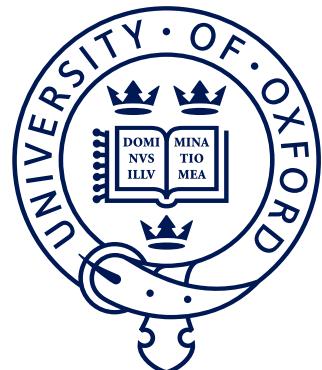


<sup>1</sup> **CP violation in  $B^\pm \rightarrow Dh^\pm$  decays**  
<sup>2</sup> **where  $D \rightarrow K_S^0 h'^+ h'^-$**

<sup>3</sup> *A measurement of the CKM angle  $\gamma$  at LHCb and understanding  
<sup>4</sup> the impact of neutral kaon CP violation*



<sup>6</sup> Mikkel Bjørn  
<sup>7</sup> St. Anne's College  
<sup>8</sup> University of Oxford

<sup>9</sup> A thesis submitted for the degree of  
<sup>10</sup> *Doctor of Philosophy*  
<sup>11</sup> Trinity 2020

<sup>12</sup>

## Acknowledgements

<sup>13</sup> suitable thank you's

# Abstract

14

<sup>15</sup> World's best measurement of  $\gamma$ . Details to be added.

# Contents

16

<sup>17</sup>	<b>Preface</b>	<sup>vii</sup>
<sup>18</sup>	<b>1 Introduction</b>	<b>1</b>
<sup>19</sup>	1.1 Structure of the thesis . . . . .	1
<sup>20</sup>	<b>2 Theoretical background</b>	<b>2</b>
<sup>21</sup>	2.1 The C, P and T symmetries and their violation . . . . .	2
<sup>22</sup>	2.2 CP violation in the Standard Model . . . . .	4
<sup>23</sup>	2.2.1 The CKM matrix and the Unitarity Triangle . . . . .	4
<sup>24</sup>	2.2.2 Measuring $\gamma$ in tree level decays . . . . .	9
<sup>25</sup>	2.3 Measuring $\gamma$ using multi-body D final states . . . . .	12
<sup>26</sup>	2.3.1 Dalitz plots and the phase space of multi-body decays . . . . .	12
<sup>27</sup>	2.3.2 The GGSZ method to measure $\gamma$ . . . . .	14
<sup>28</sup>	2.3.3 A model-independent approach . . . . .	16
<sup>29</sup>	2.3.4 Measuring strong-phase inputs at charm factories . . . . .	18
<sup>30</sup>	2.3.5 Global CP asymmetry and the relation to GLW and ADS measurements . . . . .	22
<sup>31</sup>	2.4 Strategy for the LHCb measurement . . . . .	26
<sup>33</sup>	<b>3 The LHCb experiment</b>	<b>29</b>
<sup>34</sup>	3.1 The LHCb subdetectors . . . . .	31
<sup>35</sup>	3.1.1 The VELO . . . . .	32
<sup>36</sup>	3.1.2 Magnet and tracking stations . . . . .	34
<sup>37</sup>	3.1.3 The RICH detectors . . . . .	36
<sup>38</sup>	3.1.4 Calorimeters . . . . .	38
<sup>39</sup>	3.1.5 Muon detectors . . . . .	40
<sup>40</sup>	3.2 Reconstruction . . . . .	40
<sup>41</sup>	3.2.1 Track reconstruction . . . . .	41
<sup>42</sup>	3.2.2 Particle identification . . . . .	42
<sup>43</sup>	3.3 The LHCb trigger system . . . . .	44
<sup>44</sup>	3.3.1 The level-0 hardware trigger . . . . .	45
<sup>45</sup>	3.3.2 High-level triggers . . . . .	46
<sup>46</sup>	3.3.3 Offline data filtering: the LHCb stripping . . . . .	47
<sup>47</sup>	3.4 Simulation . . . . .	48

48	<b>4 Neutral kaon <math>CP</math> violation and material</b>	49
49	<b>interaction in BPGGSZ measurements</b>	49
50	4.1 CP violation and material interaction of neutral kaons . . . . .	49
51	4.1.1 A first look at the impact on $\gamma$ measurements . . . . .	53
52	4.2 Impact on BPGGSZ measurements of $\gamma$ :	
53	principles . . . . .	55
54	4.2.1 Modified symmetry relations . . . . .	55
55	4.2.2 Relationship between the $K_S^0$ and $K_L^0$ amplitudes . . . . .	56
56	4.2.3 Modification of the BPGGSZ yield equations . . . . .	57
57	4.3 Impact on BPGGSZ measurements of $\gamma$ :	
58	LHCb and Belle II measurements . . . . .	60
59	4.3.1 Detector geometries . . . . .	61
60	4.3.2 Kaon momentum distributions . . . . .	62
61	4.3.3 Experimental time acceptance . . . . .	63
62	4.3.4 Detector material budget . . . . .	64
63	4.3.5 Calculation procedure . . . . .	66
64	4.3.6 Results . . . . .	67
65	4.3.7 Coupled $B^\pm \rightarrow DK^\pm$ and $B^\pm \rightarrow D\pi^\pm$ measurements . . . . .	69
66	4.4 Concluding remarks . . . . .	70
67	<b>5 A BPGGSZ measurement of <math>\gamma</math> with <math>B^\pm \rightarrow Dh^\pm</math> decays</b>	72
68	5.1 Candidate reconstruction and selection . . . . .	72
69	5.1.1 Initial requirements . . . . .	73
70	5.1.2 Boosted decision tree . . . . .	75
71	5.1.3 Particle-identification requirements . . . . .	81
72	5.1.4 Final requirements . . . . .	82
73	5.1.5 Selected candidates . . . . .	82
74	5.2 Signal selection efficiencies . . . . .	83
75	5.2.1 Efficiency of the PID requirements . . . . .	86
76	5.2.2 Efficiency profile over the Dalitz plot . . . . .	87
77	5.3 Background studies . . . . .	89
78	5.3.1 Charmless decays . . . . .	91
79	5.3.2 Background from four-body $D$ decays . . . . .	92
80	5.3.3 Semi-leptonic backgrounds . . . . .	93
81	5.3.4 Cross-feed from other $D \rightarrow K_S^0 h^+ h^-$ decays . . . . .	101
82	5.3.5 Swapped-track backgrounds . . . . .	101
83	5.4 Signal and background mass shapes . . . . .	102
84	5.4.1 Signal decays . . . . .	104
85	5.4.2 Cross-feed between $B^\pm \rightarrow Dh^\pm$ channels . . . . .	105

86	5.4.3	Partially reconstructed backgrounds . . . . .	107
87	5.4.4	Combinatorial background . . . . .	113
88	5.4.5	Fit results . . . . .	113
89	5.5	Measurement of the CP-violation observables . . . . .	116
90	5.5.1	Fit setup . . . . .	117
91	5.5.2	Main results . . . . .	124
92	5.5.3	Cross checks . . . . .	126
93	5.6	Systematic uncertainties . . . . .	136
94	5.6.1	Strong phase uncertainties . . . . .	136
95	5.6.2	Efficiency-profile-related systematic uncertainties . . . . .	137
96	5.6.3	Mass shapes . . . . .	140
97	5.6.4	$CP$ violation and material interaction of the $K_S^0$ . . . . .	147
98	5.6.5	Impact of $D$ mixing . . . . .	147
99	5.6.6	PID efficiencies . . . . .	148
100	5.6.7	Dalitz-coordinate resolution . . . . .	148
101	5.6.8	The fixed yield fractions . . . . .	150
102	5.6.9	Systematic uncertainty due to backgrounds that are not modelled in fit . . . . .	150
103	5.6.10	Bias correction . . . . .	152
104	5.6.11	Charmless backgrounds . . . . .	153
105	5.6.12	Summary of systematic uncertainties . . . . .	153
106	5.7	Obtained constraints on $\gamma$ . . . . .	154
107	5.7.1	Statistical approach . . . . .	155
108	5.7.2	Interpretation results . . . . .	157
109	5.7.3	Compatibility with other measurements . . . . .	160
110	<b>6</b>	<b>Summary and outlook</b>	<b>164</b>
111	<b>Appendices</b>		
112	<b>A</b>	<b>Projections of the main fit to data</b>	<b>166</b>

# Preface

<sup>115</sup> The work presented in this thesis has been resulted in two papers, either under  
<sup>116</sup> review or published in the Journal of High Energy Physics. These are

<sup>117</sup> *[?] Measurement of the CKM angle  $\gamma$  using  $B^\pm \rightarrow [K_S^0 h^+ h^-]_D h^\pm$  decays,*  
<sup>118</sup> submitted to JHEP.

<sup>119</sup> This paper describes a measurement of the CKM angle  $\gamma$  using  $pp$  collision  
<sup>120</sup> data taken with the LHCb experiment during the Run 1 of the LHC, in 2011  
<sup>121</sup> and 2012, and during the full Run 2, in 2015–2018. The measurement uses the  
<sup>122</sup> decay channels  $B^\pm \rightarrow D h^\pm$  where  $D \rightarrow K_S^0 h'^+ h'^-$ , in which  $h$  and  $h'$  denotes  
<sup>123</sup> pions or kaons. It obtains a value of  $\gamma = (69 \pm 5)^\circ$ , which constitutes the  
<sup>124</sup> world’s best single-measurement determination of  $\gamma$ . The work is the main  
<sup>125</sup> focus of this thesis and described in detail in Chapter 5.

<sup>126</sup> *[?] CP violation and material interaction of neutral kaons in measurements*  
<sup>127</sup> *of the CKM angle  $\gamma$  using  $B^\pm \rightarrow DK^\pm$  decays where  $D \rightarrow K_S^0 \pi^+ \pi^-$* , JHEP  
<sup>128</sup> 19 (2020) 106.

<sup>129</sup> This paper describes a phenomenological study of the impact of neutral-  
<sup>130</sup> kaon  $CP$  violation and material interaction on measurements of  $\gamma$ . With the  
<sup>131</sup> increased measurement precision to come in the near future, an understanding  
<sup>132</sup> of these effects is crucial, especially in the context of  $B \rightarrow D\pi$  decays; however  
<sup>133</sup> no detailed study had been published at the start of this thesis. The study is  
<sup>134</sup> the subject of Chapter 4. Some text excerpts and figures from the paper have  
<sup>135</sup> been reproduced in the thesis.

<sup>136</sup> All of the work described in the thesis is my own, except where clearly referenced  
<sup>137</sup> to others. Furthermore, I contributed significantly to an analysis of  $B^\pm \rightarrow DK^\pm$   
<sup>138</sup> decays with LHCb data taken in 2015 and 2016, now published in

<sup>139</sup> *[?] Measurement of the CKM angle  $\gamma$  using  $B^\pm \rightarrow DK^\pm$  with  $D \rightarrow K_S^0 \pi^+ \pi^-$ ,*  
<sup>140</sup>  *$K_S^0 K^+ K^-$  decays*, JHEP 08 (2018) 176.

<sup>141</sup> I was responsible for the selection and analysis of the signal channel, studies of  
<sup>142</sup> systematic uncertainties, and the interpretation of the measured observables in  
<sup>143</sup> terms of underlying physics parameters. The measurement is superseded by that  
<sup>144</sup> of Ref. [?] and is not described in detail in the thesis.

145 Within the LHCb collaboration, I took part in the preparation for Run 3 by  
146 working as a *migration coordinator* for the *B-decay-to-open charm* (B2OC) physics  
147 working group, responsible<sup>1</sup> for the migration of the wo~~o~~g group's centralised,  
148 offline selections (so called *stripping lines*) to the software-trigger framework being  
149 developed for the LHCb Upgrade.<sup>2</sup> With more than 800 lines in the current B2OC  
150 stripping module this is a major task; I took a leading role in the initial design and  
151 testing of the upgraded B2OC module, and in helping the first analysts implement  
152 their selections within it. I have also undertaken shift work as RICH piquet and  
153 Data Manager, and acted as the liason between the B2OC physics working group  
154 and the *particle-identification* performance working group.

---

<sup>1</sup>Along with Alessandro Bertolin and Shunan Shang.

<sup>2</sup>The current software trigger and *stripping* framework is presented in Section 3.3.

*Le roi est mort, vive le roi!*

— Traditional French proclamation at the death  
of one monarch and the ascension of a new

# 1

155

156

## Introduction

157 All the big picture stuff: constraints on New Physics from high precision measure-  
158 ments, a small nod to matter-antimatter asymmetry questions etc.

159 This thesis places itself at the forefront of the latter effort, presenting the  
160 World’s most precise measurement of the  $CP$ -violating phase  $\gamma$ ; a fundamental  
161 parameter in the Standard Model, in which it describes the sole source of matter-  
162 antimatter asymmetries.

163 Naturally, the work presented here is far from the last word on the value of  
164  $\gamma$ . In the very near future, several important results will be updated by the  
165 LHCb collaboration, based on the full data set collected during Run 2 of the LHC.  
166 In the slightly longer term, both the LHCb and Belle II collaborations expect  
167 to collect data samples of  $B$  decays that are orders of magnitudes larger than  
168 the current ones during the coming decade, pushing the obtainable precision on  
169  $\gamma$  towards, even below, one degree. An outlook towards this exciting future of  
170  $CP$ -violation measurements is given in Chapter 6, along with a summary of the  
171 contributions made in the present thesis.

### 172 1.1 Structure of the thesis

# 2

173

174

## Theoretical background

175 This chapter lays out the theoretical framework of the thesis. Section 2.1 introduces  
176 charge and parity symmetry violation in general, while Section 2.2 covers the  
177 description in the Standard Model and the general theory behind charge-parity  
178 symmetry violation measurements in charged  $B$  decays. Section 2.3 focuses on  
179 the theory of measurements using  $B^\pm \rightarrow Dh^\pm$  decays with multi-body  $D$  final  
180 states, after which the specific analysis strategy for the measurement described  
181 in the thesis is outlined out in Section 2.4.

### 182 2.1 The C, P and T symmetries and their vio- 183 lation

184 The concept of symmetry play a fundamental role in modern physics. By Noether's  
185 theorem [?], the simple assumption of invariance of our physical laws under universal  
186 temporal and spatial translations leads to the very non-trivial prediction of conserved  
187 energy and momentum; within the field of particle physics, the interactions and  
188 dynamics of the Standard Model (SM) follow completely simply from requiring  
189 the fundamental particle fields to satisfy a local  $U(1) \times SU(2) \times SU(3)$  gauge  
190 symmetry [?]; and one of the short-comings of the SM, is that it fails to explain  
191 the apparent *lack* of symmetry in our matter-dominated universe [?]. Indeed, it is  
192 important to experimentally establish the symmetries of our world at a fundamental  
193 level, and the degree to which they are broken.

194 Three discrete symmetries of importance are the symmetries under

I'll  
adjust  
this para-  
graph  
when  
I've  
written  
the intro-  
duction.

- 195     1. The charge operator  $C$ , which conjugates all internal quantum numbers of a  
 196       quantum state and thus converts particles into their anti-particle counter parts.  
 197       For example,  $C$  transforms the electric charge of a particle state  $Q \rightarrow -Q$ .
- 198     2. The parity operator  $P$ , which inverts the spatial dimensions of space time:  
 199        $\vec{x} \rightarrow -\vec{x}$ . As such, it transforms left-handed particle fields into right-handed  
 200       particle fields and vice versa.
- 201     3. The time-inversion operator  $T$ , which inverts the temporal dimension of space  
 202       time:  $t \rightarrow -t$ .

203     These are fundamentally related by the *CPT* theorem [?], which states that any  
 204       Lorentz-invariant Quantum Field Theory (QFT) must be symmetric under the  
 205       simultaneous application of *all* three operators. However, any one of the symmetries  
 206       can be broken individually, and experiments have shown the physical laws of our  
 207       world to violate each of the  $C$ ,  $P$ , and  $T$  symmetries.

208     Such a symmetry-breaking effect was established for the first time in 1956, when  
 209       Chien-Shiung Wu observed parity violation in weak decays of Co-60 nuclei [?], after  
 210       carrying out an experiment that was proposed by Yang Chen-Ning and Tsung-Dao  
 211       Lee [?]. While this experiment established the breaking of  $P$  symmetry, it left open  
 212       the possibility that the physical laws are invariant under a combination of a charge-  
 213       and parity in [?] sion; that they are  $CP$  symmetric. However, this was disproved in  
 214       1964 when [?] in and Fitch observed that long-lived kaons, which predominantly  
 215       decay to the  $CP$ -odd  $3\pi$  state, could also decay to the  $CP$ -even  $\pi\pi$  states [?].

216     Since then  $CP$  violation has been found in the  $B^0$  system by the BaBar and Belle  
 217       collaborations [?, ?] during the early 2000's; the  $B$  factories, along with CDF, also saw  
 218       evidence for  $CP$  violation in  $B^\pm$  decays [?, ?, ?, ?, ?, ?] [?] or confirmed by LHCb [?],  
 219       and  $CP$  violation was measured for the  $B_s^0$  meson by LHCb in 2013 [?]; within the  
 220       last year and a half, the first observation of  $CP$ -violation in  $D^0$  decays has also been  
 221       made by the LHCb collaboration [?], and most recently evidence for  $CP$ -violation in  
 222       the neutrino sector has been reported by the T2K collaboration [?]. The observed  
 223       effects can be divided into distinct classes. The conceptually simplest case is

- 224     1.  *$CP$ -violation in decay*, where  $|A/\bar{A}| \neq 1$  for some decay amplitude  $A$ , and the  
 225       amplitude  $\bar{A}$  of the  $CP$ -conjugate decay. The result is different decay rates in  
 226       two  $CP$ -conjugate decays

$$\Gamma(M \rightarrow f) \neq \Gamma(\bar{M} \rightarrow \bar{f}). \quad (2.1)$$

227     This type of  $CP$  violation was not seen until the late 1980ies [?, ?], more than  
 228       20 years after the first observation of  $CP$  violation, and only finally established  
 229       around the year 2000 [?, ?]. Also this discovery was made in  $K \rightarrow \pi\pi$  decays.

230  $CP$ -violation in decay is the only type possible for charged initial states, and  
231 it is thus the main focus of the thesis. Two additional  $CP$ -violating eff 232 are  
233 possible for neutral initial states (a situation that will be the main focus of  
Chapter 4). These effects are

- 234 2.  *$CP$ -violation in mixing*, which denotes the case where the mixing rates between  
235 the  $M^0$  and  $\bar{M}^0$  states differ

$$\boxed{\Gamma(M^0 \rightarrow \bar{M}^0) \neq \Gamma(\bar{M}^0 \rightarrow M^0)}. \quad (2.2)$$

236 The  $CP$  violation first observed by Kronin and Fitch in the neutral kaon  
237 sector [?] is (dominantly) of this type.

- 238 3.  *$CP$ -violation in interference between mixing and decay*, which can be present  
239 for a neutral initial states  $M^0$  decaying into a final state  $f$  common to both  
240  $M^0$  and  $\bar{M}^0$ . The decay rate includes an interference term between two  
241 amplitudes: the amplitude for a direct  $M^0 \rightarrow f$  decay and the amplitude  
242 for a decay after mixing:  $M^0 \rightarrow \bar{M}^0 \rightarrow f$ . Even in the absence of the two  
243 aforementioned effects, the rates  $\Gamma(M^0 \rightarrow f)$  and  $\Gamma(\bar{M}^0 \rightarrow \bar{f})$  can differ due  
244 to the interference term. Such  $CP$  asymmetries have been measured in eg.  
245  $B^0 \rightarrow J/\psi K$  by LHCb and the  $B$  factories, and in  $B_s^0 \rightarrow J/\psi \phi$  decays by the  
246 LHC and Tevatron experiments [?].

247  $CP$  violation measurements thus have a long, rich, and still-developing history.

## 248 2.2 **CP violation in the Standard Model**

249 All existing measurements of  $CP$  violation in the quark sector are naturally explained  
250 in the SM; indeed, the need to explain the observation  $CP$  violation in neutral  
251 kaons was a driving force in the development of the model in the first place, when  
252 it lead Kobayashi and Maskawa to predict the existence of then-unknown particles  
253 in 1973 [?] (now known to be the third generation quarks).

### 254 2.2.1 **The CKM matrix and the Unitarity Triangle**

255 The SM contains three generations of quarks, each consisting of an up-type quark ( $u$ ,  
256  $c$ , and  $t$ ) and a down-type quark ( $d$ ,  $s$ , and  $b$ ). The charged weak interaction of the  
257  $W^\pm$  boson couples up and down-type quarks. The quark states that couple to the  $W$   
258 are not (a priori) identical to the mass eigenstates, and can be denoted ( $u'$ ,  $c'$ , and  $t'$ )  
259 and ( $d'$ ,  $s'$ , and  $b'$ ). A basis for the quark states can be chosen such that the weakly

260 coupling up-quark states are identical to the propagating quark states,  $u = u'$ , but  
261 then the down-type quark state are different:  $d' \neq d$ . The two bases of the down-type  
262 quarks are related via the Cabibbo-Kobayashi-Maskawa (CKM) matrix [?, ?]<sup>1</sup>

$$\begin{pmatrix} d' \\ s' \\ t' \end{pmatrix} = V \begin{pmatrix} d \\ s \\ t \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ t \end{pmatrix}. \quad (2.3)$$

263 Thus the Lagrangian terms representing the coupling of a  $W^\pm$  boson with a  $u$ -  
264 and a  $d$ -type quark is

$$\mathcal{L}_{W^+} = -\frac{g}{\sqrt{2}} V_{ud} (\bar{u} \gamma^\mu W_\mu^+ d) \quad \mathcal{L}_{W^-} = -\frac{g}{\sqrt{2}} V_{ud}^* (\bar{d} \gamma^\mu W_\mu^- u) \quad (2.4)$$

265 where  $g$  is the weak coupling constant,  $\gamma_u$  are the Dirac matrices, and  $u$  and  $d$   
266 represent the left-handed components of the physical quark states.

267 The CKM matrix is a unitary complex  $3 \times 3$  matrix, and hence has  $3^2 = 9$   
268 independent, real parameters. However, 5 of these can be absorbed into non-physical  
269 phases of the quark states (both mass and weak eigenstates) and hence the matrix  
270 has 4 real, physical parameters: 3 mixing angles and a single phase. Chau and  
271 Keung [?] proposed the parameterisation

$$\begin{aligned} V &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta_{CP}} \\ 0 & 1 & 0 \\ -s_{13}e^{-i\delta_{CP}} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 1 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{CP}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{CP}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{CP}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{CP}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{CP}} & c_{23}c_{13} \end{pmatrix} \end{aligned} \quad (2.5)$$

272 which is the preferred standard by the PDG [?]. Here,  $s_{ij} \equiv \sin \theta_{ij}$  and  $c_{ij} \equiv \cos \theta_{ij}$   
273 denote the sine and cosine of three rotation angles in quark space;  $\theta_{12} = \theta_C$  being  
274 the usual Cabibbo angle [?].

275 The presence of the complex phase  $\delta_{CP}$  in the Lagrangian term of the  $W$  coupling  
276 causes  $CP$  violation because, as evident from Eq. (2.4), if  $\delta_{CP}$  enters the amplitude  
277 for some decay mediated by a  $W$  boson,  $A = |A|e^{i(\delta_0 + \delta_{CP})}$ , then it will enter the  
278  $CP$  conjugate decay amplitude with the opposite sign:  $\bar{A} = |A|e^{i(\delta_0 - \delta_{CP})}$ . In these  
279 expressions,  $\delta_0$  denotes a  $CP$  conserving phase that is not caused by complex terms  
280 in the Lagrangian, but arises due to potential intermediate states in the decay

---

<sup>1</sup> A basis for the quarks can of course be chosen, such that neither the up-quarks or the down-quarks are expressed in their mass eigenstates. In that case the CKM matrix is recovered as  $V = U_{u/d}^* U_d$ , where  $U_{u/d}$  is the unitary transformation matrices that brings the  $u/d$  quarks into their  eigenstates.

amplitude.<sup>2</sup> Usually the underlying mechanism is due to QCD effects, and these  $CP$  conserving phases are therefore generally dubbed *strong* phases, as opposed to the  $CP$  violating *weak* phase of the  $W$  coupling [?]. This terminology will be applied throughout the thesis.

Experimentally, it has been observed that the CKM matrix elements of Eq. (2.5) satisfy  $s_{13} \ll s_{23} \ll s_{12}$ . This motivates an often used, alternative parameterisation of the matrix, where the elements are expressed as power series in a parameter  $\lambda$  that naturally incorporates this hierarchy: the Wolfenstein parameterisation [?].

The definitions

$$\begin{aligned} s_{12} &\equiv \lambda \\ s_{23} &\equiv \lambda^2 A \\ s_{13} &\equiv \lambda^3(\rho - i\eta) \end{aligned} \tag{2.6}$$

are made, after which the unitarity conditions (or Eq. 2.5) determine the remaining elements to any order in  $\lambda$ .<sup>3</sup> To  $\mathcal{O}(\lambda^5)$  the Wolfenstein parameterisation of the CKM matrix is [?, ?]

$$V = \begin{pmatrix} 1 - \frac{\lambda^2}{2} - \frac{\lambda^4}{8} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda + \frac{\lambda^5}{2}A^2(1 - 2(\rho + i\eta)) & 1 - \frac{\lambda^2}{2} - \frac{\lambda^4}{8}(1 + 4A^2) & A\lambda^2 \\ A\lambda^3(1 - (\rho + i\eta)(1 - \frac{\lambda^2}{2})) & -A\lambda^2(1 - \frac{\lambda^2}{2}(1 - 2(\rho + i\eta))) & 1 - \frac{1}{2}A^2\lambda^4 \end{pmatrix}. \tag{2.7}$$

293

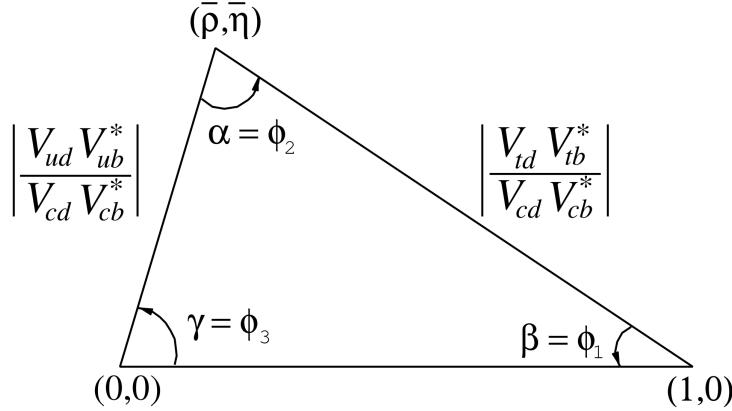
294 The unitarity condition  $V^\dagger V = \mathbb{1}$  of the CKM matrix defines 9 relations between  
295 the CKM elements of the form

$$\sum_j V_{jq}^* V_{jq} = 1 \quad , \quad q \in \{d, s, b\} \quad \text{along the diagonal} \tag{2.8a}$$

$$\sum_j V_{jq}^* V_{jq'} = 0 \quad , \quad q, q' \in \{d, s, b\}, q \neq q' \quad \text{off-diagonal.} \tag{2.8b}$$

296 The off-diagonal conditions constrain three complex numbers to sum to zero, and  
297 can thus be visualised as triangles in the complex plane, the so-called unitarity  
298 triangles. Of these, the triangle corresponding to the  $(d, b)$  elements plays a  
299 special role, because all three sides are of the same order of magnitude,  $\mathcal{O}(\lambda^3)$ .  
300 When expressed in the form

$$\frac{V_{ud}^* V_{ub}}{V_{cd}^* V_{cb}} + \frac{V_{td}^* V_{tb}}{V_{cd}^* V_{cb}} + 1 = 0, \tag{2.9}$$



**Figure 2.1:** Definition of the lengths and sides of the Unitarity Triangle. Figure is taken from the *CKM Quark-Mixing Matrix* review of the PDG [?].

it is often referred to as the singular Unitarity Triangle, illustrated in Fig. 2.1 where the usual names for the three angles are also given.

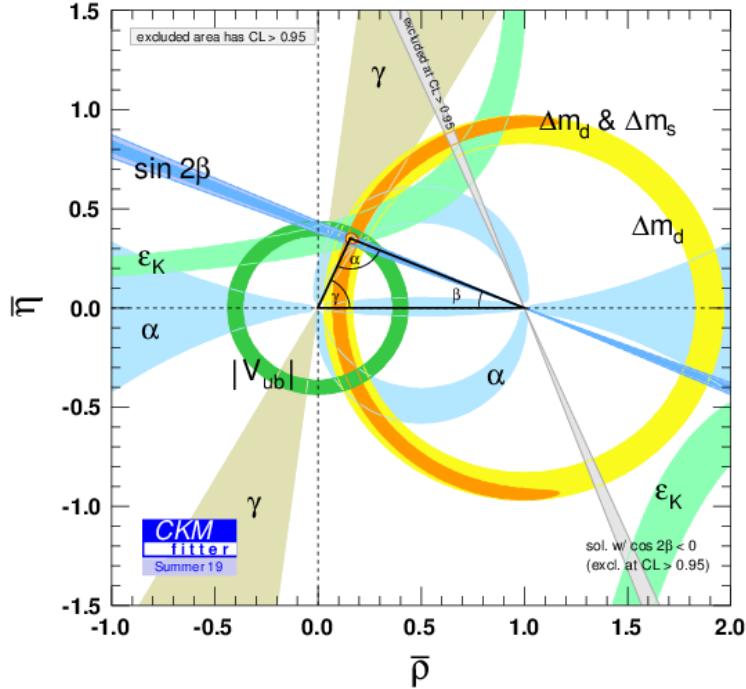
Over-constraining the unitarity triangle by making separate measurements of all sides and angles, in as many different decay channels as possible, is an important, and non-trivial test of the SM. The current experimental constraints are in agreement with the SM predictions, as visualised in Fig. 2.2. The CKM angle

$$\gamma \equiv \arg(-V_{ud} V_{ub}^* / V_{cd} V_{cb}^*) = \arg(-V_{cb} V_{cd}^* / V_{ub} V_{ud}^*) \quad \square \quad (2.10)$$

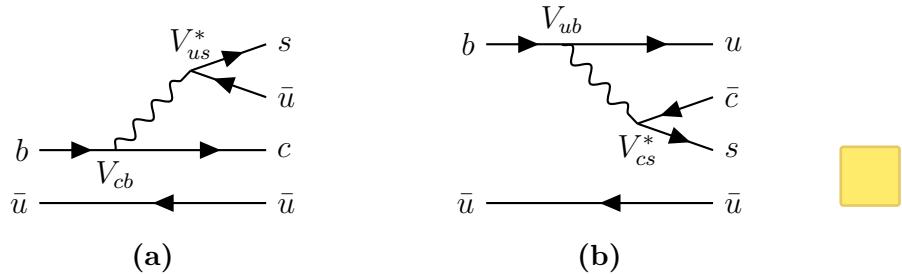
is unique among the CKM parameters, in that it can be measured in tree-level processes without significant theoretical uncertainty from lattice QCD calculations [?]. Because tree-level processes are less likely to be affected by Beyond-Standard-Model (BSM) effects, direct measurements of  $\gamma$  can be considered a SM benchmark, which can be compared to estimates based on measurements of other CKM elements that are measured in loop-level processes, and thus are more likely to be affected by BSM effects [?]. The current, worldwide combination of direct measurements, published by the CKMFitter group, is  $\gamma = (72.1^{+5.4}_{-5.7})^\circ$ , to be compared with the estimate from loop-level observables of  $\gamma = (65.66^{+0.90}_{-2.65})^\circ$  [?]. Other world averages exist [?, ?], but the overall picture is the same: the ability to constrain BSM physics is currently limited by the uncertainty of the direct measurements. Hence further precision measurements of  $\gamma$  are highly motivated. Presently, the precision is driven by time-integrated measurements of direct  $CP$ -violation in  $B^\pm \rightarrow D K^\pm$  decays;

<sup>2</sup>It is generally true that all phases of a single term in a given amplitude will be convention dependent, but that the phase differences between terms are not.

<sup>3</sup>Other variants of the Wolfenstein parameterisation do exist [?]. They all agree at the lowest orders of  $\lambda$ .



**Figure 2.2:** Current constraints on the Unitarity Triangle parameters as determined by the CKMFitter group for the EPS 2019 conference [?].



**Figure 2.3:** Tree level Feynman diagrams describing (a)  $B^- \rightarrow D^0 K^-$  and (b)  $B^- \rightarrow \bar{D}^0 K^-$  decays. The electro-weak phase difference between the two decays is  $\Delta\phi = \arg(V_{cb}V_{us}^*/V_{ub}V_{cs}^*) \simeq \gamma$ .

such a measurement is the topic of this thesis and the theory behind is treated in detail in the following section. It is also possible to measure  $\gamma$  in time-dependent mixing analyses of  $B_s^0 \rightarrow D_s^\mp K^\pm$ ,  $B^0 \rightarrow D^\mp \pi^\pm$  and related decays, by measuring  $CP$  violation in interference between mixing and decay. These modes are expected to provide competitive measurements in the future [?, ?].

### 2.2.2 Measuring $\gamma$ in tree level decays

The phase  $\gamma$  can be measured in tree-level processes with interference between  $b \rightarrow c\bar{s}u$  and  $b \rightarrow \bar{c}s u$  transitions. The canonical example, also the subject of this thesis, is based on measurements sensitive to interference between the  $B^\pm \rightarrow D^0 K^\pm$  and  $B^\pm \rightarrow \bar{D}^0 K^\pm$  decay amplitudes. As illustrated in Fig. 2.3 for the case of  $B^-$  decays, the electro-weak phase difference between the two decays is  $\Delta\phi = \arg(V_{cb}V_{us}^*/V_{ub}V_{cs}^*)$ . While  $\Delta\phi$  is not identical to the definition of  $\gamma$  in Eq. (2.10), the ratio of the involved CKM matrix elements is [?]

$$\begin{aligned} -\frac{V_{cd}^*/V_{ud}^*}{V_{us}^*/V_{cs}^*} &= -\frac{-\lambda[1 - \frac{\lambda^4}{2}A^2(1 - 2(\rho - i\eta))](1 - \frac{\lambda^2}{2} - \frac{\lambda^4}{8}(1 + 4A^2))}{\lambda(1 - \frac{\lambda^2}{2} - \frac{\lambda^4}{4})} \\ &= 1 - \lambda^4 A^2(1 - 2(\rho - i\eta)) + \mathcal{O}(\lambda^5). \end{aligned} \quad (2.11)$$

The ratio equals unity to  $\mathcal{O}(\lambda^4) \simeq 2.6 \times 10^{-3}$ , and thus  $\Delta\phi \simeq \gamma$  is a good approximation within current experimental uncertainties. For the remainder of this thesis the approximation will be used without further comment. The diagrams in Fig. 2.3 describe the leading order contributions to the two amplitudes

$$\begin{aligned} A[B^- \rightarrow D^0 K^-] &\equiv A_B \\ A[B^- \rightarrow \bar{D}^0 K^-] &\equiv \bar{A}_B \equiv r_B A_B e^{i(\delta_B - \gamma)}, \end{aligned} \quad (2.12a)$$

where the last equality introduces two new parameters: the amplitude magnitude ratio  $r_B \equiv |\bar{A}_B|/|A_B|$ , and  $\delta_B$ , the strong-phase difference between the decay amplitudes. Since all  $CP$ -violation is attributed to the electro-weak phase in the SM, the  $CP$ -conjugate decay amplitudes are [?]

$$\begin{aligned} A[B^+ \rightarrow \bar{D}^0 K^+] &= A_B \\ A[B^+ \rightarrow D^0 K^+] &= \bar{A}_B = r_B A_B e^{i(\delta_B + \gamma)}. \end{aligned} \quad (2.12b)$$

In an experimental setting, the  $D^0$  and  $\bar{D}^0$  mesons are reconstructed in some final state,  $f$  [ ] its  $CP$ -conjugate  $\bar{f}$ . In analogy with the  $B^\pm$  decays, the  $D$  decay amplitude can be related<sup>4</sup>

$$\begin{aligned} A[D^0 \rightarrow f] &= A[\bar{D}^0 \rightarrow \bar{f}] = A_D \\ A[\bar{D}^0 \rightarrow f] &= A[D^0 \rightarrow \bar{f}] = r_D A_D e^{i\delta_D}. \end{aligned} \quad (2.13)$$

where the assumption has been made that  $CP$  violation in the  $D$  decays is negligible, and  $\delta_D$  denotes a  $CP$ -conserving strong-phase difference. While  $CP$ -violation in

<sup>4</sup>In this notation  $\delta_D$  is thus phase of the suppressed  $D$ -decay amplitude minus the phase of the favoured  $D$ -decay amplitude. This is the opposite convention to that used in the LHCb measurements with the ADS technique, but aligns with the notation used in the literature on  $\gamma$  measurements in  $D \rightarrow K_S^0 \pi^+ \pi^-$  decays.

<sup>346</sup>  $D$  decays has recently been measured [?], the size of the effect is small and it is  
<sup>347</sup> considered negligible in this thesis. Based on Eqs. 2.12 and (2.13), the decay rates  
<sup>348</sup> of  $B^+$  and  $B^-$  mesons into the possible final states can be seen to satisfy

$$\Gamma(B^- \rightarrow D(\rightarrow f)K^-) \propto 1 + r_D^2 r_B^2 + 2r_B r_D \cos [\delta_B + \delta_D - \gamma], \quad (2.14a)$$

$$\Gamma(B^+ \rightarrow D(\rightarrow \bar{f})K^+) \propto 1 + r_D^2 r_B^2 + 2r_B r_D \cos [\delta_B + \delta_D + \gamma], \quad (2.14b)$$

$$\Gamma(B^- \rightarrow D(\rightarrow \bar{f})K^-) \propto r_D^2 + r_B^2 + 2r_B r_D \cos [\delta_B - \delta_D - \gamma], \quad (2.14c)$$

$$\Gamma(B^+ \rightarrow D(\rightarrow f)K^+) \propto r_D^2 + r_B^2 + 2r_B r_D \cos [\delta_B - \delta_D + \gamma]. \quad (2.14d)$$

<sup>349</sup> The processes in Eqs. (2.14a) and (2.14b) are  $CP$ -conjugate and it is clear how, in the  
<sup>350</sup> general case where  $\delta_B + \delta_D \neq 0$ , a non-zero value of  $\gamma$  leads to  $CP$  violation in the form  
<sup>351</sup> of differing decay rates. The same is true for the processes in Eqs. (2.14c) and (2.14d).  
<sup>352</sup> Depending on the choice of  $D$  final state, these expressions can be used to relate  
<sup>353</sup>  $\gamma$  to various observables that are experimentally accessible. This thesis concerns  
<sup>354</sup> the choice  $f = K_S^0 \pi^+ \pi^-$  or  $f = K_S^0 K^+ K^-$ , where the terms related to the  $D$  decay  
<sup>355</sup> all have a non-trivial variation over the phase space of the decay. However, it is  
<sup>356</sup> useful to first analyse the simpler case where  $f$  is a two-body state.

<sup>357</sup> The simplest case is when  $f$  is chosen to be a  $CP$  eigenstate, so that  $f = \pm \bar{f}$   
  and the rate equations of (2.14a)–(2.14d) simplify, because  $r_D = 1$  and  $\delta_D \in \{0, \pi\}$ .  
<sup>359</sup> Measurements of  $\gamma$  in such decay modes are denoted GLW measurements, after  
<sup>360</sup> Gronau, London, and Wyler who described the approach in the early 90ies [?, ?].  
<sup>361</sup> Experimentally it is preferable to measure yield ratios rather than absolute rates,  
<sup>362</sup> and the observables of interest are thus the  $CP$  asymmetry

$$\begin{aligned} A_{CP=\pm 1} &= \frac{\Gamma[B^- \rightarrow D_{CP} K^-] - \Gamma[B^+ \rightarrow D_{CP} K^+]}{\Gamma[B^- \rightarrow D_{CP} K^-] + \Gamma[B^+ \rightarrow D_{CP} K^+]} \\ &= \frac{\pm r_B \sin \delta_B \sin \gamma}{1 + r_B^2 \pm 2r_B \cos \delta_B \cos \gamma}, \end{aligned} \quad (2.15a)$$

<sup>363</sup> as well as the ratio

$$\begin{aligned} R_{CP=\pm 1} &= 2 \frac{\Gamma[B^- \rightarrow D_{CP} K^-] + \Gamma[B^+ \rightarrow D_{CP} K^+]}{\Gamma[B^- \rightarrow D^0 K^-] + \Gamma[B^+ \rightarrow \bar{D}^0 K^+]} \\ &= 1 + r_B^2 \pm 2r_B \cos \delta_B \cos \gamma. \end{aligned} \quad (2.15b)$$

<sup>364</sup> In practice,  $A_{CP}$  and  $R_{CP}$  are obtained from measured yield ratios that are corrected  
<sup>365</sup> with appropriate branching fractions. A measurement of  $A_{CP}$  and  $R_{CP}$  alone is not  
<sup>366</sup> sufficient to determine the underlying physics parameters  $(\gamma, r_B, \delta_B)$ , and this is  
<sup>367</sup> not solely due to the number of parameters exceeding the number of constraints:  
<sup>368</sup> the equations also allow for multiple, ambiguous solutions for  $(\gamma, \delta_B)$ . One way  
<sup>369</sup> to break the ambiguity, first noted in the original paper [?], is to make further

370 measurements in additional  $B$  decays. These can be described with the formalism  
 371 described above, but will not share the same ambiguous solutions because  $(r_B, \delta_B)$   
 372 are unique to a given  $B$  decay. Another method is to analyse  $D$  decay final states  
 373 that are not  $CP$  eigenstates.

374 A few years [?], Atwood, Dunietz, and Sonis analysed an alternative choice of  
 375  $D$  final states: a simultaneous analysis of a Cabibbo-favoured (CF) decay  $D^0 \rightarrow f$   
 376 and the doubly-Cabibbo-suppressed (DCS) decay  $D^0 \rightarrow \bar{f}$  into the  $CP$  conjugate  
 377 final state [?, ?]. Their suggested method is named the ADS method after the  
 378 authors. The classical example is to take  $f = K^-\pi^+$  and  $\bar{f} = \pi^-K^+$ . The relative  
 379 suppression means that the  $r_D$  of Eq. (2.14) is small, typically of the same order of  
 380 magnitude as  $r_B$ , and thus the  $CP$  asymmetry of the suppressed decay is  $\mathcal{O}(1)$ :

$$\begin{aligned} A_{ADS(\bar{f})} &= \frac{\Gamma[B^- \rightarrow D(\rightarrow \bar{f})K^-] - \Gamma[B^+ \rightarrow D(\rightarrow f)K^+]}{\Gamma[B^- \rightarrow D(\rightarrow \bar{f})K^-] + \Gamma[B^+ \rightarrow D(\rightarrow f)K^+]} \\ &= \frac{r_D r_B \sin(\delta_B - \delta_D) \sin \gamma}{r_D^2 + r_B^2 + 2r_D r_B \cos(\delta_B - \delta_D) \cos \gamma}. \end{aligned} \quad (2.16a)$$

381 The large  $CP$  asymmetry is a prime feature of the ADS method. However, also the  
 382 suppressed-to-favoured yield ratio is sensitive to the physics parameters of interest:

$$\begin{aligned} R_{ADS(\bar{f})} &= \frac{\Gamma[B^- \rightarrow D(\rightarrow \bar{f})K^-] + \Gamma[B^+ \rightarrow D(\rightarrow f)K^+]}{\Gamma[B^- \rightarrow D(\rightarrow f)K^-] + \Gamma[B^+ \rightarrow D(\rightarrow \bar{f})K^+]} \\ &= \frac{r_B^2 + r_D^2 + 2r_D r_B \cos(\delta_B - \delta_D) \cos \gamma}{1 + r_D^2 r_B^2 + 2r_D r_B \cos(\delta_B + \delta_D) \cos \gamma}. \end{aligned} \quad (2.16b)$$

383 The interpretation of  $A_{ADS}$  and  $R_{ADS}$  in terms of  $(\gamma, r_B, \delta_B)$  requires knowledge of  
 384 the  $r_D$  and  $\delta_D$  parameters, but these can be measured independently. In general,  
 385 the constraints from a single set of ADS observables suffer the same ambiguities as  
 386 in the GLW case. However, unlike the GLW case, each  $D$  decay mode provides an  
 387 independent set of constraints, because the parameters related to the  $D$  decay vary.

388 The discussion of this section has centred on the classical case of  $B^\pm \rightarrow DK^\pm$   
 389 decays with a two-body  $D$  final state. With minor modifications the techniques have  
 390 been used to make measurements sensitive to  $\gamma$  in  $B^0$  decays, with  $B$  decay final  
 391 states including excited  $D$  mesons or kaons, and in  $B^\pm \rightarrow D\pi^\pm$  decays (summaries of  
 392 the measurements made by the  $B$  factories and LHCb can be found in Refs. [?, ?, ?, ?]).  
 393 The  $B^\pm \rightarrow D\pi^\pm$  decay is also  $CP$ -violating, although the effect is much smaller  
 394 than in the  $B^\pm \rightarrow DK^\pm$  decay because  $r_B^{D\pi} \simeq 0.005$  [?], whereas  $r_B^{DK} \simeq 0.1$ .  
 395 Furthermore, it is possible to use multi-body  $D$  final states. However, in some  
 396 cases, a better precision can then be obtained by exploiting phase-space dependent  
 397 decay rates. This is the topic of the next section.

## 398 2.3 Measuring $\gamma$ using multi-body D final states

399 In multi-body  $D$  decays, the  $r_D$  and  $\delta_D$  parameters of the fundamental rate equations  
 400 in Eq. (2.14) vary over the phase space of the  $D$  decay. This section describes  
 401 a model-independent approach to measure  $\gamma$  in  $B \rightarrow D(\rightarrow K_S^0\pi^+\pi^-)h^\pm$  decays  
 402 by exploiting this variation. The theory is identical for  $D \rightarrow K_S^0K^+K^-$  decays,  
 403 and similar ideas have been proposed for the  $D \rightarrow K_S^0\pi^+\pi^-\pi^0$  [?] and  $D \rightarrow$   
 404  $2\pi^+2\pi^-$  modes [?]. First, however, the formalism for describing amplitudes of  
 405 multi-body decays is briefly reviewed.

### 406 2.3.1 Dalitz plots and the phase space of multi-body decays

407 In general, the phase space of the  $n$ -body decay  $P \rightarrow p_1 + p_2 + \dots + p_n$  consists of  $n$   
 408 four momenta, with a total of  $4n$  components. The requirement that each of the final  
 409 state particles is on-shell provides  $n$  constraints on these components, and energy-  
 410 momentum conservation removes a further 4 degrees of freedom. If the original  
 411 particle  $P$  is a scalar, the decay is isotropic, which removes an additional 3 degrees  
 412 of freedom, leaving the total number of degrees of freedom at  $3n - 7$ . For the specific  
 413 case of three-body decays, the available phase space can thus be parameterised with  
 414 only two parameters. A practical and often used choice is the invariant masses

$$s_{12} = m^2(p_1 p_2) = (p_1^\mu + p_2^\mu)^2, \quad s_{13} = m^2(p_1 p_3) = (p_1^\mu + p_3^\mu)^2. \quad (2.17)$$

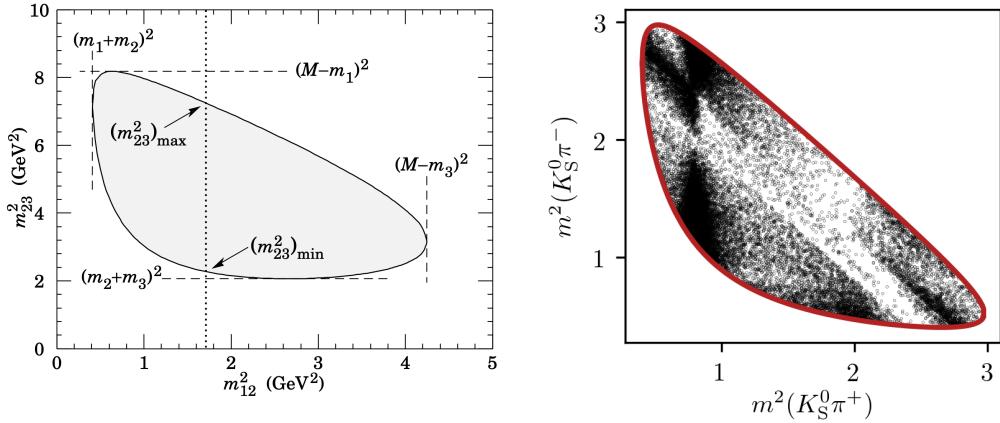
415 The choice of particle pairs is arbitrary, and the coordinates easily related

$$m_P^2 + m_{p_1}^2 + m_{p_2}^2 + m_{p_3}^2 = m^2(p_1 p_2) + m^2(p_1 p_3) + m^2(p_2 p_3). \quad (2.18)$$

416 A scatter plot of  $(s_{12}, s_{13})$  values for a sample of particle decays is denoted a Dalitz  
 417 plot [?]. It has the very useful feature that the presence of (narrow) resonances  
 418 in the decay leads to visible bands in the scatter plot. Figure 2.4 illustrates how  
 419 the limits of the Dalitz plot are defined by kinematic constraints, and shows an  
 420 example of a Dalitz plot for  $D \rightarrow K_S^0\pi^+\pi^-$  decays in which the  $K^*(892)^\pm$  and  $\rho(770)$   
 421 resonances are clearly visible. The plot shows the sample of  $B^+ \rightarrow D\pi^+$  decays  
 422 used to make the measurement described in Chapter 5 and thus the  $D$  meson is in  
 423 a superposition of  $D^0$  and  $\bar{D}^0$  states (as detailed in the following section).

424 In terms of the coordinates of Eq. (2.17) the differential decay rate is given by

$$d\Gamma = \frac{1}{32(2\pi)^3 m_P^3} |\mathcal{M}|^2 ds_{12} ds_{13}, \quad (2.19)$$

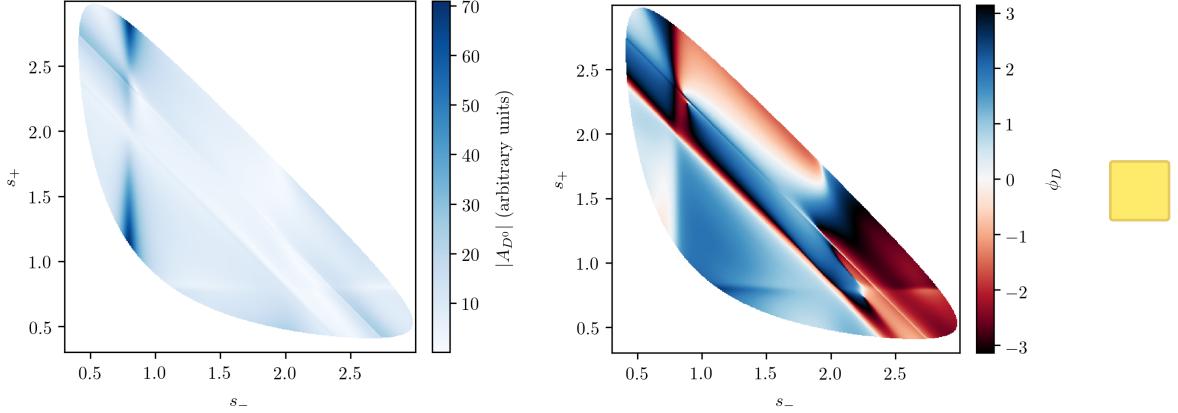


**Figure 2.4:** (Left) Schematic of a Dalitz plot and the limits of the kinematically allowed phase space limits. (Right) Example of a Dalitz plot for  $D \rightarrow K_S^0 \pi^+ \pi^-$  decays where the  $D$  meson originates in a  $B^+ \rightarrow D \pi^+$  decay; the decaying  $D$  meson is in a superposition of the  $D^0$  and  $\bar{D}^0$  states, but predominantly  $\bar{D}^0$ -like.

where  $\mathcal{M}$  is the QFT matrix element, or total decay amplitude, corresponding to the decay. In general, it is not possible to calculate  $\mathcal{M}$  from first principles. Instead, a model is defined with an empirically well motivated form, in which a number of free parameters must be determined experimentally. The simplest case is that of an *isobar* model, where it is assumed that the full decay can be composed into consecutive two-body decays of the form  $P \rightarrow R_{12}(\rightarrow p_1 + p_2)p_3$ . Thus,  $\mathcal{M}$  is expressed as a non-resonant constant amplitude term,  $k_{NR}$ , plus a sum of resonance terms

$$\mathcal{M}(s_{12}, s_{13}) = k_{NR} + \sum_r k_r \mathcal{M}^r(s_{12}, s_{13}). \quad (2.20)$$

The exact form of the  $\mathcal{M}^r$  function depends on the resonance in question. An overview is given in the PDG review on resonances and references therein [?]. The isobar formalism breaks down when resonances in the decay are not well separated. In this case, models of the form in Eq. (4.27) can still be employed, if the contribution from overlapping resonances are collected in a single term. An example of such a model, is the amplitude model for  $D^0 \rightarrow K_S^0 \pi^+ \pi^-$  decays developed by the Belle collaboration for a measurement of the CKM angle  $\beta$  in 2018 [?]. In this model, individual terms are included for  $D^0 \rightarrow K^*(\rightarrow K_S^0 \pi^\pm)\pi^\mp$  decays, whereas the  $\pi\pi$  and  $K\pi$   $S$ -wave contributions are modelled with the so-called  $K$ -matrix- and LASS formalisms [?, ?]. The amplitude and phase of  $\mathcal{M}$  as predicted by this model are shown in Fig. 2.5.



**Figure 2.5:** The (left) magnitude and (right) phase of the  $D \rightarrow K_S^0 \pi^+ \pi^-$  amplitude in the Belle 2018 model [?].

### 2.3.2 The GGSZ method to measure $\gamma$

The non-trivial phase-space dependence of the  $D \rightarrow K_S^0 \pi^+ \pi^-$  decay amplitude can be exploited to measure  $\gamma$  with  $B^\pm \rightarrow DK^\pm$  or  $B^\pm \rightarrow D\pi^\pm$  decays. This approach was proposed independently by Bondar [?], and by Giri, Grossman, Soffer, and Zupan [?] after whom it takes the commonly used acronym GGSZ. For this specific decay  $s_-$  and  $s_+$  are used to describe the Dalitz coordinates  $m^2(K_S^0 \pi^-)$  and  $m^2(K_S^0 \pi^+)$ , respectively, and the  $D$  decay amplitude is a function of these coordinates

$$A_S^{\bar{D}}(s_-, s_+) = A(\bar{D}^0 \rightarrow K_S^0 \pi^+ \pi^-). \quad (2.21)$$

To a good approximation the  $K_S^0$  meson is a  $CP$  eigenstate, meaning that the  $K_S^0 \pi^+ \pi^-$  state is self-conjugate. Assuming this approximation to be exact, and that  $CP$  violation in the  $D$  decay is negligible, the  $D$  decay amplitude satisfies the symmetry relation

$$A_S^{\bar{D}}(s_-, s_+) = A_S^D(s_+, s_-). \quad (2.22)$$

The impact of the  $K_S^0$  meson *not* being an exact  $CP$  eigenstate is treated in detail in Chapter 4. In order to simplify equations, the short-hand notation

$$(s_{-+}) = (s_-, s_+), \quad (s_{+-}) = (s_+, s_-), \quad (2.23)$$

will be employed for the remainder of the thesis, so that the relation in Eq. (2.22) can be expressed as  $A_S^{\bar{D}}(s_{-+}) = A_S^D(s_{+-})$ . Thus, the rate equations of Eq. (2.14)

458 for the  $D \rightarrow K_S^0\pi^+\pi^-$  decay mode are

$$\begin{aligned} d\Gamma^-(s_{-+}) &\propto |\mathcal{A}_S^-|^2 = |A_B|^2 |A_{K_S^0}|^2 \\ &\times [|A_S^D(s_{-+})|^2 + r_B^2 |A_S^D(s_{+-})|^2 + 2r_B |A_S^D(s_{-+})| |A_S^D(s_{+-})| \\ &\times (\cos[\delta_D(s_{-+})] \cos[\delta_B - \gamma] + \sin[\delta_D(s_{-+})] \sin[\delta_B - \gamma])], \end{aligned} \quad (2.24a)$$

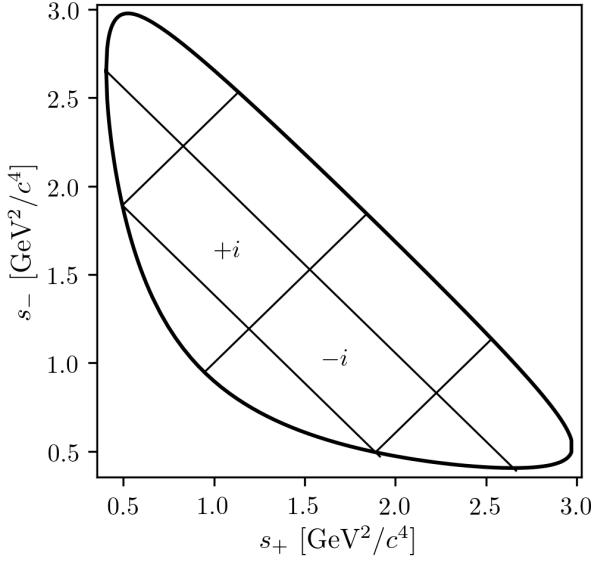
$$\begin{aligned} d\Gamma^+(s_{-+}) &\propto |\mathcal{A}_S^+|^2 = |A_B|^2 |A_{K_S^0}|^2 \\ &\times [|A_S^D(s_{-+})|^2 + r_B^2 |A_S^D(s_{-+})|^2 + 2r_B |A_S^D(s_{-+})| |A_S^D(s_{-+})| \\ &\times (\cos[\delta_D(s_{-+})] \cos[\delta_B + \gamma] - \sin[\delta_D(s_{-+})] \sin[\delta_B + \gamma])]. \end{aligned} \quad (2.24b)$$

459 Here,  $\delta_D(s_{-+}) = \phi_D(s_{-+}) - \phi_D(s_{+-}) = -\delta_D(s_{+-})$ , where  $\phi_D(s_{-+})$  denotes the  
 460 complex phase of the  $A_S^D(s_{-+})$  amplitude, and a standard trigonometric relation  
 461 have been employed to factorise the terms depending on the complex phases of the  $B$   
 462 and  $D$  decays. It can be seen that in the case where  $\gamma = 0$  the  $B^+$  and  $B^-$  decay rates  
 463 are symmetric if the Dalitz coordinates are exchanged:  $\Gamma^+(s_-, s_+) = \Gamma^-(s_+, s_-)$ .  
 464 The presence of  $CP$  violation in the  $B$  decay breaks the symmetry. Therefore it is  
 465 possible to measure  $\gamma$  (and the nuisance parameters  $r_B$  and  $\delta_B$ ) from the phase-space  
 466 distribution of  $B^\pm \rightarrow D(\rightarrow K_S^0\pi^+\pi^-)K^\pm$  decays, given knowledge of  $A_S^D(s_{-+})$ .

467 A series of measurements of  $\gamma$  have been made that use amplitude models of  
 468 the  $D$  decay [?, ?, ?, ?, ?, ?, ?, ?, ?]. However, a model-independent approach has been  
 469 proposed by Bondar and Poluektov [?, ?] that relies on binning phase-space, in  
 470 which case the necessary information on the  $D$  decay amplitude can be summarised  
 471 in a small set of coefficients that can be measured in a separate experiment. That is  
 472 the approach followed in this thesis, and has been used previously by the Belle [?]  
 473 and LHCb collaborations [?]. It is described in detail in the following section.

474 Such a model-independent approach is favourable for two reasons. Firstly,  
 475 uncertainty estimates related to model inputs and the choice of parameterisation in  
 476 an amplitude model are non-trivial, yet would become the leading systematic with  
 477 the very high precision expected for  $\gamma$  measurements in the near future. Secondly,  
 478 amplitude models are notoriously hard to reproduce, and in a high-precision era it  
 479 is favourable that any experiment is easy to reinterpret in various extensions of the  
 480 SM. This is a lot easier for an experiment that measures a small set of well-defined  
 481 observables, than for an experiment that fits a complicated amplitude model.

482 An alternative model-independent approach has recently been proposed by  
 483 Poluektov [?] where the externally measured input on the  $D$ -decay phase are Fourier  
 484 expansion coefficients, and which therefore avoids binning phase space; this approach  
 485 may have the potential to improve the obtainable precision in the future.



**Figure 2.6:** Illustration of the binning scheme used in GGSZ measurements: the bins are symmetric around the  $m^2(K_S^0\pi^+) = m^2(K_S^+\bar{K}_S^-)$  diagonal, and numbered so that opposite bins have the same number, with opposite sign.

### 2.3.3 A model-independent approach

The phase-space distribution can be analysed in a model-independent way, if the  $D$ -decay phase space is split into regions, or bins, the  $B$  decay yield in each bin determined experimentally. A measurement of  $\gamma$  using this approach is the main topic of the thesis. This section describes the fundamental principle, whereas the details pertaining to the exact experimental approach are delegated to Section 2.4.

The amplitude symmetry of Eq. (2.22) is exploited by defining  $2N$  bins to be symmetric symmetric around the  $s_- = s_+$  diagonal of the Dalitz plot, numbered  $i = -N$  to  $N$  (omitting zero) such that if the point  $(s_-, s_+)$  is in bin  $i$ , then  $(s_+, s_-)$  is in bin  $-i$ , and by convention  $i > 0$  for bins where  $s_- > s_+$ . The principle is illustrated in Fig. 2.6, but the binning schemes used in actual measurements are more complicated. The decay rates in Eq. (2.24) can be integrated over such bins, and give the bin yields

$$\begin{aligned} N_i^- &\propto h^- \left[ K_i + r_B^2 K_{-i} + 2\sqrt{K_i K_{-i}} (c_i x_- + s_i y_-) \right], \\ N_i^+ &\propto h^+ \left[ K_{-i} + r_B^2 K_i + 2\sqrt{K_i K_{-i}} (c_i x_+ - s_i y_+) \right], \end{aligned} \quad (2.25)$$

where the parameters describing the  $B$  decay have been expressed in terms of the observables

$$x_\pm = r_B \cos(\delta_B \pm \gamma), \quad y_\pm = r_B \sin(\delta_B \pm \gamma), \quad (2.26)$$

and a number of phase-space integrated quantities related to the  $D$ -decay have been introduced. The  $K_i$  parameters denote fractional yield of a flavour-tagged  $D^0$  decaying into bin  $i$ , defined as

$$K_i = \frac{1}{N_K} \int_i ds^2 |A_S^D(s_{-+})|^2, \quad N_K = \int ds^2 |A_S^D(s_{-+})|^2, \quad (2.27)$$

where  $\int_i ds^2$  denotes integration over bin  $i$  of the Dalitz plot. The  $c_i$  and  $s_i$  denote the amplitude-weighted average of  $\cos \delta_D(s_{-+})$  and  $\sin \delta_D(s_{-+})$  over bin  $i$

$$c_i = \frac{\int_i ds^2 |A_S^D(s_{-+})| |A_S^D(s_{+-})| \cos[\delta_D(s_{-+})]}{\sqrt{\int_i ds^2 |A_S^D(s_{-+})|^2} \sqrt{\int_i ds^2 |A_S^D(s_{+-})|^2}}, \quad s_i = \frac{\int_i ds^2 |A_S^D(s_{-+})| |A_S^D(s_{+-})| \sin[\delta_D(s_{-+})]}{\sqrt{\int_i ds^2 |A_S^D(s_{-+})|^2} \sqrt{\int_i ds^2 |A_S^D(s_{+-})|^2}}. \quad (2.28)$$

By the symmetry properties of  $\delta_D(s_{-+})$  these parameters satisfy  $c_i = c_{-i}$  and  $s_i = -s_{-i}$ . The normalisation constants  $h^+$  and  $h^-$  are identical in the ideal case, but it is convenient to define them separately for practical reasons: depending on the experimental setup, there may be overall production and detection asymmetries that affect the total signal yields. An experimental analysis can be made insensitive to these effects because they can be absorbed into the normalisation constants, as long as they are constant over the  $D$ -decay phase space. This comes at the cost that the information on  $x_\pm$  and  $y_\pm$  from the overall  $CP$  asymmetry is lost, but Section 2.3.5 will show the loss in precision to be minimal.

Thus, for a set of bins, the bin yields of Eqs. (2.25) provide  $4N$  constraints on a total of  $4 + 6$  parameters:  $(h^\pm, K_i, c_i, s_i, x_\pm, y_\pm)$ . However, the  $K_i$ ,  $c_i$ , and  $s_i$  parameters relate only to the  $D$  decay, and can thus, in principle, be measured in independent experiments. With such external inputs, a measurement of the  $B^\pm \rightarrow D(\rightarrow K_S^0 \pi^+ \pi^-) K^\pm$  yields in a set of bins can be used to constrain  $x_\pm$  and  $y_\pm$ , and thereby  $(\gamma, r_B, \delta_B)$ . The measurement presented in this thesis determines the  $K_i$  parameters directly, but uses externally measured values of  $c_i$  and  $s_i$  as input, as measured in quantum correlated  $D$  decays by the CLEO [?] and BESIII [?] collaborations. Because these measurements are the foundation of the approach, they are described in some detail in the following section. In the future, it is possible that the  $c_i$  and  $s_i$  parameters may be measured in quantum-correlated  $D$  decays in LHCb [?], and in charm-mixing measurements [?].

### 2.3.4 Measuring strong-phase inputs at charm factories

The strong-phase parameters  $c_i$  and  $s_i$  have been measured by the CLEO and BESIII collaborations, using quantum correlated  $D^0\bar{D}^0$  pairs from decays of the  $\psi(3770)$  resonance state, itself produced in  $e^+e^-$  collisions at the resonance energy. The  $\psi(3770)$  has quantum-number  $C = -1$ , which is conserved in the strong decay into two  $D$  mesons, and thus the two  $D$  mesons are produced in an anti-symmetric wave function. By observing the decay of one  $D$  meson into a specific final state, say a  $CP$  eigenstate, the quantum state of the other  $D$  meson can be determined. The measurement is based on decays where both  $D$  decays are reconstructed, one in the  $K_S^0\pi^+\pi^-$  final state, the other in one of several different tag categories. The main principles are outlined below, but most experimental considerations and implementation details are left out for the sake of brevity.

The simplest case is when one  $D$  meson decays into a final state that uniquely tags the flavour, such as  $\bar{D}^0 \rightarrow K^+e^-\bar{\nu}_e$ . In that case, the  $D$  meson decaying to  $K_S^0\pi^+\pi^-$  is known to be in the  $D^0$  state and the decay rate is simply determined by  $A_S^D : \Gamma(s_{-+}) \propto |A_S^D(s_{-+})|^2$ . This allows for a measurement of the  $K_i$  parameters.

If one  $D$  meson is reconstructed in a  $CP$ -even state, eg.  $K^+K^-$ , or a  $CP$ -odd state, eg.  $K_S^0\pi^0$ , the  $D$  meson decaying to  $K_S^0\pi^+\pi^-$  is known to be in a state of opposite  $CP$ . Thus, for a tag-decay of  $CP = \pm 1$  the decay rate has the form

$$\Gamma_{CP=\pm 1} \propto |A_S^D(s_{-+}) \mp A_S^D(s_{+-})|^2 \quad (2.29a)$$

and the bin yields will be given by

$$M_i^\pm \propto K_i + K_{-i} \mp 2\sqrt{K_i K_{-i}} c_i. \quad (2.29b)$$

Thus a simultaneous analysis of flavour and  $CP$  tagged decays allow for a determination of the  $K_i$  and  $c_i$  parameter sets.

Finally, the case where both  $D$  mesons, for now denoted  $D$  and  $D'$ , decay into the  $K_S^0\pi\pi$  final state can be considered. The total amplitudes have contributions from the case where  $D$  is in the  $D^0$  state and  $D'$  is in the  $\bar{D}^0$  state, as well as the opposite flavour assignment. Thus the decay rate satisfies

$$\Gamma_{CP=\pm 1} \propto |A_S^D(s_{-+})A_S^{D'}(s'_{+-}) + A_S^D(s_{+-})A_S^{D'}(s'_{-+})|^2 \quad (2.30a)$$

where  $s_{-+}$  denotes the Dalitz-plot coordinates of the  $D$  meson, and  $s'_{-+}$  those of the  $D'$  meson. Defining  $M_{ij}$  to be the yield of decays where the  $D$  decay is in bin  $i$  and the  $D'$  in bin  $j$ , the bin yields satisfy

$$M_{ij} \propto K_i K_{-j} + K_j K_{-i} - 2\sqrt{K_i K_{-i} K_j K_{-j}} (c_i c_j + s_i s_j). \quad (2.30b)$$

556 Thus, analysing these decays in addition to the  $CP$  and flavour tagged decays provide  
 557 information on all of  $K_i$ ,  $c_i$ , and  $s_i$ . Note, however, that Eqs. (2.29) and (2.30) are  
 558 invariant under the transformation  $\delta_D \rightarrow -\delta_D$ . In practice, the analysis is extended  
 559 in a number of ways to enhance the statistics: using "flavour-tag" states that are not  
 560 exact flavour tags, such as  $K^-\pi^+$ , using self-conjugate multi-body  $D$ -decay final  
 561 states that are not exact  $CP$  eigenstates, such as  $\pi^+\pi^-\pi^0$ , and using the  $K_L^0\pi^+\pi^-$   
 562 final state as well. However, the main principles are the same as described above.

563 The measurements of  $c_i$  and  $s_i$  are made for a range of different binning schemes.  
 564 It was noted already in Ref. [?] that a rectangular binning scheme, such as the  
 565 example in Fig. 2.6, does not provide the optimal sensitivity to  $\gamma$ . A better sensitivity  
 566 can be obtained if the bins are defined such that  $\delta_D$  is approximately constant over  
 567 a given bin, by defining bin  $i$  out of  $N$  via the condition

$$\text{bin}_i = \{(s_-, s_+) \mid 2\pi(i - 3/2)/N < \delta_D(s_-, s_+) < 2\pi \times (i - 1/2)/N\}. \quad (2.31)$$

568 In practice, the binning scheme is defined by splitting the  $D$ -decay phase-space  
 569 into quadratic *micro bins* with a width of  $0.0054 \text{ (GeV}/c^2)^2$  and assigning a bin  
 570 number to each micro bin via the condition in (2.31) as evaluated in an amplitude  
 571 model of choice. The obtained binning scheme when using an amplitude model  
 572 developed by the BaBar collaboration in 2008 [?] is shown in Fig. 2.7a. In Ref [?]  
 573 it was also shown that the binning can be even further optimised for sensitivity.  
 574 The suggested figure of merit is

$$Q^2 = \frac{\sum_i \left( \frac{1}{\sqrt{N_i^B}} \frac{dN_i^B}{dx} \right)^2 + \left( \frac{1}{\sqrt{N_i^B}} \frac{dN_i^B}{dy} \right)^2}{\int ds^2 \left[ \left( \frac{1}{|\Gamma^B(s_-)|} \frac{d|\Gamma^B(s_-)|^2}{dx} \right)^2 + \left( \frac{1}{|\Gamma^B(s_-)|} \frac{d|\Gamma^B(s_-)|^2}{dy} \right)^2 \right]} \quad (2.32)$$

575 which quantifies the statistical sensitivity for a given binning, relative to the one  
 576 achievable in an unbinned analysis. The CLEO collaboration defined an *optimal*  
 577 binning scheme by an iterative procedure where, starting from the equal binning  
 578 scheme, a micro-bin is randomly reassigned new bin numbers in each step, and a  
 579 step accepted if  $Q^2$  increases. The optimisation is done for the case where  $x = y = 0$   
 580 and thus  $Q^2$  simplifies to  $Q_{x=y=0}^2 = \sum_i N_i^{x=y=0} (c_i^2 + s_i^2) / N_{total}^{x=y=0}$ . The resulting  
 581 binning scheme is shown in Fig. 2.7b. An additional binning scheme is defined,  
 582 denoted the *modified optimal* scheme and shown in Fig. 2.7c, where the  $Q^2$  figure  
 583 of merit is modified to take into account the presence of backgrounds [?]. The  
 584 modified optimal binning scheme has proven beneficial to use in measurements with  
 585 small signal yields [?], but is not employed in the present thesis.

**Table 2.1:** The experimentally measured  $c_i$  and  $s_i$  values used in the thesis. The  $D \rightarrow K_S^0\pi^+\pi^-$  values are the combined values from the BESIII and CLEO measurements published by BESIII [?]. The  $D \rightarrow K_S^0K^+K^-$  values are measured by CLEO [?].

Optimal binning scheme: $D \rightarrow K_S^0\pi^+\pi^-$		
Bin $i$	$c_i$	$s_i$
1	$-0.037 \pm 0.049$	$0.829 \pm 0.097$
2	$0.837 \pm 0.067$	$0.286 \pm 0.152$
3	$0.147 \pm 0.066$	$0.786 \pm 0.154$
4	$-0.905 \pm 0.021$	$0.079 \pm 0.059$
5	$-0.291 \pm 0.041$	$-1.022 \pm 0.062$
6	$0.272 \pm 0.082$	$-0.977 \pm 0.176$
7	$0.918 \pm 0.017$	$-0.184 \pm 0.065$
8	$0.773 \pm 0.033$	$0.277 \pm 0.118$

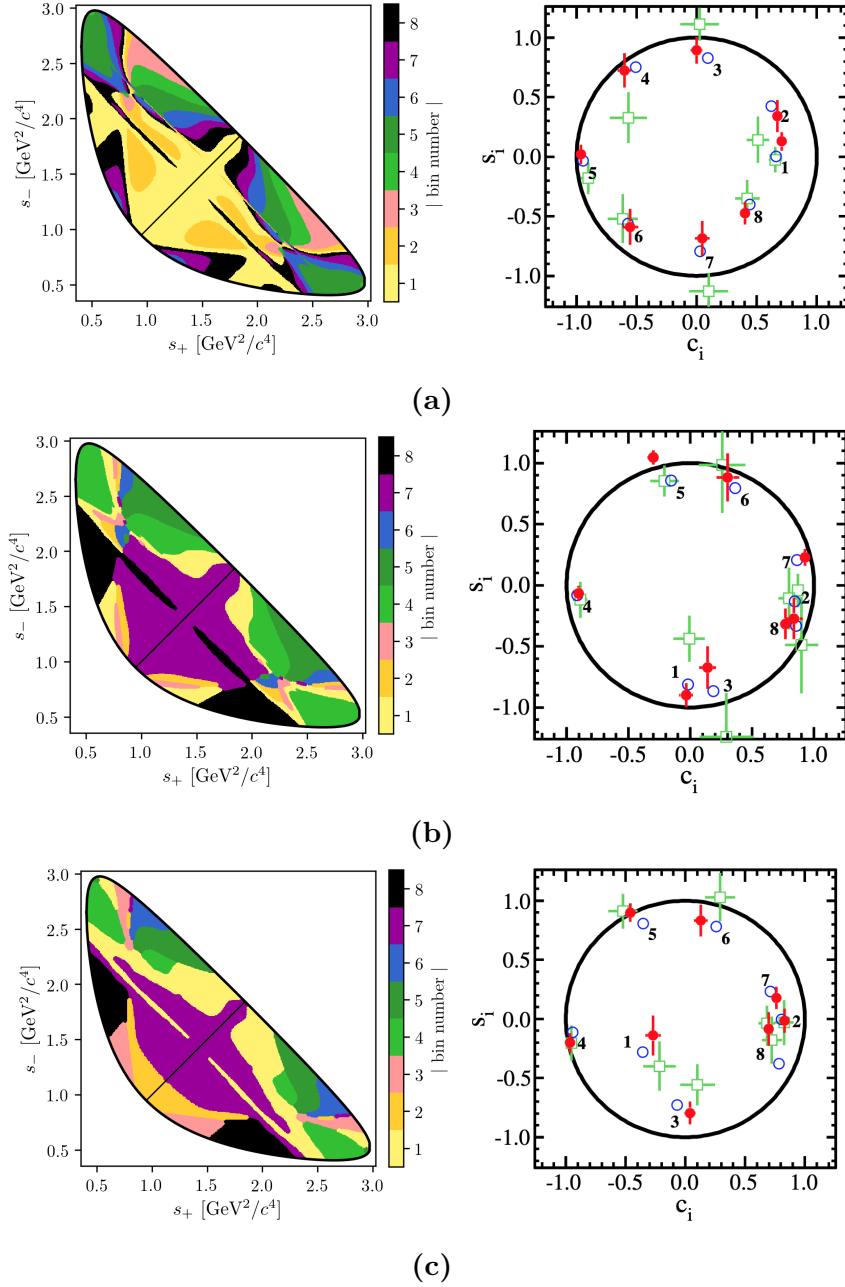
2-bins binning scheme: $D \rightarrow K_S^0K^+K^-$		
Bin $i$	$c_i$	$s_i$
1	$0.818 \pm 0.107$	$-0.445 \pm 0.215$
2	$-0.746 \pm 0.083$	$-0.229 \pm 0.220$

Both the CLEO and BESIII collaborations have measured the values of  $c_i$  and  $s_i$  for the equal, optimal, and modified optimal binning schemes. The results are also shown in Fig. 2.7, where they are compared to the expectation from the latest amplitude model [?]. The measurements presented in this thesis are based on a combination of the BESIII and CLEO results for the optimal binning scheme, made by the BESIII collaboration [?] and tabulated in Table 2.1.

While the *definition* and *optimisation* of these binning schemes depend on knowledge of  $A_S^D(s_-, s_+)$  via an amplitude model, it is important to note that no model information is needed when the binning schemes are used in the subsequent measurements of strong-phases<sup>5</sup> or  $CP$ -observables. Therefore the measurements will not be biased by any modelling imperfections, although the obtained precision might be lower than expected.

The preceding discussion has been focusing on the  $D \rightarrow K_S^0\pi^+\pi^-$  channel, however the  $D \rightarrow K_S^0K^+K^-$  channel can be analysed completely analogously. The CLEO collaboration measure  $c_i$  and  $s_i$  values for this mode as well, in three binning schemes [?]. These are all equal-phase binning schemes, with 2, 3, and 4 bins,

<sup>5</sup>With the exception of minimal model-dependence introduced when the  $K_L^0\pi^+\pi^-$  final state is employed to constrain the  $s_i$  parameters by the  $D$ -factories [?, ?], the impact of which is well under control.



**Figure 2.7:** The (left) binning schemes and (right) measured values of  $(c_i, s_i)$  for (a) equal, (b) optimal, and (c) modified optimal binning schemes for  $D \rightarrow K_S^0 \pi^+ \pi^-$  decays. The plots of the measured values are taken from Ref. [?] and show the results obtained by (red) BESIII, (green) CLEO, and (blue) the model expectation using the model from Ref. [?]. The measurement featured in this thesis used the optimal binning scheme.

respectively, shown in Fig. 2.8. The  $D \rightarrow K_S^0 K^+ K^-$  decay amplitude is almost completely dominated by two  $K^+ K^-$  resonances, the  $CP$ -odd  $\phi(1020)$  and the  $CP$ -even  $a_0(980)$ , and this means that very little gain in sensitivity can be made by altering the equal-phase binning schemes. The measured  $c_i$  and  $s_i$  values are shown in Fig. 2.8 tabulated in Table 2.1 for the 2-bins scheme, which is used in this thesis. A BESIII measurement is in preparation, but has not been finished at the time of writing.

### 2.3.5 Global CP asymmetry and the relation to GLW and ADS measurements

The introduction of separate normalisation factors  $h^+$  and  $h^-$  in Eq. (2.25) hides the fact that information on  $\gamma$  (in principle) can be obtained from the asymmetry in phase-space-integrated  $B^+$  and  $B^-$  yields. In the ideal case where  $h^- = h^+$  the total yield asymmetry is

$$\begin{aligned} A_{GGSZ} &= \frac{\sum_i N_- - N_i^+}{\sum_{i=-N}^N N_i - N_i^+} = \frac{\sum_{i=-N}^N \sqrt{K_i K_{-i}} c_i (x_- - x_+)}{1 + r_B^2 + 2 \sum_{i=-N}^N \sqrt{K_i K_{-i}} c_i (x_- + x_+)} \\ &= \frac{2 \sum_{i=1}^N \sqrt{K_i K_{-i}} c_i (x_- - x_+)}{1 + r_B^2 + 4 \sum_{i=1}^N \sqrt{K_i K_{-i}} c_i (x_- + x_+)}, \end{aligned} \quad (2.33)$$

where it has been exploited that  $\sum_{i=-N}^N \sqrt{K_i K_{-i}} s_i = 0$  by definition. The size of the asymmetry is governed by the factor  $\sum_{i=1}^N \sqrt{K_i K_{-i}} c_i$ , which is small for  $D \rightarrow K_S^0 \pi^+ \pi^-$  and  $D \rightarrow K_S^0 K^+ K^-$  decays. The underlying reason is that  $\delta_D(s_-, s_+)$  varies significantly across phase-space for these decays, as evident by the spread in the values of  $c_i$  in Table 2.1, which reduces the *average* of the asymmetry-generating  $D^0 - \bar{D}^0$  interference term to being close to zero. The value of  $\sum_{i=-N}^N \sqrt{K_i K_{-i}} c_i$  is closely related to the  $CP$  content of the final state in question: for a self-conjugate  $CP$  even (odd) final state

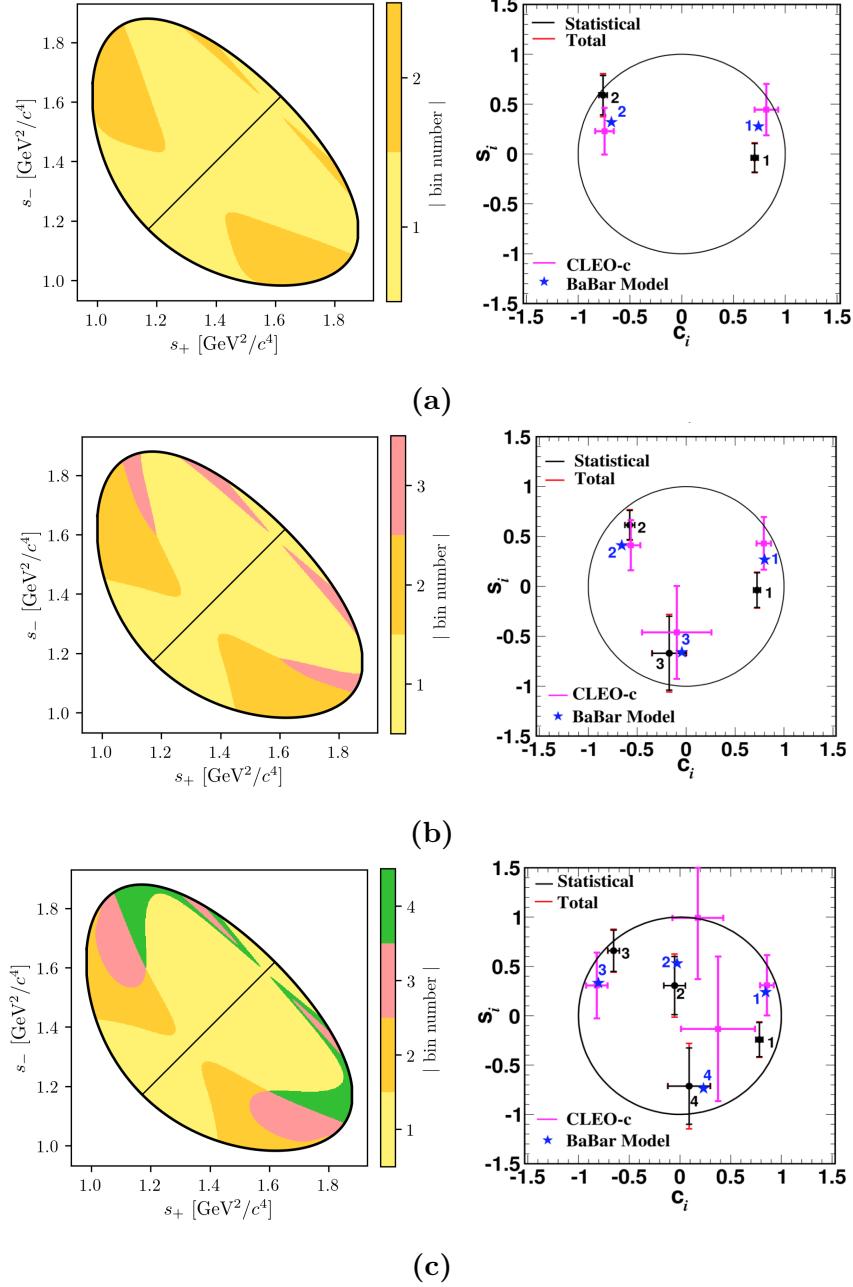
$$A_{D^0}(s_-, s_+) = {}^{(+)}_{(-)} A_{\bar{D}^0}(s_-, s_+) = {}^{(+)}_{(-)} A_{D^0}(s_+, s_-) \quad (2.34)$$

and thus  $\sum_{i=1}^N \sqrt{K_i K_{-i}} c_i = {}^{(+)}_{(-)} 1$ . This motivates the definition of the  $CP$ -even fraction of the decay

$$\mathcal{F}_+ \equiv \frac{1}{2} \left( 1 + \sum_{i=1}^N \sqrt{K_i K_{-i}} c_i \right), \quad (2.35)$$

is equivalent to the definition in Ref. [?] for the case  $N = 1$ . With  $\mathcal{F}_+$  in hand, the asymmetry in Eq. (2.33) can be rewritten

$$A_{GGSZ} = \frac{(2\mathcal{F}_+ - 1)r_B \sin \delta_B \sin \gamma}{1 + r_B^2(2\mathcal{F}_+ - 1)2r_B \cos \delta_B \cos \gamma}, \quad (2.36)$$



**Figure 2.8:** The (left) binning schemes and (right) measured values of  $(c_i, s_i)$  for the (a) 2-, (b) 3-, and (c) 4-bins binning schemes for  $D \rightarrow K_S^0 K^+ K^-$  decays. The plots of the measured values are taken from Ref. [?] and show the (error bars) results obtained by (black) BESIII, (pink) CLEO, and (blue) the model expectation using the model from Ref. [?]. The measurement featured in this thesis uses the 2-bins scheme.

which is the usual form used in quasi-GLW measurements [?, ?]. The value of  $\mathcal{F}_+$  is independent of the number and shape of bins in a given binning scheme, as long as the bin definitions follow the symmetry principles outlined in Section 2.3.3. For  $D \rightarrow K_S^0\pi^+\pi^-$  and  $D \rightarrow K_S^0K^+K^-$  decays the values of  $\mathcal{F}_+$  are

$$\begin{aligned}\mathcal{F}_+(K_S^0\pi^+\pi^-) &= X? \\ \mathcal{F}_+(K_S^0K^+K^-) &= X?\end{aligned}\tag{2.37}$$

as evaluated with the Belle 2018 model for  $D \rightarrow K_S^0\pi^+\pi^-$  decays and the BaBar 2010 model for  $D \rightarrow K_S^0K^+K^-$  decays. Since  $r_B^{DK^\pm} \sim 0.1$  the predicted global asymmetries are thus approximately 1–2 %, which is not resolvable with the current experimental yields. As shown in Chapter 4,  $CP$  violation in the  $K_S^0$  sector leads to asymmetries of a similar size, further complicating the use of global asymmetries to constrain  $x_\pm$  and  $y_\pm$ . Thus these modes are ill-suited for quasi-GLW measurements, and ignoring global asymmetries leads to a negligible loss of information on  $\gamma$  in a GGSZ measurement. The reverse is true for a well-suited quasi-GLW mode, such as  $D \rightarrow \pi^+\pi^-\pi^0$ : if  $\mathcal{F}_+$  is close to either zero or unity, it means that  $(c_i, s_i)$  will be close to  $(\pm 1, 0)$  in all bins for *any* given binning scheme, and the set of bins will provide almost identical constraints on  $x_\pm$  and  $y_\pm$ . Thus, the binning of phase space leads to no significant gain in precision compared to a global analysis.

Indeed, a crucial quality of the GGSZ method, is that exactly because each bin-pair provides independent constraints on  $x_\pm$  and  $y_\pm$ , the method provides a single solution for  $(\gamma, r_B, \delta_B)$  that does not suffer the ambiguities of the ADS and GLW approaches. In order to illustrate this further, it is useful to make one more comparison of the model-independent GGSZ formalism to the ADS and GLW formalisms. If there was no  $CP$  symmetry the  $B^+$  yield in bin  $+i$  would equal the  $B^-$  yield in bin  $-i$ . Therefore the relevant  $CP$  asymmetry for a given Dalitz bin is

$$\begin{aligned}A_{GGSZ}^i &\equiv \frac{N_i^- - N_{-i}^+}{N_i^- + N_{-i}^+} \\ &= \frac{\sqrt{K_i K_{-i}}(c_i(x_- - x_+) + s_i(y_- - y_+))}{K_i + r_B^2 K_{-i} + 2\sqrt{K_i K_{-i}}(c_i(x_- + x_+) + s_i(y_- + y_+)})\end{aligned}\tag{2.38}$$

This expression is identical to the ADS asymmetry in Eq. (2.16a) if the effective  $D$ -decay parameters  $r_D^i$  and  $\delta_D^i$  are defined via

$$\kappa_i \cos \delta_D^i \equiv c_i \quad , \quad \kappa_i \sin \delta_D^i \equiv s_i \quad , \quad r_D^i \equiv \sqrt{K_i / K_{-i}},\tag{2.39}$$

and a coherence factor,  $\kappa$ , is included in the interference terms of the ADS expression, as is standard for multi-body  $D$  decays [?]. These parameters allow us to classify

**Table 2.2:** Classification of the bins used in model-independent GGSZ measurements, in terms of whether the interplay between the  $D^0$  and  $\bar{D}^0$  amplitudes in the bin resemble typical GLW or ADS behaviour. The parameters are calculated using the 2018 Belle model [?] for  $D \rightarrow K_S^0\pi^+\pi^-$  decays and the 2010 BaBar model [?] for  $D \rightarrow K_S^0K^+K^-$  decays.

Optimal binning scheme: $D \rightarrow K_S^0\pi^+\pi^-$					
Bin $i$	$\hat{r}_D$	$\hat{\delta}_D$	$\mathcal{F}_+$	$\kappa$	Bin type
1	0.473	91.9°	48.97 %	0.81	Odd-even
2	0.164	11.1°	63.38 %	0.85	ADS-like
3	0.157	79.4°	52.50 %	0.89	ADS-like
4	0.768	175.3°	5.85 %	0.92	GLW-odd-like
5	0.759	-99.9°	42.84 %	0.87	Odd-even
6	0.223	-64.5°	57.92 %	0.87	ADS-like
7	0.651	-13.3°	89.44 %	0.89	GLW-even-like
8	1.745	21.0°	87.08 %	0.92	GLW-even-like

2-bins binning scheme: $D \rightarrow K_S^0K^+K^-$					
Bin $i$	$\hat{r}_D$	$\hat{\delta}_D$	$\mathcal{F}_+$	$\kappa$	Bin type
1	0.816	19.8°	86.14 %	0.78	GLW-even-like
2	0.775	154.5°	16.23 %	0.77	GLW-odd-like

654 a given pair of bins with number  $\pm i$  as either *GLW-like*, if  $\delta_D^i$  is close to 0 or  $\pi$   
 655 and  $r_D^i$  is close to unity, or *ADS-like* if  $0 < r_D^i < 1$ . The *CP*-even fraction of the  
 656  $D$ -decay can also be defined for a given bin-pair:

$$\mathcal{F}_+^i = \mathcal{F}_+^{-i} \equiv \frac{1}{2} \left( 1 + 2c_i \frac{\sqrt{K_i K_{-i}}}{K_i + K_{-i}} \right) = \frac{1}{2} \left( 1 + 2c_i \frac{r_D^i}{1 + r_D^i} \right). \quad (2.40)$$

657 A GLW-even-like bin pair will have  $\mathcal{F}_+^i \simeq 1$  and a GLW-odd-like bin pair will  
 658 have  $\mathcal{F}_+^i \simeq 0$ .

659 Table 2.2 summarises a classification of the bins for the optimal  $D \rightarrow K_S^0\pi^+\pi^-$   
 660 binning scheme and the 2-bins  $D \rightarrow K_S^0K^+K^-$  binning scheme following these  
 661 principles. Two bins are classified as *Odd-even*; in these bins,  $r_D^i$  is not particularly  
 662 small but  $\mathcal{F}_+^i$  is close to 0.5. The name refers to the fact that for these  
 663 bins  $A_{GGSZ}^i$ , as defined in Eq (2.38), will be positive and  $A_{GGSZ}^{-i}$  negative (or vice  
 664 versa). The fact that multiple bin types appear for both the  $D \rightarrow K_S^0\pi^+\pi^-$  and  
 665  $D \rightarrow K_S^0K^+K^-$  modes underline that each mode benefits from being analysed in  
 666 the GGSZ formalism, and that the bins provide independent constraints, allowing  
 667 for a non-ambiguous solution for  $(\gamma, r_B, \delta_B)$ .

## 2.4 Strategy for the LHCb measurement

The main topic of the thesis is a model-independent GGSZ measurement using  $B^\pm \rightarrow DK^\pm$  and  $B^\pm \rightarrow D\pi^\pm$  decays, and the two  $D$  final states  $K_S^0\pi^+\pi^-$  and  $K_S^0K^+K^-$ . The measurement uses the optimal binning scheme for the  $D \rightarrow K_S^0\pi^+\pi^-$  mode, with the combined strong-phase inputs from the BESIII [?] and CLEO [?] collaborations published in Ref. [?]. For the  $D \rightarrow K_S^0K^+K^-$  channel, the 2-bins scheme is used with the strong-phase parameters measured by the CLEO collaboration [?]. The details of the analysis are presented in Chapter (5), but the overall strategy and a few extensions of the formalism from the previous sections are given here.

Due to the geometry of the LHCb detector, the signal reconstruction efficiency for  $B^\pm \rightarrow D(\rightarrow K_S^0h^+h^-)h'^\pm$  decays varies significantly across the  $D$ -decay phase space. Denoting the efficiency profile as  $\eta(s_-, s_+)$ , the yield equations of Eq. (2.25) are therefore modified slightly

$$\begin{aligned} N_i^- &= h^{B^-} \left[ F_i + r_B^2 F_{-i} + 2\sqrt{F_i F_{-i}} (c'_i x_- + s'_i y_-) \right], \\ N_i^+ &= h^{B^+} \left[ F_{-i} + r_B^2 F_i + 2\sqrt{F_i F_{-i}} (c'_i x_+ - s'_i y_+) \right], \end{aligned} \quad (2.41)$$

where the phase-space integrated quantities now include the efficiency profile

$$F_i = \frac{1}{N_F} \int ds^2 \eta(s_{-+}) |A_S^D(s_{-+})|^2, \quad N_F = \int ds^2 \eta(s_{-+}) |A_S^D(s_{-+})|^2, \quad (2.42)$$

with an analogous definition of  $s'_i$ . At leading order, the strong-phase parameters are unaffected by the non-uniform efficiency, and, in addition, the bin definitions favour bins for which  $\cos[\delta_D(s_{-+})]$  and  $\sin[\delta_D(s_{-+})]$  take on similar values across each bin. Therefore, the  $c_i$  and  $s_i$  values reported by the charm factories are used directly in the measurement. The impact on the obtained central values is negligible, as described in detail in Section 5.6 where a systematic uncertainty is assigned.

The  $F_i$  are significantly different to the  $K_i$  due to the experimental acceptance profile in LHCb. Given external inputs for the strong-phase parameters, it is possible to fit the  $F_i$  parameters and  $x_\pm$  and  $y_\pm$  simultaneously in a fit to the LHCb  $B^\pm \rightarrow DK^\pm$  data set, in which case the obtained  $F_i$  parameters incorporate the correct acceptance profile correction by construction. However, the obtainable precision for the  $CP$  observables measured by this procedure is suboptimal. As

an alternative, the first LHCb measurement [?] made a simultaneous analysis of  $B^\pm \rightarrow DK^\pm$  and a much larger sample of  $B^\pm \rightarrow D\pi^\pm$  decays; since the  $F_i$  parameters relate to the  $D$  decay, they can effectively be obtained in the  $D\pi^\pm$  sample and shared between the two  $B^\pm \rightarrow Dh^\pm$  channels. However, there is  $CP$  violation present in the  $B^\pm \rightarrow D\pi^\pm$  decays, which led to a dominant systematic uncertainty. Later LHCb measurements [?, ?] instead relied on flavour tagged  $D$  mesons from  $\bar{B}^0 \rightarrow D^{*+}(\rightarrow D^0\pi^+)\mu^-\bar{\nu}_\mu X$  decays to obtain  $F_i$ , where no  $CP$  violation is possible. However, due to necessarily different triggering paths and selections, the acceptance profile is not exactly identical between semi-leptonic decays and the  $B^\pm \rightarrow Dh^\pm$  decays of interest. An efficiency correction based on simulation was therefore applied to obtain the correct  $F_i$ , and in this case, the uncertainty related to the correction constituted the largest systematic uncertainty on the measurement.

Both sources of systematic uncertainty can be avoided by making a simultaneous analysis of  $B^\pm \rightarrow DK^\pm$  and  $B^\pm \rightarrow D\pi^\pm$  decays, where  $CP$ -violating observables are measured in *both* channels and the  $F_i$  parameters are shared. It is a reasonable assumption that  $F_i^{DK} = F_i^{D\pi}$  to a very good approximation, given the similar kinematics of the decays. The assumption is confirmed using simulated decays in Section 5.2.2, for the candidate selection used in the measurement of the thesis. Effectively, the  $F_i$  are determined in the high statistics  $B^\pm \rightarrow D\pi^\pm$  channel, but with no systematic effect from  $CP$ -violation in that channel, since the  $CP$ -violation is incorporated in the yield description.

ADD SECTION ON INFORMATION ON GAMMA FROM DPI AND THE TWO SOLUTIONS IN THE LHCb COMBINATION?!

At the start of the work that lead to this thesis, it was not clear to what degree the measured  $CP$ -violating observables in  $B^\pm \rightarrow D\pi^\pm$  decays were affected by  $CP$  violation in the neutral kaon sector. The impact had been shown to scale as  $\mathcal{O}(|\epsilon|/r_B)$  [?], which is negligible for the  $B^\pm \rightarrow DK^\pm$  channel but suggests potentially large biases in the  $B^\pm \rightarrow D\pi^\pm$  channel, where  $r_B$  is 20 times smaller. However, the dedicated analysis presented in Chapter 4 has proved the effect on GGSZ measurements to be in fact be *smaller* than  $\mathcal{O}(|\epsilon|/r_B)$  and the simultaneous measurement is indeed viable.

The measurement is performed by making extended maximum-likelihood fits to the  $m_B$  spectra of  $B \rightarrow D(\rightarrow K_S^0 h^+ h^-)h'^\pm$  candidates split by charge and Dalitz bin. The  $B^\pm \rightarrow DK^\pm$  signal yields are parameterised using the expressions in Eq. (2.41) directly, thus obtaining values for  $x_\pm^{DK}$  and  $y_\pm^{DK}$  directly. The Cartesian  $CP$ -violating observables  $x_\pm$  and  $y_\pm$  are employed because they lead to better statistical behaviour than fits to data where the underlying parameters  $(\gamma, r_B^{DK^\pm}, \delta_B^{DK^\pm})$  are

determined, at the cost of introducing a fourth degree of freedom. With the addition of the  $B^\pm \rightarrow D\pi^\pm$  mode as a true signal channel, two new underlying parameters are introduced,  $r_B^{D\pi^\pm}$  and  $\delta_B^{D\pi^\pm}$ . There is a choice to be made, in terms of how to define the observables that are measured directly. One is to introduce an additional set of four observables,  $(x_-^{D\pi}, y_-^{D\pi}, x_+^{D\pi}, y_+^{D\pi})$ , that are analogous to the  $B^\pm \rightarrow DK^\pm$  parameters. [ ] an alternative, it is possible to introduce only two Cartesian parameters [?], defining

$$\xi_{D\pi^\pm} = \left( \frac{r_B^{D\pi^\pm}}{r_B^{DK^\pm}} \right) \exp[i(\delta_B^{D\pi^\pm} - \delta_B^{DK^\pm})] \quad (2.44a)$$

and letting

$$x_\xi^{D\pi} = \text{Re}[\xi_{D\pi^\pm}] \quad y_\xi^{D\pi} = \text{Im}[\xi_{D\pi^\pm}]. \quad (2.44b)$$

In terms of these parameters, the usual Cartesian  $x_\pm$  and  $y_\pm$  are given by

$$x_\pm^{D\pi} = x_\xi^{D\pi} x_\pm^{DK} - y_\xi^{D\pi} y_\pm^{DK}, \quad y_\pm^{D\pi} = x_\xi^{D\pi} y_\pm^{DK} + y_\xi^{D\pi} x_\pm^{DK}. \quad (2.45)$$

Using this expression, the  $B^\pm \rightarrow D\pi^\pm$  yields can also be defined via Eq. (2.41) in the maximum-likelihood fit. Note that  $\xi$  does not depend on  $\gamma$ : all information on  $CP$  asymmetries in both the  $B^\pm \rightarrow DK^\pm$  and  $B^\pm \rightarrow D\pi^\pm$  channels is encoded in  $x_\pm^{DK}$  and  $y_\pm^{DK}$ . In the thesis, the latter parameterisation is chosen, because it allows for a stable fit for all six  $x$  and  $y$  parameters and the shared  $F_i$ ; the choice is described in much greater detail in Section 5.5.1.

The combined analysis of  $B^\pm \rightarrow DK^\pm$  and  $B^\pm \rightarrow D\pi^\pm$  decays presents a significant step forward, because it solves the problem of obtaining  $F_i$  parameters for the appropriate acceptance profile in a manner that avoids leading systematic uncertainties, and almost all reliance on simulation. This is of great importance, if the large data samples that will be collected by LHCb in the future are to be exploited to their full potential.

# 3

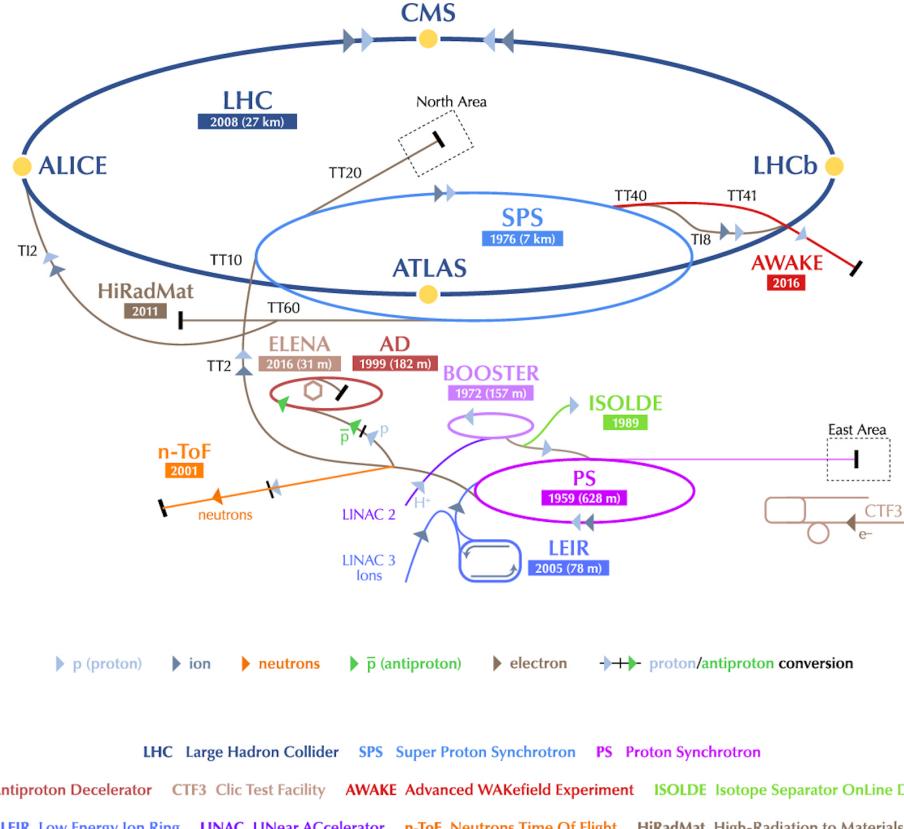
754

755

## The LHCb experiment

756 The LHCb experiment is one of the four large experiments at the Large Hadron  
757 Collider (LHC), the World's most powerful accelerator, able to accelerate protons  
758 to record centre-of-mass energies of  $\sqrt{s} = 13$  TeV in a 27 km long tunnel underneath  
759 Geneva. The LHCb experiment is specifically designed to study the large number  
760 of particles containing  $b$  or  $c$  quarks produced in such collisions, which has led  
761 to a number of design decisions that make the LHCb unique among the LHC  
762 experiments. The LHCb is not a solid-angle detector like the other three LHC  
763 experiments, CMS, ATLAS, and ALICE, but a single-arm spectrometer, instru-  
764 mented in the forward region where the majority of  $b\bar{b}$  pairs are produced. During  
765 data-taking the experiment is operated at a lower instantaneous luminosity than  
766 the other experiments, leading to far fewer  $pp$  interactions. This, in combination  
767 with a vertex detector located extremely close to the interaction point, allows for  
768 excellent resolution in the reconstruction of primary and secondary vertex locations,  
769 crucial to many of the central measurements of the experiment. Finally, dedicated  
770 particle-identification detectors allow for very efficient separation of hadron species,  
771 absolutely crucial to isolate a number important signal decays (including the  
772  $B^\pm \rightarrow DK^\pm$  decay studied in the thesis). Each of these features is described in  
773 much greater detail in the sections below.

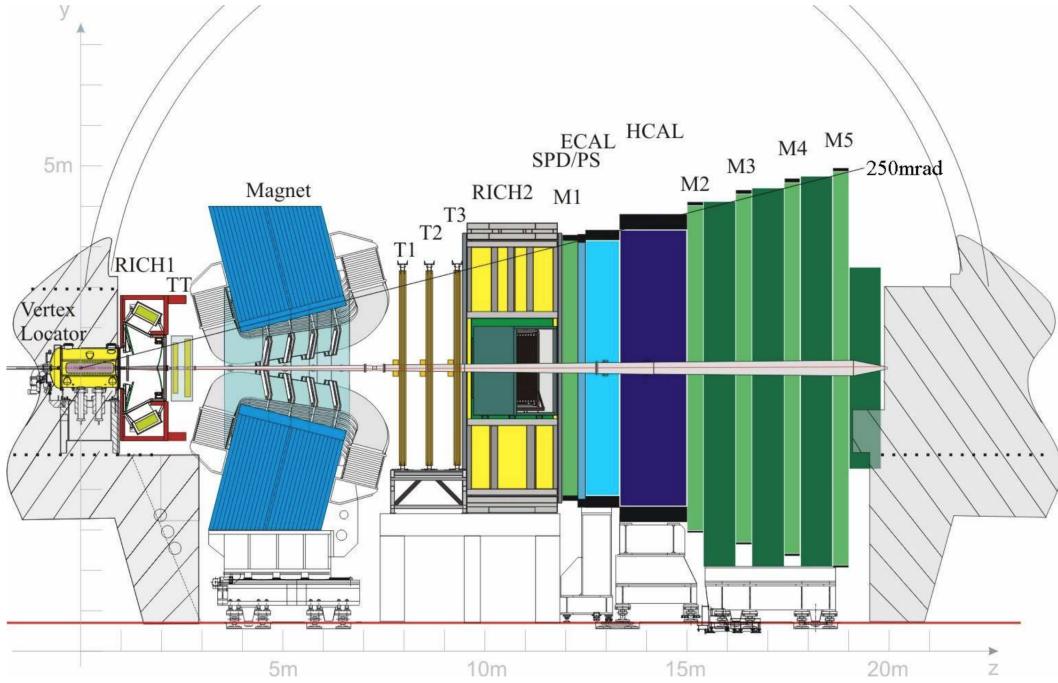
774 During operation of the LHC, bunches of about  $\mathcal{O}(10^{11})$  protons are accelerated  
775 to the desired centre-of-mass energy in a series of linear and circular accelerators,  
776 the final one being the LHC itself. This is illustrated in Fig. 3.1. The bunches  
777 remain in the LHC for the duration of a *fill*, typically about 12 hours, where they  
778 are made to collide at four distinct locations, the collision points, each home to one



**Figure 3.1:** The CERN accelerator complex, including the length and construction year for a number of accelerators, not all of which are used in  $pp$  operations. During  $pp$  operation, the proton acceleration chain is: LINAC 2 → BOOSTER → → SPS → LHC. The figure is reproduced from Ref. [?].

of the experiments. The collisions occur with a frequency of up to 40 MHz. A fill ends when the beams are dumped, typically because the average number of protons in the bunches has become too low, after which the whole process begins again.

The LHC has been providing  $pp$  collisions during two periods so far: Run 1 during 2011 and 2012, where the centre-of mass energies were  $\sqrt{s} = 7 \text{ TeV}$  and  $8 \text{ TeV}$  respectively, and Run 2 from 2015 to 2018, where  $\sqrt{s} = 13 \text{ TeV}$ . The instantaneous luminosity at the LHCb collision point has been  $4 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ , and has allowed for the collection of data set corresponding to a total of  $3 \text{ fb}^{-1}$  during Run 1 and  $6 \text{ fb}^{-1}$  during Run 2. The full data set forms the basis of the thesis. This instantaneous luminosity is significantly lower than at other collision points, for example the peak instantaneous luminosity in the ATLAS detector was about  $20 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$  in 2018 [?], but 50 times higher than in LHCb. The lower luminosity is necessary to limit the number of  $pp$  interactions per bunch crossing to an average of about 1–1.6, necessary for a vertex reconstruction with the required precision. is achieved by colliding the proton beams with a transverse off-set at the LHCb collision point.

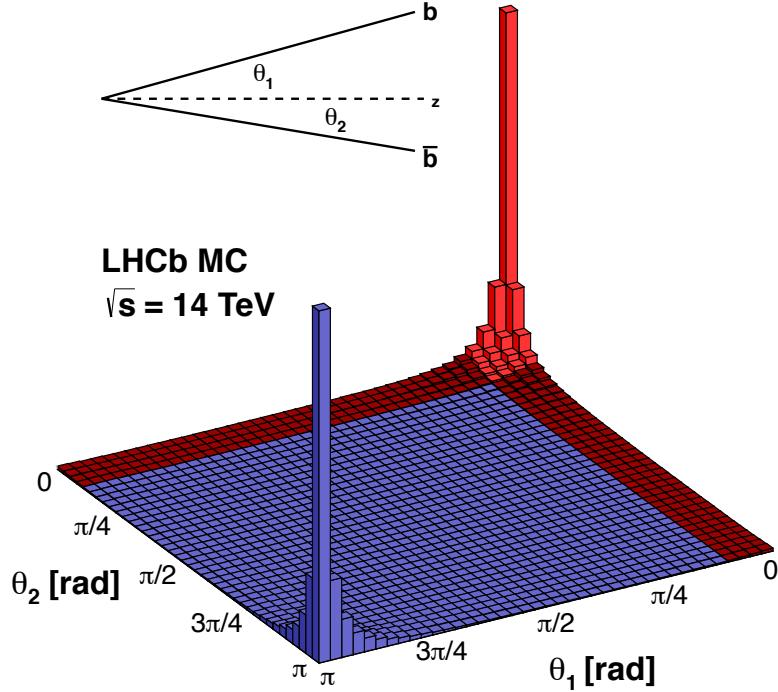


**Figure 3.2:** Overview of the LHCb detector reproduced from Ref. [?, ?]. The individual sub-detectors are described in detail in the chapter.

This has the added benefit that the offset can be continuously adjusted during a fill of the LHC, and thus all data can be taken at the same instantaneous luminosity, allowing for simpler trigger configuration, and simpler subsequent analysis because the detector occupancy is constant. The lower luminosity, of course, comes with the downside that the collected data sample is smaller.

### 3.1 The LHCb subdetectors

The LHCb detector, shown in Fig. 3.2, is able to detect particles in the forward region  $\eta \in [2, 5]$ , corresponding to an angle  $\theta$  with respect to the beam line between 15 and 300/250 mrad in the horizontal/vertical direction. As illustrated in Fig. 3.3, the  $b\bar{b}$  production cross section is very large within the LHCb acceptance: even though the acceptance covers less than 2% of the solid angle, 24% of all  $b\bar{b}$  pairs created at  $\sqrt{s} = 14$  TeV are within the acceptance. The detector is described with a coordinate system, where the  $z$ -axis is along the beam line and the  $x$  ( $y$ ) axis is in the horizontal (vertical) directions normal to the beam line. The origin is at the collision point. The experiment consists of a number of sub detectors, located in the region from around the interaction point, and up to a distance of  $z = 20$  m along the beam line (in the following, the direction from the interaction point towards



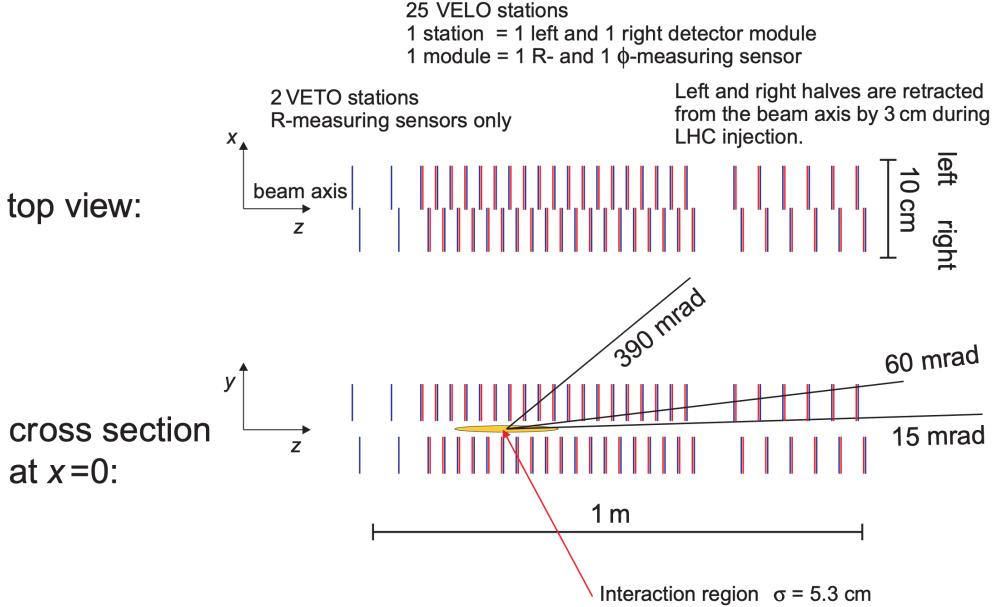
**Figure 3.3:** Production cross section of  $b\bar{b}$  pairs at a centre-of-mass energy of  $\sqrt{s} = 14$  TeV, as a function of  $\theta_1$  and  $\theta_2$ , the angle of the  $b$  and  $\bar{b}$  quark, respectively, with respect to the beam axis  $z$ . The LHCb acceptance is marked in red. The cross-section looks very similar for  $\sqrt{s} = 7, 8$  TeV. The figure is taken from Ref. [?].

the sub detectors is denoted *downstream*, and the opposite direction *upstream*).  
This section describes each of them in detail.

### 3.1.1 The VELO

The VErtex LOcator (VELO) [?] is a silicon detector located immediately around the collision point, used to provide precise measurements of the particle track coordinates in the interaction region. These are used to reconstruct the production and decay vertices of beauty and charm hadrons with a very high accuracy, allowing for an accurate reconstruction of their life times, and for efficient background rejection. The ability to distinguish tracks originating in secondary vertices also plays a crucial role in efficient triggering, as described further below.

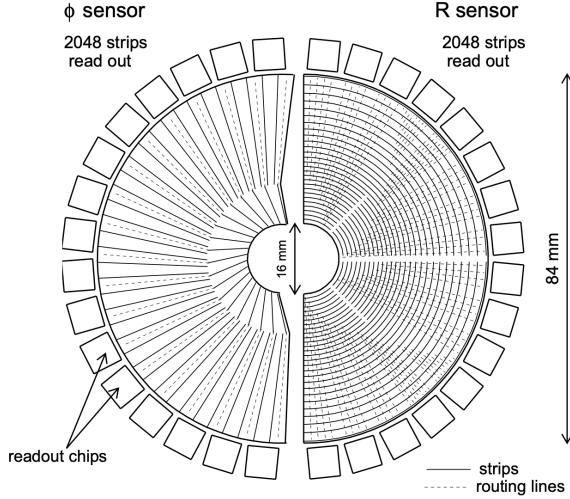
The detector consists of 21 VELO stations positioned along the beam line as illustrated in Fig. 3.4. Each station consists of two *modules*, mounted on each side of the beam line; each module, in turn, consists of two silicon strip detectors, where



**Figure 3.4:** Overview of the arrangement of VELO stations from the VELO Technical Design Report (TDR) [?]. The actual detector includes 21 stations instead of 25, but the overall design is identical [?].

the strips are oriented to provide a measurement of  $r$ , the radial distance from the beam line, and  $\phi$ , the azimuthal angle, respectively. This is illustrated in Fig. 3.5. The strip pitch varies between 40 and 100  $\mu\text{m}$  depending on the distance from the beam line. The stations are positioned such that all tracks that are within the acceptance region of the downstream detectors and originate at the interaction point are guaranteed to intersect 3 detector segments. During operation, the segments are located only 8 mm from the beam, this is achieved by mounting them on a moving frame that can be retracted during beam commissioning to avoid radiation damage. The detectors are kept in a vacuum, shielded from the beam vacuum by a 0.3 mm thick *RF foil* made of aluminium that also serves to screen the detector from electric fields induced by the proton beam. The silicon sensors were kept at an operating temperature of about -7  $^{\circ}\text{C}$ , achieved with a liquid-CO<sub>2</sub> cooling system.

The primary vertex (PV) resolution of the VELO is typically  $\sim 10 \mu\text{m}$  in the  $x$  and  $y$  directions and  $\sim 50 \mu\text{m}$  in the  $z$  direction, improving with the number of tracks originating at the PV, and deteriorating with the overall number of PVs [?]. The typical uncertainty on the decay length of a  $B$  meson is about 230  $\mu\text{m}$ , compared to a typical decay length  $O(10)$  mm. The resolution of the *impact parameter*, IP, of a track is well-described by the formula  $\sigma_{\text{IP}} = (15 + 29/[p_T/( \text{GeV}/c)]) \mu\text{m}$ . This parameter excellently distinguishes particles produced in secondary decays, from those produced in the primary interaction (for which the IP would be zero, were it not for the experimental resolution).



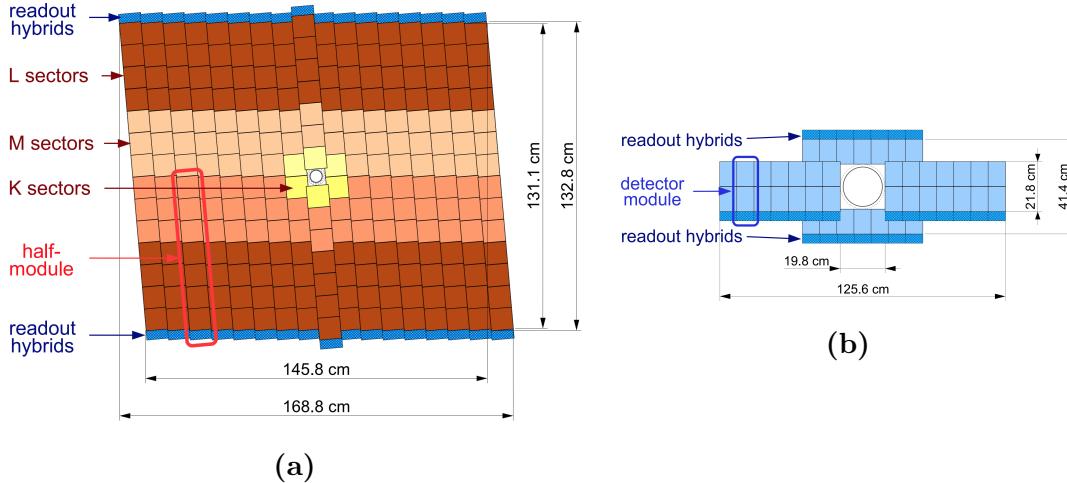
**Figure 3.5:** Illustration of the silicon strip layout in the VELO modules designed to measure (left) the azimuthal angle,  $\phi$ , of a track, and (right) the radial distance from the beam,  $r$ . Reproduced from Ref. [?].

### 3.1.2 Magnet and tracking stations

The LHCb experiments uses a warm (non-superconducting) dipole magnet to measure the momentum of charged particles, by providing a maximum magnetic field strength of approximate 1T and a total bending power of about 4 T m over the region where  $z \in [2.5, 8]$  m. The magnetic field has been measured to a relative precision of about  $4 \times 10^{-4}$  and is uniform within a percent within the tracking volume. The profile of the magnetic field along the  $z$ -axis is shown in Fig. 3.14 on page 41, where the track types within LHCb are defined. The magnet can provide a magnetic field in either vertical direction; over the span of a year of data taking approximately equal amounts of data are collected with the magnet in the "Up" and "Down" configurations; this leads to a number of large-asymmetry effects to cancel, significantly reducing potential systematic uncertainties.

The tracking system consists of the VELO, and four other tracking stations: the Tracker Turicensis (TT) upstream of the magnet, and the tracking stations 1–3 (T1, T2, T3) downstream of the magnet. The downstream tracking stations each consist of an Inner Tracker (IT) based on silicon strips, and an Outer Tracker (OT) that employs drift tubes.

Both the TT and IT are based on silicon strip detectors with a pitch of about 200  $\mu\text{m}$ ; they were developed as a single project and are collectively known as the Silicon Tracker (ST). The TT is a 140 cm wide and 130 cm tall planar tracking station, covering the whole LHCb acceptance. It is shown in Fig. 3.6a. At each

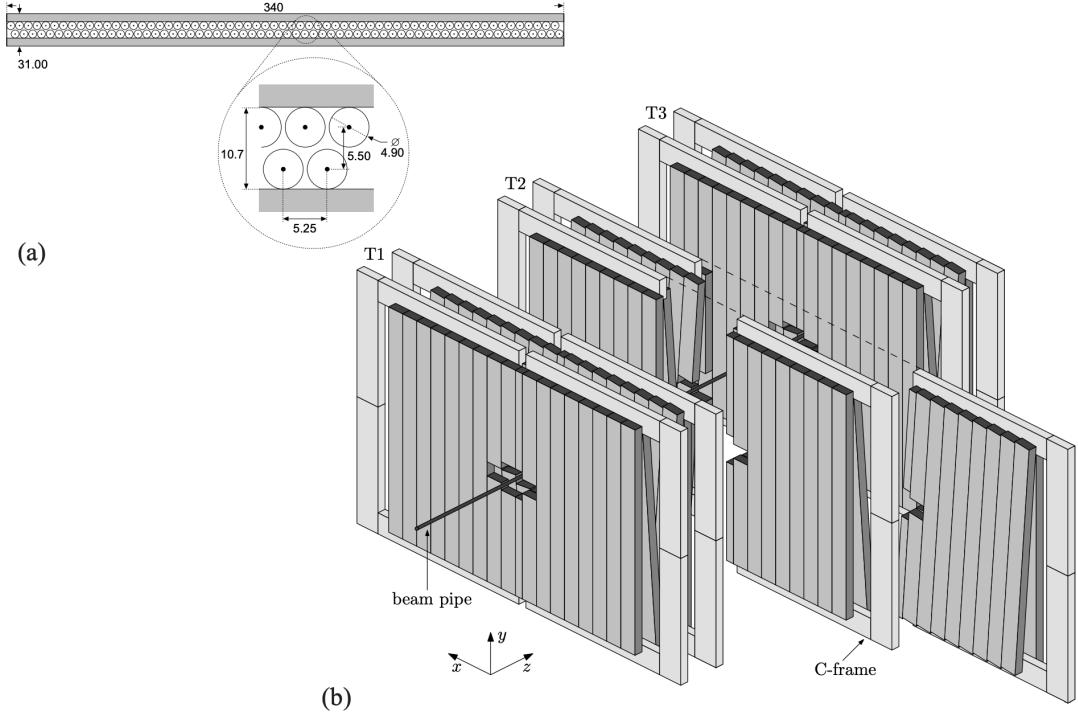


**Figure 3.6:** Overview of (a) a *v*-layer module of the IT and (b) an *x*-layer module of the IT. Reproduced from Ref. [?]

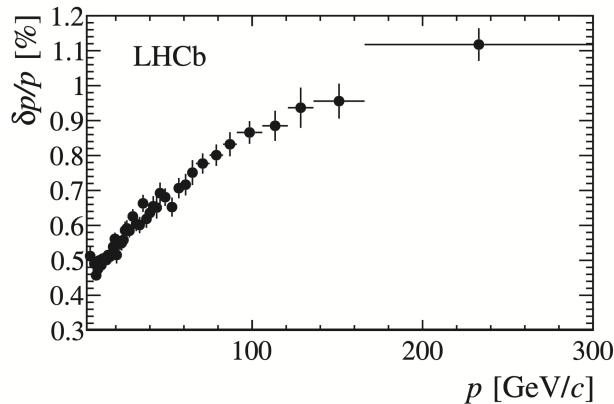
of the T1–T3 stations, the IT consist of four modules, arranged around the beam pipe as illustrated in Fig. 3.6b. They do now cover the full LHCb acceptance, only the very-forward region where the number of tracks is largest. Each TT or IT module comprises of four layers of silicon strips, where the central two layers are rotated  $\pm 5^\circ$  with respect to the first and last layer (an *x-u-v-x* geometry). The ST has a spatial resolution for a given track of approximately 50  $\mu\text{m}$ , chosen because the overall momentum resolution is then dominated by multiple-scattering effects for almost all  $\text{p}_{\text{T}}$  momenta.

At the T1–T3 stations, the OT covers the part of the overall acceptance of 300 (250) mrad in the horizontal bending (vertical non-bending) plane that is not covered by the IT. The OT consists of arrays of gas-tight drift tubes with inner diameters of 4.9 mm. The OT is shown illustrated in Fig. 3.7. An Ar/CO<sub>2</sub>/O<sub>2</sub> (70/28.5/1.5) gas mixture is used to fill the tubes that ensures a drift time below 50 ns and a drift coordinate resolution of 200  $\mu\text{m}$ . The use of a drift-chamber detector is necessary, because it was not economically feasible to instrument the whole LHCb acceptance with silicon strip detectors in T1–T3. The condition that the OT occupancy should not be above 10 % in typical run conditions determined the boundary between the IT and the OT.

The overall relative momentum resolution achieved for most charged tracks in LHCb is less than a percent, as illustrated in Fig. 3.8, where it has been determined from a fit to the mass peak in  $J/\psi \rightarrow \mu^+\mu^-$  decays in Run 1 data.



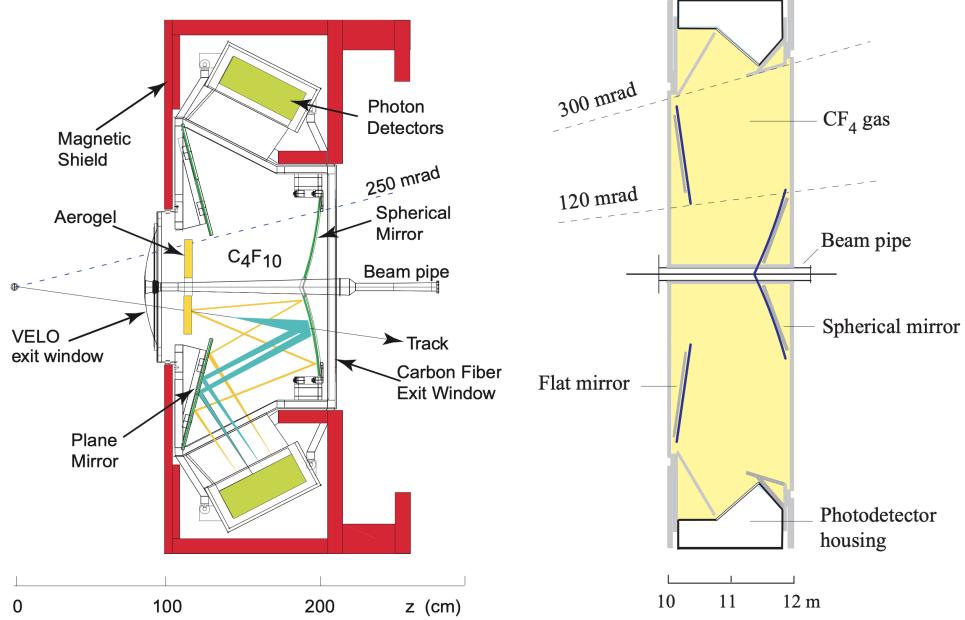
**Figure 3.7:** (a) Cross section of an OT module. (b) Arrangement of the OT modules in tracking stations. Reproduced from Ref. [?].



**Figure 3.8:** Relative uncertainty on the momentum of charged tracks (specifically long tracks, cf. the definitions in Section 3.2) in the LHCb detector, determined via the mass resolution obtained in  $J/\psi \rightarrow \mu^+\mu^-$  decays in Run 1 data. Reproduced from Ref. [?]

### 3.1.3 The RICH detectors

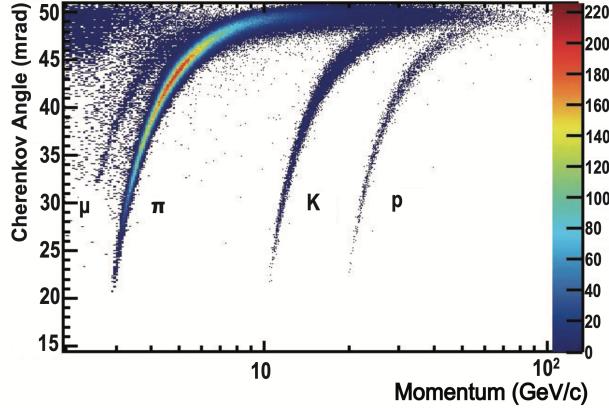
Two Ring Imaging Cherenkov detectors (RICH) provide crucial information for particle identification (PID) in LHCb, in particular the ability to separate pions and kaons that is absolutely essential for the measurement presented in the thesis. The RICH 1 detector is located upstream of the magnet, in between the VELO and the TT tracking station. It is designed to provide PID capability for tracks



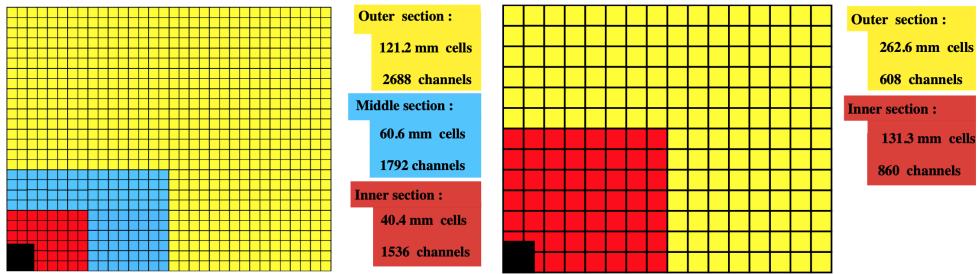
**Figure 3.9:** Overview of (left) the Rich 1 and (right) the RICH 2 detectors. Reproduced from Ref. [?, ?].

in the momentum range  $p \in [1 - 10]$  GeV/ $c$  using a C<sub>4</sub>F<sub>10</sub> radiator, and covers the full LHCb acceptance. During Run 1 the RICH 1 detector also included an Aerogel radiator designed to provide PID for very low momentum particles; however, it was removed before Run 2 because it did not meet the performance requirements during Run 1 [?, ?]. The RICH 2 detector is located downstream of the T1–T3 tracking stations. It is designed to provide PID capabilities for higher momentum tracks in the range  $p \in [15 - 100]$  GeV/ $c$  using a CF<sub>4</sub> radiator. It only covers the very forward region where  $|\theta| < 120$  mrad(100 mrad) in the horizontal (vertical) directions, as high momentum particles are produced in that region. In both RICH detectors, mirrors are used to reflect the Cherenkov photons to arrays of Hybrid Photon Detectors (HPDs) located outside the LHCb acceptance. The optics are designed such that photons originating from a given track form rings in the HPD arrays, where the radius is determined by the Cherenkov angle  $\theta_c$ . The detectors are illustrated in Fig. 3.9.

The resolution on  $\theta_c$  can be measured by fitting the obtained  $\theta_c$  distribution in high momentum tracks, where the Cherenkov angle is saturated. It is found to be  $1.618 \pm 0.002$  mrad for RICH 1 and  $0.68 \pm 0.02$  mrad for RICH 2 in Run 1 data [?], and was essentially unchanged in Run 2 [?]. Figure 3.10 shows the relation between track momentum and  $\theta_c$  in RICH 1 for *isolated tracks* in Run 1 data; these are tracks where the Cherenkov ring does not overlap with any other



**Figure 3.10:** Cherenkov angle for isolated tracks in the RICH 1 radiator as a function of track momentum. Reproduced from Ref. [?].



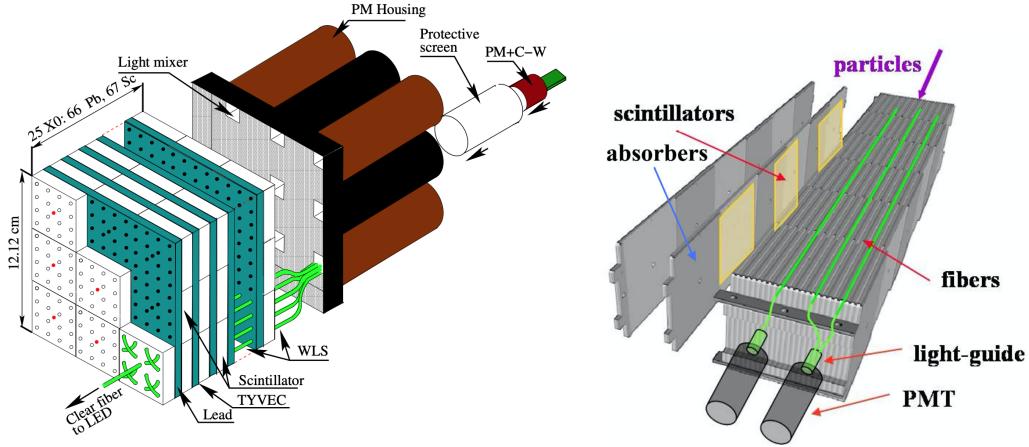
**Figure 3.11:** Illustration of the calorimeter cell size of (left) the ECAL and (right) the HCAL. Reproduced from Ref. [?].

Cherenkov rings. The bands for each hadron species are clearly visible, and it can be seen that the RICH detector also provide some ability to distinguish muons. The definition of the PID variables used in analysis is discussed in Section 3.2.2, along with the achieved PID performance.

### 3.1.4 Calorimeters

The calorimeter system of the LHCb detector has four components. Ordered from the interaction point, these are the Scintillating Pad Detector (SPD) Shower (PS), an Electromagnetic CALorimeter (ECAL), and a Hadron CALorimeter (HCAL). Information from the calorimeters also provide identification of electrons, photons, and hadrons, and measurements of their energies and positions, and also plays a crucial role in the triggering, as described below. In all four cases, light is produced in organic scintillators and transmitted to Photo Multiplier Tubes (PMTs) via optical fibres [?].

The SPD and PS detectors consist of almost identical planes of rectangular scintillator pads, with a 15 mm thick lead absorber located in between. The presence

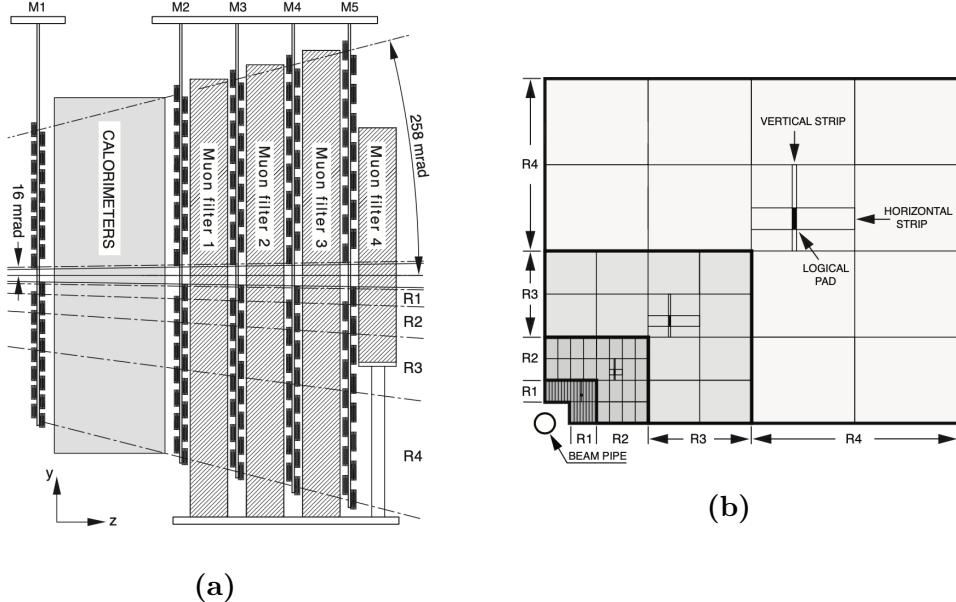


**Figure 3.12:** Illustration of (left) an ECAL and (right) a HCAL module. Reproduced from Ref. [?, ?].

of the SPD before the first absorption allows for the separation photons and charged particles using tri[ ] information alone, because only elec[ ]ic showers formed by the latter will deposit energy in the SPD. The PS allows for the separation of pion and electron tracks, as only the latter deposit significant energy in the thin lead layer. The cell divisions of the detectors closely follow that of the ECAL, shown in Fig. 3.11, to allow for the matching of energy deposits.

The ECAL has a Shashlik structure, with 66 layers consisting of 2 mm of lead absorber and 4 mm of scintillator; an example of a calorimeter module is shown in Fig. 3.12. Accurate energy measurements require that the full electronic shower is contained in the ECAL, which is achieved since the structure extends for 25 radiation lengths. The scintillators are divided into cells that allow for the determination of the location and shape of energy deposits; the cell dimensions vary as a function of radial distance from the beam pipe as shown in Fig. 3.11, to take into account the varying occupancy. The resolution of the ECAL has been measured to be  $\Delta E/E \simeq (9/\sqrt{E} \oplus 0.8)\%$  ( $E$  in  $\text{GeV}/c^2$ ) [?].

The HCAL is located downstream of the ECAL, designed to measure the energy of charged hadrons (which leave relatively little energy in the ECAL). It is constructed with layers of 1 cm iron absorbers inter-spaced with scintillators, oriented *along* the beam direction, such that a typical track will traverse 16 mm of iron per 4 mm of scintillator [?]. As for the ECAL, the cell size varies as a function of distance to the beam line, as shown in Fig. 3.11. An example of a module is shown in Fig. 3.12. The energy resolution required for efficient triggering is moderate; therefore, the HCAL only has a length of 5.6 interaction lengths and can measure the hadron energies at a resolution of  $\Delta E/E \simeq (69/\sqrt{E} \oplus 9)\%$  ( $E$  in  $\text{GeV}/c^2$ ) [?].



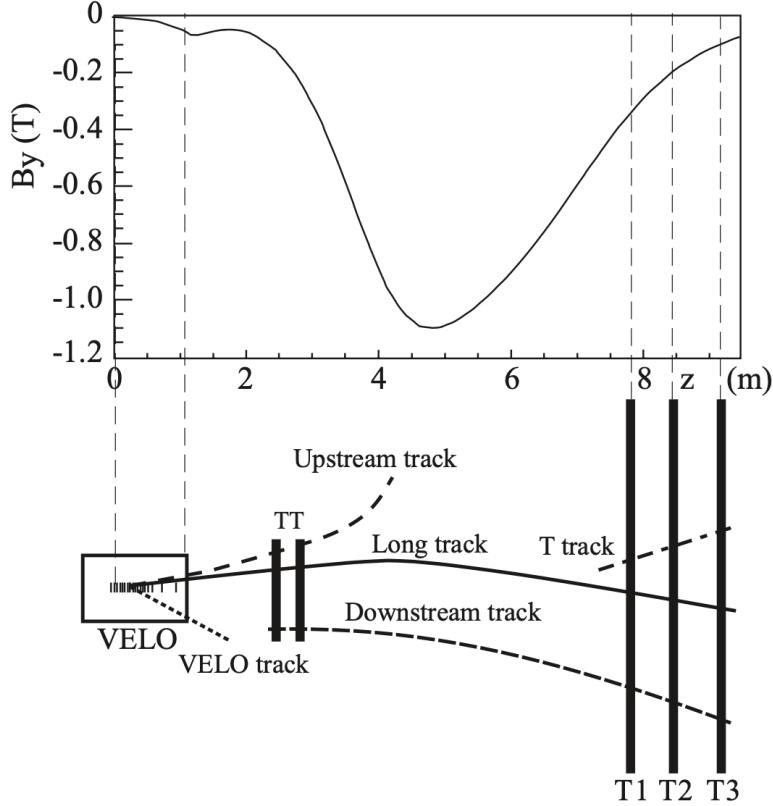
**Figure 3.13:** Illustration of (a) the location of the muon stations along the  $z$ -axis of the experiment, and (b) the geometry of the logical pads of the M3 muon station. Reproduced from Ref. [?].

### 3.1.5 Muon detectors

Muon identification and triggering is crucial for a range of high-profile LHCb measurements, such as lepton-universality tests or measurements of  $B_{(s)}^0 \rightarrow \mu^+ \mu^-$  decays. In the thesis, muon identification plays a role in suppressing a number of backgrounds. The LHCb muon system consists of 5 tracking stations, M1–M5, covering the full LHCb acceptance. M1 is located upstream of the ECAL, whereas M2–M5 are located downstream of the HCAL and inter-spaced with 80 cm thick ion absorbers in order to select penetrating muons. This is illustrated in Fig. 3.13a. The detectors are multiwire proportional chambers (MWPC), organised into logical pads, the dimensions of which define the  $(x, y)$  resolution of the measured spatial points. As for the calorimeters, the size of the pads vary as a function of the radial distance from the beam pipe, as illustrated in Fig. 3.13b. The resolution is significantly better in the bending plane ( $x$ ) than in the non-bending plane ( $y$ ). The resolution is also significantly better in the M1–3 stations than in M4 and M5, which are mostly used to identify penetrating tracks. The muon system can independently measure the  $p_T$  of a muon to within 20 %, which allows for efficient triggering.

## 3.2 Reconstruction

This section describes the reconstruction algorithms that fit the detector hits in the tracking stations to form track candidates, as well as the algorithms used to



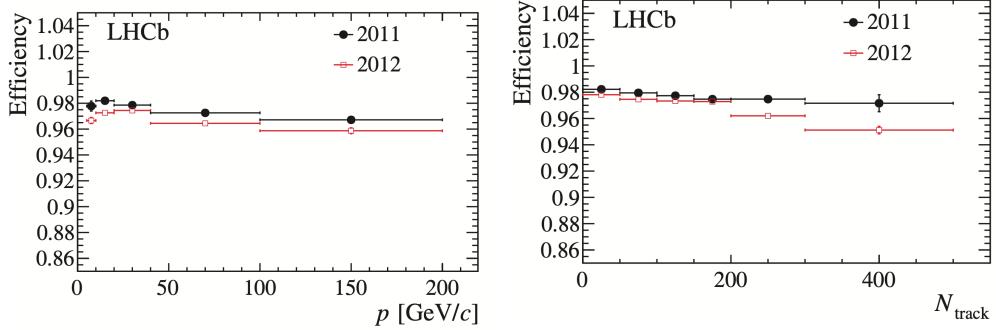
**Figure 3.14:** Definition of track types within the LHCb detector, depending on which set of tracking detectors the track intersects. The profile of the magnetic field is also shown. Reproduced from Ref. [?].

971 identify the types of the particles that formed these tracks.

### 972 3.2.1 Track reconstruction

973 The LHCb experiments operates with a number of different particle track types,  
 974 depending on which sub detectors a track intersects; these are summarised in  
 975 Fig. 3.14. The two track types that are important for this thesis are *long* tracks,  
 976 which have hits in the VELO and the TT and T1–T3 tracking stations, and  
 977 *downstream* tracks that only have hits in the TT and T1–3 tracking stations. The  
 978 analysis depends on both track types because a number of  $K_S^0$  mesons produced in  
 979 the signal decay leave the VELO before they decay into the  $\pi^+\pi^-$  final state that  
 980 is reconstructed; hence these pions necessarily form downstream tracks.

981 The first step is to form track candidates from hits in the VELO (VELO tracks)  
 982 and T1–3 stations (T tracks) separately; because the magnetic field is low in  
 983 the tracking detectors, these tracks are fairly straight. Long tracks are formed  
 984 using two separate search strategies: in one, *forward tracking* [?], VELO tracks



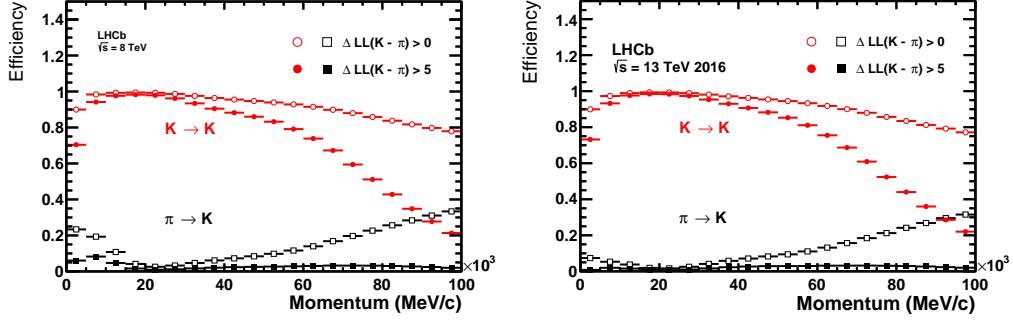
**Figure 3.15:** The long track reconstruction efficiency as a function of (left) track momentum and (right) the number of charged tracks in the event. The figure is reproduced from Ref. [?].

are used as seeds and matched with hits in the TT and T1–3 tracking stations by extrapolation. These are combined to form long tracks that are required to pass a set of quality conditions. An alternative approach, *track matching* [?, ?], matches VELO and T tracks by extrapolating both through the bending region, and deciding if they below together; finally TT hits are added. The union of tracks found via both approaches is saved, where only the track candidate with the best fit quality is kept in the case where a track appears twice. Downstream tracks are formed based on T tracks as seeds, matched with hits in the TT detector in a search region obtained by extrapolation of the seed [?]. Finally, each track is reprocessed using a Kalman filter that takes into account multiple scattering and corrects for energy loss due to ionisation [?, ?].

Many of the interesting signal decay channels of LHCb have 4–6 charged final state tracks, and therefore it is crucial to have a single-track reconstruction efficiency close to 100 %. The single-track reconstruction efficiency is shown in Fig. 3.15 as a function of track momentum and the number of tracks in an *event* (an *event* denotes a  $pp$  collision and all the particles produced therein and in subsequent decays). The efficiencies have been obtained in data, using a tag-and-probe method in  $J/\psi \rightarrow \mu^+\mu^-$  decays [?]. One muon, the *tag*, is fully reconstructed, while the other, the *probe* is only partially reconstructed, allowing for the  $J/\psi$  invariant mass to be reconstructed with reasonable resolution. If the partially reconstructed probe track is matched to a full long track, the track is classified as efficient. Similar efficiencies have been achieved in Run 2.

### 3.2.2 Particle identification

The information from the RICH detectors, the calorimeters, and the muon system is generally combined, for optimal identification of charged tracks as electrons, muons,



**Figure 3.16:** The probability to correctly identify a kaon/misidentify a pion as a kaon given two different requirements on  $\Delta LL(K)$ , as a function of track momentum in (left) Run 1 data from 2012 and (right) Run 2 data from 2016. Reproduced from Ref. [?].

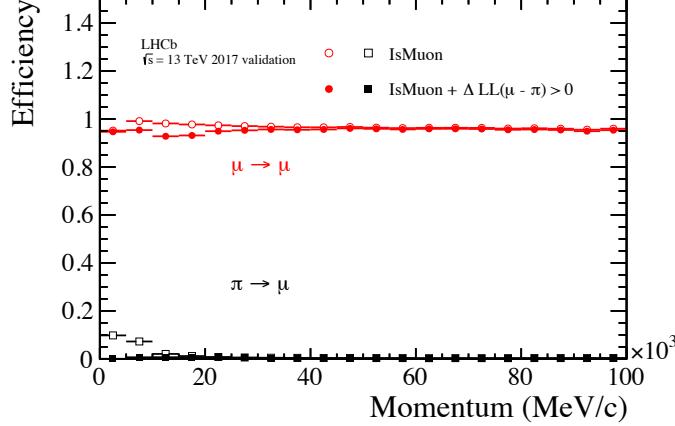
1010 pions, kaons, or protons. Photons and neutral pions are identified using the ECAL,  
 1011 but play no role in the thesis, and will not be discussed further.

1012 The ability to separate  $B^\pm \rightarrow DK^\pm$  and  $B^\pm \rightarrow D\pi^\pm$  decays is essential to the  
 1013 measurement presented in this thesis. In LHCb, hadron separation is achieved via  
 1014 information from the RICH detectors, using a likelihood method where the observed  
 1015 pattern of hit pixels in the photo detectors is compared to the expected pattern,  
 1016 given all reconstructed tracks in an event under a given set of particle hypothesis.  
 1017 The likelihood is maximised by varying the particle hypotheses for each track being  
 1018 an electron, muon, pion, kaon, or proton [?]. It is necessary to consider all tracks  
 1019 of an event simultaneously because the Cherenkov rings of different tracks overlap.  
 1020 For each track, the maximum log likelihood of a particle hypothesis, say that the  
 1021 track is a kaon, relative to the hypothesis that it is a pion

$$\Delta LL_{\text{track}_i}^{\text{RICH}}(K) = \ln \mathcal{L}_{\max}^{\text{RICH}}(\text{pattern}|\text{track}_i = K) - \ln \mathcal{L}_{\max}^{\text{RICH}}(\text{pattern}|\text{track}_i = \pi), \quad (3.1)$$

1022 is saved to inform PID decisions. In the case of pion-kaon separation, this variable  
 1023 alone is enough to achieve good separation power; in the remainder of the thesis  
 1024 it is denoted PIDK. The PID performance for pion-kaon separation has been  
 1025 measured in calibration data, following a procedure described in Section 5.2.1,  
 1026 and is illustrated in Fig. 3.16.

1027 Muons are identified by extrapolating tracks to the muon stations to define  
 1028 fields-of-interest (FOI). A track is considered as a muon candidate when a minimum  
 1029 number of stations (2–4 depending on the track momentum) have hits in the  
 1030 corresponding FOI [?, ?]. This information is encoded in a variable denoted `isMuon`  
 1031 throughout the thesis. Additional information, such as a comparison of the slopes  
 1032 of the track in the main tracker and the muon stations, and the average track-hit



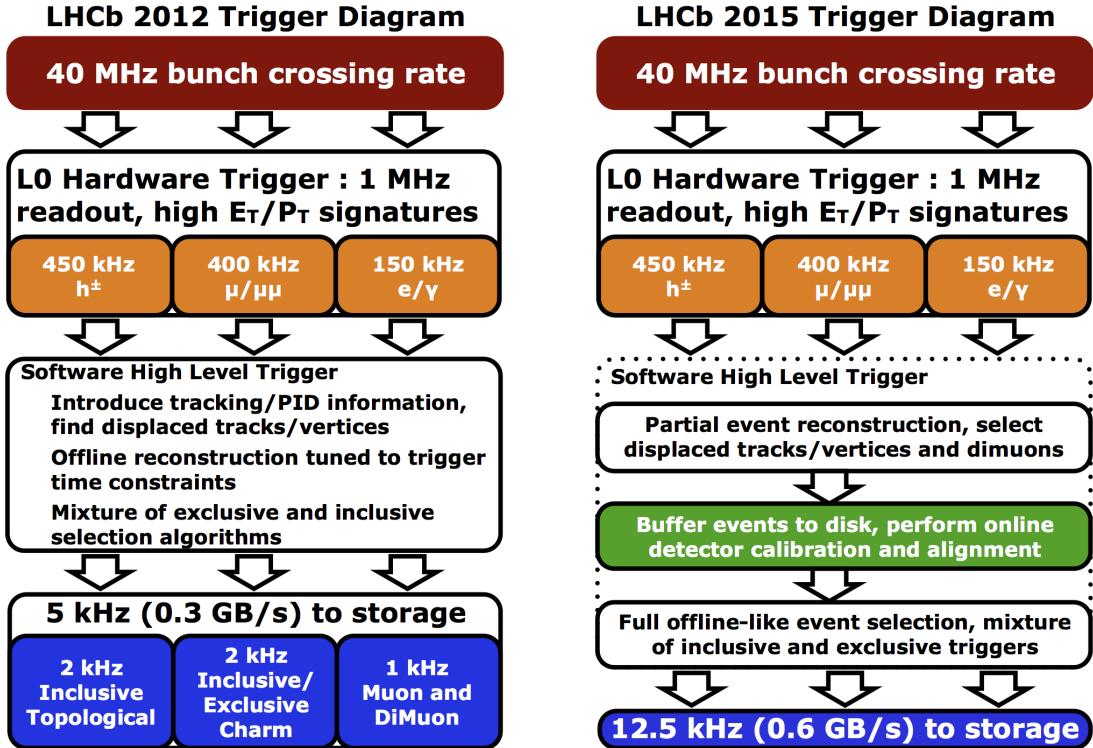
**Figure 3.17:** The probability to correctly identify a muon/misidentify a pion as a muon given requirements on either `isMuon` or  $\Delta LL(\mu - \pi)$ , as a function of track momentum in Run 2 data from 2017. Reproduced from Ref. [?].

distance in the FOI is used to form a  $\Delta LL^{\text{muon}}(\mu)$  variable analogous to the one defined in Eq. (3.1) for the RICH detectors; it can be combined with  $\Delta LL^{\text{RICH}}(\mu)$  to form a PID variable that takes information from both detectors into account, denoted  $\text{PID}_{\text{mu}}$ . The performance of the muon PID variables is shown in Fig. 3.17 as obtained in data. It can be seen that requiring `isMuon=0` rejects muon tracks efficiently at all momenta; this is used in the analysis to veto a number of semi-leptonic backgrounds.

In similar manner, a potential semi-leptonic background with electrons is also vetoed in the analysis presented in the thesis. In LHCb, electron PID is mainly based on the balance between deposited energy and track momentum in the ECAL [?]. This information is combined with information on photo-electron deposits from brehmstrahlung, and energy deposits in the PS and HCAL, as well as information from the RICH and muon detectors, to form yet another  $\Delta LL$  variable, denoted  $\text{PID}_{\text{e}}$ . As an example of the obtainable performance, an average electron selection efficiency of  $(91.9 \pm 1.3)\%$  was achieved in displaced  $J/\psi \rightarrow e^+e^-$  decays in Run 1, with a hadron misidentification rate of  $(5.54 \pm 0.02)\%$  [?].

### 3.3 The LHCb trigger system

The collision rate in the LHC is up to 40 MHz, with a visible inelastic collision rate in LHCb of up to 30 MHz. The LHCb uses a multi-stage trigger to reduce rate with which events are stored to a manageable level (of eg. 12.5 kHz during Run 2). The first stage consists of a hardware trigger that selects events with high transverse energy in the calorimeters, or hits in the muon detectors. This is followed by two software stages that rely on a reconstruction of tracks in the



**Figure 3.18:** Illustration of stages and event processing rates in the LHCb trigger during (left) Run 1 and (right) Run 2.

1055 detector to select events that are likely to include interesting physics. The overall  
 1056 trigger stages were identical in Run 1 and Run 2, however the throughput rate was  
 1057 upgrade significantly between the two data taking period, as was the quality of the  
 1058 reconstruction in the software trigger stages; in Run 2, the final software trigger  
 1059 decisions are in fact based on the full  event reconstruction [?]. The stages  
 1060 are illustrated in Fig. 3.18, and described in detail in the following.

1061 A further, offline processing and reconstruction step is applied to all events  
 1062 before they are made available to most LHCb analyses, commonly denoted as the  
 1063 *stripping* step. Although the stripping does not form part of the LHCb trigger,  
 1064 it does constitute an additional, centralised filter on the data, and a description  
 1065 is included in Section 3.3.3.

### 1066 3.3.1 The level-0 hardware trigger

1067 The level-0 (L0) triggers that select physics events are based on the calorimeters  
 1068 and the muon system. The ECAL and HCAL are divided into clusters of  $2 \times 2$   
 1069 cells, for which the transverse energy is defined as

$$E_T = \sum_j E_j \sin \theta_j, \quad (3.2)$$

1070 where  $\theta_j$  is the angle of cell  $j$  with respect to the beam axis and the average collision  
 1071 point. The trigger forms a L0Hadron candidate with the highest  $E_T$  found in the  
 1072 HCAL, combined with the ECAL cluster in front of it if such a cluster is present.  
 1073 Photon and electron candidates are formed based on [redacted] layers in the ECAL, identified  
 1074 by the presence (lack) of hits in the SPD for an electron (photon). The transverse  
 1075 energies of the candidates are compared to a fixed set of thresholds, and events  
 1076 where at least one candidate is above threshold are retained.

1077 The muon trigger searches for straight line tracks in the muon stations, [redacted] eliminating  
 1078 the associated muon  $p_T$  based on the track direction. An event is retained if either  
 1079 the largest muon  $p_T$  is above a given threshold, or the product of the two highest  
 1080 muon  $p_T$  values is above a different threshold.

1081 High-multiplicity events take a long time to process in the subsequent software  
 1082 stage; therefore it is favourable for the overall retention rate of interesting physics  
 1083 decays to put a maximum limit on the event multiplicity at the L0 stage. This  
 1084 is achieved by requiring the number of hits in the SPD detector to be below a  
 1085 threshold value in most L0 lines.

### 1086 3.3.2 High-level triggers

1087 The events that pass the L0 trigger are passed to a farm of multiprocessor computing  
 1088 node, the Event Filter Farm (EFF), tasked with bringing the rate down from  
 1089 approximately 1 MHz to the  $\mathcal{O}(1 - 10)$  kHz rate that can be saved to disk. The EFF  
 1090 consisted of 900 (1700) nodes during Run 1 (Run 2). The software-based filtering  
 1091 proceeds in two stages: a first filter (HLT1) brings the rate down to approximately  
 1092 40 (110) kHz based on a limited reconstruction of the event, after which a second  
 1093 stage (HLT2) filters the events further based on a more complete reconstruction.  
 1094 Each step executes a number of different algorithms, each of which can allow an  
 1095 event to be accepted; these are denoted *trigger lines*.

1096 During both runs, the HLT1 performed a partial event reconstruction by building  
 1097 long tracks that satisfy a  $p_T$  requirement using the forward tracking approach  
 1098 described in Section 3.2.1, and determining the location of PVs using VELO tracks.  
 1099 In both runs, the HLT1 included an inclusive trigger that selected a high  $p_T$  track  
 1100 with significant displacement of all PVs (typical of a  $b$  or  $c$  decay). This line is denoted  
 1101 `HLT1TrackAllL0` in Run 1 [?]; for Run 2 the track requirements were reoptimised  
 1102 and it is denoted `Hlt1TrackMVA`. Further, an additional inclusive trigger was added  
 1103 that forms a two-prong vertex out of high  $p_T$  tracks inconsistent with originating in  
 1104 a PV, and applies a multivariate classifier to determine if it is signal-like based on a  
 1105 number of track and vertex properties. This line is denoted `Hlt1TwoTrackMVA` [?].

1106 These lines triggered all events included in the analysis of the thesis; other lines exist  
1107 for selecting events that include muons, calibration data, low-multiplicity events, and  
1108 a number of exclusive lines, for a total of approximately 20 lines during Run 2 [?].

1109 Because the rate of events is reduced significantly by HLT1, the HLT2 decisions  
1110 can be based in a more complete reconstruction of the event. Indeed, during Run 2  
1111 it was based on a complete, fully aligned reconstruction equivalent to the offline  
1112 reconstruction. During Run 1 the HLT2 reconstruction only included long tracks and  
1113 did exclude some low momentum tracks; this was a main motivation for the upgrade  
1114 of the EFF during the shutdown period. The need for full alignment in HLT2 means  
1115 that it could not be run fully online in Run 2; instead the output event were saved  
1116 to disk in the EFF, and processed with some delay [?]. The analysis presented in  
1117 the thesis is based on a number of inclusive "topological" trigger lines, based on  
1118 combinations of 2, 3, or 4 tracks that satisfy fit quality requirements, have high  $p_T$ ,  
1119 are separated from the PVs, and have a distance-of-closes-approach below 0.2 mm.  
1120 A multivariate classifier is applied to each formed  $n$ -body object, to determine if the  
1121 event should be accepted based on the track momenta, invariant mass, a corrected  
1122 invariant mass that takes into account missing transverse momentum, distance of  
1123 closest approach, and the impact parameter and separation with the associated PV.  
1124 The resulting trigger lines were denoted `Hlt2Topo{2, 3, 4}BodyBBDT` during Run 1  
1125 and `Hlt2Topo{2, 3, 4}Body` during Run 2. A large number of other HLT2 lines  
1126 exist (more than 500 in Run 2), including a significant number of exclusive lines that  
1127 aim to select specific decays and only save information on the signal decay, not the  
1128 whole event. This was made possible by the full reconstruction within HLT2 [?], and  
1129 have allowed for larger signal yields to be collected within the data storage limits.

### 1130 3.3.3 Offline data filtering: the LHCb stripping

1131 Events that are written to disk are processed with the full detector alignment and  
1132 calibration. In a further, offline processing step denoted the *stripping*, hundreds  
1133 of different, dedicated reconstructions are performed; decay candidates for various  
1134 signal decays are built and a number of requirements are made to reject backgrounds  
1135 from random track combinations. For example, the  $B^\pm \rightarrow D(\rightarrow K_S^0 h^+ h^-) h^\pm$   
1136 candidates that are analysed in this thesis are built during the stripping stage, as  
1137 described further in Section 5.1. The stripping is a centralised computing task,  
1138 executed on the Worldwide LHC Computing Grid [?], and allows the analysts  
1139 to process much smaller data sets during their individual analysis. Because the  
1140 stripping is based on data saved to offline storage it can be repeated; however,  
1141 the processing of data collected during a year of data taking takes many weeks,  
1142 so this does not happen often.

## 1143 3.4 Simulation



1144 A centralised LHCb simulation is able to simulate  $pp$  collisions with the proper  
1145 conditions within LHCb, subsequent secondary decays, and the detector response,  
1146 and process it via the full LHCb reconstruction. In this thesis, simulated decays  
1147 are used to determine the reconstructed invariant-mass distribution of a number  
1148 of decay modes, as well as a number of relative selection efficiencies. The  $pp$   
1149 collisions are generated using PYTHIA [?] with a specific configuration specific  
1150 to LHCb [?]. The time-dependent evolution and decays of unstable particles are  
1151 described by the EVTGEN [?] package, designed specifically for  $B$  physics. Final-  
1152 state radiation is generated using PHOTOS [?]. The interaction of the generated  
1153 particles with the detector, and its response, are implemented using the GEANT4  
1154 toolkit [?] as described in Ref. [?].

1155 The most significant computational cost of the simulation is due to the detector  
1156 simulation. A single  $pp$  collision produces  $\mathcal{O}(100)$  tracks in the detector, out of  
1157 which only a handful belong to the signal decay under study. Therefore, significant  
1158 computation resources can be saved by reusing the detector simulation of non-signal  
1159 tracks a number of times, while redescaying the signal particle, say a  $B^+$ , each time.  
1160 This approach is called ReDecay [?], and has been relatively widely adopted within  
1161 LHCb. ReDecay has been used to produce simulation samples corresponding to  
1162 the conditions in 2017 and 2018 for this thesis. In some cases, the use of ReDecay  
1163 necessitates special statistical treatment due the correlated detector occupancies  
1164 between signal candidates, but for the analysis in this thesis the impact is negligible.

1165 A number of sub-dominant backgrounds are investigated using the fast-simulation  
1166 package `RapidSim` [?]. This package can decay heavy  $b$  and  $c$  hadrons with kinematic  
1167 distributions similar to those in LHCb  $pp$  collisions, or with user defined input  
1168 distributions. The decays are typically evenly distributed over phase space, but can  
1169 also be handled with EVTGEN [?] to take involved spins and resonant structure  
1170 into account. Furthermore, a smearing of the obtained momenta is implemented,  
1171 based on the LHCb resolution.

# 4

1172

1173

1174

## Neutral kaon $CP$ violation and material interaction in BPGGSZ measurements

1175 The presence of a  $K_S^0$  meson in the  $D \rightarrow K_S^0 h^+ h^-$  final states introduces a small  
1176 bias in BPGGSZ measurements due to  $CP$ -violation in the neutral kaon sector  
1177 and asymmetries caused by the interaction between the neutral kaons and detector  
1178 material. These fundamental physics effects are reviewed in Section 4.1, after which  
1179 the chapter presents a detailed analysis of the impact on the LHCb measurement  
1180 that is the subject of the thesis, as well as future  $\gamma$  measurements with the Belle II  
1181 experiment. Prior to this analysis, the only existing work on the effect on  $\gamma$   
1182 measurements suggested a small effect in  $B^\pm \rightarrow D K^\pm$  measurements but potentially  
1183 very significant effects in measurements based on  $B^\pm \rightarrow D \pi^\pm$  decays [?]. However,  
1184 as described in Section 4.1.1, the analysis in Ref. [?] does not take into account  
1185 the fundamental aspect of the BPGGSZ method: that it relies on the phase-space  
1186 distribution of signal decays, not phase-space integrated asymmetries. Furthermore,  
1187 the study only considers the  $CP$ -violation effect, not material interaction. Therefore,  
1188 a more detailed study was necessary before the  $B^\pm \rightarrow D \pi^\pm$  decay mode could  
1189 reliably be promoted to a signal channel.

1190

1191

### 4.1 $CP$ violation and material interaction of neutral kaons

1192

1193

A brief review of the general phenomenology of mixing and  $CP$  violation in the neutral kaon system is useful, before analysing the impact on  $\gamma$  measurements.

<sup>1194</sup> The presentation in this section follows the PDG review of *CP violation in the*  
<sup>1195</sup> *quark section [?]. The general theory considers any pair of neutral mesons*  $|M^0\rangle$   
<sup>1196</sup> *and*  $|\bar{M}^0\rangle$  *related by*  $CP$  *conjugation*

$$CP|M^0\rangle = e^{i\phi_M}|M^0\rangle \quad CP|\bar{M}^0\rangle = e^{-i\phi_M}|\bar{M}^0\rangle, \quad (4.1a)$$

<sup>1197</sup> where  $\phi_M$  is an arbitrary phase. In this thesis, the convention  $\phi_M = 0$  is chosen  
<sup>1198</sup> to equal zero, so that

$$CP|M^0\rangle = |\bar{M}^0\rangle \quad CP|\bar{M}^0\rangle = |M^0\rangle. \quad (4.1b)$$

<sup>1199</sup> A meson state that starts as a general superposition of  $|M^0\rangle$  and  $|\bar{M}^0\rangle$  states

$$\begin{aligned} \psi_M^0 &\equiv \psi_M(0) = a(0)|M^0\rangle + b(0)|\bar{M}^0\rangle \\ &\equiv \psi_{M^0}^0 + \psi_{\bar{M}^0}^0 \end{aligned} \quad (4.2)$$

<sup>1200</sup> will, over time, involve into a state that consists of a different superposition of  
<sup>1201</sup>  $|M^0\rangle$  and  $|\bar{M}^0\rangle$ , as well as components for all possible states the meson system  
<sup>1202</sup> can decay into

$$\begin{aligned} \psi_M(t) &= a(t)|M^0\rangle + b(t)|\bar{M}^0\rangle + \sum c_i(t)f_i \\ &\equiv \psi_{M^0}(t) + \psi_{\bar{M}^0}(t) + \sum c_i(t)f_i. \end{aligned} \quad (4.3)$$

<sup>1203</sup> For time scales that are longer than the typical strong-interaction, the time evolution  
<sup>1204</sup> of the  $M^0 - \bar{M}^0$  superposition can be described by a  $2 \times 2$  Hamiltonian

$$\frac{d}{dt} \begin{pmatrix} \psi_{M^0}(t) \\ \psi_{\bar{M}^0}(t) \end{pmatrix} = -i\mathcal{H}_0 \begin{pmatrix} \psi_{M^0}(t) \\ \psi_{\bar{M}^0}(t) \end{pmatrix} \quad (4.4)$$

<sup>1205</sup> that is *non-Hermitian* (to allow for decay) but can be parameterised in terms  
<sup>1206</sup> of two Hermitian matrices  $\mathcal{M}$  and  $\Gamma_0$

$$\mathcal{H}_0 = \mathcal{M} - \frac{i}{2}\Gamma_0. \quad (4.5)$$

<sup>1207</sup> The quantum states with well-defined (real) masses,  $m_j$ , and (real) decay widths,  
<sup>1208</sup>  $\Gamma_j$ , are the two eigenstates of  $\mathcal{H}_0$  with eigenvalues  $\lambda_j = m_j - \frac{i}{2}\Gamma_j$ . The eigenstates  
<sup>1209</sup> (of course) evolve independently in time, so that

$$\psi_j(t) = e^{-i\lambda_j t}\psi_j^0 = e^{-im_j t - \frac{\Gamma_j}{2}t}\psi_j^0. \quad (4.6)$$

<sup>1210</sup> The eigenstates are denoted  $H$  and  $L$  according to the size of  $m_j$ , the real part  
<sup>1211</sup> of the eigenvalues, such that  $m_H > m_L$ . Assuming that  $\mathcal{H}_0$  conserves  $CPT$  the  
<sup>1212</sup> eigenstates have the general form

$$\begin{aligned} |M_H\rangle &\equiv p|M^0\rangle - q|\bar{M}^0\rangle \\ |M_L\rangle &\equiv p|M^0\rangle + q|\bar{M}^0\rangle \end{aligned} \quad (4.7)$$

where  $p$  and  $q$  are complex numbers that satisfy  $|q|^2 + |p|^2 = 1$ . With the convention in Eq. (4.1b) it follows that if  $\mathcal{H}_0$  also conserves  $CP$ , so that  $|M_H\rangle$  and  $|M_L\rangle$  are  $CP$  eigenstates, then  $p = \pm q$ , where the sign depends on which of the heavy and the light meson states is  $CP$  even, and which is  $CP$  odd.

The eigenstates of the Hamiltonian governing the neutral kaon system are almost, but not exactly, equal to the  $CP$  eigenstates

$$|K_1\rangle = \frac{|K^0\rangle + |\bar{K}^0\rangle}{\sqrt{2}} \quad |K_2\rangle = \frac{|K^0\rangle - |\bar{K}^0\rangle}{\sqrt{2}}, \quad (4.8)$$

which are  $CP$  even and odd, respectively. This approximate equality leads to the most prominent feature of the neutral kaon system: the two eigenstates of  $\mathcal{H}_0$  have lifetimes that differ by orders of magnitude. This is best understood by assuming, for a moment, that the states in Eq. (4.8) *do* equal the eigenstates with definite life times. The  $K_1$  state can decay in the  $CP$  even  $\pi^+\pi^-$  and  $\pi^0\pi^0$  modes, and does so almost 100% of the time; these decay modes are not available to the  $K_2$  (in the absence of direct  $CP$  violation) which results in a much lower decay rate and much longer life time. Therefore, the eigenstates in the kaon system are labelled the *short-lived* kaon,  $K_S^0$ , which is almost  $CP$  even, and the *long-lived* kaon,  $K_L^0$ , which is almost  $CP$  odd. The life times are [?]

$$\tau_{K_S^0} = (8.954 \pm 0.004) \times 10^{-11} \text{s} \quad \tau_{K_L^0} = (5.116 \pm 0.021) \times 10^{-8} \text{s}. \quad (4.9)$$

Experimentally, it is found that the  $K_S^0$  corresponds to the light eigenstate, but that the mass splitting [?]

$$\begin{aligned} \Delta m = m_{K_L^0} - m_{K_S^0} &= (0.5289 \pm 0.0009) \times 10^{10} \hbar s^{-1} \\ &\simeq 3.5 \times 10^{-6} \text{ eV} \end{aligned} \quad (4.10)$$

is tiny compared to the neutral kaon masses of  $m_{K_S^0} = 497.6 \text{ MeV}/c^2$  [?].

However, the discovery of  $K_L^0 \rightarrow \pi\pi$  decays by Kronin and Fitch in 1964 established that the  $K_S^0$  and  $K_L^0$  are *not* exactly equal to the  $CP$  eigenstates in Eq. (4.8), because the  $\mathcal{H}_0$  relevant to the kaon system is  $CP$ -violating. The  $CP$  violation in the kaon sector is conventionally parameterised in terms of the complex parameters  $\epsilon$  and  $\epsilon'$ , in terms of which

$$\frac{A(K_L^0 \rightarrow \pi^+\pi^-)}{A(K_S^0 \rightarrow \pi^+\pi^-)} = \epsilon + \epsilon' \quad \frac{A(K_L^0 \rightarrow \pi^0\pi^0)}{A(K_S^0 \rightarrow \pi^0\pi^0)} = \epsilon - 2\epsilon'. \quad (4.11)$$

In these expressions  $\epsilon$  denotes the contribution from  $CP$  violation in mixing and  $\epsilon'$  the contribution due to direct  $CP$  violation in the decays. The  $\epsilon$  parameter has been measured to be [?]

$$|\epsilon| = (2.228 \pm 0.011) \times 10^{-3}, \quad \arg \epsilon = (43.52 \pm 0.05)^\circ. \quad (4.12)$$

1240 Direct  $CP$  violation is ignored for the remainder of the thesis, because  $\epsilon'$  is measured  
1241 to be three orders of magnitude smaller than  $\epsilon$ . In terms of the  $CP$  eigenstates  
1242 of Eq. (4.8), the mass eigenstates  $K_S^0$  and  $K_L^0$  are given by

$$\begin{aligned} |K_S^0\rangle &= \frac{|K_1\rangle + \epsilon|K_2\rangle}{\sqrt{1+|\epsilon|^2}} &= \frac{(1+\epsilon)|K^0\rangle + (1-\epsilon)|\bar{K}^0\rangle}{\sqrt{2(1+|\epsilon|^2)}} \\ |K_L^0\rangle &= \frac{|K_2\rangle + \epsilon|K_1\rangle}{\sqrt{1+|\epsilon|^2}} &= \frac{(1+\epsilon)|K^0\rangle - (1-\epsilon)|\bar{K}^0\rangle}{\sqrt{2(1+|\epsilon|^2)}}, \end{aligned} \quad (4.13)$$

1243 corresponding to the definition  $p = (1+\epsilon)/\sqrt{2(1+|\epsilon|^2)}$  and  $q = (1-\epsilon)/\sqrt{2(1+|\epsilon|^2)}$   
1244 in Eq. (4.7).

1245 In an experimental setting, the time evolution of a neutral kaon state is affected  
1246 by nuclear interactions with the detector. The interaction is governed by the strong  
1247 force, and therefore sensitive to the *flavour* of the kaon state; the interaction  
1248 strength is thus different for  $K^0$  and  $\bar{K}^0$  mesons. This difference introduces a  
1249 non-zero  $K_S^0 \leftrightarrow K_L^0$  transition amplitude for neutral kaons traversing a detector  
1250 segment. This effect was predicted early in the history of kaon physics [?]  
1251 and is commonly denoted *kaon regeneration*. The effect can be described by including a  
1252 material-interaction term in the Hamiltonian that is diagonal in the  $(|K^0\rangle, |\bar{K}^0\rangle)$   
1253 basis, so that the equation governing the time evolution is [?, ?]

$$\frac{d}{dt} \begin{pmatrix} \psi_{K^0}(t) \\ \psi_{\bar{K}^0}(t) \end{pmatrix} = -i \left[ \mathcal{H}_0 + \begin{pmatrix} \chi & 0 \\ 0 & \bar{\chi} \end{pmatrix} \right] \begin{pmatrix} \psi_{K^0}(t) \\ \psi_{\bar{K}^0}(t) \end{pmatrix}. \quad (4.14)$$

1254 The complex parameters  $\chi$  and  $\bar{\chi}$  describe the material interaction of the  $K^0$   
1255 and  $\bar{K}^0$  flavour eigenstates and are related to their scattering cross section, as  
1256 described further in Section 4.3.4. The solution of Eq. (4.14) for the time evolution  
1257 in the  $K_S^0$  and  $K_L^0$  states is [?]

$$\begin{aligned} \psi_S(t) &= e^{-i\Sigma t} \left( \psi_S^0 \cos \Omega t + \frac{i}{2\Omega} (\Delta\lambda\psi_S^0 - \Delta\chi\psi_L^0) \sin \Omega t \right), \\ \psi_L(t) &= e^{-i\Sigma t} \left( \psi_L^0 \cos \Omega t - \frac{i}{2\Omega} (\Delta\lambda\psi_L^0 + \Delta\chi\psi_S^0) \sin \Omega t \right), \end{aligned} \quad (4.15)$$

1258 in terms of the parameters

$$\begin{aligned} \Delta\chi &= \chi - \bar{\chi}, \\ \Delta\lambda &= \lambda_L - \lambda_S = (m_L - m_S) - \frac{i}{2}(\Gamma_L - \Gamma_S), \\ \Sigma &= \frac{1}{2}(\lambda_S + \lambda_L + \chi + \bar{\chi}), \\ \Omega &= \frac{1}{2}\sqrt{\Delta\lambda^2 + \Delta\chi^2}. \end{aligned} \quad (4.16)$$

1259 In the vacuum limit where  $\chi = \bar{\chi} = 0$ , the expressions in Eq. (4.6) and Eq. (4.15) are  
1260 equal.

### 4.1.1 A first look at the impact on $\gamma$ measurements

The effects described above have an impact on measurements of  $CP$  asymmetries in modes with a neutral kaon in the final state. This was analysed for the first time in relation to  $\gamma$  measurements by Grossman and Savastio in 2014 [?]. The authors point out two sources of corrections to be included:

- the fact that  $K_S^0$  is not an exact  $CP$  eigenstate can break potential symmetry relations employed in an analysis, and
- that when the neutral kaon is reconstructed in a  $\pi\pi$  final state there will be contributions from both  $K_S^0$  and  $K_L^0$  decays.

The analysis in this chapter considers yet another effect, not treated by Grossman and Savastio, namely that

- material interaction can emulate the effect of neutral kaon  $CP$  violation, because it couples the almost- $CP$ -even  $K_S^0$  and the almost- $CP$ -odd  $K_L^0$  states.

Due to the presence of  $K_L^0 \rightarrow \pi\pi$  decays, Grossman and Savastio point out that the relevant decay rates to consider in an experimental setting are of the form

$$d\Gamma(t) \propto |\psi_S(t) + \epsilon\psi_L(t)|^2. \quad (4.17)$$

The time dependence of the decay rates considered in Chapter 2 was left out because all terms shared a common time dependence. That is not the case in Eq. (4.17), due to the very different decay rates of the  $K_S^0$  and  $K_L^0$  components of the kaon state. As a consequence, the time-integrated yields have the form

$$N \propto \int dt \eta(t) |\psi_S(t) + \epsilon\psi_L(t)|^2, \quad (4.18)$$

where  $\eta(t)$  is the time acceptance in a given experimental setting. Thus, the acceptance is crucial to model in order to correctly estimate the impact of kaon  $CP$ -violation effects on a given measurement.

Considering BPGBSZ measurements, the main effect of neutral kaon  $CP$  violation is a breakdown of the fundamental Dalitz-plot symmetry that is exploited in the derivation of the bin yield equations. Extending the amplitude definition of Eq. (2.21) to include  $K_L^0$  decays

$$A_{S(L)}^{(\overline{D})}(s_-, s_+) = A((\overline{D})^0 \rightarrow K_{S(L)}^0 \pi^+ \pi^-), \quad (4.19)$$

the authors point out that  $CP$ -violation in the  $K_S^0$  system means that the relation  $A_S^{(\overline{D})}(s_{-+}) = A_S^D(s_{+-})$  is not exactly true; and in addition, there is now a

dependence on  $A_L^D(s_{-+})$  which satisfies a different approximate symmetry, namely  $A_L^{\bar{D}}(s_{-+}) \simeq -A_L^D(s_{+-})$ . Grossman and Savastio describe these symmetry breaking effects in detail, but do not explicitly derive the corrections to the yield equations of Chapter 2, nor try to quantify the potential bias on  $\gamma$  in a measurement based on the binned yields. Instead, they derive expressions for the bias in a measurement obtained from phase-space integrated  $CP$  asymmetries. This is done for both GLW measurements that use  $D \rightarrow K_S^0 X$  final states and for the  $D \rightarrow K_S^0 h^+ h^-$  final states; however, for their quantitative estimate of  $\Delta\gamma$  the authors make an approximation that corresponds to assuming that the  $D \rightarrow K_S^0 h^+ h^-$  final state is a  $CP$  eigenstate, making the two results identical. The authors find that in this case, assuming a uniform experimental acceptance for all kaon decay times, the asymmetry has the form<sup>1</sup>

$$A = \frac{2r_B \sin \gamma \sin \delta_B + 2\text{Re}[\epsilon]}{1 + r_B^2 - 2r_B \cos \gamma \cos \delta_B}, \quad (4.20)$$

If a measured value of  $A$  is interpreted to obtain  $\gamma$  without taking the  $\epsilon$  term into account, it leads to a bias of

$$\Delta\gamma = -\frac{\text{Re}[\epsilon]}{r_B \cos \gamma \sin \delta_B} + O(|\epsilon|). \quad (4.21)$$

The scaling  $\Delta\gamma \sim \mathcal{O}(r_B/|\epsilon|)$  is the main result of the analysis by Grossman and Savastio. For  $B^\pm \rightarrow D K^\pm$  decays, where  $r_B^{DK^\pm} \simeq 0.1$  this suggests a bias at the percent level, which is negligible compared to current experimental uncertainties. However, in the  $B^\pm \rightarrow D \pi^\pm$  case, where  $r_B^{D\pi^\pm} \simeq 0.005$  [?], their result suggests relative biases that are potentially of  $\mathcal{O}(1)$ .

The conclusions are lacking on two accounts, however. Firstly, as made clear in Section 2.3.5, the  $K_S^0 \pi^+ \pi^-$  and  $K_S^0 K^+ K^-$  states are *far from*  $CP$  eigenstates. From the asymmetry expression in that section, it is clear that the bias in a determination of  $\gamma$  based on phase-space asymmetries will in fact scale as

$$\Delta\gamma \sim \mathcal{O}\left(\frac{|\epsilon|}{(2\mathcal{F}_+ - 1)r_B}\right), \quad (4.22)$$

which suggests that Grossman and Savastio severely *underestimates* the potential impact. This is described in detail in Section 4.2.3. More importantly, the analysis of the phase-space integrated asymmetry is in fact *irrelevant* to BPGGSZ measurements as they are currently performed: as described in Section 2.3.5 the information from

---

<sup>1</sup>In fact the expression in Eq. (4.20) is missing a term, as will be clear when an analogous expression is derived in detail in Section 4.2.3.

the global asymmetry is completely discarded. Therefore it is necessary to analyse the effects of kaon  $CP$ -violation on a full, binned analysis of  $D \rightarrow K_S^0 h^+ h^-$  decays, which is done in detail in the following sections. While the aim is to extend the analysis if Grossman and Savastio, the treatment in the following sections is completely independent of that in Ref. [?].

## 4.2 Impact on BPGBSZ measurements of $\gamma$ : principles

The analysis of the impact on BPGBSZ measurements is carried out in two stages. This section treats the leading order effects analytically, and derives the overall order of magnitude of the expected bias in a general setting. Then Section 4.3 presents a detailed numerical study of the expected effect in measurements with the LHCb and Belle II experiments specifically, because these will be crucial to constrain  $\gamma$  during the coming decade [?, ?].

### 4.2.1 Modified symmetry relations

In order to derive the corrections to the asymmetry relation  $A_S^D(s_{-+}) \simeq A_S^{\bar{D}}(s_{+-})$ , it is beneficial to express  $A_{S(L)}^D$  in terms of the amplitudes

$$A_{1/2}^{\bar{D}} = A(\overline{D}^0 \rightarrow K_{1/2}^0 \pi^+ \pi^-), \quad (4.23)$$

because these amplitude satisfy the exact symmetries  $A_1^D(s_{-+}) = A_1^{\bar{D}}(s_{+-})$  and  $A_2^D(s_{-+}) = -A_2^{\bar{D}}(s_{+-})$ . This approach is different to that of Grossman and Savastio, but the final results are equivalent. After the decay of a  $D^0$  meson to a neutral kaon, the kaon state is

$$\begin{aligned} \psi^0 &= A_1^D |K_1\rangle + A_2^D |K_2\rangle \\ &= N \left[ (A_1^D - \epsilon A_2^D) |K_S^0\rangle + (A_2^D - \epsilon A_1^D) |K_L^0\rangle \right], \end{aligned} \quad (4.24)$$

with the normalisation constant  $N = \sqrt{1 + |\epsilon|^2}/(1 - \epsilon^2)$ . Thus it can be seen that

$$\begin{aligned} A_S^D(s_{-+}) &= N \left[ (A_1^D(s_{-+}) - \epsilon A_2^D(s_{-+})) \right], \\ A_L^D(s_{-+}) &= N \left[ (A_2^D(s_{-+}) - \epsilon A_1^D(s_{-+})) \right], \end{aligned} \quad (4.25)$$

1337 with an analogous expression for the  $\bar{D}^0$  decay amplitudes. Therefore, the generalised  
1338 relations between the  $D^0$  and  $\bar{D}^0$  amplitudes are

$$\begin{aligned} A_S^{\bar{D}}(s_{+-}) &= N[A_1^{\bar{D}}(s_{+-}) - \epsilon A_2^{\bar{D}}(s_{+-})] \\ &= N[A_1^D(s_{-+}) + \epsilon A_2^D(s_{-+})] = A_S^D(s_{-+}) + 2N\epsilon A_2^D(s_{-+}), \\ A_L^{\bar{D}}(s_{+-}) &= N[A_2^{\bar{D}}(s_{+-}) - \epsilon A_1^{\bar{D}}(s_{+-})] \\ &= -N[A_2^D(s_{-+}) + \epsilon A_1^D(s_{-+})] = -A_L^D(s_{-+}) - 2N\epsilon A_1^D(s_{-+}). \end{aligned} \quad (4.26)$$

### 1339 4.2.2 Relationship between the $K_S^0$ and $K_L^0$ amplitudes

1340 The decay amplitude  $A(D^0 \rightarrow K_S^0 \pi^+ \pi^-)$  has been carefully studied, and a number  
1341 of amplitude models have been published [?, ?, ?, ?, ?]. No models have been  
1342 published for  $D^0 \rightarrow K_L^0 \pi^+ \pi^-$  decays. However, following an approach laid out by  
1343 the CLEO collaboration [?], the two amplitudes can be related. Again, this is  
1344 most easily done by relating the  $A_1^D(s_{+-})$  and  $A_2^D(s_{+-})$  amplitudes. In the isobar  
1345 formalism, the decay amplitude  $A(D^0 \rightarrow K_1 \pi^+ \pi^-)$  is expressed as a non-resonant  
1346 constant amplitude plus a sum of resonances

$$A(D^0 \rightarrow K_1 \pi^+ \pi^-) = k_{NR} + \sum_{CF} k_i R^i(s_{K\pi^-}) + \sum_{DCS} k_j R^j(s_{K\pi^+}) + \sum_{R_{\pi\pi}} k_k R^k(s_{\pi^+\pi^-}). \quad (4.27)$$

1347 The resonances are split into Cabibbo-favoured (CF)  $K^{*-}$  resonances, doubly  
1348 Cabibbo-suppressed (DCS)  $K^{*+}$  resonances and  $\pi\pi$  resonances.<sup>2</sup> The CF resonances  
1349 couple to the  $\bar{K}^0$  component of  $K_1 (\propto K^0 + \bar{K}^0)$ , and therefore the corresponding  
1350  $k_i$  in the  $K_2 (\propto K^0 - \bar{K}^0)$  amplitude will have a relative minus sign. The DCS  
1351 resonances couple to the  $K^0$  component of  $K_1$ , and so the corresponding  $k_j$  in the  
1352  $K_2$  amplitude will have a relative plus sign. For the  $h^+ h^-$  resonances, there will be  
1353 a coupling to both the  $K^0$  and  $\bar{K}^0$  components, however the coupling to the  $K^0$   
1354 component is expected to be suppressed with a Cabibbo suppression factor  $r_k e^{i\delta_k}$ ,  
1355 where  $r_k \simeq \tan^2 \theta_C \simeq 0.05$  is determined by the Cabibbo angle  $\theta_C$  and  $\delta_k$  can take  
1356 any value. Therefore, the  $k_k$  for these resonances have a relative  $-(1 - 2r_k e^{i\delta_k})$   
1357 factor in the  $K_2$  amplitude. The same effect leads to the differences in decay rates  
1358 between  $D^0 \rightarrow K_S^0 \pi^0$  and  $D^0 \rightarrow K_L^0 \pi^0$  decays [?, ?]. Thus, given a model of the

---

<sup>2</sup>In modern models, the  $\pi\pi$  and  $K\pi$  S-wave components are modelled via the  $K$ -matrix formalism and LASS parametrisations, respectively, instead of sums of individual resonances [?]. This does not alter the arguments below, as the  $R$  functions of Eq. (4.27) can equally well represent such terms.

<sup>1359</sup> form in Eq. (4.27), a model for the  $A(D^0 \rightarrow K_2\pi^+\pi^-)$  amplitude will have the form

$$A(D^0 \rightarrow K_2\pi^+\pi^-) = k_{NR} + \sum_{CF} (-k_i) R^i(s_{K\pi^-}) + \sum_{DCS} (+k_j) R^j(s_{K\pi^+}) + \sum_{R_{\pi\pi}} (-(1 - 2r_k e^{i\delta_k}) k_k) R^k(s_{\pi^+\pi^-}). \quad (4.28)$$

<sup>1360</sup> An important consequence of these substitution rules is that

$$A_2^D(s_{+-}) = -A_1^D(s_{+-}) + r_A \Delta A(s_{+-}), \quad (4.29)$$

<sup>1361</sup> where  $r_A \simeq \tan^2 \theta_C$  and  $\Delta A(s_{+-}) \sim A_1^D(s_{+-})$  are of the same order of magnitude  
<sup>1362</sup> (at least when averaged over the bins used in  $\gamma$  measurements). This relation is  
<sup>1363</sup> sufficient to make the qualitative arguments of this section, while the full set of  
<sup>1364</sup> substitution rules above are used in the quantitative studies of Section 4.3.

### <sup>1365</sup> 4.2.3 Modification of the BPGBSZ yield equations

<sup>1366</sup> With suitable models to calculate  $A_{S(L)}^{\bar{D}}$  (or  $A_{1/2}^{\bar{D}}$ ) and knowledge of  $\Delta\chi$  for the  
<sup>1367</sup> materials relevant to an experimental setting, the relations derived in the preceding  
<sup>1368</sup> sections can be employed to calculate the expected phase-space bin yields,  $N_i^\pm$ ,  
<sup>1369</sup> including the effects of kaon  $CP$  violation and material interaction. The decay  
<sup>1370</sup> rates have additional terms compared to those in Eq. (2.24), because the  $K_L^0$   
<sup>1371</sup> contribution must be taken into account

$$d\Gamma(t, s_{+-}) \propto |\psi_S(t, s_{+-}) + \epsilon\psi_L(t, s_{+-})|^2, \quad (4.30)$$

<sup>1372</sup> where the time-dependence of  $\psi_{S/L}(t, s_{+-})$  is governed by Eq. (4.15), and the phase-  
<sup>1373</sup> space dependence is included in the state component, by defining  $\psi_{S/L}^0$  in terms of  
<sup>1374</sup>  $A_{S(L)}^{\bar{D}}(s_{+-})$ . For example, for the case of a  $B^- \rightarrow DK^-$  decay, the definition is

$$\begin{aligned} \psi_{S/L}^{0,B^-}(s_{+-}) &= A_S^D(s_{+-}) + r_B e^{i(\delta_B - \gamma)} A_S^{\bar{D}}(s_{+-}) \\ &= A_1^D(s_{+-}) - \epsilon A_2^D(s_{+-}) + r_B e^{i(\delta_B - \gamma)} (A_1^{\bar{D}}(s_{+-}) - \epsilon A_2^{\bar{D}}(s_{+-})) \\ &= A_1^D(s_{+-}) - \epsilon A_2^D(s_{+-}) + r_B e^{i(\delta_B - \gamma)} (A_1^D(s_{+-}) + \epsilon A_2^D(s_{+-})). \end{aligned} \quad (4.31)$$

<sup>1375</sup> It is useful to look at the corrections to the BPGBSZ yield expressions in Eq. (2.25)  
<sup>1376</sup> to lowest order in  $\epsilon$  and  $r_\chi = \frac{1}{2} \frac{\Delta\chi}{\Delta\lambda}$ , the dimensionless parameter governing material  
<sup>1377</sup> interactions. For LHCb and Belle II the average  $|r_\chi| \simeq 10^{-3}$ , as detailed in the  
<sup>1378</sup> Section 4.3. To first order in  $r_\chi$ , the time-dependent kaon states within a material,  
<sup>1379</sup> given in Eq. (4.15), simplify to [?]

$$\begin{aligned} \psi_S(t, s_{+-}) &= e^{-\frac{i}{2}(x+\bar{x})t} e^{-i\lambda_S t} (\psi_S^0(s_{+-}) - r_\chi (1 - e^{-i\Delta\lambda t}) \psi_L^0(s_{+-})), \\ \psi_L(t, s_{+-}) &= e^{-\frac{i}{2}(x+\bar{x})t} e^{-i\lambda_L t} (\psi_L^0(s_{+-}) + r_\chi (1 - e^{+i\Delta\lambda t}) \psi_S^0(s_{+-})). \end{aligned} \quad (4.32)$$

1380 By inserting these expressions into Eq. (4.30) and employing the definition in  
1381 Eq. (4.31) (and a similar definition for  $B^+$  decays), the binned yields can be  
1382 calculated by an integration over time and phase space. In the remainder of this  
1383 section, it is assumed that the experimental time acceptance is  $\eta(t) = 1$  for all times  
1384 and that  $r_\chi$  is constant at all times; more realistic assumptions are introduced in  
1385 Section 4.3. In this case, the binned yields are given by the expression

$$N_i^- = h_B^{-'} \left( \hat{K}_{+i} + r_B^2 \hat{K}_{-i} + 2\sqrt{\hat{K}_{+i} \hat{K}_{-i}} (x_- \hat{c}_i + y_- \hat{s}_i) + O(r\epsilon) \right), \\ N_i^+ = h_B^{+'} \left( \hat{K}_{-i} + r_B^2 \hat{K}_{+i} + 2\sqrt{\hat{K}_{+i} \hat{K}_{-i}} (x_+ \hat{c}_i - y_+ \hat{s}_i) + O(r\epsilon) \right), \quad (4.33)$$

1386 where a number of new parameters have been defined, and where  $O(r\epsilon)$  denotes terms  
1387 of  $O(r_A\epsilon)$ ,  $O(r_B\epsilon)$ ,  $O(r_A r_\chi)$ , and  $O(r_B r_\chi)$ . Since  $r_B \sim r_A \sim 10^{-1}$  (in  $B^\pm \rightarrow D K^\pm$   
1388 decays) and  $r_\chi \sim \epsilon \sim 10^{-3}$ , these terms are all of the same order of magnitude.

1389 The new normalisation constants  $h_B^{\pm'} = h_B^\pm (1 + |\epsilon + r_\chi|^2 \frac{\Gamma_S}{\Gamma_L} \mp \Delta h)$  are de-  
1390 fined in terms of

$$\Delta h = 2\text{Re}[\epsilon + r_\chi] - 4 \frac{\Gamma_S}{\Gamma_L + \Gamma_S} \frac{\text{Re}[\epsilon + r_\chi] + \mu \text{Im}[\epsilon + r_\chi]}{1 + \mu^2}, \quad \mu = 2 \frac{m_L - m_S}{\Gamma_L + \Gamma_S}. \quad (4.34)$$

1391 The  $\hat{K}_i$  parameters are defined to be

$$\hat{K}_i = \frac{1}{1 + |\epsilon + r_\chi|^2 \frac{\Gamma_S}{\Gamma_L}} \left( K_i^{(1)} + |\epsilon + r_\chi|^2 \frac{\Gamma_S}{\Gamma_L} K_i^{(2)} \right), \quad (4.35)$$

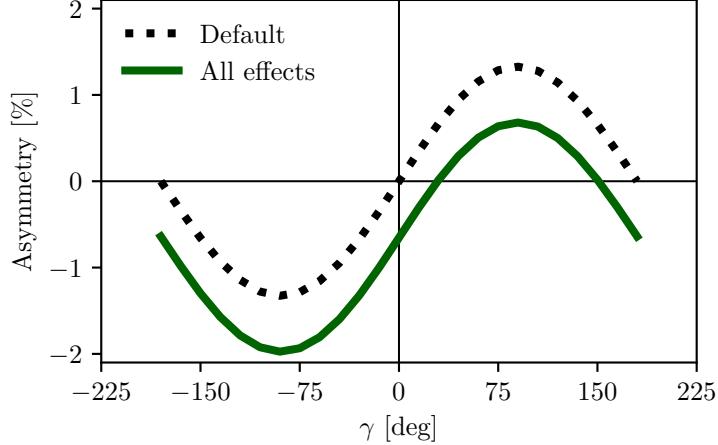
1392 in which the  $K_i^{(1/2)}$  parameters are phase-space integrals, defined as in Eq. (2.27)  
1393 but for  $A_{1/2}^D$ . To lowest order, the  $\hat{K}_i$  correspond to the fractional  $D^0$  decay yield  
1394 in each bin, as obtained in a measurement that averages  $D^0$  and  $\bar{D}^0$  decays, and  
1395 assumes the  $A_S^D(s_{-+}) = A_S^D(s_{+-})$  symmetry to be exact:

$$K_i^{\text{meas}} \equiv \frac{N_i^D + N_{-i}^{\bar{D}}}{\sum_j N_j^D + N_{-j}^{\bar{D}}} = \hat{K}_i + \mathcal{O}(r\epsilon). \quad (4.36)$$

1396 Here,  $N_i^D$  ( $N_i^{\bar{D}}$ ) is the expected yield of flavour tagged  $D^0$  ( $\bar{D}^0$ ) mesons into bin  
1397  $i$  of the  $D$  decay phase-space.

1398 In similar fashion, the parameters  $(\hat{c}_i, \hat{s}_i)$  have been introduced to denote the  
1399 *measured* average strong-phases, which are expected to differ from  $(c_i, s_i)$  at  $O(\epsilon)$ ,  
1400 since neutral kaon  $CP$  violation is not taken into account in the measurements  
1401 by CLEO. Thus, any corrections arising if  $(\hat{c}_i, \hat{s}_i)$  and  $(c_i, s_i)$  are substituted in  
1402 Eq. (4.33) will appear in the  $O(r_B\epsilon)$  terms.

1403 Two observations can be made from the expression in (4.33). The first is that  
1404 the phase-space *distribution* is only changed at  $O(r\epsilon)$  compared to the expression in



**Figure 4.1:** The asymmetry  $A_{\text{total}}$  as a function of  $\gamma$  calculated to  $O(\epsilon)$  using Eq. (4.37). The calculation is made using for (black dotted line) the default case where  $\Delta h = 0$  and (green) including neutral kaon  $CP$ -violation and material interaction with  $r_\chi = \epsilon$ .

Eq. (2.25), if the measured  $\hat{K}_i$  are used in the experimental analysis. This equally true whether the  $K_i$  are fitted in the signal channel along with  $x_\pm$  and  $y_\pm$ , as is the case in the measurement presented in the thesis, or if they are obtained in a control channel with flavour tagged  $D$  decays, according to Eq. (4.36). As the  $D^0 - \bar{D}^0$  interference term that provides sensitivity to  $\gamma$  enters at order  $O(r_B)$ , the impact on  $\gamma$  measurements can be expected to be  $\Delta\gamma/\gamma \sim O(r\epsilon/r_B)$ . For  $B \rightarrow DK$  analyses, where  $r_B \simeq 0.1$ , this is at the permille level, so the induced  $\Delta\gamma$  bias can be expected to be smaller than  $1^\circ$ . Even in the case of  $B^\pm \rightarrow D\pi^\pm$  decays, this suggests biases that are maximally a few percent. This is the main result of the chapter, because it means that the effect of neutral kaon  $CP$  violation and material interaction is small compared to the precision of the measurement that is the main subject of the thesis.

The second observation relates to potential future measurements of  $\gamma$ , which may also include sensitivity from the total, phase-space-integrated yield asymmetry

$$A_{\text{total}} = \frac{N^- - N^+}{N^- + N^+} = \frac{2(2\mathcal{F}_+ - 1)r_B \sin \delta_B \sin \gamma + \Delta h}{1 + r_B^2 + 2(2\mathcal{F}_+ - 1)r_B \cos \delta_B \cos \gamma} + O(r\epsilon), \quad (4.37)$$

where the definition of  $\mathcal{F}_+$  from Section 2.3.5 has been employed. In the limit  $r_B \rightarrow 0$  the expression agrees with the result for the analogous asymmetry in  $D^\pm \rightarrow \pi^\pm K_S^0$  decays in Ref. [?], evaluated to  $O(\epsilon)$  for an infinite and uniform time-acceptance. As hinted at above, the fact that  $\mathcal{F}_+ \simeq 0.5$  means that the asymmetry due to  $\gamma$  being non-zero is not  $\mathcal{O}(r_B)$ , but of approximately the same order of magnitude as the asymmetry due to  $CP$  violation in the neutral kaon sector, governed by  $\Delta h$ . This is illustrated in Fig. 4.1, where the expression in Eq. (4.37) is plotted in the default

case where  $\Delta h = 0$ , using the model in Ref. [?] to calculate  $K_i$  and  $c_i$ , as well as including neutral kaon  $CP$  violation and material interaction effects, calculated using  $r_\chi = \epsilon$ , with  $\epsilon$  taking the value in Eq. (4.12). The asymmetry changes significantly when including the latter effects. Therefore, measurements based only on the global asymmetry will suffer relative biases of tens of degrees, not a few degrees, if neutral kaon  $CP$  violation and material interaction is not taken into account.

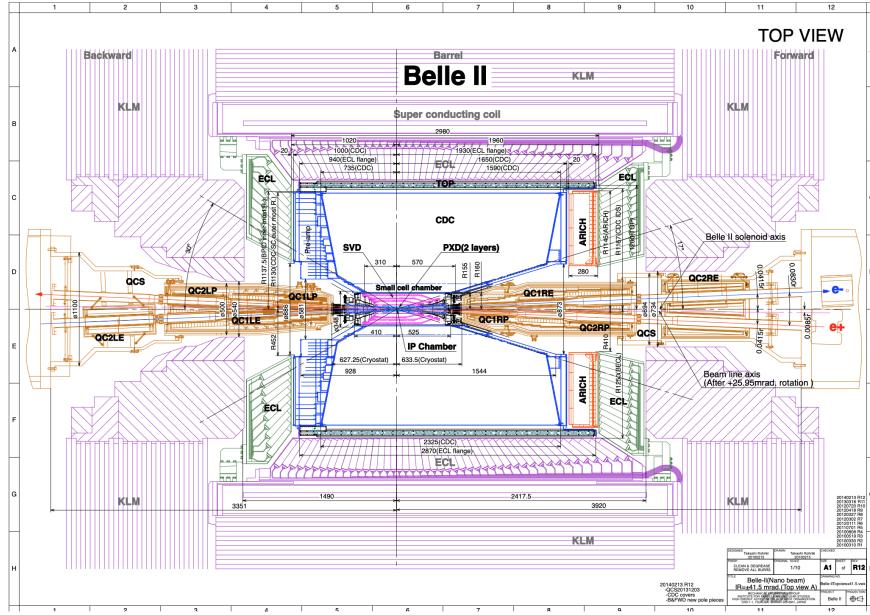
### 4.3 Impact on BPGBSZ measurements of $\gamma$ : LHCb and Belle II measurements

The previous section has established that the bias due to neutral kaon  $CP$  violation and material interaction is at the sub-percent level for measurements based on  $B^\pm \rightarrow DK^\pm$  decays, and just a few percent in  $B^\pm \rightarrow D\pi^\pm$  decays. Thus, the effects only contribute a manageable systematic uncertainty in the measurement that is the subject of the thesis. However, the expected precision on  $\gamma$  measurements will increase significantly in the coming decade, as both the LHCb [?] and Belle II [?] collaborations expect to make BPGBSZ measurements that measure  $\gamma$  with a precision of 1–3°. Therefore a deeper understanding of the expected bias for these specific experiments is important.

This section details a study, where the equations of the previous section are evaluated numerically to all orders, and care is taken to realistically model the experiment specific conditions. The scope of the original analysis, published in Ref. [?], was a stand-alone paper that covers both LHCb and Belle II, and which therefore does not rely on full detector simulation. Instead the following approaches are taken to model the necessary input

- the experimental time-acceptance is modelled based on the detector geometry and typical neutral kaon momentum spectrum
- the material interaction is included, using the material budget information available in the technical design reports on each experiment
- both the time-acceptance and material interaction depends on the neutral kaon momentum, for which realistic distributions are estimated using the `RapidSim` simulation package [?].

Each input is described in detail in the following sections. The study has been repeated to assign a systematic uncertainty to the LHCb measurement in Chapter 5, with slight adjustments to match the exact fit setup and with the inputs above extracted from full LHCb simulation. This is described further in Section 4.3.7.

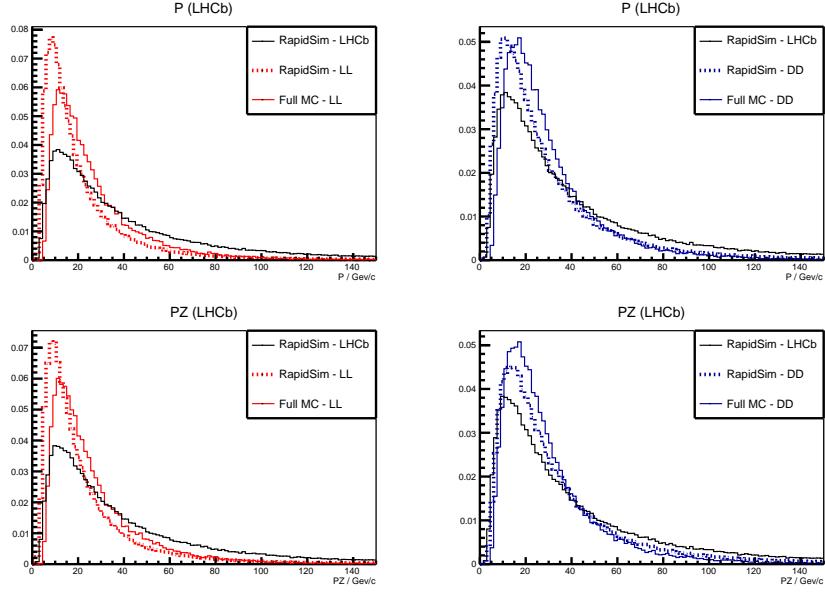


**Figure 4.2:** Schematic of the Belle II detector, reproduced from Ref. [?].

### 1459 4.3.1 Detector geometries

1460 The LHCb geometry and sub detectors are described in details in Chapter 3. In the  
 1461 LHCb measurement discussed in Chapter 5, the  $K_S^0$  mesons are reconstructed in  
 1462 the  $\pi^+\pi^-$  final state and two distinct categories of decay are considered, depending  
 1463 on where in the detector the  $K_S^0$  decay occurs. The categories have very different  
 1464 decay-time acceptance, and therefore two scenarios are considered for LHCb: one  
 1465 in which the decay products of the  $K_S^0$  leave reconstructed tracks in both the silicon  
 1466 vertex detector and downstream tracking detectors (denoted *long-long* or LL), and  
 1467 one in which the decay products of the  $K_S^0$  only leave tracks in the downstream  
 1468 tracking detectors (denoted *down-down* or DD).

1469 The Belle II detector is a general purpose spectrometer, built to collect  
 1470 data from asymmetric  $e^+e^-$  collisions provided by the SuperKEKB accelerator in  
 1471 Japan [?]. A schematic of the detector is shown in Fig. 4.2. The relevant sub  
 1472 detectors for the present study are the tracking detectors: a central silicon vertex  
 1473 detector, comprised of a total of six layers within 140 mm of the beam, and a large  
 1474 volume drift chamber with 56 wire layers, extending to a radius of 1130 mm [?].  
 1475 A single scenario is considered for Belle II, because essentially all the  $K_S^0$  mesons  
 1476 produced in signal decays in Belle II decay within the tracking volume, with more  
 1477 than 90 % decaying in the vertex detector according to the studies described below.  
 1478 Thus, three scenarios are considered in total: LL LHCb, DD LHCb, and Belle II.



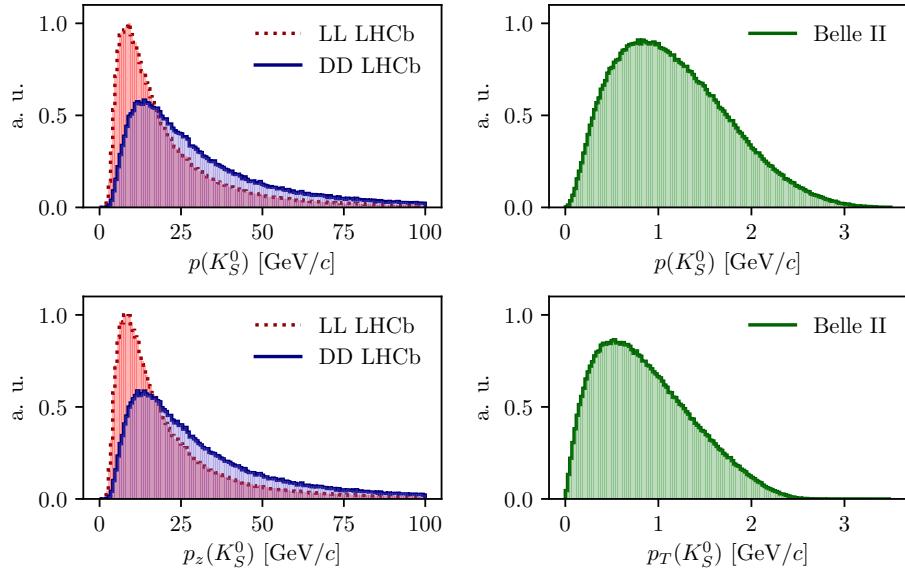
**Figure 4.3:** Momentum spectra for the  $K_S^0$  meson in LHCb, as generated using **RapidSim** (black lines) directly, as well as reweighted to match decay time acceptance in the (red) LL and (blue) DD data categories of LHCb. The LHCb spectra are compared with the spectra in fully simulated signal decays, for both (dotted red lines) LL and (dotted blue lines) DD data categories.

### 1479 4.3.2 Kaon momentum distributions

1480 The neutral kaon momentum distributions are obtained using **RapidSim** [?], a simple  
1481 tool to generate MC samples. **RapidSim** has an inbuilt capability to generate decays  
1482 of  $B$  mesons with the kinematic distribution found in LHCb collisions and falling in  
1483 the LHCb acceptance. However, the distributions need to be reweighted to take the  
1484 kaon-decay-time acceptance into account. After being reweighted, the **RapidSim**  
1485 momentum spectra are reasonably close to those found in full LHCb simulation  
1486 samples of  $B^\pm \rightarrow D(\rightarrow K_S^0\pi^+\pi^-)K^\pm$  decays, as seen in Fig. 4.3

1487 At Belle II, the signal  $B$  mesons stem from decays of  $\Upsilon(4S)$  mesons produced in  
1488 asymmetric electron-positron collisions. This leads to substantially different decay  
1489 kinematics in comparison to those found at LHCb. The momentum distribution in  
1490 Belle II is estimated by letting **RapidSim** decay  $B$  mesons with a momentum of 1.50  
1491 GeV/ $c$  along the  $z$ -axis using **RapidSim**, corresponding to the  $\gamma\beta = 0.28$  boost of  
1492 the centre-of-mass system in Belle II when operated at the  $\Upsilon(4S)$  resonance [?]. A  
1493 perfect  $4\pi$  angular acceptance is assumed. It is not necessary to reweigh the Belle II  
1494 momentum spectrum to account for the kaon-decay-time acceptance because all  
1495 produced  $K_S^0$  mesons decay in the tracking volume.

1496 The resulting momentum distributions for the three types of sample are shown in  
1497 Fig. 4.4.



**Figure 4.4:** Momentum distributions for the LHCb (red dotted line) LL and (blue) DD categories, as well as (green) Belle II, obtained using `RapidSim`.

### 4.3.3 Experimental time acceptance

In order to model the experimental time acceptance, the time-dependent decay rates are only integrated over a finite time interval  $(\tau_1, \tau_2)$ . The intervals are defined for each of the three experimental categories, by requiring that a neutral kaon, if produced at  $x = y = z = 0$  with momentum  $p = (p_T, p_z)$ , decays within the relevant part of the corresponding detector. For the LL LHCb category, it is required that the kaon decays before reaching  $z_{max} = 280$  mm, corresponding to a decay where the decay products traverse at least 3 VELO segments (ignoring a number of widely spaced VELO segments placed at a distance of up to  $z = 750$  mm from the interaction point) [?]. For the DD LHCb category a decay at  $z \in [280, 2350]$  mm is required, corresponding to decay between the LL cut-off and the first downstream tracking station [?]. The time acceptance has a significant impact for the LHCb categories, where some 20 % of the kaons escape the tracking stations completely before decaying.

For Belle II, it is assumed that the  $K_S^0$  reconstruction is similar to the Belle  $K_S^0$  reconstruction, which is based on a neural network and reconstructs  $K_S^0$  decays for which the decay product leave tracks in both the drift chamber and silicon vertex detectors, as well as decays that leave tracks in the drift chamber only [?, ?]. Therefore, the  $K_S^0$  decay is required to be within  $r_{max} = 1130$  mm of the beam axis, corresponding to a decay within the outer radius of the drift-chamber. In practice,

most of the kaons decay inside the silicon vertex detector, and requiring a decay before 1130 mm is essentially equivalent to having no time cut-off.

#### 4.3.4 Detector material budget

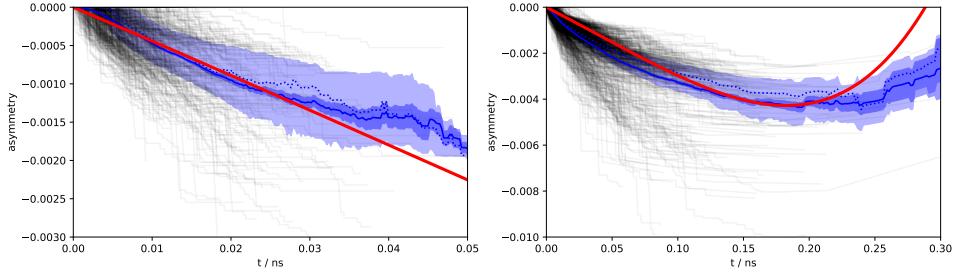
The effect of the material interaction is governed by parameter  $\Delta\chi$  of Eq. (4.16). The parameter varies along a given kaon path, as the kaon intersects detector components made of different materials. In these studies, the calculations are simplified by using a single average material parameter for each experimental scenario. The average material parameters can be estimated for a given experimental scenario by considering the type and length of material traversed by a kaon in the relevant sub-detector(s). The average value is estimated, by exploiting that  $\Delta\chi$  is related to the forward scattering amplitude  $f$  ( $\bar{f}$ ) of  $K^0$  ( $\bar{K}^0$ ) mesons in a given material [?, ?]

$$\Delta\chi = -\frac{2\pi\mathcal{N}}{m_K}(f - \bar{f}) = -\frac{2\pi(N_A\rho/A)}{m_K}(f - \bar{f}), \quad (4.38)$$

where  $\mathcal{N} = N_A\rho/A$  is the scattering centre density of the material,  $m_K$  is the mass of the kaon state,  $A$  and  $\rho$  are the nucleon number and density of the material, and  $N_A$  is Avogadro's number. Measurements made for a range of nuclei [?] show that in the momentum range  $p_K \in [20, 140] \text{ GeV}/c$

$$\left| \frac{f - \bar{f}}{p_K} \right| = 2.23 \frac{A^{0.758}}{p_K^{0.614}(\text{GeV}/c)} \text{ mb}, \quad \arg[f - \bar{f}] = -\frac{\pi}{2}(2 - 0.614), \quad (4.39)$$

where the phase of  $\Delta f$  is determined via a phase-power relation [?]. In the numerical studies presented here, Eq. (4.39) is also used for the low momentum neutral kaons in the Belle II calculations, as a more detailed modelling of the low momentum  $\Delta\chi$  based on Ref. [?] is found to yield very similar results. The scattering centre density  $\mathcal{N}$  is approximated as being constant, equal to the average density along a neutral kaon path due to its intersection with different detector segments. This average is estimated using the simplifying assumption that the total detector material budget is due to silicon. In practice,  $\mathcal{N} = N_A\rho/A$  is calculated using  $A = 28$  and  $\rho = f^{\text{Si}}\rho^{\text{Si}}$ , where  $f^{\text{Si}} < 1$  is the average fraction of a neutral kaon path length that is inside detector material, estimated via the known dimensions of the detector, the average nuclear interaction length seen by a track traversing it cf. the technical design reports [?, ?], and the nuclear interaction length of silicon  $\lambda_I^{\text{Si}} = 465.2 \text{ mm}$  [?]. The average value of  $r_\chi = \frac{1}{2}\frac{\Delta\chi}{\Delta\lambda}$ , which governs the size of the matter regeneration effect, can be calculated for the three considered experimental scenarios and satisfy  $|r_\chi^{\text{LL}}| = 2.7 \times 10^{-3}$ ,  $|r_\chi^{\text{DD}}| = 2.2 \times 10^{-3}$ , and  $|r_\chi^{\text{Belle II}}| = 1.0 \times 10^{-3}$ .



**Figure 4.5:** The asymmetry in Eq. (4.40) as a function of time for (left) LL and (right) DD  $K_S^0$  tracks in a simulated LHCb sample. The black lines show individual tracks. The light blue area is the central 50 % quantile, the dark blue area is the  $1\sigma$  uncertainty band on the mean. The red lines are calculated using the average  $\Delta\chi$  values that are also used in the calculation of biases in BPGGSZ measurements.

1548 The neutral kaon tracks in LHCb generally pass through somewhere between  
1549 zero (for a significant amount of the LL tracks) and a hundred (for some DD tracks)  
1550 distinct detector segments. Therefore it is worth examining the degree to which  
1551 using a single average  $\Delta\chi$  value, obtained following the procedure outlined above,  
1552 provides a reasonable description of the average material interaction. This can be  
1553 done using full LHCb simulation, where the kaon state for a simulated track can be  
1554 evaluated at all times, by applying Eq. (4.15) iteratively for each detector segment  
1555 the track traverses, using a  $\Delta\chi$  value appropriate for that segment. This is done  
1556 in Fig. 4.5 for a simple observable: the yield asymmetry

$$A_{K^0} = \frac{|\psi_K^0(t)|^2 - |\psi_{\bar{K}^0}(t)|^2}{|\psi_K^0(t)|^2 + |\psi_{\bar{K}^0}(t)|^2}, \quad (4.40)$$

1557 where  $\psi_K^0(t)$  ( $\psi_{\bar{K}^0}(t)$ ) is the amplitude for an initial  $K^0$  ( $\bar{K}^0$ ) to decay to two pions at  
1558 time  $t$ . In this calculation, it is assumed that  $\epsilon = 0$  to isolate the material effect with  
1559 no asymmetry contribution from the inherent  $CP$ -violation in the neutral kaon sector.  
1560 While the track-by-track asymmetries are found to differ significantly depending on  
1561 the exact detector segments a track intersects, the average asymmetry is seen to  
1562 evolve smoothly as a function of decay time, and in reasonable agreement with the  
1563 asymmetry value that is calculated using the average  $\Delta\chi$  values estimated above.

1564 The LHCb detector is undergoing a significant upgrade prior to the start of  
1565 the LHC Run 3. However, the material budget and geometry of the relevant  
1566 sub-detectors will be similar to the sub-detectors used during Run 1 and 2 [?, ?].  
1567 Hence the results of this study will be valid for measurements during the upgrade  
1568 phases of LHCb, even though the detector parameters presented in this section  
1569 relate to the original LHCb detector.

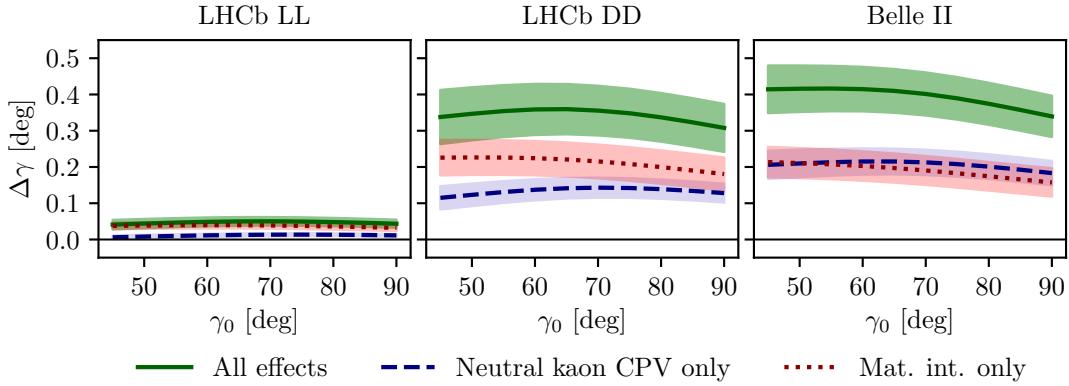
1570 **4.3.5 Calculation procedure**

1571 The main idea in the bias study is to calculate the BPGBSZ bin yields including  
1572 the full effect of neutral kaon  $CP$  violation and material, fit them using the default  
1573 equations of Chapter 2, and obtain the bias  $\Delta\gamma = \gamma - \gamma^0$  due to the kaon effects not  
1574 being considered in the parameter extraction. For the purpose of Ref. [?], a simple  
1575 fit setup of a single  $B^\pm \rightarrow Dh^\pm$  mode is investigated, where the  $K_i$  parameters are  
1576 determined in a control channel with the relevant experimental acceptance. This  
1577 setup is modified in the study used to assign a systematic uncertainty on the LHCb  
1578 measurement of Chapter 5, as described in Section 4.3.7 below.

1579 In practice, the amplitude model for  $D^0 \rightarrow K_S^0 \pi^+ \pi^-$  decays in Ref. [?] is taken  
1580 to represent the  $A_1(s_{+-})$  amplitude. Then  $A_2(s_{+-})$  is obtained as described in  
1581 Section 4.2.2. In terms of  $A_1$  and  $A_2$ , the amplitudes  $A_{S(L)}^{(\bar{D})}(s_{+-})$  can be expressed  
1582 and related via Eqs. (4.25) and (4.26), and the full signal decay amplitudes as a  
1583 function of phase-space coordinates, time, and the material interaction parameter  
1584  $\Delta\chi$  can be calculated for a given set of input parameters  $(\gamma^0, r_B^0, \delta_B^0)$ . The squared  
1585 decay amplitudes are then integrated over phase space and the kaon decay times  
1586 to obtain the binned signal yield.

1587 The signal yields depend on the momentum via the time-acceptance parameters  
1588  $\tau_1$  and  $\tau_2$ , and because the material interaction parameter  $\Delta\chi$  is momentum  
1589 dependent. Therefore, the yields are averaged over the  $K_S^0$  momentum distributions  
1590 of LHCb and Belle II.

1591 The parameters  $x_\pm$  and  $y_\pm$  are determined by a maximum likelihood fit to the  
1592 calculated yields, after which the fit result and covariance matrix are interpreted in  
1593 terms of the physics parameters  $(\gamma, r_B, \delta_B)$  using another maximum likelihood  
1594 fit [?]. In the fits, the  $K_i$  are obtained using the definition  $K_i = K_i^{\text{meas}} =$   
1595  $(N_i^D + N_{-i}^{\bar{D}})/(\sum_j N_j^D + N_{-j}^{\bar{D}})$ , in terms of the expected yields  $N_i^D$  ( $N_i^{\bar{D}}$ ) of a flavour-  
1596 tagged  $D^0$  ( $\bar{D}^0$ ) decays in bin  $i$  of the  $D$  decay phase space, calculated as described  
1597 above for  $r_B^0 = 0$ . This corresponds to experimentally measuring the  $K_i$  in a control  
1598 channel, and takes the effect of neutral kaon  $CP$  violation and material interaction  
1599 on  $K_i$  measurements into account, as well the experimental time acceptance. The  
1600  $(c_i, s_i)$  are calculated using  $A_1(s_{+-})$  and the experimental time acceptance is taken  
1601 into account in this calculation as well.

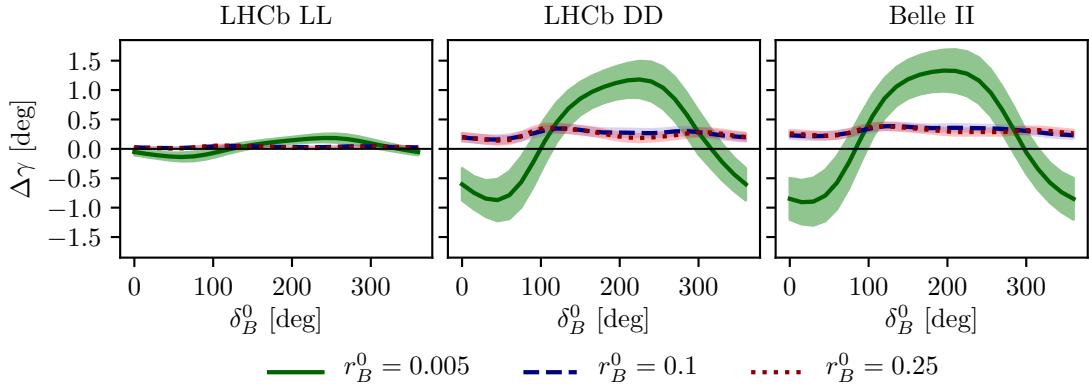


**Figure 4.6:** The bias  $\Delta\gamma$  as a function of input  $\gamma_0$  for (left) the LL LHCb category, (centre) the DD LHCb category, and (right) Belle II. The bias is calculated due to (blue, dashed line) neutral kaon  $CP$  violation alone, (red, dotted line) material interaction alone, and (green line) both effects. The shaded region shows the estimated  $1\sigma$  uncertainty band.

### 1602 4.3.6 Results

1603 The obtained bias  $\Delta\gamma$  is shown as a function of input  $\gamma^0$  for the various experimental  
1604 conditions in Fig. 4.6. The calculations are made using  $(r_B^0, \delta_B^0) = (0.1, 130^\circ)$ ,  
1605 approximately equal to the physics parameters relevant for  $B^\pm \rightarrow DK^\pm$  decays [?, ?].  
1606 The bias does not vary significantly with  $\gamma^0$  in the plotted range, which includes  
1607 the world average value of direct  $\gamma$  measurements as well as the values obtained in  
1608 full unitarity-triangle fits [?, ?, ?], and for all cases, the bias is found to be below  
1609  $0.5^\circ$ , corresponding to relative biases of about half a percent. Thus the biases are  
1610 of  $O(r\epsilon/r_B)$  as expected, given the arguments of Section 4.2.3. The contributions  
1611 from the individual  $K_S^0$  CPV and material interaction effects are also shown. It  
1612 is seen that the neutral kaon  $CP$  violation and material interaction effects leads  
1613 to approximately equal biases in all three cases.

1614 Given the decay-time acceptance and momentum distribution for each experimen-  
1615 tal category, the mean life time,  $\langle\tau\rangle$ , of the reconstructed kaons can be calculated.  
1616 In terms of the  $K_S^0$  lifetime  $\tau_{K_S^0} = (0.895 \pm 0.004) \times 10^{-11}$  s [?],  $\langle\tau_{LL}\rangle \simeq 0.1\tau_{K_S^0}$  for  
1617 the LHCb LL category,  $\langle\tau_{DD}\rangle \simeq 0.8\tau_{K_S^0}$  for the LHCb DD category, and at Belle  
1618 II  $\langle\tau_{Belle\,II}\rangle \simeq \tau_{K_S^0}$ . The difference in average kaon lifetime is reflected in the  
1619 observed biases, which are found to be larger in the samples with longer lived  
1620 kaons. The very small effect in the LL category is to be expected because the  
1621  $CP$ -violation effect due to  $K_S^0$  not being  $CP$ -even is approximately cancelled by  
1622 the  $CP$ -violation effect arising from  $K_S^0 - K_L^0$  interference for kaons with decay  
1623 times much smaller than  $\tau_{K_S^0}$  [?].



**Figure 4.7:** The bias  $\Delta\gamma$  as a function of input  $\delta_B$  for (left) the LL LHCb category, (centre) the DD LHCb category, and (right) Belle II. The bias is calculated for  $\gamma = 75^\circ$  and (green line)  $r_B = 0.005$ , (blue, dashed line)  $r_B = 0.1$ , and (red, dotted line)  $r_B = 0.25$ . The shaded region shows the estimated  $1\sigma$  uncertainty band.

1624     The uncertainty bands in Fig. 4.6 are calculated by repeating the study while  
 1625     varying some of the inputs. The model dependence of the predicted biases is  
 1626     probed by repeating the study using two other amplitude models as input for  
 1627      $A_1(s_{+-})$  and  $A_2(s_{+-})$ : the model published in Ref. [?] and the model included in  
 1628     EVTGEN [?]. hen defining  $A_2(s_{+-})$  in terms of  $A_1(s_{+-})$ , there is an uncertainty  
 1629     due to the unknown  $(r_k, \delta_k)$  parameters used to describe the  $\pi\pi$  resonance terms.  
 1630     This uncertainty is assessed by making the study with several different random  
 1631     realisations of the parameter set. The studies are repeated while varying the time  
 1632     acceptances and material densities with  $\pm 10\%$ . There is an additional uncertainty  
 1633     due to the use of simulation samples generated with `RapidSim` to describe the kaon  
 1634     momentum distribution, in lieu of full detector simulations.

1635     There is also an uncertainty from the use of  $(c_i, s_i)$  as calculated using  $A_1(s_{+-})$ .  
 1636     It is to be expected that the measured values  $(\hat{c}_i, \hat{s}_i)$  from the CLEO collaboration  
 1637     differ by those calculated using  $A_1^D(s_-, s_+)$  by terms of  $O(\epsilon)$  due to neutral kaon  
 1638      $CP$  violation, which is not taken into account in the measurement [?]. These  
 1639     corrections can be calculated via a procedure analogous to the one used to estimate  
 1640     the corrections on measurements of  $\gamma$  in this paper. However, as these corrections  
 1641     are much smaller than the experimental uncertainties in the measurement, they  
 1642     have not been studied further.

1643     For the purpose of this thesis, it is important to consider the bias in measurements  
 1644     that use  $B^\pm \rightarrow D\pi^\pm$  decays as well, and other  $B$  decay modes can also be used in  
 1645     BPFGSZ measurements, such as  $B^\pm \rightarrow D^*K^\pm$ ,  $B^\pm \rightarrow DK^{*\pm}$ , and  $B^0 \rightarrow DK^{*0}$ .  
 1646     For the purpose of the study presented here, the main difference between the decay

1647 channels is that they have different values of  $r_B$  and  $\delta_B$ . Figure 4.7 shows  $\Delta\gamma$  as  
1648 a function of input  $\delta_B^0$ , for  $\gamma^0 = 75^\circ$  and three different values of  $r_B^0$ . Aside from  
1649  $r_B^0 = 0.1$ , the results are shown for  $r_B^0 = 0.005$ , which corresponds to the expectation  
1650 in  $B^\pm \rightarrow D\pi^\pm$  decays [?] and  $r_B^0 = 0.25$ , which corresponds to  $B^0 \rightarrow DK^{*0}$   
1651 decays [?]. The most notable feature is that the biases are significantly larger in  
1652 the  $B^\pm \rightarrow D\pi^\pm$  case. This is expected: the  $r_B^0$  dependent behaviour is governed  
1653 by the relative importance of different  $O(r\epsilon)$  correction terms to the phase-space  
1654 distribution. There are terms of both  $O(r_A\epsilon)$  and  $O(r_B\epsilon)^3$ , which lead to expected  
1655 biases of size  $O(r_A\epsilon/r_B)$  and  $O(r_B\epsilon/r_B) = O(\epsilon)$ , respectively, cf. the discussion  
1656 of Section 4.2.3. In the  $B^\pm \rightarrow D\pi^\pm$  case, the  $O(r_A\epsilon)$  correction terms dominate  
1657 because  $r_A/r_B \simeq (0.05/0.005) = 10$ . This explains the relatively large bias, as  
1658  $|r_A\epsilon/r_B^{D\pi}| \simeq 4\%$ . The bias is seen to be up to  $\pm 1.5^\circ$ , but only about  $+0.2^\circ$  with  
1659 the expected value of  $\delta_B^{D\pi} \simeq 300^\circ$  [?, ?]. These biases are *much smaller* than the  
1660 precision on  $\gamma$  that is obtainable in a  $B^\pm \rightarrow D\pi^\pm$  analysis with current experimental  
1661 yields, and do thus not pose a problem. In the  $r_B^0 = 0.1$  and  $r_B^0 = 0.25$  cases the  
1662  $O(r_B\epsilon)$  correction terms dominate, and the biases are of  $O(\epsilon)$ , independent of the  
1663  $r_B^0$  value. Therefore both cases have biases of similar size.

1664 Further, it is clear that the biases depend on  $\delta_B^0$  and that the oscillation period of  
1665 the  $\delta_B$  dependence is different between the  $r_B^0 = 0.005$  case and the  $r_B^0 \in \{0.1, 0.25\}$   
1666 cases. It is to be expected that  $\Delta\gamma$  oscillates as a function of  $\delta_B^0$ , because  $\delta_B^0$  enters  
1667 the yield equations via  $\cos(\delta_B^0 \pm \gamma)$  and  $\sin(\delta_B^0 \pm \gamma)$  terms. As explained above,  
1668 the  $O(r_A\epsilon)$  terms dominate the  $B^\pm \rightarrow D\pi^\pm$  bias, and these are independent of  
1669  $\delta_B^0$ . The  $O(r_B\epsilon)$  terms, however, are important for the bias corrections for larger  
1670  $r_B$  values, and the terms include factors of  $\cos(\delta_B^0 \pm \gamma)$  and  $\sin(\delta_B^0 \pm \gamma)$ . This  
1671 explains the different bias dependence on  $\delta_B^0$ .

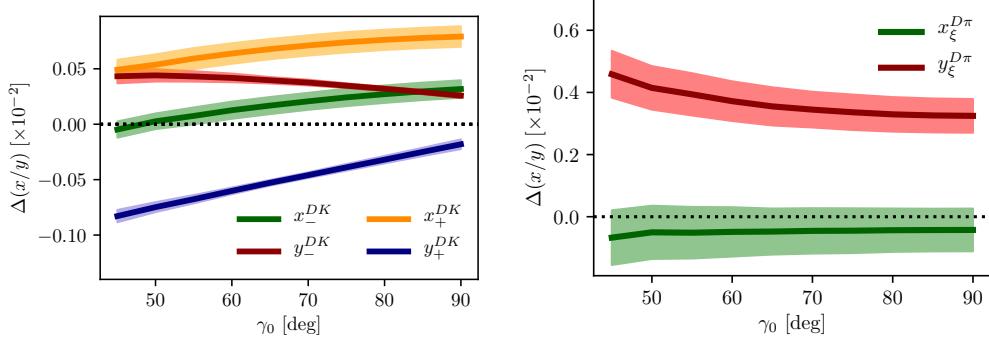
1672 While the input value of  $\gamma^0 = 75^\circ$  was chosen for these studies, there is minimal  
1673 variation in the results if another value of  $\gamma^0$  in the range  $[60^\circ, 85^\circ]$  is used.

### 1674 4.3.7 Coupled $B^\pm \rightarrow DK^\pm$ and $B^\pm \rightarrow D\pi^\pm$ measurements

1675 The studies presented above have been extended on two accounts in order to assign  
1676 a systematic uncertainty to the LHCb measurement presented in Chapter 5. Firstly,  
1677 full LHCb simulation has been used to obtain the momentum distributions, as  
1678 well as to fit a better description of the time acceptance and the reconstruction  
1679 efficiency profile over the  $D$ -decay phase space. Secondly, the fit setup is modified

---

<sup>3</sup>There are similar terms of  $O(r_A r_\chi)$  and  $O(r_B r_\chi)$ , but as  $\epsilon$  and  $r_\chi$  are of the same order of magnitude, these terms can be treated completely analogously to the  $O(r_A\epsilon)$  and  $O(r_B\epsilon)$  terms, and have been left out of the discussion for brevity.



**Figure 4.8:** The bias on (left) the  $B^\pm \rightarrow DK^\pm$  and (right)  $B^\pm \rightarrow D\pi^\pm$   $CP$ -violation observables in the LHCb DD category, evaluated in bias studies with inputs based on full LHCb simulation, calculated as a function of input  $\gamma_0$ .

to correspond to the experimental approach described in Section 2.4 and Chapter 5: the signal yields are calculated for both the  $B^\pm \rightarrow DK^\pm$  and  $B^\pm \rightarrow D\pi^\pm$  channels, and fitted in a combined fit to obtain  $(x_\pm^{DK}, y_\pm^{DK}, x_\xi^{D\pi}, y_\xi^{D\pi})$ , where the  $F_i$  parameters are allowed to float in the fit. The biases obtained for each observable are shown in Fig. 4.8, evaluated using the time-acceptance, momentum distribution, and material budget relevant for the DD category (since the effect in the LL category is much smaller). As will be clear in Chapter 5, these biases are all significantly smaller than the corresponding statistical uncertainties. Thus, the effects of neutral kaon  $CP$  violation and material interactions contribute a manageable systematic uncertainty in current BPGGSZ measurements, even if the  $B^\pm \rightarrow D\pi^\pm$  channel is promoted to a signal channel.

As the statistical uncertainty becomes comparable with the bias effects described in this chapter, the systematic uncertainty should be assigned by a more accurate study, incorporating the traversed material on a track-by-track basis in full detector simulation. Such a detailed calculations can also be used to apply a bias correction if desired.

## 4.4 Concluding remarks

The analysis presented in this chapter has shown the expected impact of neutral kaon  $CP$  violation and material interaction on current BPGGSZ measurements to be small compared to the statistical uncertainties; first by simple order-of-magnitude estimates and then by a detailed calculation of the expected effect in LHCb and Belle II.

While the calculations were made for the case of  $D \rightarrow K_S^0 \pi^+ \pi^-$  decays, the BPGGSZ approach can of course also be applied in other  $D$ -decay final states,

such as  $D \rightarrow K_S^0 K^+ K^-$  and  $D \rightarrow K_S^0 \pi^+ \pi^- \pi^0$ . The biases on measurements of  $\gamma$  based the  $D$  decay phase-space distributions should be of similar size in these decay channels. The impact on  $\gamma$  measurements based on the phase-space-integrated yield asymmetry can be expected to be tens of degrees for the  $D \rightarrow K_S^0 K^+ K^-$  channel, where the yield asymmetry is expected to be around 2 %, for the reasons explained in Section 4.2.3. The  $D \rightarrow K_S^0 \pi^+ \pi^- \pi^0$  decay, however, is dominantly  $CP$ -odd [?], and the bias in measurements based on the total asymmetry is therefore expected to be  $O(\epsilon/r_B)$ , ie. a few degrees [?]. More precise calculations of the biases would require a repeat of the study included here, with relevant amplitude models and binning schemes in place.

The chapter focuses on the model-independent, binned approach that is the subject of the thesis. However, the underlying mechanism that determines the scale of the bias, namely that the phase-space *distribution* of signal decays is unaffected at  $\mathcal{O}(\epsilon)$  and  $\mathcal{O}(r_\chi)$ , is independent on the exact measurement approach. Therefore it is expected that amplitude-model-based measurements and measurements made with new unbinned methods such as those in Ref [?] will be similarly biased if kaon  $CP$  violation and regeneration are not accounted for.

# 5

1721

1722

1723

## A BPGBGSZ measurement of $\gamma$ with $B^\pm \rightarrow Dh^\pm$ decays

1724 This chapter describes a model-independent BPGBGSZ measurement of  $\gamma$  with  
1725  $B^\pm \rightarrow DK^\pm$  and  $B^\pm \rightarrow D\pi^\pm$  decays where  $D \rightarrow K_S^0\pi^+\pi^-$  and  $D \rightarrow K_S^0K^+K^-$ ,  
1726 commonly denoted  $B^\pm \rightarrow D(\rightarrow K_S^0h^+h^-)h'^\pm$  decays. The measurement is made  
1727 with the full LHCb data set collected during Run 1 and 2 of the LHC, corresponding  
1728 to an integrated luminosity of about  $9\text{ fb}^{-1}$ . The analysis is under review for  
1729 publication in the Journal of High Energy Physics at the time of writing [?]  
1730 (one can hope).

### 1731 5.1 Candidate reconstruction and selection

1732 The  $B^\pm \rightarrow D(\rightarrow K_S^0h^+h^-)h'^\pm$  candidates are constructed during the offline *stripping*  
1733 stage described in Section 3.3.3. The candidates are defined by first combining  
1734 tracks to form a  $K_S^0 \rightarrow \pi^+\pi^-$  vertex, then a  $D \rightarrow K_S^0h^+h^-$  vertex, and finally  
1735 the  $B^\pm \rightarrow Dh'^\pm$  candidate. Each final state track is required to satisfy certain  
1736 momentum thresholds and track-quality requirements, and to be separated from all  
1737 primary interaction vertices. Each decay vertex is required to satisfy a fit-quality  
1738 threshold and to be separated from the primary vertex. Momentum thresholds  
1739 are applied to the composite particles and they are required to have reconstructed  
1740 invariant masses close to their known masses<sup>1</sup> except that the  $B$  candidate is  
1741 required to have a reconstructed invariant mass in the interval  $4750\text{--}7000\text{ MeV}/c^2$ .

---

<sup>1</sup>The exact mass window depends on the particle type and reconstruction category; narrower mass windows are applied at a later stage, as described below.

1742 The  $B$  candidate is required to satisfy  $\chi^2_{\text{IP}} < 25$ , where  $\chi^2_{\text{IP}}$  is the difference in  $\chi^2$   
1743 value of the primary vertex fit, when the vertex is formed with- and without the  $B$   
1744 candidate. As the final *stripping* stage, a multivariate algorithm is applied to the  
1745 formed  $B$  candidate to reduce the amount of random track combinations, denoted  
1746 combinatorial background, even further than the aforementioned requirements.

1747 Two data categories are defined, depending the tracks used to form the  $K_S^0$   
1748 candidate: the LL category where both pions are long tracks, and DD category where  
1749 both pions are downstream tracks, using the track classifications of Section 3.2.1.

1750 Each candidate is re-analysed with the `DecayTreeFitter` (DTF) frame work [?],  
1751 where a simultaneous fit of the full decay chain is made with a number of constraints  
1752 applied: the momenta of the composite  $D$  and  $K_S^0$  particles are required to form  
1753 invariant masses exactly equal to the known particle masses [?], and the momentum  
1754 of the  $B$  candidate is required to point in the direction defined by the  $B$  decay  
1755 vertex and the primary vertex. This refit results in improved resolution of the  
1756 invariant masses of the composite particles and, very importantly, of the Dalitz  
1757 coordinates in the  $D$ -decay phase space. It also ensures that all candidates fall in the  
1758 kinematically allowed region of the  $D$ -decay phase space. Unless otherwise specified,  
1759 all results in this chapter are based on the refitted track momenta; for reasons  
1760 explained below, some studies have to be based on parameters that are obtained  
1761 without the constraints described above, or with only a subset of them applied.

1762 Following the stripping stage, the further selection of signal candidates is  
1763 performed in three steps: an initial set of requirements that remove a large fraction  
1764 of candidates that are very likely to be background and veto a number of specific  
1765 backgrounds, the application of a multivariate analysis algorithm designed to allow  
1766 for filtering combinatorial background, and finally a set of particle-identification  
1767 requirements. The requirements are summarised in Table 5.1, and each step is  
1768 described in detail in the following sections.

### 1769 5.1.1 Initial requirements

1770 At the hardware trigger level, it is required that a particle associated with the  
1771 signal decay triggered the hadronic L0 trigger (classifying the event as *Trigger on*  
1772 *Signal*, or TOS), or that the level-0 trigger decision was caused by a particle that  
1773 is not associated with the signal decay ( $e^\mp$ *Trigger Independent of Signal*, or TIS).  
1774 The inclusion of the latter category increases the data sample about 50 %. At the  
1775 software trigger level, a particle belonging to the signal decay is required to have  
1776 caused one of each of the inclusive HLT1 and HLT2 lines to accept the events.  
1777 Specifically, the Run 1 events are required to be TOS on the `HLT1TrackAllL0` and

**Table 5.1:** Summary of requirements applied to data. The base requirements are applied to all data samples before training or applying the BDT.

Base requirements		
Variable	Cut	Comment
Bachelor momentum, $p$	$< 100 \text{ GeV}/c$	
Bachelor has RICH	<i>true</i>	
$K^\pm$ in $D$ decay: momentum, $p$	$< 100 \text{ GeV}/c$	In $D \rightarrow K_S^0 K^+ K^-$
$K^\pm$ in $D$ decay: have RICH	<i>true</i>	In $D \rightarrow K_S^0 K^+ K^-$
DecayTreeFit converged	<i>true</i>	
$D$ mass	$m_{D^0} \in m_{D^0}^{PDG} \pm 25 \text{ MeV}/c^2$	From DTF with constrained $K_S^0$ mass
$K_S^0$ mass	$m_{K_S^0} \in m_{K_S^0}^{PDG} \pm 15 \text{ MeV}/c^2$	From DTF with constrained $D^0$ mass

Background suppressing requirements		
Variable	Cut	Comment
$K_S^0$ flight distance $\chi^2$	$> 49$	for LL only
$\Delta z_{\text{significance}}^{DB}$	$> 0.5$	for all candidates

PID requirements		
Channel	Cut	Comment
$B^\pm \rightarrow DK^\pm$	PIDK > 4	for bachelor
$B^\pm \rightarrow D\pi^\pm$	PIDK < 4	for bachelor
$B^\pm \rightarrow Dh^\pm$	IsMuon = 0	for bachelor
$B^\pm \rightarrow D(\rightarrow K_S^0 \pi^+ \pi^-)h^\pm$	PIDe < 0 & IsMuon = 0	for charged $D$ decay products
$B^\pm \rightarrow D(\rightarrow K_S^0 K^+ K^-)h^\pm$	PIDK > -5 & IsMuon = 0	for charged $D$ decay products

BDT requirements		
Channel	Cut	Comment
Run 1, DD	$> 0.6$	
Others	$> 0.8$	

1778 Hlt2Topo{2, 3, 4}BodyBBDT lines and the Run 2 events are required to be TOS on  
1779 the Hlt1{Track, TwoTrack}MVA and Hlt2Topo{2, 3, 4}Body lines. These trigger  
1780 lines were described in Section 3.3.

1781 Before any processing of the data, a loose preselection is applied to remove  
1782 obvious background candidates. The reconstructed  $D$  ( $K_S^0$ ) mass is required to  
1783 be within  $25$  ( $15$ )  $\text{MeV}/c^2$  of the known values [?]. The *companion* particle, the  
1784 pion or kaon produced in the  $B^\pm \rightarrow Dh^\pm$  decay, is required to have associated  
1785 RICH information and a momentum less than  $100 \text{ GeV}/c$ ; this ensures good particle-  
1786 identification performance. Finally, all of the DTF fits of the full decay chain  
1787 are required to have converged properly.

1788 Two additional requirements are made at this stage in order to suppress specific  
1789 backgrounds. In order to suppress decays of the type  $B^\pm \rightarrow K_S^0 h^+ h^- h'^\pm$  with  
1790 no intermediate  $D$  meson, so called *charmless* decays, it is required that the  
1791 significance of the  $z$ -separation of the  $D^0$  decay vertex and the  $B^\pm$  decay vertex  
1792 is above 0.5. The significance of the  $z$ -separation of the  $D^0$  decay vertex and

1793 the  $B^\pm$  decay vertex is defined as

$$\Delta z_{\text{significance}}^{D-B} = \frac{z_{vtx}^D - z_{vtx}^B}{\sqrt{\sigma^2(z_{vtx}^D) + \sigma^2(z_{vtx}^B)}}. \quad (5.1)$$

1794 This source of background described further in section 5.3.1. In order to suppress  
1795 a background from  $D \rightarrow 4\pi$  and  $D \rightarrow \pi\pi KK$  decays, it is required that the  $K_S^0$   
1796 flight distance  $\chi_{\text{FD}}^2$  is greater than 49, where

$$\chi_{\text{FD}}^2 = \left( \frac{\Delta r}{\sigma(\Delta r)} \right)^2, \quad (5.2)$$

1797 and  $\Delta r$  is the measured flight distance of the  $K_S^0$  meson. This background is  
1798 described in further detail in section 5.3.2.

### 1799 5.1.2 Boosted decision tree

1800 A Gradient Boosted Decision Tree [?] (abbreviated BDT in the following) is  
1801 applied to class  each candidate on a scale from  $-1$  to  $+1$  as signal-like ( $+1$ )  
1802 or combinatorial-background-like ( $-1$ ), based on the values of a number of input  
1803 parameters for candidate in question. The BDT is implemented in the TMVA  
1804 frame work [?, ?].

1805 A boosted decision tree classifier consists of a number of sequentially trained  
1806 decision trees, each of which classify events as either signal or background. Each tree  
1807 bases the decision on an individual subset of variables, out of an overall set of input  
1808 variables. At each training step, the input events are weighted when training a new  
1809 tree, so that events that the already-trained trees classify incorrectly are given a  
1810 higher weight; this is denoted boosting. The term *gradient boosting* denotes a specific  
1811 weight calculation scheme [?]. The final score is the average over all decision trees.

1812 The full set of input variables are given in Table 5.2. It includes the momenta  
1813 of particles in the decay; a number of geometric parameters such a absolute and  
1814 relative vertex positions, and distances of closest approach between tracks;  $\chi_{\text{IP}}^2$   
1815 values for a number of particles in the decay chain; the  $\chi^2$  per degree of freedom  
1816 of the DTF refit; DIRA values, which denote the angle between the fitted particle  
1817 momenta and the vector spanned by it's production ad decay vertices; and finally  
1818 an isolation variable, defined as

$$A_{p_T} = \frac{p_T(B) - \sum p_T(\text{other})}{p_T(B) + \sum p_T(\text{other})} \quad (5.3)$$

1819 where the sum is over all other tracks in a cone around the  $B$ -candidate. The cone  
1820 is defined as being within a circle with a radius of 1.5 units around the  $B$  candidate

in the  $(\eta, \phi_{azim})$ -plane. This variable is highly efficient in rejecting combinatorial background. Two algorithms are trained, one for the LL category of  $K_S^0$  mesons and one for the DD category, because some input parameters relate to the  $K_S^0$  meson and have very different distributions between the two categories.

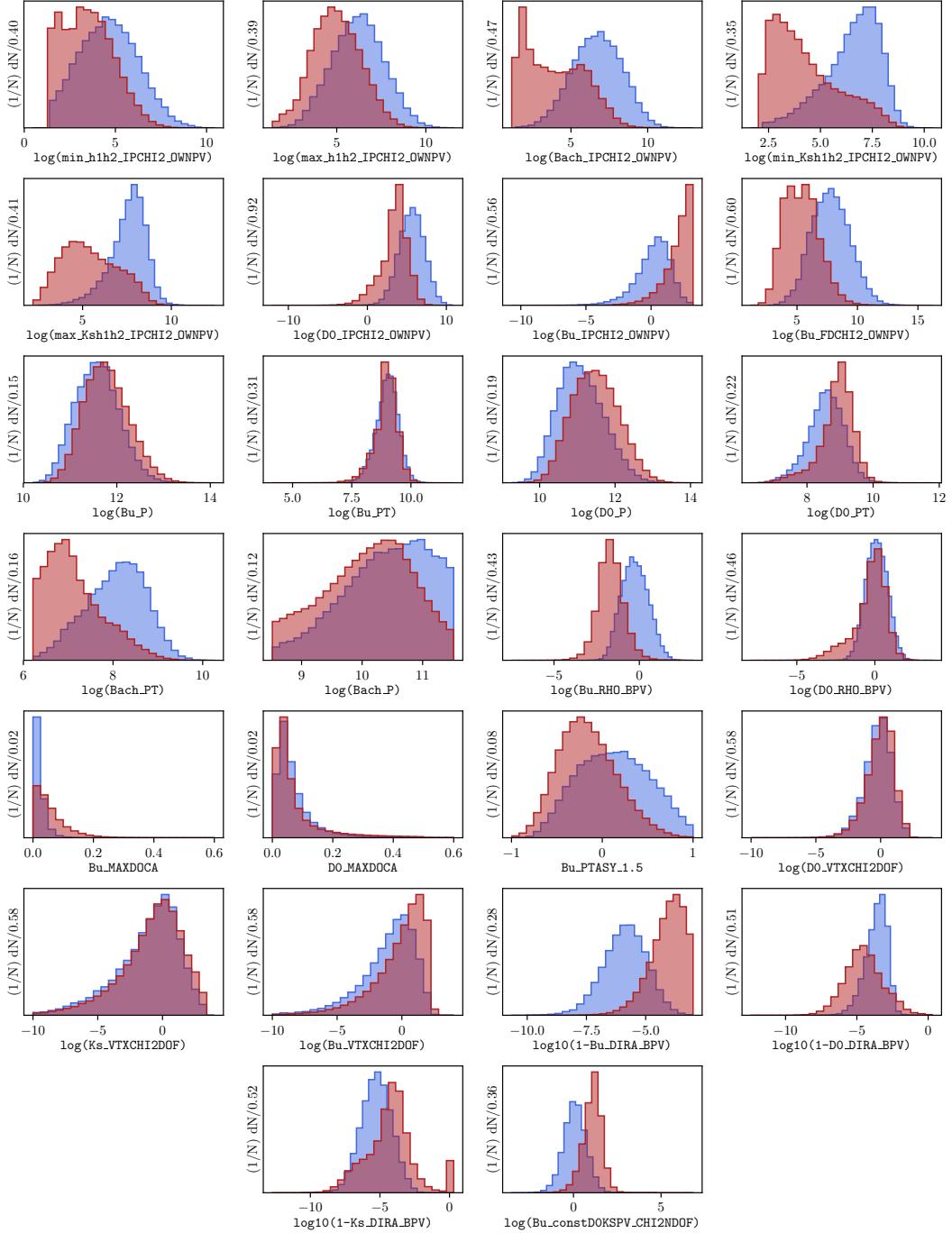


**Table 5.2:** Input parameter set used in BD<sub>LL</sub> trained to separate signal and combinatorial background, sorted according to importance in LL BDT.

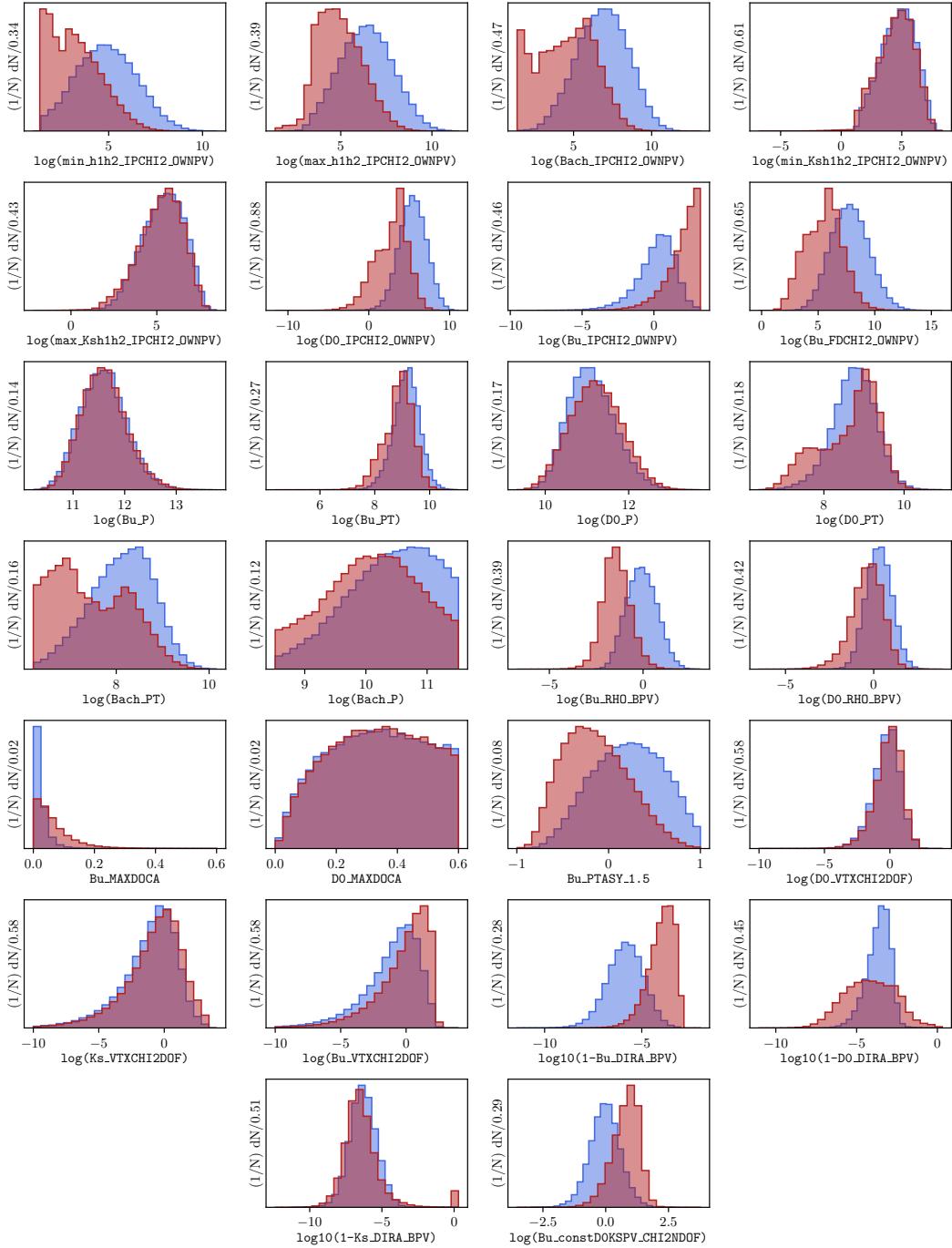
Variable name	Importance LL/DD (Rank in DD)	Description
$\log(1-Ks\_DIRA\_BPV)$	7.2 % / 3.5 % (16)	$\log \cos \theta_{DIRA}$ for $K_S^0$
$\log(Bu\_RHO\_BPV)$	5.7 % / 5.5 % (5)	Radial distance of $B$ vertex to beam line
$\log(Bach\_PT)$	5.2 % / 6.9 % (1)	$p_T$ of the bachelor particle
$\log(1-D0\_DIRA\_BPV)$	4.9 % / 5.8 % (4)	$\log \cos \theta_{DIRA}$ for $D$
$\log(1-Bu\_DIRA\_BPV)$	4.9 % / 6.4 % (3)	$\log \cos \theta_{DIRA}$ for $B^\pm$
$\log(D0\_RHO\_BPV)$	4.8 % / 5.3 % (6)	Radial distance of $D$ vertex to beam line
$Bu\_PTASY\_1.5$	4.7 % / 4.9 % (7)	Asymmetry parameters of $B^\pm$
$\log(D0\_PT)$	4.7 % / 6.6 % (2)	$p_T$ of the $D$ meson
$\log(Bu\_constDOKSPV\_CHI2NDOF)$	4.2 % / 4.5 % (9)	$\chi^2/\text{d.o.f}$ of kinematical refit with DecayTreeFitter
$\log(Bu\_FDCHI2\_OWNPV)$	3.9 % / 4.1 % (11)	Flight distance $\chi^2$ of the $B^\pm$
$\log(\max\_Ksh1h2\_IPCHI2\_OWNPV)$	3.9 % / 3.0 % (20)	Largest $\chi^2_{IP}$ of the $K_S^0$ decay products
$\log(D0\_IPCHI2\_OWNPV)$	3.8 % / 3.3 % (17)	$\chi^2_{IP}$ of the $D$
$\log(\min\_Ksh1h2\_IPCHI2\_OWNPV)$	3.7 % / 0.9 % (26)	Smallest $\chi^2_{IP}$ of the $K_S^0$ decay products
$\log(Bu\_P)$	3.7 % / 3.9 % (12)	$p$ of the $B^\pm$ meson
$\log(Bu\_IPCHI2\_OWNPV)$	3.6 % / 4.6 % (8)	$\chi^2_{IP}$ of the $B^\pm$
$Bu\_MAXDOCA$	3.6 % / 3.3 % (18)	"Distance of closest approach" for $B^\pm$ vertex
$\log(Bach\_IPCHI2\_OWNPV)$	3.3 % / 4.3 % (10)	$\chi^2_{IP}$ of the bachelor particle
$\log(Bu\_PT)$	3.3 % / 3.7 % (14)	$p_T$ of the $B^\pm$ meson
$\log(\max\_h1h2\_IPCHI2\_OWNPV)$	3.1 % / 3.8 % (13)	Largest $\chi^2_{IP}$ of the $D$ decay products
$\log(\min\_h1h2\_IPCHI2\_OWNPV)$	3.0 % / 3.4 % (19)	Smallest $\chi^2_{IP}$ of the $D$ decay products
$\log(Ks\_VTXCHI2DOF)$	2.9 % / 2.3 % (21)	$\chi^2$ of vertex fit for $K_S^0$
$D0\_MAXDOCA$	2.9 % / 1.0 % (25)	"Distance of closest approach" for $D$ vertex
$\log(D0\_VTXCHI2DOF)$	2.7 % / 1.6 % (24)	$\chi^2$ of vertex fit for $D$
$\log(D0\_P)$	2.7 % / 1.8 % (22)	$p$ of the $D$ meson
$\log(Bach\_P)$	2.2 % / 3.6 % (15)	$p$ of the bachelor particle
$\log(Bu\_VTXCHI2DOF)$	1.8 % / 1.7 % (23)	$\chi^2$ of vertex fit for $B^\pm$

The BDTs are trained and tested with input samples representing typical signal and background decay candidates: a signal sample that consists of simulated  $B^\pm \rightarrow D(\rightarrow K_S^0\pi^+\pi^-)\pi^\pm$  decays corresponding to the LHCb running conditions for the years 2012–2018, and a sample of combinatorial background candidates from real data, where the reconstructed invariant mass of the  $B$  meson is larger than 5800 MeV/ $c^2$ . The candidates in both samples were required to have passed the initial requirements described in the preceding section. The input-parameter distributions in the signal and background training samples are shown in Figs. 5.1 and 5.2. The signal and background samples are each split into two before the training stage: one sub sample, the training sample, is used to train the BDT, after which the trained algorithm is applied to the other sub sample, the test sample. The classifier is found to perform well on the test sample, not just the training sample, which ensures that it does not suffer significant overtraining. The BDT output distribution are shown for both test and training samples in Fig. 5.3, where it is clear that the classifier very effectively separates signal and background candidates.

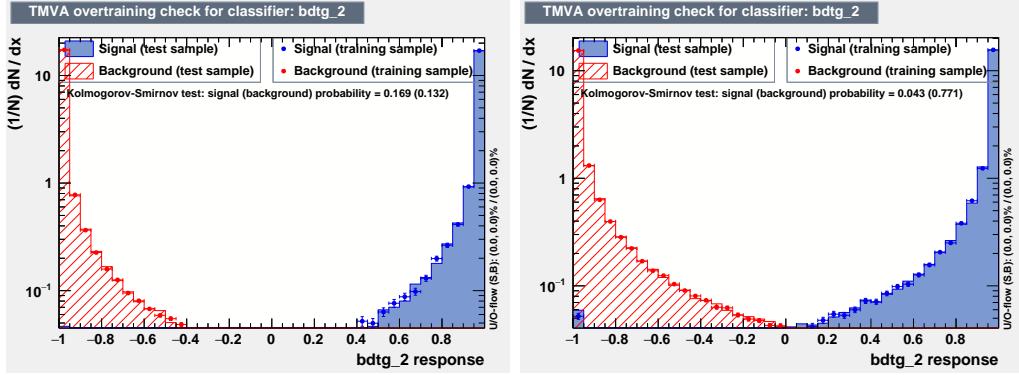
Each candidate in data is classified using the BDT, and candidates that are assigned a score below some threshold value are discarded. The threshold values are chosen in a set of pseudo experiments, such that the expected sensitivity to  $\gamma$  is maximised. This is done by performing preliminary fits to the data set for a range of different BDT threshold values, then generating many pseudo data sets with the obtained yields, and applying the full fit and interpretation procedure described in Sections 5.4–5.7 to each data set. Thus, the expected uncertainty on  $\gamma$  is obtained for a range of threshold values. The procedure is applied independently for the LL and DD categories, as well as for the Run 1 and Run 2 data sets, because some parameter distributions differ slightly between the two runs. The optimal threshold values are found to be 0.8 in all situations, except for LL candidates in Run 1 where it is 0.6. This is illustrated in Fig. 5.4 where the results of the threshold scans are shown. The same classifier is applied to both  $B^\pm \rightarrow D\pi^\pm$  and  $B^\pm \rightarrow DK^\pm$  candidates, and both  $D$  final state categories. While the classifiers were trained using samples of  $B^\pm \rightarrow D(\rightarrow K_S^0\pi^+\pi^-)\pi^\pm$  simulation and data, the decays are similar enough that no significant improvement in performance was obtained when considering a more elaborate setup. Across all categories, the requirement on the BDT output is found to remove approximately 98 % of the combinatorial background, while being approximately 93 % efficient on signal.



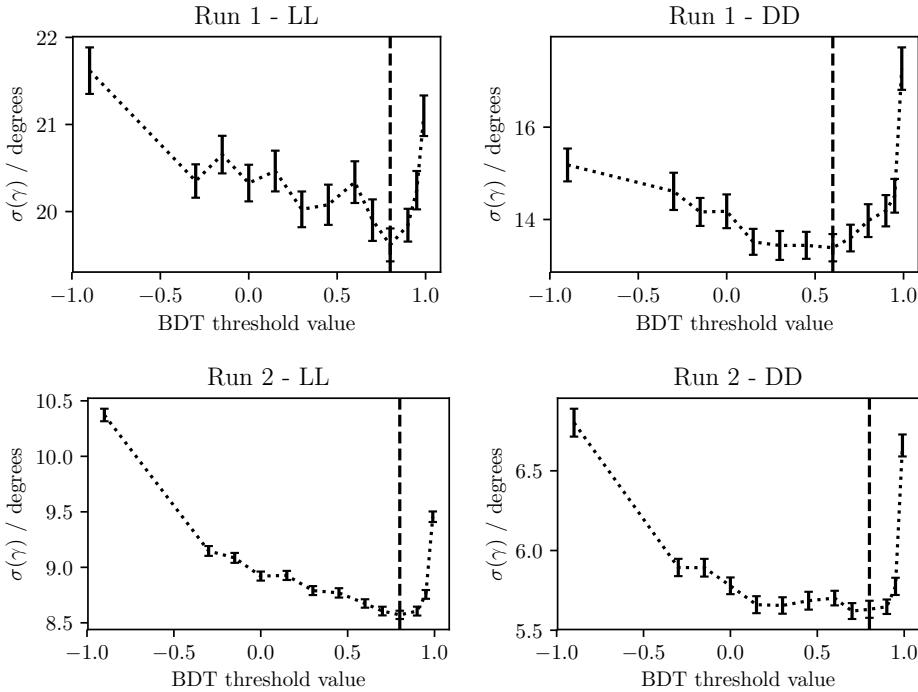
**Figure 5.1:** Distribution of input parameters in the LL training samples of (blue) signal decays from simulation and (red) background decays from the upper  $B$  sideband. The variable names described further in Table 5.2.



**Figure 5.2:** Distribution of input parameters in the DD training samples of (blue) signal decays from simulation and (red) background decays from the upper  $B$  sideband. The variable names described further in Table 5.2.



**Figure 5.3:** Distribution of BDT variable on test and training samples for (left) the LL and (right) the DD category, with logarithmic  $y$ -scale.



**Figure 5.4:** The mean uncertainty on  $\gamma$  in toy studies, performed with the signal and background yields corresponding to a given BDT requirement, using (top) the Run 1 and (bottom) Run 2 datasets, using only candidates in (left) the LL category and (right) the DD category. The dashed line shows the threshold value employed to discard background-like candidates in the selection.

### 5.1.3 Particle-identification requirements

A PID requirement is made to separate  $B^\pm \rightarrow DK^\pm$  and  $B^\pm \rightarrow D\pi^\pm$  candidates in the data sample, by requiring that the PIDK of the companion particle satisfies PIDK < 4 for  $B^\pm \rightarrow D\pi^\pm$  candidates and PIDK > 4 for  $B^\pm \rightarrow DK^\pm$  candidates. The PIDK variable was defined in Section 3.1.3. This ensures that any given candidates is selected into only one of these samples.

1865 Further to the requirement on the companion, PID requirements are made to  
 1866 suppress semi-leptonic backgrounds as well as decays where a final state particle  
 1867 decays in flight, and a loose PID requirement is made in the  $D \rightarrow K_S^0 K^+ K^-$   
 1868 channels where it leads to a higher signal purity:

- 1869 • the companion particle is required to satisfy `IsMuon` = 0.
- 1870 • For the  $B \rightarrow D(\rightarrow K_S^0 \pi^+ \pi^-) h^\pm$  samples it is required that the charged pion  
 1871 track from the  $D$  decay with opposite [redacted] sign to the companion satisfies  
 1872  $\text{PIDe} < 0 \& \text{IsMuon} = 0$ , and for the other charged pion that `IsMuon` = 0.
- 1873 • For the  $B \rightarrow D(\rightarrow K_S^0 K^+ K^-) h^\pm$  samples it is required that the charged kaon  
 1874 tracks from the  $D$  decay have RICH information, a momentum less than 100  
 1875 GeV/c and  $\text{PIDK} > -5 \& \text{IsMuon} = 0$ .

1876 These backgrounds are described in Section 5.3.3.

#### 1877 5.1.4 Final requirements

1878 For a small fraction of candidates in the final sample, it is the case that two  
 1879 or more candidates originate in the same  $pp$  collision. In order to make sure  
 1880 that all candidates are completely independent, a single, arbitrary candidate from  
 1881 each  $pp$  collision is kept for these collisions, and the other candidates discarded.  
 1882 This requirement results in the removal of less than 0.7% of candidates in  
 1883 each data category.

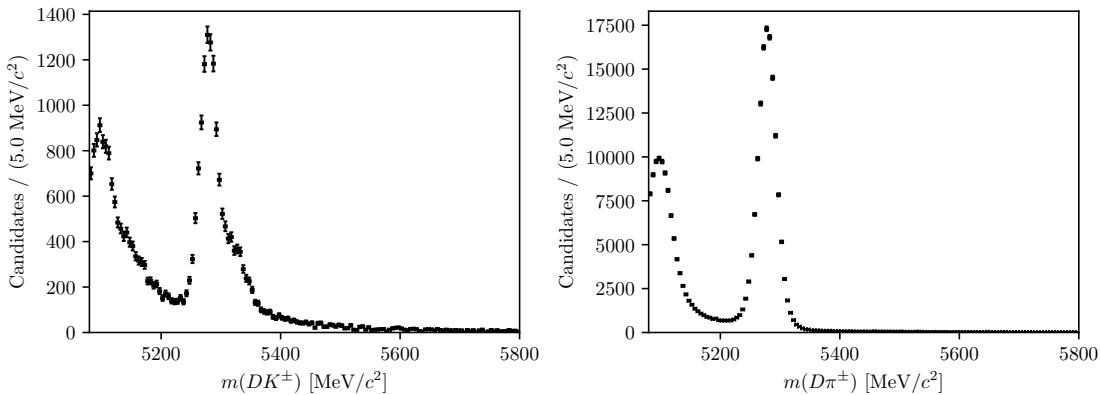
1884 Furthermore, the  $D$  mass used to define the binning schemes described in Ref. [?]  
 1885 differs slightly from the mass used in the DTF refit. Therefore a few of the decays are  
 1886 reconstructed with Dalitz coordinates outside the allowed kinematic region. Because  
 1887 this problem only concerns a handful of candidates, they are simply discarded.

#### 1888 5.1.5 Selected candidates

1889 In total, about 47,000  $B^\pm \rightarrow DK^\pm$  candidates and 400,000  $B^\pm \rightarrow D\pi^\pm$  candidates  
 1890 are selected, as summarised in Table 5.3. An example of the  $B$  mass distribution in  
 1891 one of the data categories is shown in Fig. 5.5; it is clear that a significant number  
 1892 of these candidates are background decays. The Dalitz plots for candidates in the  
 1893 signal region where  $m_B \in [5249, 5309]$  MeV/ $c^2$  are shown in Fig. 5.6 and 5.7. Due to  
 1894 the large yields in the full Run 1 and 2 LHCb data set, the asymmetries between  
 1895 the  $B^+$  and  $B^-$  distributions are visible to the eye in the  $B^\pm \rightarrow DK^\pm$  plots.

**Table 5.3:** Final candidate yield in each data category after the full selection has been applied, including removing candidates outside the region  $m_B \in [5080, 5800] \text{ MeV}/c^2$ .

$B$ Decay	$D$ final state	$K_S^0$ type	Run 1	Run 2	Total
$B^\pm \rightarrow DK^\pm$	$K_S^0\pi^+\pi^-$	LL	2275	10525	12800
		DD	5097	23508	28605
	$K_S^0K^+K^-$	LL	383	1610	1993
		DD	772	3397	4169
$B^\pm \rightarrow D\pi^\pm$	$K_S^0\pi^+\pi^-$	LL	18209	90509	108718
		DD	40167	205807	245974
	$K_S^0K^+K^-$	LL	2879	13757	16636
		DD	6033	29790	35823



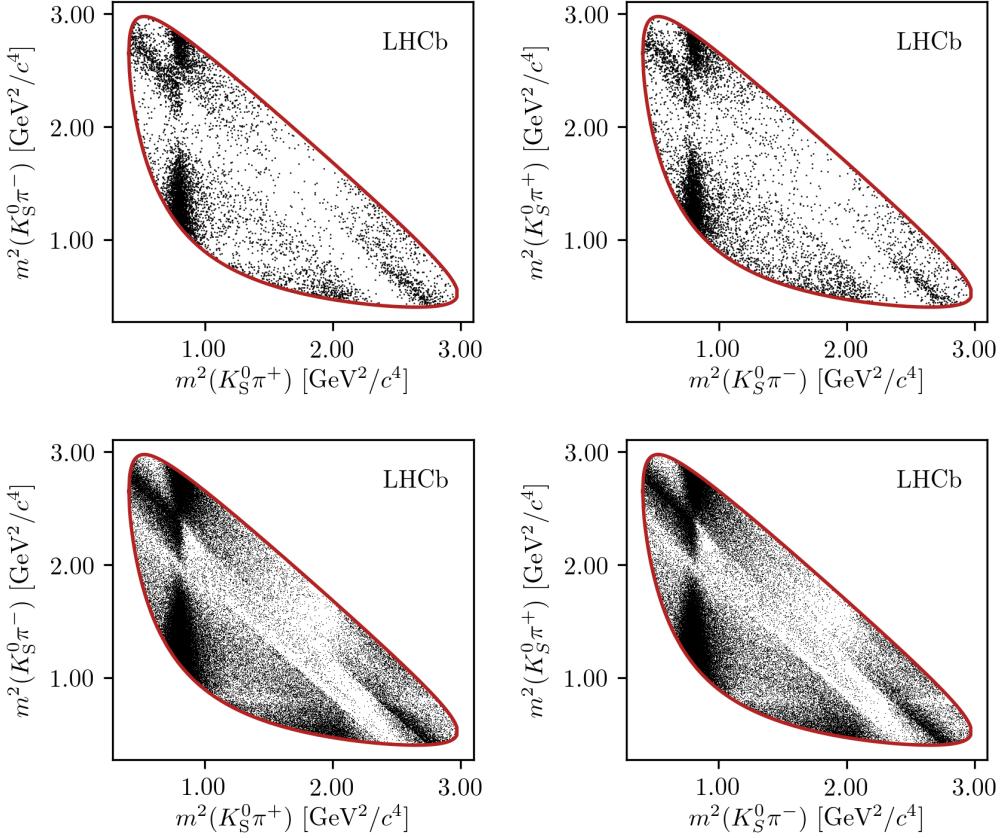
**Figure 5.5:** The spectrum of  $m_B$  in the (left)  $B^\pm \rightarrow DK^\pm$  and (right)  $B^\pm \rightarrow D\pi^\pm$  samples where  $D \rightarrow K_S^0\pi^+\pi^-$  and the  $K_S^0$  meson is reconstructed in the DD category, after the full selection has been applied.

## 1896 5.2 Signal selection efficiencies



1897 The efficiency of each step of the selection on signal decays can be investigated using  
 1898 simulated decays. In the  $B^\pm \rightarrow D\pi^\pm$  channel, only decays that were placed in  
 1899 the "test" sample when training the BDT are used, in order to avoid overesti-  
 1900 mating the efficiency.

1901 In general, the total selection efficiency up until the PID requirements, including  
 1902 the offline stage and the effect of the geometrical LHCb acceptance, is about 1  
 1903 permille, slightly higher for  $B^\pm \rightarrow DK^\pm$  than  $B^\pm \rightarrow D\pi^\pm$  due to the different  
 1904 decay channels. The PID requirements are investigated separately in Section 5.2.1 below, using samples of calibration data.  
 1905 The overall selection efficiency does not impact the measurement at all, because  
 1906 the observables of interest are sensitive *only* to the distribution of decays over the  
 1907 Dalitz plot (except, of course, in the sense that a higher signal efficiency is desirable  
 1908 for the measurement).

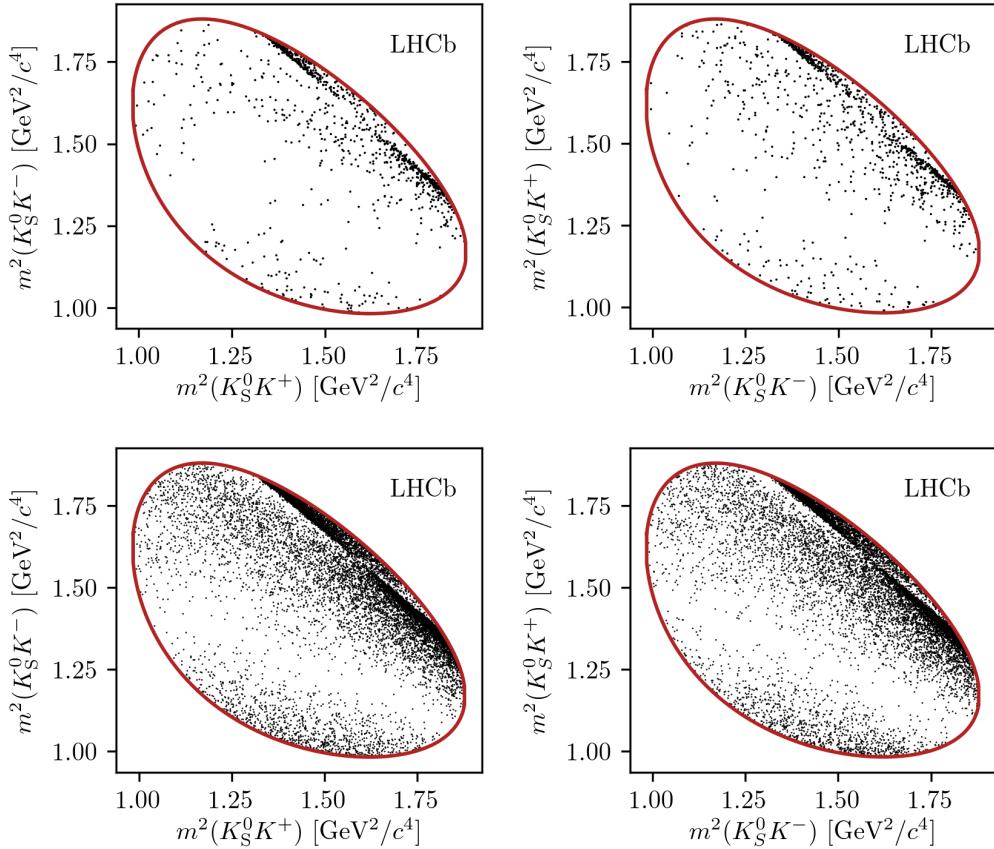


**Figure 5.6:** Dalitz plots of (left)  $B^+ \rightarrow Dh^+$  and (right)  $B^- \rightarrow Dh^-$  candidates in the signal region, in the (top)  $B^\pm \rightarrow DK^\pm$  and (bottom)  $B^\pm \rightarrow D\pi^\pm$  channels where  $D \rightarrow K_S^0 \pi^+ \pi^-$ . The LL and DD categories have been combined.

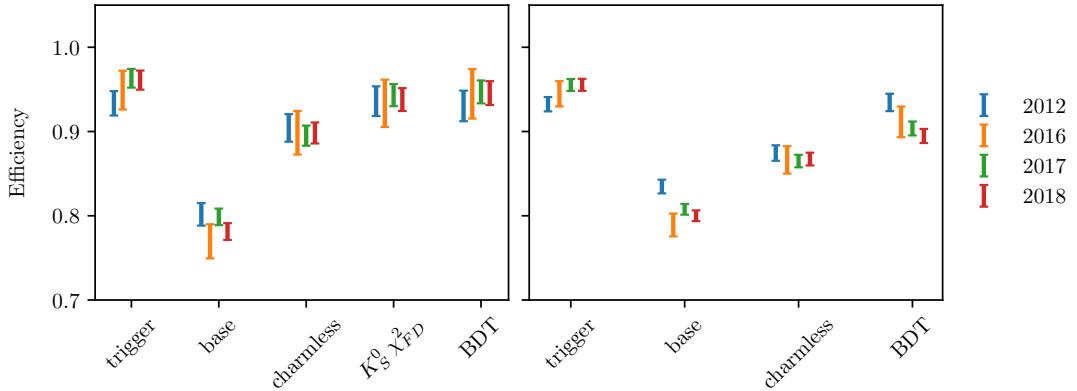
because it leads to larger signal yields). Likewise, it makes no difference that the overall selection efficiencies differ slightly between  $B^\pm \rightarrow DK^\pm$  and  $B^\pm \rightarrow D\pi^\pm$  decays, as long as the efficiency profile over the Dalitz plot is identical between the two decay channels. This is confirmed separately in Section 5.2.2 below.

The efficiencies of each individual selection step are shown in Fig. 5.8, obtained using simulated  $B^\pm \rightarrow D\pi^\pm$  decays. The main reason that some signal decays do not survive the base requirement is the  $p_{\text{companion}} < 100 \text{ GeV}/c$  requirement, which is in place to ensure that the PID performance for the companion is good. For decays with  $p_{\text{companion}} > 100 \text{ GeV}/c$ , only about 60 % of  $B^\pm \rightarrow DK^\pm$  decays survive the subsequent  $PIDK > 4$  requirement and the cross-feed from misidentified  $B^\pm \rightarrow D\pi^\pm$  decays is 50 % larger than in the current selection. Thus, loosening this requirement leads to little statistical gain, while leading to larger systematic effects from the crossfeed background.

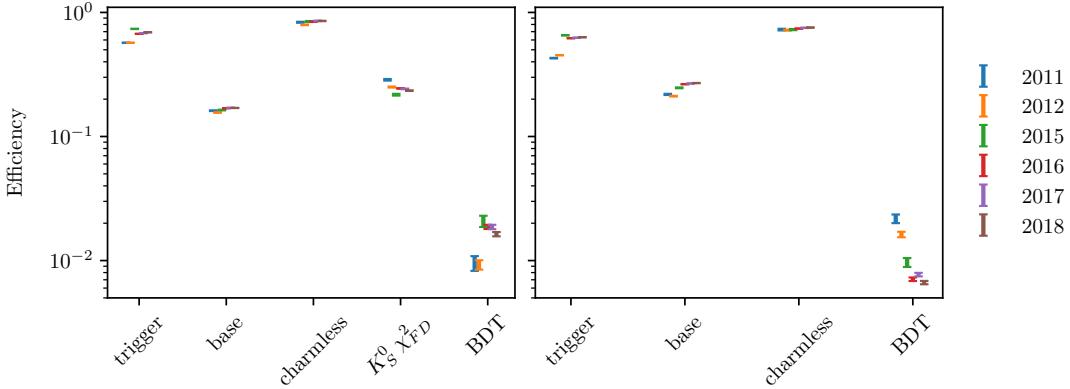
An equivalent plot for the combinatorial background is shown in Fig. 5.9, using  $B^\pm \rightarrow D(\rightarrow K_S^0 \pi^+ \pi^-)\pi^\pm$  candidates in data with a reconstructed  $B$  mass



**Figure 5.7:** Dalitz plots of (left)  $B^+ \rightarrow Dh^+$  and (right)  $B^- \rightarrow Dh^-$  candidates in the signal region, in the (top)  $B^\pm \rightarrow DK^\pm$  and (bottom)  $B^\pm \rightarrow D\pi^\pm$  channels where  $D \rightarrow K_S^0 K^\pm K^\mp$ . The LL and DD categories have been combined.



**Figure 5.8:** The efficiency of each selection step in samples of simulated  $B^\pm \rightarrow D(\rightarrow K_S^0 \pi^+ \pi^-) \pi^\pm$  signal decays in the (left) LL and (right) DD categories. The selection steps are applied on top of each other, from left to right on the horizontal axis. The samples are split by year.



**Figure 5.9:** The efficiency of each selection step in samples of  $B^\pm \rightarrow D(\rightarrow K_S^0 \pi^+ \pi^-) \pi^\pm$  candidates in data where the reconstructed  $B$  mass is above  $5600 \text{ MeV}/c^2$ , meaning the candidates stem from combinatorial background. The efficiency is shown for candidates in the (left) LL and (right) DD categories. The selection steps are applied on top of each other, from left to right on the horizontal axis. The samples are split by year. Notice the logarithmic scale on the vertical axis.

above  $5600 \text{ MeV}/c^2$ ; it can be seen that the BDT is extremely efficient at rejection combinatorial background, but that the base requirements and the requirement on the  $K_S^0$  flight distance also remove a decent amount of background.

### 5.2.1 Efficiency of the PID requirements

The efficiencies of the PID requirements on the companion enter the yield parameterisations of the mass fits in Section 5.4 and 5.5 and must therefore be known. They are determined using samples of calibration data selected without relying on PID variables, as implemented in the PIDCalib frame work [?]. Reasonably pure samples of pion and kaon tracks are obtained from  $D^0 \rightarrow K^-\pi^+$  decays, where the  $D$  meson originates in a  $D^{*+} \rightarrow D^0\pi^+$  decay and can therefore be flavour tagged. The remaining background is subtracted via the *sPlot* [?] procedure, based on a two-dimensional fit of the  $m(K^-\pi^+)$  and  $m(D^0\pi^+) - m(D^0)$  distributions. The obtained weights are employed to calculate the average efficiency of the requirement on PIDK for a number of bins in the momentum and pseudorapidity of the calibration tracks, and the number of charged tracks in the detector, thus constructing a three-dimensional efficiency lookup table. The procedure is carried out for each PID requirement, companion species, data-taking year, track charge, and magnet polarity. Based on these tables, expected PID efficiencies for the  $B^\pm \rightarrow D\pi^\pm$  and  $B^\pm \rightarrow DK^\pm$  signal decays are calculated that take the kinematical distribution and detector occupancy in the BPGGSZ data samples into account, by using the high-purity sample of  $B^\pm \rightarrow D\pi^\pm$  candidates in the signal region as a reference. The dominating

**Table 5.4:** PID efficiencies obtained with the `PIDCalib` tool. The uncertainty incorporates statistical uncertainty due to the size of the reference sample, the systematic uncertainty due to the choice of binning scheme in `PIDCalib`, and a systematic uncertainty due to the `sWeight` calculation in `PIDCalib` of 0.1 %.

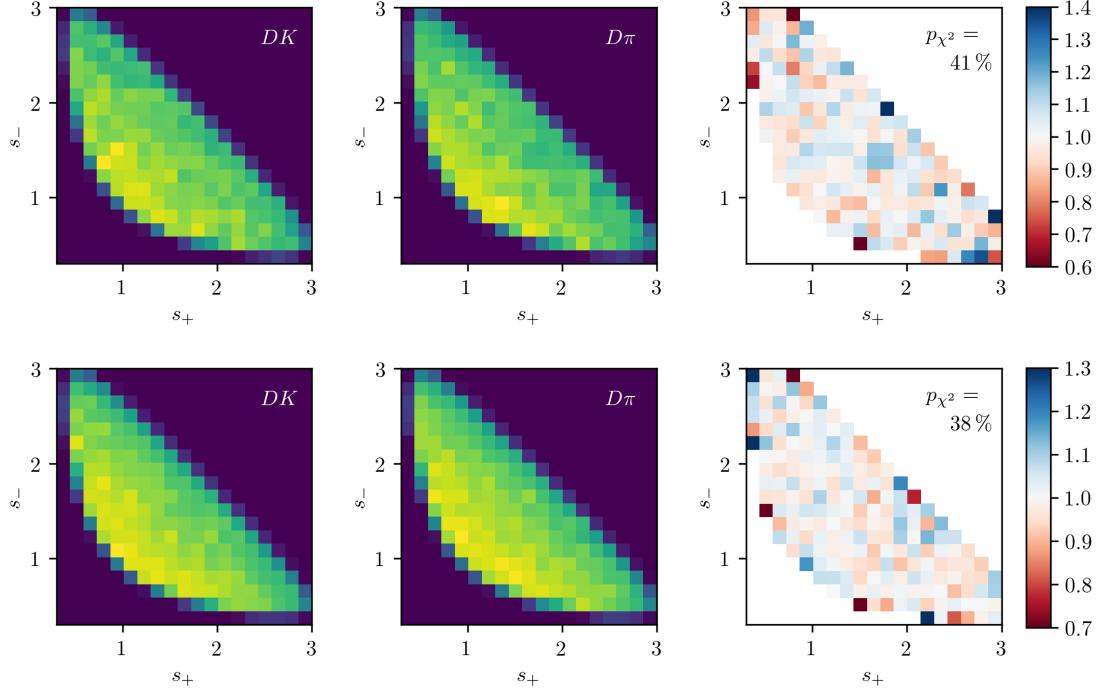
Efficiency	Particle	$D$ final state	$\varepsilon_{\text{PID}} (\%)$	
			LL	DD
Run I and II				
Correct ID	Kaon	$D \rightarrow K_S^0 \pi^+ \pi^-$	$86.74 \pm 0.13$	$86.90 \pm 0.22$
		$D \rightarrow K_S^0 K^+ K^-$	$86.22 \pm 0.26$	$86.56 \pm 0.30$
	Pion	$D \rightarrow K_S^0 \pi^+ \pi^-$	$97.11 \pm 0.11$	$97.17 \pm 0.13$
		$D \rightarrow K_S^0 K^+ K^-$	$97.07 \pm 0.11$	$97.16 \pm 0.14$

1945 uncertainty on the efficiencies is statistical in nature, due to the finite size of the  
 1946 reference sample. In addition, systematic uncertainties are included due to the `sPlot`  
 1947 procedure, estimated at 0.1 % [?], and due to the choice of binning scheme, estimated  
 1948 by repeating the procedure using a number of alternative binning schemes. The  
 1949 final efficiency estimates are shown in Table 5.4, including all sources of uncertainty.

### 1950 5.2.2 Efficiency profile over the Dalitz plot

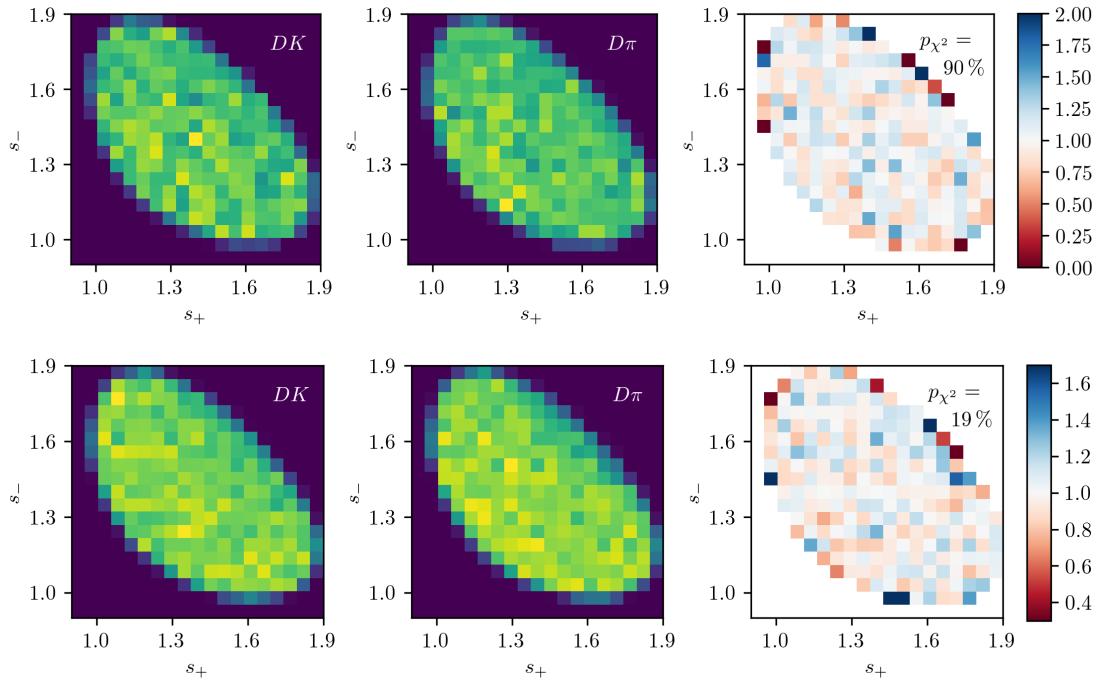
1951 The analysis strategy depends on sharing the  $F_i$  parameters between the  $B \rightarrow D\pi$   
 1952 and  $B \rightarrow DK$  channels. This is reasonable, since the phase-space dependence  
 1953 of the reconstruction efficiency is expected to be very similar between the two  
 1954 decays, given the similar kinematics; an assumption that is verified using samples  
 1955 of simulated decays. The full selection is applied to the samples. The  $B \rightarrow D\pi$   
 1956 sample of LL (DD) candidates includes about 63,000 (146,000) decays, and the  
 1957  $B \rightarrow DK$  samples include 60,000 (142,000) decays. For the  $D\pi$  mode, this is  
 1958 approximately equal to the number of decays in the full Run 1+2 data sample, and  
 1959 for  $B \rightarrow DK$  this is a factor of about 12 larger than the data sample. The decays  
 1960 were simulated with an equal decay probability across the  $D$ -decay phase space, so  
 1961 that any non-uniform distribution of reconstructed decays is completely determined  
 1962 by a phase-space dependent reconstruction and selection efficiency. Therefore the  
 1963 assumption that the phase-space dependence is identical between the  $B \rightarrow D\pi$  and  
 1964  $B \rightarrow DK$  channels is verified by seeing if the Dalitz coordinates are distributed  
 1965 differently between the samples of simulated  $B \rightarrow D\pi$  and  $B \rightarrow DK$  decays.

1966 This is investigated with two statistical tests. The first is a  $\chi^2$  comparison  
 1967 of 2D histograms of the distribution of  $m^2(K_S^0 \pi^+)$  and  $m^2(K_S^0 \pi^-)$  in the different  
 1968  $B \rightarrow D\pi$  and  $B \rightarrow DK$  channels. These histograms, and the ratio between them,  
 1969 are shown in Figs. 5.10 and 5.11, along with the  $p$ -values from the  $\chi^2$  tests. It can



**Figure 5.10:** The  $(s_+, s_-)$  distribution in simulated samples of (left)  $B \rightarrow DK$  decays and (center)  $B \rightarrow D\pi$  decays where  $D \rightarrow K_S^0\pi^+\pi^-$ , as well as (right) the ratio between the two histograms (corrected for difference in sample sizes). The plots are shown for candidates in the (top) LL and (bottom) L categories. The  $p$  values are the results of  $\chi^2$  compatibility tests between the two histograms.

be seen that, in all cases, the probability of obtaining the two histograms assuming that they share the same underlying distribution has a reasonable value, and that there is no clear trend in the ratio plots. The second test is a Kolmogorov-Smirnov test [?] of the compatibility of the one-dimensional distributions of  $m^2(K_S^0\pi^+)$ ,  $m^2(K_S^0\pi^-)$ , and  $m^2(\pi^+\pi^-)$ . These distributions, and the corresponding  $p$ -values, are shown in Fig. 5.12 and 5.13. Again, all the  $p$  values are reasonable. Therefore, it is concluded that there are no statistically significant differences between the phase-space dependence of the reconstruction and selection efficiency between the  $B \rightarrow D\pi$  and  $B \rightarrow DK$  channels, given the present sample sizes. Because the simulation samples have approximately the same amount of decays as data (or significantly more, in the  $B \rightarrow DK$  case), any potential differences will be negligible with data yields. Thus, sharing the  $F_i$  parameters between the  $B \rightarrow D\pi$  and  $B \rightarrow DK$  channels is viable, and no efficiency correction is necessary.

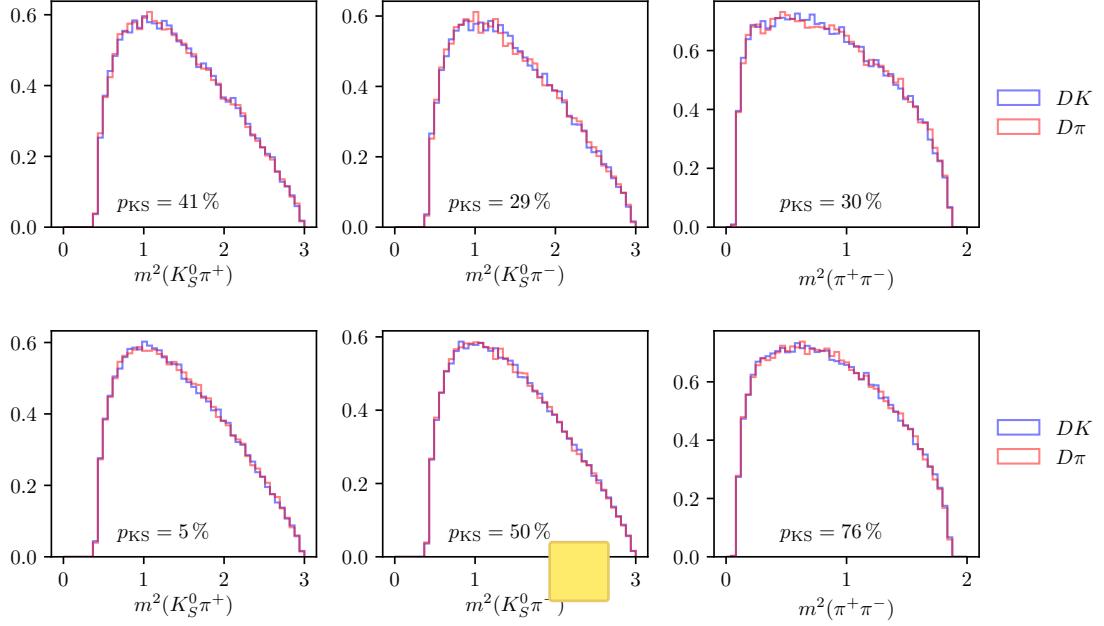


**Figure 5.11:** The  $(s_+, s_-)$  distribution in simulated samples of (left)  $B \rightarrow DK$  decays and (center)  $B \rightarrow D\pi$  decays where  $D \rightarrow K_S^0 K^+ K^-$ , as well as (right) the ratio between the two histograms (corrected for differences in sample sizes). The plots are shown for candidates in the (top) LL and (bottom) DD categories. The  $p$  values are the results of  $\chi^2$  compatibility tests between the two histograms.

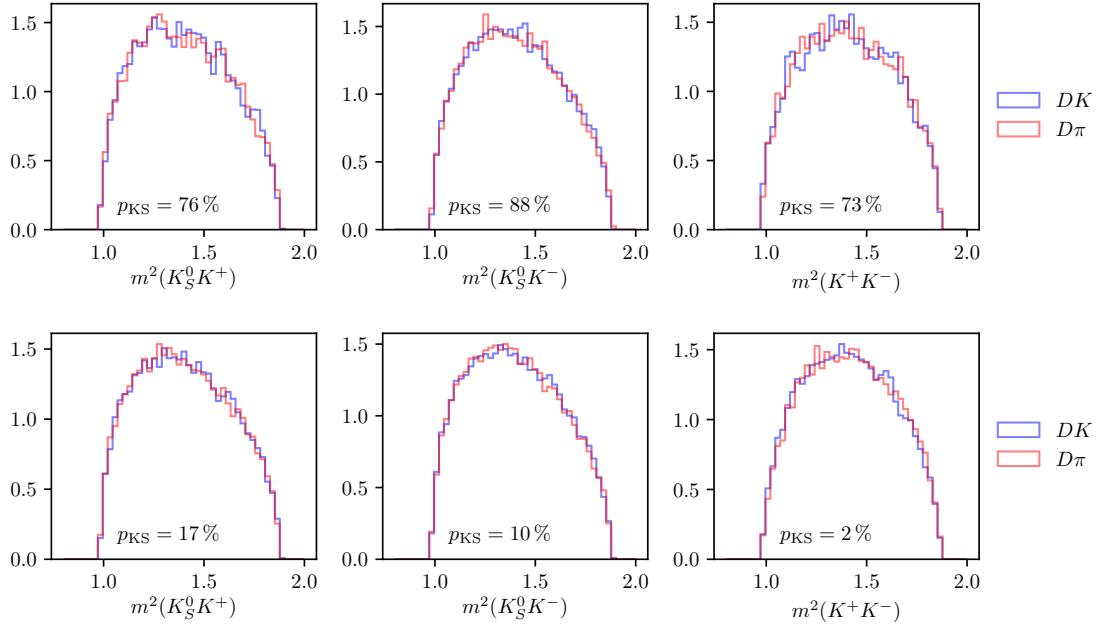
### 1983 5.3 Background studies

1984 A wide range of backgrounds can potentially pollute the sample of signal candidates.  
 1985 The backgrounds group into three categories depending on how they are treated  
 1986 in the analysis:

- 1987 • Backgrounds that can be effectively removed in the selection
  - 1988 • Backgrounds that are only present at a level where the impact on the  
 1989 measurement result is small, and which do therefore not have to be modelled
  - 1990 • Backgrounds that are present at a level where they have to be modelled in  
 1991 the fit to data, and cannot effectively be rejected further in the selection
- 1992 The latter category comprises of combinatorial background, which remains present  
 1993 at a non-negligible level after the application of the BDT described in Section 5.1.2;  
 1994 contributions from a number of partly reconstructed  $B \rightarrow Dh^\pm X$  decays, where  
 1995  $X$  denotes a pion or photon that is not included in the reconstructed decay, and  
 1996 which can only be separated from signal decays by their  $m(Dh)$  distribution; and



**Figure 5.12:** One-dimensional distributions of  $m^2(K_S^0\pi^+)$ ,  $m^2(K_S^0\pi^-)$ , and  $m^2(\pi^+\pi^-)$  in simulated (blue)  $B^\pm \rightarrow DK^\pm$  and (red)  $B^\pm \rightarrow D\pi^\pm$  decays where  $D \rightarrow K_S^0\pi^+\pi^-$  in the (top) LL and (bottom) DD categories. The  $p$  values are the results of Kolmogorov-Smirnov compatibility tests between the distributions.



**Figure 5.13:** One-dimensional distributions of  $m^2(K_S^0K^+)$ ,  $m^2(K_S^0K^-)$ , and  $m^2(\pi^+\pi^-)$  in simulated (blue)  $B^\pm \rightarrow DK^\pm$  and (red)  $B^\pm \rightarrow D\pi^\pm$  decays where  $D \rightarrow K_S^0K^+K^-$  in the (top) LL and (bottom) DD categories. The  $p$  values are the results of Kolmogorov-Smirnov compatibility tests between the distributions.

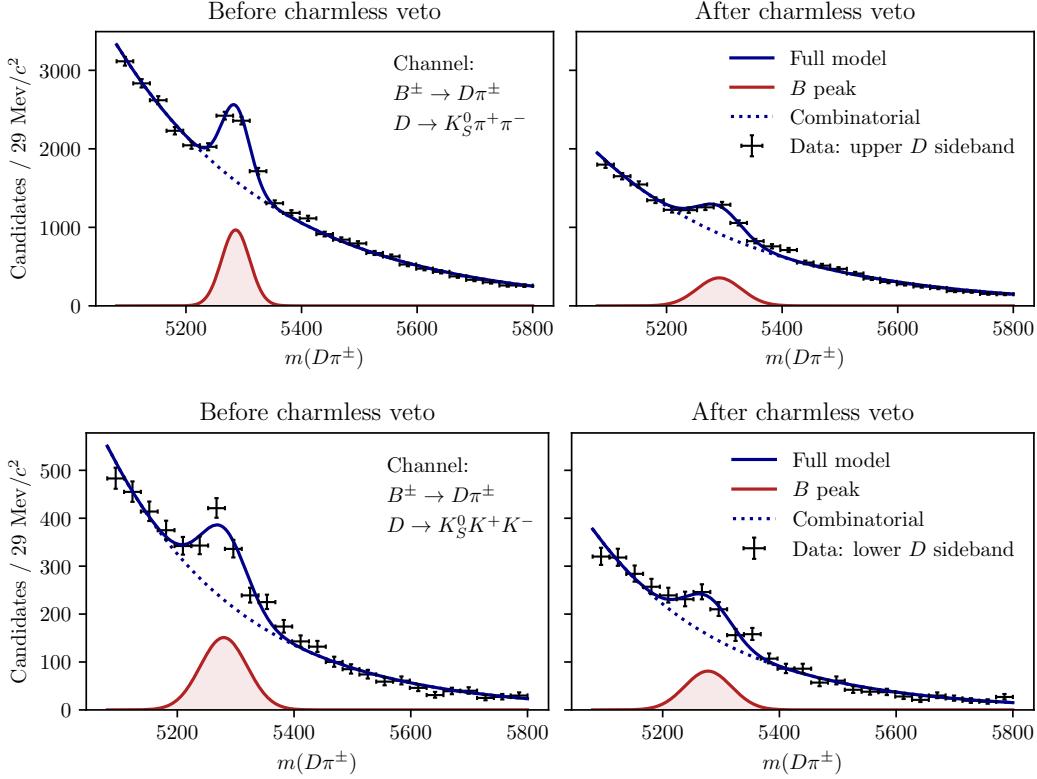
1997 finally  $B^\pm \rightarrow D\pi^\pm$  decays that are categorised as  $B^\pm \rightarrow DK^\pm$  decays in the particle-  
 1998 identification step and vice-versa. These background sources are described in detail  
 1999 in Section 5.4. This section focuses on backgrounds that led to specific requirements  
 2000 in the selection or proved to be small enough to not merit special treatment.

### 2001 5.3.1 Charmless decays

2002 There is potentially a so-called *charmless* background present in data, consisting  
 2003 of  $B^\pm \rightarrow K_S^0 h^+ h^- h'^\pm$  decays. These have the same final state as the signal decay,  
 2004 but no intermediate  $D$  meson. Because all final state particles are reconstructed,  
 2005 this background peaks in the  $B$  mass spectrum. This background is suppressed  
 2006 by requiring the reconstructed  $B$  and  $D$  decay vertices to be separated in the  
 2007  $z$  direction; specifically by requiring that  $\Delta z_{\text{significance}}^{D-B} > 0.5$ , where  $\Delta z_{\text{significance}}^{D-B}$   
 2008 was defined in Eq. (5.1). The remaining background level can be investigated  
 2009 by investigating the  $D$  mass sidebands. 

2010 However, the use of the DecayTreeFitter  $\chi^2$  as an input variable in the BDT  
 2011 removes essentially all of the  $D$  (and  $K_S^0$ ) sideband, due to the mass constraints  
 2012 in the decay chain fit. Therefore separate BDT's are trained for LL and DD  
 2013 candidates without the  $\chi^2$  as an input variable, and used when selecting candidates  
 2014 for the background studies presented in this section, and the following. In a similar  
 2015 manner, all mass window requirements are made on the *default* reconstructed  
 2016 masses, obtained with no use of DecayTreeFitter. The overlap of the two sets of  
 2017 selected candidates in the signal  $B$ -mass window is above 95 %.

2018 The reconstructed  $B$  mass spectrum is shown for  $B^\pm \rightarrow D\pi^\pm$  candidates in the  
 2019  $D$  sidebands in Fig. 5.14, both before and after making a requirement on  $\Delta z_{\text{significance}}^{D-B}$ .  
 2020 A peak is clearly visible, the size of which is reduced by the requirement. This peak  
 2021 is partly due to a contribution from  $B^\pm \rightarrow K_S^0 \pi^+ \pi^- \pi^\pm$  decays ( $B^\pm \rightarrow K_S^0 K^+ K^- \pi^\pm$   
 2022 decays) in the  $D \rightarrow K_S^0 \pi^+ \pi^-$  ( $D \rightarrow K_S^0 K^+ K^-$ ) channel, and partly due to real  
 2023 signal decays that leak into the  $D$  sidebands. The number of real signal decays can  
 2024 be calculated from the yield obtained in the fit of Section 5.4, and the reconstructed  
 2025  $m_D$  distribution in simulated signal decays. Subtracting this contribution, it is  
 2026 estimated that approximately 450 (200) charmless decays are present in the  $K_S^0 \pi^+ \pi^-$   
 2027 ( $K_S^0 K^+ K^-$ ) data samples. In similar fashion, Fig. 5.15 shows the  $m_B$  spectra for  
 2028  $B^\pm \rightarrow DK^\pm$  candidates in the  $D$  sidebands. In these plots, the peaks are at  
 2029  $m_B$  values that are lower (higher) than the  $B$  mass in the  $K_S^0 \pi^+ \pi^-$  ( $K_S^0 K^+ K^-$ )  
 2030 categories, because they stem from real  $B^\pm \rightarrow K_S^0 K^+ K^- \pi^\pm$  decays where a kaon is  
 2031 mis-reconstructed as a pion or a pion is misreconstructed as a kaon, respectively.  
 2032 The total contribution of charmless decays in the  $B^\pm \rightarrow DK^\pm$  data samples is

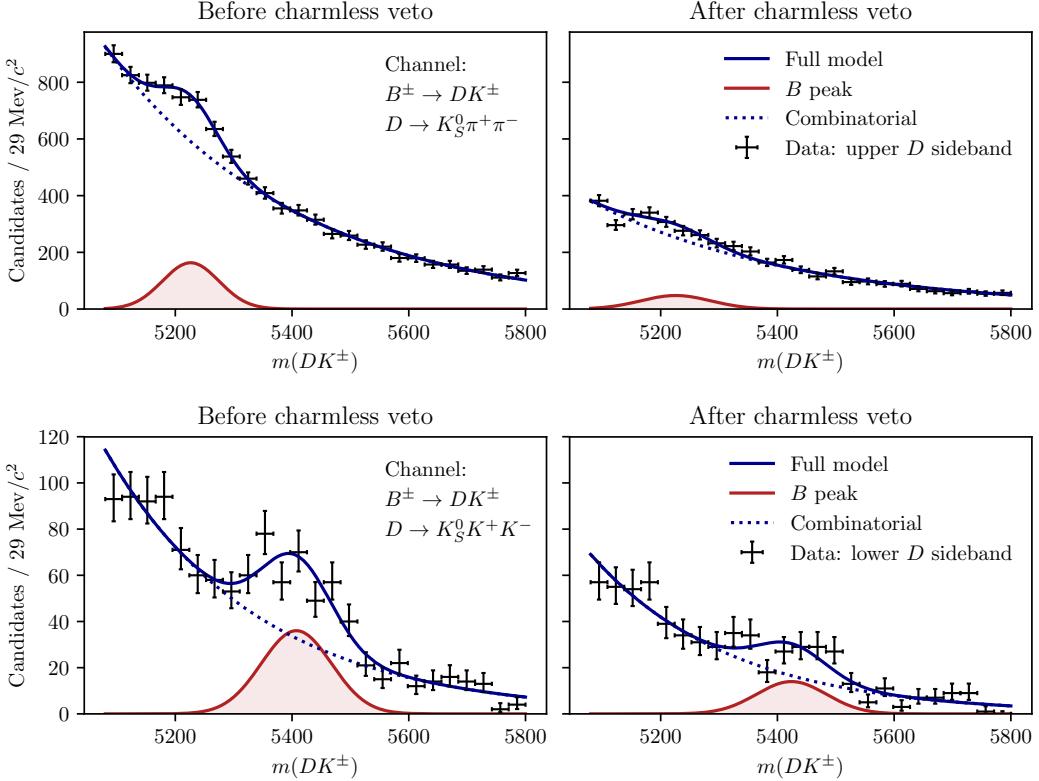


**Figure 5.14:** The  $B$  mass distribution of (top)  $B^\pm \rightarrow D(\rightarrow K_S^0 \pi^+ \pi^-)\pi^\pm$  and (bottom)  $B^\pm \rightarrow D(\rightarrow K_S^0 K^+ K^-)\pi^\pm$  candidates reconstructed in both the LL and DD categories, residing in the upper  $D$  mass sideband  $m_D \in [1910, 1960] \text{ MeV}/c^2$  for  $D \rightarrow K_S^0 \pi^+ \pi^-$  and in the lower sideband  $m_D \in [1910, 1960] \text{ MeV}/c^2$  for  $D \rightarrow K_S^0 K^+ K^-$ , with (left) no requirement on  $\Delta z_{\text{significance}}^{BD}$  and (right) after a requirement of  $\Delta z_{\text{significance}}^{BD} > 0.5$ .

estimated to be about 200 decays. As described further in Section 5.6.11, the presence of a charmless background at these levels has a negligible impact on the measurement results.

### 5.3.2 Background from four-body $D$ decays

A similar potential background is from real  $B^\pm \rightarrow Dh^\pm$  decays where the  $D$  meson decays directly to the  $\pi^+\pi^-h^+h^-$  final state, without an intermediate  $K_S^0$  meson. This background can be investigated by looking for a peak in the  $B$  mass spectrum for candidates in the  $K_S^0$  sideband, as illustrated in Fig. 5.16. The figure shows the spectrum in the final data sample, illustrating the significant effect of making the requirement on the  $K_S^0$  flight distance that was discussed in Section 5.1.1. The BDT that does *not* rely on the DTF  $\chi^2$  has been used to suppress combinatorial background. The remaining peak after requiring  $\chi_{\text{FD}}^2 > 49$  is completely accounted for by real signal decays that leak into the  $K_S^0$  sideband.

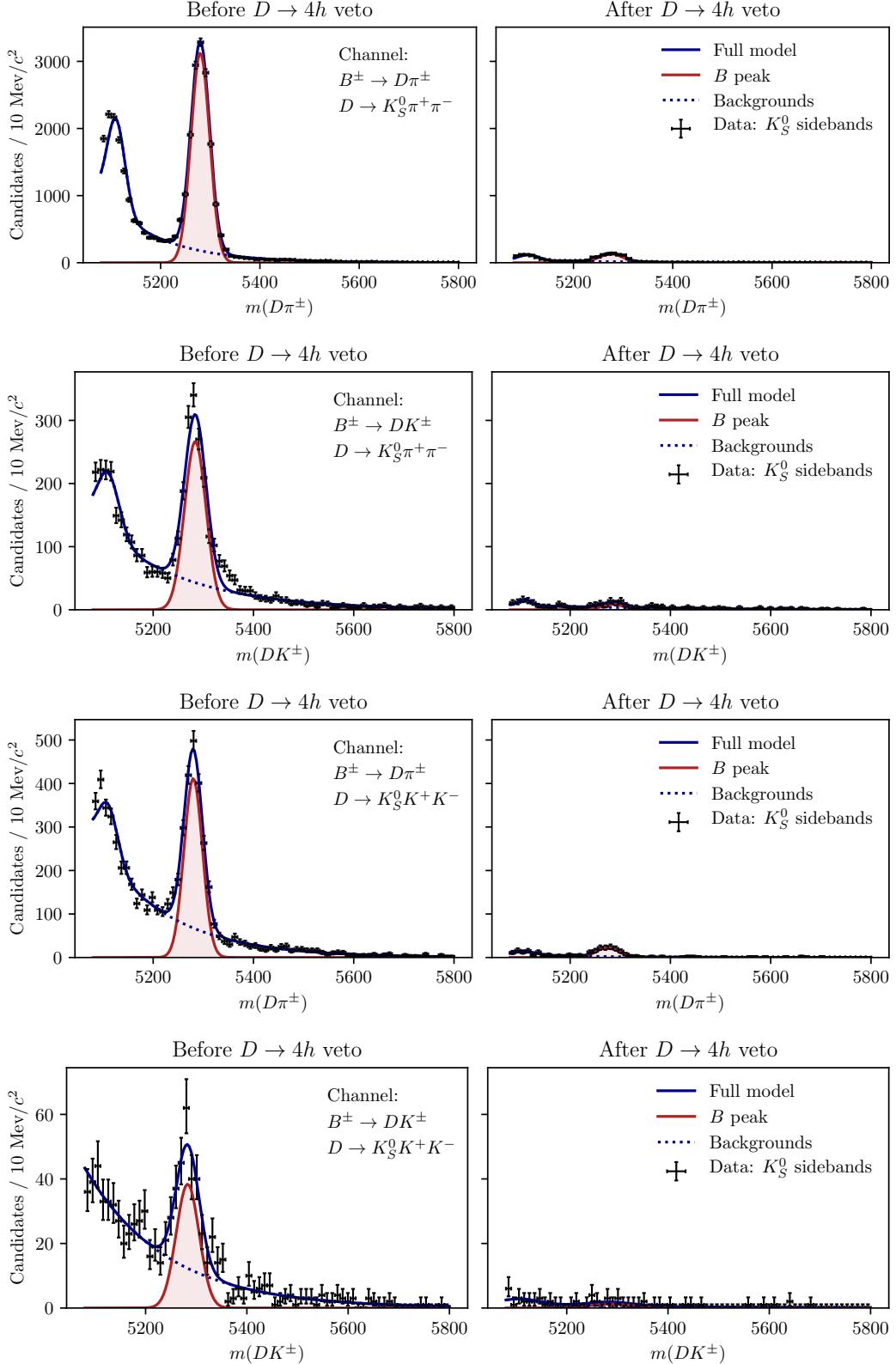


**Figure 5.15:** The  $B$  mass distribution of (top)  $B^\pm \rightarrow D(\rightarrow K_S^0 \pi^+ \pi^-)K^\pm$  and (bottom)  $B^\pm \rightarrow D(\rightarrow K_S^0 K^+ K^-)K^\pm$  candidates reconstructed in both the LL and DD categories, residing in the upper  $D$  mass sideband  $m_D \in [1910, 1960] \text{ MeV}/c^2$  for  $D \rightarrow K_S^0 \pi^+ \pi^-$  and in the lower sideband  $m_D \in [1910, 1960] \text{ MeV}/c^2$  for  $D \rightarrow K_S^0 K^+ K^-$ , with (left) no requirement on  $\Delta z_{\text{significance}}^{BD}$  and (right) after a requirement of  $\Delta z_{\text{significance}}^{BD} > 0.5$ .

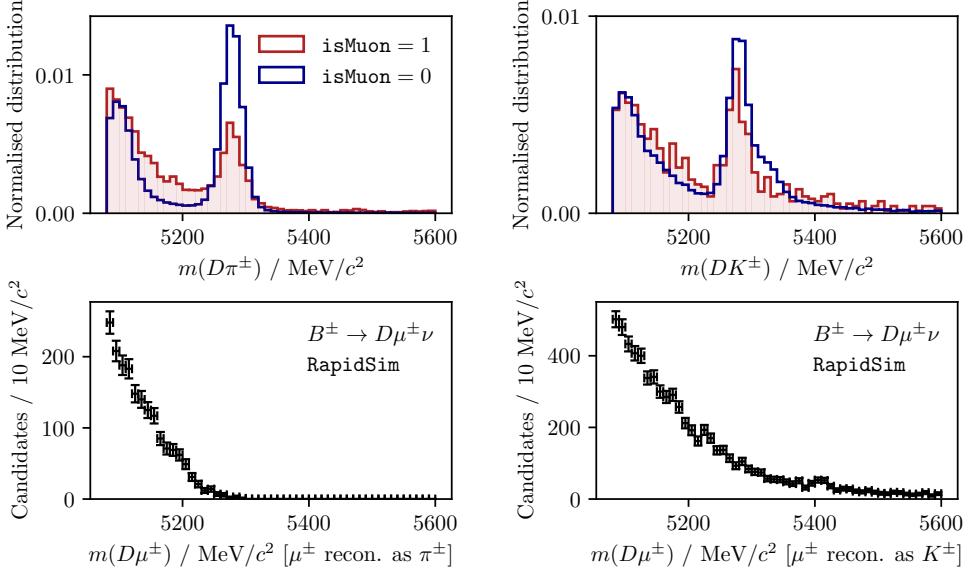
### 2046 5.3.3 Semi-leptonic backgrounds

2047 The data sample has a minor background from  $B \rightarrow D\mu\nu_\mu X$  decays, visible  
 2048 in the  $B$  mass spectrum when the companion is required to satisfy `isMuon=1`.  
 2049 This is shown in Fig. 5.17 for both the  $B^\pm \rightarrow DK^\pm$  and  $B^\pm \rightarrow D\pi^\pm$  channels  
 2050 where  $D \rightarrow K_S^0 \pi^+ \pi^-$ . The  $B$  mass spectra for simulated  $B^\pm \rightarrow D\mu^\pm \nu_\mu$  decays  
 2051 reconstructed in each category are also shown, from simulation samples produced via  
 2052 `RapidSim`. The background is very efficiently vetoed by requiring `IsMuon=0` on the  
 2053 companion. This requirement removes approximately 85 % of the background decays,  
 2054 as estimated using the `PIDCalib` calibration samples and the  $(p, p_T)$  distribution  
 2055 for the muon in the `RapidSim` samples. The fraction of signal candidates for  
 2056 which the companion satisfies `IsMuon=1` in simulated signal samples is  $\leq 0.9\%$   
 2057 so the impact on signal yield is small.

2058 The analogous  $B \rightarrow D e \nu_e X$  background is investigated by inspecting the  $B$   
 2059 mass spectra after making requirements on `PIDE` for the companion candidate, but



**Figure 5.16:** The  $B$  mass spectrum in the  $K_S^0$  sideband where  $m_{K_S^0} \in [467, 482] \text{ MeV}/c^2$  or  $m_{K_S^0} \in [512, 527] \text{ MeV}/c^2$  (left) without a requirement on the  $K_S^0$  flight distance significance, and (right) after the requirement implemented in the analysis.

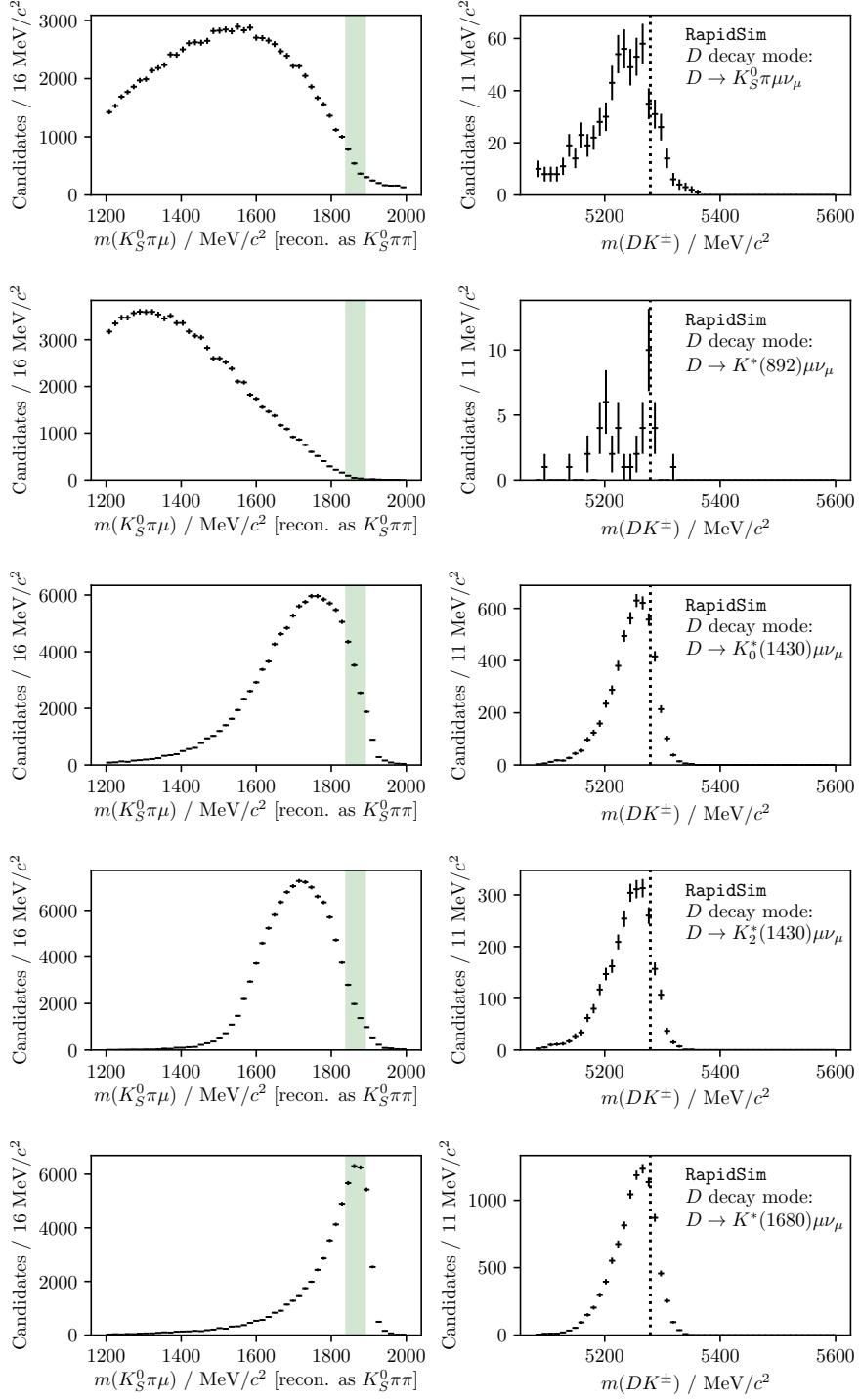


**Figure 5.17:** (Top) The  $m_B$  spectra in data split by the value of `Bach_isMuon` for (left) the  $D\pi^\pm$  and (right) the  $DK^\pm$  samples where  $D \rightarrow K_S^0\pi^+\pi^-$ . The two histograms are normalised independently, so that the distributions can be compared. The fractions of candidates in data (with  $m_B \in [5080, 5800] \text{ MeV}/c^2$ ) that satisfy `Bach_isMuon=1` are 1.6% and 1.8% for the  $D\pi^\pm$  and  $DK^\pm$  channels respectively. (Bottom) the `RapidSim` mass spectra for  $B^\pm \rightarrow D^0\mu^\pm\nu_\mu$  decays reconstructed in the (left)  $D\pi^\pm$  and (right)  $DK^\pm$  categories.

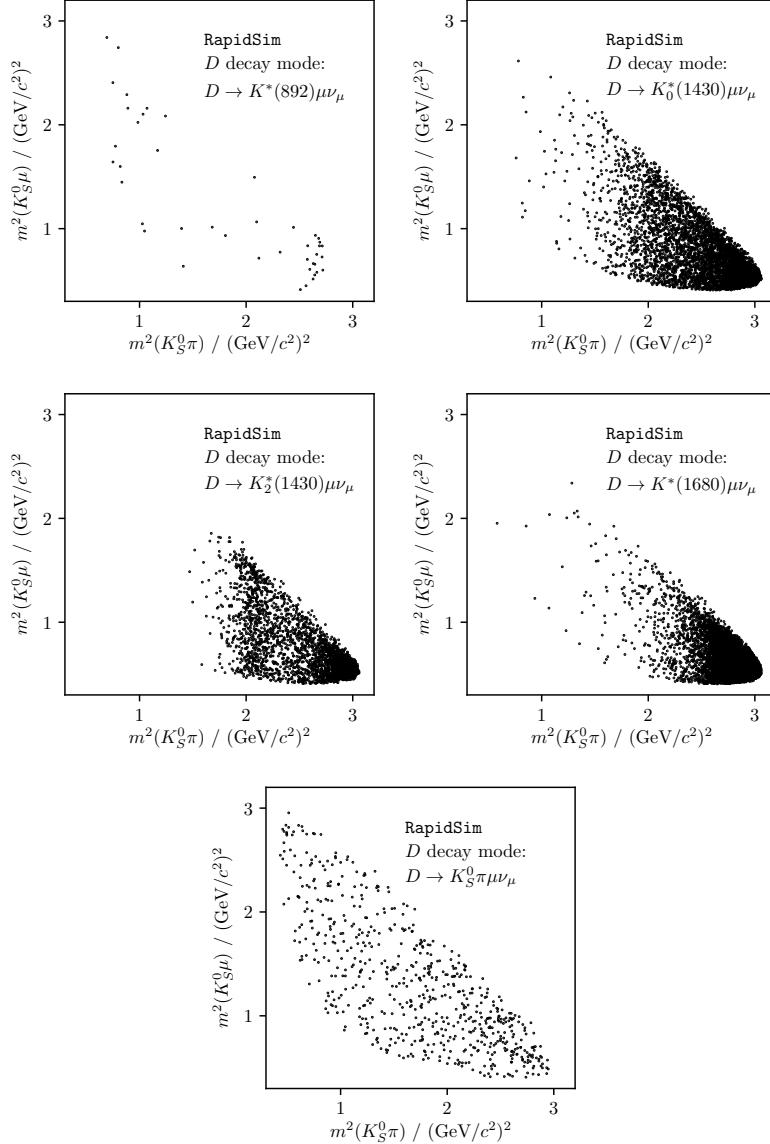
a presence of the semi-leptonic background in data is not visible and no electron veto is applied to the companion.

## Background from semi-leptonic D decays

There is a potential background from real  $B^\pm \rightarrow Dh^\pm$  decays where the  $D$  meson decays semi-leptonically:  $D^0 \rightarrow K_S^0\pi^-\ell^+\nu_\ell$ . This background is particularly dangerous because it peaks at the  $B$  mass, when the  $D$ -mass requirement is applied and it is reconstructed in the  $D \rightarrow K_S^0\pi^+\pi^-$  category. This is illustrated in Fig. 5.18 using `RapidSim` samples of  $B^\pm \rightarrow D(\rightarrow K^{*-}(X)(\rightarrow K_S^0\pi^-)\ell^+\nu_\ell)h^\pm$  decays for  $X = 892, 1430, 1680$ . The expected background yields relative to signal can be estimated by applying the  $B$  and  $D$  mass cuts to decays in the `RapidSim` samples, and using the relative branching ratios. Only the  $D^0 \rightarrow K^{*-}(892)\ell\nu_\ell$  branching fractions have been measured [?], but there is no reason to expect that higher  $K^*$  resonances should not contribute. To estimate their potential contribution,



**Figure 5.18:** The reconstructed (left)  $m(K_S^0\pi^+\pi^-)$  and (right)  $m(Dh)$  distributions in RapidSim samples of  $B^\pm \rightarrow DK^\pm$  decays where  $D^0 \rightarrow K_S^0\pi^-\mu^+\nu_\mu$ . The top plot is for PHSP decays, and the following plots show the distribution where the  $K_S^0\pi^-$  originate in the resonances  $K^{*-}(892)$ ,  $K^{*-}_0(1430)$ ,  $K^{*-}_2(1430)$ , and  $K^{*-}(1680)$ . The shapes for the  $D^0 \rightarrow K_S^0\pi^-e^+\nu_e$  case are almost identical.



**Figure 5.19:** Dalitz distribution for  $D \rightarrow K_S^0 \pi \mu \nu_\mu$  decays in RapidSim, where the  $K_S^0 \pi^-$  originate in the resonances  $K^{*-}(892)$ ,  $K^{*-}_0(1430)$ ,  $K^{*-}_2(1430)$ , and  $K^{*-}(1680)$ , as well as for a flat PHSP distribution.

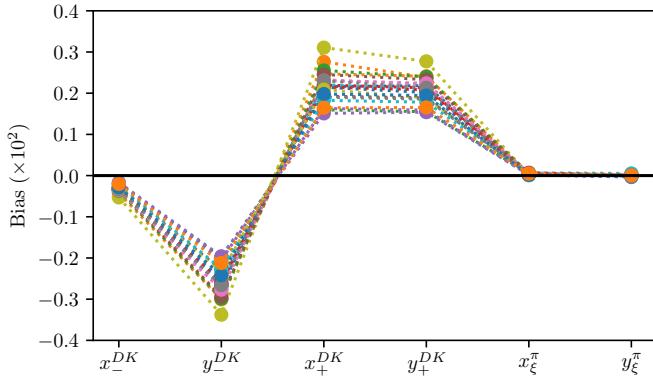
2073 the branching ratios are approximated by

$$\text{BR}[D^0 \rightarrow K^{*-}(X)(\rightarrow K_S^0 \pi^-) \ell \nu_\ell] \simeq \frac{\text{BR}[D^0 \rightarrow K^{*-}(X)(\rightarrow K_S^0 \pi^-) \pi^+]}{\text{BR}[D^0 \rightarrow K^{*-}(892)(\rightarrow K_S^0 \pi^-) \pi^+]} \text{BR}[D^0 \rightarrow K^{*-}(892)(\rightarrow K_S^0 \pi^-) \ell \nu_\ell]$$

2074 because all the relevant  $D^0 \rightarrow K^{*-}(\rightarrow K_S^0 \pi^-) \pi^+$  branching fractions are known [?].  
2075 The efficiencies and branching ratios relative to the signal channel are given in  
2076 Table 5.5. It is clear that the higher  $K^*$  resonances are important: the smaller  
2077 branching ratios are compensated for by a higher selection efficiency, due to the

**Table 5.5:** The selection efficiencies of  $B^\pm \rightarrow DK^\pm$  decays where  $D^0 \rightarrow K_S^0\pi^-\ell^+\nu_\ell$  when reconstructed in the  $D \rightarrow K_S^0\pi^+\pi^-$  mode in RapidSim relative to the signal selection efficiencies, for a number of decay modes: PHSP as well as resonant production where the  $K_S^0\pi^-$  pair originates in one of several  $K^*$  resonances. The relative branching ratios are also shown, calculated as explained in the main text, as well as the predicted relative yields.

Mode	$\epsilon_{bkg}/\epsilon_{signal}$ (%)	$\Gamma_{bkg}/\Gamma_{signal}$ (%)	$N_{bkg}/N_{signal}$ (%)
$D \rightarrow K_S^0\pi^-\mu^+\nu_\mu$ (PHSP)	$0.92 \pm 0.05$	$18.3 \pm 14.8$	$0.17 \pm 0.14$
$D \rightarrow (K_S^0\pi^-)_{K^{*-}(892)}\mu^+\nu_\mu$	$0.06 \pm 0.01$	$22.3 \pm 3.2$	$0.013 \pm 0.003$
$D \rightarrow (K_S^0\pi^-)_{K_0^{*-}(1430)}\mu^+\nu_\mu$	$7.3 \pm 0.1$	$3.7 \pm 0.8$	$0.27 \pm 0.06$
$D \rightarrow (K_S^0\pi^-)_{K_2^{*-}(1430)}\mu^+\nu_\mu$	$3.7 \pm 0.1$	$0.5 \pm 0.3$	$0.02 \pm 0.01$
$D \rightarrow (K_S^0\pi^-)_{K^{*-}(1680)}\mu^+\nu_\mu$	$24.4 \pm 0.3$	$0.6 \pm 0.5$	$0.15 \pm 0.12$
$D \rightarrow K_S^0\pi^-e^+\nu_e$ (PHSP)	$0.53 \pm 0.02$	$20.8 \pm 16.3$	$0.11 \pm 0.09$
$D \rightarrow (K_S^0\pi^-)_{K^{*-}(892)}e^+\nu_e$	$0.15 \pm 0.02$	$25.6 \pm 2.5$	$0.04 \pm 0.01$
$D \rightarrow (K_S^0\pi^-)_{K_0^{*-}(1430)}e^+\nu_e$	$6.3 \pm 0.1$	$4.2 \pm 0.8$	$0.26 \pm 0.05$
$D \rightarrow (K_S^0\pi^-)_{K_2^{*-}(1430)}e^+\nu_e$	$4.12 \pm 0.08$	$0.5 \pm 0.3$	$0.02 \pm 0.01$
$D \rightarrow (K_S^0\pi^-)_{K^{*-}(1680)}e^+\nu_e$	$10.0 \pm 0.2$	$0.7 \pm 0.5$	$0.07 \pm 0.05$
Total	-	-	$1.1 \pm 0.4$



**Figure 5.20:** Estimated biases on the measured observables due to the presence of  $D \rightarrow K_S^0\pi\ell\nu_\ell$  backgrounds, calculated while varying efficiencies and branching ratios within uncertainties.

2078 smaller phase-space of the missed neutrino. The total background yield is 1.1 %  
 2079 of the signal yield in both the  $B^\pm \rightarrow D\pi^\pm$  and  $B^\pm \rightarrow DK^\pm$  channels. However,  
 2080 there will be an additional contribution in the  $B^\pm \rightarrow DK^\pm$  channel from real  
 2081  $B^\pm \rightarrow D\pi^\pm$  decays with semi-leptonic  $D$  decays and a mis-identification of the  
 2082 companion. This background also peaks, and the yield is approximately 0.4 %  
 2083 of the  $B^\pm \rightarrow DK^\pm$  signal yield.

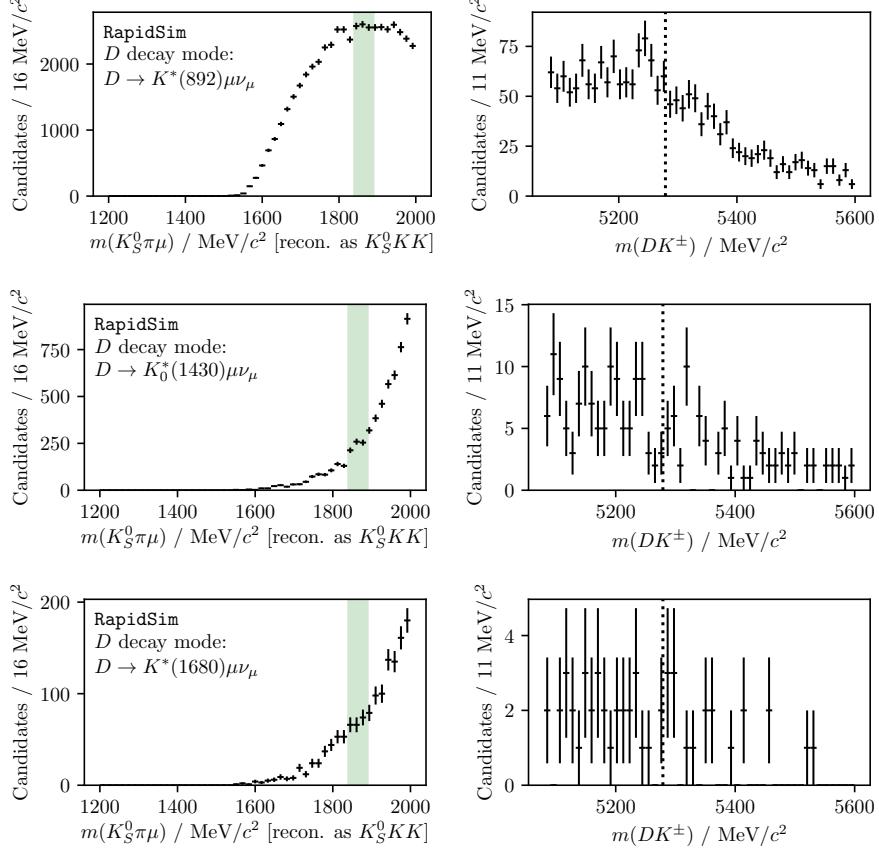
2084 The potential impact from the presence of the background is estimated by

- 2085     1. calculating the expected  $B^\pm \rightarrow D\pi^\pm$  and  $B^\pm \rightarrow DK^\pm$  yields in each bin for  
 2086       physics parameters similar to the world average values
- 2087     2. then calculating the background bin yields in each bin, using the relative  
 2088       branching fractions and efficiencies described above and taking the bin-  
 2089       distribution from the RapidSim samples. The RapidSim samples are produced  
 2090       using the `ISGW2` model in `EvtGen` [?], yielding the Dalitz distributions in  
 2091       Fig. 5.19.
- 2092     3. adding the signal and background yields, and fitting the new  $B^\pm \rightarrow D\pi^\pm$  and  
 2093        $B^\pm \rightarrow DK^\pm$  yields back with the default signal-yield expressions (including a  
 2094       fit of the  $F_i$  parameters)

2095     The obtained biases are shown in Fig. 5.20, where they are calculated a number  
 2096     of times, each time varying the efficiencies within statistical uncertainties and the  
 2097     relevant branching fractions within the measurement uncertainties. The systematic  
 2098     uncertainty due to the unknown branching fractions and the use of RapidSim in  
 2099     lieu of full simulation is not included, but is of course significant. Nevertheless  
 2100     it is clear that the potential biases are significant compared to the size of the  
 2101     systematic uncertainties of the analysis presented in Section 5.6. Therefore the  
 2102     backgrounds are vetoed by requiring `IsMuon=0` and `PIDe < 0` on the pions from  
 2103     the  $D$ -decay with opposite charge to the bachelor in the  $D \rightarrow K_S^0 \pi^+ \pi^-$  channel.  
 2104     This requirement removes 88 % of the muonic background and 99 % of the electron  
 2105     background, according to PID efficiencies obtained via the `PIDCalib` package,  
 2106     using the  $(p, p_T)$  distribution for the muon/electron in the RapidSim samples. The  
 2107     survival rate for signal decays in full simulation is 94 %, so the impact on the  
 2108     obtainable precision is only about 3 %. A systematic uncertainty is assigned to  
 2109     account for the potential remaining background.

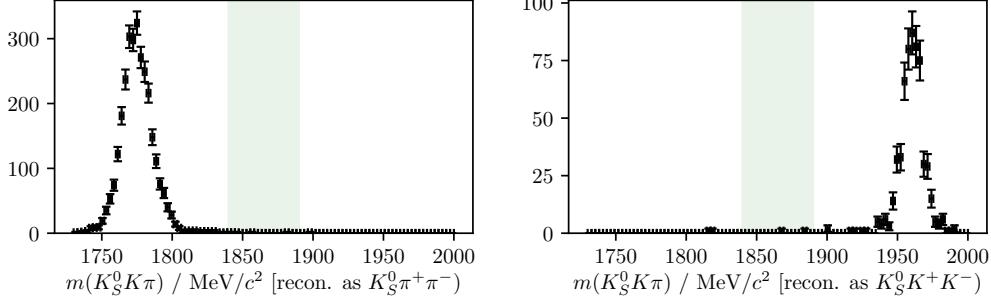
2110     In the  $D \rightarrow K_S^0 K^+ K^-$  channel an analogous study shows the relative yields  
 2111     to be similar. The selection efficiencies are higher, as are the relative branching  
 2112     ratios due to the lower  $D \rightarrow K_S^0 K^+ K^-$  branching fraction, but in this mode the  
 2113      $PIDK > -5$  requirement placed on the pion and lepton remove approximately 90 %  
 2114     of the background, leaving the relative rate similar to in  $D \rightarrow K_S^0 \pi^+ \pi^-$ . However,  
 2115     importantly, *the background is not peaking*, as shown in Fig. 5.21. The presence  
 2116     of a percent-level, *non-peaking* background in the  $D \rightarrow K_S^0 K^+ K^-$  channel is safe  
 2117     to ignore and thus no veto is applied in the  $D \rightarrow K_S^0 K^+ K^-$  channel.

2118     The muon-veto for the semi-leptonic background does remove some signal  
 2119     decays, where an original pion or kaon results in hits in the muon detectors. A



**Figure 5.21:** The reconstructed (left)  $m(K_S^0 K^+ K^-)$  and (right)  $m(Dh)$  distributions in RapidSim samples of  $B^\pm \rightarrow DK^\pm$  decays where  $D^0 \rightarrow K_S^0 \pi^- \mu^+ \nu_\mu$ , where the  $K_S^0 \pi^-$  originate in (top to bottom) the resonances  $K^{*-}(892)$ ,  $K^{*-}_0(1430)$ , and  $K^{*-}(1680)$ . The shapes for the  $D^0 \rightarrow K_S^0 \pi^- e^+ \nu_e$  case are almost identical.

2120 significant contribution is from particles that decay in flight. The track quality  
 2121 of these decays is worse than for nominal decays, which affects the resolution on  
 2122 the reconstructed Dalitz coordinates. In simulated signal decays the standard  
 2123 deviation of  $\Delta m_\pm^2 = m_{\text{reco}}^2(K_S^0 \pi^\pm) - m_{\text{TRUE}}^2(K_S^0 \pi^\pm)$  is 50 % larger for decays where  
 2124 one of the  $D$ -decay products has `IsMuon=1` than in decays where this is not the  
 2125 case. This can lead to systematic biases on the observables, as described further in  
 2126 Section 5.6.7. The overall effect is small, as evidenced by the systematic uncertainty  
 2127 described in that section; nevertheless this fact motivates removing decay-in-flight  
 2128 decays of the  $D$ -decay products. Therefore it is also required that `IsMuon=0` for  
 2129 the  $D$ -decay pion with the same charge as the companion in the  $D \rightarrow K_S^0 \pi^+ \pi^-$   
 2130 channels, and on the  $D$ -decay kaons in the  $D \rightarrow K_S^0 K^+ K^-$  channels. This veto  
 2131 removes about 2 % of signal candidates in simulation that survive the lepton vetoes  
 2132 described in the previous sections.



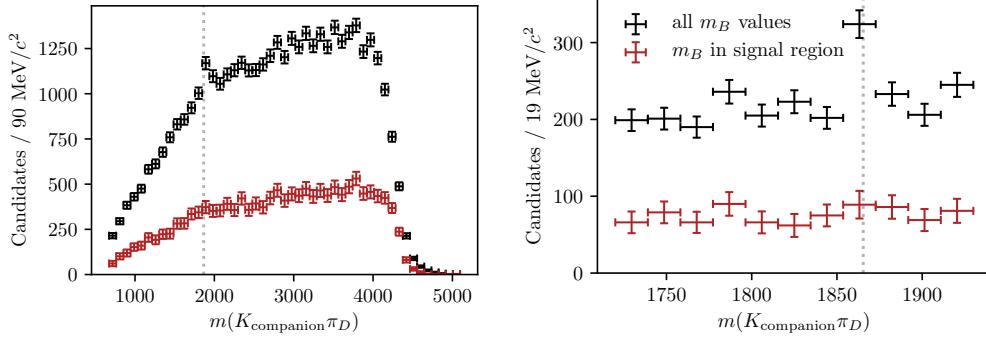
**Figure 5.22:** Simulated samples of  $B^\pm \rightarrow D(\rightarrow K_S^0 K\pi)\pi^\pm$  decays reconstructed in the (left)  $D \rightarrow K_S^0 \pi^+\pi^-$  and (right)  $D \rightarrow K_S^0 K^+K^-$  channels, combining the LL and DD categories. The  $D$ -mass region included in the selection of signal decays is illustrated with the green band. The plots in the  $B^\pm \rightarrow DK^\pm$  channels look almost identical.

### 2133 5.3.4 Cross-feed from other $D \rightarrow K_S^0 h^+h^-$ decays

2134 Misidentification of a  $D$  decay product can lead to background from cross-feed be-  
 2135 tween the  $B^\pm \rightarrow D(\rightarrow K_S^0 \pi^+\pi^-)h^\pm$  and  $B^\pm \rightarrow D(\rightarrow K_S^0 K^+K^-)h^\pm$  signal channels,  
 2136 or cross-feed from  $B^\pm \rightarrow D(\rightarrow K_S^0 K\pi)h^\pm$  decays into either of the signal channels.  
 2137 However, this background is very highly suppressed by the employed requirement  
 2138 on the  $D$  mass. This is illustrated in Fig. 5.22, where the  $D$  mass distribution in  
 2139 samples of simulated  $B^\pm \rightarrow D(\rightarrow K_S^0 K\pi)K^\pm$  decays are shown, when reconstructed  
 2140 as  $D \rightarrow K_S^0 \pi^+\pi^-$  and  $D \rightarrow K_S^0 K^+K^-$  decays. Essentially no decays that fall in  
 2141 the selected  $D$  mass window survive the full selection. Therefore this background  
 2142 is not considered further. Neither is the background due to cross-feed between  
 2143  $B^\pm \rightarrow D(\rightarrow K_S^0 \pi^+\pi^-)h^\pm$  and  $B^\pm \rightarrow D(\rightarrow K_S^0 K^+K^-)h^\pm$ , since it involves two  
 2144 misidentified particles, and therefore will result in reconstructed  $D$  masses even  
 2145 further away from the selected mass window. A very loose PID requirement on the  
 2146 charged  $D$  decay products is nonetheless included in the  $D \rightarrow K_S^0 K^+K^-$  channel,  
 2147 because it helps reduce the level of combinatorial background.

### 2148 5.3.5 Swapped-track backgrounds

2149 A possible peaking background stems from real  $B \rightarrow DhX$  decays with the same  
 2150 final state tracks as in the signal case, but where some tracks are mis-assigned in the  
 2151 reconstruction. Examples are  $B^\pm \rightarrow (K_S^0 h^+h^-)_D h^\pm$  decays where the companion  
 2152 and a  $D$ -decay product track are swapped, or  $B^\pm \rightarrow (K^- \pi^+)_D K_S^0 h^\pm$  decays, where  
 2153 the  $K_S^0$  is assigned to the  $D$  decay and the real companion is swapped with the  
 2154  $D$ -decay product of the same charge. The signature of this background type is  
 2155 a peak at the  $D$  mass, when the invariant mass corresponding to the companion



**Figure 5.23:** Invariant mass spectrum of the  $m^2(K^\pm\pi^\mp)$  combination in the  $B^\pm \rightarrow (K_S^0\pi^+\pi^-)K^\pm$  data sample for (black) all candidates and (red) candidates for which  $m_B \in m_B^{PDG} \pm 30 \text{ MeV}/c^2$ . The LL and DD categories are combined. The only difference between the left and right plots is the  $m(K\pi)$  mass range on the horizontal axis. The dotted line indicated the known  $D$  mass [?].

track and some subset of the  $D$ -decay tracks is formed. The presence of the background has been investigated by forming all such combinations, for all data categories, after the full selection has been applied. Only in a single channel is a peak visible: the  $B^\pm \rightarrow (K_S^0\pi^+\pi^-)K^\pm$  channel, where  $m(K^\pm\pi^\pm)$  has a peak, as shown in Fig. 5.23. Thus, a background is present from the favoured two-body  $D$  decay  $B^\pm \rightarrow (K^\pm\pi^\mp)_D K_S^0\pi^\pm$ , where the  $K^\mp$  is reconstructed as the companion, and the pions assigned to the  $D$  decay.

Is not favourable to veto this background, because a requirement on the invariant mass of a track combination that includes the companion track would impact the Dalitz-plot acceptance differently in the  $DK^\pm$  and  $D\pi^\pm$  channels. Thus it would break a fundamental underlying feature of the measurement: the identical selection efficiency profile between these modes. However, the yield excess in the  $m(K_{\text{companion}}^\pm\pi^\mp_D)$  range around  $m_D$ , attributed to the background, corresponds to only about 0.5% of the signal yield. A background at this level does not lead to a limiting systematic uncertainty on the measurement, as described in Section 5.6.9.

## 5.4 Signal and background mass shapes

The measurement employs *extended maximum-likelihood fits* [?] to the  $m(Dh^\pm)$  distribution of signal candidates to determine the observables of interest. The analysis implements a two-step fit procedure: first the data samples are analysed without separating the candidates by  $B$  charge or Dalitz bin, in order to determine appropriate parametrisations of the  $m(Dh^\pm)$  distribution of the signal and relevant background components. The parameterisations are then kept fixed in a subsequent

fit of the observables of interest, where the candidates are split by  $B$  charge and Dalitz bin. This section describes the first step, whereas the latter fit is the subject of Section 5.5.

In both steps, the candidates are split in 8 categories depending on whether the companion is categorised as a kaon or pion, whether the  $K_S^0$  meson is in the LL or DD category, and by whether the  $D$  meson is reconstructed in the  $K_S^0\pi^+\pi^-$  or  $K_S^0K^+K^-$  final state. In the remained of this text, these categories are indexed with the letter  $c$ . For each category,  $c$ , the expected number of observed decays at a given  $B$  mass,  $F^c(m)$ , is given by the sum of a signal contribution and a number of background distributions

$$F^c(m|\theta) = N_s^c(\theta)f_s^c(m|\theta) + \sum_b N_b^c(\theta)f_b^c(m|\theta), \quad (5.4)$$

where  $\theta$  denotes a set of parameters that describe the mass shapes and expected yields, in which some parameters are shared between categories. The distributions  $f_{s/b}^c$  are normalised to integrate to unity, and the expected signal (background) yields are denoted  $N_s^c$  ( $N_b^c$ ). A total, normalised distribution can then be defined

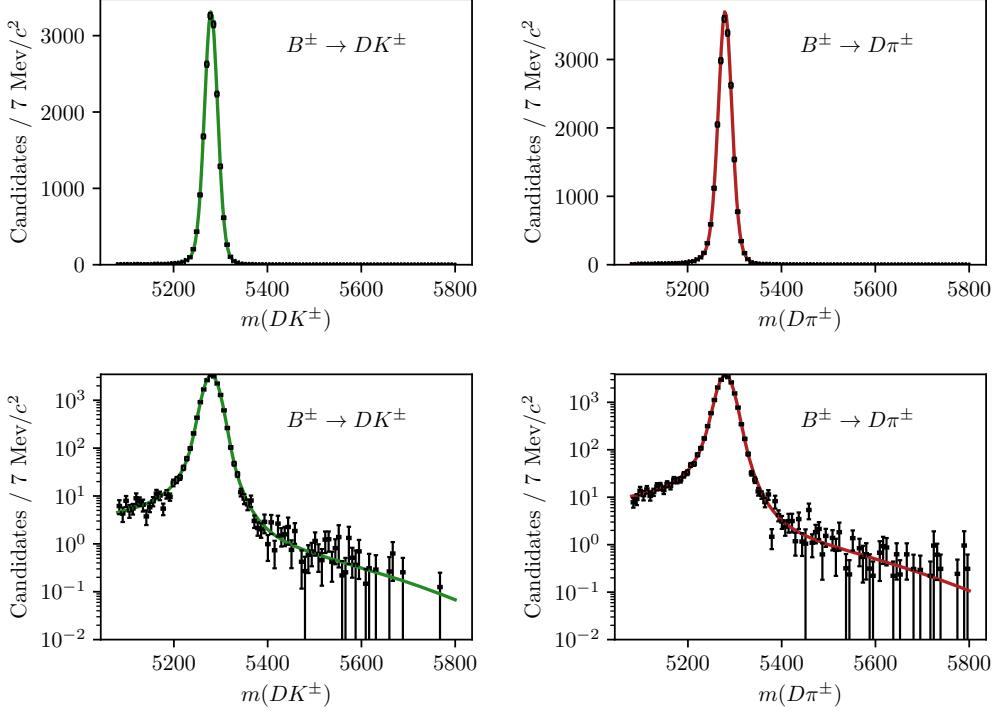
$$f^c(m|\theta) = \frac{1}{N_{\text{tot}}^c(\theta)}F^c(m|\theta), \quad N_{\text{tot}}^c(\theta) = N_s^c(\theta) + \sum_b N_b^c(\theta). \quad (5.5)$$

Given a set of  $N_{\text{observed}}^c$  measured  $B$  masses,  $\{m_i^c\}$ , in a given category, the extended log-likelihood function is defined

$$\ln \mathcal{L}_c(\theta|\{m_i^c\}) \equiv \sum_i \ln f^c(m_i^c|\theta) + \ln \text{Poisson}(N_{\text{tot}}^c(\theta), N_{\text{observed}}^c) \quad (5.6)$$

In a simultaneous fit the total, negative log-likelihood is  $-\ln \mathcal{L} = -\sum_c \mathcal{L}_c$ , and this function can be minimised to find the maximum-likelihood estimates of the parameters in  $\theta$ , as well as their confidence regions and correlation coefficients. This is handled with the `RooFit` package [?].

Apart from signal decays, the fit includes components that describe combinatorial background, backgrounds from decays where a companion pion is misidentified as a kaon or vice versa, and partially reconstructed backgrounds. Each of these components are described in detail in the following, before the results of the first-stage fit are presented in Section 5.4.5.



**Figure 5.24:** Fit projection of the signal shape to simulated  $B^\pm \rightarrow D(\rightarrow K_S^0 \pi^+ \pi^-)h^\pm$  samples reconstructed in the LL category. (Left) shows  $DK$  shapes, and (right) shows  $D\pi$  shapes. The shapes are shown with both linear and logarithmic  $y$ -axis scales.

### 5.4.1 Signal decays

The signal component is modelled with a sum of a Gaussian density function,  $f_G(m|m_B, \sigma)$ , and a modified Gaussian distribution with the parameterisation

$$f_C(m|m_B, \sigma, \alpha_L, \alpha_R, \beta) \propto \begin{cases} \exp \left[ \frac{-\Delta m^2(1+\beta\Delta m^2)}{2\sigma^2 + \alpha_L \Delta m^2} \right], & \Delta m = m - m_B < 0 \\ \exp \left[ \frac{-\Delta m^2(1+\beta\Delta m^2)}{2\sigma^2 + \alpha_R \Delta m^2} \right], & \Delta m = m - m_B > 0, \end{cases} \quad (5.7)$$

which is Gaussian when  $\Delta m^2 \ll \sigma^2/\alpha_{L/R}$  or  $\Delta m^2 \gg \beta^{-1}$  (with widths of  $\sigma$  and  $\sqrt{\alpha_{L/R}/\beta}$  respectively), with an exponential-like transition that is able to model the effect of the experimental resolution of LHCb very well. For the case  $\beta = 0$  the shape is denoted the *Cruijff* shape; however, in this case it tends to a uniform distribution for large  $\Delta m^2$  values, and cannot model the tails of the signal distribution. Thus, the full density function is

$$f_s(m|m_B, \sigma, \alpha_L, \alpha_R, \beta) = k_C f_C(m|m_B, \sigma, \alpha_L, \alpha_R, \beta) + (1 - k_C) f_G(m|m_B, \sigma). \quad (5.8)$$

The tail parameters  $(\alpha_{L/R}, \beta)$  and the constant  $k_C$  are determined in fits to simulated signal decays that have passed the full selection. The parameters are shared

2214 between the  $K_S^0\pi^+\pi^-$  and  $K_S^0K^+K^-$  channels, but otherwise independent in the  
 2215 fit categories. An example of a fit to simulated  $B^\pm \rightarrow D(\rightarrow K_S^0\pi^+\pi^-)h^\pm$  decays  
 2216 is given in Fig. 5.24. The resolution parameters  $\sigma$  are determined in the fit to  
 2217 actual data. Separate parameters are determined in the LL and DD categories,  
 2218 because the LL category has a better resolution on the  $K_S^0$  momentum, and  
 2219 therefore a narrow peak in reconstructed  $B$  mass. Likewise, separate resolution  
 2220 parameters are used for  $B^\pm \rightarrow D\pi^\pm$  and  $B^\pm \rightarrow DK^\pm$  decays, because the smaller  
 2221  $Q$  value in the latter case leads to smaller momenta of the decay products, and  
 2222 a correspondingly better resolution.

2223 The signal yields are determined independently in each  $B^\pm \rightarrow D\pi^\pm$  category.  
 2224 The yields in the  $B^\pm \rightarrow DK^\pm$  categories are then parameterised in terms of a single  
 2225 yield-ratio  $\mathcal{R}_{K/\pi}$ , and  $\epsilon^c$ , the corresponding selection efficiency for a given category

$$N_{DK^\pm}^c = \mathcal{R}_{K/\pi} \times N_{D\pi^\pm}^c \times \frac{\epsilon_{DK^\pm}^c}{\epsilon_{D\pi^\pm}^c}. \quad (5.9)$$

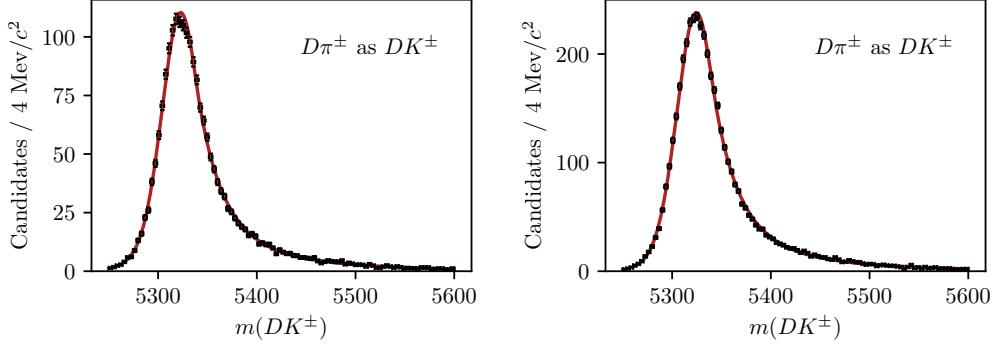
2226 The selection efficiency is obtained in simulation, except for the PID efficiencies  
 2227 which are obtained in calibration data as described in Section 5.1.3. The parameter  
 2228  $\mathcal{R}_{K/\pi}$  is shared between all categories, and corresponds to the branching ratio  
 2229 between  $B^\pm \rightarrow DK^\pm$  and  $B^\pm \rightarrow \square\pi^\pm$  decays. Therefore, it can be compared to  
 2230 the branching ratio ratio measured in dedication measurements, which serves as  
 2231 an important cross check of the efficiency determination. □

#### 2232 5.4.2 Cross-feed between $B^\pm \rightarrow Dh^\pm$ channels

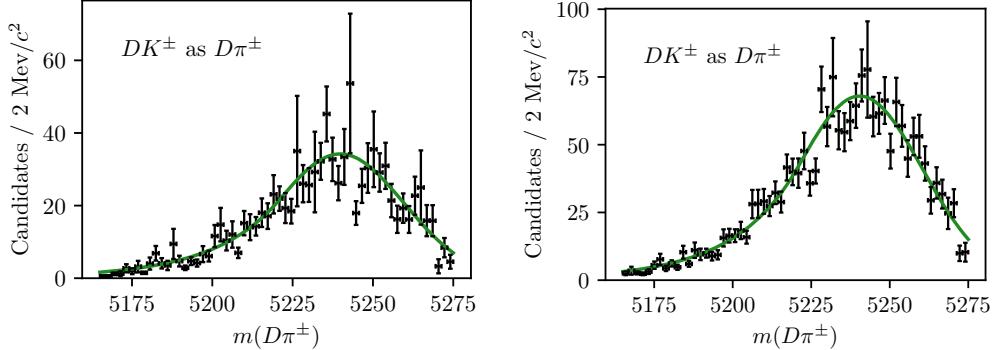
2233 There is a cross-feed between the  $B^\pm \rightarrow D\pi^\pm$  and  $B^\pm \rightarrow DK^\pm$  channels, where real  
 2234  $B^\pm \rightarrow D\pi^\pm$  decays are reconstructed as  $B^\pm \rightarrow DK^\pm$  decays, or where  $B^\pm \rightarrow DK^\pm$   
 2235 decays are reconstructed as  $B^\pm \rightarrow D\pi^\pm$  decays. Due to relative branching fractions  
 2236 the former contribution is by far the most important, but both are modelled.

2237 The cross-feed shapes are obtained in a data-driven manner using the sPlot  
 2238 method [?], and fixed in the fit to data. Separate shapes are determined for each  
 2239 category, using the following steps:

- 2240 • The procedure is based on the reasonably pure  $B^\pm \rightarrow D\pi^\pm$  sample obtained  
 2241 when the full selection is applied. A simple mass fit is performed to the  
 2242 invariant mass spectrum and the sPlot method [?] is used to obtain the  
 2243 sWeights,  $w_s$ , for the signal component. The mass fit uses the same components  
 2244 for signal, low mass shape, and combinatorial background as described in this  
 2245 section.



**Figure 5.25:** Fitted shape of the  $B^\pm$  invariant mass spectrum for  $B^\pm \rightarrow D\pi^\pm$  decays misidentified as  $B^\pm \rightarrow DK^\pm$  decays for (left) LL and (right) DD candidates in the  $D \rightarrow K_S^0 \pi^+ \pi^-$  mode.



**Figure 5.26:** Fitted shape of the  $B^\pm$  invariant mass spectrum for  $B^\pm \rightarrow DK^\pm$  decays misidentified as  $B^\pm \rightarrow D\pi^\pm$  decays for (left) LL and (right) DD candidates in the  $D \rightarrow K_S^0 \pi^+ \pi^-$  mode.

- 2246     • A set of weights are defined, based on the candidate-by-candidate PID  
2247     efficiencies obtained as described in Section 5.1.3:  
  
 2248       – The extracted PID efficiencies of the  $\text{PIDK} < 4$  cut  $\epsilon_{D\pi \rightarrow D\pi}(p, \eta, n_{\text{tracks}})$   
2249       are used to reverse-weight the  $B^\pm \rightarrow D\pi^\pm$  sample, in order to obtain the  
2250       bachelor kinematic distributions before the  $\text{PIDK} < 4$  cut is applied.  
  
 2251       – The extracted PID efficiencies of the  $\text{PIDK} > 4$  cut  $\epsilon_{D\pi \rightarrow DK}(p, \eta, n_{\text{tracks}})$   
2252       are used to obtain the bachelor kinematic distribution of the  $B^\pm \rightarrow D\pi^\pm$   
2253       candidates mis-identified as  $B^\pm \rightarrow DK^\pm$ .  
  
 2254     • The raw distribution of the invariant mass of  $B^\pm$  particles with a misidentified  
2255       bachelor,  $m_B^{mis-ID}$ , is produced by also doing the DecayTreeFit kinematic refit  
2256       while swapping the companion mass hypothesis of each  $B^\pm \rightarrow D\pi^\pm$  candidate  
2257       to a kaon hypothesis.

- 2258     • Each candidate is reweighted by the overall weight  $w = w_s^{cand.}/\epsilon_{D\pi \rightarrow D\pi}^{cand.} \cdot$   
 2259        $\epsilon_{D\pi \rightarrow DK}^{cand.}$ , and the reweighed  $m_B^{mis-ID}$  distribution is fitted to obtain the cross-  
 2260       feed mass distribution function.

2261   The distributions are modelled with a sum of two Crystal Ball density functions,  
 2262   each defined by the parameterisation [?]

$$f_{CB}(m, \mu, \sigma, \alpha, n) \propto \begin{cases} \exp \left[ -\frac{1}{2} \left( \frac{m-\mu}{\sigma} \right)^2 \right] & \text{if } (m-\mu)/\sigma > -\alpha \\ A \left( B - \frac{m-\mu}{\sigma} \right)^{-n} & \text{otherwise,} \end{cases} \quad (5.10)$$

2263   where  $\alpha > 0$ , and

$$A = \left( \frac{n}{\alpha} \right)^n \exp[-\alpha^2/2], \quad B = \frac{n}{\alpha} - \alpha. \quad (5.11)$$

2264   The obtained  $m_B^{mis-ID}$  spectrum and obtained mass shape is given in Fig. 5.25  
 2265   for the  $D \rightarrow K_S^0 \pi^+ \pi^-$  category; the  $D \rightarrow K_S^0 K^+ K^-$  shapes are very similar. An  
 2266   analogous procedure is used to obtain the mass distribution of  $B^\pm \rightarrow DK^\pm$  decays  
 2267   reconstructed in the  $B^\pm \rightarrow D\pi^\pm$  category. In the first stage where sPlots are  
 2268   extracted by a fit to the  $B^\pm \rightarrow DK^\pm$  mass spectrum, the cross-feed component  
 2269   determined as described above is included. An example of one of the resulting  
 2270   shapes is given in Fig. 5.26

2271   The yield of cross-feed from  $B^\pm \rightarrow D\pi^\pm$  decays in a given  $B^\pm \rightarrow DK^\pm$  category  
 2272   is parameterised in terms of the yield of correctly identified  $B^\pm \rightarrow D\pi^\pm$  decays and  
 2273   the mis-identification probability extracted from calibration samples as described  
 2274   in Section 5.1.3. Denoting the rate at which a pion is reconstructed as a kaon  
 2275   by  $\epsilon_{\pi \rightarrow K}^c$  in a given category,  $c$ , the yield is

$$N_{\pi \rightarrow K}^c = N_{D\pi^\pm}^c \frac{\epsilon_{\pi \rightarrow K}^c}{1 - \epsilon_{\pi \rightarrow K}^c}, \quad (5.12)$$

2276   with an analogous definition of the yield of the cross-feed component from  
 2277    $B^\pm \rightarrow DK^\pm$  decays in the  $B^\pm \rightarrow D\pi^\pm$  spectrum.

#### 2278   5.4.3 Partially reconstructed backgrounds

2279   A number of background candidates stem from partly reconstructed  $B$  decays of the  
 2280   type  $B \rightarrow DhX$ , where  $X$  denotes a photon or a pion that is not reconstructed. It  
 2281   is not possible to reject the X decays in the selection, due to the similarity to signal  
 2282   decays. However, the missing X momentum results in reconstructed  $B$  masses below  
 2283   the actual  $B$  mass, and the backgrounds are also denoted *lowmass* backgrounds.  
 2284   These mass distributions are modelled with analytic shapes, derived based on two

principles. Firstly, the kinematic endpoints of the distributions are fully defined by the particle masses in the decay. Secondly, the angular distribution of the missing particle has a one-to-one relation to the missing momentum, and therefore to the reconstructed  $B$  mass. Depending on the spin-parity of the particles and resonances involved in the decay, two different mass distributions arise.

In  $B$  decays where the missing particle is a scalar that is produced in the decay of a vector resonance (eg.  $B^\pm \rightarrow D^{*0}(\rightarrow D^0\pi^0)\pi^\pm$  decays where the  $\pi^0$  is not reconstructed), the  $m(D^0\pi^\pm)$  distribution has a double-peak structure. The  $D^{*0}$  helicity angle  $\theta$  is defined as the angle between the  $\pi^0$  momentum vector in the  $D^{*0}$  rest frame and the  $D^{*0}$  boost vector in the  $B$  rest frame. The helicity of  $D^{*0}$  means that the  $\pi^0$  will travel predominantly in the direction where  $\theta = 0$  or  $\theta = \pi$ . When  $\theta = 0$  the fraction of momentum carried by the missing  $\pi^0$  is lower, leading to a higher reconstructed  $m(D^0\pi^\pm)$ . When  $\theta = \pi$  the converse occurs. The resulting  $B$  mass distribution is a parabola  $f_{\text{HORNS}}^0(m)$  peaking near both kinematic endpoints  $a$  and  $b$

$$f_{\text{HORNS}}^0(m) = \begin{cases} (m - \frac{a+b}{2})^2, & \text{if } a < m < b \\ 0, & \text{otherwise.} \end{cases} \quad (5.13)$$

Due to the double-peaking structure, and the fact that was developed by Paolo Gandini for the two-body ADS/GLW analyses [?], this shape is denoted a *HORNSdini* shape when convolved with a resolution function as described below.

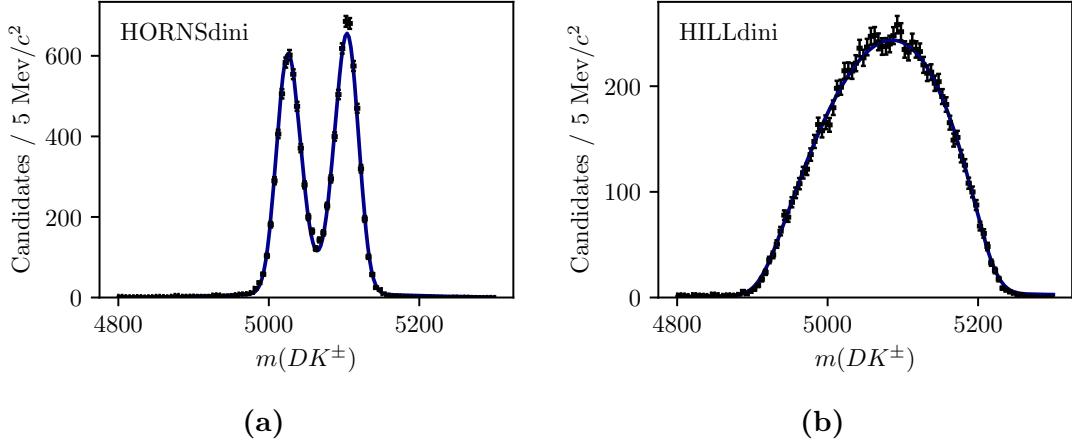
The second relevant decay situation is where the missing particle is a vector, again produced via the intermediate decay of a vector resonance (eg.  $B^\pm \rightarrow D^{*0}(\rightarrow D^0\gamma)\pi^\pm$  decays where the photon is not reconstructed). In this case, the spin-parity of the photon ( $1^-$ ) means that it will decay preferentially in the  $\theta = \frac{\pi}{2}$  or  $\theta = \frac{3\pi}{2}$  directions, and so a double-peak structure is not seen. In this case the parabolic distribution  $f_{\text{HILL}}^0(m)$  with kinematic endpoints  $a, b$  has negative curvature and can be described by

$$f_{\text{HILL}}^0(m) = \begin{cases} -(m - a)(m - b), & \text{if } a < m < b \\ 0, & \text{otherwise.} \end{cases} \quad (5.14)$$

Resolution effects mean that the parabolic shapes must be convolved with a resolution function, chosen to be a sum of two Gaussians. For a single Gaussian shape  $f_G(x|\mu, \sigma)$  with mean  $\mu$  and width  $\sigma$ , the double Gaussian is expressed as

$$f_{DG}(x) = f_G(x|\mu, \sigma)k_G f_G(x|\mu, R_\sigma\sigma). \quad (5.15)$$

where  $\sigma$  is the width of the first Gaussian, and  $k_G$  is the relative fractions between the two Gaussians and  $R_\sigma$  is their relative widths. Further, selection effects can



**Figure 5.27:** Examples of (a) the *HORNsdini* distribution fit to simulated  $B^\pm \rightarrow (D^{*0} \rightarrow D^0[\pi^0])K^\pm$  decays, and (b) the *HILLdini* distribution fit to simulated  $B^\pm \rightarrow (D^{*0} \rightarrow D^0[\gamma])K^\pm$  decays. The fits in this figure are made to illustrate the features of each shape, but do not enter the actual fit to data.

distort the horns shape such that one of the peaks is higher than the other. This is taken into account by introducing a linear polynomial with slope parameter  $\xi$ . As  $\xi \rightarrow 0$ , the left hand peak decreases in size relative to the right hand peak. The resulting *HORNsdini* and *HILLdini* distributions are therefore

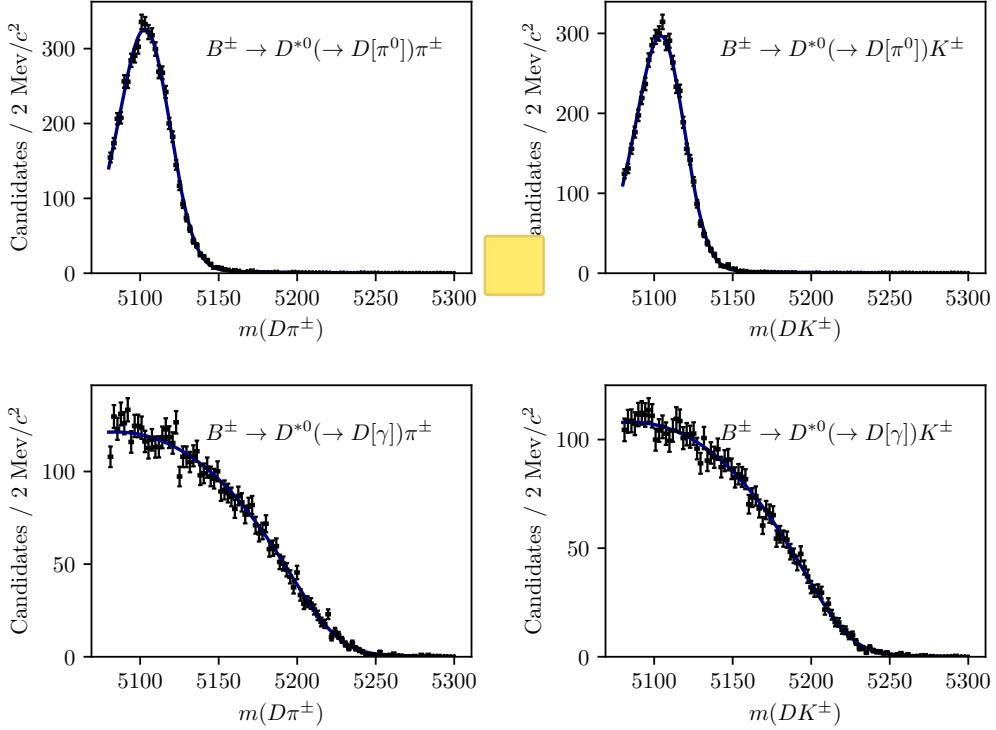
$$f_{\text{HORNs/HILL}}(m) = \int_a^b dx f_{\text{HORNs/HILL}}^0(x) f_{DG}(m|x, \sigma, k_G, R_\sigma) \left( \frac{1-\xi}{b-a}x + \frac{b\xi - a}{b-a} \right). \quad (5.16)$$

Examples of the shapes are given in Fig. 5.27. These shapes are used to fit all partially reconstructed backgrounds, as described in the following section.

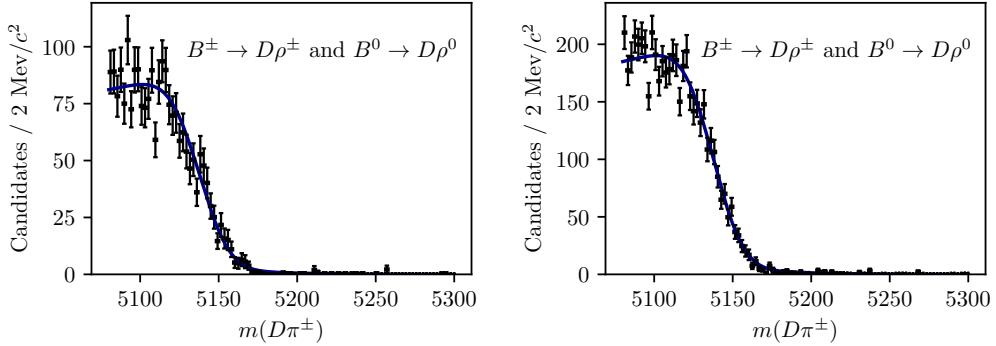
## 2320 Determination of the partially reconstructed background distributions

2321 In both the  $B^\pm \rightarrow DK^\pm$  and  $B^\pm \rightarrow D\pi^\pm$  categories, components are included to  
2322 describe contributions from the partially reconstructed decays (where the particle  
2323 in square brackets is not reconstructed)

- 2324 •  $B^\pm \rightarrow (D^{*0} \rightarrow D^0[\pi^0])h^\pm$ , described using a *HORNsdini* distribution
- 2325 •  $B^\pm \rightarrow (D^{*0} \rightarrow D^0[\gamma])h^\pm$ : described using a *HILLdini* distribution
- 2326 •  $B^0 \rightarrow (D^{*\pm} \rightarrow D^0[\pi^\pm])h^\mp$ : described using a *HORNsdini* distribution
- 2327 •  $B^{\pm(0)} \rightarrow D^0 h^\pm [\pi^{0(\mp)}]$ : described using a *HORNsdini* distribution



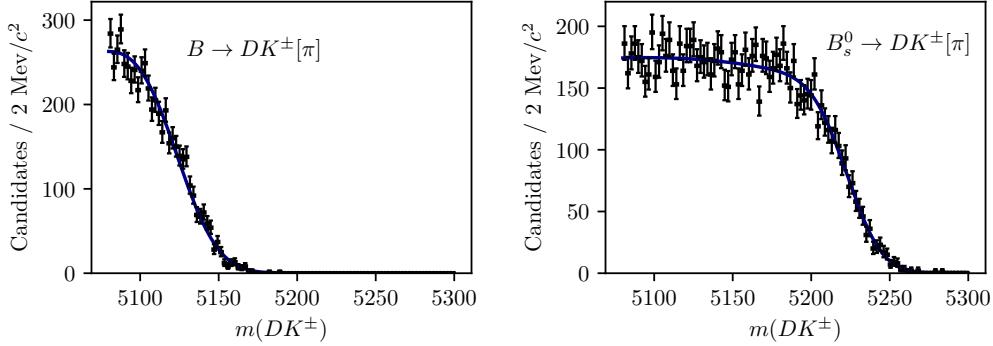
**Figure 5.28:** Fit projection of the fit to (top) simulated  $B^+ \rightarrow D^{*0}(\rightarrow D^0[\pi^0])h^\pm$  decays and (bottom) simulated  $B^+ \rightarrow D^{*0}(\rightarrow D^0[\gamma])h^\pm$  decays, all reconstructed in the DD category. Both the (left)  $DK$  and (right)  $D\pi$  shapes are shown.



**Figure 5.29:** Projections of the fit to simulated  $B^\pm \rightarrow D\rho^\pm$  and  $B^0 \rightarrow D\rho^0$  samples reconstructed as  $B^\pm \rightarrow D\pi^\pm$  decays for the (left) LL and (right) DD categories.

2328 The mass distributions of all the  $B \rightarrow D^*h^\pm$  contributions are obtained from  
 2329 fits to samples of full LHCb simulation. Examples of these fits are shown in  
 2330 Fig. 5.28. All shape parameters are kept fixed in the fit to data, except for the  
 2331 parameter  $\sigma$  of the resolution function in Eq. (5.15) which is allowed to obtain  
 2332 the value preferred by data.

2333 The mass distribution of  $B^{\pm(0)} \rightarrow D^0 h^\pm [\pi^{0(\mp)}]$  decays reconstructed in the



**Figure 5.30:** Fit projection for the fit used to obtain a shape for the partly reconstructed background from (left)  $B \rightarrow DK\pi$  decays and (right)  $B_s^0 \rightarrow DK^\pm\pi^-$  decays where a pion is not reconstructed.

2334  $B^\pm \rightarrow D\pi^\pm$  categories is obtained from full LHCb simulation samples of  $B^\pm \rightarrow D^0\rho^\pm$   
 2335 and  $B^0 \rightarrow D^0\rho^0$  decays. The shapes were compared to those predicted by an  
 2336 amplitude model for  $B^0 \rightarrow D^0 \rightarrow \pi^\pm\pi^\mp$  decays developed by LHCb [?], but found  
 2337 to be very similar for the  $m(D\pi^\pm)$  range relevant to this analysis. The obtained  
 2338 shapes are shown in Fig. 5.29.

2339 The mass distribution of  $B^{\pm(0)} \rightarrow D^0 K^\pm[\pi^{0(\mp)}]$  decays reconstructed in the  
 2340  $B^\pm \rightarrow DK^\pm$  categories, on the other hand, is obtained from a sample of signal  
 2341 decays, generated via █ amplitude model for  $B^0 \rightarrow D^0 \rightarrow K^\pm\pi^\mp$  decays developed  
 2342 by LHCb [?] and smeared to take the LHCb resolution into account. This follows  
 2343 an approach developed in the context of a GLW analysis based on partially recon-  
 2344 structed decays made within LHCb [?]. The obtained shape is shown in Fig. 5.30.

2345 The background yields of these backgrounds are parameterised in terms of  
 2346 one total yield parameter, accounting for all partially reconstructed  $B^\pm$  and  
 2347  $B^0$  decays, and a number of parameters that describe the relative rates of the  
 2348 different contributions. In the  $B^\pm \rightarrow D\pi^\pm$  channels, the relative rates of the  
 2349  $B^\pm \rightarrow (D^{*0} \rightarrow D^0[\pi^0])h^\pm$  and  $B^0 \rightarrow (D^{*\pm} \rightarrow D^0[\pi^\pm])h^\mp$  backgrounds are fixed  
 2350 from the known branching fractions, and relative selection efficiencies in simulation.  
 2351 These backgrounds have almost identical mass distributions and it is not possible to  
 2352 determine the ratio in the fit to data. The relative yield of  $B^\pm \rightarrow D^*(\rightarrow D^0[\gamma])\pi^\pm$   
 2353 compared to the  $B \rightarrow D^*(\rightarrow D^0[\pi])\pi^\pm$  is denoted  $f_{D^*\gamma}^{D\pi}$  and is floated in the fit to  
 2354 data, as is the relative yield of  $B \rightarrow D^0\pi^\pm[\pi]$  decays compared to the  $B \rightarrow D^*\pi$   
 2355 modes, denoted  $f_{D\pi\pi}^{D\pi}$ . In the  $B^\pm \rightarrow DK^\pm$  channels, all the relative background  
 2356 rates are fixed via known branching fractions and relative selection efficiencies; this  
 2357 is necessary to obtain a stable fit, due to the lower yields.

In the  $B^\pm \rightarrow DK^\pm$  categories, an additional partially reconstructed background is considered from  $B_s^0 \rightarrow \bar{D}^0[\pi^+]K^-$  ( $\pi^-$  conjugate) decays. The mass shape is obtained from simulated decays, generated using an amplitude model published by LHCb [?] and smeared to account for the experimental resolution. The obtained shape is shown in Fig. 5.30. The yield of this background component is fixed relative to the signal yields in the corresponding  $B^\pm \rightarrow D\pi^\pm$  category, taking the relative branching ratios and hadronisation factors into account.

In the  $B^\pm \rightarrow DK^\pm$  channels there is a contribution from partially reconstructed  $B \rightarrow D^*\pi^\pm X$  decays where the companion pion is misidentified as a kaon. The reverse contribution is negligible due to the relative branching fractions, and the fact that it is mostly shifted below the mass range of the fit. These are modelled using analytic, empirical mass distributions (essentially sums of a number of regular *HORN/HILLdini* distributions), with parameters that are determined in fits to simulated  $B \rightarrow D^*\pi^\pm$  and  $B \rightarrow D\rho$  decays where the pion is reconstructed with the kaon mass hypothesis. The shapes are fixed in the fit to data.

2373

#### 2374 Partially reconstructed backgrounds that are not modelled

2375 It was considered whether a background from  $\Lambda_b^0 \rightarrow D^0p\pi^-$  decays where a pion  
 2376 is not reconstructed, and the proton is misidentified as the companion, can be  
 2377 expected to contribute significantly. This has been investigated in full LHCb  
 2378 simulation samples, for the  $D$  final state  $K_S^0\pi^+\pi^-$ . Taking into account the selection  
 2379 efficiencies, branching fractions, and hadronisation fraction of this background, the  
 2380 expected relative yield of the  $\Lambda_b^0$  background compared to signal of 0.03 % in the  
 2381  $B^\pm \rightarrow D\pi^\pm$  channel, which is completely negligible. In the  $B^\pm \rightarrow DK^\pm$  channel,  
 2382 total of about 200 decays are expected, combined for the LL and DD categories.  
 2383 However, most of these lie at  $B$  masses smaller than the signal peak, and their impact  
 2384 is small. Therefore it is not necessary to model the background in the nominal fit;  
 2385 a systematic uncertainty is assigned that accounts for the small potential impact.

2386 In the analogous case of  $\Lambda_b^0 \rightarrow D^0pK^-$  decays, the missing energy of the non-  
 2387 reconstructed kaon results in a reconstructed  $B$  mass below the fit range.

2388 It has also been investigated whether a background from  $\Lambda_b^0 \rightarrow \Lambda_c^+\pi^-$  or  $\Lambda_b^0 \rightarrow$   
 2389  $\Lambda_c^+K^-$  decays can be expected, where  $\Lambda_c^+ \rightarrow pK_S^0\pi^+\pi^-$ , a pion is missed and the  
 2390 proton is misidentified as a pion or kaon from the  $D$  decay. Since no PID requirement  
 2391 is made on the  $D$  decay products in the  $D \rightarrow K_S^0\pi^+\pi^-$  channels, and only a loose  
 2392 one for  $D \rightarrow K_S^0K^+K^-$ , this is possible in principle. In practice, the background is  
 2393 sufficiently suppressed from the applied  $D$  mass requirement to have no significant  
 2394 impact, and is therefore not modelled. A systematic uncertainty is assigned that  
 2395 accounts for any potential impact on the measurement due to this choice.

2396 **5.4.4 Combinatorial background**

2397 The combinatorial background is modelled with an exponentially falling density  
2398 function, where both the yield and exponential slope are determined independently  
2399 for each category. This shape is found to model the combinatorial well in all  
2400 categories, most evident in the high- $m_B$  regions where this background dominates.

2401 **5.4.5 Fit results**

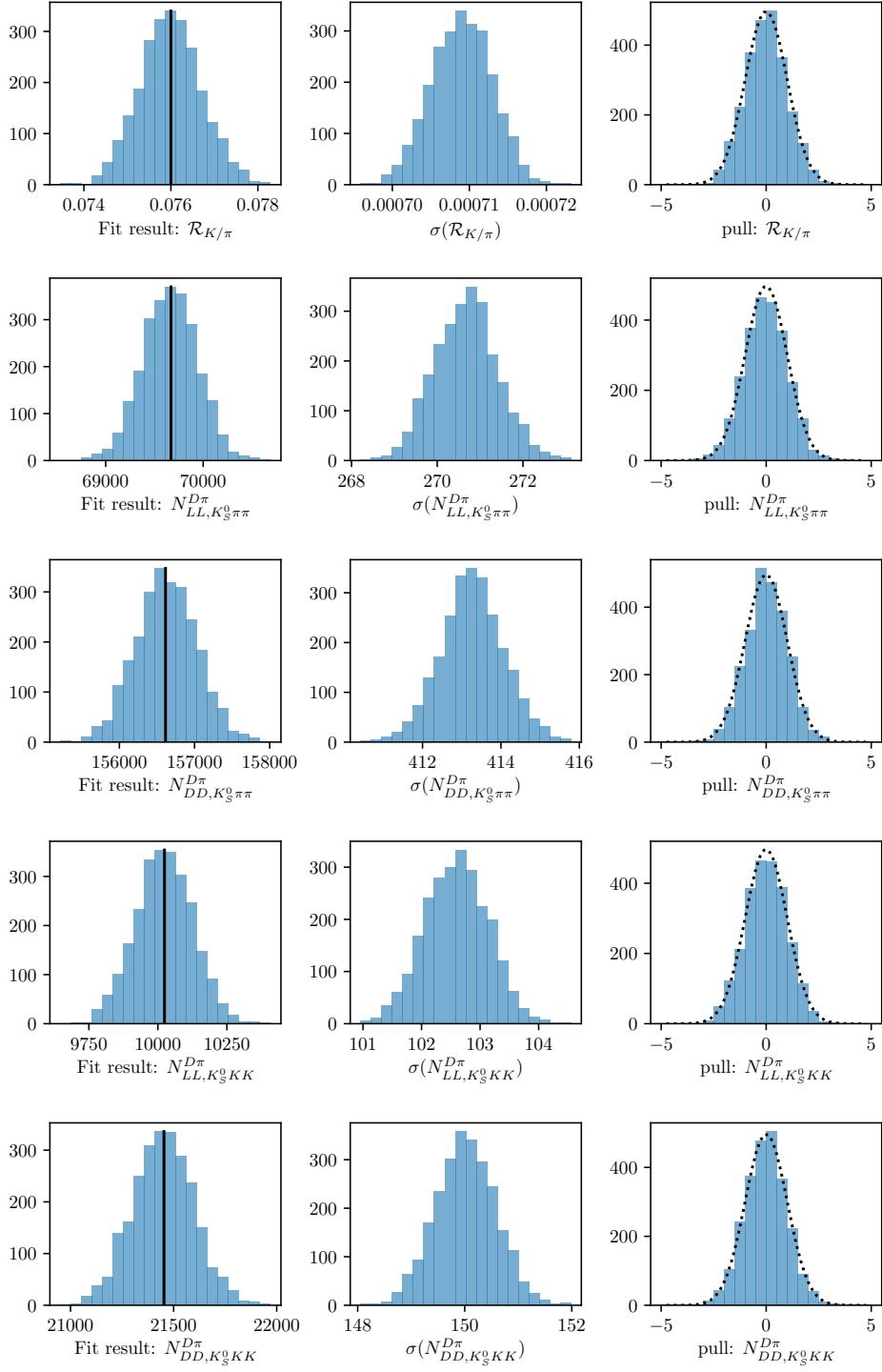
2402 The fit range is chosen to be  $m_B \in [5080, 5800] \text{ MeV}/c^2$ . The low end of this  
2403 interval includes the high-mass peak of the double-peak structure in the partially  
2404 reconstructed background, which helps the fit constrain the relative contributions of  
2405 backgrounds in the lowmass region  A number of additional backgrounds exist at  
2406 even lower  $m_B$  values, thus extending the fit range to lower masses would necessitates  
2407 an extended model, but not benefit the description of the signal region. The high  
2408 end of the interval includes enough combinatorial background to allow the fit to  
2409 determine the exponential slope parameter accurately.

2410 A large number of pseudoexperiments are carried out to verify that the fit  
2411 procedure is self-consistent, in which toy data sets are generated according to  
2412 the expected  $B$  mass distributions, and then fitted. None of the parameters  
2413 obtained in the fit exhibit a mean bias different from zero. For most parameters  
2414 the uncertainties are well estimated. This is the case for the signal yields, and  
2415 the  $DK^\pm - D\pi^\pm$  yield ratio  $\mathcal{R}$ , as evidenced by the pull plots in Fig. 5.31. The  
2416 fit underestimates the uncertainty by 10-20 % for some of the parameters related  
2417 to the partly reconstructed backgrounds, as shown in Fig. 5.32, but this is taken  
2418 into account when the uncertainties are propagated to the observables in the  
2419 second-stage fit, as described in Section 5.6.3.

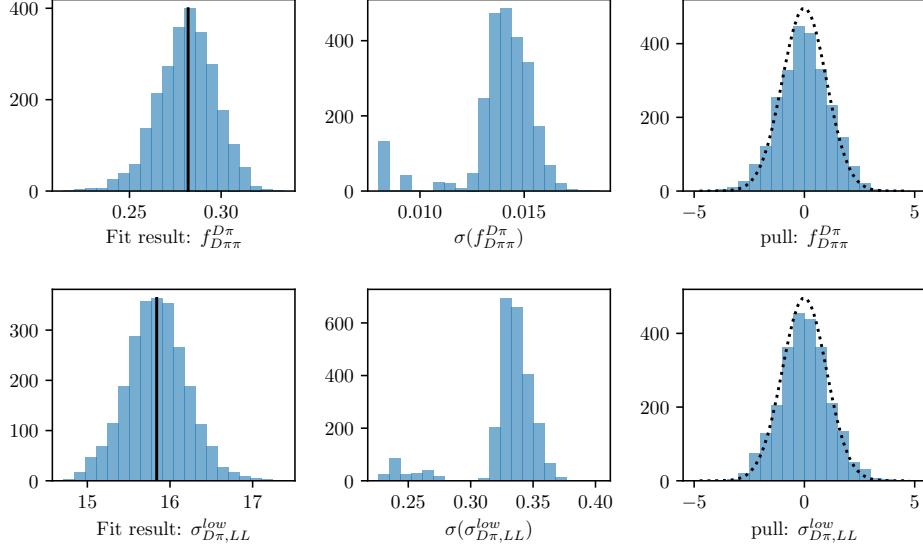
2420 The projections of the fit to data are shown in Figs. 5.33 and 5.34, for the  
2421  $D \rightarrow K_S^0 \pi^+ \pi^-$  and  $D \rightarrow K_S^0 K^+ K^-$  data sets, respectively. The obtained yields for  
2422 each fit component are given in Table 5.6. The total yield of  $B^\pm \rightarrow D\pi^\pm$  decays is  
2423 approximately 230,000 across all channels. The obtained value of the yield ratio is  
2424  $\mathcal{R}_{K/\pi} = (7.7 \pm 0.1) \%$ , corresponding to a total  $B^\pm \rightarrow DK^\pm$  yield of 16,500, of which  
2425 about 14,300 pass the PID requirement and are reconstructed in the  $B^\pm \rightarrow DK^\pm$   
2426 category. This value of  $\mathcal{R}$  is in excellent agreement with expectation from the  
2427 known branching fractions [?], which predict  $\mathcal{R}_{K/\pi}^{\text{PDG}} = (7.8 \pm 0.3) \%$ .<sup>2</sup> The shape  
2428 parameters determined in the fit to data are summarised in Table 5.7.

---

2While it would seem this measurement thus determines the yield ratio  $\mathcal{R}_{K/\pi}$  with a much better precision than the current world average uncertainty, that is because the result quoted here does not include any systematic uncertainties; it is only included to serve as a, successfully passed, cross check.



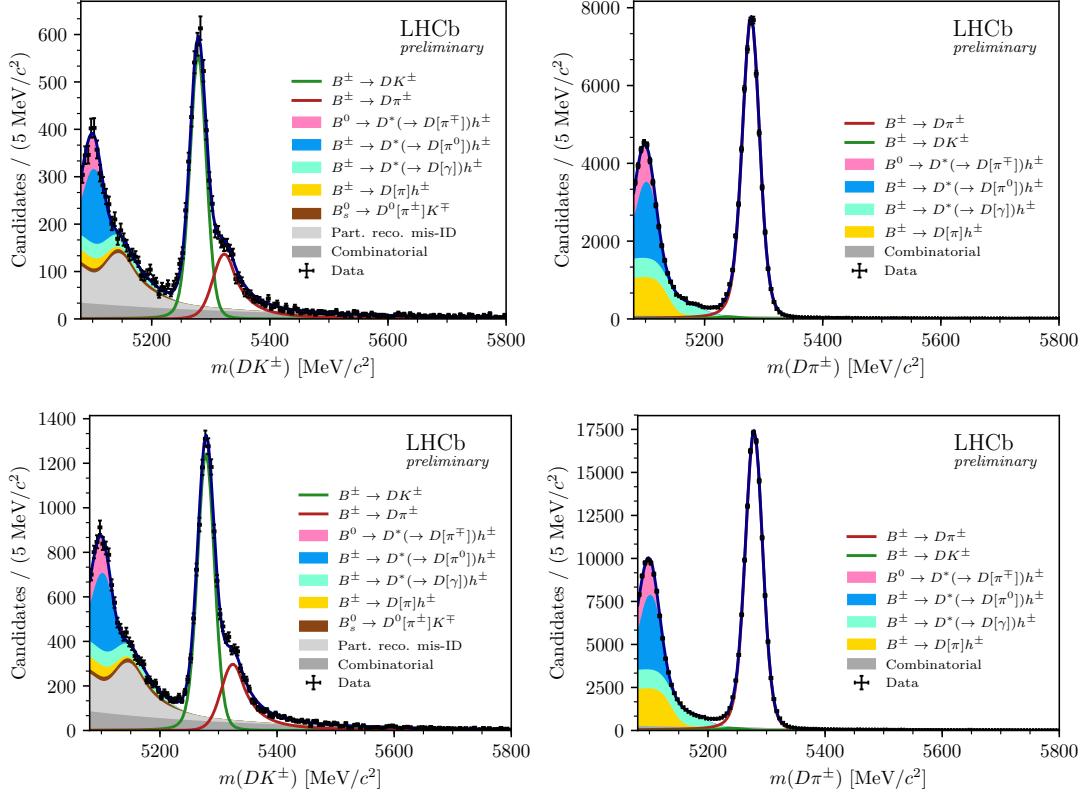
**Figure 5.31:** The (left) fitted value, (centre) estimated statistical uncertainty, and (right) pull plots for the signal yield parameters, as obtained in a number of pseudo experiments. The black line on the left shows the value used to generate the pseudo data sets; the dotted line on the right shows a Gaussian distribution with mean equal to zero and a standard deviation equal to unity.



**Figure 5.32:** The (left) fitted value, (centre) estimated statistical uncertainty, and (right) pulls obtained in a number of pseudo experiments for two examples of parameters relating to the partially reconstructed backgrounds, where the uncertainties are slightly underestimated on average. The standard deviation of the pull distributions is approximately 1.15 in both cases.

**Table 5.6:** Fitted total candidate yields. The quoted signal yields are for the number of candidates that survive the respective PID cut, whereas the  $DK^\pm$ - $D\pi^\pm$  ratio is corrected for PID and selection efficiencies so that it corresponds to the branching ratio.

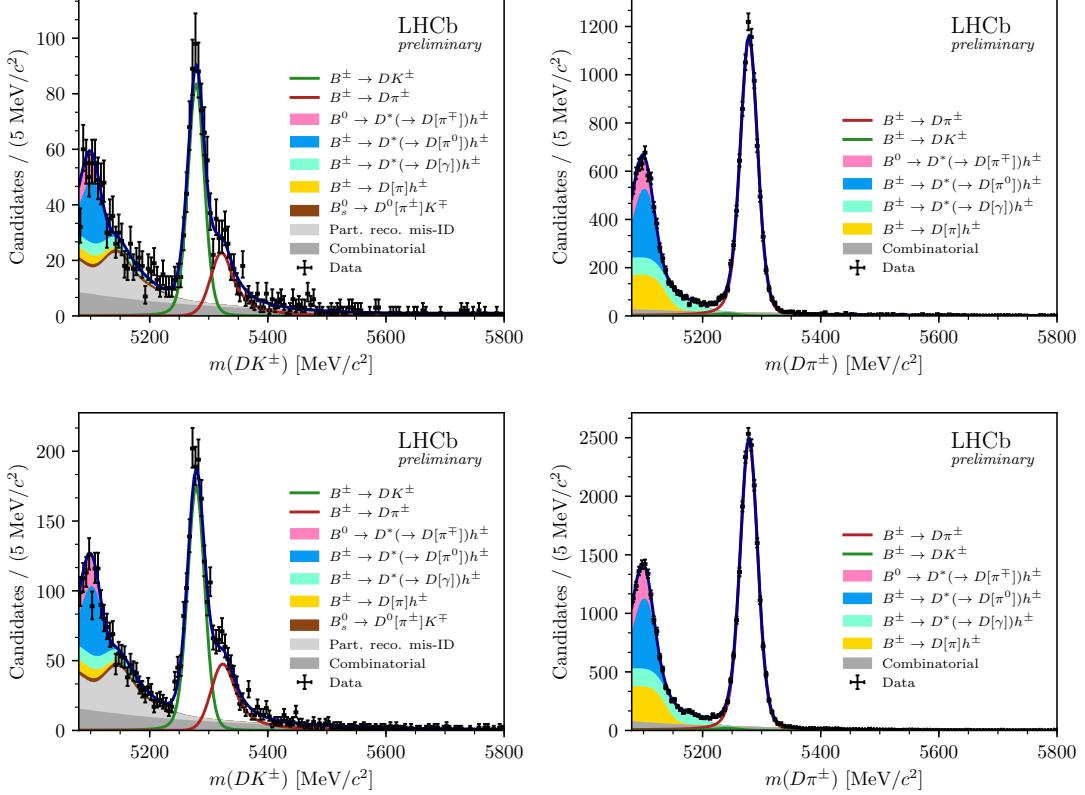
Component	LL	DD
Signal		
$B^\pm \rightarrow D(\rightarrow K_S^0\pi^+\pi^-)\pi^\pm$	$61,573 \pm 254$	$139,080 \pm 389$
$B^\pm \rightarrow D(\rightarrow K_S^0K^+K^-)\pi^\pm$	$9,160 \pm 98$	$19,910 \pm 144$
$R_{K/\pi} = n(DK)/n(D\pi)$ (%)	$7.72 \pm 0.08$	
Combinatorial		
$B^\pm \rightarrow D(\rightarrow K_S^0\pi^+\pi^-)\pi^\pm$	$3,479 \pm 198$	$9,928 \pm 376$
$B^\pm \rightarrow D(\rightarrow K_S^0K^+K^-)\pi^\pm$	$1,103 \pm 94$	$2,545 \pm 155$
$B^\pm \rightarrow D(\rightarrow K_S^0\pi^+\pi^-)K^\pm$	$1,826 \pm 107$	$3,987 \pm 177$
$B^\pm \rightarrow D(\rightarrow K_S^0K^+K^-)K^\pm$	$380 \pm 39$	$655 \pm 58$
Part. Reco.		
$B^\pm \rightarrow D(\rightarrow K_S^0\pi^+\pi^-)\pi^\pm$	$43,004 \pm 242$	$95,452 \pm 403$
$B^\pm \rightarrow D(\rightarrow K_S^0K^+K^-)\pi^\pm$	$6,247 \pm 99$	$13,241 \pm 157$
$R_{K/\pi}^{low} = n_{low}(DK)/n_{low}(D\pi)$ (%)	$6.65 \pm 0.12$	



**Figure 5.33:** The invariant mass distribution for the (left)  $B^\pm \rightarrow DK^\pm$  channel and (right)  $B^\pm \rightarrow D\pi^\pm$  channel, where  $D \rightarrow K_S^0\pi^+\pi^-$  and the  $K_S^0$  is in the (top) LL and (bottom) the DD categories. The particle within square brackets in the legend denotes the particle that has not been reconstructed.

## 5.5 Measurement of the CP-violation observables

The section describes the second fit stage, in which the  $CP$ -violation observables of interest are determined. Compared to the first fit stage, the candidates are further split by  $B$  charge, and by the assigned Dalitz bin number, making for a total of 160 categories. Another extended maximum-likelihood fit is carried out, in which shape parameters of all signal and background components are fixed to those determined in the first fit stage, and the floating parameters relate to the signal and background yields. The signal yields are expressed in terms of the observables of interest,  $(x_\pm^{DK}, y_\pm^{DK}, x_\xi^{D\pi}, y_\xi^{D\pi})$ , allowing the fit to determine their optimal values. The details of the fit setup are summarised in the following section, along with a number of studies that lead to the specific setup being chosen. The results are presented in Section 5.5.2, and a wide range of consistency checks are described in Section 5.5.3.



**Figure 5.34:** The invariant mass distribution for the (left)  $B^\pm \rightarrow DK^\pm$  channel and (right)  $B^\pm \rightarrow D\pi^\pm$  channel, where  $D \rightarrow K_S^0 K^+ K^-$  and the  $K_S^0$  is in the (top) LL and (bottom) the DD categories. The particle within square brackets in the legend denotes the particle that has not been reconstructed.

### 5.5.1 Fit setup

The basic principle of the measurement is that the signal yields in each bin (in a given category) are defined using the equations of Chapter 2, in order to allow for the determination of the  $CP$ -violation observables. In practice, a set of variables are defined

$$Y_{c,i}^- = F_{c,-i} + [(x_-^c)^2 + (y_-^c)^2]F_{c,-i} + 2\sqrt{F_{c,i}F_{c,-i}}(c_i^cx_-^c + s_i^cy_-^c), \quad (5.17)$$

$$Y_{c,i}^+ = F_{c,-i} + [(x_+^c)^2 + (y_+^c)^2]F_{c,-i} + 2\sqrt{F_{c,i}F_{c,-i}}(c_i^cx_+^c - s_i^cy_+^c), \quad (5.18)$$

for each data category,  $c$ , in terms of which the bin yields that enter the likelihood are given by

$$N_{c,i}^\pm = \frac{Y_{c,i}^\pm}{\sum_j Y_{c,j}^\pm} \times N_{c,\text{total}}^\pm. \quad (5.19)$$

This parameterisation is essentially identical to the expressions in Section 2.4, slightly modified so that the phase-space-integrated yields of  $B^+$  and  $B^-$  decays

**Table 5.7:** Fitted parameter values.

	LL	DD
$\sigma_{D\pi}$ (MeV/ $c^2$ )	$14.27 \pm 0.05$	$14.58 \pm 0.04$
$\sigma_{DK}$ (MeV/ $c^2$ )	$13.61 \pm 0.24$	$14.19 \pm 0.17$
$\mu$ (MeV/ $c^2$ )		$5278.60 \pm 0.04$
Combinatorial Slopes		
Decay mode	Slope ( $10 \times 10^{-3} GeV^{-1}c^2$ )	
$B^\pm \rightarrow D(\rightarrow K_S^0\pi^+\pi^-)\pi^\pm$	$-3.1 \pm 0.2$	$-4.0 \pm 0.1$
$B^\pm \rightarrow D(\rightarrow K_S^0K^+K^-)\pi^\pm$	$-4.1 \pm 0.4$	$-5.5 \pm 0.3$
$B^\pm \rightarrow D(\rightarrow K_S^0\pi^+\pi^-)K^\pm$	$-3.2 \pm 0.2$	$-3.9 \pm 0.2$
$B^\pm \rightarrow D(\rightarrow K_S^0K^+K^-)K^\pm$	$-4.2 \pm 0.4$	$-4.3 \pm 0.4$
Part. Reco.		
$\sigma_{D\pi}^{low}$ (MeV/ $c^2$ )	$13.73 \pm 0.33$	$13.78 \pm 0.28$
$f_{D\pi\pi}^{D\pi}$		$0.268 \pm 0.013$
$f_{D^*\gamma}^{D\pi}$		$0.317 \pm 0.005$

in a given category are determined directly, in lieu of the normalisation constants  $h^\pm$  of that section. As discussed briefly in Section 2.4, there are choices to be made in terms of how the  $x$  and  $y$  are parameterised in the  $B^\pm \rightarrow D\pi^\pm$  channel, and how the  $F_i$  parameters are determined. A series of feasibility studies were carried out to determine the optimal setup; these are presented in the following section, before the final fit setup is described in detail.

### 2457 Feasibility of alternative fit setups

2458 The motivation for promoting the  $B^\pm \rightarrow D\pi^\pm$  channel to a signal channel is two-fold:  
2459 one aim is to extract the information on  $\gamma$  from the  $B^\pm \rightarrow D\pi^\pm$  data, even the  
2460 precision gain is limited, and another is to be able to the  $F_i$  parameters directly  
2461 from the  $B^\pm \rightarrow Dh^\pm$  channels, to avoid the need for a control channel and a  
2462 simulation-reliant efficiency correction. Two different sets of observables can be  
2463 defined to describe the  $CP$ -violation effects in the  $B^\pm \rightarrow D\pi^\pm$  channel:

- 2464 • one option, defined the 8-parameters setup below, is to define a new set of four  
2465 Cartesian for the  $B^\pm \rightarrow D\pi^\pm$  mode,  $(x_-^{D\pi}, y_-^{D\pi}, x_+^{D\pi}, y_+^{D\pi})$ , defined analogously  
2466 to the  $B^\pm \rightarrow DK^\pm$  observables

$$x_\pm^{D\pi} = r_B^{D\pi} \cos(\delta_B^{D\pi} \pm \gamma), \quad y_\pm^{D\pi} = r_B^{D\pi} \sin(\delta_B^{D\pi} \pm \gamma), \quad (5.20)$$

2467 • another, proposed in Refs. [?, ?], is to introduce the parameter

$$\xi_{D\pi^\pm} = \left( \frac{r_B^{D\pi^\pm}}{r_B^{DK^\pm}} \right) \exp[i(\delta_B^{D\pi^\pm} - \delta_B^{DK^\pm})] \quad (5.21a)$$

2468 and determining the observables

$$x_\xi^{D\pi} = \text{Re}[\xi_{D\pi^\pm}] \quad y_\xi^{D\pi} = \text{Im}[\xi_{D\pi^\pm}]. \quad (5.21b)$$

2469 This is denoted the 6-parameters setup below. In terms of  $x_\xi^{D\pi}$  and  $y_\xi^{D\pi}$ , the  
2470 usual Cartesian  $x_\pm$  and  $y_\pm$  are given by

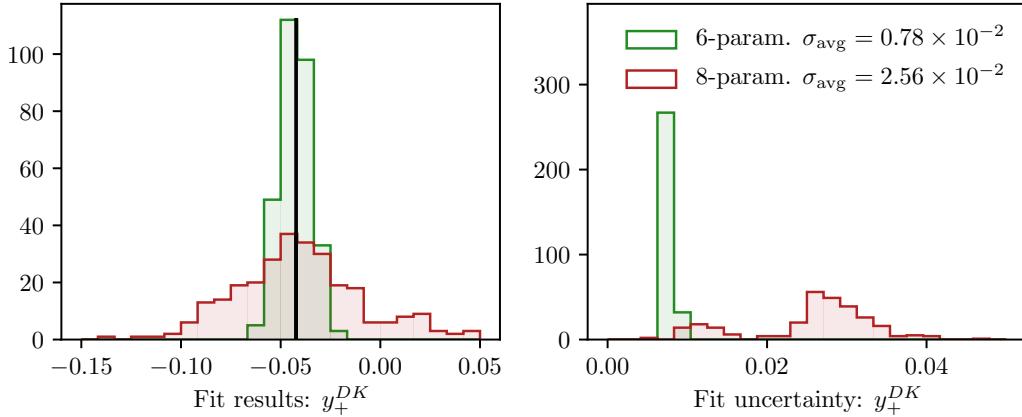
$$x_\pm^{D\pi} = x_\xi^{D\pi} x_\pm^{DK} - y_\xi^{D\pi} y_\pm^{DK}, \quad y_\pm^{D\pi} = x_\xi^{D\pi} y_\pm^{DK} + y_\xi^{D\pi} x_\pm^{DK}. \quad (5.22)$$

2471 The former parameterisation has the benefit that information on  $\gamma$  from the  
2472  $B^\pm \rightarrow DK^\pm$  and  $B^\pm \rightarrow D\pi^\pm$  channels is encoded in separate sets of observables,  
2473 whereas the latter parameterisation encodes information on  $CP$  violation from both  
2474 channel in the  $(x_\pm^{DK}, y_\pm^{DK})$  parameters. In combinations of many measurements, it is  
2475 a useful cross check to be able to compare constraints obtained from individual modes;  
2476 a good example is the LHCb combination from 2016 [?] where both  $B^\pm \rightarrow DK^\pm$  and  
2477  $B^\pm \rightarrow Dh^\pm$  combinations are made, and compared in detail. This is only possible  
2478 with the former parameterisation. On the other hand, the latter parameterisation  
2479 avoids introducing two non-physical degrees of freedom, which, as seen below,  
2480 leads to better statistical behaviour.

2481 In order to inform the choice of parameterisation, a series of pseudo experiments  
2482 have been carried out to compare the obtainable precision on  $\gamma$  (these studies were  
2483 performed, and discussed within LHCb, prior to the publication of Ref. [?]; thus, the  
2484 results presented here constitute independent work, even if there is some overlap in  
2485 scope and conclusions with that reference). Many simulated data sets were generated,  
2486 constituting of a number signal yields approximately equal to the expected yields  
2487 in the full Run 1 and 2 LHCb data set: approximately 15,000  $B^\pm \rightarrow DK^\pm$  decays  
2488 and 210,000  $B^\pm \rightarrow D\pi^\pm$  decays.<sup>3</sup> The signal decays were distributed between  
2489 Dalitz bins according to  $(\gamma, r_B^{DK}, \delta_B^{DK}) = (75^\circ, 0.1, 130^\circ)$  in the  $B^\pm \rightarrow DK^\pm$  mode,  
2490 which is to the world average values of direct  $\gamma$  measurements at the time. In the  
2491  $B^\pm \rightarrow D\pi^\pm$  mode, the behaviour is investigated for different sets of input values; of  
2492 most importance is the case  $(r_B^{D\pi}, \delta_B^{D\pi}) = (0.005, 300^\circ)$ , because it corresponds to  
2493 the solution in the LHCb combination [?] that is in agreement with the theoretical

---

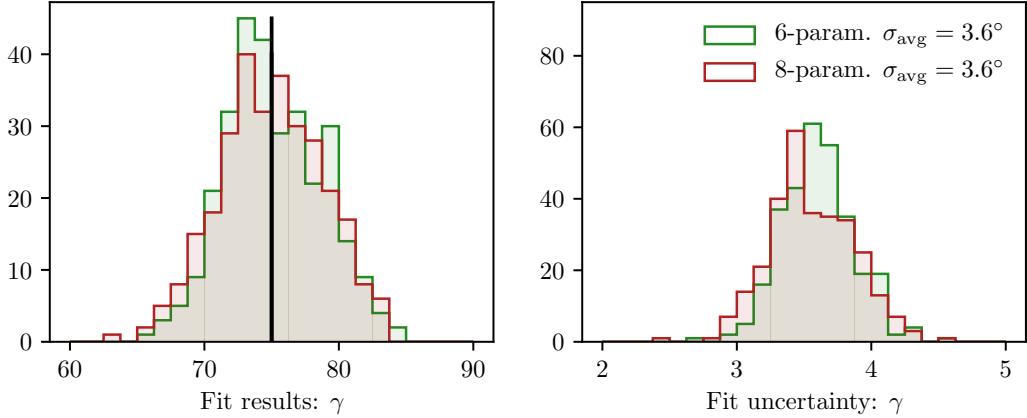
<sup>3</sup>No backgrounds were included in these studies, and thus the quoted uncertainties on  $\gamma$  are better than what is obtainable in the final measurement; a similar study including realistic backgrounds is presented for the final setup below.

**Figure 5.35:** Caption here

expectation  $r_B^{D\pi} \simeq 0.005$  [?]. The behaviour at larger  $r_B^{D\pi}$  values is also investigated. The generated data sets are fitted using both parameterisations, and the observables are interpreted in terms of the underlying physics parameters using a maximum-likelihood fit, essentially following the procedure outlined in Section 5.7.1. In the 8-parameter setup it is possible to determine  $\gamma$  using the results in either the  $B^\pm \rightarrow DK^\pm$  or  $B^\pm \rightarrow D\pi^\pm$  channels separately, consider the combined results; in the 6-parameter setup only the latter option is available. The studies are performed in two modes: with the  $F_i$  floating in the fit, emulating a realistic fit to data, as well as with the  $F_i$  fixed to the input values used in data generation. emulates a setup where the  $F_i$  parameters are determined in an ultra-high statistics control channel, and perfect efficiency corrections are applied. In all cases, a single set of  $F_i$  parameters is shared between the  $B^\pm \rightarrow DK^\pm$  and  $B^\pm \rightarrow D\pi^\pm$  modes.

The 6-parameter setup shows significantly better statistical performance than the 8-parameter setup in the realistic case where the  $F_i$  parameters are determined in the fit and  $r_B^{D\pi} < 0.03$ . The fits that employ the 6-parameter setup behave well in this case, whereas the additional degrees of freedom in the 8-parameter fit leads to essentially all parameters being 100% (anti-)correlated, and a significant number of fits not converging. For the fits that do converge, the uncertainties on the observables are significantly larger due to the large correlations, as shown exemplified with the case of  $y_+^{DK}$  in Fig. 5.35. This essentially determines the choice of parameterisation: it is possible to reliably model  $CP$  violation in the  $B^\pm \rightarrow D\pi^\pm$  channel and simultaneously determine the  $F_i$  parameters by using the 6-parameter setup, but not by using the 8-parameter setup.

Interestingly, when constraints on  $\gamma$  are . This is illustrated in Fig. ??

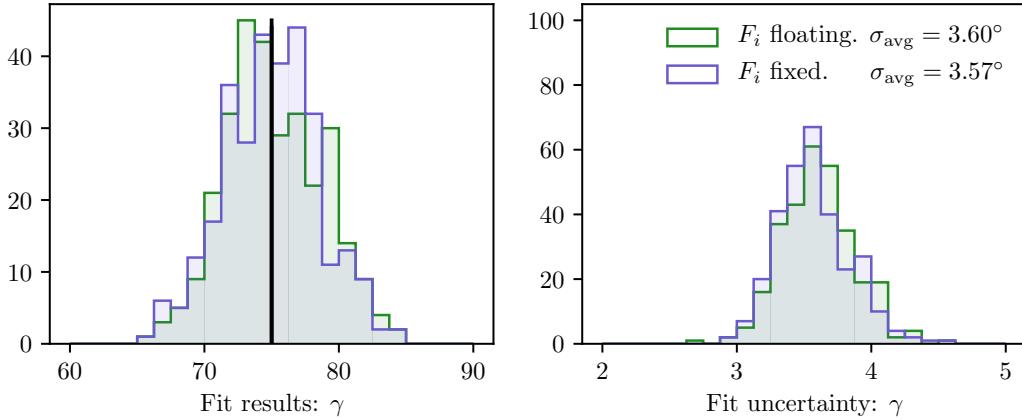
**Figure 5.36:** Caption here

Furthermore, both the 6- and 8-parameter setups lead to fits that behave well in the studies where the  $F_i$  parameters are kept fixed, and the resulting uncertainties on the Cartesian observables and  $\gamma$  are essentially identical. Thus, the 6-parameter setup does not inherently lead to a gain in precision over the 8-parameter setup; the strength of the parameterisation is that it allows for the determination of the  $F_i$  parameters. This conclusion differs somewhat from the one drawn in Ref. [?].

The fixed- $F_i$  studies allow for an assessment of the gain in precision on  $\gamma$  due to the inclusion of the  $B^\pm \rightarrow D\pi^\pm$  mode by comparing the precision obtained in the simultaneous fits, with that obtained when only incorporation information from the  $B^\pm \rightarrow DK^\pm$  channel to constrain  $\gamma$ . In the realistic case where  $r_B^{D\pi} = 0.005$ , the gain in precision is about  $0.1^\circ$ . The reason for the small impact, in spite of the yield being approximately 14 times larger in the  $B^\pm \rightarrow D\pi^\pm$  channel than in the  $B^\pm \rightarrow DK^\pm$  channel, is that  $r_B$  is 20 times smaller and the  $CP$  asymmetries are proportional to  $r_B$ . Thus, the main improvement to the analysis from including  $B^\pm \rightarrow D\pi^\pm$  as a signal channel comes from the ability to determine the  $F_i$  parameters without adding a large systematic uncertainty.<sup>4</sup>

Finally, it is worth considering whether any precision can be gained by including information on the  $F_i$  parameters from a control channel, even if the fit is well behaved without external information. The potential yield in the  $\bar{B}^0 \rightarrow D^{*+}(\rightarrow D^0\pi^+)\mu^-\bar{\nu}_\mu X$  control channel is approximately three times larger than in the  $B^\pm \rightarrow D\pi^\pm$  channel, and it does therefore offer a better statistical handle on the  $F_i$  values (at the significant cost of having to worry about efficiency).

<sup>4</sup>If this comparison is made using the alternative parameter set  $(r_B^{D\pi}, \delta_B^{D\pi}) = (0.03, 330^\circ)$ , which corresponds to the alternative, non-physical solution in the LHCb combination [?], the gain in precision is  $1.3^\circ$  instead; this fact made the statistical interpretation of the  $B^\pm \rightarrow Dh^\pm$  combination in Ref. [?] non-trivial.



**Figure 5.37:** Caption here

corrections). This question can be answered by comparing the obtained precision on  $\gamma$  in the fits where  $F_i$  parameters were floating, to the precision in the case where they were kept fixed. Such a comparison is shown in Fig. 5.37 for the realistic scenario where  $r_B^{D\pi} = 0.005$ . The difference in the average  $\sigma(\gamma)$  is *less than*  $0.05^\circ$ , which is of course completely negligible. Therefore, no gain in precision can be obtained by including the control channel in the analysis, and it is not considered further.

#### 2546 Final choice of observables and the determination of the $F_i$ parameters

2547 In the chosen setup, a single set of four parameters,  $(x_-^{DK}, y_-^{DK}, x_+^{DK}, y_+^{DK})$ , are  
 2548 shared between *all*  $B^\pm \rightarrow DK^\pm$  categories; they enter the expressions of Eq. (5.17)  
 2549 directly, and are thus determined in the fit. In the  $B^\pm \rightarrow D\pi^\pm$  categories, the  
 2550 four corresponding parameters,  $(x_-^{D\pi}, y_-^{D\pi}, x_+^{D\pi}, y_+^{D\pi})$ , are parameterised in terms  
 2551 of  $(x_\pm^{DK}, y_\pm^{DK})$  and the additional two observables  $(x_\xi^{D\pi}, y_\xi^{D\pi})$ . The  $F_i$  parameters  
 2552 are determined in the fit, being shared between the  $B^\pm \rightarrow DK^\pm$  and  $B^\pm \rightarrow D\pi^\pm$   
 2553 channels. However, separate parameter sets are determined for the LL and DD  
 2554 categories because the acceptance profile over the Dalitz plot differs between them.

2555 Because the  $F_i$  are subject to the constraint that  $\sum_{i=-N}^N F_i^c = 1$ , it is beneficial to  
 2556 introduce a reparameterisation in the likelihood function. The  $F_i$  are re-expressed  
 2557 in terms of a set of recursive fractions

$$\mathcal{R}_i = \begin{cases} F_i & , \quad i = -N \\ F_i / (\sum_{j \geq i} F_j) & , \quad -N < i < +N \end{cases} , \quad (5.23)$$

2558 for which the constraint is much simpler, namely that each  $\mathcal{R}_i$  lies in the interval  
 2559  $[0, 1]$ . This results in much better convergence behaviour in the minimisation  
 2560 of the negative log likelihood.

2561 **Strong-phase inputs**

2562 The strong-phase parameters ( $c_i, s_i$ ) are fixed in the fit to data. In the  $D \rightarrow K_S^0 \pi^+ \pi^-$   
 2563 channels, the combined CLEO [?] and BESIII [?] measurement results are used, as  
 2564 reported in Ref. [?]. The  $D \rightarrow K_S^0 K^+ K^-$  categories also use combined CLEO [?]  
 2565 and BESIII results [?], which are reported in Ref. [?]. The experimental uncertainty  
 2566 on these measurements is propagated to the measured  $CP$ -violation observables  
 2567 as part of the systematic uncertainties in Section 5.6.1.

2568 **Treatment of backgrounds**

2569 The yield of combinatorial background decays is determined independently in each  
 2570 bin. A single, overall bin yield of partially reconstructed background from  $B^\pm$  and  
 2571  $B^0$  decays is determined in each of the 160 categories; the relative contribution  
 2572 from each individual background is fixed from the results of the first-stage fit,  
 2573 corrected for the different fit region (a systematic uncertainty is assigned due to this  
 2574 choice). In the  $B^\pm \rightarrow DK^\pm$  channels, the bin yields of the partially reconstructed  
 2575 background from  $B_s^0 \rightarrow \bar{D}^0 [\pi^+] K^-$  decays are expressed via the  $F_i$ , exploiting that  
 2576 a positive companion particle is always produced along with a  $\bar{D}^0$  meson (and vice  
 2577 versa). The overall yield is fixed from the results of the first stage fit. Finally,  
 2578 the yield of the  $D\pi^\pm \leftrightarrow DK^\pm$  cross-feed components in each bin are determined  
 2579 via the obtained yield of correctly identified decays in the corresponding bin, and  
 2580 the known PID efficiencies. This is true for both fully and partially reconstructed  
 2581 decays, although only a  $D\pi^\pm \rightarrow DK^\pm$  component is included in the latter case.

2582 **The choice of fit range**

2583 The fit range is decreased to  $m_B \in [5150, 5800] \text{ MeV}/c^2$ . The information from  
 2584 candidates with lower reconstructed  $B$  masses was useful in determining the relative  
 2585 rates and free mass shape parameters of the partially reconstructed background  
 2586 components in the first-stage fit; however, with these fixed in the second-stage  
 2587 fit, this is no longer the case. Furthermore, the setup assumes that the shape of  
 2588 the partially reconstructed background is identical across the Dalitz bins. This  
 2589 assumption is not perfectly true, but the impact is minimal when the lower limit of  
 2590 the fit range is taken to be  $5150 \text{ MeV}/c^2$ , as described further in Section 5.6.3.

**Table 5.8:** Mean biases and pulls for the observables of interest in the final, binned fit, obtained in a large number of pseudoexperiments.

Parameter	Name in code	Mean bias ( $\times 10^{-2}$ )	Mean pull	Pull width
$x_-^{DK}$	A_xm_dk	$-0.018 \pm 0.022$	$-0.01 \pm 0.02$	$1.01 \pm 0.02$
$y_-^{DK}$	A_ym_dk	$-0.014 \pm 0.026$	$-0.00 \pm 0.02$	$0.99 \pm 0.02$
$x_+^{DK}$	A_xp_dk	$-0.018 \pm 0.022$	$-0.01 \pm 0.02$	$1.00 \pm 0.02$
$y_+^{DK}$	A_yp_	$-0.016 \pm 0.028$	$0.01 \pm 0.02$	$1.00 \pm 0.02$
$x_\xi^{D\pi}$	A_Re_xi_dpi	$0.029 \pm 0.052$	$0.06 \pm 0.02$	$1.00 \pm 0.02$
$y_\xi^{D\pi}$	A_Im_xi_dpi	$0.000 \pm 0.060$	$0.01 \pm 0.02$	$1.00 \pm 0.02$

### 2591 Self-consistency check

2592 In order to establish the fit stability and investigate a potential bias, a series of  
 2593 pseudoexperiments are run, in which toy datasets are generated using the model,  
 2594 and then fitted back. The total yields are taken from the first-stage fit. The signal  
 2595 yields are distributed between Dalitz bins using input physics parameters that  
 2596 approximately equal the values obtained in Section 5.5.2 from the results of the fit  
 2597 to data. The  $F_i$  parameters are taken from a fit to data. The partly reconstructed  
 2598 background is distributed as " $D^0$ -like", ie. in the  $B^\pm$  channels  $N_i^\pm \propto F_{\mp i}$ , except for  
 2599 the  $B_s^0$  background, which is " $\bar{D}^0$ -like" ( $N_i^\pm \propto F_{\pm i}$ ). The combinatorial background  
 2600 includes real  $D$  mesons paired with a random bachelor, as well as fake  $D$  mesons  
 2601 that are themselves made up of random tracks. The former is distributed as 50/50  
 2602  $D^0$ -like and  $\bar{D}^0$ -like in the toy generation, whereas the latter is assumed to be evenly  
 2603 distributed over the Dalitz plot (ie. the bin yield is proportional to the bin area).

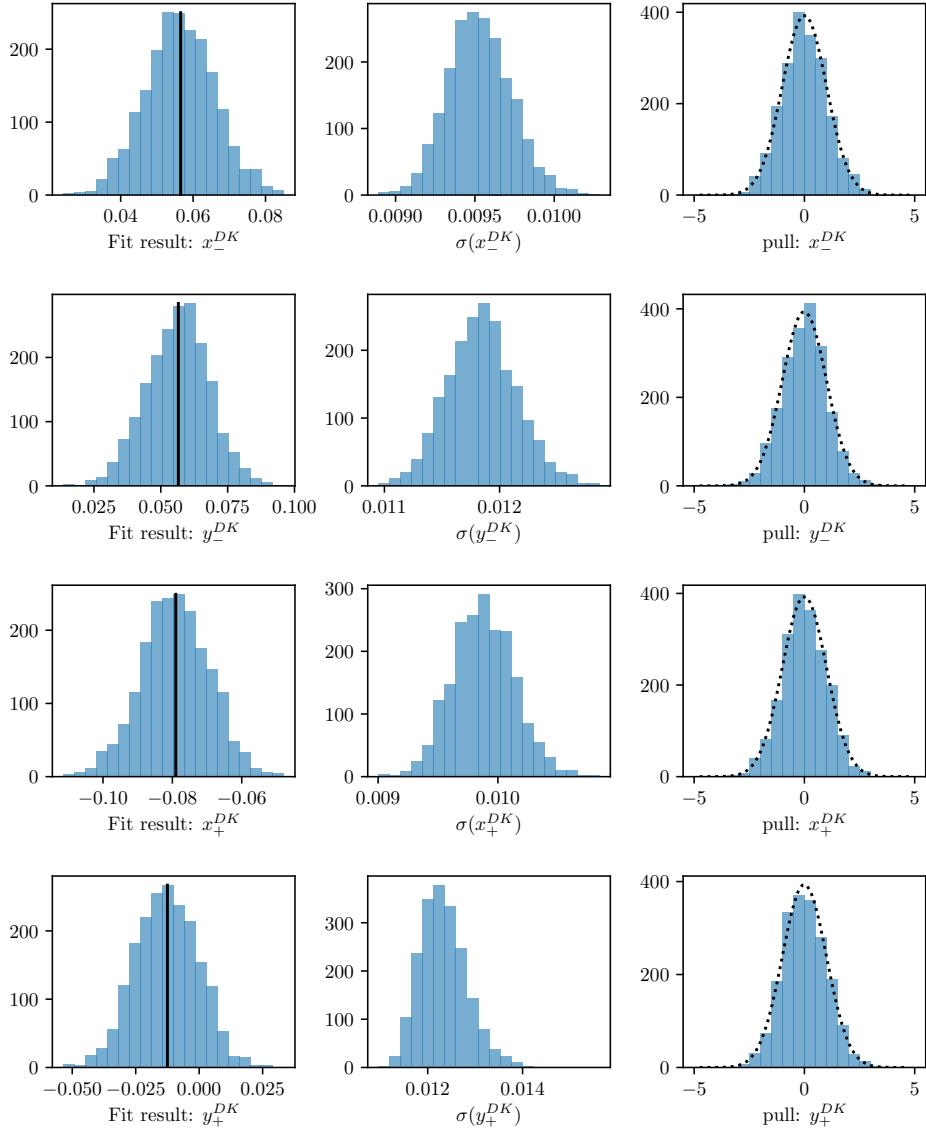
2604 A set of 2000 pseudoexperiments has been run, out of which 98.8 % con-  
 2605 verged properly. The pull plots for the observables of interest are shown in  
 2606 Figs. 5.38 and 5.39; the mean biases and pulls are summarised in Table 5.8. No  
 2607 biases are statistically significant, and the uncertainties are seen to be well estimated.

### 2608 5.5.2 Main results

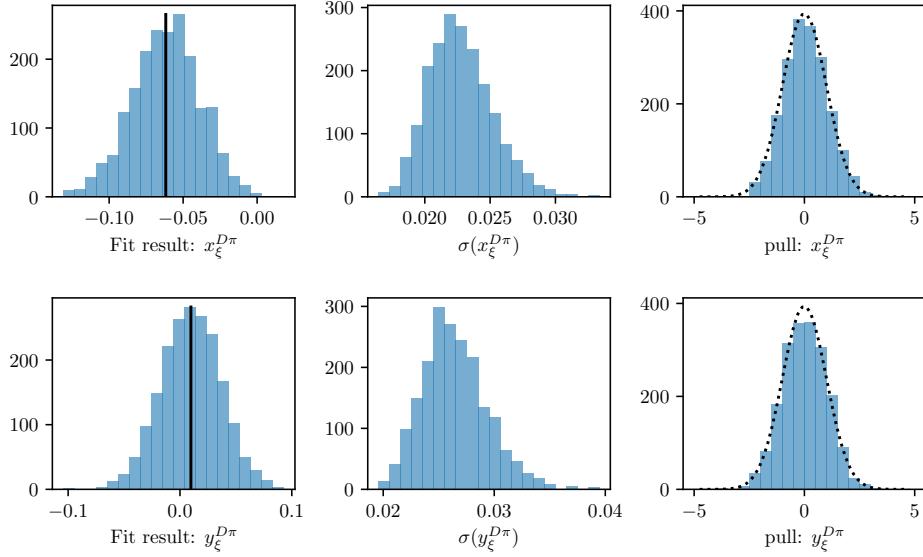
2609 The values and statistical uncertainties of observables obtained in the fit are

$$\begin{aligned} x_-^{DK} &= (-5.68 \pm 0.96) \times 10^{-2}, & y_-^{DK} &= (-6.55 \pm 1.14) \times 10^{-2}, \\ x_+^{DK} &= (-9.30 \pm 0.98) \times 10^{-2}, & y_+^{DK} &= (-1.25 \pm 1.23) \times 10^{-2}, \\ x_\xi^{D\pi} &= (-5.47 \pm 1.99) \times 10^{-2}, & y_\xi^{D\pi} &= (0.71 \pm 2.33) \times 10^{-2}. \end{aligned} \quad (5.24)$$

2610 The statistical correlation matrix for the observables is given in Table 5.9. The  
 2611 2D log-likelihood profile for the observables is shown in Fig. 5.40.



**Figure 5.38:** The (left) fitted value, (centre) estimated statistical uncertainty, and (right) pulls for the  $B^\pm \rightarrow DK^\pm$  observables, as obtained in a number of pseudo experiments. The black line on the left shows the value used to generate the pseudo data sets; the dotted line on the right shows a Gaussian distribution with mean equal to zero and a standard deviation equal to unity.



**Figure 5.39:** The (left) fitted value, (centre) estimated statistical uncertainty, and (right) pulls for the  $B^\pm \rightarrow D\pi^\pm$  observables, as obtained in a number of pseudo experiments. The black line on the left shows the value used to generate the pseudo data sets; the dotted line on the right shows a Gaussian distribution with mean equal to zero and a standard deviation equal to unity.

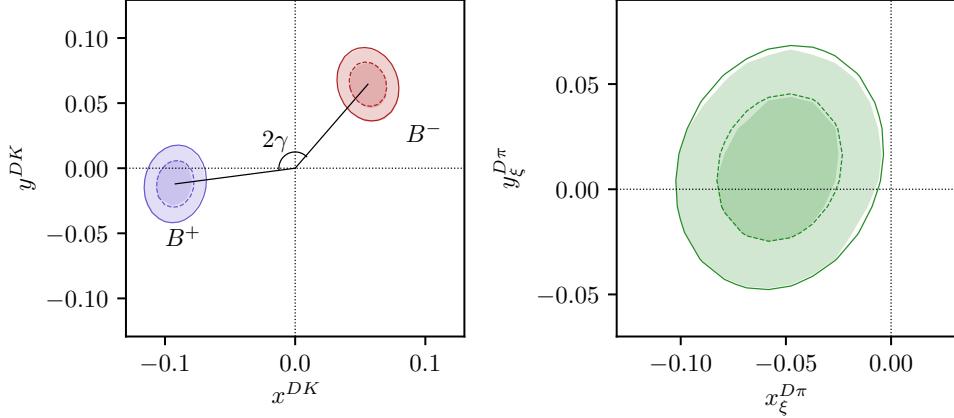
2612     The full set of fit projections in all 160 categories is included in Appendix A.  
 2613     While the  $CP$  asymmetry of the phase-space integrated yield is small, this is not the  
 2614     case for all individual bin-pairs. This is shown in Fig. 5.41 where, as an example,  
 2615     the fit projections for the  $B^+ \rightarrow DK^+$  decays in bin +2 and the  $B^- \rightarrow DK^-$   
 2616     decays in bin −2 of the  $D \rightarrow K_S^0 \pi^+ \pi^-$  Dalitz plot are compared. The presence  
 2617     of  $CP$  violation is clearly visible.

2618     The obtained  $F_i$  parameter values are shown in Table 5.10. These parameters can  
 2619     be useful in other BPGBS measurements made within the LHCb collaboration:  
 2620     it is expected that the systematic uncertainty due to differences between the  
 2621     Dalitz-plot acceptance profile in  $B^\pm \rightarrow Dh^\pm$  decays and, say,  $B \rightarrow D^*K$  or  
 2622      $B \rightarrow DK^*$  decays is smaller than the systematic arising from extracting the  
 2623     efficiency profile from simulated decays. Therefore, the obtain central values and  
 2624     uncertainties have been made public [?], including a set of systematic uncertainties  
 2625     discussed in Section 5.6.12.<sup>5</sup>

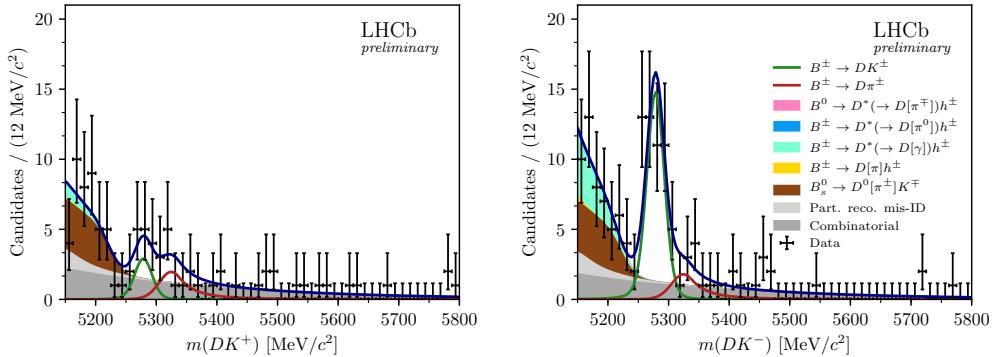
### 2626     5.5.3 Cross checks

2627     A series of cross checks are performed to verify that the fit to data is behav-  
 2628     ing as expected.

<sup>5</sup>In practice, it is the obtained  $\mathcal{R}_i$  values that are made public, related to the  $F_i$  parameters via Eq. (5.23).



**Figure 5.40:** The 68 % and 95 % confidence regions for the fitted observables. The lines show the regions estimated from the covariance matrix returned by the default fit. The shaded areas are obtained in a likelihood scan, where the binned fit is run many times with all observables held at fixed values, scanning pairs of observables over the relevant ranges. The scan is made separately for the three pairs  $(x_{-}^{DK}, y_{-}^{DK})$ ,  $(x_{+}^{DK}, y_{+}^{DK})$ , and  $(x_{\xi}^{D\pi}, y_{\xi}^{D\pi})$ , holding the four other parameters fixed at their default-fit central values during a given scan. Then the minimum log-likelihood is related to a  $\chi^2$  via  $\mathcal{L}_{\min} = \frac{1}{2}\chi^2$  (discarding an irrelevant constant), and the confidence region limits placed at  $\chi^2 = 2.30$  and  $\chi^2 = 6.18$ , yielding the relevant percentiles for a  $\chi^2$  distribution with 2 degrees of freedom.



**Figure 5.41:** The invariant mass distribution for the (left)  $B^+ \rightarrow DK^+$  candidates in bin -2 and (right) the  $B^- \rightarrow DK^-$  candidates in bin +2, where  $D \rightarrow K_S^0\pi^+\pi^-$  and the  $K_S^0$  is reconstructed in the DD category.

### 2629 Comparison to results of earlier analyses

2630 It is confirmed that the results obtained in fits of the Run 1 or 2015+16 data  
 2631 sets in isolation are compatible with the results obtained in the original LHCb  
 2632 analyses of those data sets [?, ?]. In order to do so, the whole analysis procedure  
 2633 is carried out using only the relevant subset of data, and the strong-phase inputs  
 2634 from the CLEO collaboration are used in the fit. Two effects need to be taken

**Table 5.9:** Statistical uncertainties and correlation matrix for the fit to data.

Uncertainty ( $\times 10^{-2}$ )						
	$x_-^{DK^\pm}$	$y_-^{DK^\pm}$	$x_+^{DK^\pm}$	$y_+^{DK^\pm}$	$x_\xi^{D\pi^\pm}$	$y_\xi^{D\pi^\pm}$
$\sigma$	0.96	1.14	0.96	1.20	1.99	2.34

Correlations						
	$x_-^{DK^\pm}$	$y_-^{DK^\pm}$	$x_+^{DK^\pm}$	$y_+^{DK^\pm}$	$x_\xi^{D\pi^\pm}$	$y_\xi^{D\pi^\pm}$
$x_-^{DK^\pm}$	1.000	-0.125	-0.013	0.019	0.028	-0.165
$y_-^{DK^\pm}$		1.000	-0.011	-0.009	0.105	0.030
$x_+^{DK^\pm}$			1.000	0.088	-0.099	0.038
$y_+^{DK^\pm}$				1.000	-0.076	-0.141
$x_\xi^{D\pi^\pm}$					1.000	0.146
$y_\xi^{D\pi^\pm}$						1.000

2635 into account when comparing the central values.

2636 The overlap between the samples need to be taken into account. The overlap  
 2637 is not 100 % due to changes in the candidate selection. The overlap between the  
 2638 new selection and the data set of the original analysis of Run 1 data is about  
 2639 70 %, whereas is it about 90 % for the 2015+16 data set. In order to determine the  
 2640 expected difference between the observables fitted from data sets with significant  
 2641 overlap, a large number of toy data sets were generated in sets of two, where  
 2642 70 (90) % of decays were shared between the data sets. Both data sets were  
 2643 fitted and the difference between the obtained central values for each observable  
 2644 tabulated; the standard deviation of these distributions are used to calculate the  
 2645 pulls between the old analysis results and the new fits to data. This check does  
 2646 not take into account that the semi-leptonic PID cuts were introduced to remove a  
 2647 potential peaking background, which may have had a small systematic effect on  
 2648 the earlier measurement results. Thus the expected differences are likely to be  
 2649 slightly underestimated and the check conservative.

2650 Furthermore, the  $F_i$  parameters were determined in a semi-leptonic control  
 2651 channel in the earlier analyses. Therefore, the expected difference obtained above is  
 2652 adjusted by adding the  $F_i$ -related systematic uncertainty of the original analysis  
 2653 in quadrature, when comparing the old results to those in new fits to the Run 1  
 2654 and 15+16 data sets. No further corrections have been made to the expected  
 2655 differences, which effectively assumes all other systematic uncertainties to be 100 %  
 2656 correlated. Also for this reason can the check be considered conservative. As can

**Table 5.10:** The fitted  $F_i$  values including statistical uncertainties. The associated systematic uncertainties are negligible, as discussed in Section 5.6.12.

$F_i$ values: $D \rightarrow K_S^0 \pi^+ \pi^-$		
bin	LL	DD
-8	$0.024 \pm 0.001$	$0.024 \pm 0.000$
-7	$0.127 \pm 0.001$	$0.133 \pm 0.001$
-6	$0.062 \pm 0.001$	$0.056 \pm 0.001$
-5	$0.046 \pm 0.001$	$0.042 \pm 0.001$
-4	$0.095 \pm 0.001$	$0.095 \pm 0.001$
-3	$0.160 \pm 0.001$	$0.160 \pm 0.001$
-2	$0.153 \pm 0.001$	$0.153 \pm 0.001$
-1	$0.095 \pm 0.001$	$0.097 \pm 0.001$
1	$0.022 \pm 0.001$	$0.020 \pm 0.000$
2	$0.005 \pm 0.000$	$0.005 \pm 0.000$
3	$0.004 \pm 0.000$	$0.004 \pm 0.000$
4	$0.055 \pm 0.001$	$0.056 \pm 0.001$
5	$0.027 \pm 0.001$	$0.022 \pm 0.000$
6	$0.004 \pm 0.000$	$0.003 \pm 0.000$
7	$0.055 \pm 0.001$	$0.057 \pm 0.001$
8	$0.067 \pm 0.001$	$0.072 \pm 0.001$

$F_i$ values: $D \rightarrow K_S^0 K^+ K^-$		
bin	LL	DD
-2	$0.207 \pm 0.004$	$0.202 \pm 0.003$
-1	$0.222 \pm 0.004$	$0.230 \pm 0.003$
1	$0.290 \pm 0.005$	$0.296 \pm 0.003$
2	$0.281 \pm 0.005$	$0.271 \pm 0.003$

be seen in Tables 5.11 and 5.12, neither the Run 1 and 2015+16 comparisons show unreasonable differences in central values.

### Directly fitting the signal yields

As a cross-check, the fit is run in an alternative mode, in which the signal yields of each bin are independent parameters. The obtained yields are compared to those predicted from the results of the default fit in Fig. 5.42. The yields are shown for each "effective bin", where effective bin  $i$  is defined as bin  $+i$  for  $B^+$  decays and bin  $-i$  for  $B^-$  decays; in the  $CP$  symmetric case, these bins are expected to have equal yields (modulo production and detection asymmetries). The agreement between the

**Table 5.11:** Comparison between the results on the Run 1 analysis [?] and the central values obtained when fitting the Run 1 dataset with the selection and fit setup described in this note. The pull is calculated using the  $1\sigma$  expected difference, which takes the sample overlap and the systematic uncertainty on the  $F_i$  parameters in the previous analysis into account, but assumes all other systematic uncertainties to be perfectly correlated. The new fits are performed using the CLEO strong-phase inputs.

Observable	Run 1 result [?] (central value $\times 10^{-2}$ )	New Fit (central value $\times 10^{-2}$ )	Pull
$x_-^{DK}$	2.50	4.04	0.85
$y_-^{DK}$	7.50	9.14	1.02
$x_+^{DK}$	-7.70	-9.40	-0.91
$y_+^{DK}$	-2.20	0.80	1.77
<i>p</i> -value: 0.057			

**Table 5.12:** Comparison between the results on the 2015+16 analysis [?] and the central values obtained when fitting the 2015+16 dataset with the selection and fit setup described in this note. The pull is calculated using the  $1\sigma$  expected difference, which takes the sample overlap and the systematic uncertainty on the  $F_i$  parameters in the previous analysis into account, but assumes all other systematic uncertainties to be perfectly correlated. The new fits are performed using the CLEO strong-phase inputs.

Observable	15+16 result [?] (central value $\times 10^{-2}$ )	New Fit (central value $\times 10^{-2}$ )	Pull
$x_-^{DK}$	9.00	8.36	-0.50
$y_-^{DK}$	2.10	1.16	-0.62
$x_+^{DK}$	-7.70	-8.58	-0.56
$y_+^{DK}$	-1.00	-2.82	-1.39
<i>p</i> -value: 0.239			

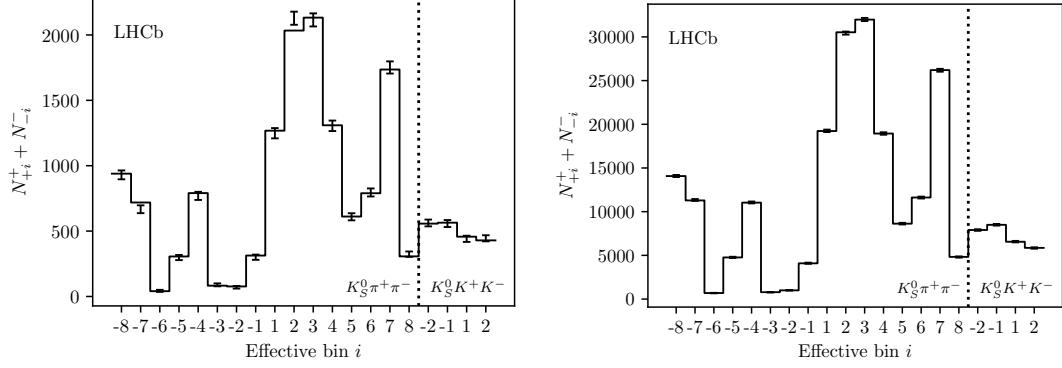
two fit set-ups is seen to be excellent. The normalised yield asymmetries, defined as

$$A^i \equiv \frac{N_{-i}^- - N_i^+}{N_{-i}^- + N_i^+} \quad (5.25)$$

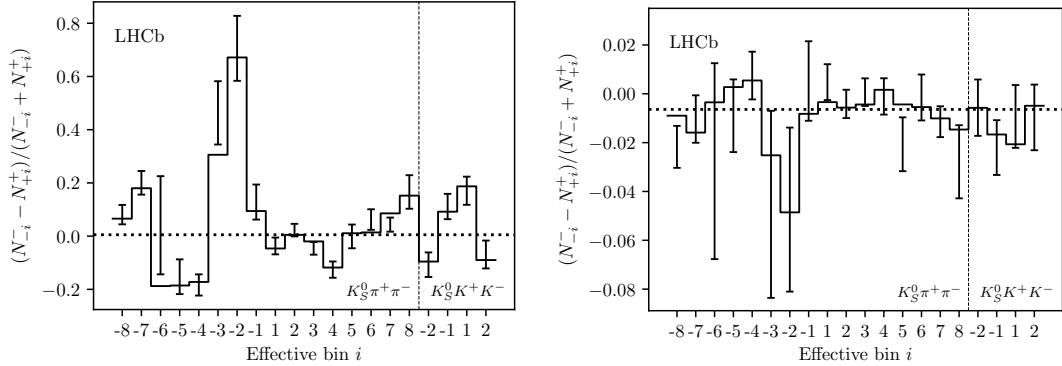
are shown in Fig. 5.43 for all decay channels. Again, the agreement between the nominal fit, and the alternative fit with independent yields is found to be excellent. It is also clear how, in the case of  $B^\pm \rightarrow DK^\pm$  decays, the asymmetry is significantly different from zero for a number of bin pairs.

### 2671 Fitting subsets of the data separately

2672 One cross check is carrying out, by determining the  $CP$  observables using a number  
2673 of independent sub samples of the data set separately. This is done for the  
2674 following following data splits

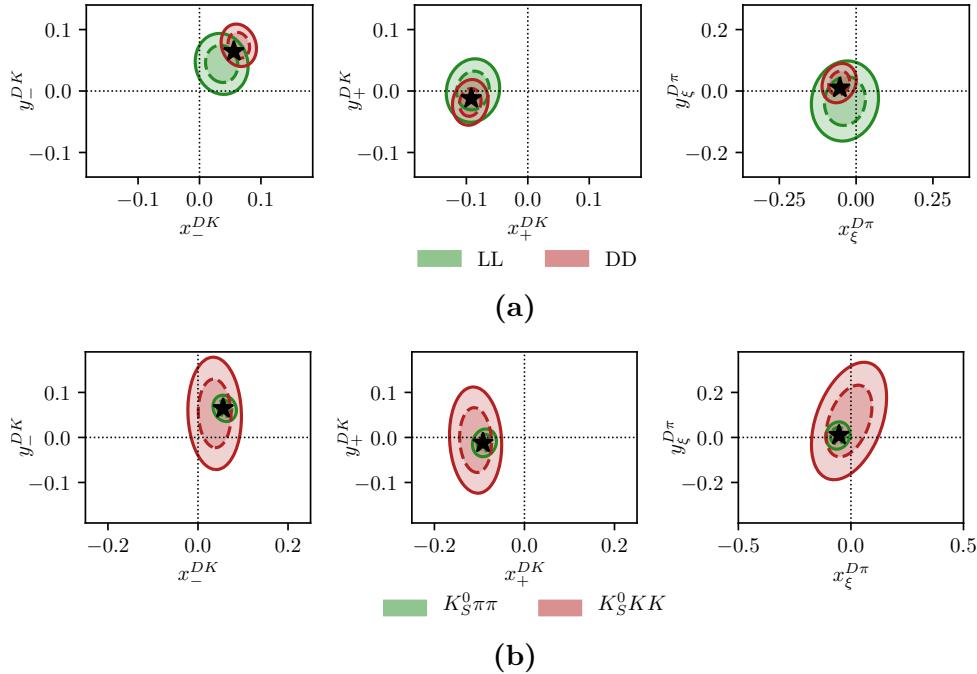


**Figure 5.42:** Comparison of (lines) the predicted yield given the determined  $CP$  observables and (error bars) the yield obtained in fits to data where each yield is an independent parameter. The yields are shown for (left)  $B^\pm \rightarrow DK^\pm$  decays and (right)  $B^\pm \rightarrow D\pi^\pm$  decays. The LL and DD categories have been combined, as has the  $B^+$  and  $B^-$  yields for each effective Dalitz bin, defined as bin  $+i$  for  $B^+$  decays and bin  $-i$  for  $B^-$  decays.



**Figure 5.43:** The bin-by-bin asymmetries  $(N_{-i}^{-} - N_{+i}^{+})/(N_{-i}^{-} + N_{+i}^{+})$  for each Dalitz-plot bin number for (left)  $B^\pm \rightarrow DK^\pm$  decays and (right)  $B^\pm \rightarrow D\pi^\pm$  decays. The prediction from the central values of the  $CP$ -violation observables is shown with a solid line and the asymmetries obtained in fits with independent bin yields are shown with the error bars. The predicted asymmetries in a fit that does not allow for  $CP$  violation are shown with a dotted line.

- 2675     • Fig. 5.44a shows the same plots, comparing the fits to the data set split by
- 2676      $K_S^0$  track type.
- 2677     • Fig. 5.44b shows the same plots, comparing the fits to the data set split by
- 2678     whether the  $D$  meson decays to the  $K_S^0\pi^+\pi^-$  or  $K_S^0K^+K^-$  final state.
- 2679     • Fig. 5.45a shows the two dimensional log likelihood contours for the observables
- 2680     for fits to the Run 1, 2015+16, 2017 and 2018 datasets separately
- 2681     • Fig. 5.45b shows the same plots, comparing the fits to the data set split by



**Figure 5.44:** Comparison of the 68 % and 95 % confidence regions for (left)  $(x_{\pm}^{DK}, y_{\pm}^{DK})$ , (centre)  $(x_{\pm}^{DK}, y_{\pm}^{DK})$ , and (right)  $(x_{\xi}^{D\pi}, y_{\xi}^{D\pi})$  obtained from fits to sub sets of the data set. The uncertainties are statistical only. The central values of the default fit are shown with a black star. The dataset is split by (a) LL and DD  $K_S^0$  types and (b)  $D$  decay mode.

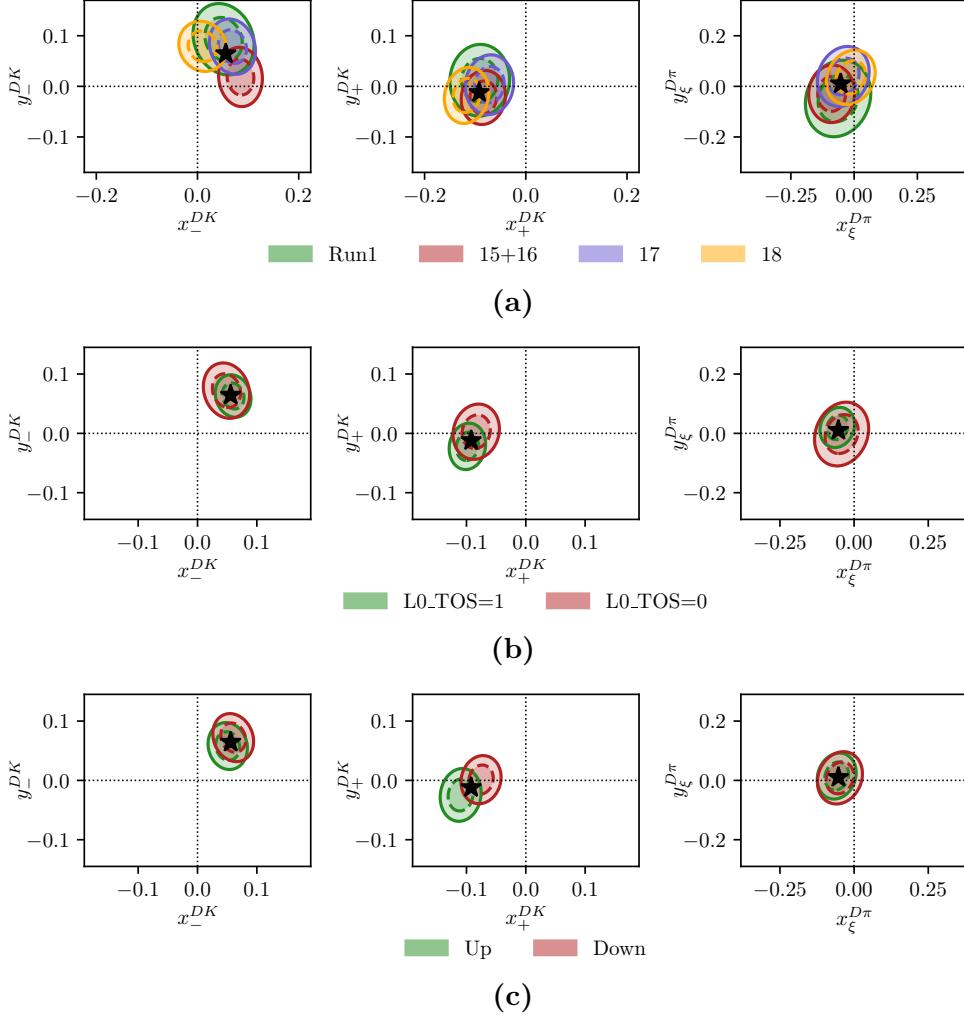
whether the candidate event was triggered by one of the signal particles at the hardware level (TOS), or by another particle in the underlying event (TIS).

- Fig. 5.45c shows the same plots, comparing the fits to the data set split the magnet polarity during data taking.

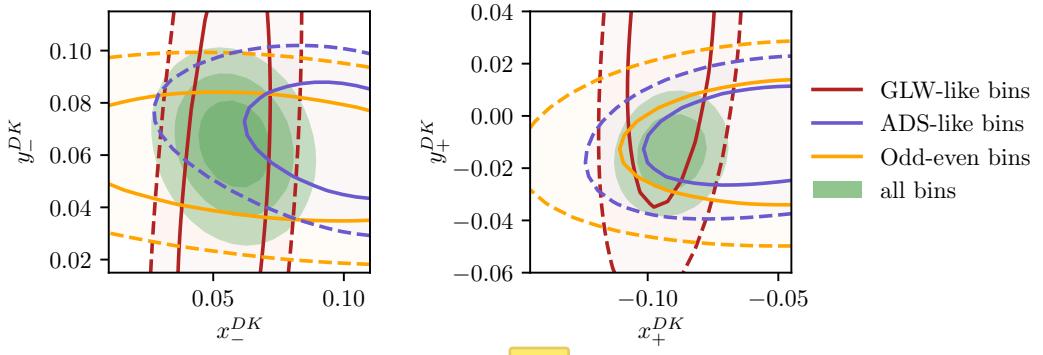
All figures show the Gaussian likelihood contours corresponding to the statistical uncertainties. There is good agreement between the results in all cases, given that in each cases the sub datasets are independent and therefore the statistical errors are uncorrelated.

## Constraints from a subset of bins

An alternative way to subdivide the data is to extract the constraints from a subset of bins individually; this forms as a cross check, seeing that the observables favoured by each sub set should be compatible, and also serves as a useful illustration of the features of the BPGBGSZ method. Likelihood contours for  $(x_{\pm}^{DK}, y_{\pm}^{DK})$  are shown in Fig. 5.46, obtained using the binned yields in the  $D \rightarrow K_S^0 \pi^+ \pi^-$  bins, determined in the fits of individual bin yields described in Section 5.5.3. The bins are split by



**Figure 5.45:** Comparison of the 68 % and 95 % confidence regions for (left)  $(x_-^{DK}, y_-^{DK})$ , (centre)  $(x_+^{DK}, y_+^{DK})$ , and (right)  $(x_\xi^{D\pi}, y_\xi^{D\pi})$  obtained from fits to sub sets of the data set. The uncertainties are statistical only. The central values of the default fit are shown with a black star. The dataset is split by (a) data taking year, (b) trigger category, and (c) magnet polarity.



**Figure 5.46:** Caption here

whether they are ADS-like, GLW-like, or Odd-even according to the classification in Section 2.3.5. It is clear that the likelihood regions show a reasonable overlap, and also how it is the GLW bins that constrain the  $x_\pm$  parameter, while the Odd-even and ADS-like bins provide the ability to constrain the  $y_\pm$  parameters.

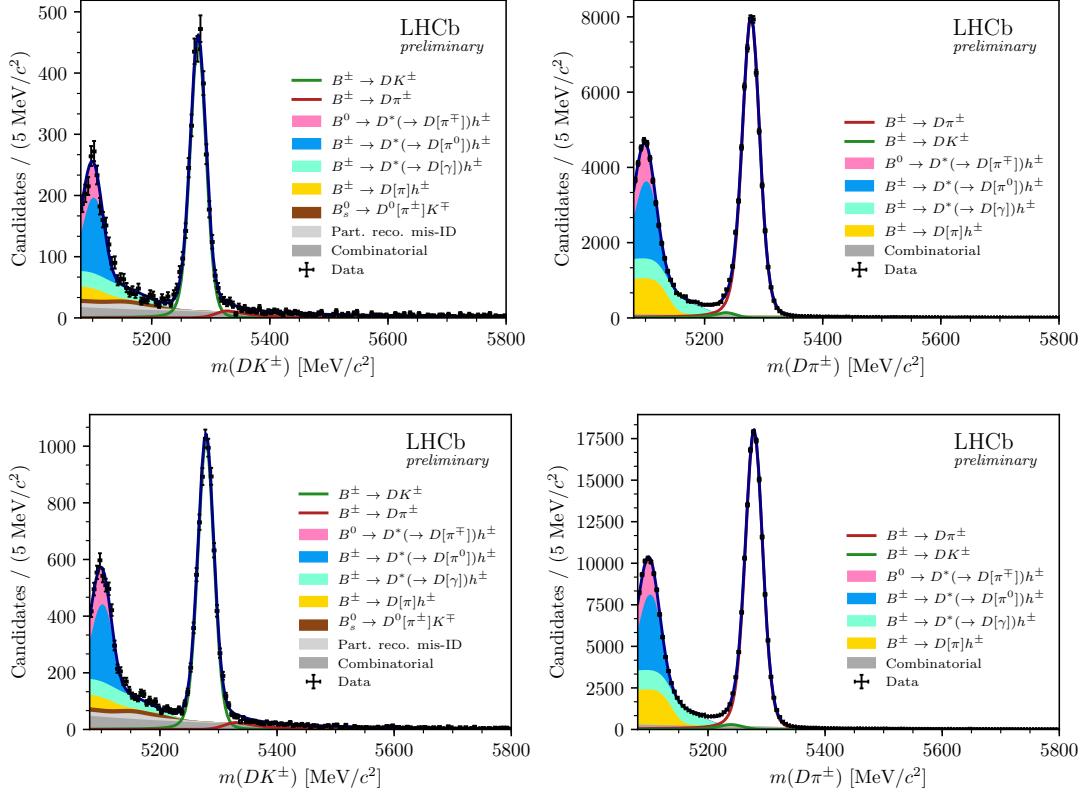
### **2701 Significantly reducing the $B^\pm \rightarrow D\pi^\pm$ to $B^\pm \rightarrow DK^\pm$ cross feed**

One of the dominant backgrounds in the signal region of the  $B^\pm \rightarrow DK^\pm$  channel is from partly reconstructed  $B \rightarrow D\pi X$  decays where the bachelor pion is misidentified as a kaon. The background mode is well described by the included shape component, and included in all relevant systematic studies. Nevertheless, an additional cross check is carried out to ensure that it is not having a significant effect on the fit: the analysis is repeated with PID requirement of  $\text{PID}_K > 12$  required to place a candidate in the  $B^\pm \rightarrow DK^\pm$  category, instead of  $\text{PID}_K > 4$ . With this requirement 99.7 % of  $B^\pm \rightarrow D\pi^\pm$  decays are correctly identified, making the cross-feed component in the  $B^\pm \rightarrow DK^\pm$  channels significantly smaller than in the default fit. This is clearly visible in Fig. 5.47, where the fit projections for the global fit of the  $D \rightarrow K_S^0 \pi^+ \pi^-$  modes are shown. In return, the probability of correctly identifying a kaon companion drops to about 68–69 %, resulting in a smaller effective signal yield.

The measurement results are compared in Table 5.13, where the differences in central value are seen to be reasonably small. It is not trivial to determine whether the difference is statistically significant or not: the same candidates are analysed in both cases, the difference being that a number of candidates that are placed in the  $B^\pm \rightarrow DK^\pm$  category in the nominal fit are placed in the  $B^\pm \rightarrow D\pi^\pm$  category in the alternative fit. The uncertainty will not be 100 % correlated because signal events that move from the  $DK$  to  $D\pi$  category are placed in a region with high background; however, this is somewhat compensated for by candidates that remain in the  $DK$  category gaining statistical power due to the increased purity. An estimate of the expected statistical fluctuation can be determined by taking the difference of the statistical uncertainties in quadrature. Using this estimate, the observed shifts are found to consistent with statistical fluctuation, and thus there is no sign of the background from  $D\pi^\pm \rightarrow DK^\pm$  cross-feed causing issues.

### **2727 Compare results obtained with different strong-phase inputs**

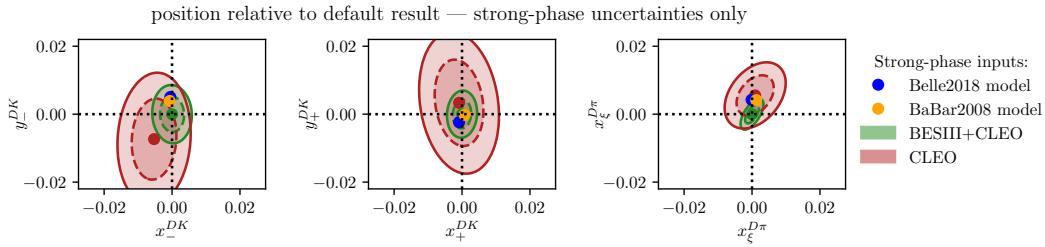
It is interesting to compare the results obtained with different strong-phase inputs. This is done in Fig. 5.48, where the default fit results are compared to those obtained if the  $CP$  fit is done with the CLEO-only inputs [?], and with the model predictions from the 2018 Belle model [?] and the 2008 BaBar model [?]. For the



**Figure 5.47:** Fit projections for fits to the  $D \rightarrow K_S^0 \pi^+ \pi^-$  candidates with a companion PIDK requirement at 12 instead of 4 used to split into (left)  $B^\pm \rightarrow DK^\pm$  and (right)  $B^\pm \rightarrow D\pi^\pm$  candidates, for the (top) LL and (bottom) DD categories.

**Table 5.13:** Results of running the measurement with the default PIDK cut at 4 used to separate  $B^\pm \rightarrow DK^\pm$  and  $B^\pm \rightarrow D\pi^\pm$  candidates, as well as with an alternative PIK cut at 12, resulting in much lower cross-feed from misidentified  $B^\pm \rightarrow D\pi^\pm$  decays. We also show the pulls, defined as  $\Delta x / \sqrt{|\sigma_{PIDK>12}^2 - \sigma_{PIDK>4}^2|}$  as described in the main text body. The comparison was made before the BESIII measurement of the  $D \rightarrow K_S^0 K^+ K^-$  strong-phase inputs became available; therefore the fits use the CLEO-only results [?] for this mode, which explains why the results quoted for  $PIDK > 4$  differ slightly from the nominal fit results.

Parameter	$PIDK > 4$	$PIDK > 12$	$\sigma = \sqrt{\sigma_{PIDK>12}^2 - \sigma_{PIDK>4}^2}$	Pull
$x_-^{DK}$	$5.59 \pm 0.96$	$5.82 \pm 1.01$	0.30	0.77
$y_-^{DK}$	$6.45 \pm 1.14$	$6.86 \pm 1.19$	0.36	1.13
$x_+^{DK}$	$-9.21 \pm 0.96$	$-8.94 \pm 1.01$	0.30	0.93
$y_+^{DK}$	$-1.21 \pm 1.20$	$-0.94 \pm 1.26$	0.37	0.71
$x_\xi^{D\pi}$	$-5.30 \pm 1.99$	$-5.13 \pm 2.02$	0.32	0.52
$y_\xi^{D\pi}$	$1.03 \pm 2.34$	$1.71 \pm 2.33$	0.28	2.40



**Figure 5.48:** Fit results for (left)  $(x_-^{DK}, y_-^{DK})$ , (centre)  $(x_+^{DK}, y_+^{DK})$ , and (right)  $(x_\xi^{D\pi}, y_\xi^{D\pi})$  depending on strong-phase inputs, shown relative to the default fit results. The included results are based on (green) the BESIII-CLEO combination, which is the default, (red) the CLEO-only results, (blue dot) the 2018 Belle model [?] and (orange dot) the 2008 BaBar model [?]. For the measurements, only strong-phase related uncertainties are included in the plotted confidence regions.

measurements, only the strong-phase-related uncertainties are included in the plot,  
since the statistical uncertainties are correlated. All results are found to agree well.

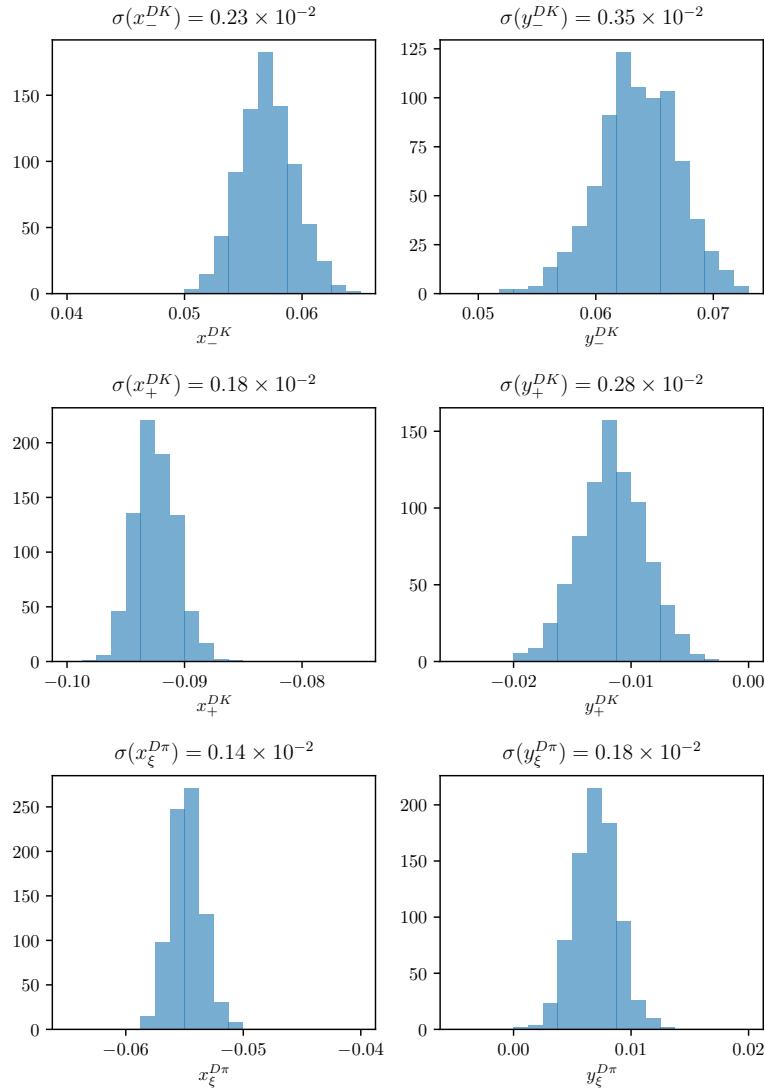
## 5.6 Systematic uncertainties

The following sections cover the suite of systematic uncertainties on the measurement that has been considered. All uncertainties are summarised in Section 5.6.12.

### 5.6.1 Strong phase uncertainties

The observables  $x_\pm^{DK}$ ,  $y_\pm^{DK}$ ,  $x_\xi^{D\pi}$  and  $y_\xi^{D\pi}$  are extracted using the central values of  $c_i$  and  $s_i$  from the BESIII–CLEO combinations [?, ?, ?]. Subsequently, the measurement uncertainty on these inputs is propagated to the observables by performing a large set of fits to data using alternative values of  $c_i$  and  $s_i$ . The new  $c_i$  and  $s_i$  values are obtained by smearing the central values by their measured statistical and systematic uncertainties while taking into account their correlation. The use of different  $c_i$  and  $s_i$  values changes the extracted  $x_\pm^{DK}$ ,  $y_\pm^{DK}$ ,  $x_\xi^{D\pi}$  and  $y_\xi^{D\pi}$  values. The width of the distributions of central values extracted from 1000 data fits are assigned as a systematic uncertainty. The distributions are shown in Fig. 5.49 and the assigned uncertainties are summarised in Table 5.14. The correlation matrix related to the strong-phase uncertainty can be obtained from the correlations observed between observables in the fits, and is also given in the table.

The set of  $(c_i, s_i)$  that was employed in this analysis will be used in a series of future BPGGSZ measurements, both with additional  $B$  decay modes within the LHCb collaboration and by the Belle II collaboration. This introduces some correlation between the measurement results. In order to allow for an estimate



**Figure 5.49:** Spread of central values for fitted observables when the input  $c_i$  and  $s_i$  from the BESIII+CLEO combination are varied according to their uncertainties and correlations.

of the degree of correlation by future analysts, the 1000 samples  $(c_i, s_i)$  values and the corresponding fit results for  $(x_\pm^{DK}, y_\pm^{DK}, x_\xi^{D\pi}, y_\xi^{D\pi})$  have been made public as supplementary material to Ref. [?].

### 5.6.2 Efficiency-profile-related systematic uncertainties

The non-trivial efficiency profile over the Dalitz plot can have a range of effects, considered in the sections below.

**Table 5.14:** Systematic uncertainties and correlation matrix due to strong-phase inputs.

Uncertainty ( $\times 10^{-2}$ )						
	$x_-^{DK\pm}$	$y_-^{DK\pm}$	$x_+^{DK\pm}$	$y_+^{DK\pm}$	$x_\xi^{D\pi\pm}$	$y_\xi^{D\pi\pm}$
$\sigma$	0.23	0.35	0.18	0.28	0.14	0.18

Correlations						
	$x_-^{DK\pm}$	$y_-^{DK\pm}$	$x_+^{DK\pm}$	$y_+^{DK\pm}$	$x_\xi^{D\pi\pm}$	$y_\xi^{D\pi\pm}$
$x_-^{DK\pm}$	1.000	-0.047	-0.490	0.322	0.189	0.144
$y_-^{DK\pm}$		1.000	0.059	-0.237	-0.116	-0.117
$x_+^{DK\pm}$			1.000	0.061	0.004	-0.139
$y_+^{DK\pm}$				1.000	0.127	-0.199
$x_\xi^{D\pi\pm}$					1.000	0.638
$y_\xi^{D\pi\pm}$						1.000

2760 **The assumption that**  $\eta^{DK}(s_-, s_+) = \eta^{D\pi}(s_-, s_+)$ 

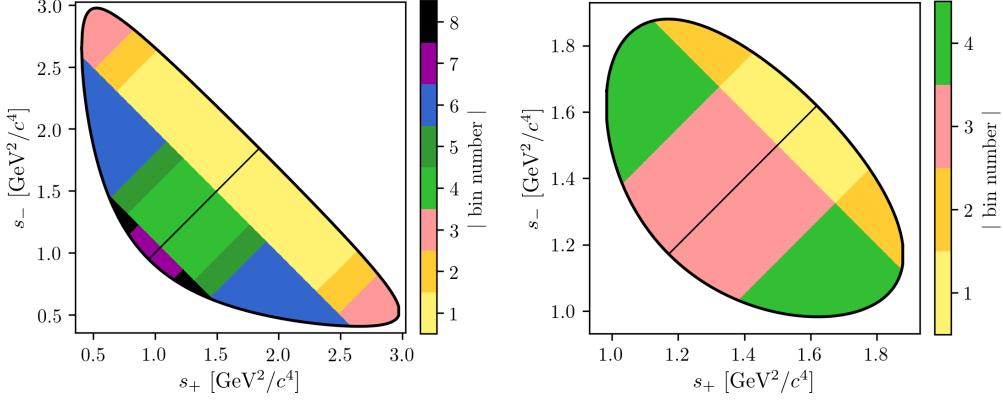
2761 The assumption that  $\eta^{DK}(s_-, s_+) = \eta^{D\pi}(s_-, s_+)$  was examined in detail in Sec-  
 2762 tion 5.2.2. It was found that with sign  yields similar to those in the data set,  
 2763 no statistically significant difference between the efficiency profiles  $\eta^{DK}(s_-, s_+)$   
 2764 and  $\eta^{D\pi}(s_-, s_+)$  was discernible, and no additional uncertainty due to this as-  
 2765 sumption is assigned.

2766 **The assumption that**  $\eta(s_-, s_+) = \eta(s_+, s_-)$ 

2767 The measurement is sensitive to effects that break the assumption  $\eta(s_-, s_+) =$   
 2768  $\eta(s_+, s_-)$ . Such a breakdown would mean that opposite points on the Dalitz plot  
 2769 have different efficiencies and can only arise through a charge detection asymmetry  
 2770 (e.g. that it is more likely to detect a  $K^+$  in the detector rather than a  $K^-$ ).<sup>6</sup>

2771 The potential size of such an asymmetry can be studied in simulation  
 2772 where the  $D$  decay has a uniform distribution over the allowed phase space; in such  
 2773 simulated samples, it would manifest itself as an observation of different fractional  
 2774 yields of  $B^-$  decays in bin  $i$  and  $B^+$  decays in bin  $-i$ . This effect has been looked for  
 2775 using the large samples of  $B^\pm \rightarrow D\pi^\pm$  decays that were generated for the analysis  
 2776 of 2015 and 2016 data. The study is performed using the rectangular binning  
 2777 schemes shown in Fig. 5.50, because this scheme is most sensitive to effects that vary

<sup>6</sup>Note that the measurement is insensitive to any asymmetry in the reconstruction of the companion track.



**Figure 5.50:** The rectangular binning schemes used to group candidates in (left) the  $D \rightarrow K_S^0 \pi^+ \pi^-$  and (right) the  $D \rightarrow K_S^0 K^+ K^-$  categories in a number of systematic uncertainty studies.

smoothly over phase space. The comparison plots are shown in Fig. 5.51, where it can be seen that the  $p$  values for the hypothesis that there is no asymmetry all take on reasonable values. Hence no further systematic uncertainty is considered.

#### Effect of phase-space efficiency profile on $c_i$ and $s_i$

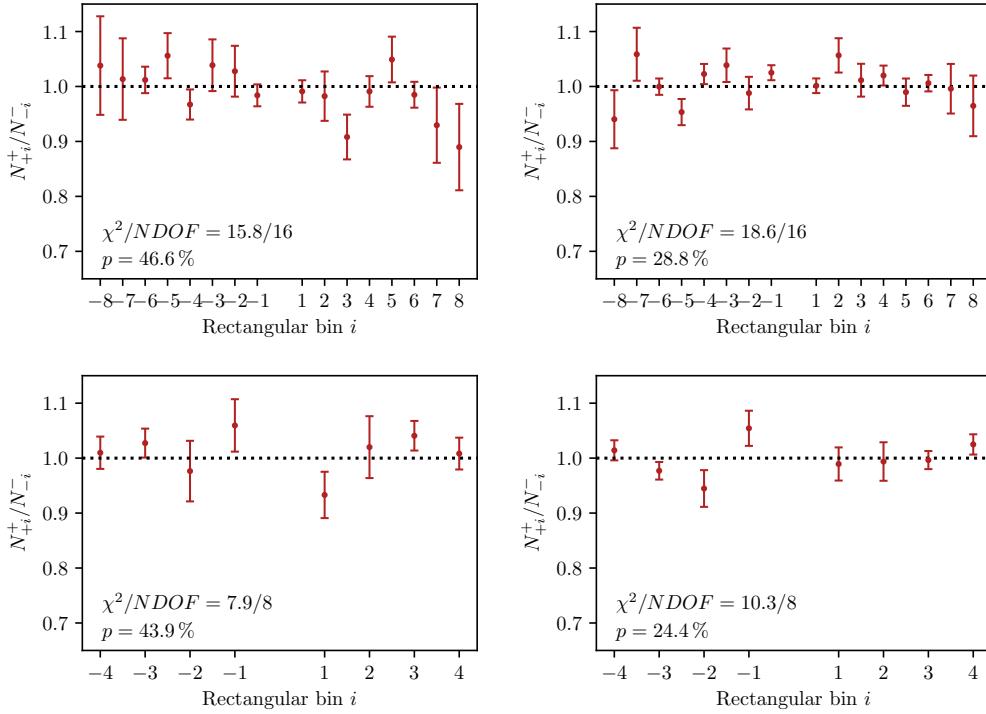
As discussed in Section 2.4 there is a small bias, because the  $c_i$  and  $s_i$  values that are used correspond to the definition

$$c_i = \frac{\int_i ds^2 |A_S^D(s_{-+})| |A_S^D(s_{+-})| \cos[\Delta\delta_D(s_{-+})]}{\sqrt{\int_i ds^2 |A_S^D(s_{-+})|^2} \sqrt{\int_i ds^2 |A_S^D(s_{+-})|^2}}, \quad (\text{and equivalent for } s_i,) \quad (5.26)$$

whereas the non-flat efficiency profile in LHCb,  $\eta(s_-, s_+) \equiv \eta(s_{-+})$ , means that the appropriate  $c_i'$  and  $s_i'$  entering the exact yield expressions are

$$c_i^{\text{eff}} = \frac{\int_i ds^2 \eta(s_{-+}) |A_S^D(s_{-+})| |A_S^D(s_{+-})| \cos[\Delta\delta_D(s_{-+})]}{\sqrt{\int_i ds^2 |A_S^D(s_{-+})|^2} \sqrt{\int_i ds^2 |A_S^D(s_{+-})|^2}}, \quad (\text{and equivalent for } s_i^{\text{eff}}.) \quad (5.27)$$

The shifts  $\Delta c_i = c_i^{\text{eff}} - c_i$ ,  $\Delta s_i = s_i^{\text{eff}} - s_i$  can be estimated using the efficiency profile in simulation and the latest amplitude models: the 2018 Belle model [?] for  $D \rightarrow K_S^0 \pi^+ \pi^-$  and the 2010 BaBar model [?] for  $D \rightarrow K_S^0 K^+ K^-$ . The strong-phase parameters are first calculated assuming a uniform reconstruction efficiency over phase space according to Eq. (5.26), obtaining a set of values  $\{c_i^{\text{model}}, s_i^{\text{model}}\}$ . Then, an alternative set is calculated,  $\{c_i^{\text{eff}}, s_i^{\text{eff}}\}$ , using the same model, and the reconstruction efficiency profile found in full LHCb simulation. The results, as well



**Figure 5.51:** Comparison of the ratio of  $B^+$  decays reconstructed in bin  $+i$  to  $B^-$  decays reconstructed in bin  $-i$  for simulated  $B^\pm \rightarrow D\pi^\pm$  decays where (top)  $D \rightarrow K_S^0 \pi^+ \pi^-$  and (bottom)  $D \rightarrow K_S^0 K^+ K^-$ , also split into (left) the LL and (right) the DD categories, using the rectangular binning schemes in Fig. 5.50. Calculated  $p$  values for the hypothesis that the ratio is flat are also shown, all of them being at least 20 %.

as their differences, are tabulated in Tables 5.15 and 5.16. The LHC reconstruction efficiency at a given point in phase-space is taken to be proportional to the yield in simulation, as the simulated decays were generated with a uniform distribution over phase space. The efficiency is averaged over the LL and DD categories in the calculation.

A systematic uncertainty due to employing the measured  $c_i$  and  $s_i$  directly in the fit is assigned by generating a large number of toy data sets where the signal yields are calculated using  $(c_i^{\text{eff}}, s_i^{\text{eff}})$ , and then fitting the data sets using  $(c_i^{\text{model}}, s_i^{\text{model}})$ . The mean bias of each observable in these toys is assigned as the systematic uncertainty, and is determined to be  $0.1 \times 10^{-2}$  or less for all observables. The smallness of the effect is the reason no effort is made to correct the  $c_i$  and  $s_i$  values in the nominal measurement.

### 5.6.3 Mass shapes

A number of uncertainties relate to the mass distributions that enter the fit model. Each is described in detail the sections below.

**Table 5.15:** The  $c_i$  and  $s_i$  values for  $D \rightarrow K_S^0 \pi^+ \pi^-$  decays calculated via the 2018 Belle model [?] in two cases: assuming a uniform reconstruction efficiency over phase space, denoted  $(c/s)_i^{\text{model}}$ , and including the LHCb efficiency profile as obtained in simulation, averaged for LL and DD, denoted  $(c/s)_i^{\text{eff}}$ . The change due to including the efficiency is also tabulated.

Bin	$c_i^{\text{model}}$	$c_i^{\text{eff}}$	$\Delta c_i$	$s_i^{\text{model}}$	$s_i^{\text{eff}}$	$\Delta s_i$
1	-0.027	-0.007	0.019	0.812	0.794	-0.018
2	0.837	0.859	0.022	0.164	0.152	-0.012
3	0.163	0.163	-0.000	0.872	0.880	0.008
4	-0.914	-0.915	-0.001	0.076	0.082	0.006
5	-0.149	-0.170	-0.021	-0.856	-0.854	0.002
6	0.373	0.362	-0.011	-0.782	-0.805	-0.023
7	0.863	0.862	-0.000	-0.203	-0.202	0.002
8	0.860	0.862	0.002	0.330	0.336	0.006

**Table 5.16:** The  $c_i$  and  $s_i$  values for  $D \rightarrow K_S^0 K^+ K^-$  decays calculated via the 2010 BaBar model [?] in two cases: assuming a uniform reconstruction efficiency over phase space, denoted  $(c/s)_i^{\text{model}}$ , and including the LHCb efficiency profile as obtained in simulation, averaged for LL and DD, denoted  $(c/s)_i^{\text{eff}}$ . The change due to including the efficiency is also tabulated.

Bin	$c_i^{\text{model}}$	$c_i^{\text{eff}}$	$\Delta c_i$	$s_i^{\text{model}}$	$s_i^{\text{eff}}$	$\Delta s_i$
1	0.738	0.735	-0.002	0.266	0.263	-0.003
2	-0.697	-0.744	-0.046	0.332	0.329	-0.003

## 2808 Determination of shape parameters

2809 The statistical uncertainties on the shape parameters that are obtained in fits  
 2810 to simulated decays and in the first stage fit to data need to be propagated to  
 2811 the uncertainty on the obtained parameters. This is done via a bootstrapping  
 2812 procedure, repeating these steps many times:

- 2813 • Each of the data sets used determine parameters of the signal, crossfeed, and  
 2814 lowmass shapes that are fixed in the first-stage fit to data of Section 5.4 are  
 2815 re-sampled with replacement, drawing a number of events equal to the original  
 2816 data-set size. These are from simulation for signal and lowmass shapes, and  
 2817 real data for the crossfeed shapes. All of the shapes are fit again, on the  
 2818 re-sampled data sets.
- 2819 • The real dataset is re-sampled with replacement, drawing a number of events  
 2820 equal to the original data-set size. Then, the first-stage fit of Section 5.4 is  
 2821 repeated with the shapes obtained as described above, obtaining values for  
 2822 the remaining shape parameters.

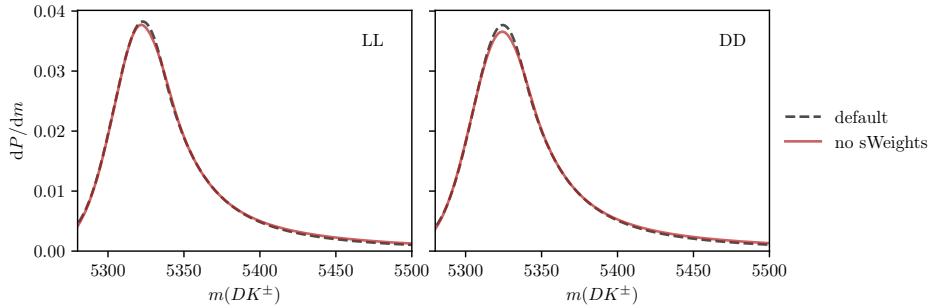
- 2823     • Finally, the  $CP$  fit is repeated using the shape parameters determined in the  
 2824       preceding steps, but *without* re-sampling the dataset (to avoid a statistical  
 2825       spread in the obtained central values that is independent of the shape  
 2826       parameters).

2827     The uncertainty on each observable is taken to be the standard deviation of the  
 2828       set of central values obtained as described above. This procedure propagates  
 2829       the statistical uncertainty on the fixed parameters to the observables, in a way  
 2830       that takes correlations into account, and which does not rely on the uncertainty  
 2831       estimates in the preliminary fits being accurate. The uncertainties are less than  
 2832        $0.1 \times 10^{-2}$  for all  $DK^\pm$  observables, in line with earlier analyses, and less than  
 2833        $0.2 \times 10^{-2}$  for all  $D\pi^\pm$  observables.

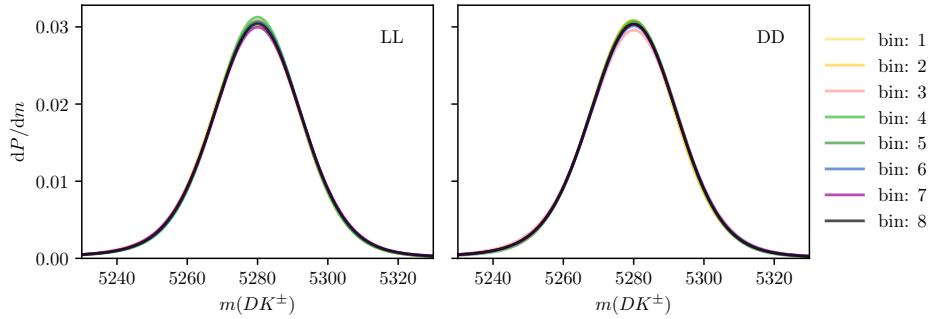
2834     A potential bias arises due the use of sWeights when obtaining the mass  
 2835       distribution of decays where a  $\pi \leftrightarrow K$  misidentification has taken place. This  
 2836       is because the  $m_{\text{swap}}(Dh^\pm)$  mass that is calculated while assuming a swapped  
 2837       companion hypothesis and the nominal  $m_{\text{default}}(Dh^\pm)$  mass are correlated (it  
 2838       is always the case that  $m_{\text{swap}} > m_{\text{default}}$  for a  $\pi \rightarrow K$  misidentification, for  
 2839       example). Thus, the assumptions of the sPlot method are not satisfied [?]. The  
 2840       correlation coefficient in the signal region is about 20% for simulated signal  
 2841       decays. In order to assess the potential impact, an alternative mass distribution  
 2842       for  $(B^\pm \rightarrow D\pi^\pm) \rightarrow (B^\pm \rightarrow DK^\pm)$  cross-feed is derived that does not rely on  
 2843       sWeights. Instead of fitting  $B^\pm \rightarrow D\pi^\pm$  sample in the whole fit range and assigning  
 2844       sWeights before recalculating the  $B$  mass under the kaon companion hypothesis,  
 2845       the shape is obtained using  $B^\pm \rightarrow D\pi^\pm$  candidates in the signal region. This is  
 2846       possible because the  $B^\pm \rightarrow D\pi^\pm$  sample is very pure. The shapes are compared in  
 2847       Fig. 5.52 and are seen to be almost identical. Thus the sWeights do successfully  
 2848       subtract the contribution of combinatorial and partially reconstructed backgrounds  
 2849       in the default setup. The impact on the obtained  $CP$ -violation observables of  
 2850       using one or the other shape in the fits is negligible, and no further systematic  
 2851       uncertainty is assigned due to this effect.

## 2852     Using the same mass shapes in all Dalitz bins

2853     The mass shapes obtained the first-stage fit where all Dalitz bins are combined,  
 2854       are used in each individual bin of the subsequent binned fit. However, there could  
 2855       be some variation in the shape over the  $D$ -decay phase space, due to correlations  
 2856       between the phase-space coordinates and particle kinematics. The potential effect is  
 2857       investigated in pseudoexperiments, where toy data sets are generated with alternative



**Figure 5.52:** Comparison of the default and no-sWeights alternative shape for the  $D\pi \rightarrow DK$  cross-feed component in the (left) LL and (right) DD categories where  $D \rightarrow K_S^0\pi^+\pi^-$ . The binned fit obtains essentially the same central values for the  $CP$ -violation observables, independently of which shape is used.

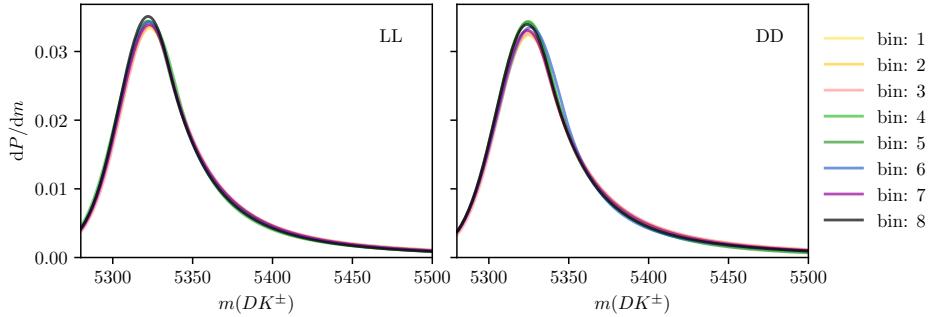


**Figure 5.53:** Signal shapes obtained in fits simulated  $B^\pm \rightarrow D\pi^\pm$  decays for individual Dalitz bins in the optimal binning scheme, for (left) LL and (right) DD candidates in the  $B^\pm \rightarrow D(\rightarrow K_S^0\pi^+\pi^-)\pi^\pm$  category.

signal, crossfeed, and combinatorial-background shapes that are allowed to differ between bins, and fitted with the default shapes. The partially reconstructed background is treated in a separate study, because further physics effects contribute to bin-by-bin variation, as described in the following section.

The alternative signal and cross-feed mass shapes are fitted independently in each bin, following identical procedures to those outlined in Sections 5.4.1 and 5.4.2. Examples of the obtained shapes are compared in Figs. 5.53 and 5.54.

The shape of the combinatorial background can also vary over the  $D$  decay phase-space; for example will the relative amount of fake  $D$  candidates versus real  $D$  decays paired with a random bachelor certainly depend on the real  $D$  decay amplitude for a given phase-space region. The effect is investigated in the high  $B$ -mass sideband  $m_B \in [5600, 6500] \text{ MeV}/c^2$ , in which the  $m(Dh^\pm)$  distribution is fitted with a single exponential distribution, in bins of the Dalitz plot. The fits combine  $B^+$  and  $B^-$  candidates and merge bins  $+i$  and  $-i$ , and are carried out for both the *optimal* binning scheme of Fig. 2.7 (on page 21) and a *rectangular*



**Figure 5.54:** Mass shapes for  $D\pi \rightarrow DK$  cross feed obtained for individual Dalitz bins in the optimal binning scheme, for (left) LL and (right) DD candidates in the  $D \rightarrow K_S^0 \pi^+ \pi^-$  category.

2873 binning scheme, shown in Fig. 5.50, which better captures continuous trends over  
 2874 the Dalitz plot. The study is done for  $D \rightarrow K_S^0 \pi^+ \pi^-$  only due to available statistics.  
 2875 The DD category of  $B^\pm \rightarrow D\pi^\pm$  decays has the largest statistics and shows the  
 2876 largest variation, and the fitted slopes for this channel are shown in Fig. 5.55. Two  
 2877 effects are visible: 1) there is some variation in the slope as a function of the Dalitz  
 2878 bin, especially visible for the rectangular scheme, and 2) the exponential slope is  
 2879 larger in general in the high  $B$ -mass sideband. The latter effect does not pose a  
 2880 problem, since the employed exponential is found to provide an excellent fit in  
 2881 the default fit region. It does however need to be taken into account when when  
 2882 deriving alternative, bin-dependent combinatorial slopes relevant for the default  
 2883 fit region. In order to do so, the alternative slope for bin  $i$  is defined

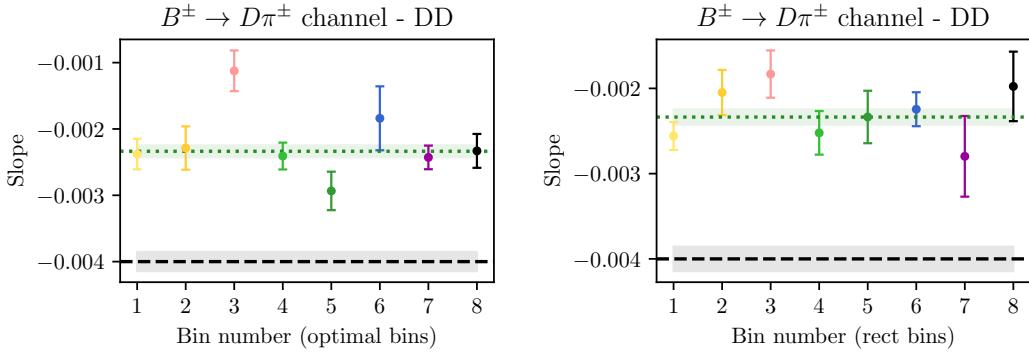
$$\alpha_{\text{default-range}}^i = \frac{\alpha_{\text{high-}m_B}^i}{\alpha_{\text{high-}m_B}^{\text{all-DP}}} \times \alpha_{\text{default-range}}^{\text{all-DP}}, \quad (5.28)$$

2884 and used when generating the combinatorial-background component of the toy  
 2885 data sets for the study.

2886 The average bias obtained for each observable in the ensemble of pseudo-  
 2887 experiments is assigned as a systematic uncertainty, found to be about  $0.1 \times 10^{-2}$   
 2888 for each observable.

### 2889 Ignoring physics effects in the lowmass background

2890 In the  $CP$  fit, the same relative fractions of partly reconstructed  $B^\pm$  and  $B^0$   
 2891 backgrounds are used in each bin, as determined in the first-stage fit described  
 2892 in Section 5.4 (whereas the partly reconstructed  $B_s^0 \rightarrow \bar{D}^0[\pi^+]K^-$  background is  
 2893 treated separately). However the distribution over the Dalitz plot depends on  
 2894 whether the partly reconstructed decays occur via an intermediate  $D^0$  meson, a  $\bar{D}^0$



**Figure 5.55:** Plot of (dots) combinatorial slope in the high  $B$  mass background for each bin in the (left) the optimal binning scheme and (right) the rectangular binning scheme, for the DD  $B^\pm \rightarrow D(\rightarrow K_S^0 \pi^+ \pi^-) \pi^\pm$  category. The slope when all bins are combined (green, dashed line) is also shown, and compared with (black dashed line) the slope in the default fit region.

meson, or and admixture of both. Consider a decay reconstructed as  $B^- \rightarrow DK^-$  but which is actually a partially-reconstructed background. There are then four types of background that should be considered:

- Decays in which the  $D$ -meson in the true decay is a  $D^0$ -meson. An example of this is  $B^- \rightarrow D^{*0}(D^0\pi^0)\pi^-$  for which the  $\pi^0$  from the  $D^{*0}$  decay is missed and the  $\pi^-$  is misidentified as the companion  $K^-$ . These are denoted ' $D^0$ -like'.<sup>7</sup>
- Decays in which the  $D$ -meson in the true decay is a  $\bar{D}^0$ -meson. An example of this is  $B_s^0 \rightarrow \bar{D}^0\pi^+K^-$  for which the  $\pi^+$  is missed and the  $K^-$  is reconstructed as the companion  $K^-$ . These are denoted ' $\bar{D}^0$ -like'.
- Decays in which the  $D$  meson in the true decay can be either flavour, and both  $D$  flavours contribute to the decay amplitude. An example of this is  $B^- \rightarrow D^*K^-$  for which the total decay amplitude into a  $D$  final state has contributions from both  $D^{*0}$  (decaying to  $D^0$ ) and  $\bar{D}^{*0}$  (decaying to  $\bar{D}^0$ ). The relative amplitude magnitude and phase between the two possible  $B$  decays are denoted  $r_B^{D^*}$  and  $\delta_B^{D^*}$  respectively. These are denoted ' $r_B$ -like'.
- Decays that can be reconstructed as both  $D^0$ - and  $\bar{D}^0$ -like but where there is no quantum-mechanical interference. An example is  $\bar{B}^0 \rightarrow D^0\pi^+\pi^-$  decays where either the  $\pi^+$  or  $\pi^-$  can be reconstructed as the bachelor. These are denoted 50/50  $D^0$ -like and  $\bar{D}^0$ -like.

<sup>7</sup>The naming convention is defined in terms of the  $D$  present in candidates reconstructed as  $B^-$  decays. For the charge conjugate case this decay would of course happen via a  $\bar{D}^0$ , but is still denoted  $D^0$ -like.

2914 For  $B^+ \rightarrow DK^+$  decays everything is CP conjugated. The Dalitz-plot distribution  
2915 for each of these cases is:

- 2916 •  $D^0$  decays (' $D^0$ -like')

$$\begin{aligned} N_{\pm i}(B^-) &\propto F_{\pm i} \\ N_{\pm i}(B^+) &\propto F_{\mp i} \end{aligned} \quad (5.29)$$

- 2917 •  $\bar{D}^0$  decays (' $\bar{D}^0$ -like'):

$$\begin{aligned} N_{\pm i}(B^-) &\propto F_{\mp i} \\ N_{\pm i}(B^+) &\propto F_{\pm i} \end{aligned} \quad (5.30)$$

- 2918 • Decays with a quantum-mechanical admixture of  $D^0$  and  $\bar{D}^0$  (' $r_B$ -like'):

$$\begin{aligned} N_{\pm i}(B^-) &\propto F_{\pm i} + (r_B^*)^2 F_{\mp i} + 2\sqrt{F_{+i}F_{-i}}[x_-^* c_{\pm i} + y_-^* s_{\pm i}] \\ N_{\pm i}(B^+) &\propto F_{\mp i} + (r_B^*)^2 F_{\pm i} + 2\sqrt{F_{+i}F_{-i}}[x_+^* c_{\pm i} - y_+^* s_{\pm i}] \end{aligned} \quad (5.31)$$

2919 where  $(x_\pm^*, y_\pm^*)$  are defined analogously to the standard  $B^\pm \rightarrow DK^\pm$  case.

- 2920 • 50/50  $D^0$ -like and  $\bar{D}^0$ -like:

$$\begin{aligned} N_{\pm i}(B^-) &\propto F_{\pm i} + F_{\mp i} \\ N_{\pm i}(B^+) &\propto F_{\pm i} + F_{\mp i} \end{aligned} \quad (5.32)$$

2921 The use of a single background shape across all bins may therefore introduce biases  
2922 because, if an admixture of these bac nds is present, such a shape has no  
2923 sensitivity to bin-to-bin variations.

2924 In the  $D\pi$  channel, the dominant backgrounds are all  $D^0$ -like ( $\bar{B}^0 \rightarrow D^{*-}\pi^0$ ,  
2925  $B^- \rightarrow D^0\rho^-$ ,  $B^- \rightarrow D^{*0}\pi^-$ ). There is a small contribution from  $\bar{B}^0 \rightarrow D^0\rho(\rightarrow$   
2926  $\pi^+\pi^-)$  decays where either the  $\pi^+$  or  $\pi^-$  from the  $\rho^0$  decay can be assigned as the  
2927 bachelor, and thus this background is 50/50  $D^0$ -like and  $\bar{D}^0$ -like. The background  
2928 only corresponds to about 0.5 % of the total partially reconstructed background and  
2929 thus the impact is small. Nevertheless it is considered in the study described below.

2930 In the  $DK$  channel all categories of background appear. In the mass region  
2931 of the  $CP$  fit approximately 75.5% of backgrounds are  $D^0$ -like ( $\bar{B}^0 \rightarrow D^{*-}K^-$ ,  
2932 mis-identified  $B^- \rightarrow D^{*0}\pi^-$ , and mis-identified  $B^- \rightarrow D^0\rho^-$ ), 7.5 % are  $\bar{D}^0$ -like  
2933 ( $B_s^0 \rightarrow \bar{D}^0\pi^+K^-$ ), 1 % is 50/50  $D^0$ - $\bar{D}^0$ -like (mis-identified  $B^0 \rightarrow D\rho^0$ ), and 16% are  
2934  $r_B$ -like ( $B^- \rightarrow D^*K^-$ ,  $B^0 \rightarrow DK^{*0}$ , and  $B^- \rightarrow DK^{*-}$ ).

2935 In order to estimate the bias due to ignoring this effect, a large number of toy  
2936 data sets are generated using the default low mass shapes and total yields from

2937 the first-stage fit in Section 5.4, but distributing each of them individually over  
2938 the Dalitz-bins according to Eqs. (5.29)-(5.31). When calculating the distribution  
2939 of  $B^+ \rightarrow D^{*0}K^+$  decays over the Dalitz plot, the values [?]

$$r_B^{D^*} = 0.191 \quad \delta_B^{D^*} = 331.6^\circ \quad (5.33)$$

2940 are used. When calculating the distribution of  $B^+ \rightarrow D^0K^{*+}$  decays over the  
2941 Dalitz plot the values [?]

$$r_B^{K^*} = 0.092 \quad \delta_B^{K^*} = 40^\circ. \quad (5.34)$$

2942 are used. The toy data sets are then fit with the default set up, and the observed  
2943 mean bias assigned as the corresponding uncertainty. The corresponding uncer-  
2944 tainties were found to be about  $0.1 \times 10^{-2}$  for all uncertainties. The variation in  
2945 the shapes is rather small in the mass range included in the fit, which explains  
2946 the small impact.

2947 If the  $B_s^0$  background is *not* treated separately in the default fit, but instead  
2948 included in a single lowmass background shape along with the  $B^0$  and  $B^\pm$  contribu-  
2949 tions, the systematic uncertainty is an order of magnitude larger when evaluated  
2950 as described above, and would be a dominating systematic. This motivates the  
2951 separate treatment of the  $B_s^0$  background.

#### 2952 5.6.4 $CP$ violation and material interaction of the $K_S^0$

2953 A systematic uncertainty due to  $CP$ -violation effects and material interaction of  
2954 the  $K_S^0$  is assign  using the results obtained in Section 4.3.7. In that section,  
2955 the expected bias on all observables in a combined  $B^\pm \rightarrow Dh^\pm$  measurement  
2956 was evaluated for the detector geometry and particle kinematics of the LHCb  
2957 experiment. The calculation was made for  $(r_B^{DK^\pm/D\pi^\pm}, \delta_B^{DK^\pm/D\pi^\pm})$  values close to  
2958 the world averages, and a number of  $\gamma$  values; the results were summarised in Fig. 4.8.  
2959 The systematic uncertainty is taken to be the largest absolute bias observed for each  
2960 parameter in the study. The largest uncertainty (on  $y_\xi^{D\pi}$  where it is  $0.46 \times 10^{-2}$ ) is  
2961 still an order of magnitude smaller than the statistical uncertainty.

#### 2962 5.6.5 Impact of $D$ mixing

2963 The effect of  $D$ -mixing is not accounted for in the measurement, which leads to  
2964 a small bias. Earlier studies have shown this to lead to a sub-degree bias on  
2965 measurements of  $\gamma$  in  $B^\pm \rightarrow DK^\pm$  decays, in the case where the  $F_i$  parameters  
2966 are determined experimentally under the same experimental conditions as the  $\gamma$

measurement [?]. A number of pseudoexperiments are carried out to verify that this is also the case for the combined  $DK^\pm - D\pi^\pm$  setup employed in the thesis. They are performed following the same procedure described in Section 5.6.4 for the case of  $K_S^0$   $CP$  violation. The yields are calculated while taking  $D$  mixing into account, using the mixing parameter values  $x = (0.39^{+0.11}_{-0.12})\%$  and  $y = (0.65^{+0.06}_{-0.07})\%$  [?], and then fitted back assuming no  $D$  mixing. The biases are found to be small, as expected, all of them smaller than  $0.05 \times 10^{-2}$ . The largest relative biases are on the  $B^\pm \rightarrow D\pi^\pm$  parameters, but even for those the relative effect is less than 2 %. In agreement with Ref. [?], it is found that the biases increase with an order of magnitude if the  $F_i$  parameters are fixed to the expected values with no  $D$ -mixing, instead of being determined as part of the fit.

### 5.6.6 PID efficiencies

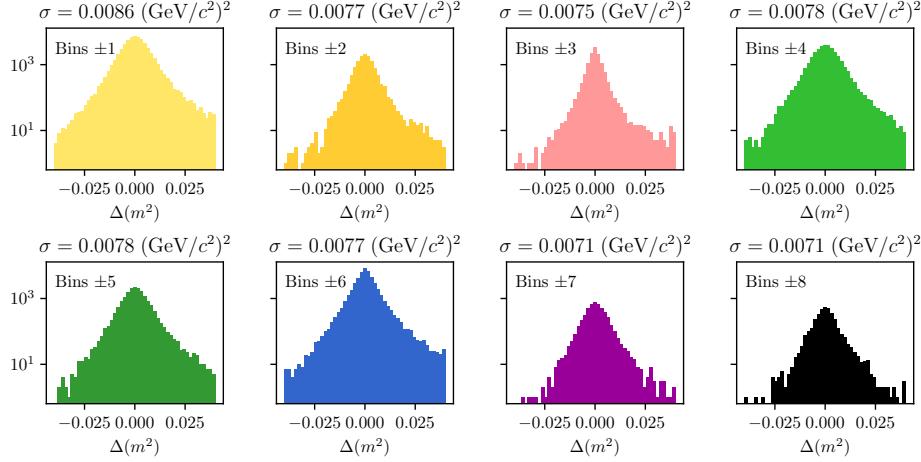
The uncertainty related to PID efficiencies is assessed by repeating the full two-stage fit procedure a number of times, each time varying the PID efficiencies within the uncertainties. The used uncertainty includes both a statistical and systematic component, as described in detail in Section 5.1.3. The standard deviations of the central values obtained for each observable are assigned as the systematic uncertainty. The uncertainties come out below  $0.1 \times 10^{-2}$  for all observables.

### 5.6.7 Dalitz-coordinate resolution

There is a small systematic uncertainty related to Dalitz-plot-bin migration, where the non-perfect resolution on the momentum measurement means that a candidate is assigned to a different bin than it truly belongs to. This leads to non-negligible net migration between bins that share a border in a region of phase space where the amplitude varies rapidly. However, since the  $F_i$  are measured in the data set, all leading order effects of migration are inherently taken into account. The measurement is only sensitive to differences in migration between the  $DK^\pm$  and  $D\pi^\pm$  channels and the effect is small.

The systematic uncertainty due to this effect is assigned using pseudoexperiments. The study is made for the  $D \rightarrow K_S^0 \pi^+ \pi^-$  mode only, which is sufficient since it completely dominates the overall sensitivity.

1. Signal  $B^\pm \rightarrow DK^\pm$  and  $B^\pm \rightarrow D\pi^\pm$  decays are generated continuously over phase space, according to the expected distribution obtained with the latest amplitude model from the Belle collaboration [?], assuming values of  $\gamma$  and  $(r_B^{DK^\pm/D\pi^\pm}, \delta_B^{DK^\pm/D\pi^\pm})$  close to the current world averages.



**Figure 5.56:** Distribution of the biases  $\Delta(m^2) = m_{true}^2 - m_{reconstructed}^2$  in simulation for  $m^2(K_S^0\pi^+)$  in bins of the rectangular binning scheme of Fig. 5.50.

- 3001     2. The Dalitz coordinates of each candidate are then smeared using the experi-  
 3002       ment resolution obtained in simulation. This is described further below.  
  
 3003     3. Finally, the generated candidates are binned and fitted back using the default  
 3004       setup.

3005       The resolution is obtained via simulation, by comparing the reconstructed  
 3006       phase-space coordinates with those calculated from the true momenta in samples of  
 3007       simulated  $D \rightarrow K_S^0\pi^+\pi^-$  decays. As can be seen in Fig. 5.56, the resolution is found  
 3008       to vary over phase space and the distribution of shifts has significant exponential  
 3009       tails. In order to take both effects into account, the smearing is done by shifting each  
 3010       generated decay with a realised coordinate shift in full simulation, for a simulated  
 3011       decay that took place at approximately the same place in the Dalitz plot. The  
 3012       shift is multiplied with 120 % to take into account that the resolution is generally  
 3013       better in simulation than data. If the shift results in Dalitz coordinates outside the  
 3014       kinematically allowed region, a different shift is applied randomly instead.

3015       The average bias seen in the pseudoexperiments is assigned as the systematic  
 3016       uncert. The uncertainties come out at about  $(0.1-0.2)\times 10^{-2}$  for all parameters.  
 3017       It is noted that for all four  $DK^\pm$  parameters the bias is towards a smaller value  
 3018       of  $r_B^{DK}$ ; this is to be expected, as bin migration washes out the asymmetries in  
 3019       different areas of the Dalitz plot.

**Table 5.17:** Summary of systematic uncertainties due to backgrounds that are potentially present with a small yield, but not included in the mass fit.

All uncertainties are quoted with implicit: $\times 10^{-2}$						
Mode	$\sigma(x_-^{DK^\pm})$	$\sigma(y_-^{DK^\pm})$	$\sigma(x_+^{DK^\pm})$	$\sigma(y_+^{DK^\pm})$	$\sigma(x_\xi^{D\pi^\pm})$	$\sigma(y_\xi^{D\pi^\pm})$
$\Lambda_b^0$ backgrounds	0.04	0.05	0.04	0.06	0.08	0.13
$B \rightarrow D\mu\nu X$	0.04	0.07	0.04	0.05	0.10	0.11
$B^\pm \rightarrow D(\rightarrow K_S^0\pi\mu\nu)h^\pm$	0.00	0.03	0.02	0.02	0.00	0.00
Swapped tracks	0.10	0.13	0.12	0.08	0.00	0.01
Total	0.11	0.16	0.13	0.12	0.08	0.13

### 3020 5.6.8 The fixed yield fractions

3021 A number of relative yields are fixed from efficiencies in simulation and branching  
 3022 fractions. In the  $DK^\pm$  modes, this is the case for all the relative yields of the  
 3023 partially reconstructed backgrounds, including partially reconstructed  $B \rightarrow D\pi[X]$   
 3024 decays where the pion is misidentified as a kaon, and the yield of the  $B_s^0 \rightarrow DK\pi$   
 3025 background relative to the  $B \rightarrow D\pi$  yield. In the  $B^\pm \rightarrow D\pi^\pm$  channel, the only fixed  
 3026 yield ratio is that of the partially reconstructed  $B^\pm \rightarrow D^{*0}(\rightarrow D\pi^0)\pi^\pm$  and  $B^0 \rightarrow$   
 3027  $D^{*\mp}(\rightarrow D\pi^\mp)\pi^\pm$  modes. The uncertainty on the observables due to uncertainties  
 3028 on these fixed fractions is assessed by repeating the two-stage fit procedure many  
 3029 times, each time shifting the yield ratios randomly within their uncertainties. The  
 3030 uncertainty on each observable is taken to be the standard deviation of the set of  
 3031 central values obtained in these fits. These are all smaller than  $0.1 \times 10^{-2}$ .

### 3032 5.6.9 Systematic uncertainty due to backgrounds that are 3033 not modelled in fit

3034 There are a number of backgrounds that are expected to be present at a small  
 3035 level, but which are not modelled in the fits to data because their impact on the fit  
 3036 results is minimal. Instead, a systematic uncertainty is assigned. Each contribution  
 3037 is described in the following sections and the related systematic uncertainties  
 3038 are summarised in Table 5.17.

#### 3039 Background from $\Lambda_b$ decays

3040 This section considers the possible impact of the two potential backgrounds from  
 3041  $\Lambda_b^0$  decays described in Section 5.4.3:  $\Lambda_b^0 \rightarrow D^0 p\pi^-$  decays where the pion is not  
 3042 included in the candidate reconstruction and the proton assigned as the companion, and  
 3043  $\Lambda_b^0 \rightarrow \Lambda_c^+(\rightarrow pK_S^0\pi^+\pi^-)\pi^-$  decays where a pion in the  $\Lambda_c^+$  decay is not reconstructed  
 3044 and the proton reconstructed as one of the  $D$  decay products. The impact of not

including these in the default fit is assessed by generating toy data sets where the backgrounds are included in the generation step, which are then fitted back with the default model. The former background is distributed over the Dalitz plot as  $\bar{D}^0$ -like, cf. the terminology of Section 5.6.3, since a positive bachelor is produced along with a  $D^0$  meson. The latter is also distributed as  $\bar{D}^0$ -like in the study; the exact distribution is unknown, but a  $\bar{D}^0$ -like background is likely to have the largest effect and thus this is a conservative choice. The total yields are taken relative to the signal yields, using the yield ratios discussed in Section 5.4.3. The  $m(Dh^\pm)$  distributions are obtained using simulated samples, produced with `RapidSim`. The mean biases come out to be less than  $0.1 \times 10^{-2}$  for each  $CP$ -violation observable, which is assigned as a systematic uncertainty.

### 3056 Semi-leptonic backgrounds

3057 The impact of remnant  $B \rightarrow D\mu\nu_\mu$  decays after requiring `isMuon=0` on the bachelor  
 3058 is assessed in pseudoexperiments. Toy datasets are generated where the background  
 3059 is added in the generation step, which are then fitted with the default model. The  
 3060 background yield relative to signal and the mass shape are obtained from a sample  
 3061 of fully simulated decays for conditions corresponding to the run conditions in  
 3062 2012. The obtained bias in the toys is assigned as the systematic uncertainties:  
 3063 it is below  $0.1 \times 10^{-2}$  for all parameters.

3064 The systematic uncertainty relating to the presence of  $D^0 \rightarrow K_S^0\pi^-\mu^+\nu_\mu$  is  
 3065 estimated by repeating the bias studies of Section 5.3.3, but scaling the background  
 3066 yields to 10% to take into account the lepton veto on the  $D$  decay products. All  
 3067 biases are less than  $0.05 \times 10^{-2}$  in this case.

### 3068 Swapped tracks

3069 There is a peaking background present from  $B^\pm \rightarrow D(\rightarrow K^\mp\pi^\pm)K_S^0\pi^\pm$  decays where  
 3070 the kaon is reconstructed as the companion and the  $K_S^0$  is assigned to the  $D$  decay.  
 3071 The yield of this background is determined to be 0.5% of the signal yield in the  
 3072  $B^\pm \rightarrow DK^\pm$  channel in Section 5.3.5. The potential impact from the presence  
 3073 of the background is estimated by

- 3074 1. Calculating the expected  $B^\pm \rightarrow D\pi^\pm$  and  $B^\pm \rightarrow DK^\pm$  signal yields in each  
 3075 bin for physics parameters similar to the world average values.

**Table 5.18:** Biases observed with alternative input parameters and the systematic uncertainty assigned for the bias correction. All numbers are quoted with an implicit  $\times 10^{-2}$ .

Input ( $\gamma, r_B^{DK^\pm}, \delta_B^{DK^\pm}, r_B^{D\pi^\pm}, \delta_B^{D\pi^\pm}$ )	$x_-^{DK^\pm}$	$y_-^{DK^\pm}$	$x_+^{DK^\pm}$	$y_+^{DK^\pm}$	$x_\xi^{D\pi^\pm}$	$y_\xi^{D\pi^\pm}$
(72, 0.080, 117, 0.005, 288)	-0.02	-0.01	-0.02	-0.02	0.03	0.00
(75, 0.100, 130, 0.005, 300)	-0.03	-0.04	-0.00	0.02	0.01	-0.03
(82, 0.112, 144, 0.005, 330)	0.00	-0.01	0.00	0.03	-0.03	0.02
(71, 0.099, 129, 0.005, 300)	-0.02	-0.04	-0.00	-0.00	0.05	-0.00
Syst. uncertainty	0.04	0.03	0.02	0.04	0.09	0.05

- 3076 2. Then calculating the background bin yields in each  $B^\pm \rightarrow DK^\pm$  bin, using  
 3077 a total yield equal to 0.5% of the signal yield, and the bin distribution  
 3078 from simulated samples of  $B^\pm \rightarrow D(\rightarrow K^\mp\pi^\pm)K_S^0\pi^\pm$  decays, produced via  
 3079 `RapidSim`. The study is carried out for multiple simulated samples, including  
 3080 decays where the  $K_S^0\pi$  pair in the  $B$  decay originate in different  $K^*$  resonances  
 3081 (generated with `EvtGen` and the proper resonance-spin models), as well as  $B$   
 3082 decays that are evenly distributed over the allowed phase space.
- 3083 3. For each sample, the signal and background yields are added, and the new  
 3084  $B^\pm \rightarrow D\pi^\pm$  and  $B^\pm \rightarrow DK^\pm$  yields are fitted back with the default signal-  
 3085 yield expressions (including a fit of the  $F_i$  parameters).
- 3086 For each parameter, the most significant bias seen across the different `RapidSim`  
 3087 samples is taken as the related systematic uncertainty. The uncertainty is below  
 3088  $0.15 \times 10^{-2}$  for all parameters.

### 3089 5.6.10 Bias correction

3090 In the default sensitivity study, the bias was found to be compatible with zero.  
 3091 However, the size of a potential bias can vary depending on the input parame-  
 3092 ters. The size of the bias has been investigated with alternate input values of  
 3093  $(\gamma, r_B^{DK^\pm}, \delta_B^{DK^\pm}, r_B^{D\pi^\pm}, \delta_B^{D\pi^\pm})$ , obtaining the results in Table 5.18. A systematic  
 3094 uncertainty due to a potential, small bias is calculated as the difference between  
 3095 the maximum and minimum bias for a given parameter. The uncertainty assigned  
 3096 in this way is very small in general, and less than  $0.1 \times 10^{-2}$  for all parameters.

3097 **5.6.11 Charmless backgrounds**

3098 As discussed in Section 5.3.1, a small number of charmless background decays  
3099 survive the  $D$  flight distance cut. In this section the systematic uncertainty related  
3100 to those is assessed, in a series of pseudoexperiments. Toy datasets are generated,  
3101 where a charmless background component is included, using the yields and shapes  
3102 obtained in the studies of Section 5.3.1. The Dalitz-bin distribution is obtained  
3103 by repeating the fits of that section for each bin individually. These datasets are  
3104 subsequently fitted back using the default model, which does not include a charmless  
3105 component. No statistically significant bias is observed.

3106 The study described above does not allow for charge-asymmetries in the charmless  
3107 backgrounds, in terms of overall yields and phase-space distributions. These  
3108 asymmetries are likely to be present, due to large local  $CP$ -violation in regions of  
3109 phase space in  $B^\pm$  decays to hadrons [?, ?]. The yields in the data-driven studies  
3110 of Section 5.3.1 are not large enough to assess asymmetries, let alone asymmetric  
3111 bin distributions with any degree of statistical precision. Instead, an extreme-case  
3112 scenario is investigated, where *all* the charmless background is added to either the  $B^+$   
3113 or  $B^-$  data sample in generation. In both cases, no statistically significant biases are  
3114 observed, and it is concluded that the impact of charmless background is negligible.

3115 **5.6.12 Summary of systematic uncertainties**

3116 The complete set of included systematic uncertainties are summarised in Table 5.19.  
3117 It can be seen that the measurement is statistically limited. The correlation matrix  
3118 pertaining to the LHCb related systematics is given in Table 5.20. For studies where  
3119 the systematic uncertainty is obtained by repeating fits to data multiple times while  
3120 varying some input, the correlation matrix from the correlations of the fitted central  
3121 values. For studies that are based on generating a large number of toy datasets and  
3122 determining the average bias, the correlation of a systematic on two observables  
3123 is taken to be +100 % if the biases are in the same direction, and -100 % if they  
3124 are in opposite directions. The total systematic correlation matrix, including both  
3125 LHCb-related systematics and that of the strong-phase inputs, is given in Table 5.21.

3126 The studies described in this section also allow for an estimate of the systematic  
3127 uncertainties on the  $\mathcal{R}_i$  parameters of Eq. (5.23) or, equivalently the  $F_i$  parameters,  
3128 in a completely analogous manner to how the uncertainty on the  $CP$ -violation  
3129 observables was assigned. In all cases, however, the systematic uncertainty is found  
3130 to be much smaller than the statistical uncertainties that were given in Table 5.10.  
3131 The central values, statistical, and systematic uncertainties of the  $\mathcal{R}_i$  parameters  
3132 have been made public in Ref. [?] because they can be employed in future LHCb  
3133 measurements, as discussed in Section 5.5.2.

**Table 5.19:** Overview of all sources of uncertainty on the measurement.

Source	All uncertainties are quoted with implicit: $\times 10^{-2}$					
	$\sigma(x_-^{DK\pm})$	$\sigma(y_-^{DK\pm})$	$\sigma(x_+^{DK\pm})$	$\sigma(y_+^{DK\pm})$	$\sigma(x_\xi^{D\pi\pm})$	$\sigma(y_\xi^{D\pi\pm})$
Statistical	0.96	1.14	0.96	1.20	1.99	2.34
Strong-Phase inputs	0.23	0.35	0.18	0.28	0.14	0.18
Efficiency correction of ( $c_i, s_i$ )	0.11	0.05	0.05	0.10	0.08	0.09
Mass-shape parameters	0.05	0.08	0.03	0.08	0.16	0.17
Mass-shape bin dependence	0.05	0.07	0.04	0.08	0.07	0.09
Lowmass physics effects	0.04	0.10	0.15	0.05	0.10	0.09
$CP$ violation of $K_S^0$	0.03	0.04	0.08	0.08	0.09	0.46
$D$ mixing	0.04	0.01	0.00	0.02	0.02	0.01
PID efficiencies	0.03	0.03	0.01	0.05	0.02	0.02
Fixed yield ratios	0.05	0.06	0.03	0.06	0.02	0.02
Dalitz-bin migration	0.04	0.08	0.08	0.11	0.18	0.10
Bias correction	0.04	0.03	0.02	0.04	0.09	0.05
Small backgrounds	0.11	0.16	0.13	0.12	0.08	0.13
Total LHCb systematic	0.20	0.25	0.24	0.26	0.32	0.54
Total systematic	0.31	0.43	0.30	0.38	0.35	0.57

**Table 5.20:** Total LHCb-related systematic uncertainties and their correlation matrix.

Uncertainty ( $\times 10^{-2}$ )						
	$x_-^{DK\pm}$	$y_-^{DK\pm}$	$x_+^{DK\pm}$	$y_+^{DK\pm}$	$x_\xi^{D\pi\pm}$	$y_\xi^{D\pi\pm}$
$\sigma$	0.20	0.25	0.24	0.26	0.32	0.54

Correlations						
	$x_-^{DK\pm}$	$y_-^{DK\pm}$	$x_+^{DK\pm}$	$y_+^{DK\pm}$	$x_\xi^{D\pi\pm}$	$y_\xi^{D\pi\pm}$
$x_-^{DK\pm}$	1.000	0.864	0.734	0.897	0.349	0.318
$y_-^{DK\pm}$		1.000	0.874	0.903	0.408	0.362
$x_+^{DK\pm}$			1.000	0.771	0.563	0.447
$y_+^{DK\pm}$				1.000	0.507	0.451
$x_\xi^{D\pi\pm}$					1.000	0.484
$y_\xi^{D\pi\pm}$						1.000

## 5.7 Obtained constraints on $\gamma$

<sup>3134</sup> The measured values of  $(x_\pm^{DK}, y_\pm^{DK}, x_\xi^{D\pi}, y_\xi^{D\pi})$  can be used to put constraints on the possible values of the CKM angle  $\gamma$  and the hadronic nuisance parameters  $r_B^{DK\pm}$ ,  $\delta_B^{DK\pm}$ ,  $r_B^{D\pi\pm}$ , and  $\delta_B^{D\pi\pm}$ . This is handled using the `gammacombo` package, which is also used to combine all measurements of  $\gamma$  made by the LHCb collaboration [?, ?].

**Table 5.21:** Total systematic uncertainties and their correlation matrix, including contributions due to strong-phase inputs as well as LHCb-related uncertainties.

Uncertainty ( $\times 10^{-2}$ )						
	$x_-^{DK\pm}$	$y_-^{DK\pm}$	$x_+^{DK\pm}$	$y_+^{DK\pm}$	$x_\xi^{D\pi\pm}$	$y_\xi^{D\pi\pm}$
$\sigma$	0.31	0.43	0.30	0.38	0.35	0.57

Correlations						
	$x_-^{DK\pm}$	$y_-^{DK\pm}$	$x_+^{DK\pm}$	$y_+^{DK\pm}$	$x_\xi^{D\pi\pm}$	$y_\xi^{D\pi\pm}$
$x_-^{DK\pm}$	1.000	0.301	0.156	0.576	0.265	0.231
$y_-^{DK\pm}$		1.000	0.437	0.218	0.183	0.170
$x_+^{DK\pm}$			1.000	0.445	0.414	0.310
$y_+^{DK\pm}$				1.000	0.353	0.243
$x_\xi^{D\pi\pm}$					1.000	0.502
$y_\xi^{D\pi\pm}$						1.000

### 3139 5.7.1 Statistical approach

3140 The optimal central values determined in a maximum likelihood fit. The set of  
 3141 all observables for which a measurement has been made is denoted  $A$ , and the set  
 3142 of underlying physics parameters is denoted  $\theta$ . The physics parameters of course  
 3143 determine the probability density function of measurement results of  $A$ ,  $f(A|\theta)$ .  
 3144 Given a specific set of measurement results,  $A_{\text{obs}}$ , a likelihood function is defined

$$\mathcal{L}(\theta|A_{\text{obs}}) = f(A_{\text{obs}}|\theta) \quad (5.35)$$

3145 and the estimate of  $\theta$  is the set of parameters that maximize the likelihood

$$\hat{\theta} = \arg \max_{\theta} \mathcal{L}(\theta|A_{\text{obs}}). \quad (5.36)$$

3146 In practice, a  $\chi^2$  function is defined

$$\chi^2(\theta|A_{\text{obs}}) = -2 \ln \mathcal{L}(\theta|A_{\text{obs}}) \quad (5.37)$$

3147 and minimized instead. In the specific case where the likelihood profile is Gaussian,  
 3148 it is given by the simple expression 

$$\chi^2(\theta|A_{\text{obs}}) = (A_{\text{obs}} - A(\theta))^T \Sigma_{A_{\text{obs}}}^{-1} (A_{\text{obs}} - A(\theta)) + c, \quad (5.38)$$

3149 where  $\Sigma_{A_{\text{obs}}}$  is the covariance matrix for the measured observables,  $A(\theta)$  denotes the  
 3150 value of the observables expressed in terms of the underlying physics parameters,

and  $c$  is a constant that is independent of  $\theta$ . In the specific case considered here

$$\begin{aligned} A &= (x_-^{DK}, y_-^{DK}, x_+^{DK}, y_+^{DK}, x_\xi^{D\pi}, y_\xi^{D\pi}) \\ \theta &= (\gamma, r_B^{DK^\pm}, \delta_B^{DK^\pm}, r_B^{D\pi^\pm}, \delta_B^{D\pi^\pm}). \end{aligned} \quad (5.39)$$

The likelihood scan presented in Section 5.5.2 proved that the Gaussian expression in Eq. (5.38) provides an excellent description of the likelihood profile of the measurement, when  $\Sigma_{A_{\text{obs}}}$  is taken to be the covariance matrix obtained in that section. Thus, the  $\chi^2$  function defined in Eq. (5.38) is minimised to determine the best estimate of  $\gamma$ .

Two different methods are employed to construct confidence regions for the observables of interest, known within the `gammacombo` framework as the PROB and PLUGIN methods. Both methods aim to construct confidence regions for some subset,  $\phi$ , of the full parameter set  $\theta$ . The remaining parameters, dubbed nuisance parameters below, are denoted  $\eta = \theta \setminus \phi$ . In practice,  $\phi$  most often denotes a single parameter, and of special interest is of course the case where  $\phi = \gamma$ . Both methods aim to solve the problem that due to the number of parameters in  $\theta$  (five in the case considered here, but up to 40 in the latest LHCb combination [?]), it is not feasible to derive the confidence regions from a full-fledged Neumann construction [?]. Under assumptions discussed below, the methods achieve reasonable coverage nonetheless, ie. had the measurement been repeated many times, the confidence region is expected to cover the true parameter(s) with a probability at least at large as the quoted confidence level (CL), independently of the true parameter value. The presentation follows the `gammacombo` manual [?].

The PROB method is a simple profile-likelihood method. The minimum value of the  $\chi^2$  function is denoted  $\chi^2_{\min} \equiv \chi^2(\hat{\theta}|A_{\text{obs}})$ . To evaluate the CL for a specific value (set of values) of  $\phi_0$ , the  $\chi^2$  function is again minimised, this time under the constraint that  $\phi = \phi_0$ , resulting in a new minimum  $\hat{\theta}' = (\phi_0, \hat{\eta}')$ . In the approximation that all likelihoods are exactly Gaussian, the variable

$$\Delta\chi^2(\phi_0|A_{\text{obs}}) = \chi^2(\hat{\theta}'|A_{\text{obs}}) - \chi^2_{\min} \quad (5.40)$$

follows a  $\chi^2$  distribution with  $n$  degrees of freedom, where  $n$  is the number of parameters in  $\phi$  [?]. This can be used to evaluate CL at that point as

$$CL(\phi_0|A_{\text{obs}}) = F_n(\Delta\chi^2(\phi_0|A_{\text{obs}})) \quad (5.41)$$

where  $F_n$  is the cumulative distribution function of a  $\chi^2$  distribution with  $n$  degrees of freedom. The method takes its colloquial name from the fact that this function is named Prob in the ROOT package. Confidence regions can be defined by scanning

the values of  $\phi_0$  over a region of interest. These confidence regions assume that the estimates  $\hat{\theta}$  follow a Gaussian distribution centred on the true values, which is the case for maximum likelihood estimates in asymptotically large samples [?]; in other cases they may not have good coverage properties. Given the Gaussian shape obtained in the likelihood scan of Section 5.5.2 the confidence regions are likely to be well behaved in the case considered here.

However, for the purpose of comparing to the combination of several LHCb measurements in Section 5.5.3 below, the PLUGIN method is necessary. It foregoes the assumption that  $\Delta\chi^2$  follows a  $\chi^2$  distribution, and instead estimates the distribution in a bootstrapping scheme. The procedure is as follows: the values of  $\hat{\theta}$ ,  $\hat{\theta}'$ , and  $\Delta\chi^2(\phi_0|A_{\text{obs}})$  are determined as described above; then the following steps are carried out a number,  $N_{\text{toys}}$ , of times

1. Generate a "toy" result,  $A_{\text{toy}}^i$ , following the distribution  $f(A|\hat{\theta}')$
2. Determine  $\Delta\chi^2(\phi_0|A_{\text{toy}}^i)$  by minimising the  $\chi^2$  function for the results  $A_{\text{toy}}^i$  twice, once where all parameters in  $\theta$  are free, and once where  $\phi = \phi_0$  is enforced

Then the CL is defined by

$$CL(\phi_0) = 1 - \frac{N(\Delta\chi^2(\phi_0|A_{\text{obs}}) < \Delta\chi^2(\phi_0|A_{\text{toy}}^i))}{N_{\text{toys}}}. \quad (5.42)$$

The method is described in Ref. [?], based on the hybrid resampling method presented in [?, ?]. While the coverage properties are not proven, evidence is presented in terms of asymptotic results and simulation studies in those references. The coverage properties have also been investigated in relation to LHCb combinations, and the intervals were found to perform well in most cases [?].

### 5.7.2 Interpretation results

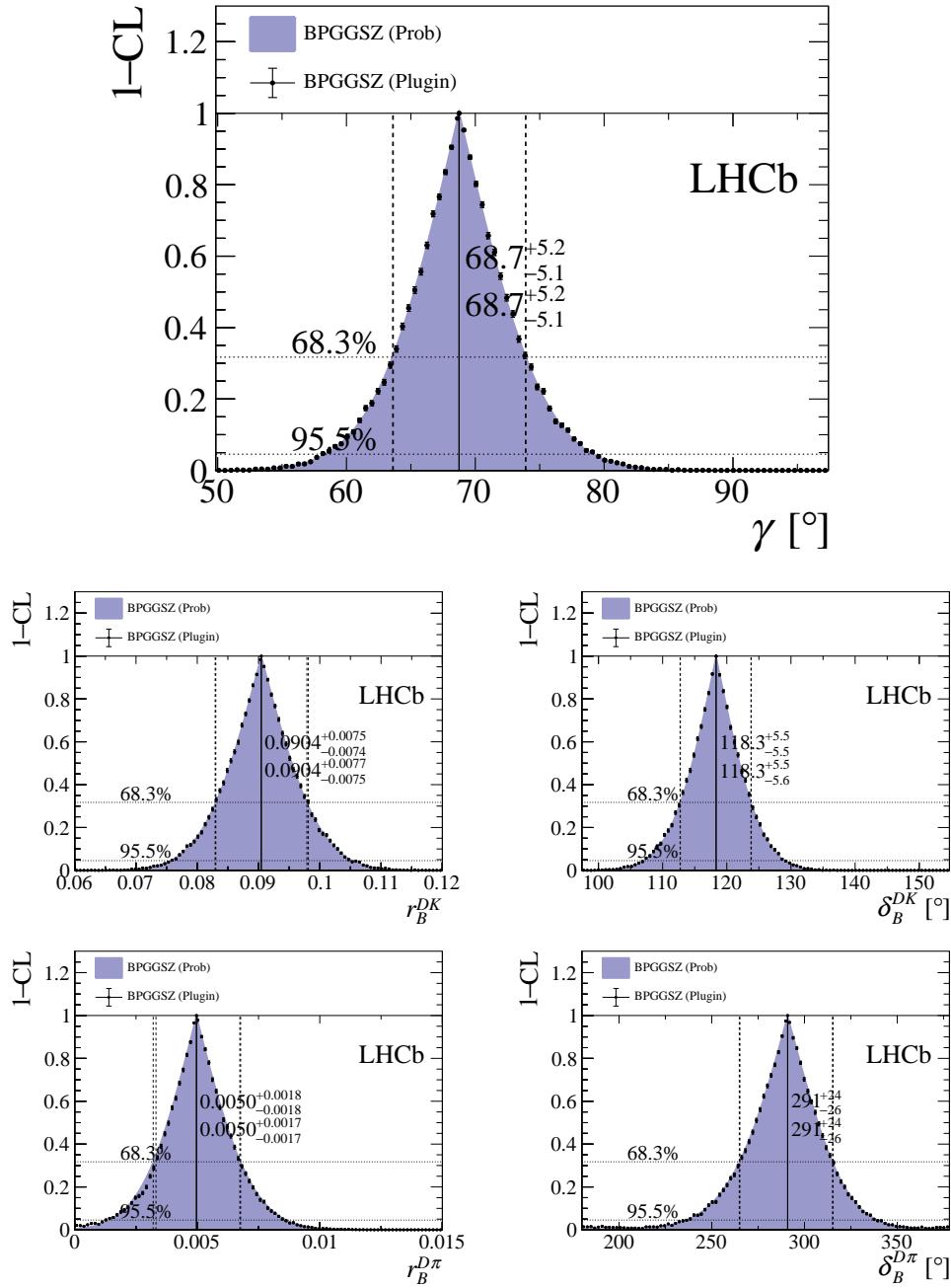
The central values and confidence regions obtained for the physics parameters are

$$\begin{aligned} \gamma &= (68.7^{+5.2}_{-5.1})^\circ, \\ r_B^{DK^\pm} &= 0.0904^{+0.0077}_{-0.0075}, \\ \delta_B^{DK^\pm} &= (118.3^{+5.5}_{-5.6})^\circ, \\ r_B^{D\pi^\pm} &= 0.0050^{+0.0017}_{-0.0017}, \\ \delta_B^{D\pi^\pm} &= (291^{+24}_{-26})^\circ, \end{aligned} \quad (5.43)$$

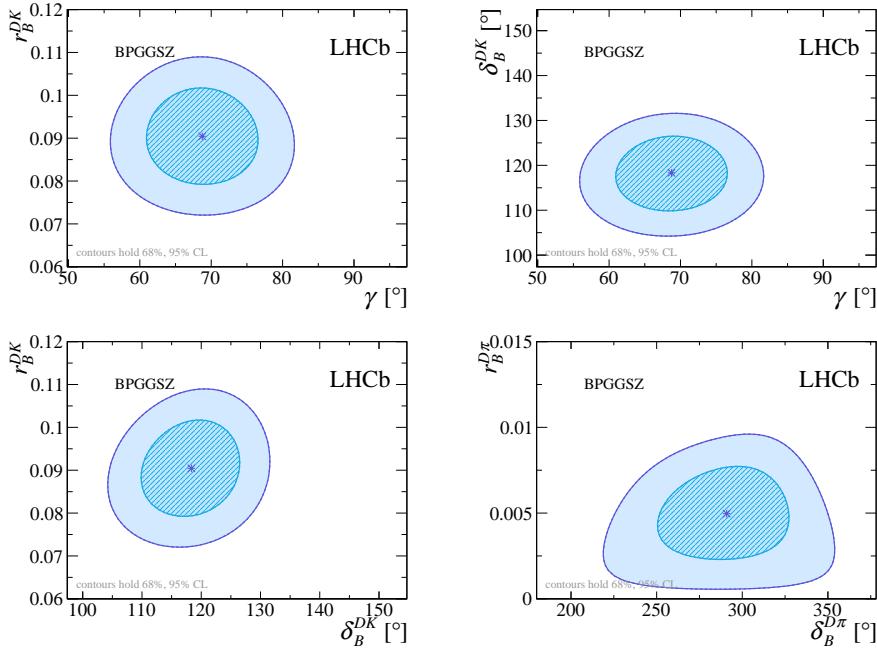
3205 where the quoted uncertainties are obtained via the **PLUGIN** method. The one-  
 3206 dimensional CL plots are shown in Fig. 5.57. It is also clear that the **PROB** and  
 3207 **PLUGIN** confidence regions agree well; this is expected given the Gaussian likelihood.  
 3208 A series of two-dimensional confidence regions are shown in Fig. 5.58, where it  
 3209 can be seen that the observables define a single solution for  $\gamma$  as expected. It  
 3210 is worth noticing that the uncertainty of this measurement alone is on par with  
 3211 the current world average, due to the increased data sample, and the significant  
 3212 reduction of systematic uncertainties due to the new measurement strategy and  
 3213 updated inputs from BESIII.

3214 The contribution to the uncertainty on  $\gamma$  from each of the statistical, strong-  
 3215 phase-related, and LHCb-related uncertainties in isolation can be estimated by  
 3216 repeating the interpretation while only including subsets of the uncertainties on  
 3217 the input parameters. Such studies have been performed using the **PROB** method.  
 3218 Running with statistical uncertainties only yields an uncertainty on  $\gamma$  of  $5.05^\circ$ .  
 3219 Including only the statistical and LHCb-related systematic uncertainties yields an  
 3220 uncertainty on  $\gamma$  of  $5.08^\circ$ , suggesting that the LHCb-related systematics contribute  
 3221 an uncertainty of  $0.6^\circ$ . This is a reduction compared to earlier analyses, where  
 3222 the contribution was about  $2^\circ$ . A significant contribution to the improvement is  
 3223 the efficiency-related systematic that has been avoided by promoting  $B^\pm \rightarrow D\pi^\pm$   
 3224 to a signal channel. Including only the statistical and the strong-phase-related  
 3225 uncertainties leads to an uncertainty on  $\gamma$  of  $5.09^\circ$ , showing the strong-phase-related  
 3226 uncertainty to be  $0.6^\circ$ , somewhat lower than the expectation of  $1.2^\circ$  presented in  
 3227 Ref. [?]. This is partly because the uncertainty estimate of that paper does not  
 3228 take into account the use of the  $D \rightarrow K_S^0 K^+ K^-$  channel, and partly because the  
 3229 uncertainty estimate depends on the specific central values.

3230 The obtained statistical uncertainty on  $\gamma$  is in excellent agreement with the  
 3231 expectation from pseudoexperiments. The interpretation procedure outlined above  
 3232 has been performed for each of the pseudoexperiments performed to establish the  
 3233 feasibility of the *CP* fit in Section 5.5.1 (including only statistical uncertainties on the  
 3234 observables) and the central 90 % interval of the obtained uncertainties  $[5.4^\circ, 6.0^\circ]$ .  
 3235 Similar studies have been carried out where no background decays are included in the  
 3236 generated toy data sets. In this case, the precision on  $\gamma$  is improved by about 30 %.



**Figure 5.57:** Confidence levels for the physics parameters of interest. The solutions are written on the plots, where the top number is given with PROB uncertainties and the bottom number with PLUGIN uncertainties.

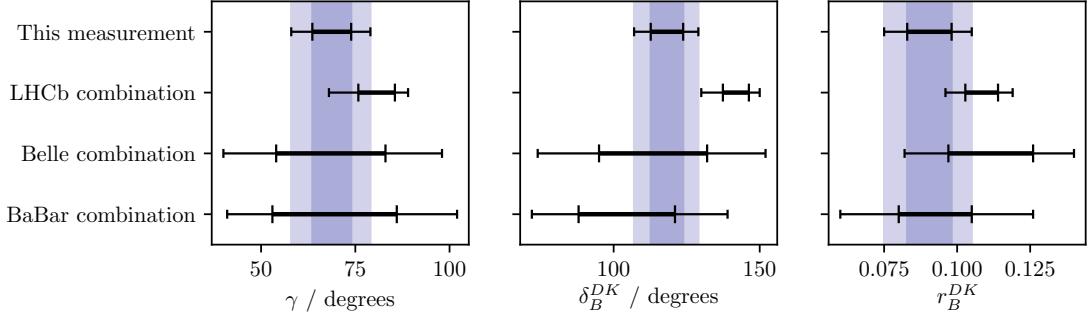


**Figure 5.58:** The 68 % and 95 % confidence regions for combinations of the physics parameters of interest, as obtained from the results of this measurement. The regions are calculated via the PROB method of `gammacombo`.

### 5.7.3 Compatibility with other measurements

It is worth comparing the obtained constraints on the physics parameters with the information available from other measurements, made at the  $B$  factories and by the LHCb collaboration using other decay channels. This comparison is made for  $\gamma$  and the hadronic parameters in the  $B^\pm \rightarrow DK^\pm$  decay in Fig. 5.59, comparing to the results of the combinations of  $\gamma$  measurements by the Belle [?] and BaBar [?] collaborations presented in 2013, and the 2018 combination of LHCb results [?]. For this purpose, the LHCb combination is re-performed, removing the input from earlier BPGBS measurements that use  $B^\pm \rightarrow DK^\pm$  decays, because they were made using data that is re-analysed in the present thesis; thus they need to be excluded to make the results that are compared independent. The combination employs the same statistical method outlined above, with the exception that the likelihood now depends on observables measured in a number of different analyses. The included measurements are summarised in Table 5.22. It can be seen in Fig. 5.59 that the results obtained in this thesis agree well with the Belle and BaBar results, but are in some tension with the 2018 LHCb combination, especially for the  $\delta_B^{DK}$  parameter.

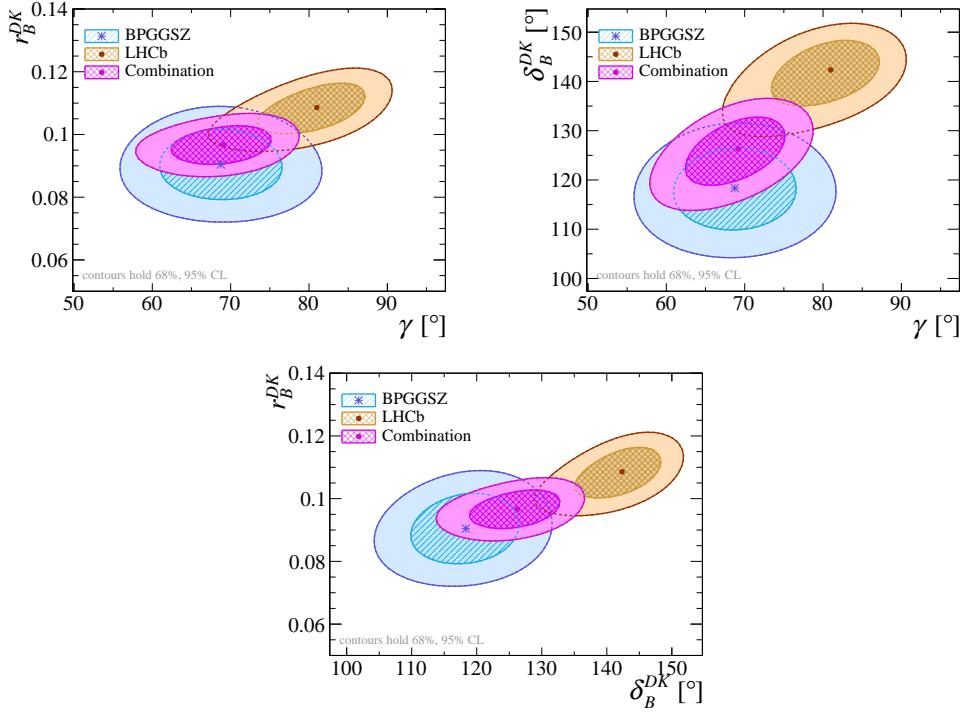
The level of compatibility can be quantified by calculating the three-dimensional  $\chi^2$  of the BPGBS results and those of the LHCb combination (without the earlier BPGBS measurements), with respect to the best fit values of  $(\gamma, r_B^{DK\pm}, \delta_B^{DK\pm})$



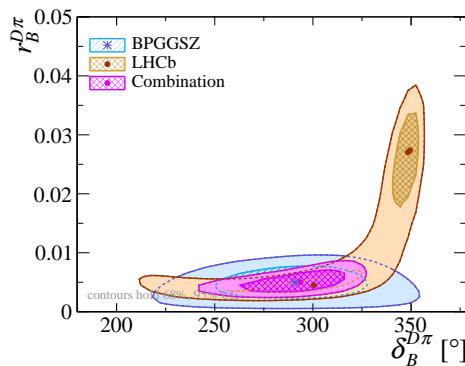
**Figure 5.59:** Comparison of the  $1\sigma$  and  $2\sigma$  confidence intervals obtained results for  $\gamma$  and the physics parameters relating to  $B^\pm \rightarrow DK^\pm$  decays, with those from the combinations of  $\gamma$  measurements by the Belle [?] and BaBar [?] collaborations, and the 2018 combination of LHCb results [?] where the BPGGSZ measurements have been excluded.

when all measurements are combined. The two-dimensional confidence regions obtained in these three cases are compared in Fig. 5.60, where some tension in  $r_B^{DK\pm}$  and  $\delta_B^{DK\pm}$  is visible again. The calculation is based on the PLUGIN uncertainties; for the LHCb combination these uncertainty estimates are slightly larger than the ones obtained via the PROB method. One obtains  $\chi^2 = \chi_{GGSZ}^2 + \chi_{\text{LHCb}}^2 = 0.7 + 9.1 = 9.8$ , which for 3 degrees of freedom correspond to a  $p$ -value of 2 %, or a  $2.3\sigma$  deviation. However, this tension is expected to be reduced when other measurements in the LHCb combination are updated to include results based on the full Run 1 and 2 data set. The most important update is that of the two-body ADS/GLW measurement in  $B^\pm \rightarrow DK^\pm$  decays because that measurement, and the BPGGSZ measurement presented in this thesis, have the largest impact in the combination.

The latest LHCb combination in which  $B^\pm \rightarrow D\pi^\pm$  parameters were determined is from 2016 [?]. Two solutions existed for  $(r_B^{D\pi^\pm}, \delta_B^{D\pi^\pm})$  which made the interpretation problematic. As can be seen in Fig. 5.61 [?], the measurement presented in this thesis picks out one of these solutions, with which it is in excellent agreement. This solution agrees with the theoretically expected value of  $r_B^{D\pi^\pm} \sim 0.005$  [?]. Thus, the inclusion of the results presented here are expected to lead to a much less problematic inclusion of the  $B^\pm \rightarrow D\pi^\pm$  channel in future LHCb combinations.



**Figure 5.60:** The 68 % and 95 % confidence regions for  $(\gamma, r_B^{DK\pm})$ ,  $(\gamma, \delta_B^{DK\pm})$ , and  $(\delta_B^{DK\pm}, r_B^{DK\pm})$  for this measurement, the 2018 LHCb combination [?] where the BPGGSZ results have been excluded, and the combination thereof, calculated via the PROB method of `gammacombo`.



**Figure 5.61:** The 68 % and 95 % confidence regions for  $(\delta_B^{D\pi\pm}, r_B^{D\pi\pm})$  obtained from the results of this measurement, in the 2016 LHCb combination [?], and the combination thereof, calculated via the PROB method of `gammacombo`.

**Table 5.22:** List of the LHCb measurements used in the combination that the results obtained in the present thesis is compared to. These correspond to the 2018 LHCb combination [?], except that the earlier BPGBS results made with  $B^\pm \rightarrow DK^\pm$  decays have not been included in the comparison. In the method column, TD stands for "time-dependent".

$B$ decay	$D$ decay	Method	Ref.	Data set
$B^+ \rightarrow DK^+$	$D \rightarrow h^+h^-$	GLW	[?]	2011-16
$B^+ \rightarrow DK^+$	$D \rightarrow h^+h^-$	ADS	[?]	2011-12
$B^+ \rightarrow DK^+$	$D \rightarrow h^+\pi^-\pi^+\pi^-$	GLW/ADS	[?]	2011-12
$B^+ \rightarrow DK^+$	$D \rightarrow h^+h^-\pi^0$	GLW/ADS	[?]	2011-12
$B^+ \rightarrow DK^+$	$D \rightarrow K_S^0 K^+ \pi^-$	GLS	[?]	2011-12
$B^+ \rightarrow D^*K^+$	$D \rightarrow h^+h^-$	GLW	[?]	2011-16
$B^+ \rightarrow DK^{*+}$	$D \rightarrow h^+h^-$	GLW/ADS	[?]	2011-16
$B^+ \rightarrow DK^{*+}$	$D \rightarrow h^+\pi^-\pi^+\pi^-$	GLW/ADS	[?]	2011-16
$B^+ \rightarrow DK^+\pi^+\pi^-$	$D \rightarrow h^+h^-$	GLW/ADS	[?]	2011-12
$B^0 \rightarrow DK^{*0}$	$D \rightarrow K^+\pi^-$	ADS	[?]	2011-12
$B^0 \rightarrow DK^+\pi^-$	$D \rightarrow h^+h^-$	GLW-Dalitz	[?]	2011-12
$B^0 \rightarrow DK^{*0}$	$D \rightarrow K_S^0 \pi^+\pi^-$	BPGBS	[?]	2011-12
$B_s^0 \rightarrow D_s^\mp K^\pm$	$D_s^+ \rightarrow h^+h^-\pi^+$	TD	[?]	2011-12
$B^0 \rightarrow D^\mp\pi^\pm$	$D^+ \rightarrow K^+\pi^-\pi^+$	TD	[?]	2011-12
Measurements included in Ref. [?] but not in the present comparison				
$B^+ \rightarrow DK^+$	$D \rightarrow K_S^0 h^+h^-$	BPGBS	[?]	2011-12
$B^+ \rightarrow DK^+$	$D \rightarrow K_S^0 h^+h^-$	BPGBS	[?]	2015-16

# 6

3274

3275

## Summary and outlook

3276 Say something clever

# Appendices

# A

3278

3279

Projections of the main fit to data