

ABSTRACT

Title of Dissertation: MEASUREMENT OF $\mathcal{R}(D^{(*)})$ IN SEMILEPTONIC
 B DECAYS AND UPGRADE OF
THE LHCb UPSTREAM TRACKER

Yipeng Sun
Doctor of Philosophy, 2023

Dissertation Directed by: Professor Manuel Franco Sevilla
Department of Physics

The LHCb experiment at the Large Hadron Collider provides a unique opportunity to study flavor physics with high luminosity. One topic in flavor physics is lepton flavor universality (LFU), a property of the standard model (SM) which requires the three generations of leptons (e, μ, τ) couple to gauge bosons of the electroweak interactions with the same strength. It is an important probe for testing the validity of the SM and possibly providing hints to new physics beyond the SM. This thesis presents a preliminary framework for the measurement of $\mathcal{R}(D^{(*)})$, a proxy to test LFU, defined as the ratio of branching fractions $\mathcal{B}(B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau)/\mathcal{B}(B \rightarrow D^{(*)}\mu^-\bar{\nu}_\mu)$, with LHCb 2016 data. Another topic of the thesis is the upgrade of the LHCb Upstream Tracker (UT) which greatly increases the readout rate of the detector and removes limitations due to hardware trigger, paving the way for future precision measurements with even higher luminosity.

MEASUREMENT OF $\mathcal{R}(D^{(*)})$ IN SEMILEPTONIC B DECAYS AND
UPGRADE OF THE LHCb UPSTREAM TRACKER

by

Yipeng Sun

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2023

Advisory Committee:

Professor Manuel Franco Sevilla, Chair/Advisor
Professor Abohassan Jawahery
Professor Zackaria Chacko
Professor Christopher Palmer
Professor Alice Mignerey, Dean's Representative

© Copyright by
Yipeng Sun
2023

Dedicate to Coca Cola, for a large portion of the thesis was written while
drinking it.

Acknowledgments

I owe my gratitude to all the people who have made this thesis possible.

First, I'd like to thank my advisor, Manuel Franco Sevilla, for his invaluable advices on my research topics as well as providing enough freedom for me to explore interesting side projects.

I would also like to thank Phoebe Hamilton for her deep insights on many problems I ran into.

My colleagues at the Maryland group, notably Alex Fernez and Svende Braun, as well as Zishuo Yang, Emily Jiang, William Parker, and Abohassan Jawahery, have provided useful support both on the analysis and hardware, to which I'm very grateful.

I also want to thank many external collaborators: on the analysis side Gregory Ciezarek, Lucia Grillo, Simone Meloni, Julián García Pardiñas, and Patrick Owen—your implementations on similar problems we ran into provided crucial insights to me. On the hardware side, I'd like to thank Mark Tobin for his extremely useful help on the topics related to the LHCb control system. On the software side, I'd like to thank Dmitry Kalinkin (`veprbl`) for introducing me to `nix`, which greatly improved the reproducibility of this analysis.

Lastly I'd like to thank my family members for being supportive.

It is impossible to remember all, and I apologize to those I've inadvertently left out.

Contents

Dedication	ii
Acknowledgements	iii
Table of Contents	iv
List of Tables	viii
List of Figures	x
List of Abbreviations	xv
Chapter 1: Introduction	1
I Preliminary measurement of $\mathcal{R}(D^{(*)})$ with LHCb 2016 data	4
Chapter 2: Analysis overview	5
Feed down	5
Background contributions	6
Signal and control skims	8
Fit overview	8
Workflow	9
Differences between the run 2 and run 1 analyses	10
Chapter 3: Theory of semileptonic B decays	12
3.1 Construction of the electroweak theory in the SM	12
3.1.1 Outline of the construction	12
3.1.2 Quantum electrodynamics (QED)	15
3.1.3 The electroweak theory	16
3.1.4 Lepton flavor universality in the SM	18
3.2 Semileptonic $B \rightarrow D l^- \bar{\nu}_l$ decays in the SM	19
3.2.1 Form factors in $B \rightarrow D^0 l^- \bar{\nu}_l$ decays	22
3.2.2 Form factors in $B \rightarrow D^{*0} l^- \bar{\nu}_l$ decays	24
3.2.3 Form factors in $B \rightarrow D^{*+} l^- \bar{\nu}_l$ decays	27
Chapter 4: The LHCb experiment	29

4.1	Reconstruction of charged particles	31
4.2	Particle identification	33
4.3	Trigger	37
Chapter 5: Event selection		40
5.1	Data samples	40
5.1.1	Right-sign samples	41
5.1.2	Wrong-sign samples	43
5.1.3	Fake muon samples	46
5.2	MC simulation	47
5.3	Selecting signal (ISO) and control background (1OS, 2OS, DD) skims with the isolation BDT	49
5.4	Auxiliary samples	53
5.5	Additional algorithms used in selection	54
5.5.1	Veto of overlapping candidates	54
5.5.2	Muon PID	55
Chapter 6: Emulation of detector responses on MC simulation		56
6.1	Trigger emulation	56
6.1.1	Emulation of L0Hadron TOS	56
6.1.2	Emulation of L0Global TIS	58
6.1.3	Correlation between L0Hadron TOS and L0Global TIS	61
6.1.4	Emulation of Hlt1TrackMVA and Hlt1TwoTrackMVA	61
6.1.5	Emulation of Hlt2XcMuXForTauB2XcMu	62
6.2	Particle identification emulation	64
Chapter 7: Correction and validation of the MC simulation		65
7.1	Form factor models	65
7.1.1	$B \rightarrow D^0$ decays	66
7.1.2	$B \rightarrow D^*$ decays	67
7.1.3	$B \rightarrow D^{**}$ decays	67
7.2	Initial reweighting	69
7.2.1	Tracking efficiency	69
7.2.2	B kinematics and multiplicity	71
7.3	Post-fit cocktail	73
7.4	Final reweighting	74
7.5	Additional correction on the B vertex resolution	84
Chapter 8: Signal extraction fit		85
8.1	Fit templates	87
8.1.1	Muon misID backgrounds	88
8.1.2	Combinatorial backgrounds	95
8.1.3	Normalization	100
8.1.4	Signal	101
8.1.5	Feed down of $B \rightarrow D^{**}l^-\bar{\nu}_l$ modes	103

8.1.6	Feed down of $B \rightarrow D_H^{**}(\rightarrow D^{(*)}\pi\pi)\mu^-\bar{\nu}_\mu$ modes	108
8.1.7	Feed down of $\bar{B}_s^0 \rightarrow D_s^{**}\mu^-\bar{\nu}_\mu$ modes	109
8.1.8	Double-charm (DDX) backgrounds	109
8.2	Shape variations in the fit	111
8.2.1	Form factor uncertainties	112
8.2.2	Data-driven Dalitz-inspired deformations to DDX model	113
8.2.3	K^* weight up/down in DDX templates	114
8.2.4	Data-driven phenomenological deformation of D_H^{**} model	114
8.2.5	Effect of kaon, pion decay-in-flight in muon misID template	115
8.3	Fit results	118
Chapter 9: Systematic uncertainties		123
9.1	$B \rightarrow (D^0 D^* D^{**})l^-\bar{\nu}_l$ form factors	124
9.2	$B \rightarrow D^{**}\tau^-\bar{\nu}_\tau$ background	125
9.3	$B \rightarrow D^{(*)}D_s^{**}(\rightarrow \tau\bar{\nu}_\tau)X$ background	125
9.4	Control sample shape parameters	126
9.5	DD model dependence	127
9.6	μ misID decay-in-flight correction	128
9.7	Combinatorial background shapes	128
9.8	μ misID unfolding algorithm	129
9.9	Coulomb correction to $\mathcal{R}(D^{*+})$ vs. $\mathcal{R}(D^{*0})$	129
9.10	Vertex resolution correction	130
9.11	Data/MC corrections	131
Chapter 10: Conclusions		132
II Upgrade of the LHCb Upstream Tracker		134
Chapter 11: Overview of the LHCb upgrade		135
11.1	Trigger schemes in run 2 and run 3	137
11.2	Tracking system	141
11.3	Particle identification system	144
Chapter 12: Upgrade of the Upstream Tracker		145
12.1	Overview of the UT detector	145
12.1.1	Stave	147
12.1.2	PEPI	150
12.1.3	LVR	154
12.2	The Data Control Board (DCB)	156
12.2.1	The GBT link	157
12.2.2	E-links	158
12.2.3	Initialization sequence after a power cycle	159
12.2.4	Test of data transmission quality	161
12.2.5	Quality assurance of DCB at CERN	162

12.3	The LHCb online system	163
12.3.1	The DIM control protocol	165
12.3.2	The UT online system	167
12.3.3	MiniDAQ	168
12.3.4	nanoDAQ: A reverse-engineered DIM client for testing UT electronics	170
Appendices		176
Appendix A: Supplemental plots		176
A.1	Change of true q^2 due to ISGW2 $\rightarrow$ BLR with verbatim parameters	176
A.2	D^* combinatorial fit to 1OS, 2OS, and DD skims	177
A.3	Compare L0Global TIS efficiencies in MC with TISTOS method	179
A.4	Initial fit results for post-fit cocktail generation	179
A.5	Pre-control fit result	185
Appendix B: Formal description of selected methods used in this analysis		189
B.1	Shift negative efficiencies	189
B.2	Apply fit yields as a weight	190
Appendix C: Rest frame approximation		192
Appendix D: Trigger categories		193
Appendix E: Unfolding procedure of fake muon control sample		194
Appendix F: Form factor variation		198
Bibliography		201

List of Tables

3.1	$B \rightarrow D^0 l^- \bar{\nu}_l$ CLN FF parameterization.	24
3.2	$B \rightarrow D^0 l^- \bar{\nu}_l$ BGL FF parameterization notation correspondence.	25
3.3	$B \rightarrow D^* l^- \bar{\nu}_l$ CLN FF parameterization.	26
3.4	$B \rightarrow D^* l^- \bar{\nu}_l$ BGL FF parameterization notation correspondence.	27
3.5	FF parameterization for $D^{1/2+}$ doublet (D_0^*, D_1')	28
3.6	FF parameterization for $D^{3/2+}$ doublet (D_1, D_2^*)	28
5.1	Trigger path for this analysis. For terminologies such as TIS and TOS, refer to Appendix D.	42
5.2	Cuts defined in <code>Hlt2XcMuXForTauB2XcMu</code>	43
5.3	Base offline selections (stripping) for both D^0 and D^* channels.	44
5.4	Additional offline selections shared among D^0 and D^* channels.	45
5.5	Offline cuts exclusive to D^0 channel.	46
5.6	Offline cuts exclusive to D^* channel.	46
5.7	List of MC sample used in this analysis. Numbers are event on disk, not total simulated numbers.	48
5.8	List of MC filtering cuts at generator level.	49
5.9	Training variables for the isolation BDT.	50
5.10	Signal and control sample isolation requirements.	52
5.11	Offline cuts on $J/\psi K$ samples.	53
5.12	Additional MC control samples. All samples are full simulations.	54
6.1	Variables used to train <code>XGboost</code> for L0 hadron emulation.	58
6.2	Global event cuts for HLT1. Taken from [35].	62
7.1	Reweighting stages and binning schemes for B kinematic and occupancy reweighting.	73
7.2	Reweighting stages and binning schemes for final reweighting.	83
8.1	Fit variables and their binnings, obtained by the rest frame approximation technique as described in Appendix C.	87
8.2	Fit templates included in the D^0 channel.	89
8.3	Fit templates included in the D^* channel.	90
8.4	Normalization factors for D^0 signal fit with ISO skim templates. Color coding: red: fixed <code>NormFactor</code> ; blue: constrained <code>OverallSys</code> ; magenta: fixed <code>OverallSys</code> loaded from control fit.	91

8.5	Normalization factors for D^* signal fit with ISO skim templates. Color coding: red: fixed NormFactor; blue: constrained OverallSys; magenta: fixed OverallSys loaded from control fit.	92
8.6	Selections for each tagged species in fake μ sample.	96
8.7	Fitted yields of the signal fit.	118
9.1	Evaluated systematic uncertainties. Run 1 uncertainties are provided as a reference inside parentheses. All uncertainties are absolute, not relative. The additive uncertainties are sorted in descending order by their run 1 significance. Currently the MC statistical uncertainties are not taken into account.	124
12.1	Possible SALT-e-link combinations per data GBTx link in the DCB.	158

List of Figures

1.1	World average of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$, as of Nov 2022. Recent result from LHCb run 1 measurement is displayed as <i>LHCb 22</i> . The plot is taken from Heavy Flavor Averaging Group [3].	3
2.1	Parton level diagrams for $B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau$ signal decays, where $\bar{q}$ can be either a $\bar{u}$ or a $\bar{d}$	6
2.2	Modified Godfrey-Isgur mass predictions for D mesons. Taken from [4].	7
3.1	Angular variables of the $\bar{B}^0 \rightarrow D^{*+}l^-\bar{\nu}_l$ decays. For $B^- \rightarrow D^0l^-\bar{\nu}_l$ decays, only the θ_l variable exists.	20
4.1	The LHCb detector. z axis indicates $b\bar{b}$ travelling direction. The impact point is at $z = 0$	30
4.2	Simulated $b\bar{b}$ production angular distribution at $\sqrt{s} = 14$ GeV. Taken from [21].	31
4.3	Reconstructed track types at LHCb.	33
4.4	PID in RICH detectors.	36
4.5	Signatures of different particles in a typical configuration of a modern detector. Taken from [28].	37
4.6	Schematic of the L0 trigger system.	39
5.1	Preliminary μ BDT study. Left: μ BDT μ selection efficiency is flat in p_T , with global cut $\text{isMu\&D}LL\mu > 2$, and $\mu\text{BDT} > 0.25$. Right: with the same global cut, μ BDT is more effective in rejecting π than LHCb official MC15TuneV1ProbNN μ . The K rejection efficiencies are similar between the two; the p rejection efficiency is lower for μ BDT, but the absolute rejection efficiency is high enough.	55
6.1	The emulated trigger probabilities applied as a weight to FullSim has excellent agreement with real L0Hadron TOS trigger <i>simulated</i> in FullSim.	59
6.2	Check that L0Global TIS is portable among signal and normalization MC modes	60
6.3	Emulated HLT1 triggers vs. real response in FullSim. For B^+ . The Hlt1TrackMVA agree very well, whereas the Hlt1TwoTrackMVA still has a constant, albeit small ($\sim 2.25\%$ in q^2), discrepancy.	63
6.4	Emulated HLT1 triggers in B^0 have similar level of agreement to real response compared to that of B^+	63

7.1	Form factor reweight for $B \rightarrow D^0 l \bar{\nu}_l$ MC samples.	66
7.2	Form factor reweight for $B \rightarrow D^* l \bar{\nu}_l$ MC samples.	67
7.3	Form factor reweight effect on $D^{**} \mu$ MC templates, with shifted parameters based an initial fit.	68
7.4	Form factor reweight effect on $D^{**} \tau$ MC templates, with shifted parameters.	69
7.5	Tracking efficiency ratios $\frac{\epsilon_{\text{data}}}{\epsilon_{\text{MC}}}$. For 2016, only the table for the <i>long</i> method is available.	70
7.6	For slow π , about 50% of tracks lies outside the binning region of the official TRACKCALIB binning range which is displayed by the dashed black box. The colored rectangles show number of events fall in each official TRACKCALIB bin. To avoid loss in selection efficiencies, correction ratios from nearby bin is applied. The study is done on $\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu$ MC sample.	71
7.7	Fit on $J/\psi K$ with a <i>sPlot</i> procedure. The signal is modelled by a double Gaussian plus a Crystal bell; the background an exponential. Left and right show the same plot, with left having a linear y axis, and right a log y axis.	72
7.8	Effect of B kinematic and occupancy reweighting, derived from $J/\psi K$ control samples. The “MC (after)” plots contain corrections from <i>all stages</i>	73
7.9	Final reweighting for D^0 channel, for a total of 11 stages. The black plots contain weights from <i>all previous</i> stages, whereas the red ones contain weights from <i>all</i> stages.	76
7.10	Final reweighting for D^0 channel, stages 2–5 (cont’d).	77
7.11	Final reweighting for D^0 channel, stages 6–9 (cont’d).	78
7.12	Final reweighting for D^0 channel, stages 10–11 (cont’d).	79
7.13	Final reweighting for D^* channel, for a total of 13 stages. The black plots contain weights from <i>all previous</i> stages, whereas the red ones contain weights from <i>all</i> stages.	79
7.15	Final reweighting for D^* channel, stages 6–9 (cont’d).	81
7.16	Final reweighting for D^* channel, stages 10–13 (cont’d).	82
7.17	Fitted m_{miss}^2 distributions in the D^* channel. Without additional vertex resolution correction, the agreement between data and fitted result becomes poorer after final reweighting, as the pulls become worse. Applying a vertex resolution smearing brings m_{miss}^2 back to good agreement. These plots are taken from an older fit configuration and do not represent final fit results.	84
8.1	Allowing the DDX relative fraction $\rho_{DDX,u/d}$ to float in the signal fit seemingly improves fit quality (smaller pull for each q^2 bin), but the signal are now completely absorbed in the DDX components. Only the D^* channel fit is shown as it shows the most improvements in terms of pulls.	86
8.2	Comparisons between misID templates. The misID templates are normalized so that only shape comparisons are meaningful.	93
8.3	misID tagged vs. unfolded.	95
8.4	Weighted yields of fake μ sample.	95
8.5	Comparison between combinatorial background fit templates.	97

8.6	D^* combinatorial auxiliary fit to the ISO skim. Right plot shows the same fit as in the left, but with a logarithmic y axis.	98
8.7	Effect of scaling for the ISO skim on the D^* combinatorial template. . .	98
8.8	B combinatorial fit result for D^* channel	100
8.9	B combinatorial fit result for D^* channel. Left: B combinatorial auxiliary fit. Right: The scaling effect on the WS μ B combinatorial template. . .	100
8.10	Comparisons between normalization templates.	102
8.11	Comparisons between the signal templates.	104
8.12	Comparison between muonic D^{**} templates.	107
8.13	Comparison between tauonic D^{**} templates.	107
8.14	Comparison between $D_H^{**}\mu$ and $D^0\tau^-$ signal templates.	108
8.15	Comparisons between $D_s^{**}\mu$ templates.	110
8.16	Comparison between DDX templates.	111
8.17	$D_1^+\mu^-$ nominal (black), $+1\sigma$ (red), and -1σ (blue) variation fit template for the parameter τ' in the BLR form factor parameterization.	113
8.18	Variations of the $B^- \rightarrow DDX$ template, with the nominal template colored black, the linear $\pm$ colored red/blue, and the quadratic $\pm$ colored orange/cyan.	114
8.19	Variations of K^* contributions of the $B^- \rightarrow DDX$ template. Color coding: nominal, up, and down.	114
8.20	Variations of $B^0 \rightarrow D_H^{**}\mu$ based on the true q^2 of the decay processes. Color coding: nominal, up, and down.	115
8.21	Effect of DiF on fit variables	117
8.22	A comparison between the smeared and unsmeared ISO skim misID templates in the D^0 channel.	117
8.23	Control fit for 1OS sample, D^0 channel.	119
8.24	Control fit for 2OS sample, D^0 channel.	119
8.25	Control fit for DD sample, D^0 channel.	120
8.26	Control fit for 1OS sample, D^* channel.	120
8.27	Control fit for 2OS sample, D^* channel.	121
8.28	Control fit for DD sample, D^* channel.	121
8.29	Signal fit for ISO sample, D^0 channel.	122
8.30	Signal fit for ISO sample, D^* channel.	122
11.1	L0 trigger yields normalized to that at $2 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$. The hadronic triggers are already saturated at run 2 luminosity. Taken from [47]. . . .	136
11.2	The LHCb detector before (left) and after (right) LS2 upgrade. Note the upgrade on the tracking system: TT $\rightarrow$ UT, T-stations $\rightarrow$ SciFi; the removal of subdetectors relevant to L0 trigger only: the M1 muon station and SPD/PS have been removed.	137
11.3	Comparison between run 2 and run 3 trigger schemes.	138
11.4	Comparison between data paths for a CPU- and GPU-based HLT1 trigger. . . .	140
11.5	GPU-based HLT1 throughputs on many GPUs with respect to their reported peak 32-bit TFLOPS (Tera Float Point Operation Per Second) performance. Taken from [54].	141

11.6	Upgraded tracking system of the LHCb detector.	142
11.7	The upgraded VELO provides better IP resolution, based on simulation. The IP resolution for current (black) and upgraded (red) VELO are both evaluated at a pile-up of 7.6 and a center of mass energy of 14 TeV, a typical run 3 condition.	142
11.8	UT provides better p_T resolution to the long tracks and reduces fake (ghost) tracks. Figures taken from [56].	143
11.9	Operation principle of SciFi, illustrated with a single fiber module viewed from top/bottom. Taken from [59].	144
12.1	Four detection layers of UT, arranged in a $x-u-v-x$ configuration.	146
12.2	Illustration of the measurement of $x-y$ track hit positions with two silicon-strip sensors. The second sensor is rotated by an angle. The intersection between the red and blue strip is the measured position of a track.	147
12.3	Overall UT configuration in the LHCb cavern. Taken from [61].	147
12.4	UT stave.	148
12.5	Production sensor modules and SALT ASICs.	149
12.6	UT data and control signal flow diagram.	150
12.7	A schematic drawing of 3 PEPI crates with DCBs and Pigtails plugged in. Taken from [61].	151
12.8	A production Backplane and its copper trace implement.	152
12.9	A production DCB with cooling backplate assembly. The master GBTx and GBT-SCA are marked. 6 data GBTxs are numbered from 1 to 6. The BP connector is marked as J3. 4 optical mezzanine connectors are marked as J6, J7, J16, J17. No optical mezzanine is connected.	154
12.10	A LVR. Taken from [68].	155
12.11	Comparison between the reference LHCb front-end electronic and DCB.	156
12.12	An eye diagram for DCB, produced overnight. Wide-open and has no noise in the eye, an indication of good signal quality.	161
12.13	DCB CERN testing panel.	163
12.14	An overview of the LHCb online system.	164
12.15	A diagram illustrating DIM control flow.	166
12.16	An overview of the UT online system.	167
12.17	A prototype PCIe40 card.	169
12.18	The WinCC OA MiniDAQ projection interface to program a single data GBTx. It supports both programming individual registers (as shown in this figure) or programming all registers with a text file.	170
12.19	A WinCC memory monitoring panel.	171
12.20	E-link phase adjust with nanoDAQ.	174
A.1	Form factor reweight effect on $D^{**}\mu$ MC templates. Parameters are taken verbatim from [18].	176
A.2	Form factor reweight effect on $D^{**}\tau$ MC templates. Parameters are taken verbatim from [18].	177
A.3	D^* combinatorial auxiliary fit to the 1OS skim.	177

A.4	D^* combinatorial auxiliary fit to the 2OS skim.	178
A.5	D^* combinatorial auxiliary fit to the DD skim.	178
A.6	Effect of scaling for the 1OS, 2OS, and DD skims on the D^* combinatorial templates.	178
A.7	Check if TISTOS method works on our MC samples.	179
A.8	Pre-control fit of <i>initial fit</i> in D^0 channel, on 1OS, 2OS, and DD samples.	180
A.9	Pre-control fit of <i>initial fit</i> in D^* channel, on 1OS, 2OS, and DD samples.	181
A.10	Control fit of <i>initial fit</i> in D^0 channel, on 1OS, 2OS, and DD samples. The Gaussian likelihood penalty on D^{**} form factor uncertainties are kept.	182
A.11	Control fit of <i>initial fit</i> in D^* channel, on 1OS, 2OS, and DD samples. The Gaussian likelihood penalty on D^{**} form factor uncertainties are kept.	183
A.12	Signal fit of <i>initial fit</i> in D^0 channel on ISO samples.	183
A.13	Signal fit of <i>initial fit</i> in D^* channel on ISO samples.	184
A.14	Pre-control fit for 1OS sample, D^0 channel.	185
A.15	Pre-control fit for 2OS sample, D^0 channel.	186
A.16	Pre-control fit for DD sample, D^0 channel.	186
A.17	Pre-control fit for 1OS sample, D^* channel.	187
A.18	Pre-control fit for 2OS sample, D^* channel.	187
A.19	Pre-control fit for DD sample, D^* channel.	188
C.1	Schematic diagram for rest frame approximation (RFA).	192
D.1	Schematic of the trigger categories.	193
E.1	Subsamples of t -enriched PIDCALIB2 samples used in unfolding.	195

List of Abbreviations

ASIC	Application-Specific Integrated Circuit
ATLAS	A Toroidal LHC ApparatuS
BP	BackPlane
CERN	European Organization for Nuclear Research (Conseil Européen pour la Recherche Nucléaire)
CMS	Compact Muon Solenoid
CPU	Central Processing Unit
DCB	Data Control Board
FF	Form factor
FullSim	Full simulation
GBLD	GigaBit Laser Driver
GBT	GigaBit Transceiver
GPU	Graphics Processing Unit
IT	Inner Tracker
LFU	Lepton flavor universality
LHC	The Large Hadron Collider
LHCb	The Large Hadron Collider beauty
LVR	Low Voltage Regulator
MC	Monte-Carlo simulation
MEP	Multiple Event Packet
NP	New physics
nTracks	Number of tracks
OT	Outer Tracker
PDG	Particle data group, often synonymous to reference values
PEPI	Peripheral Electronics
PID	Particle identification
PV	Primary vertex (pp vertex)
QED	Quantum Electrodynamics

RFA	Rest frame approximation
RICH	Ring Imaging CHerenkov detector
SCA	Slow Control Adapter
SciFi	Scintillating Fibre tracker
SIMD	Single Instruction, Multiple Data
SM	The standard model of particle physics
TO	Tracker-only
TT	Tracker Turicensis
UT	Upstream Tracker
VELO	VErtex LOcator
VEV	Vacuum expectation value

Chapter 1: Introduction

The standard model of particle physics (SM) is the quantum field theory describing the interactions of elementary particles, namely quarks, leptons, and gauge bosons. It has been successful in explaining many experimental results, such as parity violation in the weak interaction, as well as predicting the existence of new particles, such as the Higgs boson. However, there are experimental observations that cannot be explained by the SM, such as the existence of dark matter which has no viable candidate in SM. These observations demand new physics (NP). One way to probe and constrain NP is through precision measurements in which observables associated with certain processes are measured very precisely and compared with predictions from SM.

One such process is the semileptonic decay of the B mesons: $B \rightarrow D^{(*)}l^-\bar{\nu}_l$, where l^- stands for a charged lepton which can be an electron e^- , a muon μ^- , or a tau τ^- . SM predicts lepton flavor universality (LFU), that is, leptons participate in SM processes with the same strength except for the Higgs mechanism through which the leptons acquire different masses. Therefore, the ratios¹ of branching fractions $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$, collectively referred as $\mathcal{R}(D^{(*)})$, defined as:

$$\mathcal{R}(D^{(*)}) \equiv \frac{\mathcal{B}(B \rightarrow D^{(*)}\tau^-\bar{\nu}_\tau)}{\mathcal{B}(B \rightarrow D^{(*)}\mu^-\bar{\nu}_\mu)} \quad (1.1)$$

will not be 1 only because τ^- and μ^- have different masses. These ratios have precise SM predictions with uncertainties small than 4%. Thus, the ratios can serve as a probe to LFU: a significant deviation (often chosen to be at 5σ) from SM prediction is an indication of

¹Throughout the thesis, charge conjugation is assumed. Also, the convention $\hbar = c = 1$ is used unless specified explicitly.

LFU violation, which may provide hints and constraints to NP.

Measuring $\mathcal{R}(D^{(*)})$ is advantageous compared to the measurement of absolute branching fractions because several theoretical and experimental uncertainties are fully or partially cancelled in the ratios. For example, $\mathcal{R}(D^{(*)})$ have no dependence on the CKM matrix element V_{cb} , and its dependence on the muon reconstruction efficiency is largely reduced if the tau lepton in the numerator is reconstructed via the $\tau \rightarrow \mu \bar{\nu}_\mu \nu_\tau$ decay.

Since 2012, tensions have been reported between the measured $\mathcal{R}(D^{(*)})$ and the SM predictions from BaBar, Belle, and LHCb. Currently, the tension stands at about 3σ after averaging all measurements, as shown in Fig. 1.1. In 2015, the LHCb collaboration reported its first result on $\mathcal{R}(D^*)$ based on LHCb run 1 (2011–2012) data [1], overcoming the difficulties of approximating key kinematic variables that are not directly computable in a hadron collider experiment and demonstrating the viability of characterizing background decays through several control samples. Recently, a follow-up measurement [2], labelled as *LHCb22* in Fig. 1.1, reported a simultaneous extraction of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$, with the same run 1 data, based on the techniques developed in the 2015 result but with a better correction procedure for Monte-Carlo simulation samples, a novel treatment of the particles misidentified as muon, and an update of various form factor models used in the simulation. In the first part of this thesis, the author developed the framework to update the LHCb $\mathcal{R}(D^{(*)})$ measurement with the run 2 (2016–2018) data by porting and refining procedures developed in the run 1 pathfinder analyses and solving additional challenges imposed by the size of the run 2 dataset which is about 4 times larger compared to the size of run 1. The developed procedure should form the basis for the analysis of all 3 years of run 2 data.

The LHCb detector is undergoing an upgrade currently² which is expected to greatly increase the readout rate of the detector and removes the hardware triggers along with the limitations that come with them. The LHCb group at the University of Maryland is actively participating in the upgrade of the Upstream Tracker (UT), a part of the tracking

²Defined as Nov 2022 throughout the text.

system. The author was deeply involved in the design and development of the UT readout electronics, so the second part of the thesis will provide an overview of the LHCb upgrade, with a focus on the UT detector and its readout system.

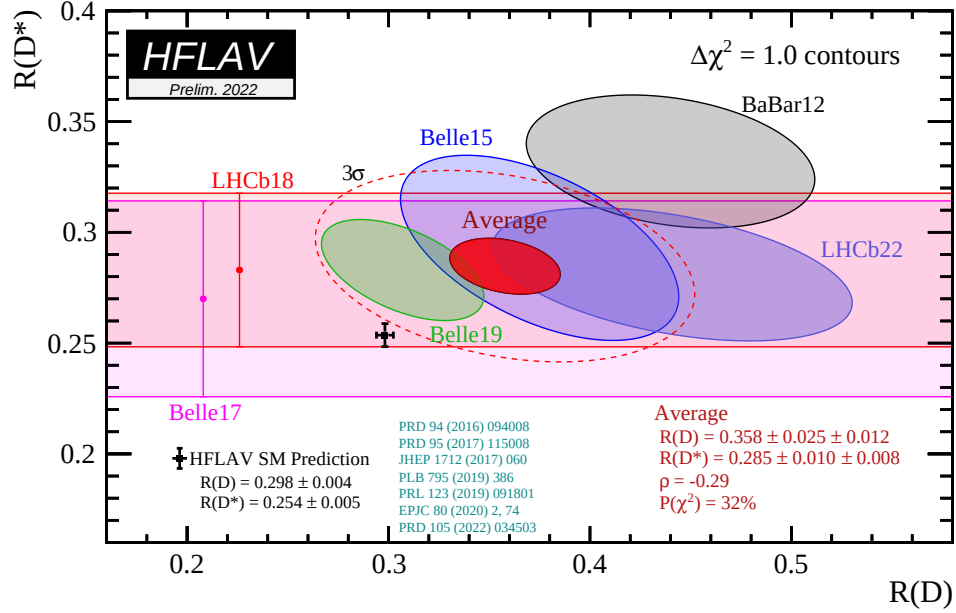


Figure 1.1: World average of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$, as of Nov 2022. Recent result from LHCb run 1 measurement is displayed as *LHCb 22*. The plot is taken from Heavy Flavor Averaging Group [3].

Part I

Preliminary measurement of $\mathcal{R}(D^{(*)})$ with LHCb 2016 data

Chapter 2: Analysis overview

This analysis aims at the measurement of $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$, with the definition reproduced from the introduction:

$$\mathcal{R}(D^{(*)}) \equiv \frac{\mathcal{B}(B \rightarrow D^{(*)} \tau^- \bar{\nu}_\tau)}{\mathcal{B}(B \rightarrow D^{(*)} \mu^- \bar{\nu}_\mu)} \quad (2.1)$$

from LHCb 2016 data¹, reconstructed in (with visible final state particles marked in red):

- $\bar{B}^0 \rightarrow D^{*+}(\rightarrow D^0(\rightarrow K^- \pi^+) \pi^+) \tau^- (\rightarrow \mu^- \bar{\nu}_\mu \nu_\tau) \bar{\nu}_\tau$ for $\mathcal{R}(D^*)$ signal decay²
- $\bar{B}^0 \rightarrow D^{*+}(\rightarrow D^0(\rightarrow K^- \pi^+) \pi^+) \mu^- \bar{\nu}_\mu$ for $\mathcal{R}(D^*)$ normalization decay
- $B^- \rightarrow D^0(\rightarrow K^- \pi^+) \tau^- (\rightarrow \mu^- \bar{\nu}_\mu \nu_\tau) \bar{\nu}_\tau$ for $\mathcal{R}(D)$ signal decay
- $B^- \rightarrow D^0(\rightarrow K^- \pi^+) \mu^- \bar{\nu}_\mu$ for $\mathcal{R}(D)$ normalization decay

such that $D^{*+} \mu^-$ and $D^0 \mu^-$ are the final reconstructed states which are often referred in this text as D^* channel and D^0 channel in later chapters. These states are chosen so that the final state particles (e.g. K^- , π^+ , μ^-) are all charged, because the LHCb detector is not well-suited for neutral particle reconstruction.

Feed down A complication of these states is that they are *inhomogeneous*, that is, not all $D^0 \mu^-$ final state events are actual decays with just $D^0 \mu^-$ in their final states: some $\bar{B}^0 \rightarrow D^{*+} l^- \bar{\nu}_l$ decays feed down into the $D^0 \mu^-$ states, because D^{*+} decays strongly to a D^0 and a slow π^+ , with the latter often missed in the reconstruction; for example, a slow π^+

¹It is planned to use full LHCb run 2 data (2016–2018) at a later date.

²A parton level diagram for the signal decays is drawn in Fig. 2.1 as a reference.

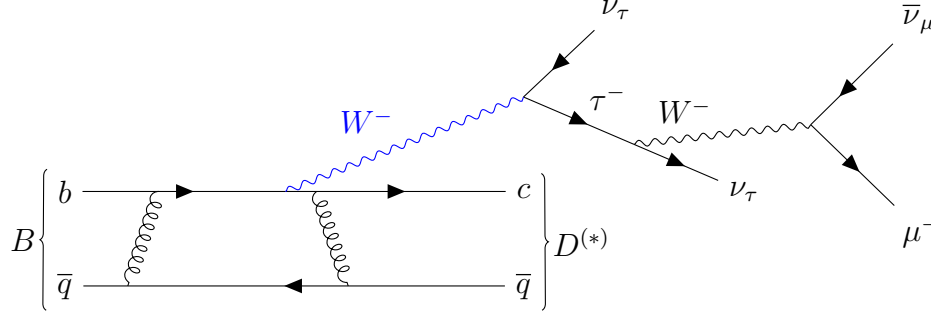


Figure 2.1: Parton level diagrams for $B \rightarrow D^{(*)} \tau^- \bar{\nu}_\tau$ signal decays, where $\bar{q}$ can be either a $\bar{u}$ or a $\bar{d}$.

with a momentum³ of 10 GeV will fail to register a well-defined track in the detector about 35% of time, thus no D^{*+} is constructed for this event. In addition, $B^- \rightarrow D^{*0} l^- \bar{\nu}_l$ decays, which have a branching fraction of about 2.5 times larger compared to $B^- \rightarrow D^0 l^- \bar{\nu}_l$ decays, contribute exclusively to $D^0 \mu^-$ states, because the neutral slow π^0 , a product of the D^{*0} decay, is entirely missed. Therefore, a *simultaneous* fit, assuming isospin symmetry, to both $D^{*+} \mu^-$ and $D^0 \mu^-$ states is needed to account for the correlations between $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ and improve the precision⁴ of $\mathcal{R}(D^*)$.

Background contributions In addition to the signal, normalization, and feed down from D^* modes, both $D^{*+} \mu^-$ and $D^0 \mu^-$ final states contain contributions from many background decay modes that are only partially reconstructed, which can be categorized as follows:

- Four $1P$ D^{**} : these are the lightest excited D mesons. These D^{**} states decay in a cascade manner: $B \rightarrow D^{**} (\rightarrow D^{0|*|**} (\rightarrow D^{0|*} \pi) \pi) l \nu_l$, contributing to both D^0 and D^* channel with additional charged or neutral pions in the final states.
- Highly excited D^{**} (D^{**} heavy, D_H^{**}): these refers to the excited D mesons that are heavier than the four $1P$ D^{**} states. The predicted masses of the known D mesons, including D^0 , D^* , the four $1P$ D^{**} , and highly excited D^{**} , are displayed in Fig. 2.2.

³The mean momentum of a slow π^+ is around 8 GeV. Generally a lower momentum implies a lower track finding efficiency.

⁴As opposed to extracting $\mathcal{R}(D^*)$ from the $D^{*+} \mu^-$ sample alone.

Similar to the four $1P$ D^{**} states, the D^{**} heavy states cascade decay into a D^0 or a D^* , with two or more extra pions in the final states: $B \rightarrow D_H^{**}(\rightarrow D^{0|*}\pi\pi)\mu\nu_\mu$.

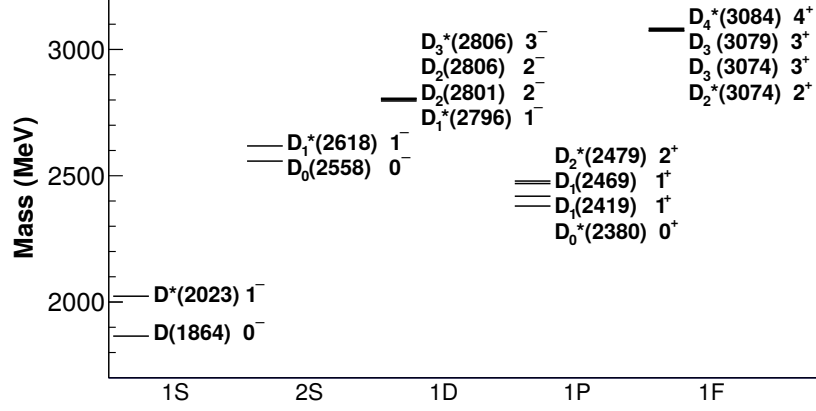


Figure 2.2: Modified Godfrey-Isgur mass predictions for D mesons. Taken from [4].

- D_s^{**} : the $B_s \rightarrow (D'_{s1}|D_{s2})(\rightarrow D^{(*)}K)l\nu_l$ decays have one extra kaon in their final states.
- DDX : the double-charm backgrounds, namely $B \rightarrow D^{(*)}D_qX$ decays where D_q stands for a generic charm meson with q being a $u/d/s$ quark and X represents extra pion(s), a kaon, or an excited K states. The lepton in such decays comes from a leptonic (when $q = s$) or a semileptonic (when $q = u$ or d) decay of the D_q . These decays contain many additional final state particles with at least one kaon.

The backgrounds due to *mis-reconstructions* of the $D^{(*)}\mu$ pair must also be taken into account, namely:

- Muon misID: sometimes the μ in the $D^{(*)}\mu$ pair is mis-identified because the particle identification algorithm is not 100% accurate. When such misidentification happens, the “ μ ” can be a pion, a kaon, a proton, an electron, or a ghost⁵.
- Combinatorial backgrounds: these refer to random combinations of the $D^0\mu$ (for the

⁵Fake tracks due to random combinations of track segments, a detector effect.

D^0 channel), $D^*\mu$, or $D^0\pi$ (the latter two are for the D^* channel) that do not originate from the decay of the same B meson.

Signal and control skims The $D^{(*)}\mu$ pairs are selected such that they are likely coming from the same B decays. However such selection criteria does not take the extra track(s) in the events from the background decays into account. The selected $D^{(*)}\mu$ samples are further split into signal-enriched and background-enriched sub-samples, named as “skims”, based on the isolation property of the $D^{(*)}\mu$ pair, that is, how likely that other well-reconstructed tracks in the same event are originating from the same B as the $D^{(*)}\mu$. There are four resulting skims, namely:

- ISO: signal-enriched. The $D^{(*)}\mu$ pair is well-isolated as there is no other track in the event that is likely coming from the same B .
- 1OS: enriched in $B \rightarrow D^{**}l\nu_l$. The 1OS skim contains events with one extra pion that is originating from the same B as the $D^{(*)}\mu$ pair (anti-isolated).
- 2OS: enriched in $B \rightarrow D_H^{**}\mu\nu_\mu$. Contain events with two anti-isolated pions.
- DD: enriched in $B \rightarrow D^{(*)}D_qX$. Contain events with one or more anti-isolated tracks with one of them being a kaon.

Fit overview The signal, normalization, and backgrounds all have differing kinematics which are for the most part modeled from Monte-Carlo simulation (MC), and sometimes from data control samples. These differences are used in the fit to determine the yields of all relevant decay modes. The differentiating kinematic variables, defined in the B rest frame, referred as “fit variables”, are the following:

- m_{miss}^2 : Defined as $(p_B - p_{D^{(*)}} - p_\mu)^2$, the missing mass (squared) is typically due to missing (unreconstructed) neutrino(s). For $B \rightarrow D\mu^-\bar{\nu}_\mu$ decays, only 1 neutrino is

missing, so m_{miss}^2 is small ($\sim$ the mass of neutrino); for $B \rightarrow D\tau^-(\rightarrow \mu^-\bar{\nu}_\mu\nu_\tau)\bar{\nu}_\tau$, 3 neutrinos are missing, and m_{miss}^2 can be large.

- E_μ^* : The energy of μ^- in the B rest frame. μ^- coming from τ^- decays are typically softer, i.e. with smaller energy, due to reduced available phase space.
- q^2 : The momentum transfer, defined as the invariant mass squared of the virtual W boson (marked in blue in Fig. 2.1) and calculated as $(p_B - p_{D^{(*)}})^2$. The q^2 spectra of $B \rightarrow P\ell\nu$ and $B \rightarrow V\ell\nu$ decays are notably different, where “P” stands for a pseudoscalar meson (e.g. D^0), and “V” a vector (e.g. D^*); the q^2 spectrum is softer (more likely to be at lower q^2) for pseudoscalar mesons. In addition, the semitauonic decays are restricted to the phase space where $q^2 > m_\tau^2$, whereas the semimuonic decays extend down to $q^2 > m_\mu^2 \approx 0$. Both modes share a common q^2 upper bound $(m_B - m_{D^{(*)}})^2$ where the D meson is produced at rest in the B rest frame.

These variables, however, cannot be deduced exactly at LHCb: because B mesons are typically produced from hadronization of pp collisions, carrying an unknown portion of pp momenta, the B momentum and its rest frame are not known. A procedure, termed “rest frame approximation” (RFA), is developed for LHCb to estimate the B rest frame, making computation of fit variables possible. More information about RFA can be found in Appendix C.

This analysis uses the same general fit strategy in the LHCb $\mathcal{R}(D^{(*)})$ run 1 analysis: fit the parameters characterizing various background decays with separate *control skims* (1OS, 2OS, and DD), then import these parameters either as constraints or fully fixed into the signal fit; this is because while the *signal skim* (ISO) contains many backgrounds, they are only a small fraction relative to the signal and normalization combined, thus no good description of the backgrounds can be extracted from the signal fit alone.

Workflow As an update to the previous analysis, this analysis shares a similar workflow with that of run 1. The workflow of this analysis is: first, $D^{*+}\mu^-$ and $D^0\mu^-$ final states are

selected with the procedure described in Chapter 5. Then, emulation of certain detector responses to MC are carried out offline with the procedure described in Chapter 6, as this analysis requires a large number of simulated events with only partial detector simulation to be computationally viable. Afterwards, various mis-modelings in the MC simulation, mainly in form factor modeling and detector response, are updated and corrected with the procedures detailed in Chapter 7. Finally, $\mathcal{R}(D^{(*)})$ are extracted from a simultaneous fit described in Chapter 8. The estimation of major systematic uncertainties is discussed in Chapter 9.

Differences between the run 2 and run 1 analyses The major differences between this analysis and the pathfinder run 1 $\mathcal{R}(D^{(*)})$ analysis are:

- Due to a larger integrated luminosity (5.4 fb^{-1} in run 2 versus 3.1 fb^{-1} in run 1), a larger B production cross section due to higher collision energy (13 TeV center of mass energy in run 2 versus 7 TeV in run 1), and a dedicated trigger resulting in a higher selection efficiency, the run 2 analysis can work with a much larger dataset: for 2016 alone (1.6 fb^{-1}), the selected $D^0\mu^-$ signal sample contains 2,178,793 events, whereas for run 1 combined (3.1 fb^{-1}) 1,734,133 events, a factor of ~ 1.26 gain.

As briefly mentioned before, a large amount of MC is needed to control the total statistical uncertainties, which makes it impractical to simulate most of the detector. Instead, only the tracking system is simulated and additional detector responses such as hardware triggers are *emulated* offline.

- In order to update (reweight) MC form factor models offline, the run 1 analysis implemented two ad-hoc approaches for form factor reweighting: one is based on **Xs1FF** reweighting class from BaBar, the other is completely in-house. A large amount of time is spent on validation of the form factor reweighting implementations.

Recently, a new software package, named **Hammer**, allows easier form factor reweighting and comes pre-validated by the **Hammer** authors. This analysis leverages on **Hammer**

for an easier to implement and more reliable form factor reweighting procedure.

Chapter 3: Theory of semileptonic B decays

3.1 Construction of the electroweak theory in the SM

Through a close collaboration between experimentalists and theorists, particle physicists developed and validated a *quantum field theory* to describe *interactions of elementary particles*, known as the standard model of particle physics (SM), the most precise theory of the sub-atomic world to date. A construction of the SM is outlined¹ first with the help of [5, 6], followed by brief discussions on quantum electrodynamics (QED), the electroweak theory, and the lepton flavor universality in the SM.

3.1.1 Outline of the construction

It is experimentally established that the speed of light in vacuum is constant in any reference frame. Then, all transformations preserving the speed of light² can be identified: rotations, Lorentz boosts, parity ($\mathbf{x} \rightarrow -\mathbf{x}$), time reversal ($t \rightarrow -t$), and spacetime translations. All but the translation transformations form a group called the “Lorentz group”. Focusing on the restricted Lorentz group $SO^+(1, 3)$, the group that preserves the direction of space and time³: it has 6 generators J_i, K_i with $i \in \{1, 2, 3\}$, which obey the following commutation relations:

$$[J_i, J_j] = i\epsilon_{ijk}J_k \tag{3.1}$$

¹“I am an experimentalist!” says the author waving his hands.

²Each of such transformations, denote as L , must leave the Minkowski metric η intact: $L^T \eta L = \eta$.

³Loosely speaking, Lorentz group contains 4 “classes”: $\{I, T, P, TP\}$. $SO^+(1, 3)$ is the one that contains the identity element. The other 3 can be obtained by applying parity (P), time reversal (T), or both (TP) to $SO^+(1, 3)$.

$$[J_i, K_j] = i\epsilon_{ijk}K_k \quad (3.2)$$

$$[K_i, K_j] = -i\epsilon_{ijk}J_k \quad (3.3)$$

such that each Lorentz transformation L of boost $\boldsymbol{\phi}$ and rotation $\boldsymbol{\theta}$ can be generated with the exponential map: $L = e^{i\mathbf{J}\cdot\boldsymbol{\theta}+i\mathbf{K}\cdot\boldsymbol{\phi}}$. A new set of generators can be defined $N^\pm = \frac{1}{2}(J_i \pm iK_i)$ such that each set of $\{N_i^\pm\}$ are generators of the $SU(2)$ group, and N_i^+, N_j^- commute. Hence $SO^+(1, 3)$ contains two copies of $SU(2)$.

As $L \in SO^+(1, 3)$ is a Lorentz transformation, it can *act* on *some objects*. If we map each L to a $n \times n$ invertible matrix M of complex numbers, the objects L can act on are represented by n -dimensional complex vectors. A map from $SO^+(1, 3)$ to the (vector) space of M is called a “representation” of $SO^+(1, 3)$; there are many representations for a single group. As an example, each generator of $SU(2)$, denote as $\{N_1, N_2, N_3\}$, can be mapped to a Pauli matrix $N_i \mapsto \sigma_i$, forming a 2-dimensional representation of $SU(2)$. More generally, for a n -dimensional representation of $SU(2)$, let N_3 be represented by a diagonal matrix and its largest eigenvalue j . Due to commutation relation $[N_i, N_j] = i\epsilon_{ijk}N_k$, only N_3 can be diagonal. It can be shown that the spectrum of the N_3 eigenvalues is discrete: $\{-j, -j+1, \dots, j-1, j\}$ with $2j+1$ elements. Since N_3 is a n -dimensional invertible matrix, $2j+1 = n$. Hence, such a representation can be labelled with $j = \frac{n-1}{2}$ which is a half-integer.

Since $SO^+(1, 3)$ contains two copies of $SU(2)$, connections between the representations of these two groups can be made and the representations of $SO^+(1, 3)$ can be labelled as (j, j') . That is, Lorentz transformations can act on many different *classes* of objects, with each class corresponds to a (j, j') representation. Some of such representations are *irreducible*⁴, which correspond to specific types of *elementary particles* of spin $j+j'$: the $(0, 0)$ representation corresponds to spin 0 scalar particles; the $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ representations both corresponds to spin $\frac{1}{2}$ particles, with the former left-handed and the latter right-handed; the $(\frac{1}{2}, \frac{1}{2})$ representation corresponds to spin 1 particles. There are many more irreducible

⁴Roughly speaking, there is no transformation that can simultaneously block-diagonalize all matrices representing the group elements.

representations, but in the SM only the spin 0, $\frac{1}{2}$, and 1 particles are present.

Now the SM can be constructed as follows: first, consider elementary particles with spin 0, $\frac{1}{2}$, and 1, which are represented by linear operators in a space. A Lorentz-invariant Lagrangian density⁵ can be built from linear, quadratic, and the lowest possible derivative terms of these operators. A Lagrangian built in this manner contains no interaction (i.e. a free Lagrangian). Then, require the Lagrangian to be invariant⁶ under certain local gauge transformations such as that from $U(1)$, $SU(2)$, and $SU(3)$. For example, a global gauge transformation $\psi \rightarrow e^{i\alpha}\psi$ where $\alpha \in \mathbb{R}$ can be promoted into a local one by changing α from a number to a function of the spacetime coordinate $\alpha(x^\mu)$. All such local gauge transformations form a $U(1)$ group. These requirements can be satisfied by adding/removing certain terms to the Lagrangian; these terms make interactions between elementary particles happen. At this point, there is nothing quantum about the theory. The theory is “quantized” by forcing canonical commutation/anti-commutation relation for integer spin and half-integer spin particles, respectively⁷. Finally, the scattering amplitude from an initial state to a final state, which is of utmost importance experimentally, can be computed in a perturbative manner⁸.

⁵Will be referred to as “Lagrangian” in later text for brevity.

⁶Local gauge invariances represent mathematical redundancy in the description of the system, and must be preserved to have a consistent theory.

⁷Note that at this point, the terms *elementary particles*, *interactive*, and *quantum field theory* are defined.

⁸To be more precise, the scattering amplitude $\langle f | i \rangle$ between an initial state $|i\rangle$ and a final state $|f\rangle$ can be calculated in the following manner [7]:

1. The amplitude $\langle f(y_1, \dots, y_r) | i(x_1, \dots, x_n) \rangle$ is related to the interactive vacuum $|\Omega\rangle$ by the LSZ reduction formula: $\langle f | i \rangle \propto \langle \Omega | T \phi(y_1) \dots \phi(y_r) \phi(x_1) \dots \phi(x_n) | \Omega \rangle$, where $\phi(x)$, an operator in the Heisenberg picture, creates and destroys particles at spacetime point x in the interactive vacuum $|\Omega\rangle$, and T stands for time-ordering.
2. Denote the free and the interaction part of the Hamiltonian as H_0, H_{int} , accordingly.
3. At a given time t , a Heisenberg operator and a interactive operator are related by $\phi(t_0, \mathbf{x}) = U^\dagger(t, t_0) \Phi_I(t, \mathbf{x}) U(t, t_0)$, where $U(t, t_0) = T e^{-i \int_{t_0}^t H_I(t') dt'}$, and $H_I(t) = e^{iH_0(t-t_0)} H_{\text{int}} e^{-iH_0(t-t_0)}$.
4. The interactive vacuum $|\Omega\rangle$ and the free vacuum $|0\rangle$ are related by $|\Omega\rangle \propto \lim_{\tau \rightarrow \infty(1-i\epsilon)} U(t_0, -\tau) |0\rangle$.
5. The scattering amplitude can now be written as $\langle f | i \rangle \propto \langle 0 | T \prod_i \Phi_I(x_i) e^{-i \int_{-\infty}^{\infty} H_I(t) dt} | 0 \rangle$, where Φ_I is the ϕ operator in the interaction picture and has a *free* mode expansion that can create and destroy particles in the *free* vacuum $|0\rangle$.
6. At this point, the exponential involving H_I are expanded in a Taylor series: $e^{-i \int_{-\infty}^{\infty} H_I(t) dt} = \sum_n \frac{(-i)^n}{n!} T \left\{ \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} H_I^1 \dots H_I^n dt_1 \dots dt_n \right\}$. Each term in the series are turned into a series of normal-ordered operators with contractions via the Wick's theorem.

3.1.2 Quantum eletrodynamics (QED)

Following the recipe provided above, first identify all⁹ possible Lorentz scalar terms in a free Lagrangian for spin $\frac{1}{2}$ particles: the second order terms $S_1 = (\chi_L)^\dagger \xi_R$, $S_2 = (\xi_R)^\dagger \chi_L$, and the first order partial derivative terms (identified as the kinetic terms) $S_3 = (\chi_L)^\dagger \partial_\mu \bar{\sigma}^\mu \chi_L$, $S_4 = (\xi_R)^\dagger \partial_\mu \sigma^\mu \xi_R$, where each χ_L, ξ_R is a 2-component Weyl spinor. This shows that both left- and right-handed particles are needed for Lorentz invariance, which motivates a 4-component Dirac spinor $\psi = \begin{pmatrix} \chi_L \\ \xi_R \end{pmatrix}$. To combine ψ into the form of S_1, S_2 , the matrix $\gamma_0 = \begin{pmatrix} 0 & I_{2 \times 2} \\ I_{2 \times 2} & 0 \end{pmatrix}$ can be used such that $\psi^\dagger \gamma_0 \psi$ is a Lorentz scalar. On the other hand, the forms of S_3, S_4 suggest the use of the matrices of the form $\begin{pmatrix} \bar{\sigma}^\mu & 0 \\ 0 & \sigma^\mu \end{pmatrix}$, which are generated by $\gamma_0 \gamma^\mu$, where $\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}_\mu & 0 \end{pmatrix}$, noting that γ^0 can be defined in the same way since $\sigma^0 = \bar{\sigma}^0 = I_{2 \times 2}$. Therefore, a Lagrangian $\mathcal{L} = a\psi^\dagger \gamma_0 \psi + b\psi^\dagger \gamma_0 \gamma^\mu \partial_\mu \psi$ can be written, where $a, b \in \mathbb{C}$. Defining a shortcut $\bar{\psi} \equiv \psi^\dagger \gamma_0$, and requiring that the Lagrangian yields the familiar Dirac equation $(i\gamma_\mu \partial^\mu - m)\psi = 0$ to set $a = -m$ and $b = i$, we arrive at the Dirac Lagrangian which describes free spin $\frac{1}{2}$ particles of mass m :

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \quad (3.4)$$

The Lagrangian is invariant under a global gauge transformation $\psi \rightarrow e^{i\alpha}\psi$ where $\alpha \in \mathbb{R}$, but *not* a local $U(1)$ one where α is replaced with $\alpha(x^\mu)$, due to the derivative term ∂_μ . To enforce such invariance, A spin 1 field operator A_μ , which satisfies the gauge transformation $A_\mu \rightarrow A_\mu + \frac{1}{g}\partial_\mu \alpha(x^\mu)$, is introduced by replacing the standard derivative ∂_μ with the covariant derivative, defined as:

$$\mathcal{D}_\mu \equiv \partial_\mu - igA_\mu \quad (3.5)$$

at this point, the Lagrangian reads $\mathcal{L} = \mathcal{L}_{\text{Dirac}} + gA_\mu \bar{\psi} \gamma^\mu \psi$. However, the equation of motion of A_μ suggests the interaction term should vanish: $gA_\mu \bar{\psi} \gamma^\mu \psi = 0$, because $0 = \partial_{A_\mu} \mathcal{L} =$

7. The contractions are Feynman propagators with the corresponding field operators which are specified by the commutation relations between the operators.

8. The first (non-trivial non-vanishing) order terms are identified as tree-level Feynman diagrams.

Higher order terms may involve infinities and require renormalization.

⁹This refers to linear, quadratic, and lowest possible derivative terms.

$g\bar{\psi}\gamma^\mu\psi$. To retain the interaction term, an additional kinetic term for spin 1 particles is needed, which is found by doing an exercise of finding all Lorentz-invariant derivative terms, as in the spin $\frac{1}{2}$ case. Such term reads $\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$, where $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$. Putting all pieces together, we obtain the Lagrangian for QED:

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\gamma^\mu \not{\partial}_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (3.6)$$

3.1.3 The electroweak theory

As demonstrated in Section 3.1.2, the local gauge invariance principle provides a method of introducing interactions to free theories. Motivated by the experimental observation that in beta decay, electrons and electron neutrinos are produced in pairs, we can add two copies of spin $\frac{1}{2}$ particles, labelled as ψ_e, ψ_{ν_e} , to the Lagrangian¹⁰ and write both in a doublet of the form $\Psi = \begin{pmatrix} \psi_e \\ \psi_{\nu_e} \end{pmatrix}$, such that one particle can be rotated into the other via a $SU(2)$ rotation $e^{i\alpha_i \frac{\sigma^i}{2}}$.

It is then natural to postulate a theory of $SU(2) \times U(1)$ local gauge symmetry. Following a similar route in QED, noting that $SU(2)$ has 3 non-commutative generators and $U(1)$ has 1 generator, three spin 1 particles W^1, W^2, W^3 are added to the theory for $SU(2)$ in addition to one spin 1 particle B for $U(1)$. After taking the fact non-commutative nature of the $SU(2)$ generators into account, a Lagrangian with correct kinetic terms can be constructed, in which the covariant derivative is defined as:

$$\not{D}_\mu = \partial_\mu - i[g_2 W_\mu^a T_{SU(2)}^a + g_1 B_\mu Y_{U(1)}] \quad (3.7)$$

where $T_{SU(2)}^a = \frac{1}{2}\sigma^a$ are the generators of $SU(2)$, and $Y_{U(1)} = C\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. However, naive mass terms $\bar{\Psi}M\Psi$ are forbidden because they spoil $SU(2)$ gauge symmetry: an $SU(2)$ transformation $e^{i\alpha_i \frac{\sigma^i}{2}}$ transforms such terms into $\bar{\Psi}e^{-i\alpha_i \frac{\sigma^i}{2}} M e^{i\alpha_i \frac{\sigma^i}{2}} \Psi$ where $e^{-i\alpha_i \frac{\sigma^i}{2}} M e^{i\alpha_i \frac{\sigma^i}{2}} \neq M$ because $[\sigma^i, M] \neq 0$ in general as $M \neq mI_{2 \times 2}$. So the theory constructed above requires *all* particles to be massless, which is inconsistent with the experimental results of massive

¹⁰The Lagrangian reads $\mathcal{L} = \bar{\Psi}(i\gamma^\mu \not{D}_\mu I_{2 \times 2} - M)\Psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$, where $M = \begin{pmatrix} m_e & 0 \\ 0 & m_{\nu_e} \end{pmatrix}$. Note that the identity matrix is often omitted.

force carrying bosons in the weak interaction¹¹, nor a massive electron.

To introduce *gauge invariant* mass-like terms, a complex scalar field is introduced with a potential $V(\phi^\dagger, \phi) = \frac{1}{4}\lambda(\phi^\dagger\phi - \frac{1}{2}v^2)^2$, which leads to a non-zero vacuum expectation value (VEV): $\langle 0|\phi|0\rangle = v$ for ϕ degenerate on a complex circle. Write the complex scalar fields as a $SU(2)$ doublet $\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$, then pick a particular vacuum and expand around it such that $\Phi = \frac{1}{\sqrt{2}}\begin{pmatrix} v+h \\ 0 \end{pmatrix}$ where $h(x^\mu)$ is a real scalar field known as the Higgs field. Doing so breaks the degeneracy (symmetry) of the vacuum, an effect called spontaneous symmetry breaking. The couplings between Φ and the spin 1 gauge fields are embedded in the kinetic term $\bar{D}_\mu \Phi D^\mu \Phi$, and a non-zero VEV leads to terms of the form $\frac{1}{2}\kappa W_\mu W^\mu$ where W is a linear combination of W^1, W^2 , and κ is a coupling constant which can be interpreted as the mass of the W . Indeed, after a field redefinition, we find three massive spin 1 fields and a massless one:

$$W_\mu^+ = \frac{1}{\sqrt{2}}(W_\mu^1 - iW_\mu^2) \quad (3.8)$$

$$W_\mu^- = \frac{1}{\sqrt{2}}(W_\mu^1 + iW_\mu^2) \quad (3.9)$$

$$Z_\mu = \cos \theta_w W_\mu^3 - \sin \theta_w B_\mu \quad (3.10)$$

$$A_\mu = \sin \theta_w W_\mu^3 + \cos \theta_w B_\mu \quad (3.11)$$

where $\theta_w = \tan^{-1}\left(\frac{g_1}{g_2}\right)$, and $m_{W^\pm} = \frac{g_2 v}{2} = m_W$, $m_Z = \frac{m_W}{\cos \theta_w}$, $m_A = 0$. The procedure above is referred to as the Higgs mechanism.

Still, the spin $\frac{1}{2}$ particles are massless. Furthermore, the weak interaction is experimentally shown to maximally violate parity, that is, only the left-handed particles interact weakly. Formally speaking, only the left-handed particles form $SU(2)$ doublets which transform as $\Psi_L \rightarrow e^{i\alpha_i \frac{\sigma^i}{2}} \Psi_L$; the right-handed particles are $SU(2)$ singlets which do not transform $\psi_R \rightarrow \psi_R$. As a side note, previously we stated that $SU(2)$ gauge symmetry forbids terms like $\bar{\Psi} M \Psi$ because $M \neq m I_{2 \times 2}$ in general, (e.g. electron and electron neutrinos have *very* different masses). Parity violation provides a different view point on the same issue: naive

¹¹If W^1, W^2, W^3 are all massless, the weak interaction would behave like electromagnetism, which it does not.

mass terms contain *only* combinations of left-handed doublets and right-handed singlets, such as $\bar{\Psi}_L \psi_R$. Inspecting the $SU(2)$ transformation property of such terms reveals that they are not $SU(2)$ invariant thus forbidden: $\bar{\Psi}_L \psi_R \rightarrow \bar{\Psi}_L e^{-i\alpha_i \frac{\sigma^i}{2}} \psi_R \neq \bar{\Psi}_L \psi_R$.

To introduce mass to spin $\frac{1}{2}$ particles, Yukawa terms $\lambda_{\text{Yuk}} \bar{\Psi}_L \Phi \psi_R$, which can be readily checked¹² to be Lorentz, $U(1)$, and $SU(2)$ invariant, are added to the Lagrangian. Through the familiar Higgs mechanism, which has a non-zero VEV for Φ , fermion mass terms such as the electron mass term $\frac{\lambda_e v}{\sqrt{2}} (\psi_{L,e}^\dagger \psi_{R,e} + \psi_{R,e}^\dagger \psi_{L,e})$ are generated.

The full SM, which includes the strong interactions on top of the electroweak interactions, can be constructed in the same way outlined above with an additional $SU(3)$ gauge group such that the full symmetry group is $SU(3) \times SU(2)_L \times U(1)$. Its construction will not be discussed in this thesis.

3.1.4 Lepton flavor universality in the SM

With the electroweak theory developed in Section 3.1.3, and accepting the experimental observation that leptons do not interact strongly (i.e. they are $SU(3)$ singlets), it is now possible to discuss lepton flavor universality (LFU): the SM permits the addition of arbitrary generations of spin $\frac{1}{2}$ $SU(2)$ lepton doublets and singlets, and each generation couples to all electroweak gauge fields, namely $W_\mu^\pm$, Z_μ , and A_μ , with the same strengths through the same covariant derivative term $\not{D}_\mu$. The LFU is manifested by the very construction of the SM. The SM does contain Yukawa terms in which the different flavors of leptons couple to the Higgs doublet Φ with different strengths, leading to different masses of the leptons.

The three flavor generations of leptons are purely determined by experiments, as the SM places no constraint on the number of flavor generations.

¹² $U(1)$: $\bar{\Psi}_L \Phi \psi_R \rightarrow (\bar{\Psi}_L e^{-i\alpha}) \Phi (e^{i\alpha} \psi_R) = \bar{\Psi}_L \Phi \psi_R$; $SU(2)$: $\bar{\Psi}_L \Phi \psi_R \rightarrow (\bar{\Psi}_L e^{-i\alpha_i \frac{\sigma^i}{2}}) (e^{i\alpha_i \frac{\sigma^i}{2}} \Phi) \psi_R = \bar{\Psi}_L \Phi \psi_R$.

3.2 Semileptonic $B \rightarrow Dl^- \bar{\nu}_l$ decays in the SM

Semileptonic $B \rightarrow Dl^- \bar{\nu}_l$ decays, where B stands for a $\bar{B}^0$ or B^- meson, D for a generic charm meson, and l^- for a e^- , μ^- or τ^- , play an central role in this analysis as the signal, normalization, and most of the background decays fall under this category. A theoretical model describing these processes enables the calculation of the decay matrix elements. This, in turn, facilitates generation of the Monte-Carlo simulation samples that are critical in several stages of the analysis. This section describes the formalism of such a modeling.

The parton level process involved in the $B \rightarrow Dl^- \bar{\nu}_l$ decays is $b \rightarrow cl^- \bar{\nu}_l$, as already shown in Fig. 2.1, which is mediated by the weak charged current in the SM. Its matrix element is given by [8]:

$$\mathcal{M}_{\lambda_D}^{\lambda_l}(q^2, \theta_l) = \frac{G_F}{\sqrt{2}} \frac{m_W^2}{m_W^2 - q^2} \sum_{\lambda_W} \eta_{\lambda_W} L_{\lambda_W}^{\lambda_l}(q^2, \theta_l) H_{\lambda_W}^{\lambda_D}(q^2) \quad (3.12)$$

in which the total amplitude is obtained from summing over the helicity states $(\pm, 0, s)$ of the off-shell virtual W , polarized in the B rest frame as:

$$\epsilon_\mu(\pm) = \frac{1}{\sqrt{2}}(0, \pm 1, -i, 0) \quad \epsilon_\mu(0) = \frac{1}{\sqrt{q^2}}(|\mathbf{p}_D^*|, 0, 0, -q^0) \quad \epsilon_\mu(s) = \frac{q^\mu}{\sqrt{q^2}} \quad (3.13)$$

where q^0 is the energy of the virtual W boson in the B rest frame, and $q^2 = (q_0, -\mathbf{p}_D^*)^2$ is the invariant mass of the W . The metric factor η_{λ_W} is given by:

$$\eta_{\pm, 0, s} = \{1, 1, \frac{q^2 - m_W^2}{m_W^2}\} \approx \{1, 1, -1\} \quad (3.14)$$

where s denotes the scalar state of the virtual W . As shown in Eq. (3.12), the leptonic and hadronic currents factorize because leptons are color singlets which do not interact strongly.

The currents are defined as:

$$L_{\lambda_W}^{\lambda_l}(q^2, \theta_l) = \epsilon_\mu(\lambda_W) \langle l(p_l, \lambda_l) \bar{\nu}_l(p_{\bar{\nu}_l}) | \bar{l} \gamma^\mu (1 - \gamma_5) \nu_l | 0 \rangle \quad (3.15)$$

$$H_{\lambda_W}^{\lambda_D}(q^2) = \epsilon_\mu^*(\lambda_W) \langle D(p_D, \lambda_D) | \bar{c} \gamma^\mu (1 - \gamma_5) b | B(p_B) \rangle \quad (3.16)$$

where p_i, λ_i refer to the four-momenta and helicities of the particle i ; θ_l refers to the angle between the lepton and the W flight direction in the $l^- \bar{\nu}_l$ rest frame, as shown in Fig. 3.1.

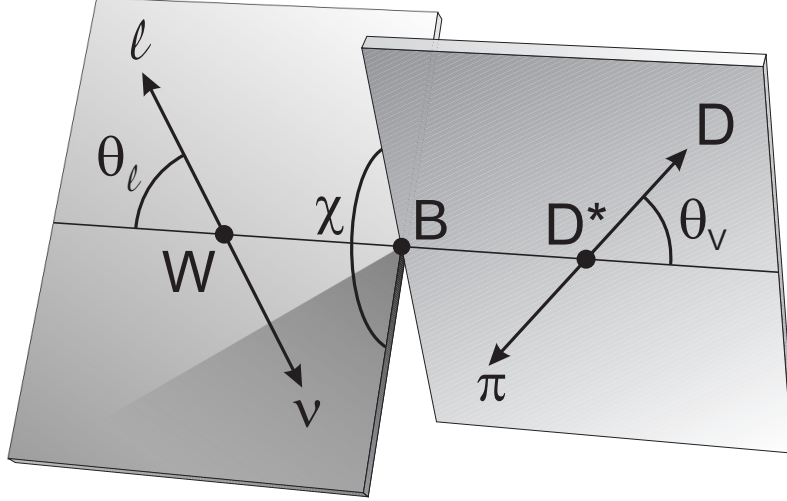


Figure 3.1: Angular variables of the $\bar{B}^0 \rightarrow D^{*+} l^- \bar{\nu}_l$ decays. For $B^- \rightarrow D^0 l^- \bar{\nu}_l$ decays, only the θ_l variable exists.

Leptonic currents The leptonic currents are straightforward to evaluate, as the weak interaction permits a perturbative calculation and the currents can be sufficiently approximated by tree-level decays. Evaluated in the B rest frame, each lepton current of the polarization combination (λ_W, λ_l) is a function of the lepton mass, the invariant mass squared q^2 of the virtual W , and θ_l as defined in Fig. 3.1. The results are listed in eqs. (2.8) and (2.9) in [9], and are copied over with minor changes on notations:

$$L_{\pm}^-(q^2, \theta_l) = -2\sqrt{q^2} v d_{\pm} \quad L_0^-(q^2, \theta_l) = -2\sqrt{q^2} v d_0 \quad L_s^-(q^2, \theta_l) = 0 \quad (3.17)$$

$$L_{\pm}^+(q^2, \theta_l) = \mp\sqrt{2} m_l v d_0 \quad L_0^+(q^2, \theta_l) = \sqrt{2} m_l v (d_+ - d_-) \quad L_s^+(q^2, \theta_l) = -2m_l v \quad (3.18)$$

where $v = \sqrt{1 - \frac{m_l^2}{q^2}}$, $d_{\pm} = \frac{1 + \cos \theta_l}{\sqrt{2}}$, and $d_0 = \sin \theta_l$.

Hadronic currents The calculation of the hadronic currents, however, involves non-perturbative quantum chromodynamics (QCD) effects as both the initial and final states are mesons which are quark-anti-quark pairs bounded by the strong interactions of QCD.

Hence, these currents are not computable in an analytical way and each of them can be re-expressed in terms of form factors (FFs), which are a function of q^2 , instead. These FFs are parameterized and estimated with relatively large theoretical uncertainties with four types of methods: use of functional properties of the matrix elements¹³, heavy quark effective theory (HQET), various quark models¹⁴, and lattice QCD [10]. As mentioned in the introduction, for $\mathcal{R}(D^{(*)})$ ratio measurements, the hadronic uncertainties are mostly factored out and the main importance of the FFs lies in the precision of the lepton universality ratio predictions and kinematic decay distributions. The estimations of FFs and the differential decay rate (which will be provided shortly after) are often performed at a particular q^2 value, such as when the D meson is produced at rest in the B rest frame¹⁵, and are extrapolated to the full q^2 range by performing fits to experimental inputs.

While not analytically calculable, the hadronic currents are constrained by the conservation of angular momentum, which implies that the D and the virtual W should have the same helicity. For $B \rightarrow D^0$ decays, whose only possible helicity state is 0, the only non-zero hadronic amplitudes are H_0^0, H_s^0 . For $B \rightarrow D^*$ decays, the amplitudes H_+^+ and H_-^- are allowed as well.

Differential decay rate The differential decay rate is found by first inserting Eqs. (3.17) and (3.18) into Eq. (3.12) then integrating over the phase space and summing over the lepton helicities:

$$\begin{aligned} \frac{d\Gamma}{dq^2 d\cos\theta_l} = & \frac{G_F^2 |V_{cb}|^2 |\mathbf{p}_D^*|^2 q^2}{256\pi^3 m_B^2} \left(1 - \frac{m_l^2}{q^2}\right)^2 \times \\ & \left[(1 - \cos\theta_l)^2 |H_+|^2 + (1 + \cos\theta_l)^2 |H_-|^2 + 2\sin^2\theta_l |H_0|^2 + \right. \\ & \left. \frac{m_l^2}{q^2} ((\sin^2\theta_l (|H_+|^2 + |H_-|^2) + 2|H_s + H_0 \cos\theta_l|^2)) \right] \end{aligned} \quad (3.19)$$

where the q^2 dependence and the helicities of the D meson in the hadronic currents $H_{\{\pm,0,s\}}$

¹³Such as analyticity, unitarity, and dispersion relations.

¹⁴For example QCD sum rule (QCDSR) and light cone sum rule (LCSR).

¹⁵Referred to as “zero recoil” due to the fact that the D meson is *not recoiling* in the B rest frame.

are omitted.

$\mathcal{R}(D^{(*)})$ calculation The parameterization of the hadronic currents H_i requires form factors, which will be described later. Once the H_i are resolved, plugging these into the differential decay rate gives an explicit expression for $\frac{d\Gamma}{dq^2}$ which can be used to compute $\mathcal{R}(D^{(*)})$:

$$\mathcal{R}(D^{(*)}) = \frac{\int_{m_\tau^2}^{(m_B - m_{D^{(*)}})^2} dq^2 \frac{d\Gamma}{dq^2}}{\int_{m_\mu^2}^{(m_B - m_{D^{(*)}})^2} dq^2 \frac{d\Gamma}{dq^2}} \quad (3.20)$$

In the next sections the form factor parameterizations relevant to this analysis for D^0 , D^* , and D^{**} mesons are discussed.

3.2.1 Form factors in $B \rightarrow D^0 l^- \bar{\nu}_l$ decays

As discussed before, a form factor parameterization of the hadronic currents is needed in order to evaluate the differential decay rate. In this analysis, simulated samples for $B \rightarrow D^0 l^- \bar{\nu}_l$ and $B \rightarrow D^* l^- \bar{\nu}_l$ are generated with the CLN parameterization and later reweighted offline to BGL (more details regarding the reweighting process will be provided in Section 7.1). The CLN and BGL parameterizations will be discussed next.

CLN For $B \rightarrow D^0 l^- \bar{\nu}_l$ decays the axial vector part of the hadronic current is vanishing due to conservation of angular momentum and parity [10]. Each hadronic current is therefore characterized by *two* FFs:

$$\begin{aligned} \langle D | \bar{c} \gamma^\mu (1 - \gamma^5) b | B \rangle &= \langle D | \bar{c} \gamma^\mu b | B \rangle \\ &= \sqrt{m_B m_D} [h_+(w)(v_B + v_D)^\mu + h_-(w)(v_B - v_D)^\mu] \end{aligned} \quad (3.21)$$

where $v_i = \frac{p_i}{m_i}$, $w(q^2) \equiv v_B \cdot v_D = \frac{m_B^2 + m_D^2 - q^2}{2m_B m_D}$. In the CLN parameterization [11], the FFs h_+ , h_- are re-expressed as $V_1(w)$, $S_1(w)$:

$$V_1(w) = h_+(w) - \frac{m_B - m_D}{m_B + m_D} h_-(w) \quad (3.22)$$

$$S_1(w) = h_+(w) - \frac{m_B + m_D}{m_B - m_D} \cdot \frac{w - 1}{w + 1} h_-(w) \quad (3.23)$$

and an expression for $V_1(w)$ is obtained by first analytically continuing w outside the physical region in the complex plane, then a conformal mapping

$$w \rightarrow z : z = \frac{\sqrt{w+1} - \sqrt{2}}{\sqrt{w+1} + \sqrt{2}} \quad (3.24)$$

is performed such that the physical region is mapped inside the unit circle of the complex z -plane. With dispersion techniques based on first principles and additional consideration from heavy quark symmetries to further reduce number of degrees of freedom, $V_1(z)$ is expressed in powers of z , known as “ z -expansion”, in the following form:

$$V_1(w) = V_1(1) \cdot [1 - 8\rho^2 z(w) + (51\rho^2 - 10)z(w)^2 - (252\rho^2 - 84)z(w)^3] \quad (3.25)$$

where $V_1(0), \rho^2$ are FF parameters, noting that $V_1(0)$ corresponds to the hadronic current at zero recoil. The factor $V_1(0)$, either fitted from external data with HQET constraints or computed with lattice QCD, represents an overall normalization which is *not cancelled* when performing CLN to BGL reweighting but *is cancelled* in the subsequent $\mathcal{R}(D)$ ratio.

The leptonic current corresponding to S_1 is helicity suppressed and, as a result, it is proportional to the lepton mass. Therefore this hadronic amplitude is insensitive to input data coming from semileptonic $B \rightarrow D^0 \ell^- \bar{\nu}_\ell$ decays to light leptons $\ell^- = e^-, \mu^-$ and is constrained based on HQET:

$$S_1(w) = \{1 + \Delta(-0.019 + 0.041(w - 1) - 0.015(w - 1)^2)\} V_1(w) \quad (3.26)$$

where Δ is a parameter with its value derived from HQET constraints.

The remaining parameter ρ^2 is obtained from a fit to experimentally measured q^2 distributions by saturating the dispersive bounds at 1σ uncertainty [10]. The values for the parameters listed above are copied from the corresponding MC generation code, and are listed in Table 3.1.

Table 3.1: $B \rightarrow D^0 l^- \bar{\nu}_l$ CLN FF parameterization.

FF parameter	Value	Hammer name
$V_1(0)$	1.035	G1
ρ^2	1.131	RhoSq
Δ	0.38	Delta

BGL Similar to CLN, the BGL parameterization [12, 13] is based on dispersion relations, analyticity, and crossing symmetry but with reduced constraints, providing more flexibility at the cost of a large number of parameters. The BGL parameterization is more suited to a data-driven approach of the determination of the FFs.

BGL adopts the following parameterization for $B \rightarrow D^0 l^- \bar{\nu}_l$ decays:

$$\langle D | \bar{c} \gamma^\mu b | B \rangle = f_+ (p_B + p_D)^\mu + (f_0 - f_+) \frac{q^\mu}{q^2} (m_B^2 - m_D^2) \quad (3.27)$$

Both f_+ , f_0 are expressed as a series of z :

$$f_{+,0}(z(w)) = \frac{1}{P_{+,0}(z(w))\phi_{+,0}(z(w))} \sum_{n=0}^{\infty} a_{+,0}^n z(w)^n \quad (3.28)$$

where $P_{+,0}(w)$ are known as *Blaschke factors* with their expressions listed in eq. (2.12) in [14], and $\phi_{+,0}$, *outer functions*, are defined in eqs. (2.13) and (2.14) of the same reference. The $a_{+,0}$ parameters are taken from Table 4 in [14] at $N = 2$, which are obtained from a fit to a combination of an unquenched lattice calculation at *non-zero recoil* and experimental data. The $N = 2$ fit result is chosen because it is the only case that the best fitted values lie within the unitary bounds and has a non-trivial correlation matrix as reported in Table 5 of [14]. The correspondence between the parameter notations used in this paper, in Ref. [14], and in **Hammer** is displayed in Table 3.2.

3.2.2 Form factors in $B \rightarrow D^* l^- \bar{\nu}_l$ decays

CLN For $B \rightarrow D^* l^- \bar{\nu}_l$ decays, the hadronic currents are parameterized by *four* FFs, V and $A_{0,1,2}$:

$$\langle D^* | \bar{c} \gamma^\mu b | B \rangle = 2iV(m_B + m_{D^*}) \varepsilon^{\mu\nu\alpha\beta} \epsilon_\nu^* (p_{D^*})_\alpha (p_B)_\beta \quad (3.29)$$

Table 3.2: $B \rightarrow D^0 l^- \bar{\nu}_l$ BGL FF parameterization notation correspondence.

FF parameter	Ref. [14]	Hammer name
a_+^n	a_n	ap , a vector
a_0^n	b_n^*	a0 , a vector

* The b_0 coefficient is fixed by the other coefficients: $b_0 = 4.99a_0 + 0.32a_1 + 0.021a_2 - 0.065b_q - 0.004b_2$ due to the constraint that $f_+ = f_0$ at $q^2 = 0$.

$$\begin{aligned}
\langle D^* | \bar{c} \gamma^\mu \gamma^5 b | B \rangle &= A_1(m_B + m_{D^*}) \epsilon^{*\mu} - \\
&= A_2 \frac{\epsilon^* \cdot \mathbf{p}_B (p_B + p_{D^*})^\mu}{m_B + m_{D^*}} + \\
&= 2m_{D^*} q^\mu (A_0 - A_3) \frac{\epsilon^* \cdot \mathbf{p}_B}{q^2}
\end{aligned} \tag{3.30}$$

in which $2m_{D^*} A_3 = A_1(m_B + m_{D^*}) - A_2(m_B - m_{D^*})$. In the CLN parameterization, these FFs are re-expressed as:

$$A_1(w) = \frac{w+1}{2} r_{D^*} h_{A_1}(w) \qquad A_0(w) = \frac{R_0(w)}{r_{D^*}} h_{A_1}(w) \tag{3.31}$$

$$A_2(w) = \frac{R_2(w)}{r_{D^*}} h_{A_1}(w) \qquad V(w) = \frac{R_1(w)}{r_{D^*}} h_{A_1}(w) \tag{3.32}$$

where $r_{D^*} = 2 \frac{\sqrt{m_B m_{D^*}}}{m_B + m_{D^*}}$. With dispersive bounds, these are expressed as:

$$h_{A_1}(w) = h_{A_1}(1) [1 - 8\rho^2 z(w) + (53\rho^2 - 15)z(w)^2 - (231\rho^2 - 91)z(w)^3] \tag{3.33}$$

$$R_1(w) = R_1(1) - 0.12(w-1) + 0.05(w-1)^2 \tag{3.34}$$

$$R_2(w) = R_2(1) + 0.11(w-1) - 0.06(w-1)^2 \tag{3.35}$$

$$R_0(w) = R_0(1) - 0.11(w-1) + 0.01(w-1)^2 \tag{3.36}$$

The parameters used in $B \rightarrow D^* l^- \bar{\nu}_l$ MC generation are listed in Table 3.3.

In the CLN parameterization, the slope and curvature, that is, the coefficients of the z and z^2 terms in Eq. (3.33), of the Isgur-Wise functions¹⁶ share the same parameter ρ , making the two too closely cross-constrained by the model. Additionally, the fit of the normalization factors of the form factor ratios at zero recoil $R_i(w=1)$ does not take the

¹⁶The form factor description in the heavy quark limit.

Table 3.3: $B \rightarrow D^* l^- \bar{\nu}_l$ CLN FF parameterization.

FF parameter	Value	Hammer name
$h_{A_1}(1)$	0.908	F1
ρ^2	1.122	RhoSq
$R_1(1)$	1.270	R1
$R_2(1)$	0.852	R2
$R_0(1)$	1.15	R0

fitted slopes $R'_i(w=1)$ into account, but uses the fixed values, such as -0.12 in Eq. (3.34), obtained from heavy quark expansion considering the contributions from $1/m_Q$ instead. Such an approach is not self-consistent [15]. Therefore, it is preferable to use the BGL parameterization.

BGL The BGL parameterization scheme for $B \rightarrow D^* l^- \bar{\nu}_l$ decays is [16]:

$$\langle D^* | \bar{c} \gamma^\mu b | B \rangle = i \sqrt{m_B m_{D^*}} h_V \varepsilon^{\mu\nu\alpha\beta} \epsilon_\nu^* (v_{D^*})_\alpha (v_B)_\beta \quad (3.37)$$

$$\begin{aligned} \langle D^* | \bar{c} \gamma^\mu \gamma^5 b | B \rangle = \sqrt{m_B m_{D^*}} \Big[& h_{A_1}(w+1) \epsilon^{*\mu} - \\ & h_{A_2}(\epsilon^* \cdot \mathbf{v}_B) (v_B)^\mu - h_{A_3}(\epsilon^* \cdot \mathbf{v}_B) (v_{D^*})^\mu \Big] \end{aligned} \quad (3.38)$$

with the FFs $\{h_V, h_{A_1, A_2, A_3}\}$ re-expressed as $\{g, f, \mathcal{F}_1, \mathcal{F}_2\}$:

$$g(w) = \frac{h_V(w)}{m_B \sqrt{r}} \quad (3.39)$$

$$f(w) = m_B \sqrt{r} (1+w) h_{A_1}(w) \quad (3.40)$$

$$\mathcal{F}_1 = m_B^2 \sqrt{r} (1+w) [(w-r) h_{A_1}(w) - (w-1)(r h_{A_2}(w) + h_{A_3}(w))] \quad (3.41)$$

$$\mathcal{F}_2 = \frac{1}{\sqrt{r}} [(1+w) h_{A_1}(w) + (rw-1) h_{A_2}(w) + (r-w) h_{A_3}(w)] \quad (3.42)$$

where $r = \frac{m_{D^*}}{m_B}$. These FFs are proportional to $H_- - H_+, H_+ + H_-, H_0, H_s$ accordingly.

With z -expansion, the FFs are expanded into the form:

$$g(w) = \frac{1}{P_{1-}(z(w)) \phi_g(z(w))} \sum_{n=0}^N a_n z(w)^n \quad (3.43)$$

$$f(w) = \frac{1}{P_{1+}(z(w))\phi_f(z(w))} \sum_{n=0}^N b_n z(w)^n \quad (3.44)$$

$$\mathcal{F}_1(w) = \frac{1}{P_{1+}(z(w))\phi_{\mathcal{F}_1}(z(w))} \sum_{n=0}^N c_n z(w)^n \quad (3.45)$$

$$\mathcal{F}_2(w) = \frac{1}{P_{0-}(z(w))\phi_{\mathcal{F}_2}(z(w))} \sum_{n=0}^N d_n z(w)^n \quad (3.46)$$

where $P_i(z(w))$ are given in eq. (61) in [16], and $\phi_i(z(w))$ in eqs. (67–70) in the same reference. The coefficients $\{a_n, b_n, c_n, d_n\}$ are given by a fit to unquenched lattice calculation at non-zero recoil and data; their values are listed in the right-most column of Table XII in the **v1** version of [16]. The parameter naming correspondence is displayed in Table 3.4.

Table 3.4: $B \rightarrow D^* l^- \bar{\nu}_l$ BGL FF parameterization notation correspondence.

FF parameter	Ref. [16]	Hammer name
a_n	a_n	avec , a vector
b_n	b_n	bvec , a vector
c_n	c_n	cvec , a vector*
d_n	d_n	dvec , a vector†

* The c_n coefficients starts from c_1 instead of c_0 .

† The c_3 and d_2 coefficients are unused in the reweighting based on a claim that they are not implemented in **Hammer**. Latest version of **Hammer** may have added support to these parameters.

3.2.3 Form factors in $B \rightarrow D^{**} l^- \bar{\nu}_l$ decays

ISGW2 The ISGW2 parameterization [17] is based on a non-relativistic constituent quark model [10] and is treated as fully predictive, that is, without any undetermined parameter. While historically important in providing a description of the FFs, it is no longer considered to provide state-of-art FFs due to its inability to describe the experimental data.

BLR The BLR parameterization [18] provides a FF parameterization for the four lightest ($1P$) excited D^{**} mesons (D_0^* , D_1' , D_1 , D_2^*), grouped into two heavy quark isospin doublets

$D^{1/2+} : (D_0^*, D_1')$ and $D^{3/2+} : (D_1, D_2^*)$. This parameterization utilizes a HQET expansion, considers all possible $b \rightarrow cl^- \bar{\nu}_l$ four-Fermi interactions, and includes $\mathcal{O}(\Lambda_{\text{QCD}}/m_{c,b})$ and $\mathcal{O}(\alpha_s)$ corrections which results in FFs that can be improved by data-driven fits. The parameterizations are listed separately in Tables 3.5 and 3.6 for the two isospin doublets.

Table 3.5: FF parameterization for $D^{1/2+}$ doublet (D_0^*, D_1') .

FF parameter	Ref. [18]	Hammer name
$\zeta(1)$	0.7	zt1
ζ'	0.2	ztp
$\hat{\zeta}_1$	0.6	zeta1

Table 3.6: FF parameterization for $D^{3/2+}$ doublet (D_1, D_2^*) .

FF parameter	Ref. [18]	Hammer name
$\tau(1)$	0.7	t1
τ'	-1.6	tp
$\hat{\tau}_1$	-0.5	tau1
$\hat{\tau}_2$	2.9	tau2

Chapter 4: The LHCb experiment

The LHCb experiment is one of the four large experiments (CMS, ATLAS, ALICE, LHCb) at the Large Hadron Collider (LHC), a superconducting circular proton-proton collider (mostly¹), built and managed by the European Organization for Nuclear Research (CERN), with a center of mass energy $\sqrt{s} = 13$ TeV during its run 2 (2016–2018) operation period.

Located 100 meters underground at Franco-Swiss border, the LHCb detector, shown in Fig. 4.1, is a forward only spectrometer with a focus on the b - and c -physics precision measurements. The same figure also defines the LHCb coordinate system: z axis is the beam axis; x -axis is the horizontal axis; y the vertical one. In terms of the spherical coordinate, the θ angle running from 0 to π is the angle respect to the $+z$ axis; the ϕ is the angle of rotation in the x - y plane which ranges from 0 to 2π . The particle production is highly non-uniform in terms of the polar angle θ of the particle 3-momentum, but is roughly uniform in pseudorapidity η , a quantity related to θ :

$$\eta \equiv -\ln \left[\tan \left(\frac{\theta}{2} \right) \right] \quad (4.1)$$

The LHCb detector covers the pseudorapidity ranging from $1.9 < \eta < 4.9$. The unusual geometry of the LHCb detector (as opposed to 4π detectors with full solid angle coverage, for example the other three large experiments at the LHC) is largely driven by the $b\bar{b}$ production mechanism at the LHC: the dominant production mode is gluon fusion in which case the momenta of the incoming partons are highly asymmetric in the laboratory frame. As a result, the $b\bar{b}$ center of mass is boosted either forward or backward in the beam direction

¹Heavy ions are also collided at the LHC, but are not relevant for this analysis.

[19], as confirmed by the simulation of $b\bar{b}$ production angular distribution, shown in Fig. 4.2. With just about 4% solid angle coverage, the LHCb detector efficiently and economically reconstructs about 20% of all produced $b\bar{b}$ pairs [20].

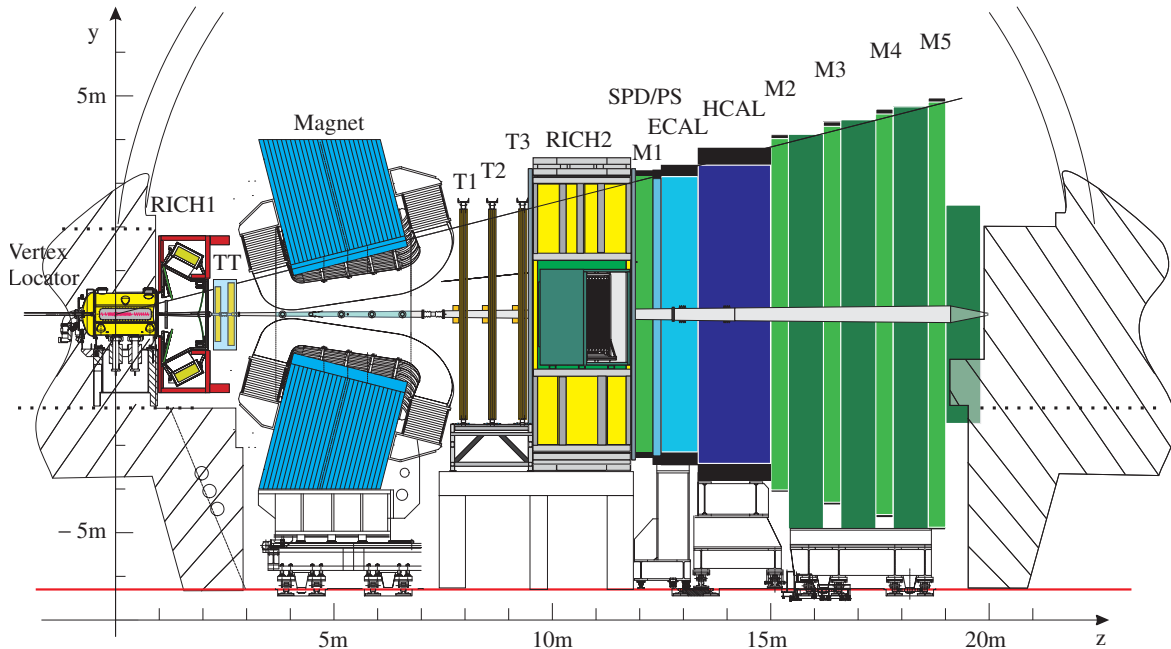


Figure 4.1: The LHCb detector. z axis indicates $b\bar{b}$ travelling direction. The impact point is at $z = 0$.

At the designed LHC luminosity, multiple pp collisions happen within the same bunch crossing (an effect called “pile-up”) which greatly complicates $b\bar{b}$ production and event reconstruction because the event² is significantly more *busy* (compared to events with single pp collision) [19]. To maintain the luminosity at optimal values for LHCb, a technique termed *luminosity levelling* is implemented to dynamically *defocus* beams by offsetting them in the transverse direction based on reported LHC instantaneous luminosity [22]. Over the course of run 1 and 2 operation periods, LHCb recorded a total integrated luminosity of $\sim 8.7 \text{ fb}^{-1}$ [23], corresponding to over 300 billion B mesons and several billions of other b hadrons, an unprecedented number for b -physics.

In general, the LHCb detector has the following features, all of which required by

²An event is defined as *everything* produced in a single bunch crossing.

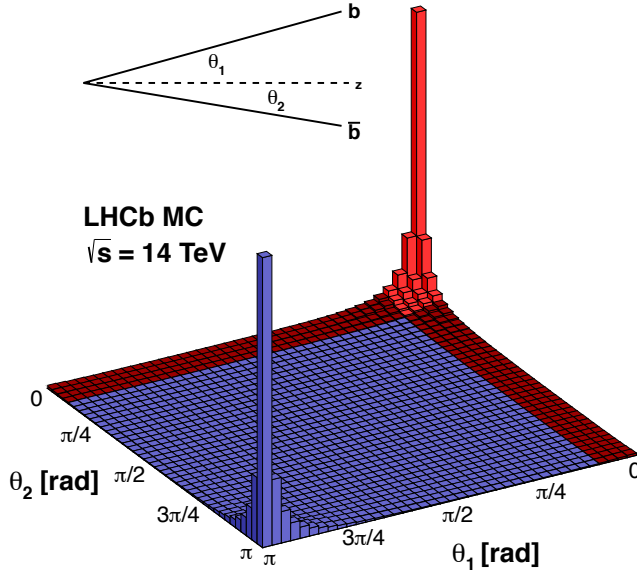


Figure 4.2: Simulated $b\bar{b}$ production angular distribution at $\sqrt{s} = 14$ GeV. Taken from [21].

b -physics measurements: a precise tracking system providing excellent track and vertex resolution; a good particle identification performance separating π , K , e , μ , and p (proton); a versatile trigger scheme efficient for both hadron and lepton with lower p_T thresholds³. The rest of chapter provides a brief introduction for the subdetectors of the LHCb detector relevant for this analysis, grouped by the features listed above. An overview of reconstruction of charged particles is provided in Section 4.1; particle identification is described in Section 4.2; the basics of trigger implementation is explained in Section 4.3.

4.1 Reconstruction of charged particles

Charged particle reconstruction relies on the tracking system, which consists of the following components (sorted based on the proximity from the pp interaction point):

- VErtext LOcator (VELO): it is a compact 21-layer silicon-strip detector placed immedi-

³Compared to CMS and ATLAS, but the lepton trigger p_T threshold is still too high for this analysis.

ately around the pp interaction point⁴ [24] that detects particles with a pseudorapidity ranging from $1.6 < \eta < 4.9$. Each layer of VELO, made of two retractable⁵ silicon half disks with a pitch⁶ ranging from 37 to 98 μm (increasing with radius), provides radial and azimuthal angular position measurements of a track. Combining positional information from all layers (with a fit), a tracking resolution of about 4 μm at best is achieved [25]. The primary vertex (PV) resolution depends on number of tracks (nTracks) used in vertex reconstruction, ranging from 35–10 μm in the transverse plane (low nTracks to high), and 260–60 μm in the beam axis.

- Tracker Turicensis (TT): located downstream of VELO and immediately before the Magnet, the TT consists of 4 planes of silicon-strip sensors, each of a size 150 cm $\times$ 130 cm, with a strip pitch of 183 μm arranged in a straight/stereo configuration⁷. It covers the full acceptance of the LHCb experiment, and has a spatial resolution of about 50 μm [24].
- T-stations: they are located immediately downstream of the Magnet. There are a total of 3 T-stations with a common design. Each of them is made up of two subdetectors: The Inner Tracker (IT), covering a cross-shaped region around the beam pipe, and the Outer Tracker (OT), covering the rest of the LHCb acceptance. The IT is made of silicon strips (same technology as in TT) whereas the OT uses gas-tight straw-tubes. The IT has a spatial resolution identical to that of TT, namely 50 μm . For the OT, its resolution is about 205 μm [26].

Due to the design of the LHCb detector, tracks are approximately straight lines separately in the upstream and downstream (in terms of the Magnet) segments because of a

⁴The first few layers (stations) are located upstream the interaction point (backward direction), whereas the last layers are downstream (forward direction).

⁵This is required because the VELO sensors are placed 8 mm from the LHC beam. The gap is smaller than the aperture required by the beam during injection, so the semi-disks are retracted in such cases.

⁶“Pitch” refers to the size of a single silicon strip, akin to a “pixel” in the LCD screen. A smaller pitch generally implies a better resolution.

⁷This will be explained in Section 12.1.

lack of magnet field. The downstream and upstream segments, labelled as VELO tracks and T-tracks, are first reconstructed separately before attempting to match them to form long tracks. If compatible hits are found in the TT during the matching, they are added to the track to further improve the momentum resolution and reduce the amount of fake tracks (ghosts). Different track types are shown in Fig. 4.3. In most analyses, this one included, only long tracks are used because they provide the best momentum resolution [27].

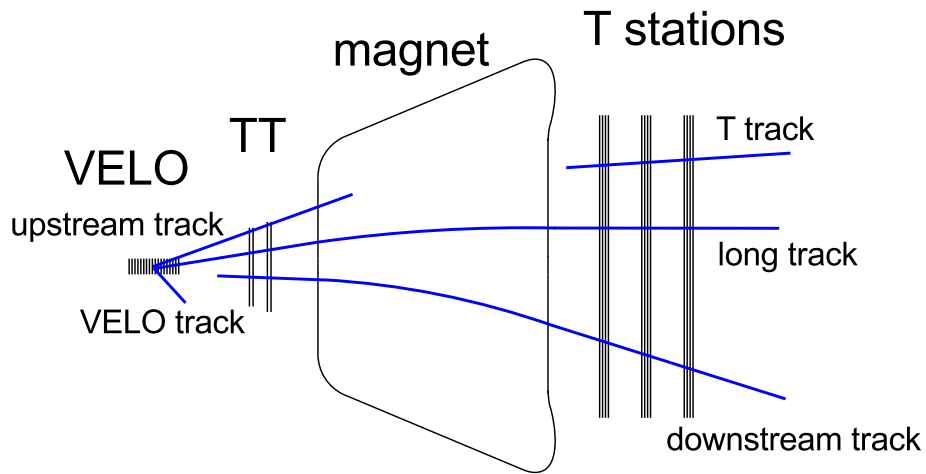


Figure 4.3: Reconstructed track types at LHCb.

4.2 Particle identification

The particle identification (PID) system consists of three subsystems:

- Two Ring Imaging CHerenkov (RICH) detectors: RICH1 and RICH2. The former is situated between the VELO and the TT, the latter after the T-stations. The RICH detectors provide identification for charged hadrons, with limited discrimination power for e and μ . These two detectors measure the Cherenkov angle of emitted photons from charged particles traversing through the medium, which depends on its refractive index n , the momentum p of the particle, and the mass m of the particle:

$$\theta_c = \frac{\sqrt{1 + (m/p)^2}}{n} \quad (4.2)$$

Particles with different masses result in different Cherenkov angles. At sufficiently large momentum, Cherenkov angles saturate and can no longer tell different particles apart. To ensure good PID performance across a wider range of momentum, materials with different refractive indices are used in RICH1 and RICH2.

RICH1 uses a gas layer of C_4F_{10} with $n = 1.0014$. This material provides hadron identification from about 1 to 60 GeV, and covers the full LHCb acceptance. RICH2 has a smaller angular acceptance corresponding to a region enriched of high-momentum particles. It uses a CF_{10} gas radiator with $n = 1.00046$, providing PID up to 150 GeV [20]. In Fig. 4.4, Cherenkov angles of different particles are shown; it also displays a real measurement of Cherenkov angles from RICH1.

- A calorimeter system composed of 4 subdetectors: the Scintillator Pad Detector (SPD), the Preshower Detector (PS), the Electromagnetic CALorimeter (ECAL), and the Hadron CALorimeter (HCAL). Located downstream of the T-stations, the calorimeter system measures energy deposit from all particles. Combining information from the tracking system, it can determine the presence of e, γ and neutral hadrons: e leaves a trajectory in the tracking system, while depositing most of its energy in ECAL; γ is invisible to tracking but deposits most energy in ECAL; while also invisible to the tracking system, neutral hadrons deposit energy in both ECAL and HCAL [20]. Note that PID in calorimeters is a destructive measurement: All particles except muons and neutrinos deposit all their energy in the calorimeters by producing electromagnetic or hadronic showers [28].

All calorimeters are made of segments of scintillators with different granularity, with ECAL and HCAL having multiple layers of alternating scintillators and absorbers. The calorimeters detect light from scintillators, emitted by particles travelling through, using photomultipliers as a way to measure energy deposit.

The SPD and PS are two layers of scintillators separated by a 12 mm thick lead sheet.

They are designed to indicate the presence of charged particles to the L0 calorimeter trigger.

The ECAL employs the *Shashlik* technology⁸ with 66 layers of independent modules constructed from 2 mm of lead plates followed by 4 mm of scintillating material with Wave-Length Shifting (WLS) optical fibers piercing through the layers and guiding scintillating light to photomultipliers. The energy resolution for the ECAL is:

$$\frac{\sigma(E)}{E} = \frac{10\%}{\sqrt{E}} \oplus 1\% \quad (4.3)$$

where E is the electron energy in GeV.

The HCAL consists of layers of iron absorber with scintillating tiles as the active material. Compared to ECAL, HCAL has a coarser granularity with a reduced energy resolution:

$$\frac{\sigma(E)}{E} = \frac{(69 \pm 5)\%}{\sqrt{E}} \oplus (9 \pm 2)\% \quad (4.4)$$

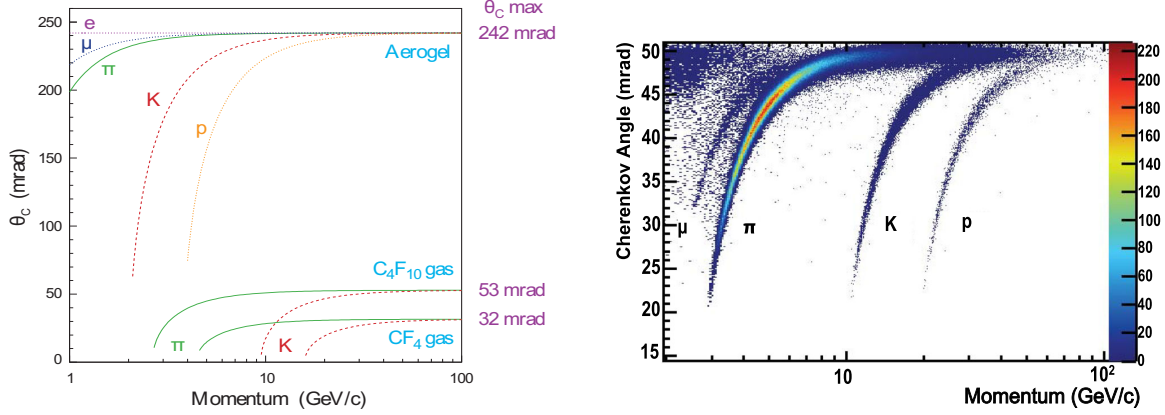
where E is the hadron energy in GeV.

- Five muon stations: labelled as M1–M5, M1 is located in front of the calorimeter system whereas M2–M5 come after. Each muon station is of rectangular shape and made up by 1368 multi-wire proportional chambers supplemented by 12 triple GEM chambers in the inner region of the first station [20]. The muon stations have an acceptance in the bending plane from 20 mrad to 306 mrad and from 16 mrad to 258 mrad in the non-bending plane, resulting in a total acceptance of about 20% for muons from semileptonic inclusive b decays.

The muon stations provide space point measurements as well as timing measurements of the remaining tracks, the vast majority of them being muons as most other particles, charged or uncharged, are absorbed by the detector before reaching the stations. By

⁸The name comes from the fact that detectors constructed this way have similar appearances to a shashlik.

requiring muon-like particles to produce hits in certain muon stations in a momentum-dependent way⁹, a binary muon PID decision, `isMuon`, is produced.



(a) Cherenkov angles for different particles traversing different media. “Aerogel” is irrelevant for run 2.

(b) Cherenkov angles measured by RICH1.

Figure 4.4: PID in RICH detectors.

As a summary, the PID system exploits the fact that different types of particles leave different signatures in the detectors, schematically shown in Fig. 4.5, to identify charged hadrons and leptons, as well as neutral hadrons. Overall, the LHCb experiment provides two types of PID variables [30]:

- DLL_X : These variables are the log likelihood difference between X and π , with the log likelihoods combined from all three PID subsystems (RICH, calorimeters, muon stations): $\mathcal{L} = \mathcal{L}_{\text{RICH}} \cdot \mathcal{L}_{\text{Calo}} \cdot \mathcal{L}_{\text{Muon}}$.
- ProbNN_X : These variables are outputs from neural networks (NN) trained to identify X . The training inputs are from both the PID and the tracking system, with data-driven corrections from calibration samples.

⁹The actual requirement can be looked up from Table 1 of [29].

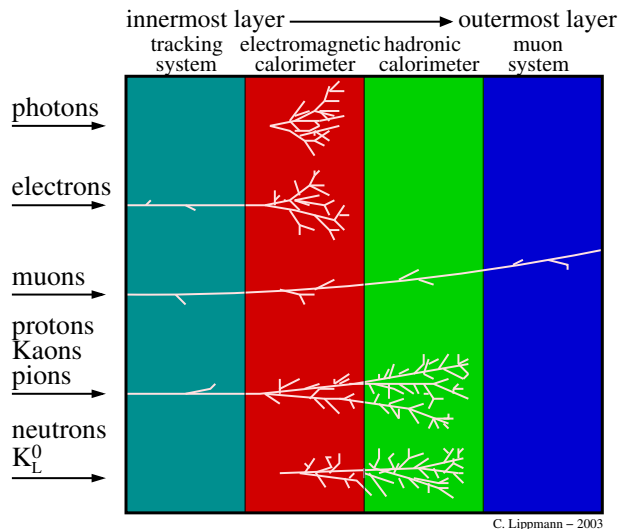


Figure 4.5: Signatures of different particles in a typical configuration of a modern detector. Taken from [28].

4.3 Trigger

Despite the fact that the LHCb experiment operates at reduced luminosity, it is still impractical to record the total output of all events due to readout rate and storage limitations. The readout rate is first reduced from 40 MHz, the LHC bunch crossing frequency, to 1 MHz by hardware triggers, referred to as Level-0 (L0) triggers. The L0 triggers reconstruct transverse energy E_T and momentum p_T in real time and select events with large E_T or p_T , a signature of B meson decays due to their large masses. In addition, the number of pp interactions in the event and the number of tracks are estimated and events with large pile-up or multiplicity are rejected. The reconstruction of these high-multiplicity events would require a disproportionately large fraction of readout bandwidth and typically results in poorer tracking quality. The selected events are further reconstructed by software high-level triggers (HLT1 and HLT2) and only events with high- p_T tracks and/or large impact parameters are further filtered with pre-defined cuts before being retained on disk for offline analysis. The HLT triggers reduce the output rate to about 12.5 kHz [31].

One of the main objective of the ongoing detector upgrade is to increase the bandwidth

of the detector readout thus removing the need of hardware triggers, and all the limitations that come with them. This section is dedicated to discussion of the hardware triggers (L0) which, while no longer are a part of the LHCb run 3 program, are of crucial importance to this analysis. The high-level triggers will be discussed and compared between run 2 and run 3 with more details in Section 11.1.

The L0 trigger is divided into three components, with each connected to a detector and its information collected by the Level-0 Decision Unit to evaluate a final trigger decision, as shown in Fig. 4.6:

- Pile-up system: located upstream of VELO, the pile-up system consists of two sets of silicon disks of the same type used in VELO. It is the only detector that can estimate the track multiplicity in the backward direction. The intended use of the system was to veto events with multiple pp interactions but is unused because efficient high level triggers made pile-up less harmful [32]. As a reference, for run 1, the average pile-up is about 1.7; for run 2, it is about 1.1 [33].
- Calorimeter trigger system: it identifies the cluster (2×2 cells of scintillators) with the largest E_T and provides a PID decision (hadron, electron, or photon) of the cluster based on the calorimeter system only. The required E_T/p_T threshold is set differently for each particle type. This analysis uses L0 hadron trigger which has a E_T threshold of 3.7 GeV for year 2016 [34].

In addition, the total number of SPD cells with a hit is counted to estimated number of tracks (charged particles) in the event. Typically events are required to have fewer than 450 SPD hits.

- Muon trigger system: The muon stations can provide an independent measurement of the track p_T with a resolution of about 25% on average [34]. The L0 muon triggers require that either the largest muon p_T exceeds 1.8 GeV or the product of the two largest muon p_T exceeds 2.25 GeV^2 (both are 2016 thresholds).

As we will discuss in Section 5.1.1, this analysis *relies* on an uniform muon selection efficiency across p_T spectrum. As the muon p_T spectra of signal decays are softer, the muon trigger cannot be used because it not only actively rejects signal, but also makes the selection efficiencies between signal and normalization different, which reduces the cancellation of experimental efficiencies in the $\mathcal{R}(D^{(*)})$ ratios.

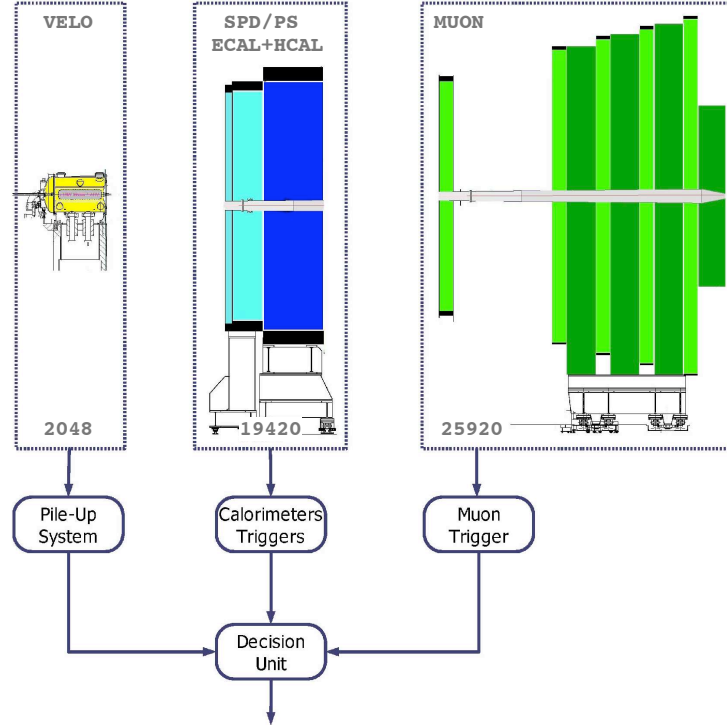


Figure 4.6: Schematic of the L0 trigger system.

Chapter 5: Event selection

This chapter describes the selection procedure for the samples used in this analysis which involves an online and offline selection. The initial online selections are a composition of pre-defined hardware and high-level software triggers provided by the LHCb experiment; some of the triggers bias important physical quantities thus cannot be used. The follow-up offline selection requirements are imposed for this analysis to further separate signal events from various backgrounds.

The events in collision data and MC samples used directly in the fit are selected with the procedure described in Sections 5.1 and 5.2, respectively. Each sample is further divided into signal-like and control background sub-samples, also known as “skims” with the additional requirements listed in Section 5.3. The selection of samples used in auxiliary studies is discussed in Section 5.4. Additional algorithms used in the selection process are described in Section 5.5.

5.1 Data samples

This analysis currently uses 1665 pb^{-1} of proton-proton collision data recorded by LHCb in 2016 at $\sqrt{s} = 13 \text{ TeV}$. It is planned to extend the analysis to employ data taken in the full 2016–2018 period (run 2). The rest of the section describes selection procedures for various data samples, which are:

- Right-sign samples: the nominal data samples to be fitted.

- Wrong-sign samples: for characterizing various combinatorial backgrounds which are random combinations of the $D^{(*)}\mu$ pairs that are *not* originating from the same B decay but pass the selections requiring so.
- Fake muon samples: for studying the contributions from other charged particles mis-identified as muon in both the nominal and the wrong-sign data samples.

5.1.1 Right-sign samples

These are the data samples used in the fit where the signal and normalization yields are extracted. The samples contain right-sign $D^{(*)}\mu$ pairs, the pairs that have the correct charges to form a legitimate B meson. That is, only $D^0\mu^-$ and $D^{*+}\mu^-$ (and their charge conjugates). The μ in these samples are required to pass μ particle identification (PID) requirements.

The online selection for the $D^0\mu^-$ and $D^{*+}\mu^-$ final states, henceforth referred as D^0 channel and D^* channel, are identical. The selection involves an online trigger path: a composition of hardware (L0) and software triggers (HLT1, HLT2), as listed in Table 5.1. Overall, the online selection expects a μ candidate to form a well-defined vertex with a high p_T $D^0(\rightarrow K^-\pi^+)$ candidate, requiring both particles to be close to each other and are not directly coming from the primary pp decay vertex (PV). More details about the trigger path are given below.

The L0 triggers are chosen to make sure events are not triggered by the μ track *alone* which would introduce a p_T bias¹ on the μ . Such a bias is very problematic for this analysis as μ from signal $\tau \rightarrow \mu\bar{\nu}_\mu\nu_\tau$ decays are typically softer and have smaller p_T , making signal events less likely to be selected. A muon trigger would, as discussed at the end of Chapter 4, actively reject signal decays as well as making the selection efficiencies for the signal and normalization decay modes different. The inclusive HLT1 triggers select events containing

¹According to table 1 in [34], the minimum p_T threshold across all L0 triggers is 1.35 GeV for 2018 L0Muon trigger.

Table 5.1: Trigger path for this analysis. For terminologies such as TIS and TOS, refer to Appendix D.

Trigger level	Requirements
L0	D^0 L0Hadron TOS B L0Global TIS
HLT1	$(K \text{ Hlt1TrackMVA TOS } \pi \text{ Hlt1TrackMVA TOS})^* $ $D^0 \text{ Hlt1TwoTrackMVA TOS}$
HLT2	$B \text{ Hlt2XcMuXForTauB2XcMu TOS}$

* This is almost equivalent to $D^0 \text{ Hlt1TrackMVA TOS}$, with a $\sim 0.0027\%$ difference in selected events in reconstructed data sample in D^0 channel. Henceforth these two trigger paths are considered equivalent.

a particle whose decay vertex is displaced from the PV, in this case a $D^0 \rightarrow K^- \pi^+$, as described in Ref. [35], Section 6.7.1. The `Hlt2XcMuXForTauB2XcMu` is specifically designed for this analysis to select $D^0 \mu^\pm$ candidates² with a high p_T requirement on D^0 , while maintain minimal p_T bias on the $\mu^\pm$. The trigger avoids cut on the invariant mass of the $D^0 \mu$ pair so that both the right-sign and wrong-sign samples can be selected with the same trigger path: the right-sign $D^0 \mu$ pairs are more likely to form a D^* meson, producing a peak around the mass of D^* , whereas the wrong-sign $D^0 \mu$ pairs are dominated by random combinations thus not peaking. The cuts imposed by the HLT2 line are listed in Table 5.2.

Additional offline selections with stricter vertex quality and PID requirements are applied to both D^0 and D^* channels. Furthermore, only the right-sign $D^{(*)} \mu$ candidates are kept. The offline selections that are shared among D^0 and D^* channels are listed in Tables 5.3 and 5.4. The additional D^0 only cuts are listed in Table 5.5, the D^* in Table 5.6.

Due to the fact that the D^0 and D^* channel have the same online selection, some events will appear in both, making the two channels correlated. This correlation is removed by a D^* veto procedure in the D^0 channel, and is discussed in Section 5.5.1.

Previous analyses ([1, 15]) showed that a tighter μ PID reduces other charged species misidentified (misID) as μ , reducing the systematic uncertainty due to misID effects. However, the standard LHCb μ PID has low efficiency at low p_T . Therefore, an offline μ

²Yes, both right-sign and wrong-signs of the $D^0 \mu$ pairs are selected.

Table 5.2: Cuts defined in H1t2XcMuXForTauB2XcMu.

Particle	Variable	Cuts
K, π	p_T	$> 200 \text{ MeV}$
	Max p_T	$> 800 \text{ MeV}$
	p	$> 5 \text{ GeV}$
	χ^2_{IP}	> 9
	DLLK	$> 2 (K), < 4 (\pi)$
D^0	p_T	$> 2000 \text{ MeV}$
	$ \pi p_T + K p_T $	$> 2.5 \text{ GeV}$
	m	1830–1910 MeV
	$\chi^2_{\text{vertex}}/\text{ndf}$	< 10
	χ^2_{FD}	> 25
	DIRA	> 0.999
	Child pair DOCA	$< 0.1 \text{ mm}$
μ	χ^2_{IP}	> 16
$D^0\mu$	$\chi^2_{\text{vertex}}/\text{ndf}$	< 15
	χ^2_{FD}	> 50
	DIRA	> 0.999
	DOCA	$< 0.5 \text{ mm}$

PID method, termed μ BDT, is developed to reduce misID while maintain a flat p_T response. A brief overview on this method is provided in Section 5.5.2.

5.1.2 Wrong-sign samples

The wrong sign samples are used to model various combinatorial backgrounds. For the D^0 channel, the $D^0\mu^+$ pairs are almost exclusively of combinatorial nature³ and is the only dominant combinatorial background; this is referred as B combinatorial. For the D^* channel, a D^{*+} may randomly combine with a *wrong-sign* μ^+ (B combinatorial); alternatively, a D^0 may combine with a *wrong-sign slow* π^- to form a D^{*-} (D^* combinatorial). The modeling of these backgrounds will be discussed in Section 8.1.2.

³The right-sign sample are reconstructed as the following decay: $B^- \rightarrow D^0(\rightarrow K^- \pi^+) \mu^-$, so $D^0\mu^+$ pairs are unphysical at tree level thus are most likely resulting from random combinations.

Table 5.3: Base offline selections (stripping) for both D^0 and D^* channels.

Event-Level	Variable	Cuts
GEC	nSPDhits	< 600
PV cut	nPV	≥ 1
Particle	Variable	Cuts
K, π	p_T	$> 300 \text{ MeV}$
	p	$> 2 \text{ GeV}$
	χ^2_{IP}	> 9
	isMuon	False
	DLLK	$> 4 (K), < 2 (\pi)$
	GhostProb	< 0.5
D^0	$ \pi p_T + K p_T $	$> 2.5 \text{ GeV}$
	$m - m_{\text{PDG}}$	$< 80 \text{ MeV}$
	$\chi^2_{\text{vertex}}/\text{nof}$	< 4
	χ^2_{FD}	> 25
	DIRA	> 0.999
μ	p	$> 3 \text{ GeV}$
	χ^2_{IP}	> 16
	isMuon	True
	DLL μ	> -200
	GhostProb	< 0.5
$D^0\mu$	m	$0\text{--}10 \text{ GeV}$
	$\chi^2_{\text{vertex}}/\text{ndf}$	< 6
	DIRA	> 0.999

The online selection for wrong-sign samples are the same as the right-sign, as described in Section 5.1.1. The offline selection is also mostly identical, with a few differences listed below:

- D^0 channel, wrong-sign μ (B combinatorial): The wrong-sign $D^0\mu^+$ candidates are used in place of the right-sign $D^0\mu^-$.
- D^* channel, wrong-sign μ (B combinatorial): The wrong-sign $D^0\mu^+$ candidates are used.

Table 5.4: Additional offline selections shared among D^0 and D^* channels.

Event-Level	Variable	Cuts
GEC	nSPDhits	$< 450^*$
Particle	Variable	Cuts
K, π	p_T	$> 500 \text{ MeV}$
	HLT1 TOS track p_T	$> 1.7 \text{ GeV}$
	p	$< 200 \text{ GeV}^\dagger$
	χ^2_{IP}	> 45
	isMuon	False
	DLL K	$> 4 (K), < 2 (\pi)$
	GhostProb	< 0.5
D^0	p_T	$> 2 \text{ GeV}$
	$ \pi p_T + K p_T $	$> 1.4 \text{ GeV}$
	$\log(\text{Impact Parameter})^\ddagger$	> -3.5
	χ^2_{IP}	> 9
	$\chi^2_{\text{vertex}}/\text{ndf}$	< 4
	χ^2_{FD}	> 250
	DIRA	> 0.9998
μ	p	3-100 GeV
	$\log_{10}(1 - \mathbf{p}_\mu \cdot \mathbf{p}_i / (\mathbf{p}_\mu \mathbf{p}_i)), i \in \{K, \pi\}$	> -6.5
	η	1.7-5
	χ^2_{IP}	> 45
	GhostProb	< 0.5
	DLL μ	> 2
	DLL e	< 1
	$\mu\text{BDT}^\S$	> 0.25

* This is to reduce correlation for *TIS* and *TOS* candidates to make trigger emulation more robust. See Section 6.1.3 for more details.

† This is to make p consistent with extended PIDCALIB2 binning.

‡ Impact Parameter in terms of the Primary Vertex. That is, the reconstructed D^0 is inconsistent of coming from Primary Vertex directly.

§ This is a BDT-based μ PID algorithm, which is discussed in Section 5.5.2.

- D^* channel, wrong-sign π (D^* combinatorial): The right-sign $D^0\mu^-$ candidates are used, but a wrong-sign slow π^- is selected so the $D^0\pi^-$ pair forms a wrong-sign D^{*-} .

Table 5.5: Offline cuts exclusive to D^0 channel.

Particle	Variable	Cuts
$D^0\mu$	m	$< 5200 \text{ MeV}^*$
	$\text{Max}(\Delta m_{\text{veto},1} - 145.454 , \dagger$	$> 4 \text{ MeV}$
	$ \Delta m_{\text{veto},2} - 145.454)$	
	$\Delta m_{\text{alt hypo}}^\dagger$	$> 165 \text{ MeV}$
	transverse flight distance	$< 7 \text{ mm}$
	$\chi^2_{\text{vertex}}/\text{ndf}$	< 6
	$DIRA$	> 0.9995

* The odd-looking 5200 is to avoid a peak around 5280 due to misID in B^- mass spectrum. For $\bar{B}^0$, the misID peak around 5280 is insignificant, so the cut is placed at 5280, at which value the neutrino is produced *at rest in B frame*, which covers a small phase space.

† To reduce overlapping events between D^0 and D^* channels. Discussed in Section 5.5.1.

Table 5.6: Offline cuts exclusive to D^* channel.

Particle	Variable	Cuts
slow π	GhostProb	< 0.25
D^{*+}	$ m - m_{D^0} - 145.454 ^*$	$< 2 \text{ MeV}$
	$\chi^2_{\text{vertex}}/\text{ndf}$	< 10
$D^0\mu$	$\chi^2_{\text{vertex}}/\text{ndf}$	< 6
	$DIRA$	> 0.999
$D^{*+}\mu$	m	$< 5280 \text{ MeV}$
	transverse flight distance	$< 7 \text{ mm}$
	$\chi^2_{\text{vertex}}/\text{ndf}$	< 6
	χ^2_{vertex}	< 24
	$DIRA$	> 0.9995

* This is to require $m_{D^{*+}}$ to close to its reference (PDG) mass, up to a reconstruction effect on m_{D^0} .

5.1.3 Fake muon samples

The fake μ samples, labeled as “misID samples”, are used to estimate the contribution from events where the μ candidate is in fact a particle of a different species. Such mis-identifications happen in both right-sign and wrong-sign nominal samples thus the fake

muon samples similarly contain two components: both right-sign and wrong-sign misID samples share the same selections as the respective nominal sample, except for μ PID which is replaced by an explicit μ veto (not `isMuon`). Also, only 10% of event passing the fake μ online selection are stored to save storage space. The contribution due to μ misID will be studied in Section 8.1.1.

5.2 MC simulation

A large number of simulated events are used to model the signal, normalization, and partially reconstructed backgrounds in data. To reduce simulation time, tracker-only (TO) MC, which has all but the tracking system of the detector turned off (set to passive material), are used. Additional small full simulation (FullSim) samples are produced for the year 2016 to aid emulations of the missing detector responses (mainly trigger). The emulation procedures are discussed in Chapter 6.

MC samples used for both D^0 and D^* channels are listed in Table 5.7, with TO and FullSim numbers listed separately. Some samples contribute to both channels due to feed down⁴, and are indicated in the same table⁵.

Both TO and FullSim MC samples are loosely filtered to keep events in LHCb acceptance, mostly to save storage space, with filtering cuts listed in Table 5.8. The online and offline selection are aligned with that of the right-sign sample, with the understanding that trigger and PID responses are emulated offline and applied as weights (for L0 trigger and PID) and cuts (for HLT1 and HLT2 trigger). Additional truth level requirements⁶ are placed to ensure the selected decay chain is compatible with simulation specification.

⁴This means that the selection procedure was unable to find the additional slow π required to reconstruct a D^* , thus only a D^0 was constructed and the candidate is *feed down* from D^* channel to D^0 channel.

⁵As a remark, the TO and FullSim have different simulation conditions, which are denoted by **SimXXX**. The differences are technical and can be ignored for this thesis.

⁶For MC samples, each reconstructed particle is associated with a true particle ID. The truth level requirements (truth-matching) demands that the true ID is the same as specified in the simulation. For example, for a candidate from MC $B^- \rightarrow D^0 \mu^-$ decay, the reconstructed μ is required to have a true ID of a MC true μ .

Table 5.7: List of MC sample used in this analysis. Numbers are event on disk, not total simulated numbers.

Group	MC ID	Decay mode	Channels	Cocktail	2016 Sim09j FullSim	2016 Sim09k tracker-only	2017 Sim09k tracker-only	2018 Sim09k tracker-only
norm.	12573012	$B^- \rightarrow D^0 \mu^- \bar{\nu}_\mu$	D^0	no	3,564,053	45,564,529	47,965,869	64,386,408
	11574021	$\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu$	both	no	3,012,029	85,470,057	81,075,745	103,168,826
	12773410	$B^- \rightarrow D^{*0} \mu^- \bar{\nu}_\mu$	D^0	no	4,003,029	129,200,391	134,602,735	169,664,396
sig.	12573001	$B^- \rightarrow D^0 \tau^- \bar{\nu}_\tau$	D^0	no	233,966	4,152,856	3,405,213	4,317,050
	11574011	$\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau$	both	no	593,350	17,217,664	18,008,069	25,341,935
	12773400	$B^- \rightarrow D^{*0} \mu^- \bar{\nu}_\mu$	D^0	no	528,616	9,813,636	10,289,156	15,091,483
D^{**}	11874430	$\bar{B}^0 \rightarrow D^{*++} \mu^- \bar{\nu}_\mu$	both	$D_1^+, D_1'^+, D_2^{*+}, D_0^{*+}$	2,865,898	46,653,556	45,469,066	58,082,278
	11874440	$\bar{B}^0 \rightarrow D^{*++} \tau^- \bar{\nu}_\tau$	both	$D_1^+, D_1'^+, D_2^{*+}, D_0^{*+}$	35,322	375,581	519,245	561,800
	12873450	$B^- \rightarrow D^{*0} \mu^- \bar{\nu}_\mu$	both	$D_1^0, D_1'^0, D_2^{*0}, D_0^{*0}$	2,333,268	37,417,148	117,729,837	48,051,754
	12873460	$B^- \rightarrow D^{*0} \tau^- \bar{\nu}_\tau$	both	$D_1^0, D_1'^0, D_2^{*0}, D_0^{*0}$	112,138	618,529	598,526	744,330
D_H^{**}	12675011	$B^- \rightarrow D_H^{*0} (\rightarrow D^0 \pi \pi) \mu^- \bar{\nu}_\mu$	D^0	Ref. [15]	419,829	6,204,217	7,215,251	8,573,968
	11674401	$\bar{B}^0 \rightarrow D_H^{*+} (\rightarrow D^0 \pi \pi) \mu^- \bar{\nu}_\mu$	D^0	Ref. [15]	353,015	6,997,221	6,746,518	12,956,078
	12675402	$B^- \rightarrow D_H^{*0} (\rightarrow D^{*+} \pi \pi) \mu^- \bar{\nu}_\mu$	both	Ref. [15]	294,872	5,560,586	4,739,776	6,162,695
	11676012	$\bar{B}^0 \rightarrow D_H^{*+} (\rightarrow D^{*+} \pi \pi) \mu^- \bar{\nu}_\mu$	both	Ref. [15]	290,791	4,824,507	4,834,264	8,353,204
	12875440	$B^- \rightarrow D_H^{*0} (\rightarrow D^{*0} \pi \pi) \mu^- \bar{\nu}_\mu$	D^0	Ref. [15]	450,907	7,840,307	7,983,973	10,039,996
D_s^{**}	13874020	$\bar{B}_s^0 \rightarrow D_s^{*+} (\rightarrow D^0 K^+) \mu^- \bar{\nu}_\mu$	D^0	no	259,166	1,654,215	1,732,571	2,214,625
	13674000	$\bar{B}_s^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu$	D^*	D_{s1}^+, D_{s2}^{*+}	142,549	1,498,067	1,531,966	2,070,074
DDX	11894600	$\bar{B}^0 \rightarrow D^0 X_c (\rightarrow X' \mu^\mp \bar{\nu}_\mu) X$	D^0	Ref. [15]	1,062,375	36,888,835	37,826,396	49,768,186
	12893600	$B^- \rightarrow D^0 X_c (\rightarrow X' \mu^\mp \bar{\nu}_\mu) X$	D^0	Ref. [15]	1,196,059	21,854,863	22,577,692	28,900,392
	11894200	$\bar{B}^0 \rightarrow D^0 D_s^- (\rightarrow X' \tau^- \bar{\nu}_\tau) X$	D^0	Ref. [15]	84,533	1,231,353	1,145,771	1,367,357
	12893610	$B^- \rightarrow D^0 D_s^- (\rightarrow X' \tau^- \bar{\nu}_\tau) X$	D^0	Ref. [15]	255,725	2,818,348	2,625,208	4,834,213
	11894610	$\bar{B}^0 \rightarrow D^{*\pm} X_c (\rightarrow X' \mu^\mp \bar{\nu}_\mu) X$	D^*	Ref. [15]	1,188,596	16,188,308	13,480,119	16,837,410
	12895400	$B^- \rightarrow D^{*\pm} X_c (\rightarrow X' \mu^\mp \bar{\nu}_\mu) X$	D^*	Ref. [15]	365,215	7,000,178	7,459,109	6,798,676
	11894210	$\bar{B}^0 \rightarrow D^{*+} D_s^- (\rightarrow X' \tau^- \bar{\nu}_\tau) X$	D^*	Ref. [15]	113,422	1,678,170	1,283,823	1,631,217
	12895000	$B^- \rightarrow D^{*+} D_s^- (\rightarrow X' \tau^- \bar{\nu}_\tau) X$	D^*	Ref. [15]	117,696	899,320	1,555,844	1,282,503

Table 5.8: List of MC filtering cuts at generator level.

Particle	Variable	Cuts
K, π	p_x/p_z	-0.38–0.38
	p_y/p_z	-0.28–0.28
	θ	> 0.01 rad
	p_T	> 250 MeV
μ	p_x/p_z	-0.38–0.38
	p_y/p_z	-0.28–0.28
	θ	> 0.01 rad
	p	> 2950 MeV
D^0	p_T	> 15 GeV
	p	> 2450 MeV

This effectively removes ghost (combinations of random track segments in the detector tracking system) and misidentified tracks, providing a consistent modeling and ensuring orthogonality between MC samples.

The highly excited (heavy) D^{**} and DDX MC are cocktails, that is, samples containing a mixture of different simulated decay modes, with their compositions listed in Ref. [15]. The heavy D^{**} MC suffer from the problem that the $D^{**} \rightarrow D^{(*)}\pi^0\pi^0$ decays are not simulated, inconsistent with the branching fraction expectations from isospin symmetry. The missing $\pi^0\pi^0$ contribution is modeled by randomly selecting $\frac{1}{3}$ of $\pi^+\pi^-$ candidates and not applying the veto on additional tracks.

5.3 Selecting signal (ISO) and control background (1OS, 2OS, DD) skims with the isolation BDT

The signal and control skims (sub-samples) are separated by inspecting if there are additional charged tracks in the event that are likely originating from the same decay vertex (typically a B decay vertex, and will be referred as so in later text) as the reconstructed $D^{(*)}\mu$ pair; if there is no such track, the $D^{(*)}\mu$ pair is considered “isolated” and is placed in

the signal skim; otherwise it is categorized as one of the control background skims.

The isolation BDT A multivariate BDT, referred as the “isolation BDT”, conceptually identical to the one used in [15], is trained on LHCb run 2 simulation with training variables listed in Table 5.9 to provide a score-based estimation on if a charged track is originating from the B ($D^{(*)}\mu$) vertex or not.

Table 5.9: Training variables for the isolation BDT.

Variable	Comment
PV χ^2_{IP}	Increase in the χ^2 of the primary pp vertex (PV) fit if the track is included. Higher χ^2 implies that the track is less likely to coming from the fitted vertex.
SV χ^2_{IP}	Increase in the χ^2 of the secondary $D^{(*)}$ vertex (SV) fit if the track is included.
track p_T	-
track opening	The cosine of the angle between the track and the flight direction of the $D^{(*)}\mu$ pair.
χ^2_{FD}	The significance of the flight distance fit ($\frac{\text{flight distance}}{\sigma(\text{flight distance})}$) with this track included.
$\Delta\chi^2_{\text{FD}}$	The difference in χ^2_{FD} with and without the track included.

The isolation BDT is applied on *all recorded* charged tracks in the event, providing isolation scores for each of them where higher score implies higher probability of coming from the B vertex, referred to as “anti-isolated” (Note: “isolated” is a property of the $D^{(*)}\mu$ pair, while “anti-isolated” is a property of the other charged tracks in the event!). Three tracks with highest isolation scores (the ones that are most likely to be anti-isolated) are saved in descending order⁷.

Skim selection criteria A brief description for the signal skim and the main physics background control skims is provided below. The dominant decay mode for each skim is also

⁷So the first one is the most anti-isolated one among all charged tracks in the event.

listed; these skims still contain a composition of many decay modes. The actual isolation selections are listed in Table 5.10.

- ISO: Enriched in $B \rightarrow D^{(*)}l\bar{\nu}_l$ decays, with $l \in \{\mu, \tau\}$. This is referred to as the signal, or the “isolated”, skim. It requires that no additional charged track compatible with coming from the B vertex (in a probabilistic sense, with probability related to the isolation scores). Such requirement translates into a cut on the maximum BDT score: $\text{BDT}_1 < 0.15$.
- 1OS: Enriched in $B \rightarrow D^{**}l\bar{\nu}_l$. This control skim, enriched in excited charm states, requires one and only one additional anti-isolated charged long track that is compatible with a π PID hypothesis and has the correct charge for a $D^{**} \rightarrow D^{(*)}\pi^-$ decay.

The BDT score requirements for this skim are: $\text{BDT}_1 > 0.15$ and $\text{BDT}_2 < 0.15$.

- 2OS: Enriched in $B \rightarrow D_H^{**}l\bar{\nu}_l$, where D_H^{**} stands for heavy D^{**} mesons. This control skim is enriched in highly excited (heavy) charm states, which is selected by requiring two and only two anti-isolated π -like long tracks of opposite charge, capable of $D_H^{**} \rightarrow D^{(*)}\pi^+\pi^-$ decay.

This skim also provides an independent selection of $B \rightarrow D^{(*)}DX$ backgrounds whose decay products contain many π .

The BDT score requirements are: $\text{BDT}_1 > 0.15$ and $\text{BDT}_2 > 0.15$ and $\text{BDT}_3 < 0.15$.

- DD: Enriched in $B \rightarrow D^{(*)}DX$, with dominant $D \rightarrow K^+\mu^-\bar{\nu}_\mu X$ sub-decays. This is a double-charm (DD) control sample, which is selected by requiring at least one anti-isolated track, a K -like long track⁸ and a high- p_T track in the three most anti-isolated tracks⁹.

The BDT score requirement: $\text{BDT}_1 > 0.15$

⁸This means the track went through both upstream and downstream trackers, and is generally of good tracking quality.

⁹Note that these three requirements can be satisfied by a single track.

Table 5.10: Signal and control sample isolation requirements.

Sample	Variable	Cuts
ISO	BDT ₁	< 0.15
1OS	BDT ₁	> 0.15
	BDT ₂	< 0.15
	Track type ₁	= 3*
	p_1	> 5 GeV
	$p_{T,1}$	> 0.15 GeV
	MC15TuneV1ProbNNK ₁	< 0.2
	$Q_1 \cdot \text{PID}_{D^0}^\dagger$	> 0
	$Q_1 \cdot \text{PID}_{D^*}^\ddagger$	< 0
2OS	BDT ₁	> 0.15
	BDT ₂	> 0.15
	BDT ₃	< 0.15
	Track type ₁	= 3
	Track type ₂	= 3
	$\text{Max}(p_1 \cdot (p_{T,1} > 0.15 \text{ GeV}), p_2 \cdot (p_{T,2} > 0.15 \text{ GeV}))$	> 5 GeV
	MC15TuneV1ProbNNK ₁	< 0.2
	MC15TuneV1ProbNNK ₂	< 0.2
	$Q_1 \cdot Q_2$	< 0
DD	BDT ₁	> 0.15
	$\text{Max}(p_1 \cdot (p_{T,1} > 0.15 \text{ GeV}),$	
	$p_2 \cdot (p_{T,2} > 0.15 \text{ GeV}) \cdot (\text{BDT}_2 > -1.1),$	> 5 GeV
	$p_3 \cdot (p_{T,3} > 0.15 \text{ GeV}) \cdot (\text{BDT}_3 > -1.1))$	
	$\text{Max}(\text{MC15TuneV1ProbNNK}_{1,2,3})$	> 0.2
	Track type _{the one passing K PID requirement}	= 3
	BDT _{the one passing K PID requirement}	> -1.1

* This means a long track.

† Apply to D^0 channel, which implies that the anti-isolated π^+ can form a D^{*+} with the D^0 .

‡ Apply to D^* channel. Here it is required that the anti-isolated π^- can form a D^{*0} with the D^{*+} .

5.4 Auxiliary samples

A $B^+ \rightarrow J/\psi K^+$ sample is used to derive B kinematic and multiplicity corrections (Section 7.2.2) to MC samples. The data sample is selected by requiring two μ at the J/ψ resonance forming a vertex with another K , and the invariant mass of the $\mu^+\mu^-K^+$ final state around the reference mass of B^+ . The offline selection cuts for $J/\psi K^+$ are listed in Table 5.11. The MC sample used to compare with $J/\psi K^+$ data is selected with the same procedure, with PID cuts applied as weights¹⁰. The number of MC events after filtering¹¹ is listed in Table 5.12.

Table 5.11: Offline cuts on $J/\psi K$ samples.

Particle	Variable	Cuts
K	p_T	> 500 MeV
	DLL K	> 4
μ^-, μ^+	p_T	> 500 MeV
	χ^2_{IP}	> 4
	DLL μ	> 2
$\mu^-\mu^+ (J/\psi)$	m_{mea}^*	3060–3140 MeV
	χ^2_{FD}	> 25
B^+	m	5150–5350 MeV
	χ^2_{vertex}	< 18
	χ^2_{IP}	< 12
	DIRA	> 0.9995

* m_{mea} refers to measured mass, which is the invariant mass given by the sum of daughters' four momenta, without any topological constraint. On the other hand, m is given by a vertex fit, which is typically of better quality.

¹⁰The weights are obtained from PIDCALIB2, with a method similar to that in Section 6.2.

An extended PID binning range is chosen to cover as much phase space of K and μ as possible. There may still be events outside the covering region, in which case the PID weights from nearby bin is applied, to avoid explicit cuts on the $J/\psi K$ data control sample.

¹¹This is also referred as “event on disk”.

Table 5.12: Additional MC control samples. All samples are full simulations.

MC ID	Decay mode	Sim09k
12143001	$B^+ \rightarrow J/\psi K^+$	4M

5.5 Additional algorithms used in selection

5.5.1 Veto of overlapping candidates

Since the selection of a $D^*\mu$ pair always requires the existence of a $D^0\mu$ pair, about 21% of events in the signal (ISO) skim of the D^* channel would contribute to the D^0 channel as well. To veto the overlapping candidates between the D^0 and D^* channels, a tool, named `TupleToolApplyIsolationVetoDst`, is added to the selection procedure, which loops over each track t in the event whose $p > 2$ GeV in the following manner:

1. A D^0t vertex is formed by combining the D^0 track and t with a fit.
2. Compute $\Delta m_{\text{veto}} \equiv m_{D^0t} - m_{D^0}$.
3. Test if $\Delta m_{\text{veto}} \in [140 \text{ GeV}, 160 \text{ GeV}]$. If it is, the track t is able to form a D^* -like particle with D^0 . Record the Δm_{veto} and isolation BDT score of the track in such a case.

It is required offline (listed in Table 5.5) that the two most anti-isolated tracks that can form a D^* , the best slow π candidates, are *incompatible* with forming a $D^0\pi$ vertex with a mass close to D^* reference mass: $|\Delta m_{\text{veto}}^{\text{candidate } 1,2} - 145.454| > 4 \text{ MeV}$.

Note that this tool does not rank tracks based *solely* on the isolation score: the poor tracking resolution of the slow π makes the association of the track more ambiguous; sometime the track may be better associated with the PV, instead of the B vertex, in which case the track would receive a small BDT score and considered *not anti-isolated*. By ranking based on the BDT scores of the tracks that fall within the $[140 \text{ GeV}, 160 \text{ GeV}]$

Δm_{veto} window only, it is more robust to retain the best slow π candidates, as reported in Ref. [15].

In addition, an alternative mass hypothesis is tested (offline, also listed in Table 5.5), where the reconstructed μ is treated as a π , and $\Delta m_{\text{alt hypo}} = m_{D^0 \mu_{\text{as } \pi}} - m_{D^0}$ is computed¹² and required to be inconsistent with the difference in D^* and D^0 PDG masses.

The veto procedures are applied on all samples of the D^0 channel only.

5.5.2 Muon PID

To maintain minimum bias on the μ p_T , a multivariate BDT, referred as μ BDT, using the uBDT method in the TMVA class, is trained to reject misidentified μ while keeping the rejection efficiency approximately flat on p_T , as described in Ref. [15].

The μ BDT has been re-trained on LHCb run 2 MC samples. There is an ongoing effort to study its efficiency against standard run 2 PID variables. Early reports suggest the flatness on p_T is maintained while rejection efficiencies relative to existing PID cut are high for all but protons (p), as shown in Fig. 5.1.

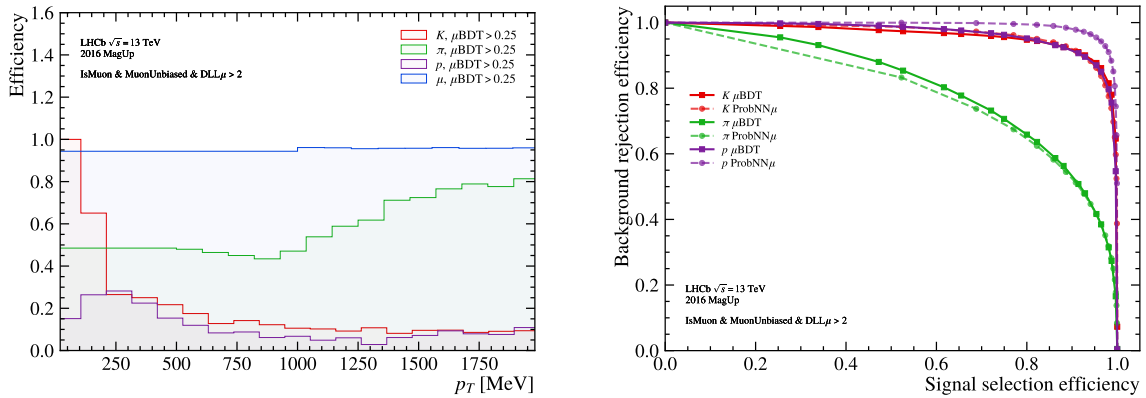


Figure 5.1: Preliminary μ BDT study. Left: μ BDT μ selection efficiency is flat in p_T , with global cut $\text{isMuon} \& \text{DLL}\mu > 2$, and $\mu\text{BDT} > 0.25$. Right: with the same global cut, μ BDT is more effective in rejecting π than LHCb official MC15TuneV1ProbNN μ . The K rejection efficiencies are similar between the two; the p rejection efficiency is lower for μ BDT, but the absolute rejection efficiency is high enough.

¹²By using the 3-momentum of the μ , while using m_π as the invariant mass of the track.

Chapter 6: Emulation of detector responses on MC simulation

To ensure that the statistical uncertainties due to the limited size of the MC samples are small compared to those of data, it is important to have MC samples with a size at least 4 times as large as data. Due to the sheer size of LHCb run 2 data (a factor of ~ 1.26 for 2016 alone compared to full run 1 despite smaller integrated luminosity due to a larger $b\bar{b}$ cross section and higher triggering efficiencies), it is not computationally possible to generate enough full simulation (FullSim) MC, where every detector response is simulated.

It is estimated in Ref. [35] that about 85% of the computation time is spent on simulating RICH and the calorimeter system. Tracker-only (TO) MC sets all but the tracking system of the detector to passive material, making the simulation about 8 times faster than FullSim. Hence, to reach the target statistical uncertainties, TO MC is (almost) exclusively used for this analysis. Therefore, detector responses like trigger and particle identification (PID) need to be emulated offline. The emulation procedures for these responses are described in the rest of the chapter.

6.1 Trigger emulation

6.1.1 Emulation of L0Hadron TOS

As discussed in Section 5.1, the L0 trigger requirements for this analysis are: either the D^0 candidate fires the L0 hadron trigger line (D^0 L0Hadron TOS) or the rest of the event aside from the $D^0\mu$ candidate (and its decay chain) triggers at least one L0 trigger (B L0Global TIS). Note that “TOS” stands for *Triggered On Signal* where *signal* in this

case is the D^0 candidate, and “TIS” means *Trigger Independent of Signal* where the *signal* refers to the $D^0\mu$ pair. For more information about trigger TIS, TOS, TISTOS, and TOB, consult Appendix D.

The L0 hadron line (L0Hadron) is triggered when the sum of the transverse energy¹ E_T deposited in a single 2×2 HCAL cluster is above a set threshold², as described in Ref. [34]. For L0Hadron TOS, it is required that for all HCAL clusters associated with the signal track (D^0 in this case), with the association predicted by the tracking system, the highest E_T is above the triggering threshold.

Tracker-only MC records the tracker-predicted position where a track hits the calorimeters. Hence, though the triggering variable E_T is not accessible, variables related to E_T , or, equivalently, to the trigger decision, are. It is conceivable that L0Hadron TOS can be emulated by providing the relevant variables and the trigger decision to train a regressor, which can in turn be used to predict trigger decisions based on the training variables alone, without requiring emulating E_T at all. The emulation procedure, inspired by Ref. [35], is described below:

1. The training variables, listed in Table 6.1, and the trigger decision from a FullSim MC of decay $B^0 \rightarrow D^{*+}\mu\nu$, a normalization mode, are used to train **XGboost**, a BDT-based regressor [36].
2. Once the model is trained, with just the training variables, **XGboost** predicts the probabilities of events passing L0 hadron trigger.
3. The probabilities are applied as weights to a separate validation sample of the same FullSim MC to check the quality of the emulation. It is shown in Fig. 6.1 that the emulated L0Hadron TOS trigger has excellent agreement with the *simulated* trigger in FullSim.

Table 6.1: Variables used to train **XGboost** for L0 hadron emulation.

Group	Variables
Event-level	nTracks
D^0	p, p_T
K^-	p, p_T , real E_T^* , HCAL x -projection, HCAL y -projection, HCAL region [†]
π^+	same as K^-

* This is the E_T measured by the tracker. In massless limit, $E_T = p_T$.

† This indicates whether the hits are in the inner or outer region of HCAL. The outer region has a coarser granularity than the inner, thus it is more likely that the energies from K^- and π^+ are deposited in the same cluster, enhancing L0Hadron TOS efficiency.

6.1.2 Emulation of L0Global TIS

As discussed before, L0Global TIS is determined by *the rest of event* (everything but the reconstructed B decay chain) which is *not saved* after event reconstruction. Therefore, a completely different approach is needed for its emulation.

Instead of predicting the event-by-event L0Global TIS probability based on the properties of the rest of event, it is possible to measure the L0Global TIS efficiencies binned in *certain reconstructed variables*³, and then apply L0Global TIS efficiencies based on the bin the event falls in.

In addition, for hadron colliders, the rest of event is *busy* due to hadronization: Indeed, for data reconstructed in the $B^- \rightarrow D^0(\rightarrow K^-\pi^+)\mu^-\bar{\nu}_\mu$ mode which is used in this analysis, the average number of tracks per event is 170, whereas the selected $D^0\mu^-$ final state contains only 3 tracks (K^- , π^+ , and μ^- , as D^0 and B^- are reconstructed from these tracks)! Therefore, it can be argued that the rest of event is essentially the same for different B decays, and L0Global TIS efficiencies are portable. Assuming the L0Global TIS efficiencies can be measured in a data sample reconstructed in a particular decay mode,

¹Defined as: $E_T \equiv \sqrt{m^2 + p^2} \sin \theta$. In massless limit, $E_T = p_T$.

²See Table 1 in [34] for year-by-year thresholds of each L0 trigger.

³The choice of variables will be discussed later.

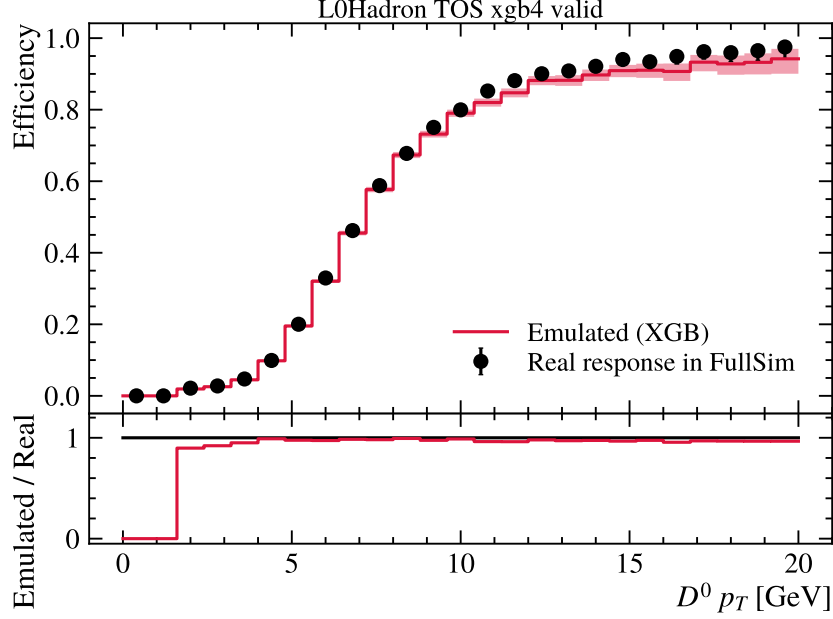


Figure 6.1: The emulated trigger probabilities applied as a weight to FullSim has excellent agreement with real L0Hadron TOS trigger *simulated* in FullSim.

the measured efficiencies can be applied to TO MC as weights regardless of the simulated decays.

To summarize, if the following assumptions hold, then the emulation outlined above should work:

1. The L0Global TIS efficiencies can be measured in data.
2. The L0Global TIS efficiencies do not depend on B decay mode.

Before we proceed, let us define L0Global TIS efficiency $\epsilon_{\text{L0Global TIS}}$ more precisely. According to [37], in LHCb $\epsilon_{\text{L0Global TIS}}$ is conventionally defined as:

$$\epsilon_{\text{L0Global TIS}} \equiv \frac{N_{\text{L0Global TIS \& selected}}}{N_{\text{selected}}} \quad (6.1)$$

which is not directly evaluable in real data, as N_{selected} refers to the number of selected events *regardless* of their trigger decisions.

However, the conditional efficiency, $\epsilon_{\text{L0Global TIS}|\text{L0Muon TOS}}$, is measurable:

$$\epsilon_{\text{L0Global TIS}|\text{L0Muon TOS}} \equiv \frac{N_{\text{L0Global TIS \& L0Muon TOS \& selected}}}{N_{\text{L0Muon TOS \& selected}}} \quad (6.2)$$

as long as L0Muon TOS and all L0 triggers (L0Global) on the rest of event are uncorrelated, $\epsilon_{\text{L0Global TIS}|\text{L0Muon TOS}}$ is equivalent to $\epsilon_{\text{L0Global TIS}}$. Such a method is called *TISTOS method*. Unfortunately, there is a correlation between the two triggers due to the fact that $b\bar{b}$ are produced in pairs, making the kinematics of the selected signal B meson and the background $\bar{B}$ meson, part of rest of event, correlated.

Still, it is argued in Ref. [37] that L0Global TIS and L0Muon TOS are *uncorrelated* in a small region of signal B phase space. In Ref. [35] it is checked that binned $\epsilon_{\text{TISTOS}} = \epsilon_{\text{TIS}}$ in $B \rightarrow J/\psi K$ FullSim MC⁴, and evaluated that $\epsilon_{\text{L0Global TIS}}$ on data reconstructed in $J/\psi K$, binned in $\log(p_z)$ and $\log(p_T)$ of the B meson. It is also checked that the efficiencies are portable among $B \rightarrow J/\psi K$, $B \rightarrow D^+ \mu \nu$, and $B \rightarrow D^+ \tau \nu$ decays with MC samples. Thus, it is possible to measure $\epsilon_{\text{L0Global TIS}}$ *binned* in B kinematic variables with the TISTOS method.

In Fig. 6.2 we check that the portability still holds for this analysis in $B \rightarrow D^{*+} \mu \nu$ and $B \rightarrow D^{*+} \tau \nu$. The efficiencies obtained in Ref. [35] are applied to all TO MC samples in this analysis as a function of the true momenta, as the reconstruction in this analysis contains missing neutrino(s), whereas $J/\psi K$ does not.

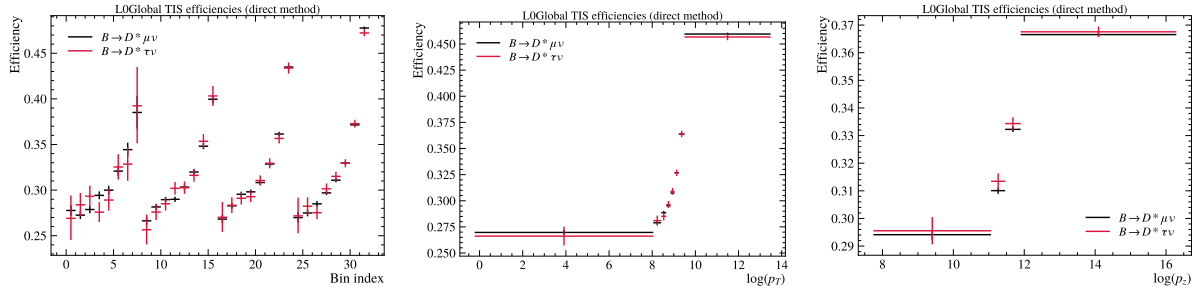


Figure 6.2: Check that L0Global TIS is portable among signal and normalization MC modes. The p_T and p_z momenta are MC true momenta of the B meson. Bin index represents the index from an unrolled 2D true- p_T -true- p_z histogram. Efficiencies are evaluated according to Eq. (6.1), as the MC samples used here are *not* filtered on trigger decision, so N_{selected} can be evaluated directly.

Efficiencies are in good agreement across all bins.

⁴TISTOS does not always work. See Appendix A.3 for a counter example.

6.1.3 Correlation between L0Hadron TOS and L0Global TIS

L0 triggers have a global event cut on the number of SPD hits. For all but the DiMuon trigger⁵, the cut is placed at 450; for DiMuon, 900, as listed in [35]. For events that are both *TIS and TOS*, a non-trivial correlation is introduced by this cut, because it affects the L0 TIS efficiencies inconsistently: On a small patch of the signal *B* phase space, for events that are TIS on *non-DiMuon*, TIS and TOS efficiency is:

$$\epsilon_{\text{TIS and TOS}} = \epsilon_{\text{L0-non-DiMuon TIS}} \cdot \epsilon_{\text{L0Hadron TOS}} \quad (6.3)$$

For L0DiMuon TIS only, however, its efficiency is affected by the SPD cut:

$$\epsilon_{\text{L0DiMuon TIS \& nSPDhits} < 450} \neq \epsilon_{\text{L0DiMuon TIS}} \quad (6.4)$$

$$\Rightarrow \epsilon_{\text{TIS and TOS}} = \epsilon_{\text{L0DiMuon TIS \& nSPDhits} < 450} \cdot \epsilon_{\text{L0Hadron TOS}} \quad (6.5)$$

$$\neq \epsilon_{\text{L0DiMuon TIS}} \cdot \epsilon_{\text{L0Hadron TOS}} \quad (6.6)$$

Therefore:

$$\epsilon_{\text{TIS and TOS}} = (\epsilon_{\text{L0-non-DiMuon TIS}} + \epsilon_{\text{DiMuon TIS \& nSPDhits} < 450}) \cdot \epsilon_{\text{L0Hadron TOS}} \quad (6.7)$$

$$\neq \underbrace{\epsilon_{\text{L0Global TIS}}}_{=\epsilon_{\text{L0-non-DiMuon TIS}} + \epsilon_{\text{L0DiMuon TIS}}} \cdot \epsilon_{\text{L0Hadron TOS}} \quad (6.8)$$

To minimize the correlation, a $\text{nSPDhits} < 450$ is applied on all data samples⁶, which removes about an additional 4.4% of candidates for both D^0 and D^* channel. After the SPD cut is applied, it is checked that $\epsilon_{\text{TIS or TOS}} \approx \epsilon_{\text{TIS}} + \epsilon_{\text{TOS}} - \epsilon_{\text{TIS}} \cdot \epsilon_{\text{TOS}}$ in [35], that is, TIS efficiencies can be considered as uncorrelated from TOS efficiencies.

6.1.4 Emulation of Hlt1TrackMVA and Hlt1TwoTrackMVA

The event selection requires that either K or π is TOS on **Hlt1TrackMVA**, or D^0 is TOS on **Hlt1TwoTrackMVA**. The single track **Hlt1TrackMVA** requires a track that has a score based on p_T and χ^2_{IP} to be above some threshold; the two track **Hlt1TwoTrackMVA** feeds the

⁵Ignore Muon high p_T line for now.

⁶It is applied on the $J/\psi K$ sample as well, the one used to obtain binned L0Global TIS efficiencies. For TO MC, there is no nSPDhits variable.

p_T and χ^2_{IP} of both tracks to a **MatrixNet** MVA, requiring the two-track combination to pass the MVA selection.

All required variables are present in TO MC, and are extracted by adding a tool⁷ to event reconstruction. The HLT1 triggers also require events to pass Global Event Cuts (GEC), as listed in Table 6.2, to remove high-pile-up events; the GEC are applied in the emulation offline. There is also an online-offline difference: The real HLT1 triggers (online) use VELO-TT tracks which require at least three hits in the TT, whereas the emulation (offline) can only use VELO tracks which have no TT requirement. Therefore, additional cuts and a 4.2% penalty factor are imposed to account for the differences. The HLT1 trigger cuts are listed Tables 3 and 4 in [35].

Table 6.2: Global event cuts for HLT1. Taken from [35].

Global event cuts (GEC)
$50 < \text{nVeloClusters} < 6000$
$50 < \text{nITClusters} < 3000$
$50 < \text{nOTClusters} < 15000$

As can be seen in Figs. 6.3 and 6.4, the emulated **Hlt1TrackMVA** triggers well-reproduce the FullSim efficiencies in all fit variables, whereas for **Hlt1TwoTrackMVA** there is a roughly constant efficiency difference which is deemed non-essential.

6.1.5 Emulation of **Hlt2XcMuXForTauB2XcMu**

The selected B is required to be TOS for **Hlt2XcMuXForTauB2XcMu**. All the selection variables listed in Table 5.2 are available in the selection procedure so these cuts are applied directly.

⁷Named **RelInfoHLT1Emulation**, available at <https://github.com/umd-lhcb/TrackerOnlyEmu/tree/master/davinci/Phys>, courtesy of the authors of [35].

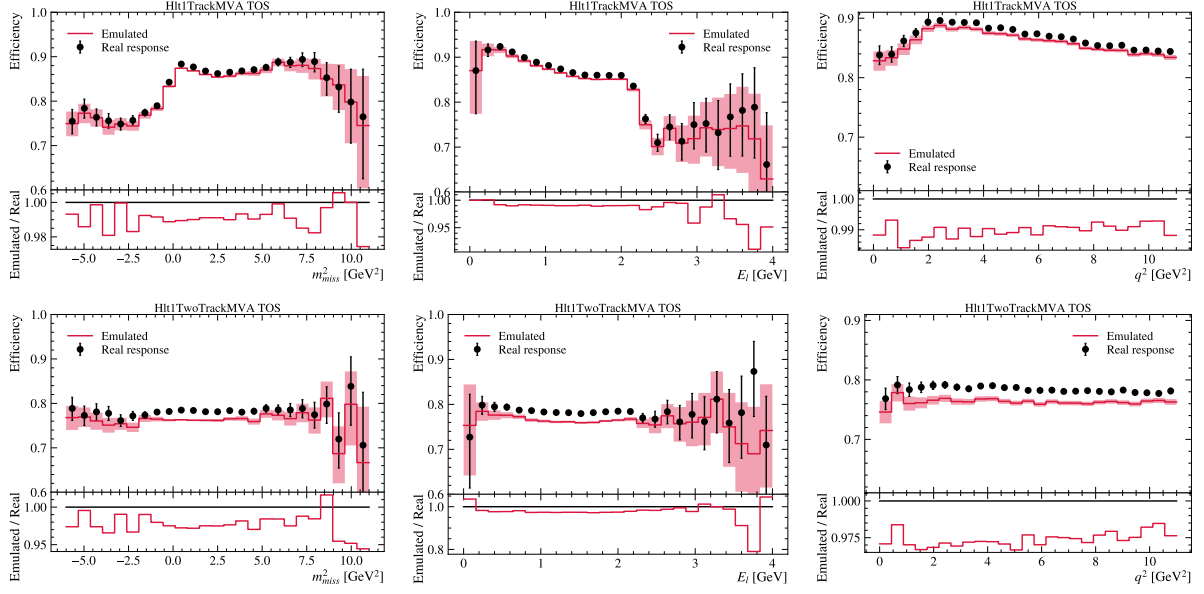


Figure 6.3: Emulated HLT1 triggers vs. real response in FullSim. For B^+ . The `Hlt1TrackMVA` agree very well, whereas the `Hlt1TwoTrackMVA` still has a constant, albeit small ($\sim 2.25\%$ in q^2), discrepancy.

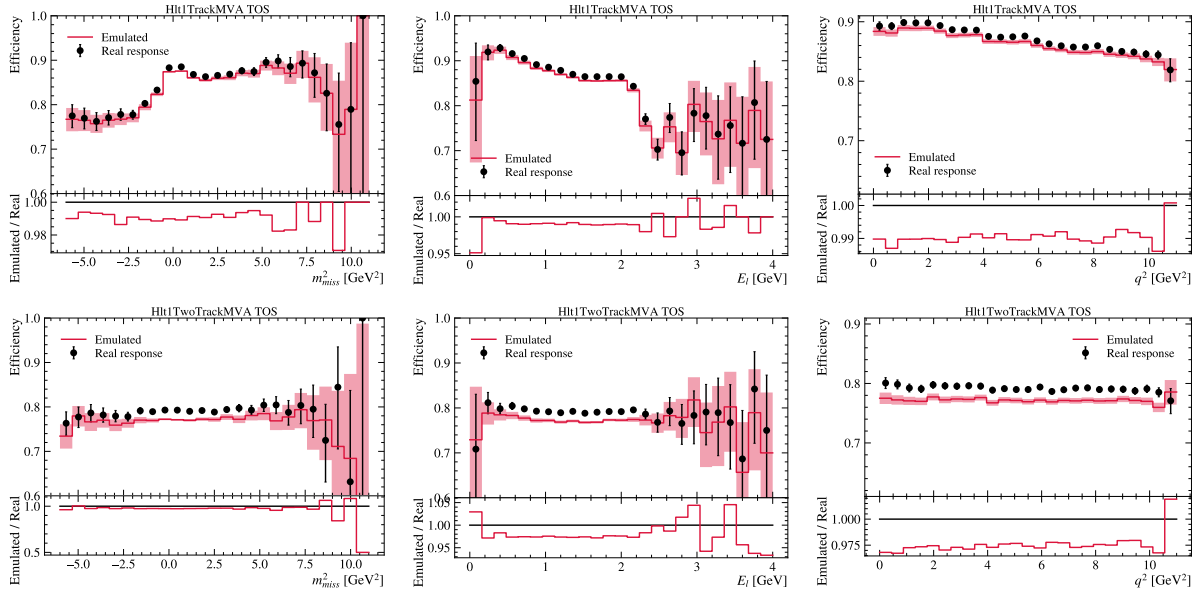


Figure 6.4: Emulated HLT1 triggers in B^0 have similar level of agreement to real response compared to that of B^+ .

6.2 Particle identification emulation

Tracker-Only (TO) MC has the RICH and calorimeters, the sub-detectors responsible for particle identification (PID), set to passive material, thus no PID requirements can be applied directly on the TO MC. Therefore, the PID requirements for MC are emulated as efficiencies (binned in p , η , and nTracks) and applied as weights.

Most of the PID efficiencies are obtained with PIDCALIB2, an official LHCb package which provides PID efficiencies evaluated at user-specified cuts on officially produced PID samples from real data with a customizable binning. The PID samples are highly enriched in charged K, π, p, e, μ , and are obtained from reconstructed $K_s^0 \rightarrow \pi^+\pi^-$, $\Lambda_c^0 \rightarrow p^+\pi^-$, $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$, $J/\psi \rightarrow \mu^+\mu^-$, and $B^+ \rightarrow J/\psi(\rightarrow e^+e^-)K^+$ data samples. These samples are shown to have high purity in [38].

It is worth noting that the PIDCALIB2 e sample lacks μ BDT, so the efficiency of all selection requirements except for the μ BDT are taken from PIDCALIB2, with the μ BDT efficiency obtained from the reconstructed events in the $\bar{B}^0 \rightarrow D^{*+}\mu^-\bar{\nu}_\mu$ MC sample whose μ is actually a true e .

The skim selection as listed in Table 5.10 contains a MC15TuneV1ProbNNK PID cut on a track of unknown particle type. To emulate this cut, the PID efficiencies from all particle types, namely K, π, p, e, μ , and ghost, are needed; the *ghost* type refers to fake tracks arising from combinations of random hits in the tracking system, and is not enriched in *any* data sample. Therefore, the ghost PID efficiency is obtained from the same $\bar{B}^0 \rightarrow D^{*+}\mu^-\bar{\nu}_\mu$ MC sample with the reconstructed μ truth-matched to ghost⁸.

⁸That is, the truth-matching failed to find a matching MC true particle.

Chapter 7: Correction and validation of the MC simulation

This chapter describes updates to form factor modeling and corrections to detector responses on MC. The procedure is the following: First, the outdated form factor models are updated with a reweighting procedure documented in Section 7.1. Then, the known problematic detector responses, such as tracking efficiency and B production kinematics, are corrected. This is termed as *initial reweighting*, and is described in Section 7.2. Afterwards, a *preliminary* fit is performed to build a *post-fit cocktail* (Section 7.3). Finally, the agreement between data and cocktail on additional kinematic and geometric variables of the B decay daughters are assessed, and a multi-stage reweighting process is performed to correct any remaining disagreements if necessary. This procedure is called *final reweighting* (Section 7.4). Additional corrections on the vertex resolution are applied after inspecting the final reweighting results. These corrections, discussed in Section 7.5, are included *before* the initial reweighting.

7.1 Form factor models

The form factor model for a $B \rightarrow Dl\nu_l$ process *defines* the truth-level q^2 distribution which in turn affects the estimated q^2 , a fit variable. Therefore, it is important to use the state-of-art form factor models. Naively this is hard, because the form model model is defined at MC generator level, any change in form factor may require a regeneration of all MC samples. In reality, however, for a given phase space point¹ $\mathcal{PS}$, the probability density that an event is generated *at* $\mathcal{PS}$ is proportional to the differential decay rate evaluated

¹The phase space is parameterized by q^2 and possibly angular variables, as defined in Eq. (3.19).

at $\mathcal{PS}$. For an event (or equivalently a given $\mathcal{PS}$ point) generated with an old form factor model, its contribution (weight) in the new model is given by the ratio of differential decay rates

$$r = \frac{d\Gamma_{\text{target}}/d\mathcal{PS}}{d\Gamma_{\text{source}}/d\mathcal{PS}} \Big|_{\text{eval at phase space point}} \quad (7.1)$$

So a change in form factor model is equivalently to a reweight, with the weight given by r .

Hammer [39], the current state-of-art form factor reweighting software, is used in this analysis to reweight $B \rightarrow (D^0|D^*|D^{**})l\bar{\nu}_l$ MC samples. **Hammer** implements a fast reweighting procedure with a multi-dimensional expansion of form factors around a central point in the form factor parameter space. The details of the form factor reweighting are given below.

7.1.1 $B \rightarrow D^0$ decays

The D^0 MC samples are generated with an up-to-date CLN form factor parameterization and reweighted to a BGL one for self-consistency and increased flexibility. Both are described in Section 3.2.1. Figure 7.1 shows the change of true q^2 distributions due to form factor reweighting for D^0 samples. The changes are minimal because both models are up-to-date.

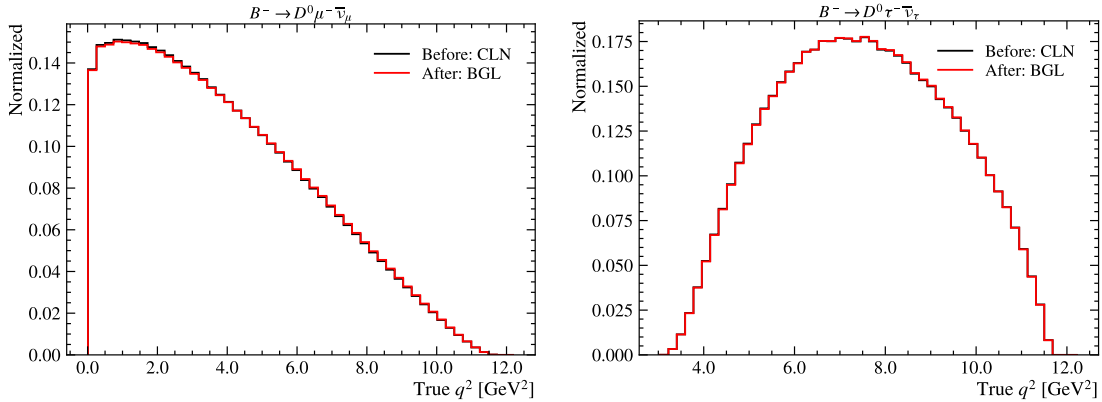


Figure 7.1: Form factor reweight for $B \rightarrow D^0 l \bar{\nu}_l$ MC samples.

7.1.2 $B \rightarrow D^*$ decays

Similar to the D^0 case, the D^* MC samples are reweighted from a CLN parameterization to a BGL one, which are described in Section 3.2.2. The change of true q^2 is shown in Fig. 7.2.

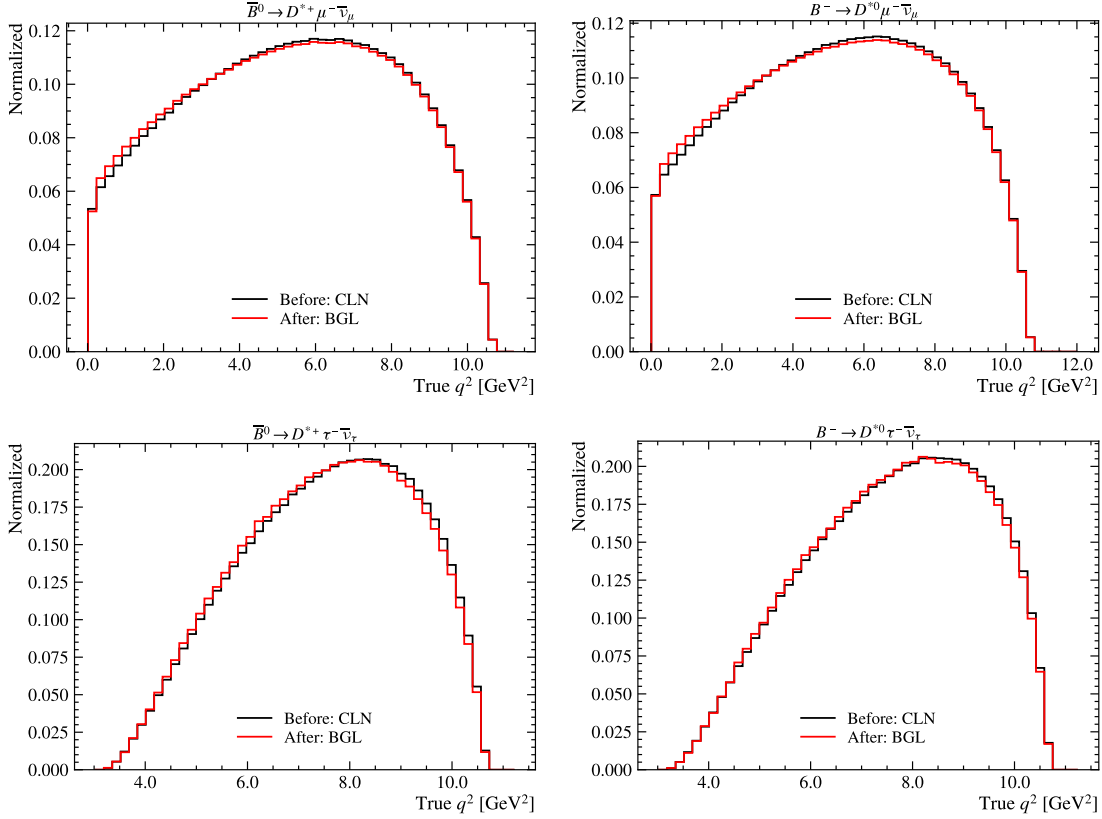


Figure 7.2: Form factor reweight for $B \rightarrow D^* l \bar{\nu}_l$ MC samples.

7.1.3 $B \rightarrow D^{**}$ decays

The four 1P D^{**} states (D_0^* , D_1 , D_1' , D_2^*) in $B \rightarrow D^{**} l \nu_l$ decays are generated with ISGW2 and reweighted to BLR. Both models are briefly discussed in Section 3.2.3.

The form factor parameters are allowed to float in the fit from the nominal values, with the shift of fit variables due to these parameters accounted for via an interpolation/ex-

trapolation mechanism². Uniquely for D^{**} , the fitted parameters have large deviations from the nominal (up to 7σ !), at which point the extrapolation becomes unphysical. A similar effect has been reported in the previous analysis.

To work around the large deviations from nominal, the *nominal* values for D^{**} form factor parameters are *shifted* based on the fit, so that the deviations are small ($\leq 2\sigma$) and the interpolation/extrapolation remains physical. The change of true q^2 with the *shifted* parameters are plotted in Figs. 7.3 and 7.4, which are taken as the nominal in the fit. The change of fit variables with the raw parameters from Ref. [18] are plotted separately in Appendix A.1.

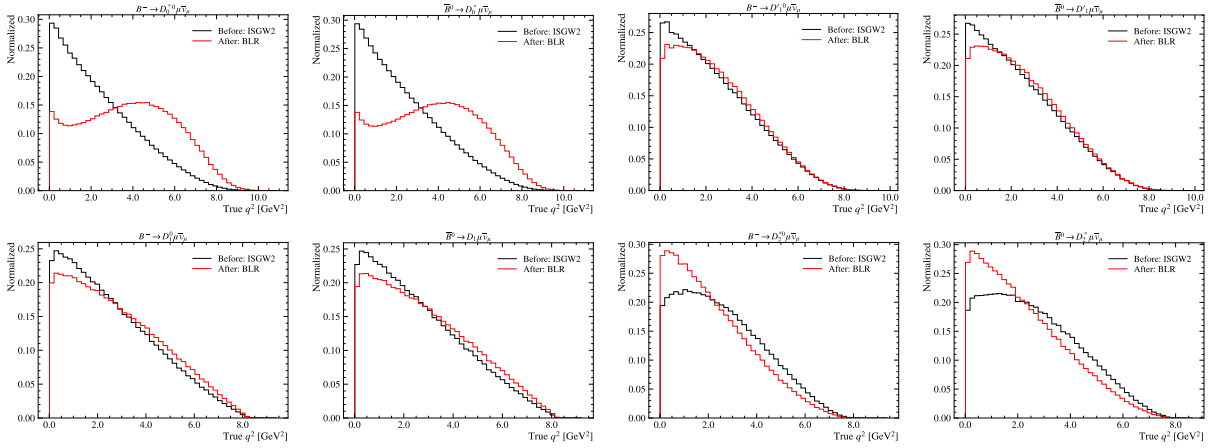


Figure 7.3: Form factor reweight effect on $D^{**}\mu$ MC templates, with shifted parameters based on an initial fit.

²More details regarding the interpolation/extrapolation are discussed in Section 8.2.

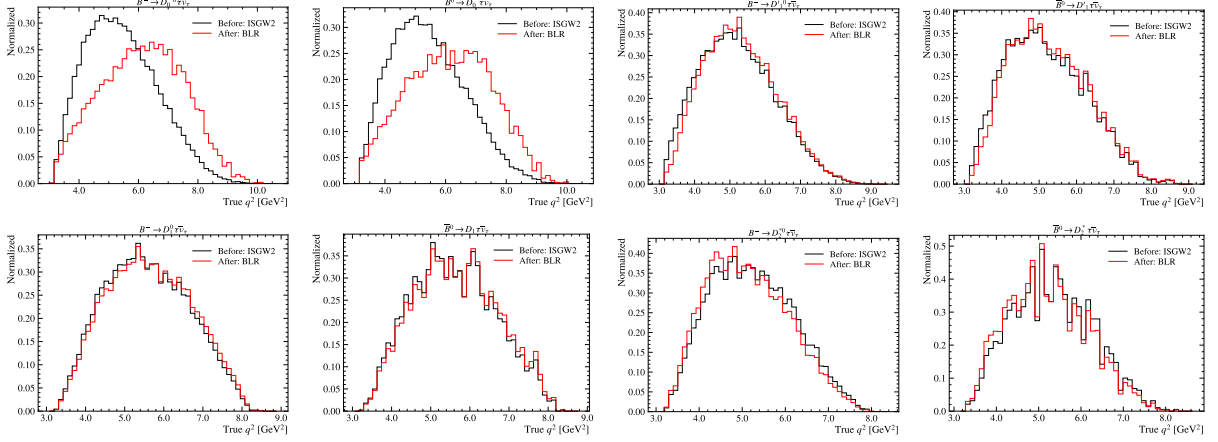


Figure 7.4: Form factor reweight effect on $D^{**}\tau$ MC templates, with shifted parameters.

7.2 Initial reweighting

7.2.1 Tracking efficiency

The track reconstruction efficiencies, defined as $\frac{N_{\text{reconstructed}}}{N_{\text{reconstructable}}}$ for *long tracks*³, are known to be different between data and MC. The efficiencies are measured with a *tag-and-probe* method on $J/\psi \rightarrow \mu^+\mu^-$ samples which involves the following steps: First, a fully reconstructed long μ -like track (tag) and a partially reconstructed⁴ μ -like track (probe) is required to form a $\mu^-\mu^+$ vertex. Then, the probe track is classified into two categories: The ones matched to a long track, defined by sharing 70% of hits with a long track in the T-stations, and the ones *not* matched to any long track. Finally, a fit on the invariant mass of the $\mu^-\mu^+$ vertex is performed to extract number of signal events separately for matched ($N_{\text{sig,matched}}$) and all ($N_{\text{sig,all}}$) probe tracks, and the tracking efficiency is computed as:

$$\epsilon_{\text{tracking}} \equiv \frac{N_{\text{reconstructed}}}{N_{\text{reconstructable}}} = \frac{N_{\text{sig,matched}}}{N_{\text{sig,all}}} \quad (7.2)$$

³Long tracks means that the tracks have hits recorded at both VELO (upstream) and T-stations (downstream). They have the best tracking quality and most analyses use them exclusively.

⁴For this analysis, partially reconstructed means that only the TT (tracker before the dipole magnet, to be replaced by UT) and T-station hits are considered for track reconstruction. This partially reconstructed track carries enough momentum information such that the invariant mass of the mother particle is reconstructed with high precision.

The efficiencies computed with TT-T-station probe tracks are referred as *long* method.

For more information, see Refs. [40, 27].

To correct for the tracking discrepancy between data and MC, the efficiency ratios ($\frac{\epsilon_{\text{data}}}{\epsilon_{\text{MC}}}$), binned in η and p of the track, between data and MC are provided by the LHCb collaboration, which are referred as TRACKCALIB efficiency ratios. The correction is applied as a weight on a per-event basis. The efficiency ratio table is shown in Fig. 7.5.

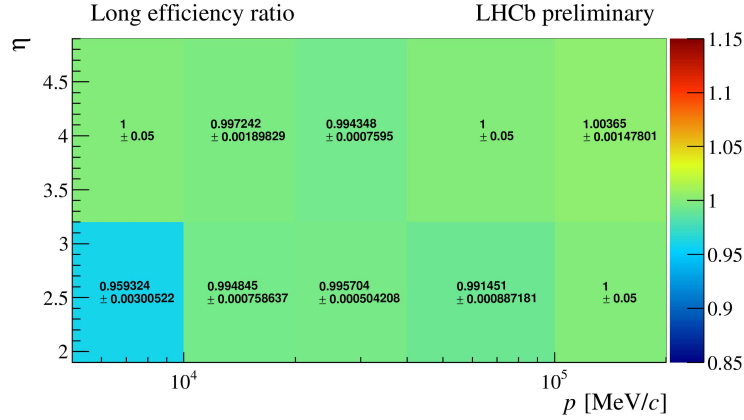


Figure 7.5: Tracking efficiency ratios $\frac{\epsilon_{\text{data}}}{\epsilon_{\text{MC}}}$. For 2016, only the table for the *long* method is available.

There are a few caveats: The efficiency ratios are available only for Sim09b MC version, but in this analysis Sim09k is used. The difference in simulation versions is ignored, given that this analysis measures the *ratio* of branching fractions, it is assumed that the discrepancy would cancel in first order. Additionally, the kinematic binning of TRACKCALIB efficiency ratios, shown in Fig. 7.6, are stricter than the signal selection cuts in this analysis. Most notably for the slow π , about 50% of the tracks falls outside of the binning range of the table for the $D^{*+}\mu^-$ MC decay mode. To circumvent the problem, one method is to apply TRACKCALIB binning range cuts on both data and MC, which leads to a significant loss in selection efficiency. It is decided that in case a track falls outside of the binning range, the efficiency ratio from its closest bin is applied. Any additional discrepancies are expected to be corrected in the *final reweighting* procedure.

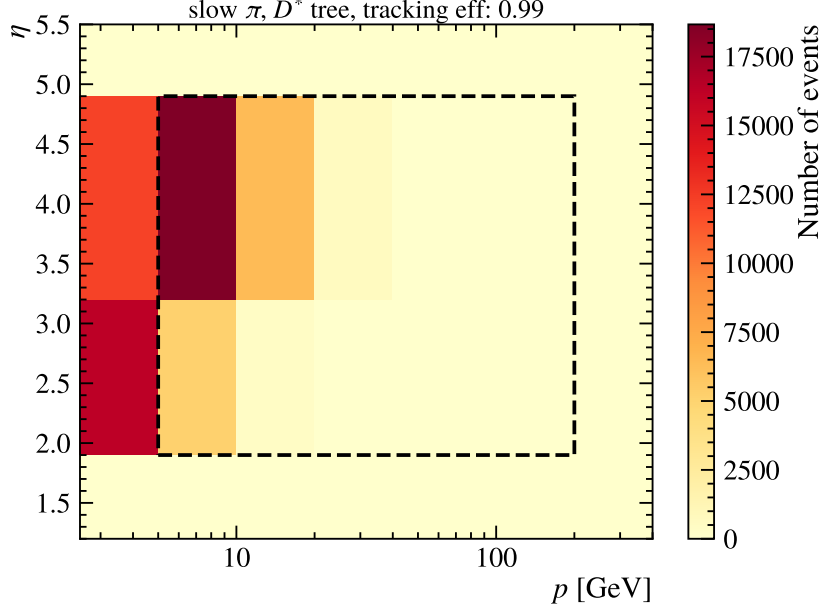


Figure 7.6: For slow π , about 50% of tracks lies outside the binning region of the official TRACKCALIB binning range which is displayed by the dashed black box. The colored rectangles show number of events fall in each official TRACKCALIB bin. To avoid loss in selection efficiencies, correction ratios from nearby bin is applied. The study is done on $\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu$ MC sample.

7.2.2 B kinematics and multiplicity

Another known source of inconsistency between data and MC is the B production kinematics and multiplicity. This is corrected by a separate set of $B^+ \rightarrow J/\psi(\rightarrow \mu^- \mu^+) K^+$ control samples, selected by the procedure described in Section 5.4. Additionally, the track efficiency correction described in the previous section is applied to the $J/\psi K$ MC sample.

To separate the $J/\psi K$ signal from the combinatorial background, a fit, shown in Fig. 7.7, is performed on invariant mass of the B meson (m_B) of $J/\psi K$ data. The fit model consists of a double Gaussian plus a Crystal bell for signal, and an exponential background. An *sPlot* procedure, a fit-based approach to subtract background as described in [41], then computes signal and background weight for each event. Distributions of variables uncorrelated with m_B , for example p_T of B , are produced by applying signal weights (*sWeight*) obtained above on all data events.

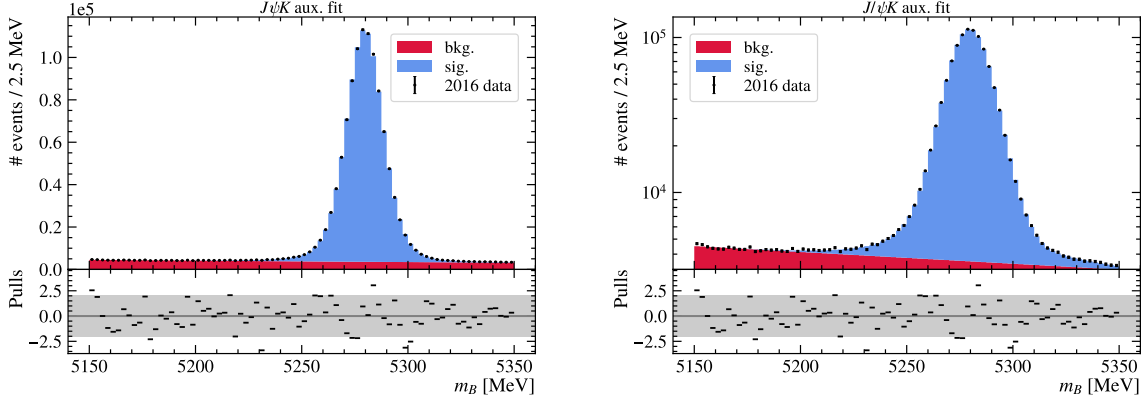


Figure 7.7: Fit on $J/\psi K$ with a *sPlot* procedure. The signal is modelled by a double Gaussian plus a Crystal ball; the background an exponential. Left and right show the same plot, with left having a linear y axis, and right a log y axis.

A multi-stage⁵ reweighting is performed on *sWeight*-ed data and MC (both are normalized, so the reweighting only corrects shape differences), with the stages listed in Table 7.1, and the results shown in Fig. 7.8. The occupancy (nTracks) has to be corrected first before B kinematics because the PID efficiency weights depend on nTracks, and the PID weights alter the shapes of the kinematic distributions of the B decay daughters, which in turn affect the B kinematic distributions themselves. The reweighting binning ranges are chosen such that:

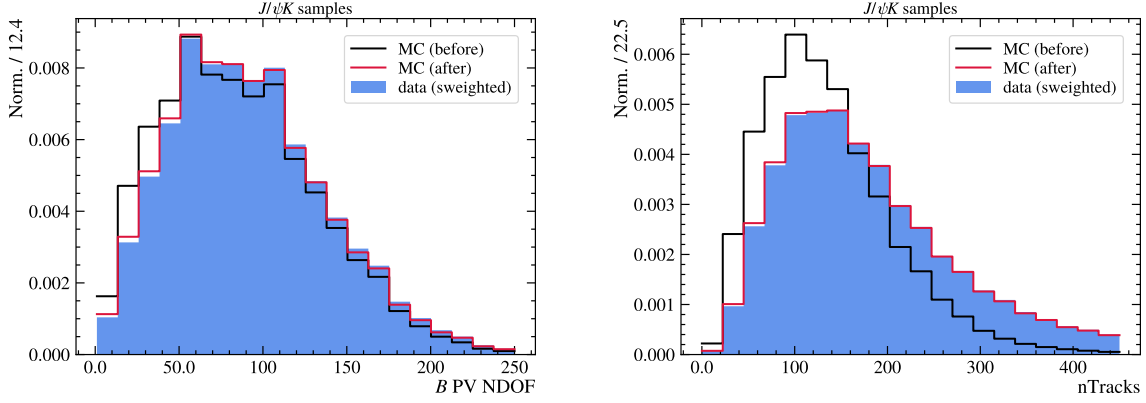
- They cover most of signal ranges of the data after applying *sWeights*.
- The MC have enough population within these binning ranges so that there is no or very few bins without a event.

When applying the $J/\psi K$ -derived weights to the nominal MC sample, for events outside the binning regions, the weights from nearby bins are applied. This is to avoid introducing explicit kinematic cuts on the B meson, which is hard to apply consistently due to the fact that for $B \rightarrow J/\psi K$ decays, there is no missing particle in the decay chain thus the B momentum is reconstructed in a more complete way, whereas for $B \rightarrow D^{(*)} l \bar{\nu}_l$, the missing neutrino(s) makes the reconstructed B momentum incomplete.

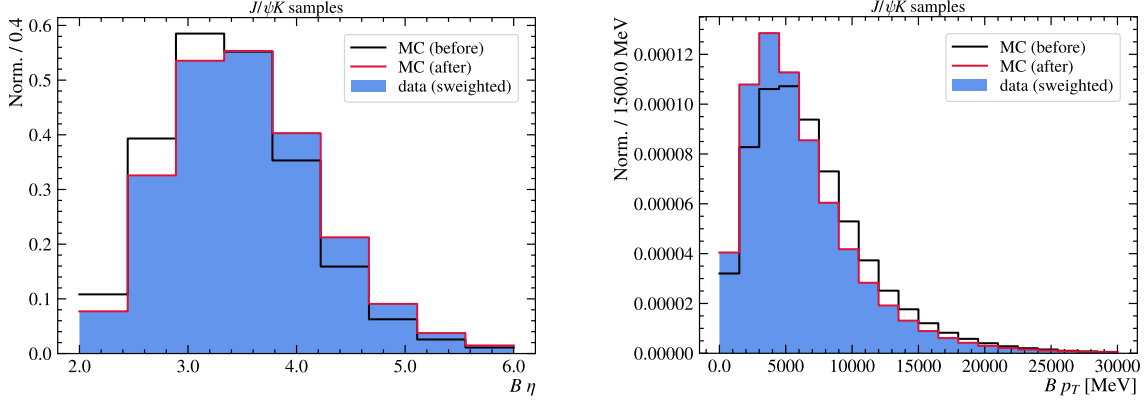
⁵“Multi-stage” means that for reweighting stage i , all previous $i - 1$ weights are applied.

Table 7.1: Reweighting stages and binning schemes for B kinematic and occupancy reweighting.

Variable 1	Binning 1	Variable 2	Binning 2	Figure
B PV ndf	20, 1–250	nTracks	20, 0–450	Fig. 7.8a
B η	9, 2–6	B p_T	20, 0–30 GeV	Fig. 7.8b



(a) Stage-1: B PV ndf and nTracks.



(b) Stage-2: B η and B p_T .

Figure 7.8: Effect of B kinematic and occupancy reweighting, derived from $J/\psi K$ control samples. The “MC (after)” plots contain corrections from *all stages*.

7.3 Post-fit cocktail

To check the data/MC agreement on many variables, a post-fit cocktail, generated from scaling each fit component based its fitted yields, is needed. This is because the selected data samples are themselves cocktails with contributions from many decay modes,

so no meaningful comparison between single decay-mode MC and data can be performed.

After the initial weights are applied to the fit templates, a multi-step fit (termed *initial fit*, identical to the *canonical* fit described in Chapter 8) is performed with initial-reweighted templates to determine the yield and variations of each template. The fit results are displayed in Appendix A.4.

Since the main objective of final reweighting is to correct the remaining inconsistencies between data and MC due to detector responses, it is important to reweight in a region where most discrepancies are induced by such responses. Therefore, a $m_{\text{miss}}^2 < 0.4 \text{ GeV}^2$ region, termed reweighting region, dominated in normalization decays, mostly free of effects due to choice of modeling on background modes, is used to derive the final weights. The post-fit cocktail is built with the following procedure:

1. Apply the $m_{\text{miss}}^2 < 0.4 \text{ GeV}^2$ cut to obtain yields in the reweighting region based on total fitted yields.
2. Some fit templates have variations which are specified by additionally providing variation templates at $\pm 1\alpha$ and an interpolation/extrapolation method. More details can be seen in Section 8.2.

After implementing possible fitted variations, each fit template is rescaled such that the scaled integral of the template matches the fitted yield in the reweighting region.

3. All MC templates are added together to form the *post-fit cocktail*. The data-driven templates are subtracted from the *data* template to form *subtracted data* template which is comparable to the *post-fit cocktail*.

7.4 Final reweighting

The final reweighting is a multi-stage reweighting process implemented by comparing selected kinematic and geometric variables of the B decay daughters, listed in Table 7.2,

between the *post-fit cocktail* and *subtracted data* in the reweighting region, selected by a $m_{\text{miss}}^2 < 0.4 \text{ GeV}^2$ cut, and generate binned $\frac{\text{data}}{\text{MC}}$ ratios as weights for each variable. The reweighting sequence is arranged such that correlated variables are handled in the same stage, and the most important variables are placed in the last stages. The final reweighting is performed separately for the D^0 and D^* channels, as the latter contains additional stages due to corrections to the slow π .

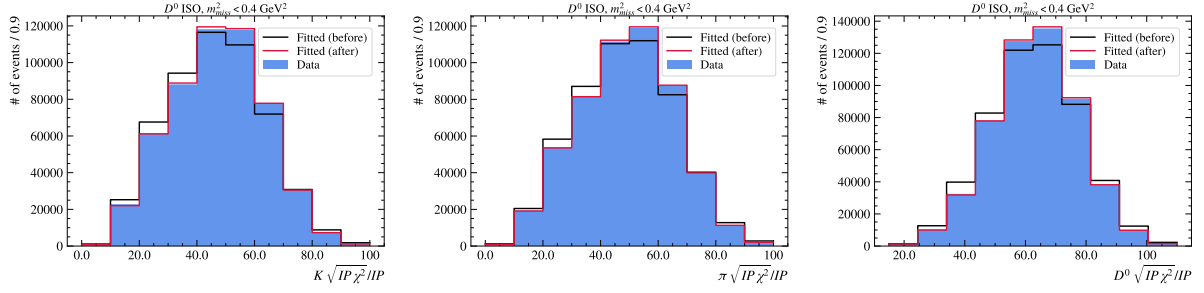
The effects of the final reweighting can be seen in Figs. 7.9 to 7.12 for the D^0 channel, and Figs. 7.13 to 7.16 for the D^* channel. Some important technical details are provided below.

Bins with 0 MC events For bins without MC events, the weights are replaced by 1, to avoid arbitrarily large weights. This is justified by the fact that if the corresponding data bins also have 0 events, no reweighting is needed; otherwise it is likely that large weights will be produced by the reweighting procedure and it is the case we want to avoid.

Treatment of large weights For large weights, defined as $w \geq 50$, the number of events n in the corresponding bin of the *subtracted data* sample is checked: If $n \leq 10$, the large weight is treated as a fluctuation and is set back to 1; otherwise, it is capped at 50. No other weight cap is imposed on the $J/\psi K$ -derived weight at this stage⁶.

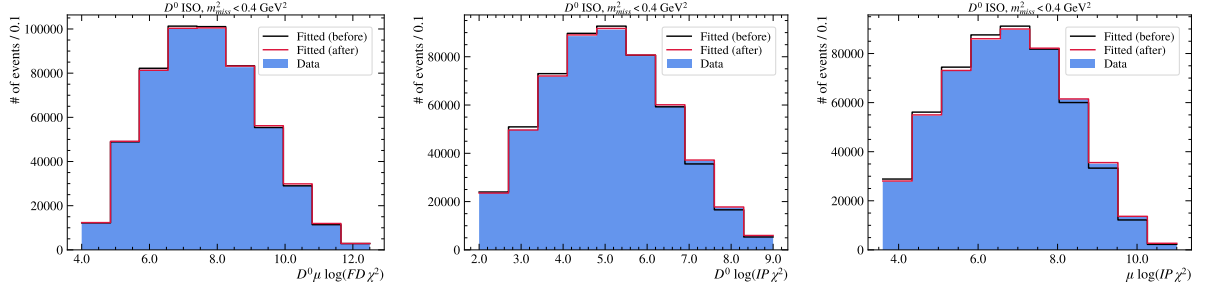
Under and overflow bins The under and overflow bins are treated as regular bins and reweighted normally during the reweighting process.

⁶There are, however, weight caps on the products of weights, which will be described in Section 8.1.

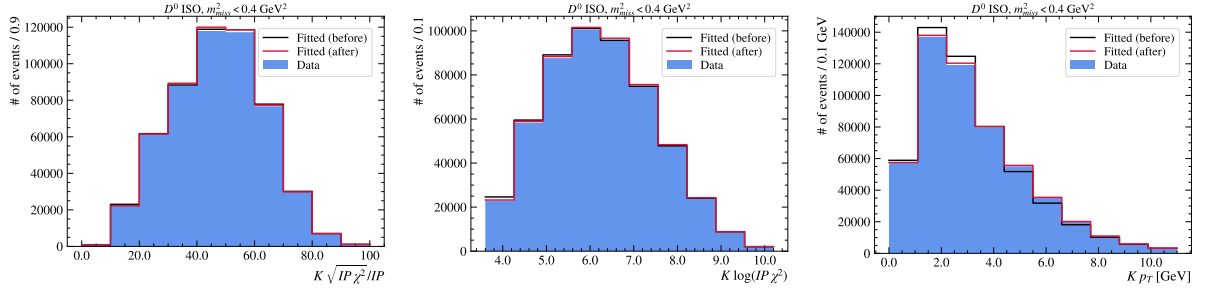


(a) Stage 1.

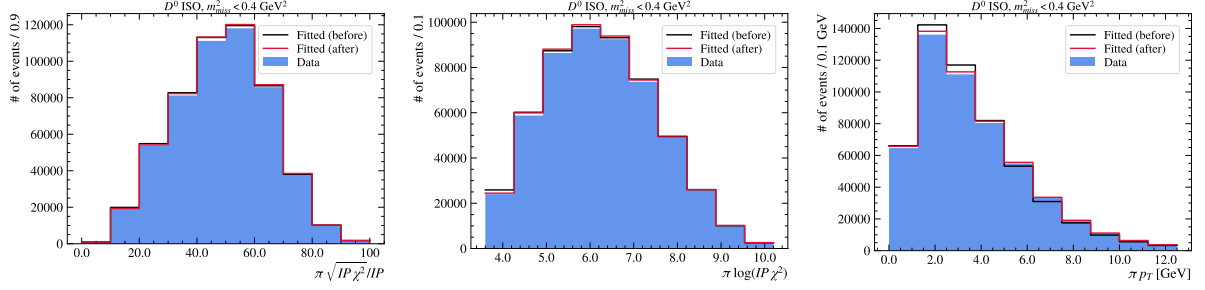
Figure 7.9: Final reweighting for D^0 channel, for a total of 11 stages. The black plots contain weights from *all previous* stages, whereas the red ones contain weights from *all* stages.



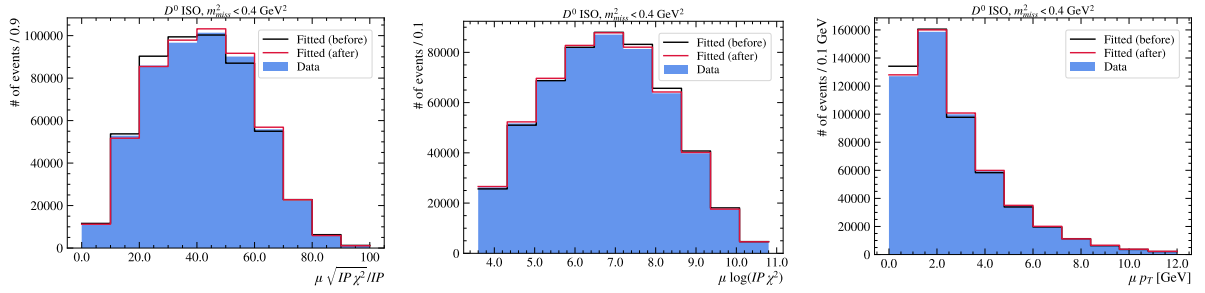
(a) Stage 2.



(b) Stage 3.

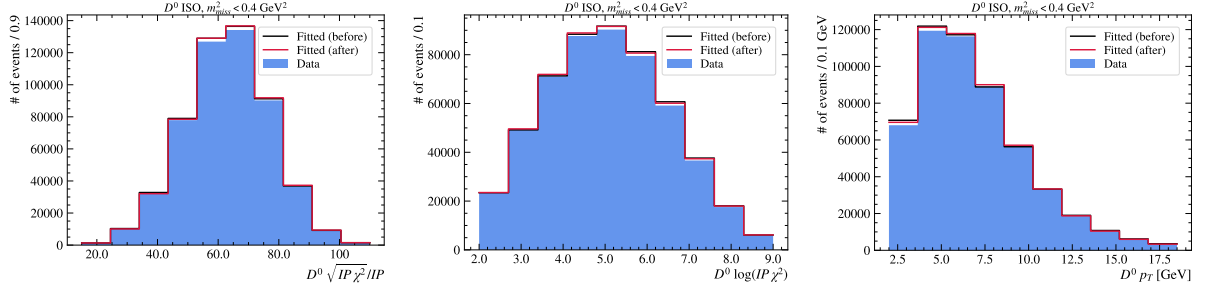


(c) Stage 4.

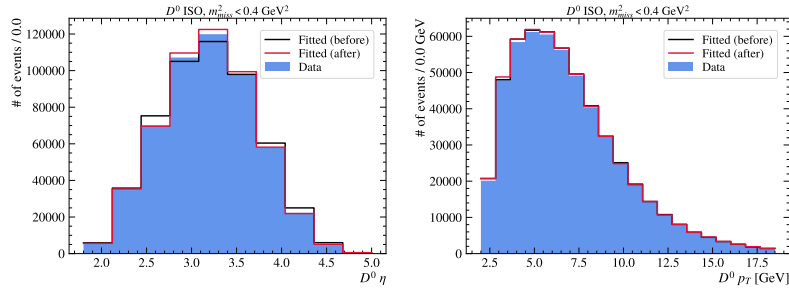


(d) Stage 5.

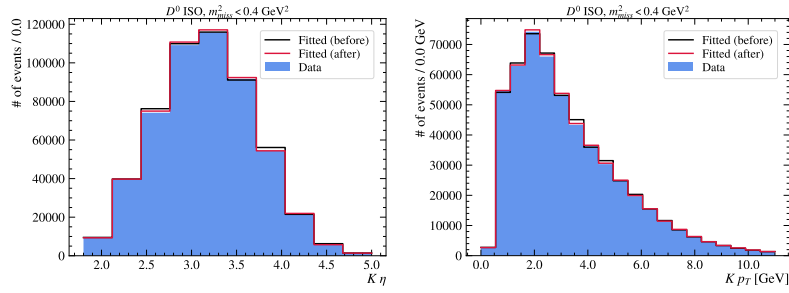
Figure 7.10: Final reweighting for D^0 channel, stages 2–5 (cont'd).



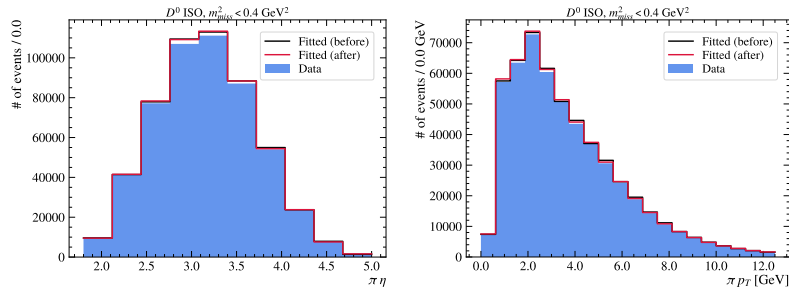
(a) Stage 6.



(b) Stage 7.

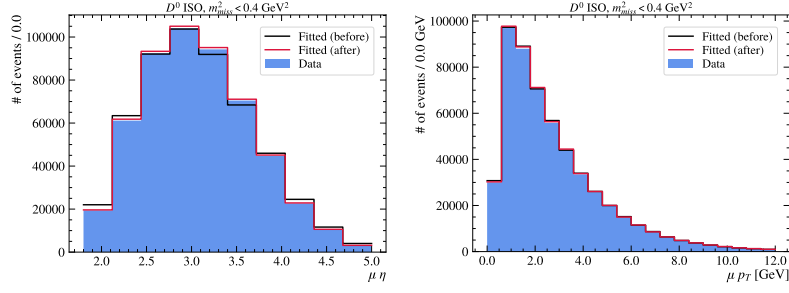


(c) Stage 8.

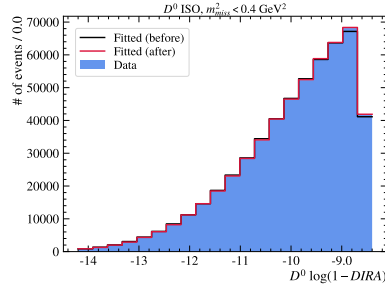


(d) Stage 9.

Figure 7.11: Final reweighting for D^0 channel, stages 6–9 (cont'd).

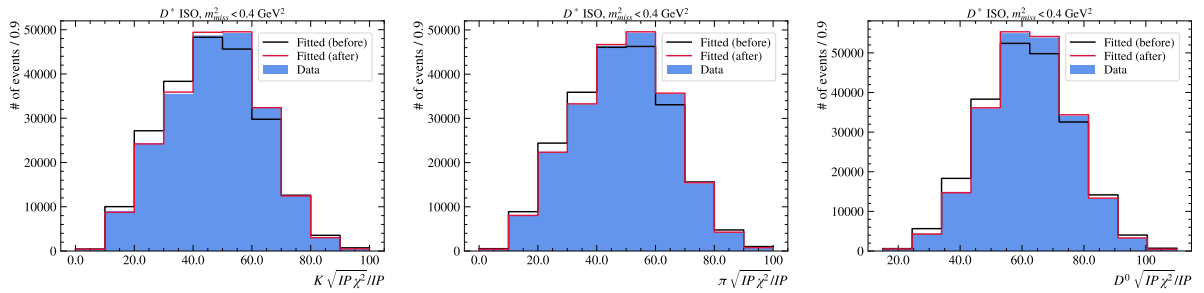


(a) Stage 10.



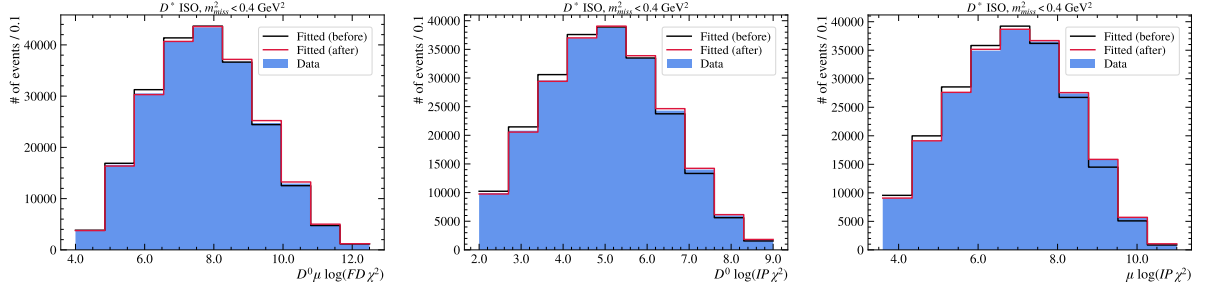
(b) Stage 11.

Figure 7.12: Final reweighting for D^0 channel, stages 10–11 (cont'd).

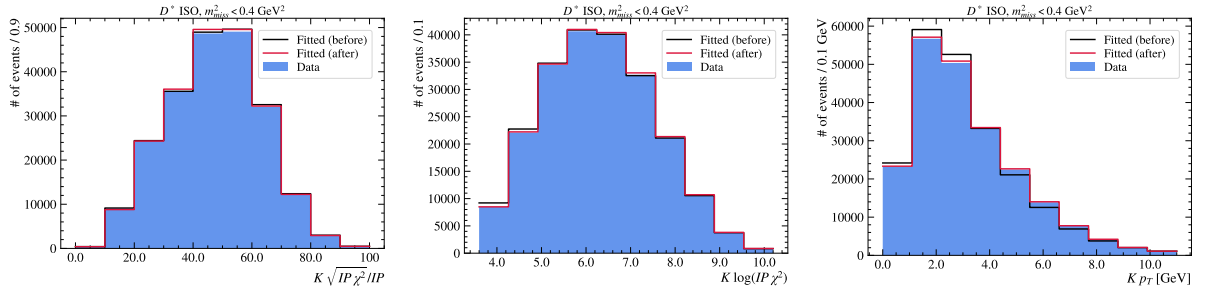


(a) Stage 1.

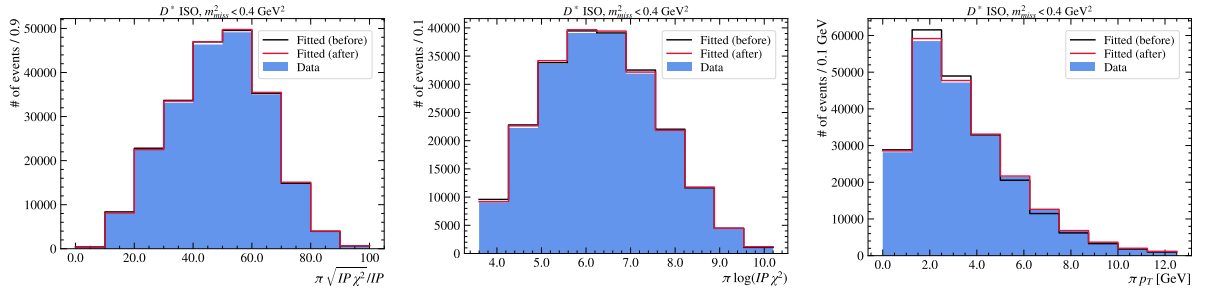
Figure 7.13: Final reweighting for D^* channel, for a total of 13 stages. The black plots contain weights from *all previous* stages, whereas the red ones contain weights from *all* stages.



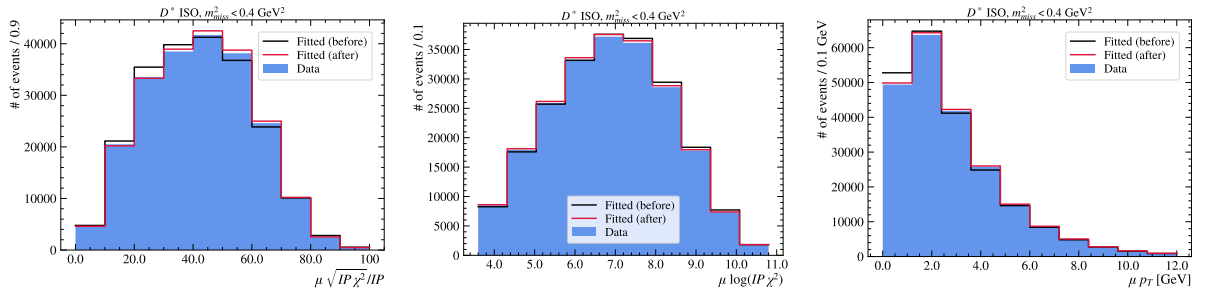
(a) Stage 2.



(b) Stage 3.

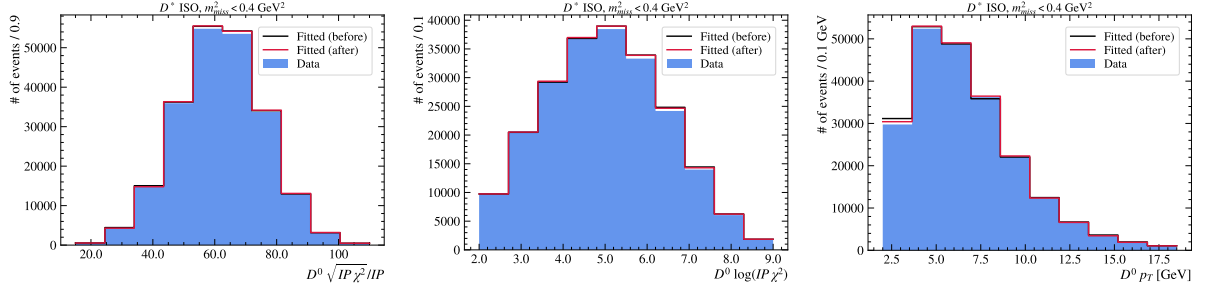


(c) Stage 4.

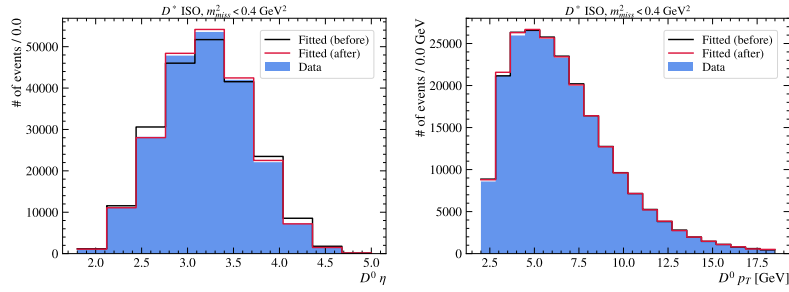


(d) Stage 5.

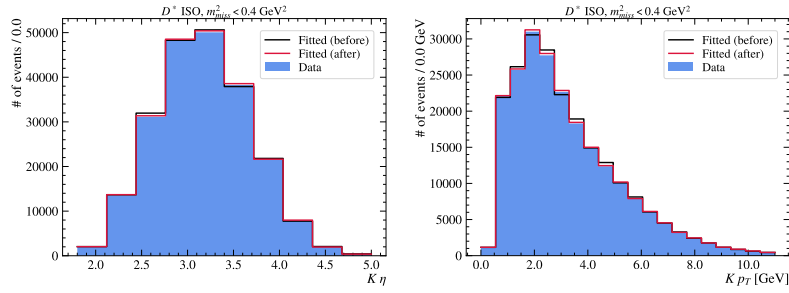
Figure 7.14: Final reweighting for D^* channel, stages 2–5 (cont'd).



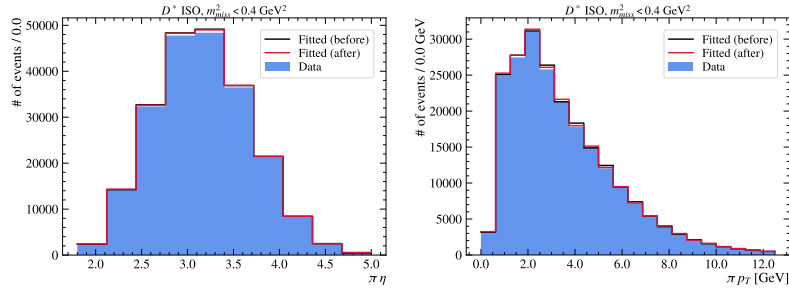
(a) Stage 6.



(b) Stage 7.

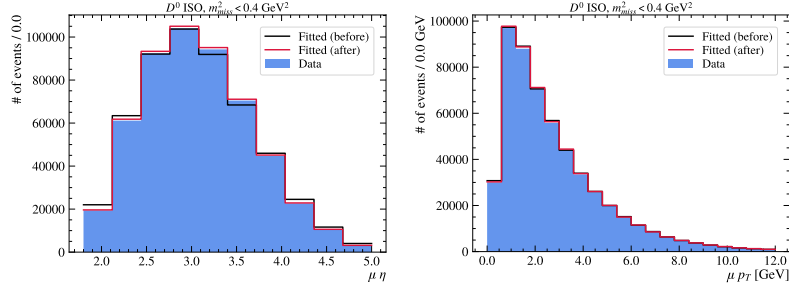


(c) Stage 8.

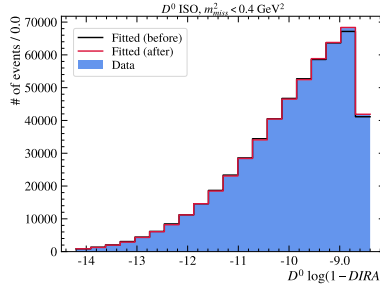


(d) Stage 9.

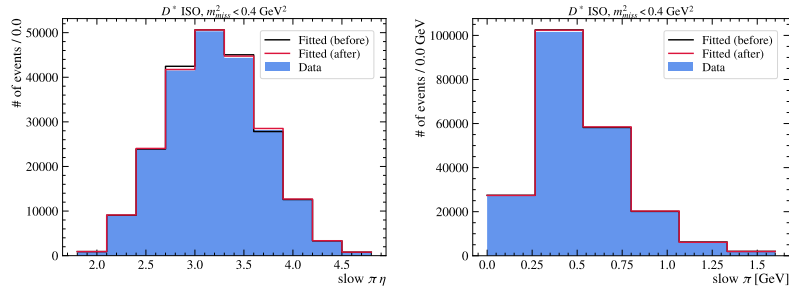
Figure 7.15: Final reweighting for D^* channel, stages 6–9 (cont'd).



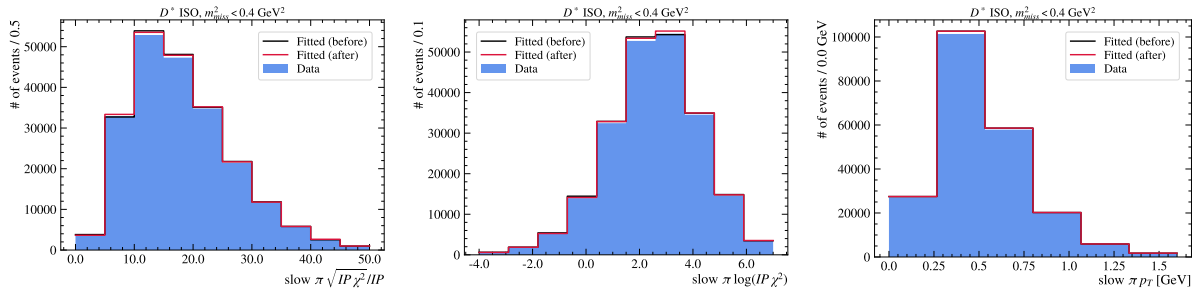
(a) Stage 10.



(b) Stage 11.



(c) Stage 12.



(d) Stage 13.

Figure 7.16: Final reweighting for D^* channel, stages 10–13 (cont'd).

Table 7.2: Reweighting stages and binning schemes for final reweighting.

Variable 1	Binning 1	Variable 2	Binning 2	Variable 3	Binning 3
$K \sqrt{IP \chi^2}/IP$	10, 0 – 100	$\pi \sqrt{IP \chi^2}/IP$	10, 0 – 100	$D^0 \sqrt{IP \chi^2}/IP$	10, 15 – 110
$D^0 \mu \log(FD \chi^2)$	10, 4 – 12.5	$D^0 \log(IP \chi^2)$	10, 2 – 9	$\mu \log(IP \chi^2)$	10, 3.6 – 11
$K p_T$ [GeV]	10, 0 – 11	$K \log(IP \chi^2)$	10, 3.6 – 10.2	$K \sqrt{IP \chi^2}/IP$	10, 5 – 100
πp_T [GeV]	10, 0 – 12.5	$\pi \log(IP \chi^2)$	10, 3.6 – 10.2	$\pi \sqrt{IP \chi^2}/IP$	10, 5 – 100
μp_T [GeV]	10, 0 – 12	$\mu \log(IP \chi^2)$	10, 3.6 – 10.8	$\mu \sqrt{IP \chi^2}/IP$	10, 0 – 100
$D^0 p_T$ [GeV]	10, 2 – 18.5	$D^0 \log(IP \chi^2)$	10, 2 – 9	$D^0 \sqrt{IP \chi^2}/IP$	10, 18 – 102
$D^0 p_T$ [GeV]	20, 2 – 18.5	$D^0 \eta$	10, 1.8 – 5	–	–
$K p_T$ [GeV]	20, 0 – 11	$K \eta$	10, 1.8 – 5	–	–
πp_T [GeV]	20, 0 – 12.5	$\pi \eta$	10, 1.8 – 5	–	–
μp_T [GeV]	20, 0 – 12	$\mu \eta$	10, 1.8 – 5	–	–
$D^0 \log(1 - DIRA)$	20, -14.2 – -8.4	–	–	–	–
slow πp_T [GeV]*	6, 0 – 1.6	slow $\pi \eta$	10, 1.8 – 4.8	–	–
slow πp_T [GeV]*	6, 0 – 1.6	slow $\pi \log(IP \chi^2)$	10, -4 – 7	slow $\pi \sqrt{IP \chi^2}/IP$	10, 0 – 50

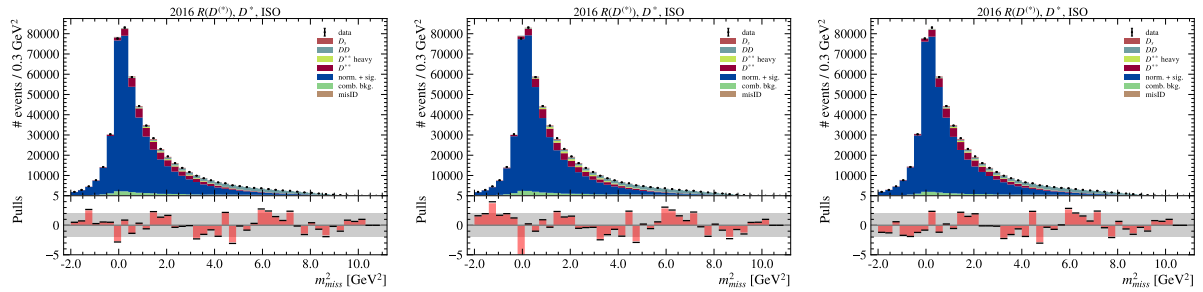
* This is for D^* channel only.

7.5 Additional correction on the B vertex resolution

The B vertex resolution is partly corrected in the final reweighting via the impact parameter (IP) variables. Although final reweighting brings these variables to good agreement between data and MC, after refitting the data, m_{miss}^2 now has worse pulls, as can be seen in Figs. 7.17a and 7.17b. This is attributed to an *incomplete correction* on vertex resolution during final reweighting, as a run 1 study (Ref. [15], Appendix G) showed.

To fully correct the vertex resolution difference between the data and MC, the B flight direction vector, which is the direction from the pp vertex to the B decay vertex, is smeared by an azimuthal angle randomly drawn from an angular distribution derived from data-driven corrections. Currently, the run 1 vertex resolution smearing is applied. This will be updated based on run 2 data at a later date. Nevertheless, this brings m_{miss}^2 back into good agreement, as can be seen in Figs. 7.17b and 7.17c.

The vertex resolution correction is applied *before* the initial reweighting, and the fit variables estimated based on smeared flight vector are used in the fit nominally. It is understood that plots shown in previous sections contain the correction described in this section.



(a) Without final reweighting. (b) With final reweighting only. (c) With both final reweighting and vertex resolution smearing.

Figure 7.17: Fitted m_{miss}^2 distributions in the D^* channel. Without additional vertex resolution correction, the agreement between data and fitted result becomes poorer after final reweighting, as the pulls become worse. Applying a vertex resolution smearing brings m_{miss}^2 back to good agreement. These plots are taken from an older fit configuration and do not represent final fit results.

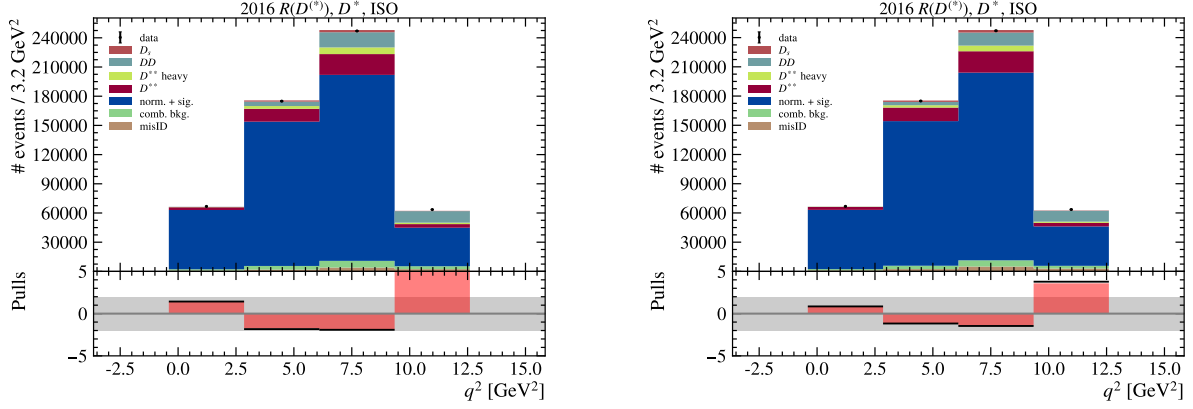
Chapter 8: Signal extraction fit

To extract $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$, the parameters of interest, a simultaneous *signal fit* to both the D^0 and D^* channels of the ISO skim¹ is performed. As briefly discussed in Chapter 2, the parameters responsible for controlling the shapes of the D^{**} and DDX backgrounds, such as the D^{**} form factors, DDX Dalitz-inspired shape variations (will be defined shortly in Section 8.2.2), and the DDX relative fraction correction (the correction to the relative fraction between the $\bar{B}^0 \rightarrow DDX$ and $B^- \rightarrow DDX$ templates), are *not well-constrained* by the signal fit as they constitute only a small fraction of the ISO data skim. As an example, allowing the DDX relative fraction correction to float in the signal fit seemingly improves the fit quality, as shown in Fig. 8.1, but one of $\mathcal{R}(D^{(*)})$ is pushed to 0. This suggests that the DDX (sometimes referred to as DD in the subsequent text) components can completely absorb the signal if their parameters are not measured in some other way.

Therefore, these background shape parameters are obtained with a *control fit*, a simultaneous fit to the 1OS, 2OS, and DD control skims with six fit channels (a D^0 and a D^* channel for each control skim), and imported back to the signal fit at the fitted values with either fitted errors as constraints (e.g. for D^{**} shape parameters) or fully fixed (e.g. for DDX shape parameters). To be more precise, the fit procedure is broken into 3 steps:

1. **Pre-control:** Performed on the control skims (1OS, 2OS, DD). In this step, all variations but the K^* weight (Section 8.2.3) in the DDX templates are turned off.

¹As a reminder, “skim” means *subsample*. The selection criteria for all four skims (ISO, 1OS, 2OS, DD) is defined at Section 5.3.



(a) The nominal fit. The pull in the last q^2 bin is above 5σ .

(b) The fit with $\rho_{DDX,u/d}$ floating. Smaller pull for each q^2 bin.

Figure 8.1: Allowing the DDX relative fraction $\rho_{DDX,u/d}$ to float in the signal fit seemingly improves fit quality (smaller pull for each q^2 bin), but the signal are now completely absorbed in the DDX components. Only the D^* channel fit is shown as it shows the most improvements in terms of pulls.

2. **Control:** Also performed on the control skims (1OS, 2OS, DD). All but the signal and normalization form factor variations are turned on. The fitted parameters from previous steps are loaded along with their errors as a starting point for the control fit.
3. **Signal:** Performed on the signal skim (ISO). The D^{**} form factors, the D_H^{**} data-driven corrections, the K^* weight in the DDX templates, and the misID DiF weight (Section 8.2.5) are loaded as constraints from the control fit; the DDX data-driven shape corrections and the DDX relative fraction correction are fixed to the control fit values.

Both the signal and the control fit are binned extended maximum likelihood template fits, with a fitter based on the ROOT HISTFACTORY fitter used in the pathfinder run 1 $\mathcal{R}(D^{(*)})$ analysis described in [15]². The fit templates are three-dimensional correlation-preserving histograms of m_{miss}^2 , E_μ^* , and q^2 , with the binning scheme listed in Table 8.1. Events outside the binning ranges of the fit variables are cut out to avoid potential impact of under/overflow

²With code refactored and dependencies updated to enhance readability and configurability, but the fit model is largely the same.

bins on the fitter. To account for external constraints on fit parameters or uncertainties on modeling of certain backgrounds, additional shape variations are introduced on most templates (i.e. all but the D^* combinatorial templates have some shape variations which will be discussed later).

Table 8.1: Fit variables and their binnings, obtained by the rest frame approximation technique as described in Appendix C.

Variable	Binning
m_{miss}^2 [GeV ²]	43, -2 – 10.9
E_μ^* [MeV]	34, 100 – 2650
q^2 [GeV ²]	4, -0.4 – 12.6

The remainder of this chapter is organized as following: a description on template generation, followed by a discussion on all components of fit templates, is described in Section 8.1. The fit template shape variations are explained in Section 8.2. The blinded fitted results are provided in Section 8.3.

8.1 Fit templates

This section describes the generation of the fit templates, followed by a distribution of the characteristics of the templates and the normalization factors controlling the yields of the templates. There are a total of 37 templates in the D^0 channel and 24 in the D^* channel. For a brief overview of the templates, Tables 8.2 and 8.3 may be used. The normalization factors for the signal fit are summarized in Tables 8.4 and 8.5 for the D^0 and D^* channel accordingly and will be documented in the next subsections. These normalization factors are color-coded as:

- **black**: an unconstrained factor, implemented with a `HISTFACTORY NormFactor`.
- **red**: a fixed `NormFactor`.
- **blue**: a factor with a Gaussian constraint where the nominal value is 1 and the standard

deviation σ . The fitter reports the *variation* of the factor with an α parameter such that when the underlying parameter is at 1, the associated $\alpha = 0$; when the underlying parameter is at $1 \pm \sigma$, the $\alpha = \pm 1$. Implemented with a HISTFACTORY `OverallSys`.

- **magenta**: a fixed `OverallSys` factor. Typically such a factor is *constrained* in the control fit and loaded as a *fixed* factor in the signal fit.

The data-driven templates, namely μ misID and combinatorial background templates, are discussed first in Sections 8.1.1 and 8.1.2. They are used to estimate both the shapes and the yields of these backgrounds.

The rest are all MC templates which can be categorized into signal, normalization, feed downs from $D^{**}/\text{heavy } D^{**}/D_s^{**}$, and contributions from DDX decay modes. Only the shapes (i.e. the distributions of the fit variables) of these templates are taken. All categories share the same selection requirement as discussed in Section 5.2. Additional global weights to account for emulation of trigger and PID (Chapter 6), form factor reweighting (Section 7.1, whenever applicable), and corrections to detector responses (Sections 7.2 and 7.4) are also applied according to Eq. (8.1), with the weight-capping strategy specified as:

$$w_{\text{tot}} = \underbrace{\left(w_{\text{trigger}} \cdot w_{\text{PID}} \cdot \overbrace{w_{\text{tracking}} \cdot w_{J/\psi K}}^{\text{initial weight}} \right)}_{\text{capped at 10}} \times \underbrace{w_{\text{form factor}}}_{\substack{\text{capped at} \\ 50 \text{ for } D^0 \text{ and } D^* \\ 10 \text{ for } D^{**}}} \times \underbrace{\prod_i w_{\text{final weight, step } i}}_{\text{capped at 10}} \quad (8.1)$$

These weight caps limit the impact of statistical fluctuations in bins with low event count.

8.1.1 Muon misID backgrounds

The μ misID fit templates, shown in Fig. 8.2, represent contributions from non- μ tracks satisfying the μ identification requirements in “real” μ samples. For *all* data-driven fit templates, misID included, their yields are defined as following:

$$\rho_{\text{data-driven}} \times N_{\text{data-driven}} \quad (8.2)$$

where $N_{\text{data-driven}}$ represents the yield to be fitted, with a Gaussian constraint $\rho_{\text{data-driven}}$

Table 8.2: Fit templates included in the D^0 channel.

Decay mode	Run 2 process	Alias in fitter	Variations	Index
$B^- \rightarrow D^0 \mu^- \bar{\nu}_\mu$	D0Mu	D_Dmu	5	1
$\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu$	DstMu	D_dDstmu	10	2
$B^- \rightarrow D^{*0} \mu^- \bar{\nu}_\mu$	Dst0Mu	D_uDstmu	10	3
$B^- \rightarrow D^0 \tau^- \bar{\nu}_\tau$	D0Tau	D_Dtau	5	4
$\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau$	DstTau	D_dDsttau	10	5
$B^- \rightarrow D^{*0} \tau^- \bar{\nu}_\tau$	Dst0Tau	D_uDsttau	10	6
$\bar{B}^0 \rightarrow D_1 \mu \bar{\nu}_\mu$	D1ststMu	D_dD1mu	3	7
$\bar{B}^0 \rightarrow D_1(\rightarrow D^0 \pi \pi) \mu \bar{\nu}_\mu$	D1ststMuD0PiPi	D_dD1mu_pipi	3	8
$\bar{B}^0 \rightarrow D_2^* \mu \bar{\nu}_\mu$	D2ststMu	D_dD2mu	3	9
$\bar{B}^0 \rightarrow D_1' \mu \bar{\nu}_\mu$	D1pststMu	D_dD1pmu	2	10
$\bar{B}^0 \rightarrow D_0^* \mu \bar{\nu}_\mu$	D0ststMu	D_dD0mu	2	11
$\bar{B} \rightarrow D^{**}(\rightarrow D^{*0} \pi \pi) \mu \bar{\nu}_\mu$	DststHMuDst0	D_Dstzpipimu	1	12
$\bar{B} \rightarrow D^{**}(\rightarrow D^* \pi \pi) \mu \bar{\nu}_\mu$	DststHMuDst	D_Dstppipimu	1	13
$\bar{B} \rightarrow D^{**}(\rightarrow D^0 \pi \pi) \mu \bar{\nu}_\mu$	DststHMuD0	D_Dpipimu	1	14
$B^- \rightarrow D_1^0 \mu \bar{\nu}_\mu$	D1stst0Mu	D_uD1mu	3	15
$B^- \rightarrow D_1^0(\rightarrow D^0 \pi \pi) \mu \bar{\nu}_\mu$	D1stst0MuD0PiPi	D_uD1mu_pipi	3	16
$B^- \rightarrow D_2^{*0} \mu \bar{\nu}_\mu$	D2stst0Mu	D_uD2mu	3	17
$B^- \rightarrow D_1'^0 \mu \bar{\nu}_\mu$	D1pstst0Mu	D_uD1pmu	2	18
$B^- \rightarrow D_0^{*0} \mu \bar{\nu}_\mu$	D0stst0Mu	D_uD0mu	2	19
$\bar{B}_s \rightarrow D_{s2}^* \mu \bar{\nu}_\mu$	Ds2Mu	D_sDs2mu	3	20
$\bar{B}_s \rightarrow D_{s1}' \mu \bar{\nu}_\mu$	Ds1pMu	D_sDs1pmu	3	21
$\bar{B}^0 \rightarrow D_1 \tau \bar{\nu}_\tau$	D1ststTau	D_dD1tau	3	22
$\bar{B}^0 \rightarrow D_1(\rightarrow D^0 \pi \pi) \tau \bar{\nu}_\tau$	D1ststTauD0PiPi	D_dD1tau_pipi	3	23
$\bar{B}^0 \rightarrow D_2^* \tau \bar{\nu}_\tau$	D2ststTau	D_dD2tau	3	24
$\bar{B}^0 \rightarrow D_1' \tau \bar{\nu}_\tau$	D1pststTau	D_dD1ptau	2	25
$\bar{B}^0 \rightarrow D_0^* \tau \bar{\nu}_\tau$	D0ststTau	D_dD0tau	2	26
$B^- \rightarrow D_1^0 \tau \bar{\nu}_\tau$	D1stst0Tau	D_uD1tau	3	27
$B^- \rightarrow D_1^0(\rightarrow D^0 \pi \pi) \tau \bar{\nu}_\tau$	D1stst0TauD0PiPi	D_uD1tau_pipi	3	28
$B^- \rightarrow D_2^{*0} \tau \bar{\nu}_\tau$	D2stst0Tau	D_uD2tau	3	29
$B^- \rightarrow D_1'^0 \tau \bar{\nu}_\tau$	D1pstst0Tau	D_uD1ptau	2	30
$B^- \rightarrow D_0^{*0} \tau \bar{\nu}_\tau$	D0stst0Tau	D_uD0tau	2	31
$\bar{B}^0 \rightarrow D^0 D_q(\rightarrow \mu \bar{\nu}_\mu X') X$	dDDMu	D_dDDmu	3	32
$B^- \rightarrow D^0 D_q(\rightarrow \mu \bar{\nu}_\mu X') X$	uDDMu	D_uDDmu	3	33
$\bar{B}^0 \rightarrow D^0 D_q(\rightarrow \tau \bar{\nu}_\tau X') X$	dDDTau	D_dDDtau	0	34
$B^- \rightarrow D^0 D_q(\rightarrow \tau \bar{\nu}_\tau X') X$	uDDTau	D_uDDtau	0	35
B^- comb. bkg.	BComb	D_comb	1	36
misID.	misID	D_misID	1	37

Table 8.3: Fit templates included in the D^* channel.

Decay mode	Run 2 process	Alias in fitter	Variations	Index
$\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu$	DstMu	Dst_sigmu	10	1
$\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau$	DstTau	Dst_sigtau	10	2
$\bar{B}^0 \rightarrow D_1 \mu \bar{\nu}_\mu$	D1ststMu	Dst_D1	3	3
$\bar{B}^0 \rightarrow D_2^* \mu \bar{\nu}_\mu$	D2ststMu	Dst_D2	3	4
$\bar{B}^0 \rightarrow D_1' \mu \bar{\nu}_\mu$	D1pststMu	Dst_D1p	2	5
$\bar{B} \rightarrow D^{*+} (\rightarrow D^* \pi \pi) \mu \bar{\nu}_\mu$	DststHMuDst	Dst_D2Smu	1	6
$B^- \rightarrow D_1^0 \mu \bar{\nu}_\mu$	D1stst0Mu	Dst_uD1	3	7
$B^- \rightarrow D_2^{*0} \mu \bar{\nu}_\mu$	D2stst0Mu	Dst_uD2	3	8
$B^- \rightarrow D_1'^0 \mu \bar{\nu}_\mu$	D1pstst0Mu	Dst_uD1p	2	9
$\bar{B}_s \rightarrow D_{s2}^* \mu \bar{\nu}_\mu$	Ds2Mu	Dst_Ds2	3	10
$\bar{B}_s \rightarrow D_{s1}' \mu \bar{\nu}_\mu$	Ds1pMu	Dst_Ds1p	3	11
$\bar{B}^0 \rightarrow D_1 \tau \bar{\nu}_\tau$	D1ststTau	Dst_D1tau	3	12
$\bar{B}^0 \rightarrow D_2^* \tau \bar{\nu}_\tau$	D2ststTau	Dst_D2tau	3	13
$\bar{B}^0 \rightarrow D_1' \tau \bar{\nu}_\tau$	D1pststTau	Dst_D1ptau	2	14
$B^- \rightarrow D_1^0 \tau \bar{\nu}_\tau$	D1stst0Tau	Dst_uD1tau	3	15
$B^- \rightarrow D_2^{*0} \tau \bar{\nu}_\tau$	D2stst0Tau	Dst_uD2tau	3	16
$B^- \rightarrow D_1'^0 \tau \bar{\nu}_\tau$	D1pstst0Tau	Dst_uD1ptau	2	17
$\bar{B}^0 \rightarrow D^* D_q (\rightarrow \mu \bar{\nu}_\mu X') X$	dDDMu	Dst_dDDmu	3	18
$B^- \rightarrow D^* D_q (\rightarrow \mu \bar{\nu}_\mu X') X$	uDDMu	Dst_uDDmu	3	19
$\bar{B}^0 \rightarrow D^* D_q (\rightarrow \tau \bar{\nu}_\tau X') X$	dDDTau	Dst_dDDtau	0	20
$B^- \rightarrow D^* D_q (\rightarrow \tau \bar{\nu}_\tau X') X$	uDDTau	Dst_uDDtau	0	21
$\bar{B}^0$ comb. bkg.	BComb	Dst_comb	1	22
misID.	misID	Dst_misID	1	23
D^* comb. bkg.	DstComb	Dst_doug	0	24

Table 8.4: Normalization factors for D^0 signal fit with ISO skim templates. Color coding: red: fixed NormFactor; blue: constrained OverallSys; magenta: fixed OverallSys loaded from control fit.

Alias	Decay mode	Normalization	Index
D_Dmu	$B^- \rightarrow D^0 \mu^- \bar{\nu}_\mu$	$N_{D\mu}$	1
D_dDstmu	$\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu$	$N_{D\mu} \times r_{D^{*+}}^{\text{isospin}} \times r_{D^{*0}}$	2
D_uDstmu	$B^- \rightarrow D^{*0} \mu^- \bar{\nu}_\mu$	$N_{D\mu} \times r_{D^{*0}}^0$	3
D_Dtau	$B^- \rightarrow D^0 \tau^- \bar{\nu}_\tau$	$N_{D\mu} \times \eta_{D^0} \times \mathcal{R}(D)$	4
D_dDsttau	$\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau$	$N_{D\mu} \times \frac{\eta_{D^{*+}}}{\eta_{D^{*+}}} \times \tilde{\eta}_{D^{*+}} \times \mathcal{R}(D^*) \times r_{D^{*0}}^0 \times r_{D^{*+}}^{\text{isospin}}$	5
D_uDsttau	$B^- \rightarrow D^{*0} \tau^- \bar{\nu}_\tau$	$N_{D\mu} \times \frac{\eta_{D^{*0}}}{\eta_{D^{*+}}} \times \tilde{\eta}_{D^{*+}} \times \mathcal{R}(D^*) \times r_{D^{*0}}^0$	6
D_dD1mu	$\bar{B}^0 \rightarrow D_1 \mu \bar{\nu}_\mu$	$\rho_{D_1}^{\text{isospin}} \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_1}^0 \times \epsilon_{D_1} \times \mathcal{B}_{D_1}^0$	7
D_dD1mu_pipi	$\bar{B}^0 \rightarrow D_1(\rightarrow D^0 \pi \pi) \mu \bar{\nu}_\mu$	$\rho_{D_1 \pi \pi}^{\text{isospin}} \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_1}^0 \times \epsilon_{D_1 \pi \pi} \times \mathcal{B}_{D_1 \pi \pi}^0$	8
D_dD2mu	$\bar{B}^0 \rightarrow D_2^* \mu \bar{\nu}_\mu$	$\rho_{D_2^*}^{\text{isospin}} \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_2^*}^0 \times \epsilon_{D_2^*} \times \mathcal{B}_{D_2^*}^0$	9
D_dD1pmu	$\bar{B}^0 \rightarrow D_1' \mu \bar{\nu}_\mu$	$\rho_{D_1'}^{\text{isospin}} \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_1'}^0 \times \epsilon_{D_1'} \times \mathcal{B}_{D_1'}^0$	10
D_dD0mu	$\bar{B}^0 \rightarrow D_0^* \mu \bar{\nu}_\mu$	$\rho_{D_1^*}^{\text{isospin}} \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_1^*}^0 \times \epsilon_{D_1^*} \times \mathcal{B}_{D_1^*}^0$	11
D_Dstzpipimu	$\bar{B} \rightarrow D^{**}(\rightarrow D^{*0} \pi \pi) \mu \bar{\nu}_\mu$	$n_{D^{*+}}^0 \times N_{D\mu} \times f_{\text{guess}} \times f_{D^{*0} \pi \pi}^0$	12
D_Dstppipimu	$\bar{B} \rightarrow D^{**}(\rightarrow D^* \pi \pi) \mu \bar{\nu}_\mu$	$n_{D^{*+}}^0 \times N_{D\mu} \times f_{\text{guess}} \times f_{D^{*+} \pi \pi}^0$	13
D_Dpipimu	$\bar{B} \rightarrow D^{**}(\rightarrow D^0 \pi \pi) \mu \bar{\nu}_\mu$	$n_{D^{*+}}^0 \times N_{D\mu} \times f_{\text{guess}} \times f_{D^0 \pi \pi}^0$	14
D_uD1mu	$B^- \rightarrow D_1^0 \mu \bar{\nu}_\mu$	$\rho_{D_1}^{\text{isospin}} \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_1^0}^0 \times \mathcal{B}_{D_1}^0$	15
D_uD1mu_pipi	$B^- \rightarrow D_1^0(\rightarrow D^0 \pi \pi) \mu \bar{\nu}_\mu$	$\rho_{D_1 \pi \pi}^{\text{isospin}} \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_1^0}^0 \times \mathcal{B}_{D_1 \pi \pi}^0$	16
D_uD2mu	$B^- \rightarrow D_2^{*0} \mu \bar{\nu}_\mu$	$\rho_{D_2^*}^{\text{isospin}} \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_2^{*0}}^0 \times \mathcal{B}_{D_2^*}^0$	17
D_uD1pmu	$B^- \rightarrow D_1'^0 \mu \bar{\nu}_\mu$	$\rho_{D_1'}^{\text{isospin}} \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_1'^0}^0 \times \mathcal{B}_{D_1'}^0$	18
D_uD0mu	$B^- \rightarrow D_0^{*0} \mu \bar{\nu}_\mu$	$\rho_{D_1^*}^{\text{isospin}} \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_1^{*0}}^0 \times \mathcal{B}_{D_1^*}^0$	19
D_sDs2mu	$\bar{B}_s \rightarrow D_{s2}^* \mu \bar{\nu}_\mu$	$\mathcal{B}_{D_{s2}^*}^* \times \frac{f_s}{f_d} \times n_{D^{*+}}^0 \times N_{s2}^0 \times N_{D\mu}$	20
D_sDs1pmu	$\bar{B}_s \rightarrow D_{s1}' \mu \bar{\nu}_\mu$	$\mathcal{B}_{D_{s1}'}^* \times \frac{f_s}{f_d} \times N_{D\mu} \times n_{D^{*+}}^0 \times N_{s1'}^0$	21
D_dD1tau	$\bar{B}^0 \rightarrow D_1 \tau \bar{\nu}_\tau$	$\rho_{D_1}^{\text{isospin}} \times \rho_{\mathcal{R}(D^{**})}^0 \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_1}^0 \times \epsilon_{D_1} \times \mathcal{B}_{D_1}^0 \times \mathcal{R}(D^{**})_{\text{avg}}^{\text{raw}} \times r(D_1)$	22
D_dD1tau_pipi	$\bar{B}^0 \rightarrow D_1(\rightarrow D^0 \pi \pi) \tau \bar{\nu}_\tau$	$\rho_{D_1 \pi \pi}^{\text{isospin}} \times \rho_{\mathcal{R}(D^{**})}^0 \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_1}^0 \times \epsilon_{D_1 \pi \pi} \times \mathcal{B}_{D_1 \pi \pi}^0 \times \mathcal{R}(D^{**})_{\text{avg}}^{\text{raw}} \times r(D_1)$	23
D_dD2tau	$\bar{B}^0 \rightarrow D_2^* \tau \bar{\nu}_\tau$	$\rho_{D_2^*}^{\text{isospin}} \times \rho_{\mathcal{R}(D^{**})}^0 \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_2^*}^0 \times \epsilon_{D_2^*} \times \mathcal{B}_{D_2^*}^0 \times \mathcal{R}(D^{**})_{\text{avg}}^{\text{raw}} \times r(D_2^*)$	24
D_dD1ptau	$\bar{B}^0 \rightarrow D_1' \tau \bar{\nu}_\tau$	$\rho_{D_1'}^{\text{isospin}} \times \rho_{\mathcal{R}(D^{**})}^0 \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_1'}^0 \times \epsilon_{D_1'} \times \mathcal{B}_{D_1'}^0 \times \mathcal{R}(D^{**})_{\text{avg}}^{\text{raw}} \times r(D_1')$	25
D_dD0tau	$\bar{B}^0 \rightarrow D_0^* \tau \bar{\nu}_\tau$	$\rho_{D_1^*}^{\text{isospin}} \times \rho_{\mathcal{R}(D^{**})}^0 \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_1^*}^0 \times \epsilon_{D_1^*} \times \mathcal{B}_{D_1^*}^0 \times \mathcal{R}(D^{**})_{\text{avg}}^{\text{raw}} \times r(D_1^*)$	26
D_uD1tau	$B^- \rightarrow D_1^0 \tau \bar{\nu}_\tau$	$\rho_{D_1}^{\text{isospin}} \times \rho_{\mathcal{R}(D^{**})}^0 \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_1^0}^0 \times \mathcal{B}_{D_1}^0 \times \mathcal{R}(D^{**})_{\text{avg}}^{\text{raw}} \times r(D_1)$	27
D_uD1tau_pipi	$B^- \rightarrow D_1^0(\rightarrow D^0 \pi \pi) \tau \bar{\nu}_\tau$	$\rho_{D_1 \pi \pi}^{\text{isospin}} \times \rho_{\mathcal{R}(D^{**})}^0 \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_1^0}^0 \times \mathcal{B}_{D_1 \pi \pi}^0 \times \mathcal{R}(D^{**})_{\text{avg}}^{\text{raw}} \times r(D_1)$	28
D_uD2tau	$B^- \rightarrow D_2^{*0} \tau \bar{\nu}_\tau$	$\rho_{D_2^*}^{\text{isospin}} \times \rho_{\mathcal{R}(D^{**})}^0 \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_2^{*0}}^0 \times \mathcal{B}_{D_2^*}^0 \times \mathcal{R}(D^{**})_{\text{avg}}^{\text{raw}} \times r(D_2^*)$	29
D_uD1ptau	$B^- \rightarrow D_1'^0 \tau \bar{\nu}_\tau$	$\rho_{D_1'}^{\text{isospin}} \times \rho_{\mathcal{R}(D^{**})}^0 \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_1'^0}^0 \times \mathcal{B}_{D_1'}^0 \times \mathcal{R}(D^{**})_{\text{avg}}^{\text{raw}} \times r(D_1')$	30
D_uD0tau	$B^- \rightarrow D_0^{*0} \tau \bar{\nu}_\tau$	$\rho_{D_1^*}^{\text{isospin}} \times \rho_{\mathcal{R}(D^{**})}^0 \times n_{D^{*+}}^0 \times N_{D\mu} \times f_{D_1^{*0}}^0 \times \mathcal{B}_{D_1^*}^0 \times \mathcal{R}(D^{**})_{\text{avg}}^{\text{raw}} \times r(D_1^*)$	31
D_dDDmu	$\bar{B}^0 \rightarrow D^0 D_q(\rightarrow \mu \bar{\nu}_\mu X') X$	$\rho_{DDX, u/d}^0 \times f_{DDX}^0 \times N_{D\mu} \times f_{DDX, d}^0$	32
D_uDDmu	$B^- \rightarrow D^0 D_q(\rightarrow \mu \bar{\nu}_\mu X') X$	$\rho_{DDX, u/d}^0 \times f_{DDX}^0 \times N_{D\mu} \times f_{DDX, u}^0$	33
D_dDDtau	$\bar{B}^0 \rightarrow D^0 D_q(\rightarrow \tau \bar{\nu}_\tau X') X$	$\rho_{\text{isolation}, \tau}^0 \times \rho_{DDX, u/d}^0 \times f_{DDX, \tau}^0 \times f_{DDX}^0 \times N_{D\mu} \times f_{DDX, d, \tau}^0$	34
D_uDDtau	$B^- \rightarrow D^0 D_q(\rightarrow \tau \bar{\nu}_\tau X') X$	$\rho_{\text{isolation}, \tau}^0 \times \rho_{DDX, u/d}^0 \times f_{DDX, \tau}^0 \times f_{DDX}^0 \times N_{D\mu} \times f_{DDX, u, \tau}^0$	35
D_comb	B^- comb. bkg.	$\rho_{B \text{ comb}}^0 \times N_{B \text{ comb}}^0$	36
D_misID	misID.	$\rho_{\text{misID}}^0 \times N_{\text{misID}}^0$	37

Table 8.5: Normalization factors for D^* signal fit with ISO skim templates. Color coding: red: fixed NormFactor; blue: constrained OverallSys; magenta: fixed OverallSys loaded from control fit.

Alias	Decay mode	Normalization	Index
Dst_sigmu	$\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu$	$N_{D^* \mu}$	1
Dst_sigtau	$\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau$	$N_{D^* \mu} \times \mathcal{R}(D^*) \times \tilde{\eta}_{D^{*+}}$	2
Dst_D1	$\bar{B}^0 \rightarrow D_1 \mu \bar{\nu}_\mu$	$\mathcal{B}_{D_1}^* \times \mathcal{B}_{D^{**}}^* \times n_{D^{**}}^* \times N_{D^* \mu} \times f_{D_1}^* \times \frac{1}{2}$	3
Dst_D2	$\bar{B}^0 \rightarrow D_2^* \mu \bar{\nu}_\mu$	$\mathcal{B}_{D_2^*}^* \times \mathcal{B}_{D^{**}}^* \times n_{D^{**}}^* \times N_{D^* \mu} \times f_{D_2^*}^* \times \frac{1}{2}$	4
Dst_D1p	$\bar{B}^0 \rightarrow D_1' \mu \bar{\nu}_\mu$	$\mathcal{B}_{D_1'}^* \times \mathcal{B}_{D^{**}}^* \times n_{D^{**}}^* \times N_{D^* \mu} \times f_{D_1'}^* \times \frac{1}{2}$	5
Dst_D2Smu	$\bar{B} \rightarrow D^{**}(\rightarrow D^* \pi \pi) \mu \bar{\nu}_\mu$	$n_{D^{**}}^* \times N_{D^* \mu} \times f_{D^{*+} \pi \pi}^*$	6
Dst_uD1	$B^- \rightarrow D_1^0 \mu \bar{\nu}_\mu$	$\mathcal{B}_{D_1^0}^* \times \mathcal{B}_{D^{**}}^* \times \epsilon_{\text{isolation}} \times n_{D^{**}}^* \times N_{D^* \mu} \times f_{D_1^0}^*$	7
Dst_uD2	$B^- \rightarrow D_2^{*0} \mu \bar{\nu}_\mu$	$\mathcal{B}_{D_2^{*0}}^* \times \mathcal{B}_{D^{**}}^* \times \epsilon_{\text{isolation}} \times n_{D^{**}}^* \times N_{D^* \mu} \times f_{D_2^{*0}}^*$	8
Dst_uD1p	$B^- \rightarrow D_1'^0 \mu \bar{\nu}_\mu$	$\mathcal{B}_{D_1'^0}^* \times \mathcal{B}_{D^{**}}^* \times \epsilon_{\text{isolation}} \times n_{D^{**}}^* \times N_{D^* \mu} \times f_{D_1'^0}^*$	9
Dst_Ds2	$\bar{B}_s \rightarrow D_{s2}^* \mu \bar{\nu}_\mu$	$\mathcal{B}_{D_{s1}^+}^* \times \frac{f_s}{f_d} \times n_{D^{**}}^* \times \frac{f_{s2}}{f_{s1'}} \times N_{s1'}^* \times N_{D^* \mu}$	10
Dst_Ds1p	$\bar{B}_s \rightarrow D_{s1}' \mu \bar{\nu}_\mu$	$\mathcal{B}_{D_{s1}^+}^* \times \frac{f_s}{f_d} \times N_{D^* \mu} \times n_{D^{**}}^* \times N_{s1'}^*$	11
Dst_D1tau	$\bar{B}^0 \rightarrow D_1 \tau \bar{\nu}_\tau$	$\mathcal{B}_{D_1}^* \times \mathcal{B}_{D^{**}}^* \times \rho_{\mathcal{R}(D^{**})}^* \times n_{D^{**}}^* \times N_{D^* \mu} \times f_{D_1}^* \times \frac{1}{2} \times \mathcal{R}(D^{**})_{\text{avg}}^{\text{raw}} \times r(D_1)$	12
Dst_D2tau	$\bar{B}^0 \rightarrow D_2^* \tau \bar{\nu}_\tau$	$\mathcal{B}_{D_2^*}^* \times \mathcal{B}_{D^{**}}^* \times \rho_{\mathcal{R}(D^{**})}^* \times n_{D^{**}}^* \times N_{D^* \mu} \times f_{D_2^*}^* \times \frac{1}{2} \times \mathcal{R}(D^{**})_{\text{avg}}^{\text{raw}} \times r(D_2^*)$	13
Dst_D1ptau	$\bar{B}^0 \rightarrow D_1' \tau \bar{\nu}_\tau$	$\mathcal{B}_{D_1'}^* \times \mathcal{B}_{D^{**}}^* \times \rho_{\mathcal{R}(D^{**})}^* \times n_{D^{**}}^* \times N_{D^* \mu} \times f_{D_1'}^* \times \frac{1}{2} \times \mathcal{R}(D^{**})_{\text{avg}}^{\text{raw}} \times r(D_1')$	14
Dst_uD1tau	$B^- \rightarrow D_1^0 \tau \bar{\nu}_\tau$	$\mathcal{B}_{D_1^0}^* \times \mathcal{B}_{D^{**}}^* \times \epsilon_{\text{isolation}} \times \rho_{\mathcal{R}(D^{**})}^* \times n_{D^{**}}^* \times N_{D^* \mu} \times f_{D_1^0}^* \times \mathcal{R}(D^{**})_{\text{avg}}^{\text{raw}} \times r(D_1)$	15
Dst_uD2tau	$B^- \rightarrow D_2^{*0} \tau \bar{\nu}_\tau$	$\mathcal{B}_{D_2^{*0}}^* \times \mathcal{B}_{D^{**}}^* \times \epsilon_{\text{isolation}} \times \rho_{\mathcal{R}(D^{**})}^* \times n_{D^{**}}^* \times N_{D^* \mu} \times f_{D_2^{*0}}^* \times \mathcal{R}(D^{**})_{\text{avg}}^{\text{raw}} \times r(D_2^*)$	16
Dst_uD1ptau	$B^- \rightarrow D_1'^0 \tau \bar{\nu}_\tau$	$\mathcal{B}_{D_1'^0}^* \times \mathcal{B}_{D^{**}}^* \times \epsilon_{\text{isolation}} \times \rho_{\mathcal{R}(D^{**})}^* \times n_{D^{**}}^* \times N_{D^* \mu} \times f_{D_1'^0}^* \times \mathcal{R}(D^{**})_{\text{avg}}^{\text{raw}} \times r(D_1')$	17
Dst_dDDmu	$\bar{B}^0 \rightarrow D^* D_q(\rightarrow \mu \bar{\nu}_\mu X') X$	$\rho_{DDX,u/d}^* \times f_{DDX}^* \times N_{D^* \mu} \times f_{DDX,d}^*$	18
Dst_uDDmu	$B^- \rightarrow D^* D_q(\rightarrow \mu \bar{\nu}_\mu X') X$	$\rho_{DDX,u/d}^* \times f_{DDX}^* \times N_{D^* \mu} \times f_{DDX,u}^*$	19
Dst_dDDtau	$\bar{B}^0 \rightarrow D^* D_q(\rightarrow \tau \bar{\nu}_\tau X') X$	$\rho_{\text{isolation},\tau}^* \times \rho_{DDX,u/d}^* \times f_{DDX}^* \times f_{DDX,\tau}^* \times N_{D^* \mu} \times f_{DDX,d,\tau}^*$	20
Dst_uDDtau	$B^- \rightarrow D^* D_q(\rightarrow \tau \bar{\nu}_\tau X') X$	$\rho_{\text{isolation},\tau}^* \times \rho_{DDX,u/d}^* \times f_{DDX}^* \times f_{DDX,\tau}^* \times N_{D^* \mu} \times f_{DDX,u,\tau}^*$	21
Dst_comb	$\bar{B}^0$ comb. bkg.	$\rho_B^* \times N_B^*$	22
Dst_misID	misID.	$\rho_{\text{misID}}^* \times N_{\text{misID}}^*$	23
Dst_doug	D^* comb. bkg.	$\rho_{D^* \text{ comb}}^* \times N_{D^* \text{ comb}}^*$	24

allowing it to vary with a pre-set uncertainty around the expected yield obtained from the relevant data-driven control sample.

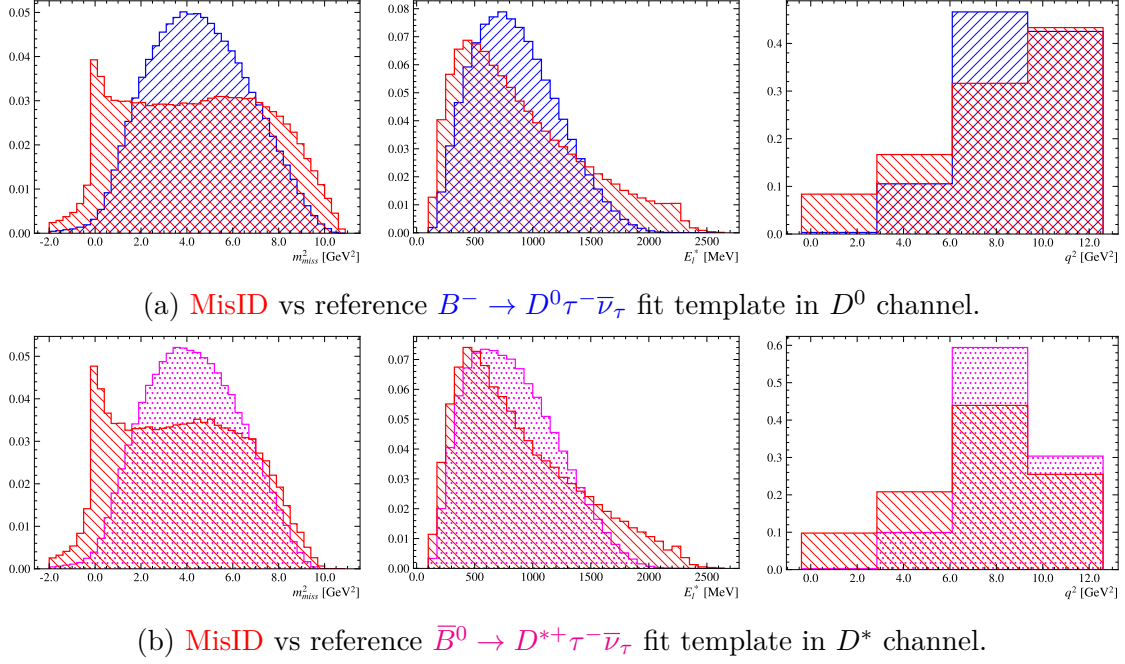


Figure 8.2: Comparisons between misID templates. The misID templates are normalized so that only shape comparisons are meaningful.

The misID templates are generated by unfolding different contributions from the fake μ control sample (Section 5.1.3). First conceived in Ref. [42], the unfolding technique consists of the following steps:

1. Tag the fake μ sample into π , K , p , e , and ghost³-like species with the selection requirements listed in Table 8.6. Denote the tagged species with a hat: $\hat{t} \in \{\hat{\pi}, \hat{K}, \hat{p}, \hat{e}, \hat{g}\}$, respectively.
2. Obtain the efficiency for a track of the true species t passing μ acceptance to be classified as a tagged species $\hat{t}'$ (where t and t' can be the same species, e.g. a true electron passing μ acceptance classified as a tagged electron): $\epsilon(t \rightarrow \hat{t}')$. These efficiencies are obtained from PIDCALIB2 for π, K, p, e and from MC for g .

³“Ghost” refers to tracks formed by random combination of hits.

3. Given the measured yields in the fake μ sample $\tilde{n}_{\hat{t}'}$ and response matrix $M_{\hat{t}',t} = \epsilon(t \rightarrow \hat{t}')$, unfold the true yields $\tilde{n}_t$. A Bayesian unfolding (iterative unfolding, [43]) algorithm is used to obtain the true yields and subsequently the efficiencies of a tagged $\hat{t}'$ being a true t : $\epsilon(\hat{t}' \rightarrow t)$.
4. Find $\epsilon(\hat{t}' \rightarrow \hat{\mu})$, the transfer factor from the fake μ sample (of tagged species $\hat{t}'$) to the “real” μ samples⁴, based on
 - the efficiencies obtained in the step above,
 - the efficiency of a track of true species t that falls within μ acceptance to also pass the μ PID cuts $\epsilon(t_{\text{acc}} \rightarrow \hat{\mu})$,
 - and the efficiency of t_{acc} to *fail* the μ PID: $\epsilon(t_{\text{acc}} \rightarrow t)$.

$$\epsilon(\hat{t}' \rightarrow \hat{\mu}) = \sum_t \frac{\epsilon(\hat{t}' \rightarrow t)}{\epsilon(t_{\text{acc}} \rightarrow t)} \epsilon(t_{\text{acc}} \rightarrow \hat{\mu}) \quad (8.3)$$

The last two efficiencies are found from PIDCALIB2.

5. Apply $\epsilon(\hat{t}' \rightarrow \hat{\mu})$ as a weight for each tagged species $\hat{t}'$ in the fake μ sample. The weighted yield of tagged species $\hat{t}$ represents the μ misID contributions of $\hat{t}$ to the “real” μ samples.

Comparisons between tagged and unfolded (true) yields of the fake μ sample are shown in Fig. 8.3. The weighted yields are projected in fit variables with finer binning; these are shown in Fig. 8.4. Some of the efficiencies obtained with PIDCALIB2 are negative; these are shifted back to a physical value between $[0, 1]$ with an algorithm described in Appendix B.1. A more detailed description of the unfolding procedure is documented in Appendix E.

⁴A subtlety is that here “real” μ refers to the samples passing μ PID cuts listed in Section 5.1.1. Because the μ -like particles are selected by a set of cuts, a hat is placed on μ ($\hat{\mu}$!) to remind the reader that this sample is nothing but a μ tag and is unrelated to the μ -enriched sample (true μ) as used by PIDCALIB2.

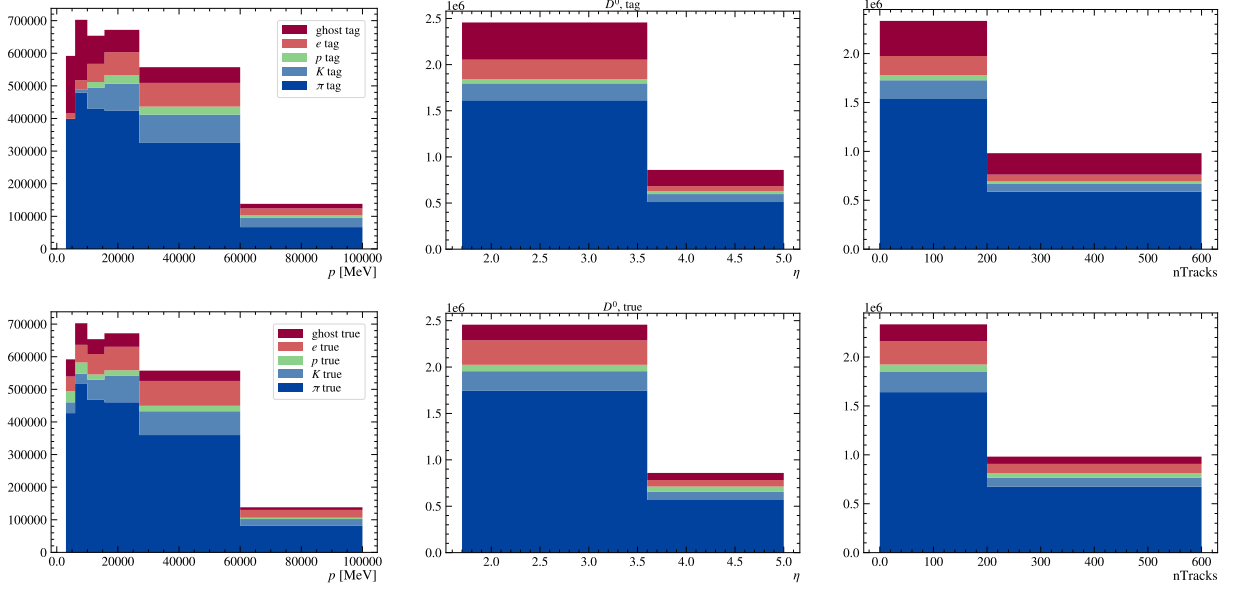


Figure 8.3: Unfolding effect displayed in p and η of the fake μ track, and nTracks of the reconstructed event. Top are the raw tagged yields, bottom are the unfolded true yields. The number of events is conserved by unfolding.

The samples displayed here are 2016 D^0 fake μ control samples $B^- \rightarrow D^0 t^-$ passing selections listed in Section 5.1.3.

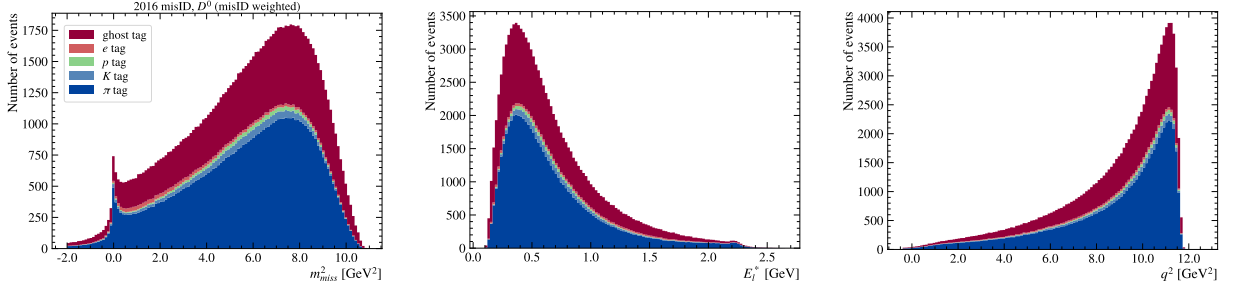


Figure 8.4: The weighted yields of each species of fake μ sample projected in fit variables. The dominate contributions to the “real” muon samples are from π -like and ghost-like tracks. The binning are finer compared to Fig. 8.2. No isolation cut applied. Actual yields are displayed without any rescaling.

8.1.2 Combinatorial backgrounds

The selection procedures for all combinatorial backgrounds are listed in Section 5.1.2. The generation procedure for these templates is documented below. A comparison between combinatorial backgrounds and signal templates is shown in Fig. 8.5.

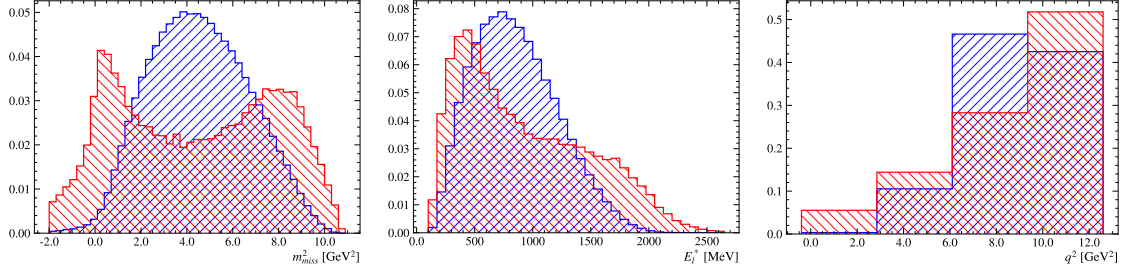
Table 8.6: Selections for each tagged species in fake μ sample.

Tagged species	Variable	Selection
π	MC15TuneV1ProbNN π	> 0.1
	DLLK	< 0.0
	DLL p	< 0.0
	DLL e	< 2.0
	MC15TuneV1ProbNNghost	< 0.25
K	MC15TuneV1ProbNNK	> 0.1
	DLLK	> 0.0
	DLL p – DLLK	< 0.0
	DLL e – DLLK	< -2.0
	MC15TuneV1ProbNNghost	< 0.25
p	MC15TuneV1ProbNN p	> 0.1
	DLL p	> 0.0
	DLL p – DLLK	> 2.0
	DLL e – DLL p	< -2.0
	MC15TuneV1ProbNNghost	< 0.25
e	DLL e	> 2.0
	DLL e – DLLK	> -2.0
	DLL e – DLL p	> -2.0
	MC15TuneV1ProbNNghost	< 0.25
ghost	Not in any species above	

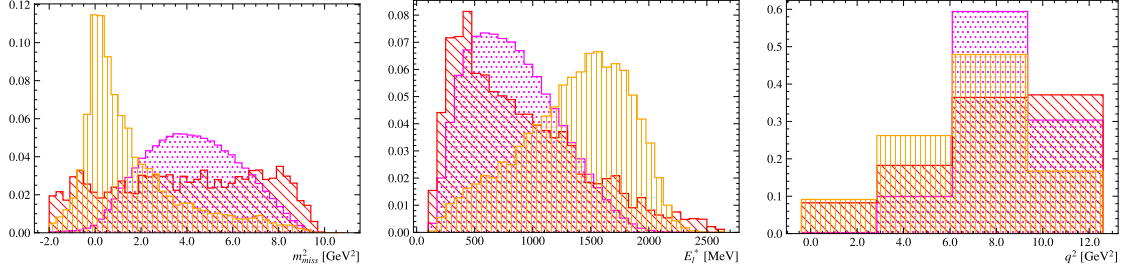
8.1.2.1 D^* combinatorial

The D^* combinatorial background arises from a reconstructed D^0 combining with an random slow π , forming a fake D^* vertex and passing all selection requirements.

The shape of the D^* combinatorial can be determined from the wrong-sign π (WS π) control sample which contains mostly $D^0\pi^-$ pairs. Still, the yields of the D^* combinatorial contributions to the nominal right-sign (RS) and WS π samples are not the same, likely due to differences in selection efficiencies. Thus a fit on the RS sample is performed to obtain the yield of D^* combinatorial, then the WS π shape is rescaled to match the fitted yield.



(a) B combinatorial vs reference $B^- \rightarrow D^0 \tau^- \bar{\nu}_\tau$ in D^0 channel.



(b) B combinatorial vs D^* combinatorial vs reference $\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau$ in D^* channel.

Figure 8.5: Comparison between combinatorial background fit templates.

The fit procedure is the following:

1. Remove the misID contribution from the RS sample.
2. Remove the misID contribution from the WS π sample⁵.
3. Fit the $m_{D^*} - m_{D^0}$ distribution (including events normally outside the D^* selection mass window, referred as “mass side-band”) on the RS sample with a double-Gaussian signal and an exponential background. The background yield under the mass window where a D^* is nominally accepted is taken as the yield of D^* combinatorial.

The fit to the ISO skim (expanded to include the mass side-band) is shown in Fig. 8.6.

The effect of scaling for the ISO skim is shown in Fig. 8.7. For fit to 1OS, 2OS, and DD skims, refer to Appendix A.2.

⁵A WS π control sample from misID sample is reconstructed to obtain the misID contribution for data WS π .

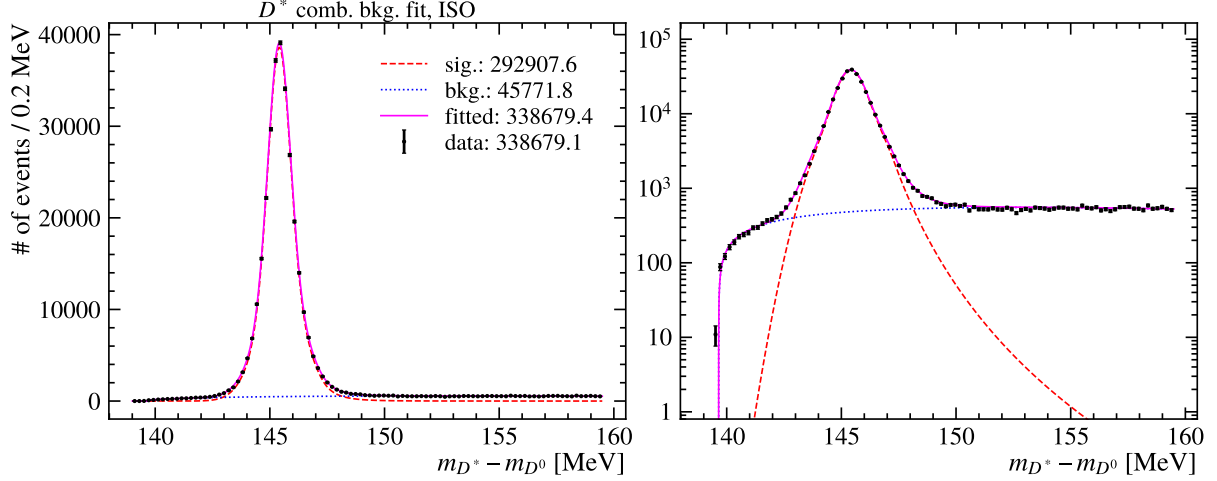


Figure 8.6: D^* combinatorial auxiliary fit to the ISO skim. Right plot shows the same fit as in the left, but with a logarithmic y axis.

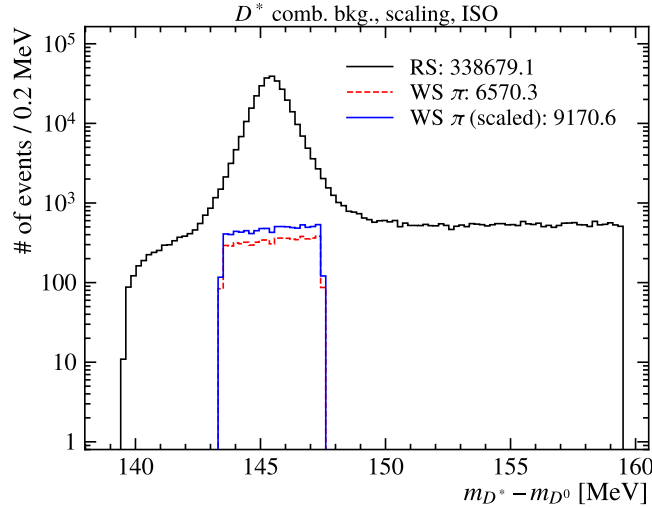


Figure 8.7: Effect of scaling for the ISO skim on the D^* combinatorial template.

8.1.2.2 B combinatorial in D^* channel

The B combinatorial background in the D^* fit channel comes from randomly combined $D^*\mu$ pairs. Again, the shape of the B combinatorial is determined from the wrong-sign μ (WS μ) control sample containing $D^{*+}\mu^+$ pairs.

As shown in Fig. 8.8, the RS and WS μ yield ratios are not one. It is empirically known that such a background can be characterized by a slow-varying exponential function

of m_B , thus the yield ratio between the two is also an exponential. Such an exponential permits Taylor expansion, and the yield ratio can be well-described by the first order term $a + bm_B$ (linear in m_B) across a sufficiently larger m_B range.

Unlike m_{D^*} in the D^* combinatorial background, m_B does not have a clear lower side-band. This is because B is only *partially* reconstructed due to missing neutrino(s). Therefore, only the upper-sideband, defined as $m_B > 5400$ MeV, which contains purely combinatorial and misID, can be used to study the B combinatorial background.

A fit is performed in the B mass upper-sideband to obtain the linear scaling factor $a + bm_B$. In the upper-sideband, it is assumed that both WS μ and RS samples contain only combinatorial backgrounds and misID. The linear scaling factor is assumed to be the same for all skims⁶. The fit is performed in the following manner:

1. Remove the misID contribution from the RS sample.
2. Remove the misID contribution from the WS μ sample.
3. Perform a fit in the upper-sideband region of the RS sample to determine the contribution from D^* combinatorial, with procedure identical to that in Section 8.1.2.1. Then remove the fitted D^* combinatorial.
4. Fit and remove D^* combinatorial from the WS μ sample.
5. Compute the ratio between RS and WS μ , perform a fit to obtain the linear scaling factor. The fit, as well as the raw RS, raw WS μ , and scaled WS μ templates, is shown in Fig. 8.8.

8.1.2.3 B combinatorial in D^0 channel

The procedure is similar to that in Section 8.1.2.2, but without D^* combinatorial removal. The fit is shown in Fig. 8.9.

⁶Within the limited statistical power of these skims, this is checked to be the case.

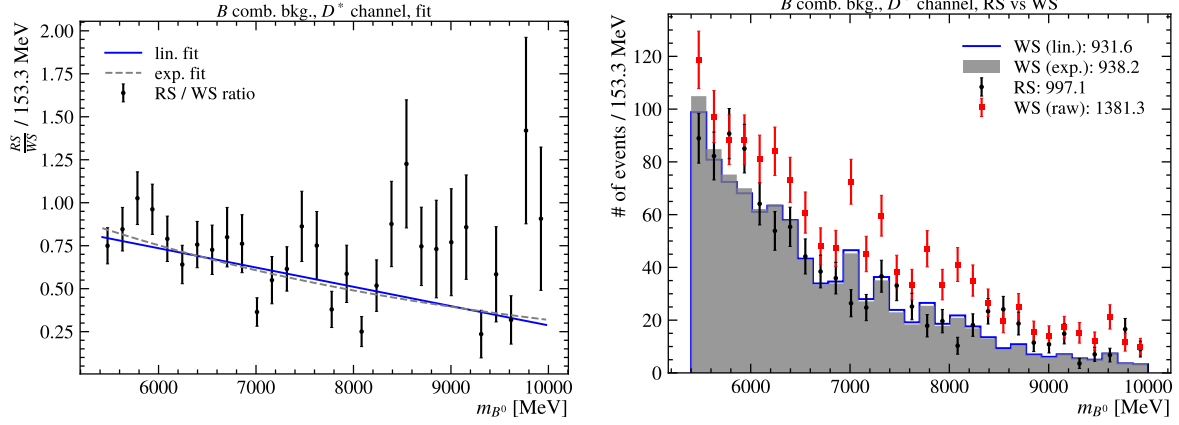


Figure 8.8: B combinatorial fit result for D^* channel.

Left: B combinatorial auxiliary fit in the m_B upper-sideband region for all skims. A quadratic fit is also performed for possible systematic studies.

Right: The scaling effect on the WS μ B combinatorial template.

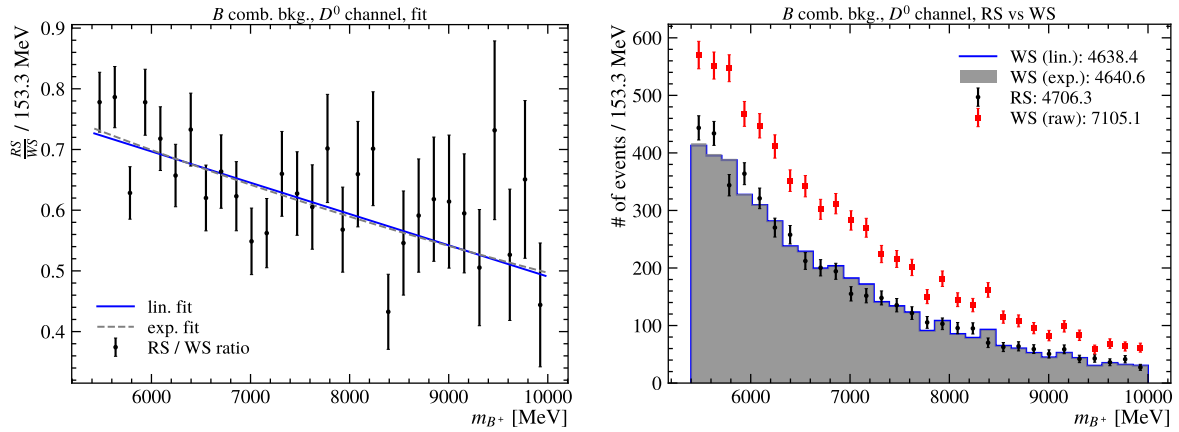


Figure 8.9: B combinatorial fit result for D^0 channel. Left: B combinatorial auxiliary fit. Right: The scaling effect on the WS μ B combinatorial template.

8.1.3 Normalization

$B \rightarrow D^{0,*+} \mu^- \bar{\nu}_\mu$ The yields of these two decay modes, denoted as $N_{D\mu}$ and $N_{D^*\mu}$, are set to be the free-floating overall normalization for all MC templates of the respective fit channel. As a side note, the $D^0\mu$ template has notably softer q^2 spectrum compared to that of $D^{*+}\mu$, as shown in Fig. 8.10.

The $D^{*+}\mu$ mode also contributes to the D^0 channel due to feed down, with its yield

normalized by:

$$N_{D\mu} \times \{r_{D^*}^{\text{isospin}} \times r_{D^{*0}}^0\} \quad (8.4)$$

where $r_{D^*}^{\text{isospin}}$ is the relative yield ratio between the $D^{*+}\mu$ and $D^{*0}\mu$ samples:

$$r_{D^*}^{\text{isospin}} = \frac{\epsilon(\bar{B}^0 \rightarrow D^{*+}\mu^-\bar{\nu}_\mu)}{\epsilon(B^- \rightarrow D^{*0}\mu^-\bar{\nu}_\mu)} \bigg|_{\text{initial value from MC}}^{\text{floating in the fit}} \times \frac{\mathcal{B}(\bar{B}^0 \rightarrow D^{*+}\mu^-\bar{\nu}_\mu)}{\mathcal{B}(B^- \rightarrow D^{*0}\mu^-\bar{\nu}_\mu)} \quad (8.5)$$

where the first ratio, a selection efficiency ratio, with its initial value obtained with MC samples satisfying the D^0 channel selection criteria, is not unity due to different isolation efficiencies for charged D^{*+} and neutral D^{*0} decays; the second ratio is the branching fraction ratio attained from external measurements. The other parameter $r_{D^{*0}}^0$ is the yield ratio between the $D^{*0}\mu$ and $D^0\mu$ samples⁷:

$$r_{D^{*0}}^0 = \frac{\epsilon(B^- \rightarrow D^{*0}\mu^-\bar{\nu}_\mu)}{\epsilon(B^- \rightarrow D^0\mu^-\bar{\nu}_\mu)} \bigg|_{\text{initially set to 1}}^{\text{floating in the fit}} \times \frac{\mathcal{B}(B^- \rightarrow D^{*0}\mu^-\bar{\nu}_\mu)}{\mathcal{B}(B^- \rightarrow D^0\mu^-\bar{\nu}_\mu)} \quad (8.6)$$

where the first term is the selection efficiency ratio between the $D^0\mu^-$ and $D^{*0}\mu^-$, which is initially assumed to be 1 and floated in the fit.

$B \rightarrow D^{*0}\mu^-\bar{\nu}_\mu$ This mode contributes to the D^0 channel exclusively due to feed down. Its yield is constrained by

$$N_{D\mu} \times r_{D^{*0}}^0 \quad (8.7)$$

A comparison between this template and the signal template in the D^0 channel is also shown in Fig. 8.10.

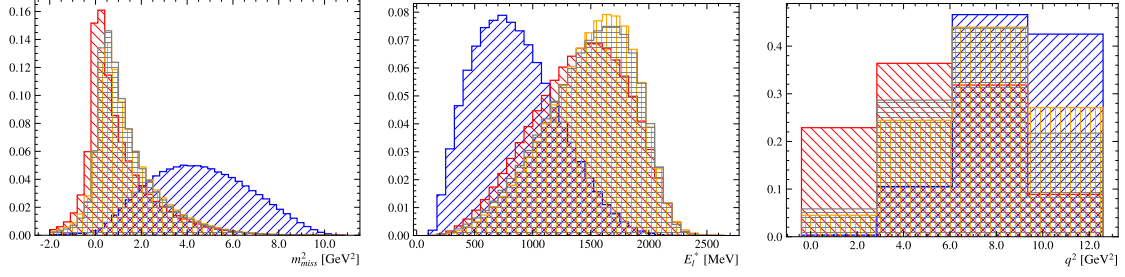
8.1.4 Signal

$B^- \rightarrow D^0\tau^-\bar{\nu}_\tau$ Its yield is related to $N_{D\mu}$ by the following expression:

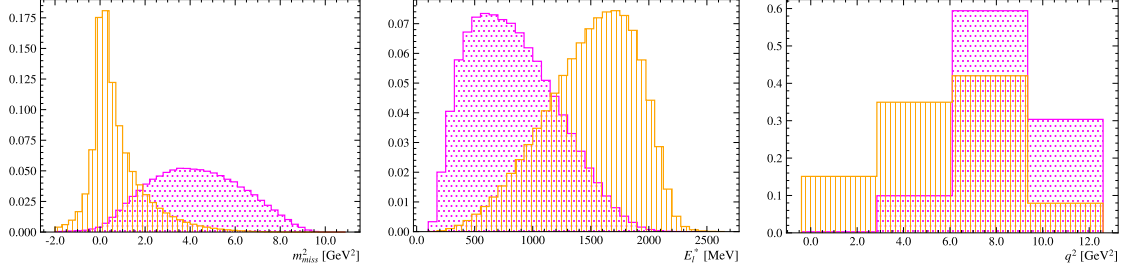
$$N_{D\mu} \times \left\{ \underbrace{\mathcal{R}(D) \times \mathcal{B}(\tau^- \rightarrow \mu^-\bar{\nu}_\mu\nu_\tau)}_{\equiv \eta_{D^0}} \times \frac{\epsilon(B^- \rightarrow D^0\tau^-[\rightarrow \mu^-\bar{\nu}_\mu\nu_\tau]\bar{\nu}_\tau)}{\epsilon(B^- \rightarrow D^0\mu^-\bar{\nu}_\mu)} \right\} \quad (8.8)$$

where $\mathcal{R}(D) \equiv \mathcal{R}(D^0)$ is the free-floating parameter of interest, and η_{D^0} is the (τ leptonic

⁷With the convention that a superscript “0/*” indicates it is a parameter in the D^0/D^* channel.



(a) $B^- \rightarrow D^0 \mu^- \bar{\nu}_\mu$ vs $\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu$ feed down vs $B^- \rightarrow D^{*0} \mu^- \bar{\nu}_\mu$ feed down. All in D^0 channel. The signal template $B^- \rightarrow D^0 \tau^- \bar{\nu}_\tau$ in the D^0 channel is included as a reference.



(b) $\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu$ vs $\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau$, both in D^* channel.

Figure 8.10: Comparisons between normalization templates.

exclusive) efficiency ratio between the signal and the normalization mode in the D^0 channel, with the efficiencies obtained with the corresponding MC samples satisfying the $D^0 \mu$ selection criteria.

$B \rightarrow D^{*+} \tau^- \bar{\nu}_\tau$ This signal mode contributes to both channels due to feed down. In the D^* channel, its normalization is given by:

$$N_{D^* \mu} \times \left\{ \mathcal{R}(D^*) \times \mathcal{B}(\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau) \times \underbrace{\frac{\epsilon(\bar{B}^0 \rightarrow D^{*+} \tau^- [\rightarrow \mu^- \bar{\nu}_\mu \nu_\tau] \bar{\nu}_\tau)_{D^*}}{\epsilon(\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu)_{D^*}}}_{\equiv \tilde{\eta}_{D^{*+}}} \right\} \quad (8.9)$$

where $\mathcal{R}(D^*) \equiv \mathcal{R}(D^{*+})$ is the other parameter of interest, and $\tilde{\eta}_{D^{*+}}$ is the efficiency ratio between the signal and the normalization in the D^* channel, with the tilde emphasizing that the efficiencies are evaluated with the $D^{*+} \mu$ selection cuts.

Its D^0 channel feed down yield can be formally written as

$$N_{D\mu} \times \underbrace{\left\{ \frac{\mathcal{B}(\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu)}{\mathcal{B}(B^- \rightarrow D^0 \mu^- \bar{\nu}_\mu)} \times \frac{\epsilon(\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu)}{\epsilon(B^- \rightarrow D^0 \mu^- \bar{\nu}_\mu)} \times \right.}_{\equiv r_{D^*}^{\text{isospin}} \times r_{D^{*0}}^0} \underbrace{\left. \frac{\mathcal{B}(\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau) \times \frac{\epsilon(\bar{B}^0 \rightarrow D^{*+} \tau^- [\rightarrow \mu^- \bar{\nu}_\mu \nu_\tau] \bar{\nu}_\tau)}{\epsilon(\bar{B}^0 \rightarrow D^{*+} \mu^- \bar{\nu}_\mu)} \times \mathcal{R}(D^*) \right\}}_{\equiv \eta_{D^{*+}}, \text{ which is not } \tilde{\eta}_{D^{*+}}!} \quad (8.10)$$

which can then be rewritten as

$$N_{D^0\mu} \times \left\{ \mathcal{R}(D^*) \times \frac{\eta_{D^{*+}}}{\tilde{\eta}_{D^{*+}}} \times \tilde{\eta}_{D^{*+}} \times r_{D^*}^{\text{isospin}} \times r_{D^{*0}}^0 \right\} \quad (8.11)$$

noting that $\eta_{D^{*+}}$ is the efficiency ratio evaluated at the D^0 channel (also notice the lack of a tilde!) and the $\mathcal{B}(\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau)$ cancels in the second parameter inside the bracket above which is a ratio.

$B^- \rightarrow D^{*0} \tau^- \bar{\nu}_\tau$ Only contributing to the D^0 channel, its yield is given by:

$$N_{D^0\mu} \times \left\{ \mathcal{R}(D^*) \times \frac{\eta_{D^{*0}}}{\tilde{\eta}_{D^{*+}}} \times \tilde{\eta}_{D^{*+}} \times r_{D^{*0}}^0 \right\} \quad (8.12)$$

where $\eta_{D^{*0}}$ is defined similarly to $\eta_{D^{*+}}$.

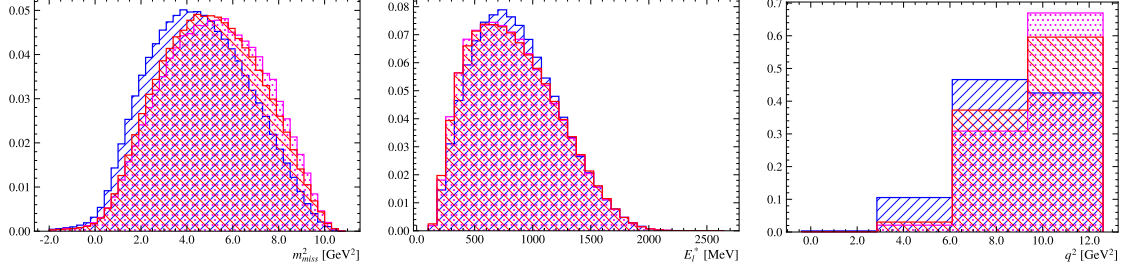
The four efficiency ratios $\eta_{D^0}, \tilde{\eta}_{D^{*+}}, \eta_{D^{*+}}, \eta_{D^{*0}}$ are all fixed in the fit, with the (multiplicative) uncertainty induced by these parameters evaluated separately (Section 9.11). Comparisons between the signal templates are shown in Fig. 8.11.

8.1.5 Feed down of $B \rightarrow D^{**} l^- \bar{\nu}_l$ modes

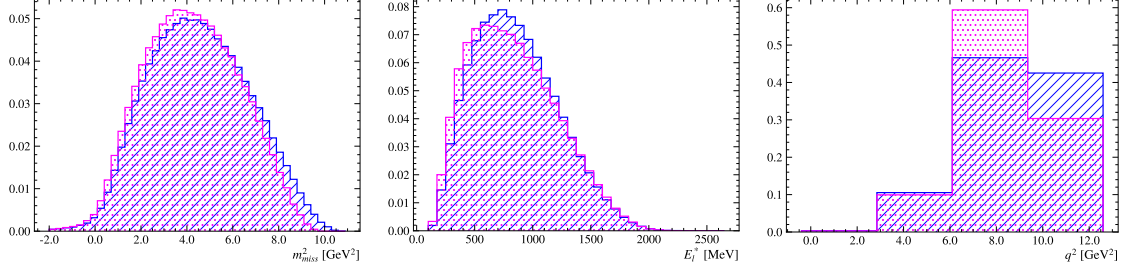
The fit includes the following feed-down templates from the four $1P$ D^{**} states: $D_0^*(2300)$, $D_1'(2430)$, $D_1(2420)$, $D_2^*(2460)$, including both isospin pairs for each D^{**} , with the subscript indicating the total angular momentum J of the meson⁸:

- **D^0 channel:** All four states contribute. Additionally, the D_1 states are reported to have the cascade decays $D_1 \rightarrow D_0^*(\rightarrow D^0 \pi) \pi$ observed; events with such decay modes

⁸Note that for $S = 1$ and $L = 1$, the allowed values for J are 2, 1, 0. That is, $1 \oplus 1 = 2 \oplus 1 \oplus 0$.



(a) $B^- \rightarrow D^0 \tau^- \bar{\nu}_\tau$ vs $\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau$ feed down vs $B^- \rightarrow D^{*0} \tau^- \bar{\nu}_\tau$ feed down.



(b) $B^- \rightarrow D^0 \tau^- \bar{\nu}_\tau$ in D^0 channel vs $\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau$ in D^* channel.

Figure 8.11: Comparisons between the signal templates.

are put into the $D_1 \rightarrow (D^0 \pi \pi)$ templates, abbreviated as $D_1 \pi \pi$.

Therefore, for the D^0 channel, there are $(4+1) \times 2$ muonic templates where the factor 2 comes from the isospin doublet (e.g. D_0^{*0} and D_0^{*+}) and $(4+1) \times 2$ tauonic ones for a total of 20 D^{**} templates.

- **D^* channel:** Only the $J = 1, 2$ states contribute. The D_0^* is assumed to *not* contribute to this channel due to conservation of angular momentum. Also, no cascade decay of the form $D^{**} \rightarrow D^{**}(\rightarrow D^* \pi) \pi$ is considered. For this channel, there are 3×2 muonic templates and 3×2 tauonic ones for a total of 12 D^{**} templates.

Both channels contain contributions from a muonic D^{**} component ($B \rightarrow D^{**} \mu \nu_\mu$) and a tauonic one ($B \rightarrow D^{**} \tau \nu_\tau$).

Muonic The normalization factor of a D^{**} “species” i in the D^0 channel is given by

$$\Omega_{D^{**}}^0 \equiv N_{D\mu} \times \left\{ \rho_{D^{**}}^{\text{isospin}} \times n_{D^{**}}^0 \times \mathcal{B}_{D^{**}}^0 \times \text{expected fraction } D_i^{**}/D^{**} \right\} \quad (8.13)$$

where the allowed species are $D_0^*, D_1', D_1, D_1 \pi \pi, D_2^*$ (both isospins), and

- $\rho_{D_i^{**}}^{\text{isospin}}$ is a Gaussian isospin constraint initially set to 1 with $\sigma = 0.5$.
- $n_{D^{**}}^0 \equiv \frac{10^{-3}}{\mathcal{B}(B^- \rightarrow D^0 \mu^- \bar{\nu}_\mu)}$ is a shared convenience factor so that the floating $\mathcal{B}_{D_i^{**}}^0$ can be interpreted as the branching fraction of D_i^{**} multiplied by its selection efficiency. $\mathcal{B}_{D_i^{**}}^0$ is initially set around 1.
- “expected fraction D_i^{**}/D^{**} ” is different for B^- and $\bar{B}^0$ modes:

- **B^- modes:** contains a single parameter $f_{D_i^{**}}^0$ defined as

$$f_{D_i^{**}}^0 = \mathcal{B}(B^- \rightarrow D_i^{**}) \times \frac{\epsilon(B^- \rightarrow D^{**} \mu)_{\text{avg}}}{\epsilon(B^- \rightarrow D^0 \mu)_{\text{avg}}} \quad (8.14)$$

where the branching fraction is computed by accounting for the relevant $B \rightarrow D^{**} \rightarrow D$ and $B \rightarrow D^{**} \rightarrow D^* \rightarrow D$ decay chains, assuming $\mathcal{B}(D^{*0} \rightarrow D^0 X) \approx 100\%$ and $\mathcal{B}(D^{*+} \rightarrow D^0 X) \approx 67\%$. Currently, these values are taken from the run 1 pathfinder analysis [15].

- **$\bar{B}^0$ modes:** contains two parameters $f_{D_i^{**}}^0 \times \epsilon_{D_i^{**}}$, where the second parameter is an efficiency ratio, defined as

$$\epsilon_{D_i^{**}} = \frac{\epsilon(\bar{B}^0 \rightarrow D_i^{**} \mu)}{\epsilon(B^- \rightarrow D^{**} \mu)_{\text{avg}}} \quad (8.15)$$

In the D^* channel, the normalization of D_i^{**} is:

$$\Omega_{D^{**}}^* \equiv N_{D^* \mu} \times \left\{ \underbrace{\mathcal{B}_{D_i^{**}}^* \times \mathcal{B}_{D^{**}}^*}_{\text{two Gaussian constraints}} \times n_{D^{**}}^* \times \text{expected fraction } D_i^{**}/D^{**} \right\} \quad (8.16)$$

where the allowed species are D_1', D_1, D_2^* . The Gaussian systematic uncertainties on the branching fraction multiplied by the (selection) isolation efficiency is split into two parameters: the correlated $\mathcal{B}_{D^{**}}^*$ and the uncorrelated $\mathcal{B}_{D_i^{**}}^*$. The shared convenience factor $n_{D^{**}}^*$ is defined similar to $n_{D^{**}}^0$, with the branching fraction replaced by $\mathcal{B}(\bar{B}^0 \rightarrow D^{*0} \mu^- \bar{\nu}_\mu)$. Again, the “expected fraction D_i^{**}/D^{**} ” is different between B^- and $\bar{B}^0$ modes:

- **B^- modes:** consists of two parameters:

$$\underbrace{\mathcal{B}(B^- \rightarrow D_i^{**}) \times \frac{\epsilon(B^- \rightarrow D^{**}\mu)_{\text{avg}}}{\epsilon(\bar{B}^0 \rightarrow D^{**}\mu)_{\text{avg}}}}_{\equiv f_{D_i^{**}}^*} \times \epsilon_{\text{isolation}} \quad (8.17)$$

where $\epsilon_{\text{isolation}}$ is an additional systematic uncertainty on BDT isolation efficiency.

- **$\bar{B}^0$ modes:** consists of two parameters:

$$\underbrace{\mathcal{B}(\bar{B}^0 \rightarrow D_i^{**})}_{\equiv f_{D_i^{**}}^*} \times \frac{1}{2} \quad (8.18)$$

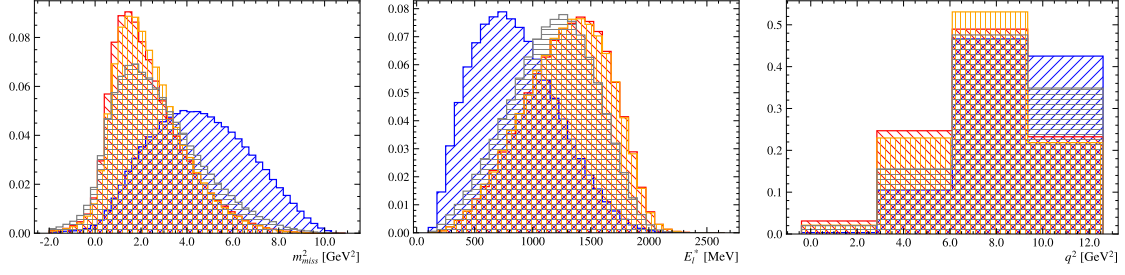
where the second parameter comes from isospin relation $\mathcal{B}(D^{**} \rightarrow D^{*+}\pi^-)/\mathcal{B}(D^{**} \rightarrow D^{*0}\pi^0) = 1/2$.

The D^{**} muonic decays often contain unreconstructed π^0 or γ which leads to a broad m_{miss}^2 distribution peaking at smaller values $\approx m_\pi^2$, as shown in Fig. 8.12. The E_μ^* spectrum is comparable, but typically softer, than that in normalization, due to D^{**} states having higher masses compared to D^0 and D^* . These decays are suppressed at zero recoil, but this only translates weakly in the reconstructed q^2 spectrum, due to rest frame approximation (Appendix C) and definition of q^2 being $(p_B - p_{D^{(*)}})^2$, instead of $(p_B - p_{D^{**}})^2$ [15].

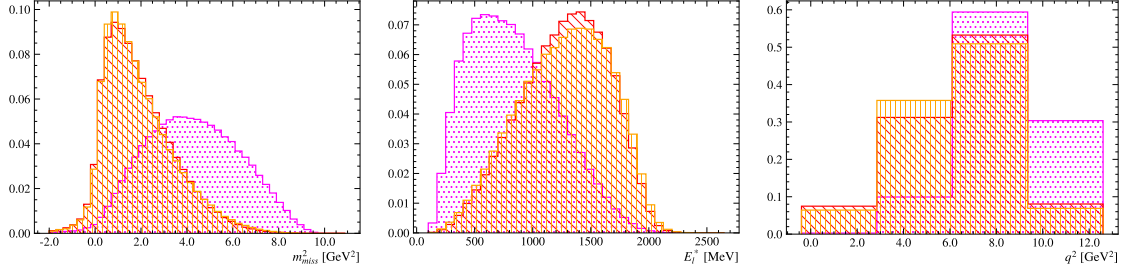
Tauonic As in run 1, the yields of tauonic D^{**} decays are constrained as a fraction of their muonic counterparts individually. In both channels, the yield with a loose 33% constraint is given by:

$$\Omega \times \mathcal{R}(D^{**})_{\text{avg}}^{\text{raw}} \times r(D_i^{**}) \times \underbrace{\left[\rho_{\mathcal{R}(D^{**})}^0 \text{ for } D^0 \text{ chan, } \rho_{\mathcal{R}(D^{**})}^* \text{ for } D^* \text{ chan} \right]}_{\text{a Gaussian constraint with a } \sigma = 0.33} \quad (8.19)$$

where Ω refers to Eq. (8.13) for D^0 channel, and Eq. (8.16) for D^* channel. The $\mathcal{R}(D^{**})_{\text{avg}}^{\text{raw}}$ is based on the average between $\mathcal{R}(D_i^{**})$, taken to be $\mathcal{R}(D^{**})_{\text{avg}} = 0.12$ which is the value used in the run 1 analysis, while also absorbing selection efficiencies. Each $r(D_i^{**}) \equiv \frac{\mathcal{R}(D_i^{**})}{\mathcal{R}(D^{**})_{\text{avg}}}$ is also taken from the previous analysis. Displayed in 8.13, the tauonic D^{**} templates have harder m_{miss}^2 and q^2 spectra compared to the signal templates.

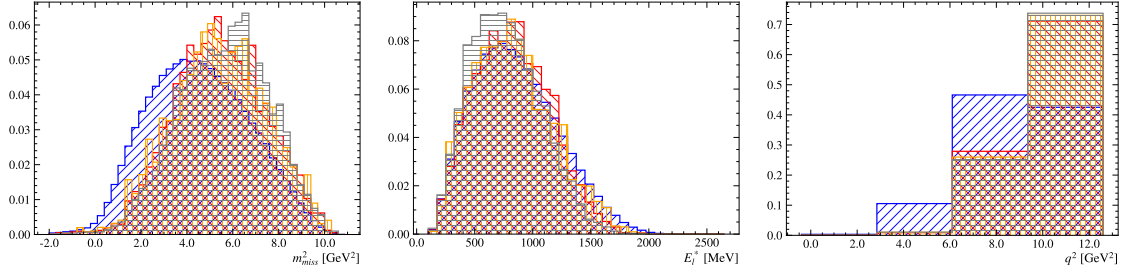


(a) $\bar{B}^0 \rightarrow D_1^+ \mu^- \bar{\nu}_\mu$ vs $\bar{B}^0 \rightarrow D_1^+ (\rightarrow D^0 \pi \pi) \mu^- \bar{\nu}_\mu$ vs $\bar{B}^0 \rightarrow D_0^{*+} \mu^- \bar{\nu}_\mu$. All in D^0 channel. The $B^- \rightarrow D^0 \tau^- \bar{\nu}_\tau$ is plotted as a reference.

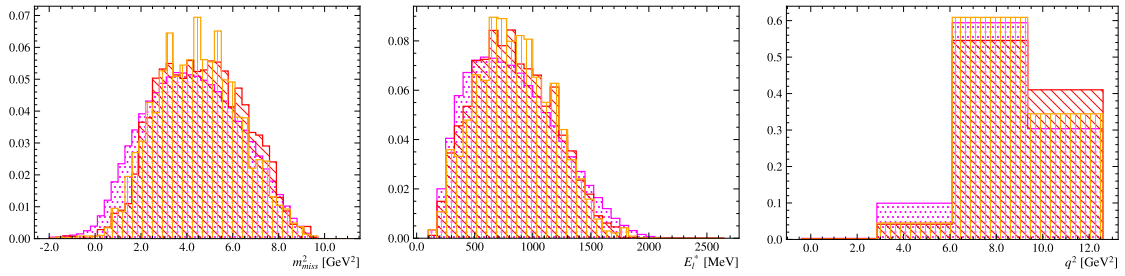


(b) $\bar{B}^0 \rightarrow D_1^+ \mu^- \bar{\nu}_\mu$ vs $\bar{B}^0 \rightarrow D_1'^+ \mu^- \bar{\nu}_\mu$ vs reference $\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau$. All in D^* channel.

Figure 8.12: Comparison between muonic D^{**} templates.



(a) $\bar{B}^0 \rightarrow D_1^+ \tau^- \bar{\nu}_\tau$ vs $\bar{B}^0 \rightarrow D_1^+ (\rightarrow D^0 \pi \pi) \tau^- \bar{\nu}_\tau$ vs $\bar{B}^0 \rightarrow D_0^{*+} \tau^- \bar{\nu}_\tau$. All in D^0 channel. The $B^- \rightarrow D^0 \tau^- \bar{\nu}_\tau$ is plotted as a reference.



(b) $\bar{B}^0 \rightarrow D_1^+ \tau^- \bar{\nu}_\tau$ vs $\bar{B}^0 \rightarrow D_1'^+ \tau^- \bar{\nu}_\tau$ vs reference $\bar{B}^0 \rightarrow D^{*+} \tau^- \bar{\nu}_\tau$. All in D^* channel.

Figure 8.13: Comparison between tauonic D^{**} templates.

8.1.6 Feed down of $B \rightarrow D_H^{**}(\rightarrow D^{(*)}\pi\pi)\mu^-\bar{\nu}_\mu$ modes

The heavy D_H^{**} decay into either a D^0 or a D^* , with two additional π , one of them possibly uncharged. The D^{*+} events contribute to both channels while the D^{*0} only to the D^0 channel. These decay modes are not well measured, and a cocktail MC is used so that the relative fractions between the D_H^{**} are fixed at simulation level. More details regarding the composition of the cocktail can be found in [15]. The yield of $D_H^{**} \rightarrow D_i$, essentially unconstrained, is defined similarly for both channels:

$$N_{\text{ch ref}} \times \left\{ n_{\text{scale}} \times f_{D_i} \underbrace{(\times f_{\text{guess}})}_{\text{only present in } D^0 \text{ channel. A fixed parameter.}} \right\} \quad (8.20)$$

where D_i can be a D^0 , a D^{*0} , or a D^{*+} . Depend on the context, $N_{\text{ch ref}} = N_{D\mu}$ or $N_{D^*\mu}$ and $n_{\text{scale}} = n_{D^{**}}^0$ or $n_{D^{**}}^*$. The parameter f_{guess} , exclusive to the D^0 channel, is an unphysical scaling factor. The $D_H^{**}\mu$ templates are plotted in Fig. 8.14.

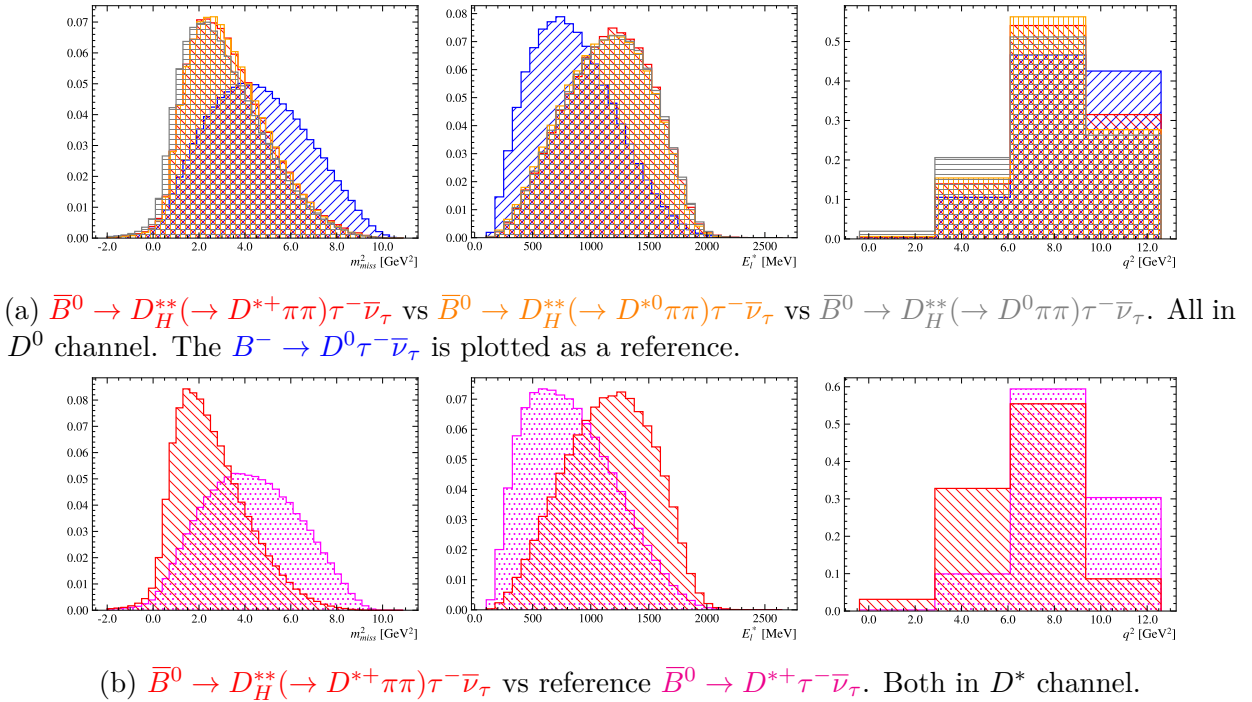


Figure 8.14: Comparison between $D_H^{**}\mu$ and $D^0\tau^-$ signal templates.

8.1.7 Feed down of $\bar{B}_s^0 \rightarrow D_s^{**} \mu^- \bar{\nu}_\mu$ modes

The contributions from the semileptonic $\bar{B}_s^0$ decays are accounted for via the $\bar{B}_s^0 \rightarrow D_{s1}^{\prime+} (\rightarrow (D^{*+} K^0 | D^{*0} K^+)) \mu^-$ and $\bar{B}_s^0 \rightarrow D_{s2}^{*+} (\rightarrow (D^{*+} K^0 | D^{*0} K^+ | D^0 K^+)) \mu^-$ simulated decays. The $D^{*+} K^0 \mu^-$ modes contribute to both channels whereas the rest to the D^0 channel only. With two templates for each fit channel, the D^0 channel templates are cocktails. The yields in the D^0 channel are constrained separately for each $D_{s^{**}i}$ species:

$$N_{D\mu} \times \left\{ \frac{f_s}{f_d} \times n_{D^{**}}^0 \times \mathcal{B}_{D_{s^{**}i}} \times N_{D_{s^{**}i}} \right\} \quad (8.21)$$

where $\frac{f_s}{f_d} = 2.6$ (the latest measured value is 2.4 at 13 TeV [44], but 2.6 is used instead) accounts for the production ratio between $\bar{B}^0$ and $\bar{B}_s^0$ mesons, $\mathcal{B}_{D_{s^{**}i}}$ is a Gaussian constraint with a 40% uncertainty, and

$$N_{D_{s^{**}i}} = \mathcal{B}(B_s \rightarrow D_{s^{**}i} (\rightarrow K^+ D^{(*)0})) \times \epsilon_{K^+ \text{ isolation}} + \quad (8.22)$$

$$\mathcal{B}(B_s \rightarrow D_{s^{**}i} (\rightarrow K^0 D^{*+})) \times r_{D^*}^{\text{isospin}} \quad (8.23)$$

is the expected yield based on measured branching fractions and selection efficiencies from MC.

The D^* channel yields for both templates are constrained with a Gaussian of a width of 24% of the expected yield of $D_{s1}^{\prime+} \mu^-$, with an additional factor $\frac{f_{s2}}{f_{s1'}} = 0.3$ to account for the relative yields for the D_{s2}^{*+} template.

Comparisons between the $D_s^{**} \mu^-$ templates can be seen at Fig. 8.15.

8.1.8 Double-charm (DDX) backgrounds

The double-charm backgrounds include contributions from $B \rightarrow D^{(*)} D_q X$ decays, where D_q represents a generic charm meson with q being a $u/d/s$ quark. When $q = s$, X is nothing, a π , or two π ; when $q = u$ or $q = d$, X is a K , a K^* , or a higher K^* resonance. For example, a muonic DDX event may come from the following decay chain: $B^0 \rightarrow D^{*+} D^- (\rightarrow K^0 \mu^- \bar{\nu}_\mu) \bar{K}^0$, whereas a tauonic one from $B^0 \rightarrow D^{*+} D_s^- (\rightarrow \tau^- \bar{\nu}_\tau)$. Many

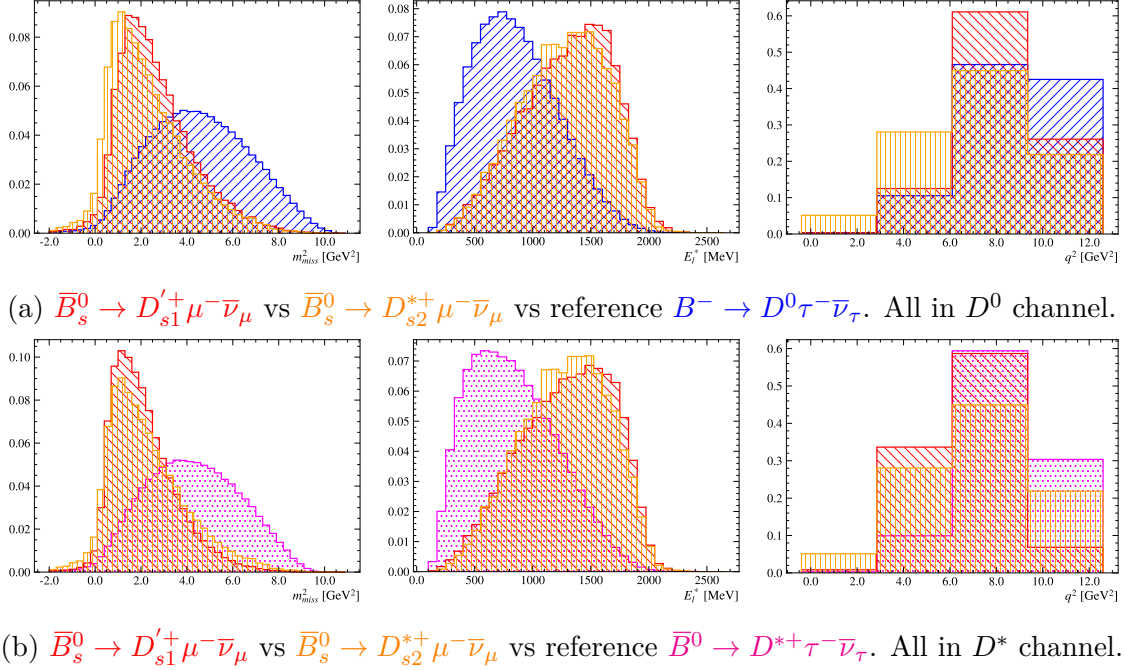


Figure 8.15: Comparisons between $D_s^{**} \mu$ templates.

of these DDX decays are poorly understood. The MC samples are generated with a phase space model for the $B \rightarrow DDX$ decays so that generated events are distributed evenly in the available phase space, without taking the resonance structure, which itself is not precisely known for $n \geq 3$ -body decays, of the Dalitz plane into account. Additionally, the $D^{(*)} \rightarrow X' l \bar{\nu}_l$ decay form factors are parameterized with ISGW2, based on a quark model, which is known to describe data insufficiently⁹ well. Both muonic and tauonic DDX fit templates have large m_{miss}^2 , similar to that of the signal, as can be seen in Fig. 8.16. This is because a large portion of the decay chain is not reconstructed. All DDX fit templates are cocktails with their relative branching fractions fixed at simulation time.

Muonic The yields that normalize muonic DDX templates are weakly constrained and are of the same form for both fit channels:

$$N_{\text{ch ref}} \times \left\{ \rho_{DDX,u/d} \times f_{DDX,q} \times f_{DDX} \right\} \quad (8.24)$$

⁹As commented in [15]: “In the current literature it is found that single-pole-dominance provides a much better description of the data than quark-model calculations.”

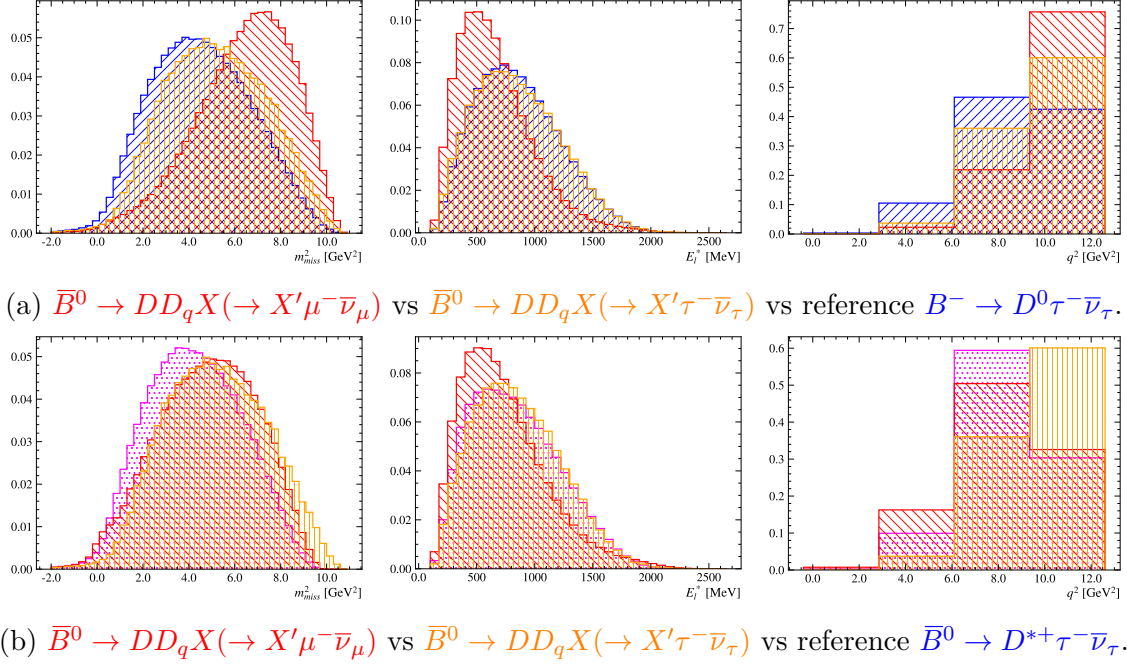


Figure 8.16: Comparison between DDX templates.

where $f_{DDX,q}$ is the expected fraction of DDX yield for a B meson type; for example, for $\bar{B}^0$ meson, $f_{DDX,q} = f_{DDX,d} = \frac{\mathcal{B}(\bar{B}^0 \rightarrow DDX) \times \epsilon(\bar{B}^0 \rightarrow DDX)}{\mathcal{B}(\bar{B}^0 \rightarrow DDX) \times \epsilon(\bar{B}^0 \rightarrow DDX) + \mathcal{B}(B^- \rightarrow DDX) \times \epsilon(B^- \rightarrow DDX)}$ where the branching fractions are deduced from MC specifications. $\rho_{DDX,u/d}$ is a fixed correction to the expected fraction and is obtained from the control fit. f_{DDX} is a free-floating parameter.

Tauonic The tauonic DDX yields are again of the same form for both channels:

$$N_{\text{ch ref}} \times \left\{ \rho_{DDX,u/d} \times f_{DDX,q,\tau} \times f_{DDX} \times \underbrace{f_{DDX,\tau} \times \rho_{\text{isolation},\tau}}_{\text{add. parameters}} \right\} \quad (8.25)$$

where $f_{DDX,\tau}$ is the expected fraction of yields for τ : $f_{DDX,\tau} = \frac{\text{expected } \tau \text{ yield}}{\text{expected } \mu \text{ yield}}$, and $\rho_{\text{isolation},\tau}$ is a Gaussian constraint to the expected fraction with a width of 30%.

8.2 Shape variations in the fit

In HISTFACTORY, the fit templates are allowed to have coherent shape variations. For a single bin, the yield after variation is defined as:

$$n_{\text{var}} = n_{\text{nom}} + \sum_i f_i(n_+, n_-, n_{\text{nom}}, \alpha) \quad (8.26)$$

where n denote yields in a single bin, and n_+, n_- are the yields when α , representing the variation of a parameter in the modeling of fit template, for example a parameter in the BGL form factor parameterization, is at ± 1 . The $n_{\pm}$ are typically evaluated at $\pm 1\sigma$ for the associated parameter such that $\alpha = \pm 1$ corresponds to a $\pm 1\sigma$ variation of the underlying parameter. Different model parameters of the same template are required by HISTFACTORY to be uncorrelated, which can be achieved by finding an orthonormal eigenbasis of the parameter covariance matrix. f is the interpolation/extrapolation function to specify variation at any value; only piecewise-linear and quadratic-interpolation linear-extrapolation functions¹⁰ in the HISTFACTORY are used. The *coherence* comes from the fact that the ± 1 yields need to be specified for *all* bins at once, in the form of histograms.

The template shape variations employed in this analysis are listed in the following paragraphs.

8.2.1 Form factor uncertainties

To take form factor uncertainties into account, the correlation matrices of the form factor parameters are transformed according to Appendix F, which produces $\pm 1\sigma$ variations in an orthonormal error eigen basis as per HISTFACTORY requirement.

Additional fit templates are generated by reweighting MC samples at the variations and are included in the fit. The number of form factor variations for D^0 , D^* are 5 and 10, consistent with their parameterizations listed in Tables 3.2 and 3.4. For D_0^* and D_1' , there are 2 variations as listed in Table 3.5 ignoring $\zeta(1)$ as it represents an overall normalization thus ineffective in shape variations. Similarly, for D_1 and D_2^* , there are 3 variations as in Table 3.6 ignoring $\tau(1)$. A sample $\pm 1\sigma$ variation of the $D_1^+ \mu^-$ template is plotted in Fig. 8.17.

¹⁰The mathematical forms of the interpolation functions f can be looked up in [45], section 4.1.

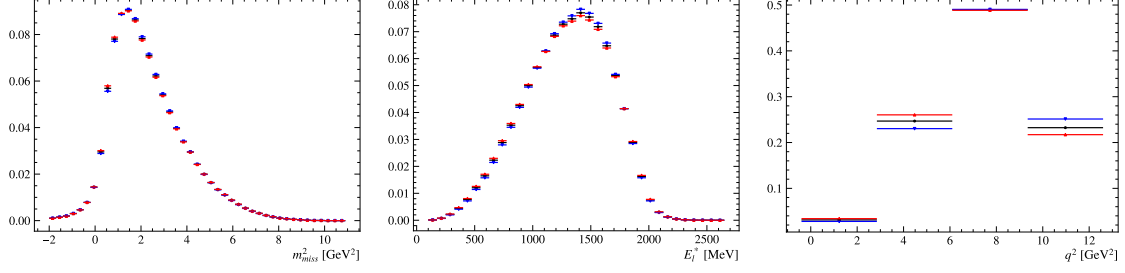


Figure 8.17: $D_1^+ \mu^-$ nominal (black), $+1\sigma$ (red), and -1σ (blue) variation fit template for the parameter τ' in the BLR form factor parameterization.

8.2.2 Data-driven Dalitz-inspired deformations to DDX model

As described in Section 8.1.8, DDX decays are not well modeled. To mitigate the physical shortcomings of the modeling of DDX decays, additional variations are introduced to give more freedom to the fit variables, allowing for a data-driven approach to determine the shapes of DDX background.

The variations are inspired by Dalitz plots: when a single scalar decays into three scalars, the phase space is uniquely determined by two variables. Hence the phase space can be deformed via one of the Dalitz variables. Here the *true* invariant mass of the $DD^{(*)}$ pair is chosen as the proxy to deform the phase space linearly and quadratically as a weight, defined in Eqs. (8.27) and (8.28) evaluated at $\alpha = \pm 1$ as usual:

$$w_{\text{lin}}^{\pm} = 1 \pm 2 \cdot \left(\sqrt{\frac{m_{DD^{(*)}}^2 - m_{\text{min}}^2}{m_{\text{max}}^2 - m_{\text{min}}^2}} - 1/2 \right) \quad (8.27)$$

$$w_{\text{quad}}^{\pm} = (1 \mp 1) \pm 8 \cdot \left(\sqrt{\frac{m_{DD^{(*)}}^2 - m_{\text{min}}^2}{m_{\text{max}}^2 - m_{\text{min}}^2}} - 1/2 \right)^2 \quad (8.28)$$

Note that the range for both sets of weights are between $[0, 2]$, $m_{\text{min}} = m_{D^{(*)}} + m_D$, and $m_{\text{max}} = m_B - m_K$. The $B^- \rightarrow DDX$ variations in the D^0 channel are shown in Fig. 8.18 as an example.

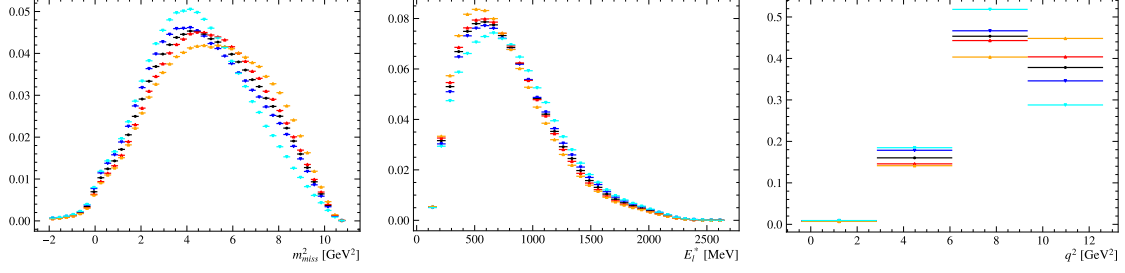


Figure 8.18: Variations of the $B^- \rightarrow DDX$ template, with the nominal template colored black, the linear $\pm$ colored red/blue, and the quadratic $\pm$ colored orange/cyan.

8.2.3 K^* weight up/down in DDX templates

An DDX event is considered to contain a K^* if $(p_B - p_{D_{1st}} - p_{D_{2nd}})^2 > 680$ MeV. The K^* events are weighted up (to 2) and down (to 0) and are included as a variation to float the relative fractions between K and K^* in the fit, as the relative selection efficiencies for these modes are not precisely known. The $B^- \rightarrow DDX$ variations of K^* in the D^0 channel are shown in Fig. 8.19

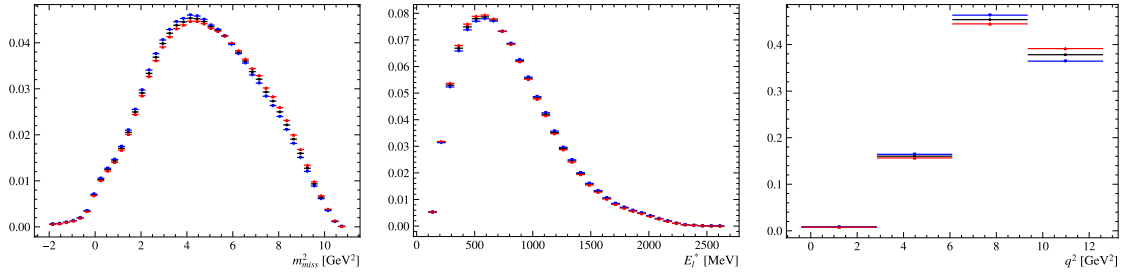


Figure 8.19: Variations of K^* contributions of the $B^- \rightarrow DDX$ template. Color coding: nominal, up, and down.

8.2.4 Data-driven phenomenological deformation of D_H^{**} model

The D_H^{**} MC samples are generated with the ISGW2 form factor model, which is known to be insufficient to describe data: similar to the run 1 analysis, the q^2 in the 2OS sample fits poorly with the ISGW2 model. Deformations based on the true q^2 in D_H^{**} with weights of the form

$$w_{\pm} = 1 \pm 2 \left(\frac{q_{\text{true}}^2 - m_{\mu}^2}{8 \times 10^6} - 0.5 \right) \quad (\text{kinematic variables in MeV}) \quad (8.29)$$

are introduced to make the q^2 distribution either harder (+, up) or softer (−, down), providing more freedom in fit variables which allows for a data-driven approach to determine the shapes of these modes. A comparison between the nominal and varied $B^0 \rightarrow D_H^{**} \mu$ templates can be seen at Fig. 8.20.

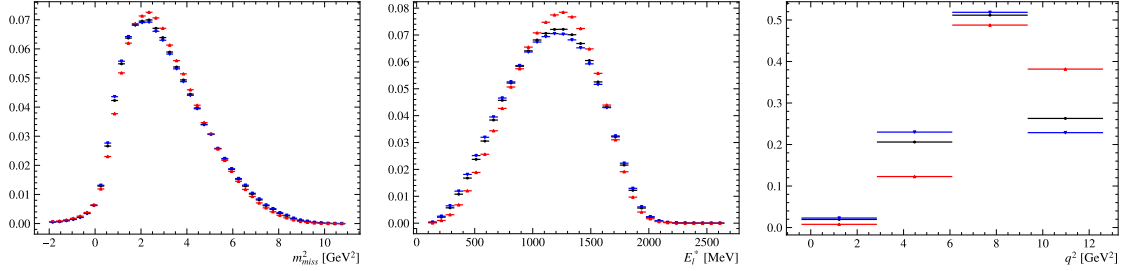


Figure 8.20: Variations of $B^0 \rightarrow D_H^{**} \mu$ based on the true q^2 of the decay processes. Color coding: nominal, **up**, and **down**.

8.2.5 Effect of kaon, pion decay-in-flight in muon misID template

K and π have some probability of decaying to a μ between upstream and downstream tracking (decay-in-flight, DiF), which leads to a mis-measurement of the track momentum, and authentic signals in the muon stations. In our fake μ sample, however, only events *without* any μ signal are selected, thus *none* of the mis-identified K and π has the DiF effect.

To account for the DiF effect, we use a smearing approach on top of the unfolding described in Section 8.1.1:

1. Select K, π passing truth-matching and nominal μ selection requirements from a MC sample.
2. Find the true and reconstructed momenta of K, π . Use ratios of the smear factor $_i \equiv \frac{\text{reco}_i}{\text{true}_i}$ as the smearing factor for the i -component of the 3-momentum, with the assump-

tion that MC models DiF sufficiently correctly.

3. Put all the smearing factors in an array, and for the fake μ track of each misID event, randomly select one to smear the reconstructed momentum:

$$p_{i,\text{smeared}} = p_{i,\text{reco}} \times \text{smear factor}_i \quad (8.30)$$

4. Estimate the B meson rest frame momentum with the rest frame approximation technique.
5. Recompute fit variables $\{m_{\text{miss}}^2, q^2, E_\mu^*\}_K, \{m_{\text{miss}}^2, q^2, E_\mu^*\}_\pi$ with the new fake μ track momentum. These sets of variables will be refer as K -smeared and π -smeared.

When building the fit variable templates, for each event of tagged species $\hat{t}'$, the nominal fit variables $\{m_{\text{miss}}^2, q^2, E_\mu^*\}$ and the two sets of the smeared fit variables are weighted according to the following rules:

- K -smeared: $\epsilon(\hat{t}' \rightarrow K)$
- π -smeared: $\epsilon(\hat{t}' \rightarrow \pi)$
- Nominal: $1 - \epsilon(\hat{t}' \rightarrow K) - \epsilon(\hat{t}' \rightarrow \pi)$

where the efficiencies are obtained from the unfolding procedure. A comparison between unsmeared and smeared fit variables in the misID templates without applying any skim cut is shown in Fig. 8.21, with the comparison of ISO skim fit templates due to DiF smearing in the D^0 channel shown in Fig. 8.22

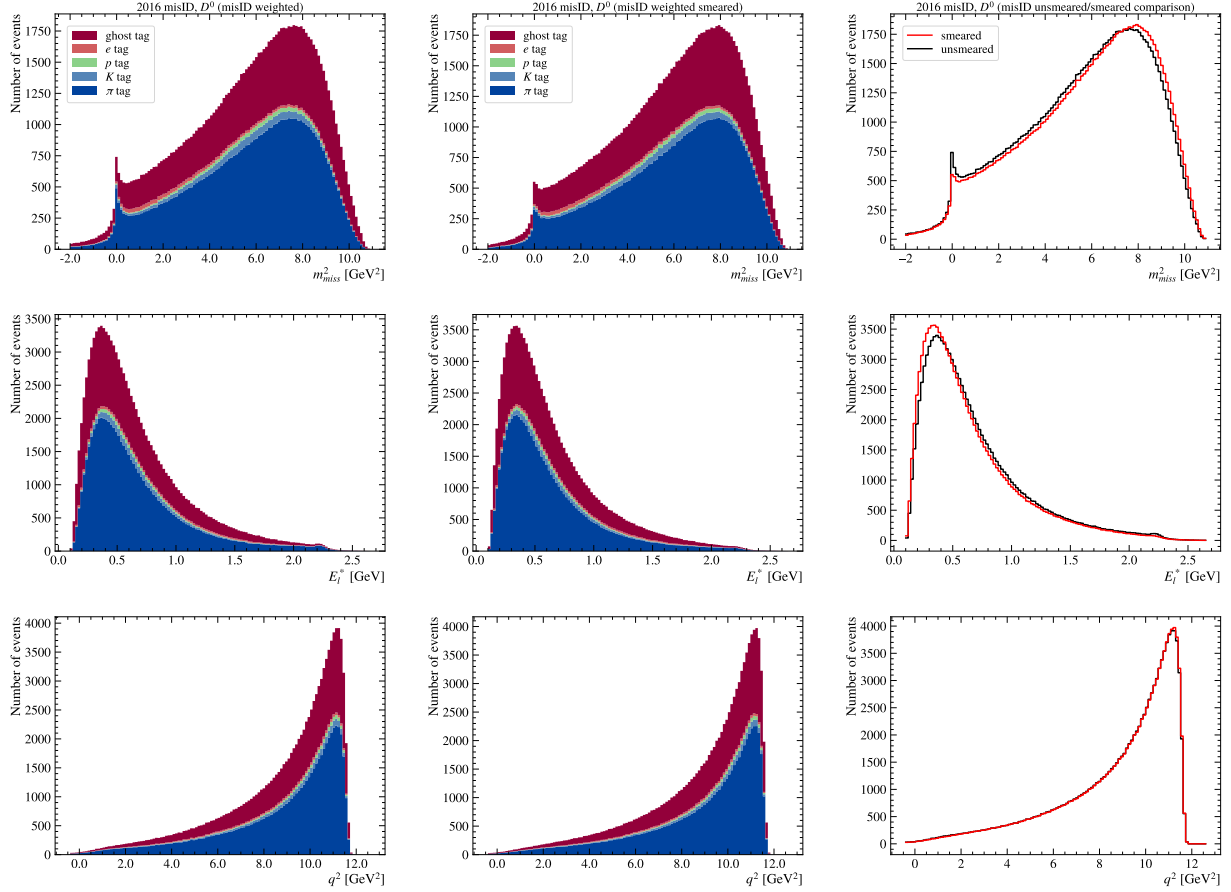


Figure 8.21: Effect of DiF on fit variables. No skim cut applied. The plots displayed here are contributions of 2016 D^0 fake μ control samples ($B^- \rightarrow D^0 t^-$) to the fit variables. The first column is unsmeared misID fit variables m_{miss}^2 , E_μ^* , q^2 ; the second column smeared; the third column comparison between the overall shape of unsmeared/smeared fit variables. Note that the total number of events is not guaranteed to be conserved, because smearing may put fit variables outside of the acceptance range.

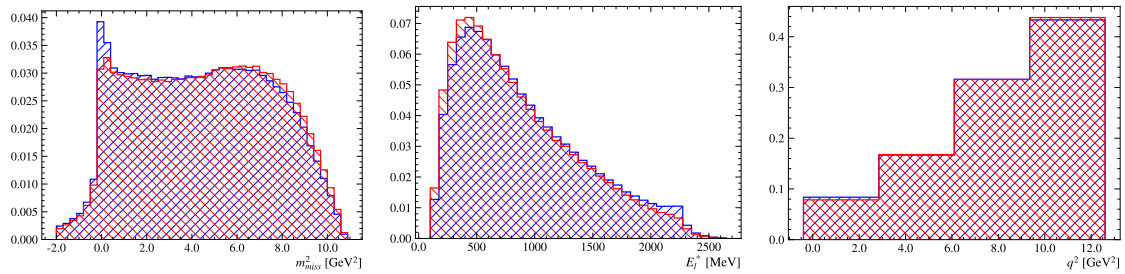


Figure 8.22: A comparison between the **smeared** and **unsmeared** ISO skim misID templates in the D^0 channel.

8.3 Fit results

The fitted yields are grouped as follows: $D^0\mu$, $D^{*+}\mu$, $D^{*0}\mu$, signal, four $1P$ D^{**} , heavy D^{**} , DD , D_s , combinatorial backgrounds, and μ misID. The combined yields of the groups of the signal fit are reported in Table 8.7 for both fit channels, with the signal yields blinded (replaced with a random number).

Table 8.7: Fitted yields of the signal fit.

Group	Yields		Group	Yields
norm. ($D^0\mu$)	445,095		norm. ($D^{*+}\mu$)	427,783
norm. ($D^{*+}\mu$)	115,124		sig. (random)	5
norm. ($D^{*0}\mu$)	1,176,363		D^{**}	37,181
sig. (random)	14		D^{**} heavy	12,804
D^{**}	167,968		D_s	3,730
D^{**} heavy	38,147		DD	33,588
D_s	5,285		comb. bkg.	16,906
DD	96,236		misID	8,175
comb. bkg.	20,596			
misID	56,596			
(a) D^0 channel			(b) D^* channel	

The pre-control fit results are plotted in Appendix A.5; the control fit results are displayed in Figs. 8.23 to 8.25 for D^0 channel; Figs. 8.26 to 8.28 for D^* channel; the signal fit results in Figs. 8.29 and 8.30 for the D^0 and D^* channel, accordingly.

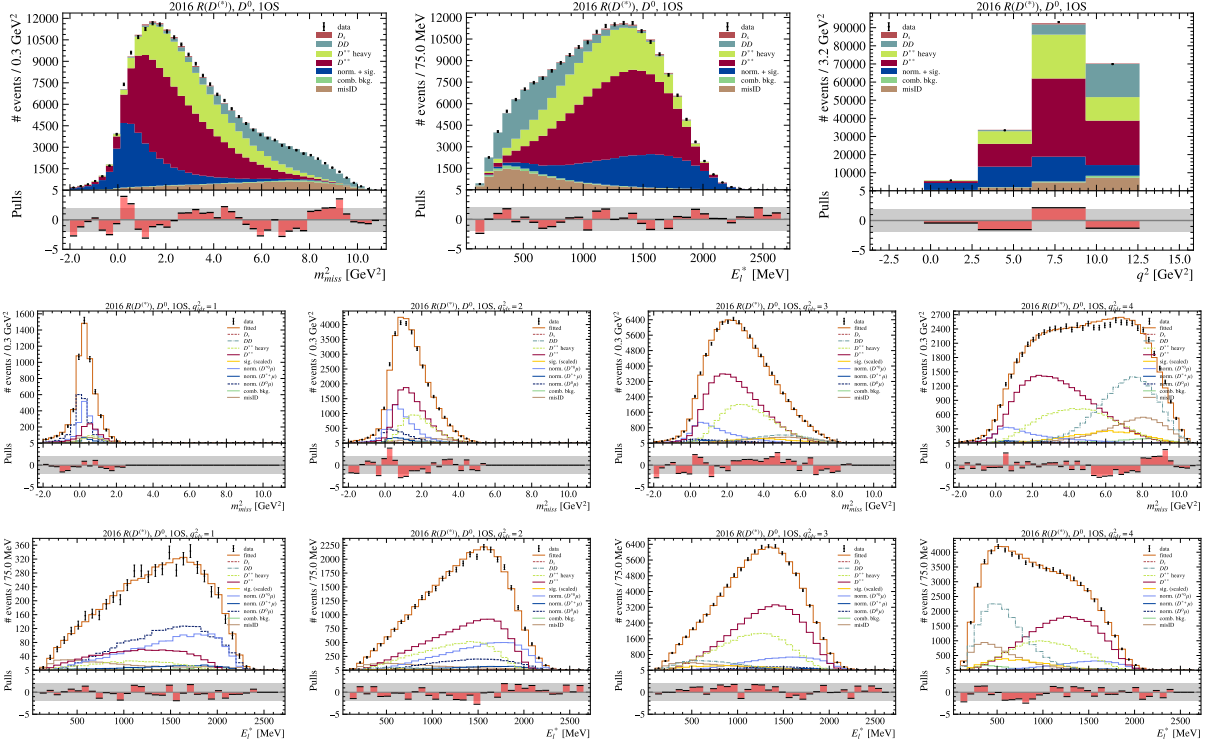


Figure 8.23: Control fit for 10S sample, D^0 channel.

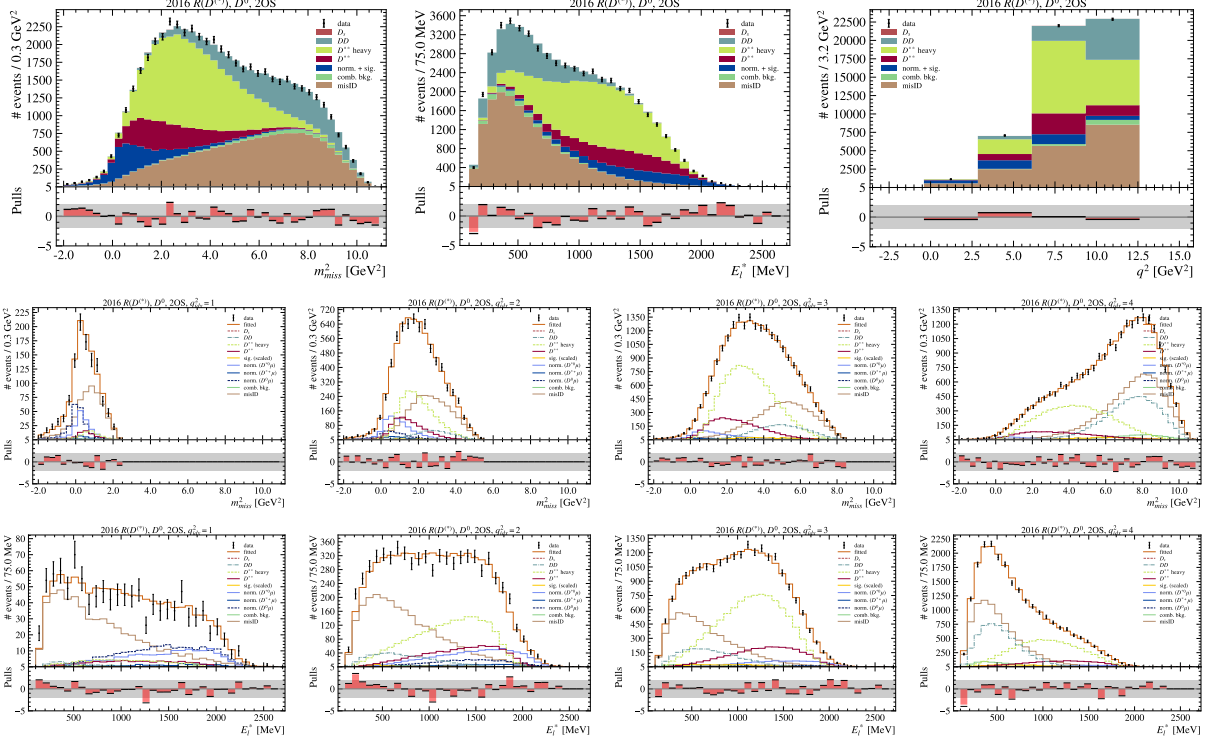


Figure 8.24: Control fit for 20S sample, D^0 channel.

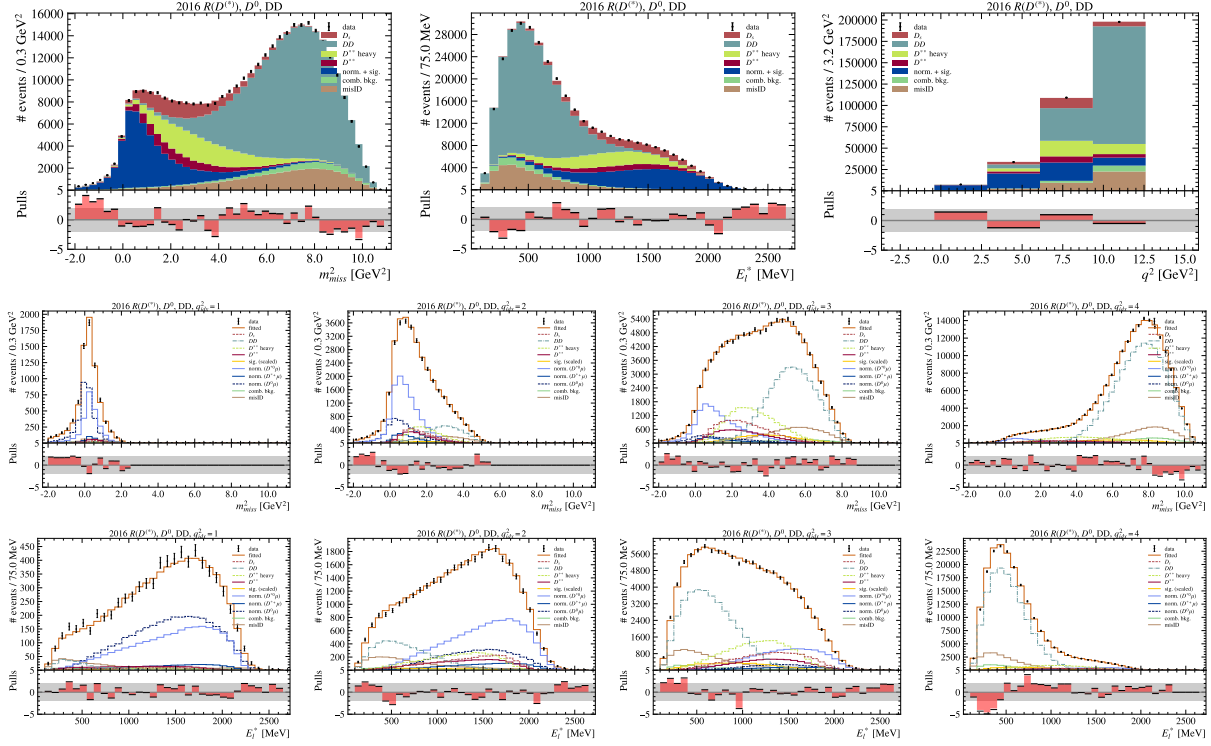


Figure 8.25: Control fit for DD sample, D^0 channel.

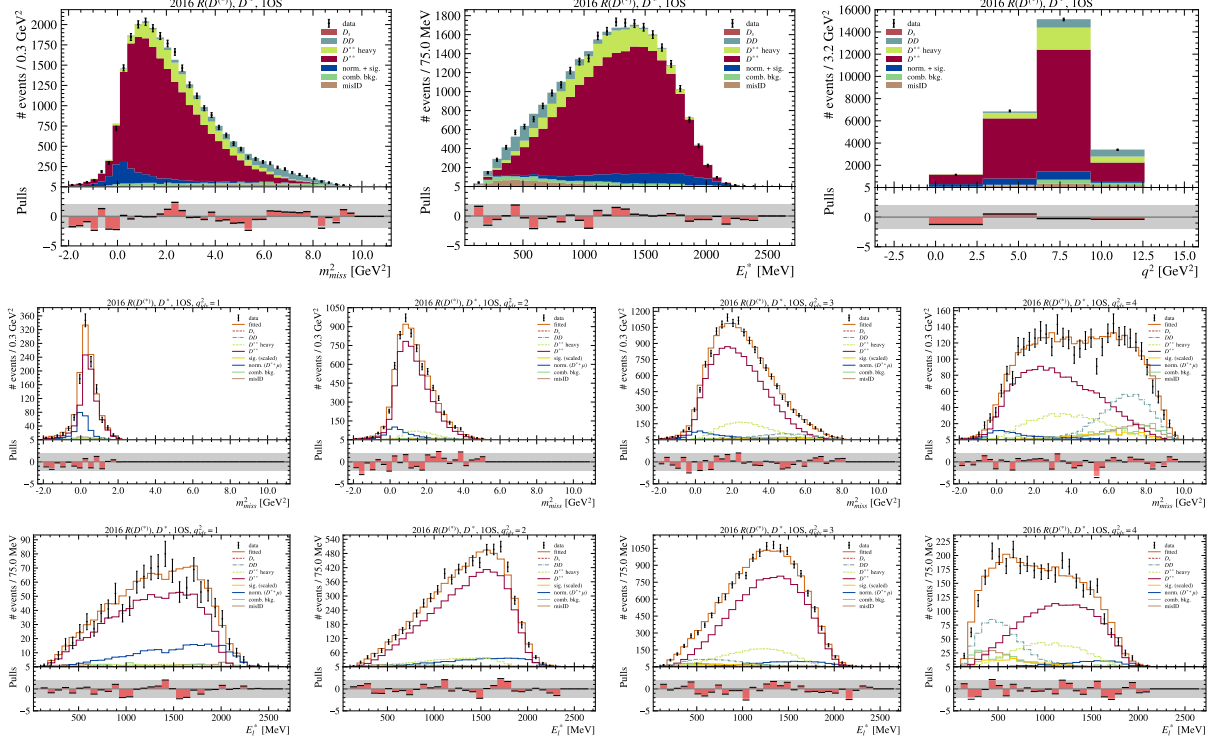


Figure 8.26: Control fit for 1OS sample, D^* channel.

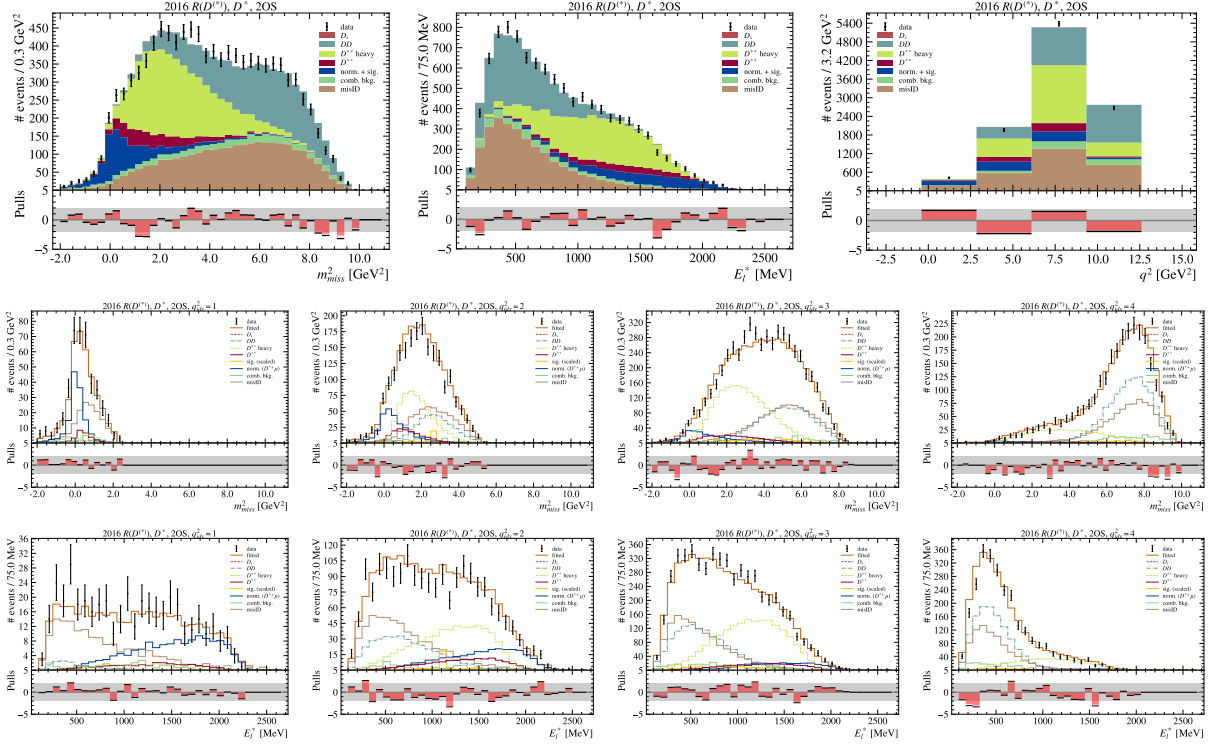


Figure 8.27: Control fit for 2OS sample, D^* channel.

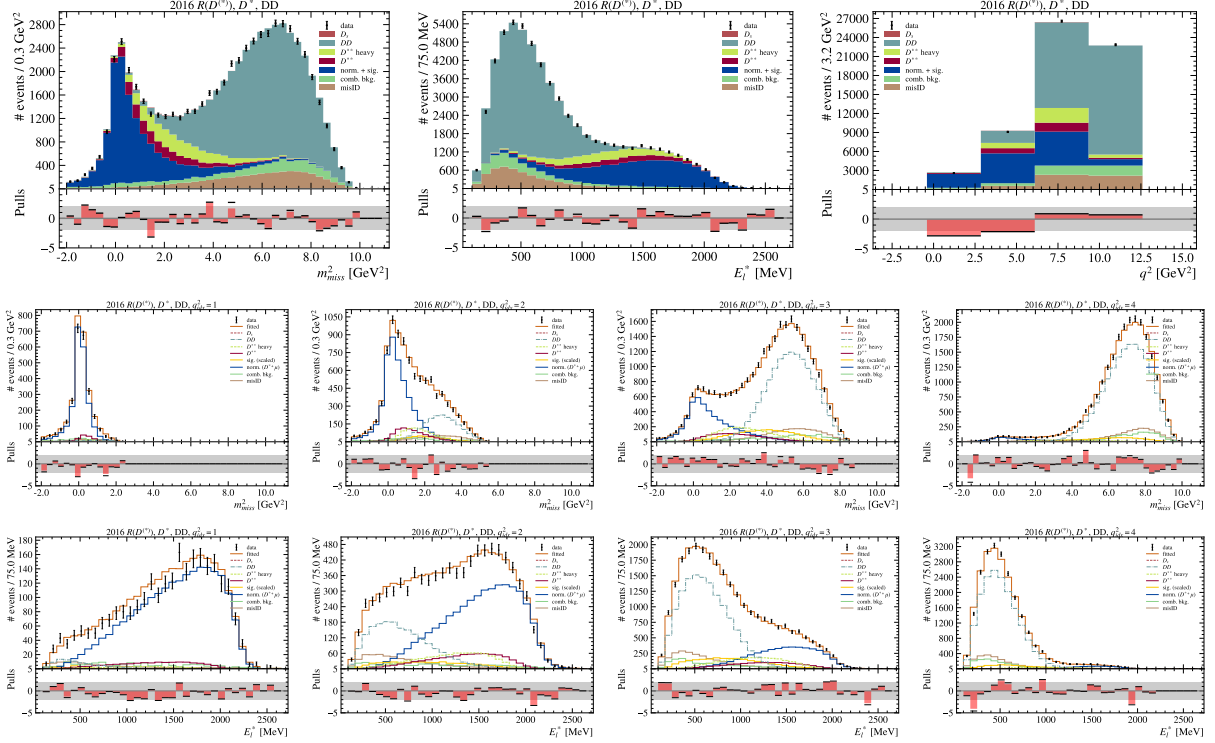


Figure 8.28: Control fit for DD sample, D^* channel.

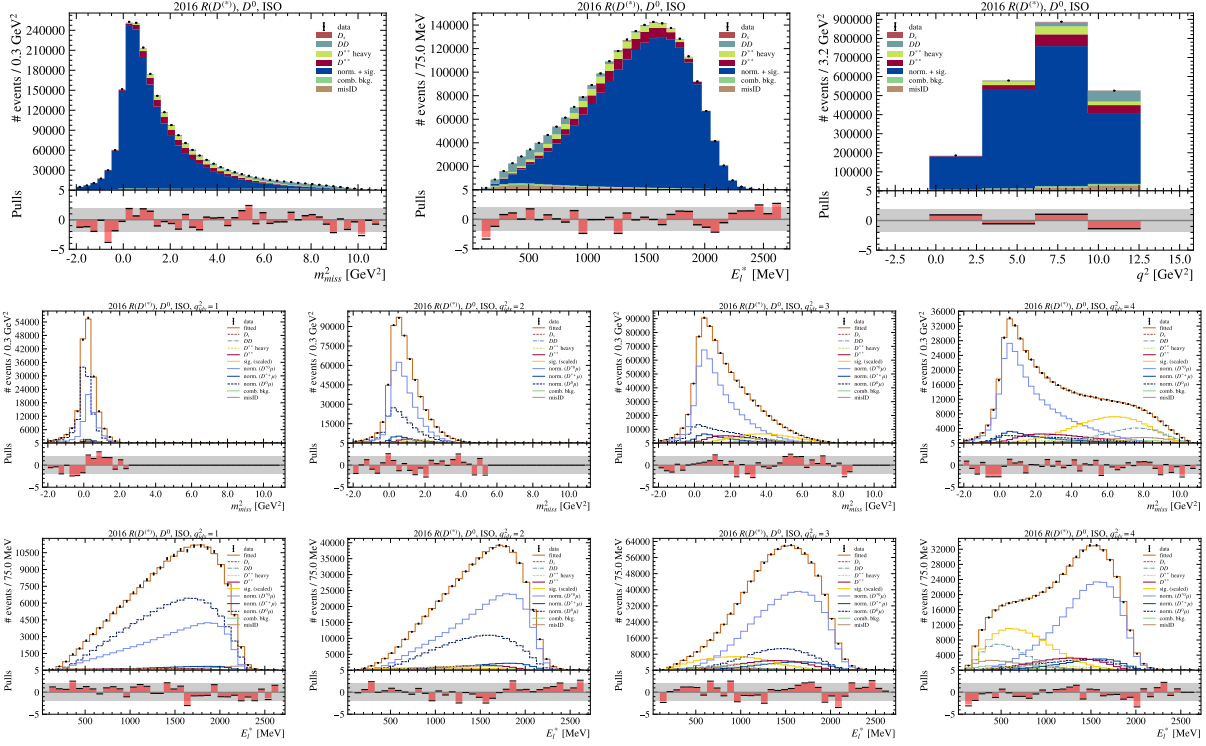


Figure 8.29: Signal fit for ISO sample, D^0 channel.

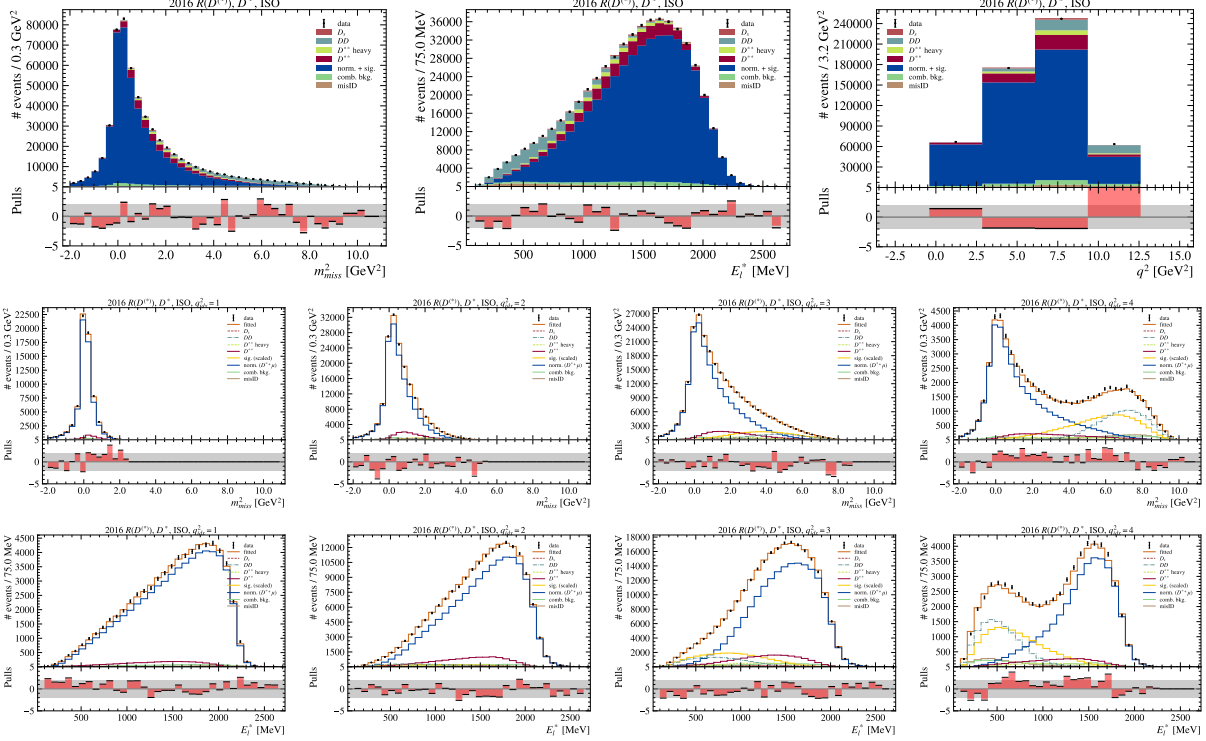


Figure 8.30: Signal fit for ISO sample, D^* channel.

Chapter 9: Systematic uncertainties

The major systematic uncertainties are evaluated as listed in Table 9.1. The corresponding uncertainties from the run 1 analysis evaluated with the HISTFACTORY fitter are also provided as reference. Unless specified, all uncertainties are *absolute* uncertainties. All uncertainties are *multiplied by 100* for convenience. To have a sense on the relative size of the systematic uncertainties, suppose the central value for $\mathcal{R}(D) = 0.441$, the same value as reported in the run 1 analysis [2]. One of the leading systematic uncertainties in the D^0 channel is the “control sample shape parameters” (Section 9.4), which is at 0.96×10^{-2} , corresponding to a relative uncertainty of about 2.2%. The overall run 1 relative uncertainty is about 20% for $\mathcal{R}(D)$, and 10% for $\mathcal{R}(D^*)$.

The remainder of the chapter is dedicated to explain the evaluation method for each tabulated systematic uncertainty. These uncertainties can be loosely classified into the following groups: Form factors (Section 9.1), under-constrained τ components (Sections 9.2 and 9.3), uncertainties on the modeling of the background components (Sections 9.4 to 9.7), uncertainties on employed algorithms (Section 9.8), theoretical corrections not included in the nominal fit (Section 9.9), and uncertainties on data/MC corrections (Sections 9.10 and 9.11).

Table 9.1: Evaluated systematic uncertainties. Run 1 uncertainties are provided as a reference inside parentheses. All uncertainties are absolute, not relative. The additive uncertainties are sorted in descending order by their run 1 significance. Currently the MC statistical uncertainties are not taken into account.

Source	$\sigma_{\mathcal{R}(D^*)} [\times 10^{-2}]$	$\sigma_{\mathcal{R}(D)} [\times 10^{-2}]$	Correlation
$B \rightarrow D^{(*)}l^- \bar{\nu}_l$ form factors	0.53 (0.58)	0.79 (2.37)	-0.71 (-0.80)
$B \rightarrow D^{**}l^- \bar{\nu}_l$ form factors	0.25 (0.78)	0.62 (1.01)	-0.85 (-0.10)
Control sample shape parameters*	0.46 (0.87)	0.96 (4.36)	(-0.49)
Λ_b^0 backgrounds	TBD (0.73)	TBD (1.16)	0.00 (0.00)
DD model dependence*	TBD (0.63)	TBD (0.74)	0.00 (0.00)
$B \rightarrow D^{**}\tau^- \bar{\nu}_\tau$ bkg	0.18 (0.17)	0.49 (0.30)	-0.94 (0.78)
$B \rightarrow D^{(*)}D_s^{**}(\rightarrow \tau^- \bar{\nu}_\tau)X$ bkg	0.18 (0.25)	0.62 (1.21)	-0.79 (0.59)
μ misID unfolding algorithm*	TBD (0.74)	TBD (1.19)	1.00 (1.00)
Coulomb correction to $\mathcal{R}(D^{*+})$ vs. $\mathcal{R}(D^{*0})$ *	0.27 (0.17)	0.21 (0.3)	-1.00 (-1.00)
μ misID decay-in-flight correction	0.14 (0.06)	0.73 (0.16)	-0.95 (0.29)
Combinatorial background shape*	TBD (0.03)	TBD (0.18)	0.00 (0.00)
Vertex resolution correction*	TBD (0.03)	TBD (0.21)	TBD (-0.25)
Data/MC corrections (add.)*	TBD (0.40)	TBD (0.75)	0.00 (0.00)
Data/MC corrections (mul.) [†]	TBD $\times \mathcal{R}(D^*)$ ($1.16 \times \mathcal{R}(D^*)$)	TBD $\times \mathcal{R}(D)$ ($0.91 \times \mathcal{R}(D)$)	0.00 (0.00)
$\mathcal{B}(\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau)$ (ext. mea.) [†]	$0.23 \times \mathcal{R}(D^*)$ ($0.23 \times \mathcal{R}(D^*)$)	$0.23 \times \mathcal{R}(D)$ ($0.23 \times \mathcal{R}(D)$)	1.00 (1.00)
Total systematics	TBD (2.03)	TBD (6.16)	TBD (-0.22)
MC statistical	0.91 (1.33)	2.62 (4.00)	-0.72 (-0.30)
Statistical	1.94 (1.88)	6.08 (6.01)	-0.07 (-0.53)
Total (sys. + stats.)	TBD (3.07)	TBD (9.49)	TBD (-0.35)

* This uncertainty is not part of the nominal uncertainties reported by the fit.

[†] This is a multiplicative uncertainty.

9.1 $B \rightarrow (D^0|D^*|D^{**})l^- \bar{\nu}_l$ form factors

The uncertainties due to $D^{(*,**)}$ form factors affect the shapes of the fit variables, as discussed in Section 7.1; the FFs are listed in Sections 3.2.1 to 3.2.3. The uncertainties due to these parameters can be determined in the following manner:

1. Perform a fit with the nominal settings and call the fitted value $\mathcal{R}(D^{(*)}) = v + \sigma_+ - \sigma_-$.
2. Fix the variations (α parameters) associated the form factor parameters at the fitted value. This removes the constrains due to these α parameters in the likelihood thus

eliminates all uncertainties due to these parameters.

3. Repeat the fit with the α parameters fixed at the fitted value. Denote the new fitted value $\mathcal{R}(D^{(*)})' = v' + \sigma'_+ - \sigma'_-$.
4. The systematic uncertainty due to the form factors are obtained from subtracting uncertainties of $\mathcal{R}(D^{(*)})$ and $\mathcal{R}(D^{(*)})'$ in quadrature. More specifically, due to the asymmetric nature of the error, the systematic uncertainty is calculated with Eq. (9.1):

$$\sigma_{\text{sys}} = \sqrt{\sigma_+ \sigma_- - \sigma'_+ \sigma'_-} \quad (9.1)$$

The systematic uncertainty due to D^0 and D^* form factors is 0.79 for $\mathcal{R}(D)$ and 0.53 for $\mathcal{R}(D^*)$; uncertainties on D^{**} form factors leads to an uncertainty of 0.62 for $\mathcal{R}(D)$ and 0.25 for $\mathcal{R}(D^*)$.

9.2 $B \rightarrow D^{**}\tau^-\bar{\nu}_\tau$ background

As discussed in Section 8.1.5, the $B \rightarrow D^{**}\tau^-\bar{\nu}_\tau$ tauonic yields are not well-constrained by external measurements, with their yields allowed to fluctuate with a Gaussian constraint of $\sigma = 0.33$ (so that an increase from the nominal yield N to $1.33N$ would result in a $1\text{-}\sigma$ penalty in the corresponding log-likelihood function). The systematic uncertainty introduced by the Gaussian constrains is evaluated based on the differences between the fit uncertainties resulting from floating and fixing these parameters at their best fit values, as described in Section 9.1. The uncertainty due to tauonic D^{**} backgrounds is 0.49 for $\mathcal{R}(D)$, and 0.18 for $\mathcal{R}(D^*)$.

9.3 $B \rightarrow D^{(*)}D_s^{**}(\rightarrow \tau\bar{\nu}_\tau)X$ background

Similar to Section 9.2, the $B \rightarrow D^{(*)}D_s^{**}(\rightarrow \tau\bar{\nu}_\tau)X$ tauonic yields are only loosely constrained with a Gaussian whose $\sigma = 0.3$, as discussed in Section 8.1.8. Using the same

method of taking the quadrature difference of floating and fixing these parameters at their fitted values, the uncertainty due to tauonic DDX backgrounds is 0.62 for $\mathcal{R}(D)$, and 0.18 for $\mathcal{R}(D^*)$.

9.4 Control sample shape parameters

As discussed in Chapter 8, in the signal (ISO) fit, the shape parameters for the background templates, namely D^{**} , D_H^{**} , and DDX , are loaded from the fitted values of the control fit, with the D_H^{**} and DDX parameters held fixed and the D^{**} shape parameters, which are form factors, constrained. These parameters are referred as the *control sample shapes parameters*.

To propagate the uncertainties from the control sample shape parameters to the signal fit, a bootstrap procedure is performed which generates alternative parameters based on the correlations among these parameters. Since the systematic uncertainties from the D^{**} form factors are studied separately, to avoid double counting, a *conditional covariance matrix*, which removes the covariance *purely due to* D^{**} FFs, is used in the bootstrap procedure. As an example, for the D^{**} form factors, the bootstrap permits variations due to correlations between them and the DDX/D_H^{**} shape parameters *only*. The removal is carried out with the following procedure:

1. Start from the full 16-dimensional space of shape variations parameters, all obtained from a fit in the control sample: 8 DDX shape parameters, 3 D_H^{**} parameters, and 5 D^{**} form factor parameters. Denote the covariance matrix as M .
2. Block decompose M into:

$$M = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \quad (9.2)$$

such that Σ_{11} , 5 dimensional, corresponds to the raw correlations between D^{**} form factors variations and Σ_{22} , 11 dimensional, DDX and D_H^{**} shape parameters.

3. The covariance purely among the D^{**} FFs, or formally the *conditional covariance* of D^{**} form factor variations conditioned on the fitted values and uncertainties of the DDX and D_H^{**} shape parameters, denoted as C , is obtained with:

$$C = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \quad (9.3)$$

4. The reduced covariance matrix R is obtained by subtracting C from Σ_{11} .

With the reduced covariance matrix R , toys of shape parameters satisfying the covariance can be generated by Cholesky decomposition and random vectors sampled from independent Gaussian distributions. Each set of the parameters is loaded in the signal fit as the best fitted control parameters and the fit is repeated normally, with the fitted value of $\mathcal{R}(D^{(*)})$ recorded. The root-mean-square of the distribution of $\mathcal{R}(D^{(*)})$ fitted this way is taken as the systematic uncertainty, which is 0.96 for $\mathcal{R}(D)$, and 0.46 for $\mathcal{R}(D^*)$.

9.5 DD model dependence

In the nominal fit, two variations on m_{DD^*} , a key proxy variable to the 3-body decay phase space, are introduced on DDX templates to allow for a data-driven approach to determine the shape of these templates, as described in Section 8.2.2.

To estimate the dependence on the choice of proxy variable to the phase space, alternative proxy variables, namely m_{DK} and p_D in the B rest frame, are deformed linearly and quadratically, the nominal DDX templates are replaced with these alternative variations, and a refit is performed. The quadratic differences between the fitted uncertainties with the nominal and alternative DD model are taken as the systematic uncertainty.

The variations of m_{DK} take the same form of Eqs. (8.27) and (8.28), with $m_{DD^{(*)}}$ replaced by m_{DK} . The mass limits are also replaced: $m_{\min} = m_{D^{(*)}} + m_K$, $m_{\max} = m_B - m_D^{(*)}$.

9.6 μ misID decay-in-flight correction

As discussed in Section 8.2.5, the decay-in-flight (DiF) effect, which stands for a K or π decaying into a μ between the upstream (VELO-TT) and the downstream (T-stations) tracking system, leading to a mis-measurement of the track momentum and an authentic μ signature, is included as a shape variation to allow the fit to determine the strength of the effect. The uncertainty due to this variation is evaluated with results from floating vs. fixing at the fitted value for the associated α parameter, which is 0.73 for $\mathcal{R}(D)$ and 0.14 for $\mathcal{R}(D^*)$.

9.7 Combinatorial background shapes

The nominal B combinatorial templates include a linear correction that depends on m_B , which is needed because this background is derived from the wrong-sign control sample and the raw yields between the wrong-sign and the right-sign samples in the B mass upper side band do not match, as discussed in Section 8.1.2.

However, as also discussed in the same section, the B combinatorial is empirically modeled as an exponential of the form $Ae^{-\lambda m_b}$, where A, λ are fitted from data. We assume both the wrong-sign and right-sign combinatorial backgrounds have the same functional form, therefore, the correction factor is the ratio between two exponential functions, which itself is an exponential, with the leading mass-dependent order correction linear.

The uncertainty on the linear correction factor is quantified by repeating the fit with the nominal B combinatorial templates substituted with the ones generated with the exponential correction, as shown in Figs. 8.8 and 8.9. The quadrature difference between $\mathcal{R}(D^{(*)})$ is taken as the associated systematic uncertainty. The exponential correction is *not taken as the nominal* because the existing data cannot resolve A, λ to high precision.

9.8 μ misID unfolding algorithm

Two main sources of uncertainties are identified to affect the μ misID unfolding algorithm, an algorithm to estimate the contributions from charged particles mis-identified as μ (Section 8.1.1):

- The limited statistics of the misID data control sample
- The accuracy of the efficiency for a ghost track to pass the μ selection requirements $\epsilon(g_{\text{acc}} \rightarrow \hat{\mu})$ (ghost mis-identification efficiency), since this is modeled from a MC sample.

The unfolding procedure utilizes an iterative algorithm that uses the unfolded results from a previous step as the input to the next one. It stops after a set number of iterations and a χ^2 statistics is reported to measure the change between the last iteration [43]. More iterations lead to an *over-fit* to the supplied data, so the nominal misID template has the number of iterations set to 5, the recommended value.

To evaluate the uncertainties due the limited statistical power, the misID algorithm is intentionally configured to *over-fit* the control sample by setting the number of iterations to 20. A fit is then repeated with the over-fit misID templates and the quadratic difference of the fit uncertainties are taken as the systematic uncertainty.

Similarly, for the ghost mis-identification efficiency, ghost tracks in an alternate MC sample with different underlying decay modes (an inclusive $J/\psi\mu$ sample, instead of the $\bar{B}^0 \rightarrow D^{*+}\mu^-\bar{\nu}_\mu$ sample discussed in Section 6.2) are used instead to produce alternative misID templates then repeat the fit.

9.9 Coulomb correction to $\mathcal{R}(D^{*+})$ vs. $\mathcal{R}(D^{*0})$

The MC simulation lacks the corrections from Coulomb interactions between the lepton and the charged hadrons in $\bar{B}^0 \rightarrow D^{*+}l^-\bar{\nu}_l$ decays. This correction leads to three

effects:

1. A deformation of the $B \rightarrow D^{*+}l^{-}\bar{\nu}_l$ fit templates, particularly for the tauonic $\bar{B}^0 \rightarrow D^{*+}\tau^{-}\bar{\nu}_\tau$ decay.
2. A shift in the efficiency ratio between $D^{*+}\tau$ and $D^{*+}\mu$ decays.
3. Breaking of the isospin symmetry between $\mathcal{R}(D^{*+})$ and $\mathcal{R}(D^{*0})$

The first two effects are checked in [15] to be negligible compared to existing additive and multiplicative uncertainties. The third effect results in about 1% relative enhancement in $\mathcal{R}(D^{*+})$ compared to $\mathcal{R}(D^{*0})$, as reported in [46], which is implemented by setting:

$$\left. \frac{\eta_{D^{*0}}}{\tilde{\eta}_{D^{*+}}} \right|_{\text{new}} = \frac{\eta_{D^{*0}}/\tilde{\eta}_{D^{*+}}}{1.01} \quad (9.4)$$

then repeat the fit, leading to an uncertainty of 0.21 for $\mathcal{R}(D)$, and 0.27 for $\mathcal{R}(D^*)$.

9.10 Vertex resolution correction

As discussed in Section 7.5, the mis-modeling of the B flight vector in the MC simulation is corrected by smearing on the azimuthal angle of the flight vector. To assess the uncertainty due to this smearing, an additional variation to either enhance or reduce the smearing effect by assigning a larger/smaller weight to events with larger smearing angles, which in turn widens or narrows the B vertex and the m_{miss}^2 resolution, is introduced *for every single MC template*.

Since such a smearing changes the shape of fit variables for all MC templates, the variation introduces a non-trivial correlation to every shape variation parameter. To evaluate the uncertainty *purely due to the variation on smearing*, all other shape variations are fixed at the best fit values before a refit is performed. The results are subtracted in quadrature as usual to obtain the systematic uncertainty due to the smearing.

9.11 Data/MC corrections

Had the data/MC corrections been perfect, after an iteration of the fit is performed, the low m_{miss}^2 region between the data and the post-fit cocktail (defined in Section 7.3) should be statistically consistent, because this normalization enriched region is insensitive to other effects due to modeling of physics backgrounds. In reality the data/MC corrections are not perfect, and the uncertainty due to this is assessed by doing an additional iteration of the final reweighting (explained in Section 7.4), regenerating the fit templates, and then repeating the fit.

The effect of such a correction is broken into an additive and a multiplicative component, which the latter affect the efficiency ratios (the 4 η parameters discussed in Section 8.1.4).

Chapter 10: Conclusions

As a proof-of-concept, this 2016 $\mathcal{R}(D^{(*)})$ analysis demonstrates that both the general fit strategy, that is, obtain the background parameters with separate control samples then import them into the signal fit, as well as most of the methods developed in the run 1 pathfinder analysis, such as the isolation BDT, the μ BDT, and the unfolding algorithm, can be carried over for run 2. The additional challenge of emulating the triggers due to the use of tracker-only MC (to have enough MC statistics) is considered to be fully solved with the method described in this thesis, with the emulation procedure portable to the year 2017 and 2018. Also, the form factor reweighting procedure has been updated to use **Hammer**, the current state-of-art form factor reweighting program, which is more robust and easier to maintain. It is expected that the 2017 and 2018 data can be processed with the same workflow developed in this analysis.

The overall fit quality is good with reasonable pulls, but, there is a large discrepancy in the high- q^2 bin for the D^* signal (ISO) fit and others in the DD control fits that need to be investigated further. Additionally, during the study of systematic uncertainties, it was discovered that by fixing the nuisance parameters of a given source of uncertainties to their fitted values, often the fitter does not yield the same minimum. The stability of the fitter needs to be improved. Therefore, the current total uncertainties cannot be determined precisely.

Since only the 2016 data is included, smaller uncertainties are to be expected once the 2017 and 2018 data is included, provided that the q^2 bin discrepancy is resolved and the fitter is converging at a more stable minimum. It is expected that the total statistical

uncertainties from the LHCb run 2 data (2016–2018) would be 7% for $\mathcal{R}(D)$ and 3% for $\mathcal{R}(D^*)$. (as a reference, the run 1 analysis has a $14\%_{\text{stat}} \pm 15\%_{\text{sys}}$ uncertainties for $\mathcal{R}(D)$ and $6.4\%_{\text{stat}} \pm 8.2\%_{\text{sys}}$ uncertainties for $\mathcal{R}(D^*)$) If the total systematic uncertainties are controlled to be below 8% for $\mathcal{R}(D)$ and 4% for $\mathcal{R}(D^*)$, the analysis could be the most precise measurement to $\mathcal{R}(D^{(*)})$ in the world.

Part II

Upgrade of the LHCb Upstream Tracker

Chapter 11: Overview of the LHCb upgrade

The run 2 of LHC reached its conclusion at the end of 2018, when it entered the second long shutdown (LS2) period. Due to luminosity levelling, since 2012, the LHCb has been collecting data at about 2 fb^{-1} integrated luminosity per year when the experiment is in a steady state, with an instantaneous luminosity of about $4 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$, far below which the LHC is capable of providing¹. More precise measurements require more data which the current LHCb detector is unable to provide, due to the fact that the detector has a limited readout rate of 1 MHz: even if the instantaneous luminosity is increased, with current hardware triggers², to maintain a constant readout rate of 1 MHz, the thresholds for p_T and E_T have to increase, which largely cancels the benefit of running at higher luminosity. As shown in Fig. 11.1, the hadronic trigger yields are saturating even at run 2 luminosity ($4 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$).

To remove the readout bottleneck, the LHCb detector readout was upgraded during 2019–2023 to achieve a readout rate of 40 MHz, the LHC bunch crossing rate, without any hardware trigger. By exposing all low-level variables to purely software-based high-level triggers with more exclusive selection criteria, events can be categorized into exclusive signal decay modes more efficiently with reasonable write throughputs [48]. After the upgrade, the detector will operate at an instantaneous luminosity of $2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$, a factor of 5 gain compared to that of run 2, without any hardware trigger, making LHCb the first (hardware) trigger-less detector in a proton collider.

¹As a reference, the CMS experiment reached a maximum instantaneous luminosity of $1.5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ in 2016.

²As discussed in Section 4.3, the hardware triggers require the event to be above a certain p_T/E_T threshold.

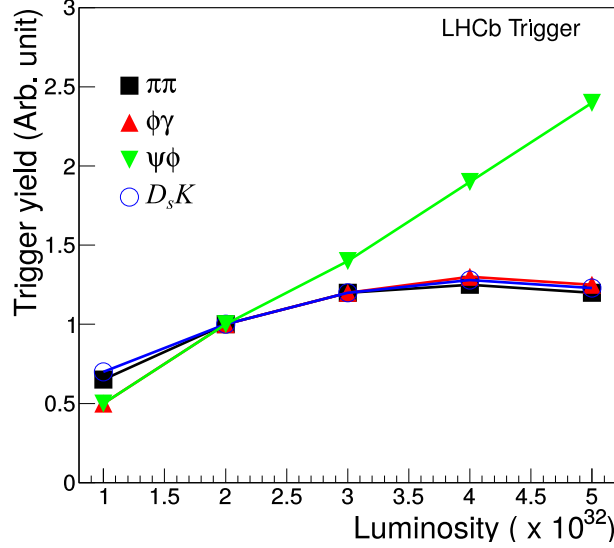


Figure 11.1: L0 trigger yields normalized to that at $2 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$. The hadronic triggers are already saturated at run 2 luminosity. Taken from [47].

Aside from the upgrade of the readout system and trigger scheme, the tracking system is upgraded to provide better resolution and more radiation tolerance to sustain the increased luminosity. The subdetectors mainly used in L0 trigger, namely Muon station 1, SPD, and PS, are removed altogether to reduce material in front of the calorimeters to improve energy resolution. The RICH1 detector which is responsible for PID is upgraded to reduce occupancy³ thus maintaining good PID efficiency. The calorimeters (ECAL and HCAL) and the muon system will not undergo substantial upgrades [50]. The changes in the LHCb detector between run 1, 2 and run 3 are indicated in Fig. 11.2.

The rest of the chapter provides an overview of the LHCb upgrade in LS2. Section 11.1 provides an comparison of the trigger schemes between run 2 and run 3. The tracking and PID subdetectors are upgraded to provide better tracking resolution and PID performance; these will be discussed in Section 11.2 and Section 11.3, respectively.

³Defined as the fraction of detected photons over the total number of channels [49].

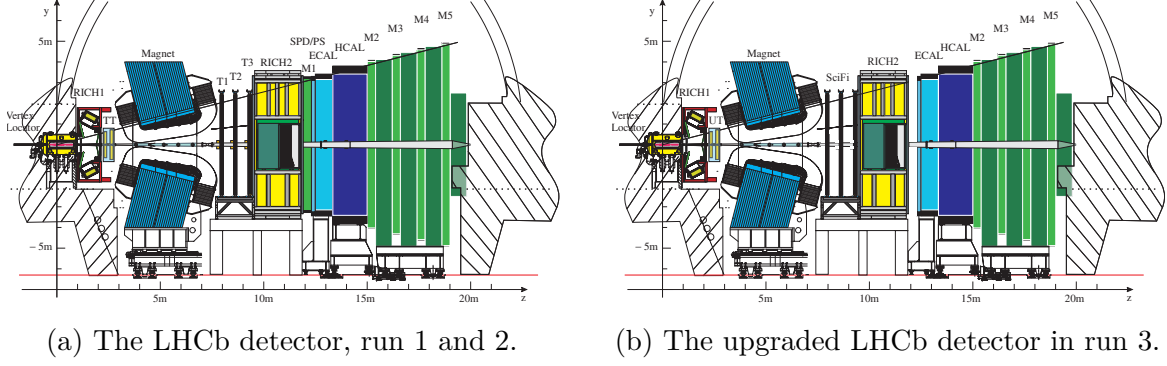


Figure 11.2: The LHCb detector before (left) and after (right) LS2 upgrade. Note the upgrade on the tracking system: $TT \rightarrow UT$, T-stations $\rightarrow$ SciFi; the removal of subdetectors relevant to L0 trigger only: the M1 muon station and SPD/PS have been removed.

11.1 Trigger schemes in run 2 and run 3

As can be seen in Fig. 11.3, the main differences between the run 2 and run 3 trigger schemes are: the limited detector readout rate of 1 MHz is increased to 40 MHz⁴; the storage write throughput is increased from 0.6 GB/s to 10 GB/s. The implications of these differences are discussed in more details in the following text.

The need for a higher readout rate at 40 MHz is due to the saturation of the hadronic L0 trigger: this trigger requires the maximum energy E_T deposited in the calorimeters to be at least 3.7 GeV (for 2016) to reduce the triggering rate at an acceptable level; the threshold is already a significant portion of the B meson mass (~ 5.3 GeV). Any further increase in luminosity requires an increase of the threshold, which removes a significant portion of the signal events that contain B mesons, resulting in an almost constant signal yield [47].

With a 40 MHz readout rate at 13 TeV and a pile-up of 5.2 (corresponding to the run 3 conditions), the run 2 inclusive trigger strategy⁵ of rejecting trivial backgrounds with minimal bias cuts on few variables in an inclusive trigger is non-sustainable for LHCb run 3:

⁴While the bunch crossing rate is 40 MHz, not every bunch is filled with particles. Hence the pp inelastic collisions happen at a frequency of about 30 MHz.

⁵To be clear, the run 2 trigger strategy contains both inclusive triggers, such as an inclusive B trigger selecting a significantly displaced high p_T track forming a displaced vertex with 1–3 other tracks (referred as “HLT2 topological trigger”), and exclusive triggers.

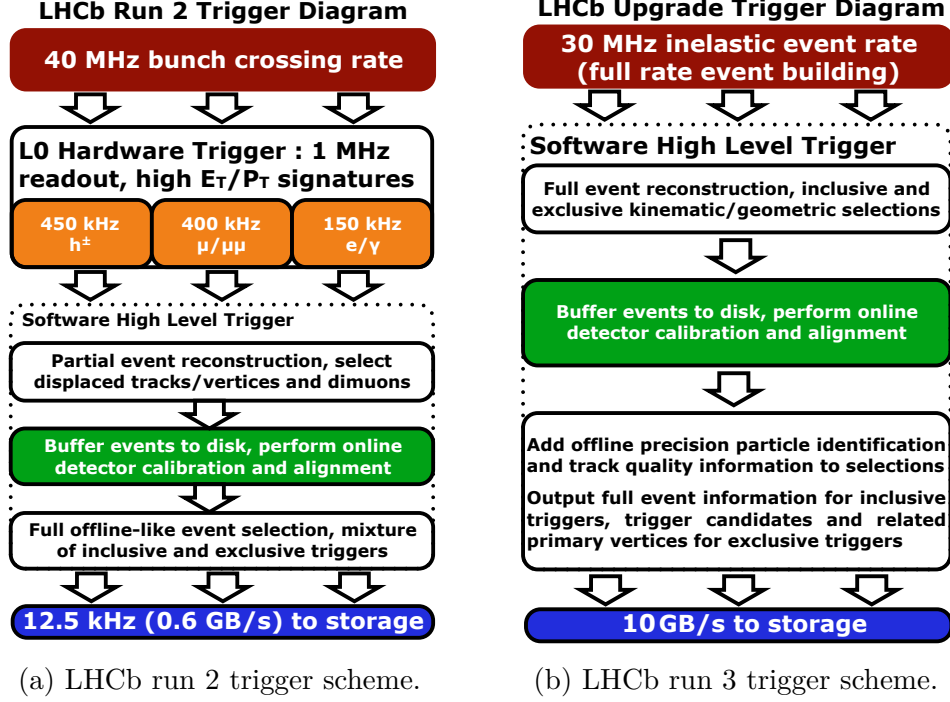


Figure 11.3: Comparison between run 2 and run 3 trigger schemes.

assuming an event size of 100 kB⁶, the write throughput for beauty hadrons is 27 GB/s, for charm it is 80 GB/s, for light long-lived particles (such as K_S^0 and Λ^0) 26 GB/s [48].

However, the write throughput is limited to 10 GB/s due to constraints from the LAN connection speed between the LHCb on-site online system (responsible for data acquisition) and the CERN's storage center (Tier0) in Meryin⁷, Switzerland. According to Ref. [51], there are more than 20 pairs of fibers, each capable of transmitting data at 1 GB/s, which, in theory, can provide a bandwidth greater than 20 GB/s. Still, in the same reference, a maximum write throughput of 10 GB/s is considered, perhaps due to concerns on robustness. In any case, the 80 GB/s write throughput for a run 2-like inclusive trigger strategy far exceeds the bandwidth capacity, mainly due to the fact that with run 3 conditions, a significant portion of collisions contains genuine B mesons. Instead, it is preferable to employ a trigger strategy that categorize different types of signals and then apply selection

⁶The original text uses kb (kilo-bit), which does not add up to 80 GB/s for a 800 kHz rate. The size should be in kB (kilo-byte), as supported in [51, 52]. The 800 kHz number is reported in [53].

⁷A small city near Geneva.

cuts with as much information from an event as possible, as early as possible [48]. Therefore, the full events are exposed directly to software HLT1 triggers which reconstruct and select tracks with higher efficiency.

In general, utilizing a software-only trigger allows more exclusive selections to reduce bandwidth, provides maximum flexibility, and makes higher-level physics variables, such as the invariant mass of a reconstructed vertex, accessible to the trigger, such that cuts can be applied directly on these variables, reducing the need of cutting on proxy variables thus the need of modeling acceptance effects offline [48].

As a side note, the run 3 HLT1 trigger is entirely GPU-based⁸, fully implemented in CUDA [52], as track reconstruction in an event is ideally suited for parallel execution⁹, to which GPU is more advantageous than CPU: while a CPU typically has highly frequencies, meaning it can execute more “commands” (instructions) in a given amount of time, a GPU, on the other hand, while having lower frequency, can execute the same command on multiple input data elements¹⁰ in parallel.

A sketch of the HLT1 workflow is the following: For a single bunch crossing, readout from all subdetectors are received separately by about 250 event builders in the counting room, each equipped with 2 GPUs, which assemble the readout fragments into a single event with a size of ~ 100 kB; multiple events are aggregated into a Multiple Event Packet (MEP) before feeding into a single GPU for execution of the full HLT1 trigger algorithms. Only events passing HLT1 requirements are copied back from the GPU and sent out for further processing by HLT2 algorithms, reducing the data throughput from 40 Tb/s to 1–2 Tb/s. As the HLT1 workflow implies, the GPU-based solution also saves bandwidth: Had the HLT1 is purely implemented with conventional CPUs, the full detector readout are required

⁸GPU: Graphics Processing Unit, as opposed to CPU: Central Processing Unit.

⁹Because a single event contains $\mathcal{O}(100)$ or more tracks, and each track are reconstructed with the same algorithms.

¹⁰Modern CPU can also do this (Single Instruction, Multiple Data, SIMD), but the level of parallelism is not comparable to that of GPU: a modern CPU may allow simultaneous precessing of 8 data elements for a single SIMD core; for a 8-core CPU, it is capable of processing 64 data elements in parallel. Whereas a single GPU (e.g. NVIDIA Tesla V100) can handle as many as 5120 data elements!

to be sent out to the HLT farms. Instead, a GPU-based HLT1 trigger allows triggering at event builder level, providing a factor of 30 reduction in output bandwidth [52], as indicated by the data transfer paths of a CPU- and GPU-based HLT1, shown in Fig. 11.4.

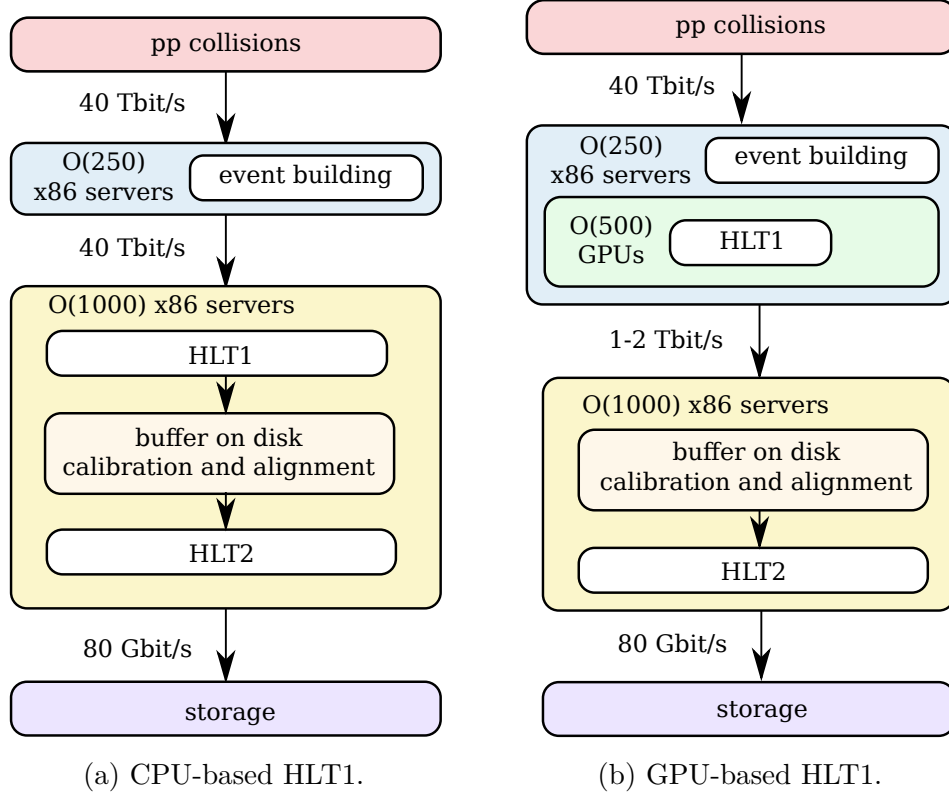


Figure 11.4: Comparison between data paths for a CPU- and GPU-based HLT1 trigger.

A recent benchmark, displayed in Fig. 11.5, shows that with about 200 high-end commercially available GPUs, the 30 MHz throughput of the LHCb readout can be sustained, which is far below the 500 (2 card per event builder for 250 builders) limit. It also shows that the current HLT1 implementation scales linearly with the performance of the GPU. As more performant GPU emerges, more powerful HLT1 algorithms may be implemented with less number of GPUs.

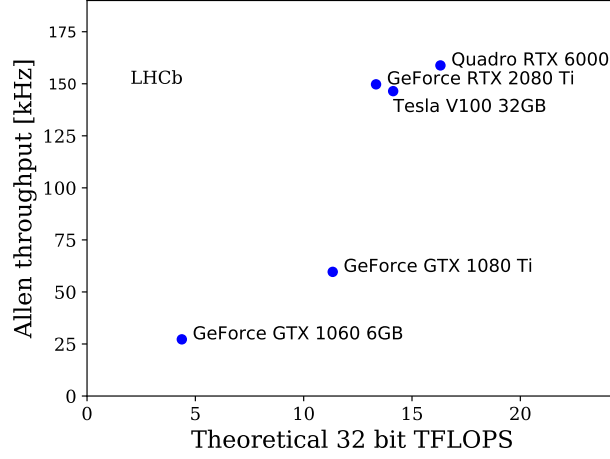


Figure 11.5: GPU-based HLT1 throughputs on many GPUs with respect to their reported peak 32-bit TFLOPS (Tera Float Point Operation Per Second) performance. Taken from [54].

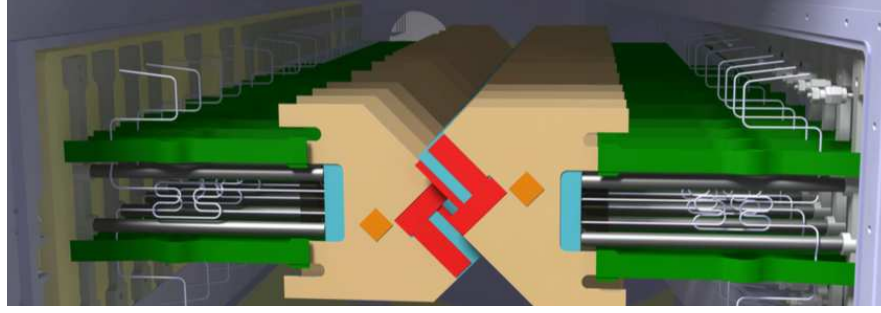
11.2 Tracking system

All subdetectors of the tracking system undergo significant upgrade to improve resolution and radiation hardness, aside from the required upgrade to their readout system.

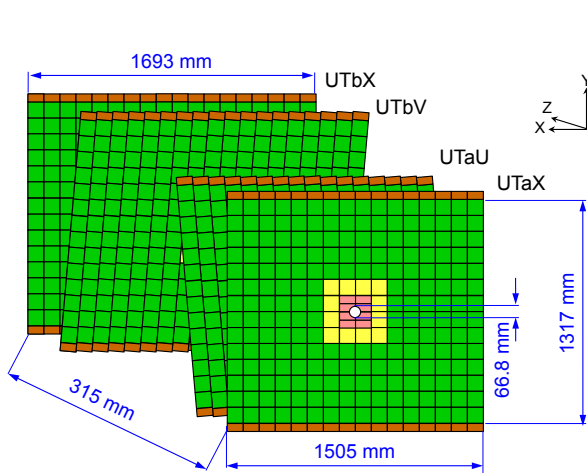
The upgraded VELO detector is now a 26-layer pixel¹¹ tracker with a pixel size of $50 \times 50 \mu\text{m}^2$, placed even closer to the beam at 5.1 mm, and has less detector materials, from 4.6% to 1.7% radiation lengths, before particles hit the first tracking layers which is crucial to improve the resolution on the impact parameters (IP) [55]. A single layer of the upgraded VELO detector is shown in Fig. 11.6a. The improvement in IP resolution based on simulation is shown in Fig. 11.7.

The Tracker Turicensis (TT) is replaced by the Upstream Tracker (UT). Located immediately before the Magnet, UT will be used for reconstructing long-lived particles decaying after VELO. It improves the momentum resolution of long tracks and trigger timing, as well as reducing fake tracks [50, 56]. Still based on silicon-strip technology, the sensors in UT has a pitch of $180 \mu\text{m}$ in the peripheral, and a smaller pitch of $95 \mu\text{m}$ compared to that of TT ($183 \mu\text{m}$) around the beam pipe. The sensors also sit closer to

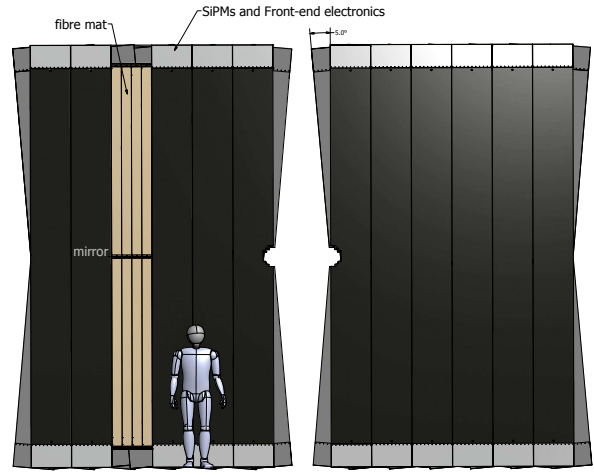
¹¹As opposed to silicon strip, which consists of thin long metal strips implant on silicon wafer, pixel are typically of finer square shape.



(a) Upgraded VELO.



(b) The UT, which replaces TT.



(c) The SciFi tracker, which replaces T-stations.

Figure 11.6: Upgraded tracking system of the LHCb detector.

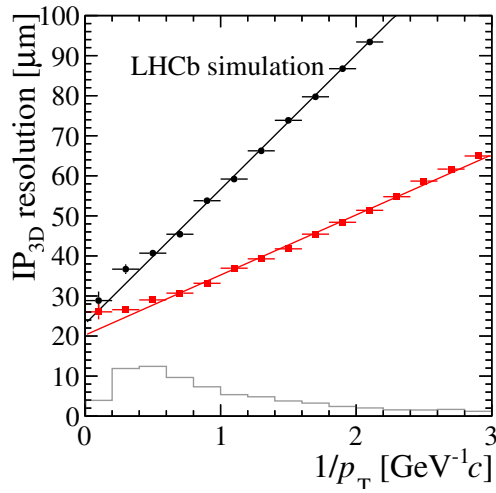
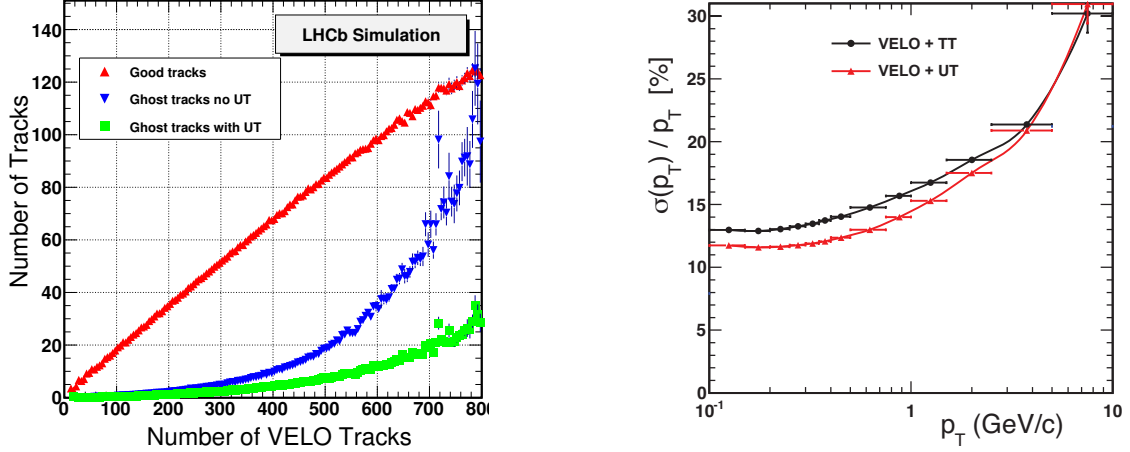


Figure 11.7: The upgraded VELO provides better IP resolution, based on simulation. The IP resolution for current (**black**) and upgraded (**red**) VELO are both evaluated at a pile-up of 7.6 and a center of mass energy of 14 TeV, a typical run 3 condition.

the beam. UT has identical geometry to that of TT: 4 planes of sensors arranged in a straight/stereo configuration¹², as shown in Fig. 11.6b. Compared to TT, UT provides sufficient radiation hardness, reduces occupancy in the critical region near the beam, and improves p_T resolution [57], as displayed in Fig. 11.8. More details regarding UT will be provided in Chapter 12.



(a) Requiring hits on at least 3 out of 4 UT planes for tracks within the UT fiducial volume reduces fake tracks significantly.

(b) Compared to TT, UT provides better p_T resolution, especially at lower p_T .

Figure 11.8: UT provides better p_T resolution to the long tracks and reduces fake (ghost) tracks. Figures taken from [56].

The Scintillating Fiber tracker (SciFi), replacing the T1–T3 stations (both IT and OT), consists of 3 detectors with 4 layers each arranged in a straight/stereo configuration, is placed after the Magnet to provide track and momentum reconstruction for the downstream tracks. Each layer of SciFi, 1.35 mm thick and has a surface area of $6 \times 5 \text{ m}^2$, is composed of minimally 10 fiber modules which is built from two groups of 6-layer 2.4 m long plastic scintillating fibers with 250 μm diameter arranged vertically, separated by a mirror in the middle. The fibers are readout with Silicon PhotoMultipliers (SiPM), located at the top and bottom of the detector, with a pixel size of $57 \times 62 \mu\text{m}^2$ [50, 58]. The operating principle of SciFi is illustrated in Fig. 11.9. The spatial resolution of the SciFi is better than 100 μm

¹²Technically termed as a x - u - v - x configuration, with u/v layers tilted by $\pm 5^\circ$ from the vertical axis. A more detailed description can be found in Section 12.1.

[58].

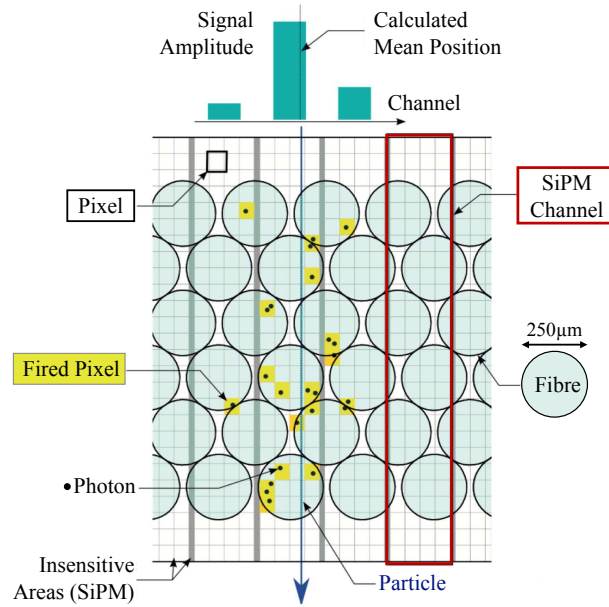


Figure 11.9: Operation principle of SciFi, illustrated with a single fiber module viewed from top/bottom. Taken from [59].

11.3 Particle identification system

In run 2 the occupancy for RICH1 can reach 30%, the upper limit of good PID performance. A factor of 5 increase in luminosity in run 3 would certainly degrade the PID performance. To mitigate the problem, the photon distribution is spread over a wider area by increasing the radius of the spherical mirrors to 3.7 m (as opposed to 2.7 m in run 2). In addition, the focal plane is moved back which leads to an enlargement of the Cherenkov ring size [49, 60].

The photon detection system for both RICH detectors are upgraded from Hybrid Photon Detectors (HPDs) to Multi-anode PhotoMultiplier Tubes (MaPMTs) to provide single-photon counting ability at 40 MHz [60].

Chapter 12: Upgrade of the Upstream Tracker

Located in the fringing fields in front of the LHCb dipole Magnet, the Upstream Tracker (UT) plays an important role in improving the tracking quality and speeding up trigger decisions: combining hits from VELO, SciFi, and UT provides the best p_T resolution and reduces the rate of random combinations from VELO and SciFi hits to form ghost tracks to about one fourth. Furthermore, combining hits from VELO and UT removes low- p_T tracks and narrows down the hit search window in SciFi, allowing the HLT track reconstruction algorithms to meet the timing constraints under the run 3 condition with high multiplicities and high pile-ups (> 5 on average).

The LHCb group at the University of Maryland is responsible for the design, testing, and commissioning of the UT peripheral and power regulating electronics. An overview of the UT, discussed in Section 12.1, summarizes the specifications and capabilities of these electronics, together with those of the silicon sensors. The author developed a quality assurance (QA) procedure for the Data Control Board (DCB), the board responsible to aggregate sensor readout and to control the UT detector, which is described in Section 12.2. A review of the software processes used in the DCB QA as well as of the online system of the LHCb experiment, responsible for data acquisition and control of the detector, is included in Section 12.3.

12.1 Overview of the UT detector

The UT consists of 4 planar detection layers with full LHCb acceptance coverage. As shown in Fig. 12.1, the first and last layers are straight, whereas the second and third are

rotated¹ by a stereo angle of $\pm 5^\circ$ to form a $x-u-v-x$ configuration. Such geometry allows for a precise determination of the position of a particle in the $x-y$ plane: a silicon-strip sensor consists of long, thin metal strips on top of a silicon base, as a charged particle hits a strip on the sensor, only a *one-dimensional* position is measured. Therefore, the first x plane can only measure the x position. A rotated sensor provides additional positional measurement with another strip that is not co-linear with the first one, forming a “X” pattern with the first strips and the $x-y$ two-dimensional position is determined as the intersection between the two. Such a process is illustrated in Fig. 12.2.

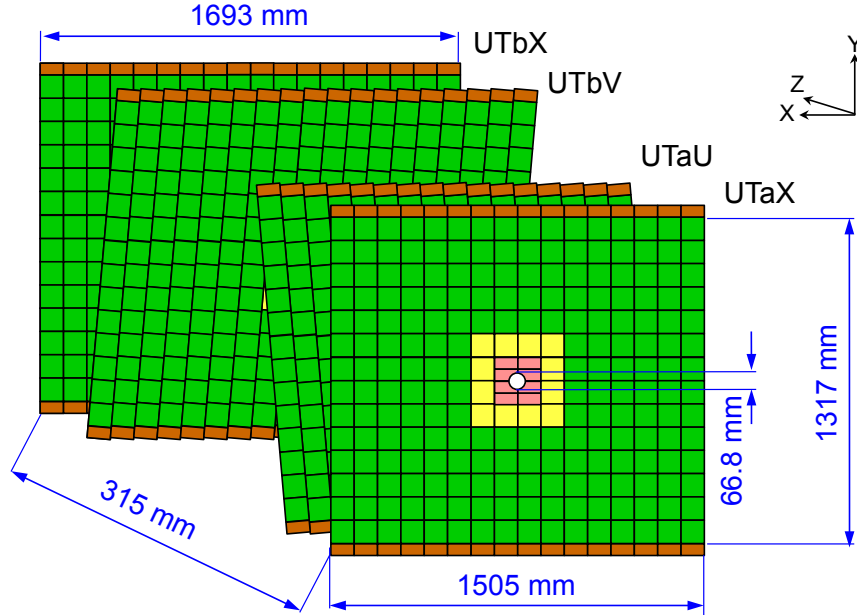


Figure 12.1: Four detection layers of UT, arranged in a $x-u-v-x$ configuration.

Each detection layer is made of a number of staves, a vertical structure of sensors and their readout electronics; the staves are discussed in Section 12.1.1. The sensor readouts are collected by the electronics hosted inside Peripheral Electronics (PEPI) crates at the top and bottom of the detection layer, and are described in Section 12.1.2. Due to space constraints, the low voltage regulators (LVRs), which supply power to the detector electronics, are placed farther away from the beam at the service bay; a brief description of the LVR is given in Section 12.1.3. An overview of overall UT configuration is shown in Fig. 12.3.

¹The rotation axis is the z -axis in Fig. 12.1.

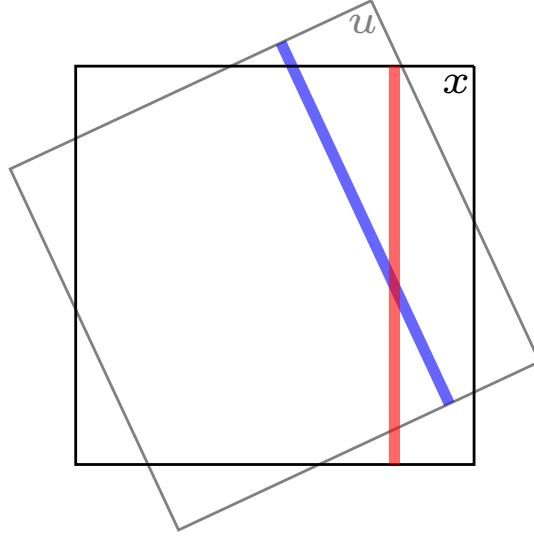


Figure 12.2: Illustration of the measurement of x - y track hit positions with two silicon-strip sensors. The second sensor is rotated by an angle. The intersection between the red and blue strip is the measured position of a track.

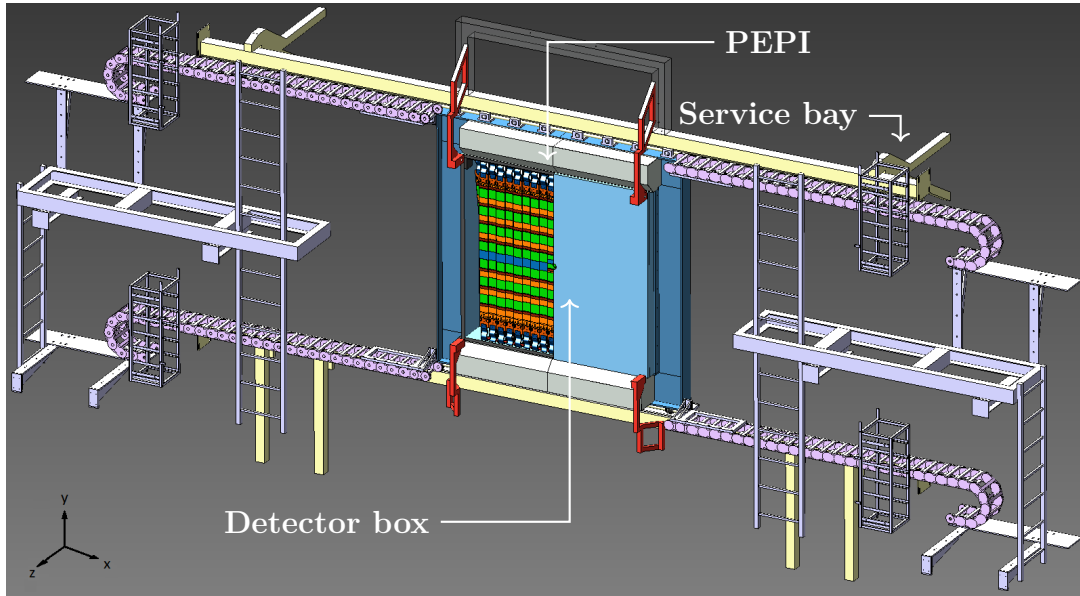
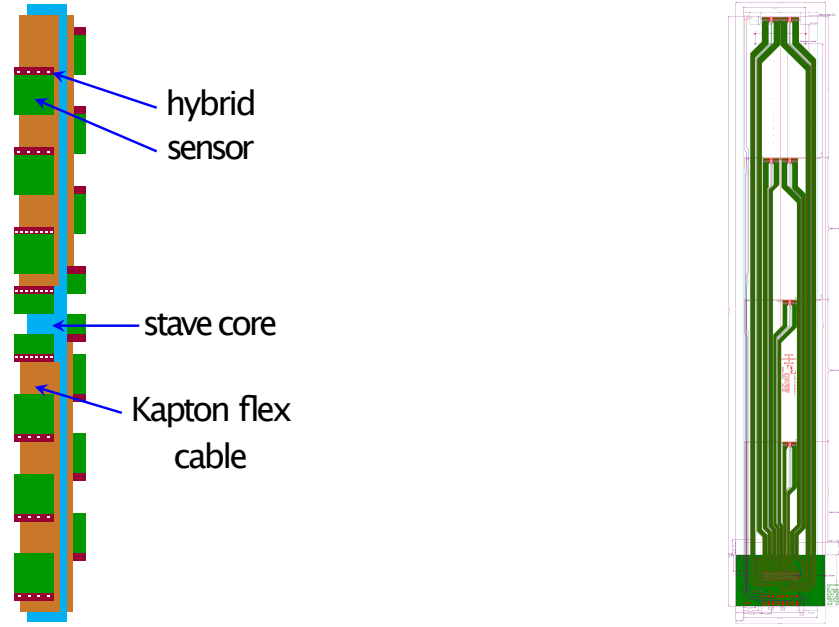


Figure 12.3: Overall UT configuration in the LHCb cavern. Taken from [61].

12.1.1 Stave

Staves consist of four flexible cables (flexes), two on each side, with sensors and hybrid readout boards wire-bonded on each flex. These flexes are laid out on a carbon fiber core with CO_2 cooling pipes. The front and back sensors are mounted in a staggered manner to

ensure coverage. The components mentioned above are illustrated in Fig. 12.4.



(a) Schematic of a UT stave. Each side of a stave contains two Kapton flex cables.

(b) A signal and power layer of the Kapton flex cable. The low voltage coppers are shown in green; the data lines are in red.



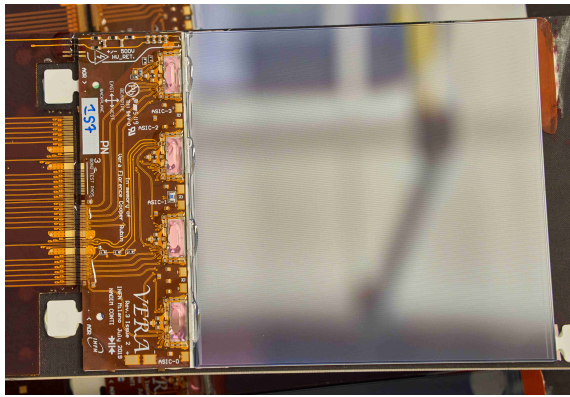
(c) A stave module with a circular cutout to be placed near the beam pipe.

Figure 12.4: UT stave.

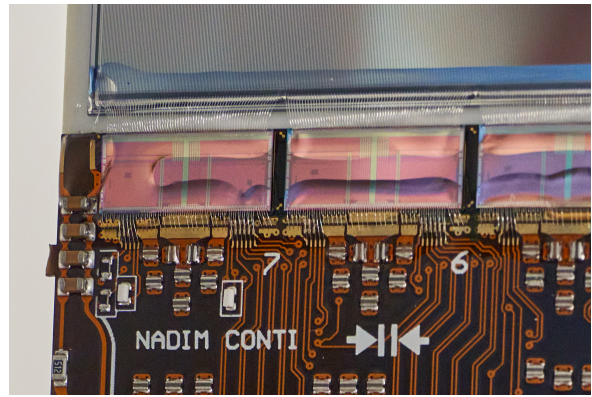
There are four types of sensors with different specifications. The sensors close to the beam pipe (yellow and pink in Fig. 12.1) have a finer silicon-strip pitch of $93.5\ \mu\text{m}$

and are made with n -in- p process for more radiation hardness, mainly due to concerns on occupancy and radiation damage. In addition, the eight half-height sensors (pink) have their strip length halved to further improve resolution near the beam pipe, with the four sensors immediately surrounding the pipe having circular cutouts, instead of rectangular gaps as in TT, to ensure maximum coverage in the very forward direction. The outer sensors (green) have a coarser pitch of $187.5\text{ }\mu\text{m}$ fabricated with p -in- n technology to reduce cost [62].

To minimize the transmission path of the analogue sensor output, readout SALT ASICs² are glued to a ceramic stiffener along with a sensor and electronic hybrid circuit to form a single UT module. The sensor strips are wire-bonded to the input pins of the SALT³ [56]. A sample UT module is shown in Fig. 12.5.



(a) Full view of a sensor module mounted on a flex.



(b) Close-up view on three SALT ASICs glued on a hybrid, with their inputs wire-bonded to a sensor.

Figure 12.5: Production sensor modules and SALT ASICs.

Fabricated with TSMC 130 nm CMOS process, the SALT ASIC extracts, shapes, and digitizes analogue signals from the sensor at 40 MHz from 128 channels simultaneously. The data is sent out at 320 Mb/s, which is achieved by increasing the master 40 MHz clock frequency four times (to 160 MHz) and using double data rate (DDR) transmission. The master clock is provided by the DCB, and the frequency increase by a phase-lock loop (PLL). The SALT chip is capable of generating pseudo-random test patterns on its own without

²SALT is developed by the UT collaboration. “ASIC” stands for Application-Specific Integrated Circuit.

³SALT ASICs are represented by small white rectangles in Fig. 12.1.

any excitation on the sensor, which is employed to test the data transmitting quality. The chip configuration is set by writing registers via the I2C protocol [63].

12.1.2 PEPI

The Peripheral electronics (PEPI) are responsible for dispatching power to the SALT chips, aggregating the sensor data readout and sending them to the counting room, and propagating control commands to the front-end electronics⁴. A diagram for data readout and control signal propagation in the UT is displayed in Fig. 12.6.

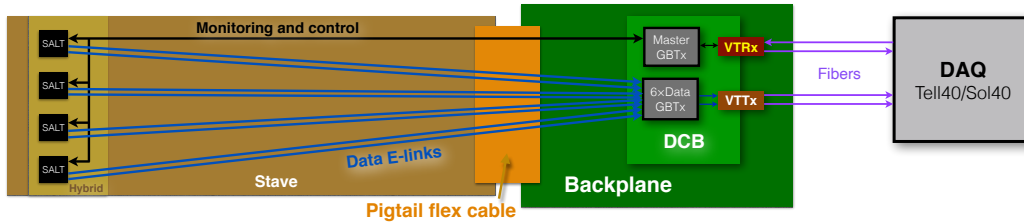


Figure 12.6: UT data and control signal flow diagram.

Up to 12 pigtails are connected to each backplane (BP), which then routes the data to the DCBs with the goal of balancing the load carried by each board (staves closer to the beam pipe have higher occupancy than those farther away). The pigtails are an engineering marvel on its own⁵, dispatching power, data, and control signals between the BP and SALTs on the stave in an *extremely* limited space. A BP can also host up to 12 DCBs. Each DCB serializes its input data and converts the digitized electric sensor readout to an optical one which is then sent to the counting room via optical fibers. The DCB is also responsible for receiving and translating control commands from the counting room to the SALT ASICs. In other words, the whole configurable part of the UT detector is exposed as DCB registers to the counting room.

The PEPI are hosted in PEPI crates, with each containing a single Backplane and up to 12 DCBs, as can be seen in Fig. 12.7. The crate is also responsible for providing a

⁴Front end electronics refer to SALTs and DCBs, which sit close to the beam.

⁵More details regarding the Pigtail can be found in [61]. It is worth noting that during the detector installation, very precise bending of the Pigtails are needed to ensure good connectivity.

common ground reference for all SALTs and DCBs.

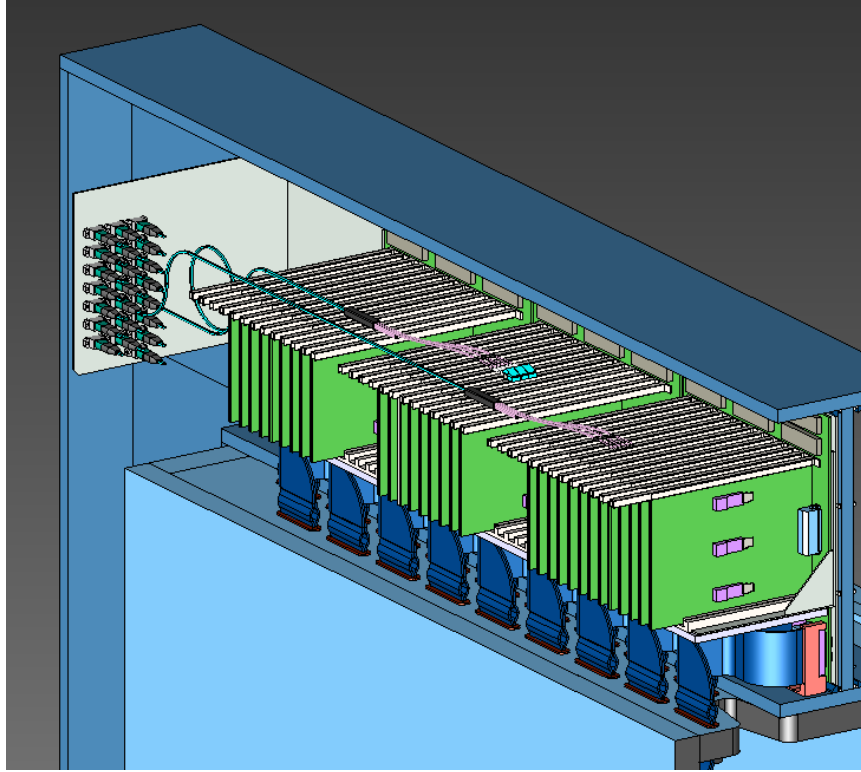
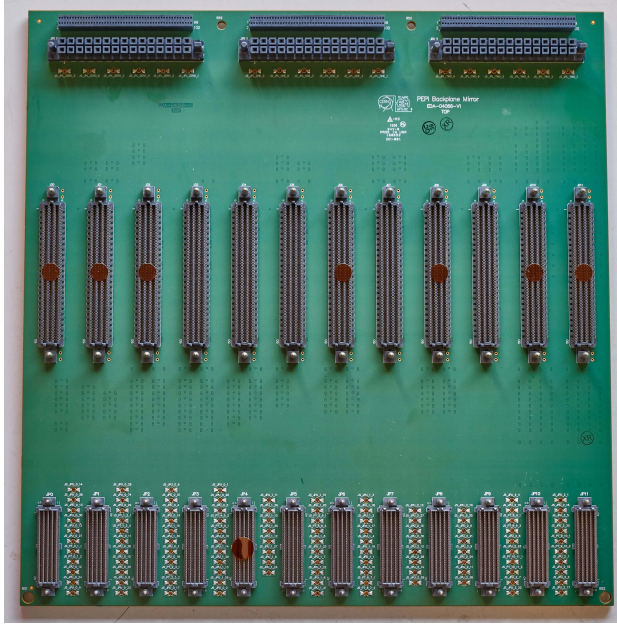


Figure 12.7: A schematic drawing of 3 PEPI crates with DCBs and Pigtails plugged in. Taken from [61].

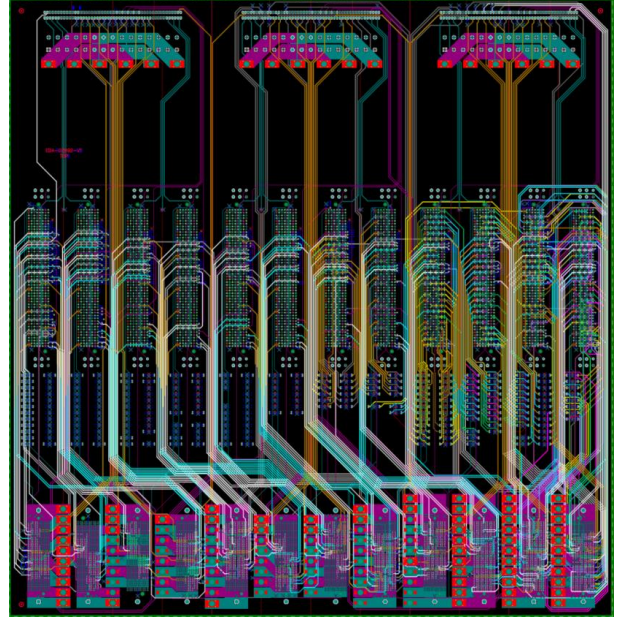
Backplanes are basically a large net of microscopic cables delivering power and facilitating data exchange to specific connector pins, and contain no active components (i.e. integrated circuits). While conceptually simple, Backplanes were very difficult to design because the trace density is extremely high. Fabricated with 28-layer PCBs⁶, BPs are at the limit of manufacturability. To ensure that the traces are routed correctly, the connectivity map (referred as a “netlist”) is exported to a machine-parsable language⁷ and checked against a set of pre-defined rules, for example one such rule ensures that the lines providing power to DCBs and SALTs have the correct voltage levels by a Python script. The quality assurance process of the BP is highly non-trivial, and will be provided on a separate document by another author.

⁶As a reference, a typical motherboard for a desktop PC has 6-layer PCBs.

⁷One such option exports the netlist to a language that is compatible with Common Lisp.



(a) A production Backplane.



(b) Traces of the backplane are of ultra high density.

Figure 12.8: A production Backplane and its copper trace implement.

As its name suggests, the Data Control Board (DCB) is responsible for aggregating SALT readout data and propagating slow control commands. The functionality of the DCB rely on the GigaBit Transceiver (GBT) chipset architecture, a set of ASICs developed by CERN to provide a high-speed optical data transmission solution in a highly radioactive environment such as in the LHC [64]. The chipset includes the following ASICs:

- **GBTx**: it is a radiation tolerant serializing/deserializing ASIC for implementing high-speed uni- or bi-directional optical links in high energy experiments [65].
- **GBLD**: the GigaBit Laser Driver is a vertical-cavity surface-emitting laser driver ⁸. [66].
- **GBT-SCA**: the GigaBit Transceiver-Slow Control Adapter provides an interface between various industrial-standard slow control protocols, namely SPI, I2C, JTAG,

⁸It has many modern applications such as in computer mice and optic fiber communication. The technical details is unknown to the author but is irrelevant for the topic.

and GPIO, to the GBT link protocol. It also has a built-in multi-channel ADC that can be configured and read via the GBT link [67].

Shown in Fig. 12.9, a DCB is a 16-layer PCB that consists of:

- **1 master GBTx:** Operating in transceiver mode, it generates a stable by phase-aligning its on-package quartz oscillator clock with the clock signal provided by the readout system, and distributes the clock signal to 6 data GBTx. The Timing and Fast Control (TFC) commands are distributed to up to 6 SALTs.
- **1 GBT-SCA:** Connected to the master GBTx via an e-port⁹, the GBT-SCA ASIC provides the I2C interfaces of the 6 data GBTx as well as the SALTs which are responsible for relaying slow control commands to these ASICs. Also, it is used to read telemetry data, namely thermistor readout on SALT, DCB, and optical mezzanines, as well as voltages at DCB, via its built-in ADCs.
- **6 data GBTx:** Operating in simplex transmitter mode with an external clock from the master GBTx, the data GBTx serialize and scramble the SALT readouts in a way suitable for optical transmission. The processed signal is sent to a Versatile link module mounted on a optical mezzanine for electro-optical conversion and transmission. One data GBTx can serialize up to twelve 320 Mb/s e-links into a single 4.8 Gb/s data stream. One single data GBTx relays its clock to up to 6 SALTs.
- **1 Versatile link transceiver (VTRx):** The VTRx contains an optical receiver and an optical transmitter with a GigaBit Laser Driver (GBLD) as the electro-optical converter.

It is used to establish a bi-directional connection to the readout system with error-correction functionality that is robust against single event upsets (SEU), suitable for relaying fast and slow control commands, as well as telemetry.

⁹The electric port for high-frequency communication with other GBT-compatible chips, also serves as the building block for e-links.

All Versatile modules are mounted on an optical mezzanine which is plugged on DCB.

- **Up to 3 Versatile link transmitter (VTTx):** It contains two GBLDs to establish two uni-directional connections, suitable for transmitting the serialized SALT readout of two data GBTx to the central LHCb readout system.

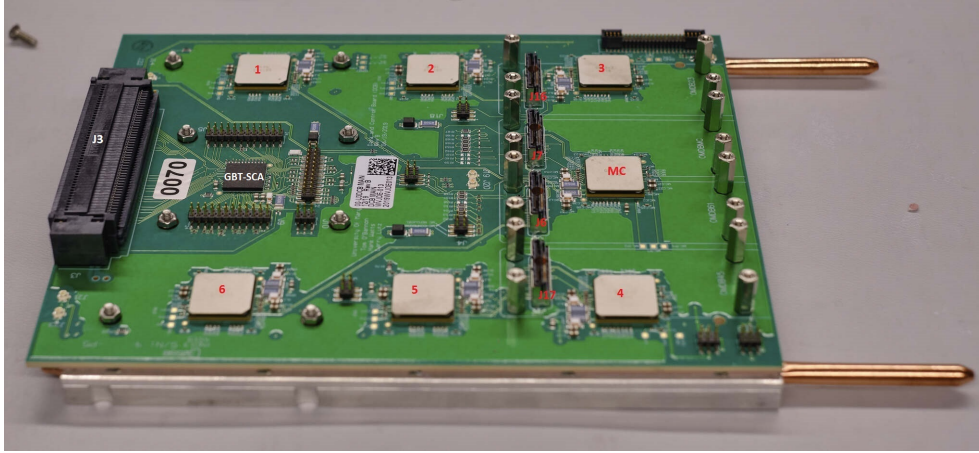


Figure 12.9: A production DCB with cooling backplate assembly. The master GBTx and GBT-SCA are marked. 6 data GBTx are numbered from 1 to 6. The BP connector is marked as J3. 4 optical mezzanine connectors are marked as J6, J7, J16, J17. No optical mezzanine is connected.

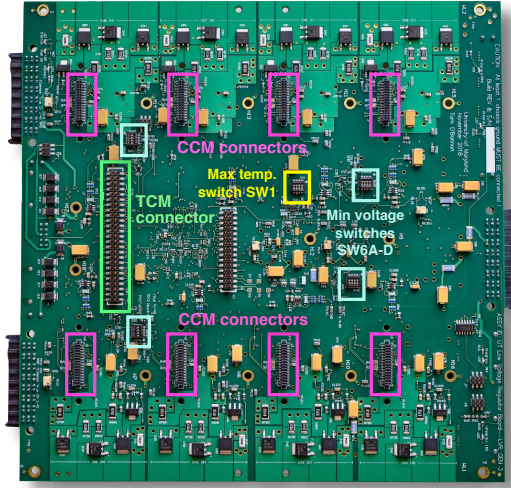
More information regarding DCBs, to which the author made significant contributions, will be provided in Section 12.2.

12.1.3 LVR

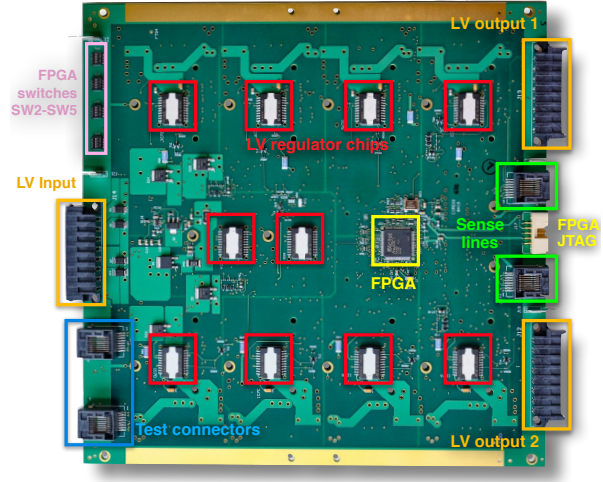
Installed inside service bays, the Low Voltage Regulator (LVR), shown in Fig. 12.10, delivers power to the SALTs and DCBs with a ~ 10 m long cable which makes the resistance of the cable non-negligible and non-trivial to account for.

To ensure correct voltages are delivered at the ASIC side, LVRs are equipped with remote sensing capabilities such that the remote voltage is constantly monitored and kept at the specified voltages¹⁰. This requires one pair of sense lines carrying little current in

¹⁰Consider the specified voltage is at 1.25 V, the LVR output voltage is at about 1.4 V. The V_{sense} vs. V_{LVR} lookup table can be found at [69].



(a) LVR top view (CCM/TCM side).



(b) LVR bottom view (FPGA side).

Figure 12.10: A LVR. Taken from [68].

addition to the pair of current-carrying lines (4 lines for a single output channel). It can be intuited as if the LVR can probe with a multimeter at the power input of the front-end electronics and adjust its output voltages accordingly. When the sense line breaks, the unsensed output voltage at LVR side is internally set at a safe value for all the loads. However, when the sense line is shorted, the LVR ramps up the voltage up to the 5.5 V rail which can damage the DCBs and SALT ASICs. Thus particular care is taken during the cabling phase to ensure no such short circuit happens.

One LVR has up to 8 output channels; each is configured by a Channel Configuration Mezzanine (CCM) with which the specified voltage is fine-tuned by a potentiometer. There are three voltage specifications: the SALT ASICs require 1.25 V; the DCB GBTs require 1.5 V; the Versatile modules require 2.5 V. A single LVR channel can output up to 2.5 A of current. When more current is needed, two channels can be configured into a master-slave pair in which case the sense lines from the master channel dictate the voltage output of the pair.

Remote channel control and monitoring via the SPI protocol is implemented by a Microsemi FPGA¹¹; such functionality can be accessed via the Telemetry Control Mezzanine

¹¹More specifically, a Microsemi ProASIC nano A3PN250-VQG100 FPGA.

(TCM) connector. A separately designed TCM board integrates the LVR into the LHCb control system.

More information regarding LVRs can be found at [68].

12.2 The Data Control Board (DCB)

The DCB is part of an implementation of the reference LHCb front-end electronic module. Shown in Fig. 12.11a, a complete LHCb front-end electronic device includes the digitization of sensor readout, serialization of digitized readout, and distribution of control commands. The DCB, schematically shown in Fig. 12.11b, implements the last two functionalities (for UT the sensor readout is implemented by the SALT ASIC).

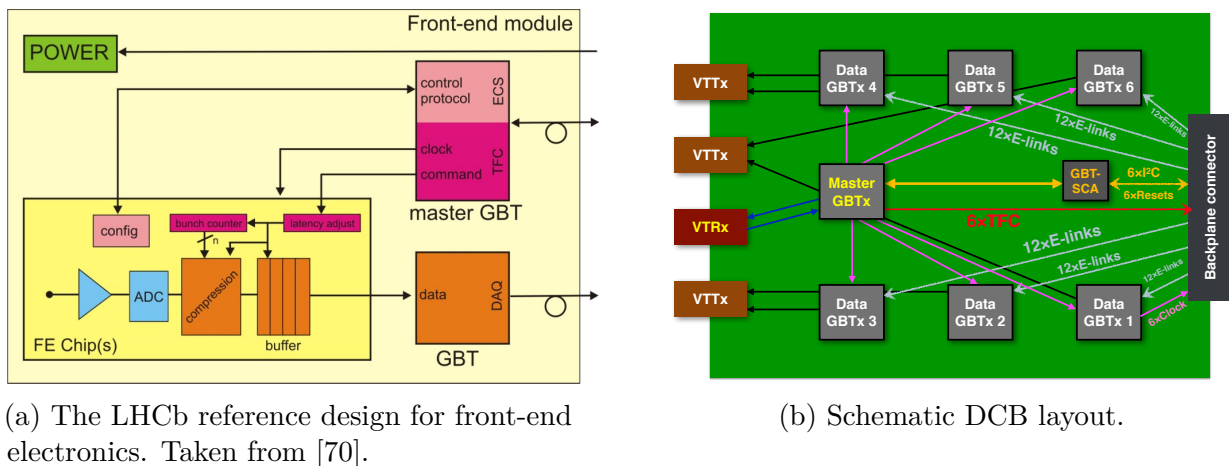


Figure 12.11: Comparison between the reference LHCb front-end electronic and DCB.

The physical layout of the DCB has been summarized in Section 12.1.2. The rest of the section provides more information on the GBT architecture which is the foundation of the DCB (Sections 12.2.1 and 12.2.2), the DCB operation (Section 12.2.3), and the testing and quality assurance of DCB (Sections 12.2.4 and 12.2.5). More details about configuring the DCB can be found in Ref. [71] which is prepared by the Maryland group.

12.2.1 The GBT link

The GBT link has two transmission modes relevant for DCB: FEC and widebus mode. The former provides error correction capability with the trade-off of a reduced data transfer speed at 3.36 Gb/s, making this mode suitable for carrying out critical timing and control data. The latter works at 4.48 Gb/s without error correction at link level, and is used for readout data transmission. In either mode, a single data frame is always 120 bit long.

Transceiver The master GBTx on the DCB is configured to work in the FEC mode for bi-directional communication with the LHCb control system. Each frame in this mode contains the following fields:

- **Header:** 8 bits. For delimitation (identification) of frame boundaries (frame synchronization).
- **Internal Control (IC):** 2 bits. Reserved for remote provisioning (i.e. change selected configuration registers and read telemetry information) of the master GBTx.
- **External Control (EC):** 2 bits. Reserved for transferring external slow control commands, such as I2C commands, to the GBT-SCA. Both IC and EC has a fixed speed of 80 Mb/s.
- **Data:** 80 bits. Contain 5 e-link groups, each 16 bits long (2 Byte). At 320 Mb/s data rate, each e-link group has 2 e-links.
- **Forward Error Correction (FEC):** 32 bits. The error correction algorithm uses a 2-time interleaved Reed-Salomon $RS(15, 11)$ code with 4-bit symbols. In practice the data is robust against up to 16 consecutive bit corruptions.

All fields above are protected by the FEC field, except for the FEC field itself.

Transmitter At the DCB, each data GBTx transmitter operates in the widebus mode. A widebus frame is identical to a FEC frame except that it has no FEC field and extends the data field to 112 bits so that the total number of e-link groups is increased to 7. As a trade-off, there is no error correction capability at link level.

As can be seen from the structure of the GBT frame, there is an overhead due to header and IC/EC fields. This is why the data transmission speed is at 4.48 Gb/s instead of 4.8 Gb/s.

12.2.2 E-links

As described in the previous section, operating in widebus mode, a data GBTx can provide up to $2 \times 7 = 14$ e-links¹², local electric serial links implemented with low-voltage differential pairs transmitting data at 320 Mb/s [65], with the transmission speed chosen by the design of the SALT. In the DCB, however, only up to 12 e-links per data GBTx are accessible to the SALTs.

A SALT chip requires 3–5 e-links [63], so a single data GBTx can aggregate readout data from up to 4 SALTs. The allowed SALT-e-link combinations per data GBTx are listed in Table 12.1.

Number of SALTs	e-links per SALT
4	3
2	3
2	4
2	5

Table 12.1: Possible SALT-e-link combinations per data GBTx link in the DCB.

It is worth pointing out that an e-link is uniquely identified by 3 numbers: the DCB index, the data GBTx index on the DCB, and the e-port number of the GBTx. In other words, e-links are *geographically addressed* without the need of additional location

¹²A GBTx contains many e-ports which are used to facilitate e-links.

information in the GBT frame: A particular range of bits from GBT frames produced by a known DCB is sufficient for identifying the geographical location of the SALT it connects to.

During data transmission via an e-link, the copper trace length between the DCB and the SALT introduces a constant phase difference between the two ASICs. Such a phase may render the data sampling unreliable as the GBTx may sample on a transitional state of the data signal. To solve this problem, each e-link on the GBTx is equipped with a phase-adjusting capability so that the sampling happens with a stable data signal in between clock transitions.

12.2.3 Initialization sequence after a power cycle

A GBTx ASIC starts up *unconfigured*: It will not communicate to the readout and control system without a configuration and the optical interface will not be functional. It is also not possible to bring the I2C interfaces via copper lines for all DCB master GBTxs to the counting room because the distance is too long for I2C to work. Therefore, the master GBTxes have their configuration fused during the QA process.

With a fused master GBTx, the front-end electronics (in this case consider a single DCB and the SALTs it connects to) have the following initialization sequence:

1. Power supplied to the DCB and SALTs.
2. The master GBTx (for short MC) pauses for 300 μs before sampling its fuses. This pause is to wait for the power to stabilize.

It is observed that with a slower power supply, such as the Rigol power supply in our laboratory, the DCB turn-on can become unstable. In such case it is advised to manually hold the master GBTx in reset until the power is stable.

3. MC samples its fused registers to configure itself.

4. MC generates a reference clock at the LHC frequency (40 MHz) with its internal Voltage-Controlled Quartz Crystal Oscillator (VCXO) by its internal phase-lock loop.
5. MC starts receiving GBT data frames with its deserializer and a Clock and Data Recovery circuit (CDR) tracks the headers of the incoming frames and phase-aligns the MC reference clock with the frame-sending frequency.

After the deserializer declares a lock, the MC's internal clock and the clock in the counting room are considered to be synchronized (they differ by a constant phase). The CDR also provides a high-frequency (320 MHz) clock to the serializer.

6. After 800 consecutive clock cycles of constant phases between its phase-lock loop and the clock provided by the CDR, the serializer considers its clock stable and start finalizing the rest of the transmitter block.

MC initialization is done and a bi-directional link in FEC mode to the counting room is established.

7. A reset signal is issued from the control system to all data GBTxes and SALTs via the GPIO interface from the GBT-SCA.

At this point, the data GBTxes have a stable clock directly from the MC.

8. The data GBTxes are programmed via the I2C interface of the GBT-SCA in a staged manner to avoid large jumps in current.

At this point SALTs also have a stable clock from a data GBTx which relays its clock.

9. The SALT chips are programmed via I2C of GBT-SCA.

Initialization is complete.

12.2.4 Test of data transmission quality

Two types of checks were performed to validate the robustness of the DCB design: an eye diagram and long-term bit error test. An eye diagram is produced from configuring the DCB to output a pseudo-random bit sequence¹³ (PRBS) by triggering on the rising and falling edges of the waveform (unit time interval) overnight and overlaying the triggered pieces together. It is a visualization of the p.d.f. of the signal per cycle, with an open “eye” indicating that the transmission is stable with low jitter. Note that the Keysight oscilloscope has to be configured to calculate the data rate *semi-automatically* instead of *automatically*. The eye diagram is displayed in Fig. 12.12.



Figure 12.12: An eye diagram for DCB, produced overnight. Wide-open and has no noise in the eye, an indication of good signal quality.

Due to the nature of PRBS, once a single bit is received from the sequence, the next

¹³A PRBS is a periodic sequence such that if current value of a is provided, the immediate next value can be computed with $f(a)$ for some known function f .

bits can be predicted and checked against indefinitely. A high-statistic check of PRBS was performed for 5 consecutive days without a single bit error, corresponding to an error-free data transmission of $\mathcal{O}(10^{15})$ bits in a low radiation environment¹⁴.

12.2.5 Quality assurance of DCB at CERN

A total of 270 DCBs were produced and tested at Maryland. Of those, 260 were shipped to CERN and 248 were installed in the UT. After the shipment, the author performed a quality assurance procedure again on all of them to ensure there was no damage during transportation. A testing panel, displayed in Fig. 12.13, was developed to perform the following tests, similar to tests performed at Maryland, with a single click¹⁵ on 2 DCBs at a time:

1. Read the status of master GBTx to ensure the bi-directional connections are established.
2. Perform a PRBS test with a DCB-generated PRBS for 2 minutes.
3. It was observed at Maryland that some of the DCBs produced random bit errors with the default optical transmission current of 5 mA. Increasing that to 6 mA solved the problem empirically.

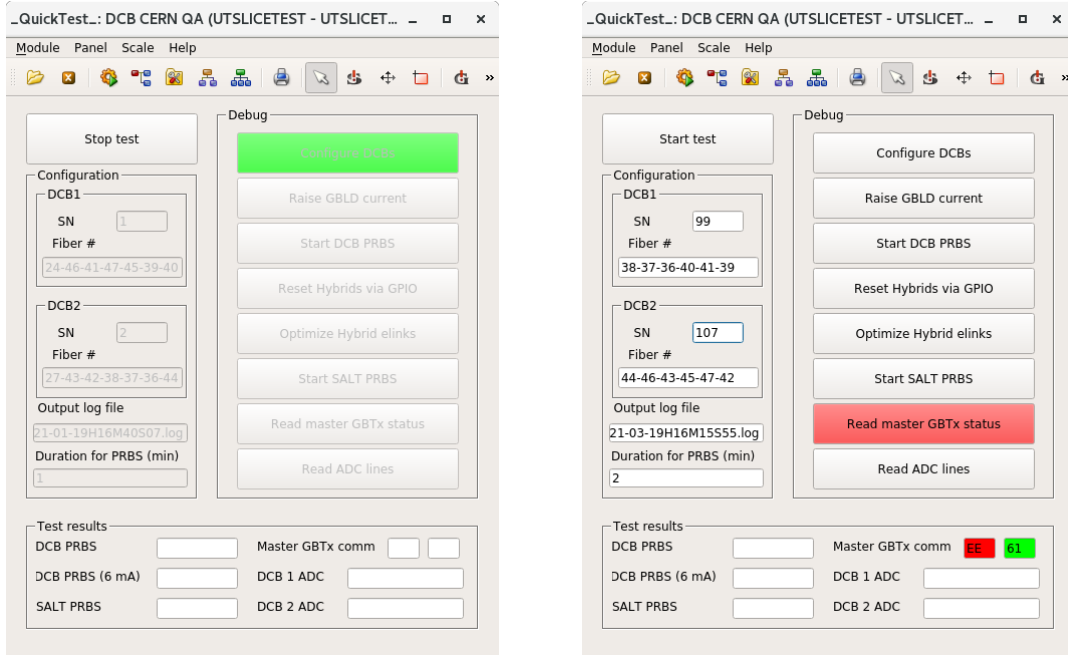
During a PRBS test, if there is any error, the VTTx transmission bias current is increased from 5 mA to 6 mA, and the PRBS test is repeated.

4. The thermistors connected to the GBT-SCA ADC are readout and checked to make sure they fall within a reasonable range.

The panel is written in a pseudo C# language used by the WinCC OA SCADA (supervisory control and data acquisition) system, the control system chosen by CERN. The

¹⁴It was performed in our laboratory on the 3rd floor of the Physical Sciences Complex at the University of Maryland.

¹⁵This made the author's life slightly easier. It also allowed the author to partially offload the task to volunteers. Thank you Roberto Piandani and Marian Stahl!



(a) A running panel.

(b) ...sometimes the master GBTx refuses to talk.

Figure 12.13: DCB CERN testing panel.

underlying functionalities are borrowed from code written by Mark Tobin and other CERN experts.

12.3 The LHCb online system

The LHCb online system is responsible for ensuring a consistent clock distribution across the whole LHCb detector, collecting readout data from front-end electronics (FE), filtering events with software-based HLT algorithms, and provisioning initialization, control, and monitoring of the detector. As shown in Fig. 12.14, aside from execution of HLT algorithms, the LHCb online system consists of three important components: [72]:

- **SODIN:** A central readout manager that receives clock from the LHC and distributes

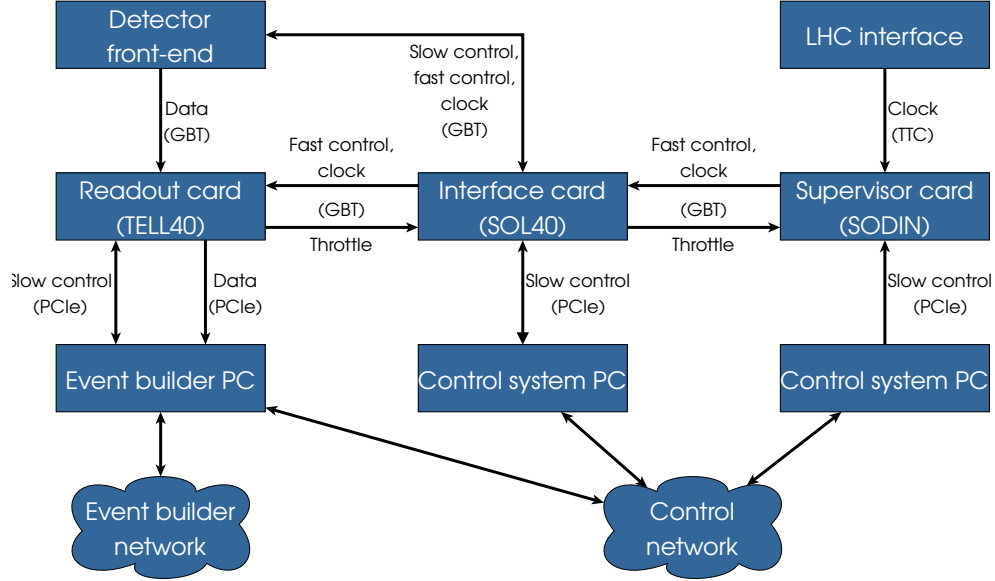


Figure 12.14: An overview of the LHCb online system.

it to SOL40. It also issues Timing and Fast Control (TFC) commands¹⁶ to the front-end to regulate data flow.

- **SOL40:** Fan out TFC commands to TELL40 and FE, and fan in throttle from TELL40 to SODIN. It also provides an interface between the FE and the control system of the experiment.
- **TELL40:** Readout from FE. Hosted inside event builders. Accept TFC commands for synchronization.

The functionalities required by the three components are implemented with a single type of hardware: the PCIe40 board, which is a PCI Express Gen 3.0 x16 add-in card based on Intel Arria 10 FPGA and high-density optical I/O with up to 48 duplex GBT link ports. By flashing a different firmware, a PCIe40 can satisfy the requirements by each component. In its TELL40 flavor, PCIe40 is capable of copying the received readout data into the main

¹⁶In general, TFC commands have the following capability: Provide a bunch-crossing ID at 40 MHz to FE which is used to ensure synchronicity across FE in the detector; allow FE to operate in special modes such as producing a calibration pulse or transmitting data non-zero-suppressed; instruct FE to reset their operation logic while retaining their configuration; among others. More information regarding TFC can be found at [73] but is not the focus of this text.

system memory at over 100 Gb/s. It is also possible to load a firmware combining all three components with reduced read bandwidth, in which case the PCIe40 forms the foundation of the MiniDAQ system used by the collaborating institutions to develop FE [74].

The remaining of this section first provides an overview of the slow control protocol (Section 12.3.1), followed by an introduction on the UT online system (Section 12.3.2). Afterwards, a brief overview of the MiniDAQ system is provided in Section 12.3.3. Finally, a program developed in-house is discussed in Section 12.3.4, showing its utilities in testing FE at Maryland.

12.3.1 The DIM control protocol

The communication between all components, for example controlling a particular SOL40 responsible for several DCBs via the control system in the counting room, of the slow control system relies on the Distributed Information Management (DIM) system, which provides a network¹⁷ based inter-process communication protocol [75], implemented in C to ensure portability across operating systems. To implement a centralized control system for the LHCb experiment hardware connected to many different controllers (servers), the DIM system requires a central name server (like a phone directory) which all DIM servers (like people offering certain services) providing device-control functionalities will register on. When the controller, a DIM *client*, needs to configure a device with a certain name, it first inquires the name server (look up a name in the phone directory) before connecting to the specific DIM server and effecting changes (call the people listed in the directory and conduct business).

To be more specific, first consider a single SOL40 connected to some FE: A part of the FPGA on the PCIe40 card implements a GBT transceiver. Further, the ability to write to and read from that transceiver is exposed to the operating system¹⁸ on SOL40. A DIM server is also running on the SOL40 to translate different types of GBT write/read commands,

¹⁷With TCP/IP protocol.

¹⁸Typically linux, more specifically CentOS 7.

such as GBT-I2C and GBT-GPIO, to DIM *services*.

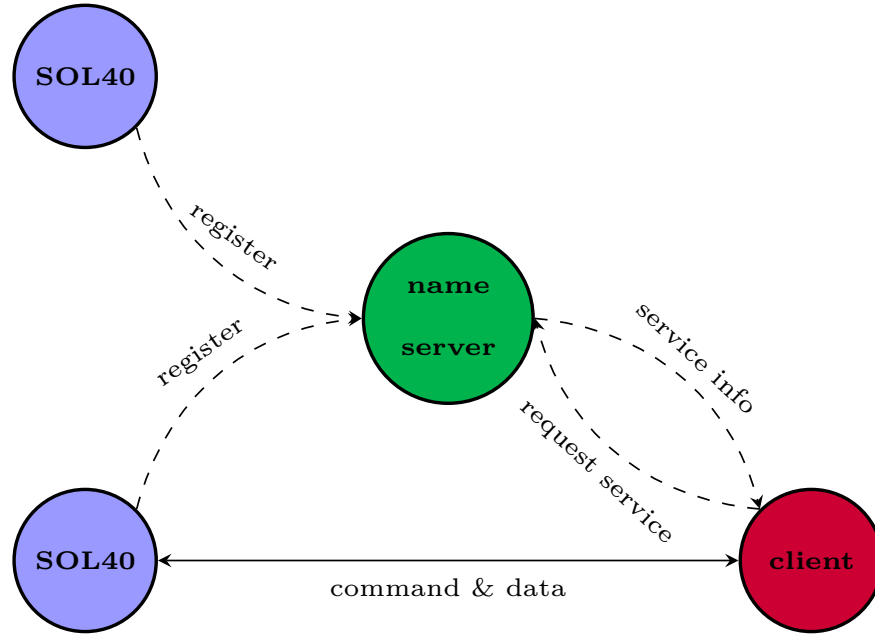


Figure 12.15: A diagram illustrating DIM control flow.

Each SOL40 then registers its DIM services to a central name server. A human controller sitting in the control room at point 8 will have a computer running a DIM client with many clickable buttons corresponding to GBT write commands to different FE. The controller can click a button in which case the DIM client will perform a lookup on the central name server and make changes to the FE via the DIM server running in the looked-up SOL40. The procedure is illustrated in Fig. 12.15. Each DIM service either accepts inputs or provides outputs, but *not both*. The format of the input/output is described by a single specification:

```
<type of 1st arg>:<num of elem in arg>;<type of 2nd arg>:<num of elem in arg>; ...
```

where:

- The following types are allowed: Char (c): 1 byte; Int (I): 4 bytes; Long (L): 4 bytes; Short (s): 2 bytes; Double (D): 8 bytes; Float (F): 4 bytes; Long Long (x): 8 bytes.
- Services with only a single argument is allowed: I:2 is legal which stands for an argument of two 4-byte integers.

- The last argument may have a variable length in which case the `num of elem in arg` is omitted. For example, `I:1;C` allows a variable number of 1-byte chars as its second argument.

As an example, consider the interface for I2C operations (`CmndI2COperation`):

```
name = CmndI2COperation, format = C:128;C
```

with the specification above, it is understood that the I2C operation takes two arguments: a 128-byte chars defining the device and register addresses, and a variable number of chars specifying the actual content to write via I2C.

12.3.2 The UT online system

The UT online system provides functionalities to control, supervise, and collect data readout from the UT detector. It consists of 3 subsystems: DAQ, DCS, and HV, as shown in Fig. 12.16.

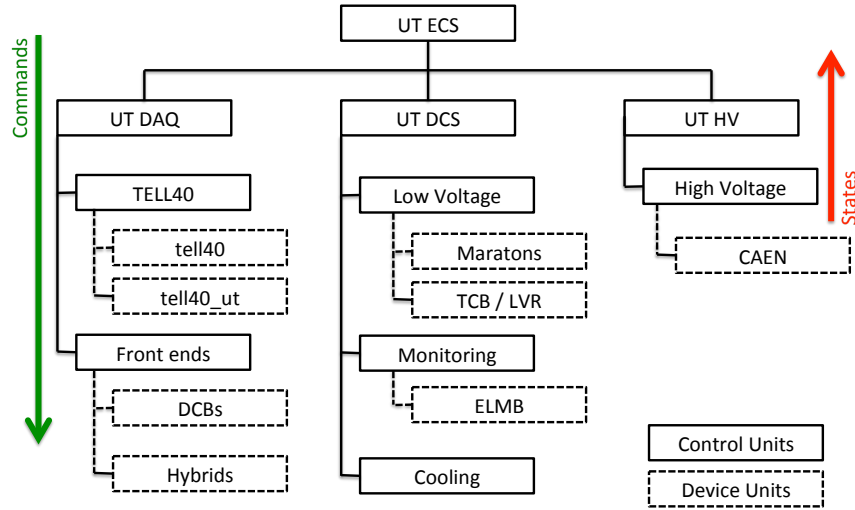


Figure 12.16: An overview of the UT online system.

The DAQ subsystem is responsible for the control and monitoring of the UT front-end electronics (FEs), namely the DCBs and SALT ASICs; it is also responsible for reading out data from the FEs. The DCS controls the LVRs and Maraton (low-voltage power supply)

that power the front-end electronics, provides an always-on monitoring of the temperature and humidity of the UT via the ELMB, and operates the UT CO₂ cooling system. The HV controls the CAEN which provide reference high-voltage lines to the silicon-strip sensors.

All three subsystems are integrated into the WinCC OA SCADA (Supervisory Control And Data Acquisition) system, the software framework for the LHCb online system that monitors the status¹⁹ of the detector continuously and provides a graphical user interface for the detector control. The FEs and LVRs are controlled by the SOL40s, which use the DIM protocol for bi-directional communication to the SCADA system. All other components, namely Maraton, ELMB, and CAEN, use OPC (OLE process control), an industrial-standard, as the communication protocol to the SCADA.

As discussed in Section 12.3, The SOL40s provide the control of FEs, whereas the TELL40s provide data readout functionality. Both are implemented by the same underlying hardware, the PCIe40 expansion card as shown in Fig. 12.17, but with different firmware. There are 7 SOL40s in total for UT, with 3 for controlling the DCBs on each side of the detector for a total of 6 SOL40s, and 1 for LVRs (via the TCBs). All 7 SOL40s (which are PCIe40 expansion cards) are installed in a single server. There are a total of 108 TELL40s with 54 for each detector side. Each data readout server is equipped with 3 TELL40s, with 2 GPUs for the HLT software triggers as discussed in Section 11.1.

12.3.3 MiniDAQ

The MiniDAQ system, a miniature LHCb online system, is provided by the LHCb collaboration as a development platform for all front-end electronics [74]. The MiniDAQ hardware is an off-the-shelf server equipped with a PCIe40 card. The PCIe40 card is loaded with a firmware capable of performing triple duty of SODIN, SOL40, and TELL40. To form a complete package, MiniDAQ provides the following software:

- A linux distribution, typically CentOS 7.

¹⁹For example, the thermistor readout from the FE and the DCB operation status.

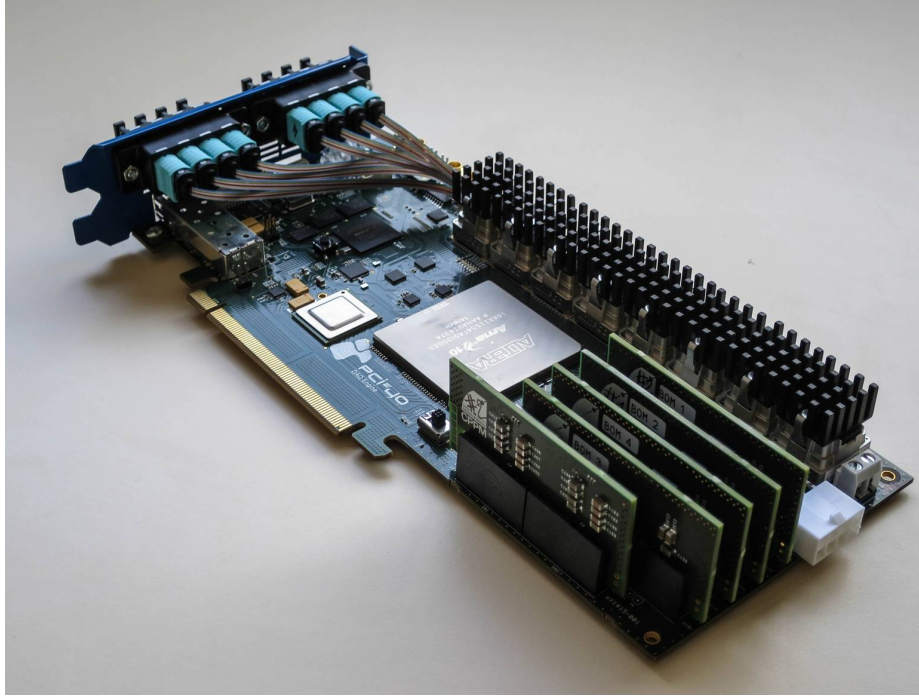


Figure 12.17: A prototype PCIe40 card.

- A linux kernel module to interface the PCIe40 FPGA to the operating system.
- A DIM server named `gbtserv`. A DIM name server is also provided.
- A WinCC OA project acting as a control system more geared toward development. The DIM client is incorporated into WinCC OA by the JCOP framework, developed in-house at CERN.

The WinCC OA project on the MiniDAQ provides functionalities such as programming the data GBTxes with GBT-I2C protocol, shown in Fig. 12.18, and monitoring the readout from them, shown in Fig. 12.19. These functionalities are crucial in the early states of the DCB development and validation, but become increasingly cumbersome as the focus shifts on quality assurance of a large number of production DCBs.

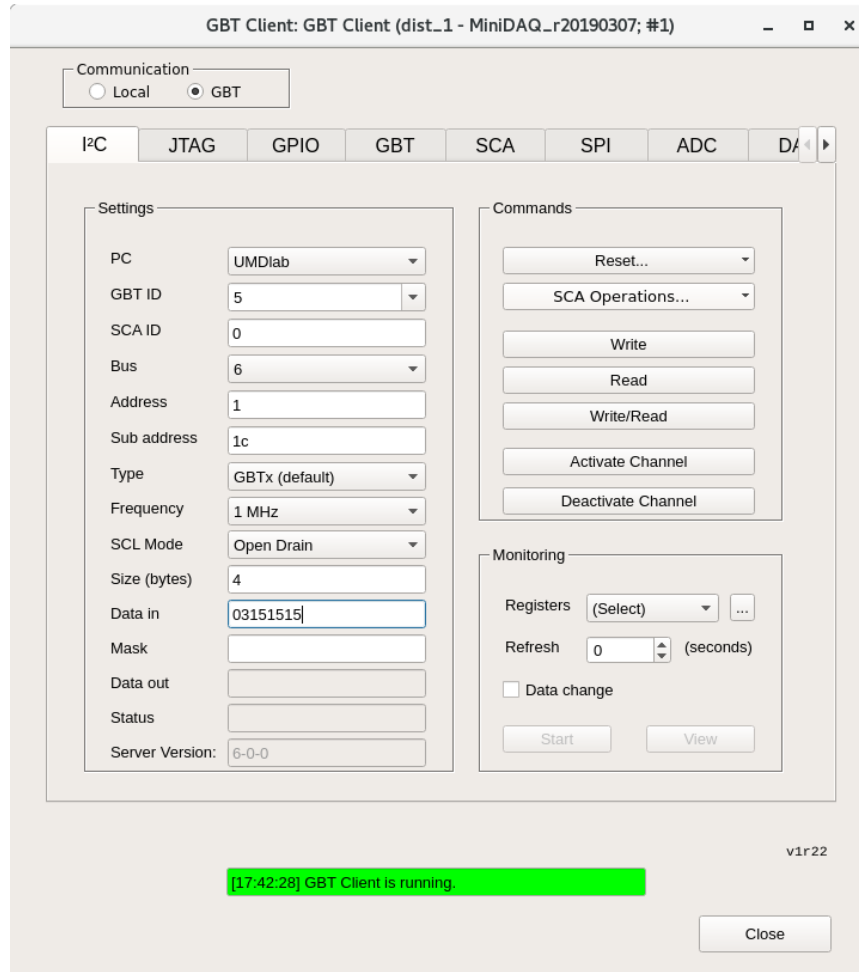


Figure 12.18: The WinCC OA MiniDAQ projection interface to program a single data GBTx. It supports both programming individual registers (as shown in this figure) or programming all registers with a text file.

12.3.4 nanoDAQ: A reverse-engineered DIM client for testing UT electronics

As stated in the previous section, repeated testing with the official MiniDAQ project is hard. Here the author supplies two concrete examples:

- A DCB contains 6 data GBTxes. With the official control panel, shown in Fig. 12.18, programming 6 data GBTxes to do PRBS tests requires the change of the “Address” field from 1 to 6 manually 6 times.
- As a stave with SALTs arrived at Maryland, a full readout test for every DCB becomes possible. As discussed in Section 12.2.1, each e-link phase needs to be adjusted to

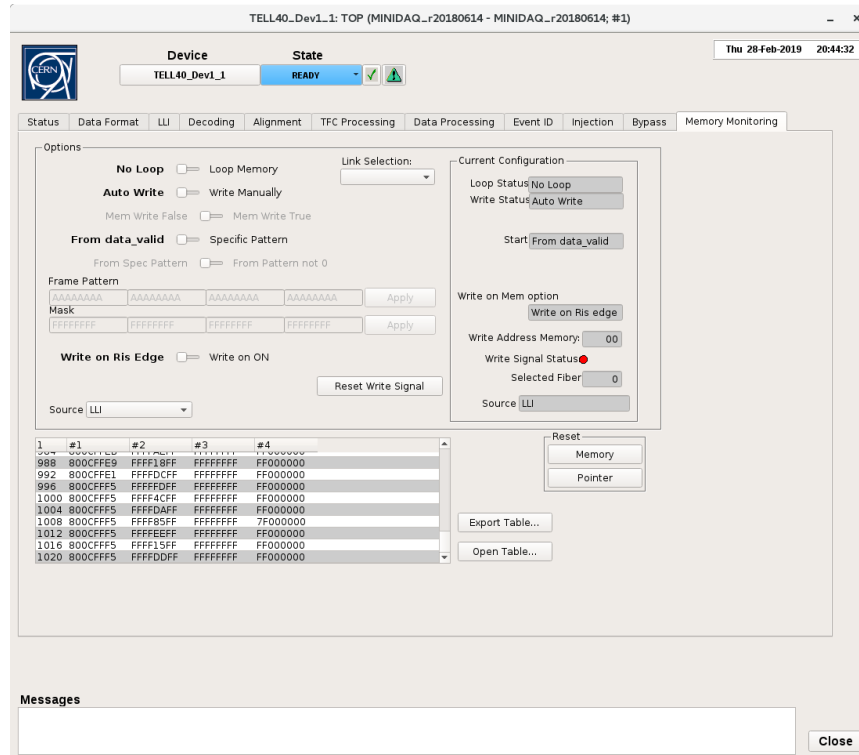


Figure 12.19: A WinCC memory monitoring panel.

ensure stable readout, which is performed on every turn on.

The procedure is the following: First configure the SALT to output a fixed pattern²⁰, then stare at the memory monitoring panel, displayed in Fig. 12.19, to spot the unstable e-links, and adjust phases for each e-link until they are stable.

The procedure was further complicated by a bug in the panel shown in Fig. 12.19: To start reading the memory which are filled with GBT frames from the GBTxes, the option “Write on Ris Edge” needs to be toggled to on. To inspect these readouts, it needs to be toggled off. There is a chance that the switch auto-toggles in which case the only known workaround to the author is clicking on the switch repeatedly until it stops.

It is evident that a batch-processing program to perform these procedures with few clicks or a single command is highly desirable. Unfortunately programming with WinCC

²⁰For some reason, the pattern C4 was chosen.

OA is hard: the development environment, called `gedi`, can auto-complete some of the keywords but cannot check basic errors such as an undefined variable. This resulted in frequent typos in the source code that are hard to spot and debug.

In the end, noting that the MiniDAQ comes with a DIM server that is responsible for performing all GBT-related tasks, a command line DIM client was developed with the Python programming language. First, all registered DIM services are dumped with `pydim` [76]. Then, the required functionalities such as I2C, GPIO, and memory monitoring are searched for based on names. For example, the I2C services have the following signatures:

```
Service 22436 : Type of service = 1 name = Gbt/UMDlab/SrvI2CWrite, format = I:1
Service 26017 : Type of service = 1 name = Gbt/UMDlab/SrvI2CRead, format = I:1;C
Service 24492 : Type of service = 2 name = Gbt/UMDlab/CmndI2COperation, format = C
               :128;C
```

where the `SrvI2CWrite` returns the write status (`I:1`) of a *previous* I2C write operation, the `SrvI2CRead` returns the I2C read status (`I:1`) and values (`C`) of a *previous* I2C read operation, and the `CmndI2COperation` accepts a I2C operation (e.g. read or write) *without returning anything*.

It became possible to identify these services in the JCOP and `gbtserv` source code, and understood the details of the required arguments. Hence the `nanoDAQ` project [77] was born. The name suggests that it only implements a small fraction of the capabilities provided by MiniDAQ, namely I2C, GPIO, and memory monitoring readout.

With `nanoDAQ`, multiple DCBs (in the example below 3) can be configured *programmatically* to generate PRBS patterns, for example by running the following `bash` script:

```
#!/usr/bin/env bash
GBTLINKS=( 3 4 5 )

for i in ${GBTLINKS[@]}; do
    echo "Reset all data GBTTxs first..."
    ./dcbutil.py gpio --reset $i

    echo "Initialize DCB on GBT link $i..."
    # here a configuration file is loaded
    ./dcbutil.py init ./gbtx_config/slave-Tx-wrong_termination.txt -g $i
```

```
echo "Configure each data GBTx to do PRBS..."
./dcbutil.py write 1c 3 -g $i -s 1 2 3 4 5 6
# to flip back to normal operation:
#./dcbutil.py write 1c 1 -g $i -s 1 2 3 4 5 6
done
```

`nanoDAQ` also comes with an e-link phase adjusting script, which performs the following tasks in sequence, with a demo screenshot provided in Fig. 12.20:

1. Configure all SALTs connected.
2. Perform an initial scan of the e-link readout, and ask the user to input the e-link indices to adjust.
3. Loop over all possible DCB e-link phases, and save the readout for each of them. The stability of the readout for a given phase is determined by checking if all readout Bytes can be shift to the fixed pattern `C4` with a common phase²¹. If not, but if the most common pattern in the readout data is a phase shift, the phase is marked in **yellow**; otherwise it is marked with a **red "X"**.
4. A stable phase, marked in **green**, is determined for each e-link by picking the central phase of consecutive stable phases.
5. Adjust both the DCB e-link phases and the SALT output phase to ensure the output pattern is `C4`.

The use of `nanoDAQ` greatly increased the efficiency of the DCB QA procedure at Maryland²².

²¹Consider the pattern `0x13`, which can be written in binary form as `0b00010011`. Compare this to the binary form of `0xC4` which is `0b11000100`, noting that the input stream is a series of Bytes of the form `0b000100[11,000100]11...`, it can be concluded that `0x13` can reproduce `0xC4` by shifting the pattern start 6 Bytes to the right.

²²Looking back, the author gained most of the online system knowledge due to the pain inflicted by DCB and SALT testing.

```

admin@UMDlab: dev+ > ./phaseadj.py -g 0 -s 3 -b 2 -a 3 -c 23
Current readings of MiniDAQ channel 23:
  tx_datavalid  header  13-12  11-8  7-4  3-0
-----
      80      0f  ff-00  13-13-13-13  26-26-26-26  26-89-89-89
      80      0f  ff-00  13-13-13-13  26-26-26-26  26-89-89-89
      80      0f  ff-00  13-13-13-13  26-26-26-26  26-89-89-89
      80      0f  ff-00  13-13-13-13  26-26-26-26  26-89-89-89
      80      0f  ff-00  13-13-13-13  26-26-26-26  26-89-89-89
      80      0f  ff-00  13-13-13-13  26-26-26-26  26-89-89-89
      80      0f  ff-00  13-13-13-13  26-26-26-26  26-89-89-89
      80      0f  ff-00  13-13-13-13  26-26-26-26  26-89-89-89
      80      0f  ff-00  13-13-13-13  26-26-26-26  26-89-89-89
      80      0f  ff-00  13-13-13-13  26-26-26-26  26-89-89-89
Input elinks to be aligned, separated by space: 11 10 9
Continue to Elink phase adjustment (y/n)? y
Generating phase-scanning table, this may take awhile...
phase  11  10  9
-----
0      26  26  26
1      26  26  26
2      26  26  26
3      26  26  26
4      26  13  26
5      13  13  X
6      13  13  13
7      13  13  13
8      13  13  13
9      13  13  13
a      13  13  13
b      13  13  13
c      13  89  13
d      89  89  89
e      89  89  89

Current fixed pattern is 13, adjusting DCB and SALT phase...
Shift SALT elink phase to 6
  tx_datavalid  header  13-12  11-8  7-4  3-0
-----
      80      0f  ff-00  c4-c4-c4-13  26-26-26-26  26-89-89-89
      80      0f  ff-00  c4-c4-c4-13  26-26-26-26  26-89-89-89
      80      0f  ff-00  c4-c4-c4-13  26-26-26-26  26-89-89-89
      80      0f  ff-00  c4-c4-c4-13  26-26-26-26  26-89-89-89
      80      0f  ff-00  c4-c4-c4-13  26-26-26-26  26-89-89-89
      80      0f  ff-00  c4-c4-c4-13  26-26-26-26  26-89-89-89
      80      0f  ff-00  c4-c4-c4-13  26-26-26-26  26-89-89-89
      80      0f  ff-00  c4-c4-c4-13  26-26-26-26  26-89-89-89
      80      0f  ff-00  c4-c4-c4-13  26-26-26-26  26-89-89-89
      80      0f  ff-00  c4-c4-c4-13  26-26-26-26  26-89-89-89
Continue to TFC phase adjustment (y/n)? n

```

Figure 12.20: E-link phase adjust with nanoDAQ.

Appendices

Appendix A: Supplemental plots

A.1 Change of true q^2 due to ISGW2 $\rightarrow$ BLR with verbatim parameters

The change of true q^2 due to form factor reweighting from ISGW2 to BLR with verbatim parameters taken from Table V of [18] are plotted in Figs. A.1 and A.2.

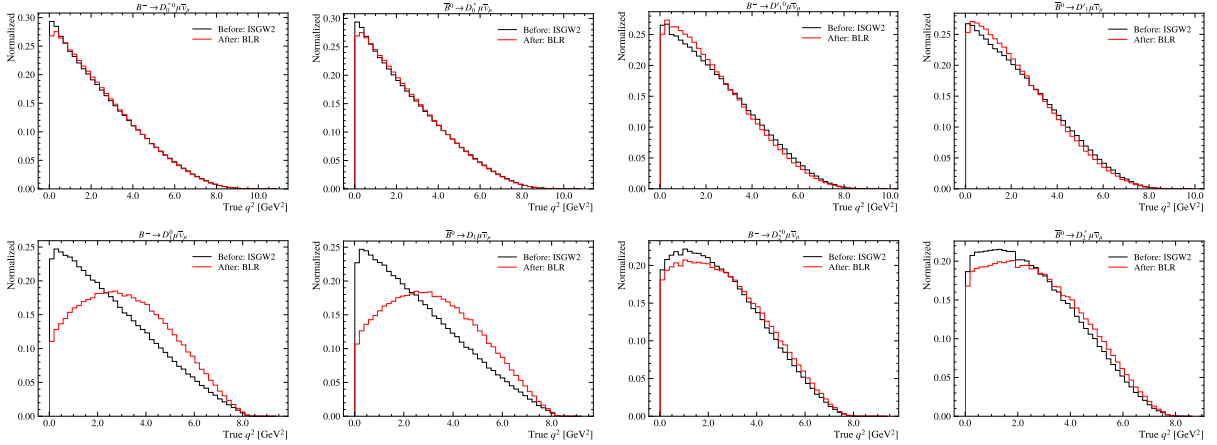


Figure A.1: Form factor reweight effect on $D^{**}\mu$ MC templates. Parameters are taken verbatim from [18].

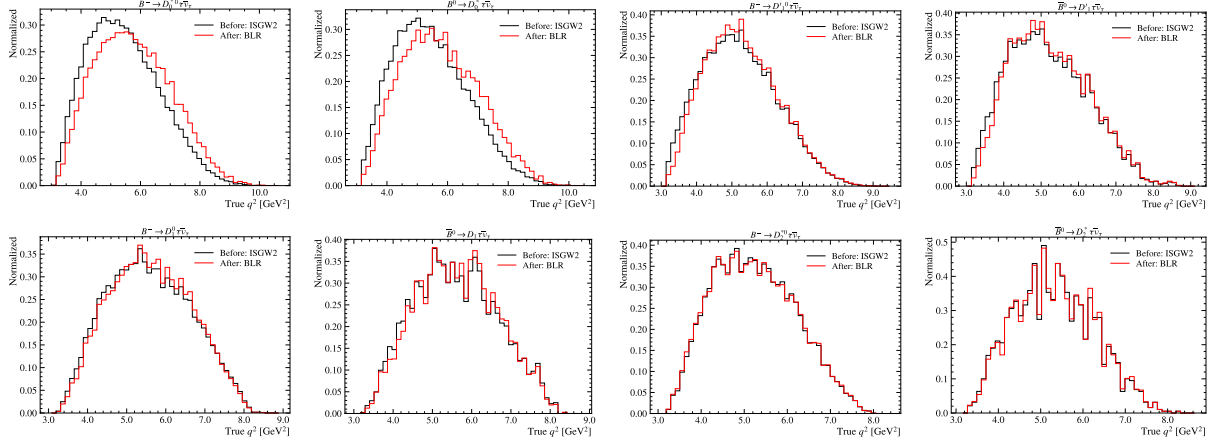


Figure A.2: Form factor reweight effect on $D^{**}\tau$ MC templates. Parameters are taken verbatim from [18].

A.2 D^* combinatorial fit to 1OS, 2OS, and DD skims

The D^* combinatorial fits to 1OS, 2OS, and DD skims are shown in Figs. A.3 to A.5.

The effects of scaling are displayed in Fig. A.6.

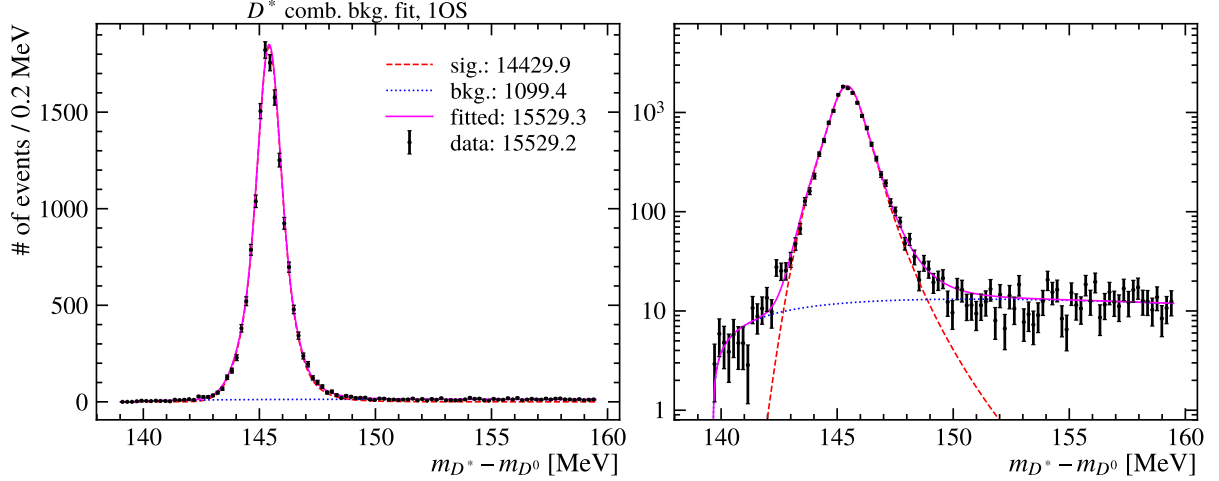


Figure A.3: D^* combinatorial auxiliary fit to the 1OS skim.

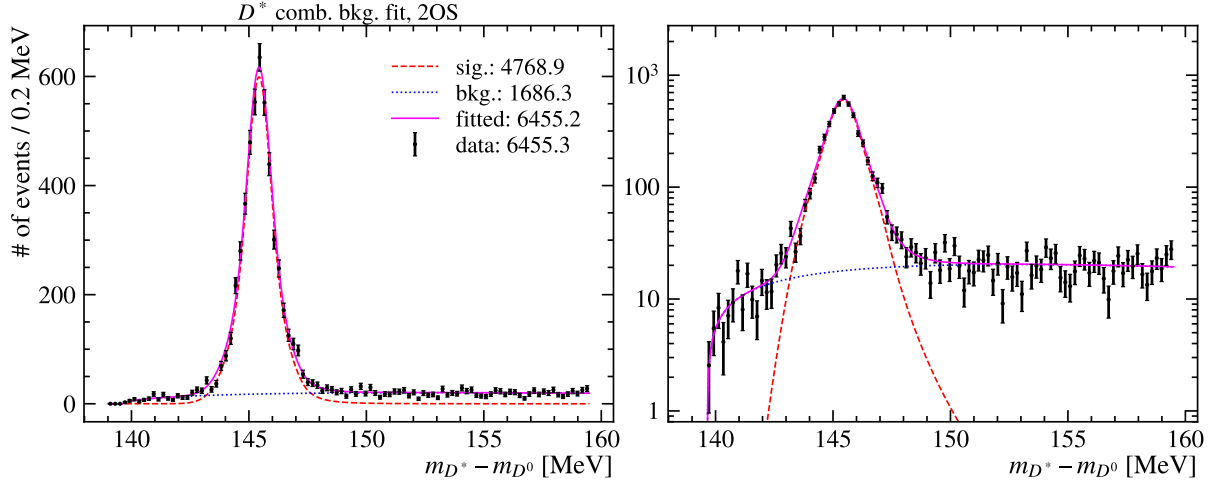


Figure A.4: D^* combinatorial auxiliary fit to the 2OS skim.

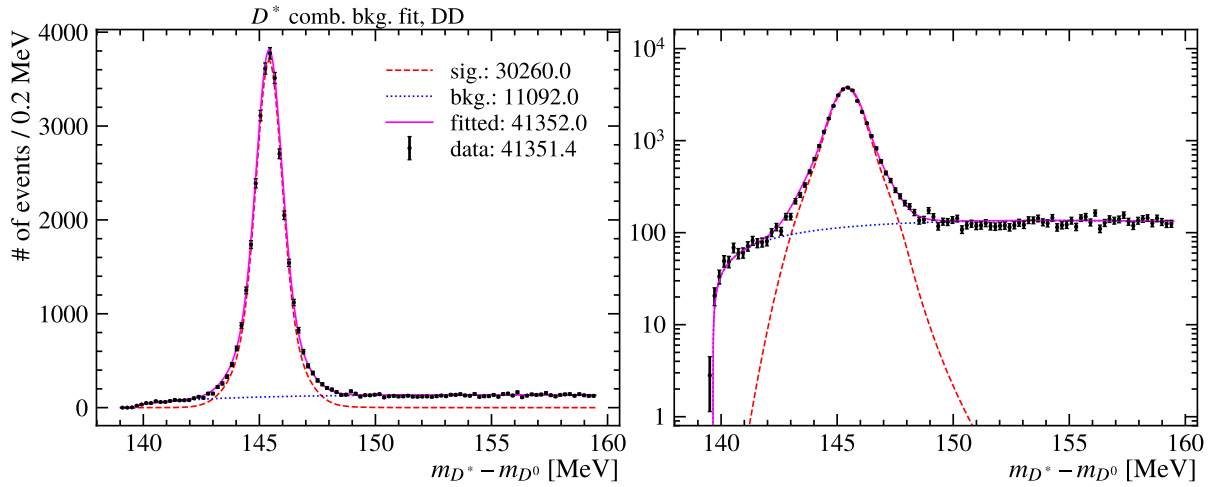


Figure A.5: D^* combinatorial auxiliary fit to the DD skim.

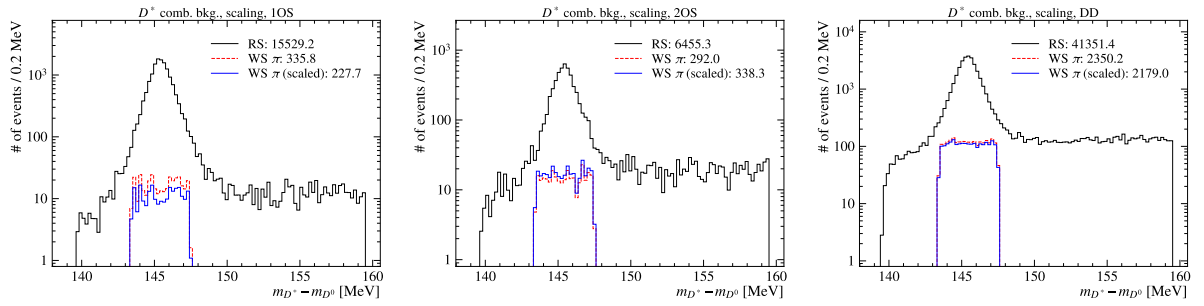


Figure A.6: Effect of scaling for the 1OS, 2OS, and DD skims on the D^* combinatorial templates.

A.3 Compare L0Global TIS efficiencies in MC with TISTOS method

It is discussed that $\epsilon_{\text{TISTOS}} = \epsilon_{\text{TIS}}$ in $B \rightarrow J/\psi K$ MC, thus TISTOS method works (on $J/\psi K$).

We show in Fig. A.7 that TISTOS method does not work in our MC samples. Therefore, in the main text, we use Eq. (6.1) to compute efficiencies in MC.

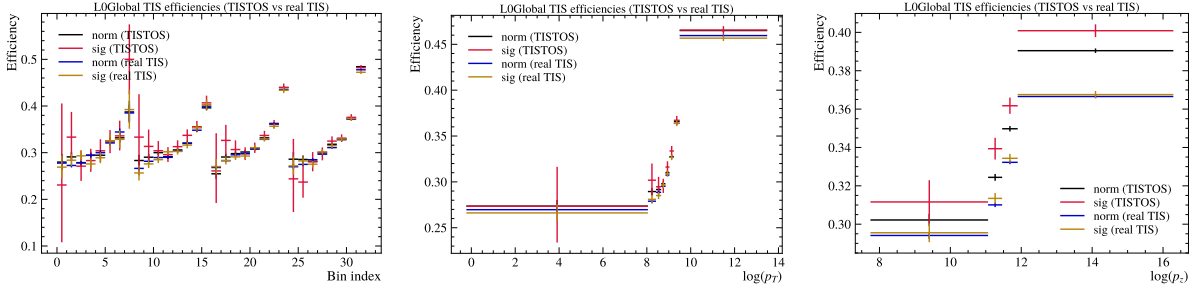


Figure A.7: Check if TISTOS method works on our MC samples.

The p_T and p_z momenta are MC true momenta of the B meson. Bin index represents the index from an unrolled 2D true- p_T -true- p_z histogram.

Efficiencies at high- $\log(p_T)$ regions do not agree. Thus, if TISTOS method is used, the conclusion that L0Global TIS is portable among MC modes cannot be drawn.

A.4 Initial fit results for post-fit cocktail generation

The pre-control fit (step-1) results for the generation of post-fit cocktail are shown in Figs. A.8 and A.9; the control fit (step-2) results are in Figs. A.10 and A.11; the signal fit (step-3) results are in Figs. A.12 and A.13.

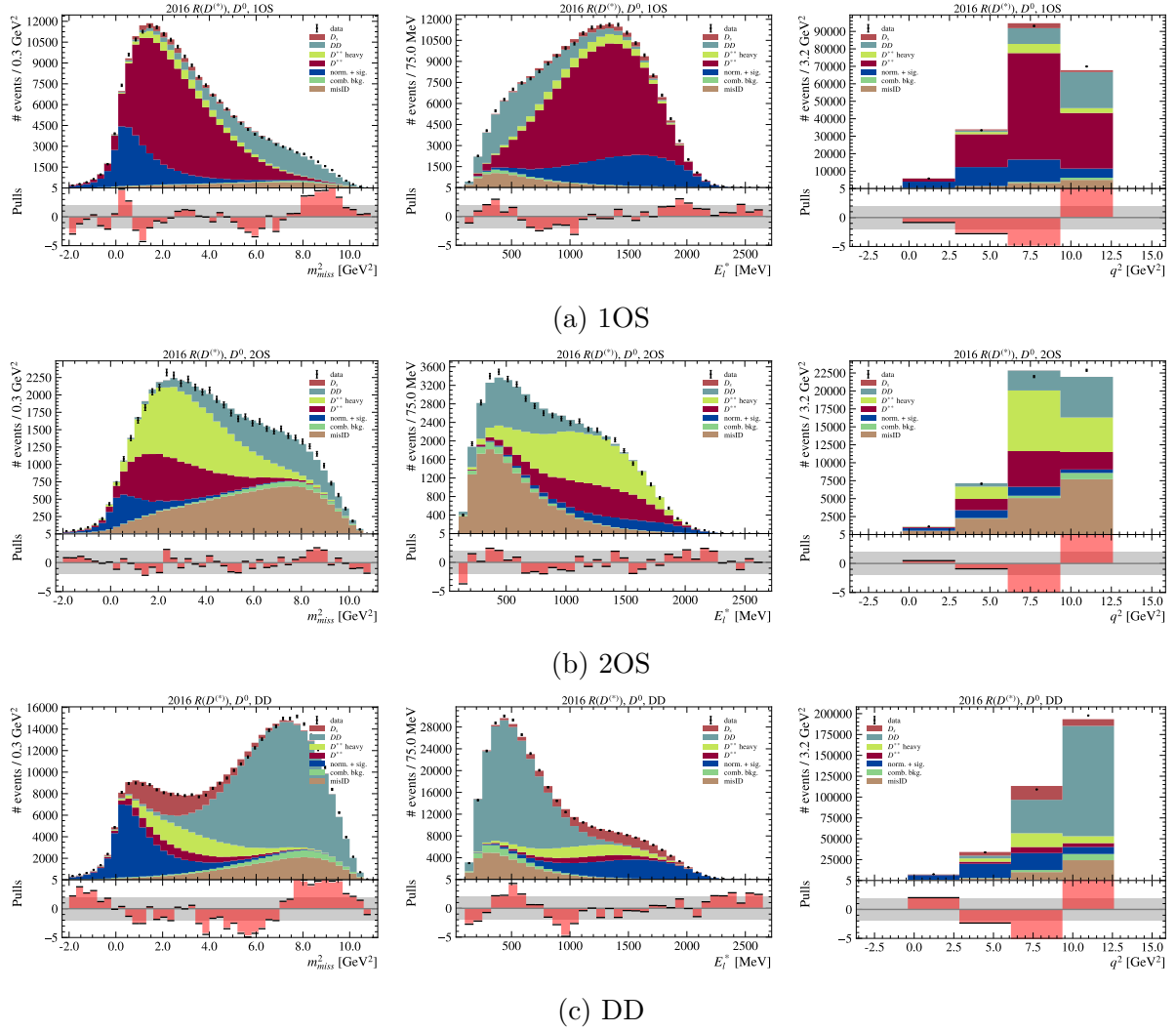


Figure A.8: Pre-control fit of *initial fit* in D^0 channel, on 1OS, 2OS, and DD samples.

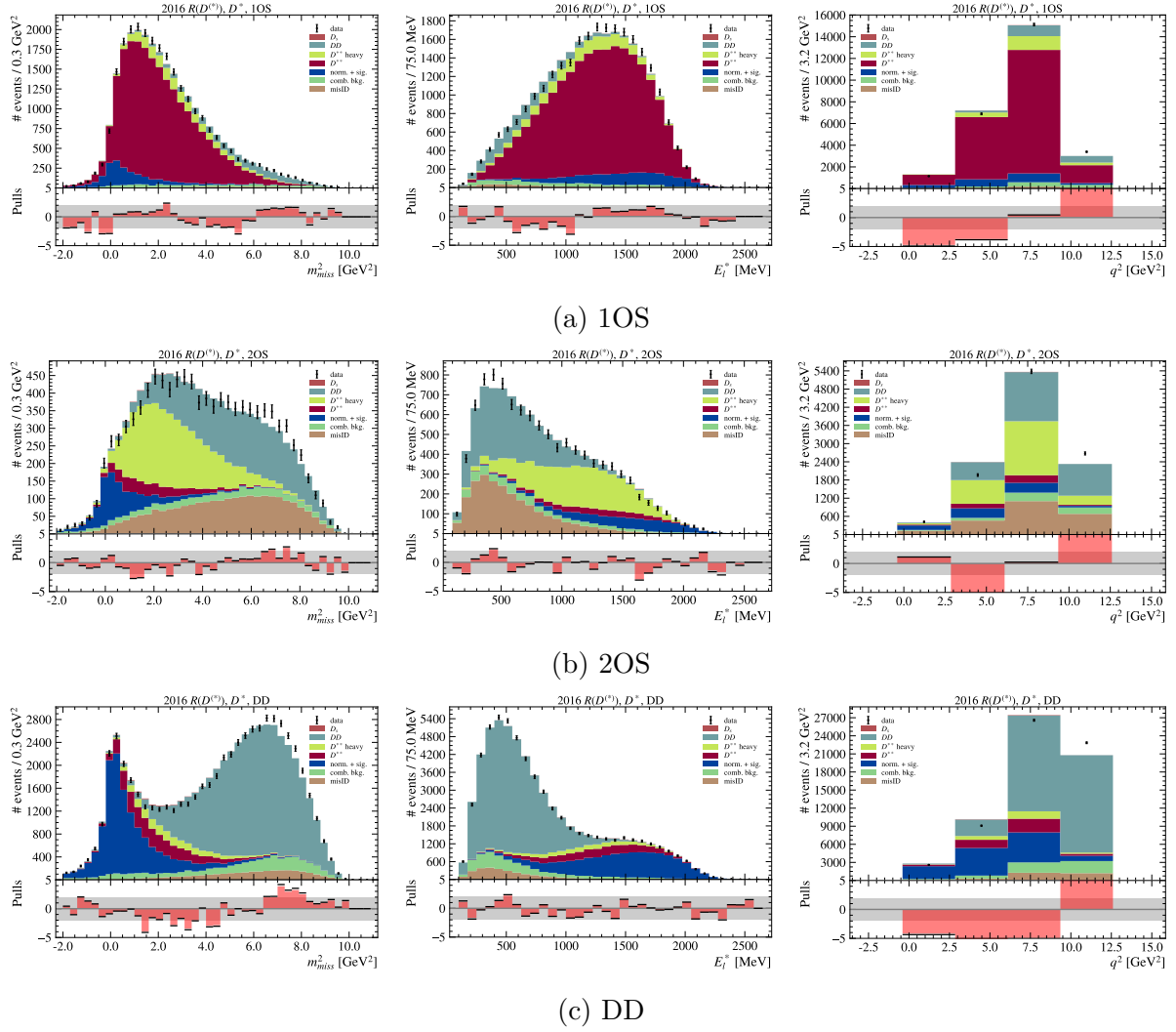


Figure A.9: Pre-control fit of *initial fit* in D^* channel, on 1OS, 2OS, and DD samples.

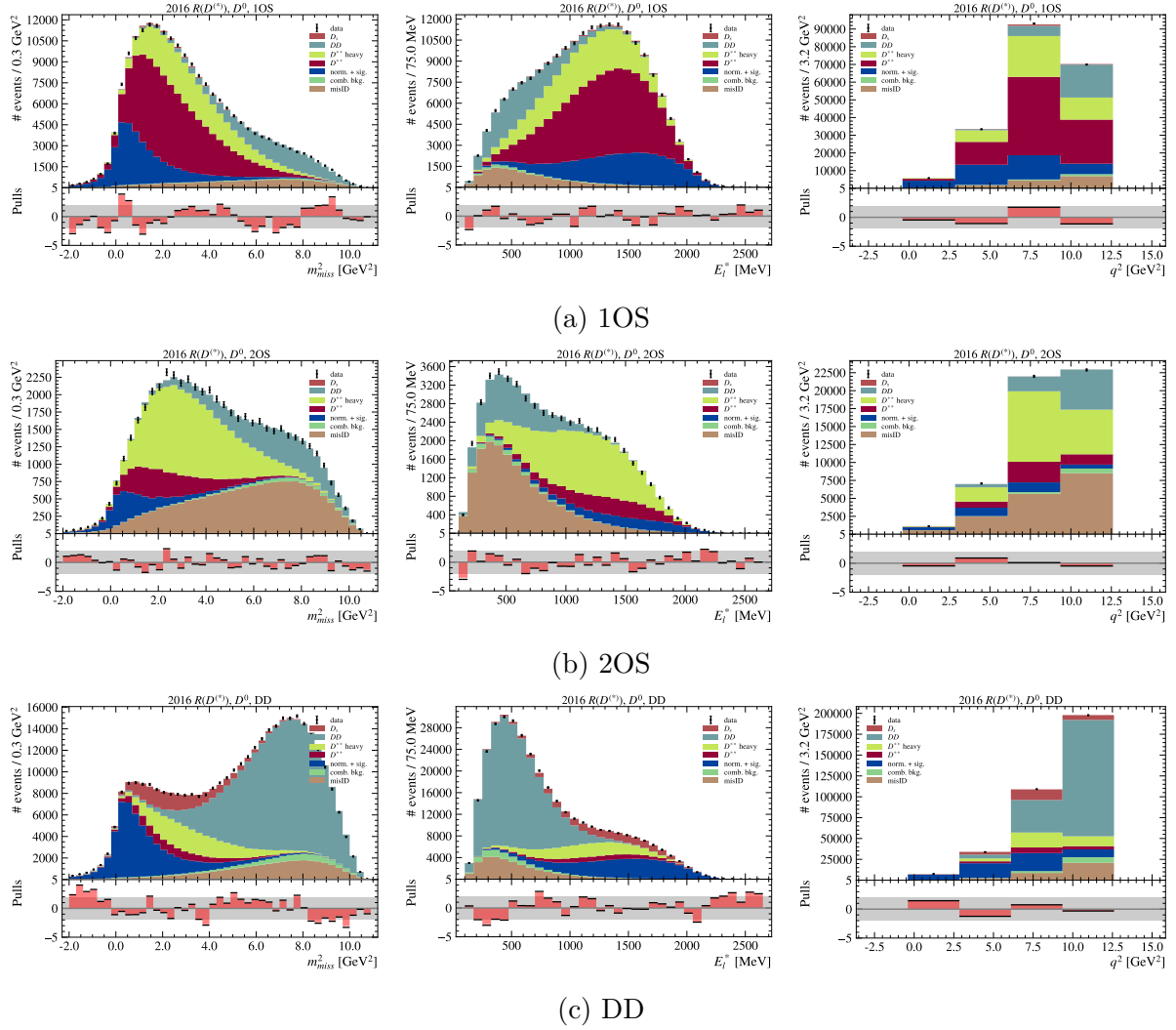


Figure A.10: Control fit of *initial fit* in D^0 channel, on 1OS, 2OS, and DD samples. The Gaussian likelihood penalty on D^{**} form factor uncertainties are kept.

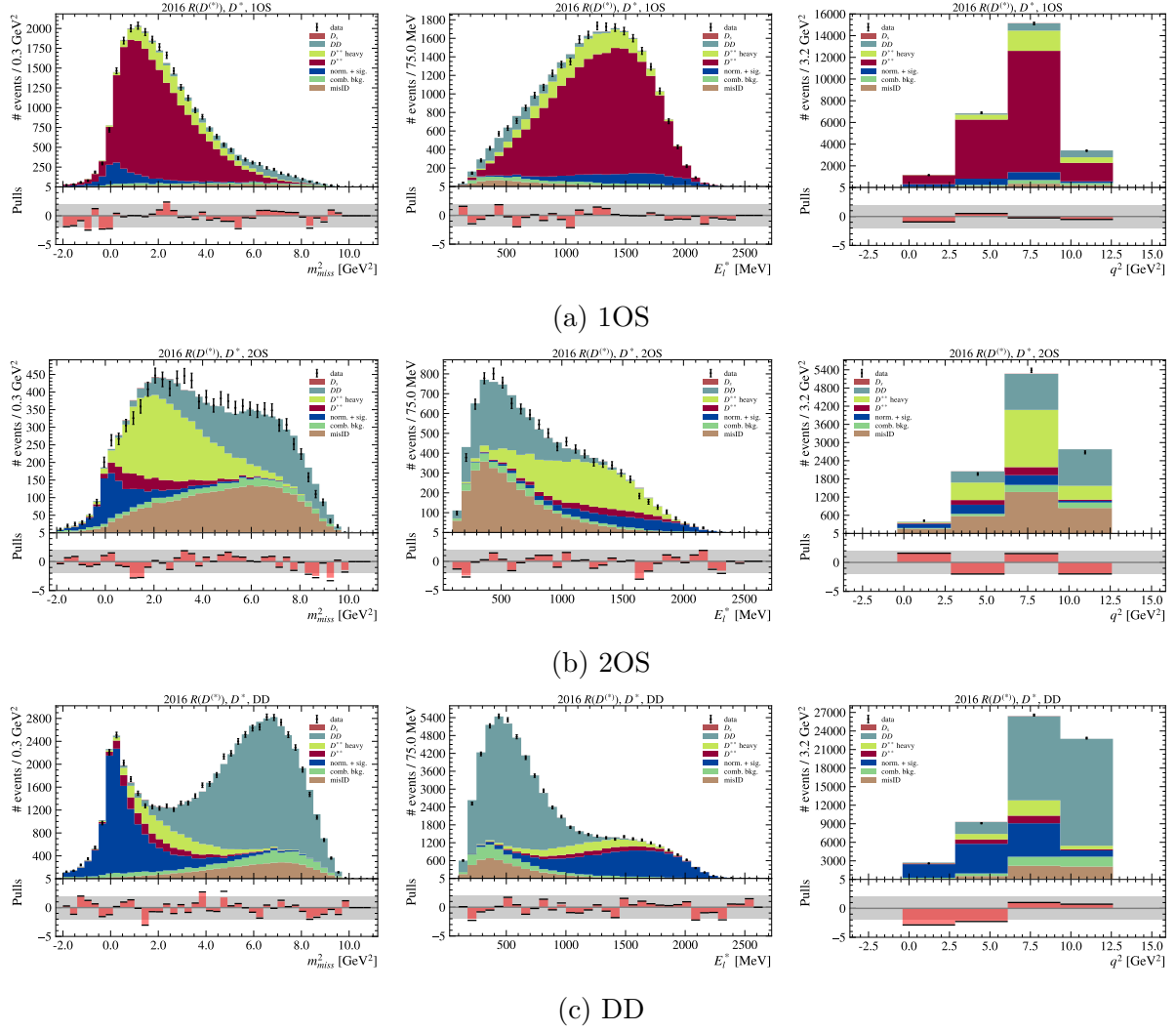


Figure A.11: Control fit of *initial fit* in D^* channel, on 1OS, 2OS, and DD samples. The Gaussian likelihood penalty on D^{**} form factor uncertainties are kept.

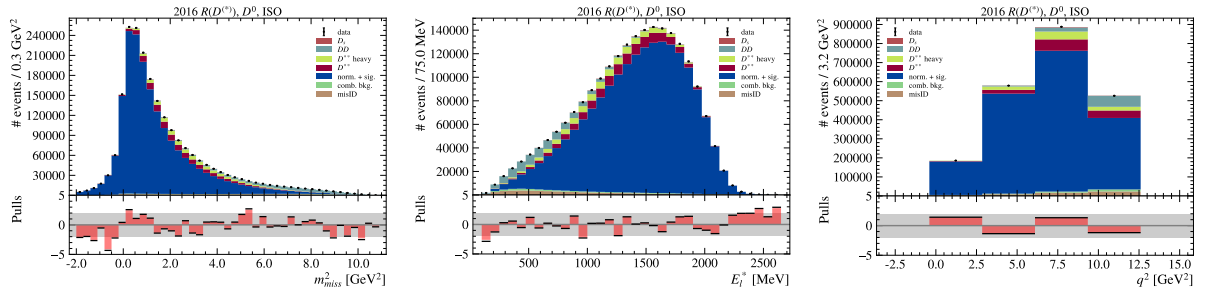


Figure A.12: Signal fit of *initial fit* in D^0 channel on ISO samples.

A.5 Pre-control fit result

The pre-control fit results are listed in Figs. A.14 to A.16 for the D^0 channel; Figs. A.17 to A.19 for the D^* channel.

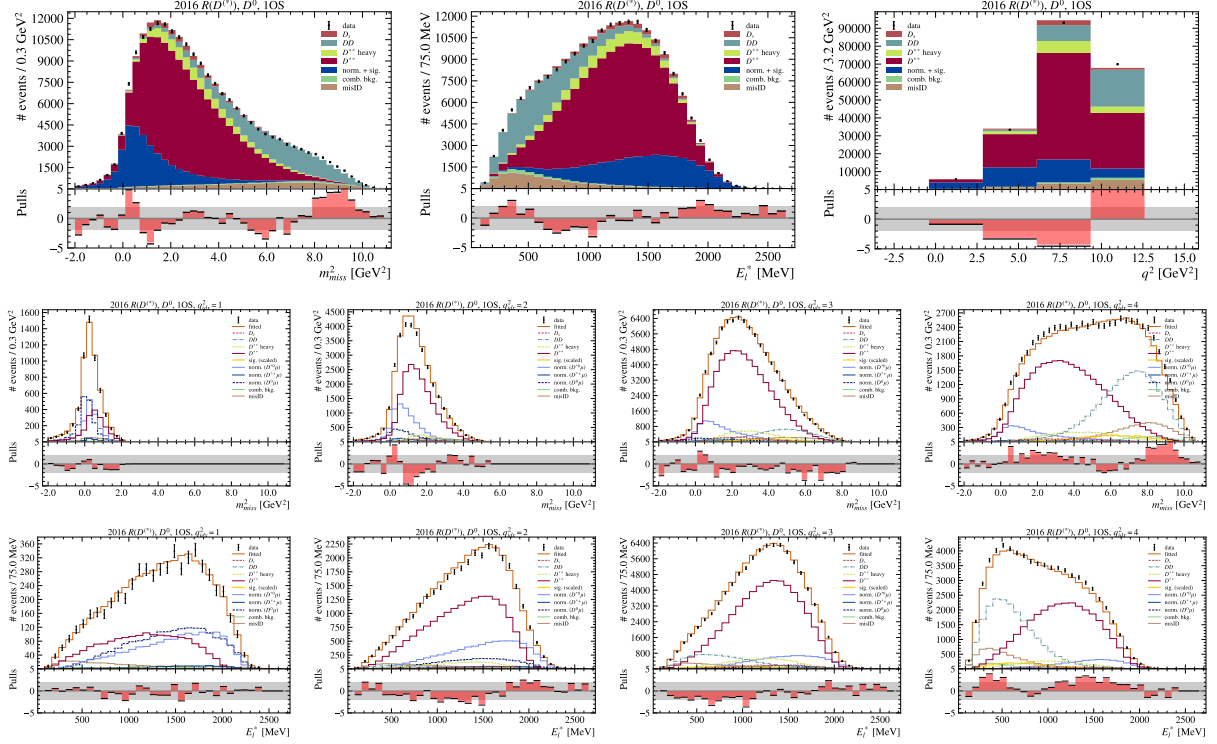


Figure A.14: Pre-control fit for 1OS sample, D^0 channel.

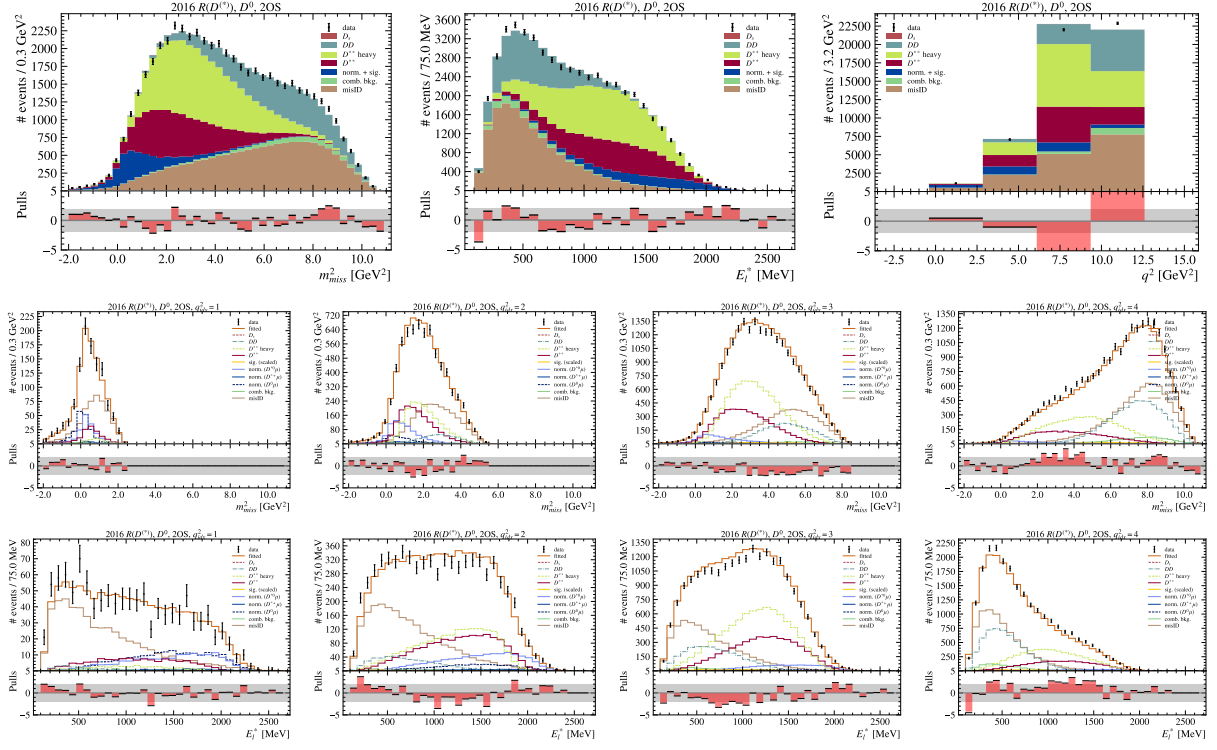


Figure A.15: Pre-control fit for 2OS sample, D^0 channel.

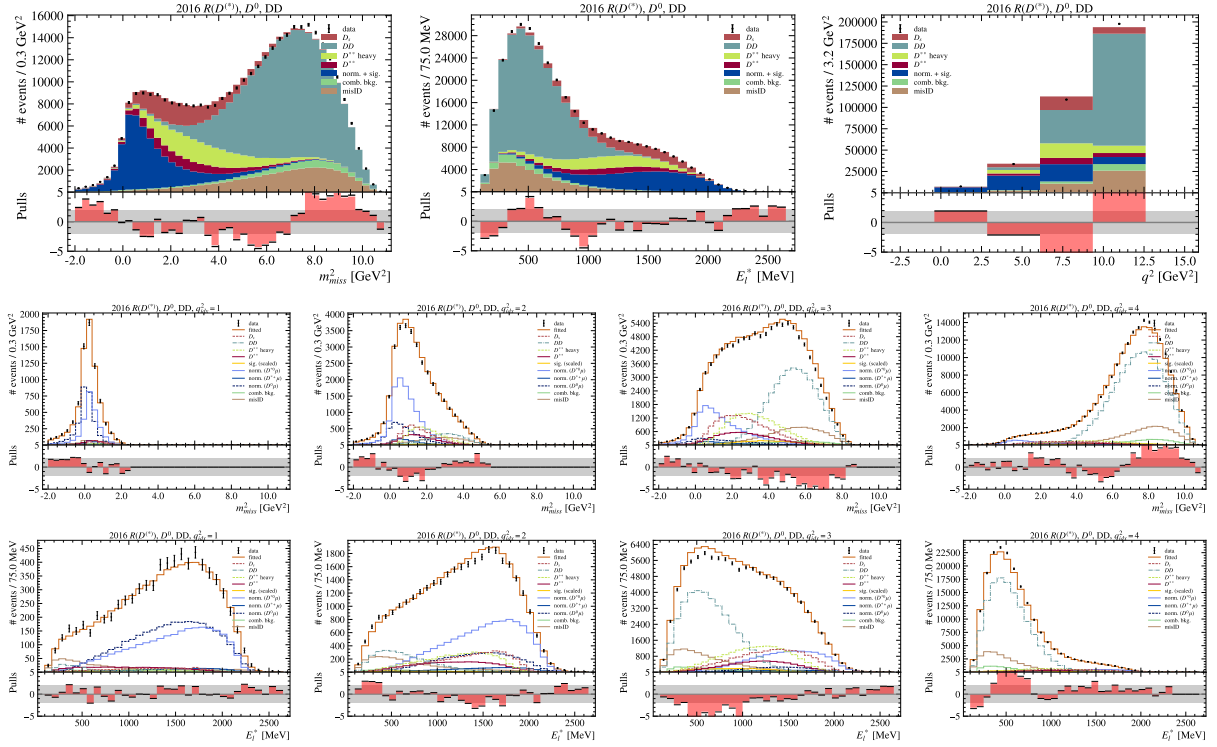


Figure A.16: Pre-control fit for DD sample, D^0 channel.

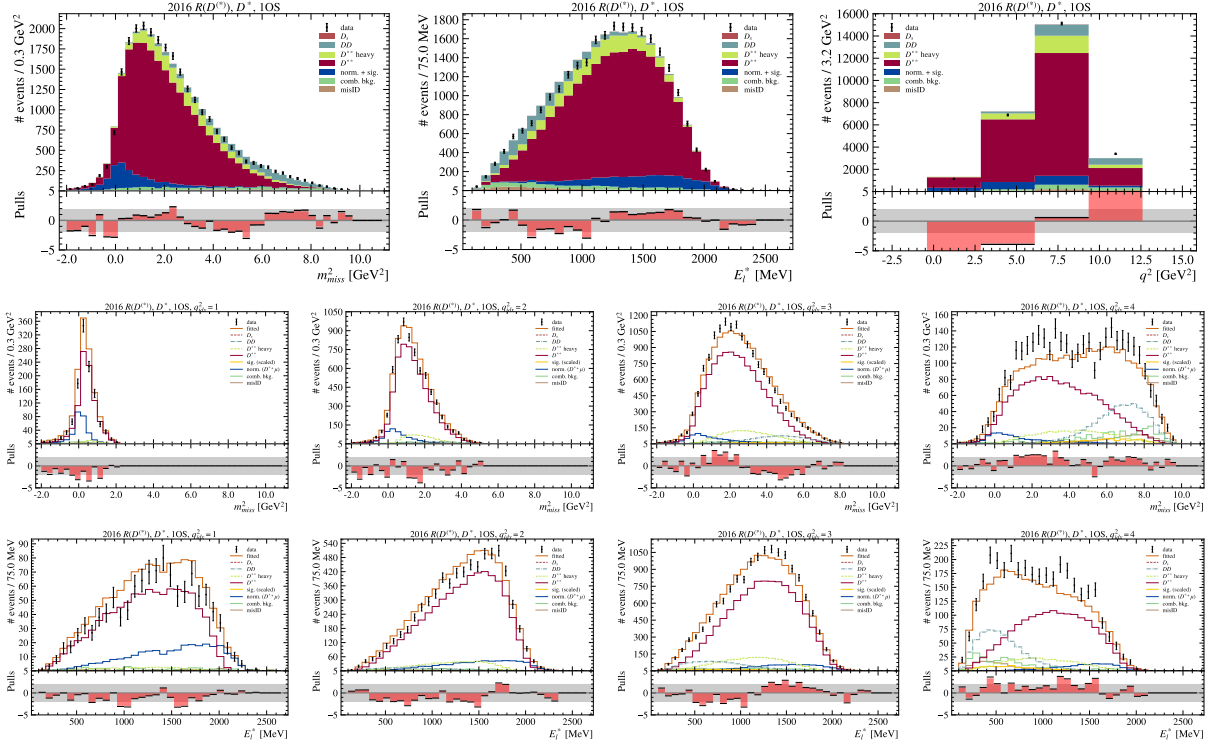


Figure A.17: Pre-control fit for 1OS sample, D^* channel.

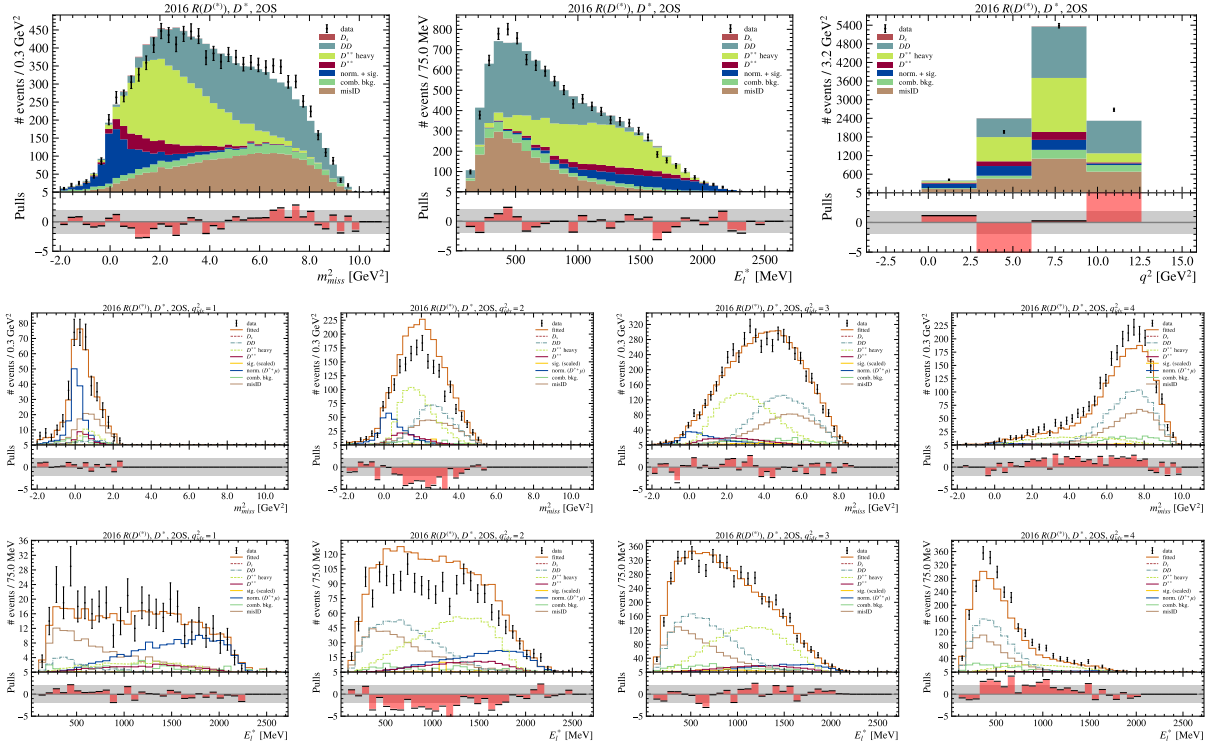


Figure A.18: Pre-control fit for 2OS sample, D^* channel.

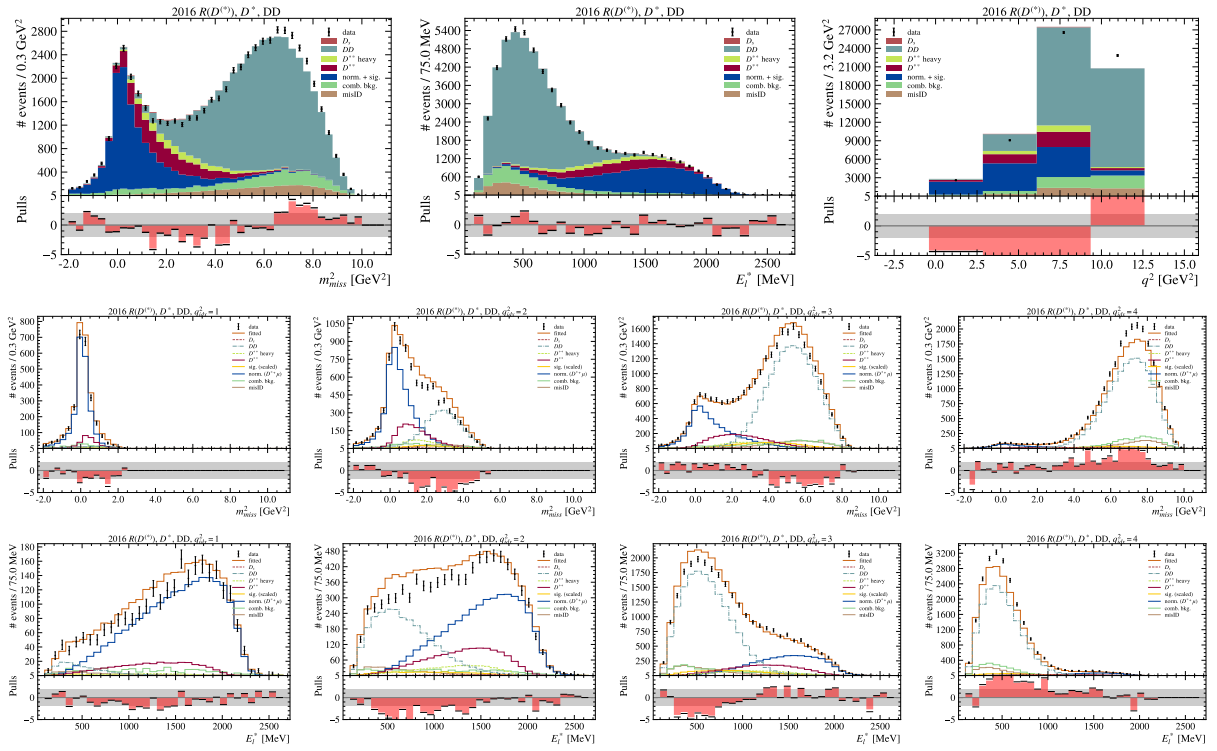


Figure A.19: Pre-control fit for DD sample, D^* channel.

Appendix B: Formal description of selected methods used in this analysis

B.1 Shift negative efficiencies

Sometimes efficiencies obtained from *sPlot* is negative, which is unphysical. Here we present a procedure to shift negative efficiencies back to positive, using the mean μ and standard deviation σ from *sPlot*.

For a given efficiency, assuming Gaussian distributed, we can compute the total probability between $[0, 1]$, the region where the efficiency have a physical meaning, then pick μ_{shifted} such that $P([0, \mu_{\text{shifted}}]) = P([\mu_{\text{shifted}}, 1])$ as the new efficiency.

For a Gaussian distribution with mean μ and standard deviation σ , its cumulative distribution function (c.d.f) is given by:

$$P([-\infty, x]) = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{x - \mu}{\sigma\sqrt{2}} \right) \right) \quad (\text{B.1})$$

where erf is the error function.

Therefore, the total probability between $[0, 1]$ is:

$$P([0, 1]) = \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{1 - \mu}{\sigma\sqrt{2}} \right) \right) - \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{0 - \mu}{\sigma\sqrt{2}} \right) \right) \quad (\text{B.2})$$

$$= \frac{1}{2} \left(\operatorname{erf} \left(\frac{1 - \mu}{\sigma\sqrt{2}} \right) - \operatorname{erf} \left(\frac{0 - \mu}{\sigma\sqrt{2}} \right) \right) \quad (\text{B.3})$$

And the c.d.f evaluated at μ_{shifted} is:

$$P([-\infty, \mu_{\text{shifted}}]) = P([-\infty, 0]) + \frac{1}{2} P([0, 1]) \quad (\text{B.4})$$

$$= \frac{1}{2} \left[1 + \frac{1}{2} \left(\operatorname{erf} \left(\frac{0 - \mu}{\sigma\sqrt{2}} \right) + \operatorname{erf} \left(\frac{1 - \mu}{\sigma\sqrt{2}} \right) \right) \right] \quad (\text{B.5})$$

Define $k \equiv \frac{1}{2} \left(\operatorname{erf} \left(\frac{0-\mu}{\sigma\sqrt{2}} \right) + \operatorname{erf} \left(\frac{1-\mu}{\sigma\sqrt{2}} \right) \right)$, using erf^{-1} , we find the expression for μ_{shifted} :

$$\mu_{\text{shifted}} = \sigma\sqrt{2}\operatorname{erf}^{-1}k + \mu \quad (\text{B.6})$$

B.2 Apply fit yields as a weight

Often times, it is desirable to look at distributions of variables other than fit variables. One can construct such distributions with nominal fit template building procedure by replacing fit variables with variables of interest, but this approach is not very flexible: For example, if a different binning scheme is required, the distributions need to be rebuilt from scratch.

It is possible to construct an event-by-event weight u such that building distributions of any variable (of arbitrary binning scheme, possibly unbinned) with u gives correct fitted yield and shape. In the following paragraphs, a method of constructing u is provided.

The weight u is constructed from fit templates and fitted yields. A typical fit template consists of a nominal template and several variational ones. Denote:

- A fit template can be represented by its yield N .¹
- A nominal fit template with yield² N is constructed with weight w and scaling factor³ $s \equiv 1$. Denote as: $(N; w, 1)$.
- Each variation is represented by a parameter α_i , with two templates specifying variations at $\alpha_i = \pm 1$.

Denote $\pm$ variation templates as $(N_{\alpha_i}^{\pm}; w_{\alpha_i}^{\pm}, s_{\alpha_i}^{\pm})$.

When α_i is not at ± 1 , an interpolation function f is used to approximate template at this given α_i . Denote the (yield of the) interpolated template as $f(\alpha_i; N, N_{\alpha_i}^+, N_{\alpha_i}^-)$.

¹When a binned template has only 1 bin, it is exactly represented by N .

²By *yield*, we mean `histogram.Integral()`.

³The scaling factor s is unabsorbed from w for convenience.

Thus, the final variated template of yield N_f with m variations can be written as:

$$N_f = N + \sum_i^m f(\alpha_i; N, N_{\alpha_i}^+, N_{\alpha_i}^-) \quad (\text{B.7})$$

Suppose the binning is fine enough such that *each event is in its own bin*, then for each bin $N = w \cdot s = w$. Thus Eq. (B.7) can be rewritten as:

$$w_f = w + \sum_i^m f(\alpha_i; w, w_{\alpha_i}^+ \cdot s_{\alpha_i}^+, w_{\alpha_i}^- \cdot s_{\alpha_i}^-) \quad (\text{B.8})$$

Noting N_f and fitted yield N_{fitted} are related by a constant scaling factor $k \equiv \frac{N_{\text{fitted}}}{N_f}$, the expression for u is obtained:

$$u = k \cdot \left(w + \sum_i^m f(\alpha_i; w, w_{\alpha_i}^+ \cdot s_{\alpha_i}^+, w_{\alpha_i}^- \cdot s_{\alpha_i}^-) \right) \quad (\text{B.9})$$

Remark Often (but not always!) variation templates are required to have the same yield as the nominal template (this is to represent a *shape-only* variation), thus a scaling factor $s_{\alpha_i}^{\pm} \equiv \frac{N}{N_{\alpha_i, \text{raw}}^{\pm}}$ is applied on top of $w_{\alpha_i}^{\pm}$.

Appendix C: Rest frame approximation

As discussed in the main text, the rest frame variables, namely m_{miss}^2 , E_μ^* , q^2 , cannot be computed exactly in a hadron collider because the B rest frame is unknown¹. To estimate these variables, an assumption is made: the *estimated proper velocity*, $\gamma\beta$, of the B meson along the beam direction (z direction) is the same as the visible one, which leads to a relationship between the estimated and visible B momenta in the z direction:

$$(p_B)_z = \frac{m_B}{m_{\text{vis}}}(p_{\text{vis}})_z \quad (\text{C.1})$$

at LHCb, the B decay vertex is measured to high precision. Therefore, the B flight direction (orange in Fig. C.1) and its angle respect to the z axis (blue α in the same figure) are both well-measured, which leads to the estimated B momentum:

$$|p_B| = \frac{m_B}{m_{\text{vis}}}(p_{\text{vis}})_z \sqrt{1 + \tan^2 \alpha} \quad (\text{C.2})$$

pointing in the B flight direction.

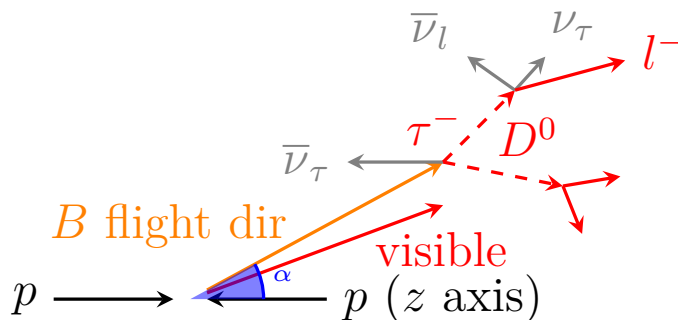


Figure C.1: Schematic diagram for rest frame approximation (RFA).

¹As a reminder, the dominating $b\bar{b}$ production mechanism in hadron colliders is gluon fusion, in which case the gluon parton momenta are unknown.

Appendix D: Trigger categories

Figure D.1 and the text in this appendix are paraphrased from [37]. It is included for completeness of this document.

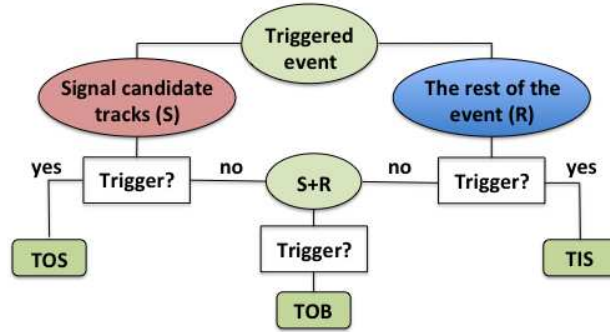


Figure D.1: Schematic of the trigger categories.

- Trigger On Signal (TOS): The signal track (and its decay chain) alone is sufficient to fire the trigger.
- Trigger Independent of Signal (TIS): The rest of event alone is sufficient to fire the trigger.
- TISTOS: Both signal and the rest of event alone are sufficient to generate a positive trigger decision.
- Trigger On Both (TOB): Neither signal nor the rest of event alone is sufficient to generate a positive trigger decision, only the combination is. Note that all event must have a positive trigger decision for it to be saved.

Appendix E: Unfolding procedure of fake muon control sample

This appendix provides more technical information on step 2–4 in μ misID unfolding.

Step 2 The efficiency $\epsilon(t \rightarrow \hat{t}')$ for a track of true species t satisfying requirements of a tagged species $\hat{t}'$ is obtained from t -enriched PIDCALIB2 samples which is taken as truth. Let us first define the subsamples used to evaluate the efficiency, as shown schematically in Fig. E.1:

- t_{PIDCalib} : The t -enriched PIDCALIB2 sample itself (or MC in the case of ghost). The sample contains minimal cuts.
- t_{acc} : The subset of t_{PIDCalib} that satisfies all μ cuts other than PID (i.e., track in μ acceptance). Note that this subsample is further divided into two subsamples: t_{fake} and $\hat{\mu}$.
- t_{fake} : The subset of t_{acc} that *fails* μ PID from Table 5.4.
- $\hat{\mu}_t$: The subset of t_{acc} that *passes* μ PID
- $\hat{t}'_{\text{fake}}$: The subset of t_{fake} that passes the $\hat{t}'$ tagging requirements from Table 8.6.

The efficiency $\epsilon(t \rightarrow \hat{t}')$ of a true t to be classified as $\hat{t}'$ can be computed as:

$$\epsilon(t \rightarrow \hat{t}') = \epsilon(t_{\text{fake}} \rightarrow \hat{t}'_{\text{fake}}) = \frac{n_{\hat{t}'_{\text{fake}}}}{n_{t_{\text{fake}}}} \quad (\text{E.1})$$

where n denote yields in the PIDCALIB2 *fake* subsamples as shown in Fig. E.1. All efficiencies are binned in p_B , η_B , and nTracks with a consistent binning scheme.

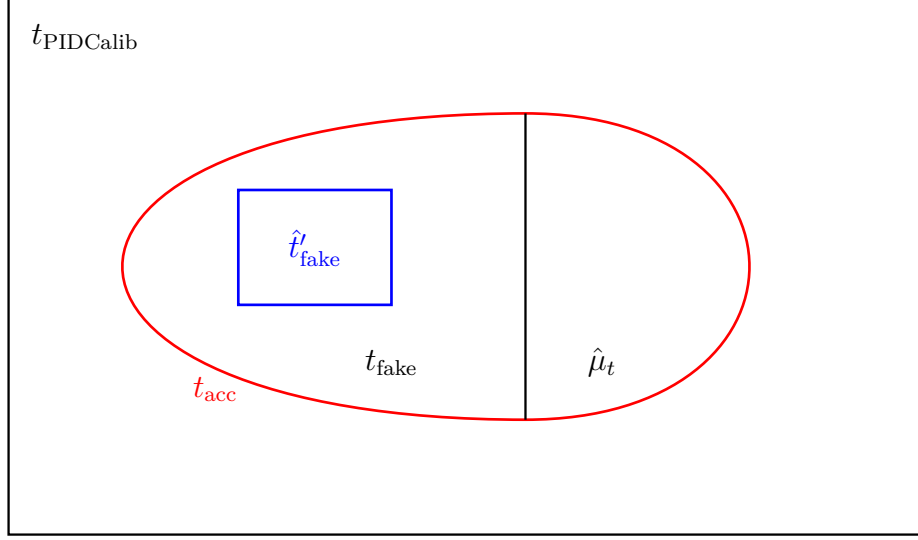


Figure E.1: Subsamples of t -enriched PIDCALIB2 samples used in unfolding.

Step 3 In a particular p_B - η_B -nTracks bin, the measured yields $\tilde{n}_{\hat{t}'}$, the true yields $\tilde{n}_t$, and the response matrix M have the following relation:

$$\begin{pmatrix} \tilde{n}_{\hat{\pi}} \\ \tilde{n}_{\hat{K}} \\ \tilde{n}_{\hat{p}} \\ \tilde{n}_{\hat{e}} \\ \tilde{n}_{\hat{g}} \end{pmatrix} = \begin{pmatrix} \epsilon(\pi \rightarrow \hat{\pi}) & \epsilon(K \rightarrow \hat{\pi}) & \epsilon(p \rightarrow \hat{\pi}) & \epsilon(e \rightarrow \hat{\pi}) & \epsilon(g \rightarrow \hat{\pi}) \\ \epsilon(\pi \rightarrow \hat{K}) & \epsilon(K \rightarrow \hat{K}) & \epsilon(p \rightarrow \hat{K}) & \epsilon(e \rightarrow \hat{K}) & \epsilon(g \rightarrow \hat{K}) \\ \epsilon(\pi \rightarrow \hat{p}) & \epsilon(K \rightarrow \hat{p}) & \epsilon(p \rightarrow \hat{p}) & \epsilon(e \rightarrow \hat{p}) & \epsilon(g \rightarrow \hat{p}) \\ \epsilon(\pi \rightarrow \hat{e}) & \epsilon(K \rightarrow \hat{e}) & \epsilon(p \rightarrow \hat{e}) & \epsilon(e \rightarrow \hat{e}) & \epsilon(g \rightarrow \hat{e}) \\ \epsilon(\pi \rightarrow \hat{g}) & \epsilon(K \rightarrow \hat{g}) & \epsilon(p \rightarrow \hat{g}) & \epsilon(e \rightarrow \hat{g}) & \epsilon(g \rightarrow \hat{g}) \end{pmatrix} \begin{pmatrix} \tilde{n}_{\pi} \\ \tilde{n}_K \\ \tilde{n}_p \\ \tilde{n}_e \\ \tilde{n}_g \end{pmatrix} \quad (\text{E.2})$$

where $\tilde{n}$ denote yields in the fake μ control sample as described in Section 5.1.3, and $\epsilon(t \rightarrow \hat{t}')$ are obtained from step-2.

To obtain the true yields $\tilde{n}_t$, naive matrix inversion cannot be used because it is sensitive to statistical fluctuations [78]. The Bayesian (iterative¹) unfolding method provided by RooUnfold [80] is used instead. The main idea is sketched below:

- ① Consider the true yield for π of a tagged species $\hat{t}'$ only.
- ② Given that all true π to a tagged species $\hat{t}''$ efficiencies $\epsilon(\pi \rightarrow \hat{t}'')$ are known, assuming

¹As described in [79], some people prefer to label this method as *iterative* rather than Bayesian as this method only uses Bayes' theorem in an intermediate step but does not explicitly produce a posterior p.d.f based on likelihood and prior p.d.f.s (both are specified by the user).

initially the true yield $\tilde{n}_\pi$ equals to the tagged yield $\tilde{n}_{\hat{\pi}}$, for a particular tagged species $\hat{t}'$, its efficiencies to a *true* π can be written as:

$$\epsilon(\hat{t}' \rightarrow \pi)_0 = \frac{\tilde{n}_{\hat{\pi}} \cdot \epsilon(\pi \rightarrow \hat{t}')}{\sum_t \tilde{n}_{\hat{t}} \cdot \epsilon(t \rightarrow \hat{t}')} \quad (\text{E.3})$$

where the subscript 0 indicates that this is an initial estimate of the efficiency.

- ③ An initial estimate of the yield of true π can be computed:

$$N_0^\pi = \sum_{\hat{t}'} \tilde{n}_{\hat{t}'} \cdot \epsilon(\hat{t}' \rightarrow \pi) \quad (\text{E.4})$$

where $\tilde{n}_{\hat{t}'}$ is the *tagged yield* as reported by the measurement.

- ④ Compute a χ^2 between N_0^π and the initial estimated true yield taken to be equal to $\tilde{n}_{\hat{\pi}}$. If the χ^2 is large-ish, go back to ② with the N_0^π as the estimated true yield, instead of $\tilde{n}_{\hat{\pi}}$.

The key insight to this unfolding approach is: The outcome (measured tagged yield) can only be estimated with a good knowledge of the true yield. Each iteration improves the knowledge on the true yield. However, exact converge should not be pursued as that likely results in overfitting. For this analysis, 5 iterations are used. Now it is time to find $\epsilon(\hat{t}' \rightarrow t)$. Recall its definition:

$$\epsilon(\hat{t}' \rightarrow t) \equiv \frac{\text{Number of } \hat{t}'\text{-tagged tracks from true } t \text{ tracks}}{\text{Total number of } \hat{t}'\text{-tagged tracks based on unfolding}} \quad (\text{E.5})$$

With the unfolded true yields, we can compute:

- Number of $\hat{t}'$ -tagged tracks in true t tracks: $\tilde{n}_t \cdot \epsilon(t \rightarrow \hat{t}')$.
- Total number of $\hat{t}'$ -tagged tracks based on unfolding: $\sum_t \tilde{n}_t \cdot \epsilon(t \rightarrow \hat{t}')$. The definition used here ensures the total number of tracks is conserved by unfolding.

Put them together (noting the similarities between this equation and Eq. (E.3)!):

$$\epsilon(\hat{t}' \rightarrow t) = \frac{\tilde{n}_t \cdot \epsilon(t \rightarrow \hat{t}')}{\sum_t \tilde{n}_t \cdot \epsilon(t \rightarrow \hat{t}')} \quad (\text{E.6})$$

Step 4 The transfer factor $\epsilon(\hat{t}' \rightarrow \hat{\mu})$ can be computed with:

$$\epsilon(\hat{t}' \rightarrow \hat{\mu}) = \epsilon(\hat{t}' \rightarrow t_{\text{acc}}) \cdot \epsilon(t_{\text{acc}} \rightarrow \hat{\mu}) \quad (\text{E.7})$$

The first efficiency $\epsilon(\hat{t}' \rightarrow t_{\text{acc}})$ is not directly measurable but can be converted from the unfolded $\epsilon(\hat{t}' \rightarrow t)$ by noting that the *fake* μ control samples, inputs to the unfolding, is directly comparable to the t_{fake} subsample in Fig. E.1. The conversion formula is:

$$\epsilon(\hat{t}' \rightarrow t_{\text{acc}}) = \frac{\epsilon(\hat{t}' \rightarrow t)}{\epsilon(t_{\text{acc}} \rightarrow t)} \approx \frac{\epsilon(\hat{t}' \rightarrow t)}{\epsilon(t_{\text{acc}} \rightarrow t_{\text{fake}})} \quad (\text{E.8})$$

note that the nominator uses t/t' as labels to species, whereas the denominator uses $t_{\text{acc}}/t_{\text{fake}}$ as labels to subsamples in the t -enriched PIDCALIB2 sample. This is unfortunate but clearer notations can be much more verbose.

The second efficiency $\epsilon(t_{\text{acc}} \rightarrow \hat{\mu})$ is measured directly in t -enriched PIDCALIB2 sample:

$$\epsilon(t_{\text{acc}} \rightarrow \hat{\mu}) = \epsilon(t_{\text{acc}} \rightarrow \hat{\mu}_t) = \frac{n_{\hat{\mu}_t}}{n_{t_{\text{acc}}}} \quad (\text{E.9})$$

Appendix F: Form factor variation

Suppose a form factor $\mathcal{F}$ can be estimated with:

$$\mathcal{F}(z) \approx f_0 + f_1 z + f_2 z^2 \quad (\text{F.1})$$

where z is some expansion parameter, and $\mathbf{f} = (f_0, f_1, f_2)^T$ are expansion coefficients.

The log-likelihood of the polynomial can be written as:

$$\log L(\mathbf{f}) = -\frac{1}{2} \underbrace{\sum_i^N \frac{(\mathcal{F}_i^{\text{mea}} - \mathbf{z}_i^T \mathbf{f})^2}{\sigma_i^2}}_{\equiv \chi^2} \quad (\text{F.2})$$

where $\mathbf{z}_i = (1, z_i, z_i^2)^T$, $\mathcal{F}_i^{\text{mea}}$ indicates each measured value of $\mathcal{F}$, and σ_i the uncertainty associated with each measurement, assuming each measurement is an independent Gaussian random variable.

To obtain best fit values of $\mathbf{f}$, we maximize χ^2 by requiring $\nabla_{\mathbf{f}} \chi^2 = \mathbf{0}$. Noting χ^2 can be rewritten more concisely in a tensorial form:

$$\chi^2 = (\mathcal{F}^{\text{mea}} - Z\mathbf{f})^T V^{-1} (\mathcal{F}^{\text{mea}} - Z\mathbf{f}) \quad (\text{F.3})$$

where Z is a $N \times 3$ matrix with $Z_i \equiv \mathbf{z}_i = (1, z_i, z_i^2)^T$, and V a $N \times N$ matrix with $V_{ij} = \delta_{ij} \sigma_j^2$.

Computing $\nabla_{\mathbf{f}} \chi^2$:

$$\nabla_{\mathbf{f}} \chi^2 = -Z^T V^{-1} (\mathcal{F}^{\text{mea}} - Z\mathbf{f}) - (\mathcal{F}^{\text{mea}} - Z\mathbf{f})^T V^{-1} Z \quad (\text{F.4})$$

$$= -Z^T V^{-1} (\mathcal{F}^{\text{mea}} - Z\mathbf{f}) - Z^T \underbrace{(V^{-1})^T}_{=V^{-1}} (\mathcal{F}^{\text{mea}} - Z\mathbf{f}) \quad (\text{F.5})$$

$$= -2Z^T V^{-1} (\mathcal{F}^{\text{mea}} - Z\mathbf{f}) \quad (\text{F.6})$$

Therefore, the best fit $\hat{\mathbf{f}}$ are linear functions of the measurements $\mathcal{F}^{\text{mea}}$:

$$-2Z^T V^{-1}(\mathcal{F}^{\text{mea}} - Z\hat{\mathbf{f}}) = 0 \quad (\text{F.7})$$

$$\implies Z^T V^{-1} Z \hat{\mathbf{f}} = Z^T V^{-1} \mathcal{F}^{\text{mea}} \quad (\text{F.8})$$

$$\implies \hat{\mathbf{f}} = \underbrace{(Z^T V^{-1} Z)^{-1} Z^T V^{-1} \mathcal{F}^{\text{mea}}}_{\equiv B} \quad (\text{F.9})$$

Using error propagation to find uncertainties of $\hat{\mathbf{f}}$, denote as U :

$$U = B V B^T = (Z^T V^{-1} Z)^{-1} \quad (\text{F.10})$$

In the case where each $\mathcal{F}_i^{\text{mea}}$ is independent and identically distributed with a variance of σ^2 , $V^{-1} = \frac{1}{\sigma^2} \text{Id}_N$, and U (a 3×3 matrix) becomes:

$$U = (Z^T Z)^{-1} \sigma^2 \quad (\text{F.11})$$

However, $\hat{\mathbf{f}}$ are often correlated, making linear expansion in terms of them hard. Still, we can always find a linear transformation A s.t. $\tilde{\mathbf{f}} \equiv A\hat{\mathbf{f}}$ has an identity covariance matrix.

1. Find the eigenvectors $\mathbf{r}_j^i$ satisfying $(\mathbf{r}^i)^T \mathbf{r}^j = \delta_{ij}$, and eigenvalues λ_i of U .

Note that λ_i are all positive because U is positive-definite.

2. For $M_{ij} = \mathbf{r}_j^i$, $M\hat{\mathbf{f}}$ has a covariance matrix element $V_{ij} = M_{ik} U_{kl} M_{lj}^T = \mathbf{r}_k^i U_{kl} \mathbf{r}_l^j = \lambda_j \delta_{ij}$.

3. Let $C_{ij} = \delta_{ij} \frac{1}{\sqrt{\lambda_j}}$, without summing over repeated indices. Then $A = CM$.

4. Note that $\hat{\mathcal{F}}(z) = \mathbf{z}^T \hat{\mathbf{f}} = \mathbf{z}^T A^{-1} \tilde{\mathbf{f}}$. Its linear variations in terms of $\tilde{\mathbf{f}}$ are $\mathbf{z}^T A^{-1} \delta \tilde{\mathbf{f}}$.

In our **Hammer**-based form factor variation, the covariance U are taken from corresponding papers directly, and the expansion coefficients are defined in terms of $\hat{\mathbf{f}}$, which are often correlated. Still, the variations must be uncorrelated due to **HISTFACTORY** requirements.

Thus, a typical variation workflow for generating $\pm 1\sigma$ form factor variations is:

1. Find A^{-1} as defined in the previous paragraph.
2. Provide form factor variation information to **Hammer**. We have 2 equivalent procedures:

- We can supply the A^{-1} to `Hammer` form factor class directly.

For example, for `BGLVar` form factor of $B \rightarrow D$ decays, the matrix is named as `apa0mat`.

In this case, the form factor variations are in an *orthonormal* error eigenbasis, so the form of variations are very simple. For example, a $+1\sigma$ variation in the first component can be supplied as¹:

```
ham.setFFEigenvectors{"BtoD", "BGLVar", {"delta_ap0", 1}};
```

- Alternatively, contract $A^{-1}\delta\hat{f}$ manually, and provide the variation in the original basis the form factors are parameterized.

This works well for `BGLVar`, because by default the covariance is set to identity.

In this case, one does not manually input the 2-dimensional covariance matrix manually in C++. The trade-off is: the variations become more complex. For example, the same $+1\sigma$ variation in the error eigenbasis now becomes:

```
ham.setFFEigenvectors("BtoD", "BGLVar",
  {"delta_ap0", -1.7339e-05}, {"delta_ap1", 0.0001643},
  {"delta_ap2", 0.000192972}, {"delta_a01", -0.100070},
  {"delta_a02", 0.002478870}, {"delta_a00", 0.0064648}}
);
```

with precisions reduced for simplicity².

In both cases, the matrix $\mathbf{z}^T A^{-1}$ ($\mathbf{z}^T \text{Id}$ in the second case) is computed only once, which can be contracted with any linear variation at a later time.

¹Here `delta_ap0` is a misnomer.

²Now `delta_ap0` etc. are consistent with the common form factor parameterization in the literature. Also note that the `delta_a00` is fixed by the unitarity requirement.

Bibliography

- [1] R. Aaij et al. “Measurement of the Ratio of Branching Fractions $\mathcal{B}(\overline{B}^0 \rightarrow D^{*+}\tau^-\overline{\nu}_\tau)$ / $\mathcal{B}(\overline{B}^0 \rightarrow D^{*+}\mu^-\overline{\nu}_\mu)$ ”. In: *Phys. Rev. Lett.* 115 (11 Sept. 2015), p. 111803. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.115.111803> (cit. on pp. 2, 42).
- [2] LHCb Collaboration. “Measurement of the ratios of branching fractions $\mathcal{R}(D^*)$ and $\mathcal{R}(D^0)$ ”. In: (2023). URL: <https://arxiv.org/abs/2302.02886> (cit. on pp. 2, 123).
- [3] Y. Amhis et al. “Averages of b -hadron, c -hadron, and τ -lepton properties as of 2021”. In: (June 2022). arXiv: 2206.07501 [hep-ex] (cit. on p. 3).
- [4] R. Aaij et al. “Study of D_J meson decays to $D^+\pi^-$, $D^0\pi^+$ and $D^{*+}\pi^-$ final states in pp collisions”. In: *Journal of High Energy Physics* 2013.9 (Sept. 2013). URL: <https://doi.org/10.1007/jhep09/82013/9145> (cit. on p. 7).
- [5] M. Robinson. *Symmetry and the Standard Model*. Springer, 2011 (cit. on p. 12).
- [6] J. Schwichtenberg. *Physics from Symmetry*. 2nd ed. Springer, 2018 (cit. on p. 12).
- [7] T. Weigand. *Lecture notes used in Physics 624: Quantum Field Theory I+II*. 2017. URL: <http://www.physics.umd.edu/grt/taj/624b/WeigandQFT.pdf> (cit. on p. 14).
- [8] M. Tanaka. “Charged Higgs effects on exclusive semi-tauonicB decays”. In: *Zeitschrift für Physik C Particles and Fields* 67.2 (June 1995), pp. 321–326. URL: <https://doi.org/10.1007/bf01571294> (cit. on p. 19).
- [9] K. Hagiwara, A. Martin, and M. Wade. “Exclusive semileptonic B-meson decays”. In: *Nuclear Physics B* 327.3 (1989), pp. 569–594. ISSN: 0550-3213. URL: <https://www.sciencedirect.com/science/article/pii/0550321389903064> (cit. on p. 20).
- [10] F. U. Bernlochner et al. “Semitauponic b -hadron decays: A lepton flavor universality laboratory”. In: *Reviews of Modern Physics* 94.1 (Feb. 2022). URL: <https://doi.org/10.1103/revmodphys.94.015003> (cit. on pp. 21–23, 27).
- [11] I. Caprini, L. Lellouch, and M. Neubert. “Dispersive bounds on the shape of form factors”. In: *Nuclear Physics B* 530.1-2 (Oct. 1998), pp. 153–181. URL: <https://arxiv.org/abs/hep-ph/9712417> (cit. on p. 22).
- [12] C. G. Boyd, B. Grinstein, and R. F. Lebed. “Constraints on Form Factors for Exclusive Semileptonic Heavy to Light Meson Decays”. In: *Physical Review Letters* 74.23 (June 1995), pp. 4603–4606. URL: <https://doi.org/10.1103/physrevlett.74.4603> (cit. on p. 24).

- [13] C. G. Boyd, B. Grinstein, and R. F. Lebed. “Precision corrections to dispersive bounds on form-factors”. In: *Phys. Rev. D* 56 (1997), pp. 6895–6911. arXiv: hep-ph/9705252 (cit. on p. 24).
- [14] D. Bigi and P. Gambino. “Revisiting $B \rightarrow D\ell\nu$ ”. In: *Physical Review D* 94.9 (Nov. 2016). URL: <https://doi.org/10.1103/physrevd.94.094008> (cit. on pp. 24, 25).
- [15] P. Hamilton, H. Jawahery, and G. M. Ciezarek. “Simultaneous extraction of the branching fraction ratios $\mathcal{R}(D)$ and $\mathcal{R}(D^*)$ with the Run 1 Dataset using the $\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$ decay”. In: LHCb-ANA-2020-056 (2020). URL: <https://cds.cern.ch/record/2728837> (cit. on pp. 26, 42, 48–50, 55, 84, 86, 105, 106, 108, 110, 130).
- [16] A. Bazavov et al. “Semileptonic form factors for $B \rightarrow D^* \ell \nu$ at nonzero recoil from $2 + 1$ -flavor lattice QCD”. In: (2021). URL: <https://arxiv.org/abs/2105.14019> (cit. on pp. 26, 27).
- [17] D. Scora and N. Isgur. “Semileptonic meson decays in the quark model: An update”. In: *Physical Review D* 52.5 (Sept. 1995), pp. 2783–2812. URL: <https://doi.org/10.1103/physrevd.52.2783> (cit. on p. 27).
- [18] F. U. Bernlochner, Z. Ligeti, and D. J. Robinson. “Model-independent analysis of semileptonic B decays to D^{**} for arbitrary new physics”. In: *Physical Review D* 97.7 (Apr. 2018). URL: <https://doi.org/10.1103/physrevd.97.075011> (cit. on pp. 27, 28, 68, 176, 177).
- [19] M. P. Altarelli and F. Teubert. “B Physics at LHCb”. In: *International Journal of Modern Physics A* 23.32 (Dec. 2008), pp. 5117–5136. URL: <https://doi.org/10.1142/s0217751x08042791> (cit. on p. 30).
- [20] I. Belyaev et al. “The history of LHCb”. In: *The European Physical Journal H* 46.1 (Mar. 2021). URL: <https://doi.org/10.1140/epjh/s13129-021-00002-z> (cit. on pp. 30, 34, 35).
- [21] The LHCb Collaboration. *$b\bar{b}$ production angle plots*. URL: https://lhcb.web.cern.ch/speakersbureau/html/bb_ProductionAngles_130424.html (cit. on p. 31).
- [22] D. Jacquet and F. Follin. *Implementation and experience with luminosity levelling with offset beam*. en. 2014. URL: <http://cds.cern.ch/record/1955354> (cit. on p. 30).
- [23] The LHCb Collaboration. *Standard set of performance numbers*. URL: <https://lhcb.web.cern.ch/speakersbureau/html/PerformanceNumbers.html> (cit. on p. 30).
- [24] A. A. A. Jr et al. “The LHCb Detector at the LHC”. In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08005. URL: <https://dx.doi.org/10.1088/1748-0221/3/08/S08005> (cit. on p. 32).
- [25] R. Aaij et al. “Performance of the LHCb Vertex Locator”. In: *Journal of Instrumentation* 9.09 (Sept. 2014), P09007–P09007. URL: <https://doi.org/10.1088/1748-0221/9/09/p09007> (cit. on p. 32).
- [26] R. Arink et al. “Performance of the LHCb Outer Tracker”. In: *Journal of Instrumentation* 9.01 (Jan. 2014), P01002–P01002. URL: <https://doi.org/10.1088/1748-0221/9/01/p01002> (cit. on p. 32).

- [27] R. Aaij et al. “Measurement of the track reconstruction efficiency at LHCb”. In: *Journal of Instrumentation* 10.02 (Feb. 2015), P02007–P02007. URL: <https://doi.org/10.1088/1748-0221/10/02/p02007> (cit. on pp. 33, 70).
- [28] C. Lippmann. “Particle identification”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 666 (Feb. 2012), pp. 148–172. URL: <https://doi.org/10.1016/j.nima.2011.03.009> (cit. on pp. 34, 37).
- [29] F. Archilli et al. “Performance of the Muon Identification at LHCb”. In: *JINST* 8 (2013). Comments: 17 pages, 10 figures, P10020. arXiv: 1306.0249. URL: <https://cds.cern.ch/record/1553139> (cit. on p. 36).
- [30] A. Seuthe. “PID performance in Run 2 at LHCb”. In: *PoS LHCP2020* (2021). Ed. by B. Mansoulie et al., p. 047. URL: <https://doi.org/10.22323/1.382.0047> (cit. on p. 36).
- [31] F. Dordei. “LHCb detector and trigger performance in Run II”. In: *EPJ Web Conf.* 164 (2017). Ed. by L. Bravina, Y. Foka, and S. Kabana, p. 01016. URL: <https://doi.org/10.1051/epjconf/201716401016> (cit. on p. 37).
- [32] S. Oggero. “Beauty in the Crowd: Commissioning of the LHCb Pile-Up detector and First evidence of $B_s^0 \rightarrow \mu^+ \mu^-$ ”. Presented 03 Oct 2013. 2013. URL: <https://cds.cern.ch/record/1635658> (cit. on p. 38).
- [33] P. d’Argent et al. “Improved performance of the LHCb Outer Tracker in LHC Run 2”. In: *Journal of Instrumentation* 12.11 (Nov. 2017), P11016–P11016. URL: <https://doi.org/10.1088/1748-0221/12/11/p11016> (cit. on p. 38).
- [34] R. Aaij et al. “Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC”. In: *Journal of Instrumentation* 14.04 (Apr. 2019), P04013–P04013. URL: <https://doi.org/10.1088/1748-0221/14/04/p04013> (cit. on pp. 38, 41, 57, 58).
- [35] M. Calvi et al. “Emulation of the hadronic trigger for tracker-only simulation”. In: LHCb-INT-2019-025 (2019). URL: <https://cds.cern.ch/record/2703802> (cit. on pp. 42, 56, 57, 60–62).
- [36] 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. San Francisco, California, USA: ACM, 2016, pp. 785–794. ISBN: 978-1-4503-4232-2. URL: <http://doi.acm.org/10.1145/2939672.2939785> (cit. on p. 57).
- [37] S. Tolk et al. “Data driven trigger efficiency determination at LHCb”. In: LHCb-PUB-2014-039 (2014). URL: <https://cds.cern.ch/record/1701134> (cit. on pp. 59, 60, 193).
- [38] M. Adinolfi et al. “Performance of the LHCb RICH detector at the LHC”. In: *The European Physical Journal C* 73.5 (May 2013). URL: <https://doi.org/10.1140/epjc/s10052-013-2431-9> (cit. on p. 64).

- [39] F. U. Bernlochner et al. “Das ist der HAMMER: consistent new physics interpretations of semileptonic decays”. In: *The European Physical Journal C* 80.9 (Sept. 2020). URL: <https://doi.org/10.1140/epjc/s10052-020-8304-0> (cit. on p. 66).
- [40] M. De Cian et al. “Measurement of the track finding efficiency”. In: LHCb-PUB-2011-025 (2012). URL: <https://cds.cern.ch/record/1402577> (cit. on p. 70).
- [41] M. Pivk and F. L. Diberder. “*sPlot*: A statistical tool to unfold data distributions”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 555.1-2 (Dec. 2005), pp. 356–369. URL: <https://doi.org/10.1016/j.nima.2005.08.106> (cit. on p. 71).
- [42] J. Wimberley, B. K. Hamilton, and H. Jawahery. “A measurement of the semitauonic branching fraction ratio $\mathcal{R}(J/\psi)$ ”. In: LHCb-ANA-2016-059 (July 2016). URL: <https://cds.cern.ch/record/2202625> (cit. on p. 93).
- [43] G. D’Agostini. “A multidimensional unfolding method based on Bayes’ theorem”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 362.2 (1995), pp. 487–498. ISSN: 0168-9002. URL: <https://www.sciencedirect.com/science/article/pii/016890029500274X> (cit. on pp. 94, 129).
- [44] The LHCb Collaboration. “Measurement of b -hadron fractions in 13 TeV pp collisions”. In: *Phys. Rev. D* 100.3 (2019), p. 031102. arXiv: 1902.06794. URL: <https://cds.cern.ch/record/2660233> (cit. on p. 109).
- [45] K. Cranmer et al. *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*. Tech. rep. New York: New York U., 2012. URL: <https://cds.cern.ch/record/1456844> (cit. on p. 112).
- [46] S. Calí et al. “Impacts of radiative corrections on measurements of lepton flavour universality in $B \rightarrow D\ell\nu_\ell$ decays”. In: *The European Physical Journal C* 79.9 (Sept. 2019). URL: <https://doi.org/10.1140/epjc/s10052-019-7254-x> (cit. on p. 130).
- [47] The LHCb experiments committe. “Letter of Intent for the LHCb Upgrade”. In: (2011). URL: <https://cds.cern.ch/record/1333091> (cit. on pp. 136, 137).
- [48] J. Albrecht et al. “The upgrade of the LHCb trigger system”. In: *Journal of Instrumentation* 9.10 (Oct. 2014), pp. C10026–C10026. URL: <https://doi.org/10.1088/1748-0221/9/10/c10026> (cit. on pp. 135, 138, 139).
- [49] C. D’Ambrosio. “The future of RICH detectors through the light of the LHCb RICH”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 876 (Dec. 2017), pp. 194–197. URL: <https://doi.org/10.1016/j.nima.2017.02.076> (cit. on pp. 136, 144).
- [50] A. Piucci. “The LHCb Upgrade”. In: *Journal of Physics: Conference Series* 878.1 (July 2017), p. 012012. URL: <https://dx.doi.org/10.1088/1742-6596/878/1/012012> (cit. on pp. 136, 141, 143).
- [51] The LHCb Collaboration. *LHCb Trigger and Online Upgrade Technical Design Report*. Tech. rep. 2014. URL: <https://cds.cern.ch/record/1701361> (cit. on p. 138).

- [52] R. Aaij et al. “Allen: A High-Level Trigger on GPUs for LHCb”. In: *Computing and Software for Big Science* 4.1 (Apr. 2020). URL: <https://doi.org/10.1007/s41781-020-00039-7> (cit. on pp. 138–140).
- [53] C. Fitzpatrick and V. V. Gligorov. *Anatomy of an upgrade event in the upgrade era, and implications for the LHCb trigger*. Tech. rep. Geneva: CERN, 2014. URL: <https://cds.cern.ch/record/1670985> (cit. on p. 138).
- [54] The LHCb Collaboration. “Performance of the GPU HLT1 (Allen)”. In: (2020). URL: <https://cds.cern.ch/record/2722327> (cit. on p. 141).
- [55] K. Hennessy. “LHCb VELO upgrade”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 845 (Feb. 2017), pp. 97–100. URL: <https://doi.org/10.1016/j.nima.2016.04.077> (cit. on p. 141).
- [56] J. Wang. “LHCb upstream tracker upgrade”. In: *PoS VERTEX2015* (2015), p. 012. URL: <https://inspirehep.net/literature/1413629> (cit. on pp. 141, 143, 149).
- [57] The LHCb Collaboration. *LHCb Tracker Upgrade Technical Design Report*. Tech. rep. 2014. URL: <https://cds.cern.ch/record/1647400> (cit. on p. 143).
- [58] A. Massafferri. “SciFi, the new Tracker of the LHCb experiment”. In: *Journal of Instrumentation* 15.08 (Aug. 2020), p. C08006. URL: <https://dx.doi.org/10.1088/1748-0221/15/08/C08006> (cit. on pp. 143, 144).
- [59] D. A. Berninghoff. “Commissioning of the Front-End Electronics of the LHCb Scintillating Fibre Tracker”. Presented 25 May 2022. 2022. URL: <https://cds.cern.ch/record/2810671> (cit. on p. 144).
- [60] S. Okamura. “The LHCb RICH detector upgrade: Improvements and expected performance”. In: *Nuovo Cimento C* 45.5 (2022), p. 118. URL: <https://cds.cern.ch/record/2837863> (cit. on p. 144).
- [61] J. E. Andrews. “Study of the decay $B^+ \rightarrow K^+\pi^0$ at LHCb and mechanical development for the design of the Upstream Tracker”. Maryland U., 2018 (cit. on pp. 147, 150, 151).
- [62] I. Carli. “Upstream Tracker - The silicon strip tracking detector for the LHCb Upgrade”. In: *PoS ICHEP2020* (2021), p. 724. URL: <https://cds.cern.ch/record/2783293> (cit. on p. 149).
- [63] C. Abellan Beteta et al. “The SALT Readout ASIC for Silicon Strip Sensors of Upstream Tracker in the Upgraded LHCb Experiment”. In: *Sensors* 22.1 (2022). ISSN: 1424-8220. URL: <https://www.mdpi.com/1424-8220/22/1/107> (cit. on pp. 150, 158).
- [64] P. Moreira et al. “The GBT-SerDes ASIC prototype”. In: *Journal of Instrumentation* 5.11 (Nov. 2010), p. C11022. URL: <https://dx.doi.org/10.1088/1748-0221/5/11/C11022> (cit. on p. 152).
- [65] GBT project. *GBTX manual*. Tech. rep. 2018. URL: <https://readthedocs.web.cern.ch/ate/files/113213790/113213791/1/1556910129000/gbtxManual.pdf> (cit. on pp. 152, 158).

- [66] GBT project. *GBLD user manual*. Tech. rep. 2015. URL: https://readthedocs.web.cern.ch/ate/files/113213790/113213794/1/1556903336000/GBLD_manual-June2015.pdf (cit. on p. 152).
- [67] GBT project. *GBT-SCA user manual*. Tech. rep. 2017. URL: https://readthedocs.web.cern.ch/ate/files/113213790/181732620/1/1611924370000/GBT-SCA-UserManual_v8.2.pdf (cit. on p. 153).
- [68] M. F. Sevilla and P. Hamilton. *LVR user's manual*. Tech. rep. 2020. URL: http://flavor.physics.umd.edu/research/ut/lvr_manual.pdf (cit. on pp. 155, 156).
- [69] UMD LHCb group. *Voltage lookup table for LVR*. 2022. URL: https://umd-lhcb.github.io/ut-hw-doc/cern/cern_qa/cern_qa/#proposed-lvr-nominal-voltages (cit. on p. 154).
- [70] K. Wyllie et al. “Electronics Architecture of the LHCb Upgrade”. In: (Mar. 2011) (cit. on p. 156).
- [71] W. Parker and S. Braun. *DCB Operation Manual*. Tech. rep. University of Maryland, 2022. URL: https://umd-lhcb.github.io/research/ut/dcb_operation_manual.pdf (cit. on p. 156).
- [72] T. Colombo et al. “The LHCb Online system in 2020: trigger-free read-out with (almost exclusively) off-the-shelf hardware”. In: *Journal of Physics: Conference Series* 1085.3 (Sept. 2018), p. 032041. URL: <https://dx.doi.org/10.1088/1742-6596/1085/3/032041> (cit. on p. 163).
- [73] F. Alessio and R. Jacobsson. *System-level Specifications of the Timing and Fast Control system for the LHCb Upgrade*. Tech. rep. Geneva: CERN, 2014. URL: <https://cds.cern.ch/record/1424363> (cit. on p. 164).
- [74] L. Granado Cardoso et al. “LHCb MiniDAQ control system”. In: *EPJ Web Conf.* 214 (2019), p. 01005. URL: <https://cds.cern.ch/record/2702137> (cit. on pp. 165, 168).
- [75] C. Gaspar, P. Charpentier, and M. Dönszelmann. “DIM, a portable, light weight package for information publishing, data transfer and inter-process communication”. In: *Comput. Phys. Commun.* 140.1-2 (2001), pp. 102–9. URL: <https://cds.cern.ch/record/559279> (cit. on p. 165).
- [76] N. Neufeld, R. Stoica, and J. M. Caicedo. *pydim3*. Version 3.0.3. 2021. URL: <https://pypi.org/project/pydim3/> (cit. on p. 172).
- [77] Y. Sun. *nanoDAQ*. Version 0.3.1. 2020. URL: <https://github.com/umd-lhcb/nanoDAQ> (cit. on p. 172).
- [78] G. Cowan. *Statistical Data Analysis*. Oxford University Press, 1998 (cit. on p. 195).
- [79] The ATLAS Collaboration. *Statistical methods for the LHC*. 2022. URL: https://statisticalmethods.web.cern.ch/StatisticalMethods/unfolding/RooUnfold_01-Methods/ (cit. on p. 195).
- [80] L. Brenner et al. “Comparison of unfolding methods using RooFitUnfold”. In: *International Journal of Modern Physics A* 35.24 (2020), p. 2050145. URL: <https://doi.org/10.1142/S0217751X20501456> (cit. on p. 195).