UNIVERSITY
of
GLASGOW

Department of Physics & Astronomy
Experimental Particle Physics Group
Kelvin Building, University of Glasgow,
Glasgow, G12 8QQ, Scotland
Telephone: +44 (0)141 339 8855 Fax: +44 (0)141 330 5881

GLAS-PPE/2008-18
25th November 2008

Measurement of the Proper Time Resolution of the LHCb Detector for $B \to h^+h'^-$ Channels from Data

LAURENCE CARSON[1]

[1]Department of Physics and Astronomy, University of Glasgow,
Glasgow, G12 8QQ

Abstract

A method to determine the parameters of the proper time resolution model for $B \to h^+h'^-$ channels at LHCb is described. The method uses only information which will be available from data. The resolution model is an important input to the fit which extracts the CKM angle γ from the channels $B_d \to \pi^+\pi^-$ and $B_s \to K^+K^-$. It is found that with one nominal year of LHCb data, non-trivial constraints can be found for both the parameters of the resolution model, and the background to signal ratios.

Contents

List of Figures

List of Tables

1 Introduction

The LHCb detector [1] will acquire very large samples of charmless 2-body B decays, which can be used for a variety of interesting physics studies. In particular, measurements of the time-dependent CP asymmetries in the channels $B_d \to \pi^+\pi^-$ and $B_s \to K^+K^-$ can be combined [2] with measurements of ϕ_d and ϕ_s from other channels to measure, under certain U-Spin assumptions, the CKM angle γ.

The measurement of the time-dependent CP asymmetries in a given channel involves a comparison of the reconstructed proper time (τ_{rec}) distribution for events where the initial flavour of the B meson has been tagged as a B, and the corresponding distribution for events where the initial flavour of the B meson has been tagged as an anti-B. Such flavour tagged proper time distributions are affected by experimental factors - namely the finite proper time resolution of the detector, and the incorrect flavour tagging of a certain percentage of events (mistag). Hence to optimise the analysis, some knowledge of both the mistag rate and the proper time resolution from data is needed.

In this study the mistag rate is assumed to be known from measurements in control channels, which will be discussed. This note concentrates on a method which assumes a model for the proper time resolution of the detector, and then extracts the values of the model parameters by fitting to the reconstructed proper time distribution of flavour tagged events for the channel $B_s \to K^-\pi^+$ (the reasons for this choice of channel are discussed in Sec. 4.2). The fit also makes use of the per-event errors ($\sigma_{\tau_{rec}}$) on τ_{rec}, but, crucially since the method is to be applied to data, does not use any information on the true lifetime (τ_{true}) of the Monte Carlo simulated events.

In section 2, two models which aim to describe the proper time resolution distribution are presented, and the models are validated by application to four different $B \to h^+h'^-$ channels. In section 4, a method to determine the parameters of the resolution model for $B \to h^+h'^-$ channels from data is discussed. In section 5, the results from applying the method from section 4 to simulated $B_s \to K^-\pi^+$ data are shown. Conclusions are given in section 6.

2 Proper Time Resolution Model

In this section two proper time resolution models are introduced, and previous work carried out using one of the models is summarised. Note that the models are designed to describe the residuals of only one channel at a time (i.e. only signal events). This is not a problem when studying simulated data, as any background events which may be present can be eliminated by looking at the simulated truth information. Of course when considering real data (i.e. the reconstructed proper time distribution), backgrounds will be present , but these can be dealt with as described in Sec. 4.4.

2.1 Composition of the Model

A model of the proper time resolution which accurately describes the distribution of the proper time residuals as seen in the Data Challenge 2004 (DC04) production of Monte Carlo simulated events has been developed by members of the LHCb group at NIKHEF, and is described fully in [3]. This model is referred to hereafter as the "full model". A simplified version of the full model, developed during this work, is also described.

2.1.1 Full Model

The model developed in [3] takes the following form (defining $x \equiv \Delta\tau \equiv \tau_{rec} - \tau_{true}$):

$$R(x) = (1 - f_1 - f_2)e^{-\frac{1}{2}(\frac{x-M}{S})^2} + f_2 e^{-\frac{1}{2}(\frac{x}{S_{fixed}})^2} + f_1 \left(e^{-\frac{1}{2}(\frac{x}{S})^2} \otimes e^{-(\frac{x}{\tau_{shift}})} \right). \quad (1)$$

As can be seen, the model consists of three distinct parts, each of which plays a different role:

- The first Gaussian, of mean M and width S, describes the bulk of events, where the reconstruction has proceeded normally without any problems.

- The second Gaussian, with mean fixed to zero and width fixed to $S_{fixed} = 10$ps, describes the very small number of events which cause wide tails in the distribution. There are several possible reasons why a signal event can still be in the tails of the $\Delta\tau$ distribution. For example, if the final state contains one or more radiated photons, the reconstructed mass value will be too low, and so this can cause the τ_{rec} value to be too low. Another possibility is that the momentum can be poorly measured (for example if one of the final state particles traverses a lot of material) even if the reconstruction correctly finds the tracks and decay vertex. If the momentum measurement is not close to the true value then the same will be true for the τ_{rec} measurement. The probability for an event to be suffer from such a problem and be in the tails of the $\Delta\tau$ distribution is parameterised by f_2.

- The exponential with decay time τ_{shift}, convoluted with a Gaussian of mean zero and width S, accounts for a bias seen in the Monte Carlo simulation, where there is a small asymmetry toward negative residuals (i.e. τ_{rec} is more likely to be smaller than τ_{true} rather than greater than it). A possible reason for this bias, proposed in Ref. [3], is that when the primary vertex is reconstructed, tracks coming from the decay of the B hadron are sometimes erroneously included. This causes the reconstructed PV to be "attracted" toward the B decay vertex, decreasing τ_{rec}. The probability for this to happen is parameterised by f_1, while the magnitude of the shift this causes in $\Delta\tau$ is parameterised by τ_{shift} (note that this parameter is *not* the proper time of the B meson). However it is not clear whether this effect is responsible for all of the observed bias, or only part of it.

2.1.2 Simplified Model

The effects of using a simplified model to describe the $\Delta\tau$ distribution is also studied. The simplified model fixes f_1 to zero, which eliminates two of the five parameters of the model (f_1 and τ_{shift}). This leaves a model with two parts:

- The main Gaussian of mean M and width S.

- The second Gaussian, with mean fixed to zero and width fixed to $S_{fixed} = 10$ps.

Hence the simplified model takes the form

$$R_{simp}(x) = (1 - f_2)e^{-\frac{1}{2}(\frac{x-M}{S})^2} + f_2 e^{-\frac{1}{2}(\frac{x}{S_{fixed}})^2}. \quad (2)$$

2.2 Previous work on $B_d \to J/\psi K^*$ and $B_u \to J/\psi K^+$

In Ref. [3] the model described in Sec. 2.1.1 was developed and applied to the decay channels $B_d \to J/\psi K^*$ and $B_u \to J/\psi K^+$, using Data Challenge 2004 (DC 04) Monte Carlo simulated data. For those channels the model was found to be a reasonable description of the proper time residual distribution (the details of how the model is applied are given in Sec. 3.1). However it is not clear a priori if the model will be as successful when applied to simulated data which is from channels (such as the $B \to h^+h^{'-}$ channels) which do not have the same topology as the $J/\psi X$ channels, and also to data using the more realistic Monte Carlo simulation which is now available. Hence there is a need for the model to be validated separately for $B \to h^+h^{'-}$ channels.

3 Validation of the Model for $B \to h^+h^{'-}$ on Full Monte Carlo Simulation

The aim of this section is to show that the resolution models introduced in Sec. 2.1 can be used to describe the proper time residual distributions of different $B \to h^+h^{'-}$ channels, making use of the Monte Carlo simulated truth information. It will also be shown that the parameters of the proper time resolution model are similar between the different $B \to h^+h^{'-}$ channels. This is necessary to allow the results from the fit to $B_s \to K^-\pi^+$ data to be applied to the other channels, where a measurement of the proper time resolution model from data will be (for reasons given in Sec. 4.2) difficult if not impossible.

3.1 Method for Validation of Model

The method described here follows the method used in Ref. [3].

For each event, τ_{rec} has a per-event error $\sigma_{\tau_{rec}}$ associated with it. This error is calculated by propagating the errors on the measurements which are combined to calculate the proper time—namely the B-meson momentum, the B-meson flight distance and the reconstructed mass of the B-meson. In LHCb a specialised tool called the

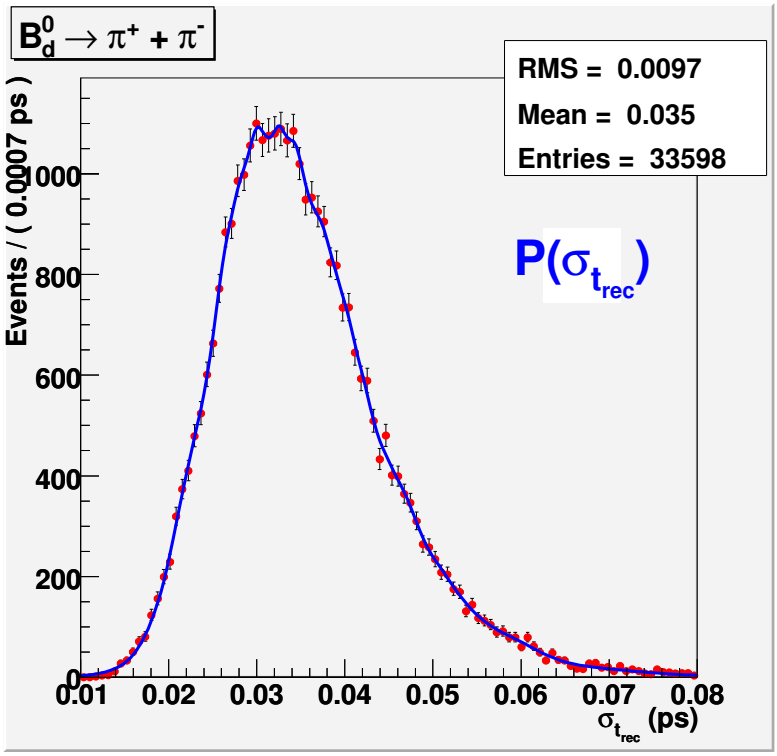

Figure 1: Proper time error distribution (red points) for $B_d \to \pi^+\pi^-$, fitted with a PDF (blue curve) built by kernel estimation.

lifetime fitter has been developed [4] which calculates the τ_{rec} and the $\sigma_{\tau_{rec}}$ for each event. In most fully reconstructed channels at LHCb, the $\sigma_{\tau_{rec}}$ distributions are broadly similar to each other. As an example the $\sigma_{\tau_{rec}}$ distribution for $B_d \to \pi^+\pi^-$ is shown in Fig. 1. The mean of this distribution, 35fs, sets the scale for the expected values for the resolution model parameter S. Although the $\sigma_{\tau_{rec}}$ distribution does not follow an analytical formula, a PDF can be constructed which follows the distribution closely. This is done using the method of kernel estimation [5], where the PDF is built by a superposition of Gaussians, one for each data point in the distribution. An example of such a PDF, along with the $\sigma_{\tau_{rec}}$ distribution it is derived from, is shown in Fig. 1.

For the model to be useful, its parameters have to be global, i.e. they have to describe the resolution model adequately for all signal events. However as signal events have a range of values of $\sigma_{\tau_{rec}}$, the same value of, for example, S is not expected to be able to describe the width of the entire distribution. The goal is to determine a globally applicable model where the parameters of $R(x)$ can be determined on an "event-by-event" basis according to some reconstructed parameter(s) of each event.

It is found that the 5 parameters of $R(x)$, which are M, S, τ_{shift}, f_1 and f_2, can be parameterised very simply. They fall into two groups: M, S and τ_{shift} are linear functions of $\sigma_{\tau_{rec}}$, while f_1 and f_2 are constants.

The parameters relating the parameters of the resolution model to $\sigma_{\tau_{rec}}$ are defined as follows:

- $F1$, which is just the value of f_1 (independent of $\sigma_{\tau_{rec}}$).
- $F2$, which is just the value of f_2 (independent of $\sigma_{\tau_{rec}}$).
- GM, which is the scaling factor between M and $\sigma_{\tau_{rec}}$.
- GS, which is the scaling factor between S and $\sigma_{\tau_{rec}}$.
- $SHIFT$, which is the scaling factor between τ_{shift} and $\sigma_{\tau_{rec}}$.

These dependences allow $R(x)$ to be made conditional on the $\sigma_{\tau_{rec}}$ distribution. If $P(\sigma_{\tau_{rec}})$ is a PDF built by kernel estimation from the $\sigma_{\tau_{rec}}$ distribution, then the equation:

$$R(x) = \int_0^\infty R(\Delta\tau|\sigma_{\tau_{rec}})P(\sigma_{\tau_{rec}})d\sigma_{\tau_{rec}} \tag{3}$$

allows $R(x)$ to be simultaneously fitted to the $\Delta\tau$ distribution for all signal events, regardless of their $\sigma_{\tau_{rec}}$ value. It is merely a question of finding the values of $F1$, GS, etc. which give the best fit to the $\Delta\tau$ distribution.

3.2 Event Selection

The proper time resolution model is applied, using the method outlined in Sec. 3.1, to Data Challenge 2006 (DC 06) Monte Carlo simulated distributions of $\Delta\tau$. The method is applied to selected events in each of the four channels which have the highest branching ratios of the possible $B \to h^+h^{'-}$ channels: $B_d \to \pi^+\pi^-$, $B_s \to K^+K^-$, $B_s \to K^-\pi^+$ and $B_d \to K^+\pi^-$.

To select the events, v19r10 of the LHCb physics analysis package DaVinci [6] is used. The standard LHCb $B \to h^+h^{'-}$ selection, with a tightened mass window and particle identification (PID) cuts added, is run on signal samples for each decay given above. The cuts used are given in Table 1.

Table 1: List of Cuts Applied by the $B \to h^+h^{'-}$ Selection. Note that p_T refers to transverse momentum, IPS refers to impact parameter significance, and FDS refers to flight distance significance. For the PID cuts, DLL$(p_1 - p_2)$ refers to the difference in the log likelihood for a particle to be of type p_1 as opposed to type p_2.

Type of Cut	Value
Mass Window for B	$\pm$50 MeV
Max χ^2 of B vertex	5.0
Min p_T of B	1.0 GeV
Max IPS of B	2.5
Min FDS of B	18.0
Min p_T for both daughters	1.0 GeV
Min p_T for (at least) one daughter	3.0 GeV
Min IPS for both daughters	6.0
Min IPS for (at least) one daughter	12.0
Minimum DLL$(K - \pi)$ for K candidates	0.0
Minimum DLL$(\pi - K)$ for π candidates	0.0

Since occasionally a background event passes the selection, and the proper time resolution model is to be studied for only one class of events at a time, the Monte Carlo simulated truth information is used to select only signal events. The number of signal tape events that the selection was run on, and the results of the selection, are summarised in Table 2.

Input DSTs for the fits to the $\Delta\tau$ distribution were taken from the DC06-phys-v2-lumi2 signal tapes, except for $B_s \to K^+K^-$ where the output of the DC06 stripping (Stripping-v31-lumi2) was used. The reason for this is that the number of $B_s \to K^+K^-$ events available from the DC06-phys-v2-lumi2 signal tape was only 22,000, which is too few to enable a useful fit to the $\Delta\tau$ distribution of selected events to be done. In order to calculate the selection efficiency for $B_s \to K^+K^-$, a small number of $B_s \to K^+K^-$ signal tape events were ran through the selection, and the results of this are used to calculate the selection efficiency for $B_s \to K^+K^-$. For the other three channels there were sufficient numbers of DC06-phys-v2-lumi2 signal events available for that sample to be used to produce the $\Delta\tau$ distribution.

The generator levels cuts are also taken into account, to give an efficiency relative to the total number of events produced, regardless of whether they fall within the LHCb acceptance or not. The generator level efficiencies are not the same across all 4 channels because for $B_s \to KK$ events are retained if the B hadron is within the LHCb acceptance, whereas for the other 3 channels the event is kept only if the decay products of the B hadron are all within the LHCb acceptance, which is a more stringent requirement.

Note that the efficiencies shown here do not include the effect of applying the trigger. For $B_d \to \pi^+\pi^-$ the fit to $\Delta\tau$ was performed on the output of the selection from both productions. Both fits had the same parameters, within errors. This confirms that there is no bias from selecting one production rather than the other.

The overall efficiencies given in Table 2 are very slightly higher than those measured in Ref. [8], which varied between 2.5% and 2.6%. This is because the PID cuts used here are marginally looser than those used in Ref. [8], where for example cuts were made on DLL$(p - \pi)$ and DLL$(\mu - \pi)$ as well as DLL$(K - \pi)$. The fact that the efficiencies have changed little from those in Ref. [8] means that the estimated event yields calculated there can safely be used here.

3.3 Results of Model Validation

Once signal events for a particular $B \to h^+h^{'-}$ channel have been selected in the way just described, the $\Delta\tau$ distribution is fitted using both the full 5 parameter model and the simplified 3 parameter model.

Figures 2 to 9 show the (separate) application of both models to each $B \to h^+h^{'-}$ channel in turn. The values of the parameters of the model are displayed in the top right corner of the graph, along with the χ^2 per degree of freedom of the fit.

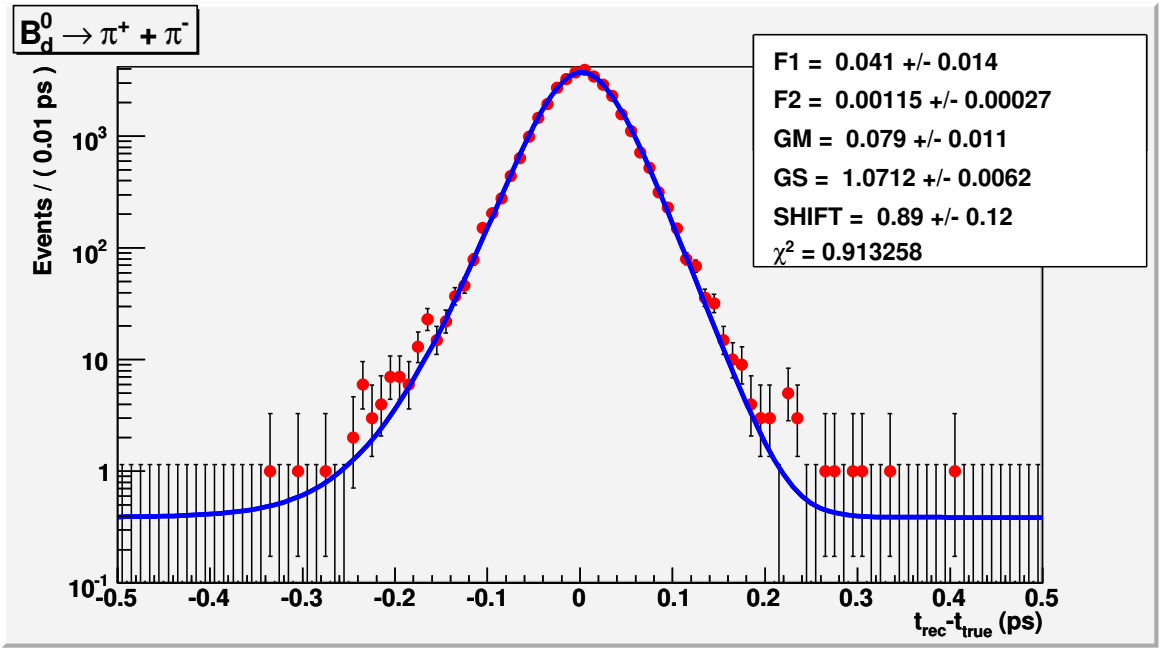

Figure 2: Full model global fit to the proper time residual distribution for $B_d \to \pi^+\pi^-$. The figure quoted as χ^2 is the χ^2 per degree of freedom of the fit.

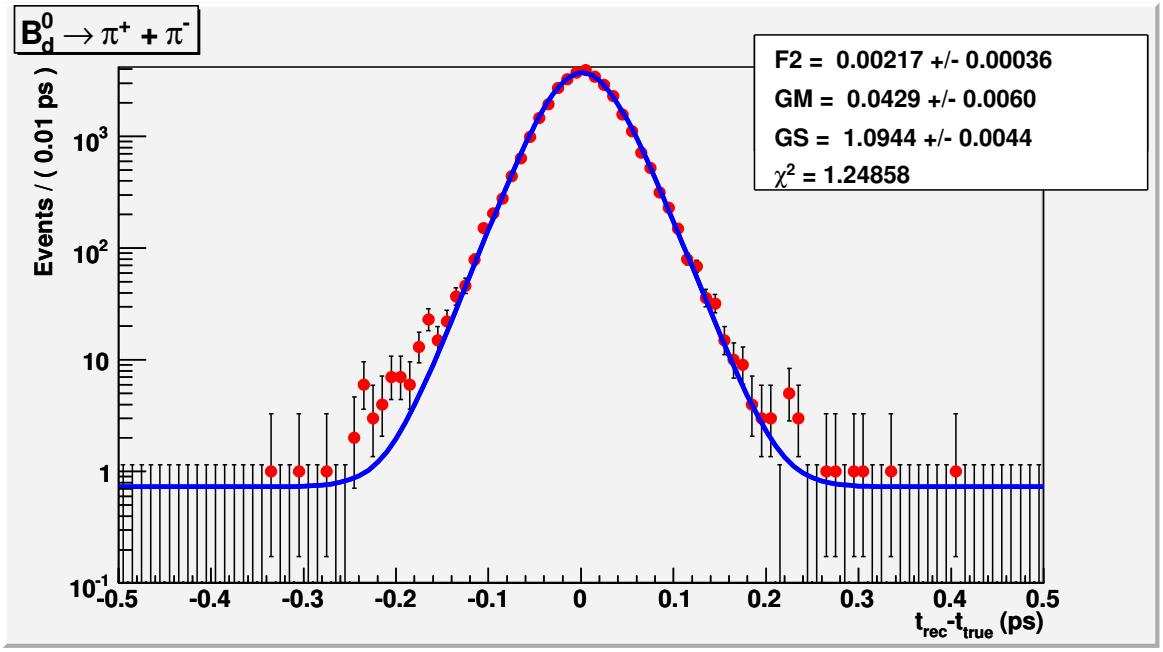

Figure 3: Simplified model global fit to the proper time residual distribution for $B_d \to \pi^+\pi^-$. The figure quoted as χ^2 is the χ^2 per degree of freedom of the fit.

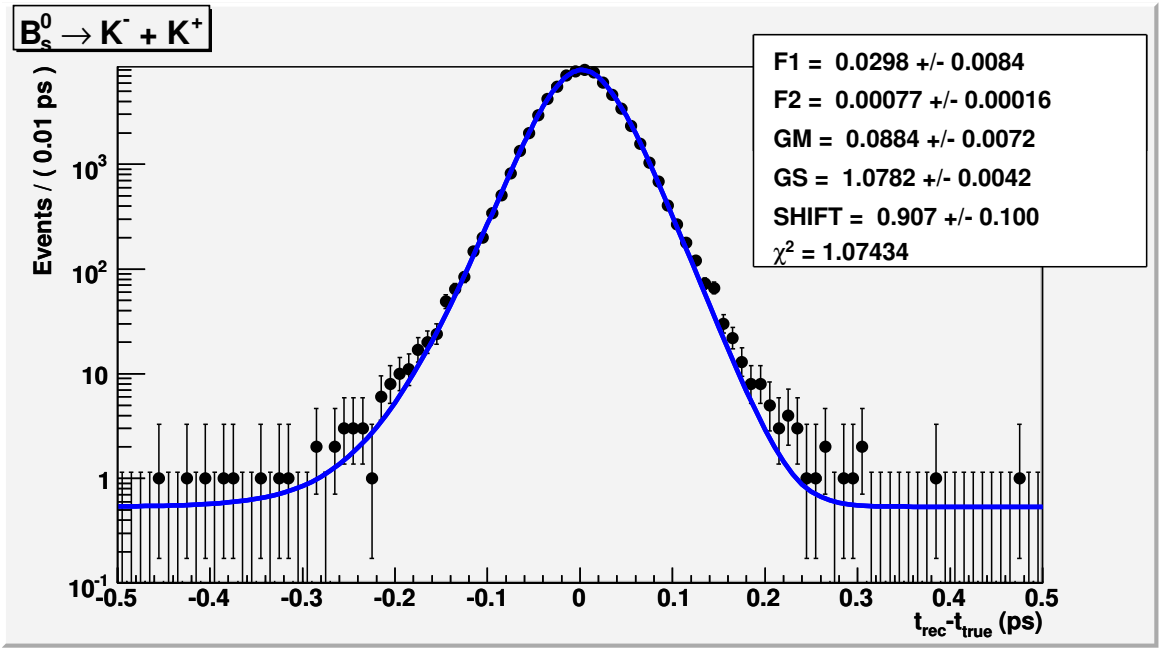

Figure 4: Full model global fit to the proper time residual distribution for $B_s \to K^+K^-$. The figure quoted as χ^2 is the χ^2 per degree of freedom of the fit.

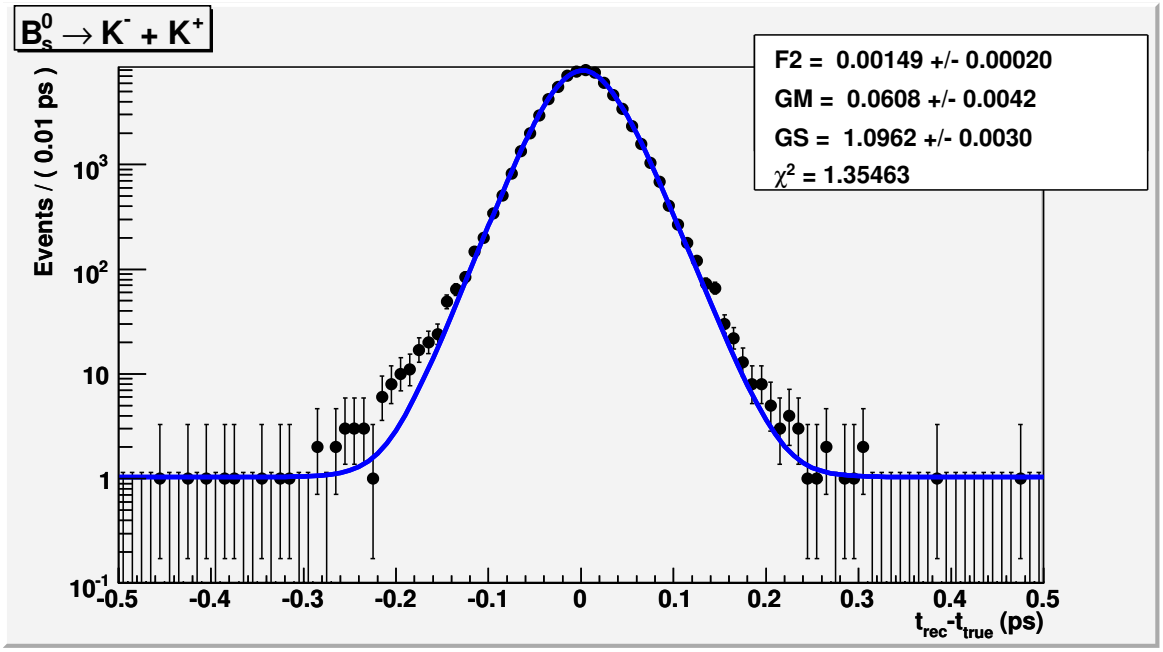

Figure 5: Simplified model global fit to the proper time residual distribution for $B_s \to K^+K^-$. The figure quoted as χ^2 is the χ^2 per degree of freedom of the fit.

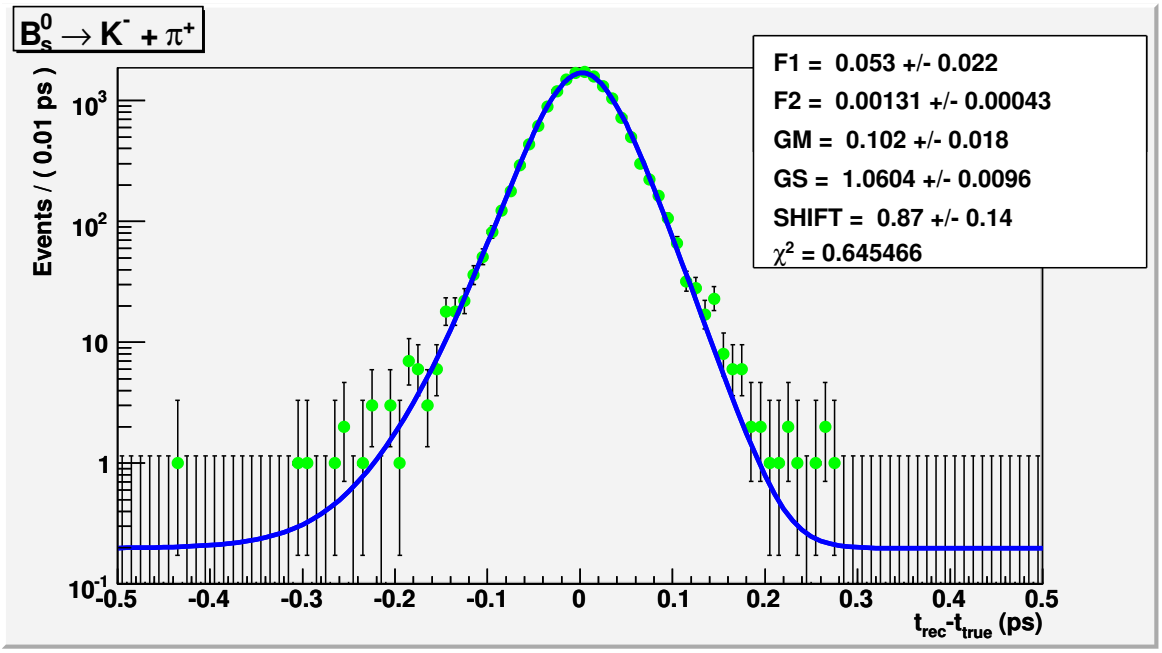

Figure 6: Full model global fit to the proper time residual distribution for $B_s \to K^-\pi^+$. The figure quoted as χ^2 is the χ^2 per degree of freedom of the fit.

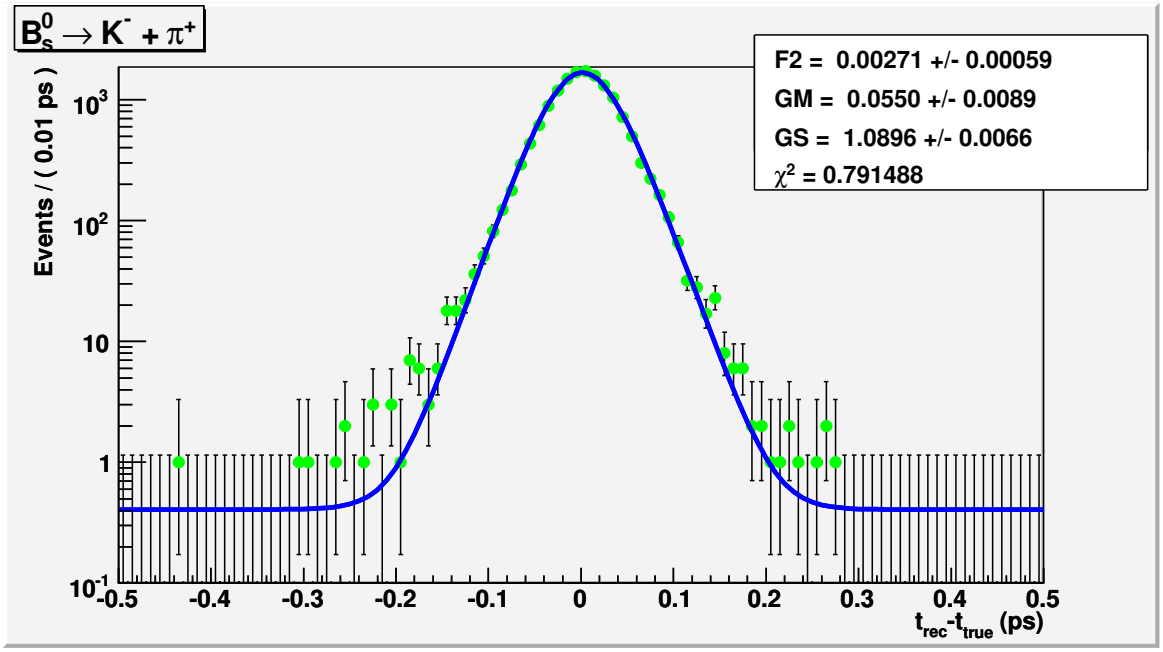

Figure 7: Simplified model global fit to the proper time residual distribution for $B_s \to K^-\pi^+$. The figure quoted as χ^2 is the χ^2 per degree of freedom of the fit.

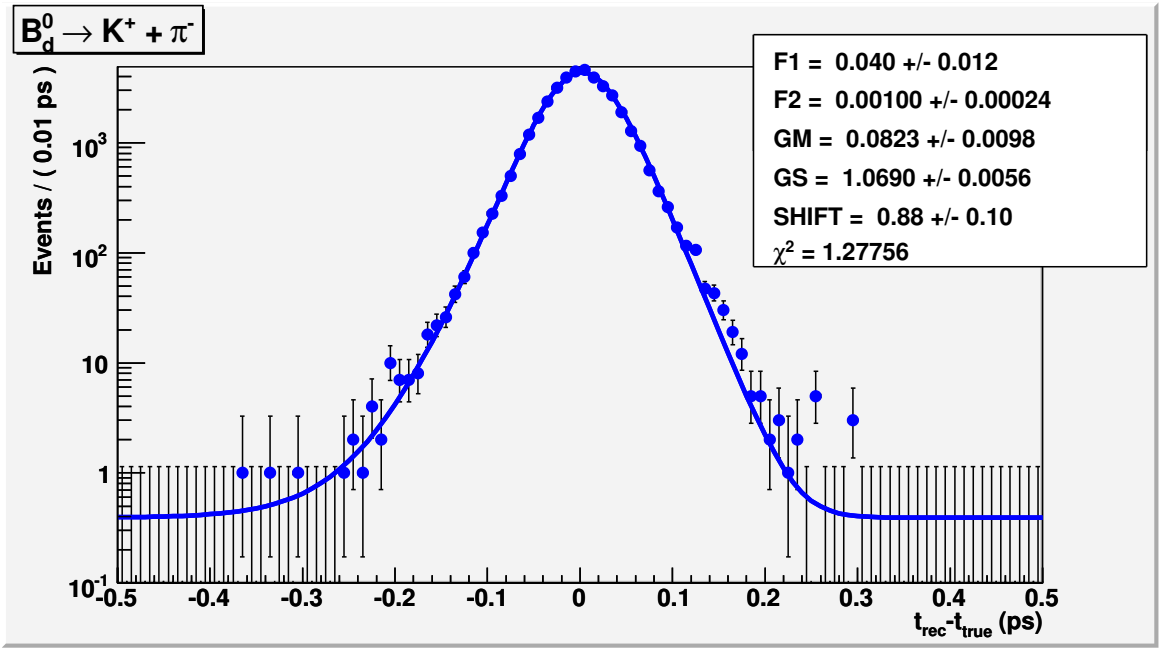

Figure 8: Full model global fit to the proper time residual distribution for $B_d \to K^+\pi^-$. The figure quoted as χ^2 is the χ^2 per degree of freedom of the fit.

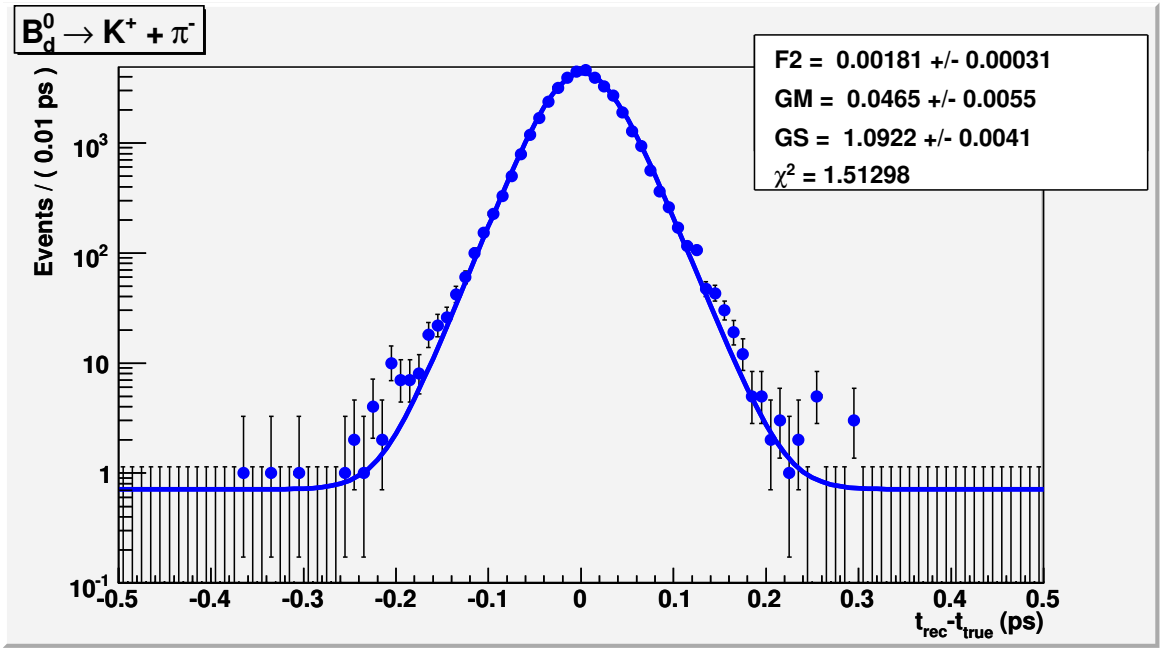

Figure 9: Simplified model global fit to the proper time residual distribution for $B_d \to K^+\pi^-$. The figure quoted as χ^2 is the χ^2 per degree of freedom of the fit.

Table 2: Comparison of Selection Efficiencies for $B \to h^+h^{'-}$ Channels. The total number of input events is given to the nearest 1,000.

Parameter	$B_d \to \pi\pi$	$B_s \to KK$ (stripped sample)	$B_s \to KK$ (signal tape)	$B_s \to K\pi$	$B_d \to K\pi$
Total Events on Input	260,000	196,000	5,000	116,000	298,000
Candidates Selected	33,674	69,789	375	15,095	39,360
Efficiency of Generator Level Cut	19.9%	34.6%	34.6%	20.4%	20.2%
Selection Efficiency	13.0%	n/a	7.5%	13.0%	13.2%
Overall Efficiency	2.6%	n/a	2.6%	2.7%	2.7%

3.4 Discussion of Model Validation

For all eight fits that were carried out, it can be seen both visually and from the fit χ^2 that very good fits are obtained over most values of $\Delta\tau$. However there tends to be, for both models, a small excess of events in the region $\Delta\tau \in [0.15, 0.25]$ps. There is also always a small excess, more significant for the simplified model fits than the full model fits, in the region $\Delta\tau \in [-0.25, -0.15]$ps.

Note also that in the simplified model fits the values of the three parameters have shifted slightly with respect to their values from the full model fit. This is because the fit is compensating for the lack of the $F1$ part in the model.

So both models provide an adequate description of the proper time residual distribution for $B \to h^+h^{'-}$ channels, with the full model giving a slightly better fit. However this comes at the price of introducing two extra parameters, $F1$ and $SHIFT$. It will be seen in Sec. 5 that these extra parameters cause problems when trying to extract the resolution model parameters from data.

Another problem with the full model fit is that the global correlation values from the fit for the $F1$ and $SHIFT$ parameters are always very high, around 0.9. For the simplified model fit the global correlation values are all very low (below 0.2). In addition, the full model fit in some channels returns a different result depending on the seed values of the fitted parameters (for example $F2$). In contrast, the simplified model fit is robust to changes in the seed values of the fitted parameters.

The results for both types of fit for all four $B \to h^+h^{'-}$ channels are summarised in Tables 3 and 4. The parameter values from the full model fit for all four channels are compatible with each other, within the fit errors. For the parameter values from the simplified model fit, the values of GS are compatible across channels, but there is some disagreement in the values of GM and $F2$ across certain channels at the level of $\approx 2\sigma$. However the absolute size of the differences is small (up to 0.12% for $F2$, and up to 0.018 for GM), so there is no evidence (within the statistical accuracy of this study) that the results from the fit to $B_s \to K^-\pi^+$ from data cannot be applied to the analysis of the other $B \to h^+h^{'-}$ channels. The lower statistics available in the $B_s \to K^-\pi^+$ channel lead to the fit errors in that channel being higher than in the other 3 channels. The lower statistics also lead to a lower χ^2 per degree of freedom for $B_s \to K^-\pi^+$. This is because none of the fits describe the tails of the distribution perfectly, and for the channels with more statistics the discrepancy between the data and the model in the tails becomes statistically more significant, increasing the χ^2 per degree of freedom.

Considering the effect of using a proper time resolution model that is a simple Gaussian with some fixed width (say the mean of the $\sigma_{\tau_{rec}}$ distribution) rather than a model (for example the full or simplified models presented here) which takes $\sigma_{\tau_{rec}}$ into account on a per-event basis, the main difference would be that the simple model would provide an adequate description of the $\Delta\tau$ distribution around the peak of the distribution, but would completely underestimate the tails of the distribution, due to the prescence of events with high values of $\sigma_{\tau_{rec}}$. If such a simple model were used in the analysis which studies the flavour tagged τ_{rec} distributions to extract CP asymmetries, extra systematic errors would most likely be introduced.

Table 3: Comparison of Resolution Model Parameters for $B \to h^+h^{'-}$ Channels, fitting 5 parameter model.

Parameter	$B_d \to \pi^+\pi^-$	$B_s \to K^+K^-$	$B_s \to K^-\pi^+$	$B_d \to K^+\pi^-$
F1	(4.1 ± 1.4)%	(3.0 ± 0.8)%	(5.3 ± 2.2)%	(4.0 ± 1.2)%
F2	(0.12 ± 0.03)%	(0.08 ± 0.02)%	(0.13 ± 0.04)%	(0.10 ± 0.02)%
GM	0.079 ± 0.011	0.088 ± 0.007	0.102 ± 0.018	0.082 ± 0.010
GS	1.071 ± 0.006	1.078 ± 0.004	1.060 ± 0.010	1.069 ± 0.006
SHIFT	0.89 ± 0.12	0.91 ± 0.10	0.87 ± 0.14	0.88 ± 0.10
χ^2/ndf of fit	0.91	1.07	0.65	1.28

Table 4: Comparison of Resolution Model Parameters for $B \to h^+h^{'-}$ Channels, fitting 3 parameter model.

Parameter	$B_d \to \pi^+\pi^-$	$B_s \to K^+K^-$	$B_s \to K^-\pi^+$	$B_d \to K^+\pi^-$
F2	$(0.22 \pm 0.04)\%$	$(0.15 \pm 0.02)\%$	$(0.27 \pm 0.06)\%$	$(0.18 \pm 0.03)\%$
GM	0.043 ± 0.006	0.061 ± 0.004	0.055 ± 0.009	0.047 ± 0.006
GS	1.094 ± 0.004	1.096 ± 0.003	1.090 ± 0.007	1.092 ± 0.004
χ^2/ndf of fit	1.25	1.35	0.79	1.51

4 Determination of Resolution Model Parameters on Data

4.1 Previous work on $B_d \to J/\psi K^*$ and $B_u \to J/\psi K^+$

In Ref. [3], a fit was made to the reconstructed untagged proper time (τ_{rec}) distribution to try to extract the parameters of the proper time resolution model in a manner that could be repeated using data. However the measured values for some of the parameters found by the fit to τ_{rec} did not match those found by the fit using the Monte Carlo simulated truth information (i.e. the fit to the proper time residual, $\Delta\tau$). Note that for the $J/\psi X$ channels there is some sensitivity to the proper time resolution model in the untagged τ_{rec} distribution, because the selection for $J/\psi X$ does not use impact parameter cuts on the B decay products. This causes the τ_{rec} distribution to extend slightly into the unphysical negative region, by an amount dependent on the proper time resolution. However for $B \to h^+h^{'-}$ channels, impact parameter cuts are an important part of the selection, so the shape of the τ_{rec} distribution at the point where events start to pass the impact parameter cuts depends on these cuts as well as on the proper time resolution. So there is no real sensitivity to the proper time resolution model in the untagged τ_{rec} distribution for a $B \to h^+h^{'-}$ channel. However the *flavour tagged* τ_{rec} distribution for a $B \to h^+h^{'-}$ channel can be sensitive to the proper time resolution model, in a way discussed in Sec. 4.2. Hence it is the flavour tagged τ_{rec} distribution which is considered in the following sections. The study in these sections will show that for the $B_s \to K^-\pi^+$ channel, a fit to the flavour tagged τ_{rec} distribution can reproduce the parameters of the proper time resolution model correctly.

4.2 Choice of Channel for Parameter Determination for $B \to h^+h^{'-}$

Of course with data the residual distribution cannot be formed. However, as described above, the parameters of the resolution model are only a function of the event-by-event error. So the parameters of the model can be extracted by fitting to the reconstructed proper time distribution for flavour tagged events, making use of the knowledge of the event-by-event error. Because the motivation is to provide input to the analysis of $B_d \to \pi^+\pi^-$ and $B_s \to K^+K^-$, and decay channels with different topologies can have different proper time resolutions, a $B \to h^+h^{'-}$ channel should be studied. The candidates are the four $B \to h^+h^{'-}$ channels which have the highest branching ratios of the possible $B \to h^+h^{'-}$ channels: $B_d \to \pi^+\pi^-$, $B_s \to K^+K^-$, $B_s \to K^-\pi^+$ and $B_d \to K^+\pi^-$ (these are the same channels which the proper time resolution model was applied to in Sec. 3).

Since the final states for the decay channels $B_d \to \pi^+\pi^-$ and $B_s \to K^+K^-$ are not flavour-specific, the B meson flavour at decay cannot be determined. Hence the oscillations between B and anti-B will not show up explicitly in the τ_{rec} distribution, even if the initial flavour of the B meson has been tagged. However for a channel with a flavour-specific final state, the flavour of the B meson at decay is automatically tagged by identifying the decay products. So each B meson can be classified as having oscillated an even or odd number of times before decaying, by using the flavour tag of the B meson at production. Hence the flavour tagged proper time distribution *will* show explicit oscillations. These oscillations will be diluted due to the non-zero mistag probability, and the finite proper time resolution of the detector. Hence, assuming the mistag value is known from other measurements (see Sec. 4.3), the explicit oscillations give information on the proper time resolution of the detector.

A gross illustration of this effect is shown in Fig. 10. Degradation of the proper time resolution is seen to affect the oscillations for the B_s channel. From this it can be concluded that the channel offering the most sensitivity to proper time resolution effects is $B_s \to K^-\pi^+$.

4.3 Mistag Value

It is not possible to fit for the mistag fraction ω and the proper time resolution model simultaneously, because they both have the effect of diluting the observed oscillations, and so will be very highly correlated in any fit. The effect of having a different mistag rate is shown in Fig. 11. An increased value of ω is seen to cause a

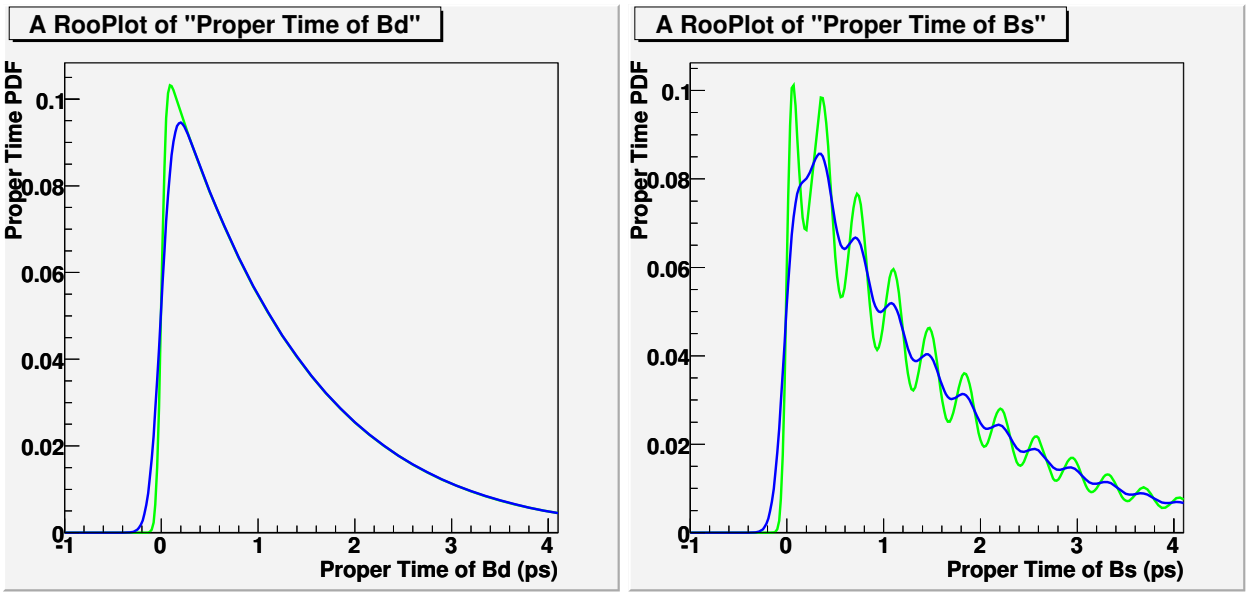

Figure 10: Illustration of the sensitivity of a flavour tagged τ_{rec} distribution to the resolution model. Two toy flavour tagged τ_{rec} distributions for a flavour-specific decay of a neutral B meson are shown. The difference between the left and right plots is the Δm_q value assumed—0.51 ps^{-1} on the left, and 17 ps^{-1} on the right. In each plot, both PDFs use a simple resolution model which is just a Gaussian of fixed width. For the PDF in green the fixed width is 40fs, for the PDF in blue it is 100fs.

small dilution in the observed depth of the oscillations, which is the same kind of effect seen in Fig. 10 when changing the width of the proper time resolution model.

Hence this method requires that the mistag already be fixed using information from control channels.

Within LHCb a standard control channels approach has been developed [7] to monitor the mistag fraction using data. This approach measures the mistag rate in a control channel and then calculates the mistag rate in the signal channel by considering the differences between the phase space of the B in the two channels (for example the p_T spectra) and the tagging category used (i.e. whether the event triggered due to the signal B or due to the other B in the event).

The control channel for measuring the mistag in $B \to h^+h^{'-}$ channels on data will be $B_d \to K^+\pi^-$. Using the standard LHCb method the mistag is expected to be measured with a precision of around 1% [7].

Another possible strategy is to start out by measuring the mistag in the channel $B_d \to K^+\pi^-$ by studying its flavour tagged τ_{rec} distribution. By the arguments set out in Sec. 4.2, the τ_{rec} PDF in this channel will be sensitive to ω, but not the proper time resolution model or CP violation effects. Hence a fit to the $B_d \to K^+\pi^-$ flavour tagged τ_{rec} distribution should give a value for ω in that channel.

The value of ω for $B_s \to K^-\pi^+$ will be lower than the ω for $B_d \to K^+\pi^-$, because B_s channels can make use of the same side kaon tagger [7]. The amount of improvement in the mistag rate that this brings can be estimated by studying a pair of channels where the size of the effect can be measured, for example $B_s \to D_s^-\pi^+$ and $B_d \to D^-\pi^+$.

These two different methods can then be applied independently to $B_s \to K^-\pi^+$ to reduce the uncertainty on the mistag rate.

4.4 Backgrounds to $B_s \to K^-\pi^+$

The size and composition of the expected backgrounds to the $B \to h^+h^{'-}$ channels have been studied in detail [8]. It is assumed that only B events will pass the selection. This is because any short-lived background should be removed by the cuts which are made on the impact parameter significance of the final state particles and flight distance significance of the B. This assumption can be tested during early data taking, when data will be taken with a random trigger to allow the nature of minimum bias events to be studied in detail. This leaves two types of background to consider—specific background and inclusive $b\bar{b}$ background.

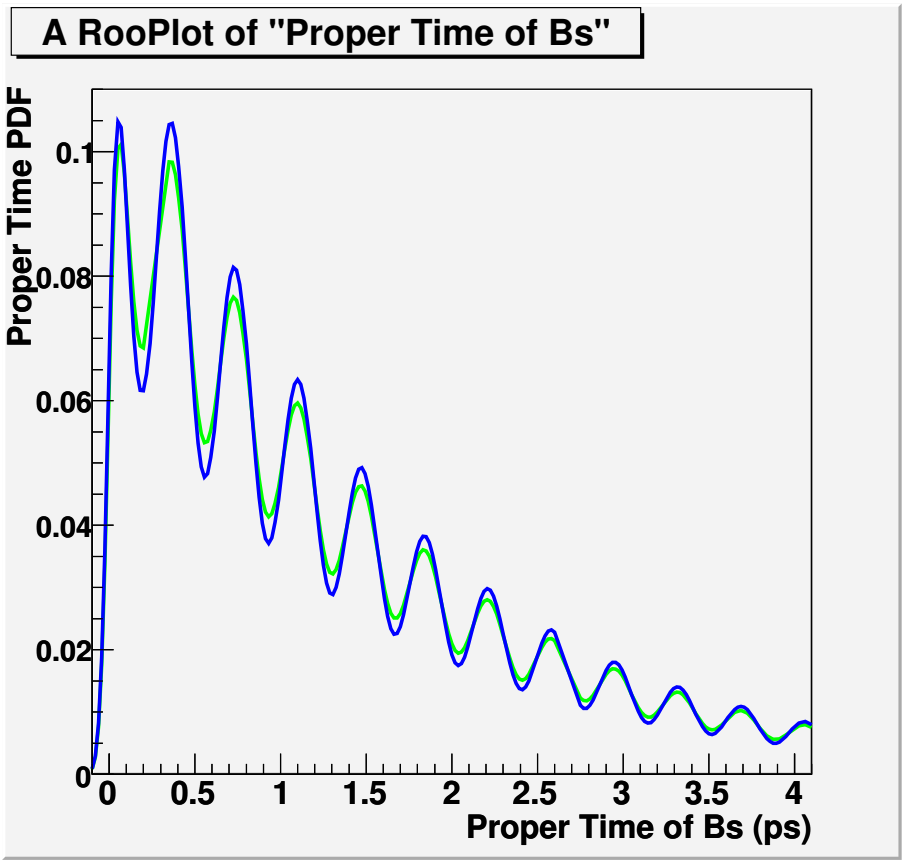

Figure 11: Illustration of the impact of ω on the flavour tagged τ_{rec} distribution for a flavour-specific decay of a B_s meson. The only difference between the PDFs is the ω value assumed—30% for the PDF in blue, and 35% for the PDF in green. Both PDFs use a simple resolution model which is just a Gaussian with a fixed width of 40fs.

4.4.1 Specific Background

There will be specific backgrounds arising from other $B \to h^+h^{'-}$ decays. The dominant such background for $B_s \to K^-\pi^+$ comes from $B_d \to K^+\pi^-$ decays. This is because this background cannot be reduced by cutting on the particle identification (PID) of the detected hadrons, since the final state is the same as the signal. The only discriminating variable available is the invariant mass of the $h^+h^{'-}$ pair. The expected background to signal ratio from specific background has been calculated to be 0.5 [8]. The level of contamination is significant because the expected yield of selected $B_d \to K^+\pi^-$ events is more than 10 times the expected yield of $B_s \to K^-\pi^+$ events (135,000 versus 9,800 in 2 fb^{-1}). This factor arises due to both the higher probability for a b quark to hadronise into a B_d meson rather than a B_s meson, and the higher branching ratio for $B_d \to K^+\pi^-$ compared to $B_s \to K^-\pi^+$.

The flavour tagged proper time distribution for $B_d \to K^+\pi^-$ will differ from that for $B_s \to K^-\pi^+$ due to the different values of Δm_q and τ_{B_q}, and from the slightly higher mistag rate. Table 5 shows the different inputs for the signal and the backgrounds. These values are assumed to have been already well measured via the application to the data of various analysis methods. For example the mistag rate can be measured as described in Sec. 4.3, and the B meson lifetimes will be measured by studying the untagged τ_{rec} distributions in high-yield channels such as the $B \to Dh$ channels.

Table 5: Comparison of inputs for the proper time distribution of the signal and both backgrounds

Input to flavour tagged τ_{rec} PDF	Signal	Specific Background	Inclusive $b\bar{b}$ Background
Proper Time Resolution Model	Full 5-Parameter Model (see Eqn. 1)	Single Gaussian with width fixed to 40fs	Single Gaussian with width fixed to 40fs
Lifetime	1.47 ps	1.53 ps	1.01ps
Mistag Rate ω	34%	37%	50% (i.e. untagged)
Oscillation Frequency	17 ps^{-1}	0.507 ps^{-1}	n/a

4.4.2 Inclusive $b\bar{b}$ Background

As well as specific background, there will also be combinatoric background coming from from inclusive $b\bar{b}$ events. These are events where the final state particles come from decays of B hadrons, but not from the same B meson. This (along with the backgrounds from other $B \to h^+h^{'-}$ decays) is expected to be a more significant background than partially reconstructed decays from the same B meson (for example $B \to K^*\pi$ and $B \to \rho\pi$). This is because partially reconstructed final states will have a measured invariant mass which sits below the B meson mass. The expected background to signal ratio from inclusive $b\bar{b}$ background has been calculated to be 1.9 [8]. This value is higher than for the other $B \to h^+h^{'-}$ channels, due to the lower number of selected $B_s \to K^-\pi^+$ events.

The proper time distribution for inclusive $b\bar{b}$ background will be similar to that for the $B \to h^+h^{'-}$ channels, because the background has to survive the same cuts as the signal. The proper time distribution for inclusive $b\bar{b}$ events passing the selection have been fitted for [8]. The inputs for inclusive $b\bar{b}$ background in Table 5 reflect the result of this fit. In data there will be many inclusive $b\bar{b}$ events in the sidebands of the invariant mass distribution which can be used to determine the proper time distribution for such events very precisely.

4.5 Construction of Flavour Tagged τ_{rec} PDFs

4.5.1 Methodology for Fit

To provide input for the fit studies, samples of simulated "toy data" were generated. The toy data consists of a given number of events which follow a distribution, or PDF, which resembles as closely as possible the distribution of reconstructed proper time (τ_{rec}) that will be seen in data for selected and flavour tagged $B_s \to K^-\pi^+$ events. This distribution will have contributions from signal events, specific background events and inclusive background $b\bar{b}$ events, and the proper time resolution model will be fixed, as described in the following sections.

To fit to these events a distribution, or PDF, is constructed which is similar to the previous one, but in which the parameters of the proper time resolution model are not fixed. Rather, in this PDF these parameters will be seeded with values which are (in general) not equal to the values used to generate the toy data. This second PDF will then be fitted to the toy data, with the goal being that the parameters of the proper time

resolution model in the second PDF after the fit will match those in the PDF which was used to generate the toy data.

Note that the fit is an unbinned minus-log-likelihood fit, carried out using the RooFit fitting package [9], which operates within the framework of the ROOT data analysis framework [10]. The minimisation of the minus-log-likelihood is carried out by the Minuit package [11].

4.5.2 Flavour Tagged τ_{rec} PDFs for Toy Data

The first step in constructing the signal ($B_s \to K^-\pi^+$) flavour tagged τ_{rec} PDF for the toy data is to define a proper time resolution model. Eqn. 1 is used, with the fixed values of $F1$, GS etc. being guided by the values found (see Sec. 3.3) by the fits to the DC06 $\Delta\tau$ distributions for $B_s \to K^-\pi^+$ and $B_s \to K^-K^+$.

This model is then used as one of the inputs to a PDF which is designed to describe the flavour tagged τ_{rec} distribution, taking into account physics parameters, detector resolution and acceptance (see Table 5). Since the backgrounds are insensitive to the details of the resolution model, they are given a simple resolution model, being a single Gaussian of width 40fs. Then a PDF similar to the signal PDF, but with different physics parameters, is created. Next, a PDF, say $P(\sigma_{\tau_{rec}})$, is formed from the $\sigma_{\tau_{rec}}$ distribution, using kernel estimation. The signal PDF described above is then made conditional on $P(\sigma_{\tau_{rec}})$. This step defines what the proper time resolution model for the toy data is, because its parameters are functions of $F1$, GS etc. and $P(\sigma_{\tau_{rec}})$.

The PDF for the toy data is formed by adding the signal PDF and the two background PDFs together, with the B/S ratios fixed to values close to those found in [8]. The resolution model parameters and B/S ratios used are summarised in Table 6.

Table 6: Inputs for Resolution Model Toy Jobs.

Input to τ_{rec} PDF	5 Parameter Fit	3 Parameter Fit
F1	3.1%	0%
F2	0.25%	0.35%
GM	0.08	0.05
GS	1.08	1.10
SHIFT	0.95	n/a
Inclusive B/S	1.94	1.94
Specific B/S	0.52	0.52

Finally the toy data are generated following this PDF. An example toy dataset is shown in Fig. 12. This dataset contains 20,300 events, 6,000 of which are signal events. This corresponds to what is expected after 2 fb^{-1} of LHCb data taking.

4.5.3 Flavour tagged τ_{rec} PDFs for Fit PDF

For the signal PDF making up part of the fit PDF, the same resolution model (Eqn. 1) is used, but the parameters are seeded using a value uniformly chosen from some "reasonable" interval for each parameter. During the fit the five resolution model parameters are allowed to float in some interval which is larger than the seed interval.

The seeding intervals and float intervals for each parameter are given in Table 7. These seeding intervals are supposed to represent what range of values it is physically reasonable for the parameter to take. The floating intervals should be as large as possible while still being logical (e.g. percentages must not be negative) and limiting the fit values to a range where it can be easily seen if the fit goes awry. The intervals are not sensitive to the details of what the resolution model in a given channel looks like. Note that the B/S values are represented in the fit by fractions of each background relative to the total number of events. Hence the values for the parameters ("Spec" for the specific background and "Incl" for the inclusive background) must lie between 0 (0%) and 1 (100%).

This resolution model is then used as an input to a flavour tagged τ_{rec} distribution with the same physics parameters and mistag as the toy data for the PDF. The difference is that the parameters of the resolution model are now floating rather than fixed. Again, the signal PDF is made conditional on $P(\sigma_{\tau_{rec}})$. This allows the PDF to be sensitive to changes in the values of $F1$, GS and so on.

The input for the background PDFs for the fit are the same as those for the background PDFs for the toy data. However when the PDFs are added together, the B/S ratios are now floating. This gives a total of seven floating parameters - the five parameters of the resolution model and the two B/S ratios.

Note that in practice a measurement of the B/S ratios will be made by studying the invariant mass distributions. Such constraints on the B/S ratios could be used as an input to the fit described here, rather than

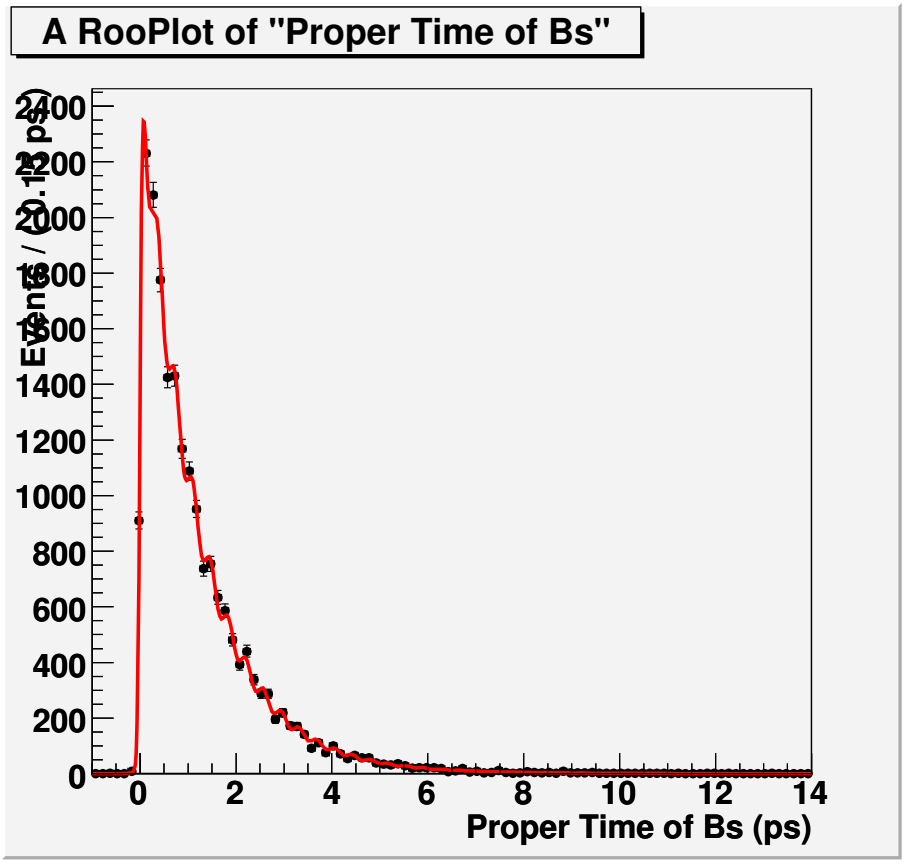

Figure 12: Example of a toy dataset including $B_s \to K^-\pi^+$ signal events, as well as specific background events and inclusive $b\bar{b}$ background events. The total number of events is 20,300. The red line shows the PDF from which the toy data is generated.

Table 7: Interval used to generate fit PDF and perform the fit.

Parameter	Seed Interval	Float Interval
F1	[0.5%,20%]	[0%,90%]
F2	[0.1%,5%]	[0%,90%]
GM	[-2.0,2.0]	[-100,100]
GS	[0.5,2.0]	[0.001,100]
SHIFT	[0.5,10]	[0.001,100]
Spec	[5%,80%]	[0%,99%]
Incl	[10%,90%]	[1%,99%]

allowing the B/S ratios to be freely fitted for. It is expected that doing this would lead to reduced errors on the measurement of the resolution model parameters. The advantage of the method used here is that it provides a check on the measurement of the B/S ratios.

An example fit to a toy dataset containing 20,300 total events is shown in Fig. 13. It can be seen that both backgrounds and the total PDF converge to the relevant PDFs used to make the toy data.

4.6 Setup for Toy Monte Carlo Simulation Study

The toy data PDF is the same for each toy experiment, but the generation of the toy data points from this PDF is done using a different seed for each toy experiment. For the fit PDF the background PDFs are the same each time, but the signal PDF has different resolution model parameters each time, uniformly drawn (according to a different seed each time) from a reasonable interval around the toy data values for each resolution model parameter. The two B/S ratios are similarly randomly seeded each time.

Since the number of selected $B_s \to K^-\pi^+$ events per nominal year of LHCb running (equivalent to 2 fb^{-1} integrated luminosity) is expected to be ∼10,000, and the tagging efficiency is around 60%, around 6,000 tagged $B_s \to K^-\pi^+$ events are expected per nominal year. It follows that around 30,000 tagged $B_s \to K^-\pi^+$ events are expected after 10 fb^{-1} (the total integrated luminosity that will be collected during the data collection period of the current LHCb detector).

As the purpose of extracting the resolution model parameters is to use them as input to the fit for γ, a fit which requires a considerable amount of data to produce a significant result, there is little point in attempting to use the method described here on very few events. Two separate sets of toy experiments are run - one corresponding to 2 fb^{-1} of data, and one corresponding to 10 fb^{-1} of data.

At each luminosity 300 sets of toy data are generated, and each set is fitted to using a PDF with different seeds as described above. For each of the seven floating parameters the distributions for the following quantities are plotted:

- The value of the parameter as returned by the fit.
- The error assigned by the fit to this value.
- The residual of the parameter, i.e the fitted value minus the value in the toy data.
- The pull of the parameter, i.e the residual divided by the error.

To monitor how the fit has performed, for each parameter single Gaussians are fitted to the fitted values, residuals and pulls distributions. (The exception to this is the parameter $F2$, which for 2 fb^{-1} of data is often fitted to zero as the size of the fitted error is about the same size as the seed value of $F2$. In this case fitting a Gaussian to the fitted values or residuals distributions is not worthwhile.)

The reliability of the fit result can also be monitored by studying the correlations between the different variables in the fit. For each variable, the minimisation package Minuit calculates a "global correlation", which combines the correlations of that variable with each other variable in the fit to estimate the overall independence of that variable in the fit. The distributions of the global correlations of each variable are also plotted.

For all plots, only fits that returned a full and accurate covariance matrix are shown.

5 Results from Fit to Toy Data

5.1 Results Fitting the Full Model to Data Made Using the Full Model

Toy data are made following the parameters from Fig. 6. Then a 7 parameter fit (the 5 resolution model parameters plus the 2 background to signal ratios) is carried out. Out of 300 fits only 143 good covariance

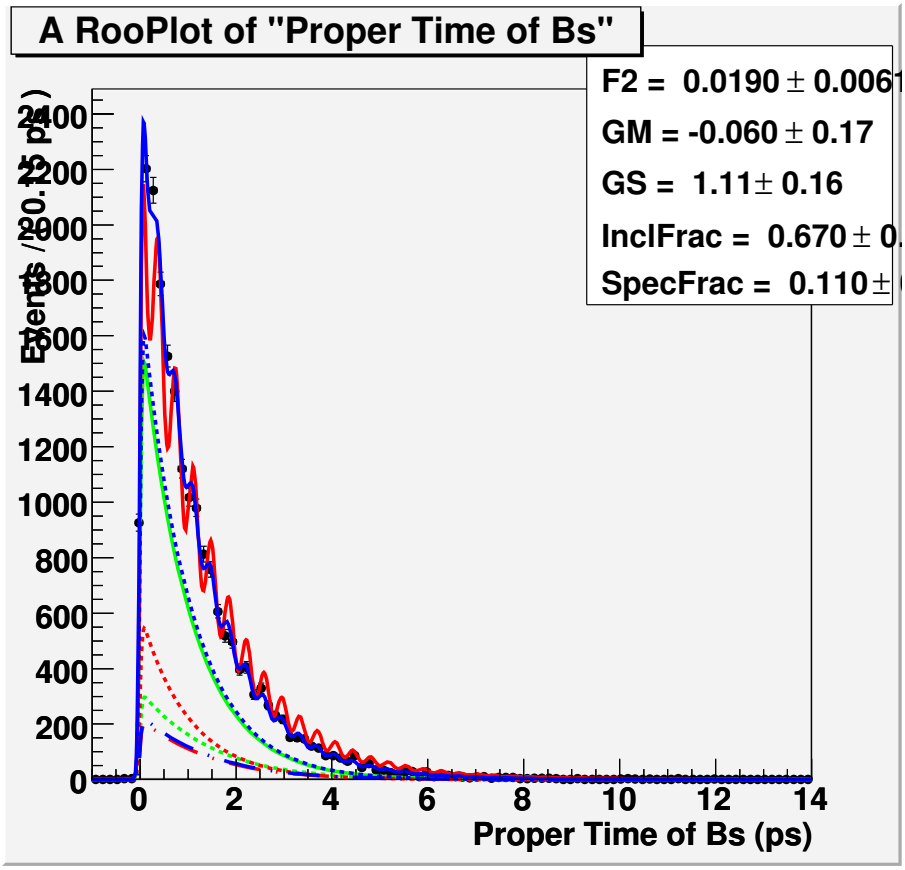

Figure 13: Example of a fit to a flavour tagged τ_{rec} distribution containing $B_s \to K^-\pi^+$ signal events, specific background events and inclusive $b\bar{b}$ background events. The total number of events is 20,300. The solid red line shows the total PDF from which the toy data is generated, and the solid blue line is the fit PDF after the fit. The background PDFs used to make the toy data are shown in green (solid line for inclusive $b\bar{b}$ background, dashed line for specific background). The background PDFs in the fit PDF *before* the fit are shown in red (dashed red line for inclusive, and broken red line for specific). The background PDFs in the fit PDF *after* the fit are shown in blue (dashed blue line for inclusive, and broken blue line for specific).

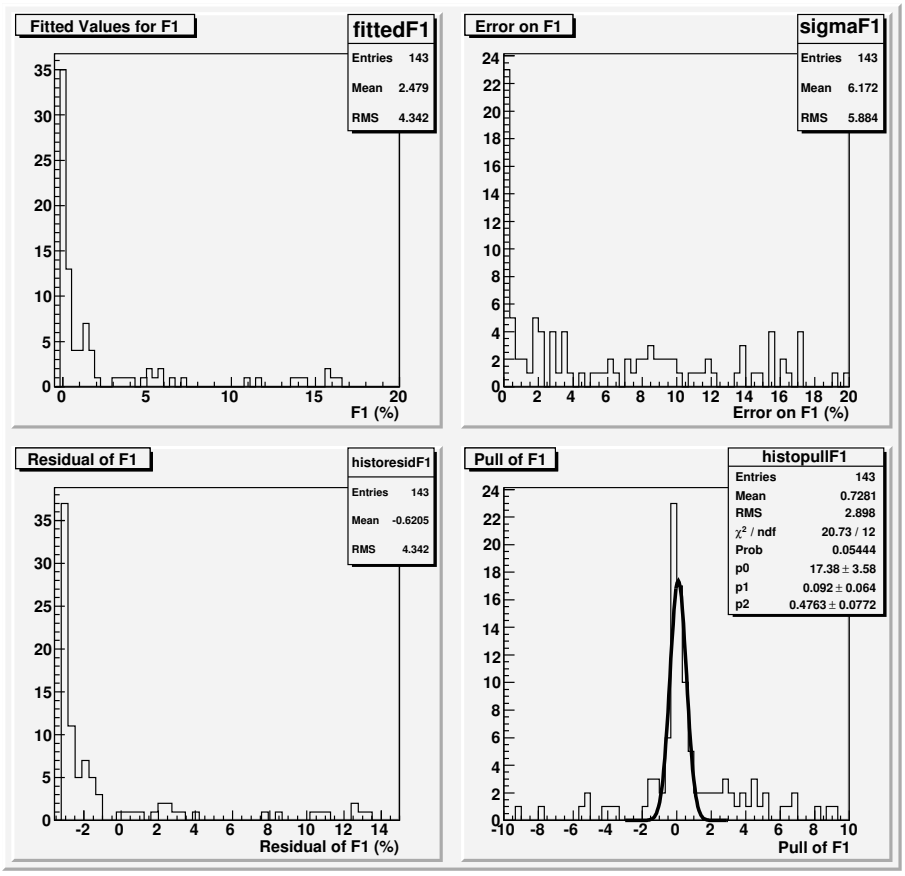

Figure 14: F1: fitting full model to 2 fb^{-1} of data simulated using the full model. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.

matrices are found. The plots are shown in Figs. 14 to 21.

5.2 Discussion

Clearly there are a lot of outliers in all the fitted values distributions, and for F1 and SHIFT the fitted values and pulls distributions are not at all Gaussian. Furthermore, the pull widths are not close to 1 for some of the other parameters (particularly GM).

Another problem with the full model fit, in addition to the poor state of many of the resulting distributions, is the fact that less than half (143 out of 300) toy fits returned an accurate covariance matrix. Having more than 50% bad covariance matrices is far above the level of a few percent that could be tolerated. Finally, the global correlations of the parameters reach very high values in some fits, with F1 and SHIFT in particular often having global correlations greater than 0.9.

The problems with these fits may be due to the fact that there are too many free parameters in the fit, and not enough independent (i.e. uncorrelated) sensitivity to them in the τ_{rec} distribution. One way to avoid this is to use the simplified model, which was shown in Sec. 3.3 to describe the proper time residual distributions almost as well as the full model.

5.3 Results Fitting the Simplified Model to Data Made Using the Full Model

To try and avoid the problems seen by fitting the full model, one can try fitting a simplified model to data made using the full model. With this configuration it is found that nearly all the covariance matrices are good (295 out of 300). This represents a very large improvement on what was seen when fitting the full model. Again, only fits which returned good covariance matrices are included in the following plots (the same is true of the plots in Sec.5.5).

The plots are given in Figs. 22 to 27.

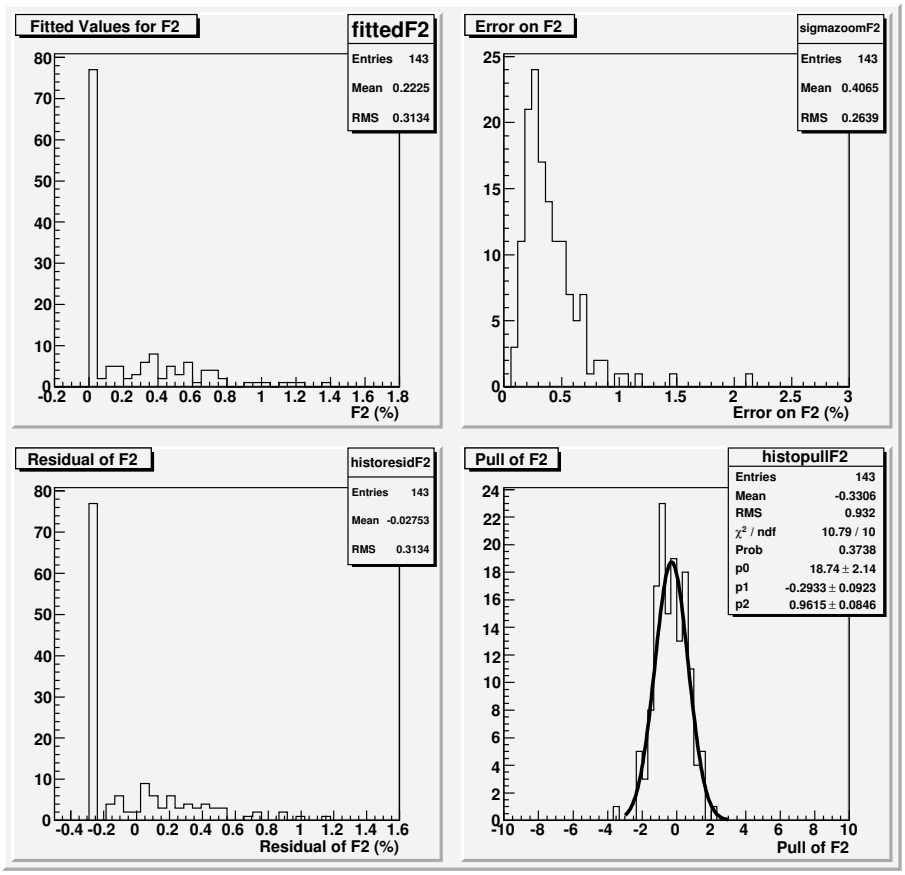

Figure 15: F2: fitting full model to 2 fb^{-1} of data simulated using the full model. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.

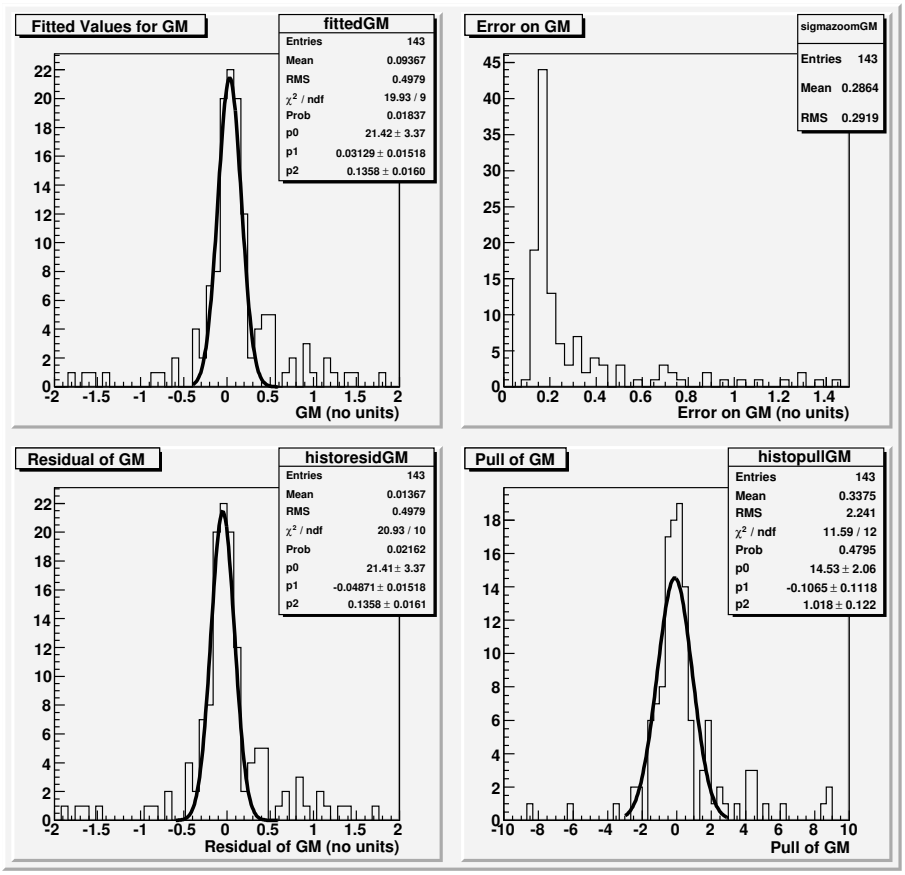

Figure 16: GM: fitting full model to 2 fb^{-1} of data simulated using the full model. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.

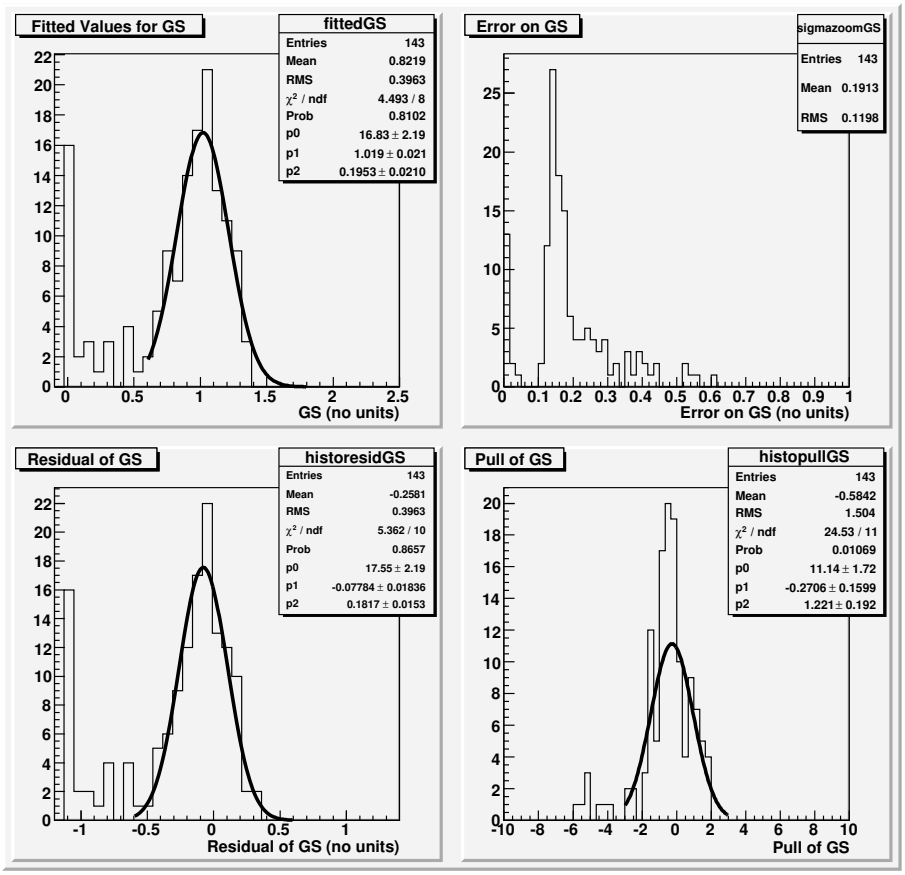

Figure 17: GS: fitting full model to 2 fb^{-1} of data simulated using the full model. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.

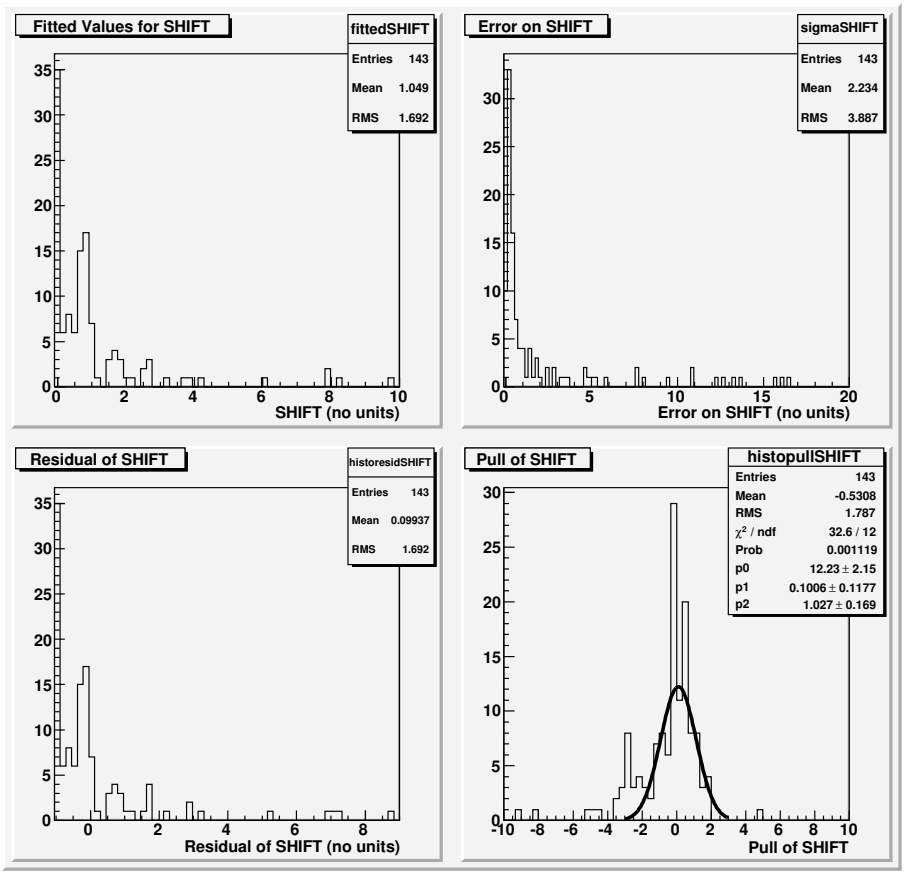

Figure 18: SHIFT: fitting full model to 2 fb^{-1} of data simulated using the full model. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.

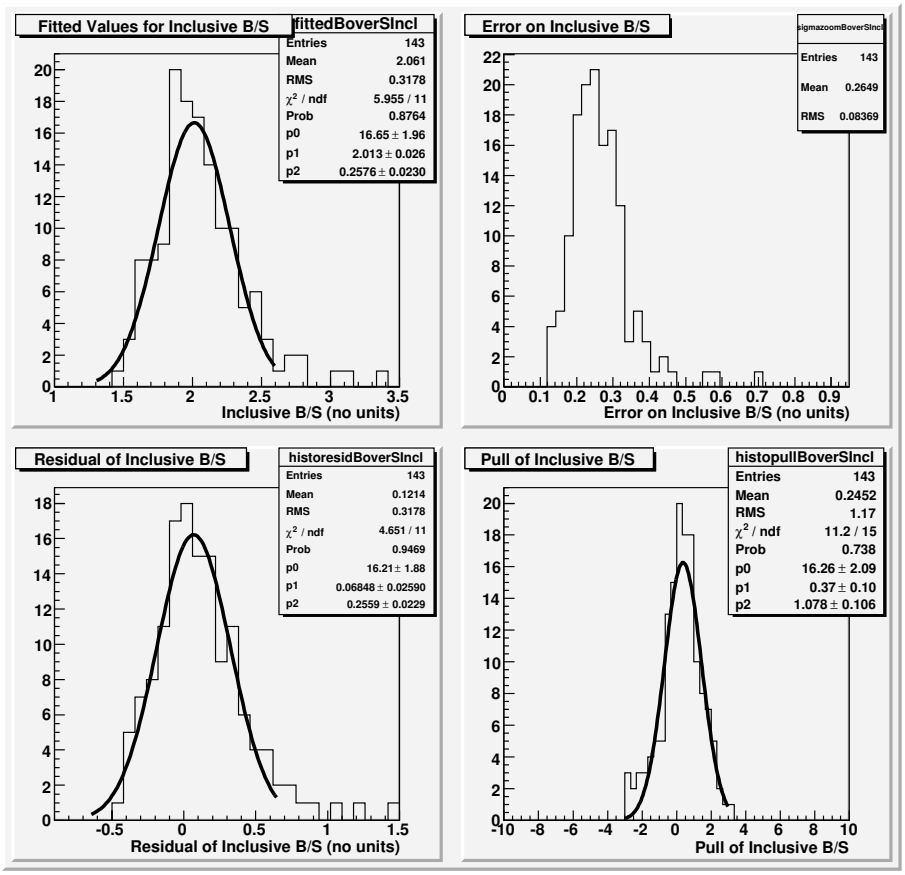

Figure 19: Inclusive Background to Signal Ratio: fitting full model to 2 fb^{-1} of data simulated using the full model. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.

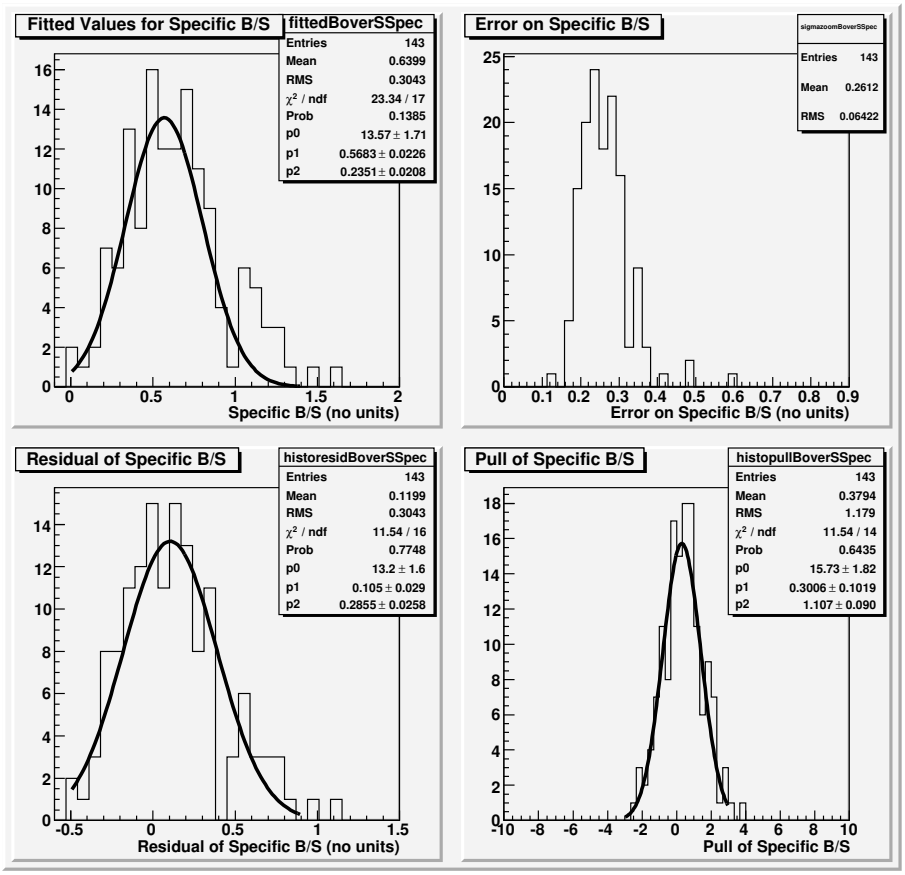

Figure 20: Specific Background to Signal Ratio: fitting full model to 2 fb^{-1} of data simulated using the full model. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.

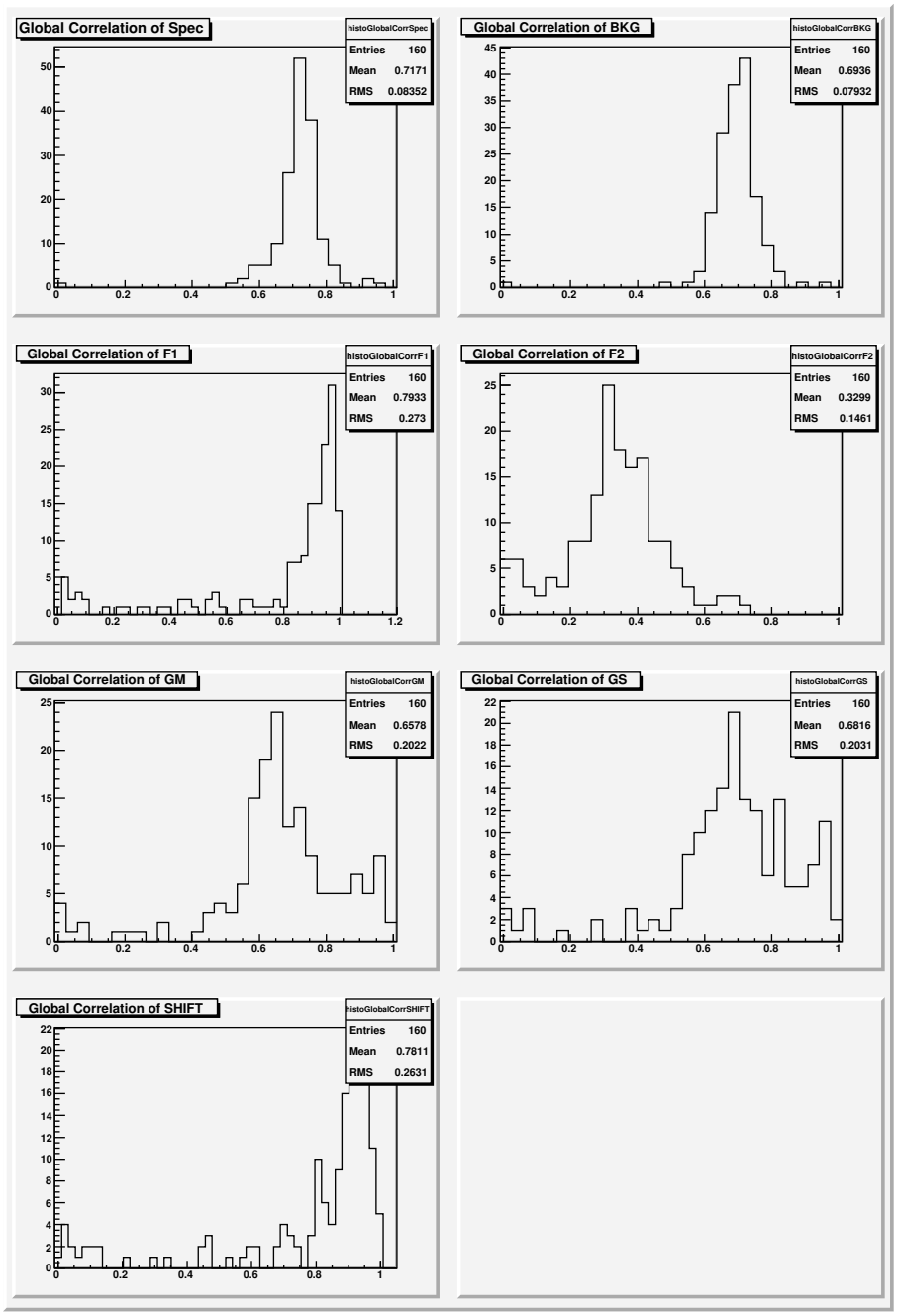

Figure 21: Global correlations: fitting full model to 2 fb^{-1} of data simulated using the full model. Clockwise from top left are the distributions for: Specific Background to Signal Ratio, Inclusive Background to Signal Ratio, F2, GS, SHIFT, GM and F1.

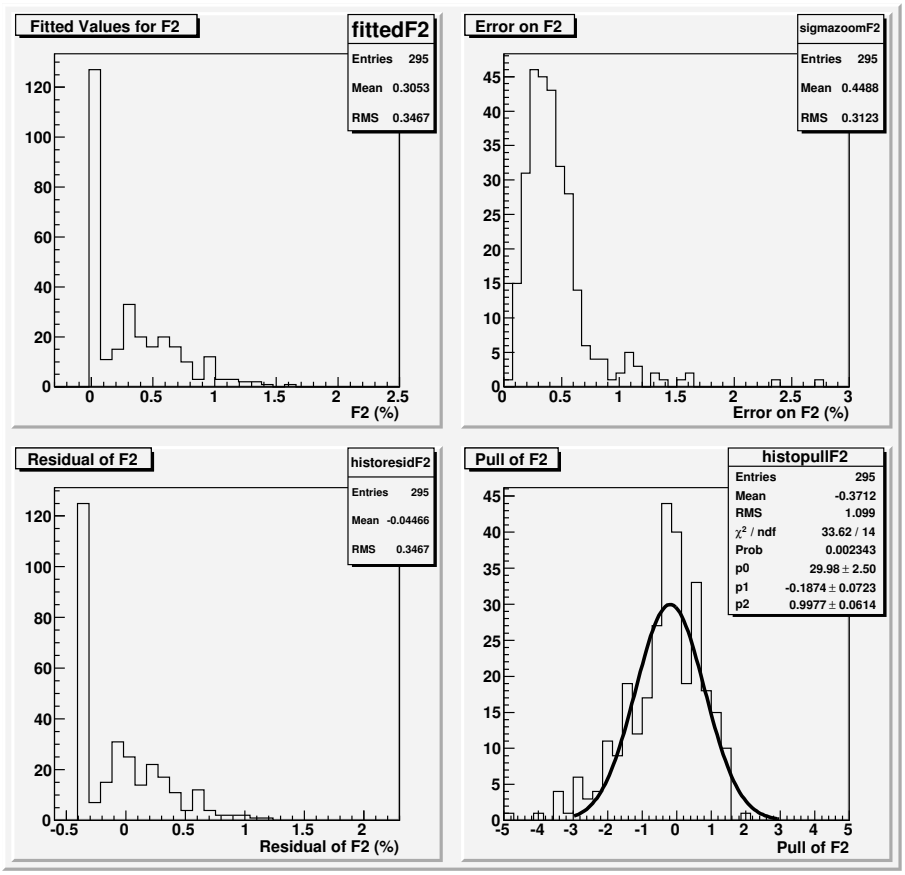

Figure 22: F2: fitting simplified model to 2 fb^{-1} of data simulated using the full model. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.

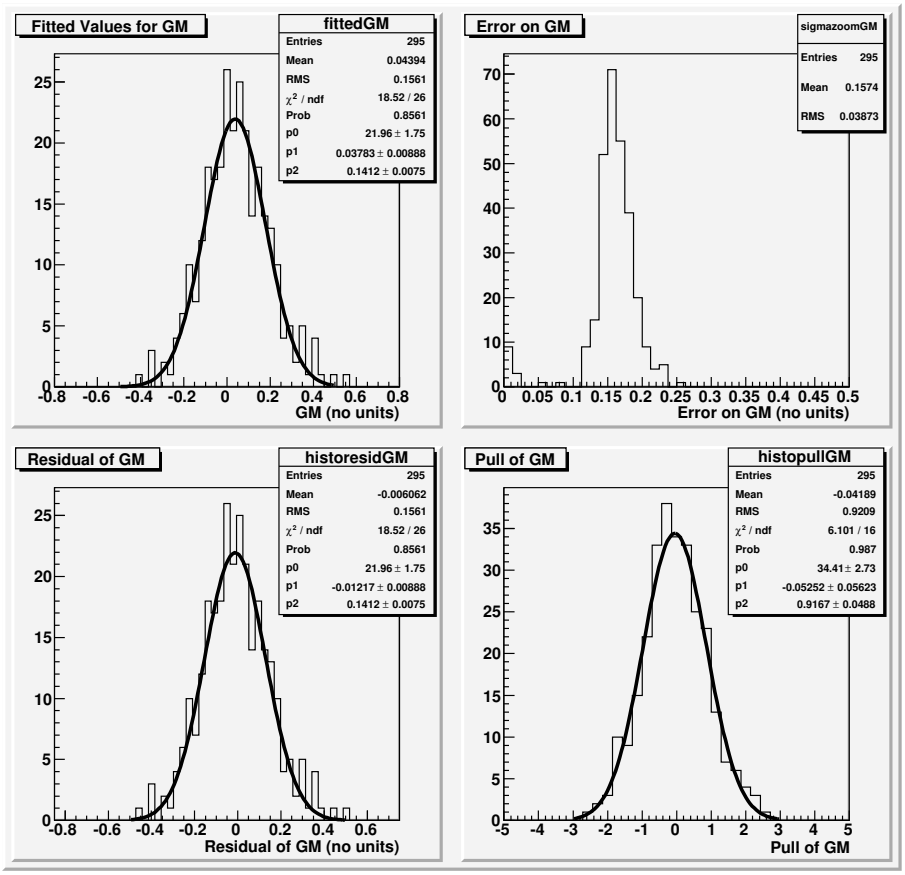

Figure 23: GM: fitting simplified model to 2 fb^{-1} of data simulated using the full model. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.

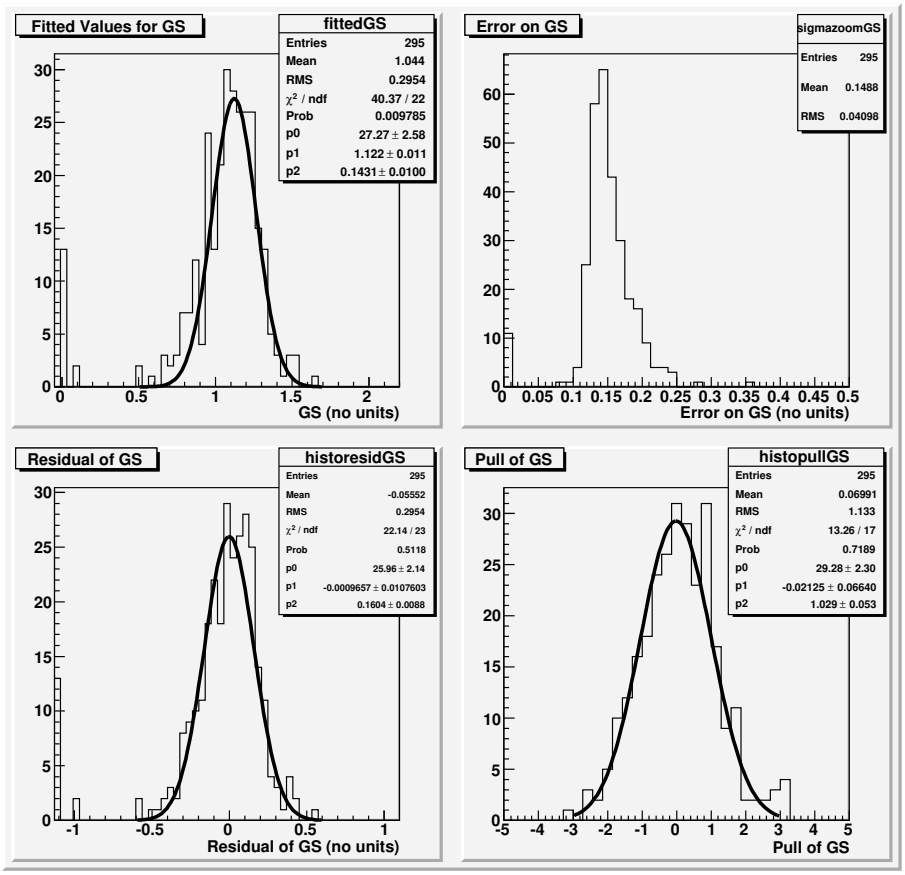

Figure 24: GS: fitting simplified model to 2 fb^{-1} of data simulated using the full model. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.

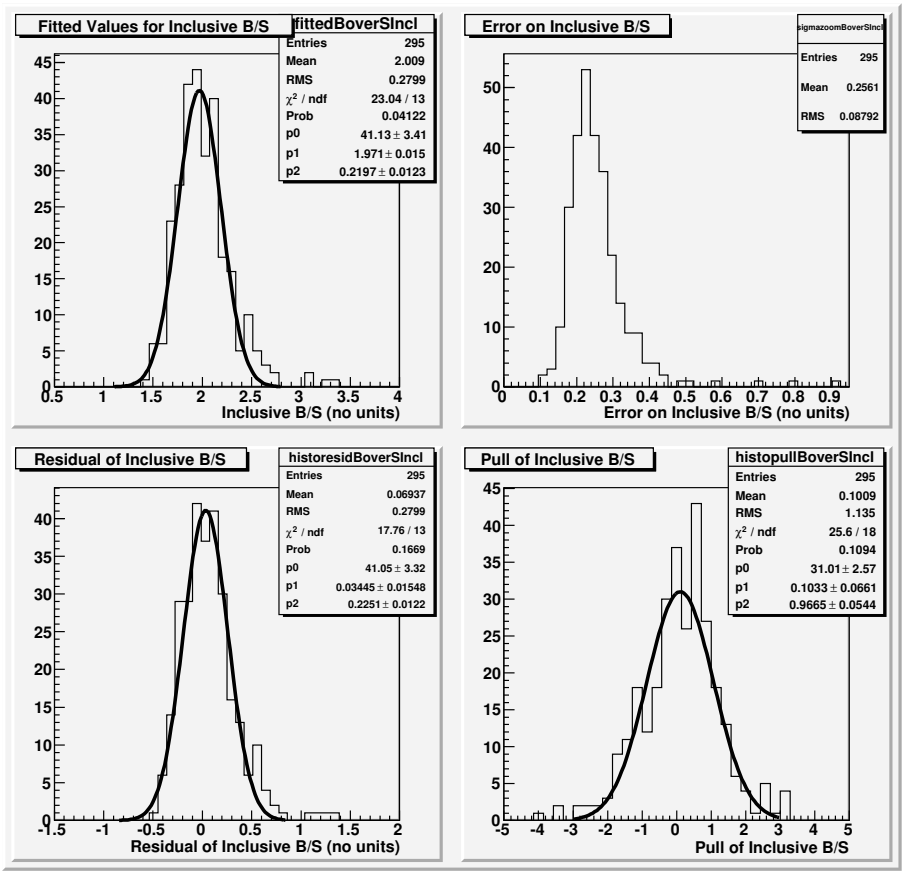

Figure 25: Inclusive Background to Signal Ratio: fitting simplified model to 2 fb^{-1} of data simulated using the full model. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.

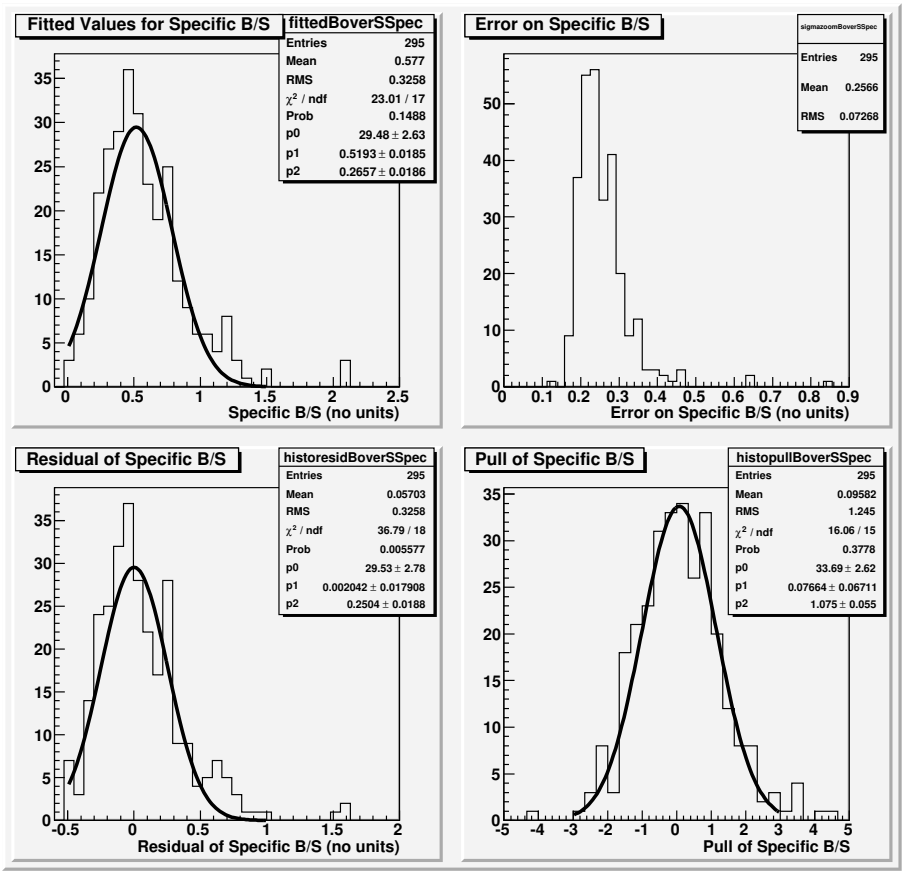

Figure 26: Specific Background to Signal Ratio: fitting simplified model to 2 fb^{-1} of data simulated using the full model. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.

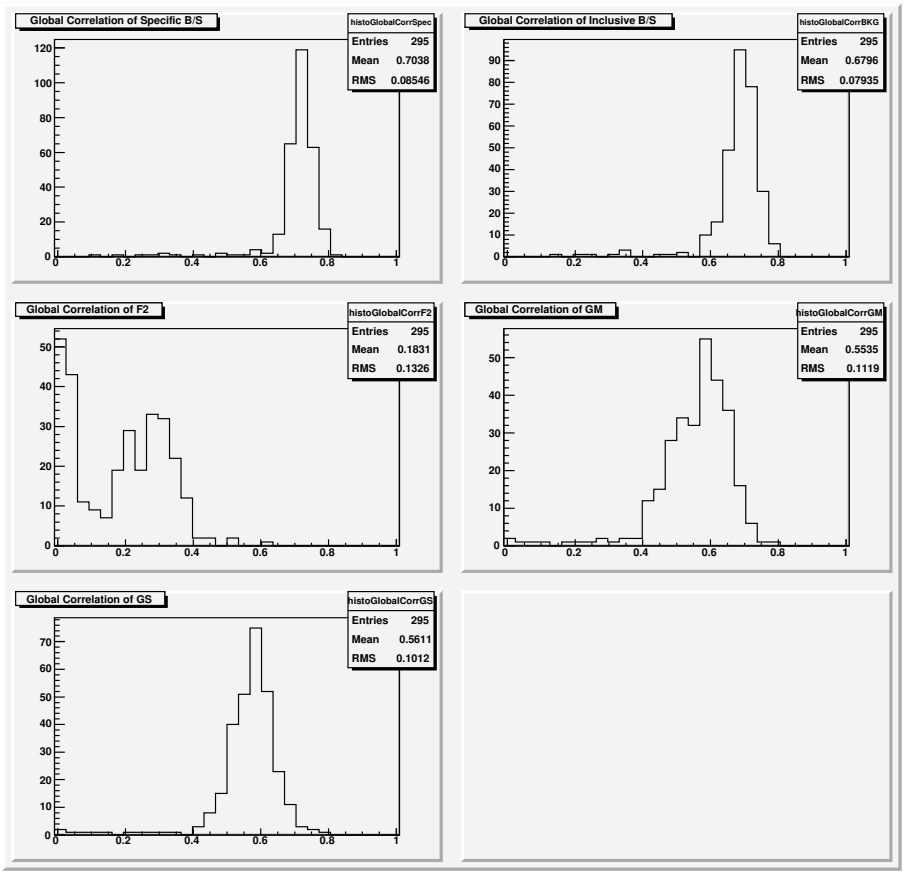

Figure 27: Global correlations: fitting simplified model to 2 fb^{-1} of data simulated using the full model. Clockwise from top left are the distributions for: Specific Background to Signal Ratio, Inclusive Background to Signal Ratio, GM, GS and F2.

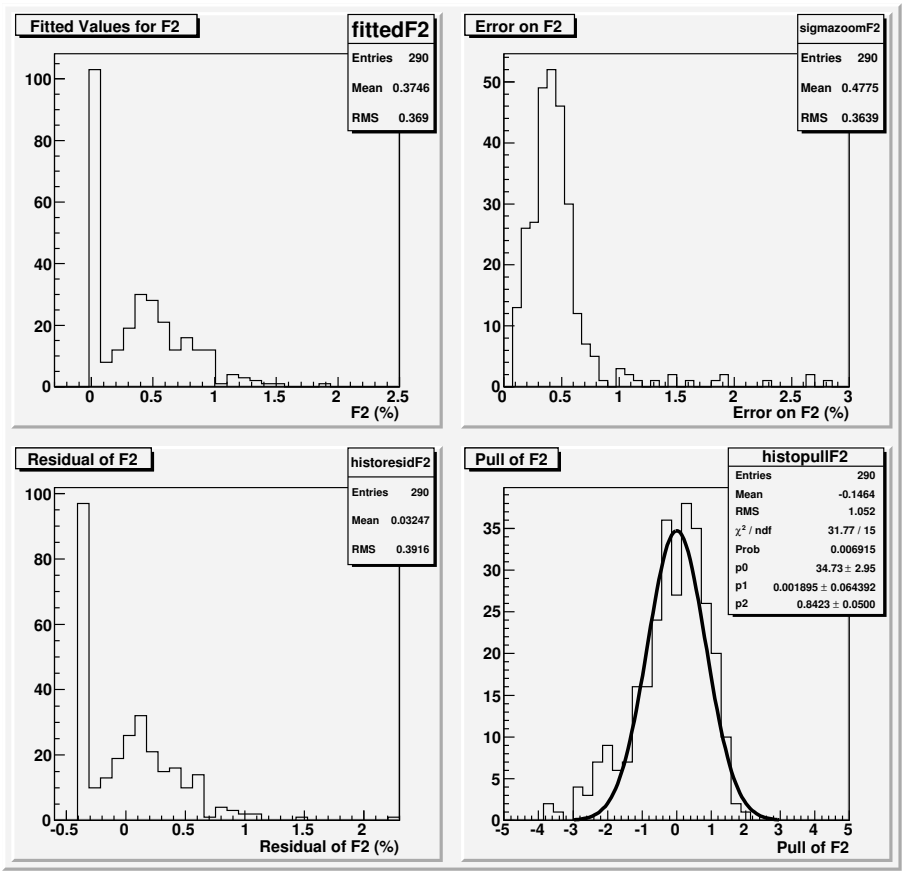

Figure 28: F2: fitting simplified model to 2 fb^{-1} of data simulated using the simplified model. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.

5.4 Discussion

Now there are very few outliers in the distributions, and the fitted values and pull distributions are all close to being Gaussian. The pull distributions all have widths which are compatible with 1 at the $\lesssim 1.5\sigma$ level. Biases have been introduced in the fitted values, but these are not too large: for GM 0.04±0.01 is found instead of 0.08 and for GS 1.12±0.01 is found instead of 1.08. The biases come from the fact that the resolution model in the data (the full model) is not quite the same as the resolution model used in the fit (the simplified model).

Also there is a spike of about 15 events appearing near zero in the fitted values graph for GS. These events are the same events that are causing spikes near zero in the fitted error graphs for GM and GS, and they are also the events causing outliers at high values in the B/S fitted values graphs.

Since there are not many of these events, and they have GS values that are clearly wrong (values like GM = 0.02 ± 0.05), if such a bad fit shows itself on data it would be spotted as wrong immediately, and careful steps could be taken to coax the fit into returning a sensible result.

5.5 Results Fitting the Simplified Model to Data Made Using the Simplified Model

One can also make toy data using the simplified model and fit to it using the same model. Again, almost all of the covariance matrices (290 out of 300) are found to be good. Results are shown in Figs. 28 to 33.

5.6 Discussion

These results are mostly very good, with one or two small issues—for example the pulls widths for F2 and GM are a bit low (0.84±0.05 and 0.90±0.05 respectively). There are also small biases in the fitted values for GM and the background to signal ratios, which are $\approx 3\sigma$ away from the seed values (here σ refers to the error on the mean of the Gaussian, *not* the width of the Gaussian, which is a much larger value). However these may simply be due to the Gaussian fits used not finding the true mean or width due to the low number of jobs. In

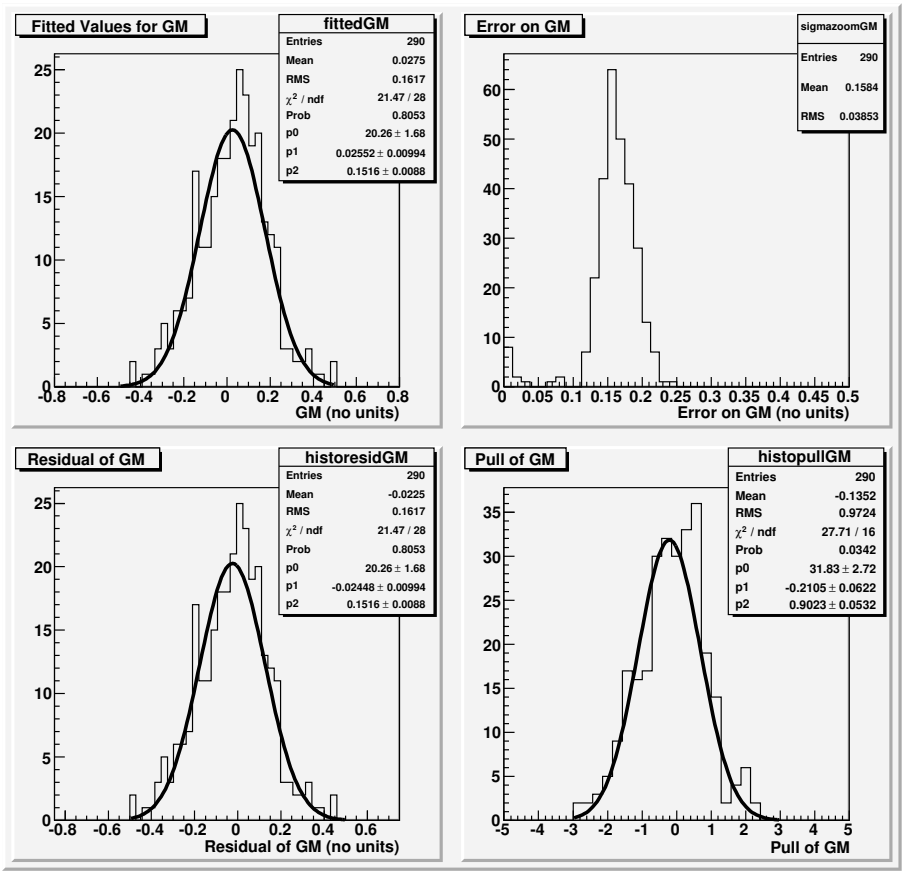

Figure 29: GM: fitting simplified model to 2 fb^{-1} of data simulated using the simplified model. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.

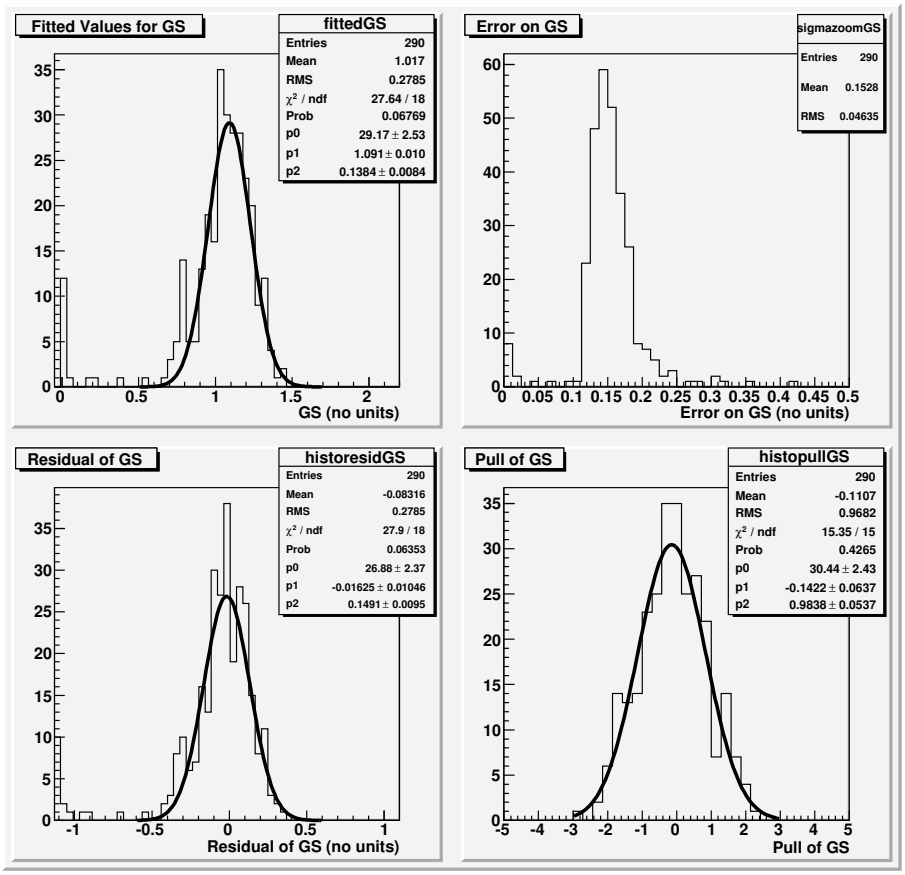

Figure 30: GS: fitting simplified model to 2 fb^{-1} of data simulated using the simplified model. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.

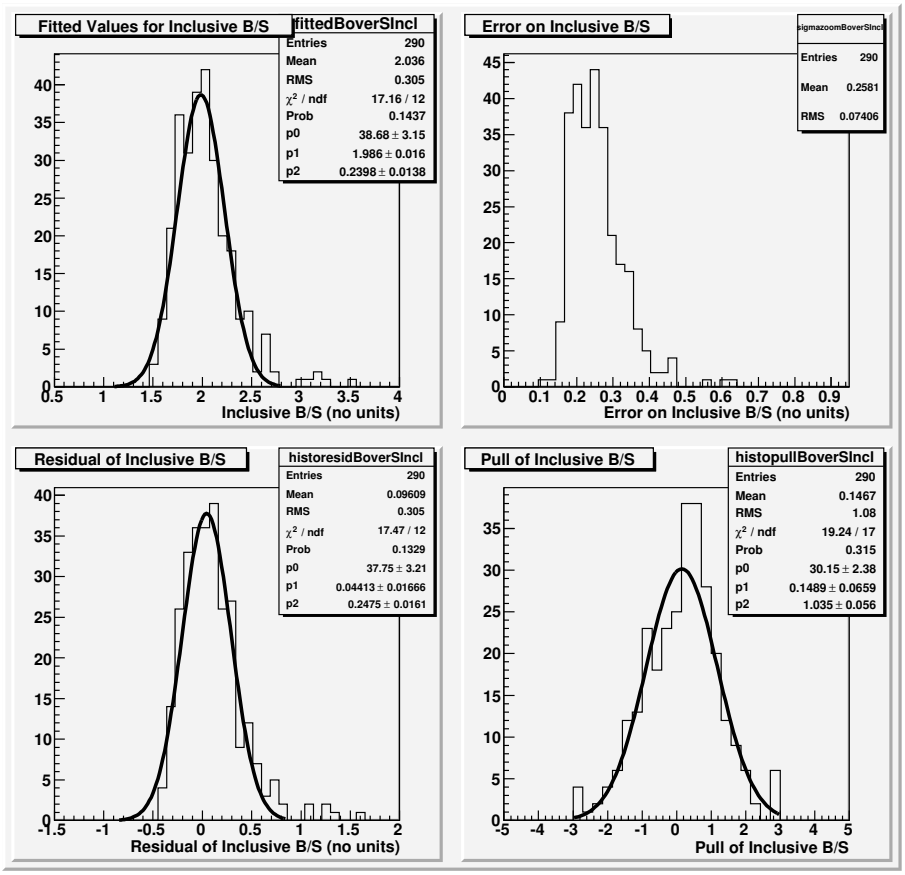

Figure 31: Inclusive Background to Signal Ratio: fitting simplified model to 2 fb^{-1} of data simulated using the simplified model. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.

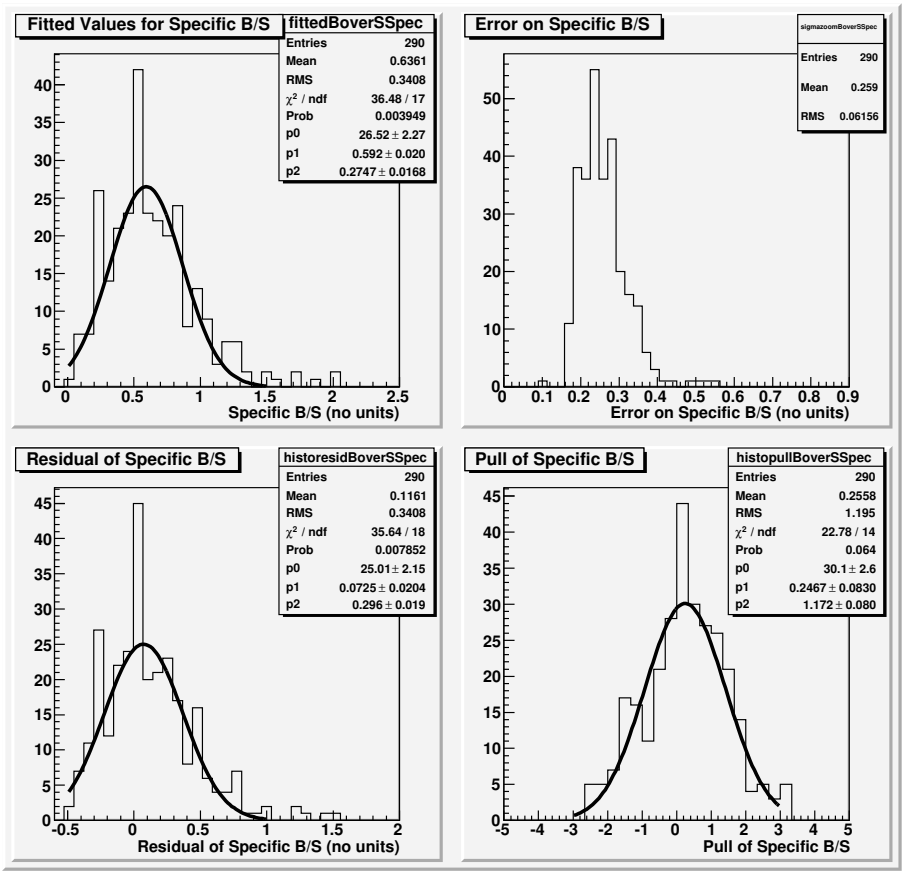

Figure 32: Specific Background to Signal Ratio: fitting simplified model to 2 fb^{-1} of data simulated using the simplified model. Clockwise from top left are the distributions of: fitted values, fitted errors, pulls and residuals.

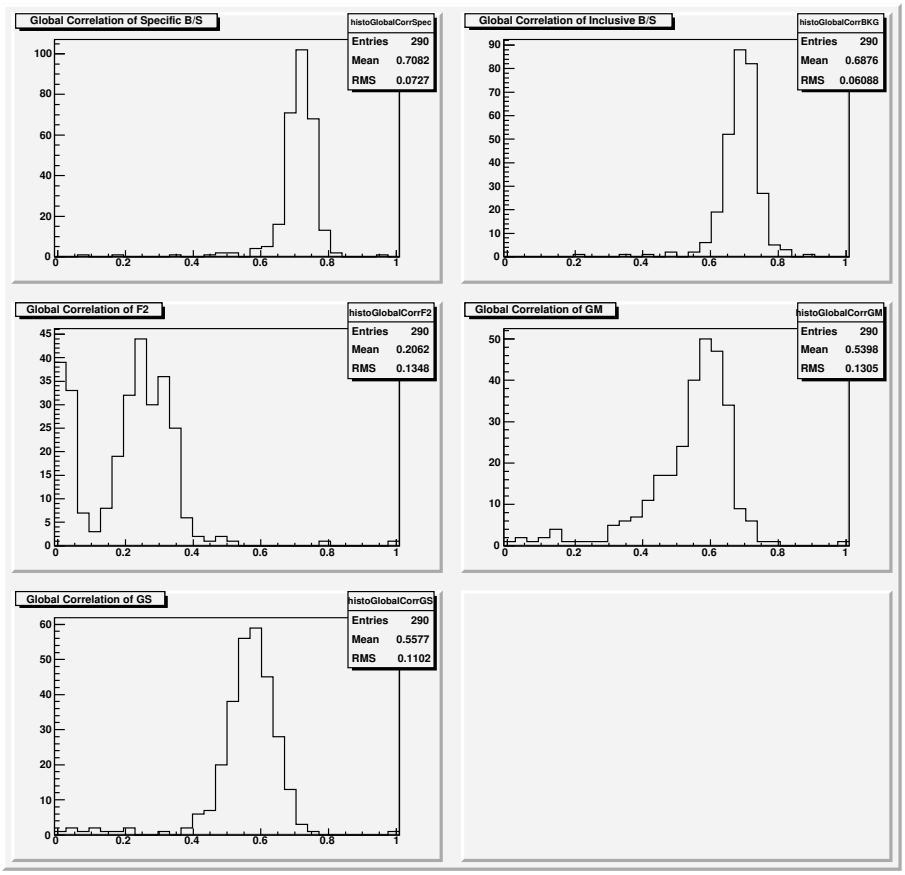

Figure 33: Global correlations: fitting simplified model to 2 fb^{-1} of data simulated using the simplified model. Clockwise from top left are the distributions for: Specific Background to Signal Ratio, Inclusive Background to Signal Ratio, GM, GS and F2.

any case the absolute size of the biases seen are small: around 10% for the pull widths, and up to 0.07 for the means of the fitted values.

The handful of events with GS fitted to near zero (see Sec. 5.3) are still there, and in the same number (15 out of 300). As already mentioned in Sec. 5.3, such bad fits would be spotted if they occurred during fits to data, and appropriate action taken to remedy the problem.

5.7 Summary of Results

All the studies reported in Secs. 5.1, 5.3 and 5.5 were repeated with toy datasets corresponding to 10 fb^{-1} of data. The results were found to be qualitatively the same, with the fitted errors (and width of the fitted values distributions) being smaller by factors of around $\sqrt{5} \approx 2.2$, which is the error reduction that would be expected in going from 2 fb^{-1} of data to 10 fb^{-1} of data.

It has been demonstrated that attempting to extract the parameters of the full resolution model does not give reliable results—the fit often returns a bad covariance matrix, and even the jobs with good covariance matrices give pathological distributions, especially for the parameters $F1$ and $SHIFT$. However if one does the fit using the simplified model, a good fit is obtained the large majority or the time, and the distributions of fitted values and pulls look normal. If the proper time residuals can be adequately described by the simplified model, then the parameters of the simplified model are recovered correctly from a fit to data. If the proper time residuals are assumed to be described by the full model, then the relevant parameters of the full model are recovered with a small bias by fitting to data using the simplified model. In particular, for all parameters the magnitude of the bias is smaller than the width of the fitted values distribution, even after 10 fb^{-1} of data has been collected (see below).

The results for fitting the simplified model to data made with the simplified model with the different dataset sizes are compared in Table 8. The results shown are the mean and width of the Gaussian fitted to the fitted values distribution of each parameter. (Note that for $F2$ with 2 fb^{-1} of data there were not enough jobs returning $F2 \neq 0$ to make fitting a Gaussian to the fitted values distribution worthwhile. However with 10 fb^{-1} of data the error was low enough that almost all jobs returned $F2 \neq 0$, and a Gaussian fit could be performed. For illustration the $F2$ results for 10 fb^{-1} are shown in Fig. 34.)

Table 8: Comparison of fit results to 2 fb^{-1} of data and 10 fb^{-1} of data. Results shown are from fitting the simplified model to data made using the simplified model.

Fit Result	Value for 2 fb^{-1}	Value for 10 fb^{-1}	Input value
F2 mean	not fitted	(0.35±0.01)%	0.35%
F2 width	not fitted	(0.16±0.01)%	n/a
GM mean	0.02±0.01	0.04±0.01	0.05
GM width	0.15±0.01	0.07±0.01	n/a
GS mean	1.09±0.01	1.09±0.01	1.10
GS width	0.14±0.01	0.06±0.01	n/a
Spec. B/S mean	0.59±0.02	0.54±0.01	0.52
Spec. B/S width	0.27±0.02	0.12±0.01	n/a
Incl. B/S mean	1.99±0.02	1.96±0.01	1.94
Incl. B/S width	0.24±0.01	0.10±0.01	n/a

It is expected that, given the magnitude of the errors (i.e the widths) in Table 8, the fit to the $B_s \to K^-\pi^+$ proper time distribution by the method laid out in this note will start to give useful input to the study of the proper time resolution model for $B_s \to K^-K^+$ (and hence to the measurement of γ) with around 2 fb^{-1} of data. By the time 10 fb^{-1} of data has been collected, the fit will be able to provide important constraints to the study of $B_s \to K^-K^+$, as well as giving another way to measure the B/S ratios for $B_s \to K^-\pi^+$.

6 Conclusions

Both the full and simplified proper time resolution models (see Eqns. 1 and 2 respectively) have been shown to describe the proper time residual distributions for Monte Carlo simulated data of the $B \to h^+h^{'-}$ channels. The parameters for all four channels are found to be compatible with each other.

The parameters obtained from the fits to the full Monte Carlo residuals have been used to guide a toy Monte Carlo study. This study requires that certain properties of the data, such as the mistag rate and the B meson lifetime, have already been well measured by other methods. The study has shown that the parameters of the simplified resolution model, as well as the background to signal ratios, can be determined from a fit to the

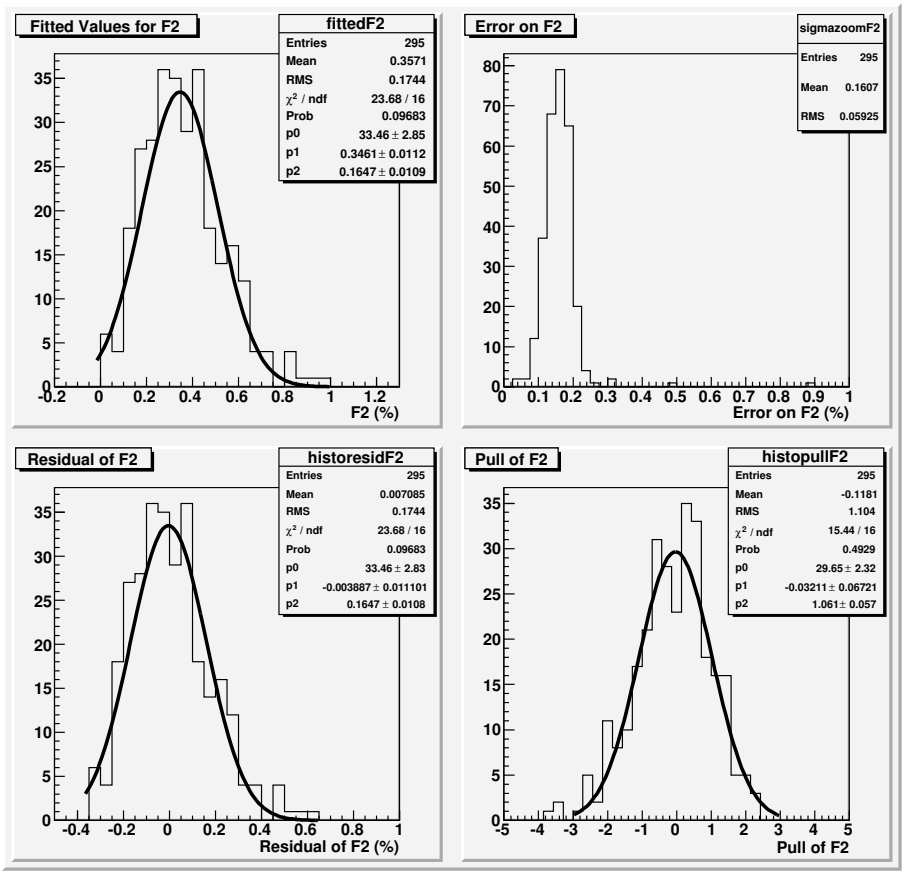

Figure 34: F2: fitting simplified model to 10 fb^{-1} of data simulated using the simplified model.

reconstructed proper time distribution, using information from the per-event proper time error. Even if the underlying data are considered to be described by the full resolution model rather than the simplified model, the parameters of the model can still be measured with biases which are smaller than the uncertainty on the parameters after 10 fb^{-1} of data. After one nominal year of LHCb data, the fit can be expected to produce results that can provide non-trivial constraints on the proper time resolution model for $B_s \to K^-K^+$, and hence form part of the input to the measurement of γ.

Another possible method of extracting the parameters of the proper time resolution model for the $B \to h^+h^{'-}$ channels is to study the reconstructed proper time distribution of prompt J/ψ events. The proper time residual for prompt J/ψ is the same as the reconstructed proper time, because the true lifetime is known to be (very close to) zero. The unphysical negative side of the reconstructed proper time distribution then gives information on the proper time resolution model for J/ψ. The open question is how similar this model would be to the model for the $B \to h^+h^{'-}$ channels. Some preliminary work on this method has been carried out by the author.

7 Acknowledgements

I would like to thank Gerhard Raven and Peter Vankov for being kind enough to invite me to NIKHEF to spend a week working with them. During this visit I was introduced to Peter's code for performing the fit to the Monte Carlo residuals, and started to develop the code used to perform the toy study of fits to data.

Thanks should also go to the authors of the DC04 study of $B \to h^+h^{'-}$ channels, as this was the source of the selection used and also the background distributions and background to signal ratios.

Finally I would like to thank all the members of the Glasgow LHCb group for their helpful advice and comments during both the carrying out of the work described here, and during the writing of this note.

References

[1] LHCb Collaboration, *LHCb Reoptimized Detector Design and Performance*, CERN LHCC 2003-030 LHCb, 9th September 2003

[2] R. Fleischer, Phys. Lett. **B459**, 306 (1999)

[3] P. Vankov and G. Raven, *Proper-Time Resolution Modeling*, LHCb 2007-055, 5th July 2007

[4] G. Raven, *Selection of $B_s \to J/\psi\phi$ and $B_u \to J/\psi K^+$*, LHCb 2003-118, 4th December 2003

[5] K.Cranmer, *Kernel Estimation in High Energy Physics*, hep-ex/0011057

[6] http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/

[7] M. Calvi, O. Leroy and M. Musy, *Flavour Tagging Algorithms and Performances in LHCb*, LHCb 2007-058, 15th May 2007

[8] A. Carbone et al., *Charmless charged two-body B decays at LHCb*, LHCb 2007-059, 8th September 2007

[9] http://roofit.sourceforge.net/

[10] http://root.cern.ch

[11] http://wwwasdoc.web.cern.ch/wwwasdoc/minuit/minmain.html