

A study of $B \rightarrow K \pi$ decays with the LHCb experiment

Andrew Gordon Buckley
of
Churchill College

A dissertation submitted to the University of Cambridge
for the degree of Doctor of Philosophy

Abstract

LHCb is a b-physics detector experiment which will take data at the 14 TeV LHC accelerator at CERN from 2007 onward. In this thesis I present two main areas of work relating to LHCb; firstly an investigation of a novel pattern recognition method in the LHCb RICH detectors and secondly a study of LHCb's potential to reconstruct and identify $B \rightarrow K \pi$ decays.

The pattern recognition studies presented here make use of a Markov Chain sampler to identify Čerenkov rings on a detector pixel array. The method presented does not use the tracking information required by the standard LHCb particle identification system. Studies have been performed to characterise the performance of the trackless ring-finder in the RICH2 detector and to investigate the sources of untracked rings. The ring-finder performance is seen to seriously degrade in high-occupancy regions, indicating that tracking information *is* required for realistic particle ID in LHCb. The sampler has also been integrated into the RICH2 background-estimation system, which permits robustness studies of the RICH particle ID system. Finally, I show that particle identification performance can be improved by use of the ring sampler for background estimation.

The study of $B \rightarrow K \pi$ decays also makes use of a multi-variable sampler, used to optimise the signal significance of signal-selection cuts. Using several types of cut definition, cut sets have been determined which show that LHCb expects to reconstruct $(90.0 \pm 6.2)\text{k}$ $B_d^0 \rightarrow K^+ \pi^-$ decays and $(6.9 \pm 0.6)\text{k}$ $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ decays in a year of operation with optimal signal significance. Robustness tests of the selection variables show the cuts to be largely resilient to systematic shifts of $\pm 20\%$ around the optimised values, although the data set is statistically limiting. It is anticipated that, in its first year of operation, LHCb will improve the statistical precision on the \mathcal{CP} observables for $B \rightarrow K \pi$ decays by an order of magnitude over that currently offered by the b-factories.

The branching ratios of the considered $B \rightarrow K \pi$ processes can be used, in principle, to both bound and to constrain the CKM weak phase γ . The predicted experimental uncertainties after one year of LHCb data-taking will tightly constrain the γ constraint, offering statistical errors of $\Delta\gamma \in [0.3^\circ, 2.9^\circ]$. However, the value of γ depends strongly on theoretical predictions of penguin-to-tree process amplitude ratios and the theoretical uncertainty on this renders the method unhelpful at present, placing an additional variation on γ of $\sim 50^\circ$. Hence, these decays are identified for now as more useful in enabling LHCb to constrain the effects of new physics in electroweak penguin loops than as a direct method to measure γ .

Declaration

This dissertation is the result of my own work, except where explicit reference is made to the work of others, and has not been submitted for another qualification to this or any other university. This dissertation does not exceed the word limit for the respective Degree Committee.

Andy Buckley

Acknowledgements

Of the many people who deserve thanks, some are particularly prominent:

My supervisor Val Gibson for extreme tolerance, many words of wisdom and of course knowing just what to write on those magic travel application forms.

The UK Particle Physics and Astronomy Research Council, for funding my PhD and enabling me to attend many useful meetings, conferences and workshops.

Chris Jones for help, concern and discussion far beyond the call of duty; Chris Lester for conceiving the ring-finding sampler and engaging my interest in topics as diverse as uni-cycling and large-number factoring; Pat Ward, Cristina Lazzeroni and Karl Harrison for advice and proof-reading; Patrick Koppenburg and Gloria Corti for help with DaVinci; Vincenzo Vagnoni and Yuehong Xie for discussions about analysis and reconstruction issues; Steve Wotton for friendship, squash advice and curry; and Martin White, whose enthusiasm for protracted musical discussions is already sorely missed.

The arXiv and SPIRES online preprint and literature databases — and especially Travis Brooks for writing the automated SPIRES bibliography generator and generously extending it to fit my needs. A multitude of free software developers deserve recognition for providing the many tools which make things like this possible.

All the other Cambridge HEP PhD students and post-docs, who have all assisted in making the last three years so enjoyable. This particularly applies to those who enthusiastically joined my campaign to play computer games across the group network twice a day instead of writing this thesis: you unwittingly saved my sanity.

And of course my parents and my girlfriend Jo, who have never been anything but supportive and generally wonderful.

Preface

This thesis describes my research on various aspects of the LHCb particle physics program, centred around the LHCb detector and LHC accelerator at CERN in Geneva. Two main areas of work are described: studies relating to pattern recognition in the LHCb ring-imaging Čerenkov (RICH) detectors and an investigation of the potential of LHCb to reconstruct and extract \mathcal{CP} violation physics from $B \rightarrow K \pi$ decay modes.

To place my own work in context, I first summarise the physics of b-hadrons and in particular their relation to Standard Model \mathcal{CP} violation: this is the subject of Chapter 1. The approach of the LHCb collaboration to investigating this physics, especially the design parameters of the LHCb detector, is described in Chapter 2. Chapter 3 focuses on new approaches to pattern recognition in the LHCb RICH detectors.

Finally, three chapters are devoted to the description of $B \rightarrow K \pi$ decays. Chapter 4 reviews the flavour physics potential of these decays and the phenomenology required to extract meaningful physics from their measurement, then Chapter 5 describes the approach taken to reconstructing and selecting these decays using LHCb. The potential for LHCb to extract \mathcal{CP} -violation physics using these decay channels is considered in Chapter 6.

Contents

1	\mathcal{CP} violation in the B-meson system	1
1.1	Neutral meson mixing	2
1.2	Introduction to \mathcal{CP} violation	5
1.2.1	Cosmological connections	6
1.3	\mathcal{CP} -violation in B-mesons	7
1.3.1	\mathcal{CP} -violation in meson mixing	7
1.3.2	\mathcal{CP} -violation in meson decay	8
1.3.3	\mathcal{CP} -violation in the interference of mixing and decay	9
1.4	\mathcal{CP} violation in the Standard Model	11
1.4.1	Unitarity triangles	16
1.4.2	$B-\bar{B}$ mixing in the Standard Model	19
1.4.3	Weak decays of B-mesons	20
1.5	Strategies to constrain CKM parameters	23
1.6	Physics from beyond the Standard Model	31
2	The LHCb experiment	35
2.1	b-physics experiments	35
2.2	The LHC	36
2.3	b-production at the LHC	38

2.4	The LHCb experiment	39
2.4.1	Interaction point	41
2.4.2	Beam pipe	42
2.4.3	Magnet	43
2.4.4	Vertex locator (VELO)	43
2.4.5	RICH system	47
2.4.6	Tracking system	54
2.4.7	Calorimeters	56
2.4.8	Muon system	58
2.5	Trigger system	59
2.5.1	L0	59
2.5.2	L1	61
2.5.3	HLT	61
2.6	Data handling	62
2.6.1	Offline analysis	63
2.6.2	LHCb software framework	63
2.7	Flavour tagging	65
3	Trackless RICH pattern recognition	67
3.1	The standard RICH particle ID algorithm	67
3.1.1	Global pattern recognition	68
3.1.2	Local pattern recognition	68
3.1.3	Other pattern recognition methods	69
3.1.4	Standard particle ID algorithm structure	69
3.1.5	Standard particle ID performance	70

3.2	Trackless ring-finding	75
3.2.1	The trackless ring-finding system	75
3.2.2	Removing ring distortions	77
3.2.3	The trackless ring-finding algorithm	79
3.2.4	Associating rings to pixels and tracks	83
3.3	Trackless ring-finder performance in RICH2	85
3.3.1	Defining MC-to-reconstruction associations	86
3.3.2	Monitoring the algorithm performance	86
3.3.3	Performance of the trackless ring-finder	88
3.4	Untracked rings	94
3.5	Impact of trackless ring-finding on particle ID	98
3.5.1	Stripping background pixels	98
3.5.2	Weighting background pixels	100
3.5.3	Simulating extra background tracks	100
3.5.4	Performance of the Markov background estimator	101
3.6	Conclusions	103
4	Phenomenology of $B \rightarrow K \pi$ decays	109
4.1	Further phenomenology of weak non-leptonic B decays	109
4.1.1	Low energy effective theories	110
4.1.2	Matrix element factorisation	111
4.2	CKM phenomenology from $B \rightarrow K \pi$ decays	113
4.2.1	Amplitudes and observables for $B \rightarrow K \pi$ channels	113
4.2.2	The “mixed” $B^\pm \rightarrow K^0 \pi^\pm$, $B_d^0 \rightarrow K^\pm \pi^\mp$ strategy	114
4.2.3	Extracting γ from $B \rightarrow K \pi$ decays	120

4.2.4	Model-dependent determination of γ	122
4.3	Hadronic uncertainties in $B \rightarrow K \pi$ decays	124
4.3.1	Effects of re-scattering on $B \rightarrow K \pi$ decay phenomenology	124
4.3.2	Theoretical approaches to re-scattering	126
4.4	Future prospects	129
5	Selecting $B \rightarrow K \pi$ decays with LHCb	131
5.1	Decay channels of interest	131
5.1.1	Signal channels	131
5.2	Sources of background	132
5.3	Simulation and initial reconstruction	133
5.4	Event reconstruction	134
5.4.1	Reconstructing $B_d^0 \rightarrow K^+ \pi^-$	135
5.4.2	Reconstructing $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$	137
5.4.3	Primary vertex selection	143
5.5	Preliminary event selection	144
5.5.1	Producing data for optimisation	145
5.5.2	Cut variable distributions	148
5.6	Optimising the selection cuts	164
5.6.1	Characterising selection performance	165
5.6.2	Preliminary data processing	166
5.6.3	Monte Carlo sampling	170
5.7	Optimised cut sets	171
5.7.1	Optimised cut sets for $B_d^0 \rightarrow K^+ \pi^-$	173
5.7.2	Optimised cut sets for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$	175

5.8	Cut performance for $B_d^0 \rightarrow K^+ \pi^-$	178
5.8.1	Cut set robustness	180
5.9	Cut performance for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$	183
5.9.1	Cut set robustness	185
5.10	Conclusions	190
6	Event yields and sensitivity to γ from $B \rightarrow K \pi$ decays	191
6.1	Efficiencies, signal yields and background estimates	191
6.1.1	Efficiencies and signal event yields	193
6.1.2	$b\bar{b}$ inclusive background constraint	194
6.1.3	Specific decay channel backgrounds	197
6.2	Sensitivity to γ	197
6.3	Sensitivity to new physics	204
7	Conclusions	205
A	Monte Carlo samplers	1
A.1	Metropolis algorithm	1
A.2	Simplex algorithm	3
B	Colophon	7
	Bibliography	9
	List of Figures	21
	List of Tables	27

*“Writing in English is the most ingenious torture
ever devised for sins committed in previous lives.”*

— James Joyce

Chapter 1

\mathcal{CP} violation in the B-meson system

“Laws were made to be broken.”

— Christopher North 1785–1854

Symmetries, either intact or broken, have proved to be at the heart of how matter interacts. The Standard Model of fundamental interactions (SM) is composed of three independent continuous symmetry groups denoted $SU(3) \times SU(2) \times U(1)$, representing the strong force, weak isospin and hypercharge respectively [1,2,3]. The breaking of such symmetries is also significant: the Higgs mechanism uses the spontaneous breaking of the Standard Model $SU(2) \times U(1)$ symmetries to the $U(1)$ of electromagnetism to generate masses for the appropriate fermion and boson fields [4].

The above examples are all continuous symmetries, which are formally connected to the gauge-current formalism around which all of modern particle physics is built [5]. But what of discrete symmetries? The Standard Model contains several conserved quantities which are not protected by any explicit “custodial” symmetry but merely happen to be conserved by every constituent interaction in the theory: baryon and lepton number are prominent examples of this.

Another set of prominent discrete symmetry operations with special position in the Standard Model are the charge conjugation (\mathcal{C}), parity (\mathcal{P}) and time-reversal (\mathcal{T}) operations [6]. These operators are defined as follows, using Lorentz four-vectors:

- \mathcal{C} : transforms particles into antiparticles while leaving e.g. spin and momentum unchanged;

- \mathcal{P} : spatial inversion such that $(ct, \mathbf{x}) \rightarrow (ct, -\mathbf{x})$; and
- \mathcal{T} : time inversion such that $(ct, \mathbf{x}) \rightarrow (-ct, \mathbf{x})$.

The transformation properties of fields under these operations are used to class them as scalars, pseudoscalars, vectors or tensors. Additionally, local quantum field theories require for consistency that the combined \mathcal{CPT} operation be a good symmetry of the theory [7]. The experimental discovery that the combined \mathcal{CP} operation is not a good symmetry of weak interactions [8] and the theoretical development of models to explain such a violation is the motivation for the work described in the rest of this thesis.

1.1 Neutral meson mixing

Before discussing \mathcal{CP} violation however, we must consider the oscillation phenomenon in neutral meson systems. Mesons are colour-singlet states composed of a quark and an anti-quark, for example $u\bar{d}$ is a π^+ meson. In neutral mesons the quark-antiquark pairs are both of either up- or down-type. For example, the K^0 meson is a $d\bar{s}$ state, the D^0 meson is $u\bar{c}$ and the B-system contains two neutral mesons: $B_d^0 = \bar{b}d$ and $B_s^0 = \bar{b}s$.

The possibility of neutral meson mixing was first proposed by Gell-Mann and Pais in 1955 [9]. Their suggestion was based on the observed lack of flavour conservation in weak decays (although at the time, this was phrased in terms of violation of the strangeness quantum number as the parton level description had not yet evolved). Their description in terms of neutral kaons freely mixing into their anti-particles can be extended to many other neutral mesons, of which the ones primarily encountered in b-physics are $B_d^0-\bar{B}_d^0$ and $B_s^0-\bar{B}_s^0$.

We can go a long way with an effective Hamiltonian approach in canonical single-particle quantum mechanics. To do this we construct a wavefunction from a combination of a generic neutral meson state $|X^0\rangle$ and its anti-state $|\bar{X}^0\rangle$:

$$|\psi(t)\rangle = a(t)|X^0\rangle + b(t)|\bar{X}^0\rangle \quad (1.1)$$

which is governed by a time-dependent matrix differential equation,

$$i\frac{\partial}{\partial t} \begin{pmatrix} a \\ b \end{pmatrix} = \underbrace{\begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12}^* - \frac{i}{2}\Gamma_{12}^* & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix}}_{\mathbf{H}} \begin{pmatrix} a \\ b \end{pmatrix}. \quad (1.2)$$

Here the Hamiltonian matrix \mathbf{H} is decomposed into the component mass and decay matrices, \mathbf{M} and $\mathbf{\Gamma}$ respectively. The assumed invariance under \mathcal{CPT} implies that the particle and antiparticle masses and lifetimes are equal, i.e. $M_{11} = M_{22} \equiv M$ and $\Gamma_{11} = \Gamma_{22} \equiv \Gamma$. Diagonalising \mathbf{H} leads to two stationary eigenstates which can be expressed as linear combinations of the input states:

$$|X_1\rangle = p|X^0\rangle + q|\bar{X}^0\rangle \quad (1.3a)$$

$$|X_2\rangle = p|X^0\rangle - q|\bar{X}^0\rangle \quad (1.3b)$$

where the admixture constants p and q are complex numbers with the normalisation condition $|p|^2 + |q|^2 = 1$. By the standard diagonalisation procedure, the energy eigenvalues obtained are

$$\begin{aligned} E_{1,2} &= M_{1,2} - \frac{i}{2}\Gamma_{1,2} \\ &= M - \frac{i}{2}\Gamma \pm \sqrt{(M_{12} - \Gamma_{12})(M_{12}^* - \Gamma_{12}^*)} \\ &= \left(M \mp \frac{\Delta M}{2}\right) - \frac{i}{2}\left(\Gamma \mp \frac{\Delta\Gamma}{2}\right). \end{aligned} \quad (1.4)$$

From this we can identify $\Delta M \equiv M_2 - M_1$ and $\Delta\Gamma \equiv \Gamma_2 - \Gamma_1$ where $M_{1,2}$ and $\Gamma_{1,2}$ are the masses and widths of the corresponding eigenstates. Similarly the eigenvector equation $(\mathbf{H} - E\mathbf{I})(p, \pm q)^T = 0$ constrains p and q :

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}}. \quad (1.5)$$

Re-arranging and substituting the inverse eigenvalues as time constants in the stationary state phase evolution of (1.3a) and (1.3b), we obtain time-evolution expressions

in terms of the initial X^0 and \bar{X}^0 states:

$$|X^0(t)\rangle = f_+(t)|X^0(0)\rangle + \frac{q}{p}f_-(t)|\bar{X}^0(0)\rangle \quad (1.6a)$$

$$|\bar{X}^0(t)\rangle = f_+(t)|\bar{X}^0(0)\rangle + \frac{p}{q}f_-(t)|X^0(0)\rangle \quad (1.6b)$$

with $f_{\pm}(t) = \frac{1}{2} [e^{-i(M_1 - \frac{i}{2}\Gamma_1)t} \pm e^{-i(M_2 - \frac{i}{2}\Gamma_2)t}]$. The probabilities of obtaining $|X^0\rangle$ and $|\bar{X}^0\rangle$ states with a measurement at time t , assuming an initially pure $|X^0\rangle$ state, are hence

$$p(X^0 \rightarrow X^0 : t) = |\langle X^0 | X^0(t) \rangle|^2 = |f_+(t)|^2 \quad (1.7a)$$

$$p(X^0 \rightarrow \bar{X}^0 : t) = |\langle \bar{X}^0 | X^0(t) \rangle|^2 = \left| \frac{q}{p} f_-(t) \right|^2 \quad (1.7b)$$

with

$$|f_{\pm}(t)|^2 = \frac{1}{4} [e^{-\Gamma_1 t} + e^{-\Gamma_2 t} \pm 2e^{-\bar{\Gamma} t} \cos(\Delta M t)] \quad (1.8)$$

where

$$\bar{\Gamma} = \frac{1}{2} (\Gamma_1 + \Gamma_2). \quad (1.9)$$

In equation (1.8), the last term describes the oscillating character of X^0 and \bar{X}^0 content in an initially pure X^0 state. The characteristics of the oscillation are determined by the oscillation parameter $x = |\Delta M|/\bar{\Gamma}$. Since x is the ratio of the oscillation and decay rates, larger values of x correspond to greater likelihoods of oscillation before decay. In Figure 1.1 the probabilities of finding an X^0 and an \bar{X}^0 from an initial X^0 state for various values of x are shown as a function of time, with the assumption that $|q/p| = 1$. As we shall see later, this assumption is equivalent to conserving \mathcal{CP} .

The sub-figures of Figure 1.1 have been chosen to correspond to K^0 – \bar{K}^0 mixing, B_d^0 – \bar{B}_d^0 mixing and B_s^0 – \bar{B}_s^0 mixing respectively. In the case of neutral kaon mixing the mass difference between the physical (K_S^0 and K_L^0) eigenstates is very small: $|\Delta M_K| \approx 3.29 \times 10^{-15}$ GeV/c [10]. The K_S^0 and K_L^0 have very different lifetimes due largely to the rapid decay of K_S^0 into two pions and so $\Gamma_S > \Gamma_L$. Overall, $x_K \approx 1$.

By contrast, for B^0 – \bar{B}^0 mixing the decay rates of the physical $B_{L,H}$ states are approximately equal so $\Gamma_L \approx \Gamma_H$. However, in direct contrast to the kaon system, the

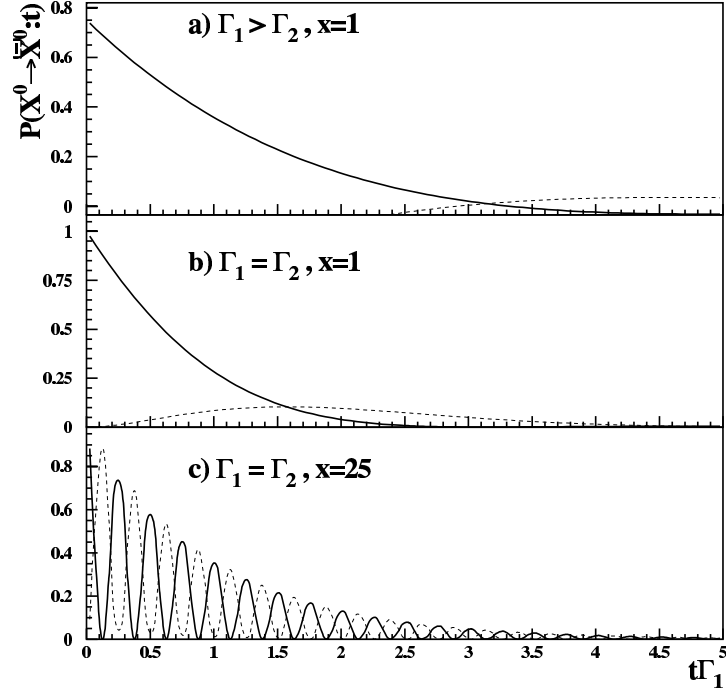


Figure 1.1: Neutral meson mixing plots, corresponding to (a) $K^0-\bar{K}^0$ mixing, (b) $B_d^0-\bar{B}_d^0$ mixing and (c) $B_s^0-\bar{B}_s^0$ mixing, starting from a pure X^0 state. The solid lines illustrate the probability of the X^0 state and the dotted lines show the probability of the \bar{X}^0 .

eigenstate mass differences are large: $\Delta M_d = (3.30 \pm 0.04) \times 10^{-13} \text{ GeV}/c$ and $\Delta M_s > 9.5 \times 10^{-12} \text{ GeV}/c$ (95% C.L.) for B_d^0 and B_s^0 respectively [10]. When combined with the measured B_d^0 and B_s^0 lifetimes, $\tau_d = (1.542 \pm 0.016) \text{ ps}$ and $\tau_s = (1.461 \pm 0.057) \text{ ps}$ respectively, we obtain corresponding oscillation parameters $x_d = 0.77 \pm 0.01$ and $x_s > 20.6$. These describe oscillation behaviours as shown in the two lower plots of Figure 1.1.

1.2 Introduction to \mathcal{CP} violation

We have already introduced the \mathcal{C} , \mathcal{P} and \mathcal{T} discrete symmetry operators and noted the dependence of local quantum field theories on symmetry under the combined \mathcal{CPT} operation. Historically it was expected that all fundamental interactions conserved all three discrete symmetries and, even when the weak force's maximal parity violation was discovered, the expectation was that the combined \mathcal{CP} operation would remain a good symmetry of the theory.

An example of this is easily seen in weak decays involving neutrinos:

$$\pi^+ \rightarrow e^+ \nu_e \quad \xrightarrow{\mathcal{C}} \quad \pi^- \rightarrow e^- \nu_e^{\mathcal{C}} \quad \xrightarrow{\mathcal{P}} \quad \pi^- \rightarrow e^- \bar{\nu}_e \quad (1.10)$$

where the significance is that the pure charge-conjugation transformation \mathcal{C} produces a left-handed $\nu_e^{\mathcal{C}}$ state which is not observed in nature. The additional parity operation \mathcal{P} resolves the problem by constructing the right-handed electron anti-neutrino $\bar{\nu}_e$.

It was therefore a surprise when Christensen, Cronin, Fitch and Turlay discovered \mathcal{CP} -violation in weak decays of neutral kaons in 1964, specifically in the observation of the decay $K_L^0 \rightarrow \pi^+ \pi^-$. Defining $|K_2^0\rangle = 1/\sqrt{2}[(K^0 - \bar{K}^0) + \epsilon(K^0 + \bar{K}^0)]$, the first measurement of \mathcal{CP} -violation established $\epsilon = 2.3 \times 10^{-3}$ — a small effect compared to the maximal weak-current \mathcal{P} -violation [8].

1.2.1 Cosmological connections

After the measurement of explicit \mathcal{C} - and \mathcal{CP} -violation in kaon decays, Sakharov argued in 1967 that \mathcal{CP} -violation is one of the minimal conditions which must be met in the early universe to explain the emergence of the observed cosmological baryon–antibaryon asymmetry from a presumably symmetric distribution at the Big Bang [11]. Sakharov stipulated three such conditions on early-universe physics:

1. \mathcal{C} and \mathcal{CP} -violating processes;
2. conditions of thermal non-equilibrium, specifically a phase-transition; and
3. baryon number violating processes: baryogenesis/leptogenesis.

These conditions are subject to several technical restrictions. Prominently, in the post-LEP era with confirmation of the Glashow-Weinberg-Salam electro-weak Standard Model [1, 2] it is tempting to assume that the electroweak phase transition is an appropriate candidate to match the second Sakharov condition. However, the phase transition must be abrupt (technically a strong “first order” transition) and for Higgs boson masses $m_H \geq 45 \text{ GeV}/c$ the electroweak phase transition is smooth (second order) [12, 13]. The rôle of the electroweak transition in generating the cosmological matter–antimatter asymmetry is unknown.

Baryon number violation is connected to new physics processes: in particular leptogenesis and lepton number violation through Majorana neutrinos and anomalous processes

(the “sphaleron”) are popular mechanisms. These both result in breaking the “accidental” B and L symmetries but retaining a combined baryon-lepton number, $B - L$ which is globally respected [14, 15, 13].

Non-hadronic \mathcal{CP} -violation (i.e. \mathcal{CP} -violation in the Higgs or lepton sectors) is probably the dominant contribution to cosmological \mathcal{CP} -violation since the hadronic sector does not contain enough \mathcal{CP} violation to account for the observed Sakharov matter-antimatter asymmetry, at least in the most standard cosmological scenarios [16, 17]. The apparent inability of Standard Model \mathcal{CP} -violation to generate a cosmological matter-antimatter asymmetry encourages searches for \mathcal{CP} -violation in new physics and in the leptonic sector following the confirmation of non-zero neutrino masses.

1.3 \mathcal{CP} -violation in B-mesons

We will now review briefly the phenomenon of neutral meson mixing and the three forms in which meson \mathcal{CP} violation can be manifest: mixing, decay and interference. For generality the description will be continued in terms of arbitrary “X” mesons but our emphasis is clearly on the more specific case of B-mesons.

1.3.1 \mathcal{CP} -violation in meson mixing

\mathcal{CP} -violation can be introduced into the formalism for neutral meson mixing already introduced by setting $|q/p| \neq 1$. This is called \mathcal{CP} -violation in meson mixing and is often referred to as *indirect* \mathcal{CP} -violation. To see how this works, we note that \mathcal{CP} invariance in the mixing is represented by equality of off-diagonal elements in the effective Hamiltonian, i.e.

$$\left| M_{12} - \frac{i}{2} \Gamma_{12} \right| = \left| M_{12}^* - \frac{i}{2} \Gamma_{12}^* \right| \Rightarrow \mathcal{CP}\text{-invariance in mixing.} \quad (1.11)$$

Physically this corresponds to equality in the rates for $X^0 \rightarrow \bar{X}^0$ and $\bar{X}^0 \rightarrow X^0$ transitions. In equation (1.5) we see that violation of the equality in equation (1.11) corresponds to $|q/p| \neq 1$ as already stated. The rate difference due to this can be seen from equation (1.6b).

For completeness we note that if \mathcal{CP} is conserved in the mixing then for the mass

eigenstates defined as

$$|X_1\rangle = \frac{1}{\sqrt{2}}(|X^0\rangle + |\bar{X}^0\rangle) \quad (1.12a)$$

$$|X_2\rangle = \frac{1}{\sqrt{2}}(|X^0\rangle - |\bar{X}^0\rangle) \quad (1.12b)$$

we obtain the standard $\mathcal{CP}|X_{1,2}\rangle = \pm |X_{1,2}\rangle$ with the conventional $|\bar{X}^0\rangle = \mathcal{CP}|X^0\rangle$ and $|X^0\rangle = \mathcal{CP}|\bar{X}^0\rangle$ definitions.

1.3.2 \mathcal{CP} -violation in meson decay

\mathcal{CP} -violation can also occur in interference between \mathcal{CP} -conjugate decay amplitudes, in general representable as $X \rightarrow f$ and $\bar{X} \rightarrow \bar{f}$. This is called \mathcal{CP} -violation in meson decay, often known as *direct* \mathcal{CP} -violation. We write the two decay amplitudes as

$$A \equiv \langle f | H | X \rangle \quad \text{and} \quad \bar{A} \equiv \langle \bar{f} | H | \bar{X} \rangle. \quad (1.13)$$

Anticipating the incorporation of this high-level description into the Standard Model, we decompose the phases in H into the *weak phases* from the weak quark mixing matrix and *strong phases* from the final state hadron re-scattering mediated by the strong force:

$$A = \sum_i A_i e^{i\phi_i} e^{i\delta_i} \quad \text{and} \quad \bar{A} = \sum_j A_j e^{-i\phi_j} e^{i\delta_j}. \quad (1.14)$$

Note here that the \mathcal{CP} -conjugation of these amplitudes inverts the sign of the weak phases ϕ_i but not the strong phases δ_i . If \mathcal{CP} is conserved then $\phi_i = \phi_j$ and, assuming the strong force to be \mathcal{CP} -invariant, $\delta_i = \delta_j$. Considering the observable ratio of decay amplitudes we see that

$$\left| \frac{\bar{A}}{A} \right| = \left| \frac{\sum_j A_j e^{-i\phi_j} e^{i\delta_j}}{\sum_i A_i e^{i\phi_i} e^{i\delta_i}} \right| = 1 \quad (1.15)$$

if \mathcal{CP} is conserved in the decay. If, however, the decay can proceed via at least two mechanisms with different phases then the interference between amplitudes leads to $|\bar{A}/A| \neq 1$. This is the only kind of meson \mathcal{CP} -violation possible for charged mesons. In principle, an asymmetry in charged decays can be constructed:

$$\mathcal{A}_{+-} = \frac{\Gamma X^+ \rightarrow f - \Gamma X^- \rightarrow \bar{f}}{\Gamma X^+ \rightarrow f + \Gamma X^- \rightarrow \bar{f}} = \frac{1 - |\bar{A}/A|^2}{1 + |\bar{A}/A|^2}, \quad (1.16)$$

but in practice predicting the strong phase difference makes this measurement difficult [18]. To date, direct \mathcal{CP} -violation has been observed in the neutral kaon system via the measurement of $\Re\{\epsilon'/\epsilon\} = (1.67 \pm 0.26) \times 10^{-3}$ [10, 19] and in the B-system via $B_d^0 \rightarrow K^+ \pi^-$ decays [20, 21].

1.3.3 \mathcal{CP} -violation in the interference of mixing and decay

A third kind of \mathcal{CP} -violation can occur when a neutral meson X^0 and its anti-particle \bar{X}^0 can decay into the same final state f or \bar{f} . In this case the \mathcal{CP} -violation occurs in the interference of mixing and decay amplitudes and is known as \mathcal{CP} -violation in the *interference of mixing and decay*.

For convenience we define the complex ratios

$$\lambda = \left(\frac{q}{p}\right)_{X^0} \left(\frac{\bar{A}}{A}\right)_f \quad \text{and} \quad \bar{\lambda} = \left(\frac{q}{p}\right)_{X^0} \left(\frac{\bar{A}}{A}\right)_{\bar{f}} \quad (1.17)$$

where the (q/p) is as defined for neutral meson mixing and evaluated for the X^0 meson. A and \bar{A} are the amplitudes for X^0 and \bar{X}^0 to decay to the final state f (or \bar{f}):

$$A_f = \langle f | H | X^0 \rangle \quad \bar{A}_f = \langle f | H | \bar{X}^0 \rangle \quad (1.18a)$$

$$A_{\bar{f}} = \langle \bar{f} | H | X^0 \rangle \quad \bar{A}_{\bar{f}} = \langle \bar{f} | H | \bar{X}^0 \rangle. \quad (1.18b)$$

\mathcal{CP} is violated in this scenario if $\lambda, \bar{\lambda} \neq 1$. This can be the case if either direct or indirect \mathcal{CP} -violating processes contribute, i.e. if \mathcal{CP} is violated either in the mixing ($|q/p| \neq 1$) or decay ($|\bar{A}/A| \neq 1$) independently. It can also be met in the case where neither direct nor indirect \mathcal{CP} -violation occurs but where the phases of q/p and \bar{A}/A interfere to give $\Im\{\lambda, \bar{\lambda}\} \neq 0$. \mathcal{CP} -violation in the interference is a combination of both the direct and indirect \mathcal{CP} -violation mechanisms.

The time-dependent rates for transitions from X^0 or \bar{X}^0 to f can be obtained, using

equations (1.6a) and (1.6b):

$$\Gamma_f = \frac{|A_f|^2}{2} e^{-\bar{\Gamma}t} [I_+(t) + I_-(t)] \quad (1.19a)$$

$$\bar{\Gamma}_f = \frac{|\bar{A}_f|^2}{2|\lambda|^2} e^{-\bar{\Gamma}t} [I_+(t) - I_-(t)] \quad (1.19b)$$

where the interference terms, $I_{\pm}(t)$ are

$$I_+(t) = (1 + |\lambda|^2) \cosh\left(\frac{\Delta\Gamma t}{2}\right) - 2 \Re\{\lambda\} \sinh\left(\frac{\Delta\Gamma t}{2}\right) \quad (1.20)$$

$$I_-(t) = (1 - |\lambda|^2) \cos\left(\frac{\Delta Mt}{2}\right) + 2 \Im\{\lambda\} \sin\left(\frac{\Delta Mt}{2}\right). \quad (1.21)$$

The corresponding expressions for the \bar{f} final state are obtained by substituting $\bar{\lambda}$ for λ in equations (1.19a)–(1.21). We can then define the time-dependent \mathcal{CP} asymmetry $\mathcal{A}_{\mathcal{CP}}(t)$ and write it in terms of λ :

$$\begin{aligned} \mathcal{A}_{\mathcal{CP}}(t) &= \frac{\Gamma_f - \bar{\Gamma}_f}{\Gamma_f + \bar{\Gamma}_f} \\ &= \frac{(1 - |\lambda|^2) \cos(\Delta Mt) - 2 \Im\{\lambda\} \sin(\Delta Mt)}{(1 + |\lambda|^2) \cosh(\Delta\Gamma t/2) + 2 \Re\{\lambda\} \sinh(\Delta\Gamma t/2)}. \end{aligned} \quad (1.22)$$

Alternatively, $\mathcal{A}_{\mathcal{CP}}(t)$ may be split into “direct” and “mixing” components with amplitudes $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$ and $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}$ respectively such that

$$\mathcal{A}_{\mathcal{CP}}(t) = \mathcal{A}_{\mathcal{CP}}^{\text{dir}} \cos(\Delta Mt) + \mathcal{A}_{\mathcal{CP}}^{\text{mix}} \sin(\Delta Mt). \quad (1.23)$$

If $\Delta\Gamma$ is small, as is the case for B_d^0 mesons, then the $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}(t)$ and $\mathcal{A}_{\mathcal{CP}}^{\text{mix}}(t)$ terms are

$$\mathcal{A}_{\mathcal{CP}}^{\text{dir}} = \frac{1 - |\lambda|^2}{1 + |\lambda|^2} \quad \text{and} \quad \mathcal{A}_{\mathcal{CP}}^{\text{mix}} = \frac{2 \Im\{\lambda\}}{1 + |\lambda|^2}. \quad (1.24)$$

A further simplification is obtained in the case where the final state f is a \mathcal{CP} eigenstate. In this case $|\lambda| = 1$ and the decay is dominated by a single \mathcal{CP} -violating

phase. $\mathcal{A}_{\mathcal{CP}}^{\text{dir}}$ is then zero and equation (1.23) reduces to

$$\begin{aligned}\mathcal{A}_{\mathcal{CP}}(t) &= \mathcal{A}_{\mathcal{CP}}^{\text{mix}} \sin(\Delta Mt) \\ &= \Im\{\lambda\} \sin(\Delta Mt).\end{aligned}\tag{1.25}$$

For decays which match the required criteria this provides a clean approach for extraction of the phase difference between the mixing and decay amplitudes since $\Im\{\lambda\}$ is the phase difference between the mixing and decay amplitudes and can be extracted without recourse to specific hadronic models. \mathcal{CP} -violation in the interference has been measured in K^0 decays where it has been observed to have approximately the same magnitude ($\mathcal{O}(10^{-3})$) as \mathcal{CP} -violation in the mixing [22]. In the B-system, \mathcal{CP} -violation in the interference has been measured by the b-factory experiments to have a relatively large asymmetry [23].

1.4 \mathcal{CP} violation in the Standard Model

The observational fact that \mathcal{CP} is violated at all in a gauge-field formalism is slightly surprising: we have to “try rather hard” to break the \mathcal{CP} symmetry [24]. This can be observed for example in the massless QED and QCD theories, where \mathcal{C} and \mathcal{P} are preserved both separately and in combination. Even when chiral symmetry is explicitly broken the combined \mathcal{CP} symmetry remains intact. This is illustrated, for example, with the linear sigma model whose Lagrangian,

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + i\sigma\bar{\psi}_L D\psi_L,\tag{1.26}$$

allows only the left-handed Weyl fermions to interact with the gauge bosons implicit in the covariant derivative D . The \mathcal{CP} symmetry remains intact in this model. Such robustness implies that \mathcal{CP} is a natural symmetry of massless gauge theories [24]. It is only possible to incorporate \mathcal{CP} -violation in theories where mass has been introduced.

Since an explicit mass term will explicitly violate gauge symmetry, we are forced to use some ingenious method to introduce mass self-consistently: the most prominent such method is the Higgs mechanism of the Standard Model, in which a scalar field couples to the fermion and gauge boson fields in a way which leaves the Lagrangian gauge-invariant. The mass is then introduced by letting the scalar (Higgs) field acquire a vacuum expectation value (VEV), which spontaneously breaks the gauge invariance and

produces mass terms as required. The Standard Model Lagrangian is then decomposable into three sectors:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{gauge}}(\psi_L, \psi_R, W, \phi) + \mathcal{L}_{\text{Higgs}}(\phi) + \mathcal{L}_{\text{Yukawa}}(\psi_L, \psi_R, \phi). \quad (1.27)$$

In the Standard Model the gauge and Higgs terms are automatically \mathcal{CP} -invariant and so the source of \mathcal{CP} -violation is narrowed down to the Yukawa sector, which gives masses to the left- and right-handed fermions.¹ The Yukawa interaction term is

$$\mathcal{L}_{\text{Yukawa}} = -\lambda_{ij}^d (\bar{Q}_L^{i,a} \Phi^a) d_R^j - \lambda_{ij}^u (\epsilon_{ab} \bar{Q}_L^{i,a} (\Phi^b)^\dagger) u_R^j + \text{h.c.}, \quad (1.28)$$

where the i, j indices are generation labels, Q_L^i represents the $\text{SU}(2)_L$ quark doublet and Φ represents the Higgs doublet (ϕ^+, ϕ^0) [24]. The components of the doublets are labelled by the a, b indices. Under \mathcal{CP} the left- and right-indices are exchanged, as are particle and anti-particle terms; the result is

$$\mathcal{CP}(\bar{Q}_L^i \cdot \Phi d_R^j) \rightarrow \bar{d}_R^j \cdot \Phi^\dagger Q_L^i \quad (1.29)$$

which maintains \mathcal{CP} -invariance if the Yukawa couplings λ are real. Hence \mathcal{CP} -violation takes place in the coupling of the scalar sector to the quark sector in the Standard Model. Noting that equation (1.28) includes generation-mixing interactions, the conventional parameterisation of the quark sector is to say that the physical particles are those which diagonalise the mass matrix and that there is no Yukawa mixing between different generations of physical quarks. The weak interaction quark states u' and d' are then superpositions of the physical quark fields u and d ,

$$u'_i = U_{ij}^u u_j \quad \text{and} \quad d'_i = U_{ij}^d d_j. \quad (1.30)$$

This transformation does not effect neutral interactions as the matrices $U^{u,d}$ are unitary, but charged weak interactions acquire a generational mixing term in the Lagrangian:

$$\mathcal{L}_{\text{CC}} = -\frac{g}{\sqrt{2}} J_{\text{CC}}^\mu W_\mu^\dagger + \text{h.c.}, \quad (1.31)$$

¹Here we are ignoring the “strong \mathcal{CP} ” problem, where \mathcal{CP} -violation occurs via a Lagrangian term proportional to $\theta_{\text{QCD}} \epsilon^{\mu\nu\rho\sigma} G_{\mu\nu}^a G_{\rho\sigma}^a$. θ_{QCD} has been experimentally constrained to be very close to zero and this term is largely considered to be unrealised in nature [25].

where the current J^μ is defined as [26]

$$J_{CC}^\mu = \begin{pmatrix} \bar{u} & \bar{c} & \bar{t} \end{pmatrix} \gamma^\mu (1 - \gamma^5) \mathbf{V}_{CKM} \begin{pmatrix} d & s & b \end{pmatrix}^T + \text{leptonic equivalent}. \quad (1.32)$$

In the hadronic part of this expression $\mathbf{V}_{CKM} \equiv (U^u)^\dagger U^d$ is the mixing matrix between down-type and up-type quarks, formed from the combination of the two matrices used to re-define the down-type and up-type fields between their mass- and weak-eigenstates. There is no equivalent construction to \mathbf{V}_{CKM} for the lepton fields due to the non-existence of right-handed neutrino states in the Standard Model. As both the component U matrices are unitary, \mathbf{V}_{CKM} is also unitary:

$$\mathbf{V}_{CKM}(\mathbf{V}_{CKM})^\dagger = \mathbf{1}. \quad (1.33)$$

\mathbf{V}_{CKM} is an extension of the 2×2 Cabbibo mixing matrix [27] to include a third generation of quarks, a step taken by Kobayashi and Maskawa, earning this the title of the Cabbibo-Kobayashi-Maskawa (CKM) matrix. It is the central construct of Standard Model \mathcal{CP} violation physics [28]. The CKM matrix can be written explicitly as

$$\mathbf{V}_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \quad (1.34)$$

where V_{ij} is the matrix element coupling the i th up-type quark to the j th down-type quark. The strengths of these couplings are not predicted by the Standard Model and must be experimentally constrained.

There are several popular parameterisations of the CKM matrix, each emphasising a different aspect. The first is the ‘‘Kang-Cheung’’ mixing-angle representation:

$$\mathbf{V}_{CKM} = \underbrace{\begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{\text{Cabibbo matrix}} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \times \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{-i\delta} & 0 & c_{13} \end{pmatrix} \quad (1.35)$$

where $c_{ij} \equiv \cos \theta_{ij}$, $s_{ij} \equiv \sin \theta_{ij}$ and δ is a non-removable phase, conventionally associated with 1–3 mixing [29]. This parameterisation is useful in that it emphasises the rôle of

the CKM matrix as a three-generation extension of the Cabibbo two-generation mixing mechanism: indeed the first term on the right hand side is identified with the Cabibbo matrix of two-generation quark mixing and θ_{12} is equivalent to the Cabibbo angle θ_c .

Since quark fields only enter in the Dirac Lagrangian in the forms $\bar{\psi}\psi$ and $\bar{\psi}\partial\psi$, the physics is invariant under field transformations of the form $q \rightarrow qe^{i\zeta}$. Such transformations can be used in the two-generation Cabibbo matrix to rotate away any “spare” phases, but for three or more generations non-removable complex phases will enter: δ is such a phase in the CKM matrix [30]. Its significance was first realised by Kobayashi and Maskawa in 1973, when they showed that such a phase could be used to incorporate direct \mathcal{CP} violation within the nascent Standard Model (the KM mechanism) [28].

Another popular parameterisation, introduced by Wolfenstein [31], is a phenomenological expansion in the Cabibbo angle, represented as $\lambda \equiv \sin \theta_{12}$ ². To write it compactly we introduce the parameters,

$$A \equiv \frac{s_{23}}{s_{12}^2}, \quad \rho \equiv \frac{s_{13} \cos \delta}{s_{12}s_{23}} \quad \text{and} \quad \eta \equiv \frac{s_{13} \sin \delta}{s_{12}s_{23}}. \quad (1.36)$$

An expansion in λ to $\mathcal{O}(\lambda^3)$ is most typically used:

$$\mathbf{V}_{\text{CKM}}^{(3)} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \quad (1.37)$$

Note that the matrix only acquires a non-zero imaginary component at 3rd order in λ , specifically in the V_{ub} and V_{td} terms — the smallness of Standard Model \mathcal{CP} -violation is made explicit in this formalism. Continuing the Wolfenstein expansion to $\mathcal{O}(\lambda^5)$ gives an extended form,

$$\mathbf{V}_{\text{CKM}}^{(5)} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 + \frac{1}{4}\lambda^4 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda + \frac{1}{2}A\lambda^4 - A^2\lambda^5(\rho + i\eta) & 1 - \frac{1}{2}\lambda^2 + \frac{1}{4}\lambda^4(1 - 2A^2) & A\lambda^2 \\ A\lambda^3(1 - \bar{\rho} - i\bar{\eta}) & -A\lambda^2 + A\lambda^4(\frac{1}{2} - \rho - i\eta) & 1 - \frac{1}{2}A^2\lambda^4 \end{pmatrix} + \mathcal{O}(\lambda^6), \quad (1.38)$$

²Note that the Wolfenstein λ parameter is not equivalent to the λ parameter used earlier to summarise \mathcal{CP} violation in the interference of mixing and decay amplitudes.

where $\bar{\rho} = (1 - \lambda^2/2)\rho$ and $\bar{\eta} = (1 - \lambda^2/2)\eta$. This extra precision will be required for the first time in LHC-era b-physics.

The existence of \mathcal{CP} violation in the Standard Model is elegantly summarised via the up-type and down-type Yukawa mass matrices, $M_i = v\lambda_i/\sqrt{2}$. It was shown by Jarlskog [32] that

$$\det [M_d, M_u] = \frac{2iJ_{\mathcal{CP}}F_dF_u}{m_b^3m_t^3} \quad (1.39a)$$

$$\text{and } \neq 0 \text{ for Standard Model } \mathcal{CP}\text{-violation,} \quad (1.39b)$$

where

$$F_\alpha = \prod_{\substack{i,j=1 \\ j>i}}^3 (m_i^2 - m_j^2). \quad (1.40)$$

The mass-squared differences in this relation represent the fact that \mathcal{CP} -violation occurs in the CKM matrix purely as a result of the single phase δ and that if quarks of the same charge also have the same mass then field re-definitions can be used to rotate this phase away. Hence Standard Model \mathcal{CP} -violation is intrinsically connected to the existence of a quark mass hierarchy.

$J_{\mathcal{CP}}$, the “Jarlskog determinant”, $J_{\mathcal{CP}} = \det \Im \mathbf{m}\{V_{i\alpha}V_{j\beta}V_{i\beta}^*V_{j\alpha}^*\}$, is also required by equation (1.39) to be non-zero. It is derived from the unitarity constraint on the CKM matrix and quantifies the strength of Standard Model \mathcal{CP} violation. In terms of the Euler and Wolfenstein parameterisations of the CKM matrix,

$$J_{\mathcal{CP}} = s_{12} s_{13} s_{23} c_{12} c_{13} c_{23} \sin \delta = \lambda^6 A^2 \eta = \mathcal{O}(10^{-5}), \quad (1.41)$$

where s_{ij} and c_{ij} respectively represent $\sin \theta_{ij}$ and $\cos \theta_{ij}$. The small size of $J_{\mathcal{CP}}$ gives quantitative support to the statement that \mathcal{CP} -violation is small in the Standard Model. However, it has recently been suggested that in cosmological scenarios of “tachyonic inflation” it is appropriate to replace the quark mass terms in (1.39) with the corresponding Yukawa couplings, greatly enhancing the effectiveness of quark sector \mathcal{CP} -violation in establishing a cosmological matter-antimatter asymmetry [33].

1.4.1 Unitarity triangles

The unitarity constraint on the CKM matrix can be expressed as a set of complex orthogonality conditions between each of the matrix columns and rows:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \quad (\text{db}) \quad (1.42a)$$

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0 \quad (\text{ds}) \quad (1.42b)$$

$$V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* = 0 \quad (\text{sb}) \quad (1.42c)$$

and

$$V_{ud}V_{td}^* + V_{us}V_{ts}^* + V_{ub}V_{tb}^* = 0 \quad (\text{ut}) \quad (1.42d)$$

$$V_{cd}V_{td}^* + V_{cs}V_{ts}^* + V_{cb}V_{tb}^* = 0 \quad (\text{ct}) \quad (1.42e)$$

$$V_{ud}V_{cd}^* + V_{us}V_{cs}^* + V_{ub}V_{cb}^* = 0 \quad (\text{uc}) \quad (1.42f)$$

where the quark pairs in parentheses on the right hand side indicates the pair of rows or columns across which the inner product is being taken. In addition there are six normalisation conditions for each of the rows and columns. The orthogonality constraints are well-known in b-physics for their convenient representation as triangles in the complex plane where each product of CKM matrix elements is a vector in the complex plane corresponding to a triangle side. The six triangles constructed this way are known as “unitarity triangles” and are shown in Figure 1.2 Each triangle has the same area, namely $J_{\mathcal{CP}}/2$, so the existence of triangles at all (as opposed to lines) is an indication of quark-sector \mathcal{CP} -violation.

The “db” and “ut” triangles differ from the other four since all their sides are of the same order in Wolfenstein’s λ parameter. These two triangles only differ at $\mathcal{O}(\lambda^5)$ and are the primary focus of \mathcal{CP} -violation measurement strategies. We define the triangles’ interior angles at order λ^3 as α , β and γ as shown in Figure 1.3. The angles are expressed in terms of CKM matrix elements as

$$\alpha \equiv \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \quad \beta \equiv \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \quad \gamma \equiv \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \quad (1.43)$$

At $\mathcal{O}(\lambda^5)$ a phase is introduced into V_{ts} such that the angles β and γ are shifted to

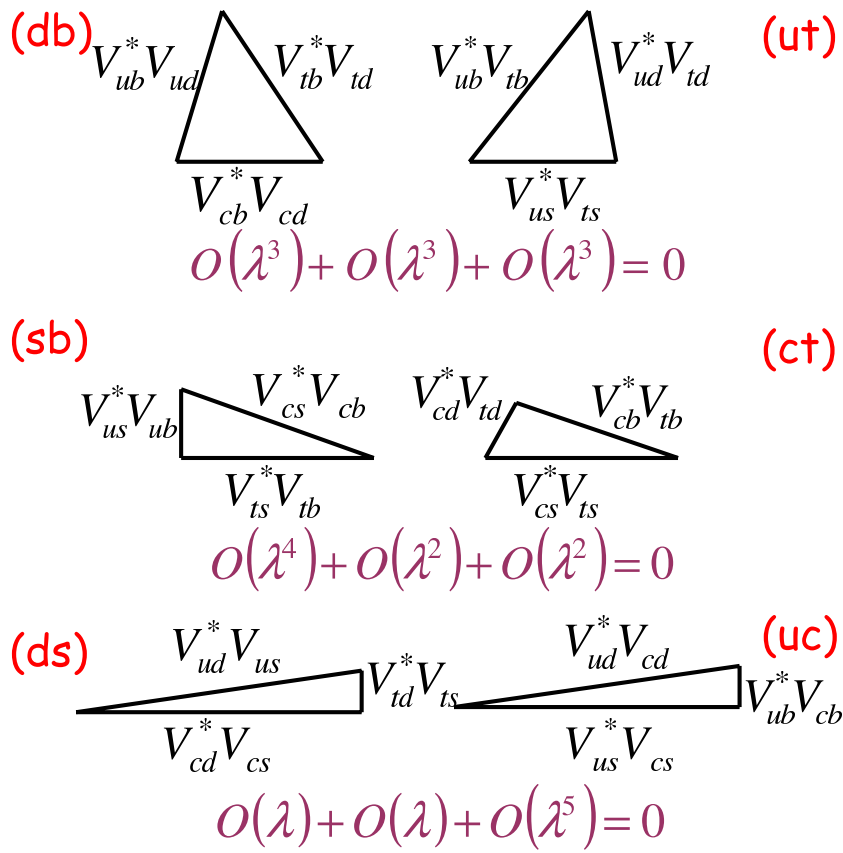


Figure 1.2: The six CKM “unitarity triangles”, labelled as in equation (1.42).

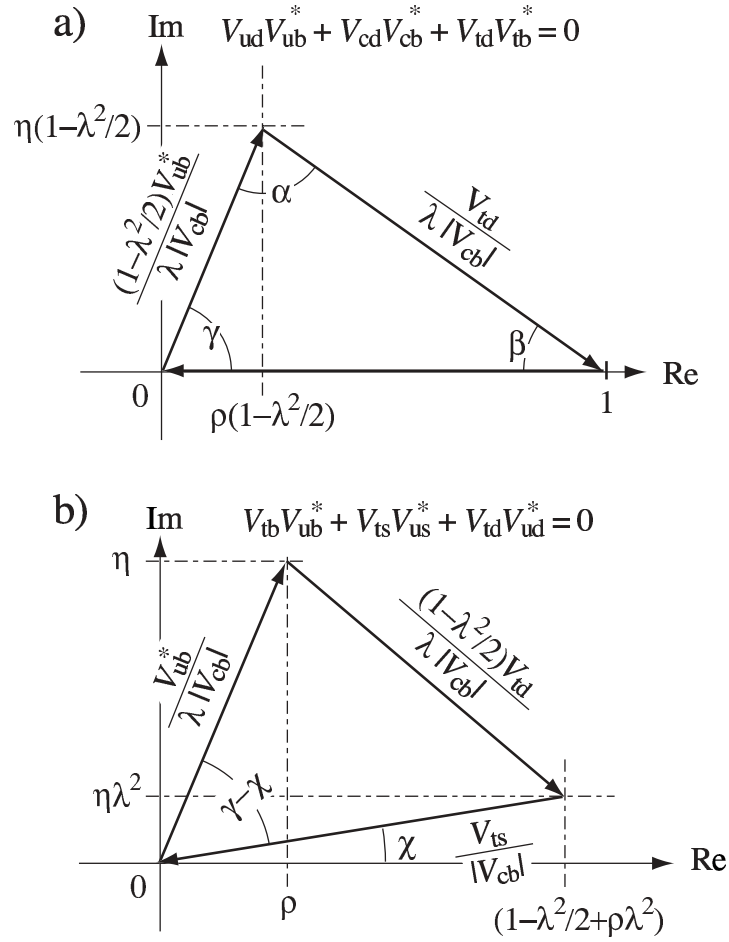


Figure 1.3: The two “unsquashed” unitarity triangles from equations (1.42a) and (1.42d) respectively. These triangles have been normalised relative to the baseline of the (db) triangle so that $V_{cd}V_{cb}^* = -1$.

β' and γ' respectively in the “ut” triangle, where

$$\beta' \equiv \arg\left(-\frac{V_{ts}V_{us}^*}{V_{td}V_{ud}^*}\right), \quad \gamma' \equiv \arg\left(-\frac{V_{tb}V_{ub}^*}{V_{ts}V_{us}^*}\right), \quad (1.44)$$

such that $\beta' = \beta + \chi$ and $\gamma' = \gamma - \chi$. The relationship between the angles and the structure of the CKM matrix is made more evident by taking the complex argument of the matrix:

$$\arg(\mathbf{V}_{\text{CKM}}) \approx \begin{pmatrix} 0 & 0 & -\gamma \\ 0 & 0 & 0 \\ -\beta & \chi + \pi & 0 \end{pmatrix} \quad (1.45)$$

where a small correction factor arising from the normalisation convention has been safely ignored in V_{cd} . By comparing equation (1.38) with this matrix of arguments we can identify the angles with the Wolfenstein parameters:

$$\beta^{(3)} = \arctan\left(\frac{\eta}{1 - \rho}\right), \quad \gamma^{(3)} = \arctan\left(\frac{\eta}{\rho}\right) \quad (1.46a)$$

and

$$\chi^{(5)} = \arctan\left(\frac{\lambda^2 \eta}{1 + \lambda^2(\rho - \frac{1}{2})}\right) \approx \lambda^2 \eta \quad (1.46b)$$

where the parenthetic superscripts indicate that β and γ are only expressed here in terms of the 3rd-order Wolfenstein parameterisation while a non-zero χ only enters at 5th order.

1.4.2 B– $\bar{\text{B}}$ mixing in the Standard Model

As the weak charged current interaction violates flavour conservation, box diagrams can be constructed which produce neutral flavour change as a second-order process. These are illustrated in Figure 1.4.

At first order in the weak coupling and ignoring the necessary QCD corrections the diagrams are mediated by CKM factors of the form $V_{ib}^* V_{iq} V_{jb}^* V_{jq}$ where $i, j \in \{\text{u}, \text{c}, \text{t}\}$ and q is the partner quark of the b in the B meson, i.e. $q = \text{d}$ in B^0 . The justification for this factor can be understood from the Feynman diagrams in Figure 1.4. The V_{ib}

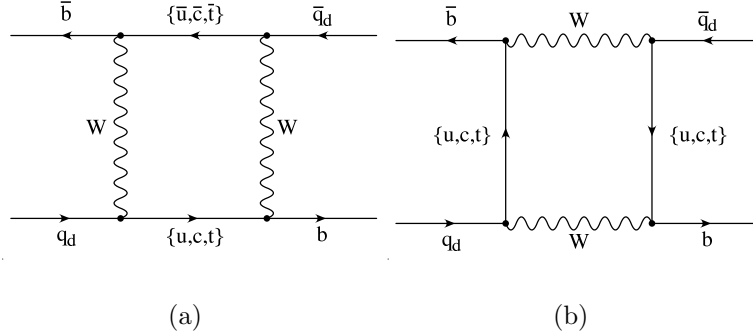


Figure 1.4: The two leading order Feynman box diagram topologies for Standard Model B - \bar{B} mixing. The q_d label refers to quarks of down-type which are not the b quark, i.e. down or strange.

factors mean that box diagrams are dominated by contributions from internal top quark propagation, since $V_{tb} \approx 1 \gg V_{cb}, V_{ub}$.

Calculation of the box diagram amplitudes requires input from several model-dependent constructions and reliance is placed on e.g. lattice QCD calculations. The indication from such calculations is that

$$\left| \frac{\Gamma_{12}}{M_{12}} \right| \approx \frac{3\pi}{2} \frac{m_b^2}{m_W^2} \frac{1}{S_0(m_t^2/m_W^2)} \sim \mathcal{O}\left(\frac{m_b^2}{m_t^2}\right), \quad (1.47)$$

which is small. This relates to the $|q/p|$ mixing-induced \mathcal{CP} -violation parameter through the expansion

$$\left| \frac{q}{p} \right|^2 = 1 + \left| \frac{\Gamma_{12}}{M_{12}} \right| \sin(\phi_M - \phi_\Gamma) + \mathcal{O}\left(\left| \frac{\Gamma_{12}}{M_{12}} \right|^2\right). \quad (1.48)$$

Hence $|q/p| \approx 1$ and mixing-induced \mathcal{CP} -violation is predicted to be very small in the b -sector. Lattice calculations predict that for the B_d^0 - \bar{B}_d^0 system $\Delta\Gamma/\Gamma \sim \mathcal{O}(1\%)$ while the value is in the region of 10% for the B_s^0 - \bar{B}_s^0 system, hence the decay rate asymmetry is negligible for B_d^0 mixing but not for B_s^0 .

1.4.3 Weak decays of B-mesons

For the purposes of this summary (and for most current b -physics phenomenology) we will restrict ourselves to the decays of B -mesons. These can be divided into three

broad categories called “leptonic”, “semi-leptonic” and “non-leptonic” for the characteristic transitions $B^- \rightarrow \ell^- \bar{\nu}_\ell$, $b \rightarrow c \ell^- \bar{\nu}_\ell$ and $b \rightarrow q_1 \bar{q}_2 q_d$ respectively (where $\ell \in \{e, \mu\}$; $q_1, q_2 \in \{u, d, s, c\}$ and $q_d \in \{d, s\}$, combined to conserve overall charge). Note that this notation is for a “black box” decay; the contributing Feynman diagram topologies will be described in Section 1.4.3.

Pure leptonic decays have branching ratios $\sim \mathcal{O}(10^{-10} - 10^{-4})$ [34, 35], with helicity conservation suppressing the lighter e^- and μ^- modes more strongly than the τ^- mode. They are largely of interest for potential observation of new physics effects and for experimentally constraining the B decay constant f_B , which is only weakly fixed by lattice gauge theory predictions [36]. Semi-leptonic decays are accessible in clean environments such as the b-factories and CLEO and are useful primarily for measuring the magnitudes of the CKM elements V_{cb} and V_{ub} [10] but they do not provide a particularly powerful mechanism for extracting CKM phases. The non-leptonic decays, which directly involve CKM transitions at the quark level and offer many possibilities for constraining phases, are the dominant channels of interest for \mathcal{CP} -violation studies at the LHC. From now on the primary emphasis of this review will be on the phenomenology of non-leptonic decays.

Tree and penguin diagrams

The leading-order Feynman diagrams which contribute to non-leptonic B-decays are divided into “tree” and “penguin” topologies, examples of which are illustrated in Figure 1.5. Tree diagrams are those where the W-boson responsible for the b-quark flavour change decays directly to a final-state quark-antiquark pair. This includes the process where the B-meson constituent quarks annihilate and the resulting W boson is responsible for producing the quarks which hadronise to the final state mesons. These processes are shown in Figures 1.5(a), 1.5(c) and 1.5(e).

Conversely, penguin diagrams are those in which the W-boson re-connects to the quark line from which it was emitted and the quark-antiquark pair in the final state is generated by the decay of a boson emitted from the quark part of the penguin loop. This is shown in e.g. Figure 1.5(b). The penguin topology can be further classified by the identity of the virtual boson emitted from the loop: if it is a gluon then the diagram is classed as a “gluonic” (or “QCD”) penguin. An internal photon or Z-boson emission characterises an “electroweak” penguin.

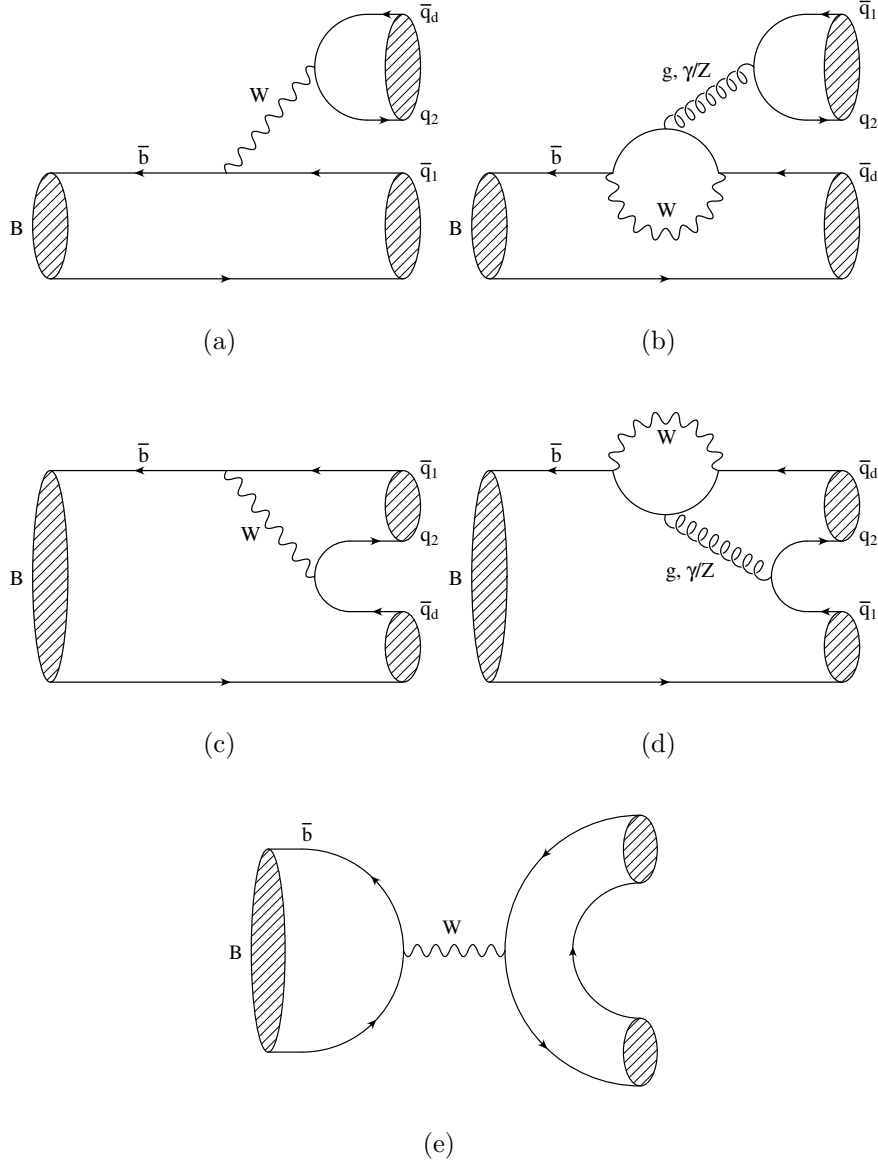


Figure 1.5: Tree and penguin Feynman diagram topologies for the (a) colour-allowed tree, (b) colour-allowed penguin, (c) colour-suppressed tree (d) colour-suppressed penguin and (e) annihilation topologies. The final state quarks are labelled as q_1 , q_2 or q_d as used for the definition of non-leptonic B-decays in all except the annihilation topology (for which the convention is not completely constraining).

A further classification may be made for both tree and penguin topologies based on the interaction of the final state quark-antiquark pair with the spectator quark and the b-quark line. If the hadronisation process forms the two final-state mesons by hadronising between the b-quark/spectator quark lines and between the remaining $q\bar{q}$ pair then the process is “colour allowed” since the colour labels of the pair are unconstrained. Conversely, hadronising to two mesons formed from the (q, q_d) and $(\bar{q}, \text{spectator})$ combinations is denoted as a “colour suppressed” decay. The differences between colour-allowed and colour-suppressed topologies are shown in Figure 1.5, with q_1 and q_2 indicated for each diagram.

The penguin and tree topologies differ in the number of weak charged interactions on the b-quark line: the tree diagram only gives a single $b \rightarrow \{u, c\}$ transition while the penguin b-line undergoes a pair of transitions: $b \rightarrow \{u, c, t\} \rightarrow \{d, s\}$. Hence the flavours of the q_1 and q_2 quarks constrain the contributing set of diagrams:

- $q_1 \neq q_2 \in \{u, c\}$: only tree diagrams contribute;
- $q_1 = q_2 \in \{u, c\}$: both tree and penguin diagrams contribute;
- $q_1 = q_2 \in \{d, s\}$: only penguin diagrams contribute.

1.5 Strategies to constrain CKM parameters

Current experimental knowledge of CKM parameters is dominated by indirect measurements of the matrix elements V_{ij} and the b-factory measurements of $\sin(2\beta)$. The angles α and γ are under active investigation by the b-factory experiments but are not expected to be truly constrained until the LHC starts running.

Current status of CKM matrix element measurements

The magnitudes of seven of the nine CKM matrix elements, $|V_{ud}|$, $|V_{us}|$, $|V_{ub}|$, $|V_{cd}|$, $|V_{cs}|$, $|V_{cb}|$ and $|V_{tb}|$ are determined from nuclear β -decay and the decay rates of pions, kaons, hyperons, D and B mesons and the top quark [37], evaluated to first order in weak interactions. The remaining two elements, $|V_{td}|$ and $|V_{ts}|$, are accessed through the neutral K and B oscillation box diagrams and rare decays which are mediated by second-order weak effects such as QCD “penguin” diagrams. The 90% confidence levels

from these measurements give ranges for the CKM matrix element magnitudes [10],

$$|\mathbf{V}_{\text{CKM}}| = \begin{pmatrix} 0.9739 - 0.9751 & 0.221 - 0.227 & 0.0029 - 0.0045 \\ 0.221 - 0.227 & 0.9730 - 0.9744 & 0.039 - 0.044 \\ 0.0048 - 0.014 & 0.037 - 0.043 & 0.9990 - 0.9992 \end{pmatrix} \quad (1.49)$$

and the corresponding 90% confidence intervals on the sines of the mixing angles [10],

$$s_{12} = 0.2243 \pm 0.0016, \quad s_{23} = 0.0413 \pm 0.0015, \quad s_{13} = 0.0037 \pm 0.0005, \quad (1.50)$$

with the CKM phase being constrained to $\delta_{13} = 1.05 \pm 0.24 \text{ rad} = 60^\circ \pm 14^\circ$ when loop level processes are included as fit constraints [10].

Indirect measurements

A standard indirect analysis of CKM parameters has now become established to constrain the apex of the unitarity triangle in the $\bar{\rho}-\bar{\eta}$ plane. This analysis uses many inputs, but the dominant contributions are

- magnitude constraints on $|V_{ub}|$ and $|V_{cb}|$ from exclusive and inclusive semi-leptonic B-decays containing $b \rightarrow u \ell \bar{\nu}_\ell$ and $b \rightarrow c \ell \bar{\nu}_\ell$ transitions. These fix a circle around $(0, 0)$;
- $B^0-\bar{B}^0$ mixing, for both d- and s-type B-mesons, can be used to fix a circle around $(1, 0)$; and
- constraints on the ϵ_K parameter, which characterises indirect \mathcal{CP} -violation in the kaon system, define a hyperbola passing through the triangle apex.

These constraints are shown schematically in Figure 1.6 and the latest global fit as prepared by the CKM Fitter group is shown in Figure 1.7 [38].

The allowed ranges for the CKM angles from this fit are, at a confidence level corresponding to 1σ ,

$$84^\circ \leq \alpha \leq 106^\circ, \quad 21.8^\circ \leq \beta \leq 25.9^\circ, \quad 50^\circ \leq \gamma \leq 72^\circ \quad (1.51)$$

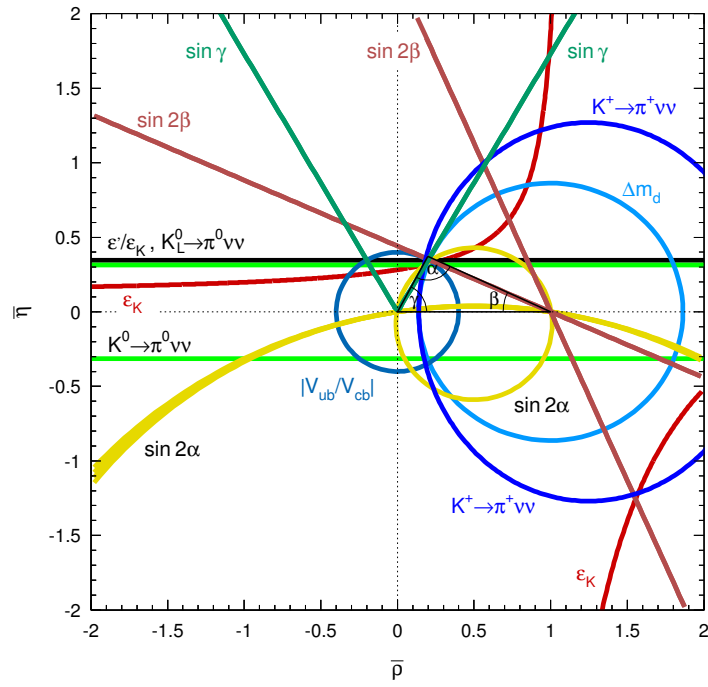


Figure 1.6: Contributions to the indirect fit of CKM parameters.

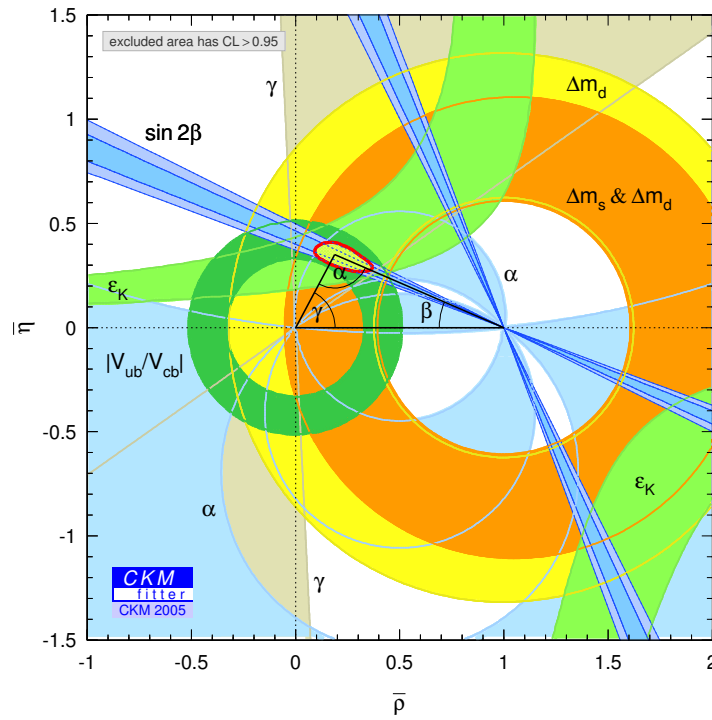


Figure 1.7: CKM Fitter group global indirect fits to the CKM parameters using the b-factory direct angle measurements as of May 2005.

where the precision of the β measurement relative to the other two angles is evident. The values for $\sin(2\beta)$ obtained from direct and indirect measurements are in excellent agreement, so there is at present no significant deviation from the Standard Model picture.

Direct measurement of β

β was the first of the CKM unitarity triangle angles to be measured directly. This was achieved through studying the \mathcal{CP} asymmetries of $B_d^0 \rightarrow J/\psi K_S^0$ decays and was the first observation of \mathcal{CP} -violation in the b-system [39]. This mode and its \mathcal{CP} conjugate are known as the “gold-plated” mode for measuring β because it is clean from both experimental and theoretical perspectives. The final state is a \mathcal{CP} eigenstate to which both B_d^0 and \bar{B}_d^0 can decay and so \mathcal{CP} -violation is observed in the interference of the direct and indirect decays via B_d^0 – \bar{B}_d^0 mixing. The phenomenological quantity of interest is the time-dependent \mathcal{CP} asymmetry [40]

$$\mathcal{A}_{\mathcal{CP}}(B_d^0 \rightarrow J/\psi K_S^0) \equiv \frac{\Gamma(B_d^0 \rightarrow J/\psi K_S^0 : t) - \Gamma(\bar{B}_d^0 \rightarrow J/\psi K_S^0 : t)}{\Gamma(B_d^0 \rightarrow J/\psi K_S^0 : t) + \Gamma(\bar{B}_d^0 \rightarrow J/\psi K_S^0 : t)} = \sin(2\beta) \sin(\Delta m_d t). \quad (1.52)$$

Here the rates are all expressed in terms of the decay occurring at proper time t and Δm_d is the B_d^0 -system oscillation frequency. Comparison with equation (1.25) reveals that

$$\Im\{\lambda\} = \sin(2\beta) \quad (1.53)$$

where λ was defined as the characteristic parameter for \mathcal{CP} -violation in mixing in equation (1.17). Measurements of β in this way have been made at the LEP experiments Aleph and OPAL [41, 42], the Tevatron CDF and DØ experiments [43, 44] and what are now the dominantly precise measurements, those from the b-factory experiments *BABAR* and *Belle* [45, 46]. The current combined measurement of $\sin 2\beta$ from the b-factories is [10]

$$\sin 2\beta = 0.736 \pm 0.049 \implies \beta = 47^\circ \pm 5^\circ \quad (1.54)$$

Direct measurement of α

α is the second CKM angle to be measured directly at the b-factories. Three sets of decay channels — $B \rightarrow \rho\rho$, $B \rightarrow \rho\pi$ and $B \rightarrow \pi\pi$ — can be used to constrain α [47]. Initially, a $B \rightarrow \pi\pi$ strategy, which is used to constrain α via CKM unitarity and neglecting penguin contributions, was thought to be the most promising route to measuring α . This predicted the following relations:

$$\mathcal{A}_{\mathcal{CP}}^{\text{dir}}(B_d^0 \rightarrow \pi^+ \pi^-) = 0, \quad (1.55)$$

$$\mathcal{A}_{\mathcal{CP}}^{\text{mix}}(B_d^0 \rightarrow \pi^+ \pi^-) = \sin(2\beta + 2\gamma) = -\sin 2\alpha \quad (1.56)$$

where the final equality makes use of CKM unitarity. This approach has been rendered secondary by the difficulties of dealing with penguin pollution. The most prominent approach to solving this problem is a general isospin analysis, requiring the averaged branching ratios and asymmetries of $B_d^0 \rightarrow \pi^+ \pi^-$, $B_d^0 \rightarrow \pi^0 \pi^0$ and $B^+ \rightarrow \pi^+ \pi^0$ decays [48, 49]. Unfortunately, the parameters describing the decay to the pair of neutral pions are plagued by experimental uncertainties which prevent the use of this method to constrain α [50, 51]. If we accept the *BABAR* results on $B \rightarrow \pi\pi$ decays as definitive [52], then the $B \rightarrow \pi\pi$ constraint on α is

$$\sin 2\alpha = -0.30 \pm 0.17 \implies \alpha = 99^\circ \pm 5^\circ, \quad (1.57)$$

but this degree of accuracy cannot be quoted when the disagreement between the *BABAR* and Belle results is accounted for.

In recent times, measurements of α via final states containing ρ mesons have become popular. These rely on the assumption that penguin contributions to charmless, strangeless B-decays are relatively small, justified by the measured relative smallness of the $B_d^0 \rightarrow \rho^0 \rho^0$ branching ratio [53, 54, 55, 56, 57]. The first such approach uses time-dependent \mathcal{CP} -asymmetries of $B_d^0 \rightarrow \rho^+ \rho^-$ decays, using

$$\frac{\Gamma(\bar{B}_d^0 \rightarrow \rho^+ \rho^-) - \Gamma(B_d^0 \rightarrow \rho^+ \rho^-)}{\Gamma(\bar{B}_d^0 \rightarrow \rho^+ \rho^-) + \Gamma(B_d^0 \rightarrow \rho^+ \rho^-)} = -C_{\rho^+ \rho^-} \cos(\Delta M \Delta t) + S_{\rho^+ \rho^-} \sin(\Delta M \Delta t), \quad (1.58)$$

where Γ is the time-dependent decay rate and

$$C_{\rho^+ \rho^-} = \frac{1 - |\lambda|^2}{1 + |\lambda|^2} \quad \text{and} \quad S_{\rho^+ \rho^-} = \frac{2 \Im\{\lambda\}}{1 + |\lambda|^2}, \quad (1.59)$$

with λ as defined earlier. From the experimental data [57],

$$C_{\rho^+\rho^-} = -0.23 \pm 0.24 \pm 0.14 \quad (1.60a)$$

$$S_{\rho^+\rho^-} = -0.19 \pm 0.33 \pm 0.11, \quad (1.60b)$$

which are compatible with zero penguin contribution ($C_{\rho^+\rho^-} = 0$) we obtain

$$S_{\rho^+\rho^-} = \sin 2\alpha = -0.19 \pm 0.35 \implies \alpha = 95^\circ \pm 10^\circ. \quad (1.61)$$

The final strategy we shall mention for directly obtaining a measurement of α is that using $B \rightarrow \rho \pi$ decays. The time-dependence of these decays is given by

$$\begin{aligned} \Gamma(\{B_d^0, \bar{B}_d^0\} \rightarrow \rho^\pm \pi^\mp) &= (1 \pm \mathcal{A}_{\mathcal{CP}}^{\rho\pi}) e^{-\Delta t/\tau} \times \\ &\quad [1 - q(C_{\rho\pi} \pm \Delta C_{\rho\pi}) \cos(\Delta M \Delta t) + q(S_{\rho\pi} \pm \Delta S_{\rho\pi}) \sin(\Delta M \Delta t)], \end{aligned} \quad (1.62)$$

where $q = \{-1, 1\}$ for B_d^0 and \bar{B}_d^0 decays at $t = 0$ respectively. The other quantities in equation (1.62) are defined as

$$\mathcal{A}_{\mathcal{CP}}^{\rho\pi} = \frac{|A^{+-}|^2 - |A^{-+}|^2 + |\bar{A}^{+-}|^2 - |\bar{A}^{-+}|^2}{|A^{+-}|^2 + |A^{-+}|^2 + |\bar{A}^{+-}|^2 + |\bar{A}^{-+}|^2}, \quad (1.63a)$$

$$C_{\rho\pi} \pm \Delta C_{\rho\pi} = \frac{|A^{\pm\mp}|^2 - |\bar{A}^{\pm\mp}|^2}{|A^{\pm\mp}|^2 + |\bar{A}^{\pm\mp}|^2}, \quad \text{and} \quad (1.63b)$$

$$S_{\rho\pi} \pm \Delta S_{\rho\pi} = \frac{2 \Im\{\lambda\}}{1 + |\lambda|^2}. \quad (1.63c)$$

In equation (1.63), terms like e.g. A^{+-} represent the decay amplitude for $B_d^0 \rightarrow \rho^+ \pi^-$ and so-on, and e.g. \bar{A}^{+-} represents the amplitude for $\bar{B}_d^0 \rightarrow \rho^+ \pi^-$. Decomposing these amplitudes into penguin and tree contributions, we can relate to physical observables and obtain the following expression (after an approximation noting that $\Delta C_{\rho\pi}$ is consistent with zero):

$$S_{\rho\pi} = \sqrt{1 - (\Delta C_{\rho\pi})^2} \sin 2\alpha. \quad (1.64)$$

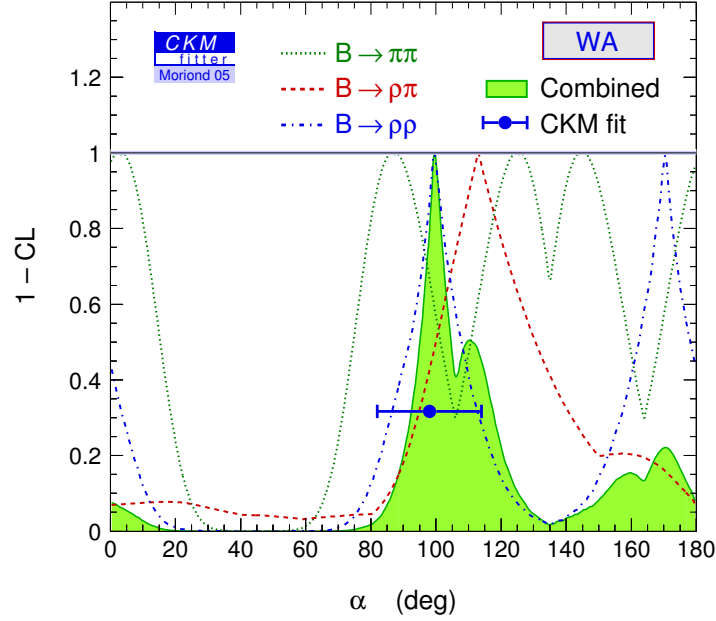


Figure 1.8: CKM Fitter confidence interval for α combined fit.

Using the averaged experimental results,

$$(S_{\rho\pi})_{\text{avg}} = -0.15 \pm 0.13, \quad (1.65a)$$

$$(\Delta C_{\rho\pi})_{\text{avg}} = 0.22 \pm 0.10, \quad (1.65b)$$

we finally obtain

$$\alpha = 94^\circ \pm 4^\circ. \quad (1.66)$$

Combining all three of these methods, we obtain the averaged result that

$$\alpha = 97^\circ \pm 3^\circ, \quad (1.67)$$

where the uncertainty is entirely experimental. It is argued in reference [47] that the theoretical uncertainty is probably on a similar scale to the experimental figure. Bearing this in mind, current precision on α is similar to that on β . The CKM Fitter confidence levels as functions of α from combining the measurements from these approaches are shown in Figure 1.8.

Direct measurement of γ

γ is unique among the CKM parameters in that it will remain only weakly probed until the LHC comes into operation. It is hence the angle on which most LHC-era b-physics is concentrated. The Belle collaboration has published results of studies to extract γ from $B_d^0 \rightarrow D^{0(*)}K$ decays via a Dalitz analysis, giving $\gamma = 68^{+14}_{-15}(\text{stat}) \pm 13^\circ(\text{syst}) \pm 11^\circ(\text{model})$ [58] from which the 95% confidence interval is $22^\circ < \gamma < 113^\circ$. *BABAR* has also carried out γ studies, but has yet to publish any measurement of γ [59]. The CKM Fitter group find $50^\circ < \gamma < 72^\circ$, derived from the best indirect fits to the Wolfenstein $\bar{\rho}$ and $\bar{\eta}$ parameters. No evidence of \mathcal{CP} -violation has been seen at the Tevatron Run-II experiments.

The standard approaches to measurement of γ can be categorised in four groups, summarised in this section. These are built around time-dependent asymmetries, time-integrated amplitude relations, isospin relations and U-spin relations. The first two of these are governed at leading order by tree-level Feynman diagram topologies, while the second pair also include contributions from penguin topologies. The symmetry relation methods are therefore potentially sensitive to new physics effects introduced by the penguin loops.

Time-dependent asymmetries: By mixing the tree diagrams $\bar{b} \rightarrow \bar{u}W^+$ and $\bar{b} \rightarrow \bar{c}W^+$ a theoretically clean measurement of the quantity $\gamma - \chi$ may be obtained, where χ is the $B_s^0 - \bar{B}_s^0$ mixing phase introduced at $\mathcal{O}(\lambda^5)$ in the CKM matrix. χ can be obtained from time-dependent asymmetries of $\{B_s^0, \bar{B}_s^0\} \rightarrow J/\psi \phi$ decays.

These tree diagrams correspond to combining the hadron-level decays of B_s^0 into $D_s^+K^-$ and $D_s^-K^+$ final states and the corresponding \mathcal{CP} -conjugate decays. LHCb studies of this approach indicate that one year of LHCb data-taking will constrain γ to within $\sim 15^\circ$, for $55^\circ < \gamma + \chi < 105^\circ$ — this channel is considered useful as a probe of new physics contributions due to the sensitivity to χ [60].

B_d^0 decays can also be used to obtain a constraint on γ . The process equivalent to $B_s^0 \rightarrow D_s^\pm K^\mp$ in the B_d^0 system is $B_d^0 \rightarrow D^\pm \pi^\mp$ which enables direct extraction of the combination $\gamma + 2\beta$, where 2β is the $B_d^0 - \bar{B}_d^0$ oscillation phase, analogous to the χ term above. This approach gives a sensitivity to γ of $\sim 14^\circ$ [60]. The angle β is expected to be strongly constrained by the “gold plated” $B_d^0 \rightarrow J/\psi K_s^0$ direct measurement.

Time-integrated amplitude relations: A second approach to determining γ is obtained via interference between the $\bar{b} \rightarrow \bar{u}W^+$ and $\bar{b} \rightarrow \bar{c}W^+$ tree amplitudes. The interference can be observed by measuring the three time-integrated decay rates for

$B_d^0 \rightarrow D^0 K^{*0}$, $B_d^0 \rightarrow \bar{D}^0 K^{*0}$ and $B_d^0 \rightarrow D_{\mathcal{CP}+}^0 K^{*0}$, where $D_{\mathcal{CP}+}^0 = (D^0 + \bar{D}^0)/\sqrt{2}$ indicates the \mathcal{CP} -even eigenstate of the D^0 – \bar{D}^0 system. The value of γ obtained by this method is therefore sensitive to new physics in D^0 – \bar{D}^0 mixing. Again, this method has been studied by LHCb with simulations indicating that a γ resolution of 7 – 8° is possible for one year of data-taking for $55^\circ < \gamma < 105^\circ$ [61, 60].

Isospin symmetry relations: The third method for determining γ is based on the approximate $SU(2)$ isospin symmetry of the strong interactions under exchange of u and d quarks. As the masses of the u and d quarks are very similar, this symmetry is considered reliable for phenomenology. $B_d^0 \rightarrow K^\pm \pi^\mp$ and $B^\pm \rightarrow K^0 \pi^\pm$ decays have both penguin and tree Feynman diagram contributions at leading order, which can be related under isospin symmetry. However, the theoretical uncertainties are dominated by the difficulty of performing long-range QCD calculations on re-scattering effects in the light hadronic final states. Due to the penguin diagram amplitudes, these decays are also a promising environment in which to observe new physics. The phenomenology of γ -determination from $B_d^0 \rightarrow K^\pm \pi^\mp$ and $B^\pm \rightarrow K^0 \pi^\pm$ decays is the subject of this thesis.

U-spin symmetry relations: The decay $B_d^0 \rightarrow \pi^+ \pi^-$ receives contributions from both the tree diagram $\bar{b} \rightarrow \bar{u} W^+$ and the penguin processes $\bar{b} \rightarrow \bar{d} \{g, \gamma/Z^0\}$. Under exchange of all d and \bar{d} quarks with s and \bar{s} quarks respectively (the “U-spin” flavour symmetry) we obtain the conjugate decay $B_s^0 \rightarrow K^+ K^-$. If the strong interaction is assumed invariant under the U-spin transformation then the ratios of tree to penguin contributions are also invariant under U-spin and γ can be extracted using constraints on χ and β from B_s^0 and B_d^0 mixing measurements respectively.

LHCb expects to obtain 4 – 6° resolution on γ from one year of data-taking, assuming the validity of U-spin symmetry which is subject to theoretical uncertainties about $SU(3)$ flavour symmetry breaking.

1.6 Physics from beyond the Standard Model

Much of the standard approach to measuring the parameters of the b -sector assumes the Standard Model formalism with the particle content of three generations of quarks and leptons coupled via four electroweak bosons. However, much effort is being devoted elsewhere to searching for physics contributions from beyond the Standard Model (BSM) at the TeV scale. Typically these take the form of extensions such as supersymmetry

(SUSY) or more exotic approaches to the hierarchy problem and associated theoretical defects of the Standard Model [62, 63, 64].

A generic prediction of such theories is the existence of flavour-changing neutral currents (FCNC) at tree level. These are only accessible via box-diagram loops in the Standard Model and so the rates for such processes are often greatly enhanced by new physics which can provide FCNC effects without the loop suppression. So-called penguin-dominated decays are often considered to be an ideal environment for observing new physics effects as they contain a loop in the leading order Feynman diagram. BSM particles can propagate off-shell in the loop, providing sensitivity to physics at higher energies than those which can be directly accessed by accelerators.

The B_s^0 meson system is a critical region of the b-sector from the point of view of BSM studies. Extraction of γ can be obtained from studies of $B_s^0 \rightarrow D_s^\pm K^\mp$ which are pure tree-level processes and unaffected by new physics or, alternatively, a combined analysis of $B_d^0 \rightarrow \pi^+ \pi^-$ and $B_s^0 \rightarrow K^+ K^-$ can be used to measure γ with loop-level sensitivity to new physics. Alternatively, the standard b-factory methods to measure γ making use of e.g. $B_d^0 \rightarrow D^0 K^*$ can be sensitive to new physics in the D^0 – \bar{D}^0 mixing. Combining these measurements should make extraction of γ possible even in the presence of new physics, but the direct measurements of the effective γ in the modes affected by new physics may be shifted substantially from the Standard Model values.

\mathcal{CP} -violation parameters themselves are not necessarily affected by the advent of new physics — for example, the minimal supergravity (mSUGRA) SUSY scenario is flavour-diagonal [65]. However, it is possible for TeV-scale new physics to have effects on certain \mathcal{CP} -violating decay channels such that inconsistent unitarity triangles are constructed. Cross-checks to measure triangle parameters via several independent methods are essential to remove systematic effects due to new BSM physics. Considering the CKM matrix in terms of the four phases, it has been shown that verifying the relationship

$$\sin \chi = \frac{\lambda^2 \sin \beta \sin \gamma}{\sin(\beta + \gamma)} \quad (1.68)$$

is a sensitive test of the validity of the Standard Model [66, 67].

b-physics is, relatively speaking, a sideline in terms of direct BSM studies but as the LHC data analyses mature, it offers a degree of precision study of the BSM flavour sector which is hard to come by elsewhere. \mathcal{CP} -violation is also considered to be of importance in, for example, SUSY scenarios where the Higgs sector is expanded to include \mathcal{CP}

eigenstate Higgses [68].

Chapter 2

The LHCb experiment

“There, sir! that is the perfection of vessels!”

— Jules Verne, 1828–1905

The LHCb experiment, to be operated at the CERN Large Hadron Collider (LHC), will be the world’s only dedicated hadron collider b-physics experiment. In this chapter we briefly describe the previous experiments to study the b-sector, introduce the LHC and describe LHCb itself in some detail.

2.1 b-physics experiments

LHCb is the latest in a line of experiments studying the physics of the b-sector. Initial studies of $b\bar{b}$ production, decay cross-sections and \mathcal{CP} -asymmetries were performed by the CLEO experiment on the CESR ring at Cornell University [69] and by the four LEP experiments [70, 71, 72, 73, 74, 75]. The symmetric e^+e^- colliders also measured the masses, lifetimes and mixing parameters of several b-hadrons, including the B_s^0 [76]. The $\Upsilon(4S)$ has been used since 1999 by the BABAR [77] and Belle [78] asymmetric e^+e^- collider experiments as a b-quark production mechanism. These “b-factory” experiments were the first to observe \mathcal{CP} -violation in the b-sector [79, 80, 81] and are currently the dominant contributors to measurements of B-meson \mathcal{CP} -violation [82].

Hadron collider b-physics has so far been performed by the general purpose CDF and DØ experiments on the Tevatron at Fermilab near Chicago which operates with a centre

of mass energy $\sqrt{s} \sim 2$ TeV. Since the $\Upsilon(4S)$ resonance is not heavy enough to produce B_s^0 mesons, the Tevatron experiments are the only place to study B_s^0 decays and mixing parameters until the start of LHC running [83]. The Tevatron is also an important place to constrain the parameters of hadron collider b-production mechanisms for use at the LHC: theoretical predictions underestimate the Tevatron b-production cross-section by a factor of ~ 2.4 due to either NNLO QCD effects, non-perturbative QCD effects or possibly new physics effects [84].

2.2 The LHC

The Large Hadron Collider (LHC) at CERN is a new hadron collider, located in the same tunnel as the Large Electron-Positron collider (LEP) [85]. Where LEP's chief task was the use of 90–207 GeV e^+e^- collisions to establish the precision physics of electroweak unification, the primary aim of the LHC is to discover the mechanism by which the electroweak symmetry is broken [86, 87]. For this, the LHC will collide protons at a centre of mass energy of 14 TeV with a luminosity of $\sim 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ [88]. A prominent difference between the LHC and the Tevatron is that the latter uses $p\bar{p}$ collisions. This is preferable logistically, since only one set of bending and focusing magnets and cavities is required to produce two counter-rotating beams, but anti-protons cannot be recycled rapidly enough for steady-state operation at the LHC. Hence a two-beam design with 10 T superconducting magnets is used instead, as shown in Figure 2.1 [88].

The 27 km-long LEP tunnel is to be re-used for the LHC machine, along with some of the experiment caverns. Four main detector experiments will take data from LHC collisions. These are ATLAS, CMS, ALICE and LHCb, of which the first two are general-purpose detectors and the remaining two are more specialised (for heavy ion and b-physics studies respectively) [89, 90, 91, 92]. The locations of the LHC experiments around the accelerator ring are shown in Figure 2.2.

The scheduled start date for LHC data-taking is Summer 2007 at which time a lower luminosity of $\sim 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ will be used [88]. This initial restricted luminosity will be upgraded to the design luminosity over the following year. Each of the counter-rotating beams contains 2835 bunches, each of which contains $\sim 10^{11}$ protons [88, 93]. The design bunch-crossing period at each of the experimental interaction points is 25 ns, but in the commissioning phase the machine will be operated with a 75 ns bunch crossing interval. To compensate for the reduced luminosity, a higher squeeze factor may be applied to

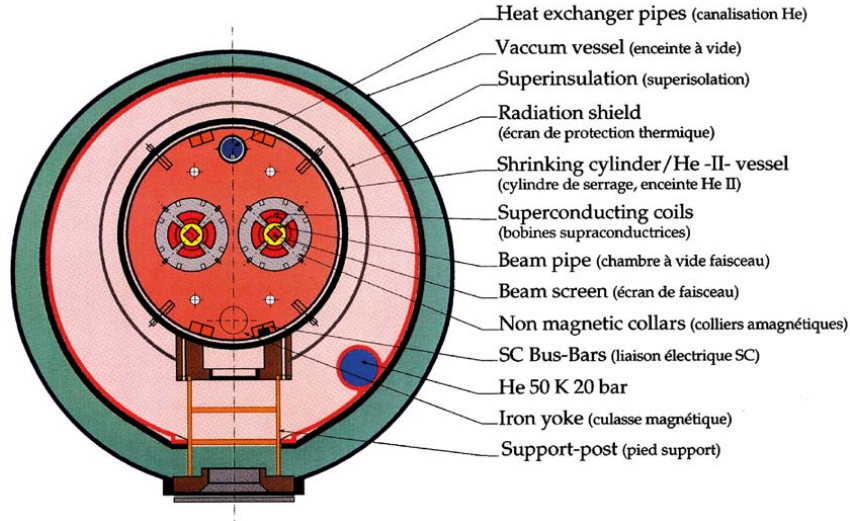


Figure 2.1: LHC beam-line cross-section, showing the double-aperture design required to accelerate two sets of like-charged particles, either protons or heavy ions.

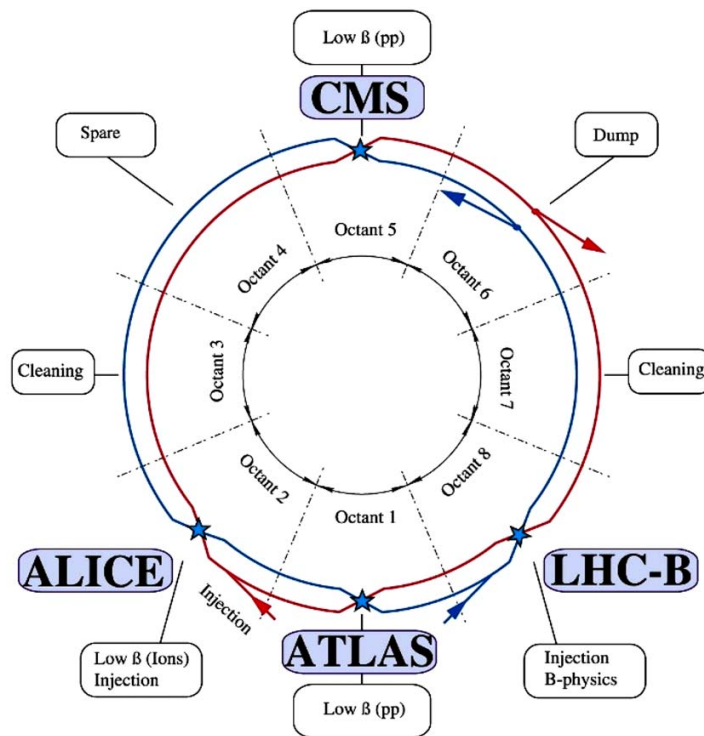


Figure 2.2: Locations of the four main LHC experimental sites around the LHC ring.

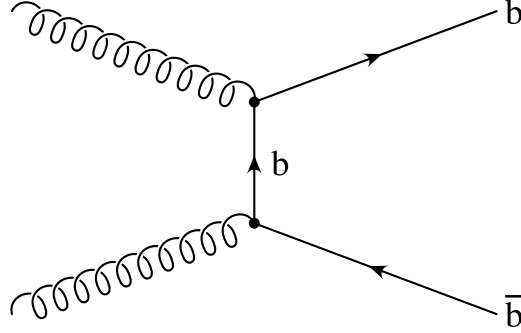


Figure 2.3: $b\bar{b}$ production by gluon fusion: this is the dominant LHC $b\bar{b}$ production mechanism.

increase the number of interactions per bunch crossing.

2.3 b-production at the LHC

The LHC will be the world’s most intense source of b-hadrons (those containing a b quark or \bar{b} anti-quark), producing $\sim 10^{12}$ $b\bar{b}$ pairs per year at the LHCb operational luminosity of $10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ [92]. This is equivalent to ~ 800 Hz during LHC running. The high centre of mass energy means that a complete spectrum of b-hadrons will be available, including B_d^0 , B^+ , B_s^0 , B_c^+ mesons and Λ_b baryons.

The dominant $b\bar{b}$ production mechanism at the LHC will be through gluon fusion [94] as shown in Figure 2.3. The momenta of the incoming partons are strongly asymmetric and are large compared to the b-quark mass. The result is that the outgoing $b\bar{b}$ pair is boosted with respect to the lab frame [92] and so both b-hadrons are dominantly produced in the same forward cone. This may be seen in Figure 2.4, which illustrates the correlation of the polar angles of the b- and \bar{b} -hadrons. The presence of both hadrons in the same cone is useful for “flavour tagging” (see Section 2.7).

V_{tb} , the dominant CKM matrix element in the weak decays of b-hadrons, is kinematically disallowed due to the large top quark mass $\sim \mathcal{O}(175 \text{ GeV}/c)$. As a result, B decays proceed at tree-level via CKM-suppressed V_{cb} and V_{ub} transitions, giving b-hadrons a relatively large proper lifetime, $\tau_b \sim \mathcal{O}(1 \text{ ps})$. Combined with the b-production Lorentz boost, this makes possible the “b-tagging” of the displacement of the b-hadron decay vertex from the interaction point. Essentially, the dynamics of proton–proton b-production has the same boosting effect as the asymmetric beam energies used by the b-factories.

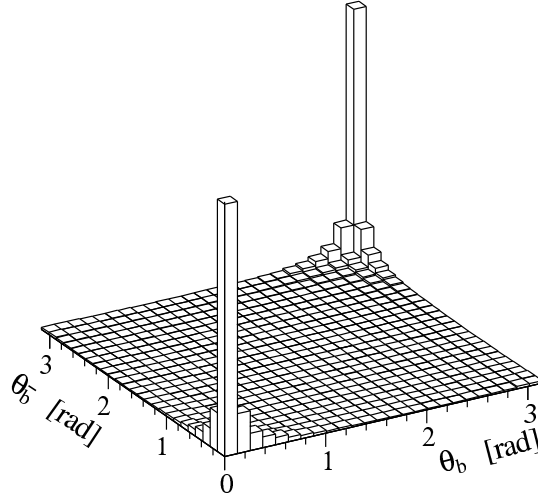


Figure 2.4: Polar angle correlations for $b\bar{b}$ production from proton collisions. θ_b and $\theta_{\bar{b}}$ are the polar production angles for the b -hadron formed from the b quark and anti-quark respectively.

2.4 The LHCb experiment

The potential of the LHC as a uniquely powerful laboratory for studying b -hadrons will be exploited by the LHCb specialist b -physics experiment.

Since both b -hadrons are preferentially produced in the same direction and are forward-boosted along the beam-pipe, the detector is not required to have full 4π solid-angle coverage. LHCb takes advantage of this by using a wedge-shaped single-arm detector with angular acceptance 10-300 mrad in the horizontal (bending) plane [92]. The maximum acceptance angle is reduced to 250 mrad in the vertical non-bending plane. The only detector component to provide coverage upstream of the interaction point is the silicon vertex detector. The detector is illustrated in Figure 2.5, showing the overall scale of the experiment and the surrounding cavern structure. The individual sub-detectors are labelled as described in Sections 2.4.1–2.4.8.

The single-sided detector design was chosen in preference to a two-armed design since the detector dimensions are restricted by the layout of the IP8 (ex-Delphi) cavern in which LHCb is located. Using all the available space for a single-arm spectrometer more than compensates in performance for the $\sim 50\%$ drop in luminosity.

A right-handed coordinate system is defined for general discussions of experiment layout. The coordinate origin is centred on the nominal interaction point and the axes are defined with the z -direction aligned along the beam axis and pointing down the

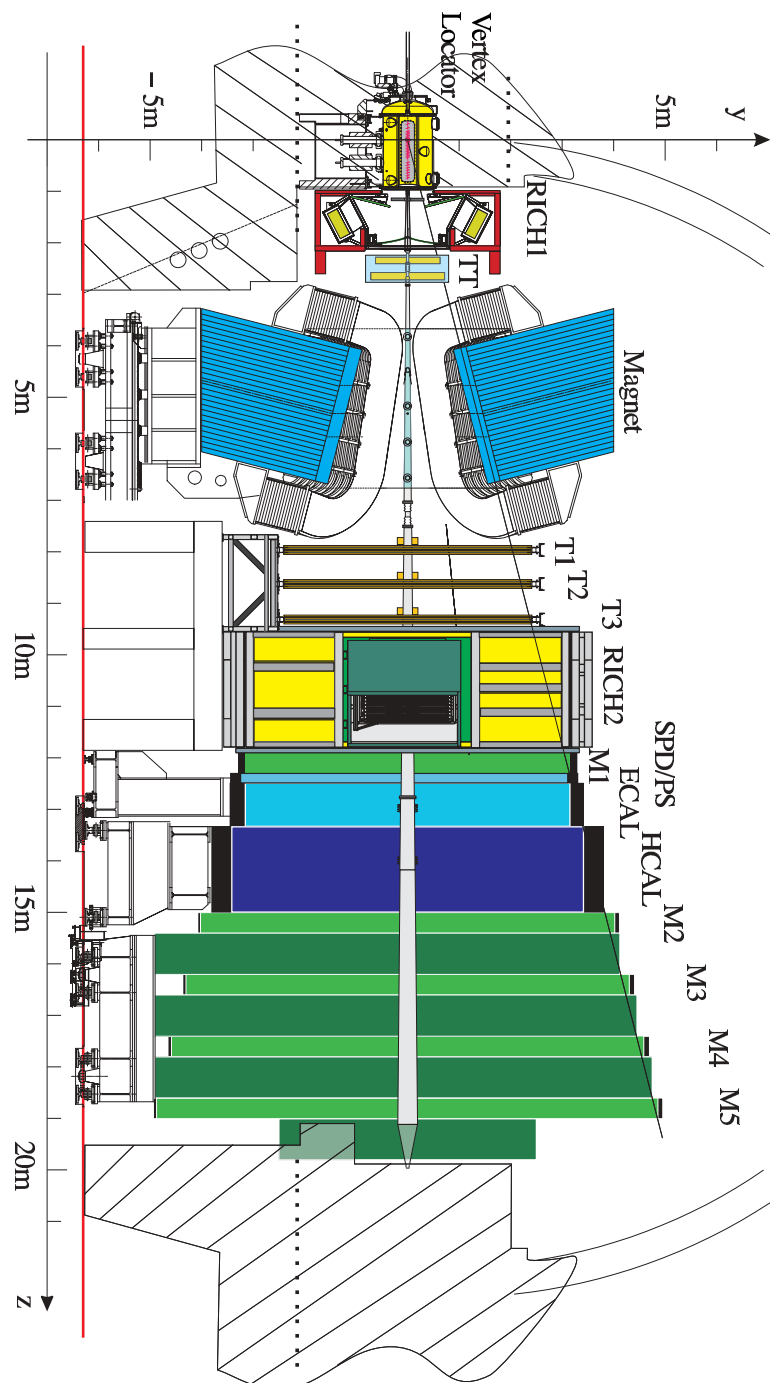


Figure 2.5: Cross-section view of LHCb, cut in the non-bending $y - z$ plane.

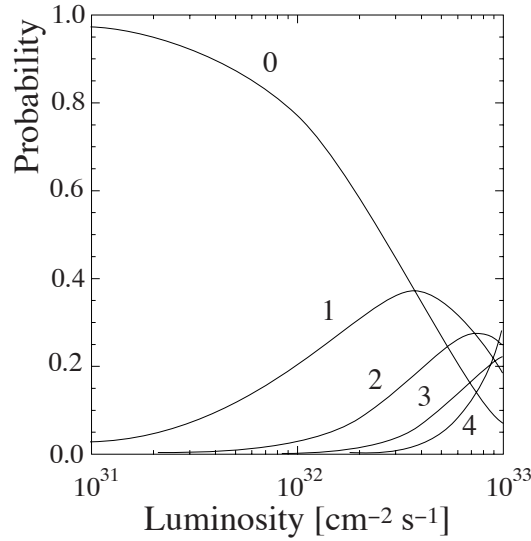


Figure 2.6: Probabilities of given numbers of pp interactions per bunch crossing as a function of LHC luminosity.

spectrometer arm and y pointing vertically upward.

2.4.1 Interaction point

The number of pp interactions in a given bunch crossing is described by a Poisson distribution with mean μ given by

$$\mu = \mathcal{L}\sigma/f \quad (2.1)$$

where \mathcal{L} is the luminosity, f is the bunch-crossing frequency and σ is the interaction cross-section [92]. The distributions of number of interactions per bunch crossing are shown in Figure 2.6.

For an inelastic pp scattering cross-section of $\sigma = 80$ mb, the LHC bunch-crossing frequency of 40 MHz and the LHC design luminosity of 10^{34} $\text{cm}^{-2} \text{s}^{-1}$ this expression evaluates to give 20 pp interactions per bunch-crossing. This detector occupancy is too high for precision b-physics as the tracking is degraded and the possibility of multiple primary vertices introduces unacceptable uncertainties into b-tagging and measurements of B proper lifetimes.

This problem is solved by defocussing the LHC beams in the IP8 interaction region, which reduces the luminosity by a factor of 100 to 10^{32} $\text{cm}^{-2} \text{s}^{-1}$ [92]. In addition, by

defocussing a quarter of the bunch-crossings will fail to interact, which reduces the active bunch crossing rate to 30 MHz. Re-calculating equation (2.1) with these new inputs gives $\mu = 0.53$, which corresponds to no interaction crossings at 17.6 MHz, one interaction crossings at 9.4 MHz and more than one > 1 interaction crossings at 3.0 MHz. The multiple-interaction events are referred to as “pile-up” events and are vetoed by the data acquisition trigger system.

The reduced luminosity at IP8 also has the advantage of reducing the radiation damage to the detector and enables LHCb to take data at the design rate from the first days of LHC operation.

2.4.2 Beam pipe

The beam pipe maintains the vacuum environment required by the beam and is shaped to present as small an interaction distance as possible to particles produced in the interaction region. This is important as excess material in the vicinity of the beam pipe makes a large contribution to the number of secondary tracks seen and in this region they cannot be identified and eliminated by e.g. track-kinking because the initial leg of the track is unresolvable.

A 1.8 m section of the beam pipe around the interaction point has an expanded diameter, of about 1.2 m, to accommodate the retractable vertex locator sensor array. This is bonded to the downstream section of the beam-pipe at the vertex locator exit window. The downstream beam pipe is composed of two conical sections, the first of which is made of 1 mm thick beryllium and has an aperture of 25 mrad for a length of ~ 1.8 m. The second section is constructed from aluminium in a 10 mrad cone which runs the full length of the detector in three mechanical sections of increasing thickness before being bonded back to the LHC pipe at a z -distance of just over 20 m from the interaction point [95].

Aluminium and stainless steel flanges and bellows on the pipe are positioned to minimise the number of low momentum material-interaction products which are able to reach the tracking stations.

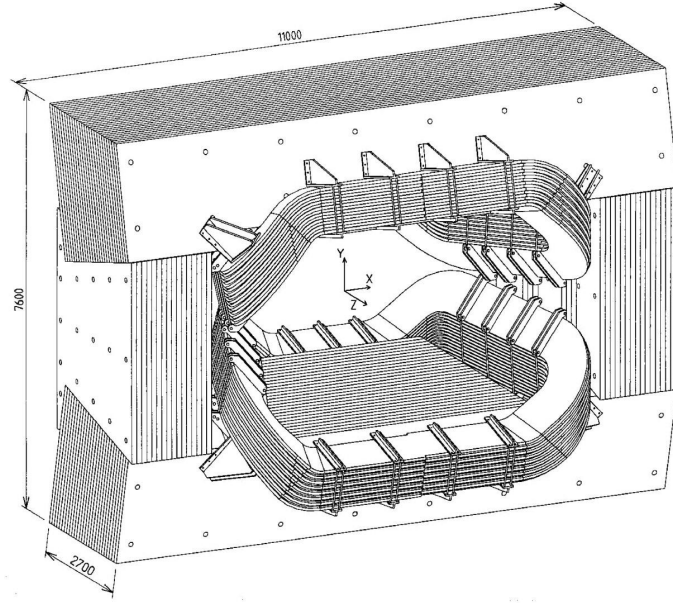


Figure 2.7: The LHCb dipole magnet.

2.4.3 Magnet

The dipole magnet in LHCb is located close to the interaction region to keep its size small. It is shown in more detail in Figure 2.7. The magnet is located immediately downstream of RICH1 and immediately upstream of the tracking stations. This positioning is designed to give low track curvature in the vertex locator, where the validity of a fast straight-track fitting algorithm is essential to trigger performance, while providing sufficient field integral before the tracking stations to give good momentum resolution from tracking. Sufficient field integral for determining track p_T is obtained by the use of a field-focusing design of the RICH1 iron housing [95].

A warm dipole magnet design is used to permit rapid field ramping and polarity switching. The magnet provides a maximum field strength of 1.1 T in the vertical direction and the polarity can be switched to reduce systematics from left-right asymmetries in the detector [96]. The total field integral of 4 Tm is sufficient for the trigger requirements [95].

2.4.4 Vertex locator (VELO)

Since the $\mathcal{O}(1\text{ mm})$ displacement of b-decay vertices from their production point is a potent identifying feature, resolution of these vertices to sub-mm accuracy is a key

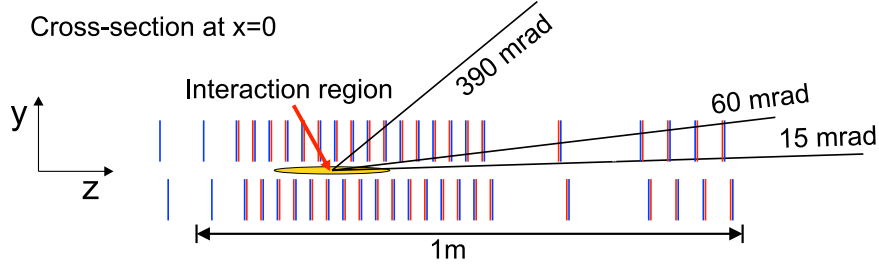


Figure 2.8: The layout of the VELO layout in the non-bending $y - z$ plane, showing the distribution of sensors around the interaction point. The red and blue lines indicate the r and ϕ sensors, respectively, with the two veto-stations at maximum negative z using only single-sided sensors.

feature of a dedicated b-physics experiment. b-tagging is required for both a high trigger efficiency and also to obtain good proper lifetime resolution: the latter is required for time-dependent studies such as those involving neutral $B - \bar{B}$ oscillations.

The LHCb silicon vertex locator system (VELO) surrounds the beam interaction region with a layout as shown in Figure 2.8. The VELO design is largely aimed at reducing the VELO material budget within the acceptance since the material concentration in the region of the interaction point significantly affects the overall detector performance [95].

A total of 21 sensor stations was chosen to provide an optimal balance between material budget and tracking performance. Of these, 8 stations are upstream of the nominal interaction point, giving a total downstream-station material budget of $\sim 16\%$ X_0 , averaged over a flat pseudo-rapidity distribution in the detector acceptance. The 8 upstream stations are used for backward tracking and for vetoing pile-up interactions — 2 of these are specifically dedicated to the pile-up veto. Tracking is an important feature of the VELO design since it is the major tracking detector prior to the TT tracking station (see Section 2.4.6).

The VELO is located in the LHC vacuum: the main VELO vacuum tank is bonded directly to the beam pipe and essentially forms an 80 cm-wide section of the pipe. A ~ 0.3 mm-thick aluminium alloy “RF foil” is located close to the beam to protect the sensors from the RF currents induced by the beam, as well as protecting the beam-line and vacuum from mechanical failures of VELO components. This foil has been designed with corrugations to give mechanical strength, to dissipate the induced RF currents and to minimise the material traversed by particles from the primary vertex [97]. The foil and its interaction with the sensors is shown in Figure 2.9.

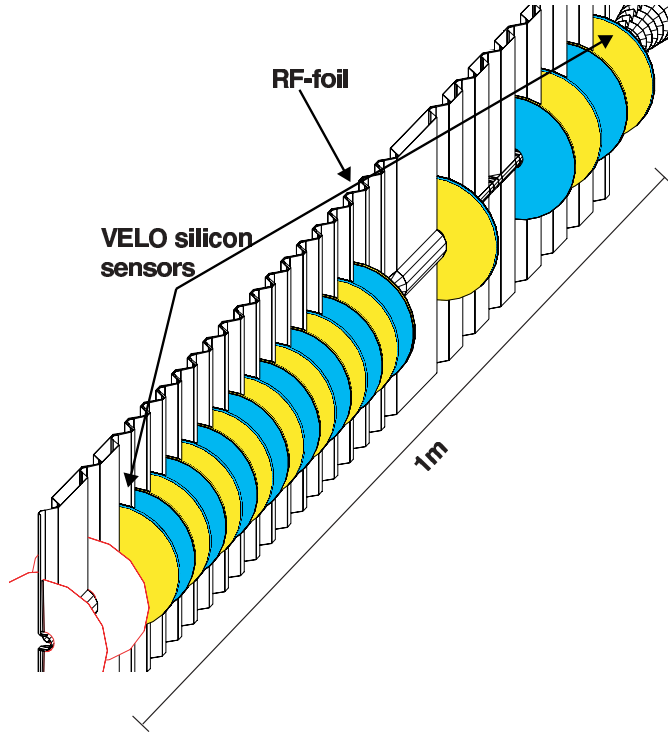


Figure 2.9: The VELO sensors and the primary vacuum RF foil.

For maximum accuracy, the sensors need to be as close to the interaction region as possible. During normal operation the beam width is narrow enough to allow sensors to be positioned within 8 mm of the nominal interaction point without damage but during injection of bunches into the machine a wider beam is generated. To be compatible with both regimes, the silicon sensors are formed from pairs of approximate semi-circles, called “half-stations”. These are mounted in pairs on opposing sides of the beam and are retractable by up to an extra 30 mm during injection — a so-called “Roman pot” design. The matching half-sensors are actually separated by ~ 1.5 cm which allows for a slight overlap and allows an RF foil corrugation to sit between them.

Each half-station is composed of two 220 μm -thick silicon sensors bonded back-to-back. One of these sensors is divided into annular strips and the other into sectors with radial boundaries: the first kind is an “ r -sensor” and provides information about the radial position of the track hit and the others are “ ϕ -sensors”, which primarily indicate the azimuthal coordinate of the detected hit but also provide some r information. The two sensor types are illustrated in Figure 2.10. As the ϕ -sensors alone provide a rough track position, they are used by the trigger system to make a quick vertex fit.

The sensors use 10 μm -wide n-doped silicon strips for readout, isolated by 5 μm -wide

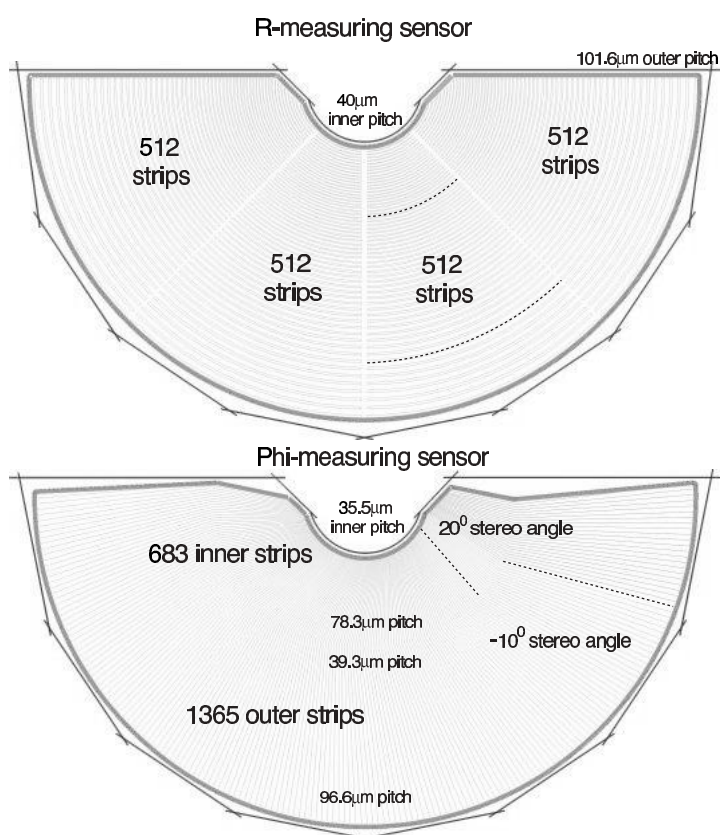


Figure 2.10: VELO r - and ϕ -sensors.

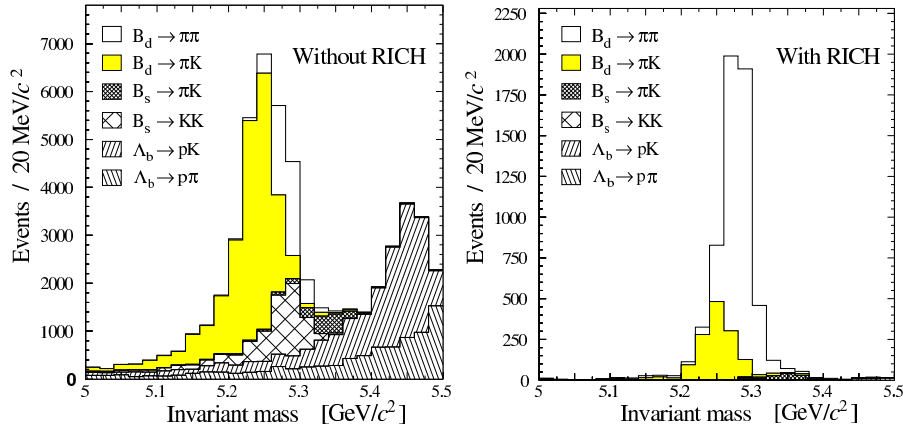


Figure 2.11: $\pi^+\pi^-$ invariant-mass spectra without (left) and with (right) RICH particle ID.

p-doped guards. When a charged particle passes through the silicon it creates electron–hole pairs which drift to the strip readout points under a reverse bias voltage. This signal is read out along 20 μm -wide double-metal strips which are routed out from the centre of the sensor on top of the other structures. Readout is controlled by a front-end “Beetle” chip, which is also used by the silicon tracker. Each half of each r -sensor is composed of four 45° sections, each incorporating 512 detector strips, while the ϕ -sensors are divided into inner and outer regions, of 683 and 1365 strips respectively, as illustrated in Figure 2.10.

The VELO is designed to give a proper time resolution of better than 45 fs, sufficient for resolving the anticipated B_s^0 oscillation rate [95].

2.4.5 RICH system

The ring-imaging Čerenkov (RICH) detector system of LHCb is one of the crucial specialisations of the detector for b-physics. Combined with the momentum measurement from track curvature the RICH system makes high-quality charged-particle identification possible. In particular the ability to distinguish between charged π and K tracks is important to precision b-physics since many decay channels of interest for \mathcal{CP} -violation studies have identical topologies and only differ in whether the final state particles are pions or kaons. The RICH system is also used to “tag” kaons as part of the strategy used in determining the flavour of the b-hadron in a reconstructed signal channel (see Section 2.7).

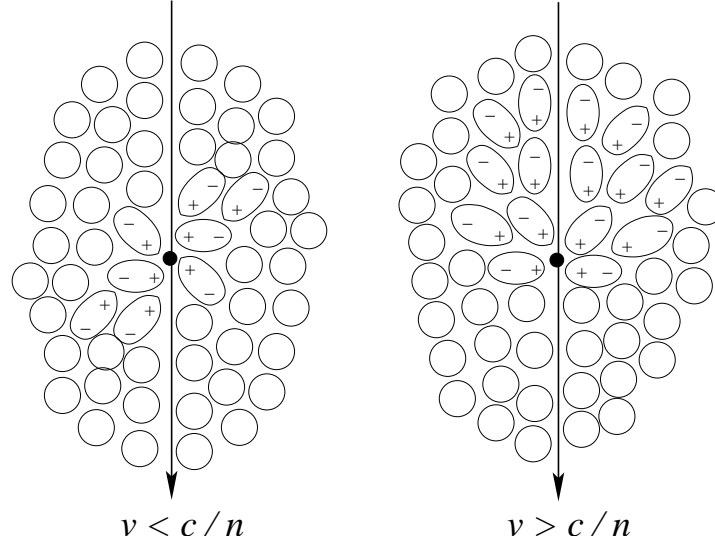


Figure 2.12: Dipole arrangements in Čerenkov radiator for a) $v < c/n$ and b) $v \geq c/n$.

The power of the particle identification supplied by the RICH system can be seen from Figure 2.11 in which the reconstructed invariant-mass spectra of $B \rightarrow \pi\pi$ decays are shown. In the left-hand plot, the “best guess” assumption is that all final-state particles are pions. In this case the signal peak is obscured by the backgrounds from mis-identified other final states. The right-hand plot shows the same quantity, but only for those decays in which both final state particles were identified as pions by the RICH particle ID [92].

The Čerenkov effect and RICH detectors

RICH detectors are based on the geometric optics of the Čerenkov effect: this occurs when a charged particle passes through a dielectric medium of refractive index n with a velocity v greater than the local speed of light, c/n . The presence of the charged particle polarises the material in its vicinity so that the surrounding atoms acquire induced dipoles. If $v < c/n$ then the arrangement of these dipoles around the particle is symmetric and so there is no net dipole and no resulting radiation. However, if $v > c/n$ then the symmetry of the dipole arrangement is broken (as there is effectively no causal connection between spatially separated dipoles), a net dipole develops and spontaneous photon emission, known as Čerenkov radiation, occurs [98]. The arrangement of dipoles for sub- and super-luminal particle motion is shown in Figure 2.12.

A Huygens construction in terms of spherical shells of probability for photon emission

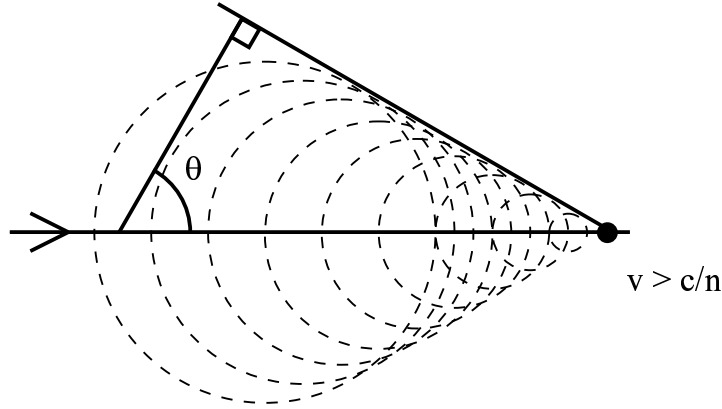


Figure 2.13: Huygens construction for Čerenkov photon emission from a superluminal charged particle.

as the particle progresses along its track is demonstrated in Figure 2.13. This shows that an effective “shock-front” of Čerenkov emission is formed, corresponding to an emission cone with opening angle θ_C around the momentum vector for each point on the track,

$$\cos \theta_C = \frac{1}{n\beta} + \frac{\hbar k}{2p} \left(1 - \frac{1}{n^2} \right) \quad (2.2a)$$

$$\sim \frac{1}{n\beta} \quad (2.2b)$$

where $\beta \equiv v/c$, the relativistic velocity fraction, p is the momentum of the radiating particle and k is the wave-number of the radiation [98]. In equation (2.2b) the small radiating particle recoil correction term is neglected — a robust approximation for all practical cases. As the dipole asymmetry only arises for $v > c/n$, Čerenkov photon production only occurs above the threshold $\beta = 1/n$ at which production is in the forward direction, $\theta_C^{\max} = 0$. The emission angle then saturates for $\beta = 1$, so that $\theta_C^{\max} = \arccos(1/n)$. An immediate consequence is that Čerenkov emission requires that $n(k) > 1$.

The central concept of RICH detectors is that by reflecting the light from a series of cones along a straight track off a spherical mirror, all photons from that track will fall on to a single characteristic ring at the mirror’s focal length (hence “ring-imaging”). The position of the centre of this ring corresponds to the projection of the track and the radius depends solely on the particle velocity. Hence tracks can be matched to rings and the combination of track momentum measured from the track curvature through the dipole magnet and the velocity measurement from the RICH detector means that the particle mass (and hence identity) may be deduced [99].

The photons produced by the Čerenkov effect have spectrum and intensity distributions described by the Frank-Tamm relation [100]:

$$\frac{dN}{dE} = \left(\frac{\alpha}{\hbar c} \right) Z^2 L \sin^2 \theta_C. \quad (2.3)$$

This predicts the number of Čerenkov photons dN to be found in an interval dE around energy E for a particle of charge magnitude $|q| = Ze$ with track-length L through the radiator. $\alpha = e^2/4\pi$ is the electromagnetic fine-structure constant. The observation that this relation is independent of E manifests itself in a flat energy spectrum for Čerenkov light. Substituting the approximate result for the Čerenkov angle from equation (2.2) we see that the number of Čerenkov photons N expected in any given energy range is maximised for ultra-relativistic particles where $\beta \rightarrow 1$:

$$N \propto 1 - \frac{1}{n^2\beta^2}. \quad (2.4)$$

The LHCb RICH system

In LHCb the RICH system is split into two detectors, RICH1 and RICH2, which are located respectively upstream and downstream of the magnet, as shown in Figure 2.5. RICH1 detects particles with relatively low momenta from 2-70 GeV/ c over a wide range of polar angles up to 300 mrad while RICH2 has a smaller angular acceptance of up to ~ 120 mrad but detects a wider range of momenta, roughly 20-100 GeV/ c . RICH1 is hence placed in front of the magnet to detect low-momentum tracks which may get swept out of the geometric acceptance of the spectrometer and also to keep its size manageable, while RICH2 is placed much further downstream. The resulting polar angle and momentum coverage of the two detectors as compared to tracks from typical B-decay events are shown in Figure 2.14. Note that RICH2's operational momentum is only expected to extend to ~ 100 GeV/ c rather than the design limit of 150 GeV/ c shown in the figure.

The nominal Čerenkov radiator gas in RICH1 is C_4F_{10} with refractive index $n = 1.0014$ for $\lambda = 260$ nm. This provides 3σ π -K separation up to momenta of ~ 50 GeV/ c [99]. A mix of C_3F_8 and C_5F_{12} in equal molar concentration provides similar optical properties and may be used in the installed system. To increase sensitivity to low momentum tracks, RICH1 also contains a 5 cm silica aerogel panel. The aerogel rings are much larger than those from the RICH1 gas radiator and have fewer photons, largely because

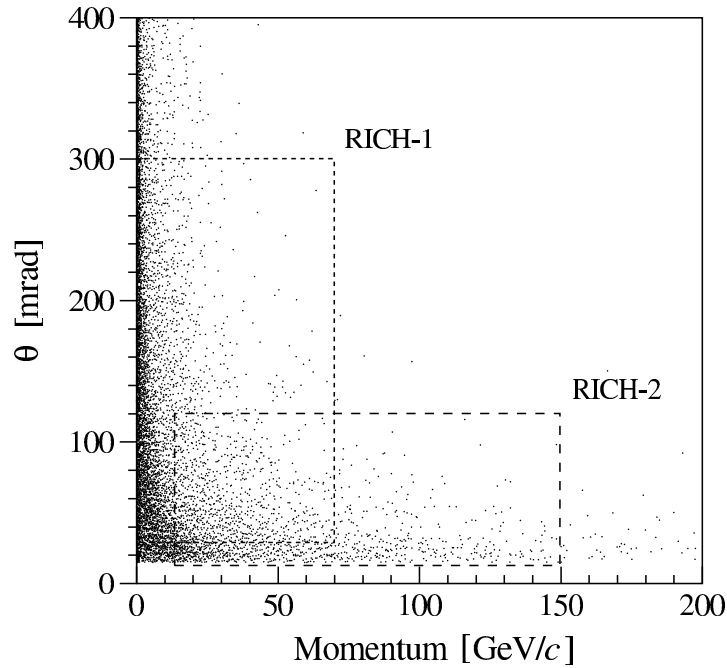


Figure 2.14: Polar angle vs. momentum coverage of RICH1 and RICH2. The two detectors are designed with complementary operational ranges.

the track-segment length in the aerogel is much less than that in the gas. RICH2 has a single radiator, CF_4 gas with a refractive index $n = 1.0005$ for $\lambda = 260$ nm, providing high-momentum π -K separation for $p \sim 20$ – 100 GeV/ c . Selected design parameters and radiation threshold momenta the RICH radiators are listed in Table 2.1. RICH1 is ~ 3 m \times 2 m in the x - y plane and extends ~ 1 m in the z -direction. RICH2 is ~ 7 m \times 7 m in the x - y plane and ~ 2 m deep.

Characteristic	Aerogel	C_4F_{10}	CF_4
L	5 cm	85 cm	167 cm
n	1.03	1.0014	1.0005
θ_C^{\max}	242 mrad	53 mrad	32 mrad
$p_{\text{thresh}}(\pi)$	0.6 GeV/ c	2.6 GeV/ c	4.4 GeV/ c
$p_{\text{thresh}}(\text{K})$	2.0 GeV/ c	9.3 GeV/ c	15.6 GeV/ c

Table 2.1: Various characteristics of the RICH radiator media at standard temperature and pressure. These are the average track length in the radiator L , the refractive indices n , the saturated Čerenkov angle θ_C^{\max} and the threshold momenta for pions and kaons.

The optical layouts of the two RICH detectors, other than in overall scale, are similar.

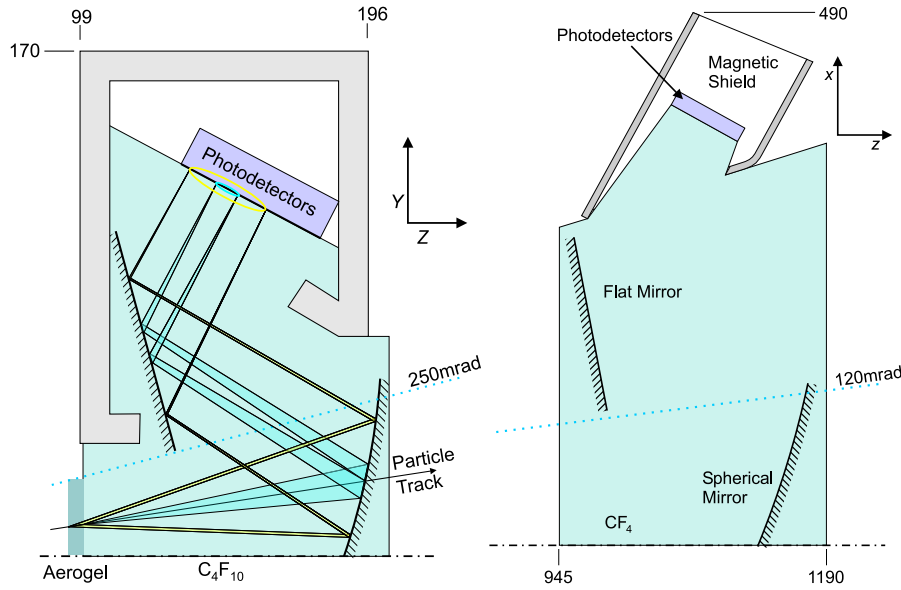


Figure 2.15: The LHCb RICH detectors, with RICH1 on the left and RICH2 on the right.

Both detectors use a flat mirror to allow the photon detector arrays to be placed outside the experiment acceptance and spherical mirrors to focus the Čerenkov emission cones as shown in Figure 2.15. The spherical mirrors have focal lengths of 0.85 m in RICH1 and 4.00 m in RICH2 [92, 99, 95]. The two RICH detectors are rotated by 90° relative to each other around the z -axis. RICH1's photon detector planes are located at the top and bottom of the detector and RICH2's photon detectors are on the sides. Again this design is motivated by constraints and requirements on the magnetic field. The mirrors in RICH1 are made of beryllium and those in RICH2 are quartz: this decision is based on a trade-off between cost and reduced radiation and interaction lengths in the geometric acceptance of the experiment.

The photon detectors in the RICH detectors are required to detect photon hits with a spatial resolution of better than ~ 2 mm and a response time compatible with the 40 MHz bunch-crossing rate (allowing for dead-time after a hit in a given pixel) [99]. The chosen photon detector technology is the pixel hybrid photon detector (HPD) shown in Figure 2.16 [101, 102, 103, 104]. These detectors have only one stage of photo-electron acceleration between cathode and anode and use a silicon matrix for detection at the anode. A direct mapping can be identified between the photon interaction point on the photo-cathode and the readout point on the silicon via the electron optics [105].

The HPD unit is an evacuated cylinder, 120 mm in length and 83 mm in diameter. Photons enter via a spherically-curved quartz window at one end and are converted to

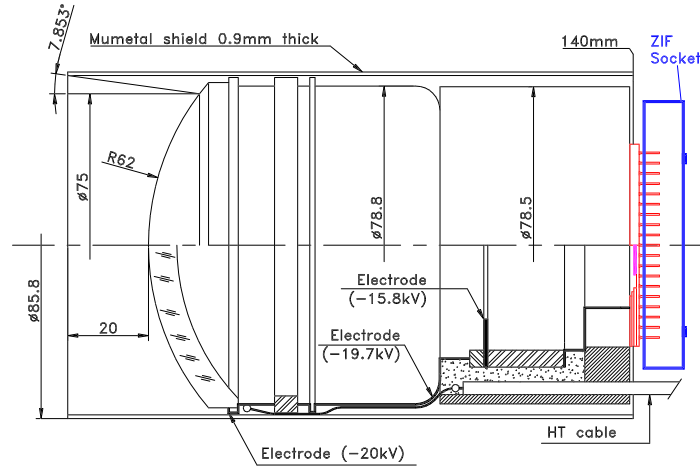


Figure 2.16: Schematic layout of a RICH system HPD, showing the magnetic shielding, the curved window cross-section, the electron optics and the silicon matrix bonded to the readout pins.

electrons at a photo-cathode deposited on the inside of the window. The peak conversion efficiency is $\sim 25\%$ at a wavelength of 300 nm [105]. The photo-electrons are accelerated through a single 20 kV potential to impact on the 1024-pixel anode assembly at the opposite end of the tube. The anode assembly consists of a silicon pixel detector bump-bonded directly to a binary readout chip. This assembly is mounted on a ceramic carrier and wire-bonded to an array of readout pins on the base of the tube from where readout to the “pixel chip” takes place. The HPDs are packed in a hexagonal array with 0.9 mm separation between tubes to allow for a thin magnetic shield of “mu-metal” round each tube. In total there are 2×98 tubes in RICH1 and 2×144 tubes in RICH2 [101] covering a total photon detector plane area of 2.6 m^2 . The interstices in the hexagonal packing are used for routing the readout cabling.

There is a delicate interaction between RICH1 and the fringe-field of the magnet, which is located only a short distance downstream. In particular, there is a requirement that the integrated on-axis B -field between the VELO and the TT tracking station be sufficiently large for the L1 trigger (see Section 2.5) and for reconstructing short tracks which do not propagate as far as the T stations. This conflicts with the strong sensitivity of the HPDs to magnetic fields: distortions in the electron optics can only be corrected for reliably up to $|B| \sim 10 \text{ G}$ [95]. An iron shielding box surrounds the RICH1 HPD system; this performs the dual rôle of shielding the HPD planes from the fringe field and focusing the B -field from the magnet poles on to the z -axis for the trigger. The latter rôle is the reason for RICH1’s vertical alignment and the addition of pole pieces to the

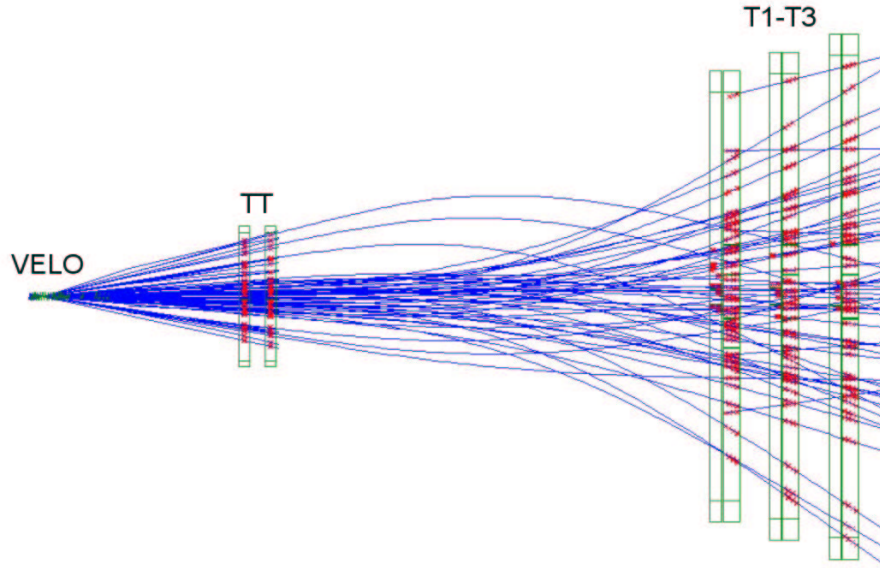


Figure 2.17: Reconstructed tracks in LHCb for an average-occupancy $b\bar{b}$ production event. Observed detector hits are shown in red and the corresponding reconstructed tracks in blue.

shielding box which improve the magnetostatic circuit. In addition, individual shields made of high-permeability iron alloy are to be used to protect the HPDs from the fringe field [95]. Structural shifts due to B -field polarity changes are simulated as being below the HPD pixel scale and hence frequent alignment calibration is not anticipated.

Ring reconstruction in the RICH system will be discussed in detail in Chapter 3.

2.4.6 Tracking system

The tracking system is composed of several distinct sub-detectors, namely the VELO, which has already been described, and the components labelled TT [106, 107] and T1–T3 in Figure 2.5 [106, 108]. The system reconstructs the tracks of individual charged particles, which directly allows the particle momentum to be determined from the track curvature in the dipole magnetic field. In addition, matching the reconstructed tracks to calorimeter clusters allows measurement of individual particle energies and the matching of tracks to the centres of RICH rings enables particle ID. About 74 reconstructed tracks are expected in an average $b\bar{b}$ production event: such an event is shown in Figure 2.17.

The silicon-based “TT” tracking station is located between RICH1 and the upstream magnet aperture. This station is used by the L1 trigger algorithm to estimate the p_T

of large impact parameter tracks. It is also used in offline analyses, being particularly useful for tracking low momentum particles which are swept out of the acceptance by the magnet and for seeding the tracking algorithm for long-lived particles such as K^0 which decay predominantly after the VELO.

The T1–T3 tracking stations are located between the magnet and RICH2 and are used to gather information about charged particles which have sufficient momentum to remain in the detector after the magnet. Each station is divided into a silicon microstrip inner tracker (IT) and an outer tracker (OT) based on straw drift chamber technology. The downstream tracking stations are sized such that the OTs extend to the LHCb acceptance and the IT–OT boundary in each station is chosen such that the occupancies in the outer trackers are kept below 10% [92]. The finer resolution of the silicon microstrip allows it to maintain tracking performance in the higher-occupancy on-axis region.

Both the TT and the IT silicon microstrip detectors are composed of four sensor planes: the two outer sensors are vertically aligned and the two inner planes are inclined by $\pm 5^\circ$ to yield stereo information. The TT is split by a 30 cm gap in the z -direction between the two inclined planes to improve the angular resolution, while the IT stations are compact in z .

The OT straw tubes are 5 mm in diameter and are filled with a mixture of argon, CO_2 and CF_4 . The admixture may be varied to fine-tune the mean drift time, which is centred on 50 ns to restrict spillover occupancy effects to two bunch crossings. Like the IT stations, the OT detectors are split into four sensor planes with the middle two inclined by $\pm 5^\circ$. Each plane is formed from two layers of straws, packed so that the spatial resolution of hits is less than the straw radius and arranged in modules of 64 straws. The only tracking in the y – z non-bending plane is the rough tracking provided by the stereo layers: this is sufficient for RICH ring-track matching and identifying calorimeter clusters and is irrelevant to the measurement of momentum components in the x – z bending plane.

The overall momentum resolution of the LHCb tracking system is $\delta p/p = 0.37\%$. The track-finding efficiency is 94% for tracks with hits in all tracking stations. The false (“ghost”) track rate is 9%, which reduces to 3% for tracks with $p_T > 0.5 \text{ GeV}/c$.

2.4.7 Calorimeters

The electromagnetic (ECAL) and hadronic (HCAL) calorimeters lie downstream of RICH2 with the ECAL located immediately upstream of the HCAL. A scintillating pad detector (SPD) and a pre-shower detector (PS) are located immediately upstream of the ECAL [92, 109]. These provide limited particle ID information for the first stages of the trigger system, at which stage the more powerful RICH particle ID is not yet available [110].

The calorimeters measure particle energy by incorporating large numbers of radiation and interaction lengths to encourage showering. They are particularly important for neutral particles which leave little other information in the detector other than the calorimeter clusters: in this case the calorimeters are actually involved in tracking to a limited extent. The calorimeters are additionally used in the trigger algorithms, where they provide information about the particles in the event with high transverse energy.

The SPD is a 15 mm-thick layer of scintillator tiles which respond to the passage of charged particles and hence allows electrons and photons to be distinguished, which cannot be done once showering has begun in the calorimeter volumes. Downstream is a 12 mm-thick layer of lead ($2.5 X_0$) followed by another layer of scintillator tiles: this combination is the pre-shower detector [109]. The PS uses the difference in interaction lengths for electrons and pions in lead to distinguish them, the electrons depositing significantly more energy in the scintillator than the pions.

The ECAL and HCAL themselves are both sampling calorimeters with an alternating structure of scintillator tiles and high- X_0 material to encourage showering, as shown in Figure 2.18. As electromagnetic showers are more collimated than hadronic showers, the ECAL is composed of alternating 2 mm-thick scintillator tiles and 4 mm-thick lead sheets with their short axes aligned along the z -direction. By comparison the HCAL is composed of scintillator tiles with average thickness of 4 mm and 16 mm-thick iron with the short axes perpendicular to the z -axis to improve the sampling for the more lateral hadronic showers. The large-scale HCAL structure in z alternates 200 mm of the above structure with 200 mm of iron [109]. The ECAL longitudinal structure corresponds to $25 X_0$.

The SPD, PS and ECAL scintillators are read out via a “shashlik” design where wavelength-shifting fibres “skewer” the scintillator tiles whereas in the HCAL the fibres are connected to the tile at one edge only. The fibres are read out to photomultiplier

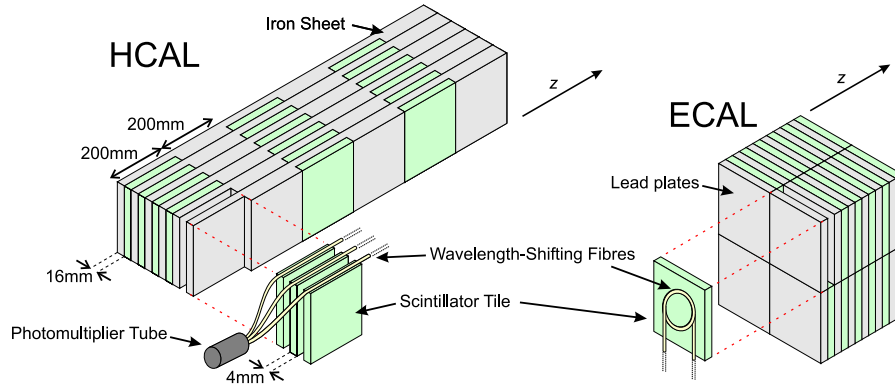


Figure 2.18: The fine structure of the LHCb calorimeters, with the single-edge readout of the iron-scintillator HCAL on the left and the shashlik readout of the lead-scintillator ECAL on the right. The different tile alignments, to handle the difference in collimation between electromagnetic and hadronic showers, is illustrated.

tubes: for the SPD, PS and ECAL multi-anode PMTs are to be used. The HCAL uses only one single-anode PMT for each set of three tiles.

The tile layout within the calorimeters is structured so as to match the geometry of tracks originating from the primary vertex, which makes combining measurements from different calorimeter components faster. The tile sizes also vary with position in the x - y plane to match the variations in particle flux: the ECAL tiles are ~ 40 mm squares at the 30 mrad closest approach to the beam but increase to first ~ 60 mm squares and finally ~ 120 mm squares further out from the beam axis. The SPD/PS tiles follow a similar pattern but are down-scaled by about 1.5% to match the projective geometry from the PV. The HCAL tiles are also sized according to position: near the beam they are ~ 130 mm square and further out the pitch is doubled to ~ 260 mm square. The trigger requirements on the HCAL do not impose a stringent shower-confinement condition and so no specific radiation length is required; the HCAL length is 5.6 interaction lengths [109].

The energy resolution provided by this calorimeter design is $\sigma(E)/E = 10\%/\sqrt{E} \oplus 1.5\%$ for the ECAL and $\sigma(E)/E = 80\%/\sqrt{E} \oplus 10\%$ for the HCAL with E in GeV; these performance figures have been verified by experimental tests on the HERA-B calorimeter modules, which are of very similar design [111].

2.4.8 Muon system

The muon detector system is the furthest downstream of the LHCb sub-detectors: it is distributed among 5 stations marked as M1–M5 in Figure 2.5. M1 is located between RICH2 and the calorimeters, while the other muon stations are all located after the HCAL. Muon detection is important for both triggering and offline analysis since muons occur in the final states of benchmark \mathcal{CP} -violating channels and of semi-leptonic B decays used by flavour tagging algorithms. The p_T information from the muon detector is used in conjunction with the calorimeter E_T measurement by the first stage of the trigger.

Of the five stations, M1–M3 are used for muon track-finding and momentum measurement and M4 & M5 are used purely for muon identification. The 310 cm separation of M1 from M2 is crucial to the precision of the muon p_T measurement. Since the M2–M5 muon chambers are separated from the interaction point by some three hundred radiation lengths and thirty interaction lengths muons are the only charged particles to reach them: being more than 100 times heavier than electrons they do not scatter or shower significantly. Between each of the M2–M5 stations is an 80 cm-thick iron plate to reduce the hadronic background [112, 113, 114].

All five muon stations are constructed as arrays of multi-wire proportional chamber (MWPC) detectors. Each chamber is constructed from four 5 mm potential gaps which are filled with Ar, CO₂ and CF₄ gases in 4:5:1 admixture and have centrally-aligned anode wires as shown in Figure 2.19. Rather than induce muon showering, the detectors operate by detecting the ~ 50 electrons produced by a typical muon as it ionises the gas: the electrons are accelerated toward the anode wires by the 3 kV potential and the resulting signal is amplified by the data acquisition electronics. The four sequentially-stacked chambers provide an overall muon-detection efficiency of $\sim 99\%$.

The muon trigger operates with an efficiency of $\sim 55\%$ (for inclusive B decays with at least one muon in the final state and fully contained in the detector acceptance). In the offline analysis the muon identification efficiency is in excess of 90% for muons with momenta above 3 GeV/ c and the mis-identification of pions as muons is restricted to less than 1.5%.

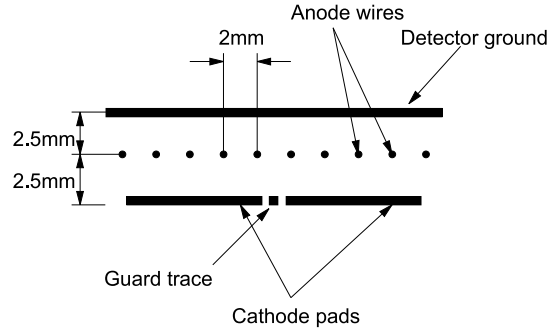


Figure 2.19: Schematic diagram of a muon system MWPC.

2.5 Trigger system

After beam de-focusing at the LHCb interaction point the active bunch crossing rate is 30 MHz, of which ~ 10 MHz are the single pp interactions of interest to LHCb. Accounting for the $b\bar{b}$ production cross-section, the experiment's geometric acceptance, the hadronisation fraction to B-mesons and the fact that the inclusive branching ratio to b-hadron decay channels of phenomenological interest is $\mathcal{O}(10^{-3} - 10^{-5})$, events of physical interest only occur at a few tens of Hz.

The solution to this huge difference between the bunch crossing and useful event rates is to implement a multi-stage trigger system: LHCb's trigger is divided into Level 0 (L0), Level 1 (L1) and the high-level trigger (HLT) [110, 115]. L0 is implemented in hardware while the L1 and HLT triggers are performed by software running on the PC farm. At each level the event data is queued in a buffer while a decision is made to accept or reject it based on the likelihood of it containing a b-decay of interest. The algorithms which make this judgement become increasingly more complex with increasing level, which is possible due to the progressively lower input rates at each stage. The separation into levels also allows a balance to be struck between minimising dead time and buffer sizes. An overview of the trigger characteristics broken down by level is shown in Table 2.2 and described in more detail below. Figure 2.20 shows the design architecture of the L0 and L1 trigger systems.

2.5.1 L0

The L0 trigger processing is performed by dedicated electronics on selected sub-detectors to simplify readout requirements. The trigger hardware operates at the 40 MHz bunch-

	L0	L1	HLT	Offline
Input rate	40 MHz	1 MHz	40 kHz	2 kHz
Output rate	1 MHz	40 kHz	2 kHz	—
Sub-detectors	VELO, ECAL, HCAL, muon	All except RICH & T1–3	All	All
Location	On detector	Counting room	Counting room	Institutes & GRID

Table 2.2: Characteristics of the trigger levels and offline analysis.

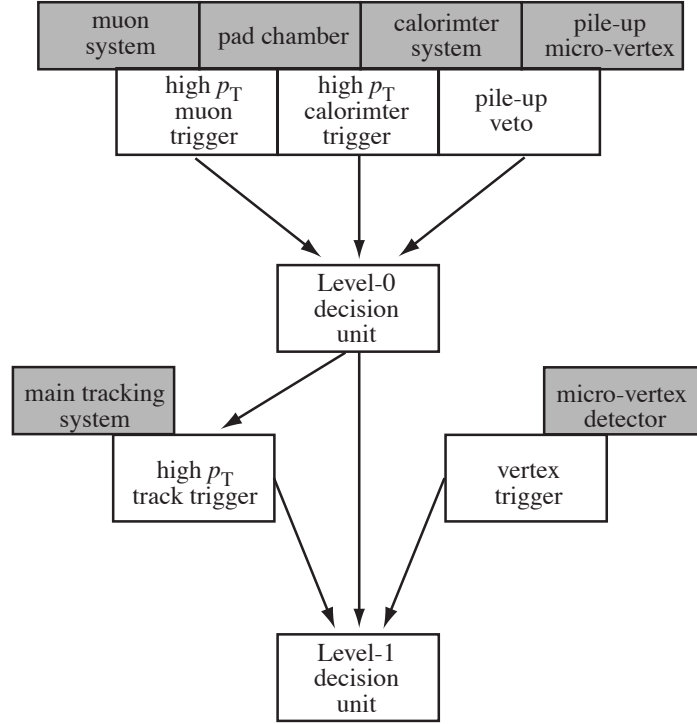


Figure 2.20: The architecture of the LHCb L0 and L1 trigger system, with the detectors involved in each stage of the trigger algorithm indicated.

crossing rate and is required to have a average acceptance rate of ~ 1 MHz. Each sub-detector is required to buffer 4 μ s of data, corresponding to 160 bunch-crossings, while the L0 decision unit (L0DU) decides to accept or reject the event.

Since b-hadron decays typically produce low-mass daughter particles with high momenta, events with high- p_T electrons, muons and hadrons are selected using rapid calorimeter and muon system measurements. Events are only accepted if the total E_T is greater than 5 GeV, to reduce the possibility of triggering on halo muons in events with no interactions. A pile-up veto is also applied at this stage: the two dedicated VELO

veto stations are used to discard the event if more than one pp interaction occurs in the bunch crossing, regardless of p_T measurements. Additionally, the L0 system applies a set of specific trigger cuts for specific benchmark physics channels: the details are documented in the LHCb trigger technical design report [110].

2.5.2 L1

The L1 trigger operates at the L0 accept rate of 1 MHz and accepts events at an average rate of 40 kHz. The trigger algorithm is performed by software and is nominally assigned 55% of the available CPUs in the counting room PC farm. The buffering is performed by dedicated hardware, also located in the counting room.

The algorithm makes its decisions based on the L0 data from the calorimeters and muon system and also incorporates tracking data from the main VELO stations and the TT detector. The VELO tracks are matched to the TT tracking data and to the muon tracks and calorimeter clusters from L0. The improved tracking is used to perform a more stringent low-cut on track p_T .

The L1 selection is primarily based around tagging displaced vertices in the VELO as a signature of b-quark decay. The event is required to have at least one VELO track with an impact parameter to the primary vertex in the range 0.15-0.3 mm. This decision is weighted by the track momentum since large impact parameters are not such a strong discriminating factor for low-momentum tracks as they are for high-momentum ones due to multiple scattering. The main L1 trigger cuts are applied in the form of a “vertical-diagonal discriminant” cut in the 2D plane of $\sum \ln p_T$ against $\sum \ln IP$: this is summarised again in the trigger TDR [110] and shown in Figure 2.21.

2.5.3 HLT

The HLT is essentially an offline selection algorithm applying loose cuts on relatively complex parameters such as impact parameters of tracks. Data from all sub-detectors is available to the HLT algorithms. The HLT receives data from the L1 accept stage at an average rate of 40 kHz and accepts events at a nominal average rate of 2 kHz. Of this, 200 Hz are dedicated to specific decay channels and the remainder is used to store generic inclusive $b\bar{b}$ production events.

The current implementation of the HLT algorithm repeats the L1 reconstruction

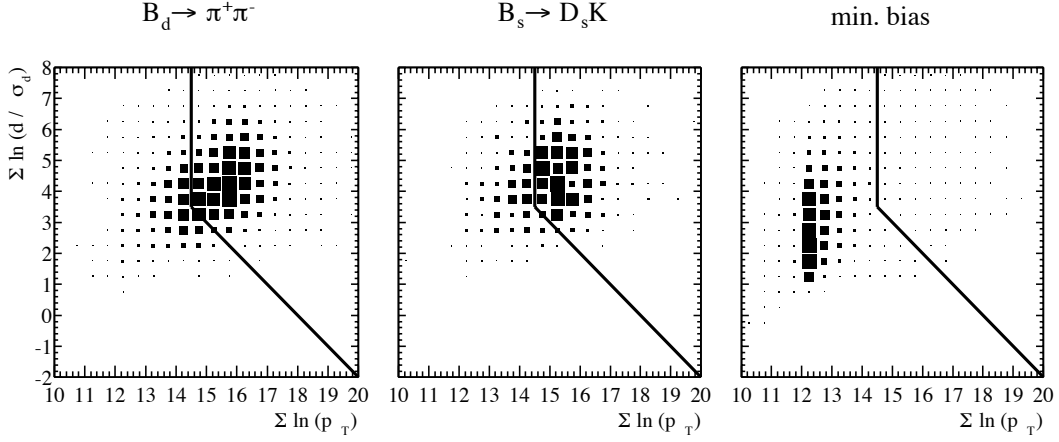


Figure 2.21: Examples of the L1 trigger vertical–diagonal discriminant cut for $B_d^0 \rightarrow \pi^+ \pi^-$, $B_s^0 \rightarrow D_s K$ and minimum bias events.

of VELO tracks and primary vertices more accurately and additionally matches the L1 tracks to the tracks reconstructed by the T1–T3 tracking detectors. The L1 cuts are then repeated with the improved tracking information. Additionally, the HLT incorporates several streams specific to particular decay channels: it is being considered that channels with neutral decay products which would normally be obliterated by the trigger (for example $B_s^0 \rightarrow K_S^0 K_S^0$) may be given special exemption via a dedicated HLT stream.

Although data from the RICH system is available to the HLT, it is currently unused. Particle ID information from the RICH is being incorporated into the HLT at the time of writing, via a fast parameterisation-based ring–track matching algorithm which avoids the computationally intensive inversion of the RICH geometric optics.

As the HLT algorithm is entirely implemented in software, it is under constant improvement and will not be fully defined until detector commissioning. The HLT is nominally assigned 25% of the counting room PC farm’s CPU power.

2.6 Data handling

LHCb data handling is largely performed by the PC cluster located in the counting room, which is in the IP8 cavern near to the the detector. This will be composed of ~ 1800 rack-mounted machines which are nominally assigned to either stages of the trigger system or to offline data analysis. Storage for triggered data will be supplied by the CERN computing centre’s high-volume data system [116].

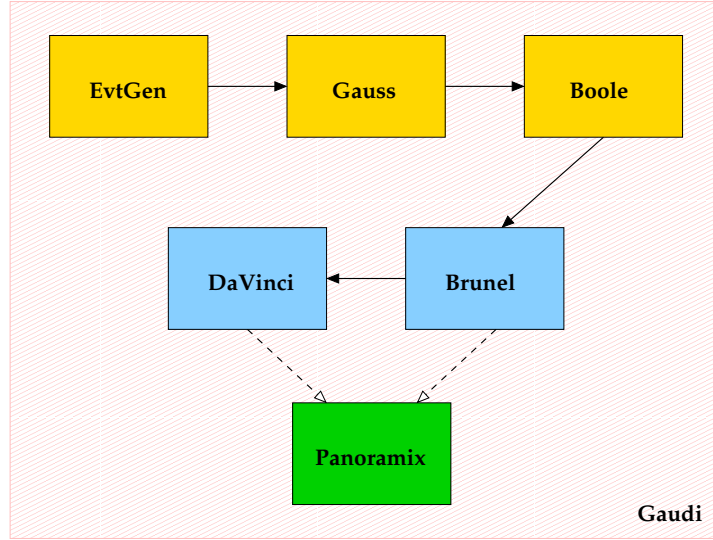


Figure 2.22: The logical structure of the LHCb software framework, divided into simulation (orange) and reconstruction (blue) components. All components make use of the high-level Gaudi framework.

2.6.1 Offline analysis

Offline data analyses are nominally assigned the remaining 20% of counting room CPUs. They will also be distributed between participating institutes via the LHC Computing Grid. The software framework used for LHCb analyses is described in Section 2.6.2 and is used in Chapter 5 for selection and analysis of $B \rightarrow K \pi$ decays.

To enable comparison of data recorded at different times, when the detector configuration is likely to have changed in-between, a detector conditions database will be generated and distributed around the offline data analysis centres.

2.6.2 LHCb software framework

The LHCb software is built on a set of object-oriented C++ frameworks, structured as shown in Figure 2.22. The “Gaudi” framework provides the administrative tasks like histogram and data persistency and run-time configuration via either Python or custom-format option files [117, 118, 119]. An algorithm structure designed for performing the same task over many events with optional initialisation and finalisation routines is also derived from a Gaudi template algorithm. Gaudi also provides hooks to enable its use in a distributed “GRID” environment and this is used by Dirac [120, 121, 122], the LHCb distributed Monte-Carlo data production system.

The LHCb software built on Gaudi can be divided largely into the simulation system and the frameworks for analysis of data: the intention is that, when data-taking begins, the analysis software chain can be applied to the detector output in precisely the same way as it is used on simulated data.

Simulation software

The simulation software is divided into several stages. “EvtGen” is the Monte-Carlo event generator, using a combination of Pythia 6 [123] and custom b-decay algorithms to perform hard-scattering calculations, hadron showers and b-hadron decays. This is executed from the context of “Gauss” [124], which also contains the detector simulator, built on Geant4 [125] and a comprehensive XML description of the LHCb detector. The simulation is a full material interaction model rather than a parameterisation approach. “Boole” is the final element in the simulation chain, simulating the digitisation process from Gauss interaction data [126]. The “digi” files produced by Boole use the same data format as the raw detector output data frames.

Analysis software

The analysis software essentially seeks to invert the simulation process and work back up the chain to the fundamental physics, removing as far as possible the effects of detector resolution. The first tool in this process is the event reconstruction system, “Brunel” [127], which analyses raw “digi” data files to build physics objects like tracks, calorimeter clusters, RICH rings etc. The output of Brunel is used finally by the analysis system, “DaVinci” [128], to create high-level objects such as particles and vertices. DaVinci provides many tools for combining particle objects to form decay chains and for identifying signal events. Finally, an event visualisation system, “Panoramix” [129], is built on a combination of the Gaudi framework and the OpenScientist [130] modelling system.

Monte-Carlo truth information

One crucial aspect of the LHCb software is not mentioned above: how are we to build and validate the Brunel reconstruction algorithms without knowing which physics objects match up with the true processes simulated in Gauss? And how can we write

analysis algorithms without knowing which events are signal and which are background according to EvtGen and Gauss? Monte Carlo truth information is recorded in all the simulation steps and is available to MC-aware tools on the reconstruction and analysis side. Development methods are in force to ensure that all MC-awareness in the analysis and reconstruction software can be globally disabled and that none of the active reconstruction algorithms *depend* on knowing about the MC truth.

2.7 Flavour tagging

In studying, for example, B^0 oscillations it is important to be able to identify not just the flavour of the neutral B-meson at decay but also the initial flavour, in particular whether the constituent b-quark is a b or \bar{b} . Although the initial flavour cannot be measured directly, it may be possible to infer it from the decay products of the b-hadron containing the other member of the $b\bar{b}$ pair. It is conventional to refer to the hadron which is being reconstructed as the “same-side” hadron and the other one as the “opposite-side” hadron. Analysing the flavour of the opposite-side hadron to infer that of the same-side particle is called “tagging” [92].

A typical tagging strategy is to observe a charged lepton which does not originate from the same-side b-hadron: for example observing an opposite-side electron or muon implies an opposite-side b-decay chain of the form $b \rightarrow c \ell^- \bar{\nu}_\ell$ or conversely a positron or anti-muon indicates a $\bar{b} \rightarrow \bar{c} \ell^+ \bar{\nu}_\ell$ decay chain. Similarly, observing a charged kaon may imply a decay chain of the form $b \rightarrow c \rightarrow s$ or $\bar{b} \rightarrow \bar{c} \rightarrow \bar{s}$ according to the kaon charge.

The principal weaknesses of the strategy are threefold:

- not all events are taggable as, for example, not all the necessary decay products of the opposite-side decay lie within the detector acceptance or fail to meet other reconstruction criteria. Alternatively the decay daughters may not unambiguously indicate the b-hadron flavour, such as in the decay pair $B_d^0/\bar{B}_d^0 \rightarrow \pi^+\pi^-$;
- not all final state particles come from the opposite-side decay: in addition the detector contains final state tracks from the signal decay and fragmentation products from the pp interaction; and
- the opposite-side b-hadron may have also oscillated if it is a neutral meson: in this case we do not absolutely know the initial flavour of the opposite-side b.

These problems, in particular the second, lead to the tagging only being performed on a subset of events and the final state particles on which it is applied have to pass a set of selection cuts to identify the relevant opposite-side candidates. As the daughter particles from decays of b-hadrons have masses much less than the b-quark they typically have higher transverse momenta p_T than particles from pp fragmentation. Additionally, as they originate from a displaced secondary vertex, their mean impact parameter with respect to the primary vertex is also large.

These properties are used as for signal selection algorithms to select the best tagging candidate. In the event that more than one tagging candidate is found for a given tag-species (kaon, muon or electron) then the candidate with the highest p_T is used. If more than one species of tag is available then they are used in decreasing order of preference: muon, then electron, then kaon.

The overall tagging “power” is represented by the quantity ϵD^2 , defined as

$$\epsilon D^2 \equiv \epsilon (1 - 2\omega)^2, \quad (2.5)$$

where ϵ is the efficiency with which tagging can be performed at all and ω is the wrong-tag fraction, enhanced by a factor of two as a wrong-tag not only fails to tag the b-hadron but actually contaminates the set of tagged events. Simulations of the tagging performance indicate that $\epsilon \sim 40\%$ and $\omega \sim 30\%$ for typical b-decays [95].

Chapter 3

Trackless RICH pattern recognition

“Science in the modern world has many uses; its chief use, however, is to provide long words to cover the errors of the RICH.”

— G. K. Chesterton, 1874–1936 (misappropriated)

In this chapter, the standard RICH pattern recognition system is described and is extended to investigate the possibility of ring-finding without use of tracking information. The potential of this approach for studying rings for which no track information is present and for identifying background hits is explored. Section 3.1 describes the standard particle ID algorithms and Section 3.2 onward describe the trackless ring-finding system and its impact on particle ID.

3.1 The standard RICH particle ID algorithm

Before considering the various facets of trackless ring-finding, we need to examine the standard “track-aware” pattern recognition and particle ID system. This consists of two main pattern recognition algorithms for matching activated pixels to reconstructed tracks, known as the “global” and “local” approaches[131]. These are used together with an algorithm for assigning a probability of being a background hit for each activated pixel to form the standard RICH particle ID system.

3.1.1 Global pattern recognition

The “global” pattern recognition method treats a RICH event as a single entity and maximises a likelihood measure across all particle ID hypotheses for each reconstructed track. The likelihood is calculated by comparing the number of activated pixels with the number expected from the Čerenkov emission signal and background distributions, including photon scattering in the RICH1 aerogel. The fitting of Čerenkov angles for all tracks is performed by inverting the RICH optics analytically. This involves the solution of a quartic equation which accounts for much of the computational time required by the global approach [99]. Comparison of the fitting function to the observed photon detector signals yields a log-likelihood function:

$$\ln \mathcal{L} = - \sum_{\text{track } j} \mu_j + \sum_{\text{pixel } i} n_i \ln \left(\sum_{\text{track } j} a_{ij} + b_i \right), \quad (3.1)$$

where a_{ij} is the expected number of photoelectrons detected in pixel i from track j under a set of hypotheses. μ_j is the expected total number of photoelectrons from track j , $\mu_j = \sum_i a_{ij}$. n_i is the number of photoelectrons observed in pixel i and b_i is the expected background in the same pixel from untracked rings. This likelihood form is derived from a consideration of the contributing Poisson probabilities of pixel occupancy [132]. The sums over i are kept efficient by only summing over *activated* pixels in the second term. Iterations of the set of particle ID hypotheses are made until no further likelihood improvement is observed.

3.1.2 Local pattern recognition

In the local pattern recognition algorithm, each reconstructed track is treated independently and each pixel’s Čerenkov angle is calculated relative to that track. Treating the tracks independently reduces the number of calculations required for each iteration and makes the local method faster than the global method. For each track a log-likelihood function is calculated, based on an assumed Gaussian distribution of hit probability in Čerenkov angle for each track and particle ID hypothesis:

$$\ln \mathcal{L}(\theta_{\mathcal{H}}) = \sum_i \ln \left(1 + \frac{1}{\sqrt{2\pi}\sigma_\theta} \exp \left(-\frac{(\theta_i - \theta_{\mathcal{H}})^2}{2\sigma_\theta^2} \right) \right). \quad (3.2)$$

In equation (3.2), θ_i is the reconstructed emission angle of hit i , $\theta_{\mathcal{H}}$ is the expected emission angle under particle ID hypothesis $\mathcal{H} \in \{\{e, \mu, \pi\}, \{K, p\}\}$ and σ_θ is the angular resolution of the detector. The function is normalised such that $\ln \mathcal{L}(\theta)$ is equal to the hit density expected at angle θ [133]. The sum is over all i , except for those where $\theta_i \gg \theta_{\text{sat}}$, the saturated emission angle, which are omitted for computational efficiency.

This function is used to calculate the number of hits for each hypothesis at the mean emission angle. From this a Poisson probability is calculated and used to discriminate between different hypotheses for each track. Since this method is much faster than the global algorithm, it is used for seeding the global approach.

3.1.3 Other pattern recognition methods

In addition to the standard global and local particle ID algorithms, several subsidiary methods are under investigation. Firstly, the global algorithm is too slow for operation in the high level trigger (HLT) and the analytic inversion of the RICH optics to calculate the Čerenkov angles of hits relative to reconstructed tracks has been identified as a processing bottleneck. As a result, HLT studies are underway to incorporate RICH particle ID into the HLT by using a simple parameterisation and a variant of the local particle ID method to map between hit position and emission angle [134, 135].

Secondly, studies have been made of using a stereographic projection to re-fit rings to tracks: the properties of the projection result in a reduction of the track error. This work has been summarised in reference [136].

Finally, work on trackless ring-finding with a Markov chain sampler is the subject of the work presented in this chapter.

3.1.4 Standard particle ID algorithm structure

A benefit of the global approach to pattern recognition is that, since the dominant background to reconstruction of a ring from a given track is the hits from other tracks, these backgrounds are naturally accounted for in the construction of the global log-likelihood function. This is much harder to do within the framework of the local algorithm where the log-likelihood constructed for each track is unaware of the specific locations of other tracks.

The approach used by the standard particle ID system is to first execute the fast local algorithm to seed the global algorithm hypotheses. This offers an improvement over the uninformed best-guess that all the tracks correspond to pions. Next a special background estimation algorithm is executed which assigns background contribution weights to each pixel based on the global tracking information. The background estimator is followed by a first run of the global particle ID algorithm.

Since the background estimation can be improved using the results of the global fit, a second round of background estimation and global particle ID is executed to obtain the final maximum-likelihood particle ID hypotheses.

3.1.5 Standard particle ID performance

The number of activated pixels per event is shown in Figure 3.1 for typical signal events which decay in the experiment's acceptance and pass the L0 trigger. It is evident that a large proportion of the hits originate from secondary interactions, in fact the secondaries account for $\sim 70\%$ of all hits. Secondly, the maximum number of hits in RICH1 (about 3000) is about twice as much as that for RICH2. This is because, since RICH2 is intended for particle ID of high-momentum particles, many of the potential radiating tracks have been swept out of the acceptance by the magnetic field by the time they reach it. These hits are distributed on the photon detector planes with maximum occupancy of $\sim 8\%$ in the central region of RICH1 and $\sim 1\%$ elsewhere, including all of RICH2 [99].

Property	Aerogel	C ₄ F ₁₀	CF ₄
N_{pe}	6.6	32.7	18.4
$\sigma_{\theta}^{\text{emission}}$ [mrad]	0.60	0.74	0.31
$\sigma_{\theta}^{\text{chromatic}}$ [mrad]	1.61	0.81	0.42
$\sigma_{\theta}^{\text{pixel}}$ [mrad]	0.78	0.83	0.18
$\sigma_{\theta}^{\text{track}}$ [mrad]	0.26	0.42	0.20
$\sigma_{\theta}^{\text{total}}$ [mrad]	2.00	1.45	0.58

Table 3.1: Number of photoelectrons per ring and contributions to Čerenkov emission angle resolution in the various RICH radiators using the global ring-finding algorithm with background estimation [99]. The various resolution contributions are respectively from the uncertainty on the emission point position, chromatic aberration, pixel discreteness and tracking errors.

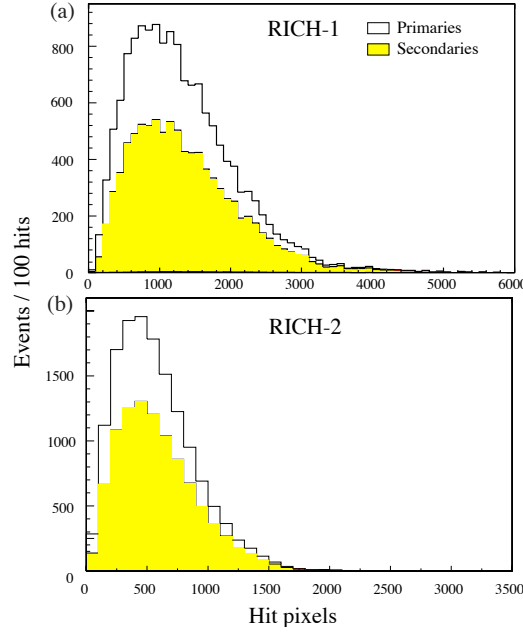


Figure 3.1: Number of activated pixels per event in RICH1 and RICH2.

The expected number of detected photoelectrons and various contributions to the Čerenkov angle resolution are listed in Table 3.1. The various contributions to the emission angle resolution for the RICH1 C_4F_{10} radiator are illustrated in Figure 3.2; the final resolution is a Gaussian of width $\sigma_\theta = 1.45$ mrad with a small tail due to tracking errors [99].

Tests of the RICH particle ID performance have been made with many decay channels: the particle ID results are shown in Table 3.2 where the matrix would ideally be diagonal [99]. It can be seen that the off-diagonal components are relatively small, indicating overall good particle ID performance. The efficiency ε (the fraction of true particles of a given type which are correctly identified) and purity p (the fraction of tracks identified as a particular particle type which are actually of that type) are also indicated. The efficiencies are typically high, in the vicinity of 80% and the purities are also high except for muons, which suffer significant pion contamination. This is compensated by the dedicated muon detectors, which are the dominant contribution to muon tagging. The most important contributions to the particle ID matrix are the performances for π and K identification, since the ability to separate these efficiently is crucial to the physics performance of LHCb.

The separation between particle ID hypotheses can be expressed in Gaussian standard

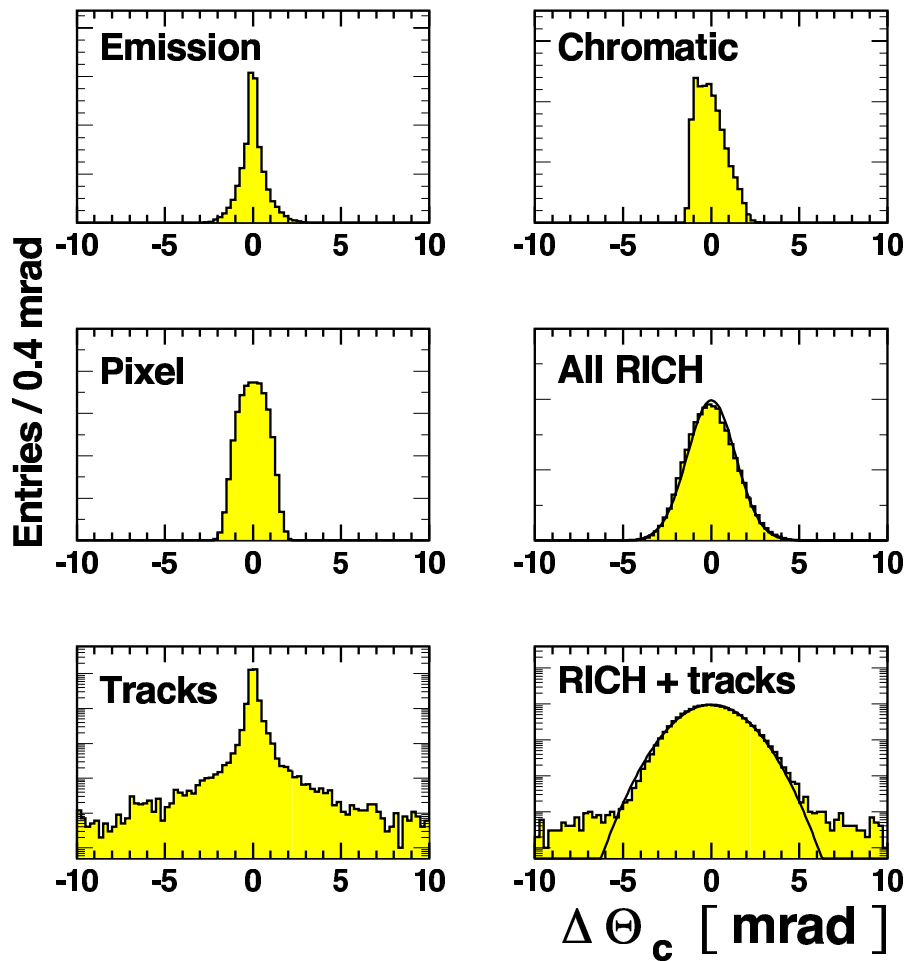


Figure 3.2: Various contributions to the Čerenkov angle resolution in the RICH1 C_4F_{10} radiator. The emission point, chromatic and pixel error distributions are shown with linear scales, as is their convolution which is marked as “all RICH”. The contribution from the tracking resolution and the final convolution are shown with logarithmic scales.

Rec.	True particle type						p
	e	μ	π	K	p	X	
e	97.4	0.7	24.6	1.4	0.5	3.1	0.76
μ	4.0	8.8	69.5	2.0	0.5	4.9	0.10
π	2.5	1.3	545.7	3.3	0.7	5.1	0.98
K	0.3	0.1	12.7	70.6	4.8	4.3	0.76
p	0.2	0.0	1.7	4.3	35.9	0.0	0.85
X	9.9	0.8	19.8	3.2	0.0	55.6	0.62
ε	0.85	0.76	0.81	0.83	0.85	0.76	

Table 3.2: Particle ID matrix for the standard RICH particle ID system. The entries in the matrix correspond to tracks of all momenta from $\sim 80k$ typical events but have been normalised to 1000. The efficiency and purity numbers for each particle type are shown at the bottom and the right-hand side respectively. The X label is used to indicate a track below the Čerenkov-radiating threshold.

deviations via the relation $N_\sigma = \sqrt{2\Delta \ln \mathcal{L}}$, where the likelihood \mathcal{L} is the global algorithm definition in equation (3.1) and the difference is obtained by changing the particle ID hypothesis of one track [99]. The average separation between pion and kaon hypotheses for true pions in triggered signal events in the acceptance is shown in Figure 3.3. Better than 3σ separation is achieved for particles with momenta between 2 and 100 GeV/c. Threshold effects and resolution restrictions reduce the separations at low and high momenta respectively but useful discriminating power is available from ~ 1 –150 GeV/c. Figure 3.3 breaks down this separation by event multiplicity, defined by bins of 0–1000, 1000–2000 and > 2000 hits. As expected, high multiplicity events give poorer separation in the central momentum region but performance is not strongly affected by multiplicity at low and high momenta.

The particle ID efficiency and mis-ID rates are often used as characteristic measures of RICH performance. The usual way it is considered is to only determine a track as being “light” (e, μ , π) or “heavy” (K, p) to reduce the matrix in Table 3.2 to a binary choice. Figure 3.4 shows the light/heavy efficiencies and mis-ID rates against momenta for the standard particle ID algorithms, where again the performance reductions at low and high momenta are due to threshold and resolution effects. Example total particle ID efficiencies for the standard particle ID system with different RICH detectors in use, obtained from 8k $B_d^0 \rightarrow K^+ K^-$ signal events, are listed in Table 3.3. The particle ID rates are generally higher in the cleaner environment of RICH2.

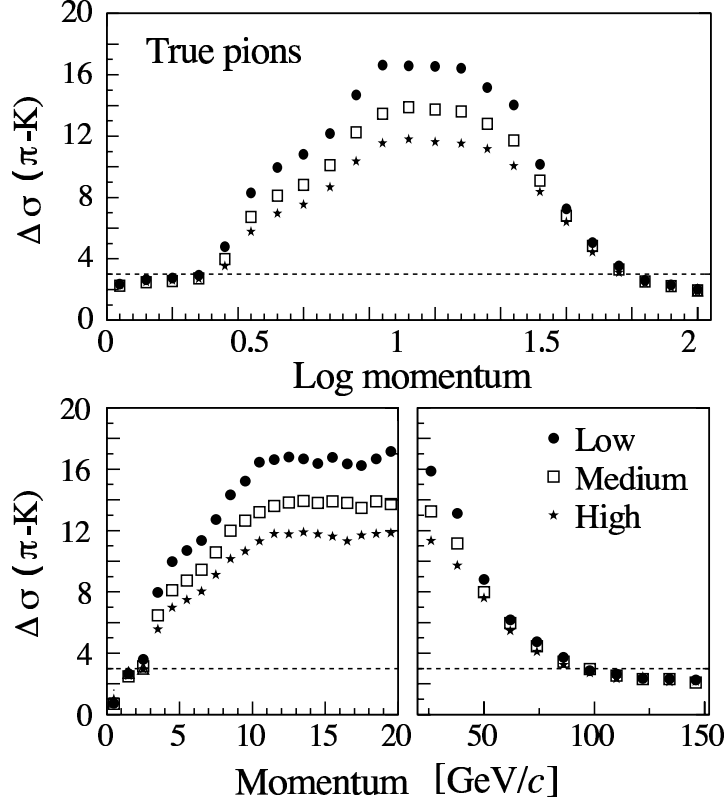


Figure 3.3: The mean particle ID likelihood π -K separation for true pions, separated into curves for low, high and medium multiplicity events as defined in the text.

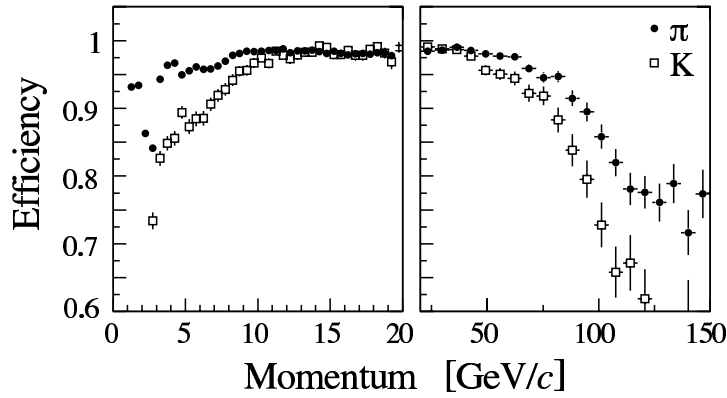


Figure 3.4: Particle ID efficiencies for pions and kaons using the standard particle ID system, binned by momentum.

Detectors used	“K” efficiency [%]	“ π ” efficiency [%]
RICH1 only	86.37 ± 0.27	89.51 ± 0.10
RICH2 only	92.01 ± 0.32	96.71 ± 0.07
RICH1 & 2	88.33 ± 0.24	92.52 ± 0.07

Table 3.3: Particle ID rates for “kaons” (heavy ID) and “pions” (light ID) in 8k $B_d^0 \rightarrow K^+ K^-$ signal events for particles in the acceptance of either or both of the RICH detectors. The rates are defined as the ratio of correctly identified particles of each class (“heavy”/“light”) to the number of MC particles of that class which are reconstructible using the given detector combination. The particle ID rates are generally higher in RICH2 due to the lower occupancy.

3.2 Trackless ring-finding

An alternative approach to the problem of reconstructing Čerenkov rings from pixel data is to fit circles directly to the pixel position distributions without the use of information from the tracking system. Trackless ring-finding is largely motivated by the failure of the global particle ID to find rings where the tracking system has failed to track a physical radiating particle. Such rings are evident to human observers, as can be seen from studying Figure 3.5. Identifying why the tracking system fails to identify the charged particles which produce untracked rings is of interest for ensuring the robustness of the standard reconstruction system ¹.

As a significant amount of information is being discarded in the trackless ring-finding approach, the trackless fitting procedure must scan a larger parameter space than that required by the standard procedure. The non-reliance on tracking information means that the trackless ring-finder can potentially be a useful tool for characterising the standard ring-finding method in cases where rings are produced by untracked particles.

3.2.1 The trackless ring-finding system

The trackless ring-finding system described here is based on a multi-dimensional Markov chain sampler which is designed to efficiently explore the parameter space describing ring sizes and positions. The ring-finder finds sets of rings which describe the hit pixel

¹Note that we are using “trackless” to describe the ring-finding procedure which does not depend on tracking information and “untracked” to characterise RICH rings without any corresponding reconstructed track. The two usages are quite distinct.

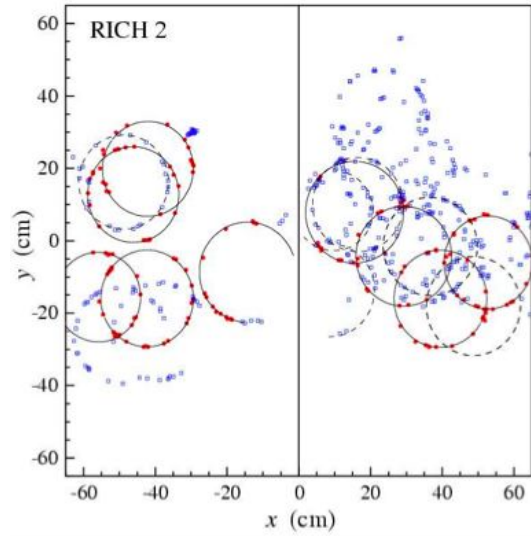


Figure 3.5: Čerenkov rings reconstructed using the standard (global) pattern recognition system in RICH2. Solid rings with red pixels are those identified by the global pattern recognition; blue pixels are considered as background. There are clear sets of pixels at the bottom of the left-hand panel and the top of the right-hand panel which have not been reconstructed by the track-based ring-finder.

distributions well and which are compatible with supplied prior probabilities for the ring parameters. However, it does not attempt to associate pixels to rings or consider the matching of tracks and rings at all — these must be done by hand after the sampler has returned a ring set. In this study we first describe the algorithm used to identify RICH rings without using tracking information, then consider how to associate hits and tracks to the resulting rings. The effects of this system on particle ID performance are considered in Section 3.3.3 onward.

The trackless ring-finder system is written in object-oriented C++ and incorporated into the Gaudi framework as a set of Brunel algorithms. The development has been performed with Brunel v24.4 reconstruction framework and also makes use of some features from DaVinci v12.6. Extensive object templation is used to allow a single sampler algorithm to be instantiated four times for each event, each instantiation corresponding to one of the four RICH photon detector panels with appropriate angle-to-length scaling parameters, geometric constraints and prior distributions.

The first stage of processing extracts the pixel hit positions for each event from the digitisation data produced by Boole and converts these positions into a form used by the Markov chain ring-finding sampler. The conversion from reconstruction pixels to

sampler hits includes a transformation for RICH2 to correct for ring distortions: this transformation is described in Section 3.2.2. The ring-finding algorithm is described in Section 3.2.3.

The stages after the sampler use the circle-to-hit association probabilities from the sampler to assign pixels to each ring and also use tracking information from another part of Brunel to match ring candidates to track-segments in the RICH radiators. These aspects of the ring-finding algorithms are discussed in Section 3.2.4.

Finally, an optional monitor algorithm is executed. This makes use of the reconstruction objects produced by the ring-finder algorithm and MC truth information to analyse the performance of the ring-finder and the association of its output to reconstructed tracks. The structure of the monitor algorithm and the information which it produces are discussed in Section 3.3.2 and Section 3.3.3 respectively.

Due to the parameterisation of rings as circles in the sampler, it is naturally best-suited to use in RICH2 where optical distortions are small and the activated pixel density is relatively low. The higher occupancy and more substantial optical distortions in RICH1 make it less-suited to a trackless approach. RICH1 also has two radiator media, the aerogel and the C_4F_{10} gas, and therefore the ring radius distribution is bimodal. This presents some difficulties to the sampler, especially as the aerogel rings, with a mean radiator length of only 5 cm, are very sparsely populated. For these reasons, the results shown here only apply to ring-finding in RICH2.

3.2.2 Removing ring distortions

Due to the realities of the RICH optics, the rings projected on to the detector planes are not in general perfect circles but instead are distorted to form rough ellipses. The variation in radius of photon hits from the projected radiating track with azimuthal angle as simulated by Brunel is shown in Figure 3.6. These data were obtained by projecting all the long tracks in 100 typical events on to the photon detector planes and ray-tracing photons uniformly around the track segment vectors at $\theta_C = 30$ mrad [135].

The RICH2 distortions are described very well by a sinusoidal variation as indicated in Figure 3.6(b):

$$\Delta r = A \cos 2\phi, \quad (3.3)$$

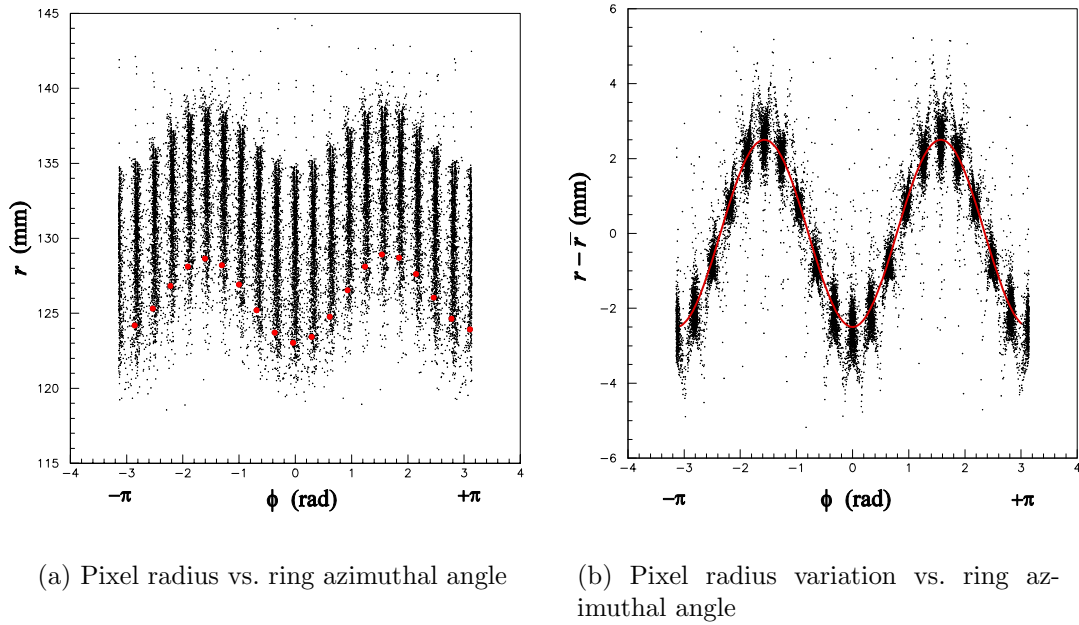


Figure 3.6: Čerenkov ring distortions in RICH2, expressed as the variation in ring radius with azimuthal angle for long tracks in 100 typical events. Figure (a) shows the raw pixel radii and Figure (b) shows the deviations of pixel radius from the mean pixel radius per ring. The red dots in (a) show the azimuthal variation in ring radius around a single ring.

where $A = 2.5$ mm and ϕ is the azimuthal angle of ray-tracing emission. This expression corresponds to a “squeezing” in the x direction and “stretching” in the y direction which can be parametrically removed. The hits which are supplied to the Markov chain ring finder are defined by a 2D vector (\tilde{x}, \tilde{y}) in the photon detector plane: in RICH2, \tilde{x} and \tilde{y} are obtained from the pixel position (x, y) by the transformations

$$\tilde{x} = \frac{x}{1 + \frac{2.5}{131}} \quad \tilde{y} = \frac{y}{1 - \frac{2.5}{131}}, \quad (3.4)$$

where the 2.5/131 factors come from the relative magnitudes of the sinusoidal radius variation and the mean radius in Figure 3.6.

After the Markov sampler has run and hits have been nominally assigned to the resulting set of circles, the reconstruction rings are built with the physical pixel positions denoting each hit, so the transformation is only applied to the data as seen by the internal mechanisms of the trackless ring-finder.

3.2.3 The trackless ring-finding algorithm

The approach used for trackless RICH ring-finding is an iterative process based on a Markov chain sampler. The Markov chain approach involves evaluating a characteristic function at various points in the space of configuration parameters and making the decision about which point in the configuration space to “visit” next based on the characteristics of the current position and an algorithm specific to the sampling method. The intention is not to evaluate “all” points in the space or even a granular approximation to that as might be attempted by a regularly-spaced grid scan of the parameter space. Instead, the Markov chain samples the space with a position-dependent density which represents the structure of the function being evaluated, such that most visited points are in the “typical set” of the space [137]. Sampling of variables such as the ring radius and position distributions (and the number of rings) is assisted by the use of prior probability distributions in the sampler.

In defining our sampling space, a ring-finder circle is represented by 3 parameters: two to describe the ring-centre position on the detector plane and one for the ring radius. This parameterisation assumes that the rings are perfectly circular for all centre positions on the detector planes, that each detector plane has 100% photon acceptance across its area and that emission of a Čerenkov photon is equally likely for all azimuthal angles, all of which are reasonable assumptions in RICH2.

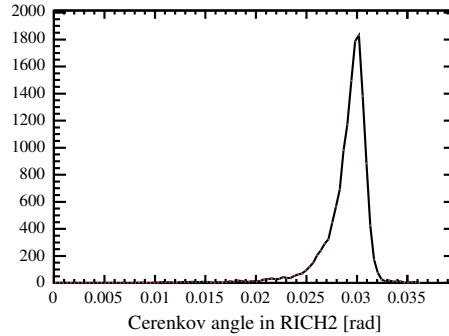


Figure 3.7: True Čerenkov emission angle distribution for RICH2, used as a prior in the ring-finding sampler. The y-values are the number of rings per bin, derived from a sample of typical B events.

It should be noted at this point that this sampling problem is different from more typical minimisation scenarios in that the parameter space has a different number of dimensions depending on the number of rings being considered at any sampler iteration ($3N$ dimensions for N rings). The approach used in this sampler is to include the number of rings as a sampling parameter and to naïvely change the number of dimensions: this introduces a strictly undefined sampler behaviour which in practice does not cause problems.

The core of the trackless ring-finder is an implementation of the Metropolis-Hastings sampling algorithm [137] — the general M-H algorithm is described in Appendix A. The characteristic function used by the sampler is a measure of the probability of the proposed rings to the measured hit distribution. This is calculated over a set of Bayesian priors on the reconstructed emission angle, the average hit densities on the photon detector planes and the number of pixels per ring against ring radius. Of these priors, only the emission angle distribution is directly derived from the simulation: Figure 3.7 shows the RICH2 emission angle prior distribution.

Ring-fitting is performed via an approach where the number of rings is sampled from a prior distribution generated from the detector simulation Monte Carlo program and the ring distributions are then proposed by iterative “jittering” of the circle centre positions and radii to match at least three hits. As the procedure is computationally intensive, caching is used extensively to allow “dead-end” sampling trajectories to be rapidly discarded for a previous known-good configuration. The details of the sampling procedure are described in reference [138].

The probability assigned to a given parameter space point is generated by assuming

that Čerenkov emission is azimuthally uniform around the radiating track segment. The prior probability of a hit occurring at a given point due to a particular Markov circle is hence only a function of the radial position of that hit relative to the ring centre. The choice of functional form is similar to a Gaussian around the ring's nominal radius r_0 but since a naïve definition such as $p(r|r_0) \propto \frac{1}{2\pi r} \exp\left(-\frac{(r-r_0)^2}{2\sigma^2}\right)$ has continuity problems at $r = 0$ we instead use a modified Gaussian in $\log r$ space, with a parameter ϵ describing the width of the peak and an additional n parameter to correct for the log-space skew:

$$p(r|r_0, \epsilon, \alpha) \propto r^n \exp\left(-\frac{\log^2(r/r_0)}{2\epsilon}\right). \quad (3.5)$$

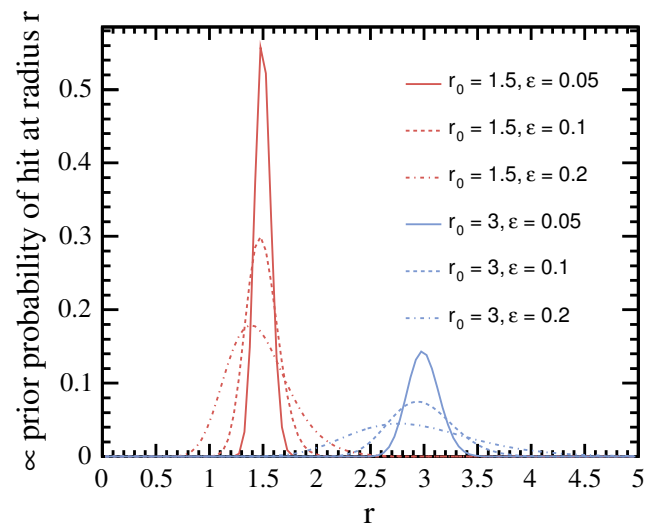
The form of the function is then constrained to ensure that the mean of the distribution is always located at r_0 , introducing $\alpha = n + 2$ and incorporating a r^{-2} dependence from area-wise normalisation:

$$p(r|r_0, \epsilon, \alpha) = \frac{1}{Z} \frac{1}{2\pi r^2} \frac{e^{-\alpha^2 \epsilon^2 / 2}}{\sqrt{2\pi \epsilon^2}} \left(\frac{r}{r_0}\right)^{\alpha-2} \exp\left(-\frac{\log^2(r/r_0)}{2\epsilon^2}\right) \quad (3.6)$$

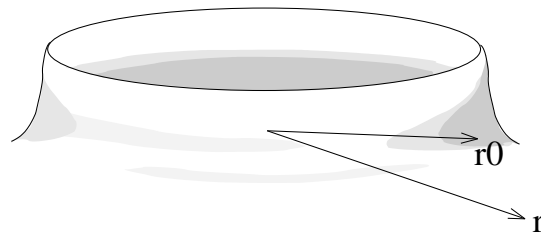
where Z is a normalisation factor. The values of α and ϵ are fixed for a given run and so the only variations are in r_0 . The function is plotted in Figure 3.8 for values of ϵ between 0 and 0.3 with $\alpha = 2$ and $r_0 = \{1.5, 3\}$. In practice, $\alpha = 2$ was used for sampling as it provides a good fit to the Čerenkov angle distribution.

Unlike the standard particle ID algorithm, which performs its calculations in terms of Čerenkov angles, the trackless algorithm performs its manipulations on the photon detector plane. As rings can fall partially on the plane, ring-fitting near the edges incorporates the concept that truncated rings may be seen and adapts the likelihood measure accordingly, as very asymmetric hit distributions with a fraction of the normal hits are normally disfavoured. However, due to the implementation of the algorithm as four independent samplers on four separate photon detector panels, communication between samplers is not possible and rings which are split across the boundary between two panels are not recognised as a single entity. This is a reasonably safe approximation in RICH2.

The result of the sampler is a best fit of rings to data as weighted by the priors. For example, single-hit rings are not unfeasible but an entire detector full of them is *very* unlikely and so the “one ring for each pixel” scenario, despite offering an excellent fit to the data, never occurs in a reasonable sampling scenario.



(a)



(b)

Figure 3.8: The “Yorkshire pudding” function, parameterising hit probability on a Čerenkov ring in terms of ring radius.

3.2.4 Associating rings to pixels and tracks

The collection of sampler-proposed rings can now be matched against pixels and reconstructed track segments. The resulting set of reconstructed rings, each of which “owns” a set of pixels and is optionally matched to a reconstructed radiating track segment, is similar to the output of the standard particle ID system. The major difference is that rings without matched track segments cannot occur in the standard approach.

Matching pixels to rings

A decision has to be made to determine which pixels are to be considered as “belonging” to each prototype ring. The sampler algorithm has been written to provide some flexibility at this point and run-time configuration directives can be used to specify whether the matching is to be purely based on geometric proximity of the hit and the ring or whether probabilities are to be used. Whether or not each pixel must be *uniquely* associated to only one ring (or to no ring at all) or if a pixel may be considered as belonging to any number of rings is also specified at run-time.

We consider first the geometric matching procedure: in this the radial distance from a hit to the centre of each ring is calculated and whether the hit is considered as being due to that ring is decided by whether it lies within a small range of the ring’s radius. To make the association algorithm relatively independent of ring size, this range is specified in units of the ring radius, so to pass this cut the following inequality must evaluate as true:

$$|r_{\text{hit}} - r_{\text{ring}}| \leq a_{\text{geom}} r_{\text{ring}}, \quad (3.7)$$

where a_{geom} is the tuneable cut parameter. By default this is set to 10%, after some exploratory “by-eye” tuning with the Panoramix visualisation system. In the unique-matching configuration, each hit will only be assigned to one ring: that which meets this constraint and which produces the smallest term on the left-hand side of equation (3.7).

By contrast, the probabilistic ring–pixel matching procedure uses the sampler probabilities directly: while these are fundamentally derived from the relative geometry of hits and circles they also incorporate information about prior probability distributions and are a more informed measure of association likelihood.

There is a catch in using raw probabilities as a measure, however. This can be

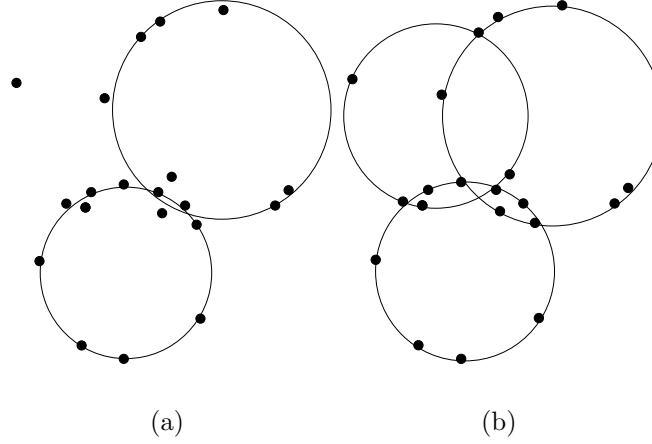


Figure 3.9: Two different contexts in which hits can be interpreted. In (a) the indicated hits are produced by two rings while in (b) they are in the vicinity of three rings. Due to probability normalisation, the hits in case (b) will have a reduced probability of association to any given ring, even if the geometric match is just as good. This illustrates the danger of using absolute probabilities to perform pixel–ring matching.

illustrated by considering an identical hit in two contexts: one where only one circle is a realistic candidate for having produced the hit and one where several rings are equally probable, as illustrated in Figure 3.9. Since the hit-to-circle association probabilities are normalised, the hit in the vicinity of several circles will have a fraction of the probability of having come from any one ring than it would have done if the two other rings had not been there. Hence we have to be careful not to use absolute probabilities: the method used is to record the association probabilities for all combinations of hits and circles and to identify the most probable circle for each hit.

An initial cut is then applied which requires that the pixel must have a background probability (i.e. the probability of not having come from any rings at all) less than 50%. Rings are then associated to the hits whose associated pixel passes this cut. A given ring is associated to a pixel if the corresponding circle has a match to the hit which compares favourably to the matching probability of the best match combination. In other words, for a circle C_i to be associated to a hit h_j , the following inequality must be satisfied:

$$p(C_i | h_j) \geq a_{\text{prob}} p(C_{\text{best}} | h_j), \quad (3.8)$$

where the a_{prob} parameter is used to parameterise the acceptance of the association procedure. Reasonable results are found for low values of a_{prob} , between $\sim 1\text{--}5\%$, established

primarily via “by-eye” tuning using the Panoramix system. In the results to be shown, non-unique probabilistic matching is used, with a tuneable value for a_{prob} .

Matching track segments to rings

As the sampler operates purely in terms of hits, there is no probabilistic matching possible from sampler circles to track segments in the Čerenkov radiators. We find the track segment for each ring by projecting all segments through the RICH optics to produce a set of track segment positions on the photon detector plane. These should then correspond to the centres of reconstructed Čerenkov rings and for each ring the associated reconstructed track segment is taken to be that which is closest matched to the centre of the ring, provided that the following inequality is satisfied:

$$d_{\text{seg}} \leq a_{\text{seg}} r_{\text{ring}}. \quad (3.9)$$

where d_{seg} is the distance from the ring centre to the projected segment position on the detector plane.

Here the a_{seg} parameter characterises a degree of tolerance in the association procedure: values of a_{seg} greater than 1 mean that a track segment whose projection lies outside of the reconstructed circle may be considered as having produced that ring. While reconstruction errors mean that this scenario is possible, it is not common and so a_{seg} is taken to have a default value of 50%. Circles which do not match any reconstructed segments are turned into reconstruction rings without a matched segment: these are termed “untracked” rings and can be divided into sampler artifacts and genuine untracked physical rings. The latter set are of great interest, as their existence motivates the trackless ring-finding approach.

3.3 Trackless ring-finder performance in RICH2

There are many difficulties in objectively assessing the performance of the trackless ring-finding algorithm. The principal difficulty is the lack of a robust definition of a “good” match from a ring to a track: as ever we must consider the reconstruction performance in terms of how closely it matches the MC event but the measure of closeness is ill-defined. The definitions used in this study are described in Section 3.3.1. The algorithm used to monitor the ring-finder’s performance is described in Section 3.3.2. Finally, the

ring-finder performance is summarised in Section 3.3.3.

3.3.1 Defining MC-to-reconstruction associations

The central problem of characterising the trackless ring-finder’s performance is that matching reconstructed Čerenkov rings to their MC counterparts is not well-defined: reconstructed rings may be comprised of pixels from several MC rings or from background processes like direct activation of pixels by tracks. A characteristic “goodness of fit” measure is required, of which there may be several candidates.

The main object associations which are required in RICH reconstruction are those from pixel to ring and from a ring to a track segment. For pixels, there is the possibility that one activated pixel is actually the result of more than one Čerenkov photon hit within the photon detector’s spatial resolution. For rings, it is possible to reconstruct sets of rings which appear plausible in terms of the observed pixels but none of which correspond to the idealised pixel probability distributions from MC radiator segments. And finally, a given reconstructed ring may be due to several segments but is unlikely to exactly match the parameters of an idealised ring from that segment.

From the point of view of reconstruction, what we really want to do is to form an association between Markov rings and MC particles. We have already described in Section 3.2.4 the matching of a ring to a reconstructed track segment by projection of the segment through the RICH optics. The Monte Carlo analogue of this matching is to follow the MC association chain: starting with a Markov ring we obtain the MC pixel associated to each reconstructed pixel on the ring. The MC pixel then provides a link to the MC Čerenkov photon which activated it; from there the MC segment can be found and finally the MC track-to-particle association can be made, linking each reconstructed pixel to an MC particle.

3.3.2 Monitoring the algorithm performance

Assessing the performance of the ring-finding algorithm requires an external monitoring system with access to the “true” Monte Carlo version of the event as well as the output of the reconstruction system. This is achieved in the Gaudi framework by placing a second algorithm in the execution chain after the ring-finder has saved its reconstruction objects to the reconstructed event store.

The initial stages of the monitor algorithm are devoted simply to retrieving the event descriptions according to the MC event and from the reconstruction chain, which includes the ring-finder algorithms either trackless or track-aware. Once these have been retrieved, some basic event-level statistics such as the number of rings and the number of pixels in the event are recorded.

Secondly, each Markov ring is examined and the MC counterparts of the assigned ring pixels are used to make a link to the MC particle responsible for that pixel's activation. The list of MC particles for each ring is then used to establish the MC particle responsible for the majority of pixels on the ring and to record some more statistics about the quality of the MC matching for each ring.

Before any further comparisons between the MC event and the Markov-reconstructed event are made, it is required that a ring has at least one non-null association to an MC particle. In reality, there is no significant attrition rate due to this requirement: almost all reconstructed pixels have at least one associated MC pixel, from which a good link to an MC particle is guaranteed: the only notable exceptions are those pixels which are activated from spillover events and those are unlikely to be responsible for *all* the pixels on a ring.

At this point the best match from a given Markov ring to MC particle can be taken to be that MC particle which produced most of the MC pixels on the ring. In addition, similarly to the criteria for making reconstructed ring-to-pixel associations, a cutoff on the minimum fraction of a ring's particles which must be associated to a single MC particle is applied: this is used to restrict the comparisons of untracked Markov rings with Monte Carlo to Markov rings which are genuinely produced by one MC particle. The cutoff fraction of MC pixels from a single MC particle to define a Markov ring as "good" or "bad" is set as 70%. We should note that the earlier criteria for binding particular pixels to a Markov ring mean that this definition is not robust at some level, but useful results can be extracted for the given cutoff values.

"Good" rings with a good majority association to a single MC particle can now be sub-divided into the tracked and untracked sets by observing whether the algorithm has managed to geometrically match a reconstructed track segment to the ring. If there is no such segment, then the ring must be considered untracked and the associated MC particle decay tree is analysed to evaluate why the ring has not been tracked.

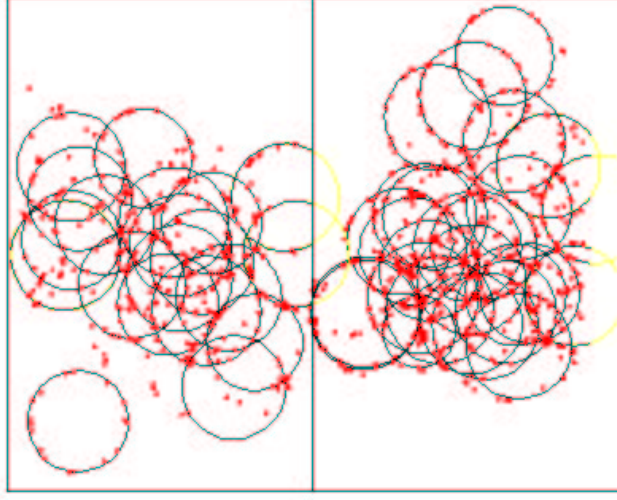


Figure 3.10: Čerenkov rings reconstructed using the trackless ring-finder in RICH2. The yellow ring-segments are those parts which are outside the photon-detector acceptance. Note that yellow ring-segments are observed *in* the acceptance of the right-hand panel: this is because the algorithm used to construct a ring based on pixels in the left-hand panel considers pixels in the right-hand panel to be out of its acceptance. In fact, no pixels are observed on the yellow ring segments, due to `CORRECTION_TODO`

3.3.3 Performance of the trackless ring-finder

The trackless ring-finder has been tested for RICH2 only, with the RICH1 radiators and reconstruction disabled. Studies have been performed using 8k $B_d^0 \rightarrow K^+ K^-$ signal events, which produce sizable populations of both signal kaons and background pions. These events were generated using the Gauss/Boole simulation system described in Section 2.6.2. No random noise background was added to the HPD output, since testbeam studies of HPDs indicate that the rate of dark count electrons hitting the pixel chip is ~ 0.005 electrons per event per HPD at $25^\circ C$: a negligibly small rate [139]. A typical event reconstructed in RICH2 using the trackless ring-finder is shown in Figure 3.10, where the main significant feature is the absence of any obvious unreconstructed rings. It will also be useful to note the complexity and incoherence of the ring structure in the busy central regions of each panel.

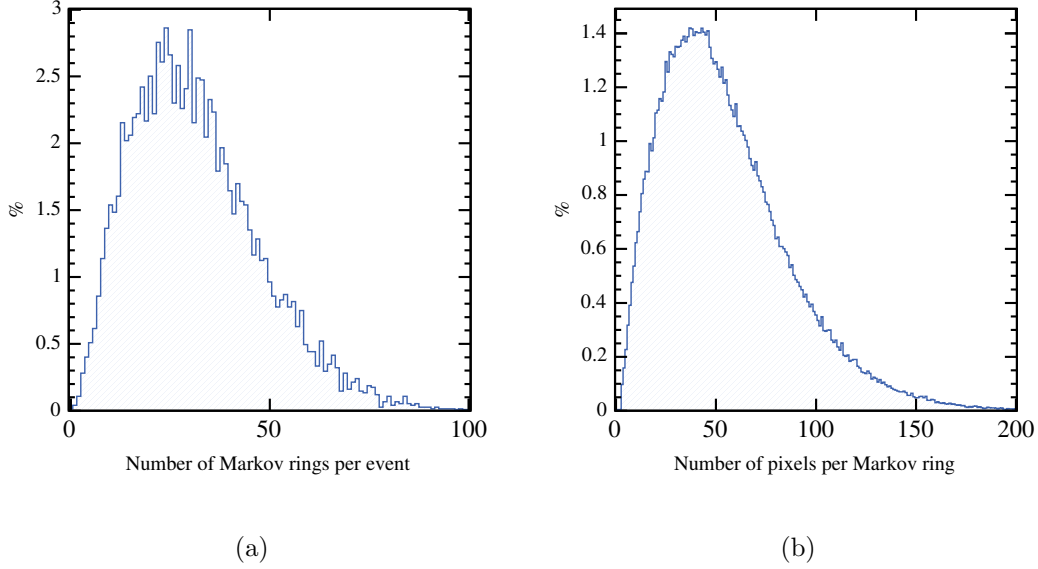


Figure 3.11: Characteristic numbers of rings and pixels in RICH2: (a) the number of Markov rings per event and (b) the number of pixels per Markov ring. The vertical scale represents the fraction of the total number of rings/pixels per bin, with 100 bins in plot (a) and 200 bins in plot (b).

Numbers of rings and pixels

The number of Markov rings found per event is shown in Figure 3.11(a). Similarly, the number of pixels per ring are shown in Figure 3.11(b). About 20 rings are found per event, in keeping with the mean of 19.6 tracks passing through the CF_4 radiator [140]. The mean number of pixels per ring is around 50.

Purities and efficiencies

Two “levels” of purity and efficiency number can be calculated for the trackless ring-finder: a “ring” level and an “event” level. The ring purity is defined as the fraction of ring pixels which are associated to MC pixels which were activated by the majority MC particle for that ring. Similarly, the ring efficiency is defined as the fraction of pixels associated to MC pixels from the dominant MC particle for the ring (i.e. “true” pixels) which are associated to that ring. These numbers are plotted in Figures 3.12(a) and 3.12(b), where it can be seen that both have low central values. The efficiency has a more uniform distribution than the purity and studies have shown that there is no strong

correlation between the efficiency and purity of a given ring.

The “event-level” purity and efficiency definitions are formed in terms of the fraction of *rings* in the event which meet a criterion (as opposed to the number of pixels per ring, as used for the ring-level definitions). The non-absolute nature of the matching between a proposed ring and the MC ring means that event-level figures are open to some interpretation.

To calculate the event purity we define “good rings” as those Markov rings in an event for which the MC associations obtained by matching projected track segments to rings and by using MC truth from pixels both return the same non-null MC particle. The event purity is then the fraction of Markov rings which are “good”. For 8k $B_d^0 \rightarrow K^+ K^-$ signal events the RICH2 matching rate between projected segment MC association and pixel-wise MC association is 7236 agreements and 326 disagreements (95.7% agreement) for cases where both methods return an MC-association. This indicates that rings where both methods agree are very accurately identified by the sampler, but only $\sim 5\%$ of rings meet this criterion (about one ring per event).

The event efficiency is defined as the number of MC particles identified by pixel-wise MC association divided by the total number of reconstructible MC particles in the event. “Reconstructible” is defined (for track-types “VELO”, “T”, “long”, “upstream” and “downstream”) as follows [95]:

- for “VELO” tracks the particle must give at least 3 r hits and 3 ϕ hits;
- for “T” tracks the particle must give at least 1 x and 1 stereo hit in each of the T1–T3 stations;
- for “long” tracks the particle must be reconstructible as a VELO and a T track;
- for “upstream” tracks the particle must be reconstructible as a VELO track and give at least 3 hits in TT;
- for “downstream” tracks the particle must be reconstructible as a T track and give at least 3 hits in TT.

Note that this defines the track-types: a track may be considered as either “VELO-reconstructible” or “T-reconstructible” if it meets the first two criteria above. A track is considered to be classed according to the most stringent form of reconstructibility that it satisfies. In using this information, no further reconstruction criteria are required of

the MC particles: they do not have to be specifically reconstructible in the vicinity of RICH2, for example. As a result, the “event efficiency” quoted here is likely to be an underestimate compared with a more exact definition.

The event-level efficiency and purity for 8k $B_d^0 \rightarrow K^+ K^-$ events are plotted in Figures 3.12(c) and 3.12(d). Again the distributions are localised in the region between 10–20%. Despite the encouraging appearance of Figure 3.10, neither the ring-level nor the event-level measures of ring-finder performance indicate reliable ring-identification.

The principle reason for this disparity between qualitative indications from e.g. Figure 3.10 and the quantitative purity and efficiency figures shown in this section is that the regions of the RICH2 detector with the highest pixel densities are compatible with many ring-placements generated by the sampler. Images like Figure 3.10 are partnered by a multitude of similar “solutions” in which the well-defined outer rings remain static but those in high density regions shift and change. Only one such solution is “correct”, but it cannot be identified purely based on sampler probabilities. The indication from this study is that tracking information *is* required to break the degeneracy of the solutions identified by a trackless ring sampler.

Note that as stated in Section 3.3.2 all further discussion of “good” Markov rings is derived from rings with a ring purity greater than 70%, corresponding to a tiny fraction of the rings represented by these plots.

Checking for biases

As the ring-finder algorithm relies on a Bayesian measure of fit probability, we need to check that the Bayesian prior distributions do not introduce any substantial biases into observable features of the fit. Figure 3.13 shows the deviation of the sampler-reconstructed Čerenkov emission angle from the true emission angle. The width of this distribution is ~ 1 mrad, which is comparable with the RICH2 Čerenkov angle resolution of ~ 0.7 mrad, and the peak of the distribution shows no significant bias away from the true value.

The distribution is composed of a sharp peak of very well-reconstructed rings and a wider distribution of imperfectly-matched rings with a slight bias favouring higher emission angles. Referring again to Figure 3.10, the sharp peak at zero bias is due to the isolated rings in the outer regions of the detector panels where the ring-finder can operate well, but the wider part of the distribution is due to the higher-occupancy

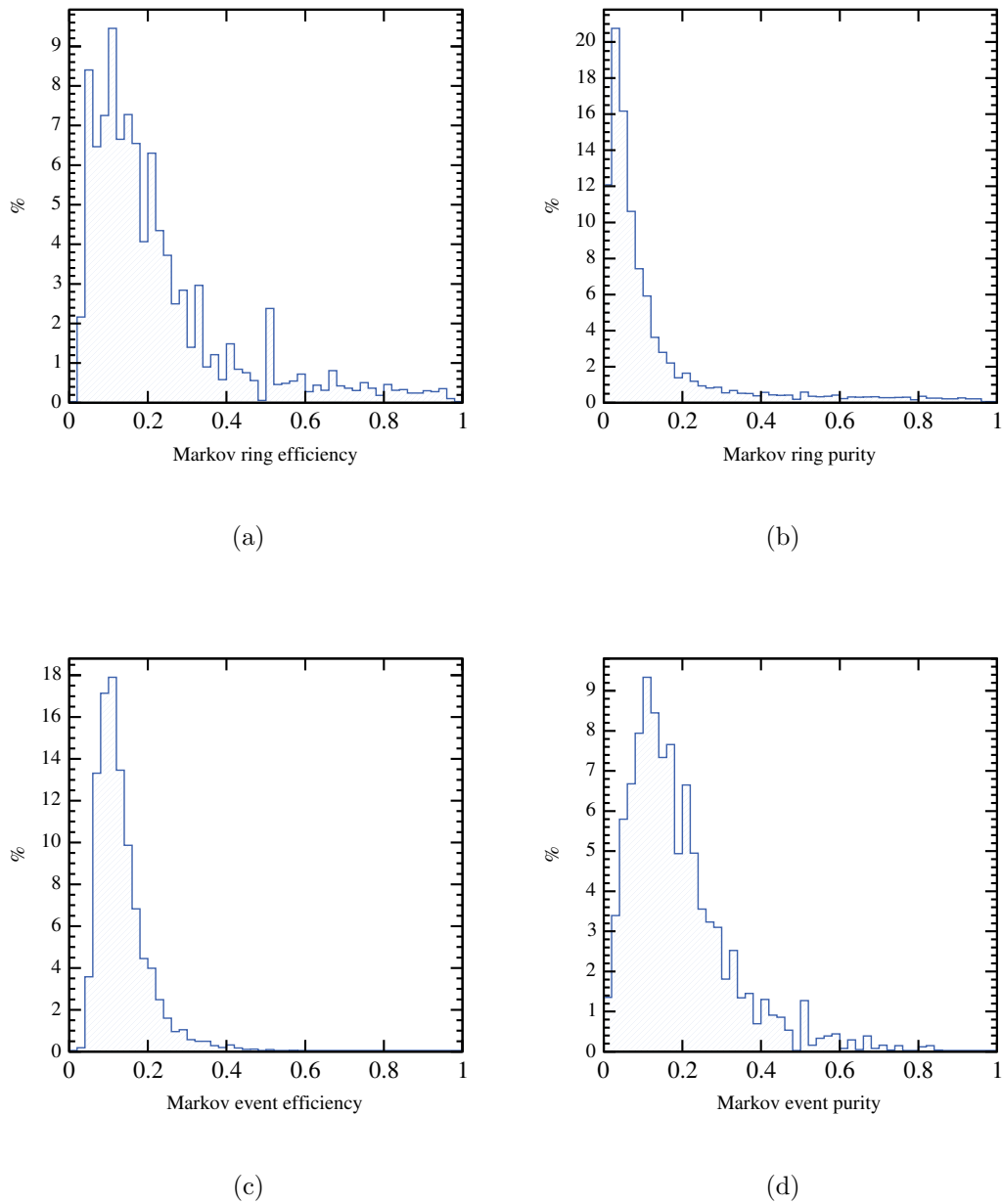


Figure 3.12: Ring and event efficiencies and purities in RICH2. The figures are (a) the ring pixel efficiency, (b) the ring pixel purity, (c) the event efficiency and (d) the event purity as defined in the text. The vertical scales represent the percentage of the total distribution per bin, with 50 bins per plot.

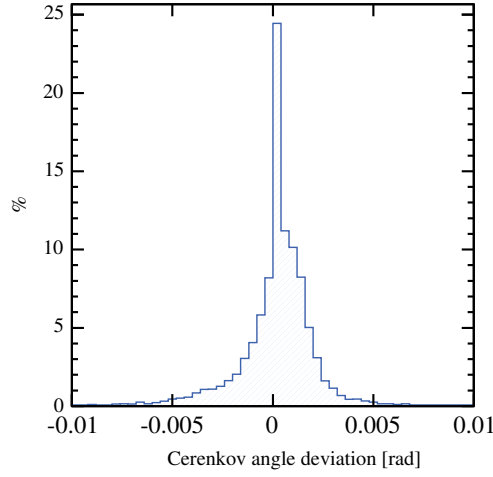


Figure 3.13: Pull distribution of the Čerenkov emission angle in RICH2 due to Markov priors. There is no strong bias away from the MC value. The vertical scale represents the fraction of entries per bin, with 50 bins in the plot.

regions of the detector where rings can be drawn through many essentially equivalent sets of pixels. The indication is that this region of the detector results in qualitatively good but quantitatively bad trackless ring-finding and that the tracking information is required to break the degeneracy of the *a priori* equivalent ring configurations.

Comparing tracked and untracked reconstruction

It is useful to compare the accuracy of ring locations determined by the trackless ring-finder to the standard track-aware procedure. For each Markov ring with a matching track segment, the MC segment is obtained and projected on to the appropriate detector plane. The distance from this point to the corresponding reconstructed track segment projection is shown for 8k $B_d^0 \rightarrow K^+ K^-$ events in Figure 3.14(a), equivalent to the ring-centre error of the standard global particle ID method. Correspondingly, the distance from the Markov ring centre to the MC segment point is plotted in Figure 3.14(b).

Similarly to the distribution obtained for the Čerenkov emission angle bias, the trackless method produces a generally wide distribution extending beyond 30 mm but also has a sharp peak at or near zero. Again, this indicates that in some cases the ring-finder is able to get an almost perfect match but that most of the time it is quantitatively unreliable.

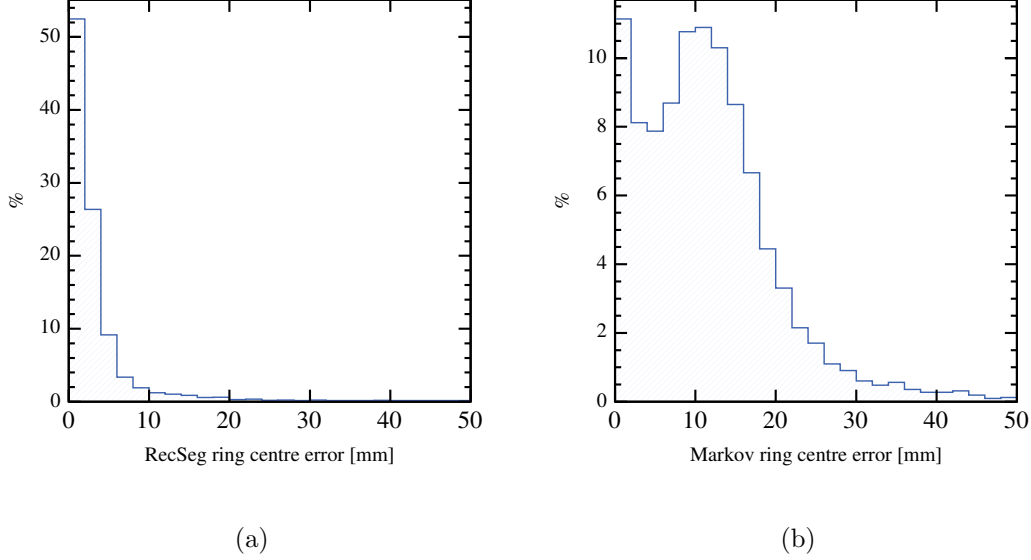


Figure 3.14: Magnitude of the ring-centre position error from (a) projection-matching of reconstructed track-segments and (b) from Markov chain circle-fitting. The vertical scale represents the fraction of entries per bin, with 50 bins per plot.

Processing time

Again running on 8k $B_d^0 \rightarrow K^+ K^-$ signal events, the Gaudi system allows comparison of the processing time devoted to the standard particle ID system and the trackless ring-finder. This indicates that, only considering RICH2, the trackless sampler is ~ 30 times slower than standard particle ID. Combined with the low purities and efficiencies, this means that trackless ring-finding is not a viable method for the mainstream particle ID system. However, the trackless method still has the ability to characterise the untracked rings which cannot be found using the global particle ID procedure and it is to these that we turn our attention in the next section.

3.4 Untracked rings

In certain cases isolated rings are found around the edges of the detector planes with a well-fitted Markov ring but no corresponding reconstructed track segment and hence no standard reconstructed ring. These “untracked” rings are a large part of the motivation

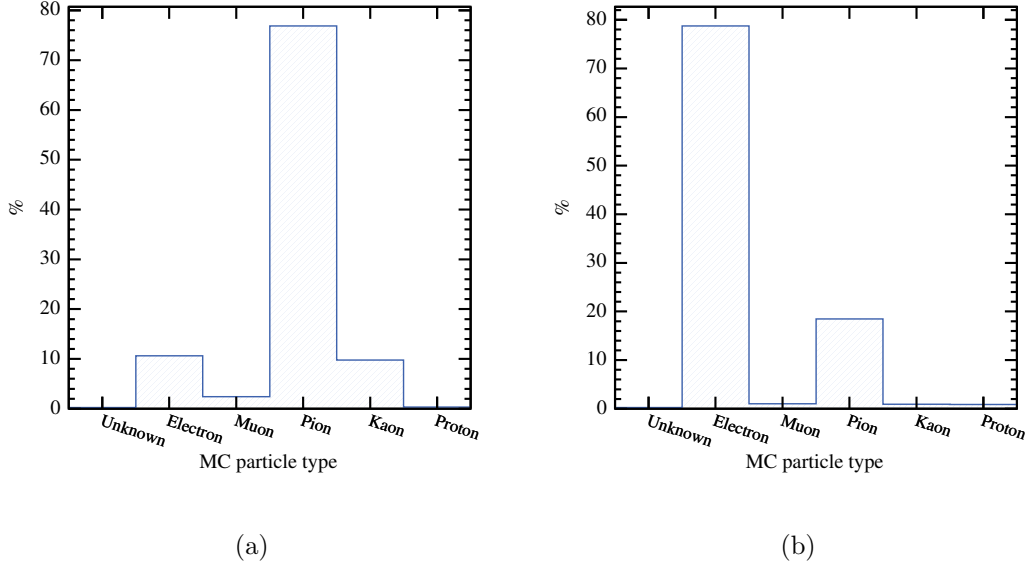


Figure 3.15: MC particle ID fractions for tracks from (a) tracked and (b) untracked Markov rings.

for a trackless ring-finder algorithm and studying their properties is a unique possibility offered by the sampler.

Firstly, we are interested in knowing what particles produce the untracked rings. As usual the MC associator system comes in useful at this point: by using the MC pixel matching method and associating to the majority MC particle we can identify the particle type, the decay chain that the radiating particle lies within and several characteristics of the associated track.

The MC particle IDs for tracked and untracked rings are plotted in Figure 3.15. From this it can be clearly seen that, while pions are the most common particle type for tracked rings, the untracked rings are dominated by electron/positron contributions: this is immediately interesting since LHCb's electron tracking is poorer than for other charged tracks due to the low mass, which increases the importance of secondary scattering and bremsstrahlung processes [141].

An immediate follow-up is to ask from where the MC tracks responsible for producing untracked rings actually originate. This is obtained by making the association to the MC track and then requesting the origin position of the track. These locations are plotted for tracked and untracked rings in Figures 3.16 and 3.17, the first of which is a raw

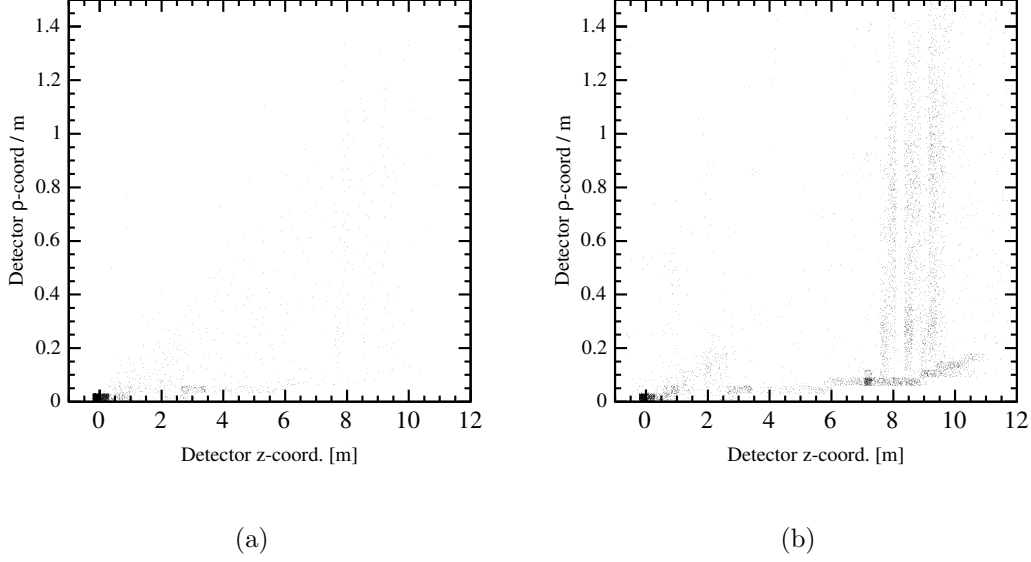


Figure 3.16: Origin vertex positions of MC tracks which produce (a) tracked and (b) untracked RICH rings in ρ - z projection.

scatter-plot in ρ - z co-ordinates ($\rho = \sqrt{x^2 + y^2}$) and the second is a binned histogram purely in terms of number of track origins per z -interval.

The main feature of note in Figure 3.16(b) (the scatter-plot of origin vertices for untracked rings) is that the origin vertex positions trace out the sub-detector structure of LHCb with the beam-pipe (and flanges) and the T1–3 tracking stations being particularly clearly visible. As no such structure occurs for the tracked rings, which Figure 3.17(a) shows are dominantly produced in the VELO and the immediately downstream region, the untracked rings are clearly dominated by secondary interactions. From Figure 3.15 it is apparent that the untracked rings are dominantly produced by electrons: in particular this indicates that the secondary interactions are photon conversions to e^+e^- pairs, with the photons originating from pion decays.

Finally, we wish to understand the origin of the strong origin vertex peak for untracked rings located at the interaction vertex in Figure 3.17(b). The question is how these tracks from the origin which pass through all the tracking stations to reach RICH2 are still not present in the set of tracks passed to the standard RICH particle ID system.

Studying the explicit structure of the particle decay chains for the untracked MC particles whose origins are in the VELO indicates that this peak is not physical but a

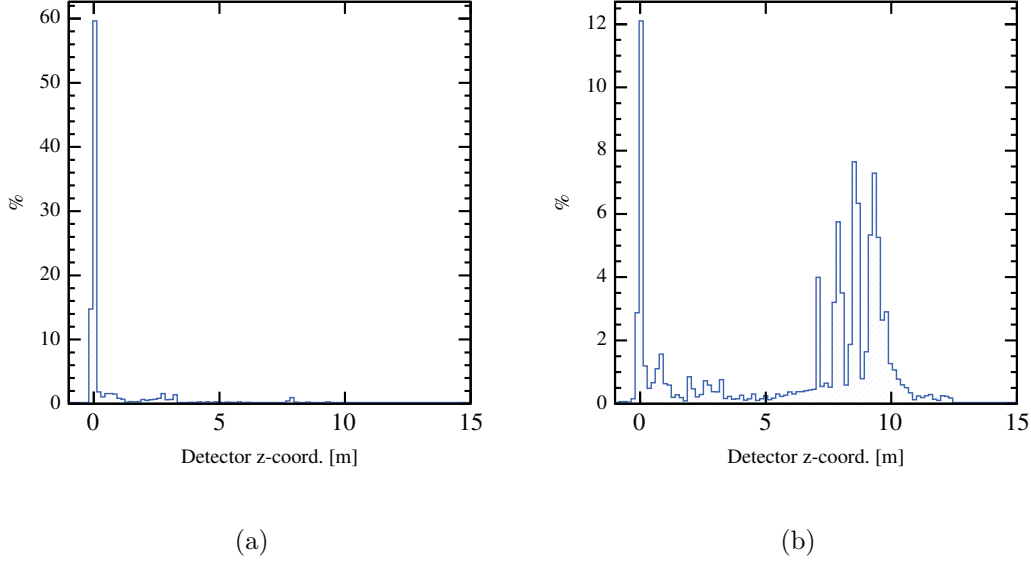


Figure 3.17: Origin vertex z -positions of MC tracks which produce (a) tracked and (b) untracked Markov rings. The vertical scale represents the fractional bin occupancy.

software artifact. Most of the time these particles have *decay* vertices which are also upstream of RICH2. This seemingly paradoxical behaviour is thought to be because decay daughters of these particles directly produce hits by hitting the RICH2 photon detectors (as opposed to the standard excitation by Čerenkov emission). The MC association reports the wrong responsible particle because there is no intermediate Čerenkov photon involved. A fix for this anomalous behaviour will be applied to the MC associator system.

The overall indication of this study is that the real RICH2 untracked rings are produced by electrons and positrons from photon conversions due to the detector material. These conversions primarily take place in the beam pipe and the ‘T’ tracking stations. Although it would be useful to identify the cause of these rings to improve the overall particle ID, they are not of interest except as a source of background to physically interesting tracks and the global particle ID system already treats them as such. The final aspect of this study is to determine whether using the trackless ring-finder to identify pixels from untracked rings as background can lead to better particle ID performance than the standard background estimator algorithm.

3.5 Impact of trackless ring-finding on particle ID

An interesting test of the potential of the trackless ring-finder algorithm is to use it for identifying pixels which come from background rather than from Čerenkov rings and to use this information to “clean” the detector of these background pixels before running the standard particle ID algorithms.

As noted in Section 3.1, the standard particle ID is made up of first a local algorithm to seed rings, then two runs of a background estimator algorithm and the global ring-finder. The pixel background weight factor set by the background estimator indicates how likely a pixel is to be a background hit based on track projections — the idea of this section is to investigate using the trackless ring-finder as a more intelligent substitute for the background estimator. In this we have investigated two approaches: explicitly removing pixels identified as background from the detector before running the global algorithm and, less drastically, using the Markov chain finder to apply the pixel weight factors.

3.5.1 Stripping background pixels

For the stripping approach, we still run the standard background estimator but first the Markov ring-finder is run and each pixel is identified with a set of Markov rings by probabilistic matching. We then have the choice of what to consider as background: do we just consider pixels which are not associated to any Markov ring or do we also regard those pixels which are only bound to rings with no associated reconstructed track segment as background? The first definition is the more “correct”, in that it recognises the existence of untracked rings, while the second mimics more closely the behaviour of the standard (track-based) background estimator. We will use the second definition, to emulate the standard background estimator.

The best possible particle ID efficiencies obtainable with a background stripping tool can be obtained by using an MC-aware “pixel maker” which only creates pixel objects which have good MC associations to real MC particles. Again only using RICH2 for particle ID studies, a “perfect” background stripper was used, followed by the usual two-times sequence of the local particle ID algorithm, background estimator and global particle ID algorithm. Using 8k $B_d^0 \rightarrow K^+ K^-$ events over all momenta, this leads to kaon and pion particle ID efficiencies of $\varepsilon(K) = (98.06 \pm 0.17)\%$ and

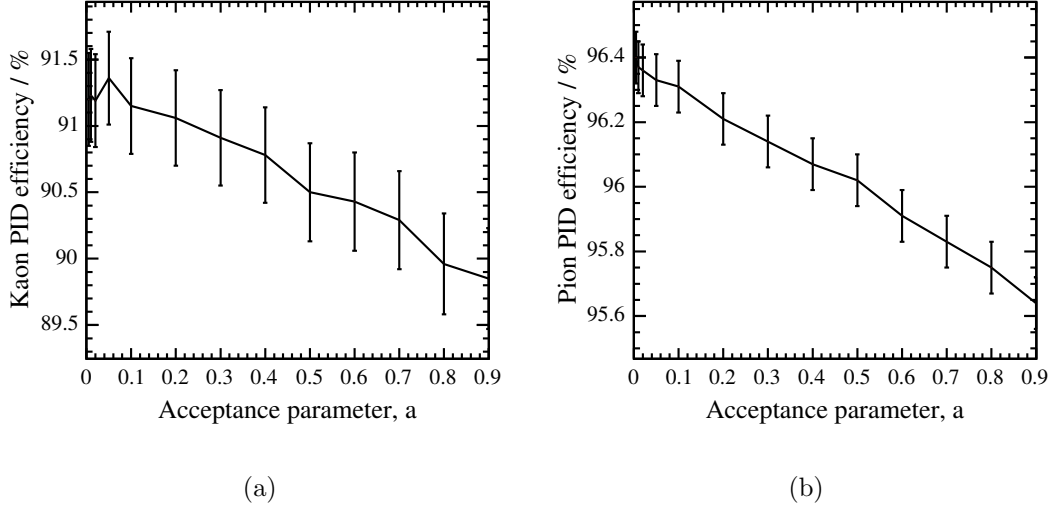


Figure 3.18: Particle ID efficiencies as a function of the acceptance parameter a_{prob} for (a) kaons and (b) pions.

$\varepsilon(\pi) = (98.22 \pm 0.05)\%$, compared with the standard particle ID system efficiencies in RICH2: $\varepsilon(K) = (92.01 \pm 0.32)\%$ and $\varepsilon(\pi) = (96.71 \pm 0.07)\%$. This indicates that there is the potential for background stripping to improve particle ID performance in RICH2.

An implementation of the background stripper using the trackless ring-finder has been written. This performs the usual ring-finding and considers Markov rings without a projection-matched track segment to be untracked. The usual process to match pixels to rings is then performed using the probabilistic matching procedure and any pixels which are not associated to a tracked ring are removed from the system. The standard particle ID system is then executed. The only parameter introduced in the process is that for the acceptance parameter a_{prob} which determines the fraction of the best-match pixel-ring probability for each pixel that a pixel must have to be considered matched to the ring. The variations of kaon and pion particle ID efficiency with a_{prob} between 0.005 and 1 are shown in Figure 3.18.

In all cases the stripping produces lower particle ID efficiencies than those from the standard RICH2 particle ID and low values of a_{prob} are favoured: for $a_{\text{prob}} = 0$ all pixels are matched to all rings and no stripping will take place. The indication is that stripping background pixels does not improve the particle ID, probably again due to mis-identification of rings in the high-occupancy regions of the detector which result in “good” hits being removed.

3.5.2 Weighting background pixels

The second approach to background removal is to replace the standard background estimator with a more “ring-aware” system based on the trackless ring finder. A conversion is needed internally to convert the sampler probability of a pixel not being associated to any (tracked) rings into a background likelihood contribution which is used as the background weight-factor by the global particle ID. The conversion relation is to multiply the background probability by the effective fraction of HPD pixels which are expected to be background, $\bar{n}_{\text{bgpix}}/n_{\text{pix}}$. The value of n_{pix} used by the standard background estimator is $n_{\text{pix}} \sim 785$, based on the fraction of the 1024 pixels on to which the circular HPD window projects and the expected number of background pixels, $\bar{n}_{\text{bgpix}} \sim \mathcal{O}(10)$ [140].

The background contribution distributions from the standard background estimator and from the trackless ring-finder system are shown in Figure 3.19, with $\bar{n}_{\text{bgpix}} = 7.3$ ² in Figure 3.19(b). The difference in distribution shapes is evident, with the trackless system again showing a sharp peak for pixels which it “believes” to be almost certainly from background or untracked rings.

3.5.3 Simulating extra background tracks

In addition to the a_{prob} acceptance parameter and the \bar{n}_{bgpix} parameter already introduced, it is of interest to study the robustness of both the standard particle ID system and the system with Markov-enhanced background estimation to extra ring-like background. This is achieved by storing the MC tracks which produce pixels in RICH2 from chosen events and adding the pixels for each track to the remaining events. If the reconstruction system is run with a requirement to add n_b sets of background pixels to each studied event then, on “visiting” an event and finding that not enough such pixels have yet been stored, the system will take all the pixels from that event (indexed by the responsible MC particle) and store them. The event is then aborted so that it does not contribute to the particle ID efficiency count. On visiting an event and finding more than n_b tracks’ worth of pixels in the background cache, n_b sets of pixels will be added to the event without the accompanying tracks.

RICH2 has a mean ~ 20 tracks per event [140], so every ~ 20 tracks’ worth of untracked pixels corresponds to an entire untracked event spilling over. There is no current LHCb

²The non-integer values of \bar{n}_{bgpix} in this section are a result of an alternative approach taken to deriving the conversion relation.

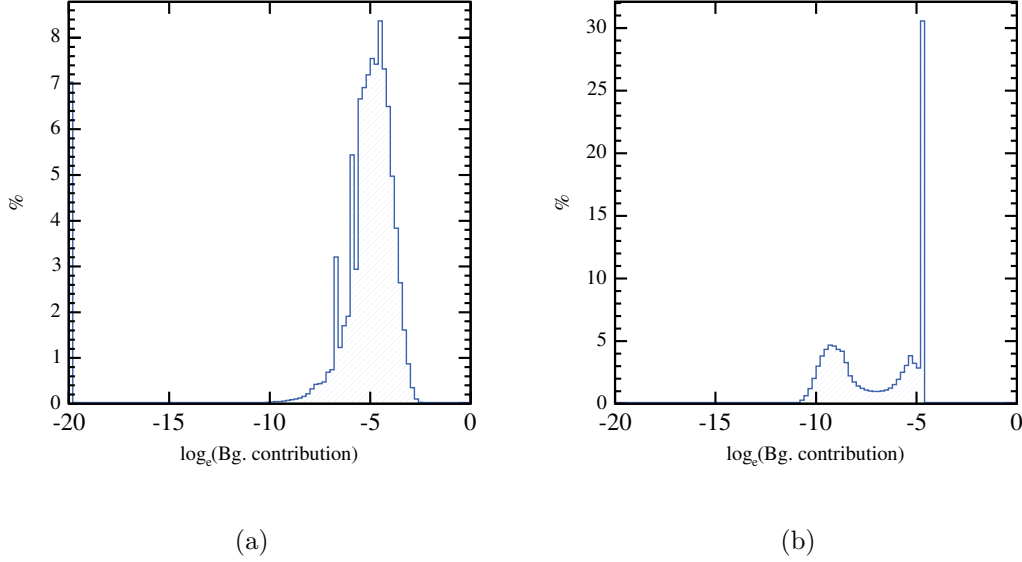


Figure 3.19: Pixel background contribution distributions ($p(bg) \times \bar{n}_{\text{bgpix}}/n_{\text{pix}}$, as described in the text) for (a) the standard background estimator and (b) the background estimator based on the trackless ring-finder with $\bar{n}_{\text{bgpix}} = 7.3$. The vertical scale represents the fraction bin occupancy.

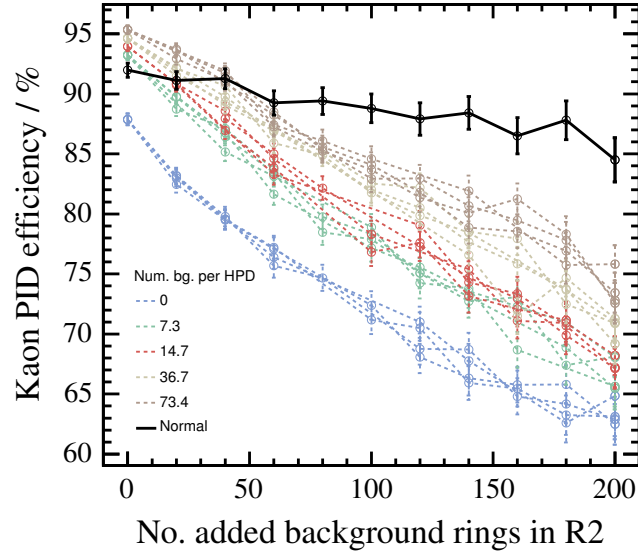
estimate of what constitutes a “reasonable” amount of excess background so we study the effect of extra backgrounds up to 200 tracks, with the region up to 50 tracks being considered most realistic.

3.5.4 Performance of the Markov background estimator

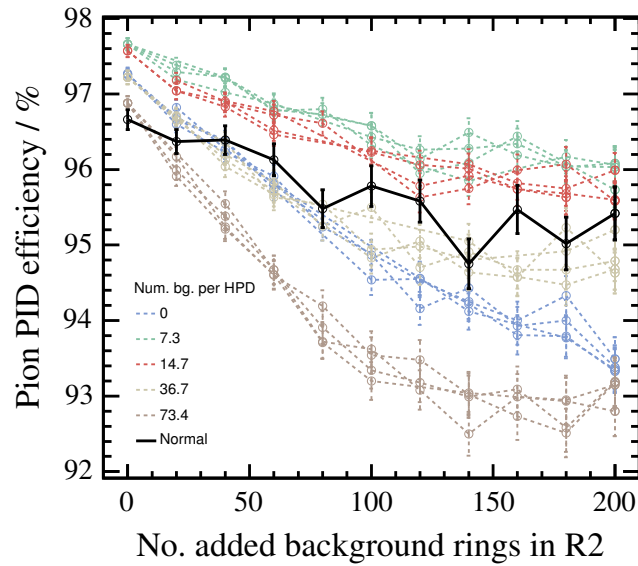
A set of parameter scans, using 8k $B_d^0 \rightarrow K^+ K^-$ events, have been performed on the variables $a_{\text{prob}} \in \{0.005, 0.01, 0.02, 0.05\}$, $\bar{n}_{\text{bgpix}} \in [10^{-3}, 74]$ and extra pixel backgrounds corresponding to 0–200 tracks. The particle ID efficiencies for RICH2 from this initial scan are shown in Figure 3.20.

The general trend is that a_{prob} , which produces the spread within the coloured bunches of lines, is relatively unimportant in determining the overall particle ID efficiencies. Studies of the particle ID performance as a function of a_{prob} indicate that there is no strong preference for any of the values of a_{prob} in the given range, so from now on only $a_{\text{prob}} = 0.01$ will be used.

The major impact is made by varying \bar{n}_{bgpix} , which is used to index the different



(a)



(b)

Figure 3.20: RICH2 particle ID efficiencies as a function of number of added background rings for (a) kaons and (b) pions. The different colour groupings indicate the choice of \bar{n}_{bgpix} as described in the text and the variations within the colour groups correspond to variation of the a_{prob} parameter.

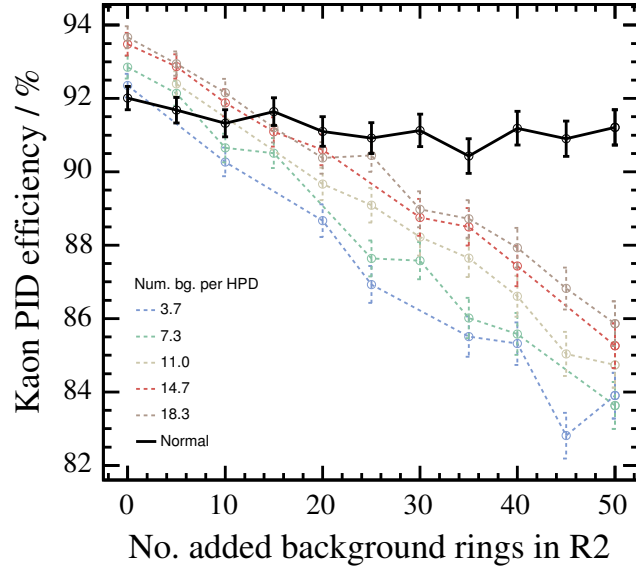
groups of coloured lines in Figure 3.20. The best performance for kaons is given by the largest values of \bar{n}_{bgpix} , but this corresponds to the lowest performance in the pion ID efficiency. The best performance trade-off is seen for $\bar{n}_{\text{bgpix}} \in [7, 15]$, with the precise value depending on the chosen measure of overall performance. Both these values of \bar{n}_{bgpix} give kaon performances which exceed that of the standard background estimator for backgrounds of up to ~ 20 tracks and which surpass the pion ID efficiencies by $\sim 1.5\%$ for all levels of extra background. It should be noted that for higher background levels, the sampler priors are not optimised and so the trackless finder performance is underestimated. Secondly, as higher levels of background require the addition of more tracks, fewer events contribute to the particle ID efficiencies for high values of n_b and the uncertainties are correspondingly larger.

Using the first parameter scan as a starting point, a second scan over a tighter set of parameters has been performed for $\bar{n}_{\text{bgpix}} \in [3, 19]$ and for backgrounds of up to 50 tracks with a_{prob} set to 0.01. The kaon and pion ID efficiencies are shown in Figure 3.21, again showing that the best performance is obtained for values of \bar{n}_{bgpix} between 7 and 15. It is desirable that future studies should attempt to predict \bar{n}_{bgpix} rather than parameter scanning across a range of possible values, but it is also reassuring that the best performance mixtures are obtained for a reasonable set of values of \bar{n}_{bgpix} .

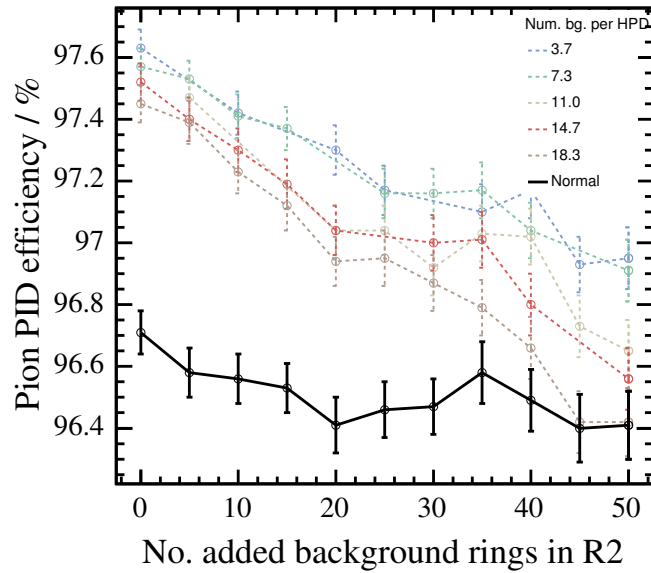
For the configuration $a = 0.01$ and $\bar{n}_{\text{bgpix}} = 7.3$, Figure 3.22 shows the light and heavy particle ID efficiencies and mis-ID rates as functions of momentum relative to the standard particle ID system (also running only on RICH2) with no added background. The results show improvements from trackless background weighting in the middle and high momentum regions but the standard background estimator still performs better in the low momentum region near the Čerenkov emission threshold. The fall in performance at low momenta is because of the low detector occupancy in this region: wrongly weighting pixels as background has a larger adverse effect when there are fewer pixels available for that track. In any case, the fall-off at low momentum is not important as particle ID in this region is mainly supplied by RICH1.

3.6 Conclusions

In this chapter we have presented an alternative approach to LHCb RICH particle ID which uses a Markov Chain sampler to find Čerenkov rings without using tracking information. Studies of the sampler indicate that it performs well on sparse pixel arrays



(a)



(b)

Figure 3.21: RICH2 particle ID efficiencies as a function of number of added background rings for (a) kaons and (b) pions, with a tighter range of parameters than in Figure 3.20. The different colour groupings indicate the choice of \bar{n}_{bgpix} as described in the text and the variations within the colour groups correspond to variation of the a_{prob} parameter.

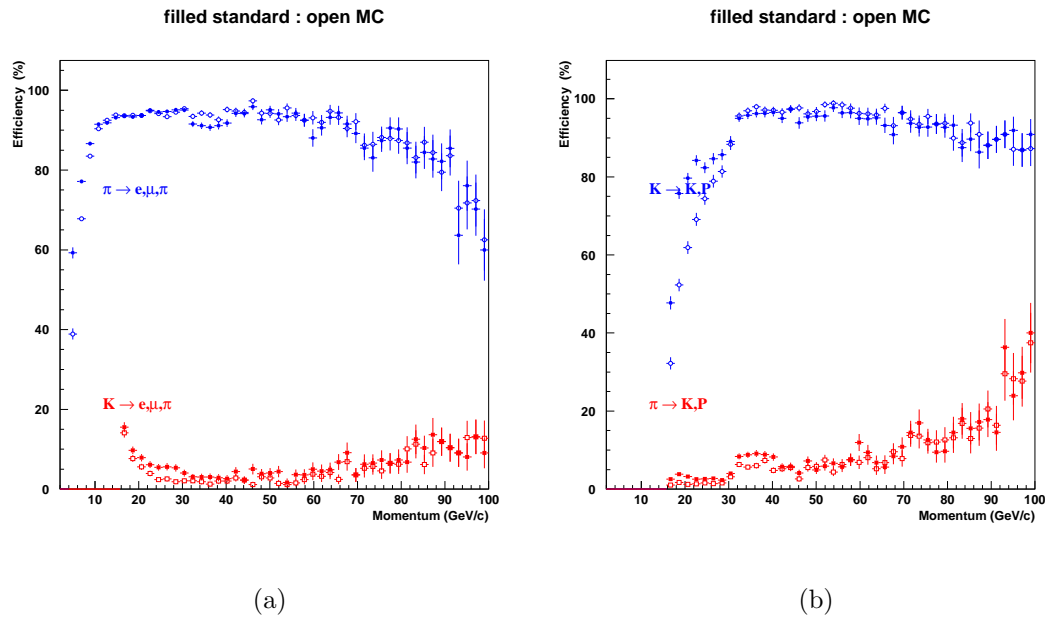


Figure 3.22: Particle ID efficiency and mis-ID rate comparisons between the standard and Markov background estimation systems. The filled points represent the standard system and the open points are from the Markov-based system with (a) showing the “light” ID performance (identifying particles as e , μ or π) and (b) shows the performance for the “heavy” (K , p) ID.

but that it breaks down in high-occupancy environments such as the central regions of the LHCb RICH detectors. In such regions, the tracking information again becomes essential for reliable particle ID. This tool has also been used to study the origins of real Čerenkov rings which are not found by the standard (track-based) particle ID system. These were found to be dominated by photon conversions due to interactions with the detector material, but are not significant for physics purposes.

The final section of this chapter studied the behaviour of the LHCb RICH particle ID system when the trackless-ring finder is integrated into the RICH background estimator algorithms. In this case it was found that the trackless ring-finder can improve the kaon and pion ID efficiencies by 1–2% when used to determine background weightings for the pixels. Robustness studies were performed by adding RICH events as background: these showed good robustness of the standard RICH particle ID system and of the pion efficiency with the trackless background estimator, but the kaon efficiency is adversely affected by the trackless weighting system for high background levels.

Studying the behaviour and application of the trackless ring-finding approach is a seemingly bottomless research topic: there are so many possible variations that many must by necessity go without investigation. Despite the poor performance of trackless ring-finding on whole RICH events, the ability to make good fits to isolated rings on the outer edges of the detector planes means that there is still value to be found in future trackless ring-finding studies.

The primary extension of the work presented here will be to explicitly restrict the pattern recognition to regions of the detector planes with low hit densities. This would increase the speed of the sampler, having a much smaller phase space from which to sample, and would focus on the rings which the sampler is good at finding. Alternatively, the ring-finding sampler could be applied only to hits which the standard tracking has identified as having high background weights: this would again reduce the phase space and would focus the sampler on studying untracked rings. If approaches of this sort can be made fast enough and can identify untracked rings reliably then this approach may be useful as part of the standard offline analysis system.

An additional use of the trackless ring-finding method, which has not yet been investigated in sufficient depth to report here in detail, is that of using the ring-finder's high-quality matches to untracked rings to improve the tracking of such rings. While the studies of untracked rings presented here in Section 3.4 confirm that the majority of RICH2 untracked rings are from secondary interactions of tracks with the detector ele-

ments, making the tracking aware of such tracks is certainly useful. Preliminary studies of this approach with a reversed RICH2 optics ray-tracing algorithm have indicated that the distribution of projected track segment impact vectors on to the detector planes is too broad for useful tracking when all rings are considered. However, this distribution could be narrowed and used for reconstruction by parameterising the impact vector in terms of the position on the detector plane and by restricting interest to the untracked rings only. There is plenty of scope for more work to be done on trackless ring-finding in LHCb.

Chapter 4

Phenomenology of $B \rightarrow K \pi$ decays

*“Beauty is truth, truth beauty, ’—that is all
Ye know on earth, and all ye need to know.”*

— John Keats, 1795–1821

$B \rightarrow K \pi$ decays are of interest in b-physics primarily because they can be used in strategies to bound or determine the CKM weak decay phase γ . They are also interesting from both QCD and new physics perspectives since the charmless mesons in the final state are susceptible to non-perturbative QCD effects and the weak component of the decay amplitude is sensitive to contributions from potential new physics. In this chapter we review the status of measurements and methods to obtain constraints on γ and concentrate in particular on one strategy which uses an isospin argument to relate observables in $B_d^0 \rightarrow K^+ \pi^-$ and $B^+ \rightarrow K_S^0 \pi^+$ decays.

4.1 Further phenomenology of weak non-leptonic B decays

The phenomenology of weak B decays has already been introduced in Chapter 1, where we introduced tree and penguin Feynman diagram topologies. In addition to this, we require robust modelling of hadronic processes to describe the phenomenology of $B \rightarrow K \pi$ decays. As QCD is largely non-perturbative at the energy scales involved in non-leptonic B-decays, we must introduce models to describe the hadronic processes. The starting

point for defining such models is the formulation of *effective theories*.

4.1.1 Low energy effective theories

In practice, non-leptonic B-decays are not analysed from a first-principles expansion of Feynman diagrams to arbitrary order but via the formalism of constructing a low energy effective Hamiltonian which encapsulates the required physics [142, 143]. The operator product expansion is used to construct transition matrix elements of the following form:

$$\langle f | \mathcal{H}_{\text{eff}} | i \rangle = \frac{G_F}{\sqrt{2}} \lambda_{\text{CKM}} \sum_k C_k(\mu) \langle f | Q_k(\mu) | i \rangle, \quad (4.1)$$

where μ is an appropriate renormalisation scale $\sim \mathcal{O}(m_b)$ and λ_{CKM} is a general CKM factor [142].

The rôle of this expansion has been to separate the short and long-distance contributions to the transition amplitude. The short distance physics, corresponding to the electroweak scale effects, has been absorbed into the Fermi constant $G_F/\sqrt{2}$ by integrating out the weak boson propagators so that $G_F \sim \mathcal{O}(g_2^2/M_W^2)$ where g_2 is the weak coupling constant. The long-range physics is itself split into perturbative Wilson coefficients $C_k(\mu)$ and non-perturbative hadronic matrix operators $Q_k(\mu)$. The Wilson coefficients are calculated by evaluating QCD corrections to the leading order diagrams in both the effective and the full theory and then matching the two with the assistance of the renormalisation group formalism which relies on the physical insignificance of the renormalisation scale [144].

The conventional assignment of Q_k operators is that Q_1 and Q_2 represent tree-topology QCD and electroweak processes, composed of a pair of colour singlet currents. For a $b \rightarrow c \bar{u} d$ transition,

$$Q_1 = (\bar{d} u)_{V-A} (\bar{c} b)_{V-A}, \quad (4.2a)$$

$$Q_2 = (\bar{c} d)_{V-A} (\bar{u} b)_{V-A} \quad (4.2b)$$

where

$$(\bar{d} u)_{V-A} \equiv \bar{d} \frac{\gamma_\mu (1 - \gamma_5)}{2} u. \quad (4.3)$$

The remaining Q_k are QCD penguin operators for $k = 3, \dots, 6$ and electroweak penguin operators for $k = 7, \dots, 10$. Significant electroweak penguin effects may be observed in certain channels because the Wilson coefficients are enhanced when evaluated at a renormalisation scale $\sim \mathcal{O}(m_t)$ as appropriate for internal top quark exchange.

The formalism described so far has provided a set of four-quark local operators effectively describing the electroweak and QCD effects at a given scale. However, the calculation of the non-perturbative hadronic matrix elements is fraught with theoretical problems and another approximation is required to proceed. The basic approach is that of “naïve factorisation”, which has a long and controversial history.

4.1.2 Matrix element factorisation

The key idea of naïve factorisation is that transition matrix elements containing the four quark operators already described may be explicitly split into a pair of independent amplitudes at first order. For the tree decay $\bar{B}_d^0 \rightarrow D^+ \pi^-$, diagrams with gluons connecting the D and π final states are anticipated to be a small correction since the B is heavy. The tree diagram is then “cut” to give the factorised amplitude [143],

$$\langle D^+ \pi^- | (\bar{d}u)_{V-A} (\bar{c}b)_{V-A} | \bar{B}_d^0 \rangle_{\text{fact}} = \langle \pi^- | (\bar{d}u)_A | 0 \rangle \langle D^+ | (\bar{c}b)_V | \bar{B}_d^0 \rangle \quad (4.4a)$$

$$= f_\pi (\text{“decay constant”}) \times F_{B \rightarrow D} (\text{“form factor”}) \quad (4.4b)$$

where the last line shows the conventional representation of the factorised matrix element as the product of the pion “decay constant” and a $\bar{B}_d^0 \rightarrow D^+$ “form factor”. The reliability of this approach, which is based on the algebraic structure of large- N $SU(N)$ group theory, is much debated and varies according to the decay being described [143, 144]. More recently an extended approach, “QCD factorisation” has been developed which incorporates both aspects of the naïve factorisation formalism and the hard-scattering picture [145, 146, 147].

Let us consider the QCD factorisation approach to two-body non-leptonic B -decays, $\bar{B} \rightarrow M_1 M_2$ where M_1 picks up the spectator quark. If M_1 is either a heavy (D) or a light (π , K) meson and M_2 is a light meson then the transition amplitude is of the form [144]

$$\mathcal{A}(\bar{B} \rightarrow M_1 M_2) = [\text{“naïve factorisation”}] \times [1 + \mathcal{O}(\alpha_s) + \mathcal{O}(\Lambda_{\text{QCD}}/m_b)]. \quad (4.5)$$

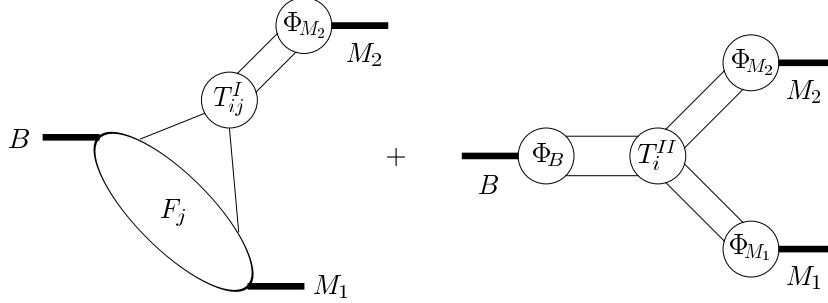


Figure 4.1: Schematic representation of the QCD factorisation formula in equation (4.6).

Here the $\mathcal{O}(\alpha_s)$ term represents the correction to the naïve factorisation amplitude by systematically calculating the non-factorisable QCD radiative contributions and the $\mathcal{O}(\Lambda_{\text{QCD}}/m_b)$ terms represent non-perturbative QCD effects. The latter set of corrections are the main source of theoretical limitation for the QCD factorisation formalism. The QCD factorisation amplitude for each local operator Q_i is written as

$$\langle M_1 M_2 | Q_i | \bar{B} \rangle = \sum_{\substack{\{M_a, M_b\} \\ \in \{M_1, M_2, \}}} F_0^{B \rightarrow M_a} T_{M_b, i}^I * f_{M_b} \Phi_{M_b} + T_i^{II} * f_B \Phi_B * f_{M_1} \Phi_{M_1} * f_{M_2} \Phi_{M_2} \quad (4.6)$$

where the F terms again represent transition form factors, the T terms are hard-scattering kernels, f_i are the decay constants for the i meson, the Φ_i are the corresponding light cone distribution amplitudes and the star-products indicate integrations across the light-cone momentum fractions of the quarks inside the mesons.

We do not need to investigate further the technical details of this formalism; equation (4.6) is schematically illustrated in Figure 4.1, where only one of the form-factor summation terms is shown for simplicity. The summation in the first term of equation (4.6) is only present in general to account for cases where the spectator quark can end up in either final state meson [148]. The QCD factorisation formalism is the main method in use at present for describing the hadronic decays of B-mesons.

4.2 CKM phenomenology from $B \rightarrow K \pi$ decays

The standard direct methods for measuring the γ CKM phase have already been summarised in Chapter 1. The focus of this chapter and the study which occupies the remainder of this thesis is on an alternative method to constrain γ , based on the SU(2) isospin symmetry and CKM unitarity.

4.2.1 Amplitudes and observables for $B \rightarrow K \pi$ channels

In 1997, the first CLEO measurements of $B \rightarrow K \pi$ decays [149, 150] re-invigorated the suggestion that such decay channels have potential for performing \mathcal{CP} asymmetry measurements [151, 152]. These decays are of interest not just because of their potential for CKM constraints but because they are significantly affected by electroweak penguins. These provide a window to beyond Standard Model (BSM) physics via heavy particles propagating in the penguin loops. In current measurements of $B \rightarrow K \pi$ and $B \rightarrow \pi \pi$ decays, discrepancies have been observed between certain observables and the Standard Model theory predictions for channels like $B^\pm \rightarrow K^\pm \pi^0$ [153, 154, 155, 156] — this has been called “the $B \rightarrow K \pi$ puzzle” by some authors [157]. $B^\pm \rightarrow K^\pm \pi^0$ decays admit colour-allowed electroweak penguin contributions, while decay modes such as $B^\pm \rightarrow K^0 \pi^\pm$ and $B_d^0 \rightarrow K^\pm \pi^\mp$ (to which only *colour-suppressed* electroweak penguins contribute) remain compatible with the Standard Model description. As a result, this region of b-physics may be tentatively described as indicating new physics in electroweak penguins, and attention has been focused on it accordingly.

The QCD isospin symmetry can be used to obtain a value for the CKM γ angle based on the relative branching fractions to charged and uncharged combinations of decay products. Three such strategies have been suggested: a “mixed” strategy using $B^\pm \rightarrow K^0 \pi^\pm$ and $B_d^0 \rightarrow K^\pm \pi^\mp$ decay rates, a “charged” strategy using $B^\pm \rightarrow K^0 \pi^\pm$ and $B^\pm \rightarrow K^\pm \pi^0$ decays, and a “neutral” strategy with $B_d^0 \rightarrow K_S^0 \pi^0$ and $B_d^0 \rightarrow K^\pm \pi^\mp$ decays. Of these the mixed strategy is the most experimentally accessible to LHCb, not being burdened with the difficulties of reconstructing a neutral pion in the final state. The other two approaches have theoretical advantages with respect to final state re-scattering effects. The main emphasis in this chapter will be on the mixed strategy.

For reference, the current set of measurements of \mathcal{CP} -asymmetries in charmless two-body B-decays as fitted by the CKM Fitter group is shown in Figure 4.2 with compar-

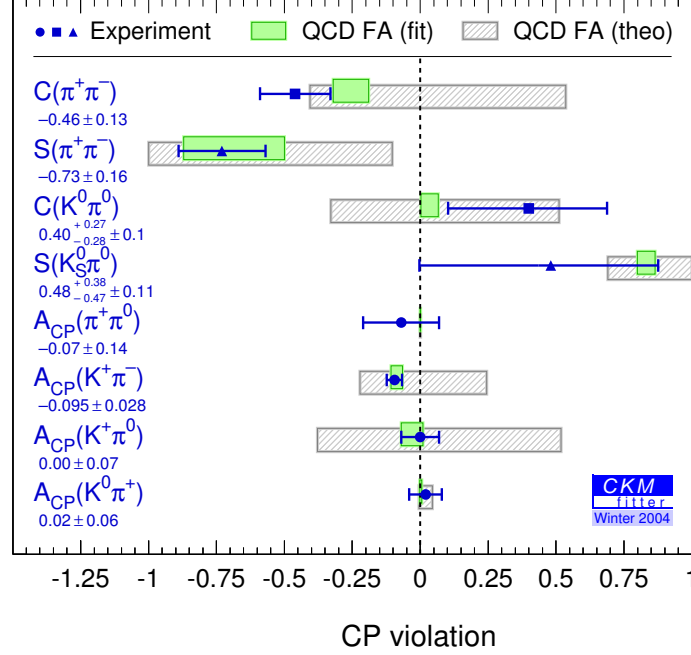


Figure 4.2: Fits to $\pi\pi$ and $K\pi$ \mathcal{CP} asymmetries by the CKM Fitter group using the QCD factorisation approximation (QCD FA). The marked experimental results are all world-averages taken from combining *BABAR*, *Belle* and *CLEO* results: the details can be found in reference [82].

isons to the QCD factorisation predictions. The $B_d^0 \rightarrow K^+ \pi^-$ channel has been recently measured to have a large direct \mathcal{CP} -violation while $B^+ \rightarrow K_S^0 \pi^+$ does not show any significant \mathcal{CP} -violation.

4.2.2 The “mixed” $B^\pm \rightarrow K^0 \pi^\pm$, $B_d^0 \rightarrow K^\pm \pi^\mp$ strategy

In the Standard Model, the decays $B^+ \rightarrow K^0 \pi^+$ and $B_d^0 \rightarrow K^+ \pi^-$ are mediated by dominant Feynman diagram contributions of the type indicated in Figures 4.4 and 4.3 respectively. $B^+ \rightarrow K^0 \pi^+$ contains leading-order contributions from penguin topologies and annihilation diagrams as shown in Figure 4.4(b) and $B_d^0 \rightarrow K^+ \pi^-$ decays are driven by penguin topologies and the tree diagram shown in Figure 4.3(b).

From the charged-current weak vertices in Figures 4.4 and 4.3 we see that the tree and annihilation diagrams have CKM contributions of $|V_{us}V_{ub}^*|$ while the penguin diagrams have a CKM contribution $\sim |V_{ts}V_{tb}^*|$, assuming the amplitude to be dominated by the contribution from the internal top quark in the penguin loop. As the ra-

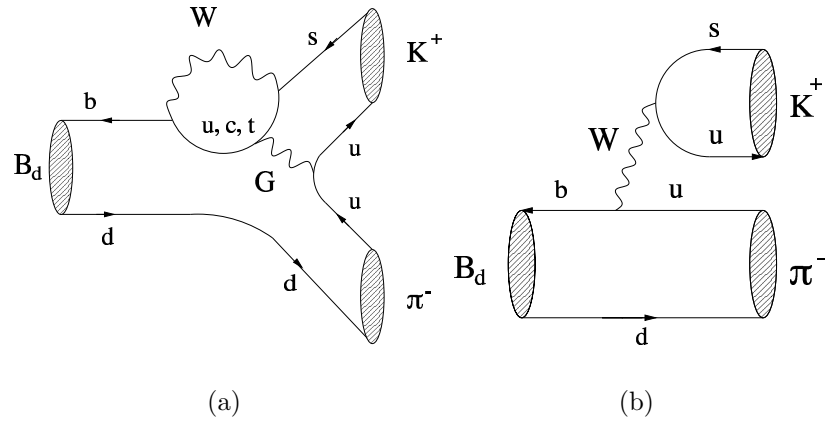


Figure 4.3: Feynman diagram topologies contributing to $B_d^0 \rightarrow K^+ \pi^-$ decays: (a) QCD penguin diagram (b) tree diagram. Colour-suppressed electroweak penguin diagrams have been omitted.

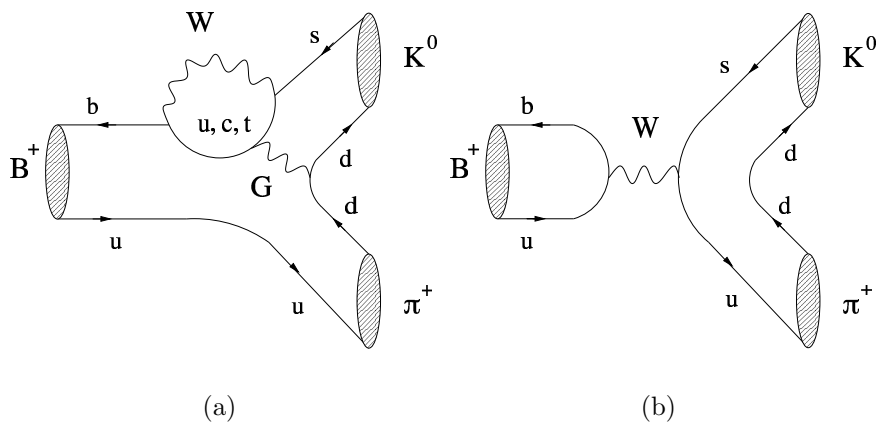


Figure 4.4: Feynman diagram topologies contributing to $B^+ \rightarrow K^0 \pi^+$ decays: (a) QCD penguin diagram (b) annihilation diagram. Colour-suppressed electroweak penguin diagrams have been omitted.

tio $|V_{us}V_{ub}^*|/|V_{ts}V_{tb}^*| \approx 0.02$ these decays are dominated by QCD penguin amplitudes even though they are loop-suppressed. Colour-suppressed electroweak penguins are also present, although they are restricted to a level assumed less than the tree graphs.

$B \rightarrow K \pi$ decay amplitudes

The basic quantities of the mixed $B \rightarrow K \pi$ decays into which γ and associated phases enter are the decay amplitudes. These incorporate the CKM phases, strong phases, re-scattering phases and contributions from various decay topologies. Of the latter contributions, we identify the previously-mentioned QCD penguin, tree, annihilation and EW penguin topologies. The $B^+ \rightarrow K^0 \pi^+$ amplitude is then

$$\mathcal{A}(B^+ \rightarrow K^0 \pi^+) = \lambda_u^{(s)}(\mathcal{P}^u + \mathcal{P}_{\text{ew}}^u + \mathcal{A}) + \lambda_c^{(s)}(\mathcal{P}^c + \mathcal{P}_{\text{ew}}^c) + \lambda_t^{(s)}(\mathcal{P}^t + \mathcal{P}_{\text{ew}}^t), \quad (4.7)$$

where \mathcal{P}^i and $\mathcal{P}_{\text{ew}}^i$ respectively represent QCD and electroweak penguin amplitudes with internal i -type quarks and \mathcal{A} is the annihilation amplitude which only contributes to the u term. The term

$$\lambda_q^{(s)} = V_{qs} V_{qb}^* \quad (4.8)$$

is the contributing combination of CKM factors. The corresponding amplitude for the $B_d^0 \rightarrow K^+ \pi^-$ decay is

$$\mathcal{A}(B_d^0 \rightarrow K^+ \pi^-) = -\left[\lambda_u^{(s)}(\tilde{\mathcal{P}}^u + \tilde{\mathcal{P}}_{\text{ew}}^u + \tilde{\mathcal{T}}) + \lambda_c^{(s)}(\tilde{\mathcal{P}}^c + \tilde{\mathcal{P}}_{\text{ew}}^c) + \lambda_t^{(s)}(\tilde{\mathcal{P}}^t + \tilde{\mathcal{P}}_{\text{ew}}^t)\right], \quad (4.9)$$

with notation as the extension of that used in equation (4.7) and $\tilde{\mathcal{T}}$ representing the contribution of tree diagrams to the u term. The overall minus sign arises from the definitions of B-meson particle states as those containing a \bar{b} anti-quark.

From CKM unitarity and application of the Wolfenstein CKM parameterisation from equation (1.37) we can express these amplitudes more explicitly. For $B^+ \rightarrow K^0 \pi^+$,

$$\mathcal{A}(B^+ \rightarrow K^0 \pi^+) = -\lambda^2 A \left(1 - \frac{\lambda^2}{2}\right) \left[1 + \rho_t e^{i\theta} e^{i\gamma}\right] \mathcal{P}_{\text{tc}} \quad (4.10)$$

where the set of parameters \mathcal{P}_{ij} represent the differences between penguin topologies

with different $i \rightarrow j$ internal quark transitions:

$$\mathcal{P}_{ij} \equiv |\mathcal{P}_{ij}|e^{i\delta_{ij}} = (\mathcal{P}^i - \mathcal{P}^j) + (\mathcal{P}_{\text{ew}}^i - \mathcal{P}_{\text{ew}}^j) \quad (4.11)$$

and λ and A are the Wolfenstein CKM parameters. The phenomenological term $\rho_r e^{i\theta}$ is defined as

$$\rho_r e^{i\theta} \equiv \frac{\lambda^2 R_b}{1 - \lambda^2/2} \left[1 - \frac{\mathcal{P}_{\text{uc}} + \mathcal{A}}{\mathcal{P}_{\text{tc}}} \right] \quad (4.12)$$

and encapsulates the long-range physics of final state re-scattering. Its significance will be returned to in Section 4.3.1.

The amplitudes derived so far have been expressed in terms of the Wolfenstein λ parameter, and the parameters A and R_b . These latter two are defined as

$$A \equiv \frac{1}{\lambda^2} |V_{\text{cb}}| = 0.853 \pm 0.031, \quad R_b \equiv \frac{1}{\lambda} \left| \frac{V_{\text{ub}}}{V_{\text{cb}}} \right| = 0.404 \pm 0.093. \quad (4.13)$$

where the values are taken from the Particle Data Group's 2004 compilation [10].

The $\rho_r e^{i\theta}$ term, through which the \mathcal{CP} -conserving strong phase θ enters equation (4.12), is related to final state re-scattering and is CKM-suppressed by $\lambda^2 R_b \approx 0.02$. Note that ρ_r is a conventional nomenclature for expressing the strength of re-scattering amplitudes and that the quantity is not related in any simple way to the Wolfenstein ρ parameter.

Using the $SU(2)$ isospin symmetry of the strong interactions we can relate QCD penguin diagrams under exchange of u and d quarks: $\mathcal{P}^c = \tilde{\mathcal{P}}^c$ and $\mathcal{P}^t = \tilde{\mathcal{P}}^t$. The \mathcal{P}^u and $\tilde{\mathcal{P}}^u$ terms cannot be included in this procedure as the transformation would then apply to the external u quark as well as the internal one, relating to a different observable process. These lead to the amplitude relation,

$$\mathcal{A}(B_d^0 \rightarrow K^+ \pi^-) = -(P + P_{\text{em}} + T) \quad (4.14)$$

where

$$P \equiv \mathcal{A}(B^+ \rightarrow K^0 \pi^+), \quad (4.15a)$$

$$\begin{aligned} T &\equiv |T|e^{i\delta_T}e^{i\gamma} \\ &= \lambda^4 A R_b \left[\tilde{T} - \mathcal{A} + \left(\tilde{\mathcal{P}}^u - \mathcal{P}^u \right) + \left(\tilde{\mathcal{P}}_{\text{ew}}^u - \tilde{\mathcal{P}}_{\text{ew}}^t \right) - \left(\mathcal{P}_{\text{ew}}^u - \mathcal{P}_{\text{ew}}^t \right) \right] e^{i\gamma} \end{aligned} \quad (4.15b)$$

and

$$\begin{aligned} P_{\text{em}} &\equiv -|P_{\text{em}}|e^{i\delta_{\text{em}}} \\ &= -\left(1 - \frac{\lambda^2}{2}\right)\lambda^2 A \left[\left(\tilde{\mathcal{P}}_{\text{ew}}^{\text{t}} - \tilde{\mathcal{P}}_{\text{ew}}^{\text{c}} \right) - (\mathcal{P}_{\text{ew}}^{\text{t}} - \mathcal{P}_{\text{ew}}^{\text{c}}) \right]. \end{aligned} \quad (4.15\text{c})$$

Here we have represented the “tree” and electro-weak penguin topologies’ amplitudes (T and P_{em} respectively) in terms of the QCD penguin amplitude, P . δ_T and δ_{em} are \mathcal{CP} -conserving strong phases and only colour-suppressed topologies contribute to the electro-weak penguin processes, i.e. the electro-weak penguin propagator’s daughter quarks are split between the K and π mesons and so their colour quantum numbers must match those of the s quark from the penguin loop and of the spectator d quark.

It is worth noting from equation (4.15) that T , usually referred to as a “tree” amplitude, is not purely composed of tree contributions from $\tilde{\mathcal{T}}$ but also incorporates both penguin and annihilation amplitudes.

$B \rightarrow K \pi$ decay observables

To parameterise observable quantities we introduce the ratios of tree and electroweak penguin amplitudes to the r.m.s. gluonic penguin amplitude,

$$r \equiv \frac{|T|}{\sqrt{\langle |P|^2 \rangle}}, \quad (4.16)$$

$$\epsilon \equiv \frac{|P_{\text{em}}|}{\sqrt{\langle |P|^2 \rangle}} \quad \text{with} \quad \sqrt{\langle |P|^2 \rangle} \equiv \frac{1}{2} \left(|P|^2 + |\bar{P}|^2 \right), \quad (4.17)$$

and define the strong phase differences between penguin topologies,

$$\delta \equiv \delta_T - \delta_{\text{tc}}, \quad \Delta \equiv \delta_{\text{em}} - \delta_{\text{tc}}. \quad (4.18)$$

From an experimental point of view, the main quantities of interest are the ratio and “pseudo-asymmetry” of \mathcal{CP} -conjugate branching ratios,

$$\left\{ \begin{array}{c} R \\ A_0 \end{array} \right\} = \frac{\mathcal{B}(B_d^0 \rightarrow K^+ \pi^-) \pm \mathcal{B}(\bar{B}_d^0 \rightarrow K^- \pi^+)}{\mathcal{B}(B^+ \rightarrow K^0 \pi^+) + \mathcal{B}(B^- \rightarrow \bar{K}^0 \pi^-)}, \quad (4.19)$$

which are important for extracting γ . By applying the \mathcal{CP} transformation $\gamma \xrightarrow{\mathcal{CP}} -\gamma$ to obtain the amplitude for the \mathcal{CP} -conjugate decays and substituting the amplitudes into equation (4.19) in the form $\mathcal{B} \propto |\mathcal{A}|^2$, R and A_0 can be explicitly written as

$$\begin{aligned} R = 1 - 2r \frac{\cos \delta \cos \gamma + \rho_r \cos(\delta - \theta)}{\omega} + r^2 \\ + 2\epsilon \frac{\cos \Delta + \rho_r \cos(\Delta - \theta) \cos \gamma}{\omega} + \epsilon^2 \\ - 2r\epsilon \cos(\delta - \Delta) \cos \gamma \end{aligned} \quad (4.20)$$

and

$$A_0 = A_+ + \frac{2r \sin \delta \sin \gamma}{\omega} + 2r\epsilon \sin(\delta - \Delta) \sin \gamma + \frac{2\epsilon \rho_r \sin(\Delta - \theta) \sin \gamma}{\omega} \quad (4.21)$$

where A_+ is the “true” $B^\pm \rightarrow K^0 \pi^\pm$ asymmetry,

$$A_+ \equiv \frac{\mathcal{B}(B^+ \rightarrow K^0 \pi^+) - \mathcal{B}(B^- \rightarrow \bar{K}^0 \pi^-)}{\mathcal{B}(B^+ \rightarrow K^0 \pi^+) + \mathcal{B}(B^- \rightarrow \bar{K}^0 \pi^-)} = -\frac{2\rho_r \sin \theta \sin \gamma}{\omega^2}, \quad (4.22)$$

and

$$\omega \equiv \sqrt{1 + 2\rho_r \cos \theta \cos \gamma + \rho_r^2}. \quad (4.23)$$

A non-zero measurement of A_+ implies strong re-scattering in $B^+ \rightarrow K^0 \pi^+$ final states [158].

Finally, we note the definition of the neutral channel asymmetry,

$$A_{K\pi} \equiv \frac{\mathcal{B}(\bar{B}_d^0 \rightarrow K^- \pi^+) - \mathcal{B}(B_d^0 \rightarrow K^+ \pi^-)}{\mathcal{B}(\bar{B}_d^0 \rightarrow K^- \pi^+) + \mathcal{B}(B_d^0 \rightarrow K^+ \pi^-)}, \quad (4.24)$$

which at first-order is the direct \mathcal{CP} -asymmetry in $B_d^0 \rightarrow K \pi$ decays, assuming that the charge of the kaon directly corresponds to the flavour of the b quark, with $K^+ \Rightarrow B_d^0$ and vice-versa.

Present experimental constraints on R and A_0 are available from the branching ratios and \mathcal{CP} -asymmetries measured by CLEO and the b -factory experiments. Taking the Particle Data Group values for the zero-compatible $B^+ \rightarrow K^0 \pi^+$ \mathcal{CP} -asymmetry and the associated branching ratio and taking the unweighted mean of the *BABAR* and *Belle*

measurements for $B_d^0 \rightarrow K^+ \pi^-$ the values are [10, 20, 21, 159]

$$R = 0.73 \pm 0.20 \quad (4.25a)$$

$$A_0 = 0.086 \pm 0.024. \quad (4.25b)$$

The corresponding averaged measurements for A_+ and $A_{K\pi}$ are [10, 20, 21]

$$A_+ = 0.03 \pm 0.08 \quad (4.26a)$$

$$A_{K\pi} = -0.12 \pm 0.04. \quad (4.26b)$$

4.2.3 Extracting γ from $B \rightarrow K \pi$ decays

The above derivations have produced a model-independent description of $B \rightarrow K \pi$ decays in the Standard Model, incorporating not only the well-described processes from tree and gluonic penguin topologies but also including a phenomenological parameterisation of final state re-scattering and electroweak penguin effects in terms of the amplitudes ρ_r , ϵ and the associated phases θ and δ_{em} .

The approaches to extracting γ from the measured R and A_0 observables split neatly into those which introduce no model-dependence but which can only place bounds on γ , and those which are model-dependent but permit a direct measurement of γ . This model-dependence is introduced in the value of r , which characterises the ratio of tree to penguin amplitude contributions.

Model-independent bounds on γ

A measurement of A_0 allows us to eliminate δ by algebraic constraint from the expression for R given in (4.20). The detailed procedure of this elimination is given explicitly, including re-scattering and electroweak penguin terms, in reference [158]. By treating r as a free variable in the resulting expression we find a model-independent minimum value for R :

$$R_{\min} = \kappa \sin^2 \gamma + \frac{1}{\kappa} \left(\frac{A_0}{2 \sin \gamma} \right)^2 \quad (4.27)$$

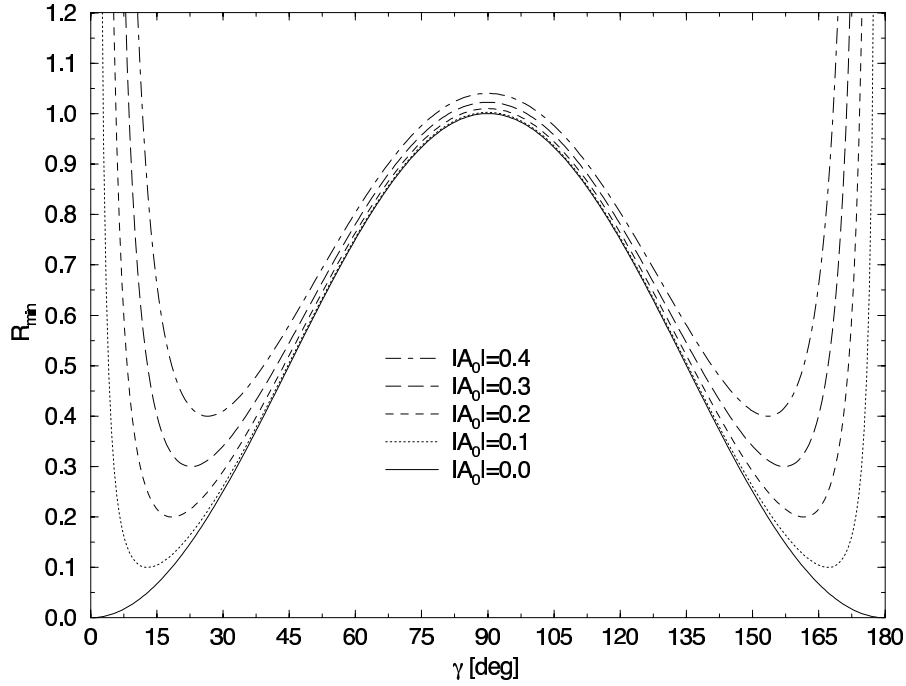


Figure 4.5: The dependence of R_{\min} on γ for various values of A_0 with neglected re-scattering and electroweak penguin effects.

where the re-scattering and electroweak penguin effects are absorbed into the parameter κ , given by

$$\kappa = \frac{1}{\omega^2} [1 + 2\epsilon\omega \cos \Delta + (\epsilon\omega)^2]. \quad (4.28)$$

Note that if we ignore re-scattering and EW penguin effects, then $\kappa = 1$.

Neglecting re-scattering and electroweak penguin effects ($\rho_r = \epsilon = 0$), R_{\min} is plotted in Figure 4.5. If $R < 1$ then we can exclude the regions of γ for which $R_{\min} > R$. If the errors R become small in the future, the current central value of $R \sim 0.73$ will result in the exclusion of a large region around 90° . Constraining the value of A_0 will exclude two smaller regions near 0° and 180° , where the dependence of R_{\min} on γ is much stronger. The current value of $A_0 \sim 0.085$ would give these exclusion regions a maximum width of $\sim 5^\circ$. This is the generalisation of the ‘‘Fleischer–Mannel bound’’ on γ [152, 158, 160].

The values of γ from 0° – 90° correspond to $\cos \delta > 0$ and those between 90° and 180° correspond to $\cos \delta < 0$: the former range is preferred theoretically.

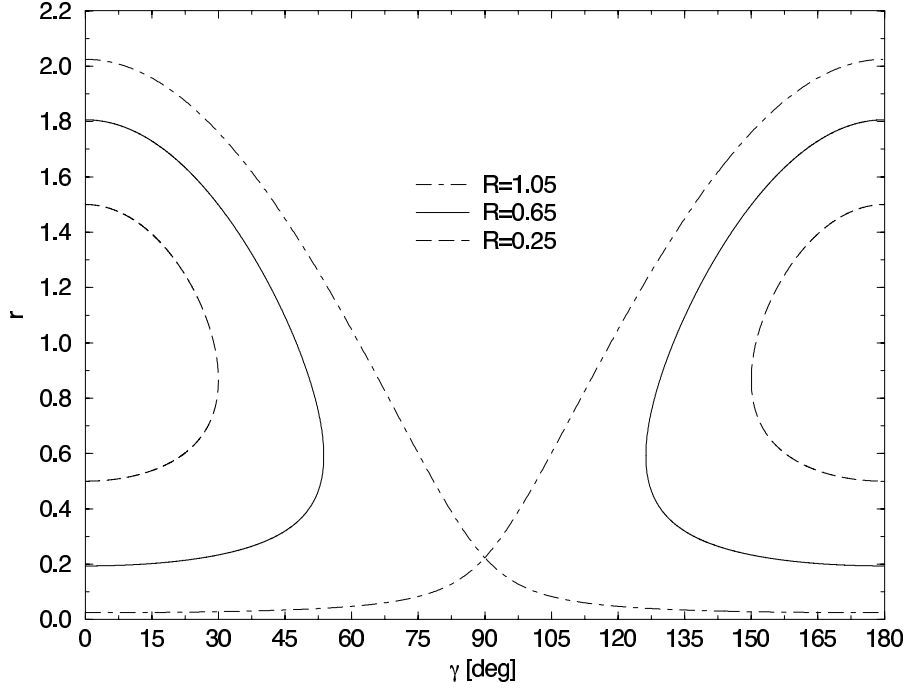


Figure 4.6: r upper and lower bounds as a function of γ and R .

4.2.4 Model-dependent determination of γ

Another useful constraint can be obtained from a measurement of R . Treating δ as a free parameter we obtain a bound on r of the form

$$r_{\min} \leq r \leq r_{\max} \quad (4.29)$$

with

$$r_{\min, \max} = \left| \sqrt{R_0 - \kappa \sin^2 \gamma} \pm \sqrt{R - \kappa \sin^2 \gamma} \right|, \quad (4.30)$$

where

$$R_0 = 1 + \frac{2\epsilon}{\omega} [\cos \Delta + \rho_r \cos(\Delta - \theta) \cos \gamma] + \epsilon^2. \quad (4.31)$$

This pair of bounds is shown in Figure 4.6 for values of R from 0.25–1.05 (the current central value, from *BABAR* and *Belle*, is $R \sim 0.73$). Theoretical central values of r lie between about 0.1 and 0.2 [147, 161].

Measurement of the pseudo-asymmetry A_0 allows for a much stronger constraint on

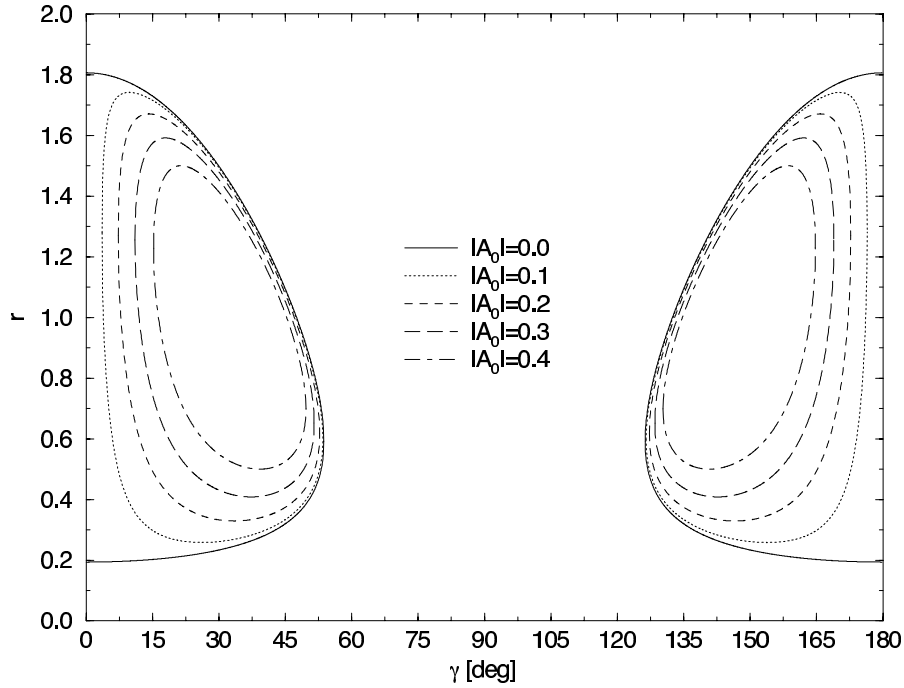


Figure 4.7: r upper and lower bounds as a function of γ and A_0 .

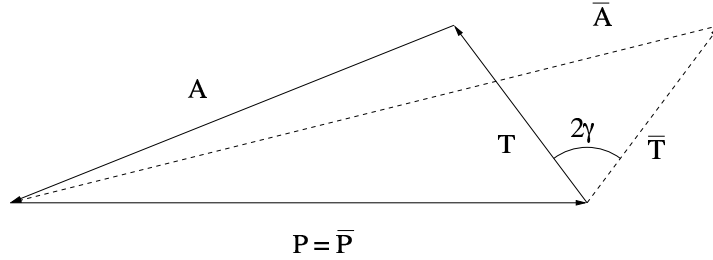


Figure 4.8: The triangle construction corresponding to γ -determination via the “mixed” $B \rightarrow K \pi$ decay strategy.

r :

$$r = \sqrt{r_0^2 + (R - R_{\min}) \pm \sqrt{K(\rho_r, \theta, \epsilon, \Delta, \gamma)(R - R_{\min})}}, \quad (4.32)$$

where K is a function obtained in the process of using A_0 to eliminate the strong phase δ . Ignoring re-scattering and electroweak penguin effects again by setting $\rho_r = \epsilon = 0$, it reduces to simply $K = 2 \cos \gamma$. Plotting equation (4.32) produces a set of contours in the γ - r plane, as shown for a range of values of A_0 in Figure 4.7. This corresponds to a simple triangle construction composed of $B \rightarrow K \pi$ tree and penguin amplitudes, which is shown in Figure 4.8.

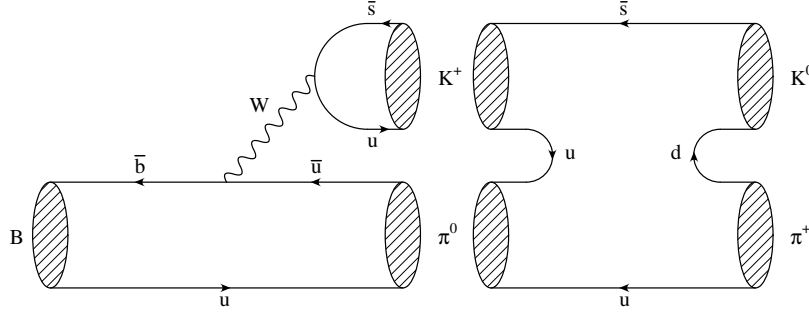


Figure 4.9: An explicit re-scattering contribution to $B^+ \rightarrow K^0 \pi^+$.

4.3 Hadronic uncertainties in $B \rightarrow K \pi$ decays

The majority of the model-dependency of the strategy described so far enters in knowing the value of r , the ratio of tree to penguin amplitudes, and in the effects of final state re-scattering. In fact, re-scattering processes may also significantly shift r , making a precision measurement of γ rather difficult with this strategy.

4.3.1 Effects of re-scattering on $B \rightarrow K \pi$ decay phenomenology

Re-scattering effects in this strategy are encoded in the phenomenological parameters ρ_r and the associated phase θ . A set of re-scattering topologies for $B^+ \rightarrow K^0 \pi^+$ channels, which illustrate the general principle, are shown in Figure 4.9.

Significantly, $\rho_r \neq 0$ and $\theta \notin \{0, \pi\}$ immediately implies direct \mathcal{CP} -violation in the mode $B^+ \rightarrow K^0 \pi^+$, as can be seen from equation (4.22). Re-scattering effects are CKM suppressed, with $\rho_r \sim \lambda^2 R_b \approx 0.02$, where only tree and top-dominated penguin amplitudes have been considered: this implies that re-scattering effects should be small. A more careful perturbative treatment, including the charm and up-quark penguin contributions and annihilation topologies, does not reveal any significant compensation for this CKM suppression. However, if re-scattering processes such as $B^+ \rightarrow \left\{ F_u^{(s)} \right\} \rightarrow K^0 \pi^+$ (where $F_u^{(s)} \in \{K^+ \pi^0, K^{*+} \pi^0, \dots\}$) dominate over charmed intermediate states then the ratio of penguin amplitude differences, $|\mathcal{P}_{uc}|/|\mathcal{P}_{tc}|$ can become large (of order 10) and re-scattering may be correspondingly enhanced to give $\rho_r \sim \mathcal{O}(0.1)$. A similar magnitude may be found by considering a “factorised” expression for r to act as an upper bound on ρ_r .

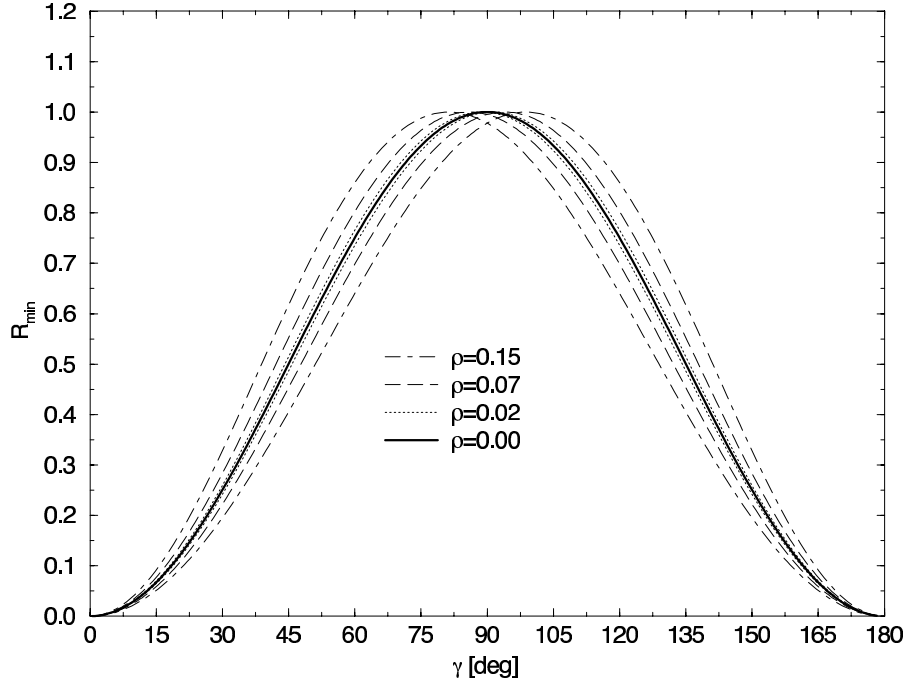


Figure 4.10: Effects of final state re-scattering on the R_{\min} bound.

Effect of re-scattering on γ bounds

If we again ignore the effects of electroweak penguins, then the expression for R_{\min} given in equation (4.27) simplifies to

$$R_{\min} = \left(\frac{\sin \gamma}{\omega} \right)^2 + \left(\frac{\omega A_0}{2 \sin \gamma} \right)^2, \quad (4.33)$$

which shifts the central value of R_{\min} depending on the re-scattering phase θ . The range of shifts obtained for $A_0 = 0$ is illustrated in Figure 4.10, with the maximal shifts occurring for $\theta \in \{0, \pi\}$. The observed maximal shifts, for $\rho_r \sim \mathcal{O}(0.15)$, only lead to shifts in the bound on γ of order 10% and around $\gamma = 90^\circ$ where the re-scattering effects are minimised a large region may still be excluded.

It is also notable that the maximal re-scattering shifts are obtained for the phases which also preclude direct \mathcal{CP} -violation in $B^\pm \rightarrow K^0 \pi^\pm$ decays — a measurement of non-zero direct \mathcal{CP} -violation in this channel allows re-scattering effects on the bounds to be further constrained. It has been further suggested that branching ratio measurements from $B^\pm \rightarrow K^\pm K^0$ channels may be used to factor out re-scattering effects via a U-spin symmetry argument — this is considered in detail in references [158, 162, 160].

Effects of re-scattering on γ -determination

For direct comparison with the shifts to bounds on γ via R_{\min} , Figure 4.10 gives the maximal shifts of the r - γ contours: uncertainties on the γ contours furthest from $\gamma = 90^\circ$ are $\mathcal{O}(3^\circ)$ while the other set of contours, including that around $\gamma = 50^\circ$ which is favoured by indirect CKM fits, have uncertainties $\mathcal{O}(15^\circ)$.

Unfortunately, constraints on the amplitude ratio r (whose value is needed to use these contours for a determination of γ) are affected by re-scattering in a more problematic way. This is because the terms which cancel in the hard-scattering diagrams cannot be easily removed when they enter as long-range re-scattering effects: these topologies are shown in Figure 4.11. Approaches to control these effects are discussed in Section 4.3.2. It should be noted that until these effects are well-understood, this pair of decay channels are of more interest for placing weak bounds on γ and for constraining the potential new-physics behaviour of electroweak penguins.

4.3.2 Theoretical approaches to re-scattering

The large non-perturbative contributions to non-leptonic B-decays mean that they are considered the most awkward kind of decay to study from first principles. As already highlighted, the effects of re-scattering effects on CKM phenomenology mean that good models of long-range effects are desperately required. At present there are four main theoretical approaches to calculate and/or parameterise the hadronic matrix elements in $B \rightarrow M_1 M_2$ decays:

- $SU(2)/SU(3)$ symmetries and phenomenological ansätze,
- Dynamical approaches based on perturbative QCD, such as “QCD factorisation” and the competing “pQCD” approach,
- A topological invariant approach, “charming penguins”, based on symmetries of the renormalisation group,
- Soft co-linear effective theory (SCET), for which several formulations exist.

Concerning non-leptonic decays, the largest current discrepancy between the data and the SM is in the decays involving QCD penguins. These include CP violation in the $b \rightarrow s\bar{s}s$ penguins, where data show a deviation of about 3σ from the SM. Also the

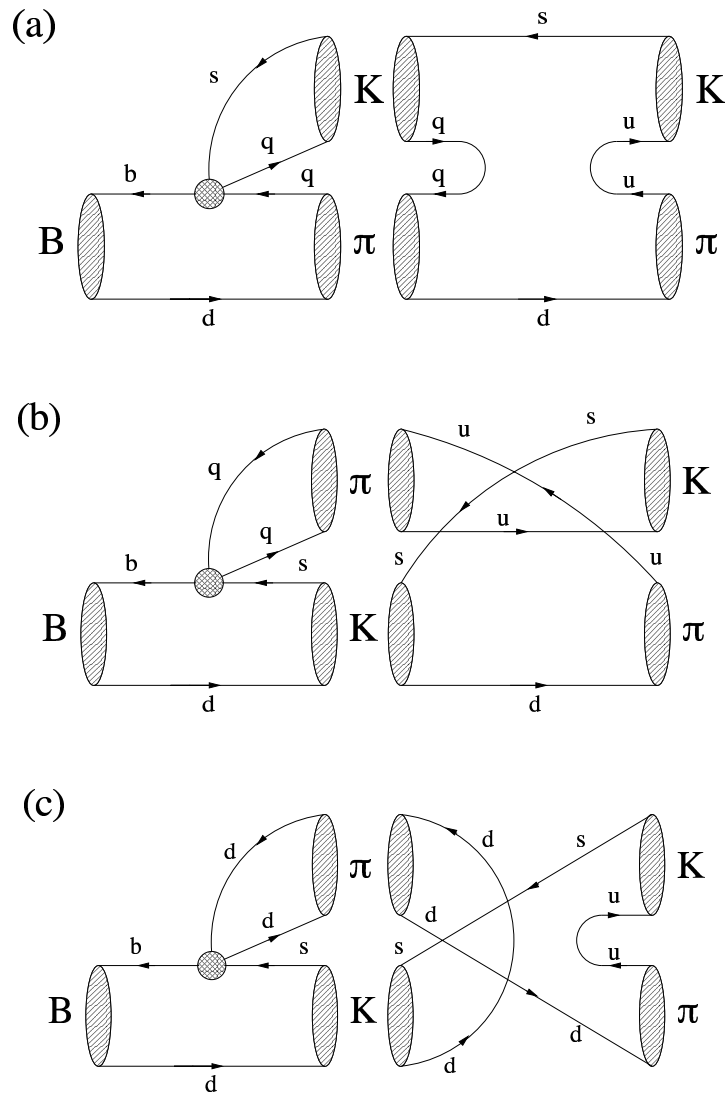


Figure 4.11: Re-scattering topologies for general $B \rightarrow K \pi$ decays.

ratio R_n in the $B \rightarrow K \pi$ decays is out of line with the SM-estimates by slightly more than 3σ . These deviations are not yet significant enough to announce physics beyond the SM [163].

Heavy quark effective theory (HQET)

HQET is an $SU(2N_H)$ symmetry of heavy quarks which can be used in charmed B decays, since heavy, non-perturbative degrees of freedom can be integrated out of the Lagrangian: a convenient analogy is with the atomic proton–electron system where isotopes have the same chemistry since the heavy degrees of freedom are very restricted relative to the electron ones. This is a standard approach to constraining hadronic characteristics in charmed B-decays like $B_d^0 \rightarrow D^+ D^-$ but for charmless decays the heavy quark symmetry from $b \rightarrow \{u, d, s\}$ is very broken and so HQET cannot be used.

As the K and π states are charmless light mesons, HQET cannot be reliably applied to these channels and will not be reviewed in any further detail.

Approximate symmetries and phenomenological ansätze

The simplest and perhaps most reliable approach in constraining non-factorisable re-scattering processes is to make use of approximate symmetries such as nuclear isospin, nuclear $SU(3)$ and U-spin. These postulate approximate symmetries under exchange of u/d quarks, u/d/s quarks and d/s quarks respectively. For the channels described here, the $SU(3)$ symmetry suggests $B^\pm \rightarrow K^\pm K^0$ as a conjugate version of the $B^\pm \rightarrow K^0 \pi^\pm$ channels through $d \rightarrow s$ exchange [158, 162]. $B^\pm \rightarrow K^\pm K^0$ is very sensitive to re-scattering effects and, combined with a measurement of A_+ , allows the $B^\pm \rightarrow K^0 \pi^\pm$ re-scattering to be controlled. The main difficulty in this procedure is that $B^\pm \rightarrow K^\pm K^0$ has not yet been observed, with the Belle collaboration having set an upper limit of $\mathcal{B}(B^+ \rightarrow K^+ K^0) < 3.3 \times 10^{-6}$ [164]. Re-scattering effects may be expected to enhance the $B^\pm \rightarrow K^\pm K^0$ branching ratio by a factor of up to five times above the quark-level calculation [165, 166, 167, 168], so observation of this decay may yet be possible and will permit application of this method.

Soft co-linear effective theory (SCET)

SCET is an extension of the QCD factorisation formalism already described, which attempts to incorporate yet more long-range QCD effects, for example re-scattering. This is achieved by explicitly identifying the degrees of freedom of the spectator quark, which tends to be a soft (i.e. long-range) particle due to not interacting directly with the b-quark line [169].

The formalism of SCET is similar in character and complexity to that of QCD factorisation and will not be discussed in detail. SCET is considered to be a very promising path for evaluating light non-leptonic decays in the future but until physics input can be provided to predict its free parameters characterising non-factorisable effects it is no more predictive than QCD factorisation. Research into the SCET description of B-decays is relatively undeveloped at present and there are many formal properties of the competing formalisms under discussion. The intense activity adds hope that SCET will provide more insights into non-leptonic decays in the near future.

Perturbative QCD (pQCD)

The perturbative QCD (pQCD) approach to B-decays is in the same vein as that described already. It emphasises that the spectator quark requires a hard gluon exchange to provide a boost to a velocity comparable with that of the B quark in order for a coherent hadronisation to occur. Three scales are then identified in the problem: M_W , the electroweak scale as encountered at Feynman vertices introducing a CKM factor; $t \sim \mathcal{O}(\Lambda_{QCD} m_b)$, the hard QCD interaction scale; and $1/b$, the factorisation scale, where b is the conjugate variable to the parton transverse momenta [170]. All interactions below the factorisation scale are non-perturbative and are absorbed into light-cone distribution amplitudes while the hard scattering is represented by hard scattering kernels (cf. QCD factorisation) with Wilson factors as described in Section 4.1.1.

Analysis of $B \rightarrow K \pi$ decays in the pQCD formalism predicts that $r = 0.201 \pm 0.037$ [171].

4.4 Future prospects

Experimental evidence is now mounting that the strong QCD phases contributing to these processes are large, at a level unlikely to be generated by perturbative QCD alone.

In addition, colour-suppressed decays are not parametrically suppressed, as opposed to their estimates in the QCD factorisation and pQCD approaches.

Consensus in the field is heading toward the decision that $B \rightarrow K \pi$ decays are crippled theoretically by the increasing evidence for non-factorisable contributions to the decay amplitudes. Until theoretical control is achieved over these modes, precision extraction of γ is not likely to be obtained with this channel. However, theoretical technologies in active development, such as the SCET formalism offer hope for the future. Experimentally, the focus for extraction of γ is shifting to strategies which rely on mode combinations like $B_d^0 \rightarrow K^+ K^-$ and $B_s^0 \rightarrow \pi^+ \pi^-$, which are experimentally cleaner than those channels containing neutrals and are relatively free from theoretical difficulties.

$B \rightarrow K \pi$ is, however, interesting as a laboratory in which to study the world of electroweak penguins. These clearly “operate” at the electroweak symmetry breaking scale $\sim \mathcal{O}(100 \text{ GeV})$ and are thus more sensitive to BSM physics effects than the topologically equivalent gluonic penguins. As already mentioned, the “ $B \rightarrow K \pi$ puzzle” of the discrepancy in decay rates between channels with colour-allowed and colour-suppressed EW penguins indicates that BSM physics signatures may be at work and that the further study of these channels may offer insight into an entirely new regime of fundamental physics.

Chapter 5

Selecting $B \rightarrow K \pi$ decays with LHCb

“Science may be described as the art of systematic oversimplification.”

— Karl Popper, 1902–1994

In this chapter we discuss the procedures used for reconstruction of $B \rightarrow K \pi$ decays with the LHCb detector and the automated procedure developed to optimise the signal/background discrimination process. The raw results of applying these tuned cut sets to the simulated data sets are presented in this chapter; their interpretation is dealt with in Chapter 6.

5.1 Decay channels of interest

In this section we briefly summarise the branching ratios for signal and background data sets.

5.1.1 Signal channels

The two B-decay channels of interest, whose phenomenology was described in the previous chapter, are $B_d^0 \rightarrow K^+ \pi^-$ and $B^+ \rightarrow K^0 \pi^+$. The current branching fractions into these channels are $\mathcal{B}(B_d^0 \rightarrow K^+ \pi^-) = (1.85 \pm 0.11) \times 10^{-5}$ and $\mathcal{B}(B^+ \rightarrow K^0 \pi^+) = (2.20 \pm 0.19 \pm 0.11) \times 10^{-5}$

where the uncertainties in the second case are first statistical and secondly systematic. The $B_d^0 \rightarrow K^+ \pi^-$ branching fraction is taken from the Particle Data Group world average [10] and is in agreement with the latest measurements from the Belle collaboration [159], which are not included in the PDG average. The $B^+ \rightarrow K^0 \pi^+$ fraction is taken purely from the Belle measurement [159] as the most recent and accurate value; there is $\sim 1\sigma$ disagreement between the Belle measurement and the PDG average. These branching ratios are in agreement with Standard Model theory, with a large \mathcal{CP} asymmetry observed in the $B_d^0 \rightarrow K^\pm \pi^\mp$ channel [20, 21].

From the K^0 produced in the B^+ decay we only observe the decay products from the \mathcal{CP} eigenstate K_S^0 , with

$$|K_S^0\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle + |\bar{K}^0\rangle). \quad (5.1)$$

The relevant branching ratio is half of that into K^0 : $\mathcal{B}(B^+ \rightarrow K_S^0 \pi^+) = \frac{1}{2} \times \mathcal{B}(B^+ \rightarrow K^0 \pi^+)$. Additionally, we will only attempt to reconstruct the $K_S^0 \rightarrow \pi^+ \pi^-$ final state, since the LHCb calorimeter is not considered sufficiently accurate to study the sub-dominant $K_S^0 \rightarrow \pi^0 \pi^0$ process. There is therefore another additional branching fraction of 68.61% (with negligible uncertainty) to be accounted for. In total $\mathcal{B}(B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+) \approx (7.5 \pm 0.7 \pm 0.4) \times 10^{-6}$.

5.2 Sources of background

The most significant sources of background to our signal channels are expected to be inclusive $b\bar{b}$ events, which outnumber the signal by a factor of about 10^6 , and those decays which only differ from our signal channels by the identity of the light charged mesons. The latter must be resolved by the RICH particle ID system. No explicit studies are made of minimum bias ($pp \rightarrow \text{anything}$) events, as they have been previously shown to be unproblematic [95].

The topologically-similar decay channels which constitute our specific backgrounds are $B^+ \rightarrow K^+ K^0$, $B_d^0 \rightarrow \pi^+ \pi^-$ and $B_d^0 \rightarrow K^+ K^-$. $B^+ \rightarrow K^+ K^0$ is not expected to be a particularly strong source of contamination since the Belle collaboration have recently placed an upper limit on the branching ratio of $\mathcal{B}(B^+ \rightarrow K^+ K^0) < 3.3 \times 10^{-6}$, only seeing a signal significance $\sim \mathcal{O}(2\sigma)$ [159]. This limit does not incorporate the $K^0 \rightarrow K_S^0$ and $K_S^0 \rightarrow \pi^+ \pi^-$ branching fractions and so should be expected to only enter with at most a

tenth of the size of the $B^+ \rightarrow K_s^0 \pi^+$ yield. Similarly $B_d^0 \rightarrow K^+ K^-$ has an upper limit of $\mathcal{B}(B_d^0 \rightarrow K^+ K^-) < 0.7 \times 10^{-6}$ and only $B_d^0 \rightarrow \pi^+ \pi^-$ has been measured with a significance of greater than 5σ : $\mathcal{B}(B_d^0 \rightarrow \pi^+ \pi^-) = (4.4 \pm 0.7) \times 10^{-6}$ [159].

An additional specific source of background is the $B_s^0 \rightarrow K^+ K^-$ decay channel. This has the potential to be more problematic than the B_d^0 and B^+ specific backgrounds, because, while mis-identification of one of the kaons as a pion would shift the invariant mass peak of the combination to a lower value, the B_s^0 mass is already slightly higher than that for B_d^0/B^+ . The danger is that this mis-ID can place the invariant mass of the $K^+ K^-$ very close to the desired B_d^0/B^+ mass. However, the combination of the low relative B_s^0 fragmentation fraction, $f_{B_s^0}/f_{B_d^0} \sim 0.27$ and the low $K \rightarrow \pi$ mis-ID rate (typically less than 5%) mean that this contamination is expected to be a $\sim 1\%$ effect and so it has not been included in this study.

Unfortunately no $B^+ \rightarrow K^+ K^0$ events were produced as part of the LHCb 2003 Monte Carlo production, on which this analysis is based, and so we can only study the contamination from competing signal channels in the $B_d^0 \rightarrow K^+ \pi^-$ decay channel, for which the $B_d^0 \rightarrow K^+ K^-$ and $B_d^0 \rightarrow \pi^+ \pi^-$ data samples will be checked against the selection cuts (which, as will be shown, are optimised solely against inclusive $b\bar{b}$ events: the optimisation does not account for these specific backgrounds).

5.3 Simulation and initial reconstruction

The events used in this study were not simulated using the EvtGen package described in Chapter 2; instead data from the DC03 Monte Carlo production is analysed here. In this production, the Pythia 6.2 event generator was used to generate the hard process and initial parton showers, with the configuration set to include hard QCD processes, single and double diffraction and elastic scattering. These events do not contain charge conjugates of signal decay channels, so all considerations of systematic uncertainties to follow cannot account for e.g. real charge asymmetries in reconstruction. The output from Pythia is a set of 4-momenta of the products of the primary interaction, the decays of which were simulated with the QQ package [172]. The interaction of these particles with the detector material was simulated with the Geant3 framework [173], in the context of the LHCb SICB software [174]. It should be noted that neither Geant3 nor SICB are included in the object-oriented implementation of the LHCb software but that the newer versions have largely been tuned and debugged by matching the performance of the DC03

programs. The reconstruction from the Geant3 output was performed by Brunel v17.4, which includes the digitisation stages which were later moved to the Boole system [175].

A geometric selection cut was applied at generator level to the hadronised particle of interest, which is the signal B^+ or B_d^0 in the case of signal events or at least one of the b-hadrons in inclusive $b\bar{b}$ ($b\bar{b} \rightarrow \text{anything}$) events. This cut is for computational efficiency, as events where the particle of interest is produced outside of the experiment’s geometric acceptance will anyway not be reconstructible. The particle of interest must have a polar angle within 400 mrad of the beam-line to fall within the LHCb acceptance and so only events which match this criterion are considered. The efficiency of this generator-level cut is 34.7% for signal events and 43.2% for generic $b\bar{b}$ events [95]. The Monte Carlo event samples used in this study are 54,500 $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ events, 45,900 $B_d^0 \rightarrow K^+ \pi^-$ events and 10^7 inclusive $b\bar{b}$ events.

An extra efficiency factor is that not all of the generated events contain the final state tracks as desired (the “decay of interest”, denoted “DoI”). For neutral B decays, B-oscillations can result in the event not containing the desired decay chain, resulting in a DoI rate of less than 100%. For the channel containing a K_S^0 , the K_S^0 can interact hadronically with the detector before decaying and as a result the correct final-state particles (the two pions) are not present, again leading to a loss of efficiency. For the $B_d^0 \rightarrow K^+ \pi^-$ channel, use of the DaVinci “efficiency algorithm” identifies 81.9% of the events as containing the decay of interest due to oscillations. The comparable number for the $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ channel is 89.5%, due to K_S^0 interactions. Both these efficiencies are to be absorbed into the detector efficiency when we consider it in Chapter 6.

5.4 Event reconstruction

The details of reconstructing the $B_d^0 \rightarrow K^+ \pi^-$ and $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ signal events will now be described. All the advanced reconstruction tasks, such as combining tracks to “make” parent particles and reconstructing the K_S^0 , are carried out in the object-oriented DaVinci framework with a set of custom algorithms.

The two algorithms used for this reconstruction, `Bu2KsPiPi_Pi` and `Bd2KPi` are located in the `Bu2KsPi` v1.4 DaVinci package and are derived from a common `B2KPi` algorithm to make direct comparison of the results easier and more robust. Both selection algorithms take the form of identifying K and π candidates and then applying a

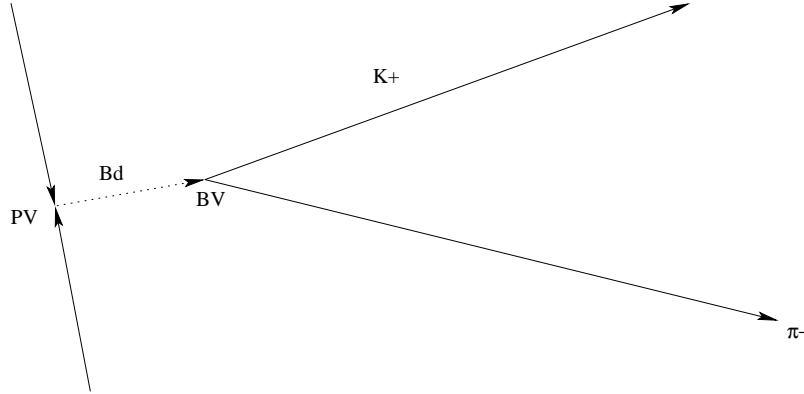


Figure 5.1: $B_d^0 \rightarrow K^+ \pi^-$ event topology. The B_d^0 propagates unseen for several millimetres before decaying.

set of selection cuts to each. The daughter candidates which pass these cuts are then combined to make a candidate B, which must again pass selection cuts before being selected.

5.4.1 Reconstructing $B_d^0 \rightarrow K^+ \pi^-$

The topology of the $B_d^0 \rightarrow K^+ \pi^-$ decay is shown in Figure 5.1: the salient features are the small displacement of the B decay vertex (secondary vertex) from the interaction (primary) vertex and the observed K^+ and π^- tracks originating directly from the secondary vertex.

Reconstruction of this channel uses a DaVinci “combined particle maker” algorithm to construct the K and π candidates. In choosing which tracks from which to construct particles, the particle maker can apply a confidence level cut on the particle ID of the charged tracks which it assigns to be of a particular type, based on information from the tracking system and the RICH detectors. The confidence level distributions of signal and background candidates for the π^- and K^+ are shown in Figure 5.2. These distributions offer no obvious opportunities for making selection cuts so, to allow for more flexible tuning later, the particle maker was configured to make no cut at this stage.

The sets of K and π candidates are then individually filtered by a set of selection cuts. The properties for which distributions are available at this stage are the reconstructed total momenta p , transverse momenta p_T and the impact parameters (IP) with respect to the primary vertex. In the absence of better information, the particles are assigned the nominal K^+ and π^- masses as appropriate.

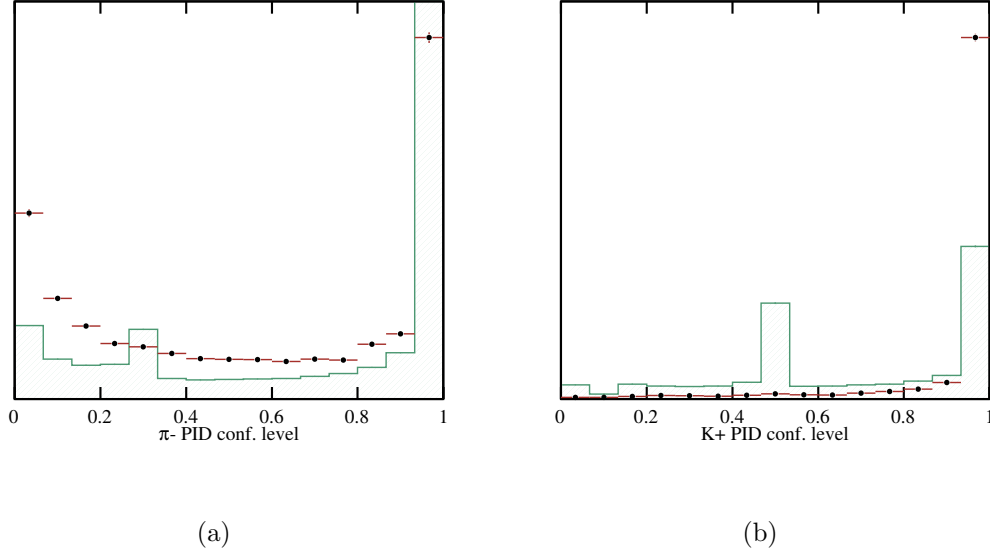


Figure 5.2: Particle ID confidence levels for K^+ and π^- particle candidates in 46k $B_d^0 \rightarrow K^+ \pi^-$ signal events. The filled distributions are from background candidates while the data points are from signal. The normalisation is arbitrary, to show relative distribution shapes rather than magnitudes, in which the background is strongly dominant.

After the individual K and π selection cut filtering, all unique K – π candidate pairs are considered. The first step is to combine the two reconstructed tracks to produce a vertex (defined by a position vector and an error matrix) by tracing through the magnetic field map. The rest mass of the particle which decayed at the reconstructed vertex is obtained from the reconstructed momentum 4-vectors of the daughter tracks. The resulting raw invariant-mass distribution arising from the unconstrained vertexing is shown in Figure 5.3, with the small “true” B peak at $5280 \text{ MeV}/c^2$ just visible. For computational convenience, the huge background peak in the invariant-mass at $\sim 1 \text{ GeV}/c$ is removed at this stage by requiring that all B candidates have a mass greater than $4 \text{ GeV}/c$.

Finally, the B candidates are also filtered. In addition to the properties already mentioned, the B candidates also have distributions in the χ^2 of the vertex fit, the invariant-mass m_{B^0} , the significance of the secondary vertex’s displacement from the primary vertex ℓ/σ_ℓ and the lifetime τ_{B^0} . Since the initial use of DaVinci is for tuning the selection cuts, more than one selection per event is allowed. However, since the phenomenology motivating this channel is simply dependent on branching ratios rather than any details of the reconstructed event, we should only consider one true signal

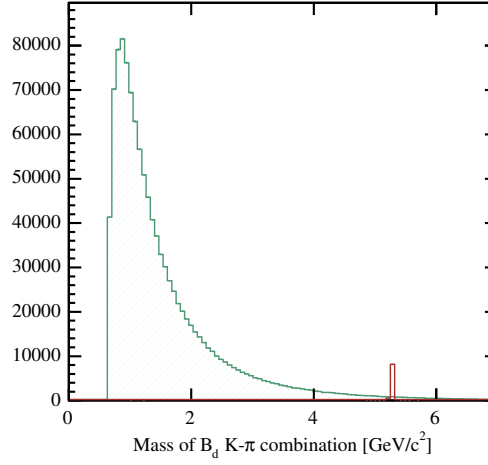


Figure 5.3: Raw invariant-mass of B_d^0 candidates from 50k signal events before the first B_d^0 mass cut. The huge peak at ~ 1 GeV/c is removed as soon as possible for computational efficiency. The sharp B_d^0 mass peak is located at ~ 5.2 GeV/c . The vertical scale represents the number of candidates per bin.

decay per event to be possible. This point will be returned to later.

5.4.2 Reconstructing $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$

The structure of the algorithm to reconstruct the $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ channel is very similar to that for $B_d^0 \rightarrow K^+ \pi^-$. However, the K_S^0 produced at the B decay vertex propagates unseen for an average distance of 80 cm (corresponding to a mean lifetime of 90 ps) before decaying to the observed $\pi^+ \pi^-$ pair as shown in Figure 5.4. This means that the K_S^0 must first be explicitly reconstructed before attempting to reconstruct the B^+ and that the additional uncertainties arising from the K_S^0 reconstruction will also propagate through the reconstruction chain.

The final-state pions can each be reconstructed from any track of the appropriate charge (at this stage referred to in the software as a “proto-particle”). A restriction is placed at the end of the event reconstruction to ensure that the lone pion from the B-decay and the π^+ from the K_S^0 decay are derived from different proto-particles.

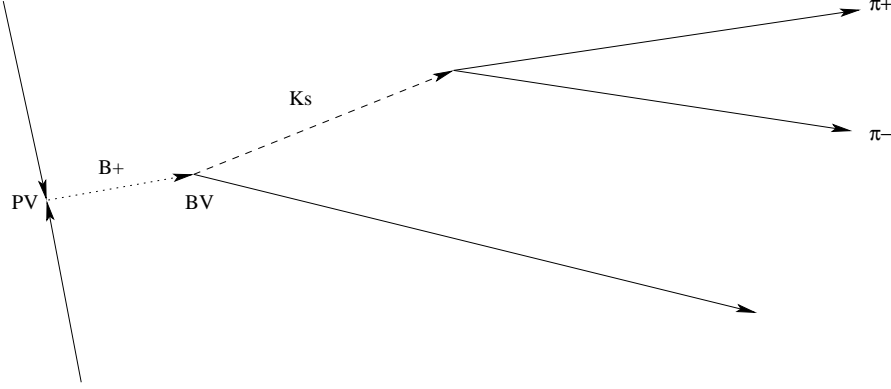


Figure 5.4: $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ event topology. The solid decay product lines indicate the trackable particles while the intermittent lines are unobserved directly. The B^+ propagates for several millimetres before decaying and the K_S^0 propagates for a mean distance of ~ 80 cm.

K_S^0 reconstruction

The K_S^0 reconstruction from the charged pion pair was performed using the `Ks2PiPiSel v1.0` [176] DaVinci package. Before continuing, it is useful to class the reconstructed tracks into five classes, named according to the positions where their track hits are recorded and illustrated in Figure 5.5:

- “Long” (L) tracks, which traverse all tracking systems from the VELO to the T stations.
- “Upstream” (U) tracks, which traverse the VELO and TT stations but leave no hits downstream of the magnet.
- “Downstream” (D) tracks, which traverse only the TT and T stations.
- “VELO” (V) tracks, which traverse only the VELO.
- “T” tracks, which traverse only the T stations (not normally used for analysis purposes).

The significance of these track type definitions is that there are no tracking stations located between the TT station and the T1–3 tracking stations about 6 m further downstream. This z -interval also has the misfortune of including the central region of the magnet and so tracks which are only seen from T1 onward have very poor vertexing potential. Downstream tracks are much poorer in this respect than long tracks but the

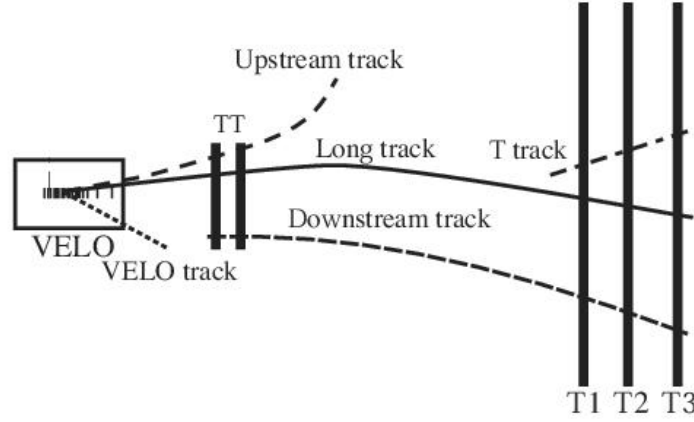


Figure 5.5: Track-type definitions relative to the upstream detector components.

K_S^0 mean flight distance of ~ 80 cm means that they are more numerous by about a factor of two.

The $K_S^0 \rightarrow \pi^+ \pi^-$ reconstruction package produces a collection of K_S^0 candidate particles for each event, classed according to the $\pi^+ \pi^-$ track type combinations: LL, DD, LD and LU. The remaining combinations, such as UU, are not considered as they only account for a small fraction of the K_S^0 decay products. As both pions are produced at the K_S^0 decay vertex, the track combinations are naturally correlated: the bulk of events fall into the LL and DD categories.

The cuts used by the K_S^0 reconstruction are listed in Table 5.1 for LL and DD K_S^0 . All further selection criteria applied to the K_S^0 selection must be considered as a super-set of these cut sets. Notably, a cut is made on the impact parameter significance of the pions since they are expected to point toward the K_S^0 decay vertex which is itself significantly displaced from the primary vertex.

Invariant-mass distribution plots for the 4 main classes of K_S^0 after reconstruction are shown in Figure 5.6: the mass resolutions are respectively $\sigma_m(\text{LL}) = 3.7 \text{ MeV}/c^2$, $\sigma_m(\text{DD}) = 7.5 \text{ MeV}/c^2$, $\sigma_m(\text{LD}) = 6.4 \text{ MeV}/c^2$ and $\sigma_m(\text{LV}) \sim 14.9 \text{ MeV}/c^2$ (all uncertainties less than $0.1 \text{ MeV}/c^2$). The background is a more significant problem than mass resolution for LD K_S^0 , which actually have a better resolution than the DD K_S^0 . Background contamination is also significant for LV K_S^0 . To get a clean K_S^0 input to the B^+ reconstruction, we only use LL and DD classes of K_S^0 in further reconstruction.

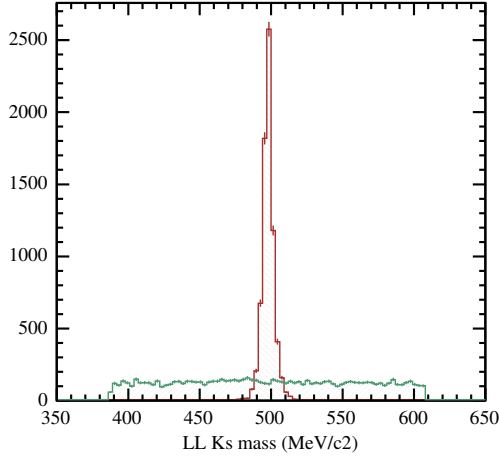
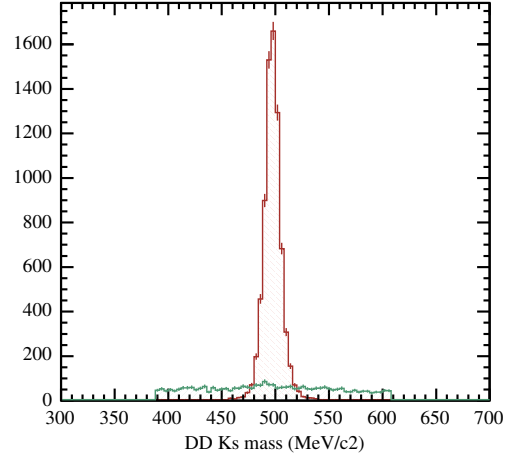
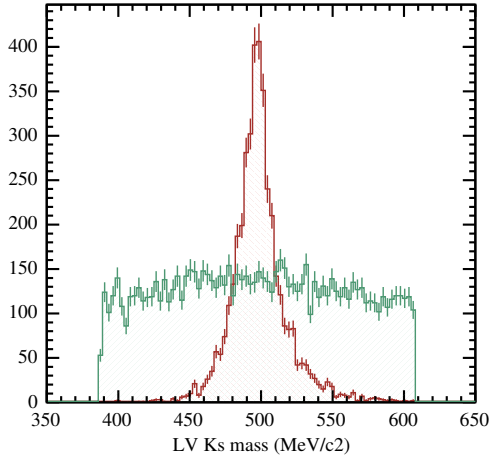
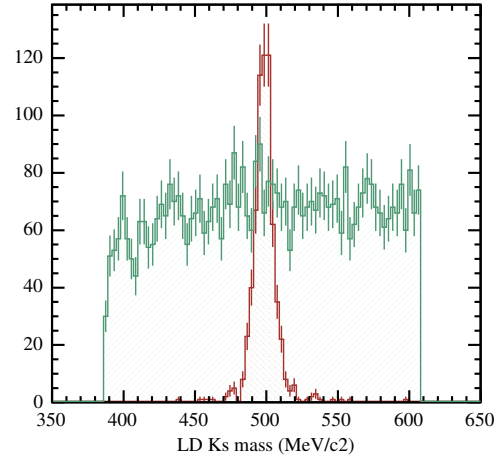
(a) for LL K_S^0 (b) for DD K_S^0 (c) for LV K_S^0 (d) for LD K_S^0

Figure 5.6: K_S^0 mass distributions (in MeV/c^2) from 50k signal events: the track-type combinations are for (a) LL, (b) DD, (c) LV & (d) LD $\pi^+\pi^-$ combinations. The peaked distributions are the signals and the flat distributions are the backgrounds, with the vertical scale representing the number of candidates per bin.

Cut variable	DD	LL
$ m - m_{K_S^0} $	$< 110.0 \text{ MeV}$	$< 110.0 \text{ MeV}$
z_{decay}	$> 200 \text{ mm}, < 3000 \text{ mm}$	$> 0 \text{ mm}, < 650 \text{ mm}$
χ^2_{vtx}	< 20.0	< 20.0
p_T	$> 800.0 \text{ MeV}/c$	$> 50.0 \text{ MeV}/c$
$\text{IP}_{\pi\text{-PV}}$	$> 3.0 \text{ mm}$	$> 6.0 \text{ mm}$

Table 5.1: Cuts used by the K_S^0 reconstruction package, given in terms of the condition required to pass the cut.

Table 5.2 shows the fractions of signal events of each K_S^0 type $\in \{\text{LL}, \text{DD}\}$ which are “reconstructible” and the fractions which are “reconstructed” as well as the corresponding reconstruction efficiencies.

Property	DD	LL
Number of decays of interest (DoIs)	—— 48919 (from 54500 events) ——	
Reconstructible no. & fraction of DoIs	6082 \rightarrow 12.4%	1748 \rightarrow 3.6%
Reconstructed no. & fraction of DoIs	4285 \rightarrow 8.6%	1484 \rightarrow 2.9%
Strict reconstruction efficiency	3977/6082 = 65.4%	1393/1748 = 79.7%
Effective reconstruction efficiency	4285/6082 = 70.5%	1484/1748 = 84.9%

Table 5.2: Reconstruction performance for K_S^0 in $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ decays on 54.5k events.

Again, we need some definitions to understand what we consider as a “reconstructible” track and what criteria a set of hits must meet before being considered as “reconstructed”. As in Chapter 3, the definition of “reconstructible” used is [95]

- for VELO tracks the particle must give at least 3 r hits and 3 ϕ hits;
- for T tracks the particle must give at least 1 x and 1 stereo hit in each of the T1–T3 stations;
- for long tracks the particle must be reconstructible as a VELO and a T track;
- for upstream tracks the particle must be reconstructible as a VELO track and give at least 3 hits in TT;

- for downstream tracks the particle must be reconstructible as a T track and give at least 3 hits in TT.

Similarly, “reconstructed” events are defined as those which satisfy [95]:

- for VELO or T tracks the track must have at least 70% of its associated hits originating from a single Monte Carlo particle;
- an upstream or downstream track must satisfy the $\geq 70\%$ unique association criterion and also have a correctly-assigned TT hit;
- long tracks must satisfy the $\geq 70\%$ unique association criterion and have both correctly identified VELO- and T-track segments.

The reconstruction efficiencies in Table 5.2 are calculated in two ways, as from the above definitions the reconstructed set for a given K_S^0 class is not necessarily a subset of the corresponding reconstructible set — for example an LL K_S^0 may be accidentally reconstructed as a DD K_S^0 , resulting in a transfer of entries between categories.

The strict efficiency in a given track category is the fraction of events identified as reconstructible in a given K_S^0 category which are reconstructed in that category. By comparison, what we have referred to in Table 5.2 as the “effective efficiency” is the ratio of the number of events reconstructed in a particular category to the number identified as reconstructible in any category. The “effective” efficiency may be the more useful figure in practice: although not strictly upper-bounded at 100%, it accounts for all reconstructed events including those which are reconstructed in the wrong class. The strict efficiency is rather pessimistic in this regard.

Note that there is an implicit loss of $\sim 85\%$ K_S^0 due to the low reconstructibility. While ~ 4 times more DD than LL tracks are reconstructible, the LL K_S^0 have a higher reconstruction efficiency because they have more associated hits and are hence better-defined for reconstruction.

B^+ reconstruction

As for the $B_d^0 \rightarrow K^+ \pi^-$ channel, an initial hard cut on the B^+ mass is made immediately after filtering the K_S^0 and π^+ properties and making an unconstrained vertex fit. Again

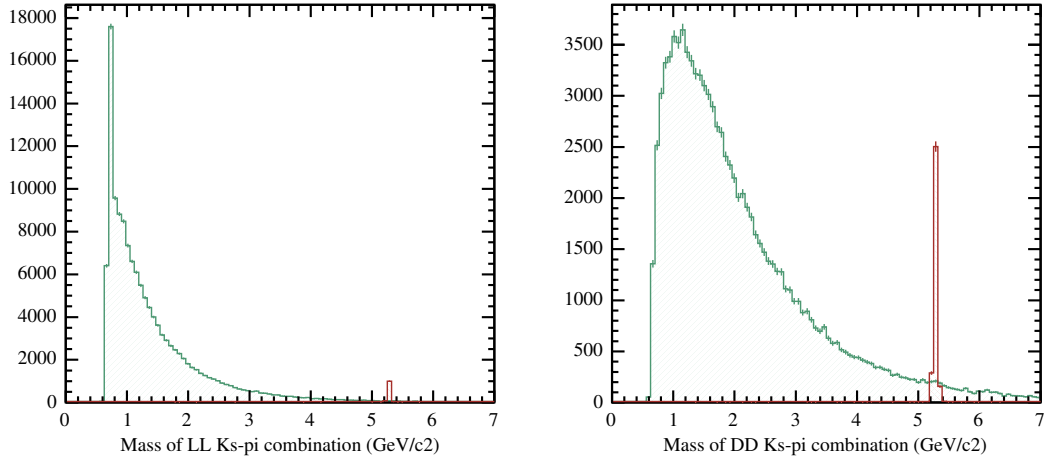


Figure 5.7: Masses of $K_S^0\pi^+$ combinations from 50k signal events before the first B^+ mass cut, using definitions of LL and DD as described in Section 5.4.2. The sharp B^+ mass peaks are again at ~ 5.2 GeV/ c : the LL set is more heavily affected by low mass combinations than the DD set as LL includes many particles from the PV. The vertical scale represents the number of combinations per bin.

the cut removes all candidates with masses lower than 4 GeV/ c to reduce the computational overhead from the huge combinatoric background peak in the invariant-mass as shown in Figure 5.7.

Figure 5.8 shows the distributions of particle ID confidence levels for the lone π^+ in signal events containing LL and DD K_S^0 . As for the pion particle ID confidence level in the $B_d^0 \rightarrow K^+ \pi^-$ decay, these distributions have peaks at both low and high confidence levels and as such are not useful for cuts: again, no restriction is placed on the pion particle ID confidence level. The charged pions used as input to the K_S^0 reconstruction will be considered separately in the next section.

5.4.3 Primary vertex selection

In general there may be several candidate primary vertices in an event, as displayed in Figure 5.9, and so a choice of the “best” one must be made using those quantities which depend on the position of the primary vertex. From the plots in Figures 5.10, it can be seen that a marginal preference can be found for the vertex candidates with the highest number of associated tracks: this is the criterion used to select the primary vertex if there is more than one candidate. A primary vertex is used for every event where the search algorithm finds one: the properties of the primary vertex are not used to reject

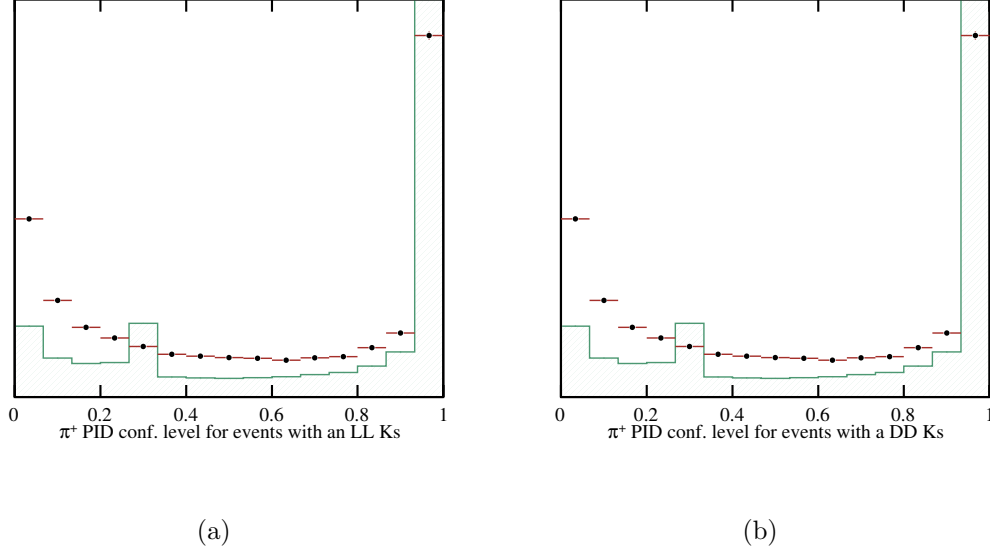


Figure 5.8: Particle ID confidence levels for LL and DD π^+ particle candidates in 54k $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ signal events. The filled distribution is from background candidates and the points are from true signal π^+ . The normalisation between signal and background is arbitrary.

events.

5.5 Preliminary event selection

The system described so far and implemented as a DaVinci algorithm was used to reconstruct candidate decays, as described by a set of characteristic variables, and to perform some simple data reduction. The data reduction is particularly important when considering the size of inclusive $b\bar{b}$ background event samples to be studied: $\sim 10^7$ events are available and iterative tuning of selection cuts cannot be applied to such a large sample. By placing emphasis on a very loose set of initial cuts to maintain high efficiencies where possible, this huge sample can be reduced to a more manageable one as a one-off procedure, with all subsequent cuts to be again considered as also including the pre-selection “stripping” cuts.

The $B^+ \rightarrow K_S^0 \pi^+$ selection algorithm is executed several times per event, each instantiation using a different set of cuts according to the class of K_S^0 involved. As already described, the selection aspects of the algorithms include three sets of iterations over

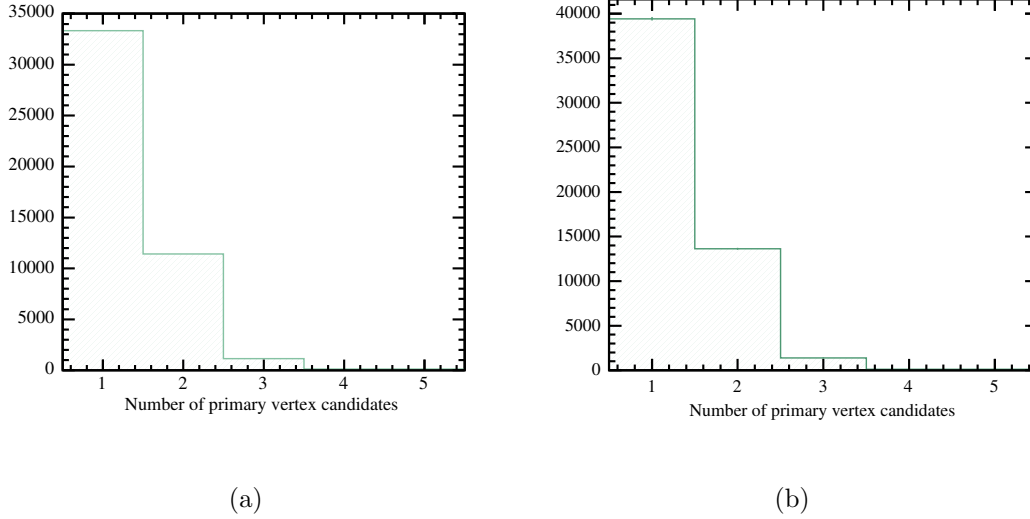


Figure 5.9: Number of events containing $n \in [0, 5]$ primary vertex candidates in (a) $\sim 46k$ $B_d^0 \rightarrow K^+ \pi^-$ signal events (b) $\sim 50k$ $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ signal events. There is no significant difference due to the signal types.

candidate particles: first the sets of K and π candidates are selected independently and then all pairwise combinations are filtered. It should also be noted that there are some implicit cuts such as the ability to form a K_S^0 or B vertex in the acceptance region of the vertexing tool and of course the track reconstructibility conditions.

The criteria for choosing the pre-selection cuts was that they should give a $b\bar{b}$ retention rate of less than 0.5%, while leaving reconstructible signal events unaffected. The pre-selection cuts for $B_d^0 \rightarrow K^+ \pi^-$ are in Table 5.3 and those for the $B^+ \rightarrow K_S^0 \pi^+$ channel, split by K_S^0 -type are listed in Table 5.4. The latter set are actually those used by the B2HH v1.5 DaVinci selection package, which was used to study generic B^0 decays to two charged light mesons: the division of some cuts into those to be applied to the K or π with the smaller or greater value of the cut variable is performed in their analysis but the distinction is not made again in the approach described here.

5.5.1 Producing data for optimisation

Data for cut optimisation is produced by applying the above pre-selection cuts to the signal data set as well as to the $b\bar{b}$ sample and the competing signal channels. Recon-

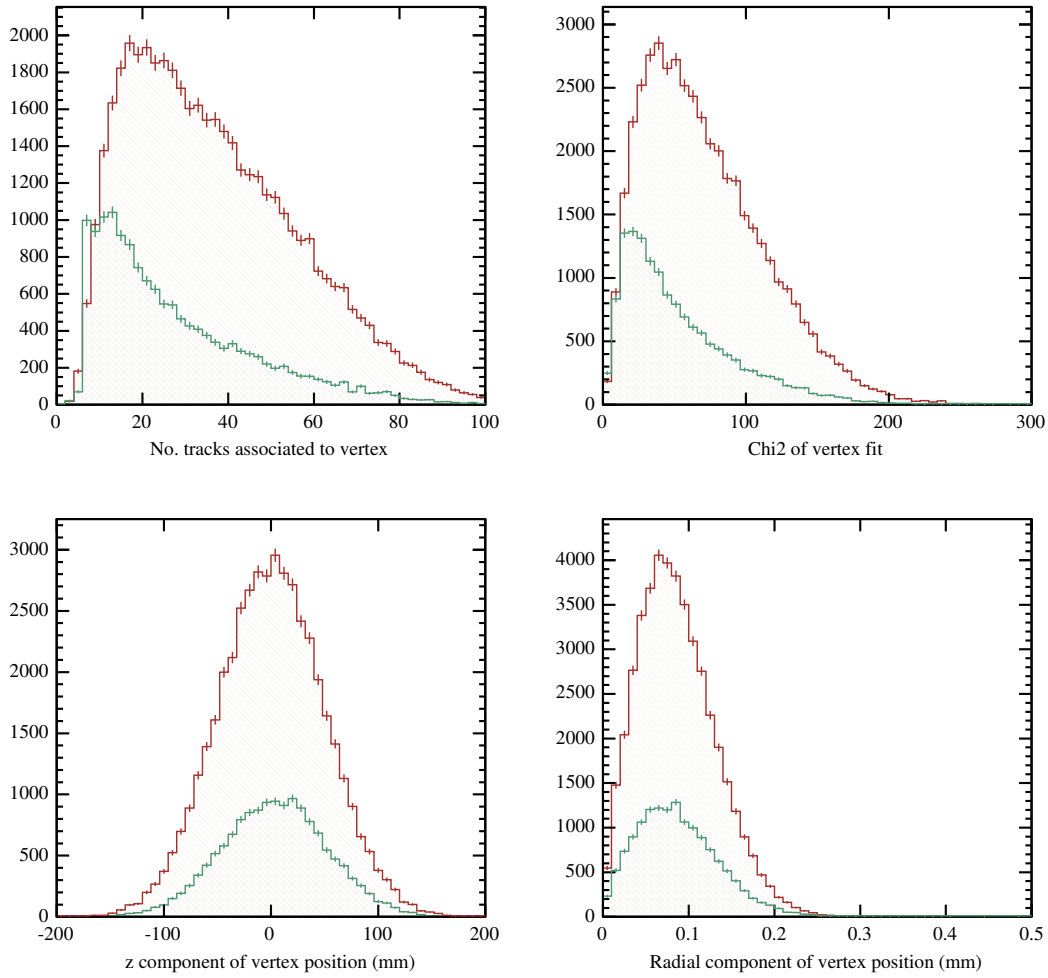


Figure 5.10: Properties of primary vertex candidates in 46k $B_d^0 \rightarrow K^+ \pi^-$ signal events. The upper distribution in each case contains the real primary interaction vertices; the lower distributions are false PVs. Almost identical distributions are obtained for the $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ events. The vertical scale represents the number of candidates per bin.

Cut variable	
$p(\pi)$	$< 300 \text{ GeV}/c$
$p_T(K/\pi \text{ with highest } p_T)$	$> 1000 \text{ MeV}/c$
$IP/\sigma_{IP}(K/\pi \text{ with smaller } IP/\sigma_{IP})$	> 2
$IP/\sigma_{IP}(K/\pi \text{ with larger } IP/\sigma_{IP})$	> 3
$IP/\sigma_{IP}(B_d^0)$	< 6
$l/\sigma_l(B_d^0)$	> 3.5
$\chi^2(B_d^0)$	< 20
$ m - m_{B_d^0} $	$< 2000 \text{ MeV}/c^2$

Table 5.3: Cuts used for $B_d^0 \rightarrow K^+ \pi^-$ pre-selection, given in terms of the condition required to pass the cut.

Cut variable	DD	LL
Main K_S^0 cuts	See Table 5.1	See Table 5.1
$ m - m_{K_S^0} $	$< 50.0 \text{ MeV}/c^2$	$< 23.0 \text{ MeV}/c^2$
$ m - m_{B^+} $	$< 150.0 \text{ MeV}/c^2$	$< 200.0 \text{ MeV}/c^2$

Table 5.4: Cuts used for $B^+ \rightarrow K_S^0(\pi^+\pi^-)\pi^+$ pre-selection, given in terms of the condition required to pass the cut.

structured properties for each cut set are written to an analysis n -tuple based on the ROOT [177, 178] TTree and derived classes. By this stage much of the low-mass combinatoric peak has been removed by the individual π and K selection loops. It can be seen in Figure 5.7 that for the $B^+ \rightarrow K_S^0 \pi^+$ channel the LL K_S^0 combinations are subject to more low-mass background combinations than the DD combinations are, since they include all charged tracks from the primary vertex.

After this loose mass cut, the set of properties listed in Tables 5.5 and 5.6 are calculated for all combinations which pass the loose B-mass cut and are stored in the n -tuples. Notably, properties such as $l/\sigma_l(K_S^0)$, despite nominally being a K_S^0 property, is actually a cut on both daughters' properties since the lone π^+ is also required in order to reconstruct the B^+ vertex from which the K_S^0 decay vertex is displaced.

In addition to these variables, the MC truth, the event and run numbers and the L0 and L1 trigger results for the parent event are stored. Single events can be associated

Cut variable	Description
$IP/\sigma_{IP}(\pi)$	π^- impact parameter significance w.r.t. primary vertex
$p(\pi)$	π^- momentum
$p_T(\pi)$	π^- transverse momentum w.r.t. detector z -axis
$IP/\sigma_{IP}(K)$	K^+ impact parameter significance w.r.t. primary vertex
$p(K)$	K^+ momentum
$p_T(K)$	K^+ transverse momentum w.r.t. detector z -axis
$\chi^2(B_d^0)$	B_d^0 vertex-fit χ^2
τ_B	B_d^0 lifetime
$l/\sigma_l(B_d^0)$	B_d^0 displacement significance from primary vertex
$IP/\sigma_{IP}(B_d^0)$	B_d^0 impact parameter significance w.r.t. primary vertex
m_B	B_d^0 mass
$p(B_d^0)$	B_d^0 momentum
$p_T(B_d^0)$	B_d^0 transverse momentum w.r.t. detector z -axis

Table 5.5: Cut variables retrieved from $B_d^0 \rightarrow K^+ \pi^-$ reconstruction for use in cut optimisation. There are a total of 13 $B_d^0 \rightarrow K^+ \pi^-$ variables.

with multiple B candidates in the n -tuples. The MC truth is determined via the DaVinci “associator” system, which uses loose links between the reconstructed and Monte Carlo elements of the analysis framework to decide if a good association between MC and reconstruction objects exists. The definition used is uniformly that to be considered “true” a particle must really be reconstructed as the right type and must occupy the expected place in a signal decay chain. Explicitly, a real K_S^0 which comes from the primary pp collision rather than from the B -decay will be considered as false even if the reconstruction declares that it is a K_S^0 .

5.5.2 Cut variable distributions

Figures 5.11–5.13 illustrate the raw cut variables as recorded for the $\sim 46k$ $B_d^0 \rightarrow K^+ \pi^-$ decays and Figures 5.14–5.16 show the variables for both the LL and DD classes of reconstructed $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ decays from the $\sim 54k$ signal set. In both sets of plots the $\sim 10^7$ inclusive $b\bar{b}$ set is also included: the backgrounds shown in the plots are the sum of inclusive $b\bar{b}$ events and the combinatoric background from signal events, combined with appropriate relative weightings to be discussed in Section 5.6.2. The background

Cut variable	Description
$IP/\sigma_{IP}(\pi)$	Lone π^+ impact parameter significance w.r.t. primary vertex
$p(\pi)$	Lone π^+ momentum
$p_T(\pi)$	Lone π^+ transverse momentum w.r.t. detector z -axis
$IP/\sigma_{IP}(K_S^0)$	K_S^0 impact parameter significance w.r.t. primary vertex
m_K	K_S^0 mass
$p(K_S^0)$	K_S^0 momentum
$p_T(K_S^0)$	K_S^0 transverse momentum w.r.t. detector z -axis
$\chi^2(K_S^0)$	K_S^0 vertex-fit χ^2
τ_K	K_S^0 lifetime
$z_{\text{dec}}(K_S^0)$	K_S^0 decay vertex position z component
$l/\sigma_l(K_S^0)$	K_S^0 displacement significance from B^+ decay vertex
$\chi^2(B^+)$	B^+ vertex-fit χ^2
τ_B	B^+ lifetime
$l/\sigma_l(B^+)$	B^+ displacement significance from primary vertex
$IP/\sigma_{IP}(B^+)$	B^+ impact parameter significance w.r.t. primary vertex
m_B	B^+ mass
$p(B^+)$	B^+ momentum
$p_T(B^+)$	B^+ transverse momentum w.r.t. detector z -axis

Table 5.6: Cut variables retrieved from $B^+ \rightarrow K_S^0 \pi^+$ reconstruction for use in cut optimisation. There are a total of 18 $B^+ \rightarrow K_S^0 \pi^+$ variables.

candidates are represented by the filled histogram and the signal set by the points. As the backgrounds are massive compared to the signal, the signal and background histograms have been independently normalised with arbitrary vertical scales.

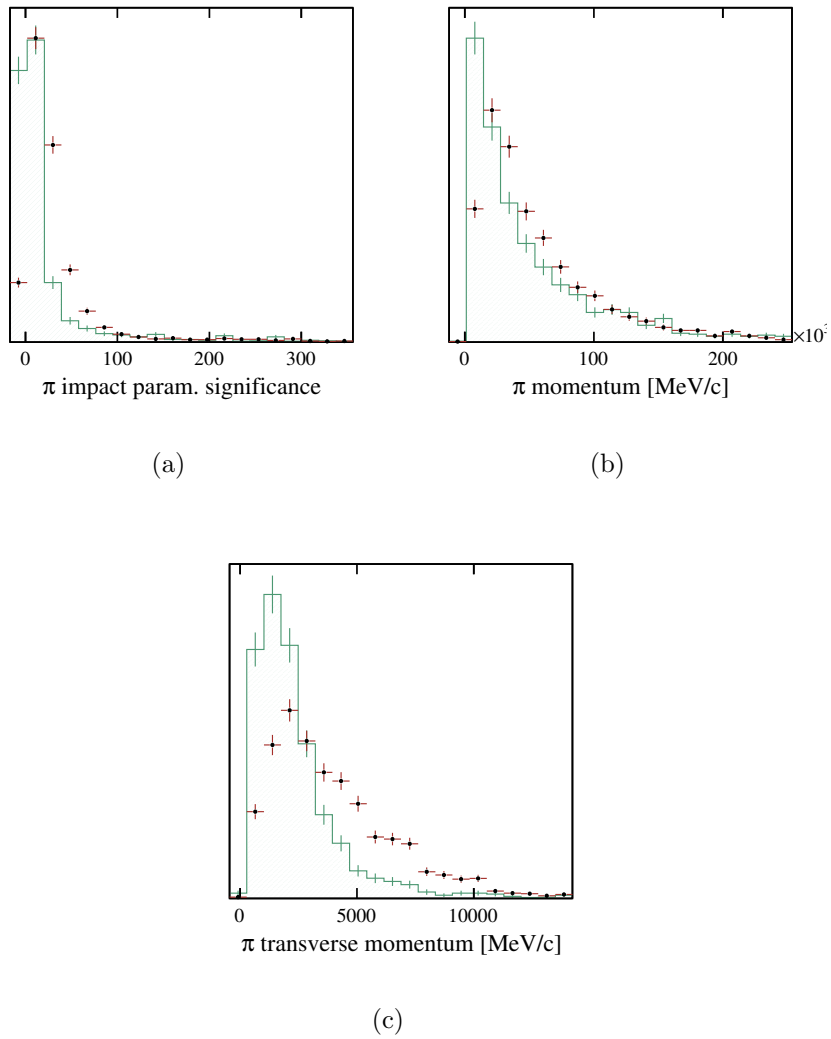


Figure 5.11: π^- cut variable distributions for $B_d^0 \rightarrow K^+ \pi^-$. The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.

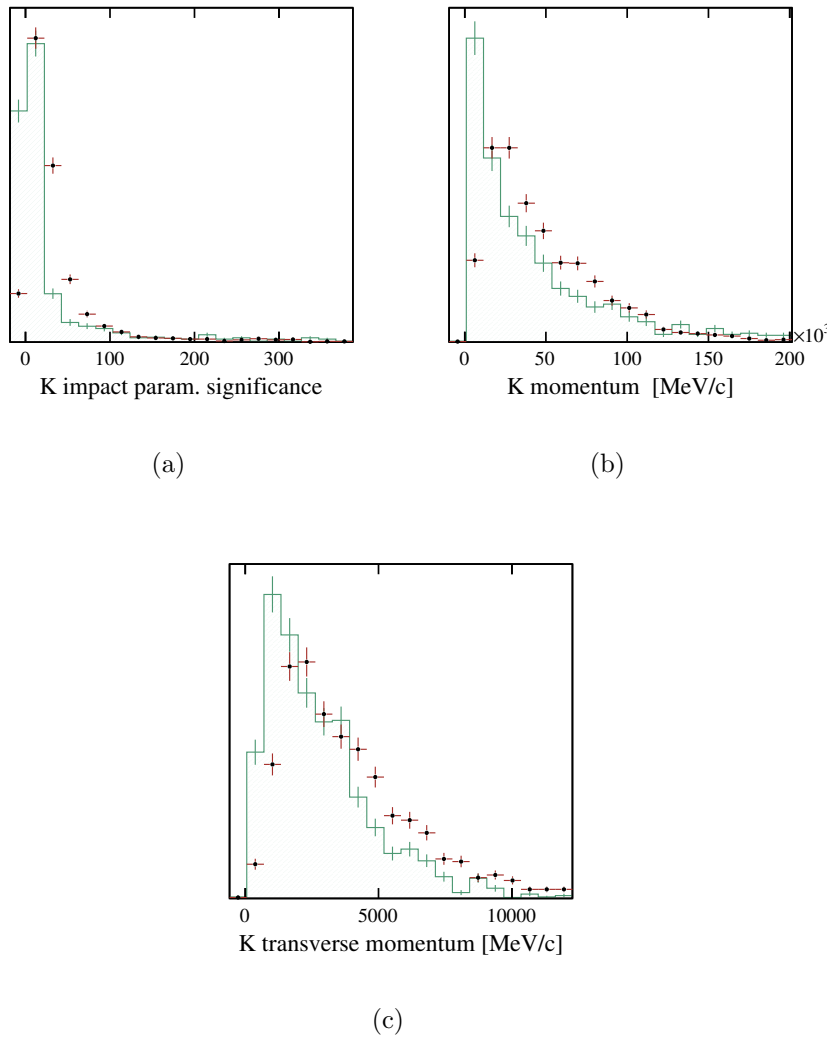


Figure 5.12: K^+ variable distributions for $B_d^0 \rightarrow K^+ \pi^-$. The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.

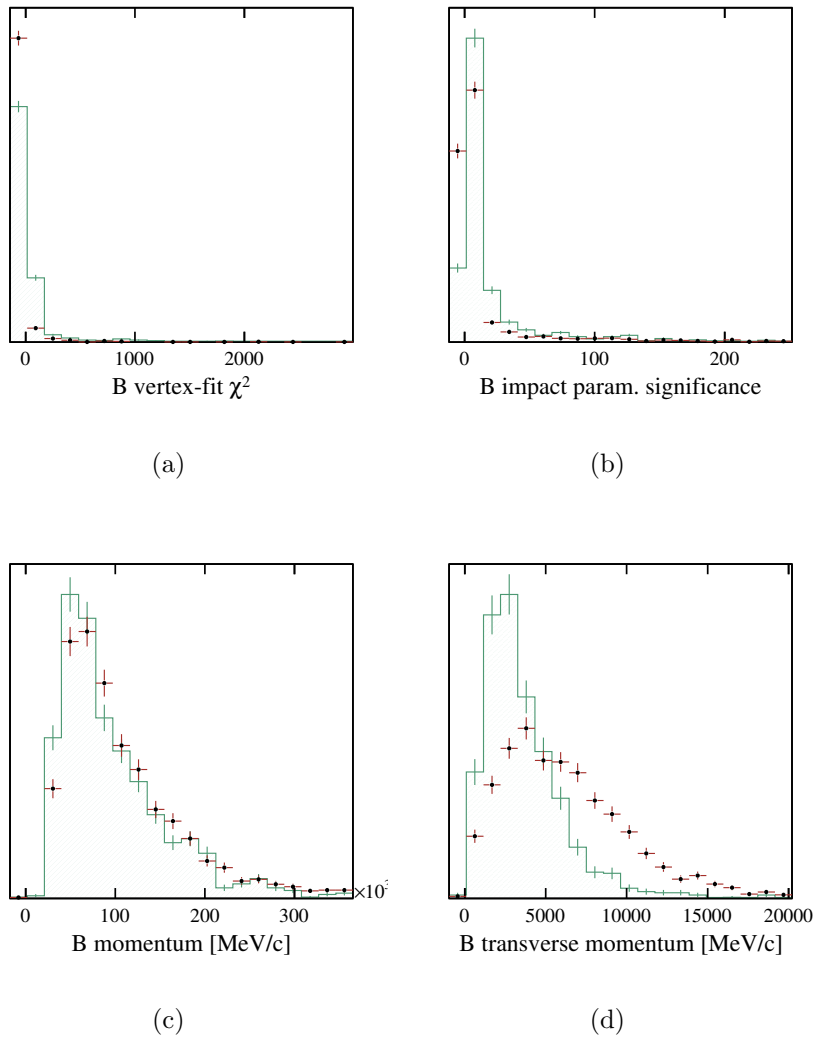


Figure 5.13: B_d^0 cut variable distributions for $B_d^0 \rightarrow K^+ \pi^-$. The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.

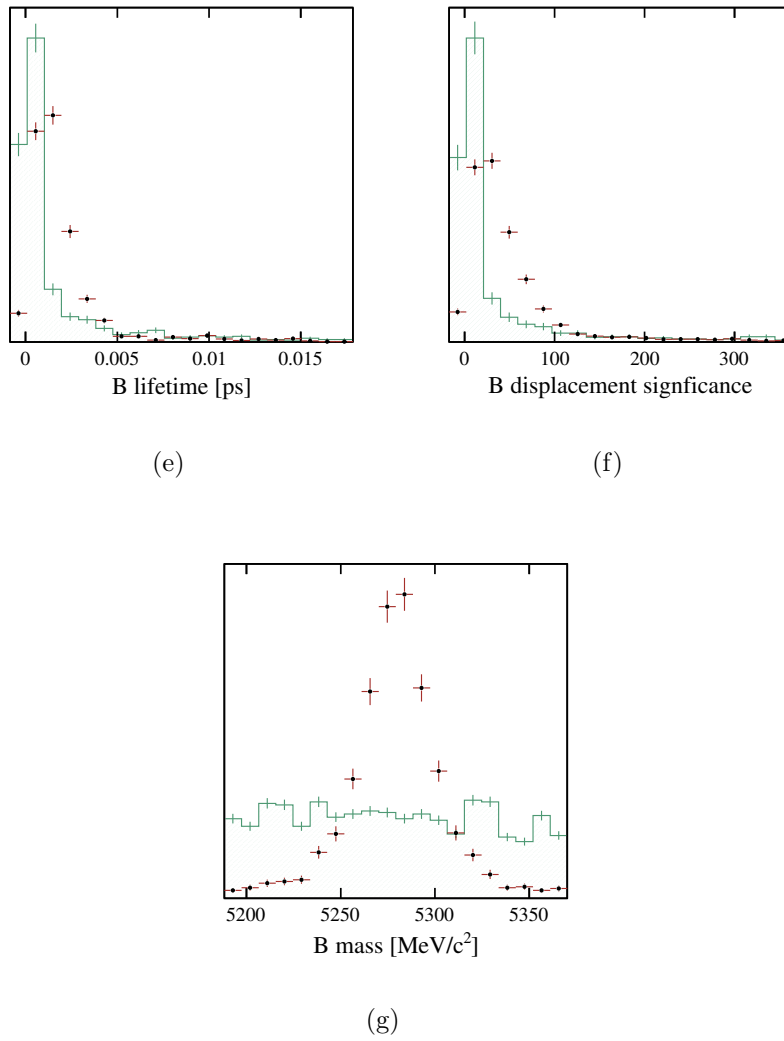


Figure 5.13: B_d^0 cut variable distributions for $B_d^0 \rightarrow K^+ \pi^-$ (contd.). The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.

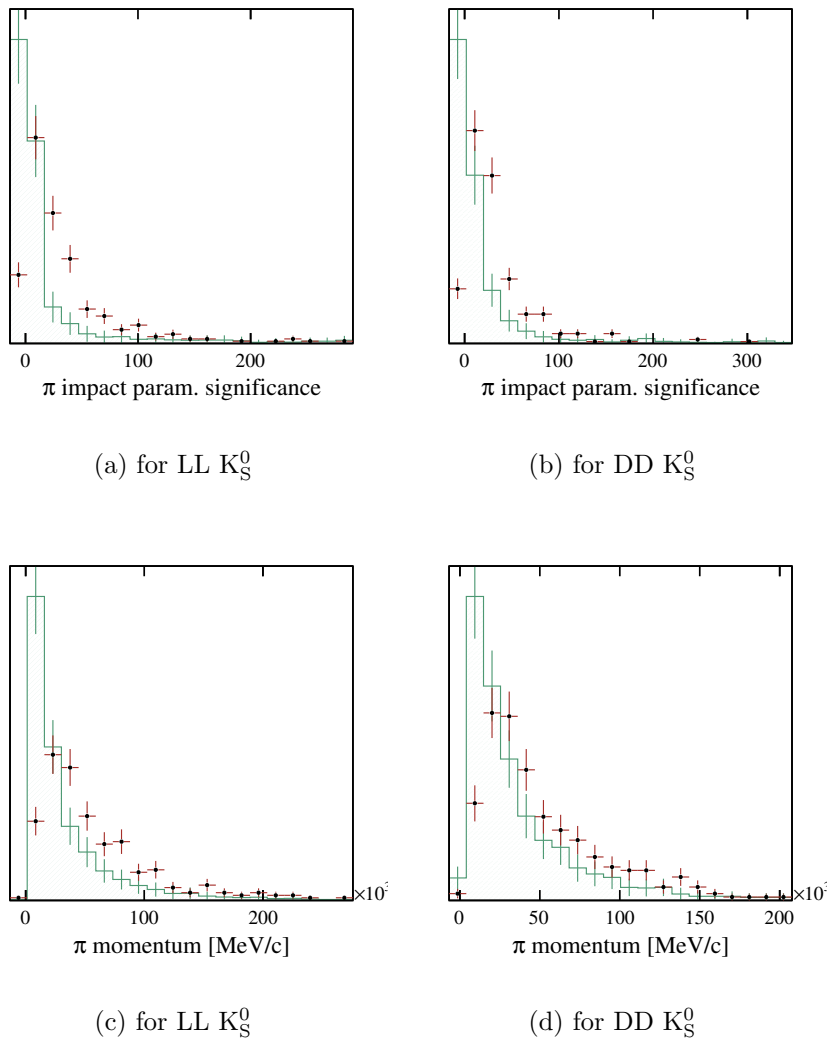


Figure 5.14: Lone π^+ cut variable distributions for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$. The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.

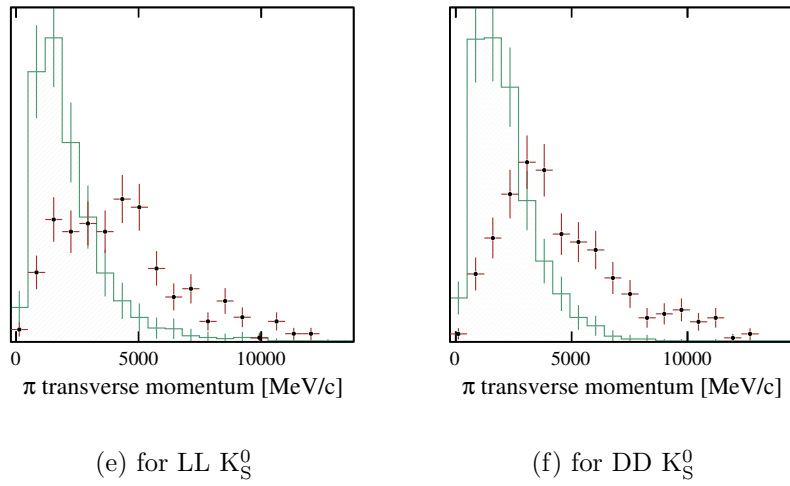


Figure 5.14: Lone π^+ cut variable distributions for $B^+ \rightarrow K_S^0(\pi^+\pi^-)\pi^+$ (contd.). The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.

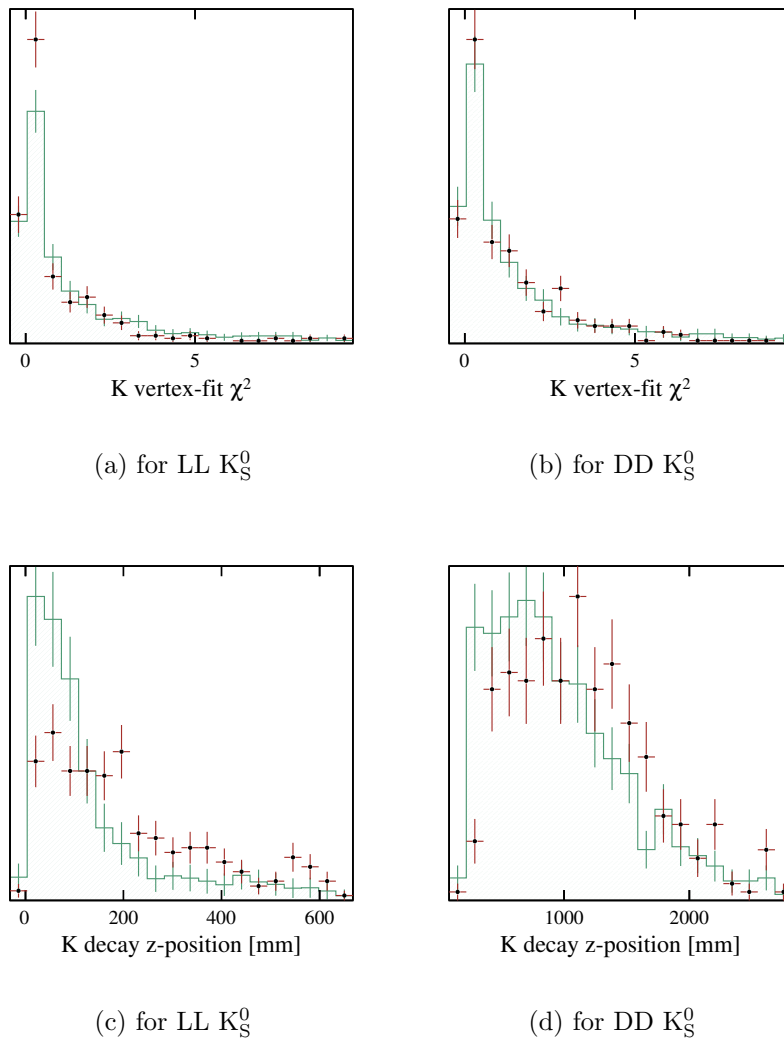


Figure 5.15: K_S^0 cut variable distributions for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$. The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.

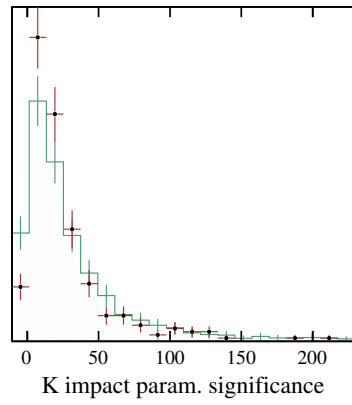
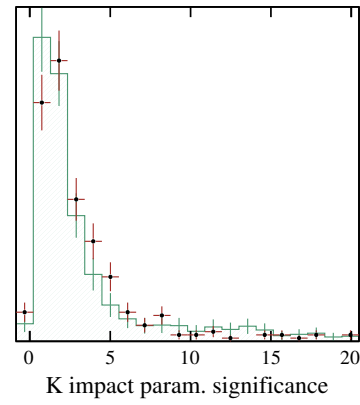
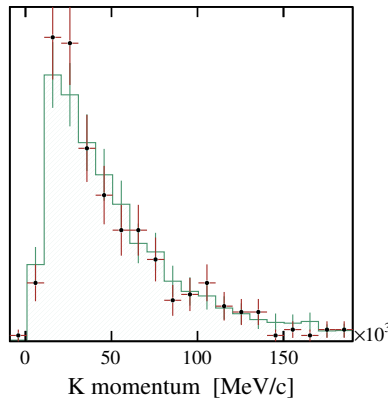
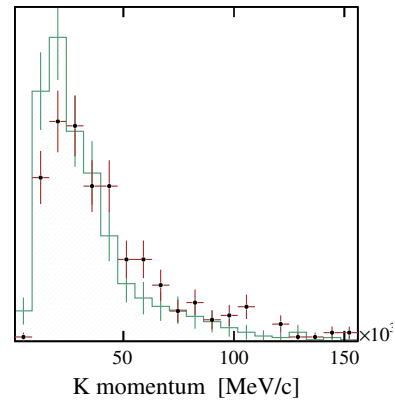
(e) for LL K_S^0 (f) for DD K_S^0 (g) for LL K_S^0 (h) for DD K_S^0

Figure 5.15: K_S^0 cut variable distributions for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ (contd.). The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.

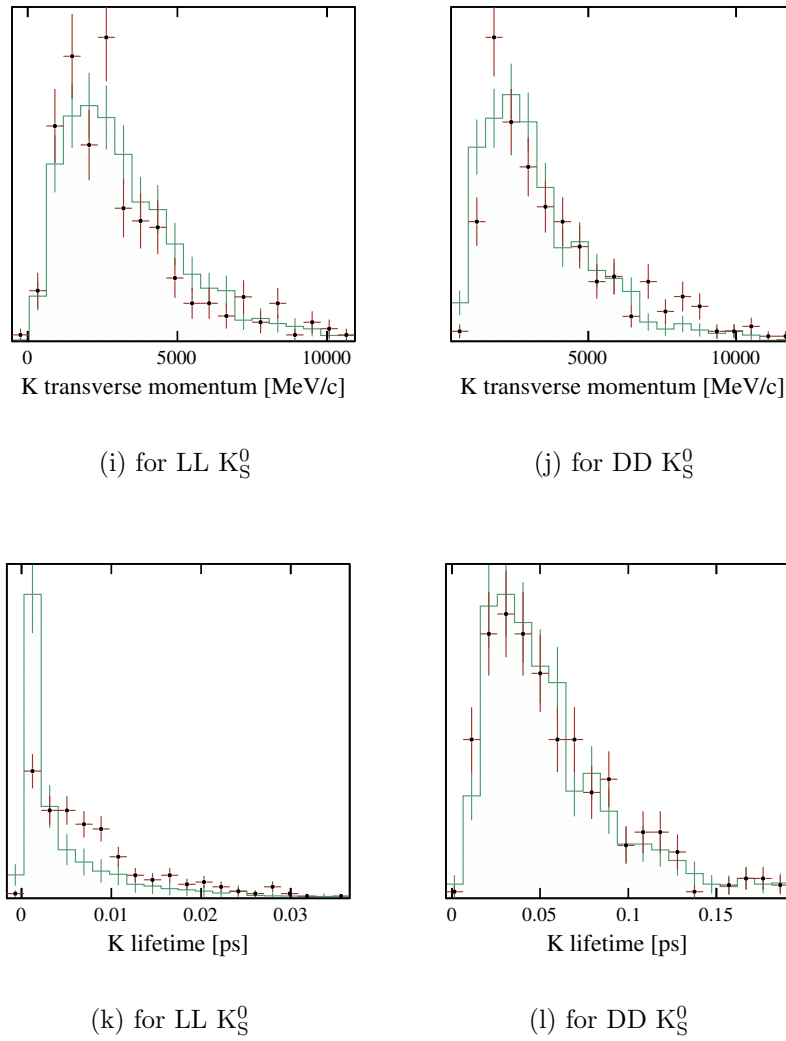


Figure 5.15: K_S^0 cut variable distributions for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ (contd.). The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.

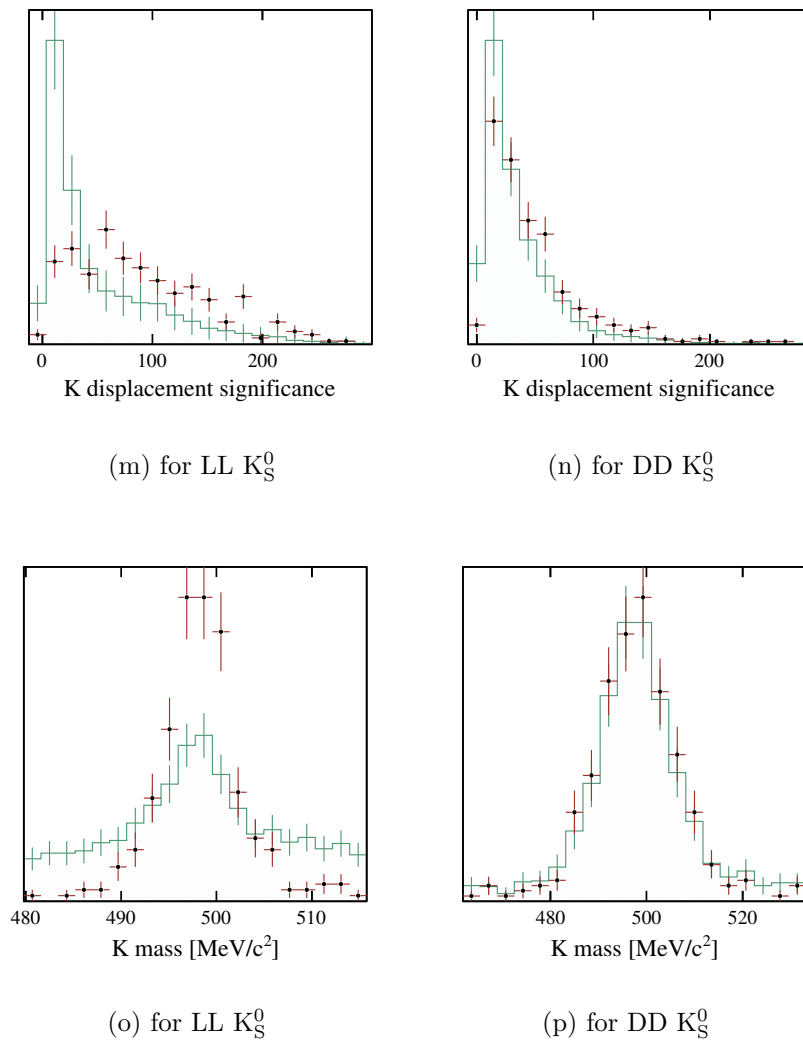


Figure 5.15: K_S^0 cut variable distributions for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ (contd.). The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.

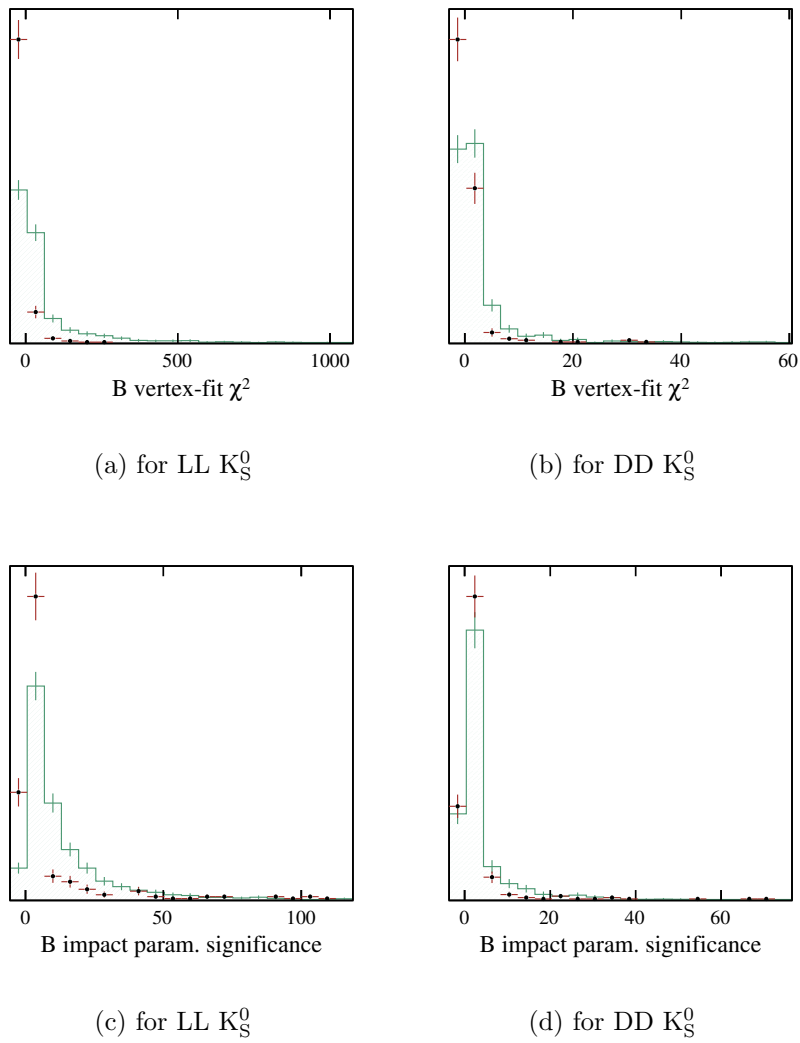


Figure 5.16: B^+ cut variable distributions for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$. The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.

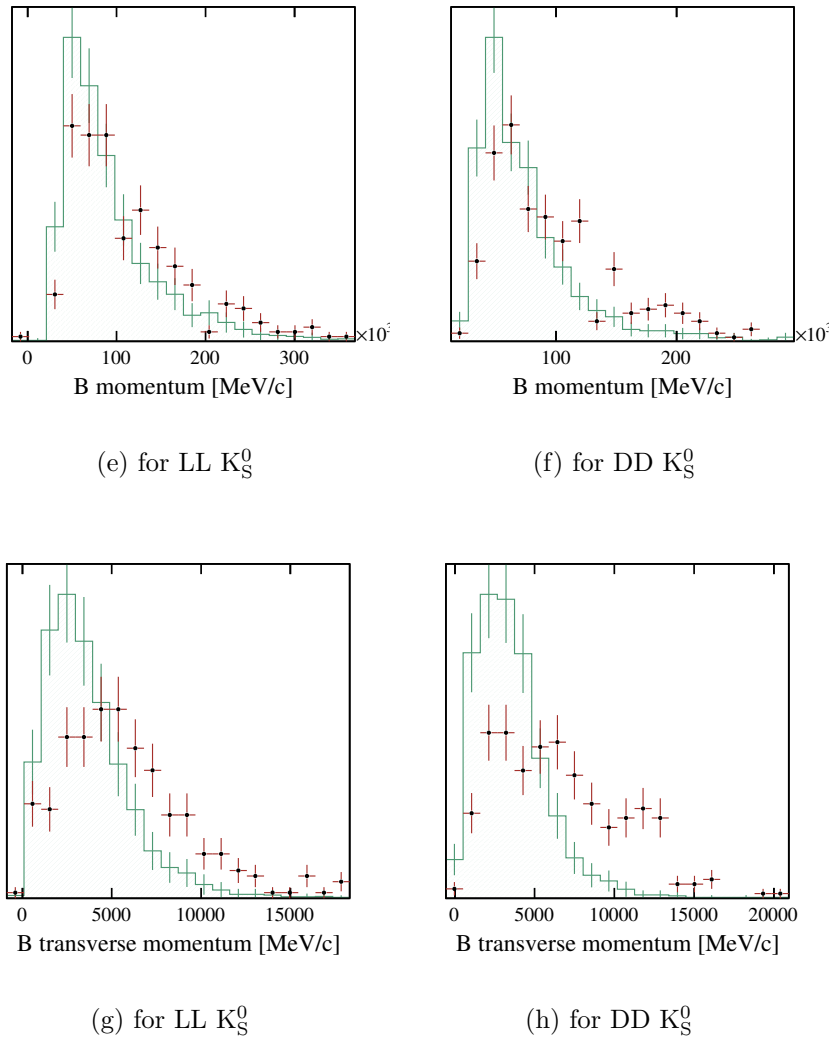


Figure 5.16: B^+ cut variable distributions for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ (contd.). The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.

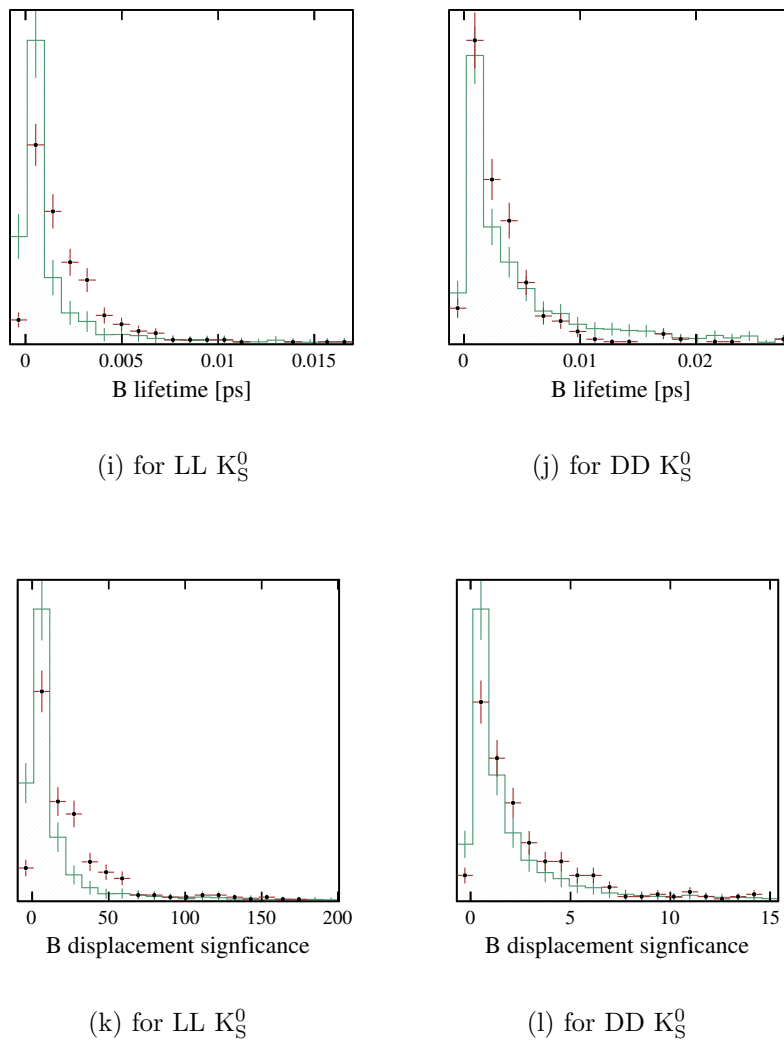


Figure 5.16: B^+ cut variable distributions for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ (contd.). The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.

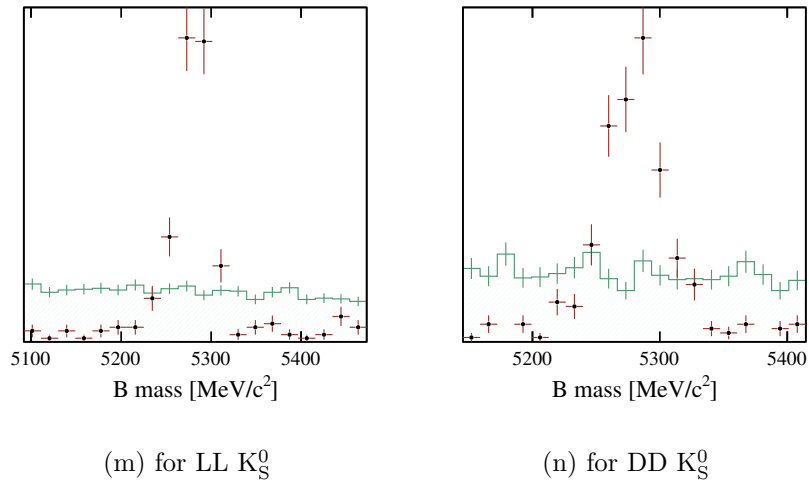


Figure 5.16: B^+ cut variable distributions for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ (contd.). The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.

5.6 Optimising the selection cuts

The optimisation of event selection by varying cut parameters is a computational problem of great depth and many methods exist for approaching it. It is clearly agreed that the *ideal* result of applying a set of selection cuts to a data set is that all false event candidates are rejected and all true candidates are accepted, but how to define “closeness” to this scenario is less absolute: it depends on what we want our data to be used for. Additionally, how to maximise this concept of closeness to the usually unobtainable ideal is a significant computational problem.

To optimise a selection process it is useful to formally visualise the problem as one of trying to optimally separate (or *discriminate*, the preferred terminology) two populations in an n -dimensional vector space with dimensions corresponding to all input parameters. Two immediate consequences of this formulation are to ask if the given set of parameters actually span the space in which the signal and background populations are embedded and to ask whether a clustering structure can be found which optimally separates the two populations given the supplied set of variables.

A selection of methods are designed to answer the second of these questions [179]. These include Monte Carlo methods, kernel methods [137], hierarchical and K-means clustering, neural nets [180] and a variety of eigenvector methods such as principal and independent component analysis (PCA/ICA) [181, 137], Fisher discriminants and quadratic discriminant analysis (QDA) [182], as well as the brute force “grid scan” approach. The most naïve approach is that of applying “selection cuts”, which remove data points failing to be either greater or less than a specified number in a characteristic variable. Although the use of cuts is embedded in the high-energy physics consciousness it is often non-optimal [179], relying on cut tuning by eye and failing to account for correlations between variables. Additionally, the variables to be “cut” are usually the physical parameters such as impact parameters, invariant-masses, momenta etc.—diagonalisation approaches based on eigenvector analysis of the distributions can improve this by building a new set of variables which minimise the correlations. The benefit of cut approaches, however, is that they are easily manipulated to provide information about performance under systematic uncertainties, something which is notoriously hard to extract from “black box” methods like neural nets.

The method to be described uses naïve cuts on physical variables but the tuning is performed by a Monte Carlo sampler which explores the space of cut parameters.

This is conceptually simple in that it performs the same rôle as a grid scan would but because the sampling is more informed as to the structure of the function, sampling time is concentrated in regions of near-maximal signal significance. As a result the computational time is much reduced relative to a comparable grid scan and the number of variables can be correspondingly increased. More cut variables increases the potential for spanning the distribution space and the background reduction can be achieved by making many small cuts rather than a small number of large ones. This allows the cuts to be chosen to impact less on signal distributions and so increases the signal efficiency for a given level of background suppression.

The framework used for this approach is a set of C++ classes implementing analysis concepts such as a `Dataset`, `Candidate`, `Cutset` and so on. Containers such as the `Dataset` provide methods to calculate basic statistical properties of the variable sets which define their collections of `Candidates`. Multi-dimensional minimisation and eigen-system routines are incorporated via the C-based GNU Scientific Library (GSL) [183].

5.6.1 Characterising selection performance

In this analysis, a characteristic function \mathcal{F} (data | cuts) is used as a measure of selection performance. The form of \mathcal{F} should ideally be matched to the phenomenological requirements of the study; however, it is usual to apply a much simpler and well-understood functional form instead. A variety of common characteristic functions are shown in Table 5.7, where the conventional extremisation sense is taken to be maximisation. The subscripts “tot” and “sel” indicate if the number of candidates are counted respectively before or after selection.

Form of \mathcal{F}	Motivation
$S_{\text{sel}}/B_{\text{sel}}$	Simple maximisation of signal to background ratio
$S_{\text{sel}}/\sqrt{B_{\text{sel}}}$	Maximisation of signal significance
$S_{\text{sel}}/\sqrt{S_{\text{sel}} + B_{\text{sel}}}$	Maximisation of signal significance for non-discovery
$S_{\text{sel}}^2/S_{\text{tot}}(S_{\text{sel}} + B_{\text{sel}})$	Maximisation of purity \times efficiency

Table 5.7: Examples of functions used to characterise selection performance.

The main characteristic function of interest is that of the signal significance: the number of observed events divided by the expected uncertainty [184, 185]. In “discovery”

measurements, where the existence of the signal is unconfirmed, the appropriate form of the significance function is $\tilde{\mathcal{F}} = S/\sqrt{B}$ where S is the number of selected signal events and B is the number of selected background. However, for systems such as $B \rightarrow K \pi$ selection where the existence of the signal is already known, the uncertainty must also incorporate the expected signal population and so the appropriate form is $\mathcal{F} = S/\sqrt{S+B}$ [186].

In practice, we used this function in a regularised form, $\mathcal{F} = S/\sqrt{S+B+\epsilon}$, where ϵ is a regularisation constant which suppresses the divergence of \mathcal{F} . This can occur in regions of parameter space where cuts are so tight that $\{S, B\} \rightarrow 0$. In the implementation, we set $\epsilon = 1$, justified by the observations that it is the fundamental scale in the problem and that if the statistics are severely affected by the addition of a single event then they are anyway unreliable. This regularised form proves to be a useful addition but we will nevertheless refer to the function without the ϵ term for convenience and familiarity.

A final point is to note that the signal significance differs from maximising the product of purity and efficiency (the final term in Table 5.7) in that there is an *a priori* reason for its comparative weightings of background and signal. The purity/efficiency product is often used as a visual measure of the optimality of a tuning but this fails to recognise that for a given task (say, signal discrimination) different levels of importance may be attached to the purity and the efficiency measures, skewing the distribution. For this reason, we will not consider the purity/efficiency product further as a characteristic function.

5.6.2 Preliminary data processing

Before proceeding to run the optimising sampler on our data sets, there are several subtleties which must be considered, to do with accounting for the statistical mismatch between signal and inclusive $b\bar{b}$ data sets and how we are to define the data set to be tuned and what we are to regard as a signal candidate.

Weight factors

A tuned selection represents a trade-off between closing the cut windows to reduce the background fraction and opening them to increase the signal yield. The admixture of these effects is affected by the relative quantities of the signal and background and so

the numbers of input candidates to the tuning procedure must be weighted. The ratio of effective (post-weighting) numbers of $b\bar{b}$ inclusive and signal events must be equal to the $B^+ \rightarrow K_S^0 \pi^+$ branching ratio, multiplied by the hadronisation fraction from a b-quark to a B_d^0 or B^+ meson. The hadronisation fractions for various b-hadrons are listed in Table 5.8, giving a B_d^0/B^+ fraction of 39.7% [10].

Hadron	Fraction [%]
B_d^0, B^+	39.7 ± 1.0
B_s^0	10.7 ± 1.1
b-baryons	9.9 ± 1.7

Table 5.8: Hadronisation fractions from a b-quark to the different b-hadrons.

Taking $w_{\text{sig}} = 1$, the weighting factor $w_{b\bar{b}}$ for the $b\bar{b}$ inclusive events is given by:

$$w_{b\bar{b}} = \frac{S_{\text{tot}}}{B_{\text{tot}}} \times \frac{\varepsilon_{b\bar{b}}^{\text{geom}}}{\varepsilon_{\text{sig}}^{\text{geom}}} \times \frac{1}{2 \times f_B \times \mathcal{B}_{\text{vis}}}, \quad (5.2)$$

where S_{tot} and B_{tot} are the number of signal and background events in the uncut data sets: the original numbers of signal events are 45900 for the $B_d^0 \rightarrow K^+ \pi^-$ channel and 54500 for the $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ channel, while the number of inclusive $b\bar{b}$ background is 10^7 for both signal channels. $\varepsilon_{b\bar{b}, \text{sig}}^{\text{geom}}$ are the geometric efficiencies (43.2% and 34.7% for $b\bar{b}$ and signal respectively), f_B is the hadronisation fraction and \mathcal{B}_{vis} is the visible branching ratio for the signal decay. The factor of 2 arises since the signal decay data sets do not contain any conjugate $\bar{B}_d^0 \rightarrow K^- \pi^+$ or $B^- \rightarrow K_S^0 \pi^-$ decays, so the number of selected B decays actually represents the total yield of B *and* \bar{B} decays. In this case both the b and the \bar{b} quark can hadronise and decay in signal channels and, ignoring the highly suppressed double-signal contribution, this is accounted for by a relative halving in $b\bar{b}$ weighting.

Given the input numbers above,

$$w_{b\bar{b}}(B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+) \sim 1150, \text{ and} \quad (5.3)$$

$$w_{b\bar{b}}(B_d^0 \rightarrow K^+ \pi^-) \sim 500. \quad (5.4)$$

This is problematic, as it tells us that our inclusive background sample is statistically inadequate by a large factor and will introduce strongly stepped features into the char-

acteristic function: whether or not a single $b\bar{b}$ event falls within the cut set will shift B_{sel} by either 1150 or 500 depending on the channel being analysed.

In practice, the main effect of this over-weighting is to make the convergence of the sampler less smooth and to effectively insist that all $b\bar{b}$ inclusive candidates be removed from the sample by the cuts. This is equivalent, then, to the established procedure of an initial cut tuning which removes all $b\bar{b}$ events and subsequent fine tuning within the boundaries established by that initial cut set. By removing all over-weighted $b\bar{b}$ events we are then in a region of no sensitivity to the $b\bar{b}$ distributions with adequate statistics and so the best we can do is to assume that our cuts constitute a hyper-surface which is just on the limit of enclosing another $b\bar{b}$ candidate and set a Poisson limit accordingly.

It is also worth noting that characteristic functions of the form $\mathcal{F} \sim S/\sqrt{B}$ scale more smoothly with background weighting than those with linear denominators and so we expect these functions to behave themselves better in the sampler than linear-denominator functions when background samples are heavily weighted.

Defining “signal” candidates

We need to be careful to remove duplicates from each event after selection, since only one candidate per event can really be true (here we ignore the tiny probability for multiple reconstructed signal decays per event). This “duplicate stripping” is used before tuning to set the representative candidate for the event to be a randomly-chosen true candidate, or a random false candidate if no true candidates are found in the event. Note that the $B^+ \rightarrow K_S^0 \pi^+$ channel LL and DD tunings are performed separately and that this stripping will not remove duplicates where LL and DD candidates are found in the same event; this is a relatively unimportant correction. Duplicate stripping is only performed on the signal data set, to avoid weighting the $b\bar{b}$ set even more intensely. This means that the $b\bar{b}$ weighting is an over-estimate, which errs in the acceptable direction when we come to set bounds on the $b\bar{b}$ background. The inclusive $b\bar{b}$ weight factor is left conservatively uncorrected for this.

After removal of duplicates from the signal decay data set, there is a degree of freedom in what we regard as a signal candidate. As the phenomenology of this analysis depends purely on branching ratios rather than accurate reconstruction of signal event characteristics (e.g. , a $B-\bar{B}$ mixing analysis requires good proper time resolution on the B decay) every Monte Carlo *event* corresponds to a real $B \rightarrow K \pi$ decay, regardless of

whether or not it is correctly reconstructed. There is therefore potential for increasing the selection efficiency if we count every candidate from a signal event as “true”. Alternatively, we have more confidence in those candidates which are regarded as true because of genuinely reliable reconstruction and so both approaches are followed independently. We define the terms “exclusive signal selection” and “inclusive signal selection” (after duplicate event stripping) as

- an “exclusive signal selection” considers a reconstructed decay to be truly signal only if the DaVinci associator system is completely consistent with that decay tree in terms of particle IDs and relationships;
- an “inclusive signal selection” considers a reconstructed decay to be truly signal if it was generated in the signal data set, regardless of MC truth information from the associator system.

Use of trigger results

It is worth considering the rôle of the trigger when optimising selection cuts: the primary reason for the trigger’s existence is to reduce readout data rates to a manageable level and it is to be expected that the trigger will not be optimised for a given channel. In particular, the L1 trigger requires the presence of a fast charged particle, which for the $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ channel is only likely to be provided by the lone π^+ . A high attrition rate is to be expected in this case.

The more compatible the off-line selection cuts can be with the on-line triggering algorithm, the better the yield can be; this compatibility can be obtained by tuning only on those events which have passed the L0 and L1 triggers, as is anyway the case with the running experiment. Here there is a tacit assumption that the yet-undefined high level trigger (HLT) is 100% efficient. The actual attrition rate will be analysed when the HLT algorithm is finalised. For compatibility with the LHCb re-optimization TDR, as well as to get the benefit of higher statistics in the sampling, we have tuned the cuts both with and without pre-applying the triggers.

We should also note that by design the inclusive backgrounds are more heavily attenuated by triggering than the signal channels and so tuning after applying the triggers affects the effective $b\bar{b}$ weighting factor. If anything, the pre-application of the trigger to the $b\bar{b}$ set will worsen the risk of over-tuning, as the resulting tuning will use even fewer background events.

5.6.3 Monte Carlo sampling

After duplicate stripping, the appropriate modifications of truth records and weight factors and, optionally, the application of the L0 and L1 triggers, the new combined and weighted data set contains candidates which are simply marked as true or false, rather than discriminating between their origin. These distributions are then transformed to make the variables more homogeneous in extent by first shifting each by the mean of its true signal component and then scaling down by the standard deviation of the signal distribution. In practical terms this resulted in the symmetric signal distributions occupying a region $\in [-0.5, 0.5]$ while less symmetric distributions assumed tails out to $\pm \mathcal{O}(1)$.

The cut set was defined internally to the sampler as a set of paired low and high cuts. Hence, for tuning purposes each candidate variable was doubled, corresponding to selection windows in each input variable. Cuts were initially placed roughly “by eye” and the sampler was run.

The sampler used was an implementation of the Nelder-Mead “simplex” sampler from the GSL [187, 183]. This sampler constructs $n + 1$ points in the n -dimensional space (a simplex) and then samples in the space by repeated geometric transformations of the simplex points. The next step in the chain is taken based on the progression toward minimum of the “score”, here $-\mathcal{F}$, the negative of the characteristic function. The algorithm has the effect of “tumbling” the simplex into a minimum of the score function, shrinking in volume as it progresses. Primarily this sampler was chosen as it does not require gradient information from the minimising function, this being difficult to define in a robust fashion given the ad hoc transformation of the input variables described above. The only additional configuration choice required is a size for the first step, which was set to 0.2 after some trials with other values.

The convergence rate and stagnation point of the sampler is hard to predict (indeed, is an unsolved hard problem in its own right) and rather than attempt to use the simplex volume as a convergence criterion we use a fixed number of iterations. In practice, convergence tends to occur between 200 and 500 successful iterations and is characterised by a long-term alternation between two equivalent states. There is definitely a danger of over-tuning but with the limited statistics available the best way out of this problem is not to work harder on the sampler convergence problem but to tune and cut on different data sets. An example of the score convergence and the resulting distribution of purity vs. efficiency are shown in Figure 5.17. The score can be seen to suddenly

improve at about 100 iterations after removal of the heavily-weighted inclusive $b\bar{b}$ events. Thereafter two distinct classes of solution can be seen, one performing significantly better than the other. The two sets of solutions are an artifact of the simplex sampler. The purity–efficiency plot shows an obvious banding structure. This is the effect of the high inclusive $b\bar{b}$ weight factors restricting the available portions of the purity–efficiency space to discrete contours. The separation of these contours can be interpreted as giving some idea of the $b\bar{b}$ uncertainty on any given tuning.

Once convergence has been achieved, we are still left with a cut set of 26 parameters on 13 variables or 36 parameters on 18 variables for the $B_d^0 \rightarrow K^+ \pi^-$ or $B^+ \rightarrow K_S^0 \pi^+$ channels respectively. However, we really want a smaller core set of “good” cuts. This is obtained by systematically extending each of the tuned cuts to “infinity”¹ such that they do not exclude any candidates. The effect of this extension on the characteristic function is the observed; extensions which result in an improvement of the score are clearly “bad” cuts and their initial positions are set to a large value for a second tuning, while cuts whose extension results in no change or a degradation of performance are marked as “indifferent” or “good” respectively. After a second tuning, if no bad cuts are reported, the subset of “good” cuts is taken as the tuned selection cuts. Finally, a by-eye check of the final cut performance is performed by applying these cuts and the triggers to the data sets and checking the variable distributions to see if any obvious small shifts can be made to remove the last background events. In the cases where the cut-tuning is performed with the data before triggering this by-eye approach was used to fine-tune cut parameters which had not quite sampled to the optimum position, due to being “unaware” of the triggers’ effect.

5.7 Optimised cut sets

In this section the sets of cuts obtained from the sampler for the various tuning definitions are presented, after testing for bad and irrelevant cuts. It should be remembered that these are in addition to the pre-selection cuts already listed for each channel.

¹In reality, the extension will only be as far as any pre-selection cut on that variable. Since the pre-selection cuts were intentionally left loose, this should not impact on the reliability of the method.

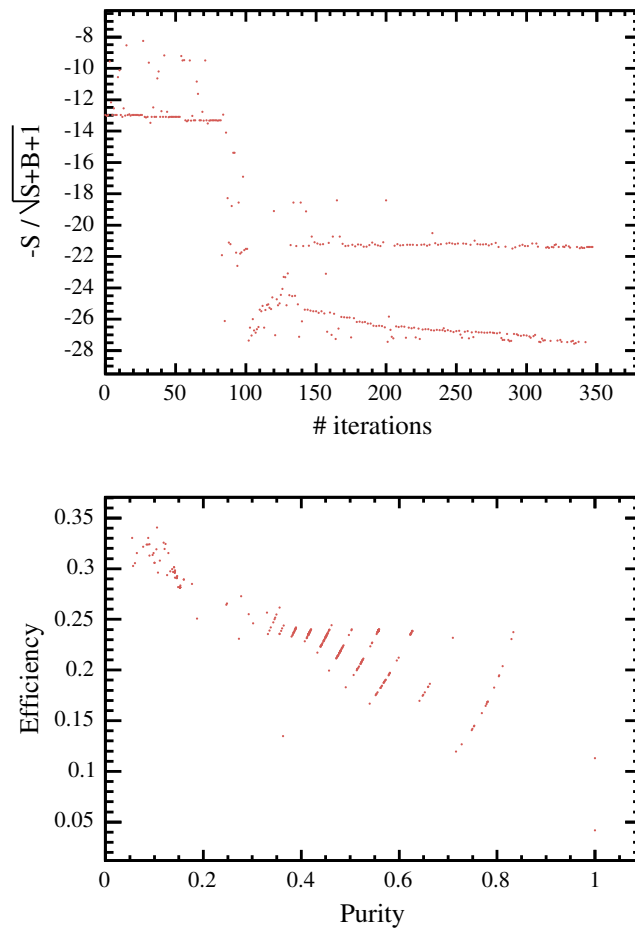


Figure 5.17: Score progression with iteration number and the purity-efficiency trade-off for the $B_d^0 \rightarrow K^+ \pi^-$ selection using the trigger in the tuning and the inclusive definition of signal.

5.7.1 Optimised cut sets for $B_d^0 \rightarrow K^+ \pi^-$

The “useful” subset of the 26 $B_d^0 \rightarrow K^+ \pi^-$ cuts for the different definitions of “signal” and the two uses of triggering in the tuning are shown in Tables 5.9, 5.10, 5.11 and 5.12. In general the forms of the cuts are sensible, requiring the decay daughters to have large impact parameters, their combination to have a small impact parameter, low χ^2 values for the K - π vertex fit and typically high (transverse) momenta. The construction of a two-sided cut on the B invariant-mass is an expected outcome of the tuning, which is used for almost all cut sets (although the trigger-aware $B_d^0 \rightarrow K^+ \pi^-$ selection only places a low cut on the B_d^0 mass). The only truly surprising cut is the high cut placed on p (K) in the tuning with an exclusive signal definition and no pre-triggering (Table 5.9). As we shall see, this tuning does not give particularly good performance anyway.

A final comment to make on the cut sets is that there is no particular pattern to be observed between the different tunings: it might be expected that $\pi^- p_T$ cuts would change between the no trigger and pre-triggered tunings but the reality is not so simple. Again, it should not be surprising that the naïve expectations do not play out: each tuning entirely re-defines the correlated structure of the signal and background populations in a 26-dimensional space and the differences between the tunings reflect this complexity.

Cut variable	Cut
$IP/\sigma_{IP}(\pi)$	> 5.86
$p_T(\pi)$	$> 1.99 \text{ GeV}/c$
$p(K)$	$< 134.36 \text{ GeV}/c$
$p_T(K)$	$> 1.67 \text{ GeV}/c$
$IP/\sigma_{IP}(B)$	< 2
$p(B)$	$> 25 \text{ GeV}/c$
$p_T(B)$	$> 0.93 \text{ GeV}/c$
m_B	$\in [5.23 \text{ GeV}/c, 5.33 \text{ GeV}/c]$

Table 5.9: $B_d^0 \rightarrow K^+ \pi^-$ tuned selection cuts for an optimisation tuning which is unaware of the effects of the L0 and L1 triggers and uses an exclusive definition of a “signal” event.

Cut variable	Cut
$p_T (\pi)$	$> 1.84 \text{ GeV}/c$
$p (K)$	$> 10.09 \text{ GeV}/c$
$p_T (K)$	$> 1.25 \text{ GeV}/c$
$IP/\sigma_{IP}(B)$	< 4.87
$p_T (B)$	$> 1.91 \text{ GeV}/c$
τ_B	$> 0.558 \text{ ps}$
m_B	$\in [5.25 \text{ GeV}/c, 5.31 \text{ GeV}/c]$

Table 5.10: $B_d^0 \rightarrow K^+ \pi^-$ tuned selection cuts for an optimisation tuning which is unaware of the effects of the L0 and L1 triggers and uses an inclusive definition of a “signal” event.

Cut variable	Cut
$IP/\sigma_{IP}(\pi)$	> 5.74
$p_T (\pi)$	$> 1.20 \text{ GeV}/c$
$p_T (K)$	$> 0.58 \text{ GeV}/c$
$IP/\sigma_{IP}(B)$	< 2.5
$p_T (B)$	$> 1.40 \text{ GeV}/c$
m_B	$> 5.23 \text{ GeV}/c$

Table 5.11: $B_d^0 \rightarrow K^+ \pi^-$ tuned selection cuts for an optimisation tuning which *is* aware of the effects of the L0 and L1 triggers and uses an exclusive definition of a “signal” event.

Cut variable	Cut
$p_T (\pi)$	$> 1.68 \text{ GeV}/c$
$p (K)$	$> 11.01 \text{ GeV}/c$
$p_T (K)$	$> 1.15 \text{ GeV}/c$
$IP/\sigma_{IP}(B)$	< 3.16
$p_T (B)$	$> 1.67 \text{ GeV}/c$
τ_B	$> 0.510 \text{ ps}$
m_B	$> 5.24 \text{ GeV}/c$

Table 5.12: $B_d^0 \rightarrow K^+ \pi^-$ tuned selection cuts for an optimisation tuning which *is* aware of the effects of the L0 and L1 triggers and uses an inclusive definition of a “signal” event.

5.7.2 Optimised cut sets for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$

The “useful” subset of the 36 $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ cuts are shown in Tables 5.13, 5.14, 5.15 and 5.16. Again, the forms of the cuts are largely sensible, given the geometry of the signal decay and typical requirements like high momenta, displacement significances and lifetimes for the B^+ candidates and the usual mass windowing. Notably, only one K_S^0 mass cut is made (for the LL cuts in Table 5.14), indicating that a future study should perhaps relax the K_S^0 reconstruction cuts before tuning the final selection cuts. Again, there is no simple relation between the tunings and the resultant cuts but this is of little significance.

Cut variable	LL cut	DD cut
$p_T (\pi)$	$> 3.48 \text{ GeV}/c$	$> 2.74 \text{ GeV}/c$
$p (K)$	$> 19.22 \text{ GeV}/c$	—
$p_T (K)$	$> 1.71 \text{ GeV}/c$	$> 1.50 \text{ GeV}/c$
$l/\sigma_l(K)$	> 13.3	—
$\chi^2 (B)$	< 3.46	—
$IP/\sigma_{IP}(B)$	< 58.8	< 82.1
$p_T (B)$	$> 7.82 \text{ GeV}/c$	$> 13.56 \text{ GeV}/c$
m_B	$\in [5.17 \text{ GeV}/c, 5.39 \text{ GeV}/c]$	$< 5.38 \text{ GeV}/c$

Table 5.13: $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ tuned selection cuts for an optimisation tuning which is unaware of the effects of the L0 and L1 triggers and uses an exclusive definition of a “signal” event.

Cut variable	LL cut	DD cut
$IP/\sigma_{IP}(\pi)$	—	> 9.3
$p(\pi)$		$> 10.74 \text{ GeV}/c$
$p_T(\pi)$	$> 2.01 \text{ GeV}/c$	$> 2.40 \text{ GeV}/c$
$\chi^2(K)$	< 2.82	< 6.69
$IP/\sigma_{IP}(K)$	—	> 0.5
$p(K)$	$> 16.35 \text{ GeV}/c$	—
$p_T(K)$	$> 0.84 \text{ GeV}/c$	$> 1.51 \text{ GeV}/c$
$z_{\text{decay}}(K)$	$< 45 \text{ cm}$	—
m_K	$< 502 \text{ MeV}/c^2$	
$\chi^2(B)$	—	< 0.45
$IP/\sigma_{IP}(B)$	< 4.01	< 30
$p_T(B)$	$> 2.98 \text{ GeV}/c$	$> 3.00 \text{ GeV}/c$
$l/\sigma_l(B)$	—	> 0.87
τ_B	$> 0.267 \text{ ps}$	—
m_B	$\in [5.24 \text{ GeV}/c, 5.32 \text{ GeV}/c]$	$\in [5.25 \text{ GeV}/c, 5.31 \text{ GeV}/c]$

Table 5.14: $B^+ \rightarrow K_S^0(\pi^+\pi^-)\pi^+$ tuned selection cuts for an optimisation tuning which is unaware of the effects of the L0 and L1 triggers and uses an inclusive definition of a “signal” event.

Cut variable	LL cut	DD cut
$IP/\sigma_{IP}(\pi)$	—	> 2.6
$p_T(\pi)$	$> 2.40 \text{ GeV}/c$	$> 2.80 \text{ GeV}/c$
$p_T(K)$	$> 0.58 \text{ GeV}/c$	—
$z_{\text{decay}}(K)$	$> 19 \text{ mm}$	—
$l/\sigma_l(K)$	> 27	—
$\chi^2(B)$	< 2	< 3.1
$IP/\sigma_{IP}(B)$	—	< 94
$p_T(B)$	$> 2.70 \text{ GeV}/c$	$> 6.50 \text{ GeV}/c$
$l/\sigma_l(B)$	> 25.3	> 1.6
τ_B	$< 6.79 \text{ ps}$	—
m_B	$\in [5.18 \text{ GeV}/c, 5.39 \text{ GeV}/c]$	—

Table 5.15: $B^+ \rightarrow K_S^0(\pi^+\pi^-)\pi^+$ tuned selection cuts for an optimisation tuning which is aware of the effects of the L0 and L1 triggers and uses an exclusive definition of a “signal” event.

Cut variable	LL cut	DD cut
$IP/\sigma_{IP}(\pi)$	> 3.00	—
$p_T(\pi)$	$> 2.00 \text{ GeV}/c$	$> 3.72 \text{ GeV}/c$
$z_{\text{decay}}(K)$	—	$> 67 \text{ cm}$
$p_T(K)$	$> 0.91 \text{ GeV}/c$	$> 1.81 \text{ GeV}/c$
τ_K	$< 22 \text{ ps}$	$< 128 \text{ ps}$
$\chi^2(B)$	—	< 3.0
$IP/\sigma_{IP}(B)$	< 4.05	—
$p_T(B)$	$> 2.34 \text{ GeV}/c$	$> 2.31 \text{ GeV}/c$
τ_B	$> 0.238 \text{ ps}$	
m_B	$\in [5.24 \text{ GeV}/c, 5.32 \text{ GeV}/c]$	$\in [5.25 \text{ GeV}/c, 5.31 \text{ GeV}/c]$

Table 5.16: $B^+ \rightarrow K_S^0(\pi^+\pi^-)\pi^+$ tuned selection cuts for an optimisation tuning which *is* aware of the effects of the L0 and L1 triggers and uses an inclusive definition of a “signal” event.

5.8 Cut performance for $B_d^0 \rightarrow K^+ \pi^-$

The numbers of signal and background events to pass various stages of the $B_d^0 \rightarrow K^+ \pi^-$ selection procedure are listed in Table 5.17. We see that the tunings which used an exclusive signal definition pass 5 combinatoric background events in both the un-triggered and pre-triggered tuning configurations and no background events in the tunings with an inclusive signal definition. These 5 candidates represent signal events and can be included in the signal yield. In all cases the sampler found cut tunings which removed all inclusive $b\bar{b}$ background events.

The best signal yield is found for the tuning where the trigger is applied to the data set before tuning and where the inclusive definition of signal has been used. With this configuration, 824 signal events are selected, while less generous configurations go as low as 581 (including the 5 combinatoric signal candidates as signal events). The two intermediate configurations (exclusive signal with pre-applied triggers & inclusive signal with no triggers) are very close in performance with the pre-triggered exclusive configuration performing just slightly better after application of triggers.

Analysis stage	Classification	
Generation:	#signal events	45900
	# $b\bar{b}$ events	10^7
	#DoIs	37602
Reconstruction:	#true	13584
	#false from sig evt	223
Pre-selection:	#true	9314
	#false from $b\bar{b}$	8326×488
	#false from sig evt	2029
	#true passing triggers	1724
	#false passing triggers	562243
Final selection 1: (no trigger, exclusive signal)	#true	2186
	#false	985
	#true passing triggers	576
	#false passing triggers	5
Final selection 2: (no trigger, inclusive signal)	#true	2243
	#false	976
	#true passing triggers	700
	#false passing triggers	0
Final selection 3: (trigger used exclusive signal)	#true	2637
	#false	10748
	#true passing triggers	701
	#false passing triggers	5
Final selection 4: (trigger used inclusive signal)	#true	2659
	#false	2440
	#true passing triggers	824
	#false passing triggers	0

Table 5.17: Number of candidates at each stage in the reconstruction and selection process for $B_d^0 \rightarrow K^+ \pi^-$ selection.

5.8.1 Cut set robustness

A major benefit of using a cut-based selection rather than a more abstracted method is that assessing the effect of systematic shifts in measured distributions is very easy. To ensure that the tuned cuts are robust (i.e. that they are not hovering on the edge of a region with much poorer performance) and to gain some understanding of the effect of measurement systematics on the selection cuts, a parameter scan has been performed around the best-tuning point for each of the cuts independently. The scan range was determined by finding the standard deviation of the signal distribution for each cut and then making variations up to 20% of the standard deviation in each direction. The results of these scans on signal and background yields after trigger can be seen for each tuning in Figure 5.18.

The initial conclusion from these scans is that none of the tunings are near particularly sensitive areas of the parameter space, i.e. there are no rapid fall-offs in signal yields and while the backgrounds show almost immediate increases of background from 0 to ~ 500 by varying some parameters, this just represents passing one inclusive $b\bar{b}$ event. Notably, and consistent with expectations, the tighter cuts on the tunings without pre-applied triggers result in lower susceptibilities to background, only passing one inclusive $b\bar{b}$ event per variable within the shift range while the trigger-aware tunings pass up to 3 inclusive $b\bar{b}$ events per variable. Given that we have optimised to construct a cut set which *minimally* excludes inclusive $b\bar{b}$ events, this is unsurprising.

The variations in signal yields due to this systematic study will be used in Chapter 6 to assess the annual yield uncertainties of LHCb on these channels and the implications for measurements of γ .

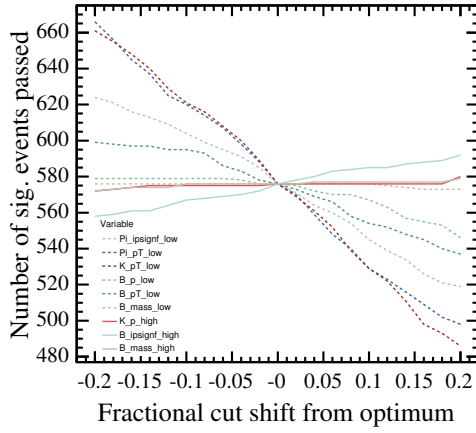
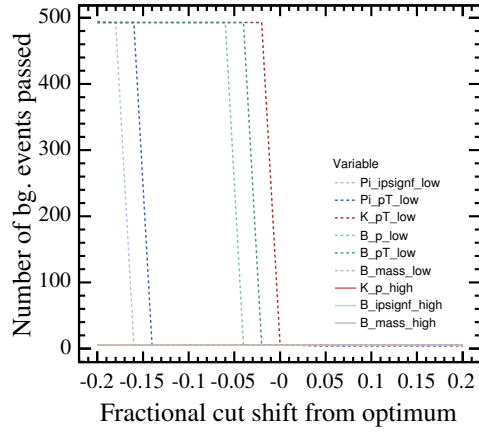
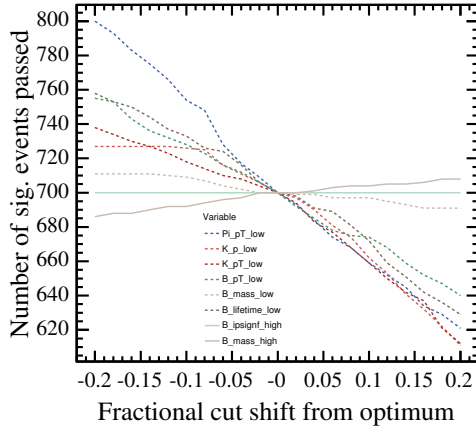
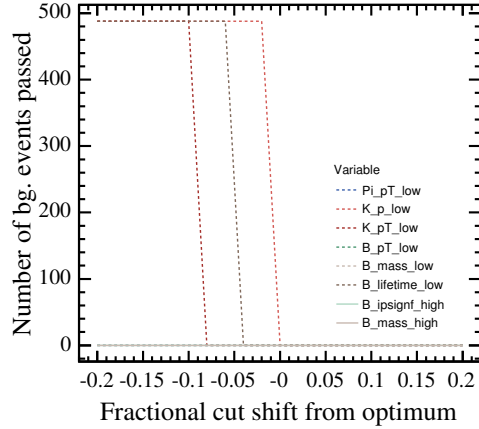
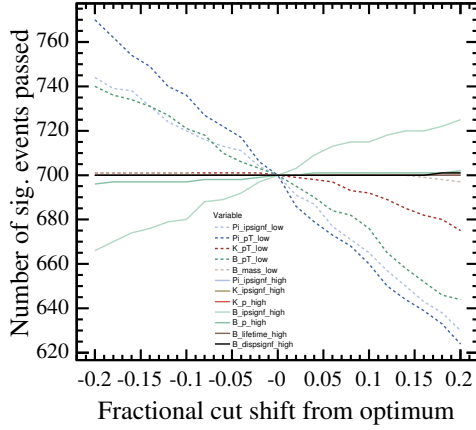
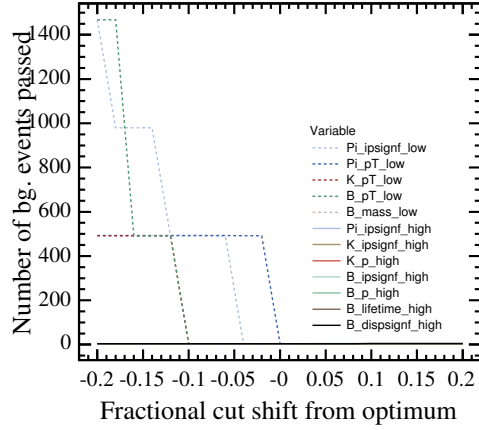
(a) B_d^0 signal (no trigger & exclusive signal)(b) B_d^0 background (no trigger & exclusive signal)(c) B_d^0 signal (no trigger & inclusive signal)(d) B_d^0 background (no trigger & inclusive signal)

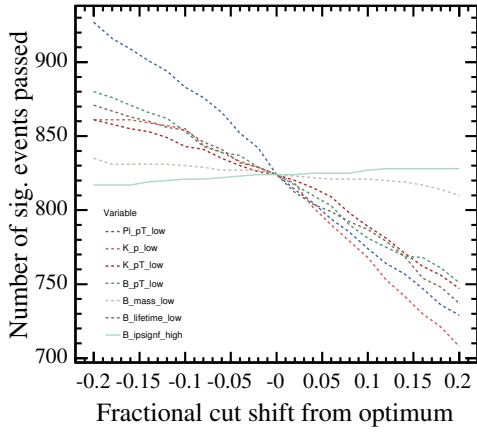
Figure 5.18: Shifts in number of signal and background events passing $B_d^0 \rightarrow K^+ \pi^-$ selection cuts and triggers due to variation of the cut values. The horizontal axis is in units of the standard deviation of the signal distribution for each cut variable so that the horizontal range represents systematic uncertainties by $\pm 20\%$ as described in the text.



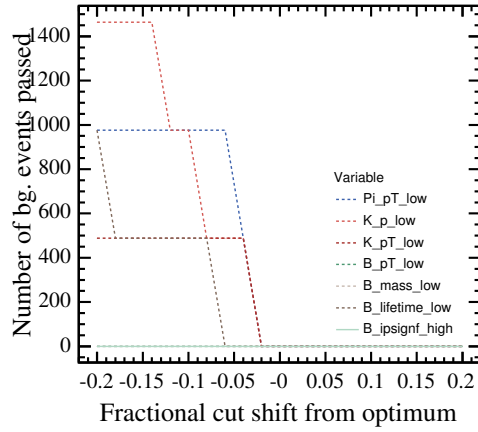
(e) B_d^0 signal (with trigger & exclusive signal)



(f) B_d^0 background (with trigger & exclusive signal)



(g) B_d^0 signal (with trigger & inclusive signal)



(h) B_d^0 background (with trigger & inclusive signal)

Figure 5.18: Shifts in number of signal and background events passing $B_d^0 \rightarrow K^+ \pi^-$ selection cuts and triggers due to variation of the cut values (contd.)

5.9 Cut performance for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$

The numbers of events to pass various stages of the $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ selection procedure are listed in Table 5.18. In this case all tuning methods and K_S^0 -type combinations result in removing all background candidates, both from the $b\bar{b}$ inclusive sample and the false combinatoric candidates from the signal distribution. Again this suggests that removing the heavily-weighted $b\bar{b}$ inclusive candidates requires tight enough cuts that the remaining signal combinatorics are either automatically removed or can be easily removed without affecting any true signal events.

Again, the performance can be improved by applying the trigger to the events before tuning and by slackening the definition of “signal” to include the events for which only a combinatoric candidate is available. For the $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ channel, the best performance is for the inclusive pre-triggered tuning, which yields 108 LL and 77 DD candidates after triggers: a total of 185 signal events. The lowest performance is again for the exclusive un-triggered sample, which yields 31 and 17 LL and DD events respectively from the same data set, a total of 48 signal events.

Comparing the numbers of background events before trigger application, we see that the pre-triggered sets allow much more background to pass as they are “aware” that the trigger will remove the remainder. The inclusive tuning results in lower pre-trigger backgrounds with the pre-triggered tuning. For the tunings where no knowledge of the trigger was available to the sampler the pre-trigger backgrounds are lower but non-zero: the weight factor means that the LL and DD data sets have backgrounds of 0 and 1 $b\bar{b}$ event respectively for the exclusive signal definition and 8 $b\bar{b}$ events each for the inclusive definition. In a sense the total removal of background is “lucky” in these cases, since the tuning was not aware that the trigger would remove the remaining background events.

Analysis stage	Classification	LL	DD
Generation:	#signal events	— 54500 —	—
	# $b\bar{b}$ events	— 10^7 —	—
	#DoIs	— 48919 —	—
Reconstruction:	#true	1402	4000
	#false from sig evt	91	309
Pre-selection:	#true	1155	2769
	#false from $b\bar{b}$	21543×1150	13965×1150
	#false from sig evt	388	589
	#true passing triggers	243	292
	#false passing triggers	2.44×10^6	1.02×10^6
Final selection 1: (no trigger, exclusive signal)	#true	94	116
	#false	0	1150
	#true passing triggers	31	17
	#false passing triggers	0	0
Final selection 2: (no trigger inclusive signal)	#true	212	349
	#false	8050	8050
	#true passing triggers	71	68
	#false passing triggers	0	0
Final selection 3: (trigger used, exclusive signal)	#true	149	326
	#false	39104	35666
	#true passing triggers	48	68
	#false passing triggers	0	0
Final selection 4: (trigger used, inclusive signal)	#true	385	338
	#false	2300	9200
	#true passing triggers	108	77
	#false passing triggers	0	0

Table 5.18: Number of candidates at each stage in the reconstruction and selection process for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ selection.

5.9.1 Cut set robustness

As for the $B_d^0 \rightarrow K^+ \pi^-$ channel, systematic shifts of each cut parameter of $\pm 20\%$ from the optimised values were performed for each tuning of the LL and DD K_S^0 selections. The results can be seen in Figures 5.19 and 5.20 for the LL and DD data sets respectively.

Again, there are no significant problems revealed by this systematic scan: the granularity is more obvious in the signal distributions this time due to the overall small number of events to pass the cuts and triggers but the best-performance cut tuning (inclusive signal, with pre-application of the triggers) shows a lowest value at $+20\%$ with variation of the π^+ IP significance cut which is still better than the central values of the other tunings for the LL set. For the DD data set, $+20\%$ variations of the K^+ and π^- p_T parameter cuts produce only marginally worse signal yields than the central values of the intermediate tunings. Looking at the background distributions, no variables in any tuning pass more than one inclusive $b\bar{b}$ event within the $\pm 20\%$ scan range.

Again, the indication is that the sampler approach under all tunings has produced a robust set of cuts in terms of both signal yield and background contamination. The results of the above scan will be used as an input to the γ sensitivity studies in the next chapter.

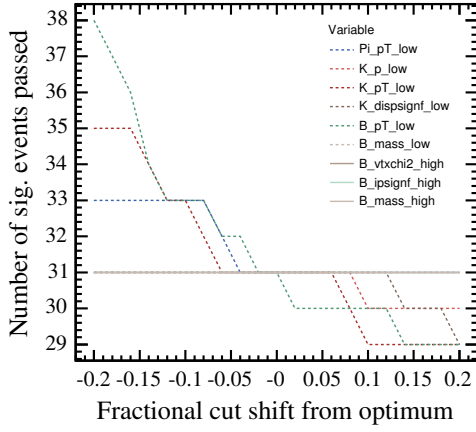
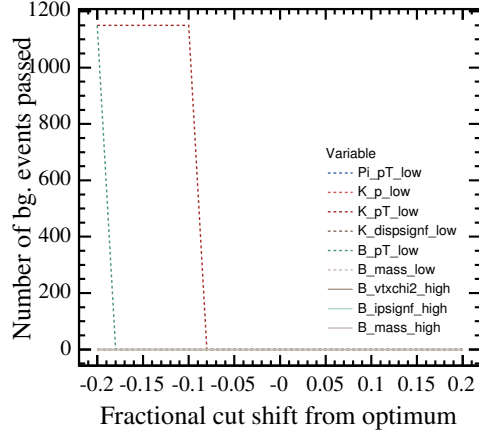
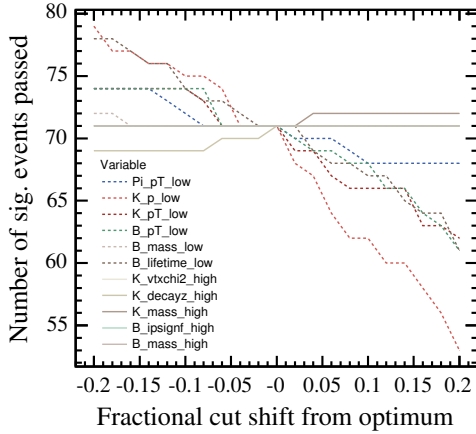
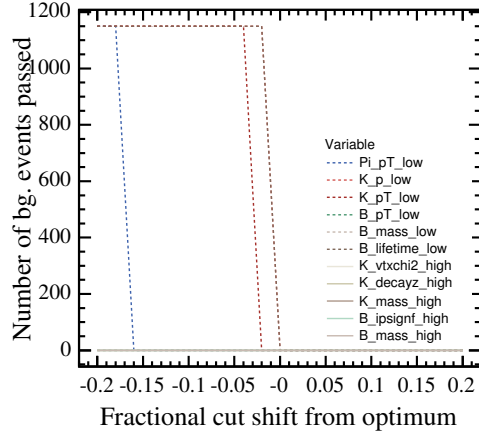
(a) B^+ LL signal (no trigger & exclusive signal)(b) B^+ LL background (no trigger & exclusive signal)(c) B^+ LL signal (no trigger & inclusive signal)(d) B^+ LL background (no trigger & inclusive signal)

Figure 5.19: Shifts in number of signal and background events passing $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ LL selection cuts and triggers due to variation of the cut values. The horizontal axis is in units of the standard deviation of the signal distribution for each cut variable so that the horizontal range represents systematic uncertainties by $\pm 20\%$ as described in the text.

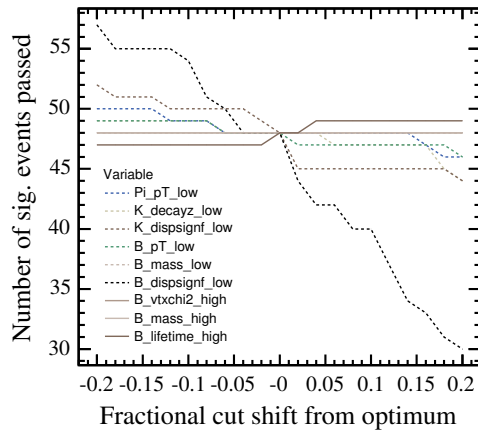
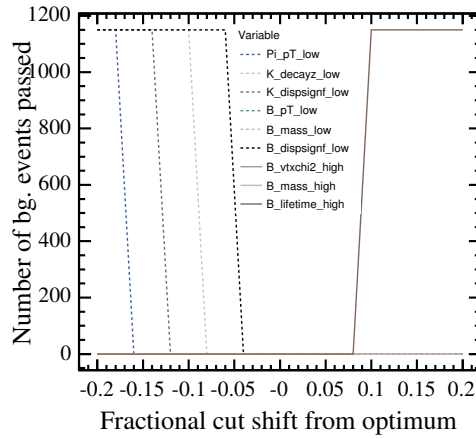
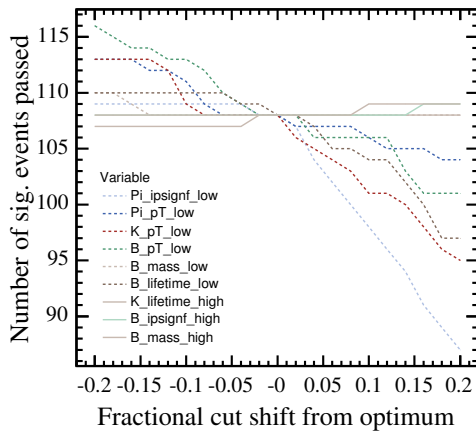
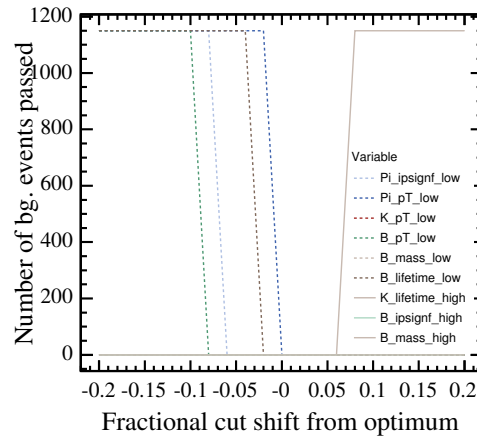
(e) B^+ LL signal (with trigger & exclusive signal)(f) B^+ LL background (with trigger & exclusive signal)(g) B^+ LL signal (with trigger & inclusive signal)(h) B^+ LL background (with trigger & inclusive signal)

Figure 5.19: Shifts in number of signal and background events passing $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ LL selection cuts and triggers due to variation of the cut values (contd.)

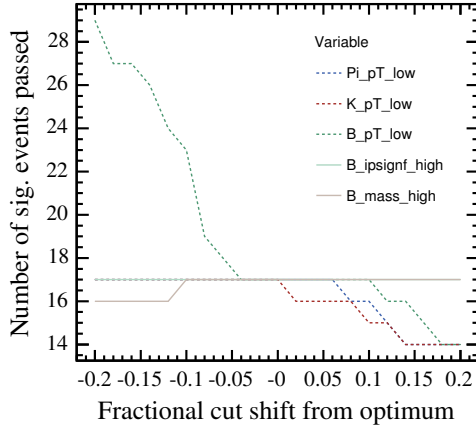
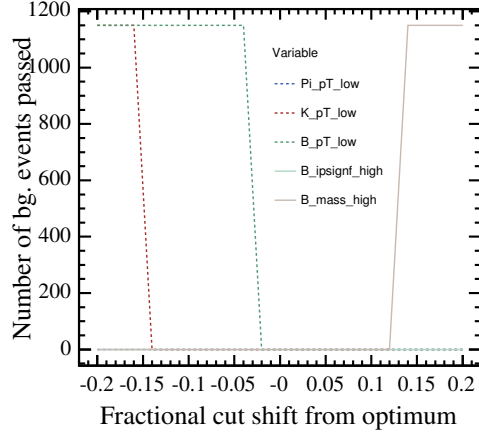
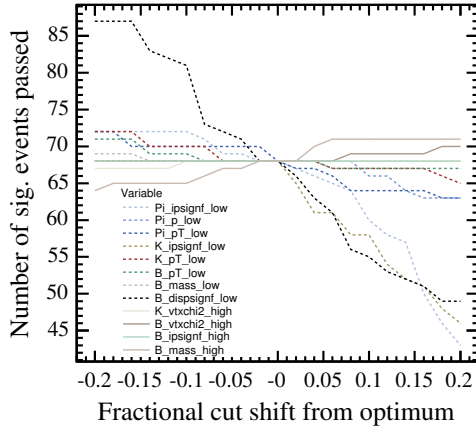
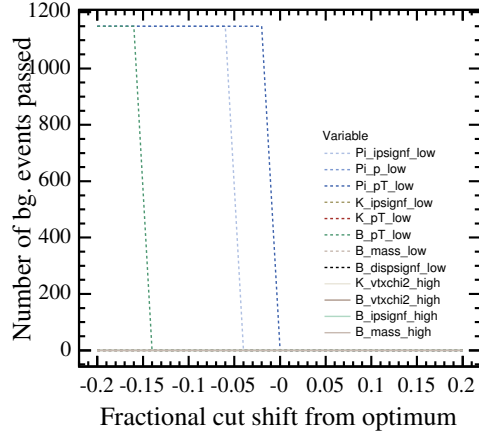
(a) B^+ DD signal (no trigger & exclusive signal)(b) B^+ DD background (no trigger & exclusive signal)(c) B^+ DD signal (no trigger & inclusive signal)(d) B^+ DD background (no trigger & inclusive signal)

Figure 5.20: Shifts in number of signal and background events passing $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ DD selection cuts and triggers due to variation of the cut values. The horizontal axis is in units of the standard deviation of the signal distribution for each cut variable so that the horizontal range represents systematic uncertainties by $\pm 20\%$ as described in the text.

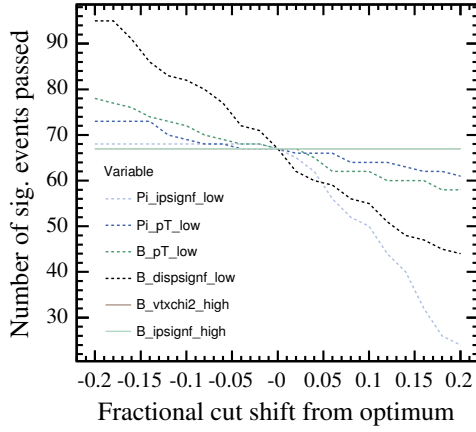
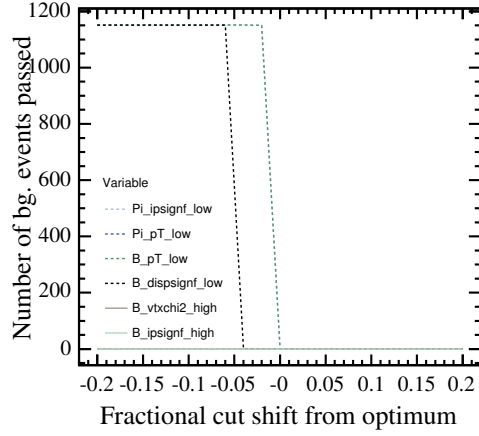
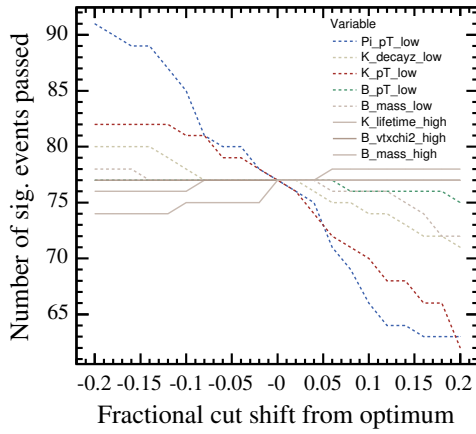
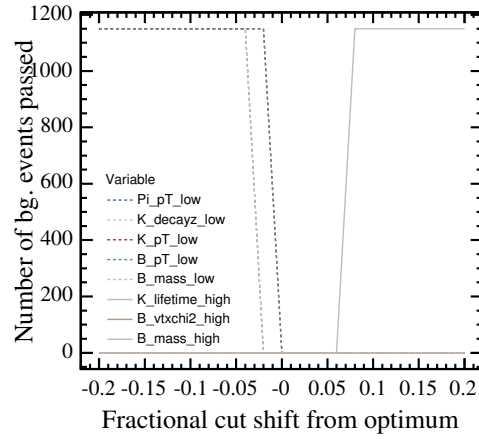
(e) B^+ DD signal (with trigger & exclusive signal)(f) B^+ DD background (with trigger & exclusive signal)(g) B^+ DD signal (with trigger & inclusive signal)(h) B^+ DD background (with trigger & inclusive signal)

Figure 5.20: Shifts in number of signal and background events passing $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ DD selection cuts and triggers due to variation of the cut values (contd.)

5.10 Conclusions

The Markov Chain sampler and pairs of low/high cuts on physical variables has been shown to be a robust approach for cut tuning, with the added benefits of checking for which cuts are useful and scanning around the optimised cut-space point for robustness and systematics checks. While the Markov Chain method does not offer the “guaranteed” optimisation of a parameter scan, it permits the use of a much larger number of characteristic variables for comparatively negligible processing time and is not limited by the finite granularity of a typical scan.

The final results shown above will be used in Chapter 6 to produce an estimated pair of annual yields and purities for the signal channels and an estimate on the γ -sensitivity of the $B \rightarrow K \pi$ strategy. It should be noted, however, that this study is statistics-limited. In particular we have been unable to study the systematic differences in efficiency between charge-conjugate channels and the signal data-sets are not large enough to tune on a different set from that to be cut.

Chapter 6

Event yields and sensitivity to γ from $B \rightarrow K \pi$ decays

*“There is nothing new to be discovered in physics now,
All that remains is more and more precise measurement.”*

— William Thompson, Lord Kelvin, 1824–1907

The methods described in the previous chapter lead us to a set of annual yield estimates for $B \rightarrow K \pi$ decays at the LHCb design luminosity and estimates of the contamination by background of the selected events. Finally, we explore the effects of statistical and systematic uncertainties on the CKM phenomenology described in the Chapter 4, to show LHCb’s sensitivity to measuring γ .

6.1 Efficiencies, signal yields and background estimates

To estimate the annual event yields for the selected signal and background channels, we choose to factorise the total efficiency in terms of the conditional probabilities of an event passing each stage of reconstruction and selection:

$$\varepsilon(\text{tot}) = \varepsilon(\text{det}) \times \varepsilon(\text{rec} \mid \text{det}) \times \varepsilon(\text{sel} \mid \text{rec}) \times \varepsilon(\text{trig} \mid \text{sel}) \quad (6.1)$$

where $\varepsilon(\text{det})$ represents the efficiency of detector effects like geometric acceptance, $\varepsilon(\text{rec} | \text{det})$ incorporates effects from the reconstruction algorithms, $\varepsilon(\text{sel} | \text{rec})$ corresponds to the selection algorithm efficiency and $\varepsilon(\text{trig} | \text{sel})$ represents the effect of the triggers.

In the case where the optimisation strategy makes use of the L0 and L1 trigger results in tuning selection cuts, the final pair of efficiencies may be combined if we ignore the effect of the high-level trigger (HLT):

$$\varepsilon(\text{sel} | \text{rec}) \times \varepsilon(\text{trig} | \text{sel}) \rightarrow \varepsilon(\text{sel}, \text{L0}, \text{L1} | \text{rec}) \sim \varepsilon(\text{sel}, \text{trig} | \text{rec}). \quad (6.2)$$

In practice $\varepsilon(\text{det})$ contains all efficiencies not accounted for by the other terms and is constrained by them:

$$\varepsilon(\text{det}) = \frac{S_{\text{sel}}/S_{\text{tot}} \times \varepsilon_{\text{sig}}^{\text{geom}}}{\varepsilon(\text{rec} | \text{det}) \times \varepsilon(\text{sel} | \text{rec}) \times \varepsilon(\text{trig} | \text{sel})}, \quad (6.3)$$

where S_{sel} is the number of signal events selected from an initial signal population of size S_{tot} and the efficiency factor $\varepsilon_{\text{sig}}^{\text{geom}} = 34.7\%$ is used to correct for the geometric 400 mrad cut on the Monte Carlo event generator.

The total efficiency, $\varepsilon(\text{tot})$, can be used as input to the calculation of annual yields, using the following relation:

$$S_{\text{ann}} = N_{\text{sig,ann}} \times \varepsilon(\text{tot}) \quad (6.4)$$

$$= \mathcal{L}_{\text{int}} \times \sigma_{b\bar{b}} \times 2 \times f_B \times \mathcal{B}_{\text{vis}} \times \varepsilon(\text{tot}) \quad (6.5)$$

where $\mathcal{L}_{\text{int}} = 10^7 \text{ s} \times (2 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}) = 2 \text{ fb}^{-1}$ is the nominal annual integrated luminosity and the $b\bar{b}$ production cross-section $\sigma_{b\bar{b}} = 500 \text{ } \mu\text{b}$. The hadronisation probability to a B^+ or B_d^0 , $f_B = 39.1\%$ and the factor of 2 accounts for the fact that both the b and the \bar{b} can hadronise. \mathcal{B}_{vis} is the sum of the visible branching ratios of interest, $\mathcal{B}_{\text{vis}}(B_d^0 \rightarrow K^+ \pi^-) = 1.85 \times 10^{-5}$ and $\mathcal{B}_{\text{vis}}(B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+) = 15.1 \times 10^{-6}$. Substituting these values and the expanded form of $\varepsilon(\text{tot})$, we obtain

$$S_{\text{ann}} = N_{\text{sig,ann}} \times [\varepsilon(\text{det}) \times \varepsilon(\text{rec} | \text{det}) \times \varepsilon(\text{sel} | \text{rec}) \times \varepsilon(\text{trig} | \text{sel})] \quad (6.6)$$

$$\text{with } N_{\text{sig,ann}} = \begin{cases} (1.5 \pm 0.1) \times 10^7 \text{ for } B_d^0 \rightarrow K^+ \pi^- \\ (5.9 \pm 0.5 \pm 0.3) \times 10^6 \text{ for } B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+ \end{cases}, \quad (6.7)$$

where the uncertainties on $N_{\text{sig,ann}}$ are directly correlated to the statistical and systematic (for $B^+ \rightarrow K_S^0(\pi^+\pi^-)\pi^+$) errors on the measured visible branching ratios from Section 5.1.1.

Using the numbers of signal events to pass reconstruction and triggering from Tables 5.17 and 5.18, we can directly obtain the total efficiency $\varepsilon(\text{tot}) = S_{\text{sel}}/S_{\text{tot}} \times \varepsilon_{\text{sig}}^{\text{geom}}$.

6.1.1 Efficiencies and signal event yields

The total selection efficiencies for both signal channels, using all four selection tunings, are shown in Table 6.1.

Tuning	B_d^0	B^+
No trig., excl.	$(4.35 \pm 0.17) \times 10^{-3}$	$(3.05 \pm 0.42) \times 10^{-4}$
No trig., incl.	$(5.29 \pm 0.19) \times 10^{-3}$	$(8.85 \pm 0.76) \times 10^{-4}$
Trigd., excl.	$(5.33 \pm 0.18) \times 10^{-3}$	$(7.38 \pm 0.66) \times 10^{-4}$
Trigd., incl.	$(6.23 \pm 0.21) \times 10^{-3}$	$(11.77 \pm 0.89) \times 10^{-4}$

Table 6.1: Total selection efficiencies, $\varepsilon(\text{tot})$, (including reconstruction, trigger and selection) for the four selection tunings on $B_d^0 \rightarrow K^+\pi^-$ and $B^+ \rightarrow K_S^0(\pi^+\pi^-)\pi^+$ decay channels.

The uncertainties in this table are binomial uncertainties on the number of observed signal events, calculated as

$$\begin{aligned} \Delta S_{\text{sel}} &= \sqrt{S_{\text{tot}} p_{\text{sel}} q_{\text{sel}}} \\ &= \sqrt{S_{\text{sel}}(1 - S_{\text{sel}}/S_{\text{tot}})}, \end{aligned} \quad (6.8)$$

where $p_{\text{sel}} = S_{\text{sel}}/S_{\text{tot}}$ is the probability of a randomly chosen signal event being selected and $q_{\text{sel}} = 1 - p_{\text{sel}}$. As a result, $\Delta\varepsilon(\text{tot}) = \sqrt{\varepsilon(\text{tot})(1 - \varepsilon(\text{tot}))/\Delta S_{\text{tot}}}$. In general, the $B^+ \rightarrow K_S^0(\pi^+\pi^-)\pi^+$ total efficiencies are an order of magnitude lower than the efficiencies for the $B_d^0 \rightarrow K^+\pi^-$ channel.

Using these total efficiencies and equation (6.6), the annual yield estimates can be computed for each channel and tuning. These are listed in Table 6.2 where the binomial uncertainties in the efficiencies have been combined in quadrature with the statistical uncertainties on the branching ratios. The $B^+ \rightarrow K_S^0(\pi^+\pi^-)\pi^+$ yield estimates also include

the effect of the Belle systematic uncertainty.

Tuning	B_d^0	B^+
No trig., excl.	$(62.99 \pm 4.51)\text{k/yr}$	$(1.79 \pm 0.15 \pm 0.26)\text{k/yr}$
No trig., incl.	$(76.58 \pm 5.36)\text{k/yr}$	$(5.19 \pm 0.45 \pm 0.52)\text{k/yr}$
Trigd., excl.	$(77.17 \pm 5.43)\text{k/yr}$	$(4.34 \pm 0.38 \pm 0.45)\text{k/yr}$
Trigd., incl.	$(90.02 \pm 6.17)\text{k/yr}$	$(6.91 \pm 0.60 \pm 0.79)\text{k/yr}$

Table 6.2: Annual signal yield estimates for $B_d^0 \rightarrow K^+ \pi^-$ and $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ channels with LHCb. The uncertainties are first statistical and second systematic. Uncertainties on the B production cross-section ($\sim 10\%$ [188]) and the fragmentation factors ($\sim 1\%$ for B_d^0 and B^+) are not included.

For $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$, using the total best efficiency of 0.118% (for combined LL and DD K_S^0 events) gives an annual signal yield estimate of 6.9k events per year. For $B_d^0 \rightarrow K^+ \pi^-$, the maximum estimated annual signal yield is 90k events per year.

Detector efficiency estimates

A breakdown of selection efficiencies at each stage of the reconstruction, selection and triggering process for the best-tuning (inclusive signal definition with pre-applied triggers) is shown in Table 6.3.

From Table 6.3, we calculate $\varepsilon(\text{det})$ as described above to be $\varepsilon(\text{det}, B_d^0) \sim 10.44\%$ for the $B_d^0 \rightarrow K^+ \pi^-$ channel and $\varepsilon(\text{det}, B^+) \sim 5.01\%$ for the $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ channel. This is in line with expectations, given the low reconstructibility of the K_S^0 and the fact that we have only used the two dominant K_S^0 reconstruction classes in this study. It also agrees with the LHCb detector efficiencies quoted for other $B \rightarrow h^+ h^-$ and $B \rightarrow X K_S^0$ [95].

6.1.2 $b\bar{b}$ inclusive background constraint

The background from $b\bar{b}$ inclusive decays to these signal channels can also be estimated. Although all of the $b\bar{b}$ event candidates have been removed by the combination of selection cuts and low-level triggers, the inadequacy of the statistics means that we can

Efficiency	B_d^0	B^+
$\varepsilon(\text{rec} \mid \text{det})$	13584/13807 = 98.4%	5402/7871 = 68.6%
$\varepsilon(\text{presel} \mid \text{rec})$	9314/13584 = 68.6%	3601/5402 = 66.7%
$\varepsilon(\text{sel} \mid \text{rec})$	2659/13584 = 19.6%	723/5402 = 13.4%
$\varepsilon(\text{sel} \mid \text{presel})$	2659/9314 = 28.5%	723/3601 = 20.1%
$\varepsilon(\text{trig} \mid \text{sel})$	824/2659 = 31.0%	185/723 = 25.6%
$\varepsilon(\text{tot})$	824/45900 \times 34.7% = 0.623%	185/54500 \times 34.7% = 0.118%

Table 6.3: Components of the selection efficiencies for the $B_d^0 \rightarrow K^+ \pi^-$ and $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ channels. The “presel” efficiency accounts for pre-selection signal losses, predominantly due to failed K – π vertexing and, in the $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ case, that only LL and DD K_S^0 classes are considered viable for selection.

only set a Poisson upper limit on the background level. To estimate the $b\bar{b}$ background with more realistic statistics we must assume that the chosen cuts are on the boundary for $b\bar{b}$ exclusion (a realistic estimate, given the nature of the optimisation procedure).

Using the procedure described by Feldman and Cousins [189], a 90% confidence band for a Poisson process with mean rate 0 and no observed events has its upper limit at 2.44. Hence, using the appropriate inclusive $b\bar{b}$ weight factors, $w_{b\bar{b}}$, (1150 for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ and 500 for $B_d^0 \rightarrow K^+ \pi^-$, as documented in Section 5.6.2), the 90% confidence upper limit on the inclusive $b\bar{b}$ background level is given by

$$\frac{B_{\text{sel}, b\bar{b}}}{S_{\text{sel}}} \sim w_{b\bar{b}} \times \frac{2.44}{S_{\text{sel}}}. \quad (6.9)$$

Using this relation, the estimated $b\bar{b}$ -to-signal ratios are listed in Table 6.4.

Tuning	B_d^0	B^+
No trig., excl.	(2.07 \pm 0.15)	58.72 \pm 4.95 \pm 8.62
No trig., incl.	(1.70 \pm 0.12)	20.18 \pm 1.83 \pm 2.39
Trigd., excl.	(1.69 \pm 0.12)	24.04 \pm 2.20 \pm 2.57
Trigd., incl.	(1.44 \pm 0.10)	15.23 \pm 1.28 \pm 1.10

Table 6.4: $B_{b\bar{b}}/S$ estimates for $B_d^0 \rightarrow K^+ \pi^-$ and $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ channels with LHCb. The uncertainties are first statistical and second systematic.

The best 90% confidence limits obtained give $B_{b\bar{b}}/S(B_d^0) < 1.44$ for $B_d^0 \rightarrow K^+ \pi^-$ and $B_{b\bar{b}}/S(B^+) < 15.23$ for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$. The weakness of these limits is again an effect of statistical mismatch between the signal and inclusive $b\bar{b}$ data sets. The expectation is that a scaling analysis of the B invariant-mass side-bands will allow a reduction of order 10 on these estimates [95], but the tightness of the pre-selection mass sidebands for the $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ channel means that this analysis is not possible with DC03 data.

A trial relaxation of the B_d^0 channel's best-tune low cut on the B_d^0 mass from $> 5244 \text{ MeV}/c^2$ to $> 5000 \text{ MeV}/c^2$ selects 3 inclusive $b\bar{b}$ events as opposed to 0 for the tight cut. This corresponds to a 90% confidence range of 1.10–7.42 events by the Feldman-Cousins procedure. Taking the B_d^0 mass of $5279.4 \text{ MeV}/c^2$ as the “central” reference value for the B_d^0 mass cut, this corresponds to a factor of $(5279 - 5000)/(5279 - 5244) \sim 7.68$ increase in the cut acceptance with only a factor $7.42/2.44 \sim 3.04$ increase in background, giving a factor of 2.5 decrease in the upper-limit $B_{b\bar{b}}/S(B_d^0 \rightarrow K^+ \pi^-)$ figure to < 0.57 . This figure should be taken as indicative rather than absolute, since variable correlations mean that a better tuning point could almost certainly be found with the mass window constrained to be artificially wide. It is uncertain what value a study of that nature would add to the existing analysis. It is also worth noting that the maximisation of signal significance for a given data set should, when fully converged, always place the cuts on the boundary of just selecting a background event as this will maximise the signal acceptance: the passing of inclusive $b\bar{b}$ events when cuts are widened is unsurprising when this is borne in mind and it should again be emphasised that more statistics are required in order to tune and evaluate cuts on independent data sets.

A similar extension of the LL and DD $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ B mass windows to their pre-selection limits (a factor of 5 widening) selects 13 and 14 background events from the inclusive $b\bar{b}$ data set for the LL and DD selections respectively. Using the Feldman-Cousins method, this corresponds to an increase rather than a further constraint on the background and so the principal conclusion for this channel must be that it has not yet been analysed with statistically sufficient data sets. The comments on the validity of the mass-window-extension method for the $B_d^0 \rightarrow K^+ \pi^-$ selection also apply to $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$.

6.1.3 Specific decay channel backgrounds

We also consider the effect of contamination of the $B_d^0 \rightarrow K^+ \pi^-$ selection with topologically similar decays which only differ in the identity of a charged pion or kaon. The specific backgrounds studied are $B_d^0 \rightarrow K^+ K^-$ and $B_d^0 \rightarrow \pi^+ \pi^-$, both of which were studied on data sets containing 50k events. The relevant \mathcal{B}_i branching ratios are $\mathcal{B}(B_d^0 \rightarrow K^+ K^-) < 0.7 \times 10^{-6}$ and $\mathcal{B}(B_d^0 \rightarrow \pi^+ \pi^-) = (4.4 \pm 0.7) \times 10^{-6}$. The numbers of background events from these data sets to pass selection cuts, as seen before the trigger application and after the L0 and L1 triggers separately, are

$$\text{KK: } 418 \xrightarrow{\text{L0}} 269 \xrightarrow{\text{L1}} 158 \quad (6.10)$$

$$\pi\pi: 186 \xrightarrow{\text{L0}} 119 \xrightarrow{\text{L1}} 66. \quad (6.11)$$

The specific background to signal ratios can now be calculated, using a procedure similar to that for the inclusive $b\bar{b}$ background:

$$\frac{B_{\text{sel}, i}}{S_{\text{sel}}} \sim w_i \times \frac{B_{\text{sel}, i}}{S_{\text{sel}}} \quad \text{for } i \in \{\text{KK}, \pi\pi\}, \quad (6.12)$$

where the new weight factors w_i are

$$w_i = \frac{S_{\text{tot}}}{B_{\text{tot}, i}} \times \frac{\mathcal{B}_i}{\mathcal{B}_{\text{sig}}}. \quad (6.13)$$

Adding the statistical uncertainties from the $\pi\pi$ branching ratio and the binomial sample uncertainties in quadrature, we obtain specific background contaminations of $B_{\text{KK}}/S = (7.25 \pm 0.85) \times 10^{-3}$ and $B_{\pi\pi}/S = (1.91 \pm 0.49) \times 10^{-2}$. In these, the relatively small contamination rates compared with the inclusive $b\bar{b}$ backgrounds are primarily due to the statistical comparability of the signal and specific background data sets.

6.2 Sensitivity to γ

Using the uncertainties on the annual yield estimates calculated in Section 6.1 and the studies of systematic shifts, we can extrapolate to the sensitivity to γ which LHCb will provide in one year of data-taking. Assuming binomial statistics again, the fractional uncertainties on the yields are $\Delta S/S \sim 1/\sqrt{S}$. Using the best yields from Table 6.2, these evaluate to $\Delta S/S(B_d^0) \sim 0.003$ for the $B_d^0 \rightarrow K^+ \pi^-$ channel and $\Delta S/S(B^+) \sim 0.012$ for

$B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$. These uncertainties do not include the contribution from the inclusive $b\bar{b}$ background constraints tabulated in Table 6.4. Including these, and recalculating the uncertainties as $\Delta(S+B)/S \sim \sqrt{S+B}/S$, we obtain $\Delta(S+B)/S(B_d^0) \sim 0.005$ for the $B_d^0 \rightarrow K^+ \pi^-$ channel and $\Delta(S+B)/S(B^+) \sim 0.048$ for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$. However, since these background bounds are a pessimistic upper limit, reflecting the statistical limits of our background study rather than a real limitation of the experiment, the yield uncertainties used to obtain LHCb sensitivities to γ will be the comparatively optimistic $\Delta S/S$ figures.

To study the LHCb sensitivity to γ , we will use the current-best measured values of branching ratios and \mathcal{CP} -asymmetries as central values. Constraints on γ are obtained via the contour relations from Chapter 4; estimates of LHCb sensitivity to γ can then be obtained by varying the measured observables within the statistical and systematic uncertainties already obtained in this chapter.

\mathcal{CP} -asymmetries are defined in terms of the ratio

$$\mathcal{A}_{\mathcal{CP}}(B \rightarrow f) = \frac{\mathcal{B}(\bar{B} \rightarrow \bar{f}) - \mathcal{B}(B \rightarrow f)}{\mathcal{B}(\bar{B} \rightarrow \bar{f}) + \mathcal{B}(B \rightarrow f)}, \quad (6.14)$$

which has been measured for the $B_d^0 \rightarrow K^+ \pi^-$ and $B^+ \rightarrow K^0 \pi^+$ signal channels by the *BABAR* and Belle experiments. Their results are listed in Table 6.5. The values for the $B_d^0 \rightarrow K^+ \pi^-$ direct \mathcal{CP} -asymmetry are more recent than the most recent Particle Data Group average, and so we will use the PDG's value of the zero-consistent $B^+ \rightarrow K^0 \pi^+$ \mathcal{CP} -asymmetry and calculate separate values of R and A_0 for *BABAR* and Belle, whose latest results differ by about 1σ .

Experiment	$\mathcal{A}_{\mathcal{CP}}(B_d^0 \rightarrow K^+ \pi^-)$	$\mathcal{A}_{\mathcal{CP}}(B^+ \rightarrow K^0 \pi^+)$
<i>BABAR</i>	-0.133 ± 0.030 (stat) ± 0.009 (syst) [20]	-0.21 ± 0.18 (stat) ± 0.03 (syst) [10]
Belle	-0.101 ± 0.025 (stat) ± 0.005 (syst) [21]	$0.03_{-0.08}^{+0.09}$ (stat) $_{-0.03}^{+0.01}$ (syst) [10]
PDG average	—	0.03 ± 0.08

Table 6.5: Values of $B \rightarrow K \pi$ \mathcal{CP} -asymmetries derived from measurements by *BABAR* and Belle. The Particle Data Group averaged value (dominated by the Belle measurement) is used for the zero-compatible \mathcal{CP} asymmetry in $B^+ \rightarrow K^0 \pi^+$ but not for the $B_d^0 \rightarrow K^+ \pi^-$ channel in which both experiments measure a large \mathcal{CP} violating component.

From the definition of a \mathcal{CP} -asymmetry and the average value for $\mathcal{A}_{\mathcal{CP}}(B^+ \rightarrow K^0 \pi^+)$

above, and using the $B^+ \rightarrow K^0 \pi^+$ branching ratio we obtain the denominator of the R and A_0 expressions in equation (4.19) as

$$\mathcal{B}(B^+ \rightarrow K^0 \pi^+) + \mathcal{B}(B^- \rightarrow \bar{K}^0 \pi^-) = (4.54 \pm 0.83 \text{ (stat)} \pm 0.23 \text{ (syst)}) \times 10^{-5}. \quad (6.15)$$

Keeping the *BABAR* and Belle $B_d^0 \rightarrow K^\pm \pi^\mp$ \mathcal{CP} -asymmetry values separate, we obtain central values for R and A_0 :

$$R(\text{BABAR}) = 0.72 \pm 0.14 \text{ (stat)} \pm 0.05 \text{ (syst)}, \quad (6.16)$$

$$R(\text{Belle}) = 0.74 \pm 0.15 \text{ (stat)} \pm 0.05 \text{ (syst)} \quad (6.17)$$

and

$$A_0(\text{BABAR}) = 0.096 \pm 0.019 \pm 0.006, \quad (6.18)$$

$$A_0(\text{Belle}) = 0.075 \pm 0.015 \pm 0.004 \quad (6.19)$$

where the systematic uncertainties have been conservatively added directly rather than in quadrature when calculating the ratios.

We now use the central values of these observables, combined with various predicted LHCb uncertainties, to estimate the LHCb precision on these quantities:

Statistical uncertainties: Taking a naïve unweighted average of the *BABAR* and Belle values for these ratios gives values $\langle R \rangle = 0.730 \pm 0.011$ and $\langle A_0 \rangle = (8.5 \pm 0.13) \times 10^{-2}$ where the uncertainties are derived from the anti-correlated limits of the one-year yield binomial uncertainties (again a conservative estimate). This corresponds to an improvement in precision of an order of magnitude over the values determined from the b-factory measurements. These central values, combined with studies of the systematic shifts shown in Chapter 5, allows us to study the sensitivity to γ with the method described in Chapter 4.

Tagging uncertainties: Separate studies of the $B^0 \rightarrow h^+ h^-$ decay channels have measured the effective flavour tagging efficiency of $B_d^0 \rightarrow K^+ \pi^-$ as $\epsilon D^2 = (4.8 \pm 1.0)\%$ for a sample size equivalent to ours. The large uncertainty on this number is indicative of the small sample size used to obtain the result and preliminary studies based on the LHCb DC04 Monte Carlo data already indicate much more precise flavour tagging efficiencies. For this reason we do not propagate this uncertainty into the one-year γ constraint, as it is not indicative of the expected knowledge of the flavour tagging efficiency at LHCb

by the time data-taking starts. We also note that, if re-scattering and particle ID inefficiencies can be ignored, then the $B_d^0 \rightarrow K^+ \pi^-$ channel is self-tagging. Making this assumption raises the tagging efficiency to $\sim 90\%$, the particle ID efficiency for a $K^\pm \pi^\mp$ final state.

Systematic uncertainties: The effect of the combined systematic uncertainties has been treated by finding all the yield shifts for $\pm 10\%$ systematic variations on the best selection tunings, where 10% variation has been chosen as a popular estimate of the systematic range. Since it is unlikely that all the systematic shifts will themselves be correlated, an estimate of their combined effect has been obtained by adding all the systematic shifts in each direction (increasing or decreasing the signal yield) in quadrature: i.e. incoherently summing all up-going or down-going shifts. This gives 10% systematic uncertainties on the central values of the two ratios: $\langle R \rangle = 0.730_{-0.06}^{+0.11}$ and $\langle A_0 \rangle = (8.5_{-0.7}^{+1.3}) \times 10^{-2}$. These 10% uncertainties are roughly an order of magnitude larger than the statistical uncertainties. Again we emphasise that this is only an exploratory estimate of the magnitude of systematic effects and that much stronger systematic constraints should be possible.

Applying the one-year statistical uncertainties to the \mathcal{CP} -asymmetries A_+ and $A_{K\pi}$, defined as

$$A_+ \equiv \frac{\mathcal{B}(B^+ \rightarrow K^0 \pi^+) - \mathcal{B}(B^- \rightarrow \bar{K}^0 \pi^-)}{\mathcal{B}(B^+ \rightarrow K^0 \pi^+) + \mathcal{B}(B^- \rightarrow \bar{K}^0 \pi^-)} \quad A_{K\pi} \equiv \frac{\mathcal{B}(\bar{B}_d^0 \rightarrow K^- \pi^+) - \mathcal{B}(B_d^0 \rightarrow K^+ \pi^-)}{\mathcal{B}(\bar{B}_d^0 \rightarrow K^- \pi^+) + \mathcal{B}(B_d^0 \rightarrow K^+ \pi^-)}, \quad (6.20)$$

we obtain precisions of $\langle A_+ \rangle = 0.030 \pm 0.003$ and $\langle A_{K\pi} \rangle = -0.117 \pm 0.014$, offering again an improvement of an order of magnitude over the current b-factory results.

γ bounds and constraints

The plots used in Chapter 4 to place bounds and constraints on γ are re-plotted showing the one-year statistical uncertainties and the effects of partially-correlated 10% systematics in Figures 6.1, 6.2 and 6.3.

The most significant feature of these plots is that the variations due to both the statistical and systematic variations are very small compared to the re-scattering effects presented in Section 4.3.1. For this reason, the limiting aspect of an attempt to constrain γ this way is likely to be the theoretical uncertainty arising from the hadronic effects

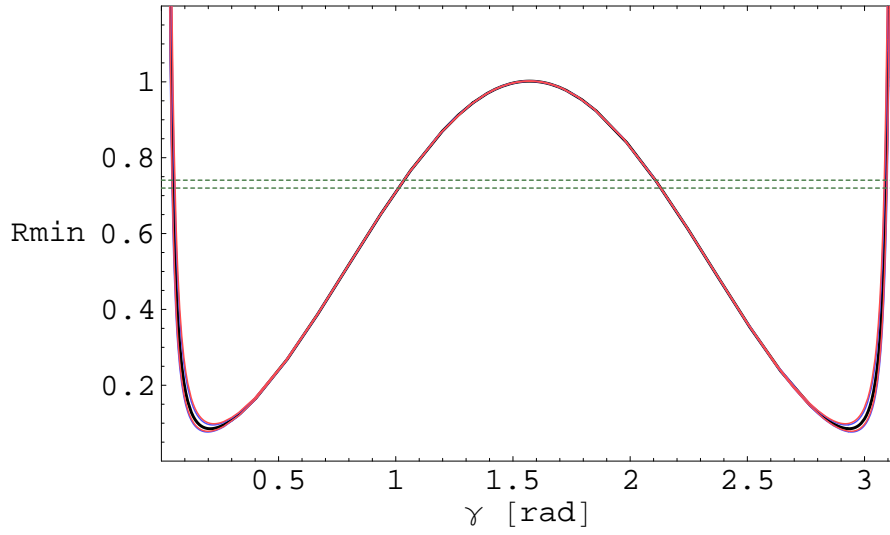


Figure 6.1: Variation of R_{\min} with γ for various values of A_0 , using the *BABAR*/*Belle* central value of $R = 0.73$. The black curves show the variation using the *BABAR* and *Belle* values of A_0 , the blue curves indicate the effect of the LHCb statistical uncertainty on A_0 after one year of data-taking and the red curves are a conservative estimate of the effects of 10% systematic uncertainties on the measurement. The horizontal green dotted lines are the current *BABAR* and *Belle* values of R , excluding values of γ in the regions where the R_{\min} curves are greater than the R bounds.

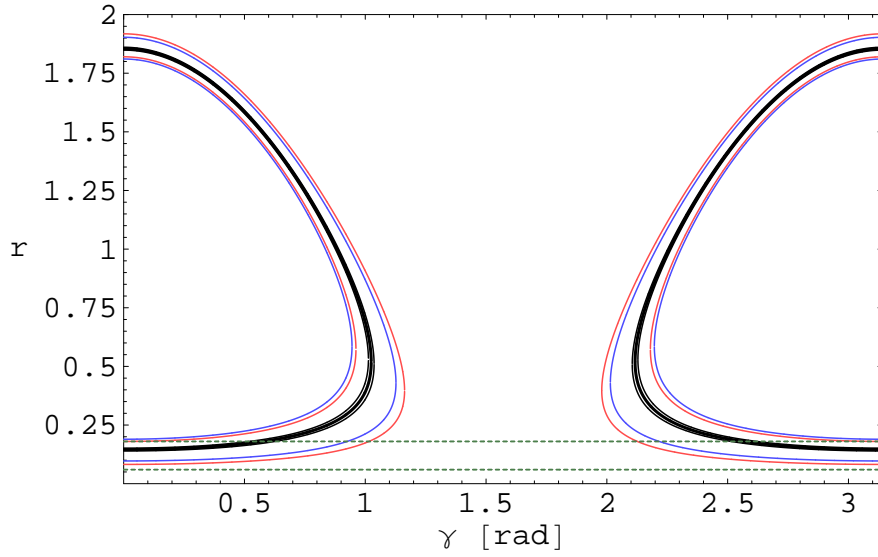


Figure 6.2: Limits on r as a function of γ , for various values of R , using the *BABAR*/*Belle* central value of $A_0 = 0.085$. The black curves show the limits obtained from using the *BABAR* and *Belle* values of R , the blue curves are from the estimated LHCb statistical uncertainty on R after one year of data-taking and the red curves are again from a conservative estimate of the systematic uncertainties. The horizontal green dotted lines indicate the typical range of r obtained from QCD factorisation and spin-ansatz analyses. Note that the central region of the γ range is not consistent with any real r value in this scheme — this is consistent with the R_{\min} bounds on γ shown in Figure 6.1.

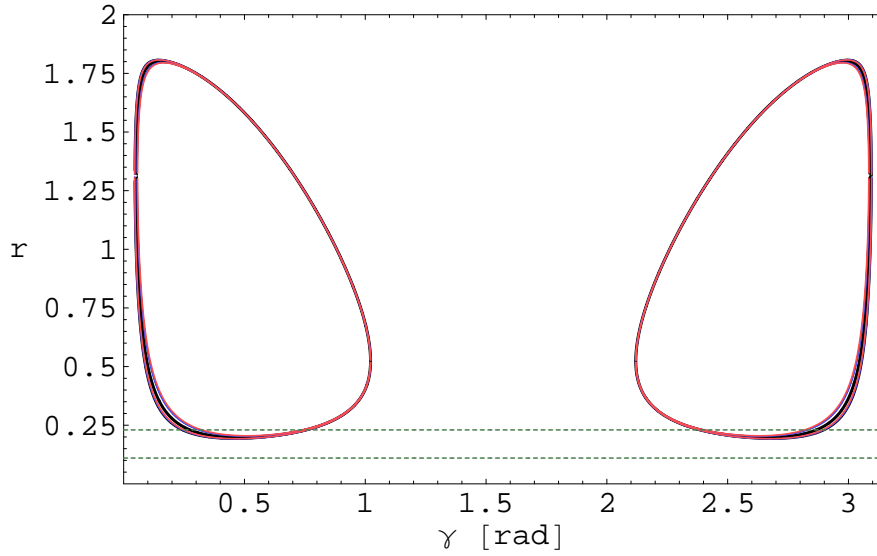


Figure 6.3: Limits on r as a function of γ , for various values of A_0 , using the *BABAR*/*Belle* central value of $R = 0.73$. The black curves show the limits obtained from using the *BABAR* and *Belle* values of A_0 , the blue curves are from the estimated LHCb statistical uncertainty on A_0 after one year of data-taking and the red curves are again from a conservative estimate of the systematic uncertainties. The horizontal green dotted lines again indicate the typical theoretical range of r .

rather than any systematic or statistical problems.

From Figure 6.1, we obtain a symmetric pair of allowed ranges for γ . Using the *BABAR*/*Belle* central value of $R = 0.73$, these ranges are $\gamma_{\text{bound1}} \in [3.0^\circ \pm 0.8^\circ, 58.7^\circ \pm 0.6^\circ]$ and $\gamma_{\text{bound2}} = 180^\circ - \gamma_{\text{bound1}}$, where the uncertainties are entirely due to the LHCb one-year uncertainty. The small experimental uncertainties and theoretical robustness (no constraint on r is required) of this bound are, however, offset by the wide range of γ values which are still allowed.

Using the contours in Figure 6.3, we obtain the LHCb one-year sensitivities to γ with a required input of r . The central values and experimental uncertainties are shown in Table 6.6 for various values of r in the 1σ allowed range 0.2–1.8. We should note, however, that r is not strongly predicted by any of the current theoretical models and that the favoured range is for values of r between about 10–20%. This latter point presents great difficulties for this approach, as this is aligned with the areas of the contours in Figure 6.3 which are least sensitive to variations in γ . For values of r less than ~ 0.25 , the uncertainty can easily grow as large as $\pm 15\%$. When this is the case, the two solutions within each contour become effectively degenerate and all sensitivity

is lost.

r	γ_1	γ_2
0.25	$15.8^\circ \pm 2.9^\circ$	$43.3^\circ \pm 1.5^\circ$
0.5	$5.9^\circ \pm 1.1^\circ$	$58.4^\circ \pm 0.4^\circ$
1.0	$3.3^\circ \pm 0.7^\circ$	$50.7^\circ \pm 0.3^\circ$
1.5	$3.3^\circ \pm 0.7^\circ$	$32.7^\circ \pm 0.3^\circ$

Table 6.6: Sensitivities to γ obtained using LHCb one-year uncertainties for various values of r and $A_0 = 0.085$. The construction of the contours means that in general four values of γ are obtained for each set of input parameters. These are $\gamma_{1,2}$ as listed and $\gamma_3 = 180^\circ - \gamma_2$, $\gamma_4 = 180^\circ - \gamma_1$.

The values of r used are for the “ansatz” fit used by Fleischer, Buras et. al. [161] at the low end ($r = 0.11 \pm 0.03$) and from the QCD factorisation approximation at the upper end of the range ($r \approx 0.2$). If r is non-factorisably enhanced into the $\sim 50\%$ region and the theoretical predictions can be strongly constrained then this strategy may have a strong contribution to make to measurements of γ , but for the current central values it is not particularly useful.

6.3 Sensitivity to new physics

The results presented so far indicate that while $B \rightarrow K \pi$ decays can be well reconstructed by LHCb, the limiting factor on using them to extract information about the CKM γ phase is the lack of theoretical control of the hadronic processes in charmless non-leptonic B-decays. However, the decays considered are also of interest because they contain electroweak penguin topologies which may be sensitive to new high-scale physics. This has been considered extensively in recent literature by Fleischer, Buras and others.

Electroweak penguin diagrams enter into the diagrams here in colour-suppressed form and are hence expected to be control channels for new physics loop effects: branching ratio enhancements in decays with colour-allowed electroweak penguin topologies relative to the channels studied here will be of great interest as new physics signals.

Chapter 7

Conclusions

“If economists were laid end to end, they would not reach a conclusion”

— George Bernard Shaw

This thesis has described two major elements of research relating to the LHCb detector under construction at CERN. These are the development and consequences of a new trackless approach to pattern recognition in ring-imaging Čerenkov (RICH) detectors; and the automated optimisation and resulting physics potential of LHCb event-selection algorithms for $B \rightarrow K \pi$ decays.

In the first case, I have described the standard “track-aware” approaches to RICH pattern recognition and particle identification and then motivated the development of a Markov Chain algorithm which attempts to do much the same job but without tracking information. The benefits of this approach are that the trackless algorithm can highlight flaws in the current tracking system and that it can be used to provide a slightly more informed approach to finding genuine background hits on the photon detector planes. This study has revealed that untracked rings are largely due to secondary interactions of particles with detector elements and that electron–positron pairs from the primary vertex are also major contributors to untracked rings. Since most such rings are either from primary pp interactions or from secondary-scattered tracks, their loss by the standard particle ID system is not a worry to LHCb from the point of view of physics potential. Incorporating the trackless ring-finder into the background estimator of the standard particle ID also produces interesting results, showing that the pion and kaon particle ID efficiencies may be improved with the use of an appropriately-tuned trackless finder,

but the computational overhead of such an approach renders this approach unrealistic for large-scale event reconstruction.

In my studies of $B \rightarrow K \pi$ decays with LHCb, I have developed a system for automatic optimisation of event selection criteria in terms of upper and lower cut pairs on physical variables. This framework again makes use of a sampler algorithm, this time to find an optimal set of cuts for a given measure of selection quality. I have studied several approaches to reconstructing $B_d^0 \rightarrow K^+ \pi^-$ and $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ decays, according to various degrees of inclusiveness in the definition of a “true” event and in the mode of application of the trigger algorithms. It has been found that the best performance was obtained for an inclusive definition of event truth when the optimisation was performed on data sets to which the trigger had already been applied. This is in contrast to the LHCb re-optimization TDR[95] assumption that tuning on un-triggered data would give better performance due to increased statistics. The available statistics and the absence of charge-conjugate channels from the 2003 data sample restricts studies of systematic effects in these channels. Robustness studies based on parameter scans around the optimal point in cut parameter space indicate that the chosen cuts are robust. The chosen selection cuts successfully reject all background from inclusive $b\bar{b}$ decays but studies of the purities of these data samples using the Feldman-Cousins 90% bounds on the observed zero-background only serve to highlight the limited statistics available in the studied data sets.

Interpreting the signal yields from Data Challenge 2003 data sets and extrapolating via the nominal LHCb luminosity predicts annual yields of $(90.0 \pm 6.2)\text{k}$ $B_d^0 \rightarrow K^\pm \pi^\mp$ events and $(6.9 \pm 0.6)\text{k}$ $B^\pm \rightarrow K^0 \pi^\pm$ events using the selection tunings presented here. These yields correspond to one-year statistical uncertainties on the R and A_0 ratios of $\langle R \rangle = 0.730 \pm 0.011$ and $\langle A_0 \rangle = (8.5 \pm 0.13) \times 10^{-2}$ where the current central values of the $B_d^0 \rightarrow K^\pm \pi^\mp$ and $B^\pm \rightarrow K^0 \pi^\pm$ \mathcal{CP} -asymmetries have been used. Similarly, the \mathcal{CP} asymmetries $\langle A_+ \rangle = 0.030 \pm 0.003$ and $\langle A_{K\pi} \rangle = -0.117 \pm 0.014$. All four branching ratio combinations correspond to an improvement in statistical precision of an order of magnitude compared to the b-factories after one year of LHCb data taking.

The above yields, when used to construct contours relating to the “mixed” $B \rightarrow K \pi$ channel strategy to constrain the CKM angle γ , indicate that tight constraints may be placed on the contour system but that theoretical uncertainties on the hadronic aspects of the decays strongly limit the value of this approach to constrain γ . The experimental statistical (one-year) and (estimated) systematic errors on measurements of γ , using current experimental central values for the relevant parameters, range from $\Delta(\gamma_1) = 2.9^\circ$,

$\Delta(\gamma_2) = 1.5^\circ$ for $r = 0.25$ to $\Delta(\gamma_1) = 0.7^\circ$, $\Delta(\gamma_2) = 0.3^\circ$ for $r \geq 1$. Here, $\gamma_{1,2}$ refer to the independent degenerate solutions for γ obtained from the phenomenological strategy described in Chapter 4. However, the tightness of these constraints belies the true nature of the result. The favoured range of values for the crucial tree-to-penguin ratio, r , are typically $r < 0.2$ and in this region the values of γ vary much more rapidly than the experimental errors grow — it is precisely the region where this strategy is least sensitive to measuring γ . Hence, assuming $r \in [0.1, 0.4]$, the corresponding theoretical uncertainty predicts $\gamma_{1,2} \in [6^\circ, 57^\circ]$, with the additional solutions $\gamma_{4,3} = 180^\circ - \gamma_{1,2}$. This dominance of the measurement precision by uncontrolled theoretical uncertainties indicates that the emphasis on the $B^\pm \rightarrow K^0 \pi^\pm$ decay channel should be shifted from a mode for measuring γ to being a control channel in searches for new physics in electroweak penguins.

Appendix A

Monte Carlo samplers

A.1 Metropolis algorithm

The Metropolis algorithm, more properly referred to as the Metropolis-Hastings method, was originally developed by Nicholas Metropolis in the 1950's and was extended in 1970 by W. K. Hastings [190, 191]. It avoids several problems of the simpler “importance sampling” and “rejection sampling” methods by use of a proposal function whose form is determined by the current state of the sampler [137]. It is nowadays a very common sampling method and has been used in Chapter 3 of this thesis in the form of the trackless ring-finding sampler.

The Metropolis-Hastings algorithm can draw samples from any probability distribution $P(x) \equiv P^*(x)/Z$, requiring only that the probability density can be calculated at x . For computational purposes, $P^*(x)$ is the important quantity, since the normalisation of the probability distribution is in general unknown and its calculation is computationally a highly non-trivial procedure. It generates a set of states $x^{(t)}$ where each state depends only on the previous one, hence it is a Markov chain. The algorithm creates a proposal density $Q(x'; x^{(t)})$ which is the probability density of generating a new proposed sample point x' . This is illustrated in Figure A.1.

We then calculate a value $a = a_1 a_2$ where

$$a_1 = \frac{P^*(x')}{P^*(x^{(t)})} \tag{A.1}$$

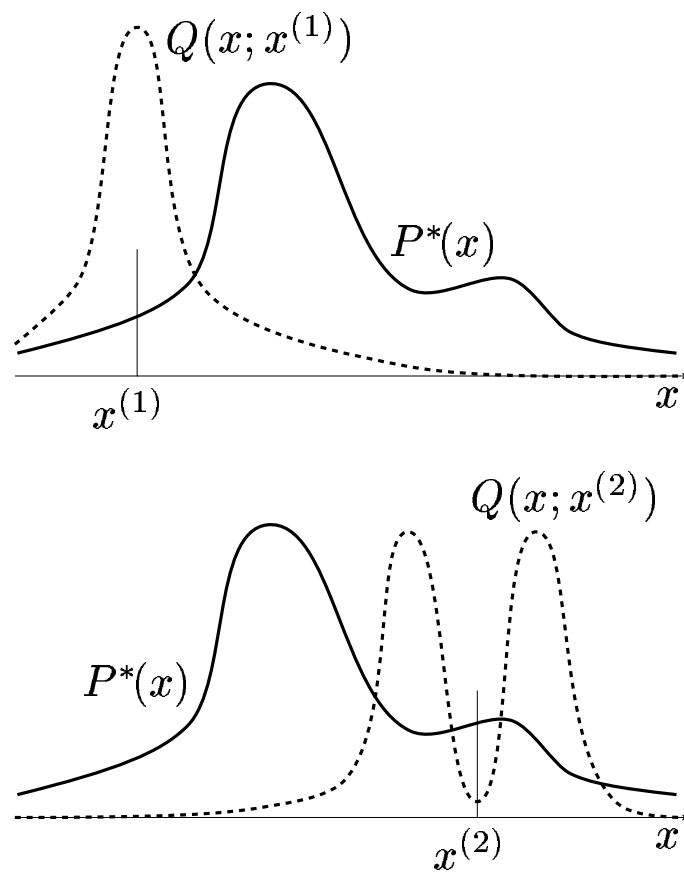


Figure A.1: Metropolis–Hastings proposal and unnormalised target distributions $Q(x'; x^{(t)})$ and $P^*(x)$. The proposal distribution is shown as having a shape which varies with x , though this is not typical in practice. Figure used courtesy of David MacKay.

is the likelihood ratio between the proposed sample x' and the previous sample $x^{(t)}$, and

$$a_2 = \frac{Q(x^{(t)}; x')}{Q(x'; x^{(t)})} \quad (\text{A.2})$$

is the ratio of the proposal density from $x^{(t)}$ to x' and vice versa. This is equal to 1 if the proposal density is symmetric: a property known as *detailed balance* which is desirable in practical proposal densities due to the simplification it offers. Then the new state $x^{(t+1)}$ is chosen with the rule

$$x^{(t+1)} = \begin{cases} x' & \text{if } a > 1 \\ x' \text{ with probability } a & \text{if } a < 1 \end{cases}. \quad (\text{A.3})$$

If $a < 1$ and the proposed new sample point is rejected then $x^{(t+1)} = x^{(t)}$, i.e. the existing sample point is repeated. Typically the algorithm is started at a random initial point $x^{(0)}$ and then a few thousand samples are calculated and discarded to “forget” the initial point: this period is known as the sampler *burn-in* time. The convergence time of such a sampler is not known in general, but in general it is improved if the shape of the proposal density matches the local shape of the target distribution $P^*(x)$: this is usually unknown. A robust rule of thumb is that if the largest length scale of the space is L and the typical step size in that direction is ϵ then at least $\sim (L/\epsilon)^2$ iterations must be sampled to obtain an independent sample of the distribution [137]. The random walk behaviour of the Metropolis sampler is reduced by use of gradient information in the proposal density: this broad class of enhancements to Markov chain samplers are called “hybrid Monte Carlo” samplers.

A.2 Simplex algorithm

The simplex algorithm used in Chapter 5 to optimise the $B \rightarrow K \pi$ selection cuts by minimising the function $-\mathcal{F} = -S/(S+B)$ was developed in the 1960’s by J. A. Nelder and R. Mead [192]. The property that it does not require gradient information to sample from the distribution places it in the class of samplers called *direct search* methods.

At each iteration t of the sampler, operating in an n dimensional vector space over the real numbers, $n+1$ points $x_1^{(t)}, x_2^{(t)}, \dots, x_{n+1}^{(t)}$ are calculated. These points describe a simplex in the space for each t . For convenience we order the points such that for a

given t and function to be minimised, $f(x)$,

$$f(x_1) \leq f(x_2) \leq \cdots \leq f(x_{n+1}). \quad (\text{A.4})$$

Trial vertices, denoted with a tilde cf. \tilde{x} , are calculated by a series of geometric transformations (reflection, expansion, contraction and shrinkage) applied to the simplex. The reflection transformation consists of reflecting the worst vertex (i.e. the one with the highest value of f) through the centroid of the other vertices:

$$\tilde{x}' = 2\langle x \rangle - x_{n+1} \quad (\text{A.5})$$

where

$$\langle x \rangle = \frac{1}{n} \sum_{i=1}^n x_i \quad (\text{A.6})$$

is the centroid of the n best vertices. The reflected vertex \tilde{x}' is accepted if $f(x_1) \leq f(\tilde{x}') < f(x_n)$ and x_{n+1} is updated without re-ordering the indices such that $x_{n+1} = \tilde{x}'$. The next iteration then begins. We will now consider the behaviour if the new point satisfies $f(\tilde{x}') < f(x_1)$ and will return shortly to what happens if the reflection vertex gives $f(x') \geq f(x_n)$.

If the reflection point satisfies $f(\tilde{x}') < f(x_1)$ then the trial step has produced a good point and the transformation is extended. The next transformation is an expansion of the trial point, computed as

$$\tilde{x}'' = 2\tilde{x}' - \langle x \rangle \quad (\text{A.7})$$

and this is accepted if $f(\tilde{x}'') < f(x_1)$, otherwise the reflected vertex \tilde{x}' is accepted. Either way, the next iteration begins.

If the reflection-proposed vertex, \tilde{x}' satisfies $f(\tilde{x}') \geq f(x_n)$ then it is not accepted and a contraction step is computed. The new contraction vertex \tilde{x}''' is

$$\tilde{x}''' = \begin{cases} \frac{1}{2} x_{n+1} + \frac{1}{2} \langle x \rangle & \text{if } f(x_{n+1}) \leq f(x') \\ \frac{1}{2} \tilde{x}' + \frac{1}{2} \langle x \rangle & \text{if } f(x_{n+1}) > f(x') \end{cases} \quad (\text{A.8})$$

and is accepted if $f(\tilde{x}''') < f(x_n)$. If this step also fails to produce an accepted point then the entire simplex is shrunk towards the best value by replacing each vertex by the

point between it and x_1 :

$$\tilde{x}_i = \frac{1}{2}(x_i + x_1). \quad (\text{A.9})$$

At this stage, all the transformations have been performed and so the vertex indices are re-ordered and the next iteration begins. The effect of the algorithm is that the vertex tumbles its way towards the function minimum, contracting as it does so. Various convergence criteria have been suggested, either in terms of the number of iterations, the convergence rate of simplex size, the standard error of the function and so-on. As mentioned in Chapter 5, we use the number of iterations as we are only analysing a single problem and are able to observe the convergence of the function value explicitly for each case.

Appendix B

Colophon

This thesis has been typeset in \LaTeX 2 ϵ , mainly written using the `emacs` editor on Linux systems. I have written several \LaTeX packages and a document class to assist in typesetting high-energy physics concepts, such as particle symbols and our rather peculiar units. These packages are as follows:

- **maybemath**: allows math mode symbols to assume the style (boldness, italic/upright shape, serif/sans-serif font etc.) of the surrounding text;
- **hepparticles**: a set of macros for typesetting HEP particle symbols in a flexible way, using **maybemath**;
- **hepname**s: a pre-defined set of particle names, using **hepparticles** and re-implementing the Particle Entity Notation scheme [193];
- **hepunits**: an extension of the **SIunits** package to include common HEP units;
- **hepthesis**: a document class based around the standard “book” class, with optional HEP-specific features.

These packages have been released through the \LaTeX “CTAN” system (available via the Web at, for example, <http://www.tex.ac.uk>) and are publicly available. I hope they will be of use to others wishing to write PhD theses and HEP-oriented documents.

Bibliography

- [1] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).
- [2] S. L. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev. **D2**, 1285 (1970).
- [3] J. Goldstone, A. Salam, and S. Weinberg, Phys. Rev. **127**, 965 (1962).
- [4] P. W. Higgs, Phys. Lett. **12**, 132 (1964).
- [5] G. Rajasekaran, Building up the standard gauge model of high-energy physics, in *Gravitation, gauge theories and the early universe*, edited by B. R. Iyer *et al.*, pp. 185–236, Kluwer Academic Pub., London, 1989.
- [6] S. Willenbrock, Symmetries of the standard model, 2004, hep-ph/0410370.
- [7] N. E. Mavromatos, \mathcal{CPT} violation: Theory and phenomenology, 2005, hep-ph/0504143.
- [8] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, Phys. Rev. Lett. **13**, 138 (1964).
- [9] M. Gell-Mann and A. Pais, Phys. Rev. **97**, 1387 (1955).
- [10] Particle Data Group, S. Eidelman *et al.*, Phys. Lett. **B592**, 1 (2004).
- [11] A. D. Sakharov, Pisma Zh. Eksp. Teor. Fiz. **5**, 32 (1967).
- [12] A. Kusenko, \mathcal{CP} violation and cosmology, 2002, hep-ph/0207028.
- [13] V. A. Kuzmin, V. A. Rubakov, and M. E. Shaposhnikov, Phys. Lett. **B155**, 36 (1985).
- [14] R. E. Marshak and R. N. Mohapatra, Phys. Lett. **B91**, 222 (1980).
- [15] R. N. Mohapatra and R. E. Marshak, Phys. Rev. Lett. **44**, 1316 (1980).

- [16] S. M. Barr, G. Segre, and H. A. Weldon, Phys. Rev. **D20**, 2494 (1979).
- [17] G. C. Branco, ECONF **C030603**, VEN05 (2003), hep-ph/0309215.
- [18] C. T. H. Davies and S. M. Playfer, editors, *Heavy flavor physics: Theory and experimental results in heavy quark physics and \mathcal{CP} violation*, 55th Scottish Universities Summer School in Physics, 2001.
- [19] NA48, A. Lai *et al.*, Eur. Phys. J. **C22**, 231 (2001), hep-ex/0110019.
- [20] BABAR, B. Aubert *et al.*, Phys. Rev. Lett. **93**, 131801 (2004), hep-ex/0407057.
- [21] Belle, Y. Chao *et al.*, Phys. Rev. Lett. **93**, 191802 (2004), hep-ex/0408100.
- [22] T. Nakada, Surveys High Energ. Phys. **17**, 3 (2002).
- [23] BABAR, B. Aubert *et al.*, (2000), hep-ex/0008048.
- [24] P. Ball, The theory of \mathcal{CP} -violation: In as much of a nutshell as will fit on 8 pages, 2004, hep-ph/0406326.
- [25] H. R. Quinn, The \mathcal{CP} puzzle in the strong interactions, 2001, hep-ph/0110050.
- [26] P. Ball *et al.*, B decays at the LHC, 2000, hep-ph/0003238.
- [27] N. Cabibbo, Phys. Rev. Lett. **12**, 62 (1964).
- [28] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
- [29] LHCb, V. Gibson, (1996), Prepared for 28th International Conference on High-energy Physics, Warsaw, Poland.
- [30] B. Kayser, \mathcal{CP} violation in the K and B systems, 1996, hep-ph/9702264.
- [31] L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).
- [32] C. Jarlskog, (1986), Contributed to 23rd International Conference on High-Energy Physics, Berkeley, CA.
- [33] J. Smit, JHEP **09**, 067 (2004), hep-ph/0407161.
- [34] Belle, K. Abe *et al.*, (2005), hep-ex/0507034.
- [35] BABAR, B. Aubert *et al.*, (2005), hep-ex/0507069.

- [36] F. Bellini, (2004), Prepared for 16th Conference on High Energy Physics, Turin, Italy.
- [37] J. L. Rosner, Status of the CKM matrix, 2004, hep-ph/0410281.
- [38] T. G. Mackay, S. Setiawan, and A. Lakhtakia, Eur. Phys. J. **C41**, 1 (2005), astro-ph/0411754.
- [39] *BABAR*, C. Touramanis, Observation of \mathcal{CP} violation in the B^0 system, 2001, hep-ex/0110064.
- [40] *BABAR*, The *BABAR* physics book: Physics at an asymmetric B factory, 1998.
- [41] Aleph, R. Barate *et al.*, Phys. Lett. **B492**, 259 (2000), hep-ex/0009058.
- [42] OPAL, E. Barberio, (1998), Prepared for 29th International Conference on High-Energy Physics (ICHEP 98), Vancouver, British Columbia, Canada.
- [43] CDF, G. Bauer, Measurement of the \mathcal{CP} -violation parameter $\sin(2\beta)$ in $B^0 \rightarrow J/\psi K_S^0$ decays, 1999, hep-ex/9908055.
- [44] CDF, G. Bauer, Mixing and \mathcal{CP} -violation measurements of B^0 mesons from the Tevatron collider, 1999, hep-ex/9904017.
- [45] *BABAR*, S. Rahatlou, Measurement of the \mathcal{CP} -violating asymmetry amplitude $\sin(2\beta)$ with the *BABAR* detector, 2002, hep-ex/0207088.
- [46] Belle, K. Abe *et al.*, Phys. Rev. **D71**, 072003 (2005), hep-ex/0408111.
- [47] E. A. Andriyash, G. G. Ovanesyan, and M. I. Vysotsky, Hunting for the α : $B \rightarrow \rho \rho$, $B \rightarrow \pi \pi$, $B \rightarrow \pi \rho$, 2005, hep-ph/0502149.
- [48] M. Gronau and D. London, Phys. Rev. Lett. **65**, 3381 (1990).
- [49] J. Zupan, Determining α and γ : Theory, 2004, hep-ph/0410371.
- [50] *BABAR*, B. Aubert *et al.*, Improved measurements of \mathcal{CP} -violating asymmetry amplitudes in $B^0 \rightarrow \pi^+ \pi^-$ decays, 2005, hep-ex/0501071.
- [51] Belle, K. Abe *et al.*, Phys. Rev. Lett. **93**, 021601 (2004), hep-ex/0401029.
- [52] R. Aleksan *et al.*, Phys. Lett. **B356**, 95 (1995), hep-ph/9506260.
- [53] Y. Grossman and H. R. Quinn, Phys. Rev. **D58**, 017504 (1998), hep-ph/9712306.

- [54] *BABAR*, B. Aubert *et al.*, Phys. Rev. Lett. **93**, 231801 (2004), hep-ex/0404029.
- [55] *BABAR*, B. Aubert *et al.*, Phys. Rev. Lett. **91**, 171802 (2003), hep-ex/0307026.
- [56] Belle, J. Zhang *et al.*, Phys. Rev. Lett. **91**, 221801 (2003), hep-ex/0306007.
- [57] *BABAR*, L. Roos, Measurements related to CKM angle α in *BABAR*, 2004, hep-ex/0407051.
- [58] Belle, K. Abe *et al.*, Measurement of $\phi(3)$ with Dalitz plot analysis of $B^+ \rightarrow D^* K^\pm$ decay at Belle, 2004, hep-ex/0411049.
- [59] *BABAR*, S. Schrenk, Nucl. Phys. Proc. Suppl. **142**, 279 (2005).
- [60] LHCb, M. Szczekowski, LHCb experiment sensitivity to CKM phases and new physics, 2004, Prepared for 12th International Workshop on Deep Inelastic Scattering (DIS 2004), Strbske Pleso, Slovakia.
- [61] LHCb, J. Rademacker, LHCb: Status and physics prospects, 2005, hep-ex/0503001.
- [62] S. Dimopoulos and H. Georgi, Phys. Lett. **B117**, 287 (1982).
- [63] Z. G. Berezhiani and G. R. Dvali, Bull. Lebedev Phys. Inst. **5**, 55 (1989).
- [64] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, Phys. Lett. **B429**, 263 (1998), hep-ph/9803315.
- [65] L. Alvarez-Gaume, J. Polchinski, and M. B. Wise, Nucl. Phys. **B221**, 495 (1983).
- [66] R. Aleksan, B. Kayser, and D. London, Phys. Rev. Lett. **73**, 18 (1994), hep-ph/9403341.
- [67] J. P. Silva and L. Wolfenstein, Phys. Rev. **D55**, 5331 (1997), hep-ph/9610208.
- [68] J. F. Gunion and H. E. Haber, Nucl. Phys. **B272**, 1 (1986).
- [69] B. Gittelmann and P. Skubic, (1980), Published in Proceedings of 1980 Moriond Conference, Les Arcs, France.
- [70] Aleph, D. Decamp *et al.*, Phys. Lett. **B258**, 236 (1991).
- [71] D. Andrews *et al.*, Phys. Rev. Lett. **45**, 219 (1980).

- [72] Delphi, P. A. Aarnio *et al.*, Nucl. Instrum. Meth. **A303**, 233 (1991).
- [73] L3, O. Adriani *et al.*, Phys. Rept. **236**, 1 (1993).
- [74] OPAL, K. Ackerstaff *et al.*, Eur. Phys. J. **C5**, 379 (1998), hep-ex/9801022.
- [75] ALEPH, DELPHI, L3, OPAL, Combining heavy flavour electroweak measurements at LEP, 1996.
- [76] Delphi, A. Borgland *et al.*, Prepared for International Europhysics Conference on High- Energy Physics (EPS-HEP 99), Tampere, Finland, 15-21 Jul 1999.
- [77] BABAR, B. Aubert *et al.*, Nucl. Instrum. Meth. **A479**, 1 (2002), hep-ex/0105044.
- [78] Belle, Y. Chao *et al.*, Nucl. Instrum. Meth. **A479**, 117 (2002).
- [79] BABAR, B. Aubert *et al.*, Phys. Rev. Lett. **86**, 2515 (2001), hep-ex/0102030.
- [80] Belle, A. Abashian *et al.*, Phys. Rev. Lett. **86**, 2509 (2001), hep-ex/0102018.
- [81] Belle, K. Abe *et al.*, Phys. Rev. Lett. **87**, 091802 (2001), hep-ex/0107061.
- [82] CKMfitter Group, J. Charles *et al.*, Eur. Phys. J. **C41**, 1 (2005), hep-ph/0406184.
- [83] CDF - Run II and D0 - Run II, S. D'Auria, (2005), hep-ex/0506049.
- [84] G. P. Salam, Acta Phys. Polon. **B33**, 2791 (2002), hep-ph/0207147.
- [85] G. Brianti, Phys. Rept. **403-404**, 349 (2004).
- [86] ATLAS and CMS, D. Pallin, (2002), Prepared for 14th Topical Conference on Hadron Collider Physics, Karlsruhe, Germany.
- [87] ATLAS, G. Azuelos, (1996), Prepared for 11th Topical Workshop on Proton-Antiproton Collider Physics, Padua, Abano Terme, Italy.
- [88] LHC, LHC design report. Vol. I: The LHC main ring, 2004, CERN-2004-003.
- [89] ATLAS, W. W. Armstrong *et al.*, ATLAS: Technical proposal for a general-purpose pp experiment at the Large Hadron Collider at CERN, 1994, CERN-LHCC-94-43.
- [90] CMS, M. Della Negra *et al.*, CMS: The Compact Muon Solenoid: Letter of intent for a general purpose detector at the LHC, 1992, CERN-LHCC-92-3.
- [91] ALICE., P. Giubellino, Nucl. Instrum. Meth. **A344**, 27 (1994).

- [92] LHCb, S. Amato *et al.*, LHCb technical proposal, 1998, CERN-LHCC-98-4.
- [93] L. Maiani, Int. J. Mod. Phys. **A19**, 1229 (2004).
- [94] P. Nason *et al.*, Bottom production, 1999, hep-ph/0003142.
- [95] LHCb, LHCb reoptimized detector design and performance, 2003, CERN-LHCC-2003-030.
- [96] LHCb, LHCb magnet: Technical design report, 2000, CERN-LHCC-2000-007.
- [97] LHCb, LHCb VELO TDR: Vertex locator. technical design report, 2001, CERN-LHCC-2001-011.
- [98] D. Green, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. **12**, 1 (2000).
- [99] LHCb, LHCb RICH: Technical design report, 2000, CERN-LHCC-2000-037.
- [100] S. J. E. Jolly, *Development of ring imaging Cherenkov detectors for the LHCb experiment*, PhD thesis, Imperial College, London, 2005, CERN-THESIS-2005-016.
- [101] LHCb, LHCb: Addendum to the LHCb RICH TDR. photon detectors for the LHCb RICH, 2003, CERN-LHCC-2003-059.
- [102] M. Campbell *et al.*, Nucl. Instrum. Meth. **A504**, 286 (2003).
- [103] M. Moritz *et al.*, IEEE Trans. Nucl. Sci. **51**, 1060 (2004).
- [104] LHCb, M. Adinolfi, AIP Conf. Proc. **722**, 119 (2004).
- [105] LHCb, M. Campbell *et al.*, Performance of hybrid photon detector prototypes with encapsulated silicon pixel detector and readout for the RICH counters of LHCb, 2002, CERN-LHCb-2002-048.
- [106] LHCb, LHCb: Inner tracker technical design report, 2002, CERN-LHCC-2002-029.
- [107] LHCb, U. Parzefall, Nucl. Phys. Proc. Suppl. **120**, 230 (2003).
- [108] LHCb, L. B. A. Hommels, (2005), CERN-LHCB-2005-014.
- [109] LHCb, LHCb calorimeters: Technical design report, 2000, CERN-LHCC-2000-036.

- [110] LHCb, LHCb trigger system technical design report, 2003, CERN-LHCC-2003-031.
- [111] HERA-B, E. I. Tarkovsky, Nucl. Instrum. Meth. **A379**, 515 (1996).
- [112] LHCb, LHCb muon system technical design report, 2001, CERN-LHCC-2001-010.
- [113] LHCb, LHCb: Addendum to the muon system technical design report, 2003, CERN-LHCC-2003-002.
- [114] LHCb, LHCb: Second addendum to the muon system technical design report, 2005, CERN-LHCC-2005-12.
- [115] LHCb, H. Dijkstra, Nucl. Phys. Proc. Suppl. **120**, 160 (2003).
- [116] LHCb, P. R. Barbosa-Marinho *et al.*, LHCb online system technical design report: Data acquisition and experiment control, 2001, CERN-LHCC-2001-040.
- [117] LHCb, G. Corti, Nucl. Instrum. Meth. **A462**, 265 (2001).
- [118] G. Barrand *et al.*, Gaudi - the software architecture and framework for building LHCb data processing applications, 2000, Prepared for International Conference on Computing in High-Energy Physics and Nuclear Physics (CHEP 2000), Padova, Italy.
- [119] Gaudi data analysis framework, <http://cern.ch/proj-gaudi>.
- [120] LHCb, LHCb Dirac distributed production system, <http://cern.ch/lhcb-comp/Production>.
- [121] N. Brook *et al.*, Nucl. Instrum. Meth. **A502**, 334 (2003).
- [122] K. Harrison *et al.*, ECONF **C0303241**, TUCT002 (2003), cs.se/0306085.
- [123] Pythia 6 event generator, [http://www.thep.lu.se/\\$\protect\protect\protect\\$\relax\protect{\protect\protect\\$\relax{\protect\protect\\$\relax{\protect\kern+.1667em\relax\@oldsim\protect\kern+.1667em\relax}}}\\$\torbjorn/pythiaaux/introduction](http://www.thep.lu.se/$\protect\protect\protect$\relax\protect{\protect\protect$\relax{\protect\protect$\relax{\protect\kern+.1667em\relax\@oldsim\protect\kern+.1667em\relax}}}$\torbjorn/pythiaaux/introduction).
- [124] LHCb, LHCb Gauss simulation framework, <http://cern.ch/lhcb-comp/Simulation/>.
- [125] Geant 4 physics simulation tool, <http://cern.ch/wwwasd/geant4/>.

- [126] LHCb, LHCb Boole digitisation framework, <http://cern.ch/lhcb-comp/Digitization/>.
- [127] LHCb, LHCb Brunel reconstruction framework, <http://cern.ch/lhcb-comp/Reconstruction/>.
- [128] LHCb, LHCb DaVinci analysis framework, <http://cern.ch/lhcb-comp/Analysis>.
- [129] LHCb, LHCb Panoramix visualisation framework, <http://cern.ch/lhcb-comp/Frameworks/Visualisation>.
- [130] OpenScientist framework, <http://www.lal.in2p3.fr/OpenScientist/>.
- [131] LHCb, R. Forty, Nucl. Instrum. Meth. **A433**, 257 (1999).
- [132] LHCb, R. Forty and O. Schneider, RICH pattern recognition, 1998, LHCb-98-040.
- [133] LHCb, A. Schnig, Particle identification in the RICH detectors and study of impact parameters, 1997, LHCb-97-018.
- [134] LHCb, C. Jones, RICH reconstruction software for the HLT, Presentation to LHCb RICH group, 2004.
- [135] LHCb, R. Forty and M. Patel, RICH particle ID algorithms for the HLT, Presentation to LHCb RICH group, 2004.
- [136] LHCb, M. Benayoun and C. Jones, RICH reconstruction and particle identification using ring fit methods: Application to the RICH2 detector, 2004, LHCb-2004-057.
- [137] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms* (Cambridge University Press, 2003), Available from <http://www.inference.phy.cam.ac.uk/mackay/itila/>.
- [138] C. G. Lester, (2005), physics/0508034.
- [139] LHCb, M. Moritz *et al.*, IEEE Transactions on Nuclear Science **51** (2004).
- [140] C. Jones, 2005, Private communication.
- [141] LHCb, M. Needham, Electron reconstruction studies, 2001, LHCb-2001-102.
- [142] R. Fleischer, Flavour physics and \mathcal{CP} violation, 2004, hep-ph/0405091.

- [143] D. Wyler, Flavour physics: What, why and how, 2002, Published in Processings of 2001 European School of High-Energy Physics, Beatenberg, Switzerland.
- [144] R. Fleischer, Lect. Notes Phys. **647**, 42 (2004), hep-ph/0210323.
- [145] M. Beneke, G. Buchalla, M. Neubert, and C. T. Sachrajda, Phys. Rev. Lett. **83**, 1914 (1999), hep-ph/9905312.
- [146] M. Beneke, G. Buchalla, M. Neubert, and C. T. Sachrajda, Nucl. Phys. **B591**, 313 (2000), hep-ph/0006124.
- [147] M. Beneke, G. Buchalla, M. Neubert, and C. T. Sachrajda, Nucl. Phys. **B606**, 245 (2001), hep-ph/0104110.
- [148] M. Beneke and M. Neubert, Nucl. Phys. **B675**, 333 (2003), hep-ph/0308039.
- [149] CLEO, Y. Kwon *et al.*, Study of charmless hadronic B decays into the final states $K \pi$, $\pi \pi$, and $K K$, with the first observation of $B \rightarrow \pi^+ \pi^-$ and $B \rightarrow K^0 \pi^0$, 1999, hep-ex/9908039.
- [150] CLEO, A. Bornheim *et al.*, Phys. Rev. **D68**, 052002 (2003), hep-ex/0302026.
- [151] R. Fleischer, Phys. Lett. **B365**, 399 (1996), hep-ph/9509204.
- [152] R. Fleischer and T. Mannel, Phys. Rev. **D57**, 2752 (1998), hep-ph/9704423.
- [153] CLEO, A. Bornheim *et al.*, Phys. Rev. **D68**, 052002 (2003), hep-ex/0302026.
- [154] BABAR, B. Aubert *et al.*, (2002), hep-ex/0207065.
- [155] BABAR, B. Aubert *et al.*, Phys. Rev. Lett. **91**, 021801 (2003), hep-ex/0303028.
- [156] T. Tomura, (2003), hep-ex/0305036.
- [157] A. J. Buras, R. Fleischer, S. Recksiegel, and F. Schwab, Eur. Phys. J. **C32**, 45 (2003), hep-ph/0309012.
- [158] R. Fleischer, Eur. Phys. J. **C6**, 451 (1999), hep-ph/9802433.
- [159] Belle, Y. Chao *et al.*, Phys. Rev. **D69**, 111102 (2004), hep-ex/0311061.
- [160] A. J. Buras and R. Fleischer, Eur. Phys. J. **C11**, 93 (1999), hep-ph/9810260.
- [161] A. J. Buras, R. Fleischer, S. Recksiegel, and F. Schwab, Nucl. Phys. **B697**, 133 (2004), hep-ph/0402112.

- [162] R. Fleischer, Phys. Lett. **B435**, 221 (1998), hep-ph/9804319.
- [163] A. Ali, Review of heavy quark physics: Theory, 2004, hep-ph/0412128.
- [164] V. Gogohia, Intrinsically nonperturbative QCD. I: The yang-mills sector, 2003, hep-ph/0311061.
- [165] J. M. Gerard and J. Weyers, Eur. Phys. J. **C7**, 1 (1999), hep-ph/9711469.
- [166] M. Neubert, Phys. Lett. **B424**, 152 (1998), hep-ph/9712224.
- [167] A. F. Falk, A. L. Kagan, Y. Nir, and A. A. Petrov, Phys. Rev. **D57**, 4290 (1998), hep-ph/9712225.
- [168] D. Atwood and A. Soni, Phys. Rev. **D58**, 036005 (1998), hep-ph/9712287.
- [169] M. Beneke and T. Feldmann, Nucl. Phys. **B685**, 249 (2004), hep-ph/0311335.
- [170] Y. Y. Keum and A. I. Sanda, Phys. Rev. **D67**, 054009 (2003), hep-ph/0209014.
- [171] Y.-Y. Keum and H.-n. Li, Phys. Rev. **D63**, 074006 (2001), hep-ph/0006001.
- [172] QQ program, <http://www.lns.cornell.edu/public/CLEO/soft/QQ>.
- [173] GEANT 3 detector description and simulation tool, <http://www.lns.cornell.edu/public/CLEO/soft/QQ>, 1994.
- [174] SICB program, <http://cern.ch/lhcb-comp/SICB>.
- [175] LHCb, Brunel v17r4 framework, http://cern.ch/lhcb-release-area/BRUNEL/release_notes.html#v17r4.
- [176] LHCb, Y. Xie, K_S^0 reconstruction, 2003, LHCb-2003-088.
- [177] R. Brun and F. Rademakers, Nucl. Instrum. Meth. **A389**, 81 (1997).
- [178] R. Brun *et al.*, ROOT data analysis system, <http://root.cern.ch>.
- [179] R. Barlow, ECONF **C030908**, MOAT002 (2003), physics/0311105.
- [180] C. M. Bishop, *Neural Networks for Pattern Recognition* (Oxford University Press, 1995).
- [181] J. W. Miskin, *Ensemble Learning for Independent Component Analysis*, PhD thesis, University of Cambridge, 2000.

-
- [182] Delphi, D. Jeans, A discussion of discriminant techniques, 2001, Delphi-2001-135.
 - [183] B. Gough *et al.*, *GNU Scientific Library Reference Manual: Multidimensional Minimisation*.
 - [184] R. Barlow, *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences* (John Wiley & Sons, 1993).
 - [185] P. K. Sinervo, Signal significance in particle physics., 2002, hep-ex/0208005.
 - [186] R. Barlow, 2004, Private communication.
 - [187] J. Dennis Jr. and D. J. Woods, Optimization on microcomputers: The Nelder-Mead simplex algorithm, ARO Workshop on Microcomputers, 1985.
 - [188] *Bottom Production*, 1999.
 - [189] G. J. Feldman and R. D. Cousins, Phys. Rev. **D57**, 3873 (1998), physics/9711021.
 - [190] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, Journal of Chemical Physics **21**, 1087 (1953).
 - [191] W. K. Hastings, Biometrika **57**, 97 (1970).
 - [192] J. A. Nelder and R. Mead, The Computer Journal **7**, 308 (1965).
 - [193] M. Goosens and E. van Herwijnen, CERN Computer Newsletter (1995).

List of Figures

1.1	Neutral meson mixing plots for $K^0-\bar{K}^0$, $B_d^0-\bar{B}_d^0$ and $B_s^0-\bar{B}_s^0$ mixing	5
1.2	The six CKM “unitarity triangles”	17
1.3	The two “unsquashed” CKM unitarity triangles	18
1.4	Leading order Feynman diagrams for Standard Model $B-\bar{B}$ mixing	20
1.5	Tree and penguin Feynman diagram topologies	22
1.6	Contributions to the indirect fit of CKM parameters.	25
1.7	Global indirect fits to the CKM parameters	25
1.8	CKM Fitter confidence interval for α combined fit.	29
2.1	LHC beam-line cross-section	37
2.2	Locations of the four main LHC experimental sites around the LHC ring.	37
2.3	$b\bar{b}$ production by gluon fusion	38
2.4	Polar angle correlations for $b\bar{b}$ production from proton collisions	39
2.5	Cross-section view of LHCb, cut in the non-bending $y-z$ plane	40
2.6	Probabilities of given numbers of pp interactions per bunch crossing	41
2.7	The LHCb dipole magnet.	43
2.8	The layout of the VELO layout in the non-bending $y-z$ plane	44
2.9	The VELO sensors and the primary vacuum RF foil.	45
2.10	VELO r - and ϕ -sensors.	46

2.11	$\pi^+\pi^-$ invariant-mass spectra with and without RICH particle ID	47
2.12	Dipole arrangements in Čerenkov radiator for $v < c/n$ and $v \geq c/n$	48
2.13	Huygens construction for Čerenkov photon emission	49
2.14	Polar angle vs. momentum coverage of RICH1 and RICH2	51
2.15	The LHCb RICH detectors	52
2.16	Schematic layout of a RICH system HPD	53
2.17	Reconstructed tracks in LHCb for a b-production event	54
2.18	The fine structure of the LHCb calorimeters	57
2.19	Schematic diagram of a muon system MWPC.	59
2.20	Architecture of the LHCb L0 and L1 trigger system	60
2.21	Examples of the L1 trigger vertical–diagonal discriminant cut	62
2.22	The logical structure of the LHCb software framework	63
3.1	Number of activated pixels per event in RICH1 and RICH2.	71
3.2	Contributions to Čerenkov angle resolution in the RICH1 C_4F_{10} radiator . .	72
3.3	The mean particle ID likelihood π –K separation for true pions	74
3.4	Standard-reconstruction particle ID efficiencies vs. momentum	74
3.5	Standard-reconstructed Čerenkov rings in RICH2	76
3.6	Čerenkov ring distortions in RICH2	78
3.7	True Čerenkov emission angle distribution for RICH2	80
3.8	The Čerenkov ring “Yorkshire pudding” function	82
3.9	Two different contexts in which RICH hits can be interpreted	84
3.10	Čerenkov rings reconstructed using the trackless ring-finder in RICH2 . .	88
3.11	Characteristic numbers of rings and pixels in RICH2	89
3.12	Ring and event efficiencies and purities in RICH2	92

3.13	Pull distribution of the Čerenkov emission angle in RICH2	93
3.14	Magnitudes of tracked/untracked ring-centre position error	94
3.15	MC particle ID for tracks from tracked and untracked Markov rings . . .	95
3.16	Origins of MC tracks with tracked/untracked Markov rings	96
3.17	Origin z -positions of MC tracks with tracked/untracked Markov rings . .	97
3.18	Particle ID efficiencies as a function of the acceptance parameter a_{prob} . .	99
3.19	Background contributions for standard/trackless estimators	101
3.20	RICH2 particle ID efficiencies vs. no. added background rings	102
3.21	RICH2 particle ID efficiencies vs. no. added background rings (tight) . .	104
3.22	PID efficiency and mis-ID rates for standard/Markov bg. estimators . . .	105
4.1	Schematic representation of the QCD factorisation formula	112
4.2	Fits to $\pi\pi$ and $K\pi$ \mathcal{CP} asymmetries using QCD factorisation	114
4.3	Feynman diagram topologies contributing to $B_d^0 \rightarrow K^+ \pi^-$ decays	115
4.4	Feynman diagram topologies contributing to $B^+ \rightarrow K^0 \pi^+$ decays	115
4.5	The dependence of R_{min} on γ for various values of A_0	121
4.6	r upper and lower bounds as a function of γ and R	122
4.7	r upper and lower bounds as a function of γ and A_0	123
4.8	Triangle construction for the “mixed” $B \rightarrow K \pi$ strategy	123
4.9	An explicit re-scattering contribution to $B^+ \rightarrow K^0 \pi^+$	124
4.10	Effects of final state re-scattering on the R_{min} bound.	125
4.11	Re-scattering topologies for general $B \rightarrow K \pi$ decays.	127
5.1	$B_d^0 \rightarrow K^+ \pi^-$ event topology	135
5.2	PID confidence levels for K^+ and π^- particle candidates	136
5.3	Raw invariant-mass of B_d^0 candidates before the first B_d^0 mass cut	137

5.4	$B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ event topology	138
5.5	Track-type definitions relative to the upstream detector components. . .	139
5.6	K_S^0 mass distributions (in MeV/c^2) from 50k signal events	140
5.7	Masses of $K_S^0 \pi^+$ combinations before the first B^+ mass cut	143
5.8	PID confidence levels for π^+ candidates in $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ events . .	144
5.9	No. PV candidates in $B_d^0 \rightarrow K^+ \pi^-$ and $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ signal events .	145
5.10	Properties of primary vertex candidates in $B_d^0 \rightarrow K^+ \pi^-$ signal events . . .	146
5.11	π^- cut variable distributions for $B_d^0 \rightarrow K^+ \pi^-$. The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.	150
5.12	K^+ variable distributions for $B_d^0 \rightarrow K^+ \pi^-$. The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.	151
5.13	B_d^0 cut variable distributions for $B_d^0 \rightarrow K^+ \pi^-$. The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.	152
5.14	Lone π^+ cut variable distributions for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$. The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.	154
5.15	K_S^0 cut variable distributions for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$. The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.	156
5.16	B^+ cut variable distributions for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$. The filled histograms represent the background and the points represent the signal. The normalisation is arbitrary.	160
5.17	Score progression and purity-efficiency distributions from cut tuning . . .	172
5.18	Systematic shifts in no. events passing $B_d^0 \rightarrow K^+ \pi^-$ cuts	181
5.19	Systematic shifts in no. events passing $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ LL cuts . . .	186

5.20	Systematic shifts in no. events passing $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ DD cuts . . .	188
6.1	Variation of R_{\min} with γ for various values of A_0	201
6.2	Limits on r as a function of γ , for various values of R	202
6.3	Limits on r as a function of γ , for various values of A_0	203
A.1	Metropolis–Hastings proposal and unnormalised target distributions . . .	2

List of Tables

2.1	Various characteristics of the RICH radiator media at STP	51
2.2	Characteristics of the trigger levels and offline analysis.	60
3.1	Contributions to emission angle resolution in the RICH detectors	70
3.2	Particle ID matrix for the standard RICH particle ID system	73
3.3	Particle ID rates for different combinations of RICH detectors	75
5.1	Cuts used by the K_S^0 reconstruction package	141
5.2	Reconstruction performance for K_S^0 in $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ decays	141
5.3	Cuts used for $B_d^0 \rightarrow K^+ \pi^-$ pre-selection	147
5.4	Cuts used for $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ pre-selection	147
5.5	Cut variables retrieved from $B_d^0 \rightarrow K^+ \pi^-$ reconstruction	148
5.6	Cut variables retrieved from $B^+ \rightarrow K_S^0 \pi^+$ reconstruction	149
5.7	Examples of functions used to characterise selection performance.	165
5.8	Hadronisation fractions from a b-quark to the different b-hadrons.	167
5.9	$B_d^0 \rightarrow K^+ \pi^-$ tuned selection cuts (1)	173
5.10	$B_d^0 \rightarrow K^+ \pi^-$ tuned selection cuts (2)	174
5.11	$B_d^0 \rightarrow K^+ \pi^-$ tuned selection cuts (3)	174
5.12	$B_d^0 \rightarrow K^+ \pi^-$ tuned selection cuts (4)	174

5.13	$B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ tuned selection cuts (1)	175
5.14	$B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ tuned selection cuts (2)	176
5.15	$B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ tuned selection cuts (3)	176
5.16	$B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ tuned selection cuts (4)	177
5.17	Number of candidates at each stage of $B_d^0 \rightarrow K^+ \pi^-$ selection	179
5.18	Number of candidates at each stage of $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ selection	184
6.1	Total selection efficiencies for $B_d^0 \rightarrow K^+ \pi^-$ and $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$	193
6.2	Annual signal yield estimates for $B_d^0 \rightarrow K^+ \pi^-$ and $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$	194
6.3	Selection efficiency components for $B_d^0 \rightarrow K^+ \pi^-$ and $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$	195
6.4	$B_{b\bar{b}}/S$ estimates for $B_d^0 \rightarrow K^+ \pi^-$ and $B^+ \rightarrow K_S^0 (\pi^+ \pi^-) \pi^+$ channels	195
6.5	Values of $B \rightarrow K \pi$ \mathcal{CP} -asymmetries from b-factory measurements	198
6.6	LHCb one-year sensitivities to γ for various values of r	204