

UNIVERSITY OF LJUBLJANA
FACULTY OF MATHEMATICS AND PHYSICS
DEPARTMENT OF PHYSICS

Andrej Lozar

**MEASURING BRANCHING FRACTION OF
 $D^+ \rightarrow \pi^+ \ell^+ \ell^-$ RARE DECAY WITH BELLE II
DETECTOR**

DOCTORAL THESIS

ADVISER: Prof. Dr. Rok Pestotnik

Ljubljana, 2026

UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO
ODDELEK ZA FIZIKO

Andrej Lozar

**MERJENJE RAZVEJITVENEGA RAZMERJA
REDKEGA RAZPADA $D^+ \rightarrow \pi^+ \ell^+ \ell^-$ Z
DETEKTORJEM BELLE II**

DOKTORSKA DISERTACIJA

MENTOR: prof. dr. Rok Pestotnik

Ljubljana, 2026

Acknowledgements

Iskreno se zahvaljujem svojemu mentorju prof. dr. Roku Pestotniku za strokovno vodenje, nasvete in podporo tekom doktorskega študija.

Zahvala gre tudi dr. Luki Šantlju za strokovne razprave, nasvete in usmeritve pri raziskovalnem delu.

Hvala tudi sodelavcem, ki ste poskrbeli za prijetno vzdušje in za izmenjavo idej, ki so pomembno prispevale k nastanku tega dela.

Iskrena hvala moji družini in vsem bližnjim, ki so me v vseh letih študija in raziskovanja podpirali, spodbujali in verjeli vame.

Največja zahvala pa gre Ani, moji ženi, ki mi vedno stoji ob strani, tako pri študijskih kot tudi pri drugih življenjskih izzivih, ter mi daje občutek, da je vse mogoče.

MERJENJE RAZVEJITVENEGA RAZMERJA REDKEGA RAZPADA $D^+ \rightarrow \pi^+ \ell^+ \ell^-$ Z DETEKTORJEM BELLE II

IZVLEČEK

V delu predstavljamo iskanje redkega razpada čarobnega mezona $D^+ \rightarrow \pi^+ \ell^+ \ell^-$ v elektronskem kanalu. Meritev smo opravili na vzorcu podatkov, zajetih z eksperimentoma Belle in Belle II, ki ustreza integriranima luminoznostma 980 fb^{-1} in 428 fb^{-1} . Podatke smo zbrali z asimetričnima trkalnikoma delcev e^+e^- KEKB in SuperKEKB na Japonskem. Ker signala nismo opazili, smo izpeljali zgornje meje razvejitvenega razmerja tega razpada, ki so bile tako prvič predstavljene s strani kolaboracij Belle in Belle II. Pri 90% stopnji zaupanja znaša zgornja meja $\mathcal{B}(D^+ \rightarrow \pi^+ e^+ e^-) < 4,7 \times 10^{-7}$, s čimer smo izboljšali prejšnjo mejo eksperimenta BABAR za faktor 2,3. Meje smo prvič določili tudi v treh intervalih invariantne mase elektronskega para. Za validacijo analize smo izmerili razvejitveno razmerje razpada $D^+ \rightarrow \pi^+ \phi(\rightarrow e^+ e^-)$, ki znaša $\mathcal{B} = (1,75 \pm 0,22 \text{ (stat.)} \pm 0,12 \text{ (sist.)}) \times 10^{-6}$. Izmerili smo ga s statistično značilnostjo 12,0 standardnih odklonov in tako natančnost prejšnje meritve kolaboracije CLEO izboljšali za približno faktor pet.

Ključne besede: Redki razpadi kvarka c , FCNC, Belle II, Belle, $D^+ \rightarrow \pi^+ e^+ e^-$, Razvejitveno razmerje, Zgornja meja

MEASURING BRANCHING FRACTION OF $D^+ \rightarrow \pi^+ \ell^+ \ell^-$ RARE DECAY WITH BELLE II DETECTOR

ABSTRACT

This thesis presents a search for the rare charm-meson decay $D^+ \rightarrow \pi^+ \ell^+ \ell^-$ in the dielectron channel. The measurement was performed on data samples recorded by the Belle and Belle II experiments, corresponding to integrated luminosities of 980 fb^{-1} and 428 fb^{-1} . The data were collected with the asymmetric-energy e^+e^- colliders KEKB and SuperKEKB in Japan. As no signal was observed, we derive upper limits on the branching fraction of this decay, the first such limits reported by the Belle and Belle II collaborations. At the 90% confidence level the limit is $\mathcal{B}(D^+ \rightarrow \pi^+ e^+ e^-) < 4.7 \times 10^{-7}$, improving on the previous limit from the *BABAR* experiment by a factor of 2.3. We also determine limits for the first time in three intervals of the dielectron invariant mass. To validate the analysis, we measure the branching fraction of the decay $D^+ \rightarrow \pi^+ \phi(\rightarrow e^+ e^-)$, obtaining $\mathcal{B} = (1.75 \pm 0.22 \text{ (stat)} \pm 0.12 \text{ (syst)}) \times 10^{-6}$. It is measured with a statistical significance of 12.0 standard deviations, improving the precision of the previous CLEO measurement by about a factor of five.

Keywords: Rare charm decays, FCNC, Belle II, Belle, $D^+ \rightarrow \pi^+ e^+ e^-$, Branching fraction, Upper limit

Contents

List of figures	15
List of tables	19
1 Introduction	21
2 Theoretical and Experimental Context	25
2.1 The Standard Model	25
2.2 Electroweak Interaction	26
2.3 Flavour-Changing Neutral Currents and Rare Decays	28
2.4 Effective Field Theory Description	30
2.4.1 Effective Hamiltonian	30
2.4.2 SM Prediction	32
2.5 Sensitivity to New Physics	34
2.6 Experimental Status	36
2.7 Measurement Strategy	37
3 The Belle and Belle II Experiments	39
3.1 The KEKB and SuperKEKB Accelerators	39
3.2 The Belle II Detector	42
3.2.1 Coordinate System	44
3.2.2 Tracking System	44
3.2.3 Cherenkov Ring Imaging Counters	45
3.2.4 Electromagnetic Calorimeter	47
3.2.5 Muon and K_L^0 Identification (KLM)	48
3.2.6 Trigger and Data Acquisition	49
3.3 The Belle Detector	50
3.4 Event Reconstruction	51
3.4.1 The Belle II Analysis Software Framework	51
3.4.2 Track Reconstruction	52
3.4.3 Charged Particle Identification	52
3.4.4 Charged Particle Identification in Belle	54
4 Data Samples	55
4.1 Experimental Data Samples	55
4.1.1 Belle II Dataset	55
4.1.2 Belle Dataset	56
4.2 Monte Carlo Simulation	57
4.2.1 Event Generators and Detector Simulation	58
4.2.2 Generic MC Samples	58

4.2.3	Signal, Normalisation, and Validation MC Samples	58
5	Event Selection and Signal Reconstruction	63
5.1	Track Selection and Particle Identification	63
5.2	D^+ Candidate Reconstruction	64
5.2.1	Normalisation Channel Reconstruction	64
5.3	Bremsstrahlung Correction	65
5.4	Loose Selection	65
5.5	$M(e^+e^-)$ Selection and Photon-Conversion Veto	67
5.5.1	Photon-Conversion Veto	67
5.5.2	Low-Mass Dielectron Requirement	71
5.5.3	Resonance Exclusion Windows	71
5.6	Best-Candidate Selection	71
5.7	Individual Selection Efficiency	72
5.8	Random Forest Classifier	72
5.8.1	Training	73
5.8.2	Performance	75
5.8.3	RFC Performance in $M(e^+e^-)$ Intervals	78
5.9	η -Veto	79
5.10	Cut Optimisation	81
5.10.1	Significance Parameter	82
5.10.2	RFC Working Point	82
5.10.3	Other Selection Variables	82
5.11	$M(e^+e^-)$ Interval Definition	83
5.12	Summary	84
6	Reconstruction Efficiency and MC/Data Comparison	87
6.1	Reconstruction Efficiency	87
6.1.1	Normalisation Channel Reconstruction Efficiency	87
6.1.2	Signal and Validation Channel Reconstruction Efficiency	89
6.2	Data/MC Validation	90
6.2.1	Validation Mode	91
6.2.2	Normalisation Mode	92
6.2.3	Sidebands	93
6.2.4	Signal Mode	93
6.3	Summary	94
7	Fit Strategy and Validation	97
7.1	Branching Fraction Determination	97
7.2	Likelihood Construction	98
7.3	PDF Functional Forms	98
7.4	Fits to Signal MC	100
7.4.1	Signal Mode	100
7.4.2	Validation Mode	101
7.4.3	Normalisation Mode Cross-Check	103
7.5	Simultaneous Fit Framework	103
7.6	Signal Significance and Limit-Setting Procedure	104
7.7	Closure Tests on Generic MC	105
7.7.1	Signal Mode	105

7.7.2	Validation Mode	106
7.7.3	Normalisation Mode	108
7.8	Toy MC Studies and Linearity Tests	109
7.8.1	Signal Mode	110
7.8.2	Signal Mode in $M(e^+e^-)$ Intervals	112
7.8.3	Validation Mode	114
7.8.4	Normalisation Mode	116
7.8.5	Normalisation Mode: PDF Mismodelling	117
8	Systematic Uncertainties	121
8.1	Classification and Propagation	122
8.2	Efficiency-Related Uncertainties	123
8.2.1	Particle Identification	123
8.2.2	Random-Forest-Classifer Efficiency	125
8.2.3	Tracking and Reconstruction Efficiency	126
8.2.4	Mis-reconstructed Signal	127
8.2.5	Signal Decay Model	127
8.3	Normalisation-Yield Uncertainties	128
8.4	Fit-Related Uncertainties	128
8.4.1	Fixed Signal-PDF Parameters	128
8.4.2	Background-Model Choice	129
8.4.3	Signal-PDF Differences Between Data and Simulation	129
8.5	Summary	131
9	Results	135
9.1	Normalisation Mode	135
9.2	Validation Mode	135
9.3	Signal Mode	137
9.4	Upper Limits	138
9.4.1	Independent Fits	141
9.5	Comparison With Previous Measurements	143
10	Summary and Outlook	145
10.1	Summary of the Measurement	145
10.2	Outlook	146
10.2.1	Improvements to the Present Analysis	146
10.2.2	Extensions of the Analysis	147
10.2.3	Prospects With Larger Data Sets	148
10.3	Concluding Remarks	149
11	Bibliography	151
	Appendix A Standard Model Prediction	157
	Appendix B Supplementary Selection and Data/MC Studies	159
B.1	Distributions of Variables After Loose Selection	159
B.2	RFC Input Variables in $M(e^+e^-)$ Intervals	162
B.3	RFC Performance in the Validation and Normalisation Modes	167
B.4	Data/MC Comparisons	169

B.4.1	Validation Mode	169
B.4.2	Normalisation Mode	171
B.4.3	Sideband Regions	173
B.4.4	Signal Mode	175
Appendix C	Supplementary Toy-MC Studies	177
C.1	Independent Fits: Signal Mode	177
C.2	Simultaneous Fit in $M(e^+e^-)$ Intervals	178
C.3	Independent Fits: Validation Mode	180
C.4	Normalisation Mode	181
Appendix D	Supplementary Systematic and Result Studies	183
D.1	Normalisation-Mode Fits Without the RFC Requirement	183
D.2	Broken-Signal Component in $M(e^+e^-)$ Intervals	184
D.3	Linearity With Signal-PDF Variations	185
D.4	Background-Model Choice	186
D.5	Validation Mode: Confidence Belt Without Systematic Uncertainties	187
D.6	Validation Mode: Independent Fits	187
Razširjeni povzetek v slovenskem jeziku		189
Uvod in teoretično ozadje		189
Eksperimentalno stanje in strategija meritve		190
Eksperimenta Belle in Belle II		191
Trkalnik SuperKEKB		191
Spektrometer Belle II		192
Spektrometer Belle		192
Simulacija in programsko okolje		192
Izbor dogodkov in rekonstrukcija signala		193
Določitev razvejitvenega razmerja in zgornjih mej		195
Sistematične negotovosti		196
Rezultati		198
Zaključek		201

List of figures

2.1	The elementary particles of the Standard Model	26
2.2	Feynman diagrams for the $c \rightarrow u\ell^+\ell^-$	29
2.3	SM Wilson coefficients C_7^{eff} and C_9^{eff} as a functions of q^2	32
2.4	SM differential branching fraction of $D^+ \rightarrow \pi^+\mu^+\mu^-$	33
3.1	Schematic view of SuperKEKB	40
3.2	KEKB and SuperKEKB collision schemes	42
3.3	Schematic view of Belle and Belle II detectors	43
3.4	Schematic view of VXD	44
3.5	CDC wire layout	45
3.6	TOP schematic view	46
3.7	ARICH detector layout	47
3.8	Layout of ECL detector	48
4.1	Dalitz and $M(e^+e^-)$ distributions of generated signal MC	61
5.1	Bremsstrahlung correction	66
5.2	Selection-variable distributions after loose selection, Belle	68
5.3	Selection-variable distributions after loose selection, Belle II	69
5.4	$M(\gamma \rightarrow e^+e^-)$, $M(e^+e^-)$ and Punzi FOM for photon veto	70
5.5	Photon-veto variables by conversion multiplicity	70
5.6	Ranking of D^+ candidates by vertex-fit probability	72
5.7	RFC learning curves, Belle and Belle II	74
5.8	RFC input-variable distributions, Belle	74
5.9	RFC input-variable distributions, Belle II	75
5.10	RFC score distributions and Punzi FOM	76
5.11	RFC ROC curve and performance vs. score cut, Belle	77
5.12	RFC ROC curve and performance vs. score cut, Belle II	77
5.13	RFC feature importance, Belle	77
5.14	RFC feature importance, Belle II	78
5.15	RFC performance across $M(e^+e^-)$ intervals	78
5.16	D^+ candidate composition after RFC selection, Belle	80
5.17	D^+ candidate composition after RFC selection, Belle II	80
5.18	$M(\pi^+e^+e^-)$ and $M(e^+e^-)$ for dedicated η -background MC	81
5.19	Punzi FOM vs. signal-window width	82
5.20	Pion-identification variable distribution	83
6.1	Signal efficiency vs. $M(e^+e^-)$	90
6.2	Data/MC comparison, $M(\pi^+e^+e^-)$, validation channel	91
6.3	Data/MC comparison, $M(\pi^+K^+K^-)$, normalisation channel	92

6.4	Data/MC comparison, $M(\pi^+e^+e^-)$ sidebands, signal mode	93
6.5	Data/MC comparison, $M(e^+e^-)$ sidebands, signal mode	93
6.6	Data/MC comparison, $M(\pi^+e^+e^-)$, signal mode	94
6.7	Data/MC comparison, $M(e^+e^-)$, signal mode	95
7.1	Fit to signal MC, signal mode, full $M(e^+e^-)$ range	100
7.2	Broken-signal $M(\pi^+e^+e^-)$ distribution, signal mode	101
7.3	Fits to signal MC by $M(e^+e^-)$ interval, signal mode	102
7.4	Fit to signal MC, validation mode	103
7.5	Fit to truth-matched signal in generic MC, normalisation mode . . .	103
7.6	Simultaneous fit to generic MC, signal mode	106
7.8	Fit to generic MC, validation mode	106
7.7	Simultaneous fits to generic MC by $M(e^+e^-)$ interval	107
7.9	Simultaneous fit to generic MC, validation mode	108
7.10	Fit to generic MC, normalisation mode	108
7.11	Fit to Belle II generic MC, matched to data size, normalisation mode	109
7.12	Broken-signal $M(\pi^+K^+K^-)$ distribution, normalisation mode	109
7.13	Toy MC pull distributions, signal mode	111
7.14	Linearity test, simultaneous fit, signal mode	112
7.15	Toy MC pull distributions, signal mode, I_1	113
7.16	Linearity test, simultaneous fit, signal mode, I_1	114
7.17	Toy MC pull distributions, validation mode	115
7.18	Linearity test, simultaneous fit, validation mode	116
7.19	Linearity test, normalisation mode	117
7.20	Relative deviation of fitted from generated yield, normalisation mode	119
8.1	PID correction factors for pions, Belle II	124
8.2	Example background-model variation fit, Belle II	130
9.1	Fit to $M(\pi^+K^+K^-)$, normalisation mode, data	136
9.2	Simultaneous fit, validation mode, data	137
9.3	Neyman confidence belt, validation mode	138
9.4	Simultaneous fits, signal mode, by $M(e^+e^-)$ interval	139
9.5	Neyman confidence belt, signal mode, full range	140
9.6	Neyman confidence belts, signal mode, by interval	141
9.7	Independent fits, signal mode, data	142
9.8	Neyman confidence belts, independent fits, signal mode	143
B.1	Additional loose-selection distributions, Belle	160
B.2	Additional loose-selection distributions, Belle II	161
B.3	RFC input variables, Belle, I_1	162
B.4	RFC input variables, Belle, I_2	163
B.5	RFC input variables, Belle, I_3	164
B.6	RFC input variables, Belle II, I_1	165
B.7	RFC input variables, Belle II, I_2	166
B.8	RFC input variables, Belle II, I_3	167
B.9	RFC performance on validation mode, Belle and Belle II	168
B.10	RFC performance on normalisation mode, Belle and Belle II	168
B.11	Data/MC comparison, validation-mode variables, Belle	169

B.12	Data/MC comparison, validation-mode variables, Belle II	170
B.13	Data/MC comparison, normalisation-mode variables, Belle	171
B.14	Data/MC comparison, normalisation-mode variables, Belle II	172
B.15	Data/MC comparison, sideband variables, signal mode, Belle	173
B.16	Data/MC comparison, sideband variables, signal mode, Belle II	174
B.17	Data/MC comparison, signal-mode variables, Belle	175
B.18	Data/MC comparison, signal-mode variables, Belle II	176
C.1	Toy MC pulls, independent fits, signal mode	177
C.2	Linearity, independent fits, signal mode	178
C.3	Toy MC pulls, simultaneous fit, signal mode, I_2	178
C.4	Toy MC pulls, simultaneous fit, signal mode, I_3	179
C.5	Linearity, simultaneous fit, signal mode, I_2/I_3	179
C.6	Toy MC pulls, independent fits, validation mode	180
C.7	Linearity, independent fits, validation mode	180
C.8	Toy MC pulls, normalisation mode, Belle	181
C.9	Toy MC pulls, normalisation mode, Belle II	182
D.1	Fits without RFC requirement, normalisation mode, Belle	183
D.2	Fits without RFC requirement, normalisation mode, Belle II	183
D.3	Broken-signal fits in $M(e^+e^-)$ intervals	184
D.4	Linearity with varied signal-PDF parameters, signal mode	185
D.5	Linearity with varied signal-PDF parameters, validation mode	185
D.6	Remaining background-model variations	186
D.7	Neyman confidence belt, validation mode, no systematics	187
D.8	Independent fits, validation mode, data	187
1	Diferencialno razvejitveno razmerje SM za $D^+ \rightarrow \pi^+ \mu^+ \mu^-$	190
2	Sočasna prilagajanja signalnega kanala	200

List of tables

2.1	Electroweak quantum numbers of first-generation fermions	27
2.2	Summary of previous experimental results	36
3.1	Production cross sections	41
3.2	Machine parameters for KEKB and SuperKEKB	41
4.1	Summary of the Belle signal MC samples	60
4.2	Summary of the Belle II signal MC samples	60
5.1	PID efficiencies and fake rates at tight working point	64
5.2	Retention rate per initial-selection cut, Belle	72
5.3	Retention rate per initial-selection cut, Belle II	73
5.4	RFC signal efficiency and background retention	75
5.5	RFC performance at matched signal efficiency	76
5.6	RFC performance in $M(e^+e^-)$ intervals	78
5.7	RFC performance at matched signal efficiency by interval, Belle . . .	79
5.8	RFC performance at matched signal efficiency by interval, Belle II . .	79
5.9	Effect of the η -veto on signal and background reduction	81
5.10	Summary of signal-mode selection criteria	84
5.11	Summary of normalisation-mode selection criteria	85
6.1	Normalisation-channel reconstruction efficiencies	88
6.2	Inclusive signal and validation channel efficiencies	89
6.3	Signal efficiency by $M(e^+e^-)$ interval	90
6.4	RFC performance, validation mode	91
6.5	RFC performance, normalisation mode	92
6.6	Observed data and generic MC yields, signal mode	94
7.1	Signal and background PDFs used in each fit	100
7.2	Broken-signal fractions from signal MC	102
7.3	Signal yields in generic MC, validation mode	106
7.4	Signal yields in generic MC, normalisation mode	110
7.5	Toy MC pull means and widths, signal mode	111
7.6	Linearity-test slope and intercept, signal mode	112
7.7	Toy MC pull means and widths, signal mode, I_1	113
7.8	Linearity-test by $M(e^+e^-)$ interval	114
7.9	Toy MC pull means and widths, validation mode	115
7.10	Linearity-test slope and intercept, validation mode	116
7.11	Toy MC pull means and widths, normalisation mode, Belle	117
7.12	Toy MC pull means and widths, normalisation mode, Belle II	118

8.1	PID correction factors, signal mode, Belle	125
8.2	PID correction factors, signal mode, Belle II	125
8.3	PID correction factors, normalisation and validation modes	126
8.4	Additive systematic from background-model choice	129
8.5	Signal-PDF shape parameters, data vs. MC	131
8.6	Systematic uncertainties by $M(e^+e^-)$ interval	133
8.7	Systematic uncertainties common to all intervals	134
9.1	Efficiencies, yields and fitted $\mathcal{B}$ by interval	138
9.2	Upper limits on $\mathcal{B}(D^+ \rightarrow \pi^+ e^+ e^-)$	141
A.1	SM predictions for short-distance branching fractions	157
1	Izkoristki rekonstrukcije	195
2	Sistematične negotovosti po intervalih $M(e^+e^-)$	198
3	Sistematične negotovosti, skupne vsem intervalom $M(e^+e^-)$	199

1. Introduction

Particle physics is a field of physics devoted to the study of nature’s elementary building blocks and the interactions governing their behaviour. Currently, the Standard Model (SM) is considered the most complete theoretical framework for describing these interactions, with a wide range of experimental measurements showing excellent agreement with its predictions.

Despite its success, the SM is known to be incomplete and cannot satisfactorily explain certain phenomena, such as the asymmetry between matter and antimatter in the Universe, the nature of dark matter, the origin of neutrino masses, and the absence of a description of gravity. These shortcomings indicate that nature contains additional unexplored particles or interactions beyond those described by the SM. Proposed extensions of the SM are collectively referred to as physics beyond the Standard Model (BSM), or new physics (NP).

Experimental searches for NP commonly follow two complementary approaches. One strategy is to directly probe unexplored regions of phase space, for example, by accessing higher energies at the so-called *energy frontier*. Another approach relies on high-precision measurements of already known processes, enabled by large data samples collected with advanced detector technologies at the *luminosity frontier*. In the latter case, rare decays provide a powerful probe for NP, as even small deviations from the SM predictions may indicate the presence of NP. This thesis focuses on one such precision measurement: the search for the flavour-changing neutral-current (FCNC) decay mediated by the quark-level transition $c \rightarrow u\ell^+\ell^-$.

FCNC processes are highly suppressed in the SM due to the Glashow–Iliopoulos–Maiani (GIM) mechanism and can occur only through higher-order loop diagrams [1]. As a result, their SM-predicted branching fractions are extremely small. In many BSM scenarios, such as supersymmetry or models with leptoquarks, FCNC processes can receive substantial contributions from new particles or interactions, leading to significantly enhanced decay rates compared to SM predictions. Such contributions can interfere with the SM amplitude, leading to measurable deviations in a variety of physics observables, such as the differential branching fraction, lepton-flavour universality ratio, forward–backward asymmetry, and angular distributions. Due to this sensitivity, FCNC transitions serve as powerful probes for NP. A confirmed statistically significant deviation from an SM prediction would provide indirect evidence of new particles or interactions, even if their masses lie far beyond the energies directly accessible in the collision.

In the charm sector, FCNC transitions are significantly more suppressed than in the beauty sector, where SM branching fractions are experimentally accessible and have already been studied. Experimentally, searches for charm FCNC are challenging due to limited signal yields, dominant long-distance contributions from hadronic resonances, and theoretical difficulties in describing non-perturbative ef-

fects [2]. Nevertheless, charm FCNC processes probe flavour structures that are complementary to those in strange and beauty-meson decays.

One of the most experimentally accessible channels to study the $c \rightarrow u\ell^+\ell^-$ transition is the exclusive decay $D^+ \rightarrow \pi^+\ell^+\ell^-$. Its rate is dominated by long-distance dynamics, in which the lepton pair is produced through an intermediate hadronic resonance rather than through the loop process of interest. Away from these resonances, however, the decay retains sensitivity to short-distance physics, and its simple topology and favourable kinematics make it a natural place to look for deviations from the SM.

Prior to this work, the decay had been searched for in both the electron and muon final states, with no evidence for a non-resonant signal in either. The most stringent constraint in the muon channel has been set by the LHCb collaboration at the level of 10^{-8} [3], while the *BABAR* experiment has obtained the best limit in the electron channel at the level of 10^{-6} [4]. These limits remain several orders of magnitude above the SM short-distance prediction, especially in the electron channel, providing an opportunity to further constrain contributions from BSM physics.

This thesis presents a search for the $D^+ \rightarrow \pi^+e^+e^-$ decay with the Belle II experiment, extended by the data set collected by its predecessor, Belle, with the two samples analysed in a single simultaneous fit. We analyse the electron final state, for which the pre-existing constraint is more than an order of magnitude weaker than for the muon channel. The latter is already constrained most strongly by LHCb. The data were collected at the KEKB and SuperKEKB asymmetric-energy e^+e^- colliders, running primarily at the $\Upsilon(4S)$ resonance and providing relatively clean collision environments with excellent particle identification. Analysing the two samples together gives access to a combined integrated luminosity of 1408 fb^{-1} , of which 980 fb^{-1} were collected by Belle and 428 fb^{-1} by Belle II, nearly four times the sample used in the most sensitive previous search in the electron channel [4].

The thesis is organised as follows.

Chapter 2 presents the theoretical framework relevant to this analysis: the Standard Model description of charm FCNC transitions, the effective-field-theory treatment of the short- and long-distance contributions, the resulting predictions for $D^+ \rightarrow \pi^+\ell^+\ell^-$, the sensitivity of this mode to new physics, and the current experimental status.

Chapter 3 describes the KEKB and SuperKEKB accelerators and the Belle and Belle II detectors, with emphasis on the subsystems relevant to the reconstruction of charged pions, kaons, and electrons.

Chapter 4 introduces the data samples and the simulation and software frameworks used to develop the analysis.

Chapter 5 describes the reconstruction of D^+ candidates and the selection criteria, including particle identification, the bremsstrahlung correction, the multivariate classifier used to suppress combinatorial background, and the definition of the $M(e^+e^-)$ intervals.

Chapter 6 presents the reconstruction efficiencies for the signal, validation, and normalisation channels, and the comparison between data and simulation.

Chapter 7 describes the signal-extraction procedure, based on a simultaneous unbinned maximum-likelihood fit to the Belle and Belle II samples, together with the studies performed to validate it.

Chapter 8 evaluates the systematic uncertainties affecting the measurement.

Chapter 9 reports the results: the branching fraction of the validation channel $D^+ \rightarrow \pi^+ \phi (\rightarrow e^+ e^-)$, the signal yields in each $M(e^+ e^-)$ interval, and the resulting upper limits on $\mathcal{B}(D^+ \rightarrow \pi^+ e^+ e^-)$.

Chapter 10 summarises the work and outlines the prospects for this measurement with the full Belle II data set.

Charge-conjugated processes are implied throughout the document unless specified otherwise.

2. Theoretical and Experimental Context

2.1 The Standard Model

The Standard Model [5, 6, 7] is the current theoretical framework that describes the fundamental particles of matter and the forces acting between them: the electromagnetic, weak, and strong interactions. It is formulated as a quantum field theory, where matter particles (fermions) are excitations of fermion fields, while their interactions are mediated by force-carrying particles (gauge bosons), which are excitations of the corresponding gauge fields. In addition, a scalar field (the Higgs field) is responsible for the mechanism of mass generation [8, 9].

Fermions constitute the building blocks of matter, while gauge bosons mediate their interactions. The dynamics of the gauge fields are governed by the local gauge symmetry:

$$SU(3)_C \times SU(2)_L \times U(1)_Y,$$

where $SU(3)_C$ is the special unitary group for the strong interaction acting on the colour charge and described by quantum chromodynamics (QCD). The product $SU(2)_L \times U(1)_Y$ describes the combination of the weak and electromagnetic interactions (electroweak interaction). Here, $SU(2)_L$ is the special unitary group acting on the left-handed weak isospin doublets, while $U(1)_Y$ corresponds to hypercharge.

The gauge bosons are spin-1 particles that mediate the fundamental forces. The eight massless gluons carry colour charge and mediate the strong interaction between quarks. The weak interaction is mediated by two charged massive bosons ($W^\pm$) and one neutral massive boson (Z^0). Electromagnetic interactions occur between particles carrying electric charge and are mediated by the neutral massless photon γ .

Fermions are spin- $\frac{1}{2}$ particles and come in two types: quarks and leptons. They are arranged in three generations of increasing mass as shown in Figure 2.1. Quarks are further divided into up-type (u, c, t) and down-type quarks (d, s, b), while the leptons consist of the charged leptons ($\ell = e, \mu, \tau$) and the neutral leptons, or neutrinos ($\nu_\ell = \nu_e, \nu_\mu, \nu_\tau$). The higher generations are identical copies of the first generation but with higher mass, and each fermion has a corresponding antiparticle with opposite quantum numbers. Quarks and charged leptons acquire mass through their interaction with the Higgs field, described by Yukawa couplings whose different strengths determine the observed mass hierarchy across the generations. Neutrinos are massless in the minimal SM, but the observation of neutrino oscillations shows that they are not, and constitutes one of the clearest experimental indications that the SM is incomplete. All fermions couple to the weak interaction through their weak

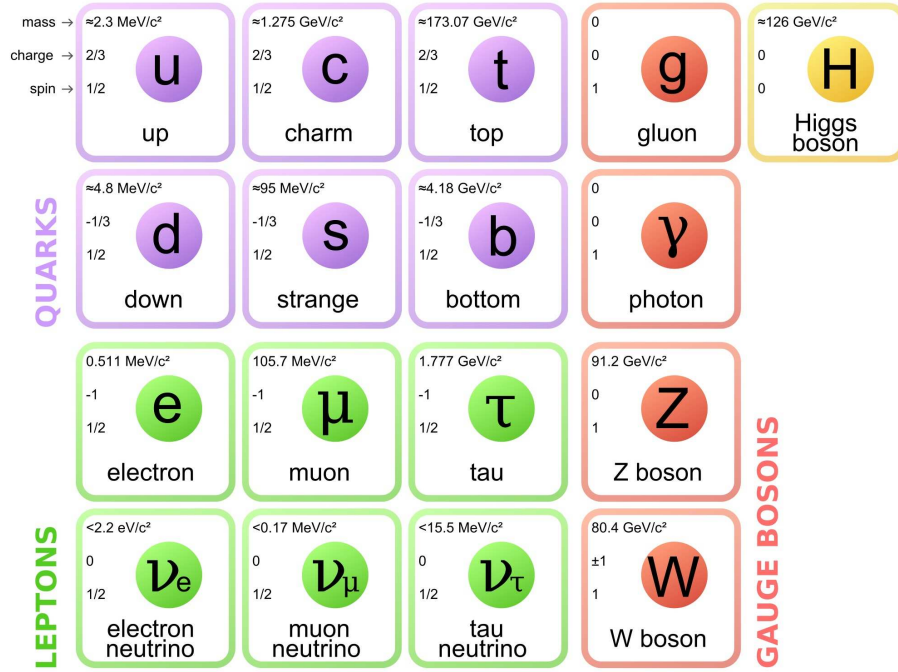


Figure 2.1: The elementary particles of the Standard Model [10].

isospin. Quarks and charged leptons additionally participate in electromagnetic interactions via their electric charge. Only the quarks carry a colour charge, making them the only fermions subject to the strong force. They form colour triplets under the $SU(3)_C$ group with three colour components: *red*, *green* and *blue*, while all leptons are colour singlets. In nature, quarks cannot exist as free particles due to colour confinement. They are always bound into colourless *hadrons*, such as *mesons* (made of a quark and an anti-quark) or *baryons* (three quarks, or three anti-quarks in the case of antibaryons). Particles of the higher generations are unstable and eventually decay into lighter particles. Ordinary matter is therefore composed predominantly of first-generation fermions.

2.2 Electroweak Interaction

The electroweak (EW) theory, developed by Glashow, Weinberg, and Salam [5, 6, 7], unifies the electromagnetic and weak interactions in a single gauge framework based on the group $SU(2)_L \times U(1)_Y$. The subscript L indicates that only left-handed fermions participated in the $SU(2)_L$ interaction, while right-handed fermions are $SU(2)_L$ singlets. The electromagnetic interaction, by contrast, couples to both chiralities.

Each fermion is assigned a weak hypercharge Y_W , related to its electric charge (Q) and the third component of the weak isospin (T_{W3}) by the Gell-Mann-Nishijima formula [11, 12], defined as

$$Y_W = 2 \times (Q - T_{W3}). \quad (2.1)$$

The left-handed fermions form weak isospin doublets, while the right-handed fermions are weak isospin singlets:

$$\begin{aligned} \text{leptons: } \ell_L &\equiv \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L & e_R, \\ \text{quarks: } q_L &\equiv \begin{pmatrix} u \\ d \end{pmatrix}_L & u_R \quad d_R. \end{aligned}$$

Quantum numbers for the first-generation fermions are summarised in Table 2.1, where subscripts L and R denote left- and right-handed particles.

Table 2.1: Electric charge, weak isospin, third component of weak isospin and weak hypercharge for different fermions

Particle	Q	T_w	T_{w3}	Y_w
u_L	$2/3$	$1/2$	$1/2$	$1/3$
d_L	$-1/3$	$1/2$	$-1/2$	$1/3$
u_R	$2/3$	0	0	$4/3$
d_R	$-1/3$	0	0	$-2/3$
$\nu_{e,L}$	0	$1/2$	$1/2$	-1
e_L	-1	$1/2$	$-1/2$	-1
e_R	-1	0	0	-2

The four gauge fields introduced in the EW theory are not the physical particles observed in nature. Through spontaneous symmetry breaking, which occurs after the introduction of the Higgs mechanism, these fields mix to form three heavy weak mediators (Z^0 , $W^\pm$) and a massless photon (γ). The same mechanism also provides for fermions to acquire mass. Their masses are set by free parameters known as Yukawa couplings, which quantify the strength of the interaction between each fermion and the Higgs field. Since the theory does not predict the actual values, the fermion masses are determined from experiments.

The charged $W^\pm$ bosons couple up-type and down-type quarks across generations. In the SM, this occurs because the weak eigenstates slightly differ from the mass eigenstates. This flavour mixing is described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix [13, 14], a unitary 3×3 matrix that relates the mass eigenstates (d , s , b) to the weak eigenstates (d' , s' , b'):

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (2.2)$$

Diagonal elements correspond to transitions within the same generation, whereas off-diagonal elements describe transitions between different generations. In Feynman

diagrams, a CKM element (V_{ij}) appears whenever there is a charged current vertex, where i and j represent the up- and down-type quark flavours. For example, the V_{cd} matrix element determines the coupling between charm and down quarks, and the probability for the transition is proportional to $|V_{cd}|^2$. The CKM elements are free parameters of the SM and have to be measured experimentally. Unitarity of the CKM matrix ($V_{\text{CKM}}V_{\text{CKM}}^\dagger = \mathbb{I}$) requires orthogonality between rows and columns:

$$\sum_i V_{ij}V_{ik}^* = \delta_{jk} \quad \text{and} \quad \sum_j V_{ij}V_{kj}^* = \delta_{ik}, \quad (2.3)$$

which leads to relations that can be represented as unitarity triangles, such as

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \quad (2.4)$$

The CKM matrix contains four free parameters, which in the standard parametrisation are three mixing angles and one CP-violating phase δ . It is the only source of quark-flavour change in the SM, and δ is the only source of CP violation in weak quark mixing. Because of the strong hierarchy, the matrix is often expressed in the Wolfenstein parametrisation [15], as an expansion in $\lambda \approx 0.225$ [16]. In this form, diagonal elements are close to one, while higher powers of λ suppress off-diagonal elements:

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \quad (2.5)$$

Here λ , A , ρ and η are all real parameters; only the combination $\rho - i\eta$ is complex, and η carries the CP-violating phase introduced above. Transitions within the same generation are most likely, those between neighbouring generations are suppressed, and transitions that skip a generation are highly suppressed. The magnitudes of the CKM matrix elements, obtained from a global fit to experimental data in which unitarity is imposed, are approximately [16]

$$|V_{ij}| \approx \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.225 & 0.973 & 0.042 \\ 0.009 & 0.041 & 0.999 \end{pmatrix}. \quad (2.6)$$

Since the CKM matrix governs flavour-changing interactions in the SM, this also controls the suppression pattern of flavour-changing neutral currents, a topic discussed in the next section.

2.3 Flavour-Changing Neutral Currents and Rare Decays

In the Standard Model, flavour-changing neutral currents (FCNCs), such as $c \rightarrow u\ell^+\ell^-$, can only occur via higher-order loop processes, such as the penguin and box diagrams shown in Figures 2.2a and 2.2b, whereas the corresponding tree-level transition mediated by the Z^0 (Figure 2.2c) is forbidden. This suppression is a

2.3. FLAVOUR-CHANGING NEUTRAL CURRENTS AND RARE DECAYS

consequence of the GIM mechanism, proposed to account for the observed smallness of FCNC rates and, in the process, predicting the existence of the charm quark [1]. Consequently, flavour-changing transitions at tree level occur only via $W^\pm$ bosons, while Z^0 conserve quark flavour. A $c \rightarrow u$ transition proceeds through an internal down-type quark $q = d, s, b$, and its CKM structure is governed by the unitarity relation $V_{cd}^*V_{ud} + V_{cs}^*V_{us} + V_{cb}^*V_{ub} = 0$. Using this relation to eliminate the d -quark term, the amplitude takes the schematic form

$$\mathcal{A}(c \rightarrow u) \propto \frac{1}{16\pi^2} \left\{ V_{cs}^*V_{us} [f(x_s) - f(x_d)] + V_{cb}^*V_{ub} [f(x_b) - f(x_d)] \right\}, \quad (2.7)$$

where $x_q \equiv m_q^2/M_W^2$, f denotes the relevant loop function and M_W is the mass of the $W^\pm$ boson. Each of the two terms is suppressed, but for a different reason. The first carries an unsuppressed CKM factor, $V_{cs}^*V_{us} = \mathcal{O}(\lambda)$, but the loop functions nearly cancel because both internal masses are far below M_W . The second involves the much heavier b quark, for which the loop functions no longer cancel, but its CKM factor $V_{cb}^*V_{ub} = \mathcal{O}(\lambda^5)$ is of order 10^{-4} . These two effects together constitute the GIM suppression of $c \rightarrow u$ transitions.

In the beauty sector, however, FCNC decays are less strongly suppressed. In $b \rightarrow s$ transitions the dominant internal quark is the top, for which $x_t = m_t^2/M_W^2 \approx 4.7$ gives no mass suppression, while the corresponding CKM factor $V_{tb}V_{ts}^* = \mathcal{O}(\lambda^2)$ is comparatively large. Neither produces the severe suppression found in charm transitions, which is why $b \rightarrow s\ell^+\ell^-$ and $b \rightarrow d\ell^+\ell^-$ decays have branching fractions within experimental reach and have been studied extensively [17, 18, 19, 20]. In the charm sector, both mechanisms act simultaneously, making $c \rightarrow u$ transitions far rarer, and correspondingly more sensitive to virtual contributions from heavy non-SM particles that need not respect the same cancellation pattern.

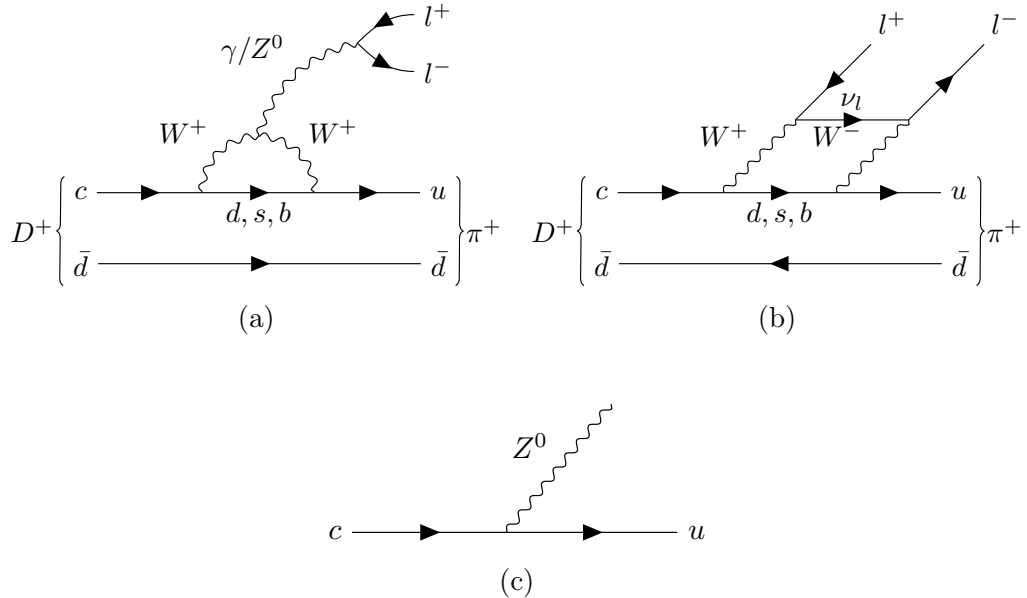


Figure 2.2: Electroweak penguin (a) and box (b) Feynman diagrams for the $c \rightarrow u\ell^+\ell^-$ transition. (c) Tree-level transition via Z^0 , which is forbidden in the SM.

2.4 Effective Field Theory Description

The decay $D^+ \rightarrow \pi^+ \ell^+ \ell^-$ receives contributions from two distinct dynamical regimes, commonly referred to as short-distance and long-distance, which differ in their theoretical descriptions and in their dependence on the dilepton invariant mass $M(\ell^+ \ell^-) = \sqrt{q^2}$, where q^2 is the squared four-momentum of the lepton pair. The short-distance contributions arise from local electroweak loop processes. They are described within an effective field theory framework and are perturbatively calculable, encoding sensitivity to potential physics beyond the SM.

Long-distance contributions originate from intermediate hadronic states that couple to the lepton pair. Rather than a short-distance FCNC transition, the dominant mechanism is a genuine long-distance process, in which an ordinary singly Cabibbo-suppressed weak transition $D^+ \rightarrow \pi^+ V$ is combined with the electromagnetic production of the lepton pair through $V \rightarrow \ell^+ \ell^-$, where $V = \rho^0, \omega, \phi$ [21]. These contributions are conventionally described by a q^2 -dependent shift of the Wilson coefficient C_9 , with each resonance modelled as a Breit–Wigner amplitude whose normalisation is fixed by the measured products $\mathcal{B}(D^+ \rightarrow \pi^+ V) \mathcal{B}(V \rightarrow \ell^+ \ell^-)$ [22, 23]. The relative strong phases between these amplitudes are not predicted and have long constituted the dominant theoretical uncertainty on the long-distance rate, although recent theoretical work has made progress in constraining them [21]. Since the vector resonances dominate the total decay rate, a branching-fraction measurement in the phase-space regions around their poles offers little direct sensitivity to short-distance dynamics. Angular and CP observables, by contrast, can be enhanced there through interference with the long-distance amplitude [24]; these are discussed in Section 2.5.

The pseudoscalar mesons η and η' also contribute, but their resonance parameters are two to three orders of magnitude smaller than those of the vector mesons, and their contribution to the decay rate is suppressed accordingly [2]. The η is nonetheless experimentally relevant because the Dalitz decay $\eta \rightarrow \gamma e^+ e^-$ produces a dielectron pair, so $D_{(s)}^+ \rightarrow \pi^+ \eta$ decays mimic the signal when the photon is not reconstructed. This background motivates the exclusion of the η mass region, discussed in Section 5.9.

Away from the resonance regions, the long-distance contribution decreases rapidly, making the low and high $M(\ell^+ \ell^-)$ regions particularly interesting for probing short-distance dynamics. Since the pion is the lightest hadronic final state accessible in charm decays, the channel $D^+ \rightarrow \pi^+ \ell^+ \ell^-$ provides the most favourable kinematic conditions for isolating such effects. The maximum kinematically allowed dilepton invariant mass squared for the exclusive decay is $q_{\text{max}}^2 = (m_{D^+} - m_{\pi^+})^2 \approx 2.99 \text{ (GeV}/c^2)^2$, corresponding to $M(\ell^+ \ell^-) < 1.73 \text{ GeV}/c^2$ (Section 5.11).

2.4.1 Effective Hamiltonian

The short-distance contribution to rare charm decays is described within the SM using the framework of effective field theory. The amplitude for a transition between an initial state ($|i\rangle$) and a final state ($|f\rangle$) is given by

$$A(i \rightarrow f) = \langle f | \mathcal{H}_{\text{eff}} | i \rangle, \quad (2.8)$$

where $\mathcal{H}_{\text{eff}}$ is the effective Hamiltonian for the transition.

Rare $c \rightarrow u \ell \ell$ processes are commonly described using the Operator Product

2.4. EFFECTIVE FIELD THEORY DESCRIPTION

Expansion (OPE), in which short- and long-distance effects are separated. The effective Hamiltonian takes the form of a sum of products of Wilson coefficients C_i and operators O_i , restricted here to those mediating the transition directly,

$$\mathcal{H}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}} \frac{\alpha_e}{4\pi} \sum_{i=7,9,10} C_i(\mu) \cdot O_i(\mu), \quad (2.9)$$

where G_F is the Fermi constant and $\alpha_e = e^2/(4\pi)$ is the fine-structure constant with the electromagnetic coupling e . In the charm sector, the CKM structure of Equation (2.7) is absorbed into the Wilson coefficients themselves, so no explicit CKM factor appears here, unlike in the corresponding $b \rightarrow s$ Hamiltonian. The contributing operators to the short-distance component are

$$\begin{aligned} O_7 &= \frac{m_c}{e} (\bar{u}_L \sigma_{\mu\nu} c_R) F^{\mu\nu} \\ O_9 &= (\bar{u}_L \gamma_\mu c_L) (\bar{\ell} \gamma^\mu \ell) \\ O_{10} &= (\bar{u}_L \gamma_\mu c_L) (\bar{\ell} \gamma^\mu \gamma_5 \ell), \end{aligned} \quad (2.10)$$

where $\sigma_{\mu\nu} = \frac{i}{2}[\gamma_\mu, \gamma_\nu]$, $F^{\mu\nu}$ is the electromagnetic field strength tensor, and $q_{L,R} = \frac{1}{2}(1 \mp \gamma_5)q$ denote the left- and right-handed chiral projections of the quark fields.

In this framework, the effects of heavy particles (W , Z , top and bottom quarks) are integrated out, and the short-distance effects are absorbed into the Wilson coefficients $C_i(\mu)$, which arise from the perturbative dynamics of the electroweak interactions and QCD renormalisation. The operators O_i describe the effective low-energy interactions among the remaining light fields, with masses below the renormalisation scale $\mu < m_b$. The coefficients are obtained by matching the effective theory onto the full theory at the electroweak scale μ_W and evolving them down to the charm scale $\mu_c = m_c$ [25, 26, 27]. At the electroweak scale, only the four-quark current-current operators $O_{1,2}$, induced by tree-level W -boson exchange and not shown in Equation (2.9), have non-vanishing coefficients, while all others are protected by the GIM mechanism. Contributions to $O_{7,9}$ are generated by operator mixing under QCD running and by integrating out the b quark at μ_b . The coefficients of $O_{1,2}$ remain of order unity at μ_c , while all others assume values at the per cent level or below. Through their matrix elements, $O_{1,2}$ also generate perturbative contributions to $O_{7,9}$, which are absorbed into effective coefficients $C_{7,9}^{\text{eff}}(q^2)$ [28]. The electromagnetic dipole operator O_7 mediates the radiative transition $c \rightarrow u \gamma^{(*)}$, where the lepton pair is produced by the subsequent conversion of the virtual photon. The operators O_9 and O_{10} are semileptonic, carrying a vector and an axial-vector lepton current, respectively.

In the SM, only the effective Wilson coefficients $C_{7,9}^{\text{eff}}$ receive non-zero contributions among the operators in Equation (2.9) at the charm scale μ_c , while C_{10} vanishes. As shown in Figure 2.3, C_9^{eff} shows noticeable q^2 dependence, while C_7^{eff} remains relatively flat across q^2 . Numerically, the effective coefficients at the charm scale are tiny. At $\mu_c = m_c$ one finds $C_7^{\text{eff}} \simeq (-0.0011 - 0.0041 i)$, and for $q^2 \gtrsim 1 \text{ GeV}^2$ the short-distance contribution is $|C_9^{\text{eff}}| \lesssim 5 \times 10^{-4}$ [25]. For comparison, the corresponding coefficient in $b \rightarrow s \ell^+ \ell^-$ transitions is $C_9 \approx 4$ [29].

Beyond the Standard Model, new physics can manifest either through modified Wilson coefficients or via additional operators such as scalar or pseudoscalar ($O_{S,P}$),

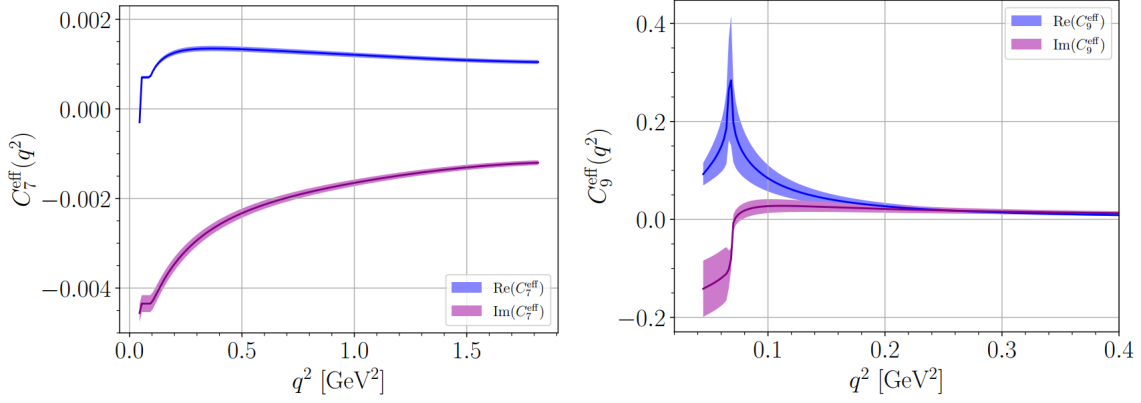


Figure 2.3: Real and imaginary parts of the SM Wilson coefficients C_7^{eff} and C_9^{eff} as a functions of q^2 . Taken from Ref [30].

tensor ($O_{T,T5}$), and chirality-flipped primed operators (O'_i). None of these contributes noticeably in the SM. The primed coefficients are suppressed by the mass ratio m_u/m_c , while the scalar, pseudoscalar and tensor coefficients vanish identically [28]. Their absence is what makes them sensitive probes of new physics, as discussed in Section 2.5.

2.4.2 SM Prediction

The short-distance transition $c \rightarrow u\ell^+\ell^-$ is calculated at the level of free quarks, and the resulting rate approximates the sum over all hadronic final states X_u^+ in $D^+ \rightarrow X_u^+\ell^+\ell^-$ rather than the single exclusive channel π^+ . The corresponding inclusive differential decay rate is [25]

$$\begin{aligned} \frac{d\Gamma(c \rightarrow u\ell^+\ell^-)}{dq^2} = & \frac{G_F^2 \alpha_e^2 m_c^3}{768\pi^5} \left(1 - \frac{q^2}{m_c^2}\right)^2 \left[\left(1 + 2\frac{q^2}{m_c^2}\right) (|C_9|^2 + |C_{10}|^2) \right. \\ & \left. + 4 \left(2\frac{m_c^2}{q^2} + 1\right) |C_7|^2 + 12\text{Re}[C_7 C_9^*] \right], \end{aligned} \quad (2.11)$$

where $q^2 = (p_c - p_u)^2 = (p_{\ell^+} + p_{\ell^-})^2$ at the parton level, with $(2m_\ell)^2 \leq q^2 \leq m_c^2$. Equation (2.11) is evaluated at leading order in the $1/m_c$ heavy-quark expansion, and lepton masses are set to zero except where they fix the phase-space boundary above. The C_{10} term is retained although it vanishes in the SM at the charm scale (Section 2.4.1), because physics beyond the SM can generate a non-zero value and the general form allows such contributions to be read off directly. In the Standard Model, the branching fractions for the electron and muon channels differ only due to kinematic effects arising from the lepton mass difference, while the τ channel is kinematically forbidden. Integrated over the accessible phase space, this parton-level rate corresponds to inclusive short-distance branching fractions of $\mathcal{B}(D^+ \rightarrow X_u^+ e^+ e^-) = 9.4 \times 10^{-10}$ and $\mathcal{B}(D^+ \rightarrow X_u^+ \mu^+ \mu^-) = 2.0 \times 10^{-10}$ [25].

This inclusive estimate does not, however, give the exclusive branching fraction of $D^+ \rightarrow \pi^+ \ell^+ \ell^-$ measured in this analysis. Obtaining the latter requires combining the same Wilson coefficients with the $D \rightarrow \pi$ transition form factors. Ref. [25] also performs this calculation and provides short-distance branching fractions separately for electrons and muons, integrated over different q^2 regions. The dominant uncertainty, reaching $\sim 250\%$, originates from the variation of the charm renormalisation

scale μ_c . The integrated short-distance branching fractions are predicted to be

$$\begin{aligned}\mathcal{B}(D^+ \rightarrow \pi^+ e^+ e^-)^{\text{SD}} &= 4.6 \times 10^{-12} (1_{-17\%}^{+211\%}), \\ \mathcal{B}(D^+ \rightarrow \pi^+ \mu^+ \mu^-)^{\text{SD}} &= 3.7 \times 10^{-12} (1_{-35\%}^{+249\%}).\end{aligned}\tag{2.12}$$

Predictions for individual q^2 bins are listed in Appendix A. These results are consistent with previous calculations [31, 32], but are a few orders of magnitude lower than other estimates [22, 23, 33, 34], mainly due to different treatment of form factors and cancellations.

The differential branching fraction of $D^+ \rightarrow \pi^+ \mu^+ \mu^-$ decay is shown in Figure 2.4, where short-distance contributions are suppressed by a few orders of magnitude relative to the long-distance ones. The widths of the bands indicate theoretical uncertainties arising from hadronic form factors, charm renormalisation scale, and resonance parameters. The largest uncertainty originates from the unknown phases in the Breit–Wigner parametrisation of the resonances.

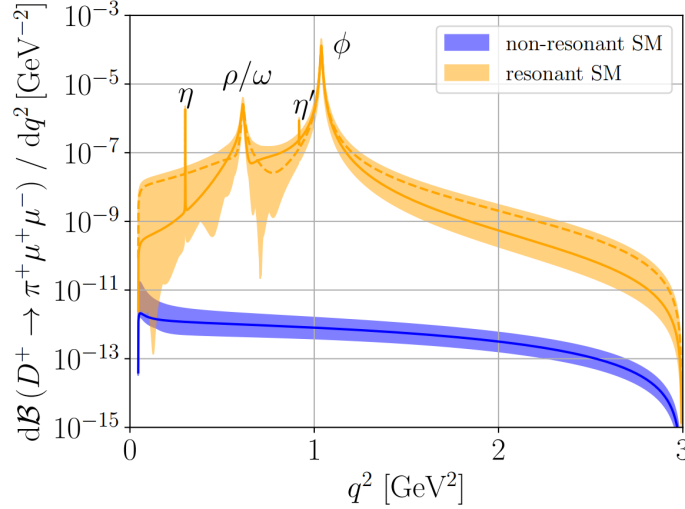


Figure 2.4: The differential branching fraction of the SM prediction for $D^+ \rightarrow \pi^+ \mu^+ \mu^-$, showing long- and short-distance contributions as yellow and blue curves. The band widths indicate theoretical uncertainties arising from hadronic form factors, resonance parameters and the charm renormalisation scale. Taken from Ref [35]. The distribution for the electron channel is qualitatively identical apart from the lower kinematic threshold.

Long-distance contributions raise the total branching fraction by several orders of magnitude relative to the short-distance rate, to the level of $\mathcal{O}(10^{-6})$ [2, 22, 23, 31, 36]. This figure refers to the rate integrated over the full $M(\ell^+ \ell^-)$ range and is almost entirely accounted for by the narrow ϕ pole. The resonant mode $D^+ \rightarrow \pi^+ \phi (\rightarrow \ell^+ \ell^-)$ alone contributes $\mathcal{B}(D^+ \rightarrow \pi^+ \phi) \times \mathcal{B}(\phi \rightarrow \ell^+ \ell^-) \approx 1.7 \times 10^{-6}$ [16]. Outside the resonance regions, the predicted rate is much lower. Recent calculations give branching fractions of $\mathcal{O}(10^{-9})$ in the dilepton-mass intervals below the ρ and above the ϕ [36]. A recent theoretical analysis, which determines the relative strong phases of the ρ , ω and ϕ contributions rather than treating them as free parameters, obtains values consistent with these in the same regions [21]. It also confirms that the long-distance amplitude dominates across the entire physical region, leaving the SM short-distance contribution unobservable anywhere within it.

The dominance of the ϕ pole is also an experimental asset. Because $D^+ \rightarrow \pi^+\phi(\rightarrow e^+e^-)$ shares the final state, topology, and kinematics of the signal decay but has a branching fraction of $(1.70 \pm 0.05) \times 10^{-6}$ [16] — three orders of magnitude larger than the short-distance rate — it provides an observable, signal-like control mode with which the full analysis chain can be validated. We exploit it for exactly this purpose (Chapters 6 and 9).

2.5 Sensitivity to New Physics

Rare charm decays occupy a unique position among FCNC processes. They are the only such transitions among up-type quarks accessible in hadron decays. The top quark decays before it hadronises, so no up-sector analogue of the kaon or B -meson systems exists. Physics beyond the SM whose flavour structure is aligned with the up-type quarks can therefore remain invisible in $b \rightarrow s$ and $s \rightarrow d$ transitions while producing observable effects in $c \rightarrow u$. Rare charm decays are in this sense a complementary probe of new physics, rather than a less sensitive repetition of the beauty-sector programme [2, 28].

New physics enters the effective Hamiltonian of Equation (2.9) either by shifting the Wilson coefficients or by generating operators absent in the SM [2, 25], extending the operator basis beyond $i = 7, 9, 10$. Because the SM short-distance contributions are strongly GIM-suppressed (Section 2.4.2), the sensitivity to NP contributions is unusually high: an enhancement that would be invisible against the much larger SM amplitudes of the beauty sector can dominate the rate here.

Ref. [37] relates new-physics effects in rare charm and kaon decays through a shared left-handed coupling structure, allowing bounds from one sector to constrain the other, and illustrates how this decay channel constrains new physics directly: an experimental upper limit on the branching fraction can be converted into a constraint on the underlying couplings without assuming a particular new-physics model. Of the two charm modes that probe these couplings, the stronger constraint depends on the lepton flavour. For muons it is the purely leptonic decay $D^0 \rightarrow \mu^+\mu^-$, whereas for electrons it is $D^+ \rightarrow \pi^+e^+e^-$, which enters their combined fit at a level comparable to that of high- p_T $pp \rightarrow e^+e^-$ production at the LHC. An improved measurement of the electron channel, such as the one presented in this thesis, would therefore directly strengthen that bound.

New-physics contributions to the branching fraction are not distributed uniformly across the decay phase space. Around the vector-meson poles, the rate is saturated by the long-distance amplitude, and although the resonance tails fall with increasing q^2 , they remain orders of magnitude above the short-distance contribution up to the kinematic endpoint [25]. Room for an observable enhancement is therefore greatest at very low and very high dilepton mass, the regions in which the theoretical description is also best controlled [36].

Of the two, the high-mass region provides the stronger constraints [38]. This motivates reporting the result separately in intervals of $M(e^+e^-)$ rather than only for the full range, as described in Section 5.11. The long-distance width predicted in Ref. [21] lies not far below the current upper limit in the muon channel, so any excess in these intervals would have to be compared against the long-distance amplitude, not only against the short-distance SM rate.

The vanishing of C_{10} , up to higher-order electromagnetic corrections at the per-mille level and far below experimental sensitivity, distinguishes charm from the kaon

and beauty sectors [2, 28]. Together with the absence of scalar, pseudoscalar and tensor operators (C_S, C_P, C_T, C_{T5}) and of the chirality-flipped coefficients C'_i , this makes the forward-backward asymmetry A_{FB} a null test [25, 36, 38], with residual SM contributions suppressed either by powers of m_D/M_W or by $\alpha_e/(4\pi)$ [35]. The flat term F_{H} of the angular distribution, introduced for $\bar{B} \rightarrow \bar{K} \ell^+ \ell^-$ decays [29] and adapted to charm in Refs. [35, 36, 38], is strongly suppressed in the SM: $\mathcal{O}(10^{-3})$ for $D^+ \rightarrow \pi^+ \mu^+ \mu^-$ at low q^2 , and even smaller for the electron mode [35]. A measurement of either observable significantly above its SM expectation would therefore be a clear sign of new physics.

Within the SM, the CP asymmetry in $D^+ \rightarrow \pi^+ \ell^+ \ell^-$ decays is expected to be very small, below 5×10^{-3} and peaking at $q^2 \sim m_\phi^2$ [25]. The CP-violating contribution is suppressed relative to the dominant CKM structure by the ratio, $|V_{cb}^* V_{ub}/(V_{cs}^* V_{us})| \sim 10^{-3}$ [2]. New physics can generate a considerably larger effect through interference between the resonant long-distance amplitude and a short-distance contribution carrying a CP-odd phase, with the resonance supplying the strong phase required for a rate asymmetry. Depending on the value of that strong phase, the asymmetry near the ϕ peak can reach the per cent level or above [24].

Several extensions of the SM have been examined in this context, and in most of them the effect in $D^+ \rightarrow \pi^+ \ell^+ \ell^-$ is small. An additional spin-1 weak triplet, introduced to account for what were then anomalies in $b \rightarrow s \ell^+ \ell^-$ data [39], shifts C_9 by too little to be observable [38]. The type-III two-Higgs-doublet model generates flavour-changing neutral currents at tree level [40], but the scalar and pseudoscalar operators it produces are constrained more strongly by $D^0 \rightarrow \mu^+ \mu^-$ [38]. In Z' extensions, the couplings are bounded by $D^0-\bar{D}^0$ mixing to such an extent that no observable signal survives [38]. Supersymmetric contributions, studied early on [22, 23], are largest in C_7 at low dilepton mass and are suppressed in $D^+ \rightarrow \pi^+ \ell^+ \ell^-$ by the q^2 factor of the $D^+ \rightarrow \pi^+ \gamma^*$ amplitude, which leaves this mode considerably less sensitive to such effects than the vector-meson mode $D \rightarrow V \ell^+ \ell^-$.

Leptoquarks are the exception. These hypothetical bosons couple quarks directly to leptons [41]. A scalar weak doublet, $(3, 2, 7/6)$, or a spin-1 weak singlet, $(3, 1, 5/3)$, generates C'_9 and C'_{10} at a level still allowed by present data, with a sensitivity at high dilepton mass comparable to that of $D^0 \rightarrow \mu^+ \mu^-$, though $D^0-\bar{D}^0$ mixing currently provides a somewhat stronger constraint [38]. In such scenarios, the ratio of the muon and electron decay widths at high dilepton mass, which equals unity in the SM up to lepton-mass corrections, is in most cases significantly enhanced [38]. Because such scenarios treat lepton flavours differently, limits from the dimuon modes do not carry over to the electron channel. A future comparison of the two channels, once both are measured with comparable precision, would constitute a direct test of lepton-flavour universality in $c \rightarrow u \ell^+ \ell^-$ transitions.

Improved limits on the branching fraction in the non-resonant regions therefore remain a direct route to constraining short-distance new physics in $c \rightarrow u \ell^+ \ell^-$ transitions. Global analyses of these transitions currently rely primarily on dimuon modes [28], while improved measurements of dielectron modes away from the long-distance resonance peaks, such as the one presented in this thesis, would provide important complementary input [28, 37].

2.6 Experimental Status

The flavour-changing neutral-current decay $D^+ \rightarrow \pi^+ \ell^+ \ell^-$ has been searched for over the past decades by a wide range of fixed-target, e^+e^- , and hadron-collider experiments. One of the earliest searches was carried out by the CLEO collaboration, which found no signal [42]. More recent searches have likewise found no signal and have instead set upper limits on the branching fractions and progressively improved them, as summarised in Table 2.2. For the electron channel, the most stringent upper limit to date, $\mathcal{B}(D^+ \rightarrow \pi^+ e^+ e^-) < 1.1 \times 10^{-6}$ at 90% confidence level (CL), was set by the *BABAR* collaboration using a data sample corresponding to an integrated luminosity of 384 fb^{-1} [4]¹. In contrast, the strongest constraint in the muon channel has been obtained by the LHCb collaboration, which reported an upper limit of $\mathcal{B}(D^+ \rightarrow \pi^+ \mu^+ \mu^-) < 6.7 \times 10^{-8}$ at 90% CL using a 1.6 fb^{-1} data set [3]. The same search also constrains the electron mode, but only to 1.6×10^{-6} , a factor of 24 weaker than its muon limit and weaker still than that of *BABAR*, which shows the difficulty of reconstructing low-momentum electrons in a hadron-collider environment. An earlier LHCb search reported upper limits separately in two intervals of the dilepton invariant mass, below the ρ and above the ϕ [45].

To date, neither the Belle nor the Belle II collaborations have published a dedicated search for the decay $D^+ \rightarrow \pi^+ \ell^+ \ell^-$. However, Belle has demonstrated its capability to study similar rare charm processes, such as $D^0 \rightarrow h(h')e^+e^-$, where $hh' = KK, \pi\pi, K\pi$ [46]. This thesis fills that gap for the electron final state, analysing the Belle and Belle II data sets together in a single simultaneous fit and extending the rare-charm programme of both experiments.

Table 2.2: Summary of experimental results of selected $D^+ \rightarrow \pi^+ \ell^+ \ell^-$ searches. All upper limits on the branching fractions are stated at the 90% confidence level and assume a three-body phase-space decay model. The limits are not defined over identical regions of the dilepton invariant mass: E791 apply no resonance veto, CLEO and *BABAR* exclude only the ϕ region and quote limits for the remaining phase space, while LHCb veto the interval $525 < M(\ell^+ \ell^-) < 1250 \text{ MeV}/c^2$ and extrapolate to the full range.

Collaboration	Sample size	Collision type	Branching fraction U.L.	
			e	μ
<i>BABAR</i> [4]	384 fb^{-1}	e^+e^-	1.1×10^{-6}	6.5×10^{-6}
LHCb [3]	1.6 fb^{-1}	pp	1.6×10^{-6}	6.7×10^{-8}
D0 [47]	1.3 fb^{-1}	$p\bar{p}$	–	3.9×10^{-6}
CLEO [48]	818 pb^{-1}	e^+e^-	5.9×10^{-6}	–
E791 [49]	2×10^{10} events	$\pi^- N$	5.2×10^{-5}	1.5×10^{-5}

¹A preliminary BES III result, $\mathcal{B}(D^+ \rightarrow \pi^+ e^+ e^-) < 3 \times 10^{-7}$ at 90% CL from 2.92 fb^{-1} of $\psi(3770)$ data, was reported in conference proceedings [43, 44] but has not been published. No restriction of the dielectron invariant-mass range is reported, so the limit is not directly comparable with those in Table 2.2.

2.7 Measurement Strategy

Two choices define the scope of this measurement. First, we reconstruct only the electron final state. The muon channel is already constrained to $\mathcal{O}(10^{-8})$ by LHCb, whose D^+ yield exceeds that of the B factories by three orders of magnitude [50]. The electron channel is a different case. Bremsstrahlung emission degrades the reconstructed invariant-mass resolution and raises the background for low-momentum electrons far more severely at a hadron collider than at a B factory [51], whereas these are precisely the final states in which an e^+e^- collider performs well, and where the largest improvement is available. Second, we analyse the Belle and Belle II samples together. The Belle data set, 980 fb^{-1} , has never been used to search for this decay, and adding it triples the sample relative to Belle II alone. The two detectors are sufficiently similar that both samples can be described in a single simultaneous fit, which constrains the shared parameters with the full data set and improves the sensitivity beyond that of separate fits combined afterwards (Chapter 9).

Two complementary sources of D^+ mesons are accessible at the Belle and Belle II experiments, operating primarily at the $\Upsilon(4S)$ resonance. D^+ mesons can be produced in the decay of a B meson, $B \rightarrow D^+ X$, or promptly in the continuum process $e^+e^- \rightarrow c\bar{c}$. Because B mesons are produced essentially at rest in the e^+e^- centre-of-mass frame at the $\Upsilon(4S)$ resonance, D^+ mesons originating from B -meson decays are kinematically limited to low momenta. In contrast, D^+ mesons produced promptly in continuum events are created nearly back-to-back with high momentum, since each charm quark carries approximately half the collision energy.

D^+ mesons from B -meson decays are less abundant but form a relatively clean sample, and the well-known $B\bar{B}$ production cross section allows their absolute branching fraction to be measured. Promptly produced D^+ mesons are far more abundant, but are accompanied by sizeable background, and their absolute branching fraction is more difficult to measure. For this reason, promptly produced D^+ decays are often studied by measuring branching fractions *relative* to a well-known reference decay of the D^+ meson, rather than as an absolute rate.

In this analysis, we target promptly produced D^+ mesons by requiring a high D^+ centre-of-mass momentum (p_{CMS}). We measure $\mathcal{B}(D^+ \rightarrow \pi^+ e^+ e^-)$ relative to the normalisation channel $D^+ \rightarrow \pi^+ K^+ K^-$, reconstructed around the ϕ resonance and with a well-known branching fraction [16]. This relative-measurement approach removes the need to know the integrated luminosity and the charm production cross-section, which cancel exactly since both modes are reconstructed from the same produced D^+ sample. Effects associated with the D^+ and the shared charged pion cancel only to first order in the difference between the two channels, since the two modes differ in their remaining final-state particles. We take the conservative approach of not crediting this cancellation as exact, and propagate the residual differences as systematic uncertainties (Chapter 8).

The analysis proceeds through the following stages, described in detail in the next chapters of this thesis. We reconstruct $D^+ \rightarrow \pi^+ e^+ e^-$ candidates in the non-resonant region of the dielectron invariant mass $M(e^+e^-)$, away from the $\phi(1020)$ resonance, and apply the centre-of-mass momentum requirement of $p_{\text{CMS}} > 2.5\text{ GeV}/c$ to suppress D^+ mesons originating from B -meson decays (Chapter 5). A random forest classifier, trained to further separate signal from combinatorial background, is subsequently applied using the significance of the D^+ flight distance, the total visible

energy of the event, p_{CMS} , and the angle between momentum and vertex vectors of the D^+ meson as input variables (Chapter 5). We extract signal candidates through an unbinned maximum-likelihood fit to the $M(\pi^+e^+e^-)$ distribution, performed simultaneously across the Belle and Belle II data samples (Chapter 7). In addition to the inclusive measurement, we repeat the fit in three $M(e^+e^-)$ intervals to probe potential q^2 -dependent effects. The decay $D^+ \rightarrow \pi^+\phi(\rightarrow e^+e^-)$, within the ϕ mass window, serves as a validation channel with a topology and kinematics closely resembling the signal (Chapter 6). The results reported in this thesis are obtained in a phase-space region from which the ϕ and η contributions have been removed. Because no significant excess of signal events is observed, we set an upper limit on $\mathcal{B}(D^+ \rightarrow \pi^+e^+e^-)$ at the 90% confidence level using the Neyman construction (Chapter 9).

We state the main results here for orientation. We observe the validation channel with a statistical significance of 12.0 standard deviations, with $\mathcal{B}(D^+ \rightarrow \pi^+\phi(\rightarrow e^+e^-)) = (1.75 \pm 0.22 \pm 0.12) \times 10^{-6}$, where the first uncertainty is statistical and the second systematic. No significant excess is observed in the signal mode, and we obtain $\mathcal{B}(D^+ \rightarrow \pi^+e^+e^-) < 4.7 \times 10^{-7}$ at 90% confidence level over the full $M(e^+e^-)$ range excluding the ϕ and η regions, improving on the previous best limit [4] by a factor of 2.3.

3. The Belle and Belle II Experiments

This chapter describes the experimental facilities that provided the data analysed in this thesis. Both the Belle and Belle II experiments are located at the KEK laboratory in Japan. We include the Belle dataset alongside Belle II to increase the overall signal sensitivity, as its sample is more than twice as large provides a substantial addition to the Belle II dataset despite the different detector configuration. The chapter begins with an overview of the SuperKEKB asymmetric-energy e^+e^- collider, followed by a detailed description of the Belle II detector and its subsystems. The chapter concludes with a summary of the Belle detector, highlighting the key hardware differences relevant to the combined analysis strategy.

3.1 The KEKB and SuperKEKB Accelerators

The SuperKEKB collider [52] is an asymmetric electron-positron collider at the High Energy Accelerator Research Organisation facility (KEK) in Tsukuba, Japan, and the successor to the KEKB collider [53, 54] built in the same tunnel. It is designed as a high-luminosity factory for producing B mesons, D mesons, and τ leptons. It operates two storage rings: the Low Energy Ring (LER) carrying 4 GeV positrons and the High Energy Ring (HER) for 7 GeV electrons, which are tuned for a centre-of-mass (CM) energy of 10.58 GeV, corresponding to the $\Upsilon(4S)$ resonance. In addition, it can also run at CM energies from the $\Upsilon(1S)$ to $\Upsilon(6S)$ resonances, enabling a broader spectroscopy programme. The beam parameters for both colliders are summarised in Table 3.2. The layout of the accelerator complex is shown in Figure 3.1.

The $\Upsilon(4S)$ resonance decays almost exclusively to $B^0\bar{B}^0$ or B^+B^- pairs, with negligible available energy to produce additional particles. This results in low detector occupancy and high reconstruction efficiency for both charged and neutral final-state particles, including photons, π^0 , and K_L^0 . The well-known initial state and near-hermetic detector coverage further enable analyses that exploit missing-energy signatures, which are relevant for decays involving neutrinos or other weakly interacting particles. The $\Upsilon(4S)$ mass lies only just above the $B\bar{B}$ threshold, so the B mesons are produced nearly at rest in the centre-of-mass system. The asymmetric beam energies boost the centre-of-mass system along the beam axis, so that the B and D mesons travel a measurable distance from the interaction point before decaying. This enables precision measurements of decay vertices, lifetimes, mixing parameters, and CP violation.

In addition to $B\bar{B}$ production, lighter-quark pair production $e^+e^- \rightarrow q\bar{q}$ (where $q = u, d, s, c$), referred to as continuum, contributes approximately 3.7 nb to the total hadronic cross-section, as summarised in Table 3.1. The $c\bar{c}$ continuum component

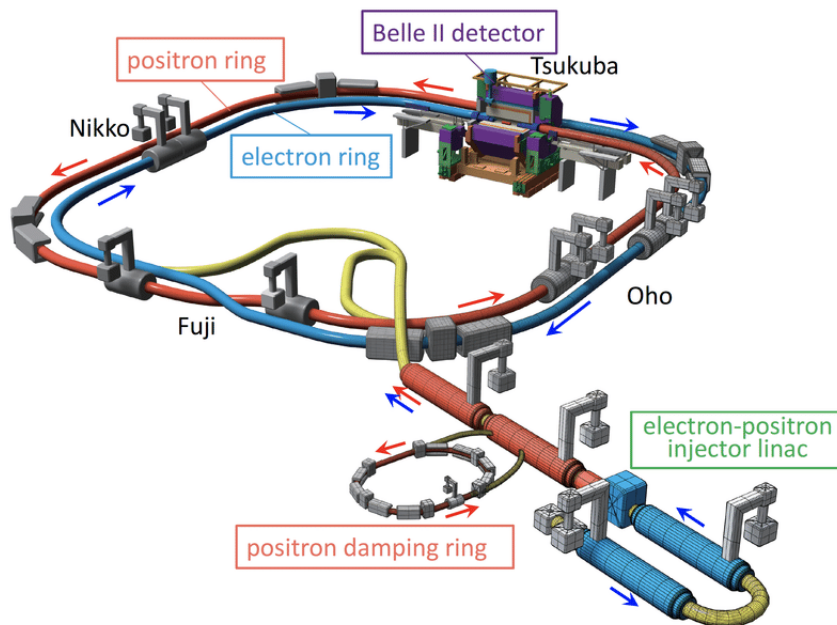


Figure 3.1: Schematic view of SuperKEKB collider facility, showing the two main storage rings (LER and HER), the linear accelerator (linac), and the interaction point where the Belle II detector is installed.

($\sigma \approx 1.3$ nb) leads to the production of charm hadrons, including D^+ mesons, which is the primary source of D^+ candidates used in this analysis. D mesons can also be produced as decay products of B mesons, but because the B mesons are nearly at rest in the centre-of-mass system, such D^+ candidates are kinematically limited to centre-of-mass momenta below about $2.5 \text{ GeV}/c$. Continuum-produced D^+ mesons inherit roughly half the collision energy and extend well beyond this limit, thus a requirement on the centre-of-mass momentum isolates them cleanly, as described in Chapter 5. Large cross-section QED processes, namely Bhabha scattering ($e^+e^- \rightarrow e^+e^-$) and digamma production ($e^+e^- \rightarrow \gamma\gamma$), are used to measure the luminosity and calibrate detector responses [55], since these have large production rates, almost no background contamination, and accurate theoretical predictions.

The instantaneous luminosity $\mathcal{L}$ reflects the rate of data collection and connects the rate of produced events to the cross-section σ of a given physics process:

$$\frac{dN}{dt} = \mathcal{L} \cdot \sigma, \quad (3.1)$$

where N is the number of produced events. The instantaneous luminosity is given by:

$$\mathcal{L} = \frac{\gamma_{\pm}}{2er_e} \left(1 + \frac{\sigma_y^*}{\sigma_x^*} \right) \frac{I_{\pm} \xi_{y\pm}}{\beta_y^*} \frac{R_{\mathcal{L}}}{R_{\xi_y}}, \quad (3.2)$$

where γ , e , and r_e are the Lorentz factor, the elementary electric charge, and the classical electron radius, respectively; σ_x^* and σ_y^* are the horizontal and vertical bunch widths at the interaction point; I is the beam current; β_y^* is the vertical betatron function at the interaction point; ξ_y is the vertical beam-beam parameter; and $R_{\mathcal{L}}$, R_{ξ_y} are geometric reduction factors accounting for the hourglass effect and

3.1. THE KEKB AND SUPERKEKB ACCELERATORS

Table 3.1: Production cross sections at $\sqrt{s} = 10.58$ GeV [56].

Physics process	Cross Section [nb]
$\Upsilon(4S)$	1.1
$u\bar{u} (\gamma)$	1.6
$d\bar{d} (\gamma)$	0.4
$s\bar{s} (\gamma)$	0.4
$c\bar{c} (\gamma)$	1.3
$\tau^+\tau^- (\gamma)$	0.9
$\mu^+\mu^- (\gamma)$	1.1
Bhabha ($e^+e^- (\theta_{\text{lab}} \geq 17^\circ)$)	44
$\gamma\gamma (\theta_{\text{lab}} \geq 17^\circ)$	2.4

the crossing angle. The product of the integrated luminosity and the cross-section then gives the number of produced events of interest:

$$N = \mathcal{L}_{\text{int}} \cdot \sigma, \quad \mathcal{L}_{\text{int}} = \int_0^t \mathcal{L}(t') dt'. \quad (3.3)$$

Table 3.2: Machine parameters for KEKB and SuperKEKB, given as LER/HER where the two rings differ. Achieved and design values are taken from Ref. [54, 56]. The current SuperKEKB values are those reported at the end of Run 1 in June 2022 [57].

	KEKB Achieved	SuperKEKB Designed	SuperKEKB June 2022
Energy [GeV]	3.5/8.0	4.0/7.0	4.0/7.0
$\beta\gamma$	0.425	0.284	0.284
Current [A]	1.6/1.2	3.6/2.6	1.3/1.1
ξ_y	0.129 / 0.090	0.090 / 0.088	0.041 / 0.028
β_y^* [mm]	5.9/5.9	0.27 / 0.41	1.0/1.0
σ_y^* [nm]	940	48/62	215
Luminosity [$10^{34} \text{ cm}^{-2} \text{ s}^{-1}$]	2.11	80	4.65

The luminosity can be increased mainly through three parameters: total beam current (I), the vertical beam-beam parameter (ξ_y), and the vertical beta function at the interaction point (IP) (β_y^*). In practice, raising the beam current is limited by hardware constraints and operating costs, so it is increased only moderately,

while values of ξ_y higher than those achieved at the KEKB collider are impractical due to beam instabilities. The dominant gain at SuperKEKB therefore comes from reducing β_y^* through the nano-beam collision scheme [58]. By design, this squeezes the vertical beam size at the IP by a factor of about 20, from approximately $1\mu\text{m}$ at KEKB to $\approx 60\text{nm}$. The beams reach their minimum size only in the immediate vicinity of the IP, growing back to larger sizes within one β_y^* of the collision point, as shown in Figure 3.2. The large half crossing angle ($\phi \approx 41.5\text{mrad}$ at SuperKEKB, compared to $\approx 11\text{mrad}$ at KEKB) combined with a small horizontal beam size ensures that only a small fraction of each bunch overlaps at any given time, giving an effective bunch length of $d = \sigma_x^* / \sin \phi \approx 0.25\text{mm}$, much shorter than the actual bunch length ($\sigma_z \approx 5\text{mm}$). Because the effective bunch length d , rather than σ_z , sets the limit on focusing, the hourglass condition becomes $\beta_y^* \gtrsim d$ instead of $\beta_y^* \gtrsim \sigma_z$, allowing β_y^* to be squeezed far below what a head-on collision would permit. Conceptually, the nano-beam scheme mimics collisions of many short micro-bunches, reducing the luminous volume to approximately 5% of that at KEKB. Combined with doubled beam currents (3.6 A and 2.6 A at SuperKEKB compared to 1.6 A and 1.2 A at KEKB), this yields a design luminosity of $8 \times 10^{35}\text{cm}^{-2}\text{s}^{-1}$, approximately 40 times higher than the peak luminosity of $2.11 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$ achieved at KEKB.

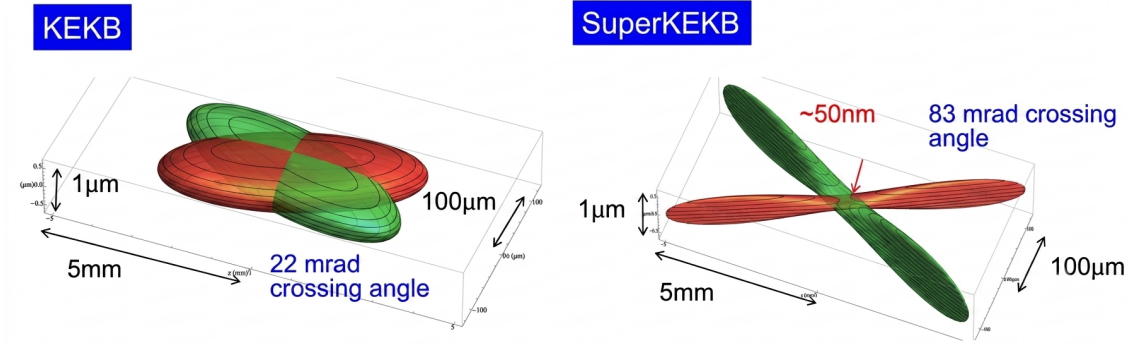


Figure 3.2: Schematic beam collision geometry at KEKB (left) and SuperKEKB (right). The nano-beam scheme uses a larger crossing angle and tighter vertical beam focusing. Taken from Ref. [59].

The KEKB collider operated from 1999 to 2010, delivering the 980fb^{-1} recorded by the Belle experiment. SuperKEKB began physics data taking in March 2019 (Run 1) and collected 428fb^{-1} before the first long shutdown (LS1) in June 2022. Run 2 started in January 2024 and finished in June 2026; a world-record instantaneous luminosity of $5.3 \times 10^{34}\text{cm}^{-2}\text{s}^{-1}$ was achieved, bringing the total collected dataset to approximately 914fb^{-1} . The analysis presented in this thesis uses the 428fb^{-1} recorded before LS1, which was the full calibrated dataset available when the analysis was performed.

3.2 The Belle II Detector

The Belle II detector is a large-solid-angle magnetic spectrometer [56] surrounding the single interaction point of the SuperKEKB collider. Like its predecessor, it is optimised for the study of B , D , and τ physics at the luminosity frontier, which sets demanding requirements: wide angular acceptance and high reconstruction efficiency for final-state particles, good momentum resolution over the full kinematic

range (from about 50 MeV/c to 8 GeV/c), precise decay-vertex reconstruction, and efficient identification of pions, kaons, protons, electrons, and muons across that range. Because the average momentum of the final-state particles is below 1 GeV/c, the material budget must be kept low to limit multiple scattering, and the material in front of the calorimeter must be minimised so that photons with energies as low as 20 MeV can be detected.

Due to the asymmetric beam energies, the centre-of-mass system is boosted along the beam axis, and the detector has extended coverage in the forward direction (the direction of the high-energy electron beam) towards which most decay products are produced. From the interaction point outward, the detector consists of a VerTeX Detector (VXD), composed of two layers of silicon-PiXeL Detector (PXD) and four layers of double-sided strip Silicon Vertex Detector (SVD); a 56-layer central drift chamber (CDC); a time-of-propagation (TOP) detector in the barrel region; an aerogel ring-imaging Cherenkov (ARICH) detector in the forward endcap; and an electromagnetic calorimeter (ECL). These subdetectors are housed inside a superconducting solenoid providing a 1.5 T magnetic field, and the iron flux return outside the solenoid is instrumented as the K_L^0 and muon detector (KLM). The charged-particle tracking covers nearly the full solid angle, spanning the polar-angle range from 17° to 150° and the complete azimuthal range.

We describe the Belle II detector in full below, while for the Belle detector only the key differences are highlighted, since the analysis uses datasets from both experiments. Figure 3.3 compares the layouts of the two detectors.

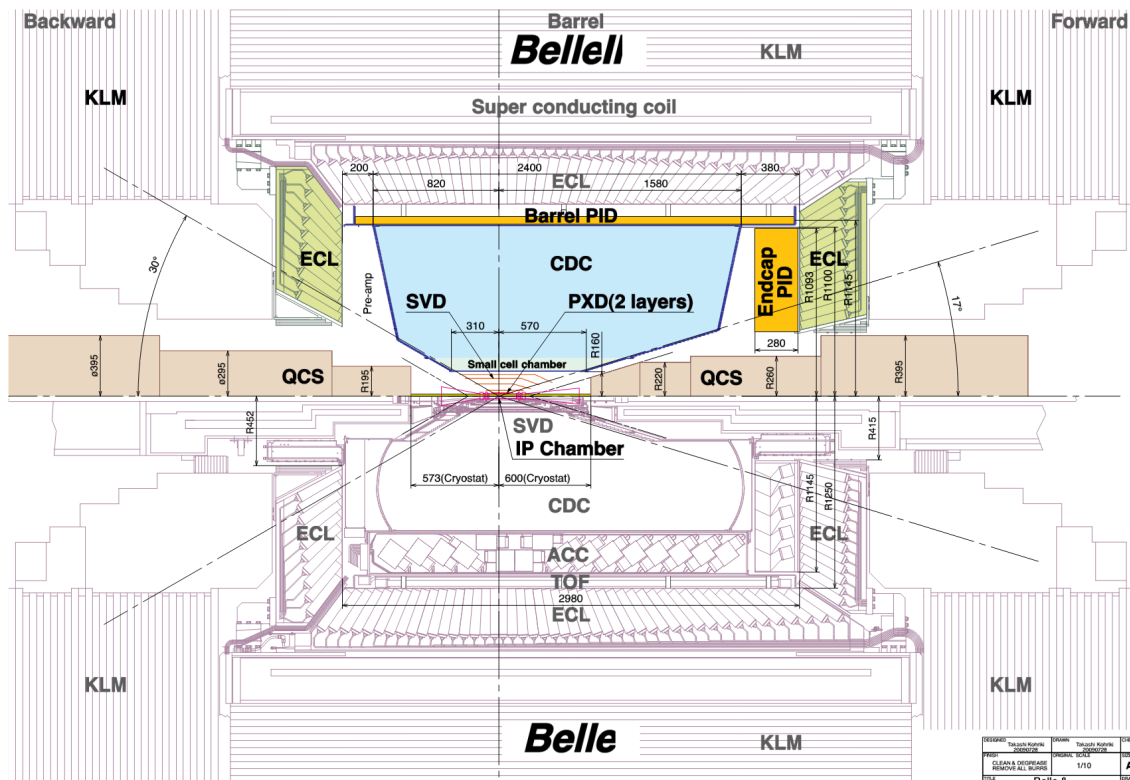


Figure 3.3: Schematic view of the Belle II detector (top half) compared with the previous Belle detector (bottom half). Taken from Ref. [60].

3.2.1 Coordinate System

The Belle II coordinate system is defined with its origin at the nominal interaction point. The z axis is aligned with the solenoid axis and bisects the crossing angle between the two beams, nearly coinciding with the high-energy electron-beam direction. The y axis points vertically upward, and the x axis is horizontal, orthogonal to y and z , pointing outward from the centre of the SuperKEKB ring. The polar angle θ is measured from the $+z$ axis and ranges from 0° to 180° , the azimuthal angle ϕ is measured around the z axis from the x axis and ranges from 0° to 360° , and the radial coordinate in the transverse plane is $\rho = \sqrt{x^2 + y^2}$. The *longitudinal* direction is parallel to the z axis and the *transverse* plane is the x - y (or ρ - ϕ) plane. Since the centre-of-mass system is boosted along $+z$, the *forward* and *backward* directions correspond to $+z$ and $-z$, respectively.

3.2.2 Tracking System

The tracking system comprises the vertex detector (VXD) and the central drift chamber (CDC). Together they reconstruct the trajectories of charged final-state particles, from which the momenta are determined through the curvature in the 1.5 T magnetic field, and the decay vertices are inferred by extrapolation towards the interaction point. The VXD consists of two silicon-based detectors: the innermost pixel detector (PXD) [61] and the surrounding double-sided strip silicon vertex detector (SVD) [62]. In both, a traversing charged particle ionises the sensitive silicon, and the liberated charge carriers are collected to form the signal.

The PXD is based on depleted field-effect transistor technology (DEPFET), with very thin $75\text{ }\mu\text{m}$ sensing layers. A thin sensor with a fine pixel granularity is essential to provide precise hits while minimising multiple-Coulomb scattering of the charged tracks. The beryllium beam pipe has an inner radius of 10 mm at the interaction point, about two-thirds of that in the final Belle configuration, which allows the innermost PXD layer to be placed at a radius of only 14 mm — closer to the interaction point than at Belle, improving the impact-parameter resolution. The PXD has two layers, at radii of 14 mm and 22 mm, comprising 8 and 12 ladders respectively, for a total of about 8 million pixels with sizes ranging from 50×50 to $50 \times 85\text{ }\mu\text{m}^2$. Figure 3.4 shows the layout of the two VXD subdetectors.

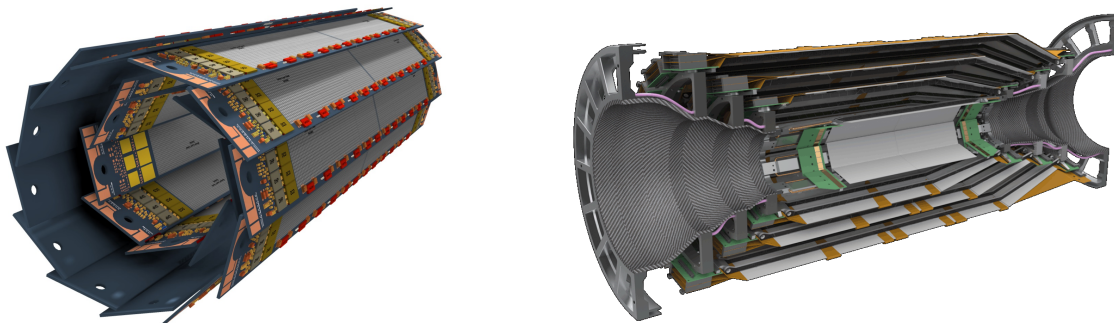


Figure 3.4: Schematic view of PXD (left) and SVD (right).

The SVD uses silicon-strip sensors rather than pixels, which reduces the number of readout channels compared to a fully pixelated outer detector. It surrounds the PXD with four layers at radii of 39 mm, 80 mm, 104 mm and 135 mm, composed of 7, 10, 12, and 16 independently read-out ladders. The double-sided SVD sensors are

320 μm thick with 512 or 768 readout strips per side, depending on the layer, and a readout pitch ranging from 50 μm to 240 μm . In the forward region the sensors are sloped at an angle, which improves the geometrical acceptance for the forward-boostered tracks and reduces the material they cross. The SVD provides the link that extrapolates tracks reconstructed in the CDC inward to the PXD, and together with the PXD it reconstructs low-momentum tracks that do not reach the outer detectors. The VXD covers the polar-angle range from 17° to 150° . The measured SVD hit resolution in the transverse plane is approximately 7 μm to 12 μm [63], which together with the PXD hits allows the D^+ decay vertex to be resolved from the interaction point and the combinatorial background to be suppressed.

The CDC is a large-volume drift chamber composed of small drift cells. It reconstructs charged-particle trajectories and measures their transverse momenta through curvature within the 1.5 T magnetic field. CDC also contributes to particle identification through the specific ionisation energy loss (dE/dx) and provides a fast hardware trigger signal for charged particles. It contains 14 336 gold-plated tungsten sense wires arranged in 56 layers, together with 42 240 aluminium field wires that shape the electric field. The wires are strung either axially, parallel to the solenoid field, or at a small stereo angle to it; combining the two orientations enables three-dimensional track reconstruction. The CDC occupies radii from 160 mm to 1130 mm, covers the polar-angle range from 17° to 150° , and is filled with a 50:50 helium–ethane ($\text{He-C}_2\text{H}_6$) gas mixture. Low-momentum tracks that do not reach the outer PID detectors can be identified using the CDC dE/dx alone. The CDC achieves a drift distance resolution of 120 μm [64]. The chamber geometry and its wire configuration are shown in Figure 3.5.

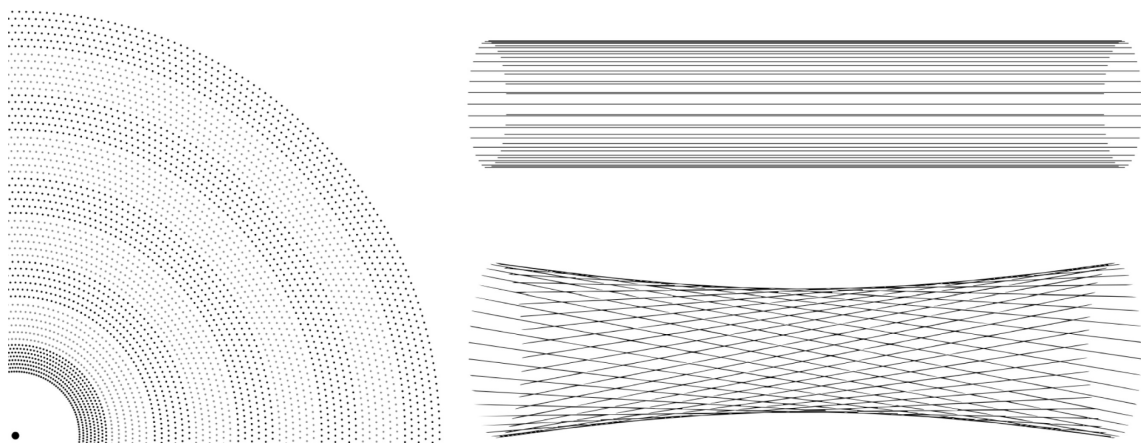


Figure 3.5: The left diagram shows a transverse cross-section of the CDC. The right diagram illustrates the axial (top) and stereo (bottom) wire orientations, with exaggerated skew on the stereo wires for visibility. Taken from Ref. [64].

3.2.3 Cherenkov Ring Imaging Counters

The Belle II detector employs two complementary Cherenkov based detectors in the barrel and forward regions to achieve robust particle identification (PID): the Time-Of-Propagation counter (TOP) [65] and the Aerogel Ring Imaging Cherenkov detector (ARICH) [66]. Both detectors exploit the principle that charged particles travelling through a medium faster than light propagates in that medium emit

Cherenkov radiation at a characteristic angle,

$$\cos \theta_c = \frac{1}{\beta n}, \quad (3.4)$$

where β is the particle velocity relative to c and n is the refractive index of the radiator. By measuring the Cherenkov angle and combining it with the particle momentum from the tracking system, the particle mass (or identity) can be inferred and tested against known particle hypotheses.

TOP Counter (Barrel Region)

The TOP counter provides particle identification in the barrel ($31^\circ < \theta < 128^\circ$) region for particles with momenta above 1 GeV/c. It consists of 16 modules arranged in a barrel geometry, with each module containing two quartz radiator bars. Each bar is 45 cm wide, 2 cm thick, and 125 cm long. When a charged particle traverses a quartz bar, it emits Cherenkov photons. Due to the high refractive index of quartz ($n \approx 1.46$), a significant fraction of these photons undergo total internal reflection at the bar boundaries and propagate along the length of the bar. This encodes the Cherenkov angle information in the pattern of photon arrival times and positions. On average, light from slower particles takes longer to reach the photo-sensors.

At one end of each bar, a prism collects light and guides it to an array of micro-channel-plate photomultiplier tubes (MCP-PMTs), which detects the arriving photons and records both their impact position (two spatial coordinates) and arrival time. A mirror on the opposite end reflects photons backward to further increase detection efficiency. The TOP achieves a single-photon time resolution of approximately 100 ps. Figure 3.6 shows the geometry of a TOP module and the propagation of Cherenkov photons within the quartz bar.

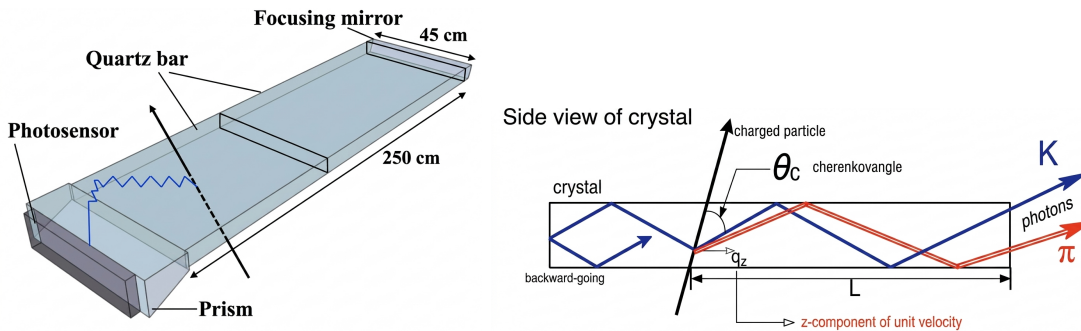


Figure 3.6: A schematic view of one TOP module showing its physical dimensions (left). Cross-section of a quartz bar illustrating Cherenkov light propagation via internal reflection for two particle hypotheses.

ARICH Counter (Forward Region)

The ARICH is the primary PID detector in the forward region ($15^\circ < \theta < 34^\circ$), designed to separate kaons from pions and protons over nearly the full momentum range (0.5–4 GeV/c), with additional separation between pions, muons, and electrons below 1 GeV/c. It is a proximity-focusing aerogel ring imaging Cherenkov counter featuring a novel dual-layer radiator design [67, 68].

The radiator consists of two 2-cm-thick layers of hydrophobic aerogel tiles with slightly different refractive indices: $n_1 = 1.045$ and $n_2 = 1.055$. Particles traversing

the radiator emit Cherenkov rings at different angles corresponding to each refractive index. The emitted photons propagate across a dedicated expansion volume of about 20 cm. This expansion distance allows the Cherenkov photons to project into fully formed rings before hitting the photon detector surface. The two refractive indices are chosen such that the rings from the first and second layers overlap on the detector plane. The detector plane comprises an array of 420 hybrid avalanche photodetectors (HAPDs). This overlapping pattern boosts the Cherenkov photon yield while keeping the ring image compact on the detector plane. Figure 3.7 illustrates this proximity-focusing arrangement and the layout of the detector.

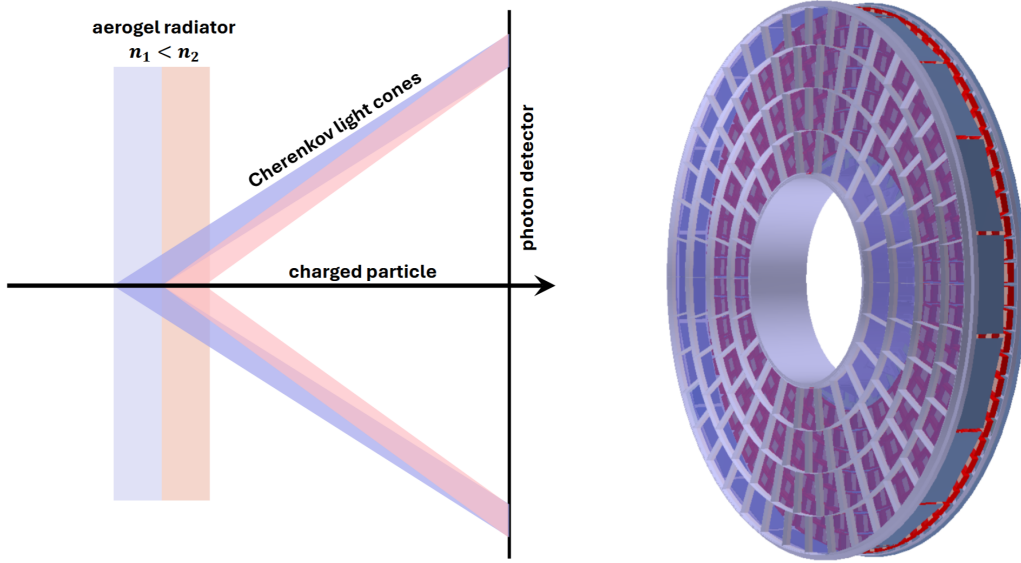


Figure 3.7: Left diagram illustrates the dual-layer aerogel radiator in a proximity-focusing configuration. Right diagram shows a schematic view of the ARICH detector, with aerogel radiator shown in light grey and HAPDs in dark red.

3.2.4 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECL) [60] measures the energy and angular coordinates of photons and acts as the primary subdetector for electron identification. In addition, it contributes to the trigger system and serves as the luminosity monitor for SuperKEKB.

The ECL consists of 8736 thallium-doped caesium iodide CsI(Tl) scintillating crystals of varying shapes, inherited from the Belle experiment. The barrel region contains 6624 crystals, while the remaining 2112 are distributed across the forward and backward endcaps, in the layout shown in Figure 3.8. Together, they provide coverage over the polar angle range 12.4° – 155.1° , covering the large majority of the solid angle and enabling reconstruction of the total visible energy of a collision event.

Each crystal is shaped as a truncated pyramid with an average cross-section of $6 \times 6 \text{ cm}^2$ and a length of 30 cm, corresponding to 16.1 radiation lengths. The crystals are oriented to point approximately towards the interaction point. Incident high-energy photons and electrons initiate electromagnetic showers through bremsstrahlung and electron–positron pair production. The resulting scintillation light is collected by Hamamatsu Photonics photodiodes mounted at the rear face

of each crystal. The photon energy resolution measured with $e^+e^- \rightarrow \gamma\gamma$ events, averaged over the calorimeter, is 1.7 % [56].

The instantaneous and integrated luminosity at SuperKEKB is monitored using the ECL endcaps alone, exploiting Bhabha scattering and two-photon QED processes. These are identified by the characteristic geometric coincidence of high-energy deposits in the forward and backward endcaps.

The primary baseline observable for distinguishing electrons from other charged particles, such as muons and pions, is the energy-to-momentum ratio, E/p . This variable is defined as the ratio between the cluster energy measured in the ECL and the reconstructed momentum of the associated charged track. While this provides excellent separation power for $p > 1$ GeV/ c , at lower momenta, bremsstrahlung energy losses degrade the E/p discriminating power.

Because electrons radiate bremsstrahlung photons while traversing the detector material, their reconstructed momentum underestimates the momentum at production. A recovery algorithm restores the radiated energy by adding back neutral ECL clusters found close to the electron trajectory, as described in Section 5.3.

BELLE Csi ELECTROMAGNETIC CALORIMETER

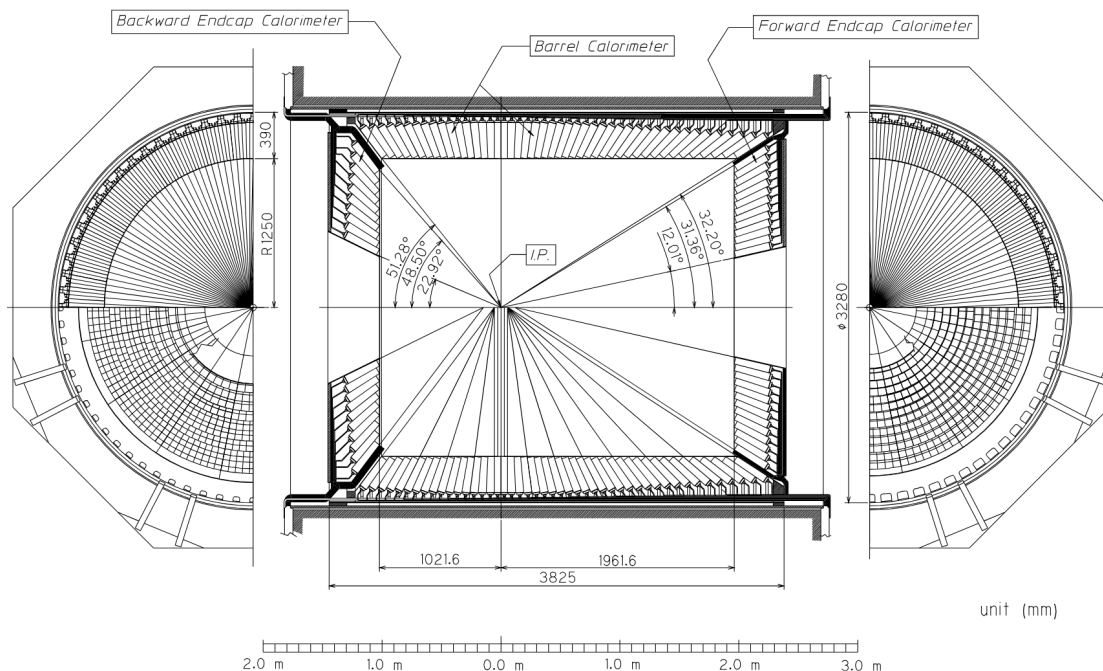


Figure 3.8: Layout of ECL detector. Taken from Ref. [69].

3.2.5 Muon and K_L^0 Identification (KLM)

The K_L^0 and muon detector (KLM) serves a dual purpose: it identifies muons and K_L^0 mesons, which travel through the inner subdetectors largely undetected, and acts as the magnetic flux return yoke for the solenoid. The KLM covers the polar angle range 20° – 155° , comprising a barrel and two endcap regions.

The detector consists of thin planar detector panels inserted between 4.7 cm thick steel plates, which provide more than 3.9 absorption lengths and serve as the solenoid flux return. In the barrel region, the outer 13 active layers consist of glass-electrode resistive plate chambers (RPCs), inherited from the Belle experiment, while the

inner two layers employ new scintillator panels coupled via wavelength-shifters to silicon photomultipliers (SiPMs). The endcaps are instrumented exclusively with scintillator panels, as RPCs suffer from long dead times following electrical discharges, leading to efficiency degradation under the elevated beam backgrounds at SuperKEKB.

3.2.6 Trigger and Data Acquisition

The extreme luminosity provided by the SuperKEKB collider produces an immense volume of collision and background events which far exceeds the capacity of permanent storage, making it impractical to record every event. A dedicated trigger and data acquisition (DAQ) system therefore performs real-time event selection, retaining only physically interesting collisions while rejecting beam-induced backgrounds. The trigger is required to achieve almost 100% efficiency for hadronic, $\mu^+\mu^-$, $\tau^+\tau^-$, Bhabha, and two-photon events, with a timing precision below 10 ns and a minimum inter-event separation of 200 ns. Bhabha and two-photon events, which would otherwise dominate the data stream, are intentionally downscaled by a factor of 100 to match data acquisition capabilities. To accomplish this, Belle II employs a two-stage trigger architecture: a hardware-based Level-1 (L1) trigger that performs a fast loose selection, followed by a software-based High-Level Trigger (HLT) for further refinement.

The L1 trigger processes inputs from the CDC, ECL, TOP, and KLM subdetectors, and operates at a maximum output rate of 30 kHz to accommodate the target SuperKEKB luminosity. All processing is implemented on field-programmable gate arrays (FPGAs) with a fixed latency of 5 μ s and a timing jitter of approximately 10 ns. The CDC provides three-dimensional track information, suppressing tracks not originating from the interaction point and supplying track momenta, angles, and production positions. The ECL contributes the energy and position of calorimeter clusters. Timing information is gathered from the CDC, ECL, and TOP, while the KLM trigger identifies muon track candidates. All inputs are processed by the Global Reconstruction Logic (GRL), which matches tracks to clusters, before passing the combined information to the Global Decision Logic (GDL), which issues the final L1 trigger signal. The L1 trigger also applies a Bhabha veto and an injection veto, the latter of which temporarily halts data taking during the severe background bursts immediately following a bunch injection into the storage ring.

Upon an L1 trigger, the DAQ system synchronises all subdetectors and transfers the full event data via high-speed optical links to the HLT computing farm. The DAQ is designed to sustain a maximum rate of 30 kHz.

The HLT is a purely software-based trigger that performs online event reconstruction using the standard Belle II software framework `basf2`, utilising information from all subdetectors except the PXD. It reduces the L1 output by approximately a factor of five, by analysing basic topological quantities and particle kinematics such as track multiplicity, vertex positions, and total ECL energy deposition. A key function of the HLT is the computation of regions of interest (ROIs) for the PXD: tracks reconstructed from SVD and CDC hits are extrapolated to the pixel detector, and only PXD hits within the corresponding ROI are read out and appended to the event record, substantially reducing the PXD data volume. The HLT additionally flags events by physics process category and provides data quality monitoring (DQM) information.

The $D^+ \rightarrow \pi^+ e^+ e^-$ final state consists of three charged tracks and is recorded by the standard hadronic trigger with an efficiency close to unity.

3.3 The Belle Detector

The Belle detector operated at the KEKB asymmetric e^+e^- collider from 1999 to 2010, accumulating a dataset of 980 fb^{-1} . It was located at the same experimental site as Belle II and shared the same design philosophy of a general-purpose collider detector, as described in Ref. [69]. We use both the Belle and Belle II datasets in this analysis, with all reconstruction and selection steps performed independently for each detector. The two samples are combined only at the fit stage, where the branching fraction enters as a parameter common to both, from which the upper limit is subsequently derived (Chapters 7 and 9). This section summarises the key hardware differences with respect to Belle II that are relevant to the present analysis.

Tracking

The Belle tracking system comprised the CDC and a silicon vertex detector (SVD) only, with no pixel detector layer. The SVD was upgraded during the lifetime of the experiment. The original three-layer detector (SVD1) was replaced in 2003 by a four-layer detector (SVD2), whose readout electronics were more radiation tolerant and which achieved better vertex resolution and improved particle identification. At the same time the beryllium beam pipe was replaced, its inner radius decreasing from 20 mm to 15 mm, which allowed the innermost SVD layer to move from a radius of 30 mm to 20 mm. Both configurations are used in this analysis, and neither reaches as close to the interaction point as the Belle II PXD at 14 mm.

Particle Identification

The Belle PID system for hadrons consisted of an aerogel Cherenkov counter (ACC) and a time-of-flight (TOF) detector, replacing the TOP and ARICH systems used in Belle II. The likelihood-based selectors built from these detectors are described in Section 3.4.4. The TOF measured the flight time T of a particle travelling a distance $L \approx 1.2 \text{ m}$ from the IP, allowing the mass to be inferred from the known momentum p as:

$$m = \frac{p}{c} \sqrt{\frac{c^2 T^2}{L^2} - 1}. \quad (3.5)$$

It comprised plastic scintillator modules read out by photomultiplier tubes (PMTs) with a timing resolution of approximately 100 ps, providing K/π separation for momenta below $\sim 1.2 \text{ GeV}/c$. A dedicated thin-trigger scintillation counter in front of each TOF module also provided fast-timing signals to the DAQ.

The ACC extended K/π separation to momenta between 1.2 and $3.5 \text{ GeV}/c$. Unlike the ARICH, which reconstructs the Cherenkov ring to measure the emission angle, the ACC operated as a threshold counter, where Cherenkov photons are emitted only when the particle momentum exceeds the threshold

$$p_{\text{th}} = \frac{mc}{\sqrt{n^2 - 1}}, \quad (3.6)$$

where n is the aerogel refractive index and m the particle mass, providing a binary above/below-threshold discrimination.

Electromagnetic Calorimeter

The Belle ECL used CsI(Tl) crystals housed in a support structure that was subsequently reused by Belle II. The principal difference lies in the readout electronics, where Belle II introduced waveform sampling to separate the signal from pile-up noise arising from the higher beam backgrounds at SuperKEKB.

K_L^0 and Muon Detector

The Belle KLM used resistive plate chambers (RPCs) throughout both the barrel and endcap regions. In Belle II, the endcap RPCs were replaced with scintillator strips read out via WLS fibres coupled to SiPMs, and the inner two barrel layers were similarly upgraded, to tolerate the higher hit rates and overcome the long dead times of RPCs under higher beam backgrounds at SuperKEKB.

3.4 Event Reconstruction

The role of event reconstruction is to transform raw electronic signals from individual subdetectors into calibrated, high-level physics variables. The initial stage transforms the raw electronic signals into low-level objects (track hits and digits). From these inputs, reconstruction algorithms build high-level variables: charged particle tracks are fitted from space points in the PXD, SVD, and CDC; energy clusters are reconstructed from energy deposits in the ECL; and unified PID variables are calculated from combined subsystem responses. The following subsections outline the software framework, tracking, and particle identification strategies used to prepare the data for physics analysis.

3.4.1 The Belle II Analysis Software Framework

The processing of both experimental data and simulated Monte Carlo (MC) samples is managed by the Belle II Analysis Software Framework (**basf2**) [70], developed by the Belle II Collaboration. This framework is used across all stages of the experiment, from online data acquisition, high-level trigger filtering, offline event reconstruction, MC generation, and high-level physics analysis. The foundation of **basf2** is written in C++, while high-level configuration and analysis workflows are steered via Python scripts, utilising a modular design. Individual processing tasks, such as tracking, clustering, or particle combination, are encapsulated into self-contained **Modules**. These modules are appended to a **Path** for sequential processing.

For high-level physics analysis, data processing goes from detector-oriented objects (tracks and clusters) to particle-oriented objects. Reconstructed tracks and calorimeter clusters, together with particle identification information, are stored in a compact data format known as the mini-Data Summary Tape (mDST), which serves as the standard input to high-level physics analysis in **basf2**. The **basf2** populates candidate lists for stable particles ($e^\pm$, $\mu^\pm$, $\pi^\pm$, $K^\pm$) based on the reconstructed tracking and PID inputs. High-level analysis modules then combine these basic building blocks to reconstruct short-lived mother particle candidates (such as $D^+ \rightarrow \pi^+ e^+ e^-$), evaluating decay kinematics and vertex physics.

In simulated MC data, this reconstruction chain is augmented with a truth-matching algorithm, which establishes relations between the reconstructed and true MC-generated particles [71]. This relationship provides a means to verify whether a reconstructed candidate represents a true signal event or a random combinatorial background, enabling the study of detector effects, analysis backgrounds, and the

calculation of selection efficiencies.

The Belle-to-Belle II (B2BII)

The Belle dataset is analysed using the Belle-to-Belle II (B2BII) conversion package [72], which translates Belle mDST files from the PANTHER [73] format into ROOT [74] objects compatible with the `basf2`. This allows all Belle II analysis tools available in `basf2` to be applied to both datasets.

3.4.2 Track Reconstruction

Charged track reconstruction, described in detail in Reference [64], is the process of estimating the kinematic properties of charged particles (momentum and position of closest approach to the IP). This task is implemented as a two-stage procedure consisting of track finding (pattern recognition) and track fitting.

Track finding begins in the CDC, where hit patterns are used to form initial track candidates. These candidates are fitted to CDC hits alone using the deterministic annealing filter (DAF) implemented in the GENFIT2 framework [75]. The fitted CDC tracks are then extrapolated inward to the SVD using a combinatorial Kalman filter (CKF), which searches for compatible SVD clusters. The combined CDC+SVD track candidates are subsequently refitted with the DAF. Finally, the tracks are further extrapolated inward to the PXD using an additional CKF step to associate pixel hits.

Track fitting determines the five parameters of the helicoidal trajectory of the particle in the solenoidal magnetic field, from which the momentum and the position of closest approach to the interaction point are extracted. Reconstructed tracks are additionally extrapolated outward to the outer subdetectors to geometrically match them with PID detector signals.

3.4.3 Charged Particle Identification

Reliable charged particle identification (PID) is essential for reconstructing hadronic final states and suppressing backgrounds from misidentified particles. In this analysis, robust separation of pions from kaons and electrons from hadrons is required to isolate the $D^+ \rightarrow \pi^+ e^+ e^-$ signal and reject combinatorial and misidentified backgrounds.

A simple approach to PID is to apply selection criteria directly on individual sub-detector observables, such as E/p for electrons or the Cherenkov angle for hadrons. However, a more powerful approach combines information from multiple subdetectors into a high-level likelihood-based discriminant, which is also more practical for use in the analysis. The likelihood $\mathcal{L}_\alpha$ represents the probability of observing the reconstructed PID information under the hypothesis that the particle is of type α ; equivalently, it quantifies how well the measured track properties are consistent with a true α particle.

In Belle II, PID is performed for the long-lived charged particle hypotheses $\alpha \in \{e, \mu, \pi, K, p, d\}$, where d denotes the deuteron. Charged hadrons are identified using the specific ionisation loss dE/dx measured in the SVD and CDC, supplemented by the Cherenkov photon timing from the TOP and the Cherenkov ring radius from the ARICH for K/π separation. While electron and muon identification rely primarily on the ECL and KLM, respectively, both lepton selectors incorporate complementary information from the SVD, CDC, TOP, and ARICH.

Each subdetector independently provides a likelihood $\mathcal{L}_\alpha^{\text{det}}$ for each hypothesis. Under the assumption that the subdetector responses are statistically independent, these are combined into a global likelihood by taking their product over all contributing subdetectors:

$$\mathcal{L}_\alpha = \prod_{\text{det}} \mathcal{L}_\alpha^{\text{det}}. \quad (3.7)$$

If a subdetector cannot provide a likelihood for a given track, for example because a low-momentum particle did not reach it, its contribution defaults to unity for all hypotheses ($\mathcal{L}_\alpha^{\text{det}} = 1$) and does not affect the combined likelihood. Individual subdetectors can also be excluded from the product when their calibration is unreliable in a given data-taking period, at the cost of some discriminating power.

The global PID variable for hypothesis α is then defined as the normalised likelihood ratio:

$$\text{PID}_\alpha = \frac{\mathcal{L}_\alpha}{\sum_i \mathcal{L}_i}, \quad (3.8)$$

where the sum runs over all six hypotheses $i \in \{e, \mu, \pi, K, p, d\}$. This variable takes values between 0 and 1, with values close to unity indicating strong consistency with hypothesis α .

To enhance π/K separation performance, modern Belle II hadron identification employs a machine-learning approach. A fully connected neural network (NN) is used to construct a high-level PID variable, πID_{NN} . This classifier combines the 36 individual subdetector likelihoods (derived from the six particle hypotheses across the six detector systems) together with the track's momentum, polar and azimuthal angles, and charge. Although other machine-learning-based quantities exist for other particle hypotheses, this multivariate approach for pions is the only one utilised in this analysis.

PID variables depend on the track momentum, polar angle, and charge, as the subdetector responses vary across the phase space, therefore the PID performance is evaluated in bins of (p, θ, q) using dedicated, highly pure data control samples (for example $D^{*+} \rightarrow \pi^+ D^0 (\rightarrow K^- \pi^+)$). The performance is quantified by two distinct metrics: the PID efficiency and misidentification rate ($\text{misID}_{\beta \rightarrow \alpha}$). The ϵ_α efficiency for hypothesis α is defined as the fraction of true α particles that are correctly identified after applying a given PID requirement:

$$\epsilon_\alpha = \frac{\text{number of } \alpha \text{ tracks identified as } \alpha}{\text{number of } \alpha \text{ tracks}}. \quad (3.9)$$

The misidentification rate is the fraction of background track species β ($\beta \neq \alpha$) incorrectly identified as α by satisfying the same requirement:

$$\text{misID}_{\beta \rightarrow \alpha} = \frac{\text{number of } \beta \text{ tracks identified as } \alpha}{\text{number of } \beta \text{ tracks}}. \quad (3.10)$$

The fake rate (f_α) represents the sample-dependent fraction of non- α tracks contaminating the final selected sample (equivalent to $1 - \text{purity}$):

$$f_\alpha = \frac{\sum_{\beta \neq \alpha} N(\beta \rightarrow \alpha)}{N_{\text{selected}}}, \quad (3.11)$$

where the denominator is the total number of all tracks satisfying the α identification criteria. The specific PID working points chosen for this analysis, along with the data-driven efficiency correction weights applied to account for simulation discrepancies, are detailed in Chapters 5 and 8, respectively.

3.4.4 Charged Particle Identification in Belle

In the Belle experiment, hadron PID was computed as a binary likelihood ratio between a candidate hypothesis α and an alternative hypothesis β , using dE/dx from the CDC, time-of-flight from the IP to the TOF, and the number of detected Cherenkov photons from the ACC:

$$\mathcal{R}(\alpha, \beta) = \frac{\mathcal{L}_\alpha^{\text{CDC}} \mathcal{L}_\alpha^{\text{TOF}} \mathcal{L}_\alpha^{\text{ACC}}}{\mathcal{L}_\alpha^{\text{CDC}} \mathcal{L}_\alpha^{\text{TOF}} \mathcal{L}_\alpha^{\text{ACC}} + \mathcal{L}_\beta^{\text{CDC}} \mathcal{L}_\beta^{\text{TOF}} \mathcal{L}_\beta^{\text{ACC}}}. \quad (3.12)$$

If a subdetector provided no information for a given track, its likelihood value defaulted to 0.5, contributing no discriminating power. Kaon and pion identification were then defined as $\mathcal{R}(K, \pi)$ and $\mathcal{R}(\pi, K) = 1 - \mathcal{R}(K, \pi)$.

Electron identification in Belle combined likelihoods from the CDC (dE/dx), ACC (light yield), TOF, and ECL (energy-momentum ratio E/p and transverse shower shape):

$$\mathcal{R}(e, \pi) = \frac{\prod_i \mathcal{L}_i^e}{\prod_i \mathcal{L}_i^e + \prod_i \mathcal{L}_i^\pi}, \quad (3.13)$$

where i runs over the CDC, ACC, TOF, and ECL.

4. Data Samples

This chapter describes the experimental data samples and the simulated Monte Carlo (MC) samples used throughout the analysis, including the event generators and decay models employed for signal, normalisation, and validation channels.

We develop the analysis procedure using simulated MC samples before applying it to data, optimising event selection, training the multivariate classifier, and estimating signal efficiencies entirely on MC. The signal region is defined by $1.70 \text{ GeV}/c^2 < M(\pi^+e^+e^-) < 1.92 \text{ GeV}/c^2$ and $0.2 \text{ GeV}/c^2 < M(e^+e^-) < 1.75 \text{ GeV}/c^2$, excluding the ϕ and η resonance regions defined in Section 5.5.3. These data were not examined until the selection and fit procedure had been fixed. The normalisation and validation channels, in contrast, are abundant, well-established decay modes with negligible signal contribution from the rare process under study; these were examined in data from the beginning, as they are used to normalise the signal branching fraction and validate the fit procedure rather than to probe the signal itself. Fixing the selection before examining the signal region reduces experimental bias, ensuring that selection criteria are not tuned to favour a particular result.

4.1 Experimental Data Samples

We use the full datasets collected by both the Belle and Belle II experiments at all available centre-of-mass energies, which are described in detail in Chapter 3. As discussed in Section 3.1, D^+ mesons used in this analysis predominantly originate from direct $c\bar{c}$ continuum production rather than from B -meson decays, and are also produced at energies below and above the $\Upsilon(4S)$ resonance. The full datasets are therefore used irrespective of centre-of-mass energy.

Both datasets are processed using **basf2**; the Belle dataset is converted into a **basf2**-compatible format using the **B2BII** package, as described in Section 3.4.1. Both samples used in this analysis were processed with the **light-2305-korat** release of **basf2**.

Data-Taking Structure

In both experiments, data-taking periods are organised into experiments and runs, each assigned incrementing integer numbers [76]. An experiment corresponds to a longer period of consistent detector conditions, typically of order months, and its number is incremented following a major detector or accelerator change or after a shutdown. A run is a shorter data-taking unit, incremented on demand by the data acquisition system, and corresponds to up to eight hours of data taking.

4.1.1 Belle II Dataset

The Belle II dataset corresponds to an integrated luminosity of 428 fb^{-1} [55], comprising 365 fb^{-1} collected at the $\Upsilon(4S)$ resonance, 43 fb^{-1} recorded 60 MeV below

it, and 20 fb^{-1} from an energy scan performed at four points spanning the centre-of-mass energy range 10.657 GeV to 10.810 GeV .

This unprecedented integrated luminosity produces immense volumes of raw data. For an individual analysis, it is computationally unfeasible to process every recorded event, as doing so would demand extremely long processing time and storage resources. To resolve this, the collaboration provides centrally produced, loosely preselected datasets—referred to as skims. These skims apply loose topological and kinematic criteria to reduce the file size to a manageable size while maintaining a near-perfect signal retention efficiency for the decay topologies of interest. The skim selection criteria for Belle II are summarised below. We require all final-state track candidates (electrons, kaons, and pions) to originate close to the nominal interaction point, with impact parameters in the transverse (dr) and longitudinal (dz) planes such that $dr < 1.0\text{ cm}$ and $|dz| < 3.0\text{ cm}$. Furthermore, the trajectories of these final-state particles must fall well within the geometric acceptance of the CDC, restricting the track polar angle. Beyond spatial restrictions, very loose particle identification requirements are applied at the skim stage to suppress combinatorial background:

- Electron candidates ($e^\pm$) are required to satisfy a global electron likelihood ratio computed without SVD or TOP information, $\text{PID}_e^{\text{noSVD,noTOP}} > 0.01$.
- Kaon candidates ($K^\pm$) are required to satisfy a global kaon likelihood ratio, $\text{PID}_K > 0.1$.
- Pion candidates are subject only to the track-quality and acceptance requirements stated above.

Finally, we require reconstructed D^+ candidates to satisfy a loose invariant-mass window, $1.67\text{ GeV}/c^2 < M(\pi^+ e^+ e^-) < 2.17\text{ GeV}/c^2$, and a centre-of-mass momentum requirement, $p_{\text{CMS}} > 2.0\text{ GeV}/c$.

4.1.2 Belle Dataset

The Belle dataset corresponds to an integrated luminosity of 980 fb^{-1} [77], collected at several centre-of-mass energies: 711 fb^{-1} at the $\Upsilon(4S)$ resonance, 121 fb^{-1} at the $\Upsilon(5S)$ resonance, 89 fb^{-1} recorded 60 MeV below the $\Upsilon(4S)$ resonance (the off-resonance sample), 28 fb^{-1} from energy scans above the $\Upsilon(4S)$ resonance, and the remainder collected at or near the $\Upsilon(1S, 2S, 3S)$ resonances.

We select Belle events using the HadronB(J) skim, processed with the latest available Belle software library (Case B), before conversion into `basf2` via B2BII. The HadronB(J) skim applies the following event-level requirements:

- at least three charged tracks;
- tracks with $p_t > 0.1\text{ GeV}/c$, $dr < 2.0\text{ cm}$, and $|dz| < 4.0\text{ cm}$;
- total visible energy of the event, defined as the sum of energies of charged tracks and reconstructed photons, satisfying $E_{\text{vis}}^* > 0.2\sqrt{s}$;
- primary event vertex with radial (V_r) and longitudinal (V_z) components satisfying $V_r < 1.5\text{ cm}$ and $|V_z| < 3.5\text{ cm}$, when the vertex quality exceeds two;
- at least two ECL clusters found within $-0.7 < \cos\theta < 0.9$;

- either $E_{\text{sum}}^* > 0.18\sqrt{s}$, $HJM > 1.8 \text{ GeV}/c^2$, or $HJM/E_{\text{vis}}^* > 0.25$, together with $E_{\text{sum}}^* < 0.8\sqrt{s}$. Here, E_{sum}^* is the total ECL cluster energy and HJM denotes the heavy jet mass, defined as the invariant mass of particles found in the hemisphere perpendicular to the event thrust axis.

4.2 Monte Carlo Simulation

MC simulation is the primary tool for developing and validating the analysis procedure before examining collision data in the signal region. In the high-energy physics community, the Monte Carlo method is a computational algorithm relying on repeated pseudo-random sampling, and the resulting simulated datasets are universally referred to simply as MC or MC samples.

These MC samples are produced using event generators: software programs that generate sets of four-momenta of final-state particles, reproducing e^+e^- collision outcomes according to theoretical models or prior experimental measurements of the relevant branching fractions and kinematic distributions. Where no measurement exists, generator inputs are based on theoretical predictions or best available estimates. The generated particles are then passed through a detailed simulation of the detector response, producing output in the same format as recorded data. This allows the full reconstruction chain to be applied identically to both simulation and real data.

In this analysis, we categorise MC samples used into two types: generic MC and signal MC. Generic MC samples are generated according to the known decay processes and branching fractions of all dominant e^+e^- collision products. They are used throughout the analysis for background modelling and for validating the agreement between simulation and data. The generic MC sample does not include the signal decay $D^+ \rightarrow \pi^+ e^+ e^-$, since this process is unmeasured and its branching fraction is unknown. The normalisation and validation channels are present within the generic MC sample, as these decays have been previously measured.

In contrast, signal MC samples are dedicated simulations of the specific decay channels studied in this analysis — the signal, normalisation, and validation modes. They are generated separately either because their branching fractions are too small for adequate representation in the generic MC sample or because dedicated statistics and control are needed for detailed studies. These dedicated samples are essential for studying signal and background shapes, estimating reconstruction efficiencies, and training the multivariate classifier described in Section 5.8.

For the Belle dataset, the MC samples are typically produced in larger, experiment-wide blocks (experiment-dependent MC), reflecting the comparatively stable detector conditions maintained during KEKB operation. Key parameters, such as the beam energy, the interaction-point profile, and the detector configuration, are averaged over all runs within a given experiment. Beam-induced background, obtained from randomly triggered data, is similarly averaged over the entire experiment and overlaid onto the simulated events. In contrast, Belle II operates in a higher-luminosity SuperKEKB environment, where beam-background conditions and calibration factors evolve more rapidly, requiring a closer reproduction of the per-run detector state. MC samples are therefore produced at the finer granularity of individual runs (run-dependent MC), with the beam background overlay extracted directly from randomly triggered data taken during the corresponding run.

4.2.1 Event Generators and Detector Simulation

We generate $B\bar{B}$ events using the EVTGEN software package [78], which handles the decay of heavy hadrons such as B and D mesons using experimentally measured branching fractions where available, and delegates unmeasured or poorly known decays to PYTHIA8.

For continuum events ($e^+e^- \rightarrow q\bar{q}$, $q = u, d, s, c$), including the signal process, we use PYTHIA6 [79] for Belle, and the KKMC generator [80] together with PYTHIA8 [81] for Belle II. In the Belle II simulation chain, KKMC generates the initial e^+e^- annihilation into a virtual photon and the subsequent splitting into a quark-antiquark pair, while PYTHIA8 simulates the hadronisation process that produces the observable final-state hadrons.

Electromagnetic final-state radiation from charged particles is simulated using the PHOTOS package [82] for all charged final-state particles. The detector response is simulated using GEANT3 [83] in Belle and GEANT4 [84] in Belle II. The combination of simulated $B\bar{B}$ and inclusive $q\bar{q}$ events constitutes the generic MC sample.

4.2.2 Generic MC Samples

For Belle II, we utilise four statistically independent replicas of the run-dependent generic MC sample from the 15th official MC production campaign (MC15rd). These samples are produced centrally at centre-of-mass energies corresponding to the $\Upsilon(4S)$, the $\Upsilon(4S)$ off-resonance, and $\Upsilon(5S)$ scan energies, with a total effective integrated luminosity of 1690 fb^{-1} , approximately four times that of the data. This sample includes $u\bar{u}$, $d\bar{d}$, $s\bar{s}$, $c\bar{c}$, $B^0\bar{B}^0$, and B^+B^- production. Using larger MC samples reduces the statistical uncertainty on background estimates and allows small contributions from rare processes to be identified and modelled more reliably in the simulated sample.

For Belle, a single replica was found to be sufficient based on experience gained from the Belle II dataset, which was processed first. This sample includes $u\bar{u}$, $d\bar{d}$, $s\bar{s}$, $c\bar{c}$, $B^0\bar{B}^0$, B^+B^- , $B_s\bar{B}_s$, and non- $B_s\bar{B}_s$ $b\bar{b}$ production and is produced centrally at all centre-of-mass energies recorded by the experiment. Due to missing generic MC files for experiment 73 within the collaboration's database, the effective integrated luminosity of the Belle generic MC sample is 958 fb^{-1} , slightly below the 980 fb^{-1} of recorded data. Because the generic MC is scaled to the data luminosity wherever it is used for background modelling, this affects only the statistical precision of the background estimate.

4.2.3 Signal, Normalisation, and Validation MC Samples

The kinematic properties of the signal decay $D^+ \rightarrow \pi^+ e^+ e^-$ are described in terms of the invariant mass squared of the dielectron system, defined via the four-momenta of the final-state particles as

$$q^2 \equiv (p_{D^+} - p_{\pi^+})^2 = (p_{e^+} + p_{e^-})^2 = M^2(e^+e^-), \quad (4.1)$$

where the equivalence follows from four-momentum conservation. While q^2 is the natural variable in the effective-field-theory description of the decay introduced in Chapter 2, we adopt the equivalent variable $M(e^+e^-) = \sqrt{q^2}$ throughout the remainder of this thesis.

We generate signal MC samples assuming a uniform three-body phase-space distribution (PHSP model), in which the decay products are distributed uniformly

across the kinematically allowed region. The true distribution is unknown, and because the reconstruction efficiency varies across the Dalitz plane, the choice of decay model affects the efficiency and hence the extracted branching fraction. We reduce this dependence by evaluating the efficiency in intervals of $M(e^+e^-)$ rather than as a single integrated value, as described in Section 6.1.2, so that the result depends on the assumed model only through the distribution within each interval. A residual dependence remains through the distribution of events within each interval, for which we assign a systematic uncertainty derived from the efficiency difference between the phase-space and resonant samples, as described in Section 8.2.5.

The same phase-space assumption underlies the previous searches summarised in Table 2.2. Figure 4.1 illustrates the uniform phase-space distribution of the signal decay. Unless otherwise specified, electromagnetic final-state radiation for all decay channels is simulated using the PHOTOS package. From this point forward, the inclusion of PHOTOS is assumed for all generated samples.

We generate the validation channel, $D^+ \rightarrow \pi^+\phi(\rightarrow e^+e^-)$, using the scalar-to-vector-scalar (SVS) model for the $D^+ \rightarrow \pi^+\phi$ decay and the vector-to-lepton-lepton (VLL) model for $\phi \rightarrow e^+e^-$, in both Belle and Belle II.

For the normalisation channel, $D^+ \rightarrow \pi^+K^+K^-$, the decay model differs between the two experiments. In Belle II, we generate the normalisation MC sample according to a Dalitz decay model (DDALITZ), which incorporates the ϕ resonance, non-resonant $D^+ \rightarrow \pi^+K^+K^-$ amplitudes, and their interference. In Belle, the Dalitz model is not implemented in the simulation framework; the normalisation channel is therefore generated in two stages via the ϕ resonance using the SVS model for $D^+ \rightarrow \pi^+\phi$ and the vector-to-scalar-scalar (VSS) model for $\phi \rightarrow K^+K^-$, restricted to the resonance region. For Belle II, in addition to the Dalitz sample, we also generate a sample transitioning via the resonance using the SVS and VSS decay models, analogous to the Belle sample, to allow a direct comparison between the two modelling approaches. Because the Belle resonance-only sample does not account for the non-resonant contribution to the inclusive $D^+ \rightarrow \pi^+K^+K^-$ branching fraction, a dedicated correction factor, derived from this Belle II comparison, is applied when determining the Belle normalisation efficiency. This derivation is described in Section 6.1.1.

We additionally generate a dedicated $D_{(s)}^+ \rightarrow \pi^+\eta(\rightarrow e^+e^-\gamma)$ MC sample, used solely for the background study associated with this decay, described in Section 5.9. Unlike the other signal MC samples, this sample is generated experiment-dependently in Belle and run-independently in Belle II.

To produce MC signal samples, we use the `mcprodzh` package, which generates events with EVTGEN, simulates the detector with `Gsim`, and reconstructs events in a single production chain, outputting directly in the Belle mDST format. For Belle II, signal MC samples are generated directly within `basf2` (release 06-01-10), following the same run-dependent production scheme as the generic MC sample.

The decay models and number of generated events for all signal MC samples are summarised in Tables 4.1 and 4.2 for Belle and Belle II, respectively. Small deviations (few-per cent level) from the target event counts arise because generated events are distributed among experiments and runs in proportion to their luminosities, which is the expected behaviour of the central production system. For the Belle signal mode, an error in the production script distributed the target of approximately two million events across both valid and non-existent experiment numbers, so that

Table 4.1: Summary of the signal MC samples generated for the Belle dataset. The table lists the specific decay channels, the total number of simulated events, and the kinematic decay models used.

Signal MC channel	No. of generated events	Decay model
$D^+ \rightarrow \pi^+ e^+ e^-$	1 700 818	PHSP
$D^+ \rightarrow \pi^+(\phi \rightarrow e^+ e^-)$	1 955 002	$D^+ \rightarrow \pi^+ \phi$ SVS $\phi \rightarrow e^+ e^-$ VLL
$D^+ \rightarrow \pi^+(\phi \rightarrow K^+ K^-)$	1 958 750	$D^+ \rightarrow \pi^+ \phi$ SVS $\phi \rightarrow K^+ K^-$ VSS
$D_{(s)} \rightarrow \pi^+ \eta(\rightarrow e^+ e^- \gamma)$	$\sim 1\text{M}$	PHSP

Table 4.2: Summary of the signal MC samples generated for the Belle II dataset. The table lists the specific decay channels, the total number of simulated events, and the kinematic decay models used.

Signal MC channel	No. of generated events	Decay model
$D^+ \rightarrow \pi^+ e^+ e^-$	1 989 837	PHSP
$D^+ \rightarrow \pi^+(\phi \rightarrow e^+ e^-)$	1 989 843	$D^+ \rightarrow \pi^+ \phi$ SVS $\phi \rightarrow e^+ e^-$ VLL
$D^+ \rightarrow \pi^+(\phi \rightarrow K^+ K^-)$	3 988 599	$D^+ \rightarrow \pi^+ \phi$ SVS $\phi \rightarrow K^+ K^-$ VSS
$D^+ \rightarrow \pi^+ K^+ K^-$	1 994 340	DDALITZ
$D_{(s)} \rightarrow \pi^+ \eta(\rightarrow e^+ e^- \gamma)$	$\sim 1\text{M}$	PHSP

1 700 818 events were retained. The surviving events remain distributed across experiments in proportion to their integrated luminosities, so the luminosity-weighted efficiency is unaffected, and only its statistical precision is reduced.

We reserve 1.58 % of the Belle generic MC and 3.88 % of the Belle signal MC for MVA training. For Belle II, we reserve 1.2 % of the generic MC and 3.2 % of the signal MC for this purpose. The remaining events in all samples are used for efficiency estimation, probability density function shape determination, fitting, etc.

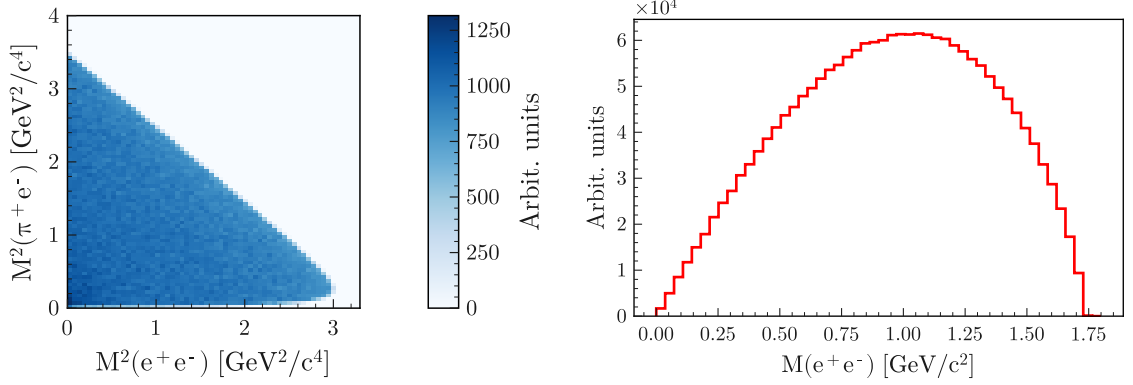


Figure 4.1: Two-dimensional kinematic distribution in Dalitz plane (left) and $M(e^+e^-)$ distribution (right) of the generated $D^+ \rightarrow \pi^+e^+e^-$ signal events at the generator-level information. The signal MC sample is simulated assuming a uniform three-body phase-space (PHSP) decay model, with electromagnetic final-state radiation modelled using the PHOTOS package.

Signal and generic MC samples are produced run- or experiment-dependently with the same centre-of-mass energy composition as the corresponding data. The resulting reconstruction efficiency therefore represents a luminosity-weighted average over the collected energies. No explicit correction for the mixture of centre-of-mass energies is therefore required.

5. Event Selection and Signal Reconstruction

Candidate selection in this analysis proceeds through four successive stages of increasing restriction. The *skim* stage, described in Chapter 4, applies a loose, centrally-produced event-level pre-selection to reduce the raw dataset to a manageable size. A *loose selection*, applied in the analysis steering script, imposes minimal track-quality and kinematic requirements sufficient to suppress the dominant combinatorial background while retaining a wide D^+ invariant-mass window. We use this stage to study variable distributions and guide the choice of subsequent cuts. The *initial selection* then applies the full set of kinematic, vertex, and particle-identification requirements summarised in Tables 5.10 and 5.11, except the multivariate analysis (MVA) requirement and the η -veto. Finally, the *final selection* adds the MVA requirement, specifically a random forest classifier (RFC) described in Section 5.8, and the η -veto (Section 5.9), yielding the sample used for signal extraction.

5.1 Track Selection and Particle Identification

We require all charged-particle tracks to originate close to the interaction point, with a distance of closest approach $|dr| < 0.5$ cm in the transverse plane and $|dz| < 2.0$ cm along the beam axis, and to fall within the geometric acceptance of the CDC. These requirements are applied uniformly at both the loose- and initial-selection stages, and are tighter than the corresponding skim-level selection (Section 4.1).

Unless otherwise noted, SVD and TOP information is excluded from the Belle II electron-identification likelihood ratio, $\text{PID}_e^{\text{noSVD, noTOP}}$. For brevity, we hereafter denote this restricted likelihood ratio as PID_e .

At the loose-selection stage, we apply a global electron PID requirement $\text{PID}_e > 0.5$ (Belle II) or a binary PID requirement $\mathcal{R}(e, \pi) > 0.6$ (Belle), to reduce the dominant background while retaining a sample broad enough for the diagnostic studies of Section 5.4. No pion PID requirement is applied at this stage. For kaons, used in the reconstruction of the normalisation channel, we require a global kaon PID requirement $\text{PID}_K > 0.5$ in Belle II and a binary requirement $\mathcal{R}(K, \pi) > 0.5$ in Belle.

In the initial-selection stage, we tighten the electron PID requirement to $\text{PID}_e > 0.9$ (Belle II) or $\mathcal{R}(e, \pi) > 0.9$ (Belle). Pion candidates are selected using a neural-network-based identification score $\pi\text{ID}_{\text{NN}} > 0.40$ (Belle II), which also serves as a binary kaon-pion separator, or a dedicated binary kaon-pion separation requirement, $\mathcal{R}(\pi, K) > 0.5$ (Belle). We further tighten the kaon PID requirements to $\text{PID}_K > 0.6$ (Belle II) and $\mathcal{R}(K, \pi) > 0.6$ (Belle). The resulting PID efficiencies, evaluated on the electron and pion tracks of the signal MC sample (Section 4.2.3), and fake rates, evaluated on the corresponding tracks of the generic MC sample (Section 4.2), are

summarised in Table 5.1. For kaons, used to reconstruct the normalisation channel, efficiencies and fake rates are evaluated analogously using the normalisation-channel signal MC and generic MC samples, respectively.

Table 5.1: Particle identification efficiencies and fake rates at the tight (initial-selection) working points, for Belle and Belle II. Efficiencies are evaluated on signal MC; fake rates are evaluated on generic MC.

Particle	Efficiency		Fake rate	
	Belle	Belle II	Belle	Belle II
Electron	81 %	76 %	12 %	21 %
Pion	95 %	96 %	21 %	16 %
Kaon	88 %	80 %	3 %	1 %

5.2 D^+ Candidate Reconstruction

We reconstruct D^+ candidates¹ by combining three charged-particle tracks satisfying the identification criteria of Section 5.1, requiring a total charge of +1. At the loose-selection stage, candidates are retained within a wide invariant-mass window of $1.67 \text{ GeV}/c^2 < M(\pi^+ e^+ e^-) < 2.17 \text{ GeV}/c^2$, used both for the diagnostic studies described in Section 5.4 and for the sideband studies of Section 6.2.3. At the initial- and final-selection stages, this window is narrowed to $1.70 \text{ GeV}/c^2 < M(\pi^+ e^+ e^-) < 1.92 \text{ GeV}/c^2$, corresponding to a $220 \text{ MeV}/c^2$ window situated around the nominal D^+ mass, $m_{D^+} = (1869.66 \pm 0.05) \text{ MeV}/c^2$ [16]. The window is deliberately asymmetric, extending further below the nominal mass than above it, to retain signal candidates in the low-mass tail left by unrecovered bremsstrahlung (Section 5.3). To suppress D^+ candidates originating from B -meson decays, we require the D^+ candidate momentum in the CMS frame to satisfy $p_{\text{CMS}} > 2.5 \text{ GeV}/c$, tighter than the corresponding $2.0 \text{ GeV}/c$ requirement already applied at the skim stage (Section 4.1). We perform a vertex fit using the TreeFitter algorithm [85], a kinematic fit that treats the full decay chain of the D^+ candidate simultaneously, rather than fitting each vertex independently. The algorithm uses a Kalman filter to solve a least-squares estimate of the decay-chain parameters, constraining each final-state particle to its nominal mass and propagating vertex and momentum information along the decay tree. This global approach allows the D^+ production vertex to be constrained to the interaction point within the same fit, improving the resolution of the fitted decay parameters. In the loose-selection stage, we require only that the fit converges, while in initial-selection, the fit probability (or p -value), $P(\chi^2)$, must satisfy $P(\chi^2) > 0.001$.

5.2.1 Normalisation Channel Reconstruction

We reconstruct the normalisation channel, $D^+ \rightarrow \pi^+ K^+ K^-$, using the same selection criteria and reconstruction procedure as the signal channel (Sections 5.1

¹Charge-conjugate decay modes are implied throughout this thesis.

and 5.2), with the dielectron pair replaced by an oppositely charged kaon pair. Kaon candidates are selected using the kaon identification requirements introduced in Section 5.1, $\text{PID}_K > 0.6$ (Belle II) or $\mathcal{R}(K, \pi) > 0.6$ (Belle). We combine an oppositely charged kaon pair with a charged pion to form the D^+ candidate, subject to the same D^+ mass window, p_{CMS} requirement, and vertex fit described in Section 5.2, and additionally require the kaon-pair invariant mass to satisfy $|M(K^+K^-) - m_\phi| < 20 \text{ MeV}/c^2$, where m_ϕ is $1019.46 \text{ MeV}/c^2$ [16]. Within this mass window, the sample is dominated by the $\phi(1020)$ resonance, but also receives non-resonant $D^+ \rightarrow \pi^+K^+K^-$ contributions and their interference with the resonant amplitude. The treatment of this non-resonant component in the efficiency determination is described in Section 6.1.1. The normalisation channel is subject to the same best-candidate selection and random forest classifier requirements as the signal channel (Sections 5.6 and 5.8); the η -veto (Section 5.9) does not apply, since the normalisation channel does not share the $e^+e^-\gamma$ final state targeted by that veto.

5.3 Bremsstrahlung Correction

As electrons traverse detector material, they lose energy through bremsstrahlung radiation. These emitted photons are collinear with the initial electron trajectory and are reconstructed as separate, neutral ECL clusters. Failing to account for this radiated energy shifts the reconstructed electron momentum towards lower values and degrades the $M(\pi^+e^+e^-)$ resolution. To mitigate this effect, we apply a bremsstrahlung correction algorithm (`correctBremsBelle`) identically across both the Belle and Belle II datasets. This algorithm identifies candidate bremsstrahlung photons within a 2.8° cone around the electron track direction and adds their four-momenta back to the electron candidate. Because beam-induced background contributes photon clusters non-uniformly across the ECL acceptance, region-dependent minimum energy thresholds are applied to candidate bremsstrahlung photons, requiring $E_\gamma > 0.075 \text{ GeV}$ in the forward endcap, 0.05 GeV in the barrel, and 0.1 GeV in the backward endcap. The bremsstrahlung correction affects 20.4 % of candidates in the Belle signal MC sample and 24 % in the Belle II signal MC sample. Figure 5.1 shows the impact of the correction on the $M(\pi^+e^+e^-)$ ² distribution in signal MC, for both experiments.

5.4 Loose Selection

Using the D^+ candidate reconstruction procedure and loose-selection working points introduced in Sections 5.1 and 5.2, we study the distributions of a broad set of candidate variables to identify an appropriate initial selection and to select the input variables used in the RFC (Section 5.8).

Among this broad set, the following observables exhibit the most distinct separation power between signal and background, from which a subset of these features is subsequently selected to train the RFC:

²Throughout this thesis we denote the invariant mass of the reconstructed D^+ candidate by the explicit final state it is built from: $M(\pi^+e^+e^-)$ for the signal and validation channels, and $M(\pi^+K^+K^-)$ for the normalisation channel. Several figures are reproduced directly from the analysis framework, where the same quantity is labelled $M(D^+)$ on the axes. The two notations denote the same reconstructed invariant mass and are used interchangeably in figure axis labels.

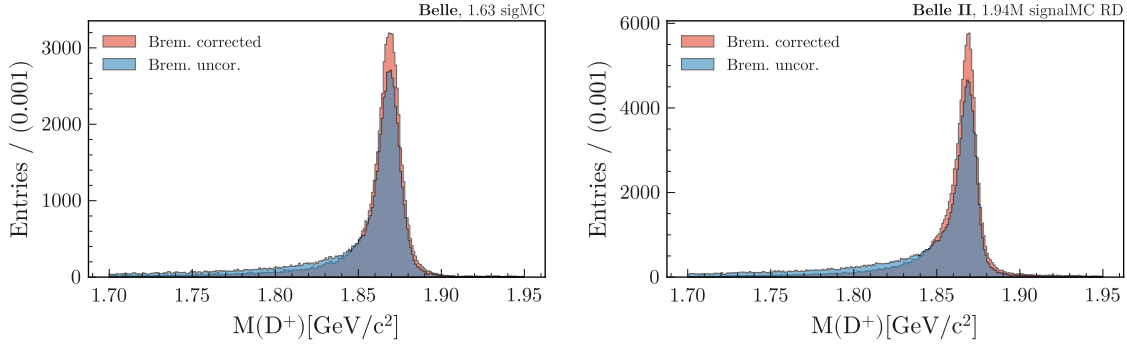


Figure 5.1: Impact of the bremsstrahlung correction on the $M(\pi^+e^+e^-)$ distribution in signal MC, for Belle (left) and Belle II (right).

- Invariant masses ($M(\pi^+e^+e^-)$ and $M(e^+e^-)$):** As expected from the absence of the signal decay in the generic MC sample (Section 4.2), the $M(\pi^+e^+e^-)$ distribution exhibits no peaking structure and is entirely dominated by combinatorial background. Even the validation channel $D^+ \rightarrow \pi^+\phi(\rightarrow e^+e^-)$, which is physically present in the generic MC sample, remains completely hidden by the background due to its small branching fraction. In contrast, the $M(e^+e^-)$ distribution reveals a large amount of background at values below $0.2 \text{ GeV}/c^2$, driven by photon conversions and π^0 decays.
- Significance of flight distance:** The distance between the fitted D^+ decay vertex and the interaction point, divided by its measurement uncertainty (see Section 5.8 for the precise definition used). The combinatorial background distribution is symmetrically centred around zero, whereas the true signal exhibits a long tail extending to large positive values, reflecting the macroscopic lifetime of the D^+ meson.
- Total visible centre-of-mass energy (`visibleEnergyOfEventCMS` or E_{CMS}^{vis}):** The total reconstructed energy of all final-state particles in the event, evaluated in the CMS frame. The signal distribution peaks at approximately the same value as the $c\bar{c}$ continuum background, while the $B\bar{B}$ background distribution is shifted towards significantly lower energies, providing discrimination against B -meson decays.
- Centre-of-mass momentum (p_{CMS}):** The D^+ candidate momentum in the CMS frame, slightly higher for signal than for generic background, aiding suppression of low-momentum combinatorial candidates.
- Pointing angle ($\cos\theta_{p,v}$):** The cosine of the angle between the D^+ candidate momentum vector and the vector connecting the interaction point to the fitted D^+ decay vertex. For true signal decays, the reconstructed momentum vector closely aligns with the flight direction, resulting in a distribution peaking near 1.
- Particle identification (PID_e and the pion neural-network score):** For correctly identified signal tracks, these scores peak near 1, while misidentified background tracks populate lower values, giving strong separation between the two.

5.5. $M(e^+e^-)$ SELECTION AND PHOTON-CONVERSION VETO

Figures 5.2 and 5.3 show the resulting distributions for Belle and Belle II, respectively, with the different generic MC background components shown as a stacked histogram and the signal MC component shown separately, unstacked. Each distribution is accompanied by the corresponding Punzi figure of merit (FOM),

$$\text{FOM}(t) = \frac{\epsilon(t)}{a/2 + \sqrt{B(t)}}, \quad (5.1)$$

evaluated as a function of the candidate cut value t on the plotted variable, where ϵ is the signal efficiency, B is the expected number of background events, and a is the target discovery significance in standard deviations. We adopt $a = 5$ throughout this analysis (see Section 5.10 for further discussion of this choice). The cut value maximising the FOM for each variable, indicated by a dashed vertical line, is not applied directly. Instead, these scans serve to define approximate initial-selection regions and identify variables with strong separation power, which are later combined in the RFC. The applied cuts are those listed in Tables 5.10 and 5.11. Additional distributions after loose selection are provided in Appendix B.1.

5.5 $M(e^+e^-)$ Selection and Photon-Conversion Veto

5.5.1 Photon-Conversion Veto

A photon converting into an e^+e^- pair ($\gamma \rightarrow e^+e^-$) in detector material can mimic the signal final state if one or both of the converted electrons are combined with an unrelated track to form a D^+ candidate. We suppress this background with a dedicated veto. For each signal candidate, we separately combine the positive and negative signal lepton with an oppositely charged track from the rest of the event (any reconstructed track not used in the D^+ candidate), forming two lists of photon-conversion candidates, one for each signal lepton. From both lists combined, we select the candidate with the smallest opening angle between the two combined tracks, and reject the D^+ candidate if the invariant mass of this photon candidate satisfies $M(\gamma \rightarrow e^+e^-) < 0.05 \text{ GeV}/c^2$. This threshold is selected using the Punzi FOM. Figure 5.4 shows the $M(\gamma \rightarrow e^+e^-)$ and $M(e^+e^-)$ distributions for generic MC background components and signal MC, together with the corresponding Punzi FOM.

Because a converted photon splits into a nearly collinear e^+e^- pair, both the opening angle and the invariant mass of the resulting photon candidate peak near zero for true conversions, whereas both quantities shift to larger values for signal candidates and for wrongly combined background tracks. This method is effective when only one of the two signal leptons originates from a photon conversion, since in that case the reconstructed $M(\gamma \rightarrow e^+e^-)$ peaks near zero. Figure 5.5 illustrates this behaviour, showing the $M(\gamma \rightarrow e^+e^-)$, opening-angle, and $M(e^+e^-)$ distributions, categorised by the number of selected signal leptons originating from a photon conversion (zero, one, or both), together with the true signal component. When both signal leptons originate from a conversion, the veto is ineffective; instead, such candidates populate the low-mass region of $M(e^+e^-)$ visible in the bottom panel and are removed by the requirement introduced in the following section.

CHAPTER 5. EVENT SELECTION AND SIGNAL RECONSTRUCTION

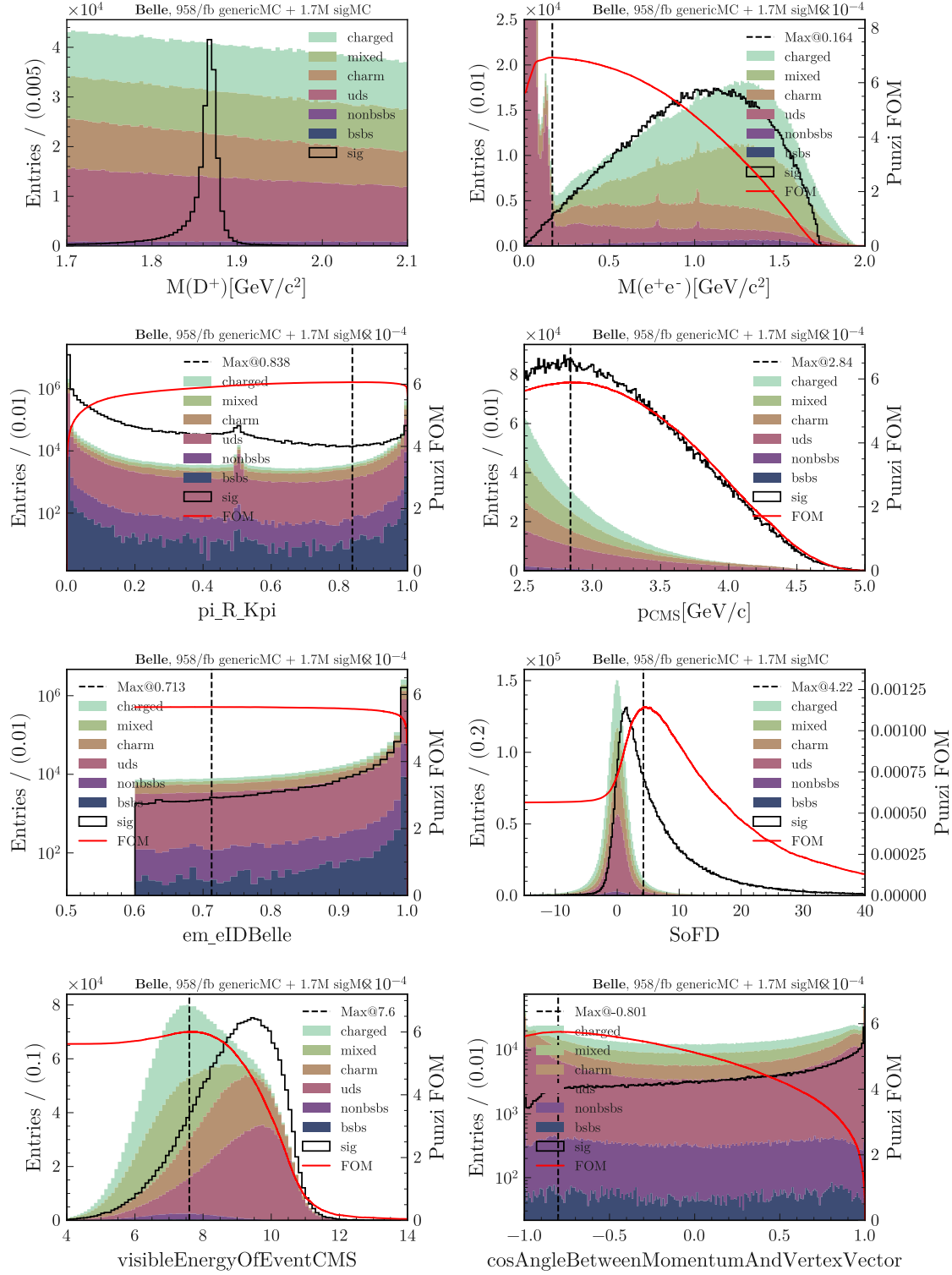


Figure 5.2: Distribution of selection variables after the loose selection, for Belle.

5.5. $M(e^+e^-)$ SELECTION AND PHOTON-CONVERSION VETO

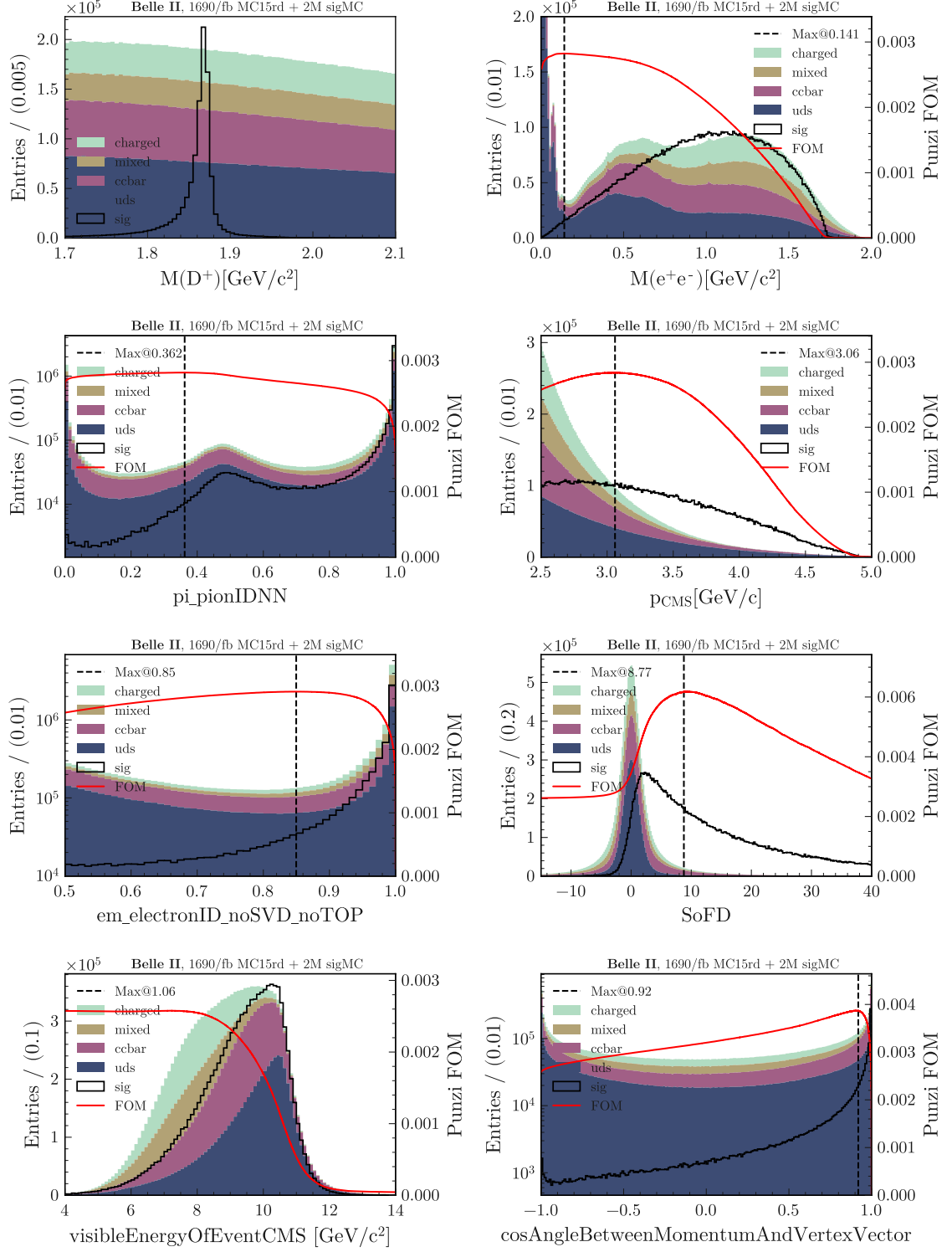


Figure 5.3: Distribution of selection variables after the loose selection, for Belle II.

CHAPTER 5. EVENT SELECTION AND SIGNAL RECONSTRUCTION

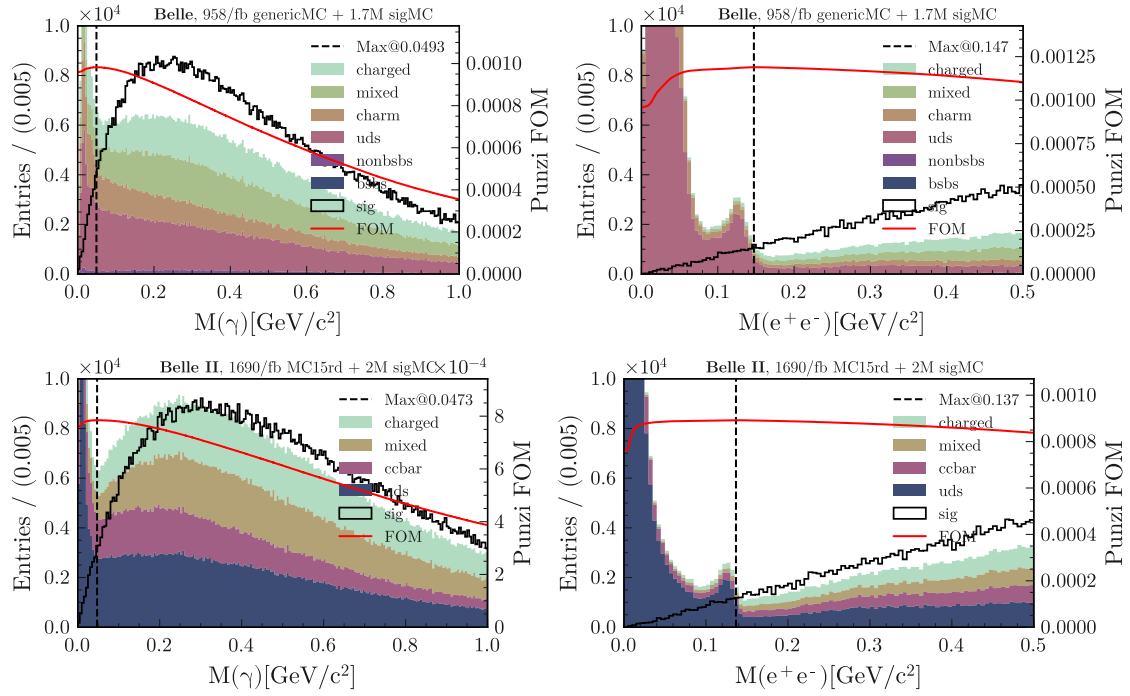


Figure 5.4: Distributions of $M(\gamma \rightarrow e^+e^-)$ and $M(e^+e^-)$ for generic MC background components and signal MC, with the corresponding Punzi FOM, for Belle (top) and Belle II (bottom).

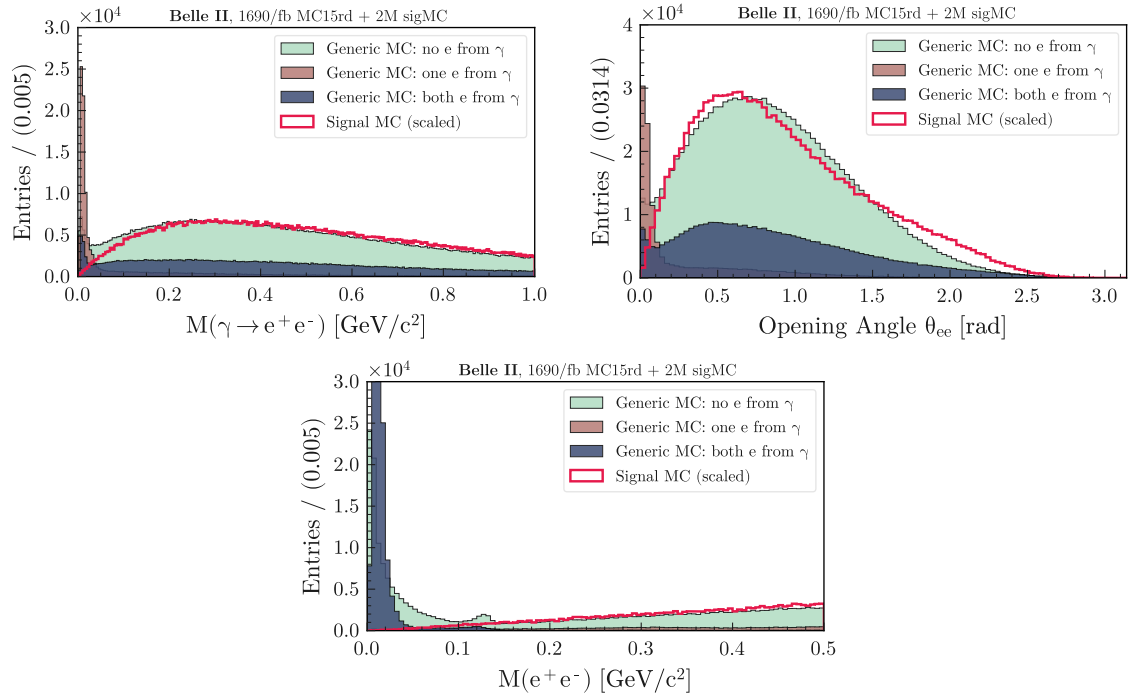


Figure 5.5: Distributions of $M(\gamma \rightarrow e^+e^-)$ (top left), the opening angle of the photon-conversion candidate (top right), and $M(e^+e^-)$ (bottom), categorised by the number of signal leptons originating from a photon conversion (zero, one, or both) and for true signal events.

5.5.2 Low-Mass Dielectron Requirement

To suppress background from low-mass dilepton sources, in particular from candidates in which both signal leptons originate from a photon conversion (Figure 5.5), we require $M(e^+e^-) > 0.2 \text{ GeV}/c^2$. At low dielectron invariant mass, the dominant background contributions arise from photon conversions and from $\pi^0 \rightarrow e^+e^-\gamma$ Dalitz decays, both of which are effectively removed by this requirement, consistent with the approach adopted in previous searches for rare charm decays [4]. We restrict the analysed phase space to an upper limit of $M(e^+e^-) < 1.75 \text{ GeV}/c^2$, chosen to fully include the kinematically allowed region for $D^+ \rightarrow \pi^+e^+e^-$ signal events, $m_{D^+} - m_{\pi^+} \approx 1.73 \text{ GeV}/c^2$.

5.5.3 Resonance Exclusion Windows

We additionally exclude two $M(e^+e^-)$ regions associated with intermediate resonances, reserving them for dedicated treatment elsewhere in the analysis.

We exclude the region $0.95 < M(e^+e^-) < 1.05 \text{ GeV}/c^2$, corresponding to the $\phi(1020)$ resonance. This region is reserved for the validation channel $D^+ \rightarrow \pi^+\phi(\rightarrow e^+e^-)$, reconstructed separately within the narrower window $0.995 < M(e^+e^-) < 1.035 \text{ GeV}/c^2$, as described in Section 5.11.

We further exclude the region $0.50 < M(e^+e^-) < 0.56 \text{ GeV}/c^2$, associated with the η resonance. This window addresses the peaking component of the background from $D_{(s)}^+ \rightarrow \pi^+(\eta \rightarrow \gamma e^+e^-)$ decays; the full mechanism behind this background, together with the dedicated veto used to suppress its non-peaking component, is described in Section 5.9.

5.6 Best-Candidate Selection

Because multiple track combinations satisfying the selection criteria of Sections 5.1 and 5.2 can occur within a single event, more than one D^+ candidate may be reconstructed per event. In cases of multiple candidates, we retain only the candidate with the highest vertex-fit probability $P(\chi^2)$. The average candidate multiplicity in events with at least one reconstructed candidate is 1.17 in signal MC and 1.46 in generic MC for Belle, and 1.17 in signal MC and 1.54 in generic MC for Belle II. This best-candidate selection retains 87.2 % of true signal events and 68.6 % of background events in Belle, and 90.9 % of true signal events and 66.5 % of background events in Belle II. The fraction of mis-reconstructed candidates relative to all reconstructed candidates in the signal MC sample (Section 4.2.3) — where a candidate does not correspond to the true generator-level decay, identified via the truth-matching procedure described in Section 3.4.1 — decreases from 30.5 % to 17.7 % in Belle, and from 19.8 % to 10.2 % in Belle II, upon applying this requirement. We quantify the related residual mis-reconstruction fraction in the fully selected sample separately as a systematic uncertainty in Chapter 8. Correspondingly, the Punzi FOM (Eq. 5.1) increases from 12.6×10^{-3} to 13.3×10^{-3} in Belle, and from 9.7×10^{-4} to 11×10^{-4} in Belle II. The absolute values are not comparable between experiments because the background yields are normalised to different effective MC luminosities; thus only the relative improvement within each experiment is meaningful here. Figure 5.6 shows the distribution of best-candidate rankings for both experiments.

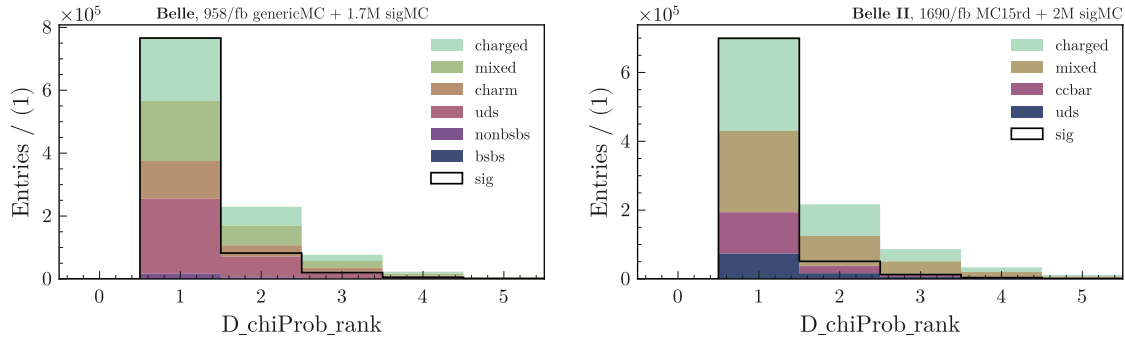


Figure 5.6: Ranking of reconstructed D^+ candidates by vertex-fit probability, for Belle (left) and Belle II (right).

5.7 Individual Selection Efficiency

Tables 5.2 and 5.3 show the effect of each initial-selection requirement, applied individually to the sample remaining after the loose selection (Section 5.4), rather than cumulatively. Retention rates therefore reflect each cut applied in isolation, not the fraction surviving all preceding cuts. Because Belle and Belle II employ different loose-selection working points (Section 5.1), the absolute retention rates at each subsequent cut are not directly comparable between the two experiments. Unlike the other rows, the best-candidate selection is applied after all preceding requirements, on the sample surviving the full initial selection.

Table 5.2: Retention rate for each initial-selection requirement in % for different Belle samples.

Cut	genericMC-bkg	genericMC-sig	signalMC-sig
$\mathcal{R}(e, \pi) > 0.9$ (e^- / e^+)	91.3 / 90.5	93.3 / 93.7	95.0 / 95.0
$\mathcal{R}(\pi, K) > 0.5$ (π)	78.8	92.5	94.3
$P(\chi^2) > 0.001$	57.8	86.6	90.3
$M(e^+e^-) < 1.75 \text{ GeV}/c^2$	98.6	100.0	100.0
$M(e^+e^-) > 0.2 \text{ GeV}/c^2$	64.0	53.0	97.9
$M(\gamma) > 0.05 \text{ GeV}/c^2$	87.8	95.5	96.5
ϕ exclusion	95.0	82.9	91.1
η exclusion	98.1	94.5	96.8
Best-candidate selection	69.0	90.0	87.2

5.8 Random Forest Classifier

A random forest is an ensemble learning method that combines the predictions of many individual decision trees and derives the final classification through majority consensus voting. Randomness is introduced at two levels during training. First,

5.8. RANDOM FOREST CLASSIFIER

Table 5.3: Retention rate for each initial-selection requirement in % for different Belle II samples.

Cut	genericMC-bkg	genericMC-sig	signalMC-sig
$\text{PID}_e > 0.9$ (both leptons)	55.9 / 53.1	86.5 / 85.3	84.3 / 84.4
$\pi\text{ID}_{\text{NN}} > 0.4$ (π)	77.9	95.7	96.3
$P(\chi^2) > 0.001$	41.1	88.8	89.1
$M(e^+e^-) < 1.75 \text{ GeV}/c^2$	98.7	100	100
$M(e^+e^-) > 0.20 \text{ GeV}/c^2$	80.7	96.9	97.9
$M(\gamma) > 0.05 \text{ GeV}/c^2$	86.1	96.7	95.5
ϕ exclusion	94.1	18.3	91.2
η exclusion	94.0	99.5	93.8
Best-candidate selection	66.5	90.5	90.9

each tree is trained on an independent bootstrap sample, obtained by sampling the training data with replacement. Second, at each split within a tree, only a randomly restricted subset of input variables is considered, rather than all available variables. Aggregating the output of many such decorrelated trees reduces the risk of overfitting relative to a single decision tree, while retaining the ability to model non-linear relationships and variable correlations [86].

We use a random forest classifier (RFC), implemented in the `scikit-learn` Python package [87], in the final stage of candidate selection. Neural-network and `XGBoost` classifiers were also tested but showed no improvement over the RFC. Event-shape variables available in `basf2` (e.g., Fox–Wolfram moment ratios, thrust) were similarly tested as additional input variables but did not improve classifier performance. We train and optimise two RFCs separately for Belle and Belle II. The RFC hyperparameters are left at their default `scikit-learn` values, except for the number of trees, $n_{\text{estimators}} = 50$, and the minimum number of samples per leaf, $\text{min_samples_leaf} = 50$. Candidates are accepted if $\text{RFC_score} > 0.97$ for Belle and $\text{RFC_score} > 0.98$ for Belle II, with the two working points optimised independently for each experiment.

5.8.1 Training

Each classifier is trained over the range $0.2 < M(e^+e^-) < 1.75 \text{ GeV}/c^2$, without excluding the η and ϕ resonance regions, so that a single classifier applies uniformly to the signal and validation channels. Performance in the three intervals used for the measurement is reported in Section 5.8.3, and the effect of optimising the working point separately in each interval is examined in Section 5.10. Following the generic and signal MC training fractions introduced in Section 4.2, the RFC training sample for each experiment corresponds to approximately 10 000 background and 10 000 signal events, while the remaining events are reserved for RFC performance evaluation, signal PDF shape determination, and the fit itself (Chapter 7). As shown

CHAPTER 5. EVENT SELECTION AND SIGNAL RECONSTRUCTION

by the learning curves in Figure 5.7, classifier performance plateaus once the training sample exceeds approximately 10 000 entries, justifying this choice of training sample size.

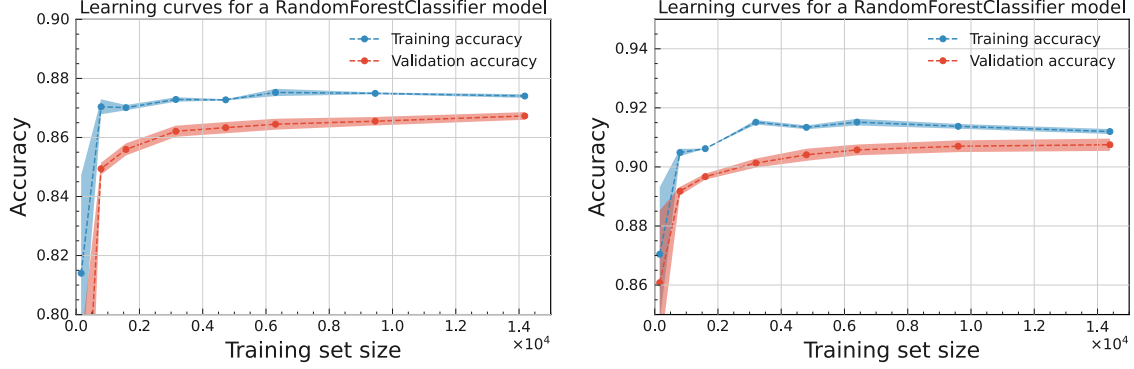


Figure 5.7: Learning curves for the training and validation samples, for Belle (left) and Belle II (right).

The RFC uses four input variables, defined and motivated in Section 5.4: the D^+ centre-of-mass momentum p_{CMS} , the pointing angle $\cos\theta_{\mathbf{p},\mathbf{v}}$, the total visible centre-of-mass energy $\text{visibleEnergyOfEventCMS}$, and the significance of flight distance. Distributions of these input variables after initial selection are shown in Figures 5.8 and 5.9 for Belle and Belle II, respectively; corresponding distributions in individual $M(e^+e^-)$ intervals are provided in Appendix B.2.

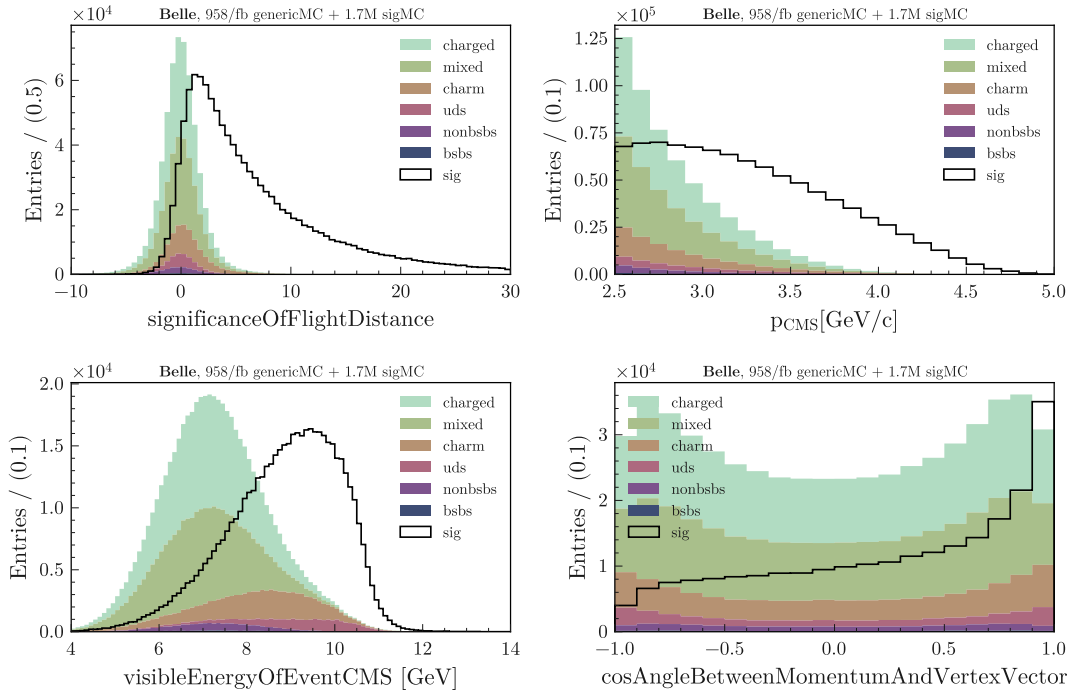


Figure 5.8: Distribution of RFC input variables after initial selection for Belle, without the pion identification requirement.

The pion identification requirement (Section 5.1) was introduced into the analysis chain after RFC training had already begun, but before the signal region in

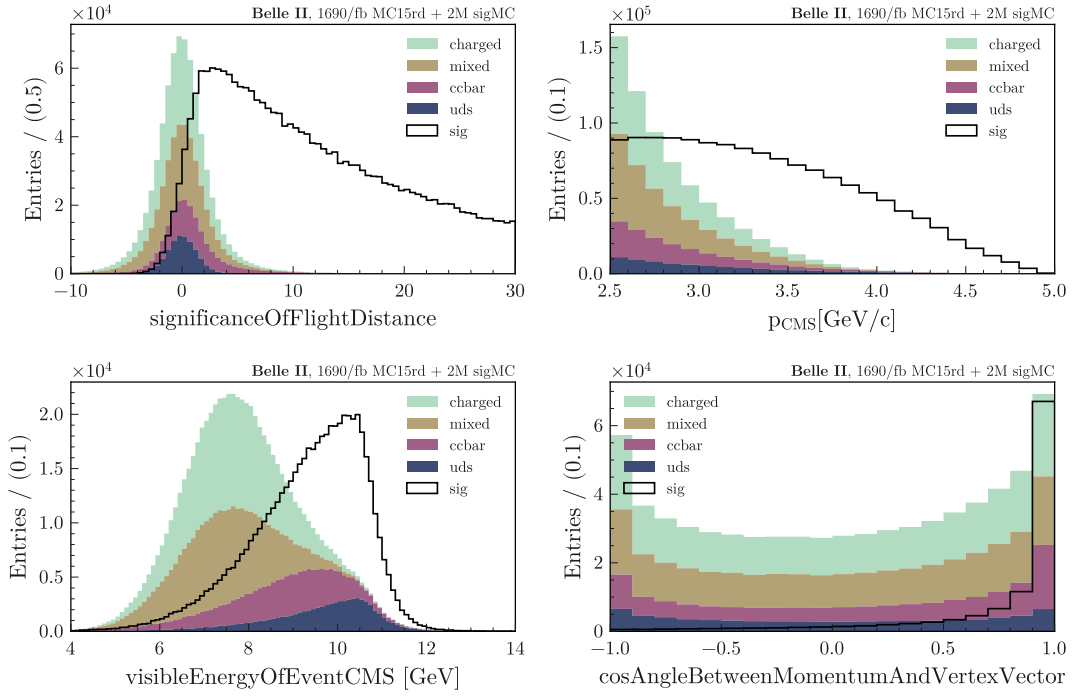


Figure 5.9: Distribution of RFC input variables after initial selection for Belle II, without the pion identification requirement.

data was examined, so the RFC training sample uses the initial selection with this single exception. We subsequently trained and evaluated the RFC on samples both with and without the pion identification requirement applied beforehand, and found marginally better performance without it. This comparison was made on simulation alone, before the signal region in data was examined. We therefore retain the RFC trained on the sample excluding this requirement, while the pion identification requirement itself continues to be applied independently as part of the full selection chain.

5.8.2 Performance

Table 5.4 summarises the RFC signal efficiency and background retention rate for each experiment, evaluated over the full $M(e^+e^-)$ region for signal mode. The performance in the normalisation and validation channels is discussed in Sections 6.2.2 and 6.2.1. Classifier performance is further quantified using the area under the receiver-operating-characteristic (ROC) curve (AUC), evaluated on the test sample via bootstrapping with 1000 repetitions and a sample size of 10 000, giving $\text{AUC} = 0.965 \pm 0.002$ for Belle II and $\text{AUC} = 0.936 \pm 0.003$ for Belle.

Table 5.4: RFC performance for Belle and Belle II, evaluated over the full $M(e^+e^-)$ region.

	Signal ϵ [%]	Bkg. retention [%]
Belle	36.8 ± 0.1	0.275 ± 0.012
Belle II	47.8 ± 0.1	0.292 ± 0.007

CHAPTER 5. EVENT SELECTION AND SIGNAL RECONSTRUCTION

To compare discrimination power directly between experiments, we evaluate the background retention rate of each experiment's RFC at the signal efficiency achieved by the other (Table 5.5). At matched signal efficiency, Belle II consistently achieves a lower background retention rate than Belle, indicating better signal-background discrimination.

Table 5.5: RFC performance for Belle and Belle II at matched signal efficiency, for the full $M(e^+e^-)$ region.

	Signal ϵ [%]	Bkg. retention [%]	RFC_score cut
Belle	47.8 ± 0.1	0.56 ± 0.01	0.9424
Belle II	36.8 ± 0.1	0.168 ± 0.005	0.9908

Figure 5.10 shows the RFC score distributions for signal and background events together with the corresponding Punzi FOM curve, while Figures 5.11 and 5.12 show the ROC curves and the true-positive and false-positive rates (TPR and FPR) as a function of the RFC score cut, for Belle and Belle II respectively. These figures compare the nominal RFC to two alternative RFCs trained with different hyperparameter configurations, which show only small performance differences relative to the nominal choice.

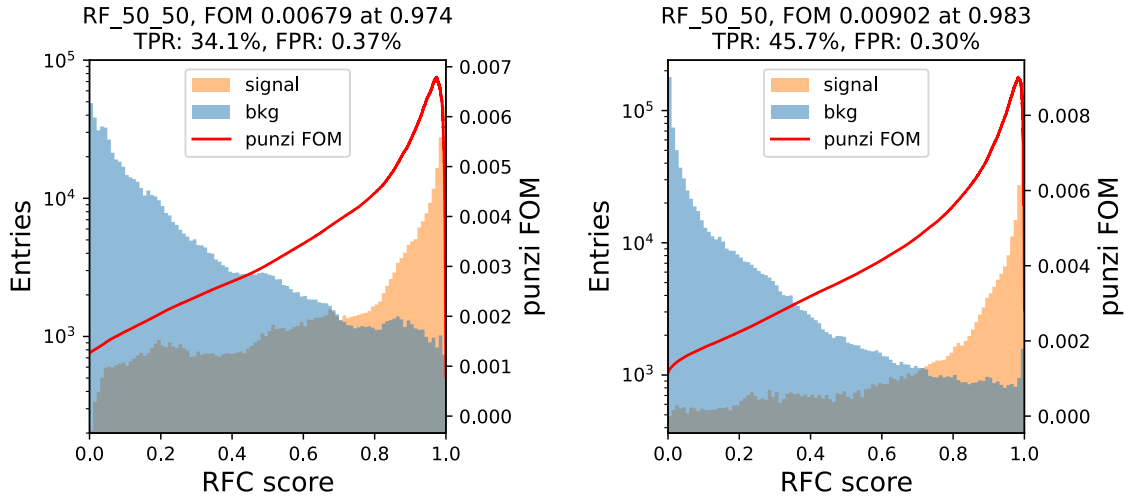


Figure 5.10: Distribution of RFC_score for signal and background events, and the corresponding Punzi FOM, for Belle (left) and Belle II (right).

Figures 5.13 and 5.14 show the ranking of input-variable importance, based on the mean decrease in impurity (MDI), for Belle and Belle II, respectively. Although the pointing angle $\cos\theta_{\mathbf{p},\mathbf{v}}$ shows low importance for Belle, we retain it as an RFC input variable for consistency between the two experiments.

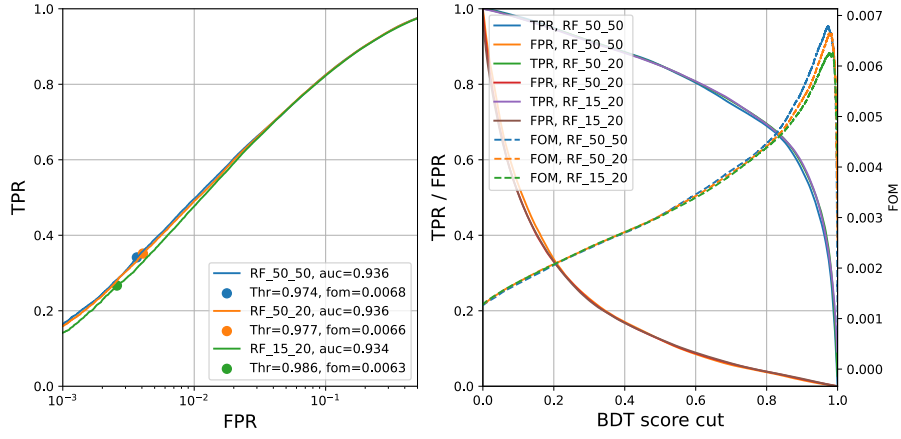


Figure 5.11: Left: ROC curves for three RFC configurations for Belle. Right: true- and false-positive rates and Punzi FOM as a function of the RFC_score cut.

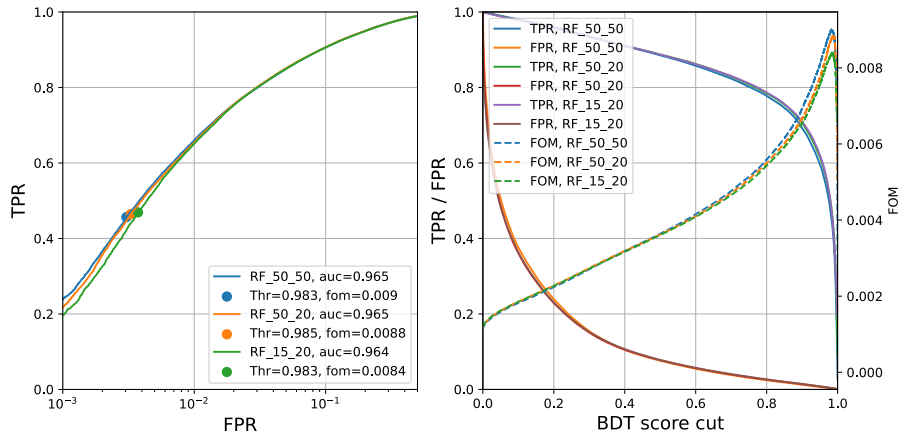


Figure 5.12: Left: ROC curves for three RFC configurations for Belle II. Right: true- and false-positive rates and Punzi FOM as a function of the RFC_score cut.

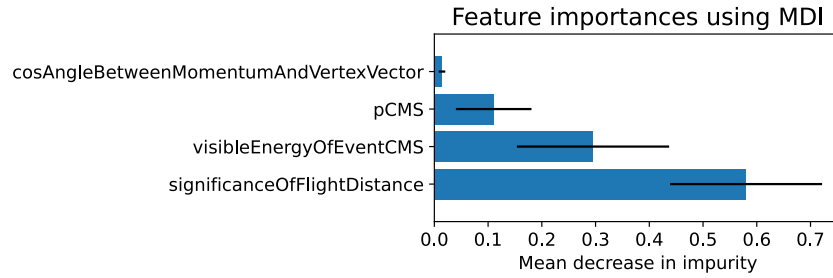


Figure 5.13: Feature importance of RFC input variables for Belle.

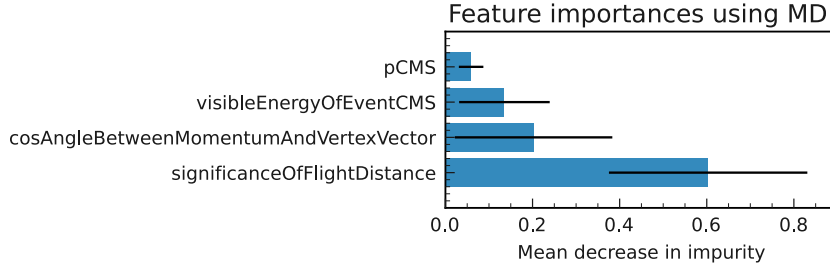


Figure 5.14: Feature importance of RFC input variables for Belle II.

5.8.3 RFC Performance in $M(e^+e^-)$ Intervals

We apply the same RFC working point uniformly across all $M(e^+e^-)$ intervals. Separate per-interval optimisation improves the FOM somewhat (Section 5.10), but has negligible impact on the final upper limit, so we retain the fixed working point for simplicity. Table 5.6 and Figure 5.15 show the resulting signal efficiency and background retention rate in each interval, for the fixed working points $\text{RFC_score} > 0.97$ (Belle) and $\text{RFC_score} > 0.98$ (Belle II).

Table 5.6: RFC performance in individual $M(e^+e^-)$ intervals, for the nominal working points.

$M(e^+e^-)$ [GeV/ c^2]	Signal ϵ [%]		Bkg. retention [%]	
	Belle	Belle II	Belle	Belle II
0.20–0.50	35.8 ± 0.1	47.1 ± 0.3	0.88 ± 0.07	1.04 ± 0.04
0.56–0.95	36.0 ± 0.2	47.4 ± 0.2	0.29 ± 0.03	0.31 ± 0.01
1.05–1.75	37.5 ± 0.2	48.1 ± 0.1	0.17 ± 0.01	0.135 ± 0.006

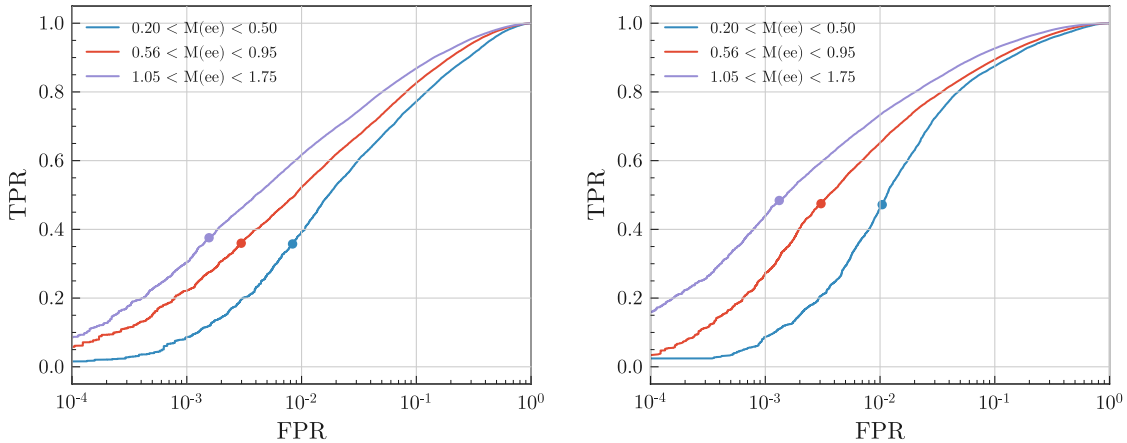


Figure 5.15: RFC performance in three $M(e^+e^-)$ intervals. Dots indicate the nominal working point, $\text{RFC_score} > 0.97$ for Belle (left) and $\text{RFC_score} > 0.98$ for Belle II (right).

As in the full-region comparison above, we evaluate the background retention rate of each experiment at the signal efficiency of the other, separately in each $M(e^+e^-)$ interval (Tables 5.7 and 5.8), confirming that Belle II provides better signal-background discrimination than Belle across all intervals.

Table 5.7: RFC performance for Belle at matched Belle II signal efficiency, in bins of $M(e^+e^-)$.

$M(e^+e^-)$ [GeV/ c^2]	Signal ϵ [%]	Bkg. retention [%]	RFC_score cut
0.20–0.50	47.1 ± 0.3	1.38 ± 0.06	0.9407
0.56–0.95	47.4 ± 0.2	0.73 ± 0.03	0.9401
1.05–1.75	48.1 ± 0.2	0.35 ± 0.01	0.9445

Table 5.8: RFC performance for Belle II at matched Belle signal efficiency, in bins of $M(e^+e^-)$.

$M(e^+e^-)$ [GeV/ c^2]	Signal ϵ [%]	Bkg. retention [%]	RFC_score cut
0.20–0.50	35.7 ± 0.3	0.67 ± 0.03	0.991
0.56–0.95	36.0 ± 0.2	0.17 ± 0.01	0.9912
1.05–1.75	37.5 ± 0.1	0.065 ± 0.004	0.99058

5.9 η -Veto

Although the random forest classifier of Section 5.8 removes the majority of combinatorial background, a substantial fraction of the background surviving the RFC selection originates from a specific background: the decay $D_{(s)}^+ \rightarrow \pi^+(\eta \rightarrow \gamma e^+e^-)$, in which the radiated photon from the η decay is not reconstructed.

Figures 5.16 and 5.17 show the composition of D^+ candidates passing the RFC selection, categorised by their generator-level (mcMother_mcPDG) origin, for Belle and Belle II, respectively. In the region $M(e^+e^-) < 0.5 \text{ GeV}/c^2$, this non-peaking $D_{(s)}^+ \rightarrow \pi^+(\eta \rightarrow \gamma e^+e^-)$ background accounts for approximately 36 % and 53 % of all background events surviving the RFC selection, in Belle and Belle II respectively.

We study this background in detail using the dedicated $D_{(s)}^+ \rightarrow \pi^+(\eta \rightarrow \gamma e^+e^-)$ signal MC samples introduced in Section 4.2.3. Figure 5.18 shows the resulting $M(\pi^+e^+e^-)$ and $M(e^+e^-)$ distributions, separated by generator-level truth-matching ($\text{isSig} == 1$ or $\text{isSig} \neq 1$) and by parent particle (D^+ or D_s^+). Candidates in which the radiated photon is recovered through the bremsstrahlung correction (Section 5.3) are correctly reconstructed, and peak near m_{D^+} in $M(\pi^+e^+e^-)$ and near m_η in $M(e^+e^-)$. Candidates in which the photon is not reconstructed instead populate both distributions broadly, reflecting the missing photon energy.

The peaking component is removed almost entirely by the resonance-exclusion window $0.50 < M(e^+e^-) < 0.56 \text{ GeV}/c^2$ (Section 5.5.3), since it falls within this

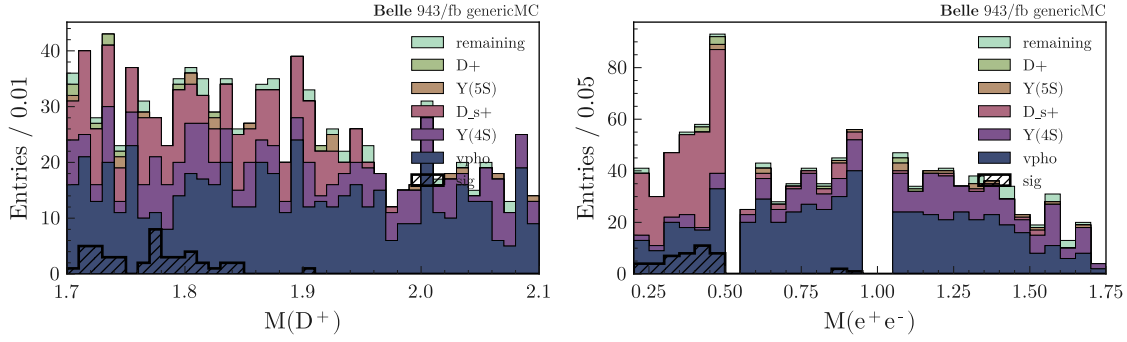


Figure 5.16: Composition of D^+ candidates in the Belle generic MC sample, based on generator-level (mcPDG) origin, after the RFC selection. Candidates labelled “sig” in the $M(e^+e^-) < 0.5 \text{ GeV}/c^2$ region are incorrectly MC-matched $D_{(s)}^+ \rightarrow \pi^+(\eta \rightarrow \gamma e^+e^-)$ decays: non-peaking, but tagged as isSig == 1.

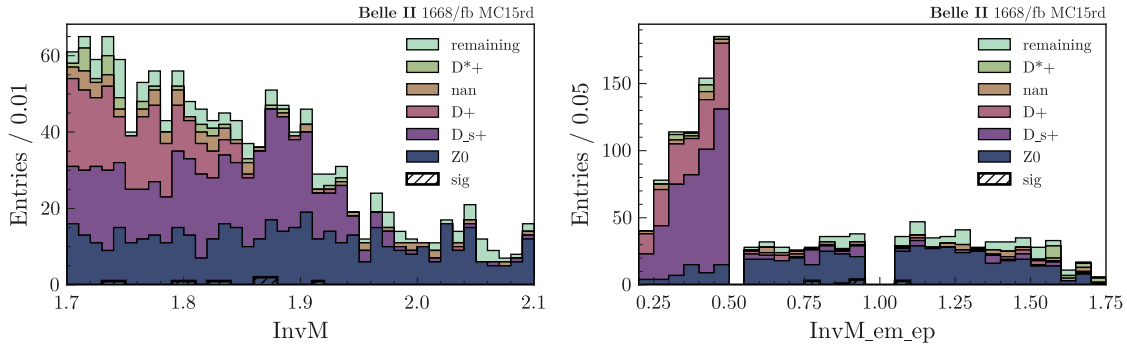


Figure 5.17: Composition of D^+ candidates in the Belle II generic MC sample, based on generator-level (mcPDG) origin, after the RFC selection.

window by construction. The non-peaking component, however, survives this simple mass-window exclusion and instead contributes a broad background below the η mass.

To suppress this non-peaking background, we apply a dedicated η -veto. For each signal candidate, we attempt to reconstruct the alternative decay chain $D_{(s)}^+ \rightarrow \pi^+(\eta \rightarrow \gamma e^+e^-)$, using the same final-state particles and selection criteria as the signal channel but excluding the requirements on $M(e^+e^-)$, and combining the selected π^+ , e^+ , and e^- tracks with an additional photon candidate drawn from the same photon list used for bremsstrahlung recovery, without applying the angular-cone restriction. We reject the D^+ signal candidate if the invariant mass of the resulting $e^+e^-\gamma$ system satisfies $|M(e^+e^-\gamma) - m_\eta| < 10 \text{ MeV}/c^2$, where m_η is taken from the PDG [16]. Over the full $M(e^+e^-)$ range, this requirement removes 8 % (12 %) of background events and 1 % (1 %) of signal events in Belle (Belle II), as estimated from MC simulation.

We further evaluate η -veto performance in the $M(e^+e^-) < 0.5 \text{ GeV}/c^2$ region using the dedicated signal MC samples, with results shown in Table 5.9.

For Belle, the truth-matching procedure incorrectly assigns isSig == 1 to non-peaking $D_{(s)}^+ \rightarrow \pi^+(\eta \rightarrow \gamma e^+e^-)$ candidates as well, preventing a similar separation between peaking and non-peaking contributions. Because the peaking-component’s contribution is negligible after removing $0.5 < M(e^+e^-) < 0.56 \text{ GeV}/c^2$, we did not

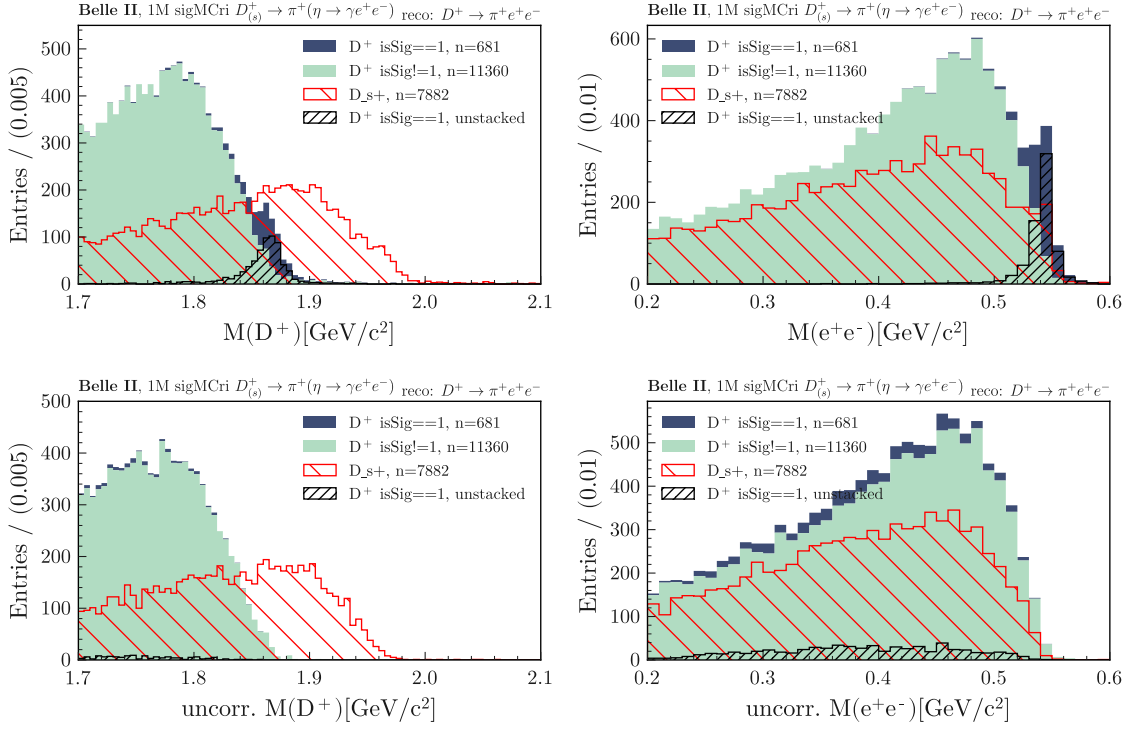


Figure 5.18: $M(\pi^+e^+e^-)$ (left) and $M(e^+e^-)$ (right) distributions for the dedicated $D_{(s)}^+ \rightarrow \pi^+(\eta \rightarrow \gamma e^+e^-)$ signal MC sample in Belle II, separated by truth-matching status and parent particle. The upper row shows the distributions with the bremsstrahlung correction applied, the lower row without it.

Table 5.9: Percentage of events removed by the η -veto within $0.2 < M(e^+e^-) < 0.5 \text{ GeV}/c^2$, evaluated on the dedicated $D_{(s)}^+ \rightarrow \pi^+(\eta \rightarrow \gamma e^+e^-)$ and signal MC samples, for Belle and Belle II.

Signal MC sample	Removal [%]	
	Belle	Belle II
Signal events	10	9
$D^+ \rightarrow \pi^+(\eta \rightarrow \gamma e^+e^-)$ (peaking)	NA	9
$D^+ \rightarrow \pi^+(\eta \rightarrow \gamma e^+e^-)$ (non-peaking)	34	32
$D_s^+ \rightarrow \pi^+(\eta \rightarrow \gamma e^+e^-)$	35	34

investigate it further, and applied the same resonance-exclusion window uniformly.

5.10 Cut Optimisation

We use the Punzi figure of merit (FOM) introduced in Eq. 5.1 to guide working-point selection throughout this chapter, evaluating it as a function of the candidate cut value t , $\text{FOM}(t) = \epsilon(t)/(a/2 + \sqrt{B(t)})$. Unless otherwise stated, B is evaluated over the wide diagnostic window $1.7 < M(\pi^+e^+e^-) < 2.1 \text{ GeV}/c^2$ introduced in Section 5.4.

5.10.1 Significance Parameter

We adopt a target significance of $a = 5$ throughout this analysis. As a cross-check against the more conventional choice of $a = 3$, typically associated with an evidence-level significance, we recomputed the FOM for the primary selection variables and found no appreciable difference in the resulting optimal working points between the two choices.

5.10.2 RFC Working Point

In the nominal analysis, the RFC score requirement is optimised inclusively over the full $M(e^+e^-)$ region rather than separately within each interval (Section 5.8). We additionally studied optimising the RFC working point separately within each $M(e^+e^-)$ interval, finding an improvement in the FOM of approximately 1% for Belle and up to 8% for Belle II relative to the inclusive optimisation. However, propagating these interval-dependent working points to the simultaneous upper-limit determination has a negligible impact on the resulting upper limits, since the larger Belle sample contributes the greater share of the sensitivity. Introducing interval-dependent working points would also require separate systematic-uncertainty evaluations for the RFC efficiency in each interval (Chapter 8), rather than a single evaluation for the inclusive selection. We therefore retain the single, inclusively-optimised RFC working point for the nominal analysis.

We further examined the sensitivity of the RFC FOM to the width of the $M(\pi^+e^+e^-)$ region used to evaluate B (Figure 5.19). The FOM is stable against this choice: the curves obtained for the wide and intermediate windows are nearly indistinguishable, and only the narrowest window around the D^+ peak shifts the FOM to slightly lower values. Together with the negligible effect of interval-dependent optimisation on the final upper limits established above, this shows that the analysis does not depend on the precise location of the RFC working point. We therefore round the final RFC working point to two significant figures, $\text{RFC_score} > 0.97$ (Belle) and $\text{RFC_score} > 0.98$ (Belle II).

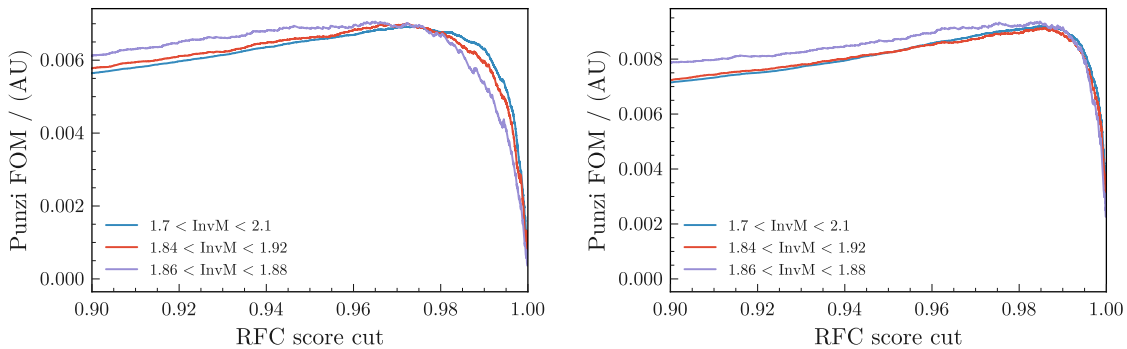


Figure 5.19: Comparison of the Punzi FOM for different widths of the $M(\pi^+e^+e^-)$ region used to evaluate the background yield B , for Belle (left) and Belle II (right). Curves are rescaled to a common peak height for ease of comparison.

5.10.3 Other Selection Variables

We also use the Punzi FOM to evaluate the discriminating power of other selection variables, as shown in Figures 5.2 and 5.3 (Section 5.4). As discussed there, the

resulting optimal cut values are not applied directly; rather, they serve as guidance for defining approximate selection regions and for identifying variables with sufficient separation power to include as RFC input variables.

For the pion and electron identification requirements (Section 5.1), the FOM is relatively flat over a wide range of candidate cut values (Figure 5.20). In this regime, we choose the working point based on the granularity of the available particle-identification correction tables (Chapter 8) rather than the location of the poorly-constrained FOM maximum. For the D^+ centre-of-mass momentum, p_{CMS} , we retain the requirement at 2.5 GeV/c, consistent with the value used in comparable analyses and sufficiently loose that the RFC performs the dominant background discrimination for this variable.

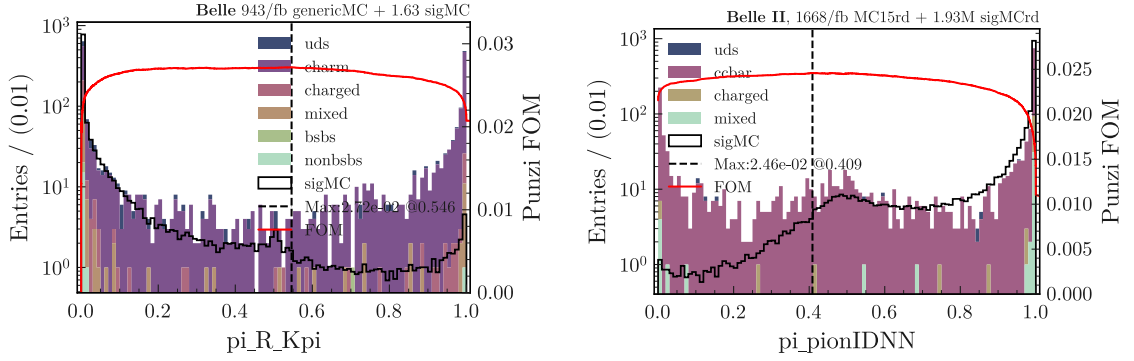


Figure 5.20: Distribution of the pion identification variable, $R(K, \pi)$ for Belle (left) and πID_{NN} for Belle II (right), on MC samples after the RFC selection.

5.11 $M(e^+e^-)$ Interval Definition

In addition to the inclusive branching-fraction measurement, we divide the selected $D^+ \rightarrow \pi^+ e^+ e^-$ candidates into three intervals of $M(e^+e^-)$, allowing a differential measurement of the branching fraction as a function of the lepton pair invariant mass. Excluding the resonance regions defined in Section 5.5.3, the accessible phase space, $0.20 < M(e^+e^-) < 1.75 \text{ GeV}/c^2$, is divided into three intervals:

$$I_1 = (0.20, 0.50) \text{ GeV}/c^2, \quad (5.2)$$

$$I_2 = (0.56, 0.95) \text{ GeV}/c^2, \quad (5.3)$$

$$I_3 = (1.05, 1.75) \text{ GeV}/c^2. \quad (5.4)$$

We denote the union of these three intervals as $I_{\text{full}} = I_1 \cup I_2 \cup I_3$. The inclusive branching-fraction measurement is performed over I_{full} , and therefore already excludes the η and ϕ resonance regions. The signal branching fraction is determined independently within each interval, as well as over I_{full} , using the fit procedure described in Chapter 7. We additionally define a validation interval,

$$I_{\phi}^{\text{Valid.}} = (0.995, 1.035) \text{ GeV}/c^2, \quad (5.5)$$

a narrower window centred on the $\phi(1020)$ resonance than the excluded region of Section 5.5.3, within which we reconstruct the validation channel $D^+ \rightarrow \pi^+ \phi(\rightarrow e^+ e^-)$ (Section 4.2.3), to test the analysis procedure on a topologically similar, previously measured signal.

5.12 Summary

The final selection combines the initial-selection requirements of Sections 5.1–5.11 with the random forest classifier requirement (Section 5.8) and, for the signal channel, the η -veto (Section 5.9). Tables 5.10 and 5.11 summarise the complete set of selection criteria applied to the signal and normalisation channels, respectively. Belle criteria are identical to Belle II except where indicated in brackets.

The resulting signal reconstruction efficiency ranges from 3.8 % to 5.9 % in Belle and from 4.7 % to 7.4 % in Belle II, depending on the $M(e^+e^-)$ interval (Section 5.11). The full signal reconstruction efficiency curve as a function of $M(e^+e^-)$, evaluated from the phase-space-generated signal MC sample, is shown in Figure 6.1.

Table 5.10: Summary of selection criteria for the signal mode. Belle criteria are the same as for Belle II, except for the cuts written in brackets.

Candidate	Selection Criteria
π^+	$ dz < 2.0, dr < 0.5$ CDC acceptance $\pi\text{ID}_{\text{NN}} > 0.40$ ($\mathcal{R}(\pi, K) > 0.5$)
$e^\pm$	$ dz < 2.0, dr < 0.5$ CDC acceptance $\text{PID}_e > 0.9$ $(\mathcal{R}(e, \pi) > 0.9)$ Bremsstrahlung correction (Section 5.3): FWD: $E_\gamma > 0.075$ GeV BRL: $E_\gamma > 0.050$ GeV BWD: $E_\gamma > 0.100$ GeV
D^+	$p_{\text{CMS}} > 2.5$ GeV/c TreeFitter: $P(\chi^2) > 0.001$, IP constrained Best-candidate selection: highest $P(\chi^2)$ Total charge of final-state particles: +1 $0.2 \text{ GeV}/c^2 < M(e^+e^-) < 1.75 \text{ GeV}/c^2$ Resonance exclusion (Section 5.5.3): η : $0.50\text{--}0.56 \text{ GeV}/c^2$ ϕ : $0.95\text{--}1.05 \text{ GeV}/c^2$ Photon-conversion veto: $M(\gamma \rightarrow e^+e^-) > 0.05 \text{ GeV}/c^2$ η -veto: $ M(e^+e^-\gamma) - m_\eta > 10 \text{ MeV}/c^2$
Random	Input variables (Section 5.4): p_{CMS} , visible CMS energy
Forest	pointing angle $\cos \theta_{p,v}$
Classifier	significance of flight distance RFC_score > 0.98 (0.97)

Table 5.11: Summary of selection criteria for the normalisation mode. Belle criteria are the same as for Belle II, except for the cuts written in brackets.

Candidate	Selection Criteria
π^+	$ dz < 2.0, dr < 0.5$ CDC acceptance $\pi\text{ID}_{\text{NN}} > 0.40$ ($\mathcal{R}(\pi, K) > 0.5$)
K^+	$ dz < 2.0, dr < 0.5$ CDC acceptance $\text{PID}_K > 0.6$ ($\mathcal{R}(K, \pi) > 0.6$)
D^+	$p_{\text{CMS}} > 2.5 \text{ GeV}/c$ TreeFitter: $P(\chi^2) > 0.001$, IP constrained Best-candidate selection: highest $P(\chi^2)$ Total charge of final-state particles: +1 $ M(K^+K^-) - m_\phi < 20 \text{ MeV}/c^2$
Random	Input variables (Section 5.4):
Forest	p_{CMS} , visible CMS energy
Classifier	pointing angle $\cos \theta_{p,v}$
	significance of flight distance
	$\text{RFC_score} > 0.98$ (0.97)

6. Reconstruction Efficiency and MC/Data Comparison

Following the event selection described in Chapter 5, we determine the reconstruction efficiency for each channel entering the branching-fraction calculation of Eq. 7.1: the signal mode $D^+ \rightarrow \pi^+ e^+ e^-$, the validation mode $D^+ \rightarrow \pi^+ \phi(\rightarrow e^+ e^-)$, and the normalisation mode $D^+ \rightarrow \pi^+ K^+ K^-$. The reconstruction efficiency of a given channel is defined as the fraction of generated signal events that pass the full selection chain and are correctly identified as signal via MC truth-matching, determined using dedicated signal MC samples. We evaluate both inclusively and, for the signal channel, differentially in bins of $M(e^+ e^-)$, to account for the dependence of the reconstruction efficiency on the electron-pair invariant mass.

Because the analysis relies on simulation both to determine these efficiencies and to model background shapes, we assess the level of agreement between data and MC used for these purposes in Section 6.2. We compare data and generic MC distributions for the validation and normalisation modes, for which the final selection can be examined without introducing analysis bias, as well as in signal-mode sideband regions outside the D^+ signal window. To avoid introducing experimenter bias into the selection or fit design, we did not examine the signal-region data in any form — no data/MC comparisons, distributions, fits, or signal extraction — until the analysis procedure was fully finalised. We performed the signal-mode data/MC comparison only after this point. Residual data/MC differences affecting the reconstruction efficiencies and the fixed fit shapes (the RFC response and the signal PDF parameters) are not used to correct the central efficiency values, but are instead propagated as systematic uncertainties in Chapter 8. The one exception is PID, which is corrected directly using dedicated data/MC correction factors (Section 8.2.1). We expect the overall yield excess seen in the data/MC comparisons to cancel in the branching-fraction ratio, and we do not assign a systematic uncertainty for it.

6.1 Reconstruction Efficiency

6.1.1 Normalisation Channel Reconstruction Efficiency

As introduced in Section 5.2.1, the normalisation-channel mass window $|M(K^+ K^-) - m_\phi| < 20 \text{ MeV}/c^2$ is dominated by the $\phi(1020)$ resonance but also receives a non-resonant $D^+ \rightarrow \pi^+ K^+ K^-$ contribution and its interference with the resonant amplitude. Because the normalisation branching fraction $\mathcal{B}_{\pi K K} = (9.68 \pm 0.18) \times 10^{-3}$ [16] refers to the fully inclusive decay $D^+ \rightarrow \pi^+ K^+ K^-$, the reconstruction efficiency entering Eq. 7.1 must account for the fraction of all such decays that fall within the ϕ mass window, rather than the efficiency for the resonant decay alone.

For Belle II, we determine the normalisation efficiency directly from an MC sample generated according to a Dalitz-plot model for $D^+ \rightarrow \pi^+ K^+ K^-$, which incorporates the ϕ resonance, the non-resonant amplitude, and their interference. Because this generator model already reproduces the full dynamics of the decay, the reconstruction efficiency obtained from it, evaluated within the ϕ mass window, is directly usable in the branching-fraction calculation without further correction for non-resonant effects.

For Belle, no such Dalitz-plot generator was available at the time the signal MC samples were produced. We instead generated the normalisation channel exclusively via the ϕ resonance, using the SVS model for $D^+ \rightarrow \pi^+ \phi$ and the VSS model for $\phi \rightarrow K^+ K^-$ (Section 4.2). The resonance-only efficiency therefore does not capture the acceptance of non-resonant events into the ϕ mass window, and a dedicated correction factor is required.

We derive this correction using the Belle II Dalitz-model sample, exploiting the fact that both the resonance-only and Dalitz-model efficiencies can be computed within Belle II, allowing their ratio to be used as a proxy correction for Belle. We estimate the Dalitz-model normalisation efficiency for Belle as

$$\epsilon(D^+ \rightarrow \pi^+ K^+ K^-; M(\phi))_{\text{Dalitz}}^B = \frac{\epsilon(M(\phi))_{\text{Dalitz}}^{B2}}{\epsilon_{\text{resonance}}^{B2}} f(M(\phi)) \epsilon_{\text{resonance}}^B, \quad (6.1)$$

where $\epsilon(M(\phi))_{\text{Dalitz}}^{B2} = (7.87 \pm 0.04)\%$ is the Belle II Dalitz-model efficiency evaluated within the ϕ mass window, $\epsilon_{\text{resonance}}^B$ and $\epsilon_{\text{resonance}}^{B2}$ are the resonance-only efficiencies obtained for Belle and Belle II, respectively, and $f(M(\phi)) = (28.49 \pm 0.03)\%$ is the fraction of generator-level $D^+ \rightarrow \pi^+ K^+ K^-$ events, generated according to the Dalitz model, that fall within the ϕ mass window, extracted at generator level from the Belle II Dalitz-model sample. The ratio $\epsilon(M(\phi))_{\text{Dalitz}}^{B2}/\epsilon_{\text{resonance}}^{B2}$, together with $f(M(\phi))$, defines an overall correction factor of 0.308 applied multiplicatively to the Belle resonance-only efficiency. The dominant factor $f(M(\phi))$ is a generator-level quantity, so only the residual 8% acceptance ratio is transferred from Belle II to Belle. The relative uncertainties of the individual factors in Eq. 6.1 are summed linearly, which is conservative given that they are partly correlated.

Table 6.1 summarises the resonance-only and final (Dalitz-corrected, for Belle; Dalitz-model, for Belle II) normalisation efficiencies used in the branching-fraction calculation. The final efficiencies are $\epsilon_{\pi K K} = (1.46 \pm 0.02)\%$ for Belle and $(2.24 \pm 0.01)\%$ for Belle II, where the quoted uncertainties are statistical only.

Table 6.1: Normalisation-channel reconstruction efficiencies. The resonance-only efficiency is obtained directly from signal MC generated exclusively via the ϕ resonance; the final efficiency additionally accounts for the non-resonant $D^+ \rightarrow \pi^+ K^+ K^-$ contribution, either directly (Belle II, Dalitz model) or via the correction of Eq. 6.1 (Belle). Uncertainties are statistical only.

	Belle [%]	Belle II [%]
$\epsilon(D^+ \rightarrow \pi^+ \phi(\rightarrow K^+ K^-))_{\text{resonance}}$	4.74 ± 0.01	7.27 ± 0.01
$\epsilon_{\pi K K}$ (used in Eq. 7.1)	1.46 ± 0.02	2.24 ± 0.01

As noted before, the efficiencies of Table 6.1 are taken directly from simulation, without any data/MC correction applied to their central values. The one exception is PID, where we apply dedicated data/MC correction factors following the procedure described in Section 8.2.1. Any other data/MC discrepancies affecting the normalisation channel, discussed in Section 6.2.2, are not corrected for directly but instead enter the analysis as systematic uncertainties, described in Chapter 8.

6.1.2 Signal and Validation Channel Reconstruction Efficiency

We determine the reconstruction efficiency of the signal and validation channels using dedicated signal MC samples, generated separately for each experiment (Section 4.2). Data and MC samples taken at different CMS energies are combined without an explicit correction. Because the signal MC samples are generated proportionally to the data luminosity of each run (or experiment), the reported efficiency represents a luminosity-weighted average across all run (or experiment) conditions. Table 6.2 summarises the inclusive reconstruction efficiencies of the validation channel, $D^+ \rightarrow \pi^+ \phi (\rightarrow e^+ e^-)$, and the signal channel, $D^+ \rightarrow \pi^+ e^+ e^-$, evaluated over the full $M(e^+ e^-)$ range (I_{full} ; for the signal channel, excluding the η and ϕ resonance regions, Section 5.11). Belle II achieves a higher reconstruction efficiency than Belle across both channels. This improvement originates from Belle II's stronger RFC signal-background separation (Section 5.8), which benefits directly from the detector's improved vertex resolution and more precise flight-distance measurements.

Table 6.2: Inclusive signal and validation channel reconstruction efficiencies. Uncertainties are statistical only.

Decay mode	Belle [%]	Belle II [%]
$D^+ \rightarrow \pi^+ \phi (\rightarrow e^+ e^-)$ (validation)	4.75 ± 0.02	6.13 ± 0.02
$D^+ \rightarrow \pi^+ e^+ e^-$ (signal, I_{full})	5.29 ± 0.02	6.57 ± 0.02

Because the signal reconstruction efficiency depends on $M(e^+ e^-)$ (Figure 6.1), we additionally determine the efficiency differentially in each of the three signal $M(e^+ e^-)$ intervals defined in Section 5.11. Table 6.3 reports these efficiencies relative to the number of generated signal events within the corresponding bin, which is the quantity used directly in the per-interval branching-fraction calculation of Eq. 7.1. The efficiency rises with increasing $M(e^+ e^-)$ in both experiments, ranging from 3.82% (4.73%) in I_1 to 5.94% (7.42%) in I_3 for Belle (Belle II).

A finer binning of the signal reconstruction efficiency versus $M(e^+ e^-)$, obtained directly from the phase-space-generated signal MC sample without any interval selection, is shown in Figure 6.1.

CHAPTER 6. RECONSTRUCTION EFFICIENCY AND MC/DATA COMPARISON

Table 6.3: Signal reconstruction efficiency in each $M(e^+e^-)$ interval, defined as the ratio of reconstructed to generated signal events within that interval. Uncertainties are statistical only.

$M(e^+e^-)$ interval [GeV/c^2]	Belle [%]	Belle II [%]
$I_1 = (0.20, 0.50)$	3.82 ± 0.04	4.73 ± 0.04
$I_2 = (0.56, 0.95)$	5.03 ± 0.03	6.14 ± 0.03
$I_3 = (1.05, 1.75)$	5.94 ± 0.03	7.42 ± 0.03

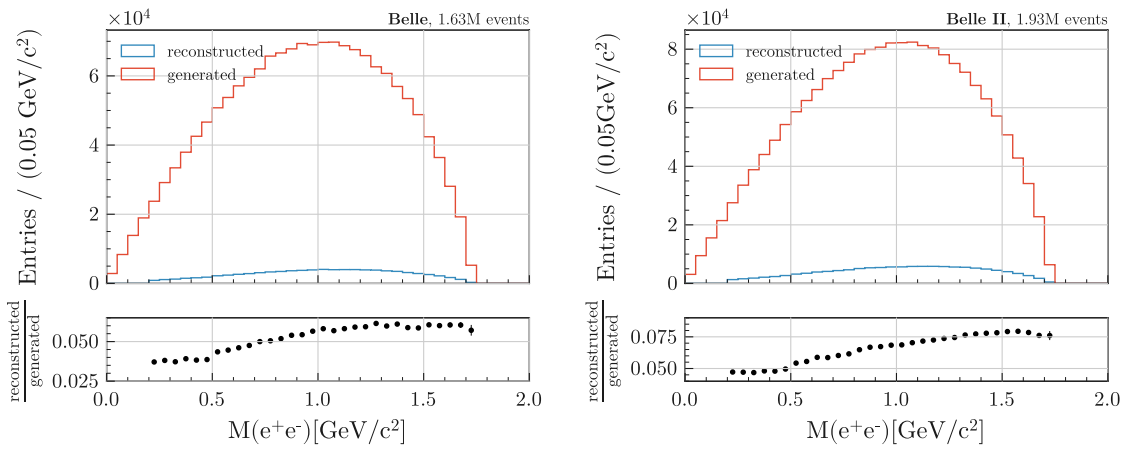


Figure 6.1: Signal reconstruction efficiency after the final selection as a function of $M(e^+e^-)$, determined from the phase-space signal MC sample, for Belle (left) and Belle II (right).

As for the normalisation channel, we apply data/MC correction to the signal efficiency values only for PID requirements. Any remaining data/MC discrepancies are addressed in Section 6.2 and propagated as systematic uncertainties in Section 8.4.3.

6.2 Data/MC Validation

Before using the generic MC sample for background modelling and for the systematic studies of Chapter 8, we compare data and generic MC distributions after the final event selection to check for qualitative disagreements in the shapes of key kinematic, PID, and RFC variables. In this section, data/MC agreement is visually investigated by overlaying distributions without directly quantifying or correcting any observed discrepancies. Because the validation channel is statistically limited, and because we left the signal-channel data in the signal region unexamined until the selection and fit procedure had been fixed, we do not use either mode to derive data/MC corrections or systematic uncertainties. Therefore, while particle-identification data/MC corrections are applied directly across all channels using centralised calibration tables (Section 8.2.1), quantitative data/MC discrepancies for the RFC response (Section 8.2.2) and signal PDF shape parameters (Section 8.4) are evaluated solely in the high-yield normalisation channel and subsequently propagated as systematic uncertainties. We perform these visual checks for the normalisation and validation modes,

as well as the signal-mode sidebands. Once the selection and fit procedure had been fixed, we also examined the signal region itself.

Throughout this section, generic MC distributions are normalised to the integrated luminosity of the corresponding data sample, and no correction for the observed yield difference is applied. The sideband ranges differ between the two experiments because they are set by the width of the centrally produced skim windows (Section 4.1), which is wider for Belle. The sideband comparisons are therefore not directly comparable between experiments.

6.2.1 Validation Mode

We compare data and generic MC distributions for the validation channel, $D^+ \rightarrow \pi^+\phi(\rightarrow e^+e^-)$, after the final selection, with $M(e^+e^-)$ restricted to the validation window $I_\phi^{\text{Valid.}} = (0.995, 1.035) \text{ GeV}/c^2$ (Section 5.11). Figure 6.2 shows the $M(\pi^+e^+e^-)$ distribution for Belle and Belle II. Good agreement is observed between data and generic MC in both experiments.

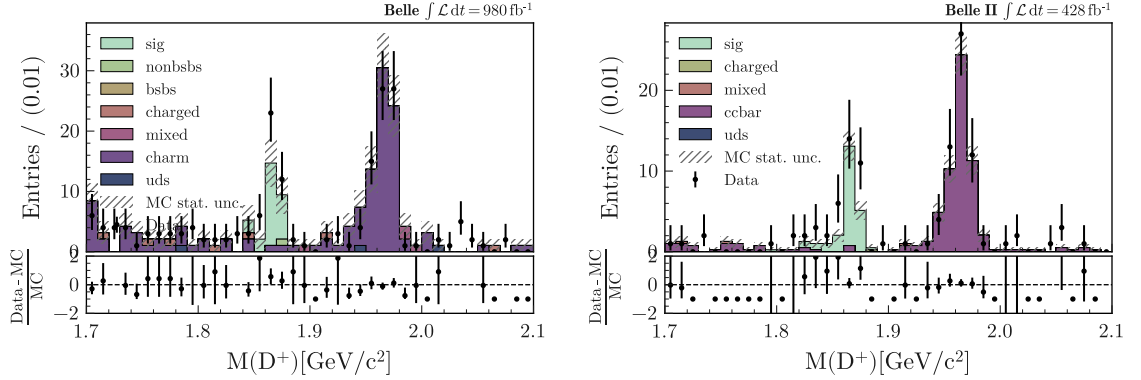


Figure 6.2: Data/generic MC comparison of the $M(\pi^+e^+e^-)$ distribution for the validation channel $D^+ \rightarrow \pi^+\phi(\rightarrow e^+e^-)$, for Belle (left) and Belle II (right).

The signal efficiencies and background retention rates of the RFC, trained on the signal mode and evaluated on the validation mode using the same working points (RFC_score > 0.97 for Belle, > 0.98 for Belle II), are summarised in Table 6.4 and are similar to the values obtained for the signal mode (Table 5.4). Additional data/MC comparisons of the validation-mode final-selection distributions, including the RFC input variables, are provided in Appendix B.4.1.

Table 6.4: RFC performance for the validation mode, evaluated using the nominal working points RFC_score > 0.97 (Belle) and > 0.98 (Belle II). The background retention is only loosely constrained by the limited size of the generic MC sample.

	Belle	Belle II
Signal efficiency [%]	38.7 ± 0.1	50.0 ± 0.1
Background retention [%]	0.3 ± 0.3	0.3 ± 0.3

6.2.2 Normalisation Mode

We compare data and generic MC distributions for the normalisation channel, $D^+ \rightarrow \pi^+ K^+ K^-$, after the final selection. Figure 6.3 shows the $M(\pi^+ K^+ K^-)$ distribution for Belle and Belle II. For Belle, the generic MC underestimates the number of entries across the full $M(\pi^+ K^+ K^-)$ range, and the ϕ resonance peak is slightly shifted relative to data. For Belle II, the MC underestimates the D^+ peak but overestimates the D_s^+ peak, while the background-dominated region of $M(\pi^+ K^+ K^-)$ shows good agreement between MC and data. This discrepancy in the reconstructed D^+ yield may originate either from an incorrect D^+ production rate assumed in simulation or from a difference in reconstruction efficiency between data and MC, since no corrections are applied at this stage. A production-rate difference cancels exactly in the branching-fraction ratio, and given the similar kinematics of the normalisation and signal channels, an efficiency difference is expected to cancel to a large extent as well.

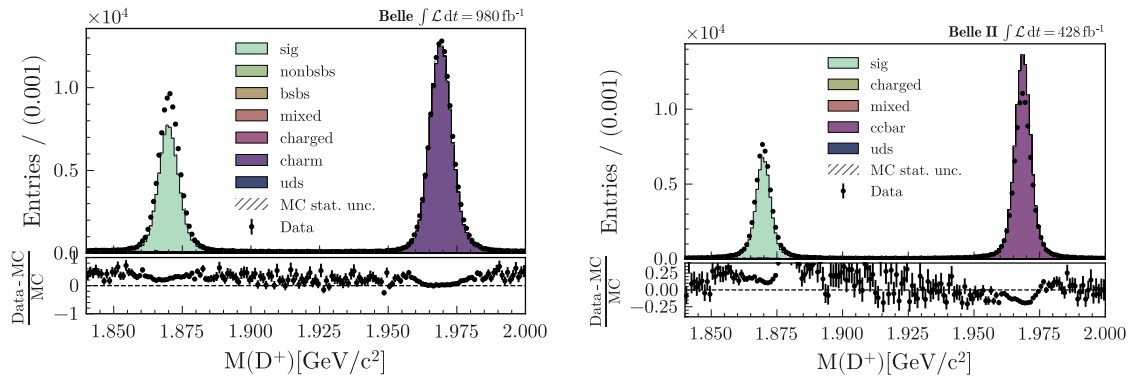


Figure 6.3: Data/generic MC comparison of the $M(\pi^+ K^+ K^-)$ distribution for the normalisation channel $D^+ \rightarrow \pi^+ K^+ K^-$, for Belle (left) and Belle II (right).

The signal efficiencies and background retention rates of the RFC, evaluated using the same working points as the signal mode (RFC_score > 0.97 for Belle, > 0.98 for Belle II, Section 5.8), restricted to events within $10 \text{ MeV}/c^2$ of the nominal D^+ mass, are summarised in Table 6.5, and are slightly worse than for the signal mode (Table 5.4). This is primarily due to the shift in the significance of the flight distance distribution towards lower values in the normalisation mode, which more closely resembles the background distribution from generic MC. At matched signal efficiency, Belle II retains substantially less background than Belle (0.86% vs. 1.51% background retention at $\epsilon \approx 37\%$), confirming the stronger discrimination power of the Belle II RFC already noted in Section 5.8.

Table 6.5: RFC performance for the normalisation mode, evaluated using the nominal working points RFC_score > 0.97 (Belle) and > 0.98 (Belle II), for events within $10 \text{ MeV}/c^2$ of the nominal D^+ mass.

	Belle	Belle II
Signal efficiency [%]	26.74 ± 0.09	37.05 ± 0.07
Background retention [%]	0.64 ± 0.02	0.86 ± 0.02

Additional data/MC comparisons of the normalisation-mode final-selection distributions, including $M(K^+K^-)$, p_{CMS} , and the RFC input variables, are provided in Appendix B.4.2.

6.2.3 Sidebands

We additionally compare data and generic MC in the $M(\pi^+e^+e^-)$ sideband regions after final selection: $1.50 \text{ GeV}/c^2 < M(\pi^+e^+e^-) < 1.70 \text{ GeV}/c^2$ and $2.10 \text{ GeV}/c^2 < M(\pi^+e^+e^-) < 2.30 \text{ GeV}/c^2$ for Belle, and $1.67 \text{ GeV}/c^2 < M(\pi^+e^+e^-) < 1.70 \text{ GeV}/c^2$ and $2.10 \text{ GeV}/c^2 < M(\pi^+e^+e^-) < 2.17 \text{ GeV}/c^2$ for Belle II. A larger discrepancy between data and generic MC is observed in the sidebands than in the validation and normalisation modes, particularly in the right sideband (Figures 6.4 and 6.5). For Belle, this discrepancy originates predominantly from lower values of the flight-distance significance.

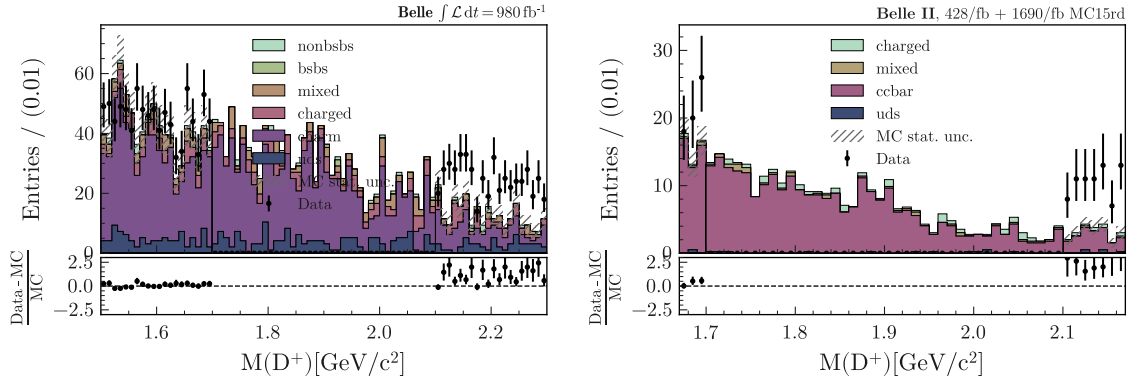


Figure 6.4: Data/generic MC comparison of the $M(\pi^+e^+e^-)$ distribution in the sideband regions after final selection, for Belle (left) and Belle II (right).

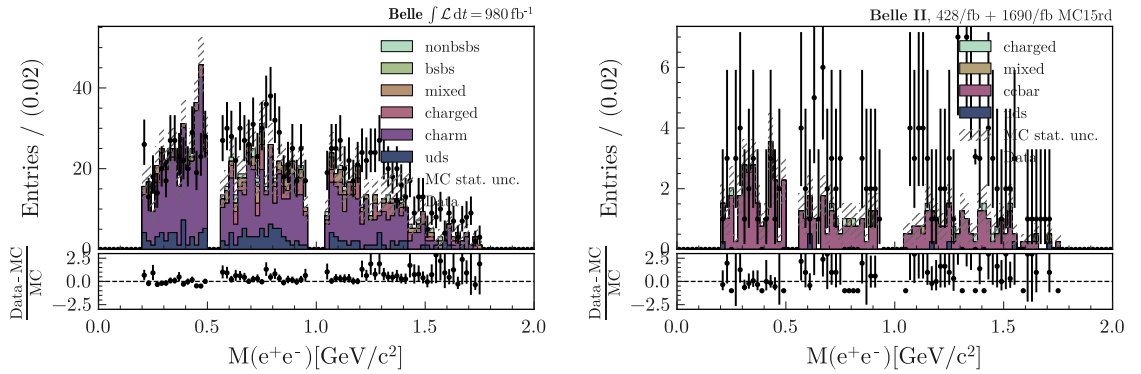


Figure 6.5: Data/generic MC comparison of the $M(e^+e^-)$ distribution in the sideband regions after final selection, for Belle (left) and Belle II (right).

Additional data/MC comparisons of the remaining selection variables in the sideband regions are provided in Appendix B.4.3.

6.2.4 Signal Mode

We perform the data/generic MC comparison for the signal mode only after the analysis procedure was fully finalised, consistent with the approach stated in the

CHAPTER 6. RECONSTRUCTION EFFICIENCY AND MC/DATA COMPARISON

chapter introduction. As observed in the sideband comparison (Section 6.2.3), data exceeds generic MC by approximately 27 % and 46 % for Belle and Belle II, respectively. In the $M(e^+e^-)$ distribution, we find fewer events in the 0.20–0.50 GeV/ c^2 interval and more events in the remaining two intervals; as discussed in Chapter 9, this leads to a tighter upper limit in the lowest $M(e^+e^-)$ interval and slightly weaker limits in the remaining two, relative to expectation.

Table 6.6: Observed data and generic MC event yields for the signal mode after final selection, within $1.7 < M(\pi^+e^+e^-) < 1.92$ GeV/ c^2 . Generic MC is normalised to the integrated luminosity of the data.

Sample	$M(e^+e^-)$ interval [GeV/ c^2]			
	0.20–0.50	0.56–0.95	1.05–1.75	0.20–1.75
Belle data	156	369	357	882
Belle MC	246.3	219.3	230.7	696.4
Belle II data	60	108	177	345
Belle II MC	122.1	48.0	67.2	237.3

Figures 6.6 and 6.7 compare the $M(\pi^+e^+e^-)$ and $M(e^+e^-)$ distributions between data and generic MC for the signal mode after the final selection.

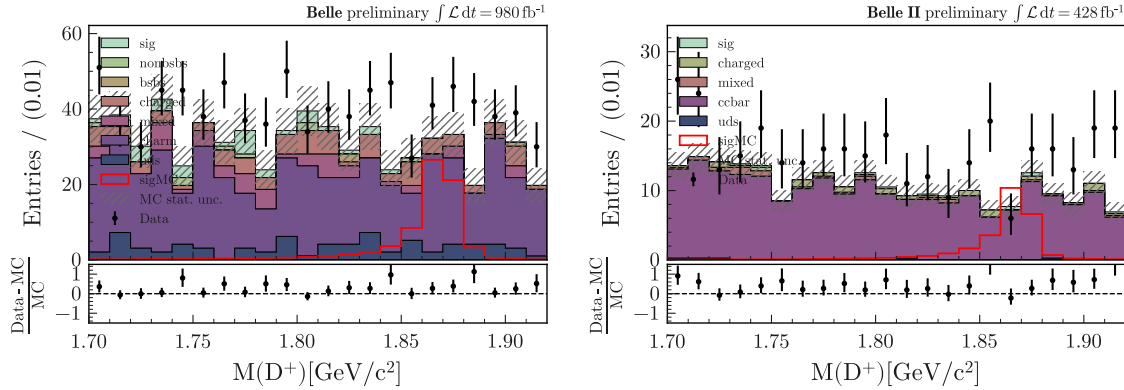


Figure 6.6: Data/generic MC comparison of the $M(\pi^+e^+e^-)$ distribution for the signal mode after final selection, for Belle (left) and Belle II (right). Generic MC is normalised to the integrated luminosity of the data. The signal MC distribution is also superimposed, shown for shape comparison only and not to scale.

Additional data/MC comparisons of the signal-mode final-selection distributions, restricted to events near the D^+ mass peak, are provided in Appendix B.4.4.

6.3 Summary

Across all channels, we observe an excess of data yields over the MC expectation (Sections 6.2.2–6.2.4), most pronounced in the sidebands and the signal mode. The excess is not uniform in $M(e^+e^-)$: in the signal mode the generic MC over-predicts

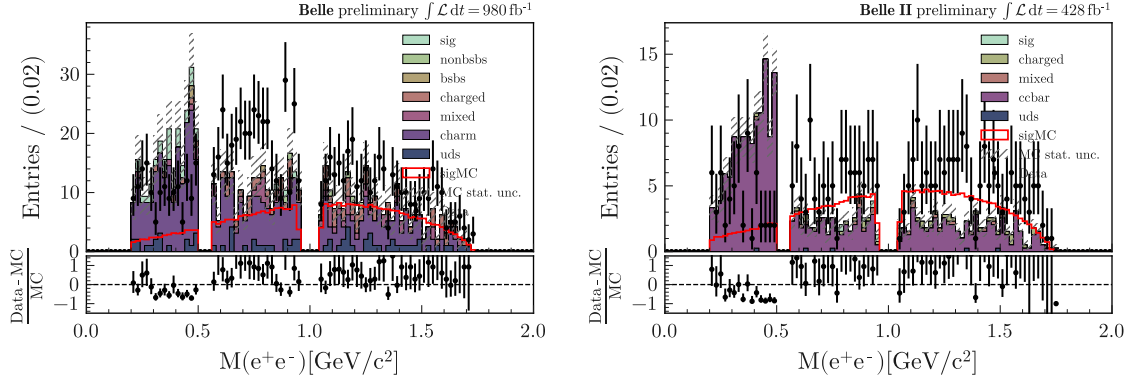


Figure 6.7: Data/generic MC comparison of the $M(e^+e^-)$ distribution for the signal mode after final selection, for Belle (left) and Belle II (right). The signal MC distribution is also superimposed, shown for shape comparison only and not to scale.

the yield in I_1 by 60 % for Belle and by roughly a factor of two for Belle II, and under-predicts it in I_2 and I_3 by 35 % to 60 % (Table 6.6). In the normalisation mode the excess appears in the reconstructed D^+ yield itself. Because the branching fraction is measured as a ratio to the normalisation channel (Eq. 7.1), the reconstruction efficiencies enter only through their ratio. Given the similar kinematics and selection of the signal, validation, and normalisation channels, an excess common to all of them is expected to cancel, and no systematic uncertainty is assigned for it. As a cross-check, the validation channel measurement of $\mathcal{B}(D^+ \rightarrow \pi^+ \phi(\rightarrow e^+ e^-))$ agrees with the value derived from the Particle Data Group averages (Chapter 9). Any data/MC difference in the combinatorial background enters the Neyman construction directly through the fitted background yield, and is therefore already reflected in the upper limit.

7. Fit Strategy and Validation

We now describe how the branching fraction is extracted from the data. This chapter establishes the full fit machinery and validates it on simulation before it is applied to data in Chapter 9. To reduce the potential for experimental bias, we fixed the selection and fit procedures before examining the signal region and did not change them afterwards. No result presented here uses signal-region data. The only input taken from data before examining the signal region is the branching fraction of the validation channel, which lies outside the signal region, used in Section 7.8.3 to set the generation point of a toy study to the value subsequently measured there.

We first define the branching-fraction relation to the normalisation channel (Section 7.1), the likelihood functions used to extract the yields (Section 7.2), and the probability density functions (PDFs) describing the signal and background in each decay mode (Section 7.3), whose shape parameters are determined from fits to signal MC (Section 7.4). We then introduce the simultaneous Belle and Belle II fit, in which the branching fraction is a single shared free parameter (Section 7.5), and the statistical procedures used to assess a possible signal and to set an upper limit in its absence (Section 7.6). Finally, we verify that this procedure extracts the branching fraction without bias, through closure tests on generic MC (Section 7.7) and toy-MC linearity and pull studies (Section 7.8).

7.1 Branching Fraction Determination

We measure the signal branching fraction relative to the normalisation channel $D^+ \rightarrow \pi^+ K^+ K^-$, reconstructed within $|M(K^+ K^-) - m_\phi| < 20 \text{ MeV}/c^2$ (Section 5.2.1). Measuring a ratio rather than an absolute rate removes the need to know the integrated luminosity and the D^+ production cross section. It cancels effects common to the two channels. In each $M(e^+ e^-)$ interval I_j the branching fraction is

$$\mathcal{B}_{\pi ee}^j = \frac{N_{\pi ee}^j}{N_{\pi KK}} \frac{\epsilon_{\pi KK}}{\epsilon_{\pi ee}^j} \mathcal{B}_{\pi KK}, \quad (7.1)$$

where $N_{\pi ee}^j$ and $\epsilon_{\pi ee}^j$ are the signal yield and reconstruction efficiency in that interval, $N_{\pi KK}$ and $\epsilon_{\pi KK}$ the corresponding quantities for the normalisation channel, and $\mathcal{B}_{\pi KK} = (9.68 \pm 0.18) \times 10^{-3}$ the inclusive branching fraction of the normalisation channel [16]. The index j runs over the three $M(e^+ e^-)$ intervals I_1, I_2, I_3 and the inclusive range I_{full} defined in Section 5.11, and, for the validation channel, over the interval $I_\phi^{\text{Valid.}}$. We define $\epsilon_{\pi KK}$ as the number of candidates reconstructed within the ϕ mass window after the full selection, divided by the number of all generated $D^+ \rightarrow \pi^+ K^+ K^-$ decays, so that the acceptance of the mass window is contained in the efficiency itself. The efficiencies entering Eq. (7.1) are those determined in Chapter 6. Since only the ratio $\epsilon_{\pi KK}/\epsilon_{\pi ee}^j$ enters, effects that shift both efficiencies by a common factor cancel, and only the residual differences between the two channels

are propagated as systematic uncertainties (Chapter 8).

7.2 Likelihood Construction

All yields in this analysis are extracted by maximum-likelihood fits to the reconstructed D^+ candidate invariant-mass distribution — $M(\pi^+e^+e^-)$ for the signal and validation modes, $M(\pi^+K^+K^-)$ for the normalisation mode. For the sparse signal and validation modes, we use an extended unbinned likelihood,

$$\mathcal{L} = \frac{e^{-\nu}}{N!} \prod_{i=1}^N \sum_c n_c f_c(x_i; \theta), \quad \nu = \sum_c n_c, \quad (7.2)$$

where N is the number of observed candidates, x_i the candidate mass of event i , c runs over the signal and background components with yields n_c and PDFs $f_c(x; \theta)$ (Section 7.3), and ν is the total expected number of events [88]. We use an unbinned fit rather than a simple event count in a mass window because the signal mode contains very few candidates. By binning, or discarding the exact candidate mass of each event, we would throw away information that is not negligible at this sample size.

For the normalisation mode, and for the fits to signal MC used to determine the signal shapes, we instead use a binned likelihood,

$$\mathcal{L} = \prod_{i=1}^{N_{\text{bins}}} \frac{\nu_i^{n_i} e^{-\nu_i}}{n_i!}, \quad (7.3)$$

where n_i is the number of observed candidates in bin i and ν_i the number expected from the fitted model in that bin. In these samples, the number of events is large enough that binning loses no appreciable information, while being considerably faster to fit.

7.3 PDF Functional Forms

The signal and background PDFs adopted for each decay mode and experiment are summarised in Table 7.1, while the functional forms and the physical reasoning behind each choice are described below.

Two-sided Crystal Ball. We model the signal-mode $M(\pi^+e^+e^-)$ shape in both experiments with the two-sided (double) Crystal Ball (CB) function implemented in RooFit as `RooCrystalBall` [89], with a Gaussian core of independent left- and right-hand widths σ_L, σ_R and independent power-law tails:

$$f(m; m_0, \sigma_L, \alpha_L, n_L, \sigma_R, \alpha_R, n_R) = \begin{cases} A_L \left(B_L - \frac{m - m_0}{\sigma_L} \right)^{-n_L}, & \frac{m - m_0}{\sigma_L} < -\alpha_L \\ \exp \left(-\frac{1}{2} \left[\frac{m - m_0}{\sigma_L} \right]^2 \right), & \frac{m - m_0}{\sigma_L} \leq 0 \\ \exp \left(-\frac{1}{2} \left[\frac{m - m_0}{\sigma_R} \right]^2 \right), & \frac{m - m_0}{\sigma_R} \leq \alpha_R \\ A_R \left(B_R + \frac{m - m_0}{\sigma_R} \right)^{-n_R}, & \text{otherwise,} \end{cases} \quad (7.4)$$

with $A_i = (n_i/|\alpha_i|)^{n_i} \exp(-|\alpha_i|^2/2)$ and $B_i = n_i/|\alpha_i| - |\alpha_i|$ ($i = L, R$). This asymmetric form is required because bremsstrahlung photon emission from the two electrons (Section 5.3) shifts events towards lower $M(\pi^+e^+e^-)$, giving the signal-mode lineshape a genuinely asymmetric core in addition to its radiative tail.

Symmetric-core Crystal Ball. The normalisation-mode shape uses the same `RooCrystalBall` [89] form of Eq. (7.4) constructed with a single common width $\sigma_L = \sigma_R = \sigma$, leaving only the tail parameters $\alpha_L, n_L, \alpha_R, n_R$ independent. The normalisation-mode final state contains no electrons, so its $M(\pi^+K^+K^-)$ resolution is close to the intrinsic detector resolution and its core is narrower and more symmetric than in the signal mode. The tails are kept independent to accommodate a small residual asymmetry observed in the signal MC fit.

Cruijff function. The symmetric Cruijff function used for the Belle II validation-mode core, implemented in RooFit through a `RooGenericPdf`, is

$$f(m; m_0, \sigma, \alpha) = \exp\left(-\frac{(m - m_0)^2}{2\sigma^2 + \alpha(m - m_0)^2}\right), \quad (7.5)$$

a Gaussian-like function whose tails widen smoothly with α rather than transitioning to a power law; it is used elsewhere in flavour physics, e.g. Ref [90].

Bifurcated Gaussian. A bifurcated (asymmetric) Gaussian, implemented in RooFit as `RooBifurGauss` [89] and sharing a common mean m_0 with the Cruijff component, is added to describe residual asymmetry in the Belle II validation-mode peak:

$$f(m; m_0, \sigma_L, \sigma_R) = \begin{cases} \exp\left(-\frac{(m - m_0)^2}{2\sigma_L^2}\right), & m < m_0 \\ \exp\left(-\frac{(m - m_0)^2}{2\sigma_R^2}\right), & m \geq m_0. \end{cases} \quad (7.6)$$

For Belle, the same validation-mode peak is instead described by a two-sided CB, as for the signal mode. For each channel and experiment, the signal shape was selected empirically by fitting the signal-MC mass distribution with several candidate forms and retaining the one that described it best. The choice was judged from the pull distribution and the χ^2/ndf shown on the fit figures, with the overlaid fit curve used as a visual cross-check, and the resulting forms therefore differ between the two experiments. Simpler models, such as a single Gaussian or a single-sided Crystal Ball, left significant residuals in the peak and were not used. All retained forms are asymmetric, reflecting the low-mass tail in $M(\pi^+e^+e^-)$ from bremsstrahlung energy loss of the electrons.

Background. We model combinatorial background with a Chebyshev polynomial, $f(x') = \sum_{i=0}^k a_i T_i(x')$, with x' the fit variable linearly mapped onto $[-1, 1]$ and T_i the Chebyshev polynomials ($T_0 = 1$, $T_1 = x'$, $T_{i+1} = 2x'T_i - T_{i-1}$). We use a first-order Chebyshev ($k = 1$) for the signal channel in both experiments and for the Belle validation channel, and a flat (zeroth-order) shape for the Belle II validation channel. For the normalisation channel, we use a second-order Chebyshev ($k = 2$): its larger number of events makes a mild curvature in the background visible and provides enough events to determine the extra coefficient reliably, whereas the smaller signal and validation samples are kept simpler so the background does not absorb genuine signal. All background-shape parameters are left floating in every fit.

Table 7.1: Signal and background PDFs used to fit each decay mode, for the Belle and Belle II samples. The signal-shape parameters are fixed to the values obtained from fits to signal MC (Section 7.4) in every mode except the normalisation mode, where they are left free; all background-shape parameters float in every fit. “Two-sided CB” denotes the double Crystal Ball function of Eq. (7.4) and “symmetric-core CB” the same function with a common Gaussian width.

Decay mode	Signal PDF	Shape	Background PDF
Belle II			
$D^+ \rightarrow \pi^+ K^+ K^-$	Symmetric-core CB	float	2nd-order Cheb.
$D^+ \rightarrow \pi^+ e^+ e^-$	Two-sided CB	fixed	1st-order Cheb.
$D^+ \rightarrow \pi^+ \phi(\rightarrow e^+ e^-)$	Symmetric Cruijff + bifurcated Gaussian with common mean	fixed	flat
Belle			
$D^+ \rightarrow \pi^+ K^+ K^-$	Symmetric-core CB	float	2nd-order Cheb.
$D^+ \rightarrow \pi^+ e^+ e^-$	Two-sided CB	fixed	1st-order Cheb.
$D^+ \rightarrow \pi^+ \phi(\rightarrow e^+ e^-)$	Two-sided CB	fixed	1st-order Cheb.

7.4 Fits to Signal MC

The fits in this section fix the signal-PDF shape parameters used in every subsequent fit to generic MC (Section 7.7) and data (Chapter 9), except the normalisation mode, whose signal shape is left floating throughout.

7.4.1 Signal Mode

We fit the Belle and Belle II signal-mode signal MC samples with the two-sided CB of Eq. (7.4), inclusively over I_{full} (Figure 7.1).

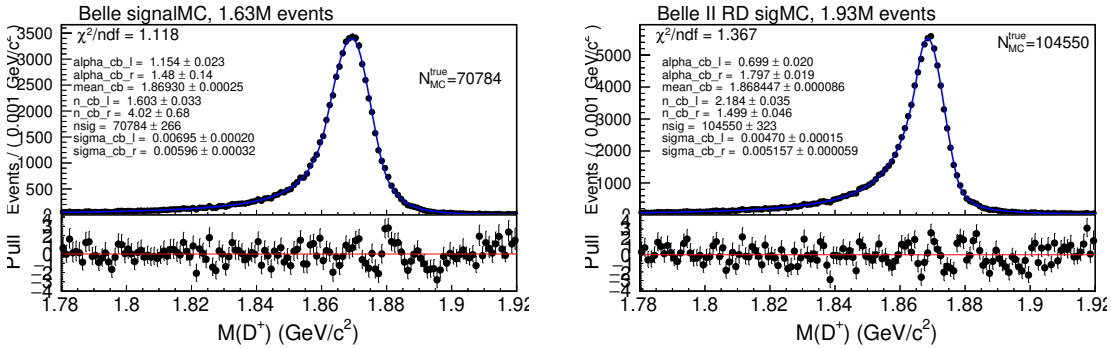


Figure 7.1: Fit to signal MC for the signal mode, Belle (left) and Belle II (right), inclusive over I_{full} .

The signal MC sample also contains mis-reconstructed candidates ($\text{isSignal} \neq 1$), which we refer to as *broken signal*. Some of these candidates are found to peak

under the D^+ mass despite being wrongly reconstructed, and are experimentally indistinguishable from correctly reconstructed signal in the $M(\pi^+e^+e^-)$ distribution. Because it cannot be separated from correctly reconstructed signal, broken signal is included in the fitted signal yield, even though it is excluded from the truth-matched reconstruction efficiency of Section 6.1.2. The same is true of the normalisation channel, so a broken-signal fraction common to both cancels in the branching-fraction ratio, and only the difference between the two fractions biases the result (Section 7.4.3).

We estimate the broken-signal yield with a separate Gaussian fit to the $\text{isSignal} \neq 1$ component (Figure 7.2), finding 1065 ± 109 events for Belle (1.4% of all signal events) and 938 ± 92 events for Belle II (0.9%). Both the fitted broken-signal yields and the truth-matched counts against which they are normalised are evaluated within the final fit region, $1.70 < M(\pi^+e^+e^-) < 1.92 \text{ GeV}/c^2$, whereas Figure 7.2 shows a wider mass range so that the shape of the mis-reconstructed component remains visible. We observe a similar peaking component from mis-reconstructed events in the normalisation channel (Section 7.4.3).

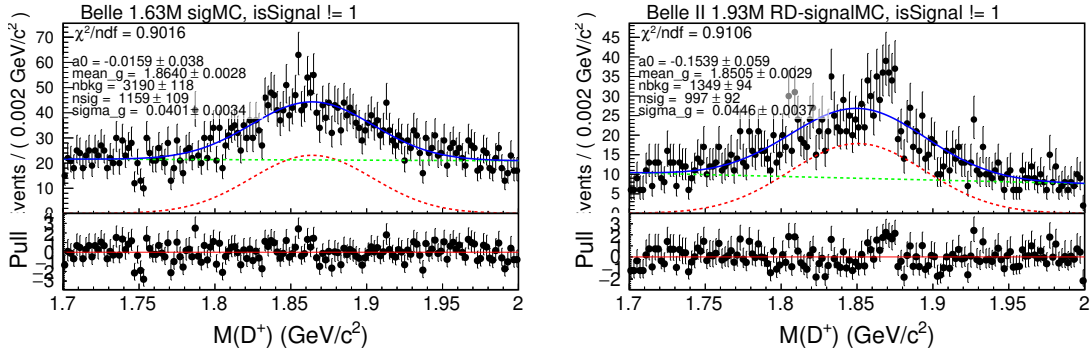


Figure 7.2: $M(\pi^+e^+e^-)$ distribution of broken-signal candidates ($\text{isSignal} \neq 1$) from signal MC, Belle (left) and Belle II (right), with the fitted Gaussian shown as a red dashed curve.

We repeat the signal-PDF fit independently in each $M(e^+e^-)$ interval I_1 , I_2 , I_3 (Figure 7.3), since the signal shape varies across intervals. We fit the mis-reconstructed signal component using the procedure described above with the fits collected in Appendix D.2, and the resulting fractions, summarised in Table 7.2, which range from 1.2–1.6% for Belle and 0.7–1.0% for Belle II.

7.4.2 Validation Mode

We fit the Belle signal MC with the two-sided CB and the Belle II signal MC with a sum of a Cruijff function and a bifurcated Gaussian, per Table 7.1 (Figure 7.4).

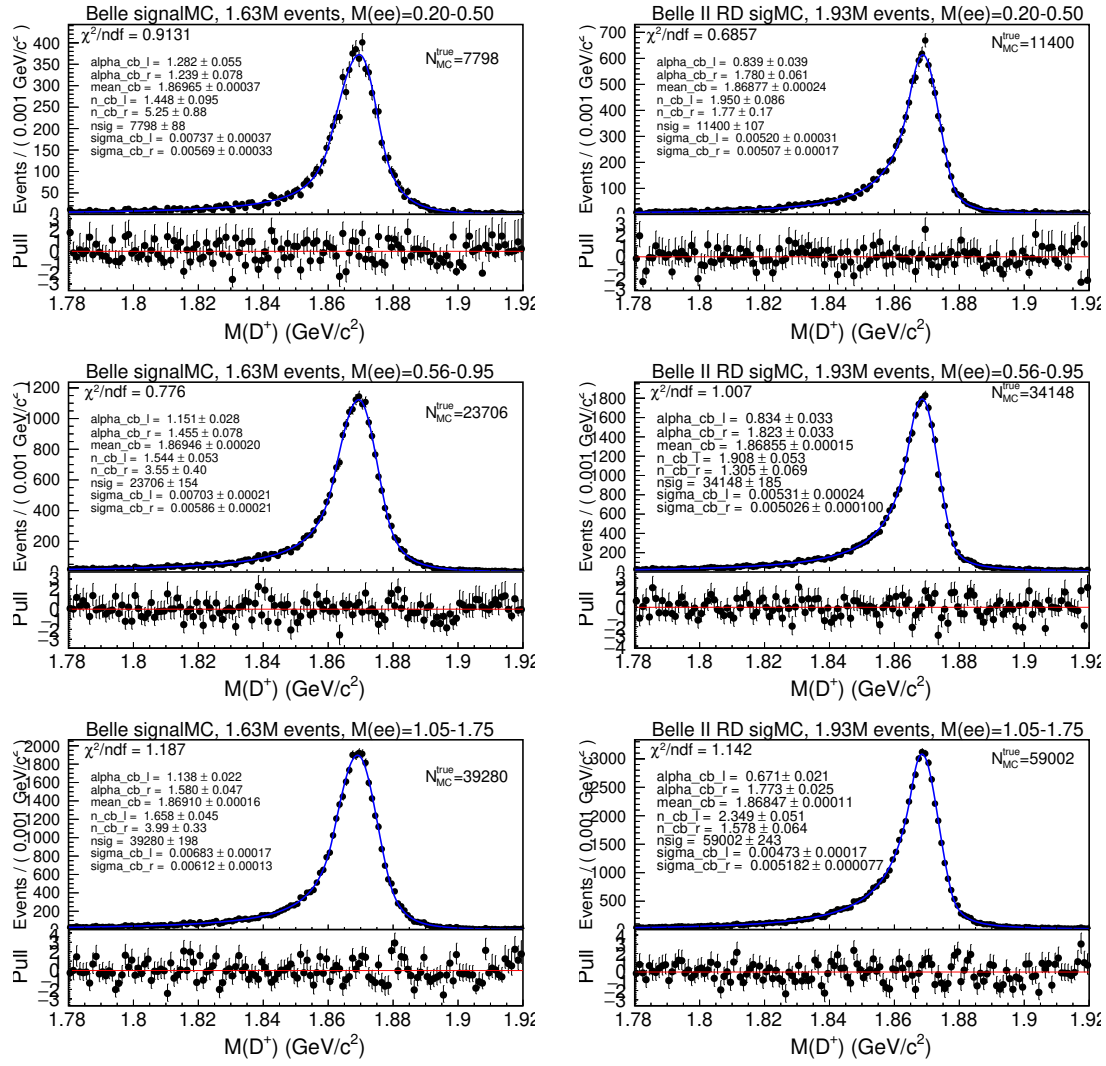


Figure 7.3: Fits to Belle (left) and Belle II (right) signal MC in the $M(e^+e^-)$ intervals I_1 , I_2 , I_3 (top to bottom).

Table 7.2: Broken-signal fractions (f_{broken}) from signal MC for the signal mode. N_{true} is the number of truth-matched signal events within $1.7 < M(\pi^+e^+e^-) < 1.92 \text{ GeV}/c^2$, and N_{broken} is the fitted broken-signal yield. The inclusive and summed-interval yields differ at the per cent level because we fit them independently.

$M(e^+e^-)$ interval [GeV/c ²]	Belle			Belle II		
	N_{true}	N_{broken}	f_{broken} [%]	N_{true}	N_{broken}	f_{broken} [%]
I_{full}	72952	1065 ± 109	1.4	107348	938 ± 92	0.9
I_1	8091	108 ± 18	1.3	11693	79 ± 30	0.7
I_2	24518	296 ± 33	1.2	35224	302 ± 72	0.9
I_3	40343	671 ± 120	1.6	60431	611 ± 67	1.0

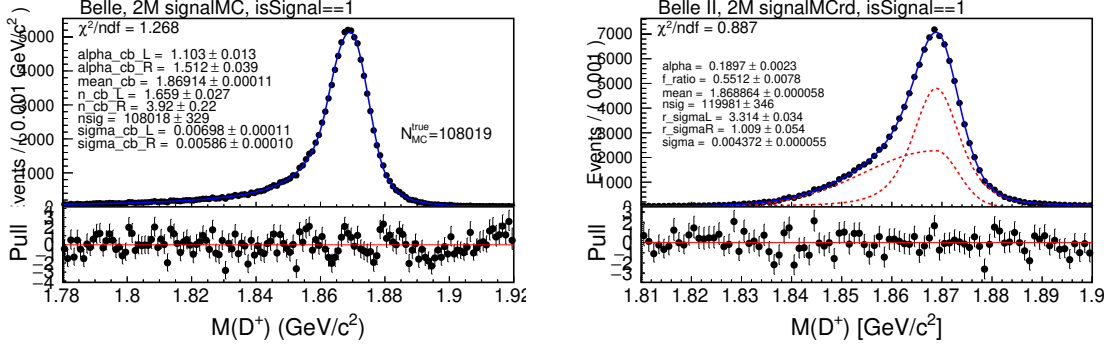


Figure 7.4: Fits to signal MC for the validation mode, Belle (left) and Belle II (right).

7.4.3 Normalisation Mode Cross-Check

Since the normalisation-mode signal shape floats freely in the fits to data, we do not fit a dedicated signal MC sample to fix it. We tested fixing individual shape parameters from a signal MC or generic-MC fit, but found that none could be fixed without visibly biasing the residual distribution of the fitted yield. Therefore, we leave the full signal shape floating in every normalisation-mode fit. As a cross-check of the symmetric-core CB choice, we also fit the truth-matched signal component within the generic MC sample (Figure 7.5).

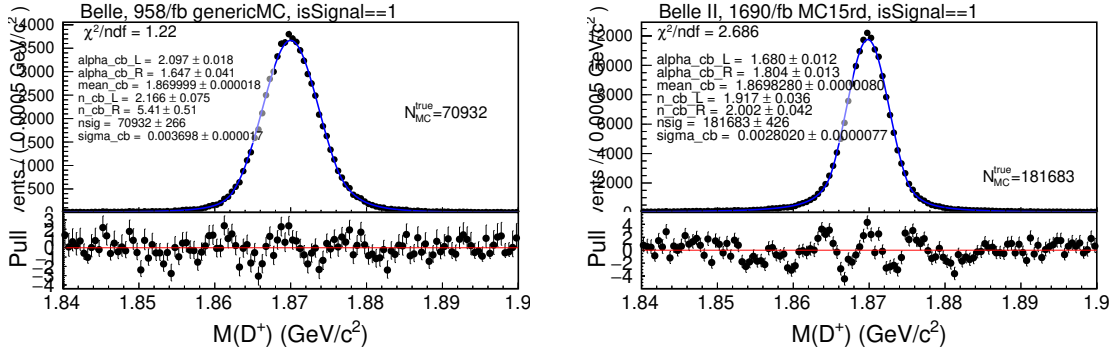


Figure 7.5: Fit to the truth-matched signal component within generic MC for the normalisation mode, Belle (left) and Belle II (right).

The same generic MC sample also exhibits a peaking, mis-reconstructed component analogous to the signal-mode broken signal, which we quantify in Section 7.7.3. Because the branching fraction is measured as a yield-and-efficiency ratio between the signal and normalisation channels, a broken-signal fraction common to both channels would cancel; it is only the difference between the two fractions that biases the ratio. This residual difference is not corrected for directly but is instead propagated as the mis-reconstruction signal systematic uncertainty of Section 8.2.4 (Chapter 8).

7.5 Simultaneous Fit Framework

For the signal and validation modes, we perform a simultaneous fit to the Belle and Belle II samples, in which the branching fraction $\mathcal{B}_{\pi\pi e}^j$ of Eq. (7.1) is a single shared

free parameter, entering the signal yield of each experiment through

$$n_{\pi ee,k}^j = \mathcal{B}_{\pi ee}^j \frac{\epsilon_{\pi ee,k}^j}{\epsilon_{\pi KK,k}} \frac{N_{\pi KK,k}}{\mathcal{B}_{\pi KK}}, \quad (7.7)$$

the inverse of Eq. (7.1), where j labels the $M(e^+e^-)$ interval, and k the experiment. The simultaneous likelihood is the product of the per-experiment likelihoods of Eq. (7.2),

$$\mathcal{L}_{\text{sim}}^j = \mathcal{L}_{\text{Belle}}^j \cdot \mathcal{L}_{\text{Belle II}}^j, \quad (7.8)$$

with $\mathcal{B}_{\pi ee}^j$ common to both factors while the background yields and shape parameters remain independent per experiment. Because the two samples are statistically independent, the probability of observing both is the product of their individual likelihoods, giving Eq. (7.8). Fitting them jointly with a shared $\mathcal{B}_{\pi ee}(I)$ lets both experiments contribute to a single measurement, each according to its own yield and efficiency. This differs from the usual approach of measuring $\mathcal{B}_{\pi ee}(I)$ separately in each experiment and combining the two results with a weighted average, which assumes Gaussian uncertainties and breaks down for the very low signal yields encountered here. The joint fit avoids this by combining the data before extracting a single value. We perform this simultaneous fit both inclusively, over I_{full} , and independently within each $M(e^+e^-)$ interval I_1, I_2, I_3 .

The normalisation-mode yields $N_{\pi KK,k}$ that enter Eq. (7.7) are instead obtained from independent binned fits to the Belle and Belle II samples, with no parameter shared between the two.

7.6 Signal Significance and Limit-Setting Procedure

For each fit, we evaluate the statistical significance of a possible signal from the likelihood-ratio test statistic

$$S = \sqrt{-2 \ln(\mathcal{L}_0/\mathcal{L}_{\text{max}})}, \quad (7.9)$$

where $\mathcal{L}_{\text{max}}$ is the likelihood of the best fit and $\mathcal{L}_0$ that of the background-only hypothesis, obtained by fixing $\mathcal{B}_{\pi ee}^j = 0$. Under Wilks' theorem, $-2 \ln(\mathcal{L}_0/\mathcal{L}_{\text{max}})$ is asymptotically distributed as χ^2 with one degree of freedom, so S can be expressed directly as the corresponding Gaussian-equivalent significance in units of standard deviations. We evaluate the significance without systematic uncertainties. It tests the background-only hypothesis, for which the systematic contributions are negligible compared with the statistical fluctuation of the yield at this sample size. We report the resulting values in Chapter 9.

The significance determines how the result is quoted. Where it reaches the conventional threshold of five standard deviations, we quote a measured branching fraction, and otherwise we set an upper limit following the procedure described below.

We set an upper limit on the branching fraction at the 90% confidence level using a frequentist Neyman construction [91] based on pseudo-experiments. For each assumed true branching fraction, we generate a large ensemble of pseudo-experiments from the fitted probability density functions of Section 7.3, refit each one as we do the data, and record the spread of fitted branching fractions $\hat{\mathcal{B}}_{\pi ee}^j$ it produces. Repeating this over a range of assumed true values maps out the

confidence belt: the band of fitted values expected for each true value. The one-sided 90% CL upper limit is the true branching fraction for which only 10% of pseudo-experiments yield a fitted value below the one observed in data. We read it directly as the point where the belt boundary meets the observed $\hat{\mathcal{B}}_{\pi ee}^j$. Since a negative branching fraction is unphysical, we restrict the construction to the physical region $\mathcal{B}_{\pi ee}^j \geq 0$. We still report the fitted central value even when it is negative, but we read the upper limit from the belt at $\hat{\mathcal{B}}_{\pi ee}^j = 0$ rather than at the unphysical negative value, which yields a weaker, conservative limit. Since a fitted deficit carries no discovery information, we set $S \equiv 0$ whenever the unconstrained best fit of the branching fraction is negative. The construction is performed separately for the inclusive range I_{full} and for each $M(e^+e^-)$ interval I_1, I_2, I_3 .

Systematic uncertainties are incorporated into this construction in two ways, depending on the source. Uncertainties that can be represented as fluctuations of the fixed fit inputs are folded into the generation of the pseudo-experiments by varying those inputs within their uncertainties before each ensemble is produced. The remaining uncertainties are applied afterwards, as multiplicative or additive variations to the branching fraction obtained from the construction. We quote the upper limit both without and with systematic uncertainties included. The individual sources, their classification into these two categories, and their magnitudes are detailed in Chapter 8.

7.7 Closure Tests on Generic MC

Before applying the fit to data, we verify that it extracts the branching fraction without bias by running the full procedure on generic MC. The signal-shape parameters are fixed to the signal-MC fits of Section 7.4 in the signal and validation modes, exactly as they will be for data, and only the yields and background shapes are free.

7.7.1 Signal Mode

Because the available generic MC luminosity does not match the analysed data in either experiment, we resample each sample with repetition to the expected data luminosity before fitting — the Belle sample upsampled, the Belle II sample downsampled. The generic MC contains no $D^+ \rightarrow \pi^+ e^+ e^-$ signal, so an unbiased fit must return a branching fraction consistent with zero. We perform the simultaneous Belle and Belle II fit described in Section 7.5, both inclusively over I_{full} (Figure 7.6) and in the three $M(e^+e^-)$ intervals (Figure 7.7). In every case the fitted branching fraction is compatible with zero within its uncertainty, confirming that the fit introduces no spurious signal.

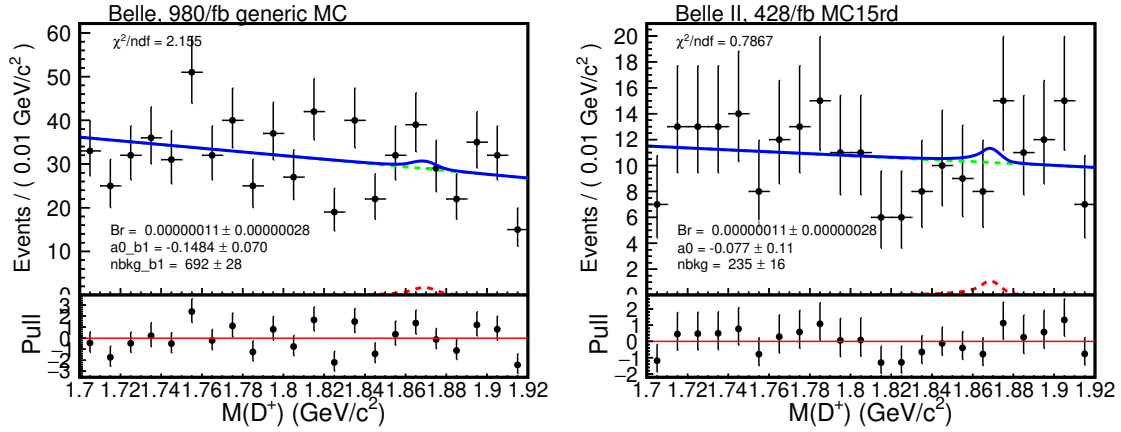


Figure 7.6: Simultaneous fit to the Belle (left) and Belle II (right) generic MC in the signal mode, inclusive over I_{full} . Only the branching fraction is shared between the two samples.

7.7.2 Validation Mode

The validation mode $D^+ \rightarrow \pi^+ \phi (\rightarrow e^+ e^-)$ is present in the generic MC, so here the closure test checks that the fit recovers the known number of signal events rather than returning zero. The signal shapes are fixed to the signal-MC fits of Section 7.4.2, while the background parameters float: a first-order Chebyshev for Belle and a flat background for Belle II. Fitting each experiment with its full available generic MC sample (Figure 7.8) gives yields consistent with the truth-matched counts within their uncertainties, as summarised in Table 7.3.

Table 7.3: Signal yields for the validation mode in generic MC. N_{fit} is the fitted signal yield and N_{true} the truth-matched count.

Dataset	N_{fit}	N_{true}
Belle	26.4 ± 6.5	25
Belle II	83.8 ± 10	88

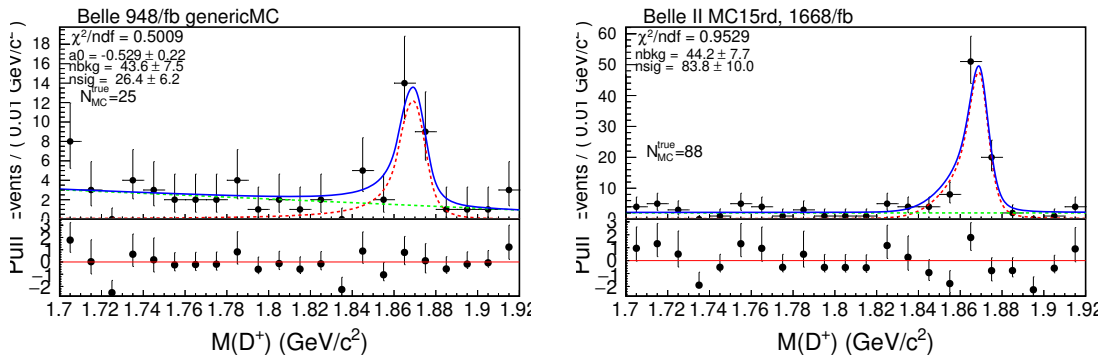


Figure 7.8: Fits to the generic MC in the validation mode for Belle (left) and Belle II (right), using the full available generic MC samples.

7.7. CLOSURE TESTS ON GENERIC MC

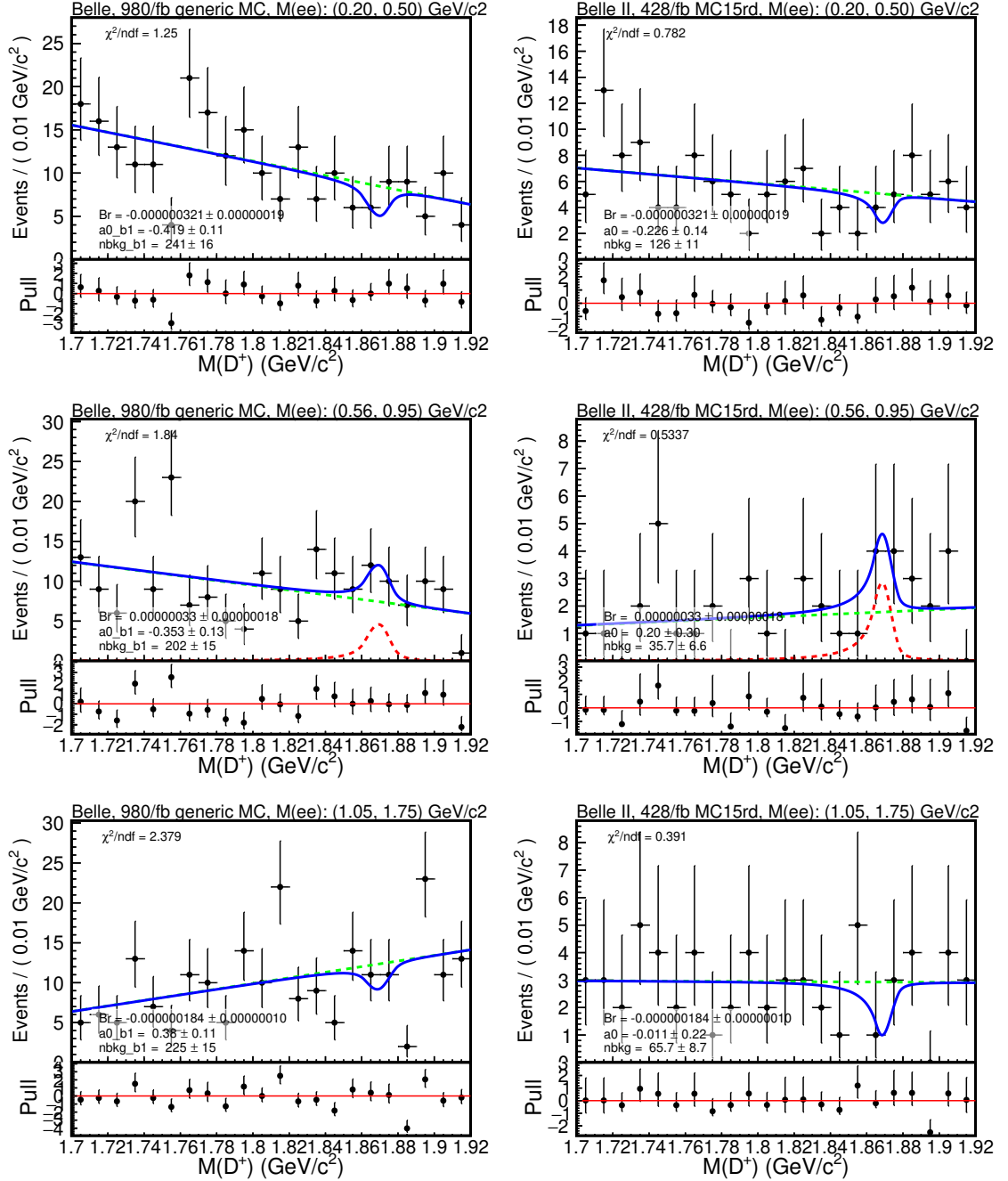


Figure 7.7: Simultaneous fits to the Belle (left) and Belle II (right) generic MC in the signal mode, in the $M(e^+e^-)$ intervals I_1 , I_2 , I_3 (top to bottom).

We then repeat the measurement as a simultaneous Belle and Belle II fit with the branching fraction as the common free parameter, on samples resampled to the expected data luminosity (Figure 7.9), reproducing the procedure that will be applied to data.

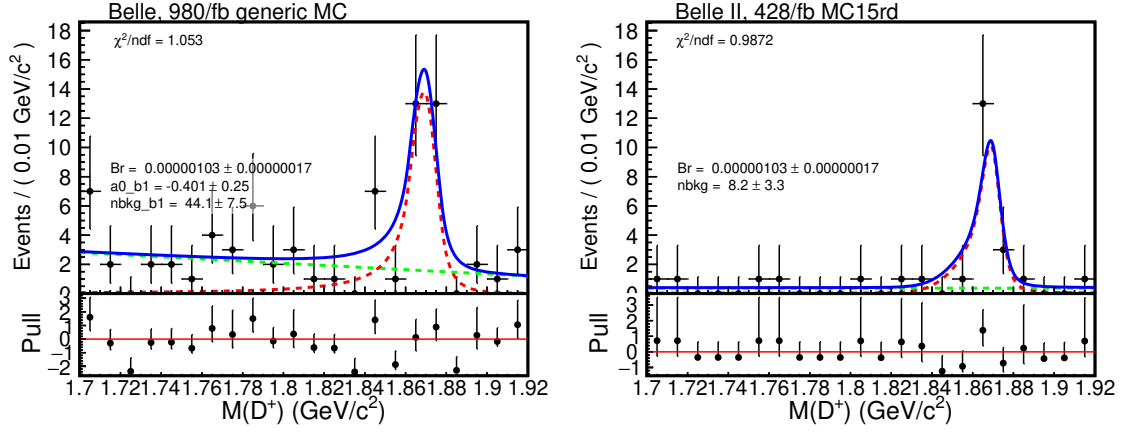


Figure 7.9: Simultaneous fit to the Belle (left) and Belle II (right) generic MC in the validation mode. Only the branching fraction is shared between the two samples.

7.7.3 Normalisation Mode

For the normalisation mode, we fit each generic MC sample with a symmetric-core CB for the peaking signal and a second-order Chebyshev polynomial for the background, with all parameters free (Figure 7.10). The fitted yields agree with the truth-matched counts to within 1.5% for Belle and 1.1% for Belle II (ratio of N_{fit} and N_{true} in Table 7.4).

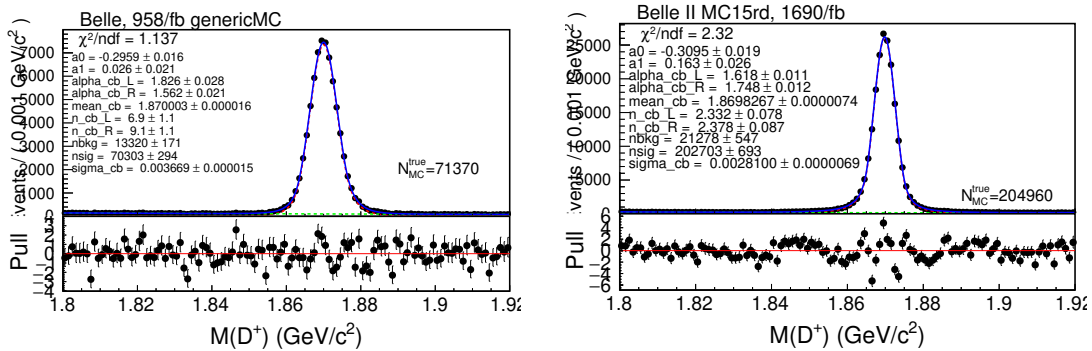


Figure 7.10: Fit to the generic MC in the normalisation mode for Belle (left) and Belle II (right).

For Belle, the available generic MC luminosity (958 fb^{-1} , Section 4.2) is close to the analysed data (980 fb^{-1}), and the sample is used without resampling. For Belle II, the available generic MC corresponds to roughly four times the analysed data luminosity (1690 fb^{-1} vs. 428 fb^{-1}). As a cross-check, we additionally fit a version resampled with repetition down to the data luminosity, alongside the full, non-resampled sample.

In the Belle II case, the fit to the larger sample clearly exhibits artefacts in the pull plot, whereas at a sample size comparable to the data (Figure 7.11), these are

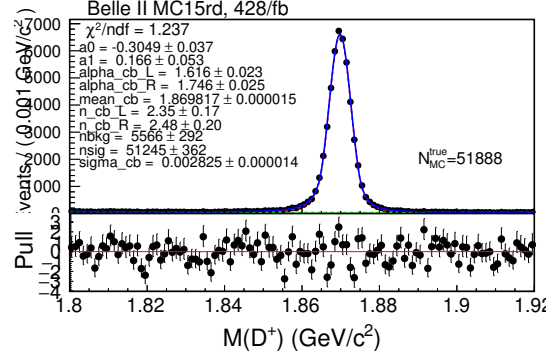


Figure 7.11: Fit to generic MC in the normalisation mode for Belle II, with sample size matched to data.

almost completely smoothed out by statistical noise. We further investigate these artefacts, together with the residual difference between the fitted and truth-matched yields, as a potential PDF mismodelling, in Section 7.8.5.

As in the signal mode, the generic MC sample contains a peaking broken-signal component ($\text{isSignal} \neq 1$), clearly visible for Belle II and only slightly for Belle (Figure 7.12).

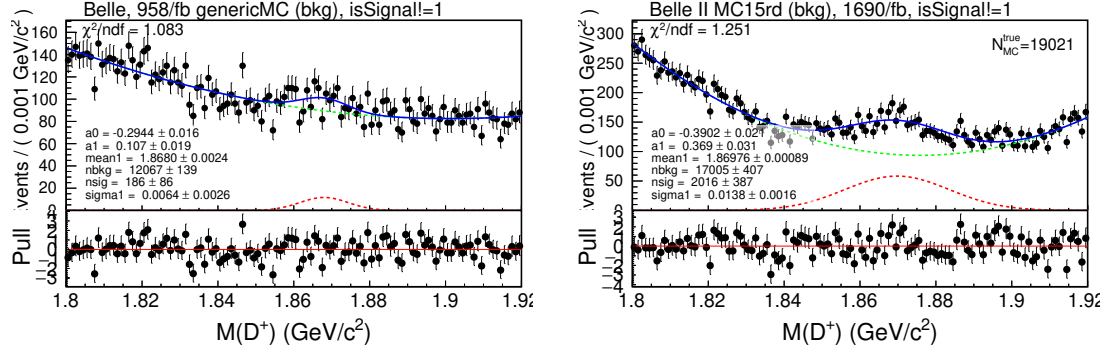


Figure 7.12: $M(\pi^+K^+K^-)$ distributions for generic MC events with $\text{isSignal} \neq 1$, Belle (left) and Belle II (right). The broken-signal component is shown as a red dashed curve, the residual background as a green dashed curve.

We estimate its yield with a Gaussian fit, resulting in 186 ± 86 events for Belle (0.3% of all signal events) and 2016 ± 387 events for Belle II (1.0%), summarised in Table 7.4. These fractions differ from the corresponding signal-mode values of Section 7.4.1 (1.6% and 0.95%), and it is this difference, which does not cancel in the branching-fraction ratio, that is propagated as the mis-reconstruction systematic in Chapter 8.

7.8 Toy MC Studies and Linearity Tests

The closure tests of Section 7.7 verify the fit on a single realisation of the generic MC. To probe its behaviour statistically, we repeat the fit many times on pseudo-data — *toy MC* samples — generated from the fitted probability density functions and refitted with the same model. We present here the studies of the simultaneous fit; the corresponding results for the independent Belle and Belle II fits, and for the remaining $M(e^+e^-)$ intervals, are collected in Appendix C.

Table 7.4: Signal yields for the normalisation mode in generic MC. N_{fit} is the fitted signal yield, N_{true} the truth-matched count, and N_{broken} the fitted broken-signal yield. The last column gives the broken-signal fraction $N_{\text{broken}}/(N_{\text{true}} + N_{\text{broken}})$.

Dataset	N_{fit}	N_{true}	N_{broken}	Fraction [%]
Belle	70303 ± 294	71370	186 ± 86	0.3
Belle II	202703 ± 693	204960	2016 ± 387	1.0

For each toy we form the pull of a fitted parameter,

$$\text{pull} = \frac{N_{\text{fit}} - N_{\text{gen}}}{\sigma(N_{\text{fit}})}, \quad (7.10)$$

shown here for a signal yield, where N_{gen} is the value used to generate the toy and $\sigma(N_{\text{fit}})$ the uncertainty returned by the fit. An unbiased fit with correctly estimated uncertainties yields a pull distribution that is Gaussian with mean zero and unit width. A non-zero mean indicates a bias in the extracted value, and a width different from unity indicates that the fit misestimates its own uncertainty.

A linearity test extends this to a range of generated signal yields: pseudo-data are produced at several values of the true signal yield or branching fraction and refitted, and the fitted values are plotted against the generated ones. A fit that responds correctly over the relevant range yields a straight line with unit slope and zero intercept.

In all studies, the PDF parameters are fixed or floated exactly as they are in the fit to data, and the toy sample sizes match the expected data sizes. We perform these studies with the `RoomCStudy` module of `RooFit` [89] and without any systematic uncertainties. By construction, the pseudo-data are drawn from the same PDFs used to fit them, so these studies test the statistical behaviour of the fitting procedure but cannot reveal a mismodelling of the underlying distribution. This limitation motivates the complementary bootstrapping study of Section 7.8.5, which resamples from data-like events rather than the fit’s own PDF.

7.8.1 Signal Mode

For the signal mode, we generate 10 000 toy samples from the PDFs fitted to the generic MC and refit each with the simultaneous Belle + Belle II fit of Section 7.5, with the branching fraction as the shared free parameter. The resulting pull distributions are summarised in Table 7.5 and shown in Figure 7.13. All parameters have pull widths compatible with unity, confirming that the fit estimates its uncertainties correctly. The branching-fraction pull has a mean of -0.06 ± 0.01 , indicating a small negative bias. As it corresponds to a shift of a few per cent of the statistical uncertainty, we neglect it and assign no associated systematic uncertainty.

The linearity of the fit is tested by generating toy samples over a range of true branching fractions and comparing the fitted values with the generated ones (Figure 7.14). The simultaneous fit responds linearly, with a fitted slope $k = 0.999 \pm 0.003$ and intercept $n = (2 \pm 10) \times 10^{-10}$, both consistent with the ideal values of unity and zero (Table 7.6). The same studies performed on the Belle and Belle II samples

7.8. TOY MC STUDIES AND LINEARITY TESTS

Table 7.5: Means μ and widths σ of the pull distributions from 10 000 toy MC studies of the simultaneous fit in the signal mode. N_{bkg} and a_0 denote the background yield and the first-order Chebyshev coefficient, with the suffix indicating the Belle sample.

Fit parameter	μ	σ
$\mathcal{B}_{\pi ee}$	-0.06 ± 0.01	1.01 ± 0.01
N_{bkg}	-0.03 ± 0.01	0.99 ± 0.01
$N_{\text{bkg}}^{\text{B}}$	-0.03 ± 0.01	0.991 ± 0.007
a_0	-0.01 ± 0.01	1.01 ± 0.01
a_0^{B}	0.00 ± 0.01	0.994 ± 0.007

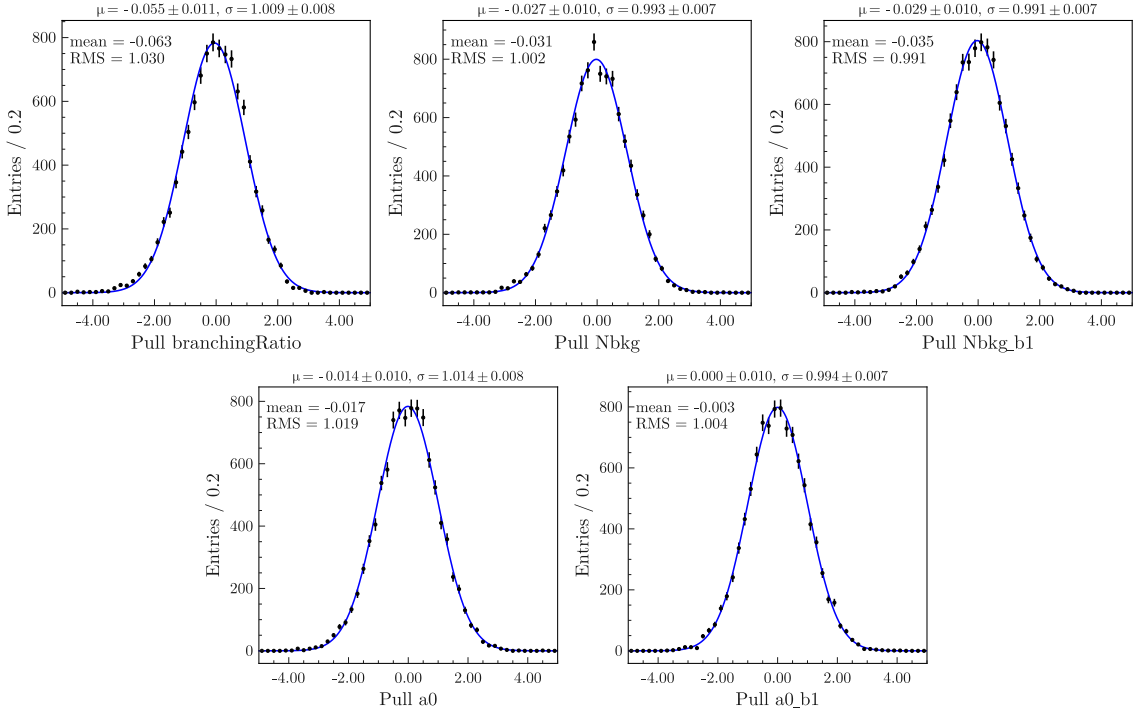


Figure 7.13: Pull distributions from 10000 toy MC studies of the simultaneous fit in the signal mode, generated at $\mathcal{B}_{\pi ee} = 0$. Black points and green bands represent the averages and standard deviations of the fitted branching fractions.

independently also show a linear response, with a minimal bias that is much smaller than the statistical uncertainty and is therefore neglected.

Table 7.6: Slope k and intercept n of the linearity tests of the simultaneous and independent fits in the signal mode, from 10 000 toy MC studies. The intercept is given in units of the branching fraction for the simultaneous fit and in units of signal yield for the independent fits.

	Simultaneous	Belle	Belle II
k	0.999 ± 0.003	1.000 ± 0.002	1.000 ± 0.003
n	$(2 \pm 10) \times 10^{-10}$	-0.04 ± 0.03	-0.04 ± 0.02

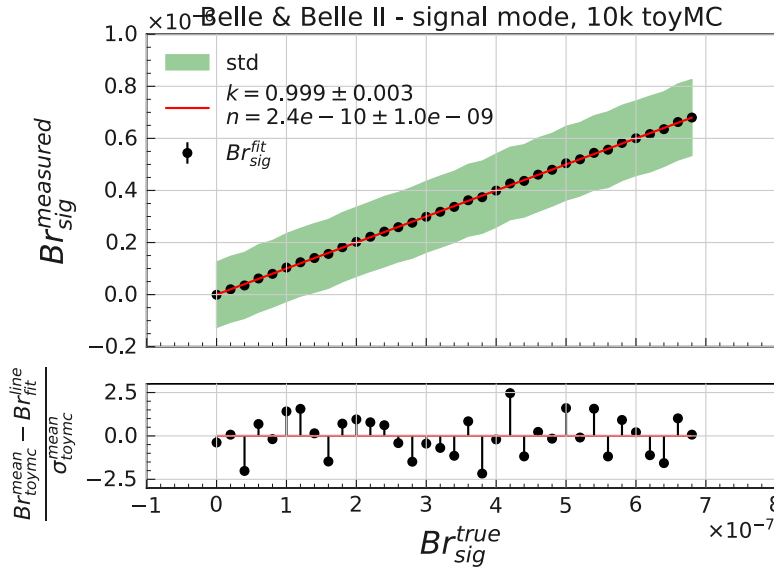


Figure 7.14: Linearity test of the simultaneous fit in the signal mode: fitted against generated branching fraction, with the fitted straight line overlaid.

As a cross-check, we repeated these studies with a binned fit in place of the unbinned one, and found a larger bias for both the signal and validation modes in each experiment. This supports the choice of an unbinned fit for the sparse signal and validation modes made in Section 7.5.

7.8.2 Signal Mode in $M(e^+e^-)$ Intervals

Since the branching fraction is also extracted independently in each $M(e^+e^-)$ interval, we repeat the toy MC and linearity studies separately in I_1 , I_2 and I_3 . The pull distributions of all free parameters follow a normal distribution with mean close to zero and unit width in every interval; the results for I_1 are given in Table 7.7 and Figure 7.15, and those for I_2 and I_3 are consistent with them. As in the inclusive case, the branching-fraction pull shows a small negative bias, of the order of a tenth of the statistical uncertainty, which we neglect. We do observe a slight deviation from a normal distribution in the left tail of the branching-fraction pull, most visible in I_3 . This has no additional impact on the UL determination. The Neyman

7.8. TOY MC STUDIES AND LINEARITY TESTS

construction of Section 7.6 is built from the actual distribution of fitted branching fractions and so inherently takes any such departure from normality into account.

Table 7.7: Means μ and widths σ of the pull distributions from 10 000 toy MC studies of the simultaneous fit in the signal mode, in the interval I_1 ($0.2 < M(e^+e^-) < 0.5 \text{ GeV}/c^2$).

Fit parameter	μ	σ
$\mathcal{B}_{\pi ee}$	-0.08 ± 0.01	1.01 ± 0.01
N_{bkg}	-0.05 ± 0.01	0.98 ± 0.01
$N_{\text{bkg}}^{\text{Belle}}$	-0.04 ± 0.01	0.99 ± 0.01
a_0	-0.04 ± 0.01	1.02 ± 0.01
a_0^{Belle}	-0.03 ± 0.01	1.01 ± 0.01

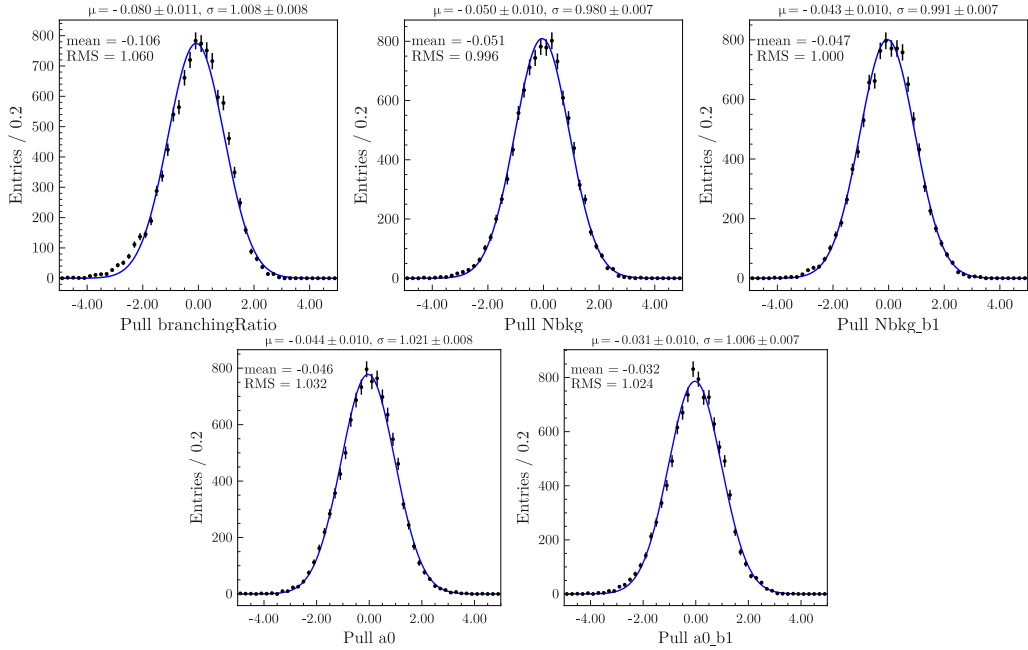


Figure 7.15: Pull distributions from 10 000 toy MC studies of the simultaneous fit in the signal mode, in the interval I_1 , generated at $\mathcal{B}_{\pi ee} = 0$.

The linearity tests show a linear response in all three intervals (Table 7.8), with the I_1 result shown in Figure 7.16. The intercept is consistent with zero in I_1 , while in I_2 and I_3 it deviates slightly. In both cases, the offset remains small compared with the statistical uncertainty on the branching fraction and is neglected.

Table 7.8: Slope k and intercept n of the linearity tests of the simultaneous fit in the signal mode, in the three $M(e^+e^-)$ intervals, from 10 000 toy MC studies. The intercept is given in units of the branching fraction.

	I_1	I_2	I_3
k	1.003 ± 0.003	1.004 ± 0.002	1.000 ± 0.003
n	$(-11 \pm 7) \times 10^{-10}$	$(14 \pm 4) \times 10^{-10}$	$(6.1 \pm 6.0) \times 10^{-10}$

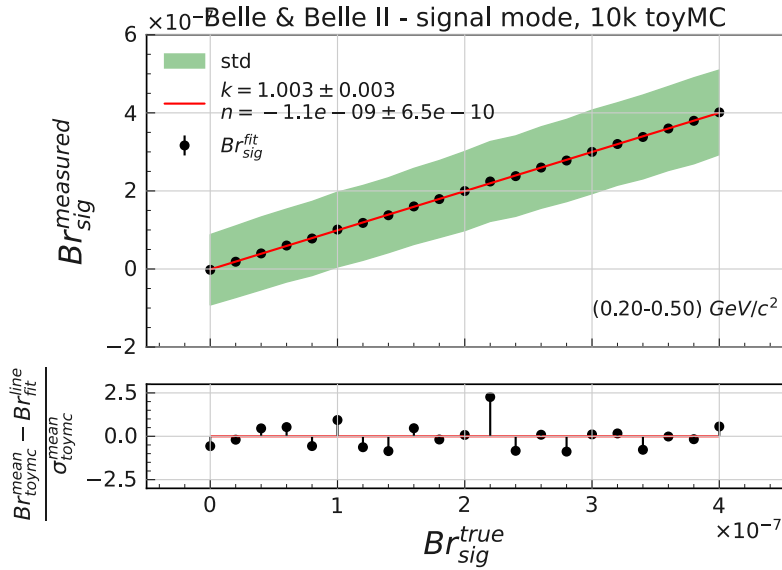


Figure 7.16: Linearity test of the simultaneous fit in the signal mode, in the interval I_1 : fitted against generated branching fraction. The points and bands show the means and standard deviations of the fitted branching fractions from 10 000 toy MC studies.

7.8.3 Validation Mode

Unlike the signal mode, the validation mode contains a measurable signal, so the pull-study toy samples are generated at the branching fraction measured in the validation channel (Section 9.2). The validation channel lies outside the signal region and was examined throughout the analysis, precisely so that the procedure could be tested on an observed signal.

The pull distributions from 10 000 toy samples of the simultaneous fit are summarised in Table 7.9 and shown in Figure 7.17. The branching-fraction pull is close to normal, with a mean of -0.05 ± 0.01 and unit width; as in the signal mode, this residual bias is small compared with the statistical uncertainty and is neglected. The background-shape parameter a_0^{Belle} shows a larger deviation from normality. This is a consequence of the very small Belle background sample in this mode: the background PDF can occasionally take negative values, causing individual fits to fail to converge. Since the branching fraction and both background yields remain well behaved, this does not affect the extracted result. The same behaviour is observed when the two experiments are fitted independently.

7.8. TOY MC STUDIES AND LINEARITY TESTS

Table 7.9: Means μ and widths σ of the pull distributions from 10 000 toy MC studies of the simultaneous fit in the validation mode, generated at the branching fraction measured in data, $\mathcal{B} = 1.75 \times 10^{-6}$.

Fit parameter	μ	σ
$\mathcal{B}$	-0.05 ± 0.01	1.00 ± 0.01
N_{bkg}	-0.10 ± 0.01	0.95 ± 0.01
$N_{\text{bkg}}^{\text{B}}$	-0.10 ± 0.01	1.00 ± 0.01

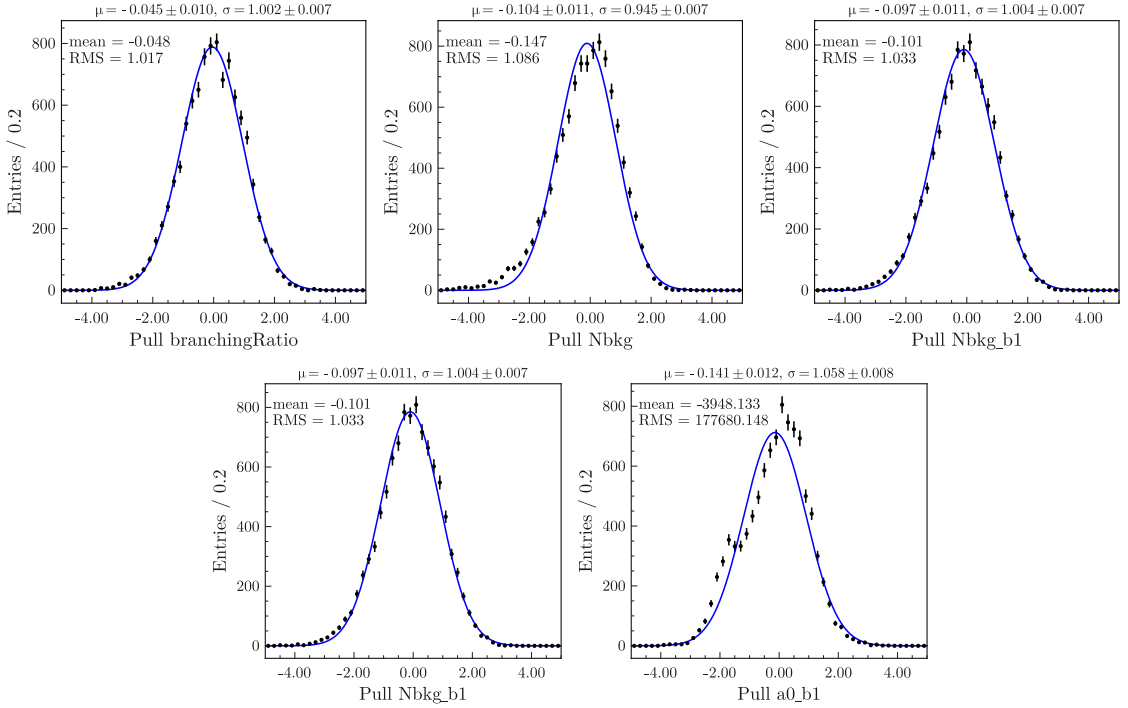


Figure 7.17: Pull distributions from 10 000 toy MC studies of the simultaneous fit in the validation mode, generated at $\mathcal{B} = 1.75 \times 10^{-6}$.

The linearity test likewise shows a good linear response (Figure 7.18), with a slope $k = 1.001 \pm 0.001$ and an intercept $n = (4.7 \pm 1.0) \times 10^{-9}$ (Table 7.10). Although the intercept is not consistent with zero, the corresponding bias is only a few per cent of the statistical uncertainty on the branching fraction, and is therefore neglected.

Table 7.10: Slope k and intercept n of the linearity tests of the simultaneous and independent fits in the validation mode, from 10 000 toy MC studies. The intercept is given in units of the branching fraction for the simultaneous fit and in units of signal yield for the independent fits.

	Simultaneous	Belle	Belle II
k	1.001 ± 0.001	1.000 ± 0.001	0.999 ± 0.002
n	$(4.7 \pm 1.0) \times 10^{-9}$	0.4 ± 0.03	-0.4 ± 0.01

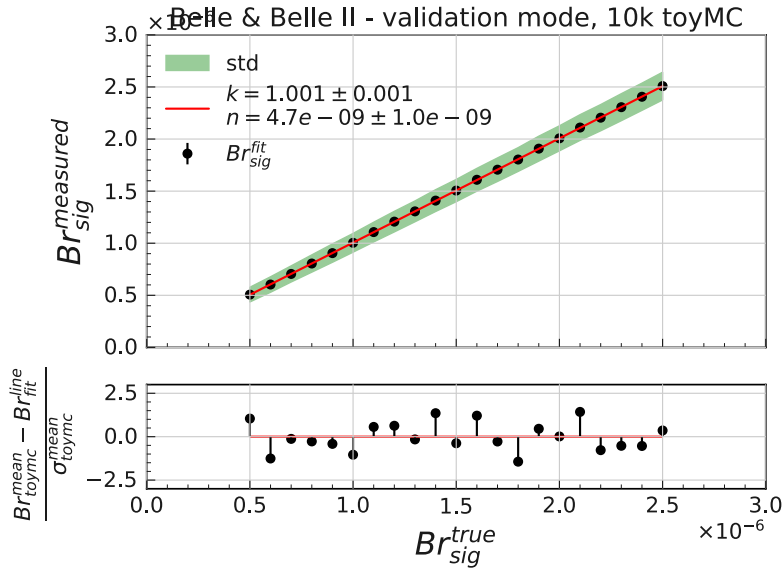


Figure 7.18: Linearity test of the simultaneous fit in the validation mode: fitted against the generated branching fraction. The points and bands show the means and standard deviations of the fitted branching fractions from 10 000 toy MC studies.

7.8.4 Normalisation Mode

We first validate the normalisation-mode yield extraction with toy MC studies, generating 10 000 toy samples from the fitted PDFs and refitting each with the nominal model. The pull distributions remain normal, but the Crystal Ball tail parameters α_L and α_R are poorly constrained by the fit, with pull widths of 0.62–0.64 for Belle and 0.73 for Belle II (Tables 7.11 and 7.12), 30–40% below the ideal unit width. This does not propagate to the signal-yield pull, whose width remains close to unity (1.006 for Belle, 1.010 for Belle II). The signal-yield pull has a mean of 0.31 ± 0.01 for Belle and 0.24 ± 0.01 for Belle II, corresponding to 116 and 78

7.8. TOY MC STUDIES AND LINEARITY TESTS

events, or about 0.1% of the respective yields. This is an order of magnitude below the normalisation PDF mismodelling bias of Section 7.8.5 and is neglected.

The linearity test shows a good linear response for both experiments (Figure 7.19), with a bias negligible compared with the signal-yield uncertainty.

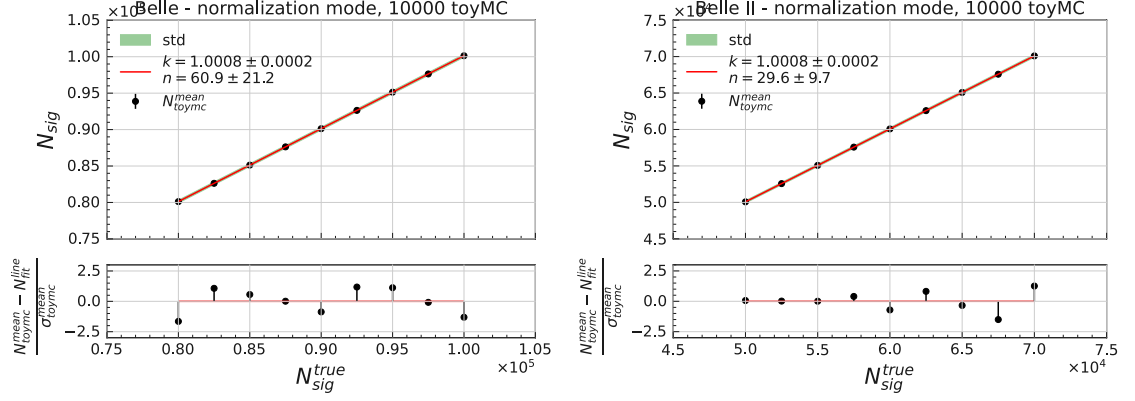


Figure 7.19: Linearity test for the normalisation mode, Belle (left) and Belle II (right). The points and bands show the means and standard deviations of the fitted yields from 10 000 toy MC studies.

Table 7.11: Means μ and widths σ of the pull distributions from 10 000 toy MC studies of the normalisation-mode fit for Belle.

Fit parameter	μ	σ
Signal yield	0.31 ± 0.01	1.006 ± 0.007
Background yield	-0.41 ± 0.01	0.999 ± 0.007
a_0	-0.27 ± 0.01	1.007 ± 0.007
a_1	0.28 ± 0.01	0.989 ± 0.007
m_0	0.07 ± 0.01	0.964 ± 0.007
α_L	0.884 ± 0.006	0.624 ± 0.004
α_R	0.976 ± 0.006	0.638 ± 0.005
n_L	-0.80 ± 0.01	1.005 ± 0.008
n_R	-0.88 ± 0.01	0.999 ± 0.008
σ	0.55 ± 0.01	0.905 ± 0.006

7.8.5 Normalisation Mode: PDF Mismodelling

The toy studies above generate pseudo-data from the fitted PDFs and show the estimator to be effectively unbiased. However, by construction they cannot reveal a mismatch between the fit model and the true distribution. The normalisation mode with its very large signal yield, of order 10^5 events, makes even a small imperfection

Table 7.12: Means μ and widths σ of the pull distributions from 10 000 toy MC studies of the normalisation-mode fit for Belle II.

Fit parameter	μ	σ
Signal yield	0.24 ± 0.01	1.010 ± 0.007
Background yield	-0.27 ± 0.01	1.021 ± 0.007
a_0	-0.09 ± 0.01	1.007 ± 0.007
a_1	0.19 ± 0.01	0.988 ± 0.007
m_0	0.078 ± 0.009	0.924 ± 0.006
α_L	0.330 ± 0.007	0.726 ± 0.005
α_R	0.384 ± 0.007	0.726 ± 0.005
n_L	-0.39 ± 0.01	0.962 ± 0.007
n_R	-0.41 ± 0.01	1.012 ± 0.008
σ	0.170 ± 0.009	0.863 ± 0.006

in the analytic signal shape visible as structure in the fit residuals (Section 7.7.3), where a pure toy study would show none. We therefore quantify the PDF mismodelling bias with a hybrid ensemble of pseudo-experiments in which the signal and background are produced independently.

In each pseudo-experiment, the background is generated from the nominal analytic background PDF, while the signal events are drawn from the truth-matched signal component ($\text{isSignal} = 1$) of the generic MC, which is the largest such sample available, four times the size of the data. Each pseudo-experiment is generated at the size of the data, with the nominal signal and background yields, and is fitted with the nominal model. For each pseudo-experiment, we compute the relative deviation of the fitted signal yield from the generated signal yield, $(N_{\text{fit}} - N_{\text{gen}})/N_{\text{gen}}$ (Figure 7.20). Drawing the signal from MC rather than from the analytic PDF is what makes the test sensitive to the peak-shape imperfection: any mismatch between the true and modelled shape now propagates into the fitted yield.

The mean relative deviation from the generated yield is -1.8% for Belle and -2.5% for Belle II. Since the normalisation yield enters the denominator of Eq. (7.1), an underestimate of $N_{\pi KK}$ of this size corresponds to an overestimate of the branching fraction by the same relative amount. Since the standard toy study above shows no such bias, the shift most likely originates from the imperfect description of the peak shape and not from the maximum-likelihood estimator itself. The method still carries a residual background contribution and a limited precision from the finite generic MC statistics, but we assume both to be small and assign the full observed shift as a systematic uncertainty, the *normalisation PDF mismodelling*, propagated in Chapter 8.

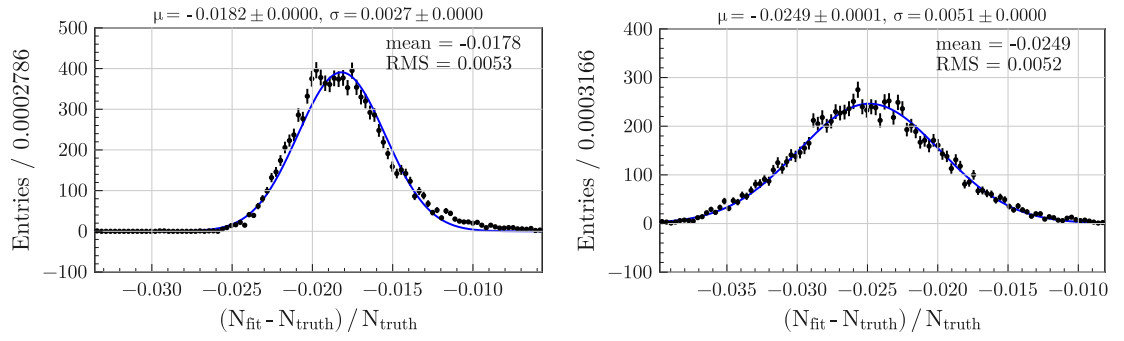


Figure 7.20: Relative deviation of the fitted signal yield from the generated signal yield, $(N_{\text{fit}} - N_{\text{gen}})/N_{\text{gen}}$, for the normalisation mode in Belle (left) and Belle II (right).

8. Systematic Uncertainties

The uncertainties treated in this chapter arise from imperfect knowledge of the detector response, of the selection efficiency, and of the models used in the fit. Unlike the statistical uncertainty of the signal yield, their dependence on sample size is source-specific. They must therefore be propagated separately and can limit the achievable precision, thus degrading the confidence intervals of Chapter 9.

The branching fraction of Eq. 7.1, defined in Section 7.1, is constructed as a ratio of the signal (or validation) yield to the normalisation yield, corrected by the corresponding reconstruction efficiencies and scaled by the well-measured branching fraction of the normalisation mode. This construction is deliberate. Any mismodelling that scales the signal and normalisation yields by the same factor leaves their ratio unchanged, and therefore does not affect the measured branching fraction. The integrated luminosity and the charm production cross-section cancel exactly, since both modes are reconstructed from the same produced D^+ sample. The effects depending on the D^+ and the charged pion (the RFC efficiency, the pion identification and the pion tracking) cancel only to first order, because the two modes differ in their remaining final-state particles and therefore populate the detector differently. We take the conservative approach and do not credit this cancellation as exact. We assign each such difference in full to the signal efficiency.

We group the systematic uncertainties by the mechanism through which they propagate to the final result rather than by their physical origin. Contributions that act on a parameter of the fit are folded directly into the generation of the pseudo-experiments, by smearing the corresponding fixed input within its uncertainty. Contributions that scale the signal efficiency are treated as multiplicative uncertainties on the expected signal yield, while contributions that shift the extracted branching fraction irrespective of its magnitude are treated as additive. Section 8.1 sets out this classification in full and specifies how each category is implemented in the upper-limit determination.

These categories differ greatly in their impact on the final result. The additive contributions dominate the upper limits, since they shift the extracted branching fraction irrespective of its magnitude, whereas the multiplicative ones, which scale the signal efficiency, matter far less near the upper limit, where the sensitivity is set by the background yield. We therefore neglect some of the smaller multiplicative contributions, as noted where they arise.

Particle identification is the only selection requirement for which we correct the simulated efficiencies, using dedicated data/MC calibration factors applied at the level of the branching-fraction calculation. For all other sources we apply no correction and instead assign the observed data/MC difference in full as a systematic uncertainty. All such contributions enter the Neyman construction as symmetric Gaussian smearing (Section 8.1), so assigning a difference in this way widens the

confidence interval without shifting it in either direction.

The chapter is organised as follows. Section 8.1 defines the propagation scheme. Section 8.2 covers the efficiency-related sources—PID, RFC, tracking, the statistical uncertainty, and the signal decay model—and Section 8.2.4 the incomplete cancellation of mis-reconstructed signal between the modes. Section 8.3 treats the uncertainties on the normalisation yield and on the external normalisation branching fraction, and Section 8.4 those arising from the fit model itself, including the data/MC differences in the signal-PDF shape parameters (Section 8.4.3). Section 8.5 summarises all contributions in Table 8.6, separately for Belle and Belle II and for each $M(e^+e^-)$ interval.

8.1 Classification and Propagation

We incorporate systematic uncertainties into the Neyman construction of the simultaneous fit, which is the fit used to extract the final result (we quote the individual per-sample fits of Chapter 7 with statistical uncertainties only). Each contribution enters at one of two points. Uncertainties on quantities held fixed in the fit are propagated during the generation of the pseudo-experiments, by smearing the corresponding input within its uncertainty. This applies to the inputs to the expected signal yield (the normalisation yield, the normalisation branching fraction and the efficiency-related uncertainties), and separately to the fixed shape parameters of the signal PDF. Uncertainties whose effect on the fitted branching fraction is a simple offset or rescaling are instead applied directly to the fitted value in each pseudo-experiment, either as an additive offset or as a multiplicative scale factor.

We implement the smearing of these inputs by rewriting the signal yield used to generate each pseudo-experiment as

$$N_{\pi ee,i}^{j,\text{eff}} = (N_{\pi KK,i} + \delta_{\pi KK,i}) \cdot \frac{\mathcal{B}_{\pi ee}^j}{\mathcal{B}_{\pi KK} + \delta_{\mathcal{B}}} \cdot \frac{\epsilon_{\pi ee,i}^j}{\epsilon_{\pi KK,i}} \cdot (1 + \delta_{m,i}^j), \quad (8.1)$$

where the index i runs over the Belle and Belle II samples, j labels the $M(e^+e^-)$ interval, and each δ is drawn in every pseudo-experiment from a Gaussian distribution centred at zero, with width equal to the corresponding uncertainty and with the correlations between the two samples specified below. The term $\delta_{\pi KK,i}$ carries the uncertainty on the normalisation yield, formed as the quadrature sum of its statistical uncertainty, the uncertainty from the choice of the normalisation background model, and the normalisation PDF mismodelling (Section 8.3). The term $\delta_{\mathcal{B}}$ carries the uncertainty on the external branching fraction of the normalisation mode and is drawn once per pseudo-experiment, common to both samples, since the same external input scales both. The term $\delta_{m,i}$ is the total multiplicative uncertainty, obtained by summing in quadrature the contributions from the RFC efficiency, PID, the mis-reconstructed signal fraction, the reconstruction efficiency, the tracking efficiency, and the signal decay model, with the PID, reconstruction and tracking contributions each including both the signal and normalisation modes. The signal-model contribution is common to the two samples: a single factor is drawn per pseudo-experiment from a standard normal distribution and scaled by each sample's own signal-model uncertainty, so that the two shifts are fully correlated but differ in magnitude. The remaining contributions to $\delta_{m,i}$ are drawn independently for each sample.

We also vary the signal-PDF parameters that are held fixed in the fit, at the generation stage, by adding to each an offset drawn from a Gaussian distribution of width equal to its fitted uncertainty. We then fit each pseudo-experiment with these parameters held at their nominal values, so that the resulting mismatch between the generated and fitted shapes broadens the distribution of the fitted branching fraction and propagates into the confidence belt.

We apply the remaining two contributions to the fitted branching fraction of each pseudo-experiment rather than to its generation. The uncertainty from the choice of the signal-mode background model enters additively, as a random offset added to the fitted branching fraction and drawn from a Gaussian centred at zero with the width given in Section 8.4. The uncertainty from residual data/MC differences in the signal-PDF shape enters multiplicatively, as a scale factor applied to the fitted branching fraction and drawn from a Gaussian centred at unity with the width determined in Section 8.4.3.

8.2 Efficiency-Related Uncertainties

The efficiency of Eq. 7.1 is affected by the modelling of particle identification, of the random-forest classifier, and of the tracking, and by the finite size of the simulated samples used to determine it. All four sources enter as multiplicative uncertainties on the signal efficiency, and we evaluate each separately for Belle and Belle II, since the two samples differ in detector, reconstruction software and calibration.

8.2.1 Particle Identification

The particle-identification efficiency differs between data and simulation, so we correct the reconstruction efficiency obtained from simulation as

$$\epsilon_{\text{data}} = \epsilon_{\text{MC}} \cdot \frac{\epsilon_{\text{data}}^{\text{PID}}}{\epsilon_{\text{MC}}^{\text{PID}}}, \quad (8.2)$$

separately for each track in the final state. We determine the efficiencies from calibration channels in which the particle species is known either from the decay topology or from a tag-and-probe technique.

For electron identification we use $J/\psi \rightarrow e^+e^-$, $e^+e^- \rightarrow e^+e^-e^+e^-$, and, for Belle II only, $e^+e^- \rightarrow e^+e^-\gamma$. For hadron identification we use $D^{*+} \rightarrow \pi^+D^0(\rightarrow K^-\pi^+)$ decays in both experiments, supplemented in Belle II by $K_S^0 \rightarrow \pi^+\pi^-$ and $\Lambda \rightarrow p\pi^-$ samples. The Belle correction tables are derived from experiments up to experiment 65, and we apply them to the later samples, which account for approximately 16 % of the Belle luminosity. The detector configuration and reconstruction software did not change significantly over this period, and any residual difference in the correction factors would affect only this fraction of the sample, with minor impact on the final result. We therefore assign no additional uncertainty.

The correction factors are provided in bins of track momentum and polar angle, since the detector response varies across the acceptance. Figure 8.1 shows the correction factors for pions in Belle II as an example; the tables for the remaining species and for Belle are structured in the same way. We combine them into a single factor per track as the track-weighted average

$$R = \frac{1}{N} \sum_l n_l R_l, \quad \Delta R^{\text{syst}} = \frac{1}{N} \sum_l n_l \Delta R_l^{\text{syst}}, \quad (8.3)$$

where $R_l = \epsilon_l^{\text{data}}/\epsilon_l^{\text{MC}}$ is the correction factor in bin l of momentum and polar angle, ΔR_l^{syst} is its systematic uncertainty, n_l is the number of signal tracks falling in that bin, and $N = \sum_l n_l$. We sum the bin uncertainties linearly rather than in quadrature, treating them as fully correlated since they originate from the same calibration samples.

We evaluate the statistical uncertainty on R with a dedicated toy study, in which the per-track weights are fluctuated within their statistical uncertainties; it is two orders of magnitude smaller than the systematic uncertainty and we therefore neglect it. We obtain the correction factor for each decay mode by multiplying the factors of all three tracks. We sum the uncertainties of the two electron tracks linearly, as they are taken from the same calibration tables and are fully correlated, while we add the pion contribution in quadrature.

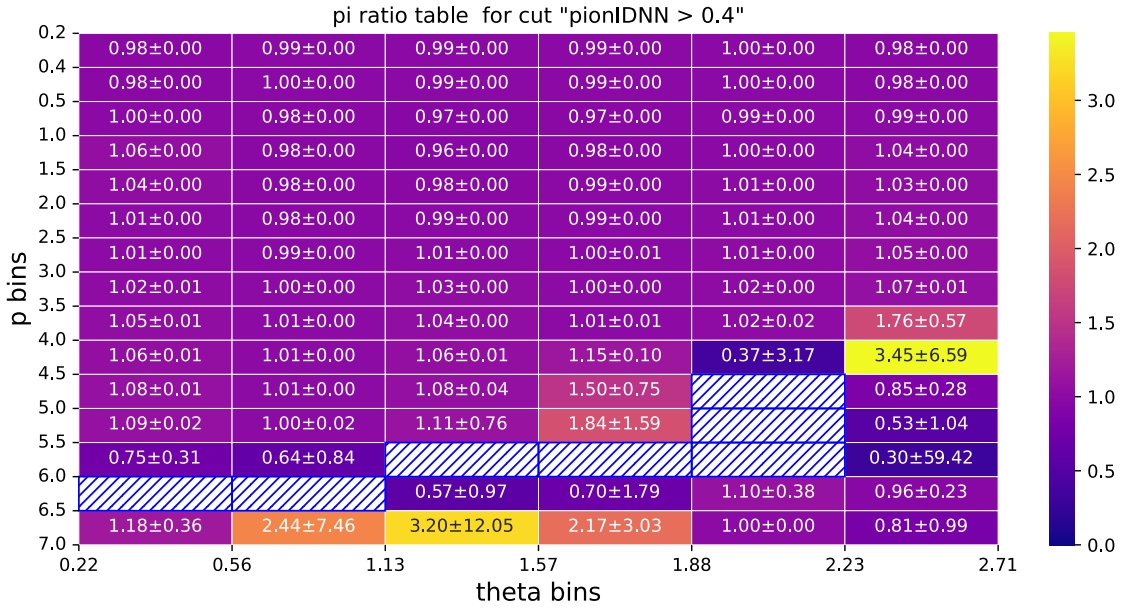


Figure 8.1: Particle-identification correction factors for pions satisfying $\pi\text{ID}_{\text{NN}} > 0.40$ in Belle II, in bins of pion track momentum and polar angle. Bins for which no correction factor is available are left blank, and the uncertainty associated with them is described in the text.

The calibration tables contain empty bins, arising from limited calibration statistics or from failed fits, for which no correction factor is available. In the signal mode, approximately 11 % of the tracks in Belle and 3 % in Belle II fall in such bins. We estimate the associated uncertainty by assigning a correction factor of unity to the empty bins and comparing the resulting mode-level correction factor with the nominal one. The differences range from 0.2 % to 0.5 % across the $M(e^+e^-)$ intervals for Belle and from 0.4 % to 0.5 % for Belle II, and we conservatively assign 0.5 % to all intervals in both samples. For the normalisation mode the uncovered fractions are 5 % in Belle and 0.4 % in Belle II, and we assign 0.1 % to Belle only, since no significant difference is observed for Belle II. For the validation mode the fractions are 8 % and 3 %, and we assign 0.3 % and 0.4 % respectively. We add this contribution linearly to the corresponding correction uncertainty rather than in quadrature, as a deliberately conservative choice, and the values quoted below and in Table 8.6 include it.

8.2. EFFICIENCY-RELATED UNCERTAINTIES

Tables 8.1 and 8.2 list the per-track correction factors and their uncertainties for the signal mode. The electron tracks dominate, contributing 2.1 % per track in Belle and 2.4 % in Belle II over the full $M(e^+e^-)$ range, compared with 0.6 % and 0.1 % for the pion. The combined correction factors are $(91.1 \pm 4.2)\%$ for Belle and $(96.2 \pm 4.8)\%$ for Belle II, and we evaluate them separately in each $M(e^+e^-)$ interval.

Table 8.1: Particle-identification correction factors and their systematic uncertainties, in per cent, for the Belle signal mode in each $M(e^+e^-)$ interval. The combined row gives the product of the three track factors, with the electron uncertainties summed linearly and the pion contribution added in quadrature.

$M(e^+e^-)$ interval [GeV/c ²]	0.20 – 1.75	0.20 – 0.50	0.56 – 0.95	1.05 – 1.75
e^-	97.2 ± 2.1	96.9 ± 2.2	97.0 ± 2.2	97.5 ± 2.0
e^+	97.2 ± 2.1	96.8 ± 2.2	97.0 ± 2.2	97.5 ± 2.0
π^+	96.3 ± 0.6	96.1 ± 0.5	96.1 ± 0.5	96.6 ± 0.6
Combined	91.1 ± 4.2	90.1 ± 4.4	90.4 ± 4.4	91.7 ± 4.0

Table 8.2: Particle-identification correction factors and their systematic uncertainties, in per cent, for the Belle II signal mode in each $M(e^+e^-)$ interval. The combined row gives the product of the three track factors, with the electron uncertainties summed linearly and the pion contribution added in quadrature.

$M(e^+e^-)$ interval [GeV/c ²]	0.20 – 1.75	0.20 – 0.50	0.56 – 0.95	1.05 – 1.75
e^-	98.6 ± 2.4	98.3 ± 2.7	98.2 ± 2.7	98.9 ± 2.3
e^+	98.7 ± 2.4	98.3 ± 2.7	98.2 ± 2.5	99.0 ± 2.2
π^+	98.8 ± 0.1	99.3 ± 0.1	99.2 ± 0.1	98.5 ± 0.05
Combined	96.2 ± 4.8	96.0 ± 5.4	95.7 ± 5.2	96.5 ± 4.4

Table 8.3 lists the corresponding factors for the normalisation and validation modes. The Belle kaon correction factors lie slightly above unity, indicating a slightly higher identification efficiency in data than in simulation, but still consistent with unity within their uncertainties.

8.2.2 Random-Forest-Classifer Efficiency

We cannot calibrate the RFC efficiency on the signal and validation modes, since their yields are too small to determine an efficiency from data with useful precision. We therefore assume that the data/MC efficiency ratio determined on the normalisation mode applies equally to the signal and validation modes. We evaluate this

CHAPTER 8. SYSTEMATIC UNCERTAINTIES

Table 8.3: Particle-identification correction factors and their systematic uncertainties, in per cent, for the normalisation and validation modes. The combined column gives the product of the three track factors, with the uncertainties of the two same-species tracks summed linearly and the remaining track added in quadrature.

Mode	Sample	Track			Combined
		h^-	h^+	π^+	
Normalisation	Belle	100.6 ± 0.5	100.5 ± 0.5	95.9 ± 0.5	96.9 ± 1.2
	Belle II	97.1 ± 1.1	97.2 ± 1.1	98.9 ± 0.1	93.4 ± 2.2
Validation	Belle	97.3 ± 2.1	97.3 ± 2.1	96.2 ± 0.5	91.1 ± 4.2
	Belle II	98.6 ± 2.1	98.6 ± 2.1	98.9 ± 0.1	96.2 ± 4.2

ratio separately for Belle and Belle II, since the two samples use different classifier working points and are calibrated on independent data.

We measure the RFC efficiency as the ratio of the normalisation signal yield after the RFC requirement to the yield before it, obtained separately from data and from generic MC. The yields before the requirement come from the fits shown in Appendix D.1. The yields after it come from the nominal normalisation-mode fits, performed on generic MC in Chapter 7 and on data in Chapter 9. We propagate the uncertainties on the RFC efficiencies from the fit uncertainties on the normalisation yields, since these exceed the purely statistical uncertainties of the samples.

For Belle, with $\text{RFC_score} > 0.97$, the efficiency is $(27.2 \pm 0.3)\%$ in data and $(26.1 \pm 0.2)\%$ in simulation, giving a relative difference of $(4.2 \pm 2.0)\%$. For Belle II, with $\text{RFC_score} > 0.98$, the efficiency is $(34.2 \pm 0.6)\%$ in data and $(36.4 \pm 0.3)\%$ in simulation, giving a relative difference of $(6.0 \pm 2.5)\%$. We apply no correction and assign the full relative difference, 4.2 % for Belle and 6.0 % for Belle II, as a multiplicative systematic uncertainty. Under the assumption above, the ratio would cancel in Eq. 7.1, but we do not credit this cancellation and assign the full difference instead.

8.2.3 Tracking and Reconstruction Efficiency

We apply no data/MC correction for the track-finding efficiency and instead assign the standard per-track uncertainties recommended by the two collaborations, 0.35 % for Belle and 0.27 % for Belle II. We sum these contributions linearly over the three tracks of the signal or validation mode and the three tracks of the normalisation mode, treating them as fully correlated. This gives 2.1 % for Belle and 1.6 % for Belle II as a multiplicative systematic uncertainty, common to all $M(e^+e^-)$ intervals. Only the charged pion is common to the two modes, and even its tracking uncertainty does not cancel fully in Eq. 7.1, since the pion kinematics differ between the modes. The electron and kaon tracks have no counterpart in the other mode and therefore cannot cancel at all.

The reconstruction efficiencies of Chapter 6 are determined from samples of finite size, and their statistical uncertainties propagate directly to the branching fraction.

We include these uncertainties as a multiplicative systematic uncertainty, evaluated separately for each mode and each $M(e^+e^-)$ interval. They range from 0.3 % to 1.0 % for Belle and from 0.4 % to 0.8 % for Belle II, with the largest values in the lowest interval, where the selection efficiency is smallest. The individual values are listed in Table 8.6.

8.2.4 Mis-reconstructed Signal

A small fraction of signal decays is reconstructed from a wrong combination of tracks, yet still produces a D^+ candidate whose invariant mass peaks in the signal region. Because this component is indistinguishable from correctly reconstructed signal in $M(\pi^+e^+e^-)$, the fit assigns it to the signal yield, as discussed in Chapter 7. It therefore inflates the measured yield in both the signal and the normalisation mode.

We quantify the effect through the fraction of mis-reconstructed events among all peaking events, taken from Table 7.2 for the signal mode and from Table 7.4 for the normalisation mode, where the corresponding fits are also described. For Belle we obtain 1.4 % and 0.3 %, and for Belle II 0.9 % and 1.0 % in the signal and the normalisation mode, respectively. Were the two fractions equal, their effect would cancel in Eq. 7.1, which is very nearly the case for Belle II but not for Belle. We take the absolute difference between the two fractions as a multiplicative systematic uncertainty, giving 1.2 % for Belle and 0.1 % for Belle II.

We repeat this procedure in each $M(e^+e^-)$ interval, while the normalisation fraction is common to all intervals. The resulting uncertainties are 1.1 %, 0.9 % and 1.4 % for Belle and 0.3 %, 0.1 % and <0.1 % for Belle II, in I_1 , I_2 and I_3 respectively. For the validation mode, we obtain 0.5 % in both samples. All values are collected in Table 8.6.

8.2.5 Signal Decay Model

The signal efficiency of Eq. 7.1 is determined from simulations generated uniformly in three-body phase space, whereas the true decay may follow a different distribution of the final-state kinematics. The dependence on the $M(e^+e^-)$ spectrum itself is only partly removed by evaluating the efficiency separately in each interval (Section 6.1.2), since the efficiency still varies within an interval, and it also depends on the decay kinematics at fixed $M(e^+e^-)$. We therefore assess the effect by repeating the efficiency determination with an alternative decay model, treating the resonant sample as a conservative bound on how far the efficiency can move under a change of decay kinematics rather than a physically preferred alternative model. We compare the reconstruction efficiency of the phase-space sample (5.73 % for Belle and 6.78 % for Belle II) with that of the validation decay $D^+ \rightarrow \pi^+(\phi \rightarrow e^+e^-)$ (4.75 % and 6.13 %), generated with an $\text{SVS} \times \text{VLL}$ model, both evaluated in the validation window $I_\phi^{\text{Valid.}}$, where the two samples populate the same $M(e^+e^-)$ region. The relative difference, $\epsilon_{\text{resonance}}/\epsilon_{\text{PHSP}} - 1$, is -17 % for Belle and -9.6 % for Belle II. Since this difference reflects a single physical assumption on the decay dynamics, common to the two experiments, we treat it as fully correlated between the Belle and Belle II samples in the upper-limit determination.

We assign this difference as a multiplicative uncertainty on the signal efficiency in all $M(e^+e^-)$ intervals. Its effect on the upper limits of Chapter 9 is below 0.05×10^{-7} in every interval and in the full range. This is expected, since the sensitivity near

the limit is dominated by the background yield rather than the signal efficiency. The validation mode is itself generated with the resonance model and therefore carries no such uncertainty.

8.3 Normalisation-Yield Uncertainties

The normalisation yield enters the denominator of Eq. 7.1, and its uncertainty propagates directly to the branching fraction. Three contributions are relevant, all expressed in units of signal yield.

The first is the statistical uncertainty of the fitted yield, 367 events for Belle and 324 for Belle II, obtained from the fits to data of Chapter 9. The second follows from the choice of the background model. We vary the order of the Chebyshev polynomial describing the combinatorial background between zeroth and second order, refit the data, and take the largest resulting change in the fitted yield, 267 events for Belle and 181 for Belle II. The third is the normalisation PDF mismodelling of Section 7.8.5, which quantifies the imperfect description of the peak shape through a hybrid ensemble of pseudo-experiments. The observed relative shifts of -1.8% and -2.5% correspond to 1650 and 1478 events respectively, and dominate the total.

We combine the three contributions in quadrature, giving 1711 events for Belle and 1524 for Belle II. This combined uncertainty is propagated through the term $\delta_{\pi KK,i}$ of Eq. 8.1, by which the normalisation yield is smeared in the generation of each pseudo-experiment.

The branching fraction of the normalisation mode is an external input, taken from Ref. [16] as $\mathcal{B}(D^+ \rightarrow \pi^+ K^+ K^-) = (9.68 \pm 0.18) \times 10^{-3}$, with a relative uncertainty of 1.9% . We propagate it through the term $\delta_{\mathcal{B}}$ of Eq. 8.1, drawn once per pseudo-experiment and applied commonly to both samples, since the same external value scales the Belle and Belle II results.

8.4 Fit-Related Uncertainties

Three sources of uncertainty arise from the fit itself: the parameters of the signal PDF that are held fixed, the choice of the background model, and residual differences between data and simulation in the signal-PDF shape.

8.4.1 Fixed Signal-PDF Parameters

The shape parameters of the signal PDF are fixed to the values obtained from signal MC, as described in Section 7.4, and their uncertainties propagate to the extracted yield. In the simultaneous fit, we account for this directly in the Neyman construction by varying each fixed parameter within its uncertainty during the generation of the pseudo-experiments, as described in Section 8.1. No separate uncertainty is therefore assigned for this source in the summary table.

As a cross-check, we also explicitly evaluate the effect for the individual fits to see how large it actually is. We set each parameter in turn to its nominal value shifted by plus or minus one standard deviation, refit the data, and retain the larger of the two resulting yield differences. Summing these differences in quadrature over all fixed parameters, and assuming them to be uncorrelated, gives 1.6 events for Belle and 0.51 for Belle II in the signal mode, and 0.3 and 0.15 in the validation mode. These shifts are small compared with the statistical uncertainties of the corresponding yields.

8.4.2 Background-Model Choice

The combinatorial background is modelled by a first-order Chebyshev polynomial in the signal and Belle validation modes, by a second-order polynomial in the normalisation mode, and by a zeroth-order polynomial in the Belle II validation mode. Over the given fit range, the polynomial orders give visually comparable descriptions of the data, so the nominal choice is not uniquely determined, and the spread between them is considered a source of uncertainty. Figure 8.2 shows an example for Belle II, while the remaining variations are shown in Appendix D.4. We therefore refit the data samples with the remaining orders in the range from zeroth to second, and take the largest resulting change as the systematic uncertainty. For the Belle validation mode, we do not consider the zeroth-order alternative, since a flat background is the one case that does not describe the observed distribution adequately. We tested several alternative descriptions of the signal component, but none reproduces the signal MC as well as the nominal PDF, and the resulting fits are visibly degraded.

In the simultaneous fit, we vary the background model of the Belle and Belle II samples independently and take the largest resulting change in the branching fraction as an additive systematic uncertainty. Across the three $M(e^+e^-)$ intervals it ranges from 0.6×10^{-7} to 0.8×10^{-7} for Belle and from 0.4×10^{-7} to 1.0×10^{-7} for Belle II, while over the full range it is 0.9×10^{-7} and 1.0×10^{-7} respectively. The individual values are listed in Table 8.4. The background-model choice is the largest single systematic uncertainty of the analysis.

Table 8.4: Additive systematic uncertainty on the branching fraction from the choice of the background model, obtained from the simultaneous fit by varying the Chebyshev polynomial order independently for the two samples.

$M(e^+e^-)$ interval [GeV/c ²]	$\Delta\mathcal{B}$ [10 ⁻⁷]	
	Belle	Belle II
0.20 – 0.50	0.8	0.5
0.56 – 0.95	0.8	1.0
1.05 – 1.75	0.6	0.4
0.20 – 1.75	0.9	1.0
$D^+ \rightarrow \pi^+\phi(\rightarrow e^+e^-)$	0.4	< 0.1

8.4.3 Signal-PDF Differences Between Data and Simulation

The shape parameters of the signal PDF are fixed to values obtained from signal MC, and Section 8.4.1 accounts for their statistical uncertainties. This does not cover a systematic difference between the shape of the peak in data and in simulation, which would bias the extracted yield in a way that the fitted uncertainties cannot reveal. As for the RFC efficiency of Section 8.2.2, we cannot measure such a difference in the signal or validation modes, and we therefore infer it from the normalisation mode.

We fit the normalisation mode separately in data and in generic MC, using the same PDF in both cases, and compare the resulting shape parameters. Table 8.5

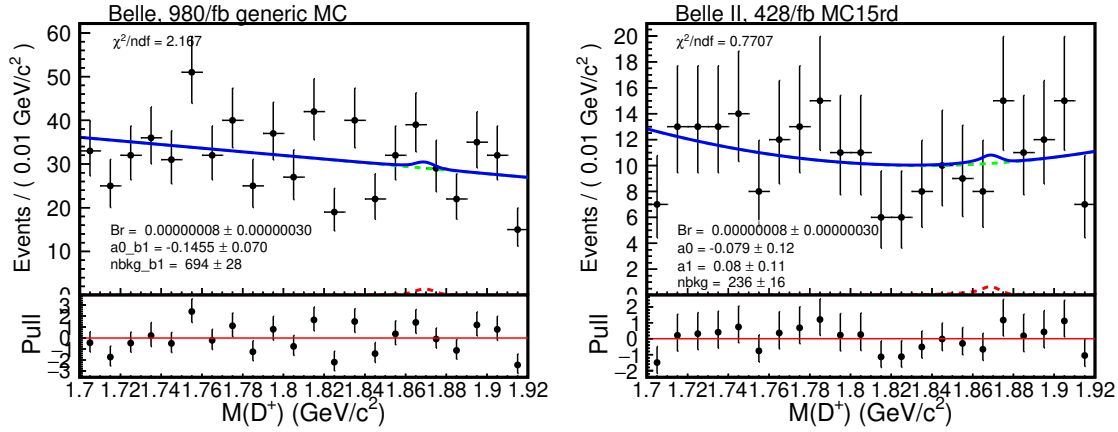


Figure 8.2: Simultaneous fit to the $M(\pi^+e^+e^-)$ distribution of the signal mode in generic MC for Belle (left) and Belle II (right), with the Belle II combinatorial background modelled by a second-order Chebyshev polynomial. The Belle background is kept at its nominal first-order description in both cases. The alternative description is barely distinguishable from the nominal one over the narrow fit range, which motivates treating the spread between them as a systematic uncertainty. The uncertainties of Table 8.6 are obtained from the corresponding refits to data.

lists them together with their relative differences. The peak position agrees to better than 0.01 % and the width to within a few per cent in both samples, while the tail parameters differ substantially, by up to 80 %. The tail parameters are strongly correlated with one another and only weakly constrained by the fit. The size of these relative differences is not in itself a measure of how much the peak shape differs, since widely separated parameter values can describe nearly the same curve. What matters is the effect on the extracted yield, which we quantify with a dedicated ensemble of pseudo-experiments.

To quantify the effect, we generate a dedicated ensemble of pseudo-experiments, following the same procedure as for the upper-limit determination but without any other systematic uncertainty included. Since the modes use different signal PDFs, we map each normalisation-mode parameter onto the analogous parameter of the target PDF. The peak position is mapped directly, the single width parameter of the normalisation PDF is applied to both the left and right width parameters, and the left and right tail parameters are mapped directly, for the signal mode and for the Belle validation mode. The Belle II validation mode is modelled by the sum of a symmetric Cruijff function and a bifurcated Gaussian sharing a common mean, so there the normalisation peak position is mapped to the common mean, the larger of the two normalisation tail parameters to the Cruijff tail parameter, and the normalisation width to the Cruijff width as well as to both widths of the bifurcated Gaussian. In each pseudo-experiment the smearing factors are drawn jointly, preserving the correlations obtained from the normalisation-mode fit, and every mapped parameter is multiplied by its factor. The sample is then generated with the smeared parameters and fitted with the parameters fixed to their nominal values. We then repeat the linearity test of Section 7.8 with this configuration. The slope of the fitted response deviates from unity by approximately -5% , consistently in all three $M(e^+e^-)$ intervals and over the full range, and by approximately -3%

for the validation mode. The linearity of the fit under these parameter variations is shown in Appendix D.3.

Varying the parameters during the fit rather than during the generation gives a deviation of the same magnitude with the opposite sign, so we treat the uncertainty as symmetric. We assign 5.0% for the signal mode and 3.0% for the validation mode as a multiplicative systematic uncertainty, common to both samples, applied to the fitted branching fraction of each pseudo-experiment in the upper-limit determination. Its effect on the upper limit is small: repeating the limit calculation with and without this source changes the result over the full $M(e^+e^-)$ range by about 0.05×10^{-7} .

In Belle II the validation mode uses a different signal PDF from the normalisation mode, so only the parameters common to both models are varied, namely the peak position, the width and the tail parameter.

Table 8.5: Signal-PDF shape parameters obtained from fits to the normalisation mode in data and in generic MC, with their relative differences. The relative differences define the widths of the scale factors applied to the parameters during the generation of the pseudo-experiments.

Parameter	Data	MC	Data/MC -1 [%]
Belle			
α_L	1.793	1.826	-1.8
α_R	1.707	1.562	9.3
μ	1.870156	1.870003	0.008
n_L	4.91	6.9	-29
n_R	5.05	9.1	-45
σ	0.003781	0.003669	3.1
Belle II			
α_L	1.591	1.618	-1.7
α_R	1.383	1.748	-21
μ	1.869785	1.869827	0.002
n_L	2.58	2.332	11
n_R	4.27	2.378	80
σ	0.002841	0.00281	1.1

8.5 Summary

Table 8.6 collects the systematic uncertainties that depend on the $M(e^+e^-)$ interval, separately for Belle and Belle II, and Table 8.7 those that are common to all

intervals. The category given with each entry indicates how the contribution enters the upper-limit determination, following the classification of Section 8.1.

Among the multiplicative uncertainties, the largest is the signal decay model, at 17 % for Belle and 9.6 % for Belle II, followed by the PID correction at 4.5 % to 5.9 % depending on the sample and interval, the signal-PDF data/MC difference at 5.0 %, and the RFC efficiency at 4.2 % for Belle and 6.0 % for Belle II. Although the signal-model uncertainty is numerically the largest multiplicative contribution, its impact on the limit is negligible, because near the upper limit the sensitivity is set by the background yield rather than the signal efficiency. The background-model choice enters additively and is the largest single contribution overall, reaching 1.0×10^{-7} in the second $M(e^+e^-)$ interval for Belle II. Its effect is largest where the statistical precision is greatest, since the additive shift does not scale with the extracted branching fraction.

The uncertainties associated with the normalisation yield are dominated by the PDF mismodelling, which exceeds the statistical uncertainty of the fitted yield by more than a factor of four in both samples. The external branching fraction of the normalisation mode contributes 1.9 % and is the only source treated as common to the two samples.

The total systematic uncertainty is small compared with the statistical uncertainty of the signal fit in every $M(e^+e^-)$ interval. However, it is not negligible, since including all sources degrades the upper limit over the full $M(e^+e^-)$ range by 12 % relative to the same construction without them, a shift driven almost entirely by the additive background-model term.

Table 8.6: Systematic uncertainties of the signal and validation modes that depend on the $M(e^+e^-)$ interval. The uncertainty from the alternative background model is given in units of branching fraction; all others are in per cent. The category in parentheses indicates whether the contribution is applied as an additive shift of the branching fraction (add.) or as a multiplicative factor (mult.), following the classification of Section 8.1. The intervals are $I_1 = (0.20, 0.50)$, $I_2 = (0.56, 0.95)$, $I_3 = (1.05, 1.75)$, $I_{\text{full}} = I_1 \cup I_2 \cup I_3$, and $I_{\phi}^{\text{Valid.}} = (0.995, 1.035) \text{ GeV}/c^2$, which denotes the validation mode.

	$M(e^+e^-)$ interval	I_1	I_2	I_3	I_{full}	$I_{\phi}^{\text{Valid.}}$
Belle	$\Delta\mathcal{B}(\text{alt. bkg.}) [10^{-7}]$ (add.)	0.8	0.8	0.6	0.9	0.4
	PID efficiency (mult.)	4.9	4.9	4.5	4.7	4.5
	Mis-reconstructed signal (mult.)	1.1	0.9	1.4	1.2	0.5
	Reconstruction efficiency (mult.)	1.0	0.6	0.5	0.4	0.3
	Signal decay model (mult.)	17	17	17	17	–
	Data–MC difference (mult.)	5.0	5.0	5.0	5.0	3.0
Belle II	$\Delta\mathcal{B}(\text{alt. bkg.}) [10^{-7}]$ (add.)	0.5	1.0	0.4	1.0	< 0.1
	PID efficiency (mult.)	5.9	5.7	4.9	5.2	4.6
	Mis-reconstructed signal (mult.)	0.3	0.1	< 0.1	0.1	0.5
	Reconstruction efficiency (mult.)	0.8	0.5	0.4	0.4	0.4
	Signal decay model (mult.)	9.6	9.6	9.6	9.6	–
	Data–MC difference (mult.)	5.0	5.0	5.0	5.0	3.0

Table 8.7: Systematic uncertainties common to all $M(e^+e^-)$ intervals, given in per cent. All are propagated by smearing the corresponding input during the generation of the pseudo-experiments. The uncertainty on the branching fraction of the normalisation mode is common to both samples, since the same external value scales them both.

Source	Belle	Belle II
<i>Normalisation channel yield uncertainty</i>		
Alternative background	0.3	0.3
Statistical uncertainty	0.4	0.5
PDF mismodelling	1.8	2.5
<i>Efficiencies, evaluated in normalisation channel</i>		
RFC efficiency	4.2	6.0
PID efficiency	1.3	2.2
reconstruction efficiency	1.0	0.4
Signal tracking efficiency	2.1	1.6
$\mathcal{B}(D^+ \rightarrow \pi^+ K^+ K^-)$	1.9	

9. Results

In this chapter, we apply the fit machinery established in Chapter 7, together with the systematic uncertainties of Chapter 8, to the Belle and Belle II data samples of 980 fb^{-1} and 428 fb^{-1} , following the procedure fixed before examining the signal region, as described at the start of Chapter 7.

We present the normalisation and validation modes first, since neither lies within the signal region and both were fit to data throughout the analysis. The normalisation mode comes first because its yield enters the branching fraction of both other modes, and the validation mode next, since it shares the final state of the signal mode and tests the reconstruction, selection and fit on an observable signal. The signal mode is presented last, together with the upper limits derived from it.

9.1 Normalisation Mode

The normalisation mode $D^+ \rightarrow \pi^+ K^+ K^-$, reconstructed within the ϕ mass window $|M(K^+ K^-) - m_\phi| < 20\text{ MeV}/c^2$, supplies the denominator of Eq. 7.1. Its branching fraction is known to 1.9% and its yield is large, so it fixes the absolute scale of the measurement while cancelling the effects common to it and the signal mode. The fit model and its closure on simulated samples are described in Section 7.7.3, and the agreement between data and simulation for this mode in Section 6.2.2.

Each experiment enters Eq. 7.1 with its own normalisation yield, so we fit the two data samples separately. Figure 9.1 shows the fitted $M(\pi^+ K^+ K^-)$ distributions. The extracted yields are 91663 ± 367 for Belle and 59119 ± 324 for Belle II, where the uncertainties are statistical.

These yields are included in the measurement along with the reconstruction efficiencies in Chapter 6. Combining the statistical uncertainties with those from the background-model choice and PDF mismodelling in quadrature, as described in Section 8.3, yields total uncertainties of 1711 and 1524 events. We propagate this combination through the term $\delta_{\pi KK,i}$ of Eq. 8.1, by which the normalisation yield is smeared in the generation of each pseudo-experiment.

9.2 Validation Mode

Unlike the normalisation mode, the decay $D^+ \rightarrow \pi^+ \phi (\rightarrow e^+ e^-)$ has the same final state as the signal mode, and a branching fraction large enough to be observed in the present data samples. It therefore provides a direct test of the reconstruction, the selection and the fit, on the same $M(\pi^+ e^+ e^-)$ distribution used for the measurement itself. We reconstruct it in the interval $0.995 < M(e^+ e^-) < 1.035\text{ GeV}/c^2$, which contains the ϕ resonance and is excluded from the signal-mode intervals. Figure 9.2 shows the simultaneous fit to the two data samples, in which the branching fraction

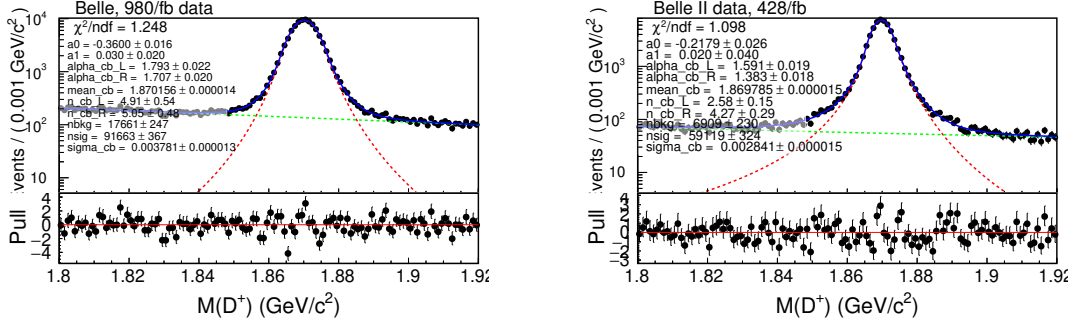


Figure 9.1: Fits to the $M(\pi^+K^+K^-)$ distribution of the normalisation mode $D^+ \rightarrow \pi^+K^+K^-$, reconstructed within the ϕ mass window, in the Belle (left) and Belle II (right) data samples.

is a common parameter. The fitted branching fraction is

$$\mathcal{B}(D^+ \rightarrow \pi^+\phi(\rightarrow e^+e^-)) = (1.75 \pm 0.22 \text{ (stat)} \pm 0.12 \text{ (syst)}) \times 10^{-6}. \quad (9.1)$$

Inverting Eq. 7.1 with the reconstruction efficiencies of Section 6.1.2, we obtain signal yields of $51 \pm 6 \text{ (stat)} \pm 3 \text{ (syst)}$ for Belle and $30 \pm 4 \text{ (stat)} \pm 2 \text{ (syst)}$ for Belle II, where the uncertainties are propagated from those of the branching fraction. The signal significance, evaluated from Eq. 7.9 without systematic uncertainties, is 12.0 standard deviations.

Since this mode has a previously measured branching fraction, we derive the one-standard-deviation confidence interval with the same Neyman construction used for the upper limits of Section 9.4, constructing a two-sided interval here rather than a limit. The validation therefore exercises the same toy generation and systematic propagation on which the signal-mode limits rely, and not only the reconstruction and the fit. We determine the two uncertainties by performing the construction twice, once including all systematic variations and once with statistical fluctuations only, and taking $\sigma_{\text{syst}} = \sqrt{\sigma_{\text{tot}}^2 - \sigma_{\text{stat}}^2}$. The quoted systematic uncertainty is therefore not the sum of the individual contributions listed in Section 8.5. The confidence belt without systematic uncertainties is shown in Appendix D.5. Comparing the 1σ interval from the confidence belt with the uncertainty returned by the likelihood fit, we find no difference, indicating that the likelihood is locally parabolic and the fit uncertainty is therefore reliable. This is expected when the signal is large, and the fitted branching fraction lies far from the physical boundary, neither of which holds in the signal mode. The result is dominated by the statistical uncertainty, as expected for a yield of this size.

The measured value agrees with the expectation derived from the known branching fractions of the two steps, $\mathcal{B}(D^+ \rightarrow \pi^+\phi) \times \mathcal{B}(\phi \rightarrow e^+e^-) = (1.70 \pm 0.05) \times 10^{-6}$ [16], which confirms that the reconstruction efficiency and the normalisation are correctly determined. It is also consistent with the previous direct measurement of this mode by the CLEO collaboration, $(1.7_{-0.9}^{+1.4}) \times 10^{-6}$ [48], and improves its total precision by approximately a factor of five. The BESIII collaboration has separately reported an observation of the corresponding $D_s^+ \rightarrow \pi^+\phi(\rightarrow e^+e^-)$ decay with a significance of 7.8 standard deviations [92].

The *BABAR* collaboration examined the same mode in their search as a cross-check but did not quote a branching fraction. Combining their published signal yield of

$21.8 \pm 5.8 \pm 1.5$ events and efficiency of 5.65 %, together with the yield of 106800 ± 500 , efficiency of $(15.44 \pm 0.07)\%$ and branching fraction of $(2.72 \pm 0.13) \times 10^{-3}$ of their normalisation mode $D^+ \rightarrow \pi^+ \phi (\rightarrow K^+ K^-)$ [4], we infer a value of $\mathcal{B}(D^+ \rightarrow \pi^+ \phi (\rightarrow e^+ e^-)) = (1.5 \pm 0.4 \pm 0.1) \times 10^{-6}$, which agrees with our measurement and with the expectation from the known branching fractions. Our result improves the statistical precision of this value by approximately a factor of two.

As a further check, we repeat the measurement fitting each sample independently. The independent fits give signal yields of 46.1 ± 8.0 for Belle and 34.2 ± 6.4 for Belle II, corresponding to branching fractions of $(1.51 \pm 0.26) \times 10^{-6}$ and $(1.99 \pm 0.37) \times 10^{-6}$, where the uncertainties follow from the fitted yields alone. The two values agree within 1.1 standard deviations, so the samples give a consistent result and the simultaneous fit does not conceal a tension between them. The corresponding fits are shown in Appendix D.6.

Taken together, these comparisons show that the analysis reproduces a known branching fraction with the same selection and fit procedure applied to the signal mode, which is the principal evidence that the measurement is free of a significant bias.

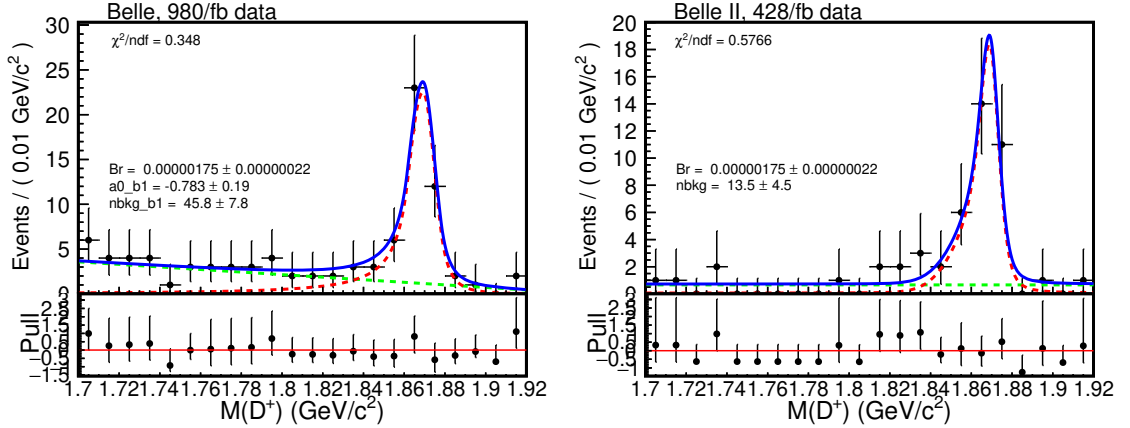


Figure 9.2: Simultaneous fit to the $M(\pi^+ e^+ e^-)$ distribution of the validation mode $D^+ \rightarrow \pi^+ \phi (\rightarrow e^+ e^-)$ in the Belle (left) and Belle II (right) data samples. The branching fraction is a common parameter of the two fits.

9.3 Signal Mode

We fit the $M(\pi^+ e^+ e^-)$ distribution of the signal mode simultaneously in the two data samples, with the branching fraction as the common parameter, separately in each $M(e^+ e^-)$ interval and over the full range. Figure 9.4 shows the fitted distributions.

The fitted branching fractions are $(-0.6 \pm 2.0) \times 10^{-7}$, $(0.0 \pm 2.1) \times 10^{-7}$ and $(-0.7 \pm 2.0) \times 10^{-7}$ in I_1 , I_2 and I_3 , and $(-0.1 \pm 3.3) \times 10^{-7}$ over the full range, where the uncertainties are statistical. Table 9.1 shows the corresponding signal yields together with the reconstruction efficiencies. All are consistent with zero, and none of the fitted distributions shows a visible enhancement at the D^+ mass.

Several fitted branching fractions are negative, which is unphysical as a decay rate but is the expected behaviour of an unbiased estimator when the true rate is small and the background fluctuates downwards. We do not constrain the fit to non-negative values, since bounding the estimator would bias it and would distort

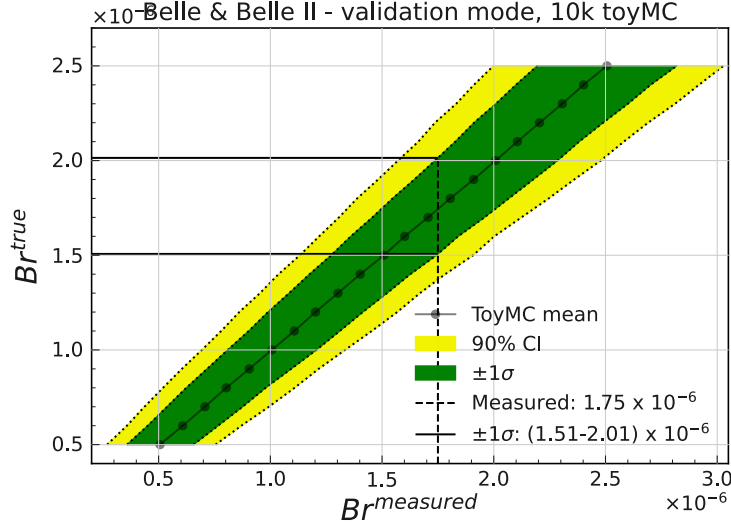


Figure 9.3: Neyman confidence belt for the validation mode, constructed from the simultaneous fit with all systematic uncertainties included. The observed branching fraction and the corresponding two-sided interval are indicated.

the coverage of the pseudo-experiment ensemble on which the Neyman construction relies. The physical constraint enters instead when the confidence interval is built, by restricting the construction to $\mathcal{B} \geq 0$, as described in Section 7.6.

Evaluated as in Section 7.6, the signal significance is consistent with zero in all three intervals and over the full range. We therefore find no evidence for the decay $D^+ \rightarrow \pi^+ e^+ e^-$ and set upper limits on its branching fraction in Section 9.4.

Table 9.1: Reconstruction efficiencies, signal yields and fitted branching fractions in each $M(e^+e^-)$ interval. The quoted uncertainties are statistical only.

$M(e^+e^-)$ interval GeV/c ²	Efficiency (%)		N_{sig}		$\mathcal{B}$ ($\times 10^{-7}$)
	Belle	Belle II	Belle	Belle II	
0.20 – 0.50	3.82 ± 0.04	4.73 ± 0.04	-1.7 ± 5.8	-1.0 ± 3.4	-0.6 ± 2.0
0.56 – 0.95	5.03 ± 0.03	6.14 ± 0.03	0.0 ± 6.0	0.0 ± 3.6	0.0 ± 2.1
1.05 – 1.75	5.94 ± 0.03	7.42 ± 0.03	-2.0 ± 5.8	-1.2 ± 3.4	-0.7 ± 2.0
Full range	5.29 ± 0.02	6.57 ± 0.02	-0.3 ± 9.6	-0.2 ± 5.7	-0.1 ± 3.3

9.4 Upper Limits

Since we observe no evidence for the decay, we set upper limits on its branching fraction at the 90% confidence level, using the Neyman construction described in Section 7.6. The systematic uncertainties of Chapter 8 enter through the mechanisms of Section 8.1, and are included in all limits quoted below unless stated otherwise.

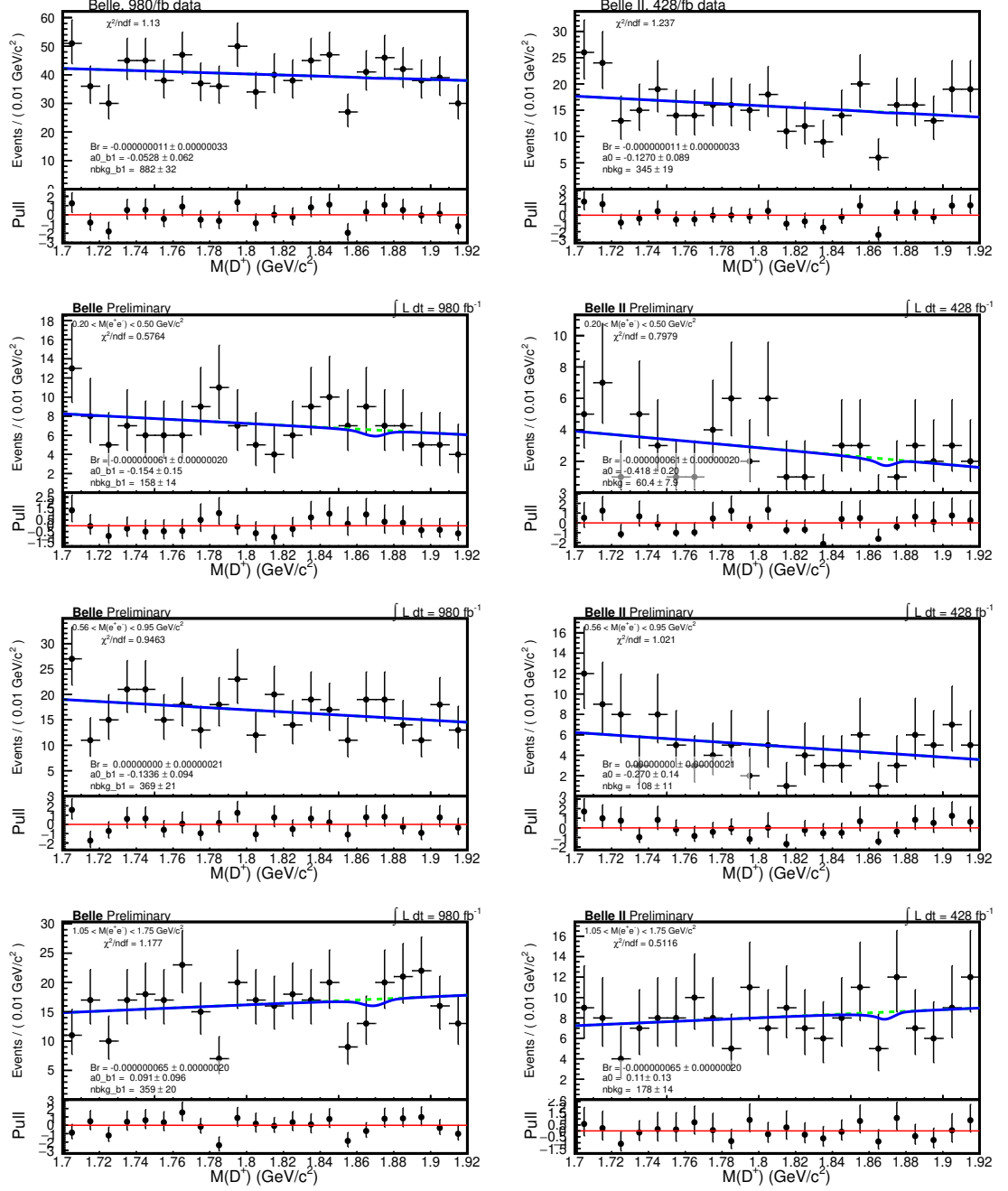


Figure 9.4: Simultaneous fits to the $M(\pi^+e^+e^-)$ distribution of the signal mode in the Belle (left) and Belle II (right) data samples, shown from top to bottom for the full $M(e^+e^-)$ range and for the intervals I_1 , I_2 and I_3 . The branching fraction is a common parameter of the two fits in each case.

Over the full $M(e^+e^-)$ range, excluding the ϕ and η regions, we obtain

$$\mathcal{B}(D^+ \rightarrow \pi^+ e^+ e^-) < 4.7 \times 10^{-7} \quad \text{at 90\% CL.} \quad (9.2)$$

In the individual intervals the limits are 3.0×10^{-7} , 3.3×10^{-7} and 2.9×10^{-7} for I_1 , I_2 and I_3 . These are limits on partial branching fractions and are therefore not directly comparable with the full-range limit, which covers the union of the three intervals. Table 9.2 lists them together with the limits obtained without systematic uncertainties. Including the systematic uncertainties raises the full-range limit from 4.2×10^{-7} to 4.7×10^{-7} . This modest shift reflects the statistical uncertainty on the fitted branching fraction, $2.0\text{--}2.1 \times 10^{-7}$ in the individual $M(e^+e^-)$ intervals and 3.3×10^{-7} over the full range (Table 9.1), which exceeds any single systematic contribution.

The median limits expected in the absence of signal, obtained from the same Neyman construction with the input parameters taken from fits to generic MC samples and no signal contribution, are 4.4×10^{-7} with systematic uncertainties and 3.7×10^{-7} without. The corresponding expected limits, including systematic uncertainties, are 4.1×10^{-7} , 2.7×10^{-7} and 2.3×10^{-7} for I_1 , I_2 and I_3 . The observed limits track the sign of the data/simulation discrepancy in the signal mode (Section 6.2.4). The limit is stronger than expected in I_1 , where simulation overestimates the background, and weaker than expected in I_2 and I_3 , where it underestimates the background. The limits therefore remain dominated by the size of the data samples rather than by the systematic uncertainties.

The limits are quoted for a signal distributed according to phase space, since the efficiency depends on the assumed $M(e^+e^-)$ spectrum and no measurement of that spectrum exists for this decay. We assign the associated model dependence as the signal-model systematic uncertainty of Section 8.2.5, which is included in the values above.

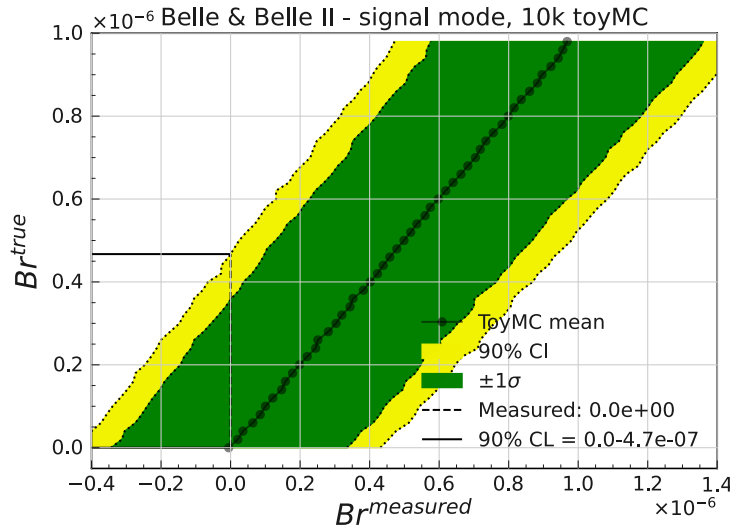


Figure 9.5: Neyman confidence belt for the signal mode over the full $M(e^+e^-)$ range, constructed from the simultaneous fit with all systematic uncertainties included. The upper limit is given by the intersection of the observed branching fraction with the upper edge of the belt, and the construction is restricted to the physical region $\mathcal{B} \geq 0$.

Table 9.2: Upper limits on $\mathcal{B}(D^+ \rightarrow \pi^+ e^+ e^-)$ at the 90 % confidence level, in units of 10^{-7} , constructed with and without systematic uncertainties. The limits are quoted for a signal distributed according to three-body phase space.

$M(e^+e^-)$ interval GeV/c^2	Statistical only	With systematics
0.20 – 0.50	2.7	3.0
0.56 – 0.95	3.0	3.3
1.05 – 1.75	2.8	2.9
Full range	4.2	4.7

Figure 9.6 shows the corresponding confidence belt for the three individual $M(e^+e^-)$ intervals.

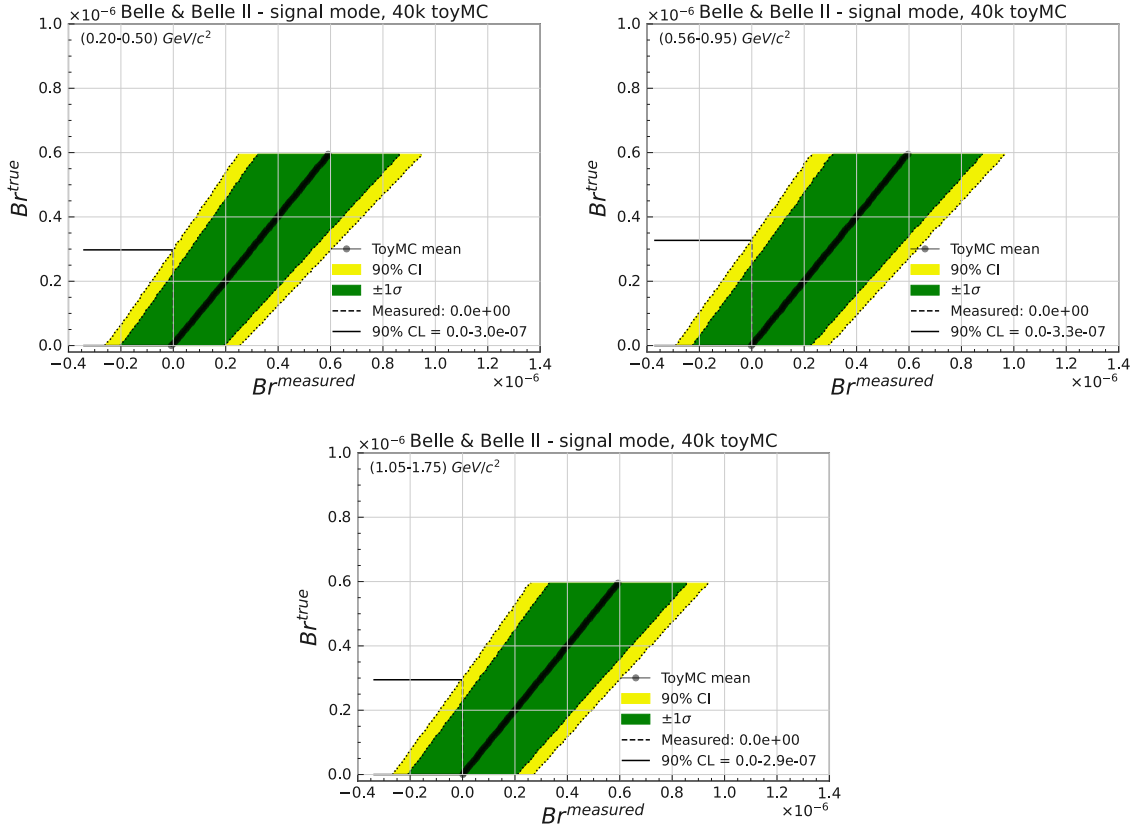


Figure 9.6: As in Figure 9.5, for the intervals I_1 (top left), I_2 (top right) and I_3 (bottom).

9.4.1 Independent Fits

The limits above come from a simultaneous fit in which the branching fraction is shared between the two experiments. To quantify what the combination provides, we repeat the analysis fitting each sample on its own, over the full $M(e^+e^-)$ range.

The independent fits (Figure 9.7) give signal yields of 19 ± 15 for Belle and

-14 ± 8 for Belle II, corresponding to branching fractions of $(5.8 \pm 4.6) \times 10^{-7}$ and $(-8.6 \pm 4.8) \times 10^{-7}$. Belle shows a fluctuation of about one standard deviation above zero, while Belle II shows one of nearly two standard deviations below zero. The two partly cancel in the simultaneous fit, which is one reason the combined branching fraction sits close to zero.

The corresponding upper limits (Figure 9.8) are 12×10^{-7} (39.4 events) for Belle and 6.9×10^{-7} (12.8 events) for Belle II. The Belle limit is significantly weaker than its statistical precision alone would suggest, since the Neyman construction assigns a wider interval to an upward fluctuation, and does not improve on the previous *BABAR* limit of 1.1×10^{-6} . Belle II shows the opposite pattern: its downward fluctuation places the fitted branching fraction below the physical boundary, so the construction returns the tightest limit it can provide.

All limits in the following comparison are obtained without systematic uncertainties. A naive inverse-quadrature combination of the two observed upper limits, 12×10^{-7} and 6.9×10^{-7} , gives approximately 6.0×10^{-7} , weaker than the 4.2×10^{-7} obtained from the simultaneous fit. The two observed limits, however, are not directly comparable measures of experimental sensitivity because they are affected differently by the statistical fluctuations described above. A fairer comparison uses the limits that each experiment would obtain for a null observation. Such limits approximately scale with the statistical uncertainties of the fits, so their inverse-quadrature combination provides an estimate of the expected combined sensitivity. For a fitted yield of zero, the Belle limit would be approximately 19.5 events, corresponding to 6.0×10^{-7} , while the quoted Belle II limit of 6.9×10^{-7} already corresponds to a null observation because its fitted branching fraction lies below the physical boundary. Their inverse-quadrature combination, approximately 4.5×10^{-7} , is close to the 4.2×10^{-7} simultaneous-fit result. This indicates that the stronger simultaneous-fit limit largely reflects the expected gain in statistical precision from combining the information in the two datasets. The simultaneous fit combines the likelihood information from both experiments before constructing a single confidence belt, allowing the opposite observed fluctuations to partly cancel before they enter the upper-limit construction.

The Belle II null-observation limit is only about 15% weaker than the Belle limit, despite its integrated luminosity being smaller by a factor of approximately 2.3. Its comparable sensitivity follows from a combination of its higher reconstruction efficiency (Table 9.1) and the stronger background rejection of its classifier (Section 5.8).

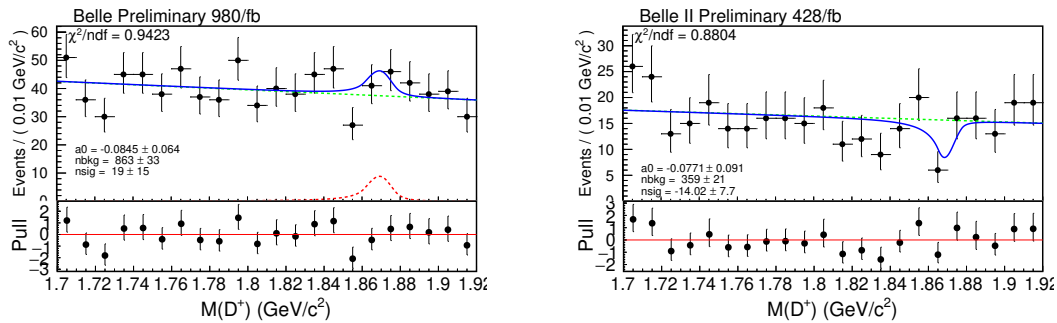


Figure 9.7: Independent fits to the $M(\pi^+e^+e^-)$ distribution of the signal mode in the Belle (left) and Belle II (right) data samples.

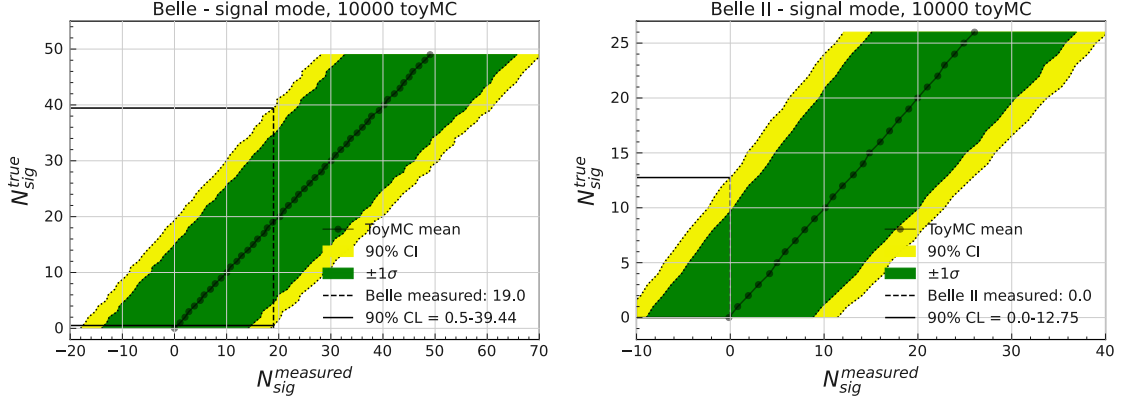


Figure 9.8: Neyman confidence belts for the signal mode over the full $M(e^+e^-)$ range, constructed from the independent fits to the Belle (left) and Belle II (right) data samples, without systematic uncertainties. Belle II’s construction is restricted to the physical region $N \geq 0$.

9.5 Comparison With Previous Measurements

The same *BABAR* analysis that observed the validation mode also set the previous best limit at the 90% confidence level on the signal mode, 1.1×10^{-6} [4], which our full-range limit improves by a factor of 2.3. The improvement follows from the larger combined data sample and from the suppression of the combinatorial background by the random forest classifier of Section 5.8.

The two limits are not defined over identical ranges. Both analyses quote a partial branching fraction, obtained with the efficiency averaged over the retained range, assuming three-body phase-space decay. The *BABAR* analysis excludes only the ϕ region, $0.95 < M(e^+e^-) < 1.05 \text{ GeV}/c^2$, while we exclude the region around the η mass as well, so our limit refers to a slightly smaller part of the phase space and the quoted improvement is marginally overstated.

We also report, for the first time in this decay, upper limits in intervals of the electron-pair invariant mass. As discussed in Chapter 2, the total rate is dominated by long-distance dynamics, in which the electron pair is produced through an intermediate vector resonance rather than through the loop process of interest. The ϕ and η regions are excluded from the measurement, but I_2 still contains the $\rho(770)$ and $\omega(782)$ resonances. I_1 and I_3 lie away from the vector poles, where the resonant amplitude falls rapidly and a short-distance contribution would be relatively enhanced; limits determined separately in them therefore carry information that the full-range limit does not.

Our limits remain well above the Standard Model short-distance rate. The exclusive short-distance contribution to $D^+ \rightarrow \pi^+ e^+ e^-$ is predicted at 4.6×10^{-12} over the full phase space (Chapter 2), five orders of magnitude below our full-range limit. Predictions binned in $M(\ell^+ \ell^-)$, with bins adaptable to the experimental selection, are already available for the muon mode [21], so limits reported in intervals can be compared with theory directly once the same becomes available for electrons.

10. Summary and Outlook

The decay $D^+ \rightarrow \pi^+ e^+ e^-$ proceeds through a flavour-changing neutral current. Such transitions are forbidden at tree level in the Standard Model, and in the up-quark sector they are suppressed further by the Glashow–Iliopoulos–Maiani mechanism and the CKM structure. The observable rate is nevertheless dominated by long-distance contributions, in which the electron pair originates from an intermediate vector resonance. They dominate not only at the resonance poles but across the whole kinematic range, exceeding the short-distance contribution by a few orders of magnitude even far from the poles. No region of $M(e^+ e^-)$ is free of them. Away from the resonance poles, however, the dominance of the long-distance contribution is least pronounced, and an enhancement of the rate from physics beyond the Standard Model would be easiest to see there. These regions had not been probed separately before this work. We searched for this decay in the data collected by the Belle and Belle II detectors at the KEKB and SuperKEKB colliders, corresponding to 980 fb^{-1} and 428 fb^{-1} .

10.1 Summary of the Measurement

We measure the branching fraction relative to $D^+ \rightarrow \pi^+ K^+ K^-$, reconstructed within $20 \text{ MeV}/c^2$ of the ϕ mass. The integrated luminosity and the charm production cross-section cancel exactly in this ratio, while the similarity of the final states leads to a partial cancellation of efficiency effects. The remaining efficiency uncertainties are treated in Chapter 8. We select candidates with requirements on track quality, particle identification and D^+ centre-of-mass momentum, with the electron candidates corrected for bremsstrahlung. Two vetoes remove specific backgrounds. The first rejects candidates in which an electron originates from a photon conversion rather than from the D^+ decay. The second targets $D_{(s)}^+ \rightarrow \pi^+ (\eta \rightarrow \gamma e^+ e^-)$ decays, in which the photon from the η decay is lost. Where that photon is instead absorbed into the bremsstrahlung correction, the candidate peaks at m_η in $M(e^+ e^-)$ and falls inside the resonance-exclusion window. Where it is lost altogether, the candidate does not peak but populates the region below m_η , and we suppress it by recombining the electron pair with an additional photon and rejecting the candidate if the $e^+ e^- \gamma$ mass is compatible with m_η . About a third of this component is removed; what remains is the dominant background at low $M(e^+ e^-)$. The random forest classifier of Section 5.8 suppresses the remaining combinatorial background, whose discrimination rests largely on the significance of the D^+ flight distance. We extract the branching fraction from an unbinned extended maximum-likelihood fit to $M(\pi^+ e^+ e^-)$, performed simultaneously in the two data samples with the branching fraction as a single shared parameter, and repeat it independently in three intervals of $M(e^+ e^-)$. We take the normalisation yields entering that fit from separate binned fits to each sample. We determine the upper limits from a frequentist Neyman con-

struction based on pseudo-experiments, into which the systematic uncertainties of Chapter 8 are folded directly.

We validate the analysis on $D^+ \rightarrow \pi^+\phi(\rightarrow e^+e^-)$. This decay has the same final-state particles as the signal mode and is selected in the same way, differing only in the retained $M(e^+e^-)$ region. We measure $\mathcal{B}(D^+ \rightarrow \pi^+\phi(\rightarrow e^+e^-)) = (1.75 \pm 0.22 (\text{stat}) \pm 0.12 (\text{syst})) \times 10^{-6}$, with a significance of 12.0 standard deviations evaluated without systematic uncertainties. The value agrees with the expectation of $(1.70 \pm 0.05) \times 10^{-6}$ calculated from the known branching fractions of the two steps [16], and with the previous direct measurement by CLEO, $(1.7^{+1.4}_{-0.9}) \times 10^{-6}$ [48], whose statistical precision it improves by approximately a factor of five. It is the first determination of the branching fraction of this decay with a significance exceeding five standard deviations. Its purpose is not to compete with the product of the two known branching fractions, which is more precise, but to establish that the normalisation, the efficiencies, the fit and the construction of confidence intervals behave as intended on an observed signal in data, and not only in simulation.

In the signal mode we find no evidence for a signal. Over the full $M(e^+e^-)$ range, excluding the ϕ and η regions, we set

$$\mathcal{B}(D^+ \rightarrow \pi^+e^+e^-) < 4.7 \times 10^{-7} \quad \text{at 90 \% CL}, \quad (10.1)$$

improving the previous best limit of 1.1×10^{-6} , set by *BABAR* [4], by a factor of 2.3, although the two limits refer to slightly different ranges of $M(e^+e^-)$ (Section 9.5). We further report limits of 3.0×10^{-7} , 3.3×10^{-7} and 2.9×10^{-7} in the intervals I_1 , I_2 and I_3 . These are the first limits on this decay determined separately in regions of the electron-pair invariant mass. Of these, I_1 and I_3 lie away from the resonance poles, while I_2 still contains the $\rho(770)$ and $\omega(782)$. The interval I_3 , above the ϕ , contains the high- q^2 region identified as particularly sensitive to physics beyond the Standard Model [36].

Until this work, the most stringent constraint on this decay in the electron channel was the *BABAR* limit, while the muon mode has been searched for repeatedly by LHCb, most recently in 2021, each time improving on the previous result. The two sectors have therefore been constrained very unevenly, and a test of lepton-flavour universality in this transition requires both. Three independent theoretical analyses of $c \rightarrow u \ell^+ \ell^-$ transitions have emphasized the potential of improved measurements of dielectron modes, including at Belle II, to strengthen experimental constraints on rare charm processes [28, 36, 37]. The limits in I_1 , I_2 and I_3 provide a more useful form of this input, since a limit quoted for a specified region of $M(e^+e^-)$ can be more easily reinterpreted under a different assumption about the signal spectrum. The limits lie several orders of magnitude above the Standard Model short-distance rate quoted in Chapter 2. The measurement therefore primarily probes possible enhancements of the rate from physics beyond the Standard Model, rather than the Standard Model contribution itself.

10.2 Outlook

10.2.1 Improvements to the Present Analysis

Several parts of the analysis could be improved without new data. The selection was optimised with the Punzi figure of merit, whereas optimising directly on the upper limit would be more appropriate. This would require the limit to be evaluated

sufficiently quickly within an optimisation loop, for example using an asymptotic approximation rather than a full pseudo-experiment construction at every point.

As reported in Section 5.10, optimising the classifier working point separately in each $M(e^+e^-)$ interval gains little, indicating that further optimisation of this requirement would improve the sensitivity by at most a few per cent. The larger gains lie in the treatment of the background, whose modelling is the largest single systematic contribution and the only one entering additively (Section 8.5). A wider fit range would constrain the shape better, but the present range is already as wide as the spectrum allows without reaching the D_s^+ peak, which would then have to be modelled alongside the signal. Separately, since generic simulation underestimates the background yield (Section 6.2.3), any future sensitivity study should extrapolate the background from the $M(\pi^+e^+e^-)$ sidebands in data rather than rely on simulation.

Changes to the selection are also worth investigating. The non-peaking $D_{(s)}^+ \rightarrow \pi^+(\eta \rightarrow \gamma e^+e^-)$ component of Section 5.9 remains the dominant background at low $M(e^+e^-)$ even after the dedicated veto. Better recovery of the lost photon, or a classifier trained specifically against this mode, would therefore target the region where the present limits are weakest. Describing this component explicitly in the fit would not reduce the background yield, but could better constrain the background shape and reduce the associated systematic uncertainty. A purer sample could be obtained by requiring the signal D^+ to originate from $D^{*+} \rightarrow D^+\pi^0$, since the slow π^0 constrains the candidate through the mass difference $M(D^+\pi^0) - M(\pi^+e^+e^-)$, at the cost of the D^{*+} production fraction, the branching fraction of the strong decay and the π^0 reconstruction efficiency. Requiring this tag would discard most of the sample, whereas fitting the tagged and untagged candidates as two mutually exclusive categories sharing a single branching fraction would retain all selected candidates while adding a high-purity category to the fit. The requirement on p_{CMS} removes D^+ mesons produced in B decays, which a classifier trained to separate the two production mechanisms could recover, although they would then require separate efficiencies and systematic treatments in the branching-fraction calculation.

10.2.2 Extensions of the Analysis

The selection developed here transfers with little modification to several related decays. The closest is $D_s^+ \rightarrow \pi^+e^+e^-$, which proceeds by weak annihilation and receives no $c \rightarrow u$ penguin contribution. It therefore probes the long-distance amplitudes rather than the Wilson coefficients, and a measurement would constrain the hadronic models entering the dispersion relation.

Its rate is predicted to be about an order of magnitude higher than that of the signal mode [21], yet the existing limit, 5.5×10^{-6} [3], is an order of magnitude weaker than the one reported here, so the mode is still under-explored relative to its expected rate.

The decay $D^+ \rightarrow \pi^-e^+e^+$ violates lepton number, while $D^+ \rightarrow \pi^+e^\pm\mu^\mp$ violates lepton flavour. Both involve the same final-state particles in different charge and flavour combinations. Neither process is allowed in the Standard Model, so the observation of either would constitute an unambiguous signal of physics beyond the Standard Model.

Besides the rate, there are other promising observables whose Standard Model values are small and bounded, making an excess above the corresponding predictions

easier to interpret as evidence for physics beyond the Standard Model. The CP asymmetry is predicted to be below 5×10^{-3} [25], but interference between the resonant amplitude and a short-distance amplitude carrying a CP -odd phase can raise it by orders of magnitude [24]. Its behaviour around the ϕ depends on the strong phase of the resonant amplitude. For phases near 0 or π the sensitivity is greatest at the peak. For phases near $\pm\pi/2$ the asymmetry is odd about the peak and a symmetric bin would average it away. The measurement is in either case made around the ϕ resonance, and the validation sample reconstructed here, with its significance of 12.0 standard deviations, is the natural starting point for it. The rate measured there fixes the magnitude of the ϕ contribution but not its strong phase, which cannot be obtained from a rate alone and requires an observable sensitive to interference between amplitudes.

Bharucha *et al.* [36] note that the angular observables F_H and A_{FB} are strongly suppressed in the Standard Model, so a sensitivity at the 10 % level would already reveal a deviation in some scenarios, while 1 % would allow a fit to the Wilson coefficients. For the electron mode, they identify F_H as the observable of interest, and note that leptoquark scenarios would require the per cent level. None of these angular observables is within reach in the signal mode, where the fitted yields of Section 9.3 are consistent with zero.

10.2.3 Prospects With Larger Data Sets

The sensitivity of this search is limited by the size of the data samples rather than by the systematic uncertainties. Because the background under the D^+ peak grows with the integrated luminosity, the limit improves approximately as $1/\sqrt{L}$. Only the Belle II sample will grow; thus, we scale the Belle II limit alone rather than the combined limit. The Belle sample is fixed at 980 fb^{-1} and would add little at these luminosities. Taking the Belle II limit from the independent fit (Section 9.4.1), 6.9×10^{-7} at 0.43 ab^{-1} , and assuming the efficiency, background composition and selection remain unchanged, gives 1.4×10^{-7} for a sample of 10 ab^{-1} and 6.4×10^{-8} for the full 50 ab^{-1} . Both projections are built without systematic uncertainties, as is the limit they scale from. The largest systematic uncertainty, the choice of background model, is the only contribution that enters additively and could therefore limit the projections. It is determined from the data and is expected to shrink as the background shape becomes better constrained, although we cannot exclude that a systematic floor is reached before the full sample is collected. The projections also assume detector performance and beam-background conditions comparable to those of the present data set, which is the least secure of these assumptions at the highest luminosities.

The Standard Model short-distance rate, of order 10^{-12} (Chapter 2), will not be reached at any of these luminosities. The long-distance contribution is closer. Bansal *et al.* [21] predict 1.5×10^{-8} for the muon mode in a bin above the ϕ that is contained within I_3 , so the corresponding prediction for I_3 itself is at least this large. Since lepton-mass effects are negligible in this region, the electron rate is expected to be nearly the same, approximately an order of magnitude below the limit reported here. The full Belle II data set would narrow this gap substantially and begin to approach sensitivity to the predicted long-distance contribution. This interval is therefore the natural focus of future work.

10.3 Concluding Remarks

The decay $D^+ \rightarrow \pi^+ e^+ e^-$ remains unobserved. We have tightened the upper limit on its branching fraction by more than a factor of two, using the combined data of two experiments rather than one, and have resolved it into regions of the electron-pair invariant mass rather than only quoting a single number across the whole range. The accompanying measurement of $D^+ \rightarrow \pi^+ \phi (\rightarrow e^+ e^-)$ provides a data-driven validation that the analysis procedure recovers a signal of the expected size. This supports interpreting the absence of a significant excess in the signal-search regions as a physical result rather than an artefact of the analysis procedure. These new limits provide input to global fits of the $c \rightarrow u \ell^+ \ell^-$ transition, for which the *BABAR* result has until now provided the most stringent constraint from this decay in the electron channel. The data still to be recorded by Belle II will probe possible contributions from physics beyond the Standard Model with substantially greater sensitivity.

11. Bibliography

- [1] S. L. Glashow, J. Iliopoulos and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2**, 1285 (1970).
- [2] H. Gisbert, M. Golz and D. S. Mitzel, *Theoretical and experimental status of rare charm decays*, Mod. Phys. Lett. A **36**, 2130002 (2021).
- [3] R. Aaij *et al.* (LHCb), *Searches for 25 rare and forbidden decays of D^+ and D_s^+ mesons*, JHEP **6** (2021).
- [4] J. P. Lees *et al.* (BABAR), *Searches for rare or forbidden semileptonic charm decays*, Phys. Rev. D **84**, 072006 (2011).
- [5] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22**, 579 (1961).
- [6] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19**, 1264 (1967).
- [7] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C **680519**, 367 (1968).
- [8] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13**, 321 (1964).
- [9] P. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12**, 132 (1964).
- [10] A. B. Arbuzov, *Quantum Field Theory and the Electroweak Standard Model*, CERN Yellow Rep. School Proc. **4**, 1 (2017).
- [11] T. Nakano and K. Nishijima, *Charge Independence for V-particles*, PTEP **10**, 581 (1953).
- [12] M. Gell-Mann, *The interpretation of the new particles as displaced charge multiplets*, Nuovo Cim. **4**, 848 (1956).
- [13] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10**, 531 (1963).
- [14] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, PTEP **49**, 652 (1973).
- [15] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51**, 1945 (1983).

- [16] S. Navas *et al.* (Particle Data Group), *Review of Particle Physics*, Phys. Rev. D **110**, 030001 (2024).
- [17] I. Adachi *et al.* (Belle II), *Evidence for $B^+ \rightarrow K^+ \nu \bar{\nu}$ decays*, Phys. Rev. D **109**, 112006 (2024).
- [18] R. Aaij *et al.* (LHCb), *First measurement of the differential branching fraction and CP asymmetry of the $B^\pm \rightarrow \pi^\pm \mu^+ \mu^-$ decay*, JHEP **10** (2015), 034.
- [19] S. Choudhury, *Rare decays from Belle and Belle II*, PoS **HQL2023**, 047 (2024).
- [20] I. Adachi *et al.* (Belle and Belle II), *Search for Rare $b \rightarrow d \ell^+ \ell^-$ Transitions at Belle*, Phys. Rev. Lett. **133**, 101804 (2024).
- [21] A. Bansal, A. Khodjamirian and T. Mannel, *$D \rightarrow P \ell^+ \ell^-$ decays assisted by QCD light-cone sum rules*, JHEP **08** (2025), 026.
- [22] S. Fajfer, S. Prelovšek and P. Singer, *Rare charm meson decays $D \rightarrow P \ell^+ \ell^-$ and $c \rightarrow u \ell^+ \ell^-$ in SM and MSSM*, Phys. Rev. D **64**, 114009 (2001).
- [23] G. Burdman, E. Golowich, J. L. Hewett and S. Pakvasa, *Rare charm decays in the standard model and beyond*, Phys. Rev. D **66**, 014009 (2002).
- [24] S. Fajfer and N. Košnik, *Resonance catalyzed CP asymmetries in $D \rightarrow P \ell^+ \ell^-$* , Phys. Rev. D **87**, 054026 (2013).
- [25] S. de Boer and G. Hiller, *Flavor and new physics opportunities with rare charm decays into leptons*, Phys. Rev. D **93**, 074001 (2016).
- [26] S. de Boer, B. Müller and D. Seidel, *Higher-order Wilson coefficients for $c \rightarrow u$ transitions in the standard model*, JHEP **08** (2016), 091.
- [27] S. de Boer, *Probing the standard model with rare charm decays*, Ph.D. thesis, Tech. U., Dortmund (main) (2017).
- [28] H. Gisbert, G. Hiller and D. Suelmann, *Effective field theory analysis of rare $|\Delta c| = |\Delta u| = 1$ charm decays*, JHEP **12** (2024), 102.
- [29] C. Bobeth, G. Hiller and G. Piranishvili, *Angular distributions of $\bar{B} \rightarrow \bar{K} \ell^+ \ell^-$ decays*, JHEP **12** (2007), 040.
- [30] M. Golz, *$|\Delta c| = |\Delta u| = 1$ transitions in and beyond the standard model*, Ph.D. thesis, Tech. U., Dortmund (main) (2022).
- [31] S. Fajfer and S. Prelovsek, *Effects of littlest Higgs model in rare D meson decays*, Phys. Rev. D **73**, 054026 (2006).
- [32] S. Fajfer, P. Singer and J. Zupan, *The Radiative leptonic decays $D^0 \rightarrow e^+ e^- \gamma$, $\mu^+ \mu^- \gamma$ in the standard model and beyond*, Eur. Phys. J. C **27**, 201 (2003).
- [33] A. Paul, I. I. Bigi and S. Recksiegel, *On $D \rightarrow X_u \ell^+ \ell^-$ within the Standard Model and Frameworks like the Littlest Higgs Model with T Parity*, Phys. Rev. D **83**, 114006 (2011).

-
- [34] R.-M. Wang, J.-H. Sheng, J. Zhu, Y.-Y. Fan and Y.-G. Xu, *Decays $D_{(s)}^+ \rightarrow \pi(K)^+ \ell^+ \ell^-$ and $D^0 \rightarrow \ell^+ \ell^-$ in the MSSM with and without R-parity*, Int. J. Mod. Phys. A **30**, 1550063 (2015).
 - [35] R. Bause, M. Golz, G. Hiller and A. Tayduganov, *The new physics reach of null tests with $D \rightarrow \pi \ell \ell$ and $D_s \rightarrow K \ell \ell$ decays*, Eur. Phys. J. C **80**, 65 (2020), [Erratum: Eur.Phys.J.C 81, 219 (2021)].
 - [36] A. Bharucha, D. Boito and C. Méaux, *Disentangling QCD and new physics in $D^+ \rightarrow \pi^+ \ell^+ \ell^-$* , JHEP **04** (2021), 158.
 - [37] S. Fajfer, J. F. Kamenik, A. Korajac and N. Košnik, *Correlating New Physics Effects in Semileptonic $\Delta C = 1$ and $\Delta S = 1$ Processes*, JHEP **7** (2023), 029.
 - [38] S. Fajfer and N. Košnik, *Prospects of discovering new physics in rare charm decays*, Eur. Phys. J. C **75**, 567 (2015).
 - [39] A. Greljo, G. Isidori and D. Marzocca, *On the breaking of Lepton Flavor Universality in B decays*, JHEP **07** (2015), 142.
 - [40] G. Branco *et al.*, *Theory and phenomenology of two-Higgs-doublet models*, Phys. Rept. **516**, 1–102 (2012).
 - [41] I. Doršner, S. Fajfer, A. Greljo, J. F. Kamenik and N. Košnik, *Physics of leptoquarks in precision experiments and at particle colliders*, Phys. Rept. **641**, 1 (2016).
 - [42] P. Haas *et al.*, *Upper limits on charm-changing neutral-current interactions*, Phys. Rev. Lett. **60**, 1614 (1988).
 - [43] M.-G. Zhao (BESIII), *D rare/forbidden decays at BESIII*, in *7th International Workshop on Charm Physics* (2016) arXiv:1605.08952 [hep-ex] .
 - [44] V. Prasad, *Experimental status of rare D decays*, PoS **CKM2016**, 060 (2017).
 - [45] R. Aaij *et al.* (LHCb), *Search for $D_{(s)}^+ \rightarrow \pi^+ \mu^+ \mu^-$ and $D_{(s)}^+ \rightarrow \pi^- \mu^+ \mu^+$ decays*, Phys. Lett. B **724**, 203 (2013).
 - [46] I. Adachi *et al.* (Belle and Belle II), *Measurement of the $D^0 \rightarrow K^- \pi^+ e^+ e^-$ branching fraction and search for $D^0 \rightarrow \pi^+ \pi^- e^+ e^-$ and $D^0 \rightarrow K^+ K^- e^+ e^-$ decays at Belle*, Phys. Rev. D **112**, L071101 (2025).
 - [47] V. M. Abazov *et al.* (D0), *Search for flavor-changing-neutral-current D meson decays*, Phys. Rev. Lett. **100**, 101801 (2008).
 - [48] P. Rubin *et al.* (CLEO), *Search for rare and forbidden decays of charm and charmed-strange mesons to final states $h^\pm e^\mp e^+$* , Phys. Rev. D **82**, 092007 (2010).
 - [49] E. M. Aitala *et al.* (E791), *Search for rare and forbidden dilepton decays of the D^+ , $D_{(s)}^+$, and D^0 charmed mesons*, Phys. Lett. B **462**, 401 (1999).

- [50] R. Aaij *et al.* (LHCb), *Measurements of prompt charm production cross-sections in pp collisions at $\sqrt{s} = 13$ TeV*, JHEP **03** (2016), 159, [Erratum: JHEP 09, 013 (2016), Erratum: JHEP 05, 074 (2017)].
- [51] S. Bifani, S. Descotes-Genon, A. R. Vidal and M.-H. Schune, *Review of lepton universality tests in B decays*, J. Phys. G: Nucl. Part. Phys. **46**, 023001 (2019).
- [52] K. Akai, K. Furukawa and H. Koiso, *SuperKEKB collider*, Nucl. Instrum. Meth. A **907**, 188–199 (2018).
- [53] K. Akai *et al.*, *Commissioning of KEKB*, Nucl. Instrum. Meth. A **499**, 191 (2003).
- [54] T. Abe *et al.*, *Achievements of KEKB*, PTEP **2013**, 03A001 (2013).
- [55] I. Adachi *et al.* (Belle II), *Measurement of the integrated luminosity of data samples collected during 2019-2022 by the Belle II experiment*, Chin. Phys. C **49**, 013001 (2025).
- [56] T. Abe *et al.* (Belle II), *Belle II Technical Design Report* (2010), arXiv:1011.0352 [physics.ins-det] .
- [57] M. Masuzawa (SuperKEKB), *SuperKEKB Status*, PoS **EPS-HEP2023**, 610 (2024).
- [58] M. Bona *et al.* (SuperB), *SuperB: A High-Luminosity Asymmetric e^+e^- Super Flavor Factory. Conceptual Design Report* (2007), arXiv:0709.0451 [hep-ex] .
- [59] M. Downes, *IPNS Seminar Report: First Measurement of Time-Dependent CP Violation in B Meson Decays into two neutral pions* (2026), <https://www2.kek.jp/ipns/en/news/8453/> [view 9. 6. 2026].
- [60] V. Aulchenko *et al.* (Belle-ECL), *Electromagnetic calorimeter for Belle II*, J. Phys. Conf. Ser. **587**, 012045 (2015).
- [61] C. Marinas, *The Belle II pixel detector: High precision with low material*, Nucl. Instrum. Meth. A **731**, 31 (2013), PIXEL 2012.
- [62] K. Adamczyk *et al.*, *The Belle II silicon vertex detector assembly and mechanics*, Nucl. Instrum. Meth. A **845**, 38 (2017), proceedings of the Vienna Conference on Instrumentation 2016.
- [63] A. B. Kaliyar *et al.*, *The Silicon Vertex Detector of the Belle II Experiment*, PoS **TIPP2023**, 114 (2025).
- [64] V. Bertacchi *et al.*, *Track finding at Belle II*, Comput. Phys. Commun. **259**, 107610 (2021).
- [65] H. Atmacan *et al.*, *The imaging Time-of-Propagation detector at Belle II*, Nucl. Instrum. Meth. A **1080**, 170627 (2025).
- [66] I. Adachi *et al.*, *Aerogel RICH Counter at the Belle II Detector*, Nucl. Instrum. Meth. A **1089**, 171592 (2026).

-
- [67] T. Iijima *et al.*, *A novel type of proximity focusing RICH counter with multiple refractive index aerogel radiator*, Nucl. Instrum. Meth. A **548**, 383–390 (2005).
 - [68] P. Križan, S. Korpar and T. Iijima, *Study of a nonhomogeneous aerogel radiator in a proximity focusing RICH detector*, Nucl. Instrum. Meth. A **565**, 457–462 (2006).
 - [69] A. Abashian *et al.*, *The Belle detector*, Nucl. Instrum. Meth. A **479**, 117 (2002), detectors for Asymmetric B-factories.
 - [70] *Belle II Software Documentation* (2023), <https://software.belle2.org/light-2305-korat> [view 9. 6. 2026].
 - [71] Y. Sato, S. Cunliffe, F. Meier and A. Zupanc, *Monte Carlo matching in the Belle II software*, EPJ Web Conf. **251**, 03021 (2021).
 - [72] M. Gelb *et al.*, *B2BII: Data Conversion from Belle to Belle II*, Comput. Softw. Big Sci. **2**, 9 (2018).
 - [73] N. Katayama, R. Itoh, A. Manabe and T. Sasaki, *Belle computing model*, Comput. Phys. Commun. **110**, 22 (1998).
 - [74] R. Brun and F. Rademakers, *ROOT — An object oriented data analysis framework*, Nucl. Instrum. Meth. A **389**, 81 (1997), new Computing Techniques in Physics Research V.
 - [75] T. Bilka *et al.*, *Implementation of GENFIT2 as an experiment independent track-fitting framework* (2019), arXiv:1902.04405 [physics.data-an] .
 - [76] D. Dossett (Belle II), *Computing and Software at Belle II*, EPJ Web Conf. **245**, 11007 (2020).
 - [77] J. Brodzicka *et al.* (Belle), *Physics achievements from the Belle experiment*, PTEP **2012**, 04D001 (2012).
 - [78] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462**, 152 (2001), BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines.
 - [79] T. Sjöstrand, S. Mrenna and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006), 026.
 - [80] S. Jadach, B. F. L. Ward and Z. Wąs, *The precision Monte Carlo event generator KK for two-fermion final states in e^+e^- collisions*, Comput. Phys. Commun. **130**, 260 (2000).
 - [81] T. Sjöstrand *et al.*, *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191**, 159 (2015).
 - [82] E. Barberio and Z. Wąs, *PHOTOS: A Universal Monte Carlo for QED radiative corrections. Version 2.0*, Comput. Phys. Commun. **79**, 291 (1994).
 - [83] R. Brun *et al.*, *GEANT: Detector Description and Simulation Tool*, CERN Program Library (CERN, Geneva, 1993) long Writeup W5013.

- [84] S. Agostinelli *et al.* (GEANT4), *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth., Sect. A **506**, 250 (2003).
- [85] J.-F. Krohn *et al.*, *Global decay chain vertex fitting at Belle II*, Nucl. Instrum. Meth., Sect. A **976**, 164269 (2020).
- [86] T. K. Ho, *Random decision forests*, in *Proceedings of 3rd International Conference on Document Analysis and Recognition*, Vol. 1 (1995) pp. 278–282.
- [87] F. Pedregosa *et al.*, *Scikit-learn: Machine Learning in Python*, J. Mach. Learn. Res. **12**, 2825 (2011).
- [88] G. Cowan, *Statistical Data Analysis* (Oxford University Press, 1998).
- [89] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling* (2003), arXiv:physics/0306116 [physics.data-an] .
- [90] P. del Amo Sanchez *et al.* (BaBar), *Study of $B \rightarrow X\gamma$ Decays and Determination of $|V_{td}/V_{ts}|$* , Phys. Rev. D **82**, 051101(R) (2010).
- [91] J. Neyman, *Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability*, Philos. Trans. R. Soc. Lond. A **236**, 333 (1937).
- [92] T. Song and Z. You (BESIII), *Search for charm rare decays at BESIII*, in *32nd International Symposium on Lepton Photon Interactions at High Energies: Lepton-Photon 2025* (2026) arXiv:2601.08681 [hep-ex] .

Appendix A

Standard Model Prediction

Table A.1: Standard Model predictions for the short-distance branching fractions of $D^+ \rightarrow \pi^+ \ell^+ \ell^-$ in different q^2 regions, summarised from Ref. [25]. Uncertainties in brackets correspond to the quadratic sum of the individual contributions reported in the cited paper.

Branching fraction	Mode	
	$D^+ \rightarrow \pi^+ e^+ e^-$	$D^+ \rightarrow \pi^+ \mu^+ \mu^-$
$\sqrt{q^2} \geq 2m_\ell$	$4.6 \times 10^{-12} (1^{+211\%}_{-17\%})$	$3.7 \times 10^{-12} (1^{+249\%}_{-35\%})$
$0.250 \text{ GeV} \leq \sqrt{q^2} \leq 0.525 \text{ GeV}$	$8.1 \times 10^{-13} (1^{+160\%}_{-20\%})$	$7.4 \times 10^{-13} (1^{+240\%}_{-35\%})$
$\sqrt{q^2} \geq 1.25 \text{ GeV}$	$7.4 \times 10^{-13} (1^{+140\%}_{-52\%})$	

Appendix B

Supplementary Selection and Data/MC Studies

This appendix collects supplementary distributions supporting the event selection of Chapter 5 and the data/MC validation of Chapter 6, the latter completed mode by mode in Section B.4. Throughout, generic MC background components are stacked and the signal MC component is overlaid unstacked; where generic MC is compared to data it is scaled to the data sample size, and all uncertainties shown are statistical only.

B.1 Distributions of Variables After Loose Selection

Figures B.1 and B.2 show the remaining candidate variables studied at the loose-selection stage of Section 5.4, beyond those reproduced in Figures 5.2 and 5.3. The dashed vertical lines mark the maximum of the Punzi FOM of Eq. 5.1 and are not applied as selection requirements.

APPENDIX B. SUPPLEMENTARY SELECTION AND DATA/MC STUDIES

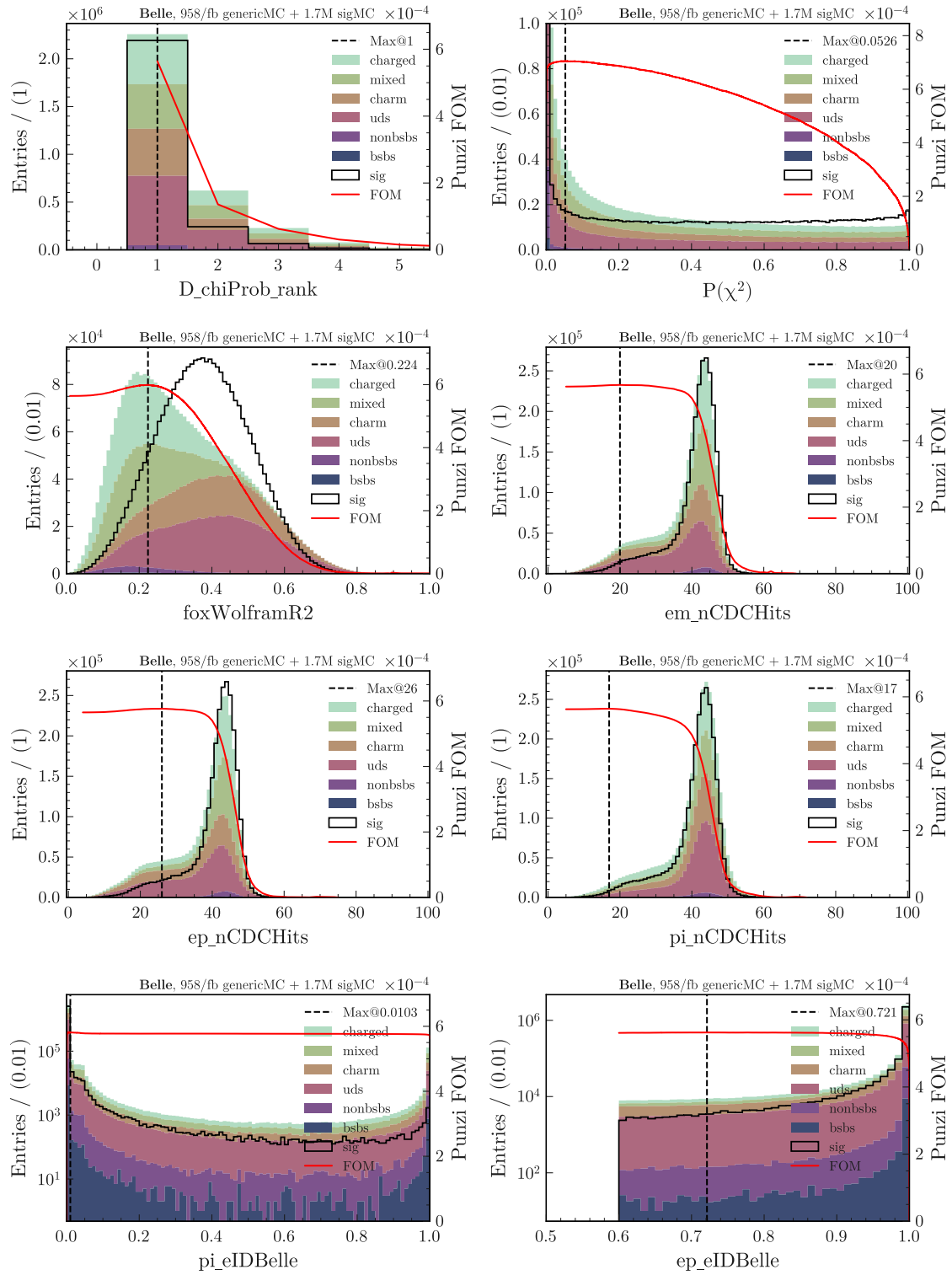


Figure B.1: Additional distributions of candidate variables after the loose selection, for Belle.

B.1. DISTRIBUTIONS OF VARIABLES AFTER LOOSE SELECTION

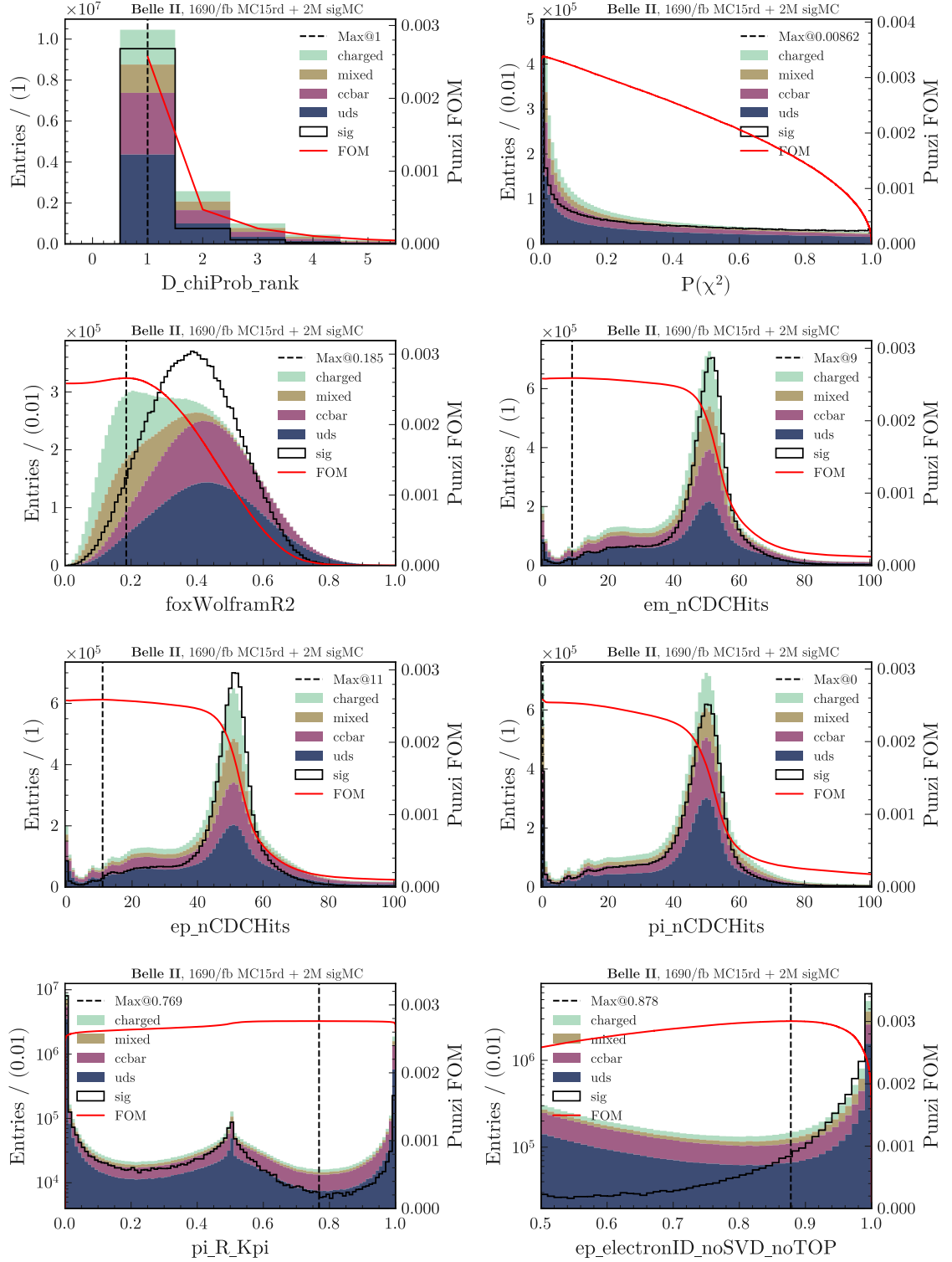


Figure B.2: Additional distributions of candidate variables after the loose selection, for Belle II.

B.2 RFC Input Variables in $M(e^+e^-)$ Intervals

Figures B.3–B.5 and B.6–B.8 show the four RFC input variables of Section 5.8 after the initial selection, separately in each of the three signal $M(e^+e^-)$ intervals defined in Section 5.11. The separation between signal and background is comparable across the three intervals, supporting the use of a single classifier trained inclusively over the full $M(e^+e^-)$ range.

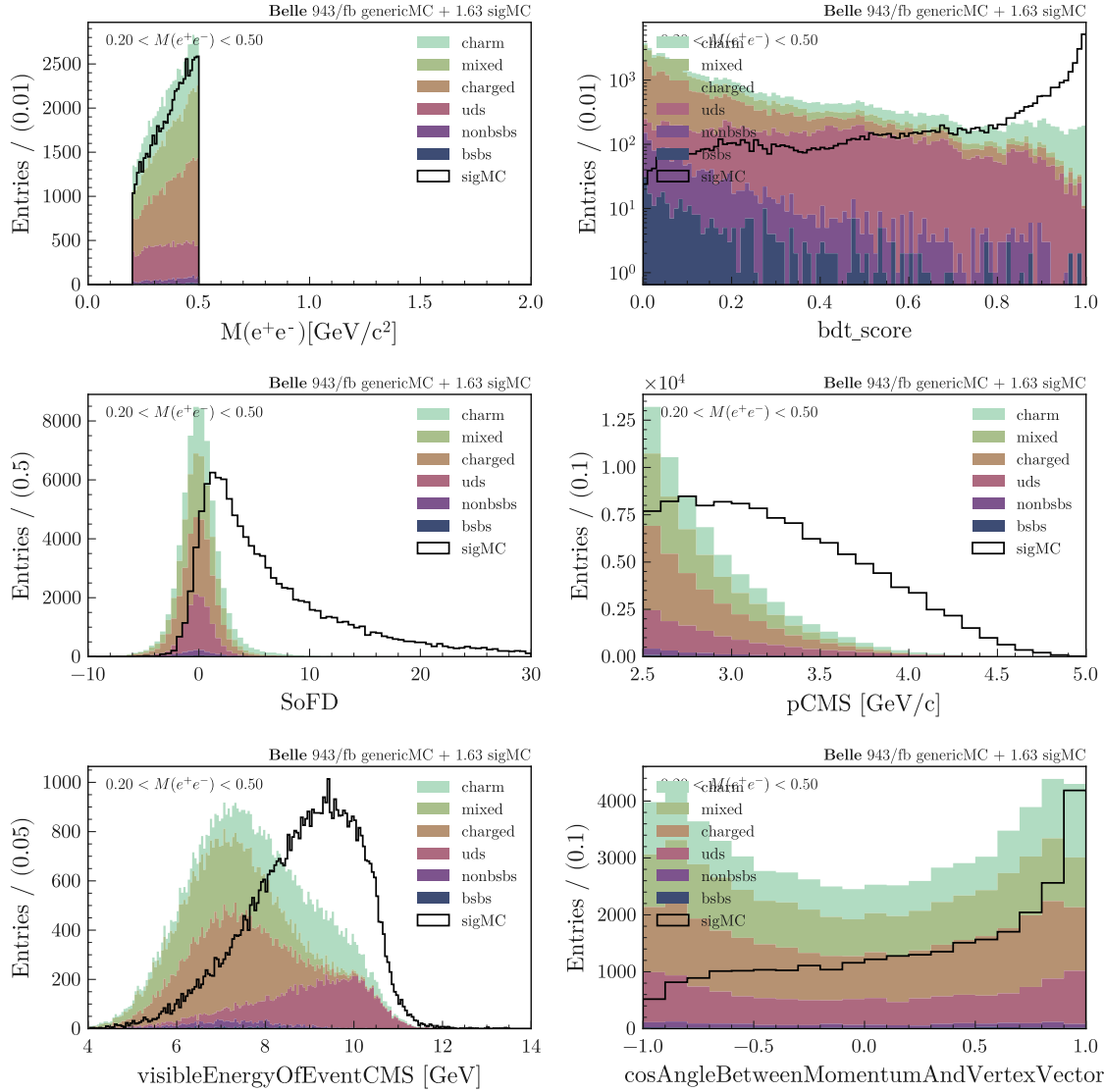


Figure B.3: Distributions of the RFC input variables after the initial selection for Belle, in the interval I_1 ($0.20 \text{ GeV}/c^2 - 0.50 \text{ GeV}/c^2$).

B.2. RFC INPUT VARIABLES IN $M(e^+e^-)$ INTERVALS

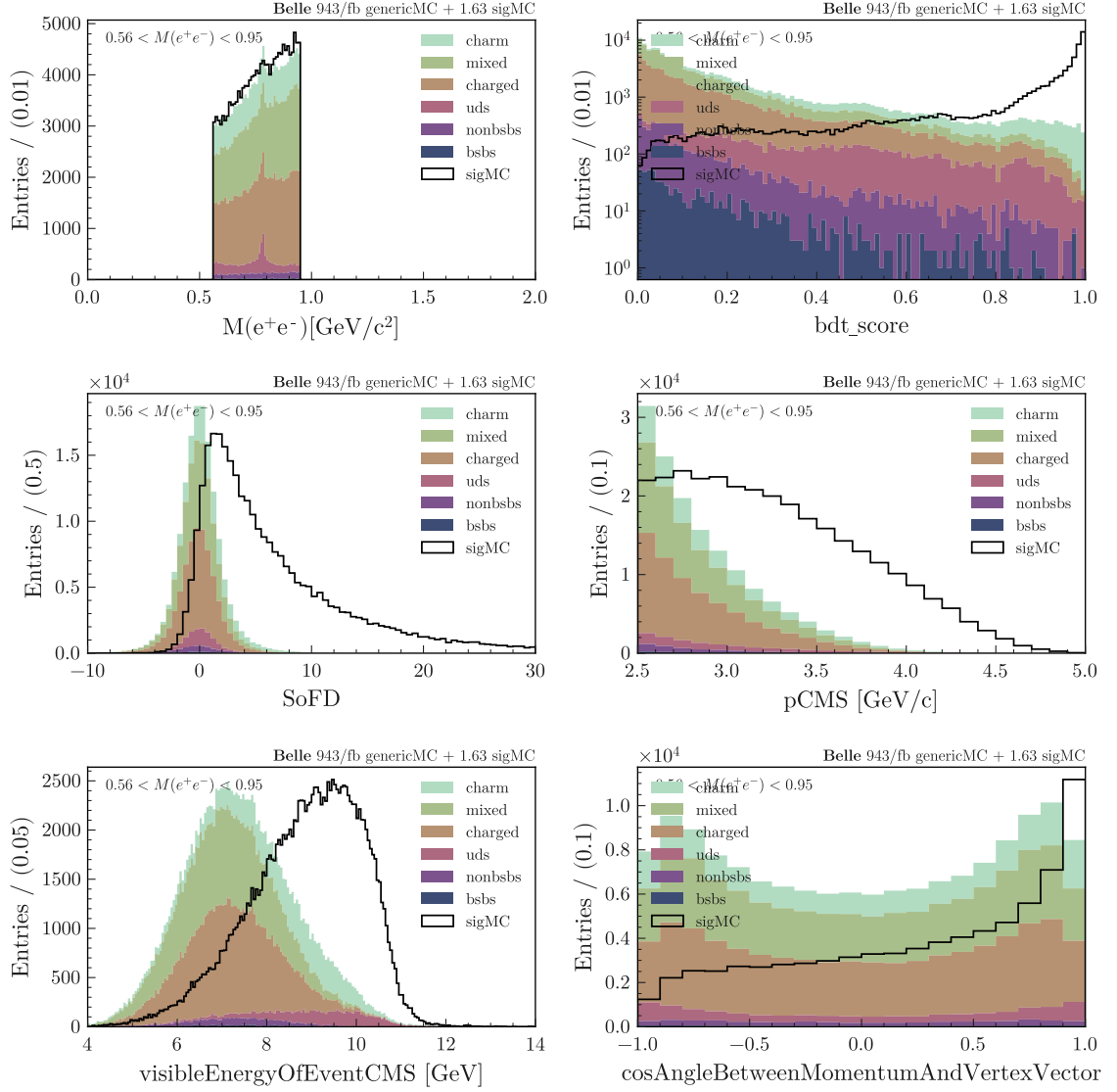


Figure B.4: Distributions of the RFC input variables after the initial selection for Belle, in the interval I_2 ($0.56 \text{ GeV}/c^2 - 0.95 \text{ GeV}/c^2$).

APPENDIX B. SUPPLEMENTARY SELECTION AND DATA/MC STUDIES

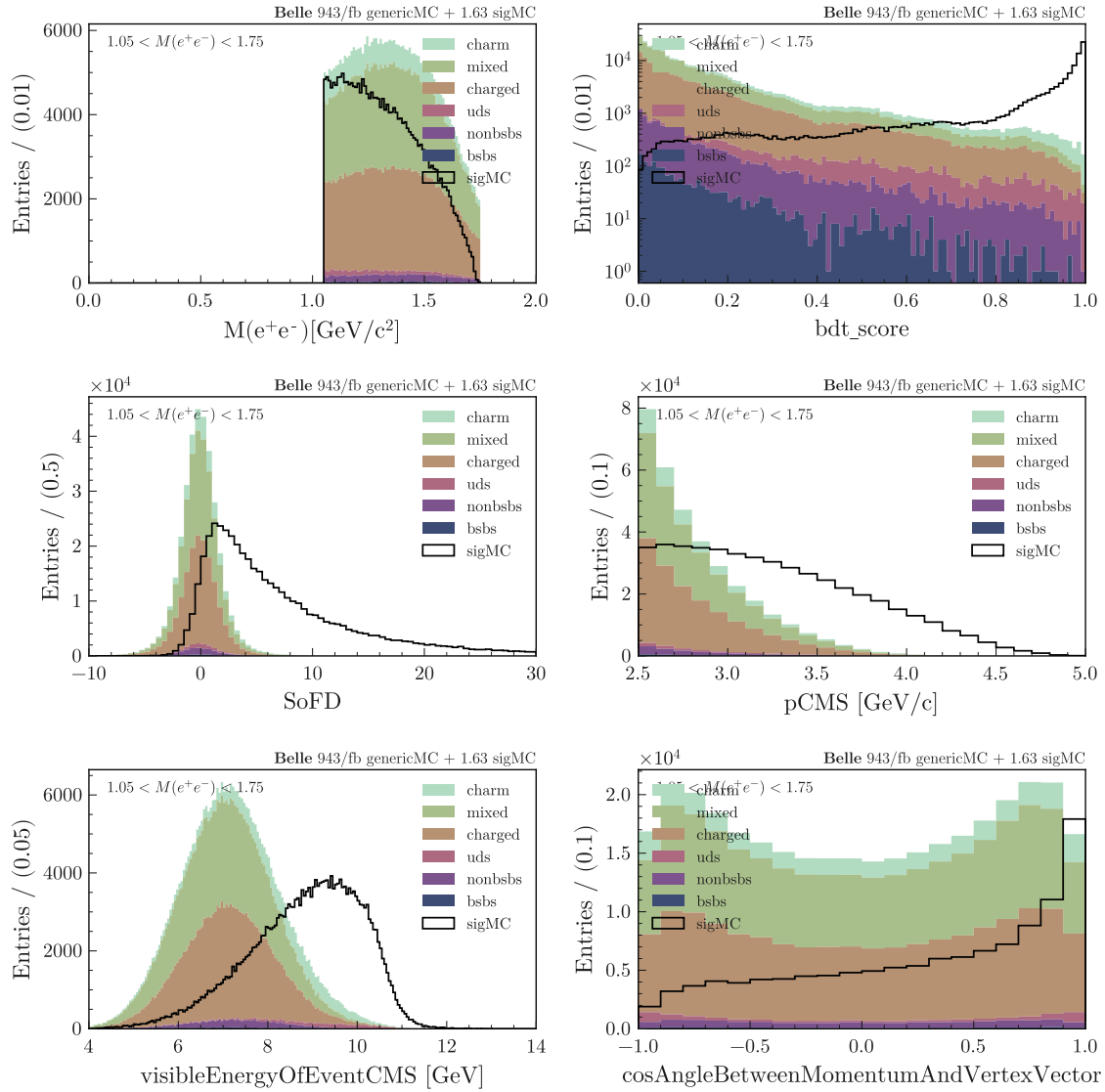


Figure B.5: Distributions of the RFC input variables after the initial selection for Belle, in the interval I_3 ($1.05 \text{ GeV}/c^2 - 1.75 \text{ GeV}/c^2$).

B.2. RFC INPUT VARIABLES IN $M(e^+e^-)$ INTERVALS

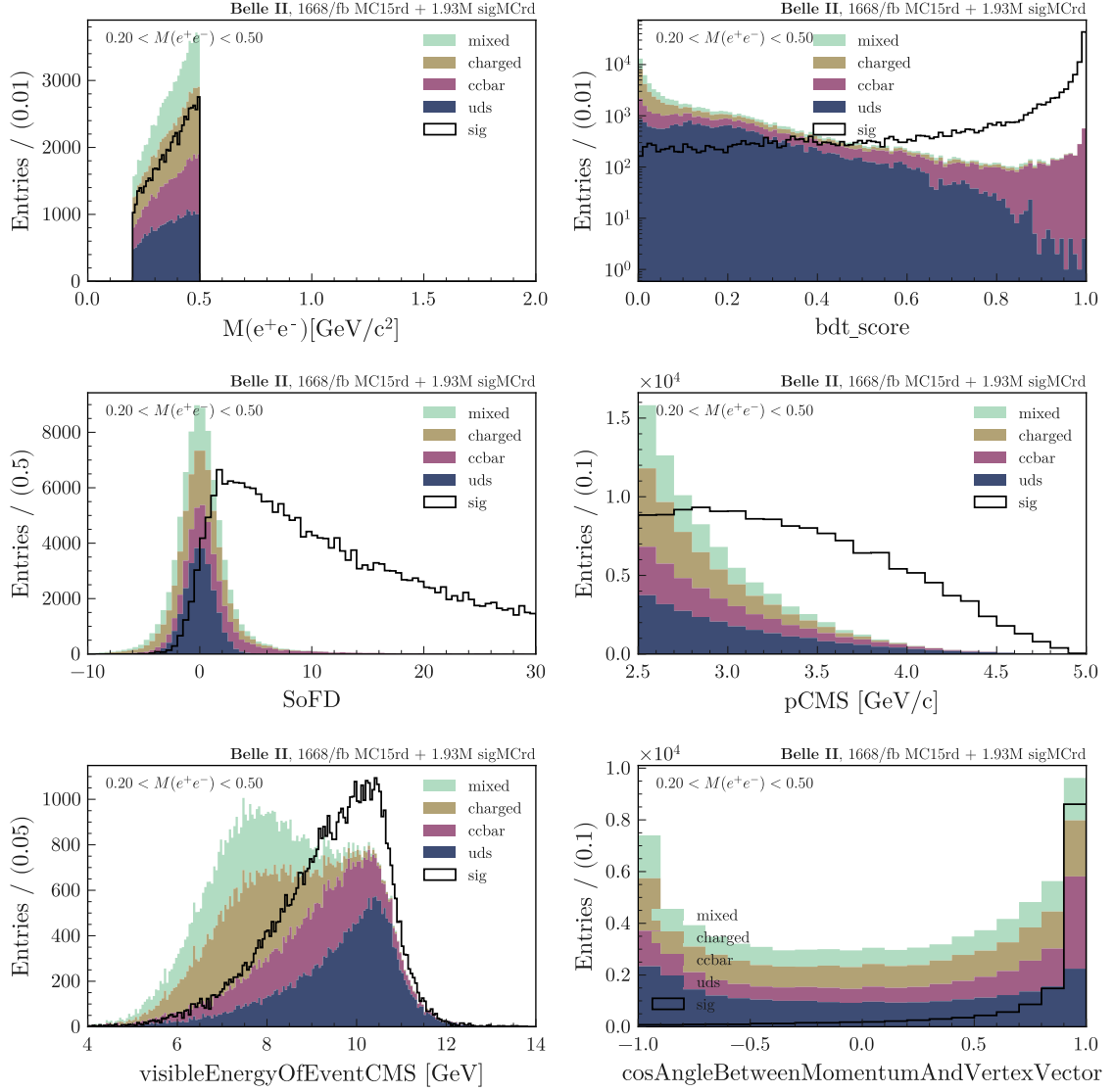


Figure B.6: Distributions of the RFC input variables after the initial selection for Belle II, in the interval I_1 ($0.20 \text{ GeV}/c^2 - 0.50 \text{ GeV}/c^2$).

APPENDIX B. SUPPLEMENTARY SELECTION AND DATA/MC STUDIES

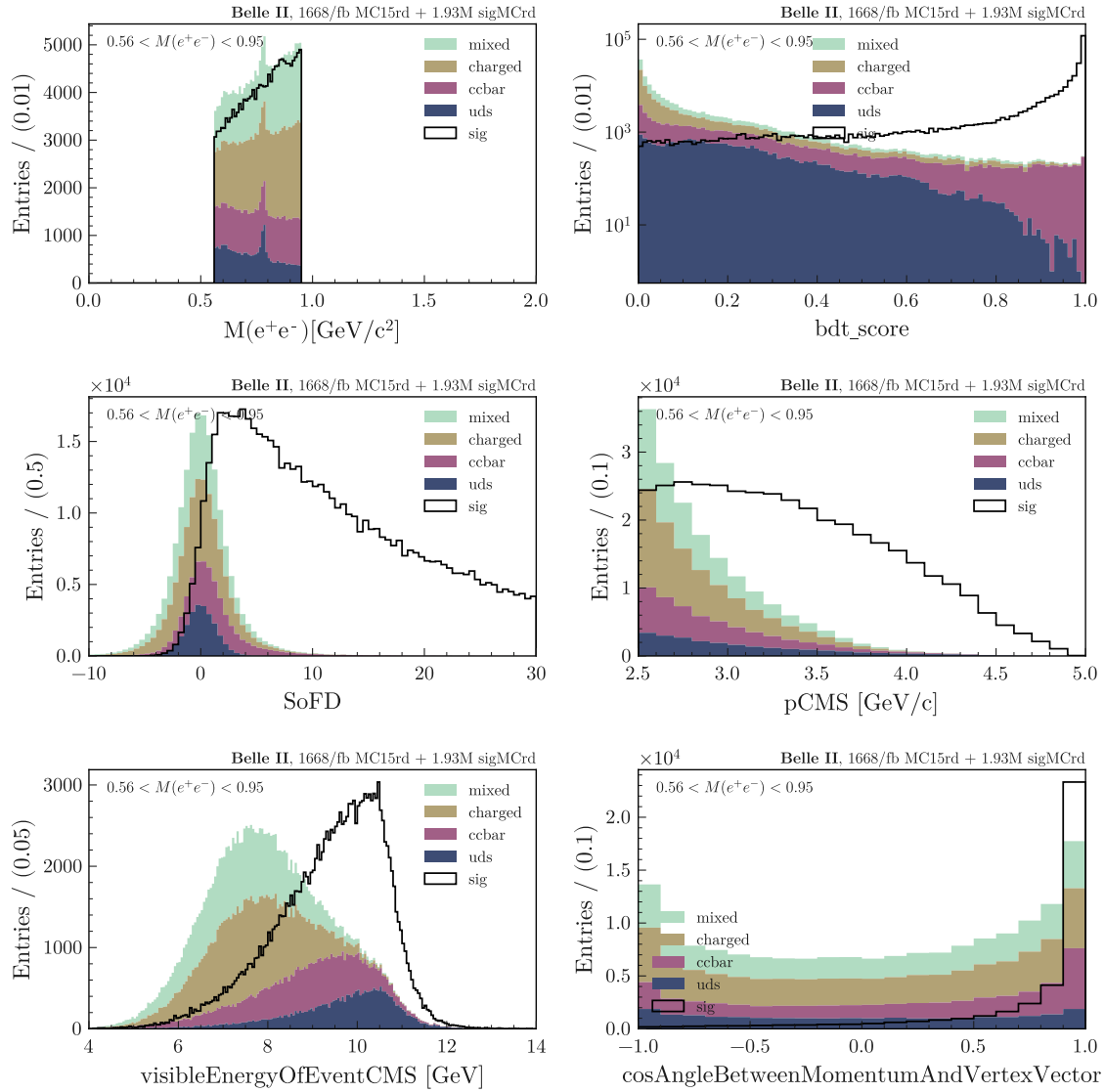


Figure B.7: Distributions of the RFC input variables after the initial selection for Belle II, in the interval I_2 ($0.56 \text{ GeV}/c^2 - 0.95 \text{ GeV}/c^2$).

B.3. RFC PERFORMANCE IN THE VALIDATION AND NORMALISATION MODES

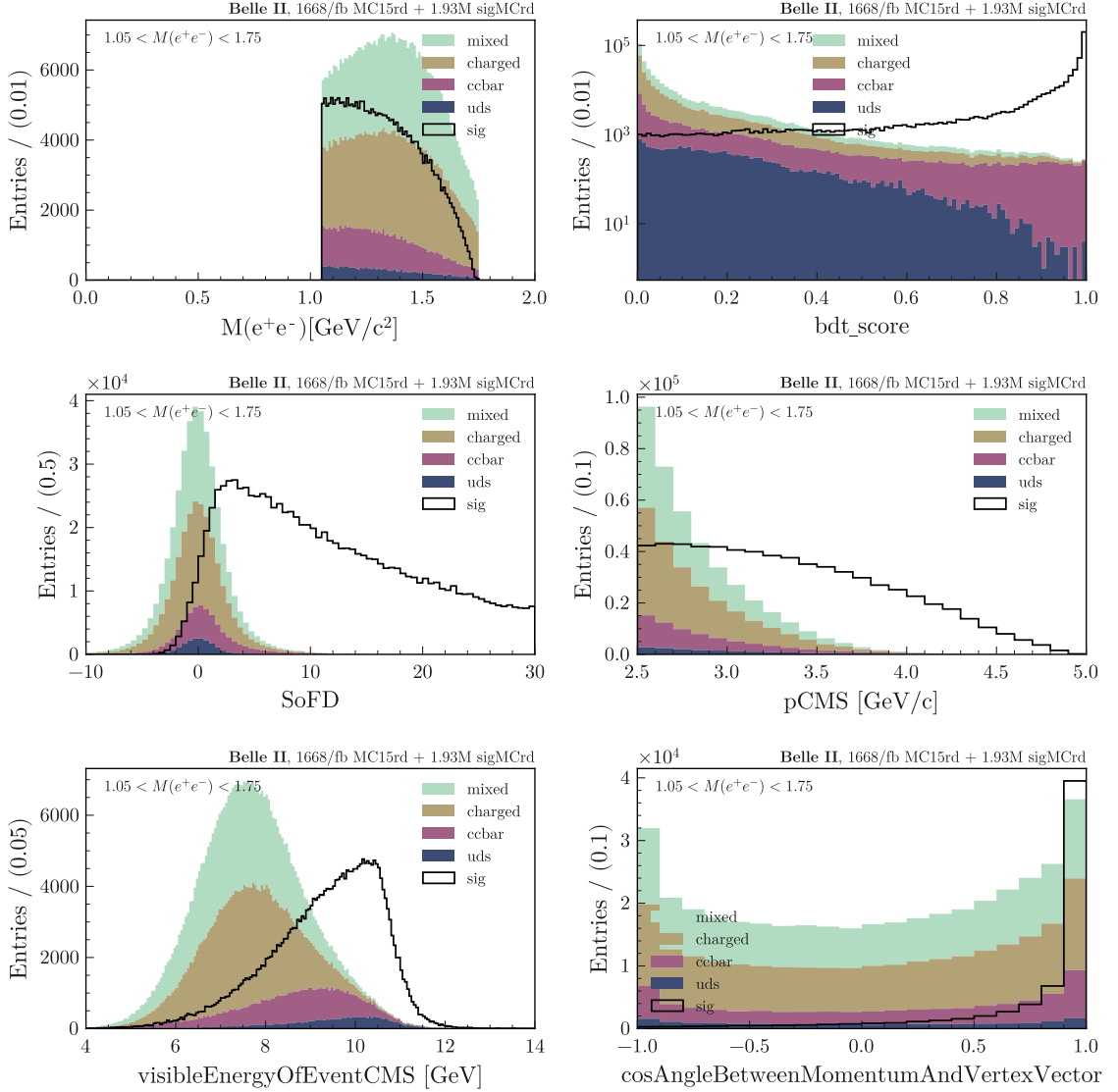


Figure B.8: Distributions of the RFC input variables after the initial selection for Belle II, in the interval I_3 ($1.05 \text{ GeV}/c^2 - 1.75 \text{ GeV}/c^2$).

B.3 RFC Performance in the Validation and Normalisation Modes

The RFC is trained on the signal mode, but the same classifier is applied to the validation and normalisation modes. Figures B.9 and B.10 show RFC performance on the validation and normalisation modes, respectively. Both are evaluated on simulated samples only.

APPENDIX B. SUPPLEMENTARY SELECTION AND DATA/MC STUDIES

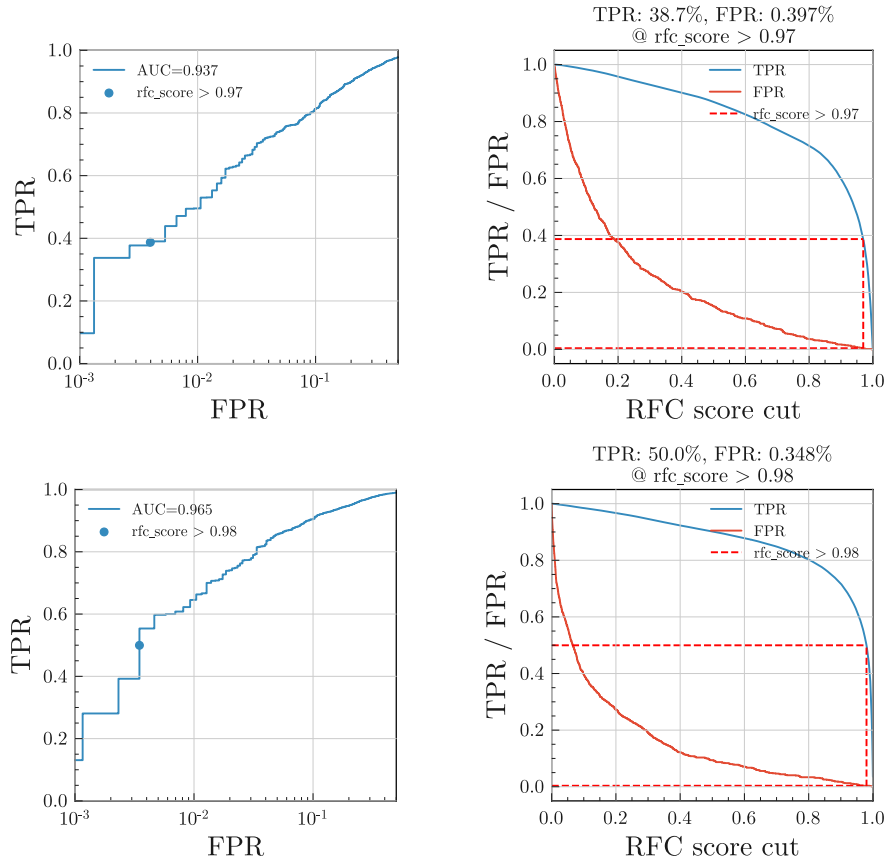


Figure B.9: RFC validation mode performance within a D^+ mass peak $10 \text{ MeV}/c^2$ window: ROC curves and TPR/FPR for Belle and Belle II simulated samples.

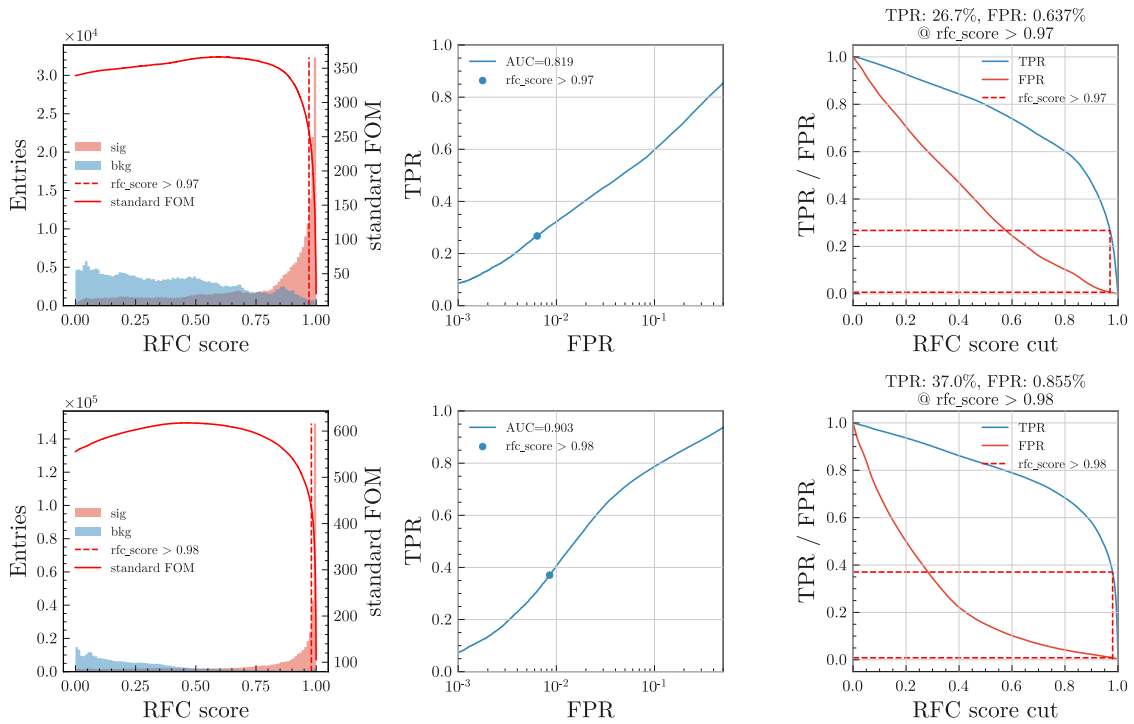


Figure B.10: RFC score distributions, TPR and FPR, Punzi FOM on normalisation mode for Belle and Belle II simulated samples.

B.4 Data/MC Comparisons

B.4.1 Validation Mode

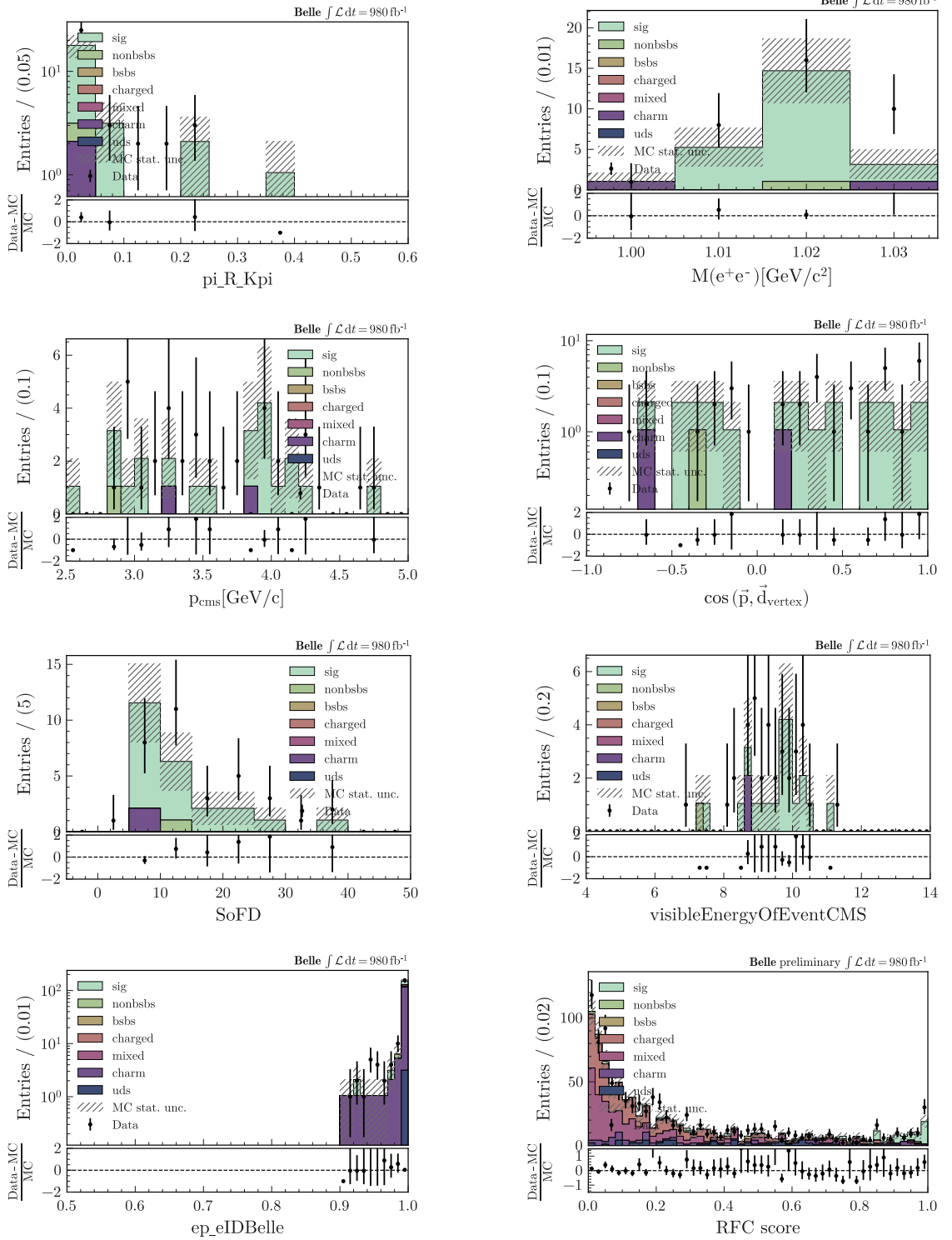


Figure B.11: Data/generic MC comparison of the validation-mode selection variables for Belle, within $|M(\pi^+e^+e^-) - m_{D^+}| < 0.01 \text{ GeV}/c^2$. The RFC score (bottom right) is shown after the initial selection, the remaining variables after the final selection.

APPENDIX B. SUPPLEMENTARY SELECTION AND DATA/MC STUDIES

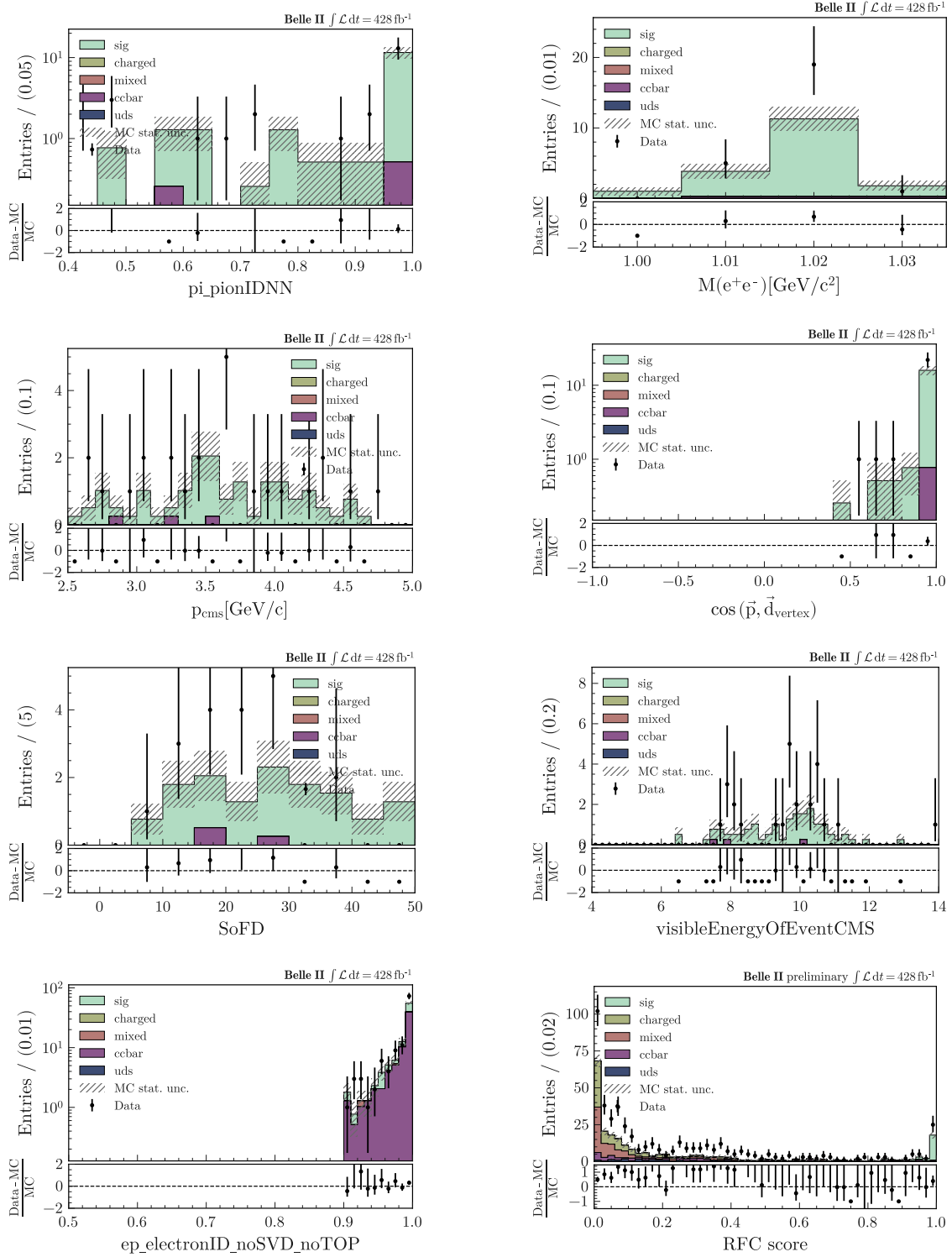


Figure B.12: Data/generic MC comparison of the validation-mode selection variables for Belle II, within $|M(\pi^+e^+e^-) - m_{D^+}| < 0.01 \text{ GeV}/c^2$. The RFC score (bottom right) is shown after the initial selection, and the remaining variables after the final selection.

B.4.2 Normalisation Mode

Figures B.13 and B.14 show the remaining selection variables for Belle and Belle II, respectively.

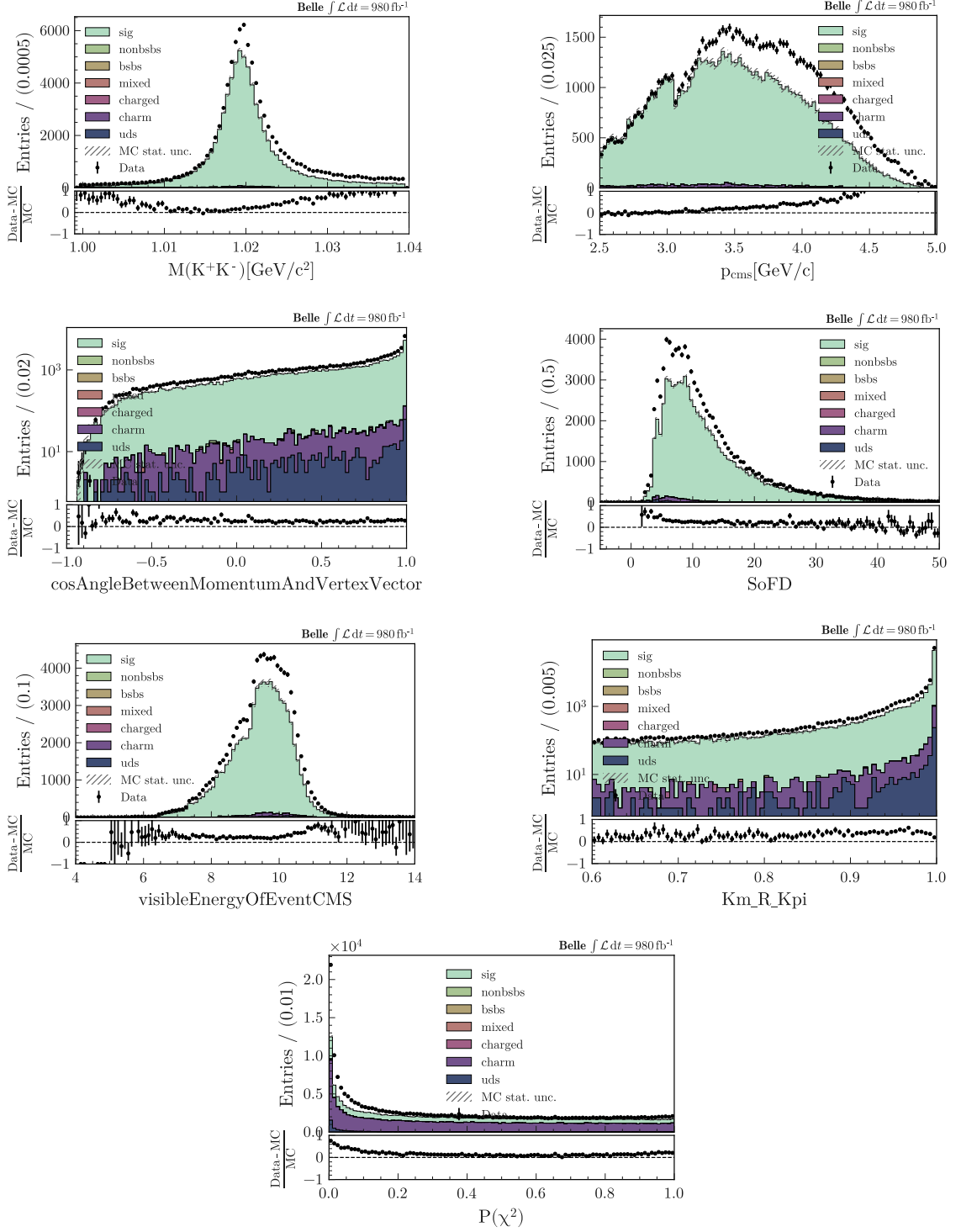


Figure B.13: Data/generic MC comparison of the normalisation-mode selection variables for Belle, after the final selection, for candidates within $|M(\pi^+K^+K^-) - m_{D^+}| < 0.01 \text{ GeV}/c^2$.

APPENDIX B. SUPPLEMENTARY SELECTION AND DATA/MC STUDIES

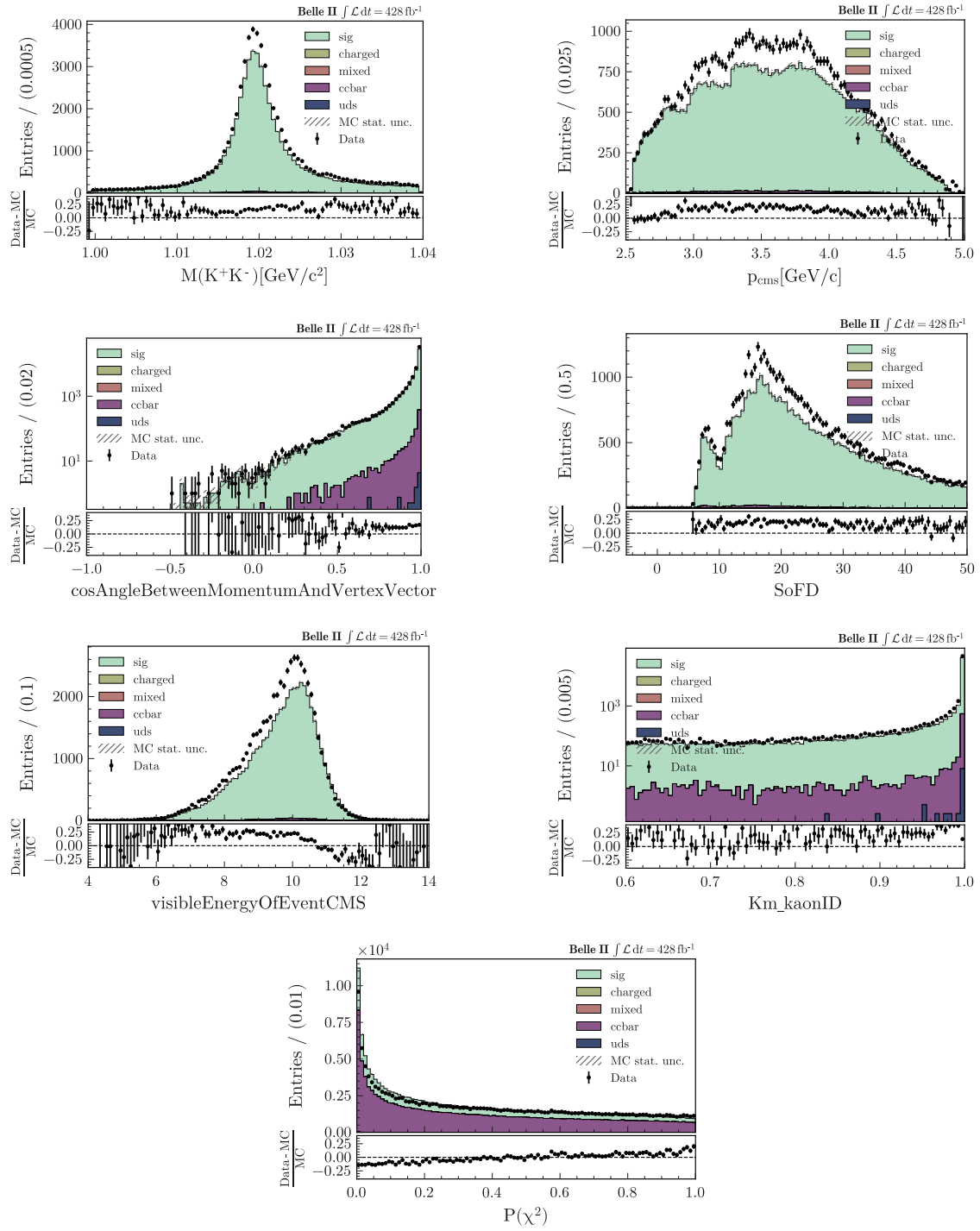


Figure B.14: Data/generic MC comparison of the normalisation-mode selection variables for Belle II, after the final selection, for candidates within $|M(\pi^+ K^+ K^-) - m_{D^+}| < 0.01 \text{ GeV}/c^2$.

B.4.3 Sideband Regions

Figures B.15 and B.16 show the remaining selection variables in the $M(\pi^+e^+e^-)$ sideband regions defined in Section 6.2.3. For Belle, the discrepancy discussed in the main text is concentrated at low values of the flight-distance significance.

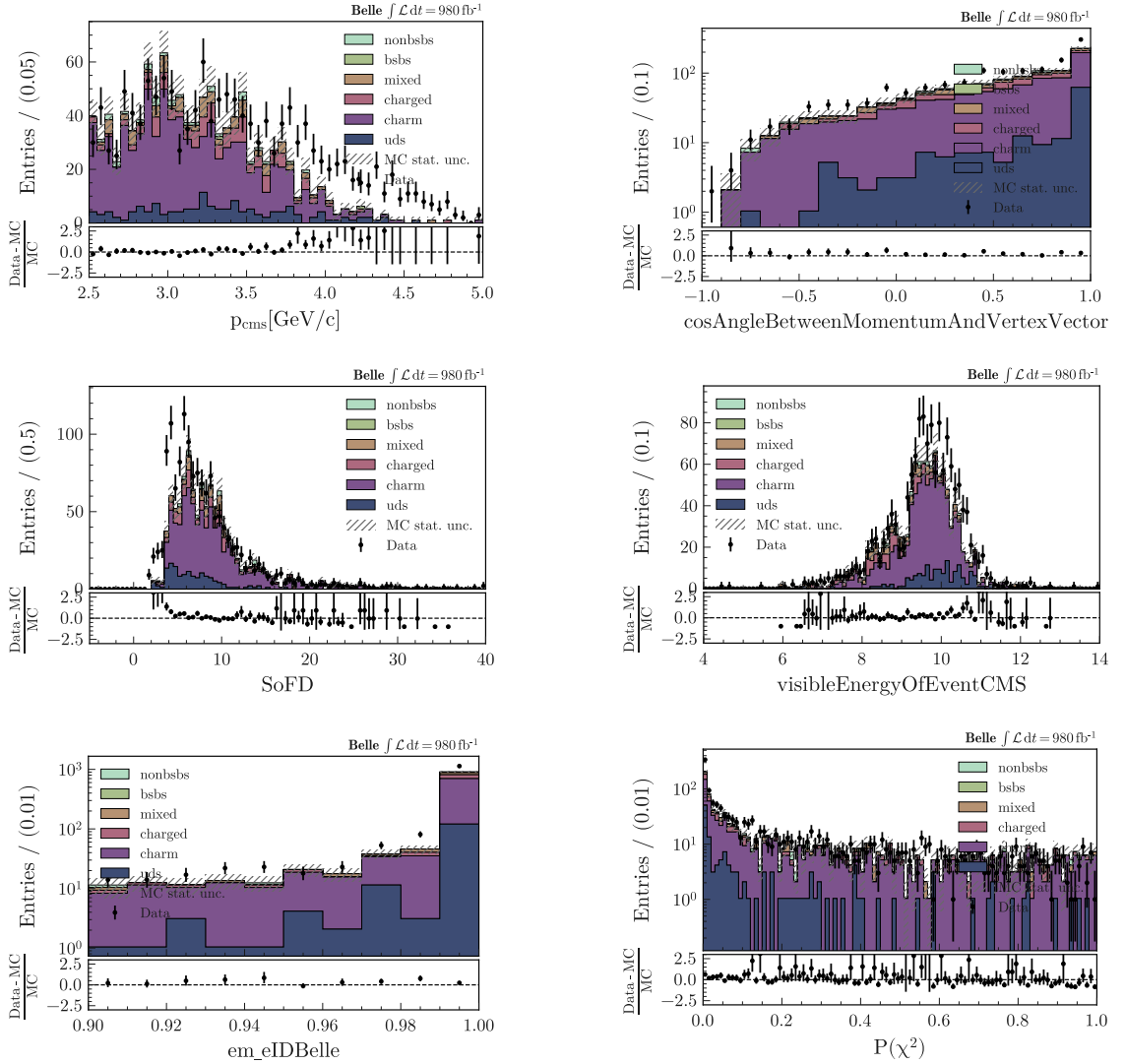


Figure B.15: Data/generic MC comparison of the remaining selection variables in the $M(\pi^+e^+e^-)$ sideband regions for Belle, after the final selection.

APPENDIX B. SUPPLEMENTARY SELECTION AND DATA/MC STUDIES

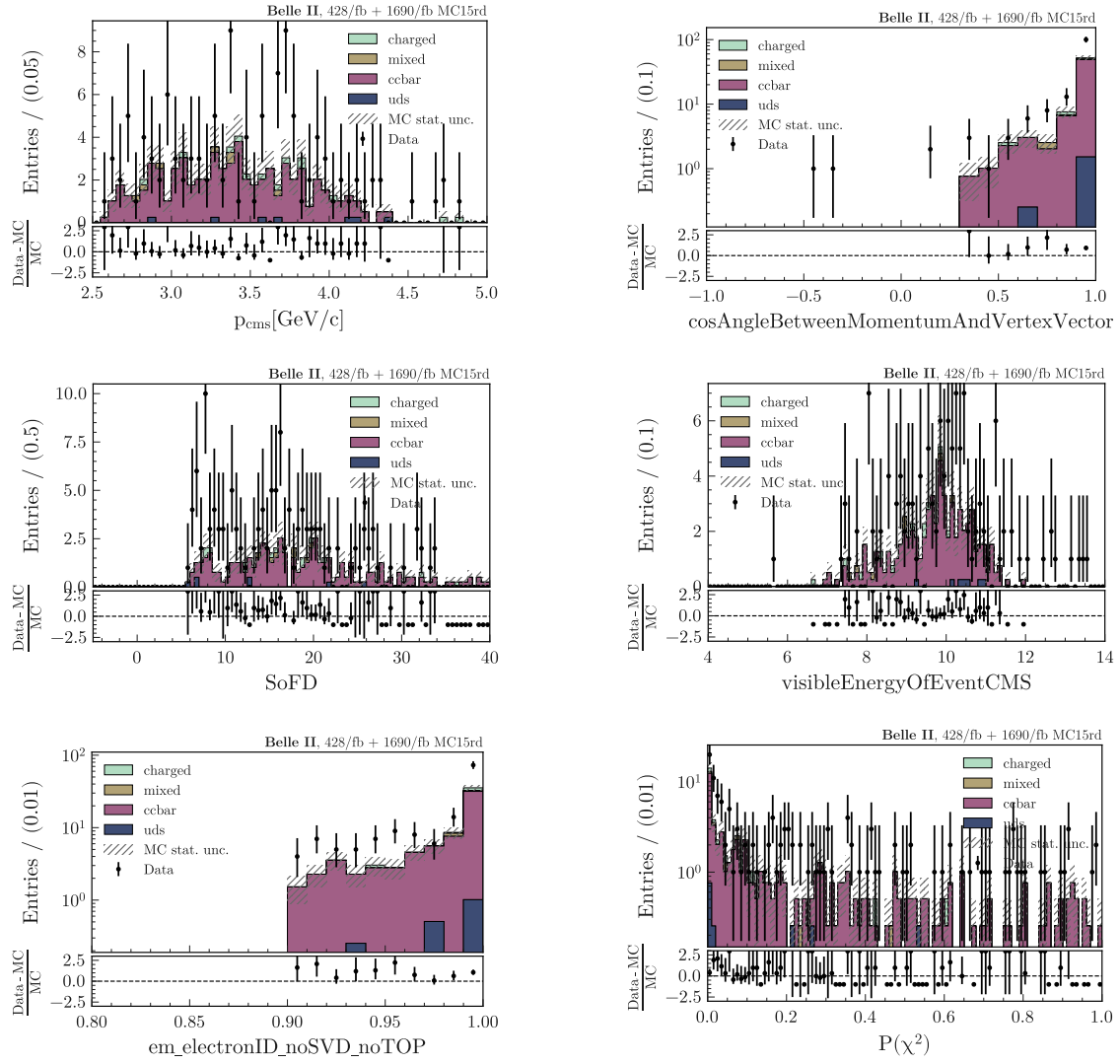


Figure B.16: Data/generic MC comparison of the remaining selection variables in the $M(\pi^+e^+e^-)$ sideband regions for Belle II, after the final selection.

B.4.4 Signal Mode

Figures B.17 and B.18 show the remaining signal-mode selection variables for Belle and Belle II respectively, for candidates near the D^+ mass peak.

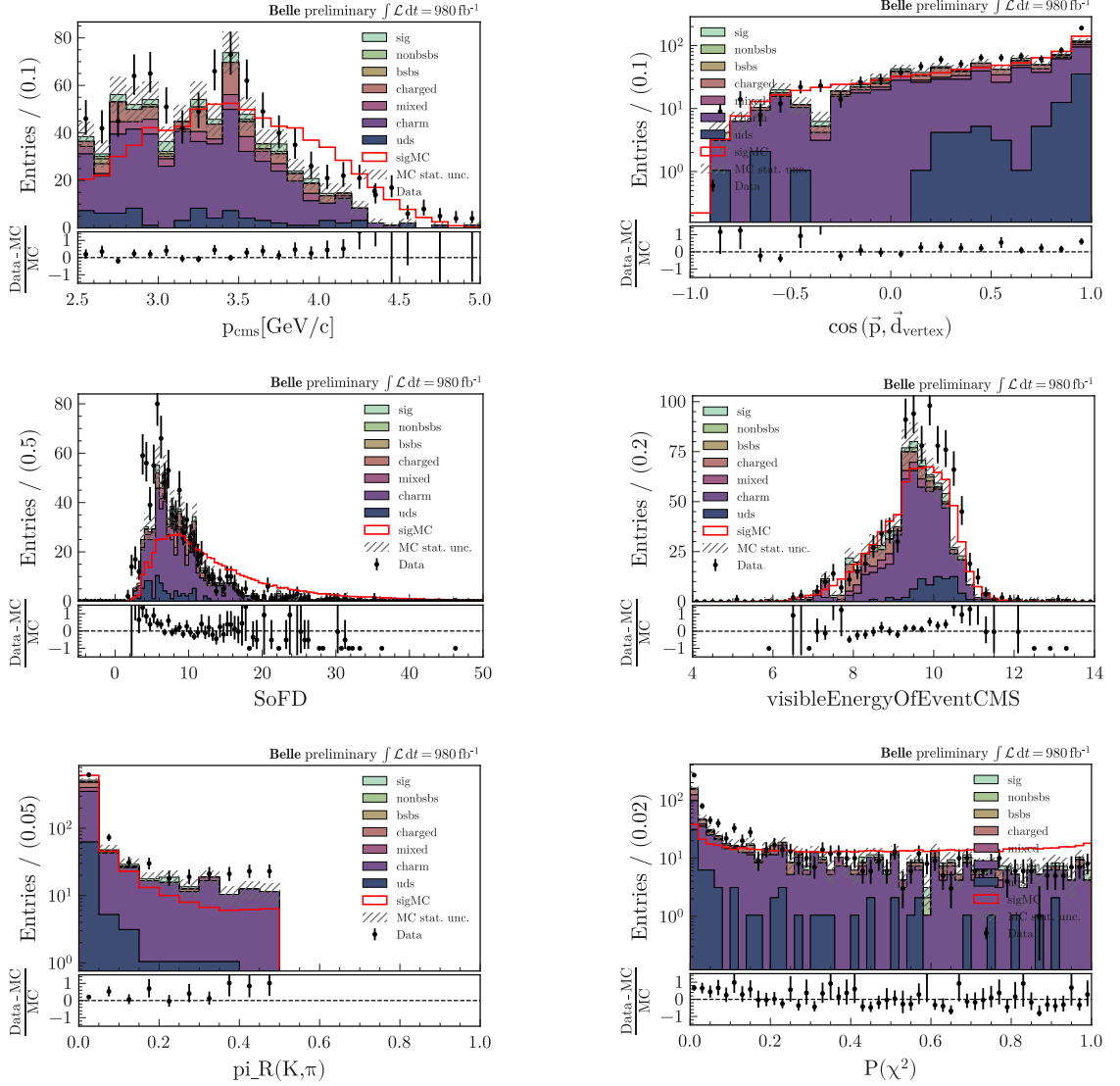


Figure B.17: Data/generic MC comparison of the remaining signal-mode selection variables for Belle, after the final selection, restricted to candidates near the D^+ mass peak. Generic MC is scaled to the data as in Section 6.2.4.

APPENDIX B. SUPPLEMENTARY SELECTION AND DATA/MC STUDIES

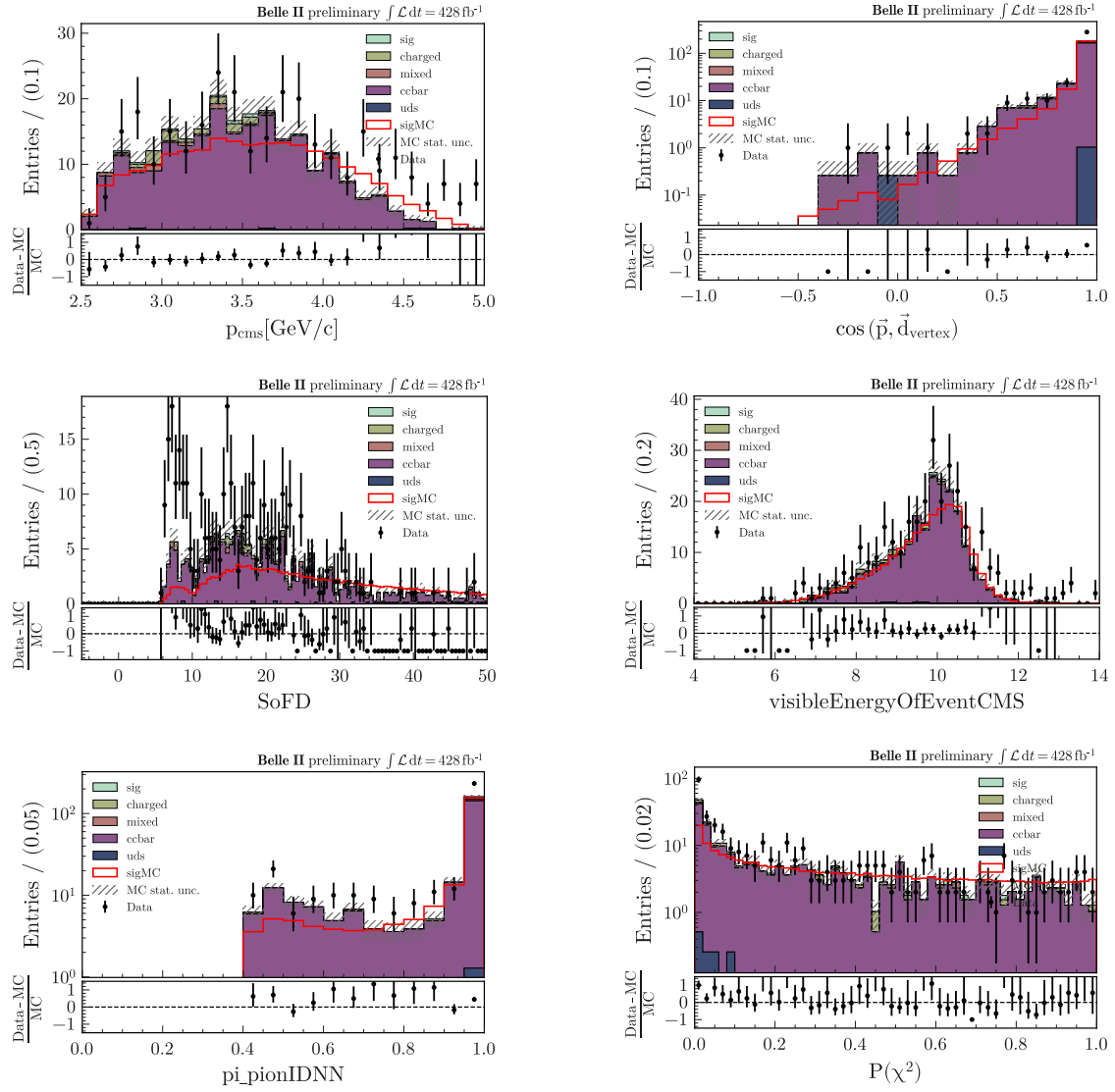


Figure B.18: Data/generic MC comparison of the remaining signal-mode selection variables for Belle II, after the final selection, restricted to candidates near the D^+ mass peak. Generic MC is scaled to the data as in Section 6.2.4.

Appendix C

Supplementary Toy-MC Studies

Chapter 7 presents the toy-MC pull and linearity studies for the simultaneous fit, and for the inclusive $M(e^+e^-)$ range and the interval I_1 . This appendix collects the corresponding studies for the individual Belle and Belle II fits, and for the remaining $M(e^+e^-)$ intervals. All ensembles comprise 10^4 pseudo-experiments.

C.1 Independent Fits: Signal Mode

Figure C.1 shows the pull distributions and Figure C.2 the linearity of the independent signal-mode fits to each sample. The pulls are centred at zero with unit width, and the fitted yields track the generated values over the full range tested.

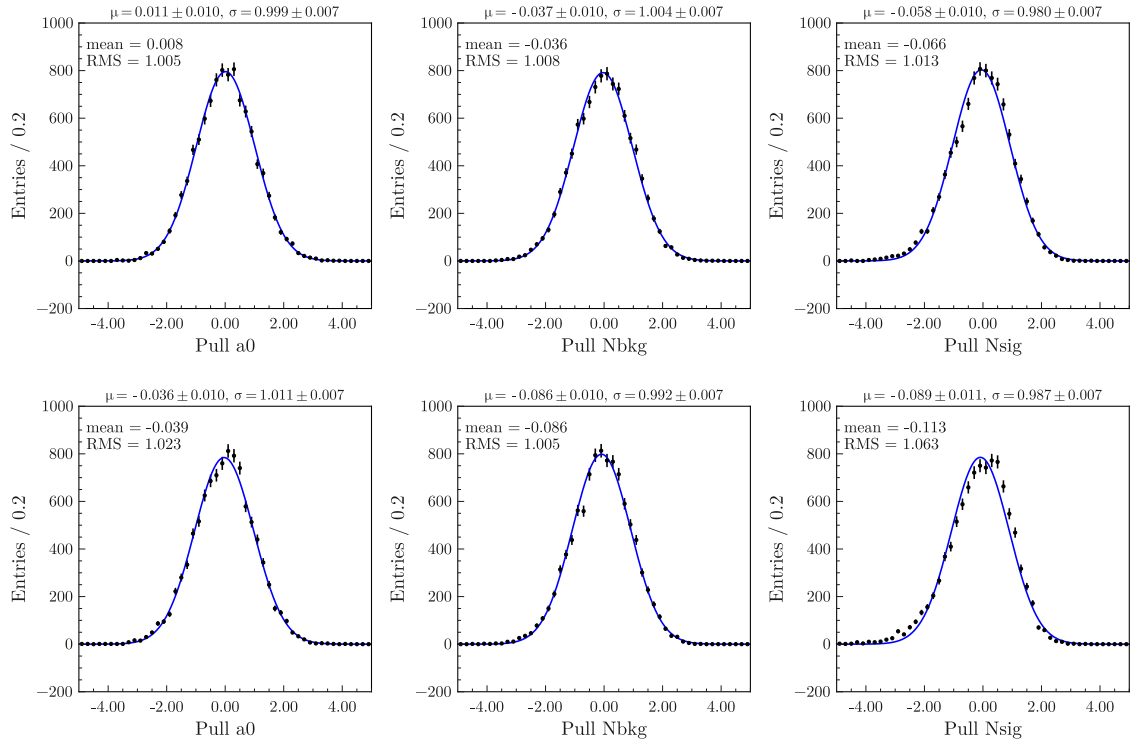


Figure C.1: Pull distributions of the fitted parameters from 10^4 pseudo-experiments, for independent signal-mode fits to the Belle (top row) and Belle II (bottom row) samples, generated with $N_{\text{sig}}^{\text{true}} = 0$.

APPENDIX C. SUPPLEMENTARY TOY-MC STUDIES

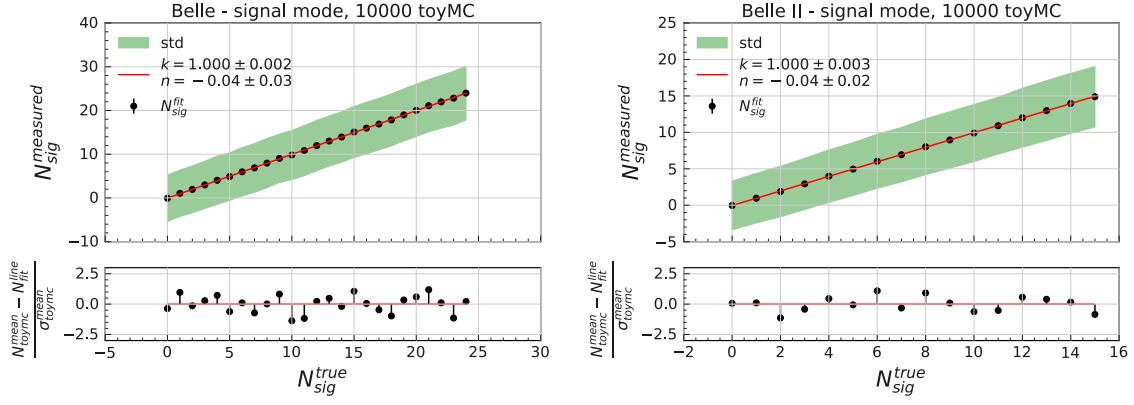


Figure C.2: Linearity of the independent signal-mode fits, for Belle (left) and Belle II (right). Black points and green bands show the mean and standard deviation of the fitted yields over 10^4 pseudo-experiments.

C.2 Simultaneous Fit in $M(e^+e^-)$ Intervals

Figures C.3 and C.4 show the pull distributions for the simultaneous fit in the intervals I_2 and I_3 , and Figure C.5 the corresponding linearity tests. The behaviour matches that observed for I_1 in Chapter 7.

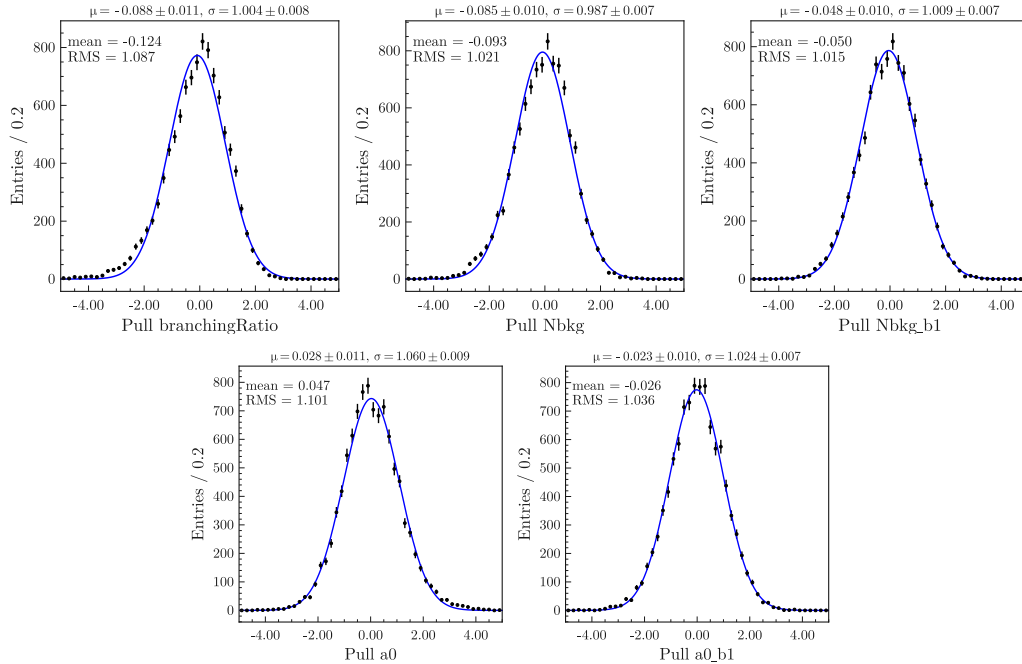


Figure C.3: Pull distributions from 10^4 pseudo-experiments for the simultaneous signal-mode fit in the interval I_2 , generated with $\mathcal{B}_{\pi ee}^{\text{true}} = 0$.

C.2. SIMULTANEOUS FIT IN $M(e^+e^-)$ INTERVALS

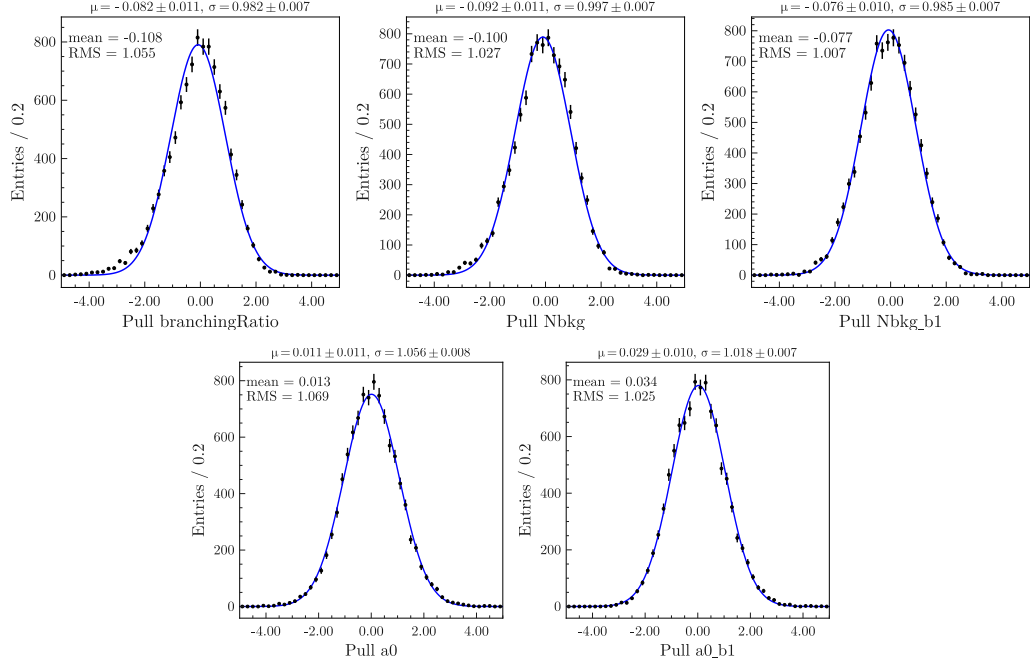


Figure C.4: Pull distributions from 10^4 pseudo-experiments for the simultaneous signal-mode fit in the interval I_3 , generated with $\mathcal{B}_{\pi ee}^{\text{true}} = 0$.

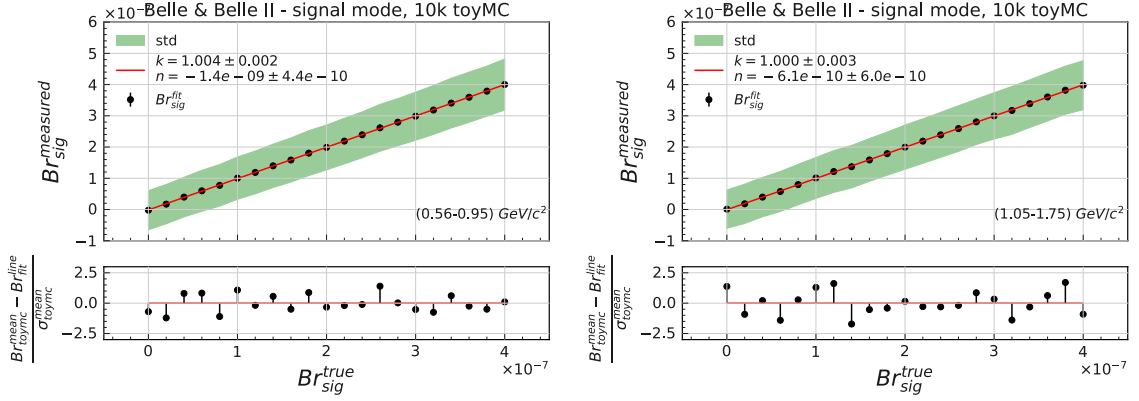


Figure C.5: Linearity of the simultaneous signal-mode fit in the intervals I_2 (left) and I_3 (right). Black points and green bands show the mean and standard deviation of the fitted branching fraction over 10^4 pseudo-experiments.

C.3 Independent Fits: Validation Mode

Figure C.6 shows the pull distributions and Figure C.7 the linearity of the independent validation-mode fits to each sample.

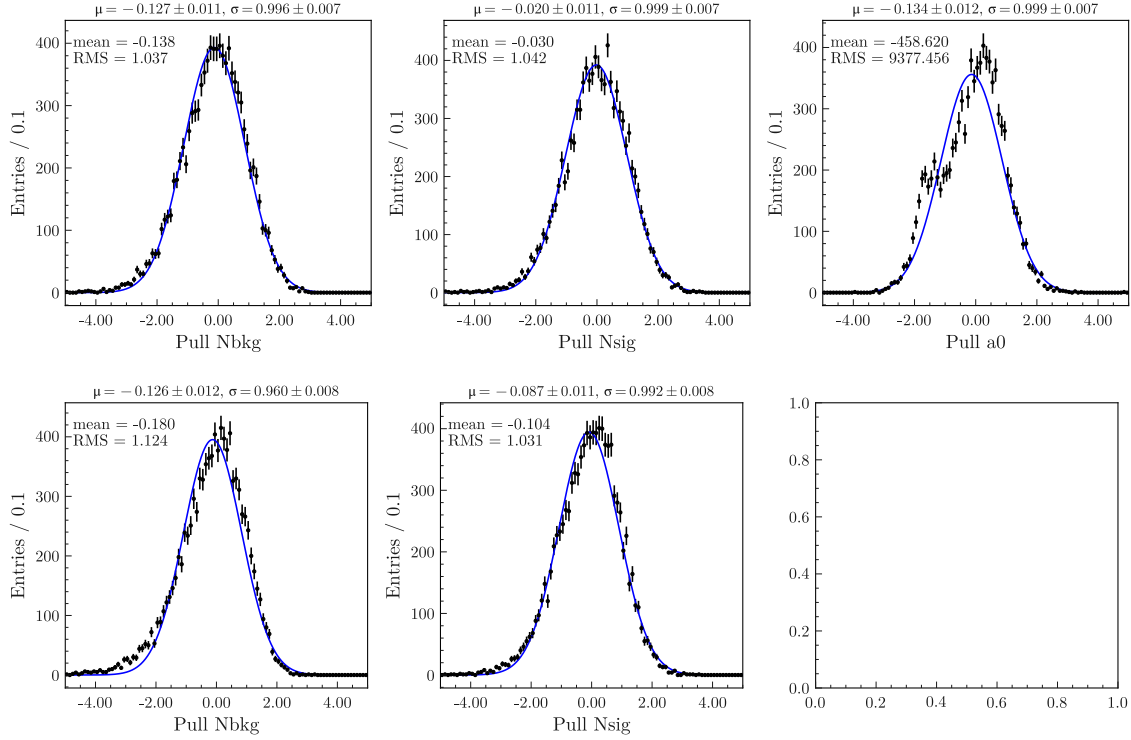


Figure C.6: Pull distributions of the fitted parameters from 10^4 pseudo-experiments, for independent validation-mode fits to the Belle (top row) and Belle II (bottom row) samples.

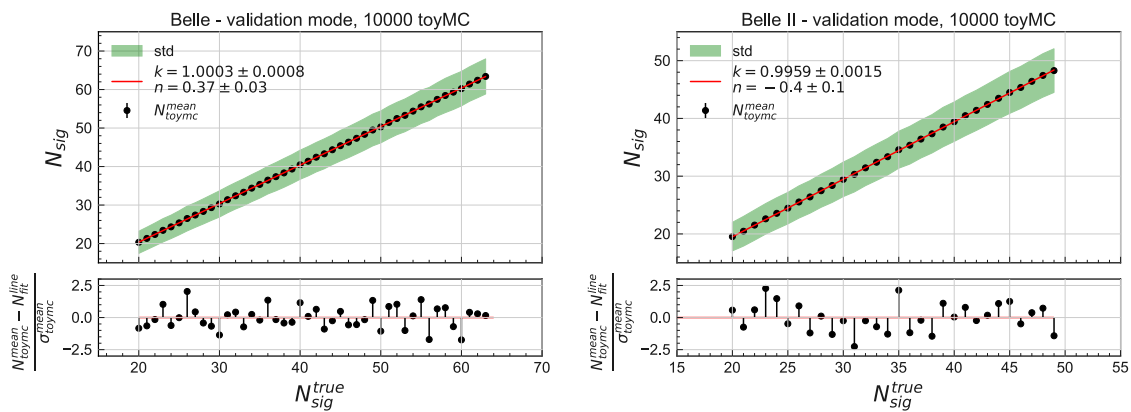


Figure C.7: Linearity of the independent validation-mode fits, for Belle (left) and Belle II (right). Black points and green bands show the mean and standard deviation of the fitted yields over 10^4 pseudo-experiments.

C.4 Normalisation Mode

Figures C.8 and C.9 show the pull distributions for the normalisation fits in Belle and Belle II, complementing the summary given in Tables 7.11 and 7.12.

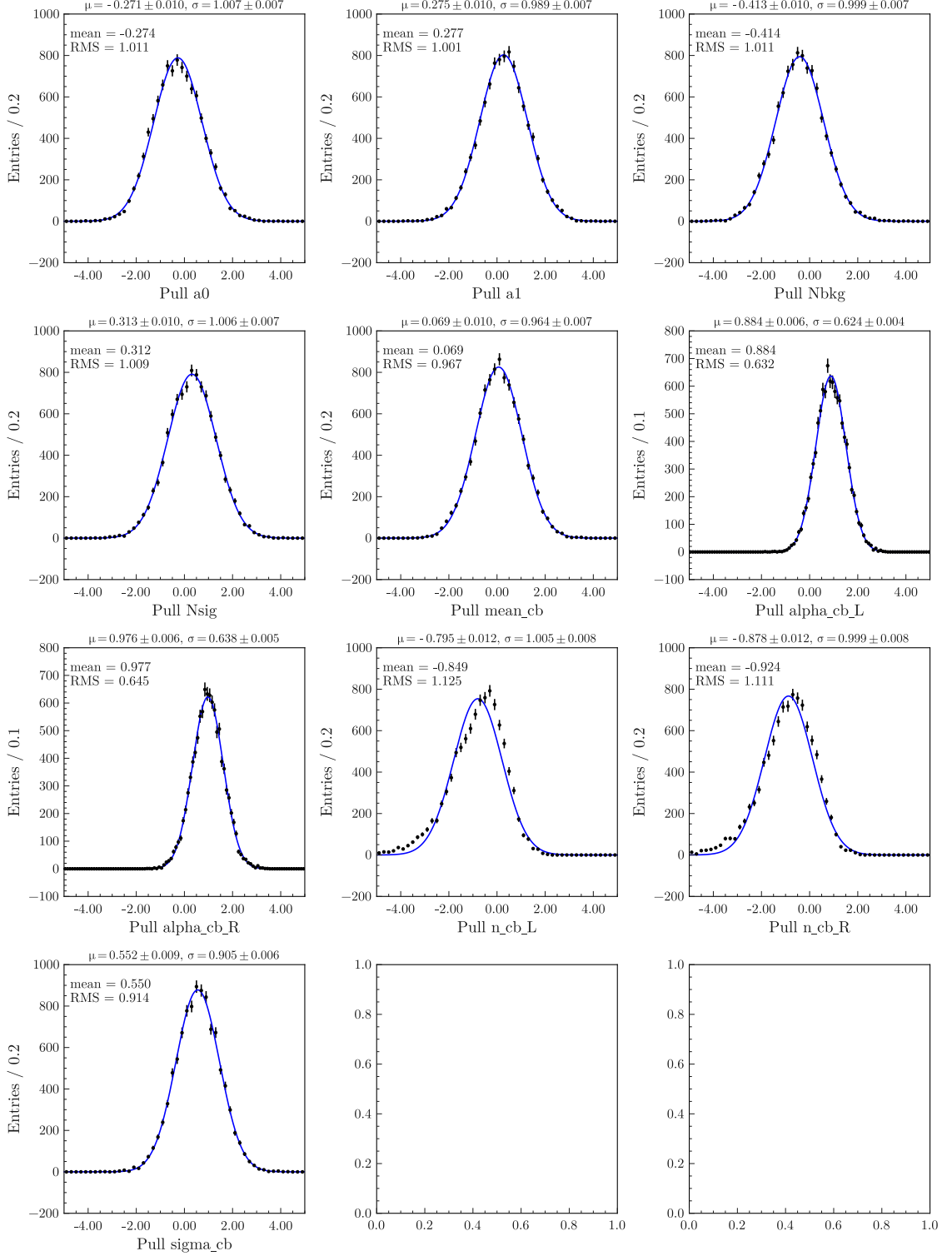


Figure C.8: Pull distributions of the fitted parameters from 10^4 pseudo-experiments for the Belle normalisation-mode fit.

APPENDIX C. SUPPLEMENTARY TOY-MC STUDIES

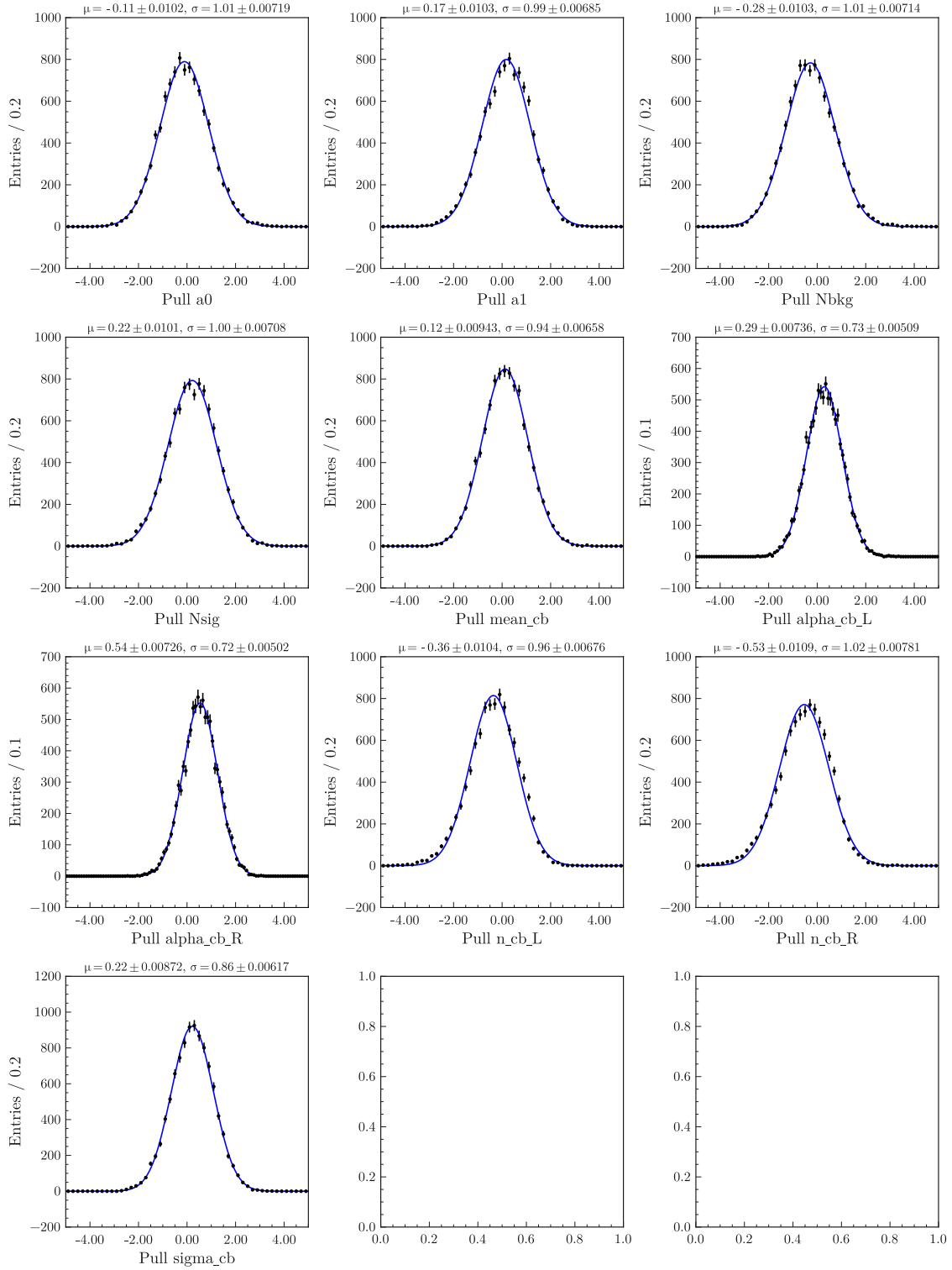


Figure C.9: Pull distributions of the fitted parameters from 10^4 pseudo-experiments for the Belle II normalisation-mode fit.

Appendix D

Supplementary Systematic and Result Studies

D.1 Normalisation-Mode Fits Without the RFC Requirement

The RFC efficiency of Section 8.2.2 is the ratio of the normalisation-mode signal yield after the RFC requirement to the yield before it, the numerator coming from the nominal fits of Chapters 7 and 9. Figures D.1 and D.2 show the fits providing the denominator, performed on generic MC and on data without any RFC requirement.

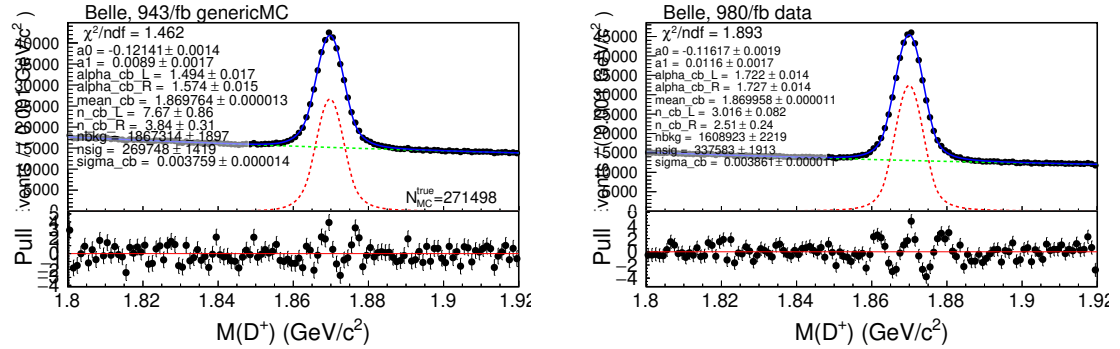


Figure D.1: Fits to the $M(\pi^+K^+K^-)$ distribution of the Belle normalisation mode without the RFC requirement, on generic MC (left) and on data (right).

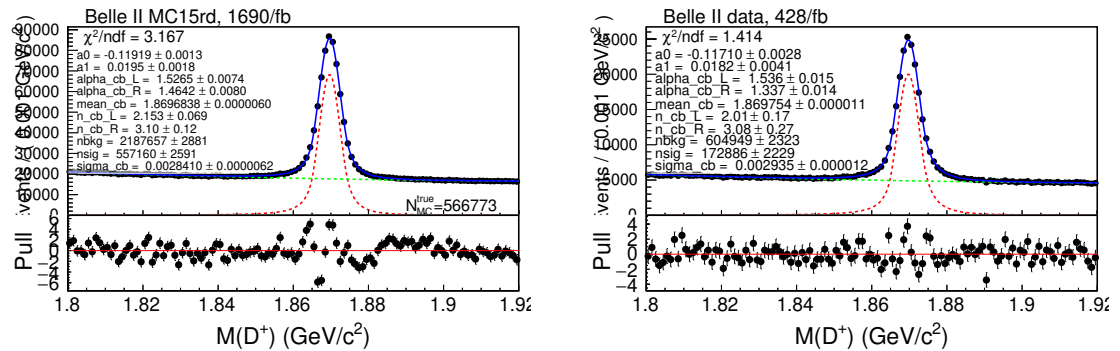


Figure D.2: Fits to the $M(\pi^+K^+K^-)$ distribution of the Belle II normalisation mode without the RFC requirement, on generic MC (left) and on data (right).

D.2 Broken-Signal Component in $M(e^+e^-)$ Intervals

Figure D.3 shows the fits to the mis-reconstructed ($\text{isSignal} \neq 1$) component of the signal-mode signal MC, performed separately in each $M(e^+e^-)$ interval with the procedure of Section 7.4.1. The resulting yields and fractions are summarised in Table 7.2.

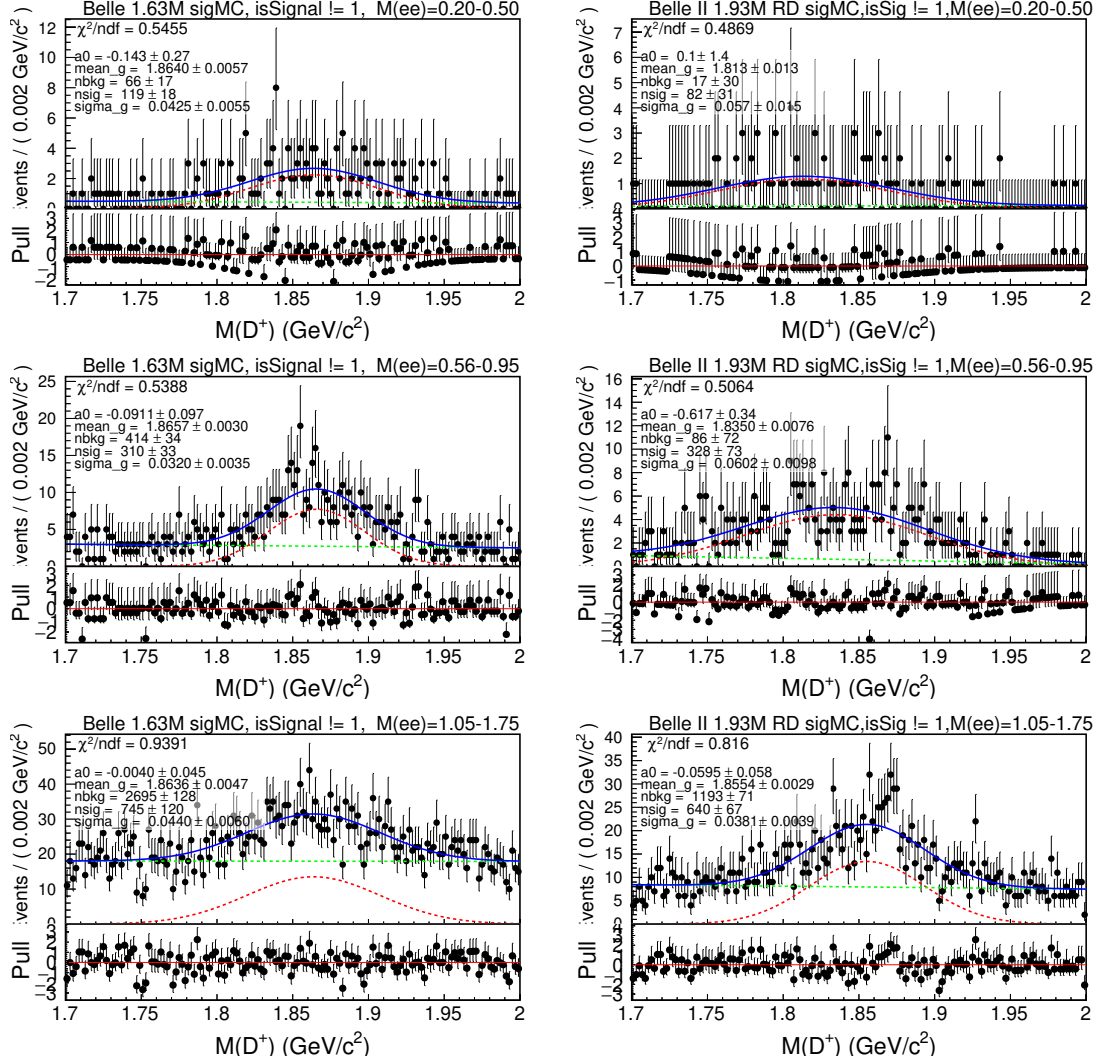


Figure D.3: Fits to the broken-signal ($\text{isSignal} \neq 1$) component of the signal-mode signal MC in the intervals I_1 (top), I_2 (middle) and I_3 (bottom), for Belle (left) and Belle II (right).

D.3 Linearity With Signal-PDF Variations

Figures D.4 and D.5 show the linearity studies of Chapter 7 with the fixed signal-PDF parameters varied within the uncertainties inferred from the normalisation mode. The resulting deviation of the response slope from unity is quantified and assigned as a systematic uncertainty in Section 8.4.3.

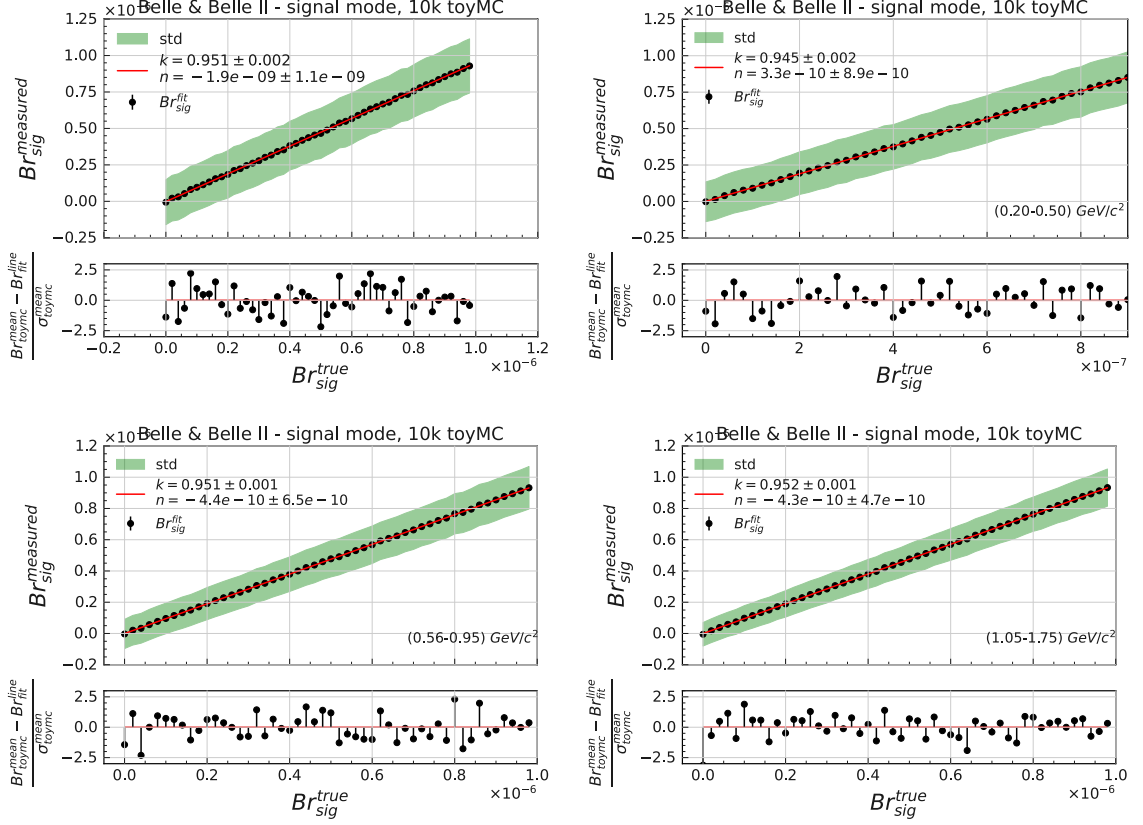


Figure D.4: Linearity of the simultaneous signal-mode fit with varied signal-PDF parameters, for I_{full} (top left), I_1 (top right), I_2 (bottom left) and I_3 (bottom right).

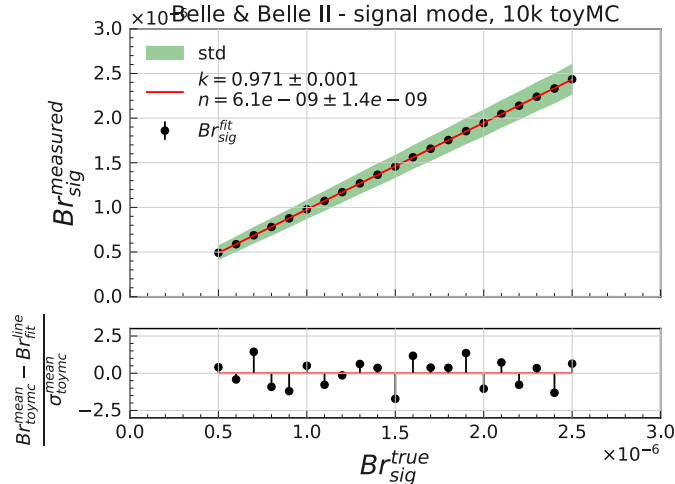


Figure D.5: As in Figure D.4, for the validation mode.

D.4 Background-Model Choice

Figure D.6 shows the background-model variations not reproduced in Section 8.4.2.

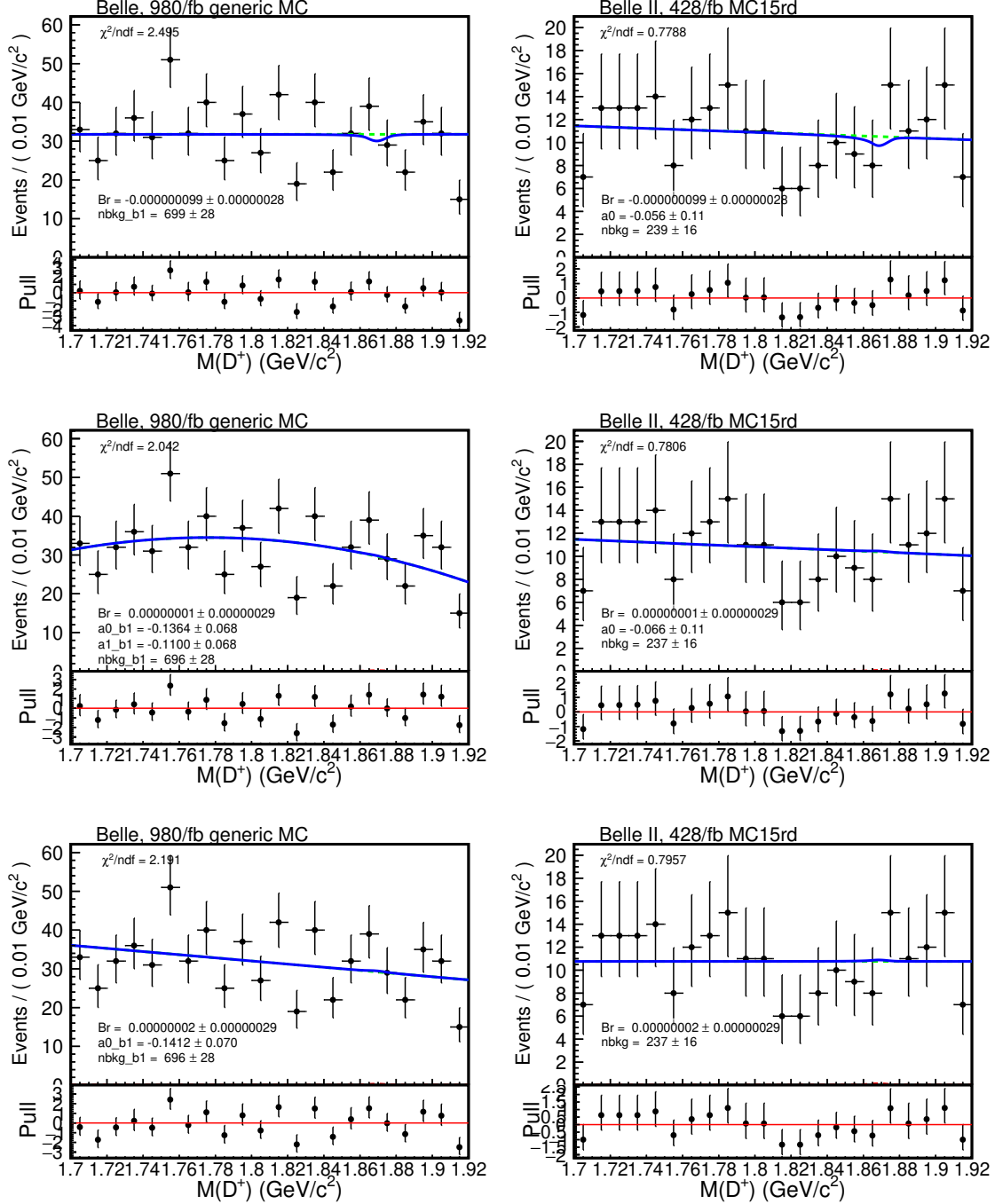


Figure D.6: Remaining background-model variations of the simultaneous fit to the $M(\pi^+e^+e^-)$ distribution of the signal mode in generic MC, shown for Belle (left) and Belle II (right) in each panel. The Belle background is varied to zeroth order (top) and to second order (middle), and the Belle II background to zeroth order (bottom), while in each case the background of the other experiment is kept at its nominal first-order description. The second-order Belle II variation is shown in Figure 8.2.

D.5 Validation Mode: Confidence Belt Without Systematic Uncertainties

Figure D.7 shows the Neyman confidence belt of the simultaneous fit constructed without systematic uncertainties. Together with the nominal belt of Figure 9.3, it gives the separation of the statistical and systematic contributions quoted in Section 9.2.

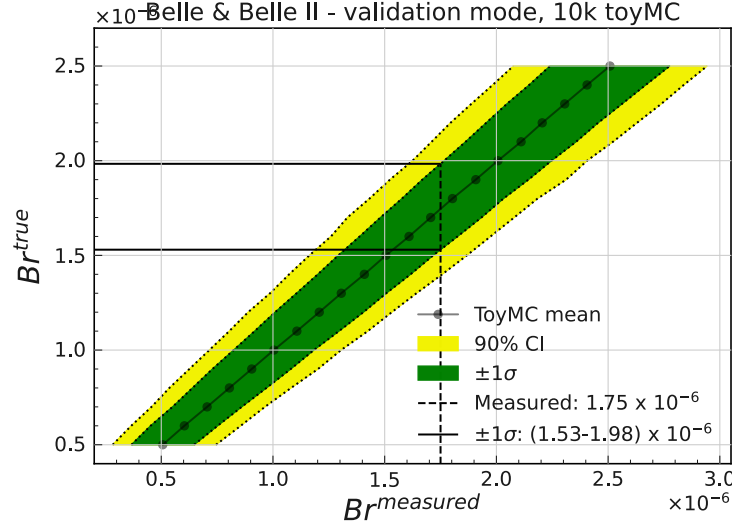


Figure D.7: Neyman confidence belt for the validation mode, constructed from the simultaneous fit without systematic uncertainties.

D.6 Validation Mode: Independent Fits

We fit each data sample independently as a cross-check of the simultaneous result of Section 9.2, with the signal and background yields floating freely rather than a common branching fraction. The resulting fits are shown in Figure D.8.

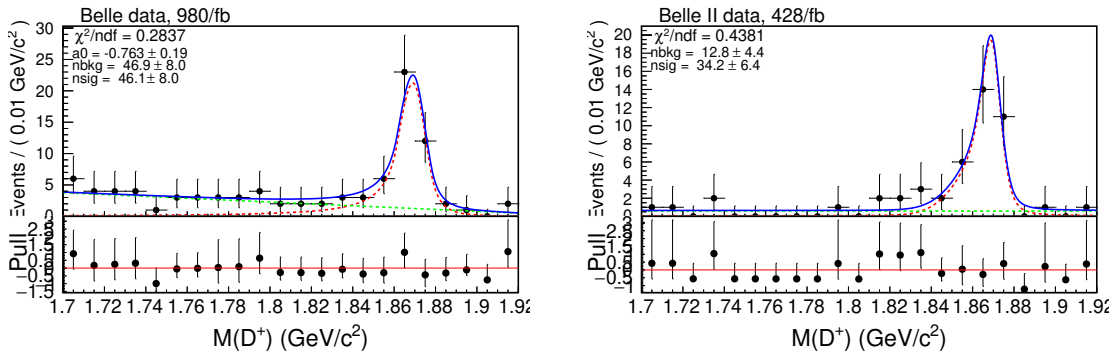


Figure D.8: Independent fits to the $M(\pi^+e^+e^-)$ distribution of the validation mode in the Belle (left) and Belle II (right) data samples.

Razširjeni povzetek v slovenskem jeziku

Uvod in teoretično ozadje

Standardni model [5, 6, 7] razvršča elementarne delce v tri generacije kvarkov in leptonov, interakcije med njimi pa posredujejo umeritveni bozoni. Okus kvarka lahko spremeni le nabiti tok, torej izmenjava bozona $W^\pm$, medtem ko nevtralni tok, ki ga posreduje bozon Z^0 , okusa ne spreminja. Prehodi z nevtralnim tokom, ki spreminja okus, (angl. *flavour-changing neutral current*, FCNC) na drevesnem nivoju zato niso mogoči in lahko potekajo le prek zank, kot so pingvinski in škatlasti diagrami.

Zankasti prispevki pri procesih tipa $c \rightarrow u \ell^+ \ell^-$ so dodatno potlačeni z mehanizmom Glashow–Iliopoulos–Maiani (GIM) [1], po katerem se prispevki posameznih kvarkov v zanki med seboj skoraj izničijo, in z majhnimi elementi matrike Cabibbo–Kobayashi–Maskawa. Pri prehodih kvarkov tipa dol, na primer $b \rightarrow s \ell^+ \ell^-$, potlačitev ni popolna, saj velika masa kvarka t v zanki mehanizem GIM učinkovito poruši. Ti razpadi so zato eksperimentalno bolj dostopni in dobro raziskani. Pri prehodu $c \rightarrow u \ell^+ \ell^-$ pa v zanki nastopajo kvarki tipa dol, katerih mase so majhne in podobne velikosti, zato mehanizem GIM deluje bistveno bolj učinkovito. Prehod je posledično izjemno redek in prav zato občutljiv na prispevke fizike onkraj Standardnega modela, ki niso nujno podvrženi enakemu vzorcu izničenja.

Kratkosežne (angl. *short-distance*, SD) prispevke opišemo v okviru efektivne teorije polja, v kateri težke prostostne stopnje izločimo, njihov učinek pa zajamemo v Wilsonovih koeficientih pred efektivnimi operatorji. V kombinaciji z oblikovnimi faktorji prehoda $D \rightarrow \pi$ dobimo kratkosežno razvejitevno razmerje [25]

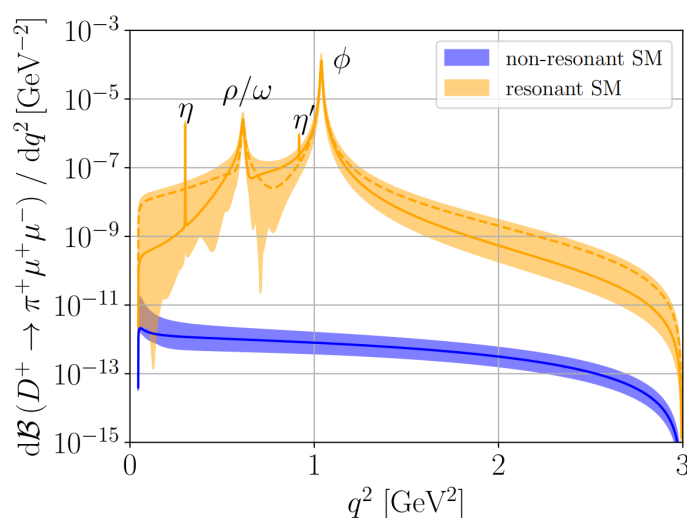
$$\mathcal{B}(D^+ \rightarrow \pi^+ e^+ e^-)^{\text{SD}} = 4.6 \times 10^{-12} \left(1_{-17\%}^{+211\%}\right) \quad (1)$$

kjer največji del negotovosti izvira iz izbire renormalizacijske skale kvarka c . Prispevki onkraj Standardnega modela se v tem opisu kažejo kot spremembe Wilsonovih koeficientov ali kot novi operatorji, ki jih Standardni model ne vsebuje. Ker je kratkosežna napoved tako majhna, meritev deluje kot test ničelne hipoteze: vsaka opažena okrepitev nad dolgosežnim ozadjem bi bila posredno znamenje nove fizike.

Merljivo razvejitevno razmerje je kljub temu za nekaj velikostnih redov večje, saj v njem prevladujejo dolgosežni (angl. *long-distance*, LD) prispevki, pri katerih elektronski par ne nastane neposredno iz prehoda $c \rightarrow u$, temveč iz vmesne vektorske resonance $V \in \{\rho, \omega, \phi\}$ v razpadu $D^+ \rightarrow \pi^+ V (\rightarrow \ell^+ \ell^-)$. Integrirano po celotnem faznem prostoru je razvejitevno razmerje reda 10^{-6} [2, 22, 23, 31, 36] in je skoraj v celoti posledica ozkega pola ϕ : sam resonančni razpad prispeva $\mathcal{B}(D^+ \rightarrow \pi^+ \phi) \mathcal{B}(\phi \rightarrow e^+ e^-) = (1,70 \pm 0,05) \times 10^{-6}$ [16]. Stran od polov je napovedano razmerje precej manjše, reda 10^{-9} [21, 36].

Slika 1 prikazuje diferencialno razvejitevno razmerje v odvisnosti od invariantne mase leptonskega para. Dolgosežni prispevek prevladuje v celotnem kinematičnem območju in kratkosežnega za nekaj velikostnih redov presega tudi daleč od polov, tako da ni območja $M(e^+e^-)$, kjer ne bi bilo dolgosežnih prispevkov. Razlika med obema prispevkoma pa je najmanjša stran od resonanc, zato bi bila tam morebitna okrepitev razvejitvenega razmerja zaradi fizike onkraj Standardnega modela najlažje opazna.

Iz tega izhaja osrednja zamisel te meritve, da razvejitvenega razmerja ne določimo le za celotno območje $M(e^+e^-)$, temveč tudi ločeno v treh intervalih. Mejo, podano za posamezno območje invariantne mase, je mogoče ob drugačni predpostavki o spektru signala lažje ponovno interpretirati kot eno samo mejo čez celotno območje. Več neodvisnih teoretičnih analiz prehoda $c \rightarrow u \ell^+ \ell^-$ [28, 36, 37] izpostavlja prav dielektronske kanale kot vhodni podatek, ki je v njihovih globalnih prilagajanjih pomanjkljiv. Za ta razpad meritev v intervalih $M(e^+e^-)$ doslej ni bilo.



Slika 1: Diferencialno razvejitevno razmerje napovedi Standardnega modela za razpad $D^+ \rightarrow \pi^+ \mu^+ \mu^-$ [35]. Rumena krivulja prikazuje dolgosežni, modra pa kratkosežni prispevek. Širine pasov ustrezajo teoretskim negotovostim, ki izvirajo iz hadronskih oblikovnih faktorjev, parametrov resonanc in renormalizacijske skale kvarka c . Porazdelitev za elektronski kanal je kvalitativno enaka, le da ima nižji kinematični prag.

Eksperimentalno stanje in strategija meritve

Razpad $D^+ \rightarrow \pi^+ \ell^+ \ell^-$ iščejo že desetletja [42], doslej pa ga še ni uspelo zaznati. Najboljšo zgornjo mejo razvejitvenega razmerja pri 90 % stopnji zaupanja je v elektronskem kanalu postavila kolaboracija BABAR, $\mathcal{B}(D^+ \rightarrow \pi^+ e^+ e^-) < 1,1 \times 10^{-6}$ [4], v mionskem kanalu pa kolaboracija LHCb, $\mathcal{B}(D^+ \rightarrow \pi^+ \mu^+ \mu^-) < 6,7 \times 10^{-8}$ [3]. V tej analizi se zato osredotočimo na elektronski kanal, kjer je prostora za izboljšavo več in kjer je trkalnik e^+e^- v prednosti, saj zavorno sevanje elektronov tu bistveno manj poslabša ločljivost invariantne mase in manj poveča ozadje kot na hadronskem trkalniku [51]. Uporabimo 980 fb^{-1} podatkov eksperimenta Belle, ki za to iskanje še niso bili uporabljeni, in 428 fb^{-1} podatkov eksperimenta Belle II. Z zahtevo po višji

težiščni gibalni količini izberemo mezone D^+ , ki nastanejo neposredno v procesu kontinuuma in ne zajamemo tistih, ki so razpadni produkti mezona B .

Razvejitveno razmerje izmerimo relativno glede na normalizacijski kanal $D^+ \rightarrow \pi^+ K^+ K^-$, rekonstruiran v ozkem območju resonance ϕ v porazdelitvi $M(K^+ K^-)$. Izbiramo ga enako kot signalni kanal, le da elektronski par nadomestimo s kaonskim, vetov, značilnih za elektrone, pa pri njem ne uporabimo. Razvejitveno razmerje izračunamo kot

$$\mathcal{B}_{\pi ee}^j = \frac{N_{\pi ee}^j}{N_{\pi KK}} \frac{\epsilon_{\pi KK}}{\epsilon_{\pi ee}^j} \mathcal{B}_{\pi KK}, \quad (2)$$

kjer sta $N_{\pi ee}^j$ in $\epsilon_{\pi ee}^j$ število signalnih dogodkov in izkoristek rekonstrukcije v intervalu j , $N_{\pi KK}$ in $\epsilon_{\pi KK}$ analogni količini normalizacijskega kanala, $\mathcal{B}_{\pi KK} = (9,68 \pm 0,18) \times 10^{-3}$ [16] pa njegovo razvejitveno razmerje, skupno obema eksperimentoma. Razmerje omogoči, da nam ni treba poznati integrirane luminoznosti in sipalnega preseka za produkcijo mezonov D^+ , poleg tega pa se nekatere sistematične negotovosti deloma pokrajšajo, saj sta si razpada podobna.

Število signalnih dogodkov določimo s sočasnim prilagajanjem porazdelitve invariantne mase $M(\pi^+ e^+ e^-)$ v obeh podatkovnih vzorcih z metodo največjega verjetja brez razvrščanja v razrede (angl. *unbinned maximum-likelihood*), pri čemer je razvejitveno razmerje skupni prosti parameter. Tako hkrati izkoristimo celoten nabor podatkov, namesto da bi naknadno združevali dva ločena rezultata. Prilagajanje opravimo na celotnem dostopnem območju $M(e^+ e^-)$ brez območij resonanc η in ϕ ter ločeno v treh intervalih $M(e^+ e^-)$.

Celoten postopek preverimo na kanalu $D^+ \rightarrow \pi^+ \phi (\rightarrow e^+ e^-)$, rekonstruiranem v ozkem oknu okoli resonance ϕ v porazdelitvi $M(e^+ e^-)$. Ta ima enako končno stanje in dovolj veliko razvejitveno razmerje, da ga lahko izmerimo.

Analizo smo razvili na simulacijah Monte Carlo in na podatkih zunaj signalnega območja, šele nato smo pogledali vanj. Postopka po tem nismo več spreminjali in tako zmanjšali možnost eksperimentalne pristranskosti.

Eksperimenta Belle in Belle II

V analizi smo uporabili celoten nabor podatkov, zajetih s spektrometrom Belle II do leta 2022 (Run 1) na trkalniku elektronov in pozitronov SuperKEKB na Japonskem. Dodali smo še celoten vzorec predhodnega spektrometra Belle, ki je na istem mestu deloval na trkalniku KEKB. Vzorca ustrezata integriranima luminoznostma 980 fb^{-1} (Belle) in 428 fb^{-1} (Belle II).

Trkalnik SuperKEKB

Asimetrični trkalnik SuperKEKB [52] trka elektrone z energijo 7 GeV in positrone z energijo 4 GeV. Težiščna energija trka $\sqrt{s} = 10,58 \text{ GeV}$ ustreza masi resonance $\Upsilon(4S)$, ki je vezano stanje kvarka b in antikvarka $\bar{b}$ ter razpade skoraj izključno v par mezonov B . Zaradi asimetrije energij se težišče trka giblje v smeri žarka elektronov, ki jo imenujemo smer naprej. Nastali mezoni zato pred razpadom prepotujejo merljivo razdaljo od mesta trka, kar omogoča določitev mest razpadov in s tem razdalje preleta.

Poleg resonance $\Upsilon(4S)$ v trkih nastajajo tudi pari lažjih kvarkov in antikvarkov, $e^+ e^- \rightarrow q\bar{q}$ ($q = u, d, s, c$), kar imenujemo kontinuum. Ti pari takoj hadronizirajo; v primeru $c\bar{c}$ pri tem nastane par čarobnih mezonov. Kontinuum je glavni vir mezonov D^+ v tej analizi, zato uporabimo podatke pri vseh težiščnih energijah, ne le tistih pri

resonanci $\Upsilon(4S)$: te prispevajo 711 fb^{-1} pri Belle in 365 fb^{-1} pri Belle II, preostanek pa je bil zajet drugod.

Spektrometer Belle II

Belle II [56] je velik magnetni spektrometer valjaste oblike, ki obdaja mesto trka in je sestavljen iz več poddetektorskih enot. Superprevodni solenoid v njem ustvarja magnetno polje $1,5 \text{ T}$ približno v smeri naprej. Najbližje mestu trka je polprevodniški točkovni detektor (PXD), sestavljen iz dveh plasti silicijevih senzorjev. Obdaja ga silicijev detektor mesta razpada (SVD) s štirimi detektorskimi plastmi. Skupaj natančno določata lego mesta razpada mezonov. Sledi osrednja potovalna komora (CDC), ki v magnetnem polju zaznava sledi nabitih delcev in iz ukrivljenosti sledi določi njihovo gibalno količino. Poleg tega meri specifično ionizacijsko izgubo dE/dx , ki prispeva k identifikaciji delcev (angl. *particle identification*, PID). Za ločevanje nabitih hadronov skrbita dva detektorja, ki izkoriščata sevanje Čerenkova. Nabit delec, ki se skozi snov giblje hitreje od svetlobe v njej, izseva svetlobo pod kotom, odvisnim od njegove hitrosti in skupaj z gibalno količino iz komore CDC to omogoča identifikacijo delca. Merilec časa propagacije (angl. *time of propagation*, TOP) pokriva cilindrični del detektorja CDC. Sestavljen je iz podolgovatih kvarčnih plošč s fotodetektorji na koncu. Čerenkova svetloba se v plošči odbija do fotodetektorjev, iz časa potovanja in mesta zaznave pa rekonstruiramo Čerenkov kot. Detektor obročev sevanja Čerenkova z aerogelom (angl. *aerogel ring-imaging Cherenkov*, ARICH) opravlja isto nalogo v smeri naprej. Sestavljata ga ravnini aerogela in fotodetektorjev. Čerenkova svetloba, ki nastane v aerogelu, na fotodetektorski ravnini nariše obroč, katerega polmer podaja Čerenkov kot. Elektromagnetni kalorimeter (ECL) meri energijo fotonov in elektronov ter je glavni detektor za identifikacijo elektronov. Kalorimeter omogoča tudi zajem fotonov zavnega sevanja, kar uporabljamo pri rekonstrukciji elektronskih parov v tej analizi. Najbolj zunaj je detektor mezonov K_L^0 in mionov (KLM), ki ti dve vrsti delcev prepozna po globini vdora v železo magnetnega jarma.

Spektrometer Belle

Spektrometer Belle [69] je bil zgrajen zelo podobno, le da je bila večina sklopov pri prehodu na Belle II nadgrajena. Detektorji SVD, CDC, ECL in KLM so obstajali že v njem, detektorja PXD pa ni imel. Identifikacijo delcev sta namesto detektorjev TOP in ARICH opravljal merilnik časa preleta (angl. *time of flight*, TOF) in pragovni števec Čerenkove svetlobe (ACC).

Simulacija in programsko okolje

Simulirane vzorce po metodi Monte Carlo uporabimo za določitev izkoristkov rekonstrukcije, ki jih iz podatkov ni mogoče izmeriti, in za preučitev sestave ozadja. Signalne dogodke generiramo ob predpostavki enakomerne porazdelitve po trodelčnem faznem prostoru. Validacijski kanal generiramo z modelom skalar $\rightarrow$ vektor-skalar (SVS) za razpad $D^+ \rightarrow \pi^+ \phi$ in z modelom vektor $\rightarrow$ lepton-lepton (VLL) za razpad $\phi \rightarrow e^+ e^-$. Normalizacijski kanal v Belle II generiramo z modelom Dalitzovega diagrama; v simulacijskem okolju Belle ta model ni na voljo, zato razpad generiramo izključno prek resonance ϕ , kar zahteva popravek, opisan v nadaljevanju. Simulirani dogodki $B\bar{B}$ in dogodki $e^+ e^- \rightarrow q\bar{q}$ skupaj tvorijo generični vzorec, ki ga uporabljamo za modeliranje ozadja in študije zmogljivosti. Odziv detektorja

simuliramo s programom GEANT3 [83] pri Belle in GEANT4 [84] pri Belle II.

Izmerjene podatke in simulirane vzorce obeh eksperimentov analiziramo v programskem ogrodju Belle II [70], pri čemer podatke in simulacijo eksperimenta Belle s paketom B2BII [72] pretvorimo v združljiv format, kar je pogoj za skupno obravnavo obeh vzorcev.

Izbor dogodkov in rekonstrukcija signala

Izbor kandidatov poteka v štirih stopnjah z vse večjimi omejitvami. Centralno pripravljen izbor (imenovan *skim*) zmanjša nabor surovih podatkov na obvladljivo velikost; *ohlapna selekcija* uporablja minimalne zahteve glede kakovosti sledi in kinematike ter se uporablja za preučevanje porazdelitev spremenljivk in usmerjanje nadaljnjih izbir; *začetna selekcija* dodaja celoten nabor kinematičnih zahtev, zahtev glede mesta razpada in identifikacije delcev; *končna selekcija* pa dodaja klasifikator naključnih gozdov in namenski η -veto, s čimer se dobi vzorec, ki se uporabi za ekstrakcijo signala.

Za vse sledi nabitih delcev zahtevamo, da izvirajo v bližini točke interakcije, $|dr| < 0,5 \text{ cm}$ in $|dz| < 2,0 \text{ cm}$, ter da ležijo znotraj območja pokritosti CDC. Elektrone identificiramo z globalnim verjetnostnim razmerjem $\text{PID}_e > 0,9$ v Belle II, ki ne vsebuje informacij iz detektorjev SVD in TOP, ter z binarnim verjetnostnim razmerjem $\mathcal{R}(e, \pi) > 0,9$ v Belle. V vzorcu Belle II pione izberemo s spremenljivko, izračunano z nevronske mreže, $\pi\text{ID}_{\text{NN}} > 0,40$, v vzorcu Belle pa z $\mathcal{R}(\pi, K) > 0,5$. Kaone, ki jih uporabljamo le v normalizacijskem kanalu, izberemo z $\text{PID}_K > 0,6$ (Belle II) ali $\mathcal{R}(K, \pi) > 0,6$ (Belle).

Kandidate za D^+ rekonstruiramo iz treh nabitih sledi s skupnim nabojem $+1$, ki se nahajajo v območju $1,70 < M(\pi^+e^+e^-) < 1,92 \text{ GeV}/c^2$. Okno je namerno asimetrično glede na nominalno maso m_{D^+} , da zajame kandidate, ki so bili slabše rekonstruirani zaradi zavornega sevanja. Za odstranitev mezonov D^+ iz razpadov mezonov B zahtevamo $p_{\text{CMS}} > 2,5 \text{ GeV}/c$. Sledi hčerinskih delcev prilagajamo z algoritmom TreeFitter, ki obravnava celotno razpadno verigo hkrati in omeji mesto razpada D^+ na točko interakcije, pri čemer zahtevamo, da p -vrednost presega 0,001. Kadar dogodek vsebuje več kot enega kandidata, obdržimo tistega z najvišjo p -vrednostjo. Normalizacijski kanal rekonstruiramo enako, pri čemer par elektronov nadomestimo s parom kaonov z nasprotnim nabojem, za katerega velja $|M(K^+K^-) - m_\phi| < 20 \text{ MeV}/c^2$.

Elektroni med potovanjem skozi material izgubljajo energijo zaradi zavornega sevanja, kar poslabša ločljivost $M(\pi^+e^+e^-)$. Izgubljeno energijo povrnemo s posebnim algoritmom, ki ponovno doda skupke fotonov, najdene v stožcu s kotom $2,8^\circ$ okoli smeri elektrona. Algoritem uporabimo pri obeh naborih podatkov; popravke spremeni 20 % kandidatov v simulaciji Belle ter 24 % v simulaciji Belle II.

Dva vira ozadja zahtevata posebno obravnavo. Naključni foton, ki se pretvori v par e^+e^- , posnema končno stanje signala, kadar se eden od pretvorjenih elektronov združi z nepovezano sledjo v signalnem razpadu. Zato vsak signalni elektron združimo z nasprotno nabito sledjo iz preostanka dogodka in kandidata zavržemo, če je tako dobljena invariantna masa manjša od $0,05 \text{ GeV}/c^2$. Kandidatov, pri katerih *oba* leptona izvirata iz pretvorbe fotona, ta veto ne zajame, temveč jih odstranimo na podlagi zahteve $M(e^+e^-) > 0,2 \text{ GeV}/c^2$, ki prav tako odstrani razpad $\pi^0 \rightarrow e^+e^-\gamma$. Drugi vir je $D_{(s)}^+ \rightarrow \pi^+(\eta \rightarrow \gamma e^+e^-)$. Kadar foton obnovimo z algoritmom za zavorno sevanje, ima kandidat vrh pri masi mezona D^+ in pri m_η v $M(e^+e^-)$, zato

ga odstranimo z izključitvijo območja $0,50 < M(e^+e^-) < 0,56 \text{ GeV}/c^2$. Ko fotona ne rekonstruiramo, kandidat nima vrha in preživi to izključitev. V območju $M(e^+e^-)$ pod $0,5 \text{ GeV}/c^2$ predstavlja približno 36 % (Belle) oziroma 53 % (Belle II) ozadja, ki prestane izbiro s klasifikatorjem, opisano v nadaljevanju. Ta prispevek delno odstranimo s posebnim η -vetom, s katerim zavržemo kandidate, pri katerih obnovljeni foton da $|M(e^+e^-\gamma) - m_\eta| < 10 \text{ MeV}/c^2$. S tem odstranimo 8 % (12 %) celotnega ozadja in 1 % celotnega signala v Belle (Belle II). Dodatno izključimo območje $0,95 < M(e^+e^-) < 1,05 \text{ GeV}/c^2$, pri čemer območje $\phi(1020)$ rezerviramo za validacijski kanal.

Prevladujoče kombinatorično ozadje dodatno odstranimo s klasifikatorjem naključnih gozdov (RFC), ki ga za vsak eksperiment posebej naučimo na približno 10 000 simuliranih signalnih in 10 000 simuliranih dogodkih ozadja. RFC uporablja štiri vhodne spremenljivke: gibalno količino mezona D^+ v težiščnem sistemu, celotno vidno energijo v težiščnem sistemu, kot med gibalno količino mezona D^+ in smerjo njegovega preleta ter značilnost razdalje preleta (angl. *significance of flight distance*). Delovni točki izberemo s Punzijevim kriterijem kakovosti $\text{FOM} = \epsilon/(a/2 + \sqrt{B})$ z $a = 5$ in ju postavimo pri vrednostih klasifikatorja 0,97 (Belle) in 0,98 (Belle II). Ohranita 36,8 % (Belle) oziroma 47,8 % (Belle II) signala, pri čemer prepustita 0,275 % oziroma 0,292 % ozadja. Pri enakem izkoristku signala Belle II ohrani približno dvakrat manj ozadja kot Belle, kar kaže na boljše ločevanje signala od ozadja. K temu najverjetneje prispeva pikselni detektor, ki izboljša ločljivost meritve mesta razpada in s tem značilnost razdalje preleta. Preizkusili smo ločeno optimizacijo delovne točke v vsakem intervalu $M(e^+e^-)$ in nismo opazili bistvenega izboljšanja zgornje meje razvejitvenega razmerja, zato uporabljamo enotno delovno točko čez vse intervale.

Dostopni fazni prostor, $0,20 < M(e^+e^-) < 1,75 \text{ GeV}/c^2$, razdelimo na tri intervale, $I_1 = (0,20, 0,50)$, $I_2 = (0,56, 0,95)$ in $I_3 = (1,05, 1,75) \text{ GeV}/c^2$, katerih unija I_{full} določa inkluzivno meritev in že izključuje območji mezonov η in ϕ . Validacijski kanal rekonstruiramo v ožjem oknu I_ϕ , $0,995 < M(e^+e^-) < 1,035 \text{ GeV}/c^2$.

Izkoristke rekonstrukcije določimo iz namenskih simuliranih vzorcev kot delež generiranih razpadov, ki prestanejo celotno izbiro (tabela 1). Izkoristek za signal narašča z $M(e^+e^-)$, in sicer od 3,82 % (4,73 %) v I_1 do 5,94 % (7,42 %) v I_3 za Belle (Belle II). Izkoristek za normalizacijski kanal zahteva dodatno pozornost: svetovno povprečje razvejitvenega razmerja se nanaša na inkluzivni razpad $D^+ \rightarrow \pi^+ K^+ K^-$, zato mora izkoristek vsebovati tudi delež vseh takih razpadov, ki padejo v masno okno mezona ϕ . Pri Belle II to neposredno določimo iz vzorca, generiranega z modelom Dalitzevega diagrama, ki vsebuje resonančno amplitudo, neresonančno amplitudo in njuno interferenco. Za Belle tak generator ni bil na voljo in je bil razpad generiran izključno prek resonance ϕ . Dobljeni izkoristek, določen le iz resonančnega vzorca, zato popravimo s faktorjem 0,308, sestavljenim iz dveh delov: iz deleža $f(M(\phi)) = 28,49\%$ vseh generiranih razpadov $D^+ \rightarrow \pi^+ K^+ K^-$, ki padejo v masno okno mezona ϕ in ga vzamemo iz vzorca Belle II z modelom Dalitzevega diagrama, pomnoženega z razmerjem $\epsilon_{\text{Dalitz}}/\epsilon_{\text{resonance}} = 1,08$ izkoristkov Belle II, določenih znotraj tega okna iz vzorca z modelom Dalitzevega diagrama in iz vzorca zgolj z resonanco.

Ker se analiza tako pri izkoristkih kot pri študijah ozadja opira na simulacijo, po končni izbiri primerjamo podatke in simulacijo v normalizacijskem in validacijskem kanalu ter v stranskih pasovih porazdelitve $M(\pi^+e^+e^-)$. Porazdelitve se po

Tabela 1: Izkoristki rekonstrukcije, uporabljeni pri določitvi razvejitvenega razmerja. Izkoristek za normalizacijski kanal vključuje neresonančni prispevek razpada $D^+ \rightarrow \pi^+ K^+ K^-$ znotraj masnega okna mezona ϕ . Navedene negotovosti so zgolj statistične.

Razpadni kanal	Belle [%]	Belle II [%]
$D^+ \rightarrow \pi^+ e^+ e^-$ (signal, I_{full})	$5,29 \pm 0,02$	$6,57 \pm 0,02$
$D^+ \rightarrow \pi^+ \phi(\rightarrow e^+ e^-)$ (validacija)	$4,75 \pm 0,02$	$6,13 \pm 0,02$
$D^+ \rightarrow \pi^+ K^+ K^-$ (normalizacija)	$1,46 \pm 0,02$	$2,24 \pm 0,01$

obliki ujemajo, kar upravičuje uporabo simulacije za določitev izkoristkov, vendar je število kandidatov v podatkih povsod večje od simuliranega, in sicer za približno 27 % pri Belle in 46 % pri Belle II. Tak presežek, ki lahko izvira iz predpostavljene velikosti produkcije mezonov D^+ ali iz skupne razlike v izkoristku, se v razmerju z normalizacijskim kanalom izniči, zato zanj ne pripišemo sistematične negotovosti. Edina količina, ki jo popravimo neposredno, je izkoristek identifikacije delcev, za kar uporabimo umeritvene tabele. Preostale razlike med podatki in simulacijo, ki se v razmerju ne pokrajšajo popolnoma — na primer odziv klasifikatorja RFC ali fiksirani parametri oblike signala —, obravnavamo kot vire sistematične negotovosti.

Določitev razvejitvenega razmerja in zgornjih mej

Število signalnih dogodkov določimo s prilagajanjem funkcij porazdelitvi $M(\pi^+ e^+ e^-)$ po metodi največjega verjetja, pri čemer kandidatov za signalni in validacijski razpad ne razvrstimo v razrede histograma, kandidate normalizacijskega kanala pa razvrstimo, saj jih je bistveno več. Vrh signala opišemo z dvostransko funkcijo »Crystal Ball«, ki ima v jedru asimetrično Gaussovo obliko in asimetrična potenčna repa. Za normalizacijski kanal uporabimo isto funkcijo, le da je jedro simetrično. Validacijski kanal v vzorcu Belle II opišemo z vsoto funkcije Cruijff, ki je Gaussovi podobna funkcija s spremenjenimi repi, in asimetrične Gaussove porazdelitve s skupno srednjo vrednostjo, v vzorcu Belle pa uporabimo dvostransko funkcijo »Crystal Ball« kot pri signalnem razpadu. Vsako obliko smo izbrali empirično, tako da smo več možnih oblik prilagodili simuliranim signalnim dogodkom in obdržali tiste z najmanjšimi odstopanji. Kombinatorično ozadje opišemo s Čebiševim polinomom, in sicer drugega reda v normalizacijskem kanalu ter nižjega reda drugod, tako da ozadje ne more absorbirati pravega signala.

Oblikovne parametre signalnih funkcij fiksiramo na vrednosti, pridobljene s prilagajanjem simuliranim vzorcem, razen v normalizacijskem kanalu, kjer je vzorec dovolj velik, da jih določimo iz podatkov samih, zato jih pustimo proste. Vsi parametri oblike ozadja so v vsakem prilagajanju prosti. Fiksiranje oblike signala temelji na predpostavki, da simulacija pravilno opiše vrh, opazen v podatkih. Preostalo razliko med njima ovrednotimo v normalizacijskem kanalu in jo upoštevamo kot sistematično negotovost.

Vzorca podatkov Belle in Belle II prilagajamo sočasno. Razvejitveno razmerje vstopa v število signalnih dogodkov posameznega eksperimenta prek njegovega izkoristka rekonstrukcije in števila dogodkov normalizacijskega kanala ter je edini pa-

parameter, ki je skupen obema funkcijama verjetja, medtem ko so števili dogodkov ozadja in oblikovni parametri neodvisni. Ker sta vzorca statistično neodvisna, je sočasno verjetje produkt obeh.

Statistično značilnost morebitnega signala ovrednotimo s testno statistiko razmerja verjetij

$$S = \sqrt{-2 \ln(\mathcal{L}_0/\mathcal{L}_{\max})}, \quad (3)$$

kjer je $\mathcal{L}_{\max}$ verjetje najboljšega prilagajanja, $\mathcal{L}_0$ pa verjetje hipoteze, da je prisotno zgolj ozadje, ki ga dobimo tako, da razvejitveno razmerje fiksiramo na nič. Ovrednotimo jo brez sistematičnih negotovosti, ki so zanemarljive v primerjavi s statističnim nihanjem števila signalnih dogodkov. Kadar S doseže konvencionalni prag petih standardnih odklonov, navedemo izmerjeno razvejitveno razmerje, sicer določimo zgornjo mejo.

Zgornje meje določimo pri 90 % stopnji zaupanja s frekventistično Neymanovo konstrukcijo, ki temelji na psevdoeksperimentih. Za vsako predpostavljeno pravo vrednost razvejitvenega razmerja generiramo množico psevdoeksperimentov, vsakega ponovno prilagodimo enako kot podatke in zabeležimo porazdelitev prilagojenih vrednosti; s ponavljanjem tega postopka po območju predpostavljenih vrednosti izrišemo pas zaupanja. Meja je tista prava vrednost razvejitvenega razmerja, pri kateri le 10 % psevdoeksperimentov da prilagojeno vrednost, manjšo od izmerjene. Ker je negativno razvejitveno razmerje nefizikalno, se omejimo na nenegativne vrednosti in mejo, kadar je prilagojena vrednost negativna, konservativno odčitamo pri nič. Sistematične negotovosti, ki jih je mogoče predstaviti kot raztros fiksnih vhodnih količin prilagajanja, vključimo že pri generiranju psevdoeksperimentov, preostale pa dodamo kot multiplikativne ali aditivne popravke dobljenih mej.

Preden postopek uporabimo na podatkih, ga preverimo na dva načina. Celotno verigo prilagodimo simuliranim vzorcem, ki ne vsebujejo signalnega razpada in morajo zato vrniti razvejitveno razmerje, skladno z nič; to se potrди tako inkluzivno kot v vsakem intervalu $M(e^+e^-)$. S študijami s psevdopodatki (angl. *toy Monte Carlo*) prilagajanje ponovimo na vzorcih, generiranih iz prilagojenih funkcij, in preverimo, ali so porazdelitve normiranih odstopkov prilagojenih parametrov (angl. *pull*) centrirane pri nič in imajo enotsko širino ter ali se prilagojeno razvejitveno razmerje linearno odziva na generirano. Za vse tri kanale je odziv linearen, porazdelitve normiranih odstopkov pa so v večini primerov skladne s pričakovanimi. Preostala neskladja so dovolj majhna, da jih zanemarimo, razen v normalizacijskem kanalu, kjer odstopanje zaradi nepopolnega opisa oblike vrha ohranimo kot sistematično negotovost.

Sistematične negotovosti

Za razliko od statistične negotovosti se sistematične negotovosti z večanjem vzorca podatkov ne manjšajo, z izjemo tistih, ki izvirajo iz končne velikosti simuliranih vzorcev. Izhajajo iz nepopolnega poznavanja odziva detektorja, izkoristka izbora in modelov, uporabljenih pri prilagajanju. Postavljajo torej mejo, pod katero natančnosti meritve ne moremo izboljšati zgolj z zbiranjem več podatkov, in razširijo intervale zaupanja ter zgornje meje.

Pri zgornjih mejah prevladujejo aditivni prispevki. V bližini zgornje mejne vrednosti občutljivost določa število dogodkov ozadja, zato največji multiplikativni viri skorajda nimajo nobenega vpliva. Pri meritvi razvejitvenega razmerja v validacijskem kanalu, kjer je signal prisoten, pa se ta faktor že rahlo pozna, ampak je

sistematična negotovost še vedno za skoraj faktor dva manjša od statistične negotovosti.

Izkoristek identifikacije delcev je edini, ki ga popravimo, in sicer z umeritvenimi faktorji iz neodvisnih študij primerjave podatkov in simulacije. Za vse druge vire ne uporabimo nobenega popravka, opaženo razliko med podatki in simulacijo pa v celoti upoštevamo kot negotovost. Vse takšne prispevke vpeljemo kot simetrično Gaussovo razmazanje, ki interval zaupanja razširi, ne da bi ga premaknilo.

Prispevke k sistematični negotovosti razvrščamo glede na to, kako se propagirajo v končni rezultat, in ne glede na njihov fizikalni izvor, in vsakega od njih naključno spreminjamo v psevdoeksperimentih, ki jih uporabimo za določitev meje. Negotovosti količin, ki so pri prilagajanju fiksirane, upoštevamo tako, da pripadajočo količino variiramo znotraj njene negotovosti že pri generiranju. Tako vpeljemo negotovosti števila dogodkov normalizacijskega kanala, njegovega razvejitvenega razmerja, skupnega multiplikativnega faktorja izkoristka in fiksiranih oblikovnih parametrov signalne porazdelitvene funkcije. Negotovosti, katerih vpliv na prilagojeno razvejitveno razmerje je preprost odmik ali skaliranje, pa uporabimo neposredno na prilagojeni vrednosti vsakega psevdoeksperimenta, in sicer kot naključen aditivni odmik ali kot naključen multiplikativni faktor. Vsi prispevki so zbrani v tabelah 2 in 3, prva navaja tiste, ki so odvisni od intervala $M(e^+e^-)$, druga pa tiste, ki so skupni vsem intervalom.

Multiplikativni faktor izkoristka vsebuje prispevke, povezane z izkoristkom: rekonstrukcije, identifikacije delcev, klasifikatorja, rekonstrukcije sledi, ostanek napačno rekonstruiranih signalnih dogodkov in model signalnega razpada za generacijo vzorcev. Napačno rekonstruirani signalni dogodki predstavljajo le majhen delež razpadov, rekonstruiranih iz napačne kombinacije sledi, ki kljub temu tvorijo vrh pri masi m_{D^+} in jih prilagajanje šteje kot signal. Ker se to pojavi v signalnem in normalizacijskem kanalu, se v izračunu razvejitvenega razmerja v veliki meri izniči, zato kot negotovost pripišemo razliko med deležema v obeh kanalih, ki znaša 1,2 % za Belle in 0,1 % za Belle II. Model signalnega razpada je med multiplikativnimi prispevki največji, 17 % za Belle in 9,6 % za Belle II, in ovrednoti predpostavko, da je signal enakomerno porazdeljen v faznem prostoru treh teles. Negotovost ocenimo tako, da znotraj $M(e^+e^-)$ validacijskega okna vzamemo razliko izkoristkov, določenih na resonančnem validacijskem simuliranem vzorcu in na signalnem simuliranem vzorcu. Negotovost števila dogodkov normalizacijskega kanala ima tri prispevke: statistična negotovost prilagojenega števila dogodkov, izbira reda polinoma za opis ozadja, in nepopolen opis oblike vrha, ki ga ovrednotimo s hibridno množico psevdoeksperimentov in prevladuje nad ostalima dvema. Razvejitveno razmerje normalizacijskega kanala je edina zunanja vhodna količina. Izžrebamo jo enkrat na psevdoeksperiment in uporabimo za oba vzorca, saj ista vrednost skalira oba. Identifikacija delcev, rekonstrukcija sledi in končna velikost simuliranih vzorcev vstopajo dvakrat, enkrat za signalni in enkrat za normalizacijski kanal. Razliko v izkoristku klasifikatorja med podatki in simulacijo določimo samo v normalizacijskem kanalu in jo pripišemo v celoti, saj je v signalnem kanalu ne moremo izmeriti.

Preostala dva prispevka pa upoštevamo po prilagajanju neposredno na prilagojenem razvejitvenem razmerju. Izbira modela ozadja v signalnem kanalu vstopa kot aditivni odmik, ki ga vzamemo kot odklon od nominalnega razvejitvenega razmerja pri prilagajanju s spremenjenim redom polinoma. Drugi prispevek je razlika med obliko signalnega vrha v podatkih in v MC simulaciji in ga upoštevamo kot multi-

plikativni faktor prilagojenega razvejitvenega razmerja. V signalnem kanalu ga ne moremo izmeriti, zato sklepamo iz normalizacijskega kanala s primerjavo njegovih prilagojenih oblikovnih parametrov v obeh, njegov vpliv pa nato ocenimo z namen-sko študijo linearnosti, v kateri generiramo psevdoeksperimente z obliko signala, ki je premaknjena za opazovane razlike, in jih prilagajamo z nominalno obliko. Prila-gojeno razvejitveno razmerje je za 5,0 % nižje od generiranega, kar opredelimo kot simetrično multiplikativno negotovost.

V vsakem intervalu $M(e^+e^-)$ in v celotnem območju je skupna sistematična nego-tovost manjša od statistične negotovosti prilagajanja signala. Ob upoštevanju vseh sistematičnih prispevkov se zgornja meja v celotnem območju $M(e^+e^-)$ poslabša za 12 %, k čemur največ prispeva aditivni prispevek zaradi izbire modela ozadja, medtem ko največji multiplikativni prispevek, model signalnega razpada, poslabša mejo za manj kot $0,05 \times 10^{-7}$.

Tabela 2: Sistematične negotovosti signalnega in validacijskega kanala, ki so odvisne od intervala $M(e^+e^-)$. Negotovost zaradi alternativnega modela ozadja je podana v enotah razvejitvenega razmerja, vse druge v odstotkih. Oznaka v oklepaju pove, ali prispevek upoštevamo kot aditivni premik (adit.) ali kot multiplikativni faktor (mult.).

	$M(e^+e^-)$ interval	I_1	I_2	I_3	I_{full}	I_ϕ
Belle	$\Delta\mathcal{B}(\text{alt. ozadje}) [10^{-7}]$ (adit.)	0,8	0,8	0,6	0,9	0,4
	Izkoristek PID (mult.)	4,9	4,9	4,5	4,7	4,5
	Napačno rekonstruirani signal (mult.)	1,1	0,9	1,4	1,2	0,5
	Izkoristek rekonstrukcije (mult.)	1,0	0,6	0,5	0,4	0,3
	Model signalnega razpada (mult.)	17	17	17	17	–
	Razlika podatki–MC (mult.)	5,0	5,0	5,0	5,0	3,0
Belle II	$\Delta\mathcal{B}(\text{alt. ozadje}) [10^{-7}]$ (adit.)	0,5	1,0	0,4	1,0	< 0,1
	Izkoristek PID (mult.)	5,9	5,7	4,9	5,2	4,6
	Napačno rekonstruirani signal (mult.)	0,3	0,1	< 0,1	0,1	0,5
	Izkoristek rekonstrukcije (mult.)	0,8	0,5	0,4	0,4	0,4
	Model signalnega razpada (mult.)	9,6	9,6	9,6	9,6	–
	Razlika podatki–MC (mult.)	5,0	5,0	5,0	5,0	3,0

Rezultati

V validacijskem kanalu, rekonstruiranem v območju $0,995 < M(e^+e^-) < 1,035 \text{ GeV}/c^2$, izmerimo

$$\mathcal{B}(D^+ \rightarrow \pi^+ \phi (\rightarrow e^+ e^-)) = (1,75 \pm 0,22 (\text{stat.}) \pm 0,12 (\text{sist.})) \times 10^{-6}, \quad (4)$$

Tabela 3: Sistematične negotovosti, ki so skupne vsem intervalom $M(e^+e^-)$, podane v odstotkih. Vse vpeljemo že pri generiranju psevdoeksperimentov. Negotovost razvejitvenega razmerja normalizacijskega kanala je skupna obema vzorcema, saj ista zunanja vrednost skalira oba.

Vir negotovosti	Belle	Belle II
<i>Število dogodkov normalizacijskega kanala</i>		
statistična negotovost	0,4	0,5
izbira modela ozadja	0,3	0,3
nepopoln opis oblike vrha	1,8	2,5
<i>Izkoristki, ovrednoteni v normalizacijskem kanalu</i>		
izkoristek klasifikatorja	4,2	6,0
izkoristek identifikacije delcev	1,3	2,2
izkoristek rekonstrukcije	1,0	0,4
Rekonstrukcija sledi signala	2,1	1,6
$\mathcal{B}(D^+ \rightarrow \pi^+ K^+ K^-)$	1,9	

in sicer s statistično značilnostjo 12,0 standardnih odklonov. Vrednost se ujema s pričakovanjem $(1,70 \pm 0,05) \times 10^{-6}$, izračunanim iz znanih razvejitvenih razmerij obeh stopenj [16], in s prejšnjo neposredno meritvijo kolaboracije CLEO, $(1,7_{-0,9}^{+1,4}) \times 10^{-6}$ [48], pri čemer smo natančnost meritve izboljšali za približno faktor pet. Ujemanje z že znanim razvejitvenim razmerjem kaže, da je postopek, ki ga uporabimo tudi v signalnem kanalu, zanesljiv in ne vsebuje znatne sistematične pristranskosti.

Prilagajanja normalizacijskega kanala dajo $91\,663 \pm 1711$ dogodkov pri Belle in $59\,119 \pm 1524$ pri Belle II, pri čemer navedena negotovost združuje statistični prispevek (367 oziroma 324 dogodkov) s sistematičnimi, ki prevladujejo. Ti števili uporabimo za normalizacijo v enačbi (2).

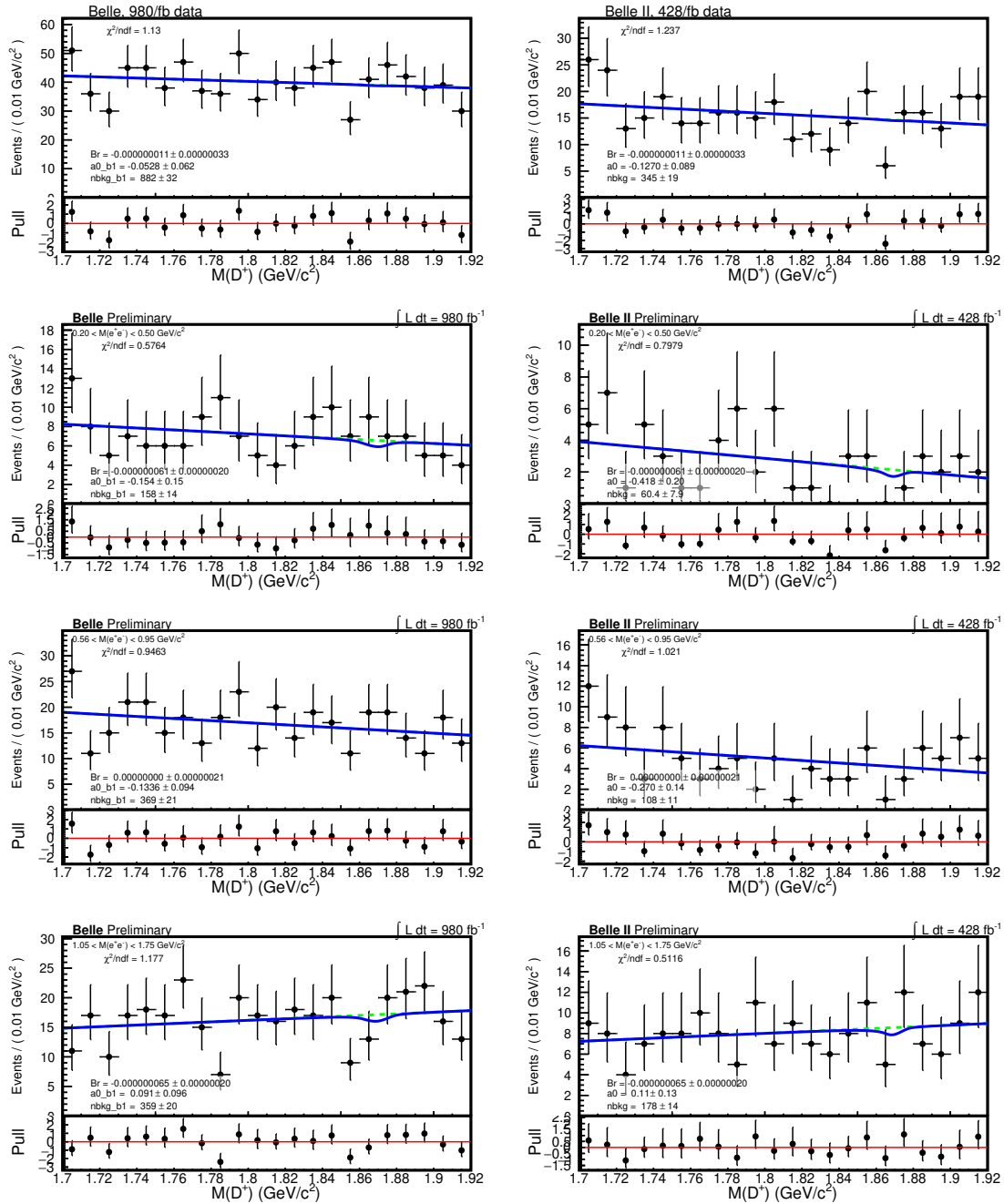
V signalnem kanalu ne opazimo znatnega presežka dogodkov glede na hipotezo, da je prisotno zgolj ozadje, niti v celotnem območju $M(e^+e^-)$ niti v katerem od treh intervalov. Prilagajanja so prikazana na sliki 2.

V celotnem območju $M(e^+e^-)$, iz katerega izvzamemo resonanci η in ϕ , dobimo

$$\mathcal{B}(D^+ \rightarrow \pi^+ e^+ e^-) < 4,7 \times 10^{-7} \quad \text{pri } 90\% \text{ stopnji zaupanja,} \quad (5)$$

s čimer prejšnjo najboljšo mejo, ki jo je postavila kolaboracija *BABAR* [4], izboljšamo za faktor 2,3. Navedena meja vključuje vse sistematične negotovosti. Brez njih bi bila meja nižja za $0,5 \times 10^{-7}$. V posameznih intervalih dobimo zgornje meje delnih razvejitvenih razmerij $3,0 \times 10^{-7}$, $3,3 \times 10^{-7}$ in $2,9 \times 10^{-7}$ za I_1 , I_2 in I_3 . To so prve meje za ta razpad, določene v intervalih $M(e^+e^-)$.

S prilagajanjem vzorcev neodvisno drug od drugega in brez upoštevanja sistematičnih negotovosti dobimo zgornji meji 12×10^{-7} v vzorcu Belle in $6,9 \times 10^{-7}$



Slika 2: Sočasna prilagajanja porazdelitvi $M(\pi^+e^+e^-)$ signalnega kanala v vzorcu Belle (levo) in Belle II (desno), prikazana od zgoraj navzdol za celotno območje $M(e^+e^-)$ ter za intervale I_1 , I_2 in I_3 . Razvejitevno razmerje je v vsakem primeru skupni parameter obeh prilagajanj. Točke s pripadajočimi negotovostmi prikazujejo podatke. Z modro krivuljo je prikazana celotna prilagojena funkcija, z zeleno pa prispevek ozadja.

v vzorcu Belle II, v primerjavi s $4,2 \times 10^{-7}$ iz sočasnega prilagajanja pod enakimi pogoji. Izboljšanje je večje, kot bi ga dala zgolj preprosta kombinacija dveh neodvisnih meritev, saj nobena od neodvisnih mej ne odraža neposredno svoje lastne občutljivosti. Prilagajanje Belle odstopa navzgor za približno en standardni odklon, prilagajanje Belle II pa da vrednost pod fizikalno mejo, zato mejo odčitamo pri nič. Sočasno prilagajanje združi obe funkciji verjetja, še preden zgradimo pas zaupanja, zato se ti odstopanji deloma izničita, namesto da bi vsako posebej popačilo svojo mejo.

Zaključek

V delu iščemo redek razpad $D^+ \rightarrow \pi^+ e^+ e^-$, ki poteka preko nevtralnega toka s spreminjanjem okusa, v združenih vzorcih podatkov, zajetih s spektrometroma Belle in Belle II na trkalnikih e^+e^- KEKB in SuperKEKB na Japonskem. Razpad je v Standardnem modelu močno potlačen in je zato občutljiv na prispevke fizike onkraj Standardnega modela, čeprav je njegovo razvejitveno razmerje zelo močno pogojeno z dolgosežnimi prispevki, ki zakrijejo kratkosežne prispevke. Znatnega signala ne opazimo, zato postavimo doslej najstrožjo zgornjo mejo njegovega razvejitvenega razmerja, skupaj s prvimi zgornjimi mejami, določenimi v intervalih invariantne mase elektronskega para. Analizni postopek preverimo z meritvijo $\mathcal{B}(D^+ \rightarrow \pi^+ \phi(\rightarrow e^+ e^-))$, ki se ujema s prejšnjo neposredno meritvijo pri občutno izboljšani natančnosti. Posodobljene meje ostajajo več velikostnih redov nad kratkosežnim razvejitvenim razmerjem, in tako dodatno omejujejo morebitne ojačitve razmerja, ki bi izvirale iz fizike onkraj Standardnega modela, ne sežejo pa do prispevka Standardnega modela samega. Občutljivost te meritve omejuje velikost vzorca in ne sistematične negotovosti. Ob skaliranju kot $1/\sqrt{\mathcal{L}}$ in ob predpostavki nespremenjenih izkoristkov, sestave ozadja in izbora bi z vzorcem Belle II velikosti 50 ab^{-1} dosegli mejo $6,4 \times 10^{-8}$, določeno brez sistematičnih negotovosti tako kot meja $6,9 \times 10^{-7}$, iz katere izhaja. Projekcija predpostavlja tudi, da bo zmogljivost detektorja pri višji luminoznosti ostala primerljiva z današnjo, pri čemer ne moremo izključiti, da bo sistematična meja dosežena prej.

