$B_s^0 \rightarrow J/\psi\phi$ with LHC-ATLAS: simulations and sensitivity studies

James Richard Catmore

Department of Physics
Lancaster University

Thesis submitted for the degree of Doctor of Philosophy
in the subject of Physics.

June 2007

Benedicite omnia opera Domini Domino: laudate et superexaltate eum in sæcula!

Acknowledgements

I would like to offer my thanks to the following:

- The Particle Physics and Astronomy Research Council for so generously funding my research and Long Term Attachment at CERN, and the Department of Physics at Lancaster University for hosting me.

- My supervisor Roger Jones, my technical supervisor Mária Smižanská and my shadow supervisor Gareth Hughes, for their continual help, support, advice and where necessary, chivvying. I wouldn't have managed without you!

- Eva Bouhova-Thacker and Rob Henderson for spending countless hours with me explaining the finer points of track reconstruction and physics analysis, and also for many enjoyable lunches at InfoLab21 Cafe. Also to Tim Brodbeck for helping me to make sense of Minuit, if such a thing is possible.

- The many ATLAS software experts whom I made a habit of pestering on a regular basis, including: Fred Akesson, Ketevi Assamagan, Frederic Brochu, Ed Moyse, Peter Sherwood, Andi Wildauer.

- Alessandro De Salvo, for showing me how to drive the Grid.

- The delightful, wonderful people with whom I have spent the last three years both at the department in Lancaster and at CERN, with whom I hiked in the summer and skied in the winter, and who have put up with my continual moaning about ATLAS software, especially (in strictly alphabetical order): Adam, Alastair, Alex, Anne, Christian, Dan H, Dan L, Darren, Dave C, Dave F, Ed, James (a.k.a. the taxi service), Janice, Ignacio, Laila, Leo, Marcus, Pavel, Rob, Ruth. Also to my house-mate Rich who inevitably bore the brunt of pre- and post-Viva paroxysms.

- My parents, for enthusiastically supporting me in everything I do

Abstract

The decay $B_s^0 \to J/\psi\phi$ is often referred to as a "golden decay mode" on account of its potential for measuring the weak mixing parameters $\Delta\Gamma_s$ and ϕ_s with little theoretical uncertainty. Being a self-conjugate process, an angular analysis is required to separate out the different contributions to the overall decay amplitude. This work describes a series of simulations of these measurements in the LHC-ATLAS detector, leading to an assessment of the precision which could be obtained after various integrated luminosities.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Since the time of the philosophers of ancient Greece and China, mankind has sought understanding of the fundamental, indivisible entities which constitute the Universe. Through a combination of observation, experiment, logic, inspired guesswork and latterly, abstract mathematics, a formidable body of knowledge has been assembled. Currently, this culminates in the **Standard Model of Particle Physics**, a theory which defines the basic constituents of matter, and describes the interactions which govern them.

The Standard Model has proved itself to be remarkably resilient. It has existed in its current form since the early 1970s, and despite thousands of experimental tests made across a wide range of energies, few, if any, deviations have been observed. Every particle predicted to exist by the model has since "checked in" at one detector or another, most recently the tau neutrino at Fermilab [1]. The only piece of the jigsaw puzzle which remains to be found is the Higgs Boson - although hints were seen at the Large Electron Positron collider at CERN, the statistics were far too low to claim a discovery [2].

So maybe this is the end of the story. Perhaps humanity has for the first time in its history gained complete knowledge of the building blocks of the universe, and perhaps the only remaining mysteries in physics concern the emergent properties of collections of these constituents.

If this is so, the laws of the Universe will be less satisfactory than many hope. The Standard Model, despite its durability, is in many ways a deeply unsatisfactory theory. Firstly, it contains nineteen free parameters - numbers that can be measured and fed into the model, but not derived from more fundamental physics. Why are these numbers as they are? Why not one percent more or three percent less? Secondly, the gravitational interaction is not accounted for at all in the theory. Thirdly, the model has a clear structure - particles are arranged into tiers or

generations according to their mass - and there is no clear reason as to why this should be so. Fourthly, the Higgs Boson has not actually been observed beyond peradventure. Fifth, are the members of the model really *fundamental* or do internal degrees of freedom exist? Finally, the model as it stands does not fit comfortably with cosmological observations, especially with regards to the current imbalance between matter and anti-matter in the observed universe - whilst such asymmetries are predicted in the model, their magnitudes are far too small to be responsible for the apparent disappearance of anti-matter just after the Big Bang.

In consequence, governments across the world are continuing to commit billions of pounds to projects which seek evidence for physics beyond the Standard Model. The Large Hadron Collider (LHC) [3], at CERN on the Franco-Swiss border between Geneva and Gex, is the most ambitious of these projects. Due to begin operations in mid-2007, this proton-proton accelerator and its four detectors will observe the remnants of hadron collisions occurring at unprecedented luminosity and centre-of-mass energies. The confirmation of the existence of the Higgs Boson and a precise measurement of its mass is but a part of a diverse scientific programme. Ultimately it is hoped that the LHC will show the Standard Model to be merely an approximation to the truth, and provide an insight into that which lies beyond it.

This thesis is a study into the potential for one of the four detectors - ATLAS[1] [4] to detect signatures of new physics from one particular channel - the decay of neutral B_s mesons to the final state $\mu^+\mu^-K^+K^-$ via the intermediate state $J/\psi\phi$. The study is necessarily based entirely on computer simulations.

The first chapter gives a brief overview of the Standard Model, paying particular attention to the concept of *discrete symmetries*. The phenomenology of neutral B-mesons is also considered here. Chapter three lays out the formalism for describing decays of particles with non-trivial spin configurations. The fourth and fifth chapters describe the hardware and computing infrastructure of the ATLAS detector. The sixth chapter takes the formalisms of earlier chapters and applies them to the decay channel of interest, and describes the implementation of this in simulation code. The seventh chapter describes the techniques for extracting the kinematics of the signal events from the background, and the eigth chapter, after presenting a method for extracting the physics parameters from this data, suggests a limit for the precision of such a measurement.

The work presented here can be assumed to be the author's, unless a specific citation is provided.

[1]A Toroidal LHC ApparatuS; also the god who, as a punishment for various misdeeds, was forced by Zeus to bear the heavens and the earth on his shoulders

Chapter 2

Theoretical Framework

2.1 Symmetry

Our ability to quantitatively comprehend nature depends crucially on the concept of *symmetry*. In the broadest possible terms, a physical system is symmetric with respect to a given transformation if, after the transformation, it behaves in identically the same way. The *conservation laws* of classical mechanics arise from the imposition of *continuous global* symmetries on the Lagrangian governing the system - in each case the postulation of invariance requires that a quantity exists which does not change with time. For instance, if one insists that the Lagrangian of a system remains invariant after a translation in space, the law of conservation of linear momentum immediately follows. Similarly, invariance with respect to rotations leads to angular momentum conservation, and energy conservation results from insisting that physical systems behave identically irrespective of translations in time. These relations are summarized in *Noether's Theorems* [6].

The Standard Model is constructed through the use of *local* symmetries (also referred to as **local gauge symmetries**. Imposing local symmetries on the quantized fields which describe the particles of the model leads to the emergence of *gauge fields*; the particles created by the operators of these fields turn out to be the force-mediating bosons (photons, W/Z bosons and gluons) [10].

The *discrete* symmetries denoted by $\mathscr{C}$, $\mathscr{P}$ and $\mathscr{T}$ are also of fundamental significance to physics.

- $\mathscr{C}$ - **charge conjugation** - replaces every particle in a system with its antiparticle.

- $\mathscr{P}$ - **parity inversion** - replaces each space co-ordinate $\hat{x}$ with its reflection $-\hat{x}$.

- $\mathscr{T}$ - **time reversal** - replaces the time co-ordinate t with $-t$. This amounts to a reversal of motion - a "re-winding" of time.

It is believed that symmetry under the combined operation $\mathscr{CPT}$ holds under all conditions, as it can be shown (e.g. [8]) that any system which did not behave identically after a $\mathscr{CPT}$ transformation would consequently violate Lorentz invariance. CPT-symmetry implies that the mass and width of a particle and its anti-particle are the same [9], and no deviations from this have been observed. Now that the manufacture of anti-hydrogen is possible (at CERN's anti-proton decelerator), opportunities have arisen to observe differences between the spectral lines of atom and anti-atom; any variance between the two would indicate CPT-asymmetry. The laser spectrometer ALPHA [5], currently under preparation at CERN, will clarify this position once it begins to take data.

For the first half of the twentieth century it was believed that all of nature was invariant to the single operations $\mathscr{C}$ and $\mathscr{P}$. However, a large body of experimental evidence (e.g. [7]) has since shown this to be a spectacularly incorrect assumption - whilst the strong and electromagnetic interactions do appear to be C- and P-invariant, the weak force is not - in fact it *maximally* violates C- and P-symmetry, in that right-handed neutrinos do not interact with the weak force at all.

Even when it became clear that the weak force violated C- and P-symmetry, there was some certainty that physics would be invariant under the combined operation $\mathscr{CP}$. Alas, in 1964, this was also proved to be an incorrect assumption when neutral kaons were observed to decay to final states whose CP eigenstates were different from the parent kaon [12]. An interesting corollary to CP-violation is the existence of an *absolute*, *convention-independent* difference between matter and anti-matter. C-violation allows us to distinguish between the two only after the imposition of a left/right-handed convention, whereas CP-violating processes permit them to be identified naturally.

For this reason, CP-violation is of cosmological interest. The Universe as we know it today is entirely made up of *matter* - it is thought that if anti-matter galaxies did exist, we would by now have observed a collision (and concomitant violent release of energy) between such a galaxy and a matter counterpart. Furthermore it is difficult to imagine a mechanism by which such galaxies could have formed in the first place. In consequence it is thought that some process occurred during the big bang which caused a slight excess of matter over anti-matter to be produced, the modern observed Universe being made up of this excess. It was pointed out by A. Sakharov [11] that CP-violation could provide such a mechanism, albeit in far greater strength than has yet been observed.

Although the observation of new sources of CP-violation will be of great interest for these cosmological reasons, this thesis is not primarily concerned with CP-violation *per se* - rather, it considers the possibility of observing additional CP-violation in a known decay process (namely, the decay of neutral B-mesons). This would immediately indicate that physics processes not described by the Standard Model were occurring. The exact nature of these processes would not be revealed - but it would be an unambiguous signal that they were there. The remainder of this chapter lays down the formal structure of the Standard Model, considers the quantum mechanical processes governing neutral mesons, and finally describes how "New Physics" could cause additional CP-violation, and consequently observable evidence, in these systems.

2.2 The Standard Model of Particle Physics

The Standard Model is constructed through use of a gauge symmetry described by the group

$$\mathscr{G}_{SM} = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{2.1}$$

and the spontaneous breaking of this symmetry described by

$$\mathscr{G}_{SM} \to SU(3)_C \otimes U(1)_Q \tag{2.2}$$

Here, Y and Q represent weak hypercharge and electric charge respectively, whilst C represents colour charge. The imposition of gauge symmetry on the group $\mathscr{G}_{SM}$ generates twelve gauge bosons: the massive $W^{\pm}$, Z^0 and the massless photon γ, which mediate electroweak interactions; and eight massless gluons, which carry the strong force and which are described by quantum chromodynamics (QCD). The spontaneous breaking of the symmetry of $\mathscr{G}_{SM}$ is the mechanism by which all of the Standard Model particles are endowed with mass; the side-effect of this is to create an additional scalar field, the operators of which generate a scalar boson referred to as the **Higgs Boson**.

The fermion fields enter the model as *representations* of the Standard Model gauge group $\mathscr{G}_{SM}$. There are five such representations in the Standard Model - one for the left-handed components of the quark fields, two for the right handed quarks, and one each for the left and right handed lepton fields. These are either triplets or singlets of $SU(3)$ and either doublets or singlets of $SU(2)$, depending on handedness.

	$SU(3)$	$SU(2)$
Quarks	$q_{L,i} = \begin{pmatrix} q_{L,r} \\ q_{L,b} \\ q_{L,g} \end{pmatrix}$	$q_{L,i} = \begin{pmatrix} u_{L,i} \\ d_{L,i} \end{pmatrix}$
	$u_{R,i} = \begin{pmatrix} u_{R,r} \\ u_{R,b} \\ u_{R,g} \end{pmatrix}$	$u_{R,i} = \begin{pmatrix} u_{R,i} \end{pmatrix}$
	$d_{R,i} = \begin{pmatrix} d_{R,r} \\ d_{R,b} \\ d_{R,g} \end{pmatrix}$	$d_{R,i} = \begin{pmatrix} d_{R,i} \end{pmatrix}$
Leptons	$L_{L,i} = \begin{pmatrix} L_{L,i} \end{pmatrix}$	$L_{L,i} = \begin{pmatrix} \nu_{L,i} \\ l_{L,i} \end{pmatrix}$
	$l_{R,i} = \begin{pmatrix} l_{R,i} \end{pmatrix}$	$l_{R,i} = \begin{pmatrix} l_{R,i} \end{pmatrix}$

In the table, $q = u, d$, $L = l, \nu$ and L, R represent left and right handedness respectively. The i index runs from 1 to 3, indicating the generation number (so $\{u_{1,2,3}\} = \{u, c, t\}$, $\{d_{1,2,3}\} = \{d, s, b\}$, $\{l_{1,2,3}\} = \{e, \mu, \tau\}$ and $\{\nu_{1,2,3}\} = \{\nu_e, \nu_\mu, \nu_\tau\}$).

From the table, we can note the following:

- The quark fields split into three states under $SU(3)$ transformations - these are the so-called "colour triplets". Consequently in strong interactions each quark is endowed with an additional quantum descriptor known as "colour"; this can be one of "red", "blue" or "green".

- The left-handed component of the fermion fields split into doublets under $SU(2)$ transformations, whereas the right-handed fields are singlets. This reflects the (experimental) fact that there are no right-handed neutrinos.

- The Standard Model provides for three generations. There is in fact no intrinsic reason why there should not be more, but there cannot be fewer.

The only member of the Standard Model "zoo" which has not yet been observed experimentally is the Higgs Boson, although its existence would seem to be necessary if the ideas of gauge symmetry are valid; introducing mass in any other way than symmetry breaking leads to a theory which is not gauge invariant [10]

2.2.1 The CKM Quark Mixing Matrix

The quark doublets as they appear in the Standard Model are in the electroweak (*interaction*) basis, and do not have definite mass. In the mass basis, the up-type quarks are paired with a state which is a mixture of down-type quarks from across the generations [23]. These flavour eigenstates are rotated into mass eigenstates by means of a unitary matrix, whose dimension is determined by the number of Standard Model generations [13]. For a two-generation Standard Model the matrix is of dimension 2×2 and is called the **Cabibbo** matrix. For the existing three-generation model, the resulting 3×3 matrix is referred to as the **Cabibbo-Kobayashi-Maskawa** (CKM) mixing matrix, given by

$$\begin{pmatrix} d_I \\ s_I \\ b_I \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d_M \\ s_M \\ b_M \end{pmatrix} \tag{2.3}$$

where I and M refer to the interaction and mass bases respectively. Using this matrix one can write the Lagrangian for the interaction between the W_μ field corresponding to the W-boson and the quark fields (the weak charged-current interaction) as [10]:

$$\mathscr{L}_{qW} = -\frac{e}{\sqrt{2}\sin\theta_W} \begin{pmatrix} u_L^\dagger & c_L^\dagger & t_L^\dagger \end{pmatrix} \gamma^\mu \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} W_\mu^\dagger + HC \tag{2.4}$$

where θ_W is the **Cabbibo Angle**. The fact that the matrix is not diagonal indicates that the $W^\pm$ gauge bosons can couple to quarks of different generations, so weak charged-current interactions can change the flavour of the quarks - flavour is not conserved in the weak interaction. Such interactions are referred to as **flavour changing charged currents**. The strength of a given coupling is determined by the magnitude of a given CKM matrix element. It should be noted that no such expression exists for the Z-boson fields; these neutral particles cannot violate flavour conservation, so there are no **flavour changing neutral currents** at tree level in the Standard Model (they are introduced through loops at higher orders [13]). It should also be noted that the organization of the elements of the CKM matrix is not unique - it is normal, however, to list the elements in order of quark mass, as shown above.

It is through this matrix that CP-violating processes can be introduced into the Standard Model without making any adjustments to the rest of the model. We note that the W and quark fields transform under the $\mathscr{C}\mathscr{P}$ operations as follows [10]:

$$\mathscr{C}\mathscr{P}\mathrm{W}_0^+\mathscr{C}\mathscr{P}^\dagger = -W_0^- \tag{2.5}$$

$$\mathscr{C}\mathscr{P}\mathrm{W}_i^+\mathscr{C}\mathscr{P}^\dagger = W_i^- \tag{2.6}$$

$$\mathscr{C}\mathscr{P} q_L \mathscr{C}\mathscr{P}^\dagger = -i\sigma^2 q_L^* \tag{2.7}$$

$$\mathscr{C}\mathscr{P} q_R \mathscr{C}\mathscr{P}^\dagger = i\sigma^2 q_R^* \tag{2.8}$$

Now we re-write equation 2.4 as [10]

$$\mathscr{L}_{qW} = -\frac{e}{\sqrt{2}\sin\theta_W}\sum_{i,j}\left[u_{Li}^\dagger \sigma^\mu V_{ij} d_{Lj} W_\mu^+ + d_{Li}^\dagger \sigma^\mu V_{ij}^* u_{Lj} W_\mu^-\right] \tag{2.9}$$

Here, i, j run through the three generations of up- and down-type quarks, V_{ij} is the ij-th CKM matrix element and the σ are the Pauli spin matrices. The superscript dagger (†) and asterisk ($*$) have their usual meanings, namely, the Hermitian and complex conjugate. Applying the $\mathscr{C}\mathscr{P}$ operator we obtain:

$$\begin{aligned}
\mathscr{C}\mathscr{P}\mathscr{L}_{qw}\mathscr{C}\mathscr{P}^\dagger &= \frac{e}{\sqrt{2}\sin\theta_W}\mathscr{C}\mathscr{P}\sum_{i,j}\left[u_{Li}^\dagger \sigma^\mu V_{ij} d_{Lj} W_\mu^+ + d_{Li}^\dagger \sigma^\mu V_{ij}^* u_{Lj} W_\mu^-\right]\mathscr{C}\mathscr{P}^\dagger \\
&= \frac{e}{\sqrt{2}\sin\theta_W}\sum_{i,j}\left[-u_{Li}^T (\sigma^\mu)^T V_{ij} d_{Lj}^* W_\mu^- - d_{Lj}^T (\sigma^\mu)^T V_{ij}^* u_{Li}^* W_\mu^+\right] \\
&= \frac{e}{\sqrt{2}\sin\theta_W}\sum_{i,j}\left[d_{Lj}^\dagger \sigma^\mu V_{ij} u_{Li} W_\mu^- + u_{Li}^\dagger \sigma^\mu V_{ij}^* d_{Lj} W_\mu^+\right]
\end{aligned} \tag{2.10}$$

This can only be the same as the original expression if all elements of V_{ij} are real. Therefore, *CP violation can occur in the Standard Model if the CKM matrix contains a complex element.*

A general $n \times n$ complex matrix contains $2n^2$ parameters. The unitarity of the matrix reduces this number by half, to n^2 parameters. As the phases of the quark fields can be rotated freely, and since the overall phase is physically unimportant, $2n - 1$ relative phases can be removed, leaving $(n-1)^2$ parameters. Now an orthogonal $n \times n$ rotation matrix contains $\frac{1}{2}n(n-1)$ rotation angles [20], so the number of independent phases is given by

$$N_{\text{phases}} = N_{\text{all}} - N_{\text{angles}} = (n-1)^2 - \frac{1}{2}n(n-1) = \frac{1}{2}(n-1)(n-2) \tag{2.11}$$

In the case of two generations of quarks, $n = 2$ so $N_{\text{phases}} = 0$. For three quark generations, $N_{\text{phases}} = 1$. The number of independent phases increases rapidly after this (3 for $n = 4$, 6 for $n = 5$).

The implication of this, therefore, is that a three-generation Standard Model can admit CP-violation through a single independent phase in the CKM matrix, whereas a two-generation model could not support such phenomena. In 1964 the third generation was not known to physics, so the observation of CP-violating processes was a powerful indicator of the existence of heavier quarks.

The magnitudes of the elements of the matrix are known with varying degrees of precision. The current state of knowledge is summarized in [107]. The unitarity of the matrix, formally expressed as

$$\mathbf{V}_{CKM}^{\dagger} \cdot \mathbf{V}_{CKM} = \mathbf{1} = \mathbf{V}_{CKM} \cdot \mathbf{V}_{CKM}^{\dagger} \tag{2.12}$$

leads to a set of twelve equations which provide a powerful constraint on the values of the individual elements:

$$\begin{aligned}
|V_{ui}|^2 + |V_{ci}|^2 + |V_{ti}|^2 &= 1,\ i = d, s, b \\
|V_{id}|^2 + |V_{is}|^2 + |V_{ib}|^2 &= 1,\ i = u, c, t \\
V_{ud}V_{us}^* + V_{cd}V_{cs}^* + V_{td}V_{ts}^* &= 0 \\
V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* &= 0 \\
V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* &= 0 \\
V_{ud}V_{cd}^* + V_{us}V_{cs}^* + V_{ub}V_{cb}^* &= 0 \\
V_{ud}V_{td}^* + V_{us}V_{ts}^* + V_{ub}V_{tb}^* &= 0 \\
V_{cd}V_{td}^* + V_{cs}V_{ts}^* + V_{cb}V_{tb}^* &= 0
\end{aligned} \tag{2.13}$$

The last six relations in particular allow the construction of the **unitarity triangles**, which are diagrams formed by representing each of the $V_{ij}V_{kl}^*$ pieces as a vector on the complex plane; the fact that each trio adds up to zero immediately leads to a triangular shape. The angles and side lengths are physical observables related directly to the moduli and phases of the CKM matrix elements. The six triangles have different shapes, but it can be shown [15] that they all have the same area.

A number of parameterizations of the CKM matrix are in use, the principal two being the PDG-approved version [107] and the Wolfenstein approximation [14]. The latter approximates the matrix by expanding each element as a power series in $\lambda \equiv |V_{us}| \approx 0.22$:

$$V_{CKM} = V_{CKM}^{(3)} + O(\lambda^4) \tag{2.14}$$

where

$$V_{CKM}^{(3)} = \begin{pmatrix} 1 - \lambda^2/2 & \lambda & A\lambda^3\,(\rho - i\eta) \\ -\lambda & 1 - \lambda^2/2 & A\lambda^2 \\ A\lambda^3\,(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} \tag{2.15}$$

2.3 The physics of neutral mesons

This work is principally concerned with the seeking of new physics, through CP-violation, in the neutral B-meson sector. A brief overview of the quantum mechanics of neutral mesons, and the means by which CP-violation exhibits itself in these systems, is therefore presented here. The formalism used here follows that of [9].

The phenomenon of particle *oscillation* occurs when the particles and their anti-particle counterparts are distinguished by an internal quantum identifier which is not conserved by the weak interaction. Following [9] we denote the particle as P^0 and the anti-particle as $\bar{P}^0$. The internal quantum number violated by the weak interaction is F, such that $\Delta F = 0$ for the electromagnetic and strong interaction Hamiltonians H_{EM} and H_S, whilst for the weak interaction H_W, $\Delta F \neq 0$. In the case of the neutral B-mesons, $F = S$, the strangeness quantum number [1].

A particle which at time $t = 0$ occupied the state $|P^0\rangle$ or $|\bar{P}^0\rangle$ will evolve into a mixed state with increasing t. We write this as a projection onto a space spanned by both $|P^0\rangle$ and $|\bar{P}^0\rangle$:

$$|\Psi(t)\rangle = a(t)|P^0\rangle + b(t)|\bar{P}^0\rangle \tag{2.16}$$

The time evolution operators a and b are governed by a coupled Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}\Psi = \mathscr{H}\Psi \tag{2.17}$$

where $\mathscr{H} = \mathscr{H}_{EM} + \mathscr{H}_W + \mathscr{H}_S$, and

$$\Psi(t) = \begin{pmatrix} a(t) \\ b(t) \end{pmatrix} \tag{2.18}$$

The Hamiltonian is a matrix of the form

$$\mathscr{H} = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \tag{2.19}$$

Following [13] we now re-write this matrix in the following way:

$$\mathscr{H} = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} = \begin{pmatrix} M_{11} - \frac{i}{2}\Gamma_{11} & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{21} - \frac{i}{2}\Gamma_{21} & M_{22} - \frac{i}{2}\Gamma_{22} \end{pmatrix} \tag{2.20}$$

This matrix can be simplified, as $\mathscr{CPT}$ invariance [9] dictates that

$$\begin{aligned} M_{11} &= M_{22} = M \\ \Gamma_{11} &= \Gamma_{22} = \Gamma \\ M_{21} &= M_{12} \\ \Gamma_{21} &= \Gamma_{12} \end{aligned} \tag{2.21}$$

[1] It should be noted that there is nothing special about the quantum number $F = S$ with respect to the weak interaction - these reactions clearly do not conserve beauty either. Many older texts make statements such as “the charged weak current does not conserve strangeness” - in fact, it does not conserve any flavour quantum numbers. There is no unique phenomenology associated with strangeness in this regard.

which leaves us with

$$\mathscr{H} = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} = \begin{pmatrix} M - \frac{i}{2}\Gamma & M_{12} - \frac{i}{2}\Gamma_{12} \\ M_{12} - \frac{i}{2}\Gamma_{12} & M - \frac{i}{2}\Gamma \end{pmatrix} \tag{2.22}$$

The task in hand is the solution of the Schrödinger equation. This is best achieved though the diagonalization of the matrix $\mathscr{H}$; the equations decouple, and the two mass eigenstates of the system are obtained. Following [9], we obtain the following for the mass eigenstates:

$$\begin{aligned} |P_1\rangle &= p|P^0\rangle + q|\bar{P}^0\rangle \\ |P_2\rangle &= p|P^0\rangle - q|\bar{P}^0\rangle \end{aligned} \tag{2.23}$$

with these eigenvalues:

$$\begin{aligned} M_1 - \frac{i}{2}\Gamma_1 &= M - \frac{i}{2}\Gamma + \frac{q}{p}\left(M_{12} - \frac{i}{2}\Gamma_{12}\right) \\ M_2 - \frac{i}{2}\Gamma_2 &= M - \frac{i}{2}\Gamma - \frac{q}{p}\left(M_{12} - \frac{i}{2}\Gamma_{12}\right) \end{aligned} \tag{2.24}$$

where

$$\frac{q}{p} = \pm\sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}} = \sqrt{\frac{H_{21}}{H_{12}}} \tag{2.25}$$

The sign ambiguity indicates the arbitrariness of the 1,2 labels on the states - selecting the opposite sign is exactly equivalent to interchanging the labels.

2.3.1 Time evolution

Solution of the Schrödinger equation 2.17 yields the following for the time evolution of the mass eigenstates:

$$|P^0_\pm(t)\rangle = |P^0_\pm\rangle e^{-i(M_\pm - i\Gamma_\pm/2)t} \tag{2.26}$$

$\Gamma_\pm$ and $M_\pm$ are the widths and masses of the mass eigenstates. We adopt the convention $M_- - M_+ = M_2 - M_1$ and $\Gamma_+ - \Gamma_- = \Gamma_1 - \Gamma_2$ [9]. It is evident that the observed particles are in fact mixed states - the entities "seen" in a detector are linear combinations of the flavour eigenstates $|P^0\rangle$ and $|\bar{P}^0\rangle$. The difference between the widths and masses of the two mass eigenstates

$$\Delta M = M_- - M_+ \tag{2.27}$$

$$\Delta\Gamma = \Gamma_+ - \Gamma_- \tag{2.28}$$

vary depending on the system in question. Through the subtraction of equations 2.24 and the taking of real and imaginary parts, we can write

$$\begin{aligned} \Delta M = M_- - M_+ = M_2 - M_1 &= -2\Re M_{12} \\ \Delta\Gamma = \Gamma_+ - \Gamma_- = \Gamma_1 - \Gamma_2 &= 2\Re\Gamma_{12}\cos\zeta \end{aligned} \tag{2.29}$$

where $\zeta = \arg\frac{\Gamma_{12}}{M_{12}}$.

The time evolution of the flavour eigenstates $|P^0(t)\rangle$ and $|\bar{P}^0(t)\rangle$ is given by

$$|P^0(t)\rangle = g_+(t)|P^0\rangle + \frac{q}{p}g_-(t)|\bar{P}^0\rangle \tag{2.30}$$

$$|\bar{P}^0(t)\rangle = g_+(t)|\bar{P}^0\rangle + \frac{p}{q}g_-(t)|P^0\rangle \tag{2.31}$$

where

$$g_\pm = \frac{1}{2}\left(e^{-i(M_+ - i\Gamma_+/2)t} \pm e^{-i(M_- - i\Gamma_-/2)t}\right) \tag{2.32}$$

From these expressions we can immediately write down mixing rates for the flavour eigenstates. Let us suppose that we require the probability (χ) of a meson in the state P^0 at $t = 0$ decaying into the state $\bar{P}^0$. This is given by:

$$\begin{aligned} \chi &= \frac{\text{Total probability of } P^0 \text{ state leaking into } \bar{P}^0 \text{ state}}{\text{Total probability of } P^0 \text{ state leaking into all possible states}} \\ &= \frac{\int_0^\infty \left|\langle\bar{P}^0|P^0(t)\rangle\right|^2 dt}{\int_0^\infty \left|\langle\bar{P}^0|P^0(t)\rangle\right|^2 dt + \int_0^\infty \left|\langle P^0|P^0(t)\rangle\right|^2 dt} \end{aligned} \tag{2.33}$$

In order to evaluate this we need the projections of P^0 and $\bar{P}^0$ onto the evolving states; these are easily found using equations 2.23, 2.30, 2.31 and 2.32.

$$\begin{aligned} \langle P^0|P^0(t)\rangle &= \frac{1}{2p}\langle P^0\Big(|P_2(t)\rangle + |P_1(t)\rangle\Big) \\ &= \frac{1}{2p}\langle P^0|P_2(t)\rangle + \frac{1}{2p}\langle P^0|P_1(t)\rangle \\ &= \frac{1}{2p}\langle P^0|e^{-(\Gamma_+/2+iM_-)t}|P_2\rangle + \frac{1}{2p}\langle P^0|e^{-(\Gamma_-/2+iM_+)t}|P_1\rangle \\ &= \frac{1}{2p}e^{-(\Gamma_+/2+iM_-)t}\Big(p\langle P^0|P^0\rangle + q\langle P^0|\bar{P}^0\rangle\Big) + \\ &\quad + \frac{1}{2p}e^{-(\Gamma_+/2+iM_-)t}\Big(p\langle P^0|P^0\rangle + q\langle P^0|\bar{P}^0\rangle\Big) \\ &= \frac{1}{2}\Big(e^{-(\Gamma_+/2+iM_-)t} + e^{-(\Gamma_-/2+iM_+)t}\Big) \end{aligned} \tag{2.34}$$

Similarly,

$$\langle\bar{P}^0|P^0(t)\rangle = \frac{q}{2p}\Big(e^{-(\Gamma_2/2+iM_-)t} - e^{-(\Gamma_1/2+iM_+)t}\Big) \tag{2.35}$$

Using 2.27 and 2.28 and writing $\Gamma = \frac{\Gamma_1+\Gamma_2}{2}$, squaring (2.34) and (2.35) and integrating over all time, we obtain:

$$\chi = \frac{\left(\frac{\Delta M}{\Gamma}\right)^2 + \left(\frac{\Delta\Gamma}{\Gamma}\right)^2}{\left(\frac{\Delta M}{\Gamma}\right)^2 + \left(\frac{\Delta\Gamma}{\Gamma}\right)^2 + \left|\frac{p}{q}\right|^2 (2 + \left(\frac{\Delta M}{\Gamma}\right)^2 - \left(\frac{\Delta\Gamma}{\Gamma}\right)^2)} \tag{2.36}$$

or, writing $x = \frac{\Delta M}{\Gamma}$ and $y = \frac{\Delta\Gamma}{\Gamma}$:

$$\frac{x^2+y^2}{x^2+y^2+\left|\frac{p}{q}\right|^2 (2+x^2-y^2)} \tag{2.37}$$

The quantity

$$x = \frac{\Delta M}{\Gamma} \tag{2.38}$$

is the *oscillation parameter*. It can be seen that the probability of state leakage is dependent on the difference in the masses of the mass eigenstates. In general, oscillations can only be observed if the period $2\pi/\Delta M$ is of the same order as the mean particle lifetime $1/\Gamma$; in other words, $x \geq 1$. The equivalent expression for the probability of a $|\bar{P}^0\rangle$ decaying as a $|P^0\rangle$ ($\bar{\chi}$) is:

$$\bar{\chi} = \frac{x^2+y^2}{x^2+y^2+\left|\frac{q}{p}\right|^2 (2+x^2-y^2)} \tag{2.39}$$

Clearly if $|\frac{p}{q}| = 1$ then $\chi = \bar{\chi}$ and an equilibrium is established.

2.3.2 CP violation in mixing

Following [9], let us assume that the P^0 and the $\bar{P}^0$ decay to some final state f through the amplitudes A_f and $\bar{A}_f$, where

$$A_f = \langle f|H_{\Delta f=1}|P^0\rangle \tag{2.40}$$

$$\bar{A}_f = \langle f|H_{\Delta f=1}|\bar{P}^0\rangle \tag{2.41}$$

$$\rho_f = \frac{A_f}{\bar{A}_f} = \frac{1}{\bar{\rho}_f} \tag{2.42}$$

Then, taking equations 2.30 - 2.32, substituting into A_f and $\bar{A}_f$ and taking the square of the modulus, we obtain:

$$\Gamma\left(P^0(t) \to f\right) \propto e^{-\Gamma_1 t} |A_f|^2 \left[K_+(t) + K_-(t)\left|\frac{q}{p}\right|^2 |\bar{\rho}_f|^2 + 2\Re\left[L(t)\left(\frac{q}{p}\right)\bar{\rho}(f)\right]\right] \tag{2.43}$$

$$\Gamma\left(\bar{P}^0(t) \to f\right) \propto e^{-\Gamma_1 t}\left|\bar{A}_f\right|^2\left[K_+(t) + K_-(t)\left|\frac{p}{q}\right|^2|\rho_f|^2 + 2\Re\left[L(t)\left(\frac{p}{q}\right)\rho(f)\right]\right] \tag{2.44}$$

where

$$\begin{aligned} |g_\pm(t)|^2 &= \frac{1}{4}e^{-\Gamma_1 t}K_\pm(t) \\ g_-(t)g_+^*(t) &= \frac{1}{4}e^{-\Gamma_1 t}L(t) \\ K_\pm &= 1 + e^{\Delta\Gamma t} \pm 2e^{\frac{1}{2}\Delta\Gamma t}\cos\Delta M t \\ L(t) &= 1 - e^{\Delta\Gamma t} + 2ie^{\frac{1}{2}\Delta\Gamma t}\sin\Delta M t \end{aligned} \tag{2.45}$$

Any difference in the decay rates 2.43 and 2.44 indicates CP-violation. There are three means by which CP-violation can appear in these mixing processes.

Direct CP violation

Let us assume there are no differences in mass or width between the eigenstates, such that $x = 0$. Then 2.43 and 2.44 become

$$\Gamma\left(P^0(t) \to f\right) \propto 4e^{-\Gamma_1 t}|A_f|^2 \tag{2.46}$$

$$\Gamma\left(\bar{P}^0(t) \to f\right) \propto 4e^{-\Gamma_1 t}\left|\bar{A}_f\right|^2 \tag{2.47}$$

It is evident that the only means for a CP-asymmetry to occur is if

$$|A_f| \neq \left|\bar{A}_f\right| \tag{2.48}$$

as shown in figure 2.1.

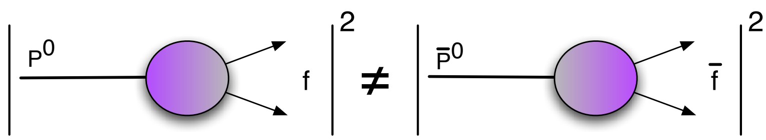


Figure 2.1: Conditions for direct CP violation

Mixing-induced CP violation

If we assume that the decay amplitudes are equal then the rate difference between 2.43 and 2.44 is given by:

$$\Gamma\left(P^0(t) \to f\right) - \Gamma\left(\bar{P}^0(t) \to f\right) \propto K_-(t)\left(\left|\frac{q}{p}\right|^2 - \left|\frac{p}{q}\right|^2\right) + 2\Re L(t)\left[\left(\frac{q}{p}\right) - \left(\frac{p}{q}\right)\right] \tag{2.49}$$

It is evident that even if the direct decay amplitudes are equal, a CP asymmetry is possible if the modulus of the ratio of the the quantities q and p is not equal to one.

$$\left|\frac{q}{p}\right| \neq 1 \Rightarrow \text{CP violation} \tag{2.50}$$

These effects are illustrated in figure 2.2.

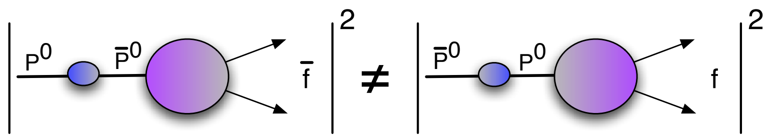


Figure 2.2: Conditions for mixing-induced CP violation

CP violation induced by the interference of mixing and decay amplitudes

An interesting case arises when both the state and the anti-state can decay into the same final state f.

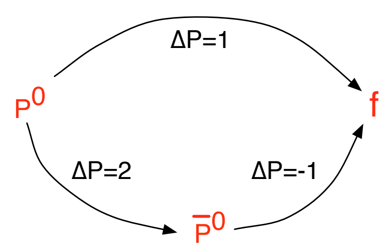


Figure 2.3: Decay of both the P and the $\bar{P}$ to the same final state

In this case it is possible that, when the direct decay amplitudes interfere with the mixing amplitudes, CP-violation arises. In this case the requirement for CP-violation is less strict than in the other two cases:

$$\frac{q}{p} \cdot \frac{\bar{A}_f}{A_f} \neq 1 \Rightarrow \text{CP violation} \tag{2.51}$$

This situation is illustrated in figure 2.4

We introduce a quantity ξ_{CP} which captures all the information on the CP-asymmetry of the system.

$$\xi_{CP} = \frac{q}{p} \cdot \frac{\bar{A}_f}{A_f} \tag{2.52}$$

This is referred to as the **CP-asymmetry parameter**.

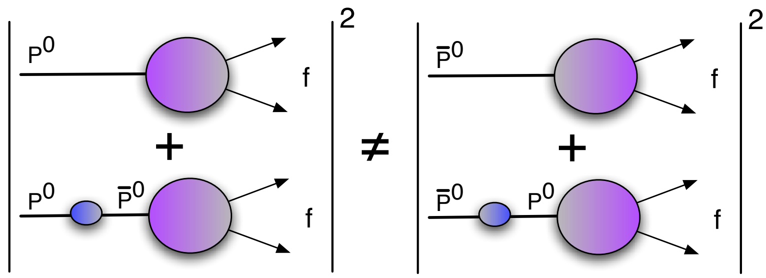


Figure 2.4: Conditions for CP violation induced by mixing and decay amplitudes

2.3.3 Mixing and CP-violation in the $B_s - \bar{B}_s$ system

We now consider a specific system - the neutral B_s mesons. In this case the P^0 mesons above become B^0, and the violated internal quantum identifier $F = S$, strangeness. The mixing interactions are described by the box diagrams in figure 2.5 [115]. The B-physics sector generally, and the B_s meson system in particular, are

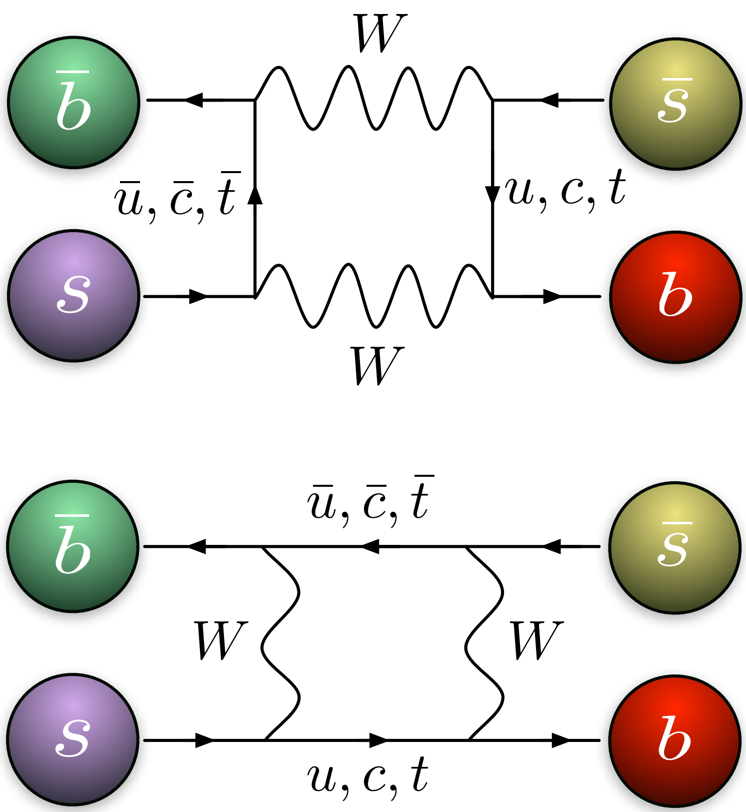


Figure 2.5: Diagrams contributing to $B_s - \bar{B}_s$ mixing

highly profitable domains in which to perform experimental tests on a wide range of phenomena. B-hadrons tend to be relatively *long lived* particles - long enough for their proper decay times to be measured by a modern high-granularity tracking detector. The high mass of the b-quark ensures a diverse range of decay channels, giving a wide scope for the study of potential new physics effects.

The two neutral B-mesons of interest here are $B_d^0 = \bar{b}d$ and $B_s^0 = \bar{b}s$. Table 2.1 summarizes the current state of knowledge for these systems, giving mean masses, lifetimes, the mass difference between the mass eigenstates and the oscillation parameter [107]. Note that the values given for the B_s^0 system are based on Tevatron (CDF) measurements [108], and whilst provided by the Particle Data Group, are not yet officially accepted. They are adopted for this work, however.

	$\bar{M}$, MeV	$c\bar{\tau}$, μm	ΔM_q, MeV	x_q
B_d^0	5279.4 ± 0.5	458.7	$(3.337 \pm 0.033) \times 10^{-10}$	0.776 ± 0.008
B_s^0	5367.5 ± 1.8	439	$(114.07^{+2.76}_{-1.38} \pm 0.46) \times 10^{-10}$	> 19.9

Table 2.1: Current state of knowledge of the neutral B-mesons

If both the B^0 and the $\bar{B}_0$ can decay to the same final state f, the direct decay amplitudes can interfere with the mixing amplitudes. This may admit CP-violation, as indicated in equation 2.51. Following [9], under the assumptions that $\Delta\Gamma \ll \Delta M$ and $\Delta\Gamma \ll \Gamma$ (which hold for both systems), we can write

$$\frac{q}{p} \propto \sqrt{\frac{M_{12}^*}{M_{12}}} \tag{2.53}$$

Now [13] assumes that the mixing process is overwhelmingly dominated by diagrams involving the top quark as the intermediary (see figure 2.5), and so we can write

$$M_{12}^q \propto \left(V_{tq}^* V_{tb}^*\right)^2 e^{i(\pi - \phi_{CP})} \tag{2.54}$$

where the V are elements of the CKM matrix and ϕ_{CP} is defined as

$$\mathscr{C}\mathscr{P}|B_q^0\rangle = e^{i\phi_{CP}}|\bar{B}_q^0\rangle \tag{2.55}$$

Re-writing this we obtain

$$M_{12}^q \propto \left|V_{tq}^* V_{tb}^*\right|^2 e^{-i\phi_q} e^{i(\pi - \phi_{CP})} \tag{2.56}$$

where

$$\phi_q = 2 \arg V_{tq}^* V_{tb}^* \tag{2.57}$$

Consequently we have

$$\frac{q}{p} \propto e^{-i\theta_{M12}^q} \tag{2.58}$$

where

$$\theta^q_{M12} = \pi - \phi_q - \phi_{CP} \tag{2.59}$$

θ^q_{M12} is referred to as the **weak mixing phase**. We now consider the special case where the final state into which the B-mesons decay is a CP-eigenstate.

$$\mathscr{C}\mathscr{P}|f_{CP}\rangle = \pm|f_{CP}\rangle \tag{2.60}$$

The CP-asymmetry parameter given by equation 2.52 becomes

$$\xi^q_{f_{CP}} = e^{-i\theta^q_{M12}} \frac{A\left(\bar{B}^0_q \to f_{CP}\right)}{A\left(B^0_q \to f_{CP}\right)} \tag{2.61}$$

The weak mixing phase θ^q_{M12} contains only CKM matrix elements and is hence "theoretically clean", whereas the ratio of the amplitudes suffers from hadronic uncertainties. The method of **low energy effective Hamiltonians** [13], is a means of approximately calculating the hadronic matrix elements, by separating out the "long" and "short" distance contributions to the decay amplitude. The ratio of amplitudes can be written as [13]

$$\frac{A\left(\bar{B}^0_q \to f_{CP}\right)}{A\left(B^0_q \to f_{CP}\right)} \propto \pm e^{-i\phi_{CP}} \frac{\sum_{j=u,c} V^*_{jq} V_{jb} \langle f_{CP}|\hat{O}|\bar{B}^0_q\rangle}{\sum_{j=u,c} V^*_{jq} V_{jb} \langle f_{CP}|\hat{O}|B^0_q\rangle} \tag{2.62}$$

where the $\hat{O}$ is the local four-quark operator (the "short-distance" contribution). It can be seen immediately that upon writing the CP-asymmetry parameter, the phase ϕ_{CP} disappears, leaving

$$\xi^q_{f_{CP}} = \mp e^{-i\phi_q} \frac{\sum_{j=u,c} V^*_{jq} V_{jb} \langle f_{CP}|\hat{O}|\bar{B}^0_q\rangle}{\sum_{j=u,c} V^*_{jq} V_{jb} \langle f_{CP}|\hat{O}|B^0_q\rangle} \tag{2.63}$$

In general this is poorly known due to the hadronic uncertainties in the amplitude ratio. However, in the special case where the decay is dominated by a single amplitude the matrix elements cancel, leading to a radically simplified expression:

$$\xi^q_{f_{CP}} = \eta_{f_{CP}} e^{-i\left(\phi_q - \phi^D_{f_{CP}}\right)} \tag{2.64}$$

$\eta_{f_{CP}} = \pm 1$ is the CP eigenvalue of the final state f_{CP}. The two phases ϕ_q and $\phi^D_{f_{CP}}$ are responsible for CP-violation in mixing and direct decay of the B-meson to its final state respectively. These decays, with one dominant amplitude which permit the cancellation of hadronic uncertainties, are of great experimental interest and are consequently referred to as "Golden Modes". One such mode, the decay of B^0_s to the final state $J/\psi\phi$, is studied in detail in this work.

2.3.4 Potential effects of New Physics

Extensions of the Standard Model introduce new processes which can increase significantly the amount of observed CP-violation. Indeed, CP-violation is often of more interest as an indicator of new physics than a phenomenon in its own right. An illustration (based on [87]) of the potential effects of new physics on mixing and CP-violation in the neutral B-meson sector is given here. No assumptions are made as to the type of process leading to the effects - the treatment is model-independent.

We assume that some new physics processes lead to additional contributions to the $B^0 - \bar{B}^0$ transition matrix element M_{12}, such that

$$M_{12} \neq M_{12}^{SM} \tag{2.65}$$

We write an expression for the ratio of the total contributions to those provided by the Standard Model:

$$\sqrt{\frac{M_{12}}{M_{12}^{SM}}} = re^{i\beta} \tag{2.66}$$

which gives (using equation 2.29)

$$\Delta M = r^2 \Delta M^{SM} \tag{2.67}$$

and

$$\Delta\Gamma = \Delta\Gamma^{SM} \cos 2\beta \tag{2.68}$$

The immediate implication of the above is that new physics contributions will lead to a *decrease* in the width difference $\Delta\Gamma$. Note that this β is *not* the same as the angle used in the unitarity triangles - it is the weak phase arising from the new physics contributions.

We introduce a quantity R, the ratio of the new physics contributions to those of the Standard Model:

$$R = \frac{M_{12}^{NP}}{M_{12}^{SM}} \tag{2.69}$$

Given that $M_{12} = M_{12}^{SM} + M_{12}^{NP}$ it follows that

$$2\beta = \arg(1 + R) \tag{2.70}$$

and the oscillation parameter x is

$$x = \frac{\Delta M}{\Gamma} = \frac{\Delta M^{SM}}{\Gamma} |1 + R| \tag{2.71}$$

From these simple equations it is possible to plot contours of constant R and $\arg R$, showing how the observable quantities can be expected to vary with differing contributions from new physics processes. Figures 2.6 and 2.7 show these plots for the B_s system. These plots will be used to devise new physics scenarios in the $B_s \to J/\psi\phi$ feasibility studies which follow in later chapters.

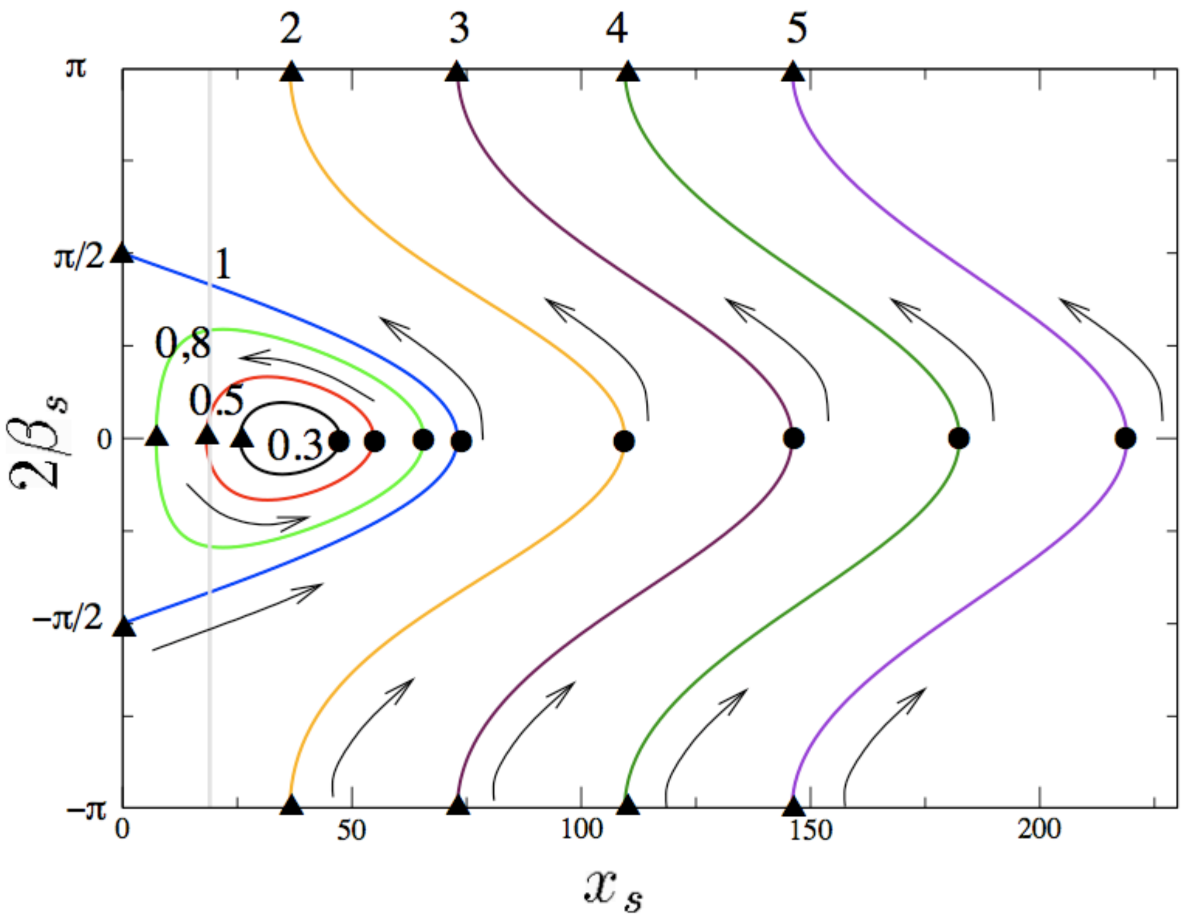


Figure 2.6: Contours of constant $|R|$ in the $x_s - \beta_s$ plane, for $|R| \in \{0.3, 0.5, 0.8, 1.0, 3.0, 5.0\}$. $\arg R$ varies from 0 to 2π along the contour lines, increasing in the direction of the arrows. At the triangles, $\arg R = 0$ and at the circles, $\arg R = \pi$. Plot taken from [87]

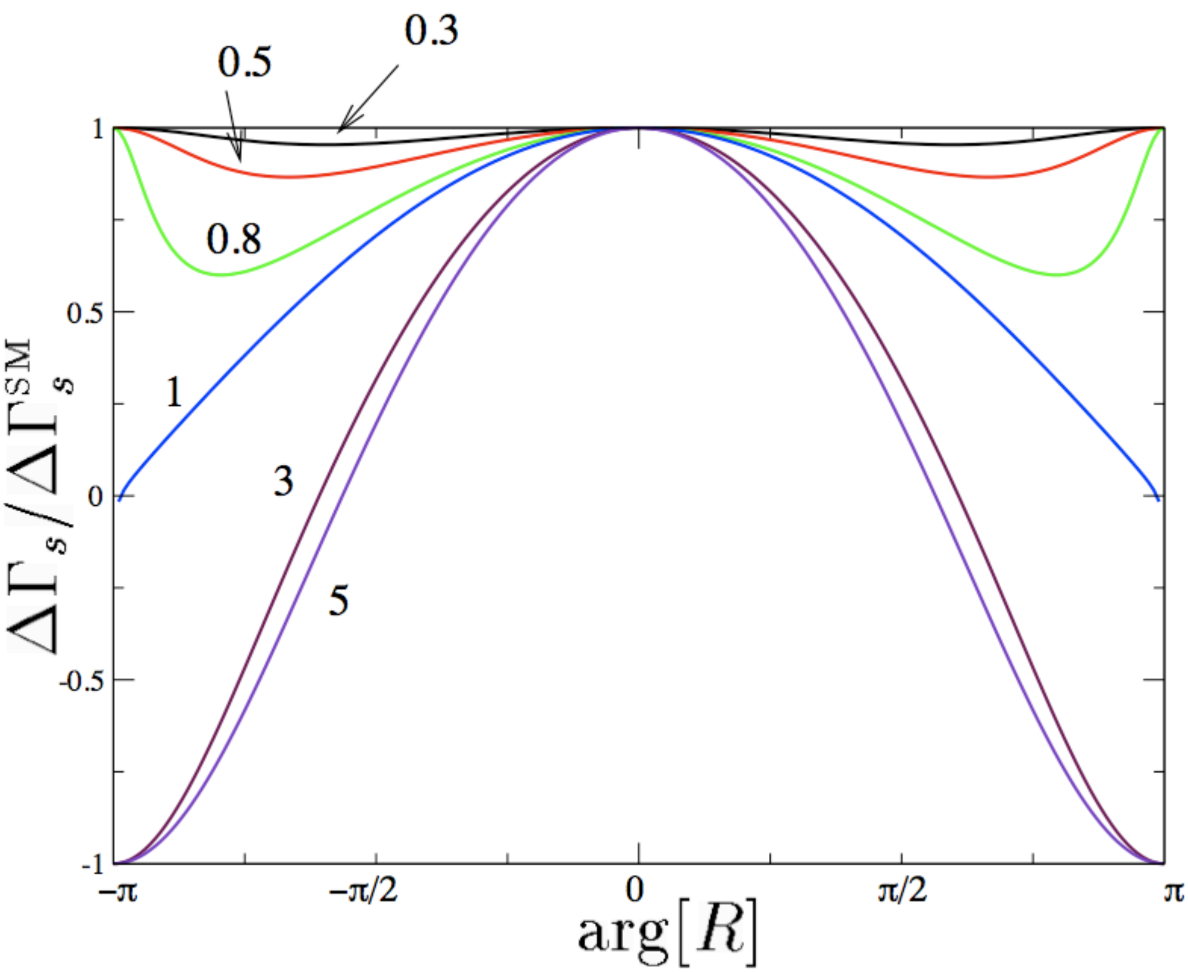


Figure 2.7: Contours of constant $|R|$ in the $\Delta\Gamma_s - \arg R$ plane, for $|R| \in \{0.3, 0.5, 0.8, 1.0, 3.0, 5.0\}$. Plot taken from [87]

Chapter 3

The Helicity Formalism

The phenomenology of decays with non-trivial spin configurations will play a crucial role in the analyses presented in this work; indeed, the central expression governing the behaviour of the signal and exclusive background channels is derived through spin formalisms. Consequently, this chapter sets out the important concepts and presents the mathematical machinery needed to perform these calculations. The material shown here is a distillation of three treatments on the topic - Richman [16], Chung [17] and Jacob and Wick [18].

3.1 Introductory concepts

The *helicity* operator is defined as

$$\hat{\lambda} = \hat{\mathbf{S}} \cdot \hat{\mathbf{p}} \tag{3.1}$$

where $\hat{\mathbf{S}}$ and $\hat{\mathbf{p}}$ are the spin and momentum operators respectively; it is the projection of the spin of a particle onto its direction of motion. The expectation value is in some sense a measure of the degree of "alignment" of the spin with the momentum vector. This is a particularly attractive quantity on which to base analyses of decays involving non-trivial spin states, as the helicity of a state is invariant *under both rotations and boosts along* $\mathbf{p}$. Of course, the normal spin-orbit formalism that is generally deployed in non-relativistic calculations, e.g. [19], can also be used; the total angular momentum of a system is conserved, so in principle the spins and the orbital angular momenta of all the particles can be summed over. In practice, however, this is a complex task - the spin operator is defined in the rest frame of the particle in question, whereas the orbital angular momentum operator is defined in the centre of mass frame for the entire system. The two are clearly not at rest relative to one another, so it is not difficult to imagine the algebraic complexity of such an analysis. The helicity formalism provides exactly the same final answer, but with far less confusion, so it is generally the adopted technique.

The route taken in this chapter is as follows: after setting up the notation, rotation operators are defined and some important properties derived. These operators are next used to construct helicity states. This equipment is then used to derive angular distributions for the simplest case of a particle decaying to two daughters, and finally, sequential decays of the type that will be considered later on in this work will be tackled.

3.2 Finite rotations and the quantum mechanics of angular momentum

The quantum mechanics of angular momentum are intimately related to rotation. The act of imposing commutator relations on the classical angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ gives rise to the operators $\hat{J}_x$, $\hat{J}_y$, $\hat{J}_y$. These define a *Lie Algebra* [23], which is a vector space using commutation as the associative operator. Every Lie Algebra has a corresponding Lie Group; in this case the algebra is $\mathcal{L} = so(3)$, and the group $\mathscr{G} = SO(3)$ [20]. $SO(3)$ is *isomorphic* to the set of proper (length preserving) rotation matrices. In short, the spin of a particle defines how the rest states of that particle transform under proper rotations, and conversely, the application of rotation operators to particle states yields information on spin-related behaviour.

3.2.1 Describing rotations

Any rotation in three dimensional space can be described by three successive rotations through the *Euler Angles* α, β and γ. The rotation is specified by a new set of axes XYZ which is attached to the rotated physical system; the angles are defined relative to the fixed original axes xyz. As can be seen from figure 3.1, the rotations are as follows:

1. A rotation through angle α about the z axis. The positions of the X and Y axes rotate away from x and y.

2. A rotation through angle β about the Y axis. The Z axis now rotates out of the z position.

3. Finally, a rotation through angle γ about the new Z-axis, taking the XYZ coordinates to their final position.

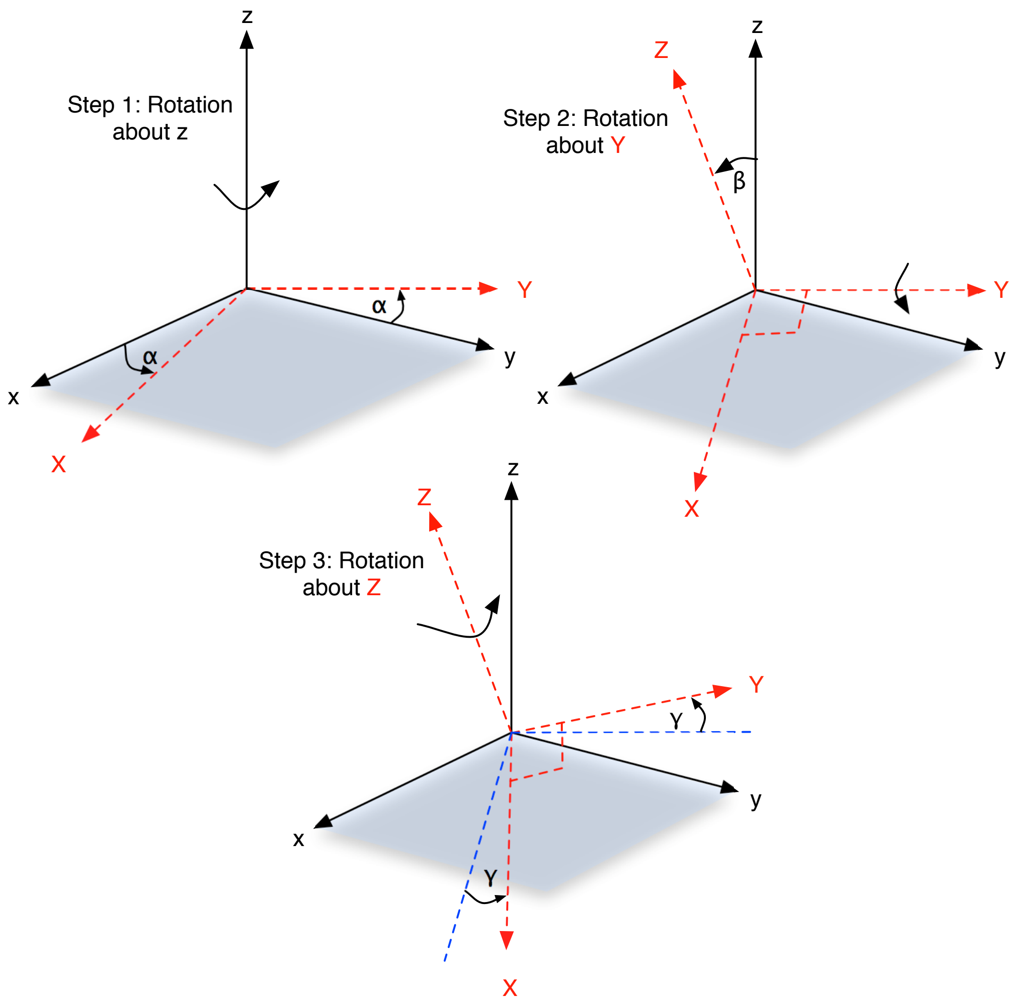


Figure 3.1: Construction of an arbitrary rotation through three successive rotations

Note that the polar coordinates φ and θ of the Z-axis in its final position with respect to the fixed xyz system, are identical to α and β (see figure 3.2). Before we can make further progress we must express all of these rotations relative to the fixed xyz system. Following [21] we denote a rotation of angle θ about an axis ξ by the operator $\hat{D}_\xi(\theta)$. We can write the three rotations

1. $\hat{D}_z(\alpha)$
2. $\hat{D}_Y(\beta)$
3. $\hat{D}_Z(\gamma)$

Now let us express rotations 2 and 3, which are currently pivoted on shifting axes, in terms of multiple rotations about the fixed xyz axis.

$$\hat{D}_Y(\beta) \equiv \hat{D}_z(\alpha)\,\hat{D}_y(\beta)\,\hat{D}_z(-\alpha)$$

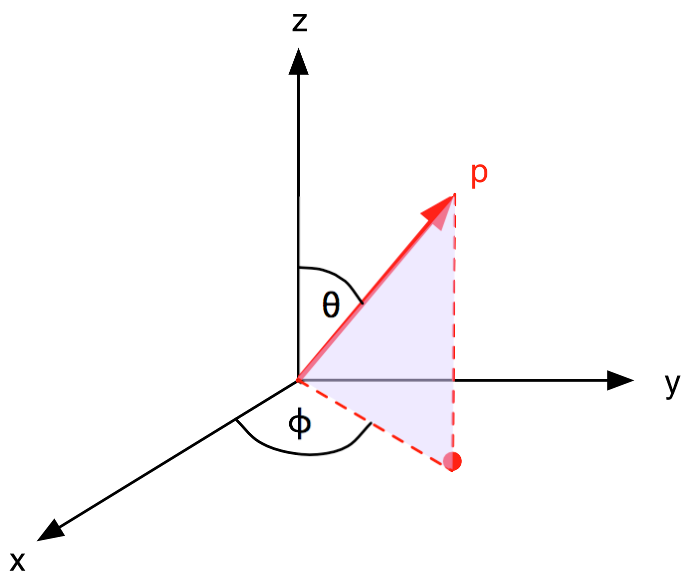


Figure 3.2: Definition of polar coordinates

$$\begin{aligned}\hat{D}_Z(\gamma) &\equiv \hat{D}_Y(\beta)\,\hat{D}_Z(\gamma)\,\hat{D}_Y(-\beta) \\ &\equiv \hat{D}_z(\alpha)\,\hat{D}_y(\beta)\,\hat{D}_z(\gamma)\,\hat{D}_y(-\beta)\,\hat{D}_z(-\beta)\end{aligned}$$

Putting all this together, we get the final result

$$\hat{D}_Z(\gamma)\,\hat{D}_Y(\beta)\,\hat{D}_z(\alpha) \equiv \hat{D}_z(\alpha)\,\hat{D}_y(\beta)\,\hat{D}_z(\gamma) \tag{3.2}$$

In other words, a rotation as described above is identical to

1. A rotation γ about the fixed z axis
2. A rotation β about the fixed y axis
3. A rotation α about the fixed z axis

which is perhaps not an immediately obvious result.

3.2.2 Angular momentum operators and single particle states

The states of a particle at rest with mass $w > 0$ may be written as $|jm\rangle$, where j is the total spin and m is its z-component, $(m = -j, -j+1, ..., j-1, j)$. The three components of the angular momentum operator $\mathbf{J}$ (denoted $\hat{J}_x$, $\hat{J}_y$ and $\hat{J}_z$), obey the commutation relations:

$$[\hat{J}_i, \hat{J}_j] = i\epsilon_{ijk}\hat{J}_k \tag{3.3}$$

where ϵ_{ijk} is the Levi-Civita tensor. The operators act on the single particle states as follows:

$$\hat{\mathbf{J}}^{\mathbf{2}}|jm\rangle = j(j+1)|jm\rangle \tag{3.4}$$

$$\begin{aligned}\hat{J}_z|jm\rangle &= m|jm\rangle \\ \hat{J}_\pm|jm\rangle &= \sqrt{(j \mp m)(j \pm m + 1)}|jm \pm 1\rangle\end{aligned} \tag{3.5}$$

where $\hat{\mathbf{J}}^{\mathbf{2}} = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2$ and $\hat{J}_\pm = \hat{J}_x \pm i\hat{J}_y$. The states are normalised thus:

$$\langle j'm'|jm\rangle = \delta_{j'j}\delta_{m'm} \sum_{jm} |jm\rangle\langle jm| = \hat{\mathbf{I}} \tag{3.6}$$

where $\hat{\mathbf{I}}$ is the identity operator.

3.2.3 Transformation of the angular momentum operators under finite rotations

Let us consider the effects of a rotation about the z axis through an angle γ, given by $\hat{D}\,(00\gamma)$, on some physical system (formally a field function) $f\,(\theta, \varphi)$.

$$\hat{D}\,(00\gamma)\,f\,(\theta,\varphi) = f'\,(\theta,\varphi) = f\,(\theta,\varphi+\gamma) \tag{3.7}$$

It follows that

$$\frac{\partial}{\partial\gamma}\hat{D}\,(\alpha\beta\gamma)\,f\,(\theta,\varphi) = \hat{D}\,(\alpha\beta\gamma)\,\frac{\partial}{\partial\varphi}f\,(\theta,\varphi) \tag{3.8}$$

Noting that, written in differential form, the operator $\hat{J}_z = -i\,(\partial/\partial\varphi)$, we can write

$$\frac{\partial}{\partial\gamma}\hat{D}\,(\alpha\beta\gamma)\,f\,(\theta,\varphi) = \hat{D}\,(\alpha\beta\gamma)\,i\hat{J}_z f\,(\theta,\varphi) \tag{3.9}$$

The solution to this differential equation is

$$\hat{D}\,(\alpha\beta\gamma) = C\,(\alpha\beta)\,e^{i\gamma\hat{J}_z} \tag{3.10}$$

Other rotations can be expressed similarly, and we arrive at

$$\hat{D}\,(\alpha\beta\gamma) = e^{i\alpha\hat{J}_z}e^{i\beta\hat{J}_y}e^{i\gamma\hat{J}_z} \tag{3.11}$$

using the results obtained in the previous sections.

3.2.4 Matrix elements of finite rotations

The matrix elements are written as follows:

$$\langle jm'|\hat{D}\,(\alpha\beta\gamma)\,|jm\rangle = \mathscr{D}^j_{m'm}\,(\alpha\beta\gamma) \tag{3.12}$$

We make the following simplification in the notation:

$$\mathscr{D}^j_{m'm}\,(0\beta0) = d^j_{m'm}(\beta) \tag{3.13}$$

such that

$$\mathscr{D}^{j}_{m'm}(\alpha\beta\gamma) = e^{im'\gamma} d^{j}_{m'm}(\beta) e^{im'\alpha} \tag{3.14}$$

where the $\hat{J}_{x,z}$ operators have acted on the states. Evaluation of the rotation matrix elements therefore reduces to the problem of calculating

$$d^{j}_{m'm}(\beta) = \langle jm'| e^{i\beta \hat{J}_y} |jm\rangle \tag{3.15}$$

Methods for evaluating these $\tilde{d}$ matrices are given in [21] and [22], and sample formulae are given in [18]. [21] in particular demonstrates that they are described by *Jacobi polynomials*; this is beyond the scope of this work, however. The results of the calculations are quoted here without proof, directly from [21].

$$\begin{aligned} d^{j}_{m'm}(\beta) = & \left(\frac{j-m'}{j-m}\right)^{\frac{1}{2}} d^{j-\frac{1}{2}}_{m'+\frac{1}{2}m+\frac{1}{2}}(\beta)\cdot\cos\frac{\beta}{2} \\ & - \left(\frac{j+m'}{j-m}\right)^{\frac{1}{2}} d^{j-\frac{1}{2}}_{m'+\frac{1}{2}m+\frac{1}{2}}(\beta)\cdot\sin\frac{\beta}{2} \end{aligned} \tag{3.16}$$

$$d^{j}_{m'j}(\beta) = (-1)^{j-m}\left[\frac{(2j)!}{(j+m)!(j-m)!}\right]^{\frac{1}{2}} \left(\cos\frac{\beta}{2}\right)^{j+m} \left(\sin\frac{\beta}{2}\right)^{j-m} \tag{3.17}$$

3.3 Helicity States

3.3.1 One particle helicity states

A single particle helicity state is defined by performing a rotation and a boost on a state vector. Rotations have been studied above; a *boost* is defined as an operator $\hat{B}(\mathbf{p})$ which takes a state from rest into momentum $\mathbf{p}$. Following [16] the state $|jm\rangle$ is now written as $|s\lambda\rangle$. The single particle helicity state is

$$\begin{aligned} |\mathbf{p}; s\lambda\rangle &= |\theta, \varphi, p; s\lambda\rangle \\ &= \hat{B}(\mathbf{p})\hat{D}(\varphi, \theta\ -\varphi)|j\lambda\rangle \end{aligned} \tag{3.18}$$

where θ and φ are the polar coordinates of $\mathbf{p}$. Note that switching the order of the boost and rotation would have no effect on the helicity state.

It was stated in the introduction to this chapter that the helicity λ is invariant under rotations and boosts, this being the main strength of the helicity method. This is now shown formally. First, consider the effect of a rotation $\hat{D}$ on the helicity state:

$$\begin{aligned} \hat{D}|\mathbf{p}; s\lambda\rangle &= \hat{D}\hat{B}(\mathbf{p})\hat{D}(\varphi, \theta\ -\varphi)|s\lambda\rangle \\ &= |\hat{D}\mathbf{p}; s\lambda\rangle \end{aligned}$$

and now consider the effect of a boost $\hat{B}'$ which takes the state from momentum $\mathbf{p}$ to $\mathbf{p}'$, $\mathbf{p}$ being parallel to $\mathbf{p}'$:

$$\begin{aligned}\hat{B}'|\mathbf{p};s\lambda\rangle &= \hat{B}'\hat{B}(\mathbf{p})\hat{D}(\varphi,\theta-\varphi)|s\lambda\rangle \\ &= \hat{B}(\mathbf{p}')\hat{D}(\varphi,\theta-\varphi)|s\lambda\rangle \\ &= |\mathbf{p}';s\lambda\rangle\end{aligned}$$

As can be seen, the helicity does not change.

Following [17], we define the normalisation of the helicity states as being

$$\langle\mathbf{p}';s'\lambda'|\mathbf{p};s\lambda\rangle = \tilde{\delta}\left(\mathbf{p}'-\mathbf{p}\right)\delta_{ss'}\delta_{\lambda\lambda'} \tag{3.19}$$

where $\tilde{\delta}\left(\mathbf{p}'-\mathbf{p}\right)$ is the Lorentz invariant delta function

$$\tilde{\delta}\left(\mathbf{p}'-\mathbf{p}\right) = (2\pi)^3\,(2E)\delta^3\left(\mathbf{p}'-\mathbf{p}\right) \tag{3.20}$$

Other properties of signle particle helicity states are given in [17].

3.3.2 Two particle helicity states

To construct states which represent two particles with momenta $\mathbf{p_1}$ and $\mathbf{p_2}$, the direct product of the two single particle states is taken [16]:

$$|\mathbf{p_1}\lambda_1;\mathbf{p_2}\lambda_2\rangle = |\mathbf{p_1};s_1\lambda_1\rangle \otimes |\mathbf{p_2};s_2\lambda_2\rangle \tag{3.21}$$

From equation 3.19 we can immediately write down the normalisation for the two particle state:

$$\begin{aligned}&\langle\mathbf{p_1'}\lambda_1';\mathbf{p_2'}\lambda_2'|\mathbf{p_1}\lambda_1;\mathbf{p_2}\lambda_2\rangle \\ &\quad = (2\pi)^6 4E_1E_2\delta^3\left(\mathbf{p_1'}-\mathbf{p_1}\right)\delta^3\left(\mathbf{p_2'}-\mathbf{p_2}\right)\delta_{\lambda_1'\lambda_1}\delta_{\lambda_2'\lambda_2}\end{aligned} \tag{3.22}$$

If the centre of mass frame is picked, such that $\mathbf{p_1} = -\mathbf{p_2} = \mathbf{p}$, the particles are back-to-back. In this frame it is possible to express the two-particle state in terms of the coordinates p,θ,φ where $p = |\mathbf{p}|$ and θ and φ are the polar angles of either of the momenta ($\mathbf{p_1}$ is chosen by convention). The state is now written $|p\theta\varphi\lambda_1\lambda_2\rangle$. It is shown in the appendix of Richman [16] that the normalization in these coordinates is

$$\begin{aligned}&\langle p\theta'\varphi'\lambda_1'\lambda_2'|p\theta\varphi\lambda_1\lambda_2\rangle \\ &\quad = (2\pi)^6\frac{4\sqrt{s}}{p}\delta^4\left(\mathbf{P}'^{\alpha}-\mathbf{P}^{\alpha}\right)\delta\left(\cos\theta'-\cos\theta\right)\delta\left(\varphi'-\varphi\right)\delta_{\lambda_1'\lambda_1}\delta_{\lambda_2'\lambda_2}\end{aligned} \tag{3.23}$$

where $\mathbf{P}^{\alpha} = \mathbf{P}_1^{\alpha} + \mathbf{P}_2^{\alpha}$ is the total 4-momentum in the centre of mass frame, such that

$$\mathbf{P}^{\alpha} = (E, 0, 0, 0) \tag{3.24}$$
$$E = E_1 + E_2 = \sqrt{p^2 + m_1^2} + \sqrt{p^2 + m_2^2} \tag{3.25}$$
$$\sqrt{s} = E_1 + E_2 \tag{3.26}$$

We can write [16]

$$|p\theta\varphi\lambda_1\lambda_2\rangle = (2\pi)^3 \left[\frac{4\sqrt{s}}{p}\right]^{\frac{1}{2}} |\theta\varphi\lambda_1\lambda_2\rangle|\mathbf{P}^{\alpha}\rangle \tag{3.27}$$

Other properties are described in [16] and [17].

The final task that needs to be performed before we can work out the angular distributions is to express the two-particle states as eigenstates of *definite total angular momentum*. This is because the calculation of the decay angles requires the imposition of angular momentum conservation on the "before" and "after" states; of course this is only meaningful if these states are eigenstates of total angular momentum. We write these new basis states as $|pJM\lambda_1\lambda_2\rangle$ where p is the magnitude of the momentum of either particle (in the rest frame of the decaying particle these will be the same), J and M are the total angular momenta and z-component of the whole system, and $\lambda_{1,2}$ are the helicities of the two particles as before. The transformation between the two bases is written [16]

$$|p\theta\varphi\lambda_1\lambda_2\rangle = \sum_{J,M} c_{JM}\left(p\theta\varphi\lambda_1\lambda_2\right)|pJM\lambda_1\lambda_2\rangle \tag{3.28}$$

It is shown in the appendix of Richman [16] that

$$c_{JM}\left(p\theta\varphi\lambda_1\lambda_2\right) = c_{J\lambda}\left(p, \theta = 0, \varphi = 0, \lambda_1, \lambda_2\right)\mathscr{D}^{J}_{M\lambda}\left(\varphi, \theta, -\varphi\right) \tag{3.29}$$

We now have all of the tools needed to calculate angular distributions.

3.4 Calculating Angular Distributions

We now come to the kernel of this chapter - the calculation of amplitudes which give the angular distribution for a particular decay. Ultimately the problem reduces to the evaluation of the amplitude

$$A = \langle\theta, \varphi, \lambda_1, \lambda_2|\hat{U}|JM\rangle \tag{3.30}$$

where J and M are the total spin and spin projection along the z-axis respectively for the decaying (parent) particle, and $\hat{U}$ is some operator which takes the initial state to the final state. The helicity states and rotation operators developed in the preceding sections will be used to facilitate this. First, we will consider a two-body decay $B \to \alpha\beta$, and then extend these results to derive the expression for sequential decays.

3.4.1 Two-body decays

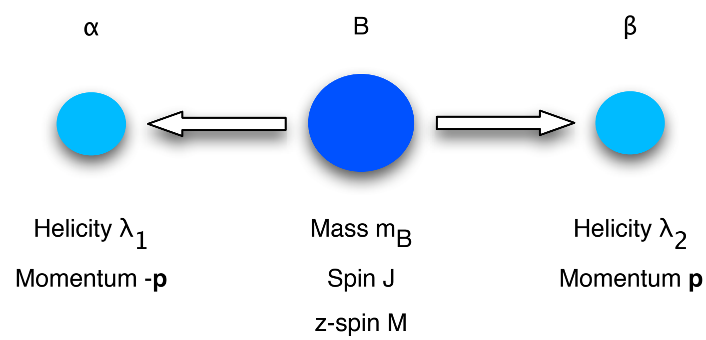


Figure 3.3: Two-body decay

Consider the situation shown in figure 3.3. The final state can immediately be written using equation 3.27

$$|f\rangle = |p_f\theta_f\varphi_f\lambda_1\lambda_2\rangle = (2\pi)^3\left[\frac{4m_B}{p_f}\right]^{\frac{1}{2}}|\theta_f\varphi_f\lambda_1\lambda_2\rangle|\mathbf{P}_f^B\rangle \tag{3.31}$$

The decay amplitude for the particle B to decay to the final state f is

$$A\left(B\to f\right) = (2\pi)^3\left[\frac{4m_B}{p_f}\right]^{\frac{1}{2}}\langle\theta_f\varphi_f\lambda_1\lambda_2|\hat{U}|JM\rangle \tag{3.32}$$

Now, suppressing the constants

$$\begin{aligned} A\left(B\to f\right) &= \langle\theta_f\varphi_f\lambda_1\lambda_2|\hat{U}|JM\rangle \\ &= \sum_{J_f,M_f}\langle\theta_f\varphi_f\lambda_1\lambda_2|J_fM_f\lambda_1\lambda_2\rangle\langle J_fM_f\lambda_1\lambda 2|\hat{U}|JM\rangle \\ &= \sum_{J_f,M_f}\left[\frac{2J+1}{4\pi}\right]^{\frac{1}{2}}\mathscr{D}^J_{M\lambda}\left(\varphi_f,\theta_f,-\varphi_f\right)\delta_{J_f,J}\delta_{M_f,M}\langle\lambda_1\lambda_2|\hat{U}|M\rangle \\ &= \left[\frac{2J+1}{4\pi}\right]^{\frac{1}{2}}\mathscr{D}^J_{M\lambda}H_{\lambda_1\lambda_2} \end{aligned} \tag{3.33}$$

where $\lambda = \lambda_1 - \lambda_2$. Summing over helicity states we get

$$A\left(B\to f\right) = \sum_{\lambda_1,\lambda_2}\left[\frac{2J+1}{4\pi}\right]^{\frac{1}{2}}\mathscr{D}^J_{M\lambda}H_{\lambda_1\lambda_2} \tag{3.34}$$

The angular distribution is given by

$$\frac{d\sigma}{d\theta_f \varphi_f} = |A\,(B \to f)\,|^2 \tag{3.35}$$

There is clearly the potential for interference terms to crop up in the angular distributions - in fact they will appear whenever there is more than one helicity state. These interference effects between helicity amplitudes have tell-tale signatures in the angular distribution, and they are seen in a wide range of decay processes.

Decay of a scalar to two vector particles

We now consider a concrete example of a single-stage decay, that of a scalar particle (spin-0) to two vectors (spin-1) - $B \to V_a V_b$. Examples of such decays include $B_s \to J/\psi\phi$, $B_d \to J/\psi K^{0\star}$ and $B \to D^\star\rho$, and the angular distributions from such decays will be of crucial importance later on in this work.

We have the following situation:

$$J = M = 0 \tag{3.36}$$

Now there is a condition on the helicities of the final state:

$$|\lambda_1 - \lambda_2| \leq J \tag{3.37}$$

To understand why this should be, one can consider the converse case - where the difference between the helicities is greater than J. Given that J is the total angular momentum - the sum of the spin S and the orbital angular momentum L - this case would imply a component of L in the direction of the decay momentum vector $\mathbf{p}$. But from the definition of the orbital angular momentum, $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, this can be seen to be impossible.

Taking 3.37 into account, the following restriction is placed on the helicities:

$$\begin{aligned} \lambda &= 0 && (3.38)\\ (\lambda_1, \lambda_2) &= (-1,-1)\,, (0,0)\,, (1,1) && (3.39) \end{aligned}$$

The decay amplitude is hence a sum of *three* helicity amplitudes:

$$A\,(B \to V_1 V_2) = \frac{1}{\sqrt{4\pi}} \mathscr{D}^0_{00}\,(H_{-1,-1} + H_{0,0} + H_{1,1}) \tag{3.40}$$

Using equation 3.14, $\mathscr{D}^0_{00}$ can be written

$$\mathscr{D}^0_{00} = e^{i0\gamma} d^0_{00}(\beta) e^{i0\alpha} \tag{3.41}$$

and the value of d^0_{00} can be read off from the tables in [21] or generated from equation 3.17. It is in fact precisely 1, giving

$$A\left(B \to V_1 V_2\right) = \frac{1}{\sqrt{4\pi}}\left(H_{-1,-1} + H_{0,0} + H_{1,1}\right) \tag{3.42}$$

Obviously when we disregard all subsequent decays we will see no interesting angular behaviour - the two vectors will decay back-to-back and that is the end of it. This is expressed by the fact that no angles appear in the decay amplitude above. The point of this exercise has been to show that three helicity amplitudes appear. In almost every paper involving angular analyses of decays of this type, statements appear such as "the angular distribution is described by bilinear combinations of decay amplitudes". The origin of these combinations is now clear - they appear when the square of the modulus of the above amplitude is calculated to evaluate the decay probability.

We shall now go on to consider the more interesting case where the two vectors subsequently decay themselves, allowing us to see non-isotropic angular distributions.

3.4.2 Sequential Decays

Let us consider the case where, in 3.3, the particles α, β decay to two further particles each:

$$\alpha \to a_1 + a_2$$
$$\beta \to b_1 + b_2$$

with helicities $\lambda_{a_1}, \lambda_{a_2}, \lambda_{b_1}, \lambda_{b_2}$. First we must carefully define a set of decay angles for the system. The following choice is made [88]:

1. The *overall rest frame* is taken to be the rest frame of the decaying B particle.

2. The polar decay angles for the particles a_1, a_2, (θ_a, φ_a) are defined in the *rest frame* of α, with the z-axis being defined by the momentum $-\mathbf{p}$ of α.

3. The polar decay angles for the particles b_1, b_2, (θ_b, φ_b) are defined in the *rest frame* of β, with the z-axis being defined by the momentum $\mathbf{p}$ of β.

4. The sum of the two angles $\varphi_a + \varphi_b$ is denoted ϕ. This angle is in fact the angle between the planes formed by the momenta of the final state particles.

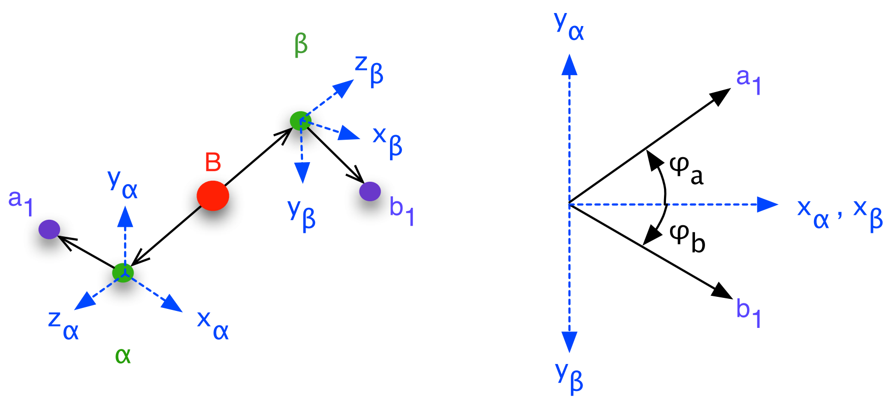


Figure 3.4: Definition of decay angles for sequential decays

This arrangement is shown pictorially in figure 3.4.

The decay amplitude is written, using equation 3.34

$$A = \sum_{\lambda_1,\lambda_2} \mathscr{D}^J_{M\lambda} H_{\lambda_1\lambda_2} \mathscr{D}^{s_\alpha}_{\lambda_1,\lambda_{a_1}-\lambda_{a_2}} (\varphi_a, \theta_a, 0) \, \mathscr{D}^{s_\beta}_{\lambda_2,\lambda_{b_1}-\lambda_{b_2}} (\varphi_b, \theta_b, 0) \tag{3.43}$$

where J, M, λ, λ_1 and λ_2 are as before, and $s_{\alpha,\beta}$ are the spins of α and β. Using equation 3.14 we write

$$A = \sum_{\lambda_1,\lambda_2} \mathscr{D}^J_{M\lambda} H_{\lambda_1\lambda_2} e^{i\lambda_1\phi} d^{s_\alpha}_{\lambda_1,\lambda_{a_1}-\lambda_{a_2}} (\theta_a) \, d^{s_\beta}_{\lambda_2,\lambda_{b_1}-\lambda_{b_2}} (\theta_b) \tag{3.44}$$

3.4.3 Angular Distribution for $B_s \to J/\psi\,(\mu\mu)\,\phi\,(KK)$

We now derive the complete angular distribution for the decay $B_s \to J/\psi\,(\mu\mu)\,\phi\,(KK)$, which is a $S \to VV$ decay as described above. Note that the decay $B_d \to J/\psi\,(\mu\mu)\,K^{0*}\,(K\pi)$ has an identical angular distribution. Using the notation from the previous sections, we have the following situation:

$$B = B_s, \quad \alpha = J/\psi, \quad a_1 = \mu^+, \quad a_2 = \mu^-, \quad \alpha = \phi, \quad b_1 = K^+, \quad b_2 = K^-$$

The final decays have the following helicity states [88]:

$$\lambda_{K^+,K^-} = 0 \tag{3.45}$$

$$(\lambda_{\mu^+}, \lambda_{\mu^-}) = \left(+\frac{1}{2}, -\frac{1}{2}\right) \text{or} \left(-\frac{1}{2}, +\frac{1}{2}\right) \tag{3.46}$$

So we have:

$$\lambda_{a_1} - \lambda_{a_2} = \pm 1, \qquad \lambda_{b_1} - \lambda_{b_2} = 0 \tag{3.47}$$

We now switch to the notation normally used in such channels, namely

$$\theta_1 = \theta_a\,(J/\psi)\,, \theta_2 = \theta_b\,(\phi) \tag{3.48}$$

The helicity states are given by (using equation 3.44)

$$A^{(\chi)} = \sum_{m=-1,0,1} H_m e^{im\phi} d^1_{m\chi}(\theta_1)\, d^1_{m0}(\theta_2) \tag{3.49}$$

where $\chi = \lambda_{\mu^+} - \lambda_{\mu^-} = \pm 1$ and where m is given by

$$m = (-1, 0, 1) = \left(\lambda_{J/\psi}\lambda_\phi = \{-1,-1\}, \{0,0\}, \{1,1\}\right) \tag{3.50}$$

The angular distribution is given by the incoherent sum of the distributions for the two final state muon configurations [88]

$$\frac{d\sigma}{d\theta_1 d\theta_2 d\phi} = |A^{(+1)}|^2 + |A^{(-1)}|^2 \tag{3.51}$$

We are now in a position to explicitly evaluate the amplitudes. Let us re-write the amplitudes as

$$A^{(\lambda)} = \sum_m H_m g_m^{(\lambda)}(\theta_1\theta_2\phi) \tag{3.52}$$

where $m \in \{-1, 0, 1\}$ and $\lambda \in \{-1, 1\}$ and

$$g_m^{(\lambda)}(\theta_1\theta_2\phi) = e^{im\phi} d^1_{m\lambda}(\theta_1)\, d^1_{m0}(\theta_2) \tag{3.53}$$

These values can be read off the table in [21]. They are:

$$\begin{aligned}
g_{+1}^{(+1)} &= -\frac{1}{2\sqrt{2}}(1+\cos\theta_1)\, e^{i\phi}\sin\theta_2 \\
g_{0}^{(+1)} &= \frac{1}{\sqrt{2}}\sin\theta_1\cos\theta_2 \\
g_{-1}^{(+1)} &= \frac{1}{2\sqrt{2}}(1-\cos\theta_1)\, e^{-i\phi}\sin\theta_2 \\
g_{+1}^{(-1)} &= -\frac{1}{2\sqrt{2}}(1-\cos\theta_1)\, e^{i\phi}\sin\theta_2 \\
g_{0}^{(-1)} &= -\frac{1}{\sqrt{2}}\sin\theta_1\cos\theta_2 \\
g_{-1}^{(-1)} &= \frac{1}{2\sqrt{2}}(1+\cos\theta_1)\, e^{-i\phi}\sin\theta_2
\end{aligned} \tag{3.54}$$

The square of the amplitude in equation 3.52 is given by

$$\begin{aligned}
|A^{(\chi)}|^2 &= \left(\sum_m H_m g_m^{(\chi)}\right)^* \left(\sum_n H_n g_n^{(\chi)}\right) \\
&= \sum_m |H_m|^2 |g_m^{(\chi)}|^2 \\
&\quad +2\sum_{m<n}\left(\Re\left(H_m^* H_n\right)\Re\left(g_m^{(\chi)*}g_n^{(\chi)}\right) - \Im\left(H_m^* H_n\right)\Im\left(g_m^{(\chi)*}g_n^{(\chi)}\right)\right)
\end{aligned} \tag{3.55}$$

The process of substituting in the appropriate values of g, taking the real and complex parts and summing over m to obtain the full expression for 3.51 is long and tedious but mathematically trivial. The result is as follows.

$$\begin{aligned}\frac{d\sigma}{d\theta_1 d\theta_2 d\phi} &= \left(|H_+|^2 + |H_-|^2\right)\left(1+\cos^2\theta_1\right)\sin^2\theta_2 + 4|H_0|^2\sin^2\theta_1\cos^2\theta_2 \\ &\quad -2\left\{\Re\left(H_+^* H_-\right)\cos 2\phi + \Im\left(H_+^* H_-\right)\sin 2\phi\right\}\sin^2\theta_1\sin^2\theta_2 \\ &\quad -\left\{\Re\left((H_+ + H_-)^* H_0\right)\cos\phi + \Im\left((H_+ - H_-)^* H_0\right)\sin\phi\right\} \\ &\quad \times\sin 2\theta_1 \sin 2\theta_2 \end{aligned} \tag{3.56}$$

The Transversity Basis

In the appendix of Richman [16] it is shown that the behaviour of a helicity state under the parity transformation is described by

$$\hat{\mathscr{P}}|JM,\lambda_a,\lambda_b\rangle = \pi_a\pi_b\left(-1\right)^{J-s_a-s_b}|JM,-\lambda_a,-\lambda_b\rangle \tag{3.57}$$

where $s_{a,b}$ and $\pi_{a,b}$ are the spins and parities of the *child* particles (that is, the intermediate states - i.e. in this case the J/ψ and ϕ). If this is applied to the $S\to VV$ case we obtain

$$\begin{aligned}\hat{\mathscr{P}}|0,0,-1,-1\rangle &= |0,0,1,1\rangle \\ \hat{\mathscr{P}}|0,0,0,0\rangle &= |0,0,0,0\rangle \\ \hat{\mathscr{P}}|0,0,1,1\rangle &= |0,0,-1,-1\rangle \end{aligned} \tag{3.58}$$

It is immediately clear that the helicity basis states are *not* eigenstates of parity. In many cases it is extremely useful that the states are parity eigenstates, so a change of basis is needed. The new basis, which is widely referred to as the *transversity basis*, is constructed as follows:

$$\begin{aligned} A_\parallel &= \frac{H_+ + H_-}{\sqrt{2}} \\ A_0 &= H_0 \\ A_\perp &= \frac{H_+ - H_-}{\sqrt{2}} \end{aligned} \tag{3.59}$$

It is clear that $A_\parallel$ and A_0 have positive parity, whilst $A_\perp$ has negative parity. The angular distribution 3.56 is easily re-expressed in terms of the transversity basis; the result is

$$\frac{d\sigma}{d\theta_1 d\theta_2 d\phi} = \frac{9}{64\pi}\Bigg[4|A_0|^2\sin^2\theta_1\cos^2\theta_2$$

$$
\begin{aligned}
&+|A_\parallel|^2 \left[\left(1+\cos^2\theta_1\right)\sin^2\theta_2 - \sin^2\theta_1\sin^2\theta_2\cos 2\phi\right] \\
&+|A_\perp|^2 \left[\left(1+\cos^2\theta_1\right)\sin^2\theta_2 + \sin^2\theta_1\sin^2\theta_2\cos 2\phi\right] \\
&+2\Im\left(A_\parallel^* A_\perp\right)\sin^2\theta_1\sin^2\theta_2\sin 2\phi \\
&-\sqrt{2}\Re\left(A_0^* A_\parallel\right)\sin 2\theta_1\sin 2\theta_2\cos\phi \\
&\left.+\sqrt{2}\Im\left(A_0^* A_\perp\right)\sin 2\theta_1\sin 2\theta_2\sin\phi\right]
\end{aligned}
\tag{3.60}
$$

where the overall normalization factor is chosen such that, after integration over all angles, $\Gamma = |A_0(t)|^2 + |A_\parallel(t)|^2 + |A_\perp(t)|^2$ [89].

This expression is identical to the equation given in [89] (page 43, equation 4.21) for decays of this type. It will form a crucial part of the analysis presented in this work.

Chapter 4

The ATLAS experiment: hardware and research programme

We now turn from theoretical to experimental considerations. Whilst only one member of the Standard Model particle set - the Higgs Boson - remains to be observed, it is strongly suspected that nature permits processes that are not described by the machinery of the Model. Colloquially referred to as "Physics beyond the Standard Model", these interactions and particles are the principal subject of contemporary experimental particle physics - put simply, the aim is to prove that our current knowledge of the subnuclear world is merely an approximation to more fundamental physics.

In order to bring this about, particle accelerators with ever-higher centre-of-mass energies and beam intensities are being constructed. The Large Hadron Collider, under construction at CERN and due to commence operations in mid-2007, will be the world's most powerful synchrotron. Combining unprecedented centre-of-mass energy and intensity, it should provide the wherewithal for its four detectors to observe these "beyond the Standard Model" entities and processes.

In this chapter the LHC machine itself is briefly discussed. The ATLAS detector, with which this work is concerned, is then described. Finally, the ATLAS scientific programme is considered, paying particular attention to the plans for B-physics measurements.

4.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a dual-beam proton synchrotron, with four collision points. It occupies the longest of the CERN tunnels, which is 26659 metres long and has a maximum depth of 175 metres. Electrical energy from the local power grid is coupled to the beam using 16 superconducting radio-frequency oscillators (eight per beam), each of which contains four cavities. The oscillation frequency of the standing wave established in these cavities is $400.8MHz$. The beam is directed in a curved trajectory through the use of 1232 superconducting dipole magnets, and is focused with approximately 860 quadrupoles. Undesirable resonances in the beam are suppressed through the use of over 6200 additional correction magnets. The total power consumption of the machine, including R.F. cavities, magnets and cooling, is around $120MW$. The equipment supplying the magnets and cavities with liquid helium coolant is the largest cryogenic system ever constructed.

The beams themselves will have a nominal energy of $7TeV$, giving a centre-of-mass energy of $14TeV$. The input beam will be provided by the existing complex of CERN accelerators, namely the $50MeV$ linear accelerators, the $1.4GeV$ Proton Synchrotron booster, the $26GeV$ Proton Synchrotron and finally the $450GeV$ Super Proton Synchroton (SPS). Due to the oscillating polarity of the electric fields providing the accelerating force, the beam is not a continuous stream of particles but a series of bunches. Each bunch will have a length of approximately one metre.

At each of the collision points, bunches pass through each other at a frequency of $40MHz$ (that is, every $25ns$). A detector is positioned at each of the interaction points: CMS[1] and ATLAS are general central detectors, LHCb is a dedicated B-physics experiment and ALICE[2] is for the study of heavy ion collisions. The diameter of the interaction regions is envisaged to be around 16 microns for ATLAS and CMS.

The **luminosity** of a collision is the number of particles passing through a unit area of the interaction region, per unit time. If bunches containing n_1 and n_2 particles pass through one another with frequency f, the luminosity is given by

$$\mathscr{L} = f\frac{n_1 n_2}{4\pi\sigma_x\sigma_y} \tag{4.1}$$

where $\sigma_{x,y}$ characterize the Gaussian profiles of the beam in the vertical and horizontal distances. It is envisaged that the LHC will operate in three luminosity phases - $5.0 \times 10^{32} cm^{-2}s^{-1}$ initially, through to the "low luminosity" phase (until 2010) throughout which the peak operating luminosity will be $2 \times 10^{33} cm^{-2}s^{-1}$. Beyond this time the luminosity will rise to its nominal peak value, $10^{34} cm^{-2}s^{-1}$. The number of particles in each bunch will therefore vary from $\sim 5.0 \times 10^{10}$ to $\sim 2.5 \times 10^{11}$ (noting that in general, only 80% of the possible bunch spaces in the LHC will actually contain protons). The arrangement of the CERN accelerator complex is shown in figure 4.1 and the layout of the LHC and its four experiments in figure 4.2. The

[1] Compact Muon Solenoid
[2] A LHC Ion Collision Experiment

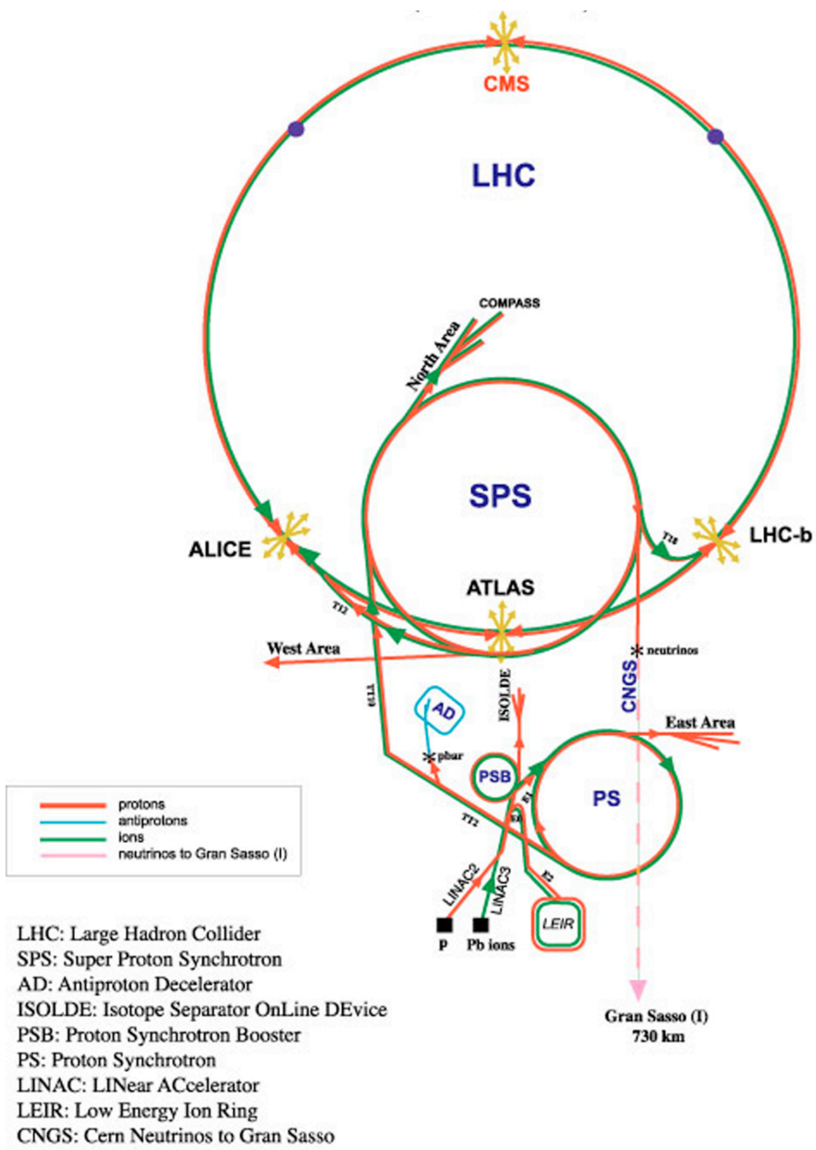


Figure 4.1: The CERN accelerator complex - taken from [75]

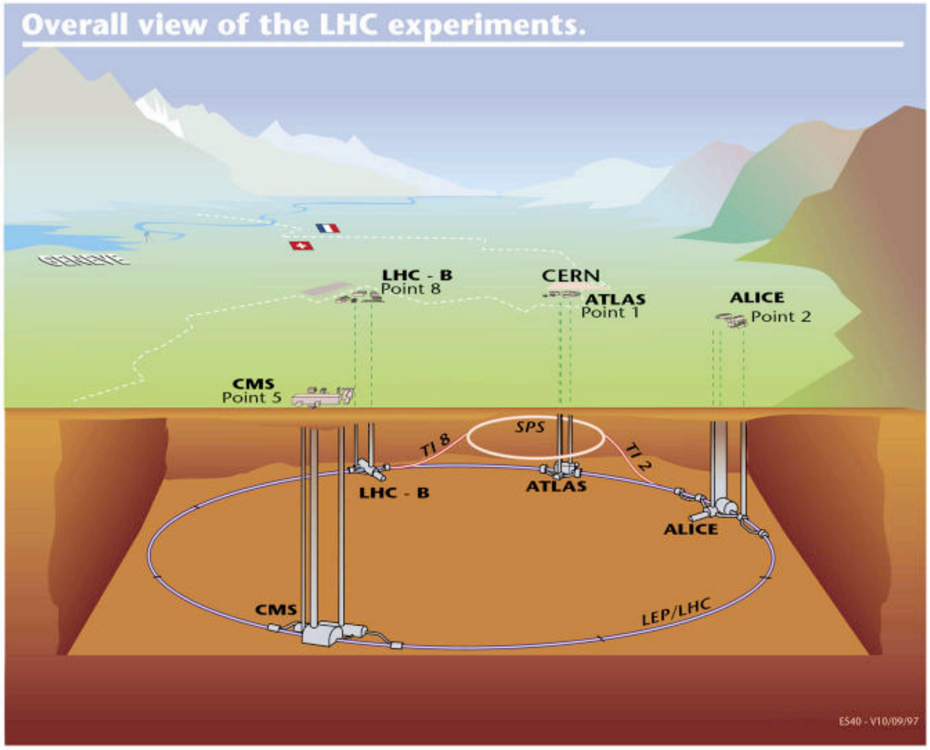


Figure 4.2: The LHC and its four detectors

number of observed events of a given signal process $P_0 \to P_1 \cdots P_N$ with branching ratio $B = B_{P_0 \to P_1 \cdots P_N}$ is

$$N_{\text{obs}} = \mathscr{L} T \sigma_{P_0} B \epsilon_r \tag{4.2}$$

where $\mathscr{L}$ is the luminosity, T is the total time during which the collisions are occurring, σ_{P_0} is the production cross section for the particle P_0 and ϵ_r is the reconstruction efficiency for the channel, which can only be determined from simulation. Current calculations are based on the assumption that during the course of a year, the LHC will run for 10^7 seconds.

During a bunch crossing, it is likely that more than one proton-proton collision will occur. The number of these **pile-up** events is given by a Poisson distribution with an average of

$$< n >= \frac{\mathscr{L} \sigma_{\text{inelastic}}}{f} \tag{4.3}$$

The expected total inelastic cross section $\sigma_{\text{inelastic}}$ is around $70mb$ [24] and so the number of pile-up events per bunch crossing can be expected to be 4 at low luminosity and 22 during the high luminosity runs. The majority are processes involving *low* energy transfer, and produce large numbers of low momentum ($\sim 0.5GeV$) hadrons. These are referred to as *minimum bias* events. The average number of charged particles per unit of pseudorapidity, for a single minimum bias events, is 7.5 [115]. The result is that many hundreds of tracks are produced per bunch crossing. This places severe requirements on the performance of the detector and the reconstruction software.

Figure 4.3 shows the total cross section along with cross sections for specific physics processes, as a function of centre-of-mass energy. It can immediately be seen that the total cross section is many orders of magnitude greater than some of the physics processes. The success of an experiment therefore depends crucially on certain signatures of interesting physics, which the detector's data acquisition systems can use as a *trigger*. Upon activation of a trigger, the event is recorded; all other data is discarded. Only in this way is it possible to avoid being overwhelmed by uninteresting minimum bias events. Typical trigger entities include high-p_T leptons, which are rare in minimum bias events but are frequently produced in the events of interest; jets and photons may also be used.

4.2 An overview of the ATLAS detector

The ATLAS experiment is the largest of the LHC experiments, both in terms of the number of collaborators and the physical size of the machine. Indeed, with a length of 46m, a radius of 11m and a mass exceeding 7000 metric tonnes, it is the largest accelerator-based particle detector ever constructed. The detector is described in detail in the experiment's Technical Design Reports [117] - here, the briefest of descriptions is given.

4.2.1 Nomenclature

The definitions of the x, y and z axes is given in figures 4.4 and 4.5. The azimuthal angle ϕ is measured around the beam axis, as shown in figure 4.5, and the polar angle θ is measured from the beam axis, as shown in figure 4.4. The **pseudorapidity** η is given by

$$\eta = -\ln\tan\frac{\theta}{2} \tag{4.4}$$

This value is equal to zero when the angle θ is 90°, and approaches infinity as θ approaches zero (that is, as it lies closer to the beam axis). Transverse momentum p_T and missing transverse energy are defined in the xy plane. The distance ΔR in $\eta\phi$ space is given by

$$\Delta R = \sqrt{\Delta^2\eta + \Delta^2\phi} \tag{4.5}$$

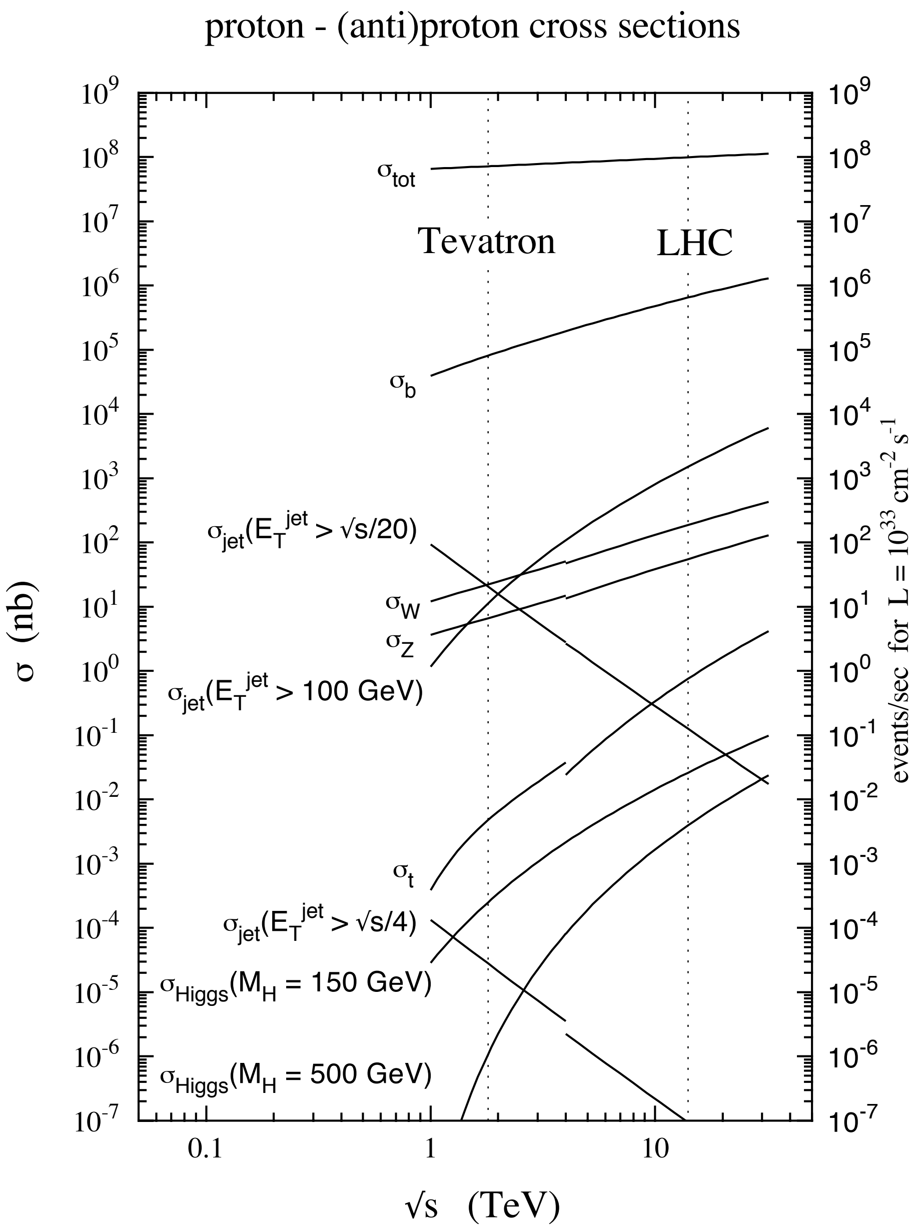


Figure 4.3: Proton-proton cross sections versus $\sqrt{s}$ (taken from [25])

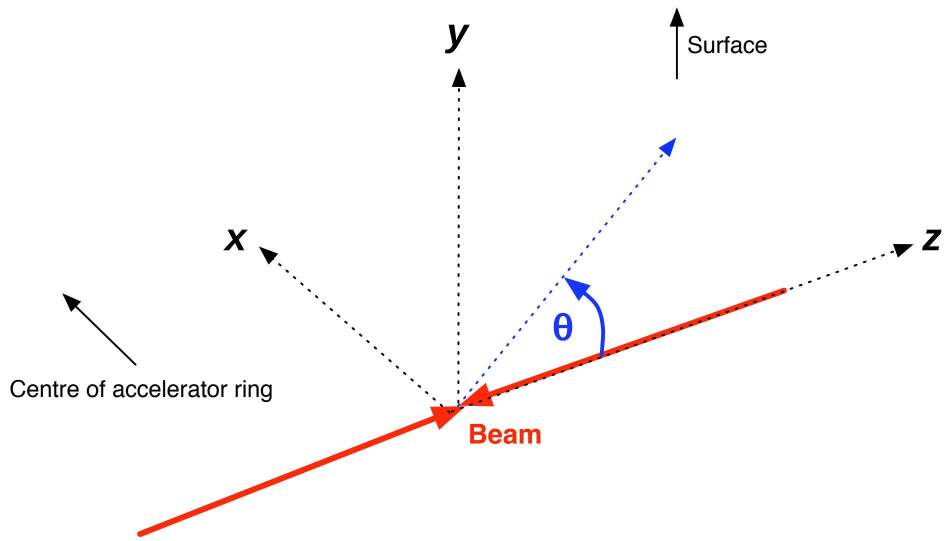


Figure 4.4: Definition of the angle θ

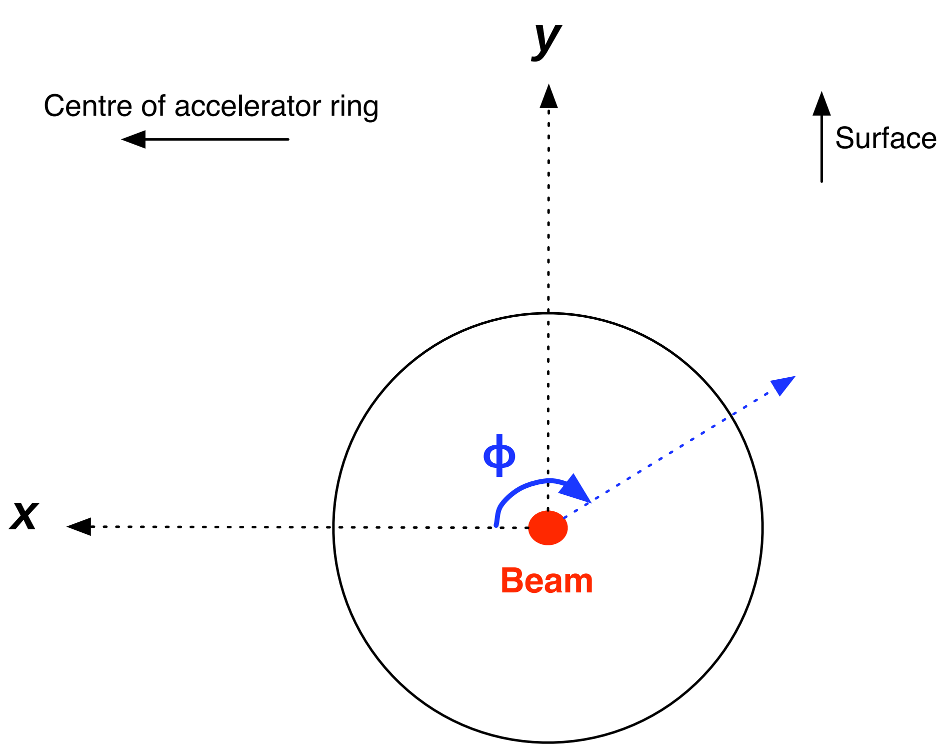


Figure 4.5: Definition of the angle ϕ

4.2.2 Design requirements

ATLAS is required to be a general-purpose detector, able to trigger on as many physics signatures as possible, over the full range of LHC luminosity conditions. As such, the design criteria include the following [117]:

- Efficient **tracking**, enabling full event reconstruction at low luminosity and identification of leptons, photons and heavy flavour jets at high luminosity;

- High-precision muon momentum measurements, using the muon spectrometer alone at high luminosities;

- Electro-magnetic calorimetry for photon and electron identification and energy measurements, and hadronic calorimetry for jet and missing energy measurements;

- Triggering on low p_T particles, enabling selection of physics events with high efficiency;

- Large acceptance in pseudorapidity η and almost full azimuthal angle coverage ϕ.

ATLAS meets these requirements through a number of features:

- A sophisticated *inner detector* which produces fine-granularity tracking measurements with up to 40 points per track;

- A system of high granularity electro-magnetic and hadronic *calorimeters* with excellent energy and position resolution;

- A full-coverage *muon* system, which, on account of the $4T$ field provided by toroidal magnets, can achieve precise muon identification and momentum measurements;

- A three-level, fast, highly selective and flexible *trigger mechanism*.

A schematic diagram of the ATLAS detector is shown in figure 4.6.

4.2.3 The magnet system

The magnets provide bending fields for the inner detector and the muon systems, facilitating the measurement of momentum. There are two magnet systems in ATLAS - the *central solenoid* and the *toroids*, which supply the inner detector and the muon systems respectively. Both are cooled to superconducting temperatures.

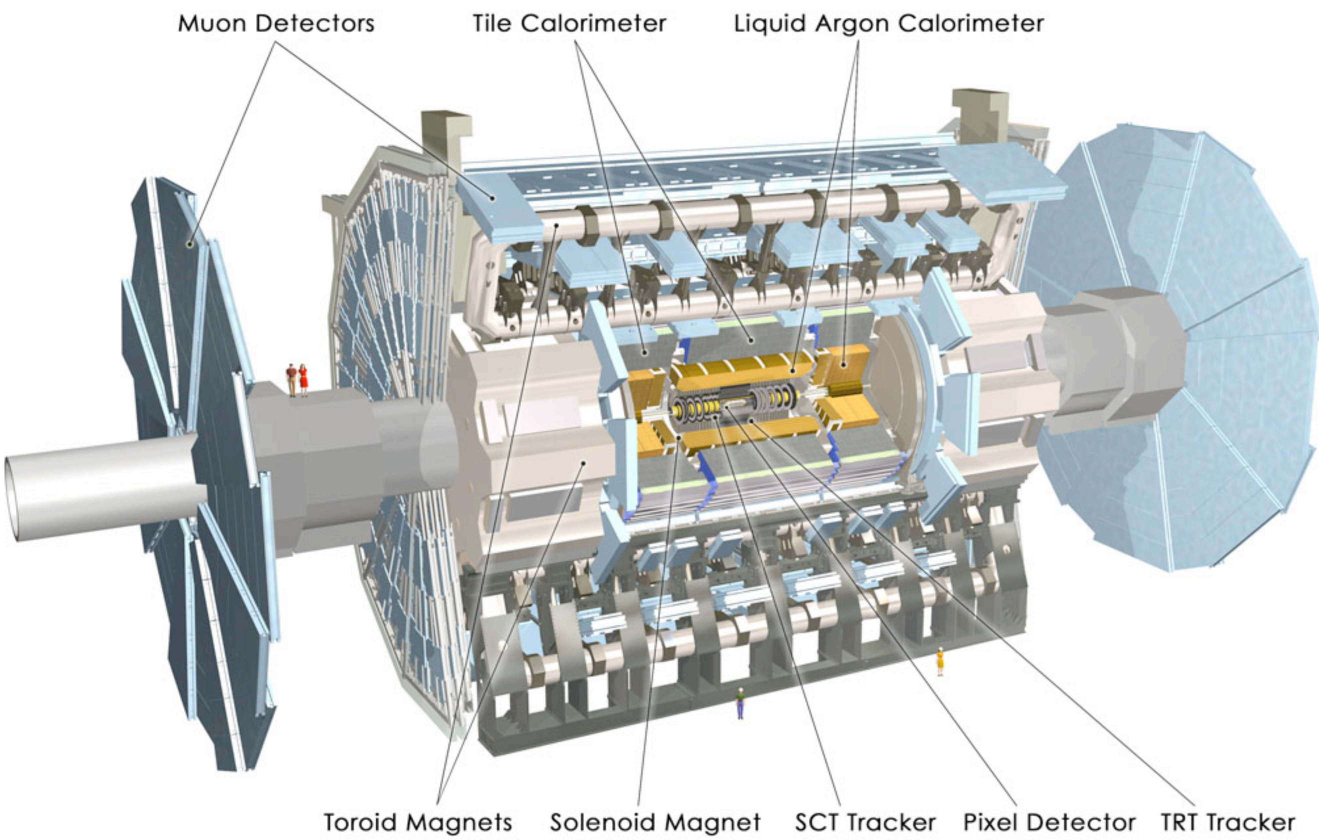


Figure 4.6: The ATLAS detector. Taken from [75]

The cylindrical central solenoid, into which the inner detector is inserted, provides a central field of 2T in the core of the inner detector, with a peak field of 2.6T at the surface of the magnet. The outer and inner bore radii are $1.32m$ and $1.22m$ respectively, and its length is $5.3m$. The total mass of the central solenoid assembly is 5.7 tonnes. It draws a current of $7.6kA$. Cooling is provided by liquid helium at $4.5K$. The central solenoid is positioned upstream of the calorimetry, so keeping the quantity of material to a minimum is an essential design requirement. This has been achieved by placing the central solenoid and the electromagnetic calorimeter in the *same* cryostat.

The toroid magnet system, which supplies the muon spectrometer with bending field, is comprised of two components - the barrel toroid, and the end-cap toroids. The barrel toroid consists of eight "race-track" coils which are arranged around the outside of the calorimetry in a torus assembly. Each coil has a length of $25.3m$ and a width of $5.4m$. The total mass of the assembly is 830 tonnes, and supplies a peak field of $3.9T$ whilst drawing a current of $20.5kA$.

There are two end-cap toroids, positioned beyond the forward hadronic calorimetry within the barrel toroid assembly. They supply a peak field of $4.1T$ from a current of $20kA$, have a mass of 239 tonnes each and have a radius of $5.4m$.

A sketch of the toroid system is shown in figure 4.7.

4.2.4 The Inner Detector

The rôle of inner detector is the provision of precise tracking information, for making high-quality measurements of particle momenta and precisely locating secondary and primary decay vertices. The device sits within the central solenoid. The total length of the inner detector cavity is $7m$, and it has a radius of $57.5cm$. It consists of three units - a barrel part extending $\pm 80cm$ from the centre, with two end-caps seated beyond this, making up the rest of the length.

Two images of the inner detector, showing the layout in 2-D profiles and as a 3-D schematic, are shown in figures 4.8, 4.9 and 4.10. All dimensions and values quoted in this section can be assumed to be from [117] or the inner detector technical design report, [72].

The inner detector is required to produce high-precision tracking measurements, and this calls for very fine granularity detection elements close to the interaction point, given the extremely high track density which is expected at the LHC. This is achieved through the use of two semi-conductor technologies - pixel detectors, and silicon microstrips (SCT, or semi-conductor tracker). The high monetary value of these elements, and the additional material that they introduce, means that their

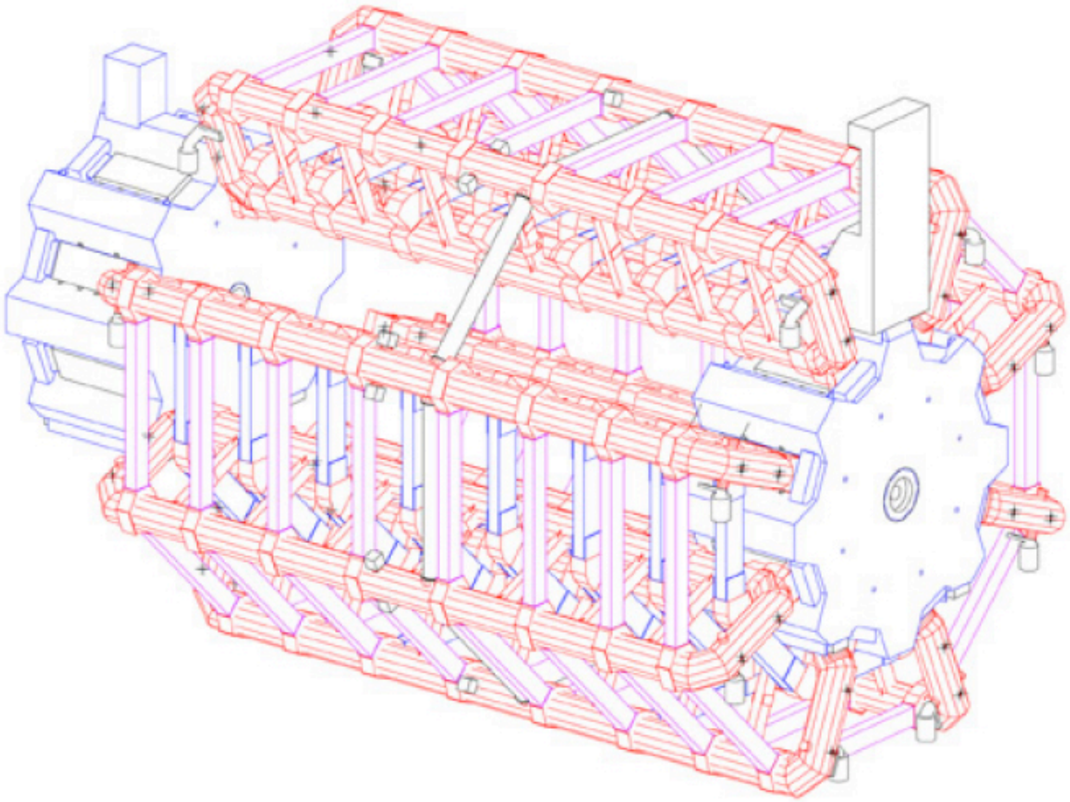


Figure 4.7: The toroid magnet system. Taken from [75]

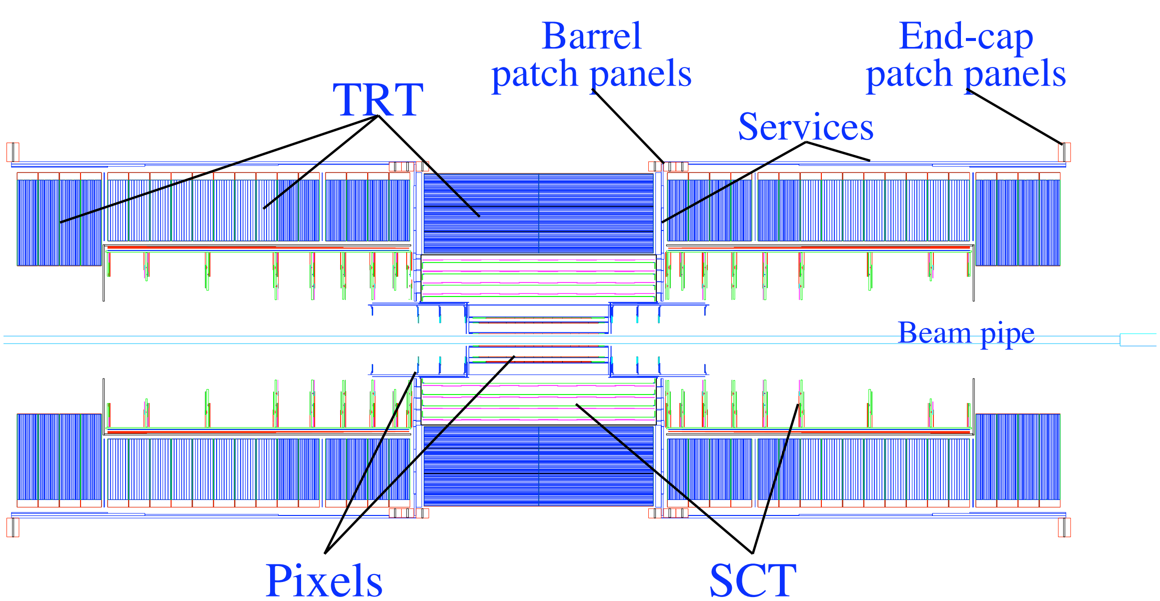


Figure 4.8: 2D schematic of the ATLAS inner detector, rz projection. Taken from [72].

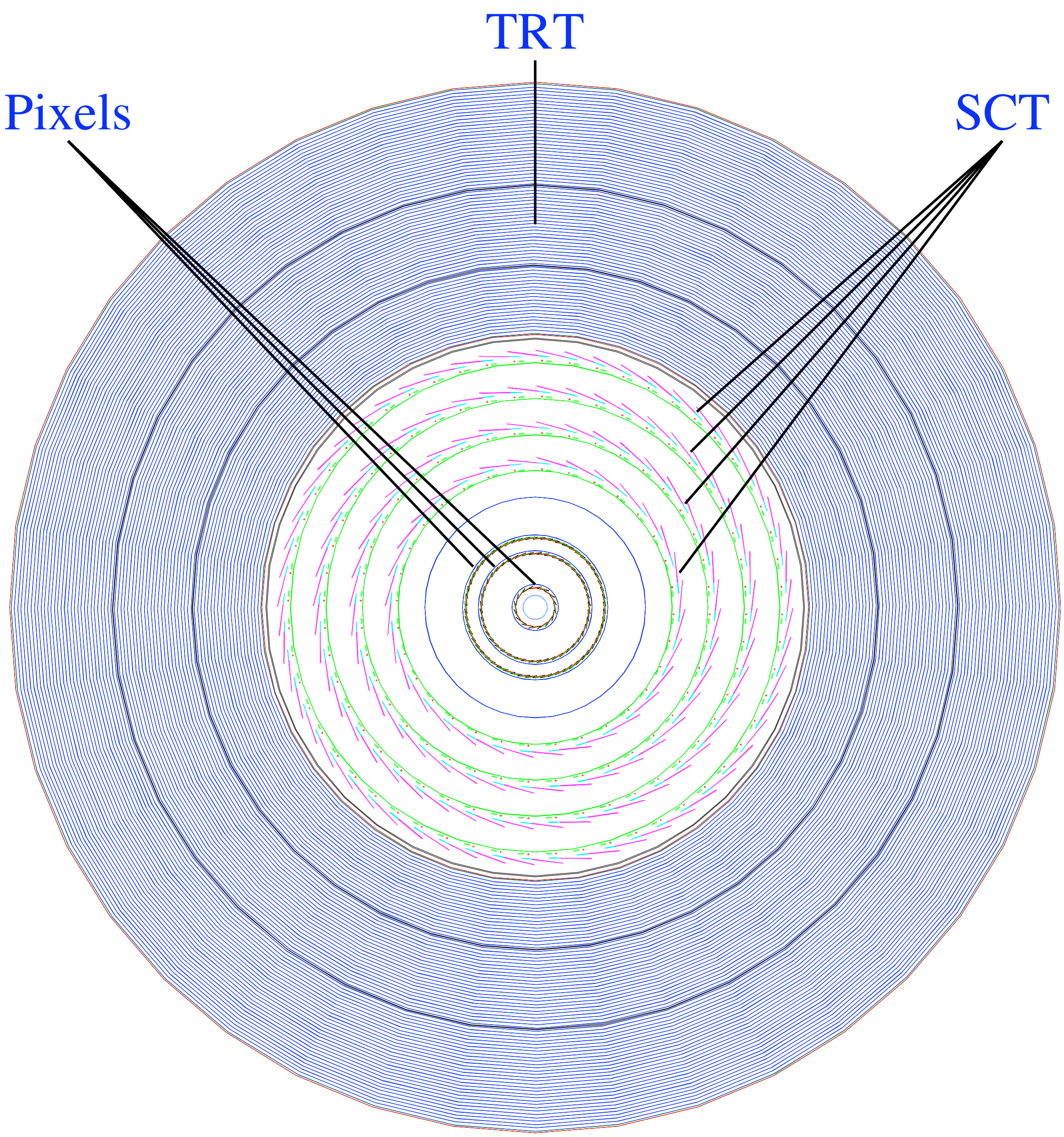


Figure 4.9: 2D schematic of the ATLAS inner detector, $r\phi$ projection. Taken from [72].

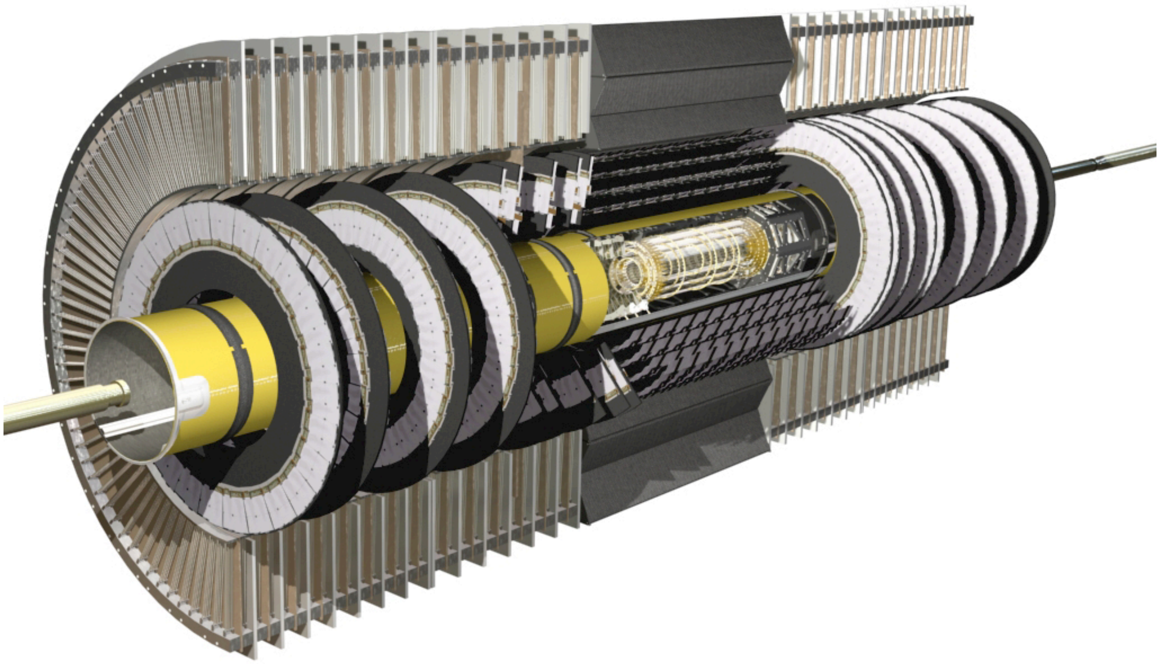


Figure 4.10: 3D schematic of the ATLAS inner detector. Taken from [75].

use must be strictly limited to regions of the highest track density. Further out straw tubes are used to provide large numbers of space points per track, at a much reduced cost.

The pixel detector

The pixel detection elements provide extremely fine granularity tracking measurements very close to the interaction point. They are segmented in a two-dimensional fashion, $50\mu m$ in the $R\phi$ direction and $300\mu m$ in the z direction, so there is little ambiguity as to the position of a hit. Given their proximity to the collision region, these components must be resistant to irradiation. The read-out electronics must be bonded directly to the detector elements due to the very brief interval between bunch crossings ($25ns$) and this places stringent demands on this equipment, given the limitations of space and the requirement that as little material as possible is located in this region (to reduce multiple scattering). The total number of pixel elements in the inner detector is approximately 140 million.

In the barrel section, the pixels are found in three concentric layers, located at approximate radii of $4.0cm$, $11.0cm$ and $14.2cm$. The innermost layer is particularly significant as a space-point this close to the interaction point allows precise secondary vertex finding, and hence lifetime measurements of short-lived heavy particles such as beauty hadrons. This is consequently referred to as the *B-layer*.

Pixels are also located in five disks on either side of the barrel, which vary in radii from 11 to 20cm, which completes the angular coverage of the system. The silicon pixels are mounted on modules, each of which contains 16 chips, with an array 24×160 pixels per chip. Each module has dimensions of $62.4mm \times 21.4mm$, and contains 61440 pixels. The modules used across the detector are identical, whether they are found in the barrel or the disks.

The semiconductor tracker (SCT)

The SCT provides an average of eight precise hits per track in intermediate-radius regions. It is again arranged in three regions - a barrel and two end-caps. The barrel SCT consists of four cylindrical layers of detector elements, which each have the dimensions $6.36 \times 6.40cm$. One element consists of 768 read-out strips. A single SCT module consists of four such detectors. The forward SCT uses 18 disks (nine on either side of the interaction point). The modules are of a similar design to the barrel SCT, except for the shape of the strips which are tapered towards the centre of the disks. The SCT contains $61m^2$ of silicon detectors, connected to 6.2 million read-out channels. The spatial resolution is $16\mu m$ in $R\phi$ and $580\mu m$ in z, meaning that two tracks can be distinguished if they are separated by more than $200\mu m$ [72].

The Transition Radiation Tracker (TRT)

The TRT, which is located in the outermost part of the inner detector, performs two rôles - providing additional tracking capability by observing hits in the outer region of the inner detector, and distinguishing pions from electrons (which leave very similar signatures in the next layer downstream, the electromagnetic calorimeter).

The TRT detector element is a single wire proportional chamber known as a **straw tube**. Each tube is $4mm$ in diameter, and contains a $30\mu m$ wire running along the length. The maximum length of such a straw is 144cm, in the barrel part of the TRT. The TRT is again divided into a barrel and two end-caps, with the barrel containing 50000 tubes and the end-caps containing 320000. There are 420000 independent channels associated with the TRT, each of which provides a drift-time measurement, and has two thresholds, allowing the distinction between tracking hits (which pass the lower threshold) and transition radiation (which passes the higher threshold). Transition radiation photons are created in a radiator which is inserted between the straws. The TRT technology is intrinsically radiation-hard (unlike the silicon elements) and is very much cheaper, but has the drawback of a high occupancy. The tubes are filled with a mixture of xenon (70%), carbon dioxide (20%) and carbon tetrafluoride (10%). The TRT provides approximately 36 hits per track.

4.2.5 Calorimetry

The calorimetry measures the energy of electrons, photons, isolated hadrons, and perhaps most importantly, jets. It is also deployed for the measurement of missing energy and the detection of low-p_T muons. A diagram of the ATLAS calorimetry is given in figure 4.11. The overall pseudorapidity coverage is $|\eta| < 4.9$.

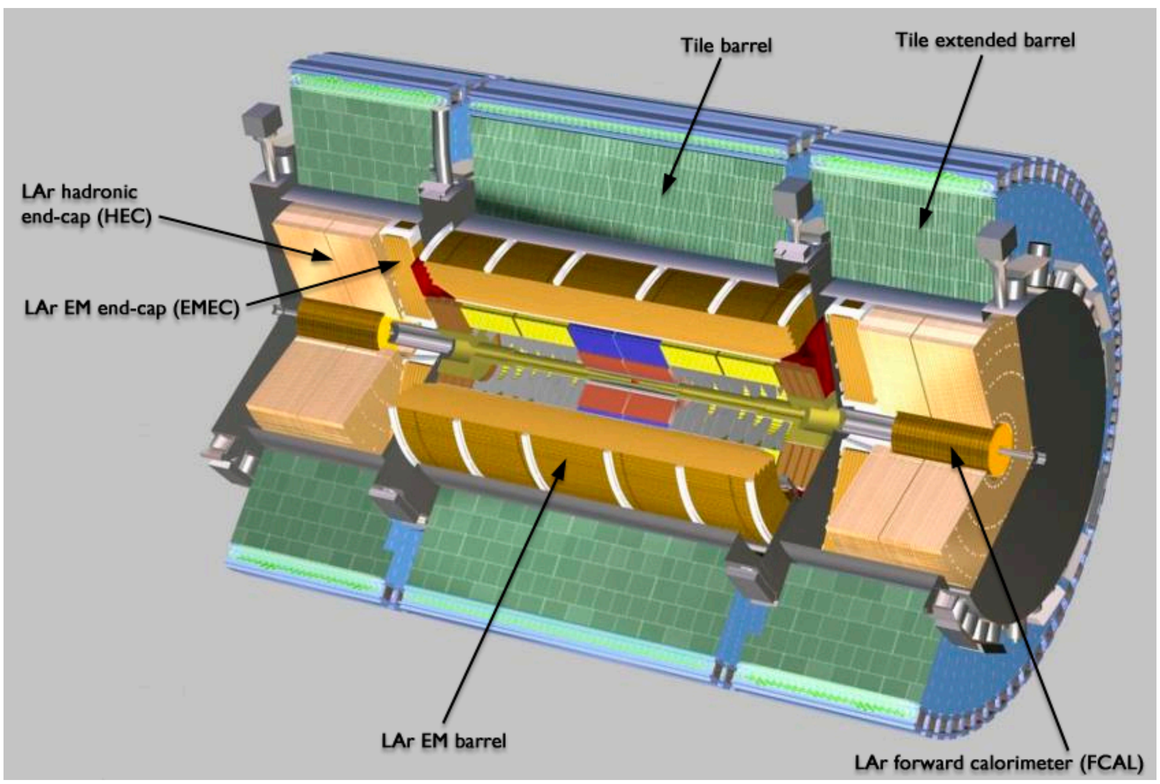


Figure 4.11: 3D schematic of the ATLAS calorimetry. Taken from [75].

There are two calorimeter systems - the **electromagnetic** calorimeter, for photons and electrons, and the **hadronic** calorimeter, for isolated hadrons and jets. Both systems are divided up into a cylindrical barrel and two end-caps for forward measurements. The calorimetry is located downstream of the inner detector and the central solenoid.

The *electromagnetic* calorimeter is a lead/liquid argon detector, with accordion-shaped electrodes and absorber plates. The barrel provides coverage in the pseudorapidity region $|\eta| < 1.475$. The length of each half-barrel is $3.2m$ and the inner and outer radii are $2.8m$ and $4m$ respectively [73]. It is in two pieces, with a $6mm$ gap at $z = 0$. The barrel electromagnetic calorimeter shares the same cryostat with the central solenoid. The end-cap E.M. calorimeters are in two discs, the inner disc covering $2.5 < |\eta| < 3.2$ and the outer wheel covering $1.375 < |\eta| < 2.5$.

The *hadronic* calorimeter uses the intrinsically radiation-hard liquid argon technology for higher pseudorapidities, where radiation levels will be higher, and plastic scintillator tiles embedded in iron absorbers for lower values ($|\eta| < 1.7$). The tile calorimeter is divided into one central barrel and two extended barrels. The inner and outer radii are $2.28m$ and $4.25m$ respectively. The scintillator tiles, $3mm$ in thickness, are arranged radially in a periodic fashion, with a total iron thickness of $14mm$ in one period. The emissions from each tile are read into two photomultipliers, one for each side, via fibre optics.

The end-caps cover the range $1.5 < |\eta| < 4.9$, where LAr calorimetry is used. They consist of inner and outer discs, the outer discs covering $|\eta|$ values below 3.2 and the inner discs dealing with values below 4.9. These inner discs are referred to as the *Forward Calorimeter*, and this must exist in a particularly challenging radiation environment.

Aside of the energy measurement rôle of the calorimetry, it should also act as a *filter* to remove all particles other than muons, such that readings from the muon spectrometer, located downstream of the hadronic calorimeter, are not polluted by hadrons. The hadronic calorimeter has a thickness of eleven radiation lengths at $\eta = 0$, which reduces punch-through of showers into the muon systems to a minimum.

4.2.6 The muon spectrometer

High p_T muons comprise the signatures of many interesting physics events, and the identification and measurement of the momentum of these particles is vital to the trigger system. These tasks fall to the **muon spectrometer**. Broadly speaking the muon system must perform two rôles - the precise measurement of muon tracks (and hence momenta) and the provision of additional information for the trigger systems, such as bunch crossing identification and rapid momentum cut-off measurements.

The muon system is located in the outermost parts of the detector, downstream of the hadronic calorimetry. It is the only detection system in ATLAS that will be visible from the outside of the detector. Momentum measurement is facilitated by the superconducting toroid magnets, where the barrel toroid provides field in the range $|\eta| < 1.0$ and the end-caps cover the range $1.4 < |\eta| < 2.7$. The region $1.0 < |\eta| < 1.4$ is covered by a combination of fields from both the barrel and the end-cap toroids.

In the barrel, the muon chambers are arranged in three cylindrical layers ("stations") whereas in the end-cap and transition region, the layers are positioned vertically. The technologies used in the detection elements differ depending on pseudorapidity and whether they are to be used for precise tracking or trigger measurements. Throughout most of the pseudorapidity range the detection elements used for tracking are **Monitored Drift Tubes** (MDTs). For high η and close to the interaction point, highly granular **Cathode Strip Chambers** (CSCs) are deployed. The trigger system, which covers the region $|\eta| < 2.4$, also uses two types of detector elements - **Resistive Plate Chambers** (RPCs) in the barrel and **Thin Gap Chambers** in the end-caps. The design and performance of these elements is covered in detail in [74].

Layout

Three images of the muon system are given in figures 4.12 - 4.14, showing a single quadrant of the muon system in the rz plane, a view of the whole system in the xy plane, and a 3D view. The system is designed to provide essentially full coverage, and particles crossing the detector from the interaction point should pass through three muon detector stations.

The barrel chambers are arranged in three concentric cylinders with radii of around $5m$, $7.5m$ and $10m$, and the end-caps are in four disks at distances of $7m$, $10m$, $14m$ and $21-23m$ from the interaction point. In the central $R\phi$ plane at $\eta = 0$ there is an opening to allow the passages of cables from and to the inner detector, the calorimetry and the central solenoid.

The chambers have been standardized to as great a degree as possible. In the barrel the chambers are rectangles with areas of between $2m^2$ and $10m^2$. In the end cap the chambers have a trapezoidal shape with tapering angles of $8.5°$ and $14°$ for the small and large chambers, respectively. The barrel and the nearer end-cap chambers are supported on the barrel toroid, whilst the outermost end-cap layers, which are separated from the rest of the detector, are mounted directly onto the cavern walls.

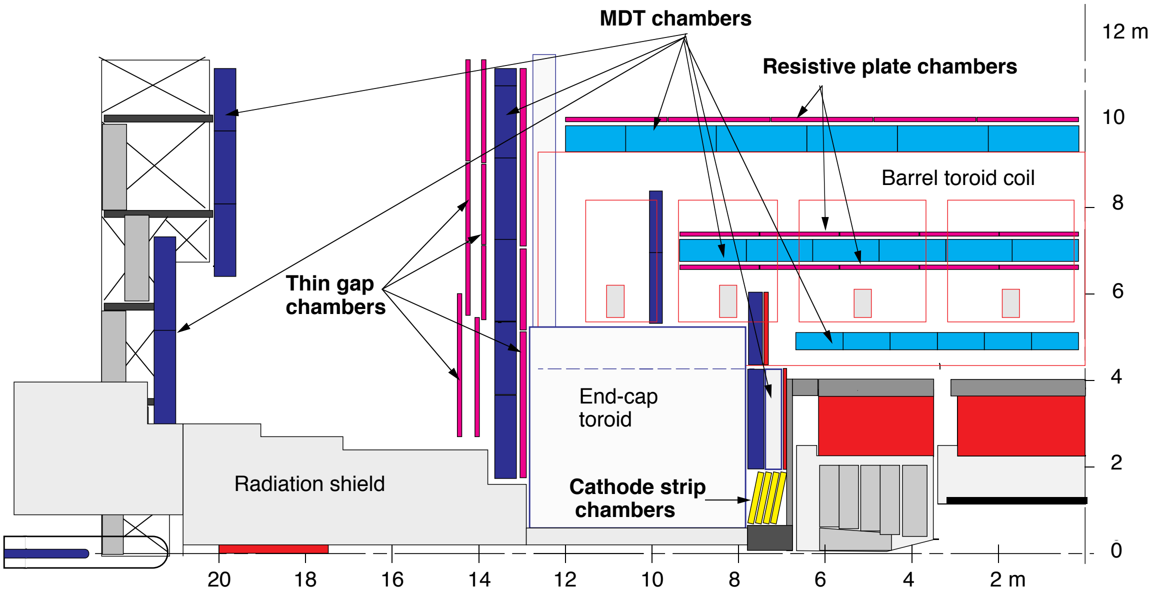


Figure 4.12: View of a single quadrant of the muon spectrometer in the rz projection. Taken from [74].

4.2.7 The data acquisition and trigger mechanism

The performance of the data acquisition (DAQ) and trigger mechanisms will ultimately determine the success of the detector in making useful physics measurements.

The requirements on these systems are severe. They must be able to cope with proton bunch crossing rates of up to $40MHz$, which poses enormous challenges in terms of identification of bunch crossings, and making trigger decisions on an acceptable time-scale. The equivalent event rate (at peak luminosity) is $1GHz$, whereas data storage resources permit data to be written to disk at a maximum rate of $200Hz$. This requires an overall rejection factor of 5×10^6, but excellent efficiency must be retained for "physics events".

A simplified flow diagram for the trigger system is given in figure 4.15. The system operates through a three level structure. Each level performs a more detailed scan on the events than the previous one, and only transfers events passing the criteria to the next level.

The data from all the detector subsystems is first written to the **pipeline memory**, at an event rate of approximately $1GHz$ at peak luminosity. The pipeline memory circuits are located close to or even bonded onto the detector element read-out circuitry, and are hence inaccessible and must operate in a high radiation environment. Each subsystem has a customized design of pipeline memory associated with it; there is no common design across the detector. The data from the calorimetry and the muon chambers is then scanned by the Level 1 (LVL1) trigger. This

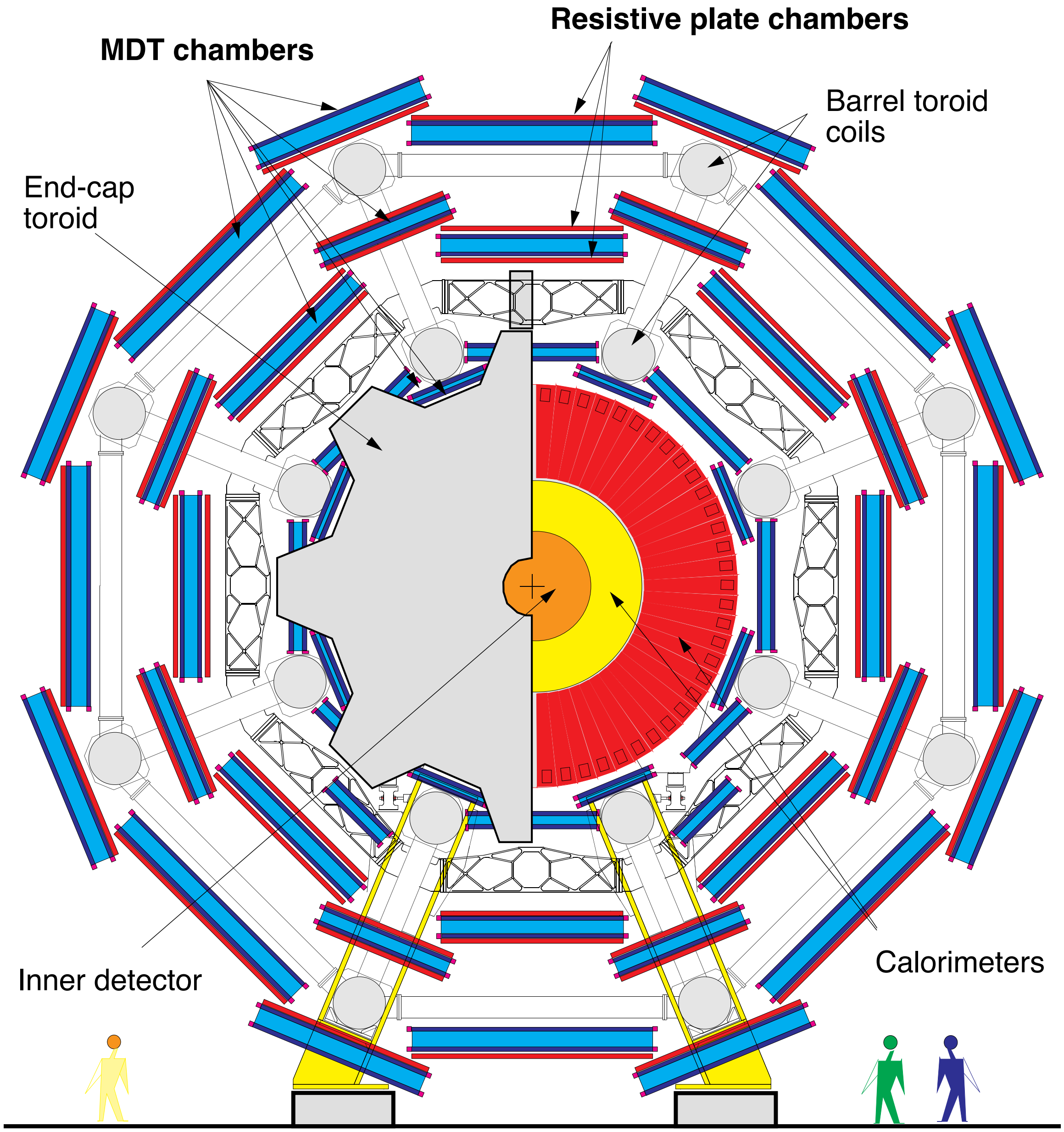


Figure 4.13: View of the muon spectrometer in the xy projection. Taken from [74].

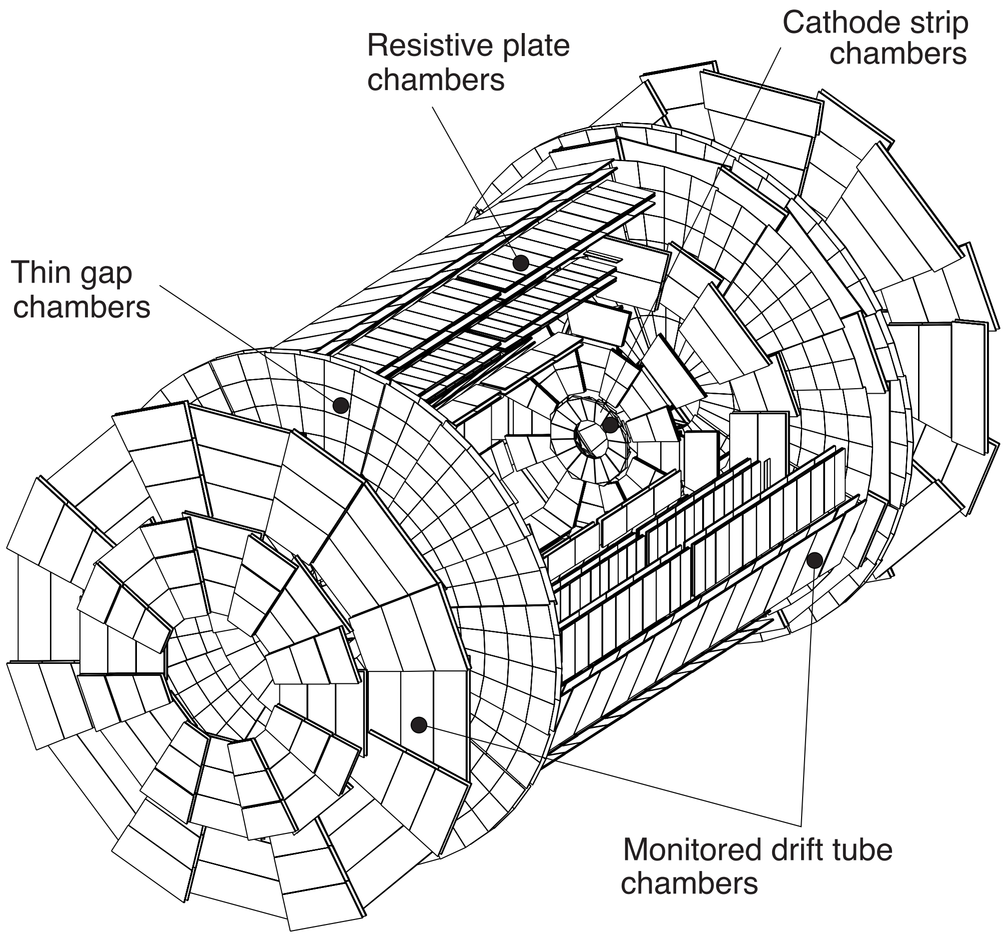


Figure 4.14: 3D view of the muon spectrometer. Taken from [74].

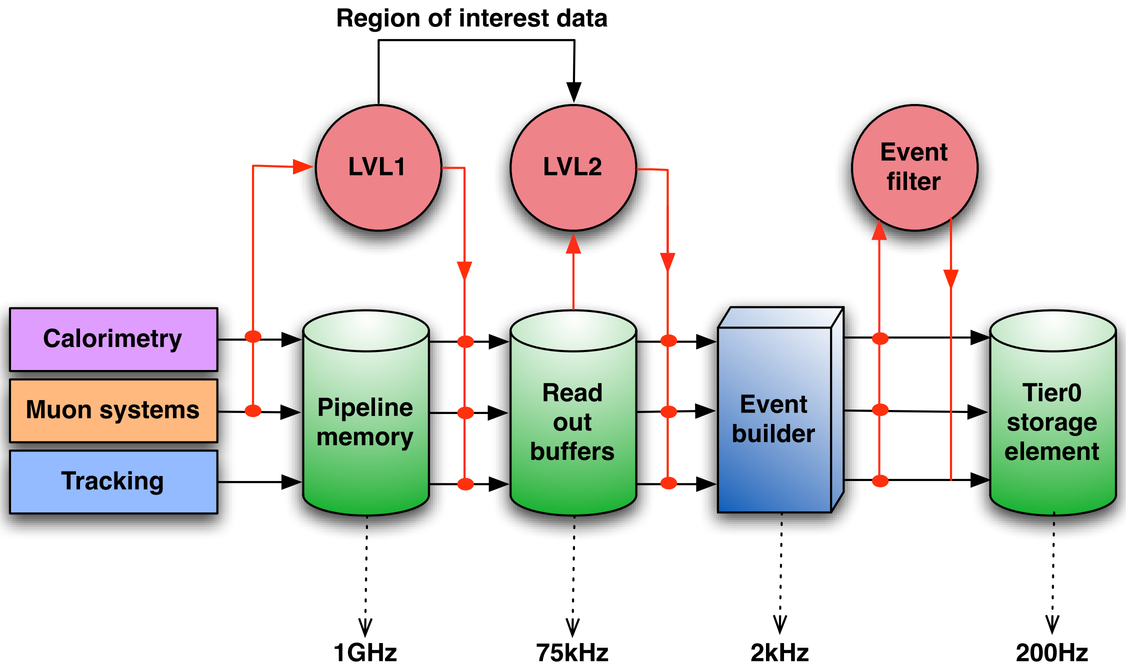


Figure 4.15: Simplified flow diagram for the ATLAS trigger

makes an initial selection based on reduced-granularity information from these two sub-detectors. High-p_T muons are identified through the use of the trigger muon chambers alone (the RPCs and the TGCs) without recourse to the precise muon tracking chambers. The full calorimeter is used, albeit at reduced granularity. The signatures which can activate the trigger are as follows:

- High p_T muons
- High p_T electrons, photons, jets
- Large missing transverse energies

The LVL1 trigger is implemented in custom electronics (FPGAs and ASICs [26]) which are located inside the detector. These will have programmable thresholds which can be modified according to the luminosity conditions and the physics requirements of the particular run. Events passing the LVL1 criteria are then written to the **read-out buffers** (ROBs) which are also located within the detector - it is foreseen that the detector will contain up to 1700 ROBs [117]. The LVL1 trigger is expected to write events at a rate of $75kHz$ [27], which represents a rejection factor of 13333.3. The LVL1 trigger must have as short a *latency* (the time required for the decision to be made and distributed) as possible - the target is $2.0\mu s$ with a $0.5\mu s$ contingency [117]. The LVL1 trigger must be able to uniquely identify a bunch crossing, which is a non-trivial task given the very short delay ($25ns$) between them - indeed the time taken for muons to cross from the interaction point to the muon trigger chambers is of the same order as the bunch crossing time, and the pulse shape of the calorimeter signals extends over many crossings [117]. The final rôle of the LVL1 trigger is the defining of a *Region of Interest* (RoI) for all events upon which it is activated. This is the geometrical region around the detector hits representing the process of interest. The RoI extends as a cone from the interaction point to the edge of the outermost muon systems.

The Level 2 Trigger (LVL2) scans the accepted events in the ROBs, reading data from all subsystems (including tracking) at full granularity. However, the scan is restricted to the region of interest, information on which is obtained from LVL1. In this way the LVL2 latency can be reduced by a very large margin. The LVL2 trigger is regarded as the lowest stratum of the *high level triggers*, and as such is implemented in software which will run on a processor farm close to, but separate from, the detector (in an adjacent cavern). Fast ethernet switches will be deployed to convey the data from ATLAS to the farm. LVL2 will first verify the LVL1 event, and will then make additional rejections through the sharpening (or raising) of cuts, the use of full granularity in the calorimetry, and matching muon and calorimeter hits with tracks in the inner detector. Implicit in this is that some limited inner detector track reconstruction will be carried out at LVL2. It is expected that this trigger will write data at a rate of $2kHz$, bringing the total rejection factor to 5×10^5. The latency time is expected to be in the range $1-10ms$.

Events accepted by the LVL2 trigger are passed to the Event Builder, which is again implemented in software run on the farm in the cavern. The event builder performs full reconstruction on the event, using the LVL2 data as a seed. This information is passed to the Event Filter (EF) which represents the final stage of the trigger system. This is capable of performing topological cuts, which can be varied depending on the physics requirements of the run. The event filter writes data at $200Hz$, representing the final rejection factor of 5×10^6. Data which passes this stage is then sent to the Tier Zero site (CERN) of the LHC distributed computing network, for further processing, distribution around the world and detailed analysis by physicists at their local institutes.

4.3 The ATLAS B-physics programme

The LHC will provide an unparalleled opportunity for the study of beauty processes. Current estimates suggest that around 1% of proton-proton collisions will result in the production of a $b\bar{b}$-quark pair. At low luminosity, this means a beauty event-rate of around $10^6 Hz$. Clearly this represents an enormous quantity of data, the exploitation of which occupies a significant part of the LHC scientific programme. However, this unprecedented B-event rate is a weakness as well as a strength. B-physics involves the precise measurement of specific decay channels, and such measurements favour a clean experimental environment - which is definitely not provided by the LHC. Accurate identification of interesting channels from the background, without introducing serious biases, is the principal issue for LHC B-physics.

One of the four LHC experiments, **LHCb**, is dedicated to B-physics. It resembles a fixed-target collision experiment, consisting of a single arm covering the forward region of the pp experiments with a pseudo-rapidity range of $2.1 < \eta < 5.3$ [[89]]. At the apex, closest to the interaction point, there is a vertex detector and a Ring Imaging Cherenchov (RICH) device, denoted RICH1. Beyond this is a tracker partly surrounded by a magnet, delivering a vertical field of 1.1T. Behind this is a second RICH detector (RICH2), followed by electromagnetic and hadronic calorimeters, and finally a muon spectrometer. The entire apparatus is some $20m$ long. Two features of this detector give it a particular capacity for B-physics measurements; firstly, the two RICH detectors will enable it to distinguish between pions and kaons, giving it a powerful tool for the identification of exclusive channels. Secondly, the luminosity of the LHC will be locally reduced at the LHCb interaction region by means of de-focusing magnets, such that, even at high luminosity, the detector will not experience a value greater than $2 \times 10^{32} cm^{-2} s^{-1}$. This will reduce pile-up to manageable levels, and will ameliorate radiation damage to the vertex detector, improving its performance in later years.

ATLAS, on the other hand, is an all-purpose detector, designed with "discovery physics" involving the explicit reconstruction of "new physics" states such as Higgs and SUSY particles, in mind. It is therefore intended that it will take measurements at the highest luminosities to maximize the search potential of the machine, with the associated problems of background. ATLAS has no RICH detectors, so hadron identification cannot be performed directly. Nevertheless, ATLAS will have a rich and varied B-physics programme. A summary is given in the remainder of this chapter.

4.3.1 B-physics triggers

The effectiveness of the B-physics triggers in removing the 99% of events which comprise the background, and in selecting particular channels of interest, will determine the success of the B-physics programme.

Throughout the high luminosity ($10^{34}cm^{-2}s^{-1}$) period and at the start of low luminosity ($2 \times 10^{33}cm^{-2}s^{-1}$) runs the B-physics measurements will essentially be limited to a di-muon trigger. This is based on the detection of muons at LVL1 with $p_T > 6GeV$ at high luminosity and, at low luminosity, $p_T > 5GeV$ in the barrel and $p_T > 3GeV$ in the end-caps [26]. These are then confirmed in the muon tracking chambers and extrapolated into the inner detector at LVL2. Topological cuts are applied in the event filter. Low p_T muons can also be identified in the tile calorimeter. The aim is to identify events containing specific exclusive decays which are indicative of the presence of B-events, such as $B \to J/\psi(\mu\mu)X$, $B \to \mu\mu X$ or the very rare $B \to \mu\mu$. The main background arises from the production of single muons from π or K mesons in flight, which have no association with a B-event. These are mostly removed at LVL2 by matching muon system hits with inner detector tracks [26].

Throughout the accelerator runs (which will typically last for twelve hours) the luminosity falls as protons are consumed in collisions. When the luminosity falls to about $10^{33}cm^{-2}s^{-1}$ the flexible nature of the trigger system allows additional trigger schemes to be "switched on". These use a combination of a single muon and at least one additional trigger from the semi-inclusive reconstruction of specific decay candidates [26], provided by additional electromagnetic or hadronic calorimeter cluster at LVL1. These are verified at LVL2 and extrapolated into the inner detector, allowing topological cuts to be made in the event filter. Such triggers are effective in identifying events containing hadronic decays such as $B_d^0 \to \pi\pi$ and $B_s^0 \to D_s\pi$. Two approaches can be utilised - the "RoI-guided" strategy, where full track reconstruction at LVL2 and EF only takes place in the RoI (as defined at LVL1), and the "full-scan" strategy which dispenses with the RoI mechanism and performs reconstruction across the whole acceptance of the Inner Detector. The latter approach is clearly more efficient in terms of identification of signal events, but budgetary constraints on the online computing resources are unlikely to allow such a method.

Table 4.1 shows the trigger rates for various important B-physics channels. The left-hand columns assume only the di-muon trigger, whilst those on the right, being for lower luminosities, include the full suite of trigger strategies. These figures are only a fraction of the total output of the ATLAS triggers - B-physics channels represent a small part of the total ATLAS programme.

4.3.2 B-physics channels in ATLAS

The processes which are currently being assessed for potential measurement in ATLAS [28] are summarized in table 4.2. The channels of interest fall into four broad

Table 4.1: Estimated B-physics trigger rates [26]

	$2 \times 10^{33} cm^{-2} s^{-1}$		$10^{33} cm^{-2} s^{-1}$	
Trigger	LVL2	EF	LVL2	EF
$B_{d,s} \to \mu\mu X$	$200 Hz$	small	$100 Hz$	small
$J/\psi \to \mu\mu$	$200 Hz$	$10 Hz$	$100 Hz$	$5 Hz$
$D_s^+ \to \phi\pi^{\pm}$	-	-	$60 Hz$	$9 Hz$
$B \to \pi^+\pi^-$	-	-	$20 Hz$	$3 Hz$
$J/\psi \to ee$	-	-	$10 Hz$	$2 Hz$

Table 4.2: The ATLAS B-physics programme

Event Type	Channel	Physics Motivation
Decays to Charm	$B_d \to J/\psi K_s^0$	$\sin 2\beta$, CP violation (CPV) from new physics
	$B_s \to J/\psi\phi$, $B_s \to J/\psi\eta$	$\Delta\Gamma_s$,ϕ_s, CPV from new physics
	$\Lambda_b \to \Lambda^0 J/\psi(\mu\mu)$	QCD, production polarization, CPV
	$B_c^+ \to J/\psi(\mu\mu)\pi^+$	QCD, heavy-heavy decays, V_{cb}
Rare decays	$B_s \to \mu\mu$, $B_d \to \mu\mu$	Branching ratio (BR)
	$B_d \to K^{0*}\mu\mu$, $B_s \to \phi\mu\mu$	BR, forward-backward asymmetry (A_{FB})
	$B_d \to K^{0*}\gamma$, $B_s \to \phi\gamma$	BR, CPV
	$\Lambda_b \to \Lambda_0\mu\mu$	BR, A_{FB}
Hadronic channels	$B_s \to D_s\pi$, $B_s \to D_s a_1$	B_s oscillations, ΔM_s, $\frac{V_{ts}}{V_{td}}$
	$B_d \to \pi\pi$	CPV
	$B^+ \to K^+K^+p^-$	BR
b-production	$\bar{b} \to B_d \to J/\psi(\mu\mu)K^0$, $b \to \mu X$	$\sigma_{production}$, $b\bar{b}$ correlations
	$\bar{b} \to B_s \to J/\psi(\mu\mu)\phi$, $b \to \mu X$	$\sigma_{production}$, $b\bar{b}$ correlations

categories:

- *Decays to charm*, involving the production of a J/ψ meson. These are an attractive set of channels, as the J/ψs decay to a pair of muons, or a pair of electrons, with branching ratios of $\sim 6\%$ each. This has clear implications for the trigger, especially in the case of the $J/\psi \to \mu\mu$ channel, which is self-triggering under the di-muon scheme. Given that the di-muon strategy will be maintained into the full luminosity period, the period over which such measurements can be made is extended significantly. Furthermore, these channels encompass a rich variety of physics processes, including the measurement of unitarity triangle angles ($\sin 2\beta$) and the detection of additional CP-violation induced by processes not described in the Standard Model. B_s mesons decay in significant numbers to J/ψs, so the mixing processes associated with this system, which are of great interest due to their susceptibility to new physics, can in principle be accessed through the di-muon trigger. The same applies to Λ_b baryons and B_c mesons.

- *Rare decays*, involving decays with extremely low branching ratios. The very high $b\bar{b}$ pair production rate of the LHC means that, at high luminosity, these decays may appear in sufficient numbers for ATLAS to make statistically significant measurements. Particularly promising are those channels where the B-hadron decays directly to two muons, as these can utilize the di-muon trigger which will operate into the high luminosity period.

- *Hadronic channels*, where the B-particle decays into two hadrons. These channels may be able to facilitate a rich selection of measurements, but triggering on such events is far more challenging - a calorimeter cluster must be used along with a single muon, and this requires the luminosity to be lower than the peak value during the "low luminosity" runs. Consequently these can only be activated later on in a run when the luminosity has diminished. Study of these channels during the high luminosity phase is out of the question.

- *b-production studies*, which use a combination of the other channels to study production cross sections and $b\bar{b}$ correlations.

Chapter 5

The ATLAS experiment: computing and software

5.1 Introduction

The software and computing infrastructure dimension to the LHC project is on an equivalent scale to its civil engineering. The sheer volume and complexity of the data, the rate at which it will emerge from the detectors, and the number and geographical dispersion of collaborators who will require access to it, pose unprecedented challenges to those involved in the development of LHC computing.

In general terms, the computing and software infrastructure of the LHC and the experiments is required to perform the following roles.

1. Monitoring and control of the accelerator and the detectors
2. Implementation of the trigger algorithms and acquisition of raw data from the detectors
3. Construction of data artefacts (event summary data) such as tracks, vertices and calorimeter clusters from the raw data
4. Recording of all data to permanent storage
5. Construction of smaller data entities, suitable for physics analysis, from the event summary data
6. Facilitation of the transmission of the data to individuals and institutes around the world
7. Provision of a set of tools for performing physics analysis

8. Provision of facilities for the simulation of event production, for performance studies

9. Provision of a single framework in which all the software operations can be executed

The first three items on the above list constitute the functions of the *online* computing; that is to say, tasks performed during the taking of data. The remaining tasks are all performed by *offline* software and computer infrastructure. This chapter will primarily be concerned with the offline computing.

The LHC will produce data at a prodigious rate. During every year of running other than 2007-2008, it is estimated that events will pass the ATLAS triggers at a rate of $200Hz$ [29]. Assuming 50000 seconds per day and 200 days of data-taking per year, the annual number of events recorded by ATLAS will be approximately 2×10^9. The estimated total size of an event, including raw and processed data, is around $2MB$ [29]. In consequence it is thought that ATLAS will produce some $4PB$ per year for committal to permanent storage. The four experiments together will write around $10PB$ annually. Storing and processing these colossal quantities of data at a single site would require unreasonable quantities of floor space, cabling, cooling, networking facilities and human resources.

In consequence CERN and the four experiments, along with numerous national bodies, have over the last few years developed the Large Hadron Collider Computing Grid (LCG) or "the Grid" as it is colloquially known. This is a world-wide network of sites offering processors and storage equipment for use by the LHC community. At the time of writing the Grid consists of 37171 processors and 19431 TB of storage at 190 sites in 46 countries [33], and it is anticipated that this number will have increased significantly by the time the LHC is switched on in 2007.

The design of the offline software structures differs from experiment to experiment. ATLAS uses an object-oriented framework known as **Athena**[1] which is itself an implementation of the **Gaudi**[2] [30] framework, used also (and originally developed by) the LHCb experiment. All of the offline operations take place within this framework, whether the data is simulated, from the test beams or from the experiment itself. Athena itself is written in C++ and is driven by Python scripts, and the bulk of the plug-in code is also in this language. Some external elements are also composed in FORTRAN, and in Java.

[1]The goddess of wisdom, war, the Arts, industry and skill

[2]Antoni Gaudi, Barcelona architect 1852 - 1926, famed for the design of Barcelona's La Sagrada Familia, an immense Basilica which has been under construction since the beginning of the 20th century and is not scheduled for completion until 2026

The structure of the offline and online software, and the means by which ATLAS will exploit distributed computing, are summarized in the ATLAS *Computing Model* [29]. This presents the anticipated state of the ATLAS computing infrastructure in 2007, and provides the programme of development which will enable this to be achieved.

This chapter will follow two broad themes. Given that all of the simulations presented in this work were performed on the Grid, the first part will give a brief exposition of this system. The lengthier second section will cover the various components of the Athena framework, including the tools used for simulating data such as event generators and detector simulation, and the software for reconstructing and analyzing real or simulated events. The author has been involved in the development of a number of software components and tools; this will be indicated clearly where appropriate.

5.2 The Grid

5.2.1 What is a grid?

"The Grid" is used as a colloquial term in the CERN community for the LHC Computing Grid, or LCG. In fact, unlike "The Internet" there is no single worldwide computing Grid; many such networks exist, although the LCG happens to be the largest at the time of writing [32]. It appears that computer scientists have some difficulty in providing an exact definition for grid [35], although the following characteristics are common to any computer network described as a grid.

1. Grids are geographically *distributed*. Whilst the "Grid" trademark is often applied to modern batch processing systems (e.g. [36]) it would seem that true grids are dispersed across several or many locations.

2. An element of *processing* is involved. This is in contrast to the world wide web, which is generally used to share static information rather than processing power. The pooling of processors is a key practical incentive for the building of grids, and the commercial potential for protocols allowing large organizations to sell idle CPU time to other companies, has certainly not been overlooked by the I.T. industry.

3. A grid should appear to the end-user to be a single computer [35]. In the same way that a user of the World Wide Web with no prior experience of information technology could reasonably assume that all the world's web sites

were stored at the local telephone exchange, a grid user should effectively see one portal of entry, and have no need to connect to the separate sites. The underlying software might well be performing some complex operation to find suitable resources across the world, but this should be invisible to the user.

4. Grids must be able to coordinate decentralized resources [35]. Even though the machines and data on a grid might be owned and administered by a myriad organizations, it must be able to manage access to them and deal with security and related matters.

5. The protocols used by a grid must be accepted by all component sites [35].

6. Grids should combine heterogeneous resources.

It should be pointed out that the processing of particle physics data is particularly well suited to the batch approach offered by a grid. In general the data is in the form of physics *events*, each event or group of events being independent of the previous or subsequent one. One proton-proton collision leading to a signal in the detector does not affect an event occuring two bunch crossings later, so the events can be dealt with as independent packages. This is in marked contrast to applications such as climate modelling, weather forecasting, finite element engineering or the simulation of nuclear weaponry. In such applications, the physical system is divided into cells, and the results of the calculations at the boundary of one cell feeds into its neighbour. In these cases the best approach is to use a true supercomputer, with parallel processors handling individual cells. Grids are less well suited to this type of task.

The components of a grid

All grids vary in composition, but some general comments can be made [35].

The grid **fabric** is a broad term covering the processors, storage facilities and the data stored on them, user interfaces, operating systems installed on the sites, local queueing systems and internet protocols.

Grid middleware is the software and hardware which binds the component sites into a single grid. This includes security systems to control access to the resources, services for providing information on available resources and data management tools such as file catalogues.

Grid tools allow applications to interface with the grid fabric. These include the **resource brokers** which decide, on the basis of information provided by middleware, where a given job should run. Grid activity monitors also form part of this class.

Applications are the programs that will run on the Grid. [35] implies that these applications will in general be written with grid use specifically in mind, and will be developed with grid developer tools. This is not in fact the author's personal experience of the LCG; in this case the software is identical to that found on the local CERN cluster, and is simply installed on every site. Where the software is not installed, the job cannot run. Naturally the LCG is still under construction so any view of its operation at this stage cannot be treated as indicative of the final state of the system.

5.2.2 The LCG

LCG Structure

The LCG sites are organized into three *tiers*, numbered from zero to two. There is only one *Tier Zero* site - CERN. Once the experiments are running, it will be this site which receives the raw data from the trigger systems, and performs full reconstruction to produce the event summary and analysis-ready data objects. It will store the raw and processed data. Only those involved with data processing and distribution will have access to the Tier Zero site.

There will be approximately ten *Tier One* sites serving ATLAS. These will be connected to the Tier Zero by fast ($10Gb/s$) network connections via routers based in the host countries. The main role of the Tier One sites is first to act as reserve storage for the raw data, and second to provide the means of distributing the reconstructed data objects to the physics collaborators. Again, access to the Tier One sites will be restricted to those involved with central production.

Tier Two encompass all other LCG sites. It is to these sites (or clusters of sites) that physicists will be connected when they perform analysis algorithms on reconstructed data. The storage at these sites will be modest - only the light-weight analysis-ready data needed for performing physics analysis will be kept here. The data stored at a given site will be in general a function of the physics interests of the local users. Tier Two sites will also perform an additional essential role - the production of Monte Carlo data.

Components of the LCG

Thus far the LCG project has not brought any completely novel software products into being, but has brought existing tools together [35]. In particular, the LCG

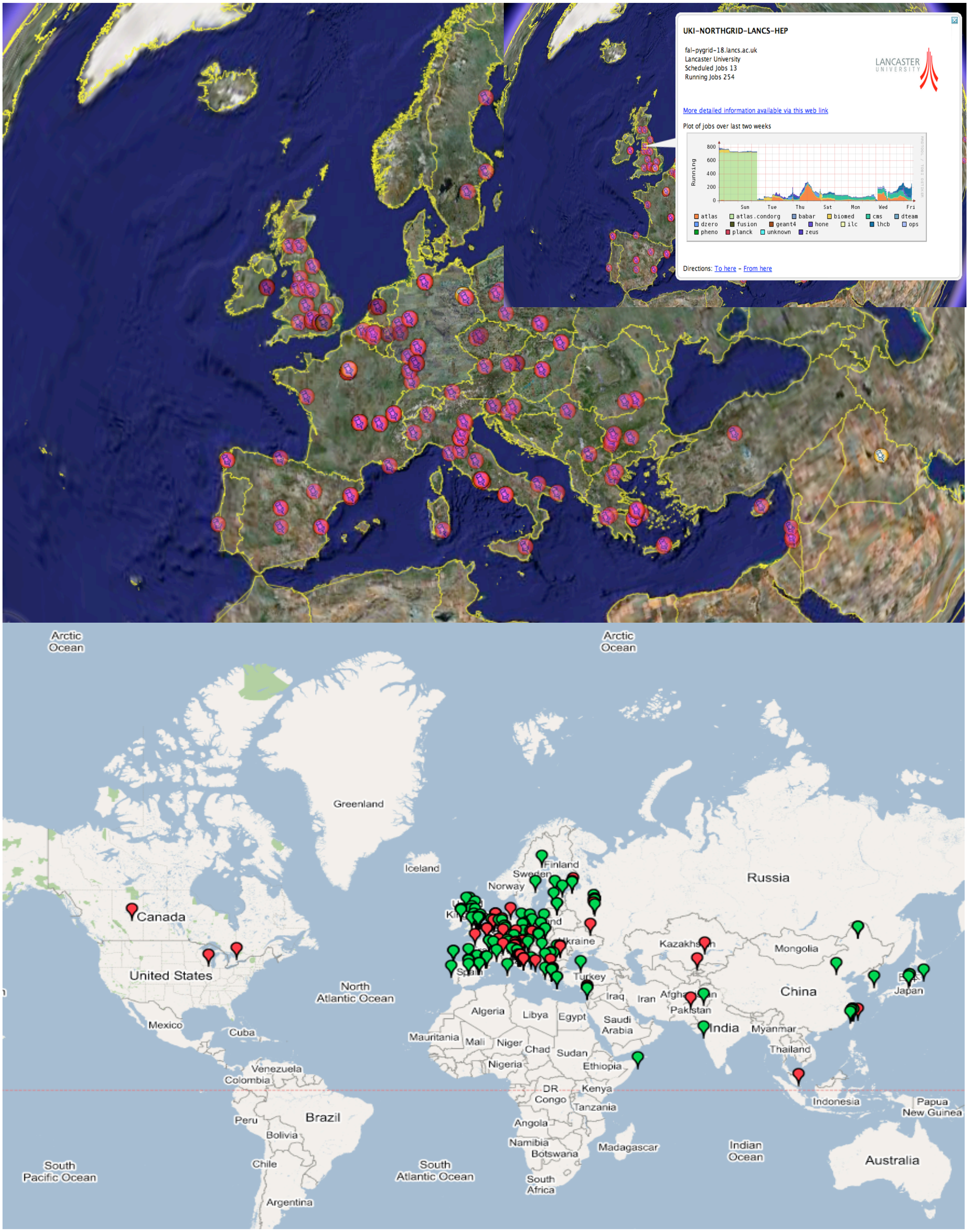


Figure 5.1: A variety of real-time LCG monitoring tools, interfaced with Google Earth (upper) and Google Maps (lower). The instantaneous status of the Lancaster site is expanded in the upper inset. Produced with tools available at [34]

middleware uses the Globus Toolkit [38] and Condor-G [40] along with components developed by the European DataGrid project [41].

The **Globus Toolkit** consists of software for resource monitoring and management, security and file handling. It is developed by the Globus Alliance [38]. Using these packages grid developers can construct the elements which together make up a grid.

Condor [39] is a project developing mechanisms to utilize spare CPU time in clusters of machines. **Condor-G** [40] combines features of both Globus and Condor to provide a complete grid resource management system [35].

The **European DataGrid** is an E.U. project developing middleware for distributed data analysis and storage, specifically for research purposes. Much of the LCG middleware is based upon EDG products. A number of terms which are now widely known by those using the LCG emerged from the EDG project, in particular:

- **Computing Element** - the interface between the processors in the grid fabric and the middleware, which allows the latter to gain knowledge of suitable processing resources and dispatch jobs accordingly.

- **Workload management system** - middleware which matches job requirements to available resources, submits and monitors jobs and maintains a database of running jobs. The tasks performed by the WMS are non-trivial, but the user sees none of this complication. He or she constructs a file written in **job description language** (JDL), enumerating the job requirements and the location of the executable file or script, and submits it to the WMS with a simple command. The **Resource Broker** (RB) then matches the job to a facility according to the instructions in the JDL. The job is then managed by Condor-G [35].

- **Storage Element** - the interface between the tapes and disks in the fabric, and the middleware.

- **Replica Management** - systems by which **replicas** of files are managed. It is a key requirement of the LCG that data will be replicated to multiple locations around the world, to enable faster access. This poses challenges of identification and cataloguing, however, and these issues are handled by the Replica Management Service (RMS) which itself commands a number of other services. Data is uniquely identified by a **Globally Unique Identifier** (GUID) which is a string of numbers and letters. The catalogues provided by these services allow more intuitive names to be attached to the files - the

Logical File Name (LFN). A single GUID may have many LFNs associated with it corresponding to replicas of the file at multiple locations.

Finally, two additional terms which have not yet been covered must be introduced. A **User Interface** (UI) is a computer which is able to communicate with the Grid middleware. It is from here that a user establishes a security proxy, interrogates Grid data management, and submits jobs. Any networked Linux machine can be a UI, given a successful installation of the UI software. In addition all of the LXPLUS machines at CERN are User Interfaces.

A **Virtual Organization** (VO) is a "dynamic collection of individuals, institutions and resources" [35][42] who use the Grid to perform their operations. In high energy physics the virtual organizations tend to be arranged by experiment (so there is one for ATLAS, another for LHCb etc). Each VO has a certain allocation of resources, and any given site providing facilities to a VO must be configured such that the VO's software can run. The same resources will often be serving several VOs, and ensuring this interoperability is a significant challenge in the development of grid technologies [35].

Current operation of the LCG: View of a physics user

Most of the people currently using the LCG are involved in its development, preparing for data taking in 2007. It has been used for some large scale physics productions, most notably for the Rome Physics Workshop which took place at Roma III University in June 2005 [29]. Most of this effort (some 500 000 job submissions leading to 6.1 million fully simulated events) was performed using the dedicated production system ProdSys [29]. This work was performed by experts who worked full-time on the production, and provided an excellent opportunity to test the LCG structures.

However, it is also possible, even at this stage in the development of the LCG, for non-experts to use the Grid for producing quantities of data which would not be practical at a single institute or on the CERN batch system. It was observed above that the user, having obtained the requisite security permissions, needs to do little more than construct an appropriate JDL file, and provide an executable script. The JDL files are written without much difficulty, but there are a number of drawbacks to working in this way.

Firstly, "providing an executable script" is often a non-trivial task. In general, most users and developers of the ATLAS software work on the LXPLUS cluster at CERN. As a result much of the documentation assumes that the user will be working in this environment, and will have access to the Andrews File System (AFS) repository of software and scripts. This is not the case when working on the Grid.

Each site will have the ATLAS software installed on it, but will not be connected to AFS. Developers should ensure that the code is written in a completely general way so that paths and environment variables transfer seamlessly onto the Grid, but this is not always the case - AFS paths can sometimes be seen in code and scripts. Such code will not run on the Grid. Setting up environment variables is similarly non-trivial.

Secondly, writing files to Grid storage is a complex task. Whilst the Grid middleware returns the standard output and error streams of a job, and any other small files of the user's choosing, to a directory on the user interface machine (the "output sandbox"), large files containing physics data must be stored on the Grid itself, so that they are available to other users. This would generally be beyond the skills of a non-Grid expert.

Thirdly, the directory on the UI (on the temporary space) into which the output sandbox is dropped will have a machine-generated name. Each sandbox from each job will go to a different place. Collecting these files would therefore be a tedious task if more than a few jobs were submitted (and after all, most of the motivation for a physicist using the Grid at this stage is the ability to submit large quantities of jobs). Noting that each job requires a JDL file, it can also be seen that writing these files will quickly become a tedious task.

Finally, keeping track of the running, completed and failed jobs can become overwhelming once there are more than a few jobs running. No method of book-keeping is provided.

In answer to these problems, a number of products have been provided by several groups to simplify and automate the setting up and execution of ATLAS software tasks on the Grid. The author has participated in the testing and documentation of some of these tools, used them to produce all of the data for this work, and led the "sub-contracted" production work given to the B-physics group for the Rome workshop, again using these tools. A brief description of these will now be given.

Job Transformations

Job transformations [43] are designed to automatically prepare a given machine session for running the ATLAS software. They consist of a tarred set of scripts and job option files which, when executed on a machine with the correct release of the software installed on it, immediately run the ATLAS software according to the specifications of the user. Although they can be used in a non-distributed context (for instance on a machine at a local institute which is not connected to AFS) they are particularly suited to the Grid. The user provides as the executable a simple script which runs the appropriate job transformation. The job transformation itself is sent to the Grid middleware along with this executable. Additional mechanisms allow the user to patch his or her own code into the transformations, and a skeleton script is provided so that, in the event that a user needs to perform a task which is not provided for in the existing code, he or she can define their own without too much difficulty. The job transformations follow the main ATLAS software releases, and are obtained from the Roma I repository.

Don Quijote

Don Quijote[3] (DQ) [44] is a package written with the intention of vastly simplifying the process of writing files to the Grid, inspecting the content of storage facilities, and copying files to a local disk. The software is written in Python and hides the many "incantations" that these tasks involve; the user then has only to type simple and intuitive commands.

Light Job Submission Framework

The Light Job Submission Framework (LJSF) [45] is a set of Python and shell scripts which allow a user to submit multiple jobs to the Grid in an efficient and rapid manner. It uses transformations to define the execution of the job and DQ to write results to the Grid storage facilities. In addition it has the facility to write multiple sets of JDL files for each stage of the simulation chain, and through a consistent file naming scheme is able to provide output files from a given stage in the chain as input for the next stage. The LJSF scripts allow a user to check the status of a job without having to type long job identifiers, and automatically collects sandbox output and places it into a common directory. The output is checked to ensure certain flags, indicative of successful execution, are present, and this information then defines the validation status of the job. The LJSF also has a dedicated MySQL database (hosted at INFN Roma I) attached to it, which registers the job name, the validation status and the names of any files written to the Grid. This information can be accessed via the command line or from a convenient web interface.

The LJSF is not intended to be a large-scale production system; it cannot, for instance, automatically re-submit failed jobs as the production system can. However it certainly allows individual physicists to make full use of the vast computing resources offered by the Grid. The ATLAS B-physics group, for instance, produced some 1.5 million fully simulated events for the Rome workshop using the LJSF, the

[3]Book by Miguel de Cervantes Saavedra (1547-1614) concerning the adventures of an itinerant (and eccentric) Spanish nobleman; it is from this book that the term *quixotic* orginates

bulk of which were performed by two individuals. All of simulations presented in this work were produced on the LCG using the LJSF as the submission tool.

5.3 The ATLAS Offline Software

The remainder of this chapter deals with the ATLAS offline software. When data-taking begins this software will handle all stages of the data processing, bar the collection from the detector and the systems monitoring. At the time of writing it is used for test beam analyses, simulations to enable feasibility studies (such as the one presented in this work and at the Rome physics workshop) and for the "data challenges" where large quantities of data are produced, moved from institute to institute and analyzed, mainly to test the robustness of the Grid.

After considering briefly the overarching framework (Athena), the simulation chain and the types of data in use, each of the plug-in physics components will be described in detail, with particular attention paid to those used to produce the results presented in this work.

5.3.1 The Athena Framework

The Athena framework [31] is an object-oriented (C++) framework which provides the wherewithal for physics data-processing applications to be run in a sequential manner. It provides a range of services and tools which facilitate the storage and transmission of data and the management of the applications and the communication of user-defined quantities to the software. The physics applications themselves "plug in" to the framework and are executed sequentially.

The design of the framework has been guided by a number of principles and requirements, namely [29]

- The data and algorithms should be separated. Classes which contain physics data (such as Tracks, Vertices) are in general independent of the code (algorithms) which performed the calculations leading to the formation of the data. The classes defining tracks, for instance, are wholly separated from the track finding algorithms. The advantages of this approach are clear - objects containing physics data are kept as simple as possible, and in the event that two or more algorithms exist for producing a given type of data (e.g. different tracking routines, different vertexing methods), switching between them is trivial - no alterations need to be made to "downstream" classes, as the data and algorithms are independent.

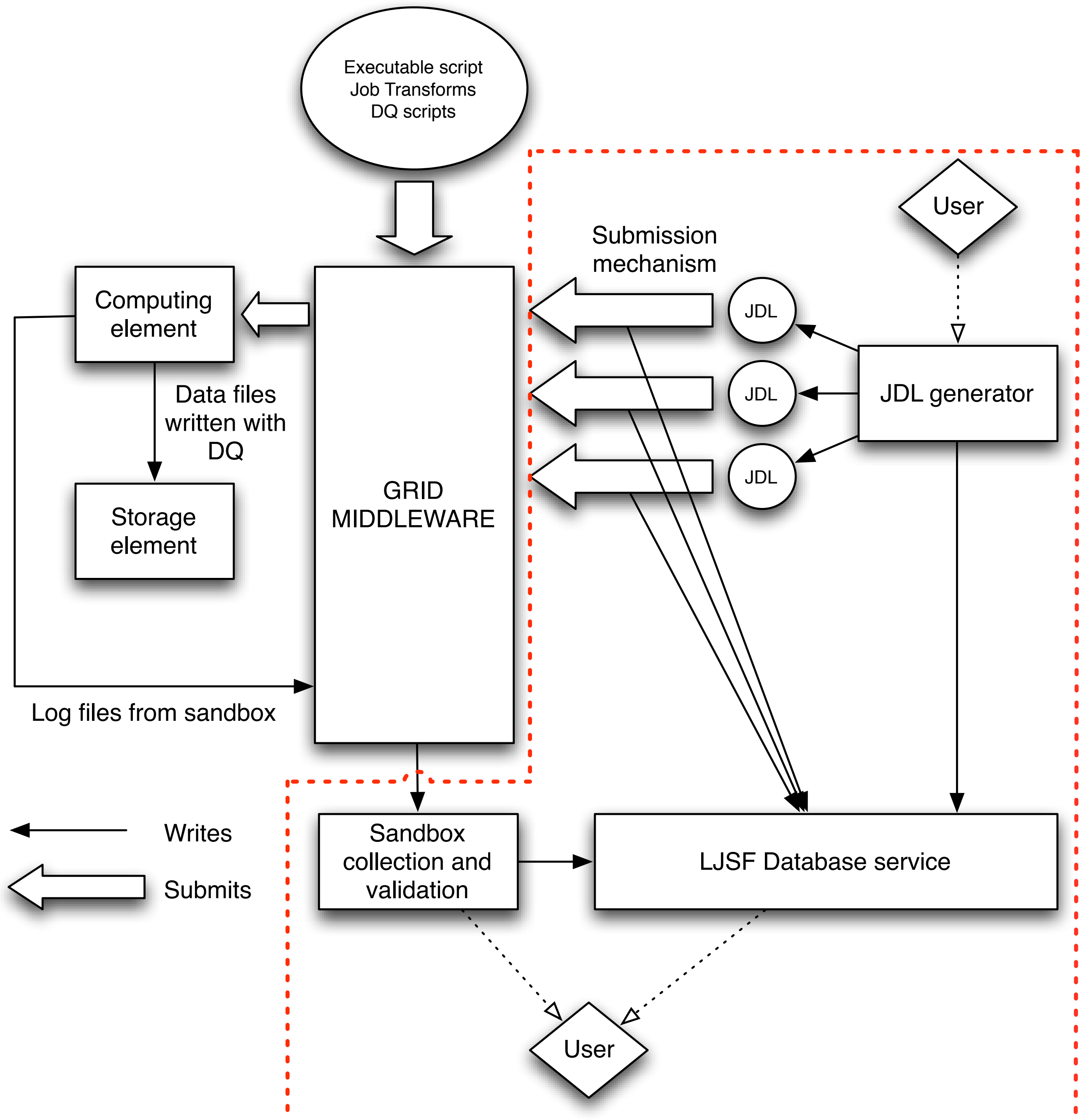


Figure 5.2: Data flow in the LJSF; red dotted line contains LJSF mechanism

- There are many types of data, having different roles and lifetimes, within the data processing chain.

- There should be a clear distinction between *transient* data, which resides in the system memory during the data processing, and *persistent* data, which is committed to disk. It is likely that the latter will evolve over the lifetime of the experiment, so changes in persistent storage must not effect the algorithmic code which deals with transient data.

- Physicists and physicist-programmers should be shielded from the full complexity of data storage and construction of data objects, by a range of simple and robust interfaces.

The important components of Athena are as follows [29]:

- **Application Manager**. This is the master class which drives and co-ordinates the entire data processing scheme. There is one instance of the Application Manager; it is a *global* class.

- **Algorithms**. These are the basic data processing class. All algorithms inherit from an Algorithm base class, which contains three methods: `initialize`, which runs once at the start of a job and is responsible for instantiating tools, services and setting up histograms and n-tuples; `execute`, which runs once per event and does the data processing; and `finalize` which runs once at the end and closes down the applications gracefully. Algorithms are often tied together in a chain or **sequence** with data passed from one to the next via the transient store. The user determines the order in which the algorithms execute. In this context **filter** algorithms are particularly important, in that events failing a given criteria can be removed from the data stream.

- **Tools**. Tools are similar to algorithms in that they process data from and write results to the transient store, but do not inherit from a common base class. They can be executed multiple times per event, in contrast to algorithms which execute once only in each event.

- **Services**. These components are used for specific purposes in an algorithm, and are generally instantiated in the `initialize` method. Examples include:

  - **Job Options Service**, which enables a user to pass requirements to Athena at run-time, through a job options file. This is written in the Python scripting language, and is interpreted by a Python driving script, which in turn activates the Athena libraries.

- **Messaging Service**, which provides a means for a user to print messages to the screen. Although this can be done using standard C++ functions in the `std` namespace, the messaging service automatically appends the name of the algorithm in question to the message, which is extremely useful if many algorithms are executing sequentially. In addition the user can increase or reduce the verbosity of the messaging service, to either permit only important or critical messages, or to let informative or debugging messages through as well.
- **Performance monitoring**, which gives information on memory and CPU usage.
- **Histogramming and n-tuple services**, which record calculated quantities over many events, allowing the presentation of analysis results.
- **Random number services**, which provide random numbers for Monte Carlo engines. The seeds are set via the job options.
- **StoreGate service**, which is the name given the the transient store. This acts as the "blackboard" onto which data produced by algorithms, tools or read out of persistency is written, and from which such data is read. Data objects are first placed into a Standard Template Library container before being written to StoreGate; the service can hold any object which is of the STL **Assignable** type [29].
- **Persistency service**, which transfers data objects from StoreGate to disk at the end of a job, and vice-versa. The format used is referred to as POOL.

Package and version management

The ATLAS software is rapidly evolving at the time of writing, as simulation software is corrected and tuned and preparations are made for data taking. Managing this evolution is therefore crucially important.

Athena algorithms and tools are divided into **packages**. The contents of a single package may be responsible for handling or processing a certain type of data object (such as tracks or vertices) or may define those classes (and in general these will be in separate packages, as noted above). Other packages may contain algorithms for performing certain types of task, for instance, high-level physics analysis, triggering simulation and so on.

Athena evolves in **releases** which are described by a string of three integers, $x.y.z$. x refers to the *major* release number, whilst z is incremented when small changes are made to fix bugs. y is incremented to allow the inclusion of experimental code which could have significant (and perhaps unwanted) effects, so in general release numbers of the form $x.0.z$ will be used for physics studies. Major releases appear on a time-scale of months, and each new major release will have significant additional functionality. The results presented in this work were produced with releases 10.0.4 and 11.0.41. It is envisaged that data will be taken under release 15. Between minor releases, small changes which are due to be implemented in the next release are tested by means of **nightly builds** which as the name suggests take place on a nightly basis.

Attached to every package is a tag - this six digit number (grouped into pairs) is incremented every time a package is modified as part of a change to the release. If a package remains unchanged when the release is updated, the tag number remains constant, so stable packages (such as those concerned with event generation) will keep the same number for several or many releases, whereas those packages undergoing rapid evolution, such as reconstruction and detector simulation, increment with almost every release.

The packages are stored in a central repository on the Andrews file-system (AFS) from where they can be copied (or "checked out") using a **configuration management tool** (CMT) [46]. The releases are also shipped around the world to Grid sites by means of **distribution kits** [29].

5.3.2 Event Generation Software

For both feasibility studies during the preparatory stage of an experiment, and also during the data taking, the production of large sets of simulated events is necessary. In the case of the LHC these simulations require the modelling of the initial proton-proton collision, the production of elementary quarks, gluons and gauge bosons from the collision, the subsequent hadronization, and the decay of these hadrons into longer-lived particles which can be observed by the detector. The output data of such simulations takes a very simple form - a list of particles, their momentum and energy, their parent particle and the daughters into which they decay, and the positions of their production and decay vertices. However, to be useful, these simulations must take account of a wide range of physical processes occurring at both the quark and meson level. The theory underpinning these processes may not be well understood, and experimental input from other measurements may be at the wrong energy scale, requiring extrapolation over several orders of magnitude. In consequence the data produced by event generation software is at best an approximation to the true behaviour of the particles, and at worst a fiction. It is clearly recommended that

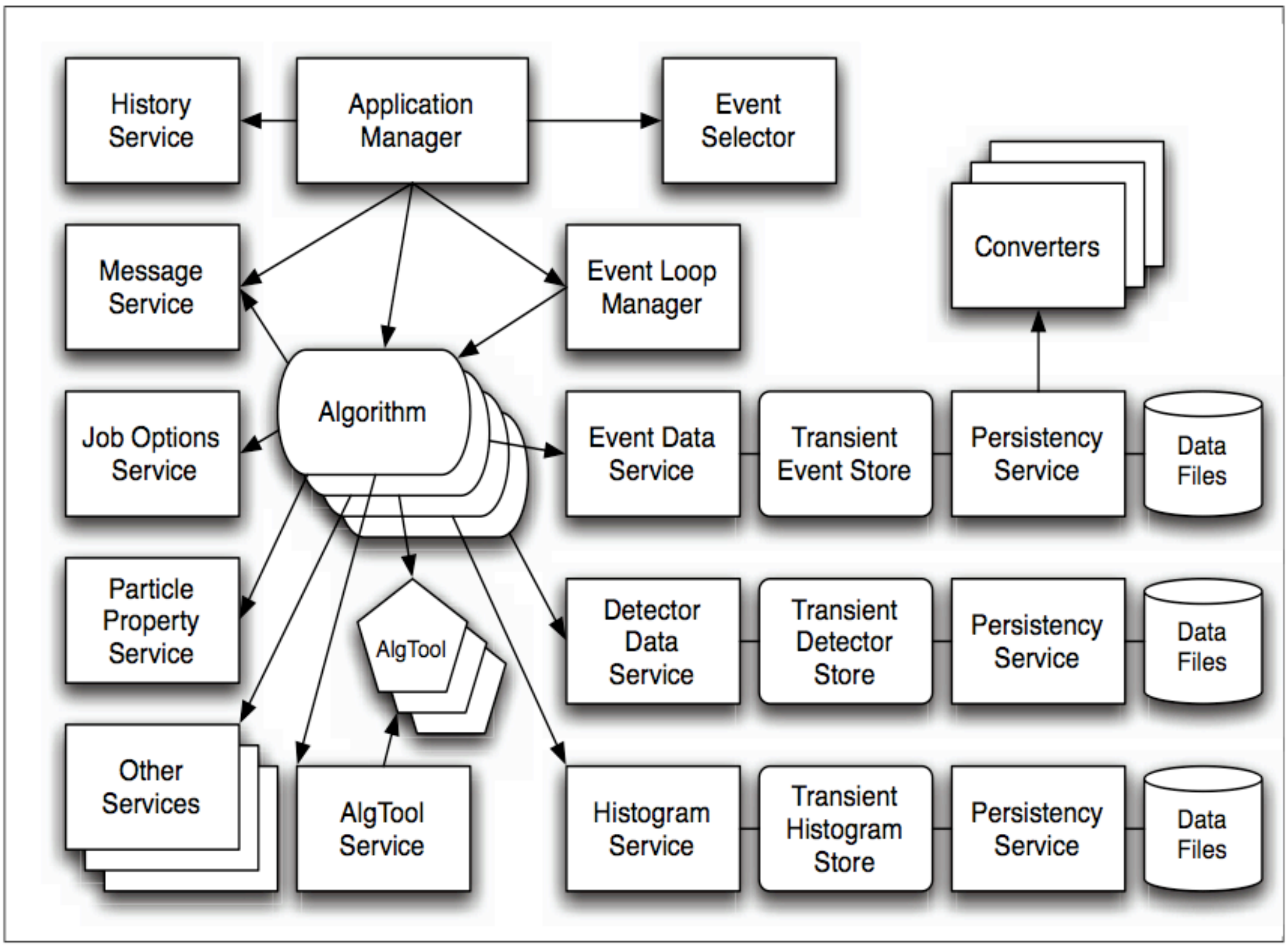


Figure 5.3: Components of Athena (taken from [29])

any study involving these generated data sets uses more than one package, and the results compared as a "reality check"

The production and decay of sub-nuclear particles is inherently random. The best any theory can give is an approximation to complex amplitude for the process. These amplitudes (or in most cases, the decay probabilities) are taken as input data by the event generator, along with branching ratio information. Monte Carlo techniques are then used by the software to decide whether a given process, with a given set of kinematic parameters, occurs at a certain time and location.

The LHC experiments have a wide range of event generators at their disposal. The development of the packages has been merged into the Generator Services (GenSer) effort [47] which stretches across the four experiments and is part of the LCG project. The core general-purpose packages are Pythia[4] [48], Herwig[5] [49] and IsaJet [50], all of which can handle the simulation of the initial partonic collision described by perturbative QCD. There are also a range of specialized packages such as Photos [51] for QED radiative corrections, Tauola [52] for τ-lepton decays, EvtGen [54] for B-meson decays and Hijing [53] for heavy ion collisions. All of these packages can be interfaced with each experiment's software framework.

The results of the event generation are written in a common format known as HepMC [56]. This object-oriented (C++) Monte Carlo event record arranges the data in a tree structure of particles and vertices (see figure 5.4). The trees are navigated and the data extracted by means of Standard Template Library iterators and simple, intuitively named function calls. HepMC data can be written to both transient and permanent storage. In many cases more than one generator may be used; for instance one package may simulate the hard process, and another may then be deployed to decay the resulting particles. In the case of the Athena framework the different packages are declared as algorithms and are then executed sequentially, with the data passed from package to package via StoreGate (in HepMC format). Additional algorithms known as **Generator Filters** may also be used; these remove generated events from StoreGate which do not pass the requirements of the users. Such applications are particularly useful for simulating the action of the trigger systems.

ATLAS B-physics studies generally use Pythia as the main event generator, and for certain studies the EvtGen package is used to decay B-mesons. In the following sections the B-event generation scheme and the software used will be described in detail.

---

[4]Title held by the Priestess of the Oracle of Apollo at Delphi. This Oracle, speaking through its priestess, was renowned for making ambiguous prophesies.

[5]Monte Carlo package for simulating Hadron Emission Radiation with Interfering Gluons

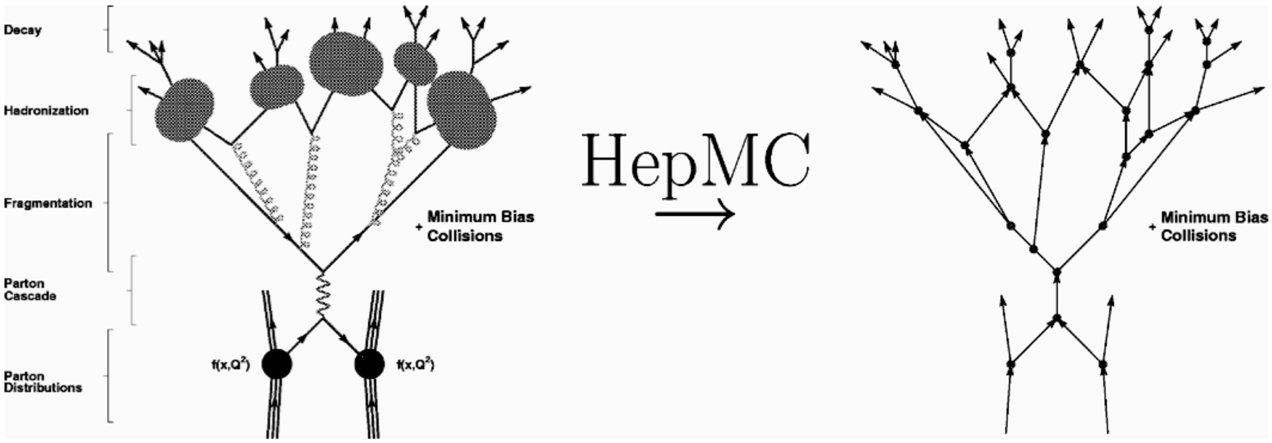


Figure 5.4: Design for the HepMC event record (taken from [56]

Pythia

Pythia [48] is a general-purpose event generator, capable of simulating the entire event production chain. The sequence of events as envisaged by the authors of Pythia is as follows:

1. The two protons approach and collide. The substructure of the protons is characterized by a set of parton distribution functions $f_i^a\,(x, Q^2)$, which parameterize the the probability of finding a parton (quark or gluon) i with a fraction x of the beam energy when the proton a is probed by a scattering of some momentum scale Q^2. The Q^2 dependence is derived through QCD calculations, whereas the x-dependence must be calculated through experiment. As the experiments which determine the parton distribution functions inevitably operate at different energy scales to the LHC, the extrapolation required is a major source of uncertainty in the accuracy of event generators such as Pythia; theoretical limitations also apply as the QCD calculations only run to leading order. The structure functions used at the time of writing are from the CTEQ3 dataset [57].

2. A **shower initiator** parton sets up a series of branchings which establish an initial state shower.

3. A parton from each initial state shower participates in the *hard* process leading to a number (usually two) outgoing partons. It is these partons which determine the overall characteristics of the event.

4. The hard process may create short lived gauge bosons which subsequently decay into partons.

5. The outgoing partons may branch to create final-state showers.

6. Other constituents of the colliding protons may interact in parallel with the main hard process.

7. The QCD confinement mechanism fragments the partons into colour-neutral hadrons.

8. The hadrons decay into final-state particles which can be observed in the detector.

Pythia is currently written in FORTRAN and passes data to the Athena framework by means of an interface algorithm. It contains probabilities for over 200 separate processes, and there are dozens of parameters which a physicist can vary to control the nature of the simulated decays. Finding settings appropriate for the simulation of B-events has been a major effort of the B-physics group [59]. In all simulations carried out up to and including the Rome Physics Workshop, the following $2 \to 2$ hard processes are switched on:

- $q + q' \to q + q'$
- $q + \bar{q} \to q' + \bar{q}'$
- $q + \bar{q} \to g + g$
- $q + g \to q + g$
- $g + g \to q + \bar{q}$
- $g + g \to g + g$

The following branching processes are included:

- $q \to qg$
- $g \to gg$
- $g \to q\bar{q}$
- $q \to q\gamma$

Finally, figure 5.5 shows in detail the various Pythia settings selected after exhaustive studies by the B-physics group [58].

MEANING	ATLAS VALUE	Pythia6 default
b-quark production-related parameters		
Structure fuction	"pypars mstp 51 1 (CTEQ3)"	CTEQ5
Min bias	"pysubs msel 1",	1
Max parton virtuality		
factor to multiply Q^2_{hard}	"pypars parp 67 1",	1
The factorization scale $Q^2_{\text{hard}} =$		
$p_T\ ^2(P_1^2 + P_2^2 + m_3^2 + m_4^2)/2$	"pypars mstp 32 8"	8
B hadron related parameters		
Spin s=1 probability	"pydat1 parj 13 0.65"	0.75
j=1 l=1 s=0	"pydat1 parj 14 0.12",	0
j=0 l=1 s=1	"pydat1 parj 15 0.04",	0
j=1 s=1 l=1	"pydat1 parj 16 0.12",	0
j=2 s=1 l=1	"pydat1 parj 17 0.2",	0
Peterson fragmentation ϵ_b	"pydat1 parj 55 -.006"	-.005
No B-oscillations	"pydat1 mstj 26 0",	
Multiple interactions parameters		
Model	"pypars mstp 82 4 (double gauss)",	1 (step fcn)
Regularization p_T scale	"pypars parp 82 1.4",	1.9
Double gauss parameters	"pypars parp 83 0.5",	0.5
"	"pypars parp 84 0.4",	0.2
Gluon probability	"pypars parp 85 0.9",	0.33
Two gluon probability	"pypars parp 86 0.95",	0.66
Energy scale for parp 82	"pypars parp 89 1800",	1000
Power of energy rescaling	"pypars parp 90 0.25".	0.16

Figure 5.5: Pythia tunings for ATLAS B-physics (taken from [59]

PythiaB

Roughly 1% of proton-proton collisions result in the production of a b- or $\bar{b}$-quark. Generating a given quantity of B-mesons therefore requires on average 100 times that number of simulated $p\bar{p}$ collisions. Furthermore, if the meson of interest has a relatively low production cross section, only a small number of the b-quarks produced will hadronize into the required particle. It can easily be seen that generating these datasets would be a laborious and CPU-intensive process, were some measure not taken to improve the generation efficiency. The most obvious method would be to intervene in the structure functions themselves to boost the probability of beauty quark production. Technical difficulties aside, this is not desirable as the consequences for the cross section would not be easily evaluated - the exercise would become completely artificial.

Instead, the approach of *repeated hadronization* is deployed. This technique interrupts the generation chain if a b- or $\bar{b}$-quark is produced. The entire event is then cloned a number of times (the number being determined by the user) and each clone is hadronized as if it were a separate independent event. In this way the efficiency of the generation is increased considerably as each hard process resulting in a beauty quark spawns a number of B-events. This scheme is implemented in the Athena algorithm **PythiaB** [59], which interfaces directly with the external Pythia code.

Usually an event generation exercise will require the production of signal processes, so PythiaB also provides a facility for selecting decay along specific channels. By means of the job options file, the user can turn off unwanted decays, leaving only the required channel or channels free. The event counter is only incremented when the required decay is produced, so the final event record contains the signal decay in every event. In all B-physics studies carried out currently, the signal B-mesons formed from the $\bar{b}$-quark are forced to decay into the signal channel of interest. The $\bar{B}$-mesons from the b-quark, on the other hand, are allowed to decay freely according to the general decay tables - no selection is applied on this side of the decay. A diagram showing the data-flow in PythiaB is shown in figure 5.6. PythiaB produces a cross section for the process at the end of the run, given by

$$\sigma_B = \frac{\sigma_{hard} N_{signal}}{N_{hard} N_{loop}} \tag{5.1}$$

where σ_B is the process cross section, σ_{hard} is the total cross section for the allowed hard processes, N_{hard} is the number of produced hard processes, N_{signal} is the number of accepted B-signal events, and N_{loop} is the number of repeat-hadronization loops (the number of clonings). This figure must be further multiplied by the cross section

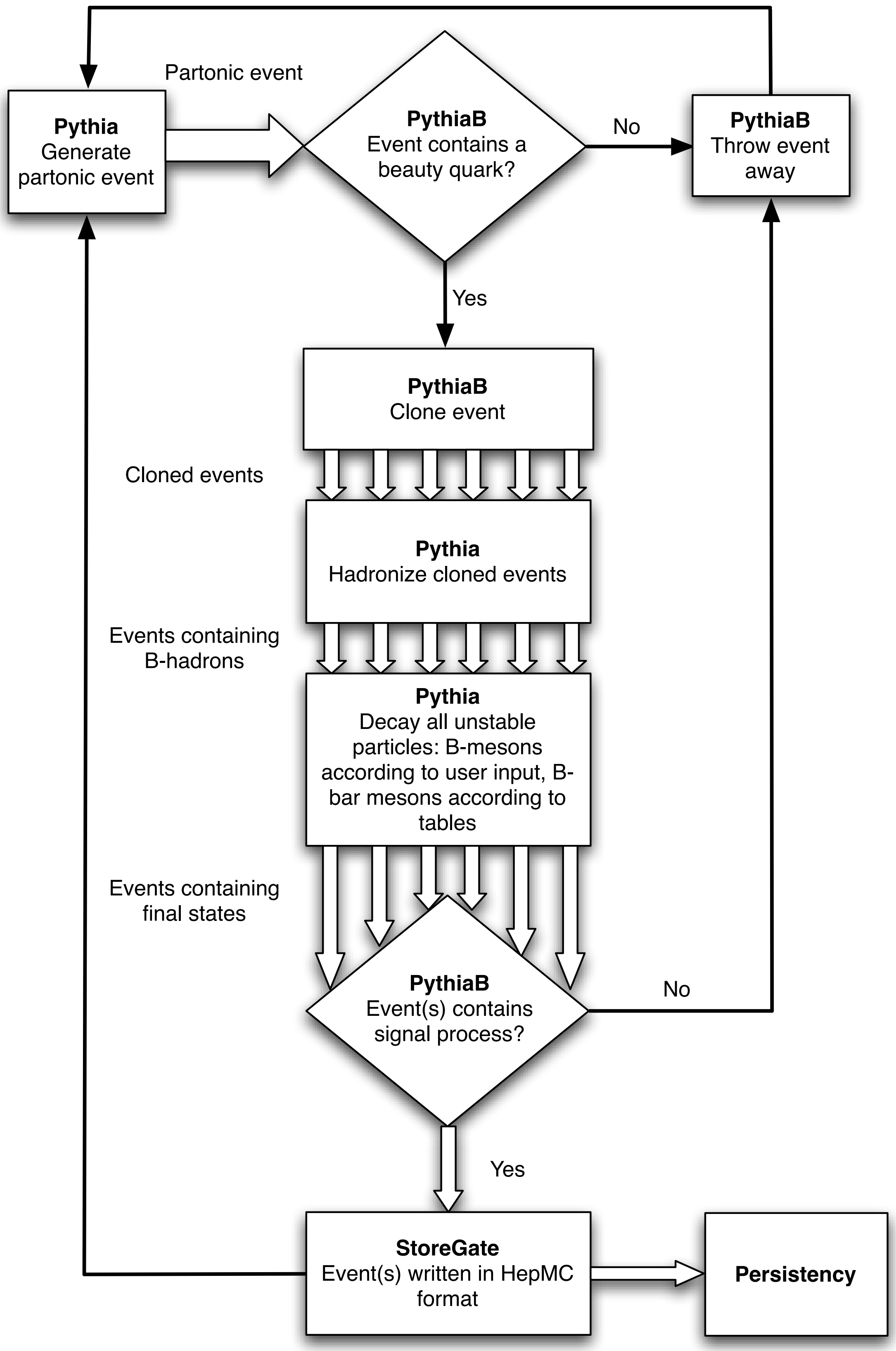


Figure 5.6: Data flow for the PythiaB algorithm

of any forced decays, and by a factor of two to reflect the fact that the other quark is allowed to decay freely.

A further value, the **cloning factor**, is also listed by the program; this gives the number of accepted signal events per set of hadronization loops. This figure should be close to 1.0. A value lower than one indicates that the generation is inefficient, whereas a number well above one indicates that too many signal processes are being produced from the same partonic event, which will clearly give unrealistic kinematics. In the former case the quantity N_{loop} should be increased, and in the latter case the number of hadronization loops should be reduced.

The author has contributed to the development and testing of the PythiaB algorithm.

BSignalFilter

As mentioned above, filters are used extensively in event generation. B-physics applications have a dedicated Monte Carlo filter, known as *BSignalFilter*. This Athena algorithm performs two rôles. First, it removes from the data stream all events containing undecayed unstable particles. The presence of such particles indicates that all possible decay routes have been switched off by the user, and are hence unwanted, or possibly some corruption in the record. Either way, events containing such particles should be removed before the data is committed to storage.

The second rôle of the filter is to simulate the action of the trigger system. It is undesirable for the Monte Carlo data to contain events that the detector cannot observe - this is wasteful of space and analysis time, and pointless where the aim of the exercise is to assess the performance of the detector in measuring particular channels. The filter therefore looks for particles in the event record that, if detected by ATLAS, would activate the triggers. Generally this involves finding one or two separate muons with certain transverse momenta and pseudorapidity, and possibly electrons or hadrons with particular kinematics. Events not containing these particles are removed from the data stream and are not committed to disk. The BSignalFilter presents a report at the end of the run to inform the user of how many events have been discarded. This information is used to re-evaluate the process cross section under trigger conditions. BSignalFilter, in common with all generator filters, reads input to and delivers output from StoreGate, in HepMC format. It is controlled from the job options script.

The author has contributed to the development, testing and documentation of this algorithm, and in particular was responsible for implementing the trigger facility [55].

EvtGen

The EvtGen package [54] is a Monte Carlo program (C++) for the simulating the decay of B-hadrons. It is not capable of modelling the proton collisions, hard processes or hadronization - these tasks must still be carried out with a general purpose event generator such as Pythia. The package was originally developed for the BaBar experiment, but has been widely utilized and is now a component of the GenSer project.

EvtGen has a number of unique features which make it particularly suitable for the simulation of B-decays. Firstly, it takes *complex amplitudes* as input, rather than the usual case where pre-evaluated decay probabilities are passed to the code. This approach is highly advantageous wherever interference between two or more processes may occur. In a conventional "probability-based" generator the best the code can do is to add up the probabilties; any cross-terms are lost. EvtGen, on the other hand, assembles the complete decay amplitude before the decay probability is evaluated by taking the square of the modulus. If this complete amplitude is a sum of several component amplitudes, as is often the case in B-hadron physics, the cross-terms will be included in the final probability. Amplitudes are encoded in C++ classes which are referred to as **models**. All models inherit from a base class which provides the common features and operations associated with complex amplitudes. In consequence the construction of a new model is a straightforward task, once the form of the complex amplitude has been derived.

Secondly, EvtGen includes *spin density matrices* in the implementation of the amplitudes. This allows the code to take account of decays involving non-trivial spin configurations, which causes angular dependences in the distribution of the decay products. Given that these angular distributions can be used to extract fundamental physical quantities, their simulation in an event generator is highly desirable.

Thirdly, the EvtGen package deploys a novel decay algorithm, in that it evaluates the decay tree using a *node-wise* approach. This method is best understood through the use of an example. Consider the decay $B_s \to J/\psi\phi$, with $J/\psi \to \mu^+\mu^-$ and $\phi \to K^+K^-$. The complete decay amplitude can be written as

$$A = \sum_{\lambda_{J/\psi}\lambda_\phi} A^{B_s\to J/\psi\phi}_{\lambda_{J/\psi}\lambda_\phi} \times A^{J/\psi\to\mu\mu}_{\lambda_{J/\psi}} \times A^{\phi\to KK}_{\lambda_\phi} \tag{5.2}$$

where $\lambda_{J/\psi}$ and λ_ϕ represent the spin of the J/ψ and ϕ respectively; the sum is over all possible spin states. The most obvious way of generating events for the process would be to produce kinematics according to phase space for the whole decay chain, that is, for each of the seven particles, and apply this data to accept-reject Monte Carlo according to the modulus-square of the amplitude. However, though trivial to describe, this is decidedly non-trivial in practice. Firstly the accept-reject Monte Carlo requires the maximum probability for the entire process, a value which can be difficult to obtain, especially if the decay chain is long. Secondly, where an event is rejected in the Monte Carlo, kinematics for the whole chain must be produced again, which is highly inefficient.

The approach used in EvtGen is radically different. Here, each node of the decay tree is treated *independently*, with accept-reject Monte Carlo carried out separately for each node point. This method avoids the two problems outlined above. The algorithm proceeds as follows [54].

1. The decay of the B-meson is first dealt with. The decay probability is given by
$$P_B = \sum_{\lambda_{J/\psi}\lambda_{\phi}} \left| A^{B_s \to J/\psi\phi}_{\lambda_{J/\psi}\lambda_{\phi}} \right|^2 \tag{5.3}$$
Kinematics are randomly generated for the decay of the B_s into $J/\psi\phi$, and evaluated according to P_B until the event passes the accept-reject algorithm.

2. A *spin density matrix*, describing the state of the J/ψ after summing over all the possible spin states of the ϕ, is now constructed. This matrix is given by
$$\rho^{J/\psi}_{\lambda^*_{J/\psi}\lambda'_{J/\psi}} = \sum_{\lambda_{\phi}} A^{B_s \to J/\psi\phi}_{\lambda_{J/\psi}\lambda_{\phi}} \left[A^{B_s \to J/\psi\phi}_{\lambda'_{J/\psi}\lambda_{\phi}} \right]^* \tag{5.4}$$

3. The J/ψ is now decayed in the same way as the B, with the probability given by
$$\frac{1}{\mathrm{Tr}\rho^{J/\psi}} \sum_{\lambda^*_{J/\psi}\lambda'_{J/\psi}} \rho^{J/\psi}_{\lambda^*_{J/\psi}\lambda'_{J/\psi}} A^{J/\psi \to \mu\mu}_{\lambda_{J/\psi}} \left[A^{J/\psi \to \mu\mu}_{\lambda'_{J/\psi}} \right]^* \tag{5.5}$$
The inverse of the trace of the spin density matrix is a scaling factor which ensures the maximum probability of the node is independent of the full decay chain. It is proportional to the decay rate. The spin density matrix also sits within the sum over spin states; this ensures the correlations between the kinematic variables in the decay are included.

4. A second spin density matrix is now formed for the ϕ decay.
$$\rho^{\phi}_{\lambda_{\phi}\lambda'_{\phi}} = \sum_{\lambda_{J/\psi}\lambda'_{J/\psi}} A^{J/\psi \to \mu\mu}_{\lambda_{J/\psi}} \left[A^{J/\psi \to \mu\mu}_{\lambda'_{J/\psi}} \right]^* A^{B_s \to J/\psi\phi}_{\lambda_{J/\psi}\lambda_{\phi}} \left[A^{B_s \to J/\psi\phi}_{\lambda'_{J/\psi}\lambda'_{\phi}} \right]^* \tag{5.6}$$

5. The ϕ is decayed in an identical manner to the J/ψ, replacing $\rho^{J/\psi}$ with ρ^{ϕ}

This scheme can be extended indefinitely to longer and more complex chains. The evaluation of the spin density matrices necessarily becomes increasingly complex as the chains lengthen, but as these calculations are all done by the computer this is of no consequence to the physicist; the generation of kinematics and the accept-reject algorithm consume far more computation time than the calculation of the spin density matrices. The reduction in the number of accept-reject cycles is considerable.

EvtGen is controlled by means of a decay table, which lists all the decays, their branchings and the name of the decay model (complex amplitude) to be used to effectuate the decay. A user decay table can be used to open or close specific channels. The decay models themselves are written in C++ so new models must be built (compiled) before they can be used. The program reads input and produces results in the HepMC format, so integrating it with other packages is straightforward. The author was involved in writing the interface between EvtGen and Athena.

The B-event generation scheme

Finally, a brief description is given of the complete generator chain used for the production of B-physics Monte Carlo data. PythiaB is always used to produce the partonic event and perform the hadronization, as described above. In some cases the studies do not require angular dependencies or interference considerations, so in this case PythiaB can safely be used to complete the chain and write the data to persistency. However, in many cases the more taxing phenomenologies are of importance, and EvtGen is deployed for such studies. The scheme then takes the following form.

1. PythiaB produces the partonic event and performs the hadronization, as described above.

2. A job options file closes all B-decay channels in Pythia, so that the program treats B-hadrons as stable particles.

3. The HepMC record of the Pythia-produced hadrons is passed to EvtGen.

4. EvtGen decays the hadrons according to its own decay table, or the user's table where one has been provided.

5. The HepMC output from EvtGen is passed to the BSignalFilter, which removes any events not passing the trigger requirements or containing undecayed unstable particles. Any B-mesons for which all channels have been closed off in the EvtGen user decay file will fall into this category, so the net effect is

that all events passing the filter are guaranteed to contain the signal processes requested in the user decay file.

6. The data is written to persistency.

5.3.3 Simulation, Digitization and pile-up

Once the events have been generated and committed to persistency in HepMC format, they are passed to detector simulation software. This takes the final states produced by the event generator and models the passage of these particles through the detector. The simulation must account for as wide a range of processes as possible and should closely reflect the actual geometry of the machine, in order to arrive at reasonable assessments of detector performance.

The simulation engine is provided by the C++ package Geant4[6] [60]. The software provides tools for setting up detector geometry with a wide range of materials, for tracking particles through materials and magnetic fields, creating new events from interactions with the materials and for simulating the response of active detector components to the particles. Services are provided for the storage of events and tracks, and powerful visualization tools allow the user to view the installed geometry on-screen (e.g. figure 5.7). At the core of the package are a set of Monte Carlo algorithms for a wide range of physics processes [62], including electromagnetic and hadronic interactions, and the decay of particles in flight. Geant4 is a generalized package, being used in space and medical applications as well as high energy physics. The specific implementation of Geant4 for ATLAS [61] takes the required simulation tools and establishes the detector geometry and composition from the main Athena geometry service (**GeoModel**) by means of a conversion tool known as **Geo2G4**. The simulation must be able to account for particles with energies as low as $10eV$, which corresponds to the ionization potential of the active gases in many of the detector tubes, and as high as a few TeV, for muons which deposit all their energy in the calorimeters [117]. The tracking detectors require a particularly detailed simulation to allow an assessment of track reconstruction efficiency and momentum measurement accuracy. Apart from handling the interactions of leptons, photons and hadrons with detector material, which lead to secondary showers, effects such as bremsstrahlung and pair production must be taken into account. Inert detector material such as cyrostats, support structures, adhesives and cabling cause signifi-

[6] *Geant* is French for *giant*, and at over one million lines of code, two thousand classes and several thousand files, this is an appropriate name. The fourth release of the code is the first version to be coded in C++, the previous versions being written in Fortran

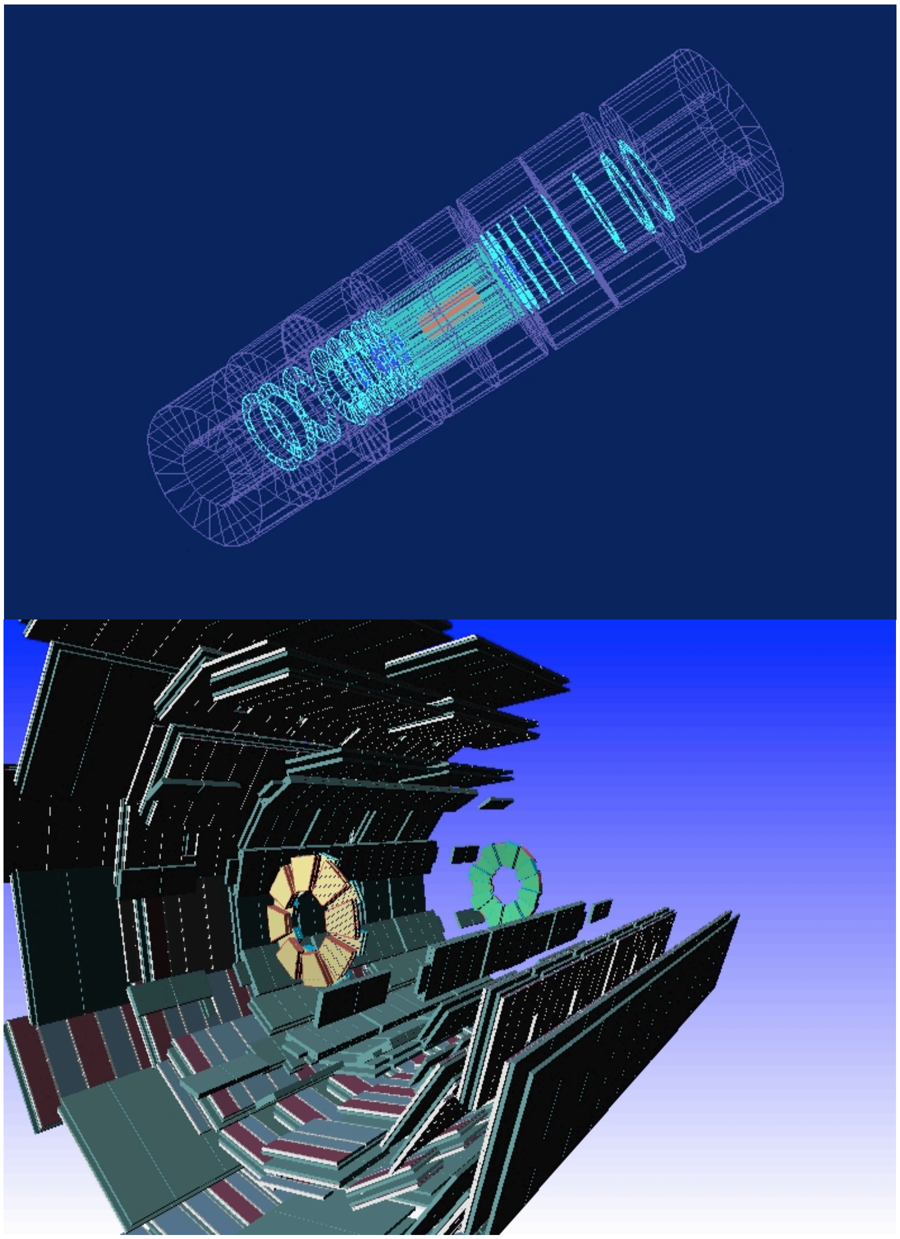


Figure 5.7: Examples of Geant4 visualization technology, showing the ATLAS inner detector and muon systems. Taken from [61].

cant showering and are included in the detector description. It is clearly particularly important that the ionization effects of particles passing through the active detector elements are well-understood and accurately simulated, given that these effects produce the ultimate output of the machine. The geometry of the magnetic field must also be well-understood to gain an insight into the momentum measuring and muon identification capacity of the detector. The final output of the simulation software is referred to as SDO (Simulation Data Objects)

The process by which data on ionization in the active components, produced by Geant4, is converted into a stream of electronic signals ("hits") is known as **digitization**. The output of the digitization process is identical in form to the data which will emerge from ATLAS in 2007. Athena performs the digitization process separately from simulation: SDOs are written to persistency before being read in by the digitization software. Each sub-detector has a different set of digitization algorithms. The final output of the digitization stage is known as RDO (Raw Data Objects), and in general these files are much smaller than the corresponding SDOs. Detector simulation tends to be the most time consuming part of the Athena chain.

When the LHC is operating at full luminosity, it is estimated that around 23 proton-proton collisions will occur per bunch crossing [117]. Natural ionizing radiation from outwith the detector will also contribute to the background. This is simulated by *piling up* the signal and background data - the events are superimposed during the digitization stage. However, this study concentrates on the low-luminosity phase of the LHC schedule, and in this case there are only two collisions per bunch crossing so pile-up is not a major consideration. It has not been considered in this work.

5.3.4 Reconstruction software

Reconstruction is the process by which pattern recognition software inspects the digital hits from the detector and attempts to deduce the physics event which produced them. This is undoubtedly the most important component of the offline software, as the success of the experiment is wholly dependent on the ability of the algorithms to reliably interpret the raw data from the detector. Indeed, the essential purpose of the simulation chain, from generation through to digitization, is the testing of the reconstruction code. The reconstruction code under development at the time of writing is the software which will be deployed "for real" in 2007.

The reconstruction remit covers all of the detectors subsystems - the inner detector, calorimetry and muon chambers. Clearly different algorithms are needed for different sub-detectors, but it must be possible to combine the output of these separate software components to obtain an overall picture of the event; for instance, the inner detector tracks must be matched up with hits in the muon chambers or clusters representing energy deposition in the calorimeters. The software which merges the output of the various algorithms is referred to as *combined reconstruction*. Some particle identification (photons, electrons, muons, taus) is carried out at this stage, although most is deferred to the later step of analysis data building.

The reconstruction software consists of dozens of packages and job options files, some of which contain finely tuned input values. To avoid confusion and in an attempt to standardize the reconstruction run by different physics groups, Athena provides two overarching packages - **RecExCommon** [63] and **InDetRecEx** [64]. As the names suggest, these packages consist of example job options scripts for running reconstruction in the entire detector and the inner detector respectively. These packages do not contain any executable code of their own - they merely call other packages. As B-physics studies currently rely solely on tracking and muon identification, and as, at the time of writing, the muon identification software is still at a very experimental stage of development, the inner detector reconstruction is used alone. Accordingly, all results presented in this work have been produced with the InDetRecEx package.

Event Summary Data

The reconstruction software produces output in the form of Event Summary Data (ESD). This takes the form of persistified C++ data objects which represent the geometric constructions produced by the pattern recognition software (for example, tracks, vertices and energy clusters). Identification of muons, electrons, photons and taus, and the building of jets, takes place once these core objects have been formed.

All current B-physics studies make use only of tracking data (although once the muon identification software is ready this information will also be used). The ESD track class is known as **Trk::Track** [65] (the initial *Trk* being the ESD tracking namespace). It is not designed for performing physics analysis, but is intended for use by other reconstruction algorithms and the combined reconstruction (another track class is used for analysis; see later). Aside of information on the hits that created it, a track is defined by five **perigee** or **helix parameters**:

- $\frac{q}{p}$, the inverse of the resultant momentum multiplied by the charge
- ϕ_0, the azimuthal angle of the tangent to the track at the point of closest approach to the nominal beam axis $(x, y) = (0, 0)$

- d_0, the transverse impact parameter, which is the minimum distance from the track to the nominal beam axis in the x, y plane
- z_0, the z-coordinate of the track at the point of closest approach to the beam axis
- θ, the slope of the track in the rz plane.

The parameters are described graphically in figure 5.8. The diagram on the left shows the track in the xy plane; the track is in red, the beam axis green and the circle (centred on C) displaying the radius of curvature of the track is in purple. The decay vertex (the point at which the track originated) is labelled V and the point of closest approach, A. On the right, the same track is shown in the rz plane, such that it appears to be a straight line. The same points V and A are labelled. It should be noted that whilst the components of the momentum of the track are not explictly part of the perigee, these quantities can be extracted from the perigee by means of simple trigonometry. In general the track fitting algorithms produce the five parameters and a 5×5 covariance matrix. The ESD Track class contains this information as a data member. The remaining data members consist of the fit quality, hit information and the **TrackStateOnSurface** which provides the state of the track (scattering angle, fit quality) on a given surface. It is important to note that the class is common - all the pattern recognition algorithms can interface with the Track constructors and so the final output is always in the same data structure, irrespective of which algorithm performed the calculations. The common ESD vertex is a very simple data object, consisting only of the position and associated covariance matrix, and information about the fitter which created it. As with the common track, all of the reconstruction packages create vertices of the same class.

Track reconstruction algorithms - xKalman and iPatrec

Currently the reconstruction of tracks from hits in the inner detector is performed in Athena by two pattern recognition programs - **xKalman** [66] and **iPatrec** [68].

The **iPatrec** algorithm begins with a *seed* defined beyond the outer surface of the inner detector (generally this seed is a cluster in the electromagnetic or hadronic calorimeter or a hit in the muon system). A “road” is then developed between the interaction point and the seed [72]. The roads are divided into cylindrical zones which radiate out from the interaction point. Combinations of hits from three or more zones constitute a track candidate, a hit in the innermost partition being compulsory. Track fitting then proceeds through a process of least-squares minimization, where the quantity to be minimized is the distance between the helix and

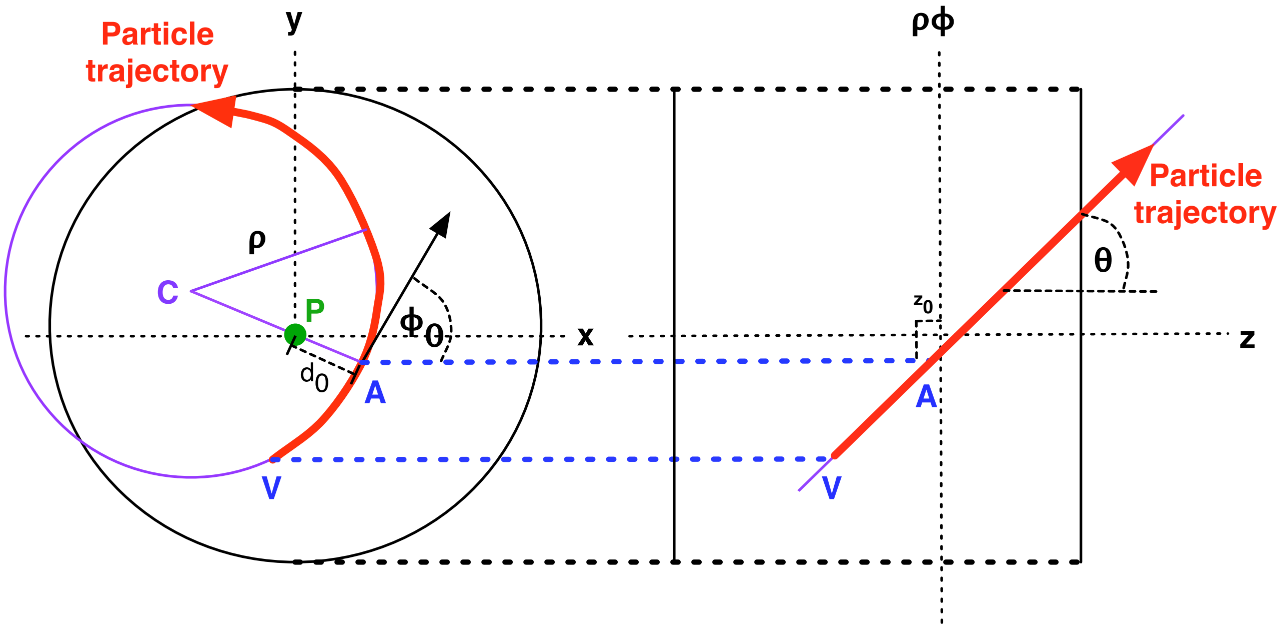


Figure 5.8: The ATLAS track helix parameters. Based on [65].

the hit centroids [68] and the minimization variables are the five helix parameters. The tracks are ranked according to the fit quality, which is determined by the fit χ^2, the number of hits and absence of "holes" (regions of active detector through which the track appears to traverse without leaving any hits) and the radial track length. In common with all studies carried out for the Rome physics workshop and its aftermath, iPatrec was used as the reconstruction package.

xKalman begins by finding track segments in the TRT. It deploys two techniques - a histogramming method for use in the TRT and the Kalman filter-smoother formalism [67] for the high-granularity sub-detectors. After the TRT search the tracks are extrapolated into the SCT, and thence to the pixels. Finally, the tracks are propagated back into the TRT, where the straw-tube drift time is accounted for.

Vertex fitting algorithms - CTVMFT and VxBilloirTools

The fitting of a set of tracks to a **vertex** (common point of origin), such that two or more tracks can be reconstructed as decay products from a single particle, is of importance at both the reconstruction and analysis stages. At the reconstruction level, the main task is to find the positions of the primary vertices, from which the entire event originated. During physics analysis, the success or failure in fitting a group of tracks to a vertex is the main means of determining whether the tracks really did emerge from the same decay. Locating the position of the vertices enables the calculation of the proper decay time. ATLAS currently uses just one set of vertexing tools at reconstruction level - the **VxBilloirTools** package [69]. At the analysis level a number of other packages are available, the main one of interest for this work being the CDF-authored package **CTVMFT** [71].

VxBilloirTools is a C++ vertex fitter which is fully integrated with the Athena reconstruction software. It is an implementation of the Billoir vertexing algorithm [70]. There are two running modes - a *fast fitter* which returns only the vertex position and the associated covariance matrix, and a *full fitter* which also returns the refitted tracks and their covariance matrices.

CTVMFT is a stand-alone Fortran package written originally by CDF for secondary vertex finding. An interface has been developed to enable it to run within Athena. It is not deployed at the reconstruction stage but is extensively used in physics analysis. The code consists of a combined geometric and kinematic fitting routine; after finding an initial approximation of the position of the intersect the code minimizes the χ^2 by adjusting the constituent track parameters. The principal constraint is that the tracks emerge from the same point, but it is also possible to constrain the tracks, or a sub-set of them, to a certain invariant mass (so if one was searching for J/ψ mesons decaying into muon pairs, one could set the mass constraint of the candidate pair to the mass of the J/ψ). Furthermore, it is possible with this code to force the resultant momentum vector of all the tracks making the vertex to point in a certain direction (usually at the primary vertex, from which the event originated). The code returns the vertex position, the perigees of the refitted tracks and the associated covariance matrices. These constraint facilities are not available in VxBilloirTools, and at the time of writing the track refitting performed by this code has not been comprehensively tested, whereas the performance of CTVMFT has been exhaustively studied [115][116]. For this reason CTVMFT is at the time of writing the package of choice for B-physics analysis in ATLAS.

5.3.5 Physics Analysis software

The final step in the simulation (or data processing) chain is physics analysis, where the reconstructed objects are studied in detail with the aim of identifying particular decay processes, and extracting the physically interesting quantities from them (decay rates, lifetimes, peak widths, branching ratios, decay angles etc). The intention is that physicists write the analysis code themselves using a range of centrally provided tools.

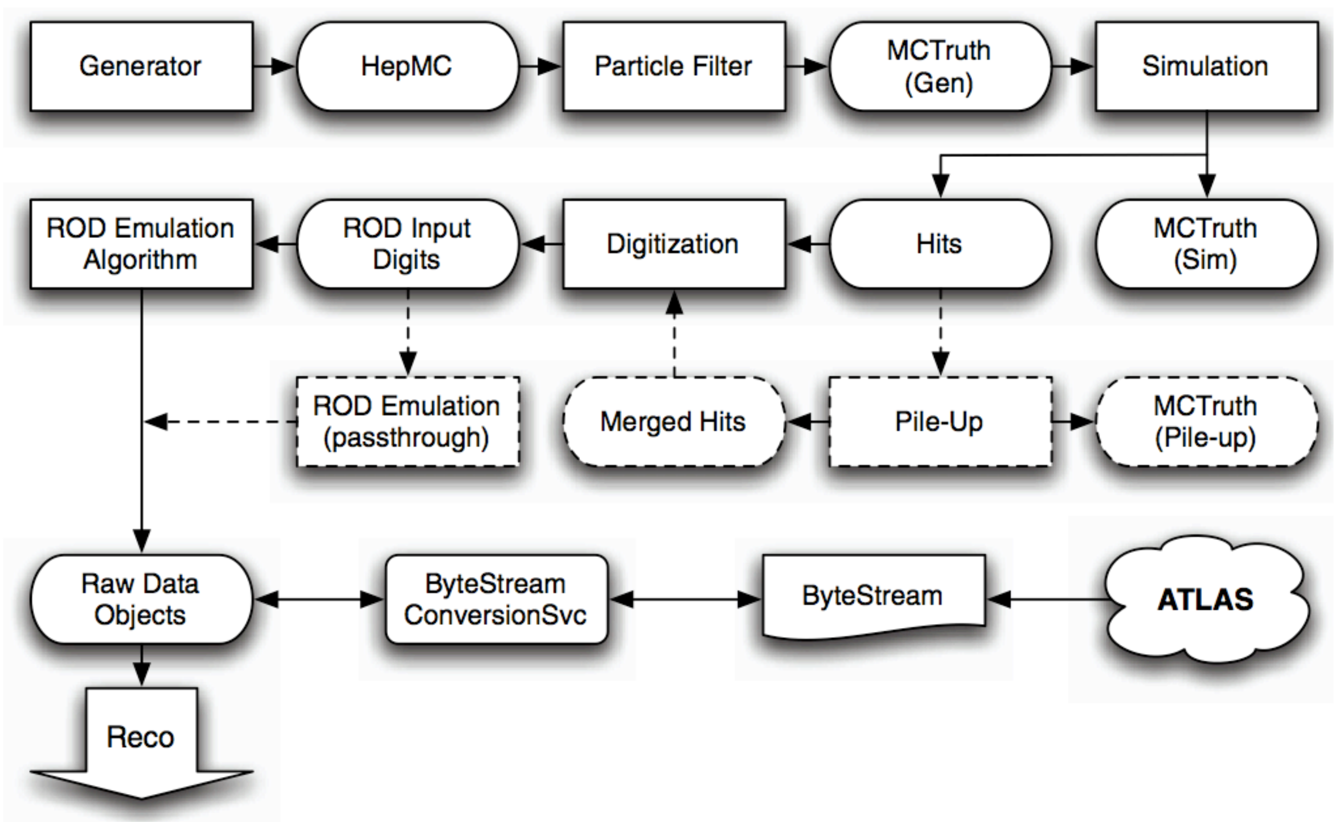


Figure 5.9: The ATLAS simulation chain from event generation to reconstruction. Sharp-cornered boxes indicate processing steps; round-cornered boxes represent data object types. The detector itself is also represented. Taken from [29].

Analysis Object Data

The output from reconstruction (ESD) contains more information than is generally required for physics analysis. The analyst will usually want simple quantities such as kinematics and particle identification rather than the panoply of data which is contained within the ESD. Furthermore, once ATLAS begins to take data it will be impractical to transmit the complete ESD around the world to the various institutes carrying out physics analysis - the reconstructed data objects are too large. In consequence the ESD is "distilled" before analysis into a simplified and lighter form, known as **Analysis Object Data** (AOD).

AOD objects differ from their ESD counterparts in that they are all direct representations of physical objects (or loci of physical objects in the case of tracks). In consequence they have a common inheritance structure - they all inherit from a **particle** class, which itself inherits from a **four momentum** class. An important feature of the AOD is *navigation* - the ability of AOD objects to point to their constituents, and possibly to the ESD objects which created them (if the data files are available). In consequence all AOD objects ultimately inherit from the **INavigable** base class. The inheritance structure of the AOD is shown in figure 5.10.

AOD is constructed from ESD in a process called **AOD building**. Loose selection cuts are placed on the reconstructed objects; it is envisaged that a number of AOD streams will be produced, each one tailor-made to specific physics groups. Examples of AOD objects include **muons**, **electrons**, **photons**, **taus** and **TrackParticles**. It is envisaged that the B-physics studies will require TrackParticles and muons only; at the time of writing the muon identification software is not ready and so current studies use only TrackParticles. The TrackParticle [76] is the AOD track class; it inherits from the particle base class (and hence the four-momentum) but also contains the track perigee and the fit quality. It is possible to navigate from the TrackParticles to the ESD Trk::Tracks which created them, if the ESD datafiles are available.

Truth Association

A vital feature of the AOD objects is *truth association*, where reconstructed objects produced from simulated data are associated with the original Monte Carlo particles (from either event generation or Geant detector simulation). Feasibility studies, the calculation of detector resolutions, reconstruction and tagging efficiencies all require knowledge of the original Monte Carlo truth at the analysis stage. The AOD allows the analyst to navigate from any object inheriting from the particle base, to the Monte Carlo particle which produced it. The Monte Carlo data is returned to the

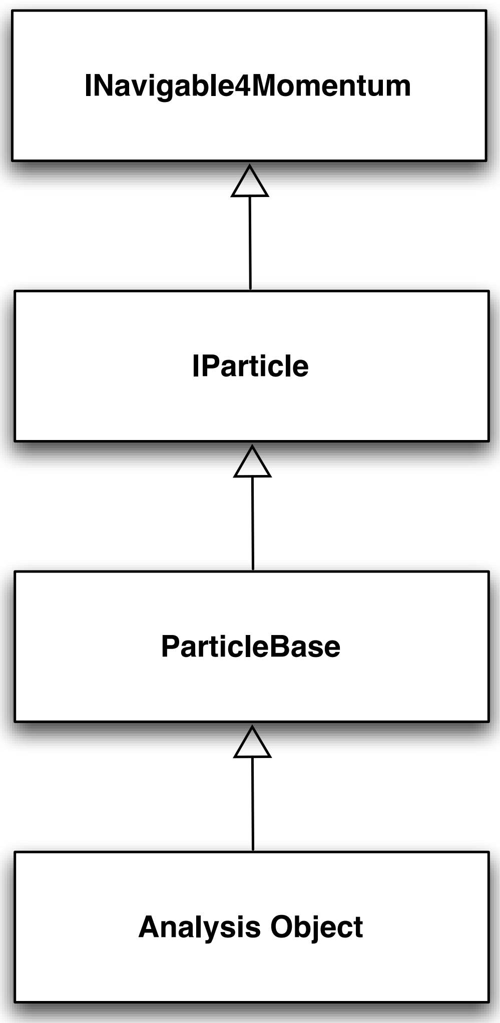


Figure 5.10: The AOD class inheritance diagram

user either in HepMC format, or alternatively in a new AOD truth object called **TruthParticle** which inherits from the particle base. Exactly what proportion of the original truth events pass through to the analysis stage depends entirely on the settings imposed at reconstruction. Clearly the more Monte Carlo truth passed into the AOD, the larger the files will be. It is also possible to summon the whole generated event (in HepMC format), independently of the reconstructed AOD objects, at the analysis stage.

B-physics analysis suite

A range of software tools are provided by Physics Analysis Tools (PAT) group [78] to allow analysts to easily construct their code. Along with this a number of sample analysis packages are provided. The tools are generic and rely heavily on **templating**, so that the same functions can be used on a variety of AOD objects. Tools for dealing with combinatorics, for making cuts, sorting and navigating to other objects are all provided. However, exercises carried out by the author led to the realization that the PAT functions were not suitable for the complex operations which are necessary to handle the chain topologies of B-physics events. Whilst the tools would certainly have been capable of performing the operations, the number of lines of code required per operation was considerable, and the code quickly became unwieldy. The tools also required the use of rather terse C++ which made the code difficult to follow - the aim was to make the analysis code as "English-like" as possible, and this was not achieved when the PAT functions were used.

Following on from this study, the author designed and wrote the **BPhysToolBox**, a class containing a range of functions specifically to facilitate B-physics analysis. Additional classes to provide users with a simple means of accessing the vertexing algorithms were also provided, along with a sample analysis algorithm (for the channel investigated in this work, $B_s \to J/\psi(\mu\mu)\phi(KK)$. The code was deployed in the PhysicsAnalysis directory of the ATLAS software [79]. The code has since evolved, with other users adding additional sample algorithms and new functions to the toolbox. It is envisaged that the code will continue to grow, with more tools and topologies added, so that when data taking begins a full suite of B-physics analysis algorithms will be available.

The package comprises three sub-packages - **BPhysAnalysisObjects**, **BPhysAnalysisTools** and **BPhysExamples**. The BPhysAnalysisObjects package contains two utility classes, **BPhys::Vertex** and **BPhys::VertexAndTracks**; the former is used for storing the output of the secondary vertex finding code, and the latter for associating such a vertex with the tracks that produced it.

The BPhysAnalysisTools package contains three classes - the BPhysToolBox, and two additional classes which provide the interfaces to the secondary vertexing programs (CTVMFT and VxBilloirTools). The functions available in the BPhysToolBox enable the analyst to:

- Calculate the transverse momentum and pseudo-rapidity of a track
- Remove tracks which do not satisfy the required kinematic (transverse momentum and pseudo-rapidity) cuts from the track collection
- Form all unique pairs or all unique triplets of tracks
- Select track pairs where the two are oppositely charged
- Calculate the invariant mass of a set of tracks
- Calculate the proper time of a decay, the impact parameter and the transverse decay length, given the position of the decay vertex and the primary vertex
- Obtain the Monte Carlo particle corresponding to a given track, or provide the closest match where no direct association exists

The BPhysExamples package currently contains algorithms for performing analysis on the following decay channels:

- $B_s \to J/\psi(\mu\mu)\phi(KK)$
- $\Lambda_b \to \Lambda_0 l^+ l^-$
- $B_s \to D_s \pi$

5.3.6 Data analysis and visualization packages

The output of the physics analysis code is generally numerical. For short tests the user may wish to write the data out explicitly in ASCI format, but for full analyses, where histograms are required, one of the data analysis packages will be used. The analysis code will generally write the numerical data to an **nTuple** (array) via the Athena service **NTupleSvc**. The data analysis package then reads the arrays and produces the plots.

There are two such packages in common use at CERN - the FORTRAN Physics Analysis Workstation (PAW) [80] and the C++ ROOT [81]. Both can operate interactively through the use of interpreters (so that FORTRAN or C++ statements can be typed into the command line without the need of a compiler) or via scripts.

Both packages have sophisticated graphical outputs, and can summon data analysis routines such as Minuit [82] for fitting. ROOT is more widely used at the time of writing, as the graphics are of a higher quality and the range of facilities is greater. The majority of the plots shown in this work were produced with ROOT, whilst a few were produced using PAW.

Finally, the event display technologies available in ATLAS should be mentioned briefly. These packages read the AOD (or ESD) files and the detector description, and produce a visual representation of the event in the detector, including hits, clusters and tracks. The chief package in use currently is Atlantis[7] [83]. This Java program shows the events in a range of 2-D projections, superimposed on top of simplified diagrams of the detector components. Other programs in use include HEPVis [84] and Persint [85]. Atlantis tends to be used for debugging reconstruction code - it is often difficult to determine why a particular event should be corrupted by inspecting four-vectors, whilst looking at the track visually will normally reveal the fault immediately. HEPVis and Persint give more sophisticated views of the detector and are used for engineering and design applications. A further use for all of these packages which should not be overlooked is their ability to create impressive-looking images for public relations ("outreach") material. Examples of the output of the three packages is shown in figure 5.11.

[7]The mythical lost island-city, domain of Poseidon and his son Atlas. Its prosperous population eventually succumbed to idleness, gluttony and decadence, and in his anger the King of the Universe, Zeus, caused the sea to swallow it up.

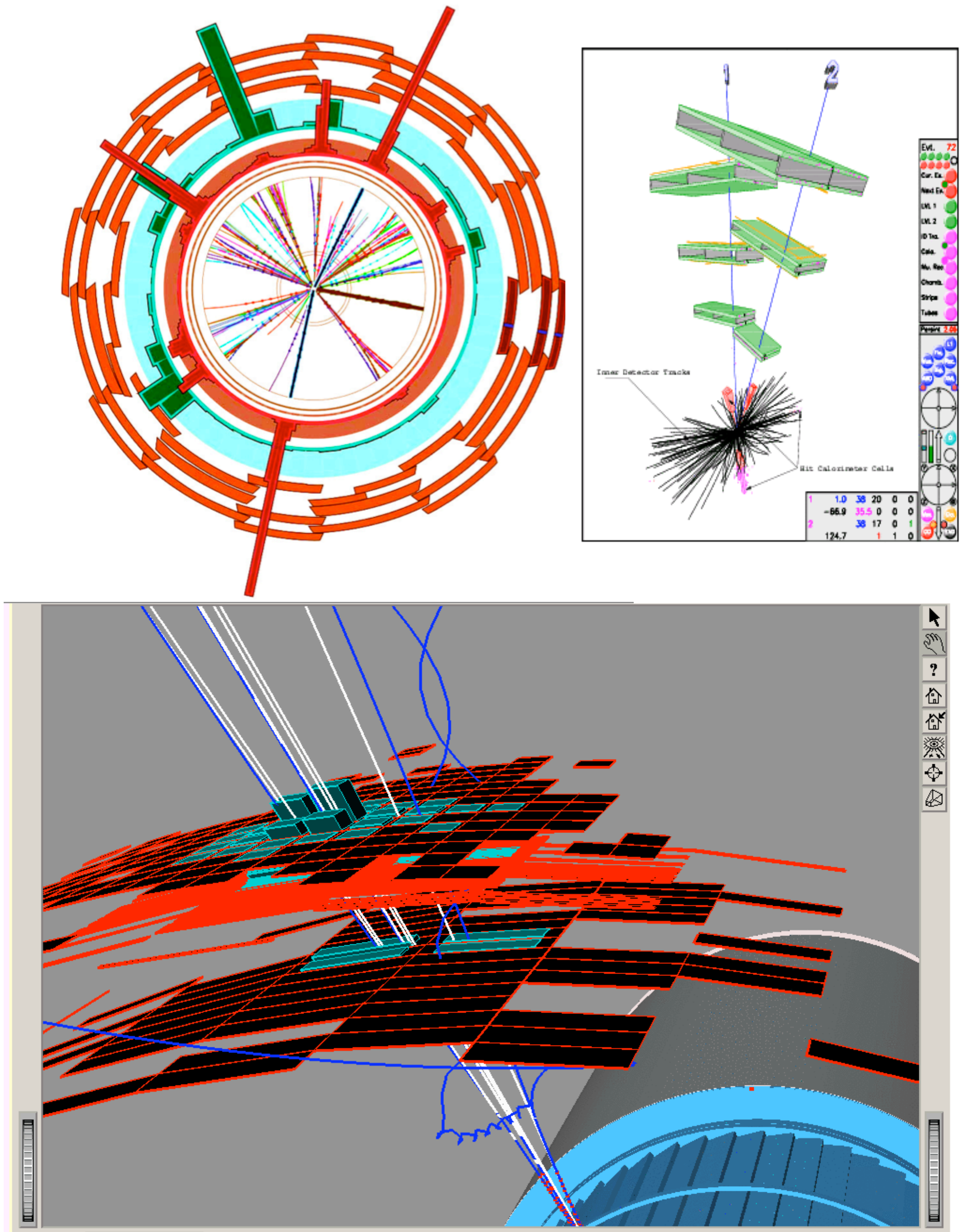


Figure 5.11: Examples of the ATLAS event displays. Clockwise from top left: Atlantis[29], Persint[29], HEPVis[84]

Chapter 6

Analysis of the decay $B_s^0 \to J/\psi\phi$ part I - Decay Modelling and Event Generation

6.1 Theory

6.1.1 General remarks

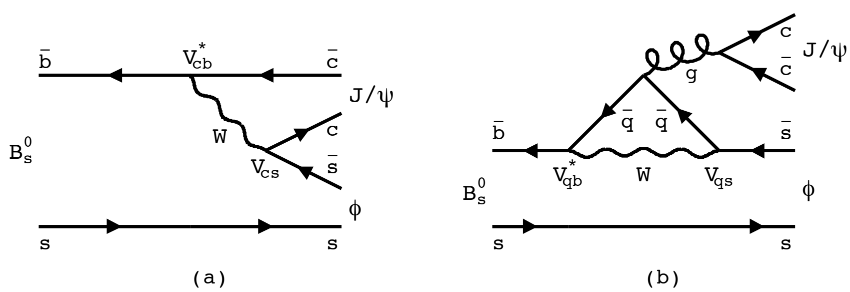


Figure 6.1: Standard Model diagrams contributing to $B_s^0 \to J/\psi\phi$: a) tree, b) penguin.

The lowest-order Feynman diagrams for the decay $B_s \to J/\psi\phi$ are given in figure 6.1. As shown, there are both "tree" and "penguin" contributions to consider. The penguin process (b) has an internal quark loop (u, c and t) which mediates the $\bar{b} \to \bar{s}$ transition. The general decay amplitude can be immediately read off these diagrams, and is given by [89]

$$A(B_s \to J/\psi\phi) = V_{cs}V_{cb}^*(A_{cur}^{c'} + A_{pen}^{c'}) + V_{us}V_{ub}^*A_{pen}^{u'} + V_{ts}V_{tb}^*A_{pen}^{t'}, \quad (6.1)$$

where $A_{cur}^{q'}$ is the amplitude for the current-current process of diagram (a), $A_{pen}^{q'}$ is the amplitude for the penguin processes with internal quark loops ($q \in \{u, c, t\}$) of diagram (b) and the V_{q1q2} are CKM matrix elements. The primes act as a reminder that the transition is $\bar{b} \to \bar{s}$.

First we recall the definition of the CP asymmetry parameter as given in chapter two.

$$\begin{aligned} \xi_{CP} = \frac{q}{p} \cdot \frac{\bar{A}_f}{A_f} \quad &\propto \quad e^{-i\theta^q_{M12}} \frac{A\left(\bar{B}^0_q \to f_{CP}\right)}{A\left(B^0_q \to f_{CP}\right)} \\ &\propto \quad \mp e^{-i\phi_q} \frac{\sum_{j=u,c} V^*_{jq} V_{jb} \langle f_{CP}|\hat{O}|\bar{B}^0_q\rangle}{\sum_{j=u,c} V^*_{jq} V_{jb} \langle f_{CP}|\hat{O}|B^0_q\rangle} \end{aligned} \tag{6.2}$$

where the various terms have the same meanings as in chapter two. We can then write an expression for the CP asymmetry parameter for this process [13]:

$$\xi^{(s)}_{J/\psi\phi} \propto e^{-i\phi_s} \left[\frac{V^*_{us} V_{ub} A^{ut'}_{pen} + V^*_{cs} V_{cb} (A^{c'}_{cur} + A^{ct'}_{pen})}{V_{us} V^*_{ub} A^{ut'}_{pen} + V_{cs} V^*_{cb} (A^{c'}_{cur} + A^{ct'}_{pen})} \right] \tag{6.3}$$

where

$$A^{ut'}_{pen} = A^{u'}_{pen} - A^{t'}_{pen} \qquad A^{ct'}_{pen} = A^{c'}_{pen} - A^{t'}_{pen} \tag{6.4}$$

The unitarity of the CKM matrix has been deployed here to eliminate $V_{ts}V^*_{tb}$. The penguin amplitudes are predicted to be suppressed by a factor of $\mathscr{O}(10^{-2})$ with respect to the tree processes, due to the loop structure of the decay [13]. Furthermore, due to the ratio

$$\left| \frac{V_{us} V^*_{ub}}{V_{cs} V^*_{cb}} \right| \approx 0.02 \tag{6.5}$$

the $A^{ut'}$ pieces are also heavily suppressed. If we make the assumption that the decay $B_s \to J/\psi\phi$ is itself CP-symmetric, that is to say, there is no direct CP-violation, we can immediately write

$$\xi^{(s)}_{J/\psi\phi} \propto e^{-i\phi_s} \tag{6.6}$$

These assumptions of zero direct CP-violation and negligible penguin contributions are followed throughout this analysis. A treatment is presented in [90] where such assumptions are not made; the calculations necessarily become more complex, but an assessment of the experimental implications should be undertaken before the measurement is made "for real".

Thus far, the theory has proceeded in line with the analogous decay in the B_d sector, $B^0_d \to J/\psi K_s$. At this point, however, the phenomenologies of the two channels diverge. $B^0_d \to J/\psi K_s$ decays into a CP-odd eigenstate (-1), whereas the final state configuration for $B_s \to J/\psi\phi$ is an *admixture* of CP-odd and CP-even eigenstates. This complicates any assessment of $\xi^{(s)}_{J/\psi\phi}$ as the contributions of the CP-odd and CP-even amplitudes have to be separated out. To achieve this it is necessary to exploit the spin structure of the process and perform an angular analysis on the final state particles. Though this technique is involved and requires large experimental statistics and good tracking precision, it facilitates the *simultaneous* measurement of the mixing parameters $\Delta\Gamma_s$, Γ_s as well as ϕ_s. For this reason the channel has been described as "Gold Plated" [89]. The remainder of this work explores the technique and presents an assessment for the feasibility of carrying out such an analysis on ATLAS data, based on currently available simulation software.

6.1.2 The structure of the angular distribution

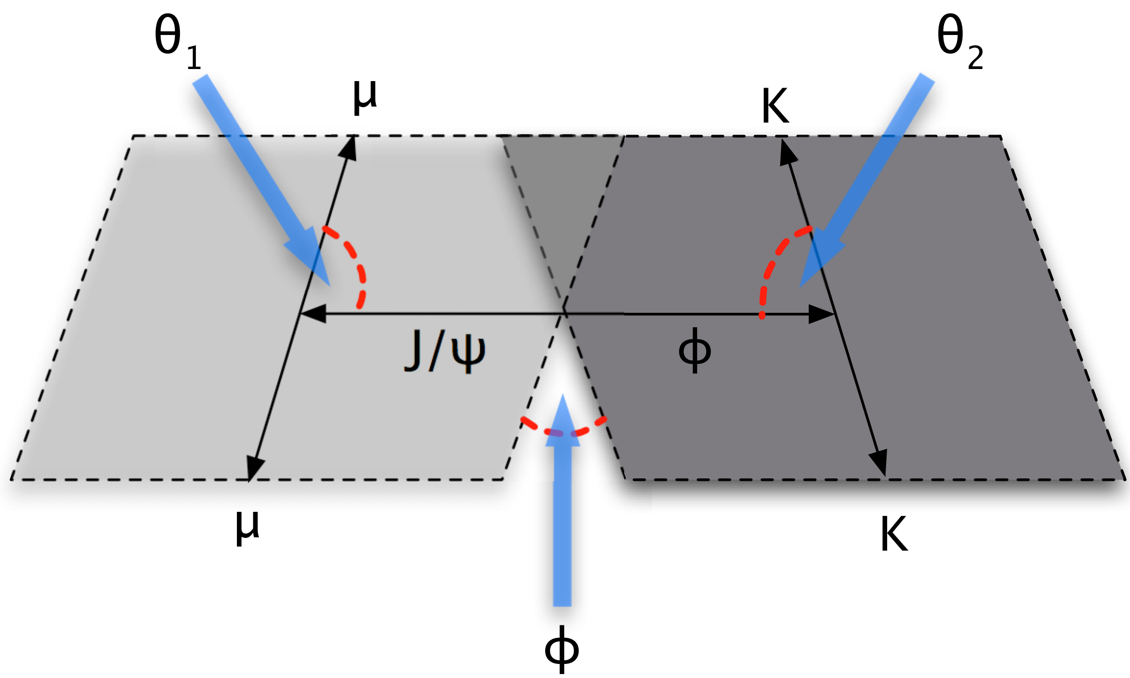


Figure 6.2: "Cartoon" defining the decay angles for $B_s^0 \to J/\psi(\mu\mu)\phi(KK)$. Note that the diagram includes three reference frames - that of the B_s^0, the J/ψ and the ϕ

Constrained by the requirement that the final state particles leave a measurable track in the inner detector, we specifically consider the case where the J/ψ and the ϕ decay to two muons and two charged kaons respectively. The distribution of the angles of the final state particles for this decay was derived in chapter 3 and given by equation 3.60. In this expression the time evolution of the states is implicit in the amplitudes. Now we must derive the explicit time dependence of these amplitudes and construct a complete probability density function for the angular distributions which can be used in a statistical fit.

The decay angles (θ_1, θ_2 and ϕ) are defined in figure 6.2. First we re-state equation 3.60.

$$\begin{aligned}\frac{d\sigma}{d\theta_1 d\theta_2 d\phi} &= \frac{9}{64\pi}\Big[4|A_0|^2 \sin^2\theta_1 \cos^2\theta_2 \\ &+|A_\parallel|^2\left[\left(1+\cos^2\theta_1\right)\sin^2\theta_2 - \sin^2\theta_1\sin^2\theta_2\cos 2\phi\right] \\ &+|A_\perp|^2\left[\left(1+\cos^2\theta_1\right)\sin^2\theta_2 + \sin^2\theta_1\sin^2\theta_2\cos 2\phi\right] \\ &+2\Im\left(A_\parallel^* A_\perp\right)\sin^2\theta_1\sin^2\theta_2\sin 2\phi \\ &-\sqrt{2}\Re\left(A_0^* A_\parallel\right)\sin 2\theta_1 \sin 2\theta_2 \cos\phi \\ &+\sqrt{2}\Im\left(A_0^* A_\perp\right)\sin 2\theta_1 \sin 2\theta_2 \sin\phi\Big]\end{aligned} \tag{6.7}$$

Recalling the formalism of chapter 2, we once again write the expression for the time evolution of a neutral B-meson state [86]:

$$|B^0_{s,phys}(t)\rangle = g_+(t)\,|B^0_s\rangle + \alpha g_-(t)\,|\bar{B}^0_s\rangle \tag{6.8}$$

$$|\bar{B}^0_{s,phys}(t)\rangle = \alpha^{-1} g_-(t)\,|B^0_s\rangle + g_+(t)\,|\bar{B}^0_s\rangle \tag{6.9}$$

where

$$\alpha = e^{-i\phi_s^{WEAK}} \tag{6.10}$$

and

$$g_\pm(t) = \frac{1}{2}\left(e^{-i\left(m_L - i\frac{\Gamma_L}{2}\right)t} \pm e^{-i\left(m_H - i\frac{\Gamma_H}{2}\right)t}\right) \tag{6.11}$$

Now we write down the amplitude for the decay of a B^0_s-meson into the $J/\psi\phi$ state, with a particular polarization λ (see chapter 3).

$$\begin{aligned}A_{B^0_s\to(J/\psi\phi)_\lambda} &= \langle (J/\psi\phi)_\lambda|H_{eff}|B^0_{s,phys}(t)\rangle \\ &= g_+(t)\langle(J/\psi\phi)_\lambda|H_{eff}|B^0_s\rangle + \alpha g_-(t)\langle(J/\psi\phi)_\lambda|H_{eff}|\bar{B}^0_s\rangle\end{aligned} \tag{6.12}$$

and similarly for the conjugate amplitude. Now the instantaneous decay amplitudes are given by

$$\langle(J/\psi\phi)_\lambda|H_{eff}|B^0_s\rangle = \langle(J/\psi\phi)_\lambda|H_{eff}|\bar{B}^0_s\rangle = |A_\lambda(t=0)|\,e^{i\delta_\lambda} \tag{6.13}$$

where $|A_\lambda(t=0)|$ is the modulus and δ_λ the phase. This is generally referred to as the "strong phase" in most literature.

The evaluation of the complete expression for the time dependent angular distribution is now a trivial, if tedious task. The time-dependent amplitudes in equation 6.7 can immediately be calculated by inserting the appropriate $\lambda = 0, \perp, \parallel$, taking the square of the modulus and taking real and complex parts where appropriate.

Recalling that only two out of the three transversity amplitudes are independent, it is possible to reduce the number of strong phases to two. This choice is arbitrary; in general the phase definition

$$\delta_1 = \arg\left\{A_\parallel(0)^* A_\perp(0)\right\} \tag{6.14}$$
$$\delta_2 = \arg\left\{A_0(0)^* A_\perp(0)\right\} \tag{6.15}$$

is taken [89]: this choice is made here.

Explicit articulation of the time dependence leads to an expression of the form

$$W^+(\theta_1, \theta_2, \phi, t) = \frac{d\sigma}{d\theta_1 d\theta_2 d\phi dt} = \sum_k \Omega^{(k)}(t)\, g^{(k)}(\theta_1, \theta_2, \phi) \tag{6.16}$$
$$W^-(\theta_1, \theta_2, \phi, t) = \frac{d\sigma}{d\theta_1 d\theta_2 d\phi dt} = \sum_k \bar{\Omega}^{(k)}(t)\, g^{(k)}(\theta_1, \theta_2, \phi) \tag{6.17}$$

where Ω and g encapsulate the angular and time dependence respectively; k runs from 1 through to 6; and W^+ represents the case where the B-meson was initially in a state $|B^0_{s,phys}(t=0)\rangle$, whilst W^- is for the Hermitian conjugate case, where the B-meson was initially in the state $|\bar{B}^0_{s,phys}(t=0)\rangle$. The components of these expressions, being somewhat lengthy, are given in the tables 6.1 and 6.2.

It is important to note that the overall *structure* of these probability density functions is not expected to be modified by new physics. The angular parts (described by the g functions) originate in simple spin considerations, and it is unlikely that any conceivable modifications to the Standard Model could affect these. The time dependent parts Ω are certainly susceptible to new physics, but only in the values of the parameters - the mathematics of quantum mechanical mixing should not change. The only caveat is the possibility that CP-symmetry may be violated in the decay of $B_s \to J/\psi\phi$ directly; in this case the simplification given by equation 6.6 does not hold true, and additional weak phases would appear in the functions.

6.1.3 Current experimental limits on mixing parameters and the importance of $SU(3)$ symmetry

The probability distribution function derived above contains eight independent parameters - two transversity amplitudes ($|A_\parallel|$, $|A_\perp|$) and two phases associated with them (δ_1, δ_2), the mass difference ΔM_s between the eigenstates, the width difference $\Delta\Gamma_s$ and the average width Γ_s of the two states (or alternatively their specific widths Γ_L, Γ_H) and the weak mixing phase ϕ_s.

Table 6.1: Tabulated components of the probability density function for the distribution of final-state decay angles for the process $B_s \to J/\psi\,(\mu\mu)\,\phi\,(KK)$

k	$\Omega^{(k)}(t)$		$g(t)$
1	$\lvert A_0(t)\rvert^2$		$4\sin^2\theta_1\cos^2\theta_2$
	$\frac{1}{2}\lvert A_0(0)\rvert^2$	$(1+\cos\phi_s)e^{-\Gamma_L^{(s)}t}+$ $(1-\cos\phi_s)e^{-\Gamma^s_H t}+$ $2e^{-\Gamma_s t}\sin(\Delta M_s t)\sin\phi_s$	
2	$\lvert A_\parallel(t)\rvert^2$		$(1+\cos^2\theta_1)\sin^2\theta_2-\sin^2\theta_1\sin^2\theta_2\cos 2\phi$
	$\frac{1}{2}\lvert A_\parallel(0)\rvert^2$	$(1+\cos\phi_s)e^{-\Gamma_L^{(s)}t}+$ $(1-\cos\phi_s)e^{-\Gamma^s_H t}+$ $2e^{-\Gamma_s t}\sin(\Delta M_s t)\sin\phi_s$	
3	$\lvert A_\perp(t)\rvert^2$		$(1+\cos^2\theta_1)\sin^2\theta_2+\sin^2\theta_1\sin^2\theta_2\cos 2\phi$
	$\frac{1}{2}\lvert A_\perp(0)\rvert^2$	$(1-\cos\phi_s)e^{-\Gamma_L^{(s)}t}+$ $(1+\cos\phi_s)e^{-\Gamma^s_H t}-$ $2e^{-\Gamma_s t}\sin(\Delta M_s t)\sin\phi_s$	
4	$\Re\{A_0^*(t)A_\parallel(t)\}$		$-\sqrt{2}\sin 2\theta_1\sin 2\theta_2\cos\phi$
	$\frac{1}{2}\lvert A_0(0)\rvert\lvert A_\parallel(0)\rvert\cos(\delta_2-\delta_1)$	$(1+\cos\phi_s)e^{-\Gamma^s_L(t)}+$ $(1-\cos\phi_s)e^{-\Gamma_H^{(s)}t}+$ $2e^{-\Gamma_s t}\sin(\Delta M_s t)\sin\phi_s$	
5	$\Im\{A_\parallel^*(t)A_\perp(t)\}$		$2\sin^2\theta_1\sin^2\theta_2\sin 2\phi$
	$\lvert A_\parallel(0)\rvert\lvert A_\perp(0)\rvert$	$e^{-\Gamma_s t}\{\sin\delta_1\cos(\Delta M_s t)-$ $\cos\delta_1\sin(\Delta M_s t)\cos\phi_s\}-$ $\frac{1}{2}\left(e^{-\Gamma_H^{(s)}t}-e^{-\Gamma_L^{(s)}t}\right)\cos\delta_1\sin\phi_s$	
6	$\Im\{A_0^*(t)A_\perp(t)\}$		$\sqrt{2}\sin 2\theta_1\sin 2\theta_2\sin\phi$
	$\lvert A_0(0)\rvert\lvert A_\perp(0)\rvert$	$e^{-\Gamma_s t}\{\sin\delta_2\cos(\Delta M_s t)-$ $\cos\delta_2\sin(\Delta M_s t)\cos\phi_s\}-$ $\frac{1}{2}\left(e^{-\Gamma_H^{(s)}t}-e^{-\Gamma_L^{(s)}t}\right)\cos\delta_2\sin\phi_s$	

Table 6.2: Tabulated components of the probability density function for the distribution of final-state decay angles for the process $\bar{B}_s \to J/\psi\,(\mu\mu)\,\phi\,(KK)$

k	$\Omega^{(k)}(t)$		$g(t)$
1	$\lvert\bar{A}_0(t)\rvert^2$		$4\sin^2\theta_1\cos^2\theta_2$
	$\frac{1}{2}\lvert A_0(0)\rvert^2$	$(1+\cos\phi_s)e^{-\Gamma_L^{(s)}t}+$ $(1-\cos\phi_s)e^{-\Gamma^s_H t}-$ $2e^{-\Gamma_s t}\sin(\Delta M_s t)\sin\phi_s$	
2	$\lvert\bar{A}_\parallel(t)\rvert^2$		$(1+\cos^2\theta_1)\sin^2\theta_2-\sin^2\theta_1\sin^2\theta_2\cos 2\phi$
	$\frac{1}{2}\lvert A_\parallel(0)\rvert^2$	$(1+\cos\phi_s)e^{-\Gamma_L^{(s)}t}+$ $(1-\cos\phi_s)e^{-\Gamma^s_H t}-$ $2e^{-\Gamma_s t}\sin(\Delta M_s t)\sin\phi_s$	
3	$\lvert\bar{A}_\perp(t)\rvert^2$		$(1+\cos^2\theta_1)\sin^2\theta_2+\sin^2\theta_1\sin^2\theta_2\cos 2\phi$
	$\frac{1}{2}\lvert A_\perp(0)\rvert^2$	$(1-\cos\phi_s)e^{-\Gamma_L^{(s)}t}+$ $(1+\cos\phi_s)e^{-\Gamma^s_H t}+$ $2e^{-\Gamma_s t}\sin(\Delta M_s t)\sin\phi_s$	
4	$\Re\{\bar{A}_0^*(t)\bar{A}_\parallel(t)\}$		$-\sqrt{2}\sin 2\theta_1\sin 2\theta_2\cos\phi$
	$\frac{1}{2}\lvert A_0(0)\rvert\lvert A_\parallel(0)\rvert\cos(\delta_2-\delta_1)$	$(1+\cos\phi_s)e^{-\Gamma^s_L(t)}+$ $(1-\cos\phi_s)e^{-\Gamma_H^{(s)}t}-$ $2e^{-\Gamma_s t}\sin(\Delta M_s t)\sin\phi_s$	
5	$\Im\{\bar{A}_\parallel^*(t)\bar{A}_\perp(t)\}$		$2\sin^2\theta_1\sin^2\theta_2\sin 2\phi$
	$-\lvert A_\parallel(0)\rvert\lvert A_\perp(0)\rvert$	$e^{-\Gamma_s t}\{\sin\delta_1\cos(\Delta M_s t)-$ $\cos\delta_1\sin(\Delta M_s t)\cos\phi_s\}+$ $\frac{1}{2}\left(e^{-\Gamma_H^{(s)}t}-e^{-\Gamma_L^{(s)}t}\right)\cos\delta_1\sin\phi_s$	
6	$\Im\{\bar{A}_0^*(t)\bar{A}_\perp(t)\}$		$\sqrt{2}\sin 2\theta_1\sin 2\theta_2\sin\phi$
	$-\lvert A_0(0)\rvert\lvert A_\perp(0)$	$e^{-\Gamma_s t}\{\sin\delta_2\cos(\Delta M_s t)-$ $\cos\delta_2\sin(\Delta M_s t)\cos\phi_s\}+$ $\frac{1}{2}\left(e^{-\Gamma_H^{(s)}t}-e^{-\Gamma_L^{(s)}t}\right)\cos\delta_2\sin\phi_s$	

Transversity amplitudes and strong phases

The *transversity amplitudes* and associated *strong phases* for the $B_s \to J/\psi\phi$ channel have only been measured at the CDF experiment [91]. However, the equivalent amplitudes for the analogous channel in the B_d system, $B_d \to J/\psi K^{0*}$, have been studied in more detail at four collaborations: BaBar [92], Belle [93], CDF [91] and CLEO [94]. A summary of these results is shown in tables 6.3 and 6.4. In all cases a maximum likelihood technique is used to fit the measured decay angles of the final states against a theoretical probability density function; such techniques will be explored in detail in later chapters of this work. As only two out of the three amplitudes are independent, there is some freedom in the choice of the fixed parameter and in the definition of the strong phases. For this reason the experimental results presented below use slightly different definitions to those presented in this work; a conversion is shown where appropriate.

For the B_d channel, the results are broadly consistent: in each case the amplitude A_0 dominates the process. At the time of writing there is little data on the B_s channel - in fact to date only the CDF paper [91] has presented results on the strengths of the amplitudes. However a number of phenomenologists (e.g. [95]) have suggested that, given that the two processes differ only by the flavour of one spectator quark, the amplitudes governing the processes at time $t = 0$ should be *identical* if one assumes exact flavour (or "$SU(3)$") symmetry [1]. If one accepts this hypothesis then the measurement of those parameters from the B_s channel, which cannot be accessed by the B_d system, is simplified considerably as the total number of unknown parameters is reduced. The tables above would suggest that there is at least some commonality - the A_0 component certainly dominates in both channels - but with the current data it is clearly not possible to state beyond peradventure that an exact or even approximate symmetry exists. In this work the approach of fixing the amplitudes to the B_d channel values according to exact $SU(3)$ symmetry is rejected; no such assumption is made and the amplitudes are allowed to float as free parameters of the model.

Decay widths

The decay widths of the two B_s states have been studied at Delphi [100], CDF [109][110][91], Aleph [111][112] and Opal [113]. Exact $SU(3)$ symmetry would imply $\tau_{B_s} = \tau_{B_d}$ (see for example [95]), and this constraint has been used in some of the analyses. The results (excluding the 2004 CDF measurement [91]) are summarized

[1] This is not the same as the $SU(3)$ symmetry of the Standard Model

Table 6.3: Current experimental limits on the transversity amplitudes for the process $B_d^0 \to J/\psi K^{0*}$; statistical errors first, systematic errors second. In the final two columns, the equivalent δ_1 and δ_2 are given (calculated from equation 6.14), these being the definitions of the strong phases used by ATLAS and throughout this work. Assuming $\arg(A_0) = 0.0$, $\delta_2 = \arg(A_\perp)$

	$\|A_0\|$	$\|A_\perp\|$	$\|A_\parallel\|$	$\arg(A_\parallel)$	$\arg(A_\perp) = \delta_2$
BaBar [92]	0.77 $\pm 0.03 \pm 0.02$	0.40 $\pm 0.03 \pm 0.01$	- -	2.50 $\pm 0.20 \pm 0.08$ $\delta_1 = -2.67$ $\pm 0.26 \pm 0.11$	-0.17 $\pm 0.16 \pm 0.07$
Belle [93]	0.79 $\pm 0.02 \pm 0.03$	0.44 $\pm 0.02 \pm 0.03$	- -	2.83 $\pm 0.19 \pm 0.08$ $\delta_1 = -2.92$ $\pm 0.23 \pm 0.10$	-0.09 $\pm 0.13 \pm 0.06$
CDF [91]	0.750 $\pm 0.017 \pm 0.012$	0.464 $\pm 0.035 \pm 0.007$	0.473 $\pm 0.034 \pm 0.006$	2.86 $\pm 0.22 \pm 0.07$ $\delta_1 = -2.71$ $\pm 0.27 \pm 0.08$	0.15 $\pm 0.15 \pm 0.04$
CLEO [94]	0.72 $\pm 0.07 \pm 0.04$	0.40 $\pm 0.08 \pm 0.04$	- -	3.00 $\pm 0.37 \pm 0.04$ $\delta_1 = -3.11$ $\pm 0.59 \pm 0.05$	-0.11 $\pm 0.46 \pm 0.03$

Table 6.4: Current experimental limits on the transversity amplitudes for the process $B_s^0 \to J/\psi\phi$; statistical errors first, systematic errors second.

	$\|A_0\|$	$\|A_\perp\|$	$\|A_\parallel\|$	$\arg(A_\parallel)$	$\arg(A_\perp)$
CDF [91]	0.784 $\pm 0.039 \pm 0.007$	0.354 $\pm 0.098 \pm 0.003$	0.510 $\pm 0.082 \pm 0.013$	1.94 $\pm 0.36 \pm 0.03$	- -

by the Particle Data Group [107] giving the following:

- Mean lifetime $\tau_{B_s} = (1.461 \pm 0.057) \times 10^{-12} s$
- Ratio of the width difference and the average width $\Delta\Gamma_s/\Gamma_s < 0.54$ without external constraints
- Ratio of the width difference and the average width $\Delta\Gamma_s/\Gamma_s < 0.29$ with the constraint $1/\Gamma_s = \tau_{B_d}$

Masses and the mixing parameter

Attempts at measuring the masses of the two eigenstates m_L and m_H (which is in effect the measuring of the frequency of the oscillations) have been made by a number of experiments, including ALEPH [96][97], CDF [98], DELPHI [99][100][101], OPAL [102][103] and SLD [104][105]. All of these analyses use two flavour tags - a "production" tag, which attempts to capture the flavour of the meson at time $t = 0$, and a "decay tag", which is the measurement of the flavour at the time of decay. A summary of these results is presented by the Particle Data Group [106], giving the following results:

- Mean mass $m_{B_s} = 5367.5 \pm 1.8 MeV$
- Mass difference $\Delta M_{B_s^0} = (114.07^{+2.76}_{-1.38} \pm 0.46) \times 10^{-10} MeV$ [2]
- Mixing parameter $x_s > 19.9$

Weak mixing phase

The weak mixing phase ϕ_s has never been measured due to insufficient statistics. In the Wolfenstein approximation of the CKM matrix (described in chapter 2) this quantity appears in the next-to-leading order terms of the expansion and is hence predicted to be small by the Standard Model. It could, however, be significantly enhanced by new physics contributions.

Choosing a parameter set

In order to carry out simulations of these decays, a set or a number of sets of parameters must be selected. Ideally one would wish to explore the whole parameter space, but this is clearly not possible due to constraints of time and computing capacity. Instead, three sample parameter sets have been constructed, assuming

[2] This is from the latest CDF measurement [108], which is not yet included in the PDG tables.

different levels of new physics contributions. Firstly we re-state the definition of the parameter R, which gives the ratio of new physics contributions to those of the Standard Model:

$$R = \frac{M_{12}^{NP}}{M_{12}^{SM}} \tag{6.18}$$

where M denotes the $B^0 - \bar{B}^0$ transition matrix, as per chapter 2. To produce the parameter sets we select three values of $|R| = \{0.0, 0.5, 1.0\}$, corresponding to the cases where there is no new physics, and new physics contributions equal to 50% and 100% of the Standard Model. There is no theoretical or experimental limits on the phase of R [87] and so the choice of this is somewhat arbitrary. The value $\arg(R) = \pi/4$ has been chosen to allow a moderate contribution by new physics, as can be seen from figures 2.6 and 2.7. The parameter sets can then be generated using equations 2.66 through to 2.71. These are displayed in table 6.5. Note that the system of units is "natural" (e.g. $c = \hbar = 1$), and the use of the slightly odd units of time, meV^{-1}, is intended to bring the parameters into "anthropomorphic" dimensions (that is to say, of the order of 1).

Table 6.5: Selected parameter sets based on the Standard Model and new physics contributions equal to 50% and 100% of the Standard Model

Quantity	$\|R\| = 0.0$	$\|R\| = 0.5$	$\|R\| = 1.0$
A_0	0.783	0.783	0.783
$A_\perp$	0.354	0.354	0.354
$A_\parallel$	0.510	0.510	0.510
Γ_L	$0.515meV$	$0.510meV$	$0.500meV$
Γ_H	$0.385meV$	$0.390meV$	$0.400meV$
Γ_s	$0.450meV$	$0.450meV$	$0.450meV$
$\Delta\Gamma_s/\Gamma_s$	0.29	0.26	0.22
ΔM_s	$16.45meV$	$25.15meV$	$30.43meV$
x_s	36.6	55.9	67.6
ϕ_s	0.04	0.48	0.39
δ_1	0.0	0.0	0.0
δ_2	π	π	π

6.1.4 Monte Carlo inspection of the theoretical distributions

In a number of studies, sets of angles and decay times were generated with accept-reject Monte Carlo programs, using the formulae presented in tables 6.1 and 6.2 as the probability density functions. In the first study, 500000 events were produced, half using tables 6.1 (corresponding to the B^0 state) and the remainder using the Hermitian conjugate expression in table 6.2 corresponding to the $\bar{B}^0$. Figures 6.3 and 6.4 show the two distributions for the angles θ_1 and θ_2. Clearly there is no statistically significant difference between the two. Figure 6.5 shows the distributions for the ϕ angle. The phase shift can clearly be seen, demonstrating a distinct difference between the decay of the state and the anti-state. Both plots use the Standard Model dataset.

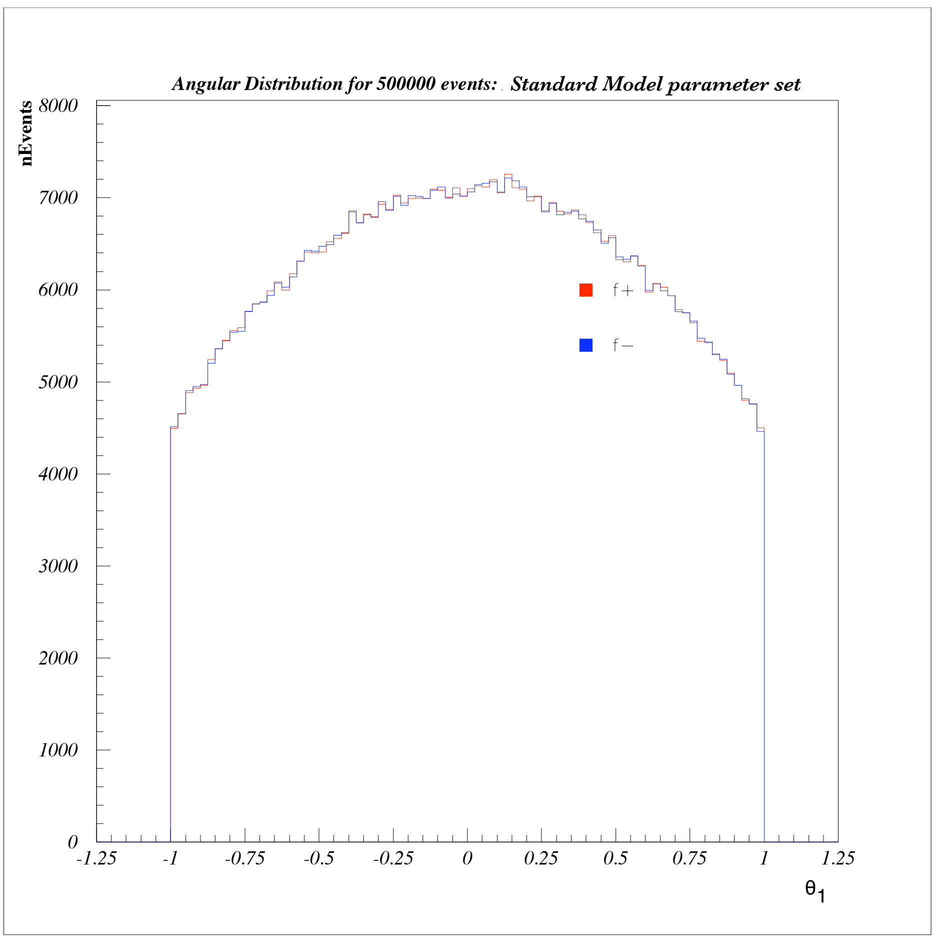


Figure 6.3: Angular distribution for $\cos(\theta_1)$, Standard Model, showing contributions from state and anti-state

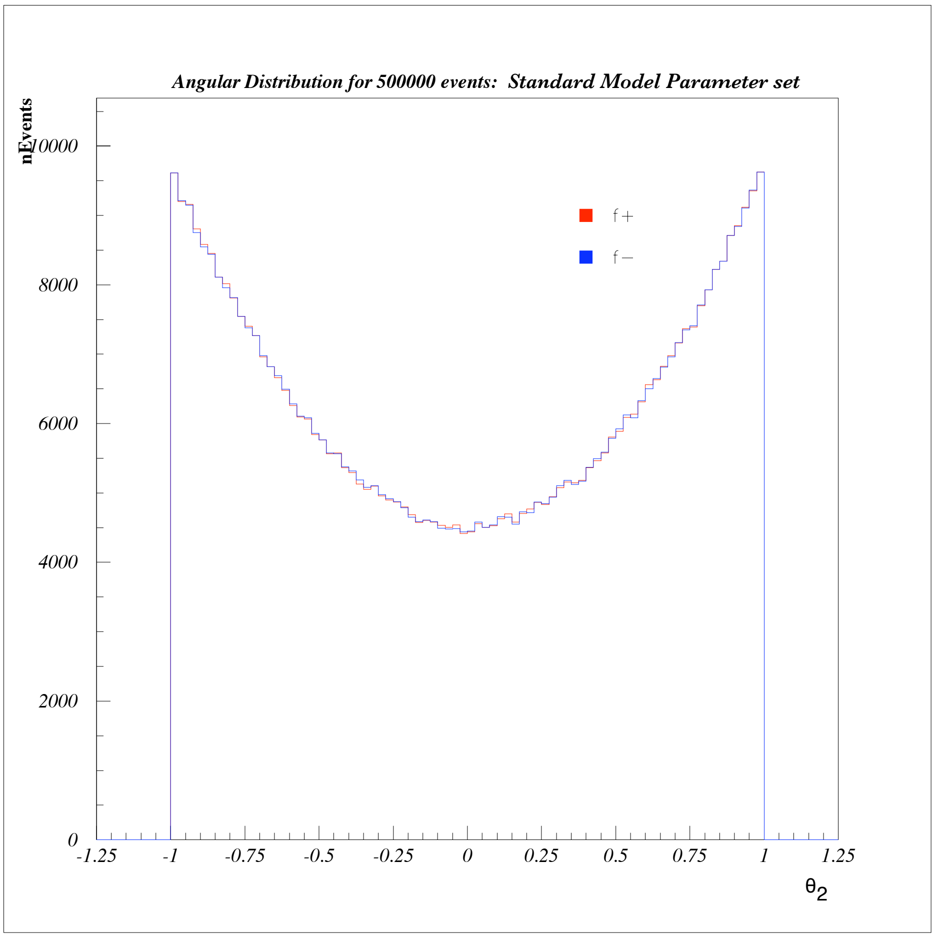


Figure 6.4: Angular distribution for $\cos(\theta_2)$, Standard Model, showing contributions from state and anti-state

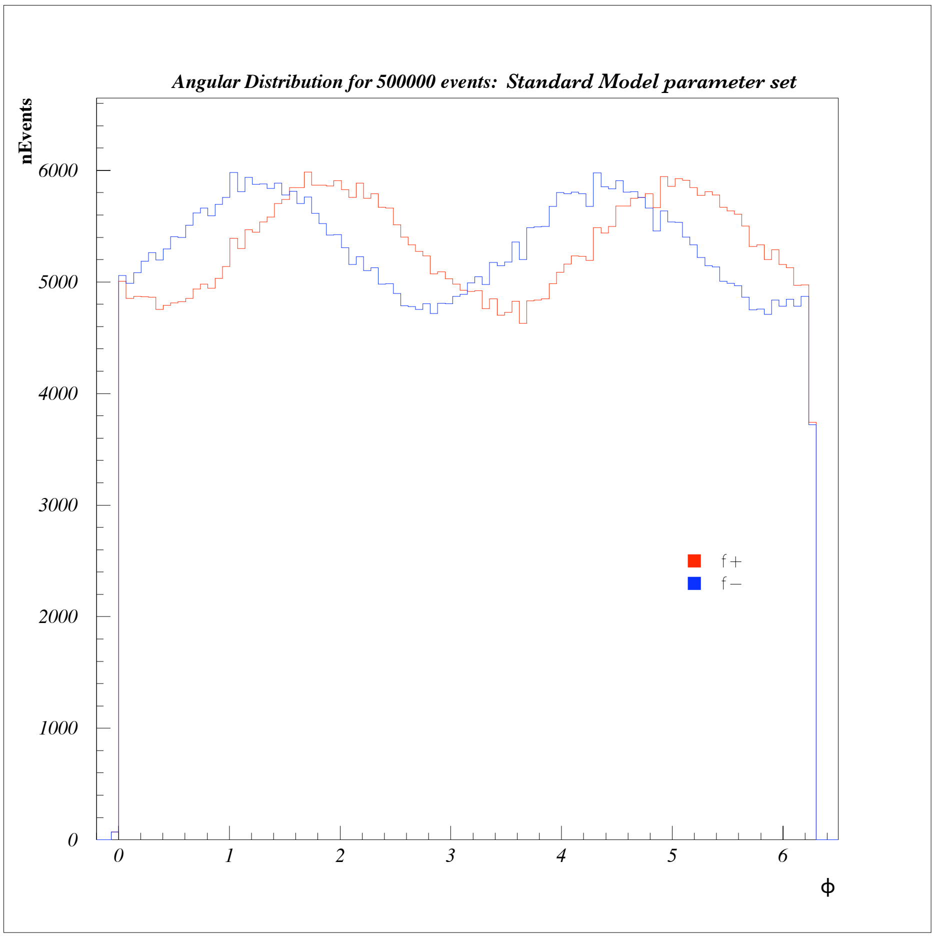


Figure 6.5: Angular distribution for ϕ, Standard Model, showing contributions from state and anti-state

In the second study, 1000000 events were generated using first the Standard Model parameter set, second for the case where $|R| = 0.5$, and third for $|R| = 1.0$. The plots are shown in figures 6.6 through to 6.9. It can be seen that the new physics has a very small effect on the time-integrated angular distributions, but a strong effect on the distribution of the proper decay times. Figure 6.9 shows in particular that increasing the weak phase ϕ_S causes the amplitude of the proper decay time oscillations to rise dramatically. All of these plots were produced with a probability density function given by table 6.1.

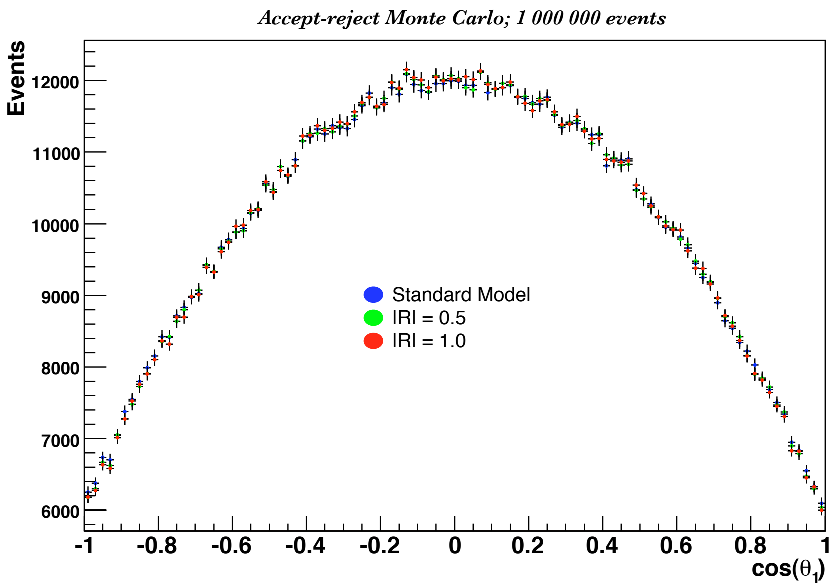


Figure 6.6: Angular distribution for $\cos(\theta_1)$

The most interesting general feature of the distributions is the ϕ plot 6.8, ϕ being the angle between the decay planes of the final states. An undulating pattern, typical of an interference effect, can clearly be seen. This appears as a result of interference between the different helicity states into which the decay amplitudes are decomposed. Note that this feature is not related to CP-violation or other weak force effects - it arises purely through spin dynamics.

6.2 Modelling the decays

We must now consider the means of simulating the decays of these neutral mesons in a manner which correctly predicts the angular distribution of the decay products, as well as the kinematics of the hard processes and the hadronization of the beauty

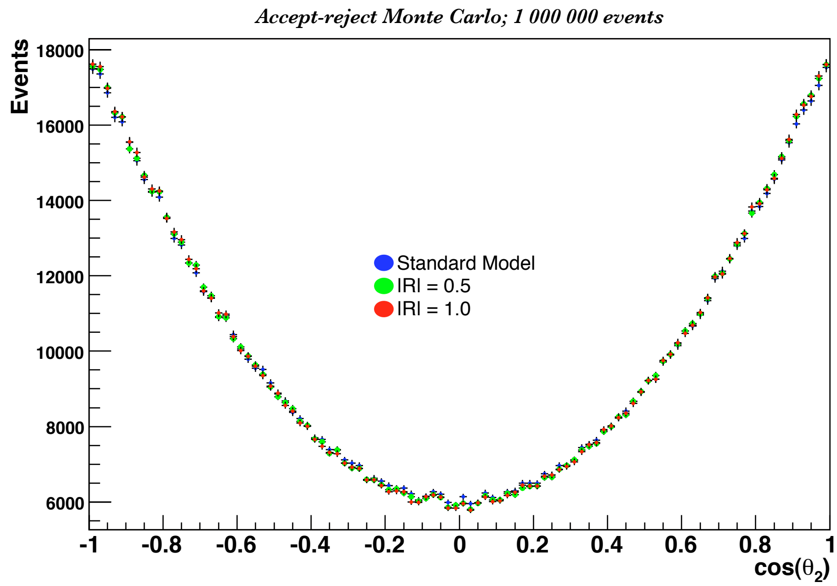


Figure 6.7: Angular distribution for $\cos(\theta_2)$

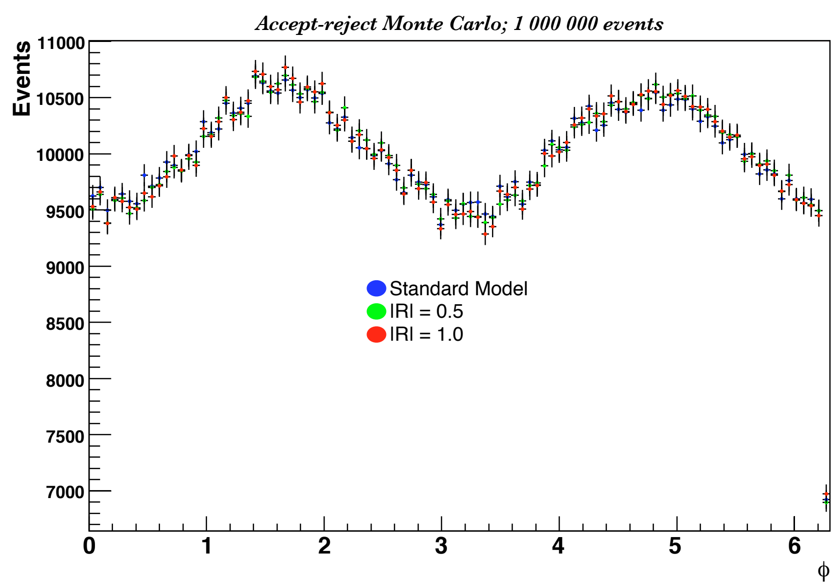
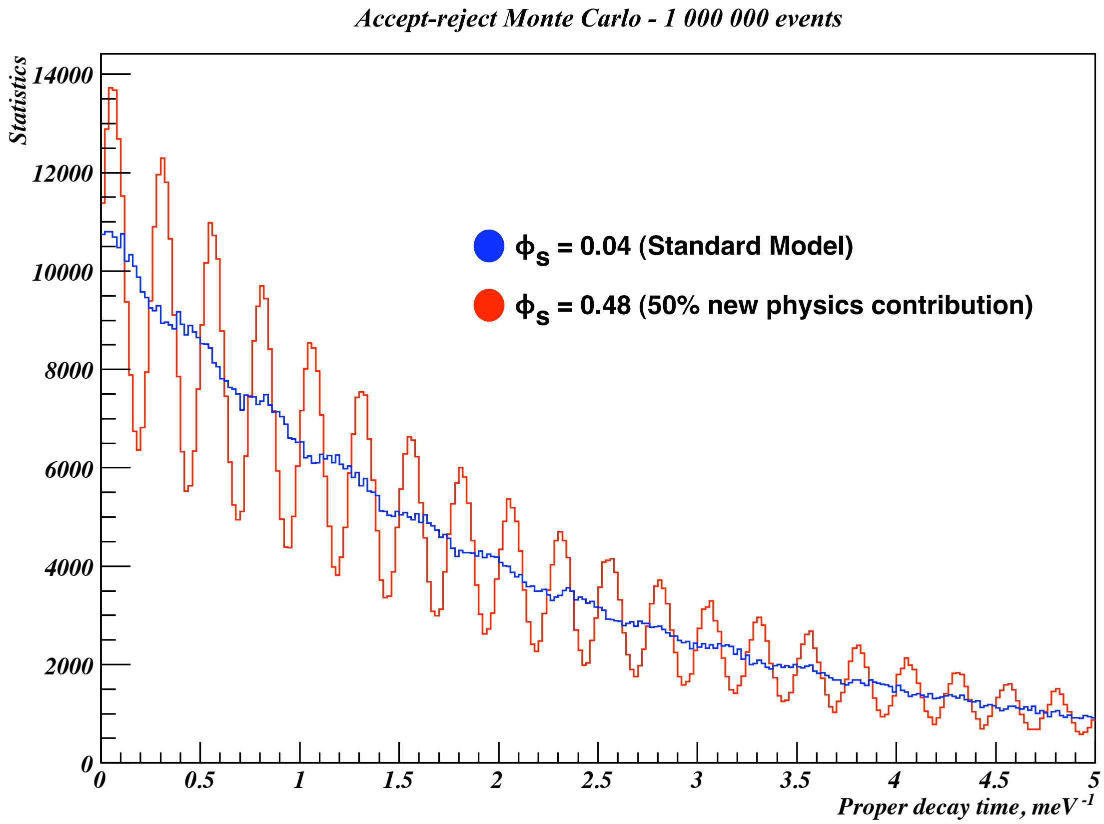


Figure 6.8: Angular distribution for ϕ



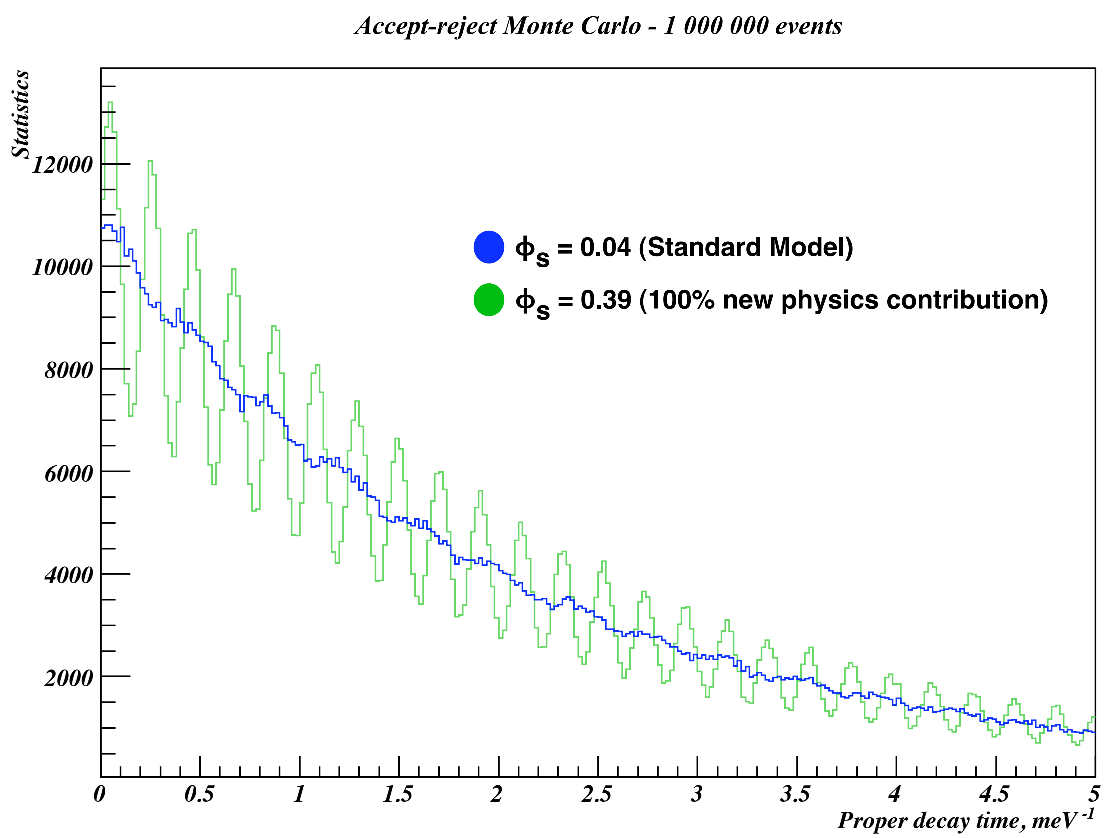


Figure 6.9: Distribution of proper decay times

quarks. As was described in chapter 5, the "work-horse" event generator is Pythia, and this is thought to give a satisfactory result for the processes up to and including the creation of the B-meson. Beyond this, however, Pythia is not sufficient. The essence of these decays is in the angular distributions, which arise from interference between helicity amplitudes. However Pythia has neither the capability to handle spin dynamics, nor the ability to deal with complex amplitudes. It uses pre-loaded decay probabilities for every decay, on which it bases the Monte Carlo procedure. In consequence any study of these decays with Pythia alone would yield flat decay angle distributions - all interference effects would be ignored. Pythia is therefore not a feasible tool with which to assess the experimental potential for this channel.

The EvtGen package, also described in chapter 5, has exactly the features required for modeling these decays. It uses representations of the particles that allow it to capture the physics of decays with non-trivial spin, and the decay engine accepts complex amplitudes rather than probabilities. If an amplitude is passed as a sum of component amplitudes, these will be mod-squared correctly, such that the final probability used in the Monte Carlo algorithm will contain all the interference terms. EvtGen is therefore the tool required for this simulation task.

EvtGen does not, however, have the capacity to simulate proton-proton collisions and hadronization; these tasks must still be done by Pythia. EvtGen takes over at the point where the B-mesons have been produced. These interfaces between event generators were described in the computing chapter.

The implementation of this specific decay process in EvtGen is now described.

6.2.1 Construction and validation of an EvtGen decay model for $B_s \to J/\psi\phi$

As has been noted in the computing chapter, the EvtGen package consists of a series of C++ classes, each of which encodes the amplitudes for a particular type of decay process. These classes are referred to as *models*. The original BaBar version of EvtGen contains a model, referred to as `EvtSVVHelAmp`, for decays of a scalar to two vectors. This model allows the user to input the magnitudes and phases of the three amplitudes in the helicity basis, but has no capacity for the simulation of mixing - only the spin physics of the process is captured. A detailed description of the model is given in pages 102, 119 and 120 of [54].

To test the accuracy of the `EvtSVVHelAmp` model, 50000 B_s mesons were generated with the repeat-hadronization PythiaB package (described in the computing chapter) and decayed with the EvtGen package using `EvtSVVHelAmp` as the decay model. Helicity amplitudes were calculated from the "Standard Model" transversity amplitudes as in table 6.5 using equation 3.59 and were used as the sole input to the model. No detector-related kinematic cuts were made on the events.

The decay angles were calculated from the final state muons and kaons produced by EvtGen, and were plotted against a set of angles produced by accept-reject Monte Carlo using the formulae in table 6.1 as the probability density function. Parameters related to weak effects (ΔM_s, ϕ_s) were set to zero in this accept-reject implementation, and all decay widths were set to the mean decay width Γ_s, to mimic the lack of weak physics in `EvtSVVHelAmp`. Figures 6.10, 6.11 and 6.12 show the angular distributions for each of the angles, normalized to the same scale. As can be seen, the software provides a close match to the theoretical Monte Carlo.

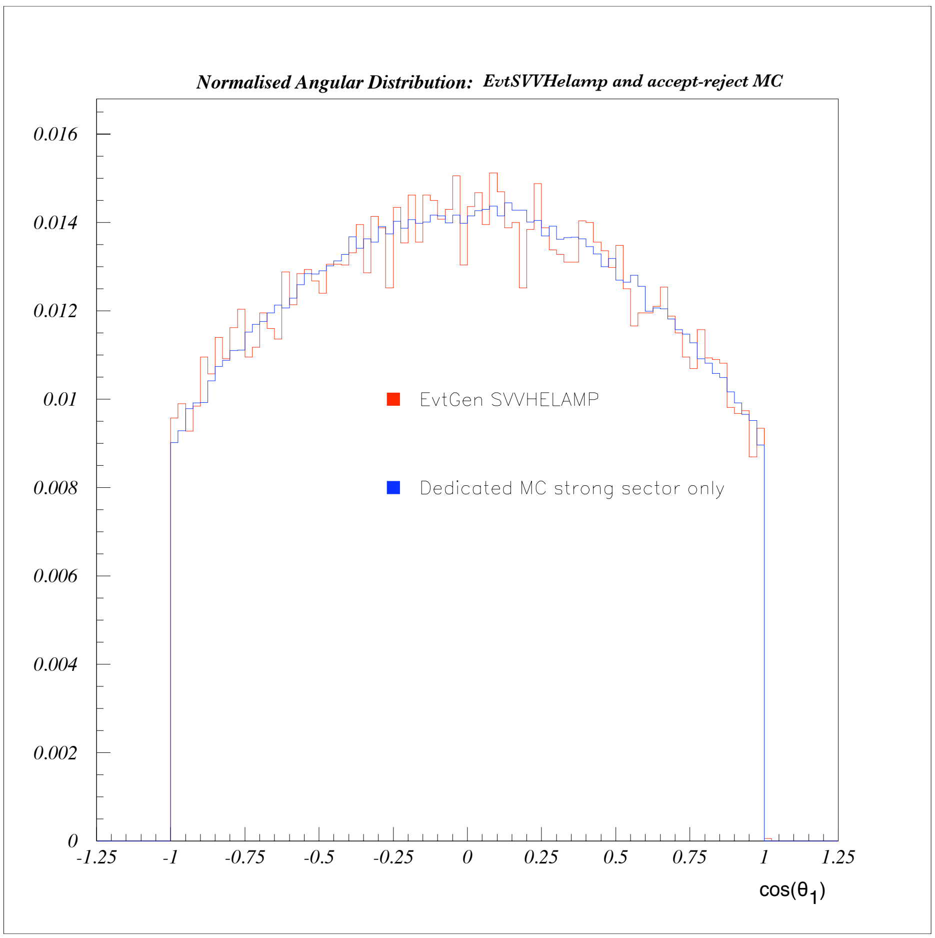


Figure 6.10: Comparison between EvtGen and dedicated Monte Carlo for strong sector effects, angle θ_1

In a second test, 50000 $\bar{B}_s$ mesons were generated and processed by identical code. The expectation was that no difference would be observed, the model containing no weak physics. This is confirmed by figures 6.13, 6.14 and 6.15. The fourth

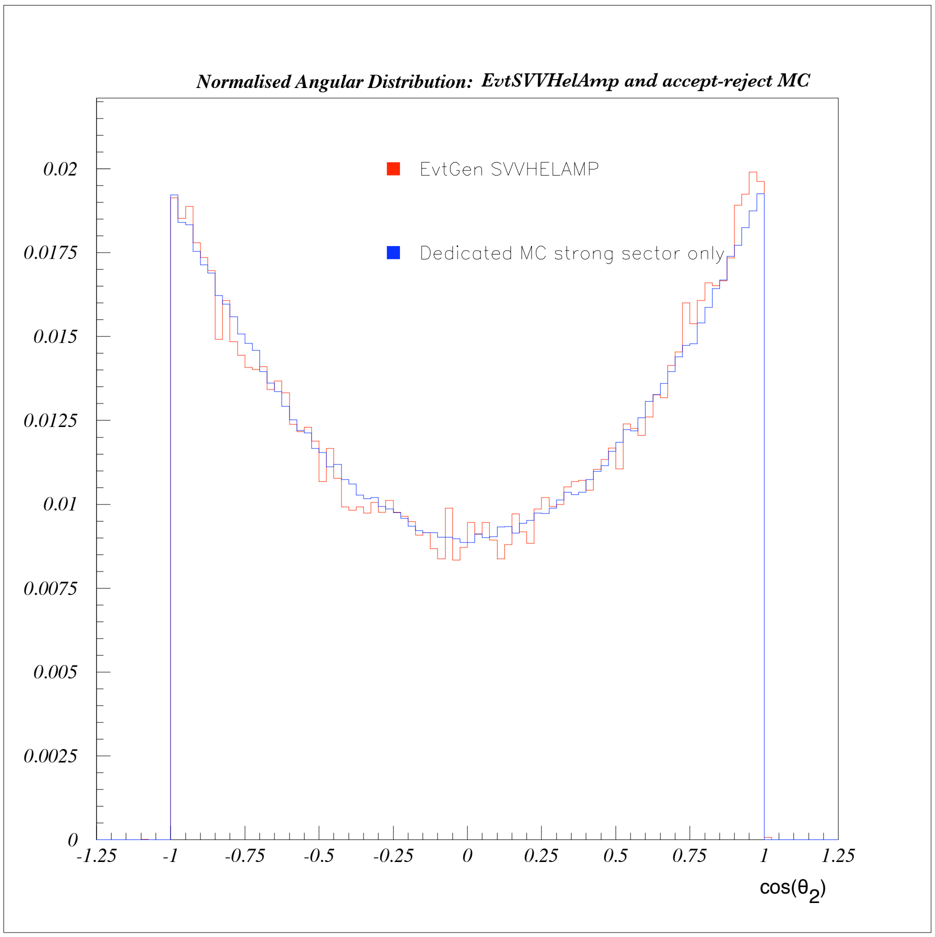


Figure 6.11: Comparison between EvtGen and dedicated Monte Carlo for strong sector effects, angle θ_2

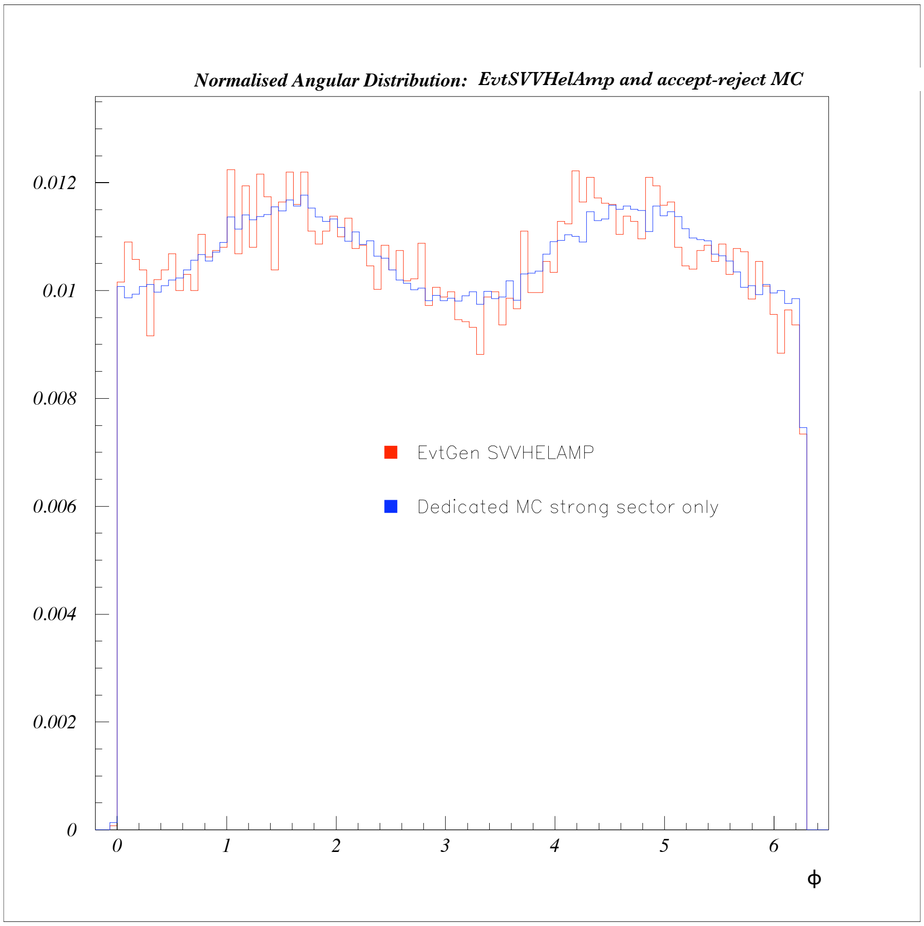


Figure 6.12: Comparison between EvtGen and dedicated Monte Carlo for strong sector effects, angle ϕ

plot 6.16 shows the ϕ distribution from `EvtSVVHelAmp` against the distributions from the accept-reject study, with the state and anti-state contributions separated. As expected, the distribution produced by EvtGen, being unaffected by weak physics, occupies an unshifted position whilst the other two are shifted to the left and the right.

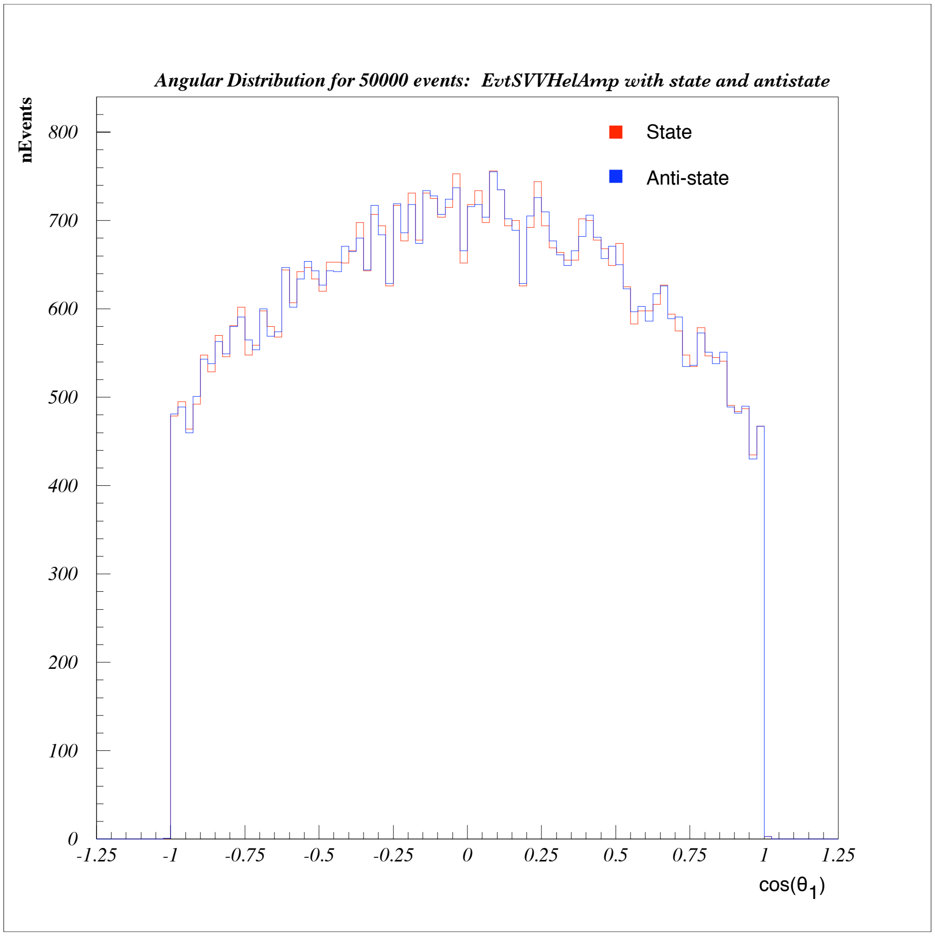


Figure 6.13: Demonstration of absence of weak effects in EvtSVVHelAmp, angle θ_1

On the strength of these results and after studying the implementation details, it was decided to use `EvtSVVHelAmp` as the base for the full mixing model.

A number of $S \rightarrow VV$ models already exist in EvtGen for simulating mixing processes (see pp 99-102 [54]) but these were not considered to be adequate, due firstly to the restrictive nature of the input parameters (unitarity triangle angles rather than physical phases), and secondly due to the use of mixing probabilities in the model. Such a probability is used by the code to "flip" the flavour of the meson as a function of time. Such a treatment would again wash out the interference terms with which this study is primarily concerned. As a consequence of these

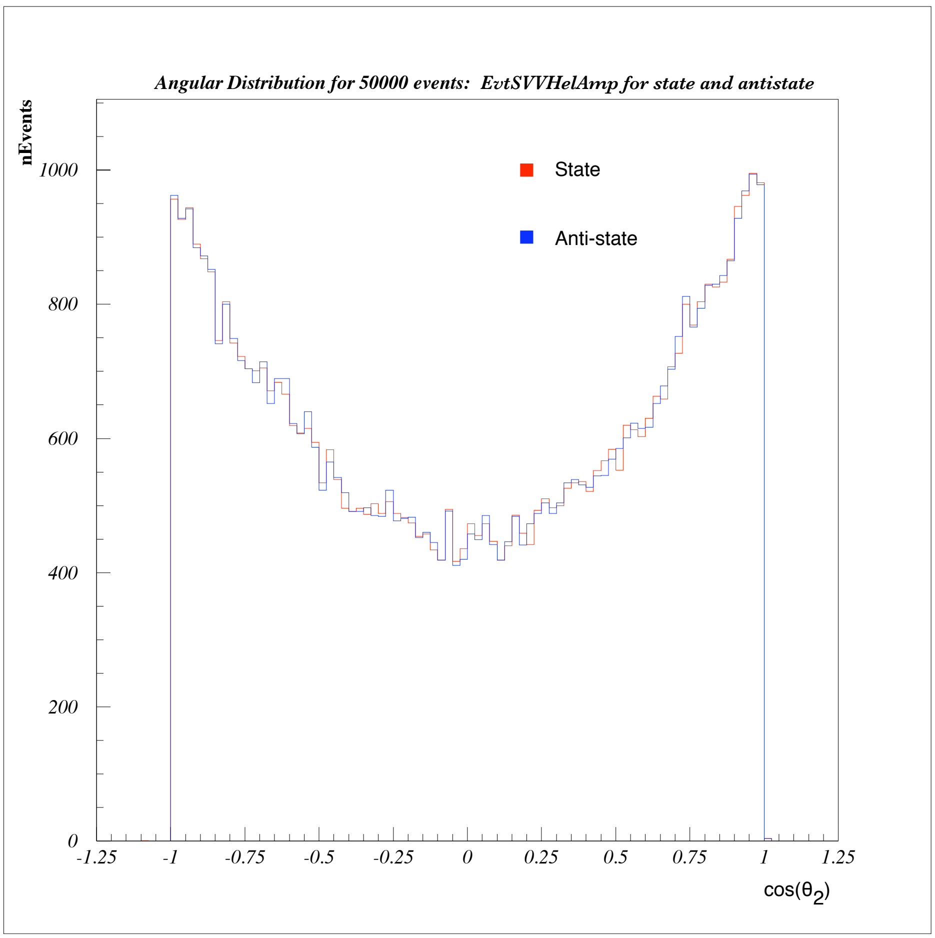


Figure 6.14: Demonstration of absence of weak effects in EvtSVVHelAmp, angle θ_2

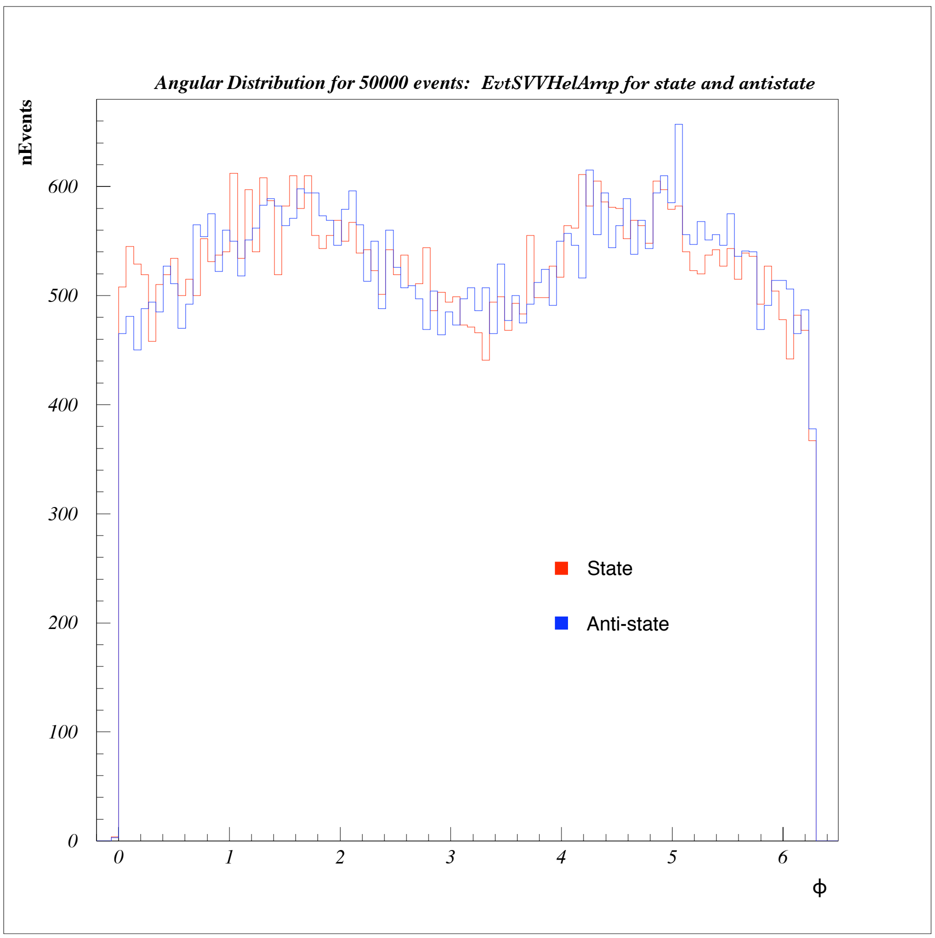


Figure 6.15: Demonstration of absence of weak effects in EvtSVVHelAmp, angle ϕ

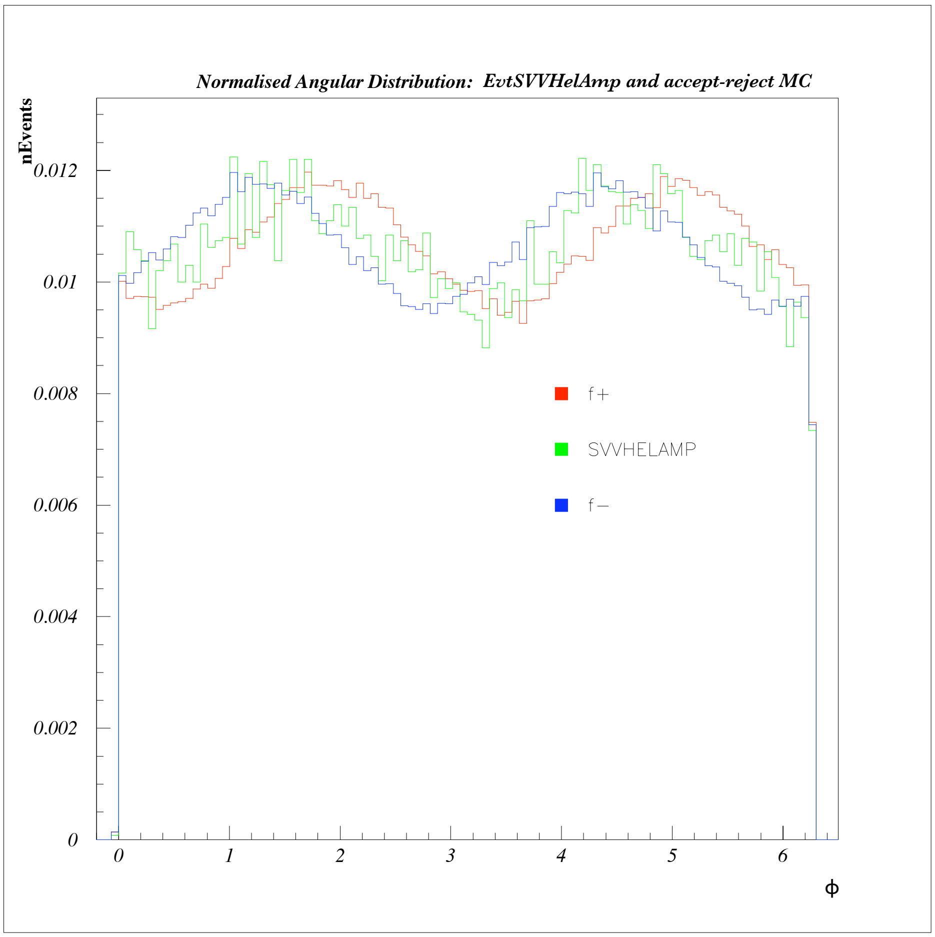


Figure 6.16: Additional demonstration of absence of weak effects in EvtSVVHelAmp, angle ϕ

considerations, a completely new EvtGen model was developed using `EvtSVVHelAmp` as the base.

Contents of the model

The requirement is to extend the existing `EvtSVVHelAmp` model to include mixing. This model takes the values of the helicity amplitudes as arguments, and then forms a transversity basis from these; the three amplitudes are then passed to the Monte Carlo engine. The new model (referred to, for want of a better name, as `EvtSVVHelCPMix`) is an extension of this - onto the amplitudes are attached the mixing parameters, according to the formalism presented above, in [86] and in chapter 2. Two cases are provided in the model - the first for a particle labelled by Pythia as a B_s^0 meson, and the second for a meson labelled as a $\bar{B}_s^0$. The decision as to which amplitude to use is made in the model, on the basis of the particle identification number in the HepMC record. In the first case the final amplitude passed to the EvtGen engine is of the form given in 6.8, whereas in the second case the form given in 6.9 is used. In each case the final amplitude is in fact a sum; the evaluation of this is automatically handled by EvtGen. The inputs to the model, provided through the user decay file, are as follows:

- The magnitudes of the helicity amplitudes H_0, H_+ and H_-
- The strong phases associated with the amplitudes δ_0, δ_+, δ_-
- Γ_s and $\Delta\Gamma_s$
- M_s and ΔM_s
- ϕ_s

The helicity basis, rather than the transversity basis, is used for the user input to remain in accord with the original model, and for similar reasons all three magnitudes and phases are provided. It is the responsibility of the user to ensure that the values are sensible, that is to say, normalized to one.

The model was found to give a close match to the directly produced accept-reject Monte Carlo events using the full probability density function in table 6.1, and was accepted into the ATLAS version of the EvtGen package.

6.3 Event Generation

6.3.1 Signal events

Within the ATHENA framework, Pythia (run via the PythiaB algorithm) was used to simulate the proton collisions. All $b\bar{b}$ quark pairs were repeatedly hadronized to increase efficiency, as described in chapter 5. All B-mesons produced in this way were then passed to EvtGen. $\bar{B}$-mesons were allowed to decay "naturally" according to Pythia tables.

Using the new `EvtSVVHelCPMix` model described above, EvtGen was used to decay all B_s mesons into the final state $\mu\mu KK$ via the intermediate state $J/\psi\phi$. Other B-mesons were not decayed, and were treated as stable. Finally, the BSignalFilter algorithm was used to remove all the undecayed B-mesons and any events not meeting the trigger requirements. The remaining events were then written to disk in the POOL format.

Using the LHC Computing Grid, approximately 200000 events were produced in this way. Table 6.6 shows the parameters used to generate the events.

Inspection of the expressions in tables 6.1 and 6.2 immediately reveals that *both* the B_s^0 and the $\bar{B}_s^0$ theoretical distributions are dependent on all the parameters, so without loss of generality it is possible to generate just one initial flavour of B-meson (in this case, the B_s^0), and still extract all parameters. Flavour tagging efficiencies can still be studied, under the assumption that the wrong-tag fraction for the $\bar{B}_s^0$ is the same as that for the B_s^0. Running through all this is the presumption of zero production asymmetry - this is not considered in this analysis.

6.3.2 Background events

Previous studies [115][117] have found that two background processes are significant for this channel and so were produced for this study; the exclusive channel (background I) $B_d \to J/\psi\,(\mu\mu)\,K^{0*}\,(K^+\pi^-)$ and the inclusive channel (background II) $b\bar{b} \to J/\psi X$, where X is anything permitted by the decay tables. The exclusive sample, being a $S \to VV$ decay with an expected angular structure in the decay products, was decayed with the default helicity model in EvtGen (`EvtSVVHelAmp`). Mixing effects were not simulated. The inclusive sample was produced with Pythia only (run in PythiaB), and used decay tables constructed as described in [114]. The LCG was again used for this production. The parameters for each channel are shown below.

Table 6.6: Signal event generation parameters

Parameter	Value
b-quark cuts $\bar{b}$-quarks cuts	None $p_T > 10.0 GeV$, $\eta < 2.5$
No. hadronization loops per $\bar{b}$-quark	14
Decay model $\lvert H_0 \rvert$ $\lvert H_+ \rvert$ $\lvert H_- \rvert$ δ_0 δ_+ δ_- M_s ΔM_s Γ_s $\Delta\Gamma_s$ ϕ_s	EvtGen `EvtSVVHelCPMix` 0.721 0.683 0.114 π 0.0 0.0 $5.370 GeV$ $16.45 \times 10^{-12} GeV$ $4.5 \times 10^{-13} GeV^{-1}$ $1.3 \times 10^{-13} GeV^{-1}$ 0.04
BSignalFilter event acceptance criteria	1 muon $p_T > 6 GeV$, $\eta < 2.5$ 1 additional muon $p_T > 3 GeV$, $\eta < 2.5$ 1 hadron $p_T > 0.5 GeV$, $\eta < 2.5$ All unstable particles decayed

Table 6.7: Background I event generation parameters

Parameter	Value
b-quark cuts $\bar{b}$-quarks cuts	None $p_T > 10.0GeV$, $\eta < 2.5$
No. hadronization loops per $\bar{b}$-quark	14
Decay model $\|H_0\|$ $\|H_+\|$ $\|H_-\|$ δ_0 δ_+ δ_-	EvtGen `EvtSVVHelAmp` 0.721 0.683 0.114 0.0 0.0 0.0
PythiaB event acceptance criteria	1 muon $p_T > 6GeV$, $\eta < 2.5$ 1 additional muon $p_T > 3GeV$, $\eta < 2.5$ 1 hadron $p_T > 0.5GeV$, $\eta < 2.5$ All unstable particles decayed

Table 6.8: Background II event generation parameters

Parameter	Value
b-quark cuts $\bar{b}$-quarks cuts	None $p_T > 10.0GeV$, $\eta < 2.5$
No. hadronization loops per $\bar{b}$-quark	14
Decay model	Pythia [114]
PythiaB event acceptance criteria	1 muon $p_T > 6GeV$, $\eta < 2.5$ 1 additional muon $p_T > 3GeV$, $\eta < 2.5$ 1 hadron $p_T > 0.5GeV$, $\eta < 2.5$
BSignalFilter event acceptance criteria	All unstable particles decayed

6.3.3 Generation results

Signal events

The production cross section for the B_s meson reported by PythiaB was $0.011mb$. As only those events containing a $B_s^0 \rightarrow J/\psi\,(\mu\mu)\,\phi\,(KK)$ decay were passed into the final sample, this figure had to be multiplied by

- 9.300×10^{-4}, the branching ratio for $B_s^0 \rightarrow J/\psi\phi$
- 0.059, the branching ratio for $J/\psi \rightarrow \mu\mu$
- 0.491, the branching ratio for $\phi \rightarrow KK$
- 0.035, the proportion of events passing the kinematic cuts on the muons and final state hadrons.
- A factor of 2 to account for the opposite side quark which is allowed to decay freely

The final cross section was therefore $2.130 \times 10^{-5}\mu b$. The total number of events passing the filter and written to disk was 190722.

Figure 6.17 shows the time-integrated angular distribution (for the ϕ angle) for these events, compared directly with an accept-reject Monte Carlo data set.

Background I results

The production cross section for the B_d meson given by PythiaB was $0.008mb$. As only those events containing a $B_d^0 \rightarrow J/\psi\,(\mu\mu)\,K^{0*}\,(K^+\pi^-)$ decay were passed into the final sample, this figure was multiplied by

- 1.310×10^{-3}, the branching ratio for $B_d^0 \rightarrow J/\psi K^{0*}$
- 0.059, the branching ratio for $J/\psi \rightarrow \mu\mu$
- 1.0, the approximate branching ratio for $K^{0*} \rightarrow K^+\pi^-$
- 0.142, the proportion of events passing the kinematic cuts on the muons and final state hadrons.
- A factor of 2 to account for the opposite side quark

The final cross section was therefore $1.789 \times 10^{-4}\mu b$. The total number of events passing the filter and written to disk was 324600, this channel being part of the effort for the 2005 Rome Physics Workshop.

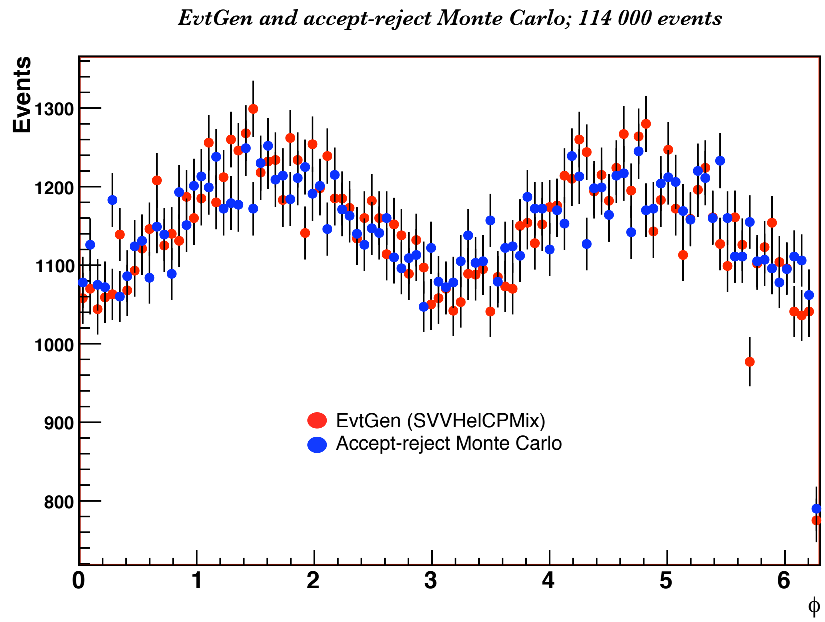


Figure 6.17: EvtGen (`EvtSVVHelCPMix`) and accept reject Monte Carlo; Standard Model parameter set

Background II results

The cross section from PythiaB for the inclusive decays $b\bar{b} \rightarrow J/\psi\,(\mu\mu)\,X$ was $1.34 \times 10^{-3}\mu b$. The total number of events written to disk was 1000000, this being part of effort for the 2005 Rome Physics Workshop.

Chapter 7

Analysis of the decay $B_s^0 \to J/\psi\phi$ II - Event Reconstruction

Following the generation, simulation and digitization processes, the resulting digit stream must be *reconstructed* such that the binary representations of simulated detector "hits", which correspond to deposition of energy by final-state particles into the detector's active components, are extrapolated into smooth loci - the particle "tracks". Following this the tracks must then be analyzed methodically to attempt to identify the signal events from the background, to determine the state of the parent B-meson, and finally to calculate the decay angles.

The software used to perform these tasks of track building and analysis was covered in the computing chapter of this work, as were the definitions of the analysis objects. This chapter follows four themes - the construction of tracks from the digits, the analysis of the tracks to determine the signal events and the capacity of the algorithms for rejecting background events, the tagging of signal events, and the calculation of the distribution of decay angles. The results laid out in this chapter were presented to the ATLAS community in preliminary form at the collaboration's physics workshop in Rome in the summer of 2005.

7.1 Track Building

As with the detector simulation, the Athena release 11.0.41 was used for the final assessment, with initial results obtained from release 10.0.4. All the fully simulated signal and background digits were processed by the full inner detector reconstruction package, using iPatrec as the track fitting package. The resulting Event Summary Data (ESD) was converted into Analysis Object Data (AOD) with further Athena packages. The AOD files, containing records of track objects produced by the fitter and the associated Monte Carlo truth, were then made available for analysis.

7.2 Track analysis

The essential information required from this analysis was as follows:

- The **track reconstruction efficiency**, being the proportion of final state tracks originating from the decay of $B_s \to J/\psi\phi$ which were successfully located in the digit stream by the reconstruction software. This is in general dependent on the momenta of the particles in question and the geometry of the machine.

- The **event reconstruction efficiency** - the proportion of successfully reconstructed signal events that were identified as such by the procedural track analysis.

- The **resolutions** in reconstructed particle mass, proper decay time and final state decay angles, for the identified signal events.

- The **background rejection efficiency**, being the proportion of background events which are successfully rejected by the analysis algorithm. In general there are two backgrounds - *real* and *combinatorial*. Real backgrounds consists of actual events which, due to similarities between them and the signal in topology or mass, can be falsely identified as signal events by the analysis algorithm. In this case there are two principal real backgrounds as identified in [115]: the exclusive channel $B_d \to J/\psi K^{0*}\,(K\pi)$ and the inclusive process $bb \to J/\psi\,(\mu\mu)\,X$, where X is any hadron (events from the exclusive B_d channel are of course a sub-set of those from the inclusive process). Combinatorial backgrounds are an unavoidable feature of the analysis algorithm, and occur where incorrect combinations of tracks (which may be from completely separate processes) nevertheless yield plausible physical features, particularly in invariant mass. Most of these "events" can be removed by attempting to fit track combinations to common vertices; those tracks which do not have an intersection will probably fail to fit. Some will inevitably succeed, so the origin of these "fakes" must be ascertained.

This information then permits an assessment, firstly, of the number of events which can be expected over time once ATLAS begins to take data and secondly, the resolutions can be fed into a simulation of an maximum likelihood fit on the decay angles, with the aim of determining the uncertainties which will be attached to measurements of the physically interesting parameters. This is deferred to the final chapter of this work.

ATLAS does not have hadron identification apparatus so the method must rely on the charge and momenta of the tracks to perform initial selections, and then make use of invariant mass and vertex position calculations to reject combinatorial and real backgrounds. ATLAS also has sophisticated muon identification apparatus and the reconstruction software is capable of matching hits in the muon chambers with tracks in the inner detector, but at the time of writing the software for simulating the action of the muon sub-detectors was still in the early stages of development and was not used.

The analysis algorithm itself was constructed as follows, and implemented as an Athena algorithm.

1. The tracks were loaded from StoreGate and the transverse momentum of each was calculated from the perigee. Those tracks with an initial p_T exceeding $3GeV$ and $|\eta| < 2.7$ were retained; all others were rejected.

2. All possible unique pairs of oppositely charged tracks were formed; those pairs containing at least one track with $p_T \geq 6.0$ were retained and all others were discarded.

3. The invariant mass of all accepted pairs was calculated, under the assumption that the tracks were from muons with a mass of $105.4MeV$. Those pairs whose invariant mass fell outside the window defined by $m_{J/\psi} \pm \sigma$, where the mass $m_{J/\psi}$ of the J/ψ meson is taken to be $3096.9MeV$ and σ was a quantity to be determined, were rejected.

4. Accepted track pairs were then fitted to a common vertex using the CTVMFT vertexing program, with the mass constrained to $m_{J/\psi}$. Where the χ^2 per degree of freedom ($\chi^2/d.o.f$) of the fit and the transverse decay length l_{xy} of the vertex passed certain cuts, the pairs were accepted. These pairs were assumed to be muons originating from a decaying J/ψ.

5. The original collection of tracks was scanned once again and those with $p_T \geq 0.5$ and $|\eta| \leq 2.7$ were accepted.

6. Unique oppositely charged pairs were once again formed and the invariant mass of each calculated, this time under the assumption that the tracks were formed by kaons ($m_K = 497.6MeV$). Those falling within a certain mass interval were accepted.

7. Each track pair was fitted to a common vertex. In this case no mass was imposed on the fit as the ϕ meson has a natural width so such a constraint would be artificial. Where the $\chi^2/d.o.f$ was less than a certain quantity which must be determined, the pairs were accepted, and were assumed to be kaons originating from a decaying ϕ meson.

8. The accepted pairs were then formed into unique quadruplets, each containing one "muon" and one "kaon" pair. Combinations where a track appears as both a "kaon" and a "muon" were discarded.

9. The tracks in each quadruplet were fitted to a common vertex (it was assumed that the decay lengths of the J/ψ and ϕ are negligible, such that the four tracks originated from one point) using a three dimensional kinematic fit. The momentum vector of the vertex was required to point at the primary vertex provided by the reconstruction software, and the two "muon" tracks were again constrained to $m_{J/\psi}$.

10. Quadruplets successfully fitted to a common vertex having p_T and $\chi^2/d.o.f$ within certain thresholds and a proper decay time $\tau > 0.5ps$ [89], were accepted and assumed to be muon and kaon pairs, originating from a B_s^0 decay with J/ψ and ϕ as the intermediate states.

11. The decay angles of the final states were calculated.

12. Using the information from the Monte Carlo truth, the resolutions for the measured decay angles, proper decay times and transverse momenta were calculated for those events confirmed as being signal processes.

13. An attempt was made to determine the flavour of the signal B_s meson.

In all cases, the cuts applied after vertexing (and the histograms allowing the calculation of the resolutions) were based on quantities formed from the *refitted* tracks formed by the vertexing algorithms, rather than those obtained directly from the reconstruction software.

7.2.1 Analysis procedure

At each stage of the analysis as itemized above, it is expected that a number of **cuts** will be applied to all track combinations, with the aim of removing real and combinatorial backgrounds and leaving only the signal events. The cuts must be methodically selected so as to maximize the signal event reconstruction efficiency whilst minimizing the number of wrongly accepted background events. This procedure must be performed in several stages.

1. Firstly, the files containing signal events must be processed in order to determine the cuts required to remove the combinatorial background, leaving only the signal events. Clearly the tracks which originated from the signal must be identified as such before the analysis can begin. This requires the use of Monte Carlo truth association apparatus, where the tracks are linked to the Monte Carlo particles which created them. Naturally this luxury will not be available with real data, so the determination of cuts is a task which can only be done with simulated events.

2. Secondly, the files containing the real background events must be scanned. In this case the combinatorial backgrounds are of no interest, so the *only* tracks which should be processed at this stage are those identified by the truth association apparatus as being from specific background processes [1]. The cuts identified in the first stage may then need to be modified to improve the rejection of real background. This operation is a compromise between rejecting as much of the background as possible, whilst throwing away as few of the signal events as can be feasibly arranged. In practice, therefore, it is usual to perform the tuning of the cuts on all samples (signal and background) simultaneously.

Only after these cuts have been determined can the signal resolutions, reconstruction efficiencies and background rejection fractions be determined.

7.2.2 Track reconstruction efficiency and selection of cuts

The samples of signal and background events were scanned by the analysis algorithm, identifying the signal (or real background) tracks for each event using the Monte Carlo association tools. It was first necessary to determine the number of the signal and exclusive background events whose final state particles were successfully reconstructed, which represents the best possible performance by the analysis code (clearly it cannot identify events whose tracks have not been reconstructed). These figures are displayed in tables 7.1 and 7.2.

The procedural track analysis was performed on all tracks for signal samples, whereas for background samples, only tracks originating from real background events were considered. The cuts, as identified in the previous section, were gradually

[1] See erratum at the end of this document

Table 7.1: Track reconstruction efficiencies for the signal process $B_s^0 \rightarrow J/\psi\,(\mu\mu)\,\phi\,(K^+K^-)$

	Number of events
Monte Carlo events	103737
Both signal $J/\psi \rightarrow \mu\mu$ tracks reconstructed	99254(95.7%)
Both signal $\phi \rightarrow KK$ tracks reconstructed	86928(83.8%)
All signal $B_s \rightarrow J/\psi\,(\mu\mu)\,\phi\,(KK)$ tracks reconstructed	86817(83.7%)

Table 7.2: Track reconstruction efficiencies for the exclusive background process $B_d^0 \rightarrow J/\psi\,(\mu\mu)\,K^{0*}\,(K^+\pi^-)$

	Number of events
Monte Carlo events	113401
Both signal $J/\psi \rightarrow \mu\mu$ tracks reconstructed	113184(99.81%)
Both signal $K^{0*} \rightarrow K\pi$ tracks reconstructed	100752(88.85%)
All signal $B_d \rightarrow J/\psi\,(\mu\mu)\,K^{0*}\,(K\pi)$ tracks reconstructed	100572(88.69%)

tightened to maximize the rejection of combinatorial and real backgrounds whilst minimizing the number of signal events rejected. The final results of this process are displayed in tables 7.3 through to 7.7. Ultimately, the optimal cuts were determined to be:

1. Oppositely charged track pairs were accepted as being from a J/ψ decay if
   - The invariant mass of the pair (under the assumption that each track is from a muon) fell within the range $2947MeV \leq M_{J/\psi} \leq 3267MeV$
   - The pair was successfully fitted to a common decay vertex, with a fit $\chi^2/d.o.f < 6$
2. Oppositely charged track pairs were accepted as being from a ϕ decay if
   - The invariant mass of the pair (under the assumption that each track was from a kaon) fell within the range $999MeV \leq M_{J/\psi} \leq 1039MeV$
   - The pair was successfully fitted to a common decay vertex, with a fit $\chi^2/d.o.f < 6$

3. Quadruplets of tracks made up of one accepted $J/\psi \to \mu\mu$ candidate and one accepted $\phi \to KK$ candidate were accepted as being from the decay of a B_s^0 meson via the $J/\psi\phi$ intermediate state if
    - They were successfully fitted to a common vertex with a fit $\chi^2/d.o.f < 5$
    - The invariant mass of the quadruplet (under the assumption that two of the tracks were muons and two were kaons) fell within the range $5250MeV \leq M_B \leq 5490MeV$
    - The proper lifetime of the meson, calculated through the vertex position, satisfied $\tau_B > 0.5ps$
    - The transverse momentum of the meson candidate satisfied $p_T > 10.5GeV$

Table 7.3: Effect of J/ψ cuts on the signal events

Cut	Signal J/ψ accepted	Signal B_s accepted
None	99254	–
Vertexing	98293	83715
$2947MeV \leq M_{J/\psi} \leq 3267MeV$	97261	82926
$\chi^2 \leq 6.0$	95783	81713

Table 7.4: Effect of J/ψ cuts on the background events. Note that the BGI component in BGII is removed prior to processing to avoid double counting.

Cut	Combinatorial J/ψ	Combinatorial B_s^0	BGI	BGII
None	529109	–	–	–
Vertexing	508881	3850673	60036	80271
$2947MeV \leq M_{J/\psi} \leq 3267MeV$	6329	3077727	59492	66327
$\chi^2 \leq 6.0$	5583	3005878	58626	64740

7.2.3 Signal acceptance and resolutions - final results

Of 103737 Monte Carlo $B_s \to J/\psi(\mu\mu)\phi(KK)$ events, 51100 were reconstructed and accepted by the analysis algorithm, giving a headline acceptance figure of 49.2%.

Table 7.5: Effect of ϕ cuts on the signal events

Cut	Signal ϕ accepted	Signal B_s accepted
None	86928	–
Vertexing	86660	83715
$999 MeV \leq M_\phi \leq 1039 MeV$	85191	83042
$\chi^2 \leq 6$	82807	80883

Table 7.6: Effect of ϕ cuts on the background. Note that the BGI component in BGII is removed prior to processing to avoid double counting.

Cut	Combinatorial ϕ	Combinatorial B_s^0	BGI	BGII
None	5.36×10^7	–	–	–
Vertexing	4.88×10^7	3850673	60036	80271
$999 MeV \leq M_\phi \leq 1039 MeV$	443426	570458	1859	11352
$\chi^2 \leq 6$	379804	506244	1788	9970

Table 7.7: Effect of B_s^0 cuts on the signal and background events. Note that the BGI component in BGII is removed prior to processing to avoid double counting.

Cut	Signal B_s accepted	Combinatorial background	BGI	BGII
Vertexing	83715	3850673	60036	80271
J/ψ cuts	81713	3005878	58619	64740
ϕ cuts	72229	387319	1742	8133
$5250 MeV \leq M_B \leq 5490 MeV$	72055	6968	1176	113
$\chi^2 < 5.0$	71578	4388	1164	68
$\tau_B > 0.5 ps$	51106	2257	876	27
$p_T > 10.5 GeV$	51100	2257	875	27

Figures 7.1 show the invariant mass distribution for the signal tracks which passed through the cuts imposed by the analysis algorithm, whilst the proper decay time resolution, obtained by subtracting the reconstructed lifetime from that given by the Monte Carlo, is shown in figure 7.2. The sum of two Gaussians, one tall and narrow for the central peak and one short and wide for the tails, were found to adequately describe the signal data. In the case of the B_s^0 mass resolution, the narrow peak had a sigma of $13.3MeV$ and the broad peak $31.4MeV$. The proper time resolution was well described by two Gaussians of sigma $71.6fs$ and $149.3fs$ respectively.

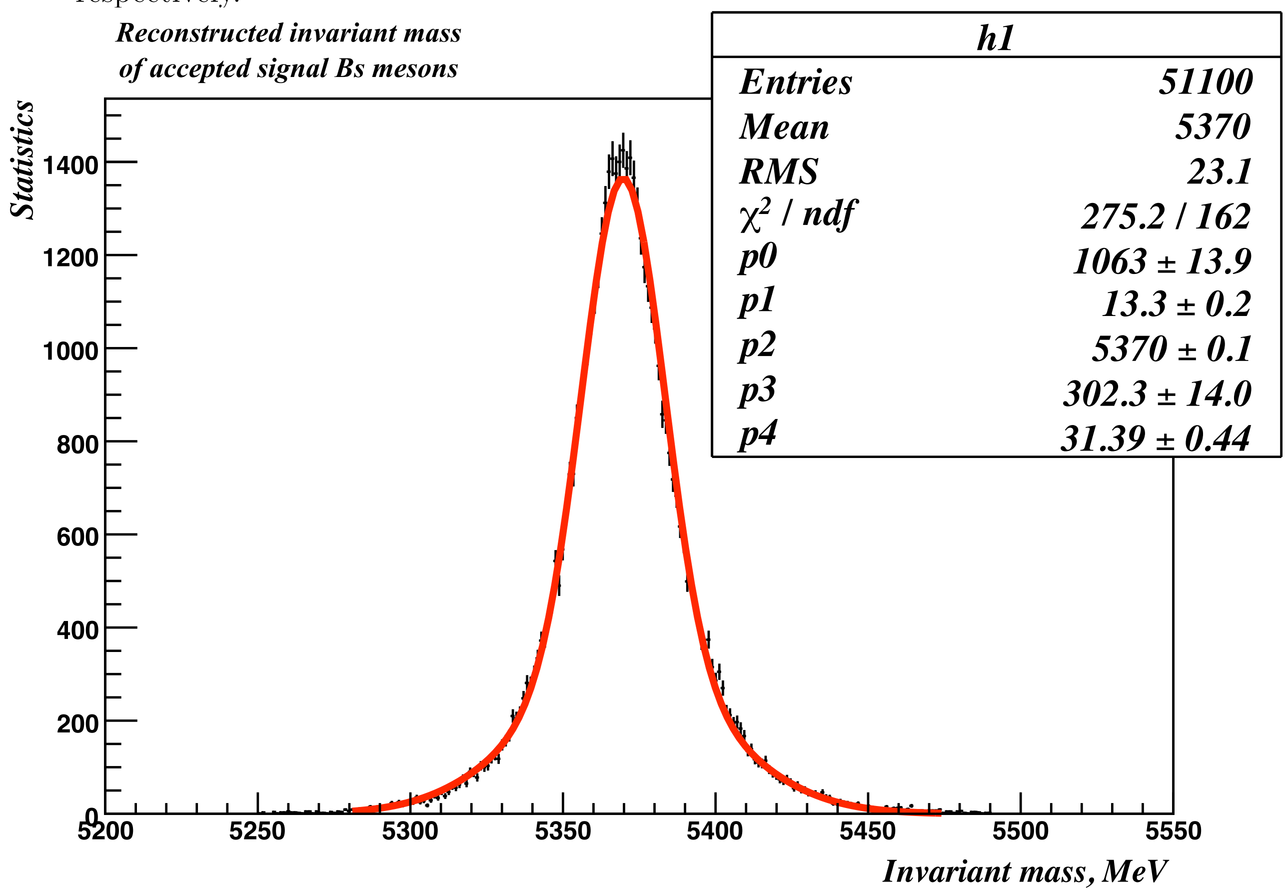


Figure 7.1: Reconstructed invariant mass for all B_s^0 candidates confirmed as being from Monte Carlo signal processes and passing the cuts. The fit is to the sum of two Gaussians, the narrower with $\sigma = 13.3MeV$ and the broader with $\sigma = 31.4MeV$

For the purposes of comparison with other previous ATLAS studies of this form, a more reproducible metric than the double Gaussian must be used. The ATLAS B-physics studies, by convention, fit to a single Gaussian which is truncated at the 10% level; that is to say, the fit is not carried out for data falling below 10% of the

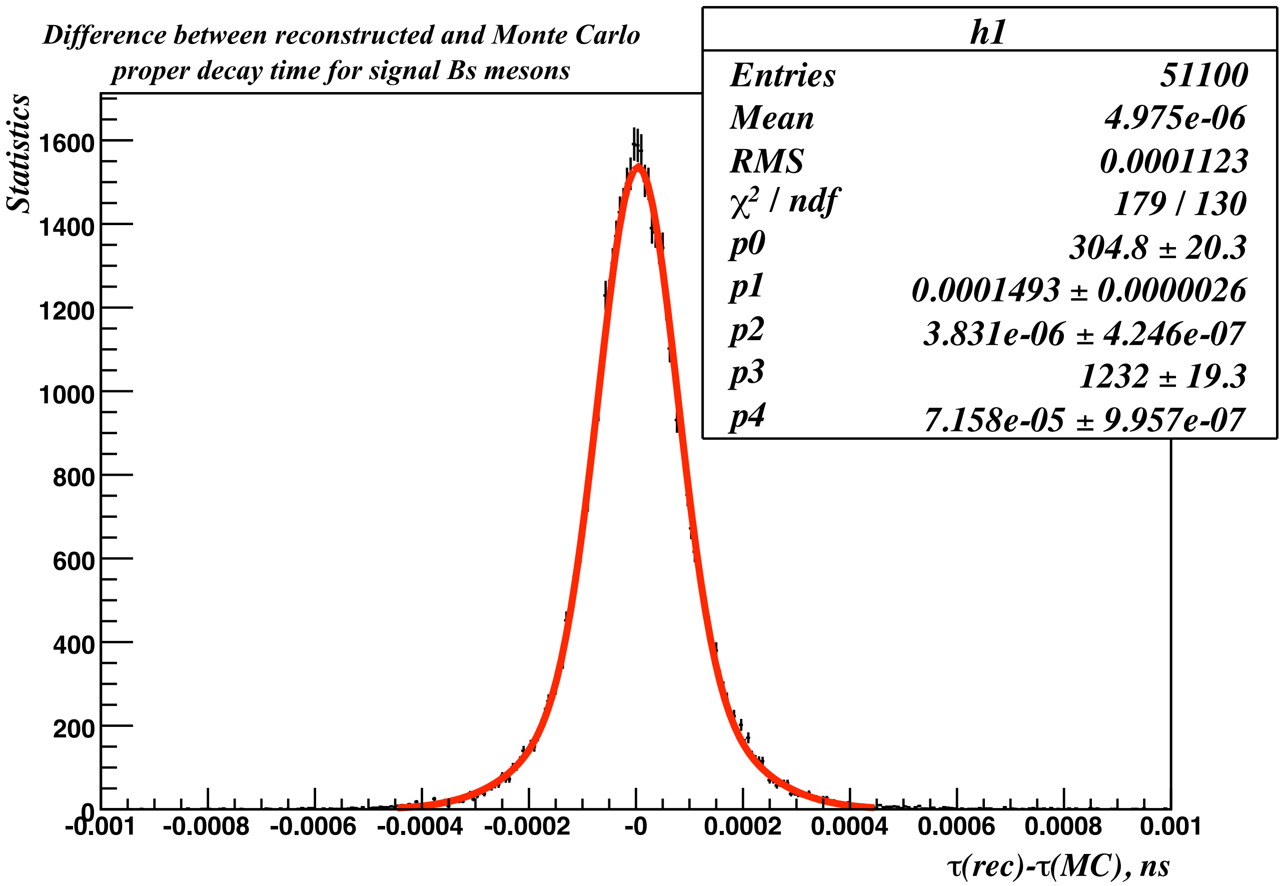


Figure 7.2: Difference between reconstructed and Monte Carlo proper decay time for all B_s^0 candidates confirmed as being from Monte Carlo signal processes and passing the cuts. The fit is to the sum of two Gaussians, the narrower with $\sigma = 71.6fs$ and the broader with $\sigma = 149.3fs$

maximum height of the central peak. The effect of this is to remove the tails from consideration. Table 7.8 compares this study with the results of the assessments carried out for the Technical Design Report [117] and Data Challenge 1 [116]. In each case the figure quoted is for the "complete" detector layout as envisaged by the designers and implemented by the authors of the simulation software at the time of the study; the inclusion of the innermost layer of pixels is the most important aspect of the simulation for such track-based exercises. The TDR studies used the original FORTRAN software, whilst the DC1 study used an early version of the C++ ATHENA code (but still using FORTRAN Geant3 for detector simulation).

Table 7.8: Comparison between this and other ATLAS studies of this channel. All resolutions are produced by fitting to a single Gaussian truncated at 10% of the peak.

	Technical design report	DC1 complete layout	This study
B_s^0 mass resolution, MeV	15.2 ± 0.2	17.2 ± 0.2	17.5 ± 0.1
B_s^0 proper decay time resolution, fs	63.2 ± 0.5	96.2 ± 1.2	86.6 ± 0.4

Plots describing the difference between the reconstructed decay angles and those produced by the Monte Carlo are given in figure 7.3. In each case this is presented in the same form as the data points which were used in the fitting exercises in the final chapter of this work - so the *cosines* of the decay angles θ_1 and θ_2 are shown, whereas the distribution for the third angle ϕ is presented in its raw form. The resolutions are given as full width at half maximum, and are shown on the plots themselves.

Finally, the acceptance of the signal reconstructed B_s^0 mesons are shown as a function of decay angle, transverse momentum and proper decay time in figures 7.4 and 7.5, where the acceptance in this case is defined as the ratio of the number of accepted signal events (i.e., after cuts) to the total number of reconstructed signal events. It is seen that no significant variation is seen across the range of values, which implies that some simplifications could be made when designing acceptance correction methods for processes such as this. This is in line with the findings of Bouhova-Thacker ([115]).

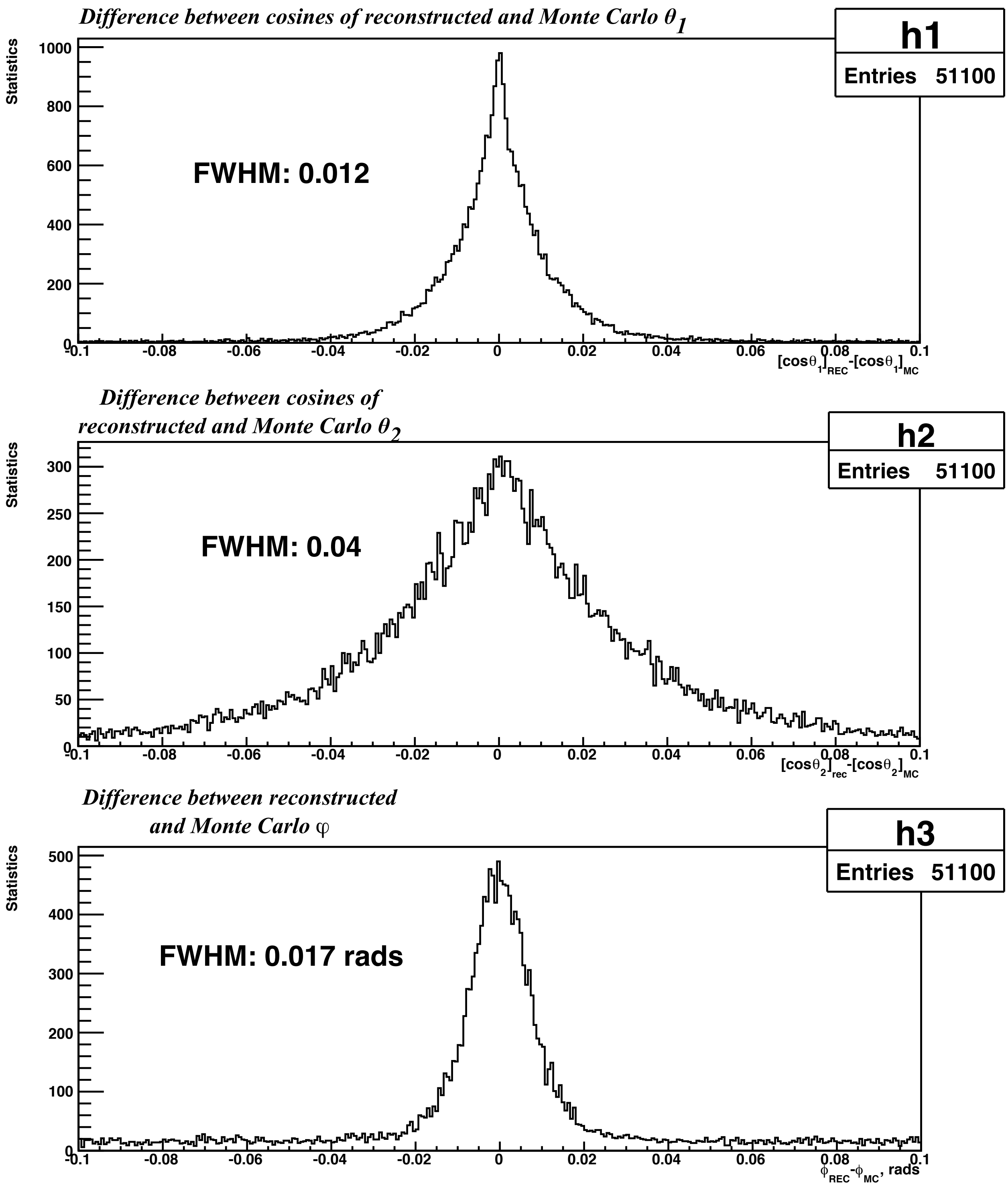


Figure 7.3: Difference between cosines of reconstructed and Monte Carlo decay angles (θ_1, θ_2 and between reconstructed and Monte Carlo ϕ. The full width at half maximum resolutions are: $0.012(\cos\theta_1)$, $0.04(\cos\theta_2)$, $0.017(\phi)$

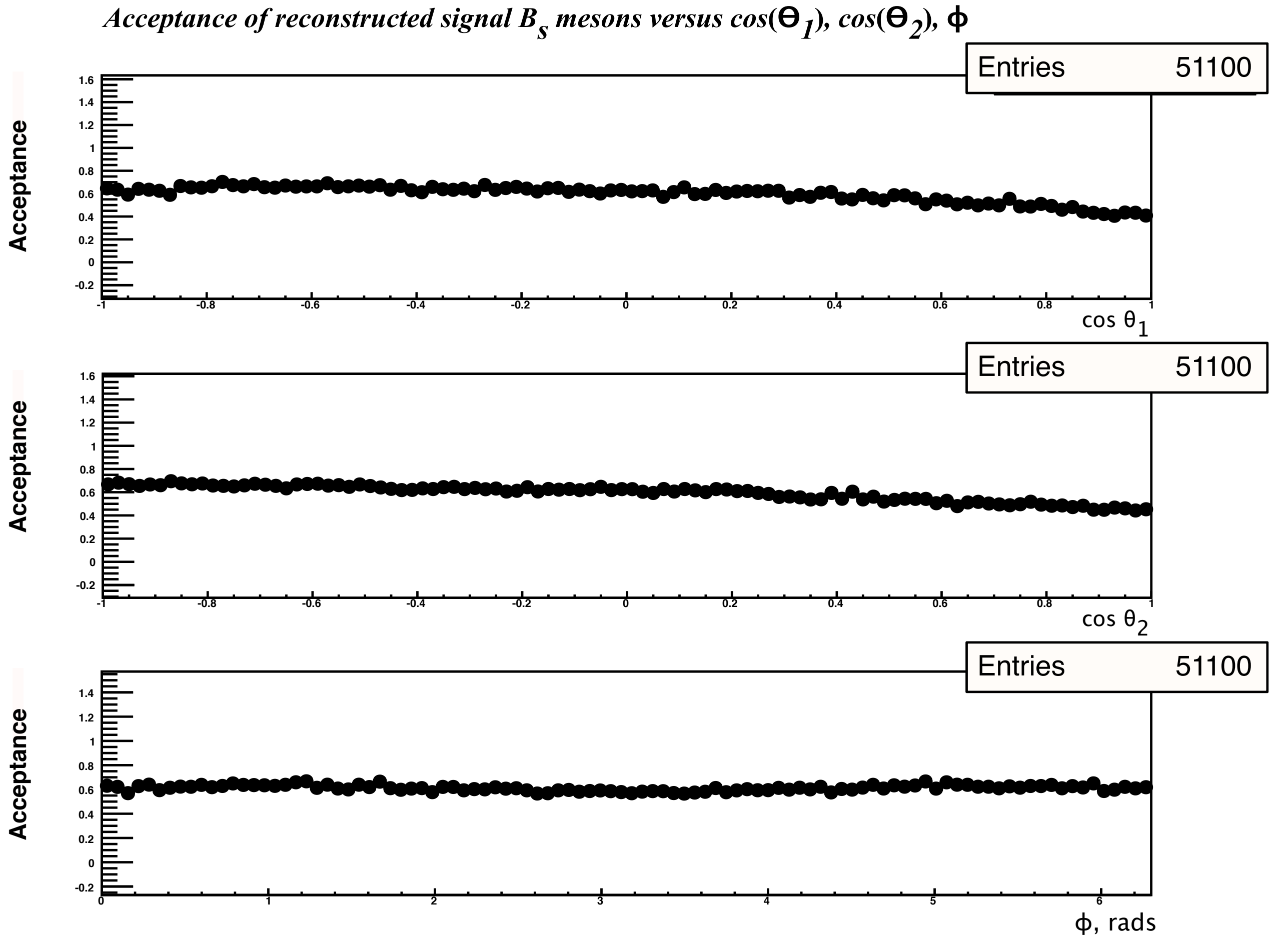


Figure 7.4: Acceptance of reconstructed signal B_s^0 mesons versus $\cos\theta_1$, $\cos\theta_2$ and ϕ

7.2.4 Background acceptance - final results

Table 7.9 gives the acceptance figures for the two backgrounds, and scales them by the process cross section to calculate the total expected contamination by each of the channels[2]. The final headline figure is 22.1%, comprising 13.1% from $B_d^0 \rightarrow J/\psi(\mu\mu)K^{0*}$ (background I), 4.6% from the inclusive process $bb \rightarrow J/\psi(\mu\mu)X$ (background II) and 4.4% from the combinatorial background. Figures 7.6 and 7.7 show the rejection of the two backgrounds with each set of cuts (the statistics for the backgrounds are scaled up according to cross section so this plot is for illustrative purposes only - in reality the cross section would not scale uniformly across the kinematic range). The apparent ineffectiveness of the cuts on the J/ψ candidates is an indication of a commonly observed feature of many of these analyses - the background largely arises from combinations of tracks which have been formed by

[2]See erratum at the end of this document

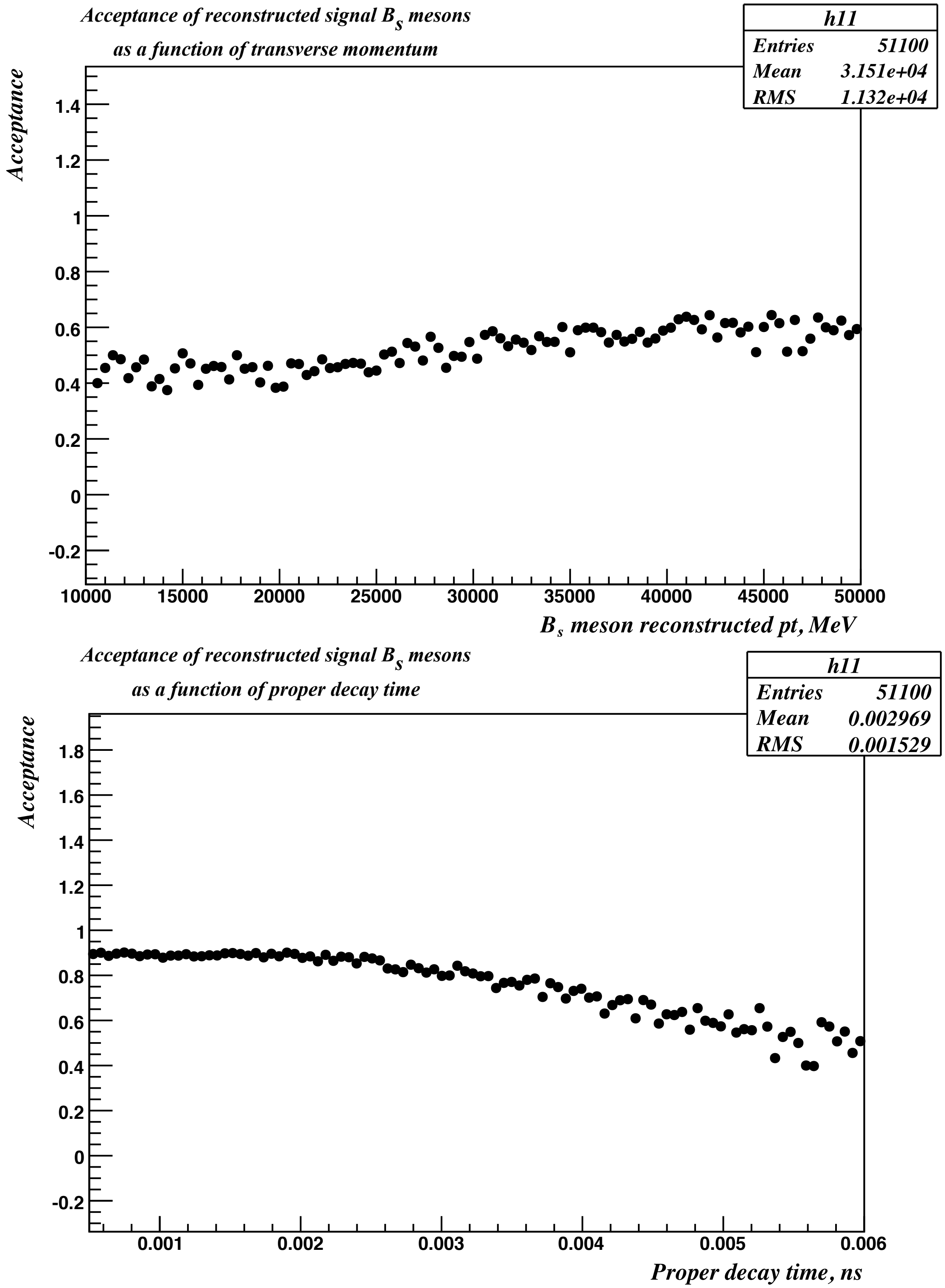


Figure 7.5: Acceptance of reconstructed signal B_s^0 mesons versus transverse momentum and proper decay time.

the decay of a genuine J/ψ and two other incorrectly assigned tracks not associated with decaying B hadron.

Table 7.9: Assessment of the levels of contamination by the two backgrounds. Note that all events of type BGI occurring in the sample of BGII are removed to avoid double counting.

	BGI	BGII
Events processed	113401	87006
Events accepted	875	27
Acceptance	7.7×10^{-3}	3.1×10^{-4}
Cross section ratio B/S	8.4	63.1
Signal events processed	103737	103737
Equivalent b.g. events	871391	7587274
Equivalent b.g. accepted	6710	2352
Signal events accepted	51100	51100
Contamination (%)	13.1	4.6

7.3 Tagging

To make full use of the information on B_s^0 oscillations carried by the angular distributions, it is necessary to determine the state of the B^0 meson - whether it is a B_s^0 or a $\bar{B}_s^0$ at its creation ($t = 0$), before mixing processes begin. This is a decidedly non-trivial exercise which is inherently prone to error. Tagging algorithms fall into two broad categories: **same side tags** (SST), where correlations between the signal B-meson and the other hadrons produced by the decaying b-quark are exploted; and **opposite side tags** (OST), where the analyst relies on the decay products of the other b-quark in the $b\bar{b}$ pair. Detailed studies of the relative performances of different tagging algorithms were performed in [115] and [117]; the results and algorithms developed for those studies are implemented here. No attempt has been made to develop the techniques further - the aim was simply to extract the information necessary to proceed with the weak mixing study.

7.3.1 Opposite Side Tags

Opposite side tags rely on the b or $\bar{b}$ quark that does *not* participate in the signal process. The assumption is made that this quark will decay in a manner which

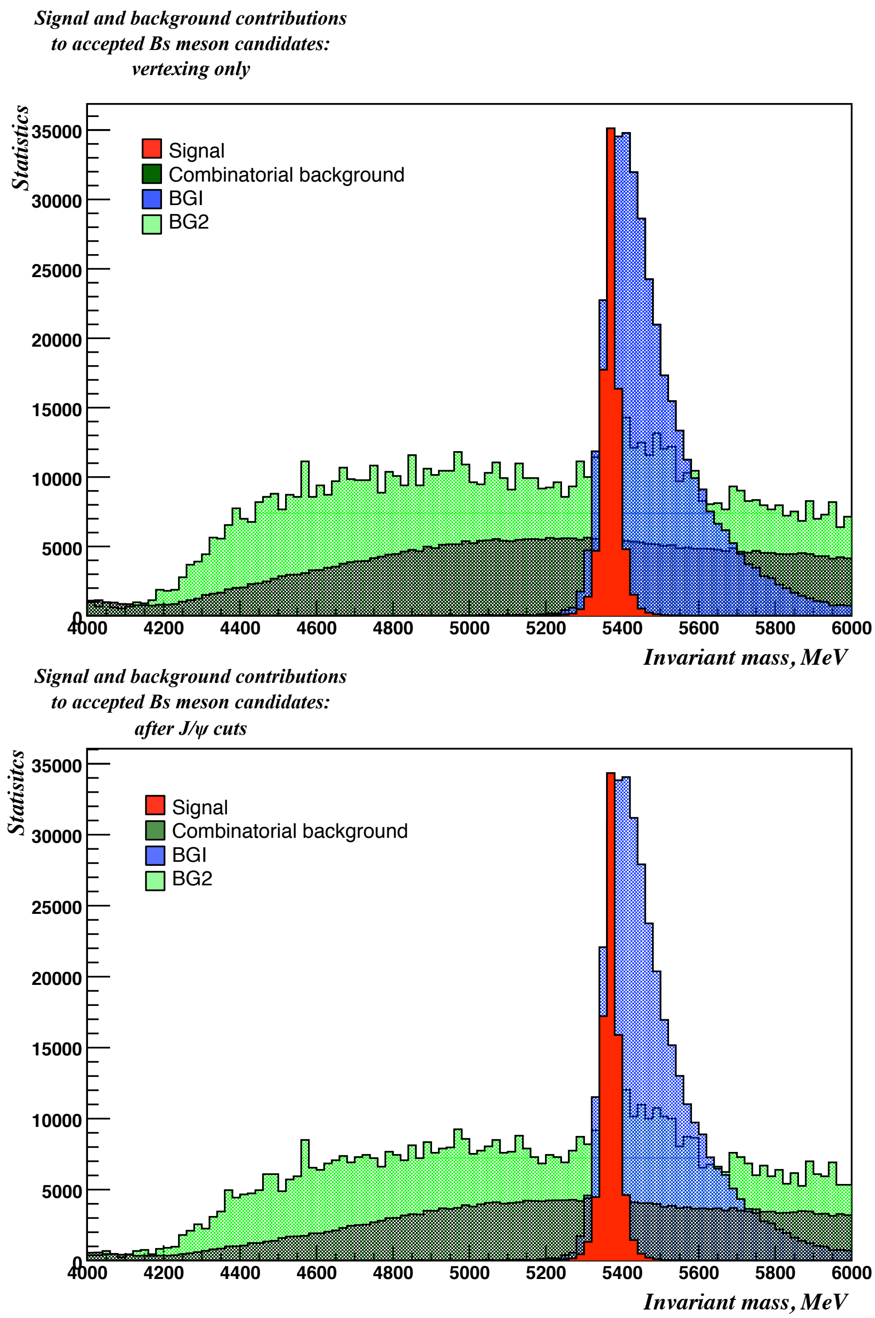
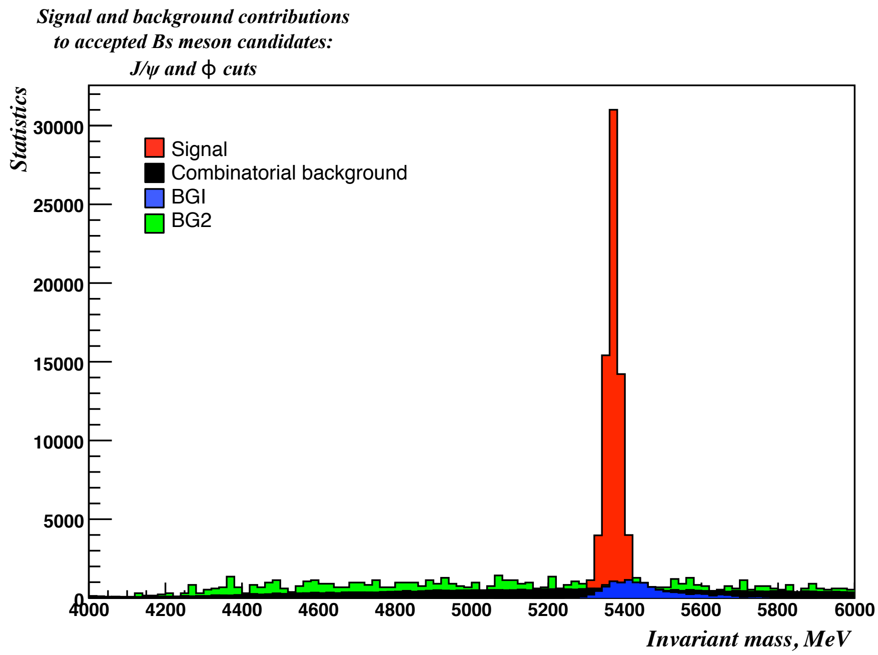


Figure 7.6: Contributions to the B_s^0 candidates from the signal and background, after vertexing and subsequent cuts on the J/ψ. Note that the real backgrounds have been scaled up according to cross section, for illustrative purposes.



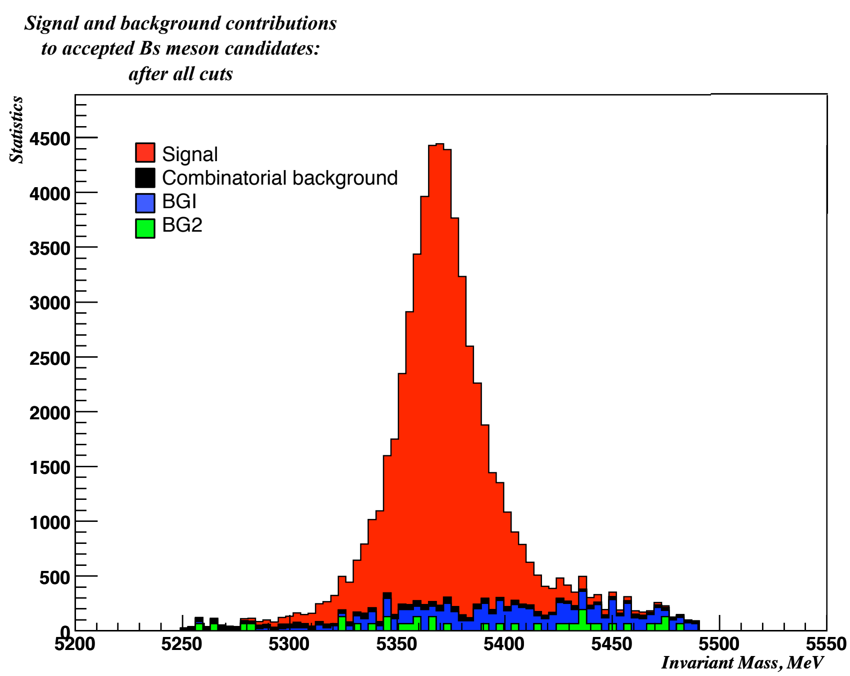


Figure 7.7: Contributions to the B_s^0 candidates from the signal and background, after the ϕ and B_s^0 cuts (the J/ψ cuts are implicit). Note that the real backgrounds have been scaled up according to cross section, for illustrative purposes. The final histogram is a stack of the different contributions, as opposed to the others which are overlaid.

permits the unambiguous determination of its state. This immediately gives the state of the signal quark and hence the signal B-meson. Typically, the technique relies on the opposite side quark decaying semi-leptonically (see figure 7.8). The lepton can be directly observed by the detector; a positively charged lepton indicates a $\bar{B}$-quark and therefore a $\bar{B}$ meson on the signal side, whereas a negative lepton is produced by a b-quark, giving a B-meson on the signal side.

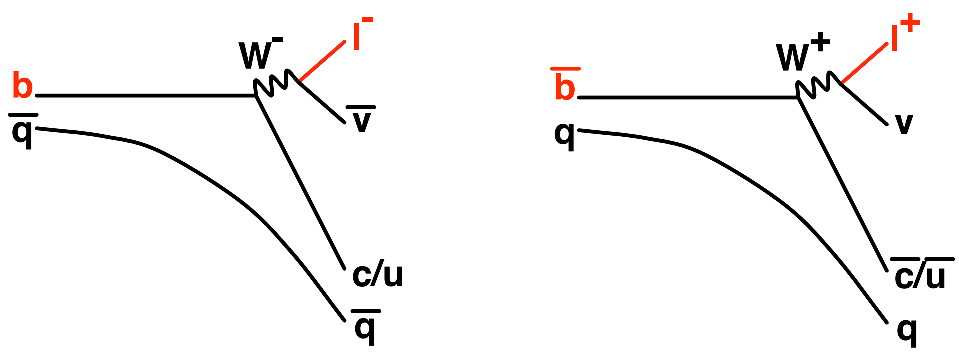


Figure 7.8: Feynman diagrams demonstrating how the semi-leptonic decay of the non-signal b-quark can be used to tag the flavour of the signal B-meson

A particular strength of this method is its simplicity - the algorithms generally involve nothing more complicated than the selection, and determination of the charge, of a track in the event with p_T greater than a certain threshold (which has not already been assigned to the signal event). This particle is taken to be from the semi-leptonic decay of the opposite side b-quark. Tags obtained from this method are also fairly *pure* - where an OST is performed, the probability of it being incorrect are rather low [117]. Errors can occur when the opposite side quark does not decay semi-leptonically and yet a high-p_T lepton is detected ($b \to c \to \mu$ and $c \to \mu$ decays [115]). If the opposite side quark forms a neutral B-meson then mixing will occur on both sides and a mis-tag is likely.

A more serious drawback of this method is the rarity of the semi-leptonic decay, on which the tag depends. The branching ratio for this process is of the order of 8% [106] so the tag *efficiency*, the proportion of events tagged, is low [115] [117]. If the analyst were to rely on this method alone, the number of events available for a full tagged analysis would be severely depleted, but those tags which were imposed would be of a high quality.

7.3.2 Same Side Tags

Same side tagging algorithms rely on correlations between the flavour of the quark which forms the B-meson and the charges of the particles produced during this fragmentation. The correlations can arise through a variety of mechanisms - for instance, figure 7.9. Essentially, the role of the algorithm is to identify which tracks are associated with the B-meson, and establish the overall charge such that the tag can be assigned

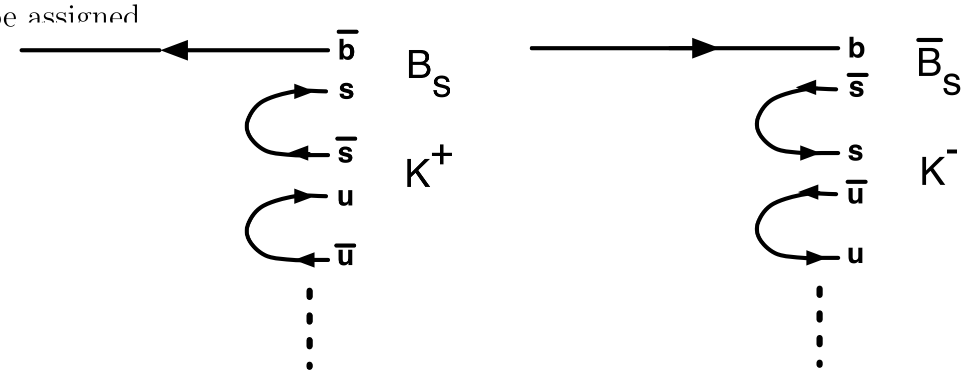


Figure 7.9: Feynman diagrams demonstrating how the charges of the particles associated with the fragmentation of a b-quark into a B-meson may be correlated with the flavour of the b-quark

A commonly used method is the **jet charge algorithm**. This makes use of a quantity referred to as the jet charge Q_{jet}, given by [115]

$$Q_{\text{jet}} = \frac{\sum_i q_i p_i^\kappa}{\sum_i |p_i|^\kappa} \tag{7.1}$$

where the index i runs over all tracks within a cone whose central axis is defined by the momentum vector of the B-meson. p_i is a measure of the momentum of the ith track, and q_i is its electric charge. The tracks within the cone are referred to as the *jet*; it is expected that, on average, sign of the jet charge associated with a $\bar{b}$-quark will be positive, and vice versa for the conjugate case (see figure 7.9), so this allows the flavour of the signal meson to be determined. The opening angle of the cone (which determines the number of tracks in the jet) and the quantity κ are free parameters and must be tuned with Monte Carlo data to give the most reliable and efficient tag. It is also possible to improve the algorithm by removing ambiguous cases (events with jet charges close to zero) from consideration, leading to an **exclusion region**. Events whose jet charges fall into this region are not tagged at all. Again, the extent of this region must be determined through tuning.

The measure of the momentum of each track in the jet can be defined in a number of ways. Following the recommendation of Bouhova-Thacker [115] this is chosen to be the component of the momentum of the track parallel with the momentum vector of the signal B-meson, and is denoted p_L.

7.3.3 Tag quality

Both of the tagging techniques deployed here have a number of parameters which need to be tuned to allow the algorithms to operate as effectively as possible. In general there are two measures of the potency of the tagging algorithm - the *efficiency* and the *purity*. The efficiency is simply the fraction of events which received a tag (correct or incorrect), and is given by

$$\epsilon_{\mathrm{tag}} = \frac{N_r + N_w}{N_t} \tag{7.2}$$

where N_r and N_w are the number of correctly and incorrectly tagged events respectively, and N_t is the total number of events processed by the tagging algorithm. Clearly a high efficiency is desirable.

The purity is given by

$$D_{\mathrm{tag}} = \frac{N_r - N_w}{N_r + N_w} = 1 - 2w_{\mathrm{tag}} \tag{7.3}$$

where w_{tag} is the wrong tag fraction, that is, the number of incorrect tags as a fraction of the total number of tagged events, given by

$$w_{\mathrm{tag}} = \frac{N_w}{N_r + N_w} \tag{7.4}$$

The requirement is that the wrong-tag fraction be as low as possible, that is, the purity to be as close to 1 as can be arranged. These two measures of tag algorithm effectiveness can be combined into a single metric, known as the quality factor Q_{tag}. This is given by

$$Q_{\mathrm{tag}} = \epsilon_{\mathrm{tag}} D_{\mathrm{tag}}^2 \tag{7.5}$$

This factor was used for the tuning of the tagging algorithms investigated in this work.

7.3.4 Results of tagging studies

Opposite side tags

The opposite side tagging algorithm was implemented alongside the main analysis code. For every event, tracks having a pseudorapidity within the standard B-physics range ($|\eta < 2.7|$) and a transverse momentum above a certain threshold value were selected, and, using Monte Carlo truth to simulate the lepton identification mechanisms of ATLAS, electron and muon tracks were identified. The signal leptons (from the decay of the J/ψ) were removed. The track with the highest remaining p_T was singled out as the "tag lepton", and the flavour of the signal B was determined according to the charge of this track. Given that the the signal samples contained only B_s events, any $\bar{B}_s$ tags were immediately evident as being incorrect. No account was taken of the possibility that non-leptons could activate the electron or muon identification apparatus, leading to a false signal (punch-through to the muon chambers being a case in point).

The transverse momentum threshold was determined by varying this quantity until the quality factor was at a maximum. This work was performed on a sub-set of 3000 events; figure 7.10 shows the quality factor obtained for a range of thresholds. 3500GeV was selected as the value. The algorithm was then run over the entire dataset, yielding the results given in table 7.10

Table 7.10: Results for opposite side tagging algorithm

Number of signal events processed	103737
Number of correct tags N_r	8066
Number of incorrect tags N_w	4014
Efficiency ϵ	11.6%
Wrong-tag fraction w_{tag}	33.2%
D_{tag}	0.335
Q_{tag}	0.013

Same side tags

The jet-charge tagging algorithm was implemented as follows:

- All tracks passing the basic cuts $|p_T| > 500MeV$ and $|\eta| < 2.7$ were accepted as potential members of the same-side jet.

- The jet was then defined as consisting of those tracks occupying a cone around the signal B_s momentum vector (in the laboratory frame). The size of the cone , denoted $\Delta R = \sqrt{\Delta\eta^2 + \Delta\varphi^2}$ where $\Delta\eta$ and $\Delta\varphi$ are the differences in

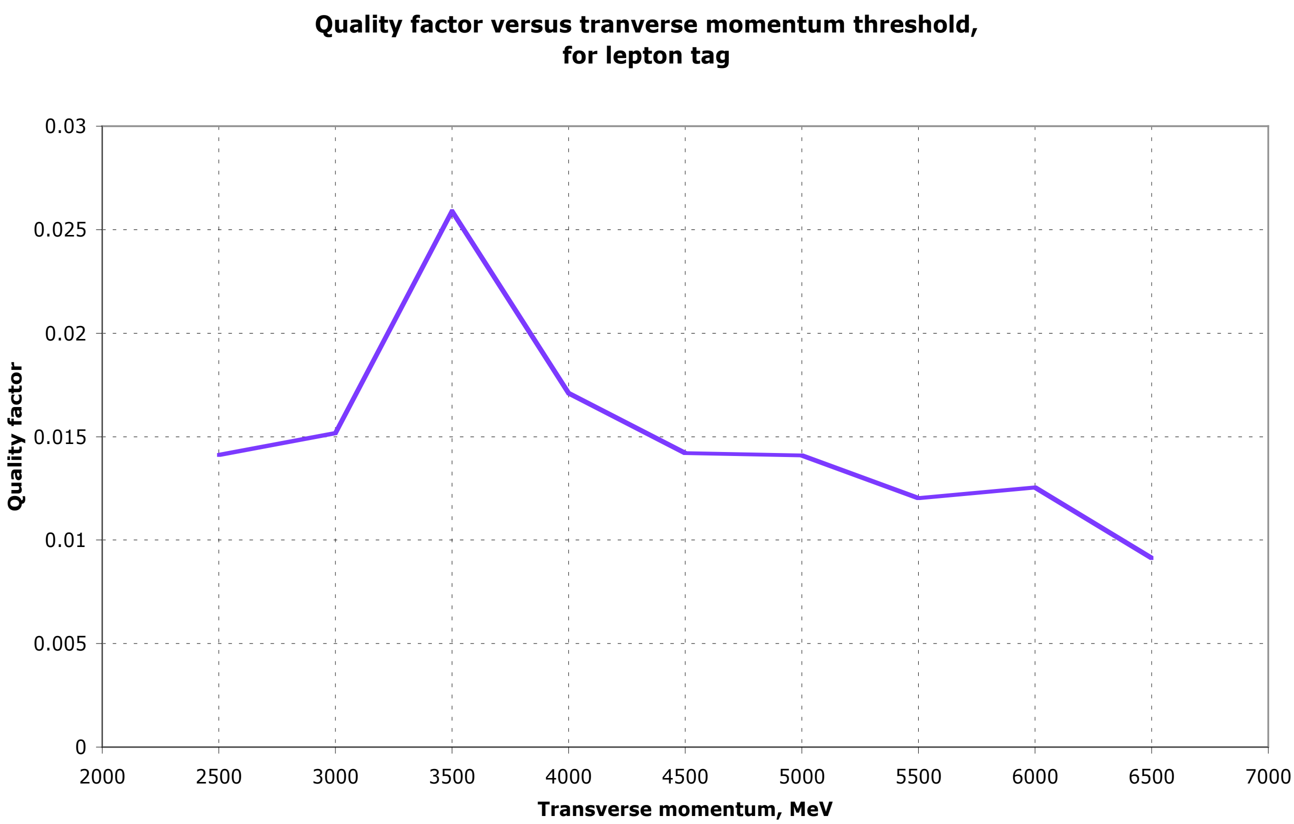


Figure 7.10: Quality factor for the opposite side lepton tagging algorithm for a variety of transverse momentum thresholds. Produced with a subset of 3000 events.

pseudorapidity and azimuthal angle between the cone wall and the B-meson, was a parameter to be determined through quality factor maximization.

- The momentum measure p_L (as described above) was calculated for each of the tracks falling within the cone, and the jet charge Q_{tag} for the event calculated according to equation 7.5, with the parameter κ to be determined through tuning.

- Those events with a jet charge within the "exclusion region" (whose size was to be determined through tuning) were denoted as having no tag.

- Outside this region, those events with a negative jet-charge were tagged as being from a $\bar{B}_s^0$ (and therefore wrong-tagged with this data) whilst those with a positive jet-charge were accorded a B_s^0 (correct) tag.

The sub-sample of 3000 events was again used to tune the parameters κ, ΔR and the exclusion region size, by varying these quantities to find the highest possible quality factor. Figures 7.11 to 7.13 show the results of this tuning exercise. The optimum cone size was found first (figure 7.11), followed by κ (7.12). No exclusion of small jet charge tags was imposed for these tests. Finally, with the optimum cone size and κ established, the quality factor was maximized with respect to varying exclusion region sizes (7.13). The optimum values were found to be $\Delta R = 0.6$ and $\kappa = 0.3$, with tags denied to all events with a jet-charge of $|Q_{jet}| < 0.1$.

The algorithm was then run over the entire dataset using these optimum values. Figure 7.14 is a histogram of the jet charge for each of the events. Those events in the red and green regions were incorrectly and correctly tagged respectively, whilst events falling into the central black region were not tagged at all. The same results are given in tabular form in table 7.11.

As the determination of the weak mixing phase relies on tagged events, as many of the events should be tagged as possible. As suggested by Bouhova-Thacker [115] it is expedient to use both tagging algorithms in tandem, using an opposite side lepton tag where available (due to the lower wrong-tag fraction) and otherwise deploying the jet-charge method. This scheme was applied to the complete dataset, yielding results as given in table 7.12. The tagging efficiencies and wrong-tag fractions shown here were then assumed in the data fitting exercises following this chapter.

Table 7.13 summarizes the findings of this chapter.

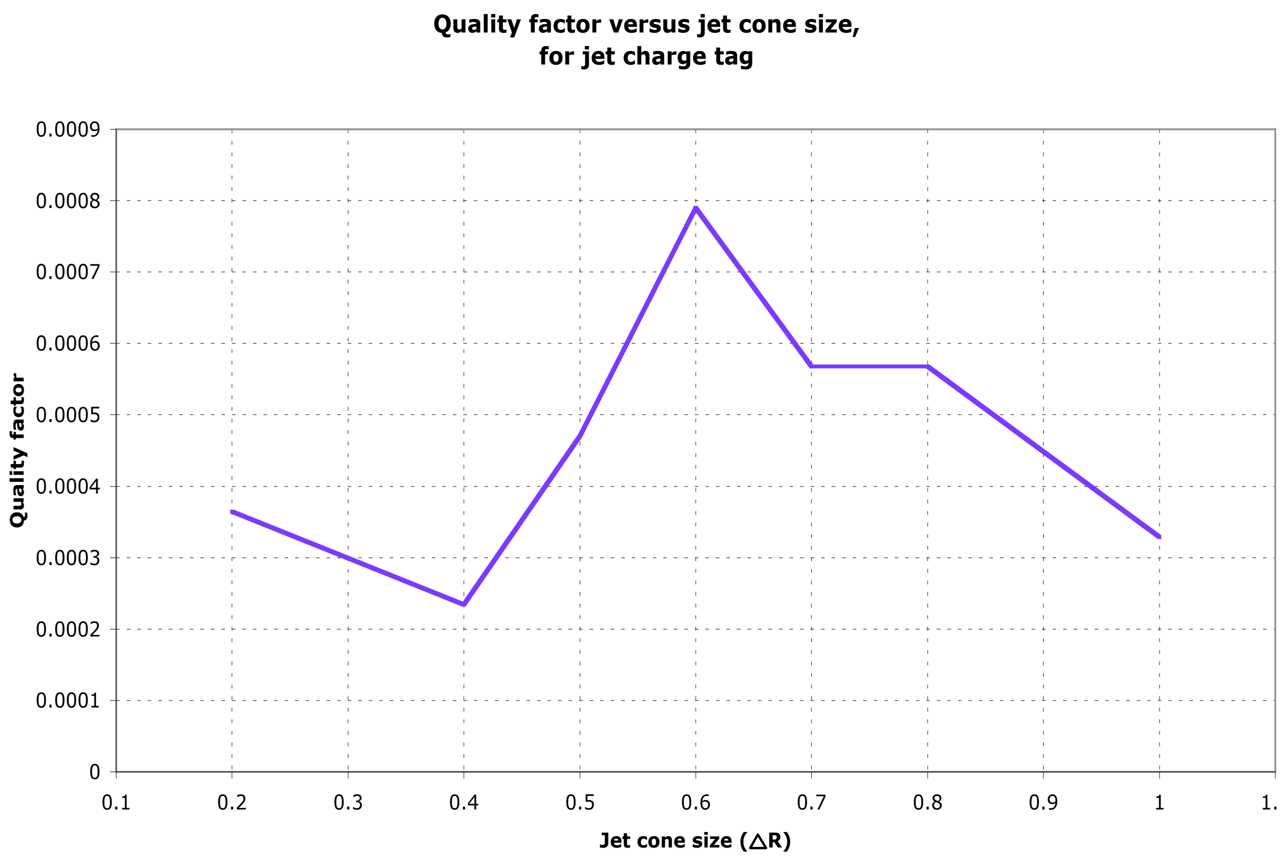


Figure 7.11: Quality factor for the same-side jet charge tagging algorithm for a variety of jet cone sizes (ΔR). Produced with a subset of 3000 events.

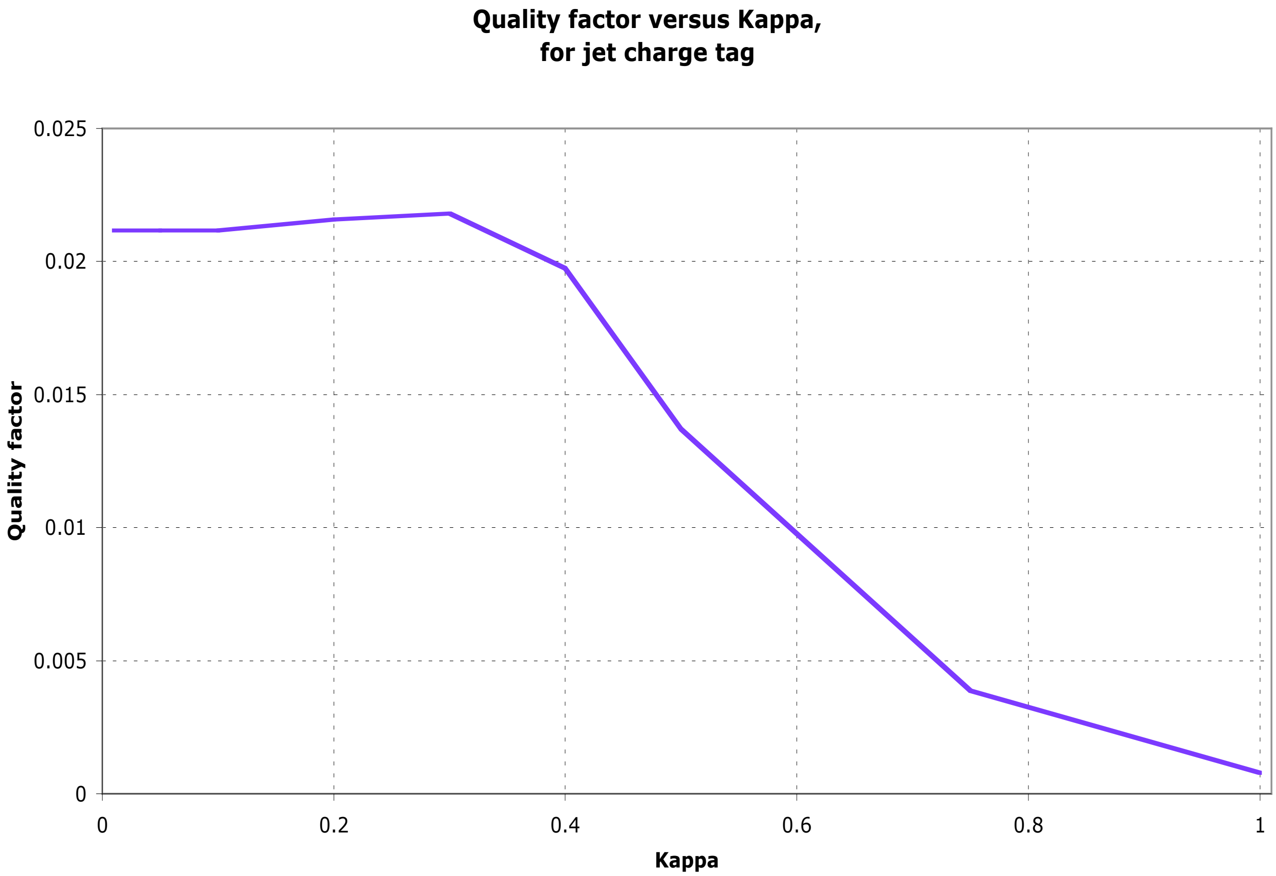


Figure 7.12: Quality factor for the same-side jet charge tagging algorithm for a variety of values of κ. Produced with a subset of 3000 events.

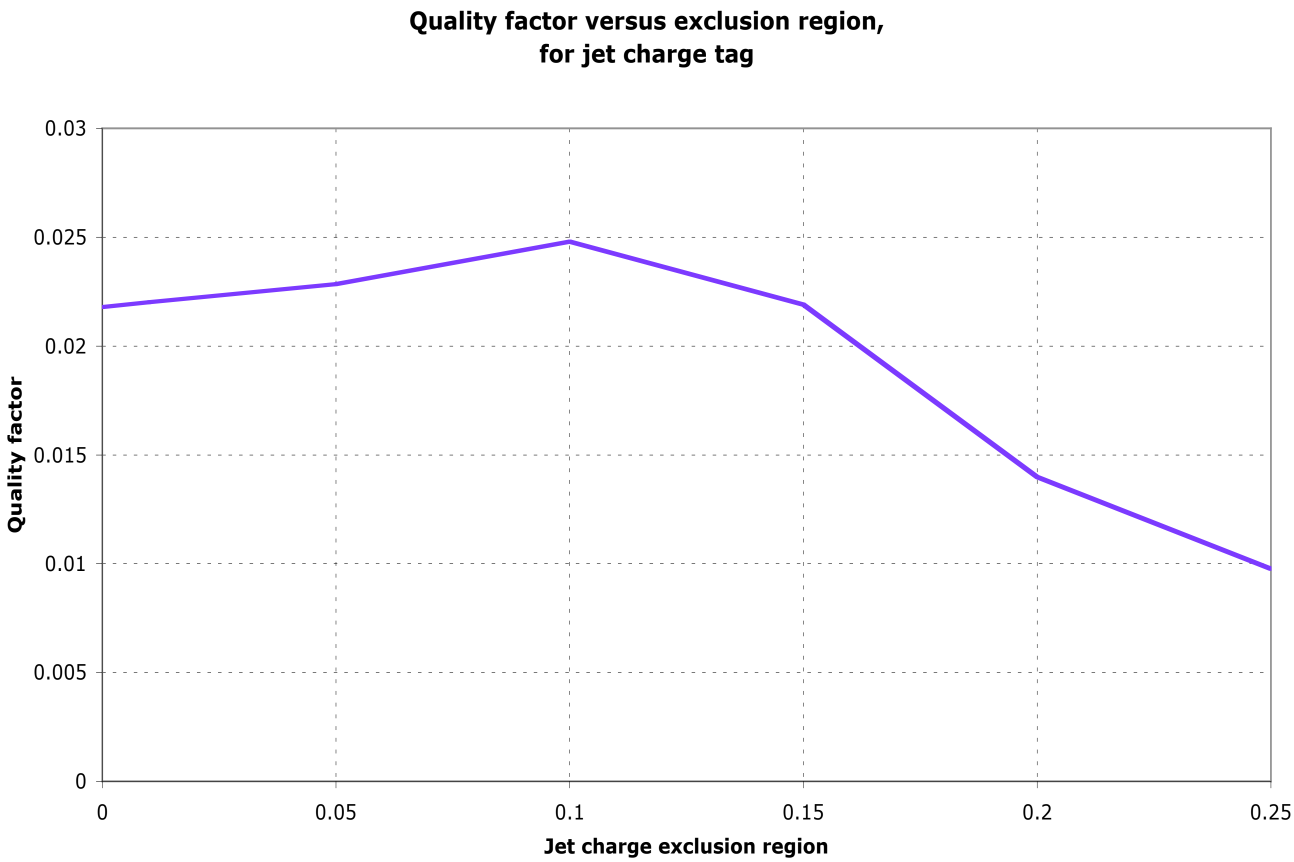


Figure 7.13: Quality factor for the same-side jet charge tagging algorithm with varying ranges of small jet charges excluded. Produced with a subset of 3000 events.

Table 7.11: Results for same side jet charge tagging algorithm

Number of signal events processed	103737
Number of correct tags N_r	28788
Number of incorrect tags N_w	18158
Efficiency ϵ	45.3%
Wrong-tag fraction w_{tag}	38.7%
D_{tag}	0.226
Q_{tag}	0.023

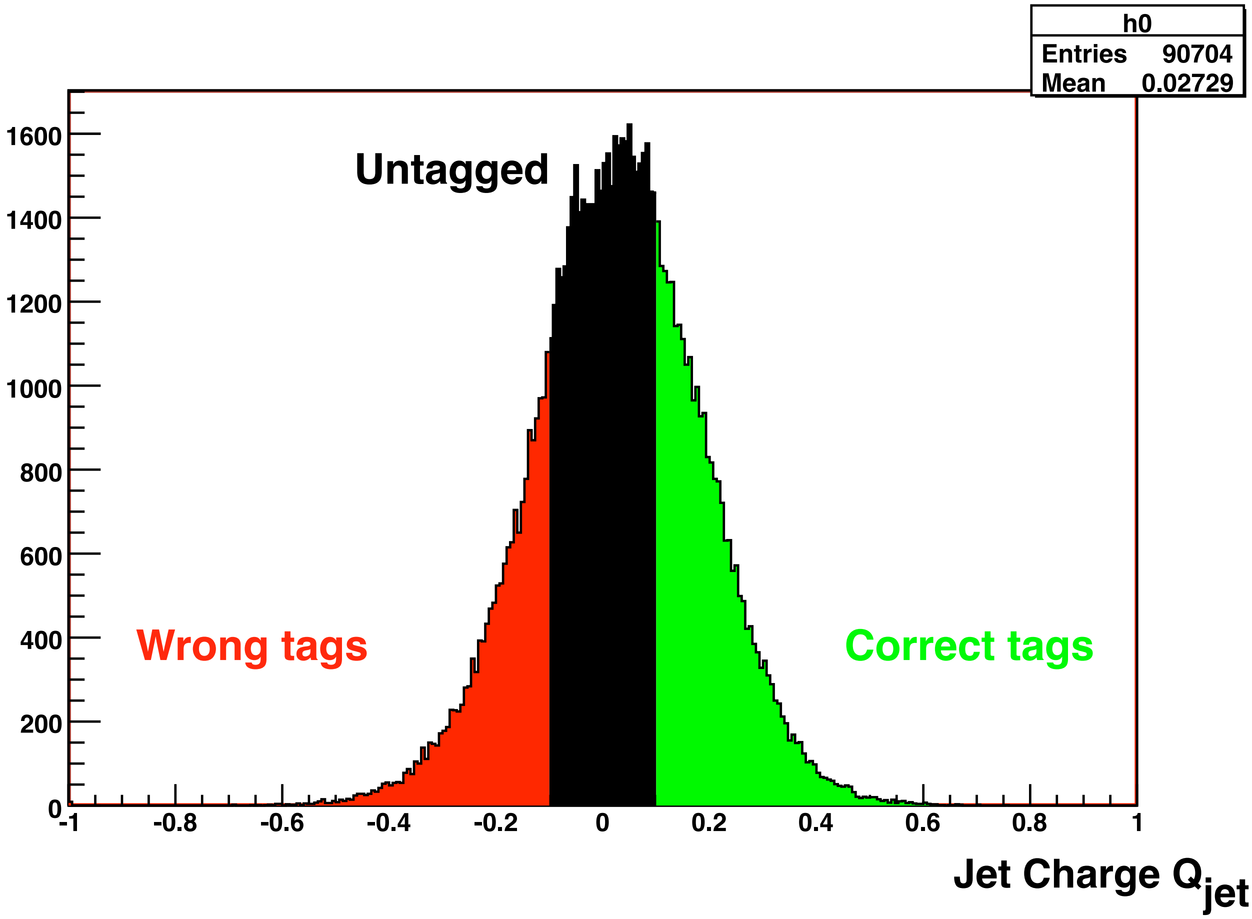


Figure 7.14: Jet charges for the complete sample of signal events, using the optimum parameters. Event in red received the wrong tag, whilst those in green were correctly tagged. The central (black) events were not tagged at all.

Table 7.12: Results for combined tagging algorithm

Number of signal events processed	103737
Number of correct tags N_r	33564
Number of incorrect tags N_w	19988
Efficiency ϵ	51.6%
Wrong-tag fraction w_{tag}	37.3%
D_{tag}	0.254
Q_{tag}	0.033

Table 7.13: Summary of chapter 7 results

Signal events acceptance	0.492
BGI acceptance	7.7×10^{-3}
BGII acceptance	3.1×10^{-4}
BGI contamination, %	13.1
BGII contamination, %	4.6
Combinatorial BG contamination, %	4.4
Overall wrong-tag fraction, %	37.3
Overall tag efficiency, %	51.6
Overall tag quality factor	0.033
B_s mass resolution, MeV (two Gauss, one Gauss)	13.3 & 31.4, 17.5
B_s proper time resolution, fs (two Gauss, one Gauss)	71.6 & 149.3, 86.6

Chapter 8

Analysis of the decay $B_s^0 \to J/\psi\phi$ III - data fitting

The ultimate aim of the analysis is to extract the quantities describing physics processes from the reconstructed angular distribution. This involves fitting the data to the theoretical expression as given in tables 6.1 and 6.2. There are a wide variety of fitting techniques in use which vary in complexity, sophistication, accuracy and the demands made on computing facilities.

In this case the chosen method is the **Maximum Likelihood** technique. In this chapter the principles of maximum likelihood and the properties of maximum likelihood estimators are described. We shall then consider how such a fit is constructed for this case, considering complications such as normalisation, resolution and background effects. The routine will then be applied directly to Monte Carlo distributions to study its properties and behaviour, with the aim of judging the feasibility of carrying out such an analysis "for real" when ATLAS begins to take data in 2007.

8.1 The technique of Maximum Likelihood

Let us suppose that some theory predicts that a measurable quantity x is distributed according to a probability density function (pdf) given by $f(x;\lambda)$, where λ is a parameter or set of parameters. The theory provides the functional form of $f(x;\lambda)$, but gives no indication as to the value(s) of λ. Now, let the measurement of x be performed n times, giving a finite set of data points. The task at hand is to extract λ from the data, under the assumption that the theory producing $f(x;\lambda)$ is correct.

Under the assumption of negligible systematic error we can say that the probability for the ith measurement to be in an interval $x_i + dx_i$ is given by $f(x_i;\lambda)\,dx_i$.

The probability that this holds for all n measurements is given by

$$P = \prod_{i=1}^{n} f(x_i; \lambda)\, dx_i \tag{8.1}$$

If the theory is correct and the value(s) of λ are close to their physical values, the set of data points which are actually measured should yield a high value of P. In contrast, if the λ are incorrect, the probability is low. Now the value of dx_i is arbitrary so we can write:

$$\mathcal{L}(\lambda) = \prod_{i=1}^{n} f(x_i; \lambda) \tag{8.2}$$

$\mathcal{L}$ is referred to as the **likelihood function**. Quite simply, the maximum likelihood method is the finding of λ for which $\mathcal{L}$ is a maximum. This value or values $\hat{\lambda}$ is known as the **maximum likelihood estimator**, and it represents the physical value or values which makes the measured set of data points most probable. The method is of course only as good as the theory which produced the pdf and is limited by the measurement technique. Note that, in general, the log of the likelihood function is taken, enabling the product to be converted into a sum. The likelihood function must also be **normalized** such that

$$\int_0^\infty \mathcal{L} dx = 1 \tag{8.3}$$

so in general it is necessary to include a normalization factor, which is found by integrating the pdf over the whole phase space.

The main challenge of the technique is the search for the estimator. In principle it is possible to find the value analytically, by locating the turning points of the likelihood estimator. In practice this is often highly problematic because of the appearance of local maxima; it can also be an algebraically daunting task for complex pdfs with more than one parameter. Some cases (those with one or two parameters) can be tackled graphically. Most often, however, it is necessary to resort to a **minimization routine** which methodically varies the parameters until the maximum likelihood estimator is found. In high energy physics (at least in connection with CERN experiments) the program **Minuit** (described in the computing chapter of this work), is overwhelmingly the most popular and is deployed here.

8.1.1 Properties of Maximum Likelihood Estimators

Consistency, Bias and Invariance

The terms **consistency** and **bias** apply to all statistical estimators. If the estimator for a quantity λ is $\hat{\lambda}$, then it is **consistent** if it converges to λ in the limit of a large number of data points. Clearly this is the essential requirement of any estimator.

An estimator is itself a random variable since it is calculated through the analysis of randomly distributed data points. If the entire experiment were to be repeated a number of times, each time taking n measurements, then the value of the estimator $\hat{\lambda}$ would be distributed according to some pdf $g\left(\hat{\lambda};\lambda\right)$, which is not known in general. Such a pdf is known as a **sampling distribution** [118]. The **expectation value** $E\left(\hat{\lambda}\right)$ of an estimator is the expected mean value of $\hat{\lambda}$ from an infinite number of identically arranged experiments and is given by [118]

$$\begin{aligned} E\left[\hat{\lambda}(\mathbf{x})\right] &= \int \hat{\lambda} g\left(\hat{\lambda};\lambda\right) d\hat{\lambda} \\ &= \int \ldots \int \hat{\lambda}(\mathbf{x}) f(x_1;\lambda) \ldots f(x_n;\lambda) dx_1 \ldots dx_n \end{aligned} \tag{8.4}$$

The **bias** is given by [118]

$$b = E\left(\hat{\lambda}\right) - \lambda \tag{8.5}$$

An **unbiased** estimator has a zero bias irrespective of the sample size; an **asymptotically unbiased** estimator has a bias which approaches zero as the sample size becomes infinitely large. An estimator can be consistent and yet biased - a single experiment with an infinitely large sample size may give an estimator which tends to the true value, but the average result from an infinite number of experiments each with a finite sample size may give a non-zero bias [118]. In general, maximum likelihood estimators are biased [119], but for sufficiently large samples this bias is small compared with the statistical error and can be safely ignored.

One of the most attractive properties of maximum likelihood estimators is their **invariance** to parameter transformations. The maximum likelihood technique involves the finding of turning points in the likelihood function - the place where this turning point occurs does not change. In other words, if the turning point is at a, then that value will still be found whether the parameter being fitted is a, a^2 or $a^{-\frac{1}{3}}$. This gives the analyst a great deal of freedom in assembling the liklihood function. It should be pointed out, however, that a change in variables can introduce a bias [119].

Variance

The variance of a maximum likelihood estimator, given by

$$V\left[\hat{\lambda}\right] = E\left[\hat{\lambda}^2\right] - \left(E\left[\hat{\lambda}\right]\right)^2 \tag{8.6}$$

is difficult to calculate analytically for all but the simplest likelihood functions, due to the complexity of evaluating equation 8.4. In some cases a Monte Carlo approach could be used to perform multiple maximum likelihood fits and thereby obtain the spread in the estimators, but this is impractical if the fit takes time to converge. Yet calculation of the variance is vitally important as it provides the statistical error of the fit. Generally, then, the **Rao-Cramer-Frechet inequality** (**RCF** or **information inequality**) is used. This is stated without proof here; the derivation can be found in [120]. It expresses the relationship between the *data*, the *bias* and the *variance*, giving a *lower bound* on the variance. This is a general expression which applies to all estimators.

$$V\left[\hat{\lambda}\right] \geq \frac{\left(1+\frac{\partial b}{\partial \lambda}\right)^2}{E\left[-\frac{\partial^2 \log L}{\partial \lambda^2}\right]} \tag{8.7}$$

If the inequality becomes an equality (such that the minimum variance is reached) the estimator is described as **efficient**. Maximum likelihood methods always locate the efficient estimator if it exists. In the limit of a large sample, ML estimators are efficient [118].

Where more than one parameter is involved in the fit the following formula describes the elements of the covariance matrix of the estimators [118]

$$\begin{aligned}\left(V^{-1}\right)_{ij} &= \operatorname{cov}\left[\hat{\lambda}_i, \hat{\lambda}_j\right] \\ &= E\left(-\frac{\partial^2 \log L}{\partial \lambda_i \partial \lambda_j}\right) \\ &= n \cdot \int -f\left(x;\lambda\right) \frac{\partial^2}{\partial \lambda_i \partial \lambda_j} \log f\left(x;\lambda\right) dx \end{aligned} \tag{8.8}$$

where there are n measurements and $f\left(x;\lambda\right)$ is the pdf for the random variable x, and where 8.4 has been used. The multiplier n leading the equation is expected, as this expresses the fact that statistical error on the standard deviation ($\sigma = \sqrt{V}$) drops off as the inverse of the square root of the sample size. The integral over x makes this impractical for most applications. However for large samples V^{-1} can be estimated by evaluating the derivative with the *measured data* and the *ML estimates* $\hat{\lambda}$ [118]

$$\left(\hat{V}^{-1}\right)_{ij} = -\left[\frac{\partial^2 \log L}{\partial \lambda_i \partial \lambda_j}\right]_{\lambda=\hat{\lambda}} \tag{8.9}$$

This approximation is used, for instance, by Minuit when it calculates the covariance matrix of the estimator. In the case of a single parameter, the variance can be assessed graphically. If we Taylor-expand a log-likelihood function about the maximum likelihood estimator $\hat{\lambda}$, we get

$$\begin{aligned}\log \mathcal{L}(\lambda) = \quad & \log \mathcal{L}(\lambda) + \left[\frac{\partial \log \mathcal{L}}{\partial \lambda}\right]_{\lambda=\hat{\lambda}} \left(\lambda - \hat{\lambda}\right) + \\ & + \frac{1}{2!}\left[\frac{\partial^2 \log \mathcal{L}}{\partial \lambda^2}\right]_{\lambda=\hat{\lambda}} \left(\lambda - \hat{\lambda}\right)^2 + \ldots \end{aligned} \tag{8.10}$$

Now $\hat{\lambda}$ is at the turning point so $\log \mathcal{L}(\hat{\lambda}) = \log \mathcal{L}_{\max}$ and the second term in the expansion is zero, the gradient being flat. If we now use equation 8.9, and disregarding terms of higher order than two, we can write

$$\log \mathcal{L}(\lambda) = \log \mathcal{L}_{\max} - \frac{(\lambda - \hat{\lambda})^2}{2\hat{\sigma}^2_{\hat{\lambda}}} \tag{8.11}$$

If we let $\lambda = \hat{\lambda} \pm \sigma^2_{\hat{\lambda}}$ then

$$\log \mathcal{L}(\hat{\lambda} \pm \sigma^2_{\hat{\lambda}}) = \log \mathcal{L}_{\max} - \frac{1}{2} \tag{8.12}$$

In other words, a change in the parameter of one standard deviation away from the maximum likelihood estimate leads to a decrease of half a unit of log-likelihood.

8.2 Development of a maximum likelihood fit for $B_s^0 \rightarrow J/\psi\phi$

The task in hand is to develop a maximum likelihood method which will provide estimators for the physically important quantities asssociated with the decays of B_s^0 mesons to the final state $J/\psi(\mu\mu)\phi(KK)$. In all cases a single data-point will consist of three angles as described in earlier chapters, and the proper decay time of the B_s^0 meson. We assume that the data points are randomly distributed according to the probability density functions given by tables 6.1 and 6.2. As mentioned previously, no physics model yet proposed modifies the structure of this pdf. The parameters themselves may be considerably modified, but the pdf, derived as it is from simple spin physics, should remain unaltered. This is of course an essential pre-requisite for any kind of statistical fit - the structure of the pdf must be fixed.

8.2.1 General comments

Probability density functions

The complete expressions 6.1 and 6.2 contain eight independent physical parameters. These are

- The magnitudes of the transversity amplitudes $|A_\perp(t = 0)|$ and $|A_\parallel(t = 0)|$ (recalling that only two of the three are independent)

- The associated strong phases δ_1 and δ_2

- The decay widths of the mass eigenstates $B^0_{s,H}$ and $B^0_{s,L}$, Γ_H and Γ_L which are re-cast into the difference $\Delta\Gamma_s$ and the average Γ_s
- The difference between the masses of $B^0_{s,H}$ and $B^0_{s,L}$, ΔM_s (this determines the rapidity of the oscillations).
- The weak phase ϕ_s

Ideally, all eight parameters should be fitted simultaneously. However, as was discussed in chapter 6, some of the parameters can be ascertained from other decay processes and treated as fixed quantities in this fit. These approaches have been studied for some cases [122]; that study is extended here. The parameter ΔM_s was not investigated at all here; the rapidly oscillating terms with which this is associated in the probability density function do not permit this to feasibly form part of the fit [122], and besides, the parameter can be accessed through hadronic channels such as $B_s \rightarrow D_s^- \pi^+$ [121]. Indeed, a simultaneous fit of these two channels will ultimately be necessary. For this exercise, however, the total number of parameters under consideration is reduced to seven - ΔM_s is fixed.

Tagging

As was discussed in chapter 6, the pdf for the angular distribution followed by a given decay depends on the initial state (either B^0_s or $\bar{B}^0_s$) of the decaying B-meson before it begins to oscillate. To determine the parameters with the best precision, both pdfs (tables 6.1 and 6.2) should be used, switching between one and the other according to the flavour tag associated with the event. However, as we have seen from chapter 7, the tags often fail, and where a tag succeeds considerable uncertainty is attached to it. In terms of the fit, therefore, two approaches are possible:

1. The tagging information is not used, and the two pdfs are added togther and treated as one. In this case, all terms of the form $2e^{-\Gamma_s t}\sin(\Delta M_s t)\sin\phi_s$ in tables 6.1 and 6.2 cancel. This simplifies the pdf considerably but such studies fail to take advantage of all the information in the distribution. Table 8.1 shows the simplified probability density function which results from the summing of the state and its Hermitian conjugate.
2. The tagging information is used, with a wrong-tag fraction ϵ appended to the alternate pdf to account for the possibility that the meson is in fact of the opposite flavour. Where no tag exists, the two pdfs are combined as in case 1 above (so both wrong-tag fractions are set to 0.5). The wrong-tag fractions and tagging efficiencies have been determined in chapter 7.

The first technique was investigated in [122]; both will be studied here.

Table 8.1: The probability density function for an untagged sample

k	$\Omega^{(k)}(t)$		$g(t)$
1	$\|A_0(t)\|^2$		$4\sin^2\theta_1\cos^2\theta_2$
	$\frac{1}{2}\|A_0(0)\|^2$	$(1+\cos\phi_s)e^{-\Gamma_L^{(s)}t}+$ $(1-\cos\phi_s)e^{-\Gamma_H^s t}$	
2	$\|A_\parallel(t)\|^2$		$(1+\cos^2\theta_1)\sin^2\theta_2 - \sin^2\theta_1\sin^2\theta_2\cos 2\phi$
	$\frac{1}{2}\|A_\parallel(0)\|^2$	$(1+\cos\phi_s)e^{-\Gamma_L^{(s)}t}+$ $(1-\cos\phi_s)e^{-\Gamma_H^s t}$	
3	$\|A_\perp(t)\|^2$		$(1+\cos^2\theta_1)\sin^2\theta_2 + \sin^2\theta_1\sin^2\theta_2\cos 2\phi$
	$\frac{1}{2}\|A_\perp(0)\|^2$	$(1-\cos\phi_s)e^{-\Gamma_L^{(s)}t}+$ $(1+\cos\phi_s)e^{-\Gamma_H^s t}$	
4	$\Re\{A_0^*(t)A_\parallel(t)\}$		$-\sqrt{2}\sin 2\theta_1\sin 2\theta_2\cos\phi$
	$\frac{1}{2}\|A_0(0)\|\|A_\parallel(0)\|\cos(\delta_2-\delta_1)$	$(1+\cos\phi_s)e^{-\Gamma_L^s(t)}+$ $(1-\cos\phi_s)e^{-\Gamma_H^{(s)}t}$	
5	$\Im\{A_\parallel^*(t)A_\perp(t)\}$		$2\sin^2\theta_1\sin^2\theta_2\sin 2\phi$
	$\|A_\parallel(0)\|\|A_\perp(0)\|$	$\left(e^{-\Gamma_H^{(s)}t}-e^{-\Gamma_L^{(s)}t}\right)\cos\delta_1\sin\phi_s$	
6	$\Im\{A_0^*(t)A_\perp(t)\}$		$\sqrt{2}\sin 2\theta_1\sin 2\theta_2\sin\phi$
	$\|A_0(0)\|\|A_\perp(0)\|$	$\left(e^{-\Gamma_H^{(s)}t}-e^{-\Gamma_L^{(s)}t}\right)\cos\delta_2\sin\phi_s$	

Background

The capacity of the analysis algorithm for rejecting the background processes has been studied in chapter 7. The possibility that a given event might not be the signal of interest but a background event must be accounted for in the fit. As has been discussed, one of the main exclusive background channels to $B_s^0 \to J/\psi(\mu\mu)\phi(KK)$ is $B_d^0 \to J/\psi(\mu\mu)K^{0*}(K\pi)$, which is also a $S \to VV$ decay. In consequence a component of the background will have a non-trivial angular distribution. Ideally this angular structure in the background should appear in the fit, but it represents a considerable complication and has not been included for this study. Instead the approach of [122] is followed; the background is assumed to have the form $be^{-\Gamma_0 t}$, where Γ_0 is the mean decay width for all the B-mesons participating in the background; b is the level of background, determined in chapter 7.

Detector Resolution

The resolution of the measurements of proper decay time and decay angles made by ATLAS have been studied in chapter 7. These effects should be included in the fit. The importance of this smearing effect on the estimators depends on how steeply the pdf changes with respect to the variable in question. The FWHM angular resolutions (determined in chapter 7; figure 7.3) are of excellent precision, and given the slow variation in the angular distribution (see the plots of the angles in chapter 6) any smearing effects are negligible, so the angular data from the detector is taken to be "perfect". In contrast the proper decay time resolution (figure 7.2) is poor in comparison with the rapidly changing terms in tables 6.1 and 6.2, so the smearing of the decay time must be accounted for in the fit.

It is assumed that the smearing effect is Gaussian with a standard deviation equal to the proper time resolution of the detector as determined in chapter 7. To include these effects in the fit, the entire pdf (that is, both the signal and the background terms) must be convolved with a Gaussian function having this standard deviation. The convolution of a function with a Gaussian is given by

$$\begin{aligned} C\left(t'\right) &= f\left(t\right) \bigotimes \rho\left(t'-t\right) \\ &= \int_0^\infty f\left(t\right) \rho\left(t'-t\right) dt \end{aligned} \tag{8.13}$$

where

$$\rho\left(t'-t\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\left(t'-t\right)^2}{2\sigma^2}} \tag{8.14}$$

The resulting function is a smeared or "blurred" representation of the original expression. This function gives a truer representation of the reconstructed data, and is thus adopted in the fit.

Acceptance Corrections

Even if the behaviour of the decaying particles is exactly described by the probability density function, the measured angular and time distributions will be distorted by trigger selections and the limitations of the detector itself. It is to be expected that this distortion will be a function of the decay angles and the proper decay time, giving a *reconstruction efficiency* $\varepsilon\left(\theta_1, \theta_2, \phi, t\right)$. The measured angular distributions should be corrected according to this efficiency before being processed by the fitting routine.

The reconstruction efficiencies were studied in chapter 7. As can be seen from figures 7.4 and 7.5, they are not in fact strongly dependent on the angles and decay times, maintaining a virtually flat profile. The effects would therefore seem to be marginal in this case, and given the considerable effort required to perform the acceptance corrections properly, this was not attempted for this study.

Normalization

The pdf must be normalized such that the integral over the entire phase space is unity. In practice this means that the pdf, once constructed, must be divided by the integral over all variables, over all space. In this case time runs from the minimum decay time permitted in the event selection to infinity. The angle ϕ runs from zero to 2π and θ_1 and θ_2 run from zero to π.

8.2.2 Construction of the Likelihood function

Heeding all the remarks made above, the following is adopted as the likelihood function

$$\mathcal{L} = \prod_{i=1}^{N} \frac{\left(\epsilon_1 W^+\left(t', \Omega_i\right) + \epsilon_2 W^-\left(t', \Omega_i\right) + be^{-\Gamma_0 t'}\right) \otimes \rho\left(t_i - t'\right)}{\int_{t_{min}}^{\infty} \int_0^{2\pi} \left(\epsilon_1 W^+\left(t', \Omega\right) + \epsilon_2 W^-\left(t', \Omega\right) + be^{-\Gamma_0 t'}\right) \otimes \rho\left(t - t'\right) d\Omega dt} \tag{8.15}$$

- N is the total number of data points (events).
- W^+ and W^- are the probability density functions of the decay angles and times for a B_s^0 and a $\bar{B}_s^0$ respectively (tables 6.1 and 6.2).
- ϵ_1 and ϵ_2 are the tagging terms. For untagged events, $\epsilon_1 = \epsilon_2 = 0.5$. Where the meson is tagged as being in the particle state at $t = 0$, $\epsilon_1 = 1 - w$ and $\epsilon_2 = w$ where w is the wrong-tag fraction as investigated in chapter 7. For the case where the meson is tagged as being initially in the anti-particle state, $\epsilon_1 = w$ and $\epsilon_2 = 1 - w$.
- b is the level of background as determined in the preceding chapter.
- Γ_0 is the average decay width for the background.
- t_{min} is the minimum proper decay time permitted in the event selection, which is $0.5ps$ for these events.
- Ω represents the three decay angles θ_1, θ_2, ϕ.
- ρ represents the Gaussian function with proper decay time resolution σ, as determined in chapter 7.

The *indexed* variables t_i and Ω_i are the data points themselves. During the fitting procedure they are discrete - each ith event has a single values for t_i and Ω_i. However during the convolution the decay time t_i is treated as a constant, whilst the convolution integral is performed over t'. The end result is that the convolved function contains only indexed variables; after convolution the t_i takes the place of the t'. In line with the discussion above, no convolution with an angular resolution is performed as this value is small in comparison with the angular functions. In consequence the Ω_i appear "as is" whilst t_i appear only in the convolution function.

The denominator provides the normalization; here there are only integration variables, so the indices disappear altogether. The convolution still has to be performed, and clearly must take place *first* - the normalization integral must be over the complete smeared expression.

8.2.3 Practical Details

The fitting was performed by the Minuit minimization program. For each event the minimization was performed by the routine `Migrad`. The data provided to the framework consisted of sets of events each containing three angles and a decay time. The fitter was provided with starting values, and step sizes which were set to 10% of the starting value. Also provided were the detector resolution σ, the wrong-tag fraction w, the background level b and and the mean decay time Γ_0. For each set of fit parameters selected by Minuit, the log of the likelihood is calculated according to equation 8.15, such that the product becomes a sum. The normalization is performed once per set of parameters, but the pdf must be evaluated for every event (in other words, N times per Minuit step, where N is the number of events). As Minuit is a minimization routine the log-likelihood result must be multiplied by -1 before being passed back to the fitter.

The **convolution** and **normalization** each require the effectuation of integrals where one of the integration variables, namely time, runs to infinity. In the case of the convolution the integral is one-dimensional (over time) whereas the normalization requires a four-dimensional integration over time and three angles. The integration over the angles is trivial and a simple analytical expression was provided. For the time-dependent parts, a numerical algorithm (**adaptive Gaussian quadrature** [123]) was used. This was already implemented in ROOT (based on older CERNLIB Fortran routines [124]) and so, with the exception of slight structural modifications, no changes were made in the algorithms themselves.

In all, one normalization integration must take place per Minuit step, but *every* evaluation of the pdf requires several convolution integrals, as time-dependent terms appear throughout the pdf. It can be seen that a very large number of these convolution integrals must be performed during a single Minuit step. With this in mind analytical forms for the convolution integrals were sought for all terms in the pdf. These were taken from standard integral tables.

In all cases the time integrals run from zero or t_{min} to ten times the mean B_s^0 lifetime, convergence having been achieved by this time.

It is necessary to worry about **machine precision**, as each Minuit step involves thousands of calculations and rounding errors can mount up and have ruinous consequences for the fit. An appropriate system of units must be chosen such that most of the quantities being handled by Minuit are as close to 1 as consistency will allow. Obviously these units must be *natural* (e.g. $\hbar = c = 1$) so that all quantities have the dimensions of energy or inverse energy. As previously discussed, it is quickly evident that the best measure to use is the milli-electron-volt meV.

For the purposes of the fit, the amplitudes A_0, $A_\perp$ and $A_\parallel$ were recast into two parameters:

$$r_\perp = \frac{|A_\perp(0)|}{|A_0(0)|} \tag{8.16}$$

$$r_\parallel = \frac{|A_\parallel(0)|}{|A_0(0)|} \tag{8.17}$$

8.3 Monte Carlo studies

The maximum likelihood apparatus was applied to directly produced Monte Carlo data (that is, sets of decay angles and times generated with an accept/reject method using equations 6.1 and 6.2 as the pdf). The proper decay time for each of these Monte Carlo "events" was smeared randomly according to a Gaussian with $\sigma = 87fs$, corresponding to the measured resolution of the detector as established in the preceding chapter (the angles were left untouched). The maximum probability for the accept-reject algorithm was selected manually, to avoid having to perform a time consuming Monte Carlo maximization.

A range of tests were performed with the aim of gauging the performance of the fits under a variety of possible analysis conditions. Of particular interest is the *precision* with which the parameters can be measured, and the *correlations* between the parameters. In principal all seven parameters are fully independent, but the finite resolution of the data (modelled by the Gaussian smearing) is expected to introduce correlations between the parameters. Some of these correlations are expected, from previous unpublished studies, to be large. The precision is expected to depend on the number of data points (that is, the number of events), the level of background, the tagging efficiency and the proper decay time resolution.

A similar study was performed in [122], where the precision on the measurements of $\Delta\Gamma_s$, Γ_s, $A_\perp$ and $A_\parallel$ was investigated for *untagged* events. In that paper, the remaining four parameters were treated as constants.

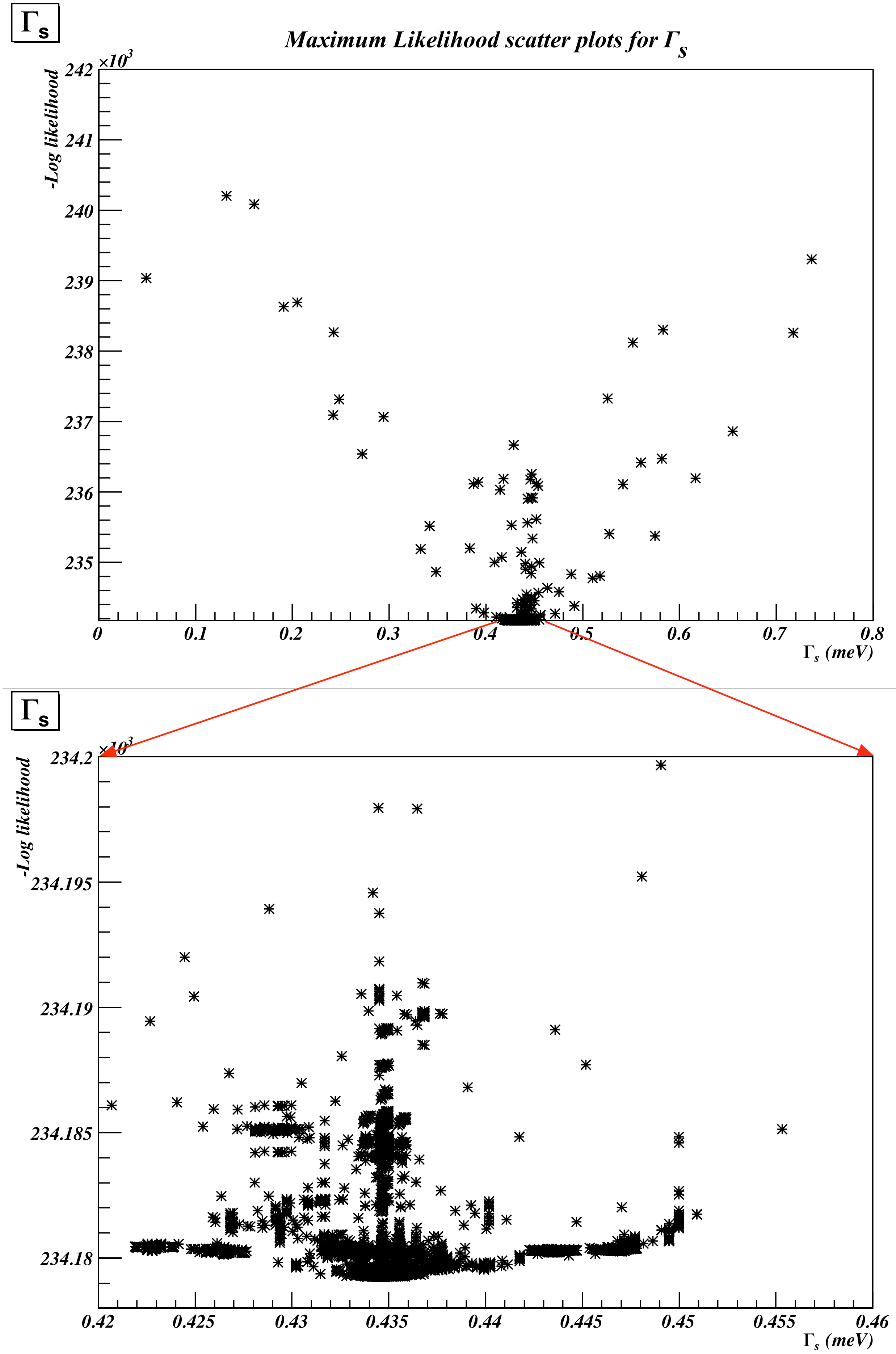
8.3.1 Fit validation

In order to ensure the validity of the fitting program, single fits were run under a range of conditions:

- Untagged data generated with the Standard Model parameter set, ϕ_s and ΔM fixed
- Tagged data generated with the Standard Model parameter set, ΔM fixed
- Tagged data generated with the 50% new physics parameter set, ΔM fixed

In each case the values for the parameters tried by Minuit were plotted against the overall value of the likelihood function as scatter graphs. Given that all the other parameters are varying at the same time as the parameter in question, such plots take into account the correlations between the parameters. This is in contrast to the normal scan plots, which fix all parameters other than the one under consideration.

The scatter plots are shown in figures 8.1 through to 8.5. Each of these plots was produced with a small sample of 10000 events. Untagged events were used to produce the plots for all the parameters, with the exception of the weak phase ϕ_s, for which tagged events were used. It can be seen that Minuit converges to values close to the input parameters for all cases, *except* for ϕ_s. It can be seen from figure 8.5 that for data produced with small values of ϕ_s (as per the Standard Model), the code does not reliably converge to any single value. If the larger values of ϕ_s predicted by the new physics models are used, the code converges more reliably. This is in line with previous ATLAS studies - if the weak phase ϕ_s is very small, ATLAS will not be able to measure it with this channel.



Γs
Maximum Likelihood scatter plots for Γs
-Log likelihood
×10^3
242
241
240
239
238
237
236
235
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Γs (meV)
Γs
-Log likelihood
×10^3
234.2
234.195
234.19
234.185
234.18
0.42
0.425
0.43
0.435
0.44
0.445
0.45
0.455
0.46
Γs (meV)

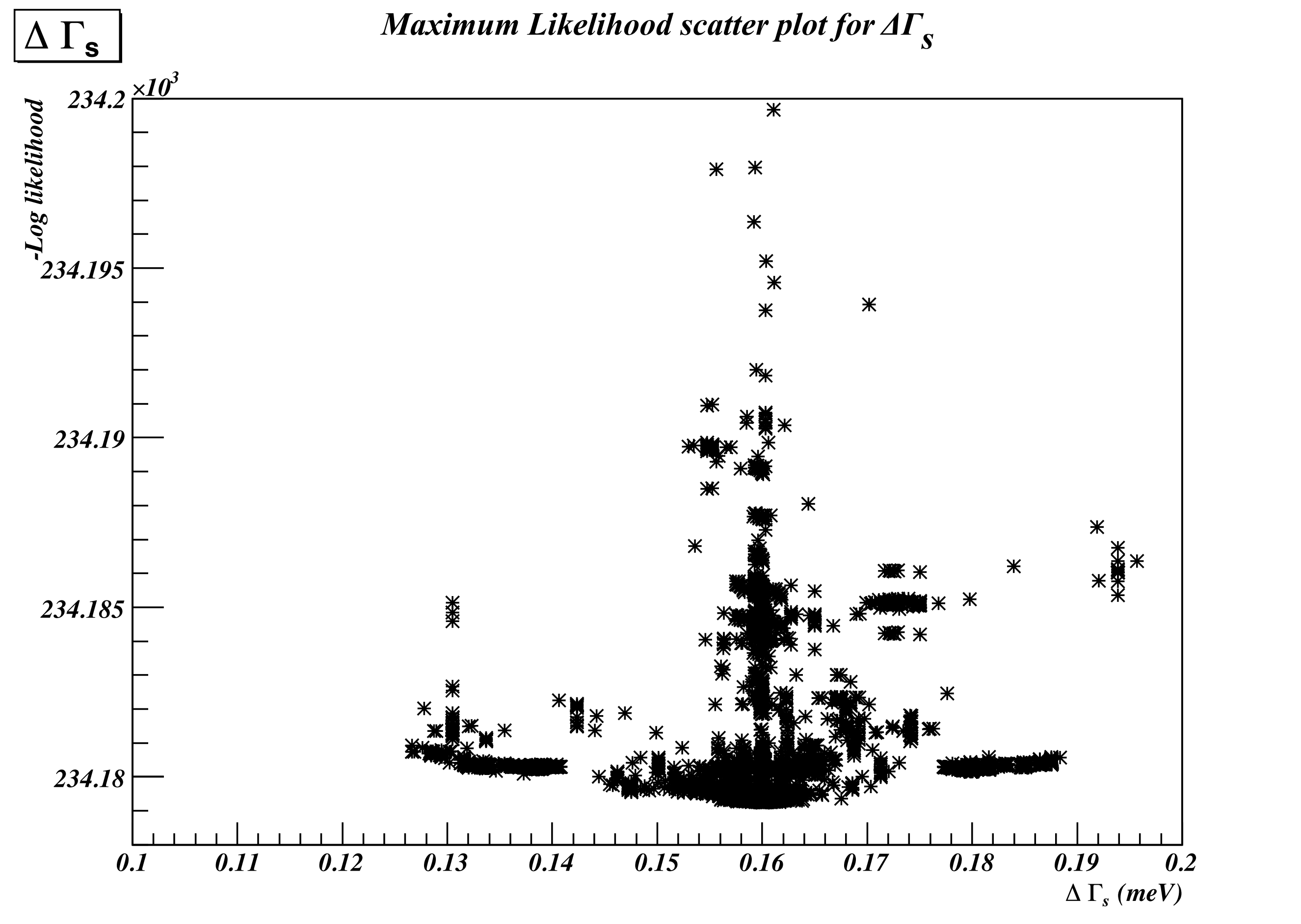
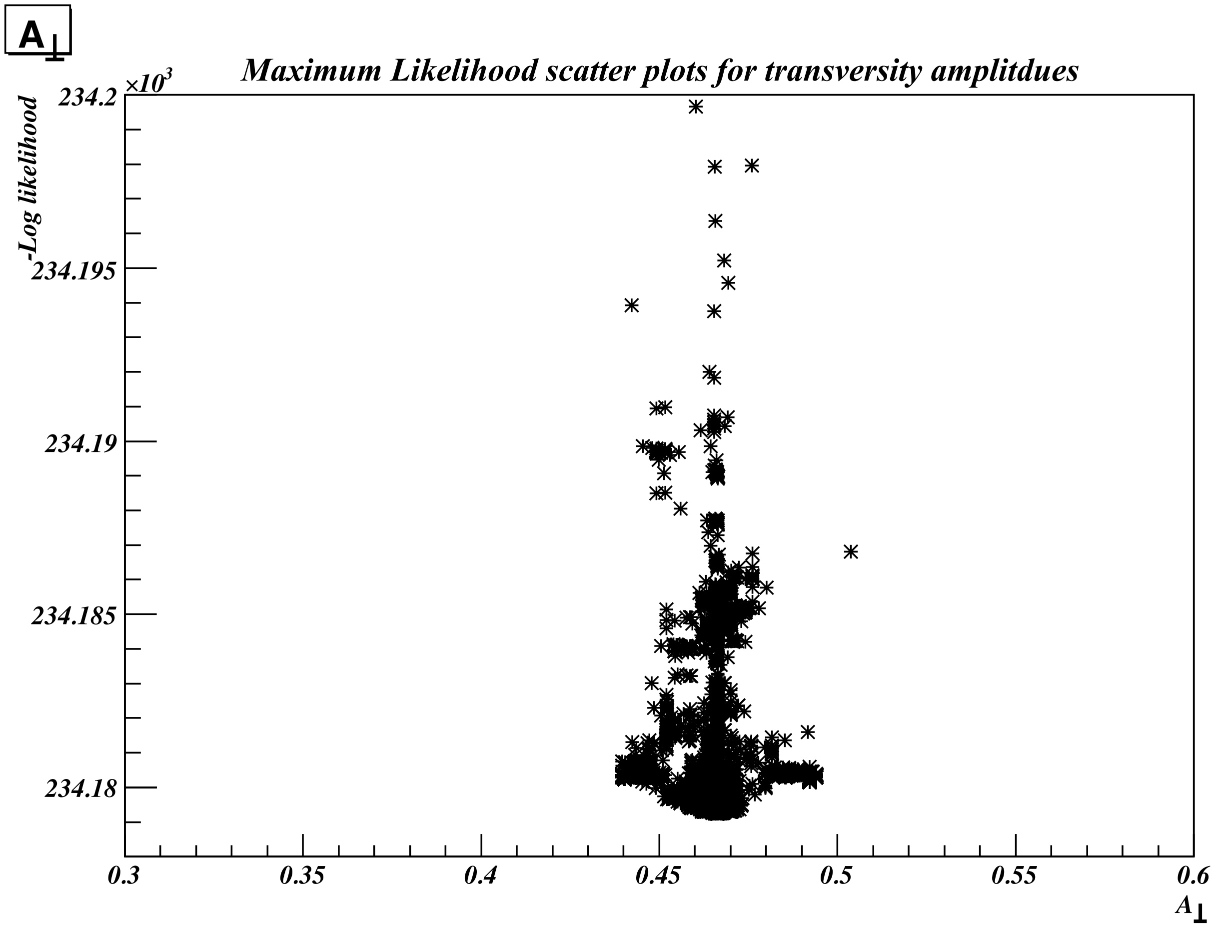
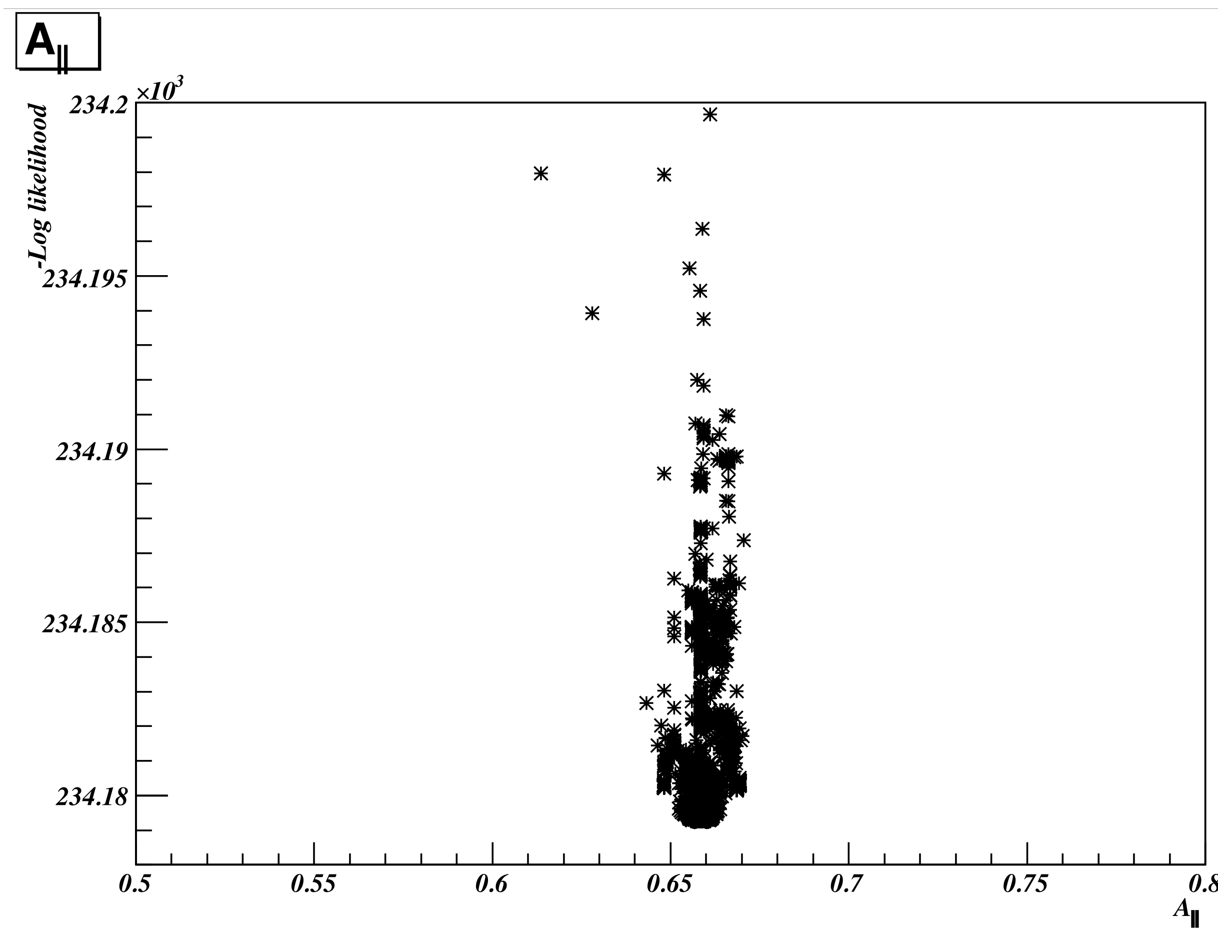
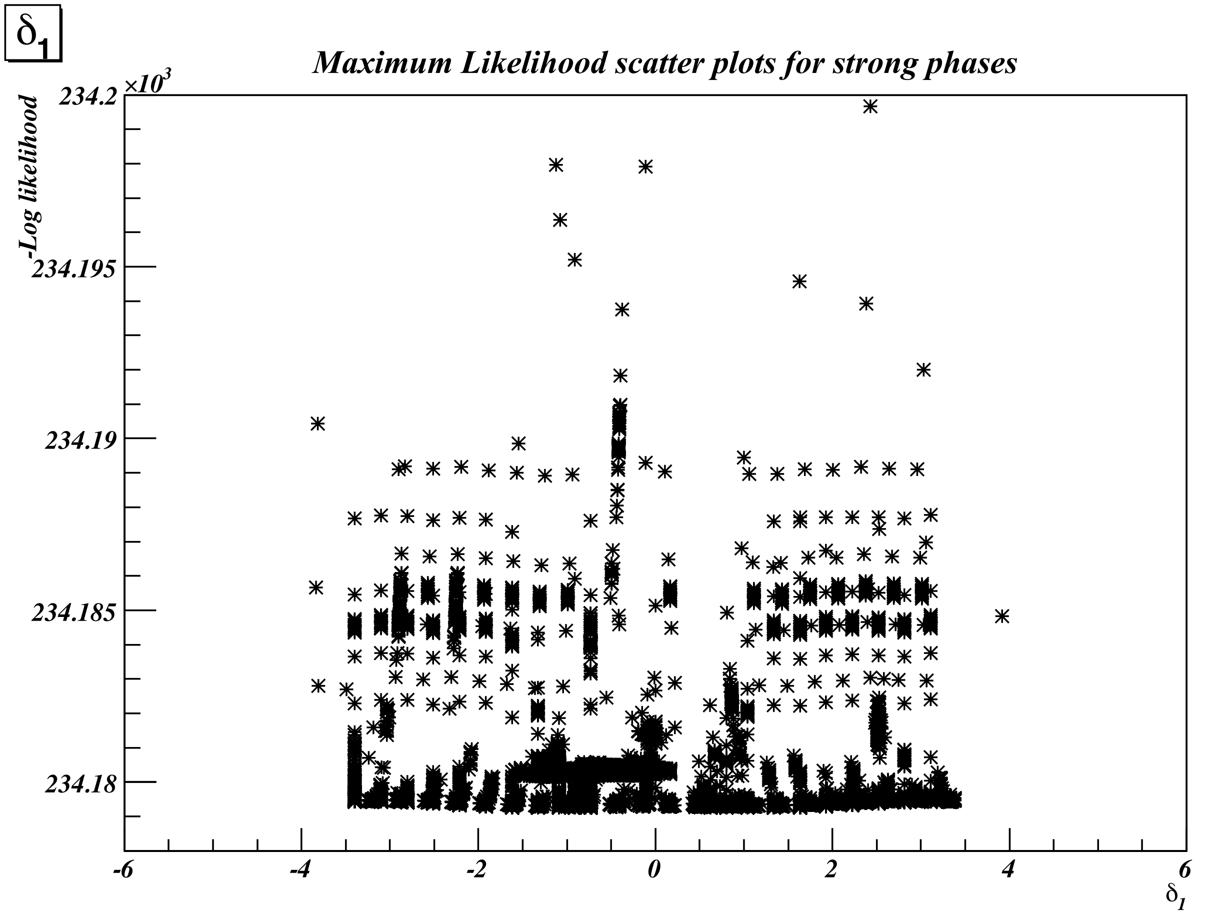
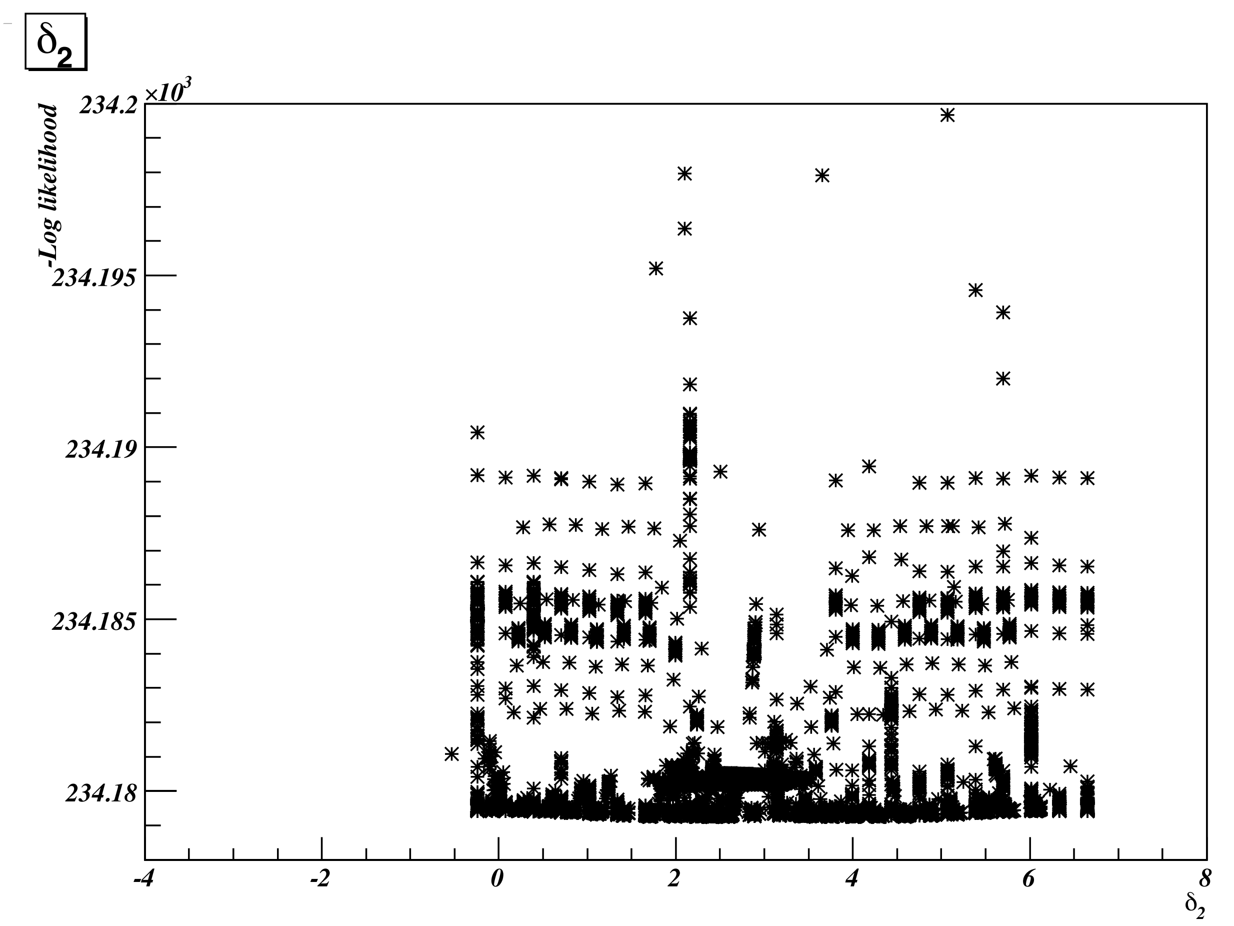


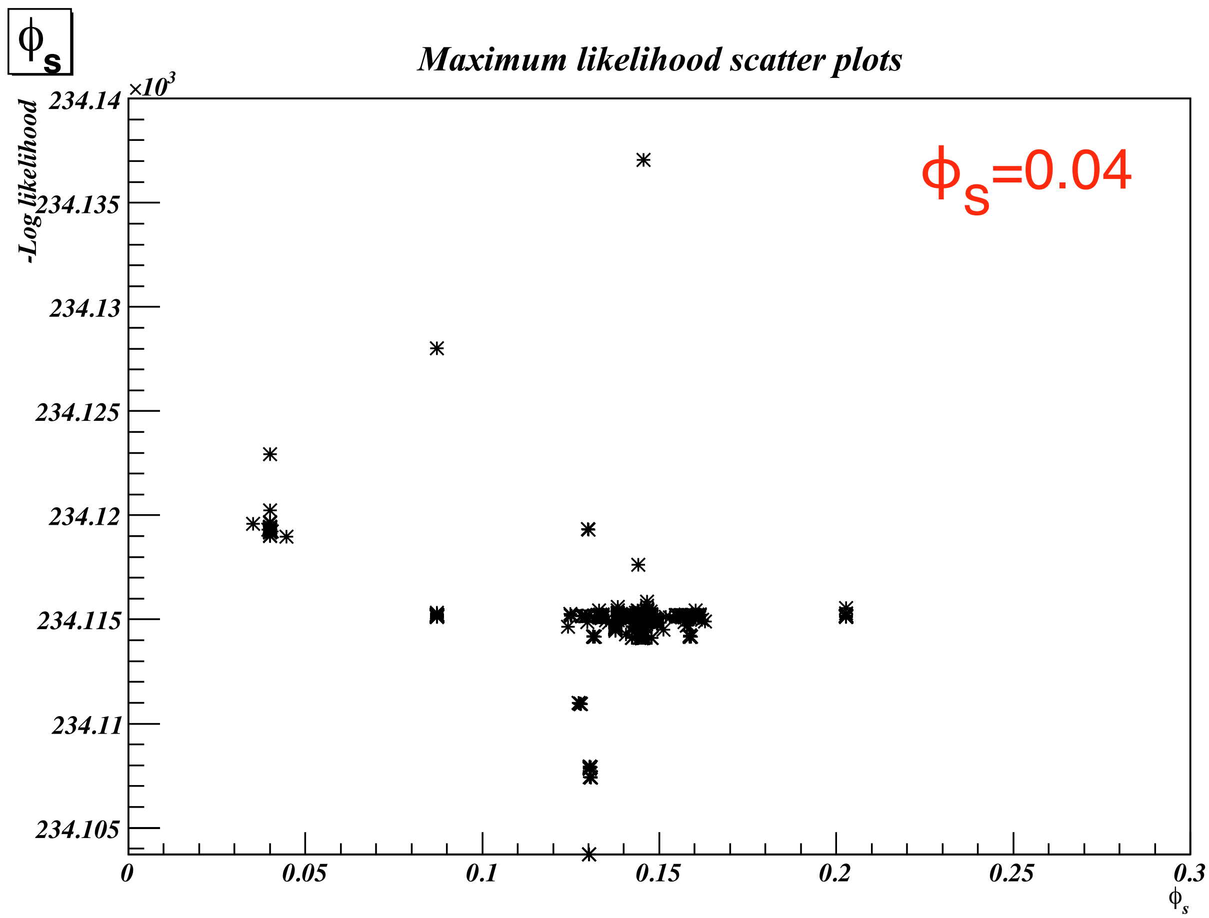
Figure 8.2: Scatter plot of values of $\Delta\Gamma_s$ tried by Minuit versus likelihood during a fit to untagged data produced with the Standard Model dataset (10000 events). ΔM and ϕ_s fixed; all other parameters free.


A⊥
Maximum Likelihood scatter plots for transversity amplitdues
-Log likelihood
×10^3
234.2
234.195
234.19
234.185
234.18
0.3
0.35
0.4
0.45
0.5
0.55
0.6
A⊥

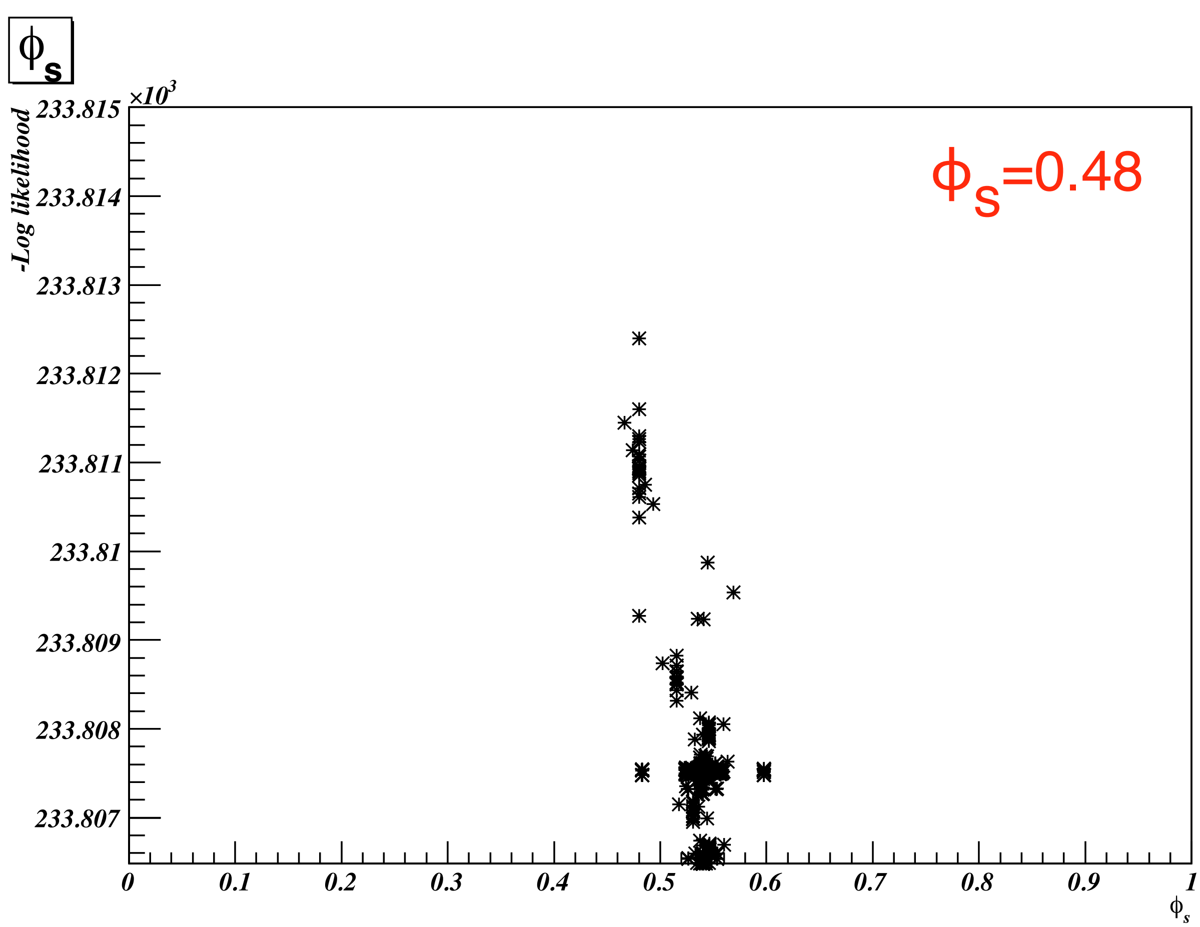

A∥
-Log likelihood
×10^3
234.2
234.195
234.19
234.185
234.18
0.5
0.55
0.6
0.65
0.7
0.75
0.8
A∥


δ1
Maximum Likelihood scatter plots for strong phases
-Log likelihood
×10^3
234.2
234.195
234.19
234.185
234.18
-6
-4
-2
0
2
4
6
δ1


δ2
-Log likelihood
×10^3
234.2
234.195
234.19
234.185
234.18
-4
-2
0
2
4
6
8
δ2



ϕ_s
Maximum likelihood scatter plots
$\times 10^3$
-Log likelihood
234.14
234.135
234.13
234.125
234.12
234.115
234.11
234.105
0
0.05
0.1
0.15
0.2
0.25
0.3
ϕ_s
ϕ_s=0.04



ϕ_s
$\times 10^3$
-Log likelihood
233.815
233.814
233.813
233.812
233.811
233.81
233.809
233.808
233.807
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
ϕ_s
ϕ_s=0.48

8.3.2 Untagged event fits

The full maximum likelihood apparatus was first used to generate, smear and fit parameters to *untagged* data (generated using table 8.1 as the probability density function). The weak phase ϕ_s cannot be accessed with such data and for these exercises was fixed.

Previous studies [122] demonstrated that better precision on the remaining values could be attained by making the assumption of "exact $SU(3)$ symmetry", where it was taken that the transversity amplitudes measured using equivalent decays in the B_d system *identical* to those of $B_s \to J/\psi\phi$. Fixing these amplitudes in the $B_s \to J/\psi\phi$ fit to the B_d values gives the increase in precision in the remaining parameters. However, as has been discussed in chapter 6 there is scant experimental evidence to support this symmetry - the error bars on these amplitudes are currently too wide to make any informed judgement - so in this work no symmetry is assumed and the amplitudes are allowed to float freely in the fit.

In each case, fits are performed for five data sample sizes, corresponding to 1, 2, 5, 10, 20 fb^{-1}. The sample sizes, estimated from the conclusions of chapter 7, are given in table 8.2. For each size of sample, 100 separate sets of data (or "experiments")

Table 8.2: Sample sizes used in untagged maximum likelihood studies

Integrated luminosity, fb^{-1}	1	2	5	10	20
Number of reconstructed events	11500	22500	56500	113000	226000

were generated and fitted to the probability density function of table 8.1, using 8.15 as the likelihood function. On the basis of the results of chapter 7, all time-dependent terms were smeared by a Gaussian function with a resolution of $90fs$. The level of background was assumed to be 22%, corresponding to the levels of the exclusive background from chapter 7 [1]. The Standard Model parameter set was used to generate the sample for each experiment. The weak phase ϕ_s and the oscillation frequency ΔM_s were fixed at their Standard Model values (this is a potential source of systematic uncertainty and will be discussed in later pages).

Table 8.3 shows the mean of the relative errors reported by Minuit. Table 8.4 gives the correlations between all pairs of variables as reported by Minuit (averaged over all experiments and sample sizes).

Figures 8.7 through to 8.11 show three types of plot for each of parameters. In each of the plots, the first column of histograms (filled in red) shows the value to

[1] See erratum at the end of this document

Table 8.3: Mean of absolute uncertainties (means of pulls in brackets) reported by Minuit on fit parameters for untagged events

Statistics	11500	22500	56500	113000	226000
$\Delta\Gamma_s$	0.030 (-0.004)	0.019 (0.006)	0.013 (-0.007)	0.010 (0.003)	0.009 (0.004)
Γ_s	0.013 (-0.0002)	0.008 (-0.003)	0.024 (0.0007)	0.004 (-0.001)	0.004 (-0.001)
δ_1	4.434 (-0.327)	3.022 (0.181)	2.051 (0.077)	1.972 (-0.015)	1.824 (-0.014)
δ_2	3.888 (-0.609)	2.649 (-0.569)	2.198 (-1.255)	2.028 (-0.041)	1.948 (0.046)
$r_\perp$	0.036 (0.002)	0.025 (0.007)	0.022 (0.005)	0.010 (0.004)	0.009 (0.004)
$r_\parallel$	0.017 (0.005)	0.017 (0.003)	0.016 (-0.001)	0.008 (-0.002)	0.007 (-0.002)

Table 8.4: Correlations between variables reported by Minuit for untagged fits, averaged over all experiments and sample sizes

	Γ_s	$\Delta\Gamma_s$	$r_\parallel$	$r_\perp$	δ_1	δ_2
Γ_s	1.000	-0.594	-0.247	-0.095	-0.301	-0.295
$\Delta\Gamma_s$	-0.594	1.000	0.348	0.307	0.332	0.329
$r_\parallel$	-0.247	0.348	1.000	-0.201	0.512	0.512
$r_\perp$	-0.095	0.307	-0.201	1.000	-0.259	-0.263
δ_1	-0.301	0.332	0.512	-0.259	1.000	0.785
δ_2	-0.295	0.329	0.512	-0.263	0.785	1.000

which Minuit converges for each of the the 100 experiments. The histograms in the second column (green) plot the uncertainties on each parameter reported by Minuit. The third column displays the **pulls** for each parameter; that is, the difference between the result found by Minuit and the orginal input value that was used to create the Monte Carlo data. Each plot has five rows of histograms, corresponding to the five sample sizes described above.

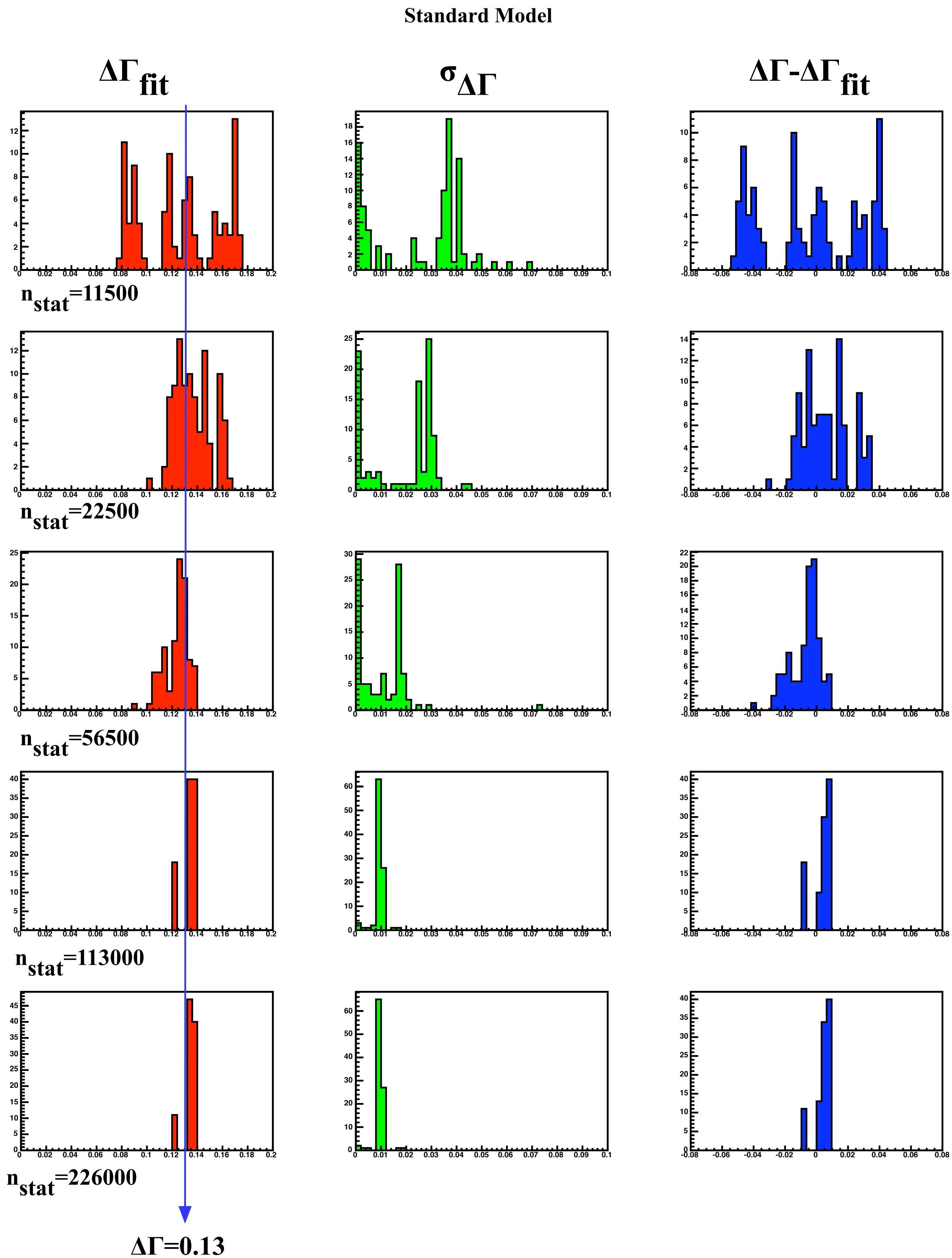


Figure 8.6: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, for the parameter $\Delta\Gamma$, using the Standard Model parameter set.

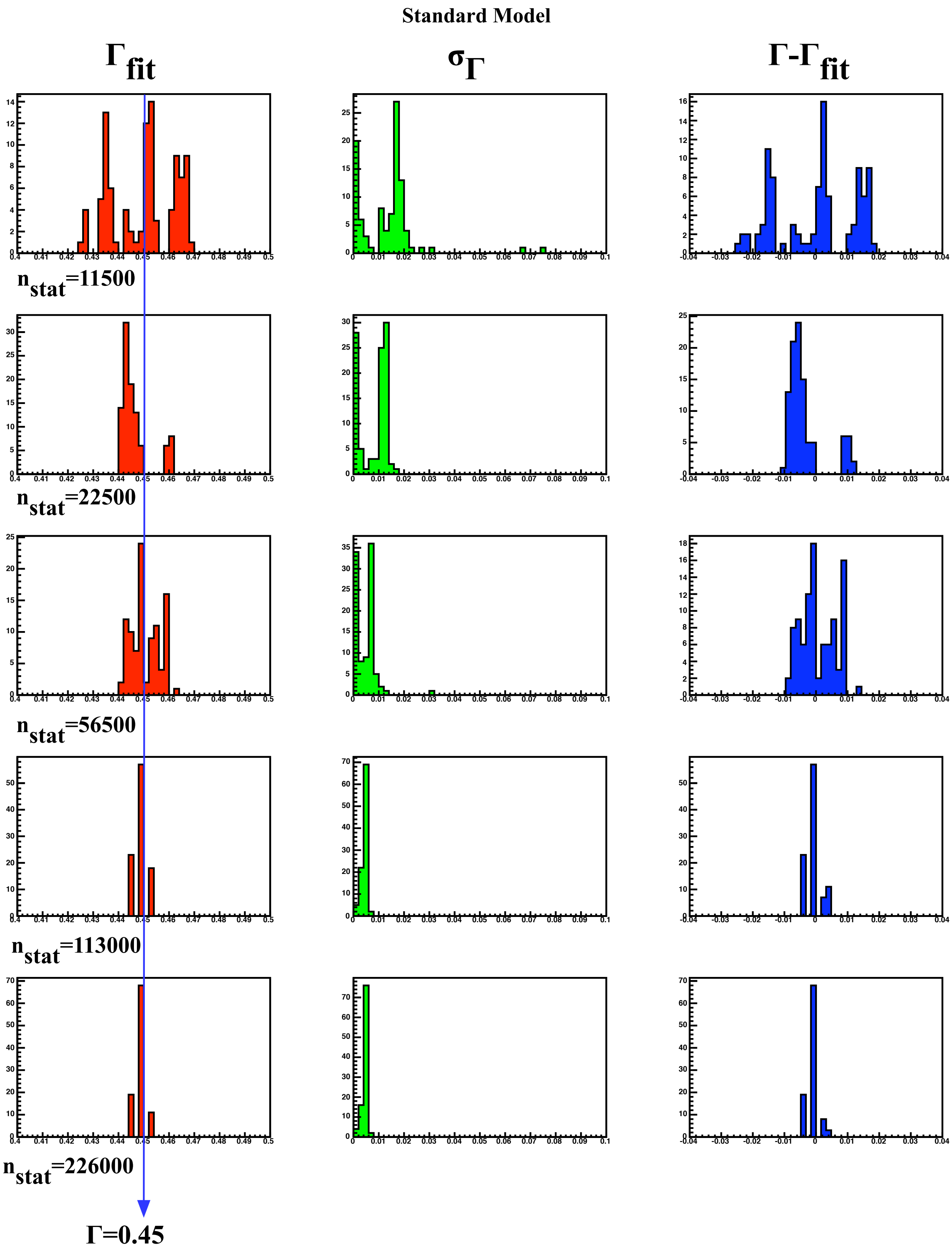


Figure 8.7: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, for the parameter *Gamma*, using the Standard Model parameter set.

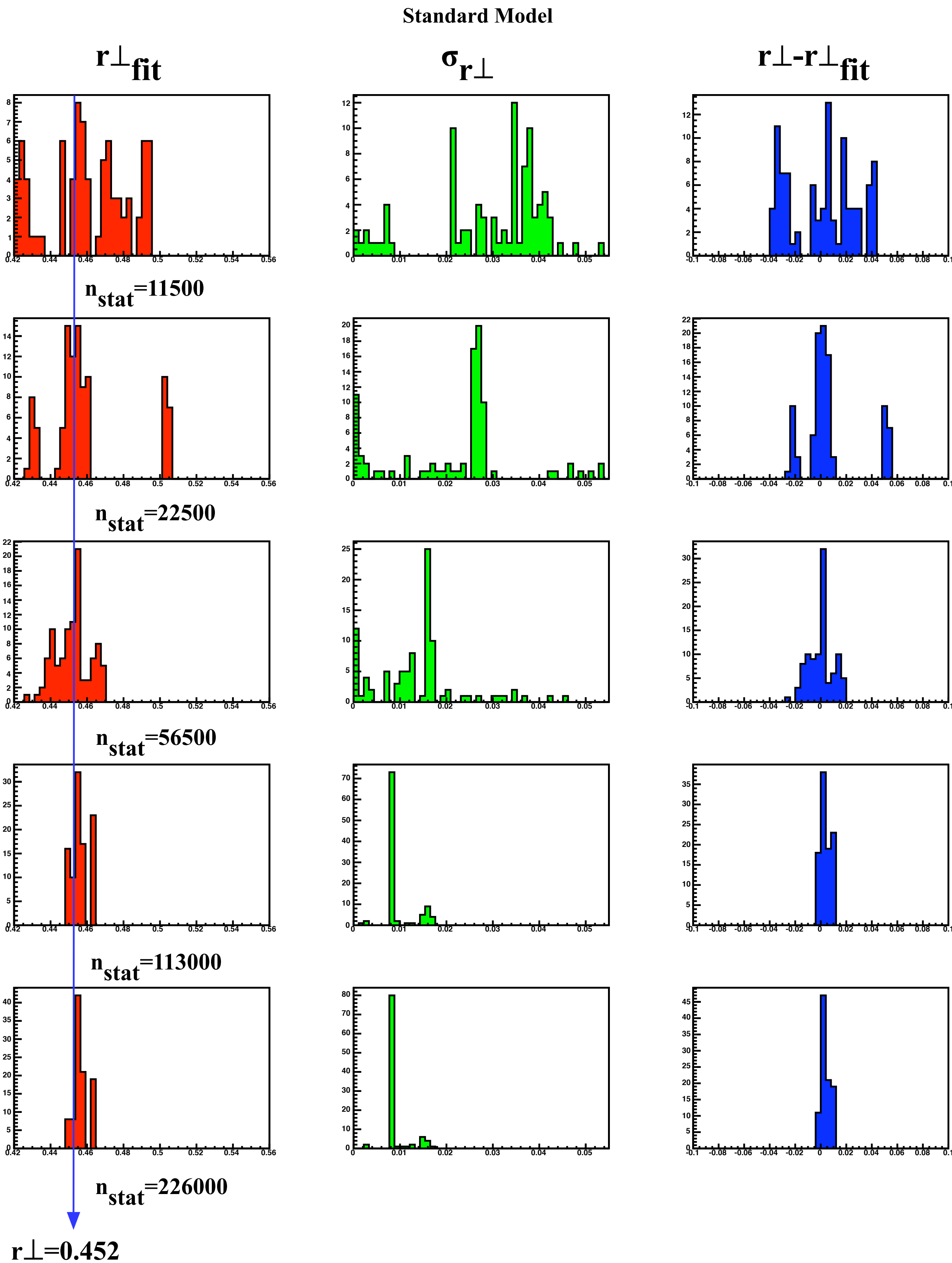


Figure 8.8: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, for the parameter $r_\perp$, using the Standard Model parameter set.

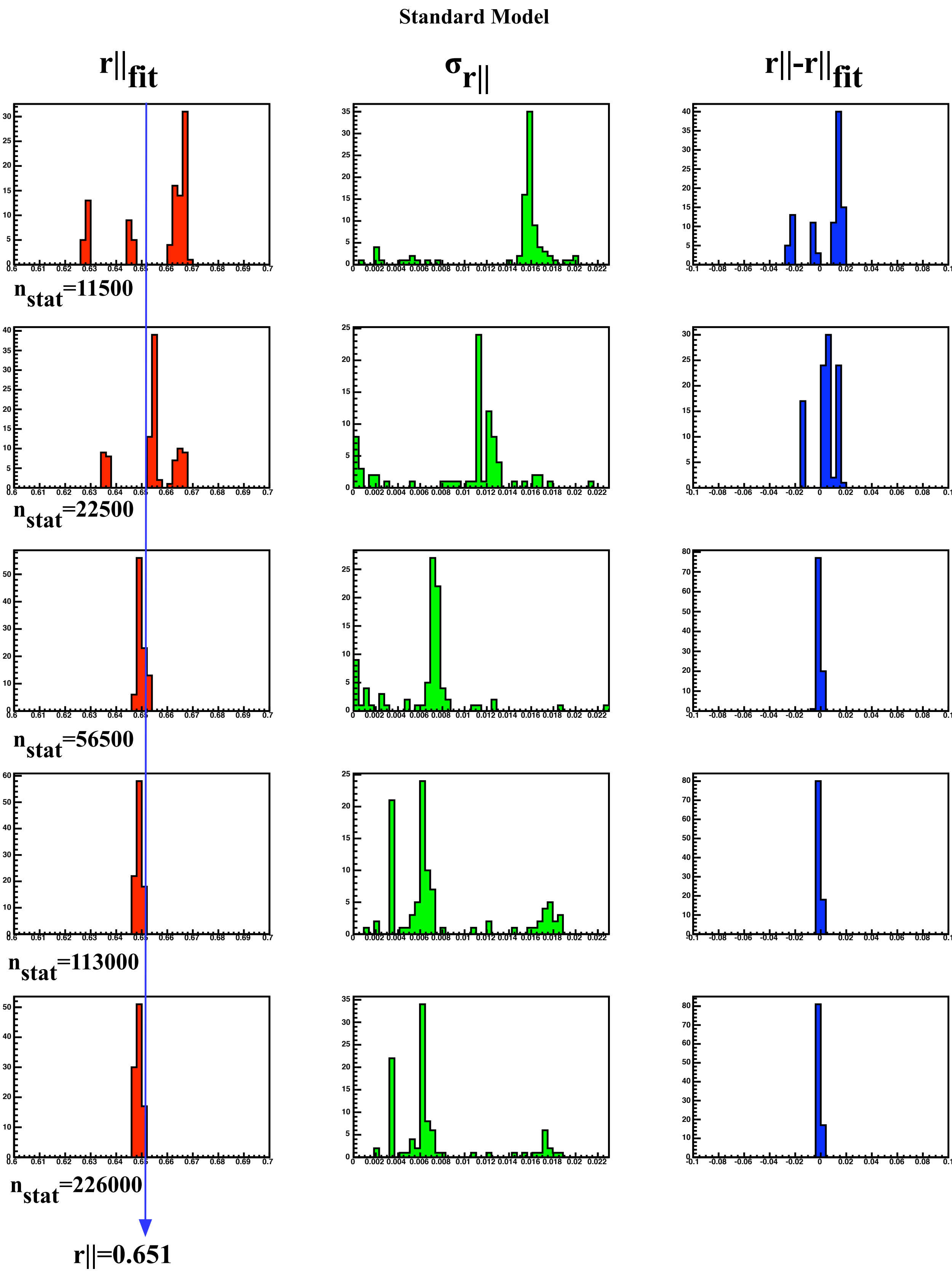


Figure 8.9: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, for the parameter $r_{\parallel}$, using the Standard Model parameter set.

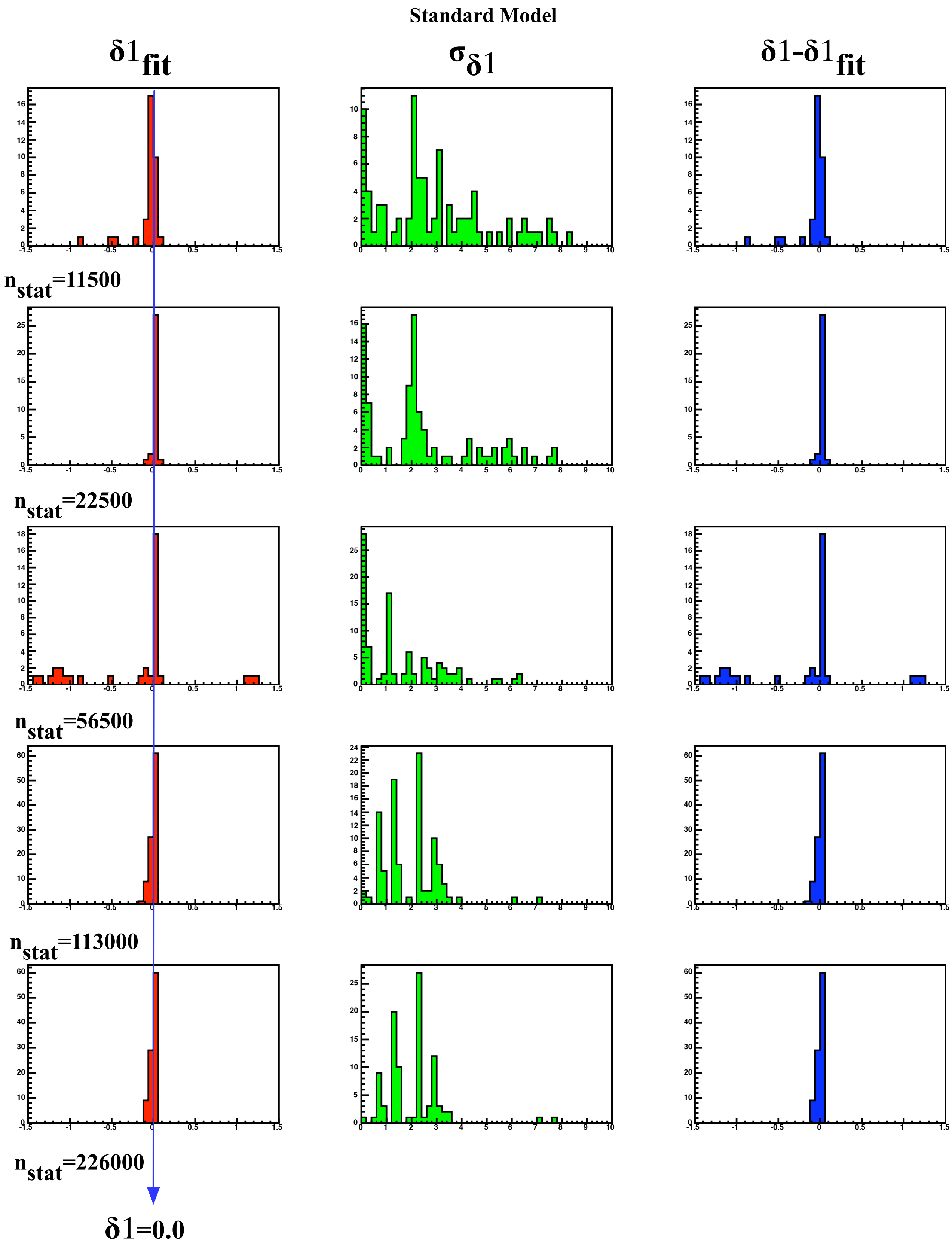


Figure 8.10: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, for the parameter δ_1, using the Standard Model parameter set.

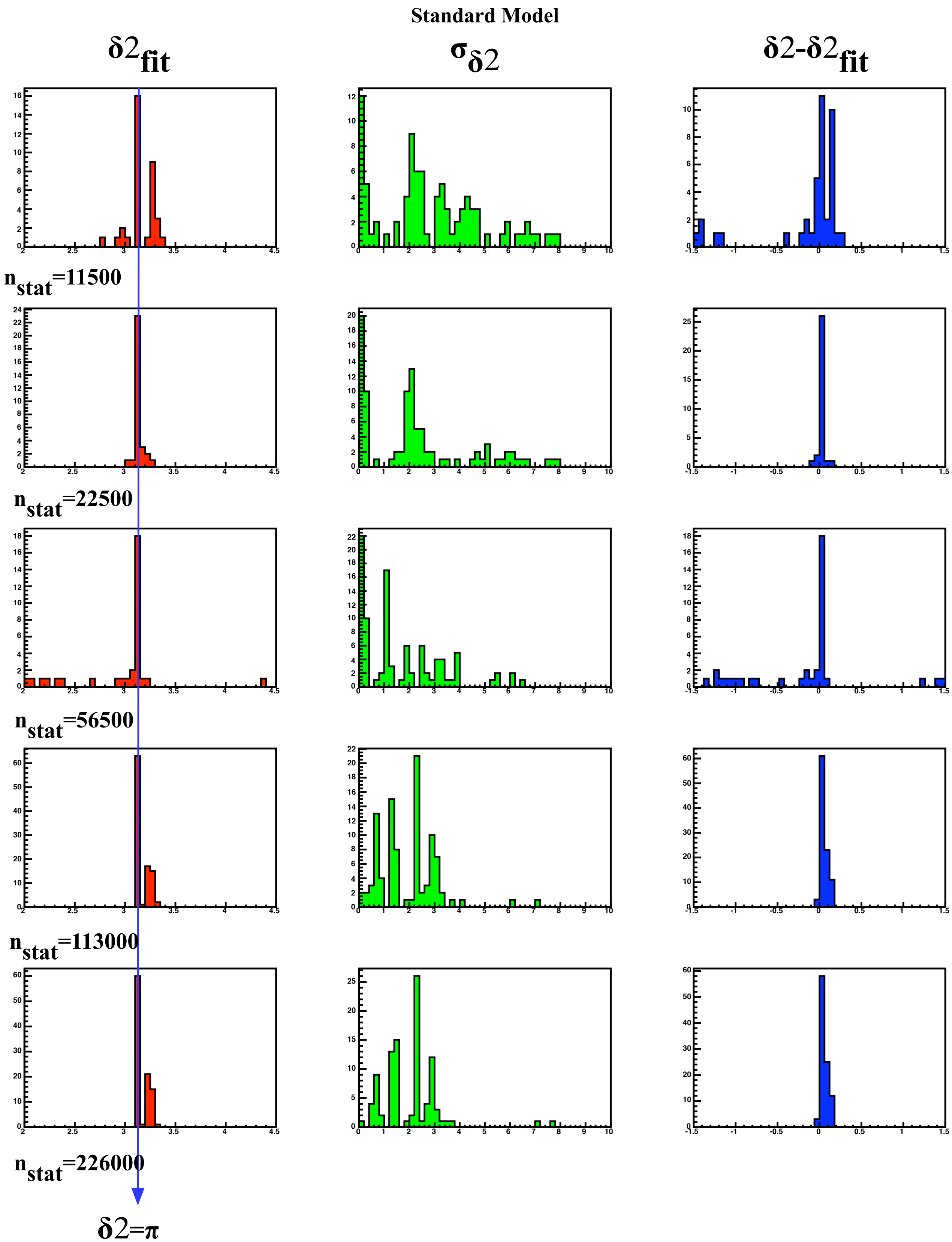


Figure 8.11: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, for the parameter δ_2, using the Standard Model parameter set.

It is evident from the table 8.3 and the plots 8.7 through to 8.11, that the behaviour of the fits is broadly in line with expectations - the values to which Minuit converges approaches the input values used to generate the data as the statistics increase. The reported error also shrinks with increasing statistics, as expected. Additionally, the mean pulls of the Minuit answers (i.e. the difference between the Minuit answer and the input value) are consistently smaller than the Minuit uncertainties. This suggests that the uncertainties reported by Minuit from these fits are trustworthy and can be used as a means of reporting the estimated ATLAS sensitivity to the parameters.

From the table showing the mean correlations between the parameters (8.4), it is apparent that some of the variables are rather strongly correlated, most notably the average lifetime and the lifetime difference of the two B_s states, $\Delta\Gamma_s$ and Γ_s, and the phases of the transversity amplitudes δ_1 and δ_2

In summary, after $20fb^{-1}$ of untagged events, ATLAS can expect to achieve the precisions as shown in table 8.11, if the Standard Model predictions are correct. Relative precisions have been calculated using the "Standard Model" input values.

Table 8.5: Summary of estimated precisions for the Standard Model after $20fb^{-1}$ untagged data

$\Delta\Gamma_s$	0.009 (6.9%)
Γ_s	0.004 (0.8%)
δ_1	1.824
δ_2	1.948
$r_\perp$	0.009 (2.0%)
$r_\parallel$	0.007 (1.1%)

8.3.3 Tagged event fits

To access the weak phase ϕ_s, a *tagged* sample must be used, where the full likelihood expression 8.15 is deployed, with appropriate values selected for the wrong tag fractions. The study was performed as follows:

1. All the B-mesons were assumed to be in the B_s^0 state at time $t = 0$. Whilst this is not physical, it is adequate for the purposes of estimating the precision on the weak phase measurement provided that the reconstruction and tagging efficiencies are the same for B_s^0 and $\bar{B}_s^0$ states

2. The wrong-tag fraction ϵ was fixed at 37%, as calculated in the preceding chapter.

3. The statistics used in the study are given in table 8.6. These figures were calculated on the basis of the tagging efficiency of 51.6% found in the preceding chapter.

Table 8.6: Sample sizes used in tagged maximum likelihood studies

Integrated luminosity, fb^{-1}	1	2	5	10	20
Number of reconstructed events	5900	11500	29100	58300	116600

4. The other experimental variables (background, resolution) were maintained at the values used in the untagged study

5. The fits were performed for two sets of input parameters - corresponding to new physics contributions of 50% and 100% of the Standard Model predictions. The input parameters can be found in the table 6.5.

6. All the parameters were allowed to float freely in the fit, aside of the oscillation parameter ΔM_s, which was fixed at the value appropriate to the physics model under consideration. This is a potential source of systematic error, and is considered later on.

100 experiments were carried out for each sample size, for the two "new physics" scenarios. Tables 8.7 and 8.8 show the mean of the relative errors reported by Minuit for the two physics scenarios. Tables 8.9 and 8.10 gives the correlations between all pairs of variables as reported by Minuit (averaged over all experiments and sample sizes).

Figures 8.19 through to 8.25 show again the Minuit "answer" (red) and the input value, the Minuit error (green) and the pull (blue) for each of the parameters, along with the input value for each parameter, for the two new physics scenarios.

Table 8.7: Mean of absolute uncertainties (and pulls) reported by Minuit on fit parameters for tagged events, for the "50% new physics" model

Statistics	5900	11500	29100	58300	116600
$\Delta\Gamma_s$	0.048 (0.004)	0.035 (−0.015)	0.020 (0.004)	0.017 (−0.003)	0.010 (0.004)
Γ_s	0.021 (−0.005)	0.016 (0.004)	0.009 (−0.003)	0.007 (0.0005)	0.004 (−0.002)
δ_1	0.403 (0.213)	0.351 (0.008)	0.294 (−0.191)	0.220 (−0.081)	0.102 (0.065)
δ_2	0.419 (0.092)	0.364 (0.037)	0.314 (−0.104)	0.230 (0.104)	0.106 (0.035)
$r_\perp$	0.057 (0.019)	0.032 (−0.002)	0.025 (0.002)	0.024 (−0.001)	0.010 (0.005)
$r_\parallel$	0.028 (0.003)	0.016 (0.004)	0.015 (0.008)	0.010 (−0.0009)	0.005 (0.0002)
ϕ_s	0.151 (0.125)	0.084 (−0.069)	0.059 (0.022)	0.052 (−0.017)	0.029 (0.009)

Table 8.8: Mean of absolute uncertainties (and pulls) reported by Minuit on fit parameters for tagged events, for the "100% new physics" model

Statistics	5900	11500	29100	58300	116600
$\Delta\Gamma_s$	0.047 (0.006)	0.033 (−0.011)	0.019 (−0.009)	0.013 (−0.004)	0.010 (0.005)
Γ_s	0.023 (−0.005)	0.021 (0.003)	0.018 (0.003)	0.006 (0.0006)	0.004 (−0.002)
δ_1	0.493 (−0.427)	0.312 (0.065)	0.272 (0.104)	0.150 (0.051)	0.113 (0.022)
δ_2	0.535 (−0.364)	0.339 (−0.067)	0.286 (−0.057)	0.169 (0.028)	0.129 (−0.016)
$r_\perp$	0.049 (0.016)	0.032 (−0.004)	0.029 (−0.004)	0.012 (−0.001)	0.010 (0.005)
$r_\parallel$	0.025 (0.009)	0.017 (0.003)	0.015 (0.005)	0.006 (−0.0008)	0.005 (0.0006)
ϕ_s	0.206 (−0.013)	0.119 (−0.016)	0.117 (0.084)	0.095 (0.024)	0.055 (0.002)

Table 8.9: Correlations between variables uncertainties reported by Minuit for tagged fits using the "50% new physics" model, averaged over all experiments and sample sizes

	Γ_s	$\Delta\Gamma_s$	$r_\parallel$	$r_\perp$	δ_1	δ_2	ϕ_s
Γ_s	1.000	−0.794	−0.080	−0.539	0.007	0.001	−0.050
$\Delta\Gamma_s$	−0.794	1.000	0.106	0.685	0.008	0.017	0.119
$r_\parallel$	−0.080	0.106	1.000	0.010	0.025	0.044	−0.005
$r_\perp$	−0.539	0.685	0.010	1.000	−0.003	−0.003	0.099
δ_1	0.007	0.008	0.025	−0.003	1.000	0.685	0.003
δ_2	0.001	0.017	0.044	−0.003	0.685	1.000	0.003
ϕ_s	−0.050	0.119	−0.005	0.099	0.003	0.003	1.000

Table 8.10: Correlations between variables uncertainties reported by Minuit for tagged fits using the "100% new physics" model, averaged over all experiments and sample sizes

	Γ_s	$\Delta\Gamma_s$	$r_\parallel$	$r_\perp$	δ_1	δ_2	ϕ_s
Γ_s	1.000	−0.633	−0.056	−0.422	−0.023	−0.022	−0.101
$\Delta\Gamma_s$	−0.633	1.000	0.099	0.616	0.026	0.023	0.141
$r_\parallel$	−0.056	0.099	1.000	0.049	0.033	0.026	−0.004
$r_\perp$	−0.422	0.616	0.049	1.000	0.002	0.012	0.137
δ_1	−0.023	0.026	0.033	0.002	1.000	0.507	0.012
δ_2	−0.022	0.023	0.026	0.012	0.507	1.000	0.011
ϕ_s	−0.101	0.141	−0.004	0.137	0.012	0.011	1.000

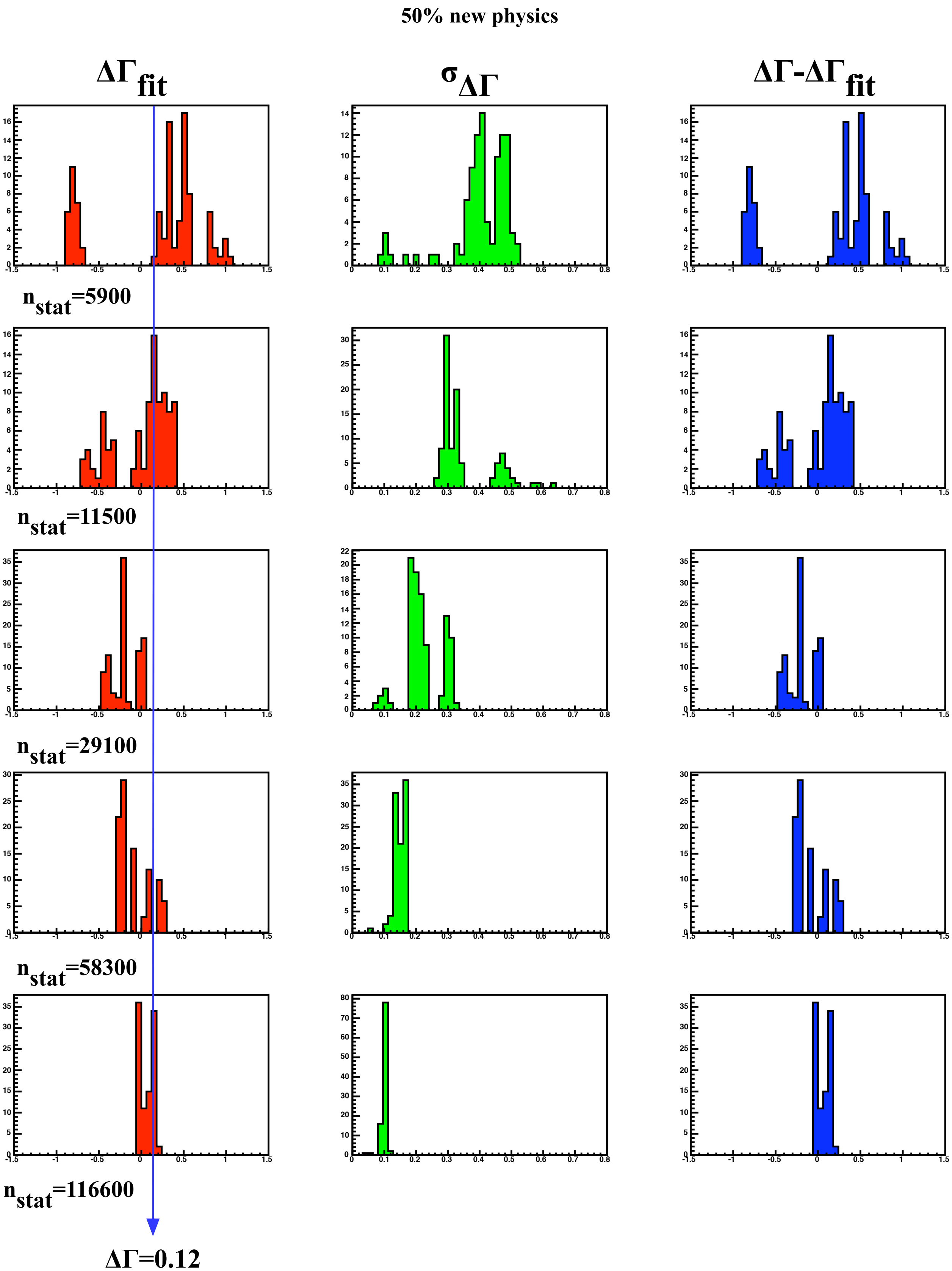


Figure 8.12: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, tfor he parameter $\Delta\Gamma$, under the 50% new physics model.

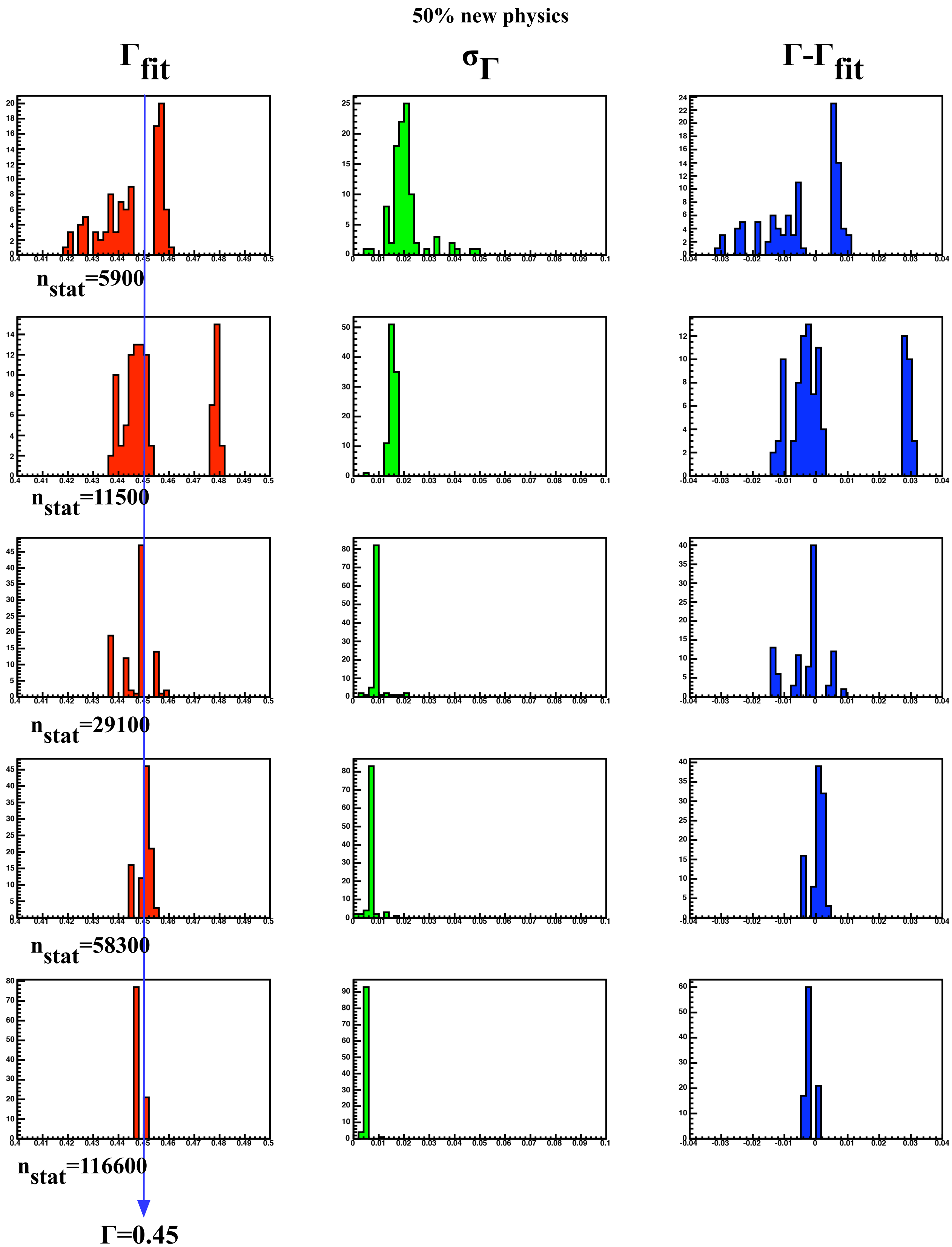


Figure 8.13: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, for he parameter Γ, under the 50% new physics model.

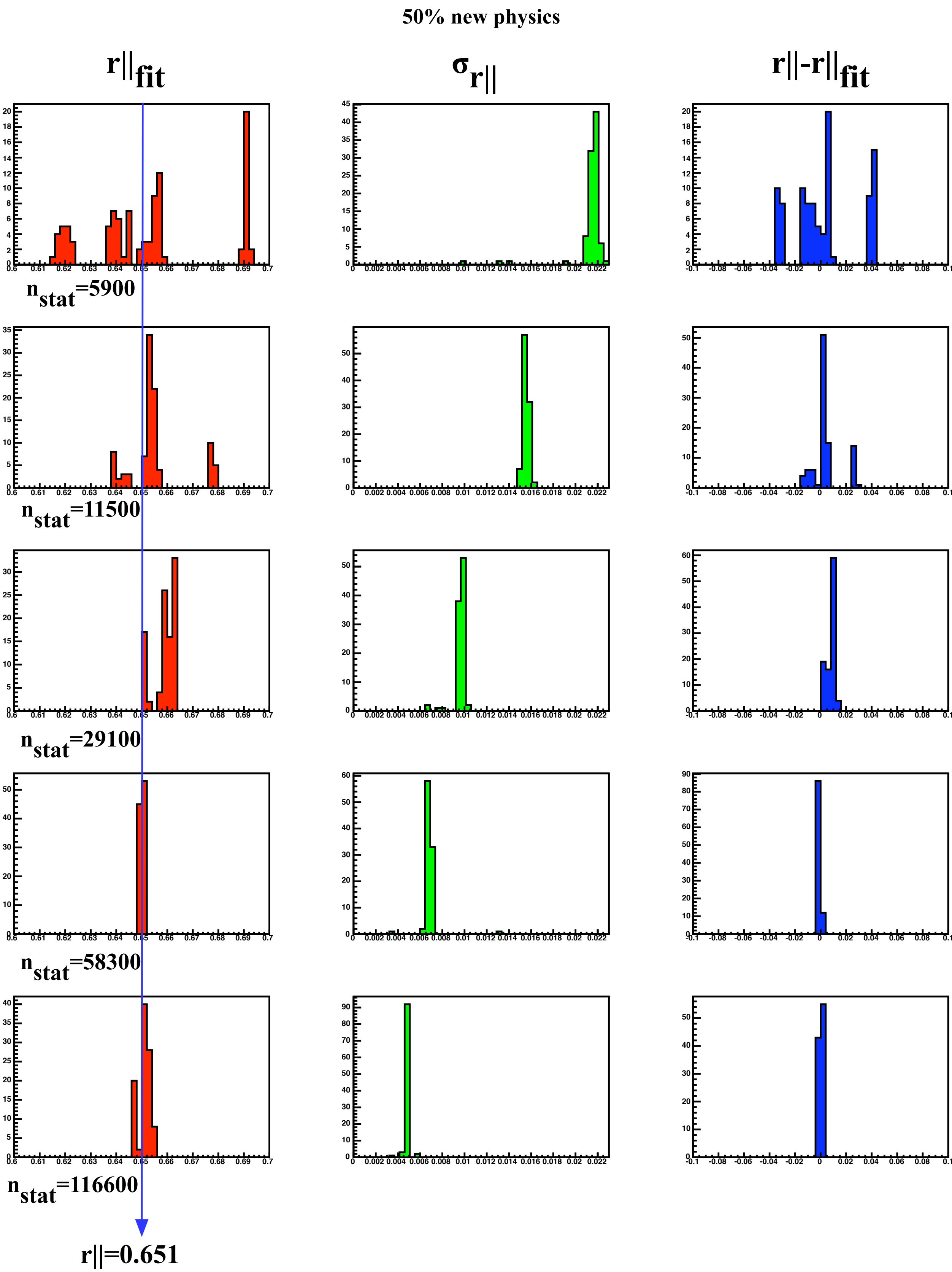


Figure 8.14: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, for the parameter $r_{\parallel}$, under the 50% new physics model.

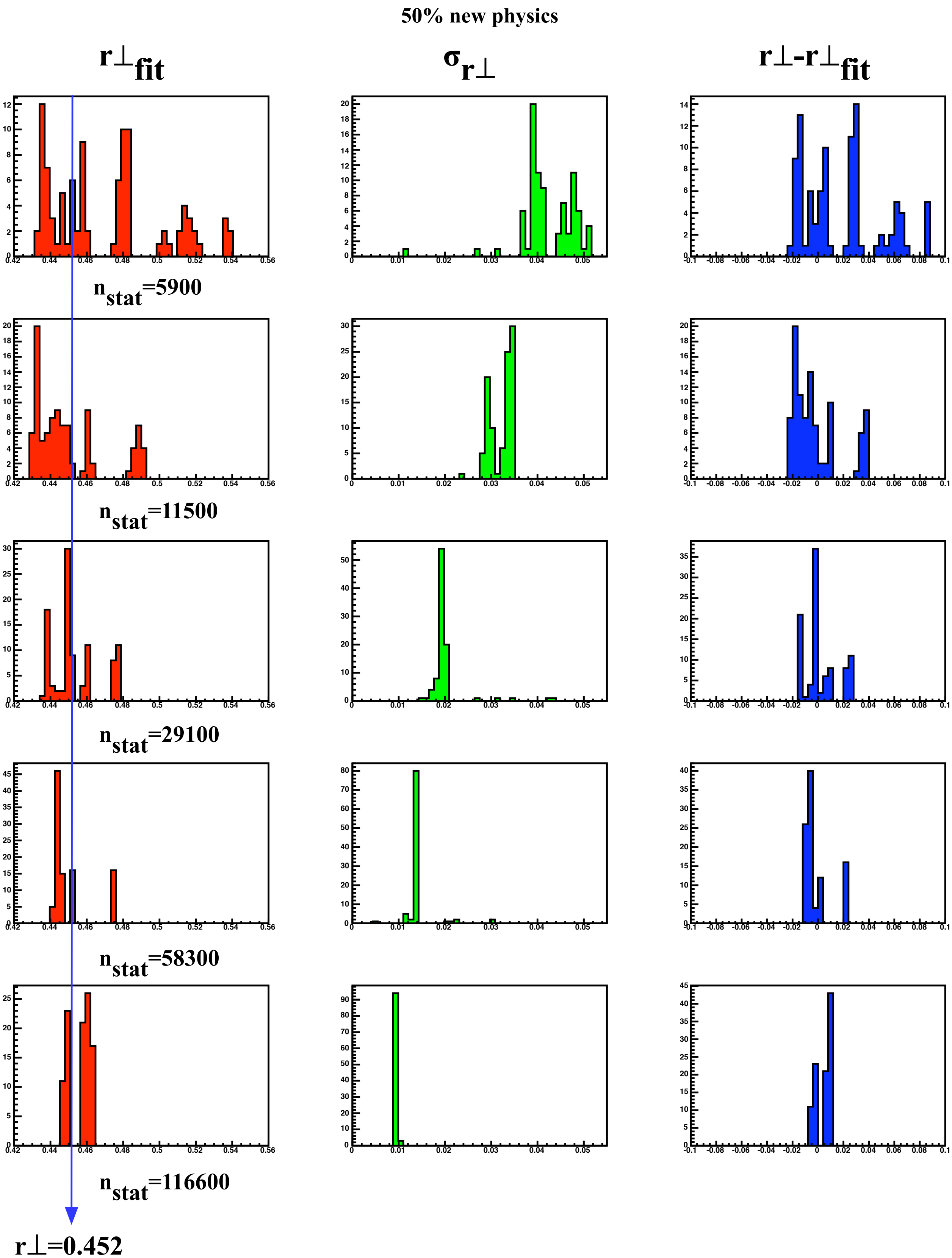


Figure 8.15: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, for the parameter $r_\perp$, under the 50% new physics model.

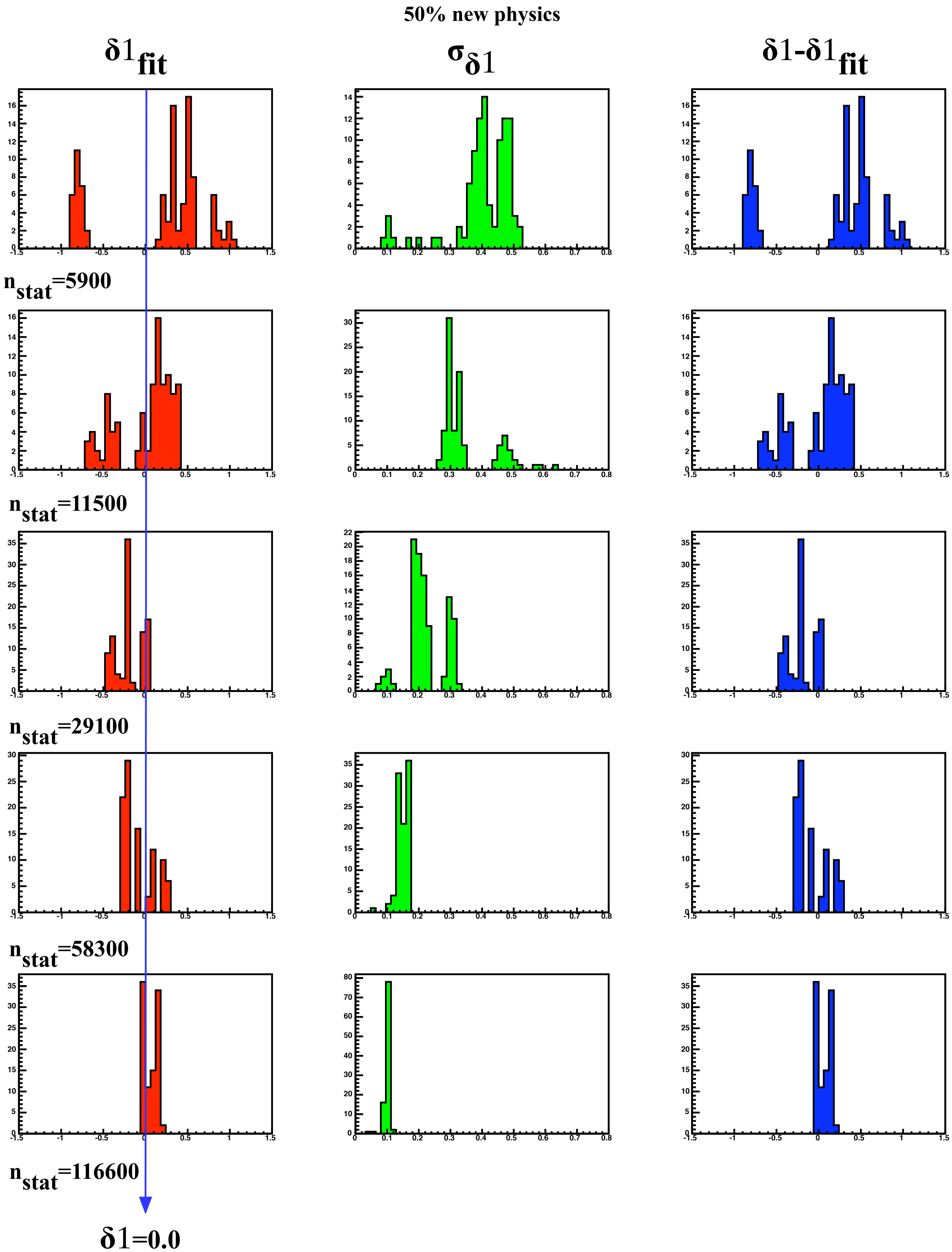


Figure 8.16: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, for the parameter δ_1, under the 50% new physics model.

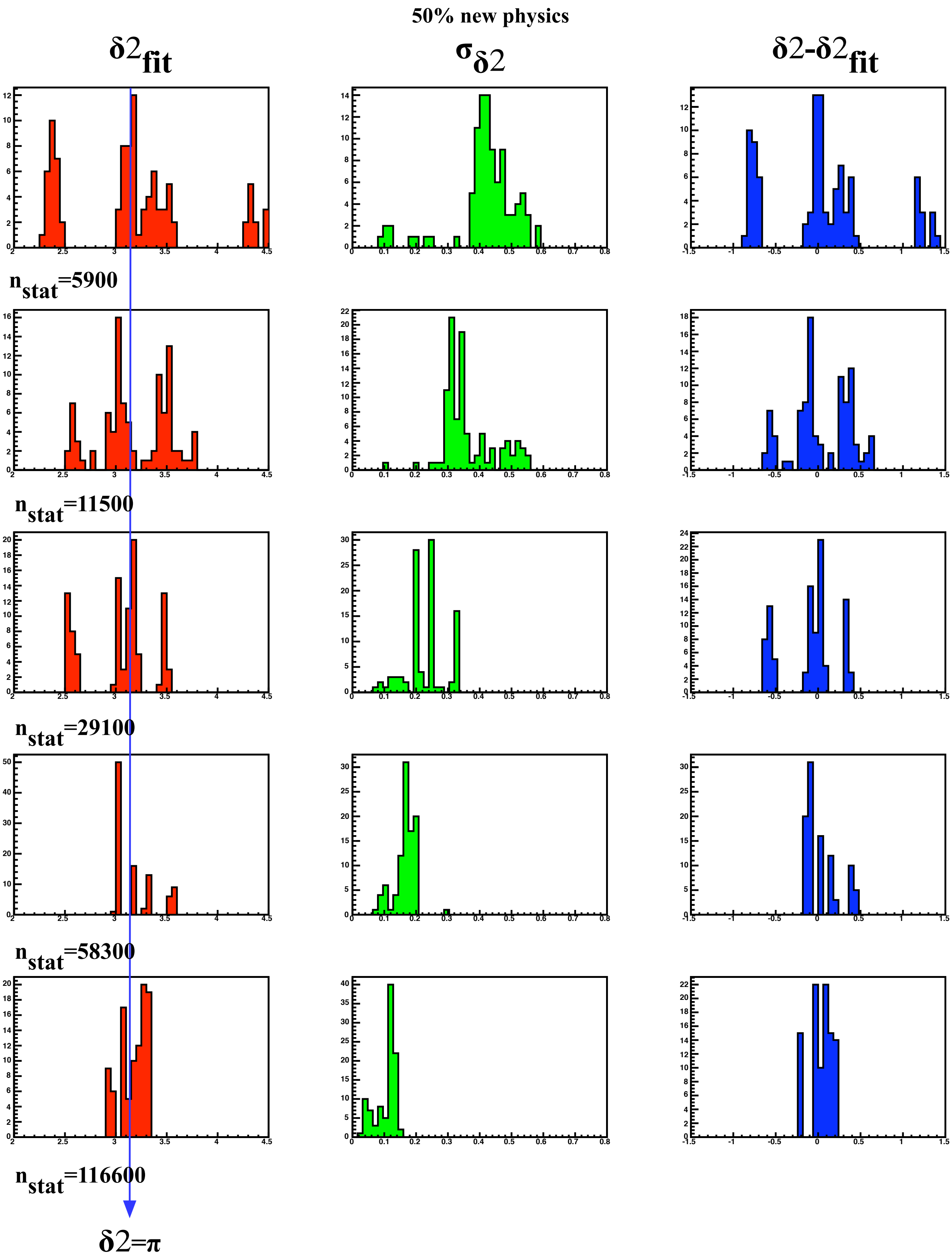


Figure 8.17: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, for the parameter δ_2, under the 50% new physics model.

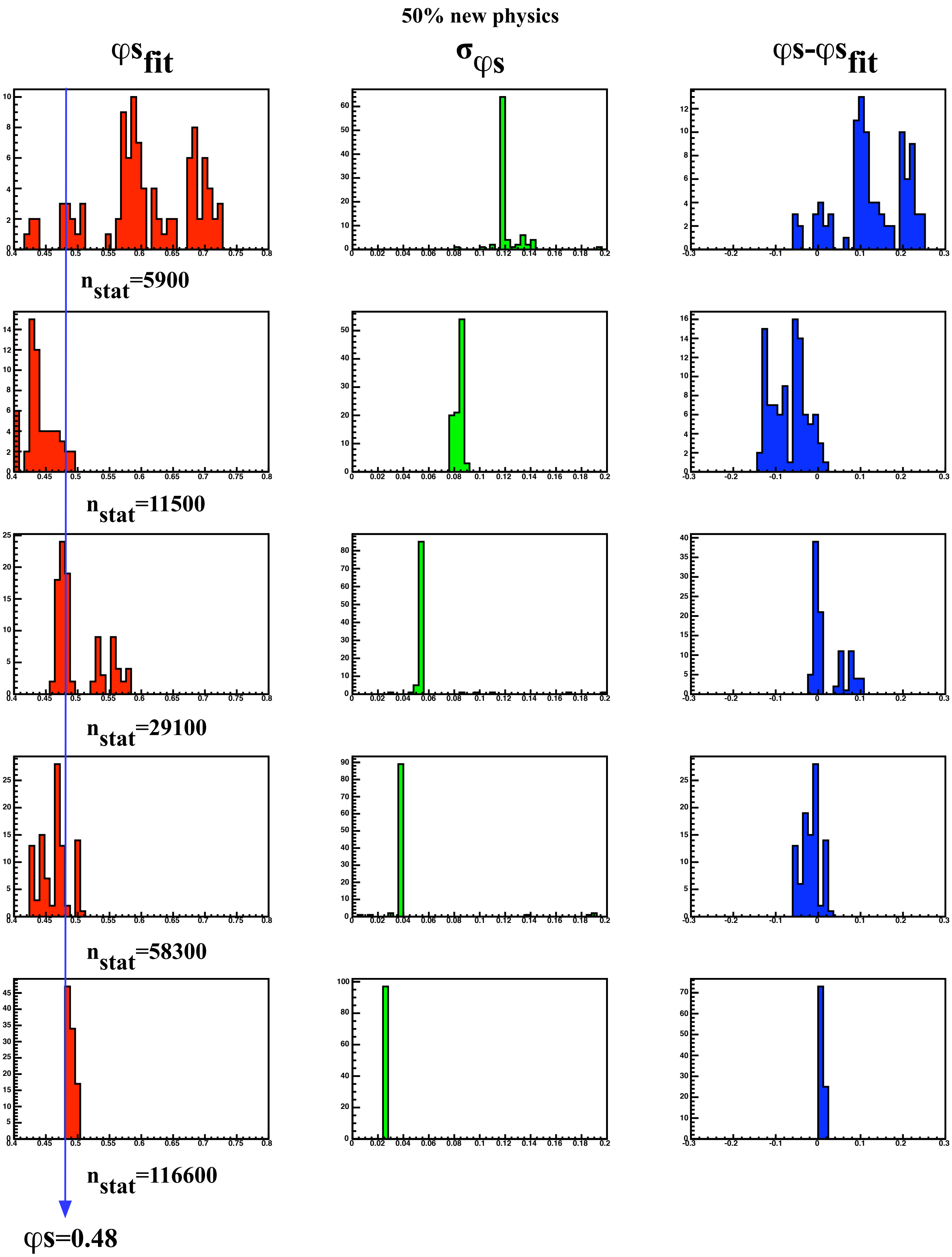


Figure 8.18: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, for the parameter ϕ_s, under the 50% new physics model.

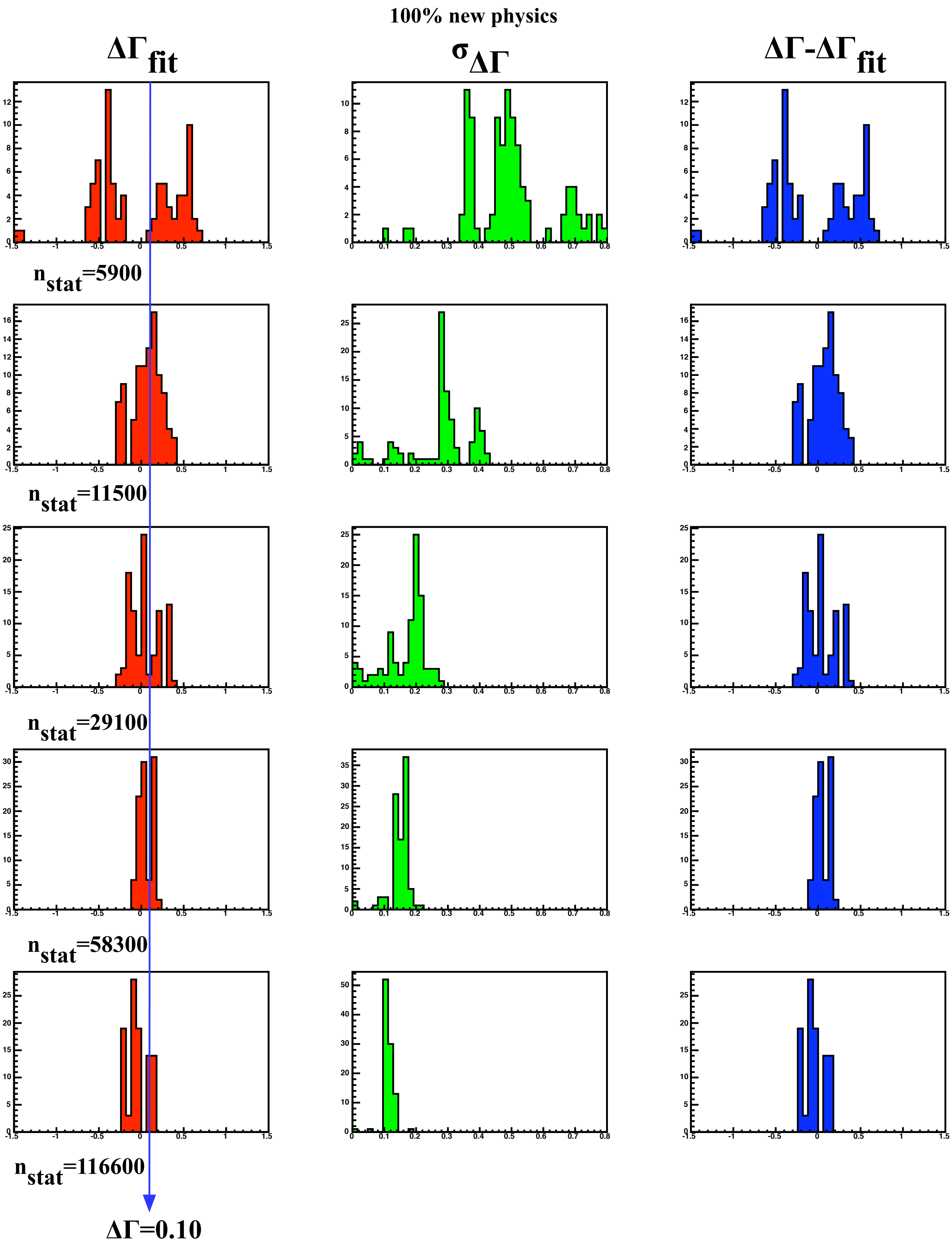


Figure 8.19: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, for the parameter $\Delta\Gamma$, under the 100% new physics model.

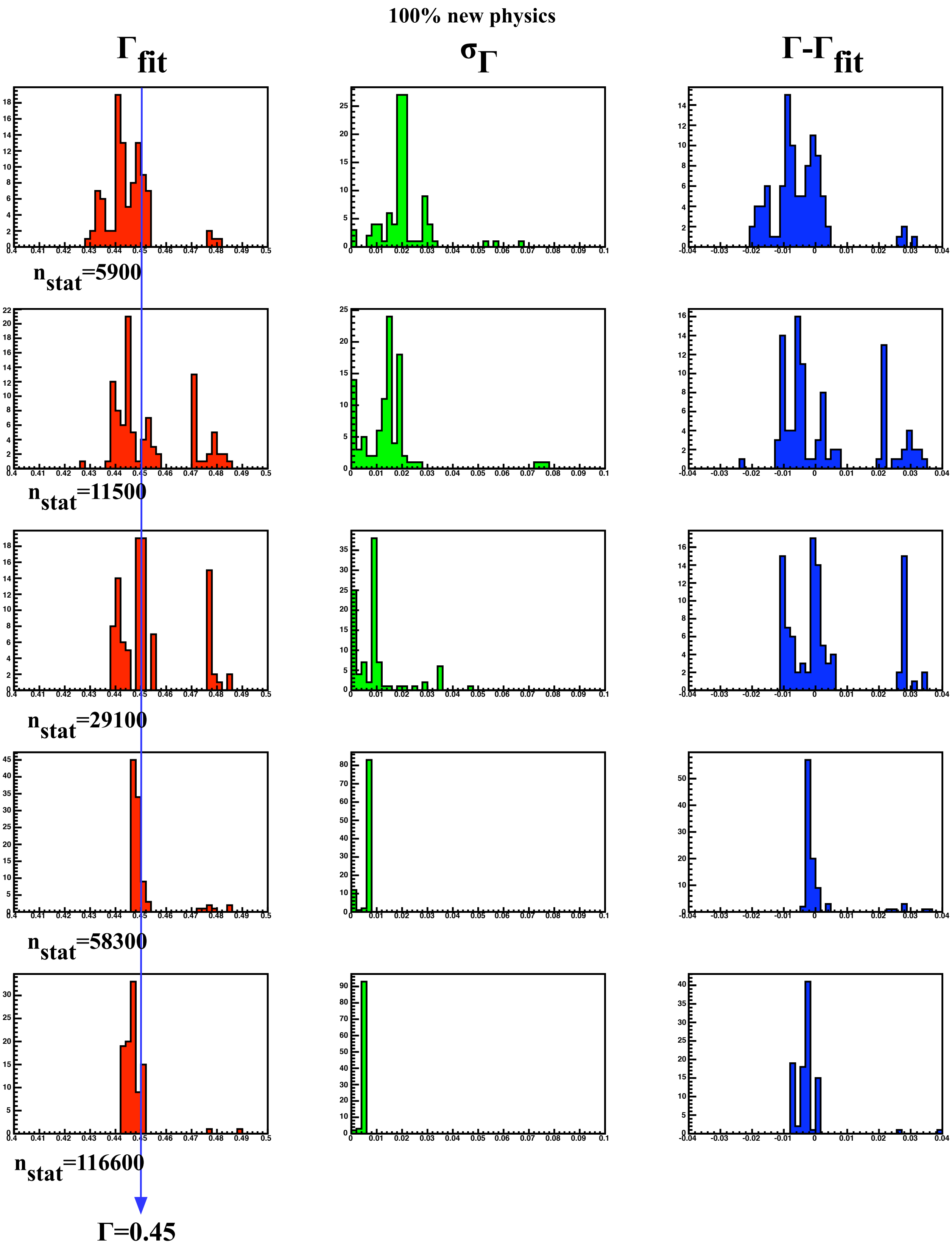


Figure 8.20: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, for the parameter Γ, under the 100% new physics model.

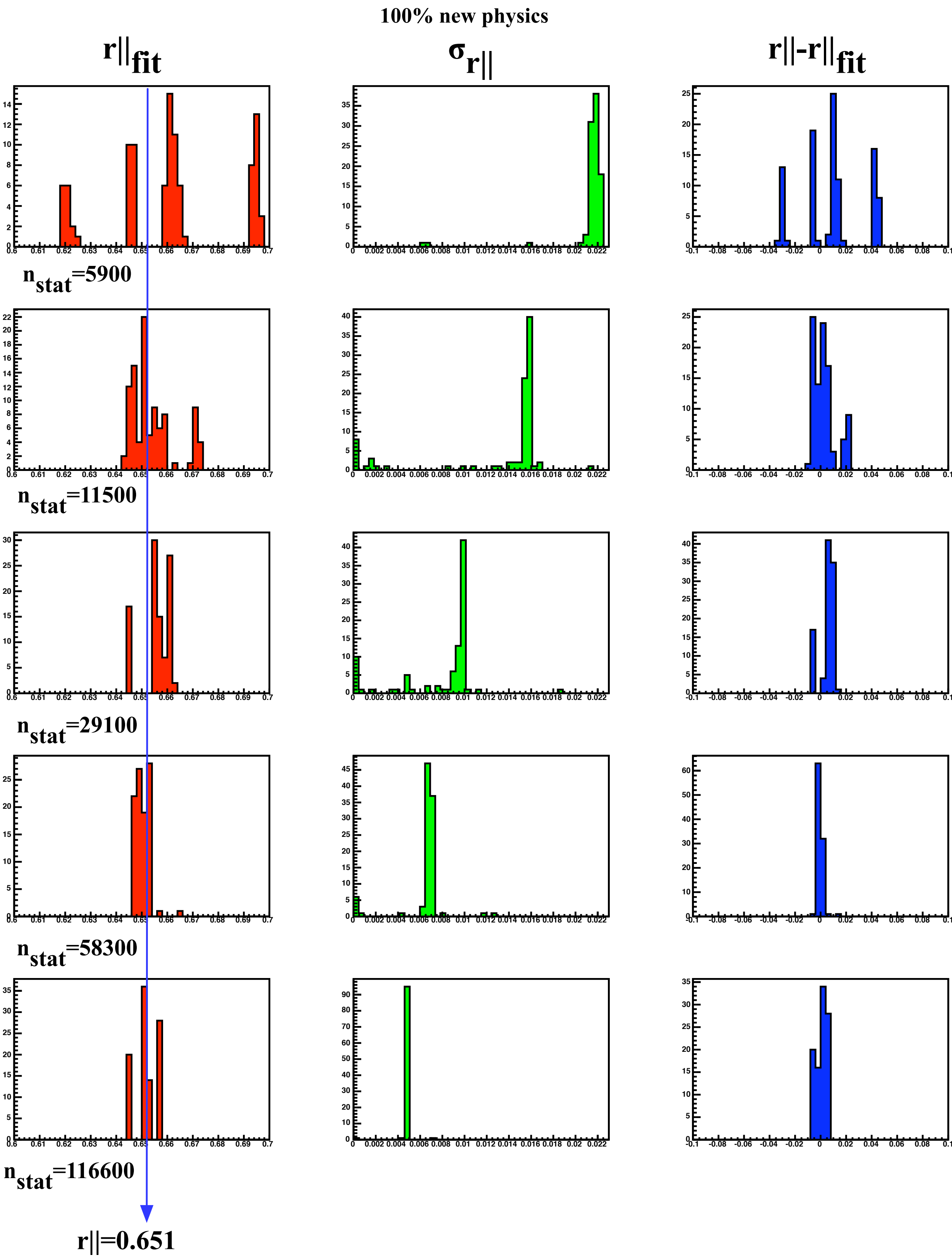


Figure 8.21: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, for the parameter $r_{\|}$, under the 100% new physics model.

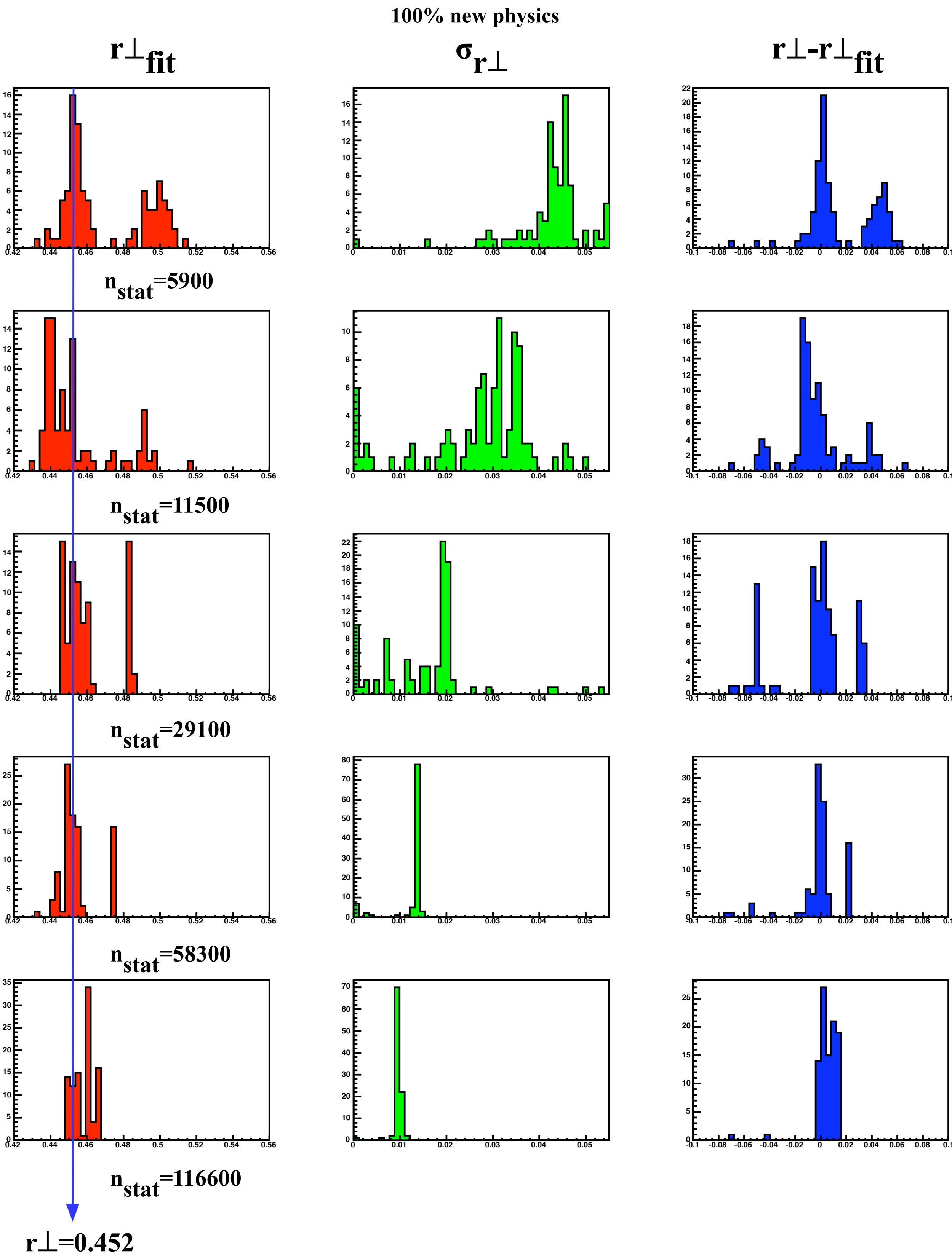


Figure 8.22: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, for the parameter $r_\perp$, under the 100% new physics model.

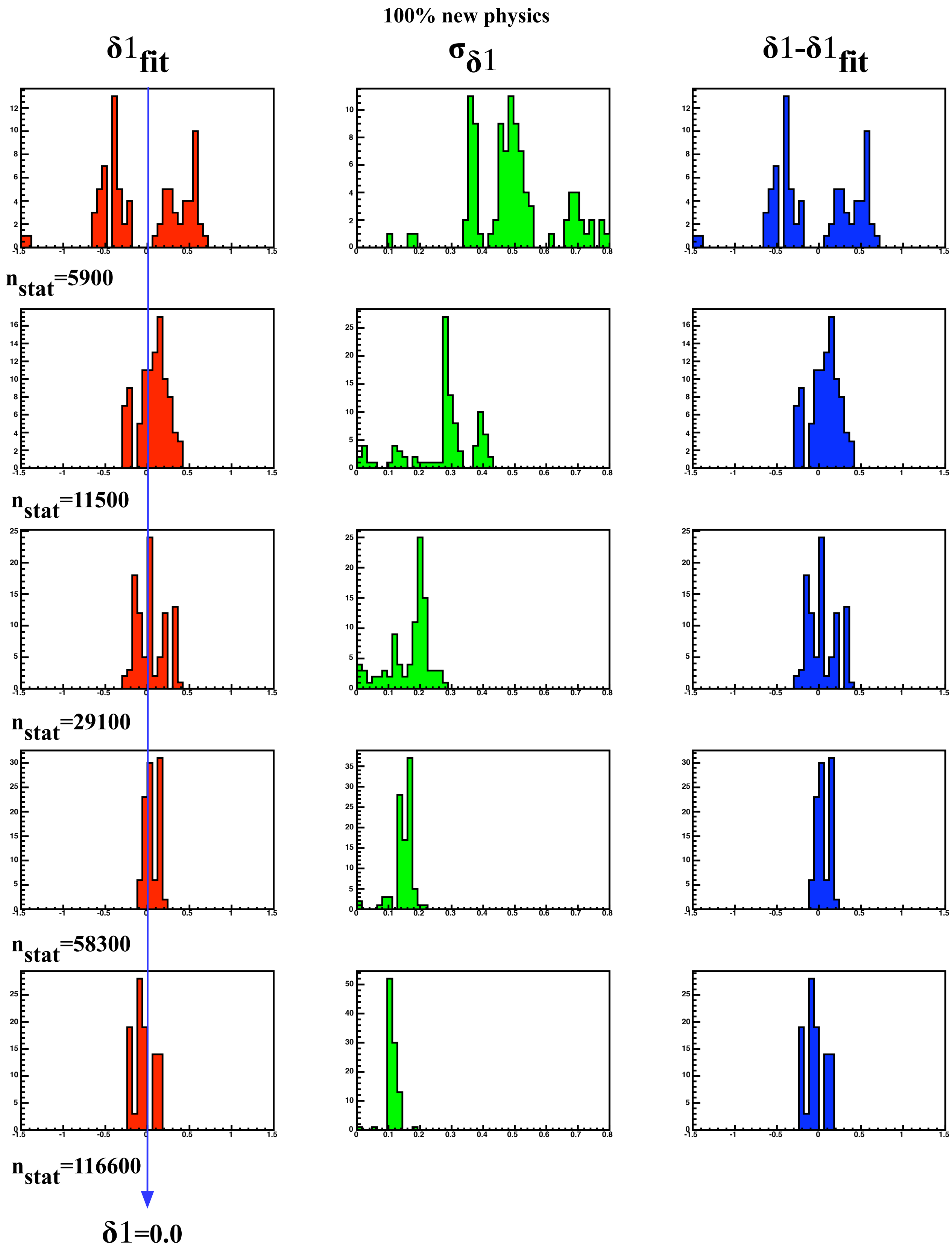


Figure 8.23: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, for the parameter δ_1, under the 100% new physics model.

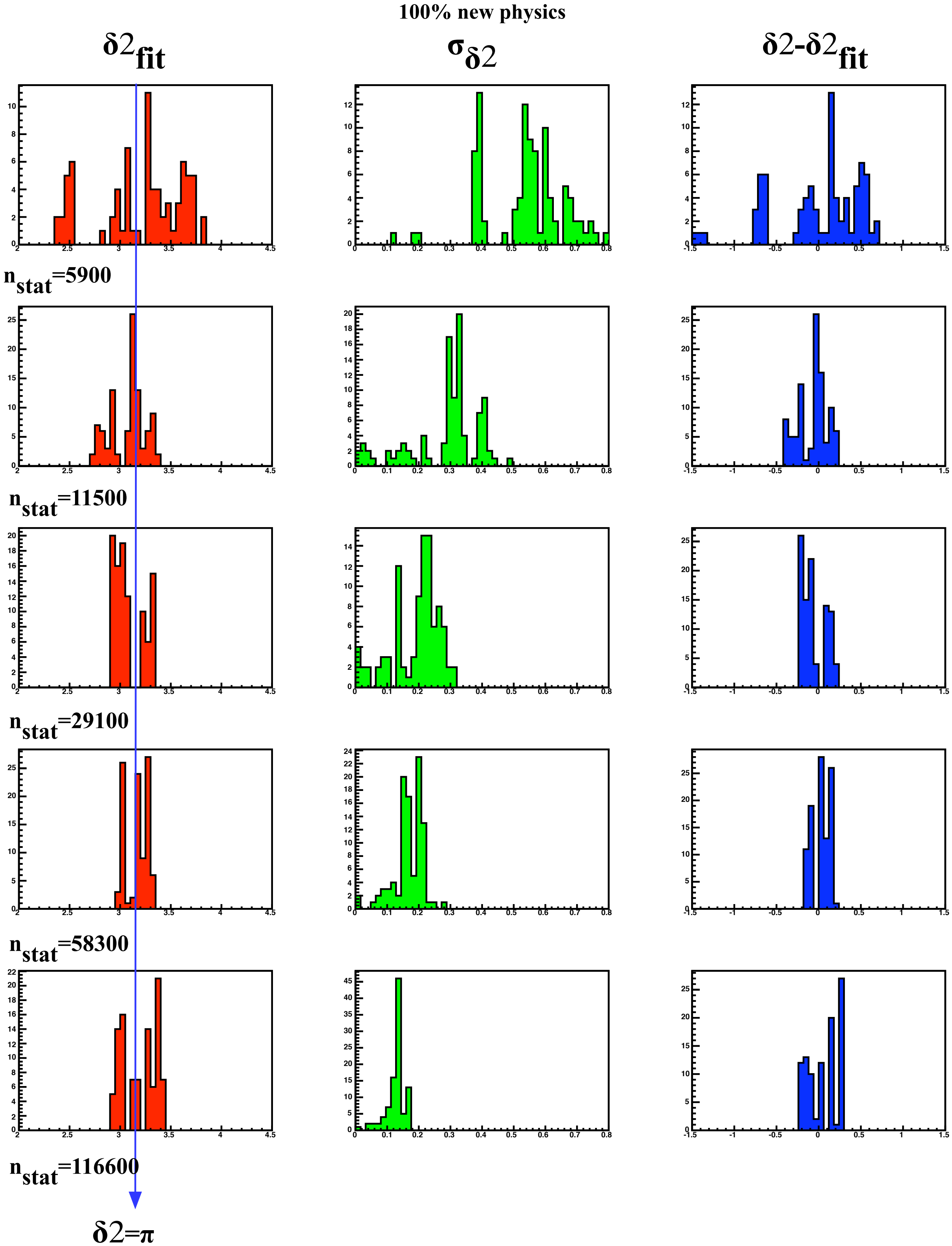


Figure 8.24: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, for the parameter δ_2, under the 100% new physics model.

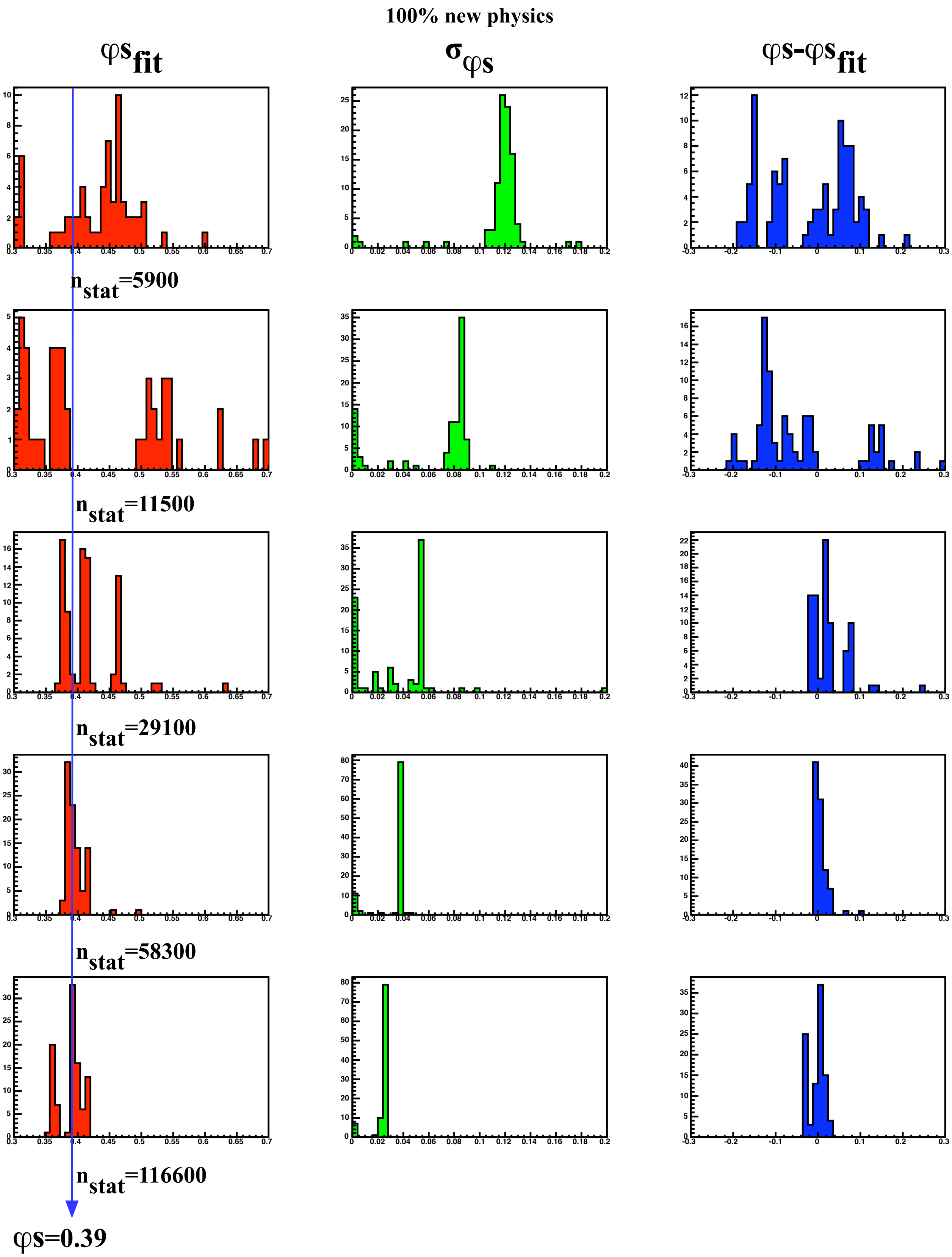


Figure 8.25: Minuit results (red), errors (green) and differences between input value and Minuit result (blue) for 100 fits, for he parameter ϕ_s, under the 100% new physics model.

Again the behaviour of the fits is reasonable, as can be seen from the tables 8.7 and 8.8 and the plots 8.19 through to 8.25.

From the tables showing the mean correlations between the parameters (8.9 and 8.10), it is apparent that some of the variables are strongly correlated, most notably the average lifetime and the lifetime difference of the two B_s states, $\Delta\Gamma_s$ and Γ_s.

In summary, after $20fb^{-1}$ of tagged events data, ATLAS can expect to achieve the following precisions on the physical parameters (relative precisions have been calculated using the input values for the respective model) as shown in table 8.11.

Table 8.11: Summary of estimated precisions for New Physics models after $20fb^{-1}$ tagged data

	50% new physics	100% new physics
$\Delta\Gamma_s$	0.010 (8%)	0.010 (10.0%)
Γ_s	0.004 (0.8%)	0.004 (0.8%)
δ_1	0.102	0.113
δ_2	0.106	0.129
$r_\perp$	0.010 (2%)	0.010 (2%)
$r_\parallel$	0.005 (0.8%)	0.005 (0.8%)
ϕ_s	0.029 (6.0%)	0.055 (14.0%)

8.3.4 Systematic errors

The significant assumption in this sensitivity study is the absence of angular structure in the background. In fact, the exclusive background $B_d^0 \to J/\psi(\mu\mu)K^{0*}(K\pi)$, being an $S \to VV$ decay, does have a very similar angular structure to the signal process. In the analysis the angles will be calculated assuming the wrong masses, and so the distribution passed to the maximum likelihood will be distorted. Furthermore, those processes which are genuinely flat on average such as the semi-inclusive $bb \to J/\psi(\mu\mu)X$, will acquire a distorted angular structure, although the pollution from these backgrounds is less than 5% and can be discounted. However, the exclusive background contributes just over 15% [2] and so this is likely to be a non-neglibigle contributor to the systematic uncertainty. The measurement made by CDF [91] estimated that such systematic errors would be around 1.5% for the transversity amplitudes and the strong phases and 2% for $\Delta\Gamma_s$. This assessment

[2] See erratum at the end of this document

has also assumed a perfect method of acceptance correction, which of course in reality will not be the case - a certain amount of detector-induced distortion will also be present in the distribution, although the CDF study did not find this to be a significant source of systematic uncertainty [91].

In the sensitivity study the oscillation frequency (mass difference) ΔM_s was fixed as previous studies had shown that analyses of this channel alone would not be able to reach it with any meaningful precision (a joint analysis with a hadronic channel such as $B_s \to D_s \pi$ being necessary). It was assumed that fixing this quantity would not cause the errors to be underestimated to any serious degree. This assumption was tested by running two identical fits (with the same random seeds), with the quantity ΔM_s fixed to values 15% above and below the value used in the main studies, to ensure there was no significant change in the uncertainties reported by Minuit. These tests were carried out for each of the parameter sets. The results are shown in table 8.12. It can be clearly seen that there is no significant difference between the errors reported in any of the parameters (less than 1% in all cases) and so fixing the oscillation frequency can therefore be neglected as a source of systematic uncertainty in this study.

Table 8.12: Uncertainties reported by Minuit for a single fit, with ΔM_s set to values 15% greater and 15% less than the values used in the main study. The same random seed was used throughout and each fit was run for 10000 events. All other settings were the same as in the main study

	SM		50% NP		100% NP	
ΔM_s	13.98	18.92	22.68	27.62	27.96	32.90
$\Delta\Gamma_s$	0.151	0.150	0.121	0.121	0.117	0.118
Γ_s	0.068	0.068	0.055	0.055	0.053	0.053
$A_\perp$	0.041	0.041	0.033	0.034	0.028	0.028
$A_\parallel$	0.017	0.017	0.016	0.016	0.015	0.015
δ_1	5.651	5.812	0.322	0.320	0.315	0.313
δ_2	5.719	5.886	0.378	0.377	0.311	0.311
ϕ_s	–	–	0.152	0.151	0.117	0.119

Chapter 9

Summary of results and conclusions

This work has described the exercises to assess the sensitivity of ATLAS to the weak mixing parameters $\Delta\Gamma_s$ and ϕ_s through an angular analysis of the decay $B_s^0 \to J/\psi(\mu\mu)\phi(KK)$. A new decay model, written within the framework of the EvtGen package, has been shown to produce decays whose angular distributions match closely with the derived expression. The performance of the ATLAS detector in collecting these decays and removing the background contributions has been assessed through simulation; it is found that some 49% of the signal events will be collected, with a background contamination of around 26% [1]. The overwhelming source of this background is from the decay $B_d^0 \to J/\psi(\mu\mu)K^{0*}(K\pi)$, and almost all of the other background events contain a real J/ψ and a fake decay on the other side. The angular resolution of the detector is excellent and should not degrade the precision of the parameter measurement to any significant degree, whilst the proper time resolution is more problematic. The events need to be tagged to measure the weak phase, and ATLAS can do this with an efficiency of about 50% and a wrong-tag fraction of about 37%. Finally, using these quantities extracted from the detector performance studies, the sensitivity to the mixing parameters was assessed with a Toy Monte Carlo model.

The final sensitivity results for the three physics scenarios (Standard Model mixing and contributions to mixing amplitudes from new physics of 50% and 100% of the Standard Model), are collected from tables 8.5 and 8.11 and re-written here in table 9.1. This assessment estimates that ATLAS, after $20fb^{-1}$ of tagged data, will be able to obtain a precision on the weak phase ϕ_s of **0.029** in the case that there is a new physics contribution of 50% of the Standard Model to the mixing amplitudes.

[1]See erratum at the end of this document

If there is additional new physics equal to the Standard Model contribution, the precision on ϕ_s will be **0.055**. The weak mixing phase cannot be accessed using untagged events, although such data can still be used to extract the other parameters of this mixing process.

Table 9.1: Summary of estimated precisions for mixing parameters after $20 fb^{-1}$ data (untagged for Standard Model and tagged for New Physics models)

	Standard Model	50% new physics	100% new physics
$\Delta\Gamma_s$	0.009 (6.9%)	0.010 (8%)	0.010 (10.0%)
Γ_s	0.004 (0.8%)	0.004 (0.8%)	0.004 (0.8%)
δ_1	1.824	0.102	0.113
δ_2	1.948	0.106	0.129
$r_\perp$	0.009 (2.0%)	0.010 (2%)	0.010 (2%)
$r_\parallel$	0.007 (1.1%)	0.005 (0.8%)	0.005 (0.8%)
ϕ_s	–	**0.029** (6.0%)	**0.055** (14.0%)

The *raison d'être* of the LHC is to disprove the Standard Model, or at least to show that it is only an approximation to deeper, more meaningful and simpler physics. Additional CP-violation in B-hadrons, aside of being an intriguing subject in its own right, will be an important indicator of the existence of New Physics. Although ATLAS has a relatively small B-physics programme in comparison with other research interests, it is clear that it can make a significant contribution to the search for new sources of CP-violation in such beauty decays. Combining the analysis of this channel with the hadronic decay $B_s \to D_s\pi$ would enable the mixing oscillation rate ΔM_s to be measured simultaneously, giving the complete set of mixing parameters. This study did not look at the issue of acceptance corrections to the angular distribution; this is expected to be non-trivial and must be addressed before any real measurement can be performed.

Erratum

In the main body of the text it was stated that the background events were treated by excluding all tracks which did not originate from the B-meson decay chain. This is not the correct approach - all tracks should be processed, and the resulting extra combinatorial events are simply an additional contribution to the backgrounds. The background quoted in the main part of this work is therefore lower than it should be.

In order to assess the extra contribution, the algorithm was re-run over smaller samples of inclusive and exclusive background events (47915 and 4898 events respectively), but this time including all tracks. The results are shown in the tables below. In each case the statistics shown include all events - no distinction is made between "real" background events and combinatorially produced events.

Cuts	Accepted events All tracks included	Accepted events b-chain tracks only
$B_d \to J/\psi\,(\mu\mu)\,\phi\,(KK)$ - BGI		
Vertexing, J/ψ and ϕ	14721	75
$5250MeV \leq M_B \leq 5490MeV$	200	56
$\chi^2/d.o.f < 5.0$	79	51
$\tau_B > 0.5ps$	49	39
$p_{t,B} > 10.5GeV$	48	39
$bb \to J/\psi X$ - BGII		
Vertexing, J/ψ and ϕ	5110	4921
$5250MeV \leq M_B \leq 5490MeV$	71	68
$\chi^2/d.o.f < 5.0$	43	41
$\tau_B > 0.5ps$	17	16
$p_{t,B} > 10.5GeV$	17	15

The revised acceptances for the two backgrounds are therefore

$$\frac{48}{4898} = 9.8 \times 10^{-3} \text{ - BGI}$$

$$\frac{17}{49715} = 3.5 \times 10^{-4} \text{ - BGII}$$

These acceptances are higher than the figures quoted in the main text. If the revised figures are inserted into Table 7.9, which summarizes the background results, it reads as follows.

	BGI	BGII
Events processed	113401	87006
Events accepted	875	27
Acceptance	9.8×10^{-3}	3.5×10^{-4}
Cross section ratio B/S	8.4	63.1
Signal events processed	103737	103737
Equivalent b.g. events	871391	7587274
Equivalent b.g. accepted	8540	2656
Signal events accepted	51100	51100
Contamination (%)	16.7	5.2

On this basis the final background contamination is likely to be around 4% higher than quoted in the main text (26.3% as opposed to 22.1%; the combinatorial contribution to the signal events was not affected by this error and is consequently unaffected, remaining at 4.4%. These revised figures are used in the conclusions to this work.

Bibliography

[1] The DONUT collaboration, Phys. Lett. B**504**, 218-244 (2001)

[2] P. Renton, Nature **428**, 141-144 (2004)

[3] CERN Yellow Book, *The Large Hadron Collider: conceptual design*, CERN-AC-95-05

[4] The ATLAS Collaboration, CERN/LHCC/94-43

[5] The ALPHA Collaboration web-site, http://alpha.web.cern.ch/alpha/

[6] F. Byron (Jr), R. Fuller, *Mathematics of Classical and Quantum Physics*, Dover Publications (1992)

[7] S. Wu, Phys. Rev. **105**, 14131415 (1957)

[8] O. Greenberg, Phys. Rev. Lett. **89**, 231602 (2002)

[9] I. Bigi, A. Sanda, *CP Violation*, Cambridge University Press (2000)

[10] W. Cottingham, D. Greenwood, *An Introduction to the Standard Model of Particle Physics*, Cambridge University Press (1998)

[11] A. Sakharov, JETP Lett. **5**, 24 (1967)

[12] J. Christenson, J. Cronin, V. Fitch, R. Turlay, Phys. Rev. Lett **13**, 138 (1964)

[13] R. Fleischer, Physics Reports **370**, 537-680 (2002)

[14] L. Wolfenstein, Phys. Rev. Lett **51**, 1945 (1983)

[15] C. Jarlskog, R. Stora, Phys. Lett. B **208**, 268 (1988)

[16] J. Richman, *An Experimenter's Guide to the Helicity formalism*, CALT-68-1148 (1984)

[17] S. Chung, *Spin Formalisms*, CERN71-8 (1971)

[18] M. Jacob, G. Wick, Ann. Phys. **7**, 404-428 (1959)

[19] B. Bransden, C. Joachain, *Introduction to Quantum Mechanics*, Longman (1989)

[20] J. Cornwell, *Group Theory in Physics - an introduction*, Academic Press (1997)

[21] A. Edmonds, *Angular Momentum in Quantum Mechanics*, Princeton University Press (1957)

[22] M. Rose, *Elementary Theory of Angular Momentum*, Wiley (1957)

[23] W. Rolnick, *The Fundamental Particles and their Interaction*, Addison-Wesley (1994)

[24] J. Cudell *et al*, the COMPETE collaboration, hep-ph/0212101

[25] S. Catani *et al*, hep-ph/0005025

[26] S. George, talk at BEAUTY 2003 conference (Carnegie Mellon University): "The ATLAS B-physics trigger", to be published by the American Institute of Physics

[27] R. Hauser, Eur. Phys. J. C**34**, s1.173-s1.183 (2004)
DOI: 10.1140/epjcd/s2004-04-018-6

[28] M. Smižanská, Eur. Phys. J. C**34**, s1.385-s1.392 (2004)
DOI: 10.1140/epjcd/s2004-04-039-1

[29] The ATLAS Collaboration (ATLAS Computing Group), *ATLAS Computing Technical Design Report*, ATLAS-TDR-017, CERN-LHCC-2005-022 (2005)

[30] The Gaudi website, http://proj-gaudi.web.cern.ch/proj-gaudi/welcome.html

[31] The Athena website,
http://atlas.web.cern.ch/Atlas/GROUPS/
SOFTWARE/OO/architecture/index.html

[32] The Economist, article *One Grid to rule them all*, 7th October 2004

[33] LCG Operations Centre web-site, http://www.grid-support.ac.uk/GOC/

[34] LCG web site, http://www.cern.ch/lcg

[35] D. Cameron, unpublished Ph.D. thesis entitled *Replica Management and Optimization for Data Grids*, University of Glasgow (2004)

[36] Apple's Xgrid website,
http://www.apple.com/macosx/features/xgrid/

[37] C. Nicholson, unpublished graduate report *Simulation of a Particle Physics Data Grid*, University of Glasgow (2004)

[38] The Globus Alliance web-page: http://www.globus.org/

[39] The Condor web site: http://www.cs.wisc.edu/condor/

[40] James Frey *et al*, *Condor-G: A Computation Management Agent for Multi-Institutional Grids*, Proceedings of the Tenth IEEE Symposium on High Performance Distributed Computing (HPDC10) San Francisco, California, August 7-9, 2001. See also http://www.cs.wisc.edu/condor/condorg/

[41] European Data Grid web-page:
http://egee-intranet.web.cern.ch/egee-intranet/gateway.html

[42] I. Foster, C. Kesselman, S. Tuecke, International Journal of Supercomputer Applications **15**(3), 2001

[43] A. de Salvo, presentation on the ATLAS job transformations mechanism, http://jcatmore.home.cern.ch/jcatmore/JobTransforms.pdf

[44] M. Branco, Don Quijote Twiki pages,
https://uimon.cern.ch/twiki/bin/view/Atlas/DonQuijote

[45] J. Catmore, A. de Salvo, LJSF Twiki pages
https://uimon.cern.ch/twiki/bin/view/Atlas/LJSF

[46] The CMT (Configuration Management Tool) web-site
http://www.cmtsite.org/

[47] The LCG Generator Services (GenSer) web-site
http://lcgapp.cern.ch/project/simu/generator/

[48] T. Sjöstrand *et al*, Computer Phys. Commun. **135** 238 (2001)

[49] G. Corcella *et al*, JHEP 0101 (2001) 010 [hep-ph/0011363]; hep-ph/0210213.

[50] F. Paige *et al*, hep-ph/312045

[51] E. Barberio, Z. Was, Comput. Phys. Commun., **79** 291 (1994)

[52] P. Golonka *et al*, CERN-TH/2003-287, hep-ph/0312240

[53] M. Gyulassy, X. Wang, Comput. Phys. Commun., **83** 307 (1994)

[54] D. Lange, Nucl. Inst. Meth A**462** 152 (2001)

[55] J. Catmore, M. Smižanská, ATLAS internal communication ATL-COM-PHYS-2004-041 (2004)

[56] M. Dobbs, J. Hansen, Comput. Phys. Commun., **134** 41 (2001)

[57] H. Lai *et al*, hep-ph/9410404

[58] Document under preparation

[59] M. Smižanská, ATLAS communication ATL-COM-PHYS-2003-038 (2003)

[60] The Geant4 Collaboration, Nucl. Inst. Meth A**506** 250-303 (2003)

[61] The ATLAS Geant4 simulation web-page:
http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/OO/
simulation/geant4/index.html

[62] The Geant4 physics reference manual:
http://wwwasd.web.cern.ch/wwwasd/geant4/G4UsersDocuments/
UsersGuides/PhysicsReferenceManual/html/PhysicsReferenceManual.html

[63] RecExample web-page:
http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/OO/domains
/Reconstruction/packages/RecExample/RecExample_archive.htm

[64] InDetRecExample wiki page:
https://uimon.cern.ch/twiki/bin/view/Atlas/InDetRecExample

[65] DOxygen documentation for the ESD Track class:
http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/OO/dist/
nightlies/rel/latest_doxygen/InstallArea/doc/TrkTrack/html/index.html

[66] I. Gavrilenko, ATLAS internal note ATL-INDET-92-016 (1992)

[67] P. Billoir, Nucl. Instr. Meth. **255** 352 (1984)

[68] R. Clifft, A, Poppleton, ATLAS internal note ATL-SOFT-94-009 (1994)

[69] VxBilloirTools web-page: http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/OO/dist/10.0.2/InstallArea/doc/VxBilloirTools/html/

[70] P. Billoir, S. Qian, Nucl. Instr. Meth. A**311** 139-150 (1992)

[71] J. Marriner, CDF Internal Note 1996 (1993)

[72] ATLAS Inner Detector Technical Design Report, CERN/LHCC 97-16

[73] B. Aubert *et al*, Nucl. Instr. Meth. A**558** 388-418 (2006)

[74] ATLAS Muon Spectrometer Technical Design Report, CERN/LHCC 97-22

[75] ATLAS web-pages for the general public:
http://atlas.ch/

[76] TrackParticle web documentation:
http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/OO/dist/nightlies/rel/latest_doxygen/InstallArea/doc/Particle/html/

[77] TruthParticle web documentation:
http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/OO/dist/latest/InstallArea/doc/ParticleEvent/html/index.html

[78] Physics analysis tools website

[79] J. Catmore, B. Epp, P. Reznicek, ATLAS internal communication ATL-COM-PHYS-2006-013

[80] CERN program library long write-up Q121; see also
http://wwwasd.web.cern.ch/wwwasd/paw/

[81] R. Brun, F. Rademakers, Proceedings AIHENP96 Workshop, Lausanne, Sep. 1996, Nucl. Inst. & Meth. in Phys. Res. A**389**, 81-86 (1997). See also http://root.cern.ch/.

[82] F. James, CERN Program Library Long Writeup D506

[83] The Atlantis web-page:
http://atlantis.web.cern.ch/atlantis/

[84] The HEPVIS web-page:
http://boudreau.home.cern.ch/boudreau/v-atlas-hepvis.htm

[85] M. Virchaux, D. Pomarede, ATLAS internal note ATL-SOFT-2001-003 (2001)

[86] I. Dunietz, J. Rosner, Phys. Rev. **D34** number 5, 1404 (1986)

[87] P. Ball, S. Khalil, E. Kou, Phys. Rev. **D69**, 115011 (2004)

[88] K. Abe, M. Satpathy, H. Yamamoto, BELLE note 419, UH-511-982-01 (2001)

[89] P. Ball, R. Fleischer *et al* (editors), *B Decays at the LHC*, CERN-TH/2000-101 (2000)

[90] R. Fleischer, Phys. Rev. D**60**, 073008 (1999)

[91] The CDF Collaboration, CDF note 7115 (2004)

[92] The BaBar Collaboration, Phys. Rev. Lett. **87** number 24 (2001)

[93] The Belle Collaboration, Phys. Lett. **B538**, 11-20 (2002)

[94] The CLEO Collaboration, Phys. Rev. Lett. **79**, number 23 (1997)

[95] A. Dighe, I. Dunietz, H. Lipkin, J. Rosner, Phys. Lett/ **B369**, 144-150

[96] The Aleph Collaboration, Eur. Phys. J. **C4**, 367 (1998)

[97] The Aleph Collaboration, Eur. Phys. J. **C29**, 143 (2003)

[98] The CDF Collaboration, Phys. Rev. Lett. **82**, 3576 (1999)

[99] The Delphi Collaboration, Eur. Phys. J. **C28**, 155 (2003)

[100] The Delphi Collaboration, Eur. Phys. J. **C16**, 555 (2000)

[101] The Delphi Collaboration, Phys. Lett. **B414**, 382 (1997)

[102] The Opal Collaboration, Eur. Phys. J. **C11**, 587 (1999)

[103] The Opal Collaboration, Eur. Phys. J. **C19**, 241 (2001)

[104] The SLD Collaboration, Phys. Rev. **D67**, 012006 (2003)

[105] The SLD Collaboration, Phys. Rev. **D66**, 032009 (2002)

[106] S. Eidelman *et al*, The Particle Data Group, Phys. Lett. **B592**, 1 (2004)

[107] W.-M. Yao *et al*, The Particle Data Group, J. Phys. G **33**, 1 (2006)

[108] The CDF Collaboration, Phys. Rev. Lett. **97** 062003 (2006)

[109] The CDF Collaboration, Phys. Rev **D59**, 032004 (1999)

[110] The CDF Collaboration, Phys. Rev. **D57**, 5382 (1998)

[111] The Aleph Collaboration, Phys. Lett. **B486**, 286 (2000)

[112] The Aleph Collaboration, Phys. Lett. **B377**, 205 (1996)

[113] The Opal Collaboration, Phys. Lett., **B426**, 161 (1998)

[114] E. Bouhova-Thacker, unpublished note *"Model" for* $bb \rightarrow J/\psi X$ *production"*, http://msmizans.home.cern.ch/msmizans/ production/dc2/generators/Btojpsix_note.pdf

[115] E. Bouhova-Thacker, unpublished Ph.D. thesis *Feasibility study for the Measuring of the CKM Phases* γ *and* $\delta\gamma$ *in Decays of Neutral B-Mesons with the ATLAS Detector*, University of Sheffield (2000).

[116] N. Benekos *et al*, ATLAS Physics Note *B-physics performance with Initial and Complete Inner Detector layouts in Data Challenge 1*, ATL-PHYS-2005-002 (2005)

[117] ATLAS - Detector and physics performance Technical Design Report, CERN/LHCC/99-14

[118] G. Cowan, *Statistical Data Analysis*, Oxford University Press (1998)

[119] R. Barlow, *Statistics*, Wiley (1989)

[120] S. Brandt, *Statistical and Computational Methods in Data Analysis*, North-Holland (1970)

[121] B. Epp, V. Ghete, A. Nairz, hep-ph/0202192 (2002)

[122] M. Smižanská, Atlas note ATL-PHYS-99-003 (1999)

[123] A. van Doren, L. de Ridder, J.Comput. Appl. Math. **2**, 207-217 (1976)

[124] A. Genz, A. Malik, J. Comput. Appl. Math. **6**, 295-302 (1980)

[125] G. Raven, LHCb internal note 2003-119

This document was produced with TEXShop on Mac OS X. All plots and images were cropped and labelled with OmniGraffle version 4.1.1, and all diagrams drawn by the author were also produced using this package. The majority of plots were produced with ROOT5, whilst a few were generated with PAW and with Microsoft Excel 2004.