4 \times 10^{-9}. \quad (7.0.2)$$

This is a conservative estimate showing the minimum improvement from uncertainties scaling by $\sqrt{\mathcal{L}_{int}}$, yet still excluding the example NP model prediction in Equation 1.3.2 [101]. This pre-empts a more refined analysis strategy and improved techniques such as a more precise determination of the backgrounds that could further enhance the limit.

Ultra-sensitive cLFV searches are one example of key LHCb observables that will be transformed by the unprecedented sensitivity LHCb Upgrade II offers, but to deliver on the proposals all subdetectors require a significant upgrade to derive benefit from operation at $\mathcal{L} = 1.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. Chapter 6 of this thesis presented progress in both R&D and simulation for VELO Upgrade II, which is required to reconstruct up

to 40 PVs and 2000 tracks per bunch crossing during Run 5 and 6 of the LHC, with the addition of precise time resolution the primary upgrade. Section 6.3 presented a feasibility investigation for using LGADs from Micron Semiconductor Ltd. as the sensor technology for VELO UII, performing characterisation of a set of devices after irradiation to $\approx 9 \times 10^{14}$ 1 MeV $n_{\text{eq}}/\text{cm}^2$. The study showed significant reduction of device performance after irradiation, with gain reduced by up to a factor of 5 and timing performance projected to degrade below the minimum requirements for VELO UII. Therefore, considering the absolute minimum fluence requirement of 8×10^{15} 1 MeV $n_{\text{eq}}/\text{cm}^2$, the LGAD design tested was deemed unsuitable for VELO Upgrade II.

The derivation and iteration of VELO UII designs requires simulation to understand how technology choices and R&D may impact detector and physics performance. Section 6.4 this thesis displayed a simplified simulation approach allowing changes in design or reconstruction to be implemented and tested rapidly. The work was instrumental in concluding that an optimised VELO UII design with an inner radius $r_{\text{min}} \approx 7.2$ mm is currently the best approach for building a performant detector for HL-LHC conditions, considering technological and mechanical limitations. This work also showed that significant descoping of the VELO UII design would dramatically degrade the performance of the entire LHCb detector, hence it is not considered as the LHCb Upgrade II design moves towards the TDR stage.

Finally, Section 6.5 presented the development of a tool to perform flexible performance optimisation for the entire LHCb Upgrade II detector, with results for the UII tracking system shown as the first investigation of inter-subdetector performance using LHCb core simulation software. The tool will be extended to the complete detector description and used to produce results that describe LHCb physics benchmarks automatically, a required progression from current individual subdetector studies. This is vital for presenting LHCb Upgrade II as the only way to exploit the full forward physics potential of the HL-LHC.

References

- [1] J. J. Thomson, *Cathode rays*, *Phil. Mag. Ser. 5* **44**, 293–316 (1897).
- [2] C. D. Anderson, *The Apparent Existence of Easily Deflectable Positives*, *Science* **76**, 238–239 (1932), eprint: <https://www.science.org/doi/pdf/10.1126/science.76.1967.238>.
- [3] P. M. S. Blackett and G. P. S. Occhialini, *Some photographs of the tracks of penetrating radiation*, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **139**, 699–726 (1933).
- [4] CERN, *Press Release. Le synchro-cyclotron européen en 1961*, CERN-PR-62-06-EN, CERN-PR-62-06-FR (1962), <https://cds.cern.ch/record/840686>.
- [5] O. S. Brüning *et al.*, *LHC Design Report*, CERN Yellow Reports: Monographs (CERN, Geneva, 2004), doi:10.5170/CERN-2004-003-V-1.
- [6] CDF Collaboration, F. Abe *et al.*, *Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab*, *Phys. Rev. Lett.* **74**, 2626–2631 (1995).
- [7] D0 Collaboration, S. Abachi *et al.*, *Observation of the top quark*, *Phys. Rev. Lett.* **74**, 2632–2637 (1995).
- [8] ALEPH, DELPHI, L3, OPAL, SLD, LEP Electroweak Working Group, SLD Electroweak Group, SLD Heavy Flavour Group, S. Schael *et al.*, *Precision electroweak measurements on the Z resonance*, *Phys. Rept.* **427**, 257–454 (2006), arXiv:hep-ex/0509008.
- [9] The ALEPH, DELPHI, L3, OPAL Collaborations, the LEP Electroweak Working Group, *Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP*, *Phys. Rept.* **532**, 119 (2013), arXiv:1302.3415.
- [10] CMS Collaboration, S. Chatrchyan *et al.*, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, *Phys. Lett. B* **716**, 30–61 (2012), arXiv:1207.7235.
- [11] 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*, *Phys. Lett. B* **716**, 1–29 (2012), arXiv:1207.7214.
- [12] E. Corbelli and P. Salucci, *The Extended Rotation Curve and the Dark Matter Halo of M33*, *Mon. Not. Roy. Astron. Soc.* **311**, 441–447 (2000), arXiv:astro-ph/9909252.

- [13] Supernova Search Team, A. G. Riess *et al.*, *Observational evidence from supernovae for an accelerating universe and a cosmological constant*, *Astron. J.* **116**, 1009–1038 (1998), [arXiv:astro-ph/9805201](#).
- [14] S. M. Carroll, W. H. Press, and E. L. Turner, *The cosmological constant*. *Annual Rev. Astron. Astrophys.* **30**, 499–542 (1992).
- [15] R. Acciarri *et al.*, *Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE) Conceptual Design Report Volume 1: The LBNF and DUNE Projects*, 2016, [arXiv:1601.05471](#).
- [16] F. D. Lodovico and on behalf of the Hyper-Kamiokande Collaboration, *The Hyper-Kamiokande Experiment*, *Journal of Physics: Conference Series* **888**, 012020 (2017).
- [17] FCC, A. Abada *et al.*, *FCC Physics Opportunities: Future Circular Collider Conceptual Design Report Volume 1*, *Eur. Phys. J. C* **79**, 474 (2019).
- [18] M. Aicheler *et al.*, *A Multi-TeV Linear Collider Based on CLIC Technology: CLIC Conceptual Design Report*, CERN Yellow Reports: Monographs (CERN, Geneva, 2012), doi:[10.5170/CERN-2012-007](#).
- [19] T. Behnke *et al.*, *The International Linear Collider Technical Design Report - Volume 1: Executive Summary*, 2013, [arXiv:1306.6327](#).
- [20] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, *Pisma Zh. Eksp. Teor. Fiz.* **5**, 32–35 (1967).
- [21] LHCb, *LHCb: letter of intent*, CERN-LHCC-95-5, CERN-LHCC-1995-005, LHCC-I-8 (CERN, Geneva, 1995), <https://cds.cern.ch/record/290868>.
- [22] LHCb Collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030 (CERN, Geneva, 2003).
- [23] LHCb Collaboration, *Framework TDR for the LHCb Upgrade: Technical Design Report*, CERN-LHCC-2012-007 (CERN, Geneva, 2012).
- [24] LHCb Collaboration, *LHCb Framework TDR for the LHCb Upgrade II Opportunities in flavour physics, and beyond, in the HL-LHC era*, CERN-LHCC-2021-012 (CERN, Geneva, 2022).
- [25] K. Akiba *et al.*, *Considerations for the VELO detector at the LHCb Upgrade II*, LHCb-PUB-2022-001, CERN-LHCb-PUB-2022-001 (CERN, Geneva, 2022).
- [26] M. Thomson, *Modern particle physics* (Cambridge University Press, New York, 2013), doi: [10.1017/CBO9781139525367](#).
- [27] R. P. Feynman, *Mathematical formulation of the quantum theory of electromagnetic interaction*, *Phys. Rev.* **80**, 440–457 (1950).
- [28] R. P. Feynman and M. Gell-Mann, *Theory of the Fermi Interaction*, *Physical Review* **109**, 193–198 (1958).
- [29] E. C. G. Sudarshan and R. E. Marshak, *Chirality invariance and the universal fermi interaction*, *Phys. Rev.* **109**, 1860–1862 (1958).
- [30] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, *Phys. Rev. D* **2**, 1285–1292 (1970).

- [31] A. Salam, *Weak and Electromagnetic Interactions*, *Conf. Proc. C* **680519**, 367–377 (1968).
- [32] R. P. Feynman, *Space–time approach to quantum electrodynamics*, *Phys. Rev.* **76**, edited by L. M. Brown, 769–789 (1949).
- [33] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13**, edited by J. C. Taylor, 508–509 (1964).
- [34] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, *Phys. Lett.* **12**, 132–133 (1964).
- [35] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13**, edited by J. C. Taylor, 321–323 (1964).
- [36] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, *Phys. Rev. Lett.* **30**, edited by J. C. Taylor, 1343–1346 (1973).
- [37] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, *Phys. Rev. Lett.* **30**, edited by J. C. Taylor, 1346–1349 (1973).
- [38] D. J. Gross and F. Wilczek, *Asymptotically Free Gauge Theories - I*, *Phys. Rev. D* **8**, 3633–3652 (1973).
- [39] H. Yukawa, *On the Interaction of Elementary Particles I*, *Proc. Phys. Math. Soc. Jap.* **17**, 48–57 (1935).
- [40] S. Weinberg, *A model of leptons*, *Phys. Rev. Lett.* **19**, 1264–1266 (1967).
- [41] P. D. Group, R. L. Workman, *et al.*, *Review of Particle Physics*, *Progress of Theoretical and Experimental Physics* **2022**, 10.1093/ptep/ptac097 (2022).
- [42] P. N. Swallow, *Searching for Charged Lepton Flavour Violation at LHCb and Long-Lived Particles with CODEX-b*, PhD thesis (Birmingham U., Aug. 2023).
- [43] T. D. Lee and C.-N. Yang, *Question of Parity Conservation in Weak Interactions*, *Phys. Rev.* **104**, 254–258 (1956).
- [44] D. Griffiths, *Introduction to elementary particles* (John Wiley & Sons, 2008).
- [45] K. G. Wilson, *Confinement of Quarks*, *Phys. Rev. D* **10**, edited by J. C. Taylor, 2445–2459 (1974).
- [46] LHCb collaboration, R. Aaij *et al.*, *Observation of the resonant character of the $Z(4430)^-$ state*, *Phys. Rev. Lett.* **112**, 222002 (2014), [arXiv:1404.1903](#).
- [47] BESIII Collaboration, M. Ablikim *et al.*, *Observation of a Charged Charmoniumlike Structure in $e^+e^- \rightarrow \pi^+\pi^- J/\psi$ at $\sqrt{s}=4.26$ GeV*, *Phys. Rev. Lett.* **110**, 252001 (2013).
- [48] Belle Collaboration, Z. Q. Liu *et al.*, *Study of $e^+e^- \rightarrow \pi^+\pi^- J/\psi$ and Observation of a Charged Charmoniumlike State at Belle*, *Phys. Rev. Lett.* **110**, 252002 (2013).
- [49] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \rightarrow J/\psi p K^-$ decays*, *Phys. Rev. Lett.* **115**, 072001 (2015), [arXiv:1507.03414](#).
- [50] Belle-II, T. Abe *et al.*, *Belle II Technical Design Report*, (2010), [arXiv:1011.0352](#).

- [51] BESIII, M. Ablikim *et al.*, *Design and Construction of the BESIII Detector*, *Nucl. Instrum. Meth. A* **614**, 345–399 (2010), [arXiv:0911.4960](#).
- [52] N. Cabibbo, *Unitary symmetry and leptonic decays*, *Phys. Rev. Lett.* **10**, 531–533 (1963).
- [53] B. P. Roe *et al.*, *New Determination of the K^+ Decay Branching Ratios*, *Phys. Rev. Lett.* **7**, 346–348 (1961).
- [54] S. L. Glashow, *Charm: An Invention Awaits Discovery*, *AIP Conf. Proc.* **21**, edited by H. C. Wolfe and D. A. Garelick, 387–392 (1974).
- [55] SLAC-SP-017, J. E. Augustin *et al.*, *Discovery of a Narrow Resonance in e^+e^- Annihilation*, *Phys. Rev. Lett.* **33**, 1406–1408 (1974).
- [56] E598, J. J. Aubert *et al.*, *Experimental Observation of a Heavy Particle J* , *Phys. Rev. Lett.* **33**, 1404–1406 (1974).
- [57] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49**, 652–657 (1973).
- [58] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, *Phys. Rev. Lett.* **51**, 1945 (1983).
- [59] E. Fermi, *An attempt of a theory of beta radiation. 1.* *Z. Phys.* **88**, 161–177 (1934).
- [60] P. A. M. Dirac, *The Principles of Quantum Mechanics*, International series of monographs on physics (Oxford, England) (Oxford University Press, 1930).
- [61] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B_s^0 \rightarrow \mu^+\mu^-$ decay properties and search for the $B^0 \rightarrow \mu^+\mu^-$ and $B_s^0 \rightarrow \mu^+\mu^-\gamma$ decays*, *Phys. Rev.* **D105**, 012010 (2022), [arXiv:2108.09283](#).
- [62] M. Yoshimura, *Origin of Cosmological Baryon Asymmetry*, *Phys. Lett. B* **88**, 294–298 (1979).
- [63] W. Buchmuller, R. Ruckl, and D. Wyler, *Leptoquarks in Lepton - Quark Collisions*, *Phys. Lett. B* **191**, 442–448 (1987), [Erratum: *Phys.Lett.B* 448, 320–320 (1999)].
- [64] S. Meinel and G. Rendon, $\Lambda_b \rightarrow \Lambda^*(1520)\ell^+\ell^-$ form factors from lattice QCD, *Phys. Rev. D* **103**, 074505 (2021).
- [65] S. Biswas, S. Mahata, B. P. Nayak, and S. Sahoo, *Imprints of new physics in Λ_b^0 decays into $\Lambda^*(1520)\ell^+\ell^-$ in non-universal Z' model*, 2023, [arXiv:2310.09887](#).
- [66] M. Bordone, O. Catà, T. Feldmann, and R. Mandal, *Constraining flavour patterns of scalar leptoquarks in the effective field theory*, *JHEP* **03**, 122 (2021), [arXiv:2010.03297](#).
- [67] Muon g-2, B. Abi *et al.*, *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm*, *Phys. Rev. Lett.* **126**, 141801 (2021), [arXiv:2104.03281](#).
- [68] Muon g-2, D. P. Aguillard *et al.*, *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.20 ppm*, *Phys. Rev. Lett.* **131**, 161802 (2023), [arXiv:2308.06230](#).
- [69] T. Aoyama *et al.*, *The anomalous magnetic moment of the muon in the Standard Model*, *Phys. Rept.* **887**, 1–166 (2020), [arXiv:2006.04822](#).

- [70] S. Borsanyi *et al.*, *Leading hadronic contribution to the muon magnetic moment from lattice QCD*, *Nature* **593**, 51–55 (2021), [arXiv:2002.12347](#).
- [71] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality in $b \rightarrow s\ell^+\ell^-$ decays*, *Phys. Rev. Lett.* **131**, 051803 (2023), [arXiv:2212.09152](#).
- [72] M. Bordone, G. Isidori, and A. Pattori, *On the Standard Model predictions for R_K and R_{K^*}* , *Eur. Phys. J. C* **76**, 440 (2016), [arXiv:1605.07633](#).
- [73] G. Isidori, S. Nabeebaccus, and R. Zwicky, *QED corrections in $\bar{B} \rightarrow \bar{K}\ell^+\ell^-$ at the double-differential level*, *JHEP* **12**, 104 (2020), [arXiv:2009.00929](#).
- [74] LHCb collaboration, R. Aaij *et al.*, *Measurement of lepton universality parameters in $B^+ \rightarrow K^+\ell^+\ell^-$ and $B^0 \rightarrow K^{*0}\ell^+\ell^-$ decays*, *Phys. Rev.* **D108**, 032002 (2023), [arXiv:2212.09153](#).
- [75] LHCb collaboration, R. Aaij *et al.*, *Tests of lepton universality using $B^0 \rightarrow K_S^0\ell^+\ell^-$ and $B^+ \rightarrow K^{*+}\ell^+\ell^-$ decays*, *Phys. Rev. Lett.* **128**, 191802 (2022), [arXiv:2110.09501](#).
- [76] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using $\Lambda_b^0 \rightarrow pK^-\ell^+\ell^-$ decays*, *JHEP* **05**, 040 (2020), [arXiv:1912.08139](#).
- [77] P. Koppenburg, *LHCb R Measurements Plot*, 2021, <https://www.nikhef.nl/~pkoppenb/anomalies.html>.
- [78] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality in beauty-quark decays*, *Nature Physics* **18**, 277 (2022), [arXiv:2103.11769](#).
- [79] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^0 \rightarrow K^{*0}\mu^+\mu^-$ decay using 3fb^{-1} of integrated luminosity*, *JHEP* **02**, 104 (2016), [arXiv:1512.04442](#).
- [80] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the rare decay $B_s^0 \rightarrow \phi\mu^+\mu^-$* , *JHEP* **11**, 043 (2021), [arXiv:2107.13428](#).
- [81] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^+ \rightarrow K^{*+}\mu^+\mu^-$ decay*, *Phys. Rev. Lett.* **126**, 161802 (2021), [arXiv:2012.13241](#).
- [82] LHCb collaboration, R. Aaij *et al.*, *Measurements of the S-wave fraction in $B^0 \rightarrow K^+\pi^-\mu^+\mu^-$ decays and the $B^0 \rightarrow K^*(892)^0\mu^+\mu^-$ differential branching fraction*, *JHEP* **11**, 047 (2016), [arXiv:1606.04731](#).
- [83] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of $\Lambda_b^0 \rightarrow \Lambda\mu^+\mu^-$ decays*, *JHEP* **06**, 115 (2015), [arXiv:1503.07138](#).
- [84] LHCb collaboration, R. Aaij *et al.*, *Branching fraction measurements of the rare $B_s^0 \rightarrow \phi\mu^+\mu^-$ and $B_s^0 \rightarrow f_2'(1525)\mu^+\mu^-$ decays*, *Phys. Rev. Lett.* **127**, 151801 (2021), [arXiv:2105.14007](#).
- [85] G. C. Branco, R. G. Felipe, and F. R. Joaquim, *Leptonic CP Violation*, *Rev. Mod. Phys.* **84**, 515–565 (2012), [arXiv:1111.5332](#).
- [86] R. D. Peccei and H. R. Quinn, *CP Conservation in the Presence of Instantons*, *Phys. Rev. Lett.* **38**, 1440–1443 (1977).
- [87] B. Pontecorvo, *Inverse beta processes and nonconservation of lepton charge*, *Zh. Eksp. Teor. Fiz.* **34**, 247 (1957).

- [88] R. Davis Jr., D. S. Harmer, and K. C. Hoffman, *Search for neutrinos from the sun*, *Phys. Rev. Lett.* **20**, 1205–1209 (1968).
- [89] Super-Kamiokande Collaboration, Y. Fukuda *et al.*, *Evidence for oscillation of atmospheric neutrinos*, *Phys. Rev. Lett.* **81**, 1562–1567 (1998), [arXiv:hep-ex/9807003](#).
- [90] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the Unified Model of Elementary Particles*, *Progress of Theoretical Physics* **28**, 870–880 (1962), doi: [10.1143/PTP.28.870](#).
- [91] Mu2e, L. Bartoszek *et al.*, *Mu2e Technical Design Report*, **10 . 2172 / 1172555** (2014), [arXiv:1501.05241](#).
- [92] Mu3e, K. Arndt *et al.*, *Technical design of the phase I Mu3e experiment*, *Nucl. Instrum. Meth. A* **1014**, 165679 (2021), [arXiv:2009.11690](#).
- [93] Mu3e, A.-K. Perrevoort, *Charged lepton flavour violation - Overview of current experimental limits and future plans*, *PoS DISCRETE2022*, 015 (2024).
- [94] NA62, E. Cortina Gil *et al.*, *Search for Lepton Number and Flavor Violation in K^+ and π^0 Decays*, *Phys. Rev. Lett.* **127**, 131802 (2021), [arXiv:2105.06759](#).
- [95] ATLAS, G. Aad *et al.*, *Searches for lepton-flavour-violating decays of the Higgs boson into $e\tau$ and $\mu\tau$ in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, *JHEP* **07**, 166 (2023), [arXiv:2302.05225](#).
- [96] ATLAS, G. Aad *et al.*, *Search for charged-lepton-flavour violating $\mu\tau qt$ interactions in top-quark production and decay in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC*, *Phys. Rev. D* **110**, 012014 (2024), [arXiv:2403.06742](#).
- [97] CMS, A. M. Sirunyan *et al.*, *Search for lepton-flavor violating decays of the Higgs boson in the $\mu\tau$ and $e\tau$ final states in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. D* **104**, 032013 (2021), [arXiv:2105.03007](#).
- [98] M. Neubert and C. T. Sachrajda, *Spectator effects in inclusive decays of beauty hadrons*, *Nucl. Phys. B* **483**, 339–370 (1997), [arXiv:hep-ph/9603202](#).
- [99] S. Jäger and J. Martin Camalich, *Reassessing the discovery potential of the $B \rightarrow K^*\ell^+\ell^-$ decays in the large-recoil region: SM challenges and BSM opportunities*, *Phys. Rev. D* **93**, 014028 (2016), [arXiv:1412.3183](#).
- [100] S. L. Glashow, D. Guadagnoli, and K. Lane, *Lepton Flavor Violation in B Decays?*, *Physical Review Letters* **114**, [10.1103/physrevlett.114.091801](#) (2015).
- [101] M. Bordone, M. Rahimi, and K. K. Vos, *Lepton flavour violation in rare Λ_b decays*, *Eur. Phys. J. C* **81**, 756 (2021), [arXiv:2106.05192](#).
- [102] M. Fedele, F. Wuest, and U. Nierste, *Renormalisation group analysis of scalar Leptoquark couplings addressing flavour anomalies: emergence of lepton-flavour universality*, *JHEP* **11**, 131 (2023), [arXiv:2307.15117](#).
- [103] A. Greljo, J. Salko, A. Smolkovič, and P. Stangl, *Rare b decays meet high-mass Drell-Yan*, *JHEP* **05**, 087 (2023), [arXiv:2212.10497](#).
- [104] M. Ciuchini *et al.*, *Constraints on lepton universality violation from rare B decays*, *Phys. Rev. D* **107**, 055036 (2023), [arXiv:2212.10516](#).

- [105] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^+\mu^-$ differential branching fraction*, *Phys. Rev. Lett.* **131**, 151801 (2023), [arXiv:2302.08262](#).
- [106] F. F. Deppisch, P. S. Bhupal Dev, and A. Pilaftsis, *Neutrinos and Collider Physics*, *New J. Phys.* **17**, 075019 (2015), [arXiv:1502.06541](#).
- [107] C. Cornella, J. Fuentes-Martin, and G. Isidori, *Revisiting the vector leptoquark explanation of the B-physics anomalies*, *JHEP* **07**, 168 (2019), [arXiv:1903.11517](#).
- [108] M. Neubert, *Effective field theory and heavy quark physics*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics in D ≥ 4* (Dec. 2005), pp. 149–194, [arXiv:hep-ph/0512222](#).
- [109] M. I. Ali, U. Chattopadhyay, N. Rajeev, and J. Roy, *SMEFT analysis of charged lepton flavor violating B-meson decays*, *Phys. Rev. D* **109**, 075028 (2024), [arXiv:2312.05071](#).
- [110] Belle Collaboration, S. Choudhury *et al.*, *Test of lepton flavor universality and search for lepton flavor violation in $B \rightarrow K\ell\ell$ decays*, *JHEP* **03**, 105 (2021), [arXiv:1908.01848](#).
- [111] LHCb collaboration, R. Aaij *et al.*, *Search for the lepton-flavour violating decays $B_{(s)}^0 \rightarrow e^\pm\mu^\mp$* , *JHEP* **03**, 078 (2018), [arXiv:1710.04111](#).
- [112] LHCb collaboration, R. Aaij *et al.*, *Search for the lepton-flavour violating decays $B^0 \rightarrow K^{*0}\mu^\pm e^\mp$ and $B_s^0 \rightarrow \phi\mu^\pm e^\mp$* , *JHEP* **06**, 073 (2023), [arXiv:2207.04005](#).
- [113] LHCb collaboration, R. Aaij *et al.*, *Search for the lepton-flavour violating decays $B^+ \rightarrow K^+\mu^\pm e^\mp$* , *Phys. Rev. Lett.* **123**, 231802 (2019), [arXiv:1909.01010](#).
- [114] N. Sahoo, *LHCb Lepton Flavour Violation Summary*, Private Communication, 2024.
- [115] LHCb collaboration, R. Aaij *et al.*, *Search for the lepton flavour violating decay $B^+ \rightarrow K^+\mu^-\tau^+$ using B_{s2}^{*0} decays*, *JHEP* **06**, 129 (2020), [arXiv:2003.04352](#).
- [116] LHCb collaboration, R. Aaij *et al.*, *Search for the lepton flavour violating decay $B^0 \rightarrow K^{*0}\tau^\pm\mu^\mp$* , *JHEP* **06**, 143 (2023), [arXiv:2209.09846](#).
- [117] LHCb collaboration, R. Aaij *et al.*, *Search for the lepton-flavor violating decay $B_s^0 \rightarrow \phi\mu^\pm\tau^\mp$* , (2024), [arXiv:2209.09846](#), Submitted to *Phys. Rev. D*.
- [118] LHCb collaboration, R. Aaij *et al.*, *Search for the lepton-flavour-violating decays $B_s^0 \rightarrow \tau^\pm\mu^\mp$ and $B^0 \rightarrow \tau^\pm\mu^\mp$* , *Phys. Rev. Lett.* **123**, 211801 (2019), [arXiv:1905.06614](#).
- [119] L. M. Greeven, *Decoding beauty: rare baryonic decays & SciFi detector commissioning*, PhD thesis (Maastricht U., 2024), doi:[10.26481/dis.20240403lg](#).
- [120] ATLAS, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3**, S08003 (2008).
- [121] CMS, S. Chatrchyan *et al.*, *The CMS Experiment at the CERN LHC*, *JINST* **3**, S08004 (2008).
- [122] ALICE, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, *JINST* **3**, S08002 (2008).

- [123] LHCb Collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, [JINST **3**, S08005 \(2008\)](#).
- [124] P. Giubellino, *Heavy Ion Physics at the LHC*, in 19th Hadron Collider Physics Symposium (Sept. 2008), [arXiv:0809.1062](#).
- [125] E. Lopienska, *The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022*, (2022) <https://cds.cern.ch/record/2800984>, General Photo.
- [126] L. Evans and P. Bryant, *LHC Machine*, [JINST **3**, S08001 \(2008\)](#).
- [127] R. Bruce *et al.*, *LHC Run 2: Results and challenges*, in [57th ICFA Advanced Beam Dynamics Workshop on High-Intensity and High-Brightness Hadron Beams](#) (2016), MOAM5P50.
- [128] LHCb Collaboration, R. Aaij *et al.*, *LHCb detector performance*, [Int. J. Mod. Phys. **A30**, 1530022 \(2015\)](#), [arXiv:1412.6352](#).
- [129] LHCb RICH, A. Papanestis and C. D’Ambrosio, *Performance of the LHCb RICH detectors during the LHC Run II*, [Nucl. Instrum. Meth. A **876**](#), edited by P. Krizan *et al.*, 221–224 (2017), [arXiv:1703.08152](#).
- [130] H1, C. Adloff *et al.*, *Measurement of open beauty production at HERA*, [Phys. Lett. B **467**, 156–164 \(1999\)](#), [arXiv:hep-ex/9909029](#), [Erratum: [Phys.Lett.B 518, 331–332 \(2001\)](#)].
- [131] D. Hutchcroft, *VELO pattern recognition*, CERN-LHCB-2007-013, LHCB-2007-013 (2007), <https://cds.cern.ch/record/1023540> (unpublished).
- [132] R. Aaij *et al.*, *Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, [JINST **14**, P04013 \(2019\)](#), [arXiv:1812.10790](#).
- [133] LHCb, A. Ryzhikov *et al.*, *Robust Neural Particle Identification Models*, [J. Phys. Conf. Ser. **2438**, 012119 \(2023\)](#), [arXiv:2212.07274](#).
- [134] *LHCb Run 2 Dataflow*, <https://lhcb.github.io/starterkit/>, Accessed: 2024-05-28.
- [135] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, LHCb-PUB-2014-039, CERN-LHCB-PUB-2014-039 (CERN, Geneva, 2014), CDS: [1701134](#).
- [136] E. Rodrigues *et al.*, *The Scikit HEP Project – overview and prospects*, [EPJ Web Conf. **245**](#), edited by C. Doglioni *et al.*, 06028 (2020), [arXiv:2007.03577](#).
- [137] R. Brun *et al.*, *Root-project/root: v6.18/02*, version v6-18-02, June 2020.
- [138] LHCb, I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, in [2010 IEEE Nuclear Science Symposium, Medical Imaging Conference, and 17th Room Temperature Semiconductor Detectors Workshop](#) (2010), pp. 1155–1161.
- [139] H.-U. Bengtsson and T. Sjostrand, *PYTHIA: The Lund Monte Carlo for Hadronic Processes*, [Conf. Proc. C **860623**](#), edited by R. Donaldson and J. N. Marx, 311 (1986).

- [140] T. Sjöstrand *et al.*, *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191**, 159–177 (2015), [arXiv:1410.3012](#).
- [141] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth. A* **462**, edited by S. Erhan, P. Schlein, and Y. Rozen, 152–155 (2001).
- [142] GEANT4 Collaboration, S. Agostinelli *et al.*, *GEANT4—a simulation toolkit*, *Nucl. Instrum. Meth. A* **506**, 250–303 (2003).
- [143] LHCb, M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, *J. Phys. Conf. Ser.* **331**, edited by S. C. Lin, 032023 (2011).
- [144] G. Corti *et al.*, *Software for the LHCb experiment*, *IEEE Trans. Nucl. Sci.* **53**, 1323–1328 (2006).
- [145] S. Descotes-Genon and M. Novoa-Brunet, *Angular analysis of the rare decay $\Lambda_b \rightarrow \Lambda(1520)(\rightarrow NK)\ell^+\ell^-$* , *JHEP* **06**, 136 (2019), [arXiv:1903.00448](#), [Erratum: *JHEP* **06**, 102 (2020)].
- [146] Y.-S. Li, S.-P. Jin, J. Gao, and X. Liu, *Transition form factors and angular distributions of the $\Lambda b \rightarrow \Lambda(1520)(\rightarrow NK^-)\ell^+\ell^-$ decay supported by baryon spectroscopy*, *Phys. Rev. D* **107**, 093003 (2023), [arXiv:2210.04640](#).
- [147] S. Meinel and G. Rendon, *$\Lambda C \rightarrow \Lambda^*(1520)$ form factors from lattice qcd and improved analysis of the $\Lambda b \rightarrow \Lambda^*(1520)$ and $\Lambda b \rightarrow \Lambda c^*(2595,2625)$ form factors*, *Phys. Rev. D* **105**, 054511 (2022), [arXiv:2107.13140](#).
- [148] Y. Amhis, M. Bordone, and M. Reboud, *Dispersive analysis of $\Lambda_b^0 \rightarrow \Lambda(1520)$ local form factors*, *Journal of High Energy Physics* **2023**, 10 (2023).
- [149] Y. Amhis *et al.*, *Prospects for new physics searches with $\Lambda_b^0 \rightarrow \Lambda(1520)\ell^+\ell^-$ decays*, *The European Physical Journal Plus* **136**, 614 (2021).
- [150] A. Beck, T. Blake, and M. Kreps, *Angular distribution of $\Lambda_b^0 \rightarrow pK^-\ell^+\ell^-$ decays comprising Λ resonances with spin $\leq 5/2$* , *Journal of High Energy Physics* **2023**, 189 (2023).
- [151] F. Volle, *Angular analysis of $\Lambda b \rightarrow \Lambda(1520)\mu^+\mu^-$ decays with the LHCb detector*, Theses (Université Paris-Saclay, Sept. 2023).
- [152] BaBar, J. P. Lees *et al.*, *Searches for Rare or Forbidden Semileptonic Charm Decays*, *Phys. Rev. D* **84**, 072006 (2011), [arXiv:1107.4465](#).
- [153] LHCb collaboration, R. Aaij *et al.*, *Study of the productions of Λ_b^0 and \bar{B}^0 hadrons in pp collisions and first measurement of the $\Lambda_b^0 \rightarrow J/\psi pK^-$ branching fraction*, *Chin. Phys.* **C40**, 011001 (2016), [arXiv:1509.00292](#).
- [154] LHCb collaboration, R. Aaij *et al.*, *Observation of a narrow pentaquark state, $P_c(4312)^+$, and of two-peak structure of the $P_c(4450)^+$* , *Phys. Rev. Lett.* **122**, 222001 (2019), [arXiv:1904.03947](#).
- [155] A. L. Read, *Modified frequentist analysis of search results (the CL_s method)*, in *Workshop on Confidence Limits: CERN, Geneva, Switzerland 17 - 18 Jan 2000*. (CERN, 2000), pp. 81–101.
- [156] V. V. Gligorov, *Reconstruction of the Channel $B_d^0 \rightarrow D^+\pi^-$ and Background Classification at LHCb (revised)*, CERN-LHCb-2007-044 (CERN, Geneva, 2007), <https://cds.cern.ch/record/1035682> (unpublished).

- [157] LHCb collaboration, R. Aaij *et al.*, *Observation of an excited B_c^+ state*, *Phys. Rev. Lett.* **122**, 232001 (2019), [arXiv:1904.00081](#).
- [158] A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, *JINST* **8**, P02022 (2013), [arXiv:1211.1346](#).
- [159] *LHCb TCKs in data*, (2020) <https://twiki.cern.ch/twiki/bin/view/Main/LHCbImperialTCKs>, Accessed: 11-03-2024.
- [160] V. V. Gligorov, *A single track HLT1 trigger*,
- [161] BESIII, M. Ablikim *et al.*, *Search for the Lepton Flavor Violation Process $J/\psi \rightarrow e\mu$ at BESIII*, *Phys. Rev. D* **87**, 112007 (2013), [arXiv:1304.3205](#).
- [162] M. N. Achasov *et al.*, *Search for Lepton Flavor Violation Process $e^+e^- \rightarrow e\mu$ in the Energy Region $\sqrt{s} = 984 - 1060\text{MeV}$ and $\phi \rightarrow e\mu$ Decay*, *Phys. Rev. D* **81**, 057102 (2010), [arXiv:0911.1232](#).
- [163] M.-H. Schune, F. Polci, and M. Borsato, *HOP an additional tool for decays involving electrons*, LHCb-INT-2015-037, CERN-LHCb-INT-2015-037 (CERN, Geneva, 2015), <https://cds.cern.ch/record/2102345>.
- [164] LHCb collaboration, R. Aaij *et al.*, *Study of the D^0p amplitude in $\Lambda_b^0 \rightarrow D^0p\pi^-$ decays*, *JHEP* **05**, 030 (2017), [arXiv:1701.07873](#).
- [165] M. Pivk and F. R. Le Diberder, *SPlot: A Statistical tool to unfold data distributions*, *Nucl. Instrum. Meth. A* **555**, 356–369 (2005), [arXiv:physics/0402083](#).
- [166] Particle Data Group, M. Tanabashi *et al.*, *Review of Particle Physics*, *Phys. Rev. D* **98**, 030001 (2018).
- [167] L. Anderlini *et al.*, *The PIDCalib package*, LHCb-PUB-2016-021, CERN-LHCb-PUB-2016-021 (CERN, Geneva, 2016).
- [168] O. Lupton, L. Anderlini, B. Sciascia, and V. Gligorov, *Calibration samples for particle identification at LHCb in Run 2*, LHCb-PUB-2016-005, CERN-LHCb-PUB-2016-005 (CERN, Geneva, 2016).
- [169] LHCb, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, *JINST* **10**, P02007 (2015), [arXiv:1408.1251](#).
- [170] A. Rogozhnikov, *Reweighting algorithms*, (2020) https://arogozhnikov.github.io/hep%5C_ml/reweight.html, Accessed: 24-03-2024.
- [171] T. Chen and C. Guestrin, *Xgboost: a scalable tree boosting system*, in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, KDD '16 (2016), pp. 785–794.
- [172] A. Hocker *et al.*, *TMVA - Toolkit for Multivariate Data Analysis with ROOT: Users guide. TMVA- Toolkit for Multivariate Data Analysis*, physics/0703039, CERN-OPEN-2007-007 (CERN, Geneva, 2007), TMVA-v4 Users Guide: 135 pages, 19 figures, numerous code examples and references.
- [173] F. Chollet *et al.*, *Keras*, <https://keras.io>, 2015.
- [174] C. Kenworthy, *Search for the Rare Lepton Flavour Violating Decay $\Lambda_b^0 \rightarrow pK\mu e$ at LHCb*, University of Birmingham MSci Masters Project, 2022.

- [175] R. Hull, *Optimisation of simulated signal to real background ratio in the search for the Lepton Flavour Violating decay $\Lambda_b^0 \rightarrow pK\mu e$* , University of Birmingham MSci Masters Project, 2022.
- [176] S. Ioffe and C. Szegedy, *Batch normalization: accelerating deep network training by reducing internal covariate shift*, in Proceedings of the 32nd international conference on international conference on machine learning - volume 37, ICML'15 (2015), pp. 448–456.
- [177] C. Cortes, M. Mohri, and A. Rostamizadeh, *L2 regularization for learning kernels*, in [UAI 2009, proceedings of the twenty-fifth conference on uncertainty in artificial intelligence, montreal, qc, canada, june 18-21, 2009](#), edited by J. A. Bilmes and A. Y. Ng (2009), pp. 109–116.
- [178] F. Mosteller and J. W. Tukey, *Data analysis, including statistics*, in Handbook of social psychology, vol. 2, edited by G. Lindzey and E. Aronson (1968).
- [179] K. Pearson, *Note on Regression and Inheritance in the Case of Two Parents*, Proceedings of the Royal Society of London Series I **58**, 240–242 (1895).
- [180] A. Geron, *Hands-on machine learning with scikit-learn, keras, and tensorflow: concepts, tools, and techniques to build intelligent systems* (O'Reilly Media, Incorporated, 2019).
- [181] F. Pedregosa *et al.*, *Scikit-learn: machine learning in Python*, Journal of Machine Learning Research **12**, 2825–2830 (2011).
- [182] A. Kolmogorov, *Sulla determinazione empirica di una legge didistribuzione*, [Giorn Dell'inst Ital Degli Att](#) **4**, 89–91 (1933).
- [183] N. Smirnov, *Table for estimating the goodness of fit of empirical distributions*, [The Annals of Mathematical Statistics](#) **19**, 279–281 (1948).
- [184] G. Punzi, *Sensitivity of searches for new signals and its optimization*, in Statistical problems in particle physics, astrophysics, and cosmology, edited by L. Lyons, R. Mount, and R. Reitmeyer (2003), p. 79, [arXiv:physics/0308063](#).
- [185] P. Koppenburg, *Statistical biases in measurements with multiple candidates*, (2017), [arXiv:1703.01128](#).
- [186] LHCb collaboration, R. Aaij *et al.*, *Measurement of b-hadron fractions in 13 TeV pp collisions*, [Phys. Rev. D](#) **100**, 031102(R) (2019), [arXiv:1902.06794](#).
- [187] LHCb collaboration, R. Aaij *et al.*, *Precise measurement of the f_s/f_d ratio of fragmentation fractions and of B_s^0 decay branching fractions*, [Phys. Rev. D](#) **104**, 032005 (2021), [arXiv:2103.06810](#).
- [188] LHCb collaboration, R. Aaij *et al.*, *Determination of the quark coupling strength $|V_{ub}|$ using baryonic decays*, [Nature Physics](#) **11**, 743 (2015), [arXiv:1504.01568](#).
- [189] E. Parzen, *On Estimation of a Probability Density Function and Mode*, [The Annals of Mathematical Statistics](#) **33**, 1065–1076 (1962).
- [190] M. Rosenblatt, *Remarks on Some Nonparametric Estimates of a Density Function*, [The Annals of Mathematical Statistics](#) **27**, 832–837 (1956).

- [191] ARGUS, H. Albrecht *et al.*, *Search for Hadronic $b \rightarrow u$ Decays*, *Phys. Lett. B* **241**, 278–282 (1990).
- [192] N. L. Johnson, *Systems of Frequency Curves Generated by Methods of Translation*, *Biometrika* **36**, 149–176 (1949).
- [193] W. D. Hulsbergen, *Decay chain fitting with a kalman filter*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **552**, 566–575 (2005).
- [194] T. B. Berrett and R. J. Samworth, *USP: an independence test that improves on Pearson’s chi-squared and the G-test*, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **477**, 10.1098/rspa.2021.0549 (2021).
- [195] J. M. Blatt and V. F. Weisskopf, *Theoretical nuclear physics* (Springer Science & Business Media, 2012).
- [196] J. Eschle, *General model fitting with zfit and hepstats*, July 2024, doi:[10.5281/zenodo.12634108](https://doi.org/10.5281/zenodo.12634108).
- [197] LHCb Collaboration, *LHCb VELO Upgrade Technical Design Report*, CERN-LHCC-2013-021 (CERN, Geneva, 2013).
- [198] LHCb Collaboration, *LHCb Upgrade GPU High Level Trigger Technical Design Report*, CERN-LHCC-2020-006 (CERN, Geneva, 2020).
- [199] K. Hennessy, *LHCb VELO upgrade*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **845**, 97–100 (2017), Proceedings of the Vienna Conference on Instrumentation 2016.
- [200] M. Brice, *LHCb’s VELO metrology*, (2022) <https://cds.cern.ch/record/2802195>, General Photo.
- [201] I. Zurbarano Fernandez *et al.*, *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*, CERN-2020-010 10/2020, edited by I. B ejar Alonso *et al.*, 10.23731/CYRM-2020-0010 (2020).
- [202] I. Efthymiopoulos *et al.*, *LHCb Upgrades and operation at $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ luminosity – A first study*, CERN-ACC-NOTE-2018-0038 (2018).
- [203] *Simultaneous determination of the CKM angle γ and parameters related to mixing and CP violation in the charm sector*, 2022.
- [204] LHCb Collaboration, R. Aaij *et al.*, *Physics case for an LHCb Upgrade II - Opportunities in flavour physics, and beyond, in the HL-LHC era*, LHCb-PUB-2018-009, CERN-LHCC-2018-027, LHCC-G-171 (CERN, Geneva, 2016), [arXiv:1808.08865](https://arxiv.org/abs/1808.08865), ISBN 978-92-9083-494-6.
- [205] L. Rossi, P. Fischer, T. Rohe, and N. Wermes, *Pixel detectors*, en, 2006th ed., Particle Acceleration and Detection (Springer, Berlin, Germany, Jan. 2006).
- [206] A. Lampis, *Innovative silicon pixel sensors for a 4D VERTeX LOcator detector for the LHCb high luminosity upgrade*, PhD thesis (2023-01-08, Cagliari U., Jan. 2023).
- [207] G. Kramberger *et al.*, *Radiation hardness of thin low gain avalanche detectors*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **891**, 68–77 (2018).

- [208] E. Bossini and N. Minafra, *Diamond detectors for timing measurements in high energy physics*, [Frontiers in Physics](#) **8**, 10.3389/fphy.2020.00248 (2020).
- [209] G. Kramberger *et al.*, *Charge collection studies on custom silicon detectors irradiated up to 1.6×10^{17} n_{eq}/cm^2* , [JINST](#) **8**, P08004 (2013).
- [210] A. Affolder, P. Allport, and G. Casse, *Collected charge of planar silicon detectors after pion and proton irradiations up to 2.2×10^{16} n_{eq}/cm^2* , [Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment](#) **623**, 177–179 (2010), 1st International Conference on Technology and Instrumentation in Particle Physics.
- [211] J. Lange *et al.*, *Radiation hardness of small-pitch 3D pixel sensors up to a fluence of 3×10^{16} n_{eq}/cm^2* , [JINST](#) **13**, P09009 (2018).
- [212] M. Manna *et al.*, *First characterisation of 3d pixel detectors irradiated at extreme fluences*, [Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment](#) **979**, 164458 (2020).
- [213] L. Anderlini *et al.*, *Intrinsic time resolution of 3D-trench silicon pixels for charged particle detection*, [JINST](#) **15**, P09029 (2020).
- [214] J. Mulvey *et al.*, *Preliminary test results of LGADs from Teledyne e2v for the LHC’s High-Luminosity upgrade*, [JINST](#) **17**, C10001 (2022).
- [215] N. Moffat *et al.*, *Low Gain Avalanche Detectors (LGAD) for particle physics and synchrotron applications*, [JINST](#) **13**, C03014 (2018).
- [216] R. Moriya *et al.*, *26.5 ps Time Resolution Using 50 μm Low Gain Avalanche Detectors Fabricated by Micron Semiconductor Ltd*, (2023), [arXiv:2310.06183](#).
- [217] F. Siviero *et al.*, *Optimization of the gain layer design of ultra-fast silicon detectors*, [Nucl. Instrum. Meth. A](#) **1033**, 166739 (2022), [arXiv:2112.00561](#).
- [218] R. Padilla *et al.*, *Effect of deep gain layer and carbon infusion on lgad radiation hardness*, [JINST](#) **15**, P10003 (2020).
- [219] R. Ballabriga, M. Campbell, and X. Llopart, *An introduction to the Medipix family ASICs*, [Radiat. Meas.](#) **136**, 106271 (2020).
- [220] R. Mulargia *et al.*, *Temperature dependence of the response of ultra fast silicon detectors*, [JINST](#) **11**, edited by C. Gemme and L. Rossi, C12013 (2016).
- [221] N. Cartiglia *et al.*, *Design optimization of ultra-fast silicon detectors*, [Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment](#) **796**, 141–148 (2015), Proceedings of the 10th International Conference on Radiation Effects on Semiconductor Materials Detectors and Devices.
- [222] P. Dervan *et al.*, *Upgrade to the Birmingham Irradiation Facility*, [Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment](#) **796**, 80–84 (2015), Proceedings of the 10th International Conference on Radiation Effects on Semiconductor Materials Detectors and Devices.

- [223] P. Dervan *et al.*, *The Birmingham Irradiation Facility*, [Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment](#) **730**, 101–104 (2013), Proceedings of the 9th International Conference on Radiation Effects on Semiconductor Materials Detectors and Devices.
- [224] V. K. Khanna, *Extreme-temperature and harsh-environment electronics (second edition)*, 2053-2563 (IOP Publishing, 2023).
- [225] G. Laštovička-Medin *et al.*, *New insight into gain suppression and single event Burnout effects in LGAD*, [JINST](#) **18**, C02059 (2023).
- [226] S. Mägdefessel, R. Mori, N. Sorgenfrei, and U. Parzefall, *Understanding the Frequency Dependence of Capacitance Measurements of Irradiated Silicon Detectors*, (2023), [arXiv:2301.09371](#).
- [227] M. Ferrero *et al.*, *Radiation resistant lgad design*, [Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment](#) **919**, 16–26 (2019).
- [228] LHCb VELO, A. Fernández Prieto, *The LHCb VELO Upgrade II: design and development of the readout electronics*, [JINST](#) **19**, C05011 (2024).
- [229] X. Llopart *et al.*, *Timepix4, a large area pixel detector readout chip which can be tiled on 4 sides providing sub-200 ps timestamp binning*, [JINST](#) **17**, C01044 (2022).
- [230] G. Borghello *et al.*, *Ionizing radiation damage in 65 nm cmos technology: influence of geometry, bias and temperature at ultra-high doses*, [Microelectronics Reliability](#) **116**, 114016 (2021).
- [231] G. Borghello, *Ionizing radiation effects in nanoscale CMOS technologies exposed to ultra-high doses* (Udine U., 2018), Presented 08 Mar 2019.
- [232] M. Mazurek, G. Corti, and D. Muller, *New simulation software technologies at the LHCb Experiment at CERN*, LHCb-PROC-2021-011 (2021), [arXiv:2112.04789](#).
- [233] A. Sailer *et al.*, *DD4Hep Based Event Reconstruction*, [Journal of Physics: Conference Series](#) **898**, 042017 (2017).
- [234] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetry in $B_s^0 \rightarrow D_s^\mp K^\pm$ decays*, [JHEP](#) **11**, 060 (2014), [arXiv:1407.6127](#).
- [235] J. Mnich, *Agenda 4th meeting of the LHCb PhaseIIb*, CERN-RRB-2023-068 (CERN, Geneva, 2023).
- [236] LHCb, D. Fazzini, *Flavour Tagging in the LHCb experiment*, [PoS LHCP2018](#), 230 (2018).
- [237] E. Buchanan, *Spatial Resolution Studies for the LHCb VELO Upgrade* (Bristol U., 2018).
- [238] M. Aleksa *et al.*, *Strategic R&D Programme on Technologies for Future Experiments*, [10.17181/CERN.5PQI.KDL2](#) (2018).
- [239] CMS, M. Manoni, *Characterization of ultra radiation hard pixel detectors for the high luminosity phase of the CMS experiment at the LHC*, [Nuovo Cim. C](#) **47**, 125 (2024).

- [240] M. Williams, *WP1.1 Silicon Hybrid Detector*, <https://indico.cern.ch/event/1395929/contributions/5867717>, Feb. 2023.
- [241] J. Troska *et al.*, *The VTRx+, an optical link module for data transmission at HL-LHC*, *PoS TWEPP-17*, 048 (2017).
- [242] C. Scarcella *et al.*, *System development of silicon photonics links for CERN experiments and accelerators*, *JINST* **18**, C03002 (2023).
- [243] K. Bos *et al.*, *LHC computing Grid: Technical Design Report. Version 1.06 (20 Jun 2005)*, Technical design report. LCG (CERN, Geneva, 2005).
- [244] I. Bird *et al.*, *Update of the Computing Models of the WLCG and the LHC Experiments*, CERN-LHCC-2014-014, LCG-TDR-002 (2014).
- [245] A. Mazurov, B. Couturier, D. Popov, and N. Farley, *Microservices for systematic profiling and monitoring of the refactoring process at the lhcb experiment*, *Journal of Physics: Conference Series* **898**, 072037 (2017).
- [246] LHCb, D. Popov, *Testing and verification of the LHCb Simulation*, *EPJ Web Conf.* **214**, 02043 (2019).
- [247] A. Soroko *et al.*, *The ganga user interface for physics analysis and distributed resources*, LHCb-2004-076, CERN-LHCb-2004-076 (CERN, Geneva, 2004).
- [248] R. Matev, *Fast distributed compilation and testing of large C++ projects*, *EPJ Web Conf.* **245**, edited by C. Doglioni *et al.*, 05001 (2020).

Appendices

Appendix A

Additional Detail for Pre-Selection

A.1 Trigger Selection Options Efficiency Summary

A plethora of different options for trigger selection were considered for the analysis, with the final choice of L0Muon_TOS only, as described in Section 4.1.4. The major component that was not included is L0E, the L0Electron_TOS selection, as shown in Table A.1.

Table A.1: Performance of possible further trigger lines in addition to the nominal trigger selections for $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$ data signal simulation and sideband data. Trigger lines refer to the TOS version unless specified. Uncertainties are statistical.

Trigger Selection for $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$	MC ϵ	Sideband Data ϵ
L0 Nominal: L0Muon	0.568 ± 0.002	0.664 ± 0.006
L0I & !Nominal	0.146 ± 0.002	0.258 ± 0.006
L0E & !Nominal	0.123 ± 0.001	0.066 ± 0.003
L0ETIS & !Nominal	0 ± 0	0 ± 0
L0mTIS & !Nominal	0.0379 ± 0.0009	0.091 ± 0.004
L0K & !Nominal	0.0156 ± 0.0006	0.02 ± 0.002
L0p & !Nominal	0.047 ± 0.001	0.033 ± 0.002
HLT1 Nominal: TrackMVA TrackMuon	0.802 ± 0.002	0.858 ± 0.005
TwoTrack & !Nominal	0.0132 ± 0.0005	0.102 ± 0.004
TrackMuonMVA & !Nominal	0 ± 0	0.005 ± 0.001
HLT2 Nominal: Topo234 TopoMu234	0.795 ± 0.002	0.74 ± 0.006
TopoMuE234 & !Nominal	0.0014 ± 0.0002	0.02 ± 0.002
TopoE234 & !Nominal	0.0024 ± 0.0002	0.022 ± 0.002

Appendix B

Further Material for the Multivariate Analysis

The final nine variables used for the analysis MVA (described in Section 4.3) were chosen by their physical motivations as either kinematic or reconstructed variables that showed high separation power between the corrected MC signal (target sample) and upper sideband data (rejection sample). All nine variables are shown in Figure B.1, with their full descriptions in Section 4.3.2.3. It was also imperative that these variables showed negligible mass bias, hence the lower sideband and signal region was checked to have a similar distribution to the upper sideband, shown in the Figure.

A second check was performed with the control mode data and simulation to ensure the nine variables were well modelled by comparing the corrected simulation to sWeighted data. All nine MVA variables are included in Figure B.2, where the good agreement shown supports the conclusion that the variables used are well modelled.

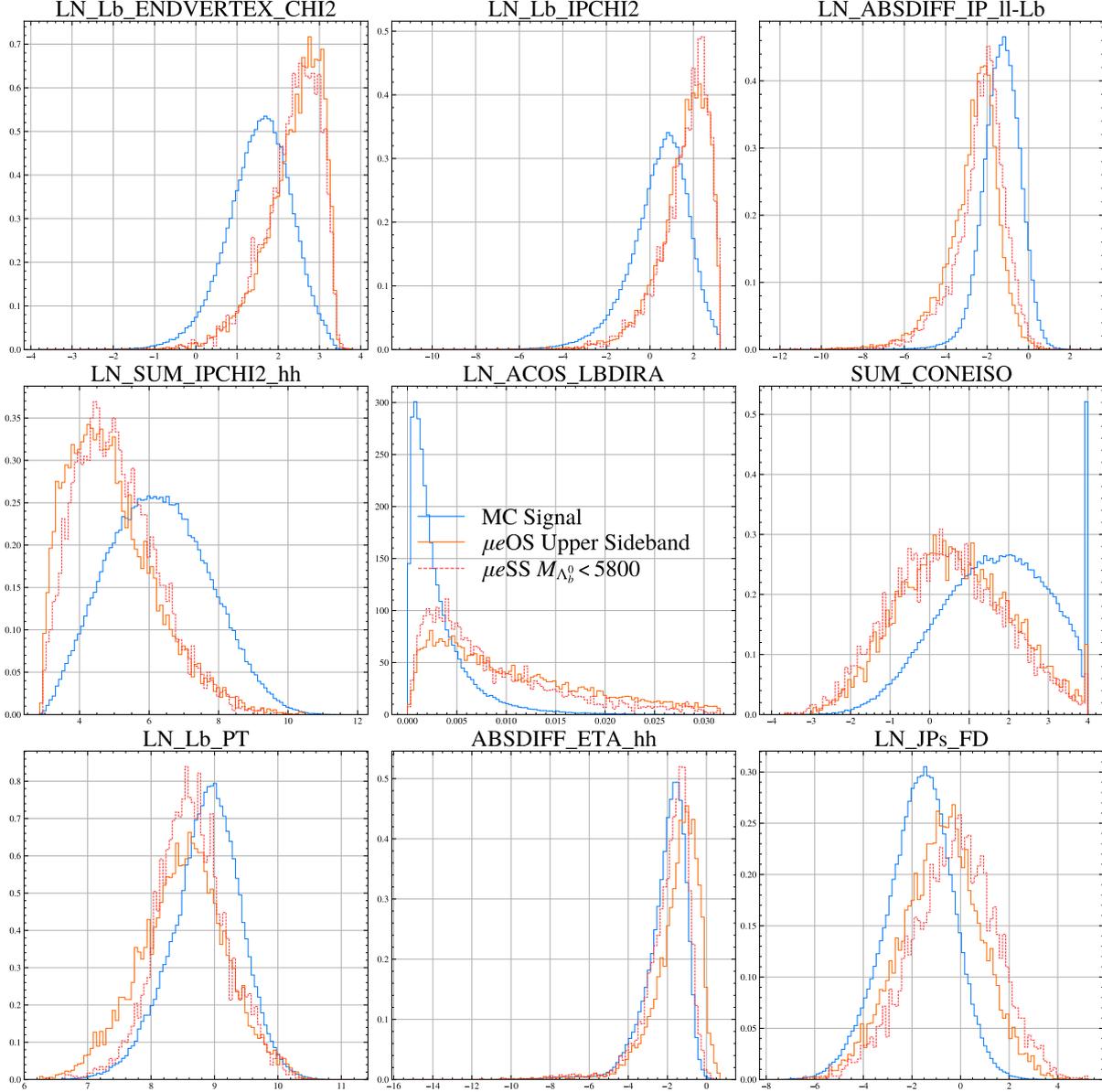


Figure B.1: Distributions of all discriminating variables used to train the MVA, with the signal samples shown. Along with the corrected signal mode MC and the μeOS upper sideband (used as the signal and background training sets), the μeSS distribution for $M(\Lambda_b^0) < 5800$ MeV is included to ensure the variable will remove background across the analysis $M(\Lambda_b^0)$ window. Full descriptions of the variables are given in Section 4.3.2.3.

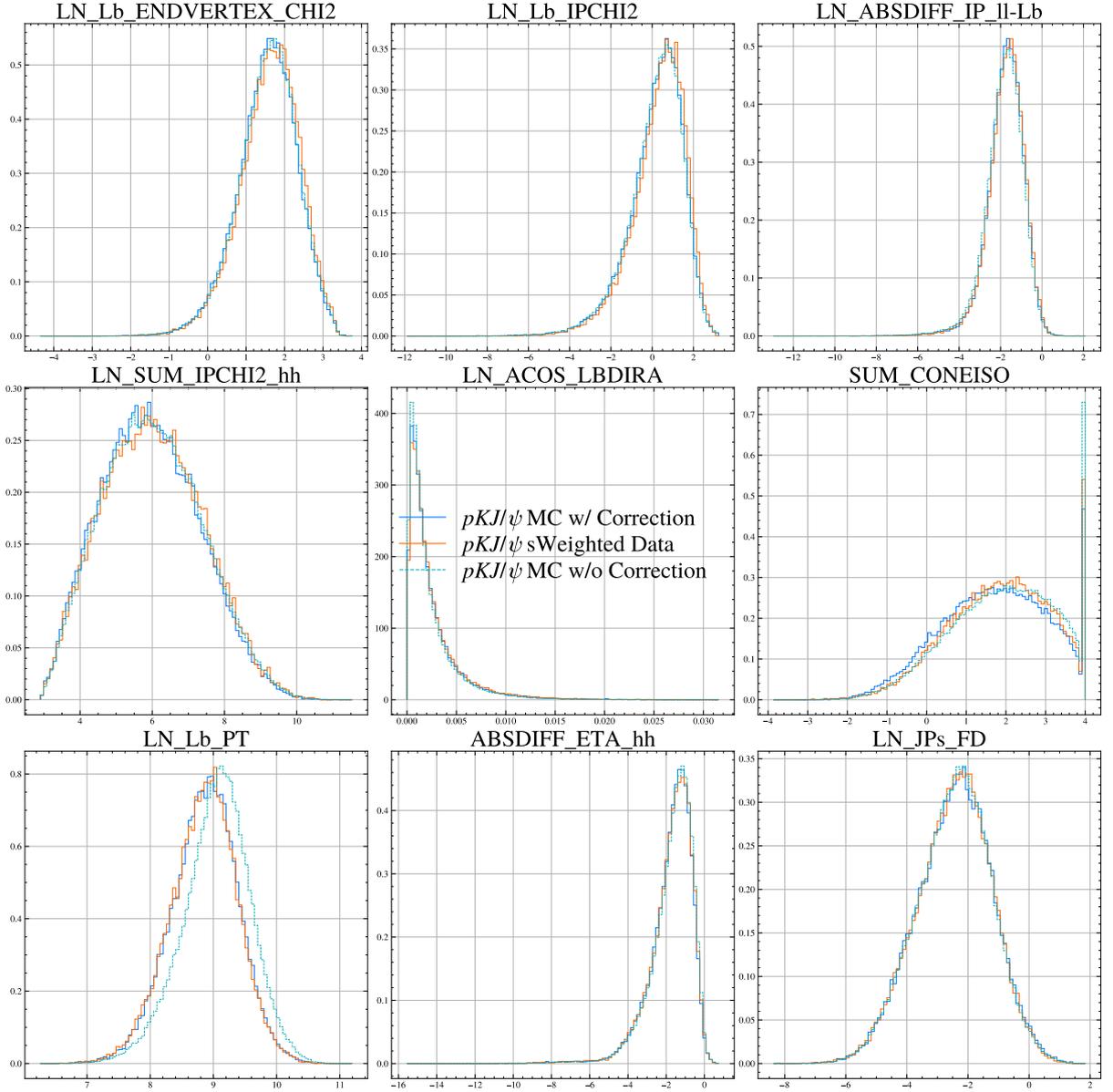


Figure B.2: Distributions of the discriminating variables used to train the MVA, with the corrected $\Lambda_b^0 \rightarrow pK^- J/\psi (\rightarrow \mu^+ \mu^-)$ simulation and sWeighted data shown to have good agreement across all variables, implying that in the signal mode these variables are also well modelled (where the same sWeight comparison is not possible). The simulation without weights is also included to show the agreement that the weights provide. Full descriptions of the variables are given in Section 4.3.2.3.

Appendix C

Complete Tables for the Background Study

The complete background study tables for the signal mode are included, to support the refined set in Table 5.1. Table C.1 shows the potential backgrounds from hadronic mis-ID ($h \rightarrow \ell$), Table C.2 from leptonic mis-ID ($\ell \rightarrow \ell'$), Table C.3 where a single-semileptonic ($\ell\nu_\ell$) decay is present and Table C.4 where a double-semileptonic (double $\ell\nu_\ell, \ell'\nu_{\ell'}$) decay is present. The full set of background for the control mode are also included in Table C.4.

Table C.1: Background yield estimates for the background samples where hadronic mis-ID ($h \rightarrow \ell$) is required in the reconstruction mode $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$. The labels PS, MVA and PID refers to the different selection stages that this specific BG estimate was calculate. If an estimate is not present the sample had no statistics remaining before that selection could be applied.

Background Sample Estimated Yield ($\mu^\pm e^\mp$)	Stage	Analysis Categories			
		Run 1 1 γ	Run 1 0 γ	Run 2 1 γ	Run 2 0 γ
$\Lambda_b^0 \rightarrow pK^- K^+ K^-$	PS:			0.0 ± 0.2	1.8 ± 0.5
	MVA:				1.7 ± 0.5
	PID:			0.0 ± 0.3	0.0 ± 0.2
$\Lambda_b^0 \rightarrow pK^- \pi^+ \pi^-$	PS:	0.006 ± 0.002	0.15 ± 0.02	0.009 ± 0.003	0.7 ± 0.1
	MVA:	0.007 ± 0.002	0.14 ± 0.02	0.007 ± 0.003	0.28 ± 0.05
	PID:	0.0003 ± 0.0005	0.006 ± 0.002	0.0007 ± 0.0007	0.012 ± 0.003
$\Lambda_b^0 \rightarrow D^0(\rightarrow K^- \pi^+) pK^-$	PS:			0.0 ± 0.06	0.0 ± 0.08
$\Lambda_b^0 \rightarrow pK^- K^+ \pi^-$	PS:			0.0 ± 0.04	0.16 ± 0.09
$\Lambda_b^0 \rightarrow \Lambda_c^+ (\rightarrow \Lambda(1520)\pi^+) \pi^-$	PS:	0.0 ± 0.002	0.0 ± 0.003		0.0 ± 0.4

Table C.2: Background yield estimate tables for the background samples where leptonic mis-ID ($\ell \rightarrow \ell'$) is required in the reconstruction mode $A_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$. The labels PS, MVA and PID refers to the different selection stages that this specific BG estimate was calculate. If an estimate is not present the sample had no statistics remaining before that selection could be applied.

Background Sample	Stage	Analysis Categories			
		Run 1 1γ	Run 1 0γ	Run 2 1γ	Run 2 0γ
$A_b^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) pK^-$	PS:	1.2 ± 0.2	7.0 ± 1.0	2.9 ± 0.4	6.5 ± 0.8
	MVA:	0.8 ± 0.1	5.8 ± 0.9	1.3 ± 0.2	2.9 ± 0.3
$B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) K^+$	PID:	0.001 ± 0.003	0.16 ± 0.04	0.013 ± 0.004	0.1 ± 0.02
	PS:	0.0 ± 0.02	0.23 ± 0.05	0.01 ± 0.01	0.25 ± 0.06
$A_b^0 \rightarrow \psi(2S)(\rightarrow \mu^+ \mu^-) pK^-$	MVA:	0.0 ± 0.0006	0.03 ± 0.01	0.01 ± 0.01	0.01 ± 0.02
	PID:	0.0 ± 0.009	0.0 ± 0.004	0.0 ± 0.006	0.0 ± 0.003
$A_b^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) pK^-$	PS:	0.032 ± 0.005	0.047 ± 0.007	0.09 ± 0.01	0.07 ± 0.008
	MVA:	0.021 ± 0.004	0.044 ± 0.007	0.049 ± 0.006	0.038 ± 0.005
$A_b^0 \rightarrow J/\psi(\rightarrow e^+ e^-) pK^-$	PID:	0.0 ± 0.0001	0.0015 ± 0.0005	0.0027 ± 0.0006	0.0024 ± 0.0004
	PS:	0.008 ± 0.005	0.003 ± 0.002	0.007 ± 0.003	0.017 ± 0.003
$A_b^0 \rightarrow \Lambda(1520)\mu^+ \mu^-$	MVA:	0.008 ± 0.006	0.003 ± 0.002	0.007 ± 0.003	0.015 ± 0.003
	PID:	0.0 ± 0.002	0.0004 ± 0.0009	0.0 ± 0.001	0.0005 ± 0.0006
$A_b^0 \rightarrow \Lambda(1520)\mu^+ \mu^-$	PS:	0.011 ± 0.002	0.3 ± 0.05	0.026 ± 0.003	0.8 ± 0.1
	MVA:	0.009 ± 0.001	0.26 ± 0.04	0.021 ± 0.003	0.74 ± 0.09
$A_b^0 \rightarrow pK^- e^+ e^-$	PID:	$1.2e - 05 \pm 8e - 06$	0.0004 ± 0.0001	$0.0001 \pm 2e - 05$	0.0035 ± 0.0004
	PS:	$6e - 05 \pm 5e - 05$	$0.00031 \pm 9e - 05$	$0.00012 \pm 5e - 05$	$0.00044 \pm 8e - 05$
$A_b^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) \pi^+ \pi^- pK^-$	MVA:	$6e - 05 \pm 5e - 05$	$0.00023 \pm 8e - 05$	$0.00012 \pm 5e - 05$	$0.00044 \pm 8e - 05$
	PID:	$1e - 05 \pm 4e - 05$	$2e - 05 \pm 2e - 05$	$1e - 05 \pm 2e - 05$	$4e - 05 \pm 2e - 05$
$A_b^0 \rightarrow \psi(2S)(\rightarrow e^+ e^-) \Lambda(1520)$	PS:	0.0 ± 9.0	0.0 ± 9.0	0.0 ± 9.0	0.0 ± 10.0
	MVA:	$0.0 \pm 2e - 05$	$0.0 \pm 2e - 05$	$0.0 \pm 2e - 05$	0.0 ± 0.3
$A_b^0 \rightarrow \psi(2S)(\rightarrow e^+ e^-) \Lambda(1520)$	PID:	$0.0 \pm 5e - 05$	$0.0 \pm 5e - 05$	$0.0 \pm 5e - 05$	$0.0 \pm 9e - 05$
	PS:	$0.0 \pm 3e - 05$	$0.0 \pm 8e - 06$	$0.0 \pm 3e - 05$	$0.0 \pm 1e - 05$
$A_b^0 \rightarrow \psi(2S)(\rightarrow e^+ e^-) \Lambda(1520)$	MVA:	$0.0 \pm 3e - 05$	$0.0 \pm 8e - 06$	$0.0 \pm 4e - 06$	$0.0 \pm 1e - 05$
	PID:	$0.0 \pm 2e - 05$	$0.0 \pm 2e - 05$	$0.0 \pm 2e - 05$	$0.0 \pm 9e - 06$

Table C.3: Background yield estimate tables for the background samples where a single-semileptonic ($\ell\nu_\ell$) decay is present in the reconstruction mode $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$. The labels PS, MVA and PID refers to the different selection stages that this specific BG estimate was calculate. If an estimate is not present the sample had no statistics remaining before that selection could be applied.

Background Sample Estimated Yield ($\mu^\pm e^\mp$)	Stage	Analysis Categories			
		Run 1 1 γ	Run 1 0 γ	Run 2 1 γ	Run 2 0 γ
$\Lambda_b^0 \rightarrow D^0 (\rightarrow K^- \mu^+ \nu_\mu) p \pi^-$	PS:	0.07 ± 0.03	2.6 ± 0.3	0.0 ± 0.7	13.0 ± 3.0
	MVA:	0.001 ± 0.006	1.4 ± 0.2		7.0 ± 2.0
	PID:	0.0 ± 0.002	0.07 ± 0.02	0.0 ± 0.8	0.3 ± 0.9
$\Lambda_b^0 \rightarrow D^0 (\rightarrow K^- e^+ \nu_e) \Lambda(1520)$	PS:	0.16 ± 0.04	0.06 ± 0.02	0.46 ± 0.09	0.47 ± 0.09
	MVA:	0.09 ± 0.03	0.05 ± 0.02	0.25 ± 0.06	0.32 ± 0.07
	PID:	0.03 ± 0.01	0.01 ± 0.01	0.06 ± 0.03	0.1 ± 0.04
$\Lambda_b^0 \rightarrow D^0 (\rightarrow K^- \mu^+ \nu_\mu) \Lambda(1520)$	PS:	0.19 ± 0.04	2.2 ± 0.4	0.37 ± 0.07	5.2 ± 0.7
	MVA:	0.11 ± 0.03	1.0 ± 0.2	0.16 ± 0.04	2.2 ± 0.3
	PID:	0.004 ± 0.007	0.04 ± 0.02	0.0 ± 0.01	0.04 ± 0.02
$\Lambda_b^0 \rightarrow A_c^+ (\rightarrow \Lambda(1520)\pi^-) \mu^- \bar{\nu}_\mu$	PS:	2.4 ± 0.7	0.3 ± 0.2	30.0 ± 20.0	0.0 ± 7.0
	MVA:	1.1 ± 0.5	0.03 ± 0.06	0.0 ± 4.0	
	PID:	0.02 ± 0.04	0.0 ± 0.03		
$\Lambda_b^0 \rightarrow D^0 (\rightarrow K^- e^+ \nu_e) p \pi^-$	PS:	0.0 ± 0.007	0.01 ± 0.02	1.0 ± 1.0	0.0 ± 0.6
	MVA:		0.0 ± 0.001		
	PID:	0.0 ± 0.004	0.0 ± 0.001		
$\Lambda_b^0 \rightarrow A_c^+ (\rightarrow \Lambda(1520)\pi^+) e^- \bar{\nu}_e$	PS:	0.0 ± 0.3	0.0 ± 0.06	0.0 ± 5.0	0.0 ± 10.0
$\Lambda_b^0 \rightarrow A_c^+ (\rightarrow \Lambda(1520)e^+ \nu_e) \pi^-$	PS:	0.0 ± 0.001	0.0 ± 0.0009	0.0 ± 0.09	0.0 ± 0.1
$\Lambda_b^0 \rightarrow A_c^+ (\rightarrow \Lambda(1520)\mu^+ \nu_\mu) \pi^-$	PS:	0.0 ± 0.001	0.0 ± 0.001	0.0 ± 0.09	0.0 ± 0.1

Table C.4: Background yield estimate tables for the background samples where a double-semileptonic (double $\ell\nu_\ell, \ell'\nu_{\ell'}$) decay is present in the reconstruction mode $\Lambda_b^0 \rightarrow \Lambda(1520)\mu^\pm e^\mp$. The labels PS, MVA and PID refers to the different selection stages that this specific BG estimate was calculate. If an estimate is not present the sample had no statistics remaining before that selection could be applied. The \dagger indicates modes that use an estimated branching fraction to determine the yields, discussed in Section 5.1.1.

Background Sample Estimated Yield ($\mu^\pm e^\mp$)	Stage	Analysis Categories			
		Run 1 1 γ	Run 1 0 γ	Run 2 1 γ	Run 2 0 γ
$A_b^0 \rightarrow D^0(\rightarrow K^- e^+ \nu_e) p \mu^- \bar{\nu}_\mu$ †	PS: MVA: PID:	29.0 ± 6.0 12.0 ± 3.0 8.0 ± 2.0	25.0 ± 5.0 9.0 ± 2.0 4.5 ± 0.9	61.0 ± 9.0 25.0 ± 4.0 14.0 ± 3.0	63.0 ± 9.0 30.0 ± 5.0 17.0 ± 2.0
$A_b^0 \rightarrow D^0(\rightarrow K^- \mu^+ \nu_\mu) p e^- \bar{\nu}_e$ †	PS: MVA: PID:	18.0 ± 4.0 7.0 ± 2.0 5.0 ± 1.0	13.0 ± 2.0 8.0 ± 2.0 3.5 ± 0.7	45.0 ± 7.0 19.0 ± 4.0 12.0 ± 2.0	38.0 ± 5.0 22.0 ± 3.0 12.0 ± 2.0
$A_b^0 \rightarrow \Lambda_c(2860)^+ \mu^- \bar{\nu}_\mu$ † $\Lambda_c(2860)^+ \rightarrow D^0(\rightarrow K^- e^+ \nu_e) p$	PS: MVA: PID:	2.5 ± 0.6 1.2 ± 0.4 0.6 ± 0.2	1.6 ± 0.4 0.8 ± 0.2 0.6 ± 0.1	5.0 ± 1.0 1.8 ± 0.5 1.1 ± 0.3	6.0 ± 1.0 3.0 ± 0.6 2.1 ± 0.4
$A_b^0 \rightarrow \Lambda_c(2860)^+ e^- \bar{\nu}_e$ † $\Lambda_c(2860)^+ \rightarrow D^0(\rightarrow K^- \mu^+ \nu_\mu) p$	PS: MVA: PID:	1.1 ± 0.3 0.5 ± 0.2 0.2 ± 0.1	1.2 ± 0.3 0.1 ± 0.1 0.08 ± 0.04	1.5 ± 0.5 0.4 ± 0.3 0.2 ± 0.1	5.1 ± 0.9 1.5 ± 0.5 1.1 ± 0.2
$A_b^0 \rightarrow \Lambda_c(2940)^+ \mu^- \bar{\nu}_\mu$ † $\Lambda_c(2940)^+ \rightarrow D^0(\rightarrow K^- e^+ \nu_e) p$	PS: MVA: PID:	0.17 ± 0.07 0.08 ± 0.04 0.03 ± 0.02	0.1 ± 0.04 0.004 ± 0.003 0.0 ± 0.002	1.2 ± 0.3 0.27 ± 0.09 0.18 ± 0.06	0.23 ± 0.07 0.003 ± 0.003 0.0 ± 0.003
$A_b^0 \rightarrow \Lambda_c^+(\rightarrow \Lambda(1520) e^+ \nu_e) \mu^- \bar{\nu}_\mu$	PS: MVA: PID:	0.01 ± 0.02 0.002 ± 0.008 0.0 ± 0.009	0.18 ± 0.07 0.05 ± 0.02 0.001 ± 0.003	0.5 ± 0.2 0.0 ± 0.04 0.0 ± 0.04	0.3 ± 0.2 0.11 ± 0.09 0.0 ± 0.05
$B^0 \rightarrow D^{*-}(2460)(\rightarrow \bar{D}^0(\rightarrow K^+ e^- \bar{\nu}_e) \pi^-) \mu^+ \nu_\mu$	PS: MVA: PID:	1.6 ± 0.3 0.51 ± 0.09 0.037 ± 0.007	1.0 ± 0.2 0.31 ± 0.06 0.019 ± 0.004	1.3 ± 0.2 0.39 ± 0.06 0.018 ± 0.003	0.67 ± 0.09 0.2 ± 0.03 0.006 ± 0.001
$A_b^0 \rightarrow \Lambda_c(2940)^+ e^- \bar{\nu}_e$ † $\Lambda_c(2940)^+ \rightarrow D^0(\rightarrow K^- \mu^+ \nu_\mu) p$	PS: MVA: PID:	0.0 ± 0.008	0.0 ± 0.004	0.3 ± 0.1 0.09 ± 0.05 0.06 ± 0.03	0.0 ± 0.008 0.0 ± 0.005 0.0 ± 0.002
$B^0 \rightarrow D^{*-}(2460)(\rightarrow \bar{D}^0(\rightarrow K^+ \mu^- \bar{\nu}_\mu) \pi^-) e^+ \nu_e$	PS: MVA: PID:	1.0 ± 0.2 0.3 ± 0.05 0.022 ± 0.004	0.8 ± 0.1 0.23 ± 0.04 0.016 ± 0.003	0.8 ± 0.1 0.23 ± 0.03 0.011 ± 0.002	0.63 ± 0.09 0.2 ± 0.03 0.009 ± 0.001
$B^0 \rightarrow D^-(\rightarrow K^{*0} e^- \bar{\nu}_e) \mu^+ \nu_\mu$	PS: MVA: PID:	1.3 ± 0.2 0.14 ± 0.04 0.01 ± 0.01	1.6 ± 0.2 0.22 ± 0.04 0.008 ± 0.005	1.2 ± 0.1 0.23 ± 0.04 0.02 ± 0.01	1.4 ± 0.1 0.19 ± 0.03 0.001 ± 0.002
$B_s^0 \rightarrow D_s^{*-}(2573)(\rightarrow \bar{D}^0(\rightarrow K^+ e^- \bar{\nu}_e) K^-) \mu^+ \nu_\mu$	PS: MVA: PID:	0.6 ± 0.1 0.2 ± 0.04 0.024 ± 0.006	0.22 ± 0.05 0.07 ± 0.02 0.007 ± 0.002	1.0 ± 0.2 0.39 ± 0.07 0.013 ± 0.006	0.26 ± 0.05 0.04 ± 0.01 0.0007 ± 0.0008
$B^0 \rightarrow D^-(\rightarrow K^{*0} \mu^- \bar{\nu}_\mu) e^+ \nu_e$	PS: MVA: PID:	0.8 ± 0.1 0.13 ± 0.03 0.008 ± 0.007	0.8 ± 0.1 0.13 ± 0.03 0.01 ± 0.004	0.8 ± 0.1 0.12 ± 0.03 0.001 ± 0.004	0.44 ± 0.06 0.04 ± 0.01 0.001 ± 0.001
$B_s^0 \rightarrow D_s^{*-}(2573)(\rightarrow \bar{D}^0(\rightarrow K^+ \mu^- \bar{\nu}_\mu) K^-) e^+ \nu_e$	PS: MVA: PID:	0.3 ± 0.07 0.09 ± 0.02 0.011 ± 0.003	0.19 ± 0.04 0.08 ± 0.02 0.009 ± 0.002	0.7 ± 0.1 0.1 ± 0.03 0.003 ± 0.003	0.31 ± 0.06 0.07 ± 0.02 0.0015 ± 0.0008
$B^0 \rightarrow D^{*-}(\rightarrow D^-(\rightarrow K^{*0} e^- \bar{\nu}_e) \pi^0) \mu^+ \nu_\mu$	PS: MVA: PID:	0.66 ± 0.08 0.11 ± 0.02 0.008 ± 0.004	0.41 ± 0.05 0.08 ± 0.01 0.004 ± 0.002	0.45 ± 0.05 0.037 ± 0.009 0.001 ± 0.002	0.42 ± 0.04 0.11 ± 0.02 0.001 ± 0.0008
$B^0 \rightarrow D^{*-}(\rightarrow D^-(\rightarrow K^{*0} \mu^- \bar{\nu}_\mu) \pi^0) e^+ \nu_e$	PS: MVA: PID:	0.36 ± 0.05 0.07 ± 0.02 0.005 ± 0.003	0.28 ± 0.04 0.04 ± 0.01 0.003 ± 0.001	0.27 ± 0.04 0.022 ± 0.007 0.0 ± 0.001	0.17 ± 0.02 0.041 ± 0.008 0.0005 ± 0.0004
$A_b^0 \rightarrow \Lambda_c^+(\rightarrow \Lambda(1520) \mu^+ \nu_\mu) e^- \bar{\nu}_e$	PS:	0.0 ± 0.01	0.0 ± 0.006	0.0 ± 0.05	0.0 ± 0.06

Table C.4: Background yield estimates for the key background samples in the control mode $\Lambda_b^0 \rightarrow pK^- J/\psi(\rightarrow \mu^+ \mu^-)$, split by LHC Run. The labels PS, MVA and PID refers to the different selection stages that this specific BG estimate was calculate. If an estimate is not present the sample had no statistics remaining before that selection could be applied.

Background Datasets Estimated Yield ($\mu^\pm e^\mp$)	Stage	Analysis Categories	
		Run 1	Run 2
$B^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) K^{*0}$	PS:	13000.0 \pm 1000.0	14000.0 \pm 1000.0
	MVA:	11000.0 \pm 1000.0	12000.0 \pm 1000.0
	PID:	710.0 \pm 80.0	1060.0 \pm 90.0
$B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) K^+ K^-$	PS:	6700.0 \pm 800.0	15000.0 \pm 1000.0
	MVA:	5500.0 \pm 700.0	13000.0 \pm 1000.0
	PID:	800.0 \pm 100.0	760.0 \pm 70.0
$\Lambda_b^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) pK^-$ (<i>pK Swap</i>)	PS:	2600.0 \pm 400.0	4700.0 \pm 500.0
	MVA:	2200.0 \pm 300.0	4000.0 \pm 400.0
	PID:	300.0 \pm 40.0	43.0 \pm 5.0
$B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) K^{*0}$	PS:		240.0 \pm 20.0
	MVA:		200.0 \pm 20.0
	PID:		18.0 \pm 2.0
$B_s^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) \pi^+ \pi^-$	PS:	47.0 \pm 6.0	41.0 \pm 4.0
	MVA:	37.0 \pm 4.0	35.0 \pm 3.0
	PID:	1.4 \pm 0.5	3.0 \pm 1.0
$B^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) \pi^+ \pi^-$	PS:	47.0 \pm 5.0	32.0 \pm 3.0
	MVA:	35.0 \pm 4.0	28.0 \pm 2.0
	PID:	1.3 \pm 0.4	2.1 \pm 0.6
$B^+ \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) K^+$	PS:	23.0 \pm 3.0	9.0 \pm 1.0
	MVA:	1.9 \pm 0.5	0.6 \pm 0.3
	PID:	0.0 \pm 1.0	0.0 \pm 2.0
$\Lambda_b^0 \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) pK^- \pi^0$	PS:	1.1 \pm 0.2	
	MVA:	0.6 \pm 0.1	
	PID:	0.31 \pm 0.06	

Appendix D

Additional for N_{Signal} and $N_{Control}$ Extraction

For the final control mode and signal fits a plethora of pseudoexperiments were performed to validate robustness and stability of the methods used, described in Sections 5.3 and 5.4 where only a selection of the full collection of results were displayed. Figure D.1 shows the stability tests for the control mode fit without exclusive backgrounds with all parameters showing good stability and correctly calculated uncertainties. Figure D.2 shows the remaining parameters from the pseudoexperiments presented in Figure 5.9. Figure D.3 shows the equivalent metric to Figure 5.10 except for Run 1, displaying that there is a significant difference between the control mode fitting techniques, with the value in b used as the systematic uncertainty from extracting $N_{Control}$. Following this, Figure D.4 shows the M_{pK} fit for Run 1, with a similar fit result.

Figure D.6 displays the remaining results from the cascade fit method described in Section 5.4.2, with the Run 2 result better describing the data, as was seen for the other categories studied in Figure 5.14. For the final background fit, the alternative option did not include an exclusive $D^0p\ell$ component. The fits for the Run 2 categories in this mode are included in Figure D.5. The stability tests for the Run 1 blinded background fits with the nominal method are shown in Figure D.7, showing similar stability to the other categories, acceptable considering the correlation between the yields. The blinded crossfit pseudoexperiments for the remaining analysis categories are shown in Figure D.8, where experiments are generated for both models where the $D^0p\ell$ is included or not, and fitting is performed with the opposite hypothesis to evaluate bias. The same conclusion is drawn for all categories that overall including the $D^0p\ell$ prevents overestimation of the background contribution in the signal region.

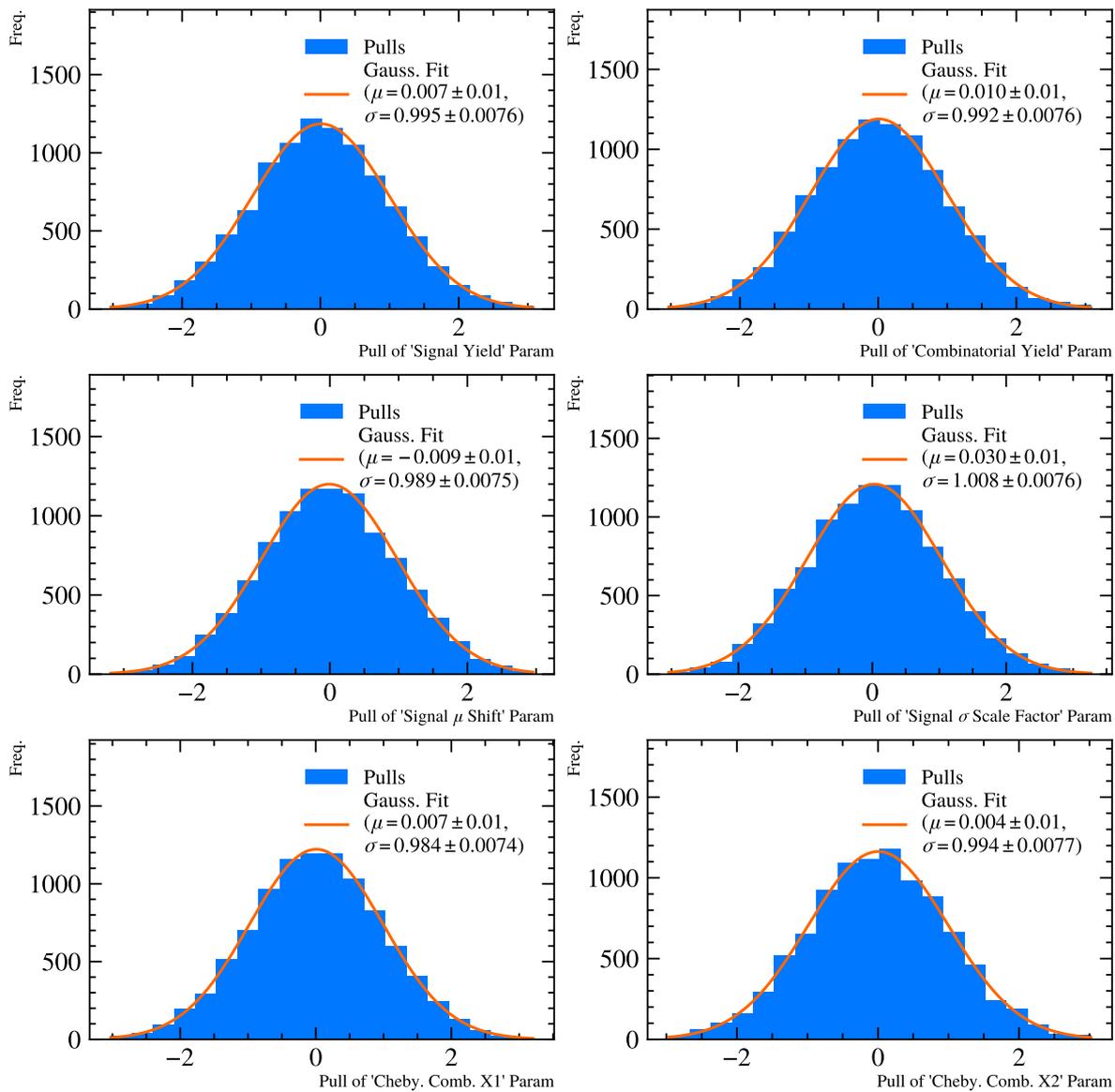


Figure D.1: Pull distributions for all floating parameters for 10000 pseudo-experiments generated for the Run 2 control-mode data fit displayed in Figure 5.5b with no exclusive background components in the fit. Simple gaussian fits with parameter results are shown in orange.

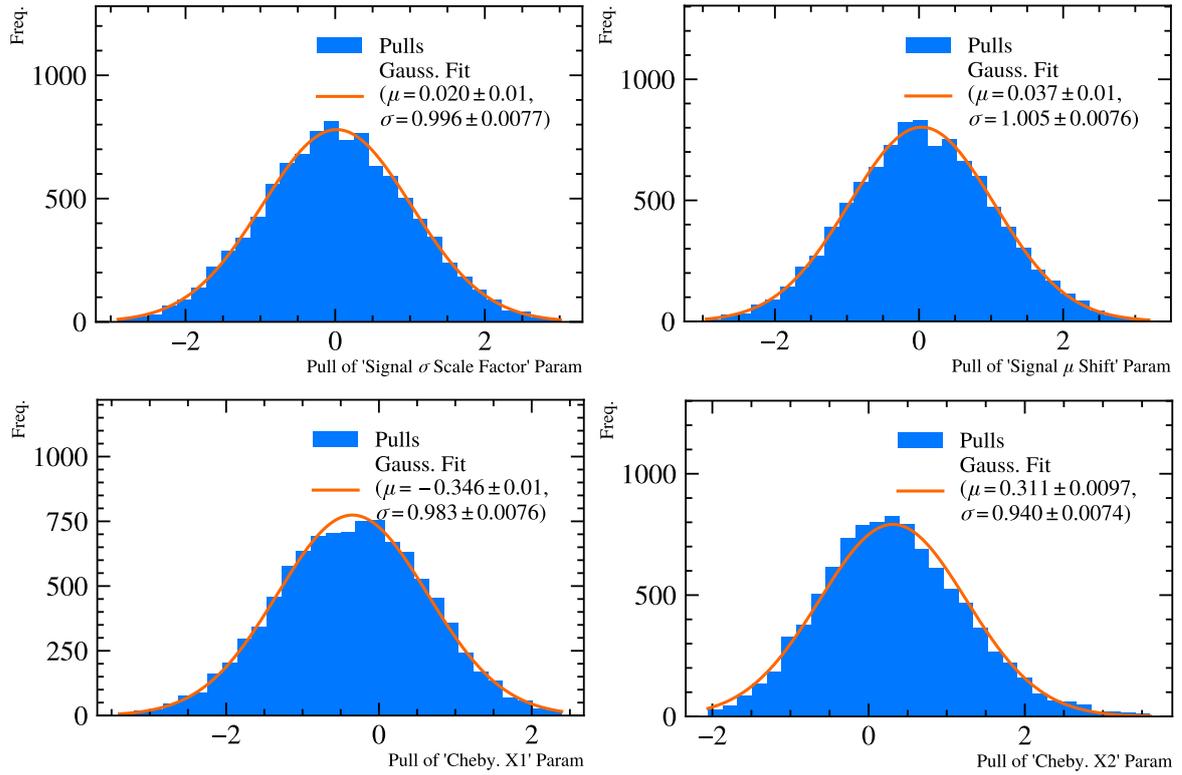


Figure D.2: Pull distributions for remaining floating parameters for 10000 pseudo-experiments generated for the Run 2 control-mode data fit displayed in Figure 5.8a. The pull distributions for the other parameters in Equation 5.3.3 are shown in Figure 5.9. The slight bias in the Chebyshev parameters was found to be due to a correlation between the X_1 and X_2 parameters as well as the exclusive background yields.

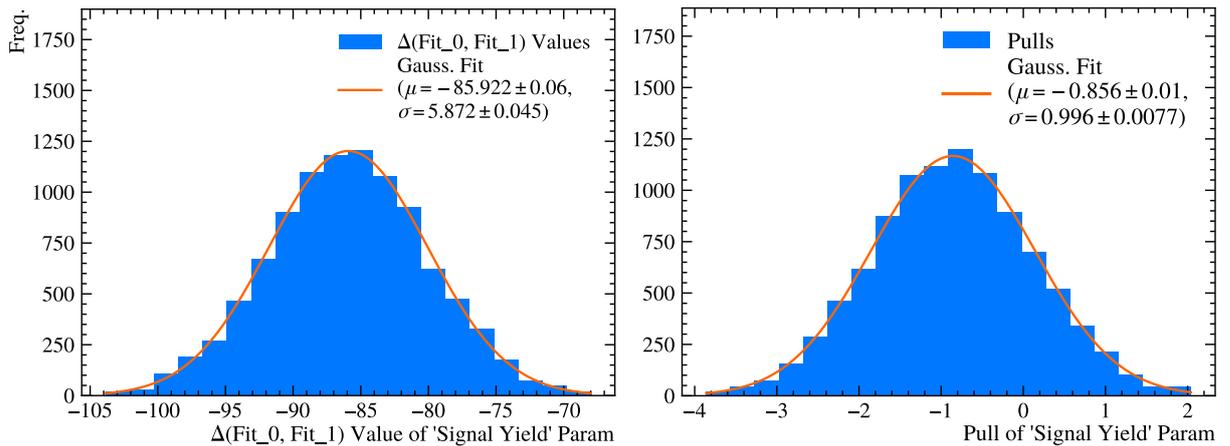


Figure D.3: Distributions from cross-fit pseudo-experiments for the Run 1 control mode, where distributions are generated from the nominal model and fitted with the nominal and alternate. (left) The raw difference in yield between the nominal (with exclusive background) and alternate (without) control mode fitter. (right) The pull distribution of the alternate model fit, using the fitted yield uncertainty only.

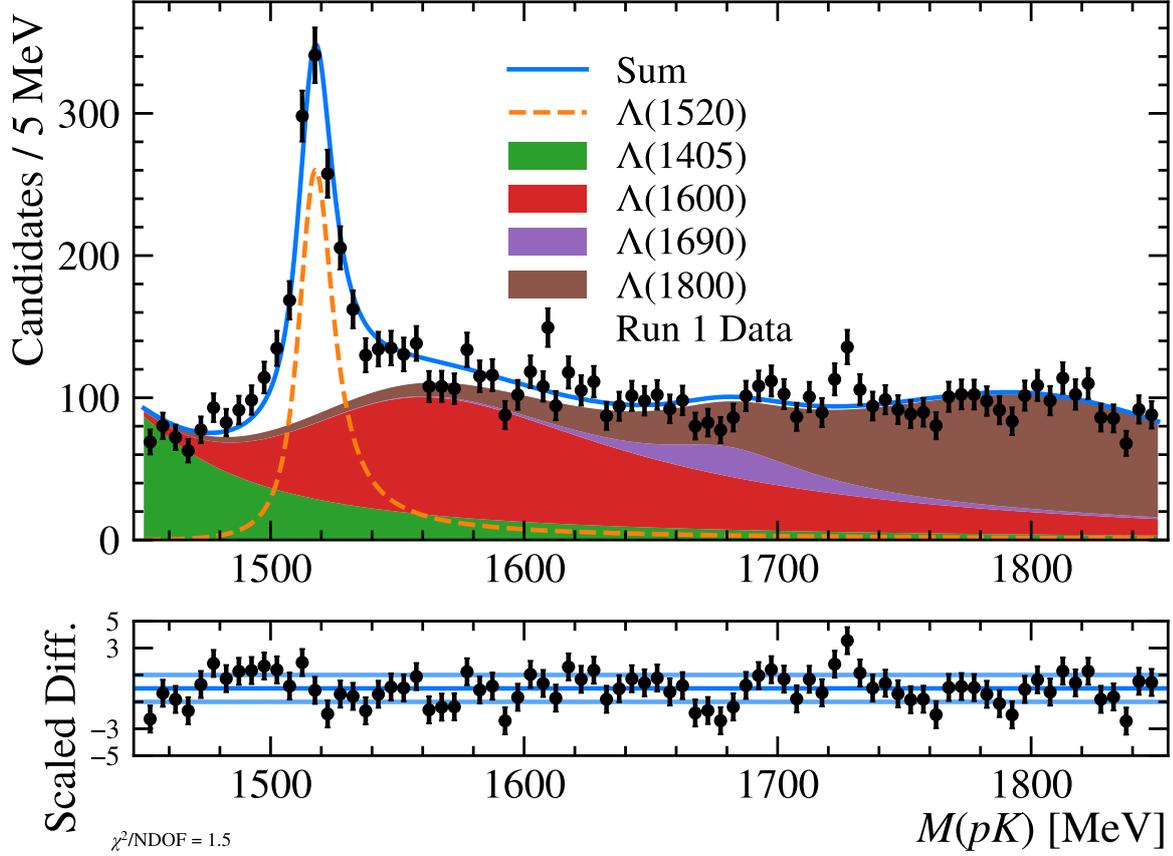


Figure D.4: Invariant mass fit to the M_{pK} spectrum for Run 1 control-mode data using the model defined in Equation 5.3.4. As before, the line shape describes the $\Lambda(1520)$ region proficiently, but the low and high mass regions show trends in the pull distribution.

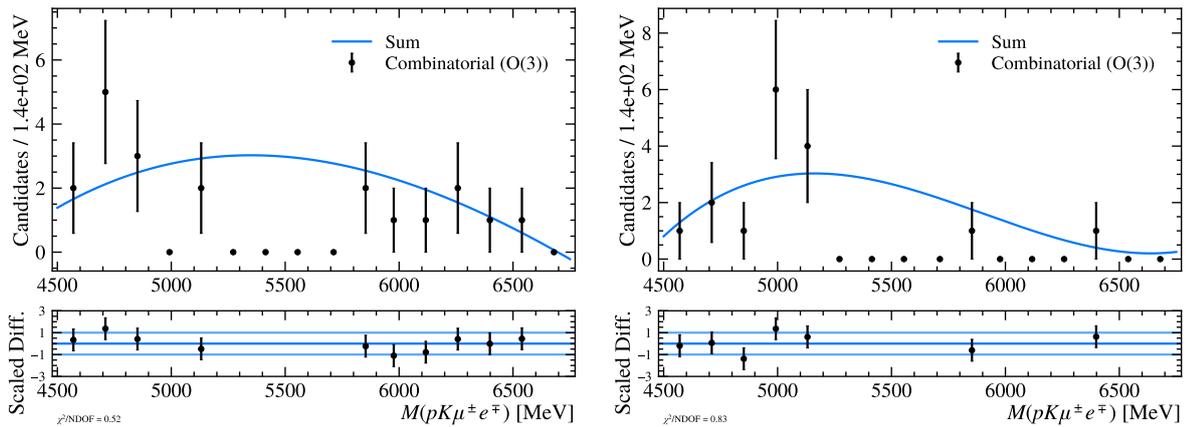


Figure D.5: Blinded fits without the exclusive $\Lambda_b^0 \rightarrow D^0(\rightarrow K^- \ell^+ \nu) p \ell^- \bar{\nu}$ component, to the $M_{\Lambda_b^0}$ sideband distribution for signal mode μe OS data split for Run 2 0γ and, Run 2 1γ . The fit is performed with the description in Equation 5.4.1 with the N_{Signal} fixed to zero, and $N_{D^0 p \ell}$ removed.

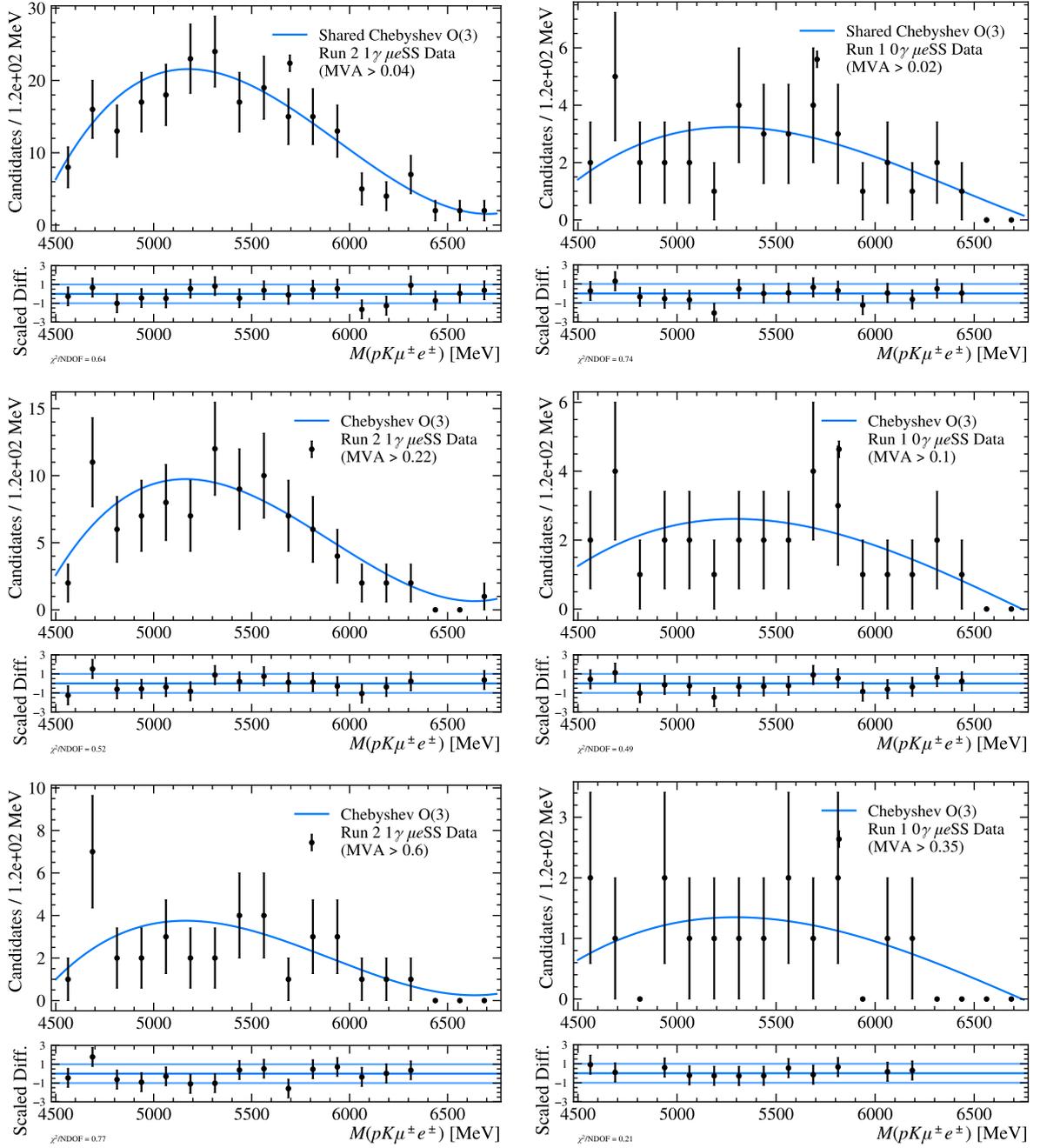


Figure D.6: Invariant mass fits to the $M_{\Lambda_b^0}$ distribution in the μeSS combinatorial proxy dataset for Run 2 0γ (left) and Run 1 1γ (right) with the plots vertically ordered by the loose, medium and nominal working points. The loose fits are simultaneously derived with the corresponding Bremsstrahlung category in the other Run.

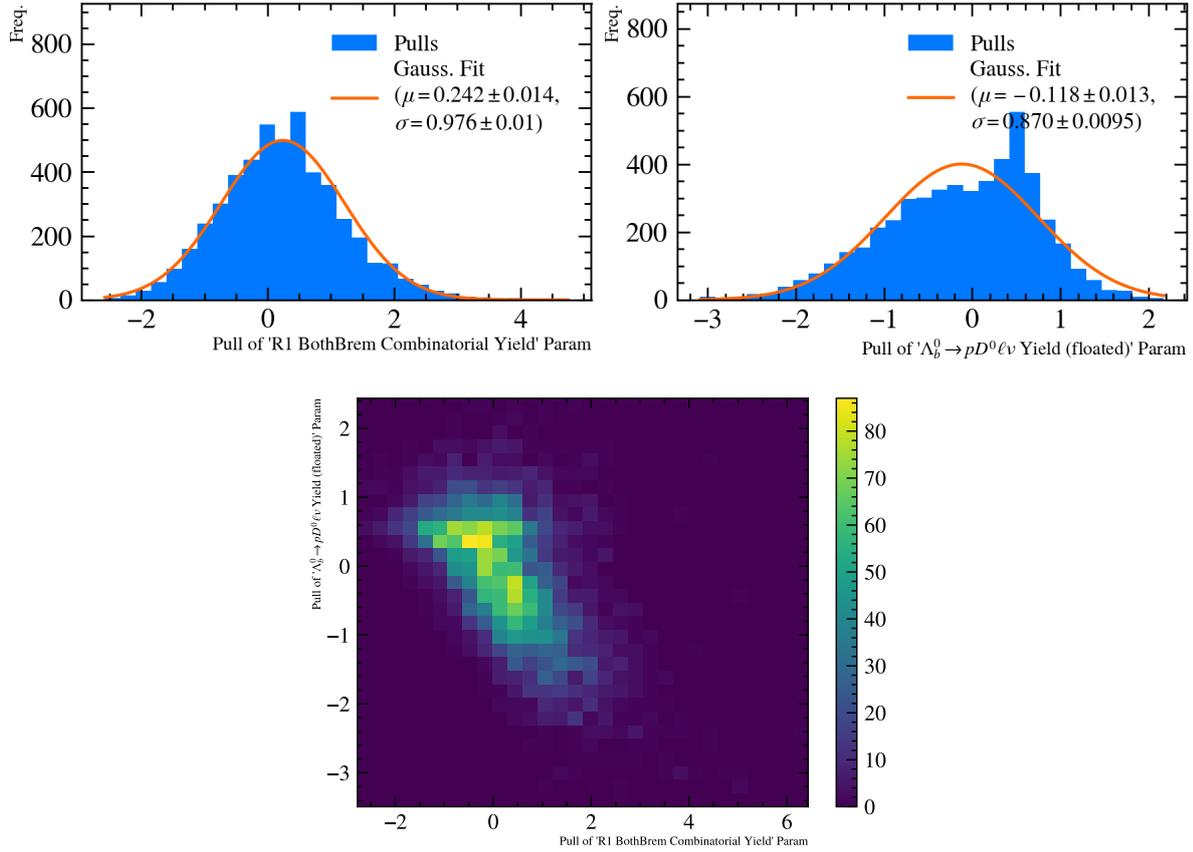


Figure D.7: Pull distributions of the floating parameters in blinded μe OS fits to the Run 1 category, from 5000 generated blind pseudo-experiments for the nominal method with the floating parameters N_{Comb} (left) and $N_{D^0 pl}$ (right) as well as a 2D pull distribution (lower) included to display correlation. The peak on the right hand side plot is from the lower limit of $N_{D^0 pl} \geq 0$. This can be seen in the correlation plot which also indicates a considerable negative correlation between the two yields.

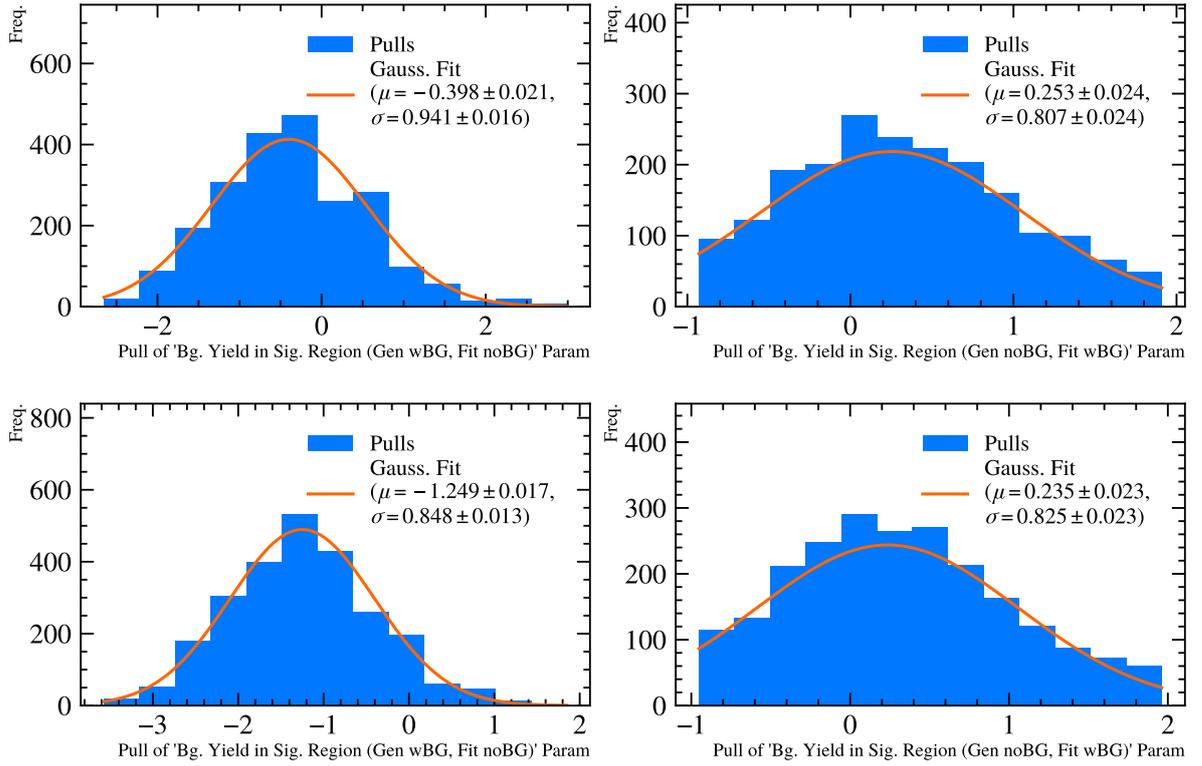


Figure D.8: Cross-compatibility of the two background fit schemes considered for the signal mode, derived by generating 5000 pseudo-experiments for both models where the $D^0 p \ell$ is included or not, and fitting with the opposite hypothesis. (left) The pull distribution of the integrated signal region background yield with the $D^0 p \ell$ generated but not included in the subsequent fit and (right) the inverse, with (upper) Run 1 and (lower) Run 2 1γ included.

Appendix E

Additional Content for Performance Optimisation Tool

An example of two configuration files used to define a build in the Custom Stack builder and to set the config for a WLCG grid job is shown in Figure E.1. Many of these are organised and maintained in files and can be set to run at regular intervals if periodic testing is required.

<pre>"name": "U2_Trk", "binaryTag": "x86_64_v2-centos7-gcc12-opt", "lcgVersion" : "103", "checkouts": { "Detector": "e33c0ce", "Gaudi": "0dc529d", "Gauss": "ecfd1ec", "Gaussino": "488dbbf", "Geant4": "5e093ad", "GaussinoExtLibs": "25e72f9", "LHCb": "9ee03eb", "Run2Support": "430100e" }, "projectsToBuild": ["Gauss"]</pre>	<pre>"jobTitle": "U2_MinBias_nu60_2p5K_TRK_TV-UP3Plane-MT_Gauss-queue", "days": ["Once"], "time": 10, "platform": "x86_64_v2-centos7-gcc12-opt", "stackName": "U2_Trk", "stackTimestamp": "", "nightlyFlavourTrk_", "nightlySlot": "", "PRJobOptions": "30000000.py 2p5KEvents.py UseBeamSpot4D_sim11.py Gauss-TVUPMP-Sim11.py Gauss-Setup-TVcylind-UP3Planes.py Beam70006V-md100-nu60-VerExtAngle.py" "inputFiles": [], "outputFileTypes": [".root"], "splitJob": 25, "handlerToRun": "U2_Trk_GaussHandler", "application": "Gauss", "applicationVer": "", "requiresJob": "", "newDir": "U2_MinBias_nu60_2p5K_TRK_TV-UP3Plane-MT",</pre>
--	---

Figure E.1: Example of configuration files used to define the (left) custom stacks and (right) the LHCbPR jobs that are submitted to the grid using the performance optimisation tool.